



# 兴泰智库研究报告

XINGTAITHINK-TANKRESEARCHREPORT

2023 年第 2 期总第 126 期 · 行业观察

## ChatGPT 及相关行业发展研究报告

合肥兴泰金融控股集团

金融研究所

徐蕾 李庆 邹干

咨询电话: 0551-63753813

服务邮箱: xtresearch@xtkg.com

公司网站: <http://www.xtkg.com/>

联系地址: 安徽省合肥市政务区

祁门路 1688 号兴泰金融广场 2602 室

2023 年 3 月 15 日

更多精彩敬请关注

兴泰季微信公众号



### 内容摘要:

- ◆ ChatGPT 是 AIGC 领域的重要成果, 是自然语言处理 (NLP) 技术突破的阶段性成果, 核心是引入了人工反馈强化学习 (RLHF)。具备出色的学术写作、对话能力, 智能化程度较高。
- ◆ ChatGPT 相关产业链雏形已基本呈现, 上游主要是数据、算法和算力等基础层, 为下游应用提供技术支持, 研发投入成本较高, 存在一定的进入壁垒。下游包含各种场景、形式的应用, 已有初创企业进行布局。
- ◆ 在核心算力和模型层面, 我国企业与国际领先水平仍存在一定的差距, ChatGPT 有望推动国产芯片替代进程加速; 在实际应用层面, 我国用户规模和数据量优势明显, 垂直应用场景丰富。
- ◆ 针对集团支持 ChatGPT 相关产业发展, 提出三方面建议: 一是提高对 ChatGPT 相关产业的关注度。二是把握 ChatGPT 相关链条国产化机遇。三是围绕核心企业开展供应链金融业务。

# ChatGPT 及相关行业发展研究报告

## 一、ChatGPT 概况及发展趋势

### （一）ChatGPT 基本概况

ChatGPT 是新型人工智能聊天机器人，属于人工智能生成内容（AIGC，AI-Generated Content）的文本生成类细分领域产品。ChatGPT 由 OpenAI 公司在 2022 年 11 月 30 日推出，也被称为 GPT-3.5，属 GPT（Generative Pre-Trained Transformer）系列产品。GPT 系列以 Transformer 为基础模型，使用预训练技术得到通用的文本模型。通过连接大量的语料库来训练模型，做到与真正人类几乎无异的聊天场景交流。

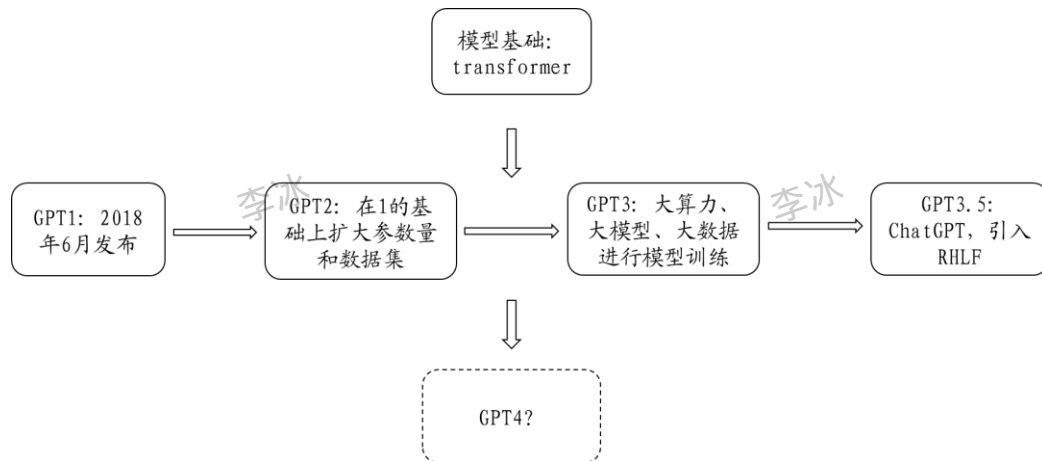


图 1 ChatGPT 技术迭代历程

1、ChatGPT 是人工智能向生成创造世界跃迁过程中的产品，是 AIGC 领域的重要成果。

传统的人工智能偏向于分析能力，通过分析一组数据，发现其中的规律和模式并应用于特定功能。现在的人工智能并不局限于分析已有的内容，而是创造、生成新的内容，由此产生了 AIGC。狭义的 AIGC 是指利用 AI 自动生成内容的生产方式，是基于以往的专业制作（PGC）和用户创作（UGC）概念提出的。广义的 AIGC 则是指像人类一样具备生成创造能力的 AI 技术，基于训练数据和生成算法模型，自主生成创造新的文本、图像、音乐、视频等各种形式的内容和数据。AIGC 产品的本质因素是互动、

数量和质量，ChatGPT 在生成和理解能力上有了显著提升，推动内容形式从单一走向多元，大幅优化了内容质量，属优质的 AIGC 的产品。

**2、ChatGPT 是自然语言处理（NLP）技术突破的阶段性成果，核心是引入了人工反馈强化学习（RLHF）。**

ChatGPT 是人工智能提升感知智能能力的阶段成果。人工智能的发展历经计算智能、感知智能和认知智能三个阶段。在计算智能阶段，仅要求机器具备存储和计算能力。在感知智能阶段，要求机器能感知外界，做到高效“看”和“听”。在这一阶段产生了两种核心技术，计算机视觉（CV）和语音识别（ASR），发展已较为成熟。在认知阶段，则要求机器可以像人类一样思考，进行独立活动。目前，人工智能已经进入提升感知智能能力的阶段，以 ChatGPT 为代表的 NLP 技术的突破是提升感知智能能力的方向之一。

