



Schumpeter

Tesla faces an identity crisis: carmaker or tech firm?

Elon Musk's fiendish conundrum

ON THE NIGHT before Elon Musk unveiled Tesla's first-quarter results on April 23rd, your columnist brought his car to a halt, noticing a futuristic vehicle hooked up to a Tesla charging station in Los Angeles. It was a dark-purple Cybertruck. Twinkling lights glittered behind the tinted windows. It looked so wedgelike, angular and otherworldly that it could have moonlighted as an armoured personnel carrier in "Civil War", a new apocalyptic film.

Its owner, Dennis Wang, is a Tesla devotee. Besides his four-month-old Cybertruck, he has owned Mr Musk's original ("sexy") quartet: the Models S, 3, X and Y. He has held shares in the company since 2018. He has full faith in Mr Musk. Despite a 40% plunge in Tesla's share price this year in the run-up to the earnings report, as well as the announcement in recent weeks of falling vehicle sales and unprecedented lay-offs, he believes the billionaire remains the best person to run the company. Even an embarrassing Cybertruck recall, caused by a stuck accelerator, was quickly fixed, he says, pointing to a new bolt in the pedal.

Yet however much Mr Wang loves Teslas, he does not think of Tesla as a car firm. He says it is a tech company. As he puts it, all electric vehicles (EVs) offer a similar driving experience. What differentiates them is the software—the brains beneath the dashboard. In Tesla's case, that is the latest version of its self-driving technology, which he calls "fantastic". His view is shared by many Tesla loyalists. It is why the company's shares trade at a multiple of earnings typical of a zippy software firm, not of a metal-basher.

Wall Street takes a different view. Though investors hope Tesla will one day make money from its snazzy artificial intelligence (AI), for now they want it to restore growth by selling more cars—the cheaper the better. Hence the sigh of relief when Tesla outlined plans within an otherwise dismal earnings report (revenues, profit margins and free cashflow all crashed) to start producing affordable vehicles by 2025 that would not rely on big new investments. Tesla's share price promptly soared more than 10%. Call that a \$5obn thumbs-up from the unit-economics guys.

Mr Musk has a history of trying to have it both ways. When investors were doubtful about demand for Tesla's EVs at the end of the 2010s, he promised shareholders that its so-called full self-driving (FSD) technology would put 1m robotaxis on the road by 2020. That did not happen, so during the pandemic, as Tesla's sales rocketed, he changed his tune. He boasted that sales were growing faster than Henry Ford's Model T, and that Tesla aspired to sell 20m EVs a year by 2030.

This year it is touch and go whether Tesla will sell more than the 1.8m cars it shipped in 2023. So Mr Musk has flipped the script again. Once more he is highlighting FSD, though this time with a twist: the latest version is so good, he told analysts this week, that it is impossible to understand the company without trying it. He went so far as to say: "If someone doesn't believe Tesla will solve autonomy, I think they should not be an investor in the company." His competing narratives create quite the conundrum among investment types. Can Tesla be a car company as well as a tech company? The answer, broadly, is yes. But it depends on which of its markets you are talking about.

From a volume-growth perspective, no country is more important than China. It is the world's biggest EV market, and though growth is slowing, sales are still rising much faster than in America. However, competition is fierce and a price war is shredding Tesla's business there. Tesla has not said

where the cheaper model it is planning will be sold. But if it is made available globally, it could help it fend off competition from BYD, a low-cost Chinese competitor that is not just the biggest EV seller in China but also has a strong presence around the world (though not in America).

Tesla's American home turf is different. Mr Musk's firm is already the market leader, so its growth prospects are probably constrained, more so because of the rising popularity of hybrids. Yet it has to sell more cars to generate cash to fund the purchase of huge volumes of AI chips that it needs to run its FSD technology. That is where a cheaper car comes in. It could help Tesla cross a bridge to the future while it attempts to overcome the huge engineering and regulatory challenges necessary for cars to drive people, rather than the other way round.

There are lots of potential roadblocks ahead. First is the risk of crumbling morale. Besides the sacking of one-tenth of its workforce, Tesla has lost several highly respected executives recently (the latest announced his departure on the quarterly earnings call). Second, trust between Mr Musk and big investors is gossamer-thin. Who knows how he will react if a majority at next month's shareholder meeting vote against the board's efforts to reinstate his \$56bn payout from 2018 that was voided by a Delaware judge. Third, the difficulty of running many businesses besides Tesla is compounded by Mr Musk's "demon mode"—irascible outbursts that can leave rubble in their wake.

| *View from the Cybertruck*

Like many Muskophiles, Mr Wang expects Tesla's boss to pull through. As a carmaker, Mr Musk excels. The Cybertruck, says its driver as his corgi scampers on the back seat, is the most comfortable car he has ever owned. As a technologist, Mr Musk continues to improve. Though Mr Wang acknowledges that the latest version of FSD requires driver supervision, he says being able to "sit back and decompress" on his commute is as valuable

to him as a pot of money. Above all, no one matches Mr Musk when it comes to turning engineering dreams into reality. As he puts it, “If Elon wants to put a chip in your head, you will get a chip in your head.” Just don’t expect it to be implanted until years after it is promised. And be prepared for its Svengali to melt down in the meantime. ■



熊彼特

特斯拉面临身份危机。它是汽车公司还是科技公司？

马斯克的大难题

在伊隆·马斯克于4月23日揭晓特斯拉一季度业绩的前一晚，本专栏作者在洛杉矶某处停车，注意到一个特斯拉充电桩上连着一辆未来感十足的汽车。那是一辆深紫色的Cybertruck。闪烁的灯光透过深色玻璃窗映射出来。楔形车身棱角分明，超凡脱俗，简直可以在新末日电影《内战》（Civil War）中充当装甲运兵车。

车主丹尼斯·王（音译）是特斯拉的“死忠粉”。除了这辆刚买了四个月的Cybertruck，他还买过马斯克原先的全部“性感”四件套：Model S、Model 3、Model X和Model Y。他自2018年起就持有该公司的股票。他对马斯克信心满满。尽管今年特斯拉的股价在临近财报发布时暴跌40%，近几周还宣布汽车销量下滑，并进行了前所未有的裁员，但他认为这位亿万富翁仍是掌舵特斯拉的不二人选。即使Cybertruck由于油门踏板会卡住而被尴尬地召回，他也指着踏板上的新螺栓说问题很快得到了解决。

不过，王先生再怎么喜欢特斯拉的车，也不认为特斯拉是一家汽车公司。他说它是一家科技公司。照他的话说，所有电动汽车的驾驶体验都差不多。不同之处在于软件，也就是仪表盘下的大脑。对于特斯拉而言，这就是它最新版的自动驾驶技术，王先生觉得“超赞”。许多特斯拉拥趸都跟他看法相同。这就是为什么特斯拉的市盈率能赶得上一家热门软件公司，而不像一家汽车制造商。

华尔街则有不同看法。尽管投资者希望特斯拉有朝一日能通过时髦的人工智能服务获利，但目前他们还是希望它能通过卖出更多车来恢复增长，而且卖得越便宜越好。因此，虽说特斯拉这次的财报原本很糟糕（收入、利润率和自由现金流都大幅下降），但它描绘了不需要大规模新投资就于2025年生产经济实惠新车型的计划，让投资者松了口气。特斯拉的股价立即大涨超过10%。可以说这是看重单位经济效益的投资者一起竖起了价值

500亿美元的大拇指。

马斯克一直试图“既要又要”。当投资者在2010年代末对特斯拉电动汽车的市场需求产生怀疑时，他向股东承诺称，特斯拉的完全自动驾驶（FSD）技术将在2020年让100万辆无人驾驶出租车上路。但这并没有实现，所以随着疫情期间特斯拉的销售一路飙升，他改变了说法。他夸耀说特斯拉销量增长的速度比福特的T型车还快，而且期望到2030年能每年销售2000万辆电动汽车。

特斯拉今年的销量是否会超出2023年的180多万辆还很悬。所以马斯克又一次改变了口风，再次开始强调FSD，只不过这次有所不同。他近日告诉分析师，最新版本太棒了，不试用一下就无法真正了解特斯拉。他甚至说：“我认为不相信特斯拉终将解决自动驾驶问题的人就不应该投资特斯拉。”他这些互相抵触的说法让各方投资者挺困惑。特斯拉可以既是汽车公司，也是科技公司吗？总的来说，答案是肯定的。但这取决于你谈论的是它的哪个市场。

从销量增长的角度来看，没有哪个国家比中国更重要。中国是世界上最大的电动汽车市场，尽管增长速度在放缓，但销量增长仍然比在美国快得多。但中国市场竞争激烈，价格战正在削弱特斯拉在中国的业务。特斯拉尚未透露计划在哪里销售它那款更便宜的车型。但如果在全球推出，可能有助于抵御来自低成本中国对手比亚迪的竞争。比亚迪不仅是中国最大的电动汽车公司，在全球各地也都很畅销（但不包括美国）。

特斯拉在美国自家地盘上的情况有所不同。马斯克的公司已经是那里的市场领军者，所以其增长前景可能受限，在混合动力车日益流行之时尤其如此。然而，它必须销售更多的汽车来筹集资金，以购买运行它的FSD技术所需的大量AI芯片。这就需要更便宜的车型发挥作用。在特斯拉努力克服无人驾驶巨大的工程和监管挑战的过程中，更便宜的车型可能可以作为过渡，帮助它迈向未来。

前方路上可能还有许多障碍。首先是士气崩溃的风险。除了裁员十分之一

之外，特斯拉最近还失去了几位德高望重的高管（最近一位是在这次季报会议上宣布离职的）。其次，马斯克与大投资者之间的信任薄如纸片。在5月的股东大会上，如果多数人投票反对董事会恢复他于2018年被特拉华州法官裁定无效的560亿美元薪酬方案，谁知道他会做何反应。第三，在特斯拉之外还要经营多家企业本就困难重重，马斯克的“恶魔模式”更让它难上加难——他的情绪失控可能留下一地鸡毛。

| Cybertruck的视野

和许多马斯克粉一样，王先生认为这位特斯拉老板能度过难关。作为汽车制造商，马斯克表现出色。王先生说Cybertruck是他开过的最舒服的车，说这话时他的柯基在后座上蹦蹦跳跳。作为技术专家，马斯克在不断改进。尽管王先生承认最新版的FSD仍需要驾驶员监督，但他说在通勤途中能够“放松和解压一下”的价值对他来说不亚于一大笔钱。最重要的是，在将工程梦想转化为现实方面，无人能与马斯克匹敌。正如王先生所说，“如果伊隆想把芯片植入你的脑袋，他说到做到。”只是不要指望很快就等到这块芯片。同时也要做好准备，在等待期间，这位造梦大师可能会失控崩溃。■



The rise of a new economic power

How strong is India's economy?

It isn't the next China, but it could still transform itself and the world

IN SIX WEEKS' time Narendra Modi is expected to win a third term as India's prime minister, cementing his status as its most important leader since Nehru. The electoral success of this tea-seller's son reflects his political skill, the potency of his Hindu-nationalist ideology and his erosion of democratic institutions. But it also reflects a sense among ordinary voters and elites that he is bringing India prosperity and power.

Mr Modi's India is an experiment in how to get richer amid deglobalisation and under strongman leadership. Whether it can grow fast and avoid unrest over the next 10-20 years will shape the fate of 1.4bn people and the world economy. As our special report explains, Mr Modi's formula is working—up to a point. But there are questions over whether India's success can last and whether it depends on him remaining in power.

India, the world's fastest-growing big country, is expanding at an annual rate of 6-7%. New data show private-sector confidence at its highest since 2010. Already the fifth-largest economy, it may rank third by 2027, after America and China. India's clout is showing up in new ways. American firms have 1.5m staff in India, more than in any other foreign country. Its stockmarket is the world's fourth-most-valuable, while the aviation market ranks third. India's purchases of Russian oil move global prices. Rising wealth means more geopolitical heft. After the Houthis disrupted the Suez canal, India deployed ten warships in the Middle East. Presidents Joe Biden and Donald Trump have courted it without disputing that it will remain an independent actor.

If you are looking for “the next China”—a manufacturing-led miracle—it

isn't India. The country is developing at a time of stagnating goods trade and factory automation. It therefore needs to pioneer a new model for growth. One pillar of this is familiar: a massive programme of infrastructure that knits together a vast single market. India has 149 airports, double the number a decade ago, and is adding 10,000km of roads and 15GW of solar-energy capacity a year. Some of this infrastructure is intangible, including digital payments, modern capital markets and banks, and a unified digital tax system. All this allows firms to exploit national economies of scale.

A second, more novel pillar is services exports, which have reached 10% of GDP. Global trade in services is still growing and Indian IT firms have marketed "global capability centres"—hubs that sell multinationals R&D and services such as law and accounting. Yet despite its slick tech campuses, India is still a semirural society. That explains the economic model's final pillar, a new type of welfare system in which hundreds of millions of poor Indians receive digital transfer-payments. New data suggest the share of the population living on less than \$2.15 a day in 2017 prices, a global measure of poverty, has fallen below 5% from 12% in 2011.

How much credit does Mr Modi deserve? His most successful policies draw on the liberal agenda that emerged in India in the 1990s and 2000s, but there is nothing wrong with that. He deserves credit for forcing through stalled reforms, personally overseeing key decisions and browbeating laggards and opponents in the bureaucracy. Some say he has fostered crony capitalism. Yet although some big firms get favours, concentration in business is falling, corruption has waned and business boasts a rich diversity. A cross between a CEO and a populist, Mr Modi relishes PowerPoint presentations as much as rallies. If he wins five more years, India will continue to grow strongly. So will its middle class: 60m people earn over \$10,000 a year; by 2027, 100m will, reckons Goldman Sachs, a bank that now has 20% of its staff in India.

Yet India faces a daunting problem. Out of a working-age population of 1bn, only 100m or so have formal jobs. Most of the rest are stuck in casual work or joblessness. Mr Modi's humble beginnings help him speak to these people. To absorb some of India's spare labour he is using a state-run incentive scheme to promote manufacturing. But even if the scheme hits its targets, it will create just 7m jobs. President Xi Jinping's plan for a Chinese export surge will only make the task harder.

India's economy must generate mass employment to sustain its growth. One path would be an even bigger IT sector, acting as a hub for a digitising world, and a cluster of export industries, including digital finance, food and defence (where stronger links with America would help). Spending by workers in these industries would in turn create more jobs in other sectors, from construction to hotels. An efficient, single domestic market would raise overall productivity and well-targeted welfare could help those who fall behind. For this, India would have to transform education and agriculture, and enable much more migration from the populous north to the big southern and western cities.

Judged by those epic standards, Mr Modi has too little to say. His Bharatiya Janata Party (BJP) has some talent and ideas but is mostly focused on ideology and Muslim-bashing. A rising illiberalism has curtailed political opposition and free speech. The fact that firms fear Mr Modi may explain why investment has yet to surge. The process of preparing the public for huge social change in the 2030s has barely begun. Remaking education, cities and agriculture will require the co-operation of state governments that are not led by the BJP and social groups that are facing disruption, but Mr Modi's rebarbative politics have left many of them estranged.

| *India's Lee Kuan Yew or its Erdogan?*

The question for India and its heavyweight economy is not whether Mr Modi wins, but whether he will evolve. Aged 73, he may find his powers of

management fade. To create a new reform agenda on a par with the one that emerged out of the 1990s, and to foster a thriving knowledge economy that rewards people for thinking for themselves, he will have to temper his autocratic impulses. To attract more local and foreign investment and to find a growth-minded successor, his party will need to curb its chauvinistic politics. If not, Mr Modi's mission of national renewal will not live up to its promise. ■



【首文】一个新经济强国的崛起

印度的经济实力有多强？

它不是下一个中国，但仍可能改变自己和全世界

六月初，莫迪有望赢得第三个印度总理任期，巩固其作为尼赫鲁以来最重要的印度领导人的地位。这位茶商之子的选举胜利反映出他的政治手腕、他的印度教民族主义意识形态的强大影响力，以及他对民主制度的侵蚀。但这也反映了在普通选民和精英阶层中存在的一种看法——他正在为印度带来繁荣富强。

莫迪治下的印度正在进行一场实验，探索如何让印度在去全球化和强人领导下变得更加富裕。印度能否在未来10到20年内实现经济的快速增长并避免动荡，将决定14亿印度人民和全球经济的命运。莫迪的手段在一定程度上卓有成效。但印度的成功能否持续，以及这是否有赖于莫迪继续掌权，这些都是问题。

印度是全球经济增长最快的大国，年增长率为6%至7%。最新数据显示，私营部门的信心达到了2010年以来的最高点。印度已经成为全球第五大经济体，到2027年可能成为美国和中国之后的第三大经济体。印度的影响力正在新的方面显现。美国公司在印度雇有150万名员工，比在其他任何一个国家都要多。印度股票市场的规模位居世界第四，航空市场位居世界第三。印度对俄罗斯石油的购买影响着全球油价。不断增长的财富意味着更大的地缘影响力。胡塞武装袭扰苏伊士运河后，印度在中东部署了十艘军舰。拜登和特朗普都曾向印度示好，而没有质疑它是否将继续保持独立性。

如果你认为印度会是“下一个中国”——一个由制造业主导的奇迹——那就错了。印度的发展恰逢商品贸易停滞和工厂自动化，因此它需要开创一种新的经济增长模式。这种模式的一个重要支柱并不让人陌生：一个连接起巨大单一市场的庞大基建计划。印度有149个机场，数量比十年前翻了一番，每年还新增一万公里道路和15GW的太阳能发电能力。其中还有一些

无形的基础设施，比如数字支付、现代资本市场和银行，以及统一的数字税收系统等。所有这些都让企业能够利用国家规模经济。

第二个支柱较为新颖，那就是服务出口，它已经占到印度GDP的10%。全球服务贸易仍然保持增长，印度的IT公司也顺势推出了很多“全球能力中心”——向跨国公司提供研发能力以及法律、会计等服务。然而，尽管印度拥有光鲜的科技园区，它仍然是一个半农村社会。这就解释了这种经济增长模式为什么还有最后一个支柱——一种新型的福利体系。在这种体系下，数以亿计的印度穷人通过数字化收到了转移支付。最新数据显示，印度在全球贫困线以下、也就是每天生活支出不足2.15美元（2017年价格）的人口比例已从2011年的12%下降至5%以下。

莫迪应该得到多少赞誉？他最成功的那些政策借力上世纪90年代和本世纪头十年在印度出现的自由主义议程，但这并没什么问题。莫迪强力推动了停滞不前的改革、亲自监督关键决策、打击官僚机构中的懒政者和反对者等，这些都值得赞扬。有人说他助长了裙带资本主义。不过，尽管一些大公司受到了偏袒，但行业集中度正在下降，腐败现象减少了，产业也呈现明显的多元化。莫迪集CEO和民粹主义者于一身，他喜欢举行集会，也喜欢PPT工作汇报。如果他再干五年，印度经济会继续强劲增长。印度的中产阶级也会稳步扩大：目前印度有6000万人年收入超过一万美元，投行高盛估计，到2027年，这一群体将增加到一亿人。高盛自己就有20%的员工在印度。

然而，印度面临着一个严峻问题。在其10亿劳动年龄人口中，只有约一亿人有正式工作。其余大多数人只能打零工或者失业。莫迪低微的出身有助于他和这些人群沟通。为了吸收印度的部分闲置劳动力，他正利用一项政府激励计划来促进制造业发展。不过，即使该计划达到预期目标，也只能创造700万个就业岗位。中国领导人习近平为大幅增加中国出口而提出的计划只会让莫迪这个计划变得更加艰巨。

印度经济要保持增长，就必须创造大量就业。其中一个途径是进一步做大IT行业，成为数字化世界的中心，并形成包括数字金融、食品和国防（加

强与美国的联系将有助于其国防产业）等出口产业集群。这些行业中从业人员的消费反过来又会为从建筑到酒店的其他行业创造更多就业机会。一个高效、单一的国内市场将提高整体生产率，而有针对性的福利政策可以帮助弱势群体。为此，印度必须改革教育和农业，并让更多人从人口稠密的北部迁移到南部和西部的大城市。

以这些宏大的标准来评判，莫迪没有什么好自吹自擂的。他所领导的印度人民党（BJP）也有一些人才和想法，但基本上都用在意识形态和打压穆斯林上。愈演愈烈的非自由主义遏制了政治反对派和言论自由。企业对莫迪的忧惧或许可以解释为什么印度的投资仍然没有大幅增长。让民众为本世纪30年代的巨大社会变革做好准备的过程才刚刚开始。要重塑教育、城市和农业，就需要其他党派执掌的邦政府以及面临冲击的各社会团体携手合作，但莫迪不得人心的政治主张已经让其中很多人疏远了他。

丨 印度李光耀还是印度埃尔多安？

对于印度及其举足轻重的经济来说，问题不在于莫迪是否会胜选，而在于他是否会有所改变。现年73岁的他可能会发现自己的掌控力已不如从前。要制定出能媲美上世纪90年代的新改革议程，促进奖励人们独立思考的知识经济的蓬勃发展，他将必须抑制自己的独裁冲动。同时为了吸引更多的国内外投资，并找到一个注重经济增长的继任者，他领导的印度人民党需要克制其政治上的沙文主义。否则，莫迪的民族复兴使命将无法兑现其承诺。 ■



Poor prescription

America's moves against Chinese biotech will hurt patients at home

The motives behind the BIOSECURE act are muddy

AMERICA'S CRACKDOWN on Chinese trade is broadening. On the campaign trail on April 17th President Joe Biden proposed tripling tariffs on steel imports, citing China's unfair trade practices. Having choked off China's access to advanced semiconductors and moved to ban TikTok, a Chinese-owned social-media app, lawmakers are eyeing a new target: biotechnology. The BIOSECURE act, which has bipartisan support in Congress, proposes to end government contracts for firms that count Chinese biotech companies as clients or suppliers. American officials have previously said they want to guard a "small yard" of sensitive technologies with a "high fence". This bill illustrates that the yard is getting bigger, with sorry consequences for American consumers.

It uses the threat of ending lucrative federal contracts to sever American firms' ties with Chinese genomic sequencers, makers of sequencing machines and makers of large-molecule drugs such as weight-loss injectables. It extends the ban to any biotech firm with its headquarters in an adversary country, and mentions four Chinese companies by name.

One target is a sequencing firm called BGI, formerly known as Beijing Genomics Institute. BGI is the largest sequencer of human DNA in the world and operates in over 100 countries. It supplies prenatal tests and diagnostic swabs for covid-19 and other diseases. The firm, like its rivals, provides health screenings on the cheap in return for keeping its patients' anonymised data. The data in turn are used in cutting-edge drug development.

Lawmakers allege that the data hoovered up by BGI are stored within reach

of nosy Communist Party officials. The firm says sensitive information is stored privately and that its American operations are limited. But lawmakers need not take BGI's word for it. In February President Joe Biden banned exports of health data to adversaries, including China, mirroring China's own controls. Lawmakers did not need to go further.

Moreover, having begun with reasonable concerns about a specific firm's handling of data, the authors of the bill have widened their scope much further, by bringing in the makers of medical equipment and drugs, such as Wuxi AppTec, which conducts research and manufactures compounds used to make drugs for clients including the world's biggest pharma firms. WuXi AppTec and its sister firm, WuXi Biologics, are by revenue the world's largest providers of contract drug discovery and manufacturing. While some diversification away from Chinese producers might have been justifiable, that is not the stated intention of the bill. In any case, the legislation would not merely enforce diversification, but wholesale decoupling.

The lawmakers claim that Chinese biotech firms have stolen intellectual property (IP) and collaborated with the People's Liberation Army (PLA) and the Chinese government's repression of Uyghurs. (WuXi AppTec says it is not aware of unauthorised transfers of IP.) Yet here too the BIOSECURE act is an overreaction. Western biopharma firms are notoriously protective of their IP and are surely best placed to decide whom to trust with their drug recipes. Chinese firms that are militarily or morally compromised should be targeted on an individual basis, not by dint of their nationality or industry.

The muddier the motivation behind the legislation seems, the harder it is to escape the conclusion that old-fashioned protectionism is at play. And that is a problem, because it means the bill would unduly hurt American consumers, without delivering any of the supposed security benefits.

| *Small gain, high price*

If the legislation passes, as seems likely, drug shortages and delayed clinical trials for medicines would probably follow. Every large Western pharma firm and many small ones would have to abandon supply chains and find new partners for trials. Biotech startups in particular rely on cheap Chinese manufacturers to bring their products to market. And that would go against another stated intention of the Biden administration: to lower drug prices. ■



糟糕的处方

美国针对中国生物技术公司的措施会伤害本国患者

《生物安全法案》背后的动机模糊不清

美国对中国贸易的打压正在扩大。在4月17日的竞选活动中，美国总统拜登提出将钢的进口关税提高两倍，理由是中国存在不公平的贸易行为。在阻断了中国获得先进半导体的渠道，并寻求封禁中资社交媒体应用TikTok之后，议员们正在瞄准一个新目标：生物技术。《生物安全法案》（BIOSECURE Act）得到了国会两党的支持，它提议终止那些客户或供应商包含中国生物技术公司的美国企业的政府合同。美国官员此前曾表示，他们希望用“高墙”保护敏感技术的“小院”。这项法案表明，这个院子正越筑越大，将给美国消费者带来糟糕的后果。

它威胁终止利润丰厚的联邦政府合同，以此切断美国公司与中国基因组测序公司、测序设备制造商和大分子药物（如减肥注射剂）制造商的联系。它将禁令扩大到所有总部设在对手国家的生物技术公司，并且点名了四家中国公司。

其一是测序公司华大基因，其前身是北京基因组研究所。华大基因是世界上最大的人类DNA测序公司，在100多个国家开展业务。它提供产前筛查以及新冠和其他疾病的诊断拭子。这家公司和它的竞争对手们一样，以低廉的价格提供健康筛查，以换取留存病人的匿名数据。这些数据继而被用于尖端药物开发。

议员们称，华大基因收集的数据存储在爱插手的共产党官员够得着的地方。华大基因则表示敏感信息是私密存储的，且它在美国的业务很有限。但议员实则并不需要信或不信。今年2月，美国总统拜登禁止向包括中国在内的对手国家转移健康数据，模仿了中国的相关管控措施。议员们实在不需要再多此一举。

此外，法案的起草者一开始对一家特定公司如何处理数据的担忧是合理

的，但现在他们大大扩展了管控范围，把医疗设备和药物的制造商也囊括了进来，比如药明康德，这家公司研究和生产用于制药的化合物，其客户包括世界上最大的那些制药公司。药明康德及其姊妹公司药明生物以营收计是全球最大的合同药物研发和生产提供商。在中国生产商之外做一些多元化也许是合理的，但这并不是该法案明确陈述的本意。无论如何，这项立法将不只是推动多元化，而是会推动大规模脱钩。

议员们声称中国的生物技术公司窃取知识产权，并且和中国军方及政府一同压制维吾尔人。（药明康德表示，它不了解有未经授权而转移知识产权的行为。）但在这方面，《生物安全法案》也反应过度了。西方生物制药公司对自己知识产权的保护是出了名的，它们当然最清楚可以放心地把自己的药物配方交给谁。应该区别对待涉军事或有道德问题的中国公司，而不是按国别或行业一竿子打翻一船人。

立法背后的动机看起来越模糊不清，就越容易让人得出这样的结论：又是保护主义的老一套在作祟。而这是一个问题，因为这意味着该法案会给美国消费者造成不必要的损害，而不会带来任何预想的安全效益。

| 收益小，代价高

如果该法案获得通过（看起来很有可能），随之而来的可能是药物短缺和药物临床试验推迟。每家西方大药厂和许多小型制药公司都将不得不放弃供应链，寻找新的合作伙伴开展试验。生物技术创业公司尤其依赖廉价的中国制造商来将自己的产品推向市场。这还将违背拜登政府宣称的另一个意图：降低药品价格。■



Social media and security

Congress tells China: sell TikTok or we'll ban it

Only America's courts can save the video app now

Editor's note (April 24th): This story was updated after Joe Biden signed the bill into law.

JOE BIDEN joined TikTok only two months ago, with a short video entitled “lol hey guys”. On April 24th America's president approved a bill that could ban the popular app. The previous day, the Senate had passed a measure to crack down on “foreign adversary controlled applications”, including TikTok, as part of a package of legislation allocating military aid to Ukraine, Israel and Taiwan. Mr Biden wasted no time signing it into law, no matter the feelings of his 300,000 followers, or TikTok's 170m American users.

The law gives ByteDance, TikTok's Chinese owner, up to 12 months to sell to non-Chinese interests. (A previous version of the bill allowed six months; the new deadline pushes the matter conveniently beyond November's elections.) A wild selection of possible buyers is touted. Microsoft, Oracle and Walmart have all shown an interest in TikTok in the past. Steven Mnuchin, a former Treasury secretary, says he is putting together a group of investors.

They may not get the chance to bid. China's government, which owns a stake in a ByteDance subsidiary, has indicated it does not want to part with TikTok. It has classified the app's recommendation algorithm as a sensitive technology, whose export needs official approval. In March a Chinese government spokesman warned that, in the case of TikTok, “the relevant party should strictly abide by Chinese laws and regulations.” The comment was read by some as a warning to ByteDance.

That would leave TikTok no choice but to shut down in America, where last year it earned revenues of \$16bn, according to the Financial Times. The company is therefore pinning its hopes on an appeal to America's mighty courts. A memo sent to staff on April 21st described the new bill, which at the time had just been passed by the House of Representatives, as a "clear violation of [users'] First Amendment rights", which enshrine freedom of speech.

TikTok has a strong case, believes Evelyn Douek of Stanford Law School. "Decades of precedent hold that the government can't ban a form of communication because they don't like the content on it, even when it involves foreign adversaries," she says. TikTok has won in court before. Last year a judge overturned a ban imposed by the state of Montana, partly on free-speech grounds. An executive order to ban the app by then-president Donald Trump was blocked by judges in 2020.

Those in favour of a ban say the problem is not the content on TikTok, but the company's conduct. It is accused of harvesting users' data and manipulating what they see, both of which it denies. If courts can be persuaded that TikTok is up to no good, a free-speech defence will not necessarily save it. In 1986 an adult-book seller in New York lost a Supreme Court appeal against its closure, when judges argued that the reason for its shutdown was not the content of its books, but other, illegal activity taking place on the premises.

If TikTok wins, it could become an even stronger force in social media. "TikTok has been fighting with one hand tied behind its back against domestic competition," argues Mark Shmulik of Bernstein, a broker. While Meta, its arch rival, has come up with technical fixes to help its advertisers get around privacy changes Apple introduced for iPhones, TikTok has played it safe. If courts remove the threat of a ban, the company "could feel empowered to step on the gas", Mr Shmulik notes. It might also stem the

exodus of senior staff. Kevin Mayer, a former chief executive hired from Disney, left amid Mr Trump's efforts to ban the app. Vanessa Pappas, its chief operating officer, departed last year. Now Erich Andersen, the chief counsel, is reportedly preparing to move on.

Whatever happens in court, TikTok is already wondering which countries might follow America's hawkish lead. India, where TikTok had 200m users, banned it in 2020 (along with several other Chinese apps) following a skirmish at the border with China. Countries including Indonesia and Pakistan have imposed and then lifted short-term prohibitions. The Taliban, naturally, outlawed TikTok on returning to power in Afghanistan.

Juicier markets look safe for now. No big European country is demanding TikTok be sold. But Europe has a record of eventually following America when it comes to China-related security matters, as in the case of its belated clampdown on Huawei, a Chinese maker of telecoms gear. Countries' willingness to act will depend partly on the closeness of their security relationship with America. America's fellow members of the Five Eyes intelligence alliance—Australia, Britain, Canada and New Zealand—have already banned TikTok on government devices.

| *Out of lip-sync*

Further restrictions on TikTok could disrupt more than the market for social media—if China chooses to retaliate. It fired a warning shot earlier this month, banning app stores from offering apps including WhatsApp and Threads, a pair of Meta products, on national-security grounds. China could make life difficult for plenty of other big American companies. Tesla is suffering because of falling car sales in China. Apple's iPhone sales in the country are ebbing. American chipmakers like AMD are being hurt, too, as China encourages its smartphone-makers to use domestic chips. America may find that banning a short-video app has long-term consequences. ■



社交媒体与安全

美国国会向中国喊话：TikTok不卖就禁

如今只有美国法院能拯救这个视频应用了

编者按（4月24日）：在拜登总统签署涉TikTok法案后，本文已作更新。

美国总统拜登两个月前才开通了TikTok帐号，发布了一段标题为“哈哈，大家好”的短视频。4月24日，他就签署了一项可能禁止这个热门应用的法案。此前一天，参议院通过了一项打击包括TikTok在内的“外国对手控制的应用”的措施，这是向乌克兰、以色列和台湾提供军事援助的一揽子立法的一部分。拜登迫不及待地将其签署为法律，毫不顾忌自己那30万粉丝或1.7亿TikTok美国用户的感受。

该法律限定TikTok的中国母公司字节跳动在12个月内把TikTok出售给非中国买家。（之前一版法案是限期六个月；新期限顺势推迟到11月美国大选之后。）被津津乐道的可能买家五花八门。微软、甲骨文和沃尔玛都曾对TikTok表示过兴趣。前财政部长史蒂文·姆努钦（Steven Mnuchin）表示正在组建一个投资者团体。

但他们可能连出价的机会都没有。持有字节跳动一家子公司股份的中国政府已经显示出不想放弃TikTok。它已将该应用的推荐算法列为敏感技术，需要官方批准才能出口。3月，中国政府发言人警告称，在TikTok的问题上，“相关方应严格遵守中国法律法规”。该表态被一些人解读为是对字节跳动的警告。

这将使TikTok别无选择，只能关闭在美业务；据英国《金融时报》报道，去年TikTok在美国的收入为160亿美元。所以TikTok把希望寄托于向美国强大的法院上诉。该公司在4月21日发给员工的一份备忘录中直指当时刚获众议院通过的新法案“明显违反[用户的]第一修正案权利”，该修正案明确保障言论自由。

斯坦福大学法学院的伊夫林·杜埃克（Evelyn Douek）认为TikTok的理据很充分。“几十年来的判例表明，政府不能因为不喜欢以某种传播形式散播的内容就将这种形式禁止，即便涉及外国对手。”她说。TikTok有过法庭胜诉的先例。去年，一名法官推翻了蒙大拿州实施的禁令，言论自由是理据之一。2020年，时任美国总统特朗普禁止该应用的行政命令被法官阻止。

支持禁令者则表示，问题不在于TikTok上的内容，而在于该公司的行为。TikTok被指控收集用户数据并操纵用户看到的内容，该公司对此均予以否认。假如能说服法院相信TikTok别有用心，以言论自由作辩护未必能挽救它。1986年，纽约一家成人书店因被封禁而上诉最高法院，但终告败诉，法官认为书店遭关闭的原因不是所售书籍的内容，而是在店内发生的其他非法活动。

如果TikTok胜诉，它可能成为社交媒体领域里一股更强大的力量。券商盛博的马克·施穆利克（Mark Shmulik）认为：“TikTok在与美国本土对手的竞争中一直受到掣肘。”其主要对手Meta已拿出技术方案帮助广告客户规避苹果为iPhone推出的新隐私条款，而TikTok不敢冒进。假如法院解除了禁令威胁，它“可能会觉得自己可以猛踩油门了”，施穆利克指出。同时它也可能止住高管人才的流失。从迪士尼被挖角到TikTok的前首席执行官凯文·梅耶尔（Kevin Mayer）在特朗普试图封杀该应用期间离职。首席运营官瓦妮莎·帕帕斯（Vanessa Pappas）在去年离职。现在，首席法律顾问埃里希·安德森（Erich Andersen）据说也正准备离职。

不论法庭最终如何判决，TikTok已经在掂量有哪些国家可能效仿美国的强硬做法。2020年，印度在与中国发生边境冲突后对TikTok（及多个其他中国应用）实施禁令，TikTok在那里有两亿用户。印度尼西亚和巴基斯坦等国曾实施短期禁令，后来撤销。塔利班在阿富汗重新掌权后，不消说，也封禁了TikTok。

更有利可图的市场目前看来是安全的。还没有哪个欧洲大国要求出售TikTok。但是，欧洲在涉华安全问题上历来最终会追随美国，比如对中国

电信设备制造商华为迟来的打压。各国采取行动的意愿在一定程度上取决于它们与美国安全关系的紧密程度。美国的“五眼联盟”伙伴（澳大利亚、英国、加拿大和新西兰）已禁止在政府设备上使用TikTok。

| 口型不同步

假如中国选择报复，那么进一步限制TikTok可能就不止会扰乱社交媒体市场了。中国在4月发出了一个警告信号：它以国家安全为由，禁止应用商店提供WhatsApp和Threads这两个Meta产品。中国可能会给其他许多美国大公司制造麻烦。特斯拉正苦于在中国市场销量下滑的问题。苹果的iPhone在中国销量不断下降。由于中国鼓励其智能手机制造商使用国产芯片，AMD等美国芯片制造商也受到影响。美国可能发现，禁掉一款短视频应用会带来长远后患。 ■



AI's next top model

Large language models are getting bigger and better

Can they keep improving forever?

IN AI-LAND, technologies move from remarkable to old hat at the speed of light. Only 18 months ago the release of ChatGPT, OpenAI's chatbot, launched an AI frenzy. Today its powers have become commonplace. Several firms (such as Anthropic, Google and Meta) have since unveiled versions of their own models (Claude, Gemini and Llama), improving upon ChatGPT in a variety of ways.

That hunger for the new has only accelerated. In March Anthropic launched Claude 3, which bested the previous top models from OpenAI and Google on various leaderboards. On April 9th OpenAI reclaimed the crown (on some measures) by tweaking its model. On April 18th Meta released Llama 3, which early results suggest is the most capable open model to date. OpenAI is likely to make a splash sometime this year when it releases GPT-5, which may have capabilities beyond any current large language model (LLM). If the rumours are to be believed, the next generation of models will be even more remarkable—able to perform multi-step tasks, for instance, rather than merely responding to prompts, or analysing complex questions carefully instead of blurting out the first algorithmically available answer.

For those who believe that this is the usual tech hype, consider this: investors are deadly serious about backing the next generation of models. GPT-5 and other next-gen models are expected to cost billions of dollars to train. OpenAI is also reportedly partnering with Microsoft, a tech giant, to build a new \$100bn data centre. Based on the numbers alone, it seems as though the future will hold limitless exponential growth. This chimes with a view shared by many AI researchers called the “scaling hypothesis”,

namely that the architecture of current LLMs is on the path to unlocking phenomenal progress. All that is needed to exceed human abilities, according to the hypothesis, is more data and more powerful computer chips.

Look closer at the technical frontier, however, and some daunting hurdles become evident.

| *Beauty's not enough*

Data may well present the most immediate bottleneck. Epoch AI, a research outfit, estimates the well of high-quality textual data on the public internet will run dry by 2026. This has left researchers scrambling for ideas. Some labs are turning to the private web, buying data from brokers and news websites. Others are turning to the internet's vast quantities of audio and visual data, which could be used to train ever-bigger models for decades. Video can be particularly useful in teaching AI models about the physics of the world around them. If a model can observe a ball flying through the air, it might more easily work out the mathematical equation that describes the projectile's motion. Leading models like GPT-4 and Gemini are now "multimodal", capable of dealing with various types of data.

When data can no longer be found, it can be made. Companies like Scale AI and Surge AI have built large networks of people to generate and annotate data, including PhD researchers solving problems in maths or biology. One executive at a leading AI startup estimates this is costing AI labs hundreds of millions of dollars per year. A cheaper approach involves generating "synthetic data" in which one LLM makes billions of pages of text to train a second model. Though that method can run into trouble: models trained like this can lose past knowledge and generate uncreative responses. A more fruitful way to train AI models on synthetic data is to have them learn through collaboration or competition. Researchers call this "self-play". In 2017 Google DeepMind, the search giant's AI lab, developed a model called

AlphaGo that, after training against itself, beat the human world champion in the game of Go. Google and other firms now use similar techniques on their latest LLMs.

Extending ideas like self-play to new domains is hot topic of research. But most real-world problems—from running a business to being a good doctor—are more complex than a game, without clear-cut winning moves. This is why, for such complex domains, data to train models is still needed from people who can differentiate between good and bad quality responses. This in turn slows things down.

| *More silicon, but make it fashion*

Better hardware is another route to more powerful models. Graphics-processing units (GPUs), originally designed for video-gaming, have become the go-to chip for most AI programmers thanks to their ability to run intensive calculations in parallel. One way to unlock new capabilities may lie in using chips designed specifically for AI models. Cerebras, a chipmaker based in Silicon Valley, released a product in March containing 50 times as many transistors as the largest GPU. Model-building is usually hampered by data needing to be continuously loaded on and off the GPUs as the model is trained. Cerebras's giant chip, by contrast, has memory built in.

New models that can take advantage of these advances will be more reliable and better at handling tricky requests from users. One way this may happen is through larger “context windows”, the amount of text, image or video that a user can feed into a model when making requests. Enlarging context windows to allow users to upload additional relevant information also seems to be an effective way of curbing hallucination, the tendency of AI models to confidently answer questions with made-up information.

But while some model-makers race for more resources, others see signs

that the scaling hypothesis is running into trouble. Physical constraints—insufficient memory, say, or rising energy costs—place practical limitations on bigger model designs. More worrying, it is not clear that expanding context windows will be enough for continued progress. Yann LeCun, a star AI boffin now at Meta, is one of many who believe the limitations in the current AI models cannot be fixed with more of the same.

Some scientists are therefore turning to a long-standing source of inspiration in the field of AI—the human brain. The average adult can reason and plan far better than the best LLMs, despite using less power and much less data. “AI needs better learning algorithms, and we know they’re possible because your brain has them,” says Pedro Domingos, a computer scientist at the University of Washington.

One problem, he says, is the algorithm by which LLMs learn, called backpropagation. All LLMs are neural networks arranged in layers, which receive inputs and transform them to predict outputs. When the LLM is in its learning phase, it compares its predictions against the version of reality available in its training data. If these diverge, the algorithm makes small tweaks to each layer of the network to improve future predictions. That makes it computationally intensive and incremental.

The neural networks in today’s LLMs are also inefficiently structured. Since 2017 most AI models have used a type of neural-network architecture known as a transformer (the “T” in GPT), which allowed them to establish relationships between bits of data that are far apart within a data set. Previous approaches struggled to make such long-range connections. If a transformer-based model were asked to write the lyrics to a song, for example, it could, in its coda, riff on lines from many verses earlier, whereas a more primitive model would have forgotten all about the start by the time it had got to the end of the song. Transformers can also be run on many processors at once, significantly reducing the time it takes to train

them.

Albert Gu, a computer scientist at Carnegie Mellon University, nevertheless thinks the transformers' time may soon be up. Scaling up their context windows is highly computationally inefficient: as the input doubles, the amount of computation required to process it quadruples. Alongside Tri Dao of Princeton University, Dr Gu has come up with an alternative architecture called Mamba. If, by analogy, a transformer reads all of a book's pages at once, Mamba reads them sequentially, updating its worldview as it progresses. This is not only more efficient, but also more closely approximates the way human comprehension works.

LLMs also need help getting better at reasoning and planning. Andrej Karpathy, a researcher formerly at OpenAI, explained in a recent talk that current LLMs are only capable of “system 1” thinking. In humans, this is the automatic mode of thought involved in snap decisions. In contrast, “system 2” thinking is slower, more conscious and involves iteration. For AI systems, that may require algorithms capable of something called search—an ability to outline and examine many different courses of action before selecting the best one. This would be similar in spirit to how game-playing AI models can choose the best moves after exploring several options.

Advanced planning via search is the focus of much current effort. Meta's Dr LeCun, for example, is trying to program the ability to reason and make predictions directly into an AI system. In 2022 he proposed a framework called “Joint Embedding Predictive Architecture” (JEPA), which is trained to predict larger chunks of text or images in a single step than current generative-AI models. That lets it focus on global features of a data set. When analysing animal images, for example, a JEPA-based model may more quickly focus on size, shape and colour rather than individual patches of fur. The hope is that by abstracting things out JEPA learns more efficiently

than generative models, which get distracted by irrelevant details.

Experiments with approaches like Mamba or JEPA remain the exception. Until data and computing power become insurmountable hurdles, transformer-based models will stay in favour. But as engineers push them into ever more complex applications, human expertise will remain essential in the labelling of data. This could mean slower progress than before. For a new generation of AI models to stun the world as ChatGPT did in 2022, fundamental breakthroughs may be needed. ■



AI的下一个顶级模型

大语言模型越来越大，越来越好

它们能永远进步吗？【深度】

在AI世界里，技术从举世瞩目到老套过气的速度堪比光速。不过18个月前，OpenAI的聊天机器人ChatGPT的发布掀起了一场AI狂潮。今天，它的本事已经变得司空见惯。几家公司（如Anthropic、谷歌和Meta）继OpenAI之后推出了自己的这类模型（Claude、Gemini和Llama），在ChatGPT的基础上朝着不同方向提升。

这种对新模型的渴望不断加速。3月，Anthropic推出了Claude 3，在各种排行榜上击败了之前OpenAI和谷歌开发的顶级模型。4月9日，OpenAI通过调整其模型夺回了（以某些指标衡量的）桂冠。4月18日，Meta发布了Llama 3，初步结果表明它是迄今为止功能最强大的开源模型。OpenAI很可能在今年某个时候制造轰动，它可能发布能力超过当前任何大语言模型的GPT-5。如果传言可信，那么下一代模型还要更加惊人——例如，能够执行多步骤任务，而不仅仅是就提示给出回应，或者能仔细分析复杂的问题，而不是直接给出通过算法得出的第一个答案。

对于那些认为这不过是那种常见的科技炒作的人，请想想这一点：投资者对支持下一代模型极度认真。训练GPT-5和其他下一代模型预计将花费数十亿美元。据报道，OpenAI还在与科技巨头微软合作新建一个造价1000亿美元的数据中心。仅从数字来看，未来似乎将保持无限的指数增长。这与许多AI研究人员所认同的“规模假说”不谋而合，即目前大语言模型的架构正在开启惊人的进步。该假说认为，要超越人类能力，它所需要的只是更多的数据和更强大的计算芯片。

然而，更近一步观察技术前沿，就很容易看到一些令人生畏的障碍。

| 好看还不够

数据很可能是最直接的瓶颈。研究机构EpochAI估计，到2026年，公共互

联网上的高质量文本数据将枯竭。研究人员为此忙不迭地想办法。一些实验室正在转向私人网络，从经纪公司和新闻网站那里购买数据。其他人则转向互联网上大量的音频和视频数据，这些数据可以在几十年里用来训练越来越大的模型。在训练AI模型了解周围世界的物理规律方面，视频特别有用。如果一个模型可以观察一个在空中飞行的球，它可能会更容易地得出描述抛物线运动的数学方程式。像GPT-4和Gemini这样的领先模型现在是“多模态”的，能够处理各种类型的数据。

当无法再找到数据时，还可以制作数据。像ScaleAI和SurgeAI这样的公司已经建立了庞大的人员网络来产生和注释数据，其中包括能解决数学或生物学问题的博士研究生。一家领先的AI创业公司的一位高管估计，AI实验室每年要为此花费数亿美元。一种更便宜的方法涉及生成“合成数据”，就是用一个大型语言模型制作数十亿页的文本来训练第二个模型。不过这种方法可能会遇到麻烦：这样训练出来的模型可能会缺少过去的知识，产生的响应缺乏创造性。用合成数据训练AI模型的一种更有效的方法是让它们通过协作或竞争来学习。研究人员称之为“自我对抗”。2017年，搜索巨头谷歌的人工智能实验室DeepMind开发了一个名为AlphaGo的模型，该模型在与自己对抗后，在围棋比赛中击败了人类世界冠军。谷歌和其他公司现在使用类似的方法来训练最新的大语言模型。

将自我对抗等想法扩展到新的领域是热门研究课题。但大多数现实世界的问题——从经营企业到成为一名好医生——都比下棋更复杂，没有明确的获胜招式。这就是为什么对于这些复杂的领域，仍然需要能够区分响应质量好坏的人提供数据来训练模型。这反过来又减慢了速度。

| 多用硅，但要做得时髦

更好的硬件是训练更强大模型的另一条途径。图形处理单元（GPU）最初是为电子游戏设计的，由于它们能够并行运行密集计算，因此已成为大多数AI程序员的首选芯片。解锁新功能的一种方法可能是使用专为AI模型设计的芯片。硅谷芯片制造商Cerebras在3月发布了一款产品，其包含的晶体管数量是最大GPU的50倍。在训练模型时，GPU需要不停地加载卸载数据，这往往会妨碍模型构建。相比之下，Cerebras的巨型芯片内置了内

存。

能够利用这些进展的新模型将更为可靠，能更好地处理来自用户的棘手的请求。要做到这一点，一种方式是使用更大的“上下文窗口”，也就是用户在发出请求时可以输入更多的文本、图像或视频到模型中。扩大上下文窗口来让用户上传更多的相关信息似乎也可以有效遏制幻觉，幻觉就是指人工智能模型往往会自信地用编造的信息回答问题。

但是，当一些模型开发者竞相争取更多资源时，其他人则看到了规模假说陷入困境的迹象。物理局限（例如内存不足或能源成本上升）是设计更大模型的实际限制。更令人担忧的是，目前尚不清楚扩大上下文窗口是否足以继续取得进展。目前在Meta任职的明星AI专家杨立昆（Yann LeCun）也和许多人一样，认为当前AI模型的局限性无法通过一味扩大规模来解决。

因此，一些科学家正在求助于AI领域中一个长久的灵感来源——人脑。普通成年人的推理和计划能力远远超过最好的大语言模型，而他们消耗的能量更少，所用数据要少得多。“AI需要更好的学习算法，我们知道这是可能的，因为人的大脑里就有这些东西。”华盛顿大学的计算机科学家佩德罗·多明戈斯（Pedro Domingos）说。

他说，一个问题是大语言模型学习所用的“反向传播”算法。所有大语言模型都是一层层叠加的神经网络，它们接收输入并进行转换以预测输出。当大语言模型处于学习阶段时，它会将其预测与其训练数据中存在的现实情况进行比较。如果存在差异，算法会对网络的每一层进行微调，以改进未来的预测。这使得它要进行密集且不断递增的运算。

当今大语言模型所用的神经网络结构效率低下。自2017年以来，大多数AI模型都使用了一种称为transformer（即GPT中的“T”）的神经网络架构，这使它们能够在数据集中相距很远的数位之间建立关系。以前的方法很难建立这种远程连接。例如，如果一个基于transformer的模型被要求为一首歌写歌词，它可以在结尾处中重复之前写下的歌词，而一个更原始模型在写到结尾时会忘记一开始写的东西。Transformer也可以同时在多个处

理器上运行，从而大大减少了训练它们所需的时间。

然而，卡内基梅隆大学的计算机科学家阿尔伯特·顾（Albert Gu，音译）认为，transformer可能很快就会过气。扩展其上下文窗口在计算效率上非常低：如果输入翻一番，那么处理输入所需的计算量就要翻两番。顾与普林斯顿大学的泰瑞·达奥（Tri Dao）一起提出了另一种名为Mamba的架构。打个比方，如果说transformer是同时阅读一本书的所有页面，那么Mamba就是逐页阅读，并随着阅读的推进更新其世界观。这不仅更有效率，也更接近人类的理解方式。

另外还需要帮助大语言模型提高推理和规划能力。前OpenAI研究人员安德烈·卡帕西（Andrej Karpathy）在最近的一次演讲中解释说，目前的大语言模型只能进行“系统1”思维。对人类来说，这是快速决策中用到的自动思维模式。相比之下，“系统2”思维的速度更慢，更加慎重，涉及反复思考。对于AI系统来说，这可能需要能够进行被称为“搜索”的算法，即能在选择最佳行动方案之前概览和检验许多不同的行动方案。这在本质上类似于下棋的AI模型在探索多个着法后如何选择最佳一着。

通过搜索做出高级规划是当前许多工作的重点。例如，Meta的杨立昆正试图设计程序将推理和预测能力直接加入到AI系统中。2022年，他提出了一个名为“联合嵌入预测架构”（JEPA）的框架，该框架经过训练，可以比当前的生成式AI模型在单一步骤中预测更大的文本或图像块。这使它能够专注于数据集的全局特征。例如，在分析动物图像时，基于JEPA的模型可以更快地关注大小、形状和颜色，而不是单块的皮毛。相比生成式模型会被不重要的细节分散注意力，JEPA有望通过把事物抽象化来更高效地学习。

对Mamba或JEPA等方法的实验仍是少数派。在数据和算力成为不可逾越的障碍之前，基于transformer的模型将继续受到青睐。但是，在工程师将它们推向越来越复杂的应用之时，人类的能力在数据标记中仍然至关重要。这可能意味着进展会比以前更慢。新一代AI模型若想像ChatGPT在2022年那样震惊世界，可能需要根本性的突破。■



Manufacturing miracles

China's better economic growth hides reasons to worry

The country's leaders are too complacent about deflation

WHEN CHINA'S leaders set an economic-growth target of “around” 5% for this year, the goal was widely agreed to be ambitious. Now the country looks increasingly likely to meet it. Several foreign banks have raised their forecasts. Data released on April 16th show the economy grew by 5.3% in the first quarter, compared with a year earlier—quicker than expected and faster than the target requires.

How is this happening? Countries at China's stage of development often shift towards services. Yet China's leaders have a soft spot for “hard” output. Xi Jinping, the country's ruler, sees manufacturing as a source of both prosperity and security. He covets what officials call a “complete” industrial chain that would free China from reliance on foreign powers for vital technological inputs. His latest five-year plan aims to stop the steady decline in manufacturing's share of GDP.

The first quarter was consistent with that goal. Manufacturing output grew by 6.7% compared with a year ago, faster than the overall economy. High-tech manufacturing fared even better. China's leaders have talked a lot about the need to cultivate “new quality productive forces”, buzzwords that appeared in the monthly statistical press release for the first time.

But even as China weans itself off foreign suppliers, it remains reliant on foreign buyers. The volume of exports grew by 14% in the first quarter compared with a year earlier, according to Zhiwei Zhang of Pinpoint Asset Management. Falling prices and a competitive currency have helped. America's Bureau of Labour Statistics reckons the price of goods from China fell by 2.9% year-on-year in the first quarter. That is the third-steepest drop

on record.

China cannot rely on strong exports for long without provoking a protectionist backlash from its trading partners. Olaf Scholz, chancellor of Germany, raised fears about Chinese overcapacity when he met Mr Xi in Beijing on April 16th. Germany used to benefit from China's economic progress. It sold sophisticated industrial goods to China, even as China's manufacturers conquered lower-end markets around the world. Now the two countries have become rivals in many industries Germany holds dear, including chemicals, machinery and, of course, cars.

China's reliance on markets abroad reflects some enduring weaknesses at home. Retail sales were surprisingly poor in March. Consumer confidence remains low. And the property market's misery continues. The price of new flats in 70 of China's biggest cities fell by 2.2% on average in March compared with a year earlier, the steepest drop since 2015, according to Reuters, a news agency. Sales of newly built residential housing fell by over a fifth.

The slump in China's property market has contributed to falling prices in many related parts of the economy, such as building materials and housing appliances. That has deepened deflation's grip on the economy. Factory-gate prices have now fallen for 18 months in a row. Consumer price inflation, after a brief uptick during the lunar new year holiday in February, remained near zero in March. Declining prices are a double-edged sword, as Ting Lu of Nomura, a bank, has pointed out. They have increased China's competitiveness abroad, which is one reason why the country's exports have been surprisingly strong. But if deflation persists it could erode revenues, making debts harder to bear. It might also force companies to cut wages, which would do nothing to restore household morale or spending.

For all the paranoia of China's leaders, they seem worryingly complacent

about the danger of deflation. Perhaps they view it as a blip, which should not distract them from long-term aims to fortify China against shifts in the global balance of power—what Mr Xi calls “changes unseen in a century”. Falling prices can, though, turn a passing downturn into a protracted slump. This week’s figures showed that China’s GDP deflator, a broad measure of prices, has fallen for four quarters in a row. That has not happened since 1999. Or to put it in terms Mr Xi might appreciate, it is a change unseen this century. ■



制造奇迹

中国经济开门红之下有隐忧

中国领导人对通缩太不在意

当中国领导人将今年的经济增长目标定在5%“左右”时，人们普遍认为这定得偏高。现在看来，中国越来越有可能实现这一目标。几家外国银行已经上调了自己的预测。4月16日公布的数据显示，与去年同期相比，中国经济在一季度增长了5.3%——高于预期，也高于实现目标所需达到的增速。

这是如何做到的？如果一个国家到了中国目前的发展阶段，它通常会向服务业转型。然而中国领导人对“硬”产出情有独钟。最高领导人习近平把制造业视为繁荣和安全之源。他渴望实现官员所称的“完整”产业链，让中国在关键技术投入方面摆脱对外部强国的依赖。在最新的五年计划中，他要停止制造业在GDP中的占比不断下滑的局面。

一季度的表现符合这一目标。制造业产出与去年同期相比增长了6.7%，高于整体经济增速。高科技制造业的数字还要更好。中国领导人多次谈到要发展“新质生产力”，这个热词此次首次出现在月度统计数据发布中。

但是，在中国逐渐摆脱对外国供应商依赖的同时，对外国买家的依赖却没有降低。保银投资的张智威称，一季度出口额同比增长了14%。价格下跌和有竞争力的汇率都起到了推动作用。美国劳工统计局（Bureau of Labour Statistics）估计，一季度来自中国的商品价格同比下降了2.9%。这是有记录以来的第三大降幅。

中国不可能长期依赖强劲出口而不致引发贸易伙伴的保护主义反弹。4月16日，德国总理朔尔茨在北京会见习主席时，提出了对中国产能过剩的担忧。德国曾是中国经济发展的受益者。当中国制造商攻克全球低端市场之时，德国向中国出售精密工业品。如今，两国在许多德国最看重的行业里已经成了竞争对手，包括化学品、机械，当然还有汽车。

中国对海外市场的依赖反映了国内一些长期存在的弱点。3月的零售数据出人意料地糟糕。消费者信心仍然低迷。房地产市场的惨淡还在继续。据路透社报道，今年3月，中国70个大城市的新房价格与去年同期相比平均下跌2.2%，是自2015年以来的最大跌幅。新建住房销售面积下降超过五分之一。

中国房地产市场低迷导致许多上下游领域价格下跌，例如建筑材料和家用电器。这继而加深了经济的通货紧缩压力。工业品出厂价已连续18个月下跌。消费者价格涨幅在2月农历新年期间短暂上扬后，在3月仍接近于零。正如野村证券的陆挺所言，价格下跌是一把双刃剑。它提升了中国在海外的竞争力，这也是中国出口惊人强劲的原因之一。但如果通缩持续下去，就有可能侵蚀营收，使债务更加难以承受，还可能迫使企业削减工资，这显然不利于恢复居民信心或消费支出。

尽管中国领导人有种种疑惧，但他们对通缩的危险似乎倒满不在乎，这让人担忧。也许他们认为这只是微不足道的小插曲，不应该为此分散对长期目标的注意力，也就是要增强中国的实力，以抵御全球力量平衡的变化——习称之为“百年未有之大变局”。然而，价格下跌可能会将短暂低迷转变为长期萧条。最近的数据显示，中国的GDP平减指数（广义的价格指标）已经连续四个季度下滑。这种情况自1999年以来还没有发生过。借用习的话来说，这也是本世纪未有之大变局。■



Degrees and difficulty

Why so many Chinese graduates cannot find work

Our number-crunching suggests that their plight could be much worse than previously thought

AROUND THIS time each year companies visit university campuses in China looking for potential employees. This year the mood is grim. At a job fair in Wuhan a firm was looking to hire management trainees—but it wanted only elite graduates and offered just 1,000 yuan (\$140) per month, claimed a post that went viral on social media. At a fair in Jilin most of the advertised positions required advanced degrees, said a soon-to-be graduate online. “Next time don’t bother inviting us.” Another griped that firms are not hiring. The recruitment process is “a lie”, she wrote.

The data paint a similarly bleak picture. The unemployment rate for people aged 16 to 24 in cities reached a record high of 21.3% last June. That was perhaps too embarrassing for the government, so it stopped publishing the data series while it rejigged its calculation to exclude young people seeking jobs while studying. (America, Britain and many other countries include such students when calculating their rates.) The new numbers are lower, but still depressing: in March 15.3% of young people in cities were unemployed. That’s nearly three times the overall jobless rate.

For young graduates the situation is probably even more dire. China does not release an unemployment rate for this cohort. But we combed through data from the country’s decennial census and its statistical yearbooks in order to produce an estimate. By our calculations (including students who are seeking jobs), the unemployment rate for 16- to 24-year-olds with a university education was 25.2% in 2020, the last year for which census data are available. That was 1.8 times the unemployment rate for all young people at the time.

It could be that things have got better since 2020 or that the variables affecting our calculations have changed in unpredictable ways. But it is also possible that things have got worse. To simplify, if we assume that the proportional relationship from 2020 still holds, over a third of young graduates might be unemployed today.

One reason to believe that things are not improving is that graduates as a share of unemployed youth are increasing faster than might be explained by broader demographic trends (see chart 1). Graduates of universities and vocational and technical colleges accounted for 70% of the unemployed young in 2022, up from 9% two decades ago. As a percentage of the youth population, those graduates amounted to 47% in 2020.

China's sluggish economy is at least partly to blame. Demand for graduates has stagnated. Meanwhile, the supply of them is growing. This year nearly 12m students are expected to graduate from higher-education institutions, an increase of 2% compared with last year. Between 2000 and 2024 the number of Chinese graduates per year grew more than tenfold (see chart 2).

The trend can be traced back to Min Tang, a Chinese economist who proposed expanding enrolment in higher education as a way of dealing with the Asian financial crisis of the late 1990s. Such a policy would postpone young people's entry into the job market and stimulate the economy by way of education spending, he said. The government adopted his plan, which coincided with societal changes that pushed in the same direction. Children born under China's one-child policy began to come of age in 1999. With family size limited, parents had more to invest in each child—and more of an incentive to encourage their studies, since these children were expected to provide for their parents in old age.

As demand for slots increased, universities grew in size and number. Laws passed in the early 2000s allowed companies to get in on the action.

Privately run institutions, called minban daxue, charge substantially higher fees than public universities and have an incentive to admit ever more students. Enrolment at these schools has ballooned, increasing by 560% since 2004. Back then one in ten students in college or university studied at a minban daxue. Now one in four does (see chart 3).

Minban daxue tend to require lower scores on China's university-entrance exam, the gaokao, than public institutions. But the acceptance rate at all colleges and universities has been rising. Before 1999 less than a quarter of gaokao-takers were accepted by these institutions. Today most make it in (see chart 4).

The rising number of graduates might not be such a problem if they were learning skills desired by employers. But Chinese companies complain that they cannot find qualified candidates for their open positions. Part of the problem are low-quality minban daxue. Yet the skills mismatch extends across higher education. For example, the number of students studying the humanities is growing even though demand for such graduates is much lower than that for specialists in other fields.

Some students are trying to dodge the tough private-sector job market. The number of people sitting for China's civil-service exam hit a record high of 2.3m in 2024, a 48% increase year on year. Others are pursuing postgraduate studies. The number of master's and doctoral students has increased by so much that some campuses have run out of housing.

Unable to find work befitting their degrees, a number of graduates are settling for low-skilled jobs, such as delivering food. Last year a memo from an airport in Wenzhou noted that it had hired architects and engineers to be its groundskeepers and bird-control personnel.

Xiaoguang Li of Xi'an Jiaotong University and Yao Lu of Columbia

University have studied underemployment in China. Using national survey data, they found that 25% of workers between the ages of 23 and 35 were overqualified for their job in 2021, up from 21% in 2015. The problem is likely to get worse, says Ms Lu, as graduates facing unemployment have no choice but to accept menial work.

As a result of all this, the returns from pursuing higher education seem to be falling. In a working paper published last year, researchers led by Eric Hanushek of Stanford University found that in China the wage premium associated with higher education dropped from 72% to 34% between 2007 and 2018 for those under the age of 35.

In 2008 an official in the education ministry seemed to admit that the state had made a mistake in expanding college and university enrolment so quickly. But the ministry quickly backtracked. Today the government seems to care more about the size of the education system than the quality of it. Sixty-one new colleges and universities opened in China last year. “Our country has built the largest higher-education system in the world,” boasted the People’s Daily, a party mouthpiece.

In his state-of-the-nation speech last month Li Qiang, the prime minister, at least paid lip service to the idea of making sure more graduates were learning the skills needed in sectors such as advanced manufacturing and elderly care. But many will continue to find that their degree is not a ticket to a good job. Told for years that higher education was a ladder to a better life, their frustrations are growing. ■



学位与难题

为什么那么多中国毕业生找不到工作

我们的数据分析表明，他们的困境可能远比先前的预计严重

每年这个时候，企业都会开展校园招聘。今年的校招阴云笼罩。在武汉的一场招聘会上，一家公司招聘管理培训生，但他们只招收优秀毕业生，而且每月工资只有1000元，此事在社交媒体上迅速传播开来。在吉林的一场招聘会上，大多数职位要求硕博学位，一名即将毕业的学生在网上说。“下次干脆别叫我们去了。”另一名学生抱怨说现在公司根本不招人。招聘就是“做做样子的”，她写道。

数据反映出的情况同样令人沮丧。去年6月，16至24岁城镇人群失业率达到了21.3%，创下历史新高。这可能让政府太过尴尬，因此不再继续发布该数据，同时调整计算方法，把正在找工作的在校生排除在统计之外。

（美国、英国和许多其他国家在计算失业率时会包括这些学生）。新计算得出的失业率数字降低了，但仍令人沮丧。今年3月，城镇青年失业率为15.3%，是整体失业率的将近三倍。

对于年轻的大学毕业生来说，情况可能更加严峻。中国没有专门公布这个群体的失业率。但为了做出估算，我们查阅了中国十年一次的人口普查和统计年鉴数据。根据我们的计算（包括正在找工作的在校生），在有最新普查数据可用的2020年，接受过大学教育的16至24岁的青年失业率为25.2%。这个数字是当时青年整体失业率的1.8倍。

可能2020年以后的情况有所改善，或者影响我们计算的变量发生了不可预测的改变。但情况也同样有可能变得更糟了。简化来说，如果我们假设上述两个失业率数据的比例关系仍然保持在2020年的水平，那么今天超过三分之一的年轻毕业生可能处于失业状态。

让人认为情况并没有在改善的一个原因是，毕业生在失业青年中占比的增长快过或许可以用更大范围人口结构趋势来解释的程度（见图表1）。

2022年，大学和职业技术学院毕业生占失业青年的70%，二十年前为9%。而毕业生占青年人口的比例在2020年为47%。

这至少在一定程度上是因为中国经济增长缓慢。对毕业生的需求停滞不前，同时供应却在不断增长。预计今年将有近1200万高校毕业生，较去年增长2%。2000年至2024年间，中国每年毕业的学生人数增长了十多倍（见图表2）。

这一趋势可以追溯到建议高校扩招的中国经济学家汤敏，他当时将扩招作为应对1990年代末亚洲金融危机的一种方式。他说，扩招将推迟年轻人口进入就业市场，并通过教育支出刺激经济。政府采纳了他的计划，此时又恰逢一些社会变迁朝着同样的方向使力。在中国的独生子女政策下出生的孩子从1999年开始步入成年。由于家庭规模受到限制，父母能把更多资金用到每个孩子身上，并且更有动力鼓励他们学习，因为父母老了要指望孩子养老。

随着上大学的需求增加，大学也越来越多、越来越大。2000年代初通过的法律允许企业参与办学。被称为民办大学的私立高教机构收取的学费比公立大学高很多，也就有动力录取更多学生。这些学校录取的人数激增，自2004年以来增长了560%。那时，每十个大学生里有一个在民办大学就读。现在，民办大学的学生占了大学生总数的四分之一（见图表3）。

民办大学的高考录取分数线往往比公办大学低。但所有高等院校的录取率都在上升。1999年之前，只有不到四分之一的高考生能上大学。如今，大多数人都能被录取（见图表4）。

如果学生们在学校学的是雇主需要的技能，那毕业生人数增加可能还不是那么大的问题。但中国企业抱怨无法找到能胜任它们空缺岗位的人选。民办大学教学质量差是原因之一。但整个高等教育领域都存在技能不匹配的情况。例如，尽管对人文学科毕业生的需求远低于对其他领域专业人才的需求，但文科生的数量仍在增长。

一些学生试图绕开艰难的私营部门就业市场。2024年，参加中国公务员考

试的人数创下230万的历史新高，同比增长48%。其他人选择考研，硕士和博士生人数大幅增加，一些学校的宿舍已经挤爆。

一些毕业生无法找到与其学位相匹配的工作，只能接受送餐等低技能工作。去年温州机场的一份公示显示，他们录取了建筑和工程专业的毕业生来担任地勤和驱鸟员。

西安交通大学的李晓光和哥伦比亚大学的陆瑶考察了中国的就业不足问题。他们通过全国调查数据发现，在2021年，23至35岁之间的就业人群中，有25%的人大材小用，高于2015年的21%。陆瑶表示，这个问题很可能会愈趋严重，因为面临失业的毕业生别无选择，只能接受低技能低报酬的工作。

这一切的结果是，接受高等教育的回报似乎正在下降。在去年发表的一份工作论文中，斯坦福大学的埃里克·哈努谢克（Eric Hanushek）领导的研究人员发现，在中国，35岁以下人群通过高等教育所得的工资溢价从2007年的72%下降到了2018年的34%。

2008年，教育部的一位官员似乎承认了政府太急促扩招是一个错误。但教育部很快又予以否认。如今，政府似乎更关心高等教育体系的规模而不是质量。去年，中国新开办了61所高等院校。“我国已建成世界最大规模高等教育体系。”党报《人民日报》夸耀道。

在上个月的政府工作报告中，总理李强至少在口头上表达了确保更多毕业生学到技能以适应先进制造和养老照护等领域人才需求的想法。但许多人仍将发现，他们的学位证书并不能保证他们找到好工作。多年来，他们一直被灌输高等教育是通向更美好生活的阶梯，现在他们的挫败感正在与日俱增。■



OK Zoomer

Reasons to be cheerful about Generation Z

They are not doomed to be poor and anxious

A VAST COHORT is coming of age. Globally, some 2bn people were born between 1997 and 2012, and so are part of “Generation Z”. In America and Britain this group makes up a fifth of the population, rivalling the share of baby-boomers; in India and Nigeria the young far outnumber the old. For each generation there is a simple narrative: that boomers were shaped by post-war plenty, for example, or millennials by the financial crisis of 2007-09. For Gen Z the popular view is that smartphones have made them miserable and they will live grimmer lives than their elders.

More and more people in the West tell pollsters that today’s children will be worse off than their parents. Youngsters themselves worry about everything from the difficulty of buying a home to the looming dangers from climate change. Social scientists fret that Gen Z-ers, having spent their formative years doomscrolling and suffering from FOMO, are now gripped by an epidemic of anxiety and depression. Politicians in America and Britain are mulling banning smartphones and restricting social media for the under-16s; parents and schoolteachers everywhere are trying to police screen time.

All this can make it hard to feel optimistic about Gen Z. But when you look around the world, and at a wider set of measures, the Zoomers are far from doomed. In many respects, they are doing rather well.

For a start, the popular narrative makes an important omission: the roughly four-fifths of the world’s 12- to 27-year-olds who live in emerging economies. Thanks to growth and the spread of technology, youngsters in places like Jakarta, Mumbai or Nairobi are far better off than their parents

were. They are richer, healthier and more educated; those who have smartphones are better informed and connected. Small wonder that, in a survey by the UN in 2021, the young in emerging economies were more optimistic than those in the rich world.

Yet in some places there is a fear that the rapid progress of recent decades might fail to repeat itself. That anxiety is evident in China. Thanks to economic uncertainty and an emphasis on quantity over quality in higher education, over a third of degree-holders there may be unemployed.

In the rich world the picture is rosier than people think. Those of Gen Z who are in work—and in America there are nearly as many of them in workplaces now as there are boomers—are doing nicely. Red-hot demand for workers helps, as does the fact that Gen Z-ers are wisely acquiring marketable skills. More of them are pursuing science, engineering and medical degrees; the humanities have fallen out of favour.

Wages for Gen Z are rising at a much faster pace than they are for older workers, and the youth-unemployment rate across the rich world is at its lowest in decades. In America the income of the average Zoomer, after adjusting for taxes and transfers, comfortably exceeds that of a millennial or a Gen X-er at the same age, in real terms. True, housing affordability has worsened since the 1980s. But, thanks to Gen Z-ers' stronger wage growth, house prices as a multiple of earnings are roughly where they were for millennials a decade ago. And young people today are at least able to put more of their salaries into savings.

Already Gen Z-ers are transforming the world of work. They have bargaining power—and they know it. Many millennials came of age in the shadow of the global financial crisis; they felt so precarious that they were afraid to ask for pay rises. Gen Z seems to have fewer qualms about quitting for a better opportunity, or taking things slowly and enjoying life. Bosses,

unused to being on the back foot, complain. But older workers will be quietly thankful if overall pay and perks go up.

Gen Z will shape society in other ways, too. Young people's concern about climate change will, as they reach voting age, make states more likely to act. More broadly, Zoomers tell pollsters they want bigger government. They may change their minds when they have to pay more taxes—or they may not.

They are a serious bunch, less given to late nights, binge drinking and promiscuity than their elders were. There is a dark side to this. They socialise in person less, have less sex, and are more likely to say they are lonely. Reported rates of anxiety and depression are rising in much of the West. Some of this probably reflects a greater willingness to open up about mental health. But other factors play a role.

The extent to which social media fuel mental distress among the young is furiously debated. In the West the rise in anxiety coincides with the adoption of social media. However, hard evidence of causation is limited, and most of it comes from studies of adults in the rich world.

What is clear is that Generation Z has been at the sharp end of a technological revolution. The speed with which smartphones and then social-media apps were adopted around the world left users, and young people in particular, scrambling to figure out how best to navigate them. Social media have brought benefits, such as entertainment and connectivity, but also costs. Some content may be harmful, and time spent scrolling could have been devoted to study or sleep.

| *OK Zoomer*

Transformative technologies often have downsides. In the past, people have adapted: think of the seat belts and regulations that made cars less

deadly. Encouragingly, there are signs that social-media habits are already shifting as users weigh the costs and the benefits. Instead of posting about themselves publicly, for example, many are retreating to private groups on messaging apps. So far, the evidence that would justify a blanket ban on smartphones for the young does not exist, though schools should be free to bar them from classrooms, and parents are right to restrict screen time.

It is only natural for the old to worry about the young. If that leads to better mental-health treatment, or fewer restrictions on building more homes, well and good. But celebrate Gen Z's resourcefulness, and its successes, too.





【首文】Z世代很OK

有理由为Z世代感到高兴

他们并非注定要陷入贫穷和焦虑

一个庞大的群体正在步入成年。全球范围内，约有20亿人出生于1997年至2012年之间，因此属于“Z世代”。在美国和英国，这一群体占总人口的五分之一，和婴儿潮一代的占比相当；在印度和尼日利亚，Z世代的人数远远超过老辈人。世人对每一代人都有个简单的叙事，例如，婴儿潮一代成长于二战后的富足年代；千禧一代深受2007至2009年金融危机的影响。对于Z世代，普遍的看法是他们被智能手机弄得郁闷困苦，生活会比老辈人更糟糕。

在西方国家的各项民意调查中，越来越多的人认为，今天的孩子们的日子将比他们的父辈更难。年轻人自己则有发不完的愁——从买不起房到迫在眉睫的气候变化威胁。社会科学家担心，在“末日刷刷刷”中成长起来且饱受“FOMO”（错失恐惧）折磨的Z世代如今正广泛受到焦虑和抑郁的困扰。美国和英国的政客们正在考虑禁止16岁以下的青少年使用智能手机，并限制他们使用社交媒体；世界各地的家长和教师也在想办法管控他们使用电子设备的时间。

所有这些都让人很难看好Z世代。但是当你环顾世界，并以一套更宽泛的标准来观察，Z世代还远没有到注定悲催的地步。在许多方面，他们过得相当不错。

首先，上述关于Z世代的流行说法忽略了一个重要因素：全球12至27岁的年轻人有约五分之四生活在新兴经济体。得益于经济增长和技术的普及，雅加达、孟买或内罗毕等地的年轻人的境况比他们父辈当年好得多。他们更富有、更健康、受教育程度更高；使用智能手机的年轻人消息更灵通、交际更广泛。难怪联合国2021年的一项调查显示，新兴经济体的年轻人比富裕国家的更乐观。

不过在一些地方，人们担心近几十年来的飞速发展可能不会重演。这种忧虑在中国很明显。由于经济的不确定性以及高等教育重数量而轻质量，中国超过三分之一的高校毕业生可能找不到工作。

在富裕国家，真实情况比人们以为的要好。那些已经工作的Z世代——在美国，目前职场中的Z世代几乎和婴儿潮一代一样多——状况不错。这得益于旺盛的用工需求，也是因为Z世代自己也识时务，正在学习市场需要的技能。他们中更多人在攻读科学、工程和医学学位；人文学科已不再受青睐。

Z世代目前的薪资增速远快于职场中更年长的员工，而富裕国家的青年失业率处于几十年来的最低水平。在美国，按实值计算（经税收和转移支付调整后），普通Z世代的收入大大超过了千禧一代或X世代在他们这个年龄时的收入。诚然，自上世纪80年代以来，住房变得更加难以负担。但由于Z世代的工资增长更为强劲，他们面临的房价收入比与十年前千禧一代所面临的水平大致相当。而且如今的年轻人至少能够存下更多的工资。

Z世代已经开始改造职场。他们拥有议价能力——他们对此也心知肚明。许多千禧一代都是在全球金融危机的阴影下成长起来的，在职场上如履薄冰，不敢要求加薪。Z世代似乎没有那么多顾虑，他们理直气壮地为更好的机会跳槽，也不怕放慢工作节奏、享受生活。不习惯陷入被动的老板们颇有微词。但如果整体的薪资福利有所提高，年长的同事们会默默感激年轻人。

Z世代还会以其他方式塑造社会。随着这批年轻人达到投票年龄，他们对气候变化的关注会让政府更有可能采取行动。在总体政策倾向上，Z世代在民意调查中表示他们想要大政府。如果必须为此缴纳更多税款，他们或许会改主意——但也可能不会。

这一代人对生活比较严肃，不像他们的长辈在年轻时那样爱熬夜、酗酒和滥交。但这也有不好的一面。他们较少参与面对面社交，性生活更少，也更可能说自己孤独。在西方很多地区，焦虑症和抑郁症的报告比例正在上

升。这一定程度上可能是因为人们更愿意公开谈论自己的心理健康了，但也存在其他因素。

关于社交媒体在多大程度上加剧了年轻人的心理困扰，人们争论不休。在西方，焦虑症的上升恰逢社交媒体的普及。不过，并没有多少确凿的证据表明两者之间存在着因果关系，而且大部分证据都来自对富裕国家成年人的研究。

可以肯定的是，Z世代一直处于技术革命的风口浪尖。智能手机以及随后到来的社交媒体应用在全球普及之迅疾，让用户、尤其是年轻人忙不迭地要弄清该如何充分使用它们。社交媒体给人们带来娱乐和易于联系等好处的同时，也让人付出代价。有些内容可能是有害的，刷社交媒体的时间也许本可以用来学习或睡觉。

| Z世代很OK

变革性技术往往有其弊端。之前人们通过调整已经适应了过来：比如座椅安全带的使用和法规的出台减少了汽车的致命事故。令人欣慰的是，有迹象表明，用户在权衡利弊后已经开始改变自己使用社交媒体的习惯。例如，许多人不再公开发布自己的信息，而是退回到即时通讯应用的私人群组中。到目前为止，还没有证据能够证明对青少年一刀切式地禁用智能手机是英明之举，不过学校有权禁止学生将智能手机带入教室，而家长控制孩子使用电子设备时间的做法也是对的。

年长者为年轻人担忧是人之常情。如果这能让年轻人得到更好的心理健康治疗，或者能让政府放宽限制来建造更多房屋，那确实很好。但不妨也为Z世代的机敏和成就叫好吧。■



All by myself

The drawbacks—and benefits—of solitude

Three books examine the perils and pleasures of being alone

All the Lonely People. By Sam Carr. Picador; 256 pages; £11.99. To be published in America in December; \$28.99

Solitude. By Netta Weinstein, Heather Hansen and Thuy-vy Nguyen. Cambridge University Press; 300 pages; \$25.95 and £20

The Triumph of the Slippers. By Pascal Bruckner. Translated by Cory Stockwell. Wiley; 118 pages; \$19.95. Polity; £16.99

WILLIAM SHAKESPEARE, that great coiner, is given credit for the word “lonely”. Coriolanus, one of his heroes, compares going into exile to a “lonely dragon” retreating to his lair. The Roman general was talking about a physical state: someone who was lonely was simply alone.

Then, thanks to the Romantic poets, the word took on emotional overtones. Loneliness became a condition of the soul. For William Wordsworth, who famously “wandered lonely as a cloud”, the natural world offered a reprieve from negative feelings of isolation—a host of daffodils could provide “jocund company”.

By the early 20th century loneliness was considered one of the defining afflictions of urban life. Hannah Arendt, a political theorist, lamented that a feeling that was “once a borderline experience usually suffered in certain marginal social conditions like old age, has become an everyday experience of the ever-growing masses”.

Her concerns resonate today, as loneliness is frequently identified as a serious public-health problem, an epidemic even, that besets the elderly

and young alike. During the covid-19 pandemic half of Britons reported often feeling lonely; those aged between 16 and 24 struggled the most. Smartphones, social media, online dating and working from home are all blamed for feelings of alienation. Three recent books have taken on the subject.

In “All the Lonely People” Sam Carr, a psychologist, collects stories of individuals who feel cut off or forsaken. A teenage Afghan refugee struggles to blend in at school in Somerset. An octogenarian languishes in a retirement home. Mr Carr handles this material sensitively, weaving their experiences together with his own, in particular his role as a single parent.

The author evokes the pain of bereavement, heartbreak and childhood trauma and underlines the stigma attached to being withdrawn and friendless. Somehow, the book is not an unbearably bleak read, but you do wonder whether loneliness is just an unavoidable part of the human condition. The range of testimonies also suggests that loneliness is not a single feeling so much as a name for a medley of emotions and unsatisfied appetites.

For “Solitude” Netta Weinstein and Thuy-vy Nguyen, two psychology professors, have teamed up with Heather Hansen, a science journalist, to ponder the rewards of time spent alone. They begin with an account of the mythology of solitude created by figures such as Michel de Montaigne, an essayist, and Edward Hopper, a painter. They then draw on laboratory work, interviews and surveys to illuminate how being alone really affects the human psyche.

It is common to treat loneliness and solitude as synonyms, but they are not. The authors suggest that what is negatively portrayed as one state can be positively reframed as the other. To this end they emphasise the restorative possibilities of being alone and include practical guidance. In a

noisy and crowded world, they argue, people should make time to be by oneself, away from attention-grabbing stimuli.

The book's interviewees mostly regard a lack of company as conducive to autonomy. But this depends on whether solitude is elective or enforced. If it is enforced, as it is for social outcasts and some prisoners, for instance, it is often wretched. Elective solitude, by contrast, above all in natural settings, affords space for reflection. It can open the door to "peak experiences" such as wonder, awe, harmony, even ecstasy. (In a hyper-connected digital age, many readers may not fancy their chances of ever being unplugged long enough to have such experiences.)

Pascal Bruckner sees a world shrinking from sociability in favour of snug seclusion. "The Triumph of the Slippers" is grounded not in research but in the French tradition of witty social criticism. Mr Bruckner, a philosopher and polemicist, dubs the present period "a Great Withdrawal". As he sees it, the openness of the late 20th century is over, and "the closing of minds and spaces is well under way." People may like living in "authorised sloppiness", yet it is hard to imagine heroes and trailblazers wearing dressing gowns: "While it's nice to be comfortable, you can't build a civilisation on softness."

With a mixture of playfulness and grandiosity, he describes a society where most tasks can be completed without leaving the house. His book overflows with soundbites—the public sphere, for example, is dominated by "orgies of pettiness"—and he peppers his short chapters with references to Nietzsche, Plato and Rousseau.

Though his argument spins off in many directions, Mr Bruckner's main theme is the need for people to relearn the art of intimacy. "Interactive solitude", enabled by technology, is no substitute for "the great theatre of the world". Mr Bruckner supplies a checklist for warding off enduring,

corrosive feelings of loneliness: “Have we loved enough, given enough, lavished enough, embraced enough?” If these books are right, many readers will find themselves answering “No.” ■



孤身一人

独处的弊端和益处

三本书探讨了独处的危险和乐趣【《孤独一众》、《独处》、《拖鞋的胜利》书评】

《孤独一众》，山姆·卡尔著。Picador出版社，256页；11.99英镑。将于12月在美国出版；28.99美元。

《独处》，奈特·温斯坦、希瑟·汉森、阮翠薇著。剑桥大学出版社，300页；25.95美元/20英镑。

《拖鞋的胜利》，帕斯卡尔·布鲁克纳著。科里·斯托克韦尔译。约翰威立国际出版社，118页；19.95美元。Polity出版社，16.99英镑。

“孤独”这个词被认为是造词大师莎士比亚所创。他的戏剧中的一位主人公科里奥兰纳斯（Coriolanus）将流亡比作“孤独的龙”退回巢穴。这位罗马将军所说的是一种物理状态，孤独其实就是独自一人而已。

之后，这个词在浪漫主义诗人的笔下开始带有情感色彩。孤独变成了一种精神状态。威廉·华兹华斯（William Wordsworth）的著名诗句感叹“我好似一朵孤独的流云”。是大自然让他暂时摆脱了与世隔绝的消极感，一簇簇水仙可以成为“快乐的旅伴”。

到了20世纪初，孤独被认为是一种典型的都市病。政治理论家汉娜·阿伦特（Hannah Arendt）悲叹，孤独感“曾是通常出现在老年等边缘社会状态中的一种边缘感受，如今却已成为日益扩大的大众群体的日常体验”。

她的担忧在今天引起了深深的共鸣，如今孤独常被划定为一个严肃的公共卫生问题，甚至是一种流行病，老年人和年轻人一样都备受其扰。在新冠疫情期间，有一半的英国人表示经常感到孤独，16至24岁之间的人群感受最深。智能手机、社交媒体、在线约会和居家办公都被认为是导致疏离感的原因。最近有三本书探讨了这一主题。

在《孤独一众》（All the Lonely People）中，心理学家山姆·卡尔（Sam Carr）收集了一些感到自己被孤立或厌弃的个体的故事。在英国萨默塞特郡（Somerset）的学校里，一名阿富汗少年难民难以融入集体。在一家养老院里，一位八旬老人颓丧凋零。卡尔以细腻敏感的笔调讲述了这些故事，将主人公和自己的经历交织在一起，尤其是他身为单身父亲的感受。

作者唤起了对丧亲、心碎和童年创伤之痛的共鸣，并着重探讨了离群索居和没有朋友带有的耻辱标签。他没有把这本书写得满目凄凉而令人难以忍受，不过你确实也会开始怀疑，孤独是否只是人类处境中不可避免的一部分。从书中各式人物的讲述来看，孤独并不只是一种单一的感觉，而更多是各种情绪和未满足的欲望混杂而成的感受。

在《独处》（Solitude）一书中，两位心理学教授奈特·温斯坦（Netta Weinstein）和阮翠薇（Thuy-vy Nguyen）与科学记者希瑟·汉森（Heather Hansen）一道思考了独处的好处。他们从散文家蒙田（Michel de Montaigne）和画家爱德华·霍普（Edward Hopper）等人创造的孤独神话谈起，再借助实验研究、访谈和调查，阐明了独处究竟会对人类心理带来怎样的影响。

人们通常把孤独和独处看做一回事，但事实并非如此。作者们认为，一种被消极描述的状态可以被积极地重构为另一种状态。为此，他们强调独处可能有助于恢复心理健康状态，并提供了实用的指导。他们指出，在一个嘈杂拥挤的世界里，人们应该抽出时间独处，远离让人分神的刺激。

书中的受访者大多认为少人陪伴有助于培养自主性。但这取决于独处是有选择的还是被强加的。例如，如果是像被社会遗弃者和一些囚犯那样被迫独处，通常就痛苦难耐。相比之下，自愿的独处，尤其是在自然环境中的独处，为沉思提供了空间。它可以打开通往惊奇、敬畏、和谐，甚至是狂喜等“巅峰体验”的大门。（在超级互联的数字时代，许多读者可能不太愿意长时间切断网络，放手让自己获得这样的体验。）

在帕斯卡尔·布鲁克纳（Pascal Bruckner）看来，越来越多的人正在远离

社交，追求舒服自在的独处。《拖鞋的胜利》（The Triumph of the Slippers）一书并非基于调查研究，而是基于法国人幽默批判社会的传统。哲学家、辩论家布鲁克纳称如今是一个“大隐退”的时代。在他看来，20世纪晚期的开放时代已经结束，“思想和空间的封闭早已开始。”人们可能喜欢“得到许可的懒散”生活，但很难想象哪个英雄和先驱会是穿着睡袍的。“舒服固然很好，但人类不能靠着柔软来建立文明。”

以一种混合了俏皮和宏大的风格，他描述了一个人类可以在家中完成大多数任务的社会。书中妙语金句不断，例如说公共领域被“琐碎偏狭的狂欢”占领；在简短的章节里，他穿插引用了大量尼采、柏拉图和卢梭的名言。

尽管布鲁克纳的论述向着四面八方发散，但他的核心主题是人们需要重新学习亲密关系的艺术。通过技术实现的“互动式独处”不能替代“世界大舞台”。为了防止持久的孤独感侵蚀心灵，布鲁克纳提供了一些问题让大家自我检验：“我们是否足够地爱过、给予过、慷慨过、拥抱过？”如果这几本书说的都没错，那么许多读者的答案应该都是“没有”。■



Free exchange

What will humans do if technology solves everything?

Welcome to a high-tech utopia

IN “PERMUTATION CITY”, a novel by Greg Egan, the character Peer, having achieved immortality within a virtual reality over which he has total control, finds himself terribly bored. So he engineers himself to have new passions. One moment he is pushing the boundaries of higher mathematics; the next he is writing operas. “He’d even been interested in the Elysians [the afterlife], once. No longer. He preferred to think about table legs.” Peer’s fickleness relates to a deeper point. When technology has solved humanity’s deepest problems, what is left to do?

That is one question considered in a new publication by Nick Bostrom, a philosopher at the University of Oxford, whose last book argued that humanity faced a one-in-six chance of being wiped out in the next 100 years, perhaps owing to the development of dangerous forms of artificial intelligence (AI). In Mr Bostrom’s latest book, “Deep Utopia”, he considers a rather different outcome. What happens if AI goes extraordinarily well? Under one scenario Mr Bostrom contemplates, the technology progresses to the point at which it can do all economically valuable work at near-zero cost. Under a yet more radical scenario, even tasks that you might think would be reserved for humans, such as parenting, can be done better by AI. This may sound more dystopian than utopian, but Mr Bostrom argues otherwise.

Start with the first scenario, which Mr Bostrom labels a “post-scarcity” utopia. In such a world, the need for work would be reduced. Almost a century ago John Maynard Keynes wrote an essay entitled “Economic Possibilities for Our Grandchildren”, which predicted that 100 years into the future his wealthy descendants would need to work for only 15 hours a

week. This has not quite come to pass, but working time has fallen greatly. In the rich world average weekly working hours have dropped from more than 60 in the late 19th century to fewer than 40 today. The typical American spends a third of their waking hours on leisure activities and sports. In the future, they may wish to spend their time on things beyond humanity's current conception. As Mr Bostrom writes, when aided by powerful tech, "the space of possible-for-us experiences extends far beyond those that are accessible to us with our present unoptimised brains."

Yet Mr Bostrom's label of a "post-scarcity" utopia might be slightly misleading: the economic explosion caused by superintelligence would still be limited by physical resources, most notably land. Although space exploration may hugely increase the building space available, it will not make it infinite. There are also intermediate worlds where humans develop powerful new forms of intelligence, but do not become space-faring. In such worlds, wealth may be fantastic, but lots of it could be absorbed by housing—much as is the case in rich countries today.

"Positional goods", which boost the status of their owners, are also still likely to exist and are, by their nature, scarce. Even if AIs surpass humans in art, intellect, music and sport, humans will probably continue to derive value from surpassing their fellow humans, for example by having tickets to the hottest events. In 1977 Fred Hirsch, an economist, argued in "The Social Limits to Growth" that, as wealth increases, a greater fraction of human desire consists of positional goods. Time spent competing goes up, the price of such goods increases and so their share of GDP rises. This pattern may continue in an AI utopia.

Mr Bostrom notes some types of competition are a failure of co-ordination: if everyone agrees to stop competing, they would have time for other, better things, which could further boost growth. Yet some types of competition, such as sport, have intrinsic value, and are worth preserving. (Humans may

also have nothing better to do.) Interest in chess has grown since IBM's Deep Blue first defeated Garry Kasparov, then world champion, in 1997. An entire industry has emerged around e-sports, where computers can comfortably defeat humans. Their revenues are expected to grow at a 20% annual rate over the next decade, reaching nearly \$11bn by 2032. Several groups in society today give us a sense of how future humans might spend their time. Aristocrats and bohemians enjoy the arts. Monastics live within themselves. Athletes spend their lives on sport. The retired dabble in all these pursuits.

| *Everyone's early retirement*

Won't tasks such as parenting remain the refuge of humans? Mr Bostrom is not so sure. He argues that beyond the post-scarcity world lies a "post-instrumental" one, in which AIs would become superhuman at child care, too. Keynes himself wrote that "there is no country and no people, I think, who can look forward to the age of leisure and of abundance without a dread. For we have been trained too long to strive and not to enjoy...To judge from the behaviour and the achievements of the wealthy classes today in any quarter of the world, the outlook is very depressing!" The Bible puts it more succinctly: "idle hands are the devil's workshop."

These dynamics suggest a "paradox of progress". Although most humans want a better world, if tech becomes too advanced, they may lose purpose. Mr Bostrom argues that most people would still enjoy activities that have intrinsic value, such as eating tasty food. Utopians, believing life had become too easy, might decide to challenge themselves, perhaps by colonising a new planet to try to re-engineer civilisation from scratch. At some point, however, even such adventures might cease to feel worthwhile. It is an open question how long humans would be happy hopping between passions, as Peer does in "Permutation City". Economists have long believed that humans have "unlimited wants and desires", suggesting there

are endless variations on things people would like to consume. With the arrival of an AI utopia, this would be put to the test. Quite a lot would ride on the result. ■



自由交流

如果科技解决了所有问题，人类还有什么事可做？

欢迎来到高科技乌托邦

在格雷格·伊根（Greg Egan）的小说《数字永生计划》（Permutation City）中，主人公皮尔（Peer）在一个完全由他掌控的虚拟现实获得了永生，却发现自己的日子无聊透顶。于是，他给自己精心安排了新的爱好。这一刻他在拓展高等数学的极限，下一刻他又在创作歌剧。“他甚至一度对人死后的极乐世界产生了兴趣。这兴趣并没持续多久。他还是更喜欢思考和桌腿有关的问题。”皮尔的变化无常触及了一个更深层次的问题：当科技解决了人类最深重的那些问题后，我们还有什么事情可做？

这是牛津大学哲学家尼克·博斯特罗姆（Nick Bostrom）在一本新书中探讨的一个问题。他在自己的上一本书中曾提出，人类有六分之一的几率会在未来100年内灭绝，原因可能是人工智能（AI）朝着危险的形态发展。在他的新书《深度乌托邦》（Deep Utopia）中，他却思索一个截然不同的结果——如果AI的发展格外理想呢？作者设想的一种情形是，AI发展到几乎能以零成本完成任何有经济价值的工作。而在一种更极端的情形下，就连那些你可能认为只有人类才能做的事情，比如养育子女，AI也能做得更好。这听起来或许更像是敌托邦而不是乌托邦，不过博斯特罗姆不这么认为。

先说第一种情形，博斯特罗姆将其称为“后稀缺”乌托邦。在这样的社会里，人们不再需要做那么多的工作。近一个世纪以前，凯恩斯就在《我们后代的经济前景》（Economic Possibilities for Our Grandchildren）一文中预言，未来100年后，后世的人们会足够富裕，每周只需工作15个小时。虽然今天这一预言并未完全实现，但人们的工作时间已经大大缩短。在富裕国家，人们每周的平均工作时间已从19世纪末的60多个小时下降到今天的不到40个小时。一个普通美国人每天除睡眠之外的时间有三分之一都花在各种休闲活动和体育运动上。将来，他们可能希望把时间花在人类目前还无法想象的一些事情上。正如博斯特罗姆所写，在强大科技的帮助

下，“我们可能获得的体验将远远超出我们目前尚未充分开发的大脑的想象”。

不过，博斯特罗姆的“后稀缺”乌托邦的说法可能稍微有点误导性：尽管超级智能会带来爆发性的经济增长，但它仍将受限于物质资源，尤其是土地。虽然太空探索可能会极大地增加可利用的建筑空间，但不会无限增加。另外还会有一些中间阶段——人类开发出了强大的新型智能，不过并未发展到太空时代。这样的社会或许积累了巨大的财富，但其中大部分财富可能都耗费在住房上——正如今天富裕国家的情况。

能够提升拥有者社会地位的“地位商品”同样很可能继续存在，而且这些商品从其本质属性上说就是稀缺的。即使AI在艺术、智力、音乐和体育方面超越了人类，人类也可能继续从“高他人一等”中获得价值，比如获得最热门赛事的门票等。1977年，经济学家弗雷德·赫希（Fred Hirsch）在《增长的社会限制》（*The Social Limits to Growth*）一书中指出，随着财富的增加，人们会更多地想要获取地位商品。人们花在竞争它们上的时间会增加，这类商品的价格会上涨，在GDP中所占份额也随之水涨船高。这种模式在AI乌托邦中可能会持续下去。

博斯特罗姆指出，一些类型的竞争是协调失败的产物：如果所有人达成一致，停止竞争，他们就有时间去做其他更美好的事情，这可能会进一步促进经济的增长。不过，某些本身就有价值的竞争，比如体育运动，却值得保留。（人类可能也找不出更好的事做了。）自1997年IBM开发的“深蓝”（Deep Blue）计算机首次击败当时的世界冠军加里·卡斯帕罗夫（Garry Kasparov）以来，人们对国际象棋的兴趣与日俱增。围绕电子竞技已经形成了一个完整的产业，而在电子竞技中，计算机可以轻松击败人类。预计未来十年，该行业的收入将以每年20%的速度增长，到2032年将达到近110亿美元。从当今社会的若干群体中，我们可以感知未来的人类可能会如何打发自己的时间。贵族和波西米亚派享受艺术。僧侣活在自己的精神世界。运动员终身从事体育运动。退休人士则和所有这些兴趣爱好都沾点边。

像养育子女这样的事情难道不会一直是人类的自留地吗？博斯特罗姆对此表示怀疑。他认为，在“后稀缺”社会之后将是“后工具”社会，其中AI在照顾孩子方面也会超越人类。凯恩斯本人曾写道：“我认为，没有哪一个国家或民族，能够毫无畏惧地期待悠闲富足时代的到来。因为太久以来，我们一直被教导要奋斗，不要享受……在世界的任何一个区域，从当今富裕阶层的行为和成就来看，前景都非常令人沮丧！”《圣经》说得更简洁：“游手好闲诸恶生。”

这些变化显现出“进步的悖论”。尽管大多数人都希望世界变得更美好，但如果科技变得过于先进，他们可能会失去目标。博斯特罗姆认为大多数人仍会享受本身就有价值的活动，比如享用美食。认为生活变得太过舒适的乌托邦主义者或许会去挑战自我，比如去殖民另一个星球，尝试从头开始再造文明。然而，到了某个时候，即使这样的冒险活动可能也让他们感觉不到意义。人类能像《数字永生计划》中的皮尔那样在不同的兴趣爱好间乐此不疲地反复横跳多久呢？这尚未可知。长期以来，经济学家一直认为人类有“无限的需求和欲望”，这暗示存在着无穷无尽的让人们乐于去消费的事物形态。随着AI乌托邦的到来，这种观点将经受检验。检验结果如何，事关重大。■



Put the axe away

When will Americans see those interest-rate cuts?

Following a nasty surprise, some now think they may come only after the presidential election

PERHAPS IT WAS always too good to be true. The big economic story of 2023 was the seemingly painless disinflation in America, with consumer-price pressures receding even as growth remained resilient, which underpinned surging stock prices. Alas, the story thus far in 2024 is not quite so cheerful. Growth has remained robust but, partly as a result, inflation is looking stickier. The Federal Reserve faces a dilemma about whether to start cutting interest rates; investors must grapple with the reality that monetary policy will almost certainly remain tighter for longer than they had anticipated a few months ago.

The latest troublesome data came from higher-than-expected inflation for March, which was released on April 10th. Analysts had thought that the core consumer-price index (CPI), which strips out food and energy costs, would rise by 0.3% month on month. Instead, it rose by 0.4%. Although that may not sound like much of an overshoot, it was the third straight month of CPI readings exceeding forecasts. If continued, the current pace would entrench inflation at over 4% year on year, double the Fed's target—based on a slightly different inflation gauge—of 2% (see chart 1).

Back in December, at the peak of optimism, most investors had priced in six or seven rate cuts this year. They have since dialled back those expectations. Within minutes of the latest inflation figures, market pricing shifted to implying just one or two cuts this year—a dramatic change (see chart 2). It is now possible that the Fed will not cut rates before the presidential election in November, which would be a blow to the incumbent, Joe Biden.

Jerome Powell, the Fed's chairman, has remained consistent. He has always insisted that the central bank will take a data-dependent approach to setting monetary policy. But rather than bouncing up and down in reaction to fresh figures, he has also counselled patience. At the start of this year, even after six straight months of largely benign price movements, he said the Fed wanted more confidence that inflation was going lower before starting to cut rates. Such caution risked seeming excessive. Today it looks appropriate.

The volatility of market pricing has also changed the Fed's positioning relative to the market. At the end of last year, when investors foresaw as many as seven rate cuts this year, officials had pencilled in just three, appearing hawkish. In their more recent projections, published less than a month ago, officials still pencilled in three cuts, which now appears doveish. The Fed will next update its projections in June.

In the meantime the Fed will be watching more than the CPI. Its preferred measure for inflation, the core personal consumption expenditures price index (PCE), will be released in a few weeks, and is expected to come closer to 0.3% month on month in March. Several of the items that drove up CPI, particularly motor-vehicle insurance and medical services, are defined differently in PCE calculations. The Fed may also be comforted by data showing wage growth has continued to moderate.

Nevertheless, trying to explain away uncomfortable numbers by pointing to this or that data quirk is redolent of 2021, when inflation denialists thought that fast-rising prices were merely a transitory phenomenon. The general conclusion today is that although growth has remained impressively strong, it now appears to be bumping up against the economy's supply limits, and is therefore translating into persistent inflationary pressure. That calls for tight, not loose, monetary policy. The Fed, already cautious about cutting rates when inflation figures were more

co-operative, is likely to be even more wary now. ■



把斧头收好

说好的降息呢？美国人何时才能看到？

在经历了恼人的意外之后，一些人现在认为降息可能要等到大选之后了

或许，当时一切就好得不真实。2023年的经济大事件是美国看似并无痛楚的通胀减缓，在消费价格压力减弱的同时经济增长却保持韧性，为股价飙升提供了支撑。可惜，2024年迄今为止的情况就没那么令人欢欣鼓舞了。经济增长依然稳健，但这也在一定程度上导致通胀看起来越发具黏性。美联储进退两难，不知是否要着手降息；而投资者必须面对一个现实，即相比他们在几个月前的预期，货币政策几乎肯定会保持更紧缩的状态更久。

令人烦恼的最新数据是4月10日公布的高于预期的3月份通货膨胀率。分析师们原本认为，剔除食品和能源成本的核心消费价格指数（CPI）将月度环比上升0.3%。但结果却是上升了0.4%。虽然听起来并没有超出很多，但这已是CPI连续三个月超出预测。如果当前的增速持续下去，通胀率将保持在同比增长超4%，是美联储2%目标（基于一个略有不同的通胀指标）的两倍（见图表1）。

去年12月乐观情绪达到顶峰时，大多数投资者都认为今年会有六到七次降息。自那以后，他们又降低了这些预期。在最新通胀数据公布后的几分钟内，市场定价转为暗示今年只有一到两次降息——这是一个戏剧性的变化（见图表2）。现在看来美联储有可能不会在11月总统大选前降息，这将对现任总统拜登造成打击。

美联储主席鲍威尔的态度始终如一。他一直坚称该行将采取以数据为准绳的方法来制定货币政策。但他并没有让政策追随新出炉的数据起起落落，而是建议大家保持耐心。今年年初，即使物价连续六个月基本呈良性波动，他仍表示美联储需要对通胀正在走低更有把握才会开始降息。之前这有过于谨小慎微之嫌，如今看来却很恰当。

市场价格的波动也改变了美联储相对于市场的立场。去年年底，当投资者

预计今年将有多达七次降息时，美联储官员却预测只有三次，显得很鹰派。在不到一个月前公布的最新预测中，美联储官员仍预计降息三次，这时倒显得鸽派了。美联储将在6月更新其预测。

与此同时，美联储将不仅仅关注CPI。美联储首选的通胀指标——核心个人消费支出价格指数（PCE）将在几周后公布，预计3月份的月度环比增长率将接近0.3%。推动CPI上涨的那几个项目，特别是机动车保险和医疗服务，在PCE计算中有着不同的定义。工资增长继续放缓的数据可能也会让美联储宽心。

然而，试图用这个或那个偶发数据异常来解释令人不安的数字的做法不禁让人想起2021年。当时，否认通胀的人认为物价快速上涨只是一种暂时现象。如今的普遍结论是，尽管经济增长仍保持不俗的强劲势头，但现在似乎已经达到了经济的供应极限，因此正在转化为持续的通胀压力。这就需要采取紧缩而非宽松的货币政策。美联储在通胀数据较为乐观时就已对降息态度谨慎，现在很可能还会更加小心。■



Buttonwood

What China's central bank and Costco shoppers have in common

Hint: it is not a fondness for cryptocurrencies

GOLD HAS always held an allure. The earliest civilisations used it for jewellery; the first forms of money were forged from it. For centuries kings clamoured to get their hands on the stuff. Charlemagne conquered much of Europe after plundering vast amounts of gold from the Avars. When King Ferdinand of Spain sent explorers to the new world in 1511, he told them to “get gold, humanely if you can, but all hazards, get gold.” Ordinary men also clamoured for it after James Marshall, a labourer, found a flake of gold while constructing a saw mill in Sacramento, California, in 1848.

People are once again spending big on the precious metal. On April 9th its spot price hit a record of \$2,364 an ounce, having risen by 15% since the start of March. That gold is surging makes a certain degree of sense: the metal is seen to be a hedge against calamity and economic hardship. It tends to rally when countries are at war, economies are uncertain and inflation is rampant.

But only a certain degree. After all, why is it surging precisely now? Inflation was worse a year ago. The Ukraine war has arrived at something of a stalemate. In the month after Hamas's attack on Israel on October 7th, the price of gold rose by just 7%—half the size of its more recent rally. Moreover, investors had only recently appeared to have gone off the stuff. Those who thought gold would act as a hedge against inflation were proven sorely wrong in 2022 when prices slipped even as inflation spiralled out of control. Cryptocurrencies like bitcoin—often viewed as a substitute for gold—have gained popularity. Longtime gold analysts are puzzled by its ascent.

An investor who cannot understand a rally on the basis of fundamentals must often consider a simpler rationale: there have been more eager buyers than sellers. So, who is buying gold in bulk?

Whoever it is, they are not using exchange-traded funds, or ETFs, the tool most often used by regular folk through their brokerage accounts, as well as by some institutional investors. There have, in fact, been net outflows from gold ETFs for more than a year. After tracking each other closely throughout 2020 and 2021, gold prices and ETF inflows decoupled at the end of 2022. Although prices are up by around 50% since late 2022, gold held by ETFs has dropped by a fifth.

That leaves three buyers. The first, and biggest, are central banks. In general, central bankers have been increasing the share of reserves that are stored in gold—part of an effort to diversify away from dollars, a move that gathered pace after America froze Russia's foreign-exchange reserves in response to its invasion of Ukraine. Nowhere is this shift clearer than in China, which has raised the share of its reserves held in gold from 3.3% at the end of 2021 to 4.3%. Trading has picked up in the so-called over-the-counter market, in which central banks buy much of their gold. China's central bank added 160,000 ounces of gold, worth \$384m, in March.

The second is big institutions, such as pension or mutual funds, which may have been making speculative bets or hedges on gold—in case inflation does come back or as protection against future calamities. Activity in options and futures markets, where they tend to do most of their trading, is elevated.

The third potential buyer is the most intriguing: perhaps private individuals or companies are buying physical gold. In August it became possible to buy hunks of the metal at Costco, an American superstore beloved by the cost-conscious middle classes for selling jumbo-size packs

of toilet paper, fluffy athletic socks and rotisserie chickens, all at super-low prices. The retailer started selling single-ounce bars of gold, mostly online, for around \$2,000—just a hair higher than the spot value of bullion at the time. It sold out almost immediately, and continues to do so whenever it restocks. Analysts at Wells Fargo, a bank, estimate that shoppers are buying \$100m-200m worth of gold each month from the superstore, alongside their sheet cakes and detergent.

That would be 40,000 to 80,000 ounces of gold each month; or, in other words, up to half as much as the Chinese central bank. Such behaviour is perhaps a harbinger of a trend. Inflation in America is creeping up again. It has overshot expectations for three consecutive months, and would reach 4% in 2024 if current trends were to continue. Medium-term expectations, which had dropped, have begun climbing. As shoppers peruse Costco's wares, worrying about the cost of living, it is it any wonder they are tempted by a bit of bullion? ■



梧桐

中国央行和开市客的购物者有什么共同点

小提示：不是都喜爱加密货币

黄金一直都充满诱惑力。最早的文明用它来制作珠宝；最初的货币就由它锻造而成。几个世纪以来，国王们吵吵嚷嚷要得到它。查理曼大帝从阿瓦尔人手中掠夺了大量黄金之后，征服了欧洲大部分地区。1511年，西班牙国王费迪南德派遣探险家前往新大陆，他要他们“拿到黄金，巧取或豪夺，拿到黄金。”1848年，在加利福尼亚的萨克拉门托，工人詹姆斯·马歇尔（James Marshall）在盖一座锯木厂时发现了一片金片，这下贩夫走卒也开始吵嚷着要淘金了。

现在，人们又一次在这种贵金属上一掷千金了。4月9日，黄金现货价格创下每盎司2364美元的纪录，自3月初以来上涨了15%。金价飙升在一定程度上是有道理的：这种金属被视为抵御灾难和经济困难的对冲工具。当国家处于战争状态、经济前景不确定、通货膨胀肆虐时，它往往会大涨。

但只是在一定程度上。为什么它会在现在这个时候飙升呢？明明通胀在一年前还更严重。乌克兰战争已经陷入了某种僵局。在哈马斯去年10月7日袭击以色列后的一个月里，黄金价格仅仅上涨了7%——是最近这次涨幅的一半。而且投资者前不久似乎才刚刚对黄金失去兴趣。2022年，事实证明那些认为黄金可以对冲通胀的人大错特错，当时通胀失控，但金价也跌了。像比特币这样的加密货币——通常被视为黄金的替代品——已经越来越受欢迎。长期跟踪黄金的分析师对金价的上涨感到困惑。

投资者如果无法从基本面上理解金价反弹，通常只能考虑一个更简单的理由了：市场上急着买进的人多于卖出的人。那么，是谁在大量购买黄金呢？

不管是谁，反正都没有用交易所交易基金（ETF），这是普通人在证券账户中最常使用的工具，也是一些机构投资者最常用的。实际上，黄金ETF

的资金净流出已经持续了一年多。在整个2020年和2021年，金价和ETF流入紧密关联，到2022年末二者脱钩了。金价自2022年末以来上涨了约50%，但ETF持有的黄金却减少了五分之一。

这样就剩下三个买家了。第一个是中央银行，也是最大的一个。总的来说，央行行长们一直在增加黄金在储备中的份额——这一定程度上是为了在美元之外实现多元化，在美国冻结俄罗斯外汇储备以回应其入侵乌克兰之后，这一动作加快了。这种转变在中国表现得最为明显，中国已将黄金在外汇储备中的占比从2021年底的3.3%提高至4.3%。在所谓的场外交易市场上，黄金交易已然增加，各国央行的大部分黄金都购自这个市场。今年3月，中国人行增持了16万盎司黄金，价值3.84亿美元。

第二个买家是大型机构，如养老基金或共同基金，它们可能一直用黄金做投机性押注或是对冲，以防通胀卷土重来，或作为对未来灾难的防范。期权和期货市场的活动在增加，它们往往在这些市场进行大部分交易。

第三个可能的买家最耐人寻味：也许个人或公司正在购买实物黄金。去年8月起，人们可以在开市客（Costco）买到金条，这家美国超市深受精打细算的中产阶级的喜爱，因为它以极低的价格出售超大包装的卫生纸、毛绒运动袜和烤鸡。它去年开始销售一盎司的金条，主要是在网上，售价约为2000美元，只比当时的现货金价高略微一点点。结果几乎顷刻间销售一空，而且之后每次补货都是如此。富国银行（Wells Fargo）的分析师估计，除了买大块蛋糕和洗衣液，消费者每月从这家超市购买价值1亿至2亿美元的黄金。

也就是说每月购入4万到8万盎司的黄金：换句话说，最多能到中国人行购买量的一半。这种行为或许是一种趋势的先兆。美国的通胀再次缓慢上扬。通胀已经连续三个月高于预期，如果目前的趋势持续下去，2024年将达到4%。曾经下降的中期预期已经开始攀升。当购物者细细浏览开市客货架上的价签而担心生活成本时，为一块金条心动，有什么奇怪的吗？■



An ageing autocracy

China's high-stakes struggle to defy demographic disaster

The Communist Party puts its faith in robots, gene-therapy and bathing services

IF CHINA'S OLD people formed their own country, it would be the fourth most populous in the world, right behind America. This silver-haired state would be growing fast, too. China's over-60 population sits at 297m, or 21% of the total. By 2050 those figures are expected to reach 520m and 38%. Yet demographers describe China's future as greyer—and smaller. While its older cohorts are growing, younger ones are not (see chart). China's total population declined for the second year in a row in 2023. Its labour force has been shrinking for most of the past decade.

China's economy risks shrinking, too, as a result. With an enormous burden of care on the horizon, the government senses an impending disaster. To date its efforts have focused on boosting the fertility rate (average births per woman), which stands at 1.2—far below the 2.1 required to keep the population stable. Now, though, it is talking about adaptation. During his state-of-the-nation speech last month, the 64-year-old prime minister, Li Qiang, sketched out what he called a “vigorous national strategy” on ageing, covering everything from insurance to pensions.

China is far from the only country facing demographic decline. The populations of Japan, Italy, South Korea and many other places are expected to shrink at an even faster rate over the next two decades. The median age in Japan has risen to nearly 50, whereas in China it is around 40. And it could be argued that China has another advantage: its autocratic government need not cater to elderly voters, whose electoral influence often distorts policy in democracies. In fact, the elderly in China expect little from the state.

Chinese officials, then, appear to have no cause for timidity. On top of being vigorous, one might have expected Mr Li's plan to be bold. So far, though, China's leaders have not pursued the types of reforms needed for their country to escape its demographic destiny. And the longer they wait, the harder their task becomes.

Broadly speaking, China faces three big challenges. The first is how to counteract the decline in the labour force in order to avoid a slowdown in GDP growth. Over the past four decades the economy has expanded by 9% a year, on average, aided by a bulge of working-age adults. Now, though, many of these workers are hitting retirement age. Worse, China is getting old before it gets rich. In 2008, when Japan's population started to fall, its GDP per person was already about \$47,500 in today's dollars. China's is just \$21,000. As more of that money is spent on protecting ageing citizens, less of it will be available for the working generation to consume or invest.

Still, there is much that China can do to keep its economy growing. To start, it could raise the retirement ages for drawing a public pension. China's are among the world's lowest. Most men can stop working at 60. Women who work in offices can retire at 55 and those in factories at 50. Since these rules were set last century, the average life expectancy in China has risen from under 60 to close to 80. "If China's elderly were to have the same labour force participation as in Japan—not a very ambitious target—by 2035 some 40m more people would be at work," writes Bert Hofman of the National University of Singapore. Officials have been floating reforms since 2008, but only now do they seem serious. Observers expect a staged increase in retirement ages to begin next year.

Meanwhile, China can make better use of its workers. Half of Chinese aged 20 to 24 have attended university or vocational college, compared with just 7% of those who are about to retire. In rural areas, though, few people make it past high school. Improving the skills of this group would go a long way.

Mr Hofman reckons that, all other things being equal, the shrinking labour force will cost China one percentage point of GDP growth each year over the next decade. But a longer-working, better-educated labour force could cancel out that effect, he says.

Officials might also think harder about what Chinese workers do and where they do it. About a quarter of them work in agriculture, compared with less than 3% in rich countries. China's leader, Xi Jinping (age 70), talks a big game about creating high-productivity jobs. These are likely to be in cities. But China's hukou, or household registration system, impedes the movement of workers. The country's urbanisation rate is about 65%. It should aim for 75-80%, says Alicia Garcia-Herrero of Natixis, a bank.

In the longer term, demographers believe more machinery and labour-saving technology will also help compensate for a smaller labour force. China has about 400 robots per 10,000 workers, which is higher than most countries. But it is still less than half of South Korea's level of automation. China is trying to catch up. In 2022 half of all the industrial robots installed worldwide were fitted in China, according to the International Federation of Robotics. Assuming the country is able to double its number of such machines to 3m by 2050, "the 'robot dividend' could make up for more than half of the future labour shortage", says a study led by researchers at Beijing Foreign Studies University.

| *Save yourselves*

If China's first challenge is how to produce as many goods and services as possible from a diminished workforce, the second is how to make sure the elderly have a sufficient claim on those items. At the moment, many old people rely on their families for support. That might please Mr Xi, who often talks about the Confucian virtue of filial piety. Shortly after he came to power in 2012, China made it illegal for children to neglect their parents. But with fewer young people around, oldsters will increasingly have to

support themselves.

Pensions are the solution in most countries, and China has its own schemes. The largest state-run fund relies on contributions from employees and companies. It pays out about 3,600 yuan (\$500) a month on average, or around 50% of what recipients were earning before they retired. That compares with over 60% in the OECD, a club of rich countries. The Chinese beneficiaries get enough to pay for basic expenses in most urban areas. But the fund only covers salaried workers in cities, less than half of the labour force. A separate scheme covering most other workers relies on state subsidies. It pays out a measly 200 yuan a month on average.

A lack of generosity is not the only problem. As the number of retirees swells, these schemes are running out of money. The larger one is on track to go broke by 2035, according to projections by a government think-tank. To maintain them both at their current benefit levels, spending would have to double by 2050, reaching 10% of GDP. (Pensioner benefit spending in Britain, by contrast, is about 5% of GDP.)

In order to ease its burden, the government is trying to convince more people to sign up for private pensions. Since 2022 workers have been able to set aside savings in tax-deferred accounts accessible upon retirement. Only 50m people have enrolled so far. Most of the accounts are still empty. Officials say workers don't understand the importance of pension planning. But a deeper problem is that many Chinese distrust financial products, preferring to put their money in property.

That leaves the government with some tough choices to make. It could press companies to contribute more to the larger state-run pension scheme (many smaller firms dodge this responsibility). Or the government could raise taxes to fill the gap in its funding. That would involve root-and-branch reform of the fiscal system, says Christine Wong of the National University

of Singapore. China's tax revenue accounts for about 20% of GDP, compared with an average of 34% across the OECD.

Another option is to make the state's pension schemes even less generous. But that risks sending more old people into poverty. In contrast to many Western countries, the elderly in China tend to be poorer than people from younger generations. A study by researchers at Peking University found that 13% of China's over-60s lived on less than 300 yuan a month in 2020. So this option appears to be a non-starter. "We will continue to increase basic pensions," said Mr Li last month.

| *Invest in silver*

It will be tricky enough for the government to ensure that the elderly have a sufficient claim on the country's output. But it must also make certain that the economy is producing the right stuff to satisfy an older population's needs. That is the final challenge—and the government is relying on market forces to play a role in allocating resources. Mr Li has promised to "develop the silver economy", a term used to describe goods and services aimed at old people. The sector will be worth 30trn yuan by 2035, up from 7trn today, according to official estimates.

Companies see opportunities, producing such things as medical equipment that can monitor old people's health at home. The demand for nutritional supplements like calcium has increased. Analysts reckon the market for adult diapers could overtake the one for infant nappies by 2025. Meanwhile, the state is subsidising private investment in everything from companion robots to anti-ageing gene therapy.

But when it comes to more substantial care services, the scale of demand is such that the government will have to step up. As the number of old people suffering from chronic diseases and disabilities increases, the potential cost of long-term care nationwide will triple to \$247bn per year by 2050,

according to estimates from researchers at Chongqing Technology and Business University.

In cities like Shanghai, where local coffers are flush, the government seems to be doing a good job. Officials have launched a programme to renovate the houses of old people, adding lifts and handrails. Cheap community canteens, which deliver food, are being built. Officials also subsidise home bathing services for those with disabilities or dementia, taking notice of a practice that is common in Japan.

Most local governments, though, are short of cash. The care homes they maintain are either shoddy or full. Staffing is a problem, too. China has 500,000 trained carers. It needs 6m, estimates Xinhua, an official news agency. In rural areas, officials are relying on what they call “mutual aid” projects, where younger old people help care for older ones.

For now few are complaining. Many oldsters have lived through far more difficult periods, such as the great Chinese famine and the Cultural Revolution. Lauren Johnston of the University of Sydney calls China’s elderly the “suck-it-up generation”. Indeed, they were forced to do just that when the government lifted most of its covid-19 restrictions in late 2022, without having adequately prepared for a big wave of the disease. More than 1m people are thought to have died in the months that followed. Most were old. The ruling Communist Party, which covered up the death toll, paid no real political price.

But China’s leaders do not completely ignore the feelings of old people. One reason so many died of covid was because the government refused to mandate vaccination, partly out of fear of a backlash from the elderly. The next generation of oldsters may have more considerable demands. They will have grown up during boom times and may expect to while away their sunset years in comfort. If the party gets things wrong, its biggest critics

may not be young liberals, but old curmudgeons. ■



威权下的老龄化

中国抵御人口灾难的抗争事关重大

党将信心寄托在机器人、基因疗法和助浴服务上【深度】

如果让中国的老人组成一个国家，它将是世界上人口第四多的国家，刚好排在美国之后。这个银发国家的人口还会快速增长。中国60岁以上人口有2.97亿，占总人口的21%；到2050年预计将达到5.2亿和38%。然而，人口学家认为中国在未来不仅会更加老龄化，人口也更少。在老年人口增长的同时，年轻人口却没有增长（见图表）。2023年，中国总人口连续第二年下降。在过去十年的大部分时间里，其劳动人口一直在萎缩。

因此，中国经济也有萎缩的风险。随着巨大的护理负担即将到来，中国政府感觉到一场灾难正在迫近。迄今为止，其工作重点是提高生育率（每名妇女的平均生育数量），中国的生育率为1.2，远低于保持人口稳定所需的2.1。不过，现在它谈论的是适应。在上个月的政府工作报告中，64岁的总理李强勾勒出“积极应对人口老龄化国家战略”，涵盖从保险到养老金的方方面面。

中国绝非唯一面对人口衰减问题的国家。未来二十年，日本、意大利、韩国和许多其他地方的人口预计萎缩得还要快。日本的中位年龄已上升到近50岁，而中国在40岁左右。一些人会说，中国还有另一个优势：其威权政府不需要迎合老年选民；换在民主国家，这个群体的选举影响力往往会扭曲政策。事实上，中国老年人对政府不抱什么期待。

这样看来，中国官员似乎没有理由胆怯。人们原本可能预期李强的计划不只是积极，还会是大胆的。然而到目前为止，中国领导人还没有寻求让中国摆脱人口困境所需要的那些改革。他们拖的时间越久，这项任务就越难。

总的来说，中国面临三大挑战。首先是如何抵消劳动力下降的影响以避免GDP增长放缓。在过去四十年里，经济平均每年增长9%，这部分得益于劳

动年龄人口的激增。然而眼下，这些工人有许多已经到了退休年龄。更糟糕的是，中国是未富先老。2008年，当日本人口开始下降时，按今天的美元计算，其人均GDP已达到47,500美元左右。中国还只有21,000美元。随着其中更多的钱用于保障老年公民，可供劳动人口消费或投资的资金就会减少。

尽管如此，中国仍可以做很多事情来保持经济增长。首先，它可以提高退休年龄，延后领取公共养老金。中国是世界上退休年龄最低的国家之一。大多数男性可以在60岁时停止工作。在办公室工作的女性可以在55岁退休，在工厂工作的女性可以在50岁退休。自上个世纪制定这些规则以来，中国的平均预期寿命已从不到60岁上升到接近80岁。“如果中国老年人的劳动参与率达到日本的水平——这不是一个非常艰巨的目标——到2035年在职人口将增加大约4000万人。”新加坡国立大学的伯特·霍夫曼（Bert Hofman）写道。自2008年以来，官员们一直在提议改革，但直到现在，他们似乎才认真起来。观察人士预计，退休年龄将从明年开始分阶段提高。

与此同时，中国可以更好地利用它的劳动人口。在20至24岁的中国人中，有一半人上过大学或职业学院，而在即将退休的人中只有7%。但是，在农村地区，很少有人能读完高中。要提高这个群体的技能有很长的路要走。霍夫曼认为，在其他条件相同的情况下，劳动人口萎缩将使中国在未来十年内每年损失一个百分点的GDP增长。但他说，一支工作年限更长、受教育程度更高的劳动力队伍可以抵消这种影响。

官员们也许还要更仔细地思考中国工人从事什么工作、在哪里工作的问题。其中约四分之一从事农业，而在富裕国家这一比例不到3%。中国领导人习近平（70岁）谈到了创造高生产率就业机会的大棋。这些就业很可能在城市里。但中国的户籍制度阻碍了工人的流动。它的城市化率约为65%。它的目标应该是75%至80%，法国外贸银行（Natixis）的艾丽西亚·加西亚-埃雷罗（Alicia Garcia-Herrero）说。

更长远看，人口学家认为，更多的机械和节省劳动力的技术也将有助于弥

补劳动力的减少。中国每10,000名工人拥有约400台机器人，高于大多数国家。但它的自动化水平仍然不到韩国的一半。中国正在努力迎头赶上。根据国际机器人联合会（International Federation of Robotics）的数据，2022年全球安装的所有工业机器人中有一半是在中国。假设到2050年，中国能够将此类机器的数量翻一番，达到300万台，“‘机器人红利’可以弥补未来劳动力短缺的一半以上”，北京外国语大学的研究人员主持的一项研究认为。

| 自己储蓄过活

如果说中国面临的第一个挑战是如何用一支缩减的劳动力队伍生产尽可能多的商品和服务，那么第二个挑战就是如何确保老年人获得足够多的商品和服务。目前，许多老人依靠家人赡养。这可能会让经常谈论儒家孝道的习高兴。2012年他上台后不久，中国就规定子女疏于照料父母是非法的。但随着年轻人的减少，老年人将越来越需要自己养活自己。

大多数国家的解决方案是养老金，中国也有自己的养老金计划。这一最大的国有基金依赖员工和公司的缴款。它支付人均每月约3600元，约为领取者退休前收入的50%。相比之下，在富裕国家俱乐部经合组织（OECD）中，这一比例超过60%。在大多数城市地区，中国养老金领取者领到的钱足够支付基本开支。但该基金只涵盖城市中的领薪工人，不到劳动人口的一半。另外一个覆盖大多数其他工人的计划依赖政府补贴，平均每月仅支付区区200元。

不够慷慨并不是唯一的问题。随着退休人员激增，这些计划的资金正在耗尽。据一家政府智库预测，到2035年，规模较大的公共养老基金可能要破产。要保持两者目前的福利水平，到2050年，总支出必须翻一番，达到GDP的10%。（相比之下，英国的养老福利支出约为GDP的5%。）

为了减轻负担，政府正试图说服更多人参加个人养老基金。从2022年开始，工人可以把储蓄放在递延纳税账户中，到退休时可以使用。到目前为止，只有5000万人参加。大多数帐户仍然空空如也。官员们说，工人们不了解养老金规划的重要性。但一个更深层次的问题是，许多中国人不信任

金融产品，宁愿把钱投入房地产。

这让政府不得不做出一些艰难的选择。它可能会迫使公司为更大的国有养老金计划缴更多的钱（许多小公司逃避了这一责任）。或者它可能提高税收来填补资金缺口。这将涉及财政体系的根本改革，新加坡国立大学的黄佩华说。中国的税收收入约占GDP的20%，而经合组织国家的平均比例为34%。

另一种选择是让中国的养老金计划更加地不慷慨。但这有可能使更多的老年人陷入贫困。与许多西方国家相反的是，中国的老年人往往比年轻世代更穷。北京大学的研究人员所做的一项研究发现，2020年，中国60岁以上人口中有13%的人每月生活费不足300元。这样看，这个选项根本行不通。我们会“继续提高退休人员基本养老金”，李强上个月表示。

｜ 投资银发

对于政府来说，要确保老年人从国家的产出中获得足够多已经够棘手了。但它还必须确保经济的产出恰好能满足日渐老化的人口需求。这是最后一个挑战，而政府正依靠市场力量在资源配置中发挥作用。李强承诺要“发展银发经济”，即针对老年人的商品和服务。根据官方估计，到2035年，银发经济的规模将从今天的7万亿元增加到30万亿元。

企业看到了机会，正在生产可以在家中监测老年人健康状况的医疗设备等产品。对钙等营养补充剂的需求增加了。分析师估计，到2025年，成人纸尿裤市场可能会超过婴儿纸尿裤市场。与此同时，政府正在补贴从陪伴机器人到抗衰老基因疗法等各个领域的私人投资。

但在更实质性的护理服务方面，需求是如此之大，政府将不得不自己站出来。根据重庆工商大学的研究人员估计，随着患有慢性病和残疾的老年人的增多，到2050年，全国长期护理的潜在成本将增加两倍，达到每年2470亿美元。

在像上海这样地方财政宽裕的城市，当地政府看上去做得很好。官员们已经启动了一项为老人翻新房屋的计划，加装电梯和扶手。正在建造的社区

食堂提供价格实惠的餐食。官员们还借鉴了日本的普遍做法，为残疾或痴呆老人的家庭沐浴服务提供补贴。

然而，大多数地方政府都缺乏资金。它们维护下的养老院要么条件恶劣，要么人满为患。人员配备也是一个问题。中国有50万名接受过训练的护理人员，而官方通讯社新华社估计的需求是600万。在农村地区，官员们依靠他们所谓的“互助”项目，让不那么老的老人帮助照顾更年长的老人。

目前，很少有人抱怨。许多老人都经历过比眼下困难得多的时期，例如三年困难时期和文化大革命。悉尼大学的劳伦·约翰斯顿（Lauren Johnston）称中国的老年人为“打落牙齿和血吞的一代”。实际上，他们在2022年底就不得不这样忍受下来，当时政府取消了大部分新冠疫情限制，却没有为这种疾病的一轮大爆发做好充分准备。在随后的几个月里，据信有超过100万人死亡。大多数都是老人。党掩盖了死亡人数，并没有在政治上付出真正的代价。

但中国领导人并没有完全忽视老人的感受。这么多人死于新冠病毒的一个原因是政府拒绝强制接种疫苗，部分原因是担心老年人的强烈反对。下一代的老人可能会有更多显见的要求。他们在繁荣时期长大，并可能期望舒适地度过自己的晚年岁月。如果党没把事情做好，它最大的批评者可能不是自由主义的年轻人，而是坏脾气的老人。■



Schumpeter

Generative AI has a clean-energy problem

What happens when the AI revolution meets the energy transition

WHEN A COLLEAGUE from this newspaper visited Sam Altman in 2022 at his home in San Francisco, he noticed two pairs of pink high-tops on a bookshelf. One had the logo of Mr Altman's machine-learning startup, OpenAI. The other bore an emblem for Helion, a nuclear-fusion company that Mr Altman also backs. The entrepreneur is obsessed with both technologies—not just as foot fashion. He believes that the cost of intelligence and the cost of energy will fall in a mutually sustainable way. He calls it a “long and beautiful exponential curve”.

Nasty, brutish and short, more like. Talk to utilities and data-centre operators and, though many share Mr Altman's excitement about artificial intelligence (AI), they are grappling with an energy conundrum on which the future of three big economic shifts partly hinges: the AI revolution; the efforts to electrify swathes of the economy; and the fight against climate change. In a nutshell, “generative” AI, the sort behind OpenAI's ChatGPT, has a ravenous appetite for electricity. It has landed, virtually out of the blue, on a global energy system that is already struggling to cope with alternative sources of power demand. As yet it is not clear whether there will be enough clean energy to meet everyone's needs.

At first glance, the solution looks simple. Data centres, such as those that companies like Alphabet, Amazon and Microsoft use to supply cloud-computing services, have over the past decade or so accounted for only 1-2% of global energy demand. For years the big-tech “hyperscalers” have harvested ever greater energy efficiencies from their server farms, even as the world's computing workloads have soared. Moreover, they have invested heavily in clean energy to offset their carbon footprints. In

America, electricity providers to the hyperscalers are only too keen to help. They have endured two decades of anaemic electricity demand and are desperate for new sources of growth. In recent earnings calls their bosses have promised tens of billions of dollars in investment over the next five years to pump more power to data centres. Last month one such firm, Talen Energy, sold Amazon a nuclear-powered data centre for \$650m. So far, so promising.

Generative AI changes the nature of the game, though. Since the days when they were the workhorses of the cryptocurrency boom, graphics-processing units (GPUs), the chips on which models like ChatGPT are trained and run, have been energy addicts. According to Christopher Wellise of Equinix, which rents out data centres, a pre-AI hyperscale server rack uses 10-15 kilowatts (KW) of power. An AI one uses 40-60KW. It is not just computation that gobbles up electricity. Keeping the racks of GPUs cool requires just as much oomph. Moreover, a lot of AI-related energy demand in the past year or so has come from trainers of “foundation” models like GPT-4, OpenAI’s latest offering. Widespread use of these as tools—for research, to make videos, to dress the Pope in Balenciaga—could put more strain on the grid. A search by ChatGPT may consume ten times the electricity of googling.

It is early days in the generative-AI boom, so it is too soon to make hard and fast predictions. But informed guesses about the related rise in energy demand are striking. At the top of its range, the International Energy Agency, a global forecaster, says that by 2026 data centres could use twice as much energy as two years ago—and as much as Japan consumes today. It expects data centres to account for a third of new electricity demand in America over the next two years. Rene Haas, chief executive of Arm, a chip-design company, told the Wall Street Journal this week that by the end of the decade AI data centres could consume as much as a quarter of all American electricity, up from 4% or less today.

In America, two things further compound the complexities. The first is timing. The rise of generative AI coincides with a booming economy, with power consumption to match. Many power consumers want their energy to be zero-carbon, creating competition for a scarce resource. So do buyers of electric vehicles (EVs), the rise of which may have slowed but has not stopped. The second is the challenge of expanding the grid. Despite support from the White House, it is not easy for utilities to build new renewable capacity quickly. They suffer from supply-chain problems; by some accounts it takes three years to deliver a transformer, up from less than a year previously. Interest rates have pushed up the cost of wind and solar projects, making them harder to fund. Building new transmission lines is fiendishly tough.

No doubt there will be creative thinking. The obvious solution is to make GPUs more energy-efficient. Nvidia, their biggest supplier, says it has already achieved this with its latest generation of AI servers. More efficient chips can, however, simply stimulate more usage. Another option, says Aaron Denman of Bain, a consultancy, is for the hyperscalers to use their deep pockets to help utilities overcome some of the grid constraints. He says that the real crunch may occur during certain parts of the year, such as unusually hot summer days when Americans switch on their air-conditioners. That means having small power plants on standby. The likelihood, though, is that these will be fuelled by natural gas, undermining the cloud providers' climate commitments.

| *The nuclear option*

If shortages of renewable energy occur, it will come at a cost. No one knows yet how generative AI will make money. What people do know is that the cost of acquiring GPUs is rocketing. If the energy costs of running them soar, too, it could put the brakes on expansion. In addition, the electrification of the rest of the economy is highly cost-dependent; an AI v

EV scramble for clean power would push up prices and serve neither industry well. By all means keep your fingers crossed for Mr Altman's rose-tinted fusion dream to become reality. But don't count on it. ■



熊彼特

生成式AI面临清洁能源难题

当AI革命遇上能源转型会发生什么

本刊有一位同事曾在2022年到山姆·阿尔特曼（Sam Altman）在旧金山的住所采访，当时他注意到一个书架上摆着两双粉红色高帮鞋。一双印有阿尔特曼创办的机器学习公司OpenAI的标志，另一双印有他投资的核聚变公司Helion的标志。这位企业家痴迷于这两种技术，可不只是为了用来装饰鞋子。他认为人工智能的成本和能源的成本将以一种互利的方式同步下降。他称之为“绵长优美的指数曲线”。

现实里，这曲线更像是丑陋、粗暴又短促的。与电力公司和数据中心运营商聊聊就会了解到，虽然他们很多都和阿尔特曼一样对人工智能（AI）兴奋不已，但正在疲于解决一个能源难题，这个问题在一定程度上关系到AI革命、经济电气化和应对气候变化这三大经济转型的未来。简单来说就是，OpenAI的ChatGPT背后的那类“生成式”AI对电力的需求庞大。全球能源系统本来就已经难以应付新兴的电力需求源头，生成式AI的从天而降更让它捉襟见肘。目前尚不清楚是否会有足够的清洁能源来满足方方面面的需求。

乍看之下，解决方案似乎很简单。过去十年左右，数据中心（如Alphabet、亚马逊和微软等公司用于提供云计算服务的那些）仅占全球能源需求的1%至2%。多年来，在全球计算量飙升的同时，这些“超云”科技巨头也在不断提升它们服务器农场的能源效率。此外，它们还大力投资清洁能源来抵消自身的碳足迹。在美国，超云巨头的供电商非常乐于配合。经历了二十年的电力需求低迷，它们迫切需要新的增长点。在最近的财报电话会议上，这些供电商的老板们承诺在未来五年投入数百亿美元，加大为数据中心供电。其中一家公司Talen Energy上月以6.5亿美元的价格向亚马逊出售了一个核电驱动的数据中心。目前看来，一切充满希望。

然而，生成式AI改变了这场游戏的性质。先前，图形处理单元（GPU）支

撑起了加密货币热潮，自那时起，这些训练和运行ChatGPT等模型所用的芯片就已成为能源消耗大户。数据中心出租公司Equinix的克里斯托弗·维利斯（Christopher Wellis）指出，之前一个非AI用途的超云服务器机架耗电10至15千瓦，而一个AI服务器机架要耗电40至60千瓦。消耗大量电力的不仅是计算，为一排排机架里的GPU降温同样耗电。此外，在过去一年左右的时间里，大量与AI相关的能源需求来自训练GPT-4（OpenAI的最新产品）等“基础”模型。广泛使用这些工具（做科研、制作视频、给教皇穿上巴黎世家的服饰）可能会给电网带来更大压力。用ChatGPT进行一次搜索可能会消耗十倍于谷歌搜索的电力。

这股生成式AI热潮仍在发展初期，要做出明确的预测为时尚早。但关于相关能源需求上升的合理猜测还是令人震惊。全球预测机构国际能源署（IEA）表示，按预测上限，到2026年，数据中心的耗电可能达到两年前的两倍，相当于目前整个日本的年用电量。该机构预计，未来两年，数据中心将占美国新增电力需求的三分之一。芯片设计公司安谋（Arm）的首席执行官雷内·哈斯（Rene Haas）最近告诉《华尔街日报》，大约五年后，AI数据中心可能消耗掉美国四分之一的电力，而目前该比例大概不到4%。

在美国，有两个因素让问题更为复杂。首先是时机。生成式AI的崛起恰逢经济繁荣，电力消耗在此期间增加。许多电力消费者希望使用零碳能源，导致对这种稀缺资源的竞争。电动汽车买家也有这样的希望，而电动车的崛起虽可能已经放缓，却并未停止。其次是扩建电网的难度。尽管得到白宫的支持，电力公司要快速建成新的可再生能源产能并非易事。它们受到供应链问题的拖累；一些人表示，现在交付一台变压器需要三年时间，而此前只要不到一年。利率推高了风能和太阳能发电项目的成本，让这些项目更难筹措资金。铺设新输电线路也极其麻烦。

无疑，人们会开动脑筋想办法。最显而易见的解决方案是提高GPU的能效。最大的GPU供应商英伟达表示，最新一代AI服务器芯片已经实现了这个目标。然而，更高效的芯片可能更会刺激更多使用。贝恩咨询公司的亚伦·登曼（Aaron Denman）表示，另一个办法是让超云巨头利用其雄厚

资金帮助电力公司克服电网的某些限制。他说，电力真正吃紧可能发生在一年中的某些时段，例如美国人都在使用空调的异常炎热夏日。这意味着要有小型发电厂备用。不过这些电厂很可能会使用天然气发电，这会破坏这些云服务供应商的气候承诺。

| 核选项

如果可再生能源出现短缺，就会产生成本。目前还没有人知道生成式AI将如何盈利。人们只知道GPU的购买成本在不断飙升。假如运行GPU的能源成本也飙升，可能就会阻碍这项技术的推广。此外，经济中其他部门的电气化进程很大程度上取决于成本；AI与电动汽车争夺清洁电力将推高价格，对两个行业都不利。我们尽可祈祷阿尔特曼的核聚变能美梦成真，但别抱太大期望。■



A numbers game

In the Premier League, data help minor clubs take on the mighty

A new book analyses Brentford's statistical shrewdness

Smart Money: The Fall & Rise of Brentford FC. By Alex Duff. Constable; 327 pages; £22

AT MOST FOOTBALL clubs, the equation is simple: you put in (vast amounts of) money, and you get out star players and trophies. Take Manchester City, the Premier League's reigning champions. Before its takeover in 2008 by Sheikh Mansour of Abu Dhabi, the club often languished in the bottom half of the table; it has won English football's top trophy seven times since 2011. Chelsea also won back-to-back titles after the club was bought by Roman Abramovich, a Russian oligarch, in 2003. (Never mind that City is accused of financial mismanagement, which it denies, and Mr Abramovich was forced to sell Chelsea after Russia invaded Ukraine.)

Brentford FC has taken a different approach. The club, based in west London, was promoted into the Premier League in 2021 after toiling for decades in the lower reaches of football. What makes its success surprising is not how much money its owner, Matthew Benham, has put into the team, but how little. In a new book Alex Duff, a lifelong Brentford fan, explains how a frugal plan paid dividends.

Mr Benham first saw "the Bees" play in 1979, when he was 11. As a young adult he studied physics at Oxford University and then went to work in the City, where he became a derivatives trader. In his early 30s, sensing correctly that bookmakers were inaccurate when setting odds for football matches, Mr Benham left banking to become a full-time gambler.

He set up his own company, Smartodds, where he applied his aptitude for

finding the underlying value of stocks to identifying the underlying strength of football teams. He competed with financial institutions to hire the best mathematicians. They were told to focus not on how many goals a team was scoring—which, according to Mr Benham, was subject to too much randomness—but on the “goalscoring opportunities” it was creating. In time, he reckoned, the goals would come.

The approach was so efficient that when Brentford, going through one of its periodic financial crises, put out a general appeal for help in 2005, Mr Benham offered his services. Within a decade he owned the club and was applying his ideas to how the team was constructed. Analytics underpinned a remarkably lucrative buy-low-sell-high transfer strategy.

Like “Moneyball” (2003), Michael Lewis’s hit book about the use of statistics in baseball, “Smart Money” is both informative and entertaining. Mr Duff introduces a disparate cast of characters, including Tony Bloom, another bettor, who took over Brighton in 2009. The book makes diverting excursions into the academic field of sports statistics and the birth of sports betting in America, among other subjects.

Mr Duff is also alert to football’s long-standing, problematic relationship with gambling. Critics have voiced concerns about the effect of near-ubiquitous advertising on supporters. Players have become addicted to betting, too. Brentford’s star striker, Ivan Toney, has just finished an eight-month ban after breaching betting rules more than 200 times.

There were plenty of sceptics within football about Brentford’s philosophy. (As one old-timer carped: “Next week, why don’t we get your computer to play against their computer and see who wins?”) But in time it worked. In their debut season in the Premier League, Mr Benham’s investment of around £100m (\$126m) in the club—a tiny sum compared with competitors—realised its first profit of £25m.

Statistics have helped the team win on the pitch, too. Mr Benham identified set pieces as an important part of creating scoring opportunities, and hired Gianni Vio, an Italian coach with 4,000 such plays in his repertoire. Players are instructed to press the opposition and tackle players within ten seconds of them receiving the ball. Though Brentford is not vying for the title, in recent years the team has beaten several of football's plutocrats—including Manchester City. What were the odds? ■



数字游戏

数据帮助英超弱旅对抗强队

一本新书分析了布伦特福德在统计方面的精明之处【《聪明钱》书评】

《聪明钱：布伦特福德足球俱乐部的衰落与崛起》。亚历克斯·达夫著。Constable出版社；327页；22英镑。

大多数足球俱乐部遵循的公式都很简单：投入（巨额）资金，就能招纳球星、斩获奖杯。以英超卫冕冠军曼城为例，在阿布扎比的谢赫·曼苏尔于2008年接管曼城之前，曼城经常在积分榜的后半段徘徊，而自2011年以来，曼城已七次赢得英格兰足球的最高奖杯。切尔西也在2003年被俄罗斯寡头罗曼·阿布拉莫维奇（Roman Abramovich）收购后连续两次夺冠。

（暂且不提曼城被指控财务管理不善——曼城对此予以否认，也别在意阿布拉莫维奇在俄罗斯入侵乌克兰后被迫出售切尔西）。

布伦特福德俱乐部采取了一种不同的方式。这家俱乐部位于西伦敦，在低级别联赛中苦苦挣扎了几十年后，于2021年升入英超联赛。它的成功之所以让人意外，并不是它的老板马修·贝纳姆（Matthew Benham）为球队投入的资金多么高昂，而是多么少。在一本新书中，布伦特福德的终身球迷亚历克斯·达夫（Alex Duff）解释了一项节俭计划是如何带来回报的。

贝纳姆第一次观看“蜜蜂”队的比赛是在1979年，当时他11岁。青年时期，他在牛津大学学习物理，随后进入伦敦金融城工作，成为一名金融衍生品交易员。30岁出头时，贝纳姆准确地意识到博彩公司在设定足球比赛赔率时存在误差，于是他离开银行业，成为一名全职赌徒。

他成立了自己的公司Smartodds，将自己发掘股票潜在价值的天赋应用于识别足球队的潜在实力。他与金融机构竞争，聘请最优秀的数学家，告诉他们不要关注一支球队进了多少球（他认为进球的随机性太大），而要关注球队创造的“进球机会”。他认为，假以时日，球就会进。

这种方法非常有效，于是，当布伦特福德在经历周期性财务危机后于2005年发出求助呼吁时，贝纳姆出手搭救。不到十年，他就拥有了这家俱乐部，并将自己的理念应用到球队建设。通过分析，他制定了利润丰厚的“低买高卖”转会策略。

与迈克尔·刘易斯（Michael Lewis）在2003年出版的关于在棒球中运用统计的畅销书《点球成金》（Moneyball）一样，《聪明钱》一书信息与趣味并重。达夫在书中介绍了一群性格迥异的人物，其中包括2009年接管布莱顿的另一名赌徒托尼·布鲁姆（Tony Bloom）。书中引人入胜地穿插了体育统计这个学术领域，以及美国体育博彩的诞生等话题。

达夫还对足球与赌博之间长期存在的麻烦关系很警觉。批评人士已经对近乎无孔不入的博彩广告对球迷的影响表示担忧。球员们自身也变得沉迷于赌博。布伦特福德队的明星前锋伊万·托尼（Ivan Toney）先前200多次违反英足总博彩条例，如今刚刚结束了为期八个月的禁赛。

足球圈内对布伦特福德的经营理念持怀疑态度的人大有人在。（一位老资格就曾这样叨叨过：“下周要不就让你的电脑跟他们的电脑对战一下吧，看看谁能赢？”）但随着时间的推移，这套理念奏效了。在英超的首个赛季，贝纳姆向俱乐部投资了约1亿英镑（1.26亿美元）——与竞争对手相比，这个数字微不足道——却首次实现了2500万英镑的盈利。

统计数字还帮助球队在球场上获胜。贝纳姆认为定位球是创造得分机会的重要一环，并聘请了拥有4000套定位球战术的意大利教练詹尼·维奥（Gianni Vio）。球员们被要求在对对手接球后10秒钟内向其施压并抢断。虽然布伦特福德没有剑指冠军，但近年来该队击败了包括曼城在内的数家足球豪门。当初谁又曾料到这个呢？ ■



Robotics

Why robots should take more inspiration from plants

They would be able to grow, grip and move in more useful ways

THE ENORMOUS titular robots of the “Transformers” universe, a popular franchise spanning toys, TV series, video games and films, move along the ground in one of two ways. On wheels, when they are shaped like vehicles; on giant humanoid feet when they are not. For decades, most real-world robots also fell neatly into these two fictional paradigms.

