



证券研究报告•行业深度报告

从CHAT-GPT到生成式AI（Generative AI）： 人工智能新范式，重新定义生产力

分析师：武超则

wuchaoze@csc.com.cn

SAC编号：S1440513090003

SFC编号：BEM208

分析师：杨艾莉

yangaili@csc.com.cn

SAC编号：S1440519060002

SFC编号：BQI330

研究助理：杨晓玮

yangxiaowei@csc.com.cn

发布日期：2023年1月29日

- 近期人工智能研究公司OpenAI推出的聊天机器人模型CHAT-GPT不断出圈，据 Semafor援引知情人士报道，微软正商谈以 290 亿美元估值，向 OpenAI 投资 100 亿美元，一切均指向人工智能模型的新范式“生成式AI模型（Generative Model）”。
- 此前的决策式AI模型（Discriminant Model）是根据已有数据进行分析、判断、预测，典型应用为内容的智能推荐（短视频）、自动驾驶等；而生成式AI更强调学习归纳后进行演绎创造，生成全新的内容，本质是对生产力的大幅度提升和创造，已催生了营销、设计、建筑和内容领域的创造性工作，并开始在生命科学、医疗、制造、材料科学、媒体、娱乐、汽车、航空航天进行初步应用，为各个领域带来巨大的生产力提升。

前言：微软，百亿美金级别的选择（1）

- 继2019年向OpenAI注资10亿美元之后，微软与OpenAI的合作进入第三阶段。
- 此前据Semafor援引知情人士报道，微软正商谈以290亿美元估值，向OpenAI投资100亿美元。2019年，微软即开始与OpenAI进行合作。到了2021年，微软向OpenAI投资10亿美金。
- 根据微软对外的声明，在与OpenAI新的合作阶段中，微软将有以下计划
 - **Supercomputing at scale:** 微软将增加对专业超算系统开发和部署的投资，以加速OpenAI突破性的独立AI研究。微软还将继续构建Azure的AI基础结构，以帮助客户在全球范围内构建和部署各自的AI应用程序。
 - **New AI-powered experiences:** 微软将在其消费者和企业产品中部署OpenAI模型，并引入基于OpenAI技术的新型数字体验。包括微软的Azure OpenAI服务，该服务使开发人员能够通过直接访问OpenAI模型来构建尖端的AI应用程序。
 - **Exclusive cloud provider:** 作为OpenAI的独家云供应商，Azure将为OpenAI在研究、产品和API服务中的所有工作负载提供支持。
- 此外据媒体报道，微软还计划将ChatGPT整合进旗下搜索引擎必应（Bing），提高必应在搜索引擎市场的市占率；同时Chat-GPT功能引入Office，用于部分文本的生成和问答。

- OpenAI表示，来自微软的投资将使其能独立研究开发出更安全、有用和强大的AI。
- 根据微软对外的声明，此前其与OpenAI的合作已取得一定成绩，并将继续向前
 - 自2016年以来，微软致力于将Azure打造成世界级AI超级计算机，微软和OpenAI推动了云超级计算技术的前沿，在2020年推出其第一台Top-5超级计算机，随后大规模构建了多个AI超级计算系统。OpenAI使用此基础结构来训练其突破性模型，这些模型现已部署在Azure中，以支持GitHub Copilot、DALL-E 2和ChatGPT。
 - 双方的创新激发了人们的想象力，并将大规模人工智能作为一种强大的通用技术平台引入，这将在个人计算机、互联网、移动设备和云的规模上产生变革性的影响。

目 录

第一章

AI：从判别决策到创造生成

第二章

生成式AI：范式升级，应用多元

第三章

生成式AI应用：内容制作，从辅助人到“替代”人

第四章

生成式AI应用：多行业垂直应用

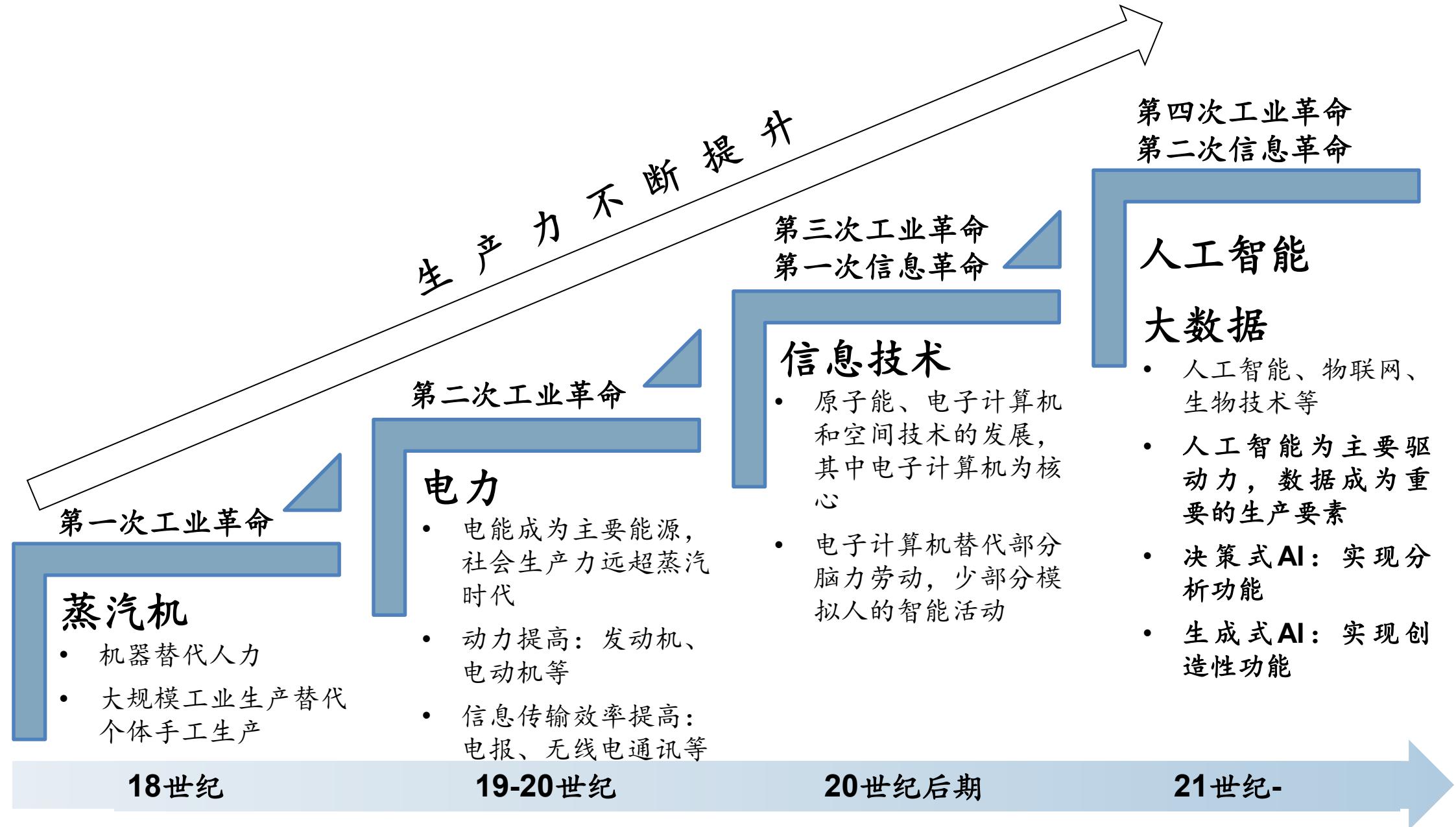
第五章

海内外发展：多家科技巨头、初创公司积极布局

第六章

综合应用：游戏行业的案例

1.1 AI正在推动第四次工业革命，进一步提升生产力



1.2 AI：从决策到生成，AI技术与应用迎来跨越发展

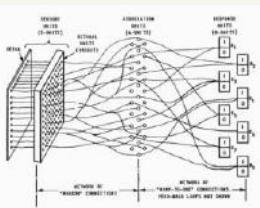
前神经网络 专家系统 概率推论

神经网络 基础模型 大规模模型

2011年以前

决策式/分析式AI时代

1956
感知器模型

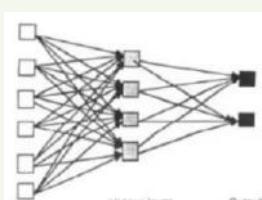


AI学科成立



1980s
Prolog &
Lisp

专家系统
解答特定领域
的问题



1994
GroupLens

首个自动
化推荐系统

推荐算法

iRobot

amazon

机器语言翻译

Google
Translate

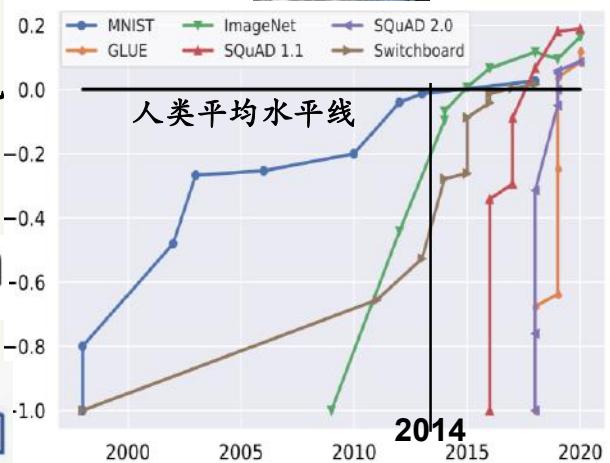
ByteDance
字节跳动

2011
大数据

IBM
沃森
益智问答
打败人类



2012
DNNs
物体识别
自动驾驶



2012年至今

生成式AI急速发展

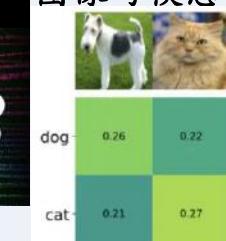
2017
Transformer
网络结构进化
为后续大型模
型打下基础



2020
GPT-3
能够完成
大多数
NLP任务



2021
CLIP
以40亿个文
本-图像对为
训练数据
实现文本到
图像跨模态



2022
DALL·E2
由文本生成
的高清具有
独创性图像

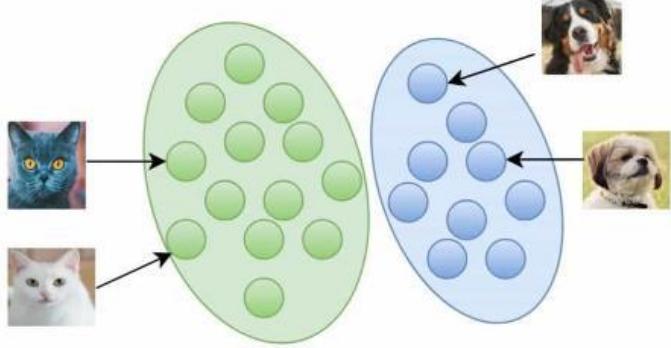
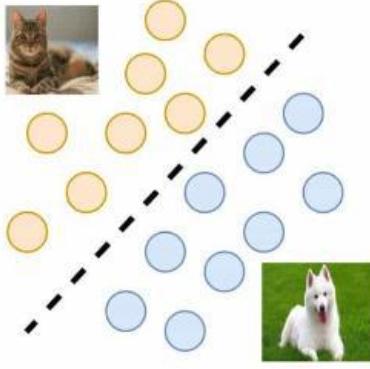


AI的急速发展使得AI模型在手写识别、语音识别、图像识别、阅读理解、语言理解等领域的表现加速超越人类平均水平

1.3 AI：决策式AI与生成式AI对比

- AI模型可大致分为决策式/分析式AI（Discriminant/Analytical AI）和生成式AI（Generative AI）两类。
 - **决策式AI：**学习数据中的条件概率分布，根据已有数据进行分析、判断、预测，主要应用模型有用于推荐系统和风控系统的辅助决策、用于自动驾驶和机器人的决策智能体。
 - **生成式AI：**学习数据中的联合概率分布，并非简单分析已有数据而是学习归纳已有数据后进行演技创造，基于历史进行模仿式、缝合式创作，生成了全新的内容，也能解决判别问题。

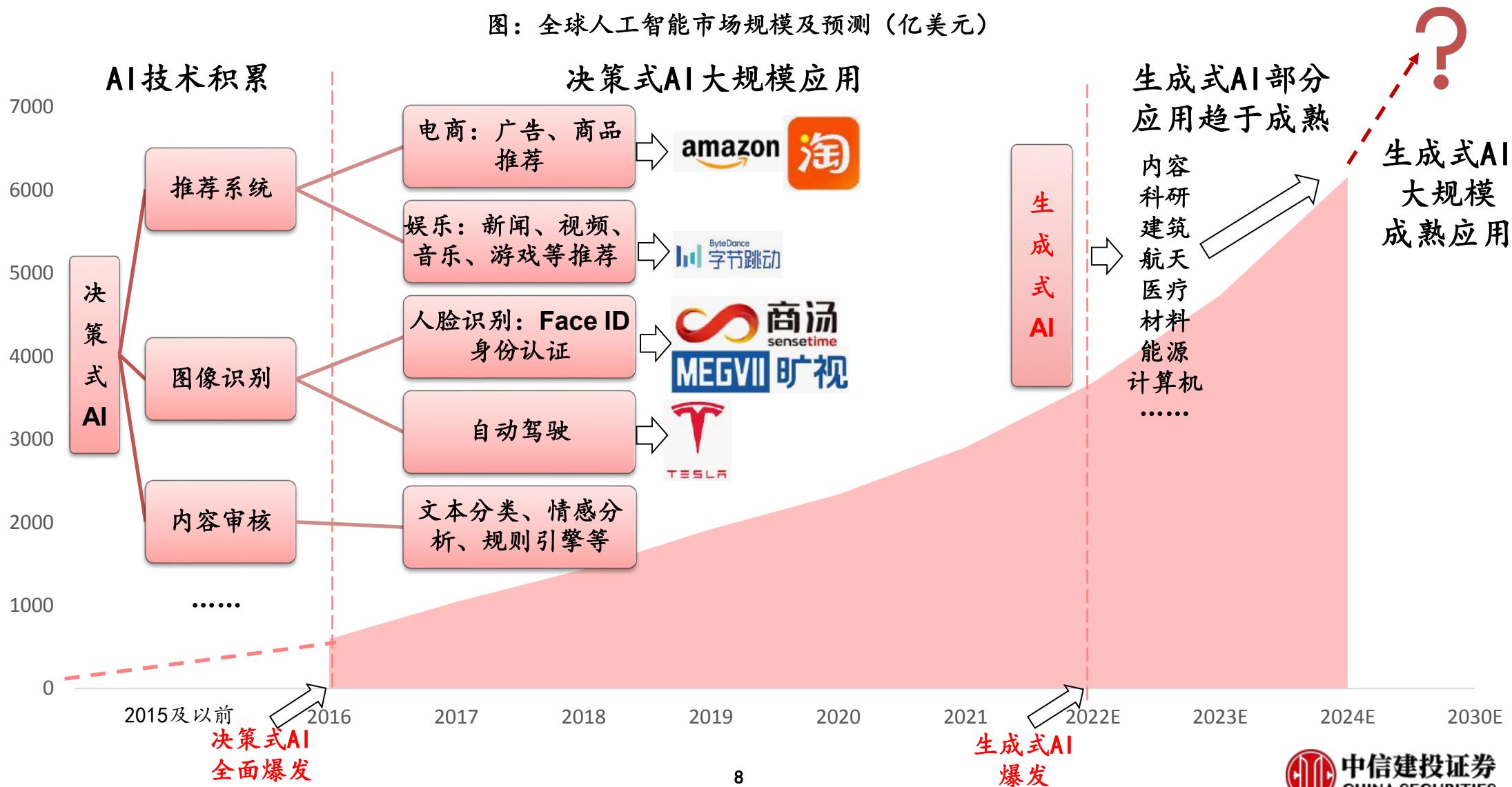
表：决策式AI与生成式AI对比

类型	决策式AI	生成式AI
技术路径	已知数据分别求解输出类别标签，区分不同类型数据，例如将图像区分为猫和狗 	分析归纳已有数据后创作新的内容，例如生成逼真的猫或狗的图像 
成熟程度	技术成熟，应用广泛，辅助提高非创造性工作效率	2014年开始快速发展，近期发展速度呈指数级爆发，部分领域应用落地
应用方向	推荐系统、风控系统、决策智能体等	内容创作、科研、人机交互以及多个工业领域
应用产品	人脸识别、精准广告推送、金融用户评级、智能辅助驾驶等	文案写作、文字转图片、视频智能配音、智能海报生成、视频智能特效、代码生成、语音人机交互、智能医疗诊断等

1.4 决策式AI：推荐系统、图像识别等已创造巨大市场

- 2016年，人工智能技术全面爆发，决策式AI开始大规模应用，包括推荐系统、计算机视觉、自然语言处理等。全球人工智能市场规模从2016年的约600亿美元发展到2021年的近3000亿美元，在推荐系统、计算机视觉、自然语言处理等技术加持下，亚马逊、字节、商汤、特斯拉等公司快速发展。
- 我们认为，生成式AI经过前期技术积累进入爆发期，随着技术与应用的进一步成熟，后续市场空间更为广阔。

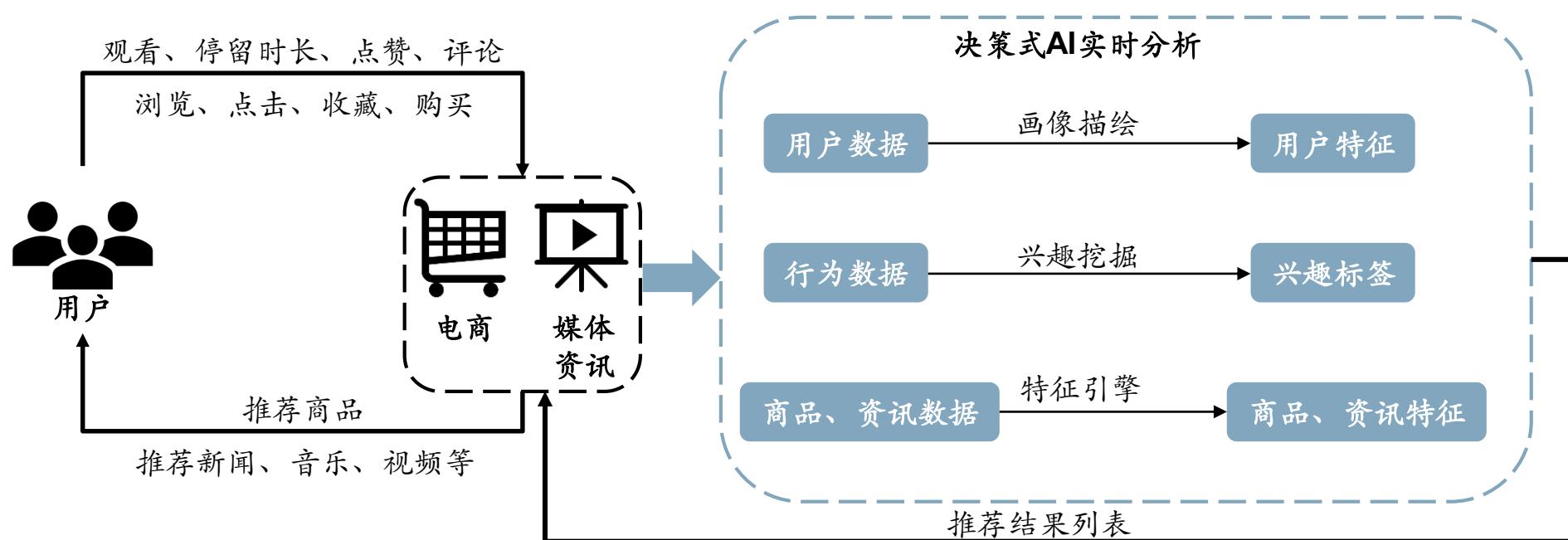
图：全球人工智能市场规模及预测（亿美元）



1.4.1 决策式AI应用：推荐系统助力电商、视频等行业快速发展

- 电商推荐：深度挖掘用户和物品的关联关系，将商品、店铺精准推送给用户。2003年，亚马逊将协同过滤推荐算法应用到电商行业，凭借推荐商品与用户需求精准匹配、大幅降低用户检索用时，推荐算法在电商交易中心逐渐成为资源配置的核心。
 - 推荐系统显著提升电商平台销售额。2006年，据VentureBeat，亚马逊的个性化推荐系统为其贡献了35%以上的销售额；2015年，京东基于大数据个性化推荐算法的订单占到总订单的10%。
- 新闻、音乐、视频等推荐：分析用户长期兴趣和短期兴趣，根据兴趣标签将内容实时推送给用户。推荐算法通过对用户观看、停留、点赞等行为的实时分析，精准刻画用户画像，将用户可能感兴趣的内容推荐给用户，降低人工运营规则的摄入，显著提高了用户粘性。

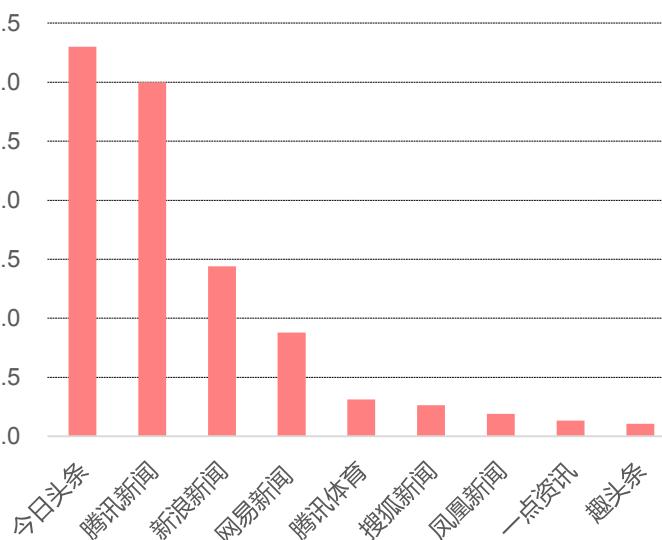
图：推荐系统精准刻画用户画像准确推送



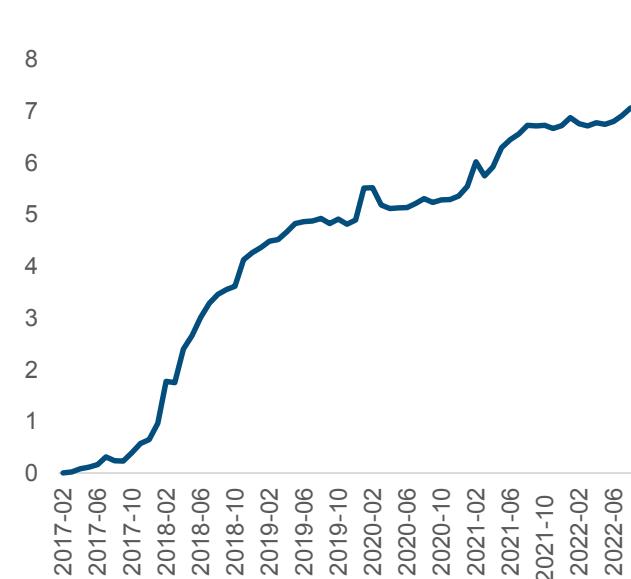
1.4.1 决策式AI应用：字节跳动凭精准推送10年内估值超3000亿美元

- **字节跳动：AI驱动抖音、头条精准推荐内容，10年内估值超3000亿美元。**凭借判别式AI等先进算法对用户画像的学习，抖音、头条等字节系列产品，将内容与广告等高效、精准推荐给用户，快速驱动自身用户数、业绩与估值增长。
 - 用户数量方面，据QuestMobile，抖音2022年9月月活7.1亿，全国移动互联网用户渗透率50%；2021年12月，今日头条月活3.3亿，远高于传统新闻资讯平台；
 - 收入方面，据晚点LatePost及华尔街日报，字节收入由2016年的60亿元，上升到2021年的4391亿元，CAGR达136%；
 - 市值方面，据《创业家》及胡润研究院《2021全球独角兽榜》，字节跳动市值由2014年的5亿美元，上升至2021年的3375亿美元。

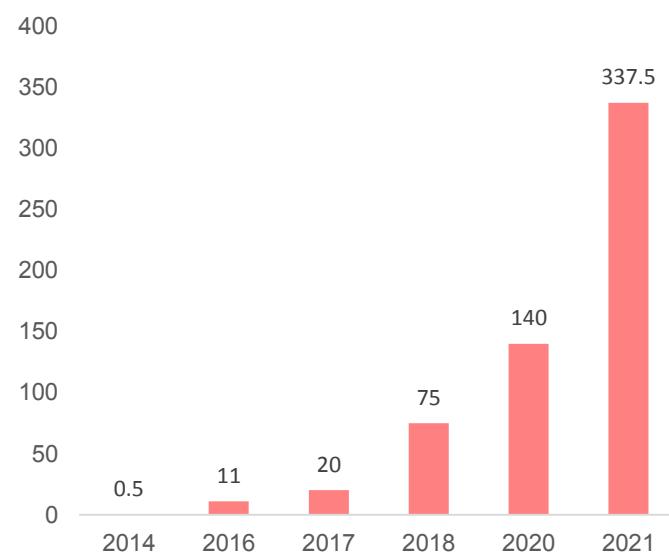
图：2021年底各大新闻资讯平台MAU（亿）



图：抖音主站月活（亿）



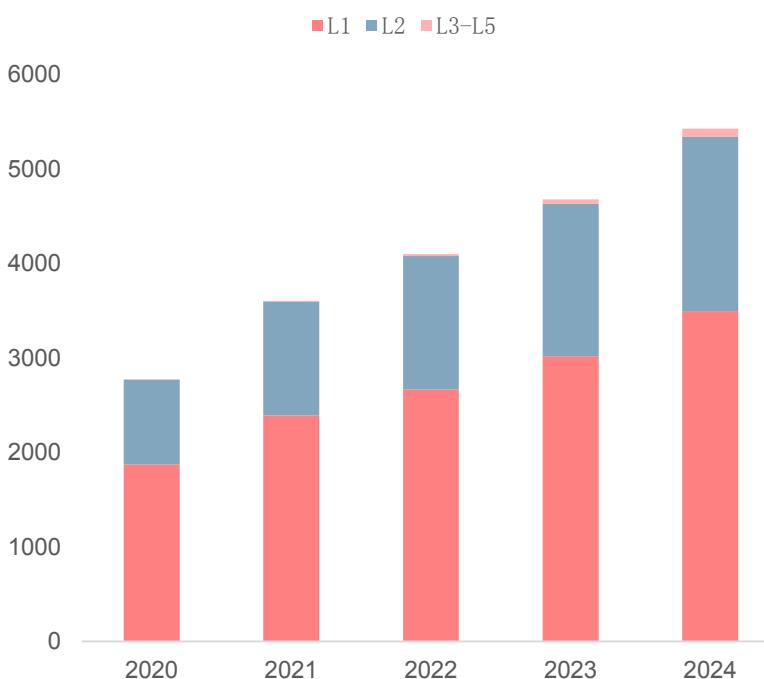
图：字节跳动市值变化（十亿美元）



1.4.2 决策式AI应用：图像识别推动自动驾驶市场快速增长

- ▶ **自动驾驶：**AI智能分析、识别路况，渗透率不断提升。借助决策式AI，自动驾驶汽车可分析判别各种路况，并可对多种物体进行识别与跟踪，从而提高行车的安全性，例如行人、空旷行驶空间、地上标注、红绿灯、车辆等。目前自动驾驶市场规模快速增长，据Research And Markets，2021年全球无人驾驶汽车的市场规模约273亿美元，2026年全球无人驾驶汽车市场规模将达到594亿美元，年复合增长率为16.84%。
- ▶ 带动产业链发展，多环节均有高市值上市公司及高估值公司。判别式AI为自动驾驶汽车提供核心技术支持，自动驾驶汽车的销售则带动整个相关产业链的发展，如车用无线通信技术V2X、高精度地图、激光雷达、芯片、ADAS驾驶辅助系统、智能车解决方案等。多个产业链环节均有高估值、大市值公司诞生，例如蔚来汽车（造车）、商汤科技（激光雷达、V2X、ADAS等）、Momenta（高精度地图）等。

图：全球自动驾驶汽车销量及预测（万辆）



表：自动驾驶产业链上市公司市值（亿美元，截至2022年11月24日）

公司名称	业务	市值
特斯拉	造车;L4自动驾驶	5785亿美元
蔚来汽车	造车;L4自动驾驶	1400亿港元
小鹏汽车	造车;L4自动驾驶	485亿港元
理想汽车	造车;L4自动驾驶	1469亿港元
商汤科技	激光雷达; V2X;ADAS算法	620亿港元
寒武纪	V2X;自动驾驶芯片;ADAS算法	255亿人民币

表：自动驾驶产业链非上市公司估值（亿美元，截至2022年2月）

公司名称	业务	估值	公司名称	业务	估值
哪吒汽车	造车	70亿	蘑菇车联	V2X;L4 自动驾驶	30亿
滴滴自动驾驶	L4 自动驾驶	34亿	地平线	V2X;自动驾驶芯片;ADAS 算法	30亿
文远知行	造车;L4自动驾驶	33亿	奇点汽车	造车;L4自动驾驶	27亿
小马智行	造车;L4 自动驾驶	31亿	酷哇机器人	造车;车联网;高精度地图	25亿
Momenta	高精度地图	31亿	千寻位置	高精度地图	21亿

1.4.2 决策式AI应用：AI四小龙多领域开创“AI+”新业态

- **AI四小龙：**在多场景探索计算机视觉应用。借助判别式AI，商汤科技、旷视科技、依图科技、云从科技对C端、B端、G端的各大应用场景进行智能分析，提高信息的判别分析效率、节省人力成本，并为用户及时提供决策反馈。四家公司均获得了较高的估值。
- **商汤科技：**将人工智能基础设施、智能驾驶、人/物检测、SLAM、图像生成、智能决策为核心，应用领域涵盖智能驾驶、身份验证、城市管理、游戏AI、VR/AR等，11月24日市值620亿港元。
- **旷视科技：**专注于图像/视频的分析与理解，以人脸识别为核心，业务场景包括城市治理、楼宇园区管理、实名认证、通行考勤等，据胡润全球独角兽榜，2021年估值270亿元。
- **依图科技：**聚焦于人脸识别、自然语言处理等领域，解决方案包括城市管理、智慧医疗等据胡润全球独角兽榜，2020年估值140亿元。
- **云从科技：**建立视觉、语言、环境认知系统，应用于智慧交通、金融、教育等领域，11月24日市值131亿元。

表：AI四小龙AI应用场景

类型	应用场景	商汤科技	旷视科技	依图科技	云从科技
C端	手机	人脸3D重建、手机图像处理、手机人脸解锁方案等	手机安全、认证、影像等解决方案		
	汽车	高级辅助驾驶系统、智能座舱	车主身份验证、司机驾驶状态分析		
B端	金融	互联网金融、线下金融、智慧网店、会员识别等	卡证信息录入、身份认证	智能网点、刷脸取款解决方案	智慧网点、智慧识人、票据OCR、风控解决方案
	医疗	智慧诊疗平台		Care.ai医疗智能全栈式产品解决方案	
	教育	AI基础教育、智慧校园	教育行业数字化		
	物流		仓储物流数字化		
	制造		工厂制造数字化		
G端	安防	方舟城市级开放视觉平台、人脸动态比对服务器等产品	城市管理、楼宇园区管理数字化解决方案	城市安全、智慧园区	AI+人像大数据、智慧社区实有人口、智慧园区安防解决方案
	交通	人群分析系统			民航旅客出港服务、航站楼管理服务、民航大数据服务
	零售	智慧商业解决方案	零售行业数字化解决方案	顾客识别、轨迹热区、4D店标准化解决方案	智慧门店综合、智慧4D专营店智慧案场人脸识别解决方案

1.4.3 决策式AI应用：内容审核显著降低人工成本

- 决策式AI通过规则引擎、文本分类、图像分类等进行内容审核，广泛应用于搜索引擎、社交平台、电子商务、新闻网站等领域，以确保内容符合法律法规和公司政策，保护用户隐私和安全，大幅降低了各公司人工审核的成本，并提高了内容质量，建立了用户信任。
 - 规则引擎：通过预定义的规则检查内容是否包含某些敏感词汇以决定是否通过审核。
 - 文本分类：按照一定的体系或标准对文本进行自动分类标记，包括情感分析、意图识别、主体分类等，例如识别评论是积极的还是消极的、将文本按照政治、军事、体育等进行分类等。
 - 图像分类：将图像内容进行分类，例如将图像分入动物、植物、风景、人像等不同的类别。