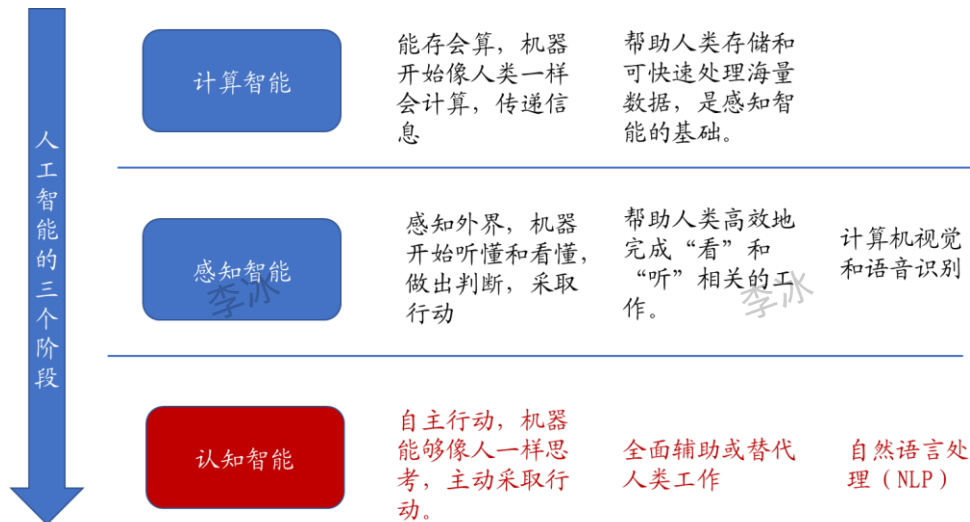


图 2 人工智能的三个阶段

ChatGPT 引入了 RLHF，表现出更强的对话能力。ChatGPT 在 GPT-3 基础上，基于人工反馈训练奖励模型，再用奖励模型去训练学习模型，这种训练模式下的模型更好地捕捉到了人工偏好，拥有出色的语言理解和文本生成的能力。具体来说，通过招募人工训练师，与 ChatGPT 进行高强度的对话，收集所有答案并对答案进行排名标注，在这个数据集上训练一个奖励模型（RM）来预测标注者更偏好的输出。最后，使用此 RM 作为奖励函数微调 GPT-3 策略，标注完成的数据再次对 ChatGPT 进行训练迭代，循环往复。GPT-3 的大模型和高质量的人工标注数据的结合，使 ChatGPT 拥有强大的对话能力。

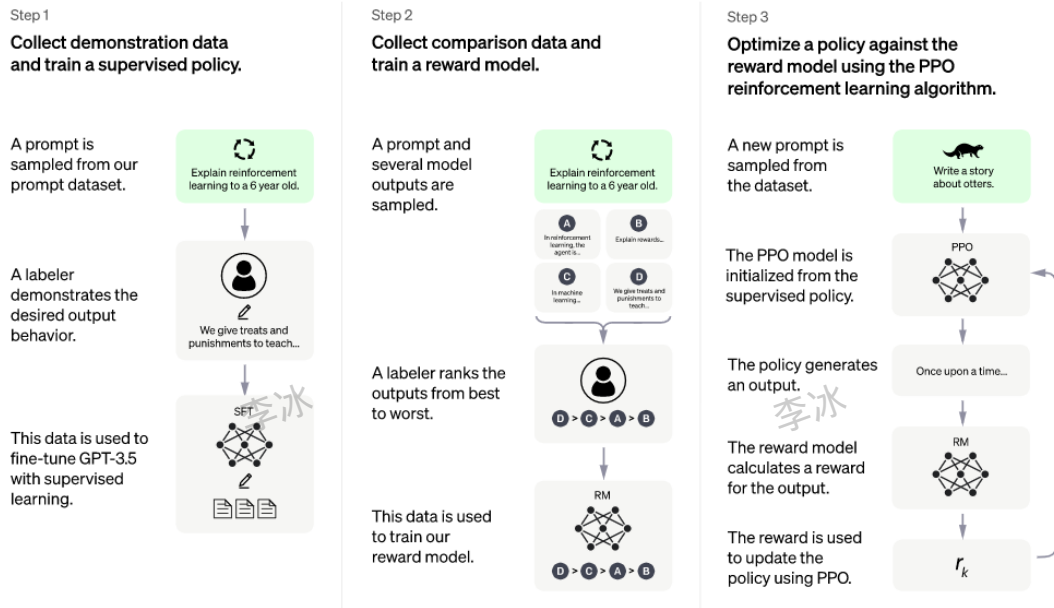


图 3 加入 RLHF 的 ChatGPT 模型训练流程

资料来源: OpenAI 官网

### 3、ChatGPT 性能优越，学术写作、对话能力出色，智能化程度较高。

由于引入 RLHF，ChatGPT 表现出了超越以往 AI 大模型的模型能力：一方面，ChatGPT 擅长提供细节丰富、篇幅较长的回答，具备一定的推理能力，在学术性问题的回答和论文写作方面具有优势；另一方面，ChatGPT 新增了主动对话的功能，能够通过上下文理解实现连续性对话，表现出与人类类似的对话逻辑，弥补了前代在自然语言理解类任务的缺陷。此外，除了人类语言的训练之外，ChatGPT 还进行了代码训练，能够输出可执行的代码。

但是，ChatGPT 也具有一定技术上的限制。首先，ChatGPT 依赖人工标记和拟合结果，回答存在一定的偏向性。其次，数据库的实时性不高，且知识的更新成本较高。最后，受到人类对大模型运作机制机理了解的限制，只能对模型进行探索优化，无法确保问题回答的准确性。这些缺陷限制了 ChatGPT 从人工辅助助手向内容自主供给方的转变。

