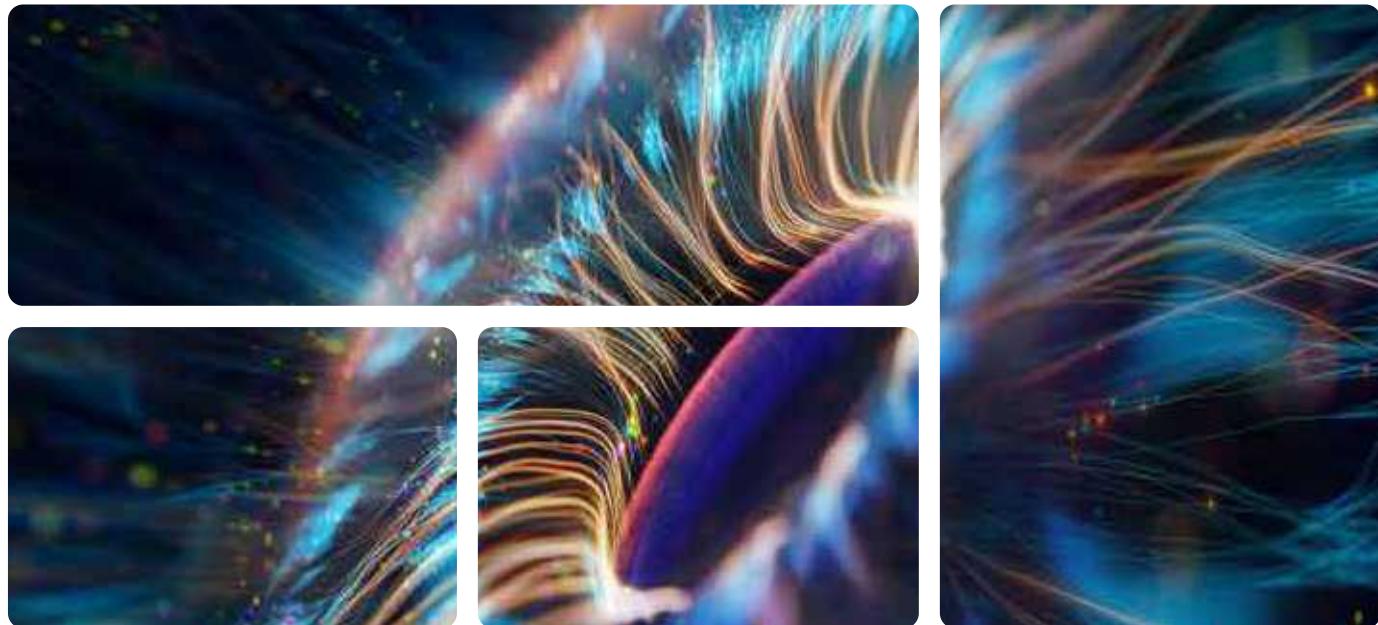


# 2020 腾讯人工智能 白皮书

WHITE PAPER

「泛在智能」



A R T I F I C I A L   I N T E L L I G E N C E

# BIMAGORA - 一站式BIM服务广场

综合BIM服务商城  
合作伙伴火热招募中...

专业平台/精准推广/高效撮合

免费入驻

- ✓ 大量采购、合作和招募信息
- ✓ 各类产品、服务和解决方案
- ✓ 专业问答、专家咨询、用户交流
- ✓ 参考文献、行业方案、知识集锦
- ✓ 领先的开放式项目管理应用系统
- ✓ 强大的GIS+BIM+X数字孪生平台

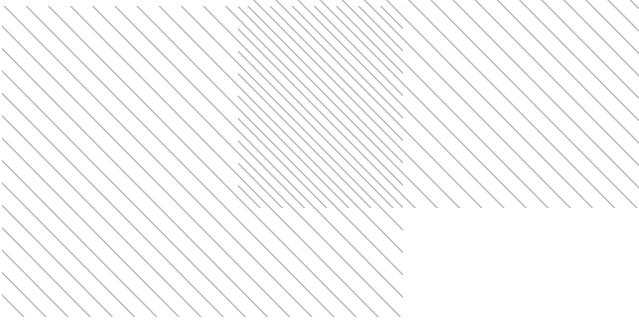
欢迎免费入驻：[BIMAGORA.COM](http://BIMAGORA.COM)

长按识别URL访问主页



咨询及合作请加微信





## 序言

人工智能的发展并不平静。自Alpha Go人机大战重新掀起人工智能热潮以来，经历了炒作与狂热、泡沫褪去后落地的艰难以及隐私伦理的挑战。而在过去的半年，人工智能与产业的结合前所未有的紧密。2020年是足够记入史册的一年，在全球抗疫的背景下，人工智能在医疗、城市治理、工业、非接触服务等领域快速响应，从“云端”落地，在疫情之中出演关键角色，提高了抗疫战争的整体效率。本次新冠肺炎疫情成为数字技术的试金石，人工智能作为新一轮科技革命和产业变革的重要驱动力量，验证了对社会的真正价值。疫情后时代，长周期的经济恢复与发展成为重点，新基建赋予了人工智能全新使命，要求人工智能技术发挥未来产业头雁效应，通过与传统产业的深度融合，助力实体经济向数字化智能化转型，催生新的业态，实现新的蜕变、新的发展。从需求侧看，长期的经济转型压力和近期的抗疫恢复形成双重牵引，各行各业充分认识到加速数字化、网络化、智能化转型已是必然趋势；从供给侧来看，人工智能技术作为国家重要战略发展至今，产业各生态层不断丰富成熟，产业参与者聚焦价值领域，去伪存真成为主旋律。

由此，我们认为人工智能正在进入技术与产业的融合发展阶段，其特征是“泛在智能”：

一是“泛”于基础设施建设。在新基建的春风下，人工智能技术将逐渐转变为像网络、电力一样的基础服务设施，向全行业全领域提供通用的AI能力，为产业转型打下智慧基座。在产业互联网时代，促进产业数字化升级和变革。

二是“泛”于更加多元的应用场景和更大规模的受众。随着技术、算法、场景和人才的不断充实，人工智能正在渗透到各个领域，在工业、医疗、城市等领域验证了AI的价值，未来会有更多产业将与智能技术进行创新融合，催生出更多新业态、新模式。同时，更多应用场景将秉承“科技向善”的信念，以更普惠、更负责任的发展为目标，泛于大众、惠于大众。

对于泛化过程中可能遇到的问题，合理的制度建设将会提供坚实的保障，多层次的治理体系、敏捷灵活的治理方式更能适应人工智能所具有的快速发展迭代、日益复杂化等特征。

经过这次疫情，中国已经不再有纯粹的传统产业，每个产业或多或少都开启了数字化、网络化、智能化进程。人工智能已“泛”起巨浪，而人人都将席卷在这巨浪之中。

腾讯研究院院长 司晓

# 主要发现

---

人工智能正在进入供需融合的创新发展期，从需求侧看，在长期的经济转型压力和近期的抗疫恢复形成双重牵引；从供给侧来看，人工智能产业各生态层不断丰富成熟，已经在工业、医疗、城市等多领域落地。

疫情是AI的试金石，是一次国家级别的数字化全面检验，从基础设施的完备性、数据的流通性到快速反应能力的一次全面测试。在疫情之下，AI公司不再是以往的旁观者，而是出演关键角色，提高抗疫战争的整体效率。经过疫情，中国已经不再有纯粹的“传统产业”，每个产业或多或少都开启了数字化进程。

受疫情用工难、成本加剧、劳动力感染等风险因素的影响，制造业和服务业正在加快人机结合的进程，向制造、服务智能化进一步转型。

此次疫情防控中，人工智能技术在城市治理方面广泛落地应用，表明我国智能社会形态正在逐渐显现。

从“互联网+”到“智能+”，中国人工智能政策越来越重视技术与实体经济的结合，新基建更是进一步让人工智能肩负重任，要求人工智能技术发挥未来产业头雁效应，通过与传统产业的深度融合，助力实体经济向数字化智能化转型。

我们认为，下一个阶段的人工智能将呈现“泛在智能”，即人工智能技术将广泛渗入新型基础设施建设，且获得越来越多元的应用场景和更大规模的受众。

随着技术、算法、场景和人才的不断充实，人工智能正在渗透到各个领域，其商业应用也在不断催生出新业态、新场景、新融合、新交互和新目标的出现。

疫情期间自动驾驶应用崭露头角的背后，是自动驾驶商业化场景开始在国内外加速落地的结果，无人经济的概念获得了新的诠释。

深度合成技术不断演进，朝着综合性的方向发展，泛娱乐和数字内容已然成为了深度合成应用的主场，数字虚拟人受到市场青睐。

日益严重的食物、能源和水问题是当前人类所面临最重要的全球性难题，腾讯相信人工智能会在这些地球级的挑战上发挥重要作用。

多层次的治理体系、敏捷灵活的治理方式更能适应人工智能所具有的快速发展迭代、日益复杂化等特征。

# 编委会

## 策划监制

姚星 司晓 张正友 吴运声 黄世飞 李学朝 罗朝亮 黄飞跃 林涛

## 主编

徐思彦

## 研究团队

**腾讯研究院:** 徐思彦 李瑞龙 吴朋阳 郜若璇 曹建峰 刘琼 Andy Wang

## 顾问团队

**腾讯优图:** 李季檩 汪铖杰 丁守鸿 李绍欣 周可菁 黄亮 高斌斌

**腾讯AI Lab:** 黄俊洲 刘威 苏丹 史树明 李蓝青 吴保元 叶德珩 张锋 赵沛霖

**腾讯云:** 王磊 周超 徐晓敏 王刚 赵明君 董婷

**腾讯云小微:** 李沐 莫小帅 李志平

## 统筹协调

王成 易潇 周威 周军军 戚蕴 梁麗汶 万琳 刘彬 吴静荷

# 目录 CONTENTS

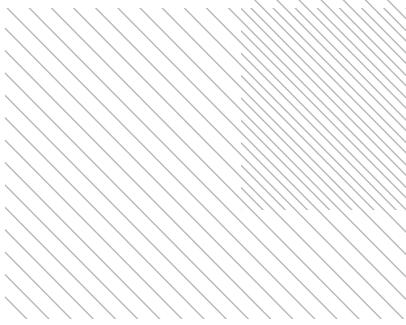
## 一、宏观环境篇

1.1 政策：支持人工智能走向“泛在”	2
1.2 需求：智能应用场景持续涌现	3
1.3 供给：智能产业生态不断丰富	4
1.4 AI 走向泛在智能	6

---

## 二、技术研究篇

2.1 机器学习	8
一、动力：学术研究与社会需求的高关联度	8
二、进展：不断拓展深度和广度，打造虚拟世界试验场	8
三、方向：从虚拟走向现实，助力传统行业实现智能化	10
2.2 计算机视觉	10
一、动力：多年的投资热点正加快寻找技术变现	10
二、进展：从“精度”到“安全”，技术应用日渐成熟	10
三、方向：提升技术可靠性，促进与社会的融合	12
2.3 语音技术	13
一、动力：语音交互成为智能时代应用的新入口	13
二、进展：不断逼近“自然交流”的三大变革实践	13
三、方向：探索产业与技术相互促进的发展模式	14
2.4 自然语言处理	15
一、动力：重点基础技术底座上的高频次需求	15
二、进展：从“用、学、传”三方面推进技术进步	15
三、方向：在变化中探索如何突破质量、场景和表达限制	16



<b>三、智慧抗疫篇</b>	18
3.1 AI+ 医疗：抗击疫情一线，缓解医疗资源压力	19
3.2 AI+ 城市：基础设施升级，融入城市治理体系	21
3.3 AI+ 复工：智能化转型助力复工复产	23
3.4 AI+ 数字内容：多点开花，赋能内容行业生产力	26
3.5 小结	28
<hr/>	
<b>四、创新经济篇</b>	29
4.1 新模式：AI 助力零售向无接触经济转型	30
4.2 新场景：加快自动驾驶的商业化进程	31
4.3 新融合：深度合成为数字内容注入新动能	33
4.4 新交互：互动感加倍，多模态虚拟人融入更多场景	34
4.5 新目标：为地球思考，AI for FEW 挑战资源难题	36
<hr/>	
<b>五、制度保障篇</b>	38
5.1 构建多层次的治理体系，采取敏捷灵活的治理方式	39
5.2 立法和监管需充分考虑国际竞争、技术的经济社会价值等视角，推动先行先试	39
5.3 推动人工智能治理的跨学科参与和国际合作	39



# — 1 —

## 宏观环境篇

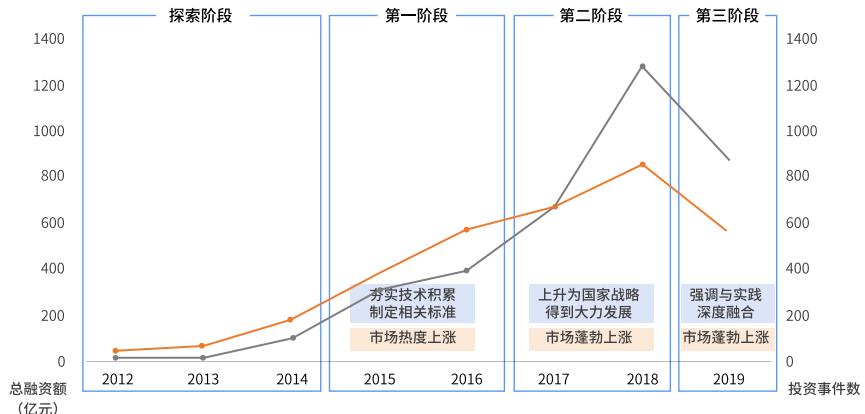
人工智能是新一轮科技革命和产业变革的重要驱动力量，其发挥作用的广度和深度堪比历次工业革命。蒸汽革命成就了英国，电气和信息化革命成就了美国。如今人工智能的历史机遇正在全面展开，也将成就新的引领者。

人工智能正成为未来经济的主要增长点之一。根据埃森哲对中国以及全球 12 个发达经济体的研究，到 2035 年，人工智能将帮助各国显著扭转经济增速近年来的下滑趋势。有效应用人工智能，中国经济增长率有望上升至 7.9%，增长额高达 7.1 万亿美元。此外根据麦肯锡的预测，人工智能将每年为中国经济增长贡献 0.8 至 1.4 个百分点。

尤其在疫情之后的近未来，将是全球经济重建的重要时期，也很可能是社会转型的重要时期。这为人工智能的发展打开了新的窗口期和丰富的实践场，一个“泛在智能”的世界正在加速成为现实。谁能更好地拥抱智能，谁就能率先擘画出未来高质量发展的蓝图。

## 「1.1 政策：支持人工智能走向“泛在”」

近年来，全球主要国家都在纷纷加大对人工智能的关注、支持和投入，这反映出人工智能的战略重要性已成为普遍共识。我国对人工智能的重视程度也在持续增加，自2015年国家产业政策正式提及人工智能以来，5年间相关政策已经历了4个阶段的发展升级。如今人工智能更成为“新基建”政策的一部分，这为我国进一步加快推进“泛在智能”提供了极为有利的条件和机遇。



(图：中国人工智能政策阶段与融资热度制图：腾讯研究院)

### 第一阶段（2015-2016）

#### 夯实技术积累，制定智能相关标准

人工智能的初期政策，主要集中在体系设计、技术研发和标准制定等方面，以尽快为后续发展奠定基本的框架和技术基础。

2015年7月，国务院出台《关于积极推进“互联网+”行动的指导意见》将人工智能纳入发展的重点任务之一，标志着专门人工智能制定产业政策的时期正式开启。2016年5月，发改委印发《“互联网+”人工智能三年行动实施方案》，提出打造人工智能基础资源与创新平台，建立人工智能产业体系、创新服务体系和标准化体系，突破基础核心技术，在重点领域培育若干全球领先的人工智能骨干企业等任务。2016年8月，国务院发布了《“十三五”国家科技创新规划》，明确把人工智能作为体现国家战略意图的重大科技项目。

### 第二阶段（2017-2018）

#### 发展智能产业，拓展智能生活

人工智能正式上升为国家战略，推出了专门的发展规划，并持续对智能产业和智能生活的重点领域进行细化

和强调，指引人工智能发展的着力方向。

2017年3月，“人工智能”首次被写入全国政府工作报告。7月，国务院发布《新一代人工智能发展规划》，人工智能全面上升为国家战略，规划明确提出“三步走”的战略目标：到2025年，新一代人工智能在智能制造、智能医疗、智慧城市、智能农业、国防建设等领域得到广泛应用。10月，人工智能被写入党的十九大报告，“推动互联网、大数据、人工智能和实体经济深度融合”。

2018年1月，《人工智能标准化白皮书(2018版)》正式发布，标准化工作进入全面统筹规划和协调管理阶段。3月，人工智能再度被列入政府工作报告，着重强调“产业级的人工智能应用”，在医疗、养老、教育、文化、体育等多领域推进“互联网+”，发展智能产业，拓展智能生活。10月，习近平总书记就人工智能专题组织中共中央政治局第九次集体学习，提出了主攻关键核心技术，围绕建设现代化经济体系，在质量变革、效率变革、动力变革中发挥人工智能作用，加强人工智能同保障和改善民生的结合，创造更加智能的工作方式和生活方式等任务。11月，工信部办公厅印发《新一代人工智能产业创新重点任务揭榜工作方案》，旨在聚焦“培育智能产品、突破核心基础、深化发展智能制造、构建支

撑体系”等重点方向，征集并遴选一批掌握关键核心技术、具备较强创新能力的单位集中攻关。

### **第三阶段（2019）：互联网+升级到智能+，强调与实体经济深度融合**

人工智能战略被升级和扩展为“智能+”，结合相关的工作指导和指引，推动人工智能与实体经济深度融合落地。

2019年，人工智能第3年出现在政府工作报告中。报告提出，要打造工业互联网平台，拓展“智能+”，为制造业转型升级赋能。促进新兴产业加快发展，深化大数据、人工智能等研发应用。3月，中央全面深化改革委员会第七次会议中，审议通过了《关于促进人工智能和实体经济深度融合的指导意见》，着重强调市场导向与产业应用，打造智能经济形态。8月，科技部《国家新一代人工智能创新发展试验区建设工作指引》提出，到2023年，布局建设20个左右试验区，创新一批切实有效的政策工具，形成一批人工智能与经济社会发展深度融合的典型模式，积累一批可复制可推广的经验做法，打造一批具有重大引领带动作用的人工智能创新高地。北京、上海、天津、深圳、杭州、合肥、德清县、济南、西安、成都、重庆等地相继获批建设国家新一代人工智能创新发展试验区。