表：内容审核应用场景

应用领域	名称	应用场景
搜索引擎	谷歌	网页敏感词审核、内容审核、广告审核等
	百度	
社交平台	Facebook	对用户发布的文字、图像、视频的进行审核，屏蔽侮辱、攻击性、暴力、色情等违法违规信息以及虚假信息等
	Twitter	
	微信	
	微博	
电子商务	京东	商品内容审核、商品宣传图审核、违法宣传标语过滤、评论审核、虚假评论及恶意评论的过滤等
	淘宝	
新闻网站	Reuters	智能新闻分类、虚假新闻识别过滤、评论审核等

1.4.4 决策式AI应用：谷歌开发多款模型替代人力分析

- 谷歌深耕人工智能领域，不仅开发出了多个图像识别、人脸识别、语音识别算法，替代了常见的人类能够完成的分析性工作，还推出了天气预测、流行病预测等模型，实现了优于常人的决策结果。

表：谷歌部分决策式AI模型梳理

名称	时间	应用
Building High-level Features Using Large Scale Unsupervised Learning	2012	训练人工神经网络从无标注图像识别特征，例如识别图像中的人脸、猫脸等
TensorFlow	2015	包含一系列机器学习算法，支持多种编程语言，方便用户训练、开发、部署机器学习模型。不仅支持判别式模型，也支持生成式模型
AlphaGo	2015	通过对大量数据学习如何做出围棋决策，击败世界顶尖围棋选手，后续用于金融领域预测股票价格、自动驾驶领域判断路况等
FaceNet	2015	使用深度卷积网络，高精度识别人脸图像之间的差别
Wide & Deep	2016	基于Embedding和多层次感知机的推荐算法，在推荐领域有广泛的应用
Joint Speech Recognition and Speaker Diarization via Sequence Transduction	2018	集成语音识别和说话人分类系统，判断谁在什么时候说了什么
Building Damage Detection in Satellite Imagery Using Convolutional Neural Networks	2019	使用卫星遥感图像分析建筑物的损坏程度，为救灾决策提供帮助
MetNet	2020	通过观测数据预测天气，能够以1km的解析度、两分钟的时间间隔预测未来8小时内的降雨量
A prospective evaluation of AI-augmented epidemiology to forecast COVID-19 in the USA and Japan	2021	模拟政策变化对新冠确诊人数和死亡病例的影响

目 录

第一章

AI：从判别决策到创造生成

第二章

生成式AI：范式升级，应用多元

第三章

生成式AI应用：内容制作，从辅助人到“替代”人

第四章

生成式AI应用：多行业垂直应用

第五章

海内外发展：多家科技巨头、初创公司积极布局

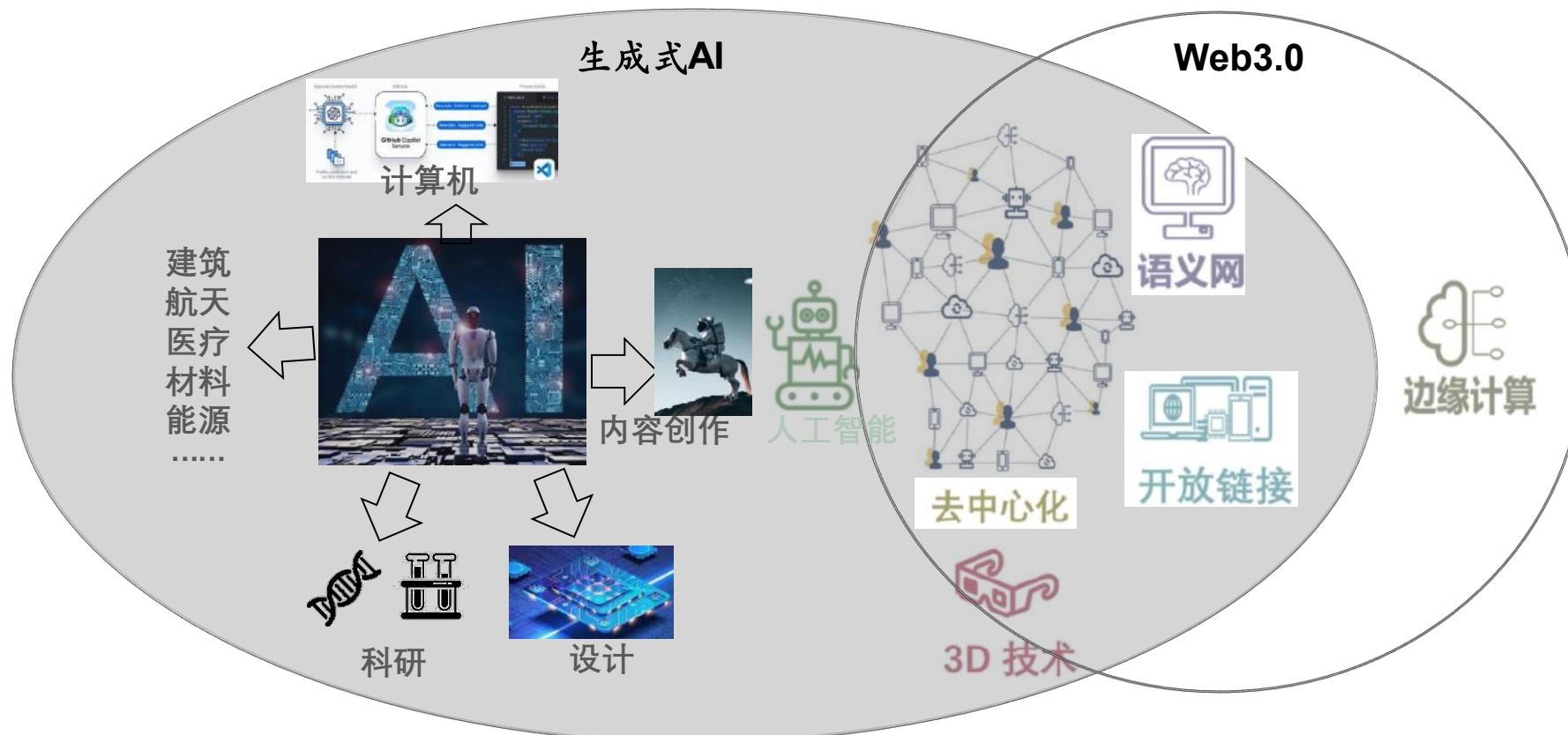
第六章

综合应用：游戏行业的案例

2.1 生成式AI VS Web 3.0：不仅是web3.0的生产工具，更是人工智能的新范式

- 生成式AI可以作为Web3.0的生产工具，但Web3.0中的应用仅是其应用的冰山一角。
- Web3.0围绕去中心化的理念展开，结合区块链、智能合约、加密货币等技术，核心是产生的数据由用户拥有，能够改变用户数据及原创内容等均由互联网中心化实体控制的现状，在创作者经济中取得更好的平衡从而提升用户创作内容的积极性。生成式AI在内容领域的应用能够满足用户不断提升的创作需求，但这仅是其应用的冰山一角。
- 生成式AI在广大垂直领域的应用带来的AI用途的结构性改变以及生产力的进一步提高，未来可能创造巨大的市场价值。

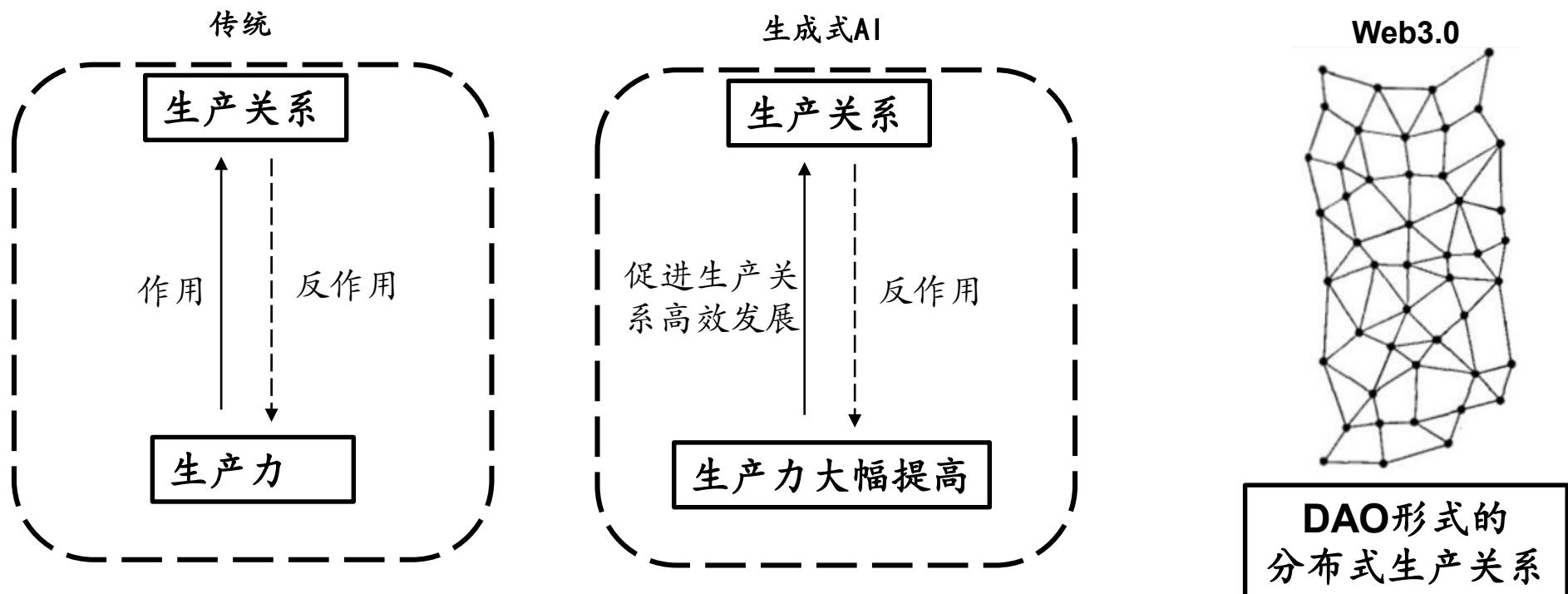
图：生成式AI与Web3.0



2.1 生成式AI VS Web 3.0：不追求生产关系的重塑，但将大幅度提升和创造生产力

- **创造是生成式AI的核心，本质是对生产力的大幅度提升和创造。**生成式AI通过从数据中学习要素，进而生成全新的、原创的内容或产品，不仅能够实现传统AI的分析、判断、决策功能，还能够实现传统AI力所不及的创造性功能。生成式AI已催生了营销、设计、建筑和内容领域的创造性工作，并开始在生命科学、医疗、制造、材料科学、媒体、娱乐、汽车、航空航天进行初步应用，为各个领域带来巨大的生产力提升。
- **生成式AI将促进生产关系高效发展，而不是像Web3.0重塑生产关系。**Web3.0通过底层的分布式与去中心化、密码学的应用、加密货币的结算方式，为集体协作、分工、收益提供了更好的结算方式，去中心化自治组织（Decentralized Autonomous Organization, DAO）将是常见的组织形态，打破了原有的生产关系。生成式AI不是打破传统的生产关系，而是通过提高生产力促进现有生产关系高效发展。

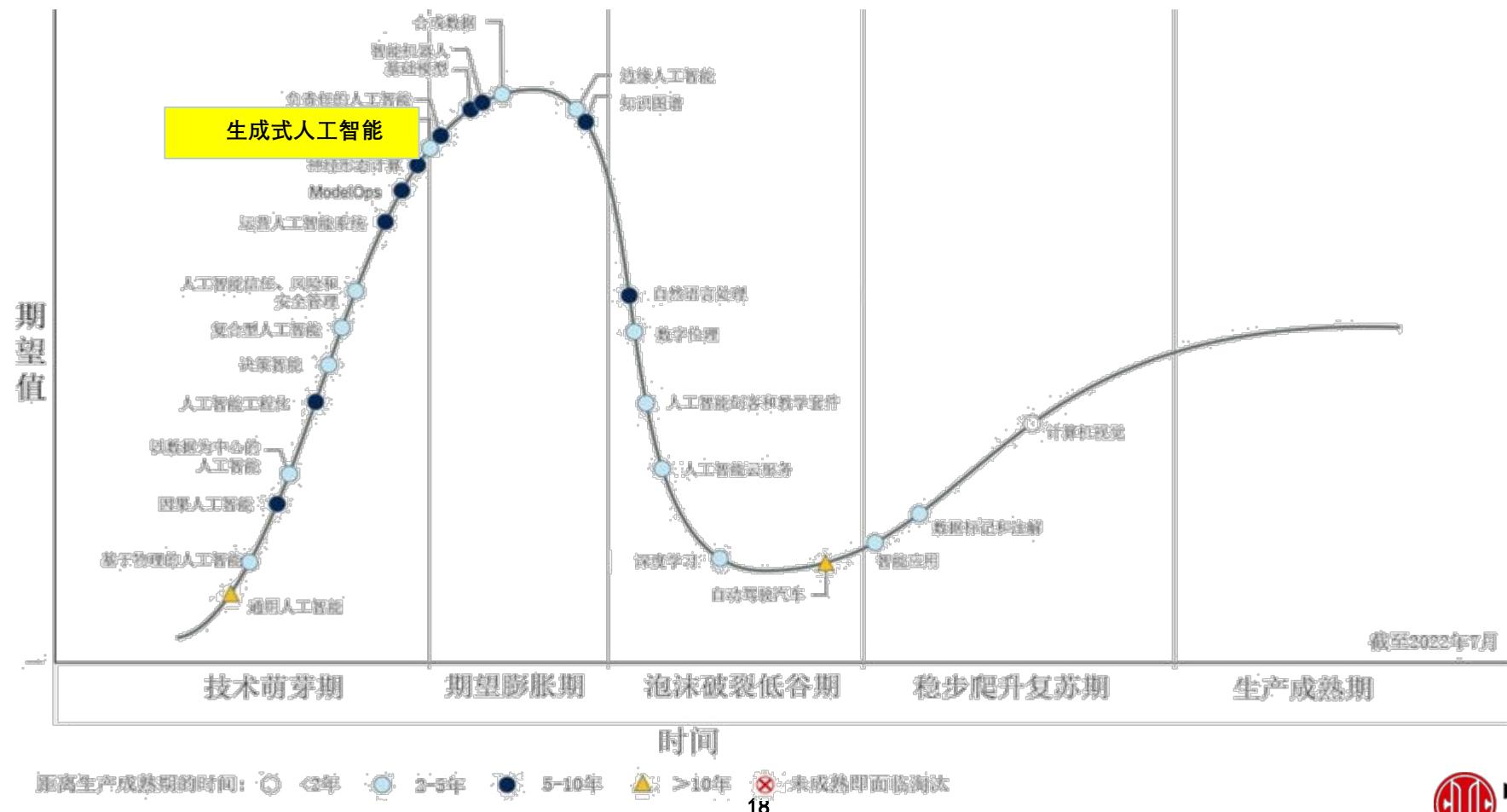
图：生成式AI与Web3.0对生产关系的影响



2.2 生成式AI：“最有商业前景的人工智能技术”

- Gartner将生成式AI列为最有商业前景的人工智能技术。根据其发布的2022年人工智能技术成熟度曲线，预计生成式AI 2-5年内将进入生产成熟期，发展潜力与应用空间巨大：
- 2025年，生成式AI产生的数据将占到所有数据的10%，而2021年生成式AI产生的数据不到所有数据的1%；
 - 2025年，30%的大型组织出站消息将由生成式AI生成；
 - 2025年，50%的药物发现与研发将使用生成式AI；
 - 2027年，30%的制造商将使用生成式AI提高产品研发效率。

图：2022年人工智能技术成熟度曲线



2.3 资本：海外一级投资涌入

➤ 多家知名风投公司看好生成式AI赛道：

- 红杉资本官网9月19日发布的文章《生成式AI：充满创造力的新世界》中提到：“生成式AI有潜力产生数万亿美元的经济价值”。
- Coatue发布的《AI2022：爆发》认为规模化突破使得AI在短时间内变得指数级强大，其应用场景迅速突破。
- 2022年10月，Stability AI完成1.01亿美元融资，估值10亿美元，投资方包括Coatue、Lightspeed Venture Partners和O'Shaughnessy Ventures。公司由前英国对冲基金经理Emad Mostaque于2020年成立。
- 2022年10月，Jasper完成1.25亿美元融资，估值达15亿美元，投资者包括Coatue、Bessemer Venture Partners、IVP等多家机构。
- 2019年，OpenAI获得微软10亿美元投资，2021年OpenAI估值已达200亿美元。

表：风投机构视角下的生成式AI

SEQUOIA

Generative AI: A Creative New World

The fields that generative AI addresses—knowledge work and creative work—comprise billions of workers. Generative AI can make these workers at least 10% more efficient and/or creative: they become not only faster and more efficient, but more capable than before.

Therefore, Generative AI has the potential to generate trillions of dollars of economic value.

COATUE

AI 2022: The Explosion

This incredible AI renaissance is happening simultaneously across many diverse fields and verticals. This is not a coincidence: **scaling breakthroughs have enabled AI to become exponentially more powerful in a very short period of time. Engineers are now taking this technology breakthrough and applying it everywhere they can.**

2.3 资本：海外一级投资涌入

- 多家生成式AI公司进入Madrona、高盛、微软、亚马逊网络服务和 PitchBook联合发布的2022年智能应用前40名榜单（Intelligent Applications 40, IA40）。IA40招募了来自40多家顶级风险投资和投资公司的50多名风险投资人，提名并投票选出塑造智能应用未来的顶级公司，这些公司自成立以来募资超160亿美元，今年募资超过50亿美元，其中包括Runway、Jasper、Copy.ai在内的14家生成式AI相关公司，占比达35%。

表：2022年IA40（生成式AI公司用虚线框出）

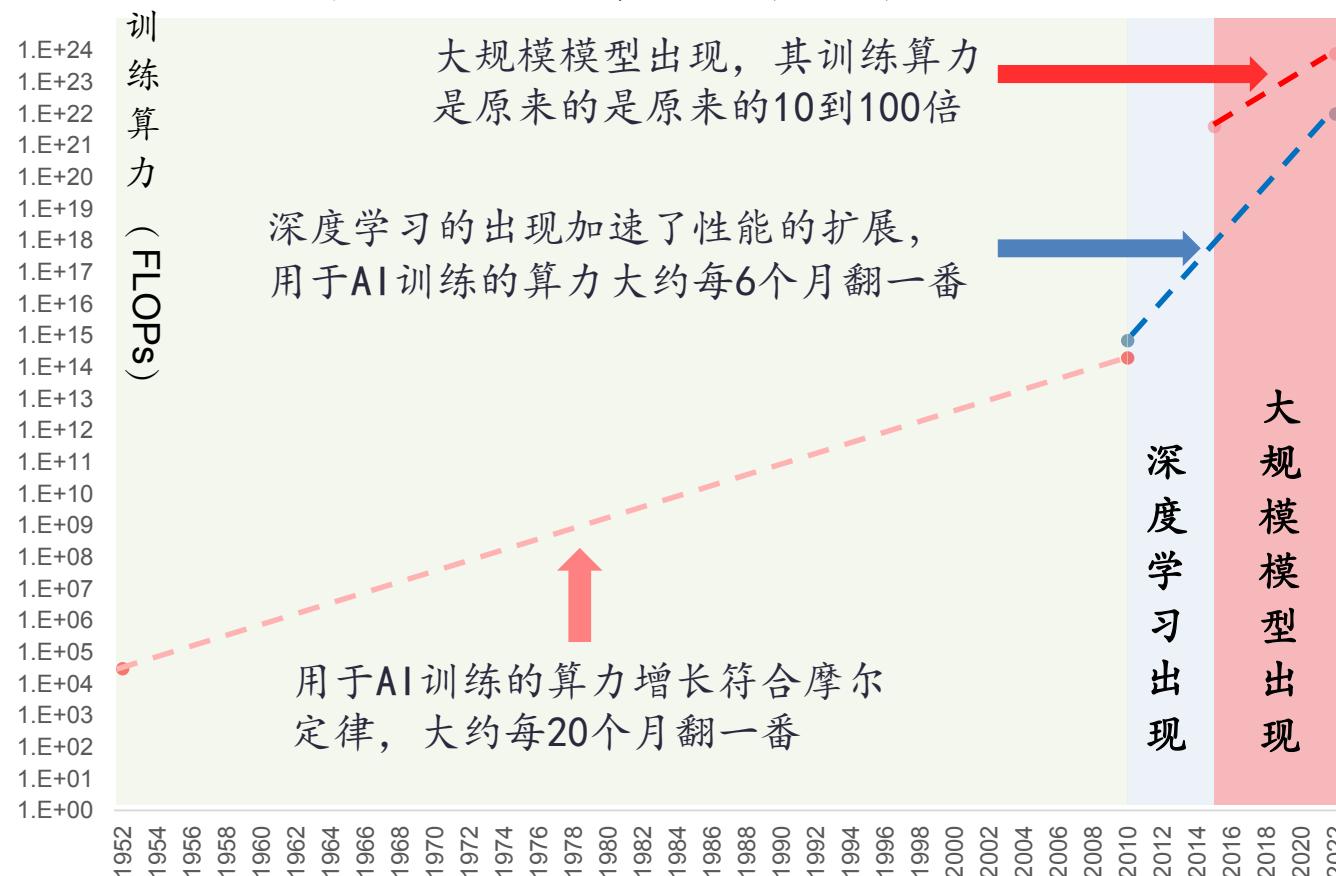
2022 Intelligent Applications 40			
EARLY STAGE	MID STAGE	LATE STAGE	ENABLER
Jasper Credo AI Linear Copy.ai Flux A-Alpha Bio Lexion DoNotPay Tabnine Vanti	Runway Cresta.ai Vanta Loom Panther Adept Gem Glean Jellyfish CaptivateIQ	Snyk Gong Airtable Calendly Abnormal Security Celonis Highspot Grammarly Clari Miro	Hugging Face Databricks dbt Labs Airbyte DataRobot Weights & Biases Anyscale OctoML ThoughtSpot Fivetran

2.4 技术：2022年是拐点

➤ 前期架构、模型、数据、算力的积累与提升

- 架构改进：深度神经网络的学习能力和模型的大小呈正相关，但规模越大训练难度越高，需要对结构进行改进，拥有更强并行性的Transformer架构带来了深度神经网络参数量从最早的几万到目前的数千亿的跃升。
- 模型发展：GPT-3、CLIP、Diffusion、DALL·E2等模型的提出极大提升了AI处理NLP、跨模态、生成问题的能力。
- 数据增多：拥有的优质训练数据越多，算法从中学习的效果越好。随着数字时代的到来，生成数据的工具和软件越来越普遍，数据总量呈指数增长，可供AI训练的数据质量和数量均有很大提升。
- 算力提升：大规模深度学习模型的参数和数据量达到了一定量级，需要相应算力的支撑，目前，大规模模型的训练算力是原来的10到100倍。

图：用于训练机器学习经典算法的算力



资料来源：COMPUTE TRENDS ACROSS THREE ERAS OF MACHINE LEARNING, Google Scholar, 中信建投

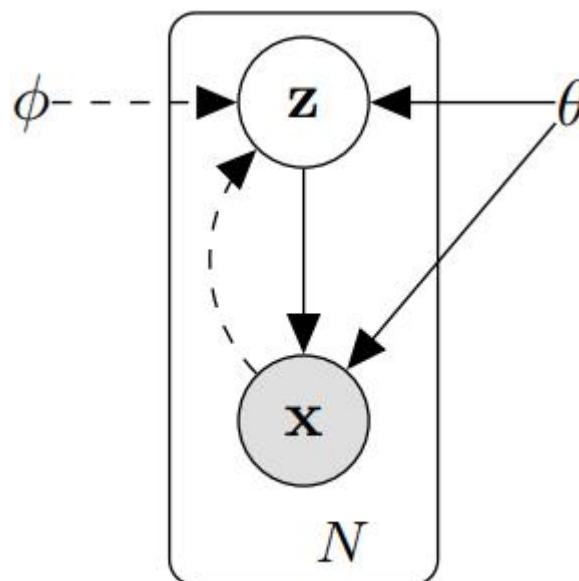
图：生成式AI关键模型/架构



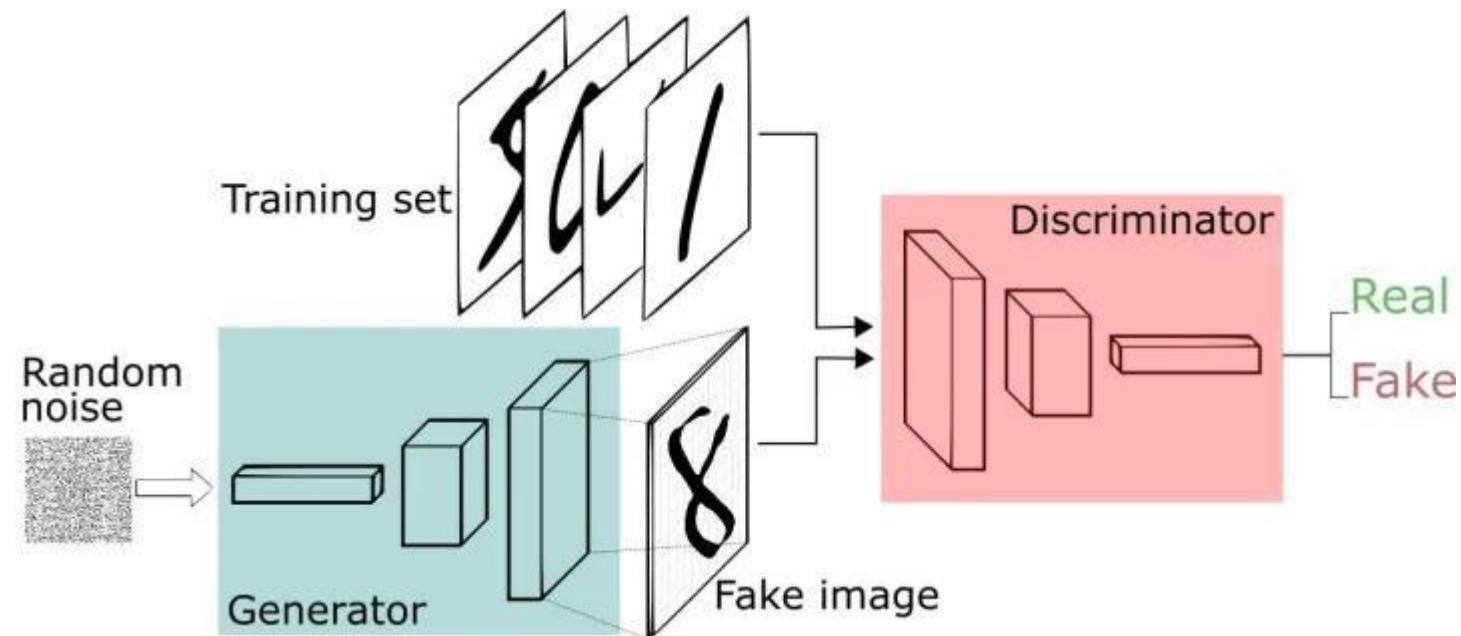
2.5 技术：模型迭代

- 变分自编码器（Variational Autoencoder, VAE）：2013年由Diederik P. Kingma和Max Welling提出，编码器将原始高维输入转换为对潜在空间的概率分布描述，从中采样输入解码器，得到新生成的结果，可用于图像生成、语音合成等，生成的图像较为模糊。
- 生成式对抗网络（Generative Adversarial Nets, GAN）：2014年由Ian J. Goodfellow等人提出，模型由生成器和判别器组成，以图像生成为例，生成器输入训练噪声后生成图像，判别器用来判断图像是真实的还是由生成器生成的，随着训练不断进行，生成器水平提升，判别器不再分辨图像真伪，固定生成器对判别器进行驯良，直到判别器能够分辨图像真伪，固定判别器再次训练生成器，不断循环，获得生成效果好的生成器。GAN模型能够生成图像、3D模型甚至视频，但对输出结果控制较弱，易产生随机结果。

图：VAE模型示意图



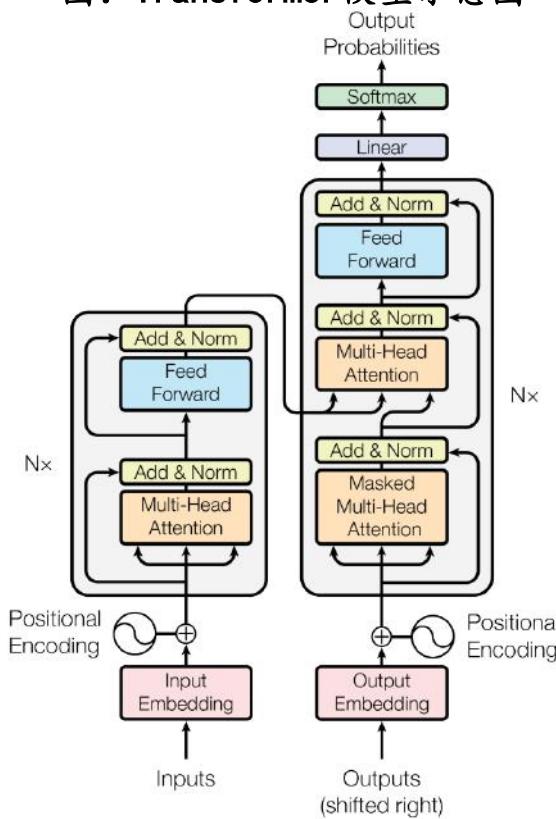
图：GAN模型示意图



2.5 技术：架构升级，能力质变

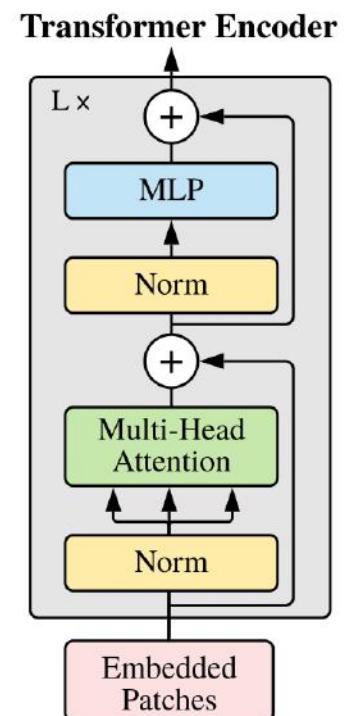
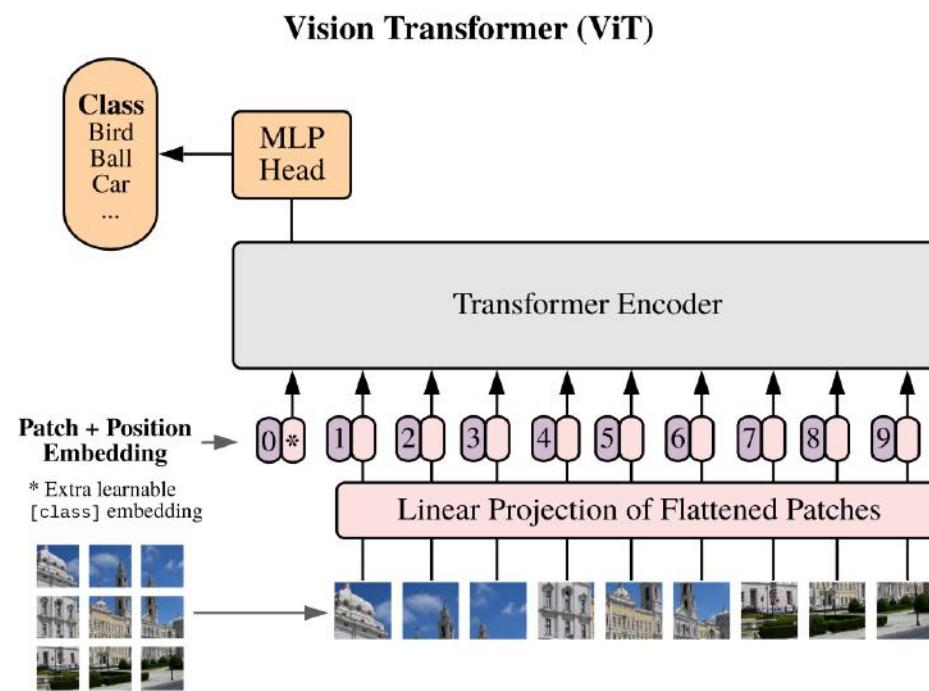
- **Transformer**: 2017年由Google团队提出，采用自注意力机制，按输入数据各部分重要性的不同而分配不同的权重，仅用attention来做特征抽取，网络结构的进化带来了参数量和模型层数的提高，引起了生成式AI技术能力的质变；并行化优势允许其在更大的数据集上进行训练，这也促成了GPT等预训练模型的发展。
- 视觉**Transformer (ViT)** : 2020年由Google团队提出，将Transformer应用在图像分类领域。ViT将输入图片分为 16×16 个patch，再将每个patch投影为固定长度的向量送入Transformer，后续操作与原始Transformer相同。ViT通过将人类先验经验知识引入网络结构设计，获得了更快的收敛速度、更低的计算代价、更多的特征尺度、更强的泛化能力，能够更好地学习和编码数据中蕴含的知识，正在成为视觉领域的基础网络架构。以ViT为代表的视觉大模型赋予了AI感知、理解视觉数据的能力，提升了AI的感知能力。