功能、优点

缺点、不足



图 4 ChatGPT 优缺点对比

(二) ChatGPT 上下游产业发展机遇

目前，产业链雏形已基本呈现，上游主要是数据、算法和算力等基础层，研发投入成本较高，存在一定的进入壁垒，为下游提供技术支持。下游包含各种不同场景、形式的应用，已有初创企业进入市场进行布局。

1、ChatGPT 的上游包括数据、算法和算力基础设施等，进入壁垒较高。

大数据体量和算法技术革新融合是发展方向，高研发投入或成为进入壁垒。从 2018 年的 GPT-1 到如今的 ChatGPT，模型参数量和数据量都发生了巨大变化，模型参数量从 1.17 亿增加到了 1750 亿，预训练数据量从 5GB 增加到了 45T。大模型的训练需要大量的资金支持，以 GPT-3 为例，根据 Alchemy API 创始人 Elliot Turner 的推测，训练 GPT-3 模型的成本可能接近 1200 万美元。此外，大模型还需要强大的 AI 团队提供技术支持，团队不仅构建模型，还要根据行业趋势做持续优化和架构升级。资金和团队资源往往集中在几家头部企业，构成行业进入壁垒，初创公司难以进入开发层面。

表 1 国外大型互联网平台模型研发情况

公司	主要模型	发布时间	参数 (单位: 个)	预训练数据量	模型类型
OPENAI	GPT	2018 年 6 月	1.17 亿	约 5GB	自然语言模型
	GPT-2	2019 年 2 月	15 亿	40GB	自然语言模型
	GPT-3	2020 年 5 月	1750 亿	45TB	自然语言模型
谷歌	BERT	2018 年 10 月	-	-	-
	SWITCH TRANSFORMER	2021 年 3 月	1.6 万亿	-	-
微软 & 英伟达	WEGATRON-TURING	2021 年 10 月	5300 亿	-	自然语言生成



配合大体量数据和大模型，高效能的算力要求更稳固的基础硬件设施。Azure AI 超算平台为 ChatGPT 提供了强大的算力支持。Azure AI 超算平台拥有超过 28.5 万个 CPU 核心，1 万个 GPU，每个 GPU 拥有 400Gbps 网络带宽。GPT-3.5 在 Azure AI 超算基础设施上进行训练，总算力消耗约 3640PF-days（即每秒一千万亿次计算，运行 3640 个整日）。因此，具有海量数据并行计算能力、能加速计算处理的硬件基础设施是高效能算力的硬基石。

## 2、ChatGPT 的下游包含各种场景下的应用。

ChatGPT 下游应用场景多元，初创企业加速进入市场。ChatGPT 可应用于传媒、办公、影视、金融等场景，可能会推动内容生产力的重构，实现企业创作的降本增效。相比于上游基础层的高进入壁垒，下游应用的进入相对容易，初创公司数量呈指数级增长。据 PinchBook 的数据，在 GPT-3 发布的两年多以来，风投资本对 AIGC 的投资增长了 400% 以上，2022 年一级市场投资规模达 21 亿美元。

表 2 ChatGPT 上下游国内外企业布局情况

产业链	产业	国外布局企业	国内布局企业	
上游	基础计算 设施	CPU	INTEL	
		GPU	NVIDIA, AMD	华为昇腾, 燧原科技, 壁仞科技
		FPGA	INTEL, XILINX	
		ASIC		百度, 复旦微电, 寒武纪
		服务器 周边		澜起科技, 中际旭创
		数据中 心	EQUINIX	万国数据, 秦淮数据
AI 模型、 算法	深度学 习框架	谷歌 (TENSORFLOW)、 FACEBOOK (PYTORCH)、 微软 (CNTK)	百度 (PADDLEPADDLE)、华 为 (MINDSPORE) 、商汤 (SENSEPARROTS)	
	大模型	微软、谷歌、FACEBOOK	百度、腾讯、阿里巴巴、商 汤	
下游	应用	微软、谷歌、FACEBOOK	百度、腾讯、阿里巴巴、宇 节跳动、华为、商汤、科大 讯飞、思必驰	

### (三) ChatGPT 未来发展趋势

得益于下一代互联网业态对内容产出效率、内容消费的需求爆发，以 ChatGPT 为代表的对话式 AI 的商业化进程提速。ChatGPT 在推出的两个月后取得了月活用户破亿的成绩，成为用户增长最快的消费级应用程序。根据 Gartner 及 Acumen Research &

Consulting 的预测，到 2030 年，对话式 AI 的市场规模将达到 34 亿美元。从 ChatGPT 商业模式来看，变现方式大概分为两种。一种是开发商 OpenAI 长期采用的商业模式，提供 API 数据接口，执行自然语言处理任务，并按字符数收费。另一种是采用订阅费的模式，即 ChatGPT Plus 计划，收取 20 美元/人/月的订阅服务费，提供服务器忙碌时段快速响应、优先获得新功能和改进等特权。这种订阅费模式打破了 AI 技术大多应用于嵌入式项目，变现困难的固有印象，拓宽了商业模式和空间。未来，不仅 B 端用户对 ChatGPT 技术存在高需求，C 端用户对技术的付费也可能形成常态化。