Then, around 15 years ago, came the realisation that other means of locomotion were possible. “Zoomorphic” robots mined the animal kingdom for inspiration, piggybacking on evolution’s millennia of research and development. One mimicked an octopus’s malleable arm, allowing it to easily grasp objects and manoeuvre into tight, tricky spaces. Another replicated the ridge-covered toes of a gecko and, consequently, its ability to scale walls.

Animal-like robots continue to be popular. In recent years both America’s and Britain’s armed forces have experimented with quadrupedal robot “dogs” for patrols and surveillance; in February a snake-shaped robot was sent into one of the damaged nuclear reactors of the Fukushima power plant in Japan to inspect the debris left inside. And yet, says Barbara Mazzolai, an Italian roboticist, the field of robotics has proved far less keen to investigate the other major category of living things—plants. She attributes the reluctance to a misconception about the usefulness of plant behaviour: that they are capable of neither motion nor perception. “It’s not true at all,” she says.

Dr Mazzolai and her team at the Bioinspired Soft Robotics Laboratory at the Italian Institute of Technology (IIT) in Genoa recently unveiled a machine

meant to uproot this dogma. Writing in *Science Robotics*, they described “FiloBot” (pictured above), a robot based on climbing plants. Like the real thing, FiloBot (from the Italian word for “tendrils”) is capable of growing, attaching to and twining around supports, and navigating through an environment in response to external stimuli.

To survive, a climbing plant must have the capacity to switch between several different modes of behaviour. In forest environments, it must first grow out of the soil and travel along the ground in search of a support to latch onto—a nearby tree, say. To do this the plant must have a structure capable of bearing its own weight. Once a support is located, though, the plant must switch strategy—anchoring itself around the object and then growing upwards towards the light. To outcompete other plants, it must move as quickly as possible, prioritising rapid growth over heft.

To help it choose the best angle at which to grow, a climbing plant uses information from light and gravity receptors distributed along each shoot. It can also modify the bulkiness of its tendrils by changing how their constituent cells divide and elongate: more padding in the middle will create a firmer tendril, while extra growth on one side will lead to curvature.

| *Move into the light*

FiloBot mimics these behaviours using sensors in its head—at the tip of the main shoot—which is also equipped with a spool of plastic and a heating element. By melting and extruding the plastic in a circular pattern, it can 3D-print its own body at a rate of between two and seven millimetres per minute (bamboo, the world’s speediest plant, cannot beat 0.1). Depending on brightness and orientation, it alters the heat that the plastic is exposed to—lower temperatures result in a more brittle body that increases in size more quickly, while higher temperatures make denser and stronger clumps that grow more slowly. By varying the amount of plastic deposited around

the circle, it can also grow in a rotating pattern to coil around a support.

The researchers found that these simple functionalities were enough to let FiloBot move through a complex, unseen environment, cross gaps and find things to attach to. The lack of heavy on-board computing hardware, they say, means that it remains nimble and requires minimal oversight, while its slow pace means that it doesn't disturb things around it. They reckon that this makes it suitable for potential applications including environmental monitoring in hard-to-reach or unknown locations (where piloting a robot along an exact course might be impossible), or monitoring disaster sites where existing infrastructure is unstable.

For now, as the researchers tweak and test it further, FiloBot's tendrils have not left the laboratory. Still, it has already been useful in deconstructing plant behaviour. For example, it was long hypothesised that climbing plants find their supports by harnessing an ability to grow towards shade, though the exact mechanism was unclear. FiloBot could replicate this behaviour by seeking out the far-red wavelengths characteristic of shaded areas, providing an insight into how plants might accomplish the same thing.

FiloBot is not the only plantlike robot the team is cultivating. Dr Mazzolai has been developing "plantoids", based on roots, since 2012 (then the first plant-inspired robotics venture in the world). These can burrow through earth, and could be used to analyse chemicals or find water. And in 2021 the group at IIT, along with European partners, started developing "I-Seed", a biodegradable mini-bot that can be moved about by wind and rain and change shape according to humidity. Based on the seeds of the South African geranium, it could be used to carry and distribute real seeds for reforestation, opening up and releasing its cargo once it hits suitable soil.

Dr Mazzolai hopes that such projects will inspire other roboticists to take

their cues from botany. The plant kingdom is another world, she says, with a completely different approach to the animal one. “And so we can develop completely new technologies and artificial solutions, because it is so different.” ■



机器人学

为何机器人应该从植物那里汲取更多灵感

这样它们就能够以更多有用的方式生长、抓握和移动

在玩具、电视剧、电子游戏和电影组成的“变形金刚”宇宙中，巨大的机器人主角们有两种在地面移动的方式：变形为车辆时用轮子，其余时候用巨大的人形双足行走。几十年来，现实世界中大多数机器人也就是按这两种虚构范式移动的。

大约15年前，人们意识到还有其他移动方式可用。“动物仿生”机器人从动物界汲取灵感，借力进化历经亿万年的研发成果。有一款仿生机器人的“手臂”模仿了章鱼可以伸展的触手，能轻松抓取物体，还能灵活探入狭小曲折的空间。还有一款复制了壁虎那满是脊状构造的脚趾，从而复制了它们攀爬墙壁的能力。

此后仿动物机器人持续流行。近年来，美国和英国军方都尝试用四足“机器狗”执行巡逻和监视任务；今年2月，一款蛇形机器人被送入日本福岛核电站中的一个受损的反应堆，检查其中损坏情况。然而，意大利机器人学专家芭芭拉·马佐莱（Barbara Mazzolai）说，机器人学界对另一大类生物就远没有那么大的研究劲头了，那就是植物。她认为这是因为人们对植物行为的用途存在误解，认为它们既不能移动也不会感知。“这完全是错误的。”她说。

在热那亚意大利技术研究所（Italian Institute of Technology，简称IIT），马佐莱和她的仿生软体机器人实验室团队最近推出了一款机器人，意在破除这种错误的信条。他们在《科学机器人学》（Science Robotics）杂志上介绍了以攀缘植物为原型设计的“FiloBot”（见上图）。与真正的攀缘植物一样，FiloBot（源自意大利语“卷须”一词）能生长，能附着并缠绕在支撑物上，并根据外部刺激在环境中探路穿行。

要生存下来，攀缘植物必须能在多种不同行为模式之间切换。在森林环境

中，它先要破土而出，然后沿地面寻找可附着的支撑物，比如附近的树木。为此，它必须拥有能支撑自身重量的结构。但一旦找到支撑物后，它就要改变策略——围绕支撑物固定自己，然后朝着能受到光照的上方生长。为了抢在其他植物前面，它必须尽快移动，尽量快速长高，而不是长粗。

为选择最佳生长角度，攀缘植物会利用分布于每根嫩枝上的光和重力受体所提供的信息。它还可以通过改变自身细胞的分裂和伸长方式来调节卷须的粗细：中部增厚会形成更坚固的卷须，而端部更多地增长会形成卷曲形态。

| 向光而行

FiloBot利用其顶部的传感器（位于“主枝”的尖端）来模仿这些行为，它还配有一卷塑料和一个加热元件。通过融化并沿环形方向挤出塑料进行3D打印，FiloBot能以每分钟两到七毫米的速度“生长”（世界上生长最快的植物竹子每分钟也长不到0.1毫米）。根据光照强度和方向，它会改变塑料所受的热量——温度越低，长出来的“身体”越脆，长得越快；温度越高，长出来的“身体”越致密结实，长得越慢。通过改变沿环形“生长”时沉积的塑料量，它还能转着圈围绕支撑物生长。

研究人员发现，这些简单功能足以使FiloBot在复杂未知的环境中移动，跨越间隙，找到物体附着。他们表示，由于没有搭载笨重的计算硬件，FiloBot能够灵活行动，无需操作人员太多干预，行进缓慢则意味着不会干扰周围事物。基于这些特点，他们认为FiloBot适用于在难以到达或未知的地点（无法控制机器人沿精确路线行进）进行环境监测，或在已有基础设施不稳固的灾难现场执行监测任务。

目前，研究人员还在进一步调整及测试FiloBot，它的“卷须”还只是在实验室里攀行。但在解构植物行为方面，它已经发挥了作用。例如，人们一直推测攀缘植物是通过一种向阴暗处生长的能力来寻找支撑物的，但确切机制尚不清楚。FiloBot可以通过探测阴暗处特有的特定波长的远红外线来复制这种行为，从而让人们了解植物如何实现这种行为。

FiloBot并非该团队正在培育的唯一仿植物机器人。自2012年起，马佐莱一直在开发模拟植物根系的“植物机器人”（当时是全球首个仿植物机器人项目）。这些机器人能钻入土壤，也许可以用来分析化学物质或寻找水源。2021年，IIT的这个团队开始与欧洲合作伙伴一起研发“I-Seed”——一款能随风雨移动并根据湿度改变形状的可生物降解的微型机器人。它以南非天竺葵种子为原型，可以携带真正的种子并播种造林，只要遇到合适的土壤，它就会自动打开并释放种子。

马佐莱希望这些项目能启发其他机器人专家从植物学中汲取灵感。她表示，植物王国是另一个世界，完全不同于动物王国。“而正因为这天壤之别，我们可以开发出全新的技术和人工模拟解决方案。”■



Free exchange

Daniel Kahneman was a master of teasing questions

How a psychologist transformed economics

WINNERS OF THE Nobel prize in economics tend to sprinkle their papers with equations. Daniel Kahneman, who died on March 27th, populated his best-known work with characters and conundrums. Early readers encountered a schoolchild with an IQ of 150 in a city where the average was 100. Later they pondered the unfortunate Mr Tees, who arrived at the airport 30 minutes after his flight's scheduled departure, and must have felt even worse when he discovered the plane had left 25 minutes late. In the 1970s readers had to evaluate ways to fight a disease that threatened to kill 600 people. In 1983 they were asked to guess the job of Linda, an outspoken, single 31-year-old philosophy graduate.

Kahneman used such vignettes to expose the seductive mental shortcuts that can warp people's thoughts and decisions. Many people, for example, think it more likely that Linda is a feminist bank-teller than a bank-teller of any kind. Presented with two responses to the disease, most choose one that saves 200 people for certain, over a chancier alternative that has a one-third chance of saving everyone and a two-thirds chance of saving no one. But if the choice is reframed, the decision is often different. Choose the first option, after all, and 400 people die for sure. Choose the second and nobody dies with a one-third probability.

Teasing questions came easily to Kahneman, even in his sleep, according to "The Undoing Project", a book by Michael Lewis. Some sprang from his teaching, which was not confined to ivory towers. He once explained the idea of "regression to the mean" to flight instructors in Israel's air force. The reason pilots tended to improve after a sloppy manoeuvre was not because the instructor screamed at them, but because the chances of an

improvement are higher if the prior performance was unusually bad.

Kahneman was a harsh grader of his own incorrigible self, attentive to his own lapses. One of his early hit papers exposed the kind of methodological muddles to which he himself was vulnerable, such as the misplaced confidence that an outlier, like a child with an IQ of 150, would not skew even a small sample.

Kahneman also had a lifelong—and life-preserving—interest in gossip. His childhood, as the son of Lithuanian Jews living a comfortable but edgy pre-war existence in Paris, was full of talk about other people, he once wrote. Jews in Europe had to “assess others, all the time,” a friend of his told Mr Lewis. “Who is dangerous? Who is not dangerous?...People were basically dependent on their psychological judgment.”

Gossip was both a source of his work and an intended target. His bestselling book, “Thinking Fast and Slow”, was written not for decision-makers, but for “critics and gossipers”. Decision-makers were often too “cognitively busy” to notice their own biases. Pilots could be corrected by observant co-pilots and overconfident bosses might be chastened by whispers around the water-cooler, especially if the whisperers had read Kahneman’s book.

To spread psychological insight, Kahneman once tried to add a course on judgment to Israel’s school curriculum. He expected the project would take a year or two. It took eight, by which time the ministry of education had lost enthusiasm; a humbling example of what he and Amos Tversky, his frequent co-author, called the “planning fallacy”. He had more success inveigling psychological wisdom into the well-guarded realm of economics, which had clung to a thin but tidy model of human decision-making.

How did he do it? One answer is that he teamed up with Tversky, whose elegant mind was as ruthlessly tidy as his desk. They incorporated the cognitive illusions they had discovered into a model called “prospect theory”. According to this theory, people’s well-being responds to changes in wealth, more than levels. The changes are judged relative to a neutral reference point. That point is not always obvious and can be recast: a bonus can disappoint if it is smaller than expected. In pursuit of gains, people are risk averse. They will take a sure win of \$450 over a 50% chance of winning \$1,000. But people gamble to avoid losses, which loom larger than gains of an equivalent size.

Prospect theory translated this model of decision-making from vignettes into the language of algebra and geometry. That made it palatable to economists. Indeed, the discipline began to claim this sort of thing as its own. Applications of psychology “came to be called behavioural economics”, lamented Kahneman, “and many psychologists discovered that the name of their trade had changed even if its content had not.”

| *The cold-hand fallacy*

Even as economics was rebranding psychology, Kahneman revived an older economic tradition: “hedonimeters”, gauges of pleasure and pain that Francis Edgeworth, a 19th-century economist, had imagined. Kahneman’s hedonimeter simply asked people to rate their feelings moment-to-moment on a scale. He found that people’s ratings were often at odds with what they later recalled. Their “remembering” selves put undue weight on the end of an experience and its best or worst moment, neglecting its duration. People would rather keep their hand in painfully cold water for 90 seconds than for a minute, if the final 30 seconds were a little less cold than the preceding 60. Likewise, people sign up for hectic tourist itineraries because they look forward to looking back on them, not because they much enjoy them at the time.

The implications of this discovery extend into philosophy. Which self counts? Despite its manifest flaws, the curatorial self, artfully arranging unrepresentative memories into a life story, is dear to people. “I am my remembering self,” Kahneman wrote, “and the experiencing self, who does my living, is like a stranger to me.” Now his experiencing self has done its living. And it is up to the many people he touched to do the remembering for him. ■



自由交流

丹尼尔·卡尼曼，脑筋急转弯大师

一位心理学家如何改变了经济学

诺贝尔经济学奖得主通常都会在他们的论文中加入一些公式。3月27日去世的丹尼尔·卡尼曼（Daniel Kahneman）则在他最著名的研究上加上了各种人物和谜题。在早期，读者会读到在一座孩童平均智商为100的城市里有一个智商高达150的学童。后来，他们又会读到倒霉的蒂斯

（Tees），他在飞机原定起飞时间的30分钟之后才赶到机场，然而当他发现飞机是在延误了25分钟后起飞的，肯定就更难过了。在1970年代，读者要面对一种可能夺走600人性命的虚构疾病，在不同的治疗方案之间做选择。到了1983年，他们要猜一猜一位哲学专业毕业、直言不讳的31岁单身女性琳达从事什么职业。

卡尼曼用这样的小故事来揭示那些会扭曲人们思考和决策的诱导性心理捷径。例如，对于琳达是一名女权主义银行出纳还是就是一名银行出纳，许多人认为前者的可能性更大。当那种疾病的两个治疗方案摆在面前时，大多数人选择了肯定能挽救200人的方案A，而不选变数更大的方案B——它有三分之一的可能挽救所有600人，但有三分之二的可能一个人也救不了。但是，如果重新设定问题，人们往往又会做出不同的决定。毕竟，选了方案A，就一定会死400人。而选择方案B，就有三分之一的可能一个也不用死。

迈克尔·刘易斯（Michael Lewis）在《思维的发现》（The Undoing Project）一书中提到，像这样逗弄一下人们大脑神经的谜题在卡尼曼是信手拈来，有些甚至是他在梦中想到的。一些问题源自他的教学，而他的教学对象并不仅限于象牙塔里的学生。他曾向以色列空军的飞行教官解释“均值回归”的概念。飞行员在做出一个很糟糕的飞行动作之后往往会有进步，这并不是因为他们被教官吼了，而只是因为，在一次异常糟糕的表现之后，下一次会有所改善的几率提高了。

卡尼曼严苛评判自己身上顽固的弱点和偏误，非常清楚地观察到自己的疏失。他早年的热门论文中有一篇揭露了一种他自己常犯的方法论错误，比如他会错误地认为一个异常值（比如一个智商为150的孩子）不会扭曲样本数据，即便是一个很小的样本。

卡尼曼一生都对八卦感兴趣，这个兴趣还能保命。他曾经写道，作为立陶宛犹太人的儿子，他的童年在战前的巴黎度过，生活舒适却也难免提心吊胆，周围的人都在谈论他人。一位朋友告诉刘易斯，欧洲的犹太人必须“时刻评估他人”。“谁会带来危险，谁不会……人们基本上依赖自己的心理判断。”

八卦既是他工作的灵感，也是他的影响目标。他的畅销书《思考，快与慢》（Thinking Fast and Slow）并不是写给决策者的，而是写给“批评家和爱八卦的人”。决策者往往因为“认知繁忙”而无法注意到自己认知偏误。飞行员的偏差可以由细心的副驾驶纠正，而自信过头的老板可能会被员工在饮水机边的窃窃私语点醒——如果这些员工读过卡尼曼的书，那效果会更好。

为了传播心理学洞见，卡尼曼曾试图在以色列的学校课程中增加一门判断力课。他预计课程准备将需要一两年的时间，结果拖了八年，到那时教育部已经失去了热情。这个令人羞惭的例子正说明了他和经常与他一起写书的阿莫斯·特沃斯基（Amos Tversky）所说的“规划谬误”。他更成功的工作是将心理学智慧渗入了经济学这个戒备森严的领域，经济学此前一直坚持着一种不太站得住脚但却条理清晰的人类决策模型。

他是如何做到的呢？一个原因是他与特沃斯基的合作。特沃斯基思维清晰简洁，一如他办公桌收拾得井井有条。他们将所发现的认知错觉纳入了一个称为“前景理论”的模型中。这个理论认为，人们的幸福感更多取决于财富的变化，而不是财富的绝对水平。财富的变化是相对于一个中立参考点来判断的。这个参考点并非总是显而易见，而且可以重新设定。如果一份奖金少于预期，就可能令人失望。在追求收益时，人们厌恶风险。在稳赢450美元和有50%的机会赢1000美元之间，他们会选择前者。但人们又会

为了规避损失而选择冒险，因为失去的痛苦要大于同等收益带来的快乐。

前景理论将这种决策模型从小故事转变成了代数和几何语言，这让经济学家接受了这种理论。事实上，经济学家开始把这种理论据为己有。对心理学的应用“开始被叫做行为经济学”，卡尼曼叹道，“许多心理学家发现，尽管他们的工作内容并没有变，他们的行当却换了个名字。”

| 冷手谬误

在经济学给心理学重新命名的同时，卡尼曼复兴了一种更为古老的经济传统——“快乐测量仪”，这是19世纪的经济学家弗朗西斯·埃奇沃思

（Francis Edgeworth）构想的测量快乐和痛苦的工具。卡尼曼的快乐测量仪直接请人们按一定分值即时评价自己的感受。他发现人们的评分经常与他们后来的回忆不符。他们的“记忆自我”过分重视一次经历的结尾及其最好或最坏的时刻，而忽视过程。人们宁可让手在冰冷刺骨的水中浸泡90秒而不是60秒——如果后30秒水温能略微调高的话。同样地，人们会报名参加行程安排得爆满的旅行，是因为他们期待自己日后回忆这些旅行，而不是因为在旅行时有多么享受其中。

这一发现的意义进入了哲学范畴。哪个自我才作数？将不具代表性的记忆精心编成人生故事，这样的“策展式”自我尽管有明显的缺陷，却受到人们的珍视。“我的记忆自我才是我，”卡尼曼写道，“而那个真实生活中的‘体验自我’对我来说倒像个陌生人。”现在，他的体验自我已经过完了一生，得由他在生命中触动到的众人来为他编写记忆了。■



Pharmacology

Could weight-loss drugs eat the world?

Scientists are finding that anti-obesity medicines can also help many other diseases

THE GILA monster is a poisonous North American lizard that measures around 50 centimetres and sports a distinctive coat of black and orange scales. This lethargic reptile, which mostly dwells underground and eats just three to four times a year, is the unlikely inspiration for one of pharma's biggest blockbusters: a new generation of weight-loss drugs that has patients—and investors—in a frenzy. Originally made for diabetes, evidence is growing that they also have benefits in diseases of the heart, kidney, liver and beyond.

Since the late 1980s scientists believed that a gut hormone called glucagon-like peptide-1 (GLP-1), which is secreted by the intestines after a meal, could help treat diabetes. GLP-1 increases the production of insulin (a hormone that lowers blood-sugar levels) and reduces the production of glucagon (which increases blood-sugar levels). But GLP-1 is broken down by enzymes in the body very quickly, so it sticks around for only a few minutes. If it were to be used as a drug, therefore, patients would have faced the unwelcome prospect of needing GLP-1 injections every hour.

In 1990 John Eng, a researcher at the Veterans Affairs Medical Centre in The Bronx, discovered that exendin-4, a hormone found in the venom of the Gila monster, was similar to human GLP-1. Crucially, the exendin-4 released after one of the monster's rare meals is more resistant to enzymatic breakdown than GLP-1, staying in its body for hours. It took more than a decade before exenatide, a synthetic version of the lizard hormone, created by Eli Lilly, an American pharma giant, and Amylin Pharmaceuticals, a biotech firm, was approved to treat diabetes in America. This breakthrough spurred other firms to develop more effective and

longer-lasting GLP-1 medications as a treatment option for diabetes, beyond injections of insulin.

Scientists had also been aware that GLP-1 had another side-effect: it slowed the rate of “gastric emptying”, which allows food to stay in the stomach for longer and suppresses appetite. But the potential weight-loss benefits were not seriously pursued at first. It was only in 2021 that Novo Nordisk, a Danish firm, showed data from a clinical trial where overweight or obese patients were put on a weekly dose of its GLP-1-based diabetic drug, semaglutide, which was then being marketed under the name Ozempic, for 68 weeks. The results were dramatic—participants had lost 15% of their body weight, on average.

| *Fat profits*

The medicines that mimic the GLP-1 hormone then became blockbusters. With close to half of the world’s population expected to be obese or overweight by 2030, according to the World Obesity Federation, demand for these drugs is surging—Bloomberg, a data provider, estimates that these medications will hit \$80bn in yearly sales by then. The market is projected to grow by 26% per year in the next five years, compared with 16% per year for oncology drugs and 4% per year for immunology medicines, the two other biggest areas.

So far only three GLP-1 drugs have been approved to treat obese or overweight individuals: liraglutide and semaglutide, developed by Novo; and tirzepatide, made by Lilly. But the market has already attracted a wave of competitors (see chart 1). Bloomberg tracks close to 100 wannabe drugs in the development pipeline. Most new therapies hope to outdo semaglutide and tirzepatide by crafting drugs that are easier to take, cause fewer side-effects or result in more effective weight loss (see chart 2).

One issue is convenience. Both semaglutide and tirzepatide are injections

that need to be taken weekly. Stop the dose and most of the weight returns within a year. Amgen, a large American biotech firm, is developing an anti-obesity drug that relies on doses once a month, and hopes the weight-loss effects will last even after treatment ends. AMG133 activates receptors for GLP-1 while blocking receptors of glucose-dependent insulintropic polypeptide (GIP), a hormone secreted in the small intestine in response to food intake that stimulates the production of both insulin and glucagon. The company is now conducting clinical trials to find out if patients can, over time, be gradually weaned towards smaller doses.

Switching from injections to pills would also make the drugs a lot more tolerable for those who dislike needles. Novo is working on an oral version of semaglutide that works just as well as its jabs. But the pill requires 20 times the amount of the active ingredient as the injection, and must be taken daily. With semaglutide in short supply, Novo has had to push back the oral version's launch. Lilly also has a daily pill that targets GLP-1 receptors called orforglipron in late-stage clinical trials.

Another drawback of GLP-1-based medicines is the nausea and vomiting that frequently accompanies their use. Zealand Pharma, a Danish biotech firm, is developing a drug that is based on a different hormone called amylin, produced in the pancreas along with insulin in response to food intake. But unlike GLP-1, which suppresses appetite, amylin induces satiety, or the feeling of fullness after a meal.

Adam Steensberg, boss of Zealand, says that in most people a hormone, leptin, is released from fat tissue that signals to the brain that the body is full. Obese individuals are insensitive to that hormone. Clinical studies have shown that analogues of amylin can make people sensitive to leptin again, helping them to stop eating earlier. Feeling full, rather than lowering appetite, may also reduce the feeling of nausea. Mr Steensberg says that results from early-stage trials suggest that its drug could achieve similar

weight loss as GLP-1 drugs, but with less nausea and vomiting.

Besides pesky injections and nausea, a bigger concern is that patients on these drugs do not just shed fat, they also lose lean muscle mass. Some patients drop almost 40% of their body weight in lean mass, a serious concern for older patients. To counter this, companies are trying out, alongside GLP-1 drugs, medicines originally designed to treat muscle atrophy.

Regeneron, an American pharma company, is testing drugs that block myostatin and activin, proteins that inhibit muscle growth in the body. Taken with semaglutide, the combination could boost the quality of weight loss by preserving lean muscle. Similarly, BioAge, a California-based biotech, is testing a drug that can be taken alongside Lilly's tirzepatide. The drug, called azelaprag, mimics apelin, a hormone secreted after exercise that acts on skeletal muscle, the heart and the central nervous system to regulate metabolism and promote muscle regeneration. In obese mice, the combination led to greater weight loss compared with tirzepatide alone, while preserving lean body tissue.

The slimming drugs aren't just for shedding pounds. Because obesity is linked to over 200 health issues, including strokes, kidney problems and fatty liver, GLP-1 drugs are proving useful in many other areas of medicine.

A recent clinical trial by Novo that ran for five years and enrolled more than 17,500 participants found that semaglutide cut the risk of serious heart issues like heart attacks, strokes, or death from heart disease by 20%. Novo believes that the heart benefits of the treatment are not due to weight loss alone, because the reduction in the risk of cardiovascular problems occurred early, before patients lost weight. In March semaglutide was approved by the US Food and Drug Administration for reducing the risk of heart disease in obese or overweight people, the first time a weight-loss

medication has been approved for this purpose. Results from another clinical trial have shown that semaglutide reduced the risk of kidney-disease-related events by 24% in patients with type-2 diabetes.

Another weight-loss drug, survodutide, being developed by Boehringer Ingelheim, a German drug company, and Zealand, has shown promising results in being able to treat a serious liver condition called metabolic dysfunction-associated steatohepatitis (MASH). This is caused by the build-up of excess fat in the liver and can lead to liver cancer or liver failure.

In a recent trial of 295 patients, 83% of them saw a significant improvement in their condition when treated with survodutide, compared with 18% of those on a placebo. Survodutide targets receptors for GLP-1 and glucagon. Waheed Jamal from Boehringer Ingelheim says that there is evidence that glucagon breaks down more fat in the liver compared with GLP-1 and reduces fibrosis (build-up of excessive scar tissue in the liver).

| *Gut meets brain*

Though a lot of focus has been on the action of these medicines on improving metabolic health, scientists are now uncovering that these drugs also engage with the brain and immune system, by interacting with GLP-1 receptors in the brain. Daniel Drucker, a diabetes researcher at Mount Sinai Hospital in Toronto, found that in mice suffering from extensive inflammation throughout the body, GLP-1 drugs reduced the condition, but only when the receptors in the brain were not blocked. When the brain receptors in mice were blocked or genetically deleted, the anti-inflammatory properties of the drugs were lost. This suggests that GLP-1 drugs tame inflammation by acting on the brain cells.

For some this suggests that these drugs might be useful for treating brain disorders that are characterised by inflammation, such as Alzheimer's disease and Parkinson's disease. Since 2021, Novo has been conducting a

clinical trial involving more than 1,800 patients to test whether semaglutide helps patients with early stages of Alzheimer's. This study is expected to be completed by 2026.

Dr Drucker sees the anti-inflammatory qualities of GLP-1 medications as key to their versatility. He notes that, besides Alzheimer's and Parkinson's, chronic inflammation is a factor in many complications for people with type-2 diabetes and obesity, and affects organs like the kidneys, heart, blood vessels, and liver. If these drugs eventually help in treating these conditions, Dr Drucker believes that their inflammation-reducing properties could explain part of their success.

The appetite-suppressing effects of these drugs has also raised interest in their ability to curb cravings more generally. Researchers in Denmark investigated the effect of GLP-1 drugs on 130 people with alcohol-use disorder. They found no overall difference in subsequent alcohol consumption between patients who used the drugs (alongside therapy) compared with those given a placebo. However, a subset of obese patients taking the drugs did end up drinking less alcohol. The researchers also looked at brain activity in the patients when they were shown pictures of alcoholic drinks—for those in the placebo groups the reward centres of their brains lit up; for patients on GLP-1 drugs, activity in the areas of the brain associated with reward and addiction was attenuated, indicating a direct brain effect. Researchers are now exploring if the drugs might have an impact on how people use other addictive substances such as tobacco or marijuana.

All these findings are still early. Developing new drugs is costly and time-consuming. There are steep failure rates. Successes in the lab may not work in people, and results in small groups may not replicate in larger ones. But with the potential to treat many conditions well beyond obesity and diabetes, hope around the new drugs will only grow. ■



药理学

减肥药能通吃天下吗？

科学家们发现，**减肥药也有助于治疗许多其他疾病【深度】**

吉拉毒蜥是北美一种有毒的蜥蜴，它体长约50厘米，全身覆盖着一层独特的黑橙相间的鳞片。这种行动迟缓的爬行动物大部分时间生活在地下，一年只进食三到四次。很难想象，它竟然会成为制药业最畅销的产品之一——新一代减肥药的灵感来源。这些减肥药令患者和投资者都为之疯狂。尽管研发它们的初衷是治疗糖尿病，不过越来越多的证据表明，它们对心脏、肾脏、肝脏等其他方面的疾病也有疗效。

上世纪80年代末以来，科学家们认为，餐后肠道分泌的一种名为胰高血糖素样肽-1（GLP-1）的肠道激素可以帮助治疗糖尿病。GLP-1能促进胰岛素（一种能降低血糖水平的激素）的分泌，并抑制胰高血糖素（会提高血糖水平）的分泌。但GLP-1很容易被人体内的酶快速分解，所以只能存留几分钟。因此，如果把它用作药物，患者每隔一小时就需要注射一次GLP-1，让人不胜其烦。

1990年，位于纽约布朗克斯区的退伍军人医疗中心（Veterans Affairs Medical Centre）的研究员约翰·恩格（John Eng）发现，在吉拉毒蜥毒液中发现的一种激素exendin-4与人体中的GLP-1相似。关键是，吉拉毒蜥在它难得一次的进食后分泌的exendin-4比GLP-1更不容易被酶分解，能在体内存留数小时。经过十多年的研究，美国制药巨头礼来（Eli Lilly）和生物科技公司Amylin Pharmaceuticals合成的类似exendin-4的药物exenatide终于在美国被批准用于糖尿病的治疗。在这一突破性进展的带动下，其他公司开始研发效果更好、作用更持久的GLP-1药物。这让糖尿病患者除了注射胰岛素之外多了一种治疗方案。

科学家们也意识到GLP-1还有另一个附带功效：它能减缓“胃排空”的速度，让食物在胃中停留更长时间，从而抑制食欲。但这种潜在的减肥功效最初并没有得到深入研究。直到2021年，丹麦诺和诺德公司（Novo Nordisk）

才公布了一项临床试验数据：超重或肥胖患者连续68周、每周服用一次GLP-1类糖尿病药物司美格鲁肽（商品名为Ozempic），结果令人吃惊——参与者的体重平均减轻了15%。

| 减肥有厚利

模拟GLP-1激素的药物随即风靡一时。世界肥胖联盟（World Obesity Federation）预计，到2030年，全球将有近一半的人口肥胖或超重，因此对这些药物的需求正在激增——数据提供商彭博估计，届时这些药物的年销售额将达到800亿美元。预计未来五年里这个市场的年增长率将达到26%，而另外两个最大的药物市场——抗肿瘤药物和免疫类药物的年增长率分别为16%和4%。

截至目前，只有三种GLP-1药物获批用于肥胖或超重人群的治疗：诺和诺德研发的利拉鲁肽和司美格鲁肽，以及礼来生产的替尔泊肽。但该市场已经吸引了一大波竞争者（见图表1）。彭博追踪了近100种此类在研药物。大多数新药物都希望实现更方便的给药、更小的副作用或者更好的减肥效果等，来超越司美格鲁肽和替尔泊肽（见图表2）。

便利性是一大问题。司美格鲁肽和替尔泊肽这两种注射剂都需要每周注射一次。停药后，大部分减掉的体重会在一年内反弹。美国大型生物技术公司安进（Amgen）正在研发一种只需每月注射一次的减肥药AMG133，并希望即使在治疗结束后减肥效果仍能保持。AMG133能在抑制葡萄糖依赖性促胰岛素多肽（GIP）受体的同时激活GLP-1受体。GIP是小肠在食物摄入时分泌的一种激素，能刺激胰岛素和胰高血糖素的分泌。安进目前正在进行临床试验，以确定患者能否在用药一段时间后逐步减小剂量。

把注射剂改成口服药也会让那些不喜欢打针的人更容易接受。诺和诺德正在研发一种口服版的司美格鲁肽，效果与注射剂一样好。不过该药片所需的活性成分含量是注射剂的20倍，且必须每天服用。由于司美格鲁肽供应短缺，诺和诺德不得不推迟了口服版的上市时间。礼来正在进行后期临床试验的奥格列龙也是一种需每天服用的GLP-1受体激动剂。

GLP-1类药物的另一个缺点是用药时经常伴有恶心和呕吐。丹麦生物技术公司Zealand Pharma正在研发一种药物，基于另一种叫作胰淀素的激素。人在进食时，胰腺不仅会分泌胰岛素，还会分泌胰淀素。但与抑制食欲的GLP-1不同，胰淀素会诱发饱腹感。

Zealand的老板亚当·斯坦斯伯格（Adam Steensberg）表示，大多数人在吃饱后，脂肪组织会分泌一种名为瘦素的激素，向大脑发出吃饱了的信号。肥胖人群对瘦素不太敏感。临床研究表明，胰淀素的类似物可以让人重新对瘦素敏感，帮助他们早一点停止进食。和抑制食欲不同，产生饱腹感还可能减少恶心的感觉。斯坦斯伯格表示，早期试验结果表明，Zealand的这种药物可以达到与GLP-1药物类似的减肥效果，而恶心和呕吐会少一些。

除了恼人的注射和恶心感之外，更大的问题是使用这些药物的患者减掉的不仅是脂肪，还有肌肉。有些患者的去脂体重下降了近40%，这对上年纪的患者来说是个严重问题。为了解决这个问题，各家公司都在尝试将原本用于治疗肌肉萎缩的药物与GLP-1药物搭配使用。

美国再生元制药（Regeneron）正在测试的药物能阻断两种抑制人体肌肉生长的蛋白质——肌肉生长抑制素和激活素。将这种药物与司美格鲁肽一起使用，可以保留肌肉，从而提高减肥的质量。同样，加州的生物技术公司BioAge也在测试一种可与礼来的替尔泊肽搭配使用的药物azelaprag。这种药物类似在运动后分泌的激素爱帕琳肽，爱帕琳肽作用于骨骼肌、心脏和中枢神经系统，能调节新陈代谢并促进肌肉再生。BioAge将它与替尔泊肽在肥胖小鼠身上搭配使用。相比于单独使用替尔泊肽，联合用药减掉了更多体重，同时保留了非脂肪组织。

减肥药的用途不仅限于减掉体重。由于肥胖与包括中风、肾脏疾病和脂肪肝等在内的200多种健康问题有关，GLP-1药物被证明还有助于治疗很多其他疾病。

诺和诺德近期一项为期五年、招募了超过1.75万名参与者的临床试验发

现，司美格鲁肽能将心肌梗塞、中风或致死性心脏疾病等严重心脏问题的风险降低20%。诺和诺德认为，该疗法之所以有益心脏，并不全然因为体重减轻，因为在患者减轻了体重之前，心血管疾病的风险就已经先一步降低。今年3月，司美格鲁肽获得了美国食品和药物管理局（FDA）的批准，可用来降低肥胖或超重人群罹患心脏病的风险，这是减肥药首次获批用于这一目的。另一项临床试验结果显示，司美格鲁肽可将二型糖尿病患者发生肾脏相关疾病的风险降低24%。

由德国勃林格殷格翰制药公司（Boehringer Ingelheim）和Zealand共同研发的另一种减肥药survodutide，在治疗一种叫作代谢功能障碍相关脂肪性肝炎（MASH）的严重肝病方面显示出了良好前景。这种疾病是由肝脏中过多脂肪的堆积所引发，可导致肝癌或肝功能衰竭。

不久前对295名患者进行的一项试验显示，在接受survodutide治疗的患者中，有83%的人病情明显改善，而在服用安慰剂的患者中的比例为18%。survodutide可以激活GLP-1和胰高血糖素受体。勃林格殷格翰的瓦希德·贾马尔（Waheed Jamal）表示，有证据表明，与GLP-1相比，胰高血糖素能分解更多肝脏中的脂肪，并减少纤维化（肝脏中形成过多瘢痕组织）。

｜ 肠脑相通

尽管对这些药物的关注多集中在它们改善新陈代谢方面的作用，但如今科学家们发现，这些药物还与大脑中的GLP-1受体相互作用，进而影响大脑和免疫系统。多伦多西奈山医院（Mount Sinai Hospital）的糖尿病研究员丹尼尔·德鲁克（Daniel Drucker）发现，在全身存在大量炎症的小鼠身上使用GLP-1药物，能减轻其病情，但前提是小鼠大脑中的受体没有被阻断。如果小鼠的大脑受体被阻断或被基因删除时，药物就会丧失抗炎作用。这表明GLP-1药物是通过作用于脑细胞来抑制炎症的。

一些人认为，这表明这些药物可能对阿尔茨海默病、帕金森病等一些带有炎症表现的脑部疾病有疗效。自2021年以来，诺和诺德一直在进行一项有1800多名患者参与的临床试验，以测试塞马格鲁肽是否对阿尔茨海默病的早期患者有帮助。这项研究预计在2026年完成。

德鲁克认为，GLP-1药物之所以用途广泛，关键在于它的抗炎作用。他指出，除了阿尔茨海默病和帕金森病之外，慢性炎症也是导致二型糖尿病和肥胖症患者患上多种并发症的因素之一，同时还影响肾脏、心脏、血管和肝脏等器官。德鲁克认为，如果GLP-1药物最终有助于治疗这些疾病，它的消炎作用应该有一定功劳。

GLP-1药物抑制食欲的作用也引发了人们探究它们能否抑制更多其他欲望。丹麦的研究人员调查了GLP-1药物对130名酒精使用障碍患者的影响。他们发现，与服用安慰剂的患者相比，在接受相关治疗的同时还服用GLP-1药物的患者的后续饮酒量总体上没有差异。不过，其中一小批服药的肥胖患者确实减少了饮酒量。研究人员还观察了患者在看到酒精饮料图片时的大脑活动——服用安慰剂的患者大脑的奖赏中心亮了起来，而服用GLP-1药物的患者大脑中与奖赏和上瘾相关的区域的活动减弱了，表明其大脑受到了直接影响。研究人员目前正在研究这些药物是否会影响人们对其他成瘾性物质的使用，比如烟草、大麻等。

所有这些都只是初步研究结果。新药研发费钱又耗时。失败率非常高。在实验室有效并不意味着一定对人体有效，在小群体中取得的成效也不一定会在大群体中复现。不过，鉴于这些药物的治疗范围可能远不止肥胖和糖尿病，人们对它们的期望只会越来越高。■



Chinese aspirations

Yu Hua on why young Chinese no longer want to work for private firms

In today's China stability trumps enterprise, argues the novelist

WHAT LIES at the far reaches of the universe? This is a question that science has yet to answer, but most young people in China today already have an answer. According to them, at the end of the universe is not the Milky Way, the Andromeda Galaxy or the Canes Venatici Constellation, but a government job.

Observing changing attitudes about government jobs among young people in China over the course of the past four decades of the reform era can reveal the deep-seated changes that have taken place in the structure of Chinese society.

During the early days of the reform era in the 1980s, there were three categories of government jobs for urban residents: the collective positions (*jiti bianzhi*), the general positions (*quanmin bianzhi*) and the cadre positions (*ganbu bianzhi*). The cadre positions were equivalent to the civil-service system of today. At the time, all young people craved an opportunity to enter that system. But by the 1990s that was no longer something they desired.

If we were to describe the pace of economic development in the 1980s as walking, by the 1990s it was akin to a horse galloping. College graduates in particular turned their backs on the bureaucratic system in favour of seeking employment with foreign firms or joint ventures, or founding their own firms. The incentive was obvious: they could earn more, possibly a lot more, in the private sector.

In 1992 alone, more than 120,000 cadre-level bureaucrats—the backbone of

various government departments at the time—resigned from their positions to work in the private sector. Seeing the tremendous opportunities of the market economy, some sought employment with foreign firms or joint ventures. Others drew on their savings and borrowed money from friends and family to start their own businesses. The most outstanding of them would later become China's best-known entrepreneurs. Those taking the private-sector plunge described it as “diving into the sea”, as the market economy was metaphorically referred to as a great sea.

But as the Chinese economy slowed, young people began to rekindle their desire for jobs within the government system—and this time their desire was even more fervent. It was referred to as “coming ashore”.

For the past decade the number of Chinese college graduates applying for jobs in the civil service has surged. In 2023 some 2.8m applicants qualified for the civil-service examinations, chasing an estimated 39,000 available positions. The media described this as “an army of soldiers and horses attempting to cross a single-plank bridge”.

What's more, having a bachelor's degree is no longer an advantage; one needs to have at least a master's degree. And for an application to be truly competitive, one needs a doctorate.

The fierce competition in the labour market has deprived young job-seekers of the ability to be selective when it comes to employment; as long as there is a secure government bureaucratic job available, no matter where or what the nature of the job, there are always large numbers of applicants. In 2020 a small subdistrict office in Hangzhou made the news when people with PhDs from some of China's most prestigious universities were among those applying for an open position. In the remote Ali region of Tibet, where living conditions are extremely difficult, a job advertised in 2021 for

a clerk at the local post office drew a staggering 20,813 applications.

Shutdowns, layoffs, unemployment and bankruptcy have all become key words to describe the economy of today. For young Chinese, high-paying jobs are no longer the be-all and end-all; what they want are stable jobs within the state bureaucratic system. After living through three years of pandemic lockdowns, it is not only the young who have realised the importance of these bureaucratic jobs. Everyone has.

No matter how long you might end up quarantined at home, government jobs are the only form of employment you can count on to continue paying you a full salary. Without a government job, the second the next lockdown comes along, your source of income may well be cut off. That is why “At the far reaches of the universe lies a government job” has become a popular phrase online in China today.

The Chinese people have always loved palm-reading, which is the simplest and most straightforward form of fortune-telling. In the past, palm-readers would base someone’s fortune on three lines which represented one’s longevity, career and love life. When young people get their fortunes told today, they care only about the career line, and all they want to know is: will there be a government job in their future.

I don’t know how these young people invented this new form of divination, but I can only marvel at the awe-inspiring power of government jobs, as they have now even entered into the long tradition of Chinese fortune-telling. ■



中国式抱负

余华谈为何中国年轻人不再想去私营部门

这位作家说，在当今的中国，稳定胜过进取

宇宙的尽头是什么？科学界对这个问题尚无定论，但今天大多数的中国年轻人已经有了答案。在他们看来，宇宙尽头不是银河系、仙女座或猎犬座，而是编制。

观察改革开放40年来中国年轻人对编制的态度变化，可以揭示中国社会结构发生的深层次变化。

在1980年代改革开放初期，面向城镇居民的体制内工作有三类：集体编制、全民编制和干部编制。干部编制相当于今天的公务员制度。当时，所有的年轻人都渴望有个干部编制。但到了90年代，他们已经不这么想了。

如果把1980年代的经济发展速度比作步行，那么90年代的发展就是骏马飞奔。大学毕业生尤其对政府工作失去了兴趣，更愿意到外资或合资企业工作，或者自己开公司。这样做的动机显而易见：在私营部门可以赚更多钱，而且可能多很多。

仅1992年一年，就有超过12万名政府中的干部（当时各级政府部门的骨干）辞职去私营部门工作。看到市场经济的巨大机遇，一些人到外企或合资企业谋职。另一些人用自己的积蓄或向亲朋好友借钱开办了自己的企业。其中最优秀的一些人后来成为了中国最知名的企业家。那些毅然投身私营部门的人把这叫做“下海”，因为市场经济被比喻为汪洋大海。

但随着中国经济放缓，年轻人又重新燃起了对编制内工作的渴望——这一次他们的渴望甚至更加强烈。这就成了所谓的“上岸”。

在过去十年里，中国大学毕业生报考公务员的人数激增。2023年，约280万报名者取得了参加公务员考试的资格，竞逐约3.9万个岗位。媒体形容这是“千军万马过独木桥”。

而且，有学士学位已经不再是个竞争优势，最起码要有硕士学位。而要想在考编中真正有竞争力，还得有博士学位才行。

劳动力市场竞争激烈，让年轻求职者没法挑挑拣拣；只要是一份稳定的政府编制工作，无论工作地点在哪里、工作性质是什么，总会吸引到大量的申请者。2020年，杭州一个小街道办的空缺岗位招聘成为了新闻热点，因为报考者当中不乏中国最知名大学的博士。在生活条件极其艰苦且偏远的西藏阿里地区，2021年当地邮政局招聘一名科员时竟吸引了20,813人报名。

停工、裁员、失业和破产均已成为如今经济的关键词。对于中国年轻人来说，高薪工作不再是终极目标，他们想要的是一份体制内的稳定工作。经历了三年的疫情封控后，不仅年轻人体会到了有编制的重要性。每个人都体会到了。

无论可能要居家隔离多久，唯有编制内工作你可以放心它会不间断全额发放工资。没有编制，下一次封锁到来时，你的收入来源很可能会被切断。正因如此，“宇宙的尽头是编制”在当今中国成了网络流行语。

中国人历来爱看手相，这是最简单、最直接的算命方式。在过去，看手相的人会根据生命线、事业线和感情线这三条掌纹来给人算命。如今年轻人在算命时只看事业线，他们只想知道：我的命里带编制吗？

我不知道这些年轻人是如何发明这种新的占卜方式的，但我只能感叹于编制的惊人魔力，甚至连中国悠久的算命传统都在为它服务了。■



Culp able

What Boeing, Disney and others can learn from General Electric

Lessons from the tenure of Larry Culp

ONLY RARELY are chief executives appointed with a mandate for dismemberment. Yet when Larry Culp assumed the top job at General Electric (GE) in October 2018, he was expected to sell parts of the 130-year-old conglomerate at a pace even faster than his empire-building predecessors had assembled them. In November 2021 he announced a radical finale: splitting the firm in three. GE's health-care business became a separate company last year. On April 2nd its power division went the same way, leaving behind GE Aerospace, the firm's engine-making operation.

Investors are reaping the rewards. After dithering during the first four years of Mr Culp's tenure, GE's shares have been on a jet-fuelled tear. The cumulative stockmarket value of GE's three successor firms is \$237bn. Although that is well below the firm's peak of \$594bn in 2000, it is more than double what Mr Culp inherited. Business-school students have spent decades dissecting the hubristic acquisitions that defined GE's life. Now they must heed the lessons from its final act.

The fact that the wind is at GE's back has undoubtedly helped. Demand for jet engines has picked up since the covid-19 pandemic ended and clean-energy projects are benefiting from government handouts in America. The bullish mood of America's stockmarket has surely eased the separation process, too. But Mr Culp also deserves praise. His tenure reflects not just a victory of focus over sprawl, but also of operational excellence at a firm that long suffered the effects of excessive financial engineering.

One advantage was that Mr Culp was an outsider—the first ever to take the

helm at GE. The turnaround probably could not have been pulled off by a lifer. GE insiders had ossified. Nothing had replaced Six Sigma, a system used to reduce manufacturing errors that was championed by Jack Welch, the firm's acquisitive boss who retired as long ago as 2001. Mr Culp brought his own Japanese-style management philosophy. The fresh but fanatical approach to continuous operational improvement helped usher in cultural change. So did getting executives onto the factory floor.

The new thinking meant that GE's assets did not languish on the chopping block, but improved on Mr Culp's watch and stopped the break-up becoming a fire sale. The programme has a familiar taste for fanciful names—the power business has been christened GE Vernova. But the markets are cheering on GE's heirs, which is striking given that spin-offs often disappoint investors. IBM's infrastructure-services business (Kyndryl) and Johnson & Johnson's consumer health-care operation (Kenvue) have both failed to shine in the public markets.

Mr Culp's success, then, is vindication for smooth operators, not firms looking to ditch problem assets or activist hedge funds in search of conglomerate discounts. Bosses preparing to lop off divisions of their own firms, including those at Unilever, will study him closely. But there are broader lessons, too. Disney, which recently prevailed in a proxy fight against Nelson Peltz, a noisy shareholder who once lobbied for change at GE, is just one American blue-chip firm that might benefit from fresh thinking.

The lessons may be most apt for Boeing, a case study for rotten corporate culture. In March its boss resigned, two months after the blowout of a fuselage panel from one of its 737 MAX planes. Some have suggested Mr Culp for the job, which would be a severe test of his operational prowess. In March Boeing confirmed that it was in discussions to acquire Spirit AeroSystems, an unsatisfactory supplier it had previously spun off. Mr Culp

says he is happy running GE Aerospace, one of Boeing's suppliers. But if he can manage a business well while pulling it apart, perhaps he could also thrive at putting one back together again. ■



【首文】卡尔普真行

波音、迪士尼和其他公司可以向通用电气学习什么

拉里·卡尔普任期内的经验

极少有首席执行官在接受任命时会被委以拆解公司的责任。不过，当拉里·卡尔普（Larry Culp）于2018年10月担任通用电气（GE）的最高职位时，人们指望他迅速出售这家有130年历史的企业集团的部分业务，甚至要比他那些建立帝国的前任们组建集团的速度更快。2021年11月，他公布了激进的最终方案：将公司一分为三。GE的医疗业务在去年成为一家独立的公司。今年4月2日，GE的电力部门也走上了同样的道路，留下的是GE航空航天（GE Aerospace），这是GE的发动机制造部门。

投资者正在收获回报。GE的股价在卡尔普任职的头四年原地打转，之后便坐上了喷气式飞机。GE拆分出的三家公司的股票市值总计为2370亿美元。尽管这一数字远低于该公司2000年时5940亿美元的峰值，却是卡尔普接手时的两倍多。几十年来，商学院的学生们一直在剖析那些贯穿GE历史的狂妄收购。现在他们还得从它的终章里吸取经验。

毫无疑问，GE是借到了一股东风。自新冠疫情结束以来，喷气发动机的市场需求回升，清洁能源项目也从美国政府的补贴中受益。美国股市的看涨情绪无疑也让分拆更为容易。不过卡尔普也值得赞扬。他的任期不仅是一场用专注取代扩张的胜利，也是在这家长久受金融工程过度之苦的公司实施卓越运营的胜利。

卡尔普的一个优势是他是外来户——他是第一个执掌GE的外人。一个在GE待了一辈子的人可能无法实现这样的逆转。它的内部人士早已思维僵化。用以减少生产错误的六西格玛迟迟未被替换，这是热衷收购的公司前老板杰克·韦尔奇倡导的体系，而韦尔奇早在2001年就退休了。卡尔普带来了自己的日式管理哲学。这种新颖而狂热的方法持续改进运营，帮助引入了文化变革。让高层人员走进工厂车间也起了同样的效果。

这种新思路意味着GE的资产没有在等待拆分时贬值，而是在卡尔普的监督下得到了提升，也避免了让拆分沦为贱卖。和人们经常看到的一样，这次拆分行动也炮制了一些别出心裁的名字——电力业务就被命名为GE Vernova。不过市场正在为GE的继承者们欢呼，考虑到分拆经常令投资者失望，这一点引人注目。IBM的基础设施服务业务勤达睿（Kyndryl）和强生的消费者保健业务科赴（Kenvue）都未能在公开市场上大放异彩。

因此，卡尔普的成功证明正确的是稳健的经营者，而不是想甩掉问题资产的公司，或是追求多元化折让的维权对冲基金。包括联合利华在内，那些准备砍掉自己公司一些部门的老板们会仔细观察他。但还有适用范围更广的经验。最近，迪士尼在代理权争夺战中战胜了尼尔森·佩尔茨（Nelson Peltz），这位好斗的股东曾游说GE进行变革。迪士尼只是有可能从转换思路中受益的美国蓝筹公司之一。

这些经验或许最适合波音公司，它是企业文化腐朽的一个典型案例。今年3月，在一架737 MAX飞机机身面板爆裂两个月后，波音的老板辞职了。有人提议由卡尔普担任这一职位，这将极大考验他的运营能力。3月，波音证实它正在讨论收购势必锐航空系统公司（Spirit AeroSystems），这家产品不如人意的供应商原本就是从波音剥离出来的。卡尔普表示，他很乐意执掌为波音供货的GE航空航天。但如果他能在拆分一家企业的时候很好地管理它，那或许他也能很好地把一家公司重新组合起来。■



Free exchange

How India could become an Asian tiger

The world's most selective bureaucracy is struggling to make it happen

IS INDIA ACHIEVING its potential? In the year to the third quarter of 2023, the country grew at a blistering rate of 8.4%. Over the next half-decade it is expected to expand at 6.5% a year, which would make it the world's fastest-growing big economy. So far, so good. The problem, as critics point out, is that China, Japan and South Korea all expanded at 10% or so a year during their periods of rapid growth. Part of the reason for India's less impressive figures is a slowdown in globalisation. But a new book by Karthik Muralidharan of the University of California, San Diego, called "Accelerating India's Development", argues that the crucial barrier to faster development is a lack of "state capacity".

Mr Muralidharan describes this concept as the "effectiveness" of government. Throwing money at a state lacking capacity is like adding fuel to a car near a breakdown: it won't get you very far. Currently, the Indian state succeeds when on "mission mode", achieving clearly defined goals. In April it should pull off the largest democratic exercise in history, as voters pick a prime minister. At the same time, it struggles with mundane, everyday aspects of governance, such as education and health. Three in five rural children in the fifth year of school cannot read at a second-year level—and in the past five years the failure rate has only worsened.

Part of the issue is the precociousness of Indian democracy. The franchise became universal in 1950, when the country was mostly impoverished. Citizens demanded that the state met their basic needs well before it had the money or capacity to do so. India launched its food-security programme in the 1960s, for instance, when it was a fifteenth as rich per person as America was when it launched its own such programme in the

1930s. This set a pattern: the Indian state does a lot, but little well.

In his 800-page tome, Mr Muralidharan lays out fixes. The book is crammed with details about how the Indian system works and could be improved. Three ideas stand out. They concern how the state should manage people, use technology and improve its federal system. And they hold lessons for other governments.

Since 2002, when Mr Muralidharan was a graduate student, he has been conducting surveys on absentee rates. It turns out that teachers skip school perhaps as often as pupils: they are absent 20-30% of the time. The problem is not pay. In 2017 a study by Rohini Pande of Yale University found that across 33 countries, India offered the second-highest wage premium to public-sector employees. Rather, the problem appears to be governance. School supervisory positions have ultra-high vacancy rates of 20-40%. Mr Muralidharan calculates that filling such roles would be ten times more cost-effective than hiring more teachers.

Such problems reflect a strange approach to civil-service management. About 1m Indians apply each year for about 1,000 slots, making it one of the most selective bureaucracies in history. Yet India has just 16 public employees per 1,000 people, one of the lowest ratios in the world. Indian officials also hop from position to position across various different tiers of government, moving every 15 months on average. As a recent retiree notes, this means that they have far too little time to develop proper expertise in any one area. Therefore problems go unresolved.

Better data would help. It should not take an academic survey to ascertain rates of teacher absenteeism. Official numbers on educational outcomes paint a far rosier picture than those collected by independent organisations. Schools and low-level bureaucrats have incentives to cheat. As such, Mr Muralidharan suggests digital data collection, audits and stiffer

accountability.

A running theme throughout the book is the need to pay attention to India's federal structure. The country has 28 hugely diverse states, 15 of which are big enough to be in the top fifth of the world's countries by population. Historically, Indian politicians have fallen prey to cycles of corruption in which companies bribe leaders in exchange for favours, with leaders then using the money to fund expensive campaigns that involve, in effect, bribing voters. In some states, this cycle is starting to break down. Campaigns promising to get stuff done are precisely what propelled Narendra Modi, India's prime minister, to fame when he ran the state of Gujarat.

Progress in the "third tier" of government, which is made up of villages, towns and cities, has been less impressive, however. Deng Xiaoping, China's leader from 1978 to 1989, introduced reforms that rewarded local officials for their economic performance. India could use something similar, especially if states granted their local authorities more resources and power. Of the 18 similar countries looked at by Mr Muralidharan, India was the least decentralised when it came to fiscal affairs. Just 3% of state spending happens at a local level, compared with 51% in China. In 2000 three Indian states—Bihar, Madhya Pradesh and Uttar Pradesh—were split in two. So was Andhra Pradesh in 2014. All four splits have led to faster economic growth.

| *Go-faster stripes*

There is much else in Mr Muralidharan's book. But changes to personnel, data and federalism would together represent the beginnings of a reform package that could do wonders for India's governance, and therefore its growth. Indeed, the sheer variation in performance across India's cities and states demonstrates the potential benefits. Karnataka, a large and reasonably well-run state, is six times richer per person than Bihar, the

country's poorest state—making India one of the world's most geographically unequal countries. Internal migration is relatively rare, which means that for incomes to grow faster, poor states will have to get their act together. The changes outlined by Mr Muralidharan will be tough to enact, but not impossible. And if some states make the leap, they may inspire others to follow their example. ■



自由交流

印度要如何成为亚洲之虎

全球最挤破头的官僚系统正在艰难追寻这一目标

印度是否正在实现自身潜力？截至2023年第三季度的一年里，该国以8.4%的惊人速度增长。未来五年，其年增长率预计将达到6.5%，使其成为全球增长最快的大型经济体。目前为止，一切顺利。但批评者指出，中国、日本和韩国在它们的快速增长期内年增长率都达到了10%左右。印度的增长数字相比之下逊色了些，原因之一是如今全球化进程放缓了。但加州大学圣地亚哥分校的卡尔希克·穆拉里塔兰（Karthik Muralidharan）在他的新书《加速印度发展》（Accelerating India's Development）中指出，阻碍印度实现更快发展的关键是缺乏“国家能力”。

穆拉里塔兰解释说，这一概念是指政府的“效能”。向一个能力不足政府投入资金就像给一辆行将出故障的汽车加油：走不了多远。目前，印度政府在“任务模式”下最易成功，能够实现界定清楚的目标。4月，随着选民选出总理，印度将完成史上规模最大的民主选举。但同时，印度在教育和医疗等不那么起眼的日常治理上却举步维艰。在农村地区，五年级学生中有五分之三达不到二年级的阅读水平，过去五年里这个比例还在上升。

这一定程度上是因为印度民主的早熟。1950年，印度开始实行普选，但当时该国大部分人口还处于贫困之中。公民要求国家满足自己的基本生活需求，而那时候的印度政府远没有这样的资金或能力。例如，印度在1960年代启动了粮食安全计划，当时其人均收入仅为美国在1930年代启动类似计划时的十五分之一。这就形成了一种模式：印度政府做了很多事，但没几件做得好。

穆拉里塔兰在他800页的长篇大论中提出了解决方案。书中充满了有关印度的体制如何运作以及如何改进的细节，其中三个想法尤为引人注目，涉及印度政府该如何管理民众、利用技术和改进联邦制度。其他政府也可以从中获得启发和借鉴。

穆拉里塔兰从2002年读研究生时起就开始调查缺勤率，他发现教师缺勤的频率堪比学生：达到20%至30%。问题不是出在工资上。耶鲁大学的罗希尼·潘德（Rohini Pande）在2017年所做的一项研究发现，在33个国家中，印度的公共部门雇员工资溢价位居第二。问题似乎是出在治理上。学校中管理职位的空缺率尤其高，达到20%至40%。按穆拉里塔兰的计算，填补这些职位的成本效益将是聘用更多教师的十倍。

这些问题反映出一种奇怪的公务员管理方式。每年约有100万印度人申请约1000个公务员职位，令印度政府成为史上入职竞争最激烈的官僚机构之一。然而，印度每一千人中只有16名公务员，比例居全球最低之列。印度官员还经常在政府各级部门之间调动，平均每15个月调动一次。正如一位最近退休的官员指出，这意味着他们在任何部门待的时间都太短了，无法发展出适当的专业能力。因此问题也就没有能力被解决。

改善数据收集会有帮助。不应依靠学术调研来弄清教师缺勤率。官方教育成果数据对现实的粉饰远远超过独立组织收集的数据。学校和基层官员有作弊的动机。因此，穆拉里塔兰建议采用数字化数据收集、审计和更严格的问责制。

贯穿全书的一个主题是需要关注印度的联邦结构。印度的28个邦差异巨大，若按人口规模算，其中15个邦足以跻身全球前五分之一国家之列。一直以来，印度政客总是陷入腐败循环——企业贿赂领导人以换取关照，然后领导人用这些钱支持成本高昂的竞选活动，实际上涉及贿赂选民。在一些邦，这个循环开始瓦解。正是承诺把事情做好的竞选活动推动了现任总理莫迪在执掌古吉拉特邦（Gujarat）时声名鹊起。

然而在包括村、镇和市在内的“第三级”政府中，进展没那么大。1978年至1989年间担任中国领导人的邓小平在任内曾推行改革，根据经济发展成绩奖励地方官员。印度可以采用类似的做法，特别是各邦可以下方更多资源和权力给下级地方政府。在穆拉里塔兰研究的18个类似国家中，印度下放的财政权力最少。地方政府只能支配邦政府支出的3%，而在中国这一比例为51%。2000年，印度有三个邦一分为二，包括比哈尔邦（Bihar）、中

央邦（Madhya Pradesh）和北方邦（Uttar Pradesh）。安得拉邦（Andhra Pradesh）也在2014年分为两个邦。这四次切分都带来了更快的经济增长。

｜ 赛车条纹

穆拉里塔兰的书中还有很多其他内容。但人事、数据和联邦制度上的变革可以成为有望给印度治理乃至增长带来奇迹的一整套改革方案的开端。事实上，印度各市和各邦之间在绩效上的巨大差异就展示了变革的潜在益处。卡纳塔克邦（Karnataka）是个管治良好的大邦，人均收入是最贫穷的比哈尔邦的六倍，印度因而成为世界上地区财富最不平等的国家之一。印度国内移民相对较少，这意味着要加快收入增长，贫穷的邦必须自己发奋图强。穆拉里塔兰提出的改革实施起来虽然困难，但并非不可能。如果一些邦能实现飞跃，就可能激励其他邦效仿。 ■



Killers' tactics

Killer whales deploy brutal, co-ordinated attacks when hunting

Their techniques are passed down through the generations

THE KILLER whales that roam the waters of the Monterey Submarine Canyon, off the coast of California, are brutal animals. When hunting grey whales, they target mother and calf pairs, chasing them until the calves begin to tire. At that point the killer whales separate the grey whales by manoeuvring between them or dragging the calves away by their tail flukes or fins. Once a calf is separated from its mother, the killer whales incapacitate it by bludgeoning it with their heads and tails before drowning it by leaping atop its blowhole to keep it below the surface.

The killer whale (*Orcinus orca*) is a diverse species, boasting several isolated populations around the world that occupy varying—and, in several cases, still little described—ecological niches. Two recent studies have shed some light on the unique hunting cultures of the killer whales that stalk the deep waters off the north-east Pacific Ocean.

Distinct killer whale populations are known as “ecotypes” and can differ in size, colour and body proportion. Where their ecological ranges overlap, ecotypes rarely interact and do not appear to interbreed. Most of what makes killer-whale ecotypes distinct is their culture—socially learned behaviours that killer whales are known to transmit from one generation to the next. These include preferences for different marine prey species and the collaborative strategies used to locate, hunt and kill them, potentially honed over millennia. Different killer-whale ecotypes even communicate using distinct “dialects” made of clicks, whistles and calls, with “accents” that differ between pods.

The most studied killer-whale ecotypes are the three that occupy the north-

east Pacific coastline, from northern Alaska to southern California. “Offshore” killer whales, up to 6.6 metres long, patrol distant waters in pods of 100 to 200 individuals, hunting pelagic fishes like sharks. “Resident” killer whales, around 7 metres in length, stalk the coastline in pods of up to 25 individuals, targeting coastal fishes (primarily salmon). Finally, “transient” killer whales, each up to 8 metres in length, are found across both shallow and deep coastal waters and pursue marine mammals, including cetaceans (whales and dolphins), as well as pinnipeds (sea lions and seals). While “inner-coast” transients stalk shallow waters near the shore, “outer-coast” transients hunt in the deep open waters along canyons at the edge of the continental shelf.

To understand these transient killer whales better, Josh McInnes from the University of British Columbia in Canada led a team that analysed 261 sightings of killer whales around the Monterey Submarine Canyon System collected by surveys and whale-watching vessels. The dataset included almost 100 predation events. Publishing their results in the journal PLOS ONE, Mr McInnes found patterns in behaviour of outer-coast transients that differed from those of their inner-coast cousins. The latter ambushed harbour seals, sea lions and the occasional otter, while the former tended to perform co-ordinated dives in their hunt for larger mammalian prey.

| *(Cetacean needed)*

Outer-coast transients appear to have developed specialised strategies—such as the repeated battering and drowning of grey whale calves—because their prey is bigger and takes more time to subdue. Like most specialised killer-whale hunting behaviours, these strategies are thought to be honed, down the generations, as a form of cultural knowledge.

While studying the outer-coast transients, Mr McInnes also stumbled upon hints of a previously undescribed population of killer whales that were big-

game hunters of an entirely different class. Writing in the journal *Aquatic Mammals*, his team described nine offshore encounters with 49 killer whales between 1997 and 2021. These killer whales were seen killing sperm whales, pygmy sperm whales and large Risso's dolphins, as well as scavenging on leatherback turtles.

The attack on sperm whales, which can grow to around ten times larger than adult killer whales, was a particularly impressive show of predatory force. On October 21st 1997, between 15 and 35 killer whales were observed hunting a herd of nine sperm whales over the course of four hours. The sperm whales were subjected to repeated waves of aggressive attack, injuring all of them and severely injuring several before one sperm whale was finally isolated from the herd and killed. Out of the surviving sperm whales, three are thought to have died from the wounds they sustained and it is possible the entire herd later succumbed to its injuries.

Little more than that is known about these killer whales for now, except that they are identifiable by characteristic circular scars left by bites from cookie-cutter sharks, which suggest these hyper-aggressive creatures stalk the deep waters of the north Pacific.

More data, especially of the genetic kind, will need to be collected from both outer-coast transient killer whales as well as their newly discovered cousins terrorising the pelagic waters of the north Pacific, before scientists can fully describe the new ecotypes and add them to marine-biology textbooks. Nevertheless, the discoveries are a reminder that biodiversity is not just about the diversity between species but also within them. Killer whales seem to exist in many more shades than just black and white. ■



杀手本色

虎鲸，协同作战的冷酷猎手

它们的技艺代代相传

在加利福尼亚沿海的蒙特雷海底峡谷（Monterey Submarine Canyon）水域游荡的虎鲸（俗称杀手鲸）是凶残的动物。它们捕猎灰鲸时，会盯着成对的母鲸和幼鲸，一路追逐直至幼鲸精疲力竭。这时，虎鲸会在两者之间机动穿梭，或者拖拉幼鲸的尾巴或鳍把它拖走。一旦幼鲸与母鲸分开，虎鲸就会用头部和尾巴猛击幼鲸，使其失去行动能力，然后跃到它的背上压住喷水孔，使其不能浮上海面而溺死。

虎鲸是一个多样化的物种，在世界各地有多个孤立的种群，占据着不同的生态位——其中一些至今仍然鲜有记载。最近两项研究探究了在太平洋东北部深海中潜行的虎鲸，揭示了它们独特的狩猎文化。

不同的虎鲸种群被称为“生态型”，体形、颜色和身型比例可能都有差异。即使在不同生态型的生存范围交集的地方，它们之间也很少互动，似乎也不会杂交。虎鲸各种独特生态型的差异主要在于文化，即虎鲸在其社会活动中习得并代代相传的行为。这些行为包括对不同海洋猎物的偏好，以及寻觅、捕猎和杀死猎物的协作策略，可能历经数千年的磨练才形成。不同的虎鲸生态型甚至使用不同的“方言”交流，包含了各种咔哒声、口哨声和呼叫声，不同鲸群还有不同的“口音”。

被研究最多的虎鲸生态型有三种，它们生活在从阿拉斯加北部到加利福尼亚南部的东北太平洋海岸线。其中“离岸型”体长可达6.6米，以100至200头为一群在远洋游弋，捕食鲨鱼等浮泳鱼。“定居型”体长约7米，每群可达25头，在近海出没，捕食近海鱼类（主要是鲑鱼）。最后是“过境型”，体长达8米，在浅海和深海水域均有出现，捕食海洋哺乳动物，包括鲸目动物（鲸和海豚）及鳍足目动物（海狮和海豹）。“内海岸”过境型虎鲸在海岸附近的浅水区出没，而“外海岸”过境型虎鲸则在大陆架边缘的峡谷附近的开阔深水区捕猎。

为了更好地了解这些过境型虎鲸，加拿大不列颠哥伦比亚大学的乔什·麦金尼斯（Josh McInnes）带领的团队分析了通过调查和观鲸船收集的在蒙特雷海底峡谷系统周围的261次虎鲸目击事件。这个数据集包括近100次捕食事件。麦金尼斯在《PLOS ONE》期刊上发表了他们的研究结果，发现外海岸过境型虎鲸的行为模式与内海岸型不同。后者伏击斑海豹、海狮，偶尔也捕食水獭，而前者倾向于协同作战，捕猎更大的哺乳动物。

｜ 鲸吞虎噬

外海岸过境型虎鲸似乎已经发展出专门的捕猎战略——例如反复击打和溺死灰鲸幼崽——因为它们的猎物更大，需要更长时间才能制服。与大多数各类虎鲸特有的狩猎行为一样，一般认为这些战略在世代相传中不断精进，好像一种文化知识。

在研究外海岸过境型虎鲸时，麦金尼斯还偶然发现了一个从未被描述过的虎鲸种群的踪迹，它们是一类完全不同的大型猎物猎手。他的团队在《水生哺乳动物》（Aquatic Mammals）期刊上发表文章，描述了1997年至2021年间与共计49头虎鲸的九次海上遭遇。人们看到这些虎鲸捕杀抹香鲸、小抹香鲸和大型里索海豚，还看到它们掠食棱皮龟。

抹香鲸的个头可以长到成年虎鲸的十倍左右，虎鲸对它们的攻击展示了令人震撼的强大猎杀实力。1997年10月21日，有人观察到15至35头虎鲸在四个小时内对有九头抹香鲸的鲸群展开捕猎。抹香鲸遭到一轮又一轮的凶猛攻击，全部受伤，其中几头伤势严重，最终一头抹香鲸落单并被杀死。在幸存的抹香鲸中，据认为有三头死于伤势过重，有可能后来整个鲸群全部因伤死亡。

除此以外，目前人们对这些虎鲸知之甚少，只知道它们身上有被雪茄达摩鲨咬伤后留下的圆形疤痕，这表明这些攻击性极强的生物在北太平洋的深海中出没。

科学家若要全面描述这些新的生态型并将它们写入海洋生物学的教科书，还需要从外海岸过境型虎鲸以及新发现的、在北太平洋中上层水域横行的

近亲那里收集更多数据，尤其是基因数据。但无论如何，这些发现都让我们认识到，生物多样性不仅仅存在于物种之间，也出现在物种之中。虎鲸似乎不仅仅只有黑白两色那么简单。 ■



Buttonwood

How the “Magnificent Seven” misleads

Forget the supergroup of stockmarket darlings

ALL MODELS are wrong, goes the statisticians’ adage, but some are useful. This time last year, plenty of pundits’ models started looking more wrong and less useful. The consensus forecast was a grim spell for economic growth and a dreary one for stock prices—and that was before a clutch of American regional banks buckled. Higher interest rates seemed set to cause pain everywhere. Instead, in the very country where the banking turmoil unfolded, the stockmarket began to soar. By the summer America’s S&P 500 index of leading shares had risen by 28% from a trough hit the previous autumn. Analysts hunted for a new model to explain what was going on, and the popular choice revolved around the “Magnificent Seven”.

The reason was that shares in this group of tech giants—Alphabet, Amazon, Apple, Meta, Microsoft, Nvidia and Tesla—were acting as if they made up a market of their own. By the start of June the S&P 500 had risen by 12% since the beginning of the year, but virtually the entire gain was down to these seven stocks, with the other 493 having collectively moved sideways. By July they were the biggest seven companies in the index. By late October the Magnificent Seven had added \$3.4trn (or 50%) to their combined market value since the start of the year, even as the other 493 had lost \$1trn (3.8%). Apple was the group’s worst performer but had nevertheless seen its share price rise by 30%.

Alas, the band has broken up. One feature of the storming bull run that began on October 27th is that the Magnificent Seven have stopped charging as a pack. As the broader S&P 500 has risen by 26%, three of the group have turned from leaders to laggards: the share prices of Alphabet and Apple have underperformed, while that of Tesla has fallen outright. The

remaining four have beaten the index, but by wildly differing amounts. Whereas Microsoft has done only slightly better than the average S&P 500 member, Nvidia's shares have rocketed by an astonishing 129%. Six of the seven are still the index's largest stocks, but Tesla has been replaced by Berkshire Hathaway, Warren Buffett's investment vehicle.

What happened? Last spring the narrative was obvious: the Magnificent Seven would end up as the winners from the artificial-intelligence (AI) gold rush. All were enormous to begin with, which would help them weather any downturn and outspend rivals on research and development. Each had already conquered its respective niche of the tech world, meaning they could probably carry on outsmarting other firms. And their big cash balances stopped rising interest rates from looking too threatening.

In part this seductive story has fallen apart because, like all models, it was oversimplified to begin with. The idea of the Magnificent Seven as "big tech" incarnate—or at least of "big tech" as a homogenous group—was always something of a stretch. Aside from trillion-dollar valuations and hordes of clever employees, a manufacturer of AI chips (Nvidia) does not have an awful lot in common with an online retailer (Amazon) or a purveyor of social networks (Meta). Nor does a software firm born in 1975 (Microsoft) necessarily share the prospects of an electric-vehicle maker established in 2003 (Tesla). In this, the Magnificent Seven concept was not alone. Think of the arbitrariness of the "BRICS" gang of countries (Brazil, Russia, India, China and South Africa), or of the even bigger and more amorphous "emerging markets" group.

Yet the disbanding of the Magnificent Seven signals some broader shifts, too. One is that, as last year's bank failures fade from memory, investors have become less anxious about the impact of higher interest rates. America's economy has defied recession forecasts and boomed instead. Firms that are insulated from shocks by big market shares, profit margins

and cash buffers have thus lost some of their “safe haven” appeal.

It also shows that, though optimism is surging through markets, not all caution has been thrown to the winds. A year ago generative AI so enthralled investors that tech giants were assumed to be winners by default. Now that automatic benefit of the doubt is being withdrawn. Nvidia has pulled away from the rest of the group as it is transforming hype into stratospheric sales and profits. Apple and Tesla are being punished, in part, for failing to communicate how they will do the same. Investors might be excited about AI but—encouragingly, for those who fear a bubble—they seem willing to be sceptical, too. ■



梧桐

“美股七巨头”的误导性

忘掉股市的这群超级宠儿吧

统计学家的格言是，所有的模型都是错误的，但有些是有用的。去年此时，许多权威专家的模型开始显得越发错误，而更没用处。各种预测的共识是未来一段时间经济增长前景黯淡，股票价格低迷——此时一批美国区域性银行的危机尚未爆发。高利率似乎很快将会让痛苦无处不在。结果，正是在美国这个银行业发生动荡的国家，股市开始飙升。到去年夏天，涵盖美国主要股票的标准普尔500指数已经从前年秋天的低谷上涨了28%。分析师们一直在寻找一种新模型来解释这一现象，最流行的一个是围绕“美股七巨头”展开的。

原因是这一组科技巨头——Alphabet、亚马逊、苹果、Meta、微软、英伟达和特斯拉——的股票表现仿佛自成一个市场。到去年6月初，标普500指数比年初上涨了12%，但实际上全部涨幅都来自这七只股票，其他493只股票则集体横盘整理。截至去年7月，它们已成为标普500指数里最大的七家公司。到10月底，“七巨头”的总市值比去年年初增加了3.4万亿美元（涨幅为50%），而其他493家公司损失了1万亿美元（3.8%）。苹果是这个群组中表现最差的，但股价也上涨了30%。

可惜啊，这支队伍已经解散了。从去年10月27日开始的一轮强劲牛市有个特点，就是七巨头已经不再集体上冲。随着标普500指数上涨26%，其中三家公司在股市上已经从领头羊变成了落后者：Alphabet和苹果的股价表现弱于大盘，特斯拉的股价则直线下跌。其余四家公司都跑赢了指数，但幅度大不相同。微软的表现仅略好于标普500成分股的平均水平，而英伟达的股价却飙升了惊人的129%。在这七只股票中，仍有六只是该指数的最大成分股，但特斯拉已被巴菲特的投资公司伯克希尔·哈撒韦取代。

到底发生了什么？去年春天的故事线很明显：“七巨头”最终会成为人工智能淘金热的赢家。它们本就体量庞大，这能帮助它们抵御任何经济衰退，

并且能在研发开支上压过竞争对手。每家公司都已在科技业中占领了各自的利基市场，这意味着它们接下来可能也会比其他公司智高一筹。而且，它们的巨额现金余额让利率上升看起来不那么凶险。

在某种程度上，这个诱人的故事已经讲不下去了，因为像所有模型一样，它一开始就过于简化了。将“七巨头”视为“科技巨头”的化身——或者至少是将“科技巨头”视为一个同质群体——从来都有些牵强。除了都有数万亿美元的估值和大批聪明的员工外，AI芯片制造商（英伟达）和在线零售商（亚马逊）或社交网络提供商（Meta）并没有太多共同之处。成立于1975年的软件公司（微软）也未必会和成立于2003年的电动汽车制造商（特斯拉）有同样的前景。在这方面，“七巨头”的概念不是孤例。想想“金砖国家”（巴西、俄罗斯、印度、中国和南非），还有更宽泛、更不固定的“新兴市场”集团，这些定义更加随意。

不过，“七巨头”的解散也标志着一些更大范围的转变。其一是，随着去年的银行倒闭风波逐渐从人们的记忆中消失，投资者已不再那么担心利率上调的影响。美国经济并没有如预测般衰退，反而再现繁荣。那些因市场份额大、利润空间大、有现金缓冲而免受冲击的公司因此失去了一些作为“避风港”的吸引力。

它也表明，尽管目前市场上乐观情绪高涨，但并非所有的谨慎都被抛到了九霄云外。一年前，投资者对生成式AI如痴如醉，科技巨头都被默认为赢家。而现在，这种不假思索的信任正在消退。英伟达正在将市场炒作转化为惊人的销售和利润，和“七巨头”中的其他公司拉开了距离。苹果和特斯拉之所以受到惩罚，部分原因是它们没有向市场传达好自己将如何取得同样的佳绩。投资者也许对AI兴奋不已，但让担心泡沫的人们欣慰的是，他们似乎也愿意保留一些怀疑。■



Health technology

The AI doctor will see you...eventually

Artificial intelligence holds huge promise in health care. But it also faces massive barriers

BETTER DIAGNOSES. Personalised support for patients. Faster drug discovery. Greater efficiency. Artificial intelligence (AI) is generating excitement and hyperbole everywhere, but in the field of health care it has the potential to be transformational. In Europe analysts predict that deploying AI could save hundreds of thousands of lives each year; in America, they say, it could also save money, shaving \$200bn-360bn from overall annual medical spending, now \$4.5trn a year (or 17% of GDP). From smart stethoscopes and robot surgeons to the analysis of large data sets or the ability to chat to a medical AI with a human face, opportunities abound.

There is already evidence that AI systems can enhance diagnostic accuracy and disease tracking, improve the prediction of patients' outcomes and suggest better treatments. It can also boost efficiency in hospitals and surgeries by taking on tasks such as medical transcription and monitoring patients, and by streamlining administration. It may already be speeding the time it takes for new drugs to reach clinical trials. New tools, including generative AI, could supercharge these abilities. Yet as our Technology Quarterly this week shows, although AI has been used in health care for many years, integration has been slow and the results have often been mediocre.

There are good and bad reasons for this. The good reasons are that health care demands high evidentiary barriers when introducing new tools, to protect patients' safety. The bad reasons involve data, regulation and incentives. Overcoming them could hold lessons for AI in other fields.

AI systems learn by processing huge volumes of data, something health-care providers have in abundance. But health data is highly fragmented; strict rules control its use. Governments recognise that patients want their medical privacy protected. But patients also want better and more personalised care. Each year roughly 800,000 Americans suffer from poor medical decision-making.

Improving accuracy and reducing bias in AI tools requires them to be trained on large data sets that reflect patients' full diversity. Finding secure ways to allow health data to move more freely would help. But it could benefit patients, too: they should be given the right to access their own records in a portable, digital format. Consumer-health firms are already making use of data from wearables, with varying success. Portable patients' records would let people make fuller use of their data and take more responsibility for their health.

Another problem is managing and regulating these innovations. In many countries the governance of AI in health, as in other areas, is struggling to keep up with the rapid pace of innovation. Regulatory authorities may be slow to approve new AI tools or may lack capacity and expertise.

Governments need to equip regulators to assess new AI tools. They also need to fill regulatory gaps in the surveillance of adverse events, and in the continuous monitoring of algorithms to ensure they remain accurate, safe, effective and transparent.

That will be hard. One solution would be for countries to work together, to learn from each other and create minimum global standards. A less complex international regulatory system would also help create a market in which small companies can innovate. Poorer countries, with less developed health infrastructure, have much to gain from introducing new tools, such as an AI-powered portable ultrasound device for obstetrics. Because the alternative to an AI tool is often no treatment at all, they may

even be able to leapfrog the entrenched health systems of rich countries—though a lack of data, connectivity and computing power will get in the way.

A final problem involves institutions and incentives. AI promises to cut medical costs by assisting or replacing workers, improving productivity, reducing errors and flattening or reducing spending, all while improving care. That is desperately needed. The world could lack 10m health-care workers by 2030, around 15% of today's workforce. And administration accounted for about 30% of America's excess health-care costs, compared with other countries, in 2022.

Yet saving money using innovation is tricky. Health systems are set up to use it to improve care, not cut costs. New technology may account for as much as half of the annual growth in health spending. Layering on new systems will increase costs and complexity. But redesigning processes to make efficient use of AI is likely to be resisted by patients and medics. Though AI may be able to triage them over the phone or provide routine results, patients may demand to be seen in person.

Worse, many health systems, such as America's, are set up to reward the volume of work. They have little reason to adopt technologies that cut the number of visits, tests or procedures. And even publicly run health-care systems may lack incentives to adopt technologies that reduce costs rather than improve outcomes, perhaps because saving money may lead to a smaller budget next year. Unless governments can change these incentives, so that AI combines better treatment with new efficiencies, innovation will increase costs. Accordingly, governments and health authorities will need to fund schemes dedicated to testing and deploying new AI technologies. Countries including America, Britain and Canada are pointing the way.

Much of the burden for boosting AI in health care falls on governments and regulators. However, companies have a part to play, too. Insurers have already used AI tools to deny care unfairly; firms have mis-sold or overstated the abilities of health AI; algorithms have made mistakes. Firms have a duty to ensure that their products are safe, reliable and accountable, and that humans, however flawed, remain in control.

These obstacles are formidable but the potential benefits of using AI in health care are so vast that the case for overcoming them should be obvious. And if AI can be made to work in medicine, it could provide a prescription for the adoption of the technology in other fields. ■



【首文】健康科技

终有一天，会由AI医生给您看病

人工智能在医疗保健领域前景广阔，但也面临巨大障碍

更精准的诊断。个性化的患者支持。更快的药物发现。更高的效率。AI在各个领域都在引发兴奋和炒作，但在医疗保健领域，它确实有可能带来变革性的影响。在欧洲，分析师预测部署AI每年可以挽救数十万人的生命；在美国，他们认为AI还能节省资金，从目前每年4.5万亿美元（占GDP的17%）的总医疗支出中省下2000亿至3600亿美元。从智能听诊器和机器人外科医生，到大数据集分析或者与有一副人类面孔的医疗AI交流，机遇无处不在。

已有证据表明，AI系统可以提高诊断准确性、方便疾病跟踪，改进预后、并建议更好的疗法。它还可以承担转录医生口述内容和监测患者等任务，简化行政流程，从而提高医院和诊所的效率。它可能已经加快了新药物进入临床试验阶段的速度。包括生成式AI在内的新工具可以进一步增强这些能力。然而，尽管AI在医疗保健领域已经运用多年，但系统整合的进展缓慢，结果往往也平平无奇。

这其中的原因有好有坏。好的原因是，为保障患者安全，在引入新的工具时，医疗行业对证据门槛的要求很高。不好的原因涉及数据、监管和激励机制。克服这些问题可能会为AI在其他领域的应用提供启示。

AI系统通过处理大量数据来学习，而医疗机构最不缺的就是数据。但这些数据高度分散，且对它们的使用有严格规定。政府认识到患者想要保护自己的医疗隐私。但他们也希望能获得更好和更个性化的护理。每年大约有80万美国人因医疗决策不当而遭受痛苦。

要提高AI工具的准确性和降低其偏误，需要用能全面体现患者多样性的大型数据集来训练它们。找到安全的途径来让医疗数据更自由地流动会有帮助。而这也可能让患者受益。他们应该有权访问自己的便携式数字病历。

消费者健康公司已经开始利用可穿戴设备的数据，取得了不同程度的成功。便携式病历将能让人们更充分地利用自己的数据，并对自己的健康承担起更多的责任。

另一个问题是对这些创新的管理和监管。在许多国家，对医疗部门AI的治理和在其他领域一样难以跟上快节奏的创新。监管机构审批新的AI工具可能需要很长时间，或者可能缺乏相关能力和专业知识。政府需要帮助监管机构提升评估新AI工具的能力，还需要填补监管空白，包括监测不良事件，以及持续监测算法以确保它们保持准确、安全、有效和透明。

这并不容易做到。一个解决方案是各国携手合作，相互学习借鉴，并制定全球最低标准。一个不那么复杂的国际监管体系也将有助于创建一个让小公司也能创新的市场。卫生基础设施欠发达的较贫穷国家将从引入新工具（例如用于产科的便携式AI超声设备）中显著受益。由于没有AI工具往往就意味着不做任何治疗，较贫穷国家甚至可能不必再建立富裕国家那种难以改变的卫生系统。不过，数据、连接性和计算能力的缺乏会成为障碍。

最后一个问题涉及制度和激励机制。AI有望通过协助或取代人力、提高生产率、减少错误、稳定或降低支出等途径来削减医疗成本，同时改善医疗水平。这是社会所迫切需要的。到2030年，全球医务人员的缺口可能会达到1000万，约为现有人员总数的15%。与其他国家相比，在2022年，行政管理占美国超额医疗成本的约30%。

然而，利用创新来节省开支并非易事。设立医疗保健系统的目的是提高医疗水平，而非降低成本。新技术可能占了卫生支出年度增长的一半之多。叠加新系统将增加成本和复杂性。但如果为了有效利用AI而重新设计流程很可能会遭到患者和医务人员的抵制。虽然AI或许能在电话上完成患者预检分诊或提供常规检查结果，但患者可能会要求线下问诊。

更糟的是，美国等许多国家的医疗系统都设计成按工作量计酬，几乎没有理由采用能够减少就诊、检查或治疗的技术。即使是公立医疗部门也可能

缺乏动力去采用会降低成本而非改善疗效的技术，这也许是因为节省资金可能会导致下一年的预算缩水。除非政府能够调整这类激励机制，让AI既能改善疗效又能提高效率，否则创新将徒增成本。因此，政府和医疗主管部门需要资助专门测试和部署新AI技术的项目。美国、英国和加拿大等国正在指明道路。

| AI，医学博士

在推动医疗保健领域的AI发展方面，政府和监管机构要负起很大一部分责任。但企业也将发挥作用。保险公司已经在利用AI工具有失公允地拒赔医疗费用；一些公司误导或夸大了医疗AI产品的能力；算法也出现过错误。企业有责任确保其产品的安全、可靠及可追责，并且应确保仍然由人类把关，不论人类有多不完美。

这些挑战是艰巨的，但在医疗保健领域利用AI的潜在好处巨大，有显而易见的理由去克服它们。而如果AI能够在医学领域发挥效力，也许也能为它在其他领域的应用提供一剂良方。■



The lost art of self-management

Have McKinsey and its consulting rivals got too big?

The golden age for CEO whisperers may be coming to an end

AN ANONYMOUS MEMO briefly circled the web in March. The authors, who claimed to be former partners at McKinsey, rebuked the illustrious strategy consultancy for its pursuit in recent years of “unchecked and unmanaged growth”, and chastised its leadership for, of all things, a “lack of strategic focus”. With humility typical of McKinseyites, they warned that “an organisation of genuine greatness” was at risk of being lost.

The memo, which was swiftly taken down, is but the latest murmur of discontent at McKinsey. In January Bob Sternfels, its managing partner, was forced into an internal contest for the top job after he failed to clinch support for re-election from a majority of senior partners in an initial round of voting. Although he ultimately prevailed, the saga hinted at the trouble brewing within the firm.

Not long ago the consulting industry looked indestructible. Fees rocketed during the covid-19 pandemic as clients sped up efforts to digitise their businesses, diversify their supply chains and respond to growing calls to bolster their environmental, social and governance (ESG) credentials. The consulting revenues of the industry’s most important firms—including the triumvirate of strategy advisers (Bain, BCG and McKinsey), the “big four” accounting giants (Deloitte, EY, KPMG and PwC) and Accenture (also the world’s largest outsourcer)—grew by 20% in 2021 and then 13% in 2022 (see chart).

Since then, however, growth has been soggy for the industry’s “great eight”, slowing to around 5% in 2023, according to estimates by Kennedy Research Reports, an industry-watcher, and calculations by The Economist, based on

company filings. Clients grappling with inflation and economic uncertainty have cut back on splashy consulting projects. A dearth of mergers and acquisitions has led to a slump in demand for support with due diligence and company integrations.

That has caused a headache for the consultancies. When demand from clients looked limitless, they recruited staff as if there were no tomorrow. Revenues at McKinsey are up by a third since 2019—but headcount is up by half, to 45,000. As job opportunities at startups and private-equity firms have withered, fewer consultants have left the firms of their own accord, reversing the spike in attrition rates during the pandemic.

Now tomorrow is here, with a vengeance. Bain and Deloitte have paid some graduates to delay their start dates. Newbie consultants at a number of firms complain that there is too little work to go around, stunting their career prospects. Lay-offs, typically rare in consulting, have become widespread. All of the big four have made cuts to their advisory teams. Last year Accenture, the only one of the eight that is publicly listed, said it would fire 19,000 staff. On March 21st the company reported that its consulting revenues for the quarter to February shrank by 3%, year on year, after flattening in the previous quarter. It also lowered its growth forecast for the year ahead, sending its share price down by 9%.

The consulting industry has made it through choppy waters before, including the dotcom crash and the global financial crisis. Yet its recovery this time will be complicated by three things. The first is geopolitics. The consulting giants, all of which are based in America or Europe, have benefited from decades of globalisation during which they spread their tendrils into every part of the globe. Deloitte, the biggest of the bunch by consulting revenues, has offices in more than 150 countries and territories.

That is now placing the firms in awkward spots. Last month it emerged that

the Urban China Initiative, a think-tank co-founded by McKinsey, provided advice to the Chinese government in 2015 that helped shape its “Made in China 2025” plan, which has sought to reduce the economy’s reliance on foreign know-how and place China at the forefront of sectors from electric vehicles to artificial intelligence (AI). Although McKinsey denied that it wrote the report, some American lawmakers have called for the firm to be barred from American government contracts. In the 12 months to September 2023 the federal government paid McKinsey more than \$100m in fees.

Now China is also starting to squeeze foreign advisers of various stripes out of its market. Last year Dentons, a global law firm, unwound its tie-up with Dacheng, a Chinese one, in response to new cyber-security and data-protection rules that made the combination unworkable. Although China is yet to produce a home-grown consulting powerhouse, it has already begun to make life difficult for foreign ones. Staff in Bain’s Shanghai office were questioned by Chinese authorities last year, for reasons unknown. On March 22nd it was reported that the Chinese government was scrutinising PwC’s auditing work at Evergrande, a bankrupt Chinese property developer that has been accused by China’s government of fraudulently inflating its revenues. That could weigh on PwC’s consulting business in the country.

It is not only the West’s relationship with China that is causing problems. In February the bosses of BCG, McKinsey and Teneo, a smaller consultancy, along with Michael Klein, a dealmaker, were hauled before a congressional committee in Washington after failing to hand over details of their work for Saudi Arabia’s Public Investment Fund. The committee is investigating Saudi Arabia’s efforts to build “soft power” in America through, for example, its investments in sports such as golf. McKinsey and BCG said that their staff in Saudi Arabia could be imprisoned if the firms divulged what they had done for their client. The Gulf has been a rare bright spot for the consultants of late, with the oil-rich states splashing out on advice as they

seek to diversify their economies.

Waning enthusiasm for ESG, denounced by critics as “woke capitalism”, presents a second threat to the industry’s recovery. In recent years the consulting giants have spent big on building out their ESG offerings, especially around decarbonisation. In 2021 McKinsey acquired three sustainability consultancies. In 2022 Accenture gobbled up five. So far these investments seem to be paying off. Christoph Schweizer, the boss of BCG, which acquired Quantis, an environmental consultancy, in 2022, says that sustainability was one of his firm’s fastest-growing areas of work last year.

Whether that growth will continue at the same pace is less clear. In America Republican-run states including Florida, Missouri and Texas have withdrawn funds parked with BlackRock, the world’s largest asset manager, in protest over its use of ESG considerations when making investments. Consulting clients surveyed in January by Source Global Research, another industry analyst, ranked sustainability projects tenth in their list of priorities for the year, down from fourth in 2023. Some consulting grandees admit that certain clients are reining in their climate ambitions. Stretched consumers are now less able to pay a green premium, suggests one bigwig.

The third and thorniest challenge for the great eight is technological change. For many years clients sought their help to modernise creaky old systems. Increasingly, the consultants themselves are grappling with digital disruption. The boss of a big buy-out firm says that his dealmakers are turning to software tools and data providers rather than pricey consultants for some of the analysis needed to evaluate a target company. Other tasks that legions of consultants once spent hours on, such as compiling and categorising data on a company’s spending habits, can now be done at the push of a button.

The consultants are not standing idle. Bain, for example, has redesigned the way it does due diligence on companies, incorporating nifty tools such as web-scraping programs. The firms are also racing to stay one step ahead of AI. Last August McKinsey launched Lilli, a ChatGPT-like bot trained on its corpus of frameworks and other intellectual property, which consultants can use to speed up their work. Others have followed suit. Manny Maceda, the managing partner of Bain, reckons these bots will free up time for consultants to better understand the “organisational realities” of their clients.

Excitement among clients over this type of “generative” AI is also creating opportunities for new work. Mr Schweizer says that BCG has already completed hundreds of projects with clients around the technology. Accenture has booked \$1.1bn-worth of generative-AI work in the past six months. Much of this is happening in collaboration with the tech companies developing the AIs. Accenture has been working with Microsoft. In March the consulting firm also announced a partnership with Cohere, an AI-model builder with which McKinsey has buddied up, too. Bain has an alliance with OpenAI, the maker of ChatGPT. BCG is collaborating with Anthropic, one more AI firm.

Such partnerships look like a welcome source of growth for the consultants. In time, though, they could become a drag—especially if they are successful. The quicker corporate clients become comfortable with chatbots, the faster they may simply go directly to their makers in Silicon Valley. If that happens, the great eight’s short-term gains from AI could lead them towards irrelevance. That is something for all the strategy brains to stew on. ■



失传的自我管理技艺

麦肯锡等咨询巨头是不是变得过于庞大？

CEO密友的黄金时代可能即将结束【深度】

今年3月，网上一度流传着一份匿名备忘录。作者自称是麦肯锡的前合伙人，他们斥责这家著名的战略咨询公司近年来追求“不受约束和监管的增长”，并出人意料地谴责其领导层“没有战略重点”。以麦肯锡人特有的谦逊，他们警告称“一个名副其实的伟大机构”正面临迷失的风险。

这份很快被撤下的备忘录不过是最新一次爆出对麦肯锡的不满声音。今年1月，麦肯锡的管理合伙人鲍勃·斯特恩费尔斯（Bob Sternfels）在竞选连任的首轮投票中未能获得多数高级合伙人的支持，不得不加入对这个最高职位的内部竞争。尽管他最终获胜，但这番波折也透露出这家公司内部的麻烦正在滋长。

就在不久前，咨询行业似乎还坚不可摧。新冠疫情期间，随着其客户加速推动业务数字化和供应链多元化，并响应日益高涨的呼声来提高环境、社会和治理（ESG）得分，咨询行业的收费一飞冲天。业内最举足轻重的八家公司——包括战略咨询三巨头（贝恩、波士顿咨询和麦肯锡）、“四大”会计师事务所（德勤、安永、毕马威和普华永道）以及埃森哲（也是全球最大的外包服务公司）——的咨询收入在2021年增长了20%，2022年增长了13%（见图表）。

然而，根据行业观察机构肯尼迪研究报告（Kennedy Research Reports）的估计，以及本刊基于公司文件的推算，自那以后，咨询行业这“八大”公司的增长就陷入疲软，在2023年放缓至5%左右。奋力应对高通胀和经济不确定性的客户削减了大手笔的咨询项目。而并购项目不足则导致对尽职调查和公司整合等支持性服务的需求大幅下滑。

这让咨询公司头疼不已。过去，当客户的需求看起来无穷无尽时，它们不顾后果地疯狂招人。自2019年以来，麦肯锡的收入增长了三分之一，员工

人数却增加了一半，达到4.5万人。随着创业公司和私募股权公司工作岗位的减少，自愿离职的咨询顾问也变少了，疫情期间员工流失率飙升的趋势转向了。

现在后果就摆在眼前，而且相当苦涩。贝恩和德勤向一些已被录用的毕业生提供补贴，让他们推迟入职。好几家公司的新手顾问抱怨无事可干，阻碍了自身的职业发展。咨询行业过去鲜有发生的裁员已变得十分普遍。四大会计师事务所都削减了各自的咨询团队。去年，“八大”中唯一的上市公司埃森哲表示将解雇1.9万名员工。3月21日，埃森哲发布报告称，其截至2月的季度咨询业务收入继上一季度同比持平后，进一步下降了3%。此外，埃森哲还下调了未来一年的增长预期，导致其股价下跌9%。

咨询行业曾经熬过了包括互联网泡沫破裂和全球金融危机在内的沟沟坎坎。然而，这一次它的复苏会因三件事情而变得复杂。首先是地缘局势。这些咨询业巨头的总部都设在美国或欧洲，几十年的全球化进程让它们获益颇多，其间它们将触角伸到全球各地。其中最大（按咨询业务收入计算）的德勤在150多个国家和地区设有办事处。

然而成也萧何败也萧何，如今这些公司也因此陷入了尴尬的境地。2月有消息称，由麦肯锡联合创办的智库“城市中国计划”在2015年向中国政府建言献策，帮助打造了“中国制造2025”计划。该计划旨在减少中国经济对外国技术的依赖，并实现中国在电动汽车和人工智能（AI）等领域的领先地位。尽管麦肯锡否认自己撰写了研究报告，但一些美国国会议员呼吁禁止麦肯锡获得美国政府合同。在截至2023年9月的12个月中，美国联邦政府向麦肯锡支付了逾一亿美元的服务费。

如今中国也开始将形形色色的外国咨询公司挤出本土市场。去年，全球知名的律师事务所Dentons与中国大成律师事务所分道扬镳，原因是新出台的有关网络安全和数据保护的条例让双方的结合实际上不再可行。尽管中国尚未打造出本土的咨询业巨头，但它已经开始让外国咨询公司的日子不好过了。去年，贝恩上海办事处的员工受到中国当局原因不明的问询。3月22日，有报道称中国政府正在审查普华永道对恒大的审计工作。中国政

府指控这家陷入破产的房地产开发商欺诈性虚报收入。这可能会影响普华永道在中国的咨询业务。

制造问题的不仅仅有中西方关系的变化。2月，美国国会的一个委员会传唤了波士顿咨询、麦肯锡和Teneo（一家规模较小的咨询公司）的老板以及交易撮合者迈克尔·克莱因（Michael Klein），原因是他们没有交出为沙特阿拉伯的公共投资基金（Public Investment Fund）工作的详细资料。沙特试图通过投资高尔夫等体育运动以及其他途径在美国建立“软实力”，该委员会正对这些做法展开调查。麦肯锡和波士顿咨询表示，如果公司泄露为这家客户所做的工作，它们在沙特的员工可能遭到监禁。由于这些产油大国在寻求经济多元化的进程中不惜重金购买咨询服务，海湾地区近年来成为咨询公司难得的亮点。

人们对ESG（它被批评者斥为“觉醒资本主义”）的热情减退，构成了该行业复苏的第二个威胁。近年来，这些咨询巨头在打造ESG服务上投入大量资金，尤其是与碳减排相关的服务。2021年，麦肯锡收购了三家可持续发展咨询公司。2022年，埃森哲吞并了五家。截至目前，这些投资似乎在收获回报。于2022年收购了环保咨询公司Quantis的波士顿咨询的老板施伟策（Christoph Schweizer）表示，可持续发展是公司去年增长最快的业务之一。

目前还不太清楚这种增长势头是否能保持下去。在美国，佛罗里达、密苏里和得克萨斯等一批共和党主政的州已经撤回了存放在全球最大资产管理公司贝莱德（BlackRock）的资金，以抗议其在投资时考虑ESG因素。另一家行业分析机构Source Global Research1月对咨询业客户企业的调查显示，它们将可持续发展项目排在本年度优先事项的第十位，而2023年时是第四位。咨询业内的一些重量级人物承认，某些客户正在收敛它们的气候雄心。一位大佬表示，如今捉襟见肘的消费者在为绿色溢价买单时没有过去那么大方了。

“八大”面临的第三个也是最棘手的挑战是技术变革。多年来，客户向它们寻求帮助，以更新老旧的系统。如今这些咨询公司自己也越来越多地面对

数字颠覆。一家大型并购投资公司的老板表示，其交易人员在获取评估目标公司所需的一些分析时，正转向软件工具和数据供应商，而不是收费高昂的咨询公司。其他一些过去需要一群咨询顾问花很多个小时才能完成的工作，比如汇编和分类有关公司开支习惯的数据，现在只需按个按钮就能搞定。

咨询公司也没闲着。例如，贝恩重新设计了对公司进行尽职调查的方式，将网络抓取程序等先进工具纳入其中。这些公司也大脚踩油门要先AI一步。去年8月，麦肯锡发布了一款类似于ChatGPT的软件工具Lilli，训练它的语料库包括麦肯锡的思考框架和其他知识产权，咨询顾问可以用它来提高工作效率。其他公司也纷纷效仿。贝恩的管理合伙人曼尼·马塞达（Manny Maceda）认为，这些机器人工具能让咨询顾问有时间更好地了解客户的“组织现实”。

客户对这类“生成式”AI的热情也给咨询公司带来了新的业务。施伟策表示，波士顿咨询已与客户一起完成了数百个围绕生成式AI的项目。在过去六个月里，埃森哲签订了价值11亿美元的生成式AI项目。其中大部分都是与开发AI的科技公司合作完成的。埃森哲一直在与微软合作。3月，它还宣布与AI模型构建公司Cohere（麦肯锡也是其合作伙伴）携手。贝恩与开发ChatGPT的OpenAI结盟。波士顿咨询正在与另一家AI公司Anthropic合作。

对于咨询公司来说，这种合作关系似乎是一个可喜的增长源。不过，它们最终可能成为拖累——尤其是在它们卓有成效的情况下。企业客户越快与聊天机器人混熟，就可能越快直接去找这些机器人在硅谷的开发商。如果真是那样，八大从AI那里获得的短期收益可能会让它们自己变得无足轻重。这一点值得所有战略顾问深思。■



Free exchange

Why “Freakonomics” failed to transform economics

The approach was fun, but has fallen out of favour

“ECONOMICS IS A study of mankind in the ordinary business of life.” So starts Alfred Marshall’s “Principles of Economics”, a 19th-century textbook that helped create the common language economists still use today.

Marshall’s contention that economics studies the “ordinary” was not a dig, but a statement of intent. The discipline was to take seriously some of the most urgent questions in human life. How do I pay my bills? What do I do for a living? What happens if I get sick? Will I ever be able to retire?

In 2003 the New York Times published a profile of Steven Levitt, an economist at the University of Chicago, in which he expressed a very different perspective: “In Levitt’s view,” the article read, “economics is a science with excellent tools for gaining answers but a serious shortage of interesting questions.” Mr Levitt and the article’s author, Stephen Dubner, would go on to write “Freakonomics” together. In their book there was little about the ordinary business of life. Through vignettes featuring cheating sumo wrestlers, minimum-wage-earning crack dealers and the Ku Klux Klan, a white-supremacist organisation, the authors explored how people respond to incentives and how the use of novel data can uncover what is really driving their behaviour.

Freakonomics was a hit. It ranked just below Harry Potter in the bestseller lists. Much like Marvel comics, it spawned an expanded universe: New York Times columns, podcasts and sequels, as well as imitators and critics, determined to tear down its arguments. It was at the apex of a wave of books that promised a quirky—yet rigorous—analysis of things that the conventional wisdom had missed. On March 7th Mr Levitt, who for many people became the image of an economist, announced his retirement from

academia. “It’s the wrong place for me to be,” he said.

During his academic career, Mr Levitt wrote papers in applied microeconomics. He was, in his own self-effacing words, “a footnote to the ‘credibility revolution’”. This refers to the use of statistical tricks, such as instrumental-variable analysis, natural experiments and regression discontinuity, which are designed to tease out causal relationships from data. He popularised the techniques of economists including David Card, Guido Imbens and Joshua Angrist, who together won the economics Nobel prize in 2021. The idea was to exploit quirks in the data to simulate the randomness that actual scientists find in controlled experiments. Arbitrary start dates for school terms could, for instance, be employed to estimate the effect of an extra year of education on wages.

Where the Freakonomics approach differed was to apply these techniques to “the hidden side of everything”, as the book’s tagline put it. Mr Levitt’s work focused on crime, education and racial discrimination. The book’s most controversial chapter argued that America’s nationwide legalisation of abortion in 1973 had led to a fall in crime in the 1990s, because more unwanted babies were aborted before they could grow into delinquent teenagers. It was a classic of the clever-dick genre: an unflinching social scientist using data to come to a counterintuitive conclusion, and not shying away from offence. It was, however, wrong. Later researchers found a coding error and pointed out that Mr Levitt had used the total number of arrests, which depends on the size of a population, and not the arrest rate, which does not. Others pointed out that the fall in homicide started among women. No-fault divorce, rather than legalised abortion, may have played a bigger role.

Other economists, including James Heckman, Mr Levitt’s colleague in Chicago and another Nobel prizewinner, worried about trivialisation. “Cute”, was how he described the approach in one interview. Take a paper

on discrimination in the “The Weakest Link”, a game show in which contestants vote to remove other contestants depending on whether they think they are costing them money by getting questions wrong (in the early portion of the game) or are competition for the prize pool by getting them right (later on). That provided a setting in which Mr Levitt could look at how observations of others’ competence interacted with racism and sexism. A cunning design—but perhaps of limited relevance in understanding broader economic outcomes.

At the heart of Mr Heckman’s critique was the idea that practitioners of such studies were focusing on “internal validity” (ensuring estimates of the effect of some change were correctly estimated) over “external validity” (whether the estimates would apply more generally). Mr Heckman instead thought that economists should create structural models of decision-making and use data to estimate the parameters that explained behaviour within them. The debate turned toxic. According to Mr Levitt, Mr Heckman went so far as to assign graduate students the task of tearing apart the Freakonomics author’s work for their final exam.

| *Did you know...*

Neither man won. The credibility revolution ate its own children: subsequent papers often overturned results, even if, as in the case of those popularised by Freakonomics, they had an afterlife as cocktail-party anecdotes. The problem has spread to the rest of the profession, too. A recent study by economists at the Federal Reserve found that less than half of the published papers they examined could be replicated, even when given help from the original authors. Mr Levitt’s counterintuitive results have fallen out of fashion and economists in general have become more sceptical.

Yet Mr Heckman’s favoured approaches have problems of their own. Structural models require assumptions that can be as implausible as any

quirky quasi-experiment. Sadly, much contemporary research uses vast amounts of data and the techniques of the “credibility revolution” to come to obvious conclusions. The centuries-old questions of economics are as interesting as they always were. The tools to investigate them remain a work in progress. ■



自由交流

为何“魔鬼经济学”没能魔改经济学

这个方法有趣，但已失宠

“经济学是一门研究人类一般生活事务的学问。”阿尔弗雷德·马歇尔（Alfred Marshall）的《经济学原理》（Principles of Economics）一书的开篇如此写道。这本19世纪的教科书帮助创造了经济学家至今仍在使用的通用语言。马歇尔认为经济学研究的是“一般”事务，这并非嘲讽，而是陈述意图。这门学科就是要严肃地研究人类生活中一些最紧迫的问题。我怎么支付生活开销？我靠什么谋生？我生病了怎么办？我有退休那一天吗？

2003年，《纽约时报》发表了一篇对芝加哥大学的经济学家史蒂芬·莱维特（Steven Levitt）的人物专访，文中莱维特表达了截然不同的观点。“在莱维特看来，”文章写道，“经济学这门科学有的是绝妙工具来找出答案，但有趣的问题却严重不足。”莱维特和该文章作者斯蒂芬·杜伯纳

（Stephen Dubner）后来共同撰写了《魔鬼经济学》（Freakonomics）。他们在书中几乎完全不谈一般生活事务。两位作者通过各种小故事，例如作弊的相扑选手、拿最低工资的毒品贩子以及信奉白人至上主义的三K党，探讨了人们如何对激励做出反应，以及如何利用新颖的数据来揭示他们行为背后的真正动机。

《魔鬼经济学》风靡一时，在畅销书排行榜上仅次于《哈利·波特》。就像漫威漫画一样，它也催生了一个扩展宇宙：有《纽约时报》专栏、播客和续作，还有各路模仿者和决心推翻其观点的批评家。它是一大波书籍的巅峰之作，这些书无不宣称要对传统观念所忽略的事物进行古怪但严谨的分析。3月7日，对许多人来说已成为经济学家形象代表的莱维特宣布从学术界退休。“这不是我该待的地方。”他说。

在他的学术生涯里，莱维特写过一些应用微观经济学的论文。他本人谦逊地说，他只是“‘可信度革命’的一个脚注”。这指的是使用各种统计技巧——如工具变量分析、自然实验和回归不连续性——从数据中梳理出因果关

系。他普及了戴维·卡德（David Card）、吉多·因本斯（Guido Imbens）和约书亚·安格里斯特（Joshua Angrist）等经济学家的方法，这些人共同获得了2021年的诺贝尔经济学奖。其设想是发掘数据中的怪异现象来模拟科学家在对照实验中发现的随机性。例如，通过任意设定学期的开始日期，可以估算出多接受一年教育对工资的影响。

正如书的副标题所称，《魔鬼经济学》方法的不同之处在于，它将这些方法应用于“隐藏在表象之下的真实世界”。莱维特着重研究的是犯罪、教育和种族歧视。书中最具争议的一章认为，1973年美国全国范围的堕胎合法化导致了1990年代犯罪率降低，因为有更多意外怀上的婴儿在成长为问题少年之前就被打掉了。这是典型的自诩高明文体：一个毫不畏缩的社会科学家利用数据得出一个违反直觉的结论，而且毫不忌讳冒犯他人。然而，这个结论是错的。后来研究人员发现了一个编号错误，指出莱维特引用的是取决于人口规模的总逮捕人数，而不是与人口规模无关的逮捕率。还有人指出，凶杀案的下降首先出现在女性群体中。无过错离婚——而不是合法堕胎——可能起到了更大作用。

其他经济学家担心这会过于琐碎化，其中包括莱维特在芝加哥大学的同事、另一位诺贝尔奖得主詹姆斯·赫克曼（James Heckman）。他在一次采访中将这种方法形容为“蛮机巧的”。以一篇研究游戏节目《智者为王》（The Weakest Link）中的歧视现象的论文为例，在这个节目中，参赛者投票决定是否淘汰其他选手，要么是因为此人答错了问题而导致损失奖金（在游戏前段），要么是因为答对了问题而成为争夺奖金池的对手（在游戏后段）。这提供了一个场景，让莱维特能够审视对他人能力的观察如何与种族主义和性别歧视相互作用。这是一个巧妙的设计——但如果要用来理解更广泛的经济结果可能用处有限。

赫克曼批评的核心是，此类研究的实践者注重“内部有效性”（在估算某种变化的影响时，确保估算的准确性）而非“外部有效性”（这些估算是否适用于更普遍的情况）。相反，赫克曼认为经济学家应该创建决策的结构化模型，然后用数据来估算出可以解释模型行为的参数。这场争论后来变得势如水火。据莱维特称，赫克曼甚至给研究生布置的期末考试就是批驳

这位《魔鬼经济学》作者的研究。

| 你知道吗.....

俩人谁都没赢。可信度革命革了自己的命：研究结果常常被后续的论文推翻，尽管这些结论还是像《魔鬼经济学》那些广为流传的故事一样，成为鸡尾酒会上的谈资。这个问题也蔓延到了经济学的其他领域。美联储的经济学家在近期的一项研究中发现，在他们查看的已发表论文中，即使得到原作者的协助，也只有不到一半可以复现结论。莱维特违反直觉的结论已经过时，经济学家普遍都已经变得更容易怀疑了。

然而，赫克曼青睐的方法也有它自己的问题。结构化模型需要假设，而这些假设可能像任何古怪的准实验一样难以令人信服。遗憾的是，许多当代研究用了大量数据和“可信度革命”的方法，只得出了一些显而易见的结论。那些已经存在了几个世纪的经济学问题一如既往地有趣。而研究这些问题的工具依然有待发展。■



AI got rhythm

A new generation of music-making algorithms is here

Their most useful application may lie in helping human composers

IN THE DYSTOPIA of George Orwell's novel "1984", Big Brother numbs the masses with the help of a "versificator", a machine designed to automatically generate the lyrics to popular tunes, thereby ridding society of human creativity. Today, numerous artificial-intelligence (AI) models churn out, some free of charge, the music itself. Unsurprisingly, many fear a world flooded with generic and emotionally barren tunes, with human musicians edged out in the process. Yet there are brighter signs, too, that AI may well drive a boom in musical creativity.

AI music-making is nothing new. The first, so-called "rules-based", models date to the 1950s. These were built by painstakingly translating principles of music theory into algorithmic instructions and probability tables to determine note and chord progressions. The outputs were musically sound but creatively limited. Ed Newton-Rex, an industry veteran who designed one such model for Jukedek, a London firm he founded in 2012, describes that approach as good for the day but irrelevant now.

The clearest demonstration that times have changed came in August 2023. That is when Meta, a social-media giant, released the source code for AudioCraft, a suite of large "generative" music models built using machine learning. AI outfits worldwide promptly set about using Meta's software to train new music generators, many with additional code folded in. One AudioCraft model, MusicGen, analysed patterns in some 400,000 recordings with a collective duration of almost 28 months to come up with 3.3bn "parameters", or variables, that enables the algorithm to generate patterns of sounds in response to prompts. The space this creates for genuinely new AI compositions is unprecedented.

Such models are also getting easier to use. In September Stability AI, a firm based in London at which Mr Newton-Rex worked until recently, released a model, Stable Audio, trained on some 800,000 tracks. Users guide it by entering text and audio clips. This makes it easy to upload, say, a guitar solo and have it recomposed in jazzy piano, perhaps with a vinyl playback feel. Audio prompts are a big deal for two reasons, says Oliver Bown of Australia's University of New South Wales. First, even skilled musicians struggle to put music into words. Second, because most musical training data are only cursorily tagged, even a large model may not understand a request for, say, a four-bar bridge in ragtime progression (the style familiar from Scott Joplin's "The Entertainer").

The potential, clearly, is vast. But many in the industry remain sceptical. One widespread sentiment is that AI will never produce true music. That's because, as a musician friend recently told Yossef Adi, an engineer at Meta's AI lab in Tel Aviv, "no one broke its heart". That may be true, but some AI firms reckon that they have found a way to retain and reproduce the "unique musical fingerprint" of their musician users, as LifeScore, a company founded near London, puts it. LifeScore's AI limits itself to recomposing the elements of a user's original recordings in ways that maintain the music's feel, rather than turning them into something radically new.

It takes about a day to plug into LifeScore's model the dozens of individually recorded vocal and instrumental microphone tracks, or stems, that go into producing an original song. Once that's done, however, the software, developed at a cost of some \$10m, can rework each stem into a new tempo, key or genre within a couple of seconds. The song's artists, present during the process, choose which remixes to keep. Manually remixing a hit track has traditionally taken one or more highly paid specialists weeks.

LifeScore, says Tom Gruber, a co-founder, is “literally swamped with requests” from clients including Sony Music, Universal Music Group and Warner Music Group. An original release is typically turned into anywhere from a handful to a dozen remixes. But one client aims to release a dizzying 6,000 or so AI versions of an original track, each targeting a different market. Artists including Pink Floyd’s David Gilmour and Tom Gaebel, a German pop singer, use LifeScore’s AI to power websites that allow fans to generate, with a few clicks, new remixes adapted to personal tastes.

| *The beat of a different drum*

If this seems like dizzying progress, it’s worth noting that AI’s impact on music is still in its early days. Legal uncertainties over the use of copyrighted recordings to train models have slowed development. Outfits that have coughed up for licensing fees note that this can get expensive. To save on that cost, MusicGen’s training set mostly sidestepped hits, says Dr Adi. Though output is pretty good, he adds, the model is not yet “artistic enough” to generate narratively complete songs. Harmonic misalignments are common. OpenAI, a San Francisco firm, for its part, says its MuseNet model struggles to pull off “odd pairings”, such as a Chopin style that incorporates bass and drums.

In time, bigger training sets of better music will largely overcome such shortcomings, developers reckon. A Stability AI spokesperson says that while Stable Audio’s top duration for coherently structured music—“intro, development and outro”—is now about 90 seconds, upgrades will produce longer pieces with “full musicality”. But judging music AI by its ability to crank out polished tracks mostly misses the point. The technology’s greatest promise, for now at least, lies elsewhere.

Part of it is the empowerment of amateurs. AI handles technical tasks beyond many people’s capabilities and means. As a result, AI is drawing legions of newbies into music-making. This is a boon for experimentation

by what Simon Cross, head of products at Native Instruments, a firm based in Berlin, calls “bedroom producers”.

Consider RX, a Native Instruments AI “assistant” that corrects errors in things like pitch and timing. For the latter, software time-shifts notes by cutting out or inserting slivers of sound with matching timbre, a process called “dynamic time-warping”. The company’s AI also determines what mixing and mastering processes were performed on a song of a user’s choosing. It then replicates, or at least approximates, the same expensive processing on the user’s own creations. Boomy, an online “music automation” platform for what Alex Mitchell, its CEO, describes as “low-friction” song production with text prompts, has more than 2m users. The company, based in Berkeley, California, uploads users’ (vetted) creations to streaming services and collects a cut of revenues.

AI serves professionals, too. The soundtracks to “Barbie” and “Oppenheimer” were cleaned up in post-production with RX, for example. Another application area is “style transfer”, in which models transform music recorded with one instrument into sounds that seem to come from a different one, often with a twist or two requested by the user. Style transfers are also used for voice. A model developed by a startup in London called Voice-Swap slices up sounds sung by (remunerated) professional singers and rearranges the slivers into lyrics written by the service’s users, who pay licensing fees for the rights to sell the resulting tracks. And AI tools already exist to recreate singers’ voices in other languages. Vocaloid, a voice-synthesising tool from Yamaha, a Japanese instrument manufacturer, is one of many that can use a translation sung by a native speaker as a template for an AI to imitate as it rearranges, modifies and stitches together tiny snippets of the original singer’s voice.

Accomplished musicians now widely tap MusicGen and its competitors as sources of “infinite inspirations”, the better to alight upon promising

composition ideas, says Meta's Dr Adi. Whether such inspiration pays off will, ultimately, be up to the listener to decide. ■



AI律动

新一代音乐制作算法流行

它们最大的用处可能是帮助人类作曲家【深度】

在乔治·奥威尔的小说《1984》所描写的反乌托邦世界中，“老大哥”利用“写诗器”来麻痹大众。这是一种可以自动为流行乐曲配词的机器，有了它，社会就再用不着人类的创造力。如今，许多人工智能（AI）模型都能自动生成乐曲，有些还是免费的。难怪许多人担心世界会充斥着情感贫瘠、千篇一律的曲调，而人类作曲家会在这个过程中被淘汰。不过，也有更光明的迹象，显示AI很可能会推动音乐创作力大爆发。

AI音乐制作并不是什么新鲜事。最早的所谓“基于规则”的音乐制作模型可以追溯到上世纪50年代。构建这些模型要费心费力地将音乐理论原理转化为算法指令和概率表，以确定音符与和弦进行。它们制作的音乐虽然悦耳动听，却没什么创造力。业内资深人士埃德·牛顿-雷克斯（Ed Newton-Rex）就曾为自己2012年在伦敦创办的公司Jukedeck设计过这样一个模型。他表示该方法在当时挺厉害，但现在已经不算什么了。

最能体现时代变化的事件发生在2023年8月。当时，社交媒体巨头Meta发布了AudioCraft的源代码，AudioCraft是用机器学习构建的一套大型“生成式”音乐模型。世界各地的AI公司立即开始使用Meta的软件来训练新的音乐生成器，其中许多还另外增加了代码。AudioCraft的模型之一MusicGen从约40万段总时长近28个月的录制音乐中分析模式，得出了33亿个“参数”（也叫变量），使得该算法能够对提示做出回应，生成不同模式的声音。这为真正全新的AI作曲创造了前所未有的空间。

这类模型也越来越容易使用。去年9月，伦敦公司Stability AI（牛顿-雷克斯不久前还在这家公司工作）发布了Stable Audio模型，使用大约80万首歌曲训练而来。用户输入文字和音频片段来引导它创作。这样就很容易上传譬如一段吉他独奏，将它改编为爵士风格的钢琴曲，或许还可以有几分黑胶唱片的感觉。澳大利亚新南威尔士大学（University of New South

Wales)的奥利弗·鲍恩(Oliver Bown)表示,能够输入音频提示之所以重要,有两个原因。首先,即使训练有素的音乐家也很难将音乐用语言表达出来。其次,因为大多数用于训练的音乐数据只做了粗略的标记,因此,即使是大型模型也可能不明白用户提出的一些需求,比如拉格泰姆曲式中的四小节间奏(斯科特·乔普林[Scott Joplin]的《The Entertainer》以这种风格著称)。

显然, AI音乐制作的潜力巨大。但许多业内人士仍持怀疑态度。一种普遍观点是, AI永远创作不出真正的音乐。这是因为“没有人伤过它的心”——Meta位于特拉维夫的AI实验室的工程师优素福·阿迪(Yossef Adi)说有一位作曲家朋友不久前这么告诉他。这话或许是对的,但一些AI公司认为自己已经找到了方法,可以保留并复制作曲家用户“独一无二的音乐指纹”,这是创立于伦敦附近的LifeScore的说法。LifeScore的AI限定自己,在改编用户原始音乐中的元素时须保留这段音乐原本的感觉,而不是把它们变成全新的东西。

制作一首原创歌曲需要数十个单独录制的人声和器乐音轨(也叫音源),而要将这些音轨或音源输入LifeScore的模型需要大约一天的时间。不过,这个过程一旦完成,耗资约1000万美元开发的LifeScore软件就能在几秒钟内对每个音源重新加工,赋予它们新的节奏、音调或风格。在场的艺术家可以选择保留哪些混音版本。用传统方式对一首热门音乐进行人工混音需要一个或几个收费不菲的行家里手花费几周时间。

LifeScore的联合创始人汤姆·格鲁伯(Tom Gruber)表示,公司对来自索尼音乐、环球音乐集团和华纳音乐集团等客户的“需求简直应接不暇”。一首已发行的原版音乐通常会被改编成几个到十几个混音版本。但有一个客户竟然打算为一首原创音乐发行6000个左右的AI版本——每个版本主打不同的市场。平克·弗洛伊德乐队(Pink Floyd)的大卫·吉尔摩(David Gilmour)和德国流行歌手汤姆·盖贝尔(Tom Gaebel)等音乐人使用LifeScore的AI来支持一些网站——粉丝只需在这些网站上点击几下,就能生成符合个人喜好的新混音版本。

如果说这样的进展似乎快到令人目眩，那么值得注意的是，AI对音乐的影响才刚刚开始。使用有版权的音乐来训练模型在法律上存在不确定性，这减缓了其进展。那些被迫支付了版权费的公司指出，模型训练可能因此变得很昂贵。阿迪表示，为节约这方面的成本，MusicGen的训练集大多避开了热门歌曲。他补充说，虽然生成的作品相当不错，但该模型还没有达到“足够的艺术性”，无法生成叙事完整的歌曲。和声不和谐是家常便饭。旧金山公司OpenAI则表示，它自己的模型MuseNet模型就很难完成“别扭的配对”，比如把低音吉他和鼓点与肖邦风格的乐曲混合在一起。

开发人员认为，假以时日，包含更优质音乐的更大训练集将在很大程度上克服这些缺陷。Stability AI的一位发言人表示，虽然目前Stable Audio制作的结构连贯的音乐（包括“前奏、主题展开和结尾”）最长可以到1分30秒左右，但升级后能制作出更长时间的“富有音乐性”的作品。不过，仅凭能否制作出优美的乐曲来判断音乐AI的未来基本上是不得要领。这项技术最大的前景——至少就目前而言——在其他地方。

其一是给业余音乐人赋能。AI所处理的技术工作是许多人凭自己的能力和财力无法完成的。因此，AI正吸引大批新手进入音乐制作领域。这对西蒙·克罗斯（Simon Cross）所称的“卧室制作人”的音乐制作尝试来说是一大福音。克罗斯是柏林公司Native Instruments的产品主管。

Native Instruments的AI“助手”RX可以纠正音高和时点等方面的错误。对于后者，软件会通过剪切或插入与原音色匹配的声音片段，对音符做时间位移，这一过程被称为“动态时间规整”。该公司的AI还能分辨用户所选的歌曲是如何进行混音和母带处理的，然后把这一成本高昂的操作复制到（或至少近似地复制到）用户自己的作品上。在线“音乐自动化”平台Boomy拥有逾200万用户，CEO亚历克斯·米切尔（Alex Mitchell）形容它根据提示词制作歌曲的过程“低摩擦”。这家加州伯克利的公司将用户（经过审核）的作品上传到流媒体服务平台，并从收入中抽成。

AI也为专业人士服务。例如，电影《芭比》和《奥本海默》在后期制作时

就使用RX修音。另一个应用领域是“风格转换”，即通过模型让一种乐器演奏的音乐变成好像是用另一种乐器演奏的，通常会根据用户的要求做一两处改动。风格转换还可用于人声。伦敦一家创业公司开发了一款名为Voice-Swap的模型，可以将专业歌手（有偿）演唱的声音分割成片段，然后按该公司用户写的歌词重新编排这些片段，用户支付版权费后就可以获得由此生成的歌曲的售卖权。现在已经有很多AI工具可以用其他语言再现歌手的声音。日本乐器制造商雅马哈公司推出的语音合成工具Vocaloid就是其中之一，它可以将母语歌手演唱的译文歌曲用作模版，供AI在重新编排、修改或拼接原唱者的声音片段时进行模仿。

Meta的阿迪表示，为了更好地激发创意，如今许多才华出众的音乐家都将MusicGen及类似的AI模型用作“无限灵感”的来源。至于这种灵感能否带来回报，最终还是要听众说了算。■



China's other TikToks

TikTok is not the only Chinese app thriving in America

What happens to them if the short-video sensation is banned?

WITH THE prospect of a ban hanging over TikTok, its 170m American users may start looking for an alternative time-sink. They could plump for Bigo Live, popular for TikTokish live-streaming. Or Likee, which offers similar video-editing and sharing options and has more than 100m users around the world. There is also Hago, which blends social media and video games, and which has clocked up some 500m downloads.

The three apps are similar to TikTok in another way, too. The firm that makes them, Joyy, is based in Singapore, just like TikTok. And like ByteDance, TikTok's parent firm, Joyy's controlling shareholder, called Huanju, is based in mainland China.

ByteDance and Huanju are among dozens of Chinese firms that are thriving on the free-wheeling Western internet, away from China's censors. They develop apps for everything from gaming and social media to fitness and photo-editing. Many have set up headquarters outside China, often in Singapore, and do not advertise their Chineseness. Plenty dispute the notion that they are Chinese at all.

According to Sensor Tower, a research firm, a Shanghainese video-gaming firm called MiHoYo last year outperformed even Tencent, a Chinese games powerhouse, when it came to foreign revenues from video games (unlisted MiHoYo does not disclose figures). The 30 biggest Chinese video-game makers account for 18% of the industry's global sales outside China. Chinese-linked e-commerce apps have been similarly successful. Shein, which sells cheap clothing, mainly to Americans, is thought to have sold garments worth tens of billions of dollars last year. Temu, which is based in

Boston but owned by a giant Chinese e-merchant, PDD, may have notched up several billion dollars. PDD, which is listed in New York, does not break out Temu's results, but on March 20th it reported that its overall sales more than doubled in the last quarter of 2023, to \$12.5bn.

Other Chinese services are also pushing into America. Daily Yoga, an app popular with exercise aficionados, is owned by a company from the Chinese city of Xi'an. Bestie-Portrait, a photo-editor, has Chinese owners. CapCut, one of the world's most popular video-editing apps, is owned by ByteDance. So is Lemon8, a photo- and video-sharing app similar to Instagram. ReelShort, which serves up soap operas to American viewers in micro-episodes, is owned by Crazy Maple Studios, a Californian subsidiary of COL Digital Publishing, whose headquarters are in Beijing.

A ban on TikTok could open the door to many of these apps being consigned to the digital dustbin. They risk being accused, as TikTok has been, of sharing users' data with Chinese authorities, being a vehicle for propaganda, or both. TikTok denies these allegations, points to its efforts to ring-fence its American data and algorithm, and says that the Communist Party holds no sway over it. Other apps with Chinese roots may soon have to start issuing similar denials. ■



中国其他的TikTok

在美国走红的中国应用不止TikTok

如果这款热门短视频应用被禁，它们会怎么样？

TikTok也许会遭到封禁，它的1.7亿美国用户可能要开始另寻一样消磨时光的东西。他们可能会选择Bigo Live，它因类似TikTok的直播而受到欢迎。或者选Likee，它提供类似TikTok的视频编辑和分享功能，全球用户超过一亿。还有Hago，它将社交媒体和电子游戏合为一体，下载量达到约5亿次。

这三款应用在另一个方面也与TikTok相似。打造它们的公司Joyy总部设在新加坡，就像TikTok一样。而且和TikTok的母公司字节跳动一样，Joyy的控股股东欢聚集团的总部也在中国大陆。

字节跳动和欢聚是在自由的西方互联网上蓬勃发展、远离中国审查的数十家中国公司之一。它们开发各种应用，从游戏、社交媒体到健身和照片编辑。许多公司把总部设在中国境外，通常是在新加坡，而且不大肆声张自己的中国出身。很多人质疑它们能不能算中国公司。

根据研究公司Sensor Tower的数据，上海一家名为米哈游的电子游戏公司去年的海外游戏营收甚至超过了中国游戏巨头腾讯（未上市的米哈游没有披露数据）。中国最大的30家电子游戏公司占到该行业中国以外全球销售额的18%。和中国有关联的电子商务应用也取得了类似的成功。据信，主要向美国人销售廉价服装的Shein去年卖出了价值数百亿美元的服装。总部位于波士顿、但属中国电子商务巨头拼多多旗下的Temu可能已实现了数十亿美元的销售额。在纽约上市的拼多多没有单独公布Temu的业绩，但在3月20日，该公司报告称其2023年最后一个季度的总销售额翻了一番不止，达到125亿美元。

其他中国服务也在进军美国。每日瑜伽（Daily Yoga）是一款受到健身爱好者欢迎的应用，由中国西安的一家公司所有。照片编辑应用Bestie-

Portrait为中国公司所有。CapCut是全球最受欢迎的视频编辑应用之一，属于字节跳动。和Instagram类似的照片和视频分享应用Lemon8也属于这家公司。ReelShort面向美国观众播放微短肥皂剧，它的所有者枫叶互动（Crazy Maple Studios）是总部在北京的中文在线集团的加州子公司。

TikTok若遭封禁可能启动许多此类应用被扔进数字垃圾箱。它们可能会像TikTok一样，被指控与中国政府共享用户数据，或是成为宣传工具，或两者兼而有之。TikTok否认了这些指控，指出它努力保护自己的美国数据和算法，并表示自己不受共产党的控制。其他源自中国的应用可能很快也将不得不开始发表类似的否认声明。■



Schumpeter

Can anything stop Nvidia's Jensen Huang?

He has become the generative-AI showman of our time

JENSEN HUANG is a man on a mission—but not so much that he does not have time to tell a good story at his own expense. Last spring, when his semiconductor company, Nvidia, was well on its way to becoming a darling of generative artificial intelligence (AI), he and his wife bought a new home in the Bay Area. Mr Huang was so busy he could not spare much time to visit it before the purchase was completed. Pity, he admitted later, sneezing heavily. It was surrounded by plants that gave him hay fever.

Mr Huang uses such self-deprecating humour often. When he took to the stage on March 18th for Nvidia's annual developers' conference, to be greeted by cheers, camera flashes and rock-star adulation from the 11,000 folk packed into a San Jose ice-hockey stadium, he jokingly reminded them it wasn't a concert. Instead, he promised them a heady mix of science, algorithms, computer architecture and mathematics. Someone whooped.

In advance, Nvidia's fans on Wall Street had dubbed it the "AI Woodstock". It wasn't that. The attendees were mostly middle-aged men wearing lanyards and loafers, not beads and tie-dyes. Yet as a headliner, there was a bit of Jimi Hendrix about Jensen Huang. Wearing his trademark leather jacket, he put on an exhilarating performance. He was a virtuoso at making complex stuff sound easy. In front of the media, he improvised with showmanship. And for all the polished charm, there was something intoxicating about his change-the-world ambition. If anyone is pushing "gen AI" to the limits, with no misgivings, Mr Huang is. This raises a question: what constraints, if any, does he face?

The aim of the conference was to offer a simple answer: none. This is the

start of a new industrial revolution and, according to Mr Huang, Nvidia is first in line to build the “AI factories” of the future. Demand for Nvidia’s graphics-processing units (GPUs), AI-modellers’ favourite type of processor, is so insatiable that they are in short supply. No matter. Nvidia announced the launch later this year of a new generation of superchips, named Blackwell, that are many times more powerful than its existing GPUs, promising bigger and cleverer AIs. Thanks to AI, spending on global data centres was \$250bn last year, Mr Huang says, and is growing at 20% a year. His company intends to capture much of that growth. To make it more difficult for rivals to catch up, Nvidia is pricing Blackwell GPUs at \$30,000-40,000 apiece, which Wall Street deems conservative.

In order to reap the fruits of this “accelerated-computing”, Nvidia wants to vastly expand its customer base. Currently the big users of its GPUs are the cloud-computing giants, such as Alphabet, Amazon and Microsoft, as well as builders of gen-AI models, such as OpenAI, maker of ChatGPT. But Nvidia sees great opportunity in demand from firms across all industries: health care, retail, manufacturing, you name it. It believes that many businesses will soon move on from toying with ChatGPT to deploying their own gen AIs. For that, Nvidia will provide self-contained software packages that can either be acquired off the shelf or tailored to a user’s needs. It calls them NIMs, Nvidia Inference Microservices. Crucially, they will rely on (mostly rented) Nvidia GPUs, further tying customers into the firm’s hardware-software ecosystem.

So far, so star-spangled. But it is not all peace and love at Woodstock. You need only to recall the supply-chain problems of the pandemic, as well as the subsequent Sino-American chip wars, to see that dangers lurk. Nvidia’s current line-up of GPUs already faces upstream bottlenecks. South Korean makers of high-bandwidth memory chips used in Nvidia’s products cannot keep up with demand. TSMC, the world’s biggest semiconductor manufacturer, which actually churns out Nvidia chips, is struggling to

make enough of the advanced packaging that binds GPUs and memory chips together. Moreover, Nvidia's larger integrated systems contain around 600,000 components, many of which come from China. That underscores the geopolitical risks if America's tensions with its strategic rival keep mounting.

Troubles may lie downstream, too. The AI chips are energy-hungry and need plenty of cooling. There are growing fears of power shortages because of the strain that GPU-stuffed data centres will put on the grid. Mr Huang hopes to solve this problem by making GPUs more efficient. He says the mightiest Blackwell system, known pithily as GB200NVL72, can train a model larger than ChatGPT using about a quarter as much electrical power as the best available processors.

But that is still almost 20 times more than pre-AI data-centre servers, notes Chase Lochmiller, boss of Crusoe Energy Systems, which provides low-carbon cloud services and has signed up to buy the GB200NVL72. And however energy-efficient they are, the bigger the GPUs, the better the AIs trained using them are likely to be. This will stoke demand for AIs and, by extension, for GPUs. In that way, as economists pointed out during a previous industrial revolution in the late 19th century, efficiency can raise power consumption rather than reduce it. "You can't grow the supply of power anything like as fast as you can grow the supply of chips," says Pierre Ferragu of New Street Research, a firm of analysts. In a sign of the times Amazon Web Services, the online retailer's cloud division, last month bought a nuclear-powered data centre.

| *'Scuse me while I kiss AI*

Mr Huang is not blind to these risks, even as he dismisses the more typical concerns about gen AI—that it will destroy work or wipe out humanity. In his telling, the technology will end up boosting productivity, generating profits and creating jobs—all to the betterment of humankind. Hendrix

famously believed music was the only way to change the world. For Mr Huang, it is a heady mix of science, engineering and maths. ■



熊彼特

还有什么能阻挡英伟达的黄仁勋？

他已经成了这个时代的生成式人工智能表演艺术家

黄仁勋当然是重任在肩，不过也还没忙到没空自嘲一把。去年春天，他的半导体公司英伟达正冉冉升起为生成式AI的宠儿，他和妻子在旧金山湾区买了新房。他太忙了，都没去好好看房就买了下来。失算啊，后来他一边打喷嚏一边承认。房子周围全是让他花粉过敏发作的植物。

黄仁勋常常玩这种自黑式的幽默。3月18日，他登上英伟达年度开发者大会的舞台，迎接他的是欢呼声、相机闪光，以及挤满了圣何塞一座冰球馆的11000名观众犹如对摇滚明星的崇拜——他开玩笑地提醒大家这不是在开演唱会。他承诺呈献给他们的是融合了科学、算法、计算机架构和数学的炫目科技。台下有人欢呼叫好。

大会之前，华尔街的英伟达粉就把这场大会称为“AI界的伍德斯托克音乐节”。倒也不是那个样子。与会者大多是脖子上挂着入场证、脚蹬乐福鞋的中年男，而非一身扎染服饰、挂珠戴串的人。不过作为压轴主讲人，黄仁勋确有几分美国摇滚歌手吉米·亨德里克斯（Jimi Hendrix）的风范。他身穿标志性的皮夹克，给出了一场令人振奋的表演。他极擅长以浅显易懂的语言表述复杂难懂的概念。面对媒体，他信手拈来，控场力十足。除了这份从容老练的魅力，他改变世界的雄心壮志也令人沉醉。如果说有谁在把“生成式AI”推向极限，那无疑是黄仁勋。这就引出一个问题：他面对什么制约（如果有的话）？

这次大会就是要给出一个简单的答案：没有制约。这是一场新工业革命的开端，按黄仁勋的话说，英伟达走在打造未来“AI工厂”的最前列。英伟达的图形处理器（GPU）最受AI建模人员的青睐，已经供不应求。没关系。英伟达宣布将在今年晚些时候推出新一代超级芯片Blackwell，性能将比现在的这些GPU强大许多倍，有望造就更大更智能的AI系统。黄仁勋表示，为了发展AI，去年全球用于数据中心的支出达到2500亿美元，并正以每年

20%的速度增长。英伟达一心要抢占这些增长的大头。为了让竞争对手更难追赶，英伟达把Blackwell GPU的价格定在每片三至四万美元，华尔街认为这定价保守了。

为收获这轮“算力加速”的果实，英伟达希望大幅扩大客户群。目前，其GPU的主要用户是Alphabet、亚马逊和微软等云计算巨头，以及开发ChatGPT的OpenAI等构建生成式AI模型的公司。但英伟达从各行各业的企业需求中看到了巨大商机：医疗保健、零售、制造业，应有尽有。英伟达相信，许多企业很快就会从尝鲜ChatGPT转向部署自己的生成式AI。为此，英伟达将提供可以即买即用或根据用户需求定制的独立软件包，名为NIM（Nvidia Inference Microservices）。至关重要的是，这些软件将依赖英伟达GPU（主要是租用），进一步把客户绑定在该公司的硬件—软件生态系统中。

目前为止，一切璀璨光明。但在“伍德斯托克”，并非只有一片祥和友爱。只要回想一下疫情期间的供应链问题以及随后的中美芯片战，就会知道有种种风险潜伏。英伟达目前的GPU产品线已面临上游瓶颈。英伟达产品要用到韩国生产的高带宽内存芯片，而那些韩国制造商无法跟上需求。负责代工生产英伟达芯片的是全球最大半导体制造商台积电，目前其集成GPU和内存芯片的先进封装产能也非常紧张。此外，英伟达更大的集成系统包含约60万个零部件，其中许多来自中国。如果美国与这个战略对手的紧张关系持续升级，其中的地缘风险就会突显。

下游可能也有麻烦。AI芯片能耗很高，需要做大量冷却。由于装载大量GPU的数据中心将给电网带来压力，人们越来越担心会出现电力短缺。黄仁勋希望通过提高GPU的效率来解决这个问题。他说，相比现有最好的处理器，代号为GB200NVL72的最强大Blackwell系统可以只用约四分之一的电力来训练比ChatGPT更大的模型。

但这仍是传统非AI数据中心服务器能耗的近20倍，克鲁索能源系统（Crusoe Energy Systems）的老板切斯·洛克米勒（Chase Lochmiller）指出。他的公司提供低碳云服务，已签约购买GB200NVL72。而无论能效

如何，GPU越多，训练出的AI模型往往表现越好。这将刺激对AI的需求，进而刺激对GPU的需求。结果，正如在19世纪末上一次工业革命时经济学家所指出的，效率的提高可能增加而非减少能耗。“你不可能像增加芯片供应那样快速增加电力供应。”分析公司新街研究（New Street Research）的皮埃尔·费拉古（Pierre Ferragu）说。亚马逊的云计算部门亚马逊云科技（AWS）在3月收购了一家核动力数据中心，很能说明时代趋势。

｜ 等等，让我先亲亲AI

黄仁勋并非看不到这些风险，不过他不在意那些更常见的担忧——生成式AI可能消灭就业或毁灭人类。在他看来，这项技术最终会提高生产率，创造利润并创造就业机会，所有这些都将是造福人类。亨德里克斯有一个广为人知的信仰：音乐是改变世界的唯一方式。对黄仁勋而言，改变世界的是融合了科学、工程与数学的炫目新技术。■



The AI pie

Just how rich are businesses getting in the AI gold rush?

Nvidia and Microsoft are not the only winners

BARELY A DAY goes by without excitement over artificial intelligence (AI) sending another company's market value through the roof. Last month the share price of Dell, a hardware-maker, jumped by over 30% in a day because of hopes that AI will boost sales. Days later Together AI, a cloud-computing startup, raised new funding at a valuation of \$1.3bn, up from \$500m in November. One of its investors is Nvidia, a maker of AI chips that is itself on an extended bull run. Before the launch in November 2022 of ChatGPT, a "generative" AI that responds to queries in uncannily humanlike ways, its market value was about \$300bn, similar to that of Home Depot, a home-improvement chain. Today it is \$2.3trn, \$500bn or so shy of Apple's.

The relentless stream of AI headlines makes it hard to get a sense of which businesses are real winners in the AI boom—and which will win in the longer run. To help answer this question The Economist has looked where value has accrued so far and how this tallies with the expected sales of products and services in the AI "stack", as technologists call the various layers of hardware and software on which AI relies to work its magic. On March 18th many companies up and down the stack descended on San Jose for a four-day jamboree hosted by Nvidia. With talks on everything from robotics to drug discovery, the shindig showed off the latest AI innovations. It highlighted furious competition between firms within layers of the stack and, increasingly, between them.

We examined four of these layers and the firms that inhabit them: AI-powered applications sold outside the stack; the AI models themselves, such as GPT-4, the brain behind ChatGPT, and repositories of them (for example, Hugging Face); the cloud-computing platforms which host many

of these models and some of the applications (Amazon Web Services, Google Cloud Platform, Microsoft Azure); and the hardware, such as semiconductors (made by firms such as AMD, Intel and Nvidia), servers (Dell) and networking gear (Arista), responsible for the clouds' computing power (see chart 1).

Technological breakthroughs tend to elevate new tech giants. The PC boom in the 1980s and 1990s propelled Microsoft, which made the Windows operating system, and Intel, which manufactured the chips needed to run it, to the top of the corporate pecking order. By the 2000s "Wintel" was capturing four-fifths of the operating profits from the PC industry, according to Jefferies, an investment bank. The smartphone era did the same to Apple. A few years after it launched the iPhone in 2007, it was raking in more than half of handset-makers' global operating profits.

The world is still in the early days of the generative-AI epoch. Even so, it has already been immensely lucrative. All told, the 100 or so companies that we examined have together created \$8trn in value for their owners since its start—which, for the purposes of this article, we define as October 2022, just before the launch of ChatGPT (see chart 2). Not all these gains are the result of the AI frenzy—stockmarkets have been on a broader tear of late—but many are.

At every layer of the stack, value is becoming more concentrated. In hardware, model-making and applications, the biggest three companies have increased their share of overall value created by a median of 14 percentage points in the past year and a half. In the cloud layer Microsoft, which has a partnership with ChatGPT's maker, OpenAI, has pulled ahead of Amazon and Alphabet (Google's parent company). Its market capitalisation now accounts for 46% of the cloud trio's total, up from 41% before the release of ChatGPT.

The spread of value is uneven between layers, too. In absolute terms the most riches have accrued to the hardware-makers. These include semiconductor firms, companies that build servers and those that make networking equipment. In October 2022 the 27 public hardware companies in our sample were worth around \$1.5trn. Today that figure is \$5trn. This is what you would expect in a technology boom: the underlying physical infrastructure needs to be built first in order for software to be offered. In the late 1990s, as the internet boom was getting going, providers of things like modems and other telecoms gubbins, such as Cisco and WorldCom, were the early winners.

So far the host of the San Jose gabfest is by far the biggest victor. Nvidia accounts for some 57% of the increase in the market capitalisation of our hardware firms. The company makes more than 80% of all AI chips, according to IDC, a research firm. It also enjoys a near-monopoly in the networking equipment used to yoke the chips together inside the AI servers in data centres. Revenues from Nvidia's data-centre business more than tripled in the 12 months to the end of January, compared with the year before. Its gross margins grew from 59% to 74%.