### **第四阶段（2020-至今）**

#### **纳入“新基建”范畴，全面加速智能普及**

人工智能被纳入“新基建”政策，成为新技术基础设施的主要支撑技术之一，背负着带动疫后经济复苏与增长的使命，也将会是新一轮产业变革的核心驱动力，肩负着推动万亿实体经济产业转型升级的目标。

早在2018年底的中央经济工作会议，习近平总书记就在会上首次提出新基建概念，强调要“加快5G商用步伐，加强人工智能、工业互联网、物联网等新型基础设施建设”。2020年初的新冠肺炎疫情，加速了国家进一步开展人工智能等新型基础设施建设的进程。同年5月，两会政府工作报告中提出，要加强新型基础设施建设，发展新一代信息网络，拓展5G应用，建设数据中心，激发新消费需求、助力产业升级。在发改委的定义中，人工智能在新基建三个层面都肩负了重要角色，既是基于新一代信息技术演化生成的信息基础设施，也在智能交通等融合基础设施中发挥重要作用，并为科技研发等提供创新基础设施支撑。长期来看，人工智能作为新技术基础设施，被视为支撑传统继承设施转型升级的融合创新工具。新基建将加速中国产业链完成数字化转型和智能化升级，实现产业要素的高效配置，助力国家经济发展新旧动能转换。

## **1.2 需求：智能应用场景持续涌现**

人工智能政策的加速推进落地，根本目标在于形成人工智能的繁荣市场，也依赖于市场需求和供给各方面所具备的资源禀赋。需求角度而言，我国庞大的人口基数和经济规模为人工智能等新技术的应用发展提供了广阔的市场基础，在长期的经济转型和近期的抗疫恢复双重牵引下，智能化相关的应用场景不断增加，有效拉动了我国人工智能的全面商业化进程。

#### **转型需要：人口红利转化为创新红利**

中国互联网大规模普及释放的人口红利已到天花板，未来国家经济发展需要从规模高速增长转向质量高效发展阶段，迫切需要人工智能等新一代技术创新添薪续力。

互联网为代表的数字经济，是过去20多年中国经济高速发展最鲜明的组成。根据CNNIC的报告，截至2020年3月，我国网民规模为9.04亿，互联网普及率达64.5%，用户规模已连续13年占据世界首位，衣食住用行各方面不断在线化。根据腾讯研究院《数字中国指数报告2019》测算，2018年中国数字经济占GDP的比重已达33.22%，为我国经济发展做出了巨大贡献。然而近年来随着互联网普及，网民整体增速在巨大的基数下已经较为缓慢，原先靠人口红利实现快速增长的模式，面临向结构和质量转型升级的迫切需求。

人工智能、大数据和云计算等新技术推动的“计算变革”，有望在互联网“连接变革”之后带动新的创新

增长。一方面在海量互联网连接的基础上，这些智能技术应用能够进一步降低成本、提升连接和各类承载线上的服务质量效率、创造新的应用和服务场景等，从而更有效地满足市场的个性化需求，扩大经济高质量增长的空间；另一方面这些新技术本身的规模发展，也将形成新的高科技产业生态，吸引资本、人才等资源向新领域聚集，从而推动经济结构向高质量发展领域转型，最终完成智能经济、智能社会的升级。

### 场景丰富：超大规模且多样的应用场景

我国经济体量大、多样性强，在网络化和数字化普及的基础上，各种经济活动形成了海量数据资源，为人工智能发展提供了丰富的应用探索和落地场景。

在规模方面，如今我国网民人数已达美国总人口的近3倍，超过欧洲的总人口数。庞大数量的网民是人工智能应用的需求来源和最终用户，提供了丰富的消费场景。不少大规模的场景应用需求，目前只发生在中国。如2019年春运期间，有4.1亿人乘坐火车，24.6亿人乘坐汽车，完成了人类历史上规模最大的周期性迁徙；同样在当年春节仅除夕到周五，就有8.23亿人次收发微信红包……这些出行、支付等海量用户应用场景，为人工智能模型训练和部署提供了丰富的大数据环境，有利于技术不断迭代升级，同时也带来了更低的创新成本和更高的创新收益。

在多样性方面，中国丰富的产业环境也为人工智能提供了各种应用场景。以工业为例，中国是全世界唯一拥

有联合国产业分类当中全部工业门类的国家。在世界500多种主要工业产品当中，有220多种工业产品中国的产量占居全球第一。在不同地区、行业和产品之间，数字化、信息化和智能化的发展水平各异，这也为人工智能在不同阶段发挥作用提供了广阔的实践空间。

### 抗疫加速：疫情等黑天鹅事件催生需求

疫情等黑天鹅事件增加全球不确定性，“物理隔离”对全球经济造成明显冲击，“无接触”的数字在线方式成为经济恢复的重要手段，也为人工智能发展创造了更多“潜在刚需”的新场景。

2020年突如其来的新冠肺炎疫情，给全球经济发展造成了很大冲击，各类生产经营活动一度按下暂停键。面对我国经济复苏的需求，“六稳六保”的重点任务，人工智能技术加速与实体经济深入融合，在新模式和新应用中发挥中坚作用。

疫情也产生了倒逼效应，教育、医疗、零售、办公等进一步向线上转移，在线教育、远程医疗、在线办公等加速崛起，线下的物流配送、无人车快递等无接触经济发展迅猛；在服务疫情防控中，AI在医疗影像识别，基因测序和疫苗研发方面效果凸显；针对安全复工复产的需求，多人快速测温、口罩识别等一批人工智能创新解决方案应运而生；面对大量工厂工人无法按时复工的痛点，依靠视觉识别的智能检测设备大显身手，未来的新型人机协作模式有望进一步加速成型。

## 「 1.3 供给：智能产业生态不断丰富 」

中国人工智能需求潜力很大，也更需要充分且有效的供给来满足。得益于这些年互联网大发展形成的数字化基础和生态环境，我国人工智能相关产业快速发展。根据中国新一代人工智能发展战略研究院发布的《中国新一代人工智能科技产业发展报告·2020》显示，我国人工智能企业数量仅次于美国排名全球第二。智能产业生态不断丰富成熟，特别体现在技术、数据、资本、市场和平台等五方面。

### 技术：从实验室走向大规模商业应用

从底层的AI芯片、深度学习框架，到图像识别、语音识别、自然语言理解和知识图谱等通用技术，人工智能技术的能力快速提升，在各类经济和社会应用中展现出巨大的经济价值。同时，人工智能与云计算、区块链、物联网、5G等新技术的相互融合，以及新技术带来海量的物联和移动等数据，正在进一步激发人工智能技术应用的广度和深度，让人工智能真正成为经济社会发展的

重要新型基础设施。如人工智能与物联网融合的AIoT，腾讯云的IoT EIDP就为万物智联提供了有效解决方案。IoT EIDP具备设备自动接入、脱机AI计算、丰富算法市场、快速二次开发、快速一键部署和主流硬件平台等6大特性，最大限度满足不同硬件、不同算法的低成本高速部署与运营，能广泛应用于智能家居、智能楼宇、智慧工厂、智慧零售等各行业场景。

### **数据：数据正式成为国家生产要素**

数据与智能密不可分。一方面有数据才能实现智能。人工智能基于数据训练，海量和优质的大数据将推动AI算法的持续优化，强化智能平台的服务能力，进一步提升各行业的智能化水平，让数据价值得以真正发挥。另一方面，人工智能也让数据更丰富。各种智能化应用的深入，将激发更多的用户使用，从而催生更多的大数据产生。

今天的中国，不仅已形成腾讯、阿里、百度等大量互联网平台应用的海量数据，各行各业、各级政府等也在持续增强数字化转型升级的力度，更广泛多元的大数据资源正在不断产生。特别2020年4月9日《中共中央国务院关于构建更加完善的要素市场化配置体制机制的意见》发布，明确数据要素市场化配置上升为国家政策，为数据的广泛流动和市场价值转化提供了依据，将有效支持人工智能在全社会的实践。

### **资本：走出炒作泡沫聚焦价值领域**

2015年以来，人工智能领域的投融资开始出现爆发式增长，相关投资金额和投资数量高速增长。根据清科集团的数据，仅2015年，投资案例就达到719起，增长了156%，投资金额升至160亿，同比增长190%。中国吸引了大量投资，根据英国Tech Nation的《UK Tech For a Changing World》报告显示，从2015年到2019年，中国人工智能公司共筹集了120亿美元，占全球22%位居第二仅次于美国。

近年来，随着人工智能走出概念炒作的风口，以及部分项目回报率不高、退出率较低等问题，加上全球经济形势的不确定性，投资界对人工智能的认知开始回归理性，从单纯的追逐热点，转向更加注重核心技术能力和应用潜力。根据中国信息通信研究院的《全球人工智能产业数据报告》，2019年一季度，中国人工智能融资

金额为30亿美元，同比下降55.8%。同时，根据德勤的《全球人工智能发展白皮书》研究，投资更加青睐底层技术创业公司以及落地性强的领域如医疗、教育、无人驾驶等创业项目。

### **市场：供需互促的正向循环基本建立**

随着部分人工智能技术应用的发展成熟，供需开始形成互相促进的正向循环形势。用户一旦开始智能化转型，就启动了“数据积累—模型优化—应用升级”不断迭代优化的循环，收益也随之不断提升，从而更多使用AI。根据麦肯锡的研究，率先应用人工智能的企业，到2030年，将会实现现金流翻倍，领跑者往往拥有强大的IT基础，更高的AI投资倾向，以及对AI业务案例的积极看法。

使用的增加带动更多AI供给，激励人工智能企业加快商业化落地步伐，以抢占市场先机。人工智能产品和服务的开展，往往涉及长时间的调研选型、规划设计、模型训练开发和应用部署，加上大量行业知识的融入，以及数据安全、系统运维等大量的保障工作。这使得人工智能转换成本较高，率先获得客户的人工智能供应商能够积累较强的先发优势。

### **平台：大厂普遍开放平台促进生态**

以大规模算力使用和大数据处理为基础的人工智能，越来越展现出规模效应的特征，即使用越多、价值越高、成本越低。这促使企业、尤其是互联网大厂积极采取开放平台的方式，将各种AI技术和资源对外开放，促进了整个AI生态的发展繁荣。典型如语音、图像等智能化处理，今天通过API方式已经能非常方便地调用，有效支持像解锁、翻译、美颜等各种创新智能应用不断涌现。

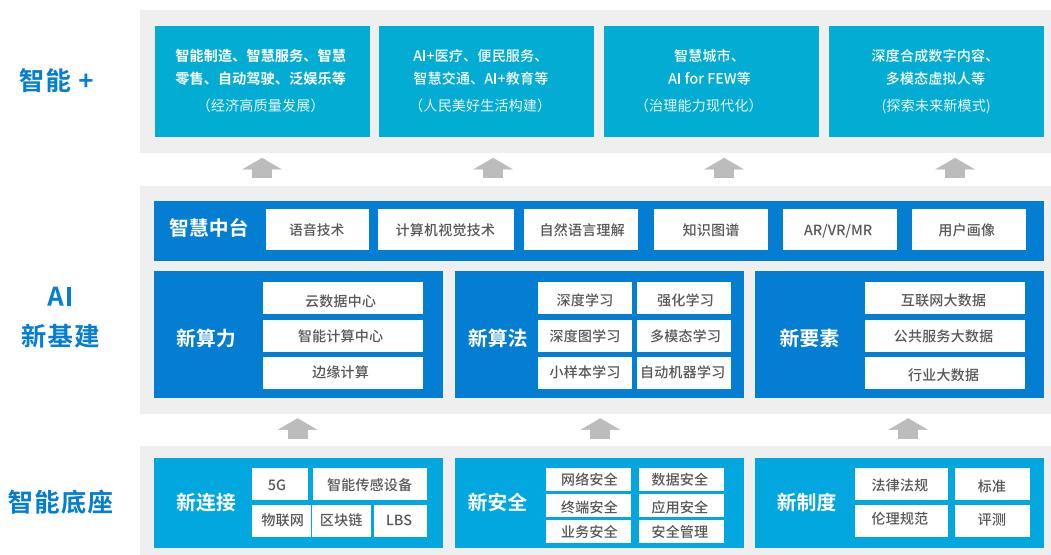
今天AI开放平台的模式，已成为行业共识。这里面既有全技术栈的腾讯AI开放平台、阿里云开放平台，百度大脑（AI开放平台）、华为HiAI能力开放平台，也有针对单一技术的开放，如科大讯飞智能语音开放平台、商汤智能视觉开放平台、依图视觉计算开放平台、海康威视视频感知开放平台、旷视图像感知开放平台等等。未来更重要的是如何让众多开放平台形成共识与合力，合作联盟、共制标准、行业自律、法律保障等制度建设的重要性将更突出。

## 「1.4 AI 走向泛在智能」

近年来，全球市值最大的科技公司苹果、谷歌、微软、亚马逊、脸书无一例外投入越来越多资源抢占人工智能市场，甚至整体转型为人工智能驱动的公司，积极扩大AI布局的巨头公司们都在争取通往AI世界的第一张“船票”。相对而言，腾讯在人工智能的战略方向上提出了一个较为平实的愿望：“AI in All”。与互联网、移动互联网一样，人工智能是基础能力。腾讯正在从技术、场景和平台三个层面打造这项基础能力。在技术层面，主动强化在核心人才方面布局，组建人工智能实验室，夯实基础研究。在场景层，腾讯广泛多元的业务内容，为机器学习、计算器视觉、语音识别及自然语言处理等领域的AI基础研究提供了丰富应用场景，使游戏、内容、社交、金融、医疗等方面的应用能够快速应用人工智能能力，带来更智能的用户体验。在平台层，推出“AI即服务”战略，满足市场对AI能力多维度的需求；并且在语音识别、图像识别、语义识别、以及深度学习框架多个维度对外开放AI能力，开源Angel、ncnn、

TNN等机器学习、深度学习平台，降低AI开发门槛，推动更多现实场景落地。

我们认为，下一个阶段的人工智能将进入“泛在智能”的阶段。泛在，即广泛存在。“泛在智能”，一是泛于基础设施建设，在“新基建”的春风下，智能技术将逐渐转变为像网络、电力一样的基础服务设施，作为新一轮产业变革的核心驱动力，人工智能算力、算法、数据三位一体的基础设施可面向全行业全领域提供通用的AI能力，促进社会经济转型升级；二是泛于越来越多元的应用场景和更大规模的受众，更多的传统产业或快或慢接入智能技术，与技术公司共同探索未来新模式，同时人们会在生活里感受到无处不在的智能应用，技术的普惠精神进一步得到体现；泛化过程中，注定会遇到同样广泛的问题，对此，与其他数字技术之间更加紧密的连接给人工智能提供了更有力的技术支持，而安全和制度保障则是不可或缺的守夜人。三个层次相互作用相互促进，共同加速人工智能在更广维度扩展。



(制图：腾讯研究院)

经此一“疫”，数字技术验证了其对社会真正的价值和未来社会发展的方向，至此，中国已经不再有纯粹的传统产业，每个产业或多或少，都主动或被动开启了数字化进程，人工智能已“泛”起巨浪，而人人都将席卷在这浪潮之中。



## —2—

# 技术研究篇

“泛智”的前提是“有智”且“智”得高效、“智”得可靠，这一切的“智”的来源离不开技术的成熟。本轮以机器学习和深度学习人工智能为主题的浪潮被认为是当前人类所面对的最为重要的技术社会变革之一，互联网存储了二十多年的数据终于找到了它的历史使命：训练机器，是互联网诞生以来的第二次技术社会形态的全球萌芽。在过去十年，用于人工智能的训练模型的计算资源激增。在2010年至2020年之间，人工智能的计算复杂度每年激增10倍，人工智能训练成本每年下降约10倍。在计算力上，得益于芯片处理能力提升，硬件价格下降使算力大幅提升。

在此基础上，各项人工智能技术不断得到突破，并找到相对明确的应用场景。据清华大学数据显示，计算机视觉、语音技术以及自然语言处理的市场规模占比分别为34.9%，24.8%和21%，是中国市场规模最大的三个应用方向。经过几年的发展，腾讯的各大AI实验室也在机器学习、计算机视觉、语音技术以及自然语言处理这四项人工智能重点技术领域不断耕耘取得突破，并持续探索未来进一步的发展方向。

## 「 2.1 机器学习 」

### 一、动力：学术研究与社会需求的高关联度

机器学习（尤其是深度学习）是催生了近年来人工智能发展热潮的最重要的技术。但是机器学习是一门历史悠久的学科，最早可以追溯1943年，Warren McCulloch和Wallter Pitts提出了神经网络层次结构模型。机器学习旨在通过计算机学习数据中的内在规律性信息，获得新的经验和知识，以提高计算机的智能性，使计算机能够像人那样去决策。这类技术最核心的意义是能让机器具备自动发现模式和寻找解决方案的能力。随着数据科学的研究的不断发展，以及各行业对数据的需求量增多，对处理和分析数据的效率要求变高，一系列机器学习算法应运而生。

**机器学习是门“全能”的基础学科。**机器学习本身就是一门综合学科，涉及统计学、神经网络、优化理论、脑科学等众多领域，由于这一显著的“交叉学科”特性，机器学习对于其它领域的科研工作，具有较好的辐射性。一方面，一些传统学科，如统计、理论计算机科学、运筹优化在发展多年之后，遭遇了一定的发展瓶颈，机器学习为其打开了新的大门；另一方面，具体的应用学科，在众多“可解释”的技术方法下难以寻找传统的量化关系，机器学习提供了解决问题的新方法与途径。例如生物信息学，相关研究涉及从“生命现象”到“规律发现”的整个过程，为机器学习的研究提供了重要的舞台。

**机器学习能较好的应用于生活的各个方面。**相对于计算机视觉、自然语言等其它应用型人工智能算法，机器学习更为底层，在各类人工智能算法的训练和推行中都有广泛应用。机器学习是统计分析时代向大数据时代发展必不可少的核心打磨，是开采大数据这一新“石油”资源的工具，比如，在基础应用领域，环境监测、能源勘探、天气预报等基础应用领域，通过机器学习，加强传统的数据分析效率，提高预报与检查的准确性；再如销售分析、画像分析、库存管理、成本管控以及推荐系统等商业应用领域，机器学习让即时响应、迭代更新的个