图：Transformer模型示意图



资料来源：中国信通院，Google Scholar，中信建投

图：ViT模型示意图

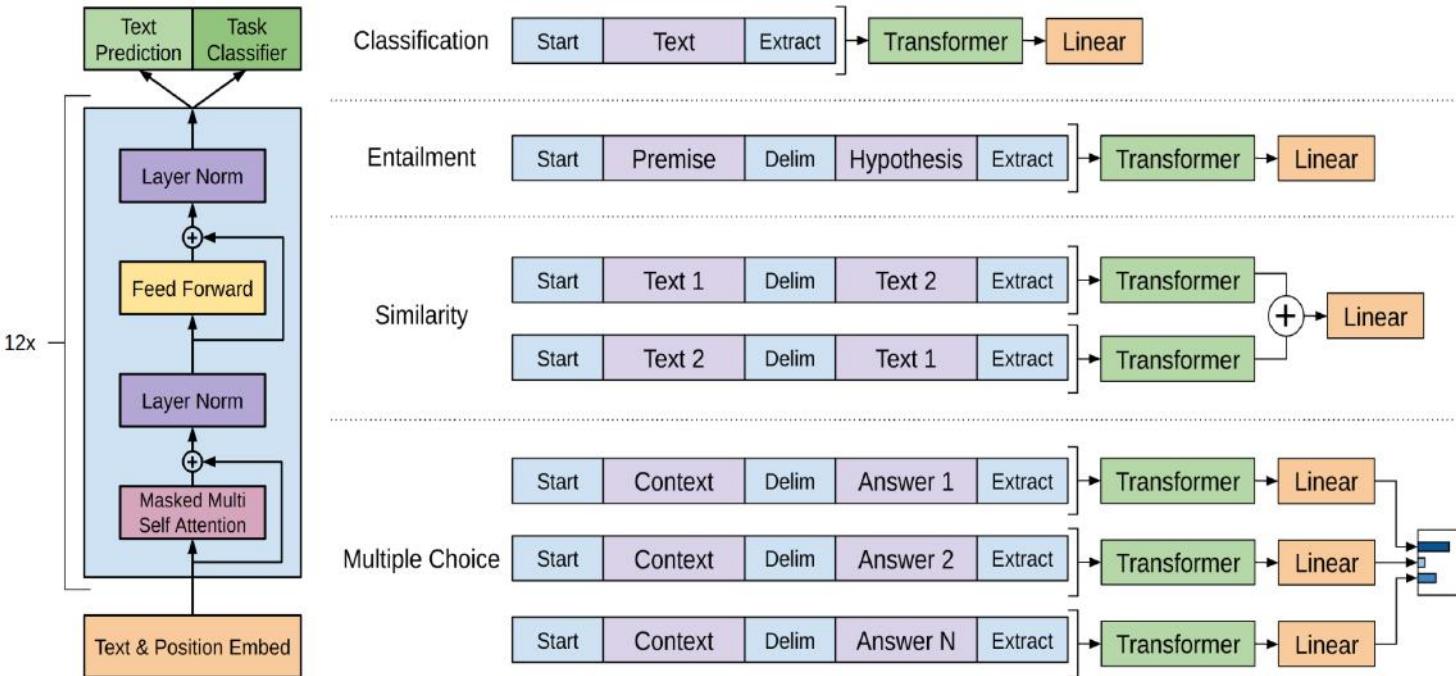


2.5 技术：快速进化

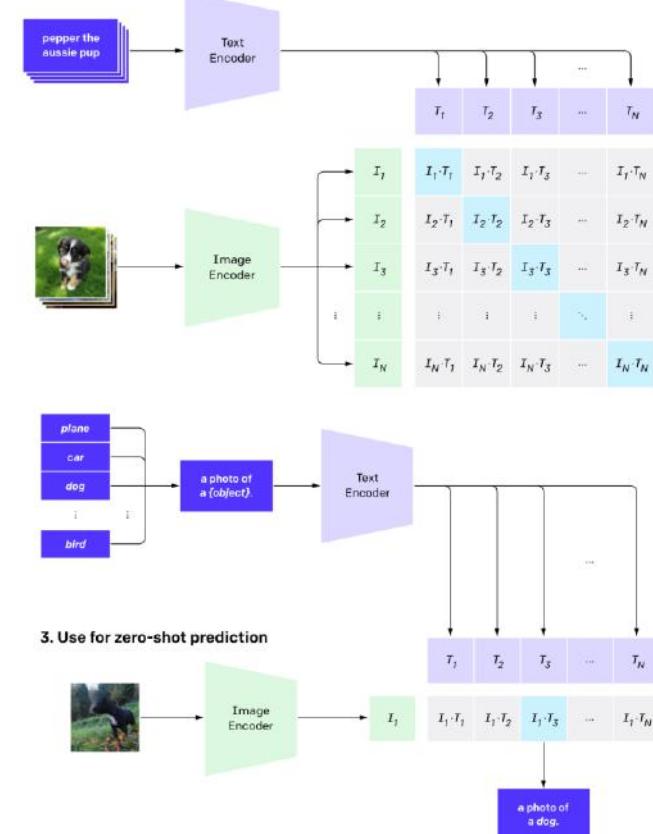
➤ GPT-3升级自然语言处理，CLIP实现文本与图像互相理解

- **GPT**: 2018年由OpenAI提出，参数量1.17亿，预训练数据量约5GB，基于Transformer模型充分利用海量无标注文本进行预训练，赋予文本大模型在小数据集、零数据集下的理解和生成能力，提升了生成式AI的认知能力。2020年GPT-3推出，参数量达1750亿，预训练数据量达45TB，除自然语言推理、句子关系判断、问答、常识推理、分类等常见NLP任务外，GPT-3在撰写文章、编写SQL语句、编写JavaScript代码等困难任务也有优异表现，入选了《麻省理工科技评论》2021年“十大突破性技术”。
- **CLIP**: 2021年由OpenAI提出，利用文本信息监督视觉任务自训练，训练数据集为40亿个“文本-图像”对，采用Transformer模型对图像的patch序列进行建模，将不同模态的原始数据映射到统一或相似的语义空间，实现不同模态信号间的相互理解，拥有寻找不同模态数据间关系的能力，基于此能够实现不同模态数据间转化与生成，进一步可以根据图片生成对应语言描述也可以根据语言提示生成对应的图片，极大丰富了生成式AI技术的应用广度，为AIGC带来了更多的可能性。

图：GPT模型示意图



图：CLIP模型示意图

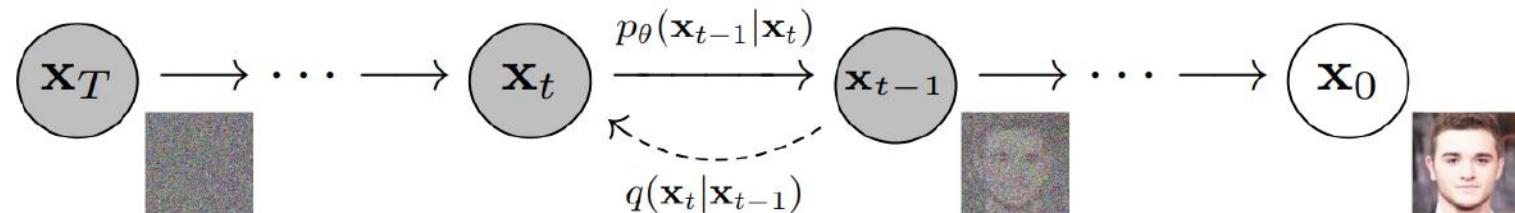


2.5 技术：快速进化

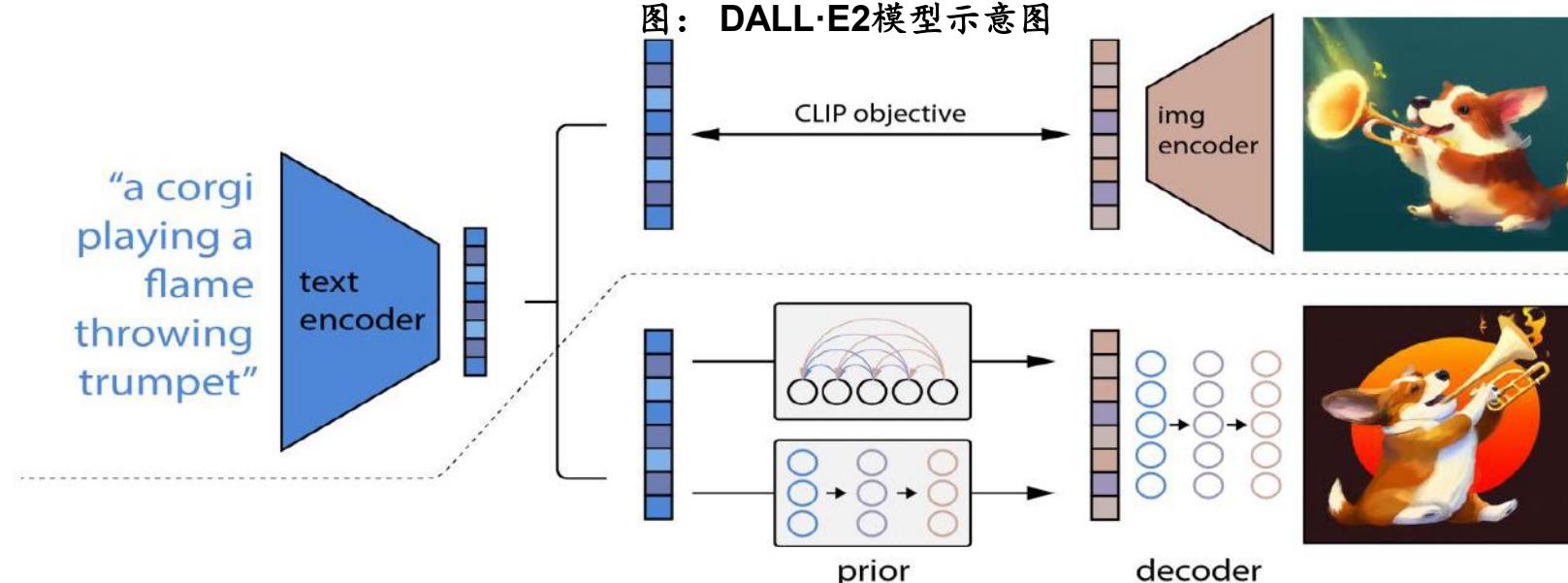
➤ Diffusion与DALL·E2实现高质量图像产出

- 扩散模型（Diffusion model）：扩散模型的概念最早在2015年“Deep Unsupervised Learning using Nonequilibrium Thermodynamics”中被提出。2020年“Denoising Diffusion Probabilistic Models”中提出DDPM模型用于图像生成。扩散模型通过给图像增加高斯噪声破坏训练数据来学习，找出逆转噪声过程的方法，利用学习到的去噪声方法实现从随机输入中合成新的图像。该算法在分子图生成中可以进行药物分子和蛋白质分子的生成。
- DALL·E2：2022年4月由OpenAI提出，基于CLIP实现文本与图像的联系，基于Diffusion从视觉语义生成图像，使用先验模型实现从文本语义到相应视觉语义的映射，最终实现以下功能：1) 根据文本生成图片；2) 将图像扩展到画布之外；3) 根据文本对图像进行编辑，实现添加或删除元素；4) 给定一张图片生成保持原风格的变体。

图：扩散模型算法示意图



图：DALL·E2模型示意图

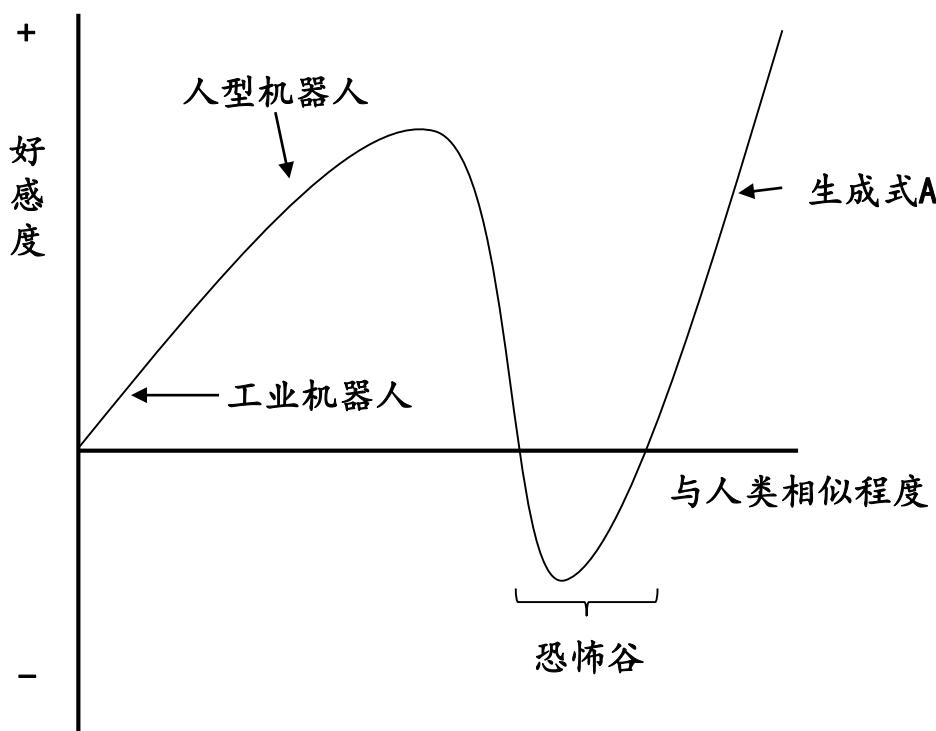


资料来源：OpenAI, 雷锋网, Google Scholar, 腾讯科技, 中信建投

2.6 跨过恐怖谷：人类对生成式AI的好感度随其拟人程度的增加而增加

- 恐怖谷效应：1970年由日本机器人专家森政弘提出的关于人类对机器人和非人类物体感觉的假设，随着类似人类物体的拟人程度增加，人类对其的好感度呈现“增-减-增”的曲线，当类似人类的物体与人类相像到一定程度后，人类对其好感度会下降呈现一个情感反应的低谷，即恐怖谷，而当其与人类的相似程度继续上升，人类对其的好感会随之增加。
- 生成式AI跨过了恐怖谷。随着技术的发展，生成式AI越来越多的能力接近人类甚至超越人类，人们对其不是反感而是热衷于使用，例如抖音中AI绘画特效使用人数已经超过2700万人，AI会话工具ChatGPT发布5天用户即超百万，人类对生成式AI的好感正随着其能力的提升、与人类相似程度的提升而增加，说明生成式AI已经成功跨过了恐怖谷。

图：生成式AI已经跨过恐怖谷



图：抖音AI绘画特效使用人数超2700万



2.7 生成式AI技术演进展望

➤ 根据红杉资本的预测：后续基本模型的发展将继续推进生成式AI应用发展，预计2023年文本领域和代码领域的应用已经成熟。图像/视频/3D/游戏领域的成熟应用仍需要一定的时间，2025年以后的应用发展进程仅为推测。

图：生成式AI相关应用发展进程时间表

	2020年之前	2020年	2022年	2023年？	2025年？	约2030年？
文本领域	诈骗垃圾信息识别 翻译 基础问答回应	基础文案撰写初稿	更长的文本二稿	垂直领域的文案撰写实现可精调（科学论文等）	终稿，水平高于人类平均值	终稿，水平高于专业写手
代码领域	单行代码补足	多行代码生成	更长的代码更精确的表达	支持更多语种领域更垂直	根据文本生成初版应用程序	根据文本生成终版应用程序，比全职开发者水平更高
图像领域			艺术图标摄影	模仿（产品涉及、建筑等）	终稿（产品涉及、建筑等）	终稿，水平高于专职艺术家、设计师和摄影师
视频/ 3D/ 游戏领域				视频和3D文件的基础版/初稿	二稿	AI版Roblox 可依个人梦想定制的游戏与电影

大规模实现难度： ● 初级尝试 ● 接近成熟 ● 成熟应用

目 录

第一章

AI：从判别决策到创造生成

第二章

生成式AI：范式升级，应用多元

第三章

生成式AI应用：内容制作，从辅助人到“替代”人

第四章

生成式AI应用：多行业垂直应用

第五章

海内外发展：多家科技巨头、初创公司积极布局

第六章

综合应用：游戏行业的案例

3.1 生成式AI应用：娱乐媒体内容制作

- 当前阶段生成式AI最常见的应用场景为娱乐媒体内容的辅助生产。以文字续写或纠错、文字转语音、文字生成语音、图像智能编辑、视频智能剪辑等方式替代既有创意产生后的专业性机械劳动，同时通过内容生成满足用户的娱乐需求或者为其提供创作灵感。
- 随着生成式AI的不断成熟，部分专业内容生产者将被替代。数据、算力的进一步提升将带来具有更强生成效果的AI，能够根据用户需求生成个性化定制内容终稿，并且AI生成内容达到专业内容生产者水平且具有独特新颖的创意，从而替代部分文字作者、翻译人员、插画创作者、配音人员、音乐制作人、视频编辑人员等等。

图：生成式AI在娱乐媒体内容领域的应用



机器
辅助人

- 辅助大批量内容生产
- 改进内容生产人工操作流程，生成内容初稿
- 替代既有创意产生后的专业性机械劳动

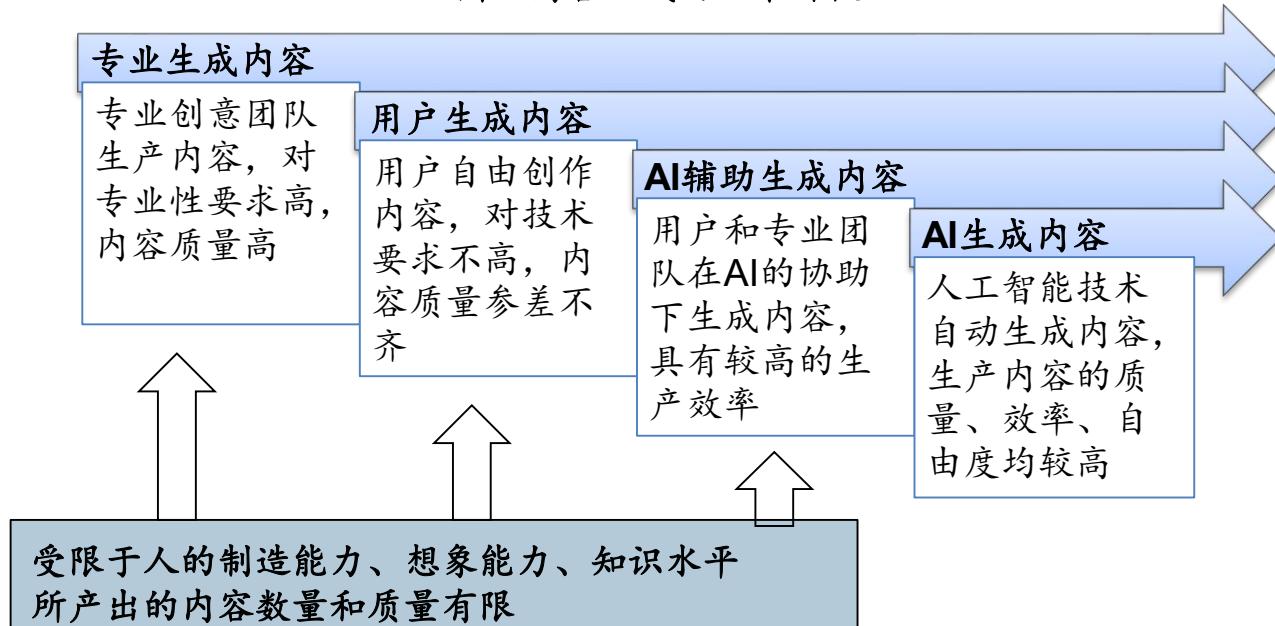
机器
“替代”人

- AI生成内容达到专业水准
- 根据用户需求生成个性化定制内容终稿
- 替代部分专业内容生产者

3.2 AIGC：生成式AI在娱乐媒体领域的应用（1）

- **AIGC (AI-Generated Content)** 指利用人工智能技术自动生成的内容，是继专业生成内容（PGC）和用户生成内容（UGC）之后一种新型生成内容的方式。国际上被称为人工智能合成媒体（AI-generated Media或Synthetic media），是通过人工智能算法对数据或媒体进行生产、操作和修改的统称。
- **AI绘画作品出圈、一级投资活跃，2022年AIGC爆发式发展。** 2022年8月，在美国科罗拉多州举办的新兴数字艺术家竞赛中，参赛者提交AIGC绘画作品《太空歌剧院》获得了此次比赛“数字艺术/数字修饰照片”类别一等奖，而参赛者没有绘画基础，通过AI绘图软件MidJourney耗时80个小时创作了该作品。AI绘画技术发展快速，从年初的技艺生疏到目前能够生成专业级别的图像仅过去几个月的时间，目前国外已有Stable Diffusion Midjourney等多个成熟的AIGC平台。其中技术领先的AI公司Open AI当前估值或达290亿美元，Stable Diffusion背后的Stability AI 估值也达到10亿美元。

图：内容生成的四个阶段



图：AI绘画作品《太空歌剧院》获奖出圈



3.2 AIGC：生成式AI在内容创作领域的应用（2）

- AIGC大幅度降低了数字内容生产的成本，打破了数字内容生产受到人类想象能力和知识水平的限制，广泛应用于文本生成、音频生成、图像生成、视频生成、跨模态生成及游戏领域，其广泛应用能够满足数字经济时代日益增长的数字内容供给需求。

表：AIGC应用场景及所处发展阶段

类型	应用场景	所处阶段
文本生成	新闻播报等结构化写作	广泛应用，技术细节待提升
	剧情续写、营销文本等非结构化写作	底层技术明确，预计1-2年将规模化应用
	内容推荐、文章润色等辅助性写作	广泛应用，技术细节待提升
	智能客服、聊天机器人等交互式文本写作	广泛应用，技术细节待提升
	文本交互游戏等	底层技术明确，预计1-2年将规模化应用
音频生成	语音克隆	广泛应用，技术细节待提升
	由文本生成播报、虚拟人歌声等特定语音	广泛应用，技术细节待提升
	作曲、编曲、自动作词等乐曲/歌曲的生成	广泛应用，技术细节待提升
图像生成	去除水印、提高分辨率、滤镜等图像编辑工具	广泛应用，技术细节待提升
	按照特定属性生成画作、根据指定要求生成功能性图像	底层技术明确，预计1-2年将规模化应用
视频生成	删除特定主体、生成特效、跟踪剪辑等视频属性编辑	广泛应用，技术细节待提升
	视频换脸等视频部分剪辑	底层技术明确，预计1-2年将规模化应用
	对特定片段进行检测及合成等视频自动剪辑	底层技术待完善，增长可期
跨模态生成	根据文字生成创意图像	底层技术明确，预计1-2年将规模化应用
	根据图片素材生成视频	底层技术明确，预计1-2年将规模化应用
	根据文字生成创意视频	底层技术待完善，增长可期
	根据图像或视频生成文字	底层技术待完善，增长可期
游戏	AI Bot	底层技术明确，预计1-2年将规模化应用
	NPC逻辑及剧情	底层技术明确，预计1-2年将规模化应用

3.2.1 文本生成：对话机器人、新闻稿撰写广泛应用

- 文本生成是生成式AI最早应用的领域之一，已经在对话机器人、内容续写、新闻稿撰写、诗歌小说创作等领域具有广泛的应用。
- 对话机器人：包括问答型机器人、闲聊型机器人、任务型机器人、知识图谱型机器人、多轮对话机器人，在智能客服场景中能够显著降低企业人力成本。
- 新闻稿撰写：在全球范围具有广泛的应用，很多新闻机构使用AI生成稿件，尤其是体育、天气、股市交易变动、公司业绩报道等结构性新闻报道。自然语言生成公司Automated Insights仅在2014年就产生了10亿篇新闻文章，每秒可撰写多达2000篇新闻报道，用户包括雅虎、美联社等。

图：腾讯云对话机器人

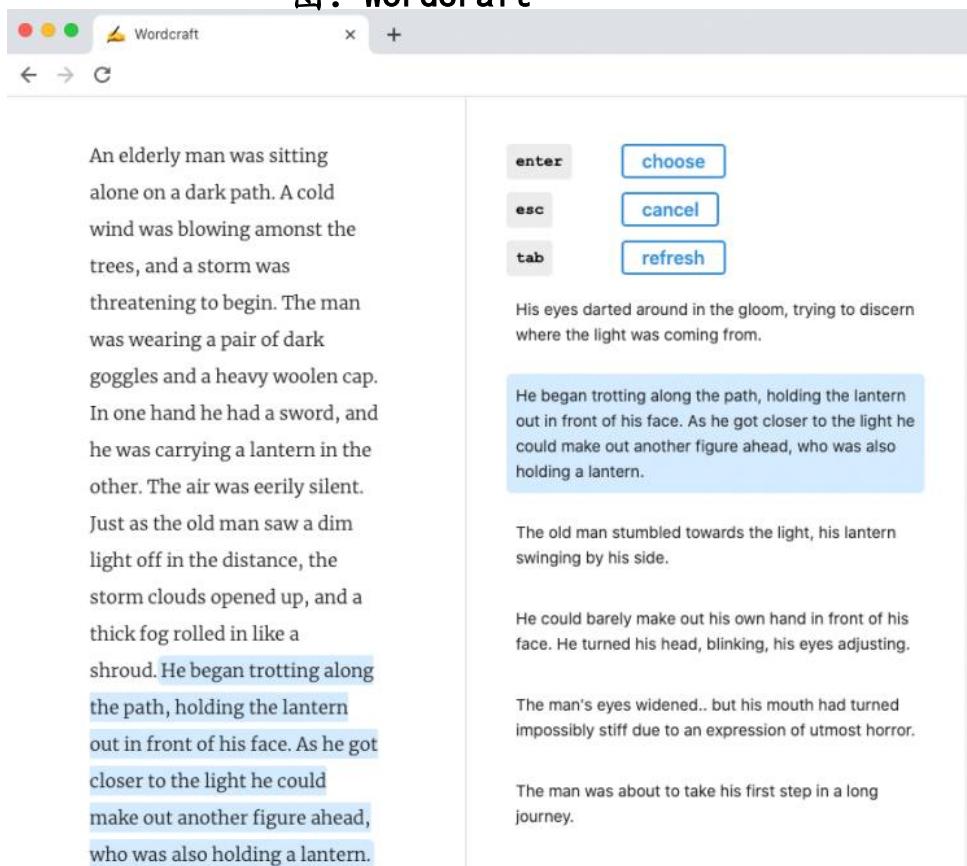


图：Automated Insights为雅虎生成的体育新闻

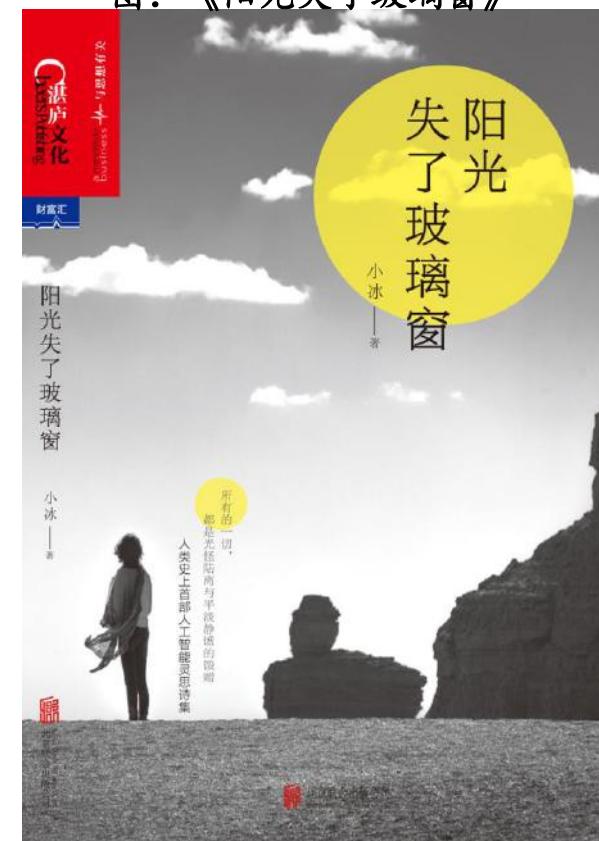
3.2.1 文本生成：内容续写、文学创作部分达到专业水平

- 内容续写：基于给定的文本续写新的内容，续写的内容包括文学创作、商业写作、教学内容等，典型应用有Google推出的Wordcraft、OpenAI推出的GPT-3等。
 - Wordcraft：2021年由谷歌推出，作者和AI以对话的形式共同编辑故事，支持续写、扩写、改写、生成大纲等功能，能够极大提升写作效率，激发创作者灵感。
- 文学创作：2017年微软研发的AI“小冰”出版了人类历史上第一部人工智能诗集《阳光失了玻璃窗》，其中包含139首现代诗；2018年，AI创作的小说《1 The Road》出版；2022年谷歌子公司DeepMind发布剧本写作AI-Dramatron，能够从一句话表述的戏剧冲突中生成剧本标题、角色、场景、对话；越来越多的AI文学创作平台进入公众视野，AI创作的内容包括剧本、诗歌、小说、对联等等。

图：Wordcraft



图：《阳光失了玻璃窗》



3.2.1 文本生成：ChatGPT集大成

- 11月30日，OpenAI发布语言模型ChatGPT，该模型采用对话的形式与人进行交互，可以回答后续问题、承认错误、挑战不正确的前提、拒绝不适当的请求。
- ChatGPT不仅在日常对话、专业问题回答、信息检索、内容续写、文学创作、音乐创作等方面展现出强大的能力，还具有生成代码、调试代码、为代码生成注释的能力。
- ChatGPT发布后5天用户已超百万，而OpenAI之前发布的GPT-3用了将近两年的时间才突破100万用户。

图：ChatGPT回答专业问题

EF 什么是物理引擎

物理引擎是一种用于模拟物理现象的计算机软件。它可以被用来模拟各种物理系统，例如运动、碰撞、重力和碰撞等，并且能够按照物理定律进行计算，以便在计算机游戏或其他软件中提供真实的物理效果。物理引擎在游戏开发、动画制作、工程模拟等领域中都有广泛应用。

EF 举个例子说明一下

例如，假设我们正在开发一款新的计算机游戏，其中玩家可以操作一个角色在游戏中进行跳跃。为了让游戏看起来更加逼真，我们可以使用物理引擎来模拟角色在空中运动的情况，包括速度、加速度、重力等。当玩家按下跳跃键时，物理引擎会按照物理定律计算出角色的运动轨迹，并将其呈现在游戏中。这样，玩家就能体验到真实的跳跃效果，而不是简单的跳跃动画。