Nvidia's chipmaking rivals want a piece of these riches. Established ones, such as AMD and Intel, are launching rival products. So are startups like Groq, which makes superfast AI chips, and Cerebras, which makes supersized ones. Nvidia's biggest customers, the three cloud giants, are designing their own chips, too—as a way to reduce reliance on one provider and to steal some of Nvidia's juicy margins for themselves. Lisa Su, chief executive of AMD, has forecast that revenues from the sale of AI chips could balloon to \$400bn by 2027, from \$45bn in 2023. That would be far too much for Nvidia alone to digest.

As AI applications become widespread, a growing share of demand will shift from chips required for training models, which consists of analysing

lots of data in order to teach algorithms to predict the next word or pixel in a sequence, to those needed to use them to respond to queries (“inference”, in tech-speak). In the past year two-fifths of Nvidia’s AI revenues came from customers using its chips for inference. Experts expect some inference to start moving from specialist graphics-processing units (GPUs), which are Nvidia’s forte, to general-purpose central processing units (CPUs) like those used in laptops and smartphones, which are AMD’s and Intel’s. Before long even some training may be done on CPUs rather than GPUs.

Still, Nvidia’s grip on the hardware market seems secure for the next few years. Startups with no track record will struggle to convince big clients to reconfigure corporate hardware systems for their untested technology. The cloud giants’ deployment of their own chips is still limited. And Nvidia has CUDA, a software platform which allows customers to tailor chips to their needs. It is popular with programmers and makes it hard for customers to switch to rival chips, which CUDA does not support.

Whereas hardware wins the value-accrual race in absolute terms, it is the independent model-makers that have enjoyed the biggest proportional gains. The collective value of 11 such firms has jumped from \$29bn to about \$138bn in the past 16 months. OpenAI is thought to be worth \$100bn, up from \$20bn in October 2022. Anthropic’s valuation rose from \$3.4bn in April 2022 to \$18bn. Mistral, a French startup founded less than a year ago, is worth \$2bn or so. Some of that value is tied up in hardware. The startups buy piles of chips, mostly from Nvidia, to train models. Imbue, which like OpenAI and Anthropic is based in San Francisco, has 10,000 such semiconductors. Cohere, a Canadian rival, has 16,000. These chips can sell for tens of thousands of dollars apiece. More sophisticated models need more chips. GPT-4 reportedly cost about \$100m to train. Some suspect that training its successor could cost OpenAI ten times as much.

Yet the model-makers’ true worth lies in their intellectual property and the

profits it may generate. The extent of those profits will depend on just how fierce competition among model providers will get—and how long it will last. Right now the rivalry is white-hot, which may explain why the layer has not gained as much dollar value.

OpenAI seized an early lead, but challengers have been catching up. They are able to tap the same data as the maker of ChatGPT (which is to say text and images on the internet) and, also like it, free of charge. Anthropic's Claude 3 is snapping at GPT-4's heels. Four months after the release of GPT-4, Meta, Facebook's parent, released Llama 2, a powerful rival that, in contrast to OpenAI's and Anthropic's proprietary models, is open and can be tinkered with at will. In February Mistral, which has fewer than 40 staff, wowed the industry with an open model that almost rivals GPT-4, despite requiring much less computing power to train and run.

Even smaller models now offer good performance at a low price, says Stephanie Zhan of Sequoia, a venture-capital (VC) firm. Some are designed for specific tasks. A startup called Nixtla developed TimeGPT, a model for financial forecasting. Another, Hippocratic AI, has trained its model on data from exams to enter medical school, to give accurate medical advice.

The abundance of models has spurred the growth of the application layer. The value of the 19 listed software firms in our application group has jumped by \$1.1trn, or 35%, since October 2022. This includes big software providers that are adding generative AI to their services. Zoom uses it to summarise video calls. ServiceNow, which provides tech and human-resources support to companies, has introduced chatbots to help resolve customers' IT queries. Adobe, maker of Photoshop, has an app called Firefly, which uses AI to edit pictures. Newcomers are adding more variety. "There's An AI For That", a website, counts over 12,000 applications, up from fewer than 1,000 in 2022. DeepScribe helps transcribe doctors' notes. Harvey AI assists lawyers. Thirty-two chatbots promise "sarcastic

conversation” and 20 generate tattoo designs. But fierce competition and low barriers to entry mean that many apps may struggle to capture value.

Then there is the cloud layer. The combined market value of Alphabet, Amazon and Microsoft has jumped by \$2.5trn during the AI boom. Counted in dollars, that is less than three-quarters of the growth of the hardware layer, and barely a quarter in percentage terms. Yet compared with actual revenues that AI is expected to generate for the big-tech trio in the near term, this value creation far exceeds that in the other layers. It is 120 times the \$20bn in revenue that generative AI is forecast to add to the cloud giants’ sales in 2024. The comparable ratio is about 40 for the hardware firms and around 30 for the model-makers.

This implies that investors believe that the cloud giants will be the biggest winners in the long run. The companies’ ratio of share price to earnings, another gauge of expected future profits, tells a similar story. The big three cloud firms average 29. That is about 50% higher than for the typical non-tech firm in the S&P 500 index of large American companies—and up from 21 in early 2023 (see chart 3).

Investors’ cloud bullishness can be explained by three factors. First, the tech titans possess all the ingredients to develop world-beating AI systems: troves of data, armies of researchers, oodles of computing power and plenty of spare cash. Second, the buyers of AI services, such as big corporations, prefer to do business with established commercial partners than with untested upstarts (see chart 4).

Third, and most important, big tech has the greatest potential to control every layer of the stack, from chips to applications. Besides designing some of their own chips, Amazon, Google and Microsoft are investing in models and apps. Of the 11 model-makers in our sample, nine have the support of at least one of the giants. That includes the Microsoft-backed OpenAI,

Anthropic (Google and Amazon) and Mistral (Microsoft again). On March 19th Microsoft announced that it had hired Mustafa Suleyman, founder of Inflection AI, another model-maker, to head a new consumer-AI division—and with him many of Inflection's staff. (Mr Suleyman sits on the board of The Economist's parent company.)

The promise of profits from controlling more layers is leading hitherto layer-specific firms to branch out. OpenAI's VC arm has invested in 14 companies since its launch in January 2021, including Harvey AI and Ambience Healthcare, a medical startup. Sam Altman, boss of OpenAI, is reportedly seeking investors to bankroll a pharaonic \$7trn chipmaking venture. Nvidia is becoming more ambitious, too. It has taken stakes in seven of the model-makers, and now offers its own AI models. It has invested in Together AI and CoreWeave, both of which compete with its big cloud customers. In San Jose it unveiled a new superchip and tools from other layers of the stack. The AI boom's biggest single value-creator is in no mood to cede its crown. ■



AI大饼

AI淘金热，企业挖到了多少？

英伟达和微软不是唯一的赢家【深度】

几乎每天都会有一次这样的兴奋时刻——又一家公司的市值因人工智能（AI）而飙升。因预期AI将提升销售额，硬件制造商戴尔的股价在3月的一天之内猛涨超过30%。几天后，云计算创业公司Together AI完成了新一轮融资，估值已从去年11月时的5亿美元上升到了13亿美元。英伟达是其投资者之一，这家生产AI芯片的公司自己的股价也在一路上涨。2022年11月，ChatGPT这个回应查询直逼真人的“生成式”AI问世，此前英伟达的市值约为3000亿美元，与家居装修连锁店家得宝（Home Depot）相当，而今它的市值已经达到2.3万亿美元，只比苹果公司低了大约5000亿美元。

AI频频登上新闻头条，让人很难辨别哪些企业是这轮AI热潮中真正的赢家——以及哪些企业会获得更长远的成功。为帮助回答这个问题，本刊研究了到目前为止价值在哪里增长最多，以及这种增长与AI“堆栈”（技术专家对AI运行所依赖的各种硬件和软件层的称法）中产品和服务的预期销售量是否相符。3月18日，AI堆栈中的众多公司齐聚圣何塞（San Jose），参加了由英伟达主办的为期四天的大会，探讨从机器人技术到药物发现的各种议题并展示了最新的AI创新。它突显出堆栈各层内部公司之间的激烈竞争，以及各层之间日益激烈的竞争。

我们研究了四个堆栈层以及其中的公司：面向栈外销售的由AI驱动的应用；AI模型本身，例如ChatGPT的大脑GPT-4，以及它们的存储库（如Hugging Face）；托管许多此类模型和一些应用的云计算平台（亚马逊云科技、谷歌云平台、微软Azure）；以及实现云计算能力的硬件，例如半导体（由AMD、英特尔和英伟达等公司制造）、服务器（戴尔）和网络设备（Arista）（见图表1）。

技术突破往往会催生新的科技巨头。上世纪80年代和90年代个人电脑的繁荣推动微软（开发Windows操作系统）和英特尔（制造运行该系统所需的

芯片）成为业界翘楚。根据投资银行杰富瑞（Jefferies）的数据，到了2000年代，“Wintel”拿走了PC行业营业利润的五分之四。智能手机时代也同样造就了苹果的至高地位。在2007年推出iPhone不过几年后，苹果的营业利润就已经占到全球手机制造商的一半以上。

生成式AI世代仍处在发展早期。尽管如此，它已经带来了巨大的获利。总体而言，自AI时代开启以来（本文将AI时代的起点划定在ChatGPT发布前夕的2022年10月），我们研究的约100家公司已为公司所有者创造了8万亿美元的价值（见图表2）。这些增长并不全是AI热潮的结果——近期股市普遍上涨，但其中很多确实是拜AI所赐。

在技术堆栈的每一层，价值都变得更为集中。过去一年半中，前三大公司在硬件、模型创建和应用层的总价值占比中位数增加了14个百分点。在云层，与ChatGPT的开发商OpenAI合作的微软已经超过了亚马逊和Alphabet（谷歌的母公司），其市值现在占到这三家云巨头总市值的46%，高于ChatGPT发布之前的41%。

价值在各层之间也分布不均。从绝对值看，价值增长最多的是硬件制造商，这包括半导体公司、服务器制造商和网络设备制造商。在2022年10月，我们调查的27家硬件上市公司的总价值约为1.5万亿美元，而今这个数字为5万亿美元。在技术发展繁荣期，这样的增长是符合预期的，因为要提供软件首先要建设底层实体基础设施。在上世纪90年代末，随着互联网繁荣期的开始，思科和世通（WorldCom）等提供调制解调器和其他电信设备的公司成了早期的赢家。

到目前为止，圣何塞大会的主办方是毫无争议的最大赢家。在我们所调查的硬件公司中，英伟达约占了市值增长的57%。市场研究公司IDC称，英伟达生产了80%以上的AI芯片。在将数据中心的AI服务器内的芯片连接在一起的网络设备方面，它也享有近乎垄断的地位。在截至1月底的12个月内，英伟达的数据中心业务的收入相比前一年增长了两倍以上。其毛利率从59%增长到了74%。

英伟达的芯片制造竞争对手希望能从这些财富中分一杯羹。AMD和英特尔这样的老牌公司正在推出竞品。像Groq这样制造超快速AI芯片的创业公司以及像Cerebras这样制造超大型芯片的公司也一样。作为英伟达最大的客户，三大云计算巨头也正在设计它们自己的芯片，以减少对单一供应商的依赖，同时也是为了从英伟达丰厚的利润中撬走一块。AMD的首席执行官苏姿丰（Lisa Su）预测，到2027年，AI芯片销售额可能会从2023年的450亿美元飙升到4000亿美元。这样的增长远非英伟达一家能独自消化。

随着AI应用的普及，越来越多的需求将从训练模型（包括分析大量数据以教会算法预测序列中的下一个单词或像素）所需的芯片转向响应查询（用技术术语来说就是“推理”）所需的芯片。在过去一年中，英伟达AI业务收入的五分之二来自于拿它的芯片用于推理的客户。专家预计，部分推理将从使用专用图形处理单元（GPU）转向使用通用中央处理单元（CPU），例如笔记本电脑和智能手机中的CPU，前者是英伟达的强项，后者则是AMD和英特尔的地盘。不久以后，甚至一些训练可能也会使用CPU而非GPU。

尽管如此，英伟达对硬件市场的控制在未来几年里看起来依然稳固。还没有什么成绩的创业公司将难以说服大客户重新配置企业硬件系统去适应它们未经检验的技术。云巨头对自家芯片的使用仍然有限。此外，英伟达还拥有CUDA软件平台，让客户能够根据自己的需求定制芯片。这一平台受到程序员的广泛欢迎，同时也让客户难以转向竞争对手的芯片，因为CUDA不支持这些芯片。

虽然从绝对值的角度看，硬件在价值增长的竞赛中胜出，但增长幅度最大的是独立模型开发商。过去16个月中，11家这类公司的总价值从290亿美元增加到约1380亿美元。据信OpenAI的估值已从2022年10月的200亿美元上升至1000亿美元。Anthropic的估值从2022年4月的34亿美元增至180亿美元。成立不到一年的法国创业公司Mistral的估值约为20亿美元。这些价值的一部分与硬件紧密相关。这些创业公司购买大量主要来自英伟达的芯片用于训练模型。和OpenAI及Anthropic一样来自旧金山的Imbue拥有1万块这样的芯片。它的加拿大竞争对手Cohere有1.6万块。这些芯片的售价可

以高达到几万美元一块。更复杂的模型需要更多芯片。据报道，GPT-4的训练成本约为1亿美元。一些人认为训练后续版本可能会让OpenAI花费十倍的成本。

然而，模型开发商真正的价值在于它的知识产权和可能产生的利润。这些利润的多少将取决于模型开发商之间的竞争会有多激烈、会持续多长时间。目前竞争白热化，这或许可以解释为什么这一层没有获得更多的绝对值增长。

推出了ChatGPT的OpenAI在早期抢得先机，但挑战者已经在迎头赶上。它们能够获得同样的数据（也就是互联网上的文本和图片），而且同样免费。Anthropic的Claude 3正在紧追GPT-4。在GPT-4发布四个月后，Facebook的母公司Meta发布了强大的竞品Llama 2。与OpenAI和Anthropic的专有模型不同，Llama 2是开源的，可以随意修改使用。2月，员工不到40人的Mistral以一款几乎能媲美GPT-4的开源模型令整个行业惊叹，而且它训练和运行模型所需的算力要少得多。

风投公司红杉资本（Sequoia）的斯蒂芬妮·詹（Stephanie Zhan）表示，现在即使更小的模型也能以低价提供良好的性能。一些模型为特定任务而设计。创业公司Nixtla开发了名为TimeGPT的金融预测模型。另一家名为Hippocratic AI的公司用医学院入学考试数据训练其模型来提供准确的医疗建议。

模型的丰富多样促进了应用层的扩张。自2022年10月以来，我们应用组中的19家上市软件公司的价值已经跃升了1.1万亿美元，增幅达35%。这其中包括将生成式AI添加到其服务中的大型软件供应商。Zoom用该技术生成视频通话内容概要。向公司提供技术和人力资源支持的ServiceNow已经推出了聊天机器人来帮助回答客户在IT方面的问题。开发Photoshop的Adobe推出了一款名为Firefly的应用，使用AI编辑图片。新进者正在把应用库变得更加丰富多彩。一个名为"There's An AI For That"的网站统计有超过1.2万个应用，而2022年时还只有不到1000个。DeepScribe帮助转录医生的证明。Harvey AI为律师提供协助。32个聊天机器人声称可以进行"讽刺性对

话”，还有20个可以生成纹身图样。然而，激烈的竞争和低准入门槛意味着许多应用可能会难以捕获价值。

然后是云层。在这轮AI热潮中，Alphabet、亚马逊和微软的总市值增长了2.5万亿美元，按绝对值计算不到硬件层增长的四分之三，按百分比算几乎只有硬件层增长的四分之一。然而，如果是相比AI在短期内预计将为这三大科技公司带来的实际收入，那么这一层的价值创造远远超过了其他层。预计在2024年生成式AI将为这些云巨头增加200亿美元的销售额；2.5万亿美元的市值增长是这个数字的120倍。相应地，硬件层公司的这一倍数约为40倍，模型开发商约为30倍。

这意味着投资者相信从长远来看，云巨头将是头号赢家。另一个衡量公司未来预期利润的指标市盈率也传达了类似的信息。三大云计算公司的平均市盈率为29。这比标准普尔500指数中的那些非科技类大公司高出约50%，也高于它们自己在2023年初时21的平均市盈率（见图表3）。

投资者看好云计算可以从三方面解释。首先，科技巨头拥有开发世界领先AI系统所需的所有要素：海量数据资源、大批研究人员、强大计算能力和大笔可用现金。其次，大公司等AI服务的买家更愿意与成熟的商业合作伙伴打交道，而不是与未经检验的新兴公司合作（见图表4）。

第三点，也是最重要的原因，大型科技公司最有潜力控制堆栈中从芯片到应用的每一层。亚马逊、谷歌和微软除了自己设计一些芯片外，还在投资模型和应用。在我们调查的11家模型开发商中有9家至少有三巨头其一的支持。其中包括微软投资的OpenAI、Anthropic（谷歌和亚马逊投资）和Mistral（也是微软投资）。3月19日，微软宣布已聘请另一家模型开发商Inflection AI的创始人穆斯塔法·苏莱曼（Mustafa Suleyman）来负责一个新的消费AI部门，他还带来了许多Inflection的员工。（苏莱曼是《经济学人》母公司的董事会成员。）

控制更多堆栈层可能带来利润，这让一些原本专注某一层的公司开始跨层拓展。自2021年1月成立以来，OpenAI的风投部门已经投资了14家公司，

包括Harvey AI和医疗创业公司Ambience Healthcare。据报道，OpenAI的老板山姆·奥尔特曼（Sam Altman）正在寻求投资者支持一个高达7万亿美元的芯片制造大计。英伟达的野心也在扩张。它已经持有七家模型开发者的股份，现在又推出了自己的AI模型。它还投资了与其大型云客户竞争的Together AI和CoreWeave。在圣何塞，英伟达推出了一款新的超级芯片和其他堆栈层的工具。作为AI热潮中创造价值最多的一家公司，英伟达可不打算把桂冠拱手让人。 ■



Silicon Valley's scribes

Why is it so hard to write a good book about the tech world?

Blame insularity, secrecy and timing

Burn Book. By Kara Swisher. Simon & Schuster; 320 pages; \$30. Piatkus; £25

Filterworld. By Kyle Chayka. Doubleday; 304 pages; \$28. Heligo Books; £22

WHEN PEOPLE ask Michael Moritz, a former journalist and prominent tech investor, what book they should read to understand Silicon Valley, he always recommends two. "They are not about Silicon Valley, but they have everything to do with Silicon Valley," he says.

One is "The Studio" (1969) by John Gregory Dunne, an American writer who spent a year inside 20th Century Fox watching films get made and executives try to balance creativity with profit-seeking. The other, "Swimming Across" (2001) by Andy Grove, a former boss of Intel, a chipmaker, is a memoir about surviving the Holocaust. It shows how adversity can engender grit, which every entrepreneur needs.

That Sir Michael does not suggest a book squarely about the tech business says a lot. Silicon Valley has produced some of the world's most gargantuan companies, but it has not inspired many written accounts with a long shelf life. Wall Street, by contrast, claims a small canon that has stood the test of time, from chronicles of meltdowns ("Too Big to Fail"), to corporate greed ("Barbarians at the Gate") to a fictionalised account ("The Bonfire of the Vanities") that popularised the term "masters of the universe".

Why not the masters of Silicon Valley? Part of the problem is access, as is often the case when writing about the powerful. Tech executives may let their guards down at Burning Man, but they have been painstakingly

trained by public-relations staff not to get burned by writers. This has been the case for a while. When John Battelle was writing “The Search” (2005), about online quests for information, he spent over a year asking to interview Google’s co-founder, Larry Page. The firm tried to impose conditions, such as the right to read the manuscript in advance and add a footnote and possible rebuttal to every mention of Google. He declined. Google ended up granting the interview anyway.

Journalists who manage to finagle access can feel they owe a company and its executives and, in turn, write meek and sympathetic accounts rather than penetrating prose. Or they cannot break in—or do not even try—and write their book from a distance, without an insider’s insights.

Two new books demonstrate how hard it is to write well about Silicon Valley. “Filterworld” is an outsider’s account of the Valley’s impact, which reads as if it was entirely reported and written in a coffee shop in Brooklyn. The book laments how “culture is stuck and plagued by sameness” and blames Silicon Valley’s algorithms, “the technological spectre haunting our own era of the early 21st century”.

This is the sort of tirade against tech that has spread as widely as Silicon Valley’s apps. It is not wrong, but nor is it insightful. The author, Kyle Chayka, who is a journalist for the New Yorker, never reconciles the tension between the cultural “sameness” he decries and the personalisation everyone experiences, with online users possessing individual feeds and living in separate information bubbles. Nor is this a wholly new phenomenon. People have been complaining about globalisation eroding local culture since “recorded civilisation” began, the author concedes. In 1890 Gabriel Tarde, a French sociologist, lamented the “persistent sameness in hotel fare and service, in household furniture, in clothes and jewellery, in theatrical notices and in the volumes in shop windows” that spread with the passenger train.

“Burn Book” is a better, though imperfect, read. Kara Swisher, a veteran chronicler of Silicon Valley, is both an insider and an outsider. She has attended baby showers for tech billionaires’ offspring and even hosted Google’s top brass for a sleepover at her mother’s apartment. But she has a distaste for the Valley’s “look-at-me narcissists, who never met an idea that they did not try to take credit for”.

In delicious detail, she offers her verdict on the techies who have become household names, such as Facebook’s founder: “As sweat poured down Mark Zuckerberg’s pasty and rounded face, I wondered if he was going to keel over right there at my feet.” That was in 2010, before he had gone through media-training galore. Much as Truman Capote, an American writer, was willing to skewer the socialite swans of New York, Ms Swisher delights in prodding some of her subjects to make readers smile and squirm, such as media mogul Rupert Murdoch (“Uncle Satan”) and Amazon’s Jeff Bezos (who has “a genuinely infectious maniacal laugh”).

Ms Swisher does not have Capote’s élan, but her book succeeds where many fail because she explores the relationship between subject and writer, which lurks in the background of most tech books. In detailing her interactions with tech bosses over three decades, she shows how the industry became more furtive and destructive, less free and fun.

While Ms Swisher uses her memoir to hold up a mirror, unfortunately she does not gaze at it long. After chronicling the internet for the Washington Post, Wall Street Journal and her own outfit, Recode, she moved from the Valley to the swamp—Washington, DC—acknowledging that “I had become too much a creature of the place” and “part of the scene in a way that was starting to feel uncomfortable”.

Still, she declines to tease out some of the more complicated aspects of covering the Valley, such as the thin line between source, friend and

adviser, and exactly how she covered the Valley dispassionately when her then-wife was a Google executive. Despite her journalistic ferocity, the reality was that Ms Swisher could not eviscerate many of her subjects, because she depended on them accepting her invitation to speak at her annual conference, one of her major sources of income, and on her podcast. She was not just “part of the scene”—she played a leading role.

Of course, journalists are not the only ones who deal with personal conflicts that affect how and what they write about tech. Too many in the Valley pursue books to buttress their personal brand, like a website or résumé that just happens to have a spine (but reads as spineless). This explains why so many venture capitalists have ventured into book-writing. The best of the lot is “Zero to One” (2014) by Peter Thiel, an early investor in Facebook, and Blake Masters, a student who took a class taught by Mr Thiel at Stanford. However, explaining how to build a monopoly, as it does with welcome and rare frankness, is probably something Mr Thiel and his peers regret, considering the scrutiny Silicon Valley has since elicited from regulators. Monopolies are not so in vogue these days.

Yet the simplest explanation for why it is so hard for a book about Silicon Valley to hit the mark is probably the most obvious: timing. The snail’s pace of research and publishing is badly suited to Silicon Valley’s speed. Today’s pressing book idea is next year’s stale one. Innovation cycles and companies’ futures often pivot too quickly.

Take Adam Lashinsky, a journalist who wrote a book about Uber. He watched as the company faltered and tried to keep his text up to date. His aptly titled “Wild Ride” was published in 2017, a month before the dramatic firing of Uber’s boss, Travis Kalanick. Mr Lashinsky has since sworn off writing about tech. His next book is about William Safire, a dead newspaper columnist. It is a subject that will not go out of date—and not try to control the narrative. ■



硅谷写书人

写一本关于科技圈的好书为什么这么难？

这要怪偏狭、保密和速度【《看完请销毁》和《过滤世界》书评】

《看完请销毁》，卡拉·斯维什尔著。西蒙与舒斯特出版社，320页；30美元。Piatkus出版社，25英镑。

《过滤世界》，凯尔·恰卡著。道布尔代出版社，304页；28美元。海力高出版社，22英镑。

每次有人问前记者、著名科技投资人迈克尔·莫里茨（Michael Moritz）应该读什么书来了解硅谷时，他总会推荐这么两本书。“它们与硅谷无关，但又与硅谷息息相关。”他说。

一本是美国作家约翰·格雷戈里·邓恩（John Gregory Dunne）的《制片厂》（The Studio, 1969）。他在20世纪福克斯电影公司待了一年，观察电影的制作过程以及高管们如何努力在鼓励创意与追逐利润之间取得平衡。另一本是芯片制造商英特尔的前老板安迪·格罗夫（Andy Grove）所著《横渡生命湖》（Swimming Across, 2001），这是一本大屠杀幸存者的回忆录。它向读者展示了逆境如何催生勇气和毅力，而这种品质是每个企业家都需要的。

莫里茨并没有推荐直接以科技行业为主题的书，这很说明问题。硅谷诞生了一些全球最庞大的公司，但并没有激发出很多经久不衰的相关著述。相比之下，华尔街倒是有一些经受住了时间考验的经典作品，从梳理经济崩溃历程的《大到不能倒》（Too Big to Fail），到描述商界贪婪的《门口的野蛮人》（Barbarians at the Gate），再到让“宇宙之王”一词流行起来的小说《虚荣的篝火》（The Bonfire of the Vanities）。

为什么“硅谷之王”这样的词没能流行起来呢？原因之一是作家们缺少接近科技大佬的机会，这是在写大人物时常见的难题。科技高管在火人节

（Burning Man）上可能会放松警惕，但他们受过公关人员的悉心培训，知道要避开写东西的人以免引火上身。很长时间以来都是这样。在约翰·巴特尔（John Battelle）撰写有关网上信息查询的《搜索》（The Search，2005年）时，他花了一年多的时间请求采访谷歌的联合创始人拉里·佩奇。谷歌提出了一些条件，例如提前阅读书稿、以及在书中每次提及谷歌时有权加脚注或可能的辩驳。巴特尔拒绝了。不过，谷歌最终还是同意了他采访。

那些使点计谋取得近身机会的记者有时会觉得对这家公司和它的高管有所亏欠，结果写出来的内容可能温和并抱有同情，而不是犀利尖锐的。或者，他们打不进内部（甚至试都不试），只是从远处观望着著书，也就不可能有对内情的洞察。

两本新书展示了要写好一本关于硅谷的书有多难。《过滤世界》（Filterworld）是一个圈外人对硅谷影响力的描述，读起来仿佛完全是在布鲁克林的一家咖啡馆里调研写就的。这本书哀叹“文化受到同质化的困扰，停滞不前”，并指责硅谷的算法“是困扰21世纪初这一时代的技术幽灵”。

这种对科技长篇累牍的批评已经和硅谷的应用一样广泛传播。所言不虚，但也没什么深刻的见解。作者凯尔·恰卡（Kyle Chayka）是《纽约客》的记者，他没有解释他所贬斥的文化“同质化”与每个人体验到的个性化（在线用户有各自的信息流，活在相互隔离的信息茧房中）又是如何并存的。而且同质化也不是一个全新现象。作者承认，自“有记载的文明”开始以来，人们一直都在抱怨全球化侵蚀本地文化。1890年，法国社会学家加布里埃尔·塔德（Gabriel Tarde）慨叹，随着客运列车的普及，“酒店菜肴和服务、家具、服装和珠宝、戏剧通告和书店橱窗中陈列的书籍始终千篇一律”。

《看完请销毁》（Burn Book）更值得一读，尽管也不完美。卡拉·斯维什尔（Kara Swisher）是一位资深的硅谷记述者，她是圈内人，又不尽是。她参加过给科技亿万富翁家的孩子举办的迎接新生儿派对，甚至还在

她母亲的公寓里招待过谷歌高层彻夜聚会。但她厌恶硅谷那些“满脸都写着‘快看我’的自恋狂，这些人碰到什么点子都想算到自己头上”。

通过生动有趣的细节描述，她对那些已经家喻户晓的科技牛人做出了自己的评判，比如Facebook的创始人。“当汗珠从马克·扎克伯格那苍白的圆润脸上不断流淌下来，我心想他会不会突然昏倒在我脚边。”这个场景发生在2010年，当时扎克伯格还没有经受过大量来自媒体的试炼。就像美国作家杜鲁门·卡波特（Truman Capote）乐于辛辣讽刺纽约的社交名媛一样，斯维什尔也喜欢调侃她的一些描写对象，比如媒体大亨默多克（“撒旦大叔”）和亚马逊的贝索斯（他有“一种真正具有感染力的疯狂笑声”），让读者会心一笑或是尴尬不安。

斯维什尔的文字没有卡波特那种张扬的活力，但她的书成功做到了其他许多作者没能做到的，因为她探索了写作对象和作者之间的关系，而这在大多数科技业书籍中从不挑明。在详细描述三十年来她与科技大佬们的互动时，她展示了这个行业如何变得更遮遮掩掩也更具破坏性，越来越缺乏自由和乐趣。

斯维什尔用她的回忆录竖起了一面镜子。可惜，她没有长时间凝视它。在为《华盛顿邮报》、《华尔街日报》和她自己的播客Recode记录了互联网的历史后，她从硅谷搬到了华盛顿“沼泽地”。她认识到“我已经过度融入那个地方了”，“成了舞台的一分子，开始感到不舒服”。

然而，她没能反思梳理自己报道硅谷的过程中更复杂的一些层面，比如消息源、朋友和顾问的一线之隔，以及在她当时的妻子担任谷歌高管时，她对硅谷的报道如何能做到客观冷静。尽管她有新闻人的犀利彪悍，在现实中却无法真正去深入剖析她的许多报道对象，因为她还要指望他们受邀到她的年会（她的主要收入来源之一）和她的播客上谈话聊天。她岂止是“舞台的一分子”，她是主角之一。

当然，并非只有新闻记者会面对影响他们如何描写硅谷的个人利益冲突。硅谷有太多人通过写书来支撑他们的个人品牌，书无异于个人网页或

简历长了个书脊（但读起来却觉得作者没长脊梁）。这也解释了为什么有这么多风投家试水写书。这其中最好的一本是彼得·蒂尔与布莱克·马斯特斯（Blake Masters）合著的《从0到1》（2014），蒂尔是Facebook的早期投资者，马斯特斯在斯坦福大学读书时听过蒂尔的课。然而，尽管这本书以难得一见的坦率解释如何建立垄断并深受读者的欢迎，鉴于在这之后监管机构就开始重点关照硅谷，蒂尔和他的同行们可能是后悔做这种吐露的。今天，垄断可不那么流行了。

但是，写一本关于硅谷的好书如此之难的最简单原因可能也是最显而易见的——时机。调研和出版的龟速完全赶不上硅谷的节奏。一个今天难以抗拒的写作创意到明年就过时了。创新周期和公司的未来往往都转向太快了。

比如曾就优步写过一本书的记者亚当·拉辛斯基（Adam Lashinsky）。他一路追踪观察这家公司如何开始掉头向下，努力让自己的记述跟上最新动向。这本书有一个贴切的书名《狂野之旅》（Wild Ride），它于2017年出版，但就在上架一个月后，优步老板特拉维斯·卡兰尼克（Travis Kalanick）突然被踢出了这家公司。这以后，拉辛斯基发誓再也不碰科技题材。他的下一本书写的是已故报纸专栏作家威廉·萨菲尔（William Safire）。这个主人公不会过气，也不会试图控制作者怎么写他。■



AI alignment

How to train your large language model

A new technique is speeding up the process

IT IS NO secret that building a large language model (LLM) requires vast amounts of data. In conventional training, an LLM is fed mountains of text, and encouraged to guess each word before it appears. With each prediction, the LLM makes small adjustments to improve its chances of guessing right. The end result is something that has a certain statistical “understanding” of what is proper language and what isn’t.

But an LLM that has only undergone this so-called “pretraining” is not yet particularly useful. When asked for a joke to cheer your correspondent up, for instance, the pretrained model GPT-2 just repeated the question back three times. When asked who the American president was, it responded: “The answer is no. The president is not the president.” Clearly, teaching an LLM to do what humans want requires something more.

One way to align such models with users’ expectations is through reinforcement learning from human feedback (RLHF). OpenAI, an American startup, introduced this technique in a preprint published in March 2022. It was a major ingredient in its recipe for ChatGPT, which was released eight months later.

RLHF normally involves three steps. First, human volunteers are asked to choose which of two potential LLM responses might better fit a given prompt. This is then repeated many thousands of times over. This data set is then used to train a second LLM to, in effect, stand in for the human being. This so-called reward model, designed to assign higher scores to responses a human would like, and lower scores to everything else, is then used to train the original LLM. As a final touch, a machine-learning

technique called reinforcement learning tweaks the knobs and levers of the original LLM to help reinforce the behaviours that earn it a reward.

This way of doing RLHF is quite involved—using two separate LLMs takes time and money, and the algorithm used for reinforcement learning is, to quote Rafael Rafailov at Stanford University, “quite painful”. This has meant that, outside of OpenAI, Google and their rivals, nobody has really exploited its full potential.

It now turns out that the same results can be achieved for a fraction of the effort. Dr Rafailov and his colleagues, including Archit Sharma and Eric Mitchell, presented this alternative in December 2023 at NeurIPS, an AI conference. Their method, Direct Preference Optimisation (DPO), relies on a satisfying mathematical trick.

This trick hinges on the observation that for every reward model there is a specific theoretical LLM that would get full marks, and every LLM likewise has a theoretical reward model that would give it flying colours. (Just as, more prosaically, every pair of trousers has a theoretical person on whom they would sit perfectly, and every person has a theoretical pair of trousers that would best fit.) This observation that each LLM conceals an implicit reward model allowed the researchers to tinker with this model directly. In the old regime, the LLM learned from the reward model, which learned from the data. Now, the LLM can learn directly from the data.

According to the authors, removing the middleman makes DPO between three and six times more efficient than RLHF, and capable of better performance at tasks such as text summarisation. Its ease of use is already allowing smaller companies to tackle the problem of alignment, says Dr Sharma. A year ago only a few world-leading models, such as Google’s Gemini and OpenAI’s GPT-4, could afford to use RLHF. But as of March 12th eight out of the ten highest-ranked LLMs on an industry leaderboard used

DPO. Mistral, the French startup seeking to rival OpenAI, uses it. Meta, a social-media giant, has integrated it into a home-grown LLM.

Further improvements are sure to come. For one thing, the consensus view is that the big AI labs have made improvements to their proprietary algorithms since they stopped publishing details in 2022. But the problem of getting an LLM to do what a human would want and expect is far from done and dusted. After all, even other humans occasionally struggle. ■



AI对齐

如何训练大语言模型

一种新方法正在加快这一过程

构建大语言模型（LLM）需要大量数据，这已不是什么秘密。在传统的训练中，要给大语言模型灌输大量文本，并鼓励它猜测下一个该出现的单词。针对每次猜测，模型都会进行一些小调整以提高猜对的几率。最终结果是，对于什么是恰当的语言，什么不是，模型有了某种统计意义上的“理解”。

但是，只经过这种“预训练”的大语言模型还不是非常好用。例如，当被要求开个玩笑来给笔者提提神时，预训练模型GPT-2只是把这条提问重复说了三遍。当被问及美国总统是谁时，它回答说：“答案是‘不’。总统不是总统。”显然，要教会大语言模型按人类的需要做事还需要做更多工作。

要让此类模型对齐用户的期望，一种方法是“基于人类反馈的强化学习”（RLHF）。美国创业公司OpenAI在2022年3月发布的预印本论文中介绍了这种方法。这是八个月后发布的ChatGPT的主要支持技术之一。

RLHF通常包括三步。首先，请人类志愿者从大语言模型给出的两个可能回应中选择更符合提示词的那一个。如此重复成千上万次。然后再使用这个数据集训练第二个大语言模型，实际上就是代替了人类的角色。这种所谓的奖励模型会给人类会喜欢的回应分配更高的分数，给其他所有回应分配更低的分数，然后用于训练原始的大语言模型。最后，用一种称为强化学习的机器学习技术调整原始大语言模型的参数，以帮助强化那些为它赢得奖励的行为。

用这种方法来做RLHF相当费劲。使用两个独立的大语言模型耗时费钱，而用于强化学习的算法——用斯坦福大学的拉斐尔·拉斐洛夫（Rafael Rafailov）的话来说——“非常麻烦”。这意味着，除了OpenAI、谷歌以及和他们同档次的对手之外，还没有人能真正完全做好。

现在，人们发现，不用花那么多功夫也可以获得同样的结果。2023年12月，在AI界大会NeurIPS上，拉斐洛夫和阿奇特·夏尔马（Archit Sharma）、埃里克·米切尔（Eric Mitchell）等同事提出了替代方案。他们的方法叫做直接偏好优化（Direct Preference Optimisation, DPO），依赖一个合用的数学技巧。

这个技巧取决于这样的观察，即对于每个奖励模型，理论上都存在一个大语言模型总能从它那里得到满分，同样对于每个大语言模型，理论上也有一个奖励模型总会给它打出满分。（用更通俗的话说就是，每条裤子理论上都能找到一个人完全合身地穿上，而每个人理论上也都能找到一条完全合身的裤子。）每个大语言模型都隐藏了一个隐性奖励模型，根据这一观察结果，研究人员能够直接修补这个隐性模型。在原来的方法中，大语言模型从奖励模型中学习，奖励模型从数据中学习。现在，大语言模型可以直接从数据中学习。

根据作者们的说法，去掉中间人使DPO的效率达到RLHF的三到六倍，并且能够在做文本摘要等任务中有更好的表现。夏尔马说，它的易用性已经使较小的公司能够解决对齐问题。一年前，只有少数几个世界领先的模型，如谷歌的Gemini和OpenAI的GPT-4，能够用得起RLHF。但到3月12日，在一个行业排行榜上排名最高的10个大语言模型中有8个使用了DPO。想要与OpenAI竞争的法国创业公司Mistral也使用了它。社交媒体巨头Meta已将其整合到自家研发的大语言模型中。

未来肯定还会有更多改进。一方面，人们普遍都认为，自2022年停止发布详细信息以来，大型AI实验室已经对其专有算法做了改进。但是，如何让大语言模型做一个人想要和期望的事情，这个问题还远未尘埃落定。毕竟，即使是其他人有时也做不到。■



Crude awakenings

Oil's endgame could be highly disruptive

The oil shocks of the future will be driven by demand, not supply

FOR DECADES, the biggest fears about oil centred on its supply. The lesson was first learnt half a century ago, when the Arab members of OPEC banned exports to America and other supporters of Israel in the Arab-Israeli war. Today you might think that the link between energy and geopolitics has been mercifully severed. Even as war has returned to the Middle East and Russia's invasion of Ukraine has made it a pariah to the West, oil markets have been largely quiescent. In fact, however, a new phase is beginning—one in which oil demand, not supply, will be the primary influence on energy markets. This shift will bring with it profound geopolitical consequences.

Governments everywhere are designing policies to reduce the demand for oil and boost alternative sources of energy, as they seek to fight climate change. Technologies such as those behind electric vehicles are only becoming cheaper and more advanced. The coming peak and subsequent decline of global demand for oil will determine prices and production over the decades to come.

Perversely, this shift will grant some producers more market power. The biggest, least carbon-intensive and cheapest reserves of petroleum by far are found in Saudi Arabia and its immediate OPEC neighbours in the Persian Gulf. As the market for oil shrinks, their share of production will soar. Depending on the pace of the energy transition, this cabal could command a market share of half or even two-thirds of global output by 2050, according to BP, an oil firm, compared with less than 40% today. Already places such as Kuwait, Saudi Arabia and the United Arab Emirates are home to some of the world's largest sovereign wealth funds and are

busily deploying capital and influence in their neighbourhood and beyond. Their piles of capital, and their desire to project their strength abroad, will only intensify.

Meanwhile, other oil powers will be left behind. Today national oil firms in several dozen countries in Africa, Latin America and Asia are pumping oil that is higher-cost and more carbon-intensive than the oil in the Gulf. By one measure, some \$1.2trn of the \$1.8trn in investments planned for the next decade by national oil companies could turn out to be unprofitable if countries make good on their official pledges to achieve net-zero emissions by 2050. Nigeria's NNPC, Mexico's Pemex and Indonesia's Pertamina are among those most at risk of being stuck with stranded assets. Because governments in many producing countries are often unduly reliant on commodity revenues, the failure of some national oil firms could lead to debt crises, bankruptcies and a decade of lost development. This would be a mirror-image of the debt crises that engulfed Latin America in the 1980s, after rising oil prices widened importing countries' trade deficits and crippled their ability to repay their debts.

How to manage this disruption? Speeding up the energy transition is necessary to tackle climate change, but the faster the transition, the worse the concentration of market power, and the greater the shock to high-cost producers. In the meantime coping mechanisms such as governments' strategic petroleum reserves could help to reduce volatility for oil consumers. These should be expanded to include the big developing countries in Asia and Africa, which will soon surpass China to become the biggest contributors to oil demand growth. The International Energy Agency, an official body created in the wake of the first oil shock, already co-ordinates strategic stocks kept by advanced economies. Its new negotiations with India should be expanded to include other big emerging economies, too.

For the unlucky producers, meanwhile, the priority must be to diversify while oil prices are relatively high and demand still strong. A few, ranging from Colombia's Ecopetrol to Malaysia's Petronas, are already spending a hefty share of their capital budgets on low-carbon technologies including renewables, hydrogen and carbon capture that could provide a hedge against an oil collapse. On average, however, national oil firms allocate barely 5% of their capital spending towards diversification; the West's oil majors, by contrast, spend 15%. Governments, too, must seek to ensure that economies can diversify away from oil, by setting business-friendly rules and spending on things like infrastructure and education, to allow private enterprise to thrive. Even so, some countries may nonetheless eventually require bail-outs, putting multilateral institutions under further strain. The supply-led oil shocks of the past half-century were a frequent source of geopolitical tumult. Unless the coming transition is approached with more foresight, the next-half century will be no less fraught. ■



【首文】原油觉醒

石油的终局之战可能极具破坏性

未来的石油危机将由需求而非供应驱动

几十年来，对石油最大的忧虑主要集中在其供应上。这个教训始于半个世纪以前，当时欧佩克（OPEC）的阿拉伯成员国对在阿以战争中支持以色列的美国和其他国家实施石油禁运。今天，你可能认为能源和地缘之间的联系已经被幸运地切断了。在中东战火重燃、俄罗斯入侵乌克兰后受到西方的排斥之时，石油市场大体上风平浪静。不过事实上，一个新的阶段已经开启。在这个阶段，影响能源市场的主要因素将是石油需求而非供应。这种转变将带来深远的地缘后果。

为了应对气候变化，各国政府眼下纷纷制定政策以求减少对石油的需求，促进替代能源的发展。电动汽车等相关技术只会变得越来越便宜、越来越先进。全球石油需求将会达峰继而下降，这将决定未来几十年的油价和产量。

有悖直觉的是，这种转变将赋予一些产油国更大的市场支配力。沙特阿拉伯及其在波斯湾的欧佩克近邻拥有的石油储量无疑是规模最大、碳强度最低、成本也最便宜的。随着石油市场萎缩，它们的产量份额将大幅增加。根据BP的数据，到2050年，该集团可能占据全球石油产量的一半甚至三分之二（取决于能源转型的速度），而当前的占比不到40%。科威特、沙特阿拉伯和阿联酋等国的主权财富基金已位列世界最大，正积极在周边及更远地区部署资本、发挥影响力。它们的资本积累以及向海外投射实力的愿望只会不断加强。

与此同时，其他产油大国将被抛在后面。如今，非洲、拉丁美洲和亚洲几十个国家的国有石油公司开采石油的成本和碳强度均高于海湾地区。根据一项测算，如果各国兑现到2050年实现净零排放的官方承诺，那么国有石油公司未来十年计划投资的1.8万亿美元中约有1.2万亿美元可能无利可图。尼日利亚国家石油公司（NNPC）、墨西哥国家石油公司（Pemex）

和印尼国家石油公司（Pertamina）等都最有可能受困于搁浅资产。由于许多产油国的政府往往过度依赖石油收入，一些国有石油公司的经营失败可能导致债务危机、破产和长达十年的发展停滞。这将与1980年代席卷拉美的债务危机如出一辙，当时油价上涨导致进口国贸易逆差扩大，重创了它们的偿债能力。

如何应对这种混乱？加快能源转型是应对气候变化的必由之路，但转型越快，市场支配力的集中度就越高，对高成本产油国的冲击也越大。与此同时，政府的战略石油储备等应对机制有助于减少石油消费国面临的波动。这些机制应该扩大到亚洲和非洲的发展中大国，这些国家将很快超过中国成为石油需求增长的最大来源。在第一次石油危机之后成立的官方机构国际能源署（IEA）已经在协调发达经济体的战略储备。它与印度的新一轮谈判也应该扩大到其他大型新兴经济体。

｜ 原油觉醒

与此同时，对那些不幸的产油国来说，当务之急是在油价相对较高、需求依然强劲之际实现多元化。从哥伦比亚国家石油公司（Ecopetrol）到马来西亚国家石油公司（Petronas）的少数几家公司已经将资本预算相当大的一部分投入到低碳技术，包括可再生能源、氢能和碳捕获，以求对冲石油需求暴跌的风险。然而，平均而言，国有石油公司用于多元化的资本支出占比还不到5%；相比之下，西方石油巨头为15%。产油国的政府也必须努力确保经济能在石油之外实现多元化，为此制定有利于营商的规则并投资于基础设施和教育等领域，以让私营企业蓬勃发展。即便如此，一些国家最终可能还是需要救助，这将使多边机构承受更大压力。过去50年来，由供应引发的石油危机频繁掀起地缘动荡。除非以更有远见的态度应对即将到来的转型，否则未来50年的麻烦也不会少。 ■



Water, water, everywhere

How to harvest moisture from the atmosphere

New technologies could provide water to Earth's most arid climates

EVEN IN THE most speculative reaches of science fiction, there is no escaping humanity's dependence on liquid water. Luke Skywalker, the hero of the original "Star Wars" trilogy, grows up on his uncle's moisture farm, extracting water from Tatooine's arid atmosphere. The residents of the desert world Arrakis, accessible to anyone with a copy of Frank Herbert's novel "Dune" (or with three hours to kill at their nearest cinema), likewise use windtraps to steal precious liquid from the air.

Engineers on Earth, too, are increasingly looking to the atmosphere for water. They have good reason to do so. Even in the depths of Chile's Atacama Desert, often called the driest place on Earth, estimates suggest that fog and dew can generate some 200ml of water per square metre. Elsewhere, the atmosphere is even more generous. Worldwide, it is estimated to contain 12,900 cubic kilometres of water, roughly the volume of Lake Superior. Moreover, models indicate that evaporation driven by global warming will increase these levels by 27% over the course of the next 50 years.

Tapping this invisible reservoir is a priority. As Earth's temperatures rise and its population grows, ever more people are likely to run short of water. More than 2.3bn are currently living in water-stressed countries and analysts predict that further droughts will force roughly a third of these to move from their homes by 2030.

Collecting water from the air is nothing new. The Inca, who are widely thought to have invented the technique, placed buckets under trees to collect the condensation from heavy fog drifting in off the sea. On the

Canary Islands laurels, junipers and pines have come to be known as “fountain trees” for their association with fog harvesting. People dwelling in the arid mountains of Oman have long built cisterns under trees for the same reason.

Modern atmospheric water harvesting follows many of the same principles. Instead of using leaves as condensation traps, however, which drip over an impractically large area, modern traps instead consist of sheets of very fine polymer mesh. As fog flows through the sheets, tiny droplets of water stick to the polymer fibres. These droplets grow until gravity pulls them into a compact trough and, thence, a reservoir. While collectors vary in size, a 40-metres-square collector in a reasonably foggy area yields around 200 litres a day, enough to supply 60 people with drinking water.

Further improvements are possible. A team led by Urszula Stachewicz at AGH University of Krakow in Poland found that the sheet could be made even more productive by changing the way in which its polymer threads were manufactured. Dr Stachewicz theorised that careful manufacture via a process known as electrospinning could lend the sheet a slight electrical charge that would prove attractive to water droplets in fog. In experiments conducted in 2021 she and a team of colleagues found that such sheets yielded 50% more water.

This past August, she and Gregory Parisi, a PhD student at Rensselaer Polytechnic Institute in New York reported yet further improvements by adding titanium dioxide (TiO_2) to the mesh. Previous work had shown that titanium dioxide could be rendered superhydrophilic (intensely attractive to water) upon exposure to ultraviolet light—a hindrance under extremely foggy conditions, as water would stick to the mesh rather than trickle into the cistern. When fog was light, however, Dr Stachewicz and Mr Parisi found that a TiO_2 -enhanced mesh became 30% more effective. Her fog collectors are now being used at sites on three continents.

Further inland, where fog is scarce, other solutions are needed. One effective approach involves harnessing the water already present in the air. When the temperature drops, the water-carrying-capacity of air decreases with it. This leads to excess water condensing onto surfaces, a process most often seen as dew. It is common in water-saturated places like Britain, but anywhere with little wind and an average relative humidity of 70% or greater can cajole water out of the air.

A key way to do this is with radiative cooling, a phenomenon that occurs at night when certain materials (like aluminium) radiate enough heat to cool below the ambient temperature of their surroundings. After sunset, water condenses on these materials, forms droplets and trickles off. Chambers built of these radiative materials sometimes include adsorbent inner surfaces to which water in the air readily sticks. When humid air drifts into such chambers, it loses its water upon exposure to the cool conditions before drifting out. One big benefit is that such techniques work best in places like deserts, with clear skies, high daytime temperatures and cool nights.

An important limitation of radiative cooling has long been its relative ineffectiveness by day. That changed in 2021 when Dimos Poulikakos and his then-doctoral-student Iwan Haechler, at ETH Zurich, crafted a piece of glass with a layer of silver at the bottom and a layer of silicon polymer, sandwiched between layers of chromium, at the top. The silver layer reflected away the incoming sunlight while the sandwiched polymer allowed the device to shed heat in the form of infrared radiation. This cooled the glass by up to 15°C below ambient temperatures, driving condensation even during the heat of the day. Paired with a heat shield, a condensation chamber built with this glass helped produce 1.2 litres of water per square metre a day.

Another challenge posed by radiative cooling systems is that water needs to

be wiped off the surface of the collection chambers. This requires power, typically from nearby turbines or solar panels, which can be expensive. To cut costs, Dr Poulikakos and Dr Haechler applied a superhydrophobic coating to the surface of the chamber, forcing water droplets off the surface and making it possible for the device to function without electricity.

Such technology is indeed affordable, with the prototype itself costing less than \$50. But in many regions where water is desperately needed, humidity levels are too low for dew harvesting to be feasible. In places like these, the most promising options are those that make use of superabsorbent materials.

Many salts, chemical cousins of the familiar sodium chloride, will readily snatch water out of the air. With this in mind, an engineering team led by Peng Wang at King Abdullah University of Science and Technology in Saudi Arabia studied the effectiveness of hollow nanocarbon capsules filled with lithium chloride. In 2020 the researchers reported that these capsules could capture more than double their weight in water vapour from ambient air when relative humidity was below 60%. Similar techniques using other salts have proved capable of gathering water in humidity levels as low as 10%.

The findings are promising, but the technology has yet to advance beyond the prototype stage. The problem is inefficiency; even Dr Wang's world-leading capsules can only produce 1.6 litres of water per kilogram of lithium chloride over the course of ten hours in very arid conditions. Better than nothing, but inadequate for sustaining a community.

Between them, though, these technologies suggest a brighter future is possible. Areas so dry as to have gone without rainfall since modern records began may one day yield enough water to sustain settlement. And not just on a fictional planet. ■



水，水，到处有水

如何从空气中取水

新技术可为地球上最干旱的地区提供水源【深度】

即使在最具想象力的科幻小说中，人类也同样离不开液态水。《星球大战》正传三部曲的主角卢克·天行者（Luke Skywalker）在他叔叔的湿气农场长大，他们从塔图因星球（Tatooine）干燥的空气中提取水分。只要你有一本弗兰克·赫伯特（Frank Herbert）的小说《沙丘》（Dune）（或者在离你最近的电影院消磨三个小时），就可以走近沙漠星球阿拉基斯（Arrakis），那里的居民同样使用捕风器从空气中收集宝贵的液态水。

如今，地球上的工程师也越来越多地转向从空气中取水。他们这么做是有充分理由的。据估计，即使在有地球最干燥地区之称的智利阿塔卡马沙漠（Atacama Desert）的腹地，雾气和露水也可以产生每平方米约200毫升的水。而在其他地方，空气能提供的水分就更多了。据估计，全世界空气中的含水量达12,900立方公里，相当于一个苏必利尔湖。此外，多个模型显示，在接下来的50年里，全球变暖导致的水体蒸发会让空气中的含水量增加27%。

利用这个看不见的水库是当务之急。随着地球温度的上升和人口的增长，将会有越来越多的人面临缺水问题。目前有超过23亿人生活在水资源短缺的国家，分析人士预测，到2030年，他们中约有三分之一的人会因更严重的缺水问题而背井离乡。

从空气中收集水分并不是什么新鲜事。印加人被公认为是这项技术的鼻祖——他们在树下放置水桶，从海上飘来的浓雾中收集凝结的水滴。在加那利群岛，月桂树、杜松和松树因能聚集雾气中的水而被称为“喷泉树”。出于同样的原因，居住在阿曼干旱山区的人们也很早就开始在一些树底下建造蓄水池。

现代大气集水技术遵循许多相同的原理。不过由于树叶的过水面积较大，不适合用作凝结水的收集器，因此新型集水器由非常细密的聚合物网片构成。当雾气飘过这些网片时，细小水滴会附着在聚合物纤维上。这些水滴不断增大，直到在重力作用下滴落到一个小水槽中，然后流入蓄水池。虽然集水器的大小各不相同，但在雾气浓度尚可的地区，一个40平方米的集水器每天可以收集200升左右的水，足够60人饮用。

这项技术还有改进的空间。波兰的克拉科夫AGH科技大学（AGH University of Krakow）由乌尔苏拉·斯塔切维奇（Urszula Stachewicz）领导的研究小组发现，通过改变网片中聚合物单丝的制造方法，可以提高网片的集水效率。斯塔切维奇推断，利用一种被称为静电纺丝的工艺精心制造的网片能带有少量电荷，应该能吸引雾气中的水滴。在2021年进行的实验中，她和同事发现这种网片让集水量增加了50%。

去年8月，她和纽约的伦斯勒理工学院（Rensselaer Polytechnic Institute）的博士生格雷戈里·帕里西（Gregory Parisi）发布报告称，通过在网片材料中添加二氧化钛（ TiO_2 ），能进一步提高吸水能力。之前的研究显示，二氧化钛在紫外线的照射下，可以产生超亲水性（对水具有超强的吸引力），这在雾气非常大的情况下是一种不利条件，因为水会粘在网片上，而不是流淌到蓄水池中。不过，斯塔切维奇和帕里西发现，当雾气不大时，添加了二氧化钛的网片的集水效率提高了30%。目前她的雾水收集器在三大洲的多个地点都有使用。

在雾气稀少的内陆则需要其他办法。一种有效方法是利用空气中已经看得到的水。当温度下降时，空气的携水能力会随之减弱。这会让多余的水分凝结在物体表面，形成我们常见的露水。这种情况在英国等空气含水量饱和的地方很普遍，但只要一个地方风少、且平均相对湿度达到或超过70%，就可以从空气中提取水。

这方面的一个关键方法是辐射冷却，这是一种发生在夜间的现象——当某些材料（比如铝）散发足够的热量，让自己的温度低于周围环境温度时，就会产生辐射冷却。日落之后，水分凝结在这些材料上形成水滴，然后滴

落下来。用这些辐射冷却材料建造的冷凝室有时会配备具有吸附功能的内壁，这样空气中的水分就很容易附着在上面。当潮湿的空气飘入这样的冷凝室，就会在它温度较低的环境中失去水分后再飘出去。这种方法有一大优势，就是在沙漠等天气晴朗、昼夜温差大的地方效果最好。

辐射冷却长期以来的一大局限性是，相对于夜间，它在白天基本不起作用。这种情况在2021年得到改变。当时苏黎世联邦理工学院（ETH Zurich）的迪莫斯·普利卡科斯（Dimos Poulikakos）和他的博士生伊万·哈伊奇勒（Iwan Haechler）制作了一块玻璃。这块玻璃的底层是一层银，上层是夹在两层铬之间的硅聚合物。银层能反射掉射入的阳光，而夹层中的硅聚合物可以让玻璃以红外线辐射的方式散热。这样一来，这种玻璃的温度可以比周围环境温度低最多15℃，即使是炎热的白天也会有水汽凝结。用这种玻璃建造的冷凝室再配上隔热罩，每天每平方米就能产生1.2升水。

辐射冷却系统面临的另一个挑战是，需要刮取冷凝室表面的水分。这需要用电，通常是由附近的风力涡轮发电机或太阳能电池板供电，这些设备可能造价不菲。为了降低成本，普利卡科斯和哈伊奇勒给冷凝室表面涂上超疏水涂层，让水滴能从表面滴落，从而让收集室不用电也能运转。

这种技术在价格上切实可行，原型机本身的成本不到50美元。但在许多极度缺水的地区，由于空气湿度太低，露水收集不太可行。在这种地方，最有前景的方法是使用超吸水性材料。

有很多与我们熟悉的氯化钠同族的盐类化合物可以轻易吸收空气中的水分。基于这一特性，沙特阿拉伯的阿卜杜拉国王科技大学（King Abdullah University of Science and Technology）的王鹏领导的一个工程团队研究了填充有氯化锂的空心碳纳米胶囊的吸水性能。2020年，他们发布报告称，当相对湿度不到60%时，这些胶囊可以从周围空气中捕获相当于自身重量两倍多的水汽。使用其他盐类化合物的类似方法已被证明能在相对湿度低至10%的情况下收集水分。

这类研究成果有很好的前景，但这项技术还处于原型阶段。目前的问题是效率低下：即使是王鹏发明的全球领先的胶囊，在非常干旱的环境下，每公斤氯化锂在10小时内也只能收集1.6升水。虽说好过一无所获，但这样的集水量还不足以供应一个社区。

不过，如果上述技术一起发力，就有可能带来更光明的前景。自有现代记录以来滴雨未下的干旱地区有朝一日也可能产生足够的水让人类定居下来。而且这可不是发生在一个虚构的星球上。■



Angry young men

Making sense of the gulf between young men and women

It's complicated. But better schooling for boys might help

MEN AND women have different experiences, so you would expect them to have different worldviews. Nonetheless, the growing gulf between young men and women in developed countries is striking. Polling data from 20 such countries shows that, whereas two decades ago there was little difference between the share of men and women aged 18-29 who described themselves as liberal rather than conservative, the gap has grown to 25 percentage points. Young men also seem more anti-feminist than older men, bucking the trend for each generation to be more liberal than its predecessor. Polls from 27 European countries found that men under 30 were more likely than those over 65 to agree that “advancing women’s and girls’ rights has gone too far because it threatens men’s and boys’ opportunities”. Similar results can be found in Britain, South Korea and China. Young women were likely to believe the opposite.

Unpicking what is going on is not simple. A good place to start is to note that young women are soaring ahead of their male peers academically. In the European Union fully 46% of them earn degrees, versus 35% of young men, a gap that has doubled since 2002. One consequence is that young women are more likely than men to spend their early adulthood in a cocoon of campus liberalism. Meanwhile, boys outnumber girls at the bottom end of the scholastic scale. Across rich countries, 28% of them fail to learn to read to a basic level. That is true of only 18% of girls.

Another big change is that, to varying degrees across the developed world, immense progress has been made in reducing the barriers to women having successful careers. College-educated men are still thriving, too—often as one half of a double-high-income heterosexual couple. Many

men welcome these advances and argue for more. However, those among their less-educated brothers who are struggling in the workplace and the dating market are more likely to be resentful, and to blame women for their loss of relative status. And young women, by and large, are glad of past progress but are keenly aware that real threats and unfairness remain, from male violence to the difficulty of juggling careers and children. In short, most young women and worryingly large numbers of young men complain that society is biased against their own sex.

Young women tend to vote for parties of the liberal left. Angry young men, sometimes dismissed as toxically masculine by those parties, are being shrewdly wooed by politicians from the right and the far right. In South Korea their support helped an overtly anti-feminist president win power. In America polls are muddy but some pollsters think young men are souring on the Democrats. In Europe, where many countries offer a kaleidoscope of political choices, young male votes have helped fuel the rise of reactionary outfits such as the AfD in Germany, Confederation in Poland and Chega, which surged at Portugal's election on March 10th.

There is no easy solution to any of this. But clearly, more should be done to help boys lagging behind at school to do better. Some policies that might work without harming their female classmates include hiring more male teachers (who are exceptionally scarce at primary schools in rich countries), and allowing boys to start school a year later than girls, to reflect the fact that they mature later. Better vocational training could encourage young men to consider jobs they have traditionally shunned, from nursing to administration. Schooling boys better would not only help boys. Increasing the supply of educated and (one hopes) less angry men would be good for the women who must share the same world. ■



【首文】愤怒男青年

理解年轻男女之间的分歧

这很复杂。但改善对男孩的学校教育成效可能会有帮助

男人和女人有不同的经历体验，所以你能想到他们的世界观会不一样。尽管如此，发达国家年轻男女之间日益扩大的分歧还是让人惊讶。来自20个发达国家的民意调查数据显示，20年前自称自由派而非保守派的18至29岁的男性和女性在比例上几乎没有差别，但现在这一比例的差距已经扩大到25个百分点。年轻男性似乎也比年长男性更反对女权主义，这与一代比一代更倾向自由主义的趋势背道而驰。来自27个欧洲国家的民意调查发现，30岁以下的男性比65岁以上的男性更有可能同意“提升女性和女孩的权利已经过头了，因为威胁到了男性和男孩的机会”。类似的调查结果也出现在英国、韩国和中国。而年轻女性的看法往往恰恰相反。

要搞清楚这到底是怎么回事并不简单。可以先来看看这样一个事实：年轻女性在学业上的成就正迅速领先于同龄男性。在欧盟足足有46%的年轻女性获得了学位，而年轻男性为35%，这一差距自2002年以来已经翻了一倍。一个结果是年轻女性比年轻男性有更高的比例在校园自由主义的茧房里度过她们成年的早期时光。同时，学业成绩垫底的学生里，男孩比女孩多。在整个富裕世界，有28%的男孩不具备基本的阅读能力，这样的女孩只占18%。

另一个重大变化是，发达国家在减少女性通向事业成功的障碍方面取得了不同程度的巨大进步。受过大学教育的男性仍然过得顺风顺水，他们往往会与一位同样高收入的女性结为伴侣。许多这样的男性欢迎这些进步，并且支持更进一步。但是，他们那些学历更低、在职场和婚恋市场中频频碰壁的兄弟们更有可能感到不满，并认为是女性造成了他们相对地位的下降。总的来说，年轻女性对过去的进步感到高兴，但也敏锐地意识到真正的威胁和不公平仍然存在，从男性暴力到难以兼顾事业和孩子等等。简而言之，大多数年轻女性和数量多到令人担忧的年轻男性都抱怨社会对自己的性别存有偏见。

年轻女性倾向于投票给自由主义左翼政党。而愤怒的男青年——他们有时被那些政党斥为“直男癌”——正在被来自右翼和极右翼的政客们不失时机地拉拢。在韩国，他们的支持帮助一位公开反对女权主义的总统赢得了权力。在美国，民调结果含混复杂，但一些民调专家认为，年轻男性对民主党渐生反感。在欧洲，许多国家存在多种多样的政治选择，年轻男性的选票帮助推动了极端保守政党抬头，如德国的选择党（AfD）、波兰的联邦党（Confederation）和在3月10日葡萄牙大选中异军突起的“够了”（Chega）党。

这些问题都没有简单的解决办法。但很明显，应该采取更多措施来帮一把学校里落后的男生。一些可能有用而又不会损害女同学的政策包括：聘请更多男教师（在富裕国家的小学里，男教师非常稀缺），以及考虑到男孩比女孩更晚熟的现实，允许男孩晚一年上学。更好的职业培训可以鼓励年轻男性考虑他们以往会绕道的工作，从护理到行政等。提升对男孩的学校教育成效不仅对男孩有益。多一些受教育程度高且（希望是）不那么愤怒的男性，对于必须和他们共享同一个世界的女性也是好事。■



The China-Mexico-US triangle

Could there be a US-Mexico trade war?

The bilateral trade deficit has exploded since Donald Trump was president

MEXICAN AND US officials were cheered by trade data released in February. It showed that Mexico overtook China in 2023 to become the number-one exporter to the United States (see chart). The value of goods sold has been rising steadily, and reached almost \$476bn for the year; the equivalent figure for Chinese goods fell sharply, from \$536bn in 2022 to \$427bn in 2023.

The last time Mexico overtook China, in 2002, it was a blip (and Canada held the top spot). Today its rise has the whiff of a new order for global trade. Data for January 2024 show the trend remains strong. United States' efforts to decouple from China and bring supply chains closer to home are intensifying, pushing trade through other countries. An updated free-trade agreement between the United States, Mexico and Canada, known as USMCA, has been in force since 2020. It supports Mexican exports of car parts, medical supplies and agricultural products to the United States.

But the third side of this trade triangle, between Mexico and China, is creating tension. China's firms did not sit idle while incentives were created for companies to look beyond its shores. They have been pushing into the same markets where the US aims to supplant China, including Mexico. This means many Chinese exports are simply taking "a slightly longer road to get to the same place", says Ana Gutiérrez of IMCO, a think-tank in Mexico City.

China appears to be promoting this strategy. In December the country's leaders said it was a priority to export products that are used to make finished goods, rather than the finished goods themselves. Mexico is an

attractive beachhead into the United States because USMCA gives tariff-free access for goods made with enough North American content.

Mexico's official customs data show no sustained influx of goods from China. But some US officials and industrialists believe Chinese inflows are being undercounted, whether deliberately or not. The suspicion is that Mexico turns a blind eye to imports from China, and that those are then re-exported to the United States.

Steel and aluminium are the main concern. In February Katherine Tai, the United States Trade Representative (USTR), noted a "lack of transparency regarding Mexico's steel and aluminium imports from third countries". In December Mexico imposed tariffs of up to 80% on some steel imports from China, but US officials remain frustrated. Setting tariff levels is one thing; actually enforcing them is another.

"What we've seen is that USMCA has really become a US, China and Mexico deal, where China is trans-shipping a lot of product through Mexico," says Jeff Ferry of the Coalition for a Prosperous America, a group that represents manufacturers.

Electric vehicles are a looming concern. The price of the average EV in China is roughly half that in the United States, and China produces more than half the world's output. Without hefty tariffs in place, Chinese EV sales in the United States would probably boom, much as they have in other countries. President Joe Biden's administration is mulling raising tariffs on the vehicles above their current level of 25%.

USMCA has rules against unfair subsidies and market practices, which are common among Chinese companies. But in many cases no law or rule is being breached. Mexico offers Chinese automakers a path around the tariff wall because USMCA's rules of origin contain what one US official calls

“loopholes” that allow for the integration of Chinese components. An importer can assemble Chinese components in Mexico and label Mexico as the country of origin, obscuring Chinese involvement. “New tools may be needed,” says the official.

There is already a large car-manufacturing industry in Mexico, and Chinese money is pouring in, especially to northern states that are the biggest exporters to the United States. In February BYD, China’s most successful EV manufacturer, said it would make 150,000 vehicles a year in Mexico. BYD says that production will serve the local market, but many companies are eyeing the larger, wealthier market north of the border.

Chinese foreign direct investment in Mexico hit \$2.5bn in 2022. Margaret Myers, of the Inter-American Dialogue, a think-tank in Washington, notes the “remarkable growth” of sophisticated manufacturing by Chinese companies in Mexico.

| *No China in the chain*

What might the United States do about this? In the case of steel and aluminium imports, it could copy its approach with the EU and Japan, where there are limits on the volumes that can be imported at lower tariff rates. Higher tariffs kick in after those limits have been reached. When the Trump administration lifted steel and aluminium tariffs on Mexico in 2019, it was on the condition that Mexico restrain export surges into the United States.

Electric vehicles are trickier. Some would like new rules to cut China out of supply chains entirely. “If you want to be a trading partner with the United States, we’re not going to allow your country to be a stopping point for goods from China,” says Mr Ferry. That would mean restrictive rules of origin and a tighter enforcement regime. It would also raise thorny questions about how to treat production by Chinese-owned factories in

Mexico.

Robert Lighthizer, who was the lead trade negotiator during Donald Trump's presidency, has said that a first step would be to strip China of its most-favoured-nation status. That would automatically ratchet up tariffs on Chinese products across the board. Enrique Dussel of UNAM, a university in Mexico City, says this would unmoor the world trade system. "The United States [would be] saying 'adopt my rules or you're against me'."

If Mr Trump wins in November, he is likely to adopt a harder stance toward Mexico. As well as Chinese trade, there is also the issue of migration. But perhaps most importantly Mr Trump loathes trade imbalances. The United States' trade deficit with Mexico rose to \$152bn in 2023, up 17% from 2022. In 2026 Mexico and the United States, along with Canada, have to discuss whether to extend USMCA's validity by another 16 years, to expire in 2052 instead of 2036. Mr Trump signed USMCA, but that is no guarantee that he would not scrap it, or use its extension as leverage to extract concessions from Mexico. He has already talked of a 10% import tariff on goods from all countries, which is not possible for Mexico and Canada under USMCA's rules.

Mexico seems unprepared, says Mr Dussel. Claudia Sheinbaum, who is expected to be elected Mexico's next president in elections on June 2nd, is promoting "nearshoring" to raise Mexico's economic growth to 5% of GDP annually. "Mexico has an enormous sign reading 'Welcome China,'" says Mr Dussel.

The question confronting Mexico is the extent to which it is willing to risk its vital relationship with the United States, to keep that welcome sign shining. ■



中墨美三角

会爆发美墨贸易战吗？

自特朗普任期以来，美墨双边贸易逆差呈爆炸式增长

二月公布的贸易数据令墨西哥和美国的官员欢欣鼓舞。数据显示，墨西哥在2023年超过中国，成为对美最大出口国（见图表）。所销售商品的价值稳步上升，全年达到近4760亿美元；而中国商品的相应数字则大幅下降，从2022年的5360亿美元降至2023年的4270亿美元。

墨西哥上一次超过中国是在2002年，当时只是昙花一现（占据榜首的是加拿大）。如今，墨西哥的崛起露出一丝全球贸易新秩序的苗头。2024年1月的数据显示，这一趋势依然强劲。美国正加紧努力与中国脱钩，使供应链更靠近本土，推动贸易通过其他国家进行。美国、墨西哥和加拿大之间新修订的《美墨加协定》（USMCA）已于2020年生效。该协定支持墨西哥向美国出口汽车零部件、医疗用品和农产品。

但是，这个贸易三角的第三条边——由墨西哥和中国构成——正在制造紧张拉力。中国企业受到激励要放眼海外，它们没有闲着。它们一直在向美国希望拿来取代中国的那些市场推进，包括墨西哥。智库墨西哥竞争力研究所（IMCO）的安娜·古铁雷斯（Ana Gutiérrez）说，这意味着许多中国出口商品只不过是“多走了一点点路，到达的还是同一个地方”。

中国似乎正在提倡这种战略。去年12月，中国领导层表示，要优先出口用于制造成品的产品，而不是成品本身。墨西哥是进入美国的一个有吸引力的滩头堡，因为《美墨加协定》为含有足够北美成分的产品提供免关税准入。

墨西哥的官方海关数据没有显示中国货物持续流入。但一些美国官员和实业家认为，无论有意无意，中国货物的流入量被少计了。他们怀疑，墨西哥对从中国进口的商品睁一只眼闭一只眼，而这些商品随后转口到了美国。

钢铁和铝是关注的焦点。2月，美国贸易代表戴琪指出，“墨西哥从第三国进口的钢铁和铝缺乏透明度”。去年12月，墨西哥对从中国进口的部分钢材征收最高80%的关税，但美国官员仍然沮丧。制定关税水平是一回事，实际执行又是另一回事。

代表制造商的团体“繁荣美国联盟”（Coalition for a Prosperous America）的杰夫·费里（Jeff Ferry）说：“我们看到的是，《美墨加协定》实际上已经成了《美墨中协定》，中国正通过墨西哥转运大量产品。”

电动汽车是一个日益引人关切的问题。中国普通电动汽车的价格大约是美国的一半，而中国产电动汽车占全球产量的一半以上。如果不征收高额关税，中国电动汽车在美国可能会像在其他国家一样热销。拜登政府正在酝酿对中国电动汽车加征关税，使之高于目前25%的水平。

《美墨加协定》有针对不公平补贴和不公平市场行为的条规，这样的补贴和操作在中国公司中很常见。但在许多案例中并没有哪条法律或规则被违反。墨西哥为中国汽车制造商提供了绕过关税墙的途径，是因为该协定的原产地规则存在一位美国官员所说的“漏洞”，让整合中国零部件的操作钻了空子。进口商可以在墨西哥组装中国零部件，并将墨西哥标为原产国，从而掩盖中国参与其中的痕迹。“可能需要新的应对办法了。”这位官员说。

墨西哥已经有了一个庞大的汽车制造业，而中国的资金也在不断涌入，特别是流向对美国出口量最大的北部各州。2月，中国最成功的电动汽车制造商比亚迪表示将在墨西哥年产15万辆汽车。比亚迪表示这部分生产将服务当地市场，但许多公司瞄准的都是边境以北那个更大、更富裕的市场。

2022年，中国在墨西哥的外国直接投资达到了25亿美元。华盛顿智库美洲对话（Inter-American Dialogue）的玛格丽特·迈尔斯（Margaret Myers）指出，中国公司在墨西哥的尖端制造业取得了“显著增长”。

｜ 产业链中没中国

美国对此可能会如何应对？在钢铁和铝进口方面，美国或许会复制对欧盟

和日本的做法，限制以较低关税税率进口的数量。达到这些上限后，就会征收更高的关税。特朗普政府曾在2019年取消对墨西哥的钢铝关税，其条件就是墨西哥克制对美钢铝出口激增。

电动汽车的问题要更棘手。一些人希望制定新规则，将中国完全排除在供应链之外。费里说：“如果你想当美国的贸易伙伴，那我们是不允许你的国家充当中国商品中转站的。”这就需要出台限制性的原产地规则和加大规则执行力度。这也会引发如何对待墨西哥中资工厂产品的麻烦问题。

曾在特朗普任总统期间担任首席贸易谈判代表的罗伯特·莱特希泽（Robert Lighthizer）说过，第一步将是剥夺中国的最惠国待遇。这将自动全面提高对中国产品的关税。位于墨西哥城的墨西哥国立自治大学（UNAM）的恩里克·杜塞尔（Enrique Dussel）说，这会松动世界贸易体系。“美国相当于在说，‘采用我的规则，否则就是跟我对着干’。”

如果特朗普在11月胜选，他很可能会对墨西哥采取更强硬的立场。除了中国贸易问题，美国与墨西哥还存在移民问题。但最重要的或许还是特朗普厌恶贸易失衡。2023年，美国与墨西哥的贸易逆差上升到1520亿美元，比2022年增长了17%。2026年，两国加上加拿大必须讨论是否将《美墨加协定》的有效期从2036年再延长16年至2052年。特朗普签署了该协定，但这并不能保证他不会废除它，或利用延期为筹码迫使墨西哥做出让步。他已经谈到要对来自所有国家的商品征收10%的进口关税，而《美墨加协定》的存在令他不可能对墨加两国做到这一点。

杜塞尔说，墨西哥似乎并无准备。预计将在6月2日的选举中当选墨西哥下一任总统的克劳迪娅·希恩鲍姆（Claudia Sheinbaum）正在推动“近岸外包”，要将墨西哥的年经济增长提高到GDP的5%。杜塞尔说：“墨西哥亮出了一块写着‘欢迎中国’的巨大招牌。”

墨西哥面临的问题是，为了让这块欢迎招牌闪闪发亮，它愿意拿自己与美国的这份重要关系冒险到何种程度。■



Pumped up

America's extraordinary economy keeps defying the pessimists

Both Donald Trump and Joe Biden have ideas that endanger it

YOU HAVE to marvel at America's economy. Not long ago it was widely thought to be on the brink of recession. Instead it ended 2023 nearly 3% larger than 12 months earlier, having enjoyed one of the boomier years of the century so far. And it continues to defy expectations. At the start of this year, economists had been forecasting annualised growth in the first quarter of 1%; that prediction has since doubled. The labour market is in rude health, too. The unemployment rate has been below 4% for 25 consecutive months, the longest such spell in over 50 years. No wonder Uncle Sam is putting the rest of the world to shame. Since the end of 2019 the economy has grown by nearly 8% in real terms, more than twice as fast as the euro zone's and ten times as quickly as Japan's. Britain's has barely grown at all.

America's expansion is all the more striking when you consider the many things that could have killed it. As the Federal Reserve has fought inflation the economy has endured the sharpest rise in interest rates since Jimmy Carter was in the White House. The covid-19 pandemic, an intensifying trade war with China and the fight against climate change have together reshaped supply chains, labour markets and consumer preferences. Wars in Ukraine and Gaza have aggravated geopolitical tensions and worsened the strains on the global trading system.

Can America's remarkable strength persist? Threats to growth still hang over the economy. The longer interest rates stay high, for instance, the more damage they could do. Although inflation has fallen, it remains sticky above the Fed's 2% target, meaning that the Fed may be unable to fulfil investors' hopes for interest-rate cuts starting in June. Geopolitical

tensions, meanwhile, look likely to spur economic fragmentation. Yet the biggest threat of all stems from November's presidential election. Neither Joe Biden nor Donald Trump seems likely to nurture the economic expansion should they return to the White House. Instead, their plans would endanger it.

To understand this, consider the reasons for the economy's extraordinary performance. A key plank was generous pandemic stimulus, which at 26% of GDP was more than double the rich-world average. This largesse fuelled inflation but also ensured fast growth: consumers have yet to spend all the cash they received in "stimmy" cheques. Even as the covid crisis passed, the government continued to borrow away. The underlying deficit over the past year was nearly 8% of GDP. That supported demand even as rates went up.

Strong demand has been met by growing supply. America has 4% more workers than it did at the end of 2019, thanks in part to rising workforce participation, but mainly owing to higher immigration. The foreign-born population is up by 4.4m, a figure which may undercount those who arrived illegally. And the expanding workforce is being put to productive use. America's flexible labour market has almost certainly made it easier for the economy to adapt fast to a changing world.

Other long-standing strengths have made America enviably placed to cope with geopolitical tumult. Its vast internal market encourages innovation and means it depends less on foreign trade than smaller rich economies do. Because the shale boom of the 2010s made America a net energy exporter, it has in aggregate benefited rather than suffered from the high energy prices that hit the wallets of Europeans.

The trouble is that each of the ingredients for growth can no longer be relied upon. It may be tempting for politicians to extrapolate from America's recent success and juice the economy with further stimulus. But

that is becoming unsustainable. Official forecasts show that America will this year spend more on debt interest than national defence. More borrowing risks building up financial perils in the future.

At the same time, both Mr Trump and Mr Biden harbour populist and protectionist instincts that will only harm America's growth potential. The sugar-rush of stimulus helped mask the damaging effects of such policies during each president's first term. This time, however, the damage will not be disguised.

Mr Trump poses the graver threat. He has entertained a blanket 10% tariff on imports, which some of his advisers see as a mere starting-point. That would triple America's existing levies on goods imports, worsen inflation and raise the cost of imported parts for manufacturers. At the same time, Mr Trump has promised the mass deportation of illegal immigrants. This goes well beyond trying to secure the border against new unauthorised entrants—a reasonable goal—and poses a risk to the labour supply: in 2021 America's 10.5m irregular migrants made up an estimated 5% of its workforce.

A Trump presidency would also threaten the Fed's independence. Mr Trump says he would not reappoint its chairman, Jerome Powell, whose term expires in 2026, and whom he often criticised when in office. A pliant Fed at a time of big deficits—which Mr Trump might increase with more tax cuts—could threaten America's macroeconomic stability.

A second Biden term promises nothing as potentially catastrophic. Mr Biden has let the Fed get on with fighting inflation, and wants to raise taxes to reduce deficits. Yet he is also an economic nationalist. His State of the Union address on March 7th was littered with protectionist promises that the government would "Buy American", and statist ideas about controlling the price of everything from housing to Snickers bars.

| *Muscle memory*

Both Mr Biden and Mr Trump harbour a misplaced nostalgia for the 1950s and '60s, and justify their policies by painting today's economy as weaker than it is. Mr Trump thinks trade and immigration have made the country poorer; Mr Biden is deeply distrustful of big business. And where they do acknowledge America's strengths, both men attribute it to their own misguided interventionism. In fact, they are chipping away at the free markets which are the ultimate source of the country's wealth.

The truth is America has thrived as its companies and workers have innovated and adapted to a rapidly changing world. If the next president does not recognise that, America's pumped-up economy will eventually deflate. ■



【首文】肌肉膨胀

美国经济的非凡表现不断打破悲观预期

特朗普和拜登的计划都可能破坏这一势头

你不得不惊叹美国经济的表现。不久前，人们普遍认为美国经济濒临衰退。但相反，在2023年结束之际，美国经济总量较12个月前增长近3%，成为本世纪较繁荣的年份之一。美国经济还在打破预期，今年年初，经济学家预测第一季度的年化增长率为1%，现在预测已翻了一番。就业市场同样强劲，失业率已连续25个月低于4%，创下50多年来最长纪录。难怪美国让世界其他国家相形见绌。自2019年底以来，美国经济实际增长近8%，是欧元区的两倍多，是日本的十倍。而英国几乎没有增长。

考虑它面对的诸多不利因素，美国的经济增长更显惊人。为对抗通胀，美联储实施了自卡特执政时期以来最大力度的加息。新冠疫情、与中国日益激烈的贸易战，以及应对气候变化的努力共同重塑了供应链、劳动力市场和消费者偏好。乌克兰和加沙的战争加剧了地缘紧张态势，给全球贸易体系带来了更大压力。

美国能否维持这非凡的强韧？威胁增长的因素依然挥之不去。例如，高利率维持的时间越久，可能造成的损害就越大。尽管通胀已有所下降，但依旧高于美联储2%的目标，这意味着美联储可能无法像投资者预期的那样在6月开始降息。与此同时，地缘紧张局势很可能刺激经济碎片化。然而，最大的威胁源自11月的总统大选。不管是拜登还是特朗普执掌白宫，似乎都不会促进美国经济增长。相反，他们的计划可能危及美国经济。

要理解这一点，来看看美国经济表现出众的原因。一大关键是美国在疫情下推出的慷慨刺激政策，规模达到GDP的26%，是富裕国家平均水平的两倍多。如此大手笔的支出推高了通胀，但也确保了快速增长：消费者到现在还没有花光从“纾困”支票里收到的钱。即使疫情危机已过，美国政府继续举债。过去一年，美国的基础赤字占GDP之比接近8%，这使得美国在持续加息之时，需求仍能得到支撑。

需求强劲又恰逢供给增加。相比2019年底，美国就业人口增加了4%，这里有劳动参与率上升的因素，不过主要原因还是移民增加。外国出生人口增加了440万，这可能还没完全计入非法移民的数量。不断扩大的劳动人口正在进入经济生产。美国灵活的劳动力市场无疑让经济能更好地快速适应这个变化中的世界。

其他长期优势也让美国在应对地缘动荡时处于令人艳羡的有利位置。美国广阔的国内市场鼓励创新，使得它相比较小的富裕经济体更少依赖对外贸易。2010年代的页岩油革命使美国成为净能源出口国，总体而言这让美国成为能源价格高涨的受益者，而不是像欧洲人那样受到冲击。

但问题是，这些推动增长的因素没有哪个可被继续依赖。政客们可能会根据美国近期的成功得出经验，再推出刺激政策来提振经济。但这已变得不可持续。官方预测显示，今年美国在债务利息上的支出将超过国防开支。进一步举债可能推高未来的金融风险。

与此同时，特朗普和拜登都带有民粹主义和保护主义倾向，而这只会危害美国的增长潜力。在这两位总统的首个任期内，刺激措施的即时强效提振都帮助掩盖了这些政策的破坏性影响。然而这一次，这种破坏将无从隐身。

两人当中，特朗普的破坏力更大。他提出对进口商品全面征收10%的关税，他的一些顾问认为这只是个起点。这将使美国的商品进口关税升至原来的三倍，加剧通胀，推高部分制造商进口零部件的成本。同时，特朗普承诺大规模驱逐非法移民。相比保护边境避免更多人非法入境这一合理目标，这种做法非常过火，可能危及劳动力供给：2021年，美国有1050万非法移民，约占其劳动人口的5%。

特朗普当选总统也可能威胁美联储的独立性。他表示不会再次任命自己在任时经常批评的鲍威尔（任期将于2026年结束）继续担任美联储主席。目前赤字高企，而特朗普还可能会通过更多减税措施来增加赤字，如果这时美联储对白宫言听计从，可能会威胁美国宏观经济的稳定。

如果拜登连任，应该不会带来灾难性后果。拜登已放手让美联储继续打击通胀，同时打算增税来降低赤字。然而他也是个经济民族主义分子。他于3月7日发表的国情咨文充斥着政府“买美国货”的保护主义承诺，以及要控制从住房到士力架等一切商品的价格的国家主义思想。

丨 肌肉记忆

拜登和特朗普都对上世纪五六十年代有一种不合时宜的怀念，并把当前经济描绘得比实际更疲弱，以此为自己的政策辩护。特朗普认为贸易和移民让美国更穷；拜登对大公司深怀疑虑。而在确实承认美国拥有的优势时，两人又都归功于自己那些错谬的干预主义政策。实际上，他们正在削弱自由市场这一美国财富的终极源泉。

事实是，美国经济欣欣向荣，是因为美国公司和工人不断创新并积极调整来适应这急速变化的世界。如果下一任总统没有意识到这一点，美国膨胀的经济终将泄气。■



And you call yourself civilised?

The history of the West is not quite what you learned in school

Josephine Quinn's new book re-examines what people think they know about civilisations

How the World Made the West. By Josephine Quinn. Random House; 592 pages; \$38. Bloomsbury; £30

ASKED WHAT he thought of Western civilisation, Mahatma Gandhi is said to have quipped that such a thing “would be a good idea”. (The West, he suggested, was not so enlightened.) But as Josephine Quinn makes clear in her new book, Western civilisation has always been a bad idea, or at any rate a wrong-headed one. To compartmentalise history into a set of distinct and essentially self-contained civilisations is a misguided quest that has dangerously distorted our understanding of the world, Ms Quinn asserts: “It is not peoples that make history, but people, and the connections that they create with one another.”

Ms Quinn, a historian and archaeologist who teaches at Oxford, does not spend 500-odd pages trashing what generations of schoolchildren have been taught to take pride in as European achievements. Instead, she demolishes the underlying concept of what she calls “civilisational thinking”. Her argument is simple, persuasive and deserving of attention.

The idea of civilisation, Ms Quinn points out, is relatively recent. The word was first used only in the mid-18th century and did not take hold of Western imaginations until the late 19th century. In that imperialist age, historians found that Greek, Roman and Christian civilisations made nice building blocks that could be stacked into a grand-looking construct, which they labelled “Western” or “European” civilisation. To this they attributed a host of inherited “classical” virtues: vigour, rationality, justice, democracy

and courage to experiment and explore. Other civilisations, by contrast, were regarded as inferior.

It does not take much unpacking by Ms Quinn to expose the folly of this approach. Behold, for instance, John Stuart Mill, a philosopher in the 19th century, claiming that the Battle of Marathon, Persia's first invasion of Greece in 490BC, was more important to English history than William the Conqueror's triumph at Hastings in 1066. (Without an Athenian victory, the logic goes, the magical seed of Greek civilisation might never have grown into Western civilisation.) Or consider "The Clash of Civilisations" (1996) by Samuel Huntington, an American historian, who declared it impossible to understand history without classifying humanity into mutually hostile civilisations between which, "during most of human existence", contact has been "intermittent or non-existent".

What is non-existent is any truth to that notion. Ms Quinn's brisk, scholarly romp across the arc of European history shows that, far from being rare, contact across and between cultures, often over surprisingly long distances, has been the main motor of human advancement in every age. Rather than being prickly and inward-looking, most societies have proved receptive to ideas, fashions and technologies from their neighbours.

Ancient Greece, for example, was less a place of origins than of transmission from Egyptian, Sumerian, Assyrian and Phoenician cultures, which themselves had mixed and exchanged ideas. And rather than being the wellspring of democracy, Athens was "something of a latecomer" to a form of governance that appears to have been first tried in Libya and on the islands of Samos and Chios. Persians, eternally cast as Greeks' polar opposites, actually imposed democracy on the Greek cities they ruled, suggesting "considerable Persian faith in popular support for their own hegemony", Ms Quinn notes.

This retelling of the West's story scintillates with its focus on the unexpected and on the interstices between realms and eras rather than on history's big, solid bits. But it is also an admirable work of scholarship. Ms Quinn's 100-plus pages of footnotes reveal that she relied not only on a wide range of primary sources, but also on scientific studies on climate change and very recent archaeological research.

Even seasoned history buffs will find much that is new and fascinating. "How the World Made the West" joins a growing sub-canon of works that explores the broad sweep of history using new intellectual framings, such as Yuval Noah Harari's "Sapiens" (2011), Peter Frankopan's "The Silk Roads" (2015) and "Fall of Civilisations", a forthcoming book by Paul Cooper, a British journalist, based on his popular podcast. Anyone who thought history was passé could not be more wrong. ■



你还自称文明人？

西方历史跟你从课堂上学到的不太一样

约瑟芬·奎恩的新书重新审视了人们对文明的认知【《世界如何缔造了西方》书评】

《世界如何缔造了西方》，约瑟芬·奎恩著。兰登书屋，592页，38美元；布鲁姆斯伯里出版社，30英镑。

据说圣雄甘地在被问及如何看待西方文明时，曾经打趣道，这东西真要有，“倒也不赖”。（在他看来，西方并没有那么文明。）不过，正如约瑟芬·奎恩（Josephine Quinn）在她的新书中阐述的那样，“西方文明”这个概念从来就不好，或至少也是有误的。试图将历史划分为一系列各不相同、本质上自成一体文明是一种误入歧途的做法，它扭曲了我们对世界的理解，这很危险。奎恩主张：“历史的创造者不是各种人群，而是所有人，以及他们彼此间建立的联系。”

奎恩是任教于牛津大学的历史学家和考古学家。她没有花500多页的篇幅驳斥一代代学童被教导要以之为荣的“欧洲成就”，而是推翻了她称之为“文明思维”的底层概念。她的论据简明有力，值得关注。

奎恩指出，“文明”是一个相对较新的概念。这个词在18世纪中期才被首次使用，直到19世纪后期才引起了西方人的兴趣。在那个帝国主义时代，历史学家发现，希腊、罗马和基督教文明是一块块好砖，可以搬来砌成一座他们标榜为“西方”或“欧洲”文明的宏伟建筑。他们为这种文明披挂上一系列承继下来的“传统”美德：活力、理性、正义、民主，以及尝试和探索的勇气。相比之下，其他文明则被视为劣等文明。

无需过多拆解，奎恩就揭示了这种观点的荒谬性。比如，19世纪哲学家约翰·密尔曾声称，论对英国历史的重要性，公元前490年波斯首次入侵希腊的马拉松战役要胜过征服者威廉在1066年黑斯廷斯战役中取得的胜利。

（这里面的逻辑是，如果没有雅典人的胜利，希腊文明的神奇种子可能永远不会成长为西方文明。）再比如，美国历史学家塞缪尔·亨廷顿

（Samuel Huntington）在其1996年出版的《文明的冲突》（The Clash of civilizations）中宣称，如果不把人类划分为相互敌对的文明，就不可能理解历史，“在人类存在的大部分时间里”，各种文明之间的交往“断断续续或者根本不存在”。

这种观点根本站不住脚。透过奎恩对欧洲历史脉络专业而利落的梳理，我们可以看到，不同文化之间的交往一点也不罕见，甚至常常跨越了难以置信的遥远距离，成为每个时代人类进步的主要动力。事实证明，大多数社会都愿意接受来自邻居的思想、时尚和技术，而不是抵触和自我封闭。

例如，古希腊与其说是西方文明的发源地，不如说是埃及、苏美尔、亚述和腓尼基文化的传播地，而这些文化本身也经历了各种思想的融合与交流。雅典与其说是民主的源泉，不如说有点像民主的“后进者”——这种治理方式似乎最早试行于利比亚，以及希腊的萨摩斯岛和希俄斯岛等地。总是被视为与希腊人截然不同的波斯人，事实上却在当年统治的希腊城邦中强制推行了民主，这表明“波斯人对民众支持自己的统治地位相当有信心”，奎恩指出。

这本书对西方历史的重述让人眼前一亮——它所关注的是人们意想不到的东西，以及不同领域和不同时代之间的断层，而不是历史上那些重大而完整的阶段。但它同时也是一部令人钦佩的学术著作。书中有100多页附有脚注，从中可以看出奎恩的论据不仅来自各种一手资料，还来自有关气候变化的科研成果和最新的考古研究。

即使是见多识广的历史爱好者也会从中发现很多让人耳目一新、饶有兴味的东西。《世界如何缔造了西方》是又一部用新思想框架探索浩瀚历史的另类力作，这个内容不断丰富的类别包括尤瓦尔·诺亚·赫拉利（Yuval Noah Harari）的《智人》（Sapiens，2011年）、彼得·弗兰科潘（Peter Frankopan）的《丝绸之路》（The Silk Roads，2015年），以及英国记者保罗·库珀（Paul Cooper）根据自己的热门播客改编的新书《文明的衰落》（Fall of Civilisations，即将出版）等等。那些认为历史学已经过时的人真是大错特错了。■



Bartleby

How can firms pass on tacit knowledge?

The problem of knowing what your co-workers know

LAST MONTH Odysseus became the first American spacecraft to land on the surface of the Moon in more than 50 years. The mission, a collaboration between NASA and a private firm called Intuitive Machines, can be counted a partial success: the craft did send back images even though its landing did not go to plan. Things might have gone better still if it had not been so long since NASA last visited the Moon. Experience usually makes things go more smoothly.

NASA does have an archive of materials from the Apollo missions. Sometimes, however, knowledge is lost for good. Gino Cattani of NYU Stern School of Business and his co-authors have looked at the violin-making family dynasties of Antonio Stradivari and others in Cremona, in Italy, in the 17th and 18th centuries. Modern players still laud the sound of the instruments made by these craftsmen. But there was a gap of about a century between the heyday of these dynasties and the rise of the public performances that showcased the instruments' qualities. In that time the techniques of the Cremonese luthiers were lost.

Most organisations do not routinely blast into orbit or wait a century for customer feedback. But all organisations face the problem of storing and transferring knowledge so that newcomers know what's what, lessons are learned from successes and failures, and wheels are not constantly being reinvented. An ageing workforce adds to the urgency of training inexperienced hires before the old hands leave the building.

Some knowledge is easier to codify than other. In the 1960s Corning, a glassmaker, had developed a particularly strong glass that was christened

Chemcor. Plans to commercialise this material faltered—among other reasons, it turned out that this was not a great windscreen for motorists to hit at speed—and Chemcor was put on the shelf.

There it remained until 2005, when the firm started to wonder whether mobile phones might provide a use for Chemcor, which was renamed Gorilla Glass. In 2007 the boss of Corning took a call from Steve Jobs, who was hunting for the right kind of glass for a new smartphone. You presume that no one at Corning has since questioned the value of keeping good records.

The tougher task is capturing “tacit knowledge”. This is the know-how born of experience, which cannot easily be documented in the manuals and is not much thought about by those who have it.

Working alongside experienced colleagues is the best way to transfer tacit knowledge but it is not always possible. Sometimes you only want your very best people working on something, especially if the stakes are high. The most valuable employees are usually the ones with the least time to mentor others. When NASA was working on a Mars rover programme in the 2010s, it gave younger engineers a smaller, parallel project: to build a rover for use in educational programmes on Earth. It wasn’t the real thing, but it was a way to give them some hands-on experience. NASA also has an emeritus programme that gets retired veterans to mentor junior staff.

Technology is both an answer and a barrier to the transfer of tacit knowledge. It is easier than ever to record and disseminate the wisdom of older hands. Unfortunately, it is easier than ever to record and disseminate the wisdom of older hands: the podcast episodes proliferate, the hours of unwatched training videos pile up. Watching someone on a screen is often less stimulating than hearing from them face-to-face. A recent study by Niina Nurmi and Satu Pakarinen, two Finnish researchers, found that

participants in virtual meetings feel drowsier than those meeting in person, which is saying something.

Christopher Myers of Johns Hopkins University is a fan of informal storytelling as a way of passing on tacit knowledge. He spent time with the crews on an air medical transport team in America, whose jobs include flying patients by helicopter from the scene of an emergency to a hospital. Crew members routinely shared stories—on shift changes, at mealtimes and at weekly meetings—in order to learn how to respond to unusual situations. (Top tip: in the event of a poisonous-snake bite, the local zoo is a good bet to get antivenom.)

Some stories are more gripping than others: people don't sit round camp fires telling each other how to get the printer to work. But managers everywhere should think about how to capture tacit knowledge. That starts by recognising the importance of retaining workers. You can't share experience if no one has any. ■



巴托比

企业如何传承隐性知识？

知道你的同事都知道些什么是个难题

上个月，奥德修斯号（Odysseus）成为50多年来第一艘登陆月球表面的美国航天器。这次任务由美国国家航空航天局（NASA）和一家名为“直觉机器”（Intuitive Machines）的私营公司合作完成，可以说取得了部分成功：尽管飞船的着陆没有按计划进行，但它确实发回了图像。要不是距离NASA上次访问月球的时间实在隔得太久，结果说不定还会更好。经验通常会让事情进行得更加顺利。

NASA确实有阿波罗计划的资料档案。然而有些时候，知识永久地丢失了。纽约大学斯特恩商学院的吉诺·卡塔尼（Gino Cattani）及合著者研究了17和18世纪意大利克雷莫纳的安东尼奥·斯特拉迪瓦里（Antonio Stradivari）等人的小提琴制造家族王朝。现代的演奏家们仍对出自这些工匠之手的乐器的音色赞不绝口。但是，从这些家族的鼎盛时期到一展乐器品质的公开表演兴起，中间相隔了大约一个世纪。在那期间，克雷莫纳制琴师的技艺失传了。

大多数组织并不会定期将产品送入轨道或花一百年等客户的反馈，但所有组织都面临知识存储和传递的问题。有了存储和传承，新人才能知道什么是，才能从成功和失败中吸取经验教训，才不会一直浪费时间从头做起。鉴于员工队伍的老龄化，在老员工离职前培训好经验欠缺的员工愈发紧迫。

有些知识比其他知识更容易整理归档。20世纪60年代，玻璃制造商康宁（Corning）研发出一种特别坚固的玻璃，并给它取名Chemcor。这种材料的商业化计划失败了，原因之一它用作挡风玻璃时耐不住驾车者在高速行驶时发生撞击，于是Chemcor被束之高阁。

它就这么被搁置，直到2005年。康宁开始琢磨手机也许能用得上

Chemcor，它也有了个新名字：大猩猩玻璃。2007年，康宁老板接到了史蒂夫·乔布斯的电话。当时乔布斯正在给一款新的智能手机寻找合适的玻璃。可以想见，自那以后在康宁没人会质疑认真记录留档的价值。

更艰巨的任务是捕捉“隐性知识”。这类专门技能源于经验，不容易记录在册，掌握这些知识的人也不怎么思考琢磨它们。

与经验丰富的同事并肩工作是传授隐性知识的最佳途径，但这并非总是可行。有时，你只想把某项工作交给最优秀的人去做，尤其是在事关重大的情况下。最有价值的员工通常也最没有工夫指导他人。2010年代NASA在开展火星探测器计划时，曾给年轻的工程师们安排了一个规模更小的并行项目：建造一个探测器，用于在地球上的科教项目。这不是真的火星探测器，但也是个让他们上手获得些许实战经验的办法。NASA还设有荣誉退休职位，请已退休的老将来指导初级工作人员。

在如何传递隐性知识的难题上，技术既是解决方案，也是障碍。记录和传播前辈的智慧从未如此容易。不幸的是，记录和传播前辈的智慧从未如此容易：播客数量激增，未观看的培训视频迅速累积。在屏幕上看一个人往往不如面对面听一个人说话那么能启迪人心。两名芬兰研究人员尼娜·努尔米（Niina Nurmi）和萨图·帕卡里宁（Satu Pakarinen）最近的一项研究发现，参加虚拟会议的人比亲身参加会议的人更容易昏昏欲睡，这说明了一些问题。

约翰·霍普金斯大学的克里斯托弗·迈尔斯（Christopher Myers）推崇的一种传承隐性知识的方法是轻松随意地讲故事。他曾与美国一支空中医疗运输队的机组人员共处过一段时间，他们的工作包括用直升机将病人从急救现场送往医院。机组人员经常在换班、用餐和每周例会时分享故事，从而学习如何应对不寻常的情况（重要小提示：如果被毒蛇咬伤，可以去当地动物园注射抗毒血清）。

有些故事比别的故事更扣人心弦，比如人们就不会围坐在篝火旁交流如何使用打印机。但天底下的管理者都应该思考如何留存隐性知识。首先就要

认识到留住员工的重要性。如果员工都是“零经验”，还谈何分享。 ■



Back to the moon

Bitcoin's price is surging. What happens next?

The cryptocurrency is up by 63% this year

FOR A BRIEF moment, everyone who owned bitcoin had made money from it. On March 5th the crypto token rose to an all-time high of just above \$69,000—a level sure to delight the meme-loving crypto-crowd—before slipping back a little. The record capped a remarkable comeback from the dark days of November 2022, when interest-rate rises were crushing risk appetite and FTX, a crypto exchange, had just gone bust. Buying bitcoin on such exchanges seemed like little more than a fun and novel way to get robbed.

Bitcoin is hardly rallying in isolation: everything is going up. Stockmarkets all over the world are near record highs. So are gold prices. Even bond prices are climbing after a miserable two-year stretch. The catalyst is a combination of hype about artificial intelligence, joy at the state of the global economy and expectations of looser monetary policy to come.

Still, bitcoin is doing better than most assets. On January 10th the Securities and Exchange Commission, an American regulator, approved applications by 11 investment firms, including BlackRock and Fidelity, to create bitcoin exchange-traded funds (ETFs). These make it easier for everyday investors to buy the cryptocurrency. Rather than setting up an account with a specialist exchange, creating a crypto wallet, making a bank transfer and then finally buying bitcoin, people can now simply log on to their brokerage accounts and purchase an ETF. Assets in the ten largest bitcoin ETFs now come to around \$50bn. And the activity appears to be self-reinforcing: the more money is poured in, the higher the price goes, the more people chatter about bitcoin ETFs, the more money pours in and so on and so on.

Bitcoin has been in existence for 14 years. The elegant mechanism by which it validates itself and supply grows has never been hacked, meaning that the token is not going anywhere. Yet it is now obvious that it is of pretty limited use for payments, as it is restricted by both the high costs and slow speed of transactions. Those trying to build applications on top of blockchains are not doing so using bitcoin either. With the creation of ETFs, bitcoin's future looks to be as an investment asset and nothing more. So after this initial surge of interest, what will its returns look like?

It would be foolish to extrapolate from bitcoin's entire history. Over the past 14 years the cryptocurrency has morphed from a niche cyberpunk idea into something approaching a mainstream financial asset. Its more recent price movements might provide some clues, however. There are two explanations for them. One is that purchases are basically a broad bet on technological progress, with variations that reflect prospects for crypto itself. For instance, even as tech stocks soared in the middle of 2021, bitcoin slumped after Elon Musk posted negative tweets about crypto payments. Prices were depressed in late 2022, too, even as stockmarkets were rallying, owing to FTX's failure.

The other theory is that bitcoin is a kind of digital gold. After all, supply is inherently limited, just as gold supply is restricted by the amount of the metal in the ground. Neither asset pays a yield nor earns profits. This theory fell out of favour in 2021 and 2022, as inflation soared and bitcoin collapsed, but last year the cryptocurrency once again moved in line with gold.

Perhaps both theories contain elements of truth. And a hybrid tech-stock-crypto-vibes-gold-bet asset could be useful in even pedestrian portfolios, especially if it is only somewhat correlated with other assets an investor might hold. Diversification among uncorrelated assets is the foundational principle of portfolio management. Reallocating, say, 1% of a fund to

bitcoin would be a low-stakes hedge.

If investors buy this argument, bitcoin's price is likely to rise for a while yet. What happens, then, when the cryptocurrency's transition into a standard financial asset is complete? Assume that bitcoin has been added to most investor portfolios. Also assume that crypto tech does not really catch on. In this world, bitcoin's returns probably do come to resemble those of gold: there is a fixed amount of it, and its price would rise over the long term roughly in line with the stock of money. That implies steady single-digit returns. The creation of a bitcoin ETF may have set off a frenzy of eye-popping gains—but the future it portends could be slower and steadier. ■



再回高位

比特币价格飙升。接下来会发生什么？

今年比特币上涨了63%

有那么一小会儿，每个持有比特币的人都从中赚到了钱。3月5日，比特币一度涨破6.9万美元，创历史新高，之后略有回落。这一高点自然让喜欢玩梗的加密爱好者感到欣喜。该纪录标志着比特币在2022年11月那段黑暗时期后强势回归，当时利率上涨压制了风险偏好，加密交易所FTX刚刚破产，在这些交易所购买比特币似乎不过是以一种新颖有趣的方式被割韭菜。

上涨的不只有比特币，一切都在涨。世界各地的股市都接近创纪录高位。黄金价格也是如此。即使是惨淡了两年的债券价格也在涨。对人工智能的炒作、对全球经济现状的欣喜之情，以及对未来货币政策更加宽松的预期结合在一起，催化了这一波行情。

然而，比特币的表现要好于大多数资产。1月10日，美国监管机构证券交易委员会（SEC）批准了包括贝莱德（BlackRock）和富达（Fidelity）在内的11家投资公司的申请，允许它们创建比特币交易所交易基金

（ETF）。这让普通投资者更容易购买这种加密货币。以前，人们需要在专门的交易所开设账户、创建加密钱包、进行银行转账，然后才能购买比特币，现在他们只需登录证券账户就可以购买这些ETF。目前，前十大比特币ETF的资产总额约达到500亿美元。这种投资方式似乎形成了一个自我强化的循环：越多资金涌入，价格就越高，就有越多人关注比特币ETF，继而吸引越多资金涌入，周而复始。

比特币已经存在了14年。它验证自身并增加供应的精妙机制从未被攻破，意味着这种代币还会这样存在下去。然而，现在显而易见的是，由于受交易成本高、速度慢的限制，它在支付方面的用途相当有限。那些试图在区块链上构建应用的人也不使用比特币。随着ETF的创建，未来比特币除了是一种投资资产外似乎没有更多用途。因此，当刚开始的兴趣高涨过后，

它的回报将会如何呢？

试图从比特币的整个历史去推断是荒诞的。在过去的14年里，这种加密货币已经从一个小众的赛博朋克理念变成了一种接近主流金融资产的东西。然而，它近期的价格波动也许能提供一些线索。关于这些波动有两种解释。一种是，买进比特币基本上反应的是整体上对技术进步的押注，其变化反映了加密货币本身的前景。例如，2021年年中，在马斯克发布了有关加密货币支付的负面推文后，尽管科技股不断飙升，比特币仍然大幅下跌。2022年底，尽管股市在上涨，但由于FTX破产，比特币价格依然低迷。

另一种理论是，比特币是一种数字黄金。毕竟，比特币的供应在本质上是有限的，就像黄金供应受到储量的限制一样。这两种资产都不产生收益，也不赚取利润。然而在2021年和2022年，由于通胀飙升和比特币崩溃，这种理论不再流行，但去年比特币的走势再次与黄金一致。

也许两种理论都有一定道理。而一种兼具科技股、加密货币、黄金投资等特点的资产即使在普通投资组合中都很有用，如果它与投资者可能持有的其他资产只在一定程度上有关联的话就更有用了。在不相关的资产中分散投资是投资组合管理的基本原则。比如，将一只基金中的1%重新配置到比特币会是一种低风险对冲。

如果投资者认同这种观点，比特币的价格很可能还会再涨一阵子。那么，到比特币完成向标准金融资产的转变时会怎样？假设比特币已被添加到大多数投资者的投资组合中，再假设加密技术并没有真正流行起来，那么在这种情况下，比特币的回报可能确实将与黄金相似：其总数是固定的，而其价格涨势从长远看将大致与货币存量一致。这意味着回报将是稳定的个位数。比特币ETF的诞生可能引发了令人瞠目结舌的疯狂上涨，但它也预示未来的上涨趋势可能会更慢、更稳定。■



Advanced materials

How medical gloves will help launch satellites

Graphene rises to new heights

GRAPHENE IS strong, lightweight, flexible and an excellent conductor of electricity. In the 20 years since it was first isolated at the University of Manchester, however, it has also proved dispiritingly light in useful applications. That is slowly beginning to change, as its remarkable properties keep researchers well-stocked with inspiration. For Krzysztof Koziol at Cranfield University in Britain, for example, what began as a covid-era plan to use graphene to improve surgical gloves has now morphed into a project to use high-altitude balloons to launch satellites into space.

Graphene, which consists of monolayers of carbon atoms bonded in a repeating hexagonal pattern, can be made in a number of ways, mostly by stripping flakes of carbon from mined graphite (sticky tape and pencil lead will do). Levidian Nanosystems, a Cambridge firm, uses a more sustainable process. It captures methane, a potent greenhouse gas, from various industrial sources, and then zaps it with microwaves inside a reaction chamber. This cracks the gas into its constituent parts, with hydrogen emerging at the top and graphene flakes at the bottom.

Dr Koziol leads a team of researchers who work with Levidian on a variety of graphene-based applications, from reinforcing aircraft, cars and wind turbines to lining gas pipelines. In 2019, they also worked with Meditech Gloves, a Malaysian firm, to improve its surgical and examination gloves. The company makes these from latex, a mixture of water and natural rubber tapped from the bark of rubber trees. As some people are allergic to proteins found in latex, most medical gloves tend to be made from petroleum-based nitrile rubber instead. As covid took hold, some 80m pairs

of gloves were being used by health workers every month, with the nitrile ones ending up in landfill. There, they could take roughly 100 years to decompose.

Cranfield helped Meditech speed up production. It developed a new latex formula to make the firm's gloves hypoallergenic and then added graphene for a version that is lighter and stronger. Both types are currently being certified for medical use. They are more sustainable than nitrile gloves because natural rubber absorbs carbon dioxide rather than producing it, and they should biodegrade in about a year.

Testing the integrity of surgical gloves involves filling them with air and looking for leaks. Treating them, in other words, like balloons. For Dr Koziol, the parallel was particularly intriguing. Most balloons used in aerospace applications (to conduct high-altitude monitoring, for example, or for communications) are produced with synthetic plastics and filled with helium, which is becoming scarce on Earth. A graphene balloon, however, could carry heavier payloads and be filled with hydrogen instead. The hydrogen might also power any on-board devices through a fuel cell.

The Cranfield team are working with Levidian and others to produce graphene-infused natural rubber aerospace balloons to operate at altitudes of 30km. Satellites could one day be sent into orbit from here on rocket-powered launch vehicles, which would need only a small amount of fuel having been lifted to a thinner atmosphere with lower gravity. The researchers are now investigating 3D-printing to make medical gloves and balloons. For safety reasons, rockets tend to be launched over remote areas or the sea. Balloons offer greater flexibility and Dr Koziol wants to build a launch pad at Cranfield, in the heart of the English countryside. There is plenty of room: as a former aeronautical college, it still operates an active airfield. ■



先进材料

医用手套如何帮助发射卫星

石墨烯升上新高度【新知】

石墨烯坚固、轻质、柔韧，是一种优秀的电导体。不过，自20年前它在曼彻斯特大学被首次分离出来后，它在实际应用上的无足轻重也令人扫兴。这种情况正慢慢开始改变，因为它的卓越特性不断带给研究人员丰富的灵感。例如，对于英国克兰菲尔德大学（Cranfield University）的克日什托夫·科齐奥尔（Krzysztof Koziol）来说，一个最初在新冠疫情期间利用石墨烯改进手术手套的方案现在已经演变成了利用高空气球将卫星发射到太空的项目。

石墨烯是由单层碳原子以六边形模式重复键合而成，可以通过多种方式制成，主要是从开采的石墨中剥离碳片（就像用胶带从铅笔芯上粘一些石墨下来一样）。位于剑桥的Levidian Nanosystems公司采用了一种更加可持续的工艺。它从各种工业排放源中捕获甲烷（一种强效温室气体），然后在反应室内用微波将其分解。这会让甲烷裂解成其组成元素，顶部生成氢，底部是石墨烯薄片。

科齐奥尔领导的研究团队和Levidian合作探索各种基于石墨烯的应用，从加固飞机、汽车和风力涡轮机，到用于天然气管道内衬等。2019年，他们还与马来西亚公司Meditech Gloves合作，改进其手术和检查用手套。Meditech Gloves用乳胶生产这些产品，乳胶是水和从橡胶树树皮中提取的天然橡胶的混合物。有些人对乳胶中的蛋白质过敏，所以大多数医用手套往往由石油基的丁腈橡胶制成。在新冠疫情蔓延期间，医疗工作者每月用掉约8000万双手套，其中的丁腈手套最终被填埋。而它们可能需要大约100年才能分解。

克兰菲尔德大学团队帮助Meditech加快了生产速度。它开发了一种新的乳胶配方，让该公司生产的手套具有低过敏性，后来又通过添加石墨烯制成了一种更轻、更坚韧的手套。这两种手套目前正在申请医用认证。它们比

丁腈手套更具可持续性，因为天然橡胶吸收二氧化碳而不是产生二氧化碳，而且它们大约一年内就能生物降解。

测试手术用手套的完好性需要给它们充气并找寻漏气点。也就是像吹气球一样。对于科齐奥尔来说，这样的类比尤其有意思。用于航空航天（例如进行高空监测或通信）的大多数气球都是用合成塑料制成的，充的是氦气，而地球上的氦气正日益稀缺。石墨烯气球则能携带更多的有效载荷，还可以充氢气。氢气或许还能通过一组燃料电池为气球携带的所有设备供电。

克兰菲尔德大学团队正与Levidian等公司合作，开发添加了石墨烯的天然橡胶航空气球，能在30千米的高空运行。有朝一日也许可以从这个位置用运载火箭把卫星发射到轨道上，由于这里的大气层更稀薄、重力更小，只需要少量燃料就能发射入轨。研究人员目前正在尝试用3D打印技术制造医用手套和气球。出于安全考虑，火箭往往在偏远地区或海上发射，而气球提供了更大的灵活性。科齐奥尔想在位于英国乡村腹地的克兰菲尔德大学建立一个发射台。空间是足够的：该大学之前是一所航空学院，现在还有一个仍在运行的机场。■



Final countdown

A new technique to work out a corpse's time of death

AI could make the work of pathologists more accurate

IN FICTION, hard-pressed pathologists presented with a corpse are able to take a bite of their sandwich and instantly pronounce a time of death. Reality is, of course, a lot messier, and the results—or lack of them—can make or break a case.

Now artificial intelligence is offering a helping hand. By analysing thousands of deaths and what follows, the technology can offer the best estimates so far of the post-mortem interval (PMI).

Working out when a person has died is the most basic (but frustratingly imprecise) part of a forensic investigator's work. For decades these specialists have had to rely on intuition, combined with observations of the state of the deceased and clues such as temperature, both of the dead body and the environment. Different bodies decay at different rates, however, and individual circumstances can throw off the most careful PMI calculations. A body found in a ditch in northern England in 2004, for example, was given an erroneously late time of death because the ditch was shielded from sunlight and the colder-than-expected conditions had helped preserve the corpse.

Forensic-science journals are full of such cases—some routine, others bizarre—while the potentially useful details of thousands more investigations are buried in case files around the world. Now forensic researchers in America are working to access and compile these valuable write-ups, and to use machine learning to analyse them. The result is an AI-powered tool, called geoFOR, that could offer the most reliable estimates of PMI so far. Sandwiches are not included.

Developed by a research team led by Katherine Weisensee at Clemson University in South Carolina, the model is based on data pooled from more than 2,500 death investigations, with more added each week. About 1,800 of these are real-world cases involving the discovery of a body. The rest are drawn from forensic experiments at so-called “body farms” in Texas and Tennessee, in which corpses are left to decay for weeks and months under varied circumstances.

For this second group, the precise PMI is known. So photos and descriptions of their various states of decomposition over time, along with information about temperature, humidity, wind, soil type and other conditions, have all helped train the AI model. Like many good detectives, the algorithm can spot patterns and make connections between the clues.

To use the tool, forensic investigators attempting to date a body simply enter some specifics of their case into an app. Entering the location allows the AI model to factor in local weather conditions. The investigators then add observations about the deceased, such as whether they were obese, whether insects have moved in and whether dogs or rats have gnawed on the body. They must also add any indications of “purging”—a particularly grim scenario in which foul-smelling fluid oozes from the nose and ears of the corpse. In return, the AI compares the case with those in its training data and offers a likely PMI and, therefore, an estimated date of death.

At present the model offers a range of days within which it is 80% confident the true PMI lies, according to Dr Weisensee’s team. This confidence level should improve as experts and investigators around the world continue to add their own cases and examples to the database.

The results could be used to check alibis and help solve crimes, but they have other uses too. Madeline Atwell, a forensic anthropologist at Clemson University who works on the project (and serves as a deputy coroner for

Richland County, South Carolina), says the model has already helped close several missing-person cases. Combining time of death with when people were last seen alive is very useful in identifying human remains. “You match it with missing-person records, and that helps narrow your time frame,” she says.

And sometimes a more reliable PMI simply gives bereaved families a better sense of a loved one’s last moments. “It helps with that process of understanding,” says Dr Atwell. ■



最后倒计时

确定尸体死亡时间的新方法

人工智能让病理学家的工作更精确

在虚构作品中，被追着要答案的病理学家面对一具尸体，咬一口三明治，就能立即宣布死亡时间。现实里自然是要难办得多，而且得出什么结果——或者没能得出什么结果——可以决定案件的成败。

现在，人工智能可以来助一臂之力。通过分析成千上万起死亡及其后续状况，这项技术可以给出迄今为止最准确的死后间隔时间（PMI）估计。

确定一个人的死亡时间是法医最基本的工作之一（但准确度低得令人沮丧）。几十年来，这些专家只能结合对死者状态的观察，以及尸体与环境的温度等线索，依靠直觉做出判断。然而，不同的尸体腐烂的速度不尽相同，最仔细的PMI计算也会因为具体情况而失准。例如，2004年在英格兰北部的一条沟渠里发现了一具尸体，对其死亡时间的估计就偏晚了，这是因为沟渠遮挡了阳光，而低于预期的环境温度有助于保存尸体。

法医学期刊里充斥着这样的案例——有的平平无奇，有的匪夷所思。同时，世界各地的案件卷宗里还写有成千上万起调查，当中可能埋藏着有用的细节。现在，美国的法医学研究人员正在努力获取和整理这些有价值的记录，使用机器学习加以分析。其成果是一个名为geoFOR的AI工具，可以提供迄今为止最可靠的PMI估计。三明治可不在内。

该模型由南卡罗来纳州的克莱姆森大学（Clemson University）的凯瑟琳·魏森西（Katherine Weisensee）带领的研究团队开发，其基础是从2500多起死亡调查中汇集的数据，而且每周都在添加更多数据。其中约1800起是发现了尸体的真实案件。其余则来自得克萨斯州和田纳西州名为“尸体农场”的法医实验，在那里，尸体会被放置在不同的环境下，在数周或数月内观察其腐烂情况。

第二组尸体的确切PMI是已知的。因此，对尸体随时间推移的腐烂情况的照片记录和描述，以及相应的温度、湿度、风力情况、土壤类型和其他环境信息，都帮助训练了这个AI模型。正如许多优秀的侦探一样，算法可以发现规律，把各种线索联系起来。

使用该工具时，试图确定尸体死亡日期的法医只需在应用中输入案件的一些具体参数。输入地点位置后，AI模型就可以将该地点的局部天气状况考虑在内。然后，法医可以输入对死者的观察结果，例如是否肥胖，有无昆虫开始寄生，有无被狗或老鼠啃咬。他们还必须输入任何“净化”的迹象——这是一种格外恐怖的情况，尸体的鼻子和耳朵会渗出恶臭的液体。然后，AI会将这起案例比对其训练数据中的案例，给出可能的PMI，从而推算出死亡日期。

据魏森西团队表示，目前该模型可以给出一个天数范围，真正的PMI落在这个范围内的可信度为80%。随着世界各地的专家和调查人员不断向数据库添加自己的案件和实例，其可信度应该还会提高。

这些结果可以用来核对不在场证明和帮助破案，不过它们也有其他用途。克莱姆森大学的法医人类学家玛德琳·阿特韦尔（Madeline Atwell，也是南卡罗来纳州里奇兰县[Richland County]的副验尸官）参与了该项目，她表示该模型已经帮助侦破了几起人口失踪案。将死亡时间与生前最后一次被看到的时间结合起来，对辨认遗骸非常有用。“将死亡时间与失踪人口记录进行比对，有助于缩小时间范围。”她说。

有时，更可靠的PMI只是为了让丧亲家属更好地了解亲人的最后时刻。“这可以帮助实现这种了解。”阿特韦尔说。■



Urban economics

The world is in the midst of a city-building boom

Everyone, from Donald Trump and Peter Thiel to Abdel Fattah el-Sisi, is getting involved

AFRICA'S TALLEST building is rising under empty skies. Beneath the Iconic Tower in northern Egypt sits a city that officials expect to one day house 6.5m people. For now, though, it is mostly empty—like the desert that came before it.

Egypt's "New Administrative Capital" is part of a rush of city-building. Firms and governments are planning more settlements than at any time in the post-war period, with many already under construction. Ninety-one cities have been announced in the past decade, with 15 in the past year alone. In addition to its new capital in the north, Egypt is building five other cities, with plans for dozens more. India is considering eight urban hubs. Outside Baghdad, Iraq, workers have just broken ground on the first of five settlements.

And it is not just emerging economies that are building. Investors in America have spent years secretly buying land for a new city in California. To the east, the deserts of Arizona and Nevada have lured Bill Gates and Marc Lore, two billionaires, each with plans for their own metropolis. Even Donald Trump, in his bid for re-election, has proposed ten "freedom cities". In their early stages, many of these projects will attract derision. History suggests that plenty will fail. But the number and diversity of settlements under construction suggests some will triumph.

That is a great thing. Edward Glaeser of Harvard University has lauded cities as mankind's greatest invention. He notes that agglomerations of money and talent make societies richer, smarter and greener. Since companies move closer to their customers and people closer to their jobs, growing

cities beget economic growth. Economists think that doubling a city's population provides a boost to productivity of 2-5%. Given both the pressing need for new urban areas and the constraints on physical growth in existing ones, starting afresh is sometimes a shrewd decision.

In much of the poor world, land disputes, shantytowns and poor infrastructure choke development. The problem will worsen as urban areas swell by an extra 2.5bn inhabitants by 2050, according to projections by the United Nations, with the new urbanites appearing in regions where cities are already under extreme stress. Builders hope that new metropolises will help relieve the pressure. In Nairobi, near where Stephen Jennings, a former private-equity boss, is building a new city called Tatu, public-transport commutes run to over an hour for most jobs. Construction is progressing nicely in Kenya's newest settlement, where 5,000 residents already live and work in a gated village. Mr Jennings is building seven other cities across five countries in the region.

Rich-world cities have problems of their own. The push for a new town outside San Francisco—a project that goes by the label of “California Forever”—came from an “epic housing shortage” on America's west coast, says Jan Sramek, who leads a group of Silicon Valley investors making it happen. The group, which includes Laurene Powell Jobs, Steve's widow; Reid Hoffman, a co-founder of LinkedIn; and Sir Michael Moritz, a venture capitalist, will put their plans for “homes, jobs and clean energy” to a public vote in November. If approved, the city will house up to 400,000 residents on 60,000 acres of what is now farmland. Starting again is a necessary part of the solution to housing shortfalls, says Mr Sramek, citing the high costs of revamping existing infrastructure.

California Forever is among a clutch of planned towns that also aim to improve urban living. The developer is promoting high-density neighbourhoods in which residents can reach schools, jobs and shops

without a car. Today's city-builders have decided that walkability—or what is sometimes called a “15-minute city”—is a crucial selling point. Some, like Dholera in India and Bill Gates's Belmont in Arizona, are pitching so-called “smart cities”, which use sensors to direct residents away from traffic or tell them the most environmentally friendly time for a shower.

A few projects double as social experiments. Mr Lore's Telosa city (adapted from the Greek word for “highest purpose”) will do away with private ownership of land, which will instead be held in a communal trust, with money generated from leasing it spent on public services. Praxis (another Greek word, meaning “theory in practice”) has raised \$19m and collected a waiting list of potential residents who want to “create a more vital future for humanity” in the Mediterranean. A private company is building Próspera, a cryptocurrency-accepting, libertarian special economic zone in the Honduras, with a mission to “maximise human prosperity”. Praxis and Próspera are funded in-part by Pronomos, a venture-capital fund established in 2019 to invest in new cities, which is run by Patri Friedman (grandson of Milton) and counts Marc Andreessen and Peter Thiel, two billionaire investors, among its supporters.

Messrs Andreessen, Lore and Thiel are among a crop of wealthy folk with ideas about how to run cities. But governments also want to experiment. Abundant capital and low interest rates in the 2010s allowed politicians to borrow cheaply. Although rates are now higher, enthusiasm for building remains, as countries copy one another. Leaders are keen on using state finances to reshape domestic economies—and believe that new cities will help.

| *Houses built on sand*

Muhammed bin Salman of Saudi Arabia hopes that several gleaming new metropolises will attract industries that his country lacks, such as financial services, manufacturing and tourism. NEOM, a city made up of a 170km-

long building in the desert, is to be the jewel in the crown. Egypt's New Administrative Capital is purpose-built for the state's bureaucratic machinery; the government hopes it will reduce congestion in Cairo. The city already includes the Ministry of Defence's imposing Octagon—not to be confused with America's Pentagon—which spreads over a square kilometre. In Indonesia workers are clearing forests for a new capital, Nusantara. For leaders such as Joko Widodo of Indonesia and Abdel Fattah el-Sisi of Egypt, a new capital promises a legacy, lots of jobs and the ability to keep voters at arm's length.

In other countries, rulers have slightly more esoteric ambitions. El Salvador is planning to sell bonds that pay out in bitcoin in order to fund a crypto-city. The Kingdom of Bhutan said in December that it would build a "mindfulness city", with neighbourhoods designed on the repeating geometric patterns of a mandala, a Buddhist symbol. The emergence of the China State Construction Engineering Corporation, whose workers are building cities in Africa, South-East Asia and the Middle East, has lowered the costs of all megaprojects, whether fanciful or prosaic.

How many of these cities will prosper? Some infrastructure, such as electricity, internet and roads, must be in place before the first resident arrives, which means that upfront costs can be staggeringly large. Mr Sramek's company has already sunk \$1bn into buying land for California Forever and will need an additional \$1bn-2bn for just the first stage of construction. Mr Lore expects to marshal \$25bn in initial investment for his city in the desert. Prince Muhammed will lean on his kingdom's oil riches to pay for NEOM at an initial cost of \$319bn. But enthusiasm, and money, can run out; grandiose projects can become white elephants. Work on Egypt's \$60bn capital city has slowed as the country's economy falters. The Chinese developer behind Malaysia's Forest City defaulted in 2023, before residents had even moved in.

History points to characteristics shared by successful projects. State institutions can help anchor cities, as Brasília (in Brazil) and Chandigarh (in India) showed in the 20th century. Although both have had problems, people in Brazil and India are voting with their feet. Brasília's population is growing at 1.2% a year, more than double the national average. Chandigarh, a state capital, is now India's fourth-richest region on a per-person basis.

The future is less certain for cities that cannot rely on taxpayers to provide jobs and pay the bills, but California Forever and Tatu seem to be based on sensible ideas. As Mr Jennings puts it, the crucial thing is to focus on getting the "boring stuff", such as roads and sewerage, right in order to create a city that is walkable and green, but not especially smart. In addition to being what he calls "a dumb city", Mr Sramek's California Forever shares another advantage with Tatu: both will piggy-back on neighbouring economies. "We are five miles away from cities on both sides," says the Californian developer. "The strength of the demand makes a big difference to how fast you can grow." In Britain, Milton Keynes—a city established in the 1960s, less than an hour by train from London—is thriving. Reston, a planned town outside Washington, DC, is another success.

Sensible city-builders are wary of taking on debt. Developers have instead started to sell stakes in projects, demonstrating buy-in for what are long-term ventures. "You are looking at a 50-year time horizon," says Mr Jennings, who admits that it "sounds insane". He has tapped friends for capital, avoiding private-equity backers and their investment horizons, which normally come in at under a decade. California Forever is entirely funded by equity investments. If the two new settlements succeed, their investors will be rewarded. But so will many others. That is the glory of cities. ■



城市经济学

全世界掀起城市建设热潮

从特朗普、彼得·蒂尔到阿卜杜勒·法塔赫·塞西，每个人都参与其中【深度】

非洲最高的建筑在空旷的天空下拔地而起。这座位于埃及北部的标志塔（Iconic Tower）俯瞰着一座城市，官员们预计有朝一日这里将容纳650万人口。不过现在，这座城市基本上还是空无人烟——一如它出现之前的那片沙漠。

埃及的这个“新行政首都”（New Administrative Capital）是一轮城市建设热的一部分。企业和政府正在规划的定居点比战后任何时候都多，其中许多定居点已经在建设中。在过去十年中已经宣布了91个城市建设计划，仅去年一年就宣布了15个。除了北部的的新首都，埃及还在建设其他五个城市，并计划再建几十个。印度正在考虑建立八个城市中枢。在伊拉克巴格达郊外，工人们刚刚破土开建五个定居点中的第一个。

而且不仅仅是新兴经济体在大兴土木。美国的投资者花了数年时间悄悄购买土地，要在加州建设一个新城市。在美国东部，亚利桑那州和内华达州的沙漠吸引了比尔·盖茨和马克·洛尔（Marc Lore）这两位亿万富翁，每个人都有自己的大都市计划。甚至特朗普在再次竞选时也提出建十个“自由之城”的计划。在早期阶段，其中许多项目会招致嘲笑。从历史经验来看，它们很多都会失败。但是，正在建设的定居点的数量和多样性表明，其中一些将取得成功。

这是件大好事。哈佛大学的爱德华·格莱泽（Edward Glaeser）称赞城市是人类最伟大的发明。他指出，资金和人才的聚集使社会更富裕、智能和环保。由于公司搬到离客户更近的地方、人们搬到离工作地点更近的地方，城市的扩张带来了经济增长。经济学家认为，城市人口增加一倍可以让生产率提高2%到5%。鉴于对新城区有迫切需求而同时现有城市的物理扩张又受到限制，另建新城有时是一个精明的决定。

在许多贫穷国家，土地纠纷、棚户区 and 基础设施薄弱抑制了发展。根据联合国的预测，到2050年，随着城市人口增加25亿，问题将进一步恶化，因为新居民将出现在本就处于极端压力之下的城市里。建筑商希望新建的大城市将有助于缓解压力。在内罗毕，大多数人的公共交通通勤时间超过一个小时；前私募股权老板斯蒂芬·詹宁斯（Stephen Jennings）正在该市附近建设一座名为塔图（Tatu）的新城。肯尼亚最新一个定居点的建设进展顺利，5000名居民已经在一个封闭式村庄里生活和工作。詹宁斯正在该地区的五个国家建设另外七座城市。

富裕世界的城市也有自己的问题。在旧金山郊外建造一个新城镇的“永远加州”（California Forever）计划便是受美国西海岸“惊人住房短缺”的推动，扬·斯拉梅克（Jan Sramek）说。他牵头一群硅谷投资者来促成这个项目，其中包括乔布斯的遗孀劳伦·鲍威尔·乔布斯（Laurene Powell Jobs）、LinkedIn的联合创始人里德·霍夫曼（Reid Hoffman）、风险投资家迈克尔·莫里茨（Michael Moritz），他们将于11月把“住房、就业和清洁能源”计划递交公开投票。如果获通过，这个新城将在现在的60,000英亩农田上容纳多达40万居民。斯拉梅克说，另建新城是解决住房短缺的必要措施之一，他指出改造现有基础设施的成本很高。

除了“永远加州”，还有一系列也意在改善城市生活的新城规划。开发商正在推广高密度社区，居民不需要用车就能到达学校、工作地点和商店。如今的城市建设者已经判定，可步行性——或者有时被称为“15分钟城市”——是一个关键的卖点。一些城市，如印度的托莱拉（Dholera）和比尔·盖茨要在亚利桑那州打造的贝尔蒙特（Belmont），正在营销所谓的“智能城市”，这些城市使用传感器引导居民避开交通拥堵，或者告诉他们最环保的淋浴时间。

一些项目同时也是社会实验。洛尔的特洛萨（Telosa，源自希腊语中“最高目标”一词）将取消土地私有，取而代之的是一个公共信托，把出租土地带来的资金用于公共服务。普莱克西斯（Praxis，另一个希腊词，意思是“付诸实践的理论”）已经筹集了1900万美元，并收集了一份潜在居民的等候名单，这些人都希望在地中海“为人类创造更有活力的未来”。一家私营

公司正在洪都拉斯建设普洛斯佩拉（Próspera），这是一个接受加密货币的自由意志主义经济特区，其使命是“充分实现人类繁荣”。普莱克西斯和普洛斯佩拉的部分资金来自Pronomos，这家成立于2019年的风险投资基金专注投资新城市，由帕特里·弗里德曼（Patri Friedman，经济学家米尔顿·弗里德曼的孙子）经营，其支持者包括两位亿万富翁投资者马克·安德森（Marc Andreessen）和彼得·蒂尔（Peter Thiel）。

包括安德森、洛尔和蒂尔在内的一群富人对如何管理城市有自己的想法。但政府也想开展实验。2010年代充裕的资本和低利率使政客们能够低成本借债。尽管现在的利率更高，但随着各国相互仿效，建设热情依旧高涨。领导人渴望利用国家财政来重塑国内经济，并相信新城市会发挥作用。

｜建在沙子上的房子

沙特王储穆罕默德·本·萨勒曼希望几座熠熠生辉的崭新大都市能够吸引来自自己国家缺乏的产业，如金融服务业、制造业和旅游业。新城NEOM是沙漠中一座170公里长的建筑，将成为皇冠上的明珠。埃及的新行政首都专门是为国家的官僚机构建设的，政府希望它能减轻开罗的拥堵。这里已经建成了雄伟的八角形国防部大楼——可不要与美国的五角大楼搞混——占地超过一平方公里。在印度尼西亚，工人们正在为新首都努桑塔拉（Nusantara）清理森林。对于印尼总统佐科·维多多和埃及总统阿卜杜勒·法塔赫·塞西等领导人来说，新首都有望成为千秋功业、带来大量就业，并能与选民保持距离。

在其他国家，统治者的宏图大计要更晦涩深奥一些。萨尔瓦多正计划发行以比特币支付的债券来为一座加密城市提供资金。不丹在去年12月表示，它将建造一座“正念之城”，其城区设计采用佛教符号曼荼罗重复的几何图案。中建集团的工人在非洲、东南亚和中东建造城市，它的出现降低了所有大型项目的成本，无论是奇特的还是平淡无奇的。

这些城市中有多少能繁荣发展？一些基础设施，如电力、互联网和道路，必须在第一批居民入住之前就位，这意味着前期成本可能高得惊人。斯拉梅克的公司已经投入了10亿美元为“永远加州”买地，而且仅第一阶段的建

设就还需要10亿至20亿美元。洛尔预计将为他的沙漠城市筹集250亿美元的初始投资。穆罕默德王子将依靠沙特的石油财富为NEOM支付3190亿美元的初始成本。但是热情和金钱可能会耗尽，宏伟的项目可能沦为面子工程。随着埃及经济陷入衰退，它耗资600亿美元的首都工程也已经放缓。马来西亚的森林城市（Forest City）背后的中国开发商在2023年违约，当时居民甚至还没有搬进来。

纵观历史，可以看到成功项目的一些共同特征。身为政府机构所在地可以帮助稳固新城，正如20世纪的巴西利亚（Brasília，巴西）和昌迪加尔（Chandigarh，印度）所彰显的。尽管两者都曾遭遇问题，但巴西和印度的人们正在用脚投票。巴西利亚的人口以每年1.2%的速度增长，是全国平均水平的两倍多。昌迪加尔是邦首府，按人均计算现在是印度第四富有的地区。

对于那些不能依靠纳税人提供就业和支付账单的城市来说，未来更不确定，但“永远加州”和塔图的思路似乎是明智的。正如詹宁斯所说，关键是要专注于做好“乏味的事情”，例如道路和污水处理系统，以创造一个适合步行和绿色的城市，而不需要特别智能。除了成为他所说的“愚钝城市”之外，斯拉梅克的“永远加州”还有另一个优势，塔图也是：两者都可以借邻近经济体的东风。“我们距离两边的城市都只有五英里远，”这位加州开发商说，“需求强度对你的增长速度有很大影响。”在英国距离伦敦不到一个小时火车车程的地方，建于1960年代的城市米尔顿凯恩斯（Milton Keynes）正在蓬勃发展。美国华盛顿特区附近的规划城镇雷斯頓（Reston）是另一个成功案例。

明智的城市建设者对举债态度谨慎。开发商转而开始出售项目的股份，以表明对这些长期项目的认同。“你要把眼光放到50年。”詹宁斯说，他承认这“听起来很疯狂”。他从朋友那里获得资金，没有去找私募股权投资者，也就避开了他们对投资期限的要求——通常不到十年。“永远加州”的资金全部来自股权投资。如果这两个新城取得成功，它们的投资者将获得回报。但其他众多人也会受益。这就是城市的荣耀。■



Schumpeter

Apple is right not to rush headlong into generative AI

One day the Vision Pro could exploit the technology to the full

IF YOU THINK Tim Cook has always led a charmed life at the helm of Apple, think again. The years straight after the death of Steve Jobs in 2011 were a trial by fire. First there was antitrust: America's Department of Justice (DoJ) sued Apple for conspiring to fix e-book prices. Then there was competition: Samsung, a South Korean rival, went to war with the iPhone with bigger, sleeker models. Then came broader concerns. Apple's new voice assistant, Siri, made rookie errors. Ditto Apple Maps, which went as far as relocating the Washington Monument to the Potomac river. At the time, the question hanging over the company was existential: could Apple's creative spark survive the death of its founder? One of Mr Cook's lieutenants was so miffed at the criticisms that he publicly retorted in 2013: "Can't innovate any more, my ass!"

A decade or so later, Mr Cook may be feeling déjà vu. On all three counts—antitrust, Asian competition, the existential question of innovation and growth—there are parallels between then and now. Competition watchdogs in the EU are demanding compliance from March 7th with rules that for the first time breach the "walled garden" which keeps users and developers bound within Apple's playpen. On March 4th they fined the company €1.8bn (\$2bn) for allegedly stifling competition in music streaming. In America the DoJ may soon launch a case against Apple. In China, Huawei, a domestic mastodon, is seizing market share. Hanging over everything is the nagging concern, amid a levelling off in iPhone sales, that Mr Cook is missing the chance to pull another rabbit out of the hat with generative artificial intelligence (gen AI).

In short, with its market value down by 10% since mid-December, and

Microsoft, thanks to gen AI, vaulting past it to become the world's most valuable company, sceptics wonder if Apple is now so dominant it has lost its mojo. So jaded is the narrative that many pay little heed to the buzz about the Vision Pro, Apple's snazzy—though lavishly priced—mixed-reality headset. What hopes they have are pinned on the company's annual developer conference in June, when they want Mr Cook to announce whizzy gen-AI upgrades proving that Apple can join the chatbot hypefest. That, though, is not how the company does things. Nor should it be.

Go back to the threat from Samsung in Mr Cook's early days. Back then investors pestered Apple to come up with a bigger phone, just as now they want it to match Samsung's models with gen-AI bells and whistles. But Apple doesn't rush things. It wasn't until the launch of the iPhone 6 in 2014 that it produced a large-screen device. When it came, it was a smash hit. Its modus operandi remains the same. It is rarely first with a product. It seeks to improve what is already out there, learning from others' mistakes and eventually trouncing the competition. Of course, that poses a risk. In theory, a scrappy upstart may produce new technology products cheaper and faster, pulling the rug from under the market leader. Perhaps a young company building a killer gadget for the gen-AI era already has Apple in its sights.

Yet you do not have to be a true believer to see why Apple may be right to take its time. First, there will be more to gen AI than chatbots. They appear revolutionary. But so far they are just a better (and accident-prone) way of putting in a query and getting an answer. That is not Apple's forte. "They are features, not products," as Horace Dediu, an expert on Apple, puts it. Nor does Apple compete with other tech giants, such as Microsoft, Amazon and Alphabet, to run cloud-computing platforms with AI models on which customers can build gen-AI apps. Instead of relying on the cloud, it seems to be working on ways to embed gen AI in its own devices, bolstering its ecosystem. Since 2017 it has been using homemade chip technology called

neural engines to handle machine-learning and AI functions its gadgets use behind the scenes.

In late February it emerged that it was scrapping its ten-year project to build an Apple car and redirecting the engineers towards gen AI. No doubt it is moving up a gear—though not from an idle start. Apple will reveal nothing about its intentions. But one of the options it has is hiding in plain sight: the Vision Pro. The most recent gen-AI launches, such as OpenAI's Sora, which converts text to video, and Groq, which speaks at humanlike speed in response to questions, suggests that eventually something other than written words could be the main gateway to gen AI. The Vision Pro is all about sounds and images.

| *Known unknowns*

In the short term none of this will resolve the growth question. In fact, the regulatory onslaught in the EU via the Digital Markets Act, which will henceforth apply to big-tech “gatekeepers” including Apple, could potentially crimp its biggest growth engine, services. For the first time Apple will be forced to allow third-party app marketplaces and alternative payment systems outside its App Store on devices in Europe. It has made no secret of its disdain for the rules. It calls them a threat to safety and privacy, and has introduced complex new fees for those who dare bypass its protective walls. Some developers have slammed its compliance measures, but they are likely to work: inertia means that many will probably stick with the status quo. As for a possible DOJ antitrust case, it would be a headache. But its scope is not yet clear.

China is a bigger problem with no clear solution. Huawei has become a formidable competitor, though in the long run it may be constrained by an America-led ban on sales to it of high-end chips. However big the geopolitical risks, Apple and China are so co-dependent that they may be stuck with one another.

Still, don't give up on Mr Cook yet. Apple is bound to be working on gen-AI products that do not leave egg on its face—just, as is its way, not in the open. At this stage, the vast sums needed to train AI models favour deep-pocketed incumbents over scrappy upstarts, which will work to Apple's advantage. You can almost hear Cupertino muttering, "Can't innovate anymore, my ass!" ■



熊彼特

苹果不急于扎进生成式AI是对的

有朝一日*Vision Pro*可能会充分利用这项技术

如果你以为库克执掌苹果的日子从来都一帆风顺，那再想想吧。2011年乔布斯去世后的那几年，库克就像是被架在火上烤。先是反垄断官司：美国司法部起诉苹果合谋操纵电子书价格。再有是竞争挑战：韩国对手三星推出更大更时尚的机型，与iPhone争夺市场。然后是更大方向上的担忧。苹果彼时新推出的语音助手Siri出现低级错误。苹果地图也有同样的问题，竟错把华盛顿纪念碑标记到波托马克河上。当时，围绕苹果的是一个生死攸关的问题：创始人去世后，它的创造力火花还能继续闪耀吗？库克的一位副手对这些批评感到非常恼火，在2013年公开反驳道：“再没法创新了？去你的吧！”

十来年过去了，今天的情形可能让库克有似曾相识之感。反垄断、来自亚洲的竞争、事关存亡的创新与增长问题——在这三方面，眼下的处境和当年都有相似之处。欧盟的竞争监督机构制定了新规则，首次打破了苹果把用户和开发者限制在自家花园里的围墙，要求苹果从3月7日起遵守这些规定。3月4日，它们对苹果处以18亿欧元（20亿美元）的罚款，理由是该公司涉嫌在音乐流媒体领域压制竞争。在美国，司法部可能很快会对苹果提起诉讼。在中国，本地巨头华为正在抢占市场份额。在iPhone销量停滞不前之际，悬于一切之上的还有那挥之不去的担忧：库克正在错失利用生成式AI创造另一次奇迹的机会。

自去年12月中旬以来，苹果的市值下跌了10%，而微软则凭借生成式AI超越苹果，晋身全球价值最高的公司。简单来说，怀疑者们猜疑现在的苹果是否因为地位过于稳固而失去了自己的魔力。这个说法听得人们耳朵都要起茧了，连苹果新推出的时髦新潮却也价格不菲的混合现实头显*Vision Pro*都挑不起他们的兴致。他们的希望都寄托在6月召开的年度开发者大会上，想看看库克会宣布什么生成式AI方面的新奇进展，证明苹果也能加入聊天机器人的热潮。不过这并不是、也不该是苹果的行事风格。

我们来回头看看库克在掌舵苹果之初面对三星的威胁时的做法。当时投资者催促苹果推出尺寸更大的手机，就像现在他们要苹果跟上三星那些配备了各种花哨AI功能的机型一样。但苹果一向不会贸然行事。直到2014年的iPhone 6，苹果才开始生产大屏幕设备，一经推出便大受欢迎。苹果的行事方式依然如故。它很少率先推出全新产品，而是力求改进市场上已有的产品，从别人的错误中吸取教训，最终击败对手。当然，这存在风险。理论上，斗志旺盛的后起之秀可能会以更低的成本和更快的速度生产出新款科技产品，把市场领导者拉下马。也许志在打造AI时代杀手级设备的年轻公司早已经瞄上了苹果。

但即便你不是果粉，也能明白为什么苹果不急于出手可能是对的。首先，生成式AI将不仅局限于聊天机器人。它可能带来革命性的改变。但迄今为止，人们得到的只是一种输入查询并获得答案的更好（也很容易出问题的）方式而已。这非苹果的强项。“这些只是功能，不是产品。”专门研究苹果的霍拉斯·德迪欧（Horace Dediu）说。苹果也不会与微软、亚马逊和Alphabet等其他科技巨头竞争，去运营AI模型的云平台，让客户可以在上面构建生成式AI应用。苹果似乎并不依赖云计算，而是在研究如何把生成式AI嵌入自己的设备中以增强自己的生态系统。自2017年以来，苹果一直在使用自家名为神经引擎的芯片技术来处理在设备后台运行的机器学习和AI功能。

2月底，有消息称苹果放弃了历时十年的造车项目，把工程师转移到了生成式AI领域。毫无疑问，苹果正在升档加速，但并不是从怠速状态启动。苹果不会透露自己的意图，但其中一个选项已昭然若揭：Vision Pro。OpenAI推出了能根据文字生成视频的Sora，还有能像人一样迅速作答的Groq，这些最新的生成式AI产品表明，最终一些书面文字以外的东西可能成为接入生成式AI的主要途径。Vision Pro就是主攻声音和图像。

| 已知的未知

短期内，这些都无法解决增长问题。事实上，欧盟通过《数字市场法案》（Digital Markets Act，今后将适用于包括苹果在内的大型科技“守门人”）带来的监管打压有可能削弱苹果最大的增长引擎——服务。破天荒

地，苹果将被迫允许欧洲的苹果设备使用其应用商店以外的第三方应用市场和其他支付系统。苹果毫不掩饰对这些规则的鄙视，称它们对安全和隐私构成威胁，并对那些胆敢绕过围墙的用户引入了复杂的新收费。一些开发者抨击苹果的合规方式，但这些措施很可能奏效：惰性可能会促使很多人选择维持现状。至于美国司法部可能提起的反垄断诉讼，则会是令人头疼的问题，但其规模尚不明确。

更大的问题是中国，而且没有明确的解决方案。华为已成为强大的竞争对手，尽管长远来看它可能受到美国主导的高端芯片销售禁令的限制。无论地缘风险有多大，苹果和中国相互依赖的程度可能令它们继续捆绑在一起。

不过，不要就此对库克丧失信心。苹果多半是在开发不会让自己丢脸的生成式AI产品，只是按其一贯作风暂不公开而已。在现阶段，训练AI模型耗资巨大，更适合资金雄厚的老牌公司而非跃跃欲试的新公司，这将对苹果有利。你几乎都听到库比蒂诺在念叨了：“再没法创新了？去你的吧！”■



Colour by numbers

Can a dozen shipwrecks tell the history of the world?