性化推荐变得更为轻松，渗透至我们生活的方方面面。

### 二、进展：不断拓展深度和广度，打造虚拟世界试验场

腾讯 AI Lab 在机器学习方面的研究主题，包括：强化学习、自动机器学习、深度图学习、小样本学习等。近年来，腾讯AI Lab探索了机器学习在多个场景的应用，例如游戏、医疗、农业等。游戏作为对真实世界的模拟，而成为了人工智能最好的试验场；医疗一方面是机器学习应用的现实场景，另外一方面是腾讯对于科技向善最好的诠释。

**强化学习。**是近年来大放异彩的机器学习技术之一，基于该技术开发的人工智能模型已经在围棋、扑克、视频游戏和机器人等领域取得了非常多的里程碑式进步。在强化学习方面的研究大都立足于游戏环境，如围棋、电子游戏《王者荣耀》和《星际争霸》等。

腾讯AI Lab的强化学习智能体绝艺和绝悟已经取得很多重大的突破，并且通过与顶级人类玩家比拼的方式取得了多个亮眼的里程碑。2019 年我们在该领域的前沿探索还有：

(1) 为多智能体强化学习环境提出了一种新的元学习方法 LIIR，通过优化各个智能体的内在奖励值来实现对总体目标的整体优化，以激励不同智能体采取多样化的有利于团队的行为。我们用这种方法让《星际争霸2》智能体学会了更多样化的决策策略。

(2) 提出了一种散度增强的策略优化算法，能在重复使用离线数据时实现对策略优化训练的稳定化。这种方法有利于在数据不足的环境中进行学习。

(3) 提出了一种针对训练数据选择方法的强化学习框架，可以很好地处理数据的领域迁移任务。

(4) 提出了一种基于强化学习的端到端主动目标跟踪方法，通过自定义奖赏函数和环境增强技术在虚拟环境中训练得到鲁棒的主动跟踪器，并在真实场景中对模型的泛化能力进行了进一步的验证。

若AI能在复杂的游戏环境中，学会人一样实时感知、分析、理解、推理、决策到行动，就可能在多变、复杂的真实环境中发挥更大作用。因此业界认为下一个AI里程碑，可能会在复杂策略游戏中诞生。世界顶级科技公司均在推进此类研究，如Google Deepmind（星际争霸2）、Facebook（星际争霸2）及OpenAI（Dota 2）等。

除游戏外，腾讯还有积极探索相关技术在实际应用场景中的落地，比如在农业方面的应用，腾讯AI Lab连续两年与荷兰顶尖农业研究机构合作举办了AI种植黄瓜、番茄的国际智慧温室种植挑战赛，并在净利润、成本控制、AI策略上均取得超越人类种植专家的喜人成绩。腾讯及其他参赛团队在比赛中基于真实数据、温室仿真器及专家经验，在自动控制层面利用DeePc等技术开发了精准稳定的控制算法，在温室的环境控制及决策层面，利用在线及离线强化学习方法将总成本优化10-16%，净利润提升100%以上。

**自动机器学习。**如今的机器学习模型有往越来越大、越来越深发展的趋势，也因此通过人工方法来设计模型和配置超参数的思路越来越捉襟见肘。因此，通过自动化的方式来设计/搜索模型架构和超参数的方法正受到越来越多的研究关注和应用。

在架构搜索和优化方面，腾讯AI Lab提出了一种神经网络架构变形器 NAT，可以将冗余操作替换为计算效率更高的操作，从而获得精确且紧凑型架构。而在超参数优化方面，我们提出了一种基于分布的贝叶斯优化算法 DistBO，它可以迁移历史任务超参数优化的知识，进而对新任务上的超参数优化起到热启动的效果。

在通过迁移学习来简化模型学习方面。腾讯AI Lab提出了一种基于渐进式特征对齐的无监督域自适应方法，其中包含一种由易到难的迁移策略（EHTS）和一个自适应中心向量对齐步骤（APA），可以迭代并交替地训练域适应网络；另外我们还提出了一种无标签领域自适应算法，可用以辅助癌症诊断任务的病理图像分类学习。另外值得一提的是我们还在一项研究中探索了组合使用迁移学习、多任务学习和半监督学习的方法，并研究了如何通过组合方法来提升医学分类模型的准确率。

在现实生活的实际应用中，很多时候还需要对模型进行压缩处理，以便在尽可能保证模型优良表现的同时提升模型的执行效率。我们提出了一种名为「协同通道剪枝（Collaborative Channel Pruning）」的模型压缩算法，可以保证在准确度无损的前提下有效降低模型的计算成本，从而可让深度学习模型在移动设备等更多场景中得到应用。我们还提出了一种用于压缩卷积神经网络（CNN）的方法，这种方法基于我们定义的一种可分解卷积滤波器，可以实现非常优良的模型压缩效果。

自动机器学习在腾讯的多个重要场景取得落地结果，包括：在医疗AI的智能显微镜项目的多个场景取得3% - 4%的精度提升，例如，在结直肠癌细胞分割上，自研网络结构搜索算法把F1值从92%提高到96%；在游戏AI的ViZDoom AI项目上帮助提升模型训练效率400%；在火锅视频的手机端模型压缩项目上提高调参效率100%，减少精度损失20%；在室内定位项目中，提高模型精度4%，提高模型训练效率300%等。

**其他机器学习方法。**除此之外，腾讯AI Lab也在探索其它形式的机器学习方法，比如结合图学习与深度学习的深度图学习、小样本学习、和多模态学习等方向。

**深度图学习。**是腾讯AI Lab一个非常重要的研究方向，可以帮助我们理解大型信息和知识网络（比如社交网络，这是腾讯的核心业务之一）中的关系信息。在深度图学习方面，我们专注于解决「深度」和「广度」这两个在深度图学习中具有挑战性的问题。「深度」方面，我们提出的 DropEdge 方法可让我们更好地学习超深层图神经网络，得到表现显著更佳的结果。在「广度」方面，我们基于自主研发的图采样算法AS-GCN，即自适应结构采样图卷积神经网络，开发了可以分布式学习超大规模图数据的图学习系统——在亿级别的超大规模图数据上，我们可以在不到5分钟内完成单次训练迭代。同时，我们也将深度图学习算法成功应用到不同领域，如复杂社交网络，谣言检测，对抗攻击等。

**小样本学习。**在实际应用中也具有极其重要的价值，毕竟不是每一种应用场景都存在可以轻松收集或标注的数据集。针对这一任务，我们提出了一种基于层次任务结构的元学习算法（HSML），该方法可以迅速找到与新任务最相关的聚类，然后从该聚类的任务中迁移和泛化知识。

多模态学习。腾讯的业务中包含许多需要处理多模态数据的场景，比如通过分析视频内容与用户弹幕来理解用户观看节目时的情绪。针对多模态数据，我们提出了一种高效的特征提取方法，该方法可以学习到更有信息量的特征映射，同时优化过程也更为高效。另外我们还提出了一种使用对抗样本的跨模态学习方法（CMLA），该方法创造的对抗样本能快速地骗过一个目标跨模态哈希网络，另一方面也能通过对抗训练提升该目标跨模态哈希网络的鲁棒性。

### 三、方向：从虚拟走向现实， 助力传统行业实现智能化

虽然机器学习的研究和应用，已经取得了很好的进

展。但是，还有若干问题制约其影响力的发挥，特别是在真实世界的应用方面。后续重点需要攻克的方向包括：

**提升小样本学习效率。**目前，如果想要得到性能良好的机器学习模型，则需要大量的带标签的训练数据。但是在很多现实场景中，一些数据的标签是很难获得的，例如分子的生物活性。怎么用少量的数据，仍然可以训练出良好的模型，是一个亟待研究的问题。

**发展离线强化学习。**强化学习在游戏中取得了惊人的效果，得益于游戏模拟器可以实时地产生大量的与模型交互的数据。但是在一些现实场景中，我们只能得到少量的现实人工交互数据，或者从模拟器所得的大量的虚拟交互数据。怎么利用这些数据做好离线强化学习，是一个尚无完美解决的问题。

## 2.2 计算机视觉

### 一、动力：多年投资热点正加快寻找技术变现

作为智能世界的双眸，计算机视觉是人工智能技术里的一大分支，是新基建中的重要组成部分。2018年，计算机视觉技术占中国人工智能市场规模的34.9%，位居第一，在投融资规模中更是一枝独秀。随着近几年技术的不断成熟，中国计算机视觉市场得到快速增长，艾媒数据显示，2018年中国计算机视觉市场规模为155亿元，较2017年增多了87亿元，复合增速超100%。

**视频爆炸时代，海量数据亟待处理。**人类的大脑皮层大约有 70% 的部分都是在处理我们所看到的内容，即视觉信息。在计算机视觉之前，图像对于机器是处于黑盒状态，就如同人没有视觉这一获取信息的主要渠道。计算机视觉的出现让计算机能够看懂图像，并能进一步分析图像。从4G到5G，正进一步引发互联网里的视频流量爆炸，视频以各种形式几乎参与了所有应用，从而产生的海量视频数据以指数级的速度在增长。在对这一新型数据类型进行更精准的处理，那么推动计算机视觉的发展是必经之路。

**重点落地场景正逐步突破工业化红线。**2016 年以来，计算机视觉相关技术的工作业应用加快，率先敲响了计算机视觉行业市场大规模爆发的前奏。伴随人脸识别、物体识别、图像识别等计算机视觉细分领域的不断突破，在零售、泛娱乐等多领域逐步加快解锁，成为行业整体快速发展的支撑。在越来越多的行业运用中，特别是医疗影像、工业制造、金融认证等对技术要求高、影响大且应用范围广的重点落地场景，对于计算机视觉的精度提出更高的要求，并希望减少安全方面的隐患。

### 二、进展：从“精度”到“安全”， 技术应用日渐成熟

市场对于技术的推进作用是巨大的，目前，计算机视觉已经在人脸识别、工业视觉、OCR以及内容理解等领域获得重大突破，面对不同场景需求，计算视觉准确度不断提升：

**人脸识别技术是主攻方向。**目前基于深度学习的人脸识别系统精度不断提升，已被广泛应用于零售及金融民生等各类场景，但如何实现统一模型同时满足不同场景对人脸识别精度的要求仍是一个难点。腾讯优图实验室打造了命名为“祖母模型”的人脸识别模型，该模型采用业界首创分支轮替深度网络训练方案，理论上可支持场景分支无上界，实现历史场景Never-Forget的全面兼容；创新性的提出分布蒸馏损失函数，自动发现新场景和已学习场景数据相似度的差异，对深度网络进行端到端训练，从而充分利用多场景数据，提升了模型泛化能力；首次提出人脸课程学习训练算法，根据图片的不同质量和识别的难易程度自动调整训练顺序，提高了模型对难样本的识别能力，避免了对噪声样本的错误过拟合。通过上述技术突破，在1:1人脸比对场景中，正确通过率从早期的80%提升到99%；在千万规模1:N检索中Top1命中率从不到20%提升至95%以上。相关技术在实际场景中，得到了广泛的应用，比如，微信刷脸支付可以实现在极低错误接受率（万亿分之一）的人脸支付场景，正确通过率从80%提高到95%以上，已稳定支持微信每天百万级的刷脸支付请求。利用“跨年龄人脸识别”技术，腾讯守护者计划团队协助警方破获数起沉寂十年的特大拐卖案，累计协助四川、广东公安机关成功比对找到15名被拐超过十年以上的儿童。

**工业视觉正成为重要生产力。**中国作为全球最大的制造业产业国，制造业既是国家实体经济的重要组成部分，也是国家的核心支柱产业。制造业中的质量检查、质量控制，是提升我国工业制品国际竞争力的重要环节。以液晶面板生产为例，其生产过程是一个光、电、化学综合的过程，至少需要300多道工艺流程，工序非常繁复。针对这一专业领域，腾讯优图实验室以视觉AI技术为基础，实现了算法平台化，加速产品工具开发，可以让普通工程师从0到1创建出一个全新的模型，完成技术内化，同时借助腾讯云开发了自动化的训模平台，集数据管理、模型训练、模型部署于一体，让普通管理人员也能轻松实现从1到N的拓展建模，大大降低模型开发人员的门槛。以华星光电为例，之前一个质检员从入职到上岗，需要培训2到3个月。如今，通过使用腾讯优图提供的缺陷检测解决方案，判片速度提升了约10

倍，人力成本缩减约50%。

**OCR实现物品的数据化。**OCR技术，是从图像中识别文字的方法，在现实中有很广泛的应用场景，如车牌识别、身份证识别、护照识别等。腾讯优图基于在OCR领域的深厚技术积累和丰富的实战场景经验，自主研发了高精度的通用OCR引擎，包括多尺度的任意形状文本检测和融合语义理解的文字识别两大核心算法，结合自研数据仿真算法生成的数千万训练集，有效解决了文本畸变、密集排布、复杂背景干扰、手写、小字模糊字等OCR方向的经典难题。为了充分验证算法的性能，腾讯优图OCR在包括文档、路标、书本、试卷、快递单等涵盖数十种场景的数千张图片上全面测试，准确率达到95%。基于自研的高精度通用OCR技术，腾讯优图进一步研发了证照类、教育试题类、票据类等50多种垂直场景的OCR能力，关键字段准确率达到98%，并通过腾讯云文字识别OCR在金融、保险、财务、物流、教育等领域得到广泛应用，信息录入速度提升90%以上，在业务处理效率提升的同时也极大节省了人工录入成本。

**内容理解助力网络环境净化。**内容理解可以应用于内容识别分类、恶意图片审核、精准广告推送和相似内容推荐等场景。互联网时代，线上信息鱼龙混杂，为实现低质内容有效剔除，很多企业安排大量的人力进行人工内容审核，却依然存在效率低、成本高的问题。腾讯优图实验室联合腾讯云，依托先进的人工智能技术，打造了涵盖音视图文的全方面多维度内容理解解决方案，包括鉴黄识别、广告识别、违法违规行为识别等。腾讯优图研究了基于自适应注意力机制和动态图卷积融合的多标签算法，丰富了图像识别的标签粒度，并提高了细小元素的识别效果；提出了基于旋转卷积和动态修正网络的多角度目标检测算法，能够精确定位各种角度的目标，在多个公开数据集上取得了不错的效果。腾讯优图打造的内容理解解决方案通过腾讯云天御出口，目前已广泛应用于UGC、直播、点播、教育和电商等多个场景，日调用量超过1.5亿，为客户降低约95%的人工审核成本。

**安全可信的AI技术已经成为业内关注的焦点。**随着对抗攻击技术的发展，深度模型的安全受到了极大的挑战，无论是人脸识别模型还是目标检测模型，对抗样本带来的扰动可能会导致这些模型进行错误的预测。作为业内较早从事该领域研究的研究机构之一，腾讯在该领域取得了一系列原创和世界领先的科研成果，涵盖了黑盒对抗攻击，精准结构化模型对抗攻击，稀疏对抗攻击，对抗检测，对抗训练等诸多重要子课题。

(1) 腾讯AI Lab于2019年首次以完全黑盒的形式成功攻破了业内主流的人脸识别API，该成果发表于人工智能顶级国际会议CVPR-2019；2020年推出最新版本的黑盒攻击算法则进一步大幅度提高了黑盒攻击的效率，继续在业内保持着领先的技术优势。另外，在1600万的大规模人脸数据集上完成了对抗训练，这是业内首次将对抗训练拓展到如此大规模数据集。利用原创的高效对抗训练算法，不仅克服了对抗训练容易导致正常人脸识别精度明显下降的问题，并明显提高了对抗训练的速度。

(2) 腾讯优图实验室，研发了FaceDefense防御系统，基于随机化的图像patch进行网格化处理，以及深度神经网络特征和传统统计特征的融合，可有效对抗Adv-Hat和Adv-Glasses等可打印佩戴图案的攻击。此外，还不断从对抗训练，人脸质量估计和多模态融合三个方面全面提升人脸识别系统的安全性。模型训练过程中自动构建对抗样本加入训练，在基本保持正常样本性能的同时，对抗样本攻击成功率降低90%；人脸质量估计方面创新性的提出质量总分与质量归因交叉验证的方案，可以在综合判断人脸可识别性的基础上给出归因提示，通过交互和优选相结合的方式提升注册人脸质量，降低误注册和低质量注册造成的系统性风险；多模态融合方面，在RGB模型基础上引入深度图信息，采用曲面法向量投影法计算稳定的深度信息并通过输入通道拼接以及输出深度特征拼接两个层次对RGB图和深度图的信息进行多模态融合，显著提高系统的安全性。

### 三、方向：提升技术可靠性，促进与社会的融合

随着技术的成熟和先发行业的经验积累，计算机视觉在未来落地速度会进一步加快，落地领域也会更加广阔。为了促进技术和社会进一步融合，发展更大的社会价值，需要在以下方面得到突破。

**技术融通，加强高阶认知智能开发。**现在人工智能技术正从语音、文字、视觉等单模态智能朝着多模块智能发展，如何从端到端打通各个模态之间的关系形成可以真正多维度交互的智能机器，增强对功能属性、物理关系、因果逻辑、动机预测等的认知进行逻辑推断，让感知智能升级为认知智能。除了人工智能范围内的技术，其他数字技术与计算机视觉的交互也会碰撞新的火花，比如与物联网的交互将为计算机视觉带来更多由硬件采集的视频数据等。

**进一步提升对抗攻击和防御技术能力。**在对抗攻击和防御技术发展的今天，很多高效的方法已经被提出用于对深度模型进行鲁棒性的测试和提升，特别是对现实物理场景的攻击和防御，更加是未来研究的重点。因此未来会从以下几个方面深入探究对抗攻击和防御方法：1) 进一步研发隐蔽性较强的可以用于物理场景的攻击方式；2) 进一步提升攻击的迁移性，即较高的攻击成功率能够迁移到不同的深度模型上；3) 研究有较强通用性的防御方法，能够提升模型对于不同攻击方式的防御能力。

**辨别深度合成，鉴伪技术遏制AI滥用。**近年来，大量的人脸编辑检测算法被提出，在人脸鉴伪上取得了良好的性能，一定程度上遏制了AI技术的滥用。但是整体而言，目前的检测算法尚处于初级阶段，存在不少的局限性，数据集质量有待提升，可解释性、可迁移性等多个问题有待深入研究，以得到一个通用、高效、高精度的人脸编辑检测框架。