表 3 AIGC 变现模式及收费方式

AIGC 变现模式	案例/收费方式
作为底层平台接入其他产品对外开放,按照数据请求量和实际计算量计算。	GPT-3 对外提供 API 接口,采用的四种模型分别采用不同的按量收费方式。
按产出内容收费。	包括 DALL. E、Deep Dream Generator 等 AI 图像生成平台大多按照图像张数收费。
模型训练费用,适用于 NPC 训练等个性化定制需求较强领域。	根据具体训练量收费。
根据具体属性收费。	版权授予(支持短期使用权,长期使用权,排他性使用权和所有权多种合作模式,用友设计图案的版权),是否支持商业用途(个人用途、企业使用、品牌使用等)、透明框架和分辨率等。
直接对外提供软件(SaaS)。	例如个性化营销文本写作工具 AXSemantics 以约 1900 人民币/月价格对外出售,并以约 4800 欧元/月的价格提供支持定制的电子商务版本。大部分 C 端 AIGC 工具以约 80 人民币/月价格时外出售,最新 ChatGPT 订阅计划售价为 20 美元/月。

但是, ChatGPT 存在数据时效性低、可靠性不足和模型成本较大等问题,大规模的商业化进程可能需要较长时间。ChatGPT 的核心价值在于多元的应用场景,不同场景下对错误的容忍程度不同, ChatGPT 在对准确度要求较高,对错误较为敏感的行业里商业落地进程较慢。高昂的模型成本也导致了场景无法下探,难以服务长尾客户。通用大模型成本高昂和模型与下游垂直场景需求匹配度不足的问题限制了 ChatGPT 的大规模推广。ChatGPT 现阶段的商业应用主要依赖于微软、BuzzFeed 等下游大厂持续投入的推动,集中在搜索、办公领域以及文本生成量较大的传媒行业。

### 1、渗透搜索引擎,或将带来搜索领域的革新。

ChatGPT 改变了搜索引擎的模式,实现从“模糊搜索”到“精准推送”的转变。用户可以以自然语言的方式获得准确的信息,实现“场景式”搜索,达到“精准推送”的效果。

在这一领域，微软于 2023 年 2 月 7 日，推出了基于 GPT-3.5 模型推出了 AI 版的 Bing 搜索和 Edge 浏览器率先实现了“ChatGPT+搜索引擎”的应用。

## 2、赋能办公领域，已初步迈入商业化阶段。

微软在 2023 年 1 月宣布将 ChatGPT 整合至全线办公产品中，加速了 ChatGPT 在办公领域应用的落地进程。目前微软已推出两款基于 GPT 技术的升级版软件，包括 VivaSales 以及高级 Teams 产品。其中 VivaSales 可自动生成场景化的电子邮件内容，比如回复询问或创建提案、提供与收件人具体相关的数据，包括定价、促销和截止日期等。高级版 Teams 可以提供会议记录自动生成、推荐任务、帮助参会者创建会议模板等帮助。其“智能回顾”功能后续将以订阅制收费，标志着 ChatGPT 相关的应用正式进入了商业化的阶段。

## 3、适配资讯平台，降低内容生产的边际成本。

可有效替代人类对已有信息进行语言整合、文字输出等，从人力驱动向全自动算法驱动转变。结合成熟的灰度测试发布，ChatGPT 能够从创作内容到发布内容形成全自动闭环，进而大幅度提高传媒平台的生产能力。新闻聚合类平台 BuzzFeed，计划今年将 AI 融合在平台的编辑和业务运行中，让 AI 协助创意过程并增强平台的内容，人类仅扮演提供创意和灵感的角色，以加强在测试、受众个性化内容生成方面的表现。

## 4、展望未来，“ChatGPT+Bot”有望实现面向 C 端的终极梦想。

在面向 C 端用户时，人形机器人的人机交互能力是刚需，需要硬件结合软件算法形成综合协同。统一的硬件是机器人作业的基础，功能则由综合算法来实现，需要智能感知能力、运动控制能力、感控一体技术、AI 算法等软硬件能力的综合配合。ChatGPT 的出现进一步提升了 NLP 技术前沿，使得人机对话体验持续优化，若接入机器人应用，有望带来更好的人形机器人交互体验，加速人形机器人的落地进程。

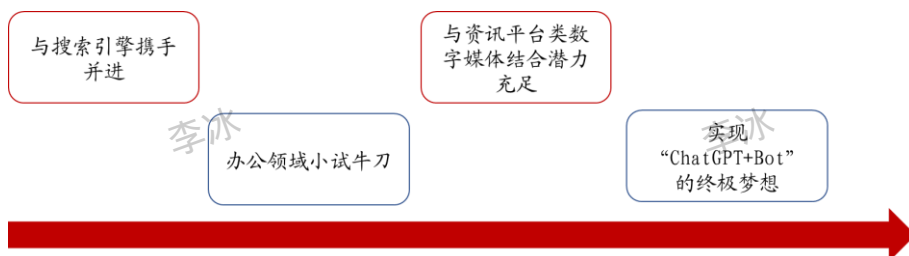


图 5 “ChatGPT+” 应用进程



## 二、国内相关政策梳理及产业发展分析

### （一）国内支持政策整理

2023年2月13日，北京市经济和信息化局发布《2022年北京人工智能产业发展白皮书》，明确支持头部企业打造对标 ChatGPT 的大模型，是国内首个明确政策支持类 ChatGPT 技术的地方政府。预计后续各地方政府将陆续发布政策支持类 ChatGPT 技术，迎来支持政策的集中发布。目前，国家层面并未有专门针对 AIGC 的政策，但从顶层设计上对人工智能领域进行了布局，相关的战略、规划或重大计划涵盖算力、算法、数据、应用场景等多个方面，为 AIGC 的发展落地保驾护航。