Historical listicles are in vogue

A History of the World in 12 Shipwrecks. By David Gibbins. St Martin's Press; 304 pages; \$32. W&N; £25

A SPECTRE IS haunting history—the listicle. You know the formula: a history of something enormous in an intriguingly specific number of unexpected things. In 2023 alone readers were met with new books about the history of baseball in 50 moments, the West in 14 lives, the information age in five hacks, women in 101 objects and the world in eight plagues, ten dinners or 50 lies. Call it history-by-numbers, or, if you must, the “histicle”.

The oldest histicle of all may be the seven wonders of the world, a list of marvels from around 300BC. But the modern vogue for history-by-numbers began in 2010, when Neil MacGregor's “A History of the World in 100 Objects” became a bestseller. Mr MacGregor, who was then director of the British Museum, highlighted treasures from the collection.

Breaking down history is appealingly accessible—no need to read a tome to understand the Silk Road, just look at this Korean roof tile. But without care and a certain amount of panache, it risks giving the impression that history is merely, as Arnold Toynbee, a historian, characterised the views of less diligent colleagues, “One damned thing after another.” Or one damned roof tile after another.

Fortunately “A History of the World in 12 Shipwrecks” does not fall into that trap. The book works in no small part because it is more modest than its title suggests, sailing between the Scylla of exaggeration and the Charybdis of triviality and avoiding a wreck along the way. For starters, it is not a

history of the world. Rather, it is a series of interconnected essays on notable shipwrecks. But in recounting maritime misadventure, David Gibbins paints a picture of how various people—from sailors in Britain in 2500BC, to Muslims in the ninth century and British merchant mariners in the 20th—pushed the borders of their known worlds.

The book's itinerary through the flotsam and jetsam of history is avowedly idiosyncratic and makes no claim to objectivity or completeness. Mr Gibbins is best known for writing historical fiction, but he is also a professional maritime archaeologist. The skeleton of the book mirrors Mr Gibbins's own career: he has worked on more than half of the 12 wrecks he writes about. His descriptions of snaking through kelp and sucking up sand to fish out lost treasure are vivid. Shipwrecks are "catastrophic events", he writes, but the challenge of diving them is "life-affirming". (After he swims into the "dead pool" off the Cornish coast, though, you may prefer to take him at his word.)

The sailors' misfortune aside, a shipwreck serves as a lucky stroke for a historian, especially if it is left intact. Each wreck is a microcosm of a lost world. As Mr Gibbins writes of the *Mary Rose*, Henry VIII's flagship vessel that sank off the Isle of Wight in 1545, a wreck "can be seen as an unfolding series of contexts", giving details not just of how ships were run, but of what its sailors took with them, and, in the cargo hold, what its culture needed or prized. For King Henry's sailors, it might have been the wooden tankards in which they drank their gallon of weak beer a day.

In the case of the *Santo Cristo di Castello*, a Genoese ship that sank in 1667 on the way from Amsterdam to England, Mr Gibbins discovered that the precious cargo included the studies for two paintings by Rembrandt—an artist who himself benefited from maritime trade and the wealthy class of collectors it created.

Ships are not just miniatures of the worlds they sailed from. Many were the engines of history themselves, carrying not just people and goods but ideas, religions and technologies. The wreck of an Arabic ship near Indonesia disclosed a cargo of Tang porcelain, as well as an inkstone, suggesting the ship might have been carrying early paper back from China.

Diving for shipwrecks, then, is an extreme version of what any historian does: plunging into the darkness to discover the missing links between events and cultures. And what is history if not a network of wrecks, accidents, hidden treasures and unexpected consequences? At its best, history-by-numbers embraces the fact that historical narratives, like history itself, are always somewhat arbitrary. This is history not “as it actually happened”, as Leopold von Ranke, a German historian, famously put it, but history as it happened to happen.

This vision of the past as a loosely connected chain of events lacks the storybook comfort of beginning, middle and end. But, at the same time, the historical list offers solace: it suggests that history is finite, knowable and manageable. The list does not just appeal to modern brains, fed on news from BuzzFeed. It also appeals to a desire for control, promising that history can be broken down into discrete parts and understood, and that five or 100 things is all you need to know. This approach to the past suits the present perfectly. ■



数字填色彩

十二艘沉船能否讲述世界历史？

清单体史书正时兴（《十二艘沉船中的世界史》书评）

《十二艘沉船中的世界史》，大卫·吉宾斯著。圣马丁出版社；304页；32美元。W&N出版社，25英镑

一个幽灵正徘徊于历史海洋中，那就是清单体。大家都知道这个套路：找到一个意想不到的事物，给它设定一个有趣的数量，以此来讲述某段宏大的历史。仅在2023年，读者就迎来了诸多这类新书：《50个瞬间看棒球史》，《14段生命述说西方历史》，《五次黑客攻击讲述信息时代》，《从101件物品看女性》，《八场瘟疫、十次晚餐和五十个谎言中的世界》。我们可以称之为数字讲史，如果你非要给它造一个词的话，也可以叫“清单体史书”。

最古老的清单体史书可能是“世界七大奇迹”，列举了公元前300年左右的一批建筑和雕塑奇观。现代这股数字讲史风尚则始于2010年，时任大英博物馆馆长的尼尔·麦克格雷格（Neil MacGregor）以该馆部分馆藏为视角著成的《大英博物馆100件文物中的世界史》（A History of the World in 100 Objects）成了畅销书。

分解讲述让历史通俗易懂，读来轻松愉快——要了解丝绸之路，无需死啃大部头，看看这块韩国的瓦片就行。但如果不用心，没有一定的派头气势，清单体史书有可能给人留下这样的印象：历史不过是“一件破事接着另一件”——按历史学家阿诺德·汤因比（Arnold Toynbee）所述，他那些不太勤奋的同行们便是持这样的见解。也可以说是一片破瓦叠着另一片。

幸好，《十二艘沉船中的世界史》（A History of the World in 12 Shipwrecks）没有落入这一窠臼。这本书之所以成功，很大程度上是因为它的内容比书名要谦和，在夸夸其谈和鸡毛蒜皮之间恰到好处地航行而能不翻船。首先，这并非一部世界史，而是一系列关于著名沉船事件的相互

关联的文章。但在对海难的叙述中，大卫·吉宾斯（David Gibbins）描绘了其中各色人物（从公元前2500年的英国船员到公元9世纪的穆斯林以及20世纪的英国商船船员）如何突破他们已知世界的边界。

这本书在历史海洋中漂浮的残骸和弃物间展开叙事，毫不掩饰个人色彩，也未声称要做到客观或全面。吉宾斯最为人称道的是他的历史小说，但他也是专业的海洋考古学家。该书的框架正是吉宾斯自己职业生涯的写照：书中写到的12艘沉船中，他亲自参与了其中半数以上的考古工作。他活灵活现地描述了在海藻中蜿蜒穿行和淘沙寻宝的过程。沉船是“灾难性事件”，他写道，但在沉船中潜水的挑战却是“振奋生命的”。（在他游进康沃尔海岸附近的“死亡水域”后，你可能更愿意相信他的话）。

撇开船员的不幸不谈，沉船对历史学家来说是件幸事，尤其是当沉船得以完整保存下来。每艘沉船都是一个失落世界的缩影。正如吉宾斯谈到1545年在怀特岛（Isle of Wight）附近沉没的亨利八世舰队的旗舰“玛丽玫瑰号”（Mary Rose）时所写的，沉船“可被视作一系列背景故事在徐徐展开”，不仅透露了船只运作的细节，还有船员携带的物品，以及货舱中的货物其文化所需要或珍视的东西。对亨利八世的船员而言，那可能是他们每天喝掉一加仑淡啤酒所用的木制大啤酒杯。

说到其他船上的宝物，1667年，一艘名为“圣克里斯托·迪卡斯特洛号”（Santo Cristo di Castello）的热那亚船只在从荷兰阿姆斯特丹驶往英国的途中沉没。吉宾斯发现，这艘沉船的珍贵货物中有伦勃朗两幅画作的草图，而这位画家本身就受益于海上贸易和由此产生的富裕收藏家阶层。

船只不仅仅是它们所属世界的微缩模型。许多船只本身就是历史的引擎，不仅运载人员和货物，还传送思想、宗教和技术。在印尼附近一艘阿拉伯船只的残骸中，人们发现了一批中国唐代瓷器，还有一方砚台，表明这艘船可能是从中国运回早期纸张。

因此，沉船潜水是历史学家工作的极限状态：潜入黑暗中，发现事件和文化之间缺失的关联。而如果历史不是由沉船、事故、隐藏的宝藏和意想不

到的后果交织而成的网络，那它又是什么？“数字讲史”的方式最为极致地展现了一个事实：历史叙事，正如历史本身，总归带有某种程度的主观随意性。它所记述的不是德国历史学家利奥波德·冯·兰克（Leopold von Ranke）所谓的“如实发生”的历史，而是碰巧如是的历史。

这种把过去视为一连串松散关联的事件的方式缺乏人们习惯了的那种有头有尾有经过的故事书叙事。但是，与此同时，这种历史的清单也给人们以慰藉：它暗示历史是有限的、可知的、可掌握的。这样的清单不仅吸引从 BuzzFeed 上获取新闻的现代人大脑，也迎合了控制欲，承诺人们历史可被分解成独立的部分而被理解，只需要知道五件或一百件事物就行了。这种看待过去的方式完美契合了当下。■



Bartleby

Why you should lose your temper at work

Sometimes. And without throwing anything

AWARENESS DAYS are meant to remind people of important causes and desirable behaviour. Among other things, February sees the International Day of Human Fraternity, World Day of Social Justice and—everyone's favourite until it became a bit too commercialised—World Pulses Day. International Day of Happiness falls in March; you have to wait until November for World Kindness Day.

Anger is far too objectionable to be celebrated with a special day of its own. There is an anger-awareness week in Britain, but the emphasis is on controlling tempers, not giving in to them. Yet in the workplace, as elsewhere, anger is more ambiguous than it seems.

Its destructive side is obvious. Furious people are not much fun to work with, and less fun to work for. A short-fused boss is likely to instil fear among employees and to discourage people from speaking up. Anger can also engender poor performance. Anyone who has ever been riled by a rude email or uncivil colleagues knows how in such circumstances suddenly nothing else matters. Every spare bit of cognitive power is redirected to thinking of devastating put-downs from which the offender will never recover; other tasks can wait.

In one paper on the effects of rudeness on medical professionals, Arieh Riskin of Bnai Zion Medical Centre in Haifa and his co-authors describe a training exercise in which teams of Israeli physicians and nurses treated a mannequin of a baby. The teams were joined by someone billed as a visiting expert from America, who offered studiously neutral comments to some groups and made unprompted and disparaging remarks about the

quality of medical care in Israel to others. The teams that had suffered rudeness performed significantly worse.

Being angry all the time is bad news for individuals and organisations alike. But so is being tremendously satisfied by everything all the time. Jeffrey Pfeffer, a professor at Stanford University who teaches a course on how to acquire power, reckons that displaying anger is an important skill for those who want to rise up the corporate ladder. It is associated with decisiveness and competence (though angry women are more likely to evoke negative emotions among other people than angry men do). Doctors who get angry if they are challenged about their medical advice are not judged to be less competent; if they show shame, patients take a dimmer view.

Anger can have a galvanising effect in specific circumstances. A study by Barry Staw of the University of California, Berkeley, and his co-authors analysed half-time team talks by college and high-school basketball coaches in America, and found that expressions of negative emotions such as anger and disappointment were associated with better second-half outcomes—up to a point. When coaches reached the bulging-eyeballs stage, rage started to have the opposite effect.

There are similar nuances in negotiations. A paper by Hajo Adam of Rice University and Jeanne Brett of Northwestern University found that as people got more upset, they were more likely to extract concessions. But being too angry was seen as inappropriate. And although displays of anger can work in one-off negotiations, they also invite retaliation in subsequent interactions.

Anger has different effects on different types of people. Agreeableness is one of the “Big Five” personality traits recognised by most psychologists. Agreeable sorts value co-operation and courtesy; disagreeable ones are more cynical and more comfortable with conflict.

In an experiment by Gerben Van Kleef of the University of Amsterdam and his co-authors, teams comprised of agreeable and disagreeable people were given feedback on their performance by an actor. The words were the same each time, but in some instances the actor looked and sounded happy and in others they looked and sounded angry. An angry evaluation spurred the more disagreeable teams to do better than a happy (or poker-faced) one; the reverse applied to the more agreeable teams.

By now the problem should be obvious. Anger involves a loss of control. But to be effective in the workplace, it needs to be carefully modulated. That means volcanic people need to find ways to rein themselves in before they spew invective everywhere. It also means that equable people need to learn to let fly occasionally. If there is room in the calendar for International Jazz Day, then there is certainly a case for World Calibrated Displays of Anger Day. ■



巴托比

职场发火有理

偶尔为之。而且不能扔东西

宣传日是为了提醒人们记得重要的事业和理想的行为。2月里有很多这样的日子，比如国际人类博爱日、世界社会公正日，以及大家原本最喜欢的世界豆类日——可惜它已经有点过于商业化了。国际幸福日在3月；世界友善日则要等到11月。

发脾气实在令人不快，不配拥有一个专属节日。英国有“愤怒情绪意识周”，但强调的是控制脾气，而不是乱发脾气。然而在职场上，就和在其他地方一样，发脾气并不是表面看起来那么简单。

它破坏性的一面是显而易见的。和容易发脾气的人共事没多少乐趣可言，在他们手下工作就更没意思了。脾气暴躁的老板可能会让员工心生恐惧，不敢直言不讳。发脾气也会导致工作表现不佳。任何曾被无礼的电邮或粗鲁的同事惹恼的人都知道，在这个时候，其他事情突然间都变得无关紧要了。那点残存的认知力全都被用来琢磨这些令人万分难堪的奚落，冒犯者将永远无法收回；其他事情都得靠边站。

在一篇关于粗鲁言行对医护人员的影响的论文中，以色列海法市（Haifa）布纳伊锡安医疗中心（Bnai Zion Medical Centre）的阿里耶·里斯金（Arieh Riskin）等人描述了一次培训活动，由一些以色列医生和护士小组对婴儿假人进行治疗操作。一位自称是美国客座专家的人加入了活动，他刻意对一些小组给出客观中立的点评，但在另外一些小组那里则无端地贬低了以色列的医疗质量。受到无礼对待的小组接下来的表现要差得多。

随时随地发脾气对个人和组织来说都是个麻烦。但对任何事情都总是心满意足也不是件好事。斯坦福大学教授杰弗里·普费弗（Jeffrey Pfeffer）有一门关于如何获得权力的课程，他认为，想要在公司晋升，发脾气是一项

重要的技能。它被认为是果断和能力的体现（不过，脾气大的女性比脾气大的男性更容易引起别人的负面情绪）。如果医生因其医嘱受到质疑而发怒，病人并不会认为他们的能力不足；但如果他们表现出羞愧，就会受到病人的质疑。

在特定情况下，发脾气可能激发人的斗志。加州大学伯克利分校的巴里·斯托（Barry Staw）等人分析了美国大学和高中篮球教练在中场休息时的训话，发现愤怒和失望等负面情绪表达与下半场更好的战绩有关——但仅限于一定程度。当教练达到怒目圆睁的阶段时，愤怒就会开始产生反作用。

在谈判中也存在类似的程度问题。莱斯大学的哈乔·亚当（Hajo Adam）和西北大学的珍妮·布雷特（Jeanne Brett）的一篇论文发现，人们越生气，就越有可能获得对方的让步。但过于愤怒也被认为是不适宜的。而尽管发脾气在一次性的谈判中可能奏效，在随后打交道时也会招致报复。

愤怒会给不同类型的人造成不同的影响。亲和性是大多数心理学家公认的“五大”性格特质之一。随和的人注重合作和礼貌；坏脾气的人更加愤世嫉俗，更适应冲突。

在阿姆斯特丹大学的赫尔本·范克里夫（Gerben Van Kleef）等人所做的实验中，由一名演员向几个团队的表演提出反馈意见。一些团队的成员性格随和，另一些脾气暴躁。演员每次说的话都一样，但有时表情和语气显得愉快，有时表现得生气。对于坏脾气的团队，愤怒的评价比愉快（或面无表情）的评价更刺激了他们表现得更好；而在较为随和的团队中，情况刚好相反。

现在，难点应该很清楚了。发脾气是一种失控。但要在职场有效地发脾气，其程度却需要小心调校。这意味着那些行走的活火山们需要在动辄破口大骂之前设法控制自己。而那些平心静气的人则需要学会偶尔放飞一下脾气。如果国际爵士乐日都能出现在日历上，那么当然也有理由设立一个世界校准发怒日。■



Silicon dreamin'

AI models make stuff up. How can hallucinations be controlled?

It is hard to do so without also limiting models' power

IT IS AN increasingly familiar experience. A request for help to a large language model (LLM) such as OpenAI's ChatGPT is promptly met by a response that is confident, coherent and just plain wrong. In an AI model, such tendencies are usually described as hallucinations. A more informal word exists, however: these are the qualities of a great bullshitter.

There are kinder ways to put it. In its instructions to users, OpenAI warns that ChatGPT "can make mistakes". Anthropic, an American AI company, says that its LLM Claude "may display incorrect or harmful information"; Google's Gemini warns users to "double-check its responses". The throughline is this: no matter how fluent and confident AI-generated text sounds, it still cannot be trusted.

Hallucinations make it hard to rely on AI systems in the real world. Mistakes in news-generating algorithms can spread misinformation. Image generators can produce art that infringes on copyright, even when told not to. Customer-service chatbots can promise refunds they shouldn't. (In 2022 Air Canada's chatbot concocted a bereavement policy, and this February a Canadian court has confirmed that the airline must foot the bill.) And hallucinations in AI systems that are used for diagnosis or prescription can kill.

| *All the leaves are brown*

The trouble is that the same abilities that allow models to hallucinate are also what make them so useful. For one, LLMs are a form of "generative" AI, which, taken literally, means they make things up to solve new problems. They do this by producing probability distributions for chunks of

characters, or tokens, laying out how likely it is for each possible token in its vocabulary to come next. The mathematics dictate that each token must have a non-zero chance of being chosen, giving the model flexibility to learn new patterns, as well as the capacity to generate statements that are incorrect. The fundamental problem is that language models are probabilistic, while truth is not.

This tension manifests itself in a number of ways. One is that LLMs are not built to have perfect recall in the way a search engine or encyclopedia might. Instead, because the size of a model is much smaller than the size of its training data, it learns by compressing. The model becomes a blurry picture of its training data, retaining key features but at much lower resolution. Some facts resist blurring—“Paris”, for example, may always be the highest-probability token following the words “The capital of France is”. But many more facts that are less statistically obvious may be smudged away.

Further distortions are possible when a pretrained LLM is “fine-tuned”. This is a later stage of training in which the model’s weights, which encode statistical relationships between the words and phrases in the training data, are updated for a specific task. Hallucinations can increase if the LLM is fine-tuned, for example, on transcripts of conversations, because the model might make things up to try to be interesting, just as a chatty human might. (Simply including fine-tuning examples where the model says “I don’t know” seems to keep hallucination levels down.)

Tinkering with a model’s weights can reduce hallucinations. One method involves creating a deliberately flawed model trained on data that contradict the prompt or contain information it lacks. Researchers can then subtract the weights of the flawed model, which are in part responsible for its output, from those of the original to create a model which hallucinates less.

It is also possible to change a model's "temperature". Lower temperatures make a model more conservative, encouraging it to sample the most likely word. Higher temperatures make it more creative, by increasing the randomness of this selection. If the goal is to reduce hallucinations, the temperature should be set to zero. Another trick is to limit the choice to the top-ranked tokens alone. This reduces the likelihood of poor responses, while also allowing for some randomness and, therefore, variety.

Clever prompting can also reduce hallucinations. Researchers at Google DeepMind found that telling an LLM to "take a deep breath and work on this problem step-by-step" reduced hallucinations and improved problem solving, especially of maths problems. One theory for why this works is that AI models learn patterns. By breaking a problem down into smaller ones, it is more likely that the model will be able to recognise and apply the right one. But, says Edoardo Ponti at the University of Edinburgh, such prompt engineering amounts to treating a symptom, rather than curing the disease.

Perhaps, then, the problem is that accuracy is too much to ask of LLMs alone. Instead, they should be part of a larger system—an engine, rather than the whole car. One solution is retrieval augmented generation (RAG), which splits the job of the AI model into two parts: retrieval and generation. Once a prompt is received, a retriever model bustles around an external source of information, like a newspaper archive, to extract relevant contextual information. This is fed to the generator model alongside the original prompt, prefaced with instructions not to rely on prior knowledge. The generator then acts like a normal LLM and answers. This reduces hallucinations by letting the LLM play to its strengths—summarising and paraphrasing rather than researching. Other external tools, from calculators to search engines, can also be bolted onto an LLM in this way, effectively building it a support system to enhance those skills it lacks.

Even with the best algorithmic and architectural antipsychotics available, however, LLMs still hallucinate. One leaderboard, run by Vectara, an American software company, tracks how often such errors arise. Its data shows that GPT-4 still hallucinates in 3% of its summaries, Claude 2 in 8.5% and Gemini Pro in 4.8%. This has prompted programmers to try detecting, rather than preventing, hallucinations. One clue that a hallucination is under way lies in how an LLM picks words. If the probability distribution of the words is flat, ie many words have similar likelihoods of being chosen, this means that there is less certainty as to which is most likely. That is a clue that it might be guessing, rather than using information it has been prompted with and therefore “knows” to be true.

Another way to detect hallucination is to train a second LLM to fact-check the first. The fact-checker can be given the “ground truth” along with the LLM’s response, and asked whether or not they agree. Alternatively, the fact-checker can be given several versions of the LLM’s answer to the same question, and asked whether they are all consistent. If not, it is more likely to be a hallucination. NVIDIA, a chipmaker, has developed an open-source framework for building guardrails that sit around an LLM to make it more reliable. One of these aims to prevent hallucinations by deploying this fact-checking when needed.

Although such approaches can decrease the hallucination rate, says Ece Kamar, head of the AI frontiers lab at Microsoft, “it is unclear whether any of these techniques is going to completely get rid of hallucinations.” In many cases, that would be akin to self-sabotage. If an LLM is asked to generate ideas for a fantasy novel, for example, its output would be disappointing if limited to the world as it is. Consequently, says Dr Kamar, her research aims not to get rid of all hallucinations, but rather to stop the model from hallucinating when it would be unhelpful.

The hallucination problem is one facet of the larger “alignment” problem in the field of AI: how do you get AI systems to reliably do what their human users intend and nothing else? Many researchers believe the answer will come in training bigger LLMs on more and better data. Others believe that LLMs, as generative and probabilistic models, will never be completely rid of unwanted hallucinations.

Or, the real problem might be not with the models but with its human users. Producing language used to be a uniquely human capability. LLMs’ convincing textual outputs make it all too easy to anthropomorphise them, to assume that LLMs also operate, reason and understand like humans do. There is still no conclusive evidence that this is the case. LLMs do not learn self-consistent models of the world. And even as models improve and the outputs become more aligned with what humans produce and expect, it is not clear that the insides will become any more human. Any successful real-world deployment of these models will probably require training humans how to use and view AI models as much as it will require training the models themselves. ■



硅之梦

人工智能模型有时胡言乱语。如何控制幻觉？

不限制模型的能力很难做到【深度】

我们已经越来越熟悉这样的情形。向OpenAI的ChatGPT这样的大语言模型提问时，模型会马上给出言之凿凿、流畅连贯却完全错误的回复。在AI模型中，这种倾向通常被称为幻觉。不过还有一个更通俗的说法：其实就是胡说八道，满嘴放炮。

当然，也有更委婉的说法。OpenAI在给用户的说明中警告ChatGPT“可能会出错”。美国AI公司Anthropic表示，其大语言模型Claude“可能会展示不正确或有害的信息”，谷歌的Gemini则提醒用户“要复核它给出的回复”。总的主题就是，无论AI生成的文本有多流利自信，你都不能轻易相信。

幻觉让人们很难在现实世界中依赖AI系统。生成新闻的算法中的错误可能会传播不实信息。图像生成器生成的画作可能会侵犯版权，即使指令明确要求不能侵权。客户服务聊天机器人可能会在没有权限的情况下向客户承诺退款。（2022年，加拿大航空公司的聊天机器人自己编造了一项丧亲机票优惠政策。今年2月，加拿大的一个法院最终判决该航空公司必须承担这笔承诺退款。）此外，用于诊断或开药的AI系统产生的幻觉还可能害死人。

| 树叶已枯黄

麻烦在于，让模型产生幻觉的能力也正是模型能发挥巨大作用的原因。具体来说，大语言模型是一种“生成式”AI，顾名思义就是它们通过生成内容来解决新问题。具体做法是预测大量字符块也就是“token”的概率分布，列出其词汇库中每一个可能的字符块接下来出现的概率。数学原理决定了每个字符块被选中的概率都不会是零，这让模型具备了灵活性，能够学习新模式，但同时也能生成不正确的内容。这里的根本问题是语言模型是基于概率的、或然的，而事实不是。

这种矛盾表现在多个方面。其中之一是大语言模型没有搜索引擎或百科全书那样的完美查全率。相反，由于模型的大小要比其训练数据小得多，它是通过压缩来学习的。模型这样就成了其训练数据的一张模糊化的图片，保留了原图的关键特征，但分辨率要低得多。有些事实不容易受模糊化的影响——例如在“法国的首都是”之后，“巴黎”可能总是出现概率最高的字符块。但许多统计概率没那么高的事实可能会被模糊掉。

在对预训练的大语言模型做“微调”时，可能会出现更多扭曲。微调是训练的后期阶段，在此阶段，要根据拟完成的任务更新模型的权重（对训练数据中单词和短语之间的统计关系所做的编码）。举例来说，如果大语言模型的微调是基于对话笔录，幻觉可能就会增加，因为模型可能会凭空捏造一些内容来让自己显得风趣，就像一个爱聊天的人可能会信口开河那样。（只需增添模型表示“我不知道”的微调示例，似乎就能减少幻觉的出现。）

调整模型权重可以减少幻觉，一种方法是使用与提示相矛盾或包含其所缺乏的信息的数据，特意训练出一个缺陷模型。研究人员随后可以从原始模型中剔除掉缺陷模型的权重（这些权重对输出有一定影响），从而创造出幻觉更少的模型。

另外，也可以通过调整模型的“温度”来减少幻觉。降低温度会鼓励模型选择出现概率最高的词语，让它更保守。提高温度会提高模型选择的随机性，让它更具创造性。如果目标是减少幻觉，那么温度应该设定为零。另一种做法是将模型的选择范围限于概率最靠前的字符块。这样可以降低出现低质量回复的可能性，同时也能允许一定的随机性，进而也就有了多样性。

此外，巧妙的提问也可以减少幻觉。谷歌DeepMind的研究人员发现，告诉一个大语言模型“深呼吸，一步步地解决这个问题”减少了幻觉，提高了问题解决能力，尤其是数学问题。这种方法之所以奏效，一种猜测是AI模型会学习模式。通过将问题分解成小问题，模型更有可能识别并应用正确的模式。但是，爱丁堡大学的爱德华多·庞蒂（Edoardo Ponti）认为，像

这样优化提示是治标不治本。

那么，问题也许在于，希望大语言模型自己来保证准确性是要求太高了。其实它们应该作为一个更大的系统的一部分——它们是引擎而不是整辆车。一种解决方案是检索增强生成（RAG），它把AI模型的任务分成检索和生成两个部分。一旦接收到提示，检索器模型就会马上从诸如报刊档案的外部信息源中提取相关的情境信息。这些信息与原始提示一起输入到生成器模型，还要在开头加上不要依赖先验知识的指示。接下来生成器就像一个普通大语言模型那样回答问题。RAG让大语言模型发挥自身强项——总结和转述，而非研究，从而减少了模型的幻觉。计算器和搜索引擎等其他外部工具也可以以这种方式附加到大语言模型上，相当于构建了一个支持系统来增强模型所缺乏的技能。

然而，即使有了最好的算法和架构设计来防范，大语言模型仍然会出现幻觉。美国软件公司Vectara制作的排行榜追踪不同模型出现这类错误的频率。排行数据显示，GPT-4在做摘要时的幻觉率为3%，Claude 2为8.5%，Gemini Pro为4.8%。这已经促使程序员更关注发现幻觉而不是阻止幻觉的发生。幻觉可能出现的一个线索是大语言模型选择词语的方式。如果概率分布曲线平坦（即多个词语被选中的可能性相近），这时模型不太确定最应该选择哪个词语。这个线索表明模型可能正在猜答案，而不是根据所提示的信息而“知道”正确答案。

发现幻觉的另一种方法是训练另一个负责核查事实的大语言模型。可以将“基准事实”和第一个大语言模型的回复一起输入给第二个模型，并询问它们是否相一致。或者，也可以把第一个大语言模型对同一个问题的多个回复输入给第二个，询问它这些回复彼此是否统一。如果不一致，那么第一个大语言模型就很可能产生了幻觉。芯片制造商英伟达开发了一个开源框架，给大语言模型安装“护栏”以提高其可靠性。目标之一就是在需要时做这样的事实核查，防止模型产生幻觉。

微软AI前沿实验室（AI frontiers）的负责人艾斯·卡玛（Ece Kamar）认为，尽管这些方法可以降低幻觉率，但“尚不清楚是否其中任何一种方法

能够完全消除幻觉”。在许多情况下，完全消除幻觉无异于自我破坏。例如，假设让一个大语言模型为一部奇幻小说提供思路，如果限制它只能依据现实世界提供输出，其结果将会令人失望。因此，卡玛说她的研究目的不是为了消除所有幻觉，而是在幻觉无益时阻止模型产生幻觉。

｜安全又温暖

幻觉问题是AI领域更大的“对齐”问题的一个方面，“对齐”指如何让AI系统可靠地执行人类用户的意图，而不会擅作主张。许多研究人员认为用更多更好的数据训练更大的大语言模型能为这个问题提供答案。其他人则认为大语言模型作为生成和概率模型，将永远无法完全消除无用的幻觉。

或者，真正的问题可能不在于模型，而在于人类用户。产生语言曾经是人类独一无二的 ability。大语言模型颇具说服力的文本输出使它们很容易被赋予人性，让人以为它们也像人类那样运作、推理和理解。但目前还没有确凿证据表明事情确实如此。大语言模型并不学习关于整个世界的自洽模型。而且，在模型得到改进、其输出越来越对齐人类的产出结果和期望的同时，并不清楚模型内部是否会变得更富人性。要在现实世界中成功部署大语言模型，可能不仅需要训练模型本身，还需要训练人类如何使用和看待模型。■



Motor no-shows

Car shows in the West are in terminal decline

Chinese firms are keeping them on life support

SHIFTING TRENDS in the car industry are best illustrated by the plight of Geneva's hoteliers. Securing a place to stay during the city's annual motor show once required booking at least a year ahead and paying an extortionate rate for even basic accommodation. This year rooms in the city were plentiful. Organisers' hopes for 200,000 visitors seem optimistic. An event that once no car firm could afford to miss, and that attracted crowds of over 700,000, was attended sparsely—chiefly, it seemed, by journalists who got a preview before the show opened to the paying public on February 28th. The few companies that turned up were characterised by one car executive as “China, China, China and Renault”. They occupied less than a quarter of the floor space of previous years.

Geneva is not the only big car event in decline. Frankfurt and Paris, which took it in turns to host Europe's other big annual show, have dramatically downsized. The Paris jamboree of 2022 was a shadow of its former self. The German bash moved to Munich and became smaller. In both cases domestic firms were outshone by Chinese rivals such as BYD and Great Wall Motor.

America's premier event in Detroit has also suffered an exodus of firms. A scheduling shift from the depths of Michigan's harsh winter to the summer months has failed to revive interest. It will now revert to its winter date in 2025. But the changing nature of cars, which rely on software to lure buyers as much as sleekly bent metal, makes events such as the Consumer Electronics Show in Las Vegas more appealing showcases for manufacturers.

Reduced relevance makes the high price tag of turning up at shows harder for carmakers to stomach. Elaborate stands cost millions to build and then reconstruct at the next event halfway across the world. Car firms started to question the logic of competing to unveil the latest models at a single event when they could command more attention with dedicated launches or significant announcements at a time and place of their choosing.

Often such occasions now coincide with motor shows. On the eve of Geneva's opening, Fiat (part of Stellantis, whose biggest shareholder part-owns The Economist's parent company) released a video of its new Panda concept cars. (Reports on February 27th that Apple was abandoning its ten-year effort to build an electric car were probably not intended as a spoiler.) As a way of attracting buyers, cash can these days be far more effectively spent on marketing through other means, especially on social media.

In the absence of prominent Western marques—many of which still bitterly resent the Swiss organisers for calling off the show in 2020 at the last minute as covid-19 was spreading and then refusing to issue refunds—the atmosphere in Geneva this year was dull even by Switzerland's exacting standards. Renault drummed up some buzz with its new electric Renault 5, an affordable, neatly styled runaround. Another company to make a splash was BYD, with the European debut of the Yangwang U8, a luxury hybrid SUV with the amphibious ability to move on water. These days it is mostly Chinese carmakers that are keeping Europe's big shows afloat. ■



无车可秀

西方车展正无可挽回地走向衰落

中国车企正在为它们“续命”

日内瓦酒店经营者的窘境最能说明汽车行业的变化趋势。过去，要在日内瓦一年一度的车展期间找到一个住处至少需要提前一年预订，即使是最基本的住宿也价格高昂。今年，日内瓦的客房却非常充裕。组织者们对20万参观者的预期已是显得乐观。曾经这个任何车企都不容错过的盛会能吸引70多万人，今年却乏人问津——来的人似乎主要是记者，他们在展会于2月28日向公众收费开放之前已先睹为快。寥寥几家参展公司按一位汽车行业高管的描述是“中国、中国、中国和雷诺”。它们所占的展厅面积还不到往年的四分之一。

日内瓦车展并不是唯一走向衰落的汽车盛会。法兰克福和巴黎曾轮流主办欧洲另一个大型年度车展，如今也大幅缩减了规模。2022年的巴黎车展大不如前。而德国将车展移师慕尼黑，规模也变得更小。在这两个展会上，本土企业都被比亚迪和长城汽车等中国竞争对手盖过了风头。

在底特律举行的美国顶级车展也遭遇了车企流失。车展将举办时间从密歇根州的严冬改到了夏季月份，仍未能让人们重燃兴趣。如今，该车展将于2025年恢复到冬季举办。但是，由于汽车的性质正在发生变化，要吸引买家不仅要靠线条优美流畅的金属车身，也要靠软件，这就使得拉斯维加斯消费电子展等活动对制造商来说更具吸引力了。

车展热度降低，汽车制造商也就更难接受参展的高昂费用。精心布置的展台要耗费数百万美元，下一次跨越半个地球参加另一场展会时又要重新搭建。车企开始质疑竞相在同一个展会上发布最新车型是否明智，因为它们大可自行选择时间和地点举行专门的发布会或宣布重大消息，这样能吸引更多关注。

这样的活动如今常与车展撞期。在日内瓦车展开幕前夕，菲亚特

（Stellantis旗下品牌，Stellantis的最大股东部分持股本刊母公司）发布了新款熊猫概念车的视频。（2月27日有报道称苹果公司将放弃十年来研发电动汽车的努力，可能并不是有意抢戏吧。）如今要吸引买家，把钱花在其他营销手段上（尤其是社交媒体）效果可能要好得多。

由于西方著名品牌的缺席（它们当中有许多仍对2020年瑞士主办方因新冠肺炎蔓延而在最后一刻取消车展并拒绝退款的做法极其反感），即使以瑞士的严格标准来看，今年日内瓦车展的气氛也堪称沉闷。雷诺推出的新款电动车雷诺5引起了一些反响，这款小轿车价格实惠、造型简洁。另一家引起轰动的公司是比亚迪，它首次向欧洲消费者展示了仰望U8车型，这款具水陆两栖能力的豪华混合动力SUV能在水上行进。如今，主要还是靠着中国车企这跟浮木，欧洲的大型车展才不至于全沉了。■



Your brain on music

Why recorded music will never feel as good as the real thing

The answer, according to neuroscience

IN A WORLD of music streaming services, access to almost any song is just a few clicks away. Yet, the live gig lives on. People still fill sweaty basements, muddy fields and gilded concert halls to hear their favourite musicians play. And now neuroscientists might know why: live music engages the brain's emotion centres more than its recorded counterpart.

Concerts are immersive social experiences in which people listen to and feel the music together through crescendos, key changes and drops. They are also dynamic—artists can adapt their playing according to the crowd's reaction.

It was this last difference that led neuroscientists, based at the Universities of Zurich and Oslo, to study the brain responses of people listening to music. In the “live” experiment, participants lay in an MRI scanner listening to the music through earphones, while a pianist was positioned outside the room. The pianist was shown the participant's real-time brain activity as a form of feedback. In the recorded condition, participants listened to pre-recorded versions of the same tunes.

The scientists were interested in how live music affected the areas of the brain that process emotions. In the live condition pianists were instructed to try and modulate their playing in order to drive the activity in one of these regions known as the amygdala, an almond-shaped area deep inside the brain.

The results, just published in the journal PNAS, showed that live music had far more emotional impact. Whether the music was happy or sad, listening

to the pianist playing in a dynamic way generated more activity in both the amygdala and other parts of the brain's emotion processing network. The researchers also found that participants' brain activity tracked the acoustic features of the music, like tempo and pitch, far more closely when it was played live.

The study was far from replicating the real experience of a gig, and the authors noted that the live music ended up sounding quite different from the recorded tracks, which may have driven some of the differences in participant's brain activity. But the results indicate that the ability of artists to change the way they play in response to the audience may be one aspect of what makes live music special. Some musical acts now attempt to recreate the real gig experience with everything but the artist—ABBA Voyage is a social, immersive show performed entirely by pre-recorded hologram avatars. But without Benny's ability to read the mood of the room, it will never quite match the real thing. ■



你的音乐大脑

为什么听录制音乐的感受永远比不上听现场

来自神经科学的答案

在一个音乐流媒体服务大行其道的世界里，点几下鼠标就能听到几乎任何歌曲。但现场演出仍在继续。人们仍然挤在闷热的地下室、泥泞的户外和金碧辉煌的音乐厅里欣赏他们钟爱的音乐家演奏。现在神经科学家可能找到了原因：现场音乐比录制音乐更能调动大脑的情感中心。

音乐会是一种沉浸式的社交体验，在声音渐强、音调变化和高潮部分的变换之中，人们共同聆听并感受音乐。音乐会还是动态的——艺术家可以根据听众的反应调整他们的演奏。

正是这最后一个不同促使苏黎世大学和奥斯陆大学的神经科学家们研究人们听音乐时大脑的反应。在“现场”实验中，参与者躺在核磁共振扫描仪中，戴着耳机听音乐，同时一位钢琴家坐在房间外面演奏。参与者的实时大脑活动作为反馈被展示给钢琴家。在录音组，参与者听的是预先录好的相同的曲子。

科学家们对现场音乐如何影响大脑中处理情绪的区域很感兴趣。在现场组，研究人员指示钢琴家试着调整演奏来驱动这些区域中一个叫做“杏仁核”的部分，这是大脑深处一个杏仁状的区域。

这一刚刚发表在《美国国家科学院院刊》（PNAS）上的研究表明，现场音乐的情感冲击力要大得多。无论音乐是快乐还是悲伤，钢琴家的动态演奏都会让杏仁核和大脑情绪处理网络的其他部分产生更多的活动。研究人员还发现，在听现场演奏时，参与者的的大脑活动对音乐声学特征的追踪要紧密得多，比如对节奏和音高。

这项研究离复制演出的真实体验还差得远，而且作者指出，现场版音乐最终听起来与录音室版本差异很大，这可能也导致参与者的的大脑活动产生了

一些不同。但研究表明，艺术家能够根据观众的喜好改变演奏方式可能是令现场音乐与众不同的一个方面。现在，有些音乐演出试图“再造”真实的演出体验，它们具备现场演出的一切要素，独独省去了艺术家真身——ABBA乐队的Voyage演唱会是一场联欢式的沉浸式演出，完全由预先录制的全息图像作为替身表演。但没了乐队主唱班尼体察现场气氛的能力，它永远没法跟真人演出相比。 ■



Meet your new copilot

How businesses are actually using generative AI

Some experiments with chatbots are more useful than others

IT HAS BEEN nearly a year since OpenAI released GPT-4, its most sophisticated artificial-intelligence model and the brain-of-sorts behind ChatGPT, its groundbreaking robot conversationalist. In that time the market capitalisation of America's technology industry, broadly defined, has risen by half, creating \$6trn in shareholder value. For some tech firms, growing revenue is starting to match sky-high share prices. On February 21st Nvidia, which designs chips used to train and run models like GPT-4, reported bumper fourth-quarter results, sending its market value towards \$2trn. AI mania has also lifted the share prices of other tech giants, including Alphabet (Google's corporate parent), Amazon and Microsoft, which are spending big on developing the technology.

At the same time, big tech's sales of AI software remain small. In the past year AI has accounted for only about a fifth of the growth in revenues at Azure, Microsoft's cloud-computing division, and related services. Alphabet and Amazon do not reveal their AI-related sales, but analysts suspect they are lower than those of Microsoft. For the AI stockmarket boom to endure, these firms will at some point need to make serious money from selling their services to clients. Businesses across the world, from banks and consultancies to film studios, have to start using ChatGPT-like tools on a large scale. When it comes to real-world adoption of such "generative" AI, companies have trodden gingerly. Yet even these baby steps hint at the changing nature of white-collar work.

Previous technological breakthroughs have revolutionised what people do in offices. The spread of the typewriter put some workers out of a job: "With the aid of this little machine an operator can accomplish more

correspondence in a day than half a dozen clerks can with the pen, and do better work,” said an observer in 1888. The rise of the computer about a century later eliminated some low-level administrative tasks even as it made highly skilled employees more productive. According to one paper, the computer explains over half the shift in demand for labour towards college-educated workers from the 1970s to the 1990s. More recently the rise of working from home, prompted by the covid-19 pandemic and enabled by video-conferencing, has changed the daily rhythms of white-collar types.

Could generative AI prompt similarly profound changes? A lesson of previous technological breakthroughs is that, economywide, they take ages to pay off. The average worker at the average firm needs time to get used to new ways of working. The productivity gains from the personal computer did not come until at least a decade after it became widely available. So far there is no evidence of an AI-induced productivity surge in the economy at large. According to a recent survey from the Boston Consulting Group (BCG), a majority of executives said it will take at least two years to “move beyond the hype” around AI. Recent research by Oliver Wyman, another consultancy, concludes that adoption of AI “has not necessarily translated into higher levels of productivity—yet”.

That is unsurprising. Most firms do not currently use ChatGPT, Google’s Gemini, Microsoft’s Copilot or other such tools in a systematic way, even if individual employees play around with them. A fortnightly survey by America’s Census Bureau asks tens of thousands of businesses whether they use some form of AI. This includes the newfangled generative sort and the older type that companies were using before 2023 for everything from improving online search results to forecasting inventory needs. In February only about 5% of American firms of all sizes said they used AI. A further 7% of firms plan to adopt it within six months (see chart). And the numbers conceal large differences between sectors: 17% of firms in the information

industry, which includes technology and media, say they use it to make products, compared with 3% of manufacturers and 5% of health-care companies.

When the Census Bureau began asking about AI in September 2023, small firms were likelier to use the technology than big ones, perhaps because less form-ticking made adoption easier for minnows. Today AI is most prevalent in big companies (with more than 250 employees), which can afford to enlist dedicated AI teams and to pay for necessary investments. A poll of large firms by Morgan Stanley, a bank, found that between the start and end of 2023 the share with pilot AI projects rose from 9% to 23%.

Some corporate giants are frantically experimenting to see what works and what doesn't. They are hiring AI experts by the thousand, suggest data from Indeed, a job-search platform (see chart). Last year Jamie Dimon, boss of JPMorgan Chase, said that the bank already had "more than 300 AI use cases in production today". Capgemini, a consultancy, says it will "utilise Google Cloud's generative AI to develop a rich library of more than 500 industry use cases". Bayer, a big German chemicals company, claims to have more than 700 use cases for generative AI.

This "use-case sprawl", as one consultant calls it, can be divided into three big categories: window-dressing, tools for workers with low to middling skills, and those for a firm's most valuable employees. Of these, window-dressing is by far the most common. Many firms are rebranding run-of-the-mill digitisation efforts as "gen AI programmes" to sound more sophisticated, says Kristina McElheran of the University of Toronto. Presto, a purveyor of restaurant tech, introduced a gen-AI assistant to take orders at drive-throughs. But fully 70% of such orders require a human to help. Spotify, a music-streaming firm, has rolled out an AI disc-jockey which selects songs and provides inane banter. Recently Instacart, a grocery-delivery company, removed a tool that generated photos of vendors' food,

after the AI showed customers unappetising pictures. Big tech firms, too, are incorporating their own AI breakthroughs into their consumer-facing offerings. Amazon is launching Rufus, an AI-powered shopping assistant that no shopper really asked for. Google has added AI to Maps, making the product more “immersive”, whatever that means.

Tools for lower-skilled workers could be more immediately useful. Some simple applications for things like customer service involve off-the-shelf AI. Most customers’ questions are simple and concern a small number of topics, making it easy for companies to train chatbots to deal with them. A few of these initiatives may already be paying off. Amdocs produces software to help telecoms companies manage their billing and customer services. The use of generative AI, the company says, has reduced the handling time of customers’ calls by almost 50%. Sprinklr, which offers similar products, says that recently one of its luxury-goods clients “has seen a 25% improvement” in customer-service scores.

Routine administrative tasks likewise look ripe for AI disruption. The “top examples” of Bayer’s 700 use cases include mundane jobs such as “easily getting data from Excel files” and “creating a first draft in Word”. Some companies are using generative AI as cleverer search. At Nasdaq, a financial-services firm, it helps financial-crime sleuths gather evidence to assess suspicious bank transactions. According to the company, this cuts a process which can take 30-60 minutes to three minutes.

Giving AI tools to a firm’s most valuable workers, whose needs are complex, is less widespread so far. But it, too, is increasingly visible. Lawyers have been among the earliest adopters. Allen & Overy, a big law firm, teamed up with Harvey, an AI startup, to develop a system that its lawyers use to help with everything from due diligence to contract analysis. Investment banks are using AI to automate part of their research process. At Bank of New York Mellon an AI system processes data for the bank’s analysts overnight and

gives them a rough draft to work with in the morning. “So rather than getting up at four in the morning to write research, they get up at six,” the bank says. Small mercies. Sanofi, a French drugmaker, uses an AI app to provide executives with real-time information about many aspects of the company’s operations.

Some companies are using the technology to build software. Microsoft’s GitHub Copilot, an AI coding-writing tool, has 1.3m subscribers. Amazon and Google have rival products. Apple is reportedly working on one. Fortive, a technology conglomerate, says that its operating companies “are seeing a greater-than-20% acceleration in software-development time through the use of gen AI”. Chirantan Desai, chief operating officer of ServiceNow, a business-software company, has said that GitHub Copilot produces “single-digit productivity gains” for his firm’s developers. With the help of AI tools, Konnectify, an Indian startup, went from releasing four apps per month to seven. Surveys from Microsoft suggest that few people who start using Copilot want to give it up.

Pinterest, a social-media company, says it has improved the relevance of users’ search results by ten percentage points thanks to generative AI. On a recent earnings call its boss, Bill Ready, said that new models were 100 times bigger than the ones his firm used before. L’Oréal, one of the world’s largest cosmetics firms, has caught the eye of investors as it improves BetIQ, an internal tool to measure and improve the company’s advertising and promotion. L’Oréal claims that generative AI is already generating “productivity increases of up to 10-15% for some of our brands that have deployed it”.

This does not mean that those brands will need 10-15% fewer workers. As with earlier technological revolutions, fears of an AI jobs apocalypse look misplaced. So far the technology appears to be creating more jobs than it eliminates. A survey published in November by Evercore ISI, a bank, found

that just 12% of corporations believed that generative AI had replaced human labour or would replace it within 12 months. Although some tech firms claim to be freezing hiring or cutting staff because of AI, there is little evidence of rising lay-offs across the rich world.

Generative AI is also generating new types of white-collar work. Companies including Nestlé, a coffee-to-cat-food conglomerate, and KPMG, a consultancy, are hiring “prompt engineers” expert at eliciting useful responses from AI chatbots. One insurance firm employs “explainability engineers” to help understand the outputs of AI systems. A consumer-goods firm that recently introduced generative AI in its sales team now has a “sales-bot manager” to keep an eye on the machines.

Though such developments will not translate into overall productivity statistics for a while, they are already affecting what white-collar workers do. Some effects are clearly good. AI lets firms digitise and systematise internal data, from performance reviews to meeting records, that had previously remained scattered. Respondents to surveys conducted by Randy Bean, a consultant, reported big improvements in establishing an internal “data and analytics culture”, which plenty of businesses find stubbornly difficult to nurture.

AI adoption may also have certain unpredictable consequences. Although AI code-writing tools are helping software engineers do their jobs, a report for GitClear, a software firm, found that in the past year or so the quality of such work has declined. Programmers may be using AI to produce a first draft only to discover that it is full of bugs or lacking concision. As a result, they could be spending less time writing code, but more time reviewing and editing it. If other companies experience something similar, the quantity of output in the modern workplace may go up—as AI churns out more emails and memos—even as that output becomes less useful for getting stuff done.

Polling by IBM, a tech firm, suggests that many companies are cagey about adopting AI because they lack internal expertise on the subject. Others worry that their data is too siloed and complex to be brought together. About a quarter of American bosses ban the use of generative AI at work entirely. One possible reason for their hesitance is worry about their companies' data. In their annual reports Blackstone, a private-equity giant, and Eli Lilly, a pharmaceutical one, have warned investors about AI-related risks such as possible leakage of intellectual property to AI model-makers. Last year Marie-Hélène Briens Ware, an executive at Orange, a telecoms company, explained that the firm had put data guardrails in place before commencing a trial with Microsoft's Copilot.

Ultimately, for more businesses to see it as an open-and-shut case, generative AI still needs to improve. In November Microsoft launched a Copilot for its productivity software, such as Word and Excel. Some early users find it surprisingly clunky and prone to crashing—not to mention cumbersome, even for people already adept at Office. Many bosses remain leery of using generative AI for more sensitive operations until the models stop making things up. Recently Air Canada found itself in hot water after its AI chatbot gave a passenger incorrect information about the airline's refund policy. That was embarrassing for the carrier, but it is easy to imagine something much worse. Still, even the typewriter had to start somewhere. ■



见见您的新副手

企业对生成式AI的实际使用情况如何

一些运用聊天机器人的尝试比其他的更有用【深度】

距OpenAI发布GPT-4已将近一年，GPT-4是其最先进的人工智能模型，也是其开创性的聊天机器人ChatGPT背后的机器大脑。在那段时间里，广义上的美国科技业的市值增长了一半，创造了6万亿美元的股票价值。对于一些科技公司来说，不断增长的收入开始与天价股价相匹配。2月21日，设计芯片供训练和运行GPT-4等模型的英伟达公布了出色的第四季度业绩，将其市值推至近2万亿美元。AI热潮也推高了其他科技巨头的股价，包括Alphabet（谷歌的母公司）、亚马逊和微软，它们正在砸巨资开发这项技术。

与此同时，科技巨头卖出的AI软件仍然不多。过去一年里，AI仅贡献了微软云计算部门Azure和相关服务收入增长的五分之一左右。Alphabet和亚马逊没有透露其与AI相关的销售额，但分析师怀疑都要低于微软。要想让AI带来的股市繁荣持续下去，这些公司到了某个时间点将必须要能从向客户出售相应服务中赚到可观的收入。从银行和咨询公司到电影制片厂，世界各地的企业都必须开始大规模使用类似ChatGPT的工具。对于在现实世界里应用这种“生成式”AI，各类公司小心翼翼地迈出了步子。但即使是这些婴儿学步也暗示了白领工作的性质在发生变化。

以往的技术突破变革了办公室工作。打字机的普及使一些人失业：“借助这台小机器，一个打字员一天完成的信件比六个文员用笔写的更多，并且做得更好。”一位观察家在1888年说。大约一个世纪后，计算机的兴起消除了一些低级行政管理工作，同时也提高了高技能员工的产出。一篇论文发现，劳动力需求在1970年代到1990年代向受过大学教育的工人转移，其中一半以上是受计算机的推动。最近，在新冠疫情的推动和视频会议的助力下，居家办公兴起，改变了白领员工的日常节奏。

生成式AI能否引发类似的深刻变化？从以往的技术突破中得到的一个教训

是，从整体经济的角度来看，它们需要很长时间才能得到回报。普通公司的普通工人需要时间来适应新的工作方式。个人电脑在普及至少十年后才带来了生产率提升。到目前为止还没有迹象表明AI在经济总体中引发了生产率的大幅提升。波士顿咨询公司最近的一项调查显示，大多数高管表示，至少需要两年时间，AI才能“走出炒作阶段”。另一家咨询公司奥纬咨询（Oliver Wyman）最近的研究认为，AI的使用“未必已然转化为更高水平的生产力”。

这并不让人意外。大多数公司目前没有系统地使用ChatGPT、谷歌的Gemini、微软的Copilot或其他此类工具，即使有个别员工在摆弄它们。美国人口普查局每两周进行一次调查，询问数以万计的企业是否使用某种形式的AI。这包括新型的生成式AI，也包括那些旧式AI工具——企业在2023年前使用这些旧工具来改进在线搜索结果、预测库存需求等各种工作。今年2月，在各种规模的美国公司中，只有约5%表示它们使用了AI；另有7%的公司计划在六个月内使用AI（见图表）。这些数字掩盖了行业之间的巨大差异：在包括科技和媒体在内的信息产业中，有17%的公司表示在用AI制造产品，而制造业和医疗保健行业中的比例分别为3%和5%。

当人口普查局在2023年9月开始询问AI的使用情况时，小公司比大公司更有可能使用这项技术，这可能是因为小公司内部的管理程序更精简，因而更便于使用AI。如今，AI在大公司（拥有超过250名员工）中最为普遍，这些公司有能力和专门的AI团队并做出必要的投资。摩根士丹利对大公司的一项调查发现，从2023年初到年底，有AI试点项目的公司比例从9%上升到了23%。

一些企业巨头正在疯狂地实验，看看哪些做法有效，哪些无效。求职平台Indeed的数据显示，它们正在招聘数以千计的AI专家（见图表）。去年，摩根大通的老板杰米·戴蒙（Jamie Dimon）表示，该银行“目前在经营中已经有了300多个AI用例”。咨询公司凯捷（Capgemini）表示，它将“利用谷歌云的生成式AI来开发一个有500多个行业用例的丰富案例库”。德国大型化工公司拜耳声称拥有700多个生成式AI用例。

有顾问称之为“用例大扩张”。它可以分为三大类：装点门面、为中低技能员工提供工具，以及为公司里最有价值的员工提供工具。其中，最常见的绝对是装点门面。多伦多大学的克里斯蒂娜·麦克赫兰（Kristina McElheran）说，许多公司正在将普通的数字化工作重新包装成“生成式AI程序”，好让它听起来更高端些。餐厅科技供应商Presto推出了一个生成式AI助手，让人们不用下车就能点单，但足足70%的此类点单需要人工帮助。音乐流媒体公司Spotify推出了一个AI DJ，可以选择歌曲并开些无聊的玩笑。最近，杂货配送公司Instacart删除了一个能生成供应商食物照片的AI工具，因为它向顾客展示了倒人胃口的照片。科技巨头们也在将自己的AI突破融入到面向消费者的产品中。亚马逊推出了Rufus，这是一款AI驱动的购物助手，但实际上并没有购物者找它帮忙。谷歌在谷歌地图中添加了AI功能，使这个产品更加“身临其境”，天知道这是什么意思。

为低技能工人提供的工具可能有更直接的用处。客户服务等一些简单的应用程序用到了现成的AI。大多数客户的问题很简单，只是围绕少量主题，这使得公司很容易训练聊天机器人来处理这些问题。其中一些举措可能已经取得了成效。Amdocs开发软件来帮助电信公司管理计费 and 客户服务。该公司表示，生成式AI的使用将客户来电的处理时间缩短了近50%。提供类似产品的Sprinklr表示，最近它的一家奢侈品客户的客户服务得分“提高了25%”。

日常行政管理工作看起来也是时候被AI颠覆了。拜耳的700个用例中的“顶级示例”包括单调的工作，例如“轻松从Excel文件中获取数据”和“在Word中创建初稿”。一些公司正在使用生成式AI进行更聪明的搜索。在金融服务公司纳斯达克，生成式AI帮助金融犯罪侦探收集证据以评估可疑的银行交易。该公司称，这能把需要30到60分钟的过程缩短到3分钟。

公司里最有价值的员工有着复杂的需求，到目前为止，向这些人提供AI工具还不那么普遍，但也越来越常见了。律师是最早的使用者之一。大型律师事务所安理国际（Allen & Overy）与AI创业公司Harvey合作开发了一个系统，可以辅助其律师完成从尽职调查到合同分析等各种工作。投资银行正在使用AI来自动化一部分研究工作。在纽约梅隆银行（Bank of New

York Mellon），一个AI系统在夜间为银行分析师处理数据，为他们提供一份草稿，以便在早上使用。“这样他们就不用早上四点起床写研报了，现在六点起床就行了。”该银行说。能多睡会儿，你就偷着乐吧。法国制药公司赛诺菲（Sanofi）使用一个AI应用为高管提供有关公司运营的方方面面的实时信息。

一些公司正在使用该技术来构建软件。微软的AI编程工具GitHub Copilot拥有130万订阅用户。亚马逊和谷歌都有与之竞争的产品。据报道，苹果也在开发一个类似的产品。科技集团Fortive表示，它旗下公司“使用生成式AI，把软件开发时间加快了20%以上”。商业软件公司ServiceNow的首席运营官齐兰坦·德赛（Chirantan Desai）表示，GitHub Copilot为该公司的开发人员带来了“几个百分点的生产率提升”。在AI工具的帮助下，印度创业公司Konnnectify从每月发布四个应用提升到七个。微软的调查表明，很少有人开始使用Copilot后会想放弃它。

社交媒体公司Pinterest表示，依靠生成式AI，它已经将用户搜索结果的相关性提高了10个百分点。在最近的一次财报电话会议上，该公司的老板比尔·雷迪（Bill Ready）表示，新模型比他的公司以前使用的模型大100倍。全球最大化妆品公司之一的欧莱雅改进了它衡量和改善广告促销效果的内部工具BetIQ，引起了投资者的注意。欧莱雅声称，生成式AI已经为它“旗下部署了该技术的一些品牌带来了高达10%到15%的生产率提升”。

这并不意味着这些品牌需要的员工也会减少10%到15%。与之前的技术革命一样，对AI毁灭就业的担忧似乎是过虑了。到目前为止，这项技术创造的就业似乎多于消除的就业。投资银行Evercore ISI在去年11月发布的一项调查发现，只有12%的公司认为生成式AI已经取代或将在12个月内取代人类劳动。尽管一些科技公司声称会因为AI而冻结招聘或裁员，但几乎没有证据表明富裕国家的裁员正在增多。

生成式AI也正在催生新型白领工作。包括食品企业集团雀巢和咨询公司毕马威在内的公司正在聘请“提示工程师”，以从AI聊天机器人那里获得有用的回应。一家保险公司聘请了“可解释性工程师”来帮助理解AI系统的输

出。一家消费品公司最近在其销售团队中引入了生成式AI，现在它有一个“销售机器人经理”来监督这些机器。

尽管这些进展一时半会儿还不会体现在整体生产率的统计数据上，但它们已经影响了白领员工的工作。有一些效果显然是好的。AI使公司能够将从绩效评估到会议记录的内部数据数字化和系统化，这些数据以前一直散布在组织架构的各个角落。在咨询公司Randy Bean进行的调查中，受访者表示在建立内部“数据和分析文化”方面取得了巨大进步，许多企业都认为这种文化向来很难培养。

使用AI也可能产生某些不可预测的后果。尽管AI编程工具正在帮助软件工程师完成工作，但软件公司GitClear的一份报告发现，在过去一年左右的时间里，此类工作的质量下降了。程序员可能在使用AI生成初稿后，却发现其中满是错误或不够简洁。也因此，他们编写代码的时间可能是减少了，但审查和修改代码的时间却增加了。如果其他公司也遇到类似的情况，那么当今职场里的产出的数量可能会增加——因为AI大量生成更多的电子邮件和备忘录，而同时这些产出对于完成工作却变得不那么有用了。

科技公司IBM的调查显示，许多公司对使用AI持谨慎态度，因为它们内部缺乏这方面的专业能力。其他公司则担心它们的各类数据过于彼此孤立和复杂，无法整合在一起。大约四分之一的美国老板完全禁止在工作中使用生成式AI。他们犹豫不决的一个原因可能是担心自己公司的数据安全。私募股权巨头黑石集团（Blackstone）和制药巨头礼来（Eli Lilly）在年度报告中警告投资者注意与AI相关的风险，例如知识产权可能被泄露给AI模型开发者。去年，电信公司Orange的高管玛丽-伊莲娜·布里恩斯·韦尔（Marie-Hélène Briens Ware）解释说，该公司在开始试用微软的Copilot之前就已经设置了数据护栏。

归根结底，若要让更多企业将其视为一个简单易用的工具，生成式AI仍然需要改进。去年11月，微软在Word和Excel等生产率软件中推出了Copilot。一些早期用户发现它出奇地笨拙且容易崩溃——更别说用起来还很繁琐，即使对于已经熟练使用Office的人来说也是如此。许多老板仍然

对使用生成式AI进行更敏感的操作持怀疑态度，除非这些模型完全不会再胡说八道。最近，加拿大航空公司（Air Canada）碰到了麻烦，因为它的AI聊天机器人就其退款政策向一位乘客提供了错误信息。这让加拿大航空很尴尬，但不难想象完全可能发生比这糟糕得多的事。不过话说回来，凡事总有个开始，打字机当年也是从小范围开始被使用的。■



How high can markets go?

A golden age for stockmarkets is drawing to a close

Share prices may be surging, but even AI is unlikely to drive a repeat of the past decade's performance

STOCKMARKETS TEND to rise gradually; recently they have soared. American stocks are up by 21% since the end of October and stand roughly 5% above their vertiginous peak in January 2022. On February 22nd Europe's equities set a new record for the first time in two years. India has been enjoying a multi-year boom as optimism about its economy abounds. Even Japanese stocks—a byword for stagnation—have at last exceeded the level they reached in 1989 before a decades-long slump. It has been an extraordinary run. Since 2010, the S&P 500 index of American stocks has returned 11% a year in real terms.

These profits are all the more striking given what markets have had to contend with. The age of free money has been followed by two years of interest-rate rises—and even now bond investors are betting against imminent cuts. A trade war is raging between America and China; actual wars are raging in Ukraine, the Middle East and parts of Africa. Around the world, governments are turning away from free markets and globalisation in favour of industrial policy and protectionism. If all that has not extinguished this rally, whatever will?

One conclusion might be that a bubble is waiting to pop, especially in America. On Wall Street, valuations—the multiple by which profits are scaled up—are on average 80% as high as they were during the dotcom mania of the late-1990s and 90% as high as they climbed during 2021, before rock-bottom interest rates rose. Similar extremes are also to be seen in other measures, including concentration (the share of the stockmarket that is made up by the top firms) or value spreads (the valuation of the most

expensive companies compared with the cheapest). The value of the top 10% of American firms as a proportion of the whole market has not been as high since the crash that was one cause of the Depression in the 1930s. And don't forget the frothiest corner of the financial markets: bitcoin is trading around \$60,000 again, just shy of its peak in 2021.

Yet there are also reasons to see markets' exuberance as rational. As central banks all over the world tightened monetary policy at a pace not seen for a generation, many analysts warned about the danger of recessions and falling corporate profits. At the start of 2023 Wall Street savants predicted that in the year to come America's economy would grow by just 0.7%. In the event it achieved more than three times that amount. A broad range of firms are publishing strong results, including retailers, such as Walmart, and Japanese carmakers, such as Toyota.

The economy continues to defy gravity. A popular regular forecast of annualised American economic growth, published by the Federal Reserve Bank of Atlanta, stands at 3.2% for the first quarter of this year. Despite a slowdown in China—whose sagging markets are an exception to the global trend—the IMF has been nudging up its global growth forecasts, too.

Adding to investors' bullishness is their optimism about artificial intelligence (AI). This is not a ChatGPT-like hallucination. The event that propelled stocks into the stratosphere was the publication on February 22nd of earnings by Nvidia, which has an iron grip on the market for chips that are critical for training AI models. In October 2022, just before OpenAI released its now-celebrated chatbot, Nvidia earned around \$3bn in gross profits each quarter, mostly from selling graphics cards to gamers. In the three months to the end of January 2024 Nvidia earned \$17bn in gross profits while enjoying a margin of 76%. The company's share price has climbed five-fold over that time, but its earnings have grown even faster. In other words the enthusiasm that has lifted Nvidia close to a stockmarket

value of \$2trn is not built on dotcom-like hype, but cold, hard profit.

Judging the boom to be justified, though, does not make it wise to rush out and buy stocks. What happens next is unlikely to fill investors with glee. That is partly because the extreme excitement about AI extends beyond Nvidia to other members of the “Magnificent 7” group of tech stocks, such as Microsoft, whose eventual commercial strategies in the AI era are still far from clear. These firms are hoarding Nvidia’s chips in the belief that, one way or another, their AI businesses will boom. However, it remains to be seen how they will resolve basic issues with their large language models. Plenty of startups want to eat the Big 7’s lunch, and competition will keep profits in check—even, eventually, at Nvidia.

Techno-optimism is also the basis, in some quarters, for bullishness about economy-wide productivity growth. The lesson from other fundamental technologies is that it takes time to work out how to exploit them. Businesses talk non-stop about generative AI but it remains at the experimental stage. As a result, even if AI is destined to transform societies utterly, today’s investors may struggle to pick which companies will make money. Believers in the dotcom boom were not wrong about the transformative power of the internet—but they still lost their shirts.

If things stay sane this time, valuations will not climb much further. The trend of rising profits, as a share of the economy, also looks spent. Their outsize growth in recent decades was a one-off, caused by the falling cost of borrowing and taxes. As inflation lingers and government finances remain stretched, that fall cannot be repeated; it may even be reversed.

Under realistic assumptions about what will happen to valuations, interest and taxes, to generate even modest real equity returns of 4% a year over the next decade, America’s firms would need to increase their underlying profits by around 6% a year, close to their best ever post-war performance.

No wonder Warren Buffett, a veteran investor, sees “no possibility” of super returns for his fund.

| *The long and grinding road*

Equities could underwhelm in many ways. Perhaps AI-exhilaration will cause a dotcom-style bubble that pops. Another war or crisis could lead to a crash. Or prices may stagnate in a gentle bear market that takes years to reverse. Whatever the path to disappointment, in ten years’ time nobody will be repeating the obvious conclusion of today: that investors in equities—especially American ones—have enjoyed a golden age. ■



【首文】股市能涨到多高？

股市的黄金时代即将结束

股价可能还在飙升，但即便是AI也不太可能让股市重现过去十年的辉煌

股市通常是逐步上涨的，近年却是一路狂飙。自去年10月底以来，美国股市上涨了21%，比2022年1月令人目眩的峰值还要高出约5%。2月22日，欧洲股市两年来首次创下新高。而在印度，投资者对经济的乐观情绪已经令股市连涨了好几年。就连被视为“停滞”代名词的日本股市也终于超过了1989年的水平（那一年之后它便经历了长达数十年的低迷）。这一轮上涨非同寻常。自2010年以来，美股标普500指数的实际年回报率达到了11%。

鉴于市场遭遇了一系列麻烦，这样的利润水平愈显惊人。廉价资本时代之后是两年的利率上涨——即使到了现在债券投资者还在押注利率短期内不会下调。中美贸易战正酣，乌克兰、中东和非洲部分地区则在真枪实弹地激战。纵观全球，各国政府正在放弃自由市场和全球化，转而支持产业政策和贸易保护主义。如果这一切都没能压制股市的这波反弹，那还有什么能做到呢？

一个可能的结论是，泡沫早晚会破裂，尤其是在美国。在华尔街，市值（利润的倍数）平均而言达到了上世纪90年代末互联网狂热期的80%，以及2021年（利率触底反弹之前）的90%。类似的极端情况还体现在其他指标上，比如集中度（少数头部企业所占的股市份额）或价值差（估值最高与最低公司之间的差距）等。美国前10%公司的价值占整个市场的比例非常高，仅次于上世纪30年代股市暴跌前的水平，那次暴跌正是大萧条的成因之一。金融市场中泡沫最严重的一块也不能不提：比特币的交易价格重回六万美元左右，仅略低于2021年时的峰值。

不过，也有理由认为市场是理性繁荣。随着世界各国央行以二三十年来未曾见过的速度收紧货币政策，许多分析师警告称未来存在经济衰退和企业利润下降的风险。2023年初，华尔街的专家们预测，未来一年美国经济的

增长率只有0.7%。实际结果却是这一预测值的三倍多。包括沃尔玛等零售商和丰田等日本汽车制造商在内的一大批公司公布的业绩都很强劲。

经济仍然以无视地心引力的势头增长。亚特兰大联储定期发布的对美国经济年化增长率的预测广受认可，该预测认为今年第一季度美国的经济增长率为3.2%。尽管中国经济放缓（其股市是全球走势中的例外），国际货币基金组织也在小幅上调自己对全球经济增长的预测。

此外，对人工智能（AI）前景看好也增强了投资者的信心。这不是ChatGPT那种幻觉。英伟达牢牢控制着训练AI模型所用关键芯片的市场，它在2月22日发布的财报让股市一飞冲天。2022年10月，就在OpenAI发布其如今广受追捧的聊天机器人ChatGPT之前，英伟达每个季度的毛利润约为30亿美元——主要来自向游戏玩家出售显卡。而在截至2024年1月底的三个月里，它的毛利润达170亿美元，利润率高达76%。在此期间，英伟达的股价上涨了五倍，但利润增长得更快。换句话说，推动英伟达市值升至接近两万亿美元的投资热情并非建立在类似互联网泡沫的炒作之上，而是建立在真金白银的利润上。

不过，就算认为市场的繁荣合乎情理，也并不意味着匆匆入市购买股票是明智之举。接下来发生的事情可能并不会让投资者欢天喜地。原因之一是，对AI的过分狂热已经从英伟达蔓延到了科技股“七巨头”中的其他几家公司，比如微软，而它们还远未明确自己在AI时代的最终商业战略。它们囤积英伟达的芯片，深信自己的AI业务不管怎样总会兴旺起来。不过，它们将如何利用自己的大语言模型解决各种基本问题还有待观察。很多创业公司都想从七巨头那里抢食，而竞争也会压低利润——就算是英伟达，最终也不能幸免。

在某些群体中，技术乐观主义也是看好整个经济生产率增长的基础。其他基础性技术的发展历程告诉我们，弄清楚如何利用这些技术需要时间。虽然企业对生成式AI津津乐道，但它仍处于实验阶段。因此，即使AI注定会彻底改变社会，今天的投资者可能还是难以辨别哪些公司能赚到钱。当年相信互联网泡沫的人确实没有看错互联网的变革力量，但他们还是赔得血

本无归。

如果这一次市场保持理性，估值将不会再大幅上涨。利润占经济总规模之比不断增长的趋势似乎也后继无力。借贷成本下降和减税造就了过去几十年来这种绝无仅有的超速增长。但随着通胀挥之不去、政府财政依然捉襟见肘，那样的降息减税不但不可能重来，甚至还可能逆转。

依据对估值、利率和税收等因素未来变化的合理假设，未来十年内，美国企业即使想要让实际股本回报率达到每年4%这个不起眼的水平，也需要将基本利润每年提高6%左右——接近它们二战以来的最好表现。难怪投资大师巴菲特都认为自己的基金“绝不可能”获得超高回报。

| 道阻且长

股市可能会以多种方式令人失望。AI狂欢或许会引发像当年互联网泡沫一样的虚假繁荣，泡沫终会破裂。如果再来一场战争或者危机，股市可能崩溃；或者股价可能在温和的熊市中停滞不前，需要多年时间才能扭转。无论通往失望的大路是哪一条，十年后都不会有人再念叨今天这个显而易见的结论：股票投资者——尤其是美股投资者——经历了一个黄金时代。■



Indecipherable fingerprints

A secret room in Florence boasts drawings by Michelangelo

The artist's fingerprints are all over the walls—or are they?

STANDING IN THE mausoleum of the Medici Chapel in Florence, among tombs and elegant sculptures designed by Michelangelo, you could be excused for failing to notice the doorway to the left. It leads to a cloakroom, where a trap door opens to a flight of narrow steps. At the bottom is a small vaulted room. The eggshell-white walls are covered with graceful nude figures, a falling Phaethon (son of the sun god), a looming horse's head and several shapely legs drawn in shadowy charcoal—sketches, the museum claims, by the great Renaissance artist himself.

This stanza segreta (secret room) had long been closed. Only scholars and bigwigs like King Charles III and Leonardo DiCaprio were allowed entry by the National Museum of the Bargello, of which the chapel is a part. But in November the basement room was opened for a trial period of four months to 100 pre-booked visitors a week; they are allowed in for 15 minutes in groups of four. Tickets sold out. Officials recently extended the experiment until July.

Entering the room, which is bare except for ankle-level strips of LED lighting, is like stepping back in time—not just into the Renaissance, but also into an earlier era of tourism, when attractions lacked explanatory placards and were free from madding crowds. Most of all, it feels like entering someone's head, or their notebook, filled with drafts and doodles, some finely wrought, some cartoonish. You can imagine the middle-aged master in here, perhaps drawing by night. In several places, the walls appear to be scorched with black soot from a burning torch. Near a rendering of a muscular figure, there is even a smudged handprint.

Legend has it that Michelangelo holed up in this stanza segreta for a few months in 1530. Following a popular uprising three years before, the artist sided with those who preferred a republic to rule by the Medicis, Italy's wealthiest family, even though he had designed the mausoleum that bears their name. When the Medicis returned, the artist was in danger. This small room, conveniently furnished with a well, might have seemed like a safe haven.

The drawings were rediscovered in 1975 by the Medici Chapel's then director, who suspected that what was then a coal-storage room might be hiding something. Paola D'Agostino, director of the Bargello until mid-January, felt it was important to finally make the room accessible to the public. It shows "Michelangelo's human aspect", she says, though she admits that it is impossible to attribute the drawings securely to the master and concedes that some of the doodles are probably not his work.

Other experts doubt it was Michelangelo's hand who drew them. Antonio Forcellino, who has written two biographies of the artist, believes that none of the drawings is of high enough quality to merit firm attribution. William Wallace, a professor at Washington University in St Louis and a leading expert on Michelangelo, thinks that only a couple may be by the artist—a head that resembles the famous Laocoon statue Michelangelo had seen in Rome, as well as a standing male figure gracing a whole wall near the stairs—and that most were made by some of the 50 or so workmen employed to build the chapel above.

As for Michelangelo hiding in the room? "That's totally bogus," according to Mr Wallace. The already-famous artist is more likely to have been protected by a friend. Yet Mr Wallace argues that the room is a landmark finding nonetheless, because Michelangelo's workmen "are in some ways equal in importance to Michelangelo" himself. Greatness does not exist in isolation.

Just as the little guys are often overshadowed by the old masters they worked for, drawing often loses out to flashier forms like painting and sculpture. But the humble medium communicates the presence of hands, and minds, in motion.

Looking at these walls, you feel as if the artists who adorned them nearly five centuries ago—perhaps hiding from the summer heat, or taking a drink from the well—have just left. Even if this is not the work of Michelangelo, it evokes his world. ■



无法辨认的指纹

佛罗伦萨密室展出疑似米开朗基罗的画作

这位艺术大师的指纹遍布墙面——确定是他的吗？

置身于佛罗伦萨美第奇礼拜堂的墓室区，周围都是米开朗基罗设计的墓室和优雅的雕塑，你很容易注意不到左边的出口。它通向一个衣帽间，那里有一扇活板门连着一段狭窄的台阶。沿台阶走到底是一个带拱顶的小房间。蛋壳白色的墙壁上画满了姿态优美的裸体人物、正在坠落的法厄同（太阳神的儿子）、一个若隐若现的马头和几条匀称的腿。博物馆称，这些略显模糊的炭笔素描草图都是这位伟大的文艺复兴艺术家本人绘制的。

这个密室长期关闭。管理礼拜堂的巴杰罗国家博物馆（National Museum of the Bargello）过去只允许学者和像英国国王查尔斯三世和莱昂纳多·迪卡普里奥这样的大人物进入。但去年11月起，这个地下室开始了为期四个月的试开放，每周允许100位提前预约过的游客前来参观；每四人一组，可以观看15分钟。门票销售一空。最近官员们将试开放延长到了7月。

走进这个除了脚踝高度的LED灯条之外什么都没有的房间，就像回到了过去——不仅回到了文艺复兴时期，还回到了早期的旅游年代，那时候景点没有说明牌，也没有喧嚣的人群。最重要的是，感觉就像进入了某人的大脑，或是打开了他们的笔记本，里面满是草稿和涂鸦，有些是精雕细琢的，有些只是随手画就。你可以想象那位年届中年的大师就在这里，可能正在夜间作画。墙壁有几处似乎被燃烧的火炬产生的黑烟熏黑了。在一幅肌肉发达的人物草图旁边甚至有个模糊的掌印。

传说1530年米开朗基罗在这间密室里躲了好几个月。在三年前的一场民众起义之后，有一派人支持成立共和国、反对意大利最富有的家族美第奇的统治，这位艺术家站在了他们一边，尽管是他设计了以美第奇家族命名的墓室。当美第奇家族卷土重来时，这位艺术家就陷入了险境。这个小房间里正好有一口水井，或许看起来是一处安全的藏身之地。

1975年，美第奇礼拜堂当时的负责人怀疑一个原本放煤的储藏室里可能藏着什么东西，于是发现了这些画作。保拉·达戈斯蒂诺（Paola D'agostino）在今年1月中旬刚卸任巴杰罗博物馆的馆长，她认为最终能让公众参观这个房间很重要。这里展现了“米开朗基罗人性化的一面”，她说，尽管她承认无法确认这些画就是这位大师的亲笔，也承认其中一些涂鸦可能并不是他画的。

其他专家则怀疑这并非米开朗基罗所绘。安东尼奥·福切利诺（Antonio Forcellino）为这位艺术家写过两本传记，他认为这些画作的质量都不够高，不足以确定出处。圣路易斯华盛顿大学（Washington University in St Louis）教授、米开朗基罗研究权威威廉·华莱士（William Wallace）认为，可能只有两幅画出自米开朗基罗之手——一幅头像画，类似米开朗基罗在罗马看过的著名的拉奥孔雕像，还有楼梯附近画满整面墙的一幅男性立像——而大多数画是由建造上方的礼拜堂的五十多名工人中的一些人创作的。

那么米开朗基罗是躲在这个房间里吗？“这纯粹是瞎说。”华莱士表示。这位当时已经功成名就的艺术家更有可能是被一位朋友保护起来了。不过华莱士认为，这个房间仍然是个里程碑式的发现，因为米开朗基罗的工人们“在某些方面与米开朗基罗本人同样重要”。伟大不是孤立存在的。

正如小人物常常被带领他们工作的年长大师的光芒遮蔽，素描也往往不及油画或雕塑等更加华丽的艺术形式耀眼。但这样朴素的艺术创作展现了手和意识的运动。

看着这些墙壁，你会觉得好像近五个世纪前在上面作画的艺术家们才刚走开——他们在这里也许是为了躲避酷暑，也许是为了从井里打口水喝。即使这不是米开朗基罗的作品，也让人们感受到他身处的世界。■



Good optics

Refik Anadol's use of AI has made him the artist of the moment

His work seems to be everywhere, blurring the boundaries between art and engineering

HE IS IN high demand. Last year Refik Anadol projected luminous images of coral on to a wall at the World Economic Forum in Davos and covered the exterior screen of the Sphere, a new concert venue in Las Vegas, with animated, tumbling blue blocks. In October the Museum of Modern Art (MoMA) in New York acquired “Unsupervised—Machine Hallucinations”, in which a machine-learning model generates artworks based on those in the museum’s collection. On February 16th “Echoes of the Earth”, his largest-ever show in Britain, opened at the Serpentine North Gallery in London.

Mr Anadol, a 38-year-old Turk who lives in Los Angeles, is riding widespread public interest in artificial intelligence to become the most visible digital artist of his generation. His work reflects the innovation and anxieties of the current moment. As Mr Anadol sees it, AI is a powerful creative tool. In a world where so much of life happens in a digital realm, he argues, data has become a new “pigment”.

He is steeped in both art and science, having completed several arts degrees and a residency at Google focused on machine learning. Mr Anadol trains AI models on massive troves of data, often publicly available, to create raucously colourful animations that he calls “dreams” or “hallucinations”. They swirl on superbright screens (creating what he calls “data paintings”) and wiggle on walls (“data sculptures”); sometimes the pieces illuminate buildings, as at the Sphere.

One artwork, displayed at the Venice Architecture Biennale, drew on 70 terabytes of brain scans, allowing AI models to imagine the organ’s development. Another piece used an archive of the Los Angeles

Philharmonic's performances to imitate dreaming. A third assimilated more than 138,000 images and pieces of metadata from MoMA's collection (pertaining to provenance, for instance), along with local weather and data about noise levels. The results are churning clouds and waves, as well as abstractions evocative of Mark Rothko, a celebrated painter.

To undertake work on this scale Mr Anadol employs around 30 people, including architects, designers and engineers, half of whom work in his studio in Los Angeles. Public projects with institutions and companies have boosted his profile, but some private collectors have bought pieces, too. Mr Anadol also mints non-fungible tokens, digital artefacts that sometimes come with physical works.

The animations have proved popular: around 2.4m people came to see an exhibition of his work at MoMA in 2022. Mr Anadol's style is accessible and often beguiling. Understanding how machine learning works may help you fathom the process behind the "data paintings", but it is not essential. (In some installations, a control panel pops up to explain the model, giving the illusion of glancing under the bonnet but mostly evoking the futurism of "The Matrix".) You can be swept along by the crashing tides of colour, or watch a rose turn into a lily, without wondering whether you are "getting it".

Naturally Mr Anadol has critics as well as admirers. Some compare his animations to glorified screensavers and lava lamps, more spectacular than substantial. (Some do look as if they belong at a hotel in Las Vegas or at Burning Man.) Like anything generated by an AI model, Mr Anadol's animations raise questions about originality and whether such creations simply recycle the work of others.

Some worry that he glorifies AI while ignoring its risks, by presenting a rosy (or deep purple or yellow) view of the tech's potential. Casey Reas, one of

Mr Anadol's former teachers, says many in the art world are prejudiced against digital art, as they once were against photography, but concedes that "sometimes Refik's work can appear to have a utopian view of technology". The artist has appeared twice at TED and is fluent in the breathless Silicon Valley idiolect of "breakthroughs" and "inclusivity".

Wearing all black, in an all-black room in his studio, illuminated only by high-definition screens, Mr Anadol acknowledges that AI is changing everything—and not always for the better. But he is indeed excited about what the technology can do in the right hands. "I don't see the problem there," he says. "I see possibilities."

His latest project, on display at the Serpentine, is "Living Archive: Large Nature Model", which trained AI models on photographs, sounds and other kinds of scientific information collected at 16 rainforests across the globe. In addition to his usual sponsors, Google and Nvidia, institutions including the Natural History Museum in London and the Smithsonian have also furnished Mr Anadol with images and data.

| *Photo synthesis*

Mr Anadol prompts the model to create his trademark abstractions, as well as hyper-realistic creatures. Eagles morph into owls, which turn into toucans; the overarching point is the connectivity of the natural world. AI offers "a new brush, a thinking brush", he says.

He hopes that his work will educate people and help them "discover new worlds". A viewer might prompt a model with the name of a plant, and it will generate a new one right in front of them. The artist's ultimate goal is a place called Dataland: a fully immersive experience, including sounds and smells.

Hans Ulrich Obrist, the Serpentine's artistic director, says that Mr Anadol

“makes the invisible visible”; he captures the power of technology as he turns AI from an abstraction in the cloud into art before the eyes. Whether that art looks like a dream or a beautiful banality is up to viewers. But like it or not, people will be seeing a lot more of Mr Anadol’s work. ■



观感甚佳

用AI创作艺术，雷菲克·安纳多尔爆火

他的作品似乎无处不在，模糊了艺术和工程之间的界限

他现在炙手可热。去年，雷菲克·安纳多尔（Refik Anadol）在达沃斯世界经济论坛上将五光十色的珊瑚影像投射到了一面墙上，还给拉斯维加斯新建成的音乐场馆Sphere的外部屏幕设计了满屏翻滚的蓝色色块。去年10月，纽约现代艺术博物馆（MoMA）收藏了他的作品《无人监督——机器幻觉》（Unsupervised—Machine Hallucinations），它用机器学习模型基于该馆的馆藏不断生成新的艺术作品。2月16日，他迄今最大规模的作品《地球的回声》（Echoes of the Earth）在英国伦敦的北蛇形画廊（Serpentine North Gallery）揭幕。

现年38岁的土耳其人安纳多尔住在洛杉矶，他利用公众对人工智能的广泛兴趣进行艺术创作，成为他这一代人中最引人注目的数字艺术家。他的作品反映了当下的锐意创新和焦虑不安。在安纳多尔看来，AI是一种强大的创作工具。他认为，如今人们生活的很大一部分都是在数字领域中开展，数据已经成为一种新的“颜料”。

他深谙艺术和科学，拥有多个艺术学位，并在谷歌做过机器学习项目驻场艺术家。安纳多尔用大量数据（通常是公开数据）来训练AI模型，创作出他称之为“梦境”或“幻觉”的炫彩动画。这些动画有的在超亮屏幕上旋转涌动（创造出他称之为“数据绘画”的作品），有的在墙上摇摆不定（“数据雕塑”），有的装点建筑物，比如显示在Sphere的外屏上。

他有一件在威尼斯建筑双年展上展出的作品，利用70TB的脑部扫描数据，让AI模型想象大脑发育的过程。另一件作品用洛杉矶爱乐乐团（Los Angeles Philharmonic）的演出资料模拟梦境。还有一件作品吸纳了MoMA的13.8多万张图像和元数据（例如作品来源信息等），并加入了本地天气和噪音水平数据，最终呈现翻腾的云海和波涛，以及会让人联想到著名画家马克·罗斯科（Mark Rothko）的抽象作品。

为完成这么大规模的作品，安纳多尔雇用了包括建筑师、设计师和工程师在内的约30名员工，其中一半在他位于洛杉矶的工作室里耕耘。与机构和公司合作的公共项目提升了他的知名度，但也有一些私人收藏家购买了他的作品。安纳多尔还发行非同质化代币，有时这种数字作品也附带实体作品。

他的动画作品非常受欢迎，约有240万人参观了他2022年在MoMA的作品展。安纳多尔的风格易于理解，经常令人着迷。了解机器学习的工作原理可能有助于理解“数据绘画”背后的过程，但这并非必不可少。（一些作品展示会弹出控制面板来解释模型的创作过程，给人以参透其中玄机的幻觉，但更多是唤起了《黑客帝国》的那种未来主义感。）观众可以沉浸在翻涌不休的五彩斑斓的浪潮之中，或看着一朵玫瑰变成百合，而无需纠结于自己是否“看懂了”。

自然，对于安纳多尔既有人赞赏，也有人批评。一些人将他的动画作品比作华而不实的屏保和熔岩灯，徒有其表。（有些的确看起来像是拉斯维加斯某个酒店的展示，或者是火人节上的作品。）和任何由AI模型生成的东西一样，安纳多尔的动画引发了关于这些作品的原创性以及它们是否只是在重复他人作品的问题。

一些人担心，他以玫红（也可能是深紫或明黄）这样的明媚色调呈现AI技术的潜力，美化了AI，却忽略了风险。安纳多尔曾经的一位老师凯西·列阿斯（Casey Reas）表示，许多艺术界人士对数字艺术有偏见，就像过去的艺术界人士对摄影技术的态度，但他承认“有时雷菲克的作品会呈现出一种乌托邦式的科技观”。安纳多尔曾两次在TED上亮相，熟稔硅谷那一套自我沉浸的话术，大谈特谈“突破”和“包容”。

一身黑衣的安纳多尔身处工作室的一个漆黑房间里，仅有的光亮来自几个高清屏幕。他承认AI正在改变一切，而结果并不总是更好。但他确实因AI技术在得到善用时可能发挥的作用感到振奋。“在这方面我不觉得有问题，”他说，“我看到的是可能性。”

他在蛇形画廊展出的最新作品是《鲜活档案：大型自然模型》（Living Archive: Large Nature Model），该作品用收集自全球16处雨林的图片、声音和其他类型的科学信息训练AI模型。除了平时赞助他的谷歌和英伟达外，伦敦自然历史博物馆和史密森尼博物馆（Smithsonian）等机构也向安纳多尔提供了图像和数据。

| 图像合成

安纳多尔通过给AI模型提示词来创作他标志性的抽象图案及超现实主义生物。雄鹰变成猫头鹰，又变成巨嘴鸟。总的主题就是大自然万物相通。AI提供了“一支新画笔，一支会思考的画笔”，他说。

他希望自己的作品能够教育并帮助人们“发现新世界”。观众可以用植物的名字做提示词输入给模型，模型当场就能生成一种新植物。安纳多尔最终的目标是创建一个数据之地（Dataland），这是一种完全沉浸式的体验，用户能听到声音，闻到气味。

蛇形画廊的艺术总监汉斯·乌尔里奇·奥布里斯特（Hans Ulrich Obrist）表示，安纳多尔“将无形变为有形”，在将云端中抽象的AI变成眼前的艺术，体现了科技的力量。至于这种艺术形式看起来像是梦幻之境还是只是虚浮俗物，则取决于观众。但不论喜欢与否，人们看到安纳多尔的作品频率都会大增。■



Silicon Valley sobriety

The age of the unicorn is over

Don't expect AI to bring it back

BUSINESS HAS never been better for America's tech giants. After slumping in 2022, the combined market value of Alphabet, Amazon, Apple, Meta and Microsoft has surged by 70%, to over \$10trn, since the start of 2023 amid the hype over artificial intelligence (AI). The technology has also propelled others into the industry's upper echelons. On February 21st Nvidia, an AI-chip champion, reported that its sales rocketed by 265%, year on year, in the quarter to January. Its market value has risen from around \$500bn a year ago to \$1.7trn, making it America's fifth-most-valuable firm. OpenAI, the maker of ChatGPT, and other AI builders such as Anthropic have shot to fame, scooping up billions of dollars in funding.

Thousands of smaller AI firms have popped up, too. Enough that a small Caribbean island called Anguilla, whose internet domain suffix is ".ai", now generates around a third of its government's budget from licensing it out, according to Rest of World, an online publication. In the latest sign of madness returning to Silicon Valley Adam Neumann, the ousted founder of WeWork, an office-sharing firm that declared bankruptcy in November after years of losses, made a bid on February 5th to retake the reins of the firm.

Yet it would be a mistake to think America's startup scene is returning to its former exuberance. For one, the reception from WeWork's management and creditors to Mr Neumann's gambit has been lukewarm. Venture-capital (VC) firms invested only \$170bn in the country last year, down by half from 2021, according to PitchBook, a research firm. Bar a few high-profile exceptions, such as OpenAI, investors have been especially wary of signing cheques at lofty valuations. Throughout the 2010s the number of

unicorns—private companies with valuations above \$1bn—soared in America. Fully 344 of them were minted in 2021. Last year's figure was 45.

The end of the era of cheap money is largely to blame. In the go-go years, as investors raced to get a piece of the buzziest startups, tech firms had little need to tap public markets for capital. Crossover investors such as Tiger Global and Coatue, which operate in both public and private markets, flooded into Silicon Valley. Dharmesh Thakker of Battery Ventures, a VC firm, recalls that founders could “raise money on a Zoom call”. In 2021 crossover investors accounted for over half of startup funding. They have since retreated, last year contributing less than a third.

Now investors are mulling how to sell their stakes in the unicorns of yesteryear. Most VC funds operate on a ten-year clock, backing startups in the first five and cashing out in the second. With over 700 unicorns, at a combined valuation of \$2.4trn, a sizeable amount of money is at stake.

The first way to exit is through an initial public offering (IPO). Yet the IPO market remains at a standstill, with 83 VC-backed listings in 2023, down from 309 in 2021. Many of those that listed last year, including Instacart, a grocery-delivery business, and Klaviyo, a software firm, are trading below their initial price. Arm, a chip designer whose share price has more than doubled since its listing in September, is a rare exception. Firms that are planning for an IPO this year are often doing so at a reduced valuation: Reddit, a meme-sharing website, plans to list at \$5bn, down from a private valuation of \$10bn in 2021.

The second path to an exit—a sale to a corporate buyer—is also partly blocked. Only 698 VC-backed firms were purchased by companies last year, according to PitchBook, down from 1,311 in 2021. Trustbusters have kept big tech, once a serial acquirer, on the sidelines. Last month Amazon abandoned its bid to acquire iRobot, a maker of robo-vacuum-cleaners,

following scrutiny from European regulators.

Selling to another private investor—the third option—is not too attractive, either. Private valuations in the so-called secondary market are below those at the latest fundraising round for more than four-fifths of unicorns, according to Caplight, a data provider. Discord, a chat service popular with gamers, was most recently valued in the secondary market at \$6bn, down from a nearly \$15bn valuation when it last raised funds in 2021.

Amid the drought, some unicorns have simply collapsed. Convoy, a logistics startup that last raised funds in 2022 at a nearly \$4bn valuation, shut down in October. Veev, a unicorn dedicated to disrupting home-building, closed its doors in November and is liquidating its assets. Samir Kaji of Allocate, a firm that connects investors with VC funds, believes that many unicorns will “quietly get acquired for parts”.

| *Back to the garage*

That is a sad fate for the founders, employees and investors of those once-promising firms. But others need not be overly worried. Tom Tunguz of Theory Ventures, another VC firm, reckons the drop in funding since 2021 is merely a return to a long-run trend that was thrown off course by the pandemic. And there is plenty to celebrate in the newfound sobriety of Silicon Valley.

It is rare these days to find a startup that espouses growth at all costs. Founders have rediscovered the concept of frugality. Many are being cautious with their hiring, a striking contrast to the race for talent during the pandemic. It helps that the industry's giants have flooded the market with thousands of techies following a bout of layoffs. Over the past two years Alphabet, Amazon, Meta and Microsoft have fired more than 75,000 staff between them.

What's more, AI is providing American startups not just with new business ideas, but also ways to do more with less. Startups burn roughly half their cash on selling their products and a third on engineering. AI assistants for salesmen and coders are raising productivity by more than a third at some startups, lowering the amount of capital they need to raise. In time, the age of the unicorn may not be missed. ■



硅谷清醒

独角兽时代落幕

别指望它借人工智能之势回归

美国科技巨头的生意从未如此红火。经过2022年的低迷后，在人工智能（AI）概念大热之下，自2023年初以来，Alphabet、亚马逊、苹果、Meta和微软的总市值飙升70%，突破十万亿美元。AI技术还助推其他公司跻身行业前列。2月21日，AI芯片领军者英伟达发布报告称，截至1月的财季其销售额同比飙升265%，市值从一年前的约5000亿美元上升至1.7万亿美元，成为美国市值第五高的公司。开发ChatGPT的OpenAI以及Anthropic等其他AI公司一举成名，收获数十亿美元融资。

同时另有成千上万小型AI公司涌现。据线上科技媒体Rest of World报道，加勒比海上一个名叫安圭拉（Anguilla）的小岛因为互联网域名后缀是“.ai”而受益，现在其政府预算约有三分之一来源于授权该域名的收入。办公室共享公司WeWork在多年亏损后于去年11月宣布破产，被赶下台的创始人亚当·诺伊曼（Adam Neumann）于2月5日出价回购该公司——这是硅谷再陷疯狂的最新例证。

但要是就此以为美国创业圈正恢复昔日盛景，那就错了。首先，WeWork管理层和债权人对诺伊曼的出价反应冷淡。研究公司PitchBook的数据显示，风投公司去年在美国仅投资1700亿美元，比2021年减少一半。除了OpenAI等少数几个吸人眼球的例外，投资者对于按高昂估值开出支票尤其谨慎。整个2010年代，美国独角兽公司（估值超过十亿美元的私人公司）的数量激增。在2021年的一年里就新增了足足344家，而去年的数字是45家。

廉价资金时代的终结是主要原因。在那些兴旺的年头，投资者争抢最热门创业公司的股份，科技公司几乎不需要从公开市场融资。老虎环球（Tiger Global）和Coatue等同时在公开市场和私人市场运作的交叉投资者涌入硅谷。巴特利风投公司（Battery Ventures）的达梅什·塔克尔（Dharmesh

Thakker) 回忆说, 创始人当时“通过Zoom电话就能融到资金”。2021年, 交叉投资者为创业公司贡献了超过一半的融资。自那以后, 他们的投资减少了, 去年只占不到三分之一。

如今投资者正琢磨如何卖掉昔日买进的独角兽股份。大多数风投基金以十年为一个周期, 前五年支持创业公司, 后五年变现退出。目前美国有700多家独角兽, 总估值达2.4万亿美元, 其中牵扯的资金规模很可观。

第一条退出途径是IPO。但IPO市场仍处于停滞状态, 2023年只有83家由风投支持的公司上市, 相比之下2021年有309家。食品杂货配送公司Instacart和软件公司Klaviyo等许多去年上市的公司目前股价都低于发行价。芯片设计公司安谋(Arm)是个少有的例外, 自去年9月上市以来, 它的股价翻了一倍多。计划今年IPO的公司大多准备降低估值上市: 网络梗分享网站Reddit计划以50亿美元的价格上市, 低于2021年私人融资时100亿美元的估值。

第二条退出途径是出售给其他公司, 这条路同样不好走了。PitchBook的数据显示, 去年只有698家由风投支持的公司被其他公司收购, 低于2021年的1311家。大型科技公司以往收购成风, 如今受制于反垄断机构, 改持观望态度。今年1月, 在受到欧洲监管机构的审查后, 亚马逊放弃收购扫地机器人制造商iRobot。

第三条途径是出售给其他私人投资者, 但也不太有吸引力。数据供应商Caplight指出, 在所谓的二级市场上, 超过五分之四的独角兽的私人收购估值低于上一轮融资时的估值。游戏玩家热捧的聊天软件Discord最近在二级市场上的估值为60亿美元, 低于2021年最近一次融资时近150亿美元的估值。

资金枯竭之下, 一些独角兽直接就崩塌了。物流创业公司Convoy在去年10月倒闭, 它上一次融资是在2022年, 当时估值近40亿美元。致力颠覆住宅建设的独角兽Veev于11月关门, 正在清算资产。为投资者和风投基金牵线搭桥的公司Allocate的萨米尔·卡吉(Samir Kaji)认为, 许多独角兽公司

将“悄无声息地被分割收购”。

| 回到车库

对于这些曾经前景大好的公司的创始人、员工和投资者来说，这是悲惨收场。但其他人不必过于担心。另一家风投公司Theory Ventures的汤姆·通古斯（Tom Tunguz）认为，2021年以来的融资减退只不过是回归了因新冠疫情而偏离轨道的长期趋势。硅谷回归清醒很值得庆贺。

如今，不惜代价追求增长的创业公司已不多见。创始人重拾节俭理念。许多人在招聘时很是谨慎，这与疫情期间的人才争夺战形成鲜明对比。行业巨头的一轮裁员潮过后，成千上万技术人员涌入市场，这也让创业公司得以捡漏。在过去两年内，Alphabet、亚马逊、Meta和微软共解雇了超过75,000名员工。

此外，AI不仅为美国创业公司提供了新的商业理念，还让它们能少花钱多办事。创业公司有一半的资金会烧在产品销售上，三分之一在工程设计上。目前，在一些创业公司里，销售和编程人员能凭借AI助手把生产率提高三分之一以上，这降低了所需的融资额。假以时日，人们或许会不再怀念独角兽时代。■



Aural history

Podcasts got their name 20 years ago this month

A once-fringe medium has gone mainstream, but many wonder what the future holds

IN FEBRUARY 2004 Ben Hammersley, a British journalist, noticed a new kind of digital media: “downloadable radio” produced by geeky amateurs or public broadcasters. Mr Hammersley suggested a few names for it. Fortunately “audioblogging” and “GuerillaMedia” did not catch on. But a third did: podcasting, a portmanteau of “iPod” and “broadcasting”.

At a tech conference the next year, Steve Jobs, Apple’s boss, asked how many people in the audience had heard of podcasting, which he described as “‘Wayne’s World’ for radio”. Not a single person raised their hand. But today, 20 years after they got their name, podcasts are mainstream. Around 43% of American internet users and 30% of Britons listen to at least one a month (including some from The Economist). You can open Apple, Spotify and, increasingly, YouTube to find some 4m shows by everyone from ex-convicts to duchesses. Podcasting has turned anyone with a microphone into a talking head who can talk straight into people’s earbuds.

Like butt-dialling and doom-scrolling, the podcasting boom was facilitated by the smartphone. The iPhone was released in 2007; the first standalone podcast apps came out between 2010 and 2012. In 2014 “Serial”, a binge-worthy investigation into the botched murder trial of a teenager in Maryland, arrived; it was the first show to gain attention in mainstream American culture. (The first season has been downloaded more than 300m times.) The number of monthly podcast listeners in America doubled in the five years after 2014.

Since “Serial”, podcasting has been in a “hyper-charged, hyper-capitalist, hyper-speculative period”, says Nicholas Quah, a podcast critic. Tech

companies swooped in, such as Spotify, which was hungry for new customers and audio content it could have royalty-free. (Spotify has invested \$1bn in podcasting since 2018.)

But last year—cue the “record scratch” sound effect—the music stopped. Economic headwinds were one factor. But advertisers and investors also realised that they know relatively little about how many people actually listen to podcasts: “downloads”, the standard industry metric for engagement, do not always equate to listening. Spotify eliminated 200 podcasting jobs; others followed suit. Many well-regarded shows were cancelled, including “Stolen”, which won the Pulitzer prize for audio journalism in 2023. Podcasters learned that “excellence won’t save you”, says Neil Verma, a media-studies scholar.

Although producers are struggling to turn a profit, podcasting is in the ascendant. Advertising spending and listening time are predicted to rise. But the medium “is growing in a direction that is hard for a lot of the podcast originalists to accept”, says Ariel Shapiro, who writes Hot Pod, an industry newsletter.

What does the future hold? Some think the future of podcasting lies in video. Many podcasts, including a talk show hosted by Joe Rogan (pictured), the most popular podcast globally, now employ video to increase their audience. “We will be watching podcasts, we will be skipping chapters of podcasts, and you will be commenting on them at the end,” predicts Will Page, Spotify’s former chief economist.

Betting that some people want visual entertainment while listening, Spotify started supporting video podcasts in 2020. Already the black microphone icon has become omnipresent on YouTube and TikTok, functioning as a visual signal for podcast-like straight talk. Video will help podcasts’ further spread; rarely does audio alone go viral.

Whatever happens to the industry, its impact on audio storytelling will endure. Podcasts offer something traditional formats often work hard to leave out, but which, as trust in media has fallen, audiences crave: transparency. Instead of formal interviews, listeners get sprawling conversations. Instead of radio features measured in minutes, listeners are led through whole investigations over hours. “In exposing the vulnerabilities of the reportorial process, the news product becomes more credible and trustworthy,” writes David Dowling in “Podcast Journalism”, a forthcoming book. Podcasts often set out to explain things—and they always end up explaining themselves. ■



耳中历史

二十年前的二月，播客这个名字诞生了

曾经的边缘媒体已经成为主流，但许多人想知道未来会怎样

二〇〇四年二月，英国记者本·哈默斯利（Ben Hammersley）注意到了一种新型数字媒体：由极客业余爱好者或公共广播公司制作的“可下载的广播”。哈默斯利给它起了几个名字。还好，“音频博客”（audioblogging）和“游击媒体”（GuerillaMedia）这两个名字没有流行起来。但是第三个名字火了：播客（podcasting），一个由“iPod”和“broadcasting”组成的合成词。

在次年的一次科技会议上，苹果公司的老板乔布斯问听众中有多少人听说过播客，他形容播客是“广播界的‘反斗智多星’（Wayne's World）”。没有一个人举手。但是今天，播客这个名字诞生20年之后，它成了主流。大约43%的美国网民和30%的英国网民每月至少听一次播客（包括本刊发布的一些）。你可以打开Apple播客、Spotify，以及越来越多地从YouTube上找到大约400万个节目，主播形形色色，从公爵夫人到前科犯什么人都有了。有了播客，只要有麦克风，你就能成为在别人耳机里说话的主播。

和不小心自动拨出了电话以及狂刷坏消息停不下来这些现象一样，这股播客繁荣也是智能手机带来的结果。iPhone于2007年发布，而第一个独立的播客应用出现在2010年到2012年之间。2014年一档名为《连载》（Serial）的播客节目横空出世，讲述了对马里兰州一桩青少年谋杀案拙劣审判的调查，很值得一口气听完，是第一个获得美国主流文化关注的播客。（第一季的下载量已超过3亿次。）2014年之后的五年间，美国每月收听播客的人数翻了一番。

播客评论家尼古拉斯·奎阿（Nicholas Quah）说，自《连载》出现以来，播客便处于“高度兴奋、高度资本主义、高度投机的时期”。科技公司蜂拥而至，比如渴望获得新客户和免版税音频内容的Spotify。（自2018年以来，Spotify已在播客上投资了10亿美元。）

但到了去年——此处应配上“刮擦唱片”音效——音乐戛然而止。经济逆风是一个原因。但广告主和投资者也意识到，他们对实际上有多少人收听播客知之甚少：“下载量”——衡量用户参与度的标准行业指标——并不总是等同于收听量。Spotify在播客部门裁员200人，其他公司也纷纷效仿。许多备受好评的节目被取消，包括2023年获得普利策音频报道奖的《被盗》（Stolen）。媒体研究学者尼尔·维尔马（Neil Verma）表示，播客主播们认识到“节目做得棒并不能拯救你”。

虽然制作人还在为实现盈利犯愁，但播客仍处在上行通道中。预计广告支出和收听时长还会增加。但行业简报《热门播客》（Hot Pod）的作者艾莉儿·夏皮罗（Ariel Shapiro）表示，这一媒介“正朝着许多播客原教旨主义者难以接受的方向发展”。

未来会怎样？有些人认为播客的未来要走向视频。包括乔·罗根（Joe Rogan，见图）主持的全球最受欢迎的脱口秀在内的许多播客现在都用视频来吸引更多听众。“我们会观看播客，跳过它们的一些章节，在结尾处发表评论。”Spotify的前首席经济学家威尔·佩奇（Will Page）预测说。

Spotify认定有些人想在听节目的同时享受视觉娱乐，从2020年开始支持视频播客。在YouTube和TikTok上，黑色麦克风图标已经无处不在，表明这是像播客那样畅所欲言的视频节目。视频将有助于播客的进一步传播；单靠音频很少能大火。

无论这个行业发生了什么，它对音频叙事的影响会持续下去。播客提供了一种传统媒体形式常常极力排除的东西：透明度。而随着人们对媒体信任度的下降，他们越发渴望透明度。听众听到的不是正式的采访，而是信马由缰的漫谈。不同于以分钟为单位的广播节目，听众花几小时在主播的带领下了解整个调查过程。戴维·道林（David Dowling）在即将出版的《播客新闻学》（Podcast Journalism）一书中写道：“通过把报道过程中的漏洞弱点直接呈现给受众，新闻产品变得更加可信可靠。”播客常常是为了解释什么事情而制作的——但最终却总是在解释自己。■



Sweet science

New treatments are emerging for type-1 diabetes

The trick is to outsmart the immune system

“WHERE ARE the islets of Langerhans?” is a trick question that pops up from time to time in quizzes. The answer is to be found not in atlases of geography, but rather in those of anatomy, for the so-called islets are in fact clusters of cells scattered through the pancreas. There they synthesise and release a range of hormones, including insulin, which regulates glucose levels and thus metabolism.

The islets’ insulin producers are called beta cells. (Cell types alpha, gamma, delta and epsilon perform other tasks.) They are the only bodily sources of that hormone. So, if their number declines, trouble looms. And decline it does, in the condition known as type-1 diabetes. This happens when, in a phenomenon called autoimmunity, the body’s own immune system attacks its complement of beta cells, wiping out as many as 80%.

Without an alternative supply of insulin, someone with type-1 diabetes will die. (In type-2 diabetes, insulin continues to be produced but the body’s cells acquire resistance.) Supplementary insulin can be administered by injection or via a device called an insulin pump. But a better way might be to replace the missing beta cells and somehow protect them from immune attack.

A few lucky patients do indeed have their beta cells replaced—by transplantation from human donors. And Vertex Pharmaceuticals, a firm in Boston, is testing beta cells grown from stem cells for the same purpose. But neither approach includes immune protection. This means both require the administration of immunosuppressive drugs to prevent the rejection that follows any transplant, let alone one where autoimmunity is

at play. One of the sessions at this year's meeting of the American Association for the Advancement of Science in Denver therefore looked into how transplanted beta cells might be made hypoimmunogenic—in other words, invisible to a patient's immune system.

Sonja Schrepfer, who works for the University of California, San Francisco (UCSF), and also at Seattle-based Sana Biotechnology, proposes a twofold approach, to deal with the fact that the immune system has two arms. One, the adaptive arm, is the basis of tissue rejection. This adaptive arm can recognise the signature of “selfness” provided by an individual's HLA proteins. These molecules contain so-called hypervariable regions, which differ from individual to individual. If the immune system comes across non-self HLA proteins, it recognises the cells displaying them as interlopers and attacks, using shock troops called killer T-cells and antibodies.

The first part of Dr Schrepfer's approach is therefore to prevent the production of HLA proteins in laboratory-grown beta cells destined for transplant. This can be done by editing two genes involved in their production, theoretically rendering the cells in question invisible to the adaptive arm.

Lack of HLA proteins does, however, bring a cell to the attention of immunity's other arm, the innate system. Its troops are called NK (natural killer) cells and macrophages, and one of the red flags it reacts to is an absence of any sort of HLA. It can, however, be warded off by over-expression of a protein called CD47, something that Dr Schrepfer's team also achieved by genetic manipulation of their beta cells.

It seems to work. In an experiment whose results were announced just before the meeting, the team first induced diabetes in a laboratory monkey and then injected their modified beta cells into one of its muscles. The

diabetes went away, and stayed away for more than six months. Now they have moved on to people. In a trial about to start at Uppsala University Hospital, in Sweden, human versions of the modified cells will be transplanted into the forearm of a single patient.

Disrupting beta cells' HLA expression is not the only possible approach to deflecting the adaptive immune system. Hasna Maachi of the Technical University of Munich, in Germany, described to the meeting how she and her mentor, Matthias Hebrok, are trying to develop an alternative. This introduces a third party, called a suppressor cell, to do the warding off.

Suppressor cells “talk” to the killer T-cells, and calm them down. Dr Hebrok's group is therefore working with Wendell Lim of UCSF, who is engineering suppressor cells specifically activated by a protein on the surface of beta cells, to devise some beta-cell protection. In this case there is no need to build in a level of protection against the innate immune system, since it will not notice anything wrong.

The Hebrok group is also, Dr Maachi explained, working on a way to turbocharge beta cells. This involves a protein called MAFA, which regulates expression of the insulin gene. Suppressed levels of MAFA are a symptom of type-1 diabetes, so boosting its presence seemed a promising approach. So far, the researchers have shown that raising MAFA levels in beta cells derived from stem cells seems to increase the amount of insulin produced.

Figures presented by Lori Sussel of the University of Colorado suggest that type-1 diabetes affects one in 500 Americans. The global average may be closer to one in 1,000. That is both a great deal of human suffering and a tempting market for whoever can come up with something which resembles a cure, rather than a treatment. Though there is a way to go yet, hypoinmunogenic beta cells might, just possibly, bring that closer. ■



甜蜜的科学

一型糖尿病的新疗法

诀窍在于骗过免疫系统

知识问答中时不时会出现这么一个刁钻问题：“朗格汉斯岛在哪里？”答案不在地理地图中，而在解剖学图册里，因为朗格汉斯岛实际上是散布在胰腺中的细胞团，也称为胰岛。胰岛合成并释放多种激素，包括调节葡萄糖水平继而调节新陈代谢的胰岛素。

胰岛中分泌胰岛素的细胞称为beta细胞。（alpha、gamma、delta和epsilon等其他类型的细胞执行别的任务。）它们是人体内唯一的胰岛素来源。因此，如果beta细胞数量减少，就会出现麻烦。而它确实会减少，患上型糖尿病就会这样。这种糖尿病的发病是因为自身免疫现象：人体的免疫系统攻击自己的beta细胞，会消灭多达80%的这种细胞。

如果没有另外的胰岛素补充，一型糖尿病患者将会死亡。（在二型糖尿病中，胰岛素仍然继续分泌，但人体细胞对其产生了抵抗。）可以通过注射或使用名为胰岛素泵的装置来补充胰岛素。但更好的方法也许是替换缺失的beta细胞，并想办法保护它们免受免疫攻击。

少数幸运的患者确实通过人类供体移植替换了beta细胞。而波士顿的Vertex Pharmaceuticals正在测试由干细胞培养的beta细胞，以便能做同样的替换。但这两种方法都没有免疫保护。这意味着它们都需要使用免疫抑制药物以避免排斥反应——任何移植手术后都会有排斥反应，更不用说还有自身免疫性现象在发挥影响了。因此，今年在丹佛举行的美国科学促进会（American Association for the Advancement of Science）年会上，就有一个专场讨论如何让移植的beta细胞具有低免疫原性——换句话说，就是在患者的免疫系统面前隐形。

在加州大学旧金山分校和西雅图的Sana Biotechnology工作的索尼娅·施雷普弗（Sonja Schrepfer）提出了一种双管齐下的方法，以应对免疫系统

的两个部分。其一是后天性免疫系统，也是组织排斥的基础。这种适应性免疫系统能够识别人体HLA蛋白表达的“自我”特征。这些分子包含有名为超变异区的部分，每个人都各不相同。如果免疫系统遇到非自身的HLA蛋白，它就会将表达这种蛋白的细胞视为入侵者，并使用杀伤性T细胞和抗体作为突击部队对其发起攻击。

因此，施雷普弗的方法首先要避免在实验室培养的用于移植的beta细胞产生HLA蛋白。要做到这一点，可以对参与产生HLA蛋白的两个基因进行编辑，理论上可以让这些细胞在后天性免疫系统面前隐身。

然而，缺少HLA蛋白又会引起免疫系统的另一个部分——先天性免疫系统的注意。它的部队称为自然杀伤细胞和巨噬细胞，缺乏任何一类HLA都是一个危险信号，会让它做出反应。然而，通过一种CD47蛋白的过度表达可以避免它的攻击，施雷普弗的团队也通过对其beta细胞进行基因编辑实现了这一点。

这种方法似乎奏效了。在会议前不久公布的实验结果显示，研究团队首先在一只实验猴身上诱发了糖尿病，然后在它的一块肌肉中注入改造过的beta细胞。糖尿病消失了，而且持续了六个多月。现在，他们已经进入人体试验阶段。在瑞典乌普萨拉大学医院（Uppsala University Hospital）即将开展的一项试验中，经过改造的人体beta细胞将被移植到一名患者的前臂。

干扰beta细胞的HLA表达并不是唯一一种可能躲避后天性免疫系统的方法。德国慕尼黑工业大学的哈斯娜·马奇（Hasna Maachi）在会上介绍了她和导师马蒂亚斯·赫布鲁克（Matthias Hebrock）正在研究的另一种方法。该方法引入了一个称为抑制细胞的第三方来规避免疫系统。

抑制细胞能与杀伤性T细胞“对话”，使它们平静下来。因此，赫布鲁克的团队正与加州大学旧金山分校的林行健合作，他正在改造抑制细胞，让它们可以专门被beta细胞表面的一种蛋白激活，从而设计出一些保护beta细胞的方法。在这种情况下，无需针对先天性免疫系统建立起某种保护，因

为它不会察觉到任何问题。

马奇解释说，赫布鲁克团队还在研究一种增强beta细胞的方法。这涉及到一种称为MAFA的蛋白，它能调节胰岛素基因的表达。MAFA水平受到抑制是一型糖尿病的症状之一，因此提高其表达似乎是一种很有前景的方法。到目前为止，研究人员已经证明，在干细胞培育的beta细胞中提高MAFA水平，似乎可以增加胰岛素的分泌。

科罗拉多大学的洛里·苏塞尔（Lori Sussel）展示的数据显示，每500个美国人中就有一人患有一型糖尿病。全球平均患病率可能接近千分之一。这既是人类的巨大苦难，也是诱人的市场——只要有人能开发出更能治本而非仅仅治标的方法。虽然还有一段路要走，但低免疫原性的beta细胞或许为接近这个目标提供了一些可能。 ■



Through the looking glass

The world's biggest maker of spectacles wants to be a tech firm

It is experimenting with smart glasses and built-in hearing aids

OVER THE course of six decades Leonardo Del Vecchio, an Italian entrepreneur who died in 2022, built Luxottica into the world's largest maker of spectacles and sunglasses. In 2018 he merged his firm with Essilor, a French lens manufacturer, to create a Franco-Italian corporate giant that is today worth over \$90bn and employs some 200,000 people. The group owns iconic eyewear labels from Ray-Ban to Oliver Peoples, and also produces glasses using the brands of European luxury houses such as Armani and Chanel. On February 14th the company reported that its sales grew by 7% last year, at constant exchange rates, faster than at many other luxury businesses.

Now Francesco Milleri, a close confidant of Mr Del Vecchio's who was appointed chief executive of Luxottica in 2017 and of the combined group in 2020, is looking for the company's next act. He has two big ideas. First is to become the leader in smart glasses. The ambition is not entirely new: the firm partnered with Google, a tech giant, on its Google Glass—an unambiguous flop thanks in part to its clunky interface and dorky design. The device was discontinued in 2015.

Yet there are plenty of optimists who think that the moment for smart spectacles has arrived. In 2021 Snap, a social-media firm, unveiled a pair of augmented-reality glasses. Last month Apple began shipping its augmented-reality headset, the Vision Pro. It reportedly pre-sold 200,000 of the devices, despite a price tag of \$3,499. "I believe we are at the dawn of a product revolution akin to what we saw in the early 90s in mobile phones," reckons Luca Solca of Bernstein, a broker.

EssilorLuxottica's latest foray into the technology is through Ray-Ban, in partnership with Meta, another tech giant. The first iteration it launched in 2021 had only modest success. It will be hoping that a new version it began selling in September, which has a snazzier camera and a longer battery life, will do better. The device can capture and relay what the wearer sees. An in-built virtual assistant can hear and respond to their requests.

Mr Milleri's second idea is the one in which he seems most invested. "We have become a med-tech company," he declares. Central to that is the development of spectacles with built-in hearing aids, which it will launch in August, for those with moderate hearing loss. It reckons that is a group of some 1.25bn people worldwide, many of whom shy away from the perceived stigma and high price of conventional hearing aids. The company's spectacles, by contrast, will look no different from other glasses and cost well below the price of a regular hearing aid, says Mr Milleri.

The thinking is not altogether original. The Beltone "Hear-N-See", one such device, first debuted in 1956. Audiofon, a German maker of hearing aids, has also dabbled in spectacles. But the idea has never been adopted en masse. Mr Milleri is counting on better technology to change that. Last year his company acquired Nuance Hearing, an Israeli startup founded in 2015. Its technology uses an algorithm to detect where sounds are coming from based on when they are picked up by different microphones. The spectacles it has developed then isolate and process the voice of the person the wearer is speaking to and transmit it through tiny built-in speakers. Nothing goes in the wearer's ear.

Mr Milleri's enthusiasm for the concept is personal. Mr Del Vecchio, who was 87 when he died, suffered from hearing loss and often lamented that hearing aids were uncomfortable to wear with glasses. That inspired Mr Milleri and Leonardo Maria Del Vecchio, one of the founder's sons, who is also the firm's strategy chief, to pursue the idea. Whether they prove to be

farsighted remains to be seen. ■



镜中缘

全球最大的眼镜制造商希望变身科技公司

它正在研发智能眼镜和内置助听器眼镜

二〇二二年去世的意大利企业家莱昂纳多·德尔维奇奥（Leonardo Del Vecchio）生前用了60年的时间将陆逊梯卡（Luxottica）打造成全球最大的光学眼镜和太阳镜制造商。2018年，他将陆逊梯卡与法国镜片制造商依视路合并，创建了一家法意企业巨头。如今，这家名为依视路陆逊梯卡（EssilorLuxottica）的集团价值超过900亿美元，雇有约20万名员工。它旗下不仅拥有雷朋、奥利弗-皮珀兹（Oliver Peoples）等一些知名眼镜品牌，还为阿玛尼、香奈儿等欧洲奢侈品牌生产眼镜。2月14日，依视路陆逊梯卡公告称，其去年的销售额（按固定汇率计算）增长了7%，增速超过许多其他奢侈品企业。

如今，德尔维奇奥生前的密友弗朗西斯科·米勒里（Francesco Milleri）正在谋划这家公司的下一步发展。米勒里于2017年和2020年先后被任命为陆逊梯卡和依视路陆逊梯卡的CEO。他有两个宏伟计划。第一是成为智能眼镜的领军者。这一雄心并非前所未有的：陆逊梯卡曾与科技巨头谷歌合作开发谷歌眼镜，但以完全失败告终——难用的操作界面和不够自然的外观设计都是其槽点。谷歌眼镜于2015年停售。

不过也有很多乐观主义者认为，智能眼镜的时代已经到来。2021年，社交媒体公司Snap推出了一款增强现实AR眼镜。今年2月，苹果开始发售其AR头显Vision Pro。尽管售价高达3499美元，但据说已经预售了20万台。“我相信，我们正处于一场产品革命的开端，就像我们在上世纪90年代初经历的手机革命一样。”经纪公司盛博（Bernstein）的卢卡·索尔卡（Luca Solca）认为。

依视路陆逊梯卡最近一次进军该技术领域是它的雷朋与另一家科技巨头Meta合作推出智能眼镜。它在2021年推出的第一代产品只算小有成就。它希望今年9月开售的新一代产品会有更出色的表现。新一代产品配备了更

时尚的摄像头，电池续航也更久。它可以拍摄并转播佩戴者看到的東西。内置的虚拟助手可以听到佩戴者的指令并做出回应。

米勒里投入最多的似乎是他的第二个计划。他宣称：“我们已经成为一家医疗科技公司。”研发内置助听器眼镜是该计划的核心。这种专为轻中度听力损失人群所打造的眼镜将于今年8月推出。依视路陆逊梯卡估计，全球约有12.5亿人患有轻中度听力损失，其中许多人因为自感羞耻以及价格昂贵而对传统助听器望而却步。米勒里表示，相比之下，依视路陆逊梯卡的助听眼镜看起来就像一副寻常眼镜，价格又远低于普通助听器。

这种思路并非全新。贝尔通（Beltone）1956年问世的“Hear-N-See”就是一款这样的设备。德国助听器制造商Audiofon也曾涉足眼镜领域。只不过这一想法从未被广泛采用过。米勒里希望更先进的技术能够改变这种状况。去年，依视路陆逊梯卡收购了2015年成立的以色列创业公司Nuance Hearing。Nuance Hearing的技术使用的算法可以根据不同麦克风的拾声时间来识别声音的来源。依视路陆逊梯卡研发的眼镜又会把佩戴者交谈对象的声音分离出来处理，并通过内置的微型扬声器传输给佩戴者。佩戴者的耳朵里不用放置任何东西。

米勒里热衷于这一想法是出于个人原因。德尔维奇奥87岁去世前听力已经很不好，他经常抱怨同时戴着助听器和眼镜很不舒服。这激发了米勒里和莱昂纳多·玛丽亚·德尔维奇奥（Leonardo Maria Del Vecchio，创始人的一个儿子、同时也是公司的战略主管）去探索内置助听器眼镜的想法。他们此举是否有远见尚待分晓。■



In the nick of time

As the Nikkei 225 hits record highs, Japan's young start investing

Will more now favour domestic stocks?

SAITO MARI, a 28-year-old nurse, was frustrated. Her pay, at just ¥160,000 (\$1,100) a month, was meagre; after bills, rent, shopping and a few holidays, she had little left over. So in 2020 she decided to buy some stocks. “I used to think it was too risky,” says Ms Saito, who learned about investing via books and YouTube. “But it was amazing to see my assets grow.”

Although Ms Saito's story would be unremarkable anywhere else, it is part of a sea change in Japan. According to surveys by the Investment Trusts Association, 23% of people in their twenties invested in mutual funds last year, up from 6% in 2016. So did 29% of people in their thirties, up from 10%—a bigger rise than in any other age group. Those with exposure to the Nikkei 225, which on February 22nd passed a record high set in 1989, are reaping the rewards.

Japan's officials, who want to boost economic growth, have long desired such a shift. The public's previous aversion to retail investing dates back to the early 1990s, when a stockmarket bubble burst. In the ensuing decades, with inflation minimal or non-existent, low-risk saving came to be seen as virtuous. Some 54% of Japanese household assets are in cash or deposits, against 31% in Britain and 13% in America.

Kishida Fumio, Japan's prime minister, outlined an “Asset Income Doubling Plan” in 2022. This aims to create a virtuous cycle: companies will grow by making use of funds from retail investors; individuals will enjoy the benefits of their growth. As part of the initiative, in January the government improved the terms of its NISA programme, modelled on Britain's ISA, which exempts retail investors from capital-gains taxes. The same month

900,000 new NISA accounts were opened with the country's five biggest investment platforms.

Mr Kishida's push has been given extra oomph by economic developments. Under Japan's zero-interest-rate policy, hoarding cash in a bank brings almost no return. This has been true for a while, but inflation now stands at around 3%—a three-decade high—meaning the value of cash not put to work is being eroded. Young generations, who do not share the trauma of the burst bubble, are more inclined to act.

The number of students at ABCash, a financial school in Tokyo targeting millennials, has doubled since 2022, reaching 40,000. Shinjo Sayaka joined after seeing an influencer mention it on Instagram. "It's hard to talk about money with my family," she reports. One problem for Mr Kishida is that many youngsters favour international markets over domestic ones. For instance, Ms Saito's investments include Apple (an American tech giant), the S&P 500 (an index of big American firms) and BioNTech (a German vaccine-maker). Yet perhaps she and others will change their approach if the Nikkei continues to soar. ■



日经逆袭时刻

随着日经225指数创下历史新高，日本的年轻人开始投资股市

接下来会有更多日本人青睐本国股票吗？

28岁的护士斋藤真理有些沮丧。她工资很低，每个月只有16万日元（1100美元）；付完各种账单和房租，再买点东西、度几次假，便所剩无几了。所以在2020年，她决定买些股票。她通过书籍和YouTube学习如何投资。“我过去认为股票风险太高，”斋藤说，“但看到自己的资产在增长，真的很惊喜。”

尽管斋藤的情况在其他任何地方都可能稀松平常，但它却反映了日本的一些重大转变。日本投资信托协会（Investment Trusts Association）的调查显示，去年，20多岁人群有23%投资了共同基金，而2016年这一比例为6%；30多岁人群中做此投资的有29%，而2016年这一比例为10%——增幅高于其他年龄段。2月22日，日经225指数突破了1989年创下的历史最高点，那些持有该指数基金的投资者正在收获回报。

长期以来，希望促进经济增长的日本官员一直就想要这种转变。民众之前对散户投资的排斥可以追溯到上世纪90年代初股市泡沫破裂的时候。而在随后几十年，通胀率极低或者为零，低风险储蓄开始被视为正途。日本家庭大约54%的资产是现金或存款，而英国和美国这一比例分别为31%和13%。

日本首相岸田文雄在2022年提出了“资产收入倍增计划”。此举想要创造一种良性循环：企业可以利用散户投资者的资金获得发展，而个人则可以享受企业发展带来的好处。日本先前仿效英国的ISA（免除散户投资者的资本利得税）推出了小额投资非课税制度（NISA）。今年1月，作为资产收入倍增计划的一部分，日本政府修改了NISA的部分条款。当月，日本五大投资平台上就新开了90万个NISA账户。

经济形势为岸田文雄的努力提供了额外的助力。在日本的零利率政策下，

将现金存在银行几乎没有收益。这种情况在日本已有时日，但现在的通胀率达到3%左右，为30年来的最高水平，这意味着如果钱放着不用，就等于在贬值。没有经历过股市泡沫破裂创伤的年轻一代更愿意投资。

东京一家理财学校ABCash向千禧一代传授投资理财知识，自2022年以来它的学员数量翻了一番，达到四万人。新条沙也加在Instagram上看到一个网红提到该学校后，也成为了它的学员。“我很难和家里人谈论怎么管钱。”她说。岸田文雄面临的一个问题是，许多年轻人喜欢投资国际市场胜过国内市场。例如，斋藤的投资包括美国科技巨头苹果、美国大公司指数标普500以及德国疫苗制造商BioNTech等。不过，如果日经指数继续飙升，她和其他日本年轻人或许就会改变做法。 ■



Held in suspense

Do not expect America's interest rates to fall just yet

The risk of a second wave of inflation remains too great

HAS INFLATION been vanquished, or is it bouncing back? The question grips bond markets and governs the world's economic prospects. At the turn of the year, after the Federal Reserve all but declared victory over America's excessive price rises, bond yields collapsed in expectation of several interest-rate cuts. Today that bet looks premature. Over the past three months core consumer prices, which exclude food and energy, have risen at an annual rate of 4%, up from 2.6% in the three months to August. Producer prices have risen more than expected and consumers' expectations of inflation over the next year have gone up, too. Inflation is much lower than at its peak, but it has not yet been defeated. As a result, Treasury yields are roughly back to where they were before the Fed's doveish turn. Yields on long-dated bonds are higher still.

Inflation is also proving stickier elsewhere. The euro zone recorded large price rises in January, Swedish inflation rose in January and the Reserve Bank of Australia recently warned that inflation will take time to become "sustainably low and stable". Everywhere, and especially in America, a resurgence of inflation threatens to delay cuts to interest rates.

To understand what is going on, look at the pattern of demand. The covid-19 pandemic led to lockdowns and generous handouts, which fuelled the demand for goods and overwhelmed supply chains. In 2021 in America the price of a washing machine rose by 12%. After Russia invaded Ukraine, energy and food prices soared. Today, by contrast, global goods prices are falling, despite being buffeted by disruptions to shipping in the Red Sea. Food and energy prices are moderate. The source of today's problem is rises in the price of global services.

Inflation in services is tightly linked to local conditions. If haircuts are in short supply, you cannot import them. It is, therefore, no surprise that the cost of services is rising in America. Over the past three months its booming jobs market has created an average of 289,000 jobs a month, more than double estimates of the sustainable rate. Wages are growing at an annual pace of more than 4.5%. GDP rose by an annualised 4.1% in the second half of 2023 and real-time indicators suggest that the expansion remains healthy. Such is the economy's strength that, even if there had been no inflation in the past two years, forecasters would be expecting it to break out.

Elsewhere, the situation is more finely balanced. In Europe unemployment is low but growth less strong. Britain fell into recession at the end of 2023. Business surveys in the euro zone are gloomy. Inflation fears are limited by sharply lower natural-gas prices. That will reduce inflation and raise the rate of sustainable, non-inflationary economic growth. The Chinese economy is in a funk and prices are falling. In Japan interest rates are still below zero. It would be considered a victory if inflationary pressures stayed strong enough to let the central bank raise rates.

If a disruptive second wave of inflation is likely to strike anywhere it is in America. That means monetary policy is likely to diverge, with the Fed keeping interest rates high even as the rest of the rich world cuts them to revive growth. Such a pattern would boost the value of the dollar, which is already climbing. When America's money markets offer high returns for little risk, the many poorer countries struggling to borrow in dollars suffer most.

If interest rates do not fall there could be nasty surprises on Wall Street, too. The stockmarket appears to have barely noticed the danger that monetary policy will stay tight this year. By contrast, rising rate expectations in 2022 and 2023 often caused sell-offs as investors adjusted the discounted value

of future profits. Neither is there much sign that America's small and midsize banks have resolved the threats that high rates pose to their balance-sheets. The inflation problem is not what it was a year ago, but the world is not yet clear of the danger. ■



【首文】悬而未决

先别指望美国的利率会下降

第二轮通胀的风险仍然很大

通货膨胀是被征服了，还是在反弹？这个问题牵动着债券市场，也关系着世界经济的前景。岁末年初，在美联储近乎要宣布已经战胜了美国过高的物价上涨之后，债券收益率因人们预期会多次降息而暴跌。今天看来，这样下注似乎还为时过早。在过去三个月里，不包括食品和能源的核心消费者价格以年化4%的速度上涨，高于去年6月至8月的2.6%。生产者价格涨幅超过预期，消费者对明年通胀的预期也有所上升。通胀虽已远低于峰值，但尚未被击败。美国国债收益率随之又大致回到了美联储转向鸽派之前的水平。长期债券的收益率还要更高。

事实证明，其他地方的通胀也更具粘性。欧元区在1月录得物价大幅上涨，瑞典的通胀在1月上升，澳大利亚央行最近警告说，通胀要“持续走低并稳定”还需要时间。在各个地方，尤其是在美国，通胀的回升都有可能导致推迟降息。

要理解现在的情况，来看看需求模式。新冠疫情导致了封锁和慷慨的补贴，这刺激了对商品的需求并让供应链不堪重负。2021年，美国的洗衣机价格上涨了12%。俄罗斯入侵乌克兰后，能源和食品价格飙升。相比之下，尽管受到红海航运中断的冲击，如今全球的商品价格正在下跌。食品和能源价格适中。今天问题的根源是全球服务价格上涨。

服务业通胀与本地状况密切相关。如果理发服务供不应求，那是没法进口的。因此，美国的服务价格上升也就不足为奇了。过去三个月里，美国火热的就业市场平均每月创造28.9万个工作岗位，是估计的可持续数字的两倍多。工资以每年超过4.5%的速度增长。2023年下半年GDP的年化增长率为4.1%，实时指标显示这样的增长依旧健康。经济如此强劲，即使过去两年没有出现通胀，预测人员也会预计通胀将会爆发。

在其他地方，现状是一种更加微妙的平衡。欧洲的失业率较低，但增长不那么强劲。英国在2023年底陷入衰退。欧元区的企业景气调查结果令人沮丧。天然气价格大幅下跌限制了对通胀的担忧。这将降低通胀，提升可持续的无通胀经济增长。中国经济阴霾密布，物价正在下跌。在日本利率仍低于零，如果通胀压力持续而足以促使日本央行加息，这会被视为一场胜利。

如果说具扰乱性的第二波通胀可能会袭击什么地方，那就是美国。这意味着货币政策可能会分化：美联储把利率维持在高位，而同时其他富裕国家降息以重振经济增长。这种模式将进一步推高本已在攀升的美元价值。当美国的货币市场以很小的风险提供高回报时，许多难以借入美元贷款的较贫穷国家就会承受最大的损失。

如果利率不降，华尔街也可能出现恼人的意外。美国股市看起来几乎还没有留意到今年货币政策将保持紧缩的危险。相比之下，在2022年和2023年加息预期往往导致抛售，因为投资者调整了未来利润的贴现价值。也没有太多迹象表明美国的中小型银行已经解决了高利率对其资产负债表构成的威胁。通胀问题已不同于一年前，但世界尚未摆脱这一危险。 ■



Bartleby

How to benefit from the conversations you have at work

Stop thinking about your next point and listen to the one being made

SUCCESSFUL WORKPLACES are usually characterised by good communication. Bosses provide a clear sense of where they want the firm to go; employees feel able to voice disagreements; colleagues share information rather than hoarding it. But being a good communicator is too often conflated with one particular skill: speaking persuasively.

In a paper published in 2015, Kyle Brink of Western Michigan University and Robert Costigan of St John Fisher College found that 76% of undergraduate business degrees in America had a learning goal for presentation skills, but only 11% had a goal related to listening. Business students were being schooled to give TED talks rather than have conversations. That may have costs. Another study, conducted by Dotan Castro of the Hebrew University of Jerusalem and his co-authors, found that when people felt listened to by those in supervisory roles their creativity and sense of psychological safety improved.

A focus on talking is understandable. The set-piece moments of careers, like job interviews and big presentations, are about transmitting information. The boss gets to be at the podium, the minions get to be in the audience. Videos of someone giving a speech are much more shareable than someone silently nodding. But interest in what makes everyday communication tick has also risen, as the importance of teams grows and as conceptions of leadership increasingly emphasise softer skills.

Recent research by Beau Sievers of Stanford University and his co-authors asked groups of MBA students to discuss the meaning of ambiguous film clips. The presence of people perceived to be of high status seemed to

impede consensus: these folk spoke more and were readier to reject the explanations of others. Groups that reached consensus were more likely to have a different character in them: people who were well-connected but not dominant, who asked lots of questions and who encouraged interaction. They made everything align—even the neural activity of their groups.

Mr Sievers's research features in "Supercommunicators", a new book by Charles Duhigg, a journalist at the New Yorker. Mr Duhigg looks at how some people forge stronger connections with others and at the techniques for having better conversations. His canvas ranges more widely than the workplace but some of its lessons are applicable there.

One chapter tells the story of the Fast Friends Procedure, a set of 36 increasingly intimate questions that are particularly effective at turning strangers into friends. The questions were first put together in the 1990s by Elaine and Arthur Aron, two psychologists at the State University of New York at Stony Brook. Their survey was designed for the lab, not the workplace. You should not suddenly start asking new colleagues what their most terrible memory is or how they feel about their mother. But if it is important to build team connections fast, then—Britain, look away now—reciprocal moments of vulnerability do seem to help.

Another chapter looks at ways to bring together people with diametrically opposed views, in this case Americans on either side of the debate over gun control. The difficulty here was in persuading people that they were genuinely being listened to, not dismissed as gun-toting loons or lily-livered liberals. Mr Duhigg describes an approach called "looping for understanding", in which people ask questions and then repeatedly distil their understanding of what they have heard back to their interlocutor.

Polarised beliefs of this sort are rare inside firms. But looping techniques still have their place: when there are long-running conflicts between

individual employees, say, or in negotiations and mediation processes.

Mr Duhigg's advice can seem obvious at times. And his examples do not always translate to the workplace. Sometimes it is more important to make a decision than to excavate everyone's point of view. Reaching consensus is vital on a jury but less necessary in a corporate hierarchy. There really is a limit to how much vulnerability you want from a leader.

But his book is a useful reminder that demonstrable curiosity about other people's experiences and ideas can benefit everyone. Asking questions, not cutting people off, pausing to digest what someone has said rather than pouncing on breaks in a discussion to make your own point: these are not enough to qualify someone as a supercommunicator. But in plenty of organisations they would still represent good progress. ■



巴托比

如何从职场交谈中获益

不要只顾自己滔滔不绝，也听听别人说了什么

成功工作场所的一个常见特征是良好的沟通。老板清楚指明公司的发展方向，员工可以自如表达不同意见，同事间分享信息而不是互有隐瞒。但是，人们往往将优秀的沟通能力与另一种技能混为一谈：说话有说服力。

西密歇根大学（Western Michigan University）的凯尔·布林克（Kyle Brink）和圣约翰费舍尔学院（St John Fisher College）的罗伯特·科斯蒂根（Robert Costigan）在2015年发表的一篇论文中指出，美国76%的商科本科都有关于演讲技巧的学习目标，但只有11%设置了与倾听相关的目标。商科学生接受的教育是去做TED演讲，而不是交谈。这也许是有代价的。耶路撒冷希伯来大学的多坦·卡斯特罗（Dotan Castro）及其他人进行的另一项研究发现，当人们感到自己的意见得到管理者的倾听时，他们的创造力和心理安全感都会提高。

看重讲话能力不难理解。许多职业生涯的关键时刻，例如求职面试和大型演讲，都需要传递信息。老板可以站上讲台，下属则坐在观众席中。某人演讲的视频远远比某人默默点头的视频更容易得到传播。但是，随着团队的重要性与日俱增，且领导力的概念越来越强调软技能，如何做好日常沟通也引起了更多兴趣。

在斯坦福大学的博·西弗斯（Beau Sievers）等人近期所做的研究中，他们让几组MBA学生讨论一些模棱两可的电影片段的含义。如果有公认地位更高的人在场，似乎更难达成共识：这些人说的更多，而且动辄否定别人的解释。达成共识的小组更可能包含另一种人：这些人的人脉很广但不盛气凌人，他们经常提问并鼓励互动。这些人的存在让一切都步调一致——甚至是小组成员的神经活动。

《纽约客》记者查尔斯·杜希格（Charles Duhigg）的新书《超级沟通

者》（Supercommunicators）介绍了西弗斯的研究。杜希格观察了一些人如何与他人建立更紧密的联系，以及实现更顺畅交谈的技巧。他观察的范围不只限于职场，但其中一些经验教训也适用于职场。

其中一章介绍了“快速交友程序”，这是一套36个逐渐变得私人的问题，能够特别有效地将陌生人变成朋友。纽约州立大学石溪分校（State University of New York at Stony Brook）的两位心理学家伊莱恩·阿伦（Elaine Aron）和亚瑟·阿伦（Arthur Aron）在1990年代首次整理出了这套问题。他们的研究是面向实验室的，而不是职场。你可不要突然开始询问新同事他们最可怕的记忆是什么，或者他们对自己母亲的感觉如何。但是，如果必须要快速建立团队连结的话，那么——你要是在英国，就不用看了——相互坦诚自己的脆弱时刻似乎确实有用。

另一章探讨了如何让观点截然相反的人能够共处，所举的例子是因控枪争论而分为两派的美国人。这里的难点是要让对方相信自己的意见确实得到了倾听，而不是被蔑视为舞枪弄炮的疯子或者内心柔弱的自由主义者。杜希格描述了一种名为“循环理解”的方法——先提出问题，听到对方回答后把自己对其内容的理解提炼出来，反馈给对方，如此不断反复。

在公司里，这种观念两极分化的情况并不多见。但这种循环法仍有用武之地：例如个别员工之间长期闹矛盾时，或者在谈判和调解的过程中。

杜希格的建议有时似乎就是明摆着的。他举出的例子也并不总能延伸到职场上。有些时候，做出决定比征求每个人的意见更重要。在陪审团中，达成共识是至关重要的，但在公司层级制度中就没那么必要。你确实也不想看到你的上司展现出太多的脆弱。

但他的书给出了有益的提醒：对别人的经历和想法表现出好奇心对每个人都有好处。要提出问题，不打断别人，要停下来消化别人说的话，而不是抓住讨论的每个间隙表达自己的观点：这些都不足以让一个人成为超级沟通者。但在许多组织中，这已经是很大的进步了。■



Baby AI

Scientists have trained an AI through the eyes of a baby

“Chair” and “ball” were among little AI’s first words

FOR DECADES linguists have argued over how children learn language. Some think that babies are born as “blank slates” who pick up language simply from experience—hearing, seeing and playing with the world. Others argue that experience is not enough and that babies’ brains must be hardwired to make acquiring language easy.

AI models such as GPT-4 have done little to settle the debate. The way these models learn language—by trawling through reams of text data from millions of web pages—is vastly different to the experiences of babbling babies.

A team of scientists at New York University examined the question by training an AI model on the experiences of a single infant. Between the ages of six and 25 months, a toddler called Sam wore a head-mounted camera for an hour a week—around 1% of his waking hours. The camera recorded everything he saw and heard while he played with toys, enjoyed days at the park and interacted with his pet cats. The recordings and transcribed audio were fed into an AI, which was set up to know that images and words that appeared at the same time were related, but was otherwise left to make sense of the mess of colours and speech that Sam experienced.

Despite the limited training data, the AI was able to pick out objects and learn the matching words. The researchers tested the model by asking it to identify objects that Sam had seen before, such as a chair from his home or one of his toy balls. Given a list of four options the model picked the correct word 62% of the time, far above the chance level of 25%. To the researchers’ surprise, the model could also identify chairs and balls that Sam had never

seen. The AI learnt at least 40 different words, but it was far from matching Sam's vocabulary and language abilities by the end of the experiment.

The researchers, published recently in the journal *Science*, argue that, to match words to objects, learning from experience may well be enough. Sceptics, however, doubt that the AI would be able to learn abstract nouns or verbs, and question how similar the learning processes really are. The mystery of language acquisition lives on. ■



宝宝AI

科学家通过婴儿的视角训练AI

“椅子”和“球”是这个宝宝AI学会的首批词汇之一

几十年来，语言学家对儿童是如何习得语言的争论不休。一些人认为婴儿生来“白纸一张”，仅凭体验（听、看和与外界互动）来学习语言。其他人则认为体验还不够，婴儿的大脑必然天生具备一些条件，能让语言习得变得容易。

GPT-4等AI模型在解决这个争论方面作用甚微。这些模型通过搜罗千百万网页上的海量文本数据来学习语言，这与婴儿咿呀学语的过程截然不同。

纽约大学的一个科学家团队对这个问题做了研究，他们根据一个婴儿的学语过程来训练AI模型。在一个名叫山姆的幼儿从6个月至25个月大的这段时间里，科学家让他佩戴头戴式摄像机，记录下他玩玩具、去公园以及在家中宠物猫互动时的情景，每周共戴一小时（约占他醒着的时间的1%）。摄像机记录下来的场景和根据音频转写的文字被输入到一个AI模型中，AI根据设定知道同时出现的图像和单词有关联，但它要自己理解山姆看到的各色图案和听到的话语。

尽管训练数据有限，AI模型仍能够辨别出物体，并学习对应的单词。研究人员要求模型识别山姆曾看到过的物体，以此测试它的学习效果，比如要求它识别山姆家里的一把椅子或者他的一个玩具球。面对四个备选词，模型的选择准确率达到了62%，远高于25%的随机水平。让研究人员惊讶的是，模型还能识别山姆从未见过的椅子和球。AI模型至少学会了40个单词，但在实验结束时，它的词汇和语言能力与山姆相去甚远。

这项研究最近发表在《科学》杂志上，研究人员认为，要将单词与物体匹配起来，仅凭体验学习可能已经足够。然而，持怀疑态度的人认为AI无法学习抽象名词或动词，并质疑AI这一学习过程与幼儿学语真有多少相似性。语言习得的谜团仍然未解。■



They're on a roll

Ancient, damaged Roman scrolls have been deciphered using AI

The new techniques could help rediscover lost works from antiquity

“I CAN’T BELIEVE it worked!” says Nat Friedman, co-founder of the Vesuvius Challenge, which offered \$1m in prizes to anyone who could use artificial intelligence (AI) to decipher papyrus scrolls carbonised by the eruption of Mount Vesuvius in 79AD. But work it did. On February 5th Mr Friedman announced that a three-person team had been awarded \$700,000 for successfully extracting four passages of text, each at least 140 characters long, and with at least 85% of the characters legible, from a scroll known as Banana Boy. The three winners, Luke Farritor, Youssef Nader and Julian Schilliger, are all computer-science students.

The scroll is one of hundreds found in the library of a Roman villa in Herculaneum, which is thought to have belonged to the father-in-law of Julius Caesar. Along with hundreds of other scrolls in the villa’s library, it was damaged by scorching gases that engulfed the town during the same eruption that also buried the nearby town of Pompeii.

Reading text from the scrolls is difficult because the heat turned them into brittle charcoal logs; all efforts to unroll them physically caused them to disintegrate. So attention shifted towards finding ways to unwrap them virtually, through computer analysis of 3D scans of the scrolls made using X-rays. This turned deciphering the scrolls into a software problem—but a very complex one.

Virtual unrolling is a two-stage process pioneered by W. Brent Seales, a computer scientist at the University of Kentucky. The first stage, called segmentation, involves tracing the edges of the rolled-up papyrus sheet inside the 3D scan, then extracting 2D images of the scroll’s surface. The

second stage, ink detection, analyses the resulting images to pluck the ink of the scroll's text from the papyrus background. This is particularly tricky for the Herculaneum scrolls, which are written in carbon-based ink, so there is very little contrast with the background of carbonised papyrus.

Dr Seales, along with Mr Friedman and Daniel Gross, two technology entrepreneurs, thought AI techniques might fruitfully be brought to bear on these two problems, and launched a prize challenge to find out. A community of thousands of enthusiasts has since developed a range of tools and tricks to speed up the fiddly process of segmentation, and to detect the ink of individual letters, and then whole words. In October 2023 Mr Farritor and Mr Nader were awarded smaller prizes for independently extracting the first legible word ("porphyras", which means "purple" in ancient Greek) from the Banana Boy scroll (so named because of its size and shape).