## 「 2.3 语音技术 」

### 一、动力：语音交互成为智能时代应用的新入口

语音技术在人工智能基础技术中举足轻重。2018年中国人工智能市场规模里，语音技术占比24.8%，仅次于视觉技术。相应的，语音应用市场规模也在高速增长，2018年达到了159.7亿元，近五年的年均复合增长率预计将达34.32%，2023年国内市场规模将达到664亿元。语音技术的发展和应用是人工智能的重要发展方向之一。目前语音技术广泛服务于通信、家庭服务、家电等多个行业，典型应用包括智能客服、智能穿戴、以及智能家居等。随着人机交互模式的范式转变，以及互联网应用的增多，语音技术相关的业务需求将进一步增大，业务对技术的反向促进作用将更加明显。

**一方面，应用载体、场景呈多样化发展。**语音技术的典型应用就是手机语音助手，如Siri。近年来，随着5G等其他技术发展，万物互联的态势已经逐渐显现，终端载体越来越多样化。以中国可穿戴设备为例，其市场规模(不包括医疗设备)在2018年达到近400亿元，根据测算，将在2022年达到607亿元。智能家居也是语音技术的重要应用场景。目前很多智能家居设备都具备语音交互功能，如智能音箱、智能空调、智能电视等。随着终端载体的多样化发展，语音技术的应用场景也将更加丰富，新应用带来的业务需求也将进一步加大，对技术的要求进一步提高，反向促进技术发展。

**另一方面，多模态的交互需求不断扩大。**人机交互模式一直在改变。除了传统的键盘交互、触摸交互，目前越来越多的设备也引入了语音交互的方式。交互模式每经历一次变化，背后的驱动力，都是对人和机器之间交互的便利性、自然性以及准确性所提出的更高的要求。融合视觉，语音等多种模态信息可以使得交互更加自然，因此多模态人机交互是一个非常重要的发展趋势，而该发展趋势也对多模态融合处理技术提出了更高的要求。

### 二、进展：不断逼近“自然交流”的三大变革实践

业内专家认为，语音技术的链条日臻完善，边界在逐渐扩大，现有的研究中，更加注重端到端的建模方法，以及领域的通用性；不少领域已经有了较成熟的应用，包括传统的语音输入和语音搜索，如手机语音输入转文本、腾讯视频语音搜索等；智能家居及车载，如腾讯听听音箱、王者机器人音箱、叮当智能屏幕等；音视频内容分析审核，如广告审核、视频字幕生成、内容挖掘与推荐等。具体的演进方向如下：

**从单模态到多模态语音交互。**近年来，就语音单模态而言，从近场语音交互，到远场语音交互均有了较大的突破。目前，语音交互正在向多模态交互场景延展。

第一阶段：近场语音交互。在这一阶段，用户可以在手持设备的状态下，与设备进行交互。我们生活中常用的手机语音助手，就是典型的近场语音交互。

第二阶段：远场语音交互。在这一阶段，交互的有效性可以扩展至更大的距离空间，通过麦克风阵列前端处理算法，使得在一到十米的距离人与机器也可进行语音交互。典型应用是智能音箱，目前教育机器人、智能厨电中也在积极尝试远场语音交互。

第三阶段：多模态交互和新交互。目前，我们正迎来语音交互的重要变革阶段。在单模态系统中，如智能音箱，智能车载等，需要首先进行唤醒，然后进行单轮或多轮交互。而采用多模态引入视觉等信息后，在此基础上叠加语音技术，更准确地理解用的需求，判断用户的意图，感受用户的情感，最终为用户提供一个更自然、高效的多模态交互体验。

**语音环境适应性持续提升。**尽管当前业内语音开放API都宣称能够达到97甚至98以上的识别率，但是在有复杂背景噪声，如其他多人说话及音乐背景干扰下，其识别率仍明显低于人的水平。而且，随着5G时代的到来，场景将多样化发展，音视频内容将大量增加，多人及复杂噪声场景下语音技术应用的痛点亟需被攻克。

通过在语音分离，降噪增强等多个方向技术的持续改进，将有助于解决多人及复杂噪声场景下的语音检测、识别需求，提升系统的通用性。当前，在覆盖如朗读式，口语等不同表达方式，干扰人及音乐等复杂环境噪声，近场、远场条件下等各种条件下的不同测试集合上，通用识别性能整体相比以往都有大幅提升。

在远场语音识别场景中，实现了对不同麦克风阵列形态设备的算法支持，同时在远场多麦克风阵列信号处理方面改进，利用一种新的Multi-Tap MVDR (MT-MVDR) 波束成型算法，相比传统的MVDR，新的算法除了利用空间信息之外，还利用相邻帧之间的相关性，在确保增强信号无失真的前提下，可以有效降低残留噪声，其次还提出了复数值掩蔽替代传统实数值掩蔽来估计MT-MVDR的方差矩阵统计量。相比纯神经网络方法，远场识别错误率能够获得相对40%的降低。

**从音频内容分析到音频内容生成。**一直以来，语音技术以训练和识别为主，近年来，随着场景不断复杂化，生成已经成为了重要的方向。目前，语音合成、歌声合成上已经有一些不错的效果和应用，如虚拟偶像、虚拟解说、虚拟主播等正走入人们的视野。

歌唱合成效果改善，实现了歌唱合成系统初步产业化探索。推出的歌唱主播AI“艾灵”，拥有了近乎真人声线，能像人类一样情感充沛地歌唱。除了声线，“艾灵”还拥有用多模态智能技术搭建的，高灵活度的数字躯体。

合成效果媲拟真人，实现了能真正能在线上应用的端到端合成系统Durian，解决了端到端系统中漏字重复等不稳定的问题。在此基础上实现了综合的多模态数字人的合成技术流程：系统首先从文本中提取各种各样的信息，包括动作、表情、情感、重音位置、和激动程度等。之后将这些信息输入到动作表情模型，生成动作和表情，同时输入给多模态合成系统 Durian 同步生成语音以及口型和表情参数，进而合成真人或者卡通形象。相较于传统语音合成方法以及目前最新的端到端语音合成方法，无论是在自然度、鲁棒性、可控性、泛化能力以及实时性等方面都取得了更好的效果。

### 三、方向：探索产业与技术相互促进的发展模式

未来，多模态、多场景需求如果得到更好的满足，将给业界带来更多的想象空间，这需要从技术、业务两个方向进行努力，形成相辅相成的螺旋上升轨迹。

#### 从技术出发：归根到底是要解释更多的“复杂性”

语音技术的规模应用才刚刚开始。随着5G、云计算等新一代基础设施的完善，语音技术将在更多的新场景、新应用中落地，从而更好地服务于人类生活。虽然说当前语音技术已但是要在以上领域取得进步，语音技术还存在以下的“点”需要被攻克：

一方面是鲁棒性仍需提升，在对多种领域变化（如表达方式、噪声环境、通道失真）的鲁棒性方面与人类水平相比，进一步逼近；一方面是语音及歌声合成，包括语音合成风格应更加多元、歌声合成效果更自然动听；另一方面是多模态交互，主要是多模态融合能力提升，以便适应于不同交互方式与应用场景的多元需求。最后是成本与效率的问题，寻找更灵活轻量的采集与生成，通过少量的、少标注的语料去更快、更高质量的完成各类语音任务，将是推动应用爆发的一大关键。

#### 从业务出发：场景应用与数据反馈是最好的“产品经理”

语音技术未来发展模式将与产业发展紧密相连。产业界为语音技术的应用落地，提供真实场景和数据，而应用中的数据又对技术形成反馈机制，帮助语音技术升级优化，不断适应产业界关注的整体解决方案需求，完成行业的应用迭代，最终又回到场景落地，在这样的闭环中，共建数字生态共同体，实现技术和产业的互相促进。

一方面，挑战重点场景，进一步扩大语音技术的应用范围。在智能家居、智能政务、智能出行等经典场景的应用下，进一步探索AR/VR/MR等场景的应用方面；另一方面，语料数据是语音训练的一大门槛，随着数字生产要素的市场化机制的不断探索，开放性数据所带来的生产力，将有望为语音技术的训练、迭代，带来重大飞跃。

## 「 2.4 自然语言处理 」

### 一、动力：重点基础技术底座上的高频次需求

始于20世纪50年代，图灵提出著名的“图灵测试”引出了自然语言处理的思想，而后经过半个多世纪的跌宕起伏，历经专家规则系统、统计机器学习、深度学习等一系列基础技术体系的迭代，如今的自然语言处理技术在各个方向都有了显著的进步和提升。作为人工智能重点技术之一，自然语言处理在学术研究和应用落地等各个方面都占据了举足轻重的地位。比尔·盖茨曾说过，“语言理解是人工智能皇冠上的明珠”。谁掌握了更高级的自然语言处理技术，谁在自然语言处理的技术研发中取得了实质突破，谁就在日益激烈的人工智能军备竞赛中占得先机。

**文本理解精度提升：**如问答系统中对知识库的构建和查询，搜索引擎中对查询词的理解以及扩展，信息流推荐中关于内容的理解和分类，电商用户评论的情感理解与判断等等。随着场景的丰富、复杂化发展，对文本理解能力的提升需求进一步增强。

**机器翻译准确性增强：**自动翻译产品作为跨语言沟通的桥梁发挥着重要作用。用户对自动机器翻译系统的准确性和实时性要求越来越高，特别是长文本翻译、篇章翻译、特殊领域文本翻译等场景。此外，自动翻译有时并不能直接应用于生产。这要求我们一方面提升机器自动翻译的效率和准确率，另一方面研发更专业更好用的人机交互式翻译系统。

**对话系统的智能化：**近几年随着智能音箱等终端进入千家万户，用户对自动对话系统的兴趣也日益浓厚，对系统的智能化有了更高的期许，例如更精准的回答、定制风格与人设、完成更多的任务等。而且，对话系统也可以扩展到多个业务场景中，例如心理咨询、智能客服、在线教育、虚拟主播弹幕回复和聊天等等。

**文本生成的实用性：**实际业务需求中有着各式的文本生成任务，例如对联生成、诗词生成、歌词生成、句子改写、体育比赛或游戏竞技解说词生成等。一方面整合现有文本生成技术迎合场景需求，另一方面，多样的充满挑战的实际需求也可以反过来进一步推动新型文本生成技术的攻关和落地。

### 二、进展：从“用、学、传”三方面推进技术进步

通过在消费互联网20余年的技术积累，腾讯在学术研究、产业应用以及生态建设等方面，具有一定的平台基础与规模优势，在自然语言处理技术开发获得了一些突破：

**学术科研上的技术模型、理论推进。**2016年以来，腾讯在自然语言处理相关的顶级会议和期刊（ACL, EMNLP, NAACL, AAAI, IJCAI, TACL等）共发表论文180余篇。同时，在各项技术比赛中，均获得了多个世界顶级奖项。AI Lab在2017年国际知识图谱构建大赛（KBP）中获得实体发现与链接（EDL）任务的冠军，并曾经占据ARC、OpenbookQA等问答与阅读理解榜单的第一名。腾讯云智能平台产品部在今年的对话系统技术挑战赛（DSTC8）上，参与了三个赛道的七项任务，共获得四项指标世界第一，两项指标第二的优秀成绩。微信AI两次获得国际对话系统技术挑战赛（DSTC7, DSTC8）赛道冠军，也在RACE、DROP、QuAC等多个代表性机器阅读理解榜单上获得第一名。

腾讯在自然语言处理方向的研究主题涵盖语义理解、文本生成、智能对话、机器翻译等领域。例如，在语义理解领域，提出了一种新的词向量训练方法，为后续大规模中文词向量的发布提供了可靠的算法保证。在对话系统中，提出了一种检索和生成相结合的智能对话框架，可以自然地融合检索方法和生成方法各自的优点，大幅度提升了对话系统的质量。在文本生成中，提出了一种可以根据任意模板约束进行生成的框架，弥补

了传统模型只能进行开放式生成的不足。在机器翻译中，首次提出了篇章级神经网络机器翻译框架，极大缓解了翻译一致性和歧义性问题。在交互式翻译系统中，提出了一种新的可以容错的约束解码框架，进一步提升了交互式翻译系统易用性和准确性。

多项技术工具与服务应用。很多科研问题都来源于各个业务场景中的实际问题，科研产出的模型和策略上线后可以进一步优化和提升业务效果，达到了一个良性的循环。

在知识图谱方向，我们构建了包含上亿实体和20多亿三元组的通用领域知识图谱Topbase，应用于微信搜一搜、信息流推荐、智能问答等业务中。

在文本理解方向，推出了自然语言理解系统TexSmart，用以对中文和英文两种语言的文本进行词法、句法和语义分析。除了支持分词、词性标注、命名实体识别（NER）、句法分析、语义角色标注、文本分类、关键词提取等常见功能外，TexSmart还提供细粒度命名实体识别、语义联想、深度语义表达等特色功能。

智能对话系统方向，通过融合检索模型、生成模型、深度语义理解、知识图谱、对话管理等技术，实现了包括任务型对话、知识问答、智能闲聊、智能客服等功能服务。目前AI Lab和腾讯云智能平台产品部的智能对话服务已经接入腾讯云AI开放平台、腾讯云小微开放平台、腾讯游戏知己等平台，激活数过亿。微信AI以“微信智言”和“微信智聆”平台为基础，广泛的为智能硬件、PaaS、行业云和AI Bot等领域提供技术支持。

在自动翻译方向，腾讯AI Lab创新性的推出智能交互式辅助翻译工具 TranSmart 来帮助翻译人员提高翻译效率，专门针对自动翻译质量欠佳，需要人工参与的情况。近日也与业界各大翻译公司进行合作，将 TranSmart 的服务接入他们的系统当中。微信翻译引擎研亦从2017年开始，基于业界领先的深度神经网络机器翻译引擎技术打造，引入、形成最前沿的机器翻译算法，先后承载着微信中对话、朋友圈场景中的文本和图

片翻译，以及扫一扫翻译等重要业务，成为一个日翻译数十亿字符的大型多语言机器翻译引擎。

**在资源传递上推进数据库和模型分享。**腾讯AI Lab在2018年10月份向业界发布了一个大型的中文词向量数据，迅速被用在了多个业务系统中。该中文词向量数据包含800多万中文词汇，其中每个词对应一个200维的向量。相比现有的中文词向量数据，腾讯AI Lab中文词向量大大提升了覆盖率、新鲜度、准确性，大幅度改善了其质量和可用性。目前词向量优化工作还在进行当中，包括增量、去噪、加速等，以及准备高质量的英文词向量数据。另一方面，其它的方向也在不断尝试，如微信AI也发布了FewRel 2.0和DocRED数据集，以推动对少样本和长文本场景下的信息抽取技术的关注与研究。此外，AI Lab还在早些时候对内开源了多种大规模中文语言模型，包括可以改善多个下游任务的BERT模型，可以大幅度提升文本生成多样性的GPT2模型。相比其他的模型，这些模型最大的特点是通过在多种领域的大规模语料的训练，可以进一步提升其在下游任务的准确率。

### 三、方向：在变化中探索如何突破质量、场景和表达限制

近些年，随着神经网络技术和深度学习的发展，较统计机器学习时代，自然语言处理技术已经发生了翻天覆地的变化。从word2vec模型，到后来崭露头角的Elmo，再到后来Google提出BERT这种划时代的语言模型，大幅度提升了自然语言处理多种任务的性能，将文本理解和语义表示推向了一个新的高度。从seq2seq+attention框架在机器翻译上得到的巨大进展，到该模型迁移到对话生成、摘要生成等文本生成任务上的所带来的可喜效果，再到GPT2/GPT3模型在开放域文本生成中带来的进一步的提升，自然语言处理技术开发惊喜不断。未来，仍有许多更有价值的学术与应用方向需要我们进一步去探索：

**提升文本理解的精度及深度。**近几年文本表示学习、预训练等技术的发展使得文本理解的精度有了一个明显的提升，但是AI在文本理解的精度和深度上和人类相比仍然有很大的差距，这直接影响了上层应用的效果和智能水平。比如现在智能对话机器人还停留在“人工智障”的水平，其根本原因在于不能从语义上准确地理解用户的问题。

**提高文本生成质量。**虽然随着深度学习模型的发展，尤其是得益于注意力机制的出现，机器翻译、自动摘要以及故事生成等领域的效果已经得到了长足的进步，但是相较于人类水平还有较大差距。如机器翻译领域中的漏翻错翻现象，以及故事生成中的重复生成、逻辑错误、篇章错乱等问题。因此，文本生成的质量还需要进一步提升。

**增强低资源和多语种场景的应用效果。**目前许多自然语言处理任务为了达到较好的效果，都需要利用海量的标注数据进行训练。但是产品出海以及本地化定制等过程中，常常出现在某些语言上训练语料不足的问题。当数据资源有限时，如何在自然语言任务上取得较好的效果是一个亟待解决的问题。

**探索更丰富的语义表达方式。**深度学习的应用虽然使得众多自然语言处理的任务性能得到了提升，但是如何设计更好的语义表达方式依然没有得到解决。将一般性文本转化为合适的语义表达形式，有助于真正的提高模型的理解和推理能力。

**强化知识表达泛化能力。**神经网络在大规模语料训练的过程中，并没有显式的将知识进行合理的结构化组织，这导致了模型领域泛化的能力较弱等问题。将知识利用知识图谱等形式进行组织，在模型预测的过程中进行知识融合，将极大的帮助模型提高泛化能力，是未来一个重要探索方向。

**优化语言生成和对话的度量。**一方面，在开放域文本生成问题中，如开放域对话和故事生成，生成结果的质量仍需要人工评价，尚没有找到合理的自动化度量方法。另一方面，在自动翻译以及文本摘要等领域，虽然有了统一的自动化度量指标，但是这些指标有时并不能完全反应生成结果的质量。未来，文本生成等领域需要一系列更好的度量方法。



## — 3 —

# 智慧抗疫篇

新冠疫情爆发，为社会经济带来了严峻的考验。面对本次疫情中出现的种种问题，以人工智能为首的数字技术为战疫提供了重要的工具，在医疗、制造业、服务业、城市治理等领域的应用价值持续释放。疫情也成为了AI的试金石，是一次国家级别的数字化全面检验，从基础设施的完备性、数据的流通性到快速反应能力的一次全面测试。数字技术在疫情中经受了一轮考验，验证了其是否对社会有真正的价值。人工智能不再是以往的旁观者，而是出演关键角色，投入到战役战场，提高抗疫战争的整体效率。在疫情防控、疾病诊疗、城市治理等诸多方面都发挥了积极的作用，并在后疫情时代全面支撑复工复产和数字经济的发展。数字化、网络化、智能化的应用让中国经济、社会在疫情面前变得更有韧性。长久来看，疫情过后，人工智能可能真正逐渐转变为像网络、电力一样的基础服务设施，无所不在，无所不含，真正渗透到社会经济的方方面面。