#### 1. 密集出台政策支持产业发展，积极探索产业规范指引和治理。

在人工智能领域的前瞻性布局，使我国人工智能发展跻身世界第一梯队，与美国形成“双雄并立”的局面。早在 2016 年《“十三五”国家科技创新规划》中，就提出要适时充实完善人工智能重大项目的布局，并在随后几年的工作报告中，不断深化对人工智能发展领域的要求，凸显国家对人工智能领域的重视。

政策支持推动人工智能技术发展步入快车道，相应的产业规范指引也在不断推出，坚持协同发展和治理。在产业标准化方面，《国家新一代人工智能标准体系建设指南》做出了顶层设计，推动人工智能产业技术研发和标准制定。在治理方面，2019 年的《新一代人工智能治理原则》初步确立了和谐友好、公正公平等 8 个原则，2021 年的《新一代人工智能伦理规范》中明确了增进人类福祉、促进公平公正、保护隐私安全、确保可控可信、强化责任担当以及提升伦理素养等 6 项基本伦理要求。

#### 2. 重点鼓励算法、算力等基础设施的建设。

AIGC 的核心竞争力是算法和算力。近几年，国家出台多项政策文件鼓励在算法和算力等关键领域的研发，涵盖对基础理论的研究，关键算法的迭代应用以及高端芯片的技术突破。在基础设施建设方面，《“十四五”数字经济发展规划》指出要推动智能计算中心有序发展，打造智能算力、通用算法和开发平台一体化的新型智能基础设施。此外，政策还鼓励平台企业作为研发主体参与到人工智能、云计算等核心技术开发中。

#### 3. 推动大数据技术发展，加强数据治理。

高质量、大体量的数据是 ChatGPT 提升问答正确率的基础，但也引发了数据安全和隐私等问题。国家政策鼓励大数据技术发展，加强新型数据中心建设，也对数据使

用和数据安全做出规范指引，构建全面的治理体系，致力解决发展和治理体系之间的匹配问题。

在推动技术发展方面，政策指出要加强新型数据中心设施、IT、网络、平台、应用等多层架构融合联动，推动新型数据中心与人工智能等技术协同发展。此外，还要推进大数据、人工智能等技术在相关领域的应用。在数据治理方面，数据安全法治意识不断增强，相关法律体系逐步构建。2021年6月，全国人大通过了我国数据安全的首部法律《数据安全法》，为数据的安全使用划定了界限，为各行业的数据安全提供了监管依据。2021年8月出台了我国第一部在个人信息保护领域的综合性立法《个人信息保护法》。除明确相关保障权益外，强化了个人信息处理者的义务，特别是为大型互联网平台企业设定了特别的个人信息保护义务。

#### 4. 开拓应用场景，加强人工智能赋能经济社会发展。

推动人工智能赋能产业发展，实现产业端的降本增效始终是人工智能发展的主线。2019年8月的《关于促进人工智能和实体经济深度融合的指导意见》指出人工智能的发展是以市场需求为导向，以产业应用为目标的。相关政策鼓励人工智能与传统制造业、医疗、养老、教育、文化、体育以及绿色低碳产业等多个领域的深度融合，开发智能家电、智慧家居、服务机器人、可穿戴设备和适老化产品等更多种类的智能终端产品。应用场景创新会倒逼人工智能领域实现技术的更迭突破，2022年8月发布的《关于加快场景创新以人工智能高水平应用促进经济高质量发展的指导意见》和《关于支持建设新一代人工智能示范应用场景的通知》都指出要深入挖掘人工智能应用场景，构建应用生态，通过打造示范应用场景推动上游人工智能技术的发展创新。

表 4 国内人工智能重点政策梳理

类别	发布时间	政策名称	具体内容
顶层设计，战略布局	2022/03	《2022年政府工作报告》	加快发展工业互联网。培育壮大集成电路、人工智能等数字产业，提升关键软硬件技术创新和供给能力。
	2022/12	《扩大内需战略规划纲要（2022-2035年）》	系统布局新型基础设施，推动人工智能、云计算等广泛、深度应用，促进“云、网、端”资源要素相互融合、智能配置。
算法、算力等基础设施建设	2021/03	《中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要》	瞄准人工智能、量子信息等前沿领域，实施一批具有前瞻性、战略性的国家重大科技项目。聚焦高端芯片、人工智能关键算法、传感器等关键领域，加快推进基础理论、基础算法、装备材料等研发突破与迭代应用。
	2022/01	《“十四五”数字经济》	推动智能计算中心有序发展，打造智能算力、通用

		发展规划》	算法和开发平台一体化的新型智能基础设施,面向政务服务、智慧城市、智能制造、自动驾驶、语言智能等重点新兴领域,提供体系化的人工智能服务。
	2022/05	《扎实稳住经济一揽子政策措施》	鼓励平台企业加快人工智能、云计算、区块链、操作系统、处理器等领域技术研发突破。
大数据技术发展和数据治理	2021/12	《“十四五”机器人产业发展规划》	推进人工智能、5G、大数据、云计算等新技术融合应用,提高机器人智能化和网络化水平,强化功能安全、网络安全和数据安全。
	2022/01	《环保装备制造业高质量发展行动计划(2022-2025年)》	深入推进5G、工业互联网、大数据、人工智能等新一代信息技术在环保装备设计制造、污染治理和环境监测等过程中的应用。
	2023/01	《关于促进数据安全产业发展的指导意见》	加快数据安全技术与人工智能、大数据、区块链等新兴技术的交叉融合创新,赋能提升数据安全态势感知、风险研判等能力水平。
应用与场景	2022/08	《关于支持建设新一代人工智能示范应用场景的通知》	充分发挥人工智能赋能经济社会发展的作用,围绕构建全链条、全过程的人工智能行业应用生态,支持一批基础较好的人工智能应用场景,加强研发上下游配合与新技术集成,打造形成一批可复制、可推广的标杆型示范应用场景,首批支持建设十个示范应用场景。
	2022/11	《虚拟现实与行业应用融合发展行动计划(2022-2026年)》	强化虚拟现实与5G、人工智能、大数据、云计算、区块链、数字孪生等新一代信息技术的深度融合,叠加“虚拟现实+”赋能能力。推进云、网、边、端协同能力体系建设。
	2023/01	《“机器人+”应用行动实施方案》	推进5G、人工智能、智能语音、机器视觉、大数据、数字孪生等技术与机器人技术融合应用,积极培育机器人校园服务新模式和新形态。
	2023/01	《关于推动能源电子产业发展的指导意见》	推动互联网、大数据、人工智能、5G等信息技术与绿色低碳产业深度融合,加快智能工厂建设,推进关键工序数字化改造,优化生产工艺及质量管控系统。