The two students then teamed up and, joined by Mr Schilliger, further improved the machine-learning technique involved in ink detection. By manually labelling areas known to be ink, they could train a neural network to find more of them; these were fed back into the model to improve its detection abilities. Mr Nader also switched the neural network to a novel architecture called a TimeSformer, which produced sharper results. Mr Schilliger, meanwhile, devised a tool to automate more of the segmentation process (much of which must still be done manually).

The deadline to submit results for the grand prize was at the end of December, and the trio was awarded the prize after an assessment of the entries by a team of papyrologists. (Three runners up will receive smaller prizes of \$50,000 each.) The winning entry revealed 15 columns of text, written in Greek. Reading it was "mind-blowing", says Federica Nicolardi, a papyrologist at the University of Naples Federico II, who was one of the judges. The text is thought to be a previously unknown work on pleasure by

Philodemus, an Epicurean philosopher who lived in Herculaneum.

Mr Friedman now wants to scale up the whole process. With ink detection solved, he says, “the bottleneck is now segmentation”. Mr Schilliger’s auto-segmentation tool is a big step forward, and he has agreed to make it open source, and to collaborate with others to improve it. Further prizes are being offered as an incentive. Mr Friedman, meanwhile, aims to scan more scrolls using the Diamond Light Source, a particle accelerator in Britain, and to standardise the scanning process.

That will cost money. Having given out \$1.2m in prizes, some of it from his own pocket, Mr Friedman is looking for other backers to help support the project. He hopes that deciphering ancient scrolls will lead to the rediscovery of lost works from antiquity—“each scroll is a mystery box”, he says—and, ultimately, revive interest in further excavating the villa in Herculaneum, which may contain thousands more of them. ■



势如破卷

利用AI成功破译烧毁的古罗马卷轴

这种新方法能帮助重新发现失传的古代作品

“难以置信啊！它真的成功了！”维苏威火山挑战赛（Vesuvius Challenge）的发起人之一纳特·弗里德曼（Nat Friedman）说。该挑战赛设有多个奖项，奖金总计100万美元，只要能够利用人工智能（AI）破译在公元79年维苏威火山喷发中碳化的莎草纸卷轴，就能获得相应奖项。举办挑战赛的做法确实奏效了。2月5日，弗里德曼宣布一个三人团队获得了70万美元的奖金，因为他们从一个叫作“香蕉小子”（Banana Boy）的卷轴中成功辨认出四段文本——每段至少包含140个字母，且至少辨认出85%的字母。三位获奖者卢克·法里托（Luke Farritor）、尤瑟夫·纳德（Youssef Nader）和朱利安·席利格（Julian Schilliger）都是计算机科学专业的学生。

香蕉小子是在赫库兰尼姆（Herculaneum）一座古罗马庄园的藏书楼中发现的，人们认为这座庄园的主人是凯撒大帝的岳父。它和庄园藏书楼里的其他数百个卷轴一道，被火山喷发时吞没了整个赫库兰尼姆的灼热气体毁坏。也正是这场火山喷发埋葬了附近的庞贝城。

由于高温把卷轴变成了易碎的炭木，要阅读上面的文字并非易事；任何试图展开卷轴的举动都会让它们变成碎片。因此，人们转而想办法虚拟打开卷轴，也就是用X射线对卷轴进行3D扫描，再做电脑分析。这就把破译卷轴变成了一个软件问题，不过这个软件问题非常复杂。

最先采用两步法虚拟展开卷轴的是肯塔基大学的计算机科学家布伦特·西尔斯（W. Brent Seales）。第一步称为分割，勾画出3D扫描图里卷起的莎草纸的边缘，然后生成卷轴表面的2D图像。第二步是墨迹识别——分析生成的图像，从莎草纸底色中提取卷轴文本的墨迹。这对于赫库兰尼姆卷轴来说尤为棘手，因为它们是用碳基墨水书写的，与碳化的莎草纸底色几乎没有差别。

西尔斯与两位科技企业家弗里德曼、丹尼尔·格罗斯（Daniel Gross）都认为，AI技术或许可以有效解决这两个问题，为此他们发起了一项有奖挑战赛来一探究竟。此后，数千名爱好者陆续开发了一系列工具和方法来加快繁琐的分割过程，以及识别单个字母的墨迹，进而识别整个单词。去年10月，法里托和纳德不约而同地从香蕉小子卷轴（因其大小和形状类似香蕉而得名）中发现了第一个可辨认的单词（“porphyras”，在古希腊语中意为“紫色”），分别赢得一个金额较小的奖项。

这两名学生随后联手，席利格也加入其中，他们三人进一步改进了墨迹识别中使用的机器学习技术。通过人工标记已知是墨迹的区域，他们可以训练一个神经网络来找到更多墨迹区域；这些信息被投喂回模型中，以提高其识别能力。纳德还将这个神经网络转换成生成更清晰结果的新架构TimeSformer。与此同时，席利格设计了一种工具，让分割过程更加自动化（大量分割目前仍须人工完成）。

角逐大奖作品的最后提交日期是去年12月底，经过一个莎草纸古文稿专家小组对参赛成果的评选，大奖最终花落这个三人团队。（三名亚军将分别获得五万美元的奖项。）获奖的成果辨认出了15列用希腊语书写的文本。其中一名评委、意大利那不勒斯费德里科二世大学（University of Naples Federico II）的莎草纸古文稿专家费代丽卡·尼古拉迪（Federica Nicolardi）表示，读到这些文字“令人无比兴奋”。这些文本被认为是居住在赫库兰尼姆的伊壁鸠鲁派哲学家菲洛德摩斯（Philodemus）先前一部不为人知的有关快乐的著作。

弗里德曼现在想要扩大整个项目的规模。墨迹识别问题得以解决后，他表示，“现在的瓶颈是分割”。席利格的自动分割工具是一大进步，他已经同意将其开源，并与他人合作来改进它。作为激励，还有更多的奖项虚位以待。与此同时，弗里德曼的目标是使用英国的一种叫作钻石光源（Diamond Light Source）的粒子加速器来扫描更多卷轴，并让扫描过程标准化。

这需要资金。弗里德曼已经发放了120万美元的奖金，其中一些是他自掏

腰包，目前他正在寻找其他赞助人来帮助支持这个项目。他希望通过破译这些古卷轴，能够重新发现一些失传的古代作品——“每幅卷轴都是个盲盒”，他表示——并最终重新唤起人们对进一步发掘赫库兰尼姆庄园的兴趣，那里可能还埋藏有成千上万部作品。■



Bartleby

Why you should never retire

Pleasure cruises, golf and tracing the family tree are not that fulfilling

IN AN EPISODE of “The Sopranos”, a popular television series which started airing in the 1990s, a gangster tells Tony, from the titular family, that he wants to retire. “What are you, a hockey player?” Tony snaps back. Non-fictional non-criminals who are considering an end to their working lives need not worry about broken fingers or other bodily harm. But they must still contend with other potentially painful losses: of income, purpose or, most poignantly, relevance.

Some simply won’t quit. Giorgio Armani refuses to relinquish his role as chief executive of his fashion house at the age of 89. Being Italy’s second-richest man has not dampened his work ethic. Charlie Munger, Warren Buffett’s sidekick at Berkshire Hathaway, worked for the investment powerhouse until he died late last year at the age of 99. Mr Buffett himself is going strong at 93.

People like Messrs Armani, Buffett or Munger are exceptional. But in remaining professionally active into what would historically be considered dotage, they are not unique. One poll this year found that almost one in three Americans say they may never retire. The majority of the nevers said they could not afford to give up a full-time job, especially when inflation was eating into an already measly Social Security cheque. But suppose you are one of the lucky ones who can choose to step aside. Should you do it?

The arc of corporate life used to be predictable. You made your way up the career ladder, acquiring more prestige and bigger salaries at every step. Then, in your early 60s, there was a Friday-afternoon retirement party, maybe a gold watch, and that was that. The next day the world of meetings,

objectives, tasks and other busyness faded. If you were moderately restless, you could play bridge or help out with the grandchildren. If you weren't, there were crossword puzzles, TV and a blanket.

Although intellectual stimulation tends to keep depression and cognitive impairment at bay, many professionals in the technology sector retire at the earliest recommended date to make space for the younger generation, conceding it would be unrealistic to maintain their edge in the field. Still, to step down means to leave centre stage—leisure gives you all the time in the world but tends to marginalise you as you are no longer in the game.

Things have changed. Lifespans are getting longer. It is true that although the post-retirement, twilight years are stretching, they do not have to lead to boredom or to a life devoid of meaning. Once you retire after 32 years as a lawyer at the World Bank, you can begin to split your time between photography and scrounging flea markets for a collection of Americana. You don't have to miss your job or suffer from a lack of purpose. If you are no longer head of the hospital, you can join Médecins Sans Frontières for occasional stints, teach or help out at your local clinic. Self-worth and personal growth can derive from many places, including non-profit work or mentoring others on how to set up a business.

But can anything truly replace the framework and buzz of being part of the action? You can have a packed diary devoid of deadlines, meetings and spreadsheets and flourish as a consumer of theatre matinees, art exhibitions and badminton lessons. Hobbies are all well and good for many. But for the extremely driven, they can feel pointless and even slightly embarrassing.

That is because there is depth in being useful. And excitement, even in significantly lower doses than are typical earlier in a career, can act as an anti-ageing serum. Whenever Mr Armani is told to retire and enjoy the

fruits of his labour, he replies “absolutely not”. Instead he is clearly energised by being involved in the running of the business day to day, signing off on every design, document and figure.

In “Seinfeld”, another television show of the 1990s, Jerry goes to visit his parents, middle-class Americans who moved to Florida when they retired, having dinner in the afternoon. “I’m not force-feeding myself a steak at 4.30 just to save a couple of bucks!” Jerry protests. When this guest Bartleby entered the job market, she assumed that when the day came she too would be a pensioner in a pastel-coloured shirt opting for the “early-bird special”. A quarter of a century on, your 48-year-old columnist hopes to be writing for The Economist decades from now, even if she trundles to her interviews supported by a Zimmer frame; Mr Seinfeld is still going strong at 69, after all. But ask her again in 21 years. ■



巴托比

为何你应当永不言退

乘游船、打高尔夫和研究家谱的生活并不够充实

在1990年代首播的热剧《黑道家族》（The Sopranos）的一集里，一个黑帮分子对该家族的老大托尼说自己想退休了。“你当自己是什么人，冰球运动员吗？”托尼厉声斥道。在现实中，不混黑帮的人在考虑结束自己的职业生涯时，不必担心手指被人折断或遭受其他身体伤害。但他们仍必须应对其他可能令人痛苦的损失：收入、目标，或者最令人心酸的——社会存在感。

有些人干脆就不退下来。已经89岁的乔治·阿玛尼（Giorgio Armani）仍然拒绝辞去在自己时装公司首席执行官的职务。他是意大利第二大富豪，但这并不折损他的工作热情。巴菲特在伯克希尔哈撒韦公司（Berkshire Hathaway）的副手查理·芒格（Charlie Munger）一直为这家投资巨头工作，直至去年年底去世，享年99岁。巴菲特本人也已93岁高龄，仍然干劲十足。

像阿玛尼、巴菲特或芒格这样的人物不是一般人。但是，在以往被认为是老态龙钟的年龄仍然活跃于职场这一点上，他们并非特例。今年的一项民意调查发现，近三分之一的美国人表示自己可能永远不会退休。大部分“永不退休者”表示，他们负担不起放弃全职工作的后果，尤其是当通胀正让本就微薄的社会保障金缩水。但假设你是可以选择退位的幸运儿之一，你该退休吗？

过去，职场人生是有章可循的。你在事业的阶梯上步步攀升，每一步都获得更高的声望和薪水。然后，在你60岁出头的时候，在某个周五下午参加了一场退休派对，也许还领到了一块金表，事情就这样结束了。第二天，充满会议、目标、任务和其他繁杂的世界就离你远去了。如果你多少有点闲不住，可以打打桥牌或给带娃的子女搭把手。如果能闲得下来，陪伴你的有填字游戏、电视和毛毯。

尽管智力刺激往往可以避免抑郁和认知障碍，但科技圈里的许多专业人士都会尽早按建议退休，以便为年轻一代腾出位置；他们承认要在这个领域里保持住自己的优势是不现实的。尽管如此，退休意味着离开中心舞台——休闲给了你无尽的时间，但往往会使你在这个世界上边缘化，因为你已经“出局”了。

现在非同以往。人的寿命越来越长。诚然，退休后暮年的不断延长不一定就意味着百无聊赖或人生失去意义。在世界银行当了32年律师的你一旦退休，就可以开始花一半时间搞摄影，花另一半时间在跳蚤市场淘一整套《大美百科全书》。你不必怀念自己的工作，也不必为没了目标而痛苦。如果你不再是医院院长，你可以加入无国界医生组织做点零散的工作，在本地诊所带带新医生或帮点忙。很多地方都可以实现自我价值和个人成长，包括非营利工作或指导他人创业。

但是，有什么能真正取代参与某种事业或行动带给人的体系感和兴奋感吗？你可以把日程排得满满当当，没有最后期限、会议和电子表格，可以尽情享受午后剧场、艺术展览和羽毛球课。对很多人来说，投入兴趣爱好挺不赖的。但对于那些极为进取的人来说，兴趣爱好可能是漫无目的、毫无意义的事情，甚至略显尴尬。

这是因为做“有用”之人带来一种严肃性和厚重感。而即使兴奋感已经远远没有职业生涯早期那么强烈，它仍然是一剂抗衰老的良药。每当有人劝阿玛尼退休，好好享受自己的劳动成果时，他的回答都是“绝不”。相反，他参与到公司的日常运营，签批每一个设计、文件和数字，这显然给他带来了源源活力。

在另一部90年代电视剧《宋飞正传》（Seinfeld）中，杰瑞去看望他的父母。这对美国中产老夫妇在退休后迁居佛罗里达，下午就开始早早地吃晚饭。“我可不会为了省几块钱在4点半逼着自己吃牛排！”杰瑞抗议道。当笔者初入职场时，她以为自己朝一日也会成为一个穿着粉彩衬衫、选用“早鸟优惠”的退休人士。四分之一世纪过去了，48岁的笔者希望几十年后还能继续为本刊撰稿，哪怕她得推着一台助行架蹒跚着前去采访；毕

竟，69岁的宋飞都还劲头十足呢。不过，不妨21年后再问她一次吧。■



Notice

The Economist Group will close down the Global Business Review app

From February 28th, you will not be able to renew your subscription or start a new subscription

We are writing to make you aware that The Economist Group will be closing the Global Business Review (GBR). We wish to thank our valued subscribers and readers, as well as our GBR colleagues for their hard work and dedication.

The GBR app will be updated as normal until May 31st. From February 28th, you will not be able to renew your subscription or start a new subscription. All GBR subscribers will have continued access to the app until their next renewal, or until May 31st, whichever is soonest, at which point the app will close.

If your subscription is due to renew between February 28th and May 31st, your subscription will end at your next renewal.

If your expiration date for current subscription is after May 31st, you will receive a prorated refund to your existing payment method within 30 days of May 31st and your subscription will end on this date.

If someone currently pays for your subscription for you, the subscription will end on the next renewal date or May 31st, whichever is soonest.

If you have questions or require assistance, you can reach us via WeChat (Please search WeChat ID: ecogbr) or scan the QR code below:





通知

经济学人集团将关闭《经济学人·全球商业评论》APP

自2月28日起，您将无法续订或开始新的订阅

我们在此通知您，经济学人集团将关闭《经济学人·全球商业评论》。我们藉此向我们尊贵的订户和读者，以及《全球商业评论》同事们的辛勤工作和奉献表示感激。

《全球商业评论》APP将正常更新至5月31日。自2月28日起，您将无法续订或开始新的订阅。所有《全球商业评论》订户都可以继续使用该APP，直至下个续订日或5月31日APP停止运营之日（以较早者为准）。

如果您的订阅到期日是在2月28日至5月31日之间，您的订阅将在下次续订时终止。

如果您当前的订阅将在5月31日之后到期，您将在5月31日之后的30天内收到按比例退还至您现有付款方式的订阅款项，您的订阅也将在5月31日终止。

如果目前由他人为您代付订阅费用，您的订阅将在下一个续订日或5月31日终止，以较早者为准。

如果您有问题或需要帮助，可以通过微信（请搜索微信ID“ecogbr”）或扫描下方二维码联系我们：





Tokyo Electrified

Japan's semiconductor toolmakers are booming

How long will it last?

CORPORATE JAPAN, long snubbed by investors, has been the source of much attention lately. That is especially so for the country's manufacturers of chipmaking tools. On February 13th Tokyo Electron's share price jumped by 13% after it reported higher profits than expected for the final quarter of last year, pushing its market value above \$100bn, roughly ten times what it was worth a decade ago. It is now Japan's fourth-most-valuable company, trouncing better-known peers such as Mitsubishi, Nintendo and SoftBank Group.

Tokyo Electron is the largest among a crowd of Japanese purveyors of chipmaking tools for which business has been booming. The combined market capitalisation of the five most valuable—Tokyo Electron, Advantest, DISCO, Lasertec and SCREEN Holdings—has doubled, in dollar terms, over the past year (see chart). These companies, which suffered from a post-pandemic slump in sales of electronics that caused chipmakers to slash investments, are now on the up. The Semiconductor Equipment Association of Japan, an industry body, expects a 27% increase in sales of machinery this year and a further 10% rise in 2025, more than reversing the fall of 19% in 2023.

The latest rally is about more than the ups and downs of a notoriously cyclical industry. Demand for increasingly complex chips capable of handling artificial-intelligence functions means demand for more equipment to make them. One estimate from McKinsey, a consultancy, puts the cost to build and equip a factory that produces five-nanometre (nm) chips at around six times the cost of one that produces 28nm chips, which were cutting-edge a decade ago.

Japanese toolmakers have long played dominant roles in niche areas of the semiconductor supply chain. Tokyo Electron controls roughly 90% of the market for tools which apply photoresist coating—a light-sensitive chemical applied to a semiconductor wafer in the manufacturing process. DISCO holds a similarly dominant share of the precise cutting and grinding tools needed for chipmaking.

The revenues of these firms have thus far not been crimped by new export controls introduced by America on the sale of certain types of advanced chipmaking machinery to China. In fact, as China has raced to develop its own chipmaking capacity, sales of equipment to its neighbour to the east have surged. Almost half of Tokyo Electron's sales in the final quarter of last year came from China.

Yet that also points to risks ahead. If relations between Beijing and Washington worsen, the Japanese government could be prodded into implementing broader restrictions on the sale of its chipmaking tools to China. Meanwhile, China is expanding its effort not only to make its own chips, but also the machinery that produces them, which could squeeze out Japanese firms. For Tokyo Electron and its peers, the deteriorating relationship between America and China is a boon for business today. In the future, though, it may become a strain. ■



东京电激

日本的半导体设备制造商生意兴隆

这种情况能持续多久？

长期以来一直被投资者冷落的日本企业近来备受关注，尤其是芯片设备制造商。2月13日，就在东京电子公布了高于预期的去年第四季度利润后，其股价暴涨了13%，市值也因此超过1000亿美元，大约是十年前的十倍。如今东京电子已成为日本价值第四高的公司，超越了三菱、任天堂和软银集团等名气更大的日本企业。

日本众多芯片设备供应商近来生意红火，东京电子是其中最大的一家。价值前五大公司——包括东京电子、爱德万测试、迪思科、Lasertec和斯库林集团——以美元计的总市值在过去一年中翻了一番（见图表）。新冠疫情过后，电子产品销量骤降，导致芯片制造商大幅削减投资，令这些公司受挫，但现在它们正在企稳向好。行业组织日本半导体制造装置协会预计，今年半导体设备的销售额将增长27%，2025年将再增长10%，轻松扭转2023年下降19%的局面。

最近的反弹不仅仅源于以周期性著称的芯片行业的起伏。为实现各种AI功能需要越来越复杂的芯片，也就需要更多制造这些芯片的设备。咨询公司麦肯锡估计，建设和装备一家5纳米芯片制造工厂的成本大约是28纳米芯片工厂的六倍，而后者在十年前还是最先进的芯片。

长期以来，日本设备制造商一直在半导体供应链的细分领域占据主导。东京电子控制着光刻胶（芯片制造过程中一种涂在半导体晶片上的光敏化学物质）涂布设备市场大约90%的份额。迪思科在芯片制造所需的精密切割和研磨工具市场占有同样压倒性的份额。

美国就向中国出售某些先进的芯片制造设备实施了新的出口管制。但到目前为止，这些日本公司的收入还没有受到很大影响。事实上，随着中国急于发展自己的芯片制造能力，它从东边邻国日本购买的芯片制造设备数量

激增。东京电子去年第四季度的销售额近一半来自中国。

然而，这也预示着未来的风险。如果中美关系恶化，日本政府可能会被迫对中国实施更广泛的芯片制造设备出口限制。与此同时，中国不仅在加大力度研发自己的芯片，也在探索自己的芯片制造设备，这可能会将日本公司挤出市场。对于东京电子及其同行来说，中美关系恶化目前是生意上的一大利好。但在未来，它可能会成为一种难以承受之重。 ■



Silicon rally

As San Francisco builds the future of technology, can it rebuild itself?

People feared a doom loop. Reality has been more surprising

SAN FRANCISCO has long been a byword for municipal failure. Even as its techies minted money and transformed the world, its government was incapable of providing residents with basic shelter and security.

Homelessness, drug overdoses and property crime were rife. Then covid-19 struck. The rise of remote work threatened to sound the city's death knell, as the tech industry took to its heels. As things have turned out, however, San Francisco has become host to an artificial-intelligence boom. Having been granted this piece of good fortune, the city must seize the opportunity to reform. This might be the best chance it gets.

Only a few years ago social disorder, toxic politics, eye-watering housing costs and the pandemic were driving people away. Venture capitalists (VCs) were splashing more of their cash beyond the Bay Area. Tech workers abandoned their offices, and sometimes the city entirely. As the downtown streets emptied of workers, homelessness and public drug use became more conspicuous and intrusive. "Retail for lease" signs covered the city like wallpaper.

People keen to write San Francisco's obituary warned of a "doom loop". Abandoned downtown property would lead to budget cuts and the erosion of public services, which would in turn accelerate the exodus. Many people worried that the city could go the way of Detroit, which suffered a painful bust when America's car industry started to build more plants in the suburbs and the South.

But the power of agglomeration is such that San Francisco has been given a fresh lease on life. The strides in artificial intelligence (AI) have only

strengthened its claim to be the innovation capital of the world. Its proximity to Stanford and the University of California, Berkeley, two AI centres of excellence, has helped infuse its startup scene with companies that deal in the technology. No other place in the world has as many AI firms or as much tech talent: from OpenAI and Anthropic to Databricks and Scale AI, almost all the big startups are based in the Bay Area. Venture funding is on the rise again, and last year the vast majority of funding for firms in the region went to startups in the city itself. Despite San Francisco's problems, and there are many, it remains a magnet for capital and talent.

All this offers the city an opportunity to fix how it is run. It helps that voters are fed up with their city's inane politics (owing to oppressive permit rules, the cost of one public toilet ballooned to \$1.7m). From their posh neighbourhoods and their Silicon Valley headquarters, techies used to be content to watch San Francisco putrefy. Today, by contrast, the wealthy who live in the city feel a need to enter the political fray.

The first signs that San Francisco's residents were revolting against the status quo came in 2022 when voters ousted three members of their school board and recalled their ultra-progressive district attorney. They are not done yet. Local elections this year will be overshadowed by national races, obviously, but city contests in March and November will offer voters in San Francisco the prospect of real change.

Moderate Democrats, such as London Breed, the current mayor, are often stymied by left-wingers who resist building houses, cutting business taxes, shrinking the bloated budget or funding the police. Electing more moderates to the board of supervisors, San Francisco's name for the city council, could make all those things easier to accomplish.

Even if sensible folk get into office, making big changes will be hard. The

local NIMBY movement is entrenched, and corruption has long been a problem. And yet second chances are rare. Young people want to live and work in the City by the Bay and to be part of a technological revolution that is changing the world. San Francisco can either capitalise on their excitement and set about fixing its problems—or sink back into complacency and squander an unforeseen opportunity. ■



【首文】硅集

打造明日科技的旧金山能否重建自身？

人们曾担心它深陷厄运循环。现实却更出人意料

长期以来，旧金山就是市政失败的代名词。这里的科技狂人创造财富、改变世界，而政府这一边却无力为居民提供基本的安身之处和安全保障。无家可归、吸毒过量和财产犯罪很常见。接着，新冠来了。远程工作兴起，科技产业随之撤退，这座城市的丧钟仿佛就要敲响。然而，结果却是旧金山迎来了一波人工智能热潮。既然交上了如此好运，旧金山必须抓住机遇进行改革。这可能是它最好的机会了。

仅仅几年前，社会失序、有毒政治、令人瞠目的房价和疫情促使人们纷纷逃离。风险资本家（VC）将更多现金投向了湾区以外的地方。科技工作者放弃了他们的办公室，有时甚至干脆放弃了这座城市。随着市中心街道上的上班族越来越少，无家可归和公共场所吸毒越发惹眼和猖獗。“商铺出租”的招贴像墙纸一样糊满了整个城市。

那些等不及要给旧金山写讣告的人警告说，旧金山将陷入一个“厄运循环”。市中心房产遭废弃将导致预算削减和公共服务质量下降，这反过来又会加速人口外流。许多人担心旧金山可能会重蹈底特律的覆辙，当年美国的汽车工业开始在郊区和南部建造更多工厂时，底特律遭遇了惨痛的萧条。

但是，产业集聚的效应如此强大，旧金山又重获新生。人工智能领域的长足进步进一步巩固了它作为世界创新之都的地位。旧金山毗邻斯坦福大学和加州大学伯克利分校这两个卓越的人工智能中心，这让该市的创业圈涌现出大量人工智能技术相关公司。世界上再没哪个地方能有如此多的人工智能公司或科技人才：从OpenAI和Anthropi到Databricks和Scale AI，几乎所有的大型创业公司都位于湾区。风险投资再次兴起，去年该地区企业的绝大部分融资都流向了旧金山本地的创业公司。尽管旧金山有它的问题，而且问题还不少，但它仍然是吸引资金和人才的磁石。

这一切为旧金山提供了一个修复自身治理的机会。选民们已经受够了这座城市愚蠢空洞的政治运作（由于高压苛刻的许可规定，一个公共厕所的造价竟高达170万美元），这一点会提供助力。过去，科技从业者自己栖身高尚社区和硅谷总部，眼看着旧金山溃烂下去，心中毫无波澜。而如今，生活在这座城市的富人们却觉得有必要参与政治角力了。

旧金山居民反抗现状的最初迹象出现在2022年，当时选民们罢免了三名学校董事会成员和极端进步派的地区检察官。他们的行动还没有结束。今年的地方选举显然会被全国性选举抢去风头，但3月和11月的市政选举将有望为旧金山选民带来真正的变革。

现任市长伦敦·布里德（London Breed）这样的温和派民主党人常受到左翼人士的阻挠，后者反对建造房屋、削减商业税、缩减膨胀的预算或为警力提供资金。选举更多温和派人士进入监事会（旧金山对市议会的称呼）可能会让这些事情都更容易实现。

即使是明智务实的人当选了，要做出大的改变也会有难度。当地的邻避运动根深蒂固，腐败问题年深日久。但第二次机会不常见。年轻人想在旧金山生活和工作，希望投身一场正在改变世界的技术革命浪潮。旧金山要么善加利用他们的这股热忱，着手解决自身问题，要么就又回到沾沾自喜的状态，挥霍掉一个不期而至的良机。■



Boxing match

How worried should Amazon be about Shein and Temu?

Dirt-cheap products and marketing splurges are catching clicks

“SHOP LIKE a billionaire.” With that enticing slogan Temu touted itself to Americans watching the Super Bowl on February 11th. Football fans had been treated to a similar advert from the e-commerce company at last year’s event. But this time the message was hammered home. In all, Temu’s ad played five times. That won’t have been cheap. A 30-second slot during this year’s Super Bowl cost around \$7m. JPMorgan Chase, a bank, reckons the company will spend \$3bn on marketing this year, up from \$1.7bn in 2023.

Temu, based in Boston, is an offshoot of Pinduoduo, a Chinese e-commerce firm. It is attempting to replicate the success of Shein, a Chinese fast-fashion seller, which shot to success in America in 2021 helped by clever marketing and ultra-low prices. In their bid to win over American shoppers the duo are spending so lavishly on digital ads that their footprints show up in big tech companies’ earnings. On February 1st Susan Li, the chief financial officer of Meta, a social-media giant, said that Chinese advertisers contributed 10% of her firm’s revenue last year and five percentage points to its worldwide revenue growth. In November Josh Silverman, the boss of Etsy, an online marketplace for artisan knick-knacks, blamed Temu and Shein for pushing up the price of digital advertising.

All that spending has bought some brand recognition. Last year Temu was the most downloaded app in America, Britain, France and Germany, according to Business of Apps, a research firm. Shein is already a shopping sensation among American teenagers. But can the firms win over American shoppers en masse?

Although Temu stocks a broader range of items, from children’s toys to

industrial tools, than Shein, which mostly sells clothing, they have similar business models. For both, the main advantage is price. Michael Morton of MoffettNathanson, a research firm, estimates that the same items of women's clothing on Temu are two to four times dearer on American websites. Temu offers electric toothbrushes, sunglasses and backpacks for about \$1 each.

The pair keep costs low in many ways. One is to cut out middlemen and deal directly with Chinese factories. Another is to charge merchants lower fees than American rivals do. They also eschew enormous logistics operations in America like Amazon's. Instead they ship products from warehouses in China and have them delivered to shoppers in America by UPS, FedEx or the post. Customers thus get low prices, but not speedy delivery. Packages can take weeks to arrive.

For now, Temu is focused on market share rather than profit. Bernstein, a broker, reckons that it loses around \$10 per item sold in America. It may be able to keep this up for a while, bankrolled by Pinduoduo's domestic success. Last year the parent firm generated \$12bn in cash from operations. Shein, by contrast, is looking to raise more funds. In November it filed for an initial public offering in America. If the listing happens, it will be huge: at its most recent funding round in May 2022 the company was valued at \$66bn.

So far the pair have made only small inroads into America's e-commerce market. Temu and Shein both have shares of about 1%, according to Bernstein. Amazon has 38%. Even so, the local giant is taking them seriously. In December Amazon said it would cut merchants' fees for clothing priced under \$15, probably in response to the twin threat. In September it rolled out an end-to-end supply-chain service in which it picks up goods from merchants' factories and ships them to customers, mirroring what its Chinese rivals do.

But Amazon is unlikely to be hurt first, or most. Clothing and accessories account for only 16% of its sales, according to eMarketer, a research firm, so Shein is probably a bigger threat to fast-fashion labels such as Forever 21. And for Temu, the absence of a local logistics network will make it difficult to compete with Amazon when it comes to goods customers want quickly, such as dishwasher tablets. It is probably a bigger threat to eBay or Etsy.

That may eventually change. Both firms have bought warehouse space in America and struck partnerships with local logistics firms. Shein is reportedly poaching supply-chain specialists from Amazon. Temu may also start to sell dearer goods, such as smartphones, a shift its parent pulled off in China. That would put it into more direct competition with Amazon and the similarly mighty Walmart.

| *Cheap shots*

Plenty could still go wrong for the two. Each must compete both with American incumbents and with each other. TikTok Shop, a marketplace run by the social-media firm that launched in America in September last year, may also get in their way. Geopolitics may hurt them, too. A committee of American senators is probing their alleged links to forced labour. (Both firms deny any such connections.) Analysis by Morgan Stanley, another bank, suggests that Americans' willingness to shop at Temu may already be waning, perhaps as the novelty has worn off.

That said, Mark Shmulik of Bernstein argues that increasing market share from zero to 1% is more difficult for new e-commerce firms than raising it from 1% to 5%. Getting consumers to first take notice is tricky. Once a brand is familiar, it is easier to sell customers more things—and more expensive ones, too. Temu and Shein may have already done the hardest part of making it in America. ■



装箱擂台

亚马逊应该多担心Shein和Temu？

廉价产品和大手笔营销正在收获点击量

“像亿万富翁那样购物。”在2月11日的超级碗比赛中，Temu向美国观众打出了这个诱人的口号来宣传自己。这家电商公司在去年的超级碗比赛中也向球迷展示过类似的广告。但这一次，它火力全开。Temu的广告播放了五次。这必然价格不菲：今年超级碗的30秒广告收费约700万美元。根据摩根大通的估计，Temu今年将在营销上支出30亿美元，而2023年的支出为17亿美元。

总部位于波士顿的Temu是中国电商公司拼多多的子公司。它试图复制中国快时尚电商Shein于2021年在美国的异军突起，后者的成功得益于巧妙的营销和超低的价格。为了赢得美国消费者的青睐，这两家公司在数字广告上大手笔撒钱，在美国科技巨头的财报上留下了明显印记。2月1日，社交媒体巨头Meta的首席财务官苏珊·李（Susan Li）表示，中国广告主去年贡献了Meta营收的10%，以及5个百分点的全球收入增长。去年11月，在线手工艺品市场Etsy的首席执行官约什·希弗曼（Josh Silverman）认为是Temu和Shein推高了数字广告的价格。

大笔支出换来了一定的品牌认知度。根据研究公司Business of Apps的数据，Temu是美国、英国、法国和德国去年下载量最高的应用。Shein已经成为备受美国青少年青睐的购物应用。但这两家公司能否赢得广大美国消费者的青睐呢？

Shein主要销售服装，Temu的商品种类更多，从儿童玩具到工业工具等。但两家公司的商业模式相似。主要优势都是价格。研究公司MoffettNathanson的迈克尔·莫顿（Michael Morton）估计，同类女装在美国购物网站的价格是Temu的两到四倍。Temu上出售每件约1美元的电动牙刷、太阳镜和背包。

两家公司靠很多方法来保持低价。其一是省去中间商，直接从中国工厂进货。另一个是对商家收取的佣金比美国竞争对手收取的少。它们也没有像亚马逊那样在美国建立庞大的物流业务，而是选择从中国的仓库发货，通过UPS、联邦快递或邮政将产品送到美国消费者手中。因此顾客能享受到低价，但无法快速收货。等待包裹运抵可能需要几周。

目前而言，Temu专注于市场份额而非利润。经纪公司盛博（Bernstein）估计，Temu在美国每销售一件商品约亏损10美元。有拼多多在中国的成功提供资金支持，这种状况或许还可以维持一段时间；去年，这家母公司的经营现金流达120亿美元。相比之下，Shein正在寻求筹集更多资金。去年11月，它在美国申请上市。如果成功，上市规模将非常庞大。在2022年5月的最新一轮融资中，该公司的估值为660亿美元。

到目前为止，Temu和Shein在美国电子商务市场上的份额都还很小。根据盛博的数据，两家公司都仅有约1%的份额。亚马逊则占了38%。即便如此，这个本地巨头并没有不把它们当回事。去年12月，可能是出于一种战略应对，亚马逊宣布将减少对价格低于15美元的服装收取的佣金。9月，它推出了一项端到端的供应链服务，模仿两家中国公司的做法，从工厂提货直接运送给顾客。

但亚马逊不太可能是首当其冲受冲击的，也不大会是受冲击最大的。研究公司eMarketer的数据显示，服装和配饰仅占亚马逊销售额的16%，因此Shein对Forever 21等快时尚品牌的威胁可能更大。至于Temu，缺乏本地物流网络将使它难以在洗碗片等顾客需要快速送达的商品上与亚马逊竞争。它对eBay或Etsy的威胁可能更大。

局面最终可能会改变。两家公司都在美国购买了仓库空间，并与当地物流公司达成了合作关系。据报道，Shein正在从亚马逊的供应链专家队伍里挖人。Temu也可能开始销售智能手机等价格更高的商品，其母公司在中國就成功实现了这样的转变。这将让它与亚马逊及同样强大的沃尔玛展开更直接的竞争。

｜ 低价做起

两家公司仍然面临诸多风险。在与美国老牌企业竞争的同时，它们相互之间也要竞争。社交媒体公司TikTok去年9月在美国推出的在线市场TikTok Shop也可能来掺一脚。地缘因素也可能对它们造成损伤。它们被指控牵涉强迫劳动（两家公司均予以否认），美国一些参议员组成的委员会正在展开相关调查。另一家银行摩根士丹利的分析表明，美国消费者在Temu购物的意愿可能已经下降，也许是因为新鲜劲过去了。

尽管如此，盛博的马克·舒米利克（Mark Shmulik）认为，对于新的电子商务公司来说，实现市场份额从零到1%的增长要比从1%到5%更难。要让消费者首先能注意到自己不容易。一旦他们熟悉了一个品牌，就更容易向他们卖出更多产品了——也能卖出更贵的。Temu和Shein可能已经跨过了在美国立足的最艰难阶段。■



Pen v plague

From Napoleon to Vladimir Putin, disease has shaped history

A sacked New York Times reporter takes a global view of disease

The Wisdom of Plagues. By Donald McNeill. Simon & Schuster; 384 pages; \$28.99 and £20

EARLY ONE morning a New York Times reporter was shaken roughly awake. He was sleeping in a tiny hut, deep in the Cameroonian rainforest. Some local criminals were coming to kidnap him, his guide warned. Both men fled, just in time.

The reporter was Donald McNeill, who started at the New York Times as a copy boy in 1976 and later spent decades covering global health, especially the diseases that afflict the poor. He did so with courage, compassion and an eye for unexpectedly important details. The story he was pursuing in Cameroon was about hunting great apes and how their extinction might harm understanding of diseases that have crossed from other primates to humans. As a scientist told him, killing chimpanzees “is like burning a library full of books you haven’t read yet”.

Mr McNeill’s reportorial adventures prepared him for the biggest story of his career: he and colleagues won a Pulitzer prize for their coverage of covid-19. Times readers no longer have the benefit of his insights, however. He was pushed out in 2021 because of complaints that he had uttered a racial slur when referring to a conversation about someone else using it. On the plus side, Mr McNeill then had time to write a cracking book on pandemics.

“The Wisdom of Plagues” puts the struggle against pestilence in historical context, noting how the shift from hunter-gathering to farming made it

easier for bugs to jump from livestock to people. Mr McNeill explains how plagues have shaped history, from the typhus that crippled Napoleon's invasion of Russia to the covid-induced isolation that, he speculates, might have aggravated Vladimir Putin's paranoia.

Fear of disease has long led to scapegoating. Syphilis was "the French pox" to the English, "the Turkish disease" to Poles and "the Christian disease" to Turks. Mr McNeill recounts how a chest of gold and silver coins was recently found in the old Jewish quarter of Erfurt, Germany. The last king depicted on the coins had ruled in 1349, the year the Black Death struck the city. Locals blamed the Jews, as they did all over western Europe. The owner of the chest never got to dig it up.

With fear comes falsehood. If it is not ethnic minorities supposedly spreading plagues, it is evil tycoons profiting from them in convoluted ways. A century before Bill Gates was accused of using covid jabs to plant chips in people, John D. Rockefeller funded a campaign to urge Americans to wear shoes to stop hookworm burrowing into their bare feet. Rumours spread that "the oil magnate had secretly bought up shoe companies".

Misinformation is a huge obstacle to fighting disease. Mr McNeill saw close-up some of the estimated 300,000 extra deaths caused by a South African president, Thabo Mbeki, who concluded after surfing the internet that HIV did not cause AIDS. He watched the global campaign to eradicate polio stumble because so many Muslims believe the vaccine is a Western plot to sterilise them. Convincing them otherwise grew harder, Mr McNeill fumes, when the CIA used a fake vaccination campaign to try to trace Osama bin Laden in 2011. In one month the next year, Pakistani jihadists shot dead nine polio vaccinators.

To fight pandemics, accurate information must be rapidly disseminated. That means bureaucracies need to move faster. "I can't think of a single

serious outbreak I learned of first” from the World Health Organisation (WHO), complains Mr McNeill. It is part of the UN, and like “a London gentleman’s club”, defers to its members. ProMED, an unofficial disease-alert network, sounded the alarm about covid-19 four days before the much better-resourced WHO.

Accurate information must also be believed. Too often, it is not, because of a breakdown of public trust in authorities. When experts make mistakes, which is inevitable when grappling with a new disease, many people conclude that the professionals are useless or self-serving.

In this, they are encouraged by reckless pundits and politicians. “The storm of rumours surrounding covid was by far the most intense I ever covered—for an obvious reason: many of them were spread by the president of the United States,” Mr McNeill writes. He deplores the way masks became political badges in America, with Republicans shunning them and Democrats sporting them even when cycling outdoors without helmets. Tribalism makes for poor risk assessment.

Mr McNeill brims with ideas for curbing the next pandemic, including sampling sewage for viruses. He would, controversially, end religious exemptions to public-health rules. He omits some important things, such as the mystery of why Sweden coped reasonably well with covid despite very loose lockdown rules. But overall, this is a fascinating, ferocious fusillade against humanity’s two deadliest enemies: disease and itself. ■



挥笔写瘟疫

从拿破仑到普京，疾病塑造了历史

一位被《纽约时报》解雇的记者用全球视野看疾病

《瘟疫的学问》，唐纳德·麦克尼尔著。西蒙与舒斯特出版社；384页；28.99美元，20英镑。

一天清晨，《纽约时报》的一名记者被猛地摇醒。他正睡在喀麦隆雨林深处的一座小屋里。当地一些罪犯来绑架他了，他的向导警告说。两人差一点就来不及逃脱了。

这名记者是唐纳德·麦克尼尔（Donald McNeill）。他在1976年进入《纽约时报》，起先递送稿件和跑腿，后来几十年里一直报道全球健康议题，尤其是那些困扰穷人的疾病。他怀着勇气和同情心工作，而且善于发现意料之外的重要细节。他在喀麦隆追踪报道了对类人猿的猎杀，以及它们的灭绝如何可能影响人们了解那些从其他灵长类动物传播给人类的疾病。正如一位科学家告诉他的，杀死黑猩猩“就如同烧掉一座满是你还没读过的书籍的图书馆”。

麦克尼尔惊险的记者经历为他职业生涯中最重大的报道做了铺垫：他和同事们因报道新冠肺炎而获得了普利策奖。然而，《纽约时报》的读者不会再读到他富洞察力的报道了。2021年，有人投诉他在提及某个人使用一个种族歧视字眼的对话时自己也说了这个词，他被迫离开。好的一面是，这让他接下来有时间撰写一本关于大流行病的精彩书籍。

《瘟疫的学问》（The Wisdom of Plagues）在历史大背景下探讨人类对抗疫病的斗争，指出从狩猎采集到农耕的转变如何让病菌更容易从牲畜传播到人。麦克尼尔阐释了瘟疫如何塑造历史，从斑疹伤寒让拿破仑对俄国的入侵受挫，到他推测新冠隔离可能加剧了普京的偏执妄想。

长久以来，对疾病的恐惧促使人们迁怒于人。梅毒在英国人眼中是“法国

疮”，在波兰人眼中是“土耳其病”，在土耳其人眼中是“基督教病”。麦克尼尔记述道，多年前在德国埃尔福特市（Erfurt）的老犹太社区发现了一箱金币和银币。硬币上描画的最后一位国王1349年时在位，而在那一年黑死病袭击了这座城市。当地人像全西欧的人那样把这归咎于犹太人。箱子的主人再没有机会把它挖出来。

随恐惧而来的是谣言。要么是说少数族裔传播了瘟疫，要么就是邪恶的大亨通过各种迂回曲折从瘟疫中获利。比尔·盖茨被指利用新冠疫苗给人体注射芯片。而在一个世纪前，约翰·洛克菲勒资助了一场运动，敦促美国人穿鞋以阻止钩虫钻进他们赤裸的双脚。谣言四处流传，称“这位石油大亨秘密收购了鞋业公司”。

不实信息是对抗疾病的巨大阻碍。南非总统塔博·姆贝基在上网浏览信息后得出HIV不会导致艾滋病的结论，造成了大约30万人的额外死亡，麦克尼尔见证了其中一些悲剧。他目睹了全球根除小儿麻痹症的运动难以推进，因为许多穆斯林相信疫苗是西方的阴谋，目的是给他们绝育。麦克尼尔愤怒地指出，当美国中情局在2011年试图通过虚假的疫苗接种活动追踪本·拉登时，说服他们的难度就更大了。第二年的一个月里，巴基斯坦圣战分子射杀了九名小儿麻痹病疫苗接种人员。

要对抗大流行病，准确的信息必须能迅速传播出去。这就需要官僚机构加快行动。麦克尼尔抱怨说：“我想不出有哪一次严重疫情我是从世卫组织那里首先得知的。”世卫组织是联合国下属机构，就和“一家伦敦绅士俱乐部”一样，唯其成员马首是瞻。非官方的疾病预警网络ProMED比资源充足得多的世卫组织早四天发出了有关新冠肺炎的警报。

准确的信息还必须能让人相信。太多时候人们都不信，因为公众对权威机构的信任崩塌了。当专家犯错时——这在试图弄明白一种新疾病时是不可避免的——许多人判定这些专业人士毫无用处或自私自利。

在这一点上，他们受到了轻率鲁莽的“名嘴”和政客的鼓动。“围绕新冠病毒的谣言风暴是我所报道过的最激烈的，原因显而易见：其中许多是由美国

总统传播的。”麦克尼尔写道。他谴责口罩在美国如何成为了政治标志——共和党人避而不戴，而民主党人在户外骑车时就算不戴头盔也要戴着口罩。部落主义导致了糟糕的风险评估。

至于如何遏制下一次大流行病，麦克尼尔有很多点子，包括对污水做病毒采样。必然引发争议的是，他主张结束对公共卫生规则的宗教豁免。他漏掉了一些重要的事情，比如为什么尽管瑞典的封锁规则非常宽松，却令人费解地相当好地应对了新冠。但总的来说，这本书引人入胜而猛烈地炮轰了人类的两个最要命的敌人：疾病和人类自己。■



The AI-chip race

Could AMD break Nvidia's chokehold on chips?

Taking on the top AI chipmaker will be hard—but maybe not impossible

“IT IS THE most advanced AI accelerator in the industry,” boasted Lisa Su, boss of Advanced Micro Devices (AMD), at the launch in December of its new MI300 chip. Ms Su rattled off a series of technical specifications: 153bn transistors, 192 gigabytes of memory and 5.3 terabytes per second of memory bandwidth. That is, respectively, about 2, 2.4 and 1.6 times more than the H100, the top-of-the-line artificial-intelligence chip made by Nvidia. That rival chipmaker's prowess in the semiconductors fuelling the AI boom has, over the past year, turned it into America's fifth-most-valuable company, with a market capitalisation of \$1.5trn. Yet most experts agreed that the numbers and Ms Su weren't lying: the MI300 does indeed outshine the H100. Investors liked it, too—AMD's share price jumped by 10% the next day.

On January 30th, in its quarterly earnings call, AMD announced that it expected to sell \$3.5bn-worth of MI300s this year. It also reported strong revenues of \$23bn in 2023, four times what they had been in 2014, when Ms Su became chief executive. Its market value is up 100-fold on her watch, to \$270bn. Relative to forecast profits in the next 12 months, its valuation is richer even than Nvidia's. Last year it displaced Intel, which once ruled American chipmaking, as the country's second-most-valuable semiconductor company. Now it is taking aim at the biggest.

Such ambition would have seemed fanciful a decade ago. Back then, recalls Mark Papermaster, AMD's technology chief, AMD was facing an “existential crisis”. In 2008 it had spun off its chip-fabrication business to focus on designing processors, outsourcing manufacturing to contract chipmakers such as TSMC of Taiwan. The idea was to be better able to compete on

blueprints with Intel, whose vast fabrication capacity AMD could not hope to match.

It didn't work. Several of AMD's chips flopped. Sales of its central processing units (CPUs), mostly for personal computers, were collapsing. In 2013 it sold and leased back its campus in Austin to raise cash. A year later Ms Su inherited a net-debt pile of more than \$1bn, a net annual loss of \$400m and a market value of less than \$3bn, down from \$20bn in 2006.

She realised that the only way for AMD to get back in the game was to steer it away from the sluggish PC market and focus on more promising areas like CPUs for data-centre servers and graphics processing units (GPUs, which make video-game visuals lifelike) for gaming consoles. She and Mr Papermaster took a gamble on a new CPU architecture designed to beat Intel not just on price, but also on performance.

| *When the going got tough*

The idea was to use a Lego-like approach to chip building. By breaking a chip up into smaller parts, AMD could mix and match blocks to assemble different types of chip, at a lower cost. When the first such composite chips were released in 2017, they were zippier and cheaper than rival offerings from Intel, possibly in part because Intel was distracted by its own problems (notably repeated manufacturing slip-ups as it moved to ever tinier transistors). In the past ten years AMD's market share in lucrative server CPUs has gone from nothing to 30%, breaking Intel's monopoly.

Having faced down one giant, AMD now confronts another. The contest with Nvidia is different. For one thing, it is personal—Ms Su and Jensen Huang, Nvidia's Taiwanese-born boss, are distant relatives. In contrast to Intel, Nvidia is, like AMD, a chip designer and thus less prone to production missteps. More importantly, the stakes are higher. Nvidia's market value of \$1.5trn is predicated on its dominance of the market for GPUs—not because

of their usefulness in gaming but because they also happen to be the best type of chip to train AI models. Ms Su expects global sales of AI chips to reach \$400bn by 2027, up from perhaps \$40bn last year. Does she stand a chance against Nvidia?

Nvidia is a formidable rival. Both its revenues and operating margins are nearly three times AMD's. According to Jefferies, an investment bank, the company dominates the market for AI accelerator chips, accounting for 86% of such components sold globally; before the launch of the MI300, AMD barely registered. Nvidia also offers network gear that connects clusters of chips, and software, known as CUDA, to manage AI workloads. Nvidia has dominated AI chipmaking because it has offered the best chips, the best networking kit and the best software, notes Doug O'Laughlin of Fabricated Knowledge, a research firm.

AMD's new processor shows it can compete with Nvidia on semiconductor hardware. This, Mr Papermaster says, is the result of a ten-year investment. AMD is spending nearly \$6bn a year on research and development, nearly as much as its larger rival—and twice as much as a share of sales (see table). This has enabled it to adapt its Lego approach to GPUs. Combining a dozen blocks—or “chipselets”—into a single chip lets AMD put processors and memory close to each other, which boosts processing speed. In December OpenAI, maker of ChatGPT and the world's hottest AI startup, said it would use the MI300s for some of its training.

To outdo Nvidia on networking and software, AMD is teaming up with other firms. In December it announced a partnership with makers of networking gear, including the two largest, Broadcom and Cisco. It is also supporting an open-source initiative for chip-to-chip communication called Ultra Ethernet Consortium as an alternative to InfiniBand, a rival championed by Nvidia.

Nvidia's lead in software will be harder to close. It has been investing in CUDA since the mid-2000s, well before the current AI wave. AI developers and researchers love the platform, which allows them to fine-tune the performance of Nvidia processors. AMD hopes to tempt customers away from Nvidia by making its software, ROCm, open-source and providing tools to make the switch smoother, by translating CUDA programs into ROCm ones.

Beating Nvidia at its own game will not be easy. Mr Huang's firm is not standing still. It recently announced plans to bring out a new chip every year instead of every two years. The tech giants with the grandest AI ambitions—Alphabet, Amazon, Meta and Microsoft—are busily designing their own accelerator chips. Despite AMD's robust sales, investors were disappointed with its forecast for MI300 shipments. Its share price dipped by 3% the day after it reported its latest results.

Still, AMD has one big thing going for it. It is not Nvidia. AI companies are desperate for an alternative to its larger rival, whose dominant position allows it to charge steep prices and, with demand outstripping supply, ration chips to buyers. Despite efforts to design their own hardware, big tech firms will rely on chipmakers for a while, and AMD gives them options, notes Vivek Arya of Bank of America. Microsoft and Meta have already announced plans to use AMD's GPUs in their data centres. And if Nvidia slips up, AMD will be there to pick up the Lego pieces. Just ask Intel.





AI芯片竞赛

AMD能否打破英伟达的芯片垄断？

挑战AI芯片老大绝非易事，但未必不可能

“这是目前业界最先进的AI加速器。”去年12月，AMD的老板苏姿丰在MI300芯片的新品发布会上夸耀道。苏姿丰一气说出了这款芯片的一系列技术参数：1530亿个晶体管、192GB的内存、每秒5.3TB的内存带宽。这分别是其竞争对手英伟达制造的顶级AI芯片H100的大约2倍、2.4倍和1.6倍。英伟达凭借其在半导体领域的实力推动了AI的繁荣，并在过去一年时间里以1.5万亿美元的市值晋身美国第五大公司。不过，大多数专家都认为AMD的这些参数是真实的，苏姿丰也没有夸大其词：MI300确实比H100更胜一筹。投资者也乐见其成——AMD的股价在第二天大涨了10%。

1月30日，AMD在季度财报电话会议上宣布，预计今年MI300的销售额会达到35亿美元。它还报称2023年营收表现强劲，达到230亿美元，是2014年苏姿丰出任CEO时的四倍。在她的执掌下，AMD的市值增长了100倍，达到2700亿美元。相对于未来12个月的预期利润而言，AMD的估值甚至超过了英伟达。去年，它取代了曾是美国芯片制造业霸主的英特尔，成为美国市值第二高的半导体公司。现在它的目标是成为美国第一。

这样的雄心壮志要放在十年前会是异想天开。AMD的技术主管马克·佩珀马斯特（Mark Papermaster）回忆道，当时的AMD正面临着一场“生存危机”。2008年，为全力投入处理器设计，AMD剥离了芯片制造业务，将制造外包给台积电等芯片代工厂。AMD希望能借此在芯片的设计上更好地与英特尔竞争，毕竟英特尔强大的制造能力是自己望尘莫及的。

然而事与愿违。AMD的几款芯片都以失败告终。其主要用于个人电脑的中央处理器（CPU）的销售一落千丈。2013年，为筹集资金，AMD出售了位于得州奥斯汀的办公园区，然后再回租。一年后，苏姿丰接手AMD，当时它的净债务超过十亿美元，年净亏损四亿美元，市值已从2006年的200亿美元下跌到不足30亿美元。

苏姿丰意识到，要想打翻身仗，AMD只有从低迷的个人电脑市场转移到更有前途的领域，比如面向数据中心服务器的CPU或者游戏机的图形处理器（GPU，可以生成逼真的游戏画面）。她和佩珀马斯特把赌注押在新的CPU架构上——不仅要在价格上、还要在性能上战胜英特尔。

丨 道阻且长

他们的想法是像搭乐高那样搭建芯片。通过将芯片分解成更小的单元，AMD可以把不同的模块混搭在一起，以更低的成本组装不同类型的芯片。2017年AMD推出了第一批这样的组合芯片，它们比英特尔的竞品速度更快、成本更低——一定程度上可能也是因为英特尔被自己的问题分散了精力（尤其是在转向更小的晶体管时一再出现制造上的失误）。过去十年里，在利润丰厚的服务器CPU市场上，AMD的份额从零增长到30%，打破了英特尔的垄断。

在击败了一个巨头之后，AMD现在面对另一个巨头。与英伟达的这场竞争有所不同。首先，这涉及私人关系——苏姿丰和出生于台湾的英伟达老板黄仁勋是远亲。其次，与英特尔不同，英伟达和AMD一样也只是芯片设计公司，因此不太容易在制造上出错。更重要的是，这次竞争的押注更大。英伟达1.5万亿美元的市值是基于它在GPU市场的主导地位，而这并不是因为GPU在游戏中的用处，而是因为这种芯片恰好还最适合用来训练AI模型。苏姿丰预计，到2027年，AI芯片的全球销售额将达到4000亿美元，远高于去年的约400亿美元。她有机会战胜英伟达吗？

英伟达是一个强大的对手。它的收入和营业利润率都是AMD的近三倍。投资银行杰富瑞（Jefferies）称，英伟达主导着AI加速芯片的市场，占全球此类组件销量的86%；在MI300推出之前，AMD在这个市场几乎没有份额。英伟达还提供连接多芯片集群的网络设备，以及用来管理AI工作负载的CUDA软件。研究公司Fabricated Knowledge的道格·奥劳克林（Doug O’laughlin）指出，英伟达之所以能在AI芯片制造领域占据主导，是因为它提供了最好的芯片、最好的网络设备以及最好的软件。

AMD新的处理器表明，它有能力在半导体硬件上与英伟达一较高下。这是

十年投资的结果，佩珀马斯特说。AMD每年的研发支出接近60亿美元，与比它体格更大的对手不相上下；这部分支出占销售额的比例是英伟达的两倍（见表）。这让它能够将它的乐高搭建法应用到GPU上。由于能将12个模块（或者说“小芯片”）整合到单块芯片上，AMD可以将多个处理器和内存紧密地放置在一起，从而提高处理速度。去年12月，开发了ChatGPT的全球最热门的AI创业公司OpenAI表示，将使用MI300进行部分模型训练。

为了在联网和软件方面超越英伟达，AMD正在与其他公司展开合作。去年12月，它宣布与多家网络设备制造商建立合作，其中包括博通和思科这两个巨无霸。它还支持一个叫作“超以太网联盟”（Ultra Ethernet Consortium）的组织，这是个芯片间通信的开源项目，可以取代英伟达旗下的竞争产品InfiniBand。

| 跃跃欲试

要在软件上赶上英伟达会更难。自2000年代中期以来，英伟达就一直在投资CUDA平台，彼时当前的AI浪潮还远远没有到来。该平台让AI开发人员和研究人员能够微调英伟达处理器的性能，深受他们欢迎。AMD希望通过将自己的ROCm软件开源、并且提供更便捷的切换工具（把CUDA程序转译为ROCm程序），将英伟达的客户吸引过来。

要在英伟达的主场击败它绝非易事，黄仁勋的公司并没有按兵不动。它在不久前宣布计划每年推出一款新芯片，而不再是每两年出一款。

Alphabet、亚马逊、Meta和微软等在AI上最雄心勃勃的科技巨头都在忙着设计自己的加速芯片。尽管AMD的销售强劲，但投资者对它的MI300的预测出货量感到失望。在它公布最新业绩的第二天，股价下跌了3%。

尽管如此，AMD仍有一大优势：它不是英伟达。一家独大的英伟达要价过高，而且因其芯片供不应求，英伟达常常对买家实行限量供应，因此AI企业都迫切希望在英伟达之外还能另有选择。美国银行（Bank of America）的维韦克·阿雅（Vivek Arya）指出，尽管科技巨头都在下大力气设计自己的硬件，但在一段时间内，它们仍会依赖芯片制造商，而AMD让它们多

了一个选择。微软和Meta已经宣布，计划在自己的数据中心使用AMD的GPU。如果英伟达出了差错，它们可以用AMD的乐高来救场。看看英特尔就知道了。 ■



Chaguan

Hard times for China's micro-industrialists

A rural hub for children's bicycle-making adjusts to a world with fewer kids

THERE ARE lots of upsides to making bikes for kids, explains Mr Li, a young entrepreneur from Pingxiang, a scruffy county in northern China that has become a centre for the children's bicycle industry. For one thing, they are easy to build, he says, nodding at a toddler-sized machine parked near his desk, held upright by tiny stabilisers. Teenage mountain bikes are a bit fiddly, but smaller ones "need no special machinery at all". Also, he grins, children grow. Sell a three-year-old their first ride and two years later their parents have to buy a bigger one, and so it goes on for years to come. The downside? China is running out of children.

Pingxiang, in the province of Hebei some 400km south of Beijing, is a revealing place to see the country's demographic future playing out today. Like many industrial clusters in China, it grew over decades as businesspeople forged networks, helped by local officials offering tax breaks and other subsidies. Initially, small firms assembled frames, pedals and other parts bought from established manufacturers in coastal cities. Over time complete supply chains were created in Pingxiang. Today, the county is a sprawl of large industrial plants linked to smaller suppliers, many of them tucked away in rural sheds and barns. There are traffic jams as lorries and three-wheelers piled high with bicycle cartons inch down narrow village lanes.

County officials report that 10m bicycles a year are built there, by thousands of firms. Official media credit Pingxiang with supplying 40% of the children's bicycles sold worldwide. It also produces half the wheeled toys sold inside China, including bicycles, tricycles, scooters and ride-along toy cars. This targeted approach to globalisation made Pingxiang

prosperous, if not lovely. The county is a drab, dusty spot, though officials have painted cycling-themed murals on walls and erected a giant sculpture of a bicycle wheel in a public square. Then came China's fertility crash.

In 2023 the number of Chinese newborns hit a record low of just over 9m, after falling for seven years in a row. That compares with nearly 19m babies born in 2016. Mr Li quit his job as a quality controller in a big factory a few years ago, and began selling hand-assembled bikes in street markets. Today he rents a small factory with a former colleague. He employs 15 people who can make hundreds of bicycles a day, when orders are good. Orders are not good right now and his staff are on short hours. Huddled in a freezing office in one corner of their steel-walled factory, he and his business partner look anxious. This is understandable, given that their main customer base—namely, Chinese toddlers—has shrunk by half over the past decade. The change “started slowly, a few years ago”, he recalls. Now the impact on sales is unmistakable.

Mr Li's plight reveals a lot about the large forces that are battering China's private sector, and about the solutions being proposed by officials in both the local and central governments. Coverage of China's economy often focuses on a handful of national champions making world-class products, from smartphones to electric vehicles. But small firms with fewer than 300 employees accounted for 79% of China's job creation and 68% of exports, the OECD, a club of mostly rich countries, reported in 2022. Though the Communist Party puts great stock in large state-owned enterprises and groundbreaking technology, China needs its backyard entrepreneurs, too. The supreme leader, Xi Jinping, calls China's mastery of the complete array of industrial sectors a source of national strength. Last year he urged officials to upgrade, not eliminate, industries deemed “low-end”.

Sluggish domestic demand, notably since the end of the covid-19 pandemic, has led officials to urge manufacturers of all types to seek new

markets abroad. Still, exports are not a cure-all. A bicycle-industry veteran in Beijing notes that China's manufacturers saw roaring domestic and foreign sales during the pandemic, as people abandoned public transport for their own two wheels. The industry now faces a hangover, as inventories are cleared and many covid-era riders lose interest. In ageing societies, e-bikes for adults are selling well. But even in markets that still have children, many want to play video games, not play outdoors. Chinese consumers lack the confidence to spend but demand is weak in many foreign markets too, says the veteran. "In some of our enterprises, production is down by a third."

| *A slowing China bets on exports*

Back in Pingxiang, Mr Li is struggling to survive. During pandemic lockdowns, many Chinese consumers had bills to pay but no income. They have not yet shaken the fear that they felt then, he says. Compared with last year, his firm's sales are down by more than half. Local officials urge businesses like his to look abroad, with a focus on countries signed up to China's Belt and Road Initiative. In 2020 less than a tenth of the firm's sales went abroad. Now exports account for 40-50% of turnover, with customers in Russia, Malaysia and Indonesia. He is grateful to county officials who subsidised his stand at a trade fair in Shanghai, where he met foreign buyers. But exports are hard work. The Russians are from that country's far east near the Chinese border, he thinks. They ask for more time to pay when the rouble is weak, though helpfully they settle their bills in Chinese yuan. Europe and America are richer markets, but the firm cannot meet their product standards.

After lunch, Chaguan is taken to a nearby village to see Plan B. In a farmyard workshop guarded by honking geese, Mr Li's business partner has a team assembling pedal tricycles for old people. These are built to order in batches of 50, which middlemen sell to domestic customers online.

Compared with children, who need new bikes as they grow, the disadvantage is that pensioners “stop riding” when they get older, says the partner earnestly. But at least China will have more and more of them. A larger tricycle-assembly line is planned for next year. These are grim times for China’s micro-industrialists. Their resilience is a wonder to behold. ■



茶馆

中国小微实业家的艰难时日

一个位于农村的童车制造中心正在调整，以适应一个孩子越来越少的世界

中国北方有一个乱糟糟的县城平乡，如今已成为儿童自行车产业的中心。当地的年轻企业家李先生解释说，给小孩子造自行车有很多有利之处。首先，童车容易造，他朝桌子旁停着的一台家伙努了努下巴，这辆幼儿尺寸的车有小小的平衡轮作支撑。制造青少年骑的山地车需要点技术，但更小的童车“根本不需要特殊的机械”。还有，他咧嘴一笑道，孩子总归要长大。三岁时买了第一辆车，两年后父母就要再买辆更大的，就这样持续好些年。不利的一面呢？中国的孩子越来越少了。

隶属河北省的平乡位于北京以南约400公里，从这里可以窥见中国的人口发展趋势。和中国的许多产业集群一样，在地方官员提供税收减免和其他补贴的帮助下，商人们打造的这个产业网络在几十年里不断发展壮大。起初，小公司组装从沿海城市的老牌制造商那里购买的车架、踏板和其他部件。渐渐地，完整的供应链在平乡建立起来。如今，这里到处都是串联着较小供应商的大型工厂，其中许多供应商栖身在农村的棚屋和谷仓里。满载着装有自行车的纸箱的卡车和三轮车在狭窄的乡间小路上缓慢挪动，造成拥堵。

县里的官员报告说，数千家公司每年在这里生产一千万辆自行车。官方媒体称，平乡制造的儿童自行车占全球销量的40%。在中国国内销售的各式玩具车中，有一半是在平乡生产的，包括自行车、三轮车、滑板车和乘坐式玩具车。这种目的明确的全球化让平乡繁荣起来，虽然不够美好有趣。这个县城单调乏味、尘土飞扬，尽管官员们在墙上画了以自行车为主题的壁画，还在公共广场上竖起了一个巨大的自行车车轮造型的雕塑。后来就碰上了中国生育率暴跌。

在连续下跌七年之后，2023年中国新生儿数量创下了历史新低，仅有900多万。相比之下，2016年有近1900万婴儿出生。几年前，李先生辞去了在

一家大工厂担任质控员的工作，开始在街市上销售手工组装的自行车。如今，他和一位前同事租下了一间小厂房。他雇了15名员工，订单多的时候，他们每天可以生产数百辆自行车。现在订单少，他的员工早早就下班了。他和生意伙伴两人蜷在钢结构厂房一角的一间冰冷的办公室里，神色焦虑。这不难理解，因为他们的主要顾客群，也就是中国的幼儿，在过去十年里少了一半。他回忆说，这种变化“在几年前就慢慢开始了”。现在这对销售的影响已经确凿无疑。

李先生的困境充分揭示了正在重创中国私营部门的那些强大力量，以及地方和中央官员提出的解决方案的影响。对中国经济的报道往往聚焦于少数几家制造世界级产品的本国领军企业，从智能手机到电动汽车等。但是，成员主要为富裕国家的经合组织（OECD）在2022年的报告中称，雇员少于300人的小企业创造了中国79%的就业，68%的出口。虽然共产党非常重视大型国有企业和突破性技术，但中国也需要自己的后院型企业。最高领导人习近平称中国门类齐全的产业体系是国家实力的源泉。去年，他敦促官员们升级而不是淘汰被视为“低端”的行业。

国内需求低迷，新冠疫情结束以来尤其如此，导致官员们敦促各行各业的制造商寻求海外新市场。不过，出口并不能包治百病。北京一位自行车产业资深人士指出，疫情期间人们放弃了公共交通工具，骑上了自己的自行车，中国制造商由此迎来了国内外销量飙升。随着库存被清空，加上许多疫情时期骑车的人又失去了兴趣，眼下整个行业正在体验宿醉。在老龄化社会，成人的电动自行车卖得很好。但即便在仍有大量儿童的市场，许多孩子还是更想玩电子游戏，而不是在户外玩耍。这位资深人士表示，中国消费者缺乏消费信心，而许多外国市场的需求也很疲软。“在我们的一些企业，产量下降了三分之一。”

｜经济放缓的中国押注出口

回看平乡，李先生正在为生存而挣扎。在疫情封锁期间，许多中国消费者入不敷出。他说，他们如今仍心有余悸。与去年相比，他公司的销售额下降了一半不止。当地官员鼓励像他这样的企业将目光投向海外，重点关注中国“一带一路”倡议的共建国。2020年时他的公司只有不到一成销售额来

自海外，而现在出口占营业额的40%至50%，客户来自俄罗斯、马来西亚和印度尼西亚。他很感谢县里的官员资助他参加上海的一个展销会，在那里他遇到了国外的买家。但出口不好做。他觉得俄罗斯客户来自俄罗斯靠近中国边境的远东地区。当卢布疲软时，他们要求延长付款时间，尽管他们用人民币结算这点挺好。欧洲和美国市场更富裕，但他的公司无法达到那里的产品标准。

午饭后，笔者被带到附近的一个村子去看备选方案。在一个由嘎嘎叫的鹅群守护着的农家院子里有一个车间，李先生的商业伙伴有个团队在这里为老年人组装脚踏三轮车。这些产品按50辆一批订购，由中间商在网上卖给国内客户。相比小孩子一路长大要换新车，不利之处是老人年纪再往上去就“不骑车了”，这位合伙人一脸诚恳地说道。但至少中国会有越来越多老人。一条更大的三轮车装配线计划于明年建成。对于中国的小微实业家来说，这是一个艰难的时期。他们的韧性令人惊叹。■



Bartleby

Companies run to their own annual rhythms

Seasonality in firms, from budgeting cycles to bonus rounds

SEASONALITY IS A big part of business. For some industries, seasonal patterns are a defining feature. Agriculture is one obvious example; tourism another. Western toymakers notch up a huge proportion of their annual sales in the run-up to Christmas. Construction is harder during cold weather, which is why that industry employs fewer people in the winter.

Firms that are less obviously tied to the seasons can still be deeply affected by them, as a recent review by Ian Hohm of the University of British Columbia and his co-authors makes clear. An analysis of social-media posts on Twitter, now X, found that dieting-related tweets peak in the spring, as the season of body dysmorphia (ie, summer) approaches. Condom sales and online searches for pornography in America tend to rise in the summer and around Christmas.

Even when overall demand does not vary greatly between the seasons, preferences change. Beef-eaters buy diced meat and roasts in the slow-cooking winter season and plump for steaks during the summer grilling months. Starbucks is among those firms that make seasonality a marketing event. The pumpkin-spiced latte is a reliable sign that autumn is on its way, along with falling leaves and glum faces at condom manufacturers.

Seasonality also leaves a less obvious imprint inside organisations. Just as there are daily and weekly patterns of activity, from slumps in concentration during the late afternoons to the ebb and flow of hybrid workers coming to the office, so annual cycles leave their mark.

One is occurring this week, with the World Economic Forum's annual

shindig in Davos. Public holidays aside, in no other week in the working year are so many CEOs of large organisations reliably away. The corporate world is briefly without a government, a concentrated version of Belgium in the early 2010s. This may well be Davos's real contribution to improving the state of the world: with so many bosses stuck on a mountain for a few days, productive employees can get on with some work and lazy ones can relax.

School holidays offer an obvious form of seasonality, although in that case people throughout the organisation are off. Mass absences make it hard to schedule meetings in Brazil in the period between Christmas and the start of Carnival; it is a similar story in August in Europe.

These patterns of clustered absences show up inside organisations in big ways and small. Second-fiddle employees are more likely to get their chance to run the show; fewer big initiatives are likely to be launched when the holidays are in full swing. Employees without children are resentful that they are covering for colleagues on holiday; colleagues on holiday are resentful that they have children.

There is some evidence that people feel more creative after returning from holiday—but you need to schedule that brainstorming session quickly. A paper from 2010 by Jana Kühnel of Goethe University and Sabine Sonnentag of Universität Mannheim reckoned that the benefits of a break fade within a month.

Set-piece events mark the corporate calendar, too. Some are public: annual general meetings and shareholder letters, say. Others are internal. At many companies the annual budgeting process involves a gathering organisational effort, in which more and more people spend more and more time arguing about numbers that are certain to be wrong. It is almost a season in itself. A pre-pandemic estimate from APQC, a benchmarking

organisation, reckoned that the median firm spends around 30 days on this effort; at plenty of firms, it takes an awful lot longer.

Pay decisions are seasonal events, too. The time when employees find out their salary rises and bonuses sets off ripples of disappointment and happiness in all workplaces. In some, they are more like tsunamis. The bonus round on Wall Street, when bankers find out what they will get for their work the previous year, is under way now and is predated by months of internal wrangling and gossip. The actual date on which bonuses are paid matters, too—once the money is safely deposited in the bank, people are more likely to move jobs.

There are other forms of corporate seasonality. The office Christmas party signals another wind-down in activity. Some firms shorten the workweek during the summer months. Yearly calendars are punctuated by sales conferences and leadership retreats. There is not much research on the impact of seasonality within firms. That they have their own annual rhythms is indisputable. ■



巴托比

公司有自己的年度律动

企业里的季节性——从预算周期到奖金发放

季节性是商业的重要部分。对于某些行业来说，季节性模式更是一个典型特征。农业就是显而易见的例子，旅游业也是。西方玩具制造商在圣诞节前夕的销售额占其全年销售的很大部分。建筑业在寒冷的天气里较难施工，因此在冬季也会减少用工。

不列颠哥伦比亚大学（University of British Columbia）的伊恩·霍姆（Ian Hohm）及其合著者最近发表的一篇报告认为，看起来与季节没有明显关联的公司也仍可能深受季节的影响。一项对推特（Twitter，现已更名为X）上社交媒体帖子的分析发现，随着体型恐惧高发季（也就是夏季）即将来临，与节食减肥相关的推文在春季达到高峰。美国的避孕套销量和网上色情搜索量在夏季和圣诞前后趋于上升。

即使总需求量在不同季节间波动不大，人们的偏好也会发生变化。爱吃牛肉的人会在冬天的慢炖季购买牛肉粒和烤牛肉，在夏天的烧烤季选购牛排。许多公司会开展季节性营销活动，星巴克便是其中之一。南瓜拿铁一出，就知道秋天要来了，正如落叶和避孕套厂商阴郁的脸色告诉你的那样。

虽然不那么明显，季节性在企业内部也存在。人员活动在每天和每周都有一定的规律，从近傍晚时分人们开始心不在焉，到混合办公的员工去办公室的人流增减。同样，一些年度周期也会留下印记。

其中一个周期刚刚上演，那就是每年在达沃斯举行的盛大的世界经济论坛。除了公共假期以外，在工作时间里没有哪一个星期会有这么多大型机构的CEO肯定不在公司。企业界短暂地进入了无政府状态，就像2010年代初的比利时的浓缩版。这很可能是达沃斯论坛对改善世界状况的真正贡献：这么多老板在一座山上待几天，勤快的员工可以继续干活，懒惰的员工

工则可以放松一下。

学校假期带来了一种明显的季节性，尽管在这种情况下，企业里上上下下的人都在休假。在巴西，从圣诞节到狂欢节开始前的时间里，大规模缺勤让安排会议变得异常困难；8月的欧洲也是类似的情况。

这种成群缺席的模式在企业内部或多或少都会出现。二把手们更有可能得到主持大局的机会；大家都去度假的时候，就不太可能搞出大动作。没有孩子的员工因为要替休假的同事顶班而不爽；休假的同事又因为自己孩子而不爽。

一些证据表明，人们度假回来后会感觉更有创造力——但你需要尽快安排头脑风暴活动。歌德大学（Goethe University）的贾娜·库内尔（Jana Kühnel）和曼海姆大学（Universität Mannheim）的萨宾·索恩塔格（Sabine Sonnentag）在2010年发表的论文中指出，休假的好处在一个月内就会消失。

公司日历上也标注着定期的活动。有些是对外界公开的：比如年度股东大会和股东信。还有一些是内部活动。在许多公司，年度预算编制是一个举全公司之力的过程，越来越多的人花越来越多的时间来争论一些肯定会错掉的数字。这本身几乎就是一个季节。根据基准测试机构APQC在疫情前的估计，公司在预算编制上花费的中位数时间为30天左右；很多公司花的时间还要长得多。

薪酬调整也是季节性事件。随着员工得知自己的加薪和奖金，所有职场里都会掀起失望和喜悦的涟漪。在有些地方，这更像是一场海啸。华尔街的奖金发放正在进行，此时银行家将会得知他们前一年工作的回报，而在此之前，内部争吵和流言蜚语早已持续了数月之久。发放奖金的实际日期也很重要——奖金落袋为安之后，员工就更有可能会跳槽。

公司季节性还有其他形式。办公室圣诞晚会预示着公司活动将再次逐渐减少。一些公司在夏季缩短了每周的工作时间。每年的日历上都穿插着各种销售会议和领导务虚会。关于公司内部季节性影响的研究并不多，但它们

都有自己的年度律动是个不争的事实。■



Online media

The end of the social network

As Facebook turns 20, social apps are being transformed

FACEBOOK MAY be turning 20 on February 4th, but it is just as much of a magnet for controversy and cash today as when it was a brash, break-everything teenager. On January 31st Mark Zuckerberg, the social network's founder, was harangued by American senators over the spread of harmful material. The next day, as we published this, he was poised to announce another set of glittering results for Meta, Facebook's parent company, which is now valued at \$1trn. Yet even as social media reliably draw vast amounts of attention from addicts and critics alike, they are undergoing a profound but little-noticed transformation.

The weird magic of online social networks was to combine personal interactions with mass communication. Now this amalgam is splitting in two again. Status updates from friends have given way to videos from strangers that resemble a hyperactive TV. Public posting is increasingly migrating to closed groups, rather like email. What Mr Zuckerberg calls the digital "town square" is being rebuilt—and posing problems.

This matters, because social media are how people experience the internet. Facebook itself counts more than 3bn users. Social apps take up nearly half of mobile screen time, which in turn consumes more than a quarter of waking hours. They gobble up 40% more time than they did in 2020, as the world has gone online. As well as being fun, social media are the crucible of online debate and a catapult for political campaigns. In a year when half the world heads to the polls, politicians from Donald Trump to Narendra Modi will be busy online.

The striking feature of the new social media is that they are no longer very

social. Inspired by TikTok, apps like Facebook increasingly serve a diet of clips selected by artificial intelligence according to a user's viewing behaviour, not their social connections. Meanwhile, people are posting less. The share of Americans who say they enjoy documenting their life online has fallen from 40% to 28% since 2020. Debate is moving to closed platforms, such as WhatsApp and Telegram.

The lights have gone out in the town square. Social media have always been opaque, since every feed is different. But TikTok, a Chinese-owned video phenomenon, is a black box to researchers. Twitter, rebranded as X, has published some of its code but tightened access to data about which tweets are seen. Private messaging groups are often fully encrypted.

Some of the consequences of this are welcome. Political campaigners say they have to tone down their messages to win over private groups. A provocative post that attracts "likes" in the X bear pit may alienate the school parents' WhatsApp group. Posts on messaging apps are ordered chronologically, not by an engagement-maximising algorithm, reducing the incentive to sensationalise. In particular, closed groups may be better for the mental health of teenagers, who struggled when their private lives were dissected in public.

In the hyperactive half of social media, behaviour-based algorithms will bring you posts from beyond your community. Social networks can still act as "echo chambers" of self-reinforcing material. But a feed that takes content from anywhere at least has the potential to spread the best ideas farthest.

Yet this new world of social-media brings its own problems. Messaging apps are largely unmoderated. For small groups, that is good: platforms should no more police direct messages than phone companies should monitor calls. In dictatorships encrypted chats save lives. But Telegram's

groups of 200,000 are more like unregulated broadcasts than conversations. Politicians in India have used WhatsApp to spread lies that would surely have been removed from an open network like Facebook.

As people move to closed groups, the open networks left behind are less useful because of the decline in public posting. During the covid-19 pandemic, scientists and doctors contributed to an online debate which contained real insight as well as misinformation. Open-source intelligence flowed when Russia invaded Ukraine. Today those conversations are disappearing or moving to closed channels, slowing the spread of ideas. The people still weighing in on the public networks, meanwhile, are disproportionately male and likely to describe themselves as very left- or right-wing: bores, in plain English.

What's more, the open-network algorithms driven by users' behaviour seem primed to spread the spiciest videos. For something to go viral on a social network, people had to choose to share it. Now they endorse it simply by watching, as the algorithm rewards content that attracts the most engagement. Deliberate curation has been replaced by a system that taps straight into the id. Provocateurs like Mr Trump or Nayib Bukele, the favourite in this week's election in El Salvador, stand to benefit, as do misinformation merchants. Platforms say they are better at weeding out fakes. Taylor Swift, the latest high-profile victim of a deepfake, might disagree.

More urgent even than the rise of fake news is a lack of the real sort. Mr Zuckerberg once said he wanted Facebook to be like a personalised newspaper. But since the network's pivot to entertainment, news makes up only 3% of what people see on it. Across social media only 19% of adults share news stories weekly, down from 26% in 2018. Publications like BuzzFeed News, which relied on social distribution, have perished. That is their lookout (and ours). But it is everyone's problem when nearly half of

young people say that, just as the platforms decide news is no longer interesting, social media are their main source of news.

| *WhatsApp and WhatsDown*

Some people argue that social networks' defects can be fixed by better governance, clever coding or a different business model. Such things can help. But the problems raised by the new generation of apps suggest that social media's flaws are also the result of the trade-offs built into human communication. When platforms swing back towards private groups, they inevitably have less oversight. When people escape their echo chambers, they may well face more extreme content. When users embrace harmless entertainment, they see less news. As social networks wither, platform operators and users should devote less time to the old battles and more to grappling with the new. ■



【首文】在线媒体

社交网络的终结

在Facebook年满20之际，社交应用正在转型

虽说Facebook在2月4日这天年满20了，但它如今吸引争议和金钱的能力丝毫不亚于自命不凡、肆无忌惮的少年时期。1月31日，Facebook的创始人扎克伯格因该平台传播有害信息的问题被美国参议员围攻。翌日，也就是本文发表之时，他会宣布Facebook母公司Meta的又一轮辉煌业绩，该公司目前的市值已达1万亿美元。然而就在社交媒体一如既往地吸引成瘾者和批评者的目光的同时，它们也在经历一场深刻却鲜有人注意的变革。

在线社交网络的奇特魔力在于将人际互动与大众传播结合起来。现在这个结合体再次一分为二。陌生人的视频就像一档极度活跃的电视节目，已经取代了朋友们的状态更新。公开发帖越来越多地转移到封闭的群组，就像电子邮件一样。扎克伯格所称的数字“城市广场”正在重建，同时也带来了问题。

这很重要，因为人们是通过社交媒体来体验互联网的。Facebook本身就拥有30多亿用户。社交应用占据了近一半的手机屏幕时间，而人们清醒时有超过四分之一的时间在看手机。整个世界都已经走上网络，与2020年相比，社交应用所吞噬的时间增加了40%。除了有趣之外，社交媒体还是在线讨论的大熔炉和政治竞选的弹射器。今年世界半数人口都将参加选举投票，从特朗普到莫迪等一众政客都会忙着在网上拉票。

新社交媒体的显著特征是它们不再具有强烈的社交性。受TikTok的启发，Facebook等应用也越来越多地由人工智能根据用户的观看行为而非社交关系挑选短视频推给他们。与此同时，人们发帖也越来越少了。自2020年以来，称自己喜欢在网上记录生活的美国人比例从40%降至28%。论辩的场所正在转向封闭平台，例如WhatsApp和Telegram。

城市广场的灯光已经熄灭。社交媒体向来是不透明的，因为每条信息流都

不一样。但对研究人员来说，TikTok这个来自中国的现象级视频应用就是个黑箱。更名为X的推特公布了部分代码，但收紧了关于展示哪些推文的数据的访问权限。私人消息群组常常是完全加密的。

这些变化的某些结果是值得欢迎的。政治竞选人表示，他们必须缓和自己的调门来赢得私人群组的支持。一条挑衅性帖子可以在喧嚣的X上赢得“点赞”，却可能会遭到WhatsApp学校家长群的排斥。即时通讯应用上的发言是按时间顺序排列的，而不是依据要让参与度最大化的算法，这就减少了炒作的动力。特别值得一提的是，封闭群组可能更有利于青少年的心理健康，因为他们极不适应自己的私生活被公开解读。

在社交媒体中超级活跃的那一半区域，基于行为的算法将让你看到自己的交际圈子以外的帖子。社交网络仍然可以成为观点自我强化的“回音室”。但是，能够从任何地方获取内容的信息流至少有可能将最好的想法传播得最远。

然而，这个社交媒体的新世界也带来了自己的问题。即时通讯应用基本上是不受监督的。这对于小群组来说是件好事：正如电话公司不该监督通话一样，平台也不该监督私聊信息。在独裁国家，加密聊天可以拯救生命。但Telegram的20万个群组更像是不受监管的广播，而不是对话。印度政客曾使用WhatsApp传播谎言，而这种谎言在Facebook这样的开放网络上肯定会被移除。

随着人们转向封闭群组，公开发帖减少了，被抛下的开放网络的用处也就不如从前。在新冠疫情期间，科学家和医生参与了在线辩论，其中既有真知灼见，也有不实信息。俄罗斯入侵乌克兰时，公开网络上的情报信息源源不断。如今这些对话要么消失，要么转到封闭渠道，思想传播因而减缓。与此同时，仍然在公共网络上发表意见的人当中男性占比过高，并且多数自称很左或很右：说白了就是话痨。

此外，由用户行为驱动的开放网络算法似乎特别适合传播劲爆的视频。以前一段内容要在社交网络上走红，必须靠人们主动分享。现在，他们只需

要观看就能提供支持，因为算法会奖励那些观众沉浸度最高的内容。以前还需要精心策划，而现在的系统会直接利用用户的身份。特朗普或纳伊布·布克尔（Nayib Bukele，萨尔瓦多大选的获胜者）等煽动者将从中受益，散布谣言的信息贩子也将获利。各平台表示，它们已经在剔除虚假信息上取得了进展。最近成为AI深度造假受害者的名人泰勒·斯威夫特（Taylor Swift）对此恐怕不敢苟同。

与假新闻满天飞相比，更紧迫的问题是真新闻的匮乏。扎克伯格曾表示，他希望Facebook能像一份个性化的报纸。但自从这一社交网络转向娱乐以来，现在新闻只占人们看到内容的3%。在所有社交媒体中，只有19%的成年人每周分享新闻报道，低于2018年的26%。BuzzFeed News等依赖社交传播的新闻媒体已经销声匿迹。那是它们（也是我们）面对的问题。然而，随着平台认为新闻不再是有兴趣的内容，而同时现在近一半的年轻人表示社交媒体是他们的主要新闻来源，这就成为所有人的问题了。

｜ 社交网络的兴与衰

有人认为，社交网络的这些弊端可以通过更好的治理、巧妙的编程或换一种商业模式来修补。这些或许都有一定帮助。但新一代应用所引发的问题表明，社交媒体的缺陷也是人类交流中固有的取舍的结果。当交流从平台转回私人群组时，它们受到的监督不可避免会减少。当人们逃离自己的回音室时，他们很可能会面对更极端的内容。当用户沉溺于无害的娱乐内容时，他们看的新闻就会减少。随着社交网络的衰落，平台运营商和用户不该再纠结于旧的问题，而是应该勇于应对新的问题。■



Hazy figuring

AI could accelerate scientific fraud as well as progress

Hallucinations, deepfakes and simple nonsense: there are plenty of risks

IN A MEETING room at the Royal Society in London, several dozen graduate students were recently tasked with outwitting a large language model (LLM), a type of AI designed to hold useful conversations. LLMs are often programmed with guardrails designed to stop them giving replies deemed harmful: instructions on making Semtex in a bathtub, say, or the confident assertion of “facts” that are not actually true.

The aim of the session, organised by the Royal Society in partnership with Humane Intelligence, an American non-profit, was to break those guardrails. Some results were merely daft: one participant got the chatbot to claim ducks could be used as indicators of air quality (apparently, they readily absorb lead). Another prompted it to claim health authorities back lavender oil for treating long covid. (They do not.) But the most successful efforts were those that prompted the machine to produce the titles, publication dates and host journals of non-existent academic articles. “It’s one of the easiest challenges we’ve set,” said Jutta Williams of Humane Intelligence.

AI has the potential to be a big boon to science. Optimists talk of machines producing readable summaries of complicated areas of research; tirelessly analysing oceans of data to suggest new drugs or exotic materials and even, one day, coming up with hypotheses of their own. But AI comes with downsides, too. It can make it easier for scientists to game the system, or even commit outright fraud. And the models themselves are subject to subtle biases.

Start with the simplest problem: academic misconduct. Some journals

allow researchers to use LLMs to help write papers, provided they say as much. But not everybody is willing to admit to it. Sometimes, the fact that LLMs have been used is obvious. Guillaume Cabanac, a computer scientist at the University of Toulouse, has uncovered dozens of papers that contain phrases such as “regenerate response”—the text of a button in some versions of ChatGPT that commands the program to rewrite its most recent answer, presumably copied into the manuscript by mistake.

The scale of the problem is impossible to know. But indirect measures can shed some light. In 2022, when LLMs were available only to those in the know, the number of research-integrity cases investigated by Taylor and Francis, a big publisher of scientific papers, rose from around 80 in 2021 to about 2,900. Early figures from 2023 suggest the number was on course to double. One possible telltale is odd synonyms: “haze figuring” as another way to say “cloud computing”, for example, or “counterfeit consciousness” instead of “AI”.

Even honest researchers could find themselves dealing with data that has been polluted by AI. Last year Robert West and his students at the Swiss Federal Institute of Technology enlisted remote workers via Mechanical Turk, a website which allows users to list odd jobs, to summarise long stretches of text. In a paper published in June, albeit one that has not yet been peer-reviewed, the team revealed that over a third of all the responses they received had been produced with the help of chatbots.

Dr West’s team was able to compare the responses they received with another set of data that had been generated entirely by humans, leaving them well-placed to detect the deception. Not all scientists who use Mechanical Turk will be so fortunate. Many disciplines, particularly in the social sciences, rely on similar platforms to find respondents willing to answer questionnaires. The quality of their research seems unlikely to improve if many of the responses come from machines rather than real

people. Dr West is now planning to apply similar scrutiny to other crowdsourcing platforms he prefers not to name.

It is not just text that can be doctored. Between 2016 and 2020, Elisabeth Bik, a microbiologist at Stanford University, and an authority on dodgy images in scientific papers, identified dozens of papers containing images that, despite coming from different labs, seemed to have identical features. Over a thousand other papers have since been identified, by Dr Bik and others. Dr Bik's best guess is that the images were produced by AI, and created deliberately to support a paper's conclusions.

For now, there is no way to reliably identify machine-generated content, whether it is images or words. In a paper published last year Rahul Kumar, a researcher at Brock University, in Canada, found that academics could correctly spot only around a quarter of computer-generated text. AI firms have tried embedding "watermarks", but these have proved easy to spoof. "We might now be at the phase where we no longer can distinguish real from fake photos," says Dr Bik.

Producing dodgy papers is not the only problem. There may be subtler issues with AI models, especially if they are used in the process of scientific discovery itself. Much of the data used to train them, for instance, will by necessity be somewhat old. That risks leaving models stuck behind the cutting edge in fast-moving fields.

Another problem arises when AI models are trained on AI-generated data. Training a machine on synthetic MRI scans, for example, can get around issues of patient confidentiality. But sometimes such data can be used unintentionally. LLMs are trained on text scraped from the internet. As they churn out more such text, the risk of LLMs inhaling their own outputs grows.

That can cause “model collapse”. In 2023 Ilia Shumailov, a computer scientist at the University of Oxford, co-authored a paper (yet to be peer-reviewed) in which a model was fed handwritten digits and asked to generate digits of its own, which were fed back to it in turn. After a few cycles, the computer’s numbers became more or less illegible. After 20 iterations, it could produce only rough circles or blurry lines. Models trained on their own results, says Dr Shumailov, produce outputs that are significantly less rich and varied than their training data.

Some worry that computer-generated insights might come from models whose inner workings are not understood. Machine-learning systems are “black boxes” that are hard for humans to disassemble. Unexplainable models are not useless, says David Leslie at the Alan Turing Institute, an AI-research outfit in London, but their outputs will need rigorous testing in the real world. That is perhaps less unnerving than it sounds. Checking models against reality is what science is supposed to be about, after all. Since no one fully understands how the human body works, for instance, new drugs must be tested in clinical trials to figure out whether they work.

For now, at least, questions outnumber answers. What is certain is that many of the perverse incentives currently prevalent in science are ripe for exploitation. The emphasis on assessing academic performance by how many papers a researcher can publish, for example, acts as a powerful incentive for fraud at worst, and for gaming the system at best. The threats that machines pose to the scientific method are, at the end of the day, the same ones posed by humans. AI could accelerate the production of fraud and nonsense just as much as it accelerates good science. As the Royal Society has it, nullius in verba: take nobody’s word for it. No thing’s, either.





霾里看花

AI可能会在推动科学进步的同时助长学术不端

幻觉、深度造假和无稽之谈：其中风险多多【深度】

最近，在伦敦的英国皇家学会（Royal Society）的一个会议室里，几十名研究生被安排与一个大语言模型（LLM）斗智。LLM是一种用于进行有用对话的AI，通常会在编程时给它们设置护栏，以防它们给出被认为有害的回答，比如教人在浴缸里制造塞姆汀塑料炸药（Semtex），或者言之凿凿地给出实际上并不真实的“事实”。

此次活动是皇家学会与美国非营利组织“人道智能”（Humane Intelligence）联合组织的，目的是要冲破这些护栏。有些结果纯属搞笑。在一名研究生的引导下，聊天机器人声称鸭子可以用作空气质量的指标（据说鸭子很能吸收铅）。在另一名研究生的提示下，聊天机器人声称多国卫生部门支持使用薰衣草油治疗长新冠（实际上并没有这回事）。但最成功的一例是通过提示让机器生成了根本不存在的学术文章的标题、发表日期和期刊。“这是我们设定的最容易的挑战之一。”人道智能的朱塔·威廉姆斯（Jutta Williams）说道。

AI有可能成为科学发展的一大利好。乐观主义者大谈机器如何对复杂研究领域的著述生成易懂的摘要，不知疲倦地分析海量数据以提出新药或新材料的建议，甚至有朝一日能够自己提出假设。但AI也有不利的一面。它有可能让科学家更容易钻系统的空子，甚至彻头彻尾地搞学术欺诈。而且这些模型本身也受到不易察觉的偏见的影响。

先看看最简单的问题：学术不端。一些期刊允许研究人员利用LLM来辅助撰写论文，前提是必须对此作出声明。但并不是每个人都愿意承认自己有这么做。有时一望便知作者使用了LLM。图卢兹大学（University of Toulouse）的计算机科学家纪尧姆·卡巴纳克（Guillaume Cabanac）发现数十篇论文中包含诸如“重新生成回复”（regenerate response）之类的短语，这是ChatGPT某些版本中按钮上的文字，点击这个按钮可以命令该

程序重新编写其最近一次的回答，作者应该是不小心把它复制到了文本中。

无法确定这类问题有多普遍。但间接指示物可以透露一二。大型科学论文出版集团泰勒弗朗西斯（Taylor and Francis）调查的学术诚信案件数量从2021年的约80起增加到2022年的约2900起，而在2022年还只有内行人能够用上LLM。2023年早些时候的数据显示这个数字将要翻倍。一个可能的线索是奇怪的同义词：例如，把“云计算”也称作“霾计算”（haze figuring），或者用“伪意识”（counterfeit consciousness）代替“AI”。

即使是诚信的研究人员也可能遇到被AI污染的数据。去年，瑞士联邦理工学院（Swiss Federal Institute of Technology）的罗伯特·韦斯特（Robert West）和他的学生们通过Mechanical Turk（一个用户可以发布零工机会的网站）雇用了一些人远程工作，为长段文字撰写摘要。在一篇于6月发表的论文中（尚未经过同行评审），该团队公布的数据显示，他们收到的所有摘要中超过三分之一是在聊天机器人的帮助下生成的。

韦斯特的团队得以将收到的摘要与另一组完全由人工生成的摘要做比较，轻而易举地发现了这种造假行为。并非所有使用Mechanical Turk的科学家都能如此幸运。许多学科依赖此类平台寻找愿意回答问卷调查的受试者，特别是社会科学。如果许多答复来自机器而不是真实的人，他们的研究质量似乎不太可能提升。韦斯特现在计划对其他众包平台进行类似的检视，但他不愿透露平台的具体名称。

能造假的不仅是文本。斯坦福大学的微生物学家伊丽莎白·比克（Elisabeth Bik）是识别科学论文中的可疑图像方面的权威，她在2016年至2020年期间发现，有几十篇论文中的图像尽管来自不同的实验室，却带有相同的特点。比克和其他人随后又识别出一千多篇有问题的论文。比克猜测，最有可能的情况是这些作者是为了支持论文结论，有意借助AI生成了这些图像。

目前还没有识别计算机生成内容（无论是图像还是文字）的可靠方法。加

拿大布鲁克大学（Brock University）的研究员拉胡尔·库马尔（Rahul Kumar）在去年发表的一篇论文中发现，学术界只能正确识别出约四分之一的计算机生成文本。AI公司试图在生成内容上嵌入“水印”，但这些其实很容易绕过。比克说：“我们现在可能已经到了无法区分照片真假的阶段。”

生成造假论文并不是唯一的问题。AI模型可能带有更不易觉察的问题，尤其是在把它们用到科学发现过程中时。例如，用于训练这些模型的许多数据不可避免地会有些过时。这可能导致模型在快速发展的领域中滞后于前沿。

用AI生成的数据训练AI模型带来了另一个问题。例如，训练机器识别合成磁共振成像扫描可以解决患者隐私问题。但有时这些数据可能会被随意使用。训练LLM的文本从互联网抓取。随着这些模型大量产生更多这样的文本，它们吸收自己的输出的可能性也增加了。

这可能会导致“模型崩溃”。2023年，牛津大学的计算机科学家伊利亚·舒迈洛夫（Ilia Shumailov）与他人合著了一篇论文（尚未经过同行评审），其中一个模型被输入手写数字，要求它识别后自己生成数字，然后再将它生成的数字输入模型。经过几轮循环后，计算机生成的数字已变得差不多无法辨认。20轮循环之后，它只能生成粗略的圆圈或模糊的线条。舒迈洛夫表示，基于自产自用的结果进行训练的模型所生成的输出在丰富性和多样性上明显不如其训练数据。

有人担心，计算机生成的内容可能来自于内部运作方式未被理解的模型。机器学习系统是难以被人类解构的“黑盒子”。伦敦的AI研究机构艾伦·图灵研究所（Alan Turing Institute）的大卫·莱斯利（David Leslie）表示，无法解释的模型并非无用，但它们的输出将需要在现实世界中经过严格的测试。这或许没有听起来那么令人不安。毕竟，在现实中验证模型正是科学研究的目标。例如，鉴于没有人完全理解人体的工作机制，新药必须通过临床试验来确定是否有效。

至少在目前，问题比答案多。已经可以肯定的是，当前科学界普遍存在的许多不合理的激励机制都很容易被利用。例如，将评估学术成就与研究人员发表的论文数量挂钩，在最坏的情况下会刺激造假，在最好的情况下也会助长操纵系统。归根到底，计算机对科研方法构成的威胁与人类构成的威胁是相同的。AI可能会加速造假和无稽之谈的生成，正如它加速了有益的科学进步一样。英国皇家学会的会训是“不人云亦云”（Nullius in verba）。同样，我们也不应该物云亦云。■



Free exchange

The false promise of friendshoring

America, China and Europe appear to be trading less with their geopolitical rivals

EACH YEAR the 193 member states of the United Nations General Assembly vote on dozens of resolutions, earnestly setting the world to rights. In December, for example, they voted in favour of reducing space threats, eradicating rural poverty and combating dust storms, among other things. The votes count for little. The assembly's resolutions are not legally binding. Its budgetary powers are small. And it has as many military divisions as the pope.

But for scholars of international relations, these votes have long provided a handy, quantitative measure of the geopolitical alignments between countries. More recently, economists have also turned to them. Owing to the trade war between America and China, Russia's invasion of Ukraine, the conflict in Gaza and recent blockades in the Red Sea, geopolitics has become impossible for dismal scientists to ignore. Although their models of trade and investment typically give pride of place to the economic size of countries and the geographic distance between them, they are now considering "geopolitical distance" as well.

The latest such study was published last month by the McKinsey Global Institute, a think-tank attached to the consultancy of the same name. By analysing countries' votes on 201 of the higher-profile resolutions between 2005 and 2022, McKinsey was able to plot countries' geopolitical stances on a scale from zero to ten. America stands at one end of the spectrum, labelled zero. At the other end is Iran at ten. In between are countries like Britain at 0.3, Brazil at 5 and China at 9.6.

The authors use this measure to provide a new perspective on each

country's trade. As well as measuring the average geographical distance that a country's trade must travel, they also calculate the geopolitical distance it must traverse. In a hypothetical world in which half of Iran's trade was with America and half with Brazil, its trade would cover a geopolitical distance of 7.5.

Their results are illuminating. European countries trade mostly with one another. As a consequence, their trade for the most part flows to their friends and neighbours. Things are rather less comfortable for Australia, however. It must trade with countries that are both geopolitically and geographically remote.

America is somewhere in between. Thanks in part to its continental size, it has few prosperous neighbours. Less than 5% of global GDP is generated by countries within 5,000km of America, as McKinsey points out. Its trade travels almost 7,200km on average, compared with 6,600km for China's trade and a global average of less than 5,200km. Yet in the diplomatic realm, the world is not so far away. The geopolitical distance America's trade must cover is only a little above the global average. It is far shorter than the diplomatic distances bridged by China. Indeed, China's trade covers a greater geopolitical gap than that of any of the other 150 countries in McKinsey's data, bar Nicaragua, which resents America, but is doomed to do business with it.

The study finds some early evidence of "friendshoring". Since 2017, America has managed to shorten the geopolitical distance covered by its trade by 10%, on McKinsey's scale. It has, for example, sharply curtailed imports from China, although some of the goods it now buys from other countries, such as Vietnam, are full of Chinese parts and components. China has also reduced the geopolitical distance of its trade by 4%, although that has required it to trade with countries farther afield geographically.

Yet the report identifies several limits to this trend. Much of the trade countries carry out with ideological rivals is trade of necessity: alternative suppliers are not easy to find. McKinsey looks at what it calls “concentrated” products, where three or fewer countries account for the lion’s share of global exports. This kind of product accounts for a disproportionate share of the trade that spans long geopolitical distances. Australia, for example, dominates exports of iron ore to China. Likewise China dominates exports of batteries made from neodymium, a “rare-earth” metal.

The attempt to reduce geopolitical dangers may also increase other supply-chain risks. Friendshoring will give countries a narrower range of trading partners, obliging them to put their eggs in fewer baskets. McKinsey calculates that if tariffs and other barriers cut the geopolitical distance of global trade by about a quarter, the concentration of imports would increase by 13% on average.

For countries in the middle of the geopolitical spectrum, friendshoring has little appeal. They cannot afford to limit their trade to other fence-sitters, because their combined economic clout is still too small. Countries that score between 2.5 and 7.5 on McKinsey’s scale—a list that includes rising economies such as Brazil, India and Mexico—account for just one-fifth of global trade. To avoid falling between two stools, they must seek to trade across the geopolitical spectrum, just as they do now.

Friendshoring has limits for China as well. There are simply not enough big economies in its geopolitical orbit to compensate for reduced trade with unfriendly Western trading partners. For China, then, friendshoring is more about replacing rivals and antagonists with more neutral parties among the non-aligned world, such as in Central Asia and the Middle East.

| *Check mate*

In studying how trade might contort itself along geopolitical lines, the McKinsey study assumes that the lines themselves remain fixed. But as the report freely admits, that might not be the case. The invasion of Ukraine and the conflict between Israel and Gaza is already causing new divisions and allegiances. It is conceivable that non-aligned countries might move closer to China politically, as China embraces them economically.

Certainly, by spurning Chinese trade and investment, the West would give China added incentive to ingratiate itself with the rest of the world. After all, there are two ways to shorten the geopolitical distance of trade: trade more with friends or make more friends to trade with. ■



自由交流

友岸外包的虚假承诺

美国、中国和欧洲似乎在减少与地缘对手的贸易

每年，联合国大会的193个成员国都会对数十项决议投票，热切地要让世界走在正确的方向上。例如在去年12月，它们投票支持减少太空威胁、根除农村贫困和对抗沙尘暴等议题。但这些投票作用不大。联大的决议不具有法律约束力，它的预算权力很小，手里的军事力量也就与教皇相当。

但对于国际关系学者来说，这些投票长期以来提供了衡量国家间地缘关系的一项便捷的定量指标。更近些时候，经济学家们也开始关注这些投票。由于美中贸易战、俄罗斯入侵乌克兰、加沙冲突以及近期红海航路被封锁，地缘态势已经成为经济学家无法忽视的问题。尽管他们的贸易和投资模型通常把国家的经济规模和彼此间的地理距离放在首位，他们现在也在考虑“地缘距离”的因素了。

最新一项此类研究在1月由咨询公司麦肯锡旗下智库麦肯锡全球研究院发布。通过分析各国在2005年至2022年间对201项有较高关注度的决议的投票情况，麦肯锡得以在一个从零到十的刻度表上标记出国家的地缘立场。美国位于刻度的一端，标记为零。另一端是伊朗，标记为十。中间则有譬如英国（0.3）、巴西（5）和中国（9.6）等国。

研究人员使用这个刻度表为各国的贸易提供了一个新视角。除了测量一个国家的贸易必须经过的平均地理距离外，他们还计算了它必须跨越的地缘距离。假想一下，如果伊朗一半的贸易是与美国进行，另一半是与巴西，那么它的贸易将跨越7.5的地缘距离。

他们的研究结果很有启发性。欧洲国家主要是在彼此之间贸易，因此它们的贸易大多流向友好邻国。然而，对于澳大利亚来说情况就不太舒适了，它必须与地缘和地理距离都遥远的国家做贸易。

美国则介于两者之间。一定程度上由于其国土占据整个大洲的很大部分，美国并没有太多富裕的邻国。麦肯锡指出，美国周围5000公里范围内的国家只产出了全球GDP的5%不到。它的平均贸易距离接近7200公里。相比之下，中国的平均贸易距离为6600公里，而全球平均水平不到5200公里。然而在外交领域，世界离美国就没那么遥远了。美国贸易必须跨越的地缘距离仅略高于全球平均。这比中国所跨越的外交距离要短得多。实际上，中国的贸易要跨越的地缘距离超过了麦肯锡数据中的其他150个国家，除了尼加拉瓜（该国怨憎美国，但又注定要与之做生意）。

研究发现了“友岸外包”的一些早期迹象。自2017年以来，美国已经在麦肯锡的刻度表上将其贸易的地缘距离缩短了10%。例如，它大幅减少了从中国的进口，尽管它现在从越南等其他国家购买的一些商品中充斥着中国生产的零部件。中国也将其贸易的地缘距离缩短了4%，尽管这需要它与地理上更遥远的国家进行贸易。

然而，报告指出这个趋势的几方面有限性。与意识形态对手进行的贸易大部分都是必要的贸易——因为难以找到替代的供应商。麦肯锡研究了它称之为“集中型”的产品，即被三个或更少的国家占据了全球出口额大头的产品。在长地缘距离的贸易中，这类产品占据了尤其高的比例。例如，澳大利亚主导了对中国出口铁矿石。同样，中国主导了用钕这种“稀土”金属制造的电池的出口。

尝试减少地缘风险也可能增加其他供应链风险。友岸外包将使各国的贸易伙伴范围变窄，迫使他们把鸡蛋放在更少的篮子里。麦肯锡计算，如果关税和其他壁垒将全球贸易的地缘距离减少约四分之一，那么进口的集中度将平均增加13%。

对于处于地缘光谱中间地带的国家来说，友岸外包的吸引力不大。它们不能仅仅和其他居间观望的国家做贸易，因为大家加总的经济实力仍然太小。在麦肯锡的刻度表上处于2.5到7.5之间的国家——包括巴西、印度和墨西哥等崛起中的经济体——只占全球贸易的五分之一。为免落个两头不着，它们必须寻求跨越地缘光谱做贸易，就像它们现在所做的那样。

友岸外包对于中国来说也有局限。在它的地缘轨道上，没有足够多的大型经济体来弥补与不友好的西方贸易伙伴减少的贸易。因此中国的友岸外包更多是用更中立的不结盟国家，如中亚和中东，来替换对手和敌对者。

| 打盟友牌

在研究贸易如何沿着地缘路线扭曲自身时，麦肯锡的研究假设这些路线本身是不变的。但正如报告坦率承认的，实际情况可能并非如此。俄罗斯入侵乌克兰和以色列与加沙的冲突已经导致了新的分裂和联盟。可以想象，随着中国在经济上拥抱不结盟国家，这些国家可能在政治上向中国靠拢。当然，西方抛弃中国的贸易和投资会给中国更多动力去取悦世界其他国家。毕竟，缩短贸易的地缘距离的方式有两种：与朋友做更多贸易，或者结交更多朋友做贸易。 ■



After the iPhone

Apple's Vision Pro headset ushers in a new era of personal technology

Tech firms are racing to build the gadget that supplants the smartphone

APPLE FANS have eagerly awaited February 2nd to get their hands on the tech giant's latest gadget, a new augmented-reality headset called the Vision Pro. Some early reviewers complained that it caused headaches and had a two-hour battery life. Many potential buyers will be put off by the price tag of \$3,499. Still, perhaps 200,000 have been pre-ordered, about 40% of what Apple had reportedly expected to sell this year. Tim Cook, Apple's boss, has described trying the Vision Pro as an "aha moment". "You only have a few of those in your lifetime," he added.

Aha or not, the Vision Pro is part of a trend. In September techies got excited about a new pair of smart glasses made by Meta, Facebook's parent company, and Ray-Ban, an eyewear brand. The spectacles are controlled by voice and can play music, send texts and film everything you see. Two months later Humane, a startup founded by former Apple executives, launched the Pin, a brooch with which users interact by talking and gesticulating. In January the r1, a voice-controlled gizmo half the size of a smartphone, enthralled attendees at the Consumer Electronics Show in Las Vegas. Its maker, a startup called Rabbit, has sold nearly 100,000.

What all these devices have in common is that they mostly do away with screens, keyboards and mice. Thanks to "generative" artificial intelligence (AI), computers are getting good at listening to, reading and watching stuff—and understanding it. That means hardware can be controlled by voice, gesture or image rather than touch. AI is thus enabling new "form factors"—tech-speak for gadgets in new shapes and sizes, just as the iPhone looked different from older handsets.

Silicon Valley's elite are cheering on the potential shift. They believe AI could create a new market for consumer hardware, replacing the smartphone as everyone's essential device. Sam Altman, boss of OpenAI, the startup behind ChatGPT, is reportedly in talks to start a firm with Jony Ive, former head of design at Apple, to make a gadget purpose-built for AI. Satya Nadella, chief executive of Microsoft, an AI-ambitious tech titan, recently said that "once you have a new interface...new hardware is also possible."

One reason for all the excitement about new gadgets is that the old ones are looking unexciting. Last year 1.2bn smartphones were sold worldwide, down by 3% from the previous year and the lowest level for a decade, according to IDC, a research firm. PCs did even worse, declining by 15% in 2023 to 242m units. Cash-strapped consumers are opting for cheaper alternatives, such as second-hand devices, or holding on to their current ones for longer.

The hope is that they may be persuaded to fork out for all-new gadgets because they offer something that old ones do not. AI could, for instance, make using devices more seamless and more personal. Users can tell or gesture to the r1 to hail a ride, order food or play music without the need to toggle between apps. It also learns from users' previous actions. Until now people had to adapt to software, says Vinod Khosla, a veteran venture capitalist and early backer of Rabbit. In the r1, "the AI adapts to you."

New gadgets are also less finicky to develop and manufacture. Lior Susan of Eclipse, a venture-capital (VC) firm, says that ten years ago building a high-tech widget required hundreds of staff. Today he can do the same thing with about ten. Every step of the manufacturing process has become easier. Initial versions can be mocked up in design software. Rather than buying an industrial machine to make parts for a prototype, they can be ordered from 3D-printing firms like Shapeways. Sensors, batteries and chips can be

bought off the shelf. Contract manufacturers, such as Foxconn, no longer insist on working only for big clients like Apple. Some offer dedicated services for hardware startups.

The resulting crop of new AI-powered devices falls into two broad categories. The first is headsets for augmented or virtual reality (VR). So far they have been most popular among gaming enthusiasts. Sales of VR headsets hit around 10m units in 2020 following the release of Meta's Quest 2, estimates George Jijiashvili of Omdia, a research firm. He thinks that the Vision Pro will breathe new life into the industry by making VR appealing to non-gamers (see chart). Promotional videos depict people using the Vision Pro to watch films, work or talk to friends.

The second category consists of subtler gizmos. Some 540m "wearables" worth \$68bn were shipped last year, according to IDC. Many already incorporate AI in one way or another. They include earphones (which account for 63% of the units sold), smartwatches (another 30%), wristbands such as the Whoop, a fitness tracker, and smart glasses, like Meta's Ray-Bans (which together make up most of the remainder). Humane's Pin and AI pendants made by two startups, MyTab AI and Rewind AI, are the latest additions to this group.

All these devices are nifty. Whether they are nifty enough to dislodge the smartphone and become the next big platform is another matter. For that to happen, consumers must take to them. This requires the things, first, to look good—which some failed early efforts, such as the dorky Google Glass specs, did not. The r1 owes its sleek retro feel to Rabbit's collaboration with Teenage Engineering, a Swedish design firm. Before its launch, the Pin appeared on a Paris catwalk at an event held by Coperni, a French fashion house. Meta's glasses are a hit in part because Ray-Ban knows what makes shades stylish.

Second, the new gadgets have to be useful in ways the old ones are not. Many hardware-makers are adding AI to existing devices. On January 31st Samsung started selling an AI smartphone that can do neat tricks such as summarising text-message threads. Microsoft's next generation of laptops and tablets will reportedly include specialist AI chips and a new keyboard button to summon "Copilot", its AI chatbot. Smart speakers, such as Amazon's Alexa and Google's Nest, and earphones, such as Apple's AirPods, are getting revamped with AI features. These including chatbots and, with AirPods, the ability to let through necessary sounds and turn down volume when the wearer is speaking.

To break through, the AI hardware will have to make life either much easier (for instance by booking a whole trip, flight, car and hotel included, with a single command) or much more marvellous (inspiring Mr Cook's "aha moment").

Users will also expect them to perform more than a couple of functions. That means lots of apps. Meta's latest VR headset, the Quest 3, offers 500 or so. The Vision Pro already boasts around 350 purpose-built apps, and can run the iPhone versions of most of the roughly 2m available in the App Store. Humane's Pin, which doubles as a phone, claims to be doing away with apps, instead offering a range of "AI-powered services" from providers such as OpenAI and Google. Rabbit's r1 piggybacks on smartphones' existing app universe, at least for the time being.

Third, although manufacturing things has got easier, managing supply chains remains the hardest part of running a hardware business, notes Shaun Maguire of Sequoia, another VC firm. Suppliers may take phone calls from smaller firms but some are still reluctant to give good prices to unproven newcomers with small orders.

None of the available AI devices overcomes all three challenges. Those that

look pretty, like the r1, the Pin or Meta's Ray-Bans, seem to be peripherals more akin to AirPods than the iPhone. Independently useful ones like the Vision Pro or the Quest are dorkier than Google Glass, and much clunkier. In addition, developing apps for Apple's headset is expensive, which is putting off developers, including some video-game studios, Netflix, Spotify and YouTube (which also happen to compete with Apple's own video and music-streaming services). Production problems afflict just about everyone. Jesse Lyu, founder of Rabbit, says that it took his product becoming an overnight sensation for him to gain a bit more bargaining power over his suppliers. Even Apple, the master of supply chains, reportedly had to scale back initial plans to ship 1m Vision Pros this year because of the complex manufacturing involved.

If some gadget-makers clear all three hurdles, they may stumble on another: keeping up with the breathtaking pace of AI advances. Apple took seven years to develop the Vision Pro, aeons in AI time. Even the next generation of Rabbit's device, which Mr Khosla says will be ready as soon as this summer, may be outmoded by the time it gets into users' hands. One of today's AI gadgets may one day dethrone the smartphone. More likely, the winning form factor has yet to take shape. ■



iPhone之后

苹果的Vision Pro头显开启个人科技产品新时代

科技公司竞相制造取代智能手机的小设备【深度】

科技巨头苹果公司于2月2日推出最新产品——一款名为Vision Pro的新型增强现实头显。粉丝们翘首以待，跃跃欲试。一些先期评测者抱怨这款头显会引发头痛，而且电池续航时间只有两小时。许多有意向的买家也会因其3499美元的标价而却步。不过Vision Pro可能已经预售出20万台，约占传言中苹果预期年销量的40%。苹果老板库克形容试用Vision Pro是个让人惊艳的时刻。“这样的感觉一生中也就只有几次。”

不管是否如此惊艳，Vision Pro的确是潮流的一部分。去年9月，Facebook母公司Meta联合眼镜品牌雷朋推出了一副全新智能眼镜，让科技迷兴奋不已。这款眼镜由语音控制，可以播放音乐，发送信息，还可以将你目光所及都摄录下来。两个月后，由苹果公司前高管创办的公司Humane推出一款名为Pin的AI胸针，用户可以通过语音和手势与之互动。今年1月，在拉斯维加斯举行的国际消费电子展上，一款只有智能手机一半大小的声控电子设备让参会客商为之着迷。它的开发者、创业公司Rabbit已经卖出了近十万台设备。

这些设备的共同点是大多都摒弃了屏幕、键盘和鼠标。得益于“生成式”AI，计算设备越发擅长聆听、阅读和观看素材——并理解它们。这意味着硬件可以通过语音、手势或图像而非触摸来控制。AI就此促成新的“形式要素”（意指全新形态和尺寸的小型电子设备的行话）出现，正如iPhone看起来如此不同于它之前的手机一样。

硅谷精英们正为这可能发生的变革而欢呼。他们认为AI可以创造全新的消费硬件市场，取代智能手机成为又一个人人必备的电子产品。据称，ChatGPT背后的创业公司OpenAI的老板山姆·阿尔特曼（Sam Altman）正与曾任苹果设计主管的乔尼·艾维（Jony Ive）商谈成立新公司，打造一款专门运用AI的小设备。科技巨头微软想在AI领域大展拳脚，CEO纳德拉

最近表示：“只要有了新接口.....新硬件也就可能出现。”

人们如此热切期待新设备的原因之一是旧设备的市场已经显得乏味无趣。研究公司IDC的数据显示，去年全球共售出12亿台智能手机，比前一年下降了3%，是十年来的最低水平。个人电脑的情况更糟糕，2023年的销量下降了15%，至2.42亿台。手头拮据的消费者转向购买二手设备等更便宜的替代品，或者继续使用老设备。

业界希望，一旦全新的设备能提供旧设备所不具有的东西，消费者可能就会被说服而出手。比如说，AI能令使用设备的体验更方便流畅和个性化。用户可以用语音或手势示意打车、点餐或播放音乐，而无需在不同的应用之间切换。它还会从用户以往的操作中学习。资深风投家、Rabbit的早期投资者维诺德·科斯拉（Vinod Khosla）说，在此之前，人必须去适应软件，有了AI，那就是“AI来适应你”。

新产品的开发和制造也变得不那么繁杂棘手。风投公司Eclipse的利奥·苏珊（Lior Susan）表示，十年前制造一台高科技产品需要数百名员工，如今，同样的工作十个人左右就能完成。制造过程的每一步都已变得更容易。最初版本可以在设计软件里模拟成型。也不必购买工业机器来制造原型机的各个部件，而只要从Shapeways这样的3D打印公司订购即可。传感器、电池和芯片都可以买现成的。富士康等代工厂不再坚持只为苹果等大客户服务，有的会向硬件创业公司提供专属服务。

由此而来的新型AI设备可分为两大类。第一类是增强现实或虚拟现实（VR）头显。目前来看，最追捧这类产品的是游戏玩家。据研究公司Omdia的乔治·吉加什维利（George Jijiashvili）估计，在Meta推出Quest 2之后，VR头显的销量在2020年达到约1000万台。他认为Vision Pro能使非游戏玩家也被VR吸引，从而为该行业注入新的生机（见图表）。Vision Pro的宣传视频展现了人们使用该设备看电影、工作及与朋友聊天的场景。

第二类是相对不起眼的小设备。据IDC统计，去年“可穿戴设备”的出货量

约为5.4亿台，价值680亿美元。许多这些产品已经多多少少带有AI功能。按销量计算，其中63%为耳机，30%为智能手表，其余大部分为手环（如健身追踪器Whoop）和智能眼镜（如Meta与雷朋合作的那款智能墨镜）。Humane的Pin胸针和两家创业公司MyTab AI和Rewind AI打造的AI吊坠是该类设备的最新成员。

所有这些设备都很精巧，但是否足以取代智能手机成为下一个大平台则是另一回事了。这要有消费者买账才行。首先新设备要好看，而早期一些失败产品就没做到这一点，比如造型不够自然好看的谷歌眼镜。r1的时尚复古感归功于Rabbit与瑞典设计公司Teenage Engineering的合作。在发布之前，r1曾在法国时尚品牌哥白尼（Coperni）的一场巴黎时装秀上亮相。Meta的眼镜之所以大受欢迎，部分原因是雷朋深谙墨镜长啥样才时髦优雅。

其次，新设备必须能旧设备所不能。许多硬件制造商正在为现有设备添加AI功能。1月31日，三星开始销售一款AI智能手机，附带总结短信线程等巧妙功能。据称，微软的下一代笔记本和平板电脑将包含专门的AI芯片并新增一个用于召唤AI聊天机器人“Copilot”的按钮。亚马逊Alexa和谷歌Nest等智能音箱以及苹果AirPods等耳机都在用AI功能翻新，比如添加聊天机器人。AirPods会在佩戴者说话时让必要的音频进入并降低音量。

要取得突破，AI硬件必须能大大提升生活的便利度（例如只需一个命令就能完成整个旅程的预订，包括机票、用车、酒店等）或者精彩程度（引发库克所谓的“惊艳时刻”）。

用户还会期望它们实现更多功能。这意味着要有大量的应用。Meta最新的VR头显 Quest 3提供约500个应用。Vision Pro已有约350个专用应用，还可以运行App Store中约200万个应用中大部分的iPhone版本。Humane声称其兼具手机功能的Pin胸针摆脱了应用，而是提供一系列由OpenAI和谷歌等供应商提供的“AI驱动的服务”。Rabbit的r1则依赖于智能手机的现有应用体系，至少目前仍然如此。

第三，虽然制造电子设备变容易了，但管理供应链仍是硬件企业的经营中最艰难的部分，另一家风投公司红杉资本（Sequoia）的肖恩·马奎尔（Shaun Maguire）指出。供应商也许会接听小公司的询问电话，但有些还是不愿意给订单量小而且实力不明的新公司提供优惠价格。

现有AI设备还没有哪个克服了所有三个挑战。r1、Pin或Meta的雷朋墨镜这些外型好看的设备似乎属于和AirPods更接近的外设，而不是iPhone这样的独立产品。而Vision Pro或Quest这些独立使用的产品的外型比谷歌眼镜更突兀，而且笨重得多。此外，为苹果头显开发应用的成本很高，让开发者望而却步，包括一些电子游戏公司、奈飞、Spotify和YouTube（它们也恰好与苹果自己的视频和音乐流服务存在竞争关系）。几乎每家公司都在生产环节备受困扰。Rabbit的创始人吕聘称，产品一夜爆火后他在供应商面前的议价能力才有一点提升。据报道，因为制造工艺复杂，即便是供应链之王的苹果也不得不缩减最初要在今年出货100万台Vision Pro的计划。

假如有设备制造商扫除了这全部三个障碍，它们可能又会遇到另一个难题：跟上AI惊人的发展速度。苹果花了七年才开发出Vision Pro，这以“AI时间”衡量可说极为漫长。就连科斯拉说今年夏天便可推出的Rabbit新一代设备，等送到用户手上可能也已经过时了。今天的某个AI小设备也许有一天能取代智能手机。但更有可能的是，最终能脱颖而出的形式要素目前尚未成型。■



Schumpeter

Can MSCI drag private markets out of the shadows?

Meet the Nicaraguan revolutionary behind the world's favourite index supplier

HENRY FERNANDEZ was once a counter-revolutionary. The man who over three decades has built MSCI, a provider of stockmarket indices, into a standard bearer of financial globalisation, started his career as a Nicaraguan diplomat in the government of Anastasio Somoza, a right-wing dictator. While some of his friends flocked to the left-wing Sandinistas ahead of the revolution that toppled Somoza in 1979, he took a look at socialism in eastern Europe and decided it was doomed to fail. Instead, he embraced free-market capitalism and moved to Wall Street.

There he encountered a different revolutionary movement that he has championed ever since: the forward march of capital markets. Starting in the Reagan era, he has played a role in several of the upheavals that have swept the world of finance, from securitisation in the 1980s and the growth of emerging-market investing in the 1990s, to the rise of index tracking and exchange-traded funds (ETFs) this century. He retains an idealistic streak. While many former advocates of environmental, social and governance investing have shied away from the climate-related fad of the past half-decade, he remains a true ESG believer.

He is now betting that MSCI's indices can penetrate the opaque world of private finance—the \$12trn-plus of assets held in private equity, credit, venture capital, real estate and infrastructure. These are some of the hottest segments of the capital markets. But they are restricted to institutions and well-heeled investors. In these secretive markets asset managers are loth to encourage more transparency and liquidity lest their fees suffer as a result. Yet technology may be moving in Mr Fernandez's favour.

MSCI, which was spun off by Morgan Stanley, an investment bank, in 2007 and has a market capitalisation of \$45bn, has two main lines of business. The first is benchmarking. It has more than 280,000 equity indices around the world that tell investors what is going on in the public markets, and provide a measuring stick against which to judge fund managers' performance. If a fund puts all its money into small-cap Japanese stocks, for instance, and MSCI's medium- and large-cap Japanese equity indices do better, it underperforms. Almost \$15trn of assets are benchmarked in such a way globally.

The second line of business is enabling investment managers to sell low-cost portfolios, such as ETFs, based on its indices. Almost \$1.5trn of ETF assets are linked to MSCI's indices, a nearly five-fold increase in a decade. BlackRock, the world's largest asset manager, is the biggest client. Its boss, Larry Fink, and Mr Fernandez have been kindred spirits for decades.

MSCI's first foray into the private realm is via benchmark indices. Since 2021 it has spent almost \$2bn buying two data-gathering firms that create indices for private assets, from real estate and infrastructure to private debt. As Mr Fernandez explains, such indices enable a property investor to decide the relative merits of putting money into, for example, offices (which crashed during the pandemic) versus data centres (which soared). Gathering such information is tricky because many of the transactions are not publicly disclosed. MSCI creates indices by sourcing data from investors in private funds, who in turn receive records of those funds' quarterly performance, including valuations of the underlying assets, from the asset managers. Its most recent acquisition, for instance, gives it data from about 13,000 private funds, representing \$15trn in cumulative investments.

Could these benchmarks eventually form the basis for indices used by ETFs to bring private markets to the masses? It seems hard to imagine. Private

assets do not trade with anything like the frequency of listed assets. They also lack the liquidity necessary for passive funds whose investors may want to redeem their money at short notice.

And yet Mr Fernandez believes that some parts of this opaque hinterland, such as private loans, are more liquid than others. “My bet is that over time there will be the development of a secondary market for private credit,” he says. To explain why, he goes back to his days as a young trader at Morgan Stanley during the “Liar’s Poker” era of the 1980s. The market for mortgage loans was fledgling and illiquid until the thrifts that made home loans came under pressure to sell them. Wall Street firms like Salomon Brothers and First Boston (where Mr Fink headed the mortgage desk) scooped them up, turned them into mortgage-backed securities and sold them to investors, creating a highly liquid secondary market. Similarly, banks that underwrite loans today face regulatory pressure to limit the size of their balance-sheets, so they sell some of the loans to firms with private-credit arms, such as Apollo and Blackstone. Mr Fernandez thinks that, as with the mortgage market, this trade could lead to a secondary market that would, eventually, have enough liquidity for index funds.

A big breakthrough will require advances in technology. For that, Mr Fernandez has his eye on his veteran comrade-in-arms, Mr Fink. This month BlackRock launched its first bitcoin ETF, and Mr Fink, who seldom keeps his cards close to his chest, hinted that this may be the start of a prolonged foray into the cryptoverse that could eventually encompass private assets. “If we could ETF a bitcoin, imagine what we could do with all financial instruments,” he told Bloomberg TV. “Everything is going to be ETFed.”

| *ETFs or WTF?*

Mr Fernandez notes that Mr Fink has become an advocate of “tokenisation”—the idea that financial assets and their owners can be

registered on a blockchain-like ledger, which could make it easier to trade property and other private assets. It is an idea in its infancy. Some people think it is barmy. The MSCI boss confesses that for the time being he himself does not fully understand it. But in contrast to the Sandinistas, who have betrayed everything they once fought for, his revolutionary zeal remains as strong as ever. ■



熊彼特

MSCI能否照亮私人市场？

来会一会全球最受欢迎的指数供应商背后的尼加拉瓜革命家

亨利·费尔南德斯（Henry Fernandez）曾经是个反革命。在历经30多年把股票市场指数供应商MSCI打造成金融全球化的标杆之前，他最早是在尼加拉瓜的右翼独裁者阿纳斯塔西奥·索摩查（Anastasio Somoza）的政府里担任外交官。1979年推翻索摩查的革命爆发前夕，他的一些朋友转投左翼的桑地诺解放阵线，而他研究了一番东欧社会主义后认为这种制度注定失败。于是他转投自由市场资本主义的怀抱，搬去了华尔街。

在那里，他遇上了另一场革命运动并化身其领军人物至今：资本市场的进军。从里根时代开始，在席卷金融界的多次变局中他都发挥了作用，包括上世纪80年代的证券化和90年代投资新兴市场的热潮，以至本世纪指数跟踪和交易所交易基金（ETF）的兴起。他至今仍保留着理想主义的特质。许多当年支持环境、社会 and 治理（ESG）投资的人都回避了过去五年里关注气候的风潮，而他依然是ESG的忠实信徒。

现在，他正押注MSCI指数能穿透私人金融这个不透明的世界——总资产超过12万亿美元的私募股权、信贷、风险投资、房地产和基础设施。这些是资本市场里其中一些最热门的领域，但只有机构和富裕投资者可以参与。在这些神秘市场中，资产管理公司不愿鼓励提高透明度和流动性，以免影响它们收取费用。但技术的发展可能有利于费尔南德斯。

MSCI在2007年从投资银行摩根士丹利剥离出来，如今市值450亿美元，有两大业务线。第一是制定基准。MSCI在全球拥有超过28万个股票指数，向投资者传递公开市场的动态信息，也为判断基金管理公司的业绩提供衡量标准。举例来说，如果一只基金把所有资金投入小盘日本股票，而MSCI的中盘和大盘日本股票指数表现更好，那么这只基金就是业绩欠佳。全球有近15万亿美元的资产就是这样做基准比较的。

第二条业务线是让投资管理公司可以销售基于其指数的低成本投资组合，如ETF。与MSCI指数挂钩的ETF资产如今接近15万亿美元，十年间增至原来的近五倍。全球最大的资产管理公司贝莱德（BlackRock）是MSCI最大的客户。贝莱德的老板拉里·芬克（Larry Fink）和费尔南德斯在过去几十年里可谓志趣相投。

MSCI最开始是尝试通过基准指数打入私人金融领域。自2021年以来，MSCI已斥资近20亿美元收购了两家编制房地产、基建以及私人债务等私人资产指数的数据收集公司。费尔南德斯解释说，这些指数有助房地产投资者判断把资金投入不同领域的相对优势，比如是写字楼（疫情期间崩盘）还是数据中心（疫情期间飙升）。收集此类信息有其难度，因为许多交易都没有公开披露信息。MSCI会从私人基金投资者那里获取数据来编制指数，而这些投资者则是从资产管理公司那里获得那些基金的季度业绩记录，包括相关资产的估值。举个例子，MSCI通过最近一次收购获得了约13,000只私募基金的数据，它们的投资额总计达15万亿美元。

这些基准最终能否构成ETF采用的指数的基础，进而把私人市场推向大众？这似乎很难想象。私人资产的交易频率远不及上市资产，也缺乏被动型基金（其投资者可能想在短时间内赎回资金）所需的流动性。

然而，费尔南德斯认为，这块隐秘腹地的某些部分相较于其他有更高的流动性，比如私人贷款。“我敢打赌，以后会发展出一个私人信贷的二级市场。”他说。在解释原因时，他提到1980年代自己年轻时在摩根士丹利做交易员的日子，那正是《老千骗局》（Liar's Poker）中所讲述的时代。当时抵押贷款市场刚刚起步，流动性很差，直到后来提供住房贷款的储蓄银行在压力之下要出售这些贷款。所罗门兄弟（Salomon Brothers）和第一波士顿银行（First Boston，芬克曾管理该银行的抵押贷款部门）等华尔街金融机构吸收了这些贷款，将其转化为抵押担保证券再出售给投资者，从而打造了一个流动性极高的二级市场。同样，现在发放贷款的银行受到监管压力，需要限制资产负债表的规模，于是把部分贷款出售给阿波罗全球管理公司（Apollo）和黑石（Blackstone）等拥有私人信贷部门的公司。费尔南德斯认为，与抵押贷款市场一样，这种交易可能催生最终能为指数

基金提供足够流动性的二级市场。

要取得重大突破将有赖于技术上的进步。在这方面，费尔南德斯很关注老战友芬克的动向。1月，贝莱德推出了自家第一只比特币ETF，一向口快的芬克暗示这可能是进军加密领域的长途跋涉的开始，最终可能涵盖私人资产。“如果我们能把比特币ETF化，想想我们可以对所有金融工具做什么？”他在彭博TV的采访中表示。“一切都可以被ETF化”。

| ETF还是WTF?

费尔南德斯指出，芬克已经成了“代币化”的倡导者，即主张金融资产及其所有者可以登记在类似区块链的分类账上，这可能会让房地产和其他私人资产的交易变得更容易。这个想法仍处于起步阶段。有人觉得这是在发癫。费尔南德斯承认目前他自己也没有完全把它想明白。但有别于桑地诺主义者最后背弃了自己曾为之奋斗的信念，费尔南德斯的革命热情坚定如故。■



After the fire

Against the odds, Notre Dame cathedral will reopen this year

The rebuilding of the famous monument prompted a debate about how much should change

ON AN ICY January morning, perched at a dizzying height of nearly 100 metres above the ground, specialist roofers are covering the rebuilt oak spire of Notre Dame cathedral with layers of lead sheeting. Working on platforms reached by a perilous flight of narrow steps that cling to the soaring spire, they are putting the final touches on a 1,000-piece, solid-oak structure that will soon restore the cathedral's familiar silhouette.

Designed in 1859 by Eugène Viollet-le-Duc, an architect, and felled by the devastating fire of 2019, the new flèche remains hidden behind dense scaffolding. But five years after the world watched aghast as the gothic cathedral roof was devoured by flames, the project to rebuild Notre Dame is, astonishingly, on schedule. The cathedral doors are due to reopen in December. (Visitors for the Olympics, which Paris will host starting in late July, must wait to glimpse inside.)

The rebuilding of Notre Dame is one of the most complex and ambitious reconstruction projects that France has ever undertaken on a historic monument. Fire engulfed the entire wooden latticework that made up the medieval roof, before melting its lead casing and toppling the spire. The cathedral's nave, choir and transepts were mostly untouched by flames, thanks to the craftsmanship of the 12th- and 13th-century stonemasons. So were the 8,000-tube great organ and stained-glass windows.

However, molten lead and charred oak beams crashed through the roof, spreading embers and lead particles. Religious artefacts, paintings and sculptures had to be rescued and cleaned; the organ and 39 stained-glass

windows were dismantled and washed. It took two years to stabilise the cathedral's stone structure.

When a grave-looking President Emmanuel Macron stood outside Notre Dame on the night of the fire, he described it as “the epicentre of our lives”. The cathedral would be rebuilt, he vowed, within five years. That promise may have sounded “a bit mad”, concedes Philippe Jost, who runs the public body in charge of the reconstruction project. But, he adds, it gave everyone a clear objective.

That the project is on track is also partly due to the commanding style of General Jean-Louis Georgelin, who ran it like a military operation until his death last August. A light management team, freed from bureaucratic excess has helped, too. Gifts from rich French industrialists—the Arnault, Bettencourt and Pinault families—as well as 340,000 smaller individual donations from around the world amounted to €846m (\$921m). Unlike the usual French grands projects, this one is costing the public purse almost nothing.

Today the cramped site on the Île de la Cité, an island in the River Seine, is a veritable construction village, complete with a canteen, shower block, offices and sculpture-restoration workshop. Last summer the first eight triangular oak trusses, crafted for the new transept roof, arrived by river. Passers-by watched in awe as, one by one, each seven-tonne truss was hoisted by crane from a barge and lowered into place.

The craftsmen working on the project are specialist artisans, drawn from dozens of small firms from around France. Rather than contract the rebuilding to one company, over 140 separate tenders were put out, in order to support traditional craftsmanship. The demands were unusually high: a decision was taken to restore the cathedral to its former splendour while remaining faithful both to its original designs and the construction

techniques of the time.

For the roof above the medieval nave and choir, 1,200 oak logs were individually selected from forestry plantations in France and hewn by hand into square timber beams. Craftsmen used hand-forged axes, based on late 12th- and early 13th-century models. On the curved apse roof, finished on January 12th, wooden dowel pegs hold the trusses together without a single industrial metal piece. “The idea was to work with tools that are as close as possible to those used in the Middle Ages in order to give the wood the aspect of the time,” says Valentin Pontarollo, a carpenter from Ateliers Perrault, a firm in western France. One of the joys this brought was that there was often “no machinery noise, just the sound of the axe”, he says.

Inside the on-site restoration workshop, sculptors and stonemasons are also working by hand, with chisels and brushes. On the floor stand some of the near-finished new stone chimeras, the grotesque and menacing beasts installed in the 19th century that look down from the cathedral façade and towers. More than a dozen gargoyles, 80 fleurons, 70 small chimeras and more than 750 crockets (decorative stone hooks) are being entirely rebuilt, as are three cathedral gables. At 96 metres above ground level, a new gilt-covered copper rooster tops the needle of the spire, hoisted into place by a crane in December. It replaces the one that miraculously survived the fire, falling to the ground beneath one of the flying buttresses. (The original is on display at a museum.)

Inevitably some modernists decry the project’s conservatism. After the fire, various architects lobbied for an audacious contemporary flourish: a plate-glass spire or roof planted with vegetation. Notre Dame, after all, is an architectural hybrid, the product of different centuries, ransacked during the wars of religion and looted during the revolution. Viollet-le-Duc’s spire was added only in the 19th century, not long after Victor Hugo’s novel, “The Hunchback of Notre Dame”, revived respect for what was then an unloved

gothic masterpiece.

Today a public debate pits those who contest the authenticity of reconstruction against those who argue that authenticity is guaranteed by respect for the original, with its irregularities as well as its durability. “This is not a monument like any other, and deserves not being distorted,” says Mr Jost: “We are restoring a cathedral that is 860 years old so that it can last for at least another 860 years.”

Notre Dame’s overseers are trying to strike a balance. The pristine, newly cleaned cathedral may take some visitors by surprise. Inside the nave, where much of the scaffolding has come down, the freshly blond stone of the pillars and vaults, as it would have appeared in medieval times, is breathtaking—and a startling contrast to their previous darkened, partly pitted form.

But the cathedral has not rejected modernity altogether. The Archbishop of Paris has chosen two contemporary French designers for the furniture. Guillaume Bardet is making the liturgical furnishings, including the altar, with dark bronze, and Ionna Vautrin the new 1,500 solid-oak congregation seats, which will have slatted backs to suggest openness and light.

The rebuilding of Notre Dame is both a test and a showcase for France: of its ability to meet an ambitious deadline, to satisfy the demanding global gaze of both secularists and spiritualists and to flaunt French craftsmanship. Finishing in time is a matter of “French pride”, says Mr Jost.

For the artisans, completion will be poignant. Emile-Armand Benoit, an ornamental roofer on the cathedral’s highest point, says that they are working as quickly as possible so that the scaffolding can come down and the city “can once again see the spire reaching into the sky”. Halfway down, the carpenters contemplate the newly finished roof. “It’s a bit heartbreaking

to leave,” says Mr Pontarollo, “We’ll never have a project like this again.” ■



浴火重生

克服万难，巴黎圣母院将在今年重新开放

对这座名胜古迹的修复引发了争论：该改变它多少？【深度】

一月的一个清晨，天寒地冻。距离地面近100米的高处令人眩晕，专业屋顶工正在那里给巴黎圣母院重建中的橡木尖塔铺上一块块铅板。他们需要惊险地走过一段沿高耸尖塔向上的狭窄扶梯，才能到达各自工作的平台。他们是在对一个由1000块实心橡木组成的结构做最后的修饰。这座大教堂往日的轮廓将很快重现在世人面前。

建筑师尤金·维奥莱特-勒-杜克（Eugène Viollet-le-Duc）于1859年设计的巴黎圣母院尖塔毁于2019年的那场毁灭性大火。如今新的尖塔仍然被密实的脚手架遮挡着。但是，在世人目瞪口呆地看着大火吞噬这座哥特式大教堂屋顶的五年之后，人们惊讶地发现，巴黎圣母院的修复工程竟然在按时推进着。这座大教堂定于今年12月重新开放。（7月下旬前往巴黎观看奥运会的游客还得等一等才能一睹其内部风貌。）

巴黎圣母院的修复是法国有史以来对历史遗迹所做的最复杂、最具挑战性的修复工程之一。大火吞噬了这座中世纪教堂屋顶的木质框架，继而导致外面的铅皮熔化和尖塔坍塌。不过得益于12、13世纪石匠们的精湛技艺，大教堂的中殿、唱经堂和耳堂的大部分都未受损。有8000根音管的巨型管风琴和花窗也幸免于难。

然而，熔化的铅和烧焦的橡木横梁冲破了屋顶，余烬和铅粒四下散落。宗教器物、绘画和雕塑必须抢救和清理，管风琴和39扇花窗被拆下来清洗。加固大教堂的石质结构花费了两年时间。

火灾当晚，法国总统马克龙站在巴黎圣母院外面，表情凝重地说圣母院是“我们生活的中心”。他誓言将在五年内修复它。负责修复工程的公共机构的负责人菲利普·约斯特（Philippe Jost）承认，这个承诺听上去可能“有点疯狂”。但这给了所有人一个明确的目标，他补充道。

修复工程能够按计划进行，一定程度上也要归功于让-路易·乔治林

（Jean-Louis Georgelin）将军的严格领导——直到去年8月去世前，他都像指挥一项军事行动一样管理这个项目。一支没有繁冗官僚做派的精干的管理团队也很有助益。除了来自世界各地的34万笔小额个人捐款，修复工程还收到了来自阿尔诺、贝当古和皮诺家族等富有的法国实业家的捐赠；全部捐款共计8.46亿欧元（9.21亿美元）。与法国一般的大型项目不同，这项修复工程几乎没有动用公共资金。

如今，这块塞纳河中西堤岛（Île de la Cité）上的狭窄场地俨然成了个建筑基地，里面配有一个食堂、一个淋浴区、多个办公室和一个雕塑修复车间。去年夏天，为新耳堂的屋顶精心制作的第一批八个三角形橡木桁架通过水运抵达。每个桁架重达七吨，起重机将它们逐个从驳船上吊起，再缓缓放落到位。路人心怀敬畏地目睹了这一过程。

参与修复工程的工匠都是各有专长的手艺人，来自法国各地几十家小公司。为寻求传统技艺，修复工程经过140多次招标，发包给了不同的公司，而不是只与一家公司签订合同。各项要求异常之高——因为修复团队决定在恢复大教堂昔日辉煌的同时，还要忠于其初始设计和当时的建筑技艺。

为了制作中世纪中殿和唱经堂的屋顶，修复团队从法国的林场中一根一根地挑来共1200根橡木，手工砍凿成方形的木梁。工匠们使用的斧头是根据12世纪末和13世纪初的样式手工锻造的。1月12日完工的弧形拱顶上的桁架是用木樨钉固定在一起，没有使用任何工业金属件。“我们希望使用的工具尽可能和中世纪的一样，从而让木材呈现那时的样子。”来自法国西部的Ateliers Perrault公司的木匠瓦伦汀·庞塔罗洛（Valentin Pontarollo）说。这种做法带来的乐趣之一是，大多时候“没有机器噪音，只有斧头的砍凿声”，他说。

在现场的修复车间内，雕刻家和石匠也在用凿子和刷子进行手工作业。地板上放置着一些接近完工的新制的石像怪。原物于19世纪添装到大教堂的外立面和塔楼上，这些面目狰狞的怪兽威严地俯视着下方。正在进行彻底

重建的还有十几个滴水嘴兽、80个花形装饰物、70个小型石像怪和750多个花形浮雕（装饰性石钩），以及教堂的三面山墙。去年12月，起重机把一只新做的镀金铜公鸡吊装到距地面96米的塔尖上。原来的那只公鸡在大火中掉落到了一座飞扶壁下的地面上，奇迹般地幸存下来，目前在一家博物馆展出。

修复工程不可避免地被一些现代主义者斥为守旧。火灾发生后，各路建筑师都在游说，希望采用大胆的当代装饰手法，比如用平板玻璃建造的尖塔或者种有植物的屋顶。毕竟，巴黎圣母院本来就融合了各种建筑形式，它曾先后在法国宗教战争和大革命期间遭受劫掠，是不同世纪叠加的产物。19世纪，维克多·雨果的小说《巴黎圣母院》重新唤起了世人对这件在当时本不受喜爱的哥特式杰作的敬意。也正是此后不久，维奥莱特-勒-杜克才新建了它的尖塔。

如今在一场公开辩论中，两派观点截然对立：一些人质疑修旧如旧的做法，另一些人则主张唯有尊重原样才能保真，这既包括保留它风格不相一致的部分，也包括保存其中持久稳定的部分。“它有别于其他任何古迹，不应该被修改失真，”约斯特表示，“我们正在修复一座有着860年历史的大教堂，以期它再存续至少860年。”

巴黎圣母院的管理者正试图找到一种平衡。这座刚刚清理干净、焕然一新的大教堂可能会让一些游客大吃一惊。中殿内的大部分脚手架已经拆除，柱子和拱顶的明黄色的石头令人叹为观止，仿佛回到了中世纪的样子——与它们之前黑魑魑、有些地方还坑坑洼洼的样子形成了惊人的对比。

不过巴黎圣母院并没有完全拒绝现代元素。巴黎大主教选用了两位当代法国设计师来设计家具。纪尧姆·巴代（Guillaume Bardet）正在制作包括圣坛在内的礼拜用的陈设，采用了深古铜色；约娜·沃特林（Ionna Vautrin）正在制作1500个实心橡木的会众椅，这些新座椅的椅背将采用板条来暗示开放和光亮。

巴黎圣母院的修复工程对于法国而言既是考验也是展示：要能在如此紧迫

的最后期限前完成任务，经受住来自全世界的宗教或非宗教人士的严苛审视，并一展法国的精工细作。按期完工事关“法国人的自尊自豪”，约斯特表示。

对于工匠们而言，完工却是令人酸楚的。埃米尔-阿尔芒·伯努瓦（Emile-Armand Benoit）在大教堂的最高处装饰屋顶，他说自己的团队正在加紧工作，这样脚手架就可以拆掉了，巴黎“就能再次看到高耸入云的尖塔”。在教堂的半高处，木匠们抬头凝视新完工的屋顶。“离开总归有些难过的，”庞塔罗洛说，“我们碰不到这样的活了。”■



AI for all

AI holds tantalising promise for the emerging world

It could help boost human capital, and ultimately growth

NEW TECHNOLOGY brings with it both the sweet hope of greater prosperity and the cruel fear of missing out. Satya Nadella, the boss of Microsoft, says he is haunted by the fact that the Industrial Revolution left behind India, his country of birth. (Indian manufacturers hardly enjoyed a level playing-field—Britain was then both their rival and their ruler.) Many technologies, such as online-education courses, have generated more hype than economic growth in the emerging world. Some people worry that generative artificial intelligence (AI), too, will disappoint the global south. The big winners so far seem to be a bunch of Western early adopters, as well as startups in San Francisco and America's "magnificent seven" tech firms, which include Microsoft and have together added an astonishing \$4.6trn to their market value since ChatGPT's launch in November 2022.

Yet AI stands to transform lives in the emerging world, too. As it spreads, the technology could raise productivity and shrink gaps in human capital faster than many before it. People in developing countries need not be passive recipients of AI, but can shape it to suit their own needs. Most exciting of all, it could help income levels catch up with those in the rich world.

The promise of AI in developing countries is tantalising. As in the West, it will be a useful all-purpose tool for consumers and workers, making it easier to obtain and interpret information. Some jobs will go, but new ones will be created. Because emerging countries have fewer white-collar workers, the disruption and the gain to existing firms may be smaller than in the West. The IMF says that a fifth to a quarter of workers there are most exposed to replacement, compared with a third in rich countries.