## 「3.1 AI+ 医疗：抗击疫情一线，缓解医疗资源压力」

我国医疗行业面临诸多痛点，在供给侧，医疗资源紧张且分布不均衡、优质医生短缺导致超负荷工作、基层医院医疗水平有待提高；在需求侧，随着人口老龄化的加剧，慢性疾病患者数量逐渐增加，医疗需求进一步加大。2018年我国三级医院占定级医院总数量的19%，却承接了55.2%的就诊需求。而2020年伊始爆发的新冠肺炎疫情，导致医疗需求激增，将医疗行业供需双方矛盾进一步放大，行业痛点亟需解决。人工智能在医疗领域中图像识别、医药筛选、远程问诊场景下的落地应用，帮助抗击疫情，彰显了AI+医疗的力量。作为新一代基础设施建设，人工智能在医疗行业的应用将对传统医疗机构运作方式带来变革，从长远有效缓解医疗资源压力。后疫情时代，AI+医疗有望迎来大发展。

### 新冠疫情加速人工智能在医疗行业的应用

#### AI落地图像识别，提升医疗效率

医学影像是医疗行业的重要辅助科学。但是传统医学影像由医生读片，导致诊断速度缓慢，且对相关领域专业人才需求大，人才短缺是重大难题，并且人工工作繁重也会导致误诊漏诊的发生。AI图像识别的应用，有助于解决解决医疗影像领域存在的此类痛点。在医学影像场景下，人工智能技术可以解决三种需求：

第一，AI图像识别满足病灶识别与标注需求。AI医学影像产品可以对X线、CT、核磁共振等影像进行分割、特征提取、定量和对比分析，完成病灶自动识别与标注，发现肉眼无法发现的病灶，降低诊断结果的假阴性概率，并且，AI医学影像产品可以在几秒内处理十万张以上影像，将大大提高诊断效率。

第二，AI图像识别解决靶区自动勾画与自适应放疗需求。这一需求主要针对肿瘤放疗环节的影像处理，在这一环节，AI医学影像产品可以帮助放疗科医生对200到450张CT片进行自动勾画，将时间缩短到30分钟一套，并且在患者放疗中不断识别病灶位置变化以达到自适应放疗，减少射线对病人的伤害。

第三，AI图像识别解决影像三维重建需求。AI可以基于灰度统计量的配准算法和基于特征点的配准算法，解决断层图像配准问题，从而节约配准时间、提高效率。

对患者而言，AI图像识别可以帮助其更快完成检查流程，更早获得更可靠的诊断结果；对于医生而言，AI图像识别将大幅缩短其读片时间，提升工作效率并且降低误诊率；对医院而言，AI图像识别有助于提升医院整体诊疗水平，并且降低医院在人工方面的成本与压力。AI+医疗产品的落地应用，体现了人工智能技术在医疗发展中的重要作用，对于我国全面提升医疗资源配置效率，缓解医疗资源压力意义重大。搭载腾讯AI医学影像产品——腾讯觅影AI和腾讯云技术的人工智能CT设备在新冠肺炎疫情期间，在湖北多家医院进行部署，帮助医护人员进行诊疗。在患者做完CT检查后，设备数秒就可完成AI识别，在一分钟内为医生提供辅助诊断参考，诊断效率提升数倍，有效缓解了疫情初期医疗资源严重不足的严峻形势。

#### AI应用医药筛选，助力新药研发

新冠肺炎威胁的是全球发展，疫苗以及相关特效药的快速研发非常重要。医药研发一直以来存在研发周期长，成本高，成功率低的问题。在高投入的情况下，医药研发却很难得到理想的回报率，特别是在药物发现阶段，需要进行大量的数据分析、筛选等工作，才能有一个正确的开端。人工智能技术能够大幅缩短药物筛选时间，提升药物筛选的精准性，从而助力科研机构进行药物研发。

在药物发现阶段，人工智能主要在靶点筛选和化合物合成与筛选两个场景下发挥作用。靶点筛选是指发现能减慢或逆转人类疾病的生物途径和蛋白，是新药研发的核心之一。在这一方面，人工智能通过自然语言处理技术（NLP），深度学习海量医学文献和处理海量相关数据，进而发现化合物与疾病之间的关系，找到靶点，从而缩短靶点的发现周期。

化合物合成与筛选是指将上百万的小分子化合物进行组合实验，来研究具有某种生物活性和化学结构的化合物，用于进一步的药物研发。人工智能可以模拟小分子化合物的药物特性，从而可以在数周内完成最佳模拟化合物的挑选并进行合成试验，能够在快速过滤低质量化合物的同时富集有效分子。

目前，包括中科院在内的多个研究所已经将人工智能技术注入药物研发，新冠肺炎疫苗以及特效药的问世指日可待，人工智能在医药研发上的优势，值得更多发掘。

### AI融入远程问诊，解决就医需求

新冠肺炎传染性强，入院问诊存在交叉感染的风险，对有问诊需求的患者造成不便，并且疫情对公众带来一定的心理恐慌，导致人们害怕去医院，但是又亟需专业的医学信息来解决问题。远程问诊、线上医学科普可以有效解决这一难题，人工智能在这一过程中也扮演着重要角色。

腾讯云基于医疗行业语料及医疗专业词汇，打造医疗行业语音识别模型。医生无论在门诊、住院查房、交接班等场景下，均可使用语音输入软件，将传统的手写病历转换为语音输入，大幅度节约病历录入的时间，减轻工作负担。NLP的关键技术包括分词、句法解析、信息抽取、词性标注、指代消解、词义消歧等。在线问诊的过程中，当用户在线上问诊平台输入症状，AI系统将识别用户输入的文本，并完成分词、词性标注、句法解析、信息抽取等一系列工作，最终在知识库中进行检索，把类似信息推给用户，完成精准的信息匹配。

腾讯医典是腾讯出品的医学科普资源平台，平台内从两个层面注入了AI技术：基于深度语义分析、信息抽取技术构建的医学问答知识库；基于NLP和深度学习技术准确判断用户意图，通过追问引导用户补充有效信息，以提供精准回答。腾讯医典在疫情期间提供免费在线义诊、AI辅助自查工具、以及权威新冠肺炎医学科普等功能，一定程度上缓解了疫情期间紧急的问诊需求。

人工智能技术推进远程问诊与医学信息在线科普发展，人们可以更加高效、快捷的触及医疗资源，长远来看，在疫情结束后，也将对缓解我国紧张的医疗资源及区域分布差异起到重要作用。

### 疫情后AI医疗的应用前景

新冠肺炎疫情期间，AI应用的落地在一定程度上缓解了医疗资源紧张、不均的问题，促进了医疗技术、诊疗能力、药物研发的发展，提升了医疗体系的运转效率。疫情之后，人工智能在医疗领域也有望在以下方面进一步发展应用：

新冠肺炎疫情暴露出我国在公共卫生应急方面仍存在短板，构建完善的公共卫生应急体系至关重要。国家卫建委表示，针对公共卫生应急体系建设，要运用大数据、人工智能、云计算等数字技术，在疫情监测分析、病毒溯源、防控救治、资源配置等方面更好地发挥支撑作用。国家层面充分重视新技术的应用，人工智能将在全国公共应急体系的构建中发挥更大作用。

个人健康管理将进一步得到社会重视。人工智能在加强用户个人健康数据与医院健康数据的关联性上将发挥更大作用，构建个人健康管理平台，通过深度学习等技术对海量健康数据进行分析，实时动态监测用户健康状态，从而提供精准的健康监测方案，使得人人身边拥有虚拟医生。

基层医疗卫生机构的弊病在疫情中更加显现，而人工智能的应用将有效解决基层医院医护人员短缺，医疗水平不足的问题。未来人工智能在基层医院应用前景广阔，将服务于早筛、诊断、临床等多个应用场景，助力缩小城乡医疗水平差异，提高中国整体医疗水平。

人工智能技术已经在新冠肺炎药物研发中发力，大幅缩短药物筛选周期，而药物生产流程的提质增效也有助于提升整个生产、销售、使用的全链价值。未来，人工智能技术将不仅作用于药物研发，还将作用于药品智能销售、精准用药，促进个性化诊疗的发展。

受疫情影响，医疗保险类服务需求也将进一步增加，人工智能技术可在风险控制、医保控费、福利管理等领域发挥数字化、智能化作用，促进医疗生态圈中的多方协同发展。形成以医疗服务为主导，配套支付、福利服务为辅的智能医疗生态圈，进而优化整体医疗服务环境，提升医疗服务质量和效率。

## 「3.2 AI+ 城市：基础设施升级，融入城市治理体系」

新冠肺炎爆发突然，疫情整体呈现出传染性强、易扩散、风险大的特点。我国作为人口大国，城市化发展持续扩张，拥有人口超过一千万的超级城市13座，城市的防控工作量巨大且时间紧迫，城市治理面临重大挑战。首先，疫情具备突发性，而许多城市的应急体系并不够完善，如何在短期内做好决策、保障政务服务，是城市管理者首先需要解决的问题；其次，城市人口密集，人员流动量大，在这样的背景下，如何全面监疫情也是城市管理者面临的难题之一；最后，抗疫过程中社区管理的作用显得非常重要，但是不少社区依然采用传统方式严防死守，需要消耗大量人力，全面提升社区防疫能力、精细化管理显得十分必要。

### 人工智能帮助突破疫情下的城市治理模式

疫情之下城市治理暴露出其短板，人工智能在智慧政务、疫情监控、社区防疫等方面都在一定程度上帮助城市管理者突破了其原有的治理模式，使得智慧城市建设提速前行。智能治理需要充分利用人工智能等信息技术，对城市治理体系和社区管理方式进行智能化升级，用“网状防控系统”，有效对抗网状传播的病毒。“智能治理”和联防联控机制在中国战疫过程里，发挥了重要作用。

#### 空间下沉：人工智能提升社区防疫能力

社区防疫是城市防疫的基础环节之一，而社区封闭化管理，需要大量保安、物业、以及志愿者，在避免接触的同时，提升登记排查工作效率非常重要。而复工复产开始后，人口流动性大幅度提高，更加为社区防疫带来挑战，考验着社区的智能化管理水平。为了帮助社区防疫，避免交叉感染，腾讯优图联合腾讯海纳推出健康申报、人车管理、体温筛查以及智能隔离等抗“疫”功能，在全国多个小区进行落地应用。如：

(1) 健康告知与申报承诺。针对健康在线申报，利用腾讯优图AI核身技术，用户可通过微信在线实名验证，提高健康申报效率，且核验身份后签署承诺书更具有法律效率；

(2) 电子出入证+AI识别开门+自动体温测量。针对封闭式小区出入口、开放式重点区域，腾讯优图AI识别技术可辅助社区人员在线申请电子出入证；同时，集成该技术的热成像摄像头、热成像门禁机等设备，可实现实时身份识别、测量体温等预警记录上报；

(3) 车辆出入管理和行为预警。腾讯优图车辆识别技术能够对社区以外的出入车辆进行自动化车辆识别登记，同时对出入社区的湖北（疫区）牌照车辆进行提示，并对湖北牌照但近期并未进入湖北的车辆进行白名单处理。还能根据车辆出入的时间、频次等行为特征，对车辆出入进行分析预警；

(4) 预警疫情风险、人群聚集告警。基于腾讯优图AI识别技术，平台能够及时发现未佩戴口罩的人员，统计社区内人员口罩佩戴情况，协助社区居民管理，降低疫情传播风险。同时，腾讯优图AI识别技术还可以用于社区公共区域，自动识别人员聚集情况，设置告警阈值，提前预知疫情风险，协助社区管理者合理调度安排执勤人员等。

#### 便民服务：人工智能加速智慧政务

抗疫期间，全国为了更好的对疫情进行防控，也为了更好的服务疫情期间的出行、复工，腾讯推出“防疫健康码”，民众通过微信小程序申请涵盖自身健康信息的二维码，获得电子出行凭证，方便民众在新型冠状病毒肺炎疫情期间出行。当前，全国各城市陆续按下“复工键”，为了保证疫情期间安全复工复产，无论是跨省市返工通行，还是进出办公场所都需要出示健康证明，健康码在其中起到关键作用。

无接触办公、服务在疫情期间得到广泛推广，智慧政务的重要性再次彰显。在这一特殊时期，人工智能在智慧政务领域的应用优势很好地展现，一方面运用技术优势，保障居民所需的政务服务；另一方面，借助人工智能力量，提升政府工作效率，加强应急管理。

人工智能在智慧政务方面的应用，除了自然语言处理，计算机视觉技术外，人机交互中的语音交互技术也发挥了很大作用。腾讯推出的AI产品“政务联络机器人”，可以与辖区居民联络，进行人机对话，完成相关政务信息的传递工作，以及对居民关于疫情的困惑解答。智能机器人还可以自动生成疫情统计报告，展示通知、排查结果，在节约人力成本，避免信息采集人员与居民交叉传染的风险的同时，进一步提高防控疫情效率。

### **智慧治理：人工智能加强疫情防控能力**

新冠疫情发生时，正值春运高峰期，人员流动大，给城市疫情监控带来了巨大的困难，智慧城市管理急需进行升级，加大技术的参与度，使得疫情监控更加智能化，减轻疫情防控压力。

计算机视觉技术在红外体温监测和人群迁徙态势感知中都得到广泛应用。AI全自动红外测温仪利用普通摄像机识别并锁定人脸，而后通过红外测温仪采集红外数据，并计算出人体温度，具备非接触、大面积、快筛查的优点。人群迁徙态势感知平台，可以通过地图、3D视觉、2D图表等可视化方式，将覆盖全国地域的人群迁徙及画像统计信息可视化呈现，帮助城市治理者更好地掌握目标区域人群情况，以便制定更精准的疫情防控方案。

全国多个省份的健康码都会用到身份核验的过程，功能调用了腾讯云慧眼的后台认证能力。比如深圳公安、山西公安、云南公安等一系列的公安机关，还有人社、税务、政务综合、住建等政务机构，都对接到了腾讯云慧眼。

### **案例：AI中台助力智慧城市构建**

腾讯云AI中台是基于自有云智天枢（TI Matrix Platform）打造的开放式人工智能平台，可以快速接入各种数据、算法和智能设备，并提供可视化编排工具进行服务和资源的管理及调度，进一步通过AI服务组件持续集成和标准化接口开放，帮助开发者快速构建AI应用。在

城市治理体系和智慧城市构建过程中，涉及政务、公安、交通等多行业领域，腾讯云AI中台凭借算法/模型等快速接入能力，以及平台与应用联动能力，构建全链条AI技术与应用开发生态，推动各个行业AI中台快速落地，为城市智慧化进程提供算法+平台+应用+服务在内的一站式AI解决方案。目前腾讯云AI中台已在多个WeCity和数字政府项目中进行实践，推动智慧城市建设规模化发展。

### **智慧城市升级方向**

此次疫情防控中，人工智能技术在城市治理方面广泛落地应用，表明我国智能社会形态正在逐渐显现。而人工智能已经被明确列为我国新型基础设施建设之一，未来，随着人工智能技术与5G、云计算、大数据等新一代技术的进一步融合，将推动智慧城市的进一步发展与升级。

政府决策将更加科学化：人工智能技术在疫情防控、风险分析预警等方面的应用，展现出其在辅助政府决策方面的优势。未来，结合大数据、云计算等技术，人工智能强大的数据采集和智能分析能力将进一步展现，社会运行体系的规律性将更容易被探索，社会需求变化趋势将更容易被监测，从而辅助政府进行科学决策。

城市管理更加可视化：计算机视觉技术在防疫中广泛应用，可视化的呈现方式让城市管理更加直观。人工智能技术将城市的动态运行映射在多维、动态的数据体系中，从而满足城市管理的实时化、数量化，并最终以可视化形式呈现。城市的可视化管理将不仅停留在疫情期间，更将成为未来城市管理的常态化。

社会服务将更加智能化：疫情期间，政务机器人的应用让我们看到了智慧城市向智能化发展的曙光。未来，随着技术的进一步发展与成熟，人机交互领域将不仅局限于语音交互，脑机交互、情感交互等也将逐步成熟落地，各类设施将更加智能化，社会服务整体呈现智能化发展趋势。

## 「 3.3 AI+ 复工：智能化转型助力复工复产 」

随着疫情防控进入常态化，统筹社会经济发展修复成为了最关键的议题。时下，新冠肺炎疫情席卷全球，国际经济不确定性剧增，企业复工复产“步履蹒跚”：人员不足、资金周转困难、生产资料匮乏、防疫物资短缺等等，很多企业在疫情中也进入到生死时速的时刻，智能化转型成为了制造业和服务业提高抗风险能力的发展方向。

### 智能制造：新型人机协作模式的建构

疫情对于制造业的传统管理模式提出了很大的挑战：

首先，人员大量被隔离造成企业在正常复工时候一定会遇到人员紧缺问题，尤其在自动化程度不是很高的企业，用工荒的问题将更加显著，员工少和生产需求大产生较大矛盾；其次，半导体、液晶面板等重资产、高价值的制造行业，由于其特殊性，一旦停产损失巨大。这种情况下，外部疫情的变化以及人员的管制，对于员工的身心都是一种压力和负担，如何在情绪变化和高压下保证员工的作业以及判断的稳定性很重要；再者，尽管国内现在已经管控的不错了，但海外国家现在状况还不乐观，然而有一些高科技行业的外部专家都在国外，这一定会影响国外专家到现场进行问题的诊断和解决。

其中可以看到，劳动力这一生产要素成为突发情况中最难以把控的一点。人力短缺带来的压力不只是在疫情期间爆发出来，在未来经济周期中，这一问题会持续存在。制造业的未来发展方向，人机结合注定是大势所趋。

#### 计算机视觉快速提升缺陷检测

计算机视觉通过用摄影机和电脑代替人眼对目标进行识别、跟踪和测量等机器视觉，并进一步做图形处理，使电脑处理成为更适合人眼观察或传送给仪器检测的图像。在计算机视觉的成熟可以大范围替代工厂中缺陷检测等环节中的人力。