## (二) 国内相关产业发展现状

### 1、核心算力和模型算法层面,与国际领先水平仍存一定差距。

算力规模领先,但核心超算算力与需求匹配不足。根据中国信通院《中国算力发展指数白皮书(2022)》,中国与美国在全球算力规模中的份额较为接近,中国占比33%略低于美国的34%,同属于第一梯队。但是,在ChatGPT所依赖的超级计算综合性指标方面,中国总算力份额仅为18%,排名第三,与美国48%的份额差距较大。中国的超算算力规模为3EFlops,根据运营阶段所需算力支持测算,在每个用户每天收到

1500 字回答的情况下，仅能支持 ChatGPT 拥有约 4700 万的日均上线人数。随着国内企业布局的铺开，对超算算力的需求将会持续提升。

表 5 运营阶段所需算力支撑测算

日均上线用户数（百万人）	100	500	1000
用户日均提问次数（次/人）	10	10	10
AI 回复平均字数（字）	150	150	150
单日 AI 生成字数（百万字）	150000	750000	1500000
A100 服务器处理效率（字/秒）	25	25	25
单台服务器处理所需时间（万小时）	166.7	833.3	1666.7
所需服务器个数（万台）	6.94	34.72	69.44
服务器月租金单价（万元/月）	5.5	5.5	5.5
服务器单位算力（PFlops/台）	0.195	0.195	0.195
年租金（亿元/年）	458	2292	4583
所需算力支持（EFlops）	13.5	67.7	135

大算力依赖的高性能芯片需加速国产替代进程。算力实现的核心是 CPU、GPU、FPGA（Field Programmable Gate Array）等各类高性能计算芯片，并由计算机、服务器、高性能计算集群和各类智能终端承载。在各类高性能计算芯片市场上，国际企业具有绝对的市场份额和主导地位：在独立的 GPU 领域，英伟达高端 GPU 占据近 80% 的市场份额；在 FPGA 市场上，赛灵思和 Altera 市占率合计达 87%。

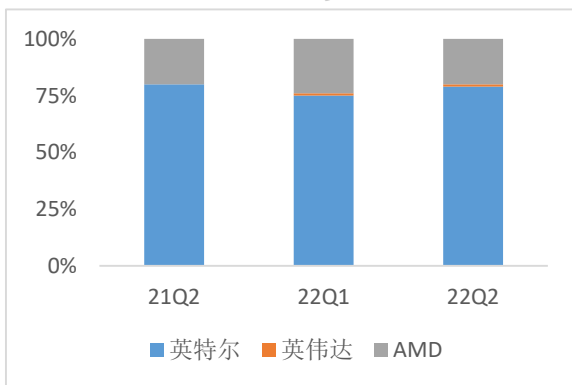


图 6 全球独立 GPU 市场份额

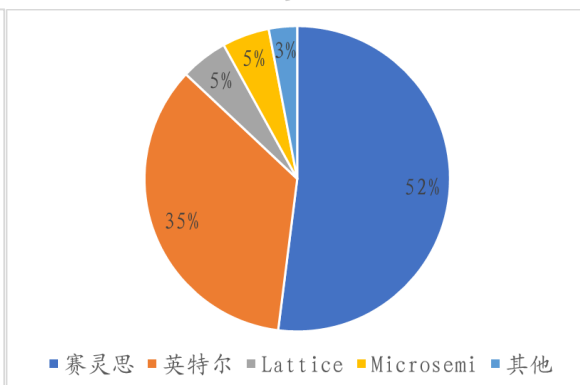


图 7 2021 年全球 FPGA 市场竞争格局

资料来源：JPR

在美国加剧对我国高端芯片进口管制下，ChatGPT 的出现进一步推动了芯片国产化进程的加速。在 GPU 市场上，景嘉微通过自主研发 GPU 并规模化商用，是国内显示控制市场龙头企业。此外，海光信息、寒武纪等公司均有布局 GPU 产品。在 FPGA 市场上，紫光国微、复旦微电和安路科技在中国市场上不断扩张，2021 年的市占率合计已超过 15%。