But a potentially transformative benefit may come from better and more accessible public services. Developing economies have long been held back by a lack of educated, healthy workers. Primary-school teachers in India have twice as many pupils as their American counterparts, but are ill-equipped for the struggle. Doctors in Africa are scarce; properly trained ones are scarcer. Whole generations of children grow up badly schooled, in poor health and unable to fulfil their potential in an increasingly global labour market.

Policymakers and entrepreneurs around the world are exploring ways that AI can help. India is combining large language models with speech-recognition software to enable illiterate farmers to ask a bot how to apply for government loans. Pupils in Kenya will soon be asking a chatbot questions about their homework, and the chatbot will be tweaking and improving its lessons in response. Researchers in Brazil are testing a medical AI that helps undertrained primary-care workers treat patients. Medical data collected worldwide and fed into AIs could help improve diagnosis. If AI can make people in poorer countries healthier and better educated, it should in time also help them catch up with the rich world.

Pleasingly, these benefits could spread faster than earlier waves of technology. New technologies invented in the early 20th century took more than 50 years to reach most countries. By contrast, AI will spread through the gadget that many people across the emerging world already have, and many more soon will: the phone in their pockets. In time, chatbots will become much cheaper to provide and acquire.

Moreover, the technology can be tailored to local needs. So far there is little sign that AI is ruled by the winner-takes-all effects that benefited America's social-media and internet-search firms. That means a variety of approaches could prosper. Some developers in India are already taking Western models and fine-tuning them with local data to provide a whizzy language-

translation service, avoiding the heavy capital costs of model-building.

Another idea that is also taking off in the West is to build smaller, cheaper models of your own. A narrower set of capabilities, rather than the ability to get every bit of information under the sun, can suit specific needs just fine. A medical AI is unlikely to need to generate amusing limericks in the style of William Shakespeare, as ChatGPT does so successfully. This still requires computing power and bespoke data sets. But it could help adapt AI in more varied and useful ways.

Some countries are already harnessing AI. China's prowess is second only to America's, thanks to its tech know-how and the deep pockets of its internet giants. India's outsourcing industry could be disrupted, as some back-office tasks are taken on by generative AI. But it is home to a vibrant startup scene, as well as millions of tech developers and a government that is keen to use AI to improve its digital infrastructure. These leave it well-placed to innovate and adapt. Countries in the Gulf, such as the United Arab Emirates and Saudi Arabia, are determined to build an AI industry as they shift from oil. They already have the capital and are importing the talent.

Each country will shape the technology in its own way. Chinese chatbots have been trained to keep off the subject of Xi Jinping; India's developers are focused on lowering language barriers; the Gulf is building an Arabic large language model. Though the global south will not dislodge America's crown, it could benefit widely from all this expertise.

| *Teaching AI*

Plenty could yet go wrong, obviously. The technology is still evolving. Computing power could become too expensive; local data will need to be gathered and stored. Some practitioners may lack the ability to take advantage of the knowledge at their fingertips, or the incentive to try new things. Although countries in sub-Saharan Africa stand to gain the most

from improvements to human capital and government services, the technology will spread more slowly there than elsewhere without better connectivity, governance and regulation.

The good news is that investments to speed AI's diffusion will be richly rewarded. Much about the AI revolution is still uncertain, but there is no doubt that the technology will have many uses and that it will only get better. Emerging countries have suffered disappointments before. This time they have a wonderful opportunity—and the power to seize it. ■



【首文】AI为人人

AI为新兴世界带来诱人前景

它有可能帮助提升人力资本，最终促进增长

新技术让人们心怀美好希望，期待更加繁荣，同时也让人心生恐惧，担心错失良机。出生于印度的微软首席执行官萨蒂亚·纳德拉（Satya Nadella）表示，印度没有赶上工业革命，这让他耿耿于怀。（当时印度制造商无法参与公平竞争，英国既是竞争对手也是它们的统治者。）在线教育等许多技术在新兴世界中引发的炒作大于实际推动的经济增长。一些人担心生成式AI也会让全球南方失望。迄今为止，最大的赢家似乎是一群西方的早期采用者，还有旧金山的创业公司和包括微软在内的美国科技“七雄”——在ChatGPT于2022年11月问世后，这七家公司的市值总计增加了惊人的4.6万亿美元。

但AI也势将改变新兴世界的生活。随着AI的普及，它可能会比之前的许多技术都更快地提高生产率和缩小人力资本差距。发展中国家的人们不必被动地接受AI，而可以按自己的需求塑造它。最令人兴奋的是，AI可以帮助发展中国家的收入水平赶上富裕世界。

AI给发展中国家带来的前景令人向往。与在西方一样，AI将成为帮助消费者和工人的多功能工具，让获取和解释信息更加容易。一些工作将会消失，但新的工作也会被创造出来。由于新兴国家的白领工人较少，AI带给现有企业的冲击和收益可能会小于在西方国家。国际货币基金组织（IMF）表示，这些国家有五分之一到四分之一的工人最容易被AI替代，而富裕国家的比例是三分之一。

但一个潜在的具变革性的益处可能来自公共服务的改善和更易获取。发展中经济体长期以来因为受过良好教育的健康劳动力不足而受制约。印度小学教师要带的学生人数是美国教师的两倍，但他们的能力不足以应对这一繁重任务。非洲急缺医生，受过正规培训的医生更是凤毛麟角。一代又一代孩子在成长时没有得到好的教育，健康状况不佳，无法在日益全球化的

劳动力市场中发挥潜力。

世界各地的政策制定者和企业家正在探索AI能够如何发挥作用。印度正在将大语言模型与语音识别软件结合起来，让不识字的农民能够向机器人咨询如何申请政府贷款。肯尼亚的学生将很快能够向聊天机器人提出有关作业的问题，而聊天机器人将相应调整和改进教学内容。巴西的研究人员正在测试一种医疗AI，辅助培训不足的初级医护人员治疗患者。从全球收集的医疗数据输入AI后可以帮助提高诊断水平。如果AI能让贫穷国家的人们更健康、接受更好的教育，那么假以时日也应该能帮助他们赶上富裕世界的生活水平。

可喜的是，这些好处可能会比之前的技术浪潮传播得更快。在20世纪初发明的新技术过了50多年才传播到大多数国家。相比之下，新兴世界中的许多人已经有了能够传播AI的设备，而且很快还会有更多人拥有这种设备——那就是他们兜里的手机。随着时间的推移，提供和购买聊天机器人的成本将大大降低。

此外，这项技术可以根据本地需求来定制。目前还没有迹象显示让美国社交媒体和互联网搜索公司受益的赢家通吃效应会支配AI的发展。这意味着各种不同的路径都有可能获得成功。印度的一些开发者已经开始把西方的模型拿来，用本地数据做优化，以提供先进的语言翻译服务，避免了建模的巨大资本成本。

另一个也开始在西方流行的思路是建造更小、更便宜的自有模型。只需开发小范围的功能，而无需拥有获得天底下所有信息的能力，将能足够应付特定的需求。医疗AI可能并不需要像ChatGPT那样能够生成莎士比亚风格的风趣打油诗。这仍然需要计算能力和定制数据集。但它可以帮助调整AI，满足更多样化、更有用的需求。

一些国家已经在利用AI。中国得益于其科技水平和互联网巨头的雄厚财力，在这方面的实力仅次于美国。印度的外包行业可能会受到影响，因为一些后台任务已经由生成式AI承担。但印度拥有充满活力的创业圈子、千

百万技术开发人员和一个热衷于利用AI改善数字基础设施的政府，这些都为它推进创新和调整适应提供了有利条件。阿联酋和沙特阿拉伯等海湾国家决心在摆脱对石油的依赖之时建立AI产业。它们已经拥有所需的资本并且正在引进人才。

每个国家都会以自己的方式塑造这项技术。中国的聊天机器人已经经过训练来避开最高领导人的话题，印度的开发者致力于降低语言障碍，海湾地区正在构建阿拉伯语的大语言模型。全球南方不会动摇美国的AI王冠，但可以从所有这些专门技术中广泛受益。

| 教学助理

显然，有很多方面可能出错。这项技术仍然在不断发展演变。算力可能变得成本过高；将需要收集和存储本地数据。一些从业者可能缺乏利用手边知识的能力，或者缺乏尝试新事物的动力。尽管撒哈拉以南非洲应该会从人力资本和政府服务的提升上获益最大，但在连通性、治理和监管没有改善的情况下，这项技术在那里的传播会比在其他地方更慢。

好消息是，加速AI传播的投资将获得丰厚的回报。AI革命的许多方面仍不确定，但毫无疑问的是，这项技术将有很多用途，而且只会越变越好。新兴国家之前历经失望。这一次它们有绝佳的机会，并有能力抓住它。■



Private assets, public interest

The risks to global finance from private equity's insurance binge

Funding pensions with private assets holds promise—but needs scrutiny

A DECADE OR so ago private equity was a niche corner of finance; today it is a vast enterprise in its own right. Having grabbed business and prestige from banks, private-equity firms manage \$12trn of assets globally, are worth more than \$500bn on America's stockmarket and have their pick of Wall Street's top talent. Whereas America's listed banks are worth little more than they were before the pandemic, its listed private-equity firms are worth about twice as much. The biggest, Blackstone, is more valuable than either Goldman Sachs or Morgan Stanley—and has the confidence of a winner. "It's the alternatives era," proclaimed the company's ebullient Taylor Swift-themed festive video in December. "We buy assets then we make 'em better."

This is not, though, the business that has recently boomed for them. Traditional private equity—using lots of debt to buy companies, improving them, and selling or listing them—has been lifeless. High interest rates have cast doubt on the value of privately held companies and reduced investors' willingness to provide new funds. It does not seem to matter. Core private-equity activity is now just one part of the industry's terrain, which includes infrastructure, property and loans made directly to companies, all under the broad label of "private assets". Here the empire-building continues. Most recently, the industry is swallowing up life insurers.

All of the three kings of private equity—Apollo, Blackstone and KKR—have bought insurers or taken minority stakes in them in exchange for managing their assets. Smaller firms are following suit. The insurers are not portfolio investments, destined to be sold for a profit. Instead they are prized for

their vast balance-sheets, which are a new source of funding.

Judged by the fundamentals, the strategy makes sense. Insurance firms invest over long periods to fund payouts, including annuities sold to pensioners. They have traditionally bought lots of government and corporate bonds that are traded on public markets. Firms like Apollo can instead knowledgeably move their portfolios into the higher-yielding private investments in which they specialise. A higher rate of return should mean a better deal for customers. And because insurers' liabilities stretch years into the future, the finance they provide is patient. In banking, long-term loans are funded with lots of instantly accessible deposits; with private assets and insurance, the duration of the assets matches the duration of the liabilities.

Yet the strategy brings risks—and not just to the firms. Pension promises matter to society. Implicitly or explicitly, the taxpayer backstops insurance to some degree, and regulators enforce minimum capital requirements so that insurers can withstand losses. Yet judging the safety-buffers of a firm stuffed with illiquid private assets is hard, because its losses are not apparent from movements in financial markets. And in a crisis insurance policyholders may sometimes flee as they seek to get out some of their money even if that entails a financial penalty. Last year an Italian insurer suffered just such a bank-run-like meltdown.

Making things harder is the complexity of the tie-ups, which involve labyrinthine interlinkages between different bits of firms' balance-sheets. Much reinsurance activity takes place in Bermuda, an offshore hub where there is more than a whiff of regulatory arbitrage. Yet compared with the zealots who police the global banking system, insurance regulators are docile.

As private assets become more important, that must change. Regulators

should co-operate internationally to ensure that the safety-buffers are adequate. High standards of transparency and capital need to be enforced by suitably heavyweight bodies. The goal should not be to crush a new business model, but to make it safer. Financial innovation often brings new benefits even as it creates new ways to blow up the system. Regulators would be making a mistake to ignore either edge of the sword. ■



【首文】私募资产，公共利益

私募股权大举收购保险公司给全球金融带来风险

用私募资产为养老金融前景良好——但需要严格审视

大约十年前，私募股权还是金融业的一个利基领域，如今它已自成一门庞大的生意。私募股权公司从银行手中抢走了业务和声誉，在全球管理着12万亿美元的资产，在美国股市的市值超过5000亿美元，并拥有华尔街最顶尖的人才。美国上市银行的市值与疫情暴发前相比仅略有增长，而美国上市私募股权公司的市值是疫情前的两倍。其中最大的黑石集团

（Blackstone）的市值比高盛或摩根士丹利都要高，并且拥有赢家的自信。“这是另类投资的时代，”该公司在12月发布的一则热情洋溢的模仿泰勒·斯威夫特巡演的节日视频里宣称，“我们购买资产，然后把它们变得更好。”

然而，他们近期十分红火的业务并不是购买资产。用大量债务购买公司，再改善它们，然后将它们出售或上市的传统私募股权投资已经了无生气。高利率让人对私人公司的价值产生了怀疑，也降低了投资者提供新资金的意愿。这似乎并不要紧。核心私募股权投资活动如今只是该行业版图的一部分，其他业务还包括基础设施、房地产和直接向企业放贷，这些都被归在“私募资产”这个大标签之下。在这方面，帝国的建设仍在继续。最近，该行业正在大举收购人寿保险公司。

私募股权投资三巨头阿波罗、黑石和KKR都收购了保险公司或持有少数股权，以换取对其资产的管理。规模较小的公司也纷纷效仿。保险公司不是投资组合，注定会为了获利而被售出。它们备受青睐其实是因为有庞大的资产负债表，可以成为新的资金来源。

从基本面来看，这一战略是合理的。保险公司进行的是长期投资，为各种赔付提供资金，包括向养老金领取者出售的年金。传统上，它们大量购买在公开市场交易的政府和公司债券。而像阿波罗这样的公司则可以游刃有余地将其投资组合转移到收益更高的私人投资领域——这是它们的专长。

回报率更高，对客户来说应该也就越划得来。而且，由于保险公司的债务会延续到未来数年，它们提供的融资是有耐心的。在银行业，长期贷款的资金源自大量即时可用的存款，而在私募资产和保险业，资产的期限与债务的期限是相匹配的。

但这种策略也带来了风险，而且不仅仅是对公司而言。养老金的承诺对社会很重要。纳税人在某种程度上直接或间接地为保险提供了支持，监管机构则强制执行最低资本要求，使保险公司能够扛住损失。然而，要判断一家拥有大量低流动性私募资产的公司的安全缓冲能力如何是很难的，因为它的损失并不能从金融市场的变动中清楚地显现出来。而在危机中，投保人有时可能会闻风而逃，设法拿回一部分钱，即使这样做会造成财务损失。去年，一家意大利保险公司就发生了类似银行挤兑的爆雷事件。

这类并购的复杂性让事情更加麻烦，因为公司资产负债表的不同部分之间会生成错综复杂的相互关联。许多再保险活动都在百慕大进行，在这个离岸中心，监管套利并不少见。但与监管全球银行系统的积极分子们比起来，保险业的监管者显得低眉顺眼。

随着私募资产变得越发重要，这种情况必须改变。监管机构应开展国际合作，确保企业有足够的安全缓冲。关于透明度和资本的高标准还需要具备足够影响力的机构来执行。其目标不应是压垮一个新的商业模式，而是让它更安全。尽管金融创新会变出新花样来破坏金融体系，但往往也会带来新的益处。监管机构忽视这把双刃剑的任何一面都将是犯错。■



A slippery concept

Common sense is not actually very common

Very few claims meet with universal agreement

IN 1776 THOMAS PAINE, a traitorous Englishman living in the American colonies, published a seditious 47-page pamphlet. Called “Common Sense”, it became a best-seller. It argued that the colonies should seek independence from British rule. Later that year they did exactly that.

Appeals to common sense are a staple of politics, especially when an insurgent wishes to distinguish himself from a supposedly aloof and out-of-touch elite. But in a paper published in Proceedings of the National Academy of Sciences, Mark Whiting and Duncan Watts, a pair of computational social scientists at the University of Pennsylvania, note that the idea has seldom been rigorously studied.

The two researchers set out to fix that. They started by noting that the standard concept of common sense has a somewhat circular definition: common sense is a set of claims that sensible people agree with, and sensible people are those who possess common sense.

To get around such philosophical tangles, the researchers turned to Mechanical Turk, a website run by Amazon, a big tech firm, that allows people to post odd jobs. They recruited 2,046 human participants and asked them to rate 50 statements from a corpus of 4,407 claims that might plausibly be seen as commonsensical.

As common sense might have predicted, the researchers found that plainly worded claims concerning facts about the real world were the most likely to be rated as demonstrating common sense (“triangles have three sides”, for example, which is true by definition, or “avoid close contact with people

who are ill”). The more abstract the claims, the less likely participants were to agree that they were common sense (“all human beings are created equal”; “perception is the only source of knowledge”).

When they split the claims by subject, the researchers found that those concerning technology and science were the most likely to be rated as commonsensical, while matters of history and philosophy were the least likely. A respondent’s age, sex, income and personal politics had little effect on what they thought counted as common sense, although psychological measures of social perceptiveness and the ability to reflect on one’s opinions did.

Having investigated individual opinions, the researchers looked at how common sense works across big groups. Here, they found much less agreement than might have been expected. Only around 44% of claims in the corpus were rated as commonsensical by at least 75% of respondents. A stricter definition of common sense, in which everyone has to agree with a claim for it to count, cut that number to just 6.6%. Where exactly a sensible cut-off lies is a matter for debate. But truly “common” sense, it seems, is an elusive thing. ■



一个含糊的概念

常识其实不太常见

很少有说法得到普遍认同

一七七六年，生活在美洲殖民地的叛徒英国人托马斯·潘恩（Thomas Paine）出版了一本煽动性的47页小册子。这本名为《常识》（Common Sense）的小册子成了畅销书。它主张殖民地应该摆脱英国的统治，寻求独立。那一年早些时候，它们确实做到了。

诉诸常识是政治的一个常见工具，特别是当一名叛乱者希望将自己与那些被认为是高高在上、脱离群众的精英区分开来时。但是在《美国国家科学院院刊》（Proceedings of the National Academy of Sciences）上发表的一篇论文中，宾夕法尼亚大学的两位计算社会科学家马克·惠廷（Mark Whiting）和邓肯·沃茨（Duncan Watts）指出，常识这个概念很少被严谨地研究。

这两位研究人员着手解决这个问题。他们首先指出，常识的标准概念有几分循环定义的味道：常识是理智人群认同的一系列主张，而理智人群则是那些有常识的人。

为了解开这团哲学上的乱麻，研究人员使用了Mechanical Turk，这家网站由大型科技公司亚马逊运营，人们可以在上面发布各种零工杂活。他们招募了2046人参与研究，让他们给50个陈述评分，它们来自一个有4407个可能被视为常识的语料库。

正如依据常识能预见到的那样，研究人员发现，关于现实世界事实的直白陈述最有可能被评断为常识（例如“三角形有三条边”这种顾名思义而正确的表述，或者“避免与生病的人密切接触”）。表述越抽象，参与者认为它们是常识的可能性就越小（“所有人生而平等”，“感知是知识的唯一来源”）。

在按主题划分这些表述时，研究人员发现，涉及技术和科学的陈述最有可能被评断为常识，而历史和哲学问题最不可能。受访者的年龄、性别、收入和个人政治观点对他们认为什么算作常识没什么影响，尽管由心理评估反映的社会洞察力和对自身观点的反思能力确有影响。

在研究了个人观点之后，研究人员又调查了群体对常识的认知。他们发现人群对于什么算常识的一致认同度要比大家原本可能以为的低得多。语料库中只有大约44%的表述被至少75%的受访者认为是常识。如果对常识做更严格的定义，即在每个人都同意的情况才能说一种表述是常识，那么这一数字将降至6.6%。把分界线划在哪里才合理是一个有待讨论的问题。但看起来，真正的“常”识近乎虚无缥缈。 ■



Live music

Does Las Vegas's Sphere reveal the future of concerts?

The venue is dazzling. But copycats are unlikely to be built soon

NESTLED BETWEEN hotels and conference centres, a short walk from the Las Vegas strip, is a giant, wide-eyed emoji. Sometimes it is an enormous, hyperrealistic eyeball, a basketball or a whorl of flames. The Sphere, a remarkable new concert venue, is 366 feet (110 metres) tall and 516 wide; an LED screen spanning almost 600,000 square feet covers the exterior.

Inside, enveloping the 17,500 seats, is another vast, ultra-high-resolution screen. This pleasuredome offers an experience unlike any other. It also raises questions about the future of live entertainment.

The high-tech arena was opened in September by U2, who remain in residence until March. The Irish band has a history of innovative concert design as well as corporate ventures, including a long partnership with Apple. Perhaps just as important, U2 is loved by middle-aged rock fans, who form the bulk of concertgoers in America and might shell out for a ticket. (Prices start at \$140 and go up to \$1,500.)

A celebration of "Achtung Baby", an album released in 1991, U2's show is an overwhelming phantasmagoria. The band plays on a stage shaped like a turntable while work by artists including Es Devlin and John Gerrard appears on the vaulting screen. During "Even Better Than the Real Thing", there are animations of slot machines and Elvis kitsch. For "Where the Streets Have No Name", the crowd is transported into the blinding daylight of the high desert. During "With Or Without You" the ceiling teems with images of endangered local species such as the leaf-nosed bat. "It's a show that's about the venue that it's in," says Willie Williams, U2's longtime production designer. "It's about us all going to experience the venue."

From one angle, the Sphere represents a major development in an existing trend of the arts becoming more immersive. Exhibitions that turn the paintings of Monet or Van Gogh into interactive, room-encompassing installations have proliferated in recent years, as have immersive theatre productions. Virtual-reality technology has improved significantly.

From another angle, it is part of the evolution of the modern rock concert, which since the 1960s has combined light and sound to transport fans into another dimension. At a show your correspondent attended at the Sphere in October, there were moments when the experience was transcendent. (There is a quiet room, filled with bean bags, for the over-stimulated.) Yet viewers still took their eyes off the big screen to reach for the smaller ones in their pockets. At times the band seemed like an expensive soundtrack to the bright lights.

Is this the future of the concert? In the short term, no. The sheer cost of the Sphere—\$2.3bn—means that the model cannot be easily reproduced. Its ostentation is also a barrier: Sadiq Khan, the mayor of London, recently vetoed a sister Sphere in the city, calling it “bulky, unduly dominant and incongruous”. Sphere Entertainment Company, the owner, hopes to build other iterations and is in “serious” talks for an arena in Abu Dhabi. But negotiations regarding Spheres in Saudi Arabia and South Korea have stalled.

Some artists and promoters are said to be wary of developing shows that cannot be taken on tour to other, standard arenas and of letting the venue outshine the music. For now, at least, what happens in Vegas is staying there.

The Sphere does mark a bullish bet on the future of live music, however. The biggest acts have long had to make do with sports stadiums with dodgy acoustics, but this is a capacious, purpose-built venue. There seems to be

plenty of demand for lavish productions by the biggest hitmakers: witness the billions of dollars in revenue made by Beyoncé's "Renaissance" and Taylor Swift's "Eras" tours.

According to Luminate, an analytics firm, in 2023 consumers spent 91% more on live music events than the year before and attended 32% more concerts. Goldman Sachs predicts the market for live music will grow by 5% this year to reach nearly \$40bn annually by 2030.

This growth is not just driven by pent-up demand from the pandemic. Youngsters, who prefer to spend their money on experiences than on items, consider concerts good value, even when they are pricey. "People still want to have that experience of liveness," says Steve Waksman, a concert historian, regardless of whether, as at the Sphere, it is "mediated" through screens. To some, the Sphere may be more bewildering than beautiful, but one thing is clear: the future of concerts is as rosy as Bono's trademark glasses. ■



现场音乐

拉斯维加斯的Sphere揭幕了演唱会的未来吗？

这个场馆令人眼花缭乱。但复刻品不太可能很快建造出来

距离拉斯维加斯大道步行不远，林立的酒店和会议中心之间有一个巨大的、瞪着大眼睛的表情符号。有时它会变成一个巨大而逼真的眼球、一颗篮球或一团火焰。这是一个名为Sphere的新奇演唱会场馆，高366英尺（110米），宽516英尺；外表覆盖着近60万平方英尺的LED屏幕。

场馆内有17,500个座位，顶棚又是一个超高分辨率的巨大屏幕。这个富丽堂皇的欢乐王国提供了一种绝无仅有的体验，也提出了关于现场娱乐未来发展的问題。

这个高科技场馆的首演是U2在去年9月举办的演唱会，他们在那里的驻场演出将一直持续到今年3月。这支爱尔兰乐队向来热衷创新演唱会设计及企业创投，包括与苹果的长期合作。也许同样重要的是，U2受到中年摇滚乐迷的喜爱，而这些人构成了美国演唱会观众的主体，可能为价格不菲的门票掏腰包。（票价140美元起，最高达1500美元。）

U2此次演唱会是对1991年发布的专辑《Achtung Baby》的纪念演绎，整个演出被打造成一个壮丽幻境。乐队在一个形似唱机转盘的舞台上表演，而艾斯·德芙琳（Es Devlin）和约翰·杰拉德（John Gerrard）等艺术家的作品出现在头顶的屏幕上。在演唱《Even Better Than the Real Thing》时，屏幕上老虎机动画和猫王俗艳风格的影像。到《Where the Streets Have No Name》时，观众被带入到高地沙漠的耀眼日光中。等到《With Or Without You》，顶棚屏幕上则满是叶鼻蝠等濒危本地物种的图像。“这场演出也是为了展现它的演出场馆，”U2的长期制作设计师威利·威廉姆斯（Willie Williams）说，“要的就是让大家都能体验这个场馆。”

从一个角度来看，在艺术变得更加沉浸式的现行趋势中，Sphere代表了一个重大发展。近些年来，将莫奈或梵高的画作变成环绕整个房间的交互式

装置的展览层出不穷。沉浸式戏剧制作也是如此。虚拟现实技术已经有了长足进步。

从另一个角度来看，它是现代摇滚演唱会演变的一部分。自1960年代起，摇滚演唱会就结合了灯光和音效来将粉丝带入另一个维度。去年10月笔者观看了在Sphere举行的一场演出，在某些时刻会有超然的体验。（那里有一个安静的房间，里面放满了豆袋沙发，专供太过兴奋的人回复平静。）不过，观众们仍然会把目光从大屏幕上移开，伸手去摸口袋里的小屏幕。有时候，乐队似乎就是绚丽灯光的昂贵配乐。

这是演唱会的未来吗？从短期来看，不是。Sphere那23亿美元的巨额造价让这种模式不易被复制。它的张扬也是一个障碍：伦敦市长萨迪克·汗（Sadiq Khan）最近否决了该市的一个Sphere姊妹项目，称它“笨重、过分突出且不协调”。Sphere的所有方Sphere娱乐公司（Sphere Entertainment Company）希望建造其他迭代版本，目前正就在阿布扎比建造一个场馆展开“严肃”谈判。但关于在沙特阿拉伯和韩国建造Sphere的谈判已经搁浅。

据称有些艺术家和推广公司对于制作无法在其他标准场馆巡演的表演持谨慎态度，也担心场馆会抢走音乐的风头。至少目前来看，发生在拉斯维加斯的还是会留在拉斯维加斯。

然而，Sphere确实对现场音乐的未来发展做出了大胆押注。长期以来，最大型的音乐活动只能将就着在音响效果不佳的体育场馆中举办，但Sphere是一个专为这类表演打造的大容量场馆。对于顶流音乐人的豪华制作似乎需求很大，碧昂丝的Renaissance巡演和泰勒·斯威夫特的Eras巡演创造出的数十亿美元收入就是例证。

据分析公司Luminate统计，2023年消费者在现场音乐活动上的支出比前一年增加了91%，观看演唱会的次数增加了32%。高盛预测，现场音乐市场今年将增长5%，到2030年将达成近400亿美元的年收入。

这一增长不仅仅是因为因疫情而受抑制的需求得到了释放。年轻人更愿意把钱花在体验而不是实物上，即使演唱会票价昂贵，他们也认为这物有所

值。“人们仍然渴望那种现场体验”，研究演唱会历史的史蒂夫·沃克斯曼（Steve Waksman）表示，无论是否像在Sphere那样通过屏幕来“传递”。对一些人来说，Sphere打造的可能更多是迷乱而非美感，但有一点是明确的：演唱会的未来就像U2主唱博诺（Bono）那标志性的粉紫色墨镜一般瑰丽。 ■



Schumpeter

Why BlackRock is betting billions on infrastructure

Demand for investment is soaring thanks to decarbonisation, digitisation and deglobalisation

THE GLOBAL economy is on the cusp of an “infrastructure revolution”, if Larry Fink is to be believed. The boss of BlackRock, the world’s largest asset manager, made the modest prediction shortly after announcing on January 12th that his firm would acquire Global Infrastructure Partners (GIP) for \$12.5bn. That company, led by Adebayo Ogunlesi, an old pal of Mr Fink’s from their banking days, is the world’s third-largest infrastructure investor, behind Australia’s Macquarie and Canada’s Brookfield. Its assets range from Gatwick Airport in London to the Port of Melbourne. Mr Ogunlesi and his fellow partners will collectively become BlackRock’s second-largest shareholder.

Mr Fink is not the only one excited about the industry. On January 16th General Atlantic, a private-equity (PE) firm, confirmed reports that it would buy Actis, an infrastructure investor focused on emerging markets. In September CVC, another PE firm, announced it was buying DIF, a Dutch infrastructure investor. Over the past decade assets under management in infrastructure funds have increased almost five-fold, to \$1.3trn, according to Preqin, a data provider. Pension funds and sovereign-wealth managers have been lured in by the industry’s returns, which are both handsome and relatively stable. More than half of such backers surveyed by Preqin intend to increase the share of their portfolios allocated to infrastructure. Some of the larger among them now invest directly in these dull assets. Why, then, all the excitement?

The infrastructure-investment business took shape in the 1990s and 2000s. Western governments with growing debts began seeking out private

investors to acquire—and help rejuvenate—ageing infrastructure from airports and railways to water pipes. Later, a growing assortment of companies from energy suppliers to telecoms operators also turned to infrastructure investors to offload assets such as pipelines and cell towers, observes Sam Pollock, boss of Brookfield's infrastructure business.

Now demand for infrastructure investment is soaring thanks to three megatrends, explains Mr Pollock. The first is decarbonisation. For the world to meet its climate goals, some \$8trn will need to be invested over the remainder of this decade in renewable energy such as solar and wind, as well as batteries to store it and transmission lines to transport it. Hefty investments will also be needed in hydrogen facilities, to produce carbon-free fuel for planes and ships, and in carbon removal. The second megatrend is digitisation. Software may well be eating the world, as a venture capitalist once predicted, but it is relying on an awful lot of physical assets to do it, from fibre-optic cables and 5G networks to data centres. Third, deglobalisation. Efforts to shift supply chains away from China are spurring demand for capital-hungry factories and new transport infrastructure to move goods over land and sea. In Europe concerns about energy security following Russia's invasion of Ukraine have also provoked a rush to build liquefied-natural-gas terminals to bring in the fuel from less belligerent places.

All that demand for investment is arriving at a time when government and corporate balance-sheets are under strain. America's \$26trn (98% of GDP) pile of federal-government debt is expected to continue expanding over the coming decade. Many governments in Europe also have weighty debt burdens. Higher interest rates are making those liabilities more expensive to service. They are also making life awkward for companies which have gorged on cheap debt to juice shareholder returns. The need to deleverage will limit their ability to make big investments in the years ahead.

Infrastructure investors are ready and willing to fill the gap. In 2022 Intel, a

big chipmaker, turned to Brookfield to fund 49% of a new \$30bn chip factory in America.

So far most infrastructure investors have concentrated on rich countries, where governments are more dependable and currencies more stable. More than four-fifths of assets under management in the industry are allocated to Western markets, according to data from Preqin. At the same time, the need for new infrastructure is most pronounced in the global south, where both populations and economies are growing faster. “Emerging-market investment is a big opportunity for us,” says Raj Rao, one of GIP’s co-founders. Leigh Harrison, who leads infrastructure investing at Macquarie, notes that his firm is increasing the share of its funds it allocates to such markets.

The industry, then, looks set to become increasingly important to the global economy. Yet it is not without its detractors. In Britain Macquarie has been criticised for its stewardship of Thames Water, which manages the water supply of London and its surrounds. During its ownership of the utility from 2006 to 2017, Macquarie tripled the company’s debts, to £11bn (\$14bn), helping to deliver a hefty return for itself and fellow shareholders. Since then the utility, weighed down by those debts, has struggled to afford necessary investments in fixing leaky pipes and reducing the sewage it pumps into rivers. Mr Harrison counters that £1bn a year was invested in the company during Macquarie’s tenure as its owner, more than in any previous period. Still, he concedes that “markets were very different” when it bought the business, and that his firm no longer loads its assets with debt to the same extent.

| *From spreadsheets to hard hats*

In a world of pricier debt, the way infrastructure investors make money is shifting from financial engineering to cleverer management of assets. Mr Harrison notes that Macquarie is bulking up the number of industry

experts in its team. “Where we really add value is when we bring greater operational rigour to an asset,” says Mr Rao of GIP. He offers the example of Gatwick, where GIP has focused on speeding up security screening, leaving travellers with more time to relax—and indulge in some pre-flight shopping. For the infrastructure firms, merely shopping around for assets is increasingly an indulgence, too. ■



熊彼特

为什么贝莱德在基建上押注上百亿美元

脱碳、数字化和去全球化导致对投资的需求飙升

如果拉里·芬克（Larry Fink）的话可信的话，那么全球经济正处于一场“基础设施革命”的前沿。这位全球最大资产管理公司贝莱德（BlackRock）的老板在1月12日宣布将以125亿美元收购全球基础设施投资公司（Global Infrastructure Partners，简称GIP），之后不久做出了这个温和的预测。由芬克在银行工作时的老朋友阿德巴约·奥贡莱西（Adebayo Ogunlesi）领导的GIP是全球第三大基建投资公司，仅次于澳大利亚的麦格理（Macquarie）和加拿大的布鲁克菲尔德（Brookfield）。从伦敦盖特威克机场（Gatwick Airport）到墨尔本港（Port of Melbourne），它的资产遍布全球。奥贡莱西和他的几个合伙人将共同成为贝莱德的第二大股东。

芬克并不是唯一一个对基建投资兴致浓厚的人。1月16日，私募股权公司General Atlantic证实了有关它将收购新兴市场基建投资公司Actis的报道。去年9月，另一家私募股权公司CVC宣布将收购荷兰基建投资公司DIF。据数据供应商Prequin称，过去十年里，基建基金管理的资产增长了近四倍，达到1.3万亿美元。养老基金和主权财富管理公司被这个行业既丰厚又相对稳定的回报所吸引。根据Prequin的调查，超过一半的此类投资机构计划增加其投资组合中分配给基建的份额。其中一些较大的机构已经在直接投资这些一点也不光鲜亮丽的资产。那么，对基建的兴趣为什么如此浓厚呢？

基建投资业务在上世纪90年代和2000年代初具规模。当时西方政府的负债不断增长，它们开始寻求私人投资者收购并帮助翻新从机场和铁路到水管等老化的基础设施。后来，从能源供应商到电信运营商等越来越多的公司也开始找基建投资公司出手管道和手机信号塔等资产，布鲁克菲尔德的基建业务负责人萨姆·波洛克（Sam Pollock）说。

如今有三个超大趋势导致了基建投资需求的飙升，波洛克解释说。首先是脱碳。为了实现全球气候目标，到2030年前需要向可再生能源（如太阳能

和风能）、储能电池和输电线路投资约8万亿美元。飞机和船用无碳燃料的氢能生产设施以及碳移除也需要巨额投资。第二个大趋势是数字化。正如一位风投家曾经预测的那样，软件可能确实在吞噬世界，但这需要通过光纤电缆、5G网络和数据中心等大量有形资产来实现。第三个趋势是去全球化。将供应链移出中国的努力推动了对工厂和新的陆海运输基础设施的需求，这些都需要大量资本。俄罗斯入侵乌克兰后，欧洲对能源安全的担忧也引发了快速建设液化天然气终端的热潮，以从不易燃起战火的地区进口燃料。

所有这些投资需求都发生在政府和企业的资产负债表承压之时。美国联邦政府的债务达到26万亿美元（相当于GDP的98%），预计未来十年将继续增长。欧洲许多政府的债务负担也非常沉重。利率上升让偿债成本变得更高。那些在过去大举借入低息贷款来提高股东回报的公司日子也不好过了。去杠杆化的需要将使这些公司在未来几年内难以大规模投资。基建投资公司能够并且也愿意填补这一缺口。2022年，芯片巨头英特尔向布鲁克菲尔德公司寻求融资，为自己将在美国投资300亿美元新建的芯片工厂提供49%的资金。

迄今为止，大多数基建投资公司都聚焦于政府更可靠、货币更稳定的富裕国家。根据Preqin的数据，基建投资行业管理的资产超过五分之四都配置在西方市场。与此同时，对新基础设施的需求在人口和经济增速都更快的全球南方地区最为凸显。GIP的联合创始人之一拉吉·拉奥（Raj Rao）表示：“投资新兴市场对我们来说是一个巨大的机会。”麦格理的基建投资主管利·哈里森（Leigh Harrison）指出，他的公司正在增加对这些市场的资金配比。

因此，看起来这个行业对全球经济的影响必将与日俱增。但对该行业也不乏指摘之声。在英国，麦格理对负责伦敦及周边地区供水的泰晤士水务公司（Thames Water）的管理饱受批评。在2006年至2017年间，在麦格理的管理下，该公司的债务增加了两倍，达到110亿英镑（140亿美元），为麦格理和其他股东带来了可观的回报。此后，由于债务负担沉重，这家水务公司难以拿出必要的投资修复漏损的管道和减少向河流排放污水。哈里森

反驳说，麦格理在其管理期内每年向该公司投资10亿英镑，投资力度比以往任何时期都大。不过他也承认，在麦格理收购该公司时“市场情况大不相同”，而且如今麦格理也不再让旗下资产有那么高的负债。

| 从电子表格到安全帽

在债务成本更高的大环境下，基建投资公司赚钱的方式正在从金融工程转向更巧妙的资产管理。哈里森指出，麦格理正在团队中加入更多的行业专家。“我们真正增加价值的地方是提升资产运营的严谨度。”GIP的拉奥表示。他以盖特威克机场为例，GIP着重加快安检速度，让旅客能有更多时间放松——也能在登机前纵情购物。而对于基建投资公司来说，只记着四下拣选划算的资产也日益是一种奢侈放纵了。■



Pics and it didn't happen

AI-generated content is raising the value of trust

Who did the posting will soon matter more than what was posted

IT IS NOW possible to generate fake but realistic content with little more than the click of a mouse. This can be fun: a TikTok account on which—among other things—an artificial Tom Cruise wearing a purple robe sings “Tiny Dancer” to (the real) Paris Hilton holding a toy dog has attracted 5.1m followers. It is also a profound change in societies that have long regarded images, video and audio as close to ironclad proof that something is real. Phone scammers now need just ten seconds of audio to mimic the voices of loved ones in distress; rogue AI-generated Tom Hankses and Taylor Swifts endorse dodgy products online, and fake videos of politicians are proliferating.

The fundamental problem is an old one. From the printing press to the internet, new technologies have often made it easier to spread untruths or impersonate the trustworthy. Typically, humans have used shortcuts to sniff out foul play: one too many spelling mistakes suggests an email might be a phishing attack, for example. Most recently, AI-generated images of people have often been betrayed by their strangely rendered hands; fake video and audio can sometimes be out of sync. Implausible content now immediately raises suspicion among those who know what AI is capable of doing.

The trouble is that the fakes are rapidly getting harder to spot. AI is improving all the time, as computing power and training data become more abundant. Could AI-powered fake-detection software, built into web browsers, identify computer-generated content? Sadly not. The arms race between generation and detection favours the forger. Eventually AI models will probably be able to produce pixel-perfect counterfeits—digital clones

of what a genuine recording of an event would have looked like, had it happened. Even the best detection system would have no crack to find and no ledge to grasp. Models run by regulated companies can be forced to include a watermark, but that would not affect scammers wielding open-source models, which fraudsters can tweak and run at home on their laptops.

Dystopian possibilities abound. It will be difficult, for example, to avoid a world in which any photograph of a person can be made pornographic by someone using an open-source model in their basement, then used for blackmail—a tactic the FBI has already warned about. Perhaps anyone will be able to produce a video of a president or prime minister announcing a nuclear first strike, momentarily setting the world on edge. Fraudsters impersonating relatives will prosper.

Yet societies will also adapt to the fakers. People will learn that images, audio or video of something do not prove that it happened, any more than a drawing of it does (the era of open-source intelligence, in which information can be reliably crowdsourced, may be short-lived). Online content will no longer verify itself, so who posted something will become as important as what was posted. Assuming trustworthy sources can continue to identify themselves securely—via URLs, email addresses and social-media platforms—reputation and provenance will become more important than ever.

It may sound strange, but this was true for most of history. The era of trusted, mass-produced content was the exception. The fact that people may soon struggle to spot the invisible hand of AI does not mean the marketplace of ideas is doomed. In time, the fakes that thrive will mostly be the funny ones. ■



【首文】有图无真相

人工智能生成的内容正在提升信誉的价值

很快，谁发布了内容将比发布了什么内容更重要

现在只需轻点鼠标，就能生成虚假但逼真的内容。可能会是好玩的东西，比如一个TikTok账号下五花八门的视频中，有一则是人工合成的汤姆·克鲁斯穿着紫色长袍，对着抱着一只小狗狗的（真人）帕丽斯·希尔顿唱《小小舞者》（Tiny Dancer），该账号已经吸引了510万粉丝关注。这也是一个深刻的社会变化：在此之前，人们长久以来都认为图像、视频和音频几乎等同于某物真实存在的铁证。现在，搞电话诈骗的人只需要十秒钟的音频就能假扮某人的亲人求救，人工智能（AI）生成的汤姆·汉克斯和泰勒·斯威夫特在网上为可疑产品代言，有政客露脸的虚假视频也是层出不穷。

根本的问题其实是个老问题。从印刷术到互联网，新技术往往使传播不实信息或冒充可信者变得更加容易。通常情况下，人类会抓住某些破绽来识别猫腻，比如，如果拼写错误太多，就表明某封电子邮件可能是网络钓鱼攻击。至于最近，AI生成的人物图像往往会因为奇怪的手部细节被识破，伪造的视频和音频有时则会出现音画不同步的问题。现在，有悖于常理的内容马上就会引起那些清楚AI本领的人的怀疑。

麻烦在于，识别虚假的速度越发难以跟上造假的速度。随着计算能力和训练数据日益充裕，AI在不断进步。网页浏览器中内置的虚假内容检测AI软件能否识别计算机生成的内容？可惜不能。生成与检测之间的军备竞赛更有利于伪造者。最终，AI模型很可能能够生成像素级还原的赝品——假如它所呈现的事情真会发生，被记录下来的话也就会是这副样子。即便最好的检测系统也可能找不出破绽，无从下手。可以强制受监管的公司在运营的模式里加入水印，但这并不会影响使用开源模型的骗子们，他们在家里的笔记本电脑上就能调整和运行这些模型。

可能出现的反乌托邦场景俯拾皆是。一个很难避免的例子就是有人可能会

在自家地下室使用开源模型，将任何人物的照片制作成色情内容，然后用于敲诈——美国联邦调查局已经对此发出警告。也许任何人都能制作出某位总统或总理宣布发起先发制人核打击的视频，让世界瞬间进入紧张状态。冒充受害者亲属的诈骗者也会兴风作浪。

然而，社会也将适应造假者的存在。人们将认识到，图像、音频或视频并不能证明某件事情发生过，就像一幅图画不能证明某事发生过一样（可以将信息收集任务可靠地众包出去的开源情报时代可能不会长久了）。在线内容将不再能自证真实性，因此是谁发布了内容将变得和发布了什么内容一样重要。假如值得信赖的信息来源能够继续通过URL、电子邮件地址和社交媒体平台安全可靠地证明自己的身份，声誉和内容出处就将变得前所未有的重要。

这听起来可能很奇怪，但这在历史上的大部分时期都成立。批量生产可信内容的时代才是个例外。人们可能很快就会难以找出AI这只看不见的手，但这并不意味着思想市场从此就要完蛋了。随着时间的流逝，能好好存活下来的虚假内容将主要是搞笑好玩的那种。■



Business and the culture wars

How to cut through the cacophony over DEI

Outrage on right and left obscures both the costs of DEI and the benefits of diversity

WHAT, IF ANYTHING, should firms do to improve the diversity of their workforce? After the murder of George Floyd in 2020, many bosses felt compelled to act. Partly out of fear of being called out for prejudice, corporate America rushed to embrace diversity, equity and inclusion (DEI) schemes. By 2022 three-quarters of the S&P 500 had a chief diversity officer; more than two-fifths of listed firms set targets to increase the racial diversity of their workforce. Now many think the pendulum has swung too far.

The resignation of Claudine Gay, a black woman, as president of Harvard University has ignited a broader debate about merit and identity. Bill Ackman and Elon Musk, two billionaires, have excoriated DEI for itself being discriminatory. After the Supreme Court's landmark decision to end affirmative action in university admissions last summer, many activists and politicians increasingly have corporate DEI schemes in their sights.

As America's culture wars rage on, bosses are being caught in the middle. Progressives argue that DEI enables companies to do their bit to tackle America's entrenched inequalities. Conservatives see it as an attack on meritocracy. One side ignores the costs of many DEI schemes, the other ignores the real benefits of diversity. How should businesses cut through the noise?

The critics are right that the thinking on DEI is muddy, and that many DEI initiatives are ineffective, even harmful. In 2015 McKinsey, a consultancy, identified a positive correlation between the gender and ethnic diversity of the workforce and firms' profitability. Although academics have since

criticised its methodology, the findings were breathlessly cited by bosses and corporate advisers, and the link was treated as causal and cast-iron. For example, from 2023 Nasdaq required firms listed on its stock exchange to have at least one board member who was not a straight white man—or explain why they do not. It was left to Jesse Fried, a professor at Harvard Law School, to point out that Nasdaq was ignoring scholarship which finds that board diversity can have a negative impact on performance.

Diversity schemes often fail. Sometimes this betrays bad faith: firms with a discrepancy between their words and actions are often accused of “diversity washing”. Some schemes are well-meant but ineffective. Research by Frank Dobbin and Alexandra Kalev showed that diversity training programmes fail to reduce bias. In the worst instances, DEI initiatives backfire. Targets can be seen as quotas, which undermine the principle of fair competition and cast a shadow over minorities who do well under them. Other research shows that adding equal-employment statements to job advertisements can put minority candidates off applying. No wonder support for votes on social issues at annual general meetings is draining away.

The case for diversity does not need dressing up in pseudoscience. The simple reason for businesses and their shareholders to care about recruiting people from a broad range of backgrounds is that they want the most able people. Mr Musk and Mr Ackman are both successful businessmen: they too want to assemble the best possible teams.

| *Opus DEI*

Diversity should be a spur to looking far and wide for talent, no matter someone’s gender, race or sexual orientation. A firm convinced that it is overlooking the best candidates from a particular demographic cohort, for example, could choose to lengthen its shortlists to include more from that group. That will not mechanically create workforces that mirror the

population, but it can maximise talent and diversity of thought. Quotas, by contrast, have the perverse effect of narrowing the search by excluding talent. As with so many areas touched by the culture wars, the row over DEI has become muddle-headed. The clear, simple argument for diversity is being drowned out. ■



【首文】商业和文化战

如何冲破“多元共融”的喧嚣争论

右翼和左翼的愤怒掩盖了DEI的代价和多元化的好处

如果真有办法的话，公司应该做些什么来提升员工队伍的多元化呢？2020年乔治·弗洛伊德（George Floyd）遇害后，许多老板感到不得不采取行动。部分出于担心被指责存在偏见，美国企业界急于拥抱“多元、公平和包容”（DEI）体系。到2022年，标准普尔500指数中四分之三的公司都设有首席多元化官，超过五分之二上市公司设定了提升员工种族多元化的目标。现在许多人认为企业在这条路上走得太远了。

黑人女性克劳丁·盖伊（Claudine Gay）辞去哈佛大学校长一职，引发了一场关于才能和身份的更广泛辩论。比尔·阿克曼（Bill Ackman）和马斯克这两位亿万富翁曾批判DEI本身具有歧视性。去年夏天，美国最高法院做出了具有里程碑意义的裁定，终止了大学招生中的平权法案。此后，越来越多的活动人士和政界人士开始关注企业的DEI制度。

随着美国的文化战愈演愈烈，老板们被夹在了中间。进步派认为，DEI让企业能够为解决美国根深蒂固的不平等问题尽自己的一份力。保守派认为，这是对优绩主义的攻击。一方忽视了许多DEI制度的代价，另一方忽视了多元化的真正好处。企业应该如何在各种声音中理清头绪？

批评DEI的人没有说错，关于DEI的思考是混乱的，许多DEI举措是无效甚至是有害的。2015年，咨询公司麦肯锡发现，员工的性别和种族多元化与公司盈利能力之间存在正相关关系。尽管学者们之后对它的研究方法提出了批评，但老板和企业顾问们急不可待地引用这项研究结果，把这种关联性视作因果关系且确凿无疑。例如，从2023年起，纳斯达克要求在其证券交易所上市的公司至少得有一名董事会成员不是异性恋白人男性，如果没有，就要做出解释。哈佛大学法学院教授杰西·弗里德（Jesse Fried）指出，纳斯达克忽略了发现董事会多元化会对业绩产生负面影响的学术研究。

多元化计划经常失败。有时这暴露了公司不诚信：言行不一的公司常被指斥搞“虚假多元化”。有些计划意图是好的，但效果不佳。弗兰克·多宾（Frank Dobbin）和亚历山德拉·卡列夫（Alexandra Kalev）的研究表明，多元化培训计划未能减少偏见。最糟糕的情况是DEI方案适得其反。目标会被视为配额，这破坏了公平竞争的原则，并给在这样的计划中表现良好的少数群体蒙上了阴影。其他研究表明，在招聘广告中加入平等就业声明可能会让少数群体求职者放弃申请。难怪在年度股东大会上对社会问题投票的支持正在减少。

想要多元化不需要用伪科学来粉饰。企业及其股东重视从广泛背景中招聘人才的一个简单原因是，他们想要最有能力的人。马斯克和阿克曼都是成功的商人，他们也想组建最好的团队。

| DEI事业

多元化应该激励人们将眼光放宽放远去寻求人才，无论他们的性别、种族和性取向如何。例如，如果一家公司确信自己忽视了某个人口群体中的最佳候选人，它可以选择加长候选名单，将更多来自该群体的人包括进来。这样不会机械地打造出完全对应人口构成的员工队伍，但可以最大限度地寻找人才并实现想法的多样性。相比之下，配额制产生了将人才排除在外、缩小搜索范围的反效果。和许多受文化战波及的领域一样，关于DEI的争执已经变成一团乱麻。支持多元化的清晰、简单的理由正在被淹没。





Chaos—or opportunity?

Donald Trump is winning. Business, beware

What a second term would mean for American business and the economy

WHEN DONALD TRUMP slunk out of the White House in 2021, executives at large American companies sighed with relief. Now that he has won Iowa's caucuses by a margin of 30 points, they are digesting the reality that this time next year Mr Trump could be behind the Resolute desk once again. The Economist has spent the past few weeks talking to these titans. Some are deeply alarmed by the prospect of Trump 2. But others quietly welcome the chaos trade.

People who run large organisations have to be optimistic. They must find opportunities when others are panicking. CEOs had an uneasy relationship with President Trump, many distancing themselves from his most outrageous pronouncements and tut-tutting about protectionism, even as they enjoyed his more conventional policies. Republicans in Congress may have talked about being the pro-worker party, but in practice they cut business taxes. It was hard for corporate America to be miserable amid a soaring stockmarket.

If Mr Trump is indeed elected again, those running big firms plan to keep their heads down ("don't be Bud Light" is a frequent refrain, after the beer brand fell victim to the culture wars). They would avoid being dragged onto Mr Trump's business councils, dodge presidential photo-ops and get on with making money. True, if Mr Trump did a deal with Russia that ended the war and sold out Ukraine, that would be bad for Western civilisation. But it would reduce energy bills.

What's more, Trump enthusiasts in the C-suite have plenty of grumbles about Joe Biden. Mention Lina Khan, who oversees the Federal Trade

Commission (the antitrust police), or Gary Gensler, who leads the Securities and Exchange Commission (the Wall Street police), and they inhale sharply. Mr Biden wants to raise taxes on companies. His administration also wants to go ahead with the Basel III “Endgame” regulations, which oblige big banks to hold perhaps 20% more capital on their balance-sheets, sedating animal spirits and damaging profitability.

Yet this bullish case for Mr Trump’s economic management is complacent. It fails to recognise how Trumponomics—a mix of deficit-funded tax cuts and tariffs—would work differently today. And it ignores the ways in which Mr Trump’s most chaotic tendencies could threaten America, including its companies.

In his first term the economy did better than many economists (including ours) expected. That was in part because Trumponomics turned out to be more moderate than the campaign had promised. The economy was also running further below capacity than thought, making it possible to cut taxes without stoking inflation. Strong overall growth and low inflation masked the damage done by Mr Trump’s protectionism.

There is no evidence that Mr Trump has updated his approach: he is still a tax-cuts-and-debt guy. But the economic conditions have changed. For the past two years the Federal Reserve has been trying to bring down inflation. Though it has nearly succeeded, the labour market remains tight. Today 2.8m more 25- to 54-year-olds are in work than would be if the employment rates of January 2017 had persisted. Then there were 1.3 unemployed workers for every job opening; today there are only 0.7. As a result the economy is more prone to overheating.

The budget is in worse shape, too. In 2016 the annual deficit was 3.2% of GDP and debt was 76% of GDP. The forecasts for 2024 are 5.8% and 100%, respectively. Should Mr Trump once again pursue tax cuts, the Fed will have

to hike up interest rates to offset the stimulus, making it costlier for businesses to raise capital and for the government to service its growing debt pile.

These are the conditions under which Latin American populists bully their central banks to keep rates low, a practice Mr Trump dabbled in last time. The Fed is supposed to be independent, but Mr Trump will have a chance to nominate a stooge as chair in May 2026 and a pliant Senate could indulge him. The risk of more inflation would surge, perhaps exacerbated by more tariffs, which would also slow growth.

On top of that big macroeconomic risk are many others. Firms would not relish further trade restrictions, but some members of Mr Trump's circle have floated a 60% tariff on imports from China. Lots of companies like the federal government's support for renewable energy (which Mr Trump calls the Green New Scam). He has promised the biggest deportation scheme in American history to reduce the number of illegal immigrants in the country. As well as causing misery, this would be a shock to that tight labour market.

As ever, saying what Mr Trump would actually do is very hard: he has few fixed beliefs, is a chaotic boss and can reverse position several times a day. In a town hall in Iowa he said he would be too busy in his second term to seek retribution against his political enemies. That was a few hours after his own campaign sent out an email with the subject line: "I am your retribution!" He could recognise Taiwan's independence, prompting a meltdown in Beijing and a blockade of the island. Or he could walk away from Taiwan in exchange for China buying more stuff from America. Businesses often say that what they fear most is uncertainty. With Mr Trump that is guaranteed.

This unpredictability could make a second Trump term very much worse

than the first. His administration would lack establishment types like Gary Cohn, once of Goldman Sachs, to shuffle the president's in-tray and hide the madder ideas from him. More moments like January 6th are possible, as is a full-on revenge presidency. The idea that in this scenario business leaders could keep a low profile and focus on EBITDA is fanciful. Employees, customers and the press would demand to know where bosses stood and what they proposed to do. The administration might in turn take exception to every whiff of criticism.

In the long run, the idea that corporate profits can be insulated from societal upheaval is a fantasy. If Mr Trump is broadly corrupting of American politics, and businesses are seen to profit from his rule, that poses a big risk to them in the future. In Latin America, when big businesses have become associated with autocrats the result was usually that capitalism was discredited and the appeal of socialism rose. That seems unthinkable in America. But so, until recently, did a second Trump term. ■



【首文】混乱还是机遇？

特朗普传捷报，商界要当心

他重回白宫将如何影响美国商业和经济

特朗普在2021年离开白宫时，美国大企业的高管们都松了一口气。现在，眼看特朗普在艾奥瓦州共和党党团初选中以30个百分点的优势击败对手，他们正在消化一个现实：明年此时，特朗普有可能重新坐到白宫的坚毅桌后。过去几周，本刊采访了这些大佬们。有人对特朗普可能的回归深感不安，但也有人暗暗期待商界乱世的来临。

做大企业的老板是必须保持乐观的。他们要在别人惊慌失措时找到机会。特朗普在任时，企业老板们和这位总统的关系尴尬，他们很享受他那些更偏保守的政策，但同时许多人会跟他的出格言论保持距离，对他的保护主义做派发出啧啧不满声。共和党人在国会里也许声称自己是为工人谋求利益，但在行动上却是给商界减税。在强劲牛市之下，美国商界那时想过苦日子都难。

假如特朗普真的再次当选，那些美国大企业的高管打算低调行事（百威的淡啤酒品牌听蓝成了文化战的牺牲品后，大家都在念叨着“别成了百威听蓝”）。他们不要被拉进特朗普的商业顾问团，不要和总统合影，而是继续埋头赚钱。诚然，假如特朗普与俄罗斯达成协议，结束俄乌战争，出卖乌克兰，对西方文明来说是件坏事。但这会降低能源价格。

而且，这些企业高管中的特朗普支持者对拜登有诸多不满。一提到反垄断机构美国联邦贸易委员会的主席丽娜·可汗（Lina Khan）或者金融监管机构美国证券交易委员会的主席加里·盖斯勒（Gary Gensler），他们都会倒吸一口冷气。拜登想对公司加税。拜登政府还计划推进实施《巴塞尔协议III》“终局”规则，要求大银行在资产负债表上多持有大约20%的资本，抑制动物精神，破坏盈利能力。

然而，如此看好特朗普的经济管理就想得太美了。这没能意识到特朗普经

济学（以赤字支持的减税和加征关税）到如今可能会有不一样的效果，还忽视了特朗普反复无常的行事风格可能如何威胁到美国，包括美国的企业。

在特朗普的第一个任期内，美国经济的表现优于许多经济学家（包括本刊）的预期。部分原因是特朗普经济学的实际执行力度比竞选时的承诺更温和。经济离满负荷运行的距离也大于预期，这使得政府可以在不刺激通胀的情况下减税。强劲的总体增长和低通胀掩盖了特朗普保护主义带来的伤害。

没有证据表明特朗普已经更新了策略：他还是主张减税和举债。但经济形势已然改变。过去两年来，美联储一直在努力降低通胀。尽管这一努力接近成功，但劳动力市场依然紧张。如果2017年1月时的就业率持续到现在，25岁至54岁的就业人口会比现实情况少280万。那时的职位空缺与失业者比率为1: 1.3，而现在仅为1: 0.7。所以现在经济更容易过热。

财政预算状况也比之前糟糕。2016年，美国的年度赤字和债务占GDP的比例分别为3.2%和76%，而对2024年的预测为5.8%和100%。如果特朗普再次推行减税政策，美联储将不得不加息以抵消这种刺激，这将使企业融资成本上升，债台高筑的政府偿债成本也会增加。

拉美的民粹主义者就是在这种情况下迫使本国央行维持低息的，特朗普在上个任期内也已经试过了。美联储本应是独立的，但特朗普将有机会在2026年5月提名一个亲信担任美联储主席，一个温顺的参议院可能对他的做派听之任之。通胀加剧的风险可能会猛增（也许是由于关税加码火上浇油），这也会拖慢经济增长。

除了这一重大宏观经济风险外，还有许多其他风险。企业不愿看到进一步的贸易限制，但特朗普圈子里的一些人已提出对从中国进口的商品征收60%的关税。很多公司希望联邦政府支持可再生能源（特朗普称之为“绿色新骗局”）。他承诺推行美国史上最大规模的驱逐计划以减少国内非法移民的数量，这不但会造成不幸，还会对美国紧张的劳动力市场造成冲

击。

与以往一样，要预测特朗普实际会怎么做非常之难：他没什么确定的信念，是个言行混乱的老板，一天之内就可能多次改变立场。在艾奥瓦州一个市政厅内，他说自己在第二个任期会忙得不可开交，没空去报复政敌。而就在几小时前，他自己的竞选团队还发出了一封邮件，主题是“我是你的报应！”他可能承认台湾独立，触怒北京并导致台湾被封锁。他也可能放弃台湾，换取中国大陆从美国进口更多商品。商界人士常常说他们最害怕的是不确定性。假如特朗普上台，肯定会带来不确定性。

这种不可预测性可能会让特朗普第二个任期的情况比第一个任期糟糕得多。他的政府将缺少像曾在高盛任职的加里·科恩（Gary Cohn）这样的建制派来梳理总统案头的公务，挡下一些他的疯癫想法。1月6日骚乱事件这样的时刻一再出现是可能的，在任期内全面复仇也是有可能的。在这样的境况下还认为企业领导者还可以保持低调、专注赚钱，那真是异想天开了。员工、客户和媒体都将会要求了解企业老板在选举中的立场和提出的商业方案。特朗普政府继而可能就嗅到的任何一丝批评气味做出强烈反应。

长远来看，认为企业利润可以不受社会动荡的影响无异于白日做梦。假如特朗普全面腐蚀美国政治，且人们认为企业从他的统治中谋利，它们的未来就会面临巨大风险。在拉丁美洲，当大企业与威权主义者搭上关系，结果通常是资本主义信誉扫地，社会主义吸引力上升。这在美国似乎是难以想象的。但要知道，“特朗普重回白宫”在不久前也是那么难以想象。■



Beyond the screen

Hollywood studios are finding new ways to bring stories to life

Watch “Stranger Things” on stage, eat at a Batman-themed restaurant—or take your chances at “Squid Game”

THE LATEST episode in Netflix’s “Stranger Things” saga was released on December 14th, featuring levitating bodies, shrieking monsters and an exploding rat. The reviews were stellar. Yet unlike the previous season of the science-fiction show, which clocked nearly 1bn hours of viewing in its first month, the most recent instalment has so far been seen by only a few thousand people. That is because Netflix’s new show is not being streamed down fibre-optic cables to television screens, but performed live on a stage in London’s West End.

“Stranger Things: The First Shadow” (pictured), the streamer’s first stab at theatre, is playing at the Phoenix, with hopes of an international run. It is not the only example of Tinseltown invading theatreland. A few streets away at the Theatre Royal, Disney offers a live version of “Frozen”, while nearby at the Adelphi there is a musical tribute to “Back to the Future”. In 2025 a live production of “Paddington” will join the London line-up. Meanwhile on Broadway, Amazon is getting ready to launch a musical of “Transparent”, a drama that first ran on its Prime Video service.

Hollywood’s turn on the stage is part of a broader shift by the movie business towards live experiences. As attendance at the cinema declines, studios are finding new ways to excite—and monetise—their fans outside their homes. From restaurants and art exhibitions to escape rooms and assault courses, film-makers are concocting novel ways to soak up demand. “There’s this insatiable appetite from those mega-fans,” says Marian Lee, Netflix’s chief marketing officer. “They want more. They’ll eat up anything you serve them.”

It is a new take on an old playbook. In the 1950s Walt Disney drew a diagram of the mutually reinforcing pillars of his business: movies promoted television spin-offs, which fed demand for theme parks, which sold merchandise, which promoted movies, and so on. Other studios such as Universal copied the blueprint, creating parks and toys from the characters made popular by films.

Today these ventures are more than sidelines. The entertainment industry's turbulent digital transition has left Hollywood more reliant than ever on rollercoasters and plastic lightsabres. In the last financial year Disney's "experiences" division, which includes five theme parks and a fleet of cruise ships, contributed \$9bn in operating profit, as its streaming business lost \$2.5bn. At the same time the creative engine at the centre of Walt Disney's diagram—the cinema—is sputtering. Worldwide box-office takings for 2023 will be 20% below their pre-pandemic level; even before covid the average American was going to the movies three and a half times a year, down from five in 2000. As the silver screen fades, studios are losing their most powerful way of exciting audiences.

So they are turning to new tactics. Some are doubling down on their parks. Warner Bros, a relative latecomer to the business, opened the world's largest indoor theme park in Abu Dhabi in 2018. Disney announced in September that it would double its investment in its parks and cruises over the next decade, promising "Frozen" lands among other attractions. Universal is building a resort in Texas and in December said it had bought land for a possible park in Britain.

Studios are also devising new kinds of entertainment. Warner opened "Visions of Magic", a Harry Potter "interactive art experience", in Cologne in December and plans a similar show based on its DC Comics series in late 2024. Universal promises an interactive "horror experience" in Las Vegas, based on its catalogue of scary movies, beginning with "Frankenstein"

(1931).

Moviemakers are experimenting with food and drink, too. Warner's eateries range from Park Row, a Batman-themed restaurant in London with a ten-course tasting menu priced at £195 (\$246), to Central Perk, a "Friends" café that opened in Boston in November, serving coffee and Joey's meatball sandwiches.

"You're trying to take the pulse of what folks are interested in," says Peter Van Roden of Warner's themed-entertainment division, who monitors merchandise sales, online clicks and box-office receipts to see which titles are good candidates for turning into such experiences. Social media, which encourage sharing and "FOMO" (fear of missing out), have made it easier for short-term "pop-up" experiences to succeed, he adds.

Netflix, the studio that is most sceptical of the cinema, has been among the most innovative at devising alternatives. Before it began organising live events, devotees of popular shows would come to premieres in home-made costumes, says Ms Lee. "We had all these breadcrumbs that were indicating to us that fans wanted a deeper way to get together...and then have something in-person to engage with," she says. Netflix tested various formats, starting with a "Stranger Things" drive-through show during the pandemic, and has since held events including "Money Heist" escape rooms and "Bridgerton" balls.

The latest is a "Squid Game" experience in Los Angeles, where participants play six games based on the show, bossed around by suitably strict guards (there is no prize money, but no risk of execution either). The Economist's contestant came second, in part because he consumed fewer pre-game Korean cocktails from the adjoining bar than some of his rivals.

Unlike Disney, whose parks are priced with an eye on the bottom line,

Netflix sees its events as ads for its shows. The brief is: “What can you give [fans] in between seasons, to really further that excitement that they have around these characters?” says Ms Lee. Prices reflect this: the cost of the 70-minute live “Squid Game” experience starts at \$39, while tickets for the “Stranger Things” play are available from £20, cheap by West End standards. (Green-and-white tracksuits and cuddly demogorgons cost extra.) The streaming company does not report revenues for its events or merchandise, but on a recent earnings call described them as “small things”.

Can the new in-person experiences make audiences fall in love with characters and stories as the cinema long has? Mr Van Roden, who declares himself a cinema fan, rejects the idea that they are substitutes. But their essential, shared characteristic is that they are communal. “Even if you’re not consciously connecting with the people in the room with you, it’s this common, human sort of electric connection, that you’re sharing it together,” he says.

At the “Squid Game” assault course, the cocktail-fuelled contestants have been having such a riotous time together that Netflix has considered adding more soundproofing to the set. The growing number of live attractions from Hollywood means that audiences have more reasons than ever to get off the sofa and have fun with other people—even if the ultimate aim is to persuade them to spend yet more time in front of the television.





超越屏幕

好莱坞制片厂找到演故事的新方法

观看舞台上的《怪奇物语》，在蝙蝠侠主题餐厅用餐，或者在“鱿鱼游戏”中搏一把

奈飞（Netflix）于去年12月14日推出了长剧集《怪奇物语》（Stranger Things）的最新一集，里面有悬在半空的人、尖叫的怪物和一只爆炸的老鼠。观众好评如潮。然而，不像该科幻剧的上一季在第一个月就有接近10亿小时的观看时长，这最新一集迄今为止只有几千人看过。这是因为奈飞的这集新剧不是通过光纤传输到电视屏幕上，而是在伦敦西区的一个舞台上现场表演。

《怪奇物语：第一道暗影》（Stranger Things: The First Shadow，如图）是奈飞首次涉足剧院，目前正在凤凰剧院（Phoenix）上演，并希望能开展国际巡演。这并不是好莱坞入侵戏剧界的唯一例子。在几条街之外的皇家剧院（Theatre Royal），迪士尼推出了《冰雪奇缘》剧场版，而附近的阿德尔菲剧院（Adelphi）正在上演致敬《回到未来》（Back to the Future）的音乐剧。2025年，剧场版的《帕丁顿熊》将登上伦敦的演出剧目名单。与此同时，在百老汇，亚马逊正准备推出音乐剧版的《透明家庭》（Transparent），这部剧集是在亚马逊的Prime Video上首播的。

好莱坞作品登上舞台反映了电影业向现场体验转变的大趋势。随着电影院上座率下降，制片公司正在找出新办法来吸引影视剧迷们走出家门和掏腰包。从餐馆和艺术展览，到密室逃脱和障碍训练场，电影制作者正在编造全新的戏法来满足观众需求。“超级粉丝的胃口永远填不饱，”奈飞的首席营销官玛丽安·李（Marian Lee）表示，“他们总要更多。你拿出什么，他们都会消化掉。”

这实是新瓶装旧酒。上世纪50年代，华特·迪士尼绘制了一张商业版图，让他的几大业务支柱相互支撑加固：电影促进电视衍生节目的开发，为主题公园带来需求，后者又销售周边商品，进而为电影增加热度，如此循环。环球影业等其他制片公司也如法炮制，用电影中的热门角色创造了主

题公园和玩具。

如今，这些不再只是副线业务。娱乐行业的数字化转型激烈动荡，让好莱坞变得比以往任何时候都更加依赖过山车和塑料光剑。在过去一个财年中，迪士尼的“体验”事业部管理的五个主题公园和游轮船队为公司带来了90亿美元的营业利润，而其流媒体业务则亏损了25亿美元。与此同时，处于华特·迪士尼的战略版图核心位置的创意引擎——电影——正在熄火。2023年全球票房收入较疫情前下降了20%；而在疫情前，美国人的年均观影次数已经从2000年的5次下降到了3.5次。随着银幕逐渐暗淡，制片公司正在失去激发观众兴趣的最有力武器。

制片公司因此正转向新策略。一些公司正加倍押注主题公园。华纳兄弟进入该领域相对较晚，于2018年在阿布扎比（Abu Dhabi）揭幕了全球最大的室内主题公园。迪士尼在去年9月宣布，未来十年在主题公园和游轮建设上的投资将翻番，承诺推出“冰雪奇缘”等主题园区。环球影业正在得克萨斯州兴建度假区，并于去年12月表示已在英国购买土地，以后可能会用来建一座主题公园。

制片公司也在设计新的娱乐形式。华纳兄弟去年12月在科隆推出了名为“魔法幻境”（Visions of Magic）的哈利·波特“互动艺术体验馆”，并计划在2024年底推出一个基于DC漫画系列的类似的互动体验馆。环球影业表示，将利用其恐怖片库在拉斯维加斯打造互动“恐怖体验”，第一个项目将是“科学怪人”（1931年）。

电影制片商也在试水餐饮业。华纳兄弟的餐馆有不同主题，搭配不同菜式。在伦敦的蝙蝠侠主题餐厅Park Row提供包含十道菜的品鉴套餐，售价195英镑（246美元）。去年11月在波士顿开业的“老友记”咖啡馆Central Perk提供咖啡和乔伊最爱的肉丸三明治。

“我们在试着摸清人们对什么感兴趣。”华纳兄弟主题娱乐部门的彼得·范罗登（Peter Van Roden）说，他通过监测周边商品销售、在线点击和票房收入来确定哪些作品适合转化为这类体验。他补充说，社交媒体鼓励分

享，也触发“FOMO”（错失恐惧）心理，让短期的“快闪”体验更容易取得成功。

最不看好院线的制片公司奈飞在设计创新替代方案方面走在前列。玛丽安·李表示，在奈飞还没开始组织现场活动时，热门剧集的粉丝就会穿着自制服装参加首映式。她说：“有很多迹象显露出粉丝希望以更深入的方式相聚……并能参与一些面对面的互动。”奈飞做了多种尝试，先是在疫情期间让剧迷们驾车勇闯“怪奇物语”小镇，后来又组织了“纸钞屋”（Money Heist）密室逃生和“布里奇顿家族”舞会等主题活动。

奈飞最新的沉浸式体验是在洛杉矶的“鱿鱼游戏”，在严厉程度适当（赢了没奖金，但输了也不会被处决）的守卫的指挥下，参与者根据剧情玩六场游戏。笔者获得了第二名，未能夺冠的部分原因是赛前在比赛场地旁的酒吧里喝的韩国鸡尾酒没有对手多。

迪士尼乐园的定价着眼盈利，奈飞不同，它将体验活动视作给自己的节目打广告。任务简介就是：“在剧歇期间，拿出什么东西延续角色（给粉丝）带来的兴奋感？”玛丽安·李说。价格就体现了这一点。70分钟的体验版“鱿鱼游戏”门票起售价39美元，剧版《怪奇物语》的门票最低20英镑，按伦敦西区的标准来说很便宜。（绿白配色的运动服和可爱的魔王公仔需要另外付费）。奈飞没有公布其主题活动或周边商品的收入，但在最近的财报电话会议上将其描述为“小意思”。

这些新的沉浸式体验能否像电影院长久以来所做的那样，让角色和故事赢得观众的喜爱？自称是影迷的范罗登否认它们是替代品。不过它们的基本共同特点是它们都是群体活动。他说：“即使你没有在有意识地与同处一室的其他人交流沟通，大家也都在一起经历着一种共同的心灵触电般的感应。”

在“鱿鱼游戏”的障碍训练场，鸡尾酒上头的参赛者在一起玩得很疯，奈飞甚至开始考虑增加隔音设施。好莱坞推出的现场体验越来越多，给了观众更多从沙发上爬起来、与其他人一起玩乐的理由，即使这些活动的最终目

的是让他们在电视机前坐更久。■



Superhero films

Marvel seems to be losing its powers

The world's mightiest movie franchise looks increasingly fragile

IN “THE AVENGERS” (2012) Nick Fury (Samuel L. Jackson), a spy, described heroes as “an old-fashioned notion”. Certainly the film’s characters, including Captain America and Iron Man, were not novel, first appearing in comic books published in the mid-20th century. But if the idea was old, the excitement around superheroes had been renewed. “The Avengers” became the first Marvel movie to make more than \$1bn at the global box office.

When Fury’s words were used in the trailer for “The Marvels” (2023, pictured), however, they took on a different tone. Heroes may seem antiquated, he argued, but “the world can still use them”. If it was an attempt to convince the viewer, it did not work. Released in November, “The Marvels”, the 33rd instalment in the Marvel Cinematic Universe (MCU), made around \$200m at the box office. It became the poorest-performing MCU film to date, and will probably lose money.

Nor was “The Marvels” a one-off disappointment. “Ant-Man and the Wasp: Quantumania” also underperformed. According to CinemaScore, an audience-rating benchmark, of the past eight MCU films, five have scored B+ or worse (see chart). Fans complain of dull characters, sloppy writing and amateurish special effects.

Marvel productions on the small screen have not fared much better. Recent MCU television series on Disney+, including “Secret Invasion”, about Fury’s character, have been poorly reviewed and, estimates suggest, little watched. It does not bode well for the shows due to be released in the coming months.

The decline is surprising: for a long time, the Marvel brand seemed invincible. Disney bought the comic-book company in 2009 and it became a prized asset. The 23 movies released between 2008 and 2019 grossed almost \$23bn in total, making Marvel the largest film franchise in history.

Marvel kept standards high even as it increased production. The company released 2.75 films, on average, in 2016-19, up from 1.2 in 2008-13. Of those 23 movies, only one ranked lower than A- on CinemaScore. Three films received an A+, awarded to fewer than 100 of over 4,000 films measured since 1979. “Black Panther” (2018) even became the first comic-book adaptation to be nominated for Best Picture at the Oscars.

Marvel pioneered an innovative “cinematic universe” model, in which plotlines and characters were shared across films. As Marvel’s universe grew, its competitors tried, and failed, to emulate its success. DC Comics—which owns Batman, Superman and Wonder Woman—set up, and recently scrapped, its “Extended Universe”. Warner Bros has turned the Harry Potter franchise into a “Wizards World”. Universal twice tried to launch a “Dark Universe” of monsters such as Dracula and the Mummy, but both attempts failed after a single release. Efforts to build out Robin Hood and his merry men (Lionsgate), Power Rangers (also Lionsgate) and King Arthur and his round table (Warner Bros) all faltered.

By the early 2020s the MCU seemed set for further dominance. In 2019 Disney acquired 20th Century Fox, which held the rights to characters including the X-Men and the Fantastic Four. The launch of Disney+ that year made it easier for fans to keep up with the ever-expanding MCU and enabled the franchise to tell new stories in a serialised format. But instead of developing its position in pop culture, Marvel has struggled creatively and financially.

Disney insiders suggest several causes for the slump. One is to do with

personnel. Several trusted writers and directors have moved on. Many of the actors playing the most popular superheroes left the MCU after “Avengers: Endgame” in 2019, and Chadwick Boseman, the star of “Black Panther”, died in 2020. Last month Disney fired Jonathan Majors after he was found guilty of assaulting and harassing his then-girlfriend. The actor played the villain at the heart of the “Multiverse Saga”, the story which would connect the films released between 2021 and 2027.

Another reason is to do with geopolitics. The first 23 films were all released in China, the world’s largest theatrical market, but between 2020 and 2022, none was. (China did not give a clear reason why, but it was probably building up its domestic film industry.) Though this de facto ban is now over, cinematic universes are hard to understand when audiences have missed several entries. Making matters worse, Disney+ is not available in China, so fans cannot watch the TV entries.

Yet part of the problem is of Marvel’s own making. Since 2021 the MCU has released an average of 3.3 films and 3.7 television series every year—a rate that seems to strain audiences, internal creative teams and special-effects departments. For prospective viewers hoping to watch a new title, 33 films and 11 seasons of television is simply too much homework. The focus on the “multiverse”, which draws on films predating the existing cinematic universe, aggravates this issue.

Audiences may yet tire of superheroes much as they tired of Westerns in the late 1960s. But for now, the genre goes on. The third “Guardians of the Galaxy” film grossed \$846m, making it the fourth-highest-grossing film of 2023, and received an A rating on CinemaScore. “Spider-Man: Across the Spiderverse”, an animated film by Sony, was also among the most popular films last year.

Bob Iger, Disney’s CEO, who initiated Marvel’s expansion, has said the

franchise can return to its former glory by slowing the pace of production. “I’ve always felt that quantity can be actually a negative when it comes to quality. And I think that’s exactly what happened. We lost some focus.” He, and Marvel’s many fans, will be holding out for the heroes. ■



超级英雄电影

漫威似乎英雄失势

世界最强大系列电影看起来日渐式微

在2012年上映的《复仇者联盟》中，塞缪尔·杰克逊（Samuel L. Jackson）饰演的特工尼克·弗瑞（Nick Fury）说，英雄是个“已经过时的概念”。毫无疑问，电影中的美国队长、钢铁侠等角色都并非新创，它们最早出现在上世纪中期出版的漫画书中。如果说英雄的概念老旧了，人们对超级英雄的热情却被重新点燃了。《复仇者联盟》成为第一部全球票房破十亿美元的漫威电影。

然而，当弗瑞的这句话用在2023年上映的《惊奇队长2》（如图）的预告片中时，话风不一样了。他说，英雄可能看上去过气了，但“世界仍然可以用用他们”。如果这话想拿来说服观众，那并没有起到效果。去年11月上映的《惊奇队长2》是漫威电影宇宙（MCU）的第33部电影，票房约为两亿美元。它也因此成为MCU有史以来票房最差的影片，而且可能还会赔钱。

《惊奇队长2》并非唯一令人失望的影片。《蚁人与黄蜂女：量子狂潮》也表现不佳。根据权威观众评分网站CinemaScore的数据，在最近上映的八部MCU电影中，有五部获得了B+或更低的评分（见图表）。粉丝们抱怨它们角色乏味、剧本潦草、特效拙劣。

漫威在小屏幕上推出的制作也乏善可陈。最近在Disney+上播出的MCU剧集，包括以弗瑞为主角的《秘密入侵》（Secret Invasion），评价都很差，而且估计收视率极低。这对计划在未来几个月上映的剧集来说不是个好兆头。

这样的衰落令人惊讶——因为在很长一段时间里，漫威品牌似乎战无不胜。2009年，这家漫画公司被迪士尼收购，成为其旗下的宝贵资产。2008年至2019年间，漫威出品了23部电影，票房总收入近230亿美元，打

造出了历史上规模最大的系列电影。

漫威在增加产量的同时保持了高水准。2016至2019年间，漫威平均每年出品2.75部电影，而2008年至2013年为1.2部。在2008年至2019年上映的23部电影中，只有一部在CinemaScore上的评分低于A-；三部电影获得了A+——在1979年以来被打分的全部4000多部电影中，获得A+的还不到100部。2018年上映的《黑豹》甚至成为第一部获得奥斯卡最佳影片提名的漫画改编电影。

漫威开创了一种“电影宇宙”的新模式，让旗下电影拥有共同的情节主线和角色。随着漫威宇宙的壮大，竞争对手们也试图仿效它的成功之道，但都以失败告终。拥有蝙蝠侠、超人和神奇女侠的DC漫画建立了“扩展宇宙”（Extended Universe），但在不久前放弃了。华纳兄弟将《哈利·波特》系列电影变成了“魔法世界”（Wizarding World）。环球影业曾两次尝试将吸血鬼德古拉和木乃伊等怪物集结在一起，推出“黑暗宇宙”（Dark Universe），但都在上映了一部电影后就难以为继。此外，狮门影业着力打造的“罗宾汉和他的快乐伙伴们”（Robin Hood and his merry men）和“超凡战队”（Power Rangers），还有华纳兄弟的“亚瑟王和他的圆桌骑士”（King Arthur and his round table），最终都归于沉寂。

一直到2020年代初，MCU似乎都势将进一步占据主导地位。2019年，迪士尼收购了拥有《X战警》和《神奇四侠》等电影角色版权的二十世纪福克斯。同年，Disney+推出，这让粉丝们更容易跟上不断扩容的MCU宇宙，也让漫威系列电影可以用剧集的形式讲述新故事。但漫威并没有巩固自己在流行文化中的地位，反而在创意和财务上陷入了困境。

针对漫威的日渐式微，迪士尼内部人士提出了几点原因。一个是人事问题。好几位值得信赖的编剧和导演相继出走。2019年《复仇者联盟4：终局之战》上映之后，许多最受欢迎的超级英雄的扮演者离开了MCU。2020年，《黑豹》的主演查德威克·博斯曼（Chadwick Boseman）去世。去年12月，迪士尼解雇了在《多元宇宙传奇》（Multiverse Saga）中饰演反派的乔纳森·梅杰斯（Jonathan Majors），原因是他因殴打和骚扰

当时的女友而被定罪。《多元宇宙传奇》在故事情节上将把2021年至2027年之间上映的多部电影串联在一起，而梅杰斯的反派是其中的核心角色。

另一个原因与地缘纷争有关。MCU最早的23部电影都在中国这个全球最大的电影市场上映，但在2020年至2022年间一部也没有。（对此中国没有给出明确解释，但可能是因为它在打造本国的电影产业。）虽然这种事实上的禁令现在已经被解除，但如果观众错过了其中几部，就很难理得清整个电影宇宙的剧情了。更糟糕的是，由于Disney+在中国无法使用，粉丝们看不到MCU的电视剧集。

但也有些问题是漫威自己造成的。自2021年以来，MCU平均每年出品3.3部电影和3.7部电视剧——这样的速度似乎让观众、内部创意团队和特效部门都难以招架。对于想要观看一部新剧的观众来说，要先了解过33部电影和11季电视剧实在是太多功课了。目前重点打造的“多元宇宙”还涉及现有电影宇宙之前的电影，更加剧了这一问题。

观众可能最终会对超级英雄失去兴趣，就像他们在上世纪60年代末对西部片产生审美疲劳一样。但就目前而言，这个影视类型还会继续下去。《银河护卫队3》的总票房为8.46亿美元，是2023年票房第四高的电影，并且在CinemaScore上获得了A的评分。索尼出品的动画电影《蜘蛛侠：纵横宇宙》也是去年最受欢迎的电影之一。

开创了漫威扩张之路的迪士尼CEO鲍勃·伊格尔（Bob Iger）已表示过，通过放慢制作速度，漫威系列电影可以重回昔日的辉煌。“我一直觉得，要说质量的话，多产实际上可能是个负面因素。我想事实正是这样。我们有些偏离重点了。”他，将和众多漫威粉丝一道，坚持等待英雄回归。■



Schumpeter

AI can transform education for the better

Meet the companies trying to make it happen

AS PUPILS AND students return to classrooms and lecture halls for the new year, it is striking to reflect on how little education has changed in recent decades. Laptops and interactive whiteboards hardly constitute disruption. Many parents bewildered by how their children shop or socialise would be unruffled by how they are taught. The sector remains a digital laggard: American schools and universities spend around 2% and 5% of their budgets, respectively, on technology, compared with 8% for the average American company. Techies have long coveted a bigger share of the \$6trn the world spends each year on education.

When the pandemic forced schools and universities to shut down, the moment for a digital offensive seemed nigh. Students flocked to online learning platforms to plug gaps left by stilted Zoom classes. The market value of Chegg, a provider of online tutoring, jumped from \$5bn at the start of 2020 to \$12bn a year later. Byju's, an Indian peer, soared to a private valuation of \$22bn in March 2022 as it snapped up other providers across the world. Global venture-capital investment in education-related startups jumped from \$7bn in 2019 to \$20bn in 2021, according to Crunchbase, a data provider.

Then, once covid was brought to heel, classes resumed much as before. By the end of 2022 Chegg's market value had slumped back to \$3bn. Early last year investment firms including BlackRock and Prosus started marking down the value of their stakes in Byju's as its losses mounted. "In hindsight we grew a bit too big a bit too fast," admits Divya Gokulnath, the company's co-founder.

If the pandemic couldn't overcome the education sector's resistance to digital disruption, can artificial intelligence? ChatGPT-like generative AI, which can converse cleverly on a wide variety of subjects, certainly looks the part. So much so that educationalists began to panic that students would use it to cheat on essays and homework. In January 2023 New York City banned ChatGPT from public schools. Increasingly, however, it is generating excitement as a means to provide personalised tutoring to students and speed up tedious tasks such as marking. By May New York had let the bot back into classrooms.

Learners, for their part, are embracing the technology. Two-fifths of undergraduates surveyed last year by Chegg reported using an AI chatbot to help them with their studies, with half of those using it daily. Indeed, the technology's popularity has raised awkward questions for companies like Chegg, whose share price plunged last May after Dan Rosensweig, its chief executive, told investors it was losing customers to ChatGPT. Yet there are good reasons to believe that education specialists who harness AI will eventually prevail over generalists such as OpenAI, the maker of ChatGPT, and other tech firms eyeing the education business.

For one, AI chatbots have a bad habit of spouting nonsense, an unhelpful trait in an educational context. "Students want content from trusted providers," argues Kate Edwards, chief pedagogist at Pearson, a textbook publisher. The company has not allowed ChatGPT and other AIs to ingest its material, but has instead used the content to train its own models, which it is embedding into its suite of learning apps. Rivals including McGraw Hill are taking a similar approach. Chegg has likewise developed its own AI bot that it has trained on its ample dataset of questions and answers.

What is more, as Chegg's Mr Rosensweig argues, teaching is not merely about giving students an answer, but about presenting it in a way that helps

them learn. Understanding pedagogy thus gives education specialists an edge. Pearson has designed its AI tools to engage students by breaking complex topics down, testing their understanding and providing quick feedback, says Ms Edwards. Byju's is incorporating "forgetting curves" for students into the design of its AI tutoring tools, refreshing their memories at personalised intervals. Chatbots must also be tailored to different age groups, to avoid either bamboozling or infantilising students.

Specialists that have already forged relationships with risk-averse educational institutions will have the added advantage of being able to embed AI into otherwise familiar products. Anthology, a maker of education software, has incorporated generative-AI features into its Blackboard Learn program to help teachers speedily create course outlines, rubrics and tests. Established suppliers are also better placed to instruct teachers on how to make use of AI's capabilities.

| *AI for effort*

Bringing AI to education will not be easy. Although teachers have endured a covid-induced crash course in education technology, many are still behind the learning curve. Less than a fifth of British educators surveyed by Pearson last year reported receiving training on digital learning tools. Tight budgets at many institutions will make selling new technology an uphill battle. AI sceptics will have to be won over, and new AI-powered tools may be needed to catch AI-powered cheating. Thorny questions will inevitably arise as to what all this means for the jobs of teachers: their attention may need to shift towards motivating students and instructing them on how to best work with AI tools. "We owe the industry answers on how to harness this technology," declares Bruce Dahlgren, boss of Anthology.

If those answers can be provided, it is not just companies like Mr Dahlgren's that stand to benefit. An influential paper from 1984 by Benjamin Bloom, an educational psychologist, found that one-to-one

tutoring both improved the average academic performance of students and reduced the variance between them. AI could at last make individual tutors viable for the many. With the learning of students, especially those from poorer households, set back by the upheaval of the pandemic, such a development would certainly deserve top marks. ■



熊彼特

人工智能可以让教育变得更好

看看正为此努力的一众公司

随着大中小学生在新的年份回到教室和讲堂，回看近几十年来的教育，其变化之小令人震惊。笔记本电脑和交互式白板基本算不上颠覆。许多父母对孩子的购物或社交方式大惑不解，但对他们的受教育方式却处之泰然。教育部门的数字化仍然落后：美国中小学和大学在技术上的支出分别约占预算的2%和5%，而美国公司的平均支出为8%。长期以来，科技公司一直觊觎着全球每年6万亿美元的教育支出，希望从中获得更大份额。

当新冠疫情迫使中小学和大学停课时，发起数字化攻势的时机似乎来到了眼前。学生纷纷涌向在线学习平台，以填补在Zoom上单调拘谨地上网课所留下的空白。在线辅导供应商Chegg的市值从2020年初的50亿美元跃升至一年后的120亿美元。它的印度同行Byju's在全球各地并购其他供应商，到2022年3月时私人市场估值已飙升至220亿美元。根据数据供应商Crunchbase统计，全球教育类创业公司拿到的风险投资从2019年的70亿美元激增至2021年的200亿美元。

然后，待疫情得到控制之后，课堂又基本恢复如初。到2022年底，Chegg的市值已暴跌至30亿美元。去年年初，随着Byju's亏损加剧，包括贝莱德（BlackRock）和Prosus在内的投资公司开始减记所持Byju's股份的价值。Byju's的联合创始人迪维亚·戈库尔纳特（Divya Gokulnath）承认：“回过头看看，我们扩张得有点太大、太快了。”

如果疫情没能战胜教育部门对数字化颠覆的抵制，那么AI可以吗？ChatGPT这样的生成式AI可以就各种各样的主题机灵地对话，显然很有胜任的气势。它如此聪明，以至于教育工作者开始恐慌，担心学生会在写论文和做作业时用它作弊。2023年1月，纽约市禁止在公立学校使用ChatGPT。然而，作为一种为学生提供个性化辅导和加快批作业试卷等繁琐工作的手段，它引发的热情日益高涨。到5月，纽约又允许这个聊天机

机器人重返课堂。

至于学生，他们正在积极采用这项技术。在去年Chegg对本科生的调查中，五分之二的受访者表示自己用AI聊天机器人辅助学习，其中半数的人每天都用。事实上，AI技术的普及让Chegg等公司陷入尴尬境地。去年5月，Chegg首席执行官丹·罗森斯威格（Dan Rosensweig）向投资者表示，ChatGPT正导致客户流失，公司股价应声暴跌。然而，有充分的理由相信，那些能够善用AI的专业教育公司最终仍将战胜ChatGPT的创造者OpenAI这样的通用型公司，以及其他对教育业务虎视眈眈的科技公司。

首先，AI聊天机器人有胡言乱语的坏习惯，这在教学中可不是个有益的特性。教科书出版商培生（Pearson）的首席教育学家凯特·爱德华兹（Kate Edwards）认为：“学生想从值得信赖的供应商那里获得内容。”该公司不允许ChatGPT和其他AI摄取自己的教材，而是用这些内容来训练自己的模型，并将它嵌入到自己的学习应用套件中。包括麦格劳希尔（McGraw Hill）在内的竞争对手也在采取类似的做法。Chegg同样开发了自己的AI机器人，并用自有的海量题库数据集训练它。

此外，正如Chegg的罗森斯威格所言，教学不仅仅是给学生一个答案，而是要以一种有助于他们学习的方式将答案呈现出来。因此，对教育学的理解使专业教育公司更具优势。爱德华兹表示，培生设计的AI工具会分解复杂的主题、测试学生的理解程度并快速提供反馈，以促进他们投入学习。Byju's在设计AI辅导工具时考虑了学生的“遗忘曲线”，按个性化的间隔来加深记忆。聊天机器人还必须针对不同年龄段的学生量身定制，以避免让他们感到内容过于难懂或过于幼稚。

教育机构往往厌恶风险，因此已经与其建立关系的专业公司将获得额外优势，能够将AI嵌入到对方已经用惯了的产品之中。教育软件开发商Anthology在其Blackboard Learn程序中加入了生成式AI功能，帮助教师快速创建课程大纲、评分标准和测试题目。老牌供应商也更容易指导教师如何利用AI的各种能力。

| 利用AI的努力值得肯定

将AI引入教育并非易事。尽管教师在疫情期间不得已上了一回教育科技速成班，但许多人还跟不上进度。去年在培生对英国教育工作者的调研中，只有不到五分之一的人表示接受过数字化学习工具的培训。许多机构预算紧张，向它们推销新技术将是一场苦战。必须说服那些对AI持怀疑态度的人，而且还需要新的AI工具来检测利用AI作弊的行为。至于这一切对教师的工作意味着什么，棘手的问题将不可避免：他们可能要将注意力转为激励学生，以及指导他们如何最好地利用AI工具。Anthology的老板布鲁斯·达尔格伦（Bruce Dahlgren）宣称：“我们应该就如何利用这项技术为业界提供答案。”

如果能够提供这些答案，受益的就不仅仅是达尔格伦的公司了。教育心理学家本杰明·布鲁姆（Benjamin Bloom）在1984年发表的一篇颇具影响力的论文发现，一对一的辅导既提高了学生的平均学习成绩，又减少了他们之间的成绩差异。AI最终可以让大多数人都负担得起私人教师了。鉴于学生们的学业因疫情干扰而退步——特别是贫困家庭的学生——这样的进展无疑该打个高分。■



Silicon lowlands

Does Europe at last have an answer to Silicon Valley?

ASML, a mighty Dutch tech firm, is at the heart of a critical supply chain

TEN TIMES a second an object shaped like a thick pizza box and holding a silicon wafer takes off three times faster than a manned rocket. For a few milliseconds it moves at a constant speed before being halted abruptly with astonishing precision—within a single atom of its target. This is not a high-energy physics experiment. It is the latest lithography machine dreamed up by ASML, a manufacturer of chipmaking tools, to project nanoscopic chip patterns onto silicon wafers. On January 5th Intel, an American semiconductor giant, became the first proud owner of this technical marvel's initial components for assembly at its factory in Oregon.

Like the outwardly unassuming machine, its Dutch maker is full of surprises. The company's market value has quadrupled in the past five years, to €260bn (\$285bn), making it Europe's most valuable technology firm (see chart 1). Between 2012 and 2022 its sales and net profit both rose roughly four-fold, to €21bn and €6bn, respectively. In late 2023 ASML's operating margin exceeded 34%, staggering for a hardware business and more than that of Apple, the world's biggest maker of consumer electronics (see chart 2).

Such stellar performance, which is set to shine brightly again when ASML reports quarterly results on January 24th, is now routine. The firm holds a monopoly on a key link in the world's most critical supply chain: without its kit it is next to impossible to make cutting-edge chips that go into smartphones and data centres where artificial intelligence (AI) is trained. With global semiconductor sales forecast to double to \$1.3trn by 2032, every big country and every big chipmaker wants ASML's gear. The company has become so important in the Sino-American techno-tussle that, as it

recently emerged, America's government pressed ASML to cancel planned deliveries of even its older machines to China.

Yet ASML's spectacular success is also underpinned by two other, less obvious factors. The company has created a network of suppliers and technology partners that may be the closest thing Europe has to Silicon Valley. And its business model ingeniously combines hardware with software and data. These unsung elements of ASML's success challenge the notion that the old continent is incapable of developing a successful digital platform.

ASML's complex machines perform a simple task. They project chip blueprints onto photosensitive silicon wafers. In 1986, when its first model was delivered, individual transistors measured micrometres and its kit was almost like a glorified photocopier, explains Marc Hijink, a Dutch journalist and author of "Focus—How ASML Conquered the Chip World", a new book. Today, with transistors shrunk by a factor of a thousand, ASML lithography gear is possibly the most sophisticated equipment ever sold commercially.

ASML and its partners pulled off this incredible shrinking trick with engineering that has a science-fiction ring to it. The process starts with powerful lasers incinerating droplets of molten tin, each no thicker than a fifth of a human hair and travelling at more than 250kph. This produces extremely short-wavelength light (extreme ultraviolet, or EUV, in the jargon) which is then reflected by a set of mirrors so smooth that the biggest imperfection is no bigger than the distance grass can grow in a millisecond. To make all this worth a chipmaker's while—the latest model costs more than \$300m—and expose enough chips, the object that holds the wafer, called a "table", has to accelerate faster than a rocket and come to a stop at exactly the right spot.

To get an idea of what it takes to build such a device, pay a visit to a

nondescript factory in Neukölln, a neighbourhood of Berlin. This is where ASML makes, among other things, “mirror blocks”, the main part of a wafer table. These are sturdy pieces of a special ceramic material, a square 8cm thick and measuring about 50cm on each side. Some get polished, measured, repolished, remeasured and so on, for nearly a year—until they are exactly the right shape, including allowances for the fact that they will sag by a few nanometres once installed.

The factory is emblematic of the company’s unusual network of suppliers. Although its owner, Berliner Glas, was acquired by ASML in 2020, it lives halfway between being an independent company and a unit of the Dutch parent. Something similar is true of the 800 or so mostly European firms that help put together ASML’s machines. ASML owns stakes in only a few of them. Yet their interdependence makes them act like a single organisation.

ASML outsources over 90% of what it costs to build one of its marvels and directly employs less than half the estimated 100,000 people the feat requires. This is partly because of its history. When it was spun out of Philips, a Dutch electronics giant, in 1984, ASML seemed stillborn. Its idea to build a “silicon stepper”, the original name of the chip-copying machine, was promising. But it had not much else going for it, in particular no production lines. It instead relied on specialist suppliers, many of them also former Philips units, such as VDL, a contract manufacturer.

The outsourcing is also a function of technology. The different parts of a lithography machine are so cutting-edge that doing it all could overwhelm one firm. “You have to decide where you add the most value and let others do the rest,” says a former ASML insider. Semiconductor economics, too, favours not doing everything yourself. The industry is prone to booms and busts, because demand moves up and down more quickly than chipmakers can install capacity. Prices rise and fall as shortages turn to gluts. Manufacturers of chipmaking gear are exposed to the same cycle. That

makes owning all the assets risky; better to shift some risk to suppliers, who can limit it by catering to customers working to different business cycles.

The required hyper-specialisation prevents the risk-reducing double sourcing that is prevalent in many other industries. In the case of ASML, technical demands are so high and production volumes so low (it shipped 317 machines in 2022) that it would be uneconomical to manage several suppliers for a single part even if they could be found. For such crucial components as lasers and mirrors, which are made by Trumpf and Zeiss, two German firms, respectively, it is impossible. Wayne Allan, who is in charge of sourcing on ASML's board, talks of "co-dependency".

The upshot is that ASML mostly limits itself to being the system's architect. It decides who does what, defines the interfaces between the main parts of its machines ("modules") and carries out research and development. This set-up makes it easier to test the pieces and transport the machines (shipping the latest model to Intel involved 250 crates and 13 containers). It also gives suppliers more freedom, including to experiment with novel technologies.

It all works because ASML has cultivated a culture of trust and transparency while preserving elements of competition. Suppliers are not squeezed to the last penny. Quite the opposite: "We need them to stay healthy," says Mr Allan. Information flows freely throughout the network, particularly between ASML, Trumpf and Zeiss. Engineering teams from different firms work together. Patents are shared, as are some financial data and, sometimes, profits. "At meetings you can't tell who is from which firm," reports a former Zeiss executive.

At the same time, many suppliers compete with each other indirectly, for instance providing similar parts for different generations of ASML's

machines. If a supplier runs into trouble, ASML dispatches a rapid intervention force, sometimes even if such help is not welcome. As a last resort, ASML can buy a supplier, as it did with Berliner Glas.

It is this loosely coupled structure that allowed ASML to outcompete more vertically integrated rivals, reckons Willy Shih of Harvard Business School. Nikon and Canon, two Japanese firms which once led the market for lithography machines, never managed to commercialise EUV kit. (Canon is trying to stage a return with “nanoimprint” lithography, which physically stamps chip designs onto wafers.)

ASML is now entrenching this dominance by complementing its hardware with software and data. When real rockets take off, their trajectory is wobbly and needs to be smoothed out by a guidance computer, which collects data to predict and adapt their course. A wafer table in a lithography machine is similarly likely to miss the mark at first. The same is true of the rest of the device. It is only with the help of lots of data and machine learning, a type of AI, that they can be fine-tuned—and made more accurate. This is rapidly turning ASML into an AI platform.

Once Intel gets all the modules for its new machine, it will take about two weeks to put the thing together. Adapting it to its new location will take a few months. Bits may have moved in transport, gravity may be slightly different in Oregon from the Netherlands and other kit nearby may create interference. Tests will collect data and trigger adjustments. “We have thousands of knobs we can turn to put it into a perfect state,” says Jos Benschop, who is in charge of technology at ASML.

ASML also uses the data from one machine to turn the knobs of others. Of the roughly 5,500 devices it has sold since its founding 39 years ago, 95% are still in operation and many send data home to headquarters. That will make its products even better, leading to more chipmaking, which

generates even more data—and so on, in a “flywheel” more typically associated with digital services such as internet search. Even if Canon, Nikon or a Chinese competitor finally managed to build EUV machines as powerful as ASML’s, it would not be able to catch up with the Dutch firm, argues Pierre Ferragu of New Street Research, a firm of analysts. “It’s mathematically impossible, as long as ASML keeps collecting data from all the installed base.”

If rivals cannot topple ASML, can anything? Maybe physics. Even with the best AI, you can’t shrink transistors for ever (certainly not in a commercially viable way). If technical requirements become too otherworldly the supplier network may unravel. Or maybe economics. Chipmakers may recoil at ASML’s data hunger, which extends to other linked devices in their factories. Some are pushing back against its digital expansion, insiders say.

Then there is geopolitics. ASML’s share price dipped after news broke about the cancelled deliveries to China. The worry is less over lower sales; ASML cannot build its machines fast enough anyway. Of greater concern is the risk that strict export controls could in time push China to build its own chipmaking-gear industry. That could one day threaten ASML’s position at the centre of the sector. For the time being, though, the company’s network and its network effects remain indomitable. Who said Europe couldn’t do tech? ■



硅低地

面对硅谷，欧洲终于有了拿得出手的东西？

强大的荷兰科技公司阿斯麦是一条关键供应链的核心【深度】

每秒十次，一个形似厚披萨盒、装载着一个硅片的物体以比载人火箭快三倍的速度起飞。它以恒定速度移动几毫秒后，便以惊人的精确度——精确到目标中的某一个原子——突然停止。这不是高能物理实验，而是由芯片制造设备生产商阿斯麦（ASML）设计的最新型光刻机，用于将纳米级芯片图案投射到硅片上。1月5日，美国半导体巨头英特尔骄傲地拿到了这项技术奇迹的初始组件，将在其位于俄勒冈的工厂进行组装。

就像这台外观不起眼的机器一样，它的荷兰制造商充满了惊喜。过去五年，这家公司的市值翻了两番，达到2600亿欧元（2850亿美元），成为欧洲最有价值的科技公司（见图表1）。在2012年至2022年间，其销售额和净利润均大约增长了四倍，分别达到210亿欧元和60亿欧元。2023年底，阿斯麦的营业利润率超过了34%，这对于硬件业务而言堪称惊人，而且超过了世界上最大的消费电子产品制造商苹果（见图表2）。

这样的亮眼表现如今已成为常态，到1月24日阿斯麦公布季度业绩时还会再次大放异彩。该公司在世界上最关键的供应链中占据了垄断地位：没有它的设备，几乎不可能制造出用于智能手机和培训人工智能（AI）的数据中心的尖端芯片。预计到2032年全球半导体销售额会翻倍至1.3万亿美元，因此每个大国和每个大型芯片制造商都想要阿斯麦的设备。这家公司在中美科技角力中变得非常关键，近日有消息称美国政府向它施压，要求它取消向中国交货的计划，哪怕是型号较旧的设备。

但阿斯麦的惊人成功还得益于另外两个不那么显见的因素。该公司建立了一个供应商和技术合作伙伴网络，它可能是欧洲所具有的最接近硅谷的东西。此外，其商业模式巧妙地将硬件与软件和数据结合在一起。阿斯麦这些鲜为人知的成功要素挑战了认为旧大陆无法开发出成功的数字平台的看法。

阿斯麦的复杂机器执行的是一个简单的任务。它们将芯片设计图投射到感光硅片上。荷兰记者、新书《聚焦——阿斯麦如何征服芯片世界》（Focus—How ASML Conquered the Chip World）的作者马克·希因克（Marc Hijink）解释说，当它的第一款机型在1986年交付时，单个晶体管的尺寸是微米级的，其设备几乎就是个名字很唬人的复印机。而今天，随着晶体管缩小了一千倍，阿斯麦的光刻机可能是有史以来最精密的商用设备。

阿斯麦及其合作伙伴利用颇具科幻色彩的工程技术完成了这一令人难以置信的微缩戏法。这个过程从高能激光燃烧熔融锡滴开始，每个锡滴的直径不超过人头发丝的五分之一，以超过250公里的时速移动。这产生了极短波长的光（行话叫极紫外光，缩写为EUV），这些光会被一组非常光滑的镜子反射——光滑到最大瑕疵不及草在一毫秒内生长的长度。为了使这一切值得芯片制造商的投资——最新型号的成本超过3亿美元——并能曝光足够的芯片，承载硅片的“台”必须比火箭加速得更快，并且精准地停在正确的位置。

要了解打造这样一台设备要费多大劲，不妨参观一下柏林新科尔恩区（Neukölln）一家不起眼的工厂。这里是阿斯麦制造硅片台的主要部分——“反射镜”等部件的地方。这些反射镜是由特殊陶瓷材料制成的坚固扁方块，厚度为8厘米，方块的每边长约50厘米。其中一些经过打磨、测量，再打磨、再测量，如此持续近一年时间——直到它们达到完全正确的形状，包括要考虑到安装后会下凹几纳米的情况。

这家工厂体现了阿斯麦不同寻常的供应商网络。尽管工厂所有者Berliner Glas在2020年被阿斯麦收购，但它的性质介于一家独立公司和荷兰母公司的一个部门之间。在大约800家帮助阿斯麦生产组件的公司中，情况都差不多如此，其中大部分是欧洲公司。阿斯麦只持有其中少数公司的股份。然而，它们彼此间的相互依赖却让它们像一个单一组织那样运作。

阿斯麦将其神奇产品超过90%的制造过程外包了出去。如果是独立完成这样的奇迹估计需要10万人，而阿斯麦的直接雇员还不到这个数字的一半。

这一定程度上是因为它的历史。1984年，当它从荷兰电子巨头飞利浦分拆出来时似乎已经夭折。它打造“硅步进机”（光刻机的最初名称）的想法很有前景。但除此之外，它几乎没有其他资源，特别是没有生产线。它转而依赖专业供应商，其中许多也是出自飞利浦，例如合同制造商VDL。

外包也是出于技术上的需要。光刻机的各个部件都是尖端产品，全部自行制造可能会压垮一家公司。“你必须决定在哪里增加最多价值，并让其他人做其余部分。”一位前阿斯麦内部人士说。半导体经济学也不赞成自己做所有事情。由于需求的上下波动比芯片制造商增加产能更快，这个行业容易出现繁荣和萧条。价格随着短缺和过剩的交替而涨跌。芯片制造设备的制造商也面临同样的周期。这使得拥有所有资产变得风险重重；更好的做法是将一些风险转移给供应商，他们可以通过服务具有不同商业周期的客户来限制风险。

由此所需的超级专业化导致在许多其他行业通行的、能降低风险的双重采购并不可行。在阿斯麦的案例中，技术要求如此之高，产量如此之低（2022年仅出货317台机器），即使能为某个部件找到多家供应商，管理这几个供应商也是不经济的。对于关键组件，比如分别由德国的通快（Trumpf）和蔡司（Zeiss）制造的激光器和反射镜，这纯粹就是不可能的。阿斯麦董事会里负责采购的韦恩·艾伦（Wayne Allan）谈到了“相互依赖”。

结果是阿斯麦基本上只让自己充当系统的架构师。它决定谁做什么，定义机器主要部分（“模块”）之间的接口，并进行研发。这种安排使得测试零件和运输机器变得更加容易（将最新型号的设备运送到英特尔用了250个板条箱和13个集装箱）。它还给予供应商更多自由，包括尝试全新技术。

这一切之所以行得通，是因为阿斯麦培养了一种信任和透明的文化，同时保留了竞争元素。供应商不会被压榨到只剩最后一分钱。恰恰相反，“我们需要它们保持健康。”艾伦说。信息在整个网络中自由流动，尤其是在阿斯麦、通快和蔡司之间。不同公司的工程团队共同工作。他们共享专利，也共享一些财务数据，有时还共享利润。“在会议上，你都分辨不出

谁是哪家公司的。”一位前蔡司高管说。

与此同时，许多供应商之间也有间接的竞争，例如为阿斯麦不同代的设备提供类似的部件。如果供应商遇到困难，阿斯麦会组织力量快速介入，哪怕有时候这种帮助并不受欢迎。万不得已时，阿斯麦可以收购供应商，就像它收购Berliner Glas那样。

正是这种松散耦合的结构使阿斯麦能够击败更垂直一体化的竞争对手，哈佛商学院的史兆威（Willy Shih）认为。曾经在光刻机市场领先的两家日本公司尼康和佳能从未能成功商业化生产EUV设备。（佳能正试图通过“纳米压印”光刻技术杀回来，这种技术把芯片设计实体压印到硅片上。）

阿斯麦现在正通过用软件和数据补充硬件来巩固自己的优势。在真正的火箭起飞时，它们的轨迹是摇摆的，需要通过导航计算机来平滑，该计算机收集数据以预测并调整火箭航向。光刻机中的硅片台一开始也可能会错过正确位置。设备的其他部分也是如此。只有在大量数据和机器学习这种人工智能的帮助下，它们才能进行微调并变得更加精确。这正在迅速将阿斯麦转变为一个AI平台。

一旦英特尔拿到其新机器的所有模块，将需要大约两周时间把它们组装起来。之后再基于新地点花几个月调试。零部件在运输过程中可能会移动，俄勒冈州的重力可能与荷兰略有不同，附近的其他设备可能会造成干扰。工作人员将通过测试收集数据做出调整。“我们有成千上万个旋钮，我们可以转动它们来让设备达到完美状态。”阿斯麦负责技术的乔斯·本斯霍普（Jos Benschop）说。

阿斯麦还用源自一台机器的数据来转动其他机器的旋钮。在其成立39年来售出的大约5500台设备中，有95%仍在运行，许多设备把数据发送回总部。这将优化其产品，从而制造出更多的芯片，继而产生更多的数据——如此循环，这是在数字服务（如互联网搜索）中更为典型的“飞轮”效应。即使佳能、尼康或中国竞争对手最终能制造出与阿斯麦一样强大的EUV光刻机，它们也无法赶上这家荷兰公司，New Street Research的分析师皮埃

尔·费拉古（Pierre Ferragu）认为。“这在数学上是不可能的，只要阿斯麦继续从所有已安装设备收集数据。”

如果竞争对手无法击败阿斯麦，那还有什么可以击败它吗？也许是物理学。即便拥有最好的AI，你也不能永远缩小晶体管（要在商业上可行是必然做不到的）。如果技术要求变得太过于超现实，供应商网络可能会解体。或者也许是经济学。芯片制造商可能会厌烦阿斯麦对数据的渴望，这种渴望扩展到了它们工厂中的其他关联设备。知情人士表示，有人在反对它的数字扩张。

然后还有地缘问题。在取消向中国交付设备的消息传出后，阿斯麦的股价下跌。人们担心的主要还不是销售额下降，因为无论如何阿斯麦也无法足够快地制造设备。更大的担忧是严格的出口管制可能最终促使中国建立自己的芯片制造设备产业。这有朝一日可能威胁到阿斯麦在该行业的核心地位。然而，就目前而言，该公司的网络及其网络效应仍然不可撼动。谁说欧洲玩不转科技的？ ■



Bartleby

When your colleagues are also your rivals

How managers should balance competition and co-operation

THE MODERN company exalts both competition and co-operation. Competition is the defining feature of markets; inside organisations, too, employees compete for limited resources. Sometimes that contest is obvious, as when performance is openly ranked or there is a race for a specific job. Sometimes it is left unspoken: there is only so much money to go round and only so many promotion opportunities on offer. Either way, competition is always there.

Yet the reason firms exist is to co-ordinate the activities of many actors in pursuit of common goals. Departments and teams are expected to work together. Collaborative behaviour is usually celebrated. Companies dole out awards for the most helpful co-workers, not the Macbeth prize for the colleague most likely to murder you in your sleep.

Rivalry and teamwork can go together nicely. A paper published in 2022 by Eric VanEpps of the University of Utah, Einav Hart of George Mason University and Maurice Schweitzer of the University of Pennsylvania looked at the best way to handle an old conundrum. To make a good impression on the higher-ups, you need to highlight your own achievements. But bragging about how great you are is not a recipe for being liked. A strategy of taking the credit for some things and doling out praise to colleagues for others resolved this problem.

It is not easy for managers to strike the right balance between encouraging contests and collaboration. (You just need to hear the word “co-opetition” to know how ugly things can get.) Competition can spur more effort but it can also have unintended consequences.

A recent study by Eddy Cardinaels of Tilburg University and Christoph Feichter of the Vienna University of Economics and Business asked supervisors to use a forced performance-ranking system to assess workers' creativity. Forced rankings require managers to assign employees to given places on a scale: if there are ten workers, say, then one must come top and one must come tenth. This approach just stressed everyone out—a bit like bellowing “relax” in someone's face. People tried harder but they also became less creative.

In a literature review published in 2020 Gavin Kilduff of the Stern School of Business at New York University, Blythe Rosikiewicz of West Chester University and Christopher To of Rutgers University concluded that competition is more likely to backfire when people feel threatened: for example, when the costs of losing are high or when people are competing against others known to be better at the task in question. But even when the stakes are low, explicit competition can backfire.

In an experiment conducted by Jeffrey Carpenter of Middlebury College and his co-authors in 2007, participants were asked to stuff envelopes. When people were paid a bonus for stuffing the most envelopes, they worked harder than if they got a flat per-envelope fee. But when they were also given a chance to sabotage their peers to get ahead, the (correct) expectation that they themselves would be sabotaged made people work less hard than if they had got piece-rate pay.

Such behaviour arises because many people—and many of them men, since women tend to be less taken by the prospect of all-out competition—like winning for its own sake. This organic aspect to competition also shows up in rivalries between individuals. Workers tend naturally to benchmark themselves against their peers in the race for status and seniority; they don't need an excuse.

These specific rivalries can be especially motivating. A paper from 2018 by Adam Galinsky and Brian Pike of Columbia Business School and Mr Kilduff found that teams in a range of American sports performed better the year after an intense rival did well in tournaments. In another study, Lisa Ordóñez of the University of Arizona and Messrs Kilduff, Schweitzer and To analysed American-football games and found that teams were more likely to take risky on-field decisions against fierce rivals. Particular opponents encourage greater risk-taking than generic competition, at least if you are a very large man in tights.

All of which argues for a restrained approach to encouraging competition. Balance individual incentives with group ones. If you are going to rate performance, make sure the measures are clear, objective and fair. Think about when risk-taking is more desirable (sales, say) and less desirable (clinical trials). By their nature organisations crackle with competitiveness. Adding a bit of fuel to the fire can be fine. Spraying petrol everywhere is unnecessary. ■



巴托比

当你的同事也是你的竞争对手

管理者应该如何平衡竞争与合作

现代公司既推崇竞争也推崇合作。竞争是市场的本质特征；在企业内部，员工也在争夺有限的资源。有时候，竞争是显而易见的，比如公开对业绩进行排名，或者角逐某个职位。有时候竞争又是心照不宣的：可供分配的钱只有这么多，摆在那里的晋升机会也只有这么多。无论显性还是隐性，竞争一直都在。

但是，公司之所以存在就是为了协调众多参与者的行动以追求共同的目标。部门和团队都是要通力合作的。协作行为通常会得到赞美。公司会给最乐于协作的员工颁奖，而不会向最有可能趁你睡着时向你捅刀子的同事颁发“麦克白奖”。

同事竞争和团队合作可以并行不悖。犹他大学（University of Utah）的埃里克·万尼普斯（Eric VanEpps）、乔治梅森大学（George Mason University）的埃纳夫·哈特（Einav Hart）以及宾夕法尼亚大学（University of Pennsylvania）的莫里斯·施魏策尔（Maurice Schweitzer）在2022年发表了一篇论文，探讨了一个老大难问题的最佳解决之道。为了给上司们留下好印象，你需要突显自己的成就。但自吹自擂并不能让别人喜欢你。解决办法是可以把一些成绩归功于自己，同时不吝把同事赞美一番。

对于管理者来说，在鼓励竞争和合作之间取得适当的平衡并非易事。（你只需要听到“竞合”这个怪词，就知道事情会变得多麻烦。）竞争可以激励人们更加努力，但也可能带来意想不到的后果。

荷兰的蒂尔堡大学（Tilburg University）的埃迪·卡迪纳尔斯（Eddy Cardinaels）和维也纳经济大学（Vienna University of Economics and Business）的克里斯托夫·费希特（Christoph Feichter）近期的一项研究

让主管使用强制绩效排名法来评估员工的创造力。强制排名要求管理者按照一定比例给每位员工排位：比如，如果有十名员工，那么必须有一人排在第一名，有一人排在第十名。这种方法只会让每个人都压力爆棚——有点像冲着某人大吼“放松”。人们是比以前更努力了，但同时创造力也下降了。

在2020年发表的一篇文献综述中，纽约大学斯特恩商学院（Stern School of Business）的加文·基尔达夫（Gavin Kilduff）、西切斯特大学（West Chester University）的布莱思·罗西凯维奇（Blythe Rosikiewicz）以及罗格斯大学（Rutgers University）的陶家页（Christopher To）得出结论，当人们感到威胁时，竞争更可能产生适得其反的后果：例如，当失败的代价很高，或者明知竞争对手比自己更擅长这项任务时。但即使风险不高，赤裸裸的竞争也可能带来反效果。

2007年，美国明德学院（Middlebury College）的杰弗里·卡彭特（Jeffrey Carpenter）及其合著者做了一个实验，让参与者往信封里装东西。当装完的信封数量最多的人能获得一笔额外的奖金时，人们工作起来会比每装一个信封得到一笔固定报酬时更努力。但是，如果他们还有机会给同事使绊子而让自己领先时，他们会想到自己也会被别人使绊子（也确实如此），这让他们的工作积极性比拿计件工资时更低。

许多人之所以有这样的行为，纯粹是好胜心使然——这其中很多是男性，因为女性往往不太喜欢你死我活的竞争。这种竞争的本能也表现在特定个体之间的对峙中。在对地位和资历的角逐中，员工往往会自然而然就拿同事当作对照标准；他们不需要理由。

这样的具体的对手有可能格外激励人。哥伦比亚大学商学院（Columbia Business School）的亚当·加林斯基（Adam Galinsky）、布莱恩·派克（Brian Pike）和基尔达夫在2018年发表的一篇论文中发现，在美国一系列不同的体育锦标赛中，参赛队伍在自己的劲敌取得好成绩后在下一年的赛事中表现更出色。在另一项研究中，亚利桑那大学（University of Arizona）的丽莎·奥多涅斯（Lisa Ordóñez）和基尔达夫、施魏策尔以及

陶家页分析橄榄球比赛后发现，球队在面对凶猛的对手时更有可能在赛场上做出冒险的决定。而当遭遇某个特定的对手时，会比泛泛的竞争激发出更多的冒险行为——至少对于穿着紧身衣裤的大块头们来说是如此。

所有这些都表明，鼓励竞争要适度。要平衡个人动机和团队目标。如果要对员工表现做评估，要确保评价标准清晰、客观和公平。要考虑什么时候比较值得冒险（比如销售）、什么时候不值得（比如临床试验）。企业的根本性质决定了它燃烧着竞争之火。往火里加点燃料可能还行，到处浇汽油就没必要了。■



The other Saudi gold

Saudi Arabia wants to be the Saudi Arabia of minerals

The kingdom plans to be digging up plenty more than oil

IN WA'AD AL-SHAMAL, 1,200km north of Riyadh, the Saudi capital, phosphate is extracted and bathed in chemicals to turn it into an acid. From there it is shipped 1,500km east by rail to the port of Ras Al-Khair. The stuff is then made into fertiliser or its precursor, ammonia, and sails west to Brazil, south to Africa and east to India and Bangladesh, where it ends up with farmers who, according to Ma'aden, the state mining firm which runs the project, grow 10% of the world's food. The venture is vast. Its sales and domestic investment are equivalent to about 2% of the kingdom's non-oil GDP. Another similar one will soon start shipping the equivalent of another 1%.

Phosphate is not the only mineral resource Saudi Arabia is eyeing to fuel its post-oil future. On January 10th the government revised its estimate of the value of its buried mineral wealth from \$1.3trn to \$2.5trn. This includes deposits of gold, copper and zinc. By the standards of Saudi oil riches, worth perhaps \$20trn at today's prices, that looks modest. By any other measure, it is gargantuan.

Muhammad bin Salman, the kingdom's crown prince and de facto ruler, wants the country to become as indispensable for minerals, including those needed for the energy transition, as it is today for black gold. He intends to achieve this without embracing the resource nationalism that has gripped other countries, from America to Chile and China. Intrigued, mining bosses and ministers from around 80 countries had assembled in Riyadh as we published this, for the country's Future Minerals Forum. As if to prove its commitment to openness, the kingdom has signed agreements both with Russia and with America's Export-Import Bank. It expects deals

worth \$20bn to be sealed at the event.

Part of the strategy looks abroad. Saudi Arabia has set up Manara Minerals, a venture backed by Ma'aden and the Saudi sovereign wealth fund. Manara will invest up to \$15bn in stakes in foreign mines. Last year it paid close to \$3bn for a 10% stake in the base metals business of Vale, a Brazilian mining giant. The Saudis are “putting their money where their mouth is,” says Eduardo Bartolomeo, Vale's boss.

The bigger bet, as the phosphate complex in Wa'ad Al-Shamal shows, is domestic. Saudi Arabia is pitching itself as an investment destination (the campaign includes ads in such unlikely places as the London Underground). In the past few years it has formed a new ministry for industry and mineral resources, waived duties on imported machinery and raw materials, reduced licence fees and royalties, offered state support for salaries and subsidised rents. It has also replaced an arcane mining law with one more like the investor-friendly codes in Australia, Botswana and Canada. Licences that took years to secure are now handed out in two months.

The result has been a sharp rise in active licences—to around 2,300, a fifth more than two years ago. About 700 of these are for exploration. Some are going to foreigners. Medium-sized or specialist outsiders such as Barrick Gold and Eurasian Resources Group have received licences to explore or have partnerships with Ma'aden. “I would rather have 50% of something than 100% of nothing,” says Robert Wilt, Ma'aden's chief executive.

“To draw big players in, Saudi Arabia will need big discoveries,” says Mark Bristow, boss of Barrick Gold. To that end it is investing over \$180m in incentives for exploration. The Saudi Industrial Development Fund, a government vehicle, offers to finance up to three-quarters of project costs. The kingdom is also bankrolling a \$200m effort to map its geology and

create a database of resources, on top of \$500m it spent on an earlier survey. Ma'aden is doing more prospecting, too, Mr Wilt says.

The government is also training a cadre of geoscientists and engineers. Such professionals are in short supply not just in Saudi Arabia but everywhere. No amount of money can get you all the people you need today, says John Bradford of the Colorado School of Mines. To ensure Saudi Arabia can get them tomorrow, it has teamed up with American think-tanks in mining research and is working with Mr Bradford's institution to create training programmes. In November Ma'aden endowed a new undergraduate degree in mining science and engineering at King Fahd University of Petroleum and Minerals.

The princely plan may misfire. Abroad, it could run into the sort of resource nationalism it itself eschews. Partners in Africa, bruised by decades of outsiders shipping off resources without boosting development, insist that this time benefits trickle down to their economies. A partnership with Saudi Arabia must be "not just extracting the ore and taking it away", says Henry Dele Alake, Nigeria's solid-minerals minister. It would require investments in Nigerian processing and factories.

At home, Prince Muhammad's short timelines are, sceptical executives note, at odds with those typical of prospecting, mine development and mining education, all of which take years. Unlike phosphate deposits, metal ores from deeper underground are harder to extract quickly. A harsh summer shuts down work for safety reasons, halting projects for three or four months a year. Little has been done to realise Saudi Arabia's potential in power-hungry processing and refining, where it could excel thanks to plentiful energy.

Last, turning the Saudi vision into reality requires a radical shift among the world's miners. In an unpredictable world, many prefer to shovel profits to

shareholders rather than into risky new projects. To change this, the prince will need all his powers of persuasion. ■



另一种沙特黄金

沙特希望成为矿业大国

这个王国计划大量开采的不止石油

在沙特首都利雅得以北1200公里处的瓦阿德·沙马尔（Wa'ad Al Shamal），人们将磷酸盐提取出来，浸泡在化学品中，使它变成一种酸。再沿铁路向东行驶1500公里，运送到达拉斯海尔港（Ras Al-Khair）。然后再把这种酸制成肥料或其前体——氨，接着扬帆向西驶向巴西，向南驶向非洲，向东驶向印度和孟加拉国，最后到达农民手中。据负责该项目的国有矿业公司沙特阿拉伯矿业公司（Ma'aden）称，这些农民种植了世界上10%的食物。这个项目规模庞大。其销售额和国内投资约相当于沙特非石油GDP的2%。另一个类似的项目不久也将开始交付另外的1%。

磷酸盐并不是沙特阿拉伯希望能为其后石油未来提供燃料的唯一矿物资源。1月10日，沙特政府将其未开发矿藏的估价从1.3万亿美元修订为2.5万亿美元。其中包括黄金、铜和锌矿。按照沙特石油财富（以今天的价格计算约为20万亿美元）的标准，这看起来并不高。但以任何其他标准衡量，这都是一笔巨大的财富。

沙特王储、实际统治者穆罕默德·本·萨勒曼希望，今后世界要获取矿物（包括能源转型所需的那些矿物）将离不开沙特，就像今天说到“黑色黄金”就绕不开沙特一样。他意图在追寻这一目标时不像从美国到智利再到中国的其他国家那样诉诸资源民族主义。这吸引了来自80个国家的矿业老板和部长们在本文发表之际齐聚利雅得，参加该国的未来矿产论坛（FMF）。似乎是为了证明自己对开放的承诺，沙特与俄罗斯和美国的进出口银行都签署了协议。它预计此次论坛将敲定价值200亿美元的交易。

其战略一定程度上着眼于国外。沙特成立了马纳拉矿业公司（Manara Minerals），该项目受到沙特阿拉伯矿业公司和沙特主权财富基金支持。马纳拉将投资最多150亿美元参股外国矿山。去年，它以近30亿美元购入巴西矿业巨头淡水河谷（Vale）基本金属业务10%的股份。淡水河谷的老

板爱德华多·巴托洛梅奥（Eduardo Bartolomeo）说，沙特人“正拿出真金白银证明自己的态度”。

正如瓦阿德·沙马尔的磷酸盐产业综合体所显示的，更大的赌注还是放在国内。沙特阿拉伯正以投资目的地的名号来推销自己，其宣传活动包括在伦敦地铁等偏门的地方投放广告。在过去几年中，沙特成立了新的工业和矿产资源部，免除了进口机械和原材料的关税，降低了许可证费用和特许权使用费，并提供政府工资资助和租金补贴。它还用一部更接近澳大利亚、博茨瓦纳和加拿大那样的投资者友好型法规取代了一部晦涩的采矿法。过去需要数年才能获得的许可证现在只需两个月。

结果就是有效许可证数量急剧上升，达到约2300个，比两年前增加了五分之一。其中约700个是勘探许可证。一些许可证由外国人获得。巴里克黄金公司（Barrick Gold）和欧亚资源集团（Eurasian Resources Group）等中型或专业的外来公司或是获得了勘探许可证，或是与沙特阿拉伯矿业公司建立了合作伙伴关系。沙特阿拉伯矿业公司首席执行官罗伯特·威尔特（Robert Wilt）说：“能得到一半比一点也得不到要好。”

“要吸引大玩家进入，沙特阿拉伯需要有重大的发现。”巴里克黄金公司的老板马克·布里斯托（Mark Bristow）说。为此，沙特正投资超过1.8亿美元激励勘探工作。政府投资工具沙特工业发展基金（Saudi Industrial Development Fund）提供多达四分之三的项目成本融资。此外，沙特还出资2亿美元绘制地质图并建立资源数据库，此前已在一项勘测工作中花费了5亿美元。威尔特说，沙特阿拉伯矿业公司也在进行更多的勘探工作。

沙特政府还在培训一批地质学家和工程师。不仅在沙特，这类专业人员在任何地方都十分紧缺。科罗拉多矿业学院（Colorado School of Mines）的约翰·布拉德福德（John Bradford）说，砸再多的钱也得不到今天就需要的所有人才。为了确保明天能得到这些人才，沙特已与美国矿业研究方面的智库合作，并与布拉德福德的学院合作制定培训计划。11月，沙特阿拉伯矿业公司在法赫德国王石油矿业大学（KFUPM）资助了一个新的采矿科学和工程本科专业。

沙特王储的宏伟计划可能不会奏效。在国外，它可能会撞上自己所回避的那种资源民族主义。数十年来，外来者将资源运出非洲，却没有促进非洲的发展，严重挫伤了非洲合作伙伴的信心。这一回，它们坚持要求合作要惠及本国经济。尼日利亚固体矿产部长亨利·德勒·阿拉克（Henry Dele Alake）说，沙特来合作“绝不能只是把矿石开采出来后运走”。它将要求对尼日利亚的加工业和工厂进行投资。

而在国内，持怀疑态度的高管指出，穆罕默德王储给出的时间表很紧迫，不符合勘探、矿山开发和矿业教育培训通常的时间表，这些工作全都需要耗时数年。与磷酸盐矿藏不同，地下深处的金属矿石更难快速开采。出于安全考虑，酷热的夏季会停工，导致项目一年中会中断三四个月。沙特在实现自身在加工和精炼上的潜力方面鲜有作为，这些工作耗能很高，能源丰富的沙特原本可以大显身手。

最后，将沙特的愿景变为现实还需要全球矿业公司做出重大转变。在这个变幻莫测的世界里，它们中的许多更愿意把利润分给股东，而不是投入到有风险的新项目中。要改变这一点，沙特王储可得使出浑身解数来说服人了。■



The Chinese EV onslaught

An influx of Chinese cars is terrifying the West

But it should keep its markets open to cheap, clean vehicles

IS CHINA ABOUT to unleash another wave of deindustrialisation on the rich world? About 1m American manufacturing workers lost their jobs to Chinese competition in 1997-2011, as the country integrated into the global trading system and began shipping cheap goods overseas. This “China shock” has since been blamed for everything from rising deaths among working-class Americans to the election of Donald Trump. The rejection of liberal attitudes to trade also explains why politicians embrace industrial policy today. Now China’s carmakers are enjoying an astonishing rise. That stokes fears of another ruinous shock. In fact, the successes of Chinese cars should be celebrated, not feared.

Just five years ago China shipped only a quarter as many cars as Japan, then the world’s biggest exporter. Earlier this month the Chinese industry claimed to have exported over 5m cars in 2023, exceeding the Japanese total. China’s biggest carmaker, BYD, sold 0.5m electric vehicles (EVs) in the fourth quarter, leaving Tesla in the dust. Chinese EVs are so snazzy, whizzy and—most important—cheap that the constraint on their export today is the scarcity of vessels for shipping them. As the world decarbonises, demand will rise further. By 2030 China could double its share of the global market, to a third, ending the dominance of the West’s national champions, especially in Europe.

This time it will be even easier for politicians to pin the blame for any Western job losses on Chinese foul play. A frosty geopolitical climate will feed the sentiment that subsidised production unfairly puts Western workers on the scrapheap. And there have certainly been subsidies. Since the launch of its “Made in China” agenda in 2014, China has brazenly

disregarded global trading rules, showering handouts on its carmakers. It is hard to be precise about the value of the underpriced loans, equity injections, purchase subsidies and government contracts Chinese firms enjoy. But by one estimate, total public spending on the industry was in the region of a third of EV sales at the end of the 2010s. These subsidies come on top of the ransacking of technology from joint ventures with Western carmakers and Western and South Korean battery-makers.

The temptation will therefore be for rich-world policymakers to shield their carmakers from the onslaught of state-backed competition. In October the European Commission opened an investigation into Chinese cars. President Joe Biden is said to be considering increasing tariffs on them, even though America's carmakers, protected by a 27.5% levy and handouts from the Inflation Reduction Act, currently face little Chinese competition. Yet locking out Chinese cars would be a mistake. The potential gains to the West from a ready supply of cheap, green vehicles are simply enormous—and dwarf the cost of disruption and the dangers it brings.

One reason is that the market for cars is going to be upended, regardless of trade with China. In 2022, 16-18% of new cars sold around the world were electric; in 2035 the EU will ban the sale of new cars with internal-combustion engines. Though firms are retaining their workers as they switch to making EVs, the process is less labour-intensive. Much as the first China shock was responsible for less than a fifth of total manufacturing job losses occurring at the time—many of which were attributable to welcome technological advances—so too there is a danger of confusing disruption caused by the shift to EVs with that caused by Chinese production of them.

Next consider the gains from letting trade flow. Vehicles are among people's biggest purchases, accounting for about 7% of American consumption. Cheaper cars mean more money to spend on other things, at a time when real wages have been squeezed by inflation. And Chinese cars are not only

cheap; they are better-quality, particularly with respect to the smart features in EVs that are made possible by internet connectivity. Nor does the existence of a carmaking industry determine a country's economic growth. Denmark has among the world's highest living standards without a carmaker to speak of. Even as cars roll off Chinese assembly lines, the economy is spluttering—in part because it has been so distorted by subsidies and state control.

Last, consider the benefits to the environment. Politicians around the world are realising just what a tall order it is to ask consumers to go green, as a backlash against costly emissions-reductions policies builds. EVs, too, are currently more expensive than gas-guzzling cars (even if their running costs are lower). Embracing Chinese cars with lower prices could therefore ease the transition to net-zero emissions. The cheapest EV sold in China by BYD costs around \$12,000, compared with \$39,000 for the cheapest Tesla in America.

What about the risks? The threat to industry from cheap imports is usually overblown. The lesson from the rise of Japanese and South Korean carmakers in the 1980s is that competition spurs local firms to shift up a gear, while the entrants eventually move production closer to consumers. Already, BYD is opening a factory in Hungary and many Chinese carmakers are scouting for sites in North America. Meanwhile the likes of Ford and Volkswagen are racing to catch Chinese firms. Last year Toyota said a breakthrough in its “solid state” technology would let it slash the weight and cost of its batteries.

Another worry is national security. Depending entirely on China for batteries, whose importance to electrified economies will go far beyond cars, would be risky. It is also possible that EVs, which are filled with chips, sensors and cameras could be used for surveillance. (China has banned even locally made Teslas from some government properties.) But so long as

presidents and spooks can travel in vehicles made in the West or by its allies, there is little reason to fear consumers sporting Chinese wheels; they can adjudicate personal-privacy concerns themselves and locally made cars will be easier to inspect.

Policymakers should therefore curb their protectionist instincts and worry only in the unlikely event that Western carmakers implode altogether. A hefty market share for Chinese carmakers that invigorates wider competition, however, is not to be feared. If China wants to spend taxpayers' money subsidising global consumers and speeding up the energy transition, the best response is to welcome it. ■



【首文】中国电动车猛攻

中国汽车大量涌入令西方恐惧

但西方市场应对廉价的绿色车辆保持开放

中国是否即将在富裕国家掀起另一波去工业化浪潮？在1997年至2011年间，随着中国融入全球贸易体系并开始向海外输送廉价商品，约有100万美国制造业工人因来自中国的竞争而失业。此后，从美国工人阶级死亡人数上升到特朗普当选，一切都被归咎于“中国冲击”。对自由贸易观的排斥拒绝也导致政客们如今纷纷采取产业政策。目前，中国的汽车制造商正在惊人崛起中。这引发了人们对另一场毁灭性冲击的担忧。事实上，中国汽车的成功值得庆祝，而不是忧惧。

仅在五年前，中国的汽车出口量还只有当时全球最大汽车出口国日本的四分之一。本月稍早时，中国汽车业称2023年中国总共出口了超过500万辆汽车，超过了日本。中国最大的汽车制造商比亚迪在第四季度售出50万辆电动汽车，让特斯拉望尘莫及。中国的电动汽车造型时尚、技术先进，而且最重要的是价格低廉，目前制约其出口的仅仅是运输船只不足。随着全球走向脱碳，需求将进一步上升。到2030年，中国在全球市场的份额可能会增加一倍，达到三分之一，从而结束西方各国龙头车企的主导地位，尤其是在欧洲。

这一次，政客将更容易把西方的一切就业流失归咎于中国的不正当竞争。冰冷的地缘氛围将强化人们的观感，认为受补贴的生产不公平地夺走了西方工人的生计。而补贴当然是真实存在的。自2014年开始实施“中国制造”计划以来，中国一直公然无视全球贸易规则，大肆补贴本国汽车制造商。它们所享受的低息贷款、股本注入、购车补贴和政府合同的价值难以确切计算。但据一项估计，到2010年代末，对汽车行业的公共支出约占电动汽车销售额的三分之一。除了补贴，中国还在与西方汽车制造商以及西方和韩国的电池制造商的合资企业中掠夺技术。

因此，富裕国家的政策制定者将会很想要保护自己的汽车制造商，让它们

免受由政府资助的竞争的猛攻。去年10月，欧盟委员会启动了对中国汽车的调查。据说美国总统拜登正在考虑对中国汽车增加关税，尽管在27.5%的汽车进口关税和《通胀削减法案》补助的保护下，美国汽车制造商目前几乎没有遭遇中国的竞争。然而，排斥中国汽车将是一个错误。西方可以通过廉价且环保的车辆稳定供应获得巨大的潜在利益，远超过它带来的破坏和危险。

一个原因是，无论与中国的贸易往来如何，汽车市场都即将被颠覆。2022年，在全球销售的新车中，电动汽车占16%至18%。2035年欧盟将全面禁止销售内燃机汽车。尽管企业在转向电动汽车生产时会保留工人，但制造电动汽车所需的劳动力将减少。第一次中国冲击造成的制造业岗位流失实际上不到流失总数的五分之一，当时许多流失是因为值得欢迎的技术进步而造成的。这一次，转向电动汽车引起的冲击与中国生产电动汽车引起的冲击同样有可能被混为一谈。

接下来想想保持贸易畅通带来的好处。汽车是人们购买的最大件商品之一，约占美国总消费的7%。在实际工资受到通胀挤压的时候，汽车更便宜就意味着人们可以把更多的钱用在别处。而且中国车不仅仅是便宜，它们的质量更好，特别是在电动汽车中通过互联网连接实现的智能功能方面。再则，一国的经济增长也不取决于是否有汽车制造业。没什么汽车制造可言的丹麦是全球生活水平最高的国家之一。大批新车从中国的生产线上下来之际，其经济增长萎靡不振，部分原因就是补贴和政府控制导致了严重扭曲。

最后看看环境效益。随着对成本高昂的减排政策的抵制越来越强烈，世界各地的政客都意识到要让消费者转向绿色生活是一项多么艰巨的任务。目前，电动车的价格也高于燃油车（尽管使用成本更低）。接受价格更低的中国汽车可能有助于向净零排放的过渡。比亚迪在中国销售的最便宜的电动汽车约为1.2万美元，而在美国最便宜的特斯拉也要3.9万美元。

那有哪些风险呢？廉价进口商品对产业的威胁往往会被夸大。从上世纪80年代日本和韩国汽车制造商的崛起中我们可以得出一个经验，就是竞争能

够促使本地企业快速自我提升，而新进入者最终都将生产转移到更贴近消费者的地方。比亚迪已经开始在匈牙利建厂，许多中国汽车制造商正在北美选址准备设厂。与此同时，福特和大众等公司正在奋起追赶中国企业。去年，丰田表示其“固态”技术的突破将大幅降低其电池的重量和成本。

另一个担忧是国家安全。在电气化的经济中，电池的重要性远远超出汽车，在电池上完全依赖中国会有风险。电动汽车里有大量芯片、传感器和摄像头，也可能被用于监视。（中国甚至已经禁止在本土制造的特斯拉汽车进入一些政府机构所在地。）但只要国家首脑和情报人员可以乘坐西方或其盟友制造的车辆出行，就没什么理由担心消费者开中国车。消费者可以自行判断个人隐私问题，并且本地生产的汽车也更容易检查。

因此，政策制定者应该抑制自己的保护主义本能。他们只在一种情况下需要担心，而它不太可能发生，那就是西方汽车制造商全面崩溃。中国车商占据重大市场份额而激发更广泛竞争的情形不需要忧惧。如果中国想用自己纳税人的钱补贴全球消费者和加速能源转型，最好的回应就是欢迎它。





Chaguan

Nostalgia for China's boom years drives a TV hit

A drama series presents 1990s capitalists as heroes, not villains

CHINA CENTRAL TELEVISION, the flagship network of the country's propaganda machine, has a new hit on its hands. "Blossoms Shanghai", a big-budget melodrama in 30 parts, has enjoyed huge audiences since its first episode aired on December 27th. The show's success—boosted by approving coverage in official and commercial media outlets—is at once unexpected and revealing. It is a surprise because its heroes are swashbuckling capitalists in the Shanghai of the early 1990s. A hard-living bunch, they cut deals, swap stock tips and scheme against rivals over an endless succession of boozy late-night banquets, filmed in demonic shades of black, gold and red. It is all a far cry from the prim, flag-waving dramas that have become the norm on state TV during Xi Jinping's rule. Typically, such series depict crime-fighting police officers, Communist Party officials toiling to serve the masses, brave Chinese soldiers or other model citizens.

Enthusiasm for the drama, the first TV series to be directed by Wong Kar-wai, a pillar of Hong Kong's film industry, sheds light on the Chinese public's mood. Much praise for the show has a distinctly backward-looking feel to it. Online, fans share their memories of boom years when ordinary Chinese could transform their fates with a lot of luck, good connections and hard work.

At the same time, the show's endorsement by government media is revealing about the hopes and fears of the country's rulers. This official embrace is rather tactical and forward-looking. Boosterish coverage of the drama is in line with a broader campaign by party leaders to cheer up Chinese consumers, whose post-pandemic caution is one reason why the economy is in a funk. Party newspapers credit the drama with sparking a

measurable surge in Shanghai hotel and restaurant bookings. Arguably, a show whose stars are heroic entrepreneurs also aligns with current official efforts to reassure China's private sector. Business types have been battered by heavy-handed regulation in recent years and left feeling generally unloved.

The hero is A Bao, a former factory worker shown making and almost losing a fortune on the stockmarket and in domestic and foreign trade. He is guided by an old man whose counsel runs from business strategy to the right cut for a three-piece suit. ("It has to be British-woven, pure wool," the sage sternly instructs a local tailor, summoned to make A Bao a new wardrobe.) The drama, adapted from a novel by Jin Yucheng, portrays capitalism as something between a test of nerves, a cruel game and a form of madness, capable of inducing a frenzy in consumers and investors alike. Characters cheat one another and commit suicide when ruined. Yet time and again the survivors are drawn to feast together at the same few restaurants, to plot and show off and drink. In contrast with the real 1990s, official corruption is nowhere to be seen. Indeed, the only important character with a public-sector job (at Shanghai's agency for foreign trade) is a paragon of honesty who uses her savings to repay businessmen for gifts they offer her.

To learn more about the show's success, Chaguan caught a fast train to Shanghai and headed to Huanghe Road, a street of restaurants and Art Deco mansions from the 1930s where much of the drama is set. He found a throng of fans taking photographs and filming themselves for social media, over the shrill, electronic whistles of police officers controlling crowds and directing traffic.

Chinese public opinion is rarely monolithic, and responses to the drama divided along lines of home town, age and social class. Locals are happy that the whole series was filmed in Shanghai dialect, with a second version

dubbed into Mandarin for nationwide release. Several Shanghainese pensioners shared strong views about the drama's realism, or lack of it. Back in the 1990s a lot of business was done over dinner, agreed an old man who worked in Shanghai's finance sector. But overall the series is a "fantasy", he scowled. "Those who went into the stockmarket and business were the rare bold ones. Most people worked in factories."

Three older women taking pictures had dressed for a fine dinner, though it was noon. They recalled neon signs that lit up Huanghe Road in those boom years. "Many businessmen gathered here, with their huge mobile phones," remembered one of the women. The trio were not among them. They were assigned jobs in a state-owned textile factory and stayed there until retirement. Modern life offers more choices but more pressure, they declared. In their telling, the series brings the Shanghai of their youth back to life. "But what use is nostalgia?" asked the same woman.

| *Nostalgia as a veiled form of complaint*

The sharpest opinions came from middle-aged fans, some of whom carried small dogs or trailed bored-looking husbands. The early years of China's "reform and opening" era were a time of hope, filled with new experiences, said a 50-year-old woman. She sighed: "We were lucky that we were born in a good age." In her view, life is very different now, and more stressful. Asked why, she replied that it is "hard to talk openly" about this. "There are so many reasons, political factors, among many others."

Young fans sounded more wistful than cross. For two female students, a lesson of the series is that there were more opportunities to move up in the world in the 1990s than now. A 25-year-old man had travelled from Hangzhou, an hour away by train, to take pictures of Huanghe Road. The series may inspire some viewers to start businesses, he enthused. Alas, capitalism is all about timing, he went on. Some may feel they have missed their moment.