图：ChatGPT创作的小说大纲

EF 以“逃离太阳系”为主题创作一篇小说的大纲

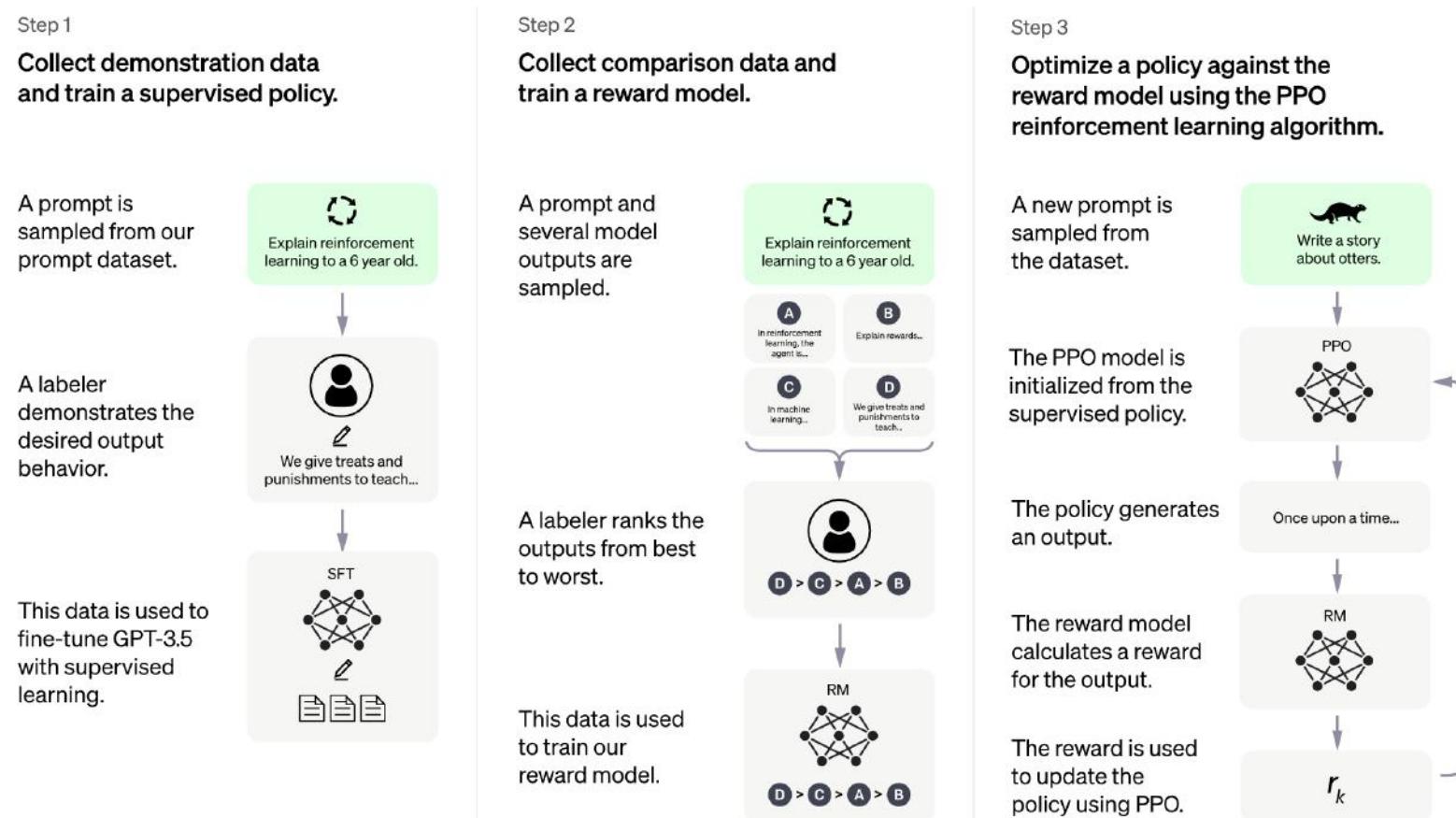
逃离太阳系的小说大纲：

1. 故事背景：在远未来的某个时刻，地球已经不再适宜人类居住，政府决定派出一支探险队，寻找新的宜居星球。
2. 主公：探险队成员之一的年轻女孩艾莉，她是一名勇敢而有智慧的宇航员，担任这次探险的领队。
3. 探险过程：艾莉带领探险队乘坐一艘高科技的宇宙飞船，逃离地球，踏上漫长的宇宙旅途。在旅途中，他们遇到了各种各样的挑战和危险，包括太阳风暴、尘埃星云、宇宙海盗等。
4. 结局：经过多年的探险，探险队终于找到了一颗适宜人类居住的星球，艾莉和其他成员安家落户，开始了新的生活。但她也意识到，这只是一个开始。

3.2.1 文本生成：ChatGPT通过基于人类反馈的强化学习训练

- ChatGPT是基于GPT-3.5系列微调的结果，通过基于人类反馈的强化学习（Reinforcement Learning from Human Feedback, RLHF）进行训练：
- 收集演示数据训练监督策略：从提示数据集中抽取提示输入并人工给出期待输出，此数据用来微调GPT-3.5；
 - 收集比较数据训练奖励模型：人工对模型输出的答案以及人工答案按照符合程度进行排序，此数据用于训练奖励模型；
 - 使用近端优化策略(PPO)优化奖励模型：在提示数据集中抽取新的提示输入，PPO模型由监督策略初始化，模型根据提示输出后由奖励模型计算奖励值，PPO使用奖励值更新策略参数，然后更新PPO，持续迭代优化。

图：ChatGPT通过RLHF训练



3.2.1 文本生成：以GPT-3和ChatGPT为例的应用及展望

- GPT-3于2020年5月发布，能够生成高质量文本，截止2021年3月，已经有超过300个应用程序提供GPT-3 驱动的搜索、对话、文本完成和其他AI功能，GPT-3平均每天生成45亿字。目前GPT-3已应用于：从客户评论中识别主题并生成摘要、为虚拟人生成对话使其与人自然交流、提高语义搜索的准确率等。
- ChatGPT相较于GPT-3具有支持多轮对话、易于修正、人机交互效果更好、更快、更高效等优势，且目前免费使用，用户已超百万，更适合应用于人工智能客服、搜索引擎、智能客服、虚拟人、代码编程、文学创作等领域。

图：Viable使用GPT-3生成客户反馈摘要

The screenshot shows a user interface for generating feedback summaries. On the left, there's a sidebar with various filters like Sentiment, Emotion, Topic (with 'calendar' selected), and Activity. The main area has a search bar with the query 'How should we improve calendars?'. Below it is a 'Feedback Summary' card with a 'BETA' badge. The summary text reads: 'We should improve calendars by allowing users to see weekly and monthly views in addition to the default daily view. We should also make it easier to add Zoom or Hangout links, and make it easier to check other people's availability.' To the left of the text is a donut chart titled '7% of feedback mentions calendar'. To the right is a list titled 'Frequently mentioned with "calendar"' with items: availability (6%), default (3%), other people (2%), and hangout links (2%). At the bottom, there's a question 'Was this summary helpful?' with upvote and downvote icons. A note at the bottom says 'We found 129 pieces of feedback to back this up.' At the very bottom, it shows 'DA Kevin@zphrr.ai' and the date '8pm 11/24/2020'.

图：GPT-3驱动的虚拟人Lucy



3.2.2 音频生成：语音生成广泛应用，大厂均有布局

- 自动语音生成：将文本转化为语音，广泛应用于新闻阅读、有声书、出行导航、通知播报、视频配音等领域。目前，谷歌、微软、亚马逊、腾讯、阿里、百度、科大讯飞等均推出了相关平台，支持在不同场景下将文本转化为流畅逼真的语音。
 - 谷歌Text-to-Speech：谷歌基于DeepMind 的语音合成专业技术打造，支持40多种语言、220多种语音和参数调整，并且支持用户上传自己录制的音频来训练自定义语音模型。
 - 科大讯飞在线语音合成：提供了100+发音人，支持多语种、多方言和中英混合，可灵活配置音频参数，只需用户上传15分钟录音便能够生成专属声音。

图：谷歌Text-to-Speech

DEMO

Put Text-to-Speech into action

Type what you want, select a language then click "Speak It" to hear.

Text to speak:

Google Cloud Text-to-Speech enables developers to synthesize natural-sounding speech with 100+ voices, available in multiple languages and variants. It applies DeepMind's groundbreaking research in WaveNet and Google's powerful neural networks to deliver the highest fidelity possible. As an easy-to-use API, you can create lifelike interactions with your users, across many applications and devices.

text ssml

Language / locale: English (United States)

Voice type: WaveNet

Voice name: en-US-Wavenet-D

Audio device profile: Default

Speed: 1.00

Pitch: 0.00

Show JSON ▾

▶ SPEAK IT

资料来源：Google, 科大讯飞, 中信建投

图：科大讯飞在线语音合成



3.2.2 音频生成：歌曲生成已有商业级模型，应用场景广阔

- 歌曲生成：AI作词、作曲、编曲等，具体包括由旋律生成歌词、由歌词生成旋律、不同曲风旋律生成、和弦生成、音乐续写等，可应用于音乐欣赏、游戏音效、实体场景配乐等多个领域。
 - 天工乐府**SkyMusic**：2022年由昆仑万维推出的商业级作曲AI模型，支持31种语种的歌词生成、多曲风旋律生成、多轨道编曲、VOCAL生成和智能缩混等，已在全球多个音视频平台发行近20首AI生成歌曲。
 - **Playform AI**：与音乐专家合作使用AI续写了贝多芬《第十交响乐》，完成了两个完整的超过20分钟的乐章，《第十交响乐》于2021年成功由波恩贝多芬管弦乐团演出。

图：天工乐府作曲专辑上线流媒体平台



图：AI续写贝多芬《第十交响乐》



3.2.3 图像生成：快速出圈，多款应用火爆

- 2022年下半年，AI由文本生成图像快速出圈，多款应用火爆。目前对大众开放的图片生成模型主要包括Stable Diffusion、OpenAI的DALL-E 2、Midjourney等，三款模型中，仅Stable Diffusion开源，累计使用人数最多，在全渠道有超1000万用户。
- 谷歌于2022年五月推出了由文本生成高清图像的模型Imagen，其官网展示了Imagen生成的高清图像，并宣称Imagen生成的图像已经全线超越了DALL-E 2，但该模型尚未对外公开使用。

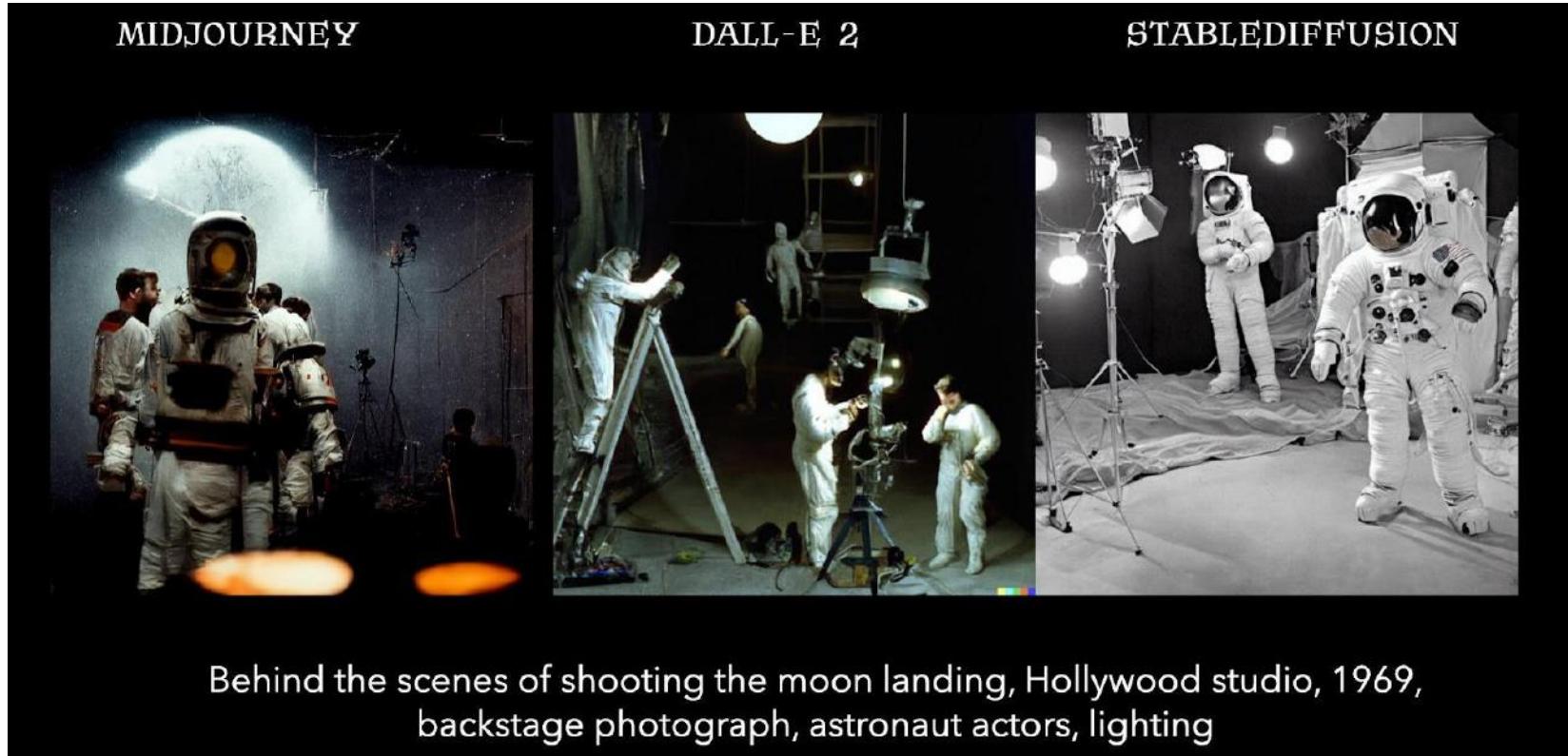
表：Stability AI、Midjourney、OpenAI、谷歌产品对比

	Stable Diffusion	Midjourney	DALL-E 2	Imagen
使用效果				
开发公司	Stability AI	Midjourney	OpenAI	谷歌
推出时间	2022. 08	2022. 07	2022. 04	2022. 05
是否开源	是	否	否	否
是否To C	是	是	是	暂未对外公开
是否有内容创作限制	否	是	是	-
图片处理	本地	云端	云端	-
使用人数 (22. 11)	超1000万	超300万	超150万	-

资料来源：The New York Times, TechCrunch, SiliconANGLE, 公司官网, 中信建投

3.2.3 图像生成：快速出圈，多款应用火爆

➤ 由于底层算法和训练数据的不同，Stable Diffusion、Midjourney和DALL-E 2创作的图片画风和效果有所差异。



Behind the scenes of shooting the moon landing, Hollywood studio, 1969,
backstage photograph, astronaut actors, lighting

Midjourney:

- 创作更像是绘画，而非照片。据公司创始人Holz介绍，这是有意为之，但也不排除公司未来会推出创作写实图片的版本。

DALL-E 2:

- 创作更写实。
- 底层算法使DALL-E 2更适合图片修复（inpainting）。

Stable Diffusion:

- 有一些突出的问题，比如人类和动物的四肢不自然；不能理解复杂的原创内容等。

3.2.4 视频生成：智能编辑为主，生成有待发展

- **视频智能编辑：**使用AI对视频进行编辑，包括删除视频特定主体、自动跟踪剪辑、视频特效生成、自动添加特定内容、视频美颜等，大大降低了视频编辑的专业门槛，提升了视频剪辑效率，目前已广泛应用于视频创作领域，具体应用包括剪映、百度智能创作平台、抖音、美图等。
- **剪映：**字节旗下视频剪辑软件，支持AI智能字幕、曲线变速、智能抠像、文本阅读等。

图：剪映视频智能抠图



3.2.4 视频生成：智能编辑为主，生成有待发展

- 视频生成：目前由智能视频生成虽有部分应用，但生成的视频时长较短，且部分视频无法准确再现真实世界的运作方式，AI视频生成技术到成熟应用还有一段距离。
 - **Make-A-Video**：2022年9月由Meta推出，支持由文字描述生成短视频，也支持输入图像制作视频及输入视频创作视频的变体。但其生成的视频存在时长较短、清晰度较低、无法复现实现中的动作或逻辑等问题。

图：Make-A-Video根据文字描述生成的视频



(a) A dog wearing a superhero outfit with red cape flying through the sky.



(b) There is a table by a window with sunlight streaming through illuminating a pile of books.



(c) Robot dancing in times square.



(d) Unicorns running along a beach, highly detailed.

图：Make-A-Video生成的部分视频存在缺陷



3.2.4 视频生成：智能编辑为主，生成有待发展

- **Imagen Video:** 2022年10月由Google推出，能根据文字描述生成 1280*768 分辨率（720P）、每秒 24 帧、长128帧的视频片段，与Make-A-Video相比生成的视频更加高清，但其生成的视频仍有部分扭曲和抖动。
- **Phenaki:** 2022年10月由Google推出，能根据 200 个词左右的提示语生成 2 分钟以上的长镜头，讲述一个连贯完整的故事，但生成的视频质量较低。

图：Imagen Video根据文字描述生成的视频



Melting pistachio ice cream dripping down the cone.



A british shorthair jumping over a couch.

图：Phenaki根据文字描述生成的长视频



3.2.4 视频生成：百度推出综合型视频智能创作平台

- **百度智能创作平台：**百度AI推出的一站式内容创作助手平台，涵盖文章创作、视频生成和视频智能编辑。落地媒体机构超过20家，包括人民日报新媒体、新华社、央视网、百家号、好看视频等。
- **视频创作：**作为综合型视频创作平台，基于多模态创作技术，为内容创作者和机构提供AI赋能创作工具，支持图文转视频、图表动画、虚拟主播、语音快剪、智能字幕等多种功能。目前，百度智能创作平台自动创作的短视频数量已超过15万条。

图：百度智能创作平台的视频创作功能



3.3 内容生产环节价值巨大：全球娱乐市场规模2.3万亿美元，各赛道均有千亿市值公司

➤ 全球娱乐及媒体市场规模超2.3万亿美元。据普华永道，2021年全球娱乐及媒体行业的总收入约为2.34万亿美元，中国市场收入达3586亿元，占比约15%。预计2026年全球娱乐及媒体行业的总收入约为2.93万亿美元，CAGR为4.6%。在各细分子行业中均有千亿级别上市公司诞生，其中综合型公司市值更高，腾讯、WALT DISNEY市值分别达3176亿美元、1802亿美元。

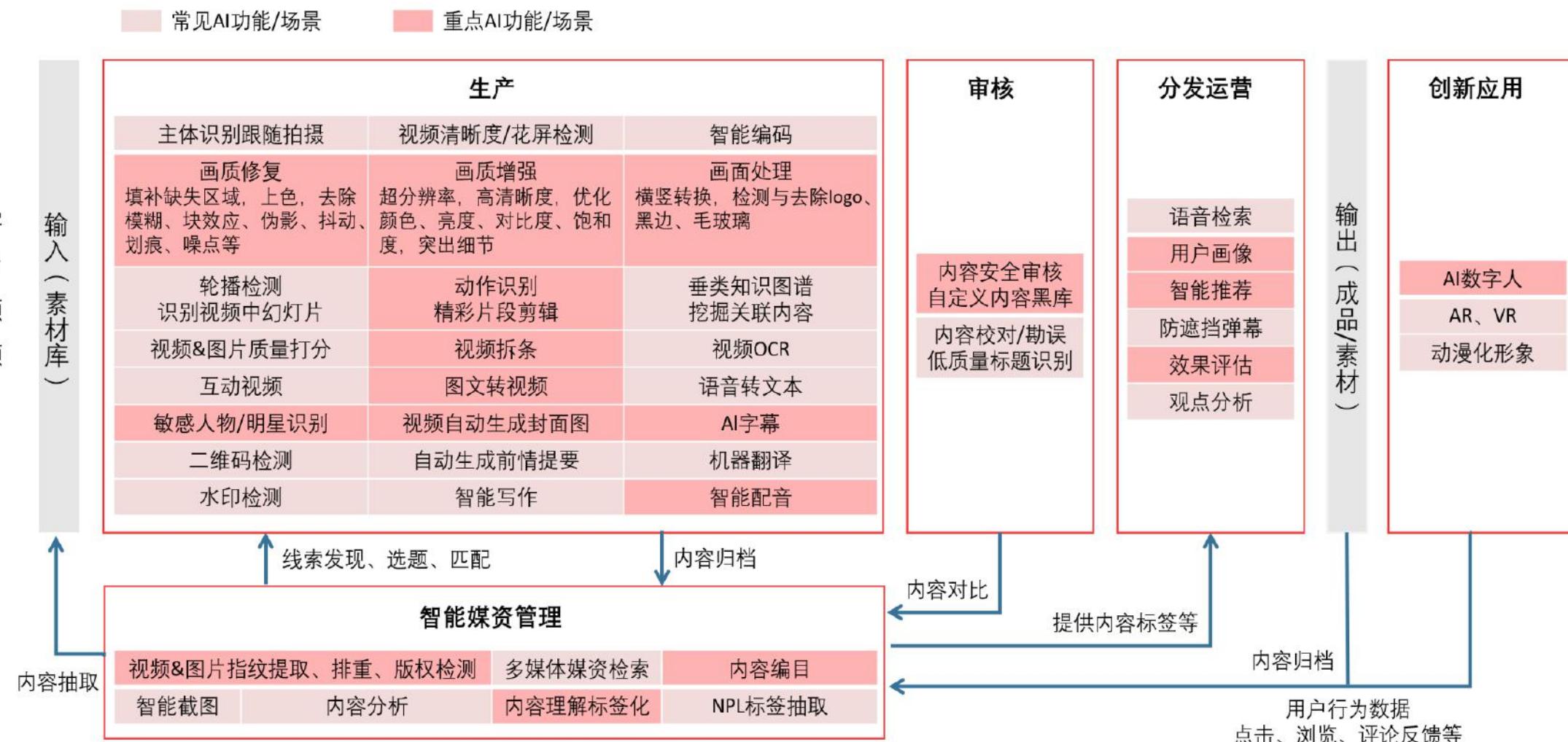
表：娱乐及媒体行业典型内容公司市值及2021年收入

细分行业	公司	市值（亿美元，截至2022年11月24日）	收入（亿美元，2021）
广播电视	SIRIUS XM	252	87
	FOX	160	136
	COMCAST CORP	1541	1164
新闻	News Corp	105	100
出版	Pearson	85	46
电影	Warner Bros	279	122
视频	快手科技	284	127
	爱奇艺	100	48
游戏	Activision Blizzard	575	88
音乐	Spotify Technology	151	109
综合型	腾讯	3176	880
	WALT DISNEY	1802	729

3.3 内容生产环节价值巨大：内容生产者处于产业链核心地位

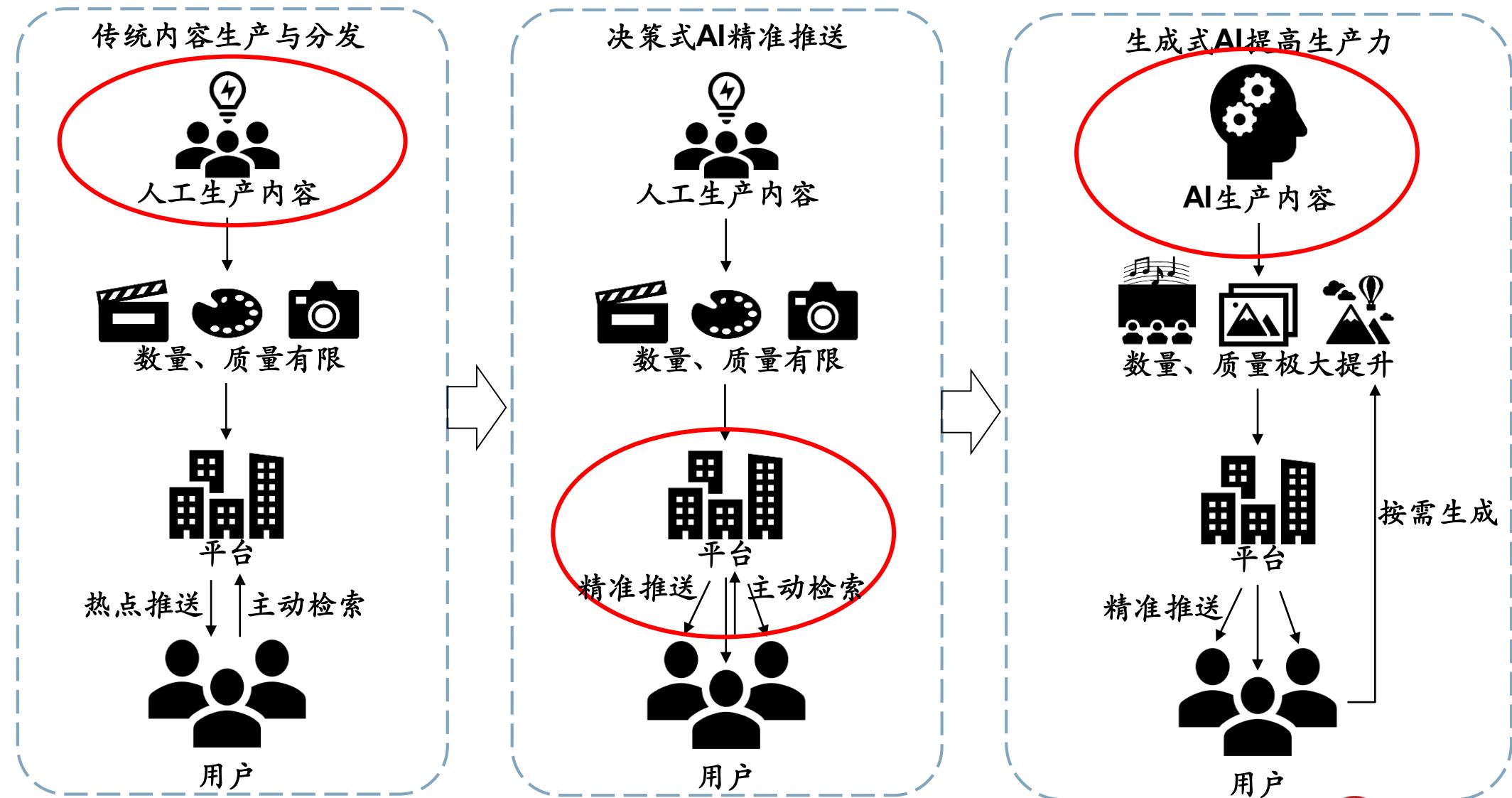
高质量内容是娱乐及媒体行业的核心。高质量内容为娱乐媒体行业的基础，内容创作者在其中扮演了重要的一环，为各类媒介带来了用户流量及使用时长的增长。以视频平台为例，IBM 商业价值研究院对 42 个国家或地区将近 21000 名消费者的调查显示，57% 的受访者认为收到高质量的专属视频内容是对视频提供商保持忠诚的关键动因。

表：AI 在媒体行业应用架构



3.4 生成式AI在娱乐媒体领域：不再是改变分发关系，而是升级生产力

- 决策式AI在娱乐媒体领域的应用集中在内容分发环节，提升分发环节的效率实现了千人千面，而不是对生产力的提升。
- 生成式AI能够创作新的内容，改变了传统的娱乐媒体内容由人工生产的方式，带来的生产力的提升与突破。

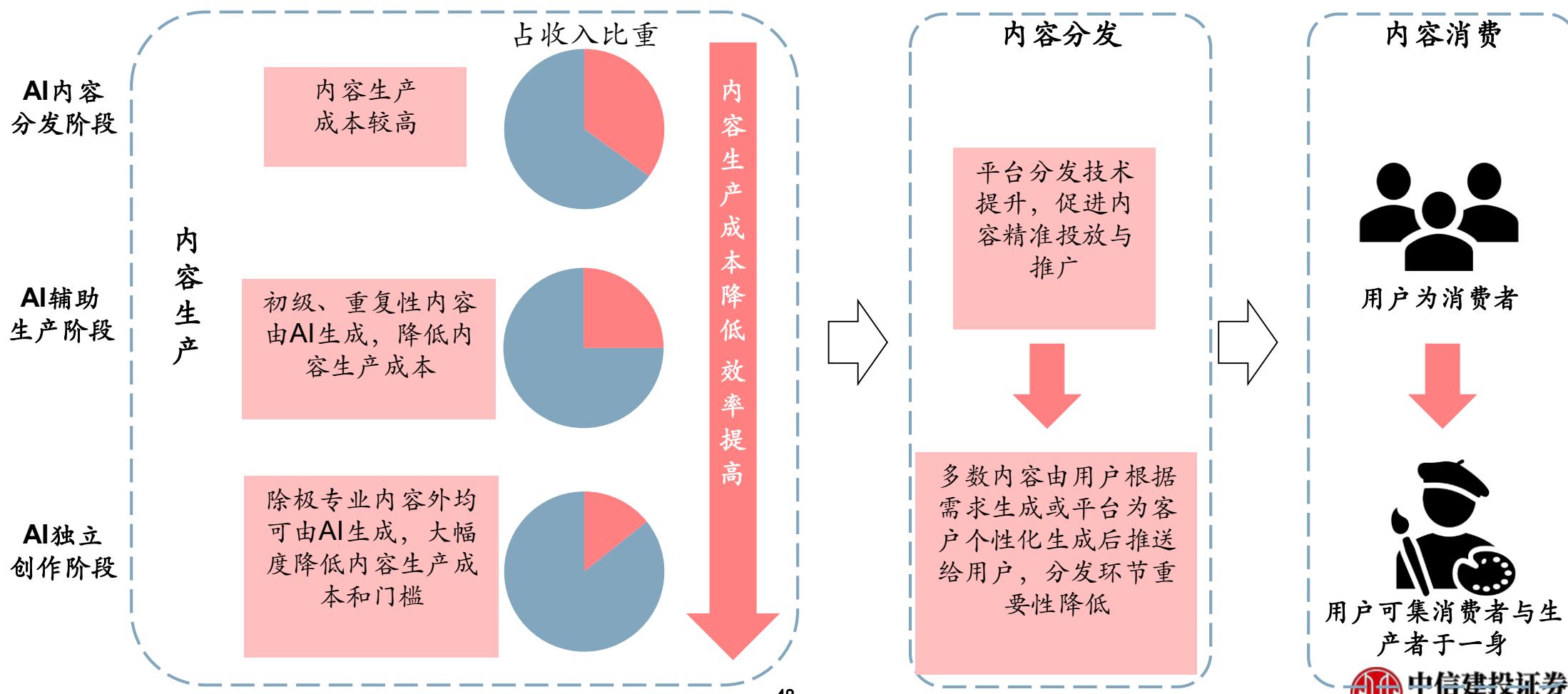


资料来源：中信建投

3.5 生成式AI为现有互联网娱乐巨头带来机遇和挑战——分发环节价值让位于内容生产环节

- AI目前在娱乐媒体领域的应用以内容分发为主，在内容生产阶段有部分辅助应用，后期将走向大规模辅助内容创作甚至大规模替代人类创作。
 - 机器辅助人阶段：生成式AI大幅度降低内容生产成本和门槛，为内容公司降本增效，现有互联网娱乐巨头有望获得更高利润。
 - 机器“替代”人阶段：用户只需要输入指令便可以得到AI创作的所需内容，内容分发环节重要性下降，现有互联网娱乐巨头面临从“精准提供符合用户需求的内容”到“提供符合用户需求的内容生产工具”的挑战。

表：AI在娱乐媒体内容领域的应用



3.5.1 阶段一：机器辅助人：大规模内容辅助创作，提质增效

➤ 生成式AI辅助创作，大幅度降低了内容生产的成本，降低了创作门槛：

- **价格方面：**人工创作价格偏高；生成式AI存在大量免费应用，部分收费的专业版本价格也远低于人工创作。
- **速度方面：**人工创作由于体力、脑力的限制，创作速度存在上限；而生成式AI辅助创作仅需几分钟甚至几秒钟便能够完成，其单纯的创作时间消耗可以忽略。
- **效果方面：**人工创作的效果因人而异，专业门槛较高，且不易统一；而生成式AI辅助创作降低了图片、音频的创作门槛，使得没有专业技能的人也能够完成自己的作品，在部分领域已经能够达到专业水准，生成式AI创作的诗歌、小说、音乐、图片具有发表，部分获得专业奖项。