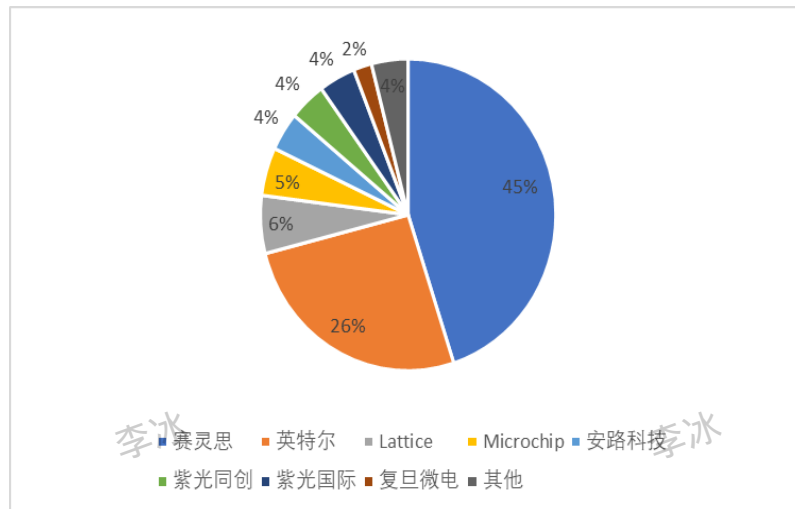


图 8 2021 年中国 FPGA 市场竞争格局

资料来源: JPR

原创模型匮乏，NLP 领域大模型算法滞后美国约 2 年，专利数量差距明显。在大模型研发方面，我国与美国 OpenAI、谷歌、微软和 Facebook 等企业存在较大差距。中国对千亿参数大模型的持有量与美国差距不大，处于国际领先地位，但原创模型相对匮乏，对算法的见解仍有待提升<sup>1</sup>。此外，我国对大模型的研发起步较晚。国外从 2017 年 OpenAI 发布 GPT-1 开始，随后几年中，几家企业陆续发布了千亿或万亿参数量的大模型，模型规模呈现不断扩张趋势。相比之下，中国的大模型集中在 2021 年到 2022 年，滞后于美国同等参数量级模型约 2 年左右。在 NLP 大模型应用方面，目前只有微软和谷歌正式发布了基于 NLP 大模型的可交互式应用，国内复旦研究团队发布的 MOSS 仍在内测阶段。对比领先 AI 公司在 NLP 领域的专利数，国内专利成果主要集中在互联网企业，腾讯、百度和阿里巴巴持有量最多，在 600 个以上。但是，微软专利持有数量超过 6000，谷歌专利持有量也超过 2000，国内外仍存在较大差距。

## 2、用户规模和数据量优势明显，垂直应用场景丰富。

我国拥有强大的基建能力和庞大的用户规模，在数据量和模型参数量上与美国差距不大，为国内企业布局类 ChatGPT 产品奠定了基础。中国的大模型主要包括百度文心、华为盘古、阿里通义、腾讯混云以及商汤的视觉模型等，百度、阿里等企业都提出要推出类 ChatGPT 产品。其中，百度计划于 3 月上线的基于文心大模型的文心一言（ERNIE Bot），目前已经有近 300 家企业宣布加入百度“文心一言”生态或将成为

<sup>1</sup>据国内 OpenBMB 开源社区统计，目前，在全球超千亿参数大模型中，中国企业和机构贡献了 1/3，美国贡献了 1/2。截至 2022 年 10 月，在拥有大模型数量及参数量前十名的组织中，中国占据 4 席，美国占据 6 席。



ChatGPT 的有力竞品。

表 6 中国大模型类型及参数

企业	预训练模型	应用	参数量	领域
百度 (文心)	文心	文本理解与创作	2600 亿	NLP
	VIMER-UFO	视觉识别	170 亿	CV
	ERNIE-ViLG2.0	语言理解和图像生成	A 作图全球最大 (目前)	多模态
	Helix	生物计算		科学计算
华为 (盘古)	NLP 大模型	文本理解与创作	千亿	NLP
	CV 大模型	视觉识别	30 亿	CV
	多模态大模型	语言理解和图像生成	千亿	多模态
	科学计算大模型	气象预测		科学计算
阿里巴巴	通义-M6	语言理解、图像生成和制造	十万亿	多模态
	通义-AliceMind	文本理解与创作	270 亿	NLP
	通义-视觉	视觉识别		CV
腾讯	混元 AI 大模型	语言理解和图像生成 (理解广告内容)	千亿	多模态
	精排大模型	精准匹配广告和人群	千亿	科学计算
商汤科技	视觉模型	视觉识别	300 亿	CV



图 9 百度文心大模型全景图

### 3、用户应用仍处于变现初期，侧重 C 端场景。

中国的 AI 应用商业化进程尚在早期，大部分产品仍处于免费试用阶段或根据使用量收费，根据内容产出量、提供服务以及使用属性等方面的多元收费模式尚未形成。

目前，国内企业围绕搜索、智能客服、语音等 C 端场景应用展开积极布局。百度的文心一言预计将融入百度搜索，提供多答案回复、智能生成等功能。阿里的云智能客服已全面应用通义大模型，在表格问答层面交付准确率达到 90% 以上，对外服务近 20 个行业。科大讯飞作为中国 AI 语音市场的龙头，具备世界领先的多语言理解、跨语言迁移和知识图谱推理等核心技术，开源中文预训练语言模型在 Github 平台的星标数位列业界第一，已去年 12 月启动生成式预训练大模型研究，预计会率先落地在教育等个别垂直领域的产品中。

### （三）合肥内相关产业发展现状

**人工智能产值增长迅速。**在政策大力支持和市场需求的双重推动下，近三年合肥市人工智能产业保持 20% 以上的高速增长。2021 年全年产值突破 815 亿元，同比增长 21.1%，集聚人工智能企业总数达 846 家，同比增长 23.2%。

**产业投融资热度维持高位。**根据 IT 桔子的统计数据，合肥市人工智能产业 2016-2022 年共发生 54 起投资事件。2020 年以来，人工智能产业投资热度攀升明显；2022 年在疫情的持续冲击下，合肥人工智能行业仍发生 11 起投融资事件，总金额达 15.3 亿元，投资热度较 2021 年收缩有限，行业扩张趋势明显。