腾讯优图实验室联合腾讯云为一家国内大型面板生产企业打造的工业AI项目，主要内容是为液晶面板做缺陷检测，利用视觉AI的算法做产品的缺陷检测，从而替代传统的人工质检。上线的模型可以达到人工检测相同水平的准确率，并大幅提升生产效率。

#### 知识图谱助力工业维保

知识图谱与机器学习技术的结合是完成技术沉淀的优先方案，将需要的知识数据（结构化或非结构化数据）以图谱的形式进行展示。通过对历史数据的学习，建立起一套可操作性强的知识图谱系统，达到即使在相关专家不在的情况下，也能实现对维修需求快速反映和指导反馈的作用，利于工人维修效率的提高。

#### 案例：一站式AI平台助力远程协作

TI-Insight是腾讯基于智能钛核心功能打造的一站式工业AI平台方案，包含AI训练系统和AI推理系统两个功能组件。主要应用于缺陷检测、视频检测、质量监控等图像视觉的应用场景。支持算法工程师、及具备有限深度学习知识的业务用户可以从0到1快速构建模型、1到N快速迭代训练模型。同时，本平台提供优化调度算法微服务能力，可以帮助团队快速地部署模型，高效利用硬件计算资源，提高生产力。

在抗疫期间，华星光电ADC项目团队在无法现场工作的情况下，与腾讯展开合作，远程安排10+位人员顺利完成模型交付、平台升级等工作，确保业务不间断、无差错。在这一平台上，团队成员可以每日查看上线模型报表、监控服务，保障训练平台和推理平台的顺利运行，避免发生生产事故；完成1支模型的算法优化和服务上线，并对15支模型进行了训练、相交断线以及服务封装等工作，使得模型覆盖率显著提升，持续为华星光电降本增效；同时，还顺利完成推理平台的扩容，提升了30%吞吐量，并及时解决扩容中出现的问题，保证工厂业务稳定运行。

## 智慧服务：产业服务智能化转型

金融、电信、呼叫中心等聚集大量业务咨询办理与客户服务场景的行业，对于服务人员的依赖性很强。这类工作的结构化流程化特征明显，工作内容并不复杂，但是重复性较强且工作量大，在面对各种情绪服务对象时十分考验客服人员的耐心。而人工智能对于这类工作的适用性极高：其结构化的特点适合人工智能进行学习，而具有规模化、不会劳累、不会有负面情绪的人工智能适合解决这类工作中存在的痛点。

近年来，随着语音识别、语音合成、NLP等技术的发展，人机对话能力持续提升，智能语音客服开始逐步在办事大厅、呼叫中心等场景投入使用，并取得了不错的效果。这不仅优化了服务体验，减少了客户等待时间，还能大幅降低企业运营成本，疫情发生后，线下服务受到极大的冲击，基于AI平台的远程智能服务发挥了重要作用。

服务业从来不是以单一形式进行呈现，热情周到的服务体现在文字、语音、交互等方方面面，而人工智能的融入也从这些角度入手，从多维切入，渗入服务业方方面面，全方位打造新型服务模式。



### 文字识别OCR提升服务连接效率

文字识别（Optical Character Recognition, OCR）基于深度学习技术，利用光学字符识别技术，将图片上的文字内容，智能识别成为可编辑的文本。常见的可识别内容包括通用文字、卡证文字、票据单据、汽车相关、行业文档等多场景下的印刷体、手写体文字识别，在政务、互联网、金融、企业等多个行业均有落地场景。在公共服务领域，需要办理相关业务时，民众不再需要去很多部门实地办理，只需在家通过扫描上传自己的证件即可，为实现政务少跑腿、办实事打下技术基础。

基于腾讯优图的底层技术，腾讯云打造的OCR技术服务发展成熟，在各行业有数十个标杆案例，落地案例丰富，以50多项的能力位居业内第一；支持19种语言识别，识别率高达95%。在疫情期间，江苏省南通市各大社区通过使用腾讯云卡证类OCR服务，让用户在家通过上传证件-识别信息-完成在线办理通行证，极大的缓解了办理通行证的压力，提升了用户在疫情期间出行体验，助力复工复产。每天的办证效率提升了10倍，而且减少了人群的聚集，效率提升的同时，还有效阻断了疫情的传播。

### 智能客服助力产业服务智能化转型

自然语言处理通过实现人机间自然语言通信，致力于让机器在不断学习中模仿人的思考，能明白人类语言表达的内容，并根据所理解的问题进行回答。在服务业中，自然语言处理往往是对话机器人的主体技术，使机器人能以对话方式提供在线咨询服务，打通人工坐席和工单系统，广泛用于客户服务场景，企业不仅可以大范围降低人工成本，还大大地提升服务客户的效率。

腾讯云智能客服系列产品基于最前沿的自然语义处理、深度学习算法以及语音识别/合成等技术，为企业级客户提供的一整套智能客服机器人产品，包括智能机器人（文本客服、智能IVR等）以及智能质检、培训等。企业外部客户可以通过微信公众号、小程序、网页、APP、H5、电话、实体机器人等渠道进行咨询，企业内部人员也可以进行业务咨询，场景上具备广泛应用。客服机器人通过与人进行智能问答互动，提升智能客服体验；通过AI语义检索引擎搜索匹配答案，准确解答用户问题；面对无法解答的问题，机器人根据语义判断，可随时快捷触发人工服务，提升机器人至人工的流转效率，升级客户体验。全天候、全通路的秒级响应服务，让客户不再排队等待，最高可节约人力成本68%。微众银行的智能客服平台服务用户总数达到百万级别，让微粒贷能7X24小时低成本服务用户，在用户高速增长的同时依旧保持用户体验，在效率和成本之间取得平衡。

## 语音助手帮助提升服务质量

基于语音识别、语音合成、声纹识别等技术，AI语音助手可以理解并执行用户的语音意图，通过语音交互获取娱乐类内容、控制智能设备，广泛应用于大屏操控、驾驶、家庭娱乐、景区旅游场景，在垂直行业领域，腾讯云小微AI助手，整合了语音交互、自然语言理解、声纹识别、人声合成等AI能力，同时连接了腾讯丰富的内容和服务生态，致力于打造拟人化、情感化、个性化的AI助手。目前，腾讯云小微AI助手已与腾讯车联、格力空调、游云南、粤税通等行业客户合作，通过语音助手带来更高的服务质量。

在客服质量监测方面，语音识别将电话录音识别为文本，再通过质检规则对文本进行分析，识别不符合规范的服务通话记录，完成人力不可能完成的超大规模电话录音质检工作，提升呼叫中心人员的服务质量。面对珍爱网业务的快速增长，智能客服平台提供了优质的服务，低成本、快速的完成3000线下红娘的电话服务质量提升和质检。通过腾讯云语音识别完成电话录音质检，珍爱网电话质量检测效率提升31倍，成本降低90%，原来需要50个人的工作量，现在只需要1个人+1台机器就可以完成，大幅提升呼叫中心工作质量管控能力。

## 多维交互更生动，虚拟人打造服务新模式

虚拟人是指具备虚拟形象的多模态交互，是文字、语音、对话AI技术融合的综合体，在传统交互基础上提供更强互动性和表现力，适合企业服务、旅游、教育等垂直行业的“窗口服务场景”如：虚拟客服、虚拟导游、AI助教等。

腾讯虚拟人运用了腾讯云小微的语音对话、腾讯AI Lab的2D虚拟人生成技术，赋予传统AI语音助手形象、情绪、多模态交互的能力，赋予IP角色“能听会说”的能力。通过虚拟人技术打造出的智能虚拟人，能支持虚拟形象离线播报视频和实时语音交互两大类使用场景，可助力企业降低成本、打造差异化产品，帮助媒体、教育、会展、客服等行业进行智能化升级。

腾讯虚拟人为全国首个省级文旅IP形象“云南云”进行虚拟人能力赋能，设计了“云南云”与景区导览、直播、线上虚拟助理、线下机器人客服相结合的应用场景，将“云南云”可爱有趣的设计理念通过动作、声音、语言，融入到游云南为游客提供的主要服务场景中，提升了用户使用体验、节省了客户服务成本、提升了云南云IP的实际服务价值。

## 案例：云智优保推动保险行业快速复工

云智优保是腾讯专门针对保险行业搭建的基于云智天枢人工智能服务平台和优图AI能力实现的人工智能保险服务平台，专注于优化传统保险业务流程中核保和理赔环节人工成本高处理效率低的业务痛点。减少保险公司核保人才的依赖，保障承保流程中录入、复核、核保环节的处理效率和处理结果稳定性，助力保险行业快速复工复产的同时，降低运营成本。同时云智优保基于AI大数据分析，能够建立个人保险信息档案，助力保险企业将存量数据价值充分挖掘，不仅可以赋能销售业务，还可以形成对外输出云数据的能力。

泰康人寿便使用了AI技术进行核保，腾讯提供“OCR+NLP+大数据分析+机器学习”四大AI基础能力，为其提供了完整的“智能核保”解决方案，其检测项结构化准确率达85%，影像检查描述和结论结构化准确率80%，核保效能提升60%，核保结论预测准确率90%。

## 工业智能升级方向

疫情发生以来，复工复产的急切需求让原本数字化较强的企业获得实实在在的优势，数字化技术正受到越来越多的认可，而这份认可使市场盼望着工业智能的快速增长。纵览制造业和服务业里的智能技术运用，我们可以发现，随着企业自动化程度提升，人机结合程度更高，人与机器的关系得到进一步的升级。从未来来看，机器与劳动力的关系将是全方位的，从低端劳动力替代，到高端劳动力结合，解决人力从数量上到质量上的依赖。从短期来看，无人化生产在疫情期间疫情爆发期间劳动力受限的情况下发挥了极大的优势；从长期来看，伴随着不断增长的退休员工人群和不断缩减的新鲜血液的人力补给，无人化生产模式不光是节省了长期的人工成本，更是减少人力数量的依赖。而对于传统工人技能与数字化转型浪潮后对人才的需求间的差距所导致的结构性失业问题，智能生产模式通过挖掘和沉淀相关知识，可以减少对于人力自身素质的要求，减少对于技术专家等特殊人才的依赖，让原有的劳动力也能完成相同的工作，从而降低对人力质量上的依赖。

## 「3.4 AI+数字内容：多点开花，赋能内容行业生产力」

疫情期间，“宅生活”开启，内容消费得以高速增长。随着内容输入渠道的日益丰富，人们对于内容的要求也水涨船高，只有更加有创意、更加个性化的内容才会对内容消费者产生持续不断的吸引力，而技术成为达到这一目标的推动力。互联网尤其是移动互联网的普及让数字内容人工智能技术已经作用于影视、社交、娱乐、游戏等诸多内容消费领域，基于深度学习的精准推荐、协同过滤等等技术正在为全球创业产业集群发展贡献力量。

### 影视：升级后期制作，扩展角色创作空间

影视作品是内容表达的经典形式，是艺术创意的集合，而技术更是为创意插上了新的翅膀，让艺术超出原有局限，变得更具有想象力。新时代的艺术注定是与技术相结合的成果。

AI在影视作品里，一方面可用来升级音视频剪辑技术，为影视制作中的特效、配音呈现更好的效果，减轻视频编辑人员的工作压力；另一方面还可以减少因为演员、拍摄场景的局限，拓展电影的创作空间，衍生出更多改编作品。例如深度合成技术可以实现逼真的“替身演出”，在《星球大战8》、《速度与激情7》等影视作品中“数字化复活”已故演员，还可以自动执行各种语言的逼真配音。此外，借助深度合成技术，影视制作方也可以针对既有内容拍摄续集、进行衍生和适应不同文化的改编等。

同样，AI也可以运用于艺术创作方面，将深度合成技术用于创造新形式的戏仿与讽刺具有很强的冲击力，包括但不限于合成明显与演讲者的语言、身份不协调的作品，讽刺、模仿、批评公众人物和社会事件，这是以往的作品无法做到的。

### 短视频：多个环节解决痛点，释放行业生产力

步入2020年，短视频行业虽然进入存量经营时期，但无疑仍将伫立于高光点上。人工智能技术等前沿技术

的应用探索为短视频提供了发展的基础。短视频在生产、分发、审核环节的智能化水平会进一步增强，有望更大程度上解决“高效剪辑”、“精准分发”、“智能把关”等短视频行业痛点。

生产方面，AI助力短视频的自动生成，大规模专业化生产时代即将来临。语音识别、图像分析、人脸识别等人工智能技术，解决传统视频剪辑中对素材初筛、整理、成片的工序繁琐问题，对素材的标签化、结构化可以缩减后期剪辑的制作时间，提高剪辑效率，同时这一自动化过程降低了视频制作的技术门槛。其基本方式有：独立根据原始素材创作短视频、从长视频中选取片段合成短视频、自动配乐、配字幕等应用。未来，视频网站、音像制作企业等基于其后台海量、结构化的素材数据库，不仅能够实现生成短视频的自动化和规模化，而且通过开放版权内容素材库和自动化创作能力，可为上游生产端提供大量数字化资源。腾讯新闻的青云智能系统，通过cross图文转短视频工具，自动做文本缩略、语音生成和视频素材生成，3分钟内合成短视频，其日均效率可达5000-7000条。短视频智能工具普及化、简单化、易用化将成为未来趋势，这会释放中长尾业务创作者的创作效能，破解创作效率和质量的瓶颈，改变现在短视频头部效应明显的生态格局，迎来大规模专业化时代。

分发方面，通过AI匹配用户和内容，做到千人千面精准推荐。个性化推荐算法可分为多种维度：用户个人数据，刻画用户画像；用户行为数据，刻画用户兴趣图谱；内容结构化数据，匹配视频内容的相似度等。在短视频领域，还将用户观看视频的完成度与“平均观看时长”的比较、是否倍速播放、是否有托拉拽、是否多次观看等指标综合衡量用户对视频的喜爱程度。AI个性化推荐算法一方面能挖掘冷门内容，避免两极分化的马太效应，为创作者找到属于其自身的长尾用户。另一方面，帮助视频平台对内容进行分类，从而利于平台针对不同用户进行精细化管理和运营，实现去中心化的内容推荐，提升分发效率。

审核方面，智能审核大大加速速度和精度，打破人工审核掣肘。人工智能审核的优势在于学习速度快、效率高、存储量大、不会疲倦。“智能在前，人工在后”模式是当前AI审核短视频的通用方式。视频平台通过人工智能模型识别视频画面，检视视频作品是否违规；在人工审核环节，审核师抽取视频中画面、关键帧，清查变相包含敏感内容的视频。但是，未来随着视频内容的爆发，人工审核仍面临巨大压力。而且，主观判断对视频审核是必要的，因为对视频内容的判断难以通过唯一的标准来划定是否存在违规，遑论通过机器就能过滤掉所有违规视频。

计算机视觉、大规模云计算、自然语言处理、深度学习等人工智能技术的突飞猛进，使得短视频与人工智能技术的结合成为短视频行业新的“天空城”，在文娱、医疗、新零售等领域开启智能化时代。

### 社交分享：人像合成带来社交新花样，激发用户好奇心

社交往往需要新奇内容推动，近年来利用人脸替换、人脸再现等AI技术，许多图像、视频合成应用在国内外市场上反响热烈，不仅可以实现图像的风格迁移，而且可以通过人脸融合、图像融合合成新的图片，还可以让用户“虚拟地参与”电影演出。例如FaceApp推出的老年滤镜、婴儿滤镜，会自动添加皱纹、白发和皮肤松弛等衰老迹象，也能基于现有面部特征合成幼化的面孔，同时保证照片的真实感。ZAO则主打视频换脸，利用特定的影视化素材进行表情迁移和头部姿势迁移，让用户与明星同框，体会“表演”的乐趣并能够展示自我。

越来越多影视作品、游戏正在使用人像合成能力，让用户与喜爱的角色融为一体，更有参与感。腾讯优图实验室也已将深度合成技术应用于社交、游戏、图像和视频制作等领域，背后所依托的是人像融合技术DittoGAN：通过自研深度解耦学习框架实现对人脸身份特征和属性特征的精细化建模；伪模板图像生成技术制作带真实标签的训练数据对，实现有监督学习，进一步提升图像的恢复质量；真人像超分辨率框架实现高清人像编辑与生成。



这一技术还能用于H5等活动营销传播、游戏影视制作、以及各类相机APP和小程序，带来更多富有创造性的体验。这类应用玩法简单、体验流畅，还能将换脸后的照片或视频分享到交友平台，满足了用户的好奇心与新鲜感，也满足了用户对于私人定制、彰显个性的心理需求。

### 游戏：从设计到玩法，AI颠覆游戏模式

游戏，是真实世界的模拟与仿真，一些AI算法直接用于现实可能会出现成本消耗过大、难以适应真实场景的情

况，而游戏由于其本身的数字性质，数据较实体产业来说更易采集，模型锻炼效率更快，成为了理想的虚拟测试场。可以说，游戏一直是检验和提升AI能力的试金石。AI技术在游戏中得到不断磨练的同时，又反哺游戏产业，开辟了游戏产业的新天地。

对游戏设计环节来说，AI颠覆了传统游戏策划模式，参与游戏辅助设计，通过无限次的模拟学习，AI在进行角色参数调整、提高测试效率、优化角色平衡性等方面达到前所未有的高效和智能，最大限度地保证游戏设计的公平合理，这一设计方式的变革将改变游戏从业者的行模式。

从游戏体验来说，人工智能玩家的出现吸引了人类玩家极大的好奇心和征服欲，从围棋界的阿法狗开始，挑战AI成为玩家的执念。腾讯旗下手游“王者荣耀”在今年五月推出了名为“绝悟”的策略协作型AI，所有玩家在限时期间都可以对它发起挑战，这一新兴玩法引起了玩家群体的极大热情。绝悟的1v1版本曾在2019年开放，在与顶级业余玩家的2100多场体验测试中胜率为99.8%，今年公开的绝悟5v5版本，最难模式代表着2019年KPL冠军队伍的战斗力水平。在腾讯AI团队对“绝悟”训练的过程中，难点在于要在不完全信息、高度复杂度的情况下作出复杂快速的决策。在庞大且信息不完备的地图上，10位参与者要在策略规划、英雄选择、技能应用、路径探索及团队协作上面临大量、不间断、即时的选择，这带来了极为复杂的局面，预计有高达10的20000次方种操作可能性。