Still, China's entrepreneurs should not become cocky about being cast as on-screen heroes. "Blossoms Shanghai" may be a runaway success, but at moments in early January the most-watched show on state TV was a documentary series about officials corrupted by business interests. Back when China first embraced market reforms, party leaders declared: "To get rich is glorious." In the Xi era, the lure of money remains distinctly dangerous. ■



茶馆

对中国经济繁荣年代的怀旧推动一部电视剧爆火

这部剧把1990年代的资本家塑造成英雄而非反派

中国宣传机器的头牌电视网中央电视台最近播出了一部大火的新剧。自12月27日开播第一集以来，30集的大制作情节剧《繁花》吸引了大批观众收看。该剧的成功——受到官方和商业媒体正面评价的推动——既出人意料，也耐人寻味。说出人意料是因为该剧的主角是1990年代初上海一些大胆冒险的传奇资本家。在黑、金、红色交织的暗黑影调中，一群恣意人生的人在无休止的深夜酒局上谈生意、交换股票内幕消息、算计对手。这一切都与习近平执政后国家电视台上常见的那些正经八百、民族情绪高涨的电视剧大相径庭。通常，这些剧集描画的都是打击犯罪的警察、为人民鞠躬尽瘁的共产党官员、勇敢的中国军人或其他模范公民。

观众对该剧（它是香港电影界中流砥柱王家卫执导的首部电视剧）的追捧反映了中国民众的心境。对该剧的赞誉大多带有明显的怀旧色彩。剧迷们在网上分享他们对那个繁荣年代的回忆，那时普通人可以凭借好运、人脉和奋斗改变自己的命运。

与此同时，官方媒体对该剧的认可透露出上层的期盼和忧虑。这种官方的肯定在某种程度上是策略性和前瞻性的。对该剧的积极报道配合了党的领导人提振消费者信心的宏观举措，中国民众在疫情后不敢消费，是如今经济低迷的原因之一。据各家党报报道，在该剧带动下，上海的酒店和餐馆的预订量显著增加。可以说，这部以英雄企业家为主角的电视剧也配合了当前官方为安抚中国私营部门所做的努力。近年来，商界人士受到监管的重拳打压，感到在社会上不受待见。

剧中主人公阿宝原本是工厂工人，后来通过炒股和从事国内外贸易发家致富，最后又几乎倾家荡产。他得到一位长者在方方面面的指点，从商场上的策略到三件套西装的正确剪裁等。（“一定要英纺、纯羊毛的。”这位智者严格叮嘱被召来给阿宝做新衣服的一位本地裁缝）。这部改编自金宇澄的

小说的电视剧把资本主义描绘成某种介于胆量的考验、残酷的博弈和某种癫狂之间的东西，能在消费者和投资者中引发狂热。剧中人物尔虞我诈，在生意失败后自杀。然而，那些熬过来的人却一次又一次回到同样那几家餐厅吃饭、密谋、炫耀、喝酒。与真实的1990年代不同，剧中看不到腐败的官员。实际上，唯一一个公营部门的重要角色（在上海外贸局工作）堪称正直的典范，商人给她送礼物，她会拿自己的积蓄来付钱。

为更多了解该剧爆火的情况，笔者乘坐高铁来到上海，直奔遍布餐馆和上世纪30年代装饰艺术风格洋房的黄河路，该剧大部分剧情都在这里展开。在控管人群和指挥交通的警察那尖锐的电子警哨声中，笔者看到一大群剧迷在那里拍摄照片和视频，准备发到社交媒体上。

中国的公众舆论很少能众口一词，人们对该剧的反应也因家乡、年龄和社会阶层的不同而各异。让上海人开心的是《繁花》全部用沪语拍摄，另有一个普通话配音的版本在全国播放。几位上海退休老人对该剧的写实或不够写实有着强烈的看法。在上世纪90年代，很多生意的确是在饭桌上谈成的，一位曾在上海金融界工作的老人对此表示赞同。但总的来说，这部电视剧是一种“幻想”，他皱着眉说。“那些炒股经商的是极少数胆子大的人。大多数人都是在工厂工作。”

虽是中午时分，有三位在那里打卡拍照的老阿姨却是一身要赴晚宴的打扮。她们回忆起那个繁荣年代黄河路上闪烁的霓虹灯。“当年有许多生意人聚在这里，手里拿着大哥大。”其中一人回忆说。三位阿姨并不在其中。她们被分配到一家国有纺织厂工作，直到退休。她们表示，现代生活提供了更多选择，但也带来了更多压力。在她们看来，《繁花》重现了她们年轻时上海的模样。“但怀旧有什么用呢？”还是那位阿姨问道。

丨 怀旧是一种隐晦的抱怨

表达最尖锐看法的是中年剧迷，他们当中有人抱着小狗，有人拽着神情无聊的丈夫。一位50岁的女士说，中国“改革开放”初期是个充满希望和新体验的时代。她叹道：“我们出生在一个好时代，那会儿是幸运的。”在她看来，现在的生活已经大不一样，压力也更大。当被问及原因时，她回答说

“很难公开谈论”这个问题。“原因太多了，比如政治因素。”

年轻剧迷表达的更多是怅然，而非怨怒。两名女学生表示，《繁花》带来的启示是，与现在相比，1990年代有更多出人头地的机会。一名25岁男子从杭州坐一小时火车来到黄河路打卡拍照。他兴奋地说，这部剧可能会激励一些观众创业。可惜啊，资本主义讲究的是时机，他接着说道。有些人可能会觉得自己已经错过时机了。

尽管如此，中国的企业家们也别因为在荧屏上被塑造成英雄而自鸣得意。

《繁花》可能一炮而红，但在1月初时，央视收视率最高的节目是一部讲述官员被商业利益腐蚀的系列专题片。在中国开展市场改革之初，党的领导人宣称“致富光荣”。而在习时代，金钱的诱惑显然还是洪水猛兽。■



Free exchange

Robert Solow was an intellectual giant

His criticisms were energetic and witty, which could make them harder to take

ENSCONCED IN A lorry, hidden from the enemy by the brow of a hill, the young Robert Solow decoded the radio signals of Nazi platoons across Italy. “We were very, very good at it,” he said. The trick was to get close to the enemy but not too close: near enough to pick up their transmissions, but not so near as to risk capture.

The codes were not fancy—it was “combat stuff”. But if they could be broken quickly, they might reveal an ammunition delivery that could be thwarted. The radiomen were not fancy either. Most were high-school graduates. Even Solow, who would go on to earn a Nobel prize in economics, the Presidential medal of freedom and a Portuguese knighthood, before his death on December 21st 2023, was “middle-middle-class”. He was educated at Brooklyn state schools. He preferred softball to books, and was destined for Brooklyn College until a teacher spotted his potential, broadened his reading, and encouraged him to apply to Harvard University, which he joined two years early and rejoined after the war.

Solow’s years as a soldier only strengthened his egalitarian streak. He declined to become an officer, so he would not have to boss anyone around. When the Massachusetts Institute of Technology (MIT) offered him a job in 1949, he asked what the lowest paid professor earned, and accepted the same. When he served in President Kennedy’s Council of Economic Advisers, the Swiss embassy wanted to know his protocol rank. His answer was that he was a full professor at MIT and the government had no rank as high. Informed in the predawn hours in October 1987 that he had won the Nobel prize, his first instinct was to go back to sleep.

What he craved was more precious than prizes: the esprit de corps that comes from membership of a small, highly motivated band of colleagues. “If you’re in a group that is doing good work, it’ll have a high morale. And if it has high morale, it’ll do good work,” he once said. As an economist, he liked formal models and mathematics. But nothing too fancy. Over-refinement reminded him of the man who knew how to “spell banana” but did not “know when to stop”. His strategy was to break big questions—about growth, resources, unemployment—into littler ones, in the hope that small answers would aggregate into larger ones.

The MIT culture he embodied disdained hierarchy, cherished collegial lunches and made time for students, many of whom became illustrious friends. Four of Solow’s students later received their own sleep-disturbing calls from Sweden. Economics, Solow maintained, was a “handicraft” industry, often driven by the “extraordinarily powerful research apparatus” of one professor and one undergraduate assistant.

Something he liked about academia was that ideas, no matter how prestigious their source, could be scrutinised by anyone. His own criticisms were energetic and witty, which could make them harder to take. He found the “freshwater” school of macroeconomics, identified with the University of Chicago, preposterous, especially in its early incarnations, which assumed a “representative agent” could stand in for the many actors in an economy. To get into a technical discussion with freshwater types was like discussing cavalry tactics with someone claiming to be Napoleon, he said. The claim is absurd, however well they know their stuff.

The work that made his name began as criticism of the growth theories of the 1930s and 1940s. In these, investment added both to national spending and the economy’s productive capacity. There was no guarantee these additions to demand and supply would stay in line with each other. Moreover, excessive spending, by boosting demand, would inspire firms to

invest even more, whereas inadequate investment would induce firms to spend still less. The economy was for ever poised on a “knife-edge” between deepening unemployment or intensifying labour shortages.

This precariousness was hard to square with the relatively stable progress of advanced economies like America, where even the Great Depression eventually ended. Solow showed that the knife-edge disappeared if economies could vary the capital-intensity of production. Strong investment would not then be destabilising. It would merely result in higher capital per worker.

High investment would not, however, result in faster growth over the long run. At some point, capital would run into diminishing returns, leaving growth to be dictated by other factors. Solow calculated that capital accumulation could explain less than 13% of the growth in income per person in America from 1909 to 1949. The remainder was attributable to other forces, which he loosely labelled “technical change”. This vast unexplained portion of growth became known as the “Solow residual”.

| *Tough paternal love*

Although his work created reams of subsequent research, the father of growth theory was not impressed by many of his progeny. He was sceptical of statistical exercises that dissected growth rates across countries at every stage of development. Nor had he intended to imply that technological progress, which he did not model, fell entirely outside economics. A lot of innovation was “dumb luck”. And much of it emerged on the factory floor, “invented” by unheralded foremen. But some was the result of profit-driven investment in research. Later attempts to create formal theories of technological progress nevertheless asked more questions than they answered, he argued.

Part of the problem was that innovation is often peculiar and particular,

whereas growth theorists strive for generality and abstraction. Solow, who had himself observed the research labs at General Motors and collaborated with the McKinsey Global Institute on industry-level studies of productivity, thought model-builders could learn from case studies and business histories. The aim was to “extract a few workable hypotheses” without getting lost in the detail. To understand how the economy works, to decode its secrets, you need to get up close, but not too close. ■



自由交流

罗伯特·索洛，一位思想巨人

他的批判活力十足、机智风趣，这可能让它们更难被接受

年轻的罗伯特·索洛（Robert Solow）躲在一辆卡车里，利用山丘的掩护避开敌人视线，抓紧破译意大利各地纳粹部队的无线电信号。“我们特别擅长干这个。”他说。诀窍是要与敌军若即若离：近到足以接收到他们的信号，但又不能太近，以免被捉住。

这些无线电密码并不复杂——就是些“作战信息”。但如果能迅速将之破解，就可能发现某个弹药运送行动而挫败它。无线电破译员也没什么神秘的，大多数是高中毕业生。即便是索洛——他后来获得诺贝尔经济学奖、总统自由勋章和葡萄牙爵位，于2023年12月21日去世——也只是一个“普通中产阶级”。他在美国布鲁克林的公立学校接受教育。比起读书，他更喜欢垒球，本来打算报读布鲁克林学院（Brooklyn College），直到一位老师发现了他的潜力，拓宽了他的阅读面，并鼓励他报读哈佛大学。他提前两年入读哈佛，战后又重返校园。

在他的从军岁月里，索洛的平等主义倾向进一步强化。他拒绝担任军官，这样他就不必对任何人发号施令。当麻省理工学院1949年聘任他时，他询问了收入最低的教授的薪资水平，并接受了同样的待遇。当他在肯尼迪总统的经济顾问委员会（Council of Economic Advisers）任职时，瑞士大使馆想了解他的礼宾级别。他答说自己是麻省理工的正教授，政府里没有比这更高的级别了。1987年10月的一个黎明时分，他得知自己获得了诺贝尔奖，第一反应是继续睡觉。

他渴求的东西比奖项更珍贵：一小群干劲十足的同事凝聚起的团队精神。他曾说过：“如果你所在的团队工作出色，士气就会高昂。如果士气高昂，工作就会出色。”作为一名经济学家，他喜欢正式的模型和数学。但他不喜欢过于繁复精致的东西。过度精细在他眼里就好像知道如何“拼写banana”却不知道“在哪里停下”。他的策略是将大问题——增长、资源、失

业——分解成若干较小问题，希望小的答案能汇聚成更大的解答。

他身上展现的麻省理工文化对等级制度不屑一顾，珍惜同事共进午餐的时光，乐意为学生抽出时间，其中许多学生日后成为了赫赫有名的友人。索洛的四名学生后来也在深夜被来自瑞典的电话吵醒。索洛认为，经济学是一种“手工艺”产业，常常由一名教授和一名本科生助理组成的“极强大研究器具”驱动。

他喜欢学术圈的一点是，无论提出观点的人地位多么显赫，任何人都可以对观点加以审视。他自己做出的批判活力十足、机智风趣，这可能让它们更难被接受。他认为以芝加哥大学为代表的宏观经济学“淡水”学派是荒谬的，尤其是其早期版本，它假定一个“代表性主体”可以代替经济中的众多参与者。他说，与淡水学派的人做学术讨论就像和某个声称自己是拿破仑的人讨论骑兵战术。无论他们多么懂行，这种自诩都是荒诞的。

他的成名始于对上世纪三四十年代经济增长理论的批判。这些理论认为，投资既增加了国民支出，也提高了经济的生产能力，但谁也不能保证这些需求和供应的增加能够保持相互匹配。此外，刺激需求导致的过度支出将激励企业进一步加大投资，而投资不足则会导致企业进一步削减支出。经济始终悬在失业恶化和劳动力短缺加剧之间的“刀口”上。

这种岌岌可危与美国等发达经济体相对平稳的发展态势并不相符合，因为即使是大萧条最终也还是结束了。索洛表明，如果经济体能够改变生产的资本密集度，刀口就会消失。这时，强劲投资就不会带来不稳定性，而只会导致工人的人均资本量增加。

然而，从长期来看，高投资并不会带来更快的增长。到了一定程度，资本将陷入收益递减，此时经济增长将受到其他因素的支配。根据索洛的计算，从1909年到1949年，美国人均收入的增长只有不到13%能用资本积累解释。其余部分归因于其他因素，他笼统地称之为“技术变革”。增长中这一无法解释的庞大部分后来被称为“索洛残差”。

尽管他的成果引发了大量的后续研究，但这位增长理论之父对众多后辈的表现不甚满意。他对剖析各国在不同发展阶段的增长率的统计学方法持怀疑态度。虽然他没有对技术进步建模，但他也无意暗示技术进步完全跌出了经济学范畴。很多创新“纯属好运”。其中大部分都是名不见经传的工头在工厂车间里“发明”的。但也有一些是源于利润驱动的对研发的投资。他认为，后来人们试图建立关于技术进步的正式理论，相比解答问题，反而提出了更多问题。

这在一定程度上是因为创新往往是特殊且具体的，而增长理论家追求普遍性和抽象性。索洛亲身观察过通用汽车的研究实验室，并与麦肯锡全球研究院合作开展行业层面的生产率研究。他认为，建模者可以从案例研究和商业史中学习。其目标是“提取几个可行的假设”，但不迷失在细枝末节之中。要了解经济的运作规律，破译其秘密，关键是若即若离：要近距离观察，但又不能太近。 ■



Machine of the state

Welcome to the era of AI nationalism

Sovereigns the world over are racing to control their technological destinies

THE HOTTEST technology of 2023 had a busy last few weeks of the year. On November 28th Abu Dhabi launched a new state-backed artificial-intelligence company, AI71, that will commercialise its leading “large language model” (LLM), Falcon. On December 11th Mistral, a seven-month-old French AI startup, announced a blockbuster \$400m funding round, which insiders say will value the firm at over \$2bn. Four days later Krutrim, a new Indian startup, unveiled India’s first multilingual LLM, barely a week after Sarvam, a five-month-old one, raised \$41m to build similar Indian-language models.

Ever since OpenAI, an American firm, launched ChatGPT, its humanlike conversationalist, in November 2022, just about every month has brought a flurry of similar news. Against that backdrop, the four latest announcements might look unexceptional. Look closer, though, and they hint at something more profound. The three companies are, in their own distinct ways, vying to become AI national champions. “We want AI71 to compete globally with the likes of OpenAI,” says Faisal al-Bannai of Abu Dhabi’s Advanced Technology Research Council, the state agency behind the Emirati startup. “Bravo to Mistral, that’s French genius,” crowed Emmanuel Macron, the president of France, recently. ChatGPT and other English-first LLMs “cannot capture our culture, language and ethos”, declared Krutrim’s founder, Bhavish Aggarwal. Sarvam started with Indian languages because, in the words of its co-founder, Vivek Raghavan, “We’re building an Indian company.”

AI is already at the heart of the intensifying technological contest between America and China. In the past year their governments have pledged

\$40bn-50bn apiece for AI investments. Other countries do not want to be left behind—or stuck with a critical technology that is under foreign control. In 2023 another six particularly AI-ambitious governments around the world—Britain, France, Germany, India, Saudi Arabia and the United Arab Emirates (UAE)—promised to bankroll AI to the collective tune of around \$40bn (see chart). Most of this will go towards purchases of graphics-processing units (GPUs, the type of chips used to train AI models) and factories to make such chips, as well as, to a lesser extent, support for AI firms. The nature and degree of state involvement varies from one wannabe AI superpower to another. It is early days, but the contours of new AI-industrial complexes are emerging.

Start with America, whose tech firms give everyone else AI envy. Its vibrant private sector is innovating furiously without direct support from Uncle Sam. Instead, the federal government is spending around \$50bn over five years to increase domestic chipmaking capacity. The idea is to reduce America's reliance on Taiwanese semiconductor manufacturers such as TSMC, the world's biggest and most sophisticated such company. Supplies from Taiwan could, fear security hawks in Washington, be imperilled if China decided to invade the island, which it considers part of its territory.

Another way America intends to stay ahead of the pack is by nobbling rivals. President Joe Biden's administration has enacted brutal export controls that ban the sale of cutting-edge AI technology, including chips and chipmaking equipment, to adversaries such as China and Russia. It has also barred Americans from sharing their AI expertise with those countries.

It is now coercing those on the geopolitical fence to fall in line. In October the American government started requiring companies in third countries, including Saudi Arabia and the UAE, to secure a licence in order to buy AI chips from Nvidia, an American firm that sells most of them. The rules have a "presumption of approval". That means the government will

“probably allow” sales to such firms, says Gregory Allen, who used to work on AI policy at the Department of Defence—as long, that is, as they do not have close ties to China. On December 6th Xiao Peng, who runs a state-backed AI startup in Abu Dhabi called G42, announced that the company would be cutting ties with Chinese hardware suppliers such as Huawei, a Chinese electronics company.

China’s AI strategy is in large part a response to American techno-containment. According to data from JW Insights, a research firm, between 2021 and 2022 the Chinese state spent nearly \$300bn to recreate the chip supply chain (for AI and other semiconductors) at home, where it would be immune from Western sanctions. A lot of that money is probably wasted. But it almost certainly helped Huawei and SMIC, China’s biggest chipmaker, to design and manufacture a surprisingly sophisticated GPU last year.

The central and local authorities also channel capital into AI firms through state-backed “guidance funds”, nearly 2,000 of which around the country invest in all manner of technologies deemed to be strategically important. The Communist Party is guiding private money, too, towards its technological priorities. Often it does so by cracking down on certain sectors—most recently, in December, video-gaming—while dropping heavy hints about which industries investors should be looking at instead. The government is also promoting data exchanges, where businesses can trade commercial data on everything from sales to production, allowing small firms with AI ambitions to compete where previously only large data-rich firms could. There are already 50 such exchanges in China.

Elements of this state-led approach are being emulated in other parts of the world, notably in the Gulf’s petrostates. Being autocracies, Saudi Arabia and the UAE can move faster than democratic governments, which must heed voters’ concerns about AI’s impact on things like privacy and jobs. Being wealthy, they can afford both the GPUs (on which the two countries have

together so far splurged several hundred million dollars) and the energy needed to run the power-hungry chips.

They can also plough money into developing human capital. Their richly endowed universities are quickly climbing up global rankings. The AI programme at King Abdullah University of Science and Technology in Saudi Arabia and the Mohamed bin Zayed University of Artificial Intelligence (MBZUAI) in Abu Dhabi, the world's first AI-focused school, have poached star professors from illustrious institutions such as the University of California, Berkeley, and Carnegie Mellon University in Pittsburgh. Many of their students and researchers come from China. And plenty stick around. Nearly all of MBZUAI's graduates, who number a couple of hundred, stay in the region to work at local firms and labs, says its provost, Timothy Baldwin (himself lured to the Middle East from the University of Melbourne).

The Gulf approach is producing results. The capabilities of the Falcon model, first built by a team of 20 or so engineers, rival those of Llama 2, the most widely used "open-source" model, devised by Meta, an American tech giant. AI71 plans to improve its open-source models using national datasets from fields including health, education and, some day, perhaps oil. "In the last 50 years, oil drove the country...now data is the new oil," says Mr al-Bannai.

| *The alignment problem*

A third group of governments is combining elements of America's approach with those of the Chinese and Emiratis. The EU has its version of America's incentives for domestic chipmaking. So do some member states: Germany is footing a third of the €30bn (\$33bn) bill for a new chip factory to be built there by Intel, an American chipmaker. Outside the bloc, Britain has promised to funnel £1bn (\$1.3bn) over five years to AI and supercomputing (albeit without going into detail about how exactly the money will be spent). India's government is promoting manufacturing,

including of semiconductors, with generous “production-linked incentives”, encouraging big cloud-computing providers to build more Indian data centres, where AI models are trained, and thinking about buying \$1.2bn-worth of GPUs.

Like China and the Gulf but unlike America, where federal and state governments are reluctant to part with public data, India and some European countries are keen on making such data available to firms. France’s government “has been very supportive” in that regard, says Arthur Mensch, Mistral’s boss. Britain’s is considering allowing firms to tap rich data belonging to the National Health Service. India’s government has enormous amounts of data from its array of digital public services, known as the “India Stack”. Insiders expect it eventually to integrate Indian AI models into those digital services.

In contrast to China, which regulates consumer-facing AI with a heavy hand, at least for the time being Britain, France, Germany and India favour light-touch rules for AI or, in India’s case, none at all. The French and German governments have soured on the EU’s AI Act, the final details of which are being hotly debated in Brussels—no doubt because it could constrain Mistral and Aleph Alpha, Germany’s most successful model-builder, which raised €460m in November.

It is natural for countries to want some control over what may prove to be a transformational technology. Especially in sensitive and highly regulated sectors such as defence, banking or health care, many governments would rather not rely on imported AI. Yet each flavour of AI nationalism also carries risk.

America’s beggar-thy-neighbour approach is likely to upset not just its adversaries but also some allies. China’s heavy regulation may offset some of the potential gains from its heavy spending. Building models for local

languages, as Krutrim and Sarvam in India plan to do, may prove futile if foreign models continue to improve their multilingual capabilities. The Gulf's bet on open-source models may misfire if other governments limit their use, as Mr Biden has hinted at in a recent executive order and the EU could do through its AI Act, out of fear that open LLMs could be put to malign uses by mischief-makers. Saudi and Emirati institutions may struggle to hold on to talent; a developer who worked on Falcon admits it greatly benefited from a partnership with a French team of engineers who have since been poached by Hugging Face, a high-flying Silicon Valley AI startup. As one sceptical investor notes, it is not yet clear how vast or useful public Emirati data actually is.

As Nathan Benaich of Air Street Capital, a venture-capital firm, sums it up, most efforts to create national models "are probably a waste of money". Mr Benaich's warning is unlikely to dissuade AI-curious governments, mindful of the rewards should they succeed, from meddling. Mr Macron will not be the only leader to greet it with a Gallic shrug. ■



国家机器

欢迎来到人工智能民族主义的时代

世界各地的主权国家正竞相控制它们的科技运数【深度】

二〇二三年最热门的技术在年末的几周里热闹非凡。11月28日，阿布扎比推出了一家由政府支持的新人工智能公司AI71，致力于把阿联酋领先的“大语言模型”（LLM）Falcon商业化。12月11日，成立七个月的法国AI创业公司Mistral宣布了一轮4亿美元的大手笔融资，知情人士称这将使该公司估值超过20亿美元。四天后，印度新兴创业公司Krutrim发布了印度首个多语言LLM，而不过一周前，创立五个月的Sarvam才刚融资了4100万美元构建类似的印地语语言模型。

自从美国公司OpenAI在2022年11月推出了对话能力接近人类的ChatGPT以来，几乎每个月都会冒出一连串类似的新闻。在这一背景下，最近的这四个公告看似并无特别之处。但仔细观察，就会发现它们透露出某种更深层的动态。这三家公司正在以各自不同的方式竞逐AI国家冠军之位。AI71背后的政府机构、阿布扎比的先进技术研究委员会（Advanced Technology Research Council）的费萨尔·班奈（Faisal al-Bannai）表示：“我们希望AI71能在全球范围里与OpenAI这样的公司竞争。”法国总统马克龙最近喜不自胜地夸赞：“为Mistral喝彩，它是法国人天分的体现。”Krutrim的创始人巴维什·阿加瓦尔（Bhavish Aggarwal）宣称，ChatGPT和其他以英语为先的LLM“无法捕捉我们的文化、语言和精神”。Sarvam的联合创始人维韦克·拉加万（Vivek Raghavan）表示“我们是在建立一家印度公司”，因此Sarvam从印地语开始。

AI已经是美国和中国日益激烈的科技竞争的核心。在过去一年中，两国政府都承诺了四五百亿美元的AI投资。其他国家不想落后或受制于一项由他国掌控的关键技术。2023年，世界上另外六个在AI上格外有抱负的政府——英国、法国、德国、印度、沙特阿拉伯和阿联酋——承诺将总共投入约400亿美元用于支持AI（见图表）。其中大部分资金将用于购买图形处理单元（GPU，用于训练AI模型的芯片）和制造此类芯片的工厂，还有较

少部分用来支持AI公司。意欲成为AI超级大国的国家在政府参与的性质和程度上各不相同。现在一切还只是个开始，但新的AI产业联合体的轮廓正在浮现。

先来看美国，它的科技公司拥有令各方羡慕的AI实力。其充满活力的私营部门正在疯狂创新，而无需山姆大叔的直接支持。美国政府转而致力提升国内的芯片制造能力，正为此在五年内投入约500亿美元。其想法是减少美国对台湾半导体制造商（如全球最大、最先进的台积电）的依赖。华盛顿的安全鹰派担心，如果中国大陆决定对台动武（中国大陆认为台湾是中国领土的一部分），那么来自台湾的供应可能会受到威胁。

美国试图保持领先的另一个方式是阻挠竞争对手。拜登政府已实施严厉的出口管制，禁止向中国和俄罗斯等对手销售尖端AI技术，包括芯片和芯片制造设备。它还禁止美国人与这些国家分享他们的AI专业知识。

现在，美国正在迫使那些在地缘上不肯选边的国家站到它这一边。去年10月，美国政府规定，包括沙特阿拉伯和阿联酋在内的第三国公司须获得许可才能从销售了全球大部分AI芯片的美国公司英伟达购货。这些规则有一个“批准假定”。也就是说，美国政府“可能会允许”向这些公司销售，曾在国防部负责AI政策的格雷戈里·艾伦（Gregory Allen）说——只要它们与中国没有密切联系。12月6日，掌管由政府支持的阿布扎比AI创业公司G42的肖鹏宣布，公司将切断与电信公司华为等中国硬件供应商的联系。

中国的AI战略在很大程度上是对美国技术遏制的回应。根据研究公司集微咨询的数据，2021年至2022年间，中国政府花费了近3000亿美元在国内重建芯片供应链（包括AI和其他半导体），以求对西方制裁免疫。其中很多钱可能被浪费掉了。但它几乎肯定帮助华为和中国最大的芯片制造商中芯国际在去年设计和制造了一款先进程度出人意料的GPU。

中央和地方政府还通过政府“引导基金”将资本导向AI公司，全国有近2000个此类基金投资于各种被认为具有重要战略意义的技术。共产党也在引导私人资金投向它要优先发展的技术。为实现这一目标，它通常会打击某些

行业——最近一次是在12月整顿电子游戏业——同时重点提示投资者应该转而关注哪些行业。中国政府还在推动建立数据交易所，企业可以在这里交易从销售到生产的各类商业数据，使怀抱AI雄心的小公司能够在过去唯数据充足的大公司才能立足的领域里竞争。中国已有50个此类交易所。

这种政府主导的模式中的一些要素正被世界其他地区仿效，尤其是海湾的石油国家。沙特和阿联酋这样的威权国家可以比民主政府行动更快，后者必须听取选民对AI影响隐私和就业等方面的担忧。而由于国家富有，它们既买得起GPU（这两个国家到目前为止已经为此花费了几亿美元），也能负担运行这些高耗能芯片所需的能源。

它们还能投资发展人力资本。它们的大学资金充裕，全球排名正在迅速攀升。沙特的阿卜杜拉国王科技大学（King Abdullah University of Science and Technology）的AI项目，以及阿布扎比的穆罕默德·本·扎耶德人工智能大学（MBZUAI，世界上第一所专注于AI的大学）已经从加州大学伯克利分校和匹兹堡的卡内基梅隆大学等声名赫赫的院校挖走了明星教授。它们的许多学生和研究人员来自中国，而且其中很多人留了下来。MBZUAI的毕业生大约有几百人，几乎全部留在了中东地区，在当地的公司和实验室工作，MBZUAI的教务长蒂莫西·鲍德温（Timothy Baldwin）说。他本人也是从墨尔本大学被吸引到中东的。

海湾国家的模式正在取得成果。最初由大约有20名工程师的团队构建的Falcon模型的性能可媲美美国科技巨头Meta设计的最广泛使用的“开源”模型Llama 2。AI71计划利用健康、教育等领域的国家数据集改进其开源模型，有朝一日或许还会包括石油领域。“过去50年里，石油驱动了这个国家……现在数据是新的石油。”班奈说。

｜ 对齐问题

第三类政府正在把美国模式与中国和阿联酋的模式中的要素结合起来。欧盟有类似美国的国内芯片制造激励措施。它的一些成员国也有这类措施：德国承担了美国芯片制造商英特尔将在该国建造的新芯片工厂300亿欧元（330亿美元）造价的三分之一。在欧盟之外，英国承诺在五年内向AI和

超级计算领域投入10亿英镑（13亿美元），尽管没有详细说明这笔钱会怎么用。印度政府正在推动包括半导体制造在内的制造业，它提供慷慨的“生产挂钩激励”，鼓励大型云计算提供商在印度建设更多可用于训练AI模型的数据中心，并考虑购买总价值12亿美元的GPU。

印度和一些欧洲国家热衷于向企业提供公共数据，这与中国和海湾国家相似，而与美国不同（联邦和州政府不愿分享这类数据）。法国政府在这方面“非常支持”，Mistral老板亚瑟·门什（Arthur Mensch）说。英国正在考虑允许企业获取国家医疗服务体系（NHS）中的丰富数据。印度政府拥有大量来自各类数字公共服务（被称为“印度堆栈”）的数据，知情人士预计它最终将把印度的AI模型整合到这些数字服务中。

中国对面向消费者的AI实行严格管制，而英国、法国、德国和印度则倾向于宽松监管AI——至少在目前是如此，印度更是完全不加管束。法国和德国政府对欧盟的《人工智能法案》（AI Act）感到不满。谈判代表们正就该法案的最终细节在布鲁塞尔激烈辩论——毫无疑问，这是因为它可能会限制Mistral和德国最成功的模型开发公司Aleph Alpha（它在去年11月融资4.6亿欧元）。

各国想要对一项可能具有变革性的技术有所控制是自然的事。尤其是在国防、银行或医疗等敏感且受高度监管的领域，许多政府不愿意依赖进口AI。然而，各色AI民族主义也会带来风险。

美国以邻为壑的政策不仅可能激怒对手，也很可能惹恼一些盟友。中国的严格监管可能会抵消其大量支出带来的潜在收益。在印度，Krutrim和Sarvam计划为本地语言构建模型，但如果外国模型继续提高其多语言能力，这些努力就可能是无用功。如果其他政府担心开放式LLM可能被图谋不轨者恶意使用而限制对开源模型的使用，海湾国家对这类模型的押注可能会落空。拜登在最近的行政命令中已暗示会限制使用开源模型，欧盟也可能通过其《人工智能法案》这样做。沙特和阿联酋的机构可能难以留住人才；一位参与Falcon项目的开发人员承认，该项目从与一个法国工程师团队的合作中受益匪浅，这些工程师后来被高歌猛进的硅谷AI创业公司

Hugging Face挖走。正如一位持怀疑态度的投资者所指出的，阿联酋公共数据实际上有多丰富和有用还是个未知数。

正如风险投资公司Air Street Capital的内森·贝奈奇（Nathan Benaich）所总结的那样，大多数创建国家模型的努力“可能是浪费钱”。贝奈奇的提醒不太可能说服那些对AI兴致盎然的政府不插手，它们一心想着如果成功将能获得的回报。马克龙不会是唯一以耸耸肩来回应的领导人。■



The Gulf between them

Can Sino-Arabian business ties replace Sino-American ones?

The Middle Kingdom gets cosy with the Middle East

WHEN CHINESE and Middle Eastern moneymen meet, it is usually behind closed doors. Last month they mingled openly in the lobby of the Hong Kong Stock Exchange, at the “FII Priority” summit, an event organised by the Public Investment Fund (PIF), a \$780bn vehicle for Saudi sovereign wealth. It was the first meeting of its kind in east Asia. It will not be the last. The PIF announced it was planning to set up an office in China. Mubadala and the Investment Corporation of Dubai, two Emirati sovereign wealth funds, the Qatar Investment Authority and Kuwait Investment Authority are all said to be preparing to deploy more capital in the world’s second-biggest economy. They think they can do this without angering the increasingly China-wary West. “We are friendly people, we are friends with everyone,” Jerry Todd, an executive at the PIF, told the conference in Hong Kong.

China’s investment firms and the companies they back need friends right now. As Sino-American geopolitics sour, American investments in China have collapsed. Chinese tech firms got \$1.2bn from American venture capitalists in 2022, down from \$14bn in 2018. Mergers and acquisitions (M&A) by American firms in China fell below \$9bn in 2023, down from \$20bn five years earlier. Meanwhile M&A deals by Gulf entities have ballooned—from next to nothing in 2019 to nearly \$9bn in 2023, according to data from LSEG, a financial-information firm (see chart).

Last month NIO, a Chinese Tesla wannabe, said it had received \$2.2bn from CYVN Holdings, a firm controlled by Abu Dhabi’s government that had previously put more than \$1bn into the electric-car maker. The NEOM Investment Fund, part of a pharaonic Saudi project to build a futuristic city

in the desert, has backed Pony.AI, a part-Chinese developer of self-driving tech. Earlier in the year Saudi Aramco, the kingdom's oil colossus, invested \$3.6bn in a Chinese petrochemical refinery called Rongsheng, and entered into a joint venture with the PIF and Baosteel, one of China's largest steelmakers, to produce high-quality metal plates in Saudi Arabia. Chinese VC firms are tight-lipped in public about their limited partners but privately confirm that in the past two years interest from Middle Eastern ones has jumped.

Tech talent, of which the Gulf is short but China has aplenty, is flowing in the other direction. The Shenzhen campus of the Chinese University of Hong Kong and the Shenzhen Research Institute of Big Data are helping Saudi Arabia's King Abdullah University of Science and Technology (KAUST) build an artificial-intelligence model to power an Arabic-language chatbot called AceGPT. Around one in five of KAUST's students and one in three of its postdoctoral researchers are Chinese.

The budding Sino-Arabic relationship will not replace the wilting Sino-American one. Dubai and Riyadh cannot match the depth of Silicon Valley's expertise and New York's capital markets. The Gulf wealth funds mostly cut cheques for a few hundred million dollars, whereas Americans also back early-stage startups in need of a few million. And for the Gulf, America remains a vital partner. In December an Emirati AI startup called G42, whose backers include Mubadala and Silver Lake, an American investor, said it would sever ties with Chinese firms rather than lose access to American technology. "We cannot work with both sides," its chief executive, Xiao Peng, told the Financial Times. So much for being friends with everyone. ■



相隔一道湾

中阿经贸关系能否取代中美经贸关系？

中国与中东关系拉近

中国和中东的金融家会面时一般都是闭门会议。上月，在港交所的大会堂里，他们在“未来投资倡议优先事项”峰会（FII Priority summit）上举行了公开会谈。组织此次活动的公共投资基金（PIF）是手握7800亿美元的沙特主权财富基金。这是东亚地区首次举行此类会议。但不会是最后一次。PIF宣布计划在中国设立办事处。据称，阿联酋两家主权财富基金穆巴达拉（Mubadala）和迪拜投资公司（Investment Corporation of Dubai）、卡塔尔投资局（Qatar Investment Authority）以及科威特投资局（Kuwait Investment Authority）都准备在这个全球第二大经济体部署更多资本。它们认为可以在不激怒日益警惕中国的西方世界的情况下做到这一点。“我们与人为善，我们和所有人都是朋友。”PIF高管杰里·托德（Jerry Todd）在香港的会议上表示。

中国的投资公司以及它们所支持的公司现在需要朋友。随着中美地缘关系恶化，美国在中国的投资已经崩溃。2022年，中国科技公司从美国风险投资机构那里仅获得12亿美元，低于2018年的140亿美元。2023年，美国公司在中国的并购交易还不到90亿美元，低于五年前的200亿美元。与此同时，根据金融信息公司伦敦交易所集团（LSEG）的数据，由海湾实体达成的并购交易激增——从2019年的几乎为零飙升到2023年的近90亿美元（见图表）。

上个月，想要成为中国版特斯拉的电动汽车制造商蔚来表示从CYVN Holdings获得22亿美元，此前它已经从这家由阿布扎比政府控制的公司获得超过10亿美元投资。NEOM投资基金（NEOM Investment Fund）是沙特一个在沙漠中建造未来城市的宏大项目的一部分，该基金投资了有中国背景的自动驾驶技术开发商小马智行。2023年更早时，沙特石油巨头沙特阿美（Saudi Aramco）向中国炼油企业荣盛石化投资36亿美元，并与PIF和中国最大钢铁企业之一宝钢成立了一家合资企业，在沙特生产高质量的金

属板材。中国的风投公司在公开场合对其有限合伙人的信息守口如瓶，但私下证实过去两年里来自中东的风投兴趣大增。

中国盛产海湾缺乏的技术人才，这些人才正涌向海湾。香港中文大学深圳校区和深圳大数据研究院正在帮助沙特阿拉伯的阿卜杜拉国王科技大学（KAUST）搭建一个人工智能模型，来驱动一个名叫AceGPT的阿拉伯语聊天机器人。KAUST大约五分之一的学生和三分之一的博士后研究人员是中国人。

萌芽中的中阿关系不会取代日渐枯萎的中美关系。迪拜和利雅得没有硅谷那么丰富的专业知识，也没有纽约那样深厚的资本市场。海湾的财富基金开出的大多是数亿美元的支票，但美国人还支持需要数百万美元的早期创业公司。而对海湾来说，美国仍然是一个至关重要的伙伴。上月，一家名为G42的阿联酋人工智能创业公司表示，宁愿切断与中国公司的联系，也不愿失去获得美国技术的机会。G42的投资者包括穆巴达拉和美国投资公司银湖（Silver Lake）。该公司首席执行官肖鹏向《金融时报》表示，“我们不能跟两边都合作。”所谓的和所有人做朋友看来也就这样了。■



Life ACWAtic

Meet ACWA Power, Saudi Arabia's unlikely solar star

The utility has green ambitions beyond its desert home

SITTING ATOP a fifth of the world's oil reserves, Saudi Arabia doesn't spring to mind when you think about renewables. Muhammad bin Salman, its crown prince and de facto ruler, would like this to change. He wants half of Saudi electricity to come from wind and solar farms by 2030. Two-thirds of that capacity, or around 40 gigawatts (GW) will, if Prince Muhammad gets his wish (as he tends to do), be courtesy of one firm: ACWA Power.

For most of its 19-year existence the utility was a relatively anonymous family-run affair. No longer. Since it went public in Riyadh in 2021 its market value has swelled nearly four-fold. It is now worth \$50bn. The Public Investment Fund (PIF), the steward of Saudi sovereign wealth, owns a 44% stake. ACWA has 24GW of green projects at home and abroad either already running or at an advanced stage, up from 0.3GW in 2014. Add its other capacity under construction and the total is 54GW. Its original business of desalinating water went from 1m cubic metres a day in 2006 to 7.6m cubic metres in December. Its newish boss, Marco Arcelli, a seasoned Italian energy executive, expects assets it has a stake in to triple between now and 2030, to \$250bn. Its projects will, he hopes, help create a broader domestic green-energy supply chain. "We are a big enabler," he says.

ACWA has thrived as many other renewables operators around the world have struggled. Whereas those rivals are seeing the cost of projects soar as a result of rising interest rates, ACWA has received non-interest-bearing loans from the PIF, in addition to debt secured against individual projects and loans from banks to tide it over while it raises more equity capital and brings in partners. Access to easy money has allowed ACWA to expand capacity, while lowering costs for customers. This has helped make the

levelised cost of Saudi solar energy, which takes into account both construction and operation of a power plant, among the lowest in the world.

Nevertheless, ACWA's returns on domestic projects are low by global standards. Mr Arcelli is thus keen to take advantage of juicier ones on offer abroad. He is investing nearby (in Bahrain, Egypt, Jordan, Oman, Turkey and the United Arab Emirates) and farther afield (Azerbaijan, Morocco, South Africa and Uzbekistan). Two-fifths of ACWA's overall capacity is to be found outside Saudi Arabia. It is also eyeing China, a highly competitive market but one where ACWA could, thinks Mr Arcelli, gain both scale and technology partners in the form of Chinese manufacturers of wind turbines and solar panels.

ACWA has its work cut out. To meet Prince Muhammad's domestic goals for it, the company must add 6-7GW of capacity—equivalent to three or four big projects—every year for the rest of the decade. It currently has just 14GW at various stages of development. Managing fast expansion will require a laser focus on costs (those of its nascent hydrogen venture has already risen by 70% from initial estimates, to \$8bn). It will also require more debt. In September ACWA was already sitting on \$7bn of it, equivalent to seven times its earnings before interest, taxes, depreciation and amortisation. Such a ratio would be considered a red flag at most firms.

ACWA may yet rise to the challenge. It can count on the PIF's deep pockets. And it is a fast learner; its domestic 1.5GW Sudair solar project may be fully up and running in just over two years, reckons Oliver Connor of Citigroup, a bank, brisk by industry standards. Mr Arcelli wants things to go faster still. Given that the prince is watching, that is no surprise. ■



ACWA生平

认识下ACWA Power，沙特冷门的太阳能明星

这家公用事业公司的绿色雄心冲出了沙漠家园

说到可再生能源，你通常不会想到坐拥全球五分之一石油储量的沙特阿拉伯。沙特的实际统治者、王储穆罕默德·本·萨勒曼想要改变你的想法。他希望到2030年，沙特一半的电力都来自风能和太阳能。如果这位王储得偿所愿（通常都能），那么其中三分之二也就是大约40GW的发电能力都将由一家公司提供：沙特国际电力和水务公司（ACWA Power）。

在成立的19年间，这家公用事业公司大多数时候都只是个不大起眼的家族企业。现在不一样了。自2021年在利雅得上市以来，ACWA的市值已经增长了近三倍，目前为500亿美元。管理着沙特主权财富的公共投资基金（PIF）持有44%的股份。ACWA在国内外已运营的或处于收尾阶段的绿色项目的总装机容量达24GW，而2014年仅为0.3GW。如果加上其他在建项目，总装机容量达到54GW。ACWA以海水淡化业务起家，其淡化水产量从2006年的每天100万立方米增加到去年12月的每天760万立方米。ACWA上任不久的意大利老板马尔科·阿尔切利（Marco Arcelli）有丰富的能源业管理经验，他预计从现在到2030年ACWA参与投资的项目资产将增长两倍，达到2500亿美元。他希望这些项目能够帮助建立一个更广泛的国内绿色能源供应链。“我们是大型赋能者。”他表示。

ACWA蓬勃扩张之时，世界各地的许多其他可再生能源运营商却在艰难求生。上升的利率让这些竞争对手的项目成本飙升，而ACWA除了以单个项目为抵押来借债和从银行贷款，还从PIF获得了无息贷款以度过难关；同时它还筹集更多股权资本、引入合作伙伴。能轻易获得资金让ACWA得以扩大发电能力，同时降低客户的成本。这使得沙特太阳能发电的平准化成本（将发电厂的建设和运营成本都考虑在内）在全球处于最低水平。

然而按照全球标准来看，ACWA国内项目的回报率处于低水平。因此阿尔切利迫切想利用在国外可以拿到的利润更高的项目。他投资的项目有近有

远，近的有巴林、埃及、约旦、阿曼、土耳其和阿联酋，远的有阿塞拜疆、摩洛哥、南非和乌兹别克斯坦。ACWA总装机量有五分之二位于沙特以外。它还瞄准了中国，尽管这是个竞争激烈的市场，但阿尔切利认为ACWA可以形成业务规模，还能找到中国的风力涡轮机和太阳能电池板制造商做自己的技术合作伙伴。

ACWA面临巨大挑战。为了实现穆罕默德王储为它设立的国内目标，它必须在2030年前每年新增6至7GW的装机量，也就是要有三到四个大型项目。而目前它在国内处于各种开发阶段的项目加在一起也只有14GW。管理快速扩张需要高度聚焦成本（其新兴的氢能业务的成本已经比最初估计高出70%，达到80亿美元）。此外也需要借更多债。去年9月，ACWA的债务就已经达到70亿美元，相当于其税息折旧及摊销前利润的七倍。这样的比率在大多数公司都会被视作危险信号。

但ACWA仍可能临危不乱。它可以倚赖PIF的雄厚财力。而且它学得很快；花旗银行的奥利弗·康纳（Oliver Connor）认为，ACWA的体量1.5GW的国内太阳能项目Sudair可能会在两年多一点的时间内就全面建成并投入运营，这以行业标准来看已经很高效率。但阿尔切利希望项目进展得更快些。这也难怪，毕竟穆罕默德王储在背后盯着呢。■



Hacker heaven

Why is Brazil a hotspot for financial crime?

Its success as a fintech hub is mostly to blame

BRAZILLIANS HAVE long been early adopters of fintech. In 2017 EY, an accounting firm, found that two-fifths of Brazilians regularly used online banking, one of the highest rates worldwide. In 2020 44% of customers had a digital-only account, compared with less than 20% in the United States and Canada, according to a survey by Accenture, a consulting firm. That year the central bank released Pix, an instant-payments platform. It has been wildly successful. Today it has 3bn transactions a month. That is five times more than transactions by debit and credit cards combined.

This bonanza has attracted cyber-criminals. Their main weapon has been the “banking trojan”, a programme that steals users’ account information. According to Kaspersky Lab, a cyber-security firm, Brazil is the top country for attacks by banking trojans, with 1.8m attempted infections from June 2022 to July 2023 (the latest data available). Globally eight of the 13 most popular types of trojans are made in Brazil.

Cyber-criminals initially focused on trojans as they require little skill to use. However, as banks developed better defences, criminals were forced to branch out into more complex and lucrative attacks. Brazil’s underworld has developed the most advanced “point of sale” malware, which scammers use to filch bank details from card readers, according to Kaspersky Lab. Known as Prilex, this application can block contactless payments by stopping the short-range connection between a credit card and the payment terminal. The terminal reads: “Error. Please Insert.” When a customer inserts her card and PIN, the malware uses the credentials to authorise a fraudulent transaction. During Rio’s carnival in 2016, a hacker used a basic version of this software to remotely take over 1,000 ATMs.

Another example is ransomware, which gangs use to scramble computers and demand money to restore them. In October last year Brazil's lawmakers met to discuss the increasing use of artificial intelligence in cyber-crime, too.

The financial losses are big. According to Andre Fleury of Accenture, Brazil is in the top five countries for the cost of cyber-crime. He estimates the figure is around \$20bn per year. That is the equivalent of 0.9% of GDP. There is some hope, though. In 2022 a hefty data-protection law came into effect, forcing companies to defend consumers' data. In 2023 Brazil's banks spent \$9bn on cyber-security, nearly double the amount in 2019, according to the Brazilian Federation of Banks. The bigger problem is naive customers who fall for scams, says Eduardo Mênaco of ClearSale, a Brazilian fraud-management company. Until they fully know the risks, there will be plenty more phish in the sea. ■



黑客天堂

巴西为什么是金融犯罪的高发地？

成功跻身金融科技中心是主要原因

在采用金融科技方面，巴西人早就走在了前面。2017年，安永会计师事务所发现，五分之二的巴西人经常使用网上银行，这一比例在全球属于第一梯队。根据咨询公司埃森哲的一项调查，2020年，巴西44%的消费者拥有一个纯数字账户，而在美国和加拿大这一比例不到20%。同年，巴西央行推出了即时支付平台Pix并大获成功。如今，每月通过Pix进行的交易达30亿笔，是借记卡和信用卡交易总和的五倍。

这一富矿带吸引了众多网络犯罪分子。他们主要利用“银行木马”程序来窃取用户的账户信息。根据网络安全公司卡巴斯基实验室（Kaspersky Lab）的数据，巴西是遭受银行木马攻击最多的国家——从2022年6月到2023年7月（可获得的最新数据），未遂的银行木马攻击有180万次之多。全球最流行的13种银行木马中有八种是在巴西编写的。

网络犯罪分子最初主要利用木马，是因为使用它们不需要什么技能。但随着银行加强了防御，犯罪分子不得不转而采取更为复杂、也更能获利的攻击手段。卡巴斯基实验室称，巴西黑社会开发出了最先进的针对POS机的恶意软件，骗子利用该软件从读卡器窃取银行账户的详细信息。这款名为Prilex的应用程序可以通过中断信用卡与支付终端之间的短程连接来阻止非接触式支付。支付终端会显示：“错误。请插卡。”当顾客插入银行卡并输入密码时，这个恶意软件就会使用认证信息授权欺诈性交易。在2016年里约狂欢节期间，一名黑客使用该软件的基础版远程控制了1000多台自动取款机。

另一种方式是勒索软件，犯罪团伙用它来扰乱电脑，然后索要赎金以换取电脑恢复。去年10月，巴西的立法者也开会讨论了不断增长的利用人工智能实施网络犯罪的问题。

这带来了巨大的经济损失。埃森哲的安德烈·弗勒里（Andre Fleury）表示，巴西是为网络犯罪付出最高代价的五个国家之一。他估计巴西每年遭受的损失约为200亿美元，相当于其GDP的0.9%。不过情况有望获得改善。2022年，一部严格的数据保护法生效，强制要求企业保护消费者的数据。巴西银行业联合会（Brazilian Federation of Banks）表示，巴西的银行在2023年为网络安全花费90亿美元，几乎是2019年的两倍。巴西一家防欺诈公司ClearSale的爱德华多·莫纳戈（Eduardo Mônaco）表示，更大的问题是消费者太轻信，容易上当受骗。在他们完全了解风险之前，网络海洋中还会有很多诱饵等着鱼儿上钩。 ■



Towering ambition

Can India, Indonesia and Saudi Arabia be the next great economies?

Meet the countries making bold—and risky—bets on growth

POLITICIANS AND policymakers all over the world share a preoccupation: how to make their countries richer. The trouble is that the route to prosperity looks ever more daunting. The global economy is changing, as new, green technologies emerge and trading relationships fragment. In countries that are already rich the state, after decades of free-market rhetoric, is back in a big way. Governments are spending hundreds of billions on handouts for industries they deem to be strategically important.

In the face of this, many developing countries' ideas for growth are staggeringly ambitious. India and Indonesia hope to become high-income countries within 25 years. Muhammad bin Salman, Saudi Arabia's crown prince, wants to diversify and develop its economy just as rapidly. Refreshingly, such plans are more outward-looking than many development strategies of old. But they contain pitfalls, too.

In many ways, the developing world is choosing to bank on globalisation. Indonesia wants a bigger role in green supply chains. It seeks to do everything from mining and refining nickel, even to building the electric vehicles that run on it. It then wants to export the finished products to the rest of the world. Countries in the Gulf want to become attractive homes for global business, and are opening up to flows of people, cargo and cash. Narendra Modi envisions India as a high-tech manufacturer for the world, churning out microchips and smartphones.

That is a welcome shift. Less than 50 years ago India hoped to grow by closing itself off from the global economy. It turned out to be an approach that failed miserably. Some still suggest that India's domestic demand

could carry its growth.

But serving foreign markets plays a vital role in development. It keeps firms honest, by forcing them to compete in markets that their governments do not control. It lets them reach the largest possible scale. And foreign customers can teach firms how to serve them better. In East Asia export performance was also a useful yardstick for policymakers, because it revealed which industries deserved their continued backing.

Nonetheless, today's development strategies also hold dangers. In many countries governments are running the risk of warping the economy in the name of nurturing it. Saudi Arabia's onslaught of industrial policy, mainly disbursed as handouts from the Public Investment Fund, exceeds the spending even of America's Inflation Reduction Act. In order to help exporters grow, India is seeking to fence off its high-tech manufacturers behind tariffs and subsidies. Indonesia's all-in bet on nickel leaves it perilously exposed, should other battery chemistries prevail.

The rich world's new-found zeal for protectionism may make it tempting for poorer countries to follow suit. Yet floods of cash and shelter from foreign competition make it impossible to know whether a government's development gambles are paying off. A bet on one technology could go wrong if others emerge.

Parts of the developing world have paid dearly to learn these lessons before. For most of the 1960s Africa's policymakers had the same ideas as East Asia's, and the continent grew as fast, until picking the wrong champions left it languishing between 1975 and 1985. It is the poorest region in the world today.

Picking winners is also harder than it was 60 years ago. Then the choice was over which form of manufacturing to promote. Cheap, abundant

workforces gave poor countries an edge. Manufacturing was the only sector in which poor countries got better faster than rich countries.

Today, however, factories have become more capital-intensive. Though manufacturing still offers a way to boost a country's productivity, it is less certain to become a poor country's comparative advantage. That makes it even harder for policymakers to spot a good industry for them to place their bets. Rather than gambling with the public's money, they would be better off keeping it off the table.

There are, after all, plenty of other worthwhile things to spend it on. The state has a vital role in providing public goods by investing in infrastructure to stitch regions together, or education to boost workers' skills. That might still favour some industries over others. But if economies stay open, then they will at least experience the disciplines and benefits of trade.

| *Vision 2050*

The stakes are high. The developing world is home to over 6bn people and some of the most fragile democracies. Getting growth wrong would keep such places poorer for longer. That would be not just a human tragedy, but also a potential source of political instability. To avert it, the developing world needs to be bold—and resist the urge to build walls around itself. ■



【首文】凌云壮志

印度、印尼和沙特能否成为下一批大经济体？

来看看这些押下大胆而冒险的赌注追求增长的国家

全世界的政客和政策制定者都操心同一件事：如何让自己的国家更加富裕。问题是，通往繁荣的道路看起来越发艰巨了。随着绿色新技术的出现和贸易关系四分五裂，全球经济正在发生变化。在那些已经富裕起来的国家，自由市场的口号喊了几十年后，政府干预又大举回归。各国政府正在斥资数千亿美元补贴它们认为具有重要战略意义的产业。

面对这样的情形，许多发展中国家关于增长的想法展现出了惊人的雄心。印度和印度尼西亚希望在25年内成为高收入国家。沙特阿拉伯王储穆罕默德·本·萨勒曼（Muhammad bin Salman）希望以同样快的速度实现经济多元化和经济发展。这些计划比以往的许多发展战略更加开放外向，令人耳目一新。但它们也存在隐患。

在许多方面，发展中国家选择指望全球化。印尼希望在绿色供应链中发挥更大作用。它希望把触角伸到方方面面，要开采和提炼镍，甚至还要制造使用镍的电动汽车。接着它还想把制成品出口到世界其他地方。海湾国家希望成为对全球商业有吸引力的家园，正在向人员、货物和现金的流动开放。而在莫迪的展望中，印度将成为世界的高科技制造商，大量生产微芯片和智能手机。

这是一个值得欢迎的转变。将近50年前，印度希望通过将自己隔绝在全球经济之外来实现增长。事实证明，这种做法一败涂地。如今仍有一些人认为印度的内需可以支撑其经济增长。

但服务外国市场在发展中起到至关重要的作用。企业将不得不在其政府无法控制的市场上参与竞争，也就不得不诚信经营。服务外国市场还能让企业达到尽可能大的规模。外国客户可以教会企业如何更好地为他们服务。在东亚，出口业绩对政策制定者来说也是个有用的衡量标准，可以揭示哪

些行业值得他们继续支持。

然而，如今的发展战略也存在危险。在许多国家，政府面临着以培育经济为名扭曲经济的风险。沙特推出了一连串产业政策，主要通过公共投资基金（PIF）拨款，其支出甚至超过了美国的《通胀削减法案》。为了帮助出口商发展壮大，印度正试图筑起关税和补贴的高墙来保护本国的高科技制造商。倘若最终是其他电池化学材料大行其道，印尼对镍的全盘押注就将使它落入岌岌可危的境地。

看到富裕国家新萌生的保护主义热情，较贫穷国家可能也会想要效仿。然而，大量的现金补贴和免受外国竞争的庇护让人无法判断政府在发展上的押注是否在收获回报。如果其他技术崭露头角，对某一种技术的押注可能就会打水漂。

部分发展中国家此前付出了惨痛的代价得到这些教训。在20世纪60年代的大部分时间里，非洲政策制定者的想法与东亚政策制定者相同，非洲大陆的发展速度也同亚洲差不多，直到因为选择了错误的扶持对象而在1975年至1985年间裹足不前。非洲如今是世界上最贫穷的地区。

挑选国家冠军的难度也比60年前更大。当时只需要选择促进哪种形式的制造。廉价、充足的劳动力给贫穷国家带来了优势。制造业是穷国唯一比富国进步更快的部门。

然而今天，工厂已变得更加资本密集。尽管制造业仍然是提高一国生产率的途径之一，但它能否成为贫穷国家的比较优势却不再那么确定无疑。这就让政策制定者更难找到一个好的行业来押宝。相比拿民众的钱去赌，他们不如谨慎行事。

毕竟，还有其他很多地方值得花钱。国家在提供公共产品方面发挥着至关重要的作用，它可以投资基础设施，将各个地区连接起来，还可以投资教育，提高劳动者的技能。这可能仍然会导致对不同的行业厚此薄彼。但如果经济体保持开放，那么它们至少会体验到贸易带来的纪律约束和好处。

这一切事关重大。发展中世界有60多亿人口和一些最脆弱的民主国家。如果没有把增长这件事做对，这些地方就会久久无法摆脱更贫困的处境。这不仅会是一种人道悲剧，也可能成为政治不稳定的源头。为了避免这种情况的发生，发展中世界需要大胆行动，还要克制在自身周围筑起围墙的冲动。■



Stockmarkets

Is America's raging bull market exhausted, or taking a breath?

Investors have a slight hangover

IF YOU HAD an overindulgent Christmas, you may have begun the new year in a more austere frame of mind. Recent goings-on in the markets may therefore seem familiar. As 2023 drew to a close the American stockmarket was on a ripping run. It ended the year with nine consecutive weeks of gains, the longest winning streak since 2004. The S&P 500 index of leading American stocks was a whisker away from its all-time high, set on January 3rd 2022, when investors thought that interest-rate rises would be small and slow. Now punters are suddenly in a more sober mood, with stocks falling by 1.4% in the first two trading days of the new year. Such modest fluctuations are hardly unusual. Nonetheless, they raise the question of whether the blistering bull market is over, or has further to go.

For the first ten months of 2023, the market rally was largely concentrated in seven tech stocks, led by Nvidia, a maker of the computer chips that are used to process artificial-intelligence (AI) algorithms. Since then, however, it broadened and gained pace. Firms that mirror the wider economy, such as retailers and banks, soared—JPMorgan Chase is up by a quarter since late October. The S&P 500 rose by 14% in the final two months of 2023, and towers 31% above its most recent trough, well above the 20% that is often used to define a bull market.

The explanation for the run was a happy mix of strong economic growth, an orderly reduction of inflation and, crucially, an enormous shift in interest-rate expectations over the past two months. America's economy expanded at an impressive annualised pace of 4.9% in the third quarter; real-time estimates suggest it grew at a still-robust 2.5% in the last three months of the year. In the past three months "core" consumer prices have

risen at an average annualised pace of just 2.2%, only a smidgen above the Federal Reserve's inflation target.

That led to a big shift in investors' expectations for interest rates. In October they thought one-year rates in a year's time would be close to 5%. Thanks to lower inflation data and a doveish set of forecasts from the Fed, that has fallen to 3.5%. Bond investors see the central bank cutting rates as soon as March—and continuing in almost every meeting in 2024. This tantalising prospect of immaculate disinflation, robust growth and the promise of easier monetary policy has underpinned the rally.

Can the bull market be sustained? Asset prices still have room to rise. Although markets are close to the heights they reached after the protracted mania of 2021, that does not mean that things are as excessive now as they were then. In real terms, stock prices remain lower; valuations are therefore not quite as elevated. Participation by retail investors, which reached a giddy peak of 24% of daily trading volumes in early 2021, was steady at around 18% in 2023.

Moreover, although tech led the charge in both 2021 and 2023, investors this time have been discerning. They have lifted up Nvidia and Microsoft but Alphabet, Amazon and Tesla are all trading below their peak valuations. It is not just Americans excited about AI who are buoying stocks: in dollar terms European and Japanese equity indices are also within touching distance of their level two years ago.

| *Rude awakening*

Yet everything hangs on whether investors' ideal economic scenario comes to pass. The expectation that it will helped lift stocks close to a record high last year. But risks to the outlook abound, and may have given investors pause in the cold light of January. Inflation in America may not be fully vanquished, not least with the economy still in rude health and the fiscal

deficit unusually wide. Strife in the Middle East could cause another commodity-price shock; the one-time easing of the supply-chain disruptions of the pandemic may be keeping inflation low only temporarily.

A downturn may merely be delayed, not dodged. Rises in interest rates may not yet have fully fed through to borrowers. Indeed, history suggests that recessions are hard to spot in real time, and tend to catch out central banks. If a recession does not arrive, it is still possible that the Fed will not move with as much alacrity as investors hope. To see what will happen in the markets in 2024, watch the real economy. ■



【首文】股票市场

美国的红火牛市是精疲力竭了，还是暂时喘口气？

投资者略感宿醉

如果你在圣诞节肆意狂欢了一番，可能就会以更加冷静节制的心态进入新年。你对近期市场上的动态可能也就不会觉得陌生。2023年接近尾声时，美国股市一路狂飙，最终以连涨九周走完了2023年，创下自2004年以来的最长连涨纪录。主要股指标普500距离2022年1月3日（当时投资者预测加息将会是“小步慢跑”）的历史最高点仅一步之遥。现在，投资者突然变得更冷静了：股市在新年的头两个交易日下跌了1.4%。这种小幅波动并不稀奇，但还是引发了一个疑问：这轮大牛市是走到头了，还是能再往前一步？

在2023年的头十个月，美股上涨主要集中在七支科技股上，由制造处理人工智能（AI）算法的计算机芯片的英伟达领跑。之后上涨范围扩大，增速也加快。零售商和银行等反映更广泛经济的公司股价飙升，摩根大通的股价自10月底以来上升了四分之一。标普500指数在2023年最后两个月上涨了14%，比最近期的低点高出31%，明显高于通常用来定义牛市的20%。

对这轮牛市的解释是几方面因素恰好叠加：经济增长强劲、通胀有序回落，最关键的是过去两个月里利率预期的巨大转变。美国经济第三季度的年化增长达到可观的4.9%；实时估算显示，2023年最后一季度的增长仍相当稳健，达到2.5%。过去三个月里，“核心”消费价格的平均年化增长只有2.2%，仅略高于美联储的通胀目标。

这让投资者对利率的预期有了巨大转变。10月时，他们认为一年后的一年期利率将接近5%。随着通胀数据下降及美联储发布一系列鸽派预测，这个预期数字已降至3.5%。债券投资者认为美联储最快会在3月降息，而且在2024年几乎每次议息会议上都会继续降息。“完美去通胀”、增长强劲，以及货币政策有望放宽的诱人前景支撑了市场上涨。

牛市能否持续？资产价格仍有上升的空间。尽管市场已接近2021年持续大热后达到的高点，但这并不意味着现在就和当年一样过热。按实际价值计算，股票价格仍然低于当时，所以估值也没当时那么高。散户投资者的交易在2021年初达到令人眩晕的高峰，占日交易量的24%，2023年则稳定保持在18%左右。

此外，虽然在2021年和2023年领涨的都是科技股，但这次投资者是挑剔的。他们推高了英伟达和微软的股价，而Alphabet、亚马逊和特斯拉的股价则均低于各自的最高点。不仅是兴奋于AI发展前景的美国人在提振股市：以美元计算，欧洲和日本的股指也回升至接近两年前的水平。

| 猛然觉醒

然而，一切取决于投资者理想中的经济情景能否实现。认为能够实现的预期推动股市在去年创下历史高点。但这样的展望面临各种风险，可能会让投资者在经历新年的冷静思考后暂时停手。美国的通胀也许不会完全受控，尤其是在经济依然强劲而且财政赤字异常庞大的情况下。中东的冲突可能引发另一轮大宗商品价格冲击；新冠疫情过后，供应链中断的缓解是一次性的，这也许只能暂时压低通胀。

经济衰退可能只是延后了，最终仍难以躲开。加息的影响可能尚未完全传导至借款人。事实上，历史表明，经济衰退很难被及时发现，而往往会把央行打个措手不及。假如衰退不来，美联储仍有可能不像投资者希望的那般欣然快速行动。想知道2024年的市场走势，请关注实体经济。■



Free exchange

How to put boosters under India's economy

With the right policies, growth could be astonishing

LAND IN ANY Indian city, such as Bangalore or Hyderabad, and you will be struck by its heady optimism. India's economy may be in the early stage of a historic boom. Recently released figures show that economic growth roared to an annualised pace of 7.6% in the third quarter of 2023. In the past few weeks four international forecasters have raised their growth projections for the year, from an average of 5.9% to one of 6.5%. The National Stock Exchange of India is now neck-and-neck with Hong Kong's stock exchange for the title of the world's seventh-largest bourse.

Pause for breath, though, and India's performance looks a little less impressive. GDP growth has been slightly slower under Narendra Modi, India's prime minister, who was elected in 2014, than in the decade before. Labour-force participation is a paltry 40-50%, and only 10-24% for women. Subsidies are distorting the economy. A semiconductor plant in Gujarat will create 5,000 jobs directly and 15,000 indirectly. But a state handout covered 70% of its \$2.7bn cost. Assuming rather generously that the factory would not have been built without government support, each job cost \$100,000—nearly 40 times India's average income per person.

Grappling with the tension between India's enormous potential and an often messy reality is the task of a new book by Raghuram Rajan, a former governor of the Reserve Bank of India, and Rohit Lamba of Pennsylvania State University. The pair sketch out a vision that amounts to an entirely new model of development for India—one that they argue is better suited to its strengths than its current model. Three lessons stand out from their work.

The first is that India should stop fetishising manufacturing—an obsession born of East Asia's growth miracle. In the 1960s India's income per person was on a par with that of China and South Korea. By 1990 South Korea had taken off, while India remained level with China. Today China is three times richer and South Korea is seven times richer, adjusted for purchasing power. The growth of India's rivals was driven by low-skilled manufacturing, which received plenty of state support. Globalisation created a vast market, leading to previously unheard of double-digit growth rates. Once workers and companies got good at the easy stuff, they began to tackle more complex tasks with their newfound skills. Why shouldn't India follow its rivals' example?

As Messrs Rajan and Lamba explain, the problem is that East Asia has made manufacturing so competitive there is little profit left to be captured. Moreover, automation has reduced the number of available jobs—and manufacturing is no longer where value is to be found. Apple is worth \$3trn because it designs, brands and distributes its products. By comparison, Foxconn, which actually makes Apple's iPhones, is worth a mere \$50bn.

The second lesson concerns the export of services, which some in India's government think is a fresh way to tap into global demand. Modern technology, especially the internet, has made services far more tradable. Remote work has accelerated this trend. Meanwhile, governments around the world are desperate to shore up domestic industries. Partly as a result, global trade in goods has declined over the past decade. Yet trade in services has continued to grow. It is hard to argue against seeking a slice of the cushiest part of the global value chain, especially when the line between services and manufacturing is blurring. Some 40% of the value-added in a Chevrolet Volt, for instance, comes from its software.

In places, India is finding success. Its famed IT service sector has moved from mostly providing back-office work to more complex front-office fare.

According to one estimate, 20% of the global chip-design workforce can already be found in the country. But profound reforms will be required if India is to succeed more broadly. Spending on education as a share of GDP is 3-4%—middling relative to others of similar income. The bigger problem is that India appears to get little bang for its buck. By the latter half of high school, around half of students have dropped out. Bosses report that many of those who graduate are still not ready for work. Getting a new business off the ground is such a nightmare that many startups incorporate in Singapore. Labour laws make workers difficult to sack once they have been employed for more than a year, which incentivises the use of intermittent contracts. France and Italy have global brands, point out Messrs Rajan and Lamba. India does not. It is these sorts of problems that help explain why.

The last big item on the authors' wishlist is liberalism—of both the economic and political varieties. Politicians should start, they write, by jettisoning protectionism. From 1991, when India opened up to global markets, to 2014, when Mr Modi took power, average tariff levels fell from 125% to 13%. They have since risen to 18%, raising the cost of intermediate inputs for producers. India has refused to join regional free-trade agreements, which inhibits the ability of its exporters to reach customers abroad. And Mr Modi's authoritarian tendencies make it difficult for business leaders to criticise the government when a change of tack is required.

| *Hear the roar*

Messrs Rajan and Lamba paint a lovely picture of what could be. A better governed, more open India would be wonderful. But whether their ambitions are politically feasible is another question. For example, better public services probably mean devolving power from the central and state governments to localities. And who wants to give up power? Certainly not Mr Modi; probably not his rivals. Moreover, a country can endure quite a lot

of illiberalism before growth starts to falter. Until recently, China was humming along just fine. The Asian tigers only became more politically free when they were rich. India's economy is already growing at north of 6% a year with a policy mix that is far from the perfect.

In a strange way, though, this ought to provide Indian reformers with encouragement. Even if only half of what would be ideal is feasible, India's boom may only just be getting started. ■



自由交流

如何给印度经济加助推器

政策若得当，增长可能是惊人的

踏上印度任何一座城市，比如班加罗尔或海德拉巴，都会感受到那里洋溢的乐观情绪。印度经济可能正处于一段历史性繁荣的开端。近期公布的数据显示，在2023年第三季度，印度经济年化增长率飙升至7.6%。过去几周，四家国际预测机构上调了对2023年印度经济增长的预测，从平均5.9%调至6.5%。印度国家证券交易所（National Stock Exchange of India）现在与香港证券交易所旗鼓相当，竞逐世界第七大证券交易所的头衔。

但停下来缓口气，就会觉得印度的表现看起来没那么惊人了。在2014年当选印度总理的莫迪治下，印度的GDP增速略低于之前十年。劳动力参与率仅在40%至50%之间，女性更是只有10%至24%。政府补贴正在扭曲经济。古吉拉特邦的一家半导体工厂将直接创造5000个工作岗位，间接创造1.5万个。但政府补贴覆盖了其27亿美元成本的70%。做个豪气的假设，如果该工厂的建设没有政府支持，每个岗位的成本就会高达10万美元，几乎是印度人均收入的40倍。

印度前央行行长拉古拉姆·拉詹（Raghuram Rajan）和宾夕法尼亚州立大学的罗希特·兰巴（Rohit Lamba）合撰的新书探讨了印度的巨大潜力与往往混乱的现实之间的矛盾。二人为印度的发展勾画愿景，搭建出一个全新模式，他们认为比当前的模式更利于印度发挥自身优势。他们的研究结果凸显了三条经验教训。

首先，印度应停止痴迷制造业，这种迷恋源于当年东亚的增长奇迹。上世纪60年代，印度的人均收入与中国和韩国相当。到1990年，韩国经济腾飞，印度仍与中国持平。而现在，按购买力调整后，中国的人均收入是印度的三倍，韩国更是印度的七倍。印度这些竞争对手的增长是由政府大力支持的低技能制造业推动的。全球化创造了巨大的市场，带来了前所未有的两位数增长率。一旦工人和企业掌握了简单的技能，就会开始利用这些

新学到的技能处理更复杂的任务。那么印度为何不该效仿竞争对手？

正如拉詹和兰巴所解释的，问题是东亚已经让制造业竞争激烈到几乎无利可图的地步。此外，自动化减少了制造业所能提供的工作岗位，而且制造本身也不再是价值所在。苹果市值三万亿美元是因为它设计、打造品牌并销售产品。相比之下，实际制造苹果iPhone手机的富士康价值仅为500亿美元。

第二条经验是关于服务出口的。印度有些政府官员认为服务出口是接入全球需求的新途径。现代技术，特别是互联网技术，大大提高了服务的可交易性。远程工作加速了这一趋势。与此同时，世界各国政府都急于扶持本国产业，这在一定程度上导致全球货物贸易在过去十年出现下降。但服务贸易却持续增长。印度想从全球价值链中最轻松的环节分一杯羹，这没什么可反驳的，尤其是如今服务业与制造业之间的界限越来越模糊。举个例子，雪佛兰Volt汽车约40%的附加值来自其软件。

在某些领域，印度正在获得成功。它著名的信息技术服务业已从主要提供后台服务转向更为复杂的前台服务。据估计，印度的芯片设计人员已占到全球的20%。但要在更广泛的领域取得成功，印度还需要深刻的改革。在印度，教育支出占GDP的比例为3%至4%，与收入相近的其他国家相比处于中等水平。更大的问题是，印度的教育投入似乎成效不大。在中学的后半段约有半数学生辍学。而雇主则表示，那些完成了学业的毕业生有许多仍不能胜任工作。在印度要创办一家公司简直就是一场噩梦，许多创业公司宁愿在新加坡注册成立。现行劳动法让雇主很难解雇受雇一年以上的员工，这就促使他们采用断断续续的短期合同。拉詹和兰巴指出，法国和意大利拥有全球品牌，印度却没有。上述这些问题可能是部分原因所在。

拉詹和兰巴的愿望清单上的最后一个大项是自由主义——既包括经济上的，也包括政治上的。他们写道，政客们应从摒弃保护主义开始做起。从1991年印度向全球市场开放到2014年莫迪上台执政时，印度的平均关税水平从125%降至13%。而自此之后又上升到18%，提高了生产商的中间投入成本。印度拒绝加入区域自由贸易协定，抑制了印度出口商争取国外客户

的能力。在需要改变政策方针的时候，莫迪的威权倾向也让商界领袖难以批评政府。

| 听到轰鸣

拉詹和兰巴描绘了一幅可能出现的美好图景。一个治理改善、更加开放的印度将是很美好的。但他们的宏伟构想在政治上是否可行就是另一个问题了。例如，改善公共服务可能意味着要把权力从中央和邦政府下放到地方。谁愿意放弃权力呢？莫迪肯定不会，他的政治对手可能也不会。而且，一个国家有可能经受相当久的不自由，一时并不会出现经济增长减速。直到不久前，中国经济一直发展得不赖。亚洲小龙们是在富起来之后政治上才变得更自由的。在政策组合远非完美的情况下，印度的经济年增长率已超过6%。

不过，这应该能以一种奇怪的方式鼓舞印度的改革者。即使理想图景只有一半有望实现，印度的繁荣也可能才刚刚开始。■



4,000 terabits under the sea

Big tech and geopolitics are reshaping the internet's plumbing

Data cables are turning into economic and strategic assets

WHEN THE navies of Britain, Estonia and Finland held a joint exercise in the Baltic Sea earlier last month, their goal was not to hone warfighting skills. Instead, the forces were training to protect undersea gas and data pipelines from sabotage. The drills followed events in October when submarine cables in the region were damaged. Sauli Niinisto, the Finnish president, wondered whether the Chinese ship blamed for the mischief dragged its anchor on the ocean bed “intentionally or as a result of extremely poor seamanship”.

Submarine cables used to be seen as the internet's dull plumbing. Now giants of the data economy, such as Amazon, Google, Meta and Microsoft, are asserting more control over the flow of data, even as tensions between China and America risk splintering the world's digital infrastructure. The result is to turn undersea cables into prized economic and strategic assets.

Subsea data pipes carry almost 99% of intercontinental internet traffic. TeleGeography, a research firm, reckons there are 550 active or planned submarine cables that currently span over 1.4m kilometres. Each cable, which is typically a bundle of between 12 and 16 fibre-optic threads and as wide as a garden hose, lines the seabed at an average depth of 3,600 metres. Close to half have been added in the past decade. Newer ones are capable of transferring 250 terabits of data every second, the equivalent of 1.3m cat videos. Data may be stored in the cloud, but it flows under the ocean.

Since 2019 demand for international internet bandwidth has tripled to more than 3,800 terabits per second, estimates TeleGeography. The boom in data-hungry artificial intelligence may strengthen this trend. Synergy

Research Group, a data firm, predicts an almost three-fold increase in big cloud providers' data-centre capacity over the next six years. To connect these data centres to the internet, between 2020 and 2025 the data-cable industry will install 440,000km of new subsea lines.

One big shift has come from big tech. Until the early 2000s subsea cables were mainly used for transporting voice traffic across the world. Telecom operators like BT and Orange (formerly France Telecom) controlled most of the capacity. By 2010 the rise in data traffic led internet and cloud-computing giants—Amazon, Google, Meta and Microsoft—to start leasing capacity on these lines.

As their data needs surged, the tech firms began investing in their own pipes. In 2012 the four companies used around a tenth of international bandwidth; nowadays they claim almost three-quarters. Big tech's deep pockets ensure that projects are completed. According to Submarine Telecoms Forum, an industry body, only about half of all announced cable systems actually get built—unless they are backed by tech firms, in which case they almost always do.

Big-tech-backed cables account for almost a fifth of the \$12bn in planned investments in new systems over the next four years. Amazon and Microsoft part-own one and four networks, respectively. Meta owns one cable system outright and is an investor in another 14. Google is the most aggressive—the search giant directly owns 12 of its 26 cables. Last year it completed Firmina, a \$360m project that stretches more than 14,000km from the east coast of North America via Brazil to Argentina.

Dedicated cables allow the tech giants to avoid competing with others for third-party bandwidth, and to react quickly to changes in user demand and to any problems (if a cable on a route is damaged, data can be redirected to another one of the firms' lines). Alan Mauldin from TeleGeography points

out that being owner-operators also gives the tech giants the luxury of designing routes that meet their specific needs. Most telecom carriers rely on public “landing stations”—which connect the cables in the sea to customers’ data centres on land. By owning their cables, the companies can plug these more directly into their own data centres, speeding up traffic.

Their bandwidth and speed is further enhanced thanks to clever technology, which ownership makes easier to deploy. In 2019 Google introduced an innovation (“space division multiplexing”) that increased the number of fibre threads in a cable from 16 to 24. Last year it went further, doubling the number of “cores”—clusters of fibre threads—in its new TPU cable system that links Taiwan, the Philippines and America, increasing capacity while lowering the operating cost per bit.

All this is transforming the business of data cables. Having begun as large buyers of bandwidth from telecom companies, big tech is now leasing capacity on some of its cables to telecom operators. Legacy telecom firms are happy with this arrangement, since they face constant pressure from consumers for more capacity but, unlike big tech, they are desperately short of capital. As for the specialist companies which supply the equipment and lay the cables, these are go-go years.

Like many other global industries, the data-cable business is also being entangled in the tech contest between America and China—a second big shift. Take the Pacific Light Cable Network (PLCN). The 13,000km data pipeline was announced in 2016, with the backing of Google and Meta. It aimed to link the west coast of America with Hong Kong. By 2020 it had reached the Philippines and Taiwan. But in 2022 America’s government denied approval for the final leg to Hong Kong, worried that this would give Chinese authorities easy access to Americans’ data. Hundreds of kilometres of cable that would link Hong Kong to the network are languishing unused on the ocean floor.

America is stymieing China in another way. Laying cables at depth is a complicated job. Only a handful of contractors have the required chops. Three—Alcatel Submarine Networks from France, NEC from Japan and SubCom from America—receive more than 80% of the spending on cable construction. HMN Tech, a Chinese challenger spun out of Huawei, China's telecoms-gear champion, claims 9% of new annual construction spending. But amid Sino-Western tensions, new cables that have links to America, which is to say most of them, avoid HMN Tech as a supplier. Telecoms executives say they are discouraged from using HMN. In 2022 a lucrative contract for SEA-ME-WE 6, a 19,000km line owned by a group of telecoms operators including India's Bharti Airtel and Singapore's SingTel, and linking South-East Asia to Europe, was awarded to SubCom, even though HMN's bid was reportedly lower.

China is responding by charting its own course. PEACE, a 21,500km undersea cable linking Kenya to France via Pakistan, was built entirely by Chinese firms as part of China's "digital silk road", a scheme to increase its global influence. Reuters reported that this year three Chinese carriers—China Telecom, China Unicom and China Mobile Limited—are investing \$500m in a cable network that connects China and France via Singapore, Pakistan and Egypt. The project, to be built by HMN Tech, will compete directly with SEA-ME-WE 6.

Despite the growing Sino-American rivalry, from 2019 to 2023 bandwidth between the two has grown by 20% a year. American and Chinese mobile operators, which also rely on cables, continue to increase network connectivity in each other's territory. The necessary licences are, however, getting harder to secure.

In March America's Federal Communications Commission issued a proposal that would require licensees to provide more information about who owns them. It also acknowledged concerns that the presence in

America of physical infrastructure of China Telecom is “highly relevant to the national-security and law-enforcement risks”. All this is making the route taken by bits and bytes more circuitous than before, and thus costlier. If transpacific tensions continue to mount, those routes may one day vanish altogether. ■



海底4000太比特

科技巨头和地缘纷争正在重塑互联网的管道

数据电缆正在变成经济和战略资产

英国、爱沙尼亚和芬兰的海军于2023年12月上旬在波罗的海进行了联合演习，他们的目标并不是磨练战斗技能，而是操练如何保护水下天然气和数据管道免遭破坏。在此次演习开展前的10月，该地区的海底电缆被损坏。芬兰总统绍利·尼尼斯托在纳闷，被指为肇事者的中国船只究竟是“故意还是因为极差劲的航行技术”在海底拖曳了船锚。

海底电缆曾被视为互联网的平凡无奇的管道。如今，正当中美紧张关系可能令全球数字基础设施分崩离析之时，亚马逊、谷歌、Meta和微软等数字经济巨头在宣示对数据流的更大控制权。其结果是，海底电缆变成了备受重视的经济和战略资产。

海底数据管道承载了几乎99%的洲际互联网流量。研究公司TeleGeography估计，目前有550条在用或计划建造的海底电缆，总长超过140万公里。每根电缆通常包裹着12到16根光纤线，粗细如同花园浇水软管，沿海底铺设，平均深度约为3600米。近一半的电缆是在过去十年里铺设的。较新的电缆能够每秒传输250太比特的数据，相当于130万段猫咪视频。数据可能被存储在云端，但它是在海底流动的。

据TeleGeography估计，自2019年以来，对国际互联网带宽的需求增长了两倍，达到每秒3800太比特以上。数据需求巨大的人工智能的腾飞可能加强这一趋势。数据公司Synergy Research Group预测，在未来六年内，大型云服务提供商的数据中心容量将增加近三倍。为了将这些数据中心连接到互联网，从2020年到2025年，数据电缆行业将新铺设44万公里的海底电缆。

一个大转变源自科技巨头。直到本世纪初，海底电缆还主要用于在全球范围内传输语音流量。像英国电信（BT）和Orange（原法国电信）这样的电

信运营商控制了大部分容量。到2010年，由于数据流量的增加，互联网和云计算巨头亚马逊、谷歌、Meta和微软开始租用这些线路上的容量。

随着它们的数据需求激增，科技公司开始投资自己的电缆。2012年，这四家公司使用了大约十分之一的国际带宽；如今它们几乎占用了三分之二。科技巨头的雄厚资金确保了项目能够完成。据行业组织海底电信论坛（Submarine Telecoms Forum）统计，所有公布投建的电缆系统中只有约一半左右最后会实际完成——除非是由科技公司支持的，那么几乎总能完成。

在未来四年总共120亿美元的新电缆投资计划中，科技巨头支持的项目占到近五分之一。亚马逊部分拥有一个电缆网，微软部分拥有四个。Meta完全拥有一个电缆系统，还是另外14个电缆系统的投资者。谷歌是最积极的——这家搜索巨头直接拥有其26条电缆中的12条。去年，谷歌完成了费尔米纳（Firmina）项目，耗资3.6亿美元，全长超过1.4万公里，从北美东海岸经巴西抵达阿根廷。

专用电缆使科技巨头能够避免与其他公司争夺第三方带宽，并迅速应对用户需求的变化和任何问题（如果某个线路上的电缆损坏，数据可以转移到这些公司的另一条电缆上）。TeleGeography的艾伦·莫尔丁（Alan Mauldin）指出，拥有并运营电缆也使科技巨头得以根据自己的需求设计线路。大多数电信运营商要依赖公共“登陆站”——它们将海底电缆连接到陆地上的客户数据中心。拥有属于自己的电缆，这些公司就可以将电缆更直接地连接到自己的数据中心，加速数据传输。

自己拥有电缆也方便了部署先进技术，进一步提升带宽和速度。2019年，谷歌推出了一项创新（“空间分集复用”），将电缆中的光纤线数量从16根增加到24根。2023年，谷歌更进一步，在其连接台湾、菲律宾和美国的新TPU电缆系统中将“核心”也就是光纤簇的数量翻了一番，增加容量的同时降低了单位流量的运营成本。

所有这些都在改变数据电缆生意。科技巨头最初是从电信公司大量购买带

宽，现在正将自家一些电缆的容量出租给电信运营商。传统电信公司乐于接受这种安排，因为它们持续面临着消费者要求更多带宽的压力，但与大科技公司不同，它们极度缺乏资金。至于那些专门提供相关设备和铺设电缆的公司，它们迎来了生意兴隆的好年景。

与许多其他全球性产业一样，数据电缆业务也被卷入了中美科技战——这是第二个大转变。太平洋光缆网络（Pacific Light Cable Network, PLCN）就是一个例子。这条长1.3万公里的数据管道项目于2016年宣布，得到谷歌和Meta的支持。它想要连接美国西海岸和香港。到2020年，它已经铺设到了菲律宾和台湾。但是2022年，美国政府拒绝批准最后一段通往香港的线路，担心这将让中国当局轻松获取美国人的数据。原本将把香港连接到该网络的数百公里电缆就此被弃用，闲置在海底。

美国也在以另一种方式阻挠中国。在深海铺设电缆是一项复杂的工作。只有少数几个承包商具备所需的技术。全球海底电缆建设支出的超过80%支付给了三家公司——法国的阿尔卡特海底网络（Alcatel Submarine Networks）、日本的NEC和美国的SubCom。中国电信设备巨头华为旗下的中国挑战者华海通信声称在新年度建设支出中拿到9%。但在中西方关系紧张的背景下，与美国有关的新电缆——也就是说大多数电缆——都在绕开华海通信这个供应商。电信业高管表示，他们被劝阻使用华海通信。2022年，连接东南亚与欧洲的SEA-ME-WE 6这个大项目的合同被授予了SubCom，尽管报道称华海通信在投标中的报价更低。SEA-ME-WE 6电缆长1.9万公里，由印度的Bharti Airtel和新加坡电信（SingTel）等一群电信运营商拥有。

中国的回应是绘制自己的电缆航道。PEACE是一条长2.15万公里的海底电缆，连接肯尼亚与法国，中间经过巴基斯坦。它完全由中国公司建设，是中国提升其全球影响力的“数字丝绸之路”的一部分。路透社报道，今年三家中国运营商——中国电信、中国联通和中国移动——正在投资5亿美元建设一条通过新加坡、巴基斯坦和埃及连接中国和法国的电缆网络。该项目将由华海通信建设，会直接与SEA-ME-WE 6竞争。

尽管中美之间的竞争日益激烈，但从2019年到2023年，两国之间的带宽每年增长20%。美国和中国的移动运营商也依赖电缆，都在继续增加在对方区域内的网络连接。然而，获取必要牌照的难度在加大。

2023年3月，美国联邦通信委员会（FCC）提出了一项提案，要求牌照持有人提供更多其所有者的信息。它还承认担忧中国电信在美国境内建设的实体基础设施“与国家安全和执法风险高度相关”。所有这些都使比特和字节的传输路径变得比以往更加迂回，因此成本更高。如果太平洋两岸的紧张局势继续升级，这些路径某天可能会完全消失。■



Desert island dismal science

Why economists love “Robinson Crusoe”

The classic yarn of a shipwrecked sailor reveals a lot about scarcity

AFTER SPENDING 28 years, two months and 19 days marooned on an island, Robinson Crusoe does not lose his nose for adventure or his “native propensity to rambling”. He crosses the Pyrenees, stalked by “hellish wolves”, witnesses the “pomp and poverty” of China and battles Tartars on the Russian steppe.

The character’s strangest adventure, however, is none of these. It is surely his centuries-long ramble through the literature of economics. Crusoe has appeared in Karl Marx’s “Das Kapital”, John Maynard Keynes’s “General Theory” and Milton Friedman’s Chicago lectures on “Price Theory”. He has an entry in the New Palgrave Dictionary of Economics. And he often washes up in economics textbooks.

Crusoe’s economic appeal is unsurprising. The sailor spends a few pages escaping pirates and shooting cannibals. But his real battle is against scarcity, which he defeats through careful deployment of the resources at his disposal, including his own labour.

After being shipwrecked, Crusoe makes his island prison habitable, even hospitable. Salvaging what he can from the wreck, he fortifies a cave (his “castle”), erects a tent (“my country house”), plants crops, tames goats (and a parrot) and fills his improvised shelves with pigeon, turtle and other foodstuffs.

Scarcity also stalked Daniel Defoe, the novelist who created Crusoe in 1719. Over a chequered career he traded in bricks, wines, pickles, tobacco and the glands of civet cats. He dabbled in horse-trading. Literally. He defaulted on

his debts. Twice. “No man has tasted differing fortunes more,” he wrote. “And thirteen times I have been rich and poor.”

He wrote allegories that turned dry economic variables into colourful characters like “Count Tariff”, an English nobleman dressed in domestically manufactured cloth, and “Lady Credit” (“if she be once Disoblig’d; no Entreaties will bring her back again). His publication “The Compleat English Tradesman” has been described as the first business textbook.

But it is his island fable that has most resonated, as Michael White of Monash University has documented. Economists are eager to find behavioural laws that apply anywhere. Crusoe’s isolation thus provides a useful thought experiment. Principles that hold true on his island must be elemental, not socially incidental.

William Forster Lloyd, for example, was keen to show that economics had something to say about value even in the absence of markets and exchange. In a publication in 1834, he pointed out that Crusoe prizes his goods more dearly as they become more scarce (“my ink beginning to fail me”, Crusoe says, “I contented myself to use it more sparingly”). He took that as evidence for the principle of diminishing marginal utility: a second bottle of ink is worth less than the first.

Most economists have turned to the tale not to corroborate a theory but merely to illustrate it. Textbook authors, for example, want to introduce the principles of supply and demand in the simplest possible case, and nothing is simpler than a one-person “Robinson Crusoe” economy.

Such an economy features in a textbook by Hal Varian, chief economist at Google. Crusoe must decide how to divide his day between gathering coconuts and working “on his tan”. In keeping with diminishing marginal utility, each extra coconut or hour of sunbathing is worth less than the last.

Each hour of work also yields fewer coconuts than the last. Under these assumptions, Crusoe should stop working at the point when an extra coconut is worth no more to him than the additional leisure he must sacrifice to gather it.

A one-person economy has several things going for it. There is no waste. If an extra coconut is not wanted, it will not be collected—supply implies its own demand. There is no unemployment. If Crusoe wants the extra coconut more than the leisure, he will employ himself to gather it. Such an economy, Keynes pointed out, cannot suffer the kind of slump that cursed the 1930s—when people fail to spend enough of their income on the goods the economy could produce.

Textbooks present Crusoe's one-man economy as a kind of benchmark, against which more sophisticated economies can be judged. Can its harmony be replicated, even when decision-making is divided up and dispersed—even when consumers and producers do not share the same mind?

The answer is yes, through the magic of flexible prices and wages. In his own more elaborate version of the parable, Daniel McFadden, a Nobel prize-winning economist who was also Mr Varian's thesis adviser, introduces a second character ("Friday"). In this version, Crusoe gathers yams not coconuts. Friday acts as a manager, hiring Crusoe's labour, paying him in yams, and giving him leftover yams as a "dividend".

Mr McFadden shows that there is an hourly wage that will reconcile the demand and supply of labour, and also, miraculously, the demand and supply of yams. But things can go wrong if wages get misaligned or expectations sink too low. If the wage gets stuck at too high a level, for example, Crusoe might find himself unable to work as long as he wants. The yams he could collect in an extra hour may be worth more to him than

the leisure he would lose. But if the wage he must receive is higher still, Friday will deny him the extra employment. The island would suffer a recession, combining unmet needs (for yams) with unused resources (Crusoe's spare labour).

If Friday worries that he will not be able to sell as many yams as he can produce, he may limit his demand for labour. That will curb his customer's purchasing power, thereby seeming to bear out his pessimistic sales forecast. Crusoe will lack work, because Friday lacks sales. And Friday will lack sales, because Crusoe lacks work.

An obvious objection to these parables is their cartoonishness. The concept of a Crusoe economy has become "another cuss-word to people who crave realism and are contemptuous of theory" noted Frank Knight, a Chicago economist, in 1960. But simplification can often aid understanding. Mr McFadden's parable, for example, illustrates that recessions are not necessary or salutary, but absurd and inefficient.

| *Never too late to be wise*

For Crusoe-lovers, however, what is most striking about these exercises is not their distance from reality, but their distance from Defoe's original tale. Neither coconuts nor yams appear in the book. And far from working on his tan, Crusoe took a "world of pains" to hide from the sun, making a "clumsy, ugly, goat's-skin umbrella" to ward off its rays. His island is not in the South Seas, as Mr McFadden maintains, but near Trinidad. And Friday and Crusoe do not bargain over labour or anything else. After Crusoe saves him from the cannibals who have carried him to the island by canoe, Friday in effect indentures himself to the sailor. One of the first English words he is taught is "Master".

Obliviousness to Defoe's tale does not invalidate the textbook parables. Little of importance hangs on whether Crusoe gathers coconuts, yams or

grapes, in the South Seas or anywhere else. But the neglect is nonetheless a missed opportunity. There is a lot of economic incident and insight in the original story. Economists might enjoy rediscovering it.

They could start not with coconut-gathering but with bread-making. “Few people have thought much upon...the strange multitude of little things necessary in the providing, producing, curing, dressing, making, and finishing [of] this one article of bread,” Crusoe says, as he struggles to make some for himself. In trying to start from near-scratch, Crusoe discovers that even the simplest product is a minor miracle of economic choreography. His thoughts resemble the classic essay, “I, Pencil”, written by Leonard Read in 1958, which details the “genealogy” of the humble pencil, with its wood from Oregon, graphite from Sri Lanka and rubber from Indonesia, all collected, transported and refined by machines that have their own even more complex genealogy.

After bread-making, economists could turn to Crusoe’s pottery. It takes him about two months to make a pair of jars—“two large, earthen ugly things”—in which to store his grain. Preserving resources is no easy matter: pests threaten his crops and decay unravels his clothes. In his 1916 book “The Natural Economic Order”, Silvio Gesell imagines how grateful Crusoe would be to lend his spare provisions to another islander, like Friday, in return for similar provisions a few years hence. He would accept the deal even if Friday pays no interest, because merely keeping wealth intact represents a victory against the relentless forces of decay. It is a useful thought experiment for anyone who resents today’s financial system, which for all its flaws, allows people to preserve their wealth in convenient savings accounts, not misshapen jars.

The Crusoe in the textbooks is a rational man, always equating marginal this with marginal that. He is the stock character of economics 101. The Crusoe in Defoe’s story is more mercurial and conflicted. As such, he lends

himself to more recent, psychologically informed theories of decision-making. He could become an icon of “behavioural economics”.

| *The want of thankfulness for what we have*

At one point, Crusoe uses his scarce ink to take stock of his predicament, drawing up a kind of balance-sheet of comforts and miseries, credits and debits. He is a lone castaway (a debit), but he is alive (a credit). The island is uninhabited, but it is not barren. He has no defences, but the island has no obvious predators. No companion survived the wreck, but provisions could be salvaged from it.

Daniel Kahneman, a psychologist who won the Nobel prize in economics, and Amos Tversky have shown that when assessing their lives, people often evaluate not their level of well-being, but their gains or losses from some “neutral” reference point.

The choice of reference point is not always obvious. On each line of his balance sheet, Crusoe entertains alternatives. His shipwrecked isolation represents a grievous loss from where he was. But it counts as a gain from an alternative scenario—not hard to imagine—in which he drowned or washed up on a more perilous shore. Mr Kahneman and Tversky point out that in dreaming up these alternative scenarios, people follow certain rules. They reimagine the chain of events leading up to their predicament, removing any strange or surprising twists of fate.

After Crusoe abandons the wrecked ship, it drifts closer to shore, allowing him to return to it and strip it bare. That, Crusoe recognises, was unlikely (100,000 to one, he says). It is therefore easy for him to imagine an alternative reference point in which he rescued nothing from the wreck. That helps him psychologically.

Behavioural economists stress that more choice is not always better. People

may be unable to resist choices they know will hurt them in the long run. Choice also invites regret. It obliges us to compare our fate with the alternative we could have chosen. Forced to stay on his island, Crusoe can be happy. But if he were to choose his isolation, he would be haunted by the alternative life he could have chosen elsewhere.

Indeed, Crusoe later concludes that he can be happier within the tight compass of his island than he would be in the outside world, where he had once lived a “wicked, cursed, abominable life”. He also immediately admits to himself that if offered the chance to escape, he would nonetheless take it.

As these examples show, economists might profit from greater familiarity with the Crusoe story. And the trade could be two-way. Defoe scholarship could and has benefited from a closer acquaintance with economics. There are several corners of Defoe’s works that require some economic knowledge to appreciate.

When Crusoe embarks on his ill-fated voyage to Guinea to buy slaves, he leaves behind a growing tobacco plantation in Brazil that would soon be worth “three or four thousand pounds”. It is hard for a reader today to make sense of such a figure. Drawing on the work of economic historians, David Spielman, formerly of Penn State University, calculates that the income on such a sum would have put Crusoe in the top 5% of English families at the time. With so much wealth in prospect, Crusoe has no reason to take risks. His voyage was as “preposterous” as he himself admits.

Economists might also resolve some other mysteries. After his return from the island, Crusoe reclaims his plantation and sells it. In the first six editions of the book, he receives 328,000 pieces of eight, worth about £72,000. But in later ones, a zero is deleted. That matters for the interpretation of the story. Does Crusoe finish the novel a rich man or a

very rich one?

Literary scholars pride themselves on their sensitivity to every nuance of a text. But the decimation of Crusoe's wealth has barely registered. "Despite the careful attention that the textual history of Robinson Crusoe has received, no one has even noticed a problem," Mr Spielman has pointed out. Economists may have lost sight of Robinson Crusoe's richness. But literary scholars have overlooked most of his riches. ■



荒岛的惨淡科学

为什么经济学家喜爱《鲁滨逊漂流记》

沉船水手的经典故事揭示了关于稀缺性的许多问题【深度】

在孤岛上度过了28年零2个月又19天后，鲁滨逊·克鲁索并没有失去对冒险的嗅觉，也没有丢掉“四处游逛的天性”。他在“地狱般的狼群”的追逐下穿越了比利牛斯山脉，目睹了中国的“浮华与贫困”，在俄罗斯大草原上与鞑靼人作战。

然而，这个角色最奇特的冒险并不是这些，而无疑是他几个世纪以来在经济学文献中的漫游。鲁滨逊曾出现在卡尔·马克思的《资本论》、约翰·梅纳德·凯恩斯的《就业、利息和货币通论》和米尔顿·弗里德曼在芝加哥主讲的《价格理论》中。《新帕尔格雷夫经济学词典》中有他的条目。而且他还经常冲上经济学教科书的岸头。

鲁滨逊对于经济学的吸引力并不让人惊讶。这个水手花了几页的时间逃离海盗并射杀食人族。但真正的战斗是对抗稀缺——他通过仔细部署他所掌握的资源（包括他自己的劳动力）来战胜稀缺性。

在遭遇海难之后，鲁滨逊把他的小岛监狱改造得适合居住，甚至可说宜人。他从沉船中打捞出一些东西，加固了一个山洞（他的“城堡”），搭建了一个帐篷（“我的乡间别墅”），种植庄稼，驯养了山羊（和一只鸚鵡），在简易架子上摆满了鸽子、乌龟和其他食物。

在1719年创作了《鲁滨逊漂流记》的小说家丹尼尔·笛福也被稀缺所困。在他坎坷的职业生涯中，他做过砖块、葡萄酒、泡菜、烟草和麝猫腺体的贸易。他还搞过马匹交易。他曾债务违约。两次。“没有人比我更饱尝命运转折，”他写道，“我经历过十三次贫富起落。”

他撰写的寓言故事将枯燥的经济变量转化为丰富多彩的人物形象，如“关税伯爵”——一位身穿家纺布料的英国贵族，以及“信用夫人”（“一旦对她

失信，再多的恳求也无法使她回心转意”）。他出版的《英格兰商人全书》（The Compleat English Tradesman）被誉为第一本商业教科书。

但正如莫纳什大学的迈克尔·怀特（Michael White）所写，他的荒岛寓言最能引起共鸣。经济学家渴望找到适用于任何地方的行为规律。因此，鲁滨逊的与世隔绝提供了一个有用的思想实验。在他的岛上适用的原则必然是基本的，而不是因社会而偶然形成的。

例如，威廉·福斯特·劳埃德（William Forster Lloyd）热衷于证明，即使在没有市场和交换的情况下，经济学也能对价值做出解释。他在1834发表的一篇著作中指出，随着商品变得越来越稀缺，鲁滨逊越来越珍视它们（“我的墨水开始不好用了，”鲁滨逊说，“我满足于更加节约地使用它”）。他将此作为边际效用递减原理的证据：第二瓶墨水的价值低于第一瓶。

大多数经济学家引用这个故事并不是为了证实某个理论，而只是为了阐释。例如，教科书的作者们希望在一个尽可能简单的案例中介绍供求原理，而最简单的莫过于一个人的“鲁滨逊·克鲁索”经济了。

谷歌首席经济学家哈尔·瓦里安（Hal Varian）在一本教科书中描述了这样的经济。鲁滨逊必须决定如何将一天的时间在采摘椰子和“晒黑皮肤”上做分配。根据边际效用递减原理，每多收一个椰子或多晒一小时太阳，其价值都会低于上一个椰子或上一小时日光浴。每工作一小时产出的椰子也比上一小时少。根据这些假设，鲁滨逊应该在一个额外椰子的价值已经比不上他为采这只椰子而必须牺牲的额外闲暇时停止工作。

单人经济有几个好处。没有浪费。如果额外的椰子不被需要，它就不会被采摘——供应量就意味着经济本身的需求量。没有失业。如果鲁滨逊想要额外的椰子而不是闲暇，他就会亲自去采摘。凯恩斯指出，这样的经济不会遭受20世纪30年代那样的衰退——当时人们没能将足够的收入用于购买经济所能生产的商品。

教科书将鲁滨逊的单人经济视为一种基准，我们可以用它来评判更复杂的

经济。即使决策被分割和分散——即使消费者和生产者的想法不同，这种和谐能否被复制？

答案是肯定的，这要利用弹性价格和工资的魔力。获得诺贝尔奖的经济学家，也是瓦里安的论文导师的丹尼尔·麦克法登（Daniel McFadden）为这个寓言写了一个更详细的版本。他引入了第二个角色（“星期五”）。在这个版本中，鲁滨逊采集的是山芋而不是椰子。“星期五”充当经理，雇用鲁滨逊的劳动力，用山芋支付报酬，并将剩余的山芋作为“红利”分给他。

麦克法登证明了，存在这样一个时薪水平，它可以协调劳动力的供求关系，还能奇迹般地协调山芋的供求关系。但是，如果工资偏离这个水平或期望值过低就会出问题。举例来说，如果工资停留在过高的水平，鲁滨逊可能会发现自己并不能想工作多久就工作多久。他多工作一小时所能收获的山芋可能对他来说比失去的闲暇更有价值。但是，如果他必须得到的工资比这个价值还要高，“星期五”就会拒绝让他额外工作。这样一来，岛上的需求（山芋）得不到满足，资源（鲁滨逊的剩余劳动力）得不到利用，经济就会出现衰退。

如果“星期五”担心自己生产的山芋卖不出去，他可能会限制对劳动力的需求。这将抑制他的客户的购买力，从而似乎证实了他悲观的销售预测。鲁滨逊缺乏工作，因为“星期五”销售不足。“星期五”销售不足，因为鲁滨逊缺乏工作。

对这些寓言的一个明显反对意见是它们太卡通了。芝加哥经济学家弗兰克·奈特（Frank Knight）在1960年指出，“克鲁索经济”的概念已成为“又一个咒骂那些渴望现实主义且蔑视理论的人的字眼”。但简化往往有助于理解。例如，麦克法登的寓言说明，经济衰退并非必要或有益，而是荒谬和低效的。

| 明智永远不嫌太晚

然而，对于鲁滨逊爱好者来说，这些思想实验最惹眼的地方不是它们与现实的距离，而是它们与笛福原著故事的差异。书中既没有椰子，也没有山

芋。鲁滨逊不仅不求晒黑皮肤，反而“费尽千辛万苦”躲避太阳，制作了一把“笨拙、丑陋的山羊皮伞”来遮挡阳光。他的岛并不像麦克法登所说的那样在南海，而是在特立尼达附近。“星期五”和鲁滨逊并没有就劳动力或任何其他问题讨价还价。食人族用独木舟把“星期五”带到了岛上，鲁滨逊把他解救出来，之后“星期五”实际上成了鲁滨逊的奴仆。他学到的头几个英语单词里就有“主人”。

对笛福原著的忽视并不意味着教科书上的寓言就失效了。鲁滨逊是在南海还是其他地方采集椰子、山芋或葡萄并不重要。但这种忽视仍然是错失了机会。原著中有很多经济事件和见解。重新发现它可能会让经济学家欣喜。

他们可以不从采集椰子开始，而是从制作面包开始。“很少有人考虑过.....在提供、生产、腌制、调味、制作和完成这一个面包时所必需的奇怪的一大堆小事。”鲁滨逊在费劲给自己制作一些面包时说道。在尝试几乎是从头开始的过程中，鲁滨逊发现，即使是最简单的产品也是经济编排的一个小小奇迹。他的想法类似于伦纳德·里德（Leonard Read）在1958年写的经典文章《我，铅笔》（“I, Pencil”），其中详细介绍了不起眼的铅笔的“谱系”：其木材来自俄勒冈州，石墨来自斯里兰卡，橡胶来自印度尼西亚，所有这些都由拥有还更复杂谱系的机器进行采集、运输和精制。

在看完了面包制作之后，经济学家可以看看鲁滨逊的陶器。他花了大约两个月的时间制作了一对罐子——“两个又大又丑的土制东西”——来储存谷物。保护资源绝非易事：害虫威胁着他的庄稼，他的衣服开始腐烂毁坏。西尔维奥·格塞尔（Silvio Gesell）在1916年出版的《自然经济秩序》一书中想象到，如果鲁滨逊能将自己的备用物资借给另一个岛民（比如“星期五”）以换取几年后类似的物资，他会多么感恩。即使“星期五”不支付利息，他也会接受这笔交易，因为仅仅保持财富完好无损就代表着对无情的自然摧毁力量的胜利。对于那些憎恨当今金融体系的人来说，这是一个有用的思想实验：尽管金融体系存在种种缺陷，但它让人们得以将财富保存在方便的储蓄账户而不是畸形的罐子里。

教科书中的鲁滨逊是一个理性人，总是在平衡边际这个和边际那个。他是经济学101的典型人物。笛福故事中的鲁滨逊更加善变和矛盾，让他更适用于阐释更新近的、受心理学启示的决策理论。他可能成为“行为经济学”的偶像。

| 对我们所拥有的事物缺乏感恩之心

有一次，鲁滨逊用他稀缺的笔墨来评估自己的困境，绘制了一份舒适与痛苦的资产负债表，有贷项有借项。他是一个孤独的漂流者（借项），但他还活着（贷项）。该岛无人居住，但并不荒芜。他没有防御措施，但岛上没有明显的掠食者。没有同伴在沉船中幸存，但可以从船上打捞出给养。

获得诺贝尔经济学奖的心理学家丹尼尔·卡尼曼（Daniel Kahneman）和阿莫斯·特沃斯基（Amos Tversky）指出，在评估自己的生活时，人们通常不是评估自己的幸福水平，而是从一些“中立”参考点出发来评估自己的得失。

参考点的选择并不总是显而易见的。在资产负债表的每一行上，鲁滨逊都考虑了其他选择。他遭遇海难，与世隔绝，这代表着他相对于之前身处之地是遭受了严重损失。但这相对于另一种情景来说则是收益——不难想象，他也可能被淹死或被冲到更危险的海岸上。卡尼曼和特沃斯基指出，在想象另外这些场景时，人们会遵循某些规则。他们重新想象导致他们陷入困境的一系列事件，消除任何奇怪或令人惊讶的命运转折。

鲁滨逊抛弃了失事的船后，船漂到了离岸更近的地方，这让他可以回到船上，把能拿的东西扒了个精光。鲁滨逊意识到这是极为难得的好运气（他说是十万分之一的几率）。这就使他很容易想象出另一个参照点：他从沉船上什么也没拿回来。这对他的心理很有帮助。

行为经济学家强调，选择并非越多越好。人们可能无法抵抗明知从长远来看会伤害自己的选择。选择的存在也会让人心生遗憾。它迫使我们把自己的命运与我们本可以选择的其他方案比来比去。当鲁滨逊被迫留在岛上，他可能会很快乐。但如果他是自己选择了与世隔绝，那他本可以选择的另

一种生活就会让他难以释怀。

事实上，鲁滨逊后来得出结论，他在自己小岛的狭小范围内比在外面的世界更快乐，因为他曾经在外面的世界过着“邪恶、受诅咒的、可憎的生活”。但他也立即向自己承认，如果有机会逃离，他还是会抓住的。

正如这些例子所显示的，经济学家可能会从更了解鲁滨逊的故事中获益。这种受益可以是双向的。对笛福的研究也可以（而且已经）受益于对经济学的更深入了解。笛福作品中有几个角落需要一定的经济学知识才能鉴赏。

当鲁滨逊踏上前往几内亚购买奴隶的命运多舛的航程时，他在巴西留下了一个正在生长的烟草种植园，很快就会价值“三四千英镑”。今天的读者很难理解这样一个数字意味着什么。曾任教于宾夕法尼亚州立大学的戴维·斯皮尔曼（David Spielman）根据经济史学家的研究，计算出这样一笔收入将使鲁滨逊跻身当时英国家庭收入的前5%。面对如此丰厚的财富，鲁滨逊没有理由冒险。他的航行就像他自己承认的那样是“荒谬的”。

经济学家还可能解开其他一些谜团。从岛上回来后，鲁滨逊收回了他的种植园并将其出售。在该书的前六版中，他收到了328,000个西班牙银洋，价值约72,000英镑。但在后来的版本中，一个零被删掉了。这对于故事的解释很重要。小说结尾时，鲁滨逊是个有钱人，还是个超级有钱的人？

文学学者为自己对文本每一处细微差别的敏感而自豪。但鲁滨逊的财富十去其九却无人在意。斯皮尔曼指出：“尽管《鲁滨逊漂流记》的文本历史受到了仔细的关注，但没有人注意到有这么个问题。”经济学家可能忽视了鲁滨逊其人的丰富性。但文学学者却忽视了他的大部分财富。 ■



Hollywood and AI

AI is stalking the last lions of Hollywood

The first actors to lose their jobs to artificial intelligence are four-legged

PADDING AROUND his spacious ranch in the hills outside Los Angeles, Luke has the swagger of an actor who has made it in Hollywood. After six successful years in the movies he lives in an open-plan home with mountain views, a large swimming pool and staff who bring him steak or smoothies when required. A driver is on hand to take him into the city when he gets a call to shoot an ad or make a paid appearance at an event, for which he commands a fee equal to a week's pay for a lowlier actor. Luke turns heads like few other artists. His agent's only complaint is that, given half a chance, he might try to eat his co-stars.

As a 500lb African lion, Luke (pictured) has to deal with challenges not experienced by other actors. But lately, he and his human colleagues have found common cause. America's screenwriters and actors spent half of 2023 on overlapping strikes. One of the reasons for their bust-up with the studios was the use of artificial intelligence (AI) in film and television production. Writers worry that AI will soon be churning out scripts, while actors fear that, as the technology improves the quality of computer-generated imagery (CGI), they may be replaced by digital stand-ins.

For most human actors that is still a hypothetical worry. But for four-legged screen stars it is already a reality. Work for animal actors is drying up as computer graphics offer an easier—and, some argue, more humane—way to portray creatures on screen. As AI makes those special effects still more realistic and easier to produce, the replacement of real animals with digital ones is accelerating, in television and advertising as well as in high-end movies. As demand for animal actors declines, their owners and trainers are leaving the business. Luke and his colleague, Maasai, who lies snoozing

in a nearby enclosure, are the last two working lions in Hollywood.

| *Rendered in tooth and claw*

Ever since a lion appeared on the title card of Goldwyn Studios, later part of MGM, animals have been an essential part of the Tinseltown talent pool. The star who saved Warner Bros from bankruptcy in the 1920s was Rin Tin Tin, a German shepherd who was in more than 20 pictures and was so bankable that producers dubbed him “the mortgage lifter”. From the 1950s an annual Oscars-style awards ceremony, first compered by an actor called Ronald Reagan, gave out gongs to stars such as Flipper the dolphin. In 1993 Entertainment Weekly ranked “the most powerful non-humans in entertainment”, including Moose, a dog-actor in “Frasier” (whose human co-stars were smeared with liver pâté to encourage him to nuzzle them).

Animals with the right skills can still find fame. Steve Martin, Luke the lion’s owner and trainer, has been in the business for more than 50 years, supplying everything from the deer in the “The Deer Hunter” to a pair of skunks named Alice and Cooper. One of the first big roles he landed for Luke was in “A Royal Romance”, a TV movie about Prince Harry and Meghan Markle. Luke played a wild lion in Botswana who, the film implied, may have been Princess Diana reincarnated. Another of Mr Martin’s lions, the late Major, had his paw prints set in Hollywood’s Walk of Fame, during an appearance alongside a nervous Sylvester Stallone to mark MGM’s 90th anniversary.

In an enclosure next to Luke and Maasai is Tag, an 11-year-old bear who lumbers over to greet Mr Martin with a lick. For a role in “Action Point”, a film released by Paramount Pictures in 2018, the 1,000lb bear was trained to drink from a beer can. “He never dented one, even,” says Mr Martin, proudly. Earlier this year Tag had to maul a dummy inside a sleeping bag, for an independent movie called “Night of the Grizzly” (Tag is a Kodiak, but he can play a grizzly). Training for that scene took two weeks, and a lot of

biscuits. Like any good film star, Tag has become somewhat fussy about his diet, shunning the raw salmon that a bear might eat in the wild in favour of cooked chicken from Walmart.

Such successful animal-acting careers are increasingly rare. Mr Martin once kept 100 or so animals at his 60-acre ranch, looked after by 15 trainers. Now his staff is down to a core of three, and most of the enclosures on the ranch are empty. Other businesses have closed or radically reduced their stock. Trainers are fewer, too. Moorpark College, on the edge of Los Angeles, has supplied animal wranglers to Hollywood through its Exotic Animal Training Management course since 1974. Gary Mui, a faculty instructor, estimates that when he graduated from the programme in 1995, about 70% of his classmates went into the entertainment industry. These days only about 10% of graduates do.

One reason for the industry's contraction is a growing public sensitivity to animal welfare, a hot topic in Hollywood since animals were first put in front of the camera. Early moviemakers treated animals as disposable props; Westerns used tripwires to make horses fall as if they had been shot. Public outcry greeted a cowboy film, "Jesse James", in 1939 in which a horse was deliberately driven off a cliff into a lake, where it drowned. The following year the American Humane Society, a charity, began providing its now-famous certification that "No animals were harmed" in the making of productions that passed its on-set inspections.

Animals' treatment in Hollywood has vastly improved since the days of "Jesse James". But abuse has not disappeared, and the risk of scandal makes studios nervous. HBO cancelled its racing drama, "Luck", after three horses died on set (poor ratings may have been another factor; one critic quipped that the animals may have died of boredom). "Gladiator 2", currently in production, has faced protests for using macaque monkeys and horses. These days many trainers are reluctant to talk about their work; one, who

says that animal-rights activists once called a police SWAT team to his home, insists (amid deafening squawking) that for security reasons he can no longer disclose whether he keeps animals on the premises.

Truly plausible digital stand-ins are now an alternative to real animals for wary studios. For many years, computer-generated creatures were convincing only in limited circumstances. “Jurassic Park” wowed audiences with its special effects in 1993, but used animatronic dinosaurs more than CGI ones. For years after that, computer models struggled to create realistic hair, which animals tend to have a lot of. Furry CGI creatures looked weirdly smooth—“like Gummi Bears”, recalls Mr Mui. Then, in 1998, came Disney’s “Mighty Joe Young”, whose star is a giant ape created from animatronics and CGI. Mr Mui watched it in dismay. “I said, ‘Oh, our days are numbered now,’” he recalls. For the first time, the hair looked good.

As the effects have got better, directors have relied on them more often. Disney has produced a series of live-action remakes of classic animal-based animations, including “The Jungle Book” (2016) and “Dumbo” (2019), in which it has cast real actors in the human roles but created the leading animals on a screen. In its blockbuster remake of “The Lion King” (2019) every creature, from warthog to wildebeest, was computer-generated. Disney is not alone. The star of Universal’s comedy-horror hit in February, “Cocaine Bear”, was a CGI creation. Tag never even got to audition. “They used to call us for stuff like that,” says Mr Martin. “Those guys really pushed us out of business.”

So far the “exotics” have lost the most work. Chimps have not acted in a big American movie since Paramount’s “The Wolf of Wall Street” in 2013. The last elephants are thought to have lumbered out of Hollywood around the same time. Studios are increasingly unwilling to work with big cats, says Mr Martin. “For wild animals, it’s basically finished,” he concludes.

Trainers of smaller stars are wondering how much further the digital trend will go. An hour or so from Luke the lion's ranch is Studio Animal Services, with an obstacle course in its front yard and multiple dog chews and fly swatters inside. Karin McElhatton, who has trained animals in Hollywood since the 1970s, is the company's owner—though Rumor, a white Pekin duck, appears to believe that he is its real proprietor, as he waddles in from a swim and preens in front of a fan, filling the room with a cloud of feathers and down.

"The business is getting more and more narrow in terms of the animals they want to use," says Ms McElhatton. She once flew a team of five squirrels, raised from babies, to Hungary for a Kinder chocolate ad, which required them to carry a coin to a vending machine. These days such jobs are nearly always computer-generated, she explains.

Ms McElhatton's 35 cats and 25 dogs still find work. Albert, a serene brown tabby on a pink lead, spends most of his time modelling for pet food commercials, though he also scored a recent role in a horror film pretending to devour a corpse. Tank, a German shepherd, is in demand as a sniffer dog on shows like "NCIS" and "911" (his trainer, David Meyers, has a Screen Actors Guild card so that he can play the accompanying police officer). Two golden retrievers, Parker and Porter, are often cast as family pets. Two years ago they went to Mexico for a comedy movie called "El Roomie"; they were flown first-class and got their own trailer.

Domestic animals have been somewhat protected against digital competition. Audiences are more familiar with the real thing and thus quicker to spot a fake dog than a fake bear. Animal-rights protesters seem somewhat less bothered by their involvement, too. Yet even man's best friend is now losing work to digital doppelgangers. "The Call of the Wild", a Disney movie released in 2020, called in computer animators to create

Buck, the St Bernard-Scotch shepherd who co-stars alongside Harrison Ford. Audiences didn't much like the digital dog; for now, directors still prefer to cast the real thing. Nonetheless, "it was extremely disconcerting for us," says Ms McElhatton, who was relieved when the movie fizzled at the box office.

Could CGI replace the two-legged animals at the top of the acting food chain? Near Seattle, an unusual team of animal trainers has seen signs that it might. The trainers recently taught a lion cub to climb onto its father's head, before rolling off his back and onto the ground. Yet the lions in question do not live in a cage, but on a hard drive, and their training is not done with biscuits, but with biomechanical elastic-body solvers and quasi-static integrators.

The digital lion-tamers work for Unity, which creates virtual environments and characters for video games and has dabbled in movies (including 2022's sequel to "Avatar"). Creating believable animals is still easier than creating realistic humans, says Allan Poore, who moved to Unity after working as an animator at Disney's Pixar. "We stare at faces and people all day...so we're going to notice."

Eyes are hard to get right, and skin can look waxy without elaborate light-diffusion techniques. But creating digital humans is getting easier. The addition of AI to the animator's toolkit has sped things up, meaning that "rigging" a digital model of a human face (think the wires on a puppet) might take an hour or two, down from a month. "We've been on this journey a long time, and I think we're getting closer and closer," Mr Poore says. "I've seen some stuff [where] it's hard to tell the difference. And you'll see more of that."

Actors are already being "de-aged" for roles which might once have gone to young lookalikes. In "Indiana Jones and the Dial of Destiny", released in

June, the octogenarian Mr Ford was rejuvenated by half a century by animators who used AI to mine footage of the actor in his “Raiders of the Lost Ark” days. Robert DeNiro underwent a similar transformation in Netflix’s “The Irishman” (2019), as did Samuel L. Jackson in Disney’s “Captain Marvel”, released in the same year. Background actors, or extras, fret that they may soon not be needed at all.

What kind of work awaits human stars then? The animals offer a preview of the coming attractions. Some still give performances that are later digitally enhanced. Disney’s live-action remake of “Lady and the Tramp” (2019) mixes footage of real dogs with CGI shots, including animating their mouths to simulate talking (an effect that trainers used to achieve by feeding the dog something chewy before the cameras rolled). Real dogs acted in Warner Bros’ “Birds of Prey” (2020), only to be morphed into a hyena in post-production. American Humane explains that a herd of 100 CGI horses is likely to be modelled on ten real ones.

Others have moved away from feature films to focus on more humdrum work. Social-media influencers, prizing high-impact, low-budget stunts, have become regular employers of exotic animals. Tag the bear has gamely “wrestled” with Logan Paul, a YouTuber, and taken on humans in a hot-dog eating contest organised by Barstool Sports, a blokeish entertainment site. Live events are also fairly safe from AI interlopers. Tag was hired in 2021 to appear at campaign rallies alongside John Cox, a California Republican who called himself “the beast”. Even this kind of work is becoming scarce, however, amid animal-welfare concerns. Mr Martin is being sued by an animal-rights group over Tag’s appearance at the rallies.

| *Circle of life*

Many of Hollywood’s last animal-handlers are now near retirement. “When this generation of trainers kicks the bucket, I don’t know that you’re going to have hardly anything in America,” says Ms McElhatton.

Animal-rights advocates are delighted. PETA, a lobby group, advocates more use of CGI creatures, arguing that “creative people can tell compelling, emotional stories using special effects and without exploiting any real animals.” “I understand some of the motivations of animal-rights people. And they need to understand that we have more in common with them than we don’t,” says Mr Mui sadly. “I want what’s best for animals, but I don’t have an issue with animals working with people as long as it’s done correctly,” he says. But now, “it’s all fading away.”

At Luke’s ranch, a for-sale sign has gone up at the end of the long driveway. Mr Martin, 76, is planning to move to Oregon with his remaining animals. He believes that their absence from the screen will ultimately harm the cause of animal welfare, by removing the audience’s connection to real, living creatures. “These guys are ambassadors for the ones in the wild,” he says. Soon, Luke and Maasai will head north with him in their air-conditioned trailer. And then the only lions left in Hollywood will be digital. ■



好莱坞和AI

人工智能盯上了好莱坞最后几头狮子

因AI失业的第一批演员是四条腿的家伙们【深度】

卢克（Luke）在洛杉矶郊外山上宽敞的牧场里转悠，一副好莱坞大腕的趾高气昂的派头。经过六年成功的电影生涯后，他住在一户开放式山景寓所，配有一个大游泳池，工作人员会在有需要时送上牛排或冰沙。当他接到拍摄广告或出席有偿活动的电话时，会有司机把他送进城，他收取的报酬相当于普通演员一周的工资。卢克所到之处，吸睛水平难有匹敌。他的经纪人唯一的抱怨是，一有机会，他可能就试图把他的合作演员吃了。

身为一头重达500磅的非洲狮，卢克（如图）必须应对其他演员没有经历过的挑战。但最近，他和他的人类同事们开始共进退。2023年的一半时间里，美国的编剧和演员的罢工此起彼伏。他们与制片厂争执的原因之一是人工智能（AI）在影视制作中的应用。编剧们担心AI很快就会开始大量创作剧本，演员们则害怕随着技术让计算机生成图像（CGI）的质量越来越好，他们可能会被数字替身取代。

这对于大多数人类演员来说仍然只是个假设性的担忧，但对于四条腿的银幕明星来说已经成为现实。动物演员的工作正趋枯竭，因为计算机图形技术为在银幕上塑造动物形象提供了更简便（而且一些人认为更为人性化）的方法。随着AI使这些特效更加逼真和易于制作，在电视、广告和高端电影中，数字动物正在加速取代真实的动物。对动物演员的需求减少后，它们的主人和驯兽师正纷纷离开这个行当。卢克和他躺在附近围栏里打盹的同事马赛（Maasai）是好莱坞最后两头还在工作的狮子。

渲染腥牙血爪

自一只狮子出现在戈德温影业（Goldwyn Studios，后来成为米高梅的一部分）的商标上起，动物就成为了好莱坞演艺人才库中不可或缺的一部分。20世纪20年代，拯救华纳兄弟免于破产的明星是任丁丁（Rin Tin Tin），他是一只德国牧羊犬，曾出演过20多部影片，票房非常出色，制

片人戏称它为“能还贷的角儿”。从1950年代起，一年一度的形同奥斯卡的颁奖礼会向海豚飞宝（Flipper）等明星颁奖，一位名叫罗纳德·里根的演员主持了第一届典礼。1993年，《娱乐周刊》发布“娱乐圈最具影响力的非人类”排行榜，上榜的包括《欢乐一家亲》（Frasier）中的狗子演员穆斯（Moose，与他合作的人类演员身上涂满了肝酱，以吸引他用鼻子蹭他们）。

拥有适当技能的动物仍有机会名声大噪。狮子卢克的主人兼驯兽师史蒂夫·马丁（Steve Martin）从事这一行已有50多年，供应过的动物林林总总，从《猎鹿人》（The Deer Hunter）中的鹿，到一对名为爱丽丝（Alice）和库珀（Cooper）的臭鼬等。他最早为卢克争取到的重要角色之一是在一部关于哈里王子和梅根·马克尔的电视电影《皇家情缘》（A Royal Romance）中。卢克在片中饰演了博茨瓦纳的一头野生狮子，影片暗示这头狮子可能是戴安娜王妃转世。马丁的另一头狮子，已故的“少校”（Major），在米高梅公司90周年庆典上与神色紧张的史泰龙一同亮相，爪印被刻在了好莱坞的星光大道上。

在卢克和马赛旁边的围栏里，11岁大的泰格（Tag）笨拙地移步过来，舔了舔马丁向他问好。为了在派拉蒙影业于2018年上映的电影《行动时刻》（Action Point）中扮演一个角色，这只重达1000磅的熊接受了从啤酒罐里喝水的训练。“他甚至没有弄瘪过一个罐头。”马丁自豪地说。今年早些时候，泰格不得不在一部名为《灰熊之夜》（Night of the Grizzly，泰格是棕熊，但它可以扮灰熊）的独立电影里袭击睡袋里的一个假人。这场戏的训练花费了两周时间和大量饼干。和其他优秀的电影明星一样，泰格变得对饮食有些挑剔，他不碰熊在野外可能会吃的生鲑鱼，而选择沃尔玛超市里的熟鸡肉。

如此成功的动物表演生涯日益罕见。马丁曾经在他占地60英亩的牧场里饲养了100多只动物，配有15名驯兽师。现在，他的员工只剩下核心三人团，牧场里的大部分围栏都空了。其他企业也纷纷关闭或大幅削减饲养的动物数量。驯兽师也越来越少。位于洛杉矶周边的莫尔帕克学院（Moorpark College）自1974年以来通过其“奇珍动物训练管理”课程为好

莱坞提供驯兽师。该学院教师加里·穆伊（Gary Mui）估计，当他1995年从该课程毕业时，约有70%的同学进入了娱乐业。如今，只有一成左右的毕业生进入该行业。

行业萎缩的原因之一是公众对动物福利日益敏感。自动物首次出现在镜头前以来，动物福利一直都是好莱坞的热门话题。早期的电影制作人将动物视为一次性道具；西部片使用绊索让马匹摔倒，制造出中弹倒地的效果。1939年，牛仔电影《杰西·詹姆斯》（Jesse James）故意将一匹马赶下悬崖，掉进湖里淹死，引发了公众强烈不满。次年，慈善机构美国人道协会（American Humane Society）开始提供到今天大家都已很熟悉的“没有动物受到伤害”认证，证明影片的制作过程通过了其现场检查。

自《杰西·詹姆斯》时代以来，好莱坞对待动物的方式有了很大改善。但虐待动物的现象并未消失，引发丑闻的风险也让制片厂感到紧张。在三匹马死在片场后，HBO中途砍掉了其赛马题材剧集《幸运》（Luck）（收视率不佳可能是另一个因素；一位剧评人打趣说它们可能是被无聊的剧情闷死的）。目前正在制作中的《角斗士2》因使用猕猴和马而遭到抗议。如今许多驯兽师都不大愿意提起自己的工作。一位驯兽师说，动物权利活动人士曾经叫了一支特警队到他家。他（在一片震耳欲聋的动物叫声中）坚持说，出于安全考虑，他不能再透露他是否在这里饲养了动物。

对于谨慎的制片厂而言，真正可信的数字替身如今已成为真实动物的一种替代选项。多年来，电脑制作的动物只有在某些情况下看着令人信服。1993年，《侏罗纪公园》的特效令观众叹为观止，但它使用的更多是仿真恐龙而不是CGI恐龙。在那之后的多年里，计算机模型一直难以制作出逼真的毛发，而动物往往有很多毛。穆伊回忆说，CGI制作出来的毛茸动物看起来很奇怪，滑溜溜的“就像小熊软糖”。1998年，迪士尼的《无敌大猩猩》（Mighty Joe Young）上映，该片的主角是一只用动画和CGI制作的巨猿。穆伊看得垂头丧气。“当时我说，‘哦，这下我们的日子可不多了’。”他回忆道。这是动物的毛发头一回看起来对劲了。

随着效果越做越好，导演们也越来越依赖它们。迪士尼已经制作了一系列

经典动物动画的真人翻拍版，包括《奇幻森林》（The Jungle Book，2016年）和《小飞象》（Dumbo，2019年），其中人类角色由真人演员出演，动物主角则是在显示屏上制作出来的。在翻拍大片《狮子王》（The Lion King，2019年）中，从疣猪到角马的每一种动物都是电脑制作的。这样做的不只是迪士尼。环球影业去年2月推出的卖座喜剧恐怖片《可卡因熊》（Cocaine Bear）的主角就是由CGI制作的。泰格甚至连试镜的机会都没有。“过去有这种角色他们会给我们打电话，”马丁说，“那些家伙真把我们逼得没事干了。”

到目前为止，“奇珍品种”丢掉的工作最多。自2013年派拉蒙出品的《华尔街之狼》（The Wolf of Wall Street）之后，黑猩猩就再没出演过美国大片。据称最后一批大象也是在同一时期沉重出走好莱坞的。马丁说，制片厂越来越不愿意与大型猫科动物合作。“对于野生动物来说，基本上已经玩完了。”他总结道。

｜ 人类最佳化身

小型动物明星的驯兽员们想知道数字化趋势会走多远。距狮子卢克的牧场约一小时车程的制片厂动物服务公司（Studio Animal Services）的前院设有障碍训练场，里头有多种狗咬胶和苍蝇拍。公司所有者卡琳·麦克尔哈顿（Karin McElhatton）自20世纪70年代以来一直在好莱坞训练动物。不过，白色的北京鸭“流言”（Rumor）似乎相信自己才是这里真正的主人，他游完泳摇摇摆摆地进屋，在一台风扇前搔首弄姿，扬起一阵羽毛和绒毛。

“娱乐业想要使用的动物品类越来越窄。”麦克尔哈顿说。她曾经带着五只从小养大的松鼠飞往匈牙利拍摄一支健达巧克力广告，广告要求它们把一枚硬币拿到自动售货机上。如今这类工作几乎都是由计算机完成了，她解释说。

麦克尔哈顿的35只猫和25只狗仍然有活干。阿尔伯特（Albert）是一只安详的棕色虎斑猫，套着粉红色的牵绳，大部分时间都在为宠物食品广告做模特，不过最近还参演了一部恐怖电影，假装吞食一具尸体。坦克

（Tank）是一只德国牧羊犬，在《海军罪案调查处》（NCIS）和《911》等节目中扮演嗅探犬（他的训练师大卫·迈耶斯 [David Meyers] 拥有美国演员工会会员证，因此可以扮演随行警官）。两只金毛猎犬帕克（Parker）和波特（Porter）常扮演家庭宠物。两年前，他们去墨西哥参演喜剧电影《室友》（El Roomie），坐头等舱，还有自己的移动化妆拖车。

家畜在某种程度上免受数字竞争的影响。观众对家畜的真身更熟悉，因此假狗比假熊更容易看出来。动物权利抗议者对它们的参与也更宽容些。但即使是人类最好的朋友现在也在被数字分身取代。2020年上映的迪士尼电影《野性的呼唤》（The Call of the Wild）请电脑动画师制作了与哈里森·福特共同出演的圣伯纳-苏格兰牧羊犬巴克（Buck）。观众们不大喜欢这只数码狗。目前而言，导演们仍然更倾向于用动物真身。不过，“我们那次可是紧张死了，”麦克尔哈顿说。直到这部片子票房惨淡，她才松了一口气。

CGI能否取代处于表演食物链顶端的两足动物？在西雅图附近，一支不同寻常的驯兽师团队已经看到了这种可能性的迹象。最近，驯兽师教一只幼狮爬到它的父亲头上，然后从父亲的背上滚到地上。不过，这些狮子并不生活在笼子里，而是生活在硬盘上，它们的训练也不是用饼干完成的，而是用生物力学弹性体解算器和准静态积分器。

这些数字驯狮者为Unity工作，该公司为视频游戏创建虚拟环境和角色，还涉足电影（包括2022年的《阿凡达》续集）。在迪士尼的皮克斯担任动画师后加入Unity的艾伦·普尔（Allan Poore）表示，创造可信的动物仍然比创造逼真的人类更容易。“毕竟我们整天都盯着面孔和人看……我们会看出来的。”

眼睛就很难做好，而如果没有精致复杂的光漫射技术，皮肤看起来会像是蜡像。但创造数字人类正变得越来越容易。在AI加入动画师的工具包后，进展已经提速，“操纵”一个人脸数字模型（想想提线木偶）现在可能只需要一两个小时，而以前要一个月。“我们已经在这条路上走了很久，我认

为离目标越来越近了，”普尔说，“我已经见过一些真假难辨的东西。而你会看到更多这种东西。”

老演员们已经在被“减龄”，去扮演那些本来可能得找长得像他们的年轻人来演的角色。在6月上映的《夺宝奇兵：命运之盘》（Indiana Jones and the Dial of Destiny）中，动画师利用AI挖掘哈里森·福特在《夺宝奇兵》第一部时期的镜头，让已经八旬高龄的福特变回三十多岁。罗伯特·德尼罗在奈飞的《爱尔兰人》（The Irishman，2019年）中实现了类似的容颜转变，塞缪尔·杰克逊在同年上映的迪士尼的《惊奇队长》（Captain Marvel）中也是如此。背景演员（群演）担心自己可能很快就完全不被需要了。

那么，什么样的工作等待着人类明星呢？这些动物预告了即将到来的精彩节目。它们有些仍在表演，但后来经过了数字增强。迪士尼的真人版《小姐与流浪汉》（Lady and the Tramp，2019）将真狗的镜头与CGI镜头混剪，包括让狗的嘴巴模拟出讲话的效果（过去，训练师会在镜头启动前给狗喂一些有嚼劲的东西来达到这种效果）。真狗出演了华纳兄弟的《猛禽小队》（Birds of Prey，2020年），但在后期制作中变成了鬃狗。美国人道协会解释说，一群100匹的CGI马很可能是以10匹真马为原型制作出来的。

其他动物已经从参演电影长片转向更日常平淡的工作。社交媒体博主们推崇高冲击力、低预算的特技表演，已经成为了奇珍动物的常见雇主。棕熊泰格曾与YouTube博主洛根·保罗（Logan Paul）英勇“搏斗”，在男性体育娱乐网站Barstool Sports组织的吃热狗比赛中挑战人类。现场活动也不会受到AI的威胁。泰格在2021年受聘与自称“野兽”的加州共和党人约翰·考克斯（John Cox）一起出席竞选集会。但由于对动物福利的担忧，连这样的工作也变得稀缺。马丁因泰格参加了这次集会被动物权利组织告上了法庭。

| 生生不息

好莱坞最后一批驯兽师中的许多人如今已近退休。麦克尔哈顿说：“等这

一代驯兽师去世了，我不知道美国还会有什么动物演员。”

动物权益倡导者很高兴。游说团体PETA主张更多使用CGI制作的动物，认为“文艺创作者可以利用特效来讲述引人入胜的情感故事，而无需利用任何动物真身”。“我理解动物权益人士的某些动机。但他们要明白，我们和他们的共识要比分歧多。”穆伊伤感地说。“我希望为动物谋福利，但我并不反对动物与人合作，只要方式正确。”但现在，“一切都快没了。”

在卢克的牧场，长长的车道尽头竖起了一块写着待售的牌子。76岁的马丁计划带着他剩下的动物搬去俄勒冈州。他认为，它们从银幕上消失最终会损害动物福利事业，因为观众与真实动物的连结断了。“这些家伙是野生动物的形象大使。”他说。很快，卢克和马赛将乘坐带空调的拖车和他一起北上。然后，好莱坞就只剩下数字狮子了。■



Robotics

Delivery robots will transform Christmas

Santa's hi-tech little helpers

A SHOP ASSISTANT leaves a Co-op convenience store in Milton Keynes and opens the lid of a white box, about the size of a small suitcase, with a red flag on top and six wheels. After the assistant drops a bag of shopping inside and scans a bar code, the box trundles off. Travelling at a brisk walking pace along the footpath, it pauses at a road junction until two cars have passed before crossing safely. Neither pedestrians nor car drivers give it a second glance. Delivery robots like this have become part of the scenery since they started work in this town, some 80km north-west of London, in 2018.

“That’s when you know a new technology is successful,” says Ed Lovelock. “People don’t notice it any more.” Mr Lovelock is product manager for Starship Technologies, a Californian firm that has so far delivered more than 5m shopping orders and restaurant meals in Europe and America using its autonomous Starships.

In some places such deliveries arrive by air. “It soon becomes a normal part of your life,” says Keller Rinaudo Clifton, the boss of Zipline, a drone-delivery firm also based in California. Zipline began using drones to deliver blood and medical products in Rwanda in 2016. It is expanding into groceries and meals and now operates in other parts of Africa as well as America and Japan. In 2024 Zipline will begin deliveries to hospitals and clinics in the north of England for Britain’s National Health Service.

| *Where’s my bot?*

Like many new technologies, delivery bots have gone through something of a hype cycle. A decade ago many predicted they would soon be everywhere.

Amazon, for one, announced with great fanfare ambitious plans for its Prime Air drone-delivery service in 2013, but progress was slow and not much happened. That is changing, and even Amazon's drones finally look like they are about to take off.

A number of things have brought this about. Companies like Starship and Zipline began modestly in areas where regulators were more comfortable with robotic deliveries. Milton Keynes, for instance, is a new town with wide paths and cycleways for bots to drive along, and with few low-flying aircraft to worry about in the sky above Rwanda, drones can operate safely.

Having steadily gained solid operating experience, officialdom is becoming more relaxed about such services. What that means is, particularly at this time of year, instead of fleets of delivery vans with drivers hauling seasonal gifts and shopping to people's front doors, an increasing number of goods are arriving by robot.

Starship has gone on to launch services in ten British towns, including older places with narrower streets like Manchester, Leeds and Cambridge. It also makes deliveries on more than 50 university campuses in America. Nor is it alone. Serve Robotics, which is backed by Uber, a ride-sharing platform, began using "sidewalk robots" for restaurant deliveries in Los Angeles in 2022 and aims to deploy some 2,000 in other American cities.

Customers typically use an app to order, with the firms adding a small delivery fee. Depending on distance, this starts at 99 pence (\$1.20) in Milton Keynes. Around 100 Starships, each able to carry up to 10kg, serve a number of stores. They navigate along pre-mapped routes using satellite positioning. Sensors, including a dozen cameras and radar, create a "bubble of awareness" around the robot. On arrival, customers use their phone to unlock the robot's storage compartment and collect their shopping.

It helps that Starships have been made cutesy. Customers in Milton Keynes can choose a song, like “Happy Birthday” for the bot to play on arrival. Some are also decorated for festive occasions, such as “pumpkinbots” during Halloween or “reindeerbots” at Christmas. As a result, adds Mr Lovelock, residents are protective of them and few are tampered with. The bots emit a high-pitched screech if anyone tries to steal them or their contents.

The growing acceptance of delivery bots is helping the idea spread. A trial scheme using sidewalk robots in Helsinki, the Finnish capital, is due to be expanded in the spring of 2024. In November the South Korean government warned people “not to be surprised” by more of them appearing in the capital, Seoul, where the 7-Eleven convenience chain has been testing four-wheel models produced by Neubility, a startup backed by Samsung.

Much the same is happening with drones, although they are not yet dressing up and singing songs. America’s Federal Aviation Administration recently allowed some firms, including Zipline, to fly “beyond-visual-line-of-sight” (BVLOS). In America and many other countries drones are not allowed to be flown out of sight of their operators unless ground observers monitor them in case other aircraft are in the vicinity. This restricted how far drones could fly and drove up costs.

Advances in technology helped win these new freedoms. For BVLOS flights, new miniaturised sensing devices can be built into drones to detect and avoid other aircraft. Zipline’s system uses specially developed microphones which can pick up the sound of an approaching aircraft and determine its position, allowing the drone to take evasive action if necessary.

| *Drone highways*

In Britain a 165-mile (265km) superhighway for drones, connecting southern England with the Midlands, will start operating in 2024. It will

rely on a series of ground stations along the route to communicate with the drones to keep them apart and avoid any other aircraft.

The ability to fly BVLOS allows Zipline to offer similar services to its African operations. For these, the company uses a fixed-wing drone capable of a round trip of some 200km. Carrying up to 1.8kg, it is launched with a giant catapult and drops its delivery using a parachute.

The company is starting to work with a number of medical centres in America. In a recent deal with the Cleveland Clinic, it will deliver medicines directly to people's homes in locations throughout Ohio. For this Zipline will use a new type of hovering drone, able to make round trips of some 30km carrying up to 3.6kg. Instead of a parachute, this drone uses a load-carrying device called a "droid". After being lowered on a cable, the droid employs a small fan motor to manoeuvre, allowing it to set packages down in precise locations, such as the front steps of a home. This drone-plus-droid system will also be used to deliver groceries and meals.

Both sidewalk robots and drones still require some level of human supervision. Usually this involves people in a control centre monitoring them and intervening if necessary. For Starships, these interventions tend only to come when a bot stops and seeks confirmation that a manoeuvre it intends to undertake, like crossing a tricky road junction, is safe. Zipline's drones can be called back to base in the event of a problem, or ordered to stop flying immediately and deploy a parachute to land in an emergency. Such events, however, are "extraordinarily rare", says Mr Rinaudo Clifton.

As for Amazon, it began a limited drone-delivery service in two small areas of California and Texas in 2022. It has since developed a new drone, called the MK30, which it plans to put into service in America, Britain and Italy by the end of 2024. These will operate out of the company's delivery centres and also use a sense-and-avoid system for BVLOS flights. It is quieter than

the firm's existing model, can carry packages up to 2.2kg and will be able to fly in unsettled weather, including light rain.

Amazon is talking about delivering millions of packages by drone every year by the end of the decade. If the giant of online retailing can finally crack the technology, then automated delivery could spread almost everywhere. If not, there are already enough firms demonstrating that, at least in some areas, delivery bots using wheels, wings or rotors are coming your way. ■



机器人技术

送货机器人带来别样圣诞节

圣诞老人的高科技小助手

一名店员从米尔顿凯恩斯（Milton Keynes）的一家Co-op便利店走出来，打开一个白色箱子的盖子。这个箱子的大小和一个小行李箱差不多，上面插着一面小红旗，底下装着六个轮子。店员把一袋商品放进箱子、扫描完条形码，箱子就缓缓滑走了。它以比路人散步略快的速度沿着人行道滑行，在一个路口停下来，等两辆车经过后，才安全穿过路口。无论是过往行人还是司机都不会多看它一眼。自2018年以来，这样的送货机器人就开始穿梭在这个位于伦敦西北向约80公里处的小镇，成为镇上的一道风景。

“当人们对它熟视无睹的时候，”埃德·洛夫洛克（Ed Lovelock）表示，“你就知道一项新技术取得了成功。”洛夫洛克是加州公司Starship Technologies的产品经理。截至目前，该公司已经使用其自主移动机器人Starship为欧美的商店和餐馆完成了500多万单外送服务。

在有些地方，这些配送是从空中送达的。“它很快就会成为你日常生活的一部分。”同样位于加州的无人机送货公司Zipline的老板凯勒·里纳乌多·克利夫顿（Keller Rinaudo Cliffton）表示。2016年，Zipline开始在卢旺达使用无人机运送血液和医疗用品。它正在向食品杂货和餐饮领域扩张，现在除了在美国和日本，它也在非洲其他地区开展业务。2024年，Zipline将为英国国家医疗服务体系（NHS）提供服务，开始向英格兰北部的医院和诊所送货。

| 我的机器人在哪里？

和许多新技术一样，送货机器人也经历了某种程度上的技术成熟度曲线。十年前，许多人预测它们很快就会无处不在。例如，亚马逊在2013年大张旗鼓地宣布了其雄心勃勃的Prime Air无人机送货服务，但该计划后来进展缓慢，并未取得多少成就。现在情况正在改变，亚马逊的无人机似乎也终于要一飞冲天了。

一些因素促成了这些改变。Starship和Zipline等公司起步都较为谨慎，选择了监管机构对机器人送货较为包容的地区。例如，米尔顿凯恩斯是新城镇，这里有宽阔步道和自行车道可供机器人行驶；在卢旺达的上空几乎没有低空飞行的飞机可担心，无人机因而可以安全飞行。

在这类公司稳步积累了扎实的运营经验后，官员们开始放宽对它们的服务的限制。这意味着越来越多的货物正由机器人送达，而不是由货车司机把节日礼物和其他商品送到居民家门口，尤其是在每年的圣诞季。

Starship后续在英国的十个城镇推出了服务，包括曼彻斯特、利兹和剑桥等一些街道狭窄的较古老城镇。它还在美国50多所大学的校园里送货。Starship并非个例。2022年，由拼车平台优步（Uber）支持的Serve Robotics公司开始在洛杉矶使用“人行道送货机器人”为餐馆送外卖，并计划在美国其他城市部署约2000台。

顾客一般会通过应用下单，送货公司会加收少许送货费。在米尔顿凯恩斯，送货费根据距离远近而定，起步价为99便士（1.20美元）。米尔顿凯恩斯大约有100台Starship，每台最多可携带十公斤物品，为好几家商店送货。它们利用卫星定位沿着预先规划好的路线行进。包括十二个摄像头和雷达在内的传感器在机器人周围形成一个“感知气泡”。到达目的地后，顾客用手机解锁机器人的储物箱，取走自己购买的物品。

Starship萌萌的外型设计也是一个加分项。米尔顿凯恩斯的顾客可以选择一首歌曲（比如《生日快乐》），让机器人在到达时播放。逢年过节时，一些机器人还会被装饰成不同造型，比如万圣节的“南瓜机器人”或圣诞节的“驯鹿机器人”。因此，洛夫洛克补充道，居民很护着它们，很少有机器人被蓄意破坏。如果有人试图偷走它们或者它们运送的物品，这些机器人会发出刺耳的尖叫声。

人们对送货机器人不断提高的接受度有助于让这个创意传播开来。2024年春季，芬兰首都赫尔辛基将扩大一项使用人行道机器人的试验计划的范围。在韩国首都首尔，7-11连锁便利店已经在测试由三星支持的创业公司

Neubility生产的四轮电动送货机器人。对于首尔街头出现的越来越多的送货机器人，韩国政府在去年11月提醒民众“无需惊讶”。

送货无人机的情况大致相同，尽管它们还没有被装扮起来或是被安排唱歌。美国联邦航空管理局（FAA）不久前允许包括Zipline在内的一些公司进行“超视距”飞行。根据美国和许多其他国家的规定，无人机不准飞出操控员的视线，除非地面观察员能够监视它们，以防它们碰到附近的其他飞行器。这一规定限制了无人机的飞行距离，推高了成本。

技术上的进步帮助无人机争取到了新的自由度。执行超视距飞行的无人机可以内置新的微型传感设备以探测和避开其他飞行器。Zipline的系统使用了专门研发的麦克风，可以捕捉到其他飞行器靠近时发出的声音并确定其位置，从而让无人机在必要时采取规避行动。

| 无人机高速通道

在英国，一条连接英格兰南部和中部地区的165英里（265公里）长的无人机高速通道将于2024年开始运营。它将依靠沿途的一系列地面站与无人机建立通信，让无人机相互保持距离，同时避开其他所有飞行器。

因为具备了超视距飞行的能力，Zipline能够为其在非洲的业务提供类似的服务。Zipline使用的是一架单次能够往返飞行约200公里的固定翼无人机。它最大载重1.8公斤，通过一个巨大的弹射器起飞，并使用降落伞投放物品。

Zipline开始与美国的一些医疗中心合作。根据不久前与克利夫兰诊所（Cleveland Clinic）达成的协议，Zipline将把药品直接送到俄亥俄州各地的居民家中。为此，Zipline将使用一种新型悬停式无人机，单次能够往返飞行约30公里，最大载重3.6公斤。这款无人机使用的是一种叫作“机器人”的载货装置，而不是降落伞。“机器人”被缆绳放下后，会启动一个小型风扇马达来让自己移动，从而将包裹精准放在指定位置，比如房门口的台阶上。这种“无人机+机器人”的系统也将被用来运送食品杂货和餐食。

无论是人行道机器人还是无人机，都还需要一定程度的人类监督。通常是

由控制中心里的人员监控，并在必要时干预。对于Starship来说，一般只有当机器人停下来并试图确认自己打算采取的行动（比如穿过一个复杂的十字路口）是否安全时，才会进行干预。Zipline的无人机在遇到问题时可以被召回基地，或者在紧急情况下按指令立即停止飞行并使用降落伞着陆。不过，这种情况“极为罕见”，里纳乌多·克利夫顿表示。

至于亚马逊，它于2022年在加州和得克萨斯州的两小块区域开始了有限的无人机送货服务。此后，亚马逊开发了一种名为MK30的新型无人机，计划在2024年底前在美国、英国和意大利投入使用。这些无人机将从亚马逊的配送中心起飞，并且同样使用了“感知回避”系统来执行超视距飞行。它的噪音比亚马逊的现有机型小，最多可携带2.2公斤的包裹，并能在小雨等不稳定的天气条件下飞行。

亚马逊正在谈论在2030年前实现每年用无人机运送数百万个包裹。如果这家在线零售巨头最终能够攻克这项技术，那么自动化配送可能会遍布几乎世界各地。就算做不到这一点，也已经有足够多的公司证明，至少在某些地方，使用轮子、机翼或螺旋桨的各种送货机器人正在向我们走来。■



Xi-3PO

China's quest to become a robot superpower

As its population shrinks, China hopes machines can pick up the slack

CHINA'S FIRST attempt at building a humanoid robot did not hit the mark. The machine produced in 2000 by a team at the National University of Defence Technology looked like a walking toaster. It had googly eyes and cannon-like protuberances near its crotch. Called Xianxingzhe, or Forerunner, it was mocked in neighbouring Japan, which at the time boasted far sleeker robots. Japanese netizens described it as China's secret weapon—designed to make its enemies die of laughter.