图：生成式AI辅助创作与单独人工创作的对比

内容	产品形态	创作方式	价格	速度	效果对比
文字	小说/新闻稿/ 聊天对话	单独人工	30-1000元/千字	5000字/天	AI纠错、补全高于人工平均水平，AI续写部分达到专业水平
		AI智能纠错、句子补全、自动续写	免费-10元/千字	最快达上万字/秒，时间几乎可忽略	
图片	绘画/图片	单独人工	几十元/张起	1小时/张起	AI编辑、创作图片优于非专业人士，部分达到专业水平
		AI编辑图片、辅助甚至单独创作	免费-0.4元/张	1张不超过1分钟	
音频/ 音乐	音乐单曲/专辑	单独人工	作曲3000元/首起 配音20/千字起	2天/首起	AI作曲作词降低创作门槛，少部分达到专业水平
		AI辅助作曲、创作歌词、智能配音	免费-10元/首	几分钟/首	

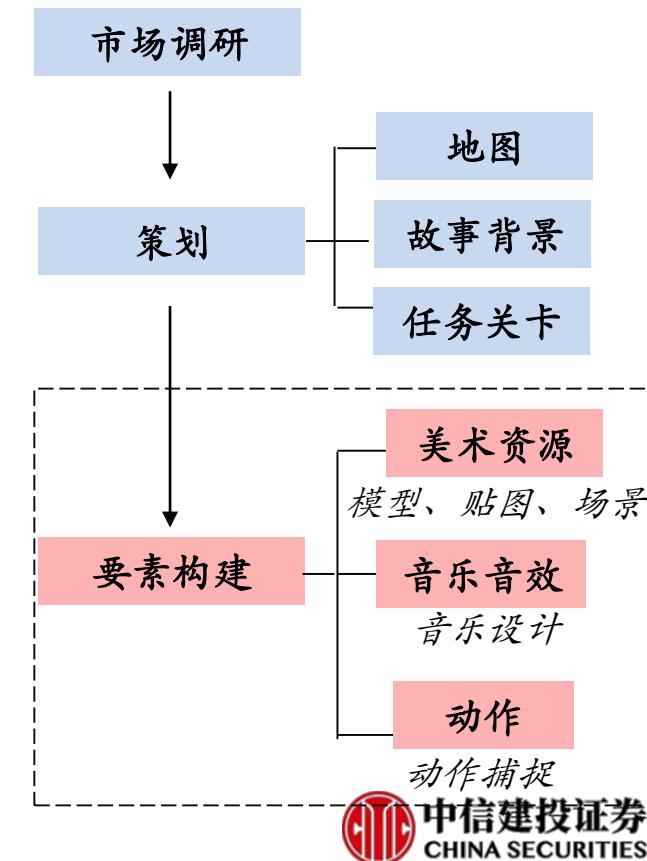
3.5.1 阶段一：机器辅助人：大规模内容辅助创作，降低人工成本

- 典型内容行业的制作、版权成本通常占比不低。如视觉中国2014-2021年累计支付给供稿方的版权费为15亿，占收入的比重为28%；爱奇艺2021年内容成本为207亿，占收入比重为68%；腾讯的VAS业务，21年成本为1386亿（包括游戏、视频、音乐等内容成本），占VAS收入的比重为48%。
- 制作/版权成本的背后通常是人力成本，AIGC有望降低人力成本。以游戏研发为例，其过程中传统的资本支出、固定投入较少，开发人员的人力成本为主要支出。除开市场调研、策划与发行阶段，在要素构建环节（如美术模型贴图、音效设计、动作设计），若用AIGC进行部分替代，则整体制作成本将明显降低。

表：典型内容公司在制作、版权端的成本及占收入比重的情况

领域	公司	成本项	2014-2021年累计				
图片	视觉中国	向供稿方支付版权许可&服务费（亿）	15				
		占收入比重	28%				
领域	公司		2017	2018	2019	2020	2021
长视频	爱奇艺	内容成本（亿）	126	211	222	209	207
		占收入比重	73%	84%	77%	70%	68%
文字/阅读	阅文集团	内容分销成本（亿）		2	6	12	12
		占收入比重		4.4%	6.8%	14.0%	13.3%
音乐	腾讯音乐	服务成本（亿，包括内容成本/分成等）	103	150	175	190	
		占收入比重		54%	59%	60%	61%
游戏/视频/音乐	腾讯	VAS业务成本（亿）	614	740	941	1213	1386
		占VAS收入比重	40%	42%	47%	46%	48%

图：游戏研发流程（虚线框中为AI可替代过程）



3.5.1 阶段一：机器辅助人：国内市场空间可达千亿

- 我们以内容制作成本可AI化制作比例，对AIGC的理论市场空间进行匡算
- 1、内容制作成本：按照各典型行业的市场规模预计的内容（或版权）成本占比进行匡算。如2021年游戏行业市场规模2965亿，游戏行业盈利能力较强&毛利高，预计其中制作成本占比25%；如直播行业主播分成比例惯例为五五分成，因此此处按照50%匡算主播分成成本；其他行业以此类推。
- 2、可AI化比例：当前AI图片、AI音乐相对较为成熟，预计未来可AI化的比例分别为55%、50%，而AI应用在视频类（直播/剧集/电影/综艺）的技术发展相对缓慢，预计可AI化的比例相对较低。综上，我们预计国内AIGC在B端市场的理论空间为1000亿左右。

表：AIGC国内理论市场空间测算

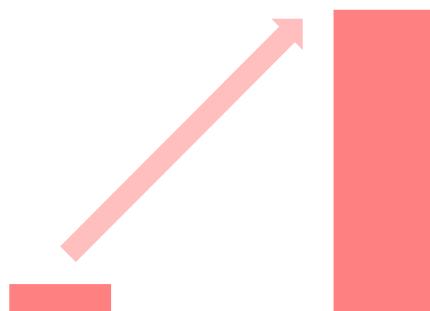
内容类型	市场规模（亿元，2021）	*内容/版权成本占收入比重	内容制作成本（亿元，匡算）	*未来可AI化制作的比例	AIGC市场测算空间（亿元）
游戏	2,965	25%	741	30%	222.38
短视频	2,255	10%	226	35%	78.93
直播	1,936	50%	968	30%	290.40
剧集	902	70%	631	10%	63.14
电影	583	60%	350	20%	69.96
综艺	564	75%	423	25%	105.75
在线音乐	440	60%	264	50%	132.00
图片版权	280	30%	84	55%	46.20
网络文学	209	35%	73	20%	14.63
合计	10,134	37%	3,760	27%	1,023

资料来源：伽马数据，中国互联网络信息中心、中国演出行业协会、国际唱片业协会、中国音乐家协会及中国电影局；弗若斯特沙利文

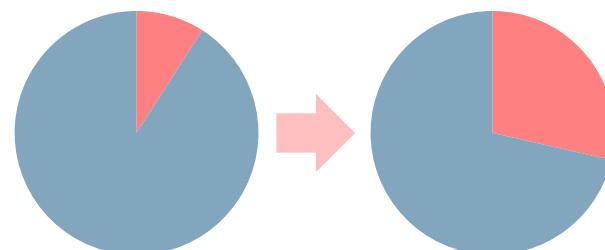
3.5.2 阶段二：机器“替代”人：为娱乐媒体领域带来颠覆性变革

- 随着技术的不断发展，生成式AI技术发展到能够替代大部分专业生产者的阶段，将为娱乐媒体领域带来颠覆性变革，我们做出以下展望：
 - 内容总量指数级增长：内容生产效率极大提升，每个用户都可以成为内容创作者，生产效率和创作人数的增长将带来内容总量的爆发。
 - 内容质量达到甚至超越专业人员创作水平：目前，生成式AI在文字生成、图像生成等领域创作的内容已经可以媲美部分专业人员，随着技术的不断发展，生成式AI创作的内容质量或超越专业人员创作水平。
 - **3D内容、交互性内容占比大大增加：**3D内容、交互性内容制作难度大、成本高，生成式AI的发展将大幅降低3D内容、交互性内容的制作门槛，叠加消费者对高质量内容的需求，3D内容、交互性内容占所有内容的比例将大大提高。
 - 决定公司核心竞争力的因素由内容质量变为生成式AI水平：目前，决定娱乐媒体领域公司核心竞争力的主要因素是内容的质量，但生成式AI将为以往的内容创作者带来更高效的创作方式，也会将部分原来的内容消费者转变为集内容创作和消费于一身的角色，届时生成式AI水平将直接决定内容质量以及创作者、消费者的体验，成为决定公司核心竞争力的因素。

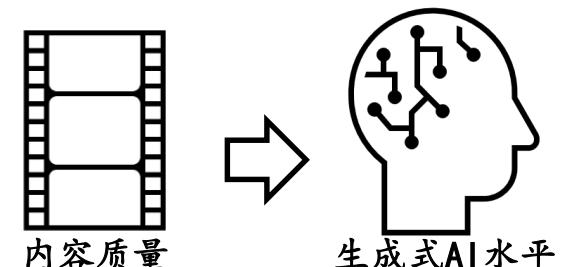
图：生成式AI成熟后为娱乐媒体领域带来的颠覆性变革



内容数量及质量



3D内容、交互性内容占比



公司核心竞争力

3.5.2 阶段二：机器“替代”人：虚拟人集成内容创作工具，成为下一代互联网应用入口

- 最早的虚拟人不具有智能交互功能。2007年初初音未来出现，能够输入音调、歌词等将人类声音合成为歌声，仅是一个虚拟的人物形象，按照人工既定的设计表演，不具有实时交互功能。
- 早期虚拟人仅具有语音交互功能，只闻其声不见其人。2011年Apple在iPhone 4S中内置Siri，用户可以通过语音与Siri聊天，用户可通过Siri进行拨打电话、设置闹钟、记录待办事项、播放音乐等收集控制。后期出现的小度、小爱等以智能音响为代表的智能家居，也是仅具有语音交互控制功能，不具有鲜活的可视形象。
- 目前的虚拟人智能程度增强，以AI驱动为核心。多数虚拟人支持捏脸、智能语音交互、表情动作变化、情绪变动等等，形象更加真实，动作自然。目前虚拟人制作以AI驱动为核心，动捕、渲染、形象库等环节由AI完成，大幅度降低了制作成本。
- 未来虚拟人集成生成式AI后，将成为元宇宙的入口，是用户创作的生产力工具集合。虚拟人集成生成式AI后将更为个性化，不仅具有交互、陪伴、控制功能，成为元宇宙的入口，还具有创作能力，成为生产力工具集合的形象代表，未来有望成为通用人工智能助手的形象代表。

图：虚拟人发展历程及展望



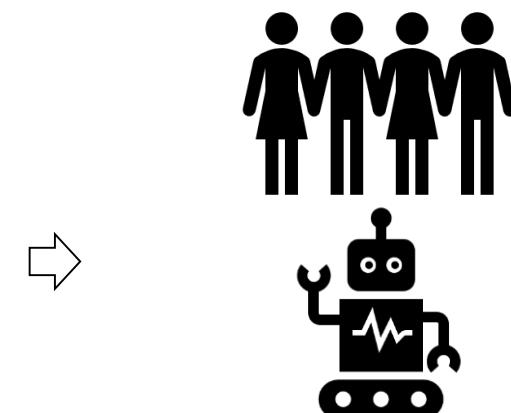
初音未来
仅为表现形象



Siri、小度等
语音交互



洛天依
AI驱动、多种交互

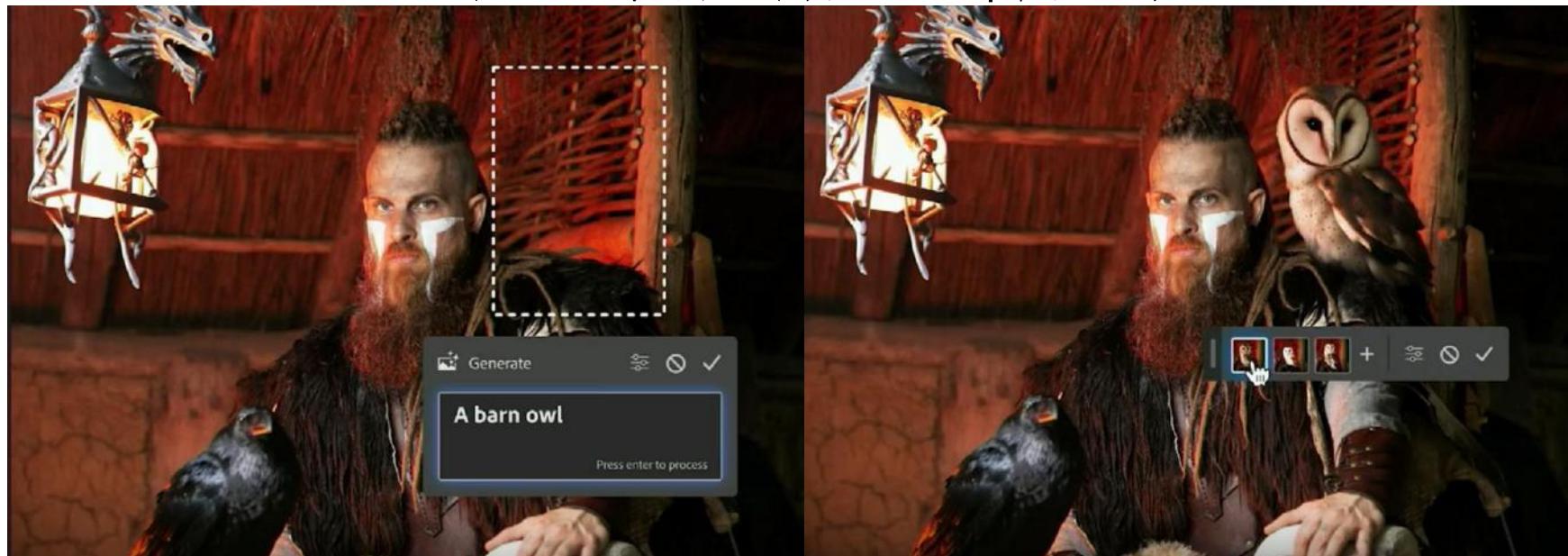


集成生成式AI的虚拟人
千人千面
元宇宙入口
生产力工具集合

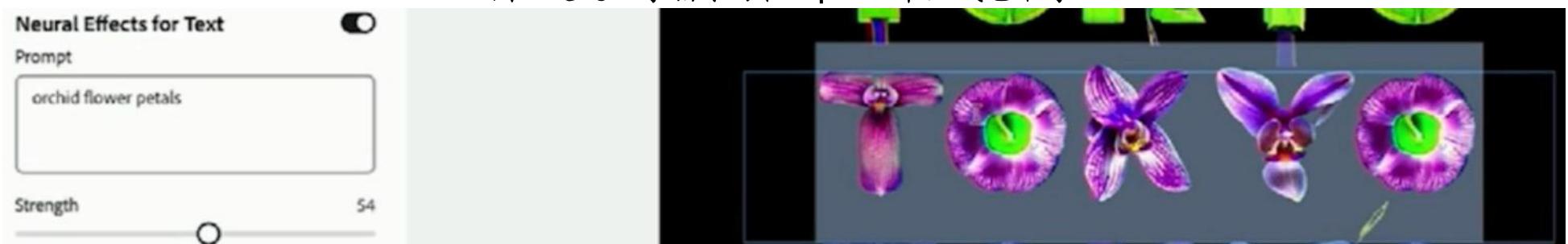
3.5.2 阶段二：机器“替代”人：未来虚拟人雏形已现——Adobe Sensei

➤ **Adobe Sensei**: 2016年由Adobe推出的人工智能平台，能够为Adobe家族中各种产品提供设计和数字支持，处于不断更新中，目前已经被嵌入到几十种Adobe产品之中，在一些产品中，人类只需要向Sensei发出指令便可以完成复杂的专业操作。我们认为，**Adobe Sensei**已经具备了未来虚拟人的雏形，可以通过人类的指令完成部分专业内容创作，随着技术的发展，**Adobe Sensei**可能具有独特的虚拟形象并支持语音交互，成为**Adobe**产品范围内的通用人工智能助手。

图：通过文字指令为图片在Photoshop中添加元素



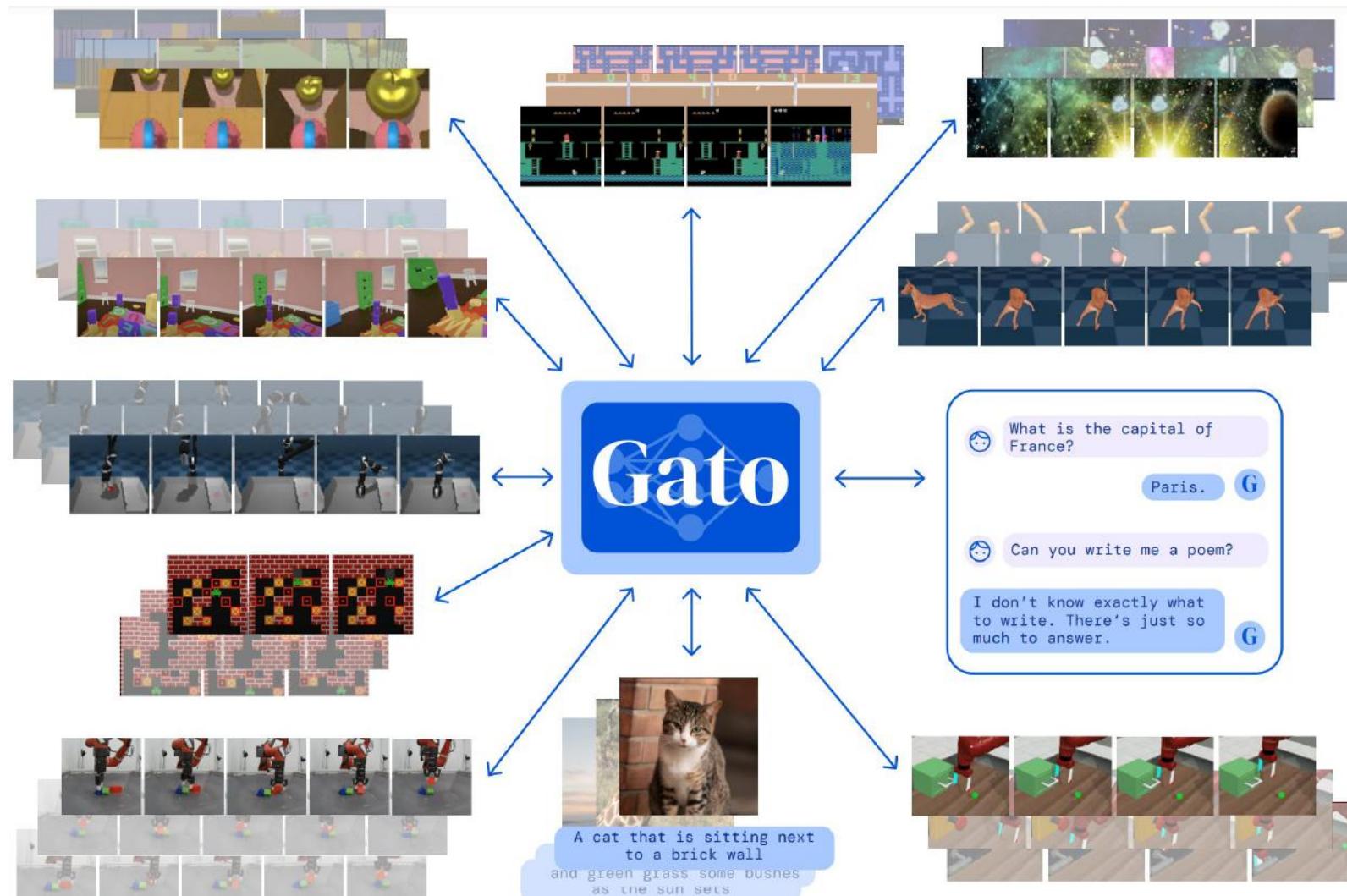
图：通过文字指令为在Express中生成艺术字



3.5.2 阶段二：机器“替代”人：未来虚拟人雏形已现——DeepMind Gato

➤ **DeepMind Gato**: 2022年由DeepMind推出的“通才”AI模型，能够根据指令完成注释图像、交互式聊天、玩小游戏、关节力矩控制、在现实中使用机械臂堆叠积木、在模拟3D环境中导航等数百种任务。相较于**Adobe Sensei**, Gato不仅集成了生成式AI，而是具备了未来集成通用人工智能的虚拟人的内核。

图：Gato能够完成的部分功能



资料来源: DeepMind, 中信建投

目 录

第一章

AI：从判别决策到创造生成

第二章

生成式AI：范式升级，应用多元

第三章

生成式AI应用：内容制作，从辅助人到“替代”人

第四章

生成式AI应用：多行业垂直应用

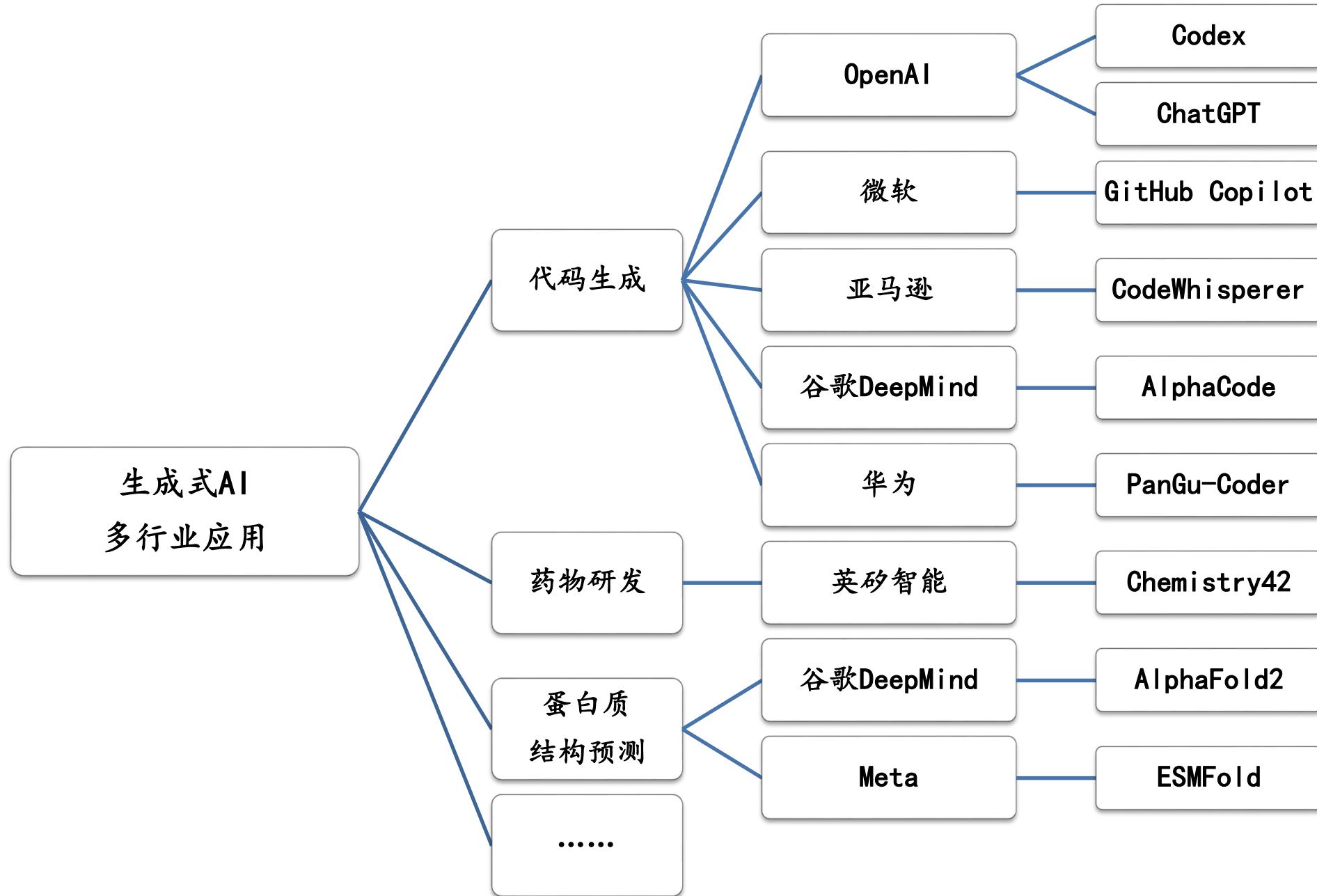
第五章

海内外发展：多家科技巨头、初创公司积极布局

第六章

综合应用：游戏行业的案例

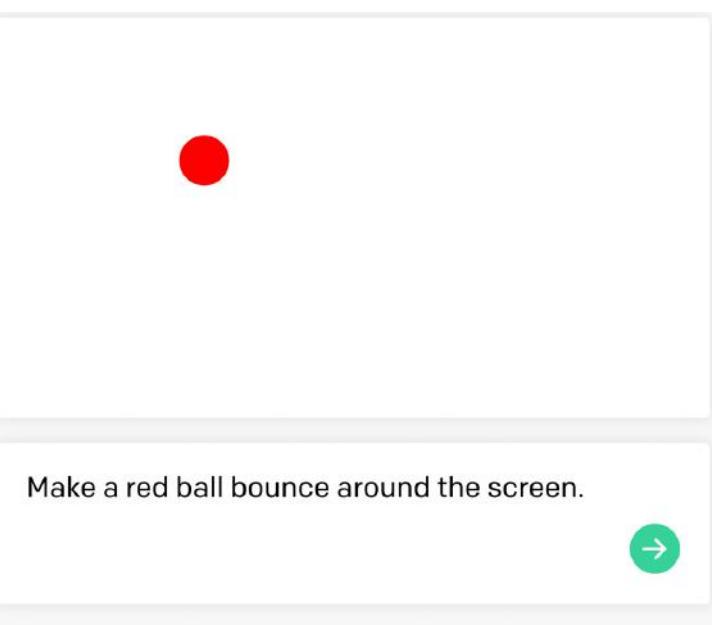
4.1 生成式AI应用：多行业垂直应用



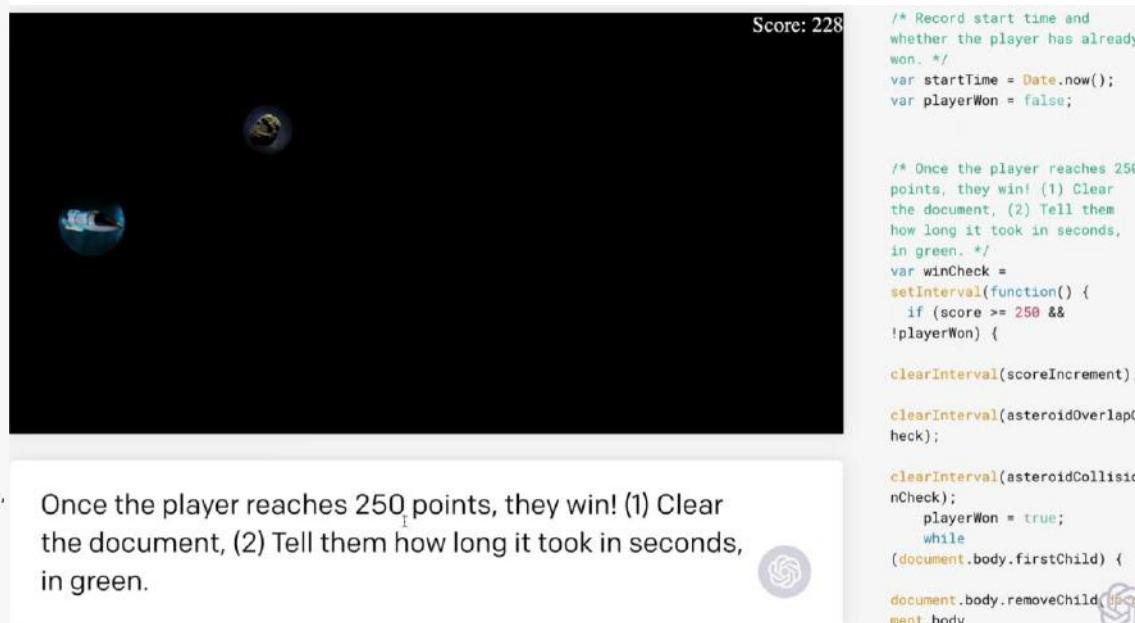
4.2 代码生成：大厂广泛布局，极大提升编程效率

- 代码生成：生成式AI将自然语言翻译成代码，极大提升了计算机编程的智能化、自动化，使得程序员能够在相同时间内编写更多的代码，同时提高debug的效率。目前，OpenAI、微软、谷歌、亚马逊、华为等均在AI代码生成领域有所布局。
- Codex：2021年8月由OpenAI基于GPT-3推出，精通Python、JavaScript、Go、Perl、PHP、Ruby、Swift、TypeScript、SQL甚至Shell等十几种编程语言，其中最为擅长Python。具有将注释变成代码、为代码生成注释、代码自动补全、查找可用库或API、代码检查、代码重构、代码语言转换等功能，是Github Copilot的底层支持，目前免费开放，需要通过API或Playground访问。

图：Codex由英文生成JavaScript代码



图：Codex由英文生成太空小游戏



4.2 代码生成：ChatGPT调试代码表现优于传统搜索引擎

- **ChatGPT：**2022年11月由OpenAI推出，是一个对话语言模型，但在代码生成领域具有优秀的表现，能够将自然语言转为代码、找出代码的问题并给出修改意见，且支持用户和AI连续对话，在调试代码过程中的使用体验优于传统的搜索引擎。

图：ChatGPT由中文生成代码并给出注释

如何使用python写冒泡排序

如果您想使用 Python 实现冒泡排序，您可以使用以下代码：

```
# 定义冒泡排序函数
def bubble_sort(nums):
    # 获取列表长度
    n = len(nums)
    # 遍历整个列表
    for i in range(n):
        # 标记是否发生了交换
        swapped = False
        # 遍历列表中未排序的部分
        for j in range(n-i-1):
            # 如果前一个数大于后一个数，则交换两个数的位置
            if nums[j] > nums[j+1]:
                nums[j], nums[j+1] = nums[j+1], nums[j]
                swapped = True
            # 如果在遍历过程中没有发生交换，说明列表已经有序，则退出循环
            if not swapped:
                break
    # 返回排序后的列表
    return nums

# 测试冒泡排序函数
nums = [5, 2, 9, 1, 4, 8, 6, 3, 7]
print(bubble_sort(nums)) # [1, 2, 3, 4, 5, 6, 7, 8, 9]
```

在上面的代码中，冒泡排序函数 `bubble_sort` 接收一个数字列表作为参数，首先遍历整

图：ChatGPT指出代码错误并给出修改意见

User this code is not working like i expect — how do i fix it?

```
resultWorkerErr := make(chan error)
defer close(resultWorkerErr)
go func() {
    defer cancel()
    resultWorkerErr <- b.resultWorker(ctx)
}()

err := b.worker(ctx)
cancel()
if err == nil {
    return <-resultWorkerErr
}
return multierror.Append(err, <-resultWorkerErr)
```

ChatGPT It's difficult to say what's wrong with the code without more context. Can you provide more information about what the code is supposed to do and what isn't working as expected? Also, is this the entire code or just a part of it?

User just part of it — the error never surfaces. I think there is something wrong with the channel

ChatGPT It's difficult to say without more information about what the code is supposed to do and what's happening when it's executed. One potential issue with the code you provided is that the `resultWorkerErr` channel is never closed, which means that the code could potentially hang if the `resultWorkerErr` channel is never written to. This could happen if `b.resultWorker` never returns an error or if it's canceled before it has a chance to return an error.

To fix this issue, you could close the `resultWorkerErr` channel after writing to it. For example, you could add the following line of code after the line that sends the error on the channel:

```
close(resultWorkerErr)
```

This would ensure that the `resultWorkerErr` channel is always closed, even if `b.resultWorker` never returns an error. Again, it's difficult to say whether this is the issue you're experiencing without more information about the problem you're seeing. Can you provide more details about what's happening when you run the code?