除此之外，游戏AI也将发力电子竞技领域。作为数字时代最受年轻人欢迎的运动，电竞已于2018年成为亚运会表演项目，中国队参赛获两金一银的佳绩。与传统体育项目一样，电竞职业选手也需要手眼脑协调、策略和操作快速反应、团队协作精神及大量刻苦训练。借助在算法和数据方面的优势，“绝悟”一类的游戏AI可为职业选手提供数据、战略与协作类实时分析与建议，及不同强度与级别的专业陪练，继续推动中国电竞在全球范

围内保持领先。

AI+游戏的探索，是希望通过游戏的虚拟世界环境助力AI的研究，始于游戏，但不限于游戏。如果在模拟真实世界的虚拟游戏中，AI学会跟人一样长期策略规划和协作能力，就能执行更困难复杂的任务并发挥更大作用，在挑战性和应用性方面存在巨大价值。游戏AI将是探索通用人工智能的关键一步，不断让AI从0到1去学习进化，并发展出一套合理的行为模式，总结中间的经验、方法与结论，长期会对实体经济带来更深远影响。

## 内容 AI 升级方向

从局部到整体，全身合成将成为新热点：在AI换脸一度引发体验热潮的同时，合成媒体的分支——针对全身的“深度合成”也悄然而至。全身合成需要的关键识别要素将更多更复杂，高质量、高仿真的深度合成内容仍然难以创建，需要掌握专业技能和专业工具的专业人员的大量投入，让合成技术的应用场景更为广泛。

从静态到动态，用AI探索视频商业价值：由于技术的限制，现在的内容研究多落地于静态的图片，近几年来大热的短视频领域虽有零散的应用出现，但还没看见大规模成熟的AI落地。未来AI技术在视频理解、视频合成方面，将会有更大的市场。例如，2018年出现的一种新的AI算法，可用于学习源人物的舞蹈动作并映射到目标人物之上，让一个不会跳舞的人也能呈现出近乎专业的舞蹈动作。

在单一的音频、图像合成之外，运用于内容领域的AI技术将朝着综合性的多模态交互方向发展。单模块的语音、图像合成已有不错的成绩，如何将不同模块相整合，在一个产品上同时具备多种AI能力成为新问题，未来研究会进一步集中于让多模块形态运行更为和谐。

AI对于内容产业的总体意义在于其改变了内容的生产方式，基于新的生产方式，以创意至上的内容产业的生产潜力得到进一步激发，注定会迸发出前所未有的生机。

## 3.5 小结

疫情之下，世界见证了数字技术的力量，人工智能在疫情中发挥的积极作用让世界重新审视人工智能的价值。除了上述领域之外，在金融、教育领域也迎来大规模落地，解决产业中面临的问题。产业互联网成为发展的共识。经过这次疫情，中国已经不再有纯粹的传统产业，每个产业或多或少都开启了数字化、网络化、智能化进程，AI赋能传统产业转型升级将迎来历史性机遇。人工智能已“泛”起巨浪，而人人都将席卷在这巨浪之中。



## — 4 —

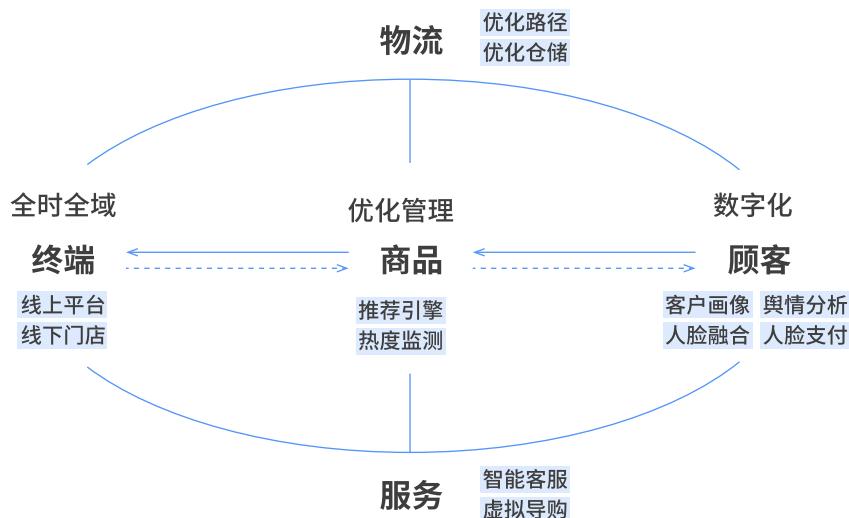
# 创新经济篇

危机总是与契机相伴而生，智慧抗疫完成优秀答卷，成为社会有序运转的强大后盾：在疫情无接触的环境下，短期内被引爆的在线经济，是对原有在线模式的深化及应用范围的增加，活跃的数字经济成为了经济增长的新支撑。制造业和服务业人机结合的进程加速，人工智能正在提升产业的智能化运营能力，甚至成为一种新的生产力，参与到实体经济和数字经济的生产环节。实体经济在智能技术的助力下变得更具韧性。

在技术和应用的叠加发展过程中，人工智能与经济社会各行业各领域融合创新水平不断提升。也在不断催生出新模式、新场景、新交互的出现。智能数字体形成了新的生产力。“泛在智能”，泛于大众，惠于大众。长期来看，发展人工智能的终极目标并不是取代人类，而是通过将人类从繁重的重复工作中解救出来，实现对人类整体更有价值的目标。以更普惠、更负责任的发展为目标，新商业和新经济将是技术发展路上附赠的礼物。

## 「4.1 新模式：AI 助力零售向无接触经济转型」

新冠肺炎疫情造成全球生产网络停摆、中断和破裂。在疫情中，仍然保持增长的新业态有一个普遍特点——“社会距离”，以“非接触”为核心的服务模式已经成为了新的经济增长点。无人零售、无人值守、物流运输等行业的非接触式服务需求大涨，迫切需要AI技术进行业务支撑。“零售+AI”正在从单点走向聚合，通过智能物资调配、智能营销推荐、智能配送实现全产业链“非接触”配置。



作为零售商业里的三个核心主体，顾客、商品、终端之间的关系也随着时代的发展而产生着变化。十年前的商业逻辑是终端-商品-顾客变成了顾客-商品-终端，顾客的需求放在了首位，商品和服务要贴近用户并满足用户需求，场也从固定场所变成线上线下相融合的消费场景。疫情期间的隔离使正常购物成为难题，线下零售受到客流骤减的冲击，这催生了新的消费习惯。新消费习惯融入人们的“宅生活”，加速消费场景的转换，促使无人经济迅速落地，非接触业态蓬勃发展，基于数字化技术的智慧零售企业获得快速增长，而人工智能技术助力零售业企业快速适应这一变化。

### 终端：打通线上线下，实现全时全域的消费场所

在消费升级的大背景下，消费者在更加追求商品品质化、个性化、多元化的同时，也更加注重购物过程的体验。与纯线上消费相比，线下实体店由于服务和体验的独特优势，在智慧零售的趋势下重新受到市场的重视。另一方面，虽然线上市场的红利逐渐衰退，电商获客成本逐步提升，但之前积累的用户数据却是有竞争力源泉。因此，对于零售商而言，通过线上线下渠道的整合，将二者

的优势相结合实现最大化，是未来发展的最佳路径。例如，深度合成技术可以提升线上购物的体验，品牌可以使用体型、身高和肤色各异的虚拟模特展示服装效果，而不用再雇佣摄影师和专业模特，甚至可以鼓励消费者自己进行数字建模，预览服装的上身效果，实现快速的数字试衣、试穿，进一步增强网购的个性化体验。

## 商品：线上预测，线下优化，商品管理愈发灵活

在商品管理方面，如今由于顾客需求日趋多样化并且期望送达时间越来越短，零售企业的库存水平和生产压力形成矛盾。基于AI技术构造门店商圈客流画像和偏好合理调配门店商品，根据顾客购物动线和停留监测，优化商品陈列，提高坪效，基于深度学习对门店和商品销量的做出预测，从而合理筹备库存和分配库存。

## 顾客：线上分析，线下追踪，比用户更懂用户

由于对于客户个性的不充分了解，零售商往往不知道谁是理想顾客、顾客想看什么、想买什么，受到不可估计的隐形损失。线上端，通过各种渠道进行用户触达，再利用AI技术来完成商圈洞察、社区洞察、人群画像、品牌舆情等用户分析，形成线上分析矩阵。腾讯珠玑这一智能AI写作工具，能够有效帮助商家运营好公众号这个触点，辅助写手或者完全自动化地在短时间内产生丰富的精彩内容，同时通过对点击和转化的精确预估，将内容精准的推送给客户，经测试，推送消息的阅读率和点击率都有20%以上的提升，购买转化率的提升更是高达三倍以上。线下端，AI通过把握客户从进店到支付离开的顾客购物全旅程，对顾客进行实时分析、锁定并抓取目标顾客、分析目标顾客潜在需求，真正做到构建每一位消费者360度全方位数字化画像；还可以引入人脸试妆、智能试衣等新鲜玩法和快捷的人脸支付等人工智能技术设备来升级购物体验。

## 物流：预先规划，快速精准送达

供应链是使人、货、场全过程顺利流动的重要环节，在即时物流情景下，短时内产生的包裹数量越来越多，并且由于客户对时间要求越来越精准，快递员已经无法依靠经验进行配送。通过深度学习等技术，AI建立高效的供应链系统，形成基于消费者、门店销售、客户一体化的供应链智能管理体系，把用户端潜在需求的判断联动到供应链、物流仓储系统，提升经营效率，降低库存和供应链成本，实现供应链管理的智能化。

## 服务：降本增效，智能客服更贴心

零售业的客服工作是一项重复性强且工作量大的工作，又是贯穿整个消费流程、影响顾客购买率和满意度的重要要素。人工智能对于这类工作的适用性极高，在购物过程中开展智能客服，使服务从售前到售后的各个环节，无限量地接待用户，成为提升服务效率、降低企业成本的不二选择。

总的来看，从渠道来看，智慧零售将从纯线上或纯线下模式发展到现在的线上与线下的结合；从环节来看，智慧零售将会串起一个个的单点智能，形成全流程全环节的智慧场景转型。新的消费模式发展需要科技创新支持，人工智能技术使打通线上线下、连接单点智能场景成为可能，用消费新体验来释放新兴消费的潜力，让实物消费和服务消费得到互补，成为消费市场的新增长点。

## 4.2 新场景：加快自动驾驶的商业化进程

2019年12月以来，新冠肺炎疫情在全球的蔓延给人们的健康和生活生产带来了极大影响，为应对疫情，宅经济、无接触服务等模式成为一段时间以来的选择，在此背景下，自动驾驶汽车的应用开始崭露头角，助力抗疫和生活生产的恢复。在国内，一众厂商开始试水自动驾驶应用，自动驾驶汽车在外卖和物资配送、物流、城市消杀作业等方面开始发挥作用。在国外，通用汽车旗下Cruise、谷歌旗下Waymo、小马智行等自动驾驶厂商开始将自动驾驶汽车应用于生鲜食品、外卖、医疗物品的配送。美国加州作为全球自动驾驶发展高地，在道

路测试和试点运营等方面积累的经验，为疫情期间探索自动驾驶应用场景提供了有力的支撑。

疫情期间自动驾驶应用崭露头角的背后，是自动驾驶商业化加速，开始催生新的无人经济。而立法的允许是必不可少的前提，例如，美国加州已于2018年允许自动驾驶汽车客运服务试点项目；2019年6月，美国佛罗里达州制定的自动驾驶立法，不仅将在道路上操作自动驾驶汽车合法化，而且提出了“自动驾驶汽车共享网络”这一全新的客运交通服务；我国的自动驾驶汽车发展也正逐步进入示范应用阶段。正因如此，自动驾驶出

租汽车（Robo-Taxi）、自动驾驶卡车和物流车、固定路线运营、最后一公里配送等商业化场景开始在国内外落地，无人经济的概念获得了新的诠释。

截至目前，腾讯通过自研了车路协同平台，扮演“生态连接器”角色，建立起应用和网络之间的连接器，通过释放5G网络潜能，连接人、车、路、云，从而搭建一系列核心能力体系，将整个产业串联起来。在视觉感知领域，腾讯有专门的研发团队正在研究包括图像识别、多传感器融合、融合决策以及改装车辆在复杂环境中测试等多项关键能力；在车与车、车与路的通信之间，腾讯发力5G标准，建立了未来网络实验室，主要研究5G车路协同落地时会面临的共性问题；在落地方面，腾讯正在考虑加强边缘计算设备的部署，目前正在与多家合作伙伴联合建设边缘计算开源平台。

### **仿真平台成自动驾驶技术发展迭代的关键，国内外厂商纷纷布局**

自动驾驶技术的发展应用离不开测试，需要通过大量测试来验证技术的安全性能和成熟度。但就目前而言，一方面监管政策对测试牌照、测试道路等有着严格的限制，另一方面实际道路测试还面临着测试成本问题，这些都不利于自动驾驶技术快速发展迭代。由于实际道路测试的诸多限制，仿真技术和仿真测试就成为了推动自动驾驶行业发展的新引擎。目前，国内外自动驾驶厂商都在加大布局仿真平台建设。

以自动驾驶的领头羊谷歌旗下Waymo为例，其很早就重视仿真技术和仿真测试，其在公开道路上的测试里程超过2000万英里，但仿真测试里程早已超过100亿英里，而且仿真测试对其自动驾驶技术的迭代和成熟发挥了重大作用。国内政策和行业也非常重视发展自动驾驶仿真技术，《深圳市关于支持智能网联汽车发展的若干措施》明确支持城市级场景仿真；腾讯结合自身的AI技术，及优势的游戏技术积累，构建起与真实世界一致的数字孪生世界，保证无限接近真实的逻辑规律，以实现仿真测试结果的真实性。腾讯TAD Sim在场景构建中，采用了三维重建和游戏引擎等技术，以小于3cm的精度误差模拟道路交通场景及虚拟城市场景，并借助游戏引擎还原出日出、日落等光照条件，以及风、霜、雨、雪等天气条件的变化，使得场景和传感器仿真的测试条件

接近真实，进而保证测试效果的真实性。在仿真世界中，交通环境车辆就好像游戏中的NPC，通过Agent AI技术，用路采数据训练交通流AI，辅以三维重建技术和虚实一体交通流，生成真实度高、交互性强的交通场景，实现自动驾驶感知、决策、控制等全部模块的闭环仿真验证。仿真平台涵盖道路场景、交通流、车载传感和车辆动力学等仿真模型，可满足研发时的精细调试和测试时批量运行的不同需求，可轻松模拟驾驶日行百万公路，相比于实际道路测试，将更能加速技术发展迭代。

### **“新基建”加持，自动驾驶行业将迎来发展黄金期**

2020年以来，国家密集部署“新基建”政策，新基建迎来风口。根据国家发改委对新基建的范围界定，新基建中的5G、人工智能、云计算、数据中心、智能计算中心等信息基础设施，以及智能交通基础设施和自动驾驶密切相关。5G将加速车联网发展，与智能交通基础设施配合，实现车路协同；人工智能、云计算是自动驾驶系统的核心，帮助自动驾驶系统实现感知、行为预测和规划，从而可以执行驾驶任务。此外，自动驾驶系统研发落地也离不开数据的支持，尤其需要大量的标记数据；在这方面，腾讯自动驾驶三大产品之一的“数据云平台”，基于云端海量存储空间与计算资源支撑，构建了数据采集、样本标注、算法训练评测、云端仿真、实车反馈闭环全流程云服务，提供自动驾驶全链路云服务和开发平台。目前已积累海量丰富样本数据，包括全要素、车道线、跨相机目标跟踪、图像分割等图像标注样本，3D激光点云标注样本，以及精准图像与点云联合标注样本。可以预见，在新基建的助攻下，未来几年自动驾驶将迎来发展黄金期，有望实现国家智能汽车战略中所要求的到2025年实现市场化应用。

自动驾驶作为综合运用各种人工智能技术的产物，其发展落地能够改善交通出行、减少事故并可能重塑我们的经济和社会。但在这一愿景达成之前，我国仍需要在安全标准、道路测试、商业试点、高精地图、责任保险等多方面配套推进政策措施，借鉴国外采取技术友好型的监管路径，并积极废除或修改阻碍自动驾驶汽车发展应用的既有法规和标准，才能确保在制造业与ICT融合的自动驾驶汽车领域抢占全球高地。

## 「4.3 新融合：深度合成为数字内容注入新动能」

随着人工智能、5G、虚拟现实等新技术融合发展，新一代媒体将由人工智能驱动，人工智能可能给数字内容领域带来重塑。其中，可以实现换脸、人脸合成、语音合成、视频生成、数字虚拟人等诸多应用形式的“深度合成”（deep synthesis）技术，作为人工智能技术发展到一定阶段的产物，逐渐从deepfake、deepnude等色情性换脸视频的阴影中走了出来，迎来商业化时代。

“深度合成”作为一种AI合成内容（AI-generated media）技术，最早引起关注是2017年美国Reddit网站上一个名为“deepfakes”的用户上传了经过数字化篡改的色情视频，即色情视频中成人演员的脸被替换成电影明星的脸。此后，媒体开始使用deepfake（“深度伪造”）来泛指这类新型的合成内容及其背后的AI技术。但实际上，deepfake这一用语并未得到技术社区的广泛认可，也未被正式纳入词典之中。“深度伪造”一词最早只用于描述实现了AI换脸的色情性视频，虽然“深度伪造”的出现让背后的AI技术获得了广泛的关注，但基于技术使用的意图（即deepfake）去定义技术，强调技术的潜在欺骗性或可能带来的负面影响，这一做法并不公正。因此，使用“深度合成”（deepfake synthesis）这一术语来描述可以实现换脸、脸部表情修改、人脸和语音合成等应用形式的AI合成内容技术，更加科学合理。“深度合成”依赖于人工智能技术，尤其是可以从大量数据中自主学习的深度学习算法模型，生成对抗网络（GAN）的出现则大大增强了“深度合成”的仿真能力。就“深度合成”的实现过程而言，一般包括数据提取、数据训练和数据转换三个步骤。“深度合成”目前主要包括四种典型方式，即人脸替换、人脸再现、人脸合成以及语音合成。

随着“深度合成”技术不断发展，新的发展趋势开始呈现：

一是深度合成技术朝着综合性的方向发展。这意味着，“深度合成”可以综合运用人脸替换、人脸再现、人脸合成、语音合成等技术，实现更加复杂的视频合成。二是面部合成之外，全身合成成为新的热点。例如，2018年出现的一种新的AI算法，可用于学习源人物的舞蹈动作并映射到目标人物之上，让一个不会跳舞的人也能呈现出近乎专业的舞蹈动作。三是2D合成之外，