China has stuck with it, though. In November the government published a plan calling for the mass production of humanoids by 2025. The country's love of robots goes beyond those that can walk and talk. Last year half of all the industrial robots installed worldwide were fitted in China, according to the International Federation of Robotics, an industry body. It is now the fifth most automated country in the world when measured by robots per worker. Motivated by pride and pressing demographic challenges, China is on a mission to become a robot superpower.

Many of the country's newly installed robots are mechanical arms that can be programmed to weld, drill or assemble components on a production line. But last year China also produced over 6m "service robots", which help humans with tasks apart from industrial automation. Such machines scoot around warehouses, moving boxes. Others clean hotels. At a restaurant in the southern city of Guangzhou meals are cooked and served by robots.

Some of this may seem gimmicky, but to the Communist Party led by Xi Jinping robots are serious business. Officials believe China fell behind and was humiliated by Western powers in the 19th century in part because it did

not embrace technological revolutions happening elsewhere. Now China aims to stay ahead of the game. Whereas officials once used steel production as a gauge of economic advancement, today they look at the number of robots installed, says Dan Wang of Hang Seng Bank.

China's impressive economic growth in recent decades was a result of three main factors: a soaring urban workforce, a big increase in the capital stock and rising productivity. Today, though, less new infrastructure is needed. And the working-age population, those between 15 and 64, is shrinking. It is projected to drop by over 20% by 2050. Earlier this year the government released a list of 100 occupations for which there is a shortage of labour. Manufacturing-related positions accounted for 41 of them. A surfeit of young and cheap workers once did these jobs; now wages are higher and workers less abundant.

As a result, Mr Xi has made boosting China's productivity a priority. The government sees robots playing a big part in this effort. For years it has pushed industry to go from being labour-intensive to robot-intensive. Provinces have spent billions of dollars helping manufacturers upgrade in this way. China's experience during the pandemic reinforced this mindset. Endless lockdowns caused factories to close and Western firms to reconsider their supply chains. When all of the controls were lifted in 2022, a wave of covid-19 again disrupted businesses as workers fell ill. With robots, health is not a concern.

Many of the challenges faced by factories apply to agriculture, too. The average Chinese farmer is in his or her 50s. Few young people want to take their place in the fields. Countries that face similar predicaments often import either their food or cheap labour. But China is paranoid about food security and uninterested in immigration. Robots could be the answer. Some aspects of agriculture, such as milking cows, can be automated fairly easily. Others are trickier, but appear possible on a small scale. The south-

western city of Chengdu has developed an unmanned vegetable farm which could, in theory, produce ten harvests a year.

In time, robots might replace ageing workers. They might also play a role in caring for them. China has far too few professionals looking after its 8.1m care-home residents. A plan from the National Health Commission, published in 2021, called for developing smart elderly care. Some of it is aspirational, such as providing frail people with electronic exoskeletons to aid their movements. But simpler robots could be used to help old folk bathe or stand up. China's tech giants are looking into the challenge. In 2022 iFlytek, a big artificial-intelligence firm, said it wanted to send robots into the homes of seniors to offer companionship and health management. Residents of a care home in Shanghai are kept happy by a robot that zips around singing revolutionary songs from their youth, according to local media.

What would make the government happy is if China's robotics industry became more self-sufficient. Local firms still rely on foreign companies for parts and know-how. China is fearful of being shut out of Western markets, for good reason. America has blocked Chinese firms from buying advanced semiconductors and the equipment used to make them (robots require chips, but usually not the most advanced kind). So the government has been trying to stimulate robotics research. In August the city of Beijing announced a 10bn-yuan (\$1.4bn) fund for robot development.

Such efforts are having some effect. Last year 36% of the industrial robots China installed were made at home, up from 25% in 2013. Shenzhen Inovance Technology, a big Chinese firm, builds robots that are used to make LED lights and mobile phones. It may be able to source all of the components it needs from Chinese companies within five years, says Zhu Xingming, its chairman.

For most Chinese robotics firms, though, self-sufficiency is still some way off. That is part of the reason why the government is pushing the development of humanoids. These may not be very practical or affordable in the near term. But officials hope the process of manufacturing them will create a domestic supply chain.

One thing the government does not have to worry about is much pushback against its plans. Surveys suggest most Chinese people think robots will create more jobs than they destroy. China, it seems, is a land of techno-optimists. It helps, of course, that independent labour unions are banned.





Xi-3PO

中国追求成为机器人超级大国

面对人口萎缩，中国希望机器人能填补空缺

中国打造人形机器人的第一次尝试并不理想。2000年，国防科技大学的一个团队开发了一台机器，看起来就像是会走路的烤面包机。它有圆溜溜的眼睛，胯部附近有炮筒一样的突起。这台名为“先行者”的机器人被当时机器人技术远远领先的邻国日本大加嘲笑。日本网民说它是中国的秘密武器——能让敌人笑死。

但中国坚持了下来。2023年11月，中国政府公布了一项计划，提出要在2025年实现大规模生产人形机器人。中国对机器人的热爱远不限于那些会走路说话的。据行业机构国际机器人联合会（International Federation of Robotics）的数据，2022年中国的工业机器人装机量占到了全球的50%。按平均每名工人配备的机器人数量计算，中国目前是全球自动化程度第五高的国家。受荣誉感的驱动，加上紧迫的人口挑战，中国正在努力成为机器人超级大国。

中国新安装的机器人有许多是经设置程序后可在生产线上焊接、钻孔或组装部件的机械臂。不过2022年中国还制造了600多万台“服务机器人”，辅助人类完成工业自动化以外的任务。这些机器人在仓库内穿行，搬运货箱。还有一些在酒店里打扫卫生。在南方城市广州的一家餐厅里，烹饪和传菜都由机器人完成。

这里面有些可能看起来只是营销噱头，但对习近平领导的共产党来说，机器人是件很严肃的事情。官员们认为，中国在19世纪落后于西方列强并遭受它们欺辱，一定程度上是因为没有接纳发生在其他地方的技术革命。现在，中国要力求在技术上保持领先。在过去，官员们用钢铁产量来衡量经济发展水平，如今他们看重的是机器人装机量，恒生银行的王丹指出。

中国近几十年来令人瞩目的经济增长归功于三大因素：城市劳动力激增，

资本存量大幅增加，以及生产率不断提升。但到了今天，对新基础设施的需求减少了。而15至64岁的劳动年龄人口日渐萎缩，预计到2050年将减少超过20%。2023年初，政府公布了一份清单，列出了“最缺工”的100个职业，其中与制造业相关的就有41个。曾经有大量年轻的廉价劳动力从事这些职业，而现在工资更高了，工人却不再充足。

因此，习已经把提高中国的生产率列为要务。政府认为机器人能在这方面发挥重要作用。多年来，政府一直在推动工业从劳动密集型向机器人密集型转型。多省投入数十亿美元计的资金帮助制造企业做这种升级。新冠疫情期间的经历更是强化了这一思路。无休止的封控导致工厂关闭，让西方企业重新考虑供应链部署。2022年所有管制解除后，一波疫情让员工纷纷病倒，再度打断了公司运作。有了机器人，就不会受制于员工的身体状态了。

工厂面临的许多挑战也是农业面对的难题。中国农民的平均年龄是50多岁。很少有年轻人愿意接手务农。陷于类似困境的国家通常会进口粮食或者引进廉价劳动力。但中国执着于保障粮食安全，也无意引进劳工。机器人可能是个出路。农业的某些工作要实现自动化并不难，比如挤奶，其他工作的自动化虽然比较棘手，但似乎也可以小规模实现。西南部城市成都开发了一个无人蔬菜农场，理论上一年可以收获十茬蔬菜。

假以时日，机器人也许能取代日渐年迈的人力，还可能在照顾他们方面发挥作用。中国入住养老院的老人达810万，而照护他们的专业人员远远还不够。国家卫健委于2021年发布的一项计划呼吁发展智慧养老服务。其中一些设想雄心勃勃，比如为体弱者提供电子外骨骼帮助其行动，但同时也可使用更简单的机器人帮助老人洗澡或站立。中国的科技巨头正迎向这一挑战。2022年，大型人工智能公司科大讯飞表示，希望把机器人送入老人家中，提供陪伴和健康管理服务。据上海媒体报道，在当地一家养老院，一台机器人会一边转悠一边唱老人们年轻时传唱的革命歌曲，把他们逗得很开心。

会让政府开心的是中国的机器人产业能变得更自给自足。本地企业仍依赖

外国公司提供零部件和专门技术。中国有充分理由担心被西方市场拒之门外。美国已禁止中国企业购买先进半导体和用于制造这些半导体的设备（机器人需要芯片，但通常不是最先进的芯片）。因此，中国政府一直努力推动机器人研究。8月，北京市宣布成立一项100亿元的基金用于机器人研发。

这些努力正初见成效。2022年中国安装的工业机器人中有36%是国产的，高于2013年的25%。大型企业深圳汇川技术制造的机器人用于生产LED灯和手机。公司董事长朱兴明表示，他们可能在五年内就能从中国企业采购到所需的全部零部件。

不过对于大多数的中国机器人公司而言，要达到自给自足仍有一段距离。这也是中国政府大力发展人形机器人的原因之一。短期内，这些机器人可能不太实用，价格也太高。但官员们希望在开发这些机器人的过程中能形成一条国内供应链。

政府不必担心其计划会遭到很多反对。调查显示，大多数中国人认为机器人创造的就业机会将多于它们会破坏掉的。由此看来，中国是一个技术乐观主义者的国度。当然了，这也得益于这里禁止成立独立工会。■



The economics of technology

A short history of tractors in English

What the tractor and the horse tell you about generative AI

IT WAS THE ChatGPT of its day. “Come and see the tractors”, entreated an article in the *Prairie Farmer* in 1915, advertising a trade show in Illinois showing off the new tech. “It will mark a new epoch in farming—the farmer’s liberation from sole dependence on the weary horse.” “Tractors are more economical than horses,” insisted an agricultural expert in a government report around the same time, “not only making farm work cheaper but easier.” The tech clearly impressed people, but it also scared them. One American observer, watching a tractor in England, said it “walked over the earth like some huge animal, puffing and snorting”. Tractors promised a revolution in American agriculture, an industry which in 1900 employed about a third of workers and produced about 15% of GDP.

Today many people expect another revolution, linked to developments in generative artificial intelligence (AI). Like then, the general public today view the technology with a mixture of awe and fear. Goldman Sachs, a bank, reckons generative AI could raise annual global GDP by 7% over ten years. Some economists now talk about “explosive growth”. Others say that before long, jobs will be eliminated in their millions. Yet the economic history of the tractor casts doubt over these predictions. Over the sweep of history the tractor has indeed had an immense impact on people’s lives. But it conquered the world with a whimper, not a bang.

Historians disagree about who invented the tractor. Some say it was Richard Trevithick, a British engineer, in 1812. Others credit John Froelich, working in South Dakota in the early 1890s. Still more point out that the word “tractor” was little used until the start of the 20th century; and that only then did people start seriously talking about the average farmer buying

one. At the time horses and mules pulled around an impressive array of farm implements, from ploughs to reapers.

The horses faced up to a bigger, more powerful beast. You can put yourself in the shoes of an early 20th-century horse by visiting Gene Jones's tractor museum in Millbrook, a small city in rural Alabama, which contains dozens of machines, including Farmalls and Fords, from the 20th century. The tractors are beautiful—and not just because Mr Jones has lovingly restored them, painting them in a variety of rich auburn-reds. They are also intimidating. Some weigh thousands of pounds. Others have cranks to start the engine which can break your arm if you don't know what you're doing.

With hindsight, it is clear that the tractor had profound impacts. It meant that a given quantity of farmland could feed more people. Tractor-owning farmers no longer needed to pasture horses, each of which required about three acres of cropland for feed each year. More intensive farming also had downsides. Some researchers have argued that tractors helped bring about the Dust Bowl of the 1930s. Their powerful ploughing techniques damaged the topsoil that had once prevented wind erosion.

The economic impact eventually became clear, too. The greater efficiency afforded by tractors allowed farmers to expand their operations, as they could manage more land with the same number of workers. Farms began to grow in size, with smaller family farms giving way to larger, mechanised operations. According to one estimate, by 1960 the average American farm was 58 acres (equivalent to the land occupied by ten big Walmarts today) larger than it would have been without tractors. The tractor also reduced the number of workers needed to produce food by about 2m, or 25% of farm employment in 1960. All these improvements added up. In a paper published in 2012 Richard Steckel and William White, two economists, argue that by the mid-1950s farm mechanisation had raised American GDP by about 8%.

And yet for much of the first half of the 20th century, tractor-induced changes did not feel very profound. This is because the tractor diffused across the American economy slower than one of Mr Jones's old Fords trying to cross a waterlogged field. In 1920, despite rave reviews in the *Prairie Farmer*, just 4% of American farms had a tractor (see chart). Even by 1940 only 23% had them. In the 1910s opportunistic businessmen had piled into the tractor-making business, hoping to make a quick buck (just as every second tech firm in Silicon Valley now describes itself as "AI-first"). Many had no customers and were forced to close.

The horse endured for a surprisingly long time. For much of the 1930s the total productive capacity of equine animals—quite literally, horsepower—across American farms still exceeded that of tractors. In 1945 a quarter of farms reported both draft animals and tractors. The slow diffusion of the tractor produced slow productivity gains. The data are spotty, but in the first half of the 20th century annual productivity growth in agriculture probably never exceeded 3%. That 8% GDP effect is real, but it made itself felt only over decades. Explosive growth? Hardly.

The tractor's plodding progress is one of the big puzzles of economic history. If they were so good, why did farmers not buy them more quickly? They were not Luddites who resisted new technologies on principle. True, anti-tractor lobby groups, such as the Horse Association of America, warned that buying one would land the farmer in unmanageable debt. But in the 1910s and 1920s many tractorless farmers did own cars, suggesting that they were willing to try new tech. In 1917 *Power Farming*, a journal, published letters from 15 farmers who used tractors. They were probably solicited, but these letters urged others to follow suit.

Three reasons explain why the triumph of the tractor took so long. First, early versions of the technology were less useful than people had originally believed, and needed to be improved. Second, adoption required changes

in labour markets, which took time. And third, farms needed to transform themselves.

Take capabilities first. The early tractors of the 1900s were behemoths. They were useful for ploughing, and a few other things, but not for cultivating fields of growing crops. Many early models had metal wheels, not tyres, so they got stuck in the mud. They were also costly. Between 1910 and 1940, however, the machines became both more versatile and smaller, making them suited to a wider range of tasks.

In 1927 John Deere released a power lift for its models. This meant that a farmer could pull a lever to raise an implement (such as a plough), rather than doing it manually. Rubber tyres came along in about 1933. For a long time the general-purpose tractor could not mechanise corn and cotton harvests, one reason why the area in which Mr Jones lives was one of the slowest to adopt tractors. But by the 1920s America had the corn-picker, followed by the mechanical cotton-picker after the second world war. By the end of the fighting, tractor prices had also fallen from their level in 1910, after adjusting for inflation, by about half.

Wages were the second factor. Horse technology was labour-intensive: horses require feeding, cleaning and medical care, even when they are not working. In the early 1930s, during the Depression, average real wages in agriculture fell. So for many farmers it became easier to hire someone to manage a horse—you could always fire them—than it was to splurge on a tractor. But by the second world war, labour shortages mounted, leading real wages to rise quickly. Suddenly, machines seemed like a better deal.

The third factor was corporate restructuring. Tractors worked best on big farms, where the farmer could spread out the expense of a huge upfront investment. As a result, enlarging the size of their holdings and buying a tractor were two sides of the same coin. In a survey in Illinois in 1916, for

instance, farmers who used tractors profitably also talked about increasing their acreage. But growing a farm takes time. A farmer looking to expand had to gather the necessary capital, and then negotiate the purchase with the owners.

The history of the tractor hints at how quickly generative AI may take over. At present most AI models still have metal wheels, not rubber tyres: they are insufficiently fast, powerful or reliable to be used in commercial settings. Over the past two years real wages have hardly grown as inflation has jumped, limiting companies' incentives to find alternatives to labour. And companies have not yet embraced the full-scale reorganisation of their businesses, and in-house data, necessary to make the most of AI models. No matter how good a new technology may be, society needs a long, long time to adjust. ■



技术经济学

英语拖拉机简史

从拖拉机和马的对峙看生成式人工智能来袭

它就是当年的ChatGPT。“来看看拖拉机吧！”1915年，《草原农场主》（Prairie Farmer）杂志上的一篇文章恳请道，它在为伊利诺伊州展示这项新技术的展销会做广告。“这将标志着农业的新纪元，让农场主从只能依靠疲倦的马匹中解放。”“拖拉机比马更经济，”一位农业专家在大约同一时间发布的一份政府报告中坚称，“它不仅让农业劳作的成本更低，也让农活更轻松。”这项技术显然让人们惊叹不已，但也让他们害怕。一位美国观察家在英国观摩拖拉机作业时说，它“喷着响鼻，像一头巨兽般在大地上横行”。拖拉机预示着美国农业的一场革命，该行业在1900年雇用了约三分之一的劳动力，创造了约15%的GDP。

如今，许多人预期另一场革命到来，这次是与生成式AI的发展相关。和当年一样，今天的公众对这项技术又敬又怕。高盛认为，生成式AI可能在十年内将全球年GDP提高7%。一些经济学家正大谈“爆炸性增长”。其他人则表示，用不了多久，千百万工作岗位将被淘汰。然而拖拉机的经济史让人对这些预测心生怀疑。纵观历史长河，拖拉机确实对人们的生活产生了巨大的影响。但它是悄无声息地征服了世界，而非轰轰烈烈地席卷而来。

历史学家对于谁发明了拖拉机存在分歧。有人说是英国工程师理查德·特里维希克（Richard Trevithick）在1812年发明的。另一些人认为是1890年代初在南达科他州工作的约翰·弗罗利希（John Froelich）。还有更多人指出，在20世纪初之前很少能看到“拖拉机”这个词。要等到进入20世纪后人们才开始认真讨论普通农场主购买拖拉机的事宜。那会儿，从犁到收割机等各种令人眼花缭乱的农具都是由马和骡子来拉动的。

马儿们遭遇了一种更高大威猛的巨兽。参观一下吉恩·琼斯（Gene Jones）位于阿拉巴马州乡村小城米尔布鲁克（Millbrook）的拖拉机博物馆，你就能体会到20世纪早期马儿们的感受了。这家博物馆内收藏了数十

台拖拉机，其中包括20世纪的Farmall系列和福特（Ford）拖拉机。它们都很漂亮，这一部分要归功于琼斯对它们的精心修复，还给它们刷上了深浅各异的赤褐色。它们也很吓人，有些重达数千磅，还有一些需要用摇动曲柄来启动发动机——要是你不知道怎么操作，可能会弄折了自己的胳膊。

事后看来，拖拉机显然产生了深远的影响。有了它，同样面积的农田就可以养活更多的人。有了拖拉机，农场主不再需要牧马，而饲养每匹马每年需要大约三英亩的耕地来种植草料。更集约化的农业也有其弊端。一些研究人员认为，拖拉机的应用是20世纪30年代沙尘碗（Dust Bowl）灾害事件发生的原因之一。它们强大的犁耕技术破坏了以往防止风蚀的表层土壤。

拖拉机的经济影响最终也变得清晰起来。拖拉机效率更高，让农场主能够用相同数量的雇工管理更多的土地，从而可以扩大经营。农场规模开始扩大，较小的家庭农场被规模更大的机械化农场取代。据一项估计，到1960年，美国农场的平均面积要比假如没有拖拉机的情况下大58英亩（相当于今天十家大型沃尔玛的占地）。拖拉机还让生产食物所需的工人减少了约200万，相当于1960年农场就业人数的25%。所有这些进步累积起来产生了巨大影响。在2012年发表的一篇论文中，经济学家理查德·斯特克尔（Richard Steckel）和威廉·怀特（William White）认为，到20世纪50年代中期，农业机械化让美国GDP提高了约8%。

然而，在20世纪上半叶的大部分时间里，拖拉机引发的变化并没有让人感觉非常深刻。这是因为拖拉机在美国经济中的普及速度比琼斯的老旧福特拖拉机穿越积水田地的速度还要慢。1920年，尽管《草原农场主》大加好评，仍只有4%的美国农场有拖拉机（见图表）。即使到了1940年也只有23%的农场有。在1910年代，投机商人纷纷涌入拖拉机制造业，希望能赚笔快钱（就像在如今的硅谷，每两家科技公司就有一家自称“以AI为先”）。许多厂家无人光顾，只能关门收场。

马儿坚持的时间之久出人意料。在20世纪30年代的大部分时间里，美国农

场上马类役畜的总生产力（可谓真正的“马力”）仍然大于拖拉机。1945年，四分之一的农场同时拥有役畜和拖拉机。拖拉机普及缓慢，导致生产率的提高同样缓慢。虽然缺乏完整数据，但可以知道在20世纪上半叶，农业生产力的年增长率可能从未高过3%。推动GDP提高8%的作用确实有，但要历经好几十年才感受得到。至于说推动爆发式增长？几乎没有。

拖拉机缓慢的普及速度是经济史上的一大谜题。如果拖拉机这么好，为什么农场主不赶紧买呢？他们并不是从原则上抵制新技术的卢德派。确实，美国马协会（Horse Association of America）等反拖拉机游说团体曾警告说，购买一辆拖拉机将使农场主陷入难以负担的债务。但在1910年代和1920年代，许多没有拖拉机的农场主却拥有汽车，表明他们是愿意尝试新技术的。1917年，《机械农业》（Power Farming）期刊发表了15位使用拖拉机的农场主的来信。虽说这些来信可能是应邀写的，但起到了敦促他人效仿的作用。

三个原因解释了为什么拖拉机的胜利花了这么长时间。首先，早期的拖拉机并没有人们最初以为的那么有用，它需要改进。其次，采用拖拉机需要改变劳动力市场，而这需要时间。第三，农场本身也需要改造。

先说说拖拉机的性能。1900年代的早期拖拉机都是些庞然大物。它们可用于开垦荒地和其他一些作业，但不适用于耕作已种植作物的田地。许多早期的机型都使用金属轮子，而不是轮胎，所以会陷在泥里。而且它们价格高昂。然而在1910年至1940年间，拖拉机变得功能更多样，体积也更小，能适用于更广泛的任务。

1927年，约翰·迪尔（John Deere）为其拖拉机设计了一个动力升降装置。这样农场主就可以拉动操作杆来拉起农具（例如犁），而无需手动操作。橡胶轮胎在1933年前后出现在拖拉机上。在很长一段时间里，通用拖拉机无法实现玉米和棉花收割的机械化，这便是琼斯居住的地区是采用拖拉机最慢的地区之一的一个原因。但到了20年代，美国有了玉米采摘机，第二次世界大战后又出现了机械采棉机。到战争结束时，拖拉机经通胀调整后的价格也比1910年的水平下降了约一半。

工资是第二个因素。用马从事农业生产要耗费很多劳动——马需要喂养、清洁和医护，即便是在它们不工作时。1930年代初经济大萧条期间，农业平均实际工资下降。因此，对于许多农场主来说，雇人管理马匹比花大钱买拖拉机更容易，雇工还可以随时解雇。但到了第二次世界大战时，劳动力变得日渐短缺，导致实际工资迅速上涨。突然间，用机器似乎更划算了。

第三个因素是农场重组。拖拉机在大型农场中效果最好，因为农场主可以摊薄巨额前期投资。因此，扩大农场规模和购买拖拉机变成了事情的一体两面。例如，在1916年伊利诺伊州的一项调查中，使用拖拉机获利的农场主也谈到了要增加农场面积。但扩大农场需要时间。想要扩大规模的农场主必须筹集所需的资金，接着还要与土地所有者协商购买事宜。

拖拉机的历史透露出生成式AI可能会以怎样的速度占据主导。目前，大多数AI模型采用的仍旧是“金属轮子”，而不是“橡胶轮胎”，它们的速度、动力或可靠性都还不足以应用于商业环境中。过去两年，随着通胀飙升，实际工资几乎没有增长，限制了企业寻找劳动力替代选择的积极性。而且企业尚未开启全面重组业务和内部数据，而这是充分利用AI模型所必需的。无论一项新技术可能有多么好，社会都需要很长很长的时间来调整适应。





Economic history

Is the age of Milton Friedman over?

Some may say so. But we are still living in it

Milton Friedman. By Jennifer Burns. Farrar, Straus and Giroux; 592 pages; \$35

IT IS VOGUISH to declare the ideas of Milton Friedman dead, whether you think they deserve damnation or eulogy. In America, prominent Democrats spit out his name contemptuously. The most influential American economist of the 20th century is routinely disparaged as a heartless fetishist of Ayn Randian capitalism, who evangelised corporate greed at home and authoritarianism abroad. Friedman is a special bugbear of President Joe Biden. While running for office in 2020, he declared that “Milton Friedman isn’t running the show anymore.”

Meanwhile, the current, populist standard-bearers of American conservatism—the political movement with which Friedman was identified throughout his life—agree, having turned their backs on fiscal discipline and open markets.

It might be tempting to buy into the thesis that Friedman’s ideas are passé and even to wish them good riddance. But that would be a mistake. Few thinkers are as important (and as grotesquely caricatured) as Friedman. His critiques of Keynesianism, his advocacy of the importance of central banks, his emphasis on the primacy of the monetary supply in explaining inflation and his prioritisation of real interest rates over nominal ones were once unorthodox. They are now mainstream.

A recent biography by Jennifer Burns, a professor at Stanford University, dispels fallacies. Because of Friedman’s long life (he was born in 1912 and

died in 2006) and prolific career involving the modern era's most important economic debates, Ms Burns's book functions as an intellectual guide to the entire 20th century, benefiting from nearly a decade of archival research.

The Friedman who emerges here is one of astonishing economic brilliance, establishing monetary policy as a field worthy of serious study. It is a portrait of a surprisingly heterodox economist, who was perhaps the last great political economist and thought deeply about the connection between political and economic freedom.

Friedman is best known for revolutionary ideas on the importance of money to the macroeconomy, which is now seen as blindingly obvious but once was not. With Anna Schwartz—his greatest intellectual partner other than his wife, Rose Director Friedman—he wrote “A Monetary History of the United States” (1963), which computed aggregate monetary supply to argue that the Federal Reserve had contributed to the Great Depression.

Ben Bernanke, the man who would lead the Fed through the global financial crisis of 2007-09, gave cheeky tribute to this argument in a speech at a gathering celebrating Friedman's 90th birthday: “Regarding the Great Depression, you're right, we did it. We're very sorry. But thanks to you, we won't do it again.”

By the 1980s Friedman was catapulted to global renown, claiming a Nobel prize, a column in Newsweek and even a popular TV series. He was a perpetual gadfly at the Fed. Were he alive today, he would probably jab at Jerome Powell for thinking that dramatic growth in the monetary supply would not manifest in higher inflation. (During his life, he sported a vanity licence plate with the formula of his quantity theory of money, $MV=PY$; the equals sign was drawn with the help of black tape, which resulted in several traffic tickets.)

He became a bogeyman of the left for a six-day trip to Chile to advise Augusto Pinochet, its dictator, though Ms Burns argues that while he “failed to appreciate the optics”, that “in truth, he played almost no role in policy design”. Other trips, to China and to countries east of the Iron Curtain, did not provoke so much conspiracism, outrage or damage to his reputation.

Although Ms Burns admires Friedman, her book is not a hagiography. She argues that his concept of freedom, the nominal core of his political philosophy, could be “woefully thin”. Friedman was vocally opposed to the Civil Rights Act of 1964, which outlawed racial discrimination, and “never revisited his position on civil rights”, she writes with disappointment.

Much as the three Abrahamic religions lay claim to one saviour, conservatives, libertarians and classical liberals all claim Friedman. But he defies easy categorisation. Unlike some libertarians, Friedman accepted the legitimacy of the state (though he fought against it and advocated eliminating some government departments). Unlike many conservatives, he deemed redistribution acceptable to alleviate poverty. Indeed, Friedman envisioned the school-voucher and health-care programmes still in effect in America, as well as the tax policies that top up working class wages—perhaps the country’s most important anti-poverty programme.

Despite consulting for Barry Goldwater, Ronald Reagan and Margaret Thatcher, Friedman considered himself a classical liberal. “Those of us who believe in liberalism...have a new faith to offer; it behoves us to make it clear to one and all what the faith is”, he wrote in 1951. He was critical of overly doctrinaire, laissez-faire philosophy that “assigned almost no role to the state other than the maintenance of order and the enforcement of contracts”.

Ms Burns insists on dubbing Friedman “the last conservative” because “the

synthesis Friedman represented—based in free-market economics, individual liberty and global co-operation—has cracked apart” in politics. Friedman may no longer be running the show, but he is still one of economics’ most influential acts. ■



经济史

弗里德曼的时代终结了吗？

也许有人这么认为。但我们仍生活在他的时代【《弗里德曼传》书评】

《弗里德曼传》，詹妮弗·伯恩斯著。Farrar, Straus and Giroux出版社，592页；35英镑。

无论你认为对米尔顿·弗里德曼（Milton Friedman）的思想是应该诅咒还是讴歌，宣称弗里德曼思想已死都很时髦。在美国，民主党要员提起他的名字时都语带轻蔑。这位20世纪最具影响力的美国经济学家经常被贬斥为安·兰德式资本主义的冷血信徒，在国内宣扬企业贪婪，在国外宣扬威权主义。弗里德曼是美国总统拜登的心头之患。在2020年竞选总统时，他宣称“米尔顿·弗里德曼已经不再是主角了”。

与此同时，美国保守主义（弗里德曼终其一生被视为这一政治运动的一员）当前的民粹主义旗手们也持相同的观点，他们已经背弃了财政纪律和开放市场。

也许人们会很容易听信弗里德曼的思想已经过时的观点，甚至庆幸终于从中解脱了。但这会是个错误。很少有思想家像弗里德曼一样重要（也很少像他那样被如此荒诞地歪曲丑化）。他批判凯恩斯主义、倡导央行的重要性、强调货币供应在解释通胀中的首要作用，以及优先考虑实际利率而非名义利率，这些观点一度都是非正统的。现在，它们已经成了主流。

斯坦福大学教授詹妮弗·伯恩斯（Jennifer Burns）最近出版的一本传记驳斥了种种谬误。由于弗里德曼的长寿（生于1912年，卒于2006年），以及学术生涯成果丰硕且参与了现代最重要的经济辩论，伯恩斯花了近十年研究档案文献写就的这本书可以充当整个20世纪的知识指南。

书中呈现的弗里德曼具有惊人的经济学才华，是他将货币政策确立为一个值得认真研究的领域。这本书描绘了一位有惊人非正统思想的经济学家，

也许是最后一位伟大的政治经济学家，对政治自由和经济自由之间的关联有深刻的思考。

弗里德曼最为人熟知的是关于货币对宏观经济的重要性的开创性思想，这一点在当今显而易见，但在当时却不然。他和除了妻子罗斯·戴瑞克特·弗里德曼（Rose Director Friedman）之外最重要的智力伙伴安娜·施瓦茨（Anna Schwartz）共同撰写了《美国货币史》（A Monetary History of the United States，1963年出版），通过计算货币供应总量来证明美联储对大萧条负有责任。

在庆祝弗里德曼90岁生日的聚会上，后来领导美联储渡过2007至2009年全球金融危机的本·伯南克（Ben Bernanke）在发言时顽皮地致敬了他的观点：“说到大萧条，你说得对，是我们干的。我们非常抱歉。但多亏了你，下次不会再犯了。”

到了1980年代，弗里德曼在全球名声鹊起，获得了诺贝尔奖，为《新闻周刊》撰写专栏，甚至还制作了一部热门电视系列片。他永远都在给美联储挑刺。如果他今天仍然在世，恐怕也会抨击鲍威尔认为大幅增加货币供应不会导致通胀上升的想法。（他一生都骄傲地挂着一张自选车牌，车牌号就是他的货币数量论公式 $MV=PY$ ，等号是用黑胶带贴出来的，还因此吃了几张交通罚单。）

他曾访问智利六天，为独裁者奥古斯托·皮诺切特（Augusto Pinochet）出谋划策，因而被左派视为妖魔。不过伯恩斯认为，尽管他“没有意识到这么做给外界的观感”，但“事实上，他在政策设计中几乎没有发挥任何作用”。他的其他出访，例如前往中国和铁幕以东的其他国家，并没有引起如此多的阴谋论、愤怒或对他声誉的贬损。

虽然伯恩斯钦佩弗里德曼，但她的书并不是一本歌功颂德之作。她认为，作为他政治哲学名义上的核心，弗里德曼的自由观可能“薄弱得可怜”。他曾公开反对1964年禁止种族歧视的《民权法案》，而且“从未重新审视过自己在民权问题上的立场”，她失望地写道。

正如三大亚伯拉罕宗教都宣称救世主在自己这边一样，保守主义者、自由意志主义者和古典自由主义者都将弗里德曼归入自己的阵营。但要给他归类并不容易。与一些自由意志主义者不同，弗里德曼接受国家的合法性（尽管他也与国家抗争并主张取消一些政府部门）。与许多保守主义者不同，他认为可以接受通过再分配来减轻贫困。事实上，弗里德曼设想的教育券和医疗保健计划仍然在美国实施，还有补贴工人阶级工资的税收政策——这也许是美国最重要的脱贫措施。

尽管曾为巴里·戈德华特（Barry Goldwater）、里根和撒切尔做过顾问，弗里德曼仍将自己视为一名古典自由主义者。他在1951年写道：“我们这些信奉自由主义的人.....有一种新的信仰要奉上；我们有责任向每个人阐明这种信仰是什么。”他批评过于教条的自由放任主义哲学，认为它“除了维持秩序和执行契约之外，几乎没有赋予国家任何作用”。

伯恩斯坚持将弗里德曼称为“最后的保守派”，因为“弗里德曼所代表的以自由市场经济、个人自由和全球合作为基础的综合理念在政治上已经分崩离析”。弗里德曼也许不再是主角，但他仍然是经济学舞台上最有影响力的角色之一。■



A dismal year for the dismal science

Economists had a dreadful 2023

Mistaken recession calls were just part of it

SPARE A THOUGHT for economists. Last Christmas they were an unusually pessimistic lot: the growth they expected in America over the next calendar year was the fourth-lowest in 55 years of fourth-quarter surveys. Many expected recession; The Economist added to the prognostications of doom and gloom. This year economists must swap figgy pudding for humble pie, because America has probably grown by an above-trend 3%—about the same as in boomy 2005. Adding to the impression of befuddlement, most analysts were caught out on December 13th by a doveish turn by the Federal Reserve, which sent them scrambling to rewrite their outlooks for the new year.

It is not just forecasters who have had a bad year. Economists who deal in sober empirical work have also had their conclusions challenged. Consider research on inequality. Perhaps the most famous economic studies of the past 20 years have been those by Thomas Piketty and his co-authors, who have found a rising gap between rich and poor. But in November a paper finding that after taxes and transfers American incomes are barely less equal than in the 1960s was accepted for publication by one of the discipline's top journals. Now Mr Piketty's faction is on the defensive, accusing its critics of "inequality denial".

Economists have long agreed that America would be richer if it allowed more homes to be built around popular cities. There is lots of evidence to that effect. But the best-known estimate of the costs of restricting construction has been called into question. Chang-Tai Hsieh of the University of Chicago and Enrico Moretti of the University of California, Berkeley, found that easing building rules in New York, San Francisco and

San Jose would have boosted American GDP in 2009 by 3.7%. Now Brian Greaney of the University of Washington claims that after correcting for mistakes the true estimated effect is just 0.02%. If builders disagreed as wildly about roof measurements, the house would collapse.

Think social mobility in America is lower than it was in the freewheeling 19th century, when young men could go West? Think again, according to research by Zachary Ward of Baylor University. He has updated estimates of intergenerational mobility between 1850 and 1940 to account for the fact that past studies tended to look only at white people, as well as correcting other measurement errors. It now looks as if there is more equality of opportunity today than in the past (albeit only because the past was worse than was thought).

A rise in suicides, overdoses and liver disease has reduced life expectancy for white Americans. Angus Deaton and Anne Case of Princeton University popularised the idea that these are “deaths of despair”, rooted in grimmer life prospects for those without college degrees. But economists have been losing faith in the idea that overdoses, which are probably the biggest killer of Americans aged 18-49, have much to do with changes in the labour market. New research has instead blamed the carnage on simple proximity to smuggled fentanyl, a powerful opioid.

Other findings are also looking shaky. The long decline in the prestige of the once-faddish field of behavioural economics, which studies irrationality, continued in 2023. In June Harvard Business School said it believed, after an investigation, that some of the results in four papers co-written by Francesca Gino, a behavioural scientist and PhD economist, were “invalid”, owing to “alterations of the data”. (Ms Gino, who has written a book about why it pays to break rules, is suing for defamation the university and the bloggers who exposed the alleged fiddling.)

What lessons should be drawn from economists' tumultuous year? One is that for all their intellectual discipline they are still human. Replicating existing studies and checking them for errors is crucial work.

Another lesson is that disdain for economic theory in favour of the supposed realism of empirical studies may have gone too far. After the global financial crisis of 2007-09, commentators heaped opprobrium on theorists' common assumption that people make rational predictions about the world; gibes about an unrealistic, utility-maximising Homo economicus helped raise the status of behavioural economics. Yet rational-expectations models allow for the possibility that inflation can fall rapidly without a recession—exactly the scenario that caught out forecasters in 2023.

A last lesson is that economists should cheer up. The research that has been called into question this year inspired much pessimism about the state of modern capitalism. But a dodged recession, flatter inequality trends and less despair would all be good news. Perhaps the dismal science should be a little less so. ■



【首文】郁闷科学的郁闷一年

经济学家经历了一个糟糕的2023年

对衰退的错误预测只是糟心事之一

体谅下经济学家的处境吧。2022年圣诞节时，这个群体异常悲观：他们对2023年美国增长的预计是55年来第四季度调查中的第四低。许多人预期会出现衰退；本刊也奉上了一些惨淡预测。到了2023年的圣诞节，经济学家们只能把圣诞布丁换成一盘尴尬的苦果，因为2023年美国的增长可能达到了3%，高于长期平均水平，大约与2005年的繁荣时期相当。再加上大多数分析师在12月13日被美联储的鸽派转向打了个措手不及，纷纷匆忙地改写自己对新一年的展望，更让人觉得这群人已经茫然无措。

不仅仅是做预测的人经历了糟糕的一年。从事严谨实证工作的经济学家们的结论也受到了挑战。看看关于不平等的研究。过去20年里最著名的经济学研究或许是由托马斯·皮凯蒂及其合著者所做的研究，他们发现富人与穷人之间的差距在扩大。但去年11月，一篇论文发现，经过税收和转移支付之后，美国的收入差距几乎与1960年代相当，该论文被一家顶级经济学期刊接受待发表。现在皮凯蒂一派摆出防御姿态，指责其批评者“否认不平等”。

经济学家长期以来一致认为，如果美国允许在热门城市的周围建造更多住房，美国会更加富裕。有很多证据支持这一点。但关于限制建筑的代价的最知名估计受到了质疑。芝加哥大学的谢长泰和加州大学伯克利分校的恩里科·莫雷蒂（Enrico Moretti）曾认为，放宽纽约、旧金山和圣何塞的建筑限制将会使2009年的美国GDP增长3.7%。现在，华盛顿大学的布赖恩·格里尼（Brian Greaney）声称，在纠正错误后，真实的影响估计仅为0.02%。如果建筑商们在屋顶测量上有这么大的分歧，那房子就得塌了。

你觉得美国的社会流动性低于自由奔放的19世纪（那时年轻人可以到西部去）？那你得再想想了——贝勒大学（Baylor University）的扎卡里·沃德（Zachary Ward）指出。他的研究更新了对1850年到1940年间代际流动性

的估计，考虑了过去的研究往往只关注白人这一问题，同时纠正了其他测量错误。现在看起来，如今的机会平等性要好于过去（尽管这只是因为过去比人们原本以为的更糟糕）。

自杀、药物过量和肝病的增加降低了美国白人的预期寿命。普林斯顿大学的安格斯·迪顿（Angus Deaton）和安妮·凯斯（Anne Case）在过去普及了一种观点：这些都属于“绝望之死”，其根源在于那些没有大学学位的人面对更加晦暗的生活前景。但经济学家开始不大相信药物过量——可能是18至49岁美国人的第一大杀手——与劳动力市场的变化有很大关系。新的研究将这类大量死亡归咎于仅仅是很容易到手的走私芬太尼，一种强效阿片类药物。

其他发现如今看起来也不大站得住脚了。研究非理性行为的行为经济学曾经风靡一时，之后声望持续下降，在2023年延续了这种趋势。哈佛商学院在6月表示，经调查认为，在行为科学家、经济学博士弗朗西斯卡·吉诺（Francesca Gino）与他人共同撰写的四篇论文中，部分结果由于“篡改数据”而“无效”。（曾著书解释为何违反规则有好处的吉诺正在控告哈佛以及曝光她涉嫌造假的博文作者诽谤。）

从经济学家动荡纷乱的一年里，应该吸取什么教训？有一点是，尽管经济学家富有智识自律，但他们终究是人。重复已有研究并检查其中的错误是至关重要的工作。

另一个教训是，蔑视经济理论、偏爱实证研究所谓的求真务实可能走过了头。在2007至2009年的全球金融危机之后，评论员对理论家认为人们会对世界做出理性预测的普遍假设大加指责；对不现实的、追求效用最大化的“理性经济人”的嘲弄帮助提升了行为经济学的地位。然而在理性预期模型中，通胀是有可能迅速下降却不引发衰退的——正是这种情景在2023年打了预测者的脸。

最后一个教训是经济学家应该振作起来。过去这一年受到质疑的这些研究激发了对现代资本主义状态的许多悲观情绪。而躲过了衰退、不平等趋势

走平，以及人们没那么绝望都应该是好消息。也许这门郁闷的科学应该变得稍微不那么郁闷些了。■



Free exchange

Where does the modern state come from?

Economists attempt to answer a profound political question

IT IS PART metaphor, part myth and part history. Thomas Hobbes thought life there was nasty, brutish and short. John Locke disagreed, proclaiming that it was where people first learnt how to own things. Jean-Jacques Rousseau described it as the place where people were born free, before they became ensnared in chains. Robert Nozick thought that people were so desperate to escape it, there was an inevitable result: the creation of a state.

Ideas about the “state of nature”—how people lived before politics organised itself into governments—have held the attention of philosophers for centuries. Discovering whether it played out as imagined was nigh-on impossible. And yet thinking about what people would do without a government helped answer profound questions. What are the limits of political power? Is the modern state something that citizens would freely choose?

Now, after all this theorising, three economists think they have some empirical answers. According to Robert Allen of New York University, Abu Dhabi, Leander Heldring of Northwestern University and Mattia Bertazzini of the University of Nottingham, the key to understanding the emergence of modern politics is not a metaphor, but the constantly shifting courses of ancient rivers in Iraq. The first states, they argue in a paper published in the *American Economic Review*, were glued together not as shelters from violence, as Hobbes believed, but by economics.

The banks of the Tigris and the Euphrates, Iraq’s two longest rivers, are home to some of the world’s oldest settlements. Mesopotamia, which 5,000 years ago refined the first known system of writing, earned the area the

reputation of “the cradle of civilisation”. The paths of these rivers shift, as floods and droughts cause their beds to flood. When a shift came, some ancient farmers were left without water for their crops.

Mr Allen and his co-authors investigate whether the timing of changes to a river’s course had anything to do with when the number and size of settlements grew. They do so by looking at the effect of the first recorded shift in 2,850BC. This presented farmers with something close to the choice imagined by philosophers when theorising about the state of nature. Those left behind by the river could revert to nomadism. Or they could band together to build irrigation systems to ferry water from distant rivers.

A philosophical question is therefore transformed into something akin to a laboratory experiment, only one set thousands of years ago and extending hundreds of miles across. Moreover, the results of the experiment are clear. A 5km-by-5km square in the basin left behind by a river was 14% more likely to have a settlement, marked by a public building such as a temple or marketplace, 150 years after the shift than in the 50 years before it. Each square was 12% more likely to have a built canal, a form of artificial irrigation that made farming far from rivers possible. Five new cities were created, and only three abandoned. Esnunna, one city along a new tributary of the river, became much bigger.

This, Mr Allen and his co-authors say, is evidence that that the first states were formed by farmers co-operating for economic reasons. A canal network would have been too large a cost for any to bear alone. But by spreading the cost, the construction was worth it for each. Such decisions were momentous. They represent some of the earliest examples of governments providing infrastructure in return for taxes, and thus the genesis of the earliest states.

The authors then divide centuries of thinking on the origins of states into

two camps. The first, which they say ranges from Daron Acemoglu, an influential economist at the Massachusetts Institute of Technology, to Karl Marx, supposes that states ultimately emerge from a process of social bargaining. The rich and high-status seize power for personal gain, and periodically dole out services, such as a road, school or police force, in order to keep populations on board. But if this had been the case in Mesopotamia then it would have been in the areas that a river shifted towards that settlements would have formed. After all, they developed richer and more fertile farmland, yielding a bigger tax take.

That Mesopotamian farmers seem to have chosen to band together as the river shifted away lends support to the second camp. Philosophers in this group, who include Locke and Rousseau, contend that governments emerged when people chose to co-ordinate themselves, swapping their freedom to do whatever they wanted for a state that mediates disputes and provides a degree of safety. Mr Allen and his co-authors analyse only Mesopotamian Iraq, but they argue that their results ought to apply more generally to other fledgling states. Governments, in other words, are chosen rather than foisted upon their citizens.

| *Meandering path*

This is quite the landgrab by economists, seizing terrain that is more commonly occupied by political theorists. The study is not flawless. Perhaps an unknown conquest explains the spread of settlements in the period under consideration. Maybe the authors are wrong and the pattern does not hold elsewhere. There were already six cities and many more settlements in the Mesopotamian Valley before its rivers really began to move, and some had existed for a thousand years. The authors insist that they are only interested in how new governments form, but there is a chance they have in fact captured older ones spreading.

The paper is nevertheless bold and valuable. Philosophers have sought for

centuries to explain why states emerge. Too little time has been spent considering whether economic factors might have been at play. Although transforming the state of nature into a specific time and place means losing some of its complexity, doing so opens the door to the sort of experiment that could only have been imagined by earlier philosophers. If Hobbes or Locke could have studied something approximating the state of nature about which they were theorising, they surely would have tried. ■



自由交流

现代国家从何而来？

经济学家试图回答一个深奥的政治问题

那种状态一部分是隐喻，一部分是神话，一部分是历史。托马斯·霍布斯认为那种状态下的生活恶劣、野蛮而又短暂。约翰·洛克不认同，他宣称正是在那种状态中人们第一次学会了拥有财产。让-雅克·卢梭称那种状态下的人们生而自由，直至被锁链束缚。罗伯特·诺齐克认为，人们如此迫切地想要逃离那种状态，而导致了不可避免的结果：国家被创造了出来。

几个世纪来，有关“自然状态”（state of nature）——也就是政治活动组织成政府之前人们的生活状态——究竟是怎么回事一直吸引了哲学家的关注。要想查明其状貌是否如人们想象的那样几乎是不可能的。然而，思考“假如没有政府，人们的生活会是什么样”有助于回答一些深刻的问题。政治权力的局限性在哪里？现代国家是公民会自愿选择的东西吗？

如今，三位经济学家经过如此种种理论分析，认为他们得出了一些实证答案。纽约大学阿布扎比分校的罗伯特·艾伦（Robert Allen）、西北大学的利安德·赫尔德林（Leander Heldring）和诺丁汉大学的马蒂亚·贝尔塔齐尼（Mattia Bertazzini）认为，理解现代政治何以出现的关键不是什么隐喻，而是伊拉克古代河流不断变化的河道。他们在《美国经济评论》（American Economic Review）上发表的一篇论文中提出，最早的国家并不是像霍布斯认为的那样是作为躲避暴力的庇护所而被组织起来，发挥粘合作用的其实是经济因素。

在伊拉克最长的两条河流底格里斯河和幼发拉底河的沿岸，世界上一些最古老的定居点曾分布于此。美索不达米亚于5000年前完善了第一个已知的文字体系，为该地区赢得了“文明的摇篮”的美誉。由于洪水和干旱导致河床被淹没，这些河流的路径发生了变化。一旦河流改道，一些古代农民就无法灌溉庄稼了。

艾伦和他的合著者调查了河道变迁的时间与定居点数量和规模发生增长的时间是否存在关联。他们的方法是研究有记录的第一次河流变道的影响，这次改道发生在公元前2850年。当时农民面临的选择与哲学家对自然状态进行理论分析时所想象的差不多。那些被改道的河流抛在身后的人们可以回归游牧生活。或者他们可以联合起来修建灌溉系统，从遥远的河流取水。

一个哲学问题就这样转化成了类似于实验室实验的东西，只不过实验时间是在数千年前，实验范围绵延数百英里。而且，实验的结果十分清晰。将河道变迁后留下的盆地划分为一个个5公里乘以5公里的方格，每个方格在河流改道150年后比改道前50年出现定居点（以拥有一个公共建筑为标志，如寺庙或市场）的可能性要高14%。每个方格修建运河的可能性也会高12%，这种人工灌溉形式使得在远离河流的地方发展农业成为可能。五个新城市被创建，只有三个被废弃。位于该河一条新支流沿岸的城市埃什努纳（Esnunna）规模大大扩张了。

艾伦及其合著者说，这证明最初的国家是由农民出于经济原因展开合作而形成的。运河网络的成本太高，任何人都无法独自承担。但分摊了成本后，这项建设对每个人来说都是划算的。这样的决定意义重大。它们代表了政府在征税后提供基础设施的一些最早的例子，也就代表着最早那批国家的起源。

作者随后将关于国家起源的几个世纪的思考分为两个阵营。他们说，从麻省理工学院颇具影响力的经济学家达龙·阿西莫格鲁（Daron Acemoglu）到卡尔·马克思都属于第一个阵营，这一派认为国家说到底是在社会层面的讨价还价过程中产生的。富人和地位高的人为了个人利益而夺取权力，并不时提供道路、学校或警力等服务，以换取民众的合作。但如果美索不达米亚是属于这种情况，那也应该发生在改道河流最新流经的地方所形成的定居点才对。毕竟，它们开垦出了更肥沃的农田，也就创造出了更多的税收。

当河流改道流走时，美索不达米亚的农民似乎选择了团结起来，这为第二

个阵营提供了支持。这一派的哲学家包括洛克和卢梭，他们认为，当人们选择互相协调、放弃为所欲为的自由而选择建立起一种能调解争端并提供一定程度安全保障的国家形态时，政府便出现了。艾伦和他的合著者只分析了伊拉克美索不达米亚，但他们认为他们的研究成果应该能更普遍地适用于其他新生国家。换句话说，政府是人们自发选择的，而不是强加给公民的。

丨 河道蜿蜒

这着实是经济学家在抢地盘了，他们抢占了更常由政治理论家占据的领地。这项研究并非毫无瑕疵。也许某次未知的征服可以解释在研究所覆盖的时期内定居点的扩展。也许作者是错的，这种模式在其他地方并不成立。在其境内的河流真正开始改道之前，美索不达米亚的大河流域就已经有六座城市和多得多的定居点，有些已经存在了一千年。作者们坚称，他们只对新政府如何形成感兴趣，但他们实际上有可能捕捉到了那些更早形成的政府的扩展情况。

尽管如此，这篇论文还是大胆且有价值的。哲学家们花了几百年的时间解释国家为何出现。而花在考虑经济因素是否可能发挥了作用上的时间仍少之又少。尽管将自然状态具体化为一个特定的时间和地点会损失一些复杂性，但这样的操作为开展对早年的哲学家而言只能想想而已的那种实验打开了方便之门。如果霍布斯或洛克当初有办法研究一下跟自己推断的自然状态近似的東西，他们肯定乐于试一试。■



Baby-making

How to entice Japanese couples to have babies

A few cities are bucking the country's low birth rate

KISHIKI NORIYO pulls up outside a house in Akashi, a city in western Japan, in a truck emblazoned with a heart logo and the slogan: “Diaper delivery: we also deliver kindness.” She steps out with two bags of nappies. Higuchi Miki, a young mother, appears at the front door with a baby on her hip. Ms Kishiki is on the front line of a ten-year push by Akashi to encourage its residents to have children. It includes delivering free baby food as well as nappies, free medical care and school lunches. The effort seems to be paying off. Akashi's population has increased for ten years in a row, to more than 300,000.

This makes the town exceptional. Japan's birth rate dipped below the replacement level of 2.1 children per woman in the mid-1970s and has been steadily declining ever since. In 2022 the total number of births dropped below 800,000 for the first time since records began in 1899. Of Japan's 1,800 municipalities, only around 200 have a rising population. According to Kishida Fumio, the prime minister, this has put the country on “the brink”.

Even in Akashi, which is an easy commute from the cities of Osaka and Kobe, most of the growth is from migration, not births. The town's birth rate is 1.65. Yet given that the national figure is 1.3, this represents success of a sort. Schools are closing across Japan for want of pupils; Akashi is short of school places.

Many of the obstacles to young Japanese forming families, from high education costs to inflexible family laws, can only be tackled with national policies. But Akashi shows how local communities can make a difference.

That starts with improving access to nurseries and day care. Smaller projects such as Ms Kishiki's, which assist parents of young children, also help.

The population of Nagareyama, a commuter town outside Tokyo, is also booming. Over the past decade it has grown by 24%, to 211,000. The opening of an express-train service to Tokyo in 2005 largely explains that. But Nagareyama's commitment to child care—the “utmost priority” of its local government—has also attracted young families. In 2007 the town launched a bus service that picks up children at train stations and takes them to day-care centres. Nagareyama has increased its number of nurseries from 17 to more than 100 in the past 15 years. “Our aim was to create an environment in which parents can continue to work while raising children,” says Izaki Yoshiharu, the town's mayor.

Lack of child care is endemic across Japan. The central government has made some improvements in this regard, reducing the number of children waiting for a nursery place by 90%. It has also introduced tax and other economic incentives to encourage people to reproduce. But, as scattered examples like Akashi and Nagareyama suggest, making a real difference requires a more fundamental shift.

Kato Hisakazu of Meiji University reckons Japan “needs to foster a culture that is generous to children”. Efforts to build more child-care facilities are often blocked by locals who worry about increased noise. When Soup Stock, a popular restaurant chain, started offering free food for babies last April, it encountered a fierce online backlash from people who objected to the prospect of sitting next to crying infants.

Tanaka Yumi, a mother of two in Nagareyama, says young Japanese parents are used to receiving a “cold look” from their unsympathetic neighbours. But the abundance of young families in the town and the supportive

policies of its local administration provide reassurance, she says. “I have many friends in Nagareyama who decided to have a second or a third.”

Akashi’s pro-child efforts are intended to drive the same cultural change. Morioka Kazumi of the town’s child-care department says the nappy deliveries are more about “alleviating loneliness” than the goods themselves. Ms Higuchi, the young mother, appreciates that. “It’s nice when someone checks on me,” she says. “It makes me feel I’m on the right track.” ■



造娃

如何说服日本夫妇生孩子

有几个城市正在对抗这个国家的低出生率

岸城乃理世（Kishiki Noriyo，音译）把车停在日本西部的明石市的一栋房子前，这是一辆绘有爱心标志的卡车，上面还印着一句口号：“送纸尿裤，也送温暖。”她走下车，拿着两袋纸尿裤。年轻的母亲樋口三希（Higuchi Miki，音译）到门口迎接，怀里抱着个宝宝。明石市十年来一直鼓励当地居民生孩子，岸城就在这项工作的第一线忙碌。该市的政策包括免费提供婴儿食品、纸尿裤、医疗保健和学校午餐。努力似乎得到了回报，明石的人口已经连续增长了十年，如今达到三十多万。

这让这个小城与众不同。日本的出生率在20世纪70年代中期降至人口更替水平（每位妇女生育2.1个孩子）以下，之后一直持续下降。2022年，出生人口自1899年有记录以来首次降至80万以下。在日本的1800个自治市中，只有200个左右的人口在增长。日本首相岸田文雄表示，这已经让日本站在了“悬崖边上”。

即便在明石市（从这里到大阪和神户的通勤都很方便），人口增长也主要来自移民，而不是本地出生。这里的生育率为1.65。不过考虑到全国的平均值是1.3，这仍代表着某种程度上的成功。因为生源不足，日本各地都有学校在关闭；明石市则是入学名额短缺。

从高昂的教育成本到僵化的家庭法，组建家庭的日本年轻人面临诸多障碍，其中许多只能通过国家政策来解决。但明石展示了地方社区能如何发挥作用。首先要做的就是提供更方便的托儿所和日托服务。类似岸城所从事的那种协助幼儿父母的小项目也能帮上忙。

在东京外围的通勤城镇流山市，人口也在迅速增长。过去的十年中，这里的人口增长了24%，达到了21.1万。这很大程度上要归功于2005年通往东京的特快列车开通。但流山对儿童保育的承诺——当地政府的“重中之重”

一一也吸引了年轻家庭。2007年，流山启动了一项公共汽车服务，从火车站接上孩子并将他们送往日托中心。在过去的15年里，流山市托儿所的数量已经从17个增加到了100多个。市长井崎义治表示：“我们的目标是创造一种环境，让父母在抚养孩子的同时能够继续工作。”

儿童保育匮乏的问题在日本普遍存在。中央政府在这方面取得了一些进步，将等待入托的儿童人数减少了90%。政府还出台了税收优惠和其他经济激励措施，鼓励人们生育。但正如明石和流山等零星例子所示，要想真正改变现状，需要一种更加根本性的转变。

明治大学的加藤久和认为，日本“要培养一种对儿童更包容的文化”。建造更多儿童保育设施的行动常常遭到本地居民的阻挠，担心会产生噪音。去年4月，当知名餐饮连锁店Soup Stock开始为婴儿提供免费食品时，遭到了一些网友的强烈反对，他们不想以后吃饭时旁边有婴儿哭闹。

居住在流山的田中佑美（Tanaka Yumi，音译）是两个孩子的母亲，她说日本的年轻父母习惯了邻居嫌幼儿太吵的“冷眼”，但流山有很多年轻家庭，加上有本地政府的支持政策，让他们感到安心。“我在流山有好多朋友都决定生二胎或第三胎。”

明石对育儿的支持意在推动同样的文化转变。该市儿童保育部门的森冈和美（Morioka Kazumi，音译）表示，送纸尿裤更多是为了“减轻孤独感”，而不是为了纸尿裤本身。年轻的母亲樋口很感激这一点。“有人来看我真好，”她说，“这让我觉得自己走在正确的路上。”■



From the ruins

Which economy did best in 2023?

Another unlikely triumph

ALMOST EVERYONE expected a global recession in 2023, as central bankers raised interest rates to cool inflation. The consensus was wrong. Global GDP has probably grown by 3%. Job markets have held up. Inflation is on the way down. Stockmarkets have risen by 20%.

But this aggregate performance conceals wide variation. The Economist has compiled data on five indicators—inflation, “inflation breadth”, GDP, jobs and stockmarket performance—for 35 mostly rich countries. We have ranked them according to how well they have done on these measures, creating an overall score. The table shows the rankings, and some surprising results.

Top of the charts, for the second year running, is Greece—a remarkable result for an economy that was until recently a byword for mismanagement. Aside from South Korea, many of the other standout performers are in the Americas. The United States comes third. Canada and Chile are not far behind. Meanwhile, lots of the sluggards are in northern Europe, including Britain, Germany, Sweden and, bringing up the rear, Finland.

Tackling rising prices was the big challenge in 2023. Our first measure looks at “core” inflation, which excludes volatile components, such as energy and food, and is a good indicator of underlying inflationary pressure. Japan and South Korea have kept a lid on prices. In Switzerland core prices rose by just 1.3% year on year. Elsewhere in Europe, though, many countries still face serious pressure. In Hungary core inflation is running at 11% year on year. Finland, which is heavily dependent on Russian energy supplies, is also

struggling.

In most countries inflation is becoming less entrenched—as measured by “inflation breadth”, a measure that calculates the share of items in the consumer-price basket where prices are rising by more than 2% year on year. Central bankers in places including Chile and South Korea increased interest rates aggressively in 2022, sooner than many of their peers across the rich world, and now appear to be reaping the benefits. In South Korea inflation breadth has fallen from 73% to 60%. Central bankers in America and Canada, where inflation breadth has dropped even more sharply, can take some credit, too.

However, in other places, the battle against inflation is not even close to being won. Take Australia. Inflation there remains firmly entrenched, with the price of close to 90% of the items in the average person’s shopping basket rising by more than 2% year on year. Worse, inflation breadth is not coming down either. France and Germany are also in trouble. And so is Spain, where inflation seems to be becoming more entrenched over time.

Our next two measures—growth in employment and GDP—hint at the extent to which the 35 economies are delivering for ordinary folk. Nowhere fared spectacularly well. Across the world productivity growth is weak, limiting potential increases in GDP. Already tight labour markets at the start of 2023 meant there was little room for improvement when it came to jobs.

Only a few countries, though, actually saw their GDP decline. Ireland was the worst performer, with a drop of 4.1% (take that with a pinch of salt: there are big problems with the measurement of Irish GDP). Estonia, another country whacked by the fallout of Russia’s invasion of Ukraine, did badly. Britain and Germany also underperformed. Germany is struggling with the consequences of the energy-price shock and rising competition from imported Chinese cars. Britain is still dealing with the aftermath of

Brexit—most economists expect the country to suffer from weak economic growth in the coming years.

America, by contrast, did well on both GDP and employment. It has benefited from record-high energy production as well as a generous fiscal stimulus implemented in 2020 and 2021. The world's largest economy may have pulled along other countries. Canada's employment has risen smartly. Israel, which counts America as its largest trading partner, comes fourth in the overall ranking, although its war with Hamas, which began in October, makes the outlook for 2024 deeply uncertain.

You might think that the American stockmarket, stuffed with firms poised to benefit from the revolution in artificial intelligence, would have done well. In fact, adjusted for inflation it is a middling performer. The Australian stockmarket, filled with commodities firms managing a comedown from high prices in 2022, underperformed. The Finnish stockmarket had a poor year, with the share price of Nokia, a national champion, continuing its long, slow decline. Japan's firms, by contrast, are experiencing something of a renaissance. For that, thank reforms to corporate governance, which finally seem to be bearing fruit. The country's stockmarket was one of the best performers in 2023, rising in real terms by nearly 20%.

But for glorious equity returns, look thousands of miles west—to Greece. There the real value of the stockmarket has increased by more than 40%. Investors have looked afresh at Greek companies as the government implements a series of pro-market reforms. Although the country is still a lot poorer than it was before its almighty bust in the early 2010s, the IMF, once Greece's nemesis, praised "the digital transformation of the economy" and "increasing market competition" in a recent statement. While underperforming Finns can console themselves this Christmas by drowning their sorrows alone in their underwear (or getting päntsdrunk, as

it is known locally), the rest of the world should raise a glass of ouzo to this most unlikely of champions. ■



从废墟中崛起

哪个经济体在2023年表现最好？

又一次意想不到的凯旋

由于各国央行加息以抑制通货膨胀，几乎所有人都预计2023年会出现全球经济衰退。这个共识是错误的。全球GDP可能增长了3%。就业市场保持坚挺。通胀正在下降。股市上涨了20%。

但这一整体表现掩盖了个体间的很大差异。《经济学人》根据五个指标——通胀、“通胀广度”、GDP、就业和股市表现——给35个以富裕经济体为主的国家编制了数据。我们根据它们在这些指标上的表现给它们排名，创建了一个综合得分。下表显示了排名，以及一些出人意料的结果。

希腊在这个排行榜上连续第二年蝉联榜首——这对于一个前些年还是管理不善的代名词的经济体来说是一个引人瞩目的成绩。除了韩国，其他排名前列的经济体大都位于美洲。美国排名第三。加拿大和智利与它相差不远。与此同时，许多表现疲软的国家位于北欧，包括英国、德国、瑞典，还有垫底的芬兰。

控制物价上涨是2023年的大挑战。我们关注的第一个指标是“核心”通胀，它不包括能源和食品等价格波动较大的项目，能很好地显示根本性的通胀压力。日本和韩国成功控制了物价。在瑞士，核心价格同比仅上涨了1.3%。然而在欧洲其他地方，许多国家仍然面临严重的通胀压力。在匈牙利，与一年前相比，核心通胀达到11%。严重依赖俄罗斯能源的芬兰也在艰难应对。

以“通胀广度”来衡量，在大多数国家，通胀正在变得不那么顽固。该指标计算了消费价格篮子中价格同比上涨超过2%的商品所占的份额。智利和韩国等国央行在2022年积极加息，动作比许多富裕国家的央行更快，现在似乎正在收获成果。在韩国，通胀广度从73%降至60%。美国和加拿大的央行官员成绩也不错，这两国的通胀广度降幅还要更大。

然而在其他一些地方，对抗通胀的战斗还远未接近胜利。比如澳大利亚。那里的通胀仍然顽固，普通人的购物篮中近90%的商品价格同比上涨超过2%。更糟糕的是，通胀广度也没能缩减下来。法国和德国也陷入了困境。西班牙也是如此，该国的通胀似乎变得越来越顽固。

接下来的两个指标是就业和GDP增长，它们显示了这35个经济体为普通人带来多少好处。没有哪个地方表现得特别出色。全球范围内生产率增长疲弱，制约了GDP的潜在增长。劳动力市场在2023年初就已经很紧张，意味着就业方面的提升空间很小。

不过只有少数几个国家实际上出现了GDP下降。爱尔兰表现最差，下降了4.1%（对此要持保留态度：爱尔兰GDP的测量存在很大问题）。另一个受俄罗斯入侵乌克兰影响的国家爱沙尼亚表现也很差。英国和德国同样表现不佳。德国正受困于能源价格冲击和中国进口汽车造成的竞争加剧。英国仍在应付脱欧的后果——大多数经济学家预计未来几年该国经济增长疲弱。

相比之下，美国在GDP和就业方面都表现良好。它受益于创纪录的能源产量以及2020年和2021年实施的慷慨财政刺激。这个世界上最大的经济体可能也拉动了其他国家的增长。加拿大的就业率大幅上升。以美国为最大贸易伙伴的以色列在总排名中位列第四，尽管该国与哈马斯于10月开始的战争让它在2024年的前景存在很大的不确定性。

你可能认为美国股市应该表现不错，因为它里头满是准备从人工智能革命中大赚一笔的公司。事实上，经过通胀调整后，它的表现只能算是中等水平。澳大利亚股市表现不佳，其中很多大宗商品公司正在应对2022年高价以后的跌落。芬兰股市度过了惨淡的一年，龙头企业诺基亚的股价延续了长期缓慢下滑的走势。相比之下，日本的公司可以说正在经历一场复兴。这要感谢企业治理方面的改革，它似乎终于开始见效。该国的股市是2023年表现最佳的股市之一，实际涨幅接近20%。

但说到亮眼的股本回报，还要看向西数千英里的希腊。在那里，股市的实

际价值增长超过40%。政府实施了一系列亲市场的改革，令投资者重新审视了希腊企业。尽管该国相比2010年代初经济大崩溃之前仍要穷得多，但希腊曾经的“报应女神”国际货币基金组织（IMF）最近在一份声明中赞扬了“该经济体的数字化转型”和“市场竞争加大”。当表现不佳的芬兰人在圣诞节时穿着内衣独自借酒消愁的时候（按当地的说法，就是 päntsdrunk[穿着内裤喝个大醉]），世界其他地方应该向希腊这个最意想不到的冠军敬上一杯茴香酒。 ■



Toko-Tok

Will TikTok's GoTo gambit save its Indonesian business?

How the video app is navigating around a digital shakedown

THE MORE the world's youngsters love TikTok's viral videos, the more their elected elders hate the app. They decry it for supposedly corroding young minds and, worse, for its links to China, home to its parent company, ByteDance. Many in America want to ban it. India already has. In October Indonesia, another big and promising market, shut down TikTok's fledgling but lucrative sideline of selling goods via its videos, by requiring social-media firms to obtain an e-commerce licence—with no guarantee of success.

Such obstacles have forced TikTok to act strategically, for instance by moving its global headquarters to Singapore and hiring a Singaporean chief executive, which has put distance between it and its Chinese parent. In another canny move, on December 11th it announced that it was paying \$840m for a 75% stake in Tokopedia, the e-commerce arm of GoTo, an Indonesian tech conglomerate. It has also pledged to invest \$1.5bn in the tie-up.

The deal is something of a shotgun marriage, but it benefits both sides. GoTo, which has struggled to turn a profit in recent years, will no longer need to subsidise its loss-making retail arm. TikTok, for its part, will be allowed to restart its e-commerce operations. Sales on TikTok's app will be fulfilled by Tokopedia's logistics network (though, like all e-merchants in Indonesia, it must now charge minimum prices for products made abroad).

TikTok and Tokopedia separately account for 10% and 28%, respectively, of Indonesia's fast-growing e-commerce market, according to Momentum Works, a data firm. Together, they are a powerhouse, matching the market

share of Shopee, hitherto the country's biggest online emporium (owned by Sea Group, a Singaporean technology conglomerate).

Most important, an intimate link with a domestic champion makes TikTok look less like a foreign interloper. If the firm can make its new partnership work in the world's fourth-most-populous country, it could use this as a model for expansion and consolidation in other countries where it is greeted with wariness, such as Malaysia and the Philippines.

It will be an uphill struggle, and not just because of challenges particular to TikTok. All over the world, the advocates of international openness in digital commerce are losing the battle for hearts and minds. Last year Sea halted its expansion to India in the face of regulatory pressure, after its popular mobile game, "Free Fire", was banned. Stringent new European rules on cloud computing, including requirements to store local users' data locally, are aimed squarely at the American tech giants.

Last month America, itself in an increasingly isolationist mood, dropped earlier demands to liberalise trade in digital goods and services as part of the Indo-Pacific Economic Framework, the already flimsy pact which President Joe Biden's administration has been negotiating with 13 Asian allies. To thrive amid rising protectionism—digital and otherwise—TikTok and its rivals will need to show plenty of delicate diplomatic footwork. ■



Toko-Tok

TikTok联姻GoTo能否拯救其印尼业务？

这个视频应用如何应对数字业监管震荡

全球年轻人越爱看TikTok上的爆红视频，他们国家掌权的长辈们就越憎恶这个应用。他们谴责它可能侵蚀年轻人的思想，更严重的是，还谴责它与中国关联，因为其母公司字节跳动来自中国。美国的许多政客想把它禁掉。印度已经这样做了。去年10月，另一个潜力大市场印尼规定社交媒体公司须获得许可（并不保证能成功获批）方可开启电子商务交易，从而叫停了TikTok新生但利润丰厚的视频带货业务。

这些障碍迫使TikTok采取战略行动，例如将其全球总部迁至新加坡，并聘请了一位新加坡籍的首席执行官，从而与中国母公司拉开距离。它还使出了另一个精明的招数，于12月11日宣布将支付8.4亿美元收购印尼科技企业集团GoTo旗下电商平台Tokopedia75%的股份。它还承诺将在这次合作中投资15亿美元。

这笔交易有点像是一场无奈的联姻，但对双方都有利。近年来一直难以盈利的GoTo将不再需要贴补亏损的零售业务。而TikTok将获准重新启动其电商业务。TikTok平台上的销售将由Tokopedia的物流网络履行（尽管与印尼的所有电商一样，它现在必须对国外制造的产品限定最低价格）。

根据数据公司Momentum Works的统计，在印尼快速增长的电商市场上，TikTok和Tokopedia分别占到10%和28%的份额。合在一起，它们是一支强大的力量，市场份额堪比目前印尼最大的在线商城虾皮购物（由新加坡科技企业集团冬海集团所有）。

最重要的是，与本土领军企业建立亲密联系使TikTok看起来不那么像是一个擅闯的外来者。如果该公司能够在全球人口第四多的国家中让这一新伙伴关系顺利推进，它就能以此为模板，在其他目前对它怀有戒心的国家扩张和巩固地位，比如马来西亚和菲律宾。

这将是一场艰苦的进军，不仅仅是因为TikTok特有的挑战。在全球各地，支持数字商务国际开放的人们都在输掉人心之战。去年，在其热门手机游戏“Free Fire”被禁后，冬海在监管压力之下暂停了向印度的扩张。欧洲对云计算出台了严厉的新规定，包括要求在本地存储本地用户数据，就是在直接针对美国的科技巨头。

美国自身的孤立主义氛围也日益浓厚。11月，它放弃了之前要求在《印太经济框架》（Indo-Pacific Economic Framework）中实现数字商品和服务贸易自由化的要求，拜登政府一直在与13个亚洲盟国谈判的这个协议本就已经脆弱不稳。想要在不断高涨的保护主义浪潮（无论是数字经济还是其他方面）中蓬勃发展，TikTok及其竞争对手将需要走出很多高难度的外交步伐。■



Bartleby

How to master the art of delegation

You can entrust decisions to subordinates without regretting it

DELEGATING WELL is the six-pack of management: widely desired and harder to achieve the older you get. In theory, handing appropriate decisions off to people lower down the corporate ladder means greater satisfaction all round. Bosses get more time to concentrate on the issues that really deserve their attention. Middle managers and workers enjoy a greater sense of autonomy. And the organisation benefits from faster decision-making on the part of people who are better informed about the matter at hand. In practice, however, delegation is a minefield.

Some bosses do not even try to delegate. They may mistrust people below them or crave control. Their career success may simply have persuaded them of their own genius. But there are kinder explanations, too. Startup founders are conditioned to do everything, at least until firms get to a certain size. Plenty of managers shoulder more work than they should in order to protect their teams from overload.

Other managers do delegate but they do so for the wrong reasons. Studies suggest that people are likely to hand off decisions when choices are hard, when the consequences affect others and when they want to avoid being blamed for a bad outcome. In a paper from 2016 by Mary Steffel of Northeastern University and her co-authors, volunteers were told that they had to book hotel rooms at a conference, either for their own use or for their boss, and asked them if they would like to reserve the rooms themselves or delegate the task to an office manager. When they were choosing for the boss and the hotels were ropery, people were more likely to pass the job to the hapless office manager.

A new study, by Victor Maas and Bei Shi of Amsterdam Business School, reaffirms this bleak picture of human motivation. It found that people were more likely to hand work off to subordinates when the performance targets for that particular task were demanding; they were much happier to keep hold of tasks with targets that were easier to attain. If a habitual micromanager unexpectedly asks you to take the lead on something, in other words, run for the hills.

The great mass of managers fall into a greyer area. They may be full of good intentions to leave decisions to others but still find it hard to do so. What if you put trust in your team members but then discover you violently dislike the choices they make? What if you want to hand over some decisions but you know that your own bosses will hold you personally responsible for them? These problems can easily result in “faux-tonomy”—a lip-service version of delegation in which managers do not actually leave their teams to get on with things or underlings use their freedom solely to guess what the boss would like.

One way to navigate such problems is to use an explicit decision-making framework that tries to make it clear who is on the hook for what. These frameworks are not perfect. Project managers often use something called the RACI model. Its first two letters sort those who are “responsible” from those who are “accountable”, a distinction which normal people may find “confusing” and “incomprehensible”. Other, clearer frameworks are available. They have punchy names like DACI, DARE and DICE: you might be choosing a cloud-computing vendor but you get to feel a little like you are in the special forces.

As well as working out who does what, it helps to have a way to parse what kinds of decision can be delegated and what not. Before Jeff Bezos started hanging out in spacesuits and doing laughable photoshoots in Vogue, he liked to articulate his management philosophy in annual letters to

Amazon's shareholders. In 2015 he made a useful distinction between type-1 decisions ("one-way doors") that are important and irreversible, and type-2 decisions ("two-way doors") that can be reversed if they do not pan out. Type-1 decisions warrant slow, deliberative processes; type-2 decisions should be taken quickly by smaller groups. Having a theory of decisions improves choices on what to delegate and reduces the chance of regrets.

Delegating well requires a lot of judgment, too. Delegation is not all-or-nothing. A detached boss can be as demotivating as a micromanager; you have to stay informed on decisions and, on occasion, override them. But checking in at the right cadence, and letting people proceed with decisions that you would not have made yourself, demands self-restraint and discipline. Just like those abs. ■



巴托比

如何掌握授权的艺术

你可以放心地让下属去做决策而不后悔

良好的授权就像保持六块腹肌：人人渴求，但随着年龄增长愈发难实现。理论上，将适当的决策交给下属能让公司里的各方皆大欢喜。老板们有更多时间专注于真正值得关注的问题。中层管理人员和员工享受更大的自主权。让那些对手头的事务进展更清楚的人更快地做出决策，整个组织都能从中受益。然而在实践中，授权是个雷区。

有些老板连试都不试一下。他们可能不信任下属，或者渴望掌控一切。他们在事业上的成功可能就是会让他们相信自己是天才。但也有更客气的解释。创业公司创始人习惯于亲力亲为，至少在公司达到一定规模之前是这样。为了不让团队不堪重负，许多管理者承担了额外的工作。

另一些管理者确实在授权，但出发点是错误的。研究表明，当难以抉择、后果影响他人，以及想避免为糟糕的结果担责时，人们更有可能把决策甩手给他人。东北大学（Northeastern University）的玛丽·斯特费尔

（Mary Steffel）及其合著者在2016年发表了一篇论文。这项研究中的志愿者被告知他们必须为一次会议预订酒店房间，有的是为自己订，有的是为他们的老板订，然后问他们是想自己预订房间还是将这件事交给办公室经理。结果发现，当他们是想要为老板选房间并且酒店条件糟糕时，他们更有可能把它推给倒霉的办公室经理去干。

阿姆斯特丹商学院（Amsterdam Business School）的维克多·马斯

（Victor Maas）和石贝所做的一项新研究再次证实了人类动机的这种阴暗面。该研究发现，当某项任务的绩效目标挑战较大时，人们更有可能把工作交给下属；而对于目标较易达成的任务，留给自己的意愿就高多了。换句话说，如果一个惯于微观管理的上司突然要求你去负责某件事情，那么最好赶紧躲开。

大多数管理者处于灰色地带。他们可能满怀好意地想授权别人做决定，但仍觉得这做起来很难。如果你信任团队成员，让他们去做决定，但后来发现自己极不喜欢他们的选择，那该怎么办？如果你想授权下属做一些决策，但又知道你自己的上司会把它们算在你的头上，那又该怎么办？这些问题很容易导致“假自主”——一种表面上的授权，即管理者实际上并没有让团队自己决定事情，或者下属拥有的自由仅够用来猜老板喜欢什么。

应对这些问题的一种方法是使用一个清楚明白的决策框架，尽量明确谁对什么事情负责。这些框架并不完美。项目管理者经常使用称为RACI模型的东西。它的前两个字母将“负责”的人与“负有责任”的人区分开来，这可能会让一般人感到“困惑”和“难以理解”。还有其他更清晰的框架。它们有响亮的名字，如DACI、DARE和DICE：你可能正在选择云计算供应商，但会感觉自己有点像在特种部队里。

除了弄清谁做什么之外，有一套区分哪些决策可以授权、哪些不可以的方法也会有帮助。在贝索斯开始穿上太空服、为《Vogue》杂志拍摄好笑的照片之前，他喜欢在每年给亚马逊股东的信中表达自己的管理理念。在2015年，他对重要且不可逆转的第一类决策（“单向门”）和可以在不奏效时撤销的第二类决策（“双向门”）做了有用的区分。第一类决策要审慎，可以做得慢一些；第二类决策应由较小的团队迅速做出。有了一套决策理论，可以在授权时做出更好的选择，减少后悔的机会。

良好的授权还需要做很多的判断。授权并非撒手不管。和微观管理者一样，冷漠的老板也可能会让人感到泄气；你必须保持对决策的了解，并在必要时出手干预。但是，以正确的节奏了解状况，并允许人们做出你自己不会做出的决策，这需要自我克制和纪律。就像保持腹肌那样。■



A surplus of anomalies

Is China understating its own export success?

The \$230bn puzzle at the heart of the country's trade figures

CHINA'S CURRENT-ACCOUNT surplus was once one of the most controversial statistics in economics. The figure, which peaked at almost 10% of GDP in 2007, measures the gap between China's earning and its spending, driven largely by its trade surplus and the income it receives from its foreign assets. For much of the past two decades, China's surpluses have left it open to the charge of mercantilism—of stealing jobs by unfairly boosting its exports. Some trading partners now worry about a similar shock if the country's output of electric vehicles grows too quickly.

But China's current-account surplus is now modest: \$312bn or 1.5% of GDP over the past year, according to the country's State Administration of Foreign Exchange (SAFE). That is below the 3% threshold that America's Treasury deems excessive.

Is the figure reliable? Some, such as Brad Setser of the Council on Foreign Relations and Matthew Klein, a financial commentator, believe that the official numbers are dramatically understated. China's true surplus, Mr Klein reckons, is now “about as large as it has ever been, relative to the size of the world economy”. They offer two arguments. First, China may be understating income from its foreign assets. Second, it may be understating exports.

According to SAFE, the income China earns on its stock of foreign assets plunged from mid-2021 to mid-2022. This seems odd given rising global interest rates. Mr Setser's alternative estimate, based on assumptions about China's assets, would add about \$200bn to the surplus.

China's goods surplus also appears smaller in SAFE's figures than it does in China's own customs data. The gap was \$230bn over the past year. "That is real money, even for China," says Mr Setser.

China might take some comfort from a bigger surplus. But it has an unsettling implication. What is happening to the additional dollars China is earning? Since they are not showing up on the books of China's central bank or its state-owned banks, they must be offset by a hidden capital outflow. Such outflows typically end up in a residual category of the ledger. Mr Setser believes this residual should be about 2% of GDP, not the official figure of near zero.

SAFE has a different explanation. It attributes the export gap largely to China's free-trade zones and similar enclaves. These lie inside China's territory but outside its official tariff border (see diagram). Goods leaving these enclaves for the rest of the world are counted as exports by customs but not by SAFE. Adam Wolfe of Absolute Strategy Research points out that these zones account for a growing share of China's exports. That may explain why the gap has emerged only in the past two years.

Mr Setser is unconvinced. If China's free-trade zones have enjoyed a dramatic export boom, it should produce ripples elsewhere. Wages earned by workers, for example, should appear as increased remittances. In fact, they have risen only a little. And as Mr Wolfe points out, even if the official current-account surplus is correctly calculated, it may be of little comfort to China's trading partners. After all, if the country's domestic demand remains weak, goods made in its free-trade zones may flood foreign markets. The rest of the world will count them, and experience them, as Chinese imports, even if SAFE does not count them as Chinese exports. ■



差异盈余

中国是否低报了它的出口成就？

中国贸易数据的核心是一个2300亿美元的谜团

中国的经常账户盈余曾经是经济学中最具争议的统计数据之一。这个数字衡量中国的收入和支出之间的差距，它在2007年达到峰值，几乎占到GDP的10%，这主要是受到中国的贸易顺差和从海外资产中获得的收入推动。在过去二十年的大部分时间里，中国的盈余让它被指责为重商主义——即通过不公平地促进出口来窃取就业机会。一些贸易伙伴现在担心，如果中国的电动汽车产量增长过快，可能会出现类似的冲击。

但是中国的经常项目盈余现在并不算高：按照国家外汇管理局的数据，去年中国的经常项目盈余为3120亿美元，占GDP的1.5%。这低于美国财政部认为过高的3%的门槛。

这个数字可信吗？有些人，比如美国外交关系协会（Council on Foreign Relations）的布拉德·塞瑟（Brad Setser）和金融评论员马修·克莱因（Matthew Klein），认为官方数据是大大低报了。克莱因认为，中国目前的真实盈余“相对于世界经济规模而言和过去任何时候都差不多大”。他们提出了两个论据。首先，中国可能算少了其海外资产的收入。其次，它对出口的统计可能打了折扣。

根据外汇局的数据，从2021年中到2022年中，中国从它的外国资产存量中获得的收入大幅下降。考虑到全球利率不断上升，这似乎有些奇怪。塞瑟根据对中国资产的假定做了另一种估算，按照他的算法，盈余要再多出约2000亿美元。

外汇局数据里中国的商品贸易顺差也小于中国海关的数据。过去一年这两个数字的差异达到2300亿美元。“即便对中国来说，这也是一大笔钱。”塞瑟表示。

盈余实则更大可能会让中国得到一些安慰。但这透露出的隐情让人不安。中国多赚的那些美元去了哪里？既然它们没有出现在中国央行或国有银行的账簿上，那么必定是被一种隐性的资本外流抵消了。这样的流出通常最终会入在分类账的一个剩余类别里。塞瑟认为这个剩余部分应该在GDP的2%左右，而不是官方数据中的接近于零。

外汇局对此有不同的解释。它把出口额上的差异主要归因于中国的自由贸易区和类似自贸区的飞地。这些区域位于中国境内，但在官方征收关税的边界之外（见图）。从这些飞地运往世界其他地方的货物在海关被算作出口，但在外汇局不算出口。Absolute Strategy Research的亚当·沃尔夫（Adam Wolfe）指出，这些自贸区在中国出口中所占的份额越来越大。这或许可以解释为什么数字差异只在过去两年才出现。

这没能让塞瑟全然信服。如果中国的自贸区出口非常繁荣，那应该会在其他地方产生连锁反应。例如，工人挣的工资应该表现为汇款增加。但实际上汇款只增长了一点点。而正如沃尔夫指出的那样，即使官方经常项目盈余的计算是正确的，可能也无法给中国的贸易伙伴带来多少安慰。毕竟，如果中国国内的需求持续疲软，其自贸区生产的商品可能会涌向国外市场。即使中国外汇局不把它们算作中国的出口，世界其他国家还是会把它们算作从中国的进口，并体验到它们的存在。■



Chaguan

China's cities compete for kids

Enlightened self-interest nudges rich places to woo rural families

AS A RULE, China's central planners do not say much about love. But look closely at recent plans from some reform-minded provinces—notably schemes that try to address a shrinking population—and appeals to hearts as well as minds leap from the page.

Take, for instance, a five-year plan to help rural migrants settle down in the cities of Zhejiang, a prosperous coastal province, and ideally to bring their young children with them. At first sight, Zhejiang's proposal, issued in July and covering 2023 to 2027, is dry stuff. One section explains how, in every city except the provincial capital, Hangzhou, recently arrived families can access places at city-funded schools and other public services. They qualify without buying a home or securing a local hukou (household registration). The hukou system has been used to regulate internal migration since Maoist times, when the Communist Party feared hungry peasants might crowd into cities. On the ground in Zhejiang the human import of these changes is well understood.

Chaguan travelled to Yiwu, a city of 1.9m in Zhejiang that is a trading hub for small commodities, supplying the world with pencils and parasols, shoelaces and shopping trolleys. He heard locals and migrants weigh the likely impact of relaxed residency rules on Yiwu's economy, on school waiting lists and on housing prices. Strikingly often, the same people then stopped talking about statistics and spoke of how the reforms make them feel.

Though Zhejiang stands out for reforming zeal, cities across China are being encouraged to hand out hukou papers more easily. Some are opening

public services to migrants who prefer to remain registered in their rural birthplaces. Both economics and demographics are driving change. Fertility rates are dropping fast and China's population declined in 2022 for the first time since the early 1960s. Natives of some of China's biggest and richest cities are proving indifferent to offers of baby-bonuses and other government incentives. Far-sighted provinces and cities are now focusing on a stock of young people who have already been born: China's 67m "left-behind children". That is the term for youngsters being raised by relatives or in boarding schools in villages, county towns or minor provincial cities, while one or both parents works as a migrant away from home.

Even some of China's biggest cities are anxious about maintaining their populations, says Lu Ming, an economist at Shanghai Jiaotong University and a prominent advocate of hukou reform. What is more, China is generating fewer of the factory jobs that can be filled by migrants straight from the countryside, and creating more service-sector jobs that require an understanding of city folk and their ways, notes Professor Lu. By way of example he cites jobs in nursing or housekeeping or as decorators, adding that workers raised and educated in cities are best placed to fill such vacancies. Cities have been offering hukous to university graduates and other skilled workers for years. Now, the contest is on for blue-collar families, the professor suggests.

Not every city has the means to compete. Yiwu, a wealthy place, has spent heavily on wooing young families this year. To help outsiders, the city closed 28 private schools that catered to migrant children, some of which charged as much as 20,000 yuan (\$2,811) a year. Others offered classes in shabby industrial premises. The city converted 24 into publicly funded schools, bringing 25,000 migrant children into the state sector. It built new primary schools, too, with one campus costing 224m yuan.

Migrant parents have mixed reactions. Yiwu's wide avenues are lined with

commercial complexes devoted to specific industries. Outside a centre for stationery merchants, your columnist found three men from the same rural corner of Hunan province. They eke out a living selling adhesive price labels from plastic crates balanced on electric scooters. One used to pay over 6,000 yuan a year to send his child to a local private school. The same school is now public and costs him a tenth of that. Yiwu “wants to hang on to more outsiders”, suggests that lucky father. A younger colleague will not be moving his 13-year-old daughter from Hunan to the city, however. “Of course, she’d prefer to live with her parents,” he admits. But he and his wife both work in Yiwu, often till midnight or later. “We don’t have time to take care of the child here,” says the label-seller, smoking as he waits for customers.

Inside the mall, a mother of one from elsewhere in Zhejiang sells children’s diaries and pens to buyers from around the world. The reforms leave her both grateful and sceptical. Migrants who rent homes and pay social-security contributions can now access city schools, even without a full hukou, she agrees. But they rarely land spots at Yiwu’s best schools. Homeowners and longstanding hukou-holders have a higher priority than newcomers who rent, she explains. In a nearby shop, a mother of two who moved to Yiwu years ago reports that she paid a hefty premium to live near a good school. She ventures that it would be “very unfair” on homebuyers if the newcomers could access the best schools.

| *A benign contest for growth*

Some migrants prefer a life in two places. Some keep a rural hukou to maintain their rights to village land. A woman from southern China may send her daughter back to her home province to take university-entrance exams in ten years’ time. Back home, the competition is less “ferocious” than in wealthy Zhejiang, she says.

In a playground near a new primary school, a retired migrant worker from

Hunan talks proudly of her grown children and the four grandchildren that she now helps to raise, each of whom has a hukou from Yiwu. A generation ago, her own children lived in her home village and she saw them twice a year. She supposes that her children missed her, she says, with a strained laugh. “But I don’t know and I would not ask.” China remains full of such sad tales. Self-interest now prompts cities and provinces to help more families stay together. Easing heartache will be one of their rewards. ■



茶馆

中国城市的抢娃大作战

开明的自利促使富裕地区想办法吸引农村家庭

中国的中央规划者一般很少谈“爱”。但仔细阅读一些具改革意识的省份近期出台的规划，尤其是那些试图应对人口减少的方案，会明显感受到字里行间不仅试图晓之以理，更试着动之以情。

以浙江省为例，这个富裕的沿海省份制定了一项为期五年的计划，旨在协助农业转移人口在该省各市安家落户，且最好他们年幼的子女也随迁。这份提案于7月发布，覆盖2023年至2027年，初看起来似乎索然无味。其中一个章节说明了除省会杭州以外，其他所有城市的新迁入家庭如何能获得公办学校入学名额和享受其他公共服务，而无需购房或获得当地户口。户籍制度自毛泽东时代起便被用来管理国家内部人口迁移，当时共产党担心饥饿的农民可能涌入城市。在浙江，民众对于目前这些变革在人口上的用意领会得非常清楚。

笔者走访了义乌。浙江这座190万人口的城市是一个小商品贸易中心，向全世界供应铅笔、遮阳伞、鞋带和购物车。笔者在那里听到本地人和外来工琢磨居住规定放宽会怎样影响义乌的经济、入学排队和房价。几无例外的是，本来说着数据的他们往往聊着聊着就开始述说起自己对这些改革的感受来。

尽管浙江的改革热情尤为突出，但全国各地的城市都受到敦促来放宽落户限制。一些城市正向更愿意保留农村出生地户籍的务工者开放公共服务。经济和人口结构变化都在推动变化。生育率迅速下降，2022年，中国人口自20世纪60年代初以来首次减少。一些中国最大、最富裕城市的居民对于生育奖励等政府激励措施反应冷淡。一些目光长远的省份和城市现在将关注的焦点放在了一个已经出生的年轻群体上，那就是中国6700万“留守儿童”。这些孩子的父母两人或之一远离家乡在外务工，孩子被留在乡村、县城或小城市里，由亲属抚养或上寄宿学校。

即便是中国最大的一些城市也在为保持人口规模而忧虑，上海交通大学经济学家、户口制度改革的知名倡导者陆铭表示。此外，他指出，中国可由直接从农村迁移而来的民工填补的工厂岗位在减少，而需要从业者懂得城里人及其生活方式的服务业岗位在增多。陆铭教授列举了护理、家政、装修这样的城市工作，并表示在城市长大和受教育的工人更适合填补这些职位空缺。多年来，城市一直向大学毕业生和其他技术工人提供户口。陆铭表示，如今争夺的焦点正转向蓝领家庭。

并非每个城市都有财力参与竞争。富裕的义乌今年投入了大量资金吸引年轻家庭。为帮助外来人口，该市关闭了28所专门面向农民工子女的民办学校，其中一些学费高达2万元一年。其他一些学校在破旧的工业场地授课。该市将其中的24所转为公办学校，将2.5万名农民工子女纳入公立教育体系。此外，义乌还新建小学，其中一个校园耗资2.24亿元。

农民工父母对此反应不一。义乌宽阔的大道两旁，聚集特定行业的商业综合体一字排开。在一个文具商场外，笔者发现了三名来自湖南省同一个村落的男子。他们以卖不干胶价签为生，“柜台”就是放在电瓶车上的塑料箱。其中一人过去每年花6000多元送孩子上义乌的民办学校。这所学校现在成了公办的，费用只有过去的十分之一。义乌“想留住更多外地人”，这位幸运的父亲表示。不过，旁边比他年轻的搭档却不准备将13岁的女儿从湖南接过来。“她当然更愿意和父母在一起。”他承认。但他和妻子两人都在义乌工作，经常会忙到半夜或更晚。“我们在这里没时间照顾孩子。”这位卖标签的父亲说着，一边抽烟一边等待顾客上前。

在商场里面，一名妇女向来自世界各地的买家销售儿童日记本和笔。她来自浙江其他地方，有一个孩子。对眼下这些改革，她既感激又怀疑。她承认，即使没有完整的户口，租房子并缴纳社保的农民工如今也能让孩子入读城市里的学校。但他们很少能够进义乌最好的学校。有房者和有常住户口的人比租房的新迁人口有优先权，她解释说。在旁边一个店铺里，一位迁居义乌多年的两个孩子的母亲称，她下了血本在一所好学校附近安了家。她提出，如果新来的人能进最好的学校，那对买了房的人就“非常不公平”。

一些外来务工者更倾向两地来回的生活。有些人保留农村户口以保持在农村的土地权益。一名来自中国南方的妇女可能会在十年后把女儿送回家乡省份参加高考。她说，在老家，竞争不像在富庶的浙江那么“激烈”。

在一所新建小学附近的操场上，一名来自湖南的退休农民工自豪地谈起她的成年子女和她现在帮忙带的四个孙辈，每个孩子都有义乌户口。二三十年前，她自己的孩子住在农村老家，她每年见他们两次。那会儿孩子们应该是想念自己的吧，她说，勉强地笑了一声。“但我不确定，也不会去问。”中国仍然有很多这样的悲伤故事。出于自身利益考虑，如今各省市正帮助更多的家庭团聚。减轻这样的内心伤痛将是它们收获的回报之一。





Future of chipmaking

Jensen Huang says Moore's law is dead. Not quite yet

3D components and exotic new materials can keep it going for a while longer

TWO YEARS shy of its 60th birthday, Moore's law has become a bit like Schrödinger's hypothetical cat—at once dead and alive. In 1965 Gordon Moore, one of the co-founders of Intel, observed that the number of transistors—a type of electronic component—that could be crammed onto a microchip was doubling every 12 months, a figure he later revised to every two years.

That observation became an aspiration that set the pace for the entire computing industry. Chips produced in 1971 could fit 200 transistors into one square millimetre. Today's most advanced chips cram 130m into the same space, and each operates tens of thousands of times more quickly to boot. If cars had improved at the same rate, modern ones would have top speeds in the tens of millions of miles per hour.

Moore knew full well that the process could not go on for ever. Each doubling is more difficult, and more expensive, than the last. In September 2022 Jensen Huang, the boss of Nvidia, a chipmaker, became the latest observer to call time, declaring that Moore's law was "dead". But not everyone agrees. Days later, Intel's chief Pat Gelsinger reported that Moore's maxim was, in fact, "alive and well".

Delegates to the International Electron Devices Meeting (IEDM), a chip-industry shindig held every year in San Francisco, were mostly on Mr Gelsinger's side. Researchers showed off several ideas dedicated to keeping Moore's law going, from exploiting the third dimension to sandwiching chips together and even moving beyond silicon, the material from which microchips have been made for the past half-century.

A transistor is to electricity what a tap is to water. Current flows from a transistor's source to its drain via a gate. When a voltage is applied to the gate, the current is on: a binary 1. With no voltage on the gate, the current stops: a binary 0. It is from these 1s and 0s that every computer program, from climate models and ChatGPT to Tinder and Grand Theft Auto, is built.

| *Small is beautiful*

For decades transistors were built as mostly flat structures, with the gate sitting atop a channel of silicon linking the source and drain. Making them smaller brought welcome side benefits. Smaller transistors could switch on and off more quickly, and required less power to do so, a phenomenon known as Dennard scaling.

By the mid-2000s, though, Dennard scaling was dead. As the distance between a transistor's source and drain shrinks, quantum effects cause the gate to begin to lose control of the channel, and electrons move through even when the transistor is meant to be off. That leakage wastes power and causes excess heat that cannot be easily disposed of. Faced with this "power wall", chip speeds stalled even as transistor counts kept rising (see chart).

In 2012 Intel began to build chips in three dimensions. It turned the flat conducting channel into a fin standing proud of the surface. That allowed the gate to wrap around the channel on three sides, helping it reassert control (see diagram). These transistors, called "finFETs", leak less current, switch a third faster and consume about half as much power as the previous generation. But there is a limit to making these fins thinner and taller, and chipmakers are now approaching it.

The next step is to turn the fins side on such that the gate surrounds them completely, giving it maximum control. Samsung, a South Korean electronics giant, is already using such transistors, called "nanosheets", in its newest products. Intel and TSMC, a Taiwanese chip foundry, are

expected to follow soon. By stacking multiple sheets and reducing their length, transistor sizes can drop by a further 30%.

Szuya Liao, a researcher at TSMC, compares going 3D to urban densification—replacing sprawling suburbs with packed skyscrapers. And it is not just transistors that are getting taller. Chips group transistors into logic gates, which carry out elementary logical operations. The simplest is the inverter, or “NOT” gate, which spits out a 0 when fed a 1 and vice versa. Logic gates are made by combining two different types of transistor, called n-type and p-type, which are produced by “doping” silicon with other chemicals to modify its electrical properties. An inverter requires one of each, usually placed side by side.

At IEDM Ms Liao and her colleagues showed off an inverter called a CFET built from transistors that are stacked on top of each other instead. That reduces the inverter’s footprint drastically, to roughly that of an individual transistor. TSMC says that going 3D frees up room to add insulating layers, which means the transistors inside the inverter leak less current, which wastes less energy and produces less heat.

The ultimate development in 3D chip-making is to stack entire chips atop one another. One big limitation to a modern processor’s performance is how fast it can receive data to crunch from memory chips elsewhere in the computer. Shuttling data around a machine uses a lot of energy, and can take tens of nanoseconds, or billionths of a second—a long time for a computer.

Julien Ryckaert, a researcher at Imec, a chip-research organisation in Belgium, explained how 3D stacking can help. Sandwiching memory chips between data-crunching ones drastically reduces both the time and energy necessary to get data to where it needs to be. In 2022 AMD, an American firm whose products are built by TSMC, introduced its “X3D” products,

which use 3D technology to stick a big blob of memory directly on top of a processor.

As with cities, though, density also means congestion. A microchip is a complicated electrical circuit that is built on a circular silicon wafer, starting from the bottom up. (Intel likens it to making a pizza.) First the transistors are made. These are topped with layers of metal wires that transport both electrical power and signals. Modern chips may have more than 15 layers of such wires.

As chips get denser, routing those power and data lines gets harder. Roundabout routes waste energy, and power lines can interfere with data ones. 3D logic gates, which pack yet more transistors into a given area, make things worse.

To untangle this mess, chipmakers are moving power lines below the transistors, an approach called “backside power delivery”. Transistors and data lines are built as before. Then the wafer is flipped and thick power lines are added to the bottom. Putting the power wires along the underside of the chip means fundamental changes to the way expensive chip factories operate. But shortening the length of the power lines means less wasted energy and cooler-running chips. It also frees up nearly a fifth of the area above the transistors, giving designers more room to squeeze in extra data lines. The end result is faster, more power efficient devices without tinkering with transistor sizes. Intel plans to use backside power in its chips from next year, though combining it with 3D transistors in full production is still a while away.

Even making use of an extra dimension has its limits. Once a transistor’s gate length approaches ten nanometres the channel it governs needs to be thinner than about four nanometres. At these tiny sizes—mere tens of atoms across—current leakage becomes much worse. Electrons slow down

because silicon's surface roughness hinders their movement, reducing the transistor's ability to switch on and off properly.

Some researchers are therefore investigating the idea of abandoning silicon, the material upon which the computer age has been built, for a new class of materials called transition metal dichalcogenides (TMDs). These can be made in sheets just three atoms thick. Many have electrical properties that mean they leak less current from even the tiniest of transistors.

Three TMDs in particular look promising: molybdenum disulphide, tungsten disulphide and tungsten diselenide. But while the industry has six decades of experience with silicon, TMDs are much less well understood. Engineers have already found that their ultra-thin profile makes it difficult to connect transistors made from them with a chip's metal layers. Consistent production is also tricky, particularly at the scales needed for reliable mass production. And the materials' chemical properties mean it is harder to dope them to produce n-type and p-type transistors.

| *The atomic age*

Those problems are probably not insurmountable. (Silicon suffered from doping problems of its own in the industry's early days.) At the IEDM, Intel was showing off an inverter built out of TMDs. But Eric Pop, an electrical engineer at Stanford University, thinks it will be a long while before they replace silicon in commercial products. For most applications, he says, silicon remains "good enough."

At some point, the day will arrive when no amount of clever technology can shrink transistors still further (it is hard to see, for instance, how one could be built with less than an atom's worth of stuff). As Moore himself warned in 2003, "no exponential is for ever." But, he told the assembled engineers,

“your job is delaying for ever”. Chipmakers have done an admirable job of that in the two decades since he spoke. And they have at least sketched out a path for the next two decades, too. ■



芯片制造的未来

黄仁勋说摩尔定律已死。其实还没有

3D元件和奇异新型材料可给它续会儿命

摩尔定律还差两年就要满60岁了，如今它已经变得有点像薛定谔的猫——既死又活。1965年，英特尔公司的联合创始人之一戈登·摩尔（Gordon Moore）提出，一块微芯片上可容纳的晶体管（一种电子元件）的数量每一年就会翻一番，后来他又将翻番的时间从一年改为两年。

摩尔的言论成了一种追求，为整个计算机行业设定了前进步调。1971年生产的芯片可以在一平方毫米中塞入200个晶体管。如今最先进的芯片可以在同样的面积上塞入1.3亿个晶体管，而且每个晶体管的运行速度也比当初快几万倍。如果汽车也以同样的速度进步，那么现如今汽车的最高时速可达到数千万英里。

摩尔很清楚这一进程不可能永远持续下去。每一次翻番都比上一次更困难、成本也更高。2022年9月，芯片制造商英伟达的老板黄仁勋成为最新一个唱衰摩尔定律的人，他宣称摩尔定律“已死”。但并非所有人都赞同他的观点。几天后，英特尔的CEO帕特·基辛格（Pat Gelsinger）表示，摩尔定律其实还“活得好好的”。

参与国际电子器件大会（IEDM）的代表大多站在基辛格一边。这一芯片行业的盛会每年在旧金山举行。研究人员在会上展示了若干让摩尔定律持续下去的创意，比如利用3D结构或把多个芯片堆叠起来，甚至不再使用硅这种过去半个世纪里用来制造微芯片的材料。

晶体管之于电流就像水龙头之于自来水一样。电流通过栅极从晶体管的源极流向漏极。当电压施加到栅极上时，电流接通——我们用二进制中的1来表示。当栅极上没有电压时，电流断开——用二进制中的0来表示。从气候模型、ChatGPT到手机交友应用Tinder、《侠盗猎车手》游戏，所有的计算机程序都由这些1和0构成。

几十年来，晶体管大多是平面结构，栅极位于连接源极和漏极的硅沟道上。制造更小的晶体管也带来了一些可喜的附带好处：晶体管越小，开关速度就越快，并且功耗也越低。这种现象被称为登纳德缩放定律。

然而到2000年代中期，登纳德缩放定律已经失效。随着晶体管源极和漏极之间距离的缩短，出现了量子效应，导致栅极开始无法控制沟道，并且即使晶体管处于关闭状态，电子也会隧穿。如此造成的漏电不仅浪费电能，还会产生难以散去的过多热量。因为这堵“功耗墙”，虽然芯片上的晶体管数量在持续增多，芯片速度却停滞不前（见图表）。

2012年，英特尔开始制造3D芯片。它把原来平面的电流通道变成了凸起于表面上的“鳍”。如此一来，栅极从三面包裹通道，帮助它增强了控制力（见图）。这些鳍式场效应晶体管（FinFET）漏电减少，开关速度提升三分之一，且功耗只有上一代晶体管的一半左右。但是这些鳍不可能无限变薄或变高，芯片制造商如今正在接近其极限。

下一步是翻转鳍，使栅极完全包裹住沟道，从而最大限度地控制沟道。韩国电子巨头三星已经在其最新产品中使用了这种叫作“纳米片”的晶体管。预计英特尔和芯片代工厂台积电也将很快跟进。通过把多个纳米片堆叠起来并缩短其长度，晶体管的尺寸可以进一步缩小30%。

台积电的研究人员廖思雅把芯片3D化比作提高城市密度——用密集的摩天大楼取代四处延伸的郊区。这不仅指晶体管在不断变高。芯片将晶体管分组组成逻辑门，这些逻辑门执行基本的逻辑运算。最简单的逻辑门是反相器，又称“非门”——如果输入为1，则其输出为0；反之亦然。逻辑门是由N型和P型两种不同类型的晶体管组成，这两种晶体管都是通过在硅中“掺杂”其他化学物质来改变其导电性能而制成的。一个反相器需要N型和P型晶体管各一个，通常并排放置。

在今年的IEDM上，廖思雅及其同事展示了一种名为互补场效应晶体管（CFET）的反相器，由晶体管相互堆叠而成。这大幅减少了反相器占用

的空间，差不多只需要一个晶体管的空间。台积电表示，芯片3D化可为增加绝缘层腾出空间，这样就可以减少反相器内部晶体管的漏电，从而降低能耗、减少发热。

研发3D芯片制造的终极目标是将整块芯片一个个堆叠起来，形成一个整体。现代处理器要从计算机其他位置的存储芯片接收所要处理的数据，接收速度是限制其性能的一大因素。计算机内部的数据传输需要消耗大量能量，并且可能需要几十纳秒（一纳秒等于十亿分之一秒）——这对计算机来说是很长的时间。

比利时微电子研究中心（Imec）的研究人员朱利安·瑞卡特（Julien Ryckaert）解释了3D堆叠技术如何可以帮助解决这个问题。将存储芯片夹在数据处理芯片之间，大大减少了将数据传输到指定位置所需的时间和能量。2022年，美国公司AMD（台积电为其代工）推出了它的“X3D”处理器，使用3D技术将一大片内存直接连接到处理器上面。

然而和城市一样，密集也意味着拥堵。微芯片是以圆形硅片为基片、自下而上构建起来的复杂电路。（英特尔将它比作做披萨。）首先要制造晶体管。它们的上端有多层用来传输电力和信号的金属线。现代芯片可能有15层以上这样的金属线。

随着芯片上的元件越来越密集，这些电源和数据线的布线变得越来越困难。线路迂回会浪费能量，而且电源线会对数据线造成干扰。3D的逻辑门因为在一定区域内塞入了更多的晶体管，布线就变得更难了。

为了解开这团“乱麻”，芯片制造商正在将电源线移至晶体管下方，这种技术被称为“背面供电”。晶体管和数据线的构建方式和以前一样。然后将晶圆片翻转过来，并在底部加上较粗的电源线。把电源线排布在芯片底部将从根本上改变造价高昂的芯片工厂的运作方式。但缩短电源线的长度意味着减少能耗，也能降低芯片运行时的温度。该技术还在晶体管上方腾出了近五分之一的面积，给设计人员留出了更多空间来塞进额外的数据线。该技术最终会制造出速度更快、能效更高的芯片，而不需要调整晶体管的尺

寸。英特尔计划从明年开始在其芯片中使用背面供电，不过把它与全面投产的3D晶体管结合起来还需要时日。

而即便是利用3D技术也有其上限。一旦晶体管的栅极长度接近10纳米，它所控制的沟道就需窄至大约在4纳米以下。如此微小的尺寸只够几十个原子通过，漏电因而会更加严重。由于硅片表面的粗糙度阻碍了电子的运动，电子的速度会变慢，也就会影响晶体管的正常开关。

因此，一些研究人员正在研究是否可以弃用硅这种让计算机时代得以建立的基础材料，转而使用一类名为过渡金属二硫属化物（TMD）的新材料。TMD可以制成只有三个原子厚的薄片。许多TMD具有良好的导电性能，即便制成最小的晶体管也能减少漏电。

三种尤其被看好的TMD材料分别是：二硫化钼、二硫化钨和二硒化钨。但是芯片制造行业与硅材料打了60年交道，对TMD的了解却要少得多。工程师们已经发现，由于TMD是一种超薄材料，用它们制造的晶体管很难与芯片的金属层连接起来。TMD也很难持续生产，尤其是要达到可靠的批量生产所需规模的话。而且TMD的化学特性决定了很难往它们掺杂别的物质来制造N型和P型晶体管。

｜ 原子时代

这些问题或许并非解决不了。（在行业发展初期，硅也曾遇到过掺杂上的问题。）在今年的IEDM上，英特尔展示了一款用TMD制造的反相器。但斯坦福大学的电气工程师埃里克·波普（Eric Pop）认为，要在商业产品中用TMD取代硅还需要相当长一段时间。他表示，在大多数应用场景下，硅仍然“足够好”。

总有一天，再巧妙的技术也无法进一步缩小晶体管的体积（例如，很难想象如何用不到一个原子大小的材料来制造晶体管）。正如摩尔本人在2003年警告的那样：“没有什么指数级增长能够永远持续。”但是，他又对当时与会的工程师们说：“你们的工作就是永远延后那一天的倒来。”自他讲这番话以来的二十年里，芯片制造商在延后上取得了令人钦佩的成绩。而且

他们至少也已蚀刻出了未来20年的发展线路。■



The wisdom of youth

Could newborn neurons reverse Alzheimer's?

Some scientists think so. Others doubt the cells even exist

ONE OF THE first signs of Alzheimer's disease is confusion. Most people can park their car in a different space every morning and find it again in the evening. Those with Alzheimer's find this type of problem much harder. Memories of things they do often, like eating or taking medication, become tangled in their minds.

The ability to distinguish between similar memories depends on a tiny strip of brain tissue called the dentate gyrus. Studies in mice have shown that the dentate gyrus is one of the few bits of the brain to generate new neurons even in adulthood. Those new neurons are thought to help keep similar memories distinct.

Whether something similar happens in humans is less clear. But a clutch of new results described recently at the Society for Neuroscience's conference in Washington, DC, suggest that it might. And if it does, then encouraging the process might offer a new treatment for Alzheimer's disease.

Until the 1960s scientists thought adult brains did not produce new neurons. Then evidence began to emerge of young neurons in the brains of adult rats and mice,—specifically in the olfactory bulb, which processes smell, and the dentate gyrus. These new neurons had developed from neural progenitor cells, a type of undifferentiated neuron akin to a stem cell, in a process scientists call neurogenesis.

And there seems to be a link, at least in mice, between these new neurons and Alzheimer's disease. Mice genetically engineered to have Alzheimer's-like symptoms also have fewer young neurons in their brains. Two studies

published in 2021 and 2022 showed that encouraging neurogenesis could improve memory in such animals. And a paper published in April this year found that boosting neurogenesis also caused other cells known as microglia to begin cleaning up tangles of protein, called amyloid-beta plaques, that are characteristic of Alzheimer's disease.

So far so promising—if you are a mouse. But studying neurogenesis in humans is tricky. MRI scanners cannot watch the growth of individual neurons. Scientists must take samples of brain tissue from cadavers and either stain them with chemicals that bind only to young neurons, or measure the genes that have been expressed within brain cells to look for the hallmarks of youth.

Dozens of such papers have been published. But for every one that has found evidence of neurogenesis, another has failed to do so. The pro-neurogenesis camp thinks the null results are down to poor-quality brain tissue or crude preservation techniques. The antis argue that either humans do not grow these neurons in the dentate gyrus, or if they do, there are too few of them to be useful.

The research presented in Washington strengthens the case for human neurogenesis. Maura Boldrini of Columbia University, Marí a Llorens-Martín of the University of Madrid, Orly Lazarov at the University of Illinois at Chicago and Hongjun Song of the University of Pennsylvania presented evidence of young neurons in the dentate gyruses of humans—though most of their findings have not yet been peer reviewed.

Three of these researchers also studied the brains of healthy older people or people with Alzheimer's disease. Their results suggested that the production of the supposed new neurons slows with age, and falls precipitously in Alzheimer's disease. Previous studies by two of the groups also found that Alzheimer's patients with more of these neurons had better

memories.

But while the new research strengthens the case for human neurogenesis, it is not yet definitive. One group of researchers from Yale University questioned whether the staining technique used in one of the studies was accidentally tagging other brain cells such as astrocytes. They are also concerned that the studies tend to find very few progenitor cells. “Where are all the mothers of these baby neurons?” asked Jon Arellano, one of the sceptics.

There are other wrinkles. Dr Song found a different set of genes expressed in the supposedly young human neurons from those seen in neurons from other animals. And the few young neurons that both Dr Song and Dr Lazarov found in the brains of Alzheimer’s patients looked very different from the same neurons in healthy brains. If the new cells are somehow defective, increasing their number may not help.

But although not all the scientists at the conference were convinced, some drug-makers appear to be. In April, Biomed Industries, a pharmaceutical firm based in California, said that results from an early clinical trial suggested that a new drug that the firm claims increases neurogenesis in mice improved memory in people with moderate Alzheimer’s. If subsequent trials prove to be equally encouraging, that could be further evidence that the neurogenesis theory of Alzheimer’s may indeed have something to it. ■



青春和智慧

新生神经元能逆转阿尔茨海默症吗？

有科学家认为可以。也有人怀疑这种细胞根本不存在【新知】

阿尔茨海默症的初期症状之一是思维混乱。大多数人可以每天早上把车停在不同地点，晚上又能找到。而阿尔茨海默症患者却很难做到。一些日常活动的记忆，比如吃饭或服药，会在他们的大脑中乱作一团。

区分相似记忆的能力取决于一小块称为齿状回的脑组织。对小鼠的研究表明，齿状回是大脑中少数几个即使在成年期也会生长出新神经元的部位之一。据信这些新神经元有助于区分相似的记忆。

在人身上是否也是如此还不太清楚。但近日神经科学学会（Society for Neuroscience）在华盛顿举行的会议上公布的一系列新研究结果表明，这是有可能的。如果确实如此，那么促进这一生长过程可能会为阿尔茨海默症带来新的治疗方法。

在1960年代以前，科学家一直认为成人脑不会产生新的神经元。后来有证据表明，成年大鼠和小鼠的大脑中会出现年轻的神经元，特别是在处理气味的嗅球以及齿状回中。这些新神经元是由神经祖细胞（一种类似于干细胞的未分化神经元）发育而来，科学家将这一过程称为神经发生。

至少在小鼠身上，这些新神经元和阿尔茨海默症之间似乎存在关联。经过基因改造而出现类似阿尔茨海默症状的小鼠，其大脑中新神经元的数量也较少。2021和2022年发表的两项研究表明，促进神经发生可以改善这些动物的记忆力。今年4月发表的一篇论文发现，促进神经发生还能使被称为小神经胶质的其他细胞开始清理蛋白质缠结，这些被称为 β -淀粉样斑块的缠结是阿尔茨海默症的症状表现。

这一切都很有希望——如果你是老鼠的话。但要研究人类的神经发生却很棘手。核磁共振扫描仪无法观察单个神经元的生长。科学家必须从尸体上

采集脑组织样本，要么用只与新生神经元结合的化学物质给它们染色，要么测量脑细胞内表达的基因来寻找年轻神经元的印记。

此类论文已发表了数十篇。但是，每有一篇论文找到神经发生的证据，就有另一篇没有找到。支持神经发生的阵营认为，这些无效结果是由于脑组织的质量太差或保存技术粗糙造成的。反方则认为，要么人类在齿状回中没有生长出这些神经元，要么就是即使生长了，也因为数量太少而没有作用。

在华盛顿会议上发表的研究进一步支持了人类神经发生的论证。哥伦比亚大学的莫拉·博尔德里尼（Maura Boldrini）、马德里大学的玛丽亚·洛伦斯-马丁（María Llorens-Martín）、伊利诺伊大学芝加哥分校的奥利·拉扎罗夫（Orly Lazarov）和宾夕法尼亚大学的宋红军展示了人类齿状回中存在新生神经元的证据——尽管他们的大部分发现尚未经过同行评议。

其中三位研究人员还研究了健康老年人或阿尔茨海默症患者的大脑。他们的研究表明，随着年龄的增长，这些所认为的新生神经元的生成速度会减慢，而在阿尔茨海默症患者身上更是急剧下降。其中两个团队之前的研究还发现，拥有更多这些神经元的阿尔茨海默患者的记忆力更好些。

但是，虽然新研究为人类神经发生提供了更多证据，却仍不足以下定论。耶鲁大学的研究团队质疑其中一项研究中使用的染色技术是否意外地标记了其他脑细胞，例如星形胶质细胞。他们也很关注这些研究往往只能发现极少量祖细胞的问题。“这些幼年神经元的所有母细胞在哪里呢？”怀疑者之一乔恩·阿雷利亚诺（Jon Arellano）问道。

还有其他疑惑。宋红军发现，在这些所认为的人类新生神经元中表达的基因与其他动物的神经元不同。此外，宋红军和拉扎罗夫都发现，阿尔茨海默症患者大脑中的少数新生神经元与健康大脑中的同类神经元看起来也有很大差异。如果这些新细胞存在某种缺陷，那么增加其数量可能也没什么好处。

但是，尽管不是所有与会的科学家都被说服了，一些制药商似乎已经信心

十足。今年4月，加州制药公司Biomed Industries表示，它声称能促进小鼠神经发生的一种新药在一项早期临床试验中改善了中度阿尔茨海默氏症患者的记忆力。如果后续的试验结果同样令人鼓舞，那将进一步证明，阿尔茨海默氏症的神经发生理论可能确有其价值。■



It does what it says on the tin

The Extremely Large Telescope will transform astronomy

It will be the world's biggest optical telescope by far—and a powerful time machine

IT IS THE colours of a sunset that inspire Joseph Anderson, an astronomer at the European Southern Observatory (ESO) in the high Atacama desert, in northern Chile. “They start off very blue and turquoise. And gradually, as it gets more purple, then we’re getting closer to observing the universe.” Once night falls the sky is dominated by the star-spangled curve of the Milky Way. If there were any extra left to give, more than two kilometres above sea level, your correspondent’s breath would have been taken away.

The Atacama is home to more than a dozen different astronomical observatories, and for good reason. It is far from the light pollution of civilisation. The air is thin and dry, which improves what astronomers call the “seeing”. And the desert averages 325 cloudless nights each year. Dr Anderson is standing on top of a mountain called Cerro Paranal, showing off the Very Large Telescope (VLT). The VLT is made up of four individual telescopes, each individually one of the largest in the world, alongside another four much smaller ones. It is Earth’s most productive astronomical facility, yielding more than one scientific paper each day. In 2004 it took the first picture of an extrasolar planet—one that orbits a star other than the Sun—and was the first instrument to track individual stars whipping around the enormous black hole at the centre of the Milky Way.

But it may not hold that title for much longer. An hour’s drive from the VLT, atop Cerro Armazones, a 3,046-metre peak, sits the half-finished bulk of the ELT, or Extremely Large Telescope. (ESO is a fan of quotidian names.) Like so many big projects, the ELT is behind schedule. But when it is finished—in 2028, on current plans, at a cost of €1.5bn (\$1.6bn)—it will be, by far, the biggest optical telescope in the known universe. The result, says Robert de

Rosa, an astronomer at ESO, will be “a step change in what we can do in terms of observational astronomy”.

Optical telescopes use a series of mirrors to capture light from space and redirect it to their various instruments. A bigger mirror can collect more light, which means it can both see dimmer things and resolve them in finer detail than a smaller one. The ELT’s main mirror will have a diameter of 39.3 metres, more than four times that of the VLT’s big telescopes (8.2 metres) and over three times that of the present record-holder, the 10.4-metre Gran Telescopio Canarias (GTC), in the Canary Islands. Since a telescope’s power depends on the area of its mirror, looking only at the diameter understates the difference. The GTC has a collecting area of around 75 square metres. The ELT will boast 978 square metres, a little smaller than four tennis courts (see diagram 1).

That size will be a boon for many branches of astronomy. The ELT will shed light on everything from the role of black holes in shaping the large-scale structure of the universe to how dark matter and dark energy affect the rate at which it is expanding, and even whether the supposed constants of physics really are constant over vast intergalactic distances.

It should also provide a big boost to the study of planets outside the solar system. These days, the existence of most exoplanets is inferred from the effects they have on the light from their parent stars. Taking pictures of them—so-called direct imaging—is rare. Of the roughly 5,500 known exoplanets, scientists have pictures of only around 1% of them.

| *Time and relative dimension in space*

The ELT’s enormous mirror will allow astronomers to separate the faint light of a planet from the overwhelming glare of its star from dozens of light-years away. The result should be a direct-imaging bonanza. And direct imaging will also help reveal the chemical composition of exoplanet

atmospheres, and whether any show signs of potential alien life.

And because telescopes are also time machines, the ELT will allow scientists more insight into what happened shortly after the universe began. Since the speed of light is finite, astronomers see distant objects not as they are today, but as they were when the light that arrives in their instruments first set out. Astronomers are keen to use the ELT to investigate how stars and galaxies formed when the universe was young. Early results from the James Webb Space Telescope, launched in 2021, have already turned up an array of stars and galaxies that seem too old to fit easily into existing theories of universal evolution. The ELT could help resolve that mystery.

Assuming, that is, that everything works. When your correspondent visited, the 80-metre-tall steel dome that will shelter the telescope was still being built. Each segment takes around 20 minutes to lift and slot into place; a further six hours are needed to tighten each of the 200 bolts. Work must finish before night, lest gusts of wind blow a panel loose. Once finished, the entire 5,500-tonne dome will be able to rotate so that the telescope can follow the stars as they move across the sky.

One of the few downsides of doing astronomy in Chile is that the country is prone to earthquakes. The telescope will therefore float atop a thin layer of oil. The oil, in turn, will sit atop hundreds of rubber shock absorbers, with the whole lot built on a 3-metre concrete foundation. That will isolate the dome both from earthquakes and any vibrations from the offices and laboratories next door.

The most impressive parts are the mirrors, of which the ELT will have five. Astronomical mirrors are precise, delicate things. Even the comparatively small mirrors of the big VLT telescopes are so heavy that, if not supported properly, they would shatter under their own weight. Engineers must lift

them with a special harness with 15 hooks when they need cleaning.

The ELT's main mirror is so big that it cannot be made as a single piece. Instead Schott, a German optics firm, will make 798 separate pieces that will act as a single mirror. Each is a slightly curved, 1.5-metre-wide hexagonal slice of high-tech glass ceramic that undergoes almost no thermal expansion. The segments are cast in Germany, polished in France, and then mated with supports produced in the Netherlands before being transported to Chile.

Each is checked to ensure it has survived the trip unscathed. Ricardo Parra, an ELT engineer, likens the process to ringing a bell. Vibrations are induced in the glass, and measurements made by accelerometers in strategic locations. The segments are finished by coating them with several further layers of chemicals, including a 100-nanometre layer of silver that provides the reflectivity. (A nanometre is a billionth of a metre.) That silver is protected from tarnishing by a layer of hard silicon nitride glass. Even so, the ESO thinks each segment will need re-coating every two years.

Getting all 798 segments to work together presents another set of difficulties. To produce a usable image each segment must be held in a precise position, with an accuracy of just tens of nanometres. Each is backed by a system of sensors and motors that can subtly deform the surface of the glass in order to correct for warping due to everything from small temperature variations to the changing angle of gravity as the mirror moves and tilts.

The primary mirror is just the first stop (see diagram 2). Light hitting it will be redirected towards a secondary and tertiary mirror, which are designed to correct various subtle optical defects. At around 4 metres across, each could be an impressive telescope main mirror in its own right.

The job of the fourth mirror is to counteract the vagaries of Earth's atmosphere. The reason stars appear to twinkle when seen from the ground is that the atmosphere is constantly churning. Frédéric Gonté, an instrumentation engineer at ESO, compares the effect to peering into water. "Try to see the ground of the swimming pool, you can see it is moving," he says. "The atmosphere is doing that to us."

Space telescopes avoid this problem by flying above the atmosphere. Ground-based ones can rely instead on a technology called adaptive optics. This involves deforming the surface of a mirror to cancel out the distortions imposed by the air. The technology is not unique to the ELT. Many modern telescopes sport it, including one of the VLT's big telescopes (it is being added to the other three). But the ELT's sheer size makes it more susceptible to atmospheric distortion than smaller telescopes. More than 5,000 actuators behind the ELT's fourth mirror will make tiny, rippling adjustments to its shape a thousand times each second. Without the adjustments, the ELT's images would be hopelessly blurred.

Working out exactly how the mirror must be deformed, millisecond by millisecond, requires the presence in the sky of something whose shape is known in advance. Comparing what the telescope actually sees with what it should see reveals the state of the atmosphere at that particular moment, allowing the system to counteract it. Often the object in question is a bright star near the object being studied. If no convenient star is available, though, astronomers can create an artificial one. "Laser guide stars" are made by firing four bright orange laser beams upwards so that they converge in a single point around 90 kilometres up, above the atmosphere's thickest layers. Because the system knows exactly what the ersatz star should look like, it can make whatever mirror-twisting adjustments are needed.

You might think that once the ELT is up and running, all other telescopes will be rendered obsolete. That is not really true, for even a machine such as

the ELT cannot do everything. The twin Keck telescopes in Hawaii, for example, once the world's largest, have mirrors that are a comparatively puny ten metres across. But they have the advantage of sitting on a substantially taller mountain, where the seeing is even better than it is in Chile. And the fact that there are two of them means they can serve twice as many astronomers at once.

The VLT, and other multi-mirror telescopes, can also use a technique called interferometry, a clever way of combining signals such that resolving power depends not on the size of individual mirrors, but on the distance between them. For the VLT that is more than 100 metres. On the other hand, that resolving power comes at the cost of a narrower field of view. The ELT is not competing with telescopes like the VLT, says Dr Gonté. "It's completing."

| *Ain't no replacement for displacement*

But when it comes to detecting the dimmest and most distant objects, there is no substitute for sheer light-gathering size. On that front the ELT looks like being the final word for the foreseeable future. A planned successor, the "Overwhelmingly Large Telescope", would have sported a 100-metre mirror. But it was shelved in the 2000s on grounds of complexity and cost. The Giant Magellan Telescope is currently being built several hundred kilometres south of the ELT on land owned by the Carnegie Institution for Science, an American non-profit, and is due to see its first light some time in the 2030s. It will combine seven big mirrors into one giant one with an effective diameter of 25.4 metres. Even so, it will have only around a third the light-gathering capacity of the ELT.

A consortium of scientists from America, Canada, India and Japan, meanwhile, has been trying to build a mega-telescope on Hawaii. The Thirty Meter Telescope would, as its name suggests, be a giant—though still smaller than the ELT. But it is unclear when, or even if, it will be finished.

Construction has been halted by arguments about Mauna Kea, the mountain on which it is to be built, which is seen as sacred by some. For the next several decades, it seems, anyone wanting access to the biggest telescope money can buy will have to make their way to northern Chile. ■



名副其实

极大望远镜将让天文学改观

它将是世界上遥遥领先的第一大光学望远镜，也是一台强大的时间机器【深度】

智利北部高海拔的阿塔卡马沙漠（Atacama）上日落降临，天空中变幻的色彩引领着欧洲南方天文台（European Southern Observatory，以下简称ESO）驻智利天文学家约瑟夫·安德森（Joseph Anderson）。“晚霞一开始是蓝绿交映，渐渐变得越来越紫，这时我们差不多可以开始观察宇宙了。”夜幕降临后，璀璨的银河横贯中天。海拔2000多米的此时此刻令笔者叹为观止。

在阿塔卡马沙漠有十多个不同的天文观测站，这是有充分理由的。这里远离人类文明的光污染。空气稀薄而干燥，可以提高天文学家所说的“视宁度”。阿塔卡马沙漠平均每年有325个夜晚晴空万里。安德森站在帕拉纳尔山（Cerro Paranal）山顶，向记者展示着“甚大望远镜”（Very Large Telescope，以下简称VLT）。VLT由四个独立的望远镜组成，每一个都是世界上最大的望远镜之一，旁边还有另外四个小得多的望远镜。VLT是地球上产出成果最多的天文设施，每天产出一篇以上的科学论文。2004年，它拍下了第一张太阳系外行星的照片（即一颗围绕太阳以外的恒星运行的行星），并且首次观测到围绕银河系中心巨大黑洞旋转的恒星。

但VLT可能没法保持世界最大望远镜的头衔太久了。在距离VLT一个小时车程、海拔3046米的阿玛索内斯山（Cerro Armazones）的山顶上，屹立着尚未完工的极大望远镜（Extremely Large Telescope，以下简称ELT）。

（ESO喜欢毫无新意的名字。）和太多大型项目一样，ELT的施工进度已经延误。但等建成后（按目前的计划将于2028年竣工，造价15亿欧元[16亿美元]），它将成为已知宇宙中遥遥领先的第一大光学望远镜。ESO的天文学家罗伯特·德·罗莎（Robert de Rosa）表示，这将使“观测天文学的能力向前迈进一大步”。

光学望远镜通过一组镜片收集来自太空的光，并将其投射到各种仪器上。

口径更大的镜面可以收集更多的光，意味着相比口径较小的镜面，它可以看到更暗的物体，成像也更清晰。ELT的主镜口径将达到39.3米，是VLT那几个大型望远镜（8.2米）的四倍多，也是加那利群岛（Canary Islands）上加那利大型望远镜（Gran Telescopio Canarias，以下简称GTC）的三倍多，口径10.4米的GTC是目前世界上最大的望远镜。由于望远镜的探测能力取决于其镜面面积，因此只看口径会低估它们之间的能力差异。GTC的镜面集光面积约为75平方米，而ELT将达到978平方米，略小于四个网球场的面积之和（见图1）。

这么大的面积对于天文学的许多分支来说都是福音。从黑洞在塑造宇宙大尺度结构中的作用，到暗物质和暗能量如何影响宇宙膨胀的速度，乃至我们所认为的物理常数放在广阔的星系际距离中是否真的是恒定的，ELT将帮助解答各种谜题。

它还将大大推动对太阳系外行星的研究。如今，大多数系外行星的存在都是根据它们对其母星发出的光的影响来推断的，很少能拍到它们本身的照片（所谓直接成像）。在大约5500颗已知系外行星中，科学家只有其中约1%的行星的照片。

| 时间和空间的相对维度

ELT的巨大镜面将使天文学家能够将数十光年之外的行星的微弱光芒与其恒星的强烈眩光区分开。这样就应该可以对系外行星大量直接成像。直接成像还将有助于揭示系外行星大气的化学成分，以及是否有任何行星存在外星生命的迹象。

由于望远镜也是时间机器，ELT将能让科学家更深入了解宇宙刚形成后不久的情况。由于光速是有限的，天文学家看到的遥远物体并不是当天的样子，而是望远镜收集到的光线从这些星体发出时的样子。天文学家热切期盼着利用ELT来研究在宇宙年轻时恒星和星系是如何形成的。2021年发射的詹姆斯·韦伯太空望远镜（James Webb Space Telescope）的初期结果已经发现了一系列恒星和星系，它们都似乎非常古老，不太符合现有的宇宙演化理论。ELT可以帮助天文学家解开这个谜团。

这一切的前提是ELT能按计划竣工。在笔者采访时，用于保护望远镜的80米高钢制圆顶仍在建造中。每一块钢板需要约20分钟才能吊起并安装到位，还需要6小时一个个拧紧200个螺栓。每天的工作必须在夜晚之前完成，以免阵风将钢板吹松。完工后，这个5500吨重的圆顶将能够整体旋转，让望远镜得以追踪在天空中移动的恒星。

在智利建设天文台的少数缺点之一是该国容易发生地震。因此，望远镜将浮在一层薄薄的油膜之上，油膜之下放置了数百个橡胶减震器，再往下是3米高的混凝土基座。这将使圆顶免受地震以及隔壁办公区和实验室产生的任何振动的影响。

ELT最令人印象深刻的部分是五个反射镜。天文镜片精密易碎。即使是VLT的那些大型望远镜中相对较小的子镜也很重，如果支撑不当，就会在自重作用下破碎。需要清洁时，工程师必须用带有15个挂钩的特殊安全带将它们吊起。

ELT的主镜太大，无法做成一整张镜面。德国光学公司肖特（Schott）将制造798个独立子镜组成一面主镜。每面子镜都是对角线1.5米的六角形，有一点曲度，用几乎不会热膨胀的高科技玻璃陶瓷制成。这些子镜的镜坯在德国制作，在法国抛光，与荷兰生产的支架组装好后运往智利。

每一面子镜都会经过检查，以确保在运输途中完好无损。ELT的工程师里卡多·帕拉（Ricardo Parra）将这个过程比作敲钟。在玻璃中引起振动，用放置在关键位置的加速度计测量振动幅度。每一面子镜都会再涂上几层化学物质，其中包括提供反射能力的100纳米厚的银。（一纳米为十亿分之一米。）最后再加一层坚硬的氮化硅玻璃以防止银失去光泽。即便如此，ESO认为每面子镜每两年就需要重新做涂层。

要让全部798面子镜协同工作还带来了另一类困难。为了生成可用的图像，每面子镜所在的位置必须十分精确，精度公差仅为数十纳米。每面子镜都由一套传感器和电机支持，它们可以让玻璃表面微微变形，以纠正从微小的温度变化到主镜移动和倾斜时重力角度变化等各种原因造成的翘

曲。

主镜只是光线到达的第一站（见图2）。照射到主镜上的光线将被折射到第二块和第三块副镜上，它们的作用是纠正各种细微的光学缺陷。每块副镜的口径约4米，本身就可以充当一个较大型望远镜的主镜。

第四块副镜的作用是抵消地球大气层的干扰。之所以从地面上会看到星星在闪烁就是大气湍流的影响。ESO的仪表工程师弗雷德里克·冈特（Frédéric Gonté）将这种效果比作向水中凝望。“试着看看游泳池的池底，你会看到它在动。”他说，“大气层就给我们造成了这种幻象。”

空间望远镜通过在大气层上方飞行来避免这个问题。地面望远镜则可以依靠一种称为自适应光学技术，即通过让镜面变形以抵消大气造成的扭曲。该技术并非ELT独有。许多现代望远镜都有自适应光学系统，比如VLT的一个独立大型望远镜（其他三个望远镜也正在配备该系统）。但ELT非常庞大，比小型望远镜更容易受到大气湍流造成扭曲的影响。ELT的第四块副镜后面有5000多个促动器，每秒都会对其形状进行一千次联动式微调。如果不做这种调整，ELT的成像将会极其模糊。

要精确计算出每一毫秒间镜面要如何调整，需要天空中有某个形状预先已知的物体。将望远镜实际看到的这个物体的成像与应该看到的成像比较，就可以揭示这一时刻的大气状态，让系统能够抵消它的光学扭曲。通常，这样的物体是正在观测的物体附近一颗明亮的恒星。不过，如果附近没有恒星方便使用，天文学家可以创造一颗人造恒星。向天空发射四束明亮的橙色激光束，让它们在距离地面约90公里处汇聚到一个点上，就形成了“激光导星”，这个高度已经高过了大气层最稠密的部分。因为系统确切地知道这颗人造星应该是什么样子，所以它可以随时根据需要调整镜面的曲率。

你可能会认为，一旦ELT投入使用，所有其他望远镜就都过时了。但事实并非全然如此，因为即使是像ELT这样的仪器也无法完成所有任务。例如，夏威夷的那两台凯克望远镜（Keck）曾经是世界上最大的望远镜，其

镜面口径相对较小，只有十米。但它们的优势是所在的山峰要高得多，那里的视宁度甚至比智利还要好。而且有两台望远镜，就可以同时为两倍之多的天文学家提供服务。

VLT和其他多镜面望远镜还可以使用“干涉测量法”，这是一种组合信号的巧妙方法，让分辨率不取决于单个子镜的大小，而是取决于它们之间的距离。对于VLT而言，这就让它的组合等效口径超过了100米。另一方面，高分辨率是以视野变窄为代价的。冈特表示，ELT并不与VLT等望远镜竞争，“而是它们的补充。”

| 无可替代

但说到观测最暗和最远的物体，巨大尺寸的聚光镜面是无可替代的。在这方面，ELT在可预见的未来似乎将无与伦比。计划中的下一个更大的“绝大望远镜”（Overwhelmingly Large Telescope）口径将达到100米。但由于复杂度和成本原因，它的建造计划在2000年代搁置。在ELT以南数百公里处，巨型麦哲伦望远镜（Giant Magellan Telescope）目前正在美国非营利组织卡内基科学研究所（Carnegie Institution for Science）拥有的土地上建造，预计将在2030年代的某个时候收集到第一束光。它将把七面大尺寸子镜组合成一面有效口径为25.4米的巨大主镜。即便如此，它的聚光能力也只有ELT的三分之一左右。

与此同时，一个成员来自美国、加拿大、印度和日本的科学家联盟直力图在夏威夷建造一台巨型望远镜，取名三十米望远镜（Thirty Meter Telescope，以下简称TMT）。听名字就知道它的尺寸很大，但仍然比ELT小。不过目前还不清楚它何时能建成，甚至是否能建成。由于TMT选址在被一些人视为圣山的莫纳克亚山（Mauna Kea），相关争议已导致该项目暂停。看来在接下来的几十年里，想要使用花钱就能建成的最大的望远镜，就只能前往智利北部了。■



Baby Berkshire

How a Canadian company became the world's best acquirer of tech firms

Meet Constellation Software, tech's Berkshire Hathaway

FOR OLDER startups these are tough times. The weak recent stockmarket debuts of Arm, a British chipmaker, Instacart, a grocery-delivery group, and Klaviyo, a software firm, have dampened enthusiasm for initial public offerings. Venture capital (VC) has dried up. Data from PitchBook, a research firm, show that late-stage startups need almost three times as much money as is available to them. Many are putting themselves up for sale. Acquisitions of private firms valued at \$100m or more are at their highest since September 2022.

One happy buyer is Constellation Software. The Canadian firm's targets must have sales of at least \$5m and show consistent revenue and profit growth. A strong management team, preferably founder-led, is a plus. Though it has splurged on larger deals, the median value of firms it acquires is around \$3m. According to Royal Bank of Canada (RBC), since 2005 Constellation has spent \$8.7bn on more than 860 firms (see chart 1). In that time its revenue has grown by about 25% a year on average. This year it could exceed \$8bn. The company's market value is up by a big-tech-like 250% in the past five years, to \$50bn, outperforming the tech-heavy NASDAQ index (see chart 2). It is now Canada's second-largest tech firm after Shopify, an e-commerce platform.

Whether by fluke or design, Constellation's dealmaking success is based on principles that look strikingly similar to those of the world's heavyweight acquirer, Berkshire Hathaway. Like Warren Buffett, Berkshire's boss, and his right-hand man, Charlie Munger (who died on November 28th), the founder and president of Constellation, Mark Leonard, seeks out

businesses with a lasting competitive edge. In Constellation's universe, such a "moat" is enjoyed by software firms that specialise in building digital wares for unsexy industries from car dealerships and builders to spas. Tech giants shun these relatively piddling markets and smaller rivals lack the requisite know-how. The result is rich profits for the incumbents.

After a deal is done Constellation, much like Berkshire, runs the business with benevolent neglect. It does not integrate newly acquired companies or parachute in fresh managers. It is content to leave day-to-day operations to the existing leadership. It does not desperately try to squeeze out inefficiencies by centralising common business functions. Constellation believes that splitting a business weakens its bond with customers, notes Paul Treiber of RBC. Cash from the subsidiaries flows to the parent company, which uses it to buy new businesses. These in turn generate more cash, and so on.

To manage over 800 firms, Constellation is structured as a holding company with six large operating groups. Businesses in similar markets are grouped together. In 2021 Constellation floated Topicus, an operating entity that generated 14% of the firm's total revenue and is now valued at \$5.8bn. When any of the other five big operating units get large enough, they, too, may be listed. As with Topicus, Constellation would retain control of the board.

Like Berkshire but in contrast to private-equity or VC funds, Constellation has no exit clock ticking. It can thus be patient with investment decisions. Mr Leonard's annual letters to investors echo Mr Buffett's in describing the company as a "good perpetual owner". This marathon mentality shapes employee pay. Bonuses are tied to returns on invested capital rather than just revenue growth. Executives must invest three-quarters of their bonus in company stock, which they cannot sell for four years. This aligns management's incentives with those of shareholders.

Constellation's success reveals an important truth about mergers and acquisitions that would also be familiar to Mr Buffett: serial acquirers tend to outdo occasional dealmakers. Tobias Lundberg of McKinsey, a consultancy, calculates that regular buyers on average generate about two percentage points more in excess total returns to shareholders annually compared with irregular ones.

Mr Lundberg puts this edge down to practice. As with exercise, the more buying a company does, the better it gets. A few firms like Tyler Technologies from Texas and Roper Technologies from Florida are trying to emulate Constellation's workout regime of picking up niche software makers. None has so far come close to matching the Canadian company's muscle. ■



小伯克希尔

一家加拿大公司如何成为全球最佳科技公司收购者

来认识下星座软件，科技界的伯克希尔·哈撒韦

对于老牌创业公司来说，这是个艰难的时期。英国芯片制造商安谋（Arm）、杂货配送公司Instacart和软件公司Klaviyo近期挂牌上市的表现不佳，打击了市场对IPO的热情。风险投资（VC）资金已经枯竭。研究公司PitchBook的数据显示，处于后期阶段的创业公司所需资金几乎是可获得资金的三倍。许多公司正在寻求出售。被以1亿美元或更高的价格收购的私人公司数量达到了自2022年9月以来的最高峰。

其中一个心满意足的买家是“星座软件”（Constellation Software）。要成为这家加拿大公司的收购目标，销售额至少要达到500万美元，收入和利润也须显示持续增长。强大的管理团队（最好由创始人领导）是个加分项。尽管该公司在较大规模的交易上花费不菲，但所收购公司的中位数价值约为300万美元。根据加拿大皇家银行（RBC）的数据，自2005年以来，星座软件已斥资87亿美元收购了860多家公司（见图表1）。在此期间，其收入年均增长约25%。今年的收入可能超过80亿美元。在过去五年中，该公司市值以堪比大科技公司的势头猛增250%，达到500亿美元，超过了科技股主导的纳斯达克指数（见图表2）。它现在是加拿大第二大科技公司，仅次于电子商务平台Shopify。

不管是因为运气还是有意为之，星座软件的成功交易所遵循的一些原则看起来与全球重量级收购者伯克希尔·哈撒韦惊人地相似。与伯克希尔的老板沃伦·巴菲特和他的得力助手查理·芒格（于11月28日去世）一样，星座软件的创始人兼总裁马克·莱昂纳德（Mark Leonard）寻觅的也是具有持久竞争优势的企业。星座软件“宇宙”中的软件公司就拥有这样的“护城河”，这些公司专门为汽车经销商、建筑商和水疗中心等不那么光鲜的行业构建数字产品。科技巨头看不上这些相对较小的市场，而较小的竞争对手又缺乏所需的专业知识。结果就是这些既有公司获得了丰厚的利润。

一项交易达成后，星座软件也如伯克希尔一样，以“善意的忽略”来经营业务。它不会整合新收购的公司，也不会空降新的经理人。它放心地让现有领导层负责日常运营，也不会集中共同业务职能以竭力铲除低效。加拿大皇家银行的保罗·特雷伯（Paul Treiber）指出，星座软件认为拆分业务会削弱与客户的联系。子公司的现金流向母公司，母公司再用这些现金收购新企业，而新企业又会产生更多的现金，如此循环。

为了管理800多家公司，星座软件采用了一个控股公司下设六个大型运营集团的结构。在相似的市场里经营的业务被组合在一起。星座软件旗下的Topicus于2021年上市，该经营实体占公司总收入14%，目前估值为58亿美元。其余五大运营单位变得足够壮大时，可能也会上市。和对Topicus一样，星座软件将保留对董事会的控制权。

与伯克希尔类似、但与私募股权基金或风投基金不同的是，星座软件没有既定的退出时间。因此，它可以在投资决策上不急不躁。莱昂纳德在每年的致投资者信中将自己的公司描述为一个“优秀的长期持有者”，与巴菲特的说法如出一辙。这种从长计议的心态也体现在员工薪酬上。奖金不仅仅要看营收增长，还与投资资本回报挂钩。高管必须将四分之三的奖金投资于公司股票，四年内不得出售。这就使管理层的激励机制与股东的利益相一致了。

星座软件的成功揭示了一个关于并购的重要真理，巴菲特对此也不会陌生：连续收购者往往胜过偶尔的交易者。咨询公司麦肯锡的托比亚斯·隆德伯格（Tobias Lundberg）估算，与不常收购的公司相比，经常收购的公司平均每年为股东带来的超额总回报要高出约两个百分点。

隆德伯格将这种优势归结为熟能生巧。正如体育锻炼一样，一家公司收购的次数越多，就越得心应手。来自得克萨斯州的泰勒科技（Tyler Technologies）和来自佛罗里达州的儒博科技（Roper Technologies）等少数公司试图效仿星座软件的“健身”方案，收购利基软件制造商。然而，迄今没有一家公司的块头能与这家加拿大公司相媲美。■