4.2 代码生成：Github Copilot每年节约开发人员成本或达百亿美元

- **Github Copilot**: 2021年6月由微软旗下GitHub推出，由OpenAI Codex 提供支持，可以集成到Neovim, JetBrains IDEs, Visual Studio, and Visual Studio Code等编辑器中，支持 Python、JavaScript、TypeScript、Java、Ruby 和 Go 等编程语言。可以根据上下文自动写代码，包括文档字符串、注释、函数名称、代码，只要用户给出提示，就可以写出完整的函数。
- **Github Copilot每年节约开发人员成本或达百亿美元**。在2021年6月-2022年6月中，超过120万开发者注册使用 GitHub Copilot 预览版，在启用的文件中，有近 40% 的代码是由 Copilot 编写的。以使用人数120万、开发人员时薪20-200美元，Copilot每年可节省开发成本48-1920亿美元。

图：Github Copilot根据提示补全代码

The screenshot shows a code editor interface with several tabs at the top: runtime.go, course.rb, time.js, and IsPrimeTest.java. The main pane displays a Go code snippet. A cursor is positioned on line 14, which contains the line 'if run.Failed {'. A tooltip from 'Copilot' appears below the cursor, showing the completed code block:

```
1 package main
2
3 type Run struct {
4     Time int // in milliseconds
5     Results string
6     Failed bool
7 }
8
9 // Get average runtime of successful runs in seconds
10 func averageRuntimeInSeconds(runs []Run) float64 {
11     var totalTime int
12     var failedRuns int
13     for _, run := range runs {
14         if run.Failed { // Copilot completion tooltip
15             failedRuns++
16         } else {
17             totalTime += run.Time
18         }
19     }
20
21     averageRuntime := float64(totalTime) / float64(len(runs) - failedRuns) / 1000
22     return averageRuntime
23 }
```

At the bottom left, there is a 'Copilot' button.

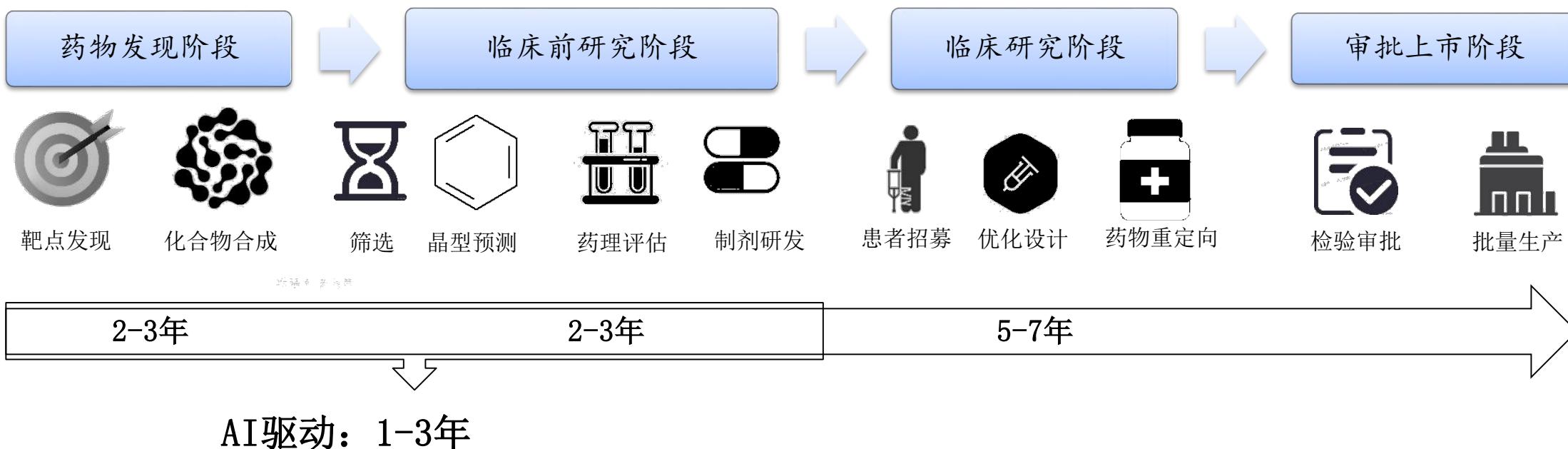
表：Github Copilot节省开发成本测算

使用人数（万人）	120
开发人员薪资（美元/小时）	20-200
使用GitHub Copilot前每年开发人员成本（亿美元）	480-4800
GitHub Copilot节省开发时间	10%
GitHub Copilot节约开发人员成本（亿美元）	48-480
GitHub Copilot节省开发时间	20%
GitHub Copilot节约开发人员成本（亿美元）	96-960
GitHub Copilot节省开发时间	30%
GitHub Copilot节约开发人员成本（亿美元）	144-1440
GitHub Copilot节省开发时间	40%
GitHub Copilot节约开发人员成本（亿美元）	192-1920

4.3 药物研发：生成式AI从头设计药物，大幅降低药物研发成本

- 传统药物研发过程需历时**10-15年**，花费数十亿美元。传统药物研发过程包括药物发现、临床前研究、临床研究、审判上市四个阶段，其中药物发现和临床前研究阶段需要4-6年。据埃森哲，一款新药问世的成本大约在26至67亿美元（包括资金成本和失败成本在内）。
- **AI从头设计药物，大幅降低药物研发成本。**生成式AI在药物发现中，能够根据有关目标结构的信息创建分子，即预测候选药物的元素（例如原子、键的类型等），并为特定目标生成化学实体，而不是通过筛选找到它们。在人工智能驱动下，药物发现和临床前研究阶段可由4-6年降低至1-3年，甚至更短，同时费用也大大降低。具体应用包括基于GAN的Reinforced Adversarial Neural Computer (RANC)，用于新型小分子有机结构从头药物设计；基于AE和GAN的LatentGAN，用于药物从头设计。

图：药物研发流程

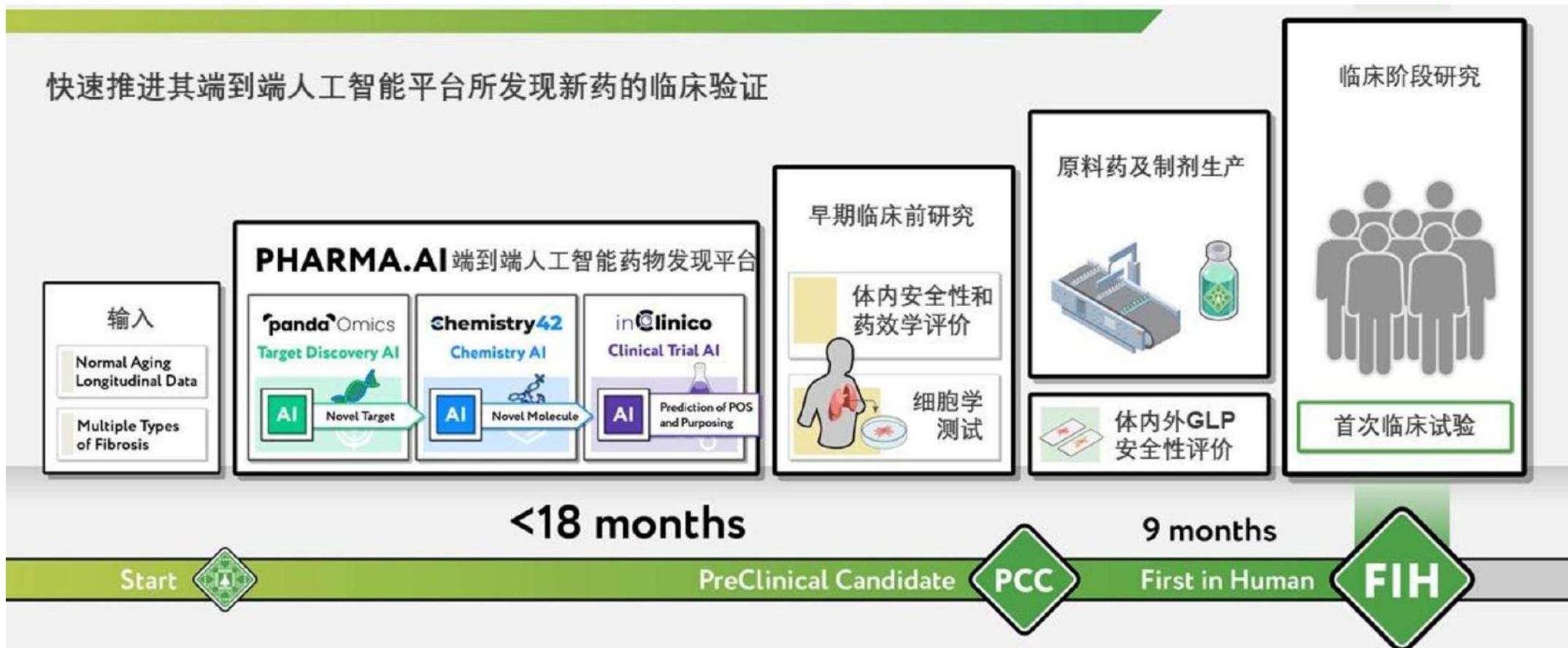


资料来源：蛋壳研究院，埃森哲，MDPI, altexsoft, 中信建投

4.3 药物研发：制药公司的AlphaGo时刻——全球首例完全由AI驱动发现的药物分子

- 全球首例完全由AI驱动发现的药物分子：2021年2月，英矽智能通过新药靶点发现平台PandaOmics和AI分子生成、设计平台Chemistry42，研发出特发性肺纤维化疾病新靶点以及全球首例完全由AI驱动发现的药物分子，用时18个月，投入270万美元，时间成本和资金成本均大大地缘传统靶点药物研发，被誉为制药公司的AlphaGo时刻。目前，该药物在进行的临床I期试验中取得积极顶线数据，在安全性、耐受性、药代动力学方面均表现良好。
- 其中，AI分子生成、设计平台Chemistry42利用生成式对抗神经网络（GAN）来进行深度学习，帮助生成新小分子化合物。

图：人工智能药物发现平台

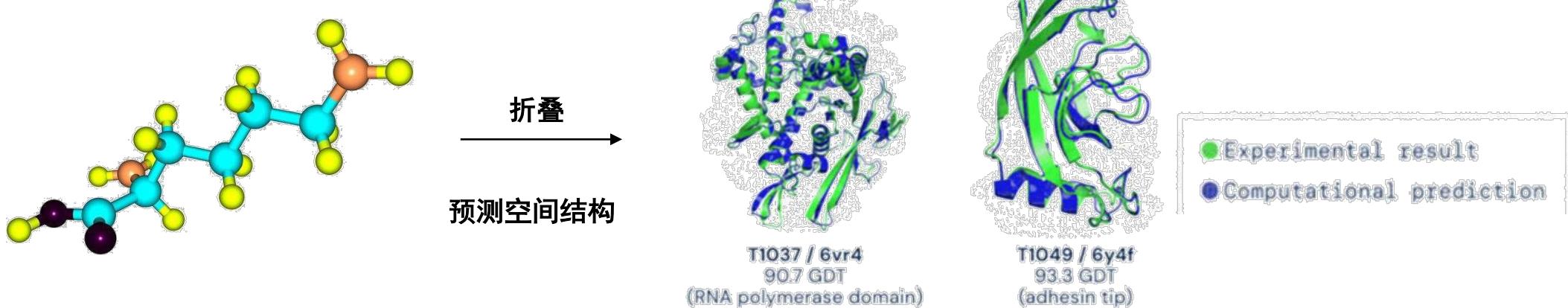


资料来源：蛋壳研究院，埃森哲，MDPI，英矽智能，altexsoft，中信建投

4.4 蛋白质结构预测：AlphaFold 2精确预测蛋白质三维结构

- 蛋白质的三维形状决定了其工作方式和作用，传统破译蛋白质三维结构的方法耗时长且成本高。蛋白质是生命的基石，支撑着每一种生物的每一个生物过程。目前，已知有超过2亿种蛋白质，每种结构不同，传统的使用X射线衍射、冷冻电子显微镜等方法探测蛋白质的确切是一个漫长且昂贵的过程。
- 1972年，美国科学家Christian Anfinsen曾提出著名推论：理论上，蛋白质氨基酸序列可完全决定其三维结构。
- 谷歌母公司 Alphabet 旗下子公司 Deepmind 推出 AlphaFold 2，实现蛋白质结构三维模型的精确预测。AlphaFold 2通过在数据库中搜寻目标氨基酸的相关序列，提取学习相邻“共进化”氨基酸的信息，从而在三维空间对蛋白质结构进行预测。
- Deepmind 2022年7月28日发布的最新版本数据库扩展到了2亿多个蛋白质结构，包括科学界已知的几乎所有蛋白质。AlphaFold团队因此获得了2023年生命科学突破奖（Breakthrough Prize in Life Sciences），这是迄今科研领域里奖金最高的生物学及医学奖项。

图：AlphaFold 2



资料来源：Deepmind, 中信建投

目 录

第一章

AI：从判别决策到创造生成，技术与应用飞速发展

第二章

生成式AI：范式升级应用多元，万亿市场空间可期

第三章

生成式AI应用：娱乐媒体内容制作，从辅助人到“替代”人

第四章

生成式AI应用：多行业垂直应用，生产力大幅提高

第五章

海内外发展：多家科技巨头、初创公司积极布局

第六章

综合应用：以娱乐媒体内容中制作最复杂的游戏为例

5.1 创业公司崭露头角，科技巨头积极布局

- 国外多家生成式AI创业公司获得大笔融资且拥有较高估值。主要包括OpenAI、Stability AI、Midjourney、Jasper等，其中OpenAI技术积累最强，发布了多款生成式AI底层算法，估值或达290亿美元。
- 国内外科技巨头在生成式AI领域多有布局。国内公司百度、腾讯、阿里、字节、网易等及国外公司谷歌、Meta、微软等均推出了生成式AI的应用产品，其中谷歌通过谷歌大脑、旗下Deepmind布局最为丰富。

表：国内外科技巨头布局生成式AI

	公司	典型产品	功能
国内公司	百度	文心	通过文字描述生成图片、图画修复、文档分析、蛋白质结构分析等
	腾讯	优图	输入多张人像图片完成对于人像面部的3D建模、提升电影分辨率等
	阿里巴巴	Lubanner	输入图片、风格、尺寸等，自动完成素材分析、抠图、配色等设计，生成多套符合要求的设计方案
	字节跳动	剪映	通过文字生成视频、AI配字幕等
	网易	网易天音	通过文字生成歌曲，支持作词、作曲、编曲、演唱全链路音乐创作
国外公司	谷歌	Imagen Video、Phenaki	通过文字描述生成视频，前者注重视频清晰度后者注重视频长度和连贯性
	Meta	Make-A-Video	通过文字描述与图片参考生成视频，提供超现实、写实及风格化三种视频类型
	微软	Github Copilot	帮助工程师自动生成代码，目前用户数超过120万，使用该工具生成的代码已超过40%

5.2 创业公司：OpenAI等专注底层系统及应用，Jasper等垂类应用四处开花

- **底层系统：**目前海外参与研发AIGC模型的公司主要包括Stability AI、OpenAI、Midjourney，三家公司的AIGC模型均已正式向C端用户发布，谷歌和Meta等公司也正在研发文字转图片的AIGC模型，但仍未对外发布。
- **垂类应用：**在图片AIGC领域，Stability AI推出了基于自有模型Stable Diffusion的DreamStudio；在文本AIGC领域，应用开发商较多，且底层模型多是基于OpenAI GPT-3。
- **目前，专注底层系统的公司估值更高。**虽然相较应用研发商，基础设施研发商的商业模式尚未成熟，多数还在探索阶段，但因其应用场景更广，且掌握核心技术能力，在未能实现稳定盈利的情况下，估值相对更高。但同样可见，海外除底层算法开发公司以外，初创公司基于开源技术，仍可开发应用并开展To C业务，具有一定商业化潜力，因此我们预计，国内尽管头部互联网公司已积极布局，但中小型公司仍有业务开展余地。



公司	OpenAI	Stability AI	Midjourney	Jasper	Copy.ai	Copysmith
开发内容	底层模型	底层模型、应用	底层模型	应用	应用	应用
AIGC内容	文本、图片	图片、音频	图片	文本	文本	文本
主要产品	GPT-3（模型） DALL-E 2（模型）	Stable Diffusion（模型） DreamStudio（应用）	Midjourney（模型）	Jasper（应用）	Copy.ai（应用）	Copysmith（应用）
营业收入（2021）	1000-5000万美元	-	-	4500万美元	240万美元	-
最新轮融资金额	-	-	-	1.25亿美元	1100万美元	1000万美元
估值	200亿美元	10亿美元	-	15亿美元	-	-

资料来源：各公司官网，TechCrunch, SiliconANGLE, The Information, 中信建投

5.2.1 Open AI：提供底层算法，开发水平强、算法通用性高的AI技术公司（1）

- OpenAI是一家大型人工智能研究公司。公司AI开发水平处于第一梯队，研发人员包括世界著名AI专家，2022年6月，量子计算专家、ACM计算奖得主Scott Aaronson宣布将加盟公司。公司以实现安全的通用人工智能（AGI）为核心宗旨。
- 2015年成立之初，公司为一家非营利性研究机构，2019年接受微软10亿美元的注资，由非营利性机构转为盈利性公司。公司的投资者包括微软、Reid Hoffman's charitable foundation和Khosla Ventures。The Information称公司2021年估值已经达200亿美元。
- 据Semafor报道，微软正商谈以290亿美元估值，向OpenAI投资100亿美元，这笔投资还将引入其他风投公司。2022年公司收入约2500万-4000万美元，以290亿美元的估值测算，市销率达725-1160倍。
- 我们认为，Open AI开发的底层算法系统，虽然在短期有较大的成本投入，但在中长期有望创造巨大的价值，类似移动互联网时期的安卓系统。根据统计，仅仅是截止2016年谷歌共通过Android操作系统获得310亿美元营收和220亿美元利润，收入体量大且盈利能力强（对应净利率为71%）

图：OpenAI发展历程

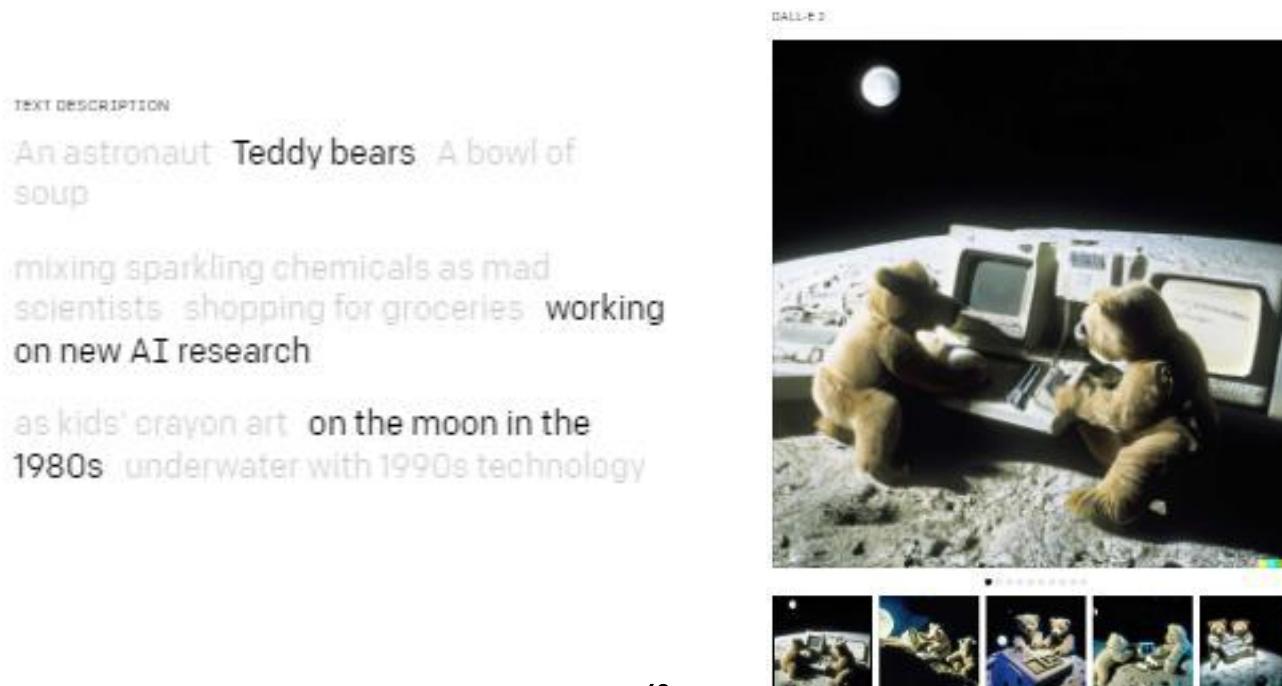


5.2.1 Open AI: 提供底层算法，开发水平强、算法通用性高的AI技术公司（2）

➤ 公司发布了多款底层算法，并推出相关应用API:

- **GPT-3**: 在NLP任务中实现重大突破，能够完成绝大多数NLP任务，入选了《麻省理工科技评论》2021年“十大突破性技术”，目前GPT-3已开源，注册人数超100万，基于该算法已经出现了Jasper、Copy.ai和Copysmith等。
- **CLIP**: 与微软合作发布，能够将图像映射到文本描述的类别中，跨越了文字语义与图像语义的鸿沟，目前已开源，Disco Diffusion、MidJourney均基于CLIP架构及生成算法实现。
- **Codex**: 基于GPT-3，具有将自然语言输入变成代码、代码自动补全等功能，是Github Copilot的底层支持。
- **DALL-E 2**: 基于CLIP和Diffusion算法，实现了有文本生成高质量图片，两个半月注册人数突破100万，公司今年7月正式开始出售DALL-E 2 的图像生成软件许可，目前用户超过150万，每天生成200多万张图片。

图：DALL • E 2由英文生成图像



5.2.1 Open AI: ChatGPT准备推出专业付费版

➤ ChatGPT:

- 强大能力且用户友好，发布5天用户超100万，引起全球关注。ChatGPT不仅在对话、检索、创作、代码等领域展现出强大的能力，还支持直接通过网站试用该文本对话模型，而不是需要专业编程技能试用，且用户界面友好，大大降低了用户门槛，推动了其在全球范围的广泛传播，发布5天用户便突破了100万。
- 计划推出无限制、性能更高的专业付费版本。1月11日，OpenAI 总裁兼联合创始人 Greg Brockman 在 Twitter 上发布 ChatGPT 专业版问卷链接，根据问卷内容，我们预计ChatGPT将推出专业付费版本。
- 据路透社报道，OpenAI预计今年的收入将达到2亿美元，到 2024 年将达到10亿美元。ChatGPT的专业付费版可能是其增收的尝试之一。

图：针对ChatGPT专业版的问卷



Working on a professional version of ChatGPT; will offer higher limits & faster performance. If interested, please join our waitlist here:

A screenshot of a Twitter post from Greg Brockman (@gdb). The post contains text and a link to a Google Doc for joining the waitlist.

Working on a professional version of ChatGPT; will offer higher limits & faster performance. If interested, please join our waitlist here:

[docs.google.com](#)
Waitlist for paid ChatGPT access (Twitter)

ChatGPT Professional (experimental) is geared towards professional use with: Always available (no blackout ...

The Google Doc page has a header "Waitlist for paid ChatGPT access (Twitter)" and a note: "ChatGPT Professional (experimental) is geared towards professional use with: Always available (no blackout windows). Responses from ChatGPT (i.e. no throttling) are selected, and tried to prioritize a balanced process and a fair. Please keep in mind that this is an early experimental program that is subject to change, and we are not making paid pro access generally available at this time."

9:33 AM · Jan 11, 2023 · 1.4M Views

1,043 Retweets 400 Quote Tweets 6,226 Likes

资料来源：OpenAI, 路透社, 中信建投

图：针对ChatGPT专业版的问卷

Waitlist for ChatGPT Professional access (Twitter)

ChatGPT Professional (experimental) is geared towards professional use with:

- Always available (no blackout windows)
- Fast responses from ChatGPT (i.e. no throttling)
- As many messages as you need (at least 2X regular daily limit)

If you are selected, we'll reach out to you to set up a payment process and a pilot. Please keep in mind that this is an early experimental program that is subject to change, and we are not making paid pro access generally available at this time.

At what price (\$ per month) would you consider ChatGPT to be so expensive that * you would not consider buying it?

您的回答

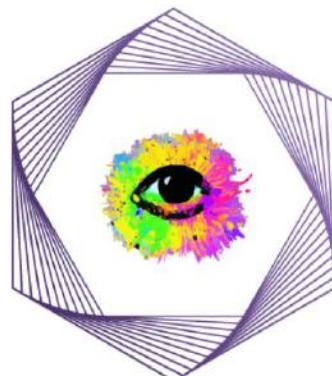
At what price (\$ per month) would you consider ChatGPT to be priced so low that * you would feel the quality couldn't be very good?

您的回答

5.2.2 Stability AI：用户数快速增长、AIGC独角兽

- Stability AI是一家致力于开发开源AI工具的公司。公司由前英国对冲基金经理Emad Mostaque于2020年成立。2022年10月，公司完成1.01亿美元融资，估值10亿美元，投资方包括Coatue、Lightspeed Venture Partners和O'Shaughnessy Ventures。以Stable Diffusion开源，累计使用人数最多，在全渠道有超1000万用户。
- 公司开发的AI工具生产的内容包括图片、语言、音频、视频、3D模型、与生物学相关的内容等。目前已推出的主要为图片AIGC模型和工具，分别为：
 - **Stable Diffusion**：开源的文本转图片AIGC模型。模型由公司和德国海德堡大学的CompVis团队、初创公司RunwayML等共同开发完成。用户可以自由下载使用、调整代码，甚至用于商业化用途。
 - **DreamStudio**：官方推出的基于Stability Diffusion的平台和API。该工具使用户不需要受限于软件安装、编程知识、较高的GPU性能要求。

图：Stability AI推出的AIGC工具



DreamStudio

图片AIGC平台
基于Stable Diffusion模型



Harmonai

音频/音乐AIGC平台
基于Dance Diffusion模型



OpenBioML

AI和生物学结合的开源研究工具
用于DNA测序、蛋白质折叠等机器学习模型

5.2.2 Stability AI: To C业务仍在早期，To B和To G预计贡献主要收入

➤ 为什么公司已提供开源模型，但客户仍有定制化服务需求？

- 开源的Stable Diffusion不提供定制化内容。训练模型的数据来自于包含英文标签的LAION 5b，用户不可以调出、调入新图片，因此不能重新训练模型并实现定制化。
- 用户需求或有差异，需要通过调整模型和训练数据，从而提供定制化服务。不同类型的公司有不同的素材要求，如游戏公司或要求画风偏二次元，广告公司或对格式有一定要求。因此需要公司对模型进行调整、重新训练模型，从而达到客户对AIGC的要求。

➤ 目前Stability AI已与多个政府和学术机构达成合作，同时通过DreamStudio服务普通用户。具体情况如下：

To B & G

已有商业模式：

- 据Forbes，公司已与多个政府和学术机构达成合作，为他们提供基于人工智能的技术服务。

预期：

- 据TechCrunch，公司计划针对企业客户开展定制化的技术服务。

To C

已有商业模式：

- DreamStudio：根据用户对图片像素和操纵步骤需求，每张图片需要消耗0.2-28.2个信用币。新注册的用户可以免费获得200个信用币，此后每100个信用币的费用为1英镑。

预期：

- 我们预计随着公司发展和网站访问量的提升，平台内将会落地更多增值服务，或增加广告业务。

5.2.3 Midjourney：小型独立研究实验室，AIGC偏向绘画作品

- **Midjourney**是一家小型的独立实验室。据官网，团队共由11位成员和多位顾问组成，且自筹资金，暂未有融资记录。团队于2022年7月推出Midjourney，是与Stable Diffusion对标的图片AIGC模型。9月，游戏设计师Jason Allen在美国科罗拉多州博览会的美术比赛中，提交使用Midjourney创作的作品“Théâtre D'opéra Spatial”，在数字艺术/数码摄影单元中获得第一名。Midjourney的具体情况如下：
- **作品类型：**创作作品均是绘画，而非照片。Midjourney尤其擅长创作奇幻、科幻类型的风景照片。据公司创始人David Holz介绍，不排除未来会推出创作写实图片的版本。
 - **服务方式：**通过第三方的Discord服务器提供付费服务。用户创作的图片会被分享至公共社区，一方面可以让新用户快速了解创作指令，另一方面其他用户可以进行二次创作。
 - **商业模式：**主要采用会员费模式。C端用户可以基于使用量的需求，选择不同套餐，每月费用在10-50美元之间；B端用户为固定价格，每年费用为600美元。
 - **用户数量：**据The New York Times，在Discord服务器中已累计有超300万用户。

图表：Jason Allen获奖作品



指令：moody, dark, chiaroscuro

图表：使用Midjourney创作的图片



指令：textured, summertime, warmth, bokeh, lotus on a pond in reflection, cinematic photograph

5.3 除技术领先的OpenAI外，其他垂直应用初创公司仍获高融资

- 在技术应用层面，海外的文本AIGC工具研发商竞争较为激烈，已领先图片AIGC公司，快速实现C端商业化。各文本AIGC工具研发商的业务形态较为相似，产品主要基于开源的OpenAI GPT-3开发，业务主要是为用户提供广告文案、博客文章、社交媒体分享等文本的内容生成服务，用户只需输入简单指令，文本AIGC工具即可生成文本。
 - 该子行业参与者较多，竞争较为激烈。Jasper是一家为主要为营销从业人员、社交媒体运营人员等提供基于AI的文本生成工具的研发商，其他同类公司还包括Copy.ai、Copysmith、WriteSonic、Peppertype等。
- 近年多家公司获得高额融资。2022年10月，Jasper完成1.25亿美元融资，估值达15亿美元，投资者包括Coatue、Bessemer Venture Partners、IVP等多家机构。其他同样基于OpenAI GPT-3、业务相近的初创公司获得较高融资，包括Copy.ai、Copysmith等。
- 由此可见，海外初创公司利用OpenAI、Stability AI等公司开发的开源技术，开发应用并开展To C业务，仍可获得一定营收规模，并获得高额融资。大型AI公司在投入人力资源、算力资源开发大规模底层算法方面优势明显，而中小型公司在垂直应用、客户资源、灵活适配等方面具有一定优势，可与大型公司合作获得算法支持在细分领域改进算法，提升用户体验，做出成绩。因此预计国内中小型公司在头部互联网公司入局后，仍能在应用层面获得发展空间。

图表：Jasper、Copy.ai、Copysmith近期融资情况

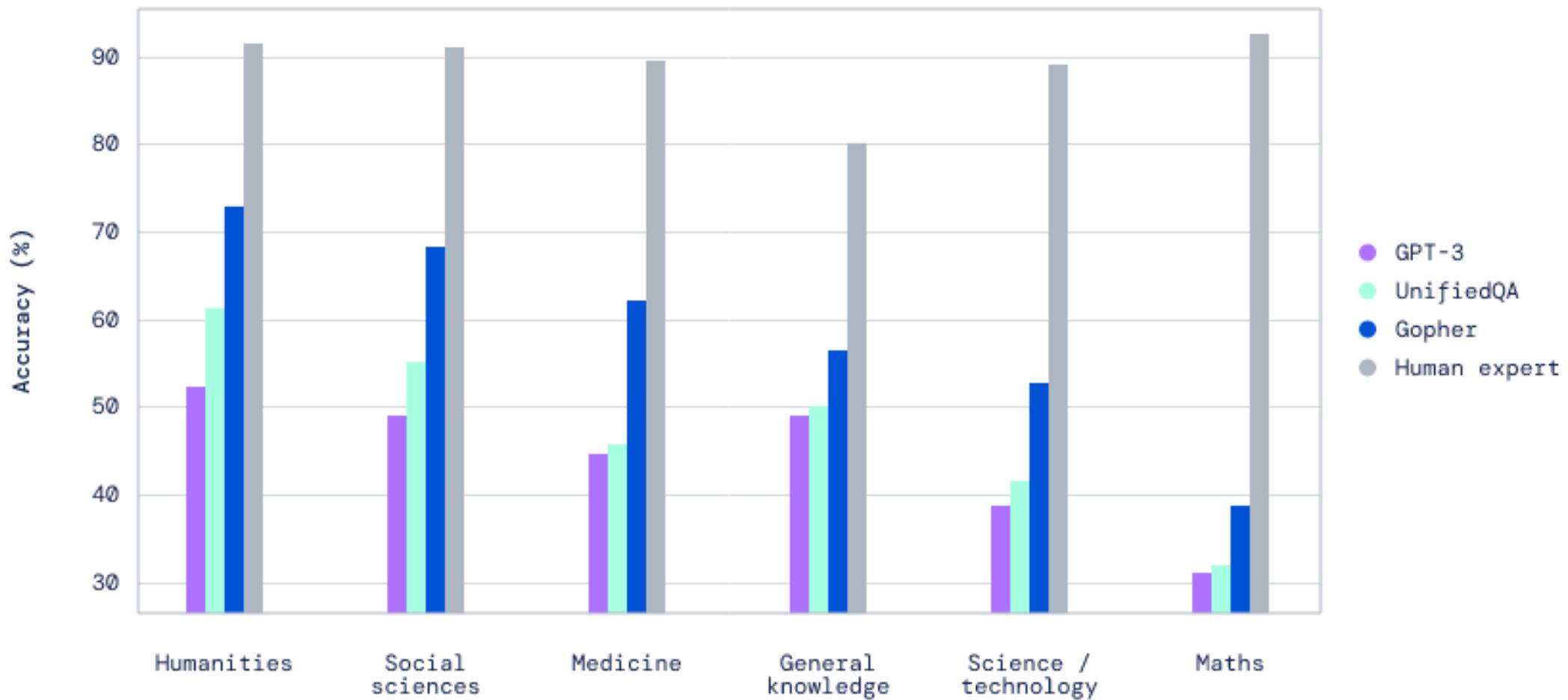
	Jasper	Copy.ai	Copysmith
成立时间	2021	2020	2020
底层AIGC模型	OpenAI GPT-3	OpenAI GPT-3	OpenAI GPT-3
融资时间	2022. 10	2021. 10	2021. 04
融资金额	1. 25亿美元	1100万美元	1000万美元
估值	15亿美元	-	-
投资机构	Coatue、Bessemer Venture Partners、IVP等	Harmony Venture Labs等	Wing Venture Capital、红杉、老虎全球等
2021年营业收入	4500万美元	240万美元	-