3D合成尤其是数字虚拟人将是下一阶段的重点，与虚拟现实的融合发展值得期待。目前，融合多种AI能力的数字虚拟人技术已成为人工智能领域的前沿阵地，在虚拟歌手、虚拟主播、虚拟教师、新闻主持和智能助手等诸多领域的应用潜力巨大，国内外科技企业纷纷加紧布局，意在攻克这一领域。例如，腾讯AI实验室从2018年起开始攻坚数字人技术，将计算机视觉、语音/歌声合成和转换、图像/视频合成和迁移、自然语言理解等多模态AI能力融合，生成可交互内容，并打造出拟人度较高的智能数字人。

### 泛娱乐和数字内容成“深度合成”技术应用的主场

随着“深度合成”技术的不断成熟和技术门槛的持续降低，国内外主流的科技公司和AI初创公司争相推出商业化的“深度合成”应用。可以预见，未来几年“深度合成”技术将持续迎来商业化时代，创新性的应用形式将持续涌现。就目前而言，泛娱乐和数字内容已然成为了“深度合成”技术应用的主场。

一是为影视创作与制作开拓新空间。例如，“深度合成”技术可以实现逼真的“替身演出”，在《星球大战8》、《速度与激情7》等影视作品中“数字化复活”已故演员，还可以自动执行各种语言的逼真配音。二是带来全新的娱乐方式。近几年来，FaceAPP、Snapchat、Face2Face、ZAO等图像、视频合成应用在国内外市场上反响热烈，不仅可以实现图像的风格迁移，而且可以通过人脸融合、图像融合合成新的图片，还可以让用户“虚拟地参与”电影演出。腾讯优图实验室也已将“深度合成”技术应用于社交、游戏、图像和视频制作等领域。三是升级电商体验和内容营销。例如，深度合成技术可以提升线上购物的体验，品牌可以使用体型、身高和肤色各异的虚拟模特展示服装效果，而不用再雇佣摄影师和专业模特，甚至可以鼓励消费者自己进行数字建模，预览服装的上身效果，实现快速的数字试衣、试穿，进一步增强网购的个性化体验。此外，在广告宣传、内容营销等领域，AI合成的人脸和虚拟形象可以替代网红、模特等。四是数字虚拟人受到市场青睐，新的应用不断涌现。数字虚拟人融合了人脸合

成、语音合成、自然语言处理等多种AI能力，目前在虚拟歌手、虚拟主播、新闻主播、影视创作等领域崭露头角。例如，虚拟歌手“洛天依”大获成功之后，业内开始大力开发游戏直播、演唱等领域的AI虚拟形象；新华社从2018年起就推出了AI合成主播，今年两会期间则推出了全球首个3D版AI合成主播。美国数字王国（Digital Domain）公司更是借助数字人技术“数字复活”了马丁·路德·金，制作了沉浸式体验展《大游行》（The March）。未来数字人技术可能融入虚拟现实应用，带来更真实的虚拟互动。

### “深度合成”技术助力科技向善， 监管需要包容审慎

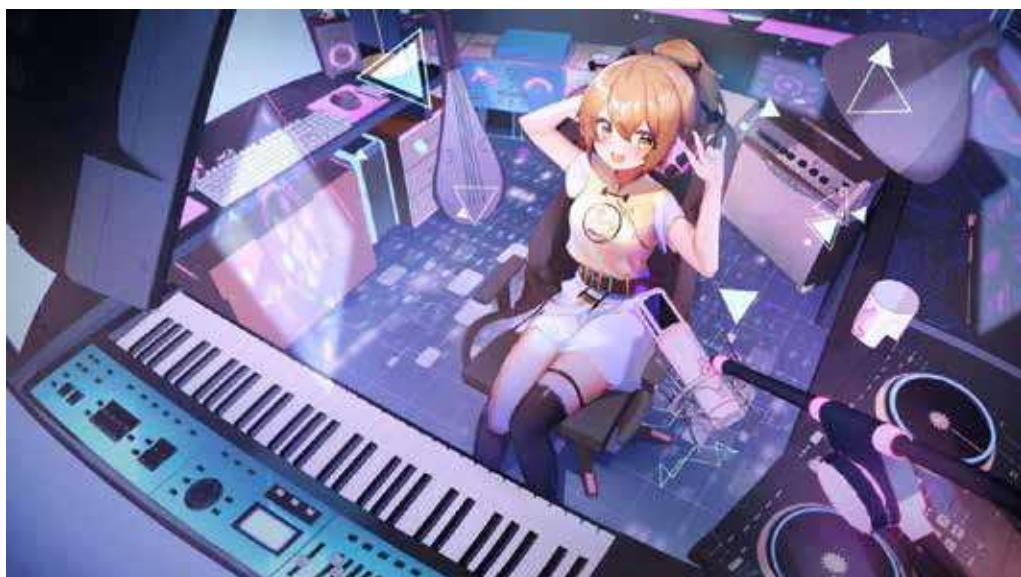
当然，“深度合成”技术也能助力科技向善。例如，语音合成软件制造商Lyrebird为渐冻症患者设计一套新的语音合成系统，用患者自身的语音数据替换以往机器

合成的语音，这意味着有失声风险的患者，可以通过提前采集自己的语音，获得能继续用“自己的声音”交流的机会。此外，该技术还可以用于帮助老年痴呆症患者与他们可能记得的年轻面孔互动，或数字化地再现截肢者的肢体。

深度合成技术在内容创意、营销、社交、娱乐、电商、通讯等诸多领域应用前景广阔，其未来应用令人期待。这也说明，深度合成并非关于“伪造”和“欺骗”的技术，而是极富创造力和突破性的技术，虽然它像其他任何技术一样，也催生了一系列必须面对的难题，但这并不会磨灭这一技术给社会带来的进步。总之，深度合成技术不会让社会真相失守，更不是世界秩序的威胁者。面对新技术的挑战，政府和监管者应当包容审慎，避免阻碍深度合成技术的有益的、创新性的应用，通过法律、技术、行业、用户的多重治理将其纳入可控的发展轨道，逐渐从deepfake、deepnude等色情性换脸视频的阴影中走了出来，迎来商业化时代。

## 「4.4 新交互：互动感加倍，多模态虚拟人融入更多场景」

面对万亿的粉丝经济蛋糕，娱乐偶像层出不穷，各路选秀节目争夺粉丝市场。然而，近年来由于丑闻曝光导致的偶像“形象崩塌”的翻车事件频频发生，这让为他们真金白银氪金的粉丝群体十分受伤。虚拟歌手“洛天依”大获成功之后，业内开始大力开发游戏直播、演唱等领域的AI虚拟形象，仿佛让这一类实力颜值永远在线、不会辜负粉丝心意并能时刻陪伴的虚拟偶像成为粉丝经济的下一站。腾讯AI Lab近来推出虚拟歌姬AI艾灵，通过基于数据依赖型的深度学习方法，不仅能作词，还能用近乎真人的声线演唱，加上用多模态智能技术搭建的数字躯体，成为唱跳原创能力兼具的全能偶像。



随着近些年机器学习技术的发展，合成语音在拟真度方面已经取得了长足的进步，机器声再也不只是人类印象中那种一板一眼的刻板模样了，现在的AI甚至还能学会像人类一样情感充沛地歌唱。不同于初音未来等虚拟歌手的“机器合成+人工调教”模式，使用了基于DURLAN声学模型的AI艾灵无须经过人工调教就能得到非常自然和拟真的歌声，甚至实现跨语种歌声合成。在与粉丝的互动中，只要用户输入的歌谱，AI艾灵就可以快速地合成出对应的歌声。

自动生成歌词方面，AI艾灵通过歌词创作深度学习模型SongNet，可以给定任意格式和模板来生成相契合的文本，完成“命题作词”。用户随机输入一个或者多个关键词，AI就可以根据这份灵感，创作出优美、恰当的歌词，生成的歌词然后会被提交给歌声合成模块，再融合对应的背景音乐，一首悦耳动听的歌曲就新鲜出炉了。

例如，给定《十年》这首歌词的格式，我们可以通过SongNet重新进行配词，并保证格式不变，可以根据原来的曲谱进行演唱：

原歌词：十年之前/我不认识你/你不属于我/我们还是一样/陪在一个陌生人左右/走过渐渐熟悉的街头  
新配词：夜深人静/思念你模样/多少次孤单/想伴在你身旁/是什么让我如此幻想/为何会对你那般痴狂

虚拟人的市场并不止于虚拟偶像，在线教育、虚拟助理、虚拟导游等领域都是虚拟人未来可能涉足的应用。例如AI虚拟主播可能成为未来突发公共事件中新闻报道的新形式。新华社从2018年起就推出了AI合成主播，今年两会期间则推出了全球首个3D版AI合成主播。美国数字王国（Digital Domain）公司更是借助数字人技术“数字复活”了马丁·路德·金，制作了沉浸式体验展《大游行》（The March）。未来数字人技术可能融入虚拟现实应用，带来更真实的虚拟互动。此类技术虽然在探索期，拟人性和实际效果还有很大提升空间，但长期看，虚拟人的应用能降低个性化内容录制和制作的时间、人力和金钱成本。并且，在疫情这种特殊情境下，虚拟人的使用还可以大大降低人与人接触的频率，从而助力无人经济的发展。

不同于四四方方的AI类产品或是产品背后的AI程序，一位能歌善舞的偶像、一位亲切耐心的老师或是一位知识渊博的导游所实现的实时动态交互，才能让人真正感受到AI，彻底让人有交互的欲望，从而让AI以更美好的形式进入每个人的生活。

## 「4.5 新目标：为地球思考，AI for FEW 挑战资源难题」

联合国预计 2050 年地球人口总数将达到 98 亿，如何有效利用地球资源成为越来越迫切的问题。遗憾的是，未来并非一片光明。据联合国粮食及农业组织统计，2014年以来，全球营养不良的人数每年都在增长；世界上有近半数人生活在可用水短缺的地区；现今的主要能源也可能无法平稳维持到下个世纪。可以说，日益严重的食物、能源和水问题是当前人类整体所面临最重要的全球性难题。为了探索解决问题的方法，不仅需要全人类共同努力，也必须借助机器智能的帮助。腾讯相信人工智能能够帮助人类找到合适的解决方案，这正是「AI for FEW」（Food、Energy、Water，用于食物、能源和水的人工智能）想要实现的目标。

### 智能分析环境、作物与市场数据， 打破作物生长瓶颈

通过智能化的数据驱动，AI可以帮助提升农作物的

产量并帮助农业从业者合理规划农作物生产种植。腾讯 AI Lab 与荷兰瓦赫宁恩大学合作探索了基于AI决策在真实世界的温室中种植黄瓜的任务。这个实验中构建的农业人工智能系统能让其根据热量、通风和二氧化碳水平等因素来优化种植策略，比如温室温度、种植密度、灌溉施肥。最后的实际种植结果表明，农业人工智能系统大幅提高了资源利用率，实现17%的净利润增长，产量媲美20年经验的农业专家组。并且，经过一定调整，这样的系统也能自动适应新的温室种植环境，能够很方便地实现大规模标准化生产，从而极大地降低相关农业生产成本和人力。腾讯AI Lab还重点推进了本地化应用，与农业大省辽宁达成合作，展开两期种植试点。一期使用当地三个日光温室种植番茄，其中两个部署了智慧温室方案iGrow IoT方案的实验组，用AI模拟上万种生态环境，决策最优生长。5月结束试点后，实验组和未改造的对照组相比，每亩每季提升数千元净利润，初步验证了iGrow的商业价值。



实际上，人工智能在农业方面的应用不仅限于优化种植和生产，还可以在作物种植和收获前后的流程中发挥重要作用。例如，在种植之前，AI模型可以分析和预测市场需求，从而指导作物种植品种选择，避免产销脱节造成经济损失和农产品浪费。另外，云计算、大数据和机器学习等技术，还可以帮助筛选和改良农作物基因，达到提升口味、增强抗虫性、增加产量的目的。

## 助力预测能源需求、优化能源供应， 促进节能与环保

以电网供能问题为例，每年夏季的丰水期，四川生产的水电会有大量富余，其中的很大一部分会因为供需不平衡而被白白浪费掉。人工智能可以分析往年同期的用电情况以及用电量随气温、湿度、节假日等因素的变化，进而帮助调整电力生产策略，实现生产资源的优化配置。对于使用煤炭、生物质或其它消耗型燃料的电力生产单位而言，这样的技术尤其具有价值——不仅能帮助节省成本，还能有效地降低环境污染。

除了在电力生产端极具价值，人工智能对使用电力的“耗电大户”也极具经济效益。开发了围棋人工智能 AlphaGo 的 DeepMind 就成功应用了机器学习算法来优化能源供给，将谷歌数据中心的冷却耗电量降低了 40%。对于谷歌和腾讯这样的拥有大量数据中心的企业而言，节能不仅是为了改善成本效益，也是大型企业所应该自觉承担的社会责任。

## 水资源运输、储备流程优化等方面 有着较大应用前景

据美国水工程协会统计，2011年到15年之间，美国投入了 1.7 万亿美元来修理和扩建饮用水基础设施，其中超过一半都用在了替换给水管道上。借助人工智能技术以及监控水质和水压数据的传感器，可以分析预测最佳的管道替换时机——如果替换太早，会浪费很多材料；如果替换太晚，管道可能就损坏了，从而浪费大量水资源。此外，随着气候变化加剧，科学家预期干旱等极端天气事件会更加频繁，通过气象数据分析提前预测干旱事件并做好水资源储备等应对措施有助于将灾害损失控制在合理范围内。

整体而言，目前的AI技术真正彻底地解决食物、能源和水等现实世界问题时还有漫漫长路，但我们相信科技的发展必须用于解决地球所面临的最大挑战，人工智能会在这些地球级的挑战上发挥重要作用，并最终将世界变得更加美好。科技向善，让科技去解决真正的“大问题”，让科技承担起“救命”的角色，让科技为地球思考，为地球造福。



## — 5 —

# 制度保障篇

人工智能的发展并不平静，围绕人工智能实用性和伦理问题的争议从未停息。疫情期间，健康码再引隐私和效率的讨论，数据治理成为公众的关切。未来人工智能的健康可持续发展，技术和市场固然很重要，但同时也更需要合理的制度建设与良好的治理来保障。良好治理旨在构建多层次的新技术治理体系，实现安全可信、负责任的人工智能。现阶段的人工智能在某些领域已经超过了人类，未来AI技术可能变得更加强大。面对正在和已经超越我们生物能力的AI技术，我们既需要充分释放AI技术在社会、经济、环境等各方面的巨大价值，还要能未雨绸缪将其风险和危害控制在能接受的限度。具体而言，未来对AI技术的良好治理需要着重考虑以下三点。

## 「5.1 构建多层次的治理体系，采取敏捷灵活的治理方式」

多层次的治理体系比单一维度的立法和监管更能适应人工智能所具有的快速发展迭代、日益复杂化等特征。顶层的法律规范可以划定技术应用的边界。在基本理念上，立法需要避免统一的、一刀切式的、激进的监管方式，而应以领域细分和风险防范管理为导向，遵循包容审慎、敏捷灵活、鼓励创新等监管理念，同时考虑不同应用场景的不同影响、监管对技术和产业的影响、技术和商业可行性、企业负担等因素，采取分类分级、分阶段的方式进行适度监管。

此外，由于技术及商业模式快速发展迭代，草率的立法不可能期待会产生正面的效果，而且成文或专门的

立法恐难跟上技术步伐，故应避免过于严格和过于细致的法律要求，而是可以采取事后监管、事后追责等轻监管的方式。中层的行业和企业自律也将发挥重要作用，需要鼓励践行科技向善、负责任创新与研究等理念，支持行业自律。具体而言，可以采取行业标准、自律公约、最佳实践做法、技术指南、伦理框架等敏捷灵活的治理方式来规范、引导人工智能的发展应用，实现科技向善。底层的教育和意识培养同样不可或缺，一方面需要提升社会公众的数字素养、算法素养等技术素养；另一方面需要针对技术人员和从业者开展伦理教育，使其做到自律。

## 「5.2 立法和监管需充分考虑国际竞争、技术的经济社会价值等视角，推动先行先试」

在当前激烈的国际竞争背景下，我国针对人工智能等新技术的立法和监管需要着重考虑国际竞争视角，避免产生阻碍、延缓技术发展应用的不利效果，削弱我国的科技和产业竞争实力。因为在全球竞争重心日益转向人工智能等新技术的大背景下，过早或过度的监管都可能削弱一国在AI领域的竞争力。与此同时，立法和监管还应充分考虑技术的经济社会价值，保证技术红利在经济社会发展中的最大释放，避免因偶发性的负面问题而“过度反应”或“因噎废食”，从而采取应激、激进的监管措施。

此外，我国正在密集推动人工智能创新发展试验区建设，自动驾驶、5G、大数据等领域也在涌现发展示范区，人工智能与实体经济融合也已成国家重要布局；这些前沿科技领域的发展和应用落地往往需要突破既有的法律和监管壁垒。因此有必要在这些领域出台更包容新技术的政策和立法，探索制定前瞻性规则，移除法律和监管障碍，推动先行先试，给予适度宽松的发展空间，给AI应用提供安全港（safe harbor），通过试验、测试、试点等方式加速AI从研发到商业落地的转变，同时审查、调整政策、监管框架和评估机制以鼓励创新和竞争。

## 「5.3 推动人工智能治理的跨学科参与和国际合作」

面向未来，AI等新技术的健康发展离不开技术、社科、人文等不同背景的人员的通力协作。科技行业需要从当前的技术中心主义模式转向技术人文协作模式，AI研究与发展需要广泛吸纳不同种族、性别、文化和社会经济阶层以及不同领域（如经济、法律、哲学、社会学、人类学、心理学等等）的人员的思考和顾虑。监管

也需要采取多利益相关方共同参与的模式，广泛听取行业主体、专家和公众的意见，避免决策者与从业者脱节。此外，在当前的全球化背景下，一个可持续发展的地球也需要各国在AI技术、产业、伦理、治理等各方面加强合作，从而让人工智能实现以人为本的、包容普惠的、安全可持续的发展。

W

H

-

T

E

P

A

P

E

R