资料来源：TechCrunch, Harvard Business Review, VentureBeat, ContentGrip., 中信建投

5.4 谷歌：DeepMind某些领域表现优于OpenAI

- DeepMind在2021年发布了拥有2800亿参数的大语言模型Gopher， Gopher在人文科学、社会科学、药学、通用知识、科学技术、数学等细分类别的大规模多任务语言理解基准性能测试中表现优于GPT-3。

图：Gopher表现



资料来源：DeepMind, 中信建投

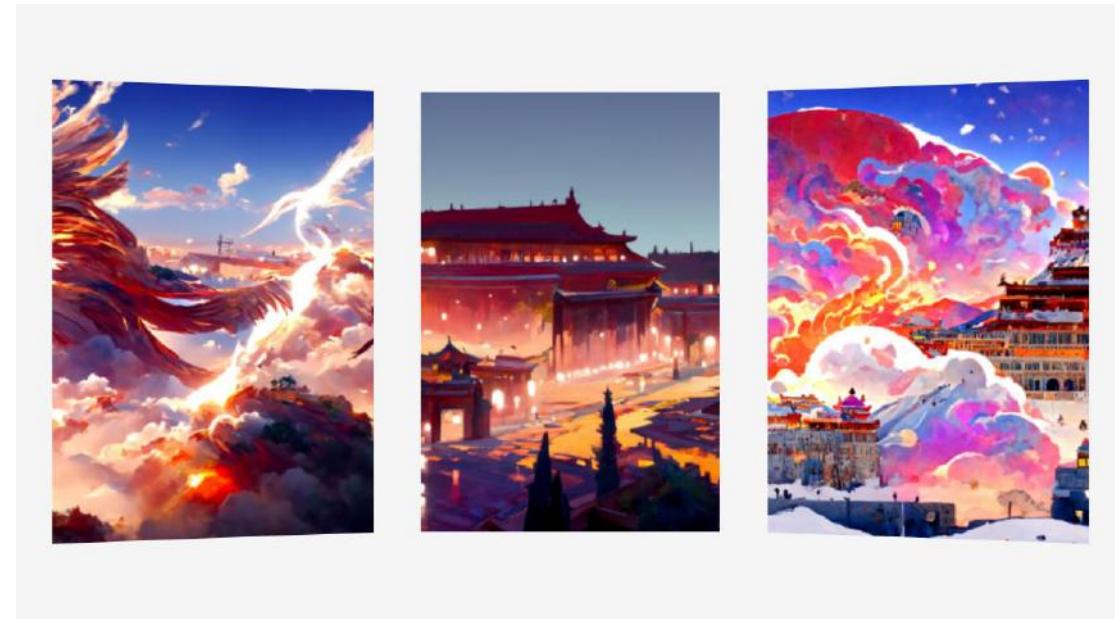
5.5 百度：今年已推出多个基于AIGC的数字人、“AI作画”平台等产品

- 百度在AIGC领域已有较多布局。自2020年9月15日百度推出首个虚拟偶像度晓晓以来，目前已有多位数字人，包括2月推出的首个实现AIGC的数字虚拟偶像希加加。今年百度的AIGC相关产品、数字人情况如下：
 - **AIGC技术创作高考作文：**度晓晓挑战高考作文，利用AIGC能力，在40秒完成40篇作文。
 - **“AI作画”文心·一格：**8月19日，百度在成都CCIG 2022大会发布AI艺术和创意辅助平台文心·一格，即系统通过用户输入中文关键词创作图片。
 - **百度输入法正式推出国内首款情感陪伴型虚拟博主林开开和叶悠悠。**8月17日，两位博主将上线百度输入法的AI侃侃功能，为用户提供24小时在线陪伴对话，可支持文字、语音、表情包聊天互动。
 - **百度的数字人家族在快手发布首个入驻视频。**11月15日，百度数字人家族四位成员分别为希加加、度晓晓、林开开、叶悠悠。未来数字人家族将参与一系列活动，包括加入爱奇艺出品的闯关游戏真人秀《元音大冒险》，以及在快手进行24小时AI直播等。

图表：百度数字人家族入驻快手



图表：用户在“AI作画”平台文心·一格上创作的国风作品



5.5 百度：技术积累构建AI基础设施

➤ 百度通过多年在人工智能领域的高投入，形成技术积累。据11月9日李彦宏在“2022联想创新科技大会”时的介绍，百度过去十年累计投入研发资金超过1000亿元，连续四年在AI专利申请量和授权量上保持国内第一，凭借丰富技术积累推出了文心大模型。

图：百度文心大模型



资料来源：百度，中信建投

5.6 小冰公司：参与底层AIGC技术研发，赋予虚拟人创作能力

- 前微软团队独立出的小冰公司完成10亿元融资。据科创板日报，11月7日，小冰公司宣布完成本次A+轮融资，投资方为高瓴投资、IDG资本，将用于加速AI Being小冰框架技术研发，推动数字员工普及。公司的业务情况如下：
 - 小冰公司目前运营多位**B端**和**C端AI Being**。AI Being是公司对虚拟人的称呼，目前已在运营的AI Being包括红杉中国虚拟分析师“Hóng”、虚拟歌手“洛天依”、万科集团数字员工“崔筱盼”等。
 - **AIGC技术是AI Being获得创造力的重要部分**。根据公司公告，近期其将对人工智能数字员工（AI Being Employee）产品线升级，包括大模型对话引擎、3D神经网络渲染、超级自然语音、AIGC人工智能内容生成。其中部分新技术已提前在招商局集团“招小影”、冬奥数字教练“观君”等数字员工中交付。
 - 公司推出的虚拟人以面向**B端**为主，预计**B端**贡献主要收入，但随着虚拟偶像的发展，预计**C端**收入将逐渐提升。

图表：2022年7月，人工智能小冰框架与Vsinger达成合作，启动为虚拟歌手洛天依定制的歌声、语音多个AI声库的制作企划，通过AIGC技术赋予虚拟人创作能力，完成从虚拟歌手向真正的AI Being的转变



图表：小冰公司融资历史

融资时间	轮次	金额	投后估值	投资方
2020. 11. 24	Pre-A轮	数亿元	50亿元	北极光创投、网易
2021. 7. 12	A轮	数亿元	70亿元	高瓴投资、五源资本、Neumann Advisors、IDG资本、GGV纪源资本、北极光创投、网易
2022. 11. 7	A+轮	10亿元	150亿元	高瓴投资、IDG资本

资料来源：IT桔子，科创板日报，澎湃新闻，wind，中信建投

5.7 其他互联网公司：积极参与底层技术研发，

➤ 国内头部互联网公司正在积极布局AIGC业务。腾讯、阿里巴巴、网易、字节跳动等国内头部互联网公司及旗下人工智能实验室正在基于自身业务情况，研发文本、图片、音乐等内容的AIGC技术和工具。其中，阿里巴巴的部分技术已运用至产品中，实现提升用户使用体验的目的；腾讯将AIGC的底层技术应用至游戏中，基于自有产品的玩家的操作数据，模拟真人玩家操作，预计未来也可用于创作虚拟玩家。各公司AIGC布局具体情况如下：

公司	业务进展
腾讯	<ul style="list-style-type: none">文本AIGC：腾讯人工智能实验室开发了一款创作辅助工具Effidit，提供智能纠错、句子补全、短语润色、例句推荐、论文搜索、短语补全以及云输入法功能，有效提高写作的水平质量。游戏：该实验室基于自有多模态学习及生成能力，在游戏领域进行全流程布局。以“绝悟”为代表，强化学习的方法来模仿真实玩家，包括发育、运营、协作等指标类别，以及每分钟手速、技能释放频率、命中率、击杀数等具体参数，让AI更接近正式服玩家真实表现。
阿里巴巴	<ul style="list-style-type: none">图片AIGC：阿里巴巴智能设计实验室自主研发的AI视觉物料生成系统鹿班。基于图像智能生成技术，鹿班可以在短时间内完成大量banner图、海报图和会场图的设计，提高工作效率。主要功能包括了一键生成海报、智能排版、设计拓展、智能创作等。图片AIGC：2020年12月，阿里巴巴在原创保护平台推出虚拟模特“塔玑”。塔玑可以利用算法生成“独一无二的虚拟人脸”，商家只要提交服装平铺图，就可以选择喜欢风格、人脸、表情、妆容、身型、背景图，让模特“穿上新款”，从提交完成到生成模特图，最快仅需1天。
网易	<ul style="list-style-type: none">音乐AIGC：1月25日，网易推出一站式音乐创作平台网易天音。用户输入祝福对象，选择祝福语，就可以自动生成拜年歌曲。用户可以更换歌词、伴奏、歌手或重写，在设定歌名、选择模板、点击发布后，将作品分享给好友或分享到朋友圈。
字节跳动	<ul style="list-style-type: none">智能创作团队：定位为AI&多媒体技术中台，在计算机视觉、音视频拍摄编辑、特效处理、语音等角度支持抖音、TikTok、剪映、西瓜等公司内众多产品线。

目 录

第一章

AI：从判别决策到创造生成

第二章

生成式AI：范式升级，应用多元

第三章

生成式AI应用：内容制作，从辅助人到“替代”人

第四章

生成式AI应用：多行业垂直应用

第五章

海内外发展：多家科技巨头、初创公司积极布局

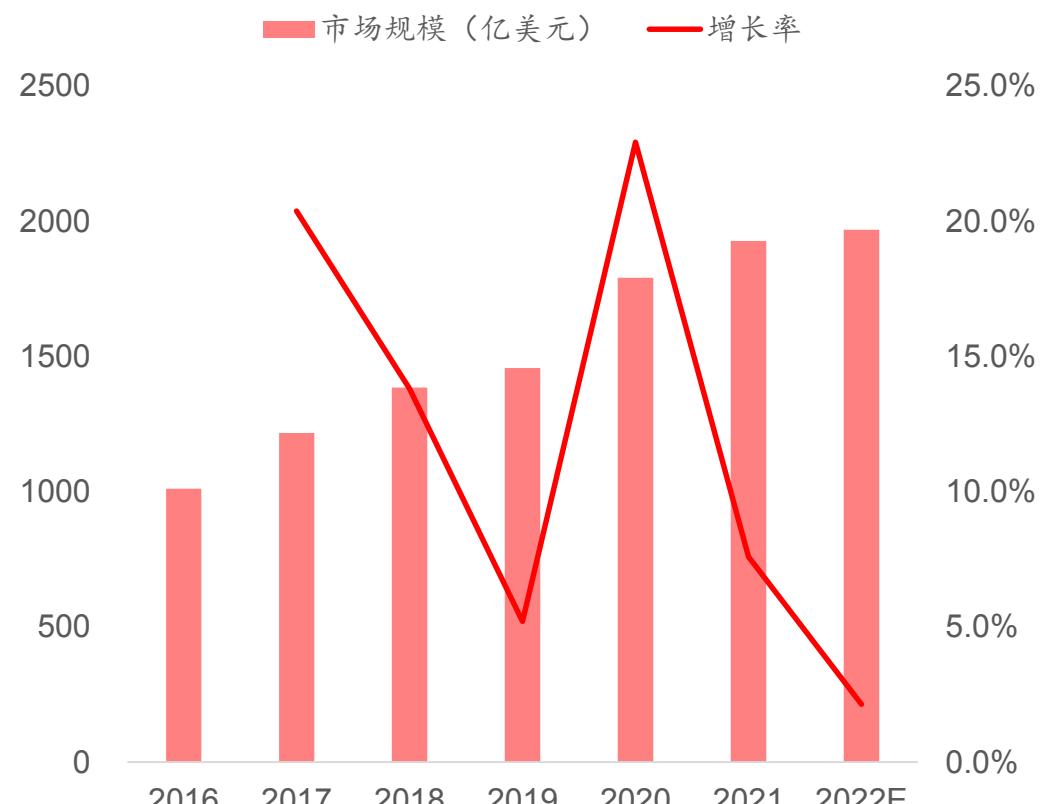
第六章

综合应用：游戏行业的案例

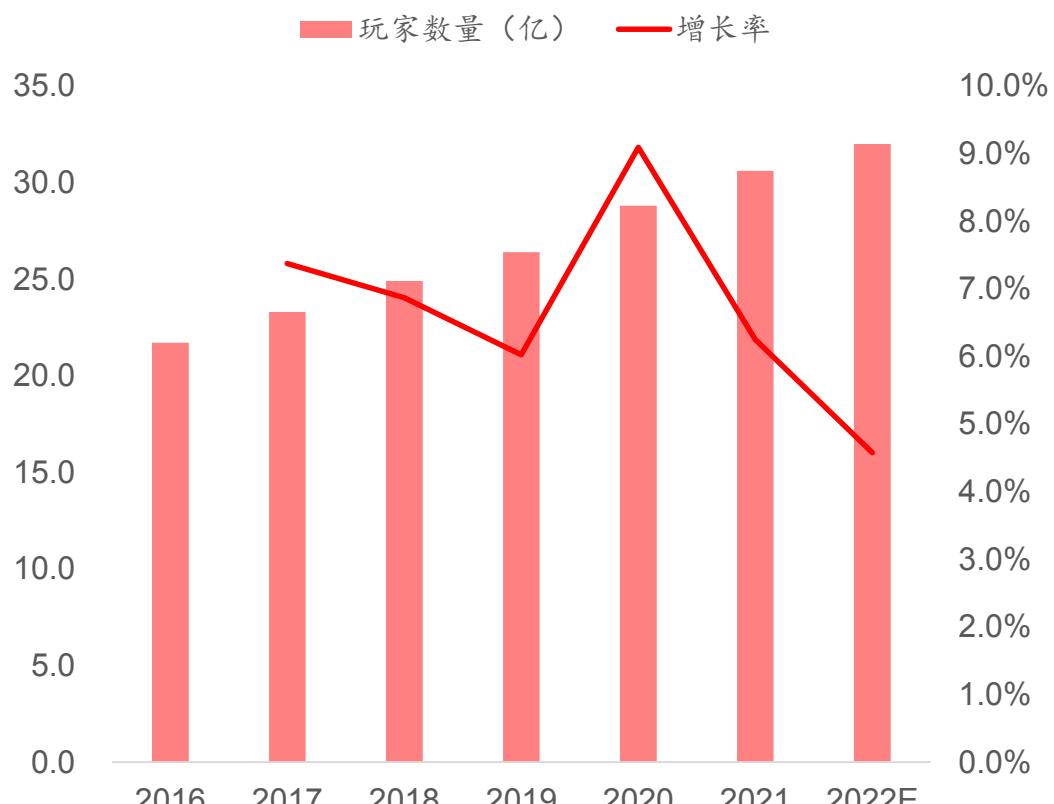
6.1 游戏：市场规模近2000亿美元，玩家数量达30.6亿近全球总人口的40%

- 全球游戏市场规模不断扩大，预计2022年将达1968亿美元。据Newzoo，全球游戏市场规模由2016年的1011亿美元增长到2021年的1927亿美元，五年市场规模增长近一倍，预计2022年将达1968亿美元，2025年将增长至2257亿美元。
- 游戏玩家数量不断增长，预计2022年达到32亿，占全球总人口的40%。据Newzoo，2021年全球游戏玩家数量为30.6亿人，预计2022年将达32亿人，以目前全球总人口80亿计算，游戏玩家占全球总人口的比例达到40%。预计2025年全球游戏玩家数量将增长至35.3亿人，2020-2025年复合增长率为4.2%。

图：全球游戏市场规模及增长率



图：全球游戏玩家及增长率

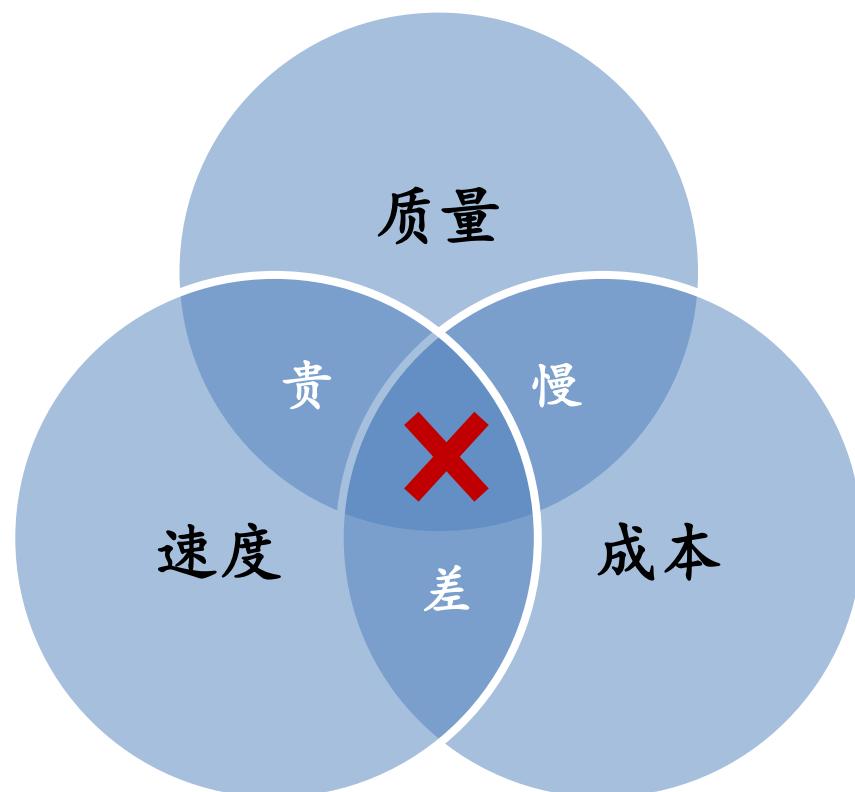


资料来源：Newzoo、中信建投

6.2 游戏制作：游戏产业的基础，受不可能三角的限制

- 游戏行业的竞争主要集中在制作环节和营销环节，优良的游戏制作是营销的基础。
- 传统游戏制作存在“质量、速度、成本”中只能有两个的不可能三角。在保证游戏质量的前提下，游戏制作只有两条路径可选，一是大量开发人员同时制作一款游戏，而大型团队管理、开发效率较低，开发成本势必提高；二是精简团队成员，制作周期长达几年甚至十几年。
- 通常优质3A大作的开发时间成本和金钱成本均很高。以销量超4600万份的《荒野大镖客：救赎 2》为例，游戏拥有超过28平方英里接近真实景象的地图和1000个NPC，即使专职开发人员超1200人，也用了8年完成，成本近3亿美元。专职开发人员少于《荒野大镖客：救赎 2》的《星际公民》，已开发10年之久，仍处于Alpha版本测试阶段，预计还需要较长时间才能完整上线。

图：传统游戏制作的不可能三角



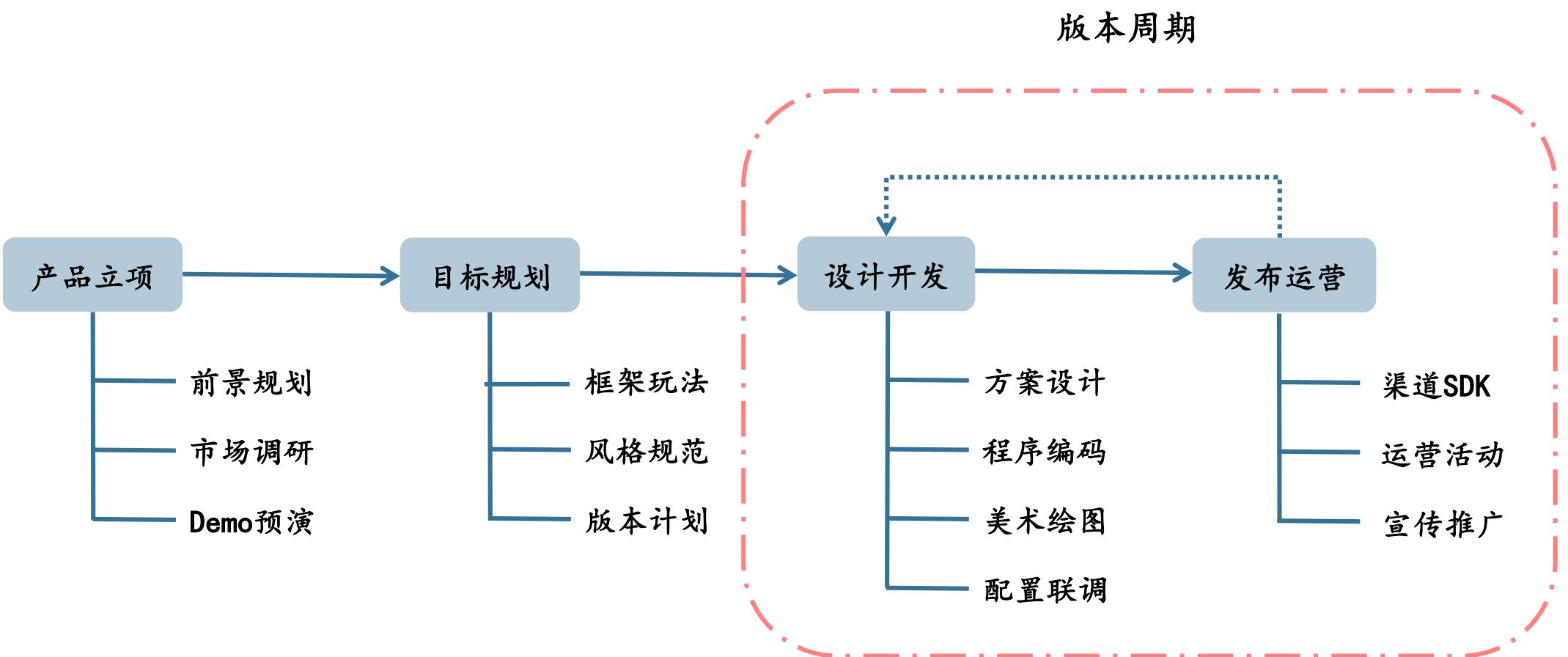
表：大型3A游戏规模及制作成本

游戏	荒野大镖客：救赎 2	星际公民
地图规模	28平方英里	陆地面积 超100万平方公里
NPC数量	超1000个	-
发售时间	2018年	暂未发售
制作周期	8年	10年以上
贡献人员	超3000人	151
开发人员 (不含外包)	约1200人	已超700人
开发成本 (不含营销)	2.4-3亿美元	已众筹5.3亿美元
销量	4600万份	-

6.2 游戏制作：流程复杂，要素齐全

➤ 游戏制作流程复杂，设计开发环节最为关键。游戏的制作一般包括“产品立项-目标规划-设计开发-发布运营”四个环节，制作周期一般以年为单位，制作流程的复杂程度和时间成本远超娱乐媒体领域的其他内容。在游戏的制作流程中，设计开发环节决定了规划的落地程度，极大程度上影响游戏质量以及玩家体验，是最为关键的一环。

图：游戏制作流程



6.2 游戏制作：流程复杂，要素齐全

- 游戏是娱乐及媒体行业中最复杂的形式，构成要素包含娱乐媒体领域内容的所有形式。游戏包含文本、图像、音效、音乐、3D模型、动画、电影、代码等多种类型的资源，涵盖了娱乐及媒体行业所有的内容形式。

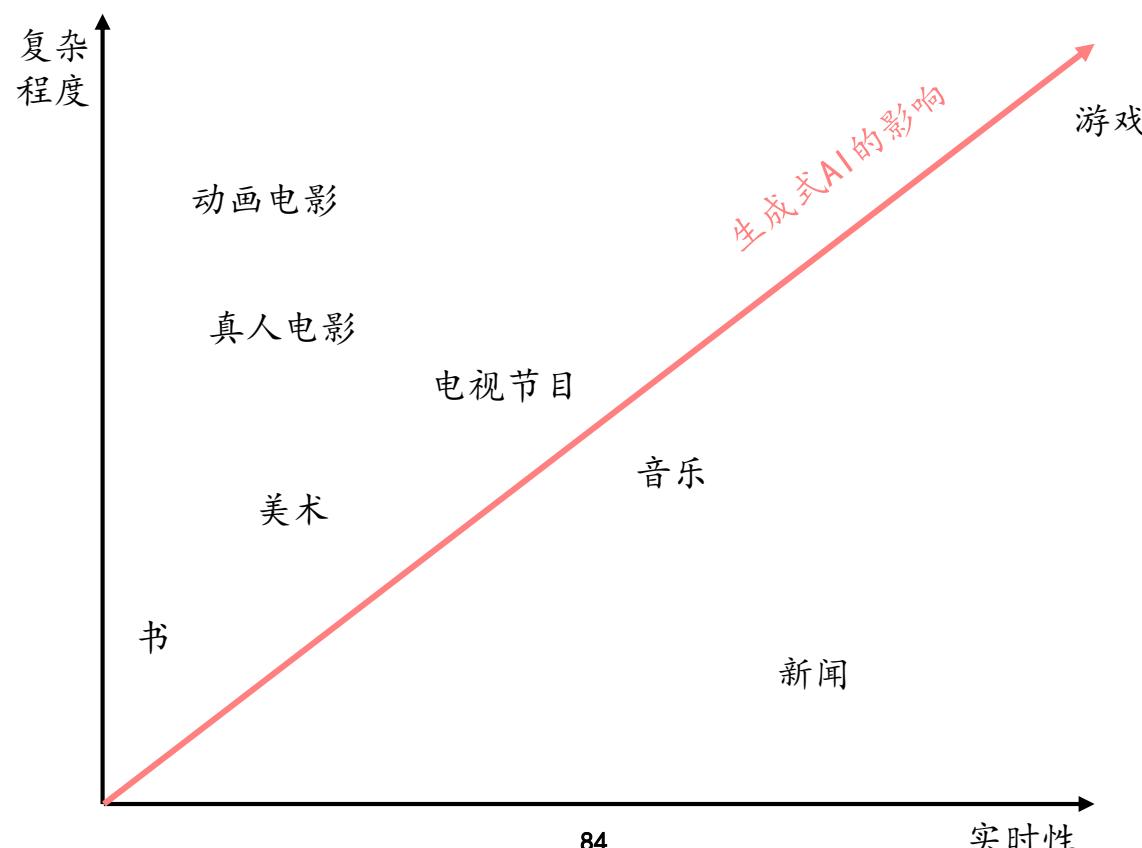
图：游戏构成要素



6.3 生成式AI将影响游戏制作，有望打破不可能三角

- 游戏是娱乐及媒体行业中对实时交互要求最高的形式。游戏强调玩家交互的实时体验，要求即时反馈，是所有娱乐形式中综合复杂程度最高的。游戏的要素复杂、制作门槛高也决定了生成式AI对游戏制作的影响是其他娱乐及媒体领域的合力，创造的机会也是最大的。
- 生成式AI带来的生产力提升，将打破游戏制作的不可能三角。生成式AI能够根据已有文本进行续写、根据文本生成语音、根据主题生成音乐、根据文本生成图像、根据二维图像生成三维模型、根据文本生成代码等，对游戏的策划、音频、美术、程序等环节的工作均带来极大的生产力提升，大幅降低游戏制作成本，打破不可能三角。以二维游戏原画为例，传统原画师根据策划内容完成设计稿需要花费数周时间，而使用生成式AI仅需要输入关键词便能得到大量设计初稿，后续进行筛选、修改即可，时间降低至几个小时。

图：生成式AI对游戏制作的影响是所有娱乐媒体内容领域最大的



资料来源：vcgamers, a16z, 中信建投

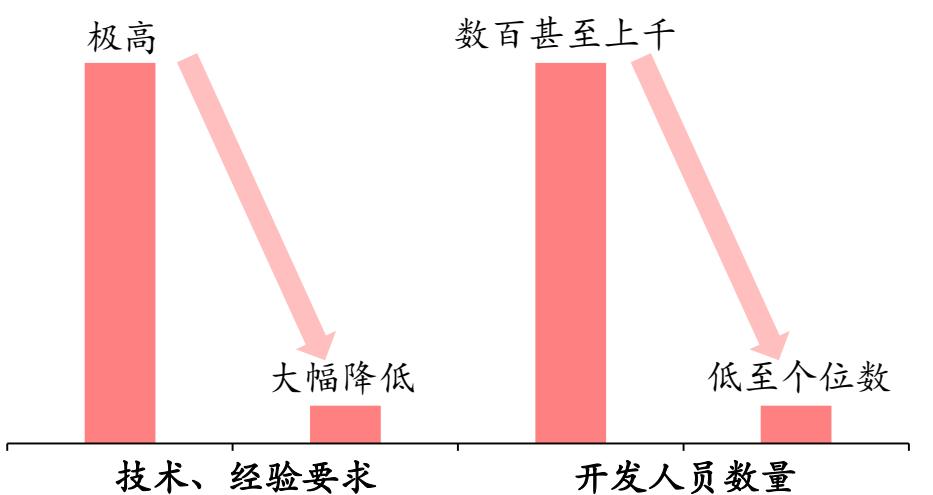
6.4 生成式AI突破专业技术、经验的限制，大幅降低游戏制作门槛

- 游戏的综合复杂程度决定游戏开发具有很高的专业技术、经验门槛，资金投入和传统AI算法积累的重要性降低。以字节跳动为例，字节2018年组建大中型游戏研发团队，公司在推荐算法、用户数据积累、资金、流量渠道等方面的优势未能让其在游戏制作领域复刻抖音、今日头条的成功，即使投入了数百亿资金，虽然推出了表现不错的《航海王热血航线》，但仍未有自研S级游戏面世。我们认为主要原因是公司后期进入游戏市场，缺乏游戏制作经验丰富、技术专精的管理人员和技术人员，据报道，公司试图以技术优势建立游戏中台加速游戏研发，但其开发人员对游戏不够熟悉导致中台偏代码思维，在便利性、使用效果方面不及网易的编辑器。
- 生成式AI制作游戏内容价格很低且不需要大量的专业技术人员，大幅降低游戏制作门槛，微型游戏制作工作室走进现实，同时缩短游戏制作周期。生成式AI制作游戏内容具有一下特点：1. 开发调试能够多次长期复用；2. 并有专业生成式模型开发公司提供B端产品，将设计人员的创意转化设计初稿；3. 大幅度降低内容制作的时间和金钱成本，同时减少对专业技术人员的需求，微型游戏制作工作室成为现实。5人组成的Innersloth工作室开发出了《Among Us》，Steam销量超2000万，全球玩家约5亿人。随着生成式AI的不断发展，微型游戏制作工作室的数量和能够开发的游戏规模将不断上升。

图：5人开发的《Among Us》



图：生成式AI降低游戏制作门槛



6.5 游戏创新：AI生成内容降本增效，产能提升实现既有创意

- 游戏的创新表现在多个方面，主要受到创意和产能两个因素的限制。策划是游戏的起点，美术设计是到游戏形象、画面的关键，游戏创新依赖于策划、美术人员的创意，而从创意到成品要经过漫长的开发过程，产能也是限制游戏创新的关键一环。
- 生成式AI提高生产力突破产能限制实现游戏创新：
 - AI生成3D模型助力飞行模拟器实现3D高清全球地图：微软的飞行模拟器计划开始于1976年，一直在向着高清、3D、全球地图努力，最新版本通过与blackshark.ai合作，使用AI从2D卫星图像生成3D模型，构建了超过15亿的3D建筑模型，突破了产能的限制，飞行区域包括整个地球、超过24000个机场，且主要地标和热门城市的地景越来越丰富。
 - AI赋能NPC打造原神高品质开放世界：开放世界游戏在上世纪80年代便出现，原神开发中使用AI赋予了每个NPC不同的个性及特殊能力，打造了更为完备的高品质开放世界，获得成功。

图：飞行模拟器真实3D地图



图：原神



6.5 游戏创新：AI创作突破个人创意局限，拓宽策划设计边界

- 生成式AI通过转型性创作、组合性创作等方式突破个人创意的局限性，拓宽游戏策划设计的边界：
- 继续性创作：学习部分完成的要素的特征，进行继续性创作，最常见的是AI续写文案、生成同一风格系列的图片等，适用于游戏策划、设计完成部分但是想象力枯竭的情况，能够通过大量继续性创作扩展人的想象力。
- 组合性创作：将不同领域、不同维度的原本无关的概念元素进行组合，创造出新的概念元素，能对游戏的关卡设计、地图设计、玩法设计、原画设计等提供全新的思路。目前图像生成算法能够根据输入的文字要素生成各种风格的图像以及超出想象的超现实动物、植物、人物形象，为游戏原画设计打开了想象空间。2021年，Anurag Sarkar和Seth Cooper使用变分自编码器（Variational auto-encoder）将不同的游戏关卡进行组合，创造出新的游戏关卡。

图：DALL·E2基于同一要素风格生成的图像



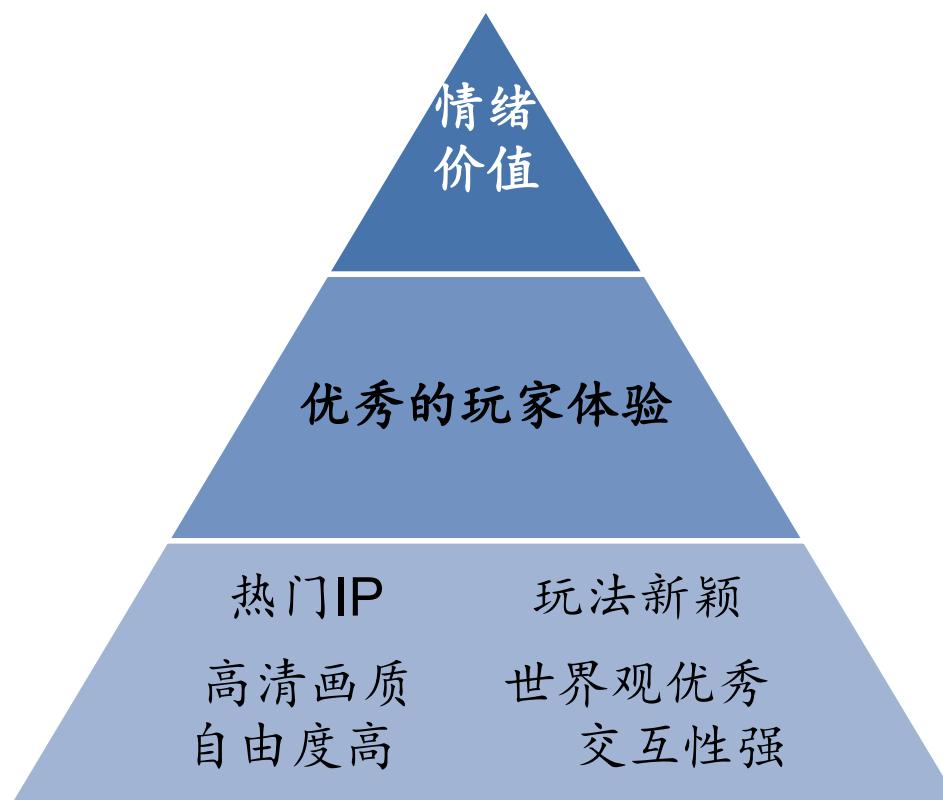
图：VAE生成的一个融合了塞尔达、银河战士、洛克人的关卡



6.6 展望：短期内游戏数量、玩法品类增加，长期有望出现定制化游戏

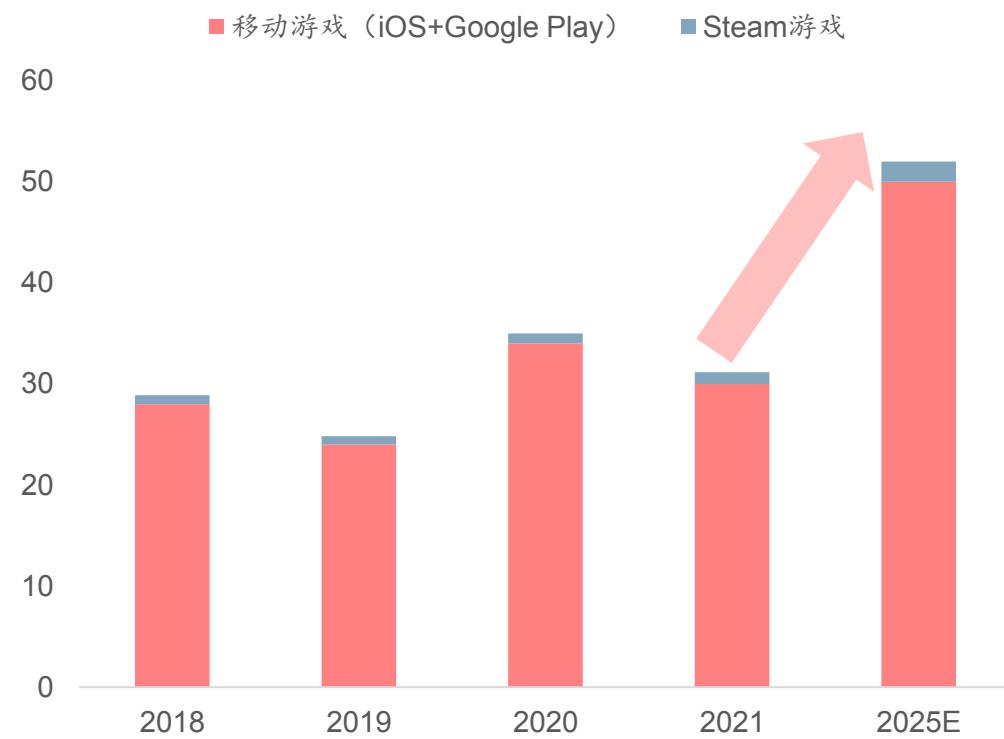
- 游戏成功与否的核心是玩家体验，本质因素则是为玩家带来的情绪价值。从表层看，决定游戏成功与否的关键因素包括游戏的IP、玩法、数值体系、画质、世界观、自由度、交互形式等，但这些因素仅是玩家体验的外在集合，优秀的玩家体验是一款游戏成功的核心，而游戏满足玩家情绪需求带来的价值是玩家体验的本质。
- 生成式AI有望带来游戏数量和玩法品类的增加，增加玩家的选择空间，带来更好的游戏体验，长期看可能出现玩家根据自身经历、情绪状态制作的个人专属游戏，深度满足情绪需求。2021年，全球iOS+Google Play上线移动游戏30万款，Steam上新发行游戏11585款，截止2022年11月29日，Steam上2022年发行的游戏数量已经超过2021年的总和。预计未来全球新游戏发行数量将持续增加。

图：游戏成功的关键因素



资料来源：Steamdb。游戏设计理论，中信建投

图：新发行的游戏数量（万款）



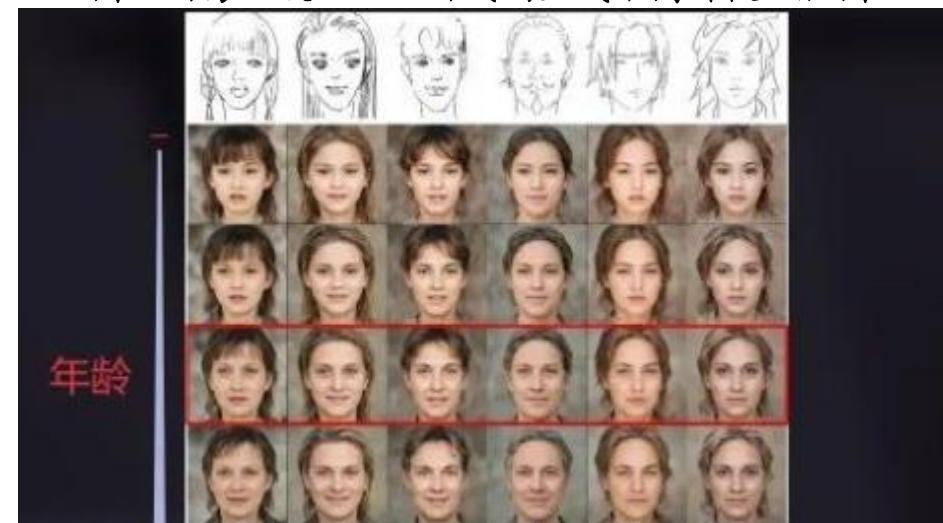
6.7 落地形式：图像、3D模型、动画、音频、NPC、玩家角色等均有落地（1）

➤ 图像：网易互娱AI Lab已经实现生成式AI辅助原画线稿上色、基于人脸线稿生成图像并编辑；Runway、Open AI等实现了有文本生成图像、拓展图像、为图像增加/替换/删除元素等。

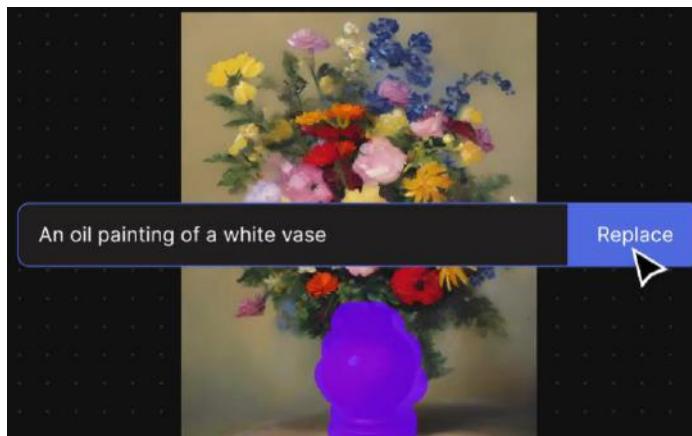
图：网易互娱AI Lab线稿上色



图：网易互娱AI Lab由线稿生成不同阶段人脸图



图：Runway由文本生成图像及为图像替换元素



Beautiful flowers in a vase, oil painting

Generate

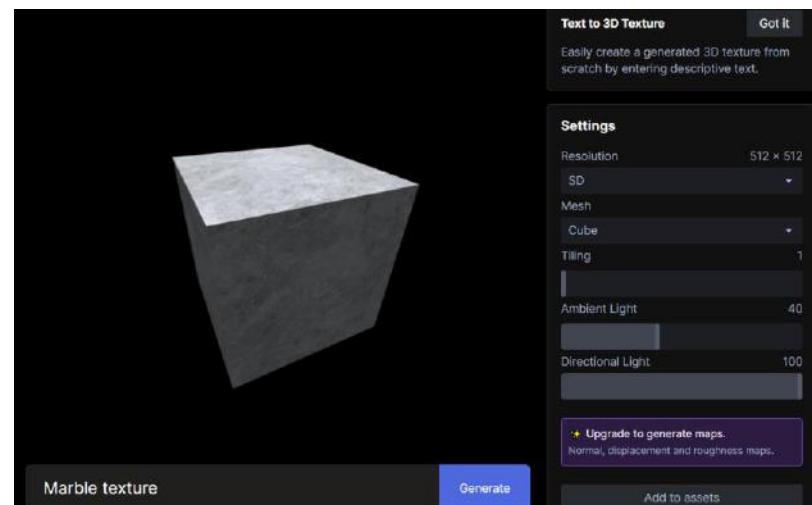
6.7 落地形式：图像、3D模型、动画、音频、NPC、玩家角色等均有落地（2）

- **3D模型：**网易互娱AI Lab通过人脸图像生成人头的3D模型，批量为游戏生成人头模型；Runway通过文本为3D模型添加纹理；Open AI开源3D模型生成器Point-E实现由文字生成3D模型，且可以在单块 Nvidia V100 GPU上在一到两分钟内生成3D模型，速度大大提升。

图：网易互娱AI Lab由图像生成3D模型



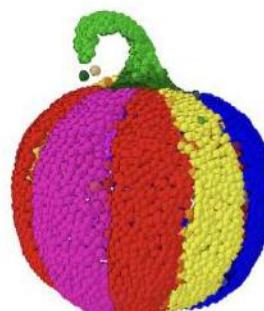
图：Runway通过文本为3D模型添加纹理



图：Open AI由文字生成3D模型



“a corgi wearing a red santa hat”



“a multicolored rainbow pumpkin”



“an elaborate fountain”



“a traffic cone”

6.7 落地形式：图像、3D模型、动画、音频、NPC、玩家角色等均有落地（3）

- 动画：网易互娱AI Lab实现了由视频、音频生成动画，大幅降低了动画捕捉成本；Runway实现了去除视频背景，为动画制作背景替换带来便利。

图：网易互娱AI Lab通过视频、音频生成动画



图：Runway去除视频背景

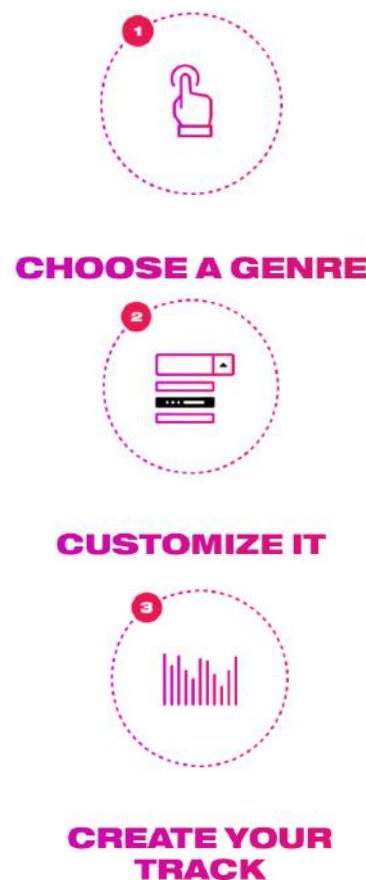


资料来源：Runway, 网易互娱, 中信建投

6.7 落地形式：图像、3D模型、动画、音频、NPC、玩家角色等均有落地（4）

- 音频：**Soundful**实现了通过用户自定义输入生成音乐。
- NPC和玩家角色：小冰成立的游戏工作室通过**AI**技术完成虚拟人的创建、驱动和内容生成，并具备感知、表达等无需人工干预的自动交互能力，降低**NPC**和游戏角色创建成本并带来更多的交互性和细节。

图：Soundful音乐生成



图：小冰虚拟人



6.8 实测效果：二次元主题游戏原画设计及生成

- 生成式AI使得没有任何游戏设计或者经验的人能够制作出游戏原画。ChatGPT强大的AI问答功能，能够在输入要求后给出文字创意，将文字输入DALL-E 2后生成游戏原画，全过程无需专业设计或者美术技能。
- 二次元主题游戏原画：ChatGPT根据需求给出“一个城市的风景，摩天大楼林立，交通繁忙。一个少年手持古剑，身着学生制服，神态憧憬。”的文字设计，将第一部分文字设计输入DALL-E 2即可生成如下游戏原画。

图：二次元主题游戏原画文字创意

我想要在DALL-E 2中生成一幅二次元主题游戏的原画，详细描述我应该输入的关键词

如果您想要在 DALL-E 2 中生成一幅二次元主题游戏的原画，您可以输入以下文本描述：

- 二次元游戏原画：一个城市的风景，摩天大楼林立，交通繁忙。一个少年手持古剑，身着学生制服，神态憧憬。
- 日式游戏原画：一个田园的景色，绿树成荫，花香四溢。一个少女身着和服，手持拂尘，神态娴静。
- 动漫游戏原画：一个森林的背景，枝叶茂密，鸟语花香。一个少女手持古剑，身着紫衣，神态迷人。
- 游戏角色原画：一个帅气的仙侠男子，身着黑色龙袍，手持紫金玉杖，脸上露出坚定的神情。
- 游戏背景原画：一个仙侠世

图：二次元主题游戏原画



6.8 实测效果：中国传统仙侠主题游戏原画设计及生成

➤ **仙侠主题游戏原画：**ChatGPT根据需求给出“中国传统仙侠游戏原画:一个仙侠世界的风景，山川清晰可辨，云雾缭绕，一个仙女站在山顶上，手持青龙偃月刀，脸上露出坚定的表情；古典仙侠游戏原画:一个仙侠世界的景色，碧水蓝天，翠竹绿草，一个道士身着青衫，手持扇子，神情淡然；二次元仙侠游戏原画:一个仙侠世界的背景，月亮悬挂在天空，星辰闪烁，一个美少女手持古剑，身着唐装，神态迷人。”等文字设计，将第一部分文字设计输入DALL-E 2即可生成如下游戏原画。

图：中国传统仙侠主题游戏原画文字创意

我想要在DALL-E 2中生成一幅中国传统仙侠主题游戏的原画，详细描述我应该输入的关键词



如果您想要在 DALL-E 2 中生成一幅中国传统仙侠主题游戏的原画，您可以输入以下文本描述：

- 中国传统仙侠游戏原画：一个仙侠世界的风景，山川清晰可辨，云雾缭绕。一个仙女站在山顶上，手持青龙偃月刀，脸上露出坚定的表情。
- 古典仙侠游戏原画：一个仙侠世界的景色，碧水蓝天，翠竹绿草。一个道士身着青衫，手持扇子，神情淡然。
- 二次元仙侠游戏原画：一个仙侠世界的背景，月亮悬挂在天空，星辰闪烁。一个美少女手持古剑，身着唐装，神态迷人。

图：中国传统仙侠主题游戏原画



6.8 实测效果：山海经异兽主题卡牌游戏原画设计及生成

- **山海经异兽主题卡牌游戏原画：**ChatGPT根据需求给出“山海经异兽卡牌游戏原画:一个古代神话世界的景色，山峰险峻，河流深邃，一个异兽在山间空中盘旋，手持古剑，神态雄健；古典神话卡牌游戏原画:一个古代神话世界的背景，云雾缭绕，水鸟飞翔，一个神仙站在天涯海角，手持权杖，神态慈祥。”等文字设计，将第一部分文字设计输入DALL-E 2即可生成如下游戏原画。
- 普通生成式AI创作的上述游戏原画仍存在部分瑕疵且画质需要提升，但已经能够为游戏策划人员、美术人员提供足够的创意及素材，随着技术的发展，针对游戏制作的专业生成式AI将带来更多的创意、素材，极大提升游戏制作的生产力。

图：山海经异兽主题卡牌游戏原画文字创意

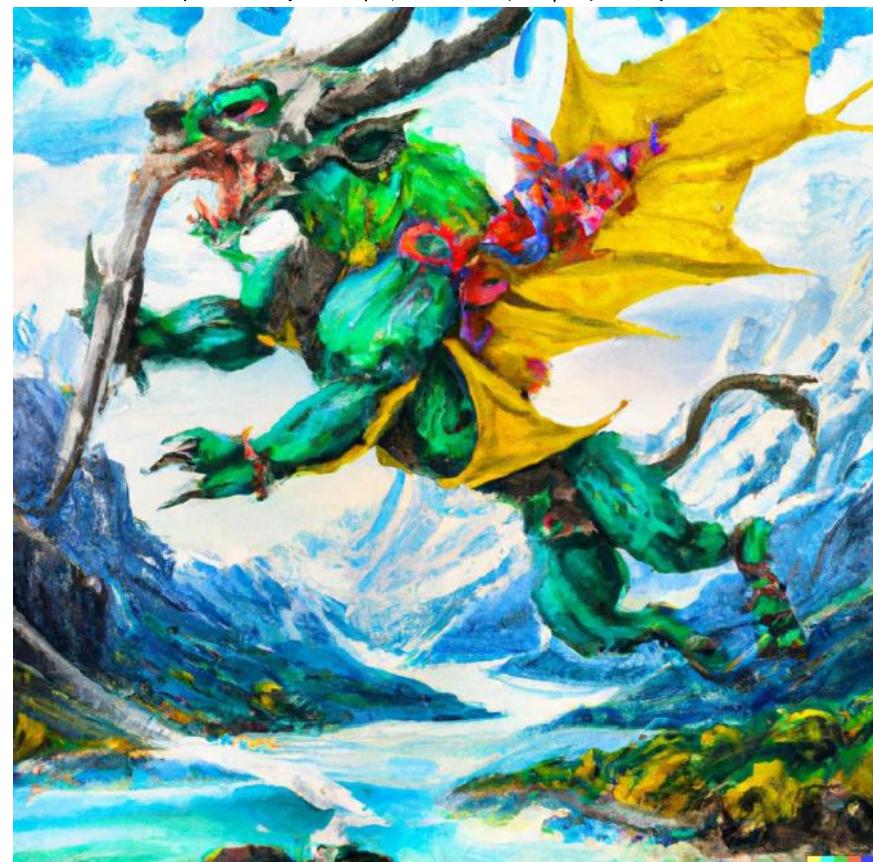
我想要在DALL-E 2中生成一幅山海经异兽主题的卡牌游戏的原画，详细描述我应该输入的关键词

如果您想要在 DALL-E 2 中生成一幅山海经异兽主题的卡牌游戏的原画，您可以输入以下文本描述：

- 山海经异兽卡牌游戏原画：一个古代神话世界的景色，山峰险峻，河流深邃。一个异兽在山间空中盘旋，手持古剑，神态雄健。
- 古典神话卡牌游戏原画：一个古代神话世界的背景，云雾缭绕，水鸟飞翔。一个神仙站在天涯海角，手持权杖，神态慈祥。
- 二次元异兽卡牌游戏原画：一个古代神话世界的景色，山川清晰可辨，云雾缭绕。一个异兽女神身着紫衣，手持银灯，神态娴静。
- 异兽角色原画：一个异兽男子，身着红衣，手持紫金玉枚，脸上露出坚定的神情。
- 异兽背景原画：一个古代神话世界的景色，山峰秀丽，溪流清澈。一座高耸的山峰上，有一座仙宫，景色宜人。

通过这样的文本描述，DALL-E 2 可以生成一幅类似的山海经异兽主题的卡牌游戏原画，并且每次生成的结果都是不同的。

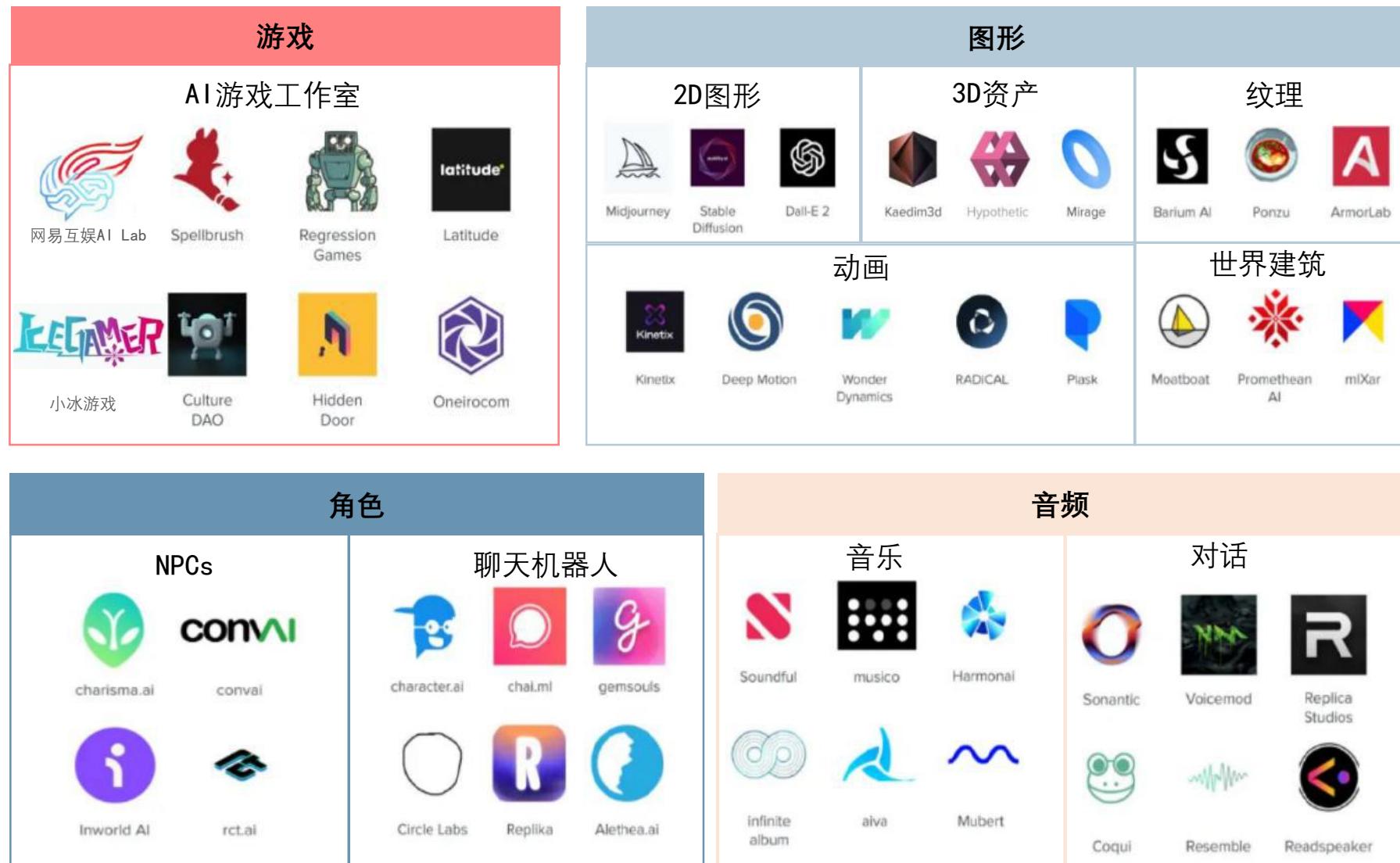
图：山海经异兽主题卡牌游戏原画



6.9 游戏领域生成式AI产业链图谱

- 目前多家国内外公司布局游戏领域的生成式AI，包括网易互娱AI Lab、小冰游戏工作室、Spellbrush等综合型AI游戏工作室以及专注于2D图形、3D资产、纹理、动画、建筑、NPC、聊天机器人、音乐、对话等领域的初创公司。

图：游戏领域生成式AI产业链图谱



风险提示

- 生成式AI技术发展不及预期：算力支持程度不及预期；数据质量及数量支持程度不及预期；用户需求不及预期；技术垄断风险；原始训练数据存在偏见风险；算法偏见与歧视风险；算法鲁棒性风险；算法透明度风险；增加监管难度风险；政策监管风险；商业化能力不及预期；相关法律法规完善不及预期；版权归属风险；深度造假风险；人权道德风险；影响互联网内容生态健康安全风险；企业风险识别与治理能力不足风险；用户审美取向发生变化的风险。

分析师介绍

武超则：中信建投证券研究所所长兼国际业务部负责人，董事总经理，TMT行业首席分析师。新财富白金分析师，2013-2020年连续八届新财富最佳分析师通信行业第一名；2014-2020年连续七届水晶球最佳分析师通信行业第一名。专注于5G、云计算、物联网领域研究。中国证券业协会证券分析师、投资顾问与首席经济学家委员会委员。

杨艾莉：传媒互联网行业首席分析师，中国人民大学传播学硕士，曾任职于百度、新浪，担任商业分析师、战略分析师。2015年起，分别任职于中银国际证券、广发证券，担任传媒与互联网分析师、资深分析师。2019年4月加入中信建投证券研究发展部担任传媒互联网首席分析师。曾荣获2019年wind资讯传播与文化行业金牌分析师第一名；2020年wind资讯传播与文化行业金牌分析师第二名；2020年新浪财经评选传媒行业新锐分析师第二名。

研究助理

杨晓玮， yangxiaowei@csc.com.cn

评级说明

投资评级标准		评级	说明
报告中投资建议涉及的评级标准为报告发布日后的6个月内相对市场表现，也即报告发布日后的6个月内公司股价（或行业指数）相对同期相关证券市场代表性指数的涨跌幅作为基准。A股市场以沪深300指数作为基准；新三板市场以三板成指为基准；香港市场以恒生指数作为基准；美国市场以标普500指数为基准。	股票评级	买入	相对涨幅15%以上
		增持	相对涨幅5%—15%
		中性	相对涨幅-5%—5%之间
		减持	相对跌幅5%—15%
		卖出	相对跌幅15%以上
	行业评级	强于大市	相对涨幅10%以上
		中性	相对涨幅-10%-10%之间
		弱于大市	相对跌幅10%以上

分析师声明

本报告署名分析师在此声明：(i) 以勤勉的职业态度、专业审慎的研究方法，使用合法合规的信息，独立、客观地出具本报告，结论不受任何第三方的授意或影响。 (ii) 本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接受到任何形式的补偿。

法律主体说明

本报告由中信建投证券股份有限公司及/或其附属机构（以下合称“中信建投”）制作，由中信建投证券股份有限公司在中华人民共和国（仅为本报告目的，不包括香港、澳门、台湾）提供。中信建投证券股份有限公司具有中国证监会许可的投资咨询业务资格，本报告署名分析师所持中国证券业协会授予的证券投资咨询执业资格证书编号已披露在报告首页。

本报告由中信建投（国际）证券有限公司在香港提供。本报告作者所持香港证监会牌照的中央编号已披露在报告首页。

一般性声明

本报告由中信建投制作。发送本报告不构成任何合同或承诺的基础，不因接收者收到本报告而视其为中信建投客户。

本报告的信息均来源于中信建投认为可靠的公开资料，但中信建投对这些信息的准确性及完整性不作任何保证。本报告所载观点、评估和预测仅反映本报告出具日该分析师的判断，该等观点、评估和预测可能在不发出通知的情况下有所变更，亦有可能因使用不同假设和标准或者采用不同分析方法而与中信建投其他部门、人员口头或书面表达的意见不同或相反。本报告所引证券或其他金融工具的过往业绩不代表其未来表现。报告中所含任何具有预测性质的内容皆基于相应的假设条件，而任何假设条件都可能随时发生变化并影响实际投资收益。中信建投不承诺、不保证本报告所含具有预测性质的内容必然得以实现。

本报告内容的全部或部分均不构成投资建议。本报告所包含的观点、建议并未考虑报告接收人在财务状况、投资目的、风险偏好等方面的具体情况，报告接收者应当独立评估本报告所含信息，基于自身投资目标、需求、市场机会、风险及其他因素自主做出决策并自行承担投资风险。中信建投建议所有投资者应就任何潜在投资向其税务、会计或法律顾问咨询。不论报告接收者是否根据本报告做出投资决策，中信建投都不对该等投资决策提供任何形式的担保，亦不以任何形式分享投资收益或者分担投资损失。中信建投不对使用本报告所产生的任何直接或间接损失承担责任。

在法律法规及监管规定允许的范围内，中信建投可能持有并交易本报告中所提公司的股份或其他财产权益，也可能在过去12个月、目前或者将来为本报告中所提公司提供或者争取为其提供投资银行、做市交易、财务顾问或其他金融服务。本报告内容真实、准确、完整地反映了署名分析师的观点，分析师的薪酬无论过去、现在或未来都不会直接或间接与其所撰写报告中的具体观点相联系，分析师亦不会因撰写本报告而获取不当利益。

本报告为中信建投所有。未经中信建投事先书面许可，任何机构和/或个人不得以任何形式转发、翻版、复制、发布或引用本报告全部或部分内容，亦不得从未经中信建投书面授权的任何机构、个人或其运营的媒体平台接收、翻版、复制或引用本报告全部或部分内容。版权所有，违者必究。

中信建投证券研究发展部

北京
东城区朝内大街2号凯恒中心B座12层
电话：(8610) 8513-0588
联系人：李祉瑶
邮箱：lizhiyao@csc.com.cn

上海
上海浦东新区浦东南路528号南塔2106室
电话：(8621) 6882-1600
联系人：翁起帆
邮箱：wengqifan@csc.com.cn

深圳
福田区益田路6003号荣超商务中心B座22层
电话：(86755) 8252-1369
联系人：曹莹
邮箱：caoying@csc.com.cn

中信建投（国际）

香港
中环交易广场2期18楼
电话：(852) 3465-5600
联系人：刘泓麟
邮箱：charleneliu@csci.hk