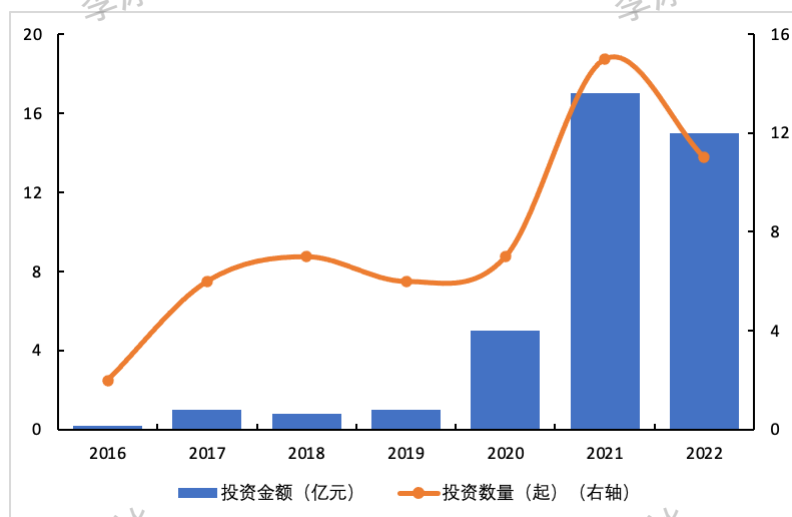


图 10 合肥人工智能产业投融资情况

资料来源：IT 桔子

**人工智能产业链条不断完善。**人工智能企业主要聚集于高新技术区(蜀山区)、庐阳区和包河区。上游基础层、中间技术层、和下游应用层环节均有企业布局。但相对来说，在中层环节发展略显薄弱，未来应加强对中游技术层企业的培养和招引。

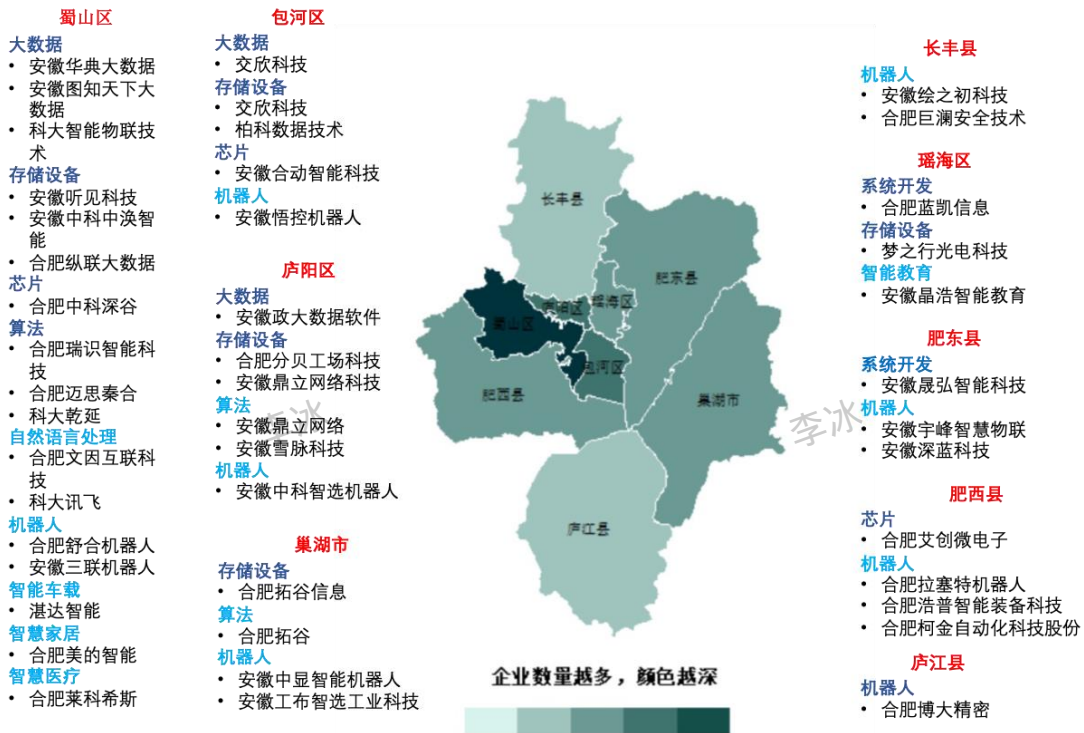


图 11 合肥人工智能企业分布

资料来源：前瞻产业研究院

### 三、促进集团相关产业业务发展的思考

一是提高对 ChatGPT 相关产业的关注度。整体来看，ChatGPT 在国内尚处于商业孵化阶段。目前，国内外主要头部科技企业凭借资金和人才资源垄断上游基础环节，行业介入机会较少。在落地场景方面，ChatGPT 仅应用在搜索、办公等文本生成类相关产品中，应用场景较为单一，多元 B 端产业落地场景仍处于构思阶段，国内尚未有成熟的类 ChatGPT 产品出现。但可预见的是，以 ChatGPT 为代表的人工智能产品未来市场空间巨大，影响领域广泛。因此在产业发展初期，应当“多看精动”，从政策引导、核心技术突破、主要成果应用、商业模式变革等角度加强对产业发展的关注，聚焦产业发展形势和未来趋势研判，为参与 ChatGPT 相关产业发展打下基础。

二是把握 ChatGPT 相关链条国产化机遇。当前，微软、谷歌等国际巨头持续加大对 AIGC 领域的投入，在扩充芯片等硬件设施需求、加速大模型发展、产业端应用落地的同时，深刻影响着国内市场的布局和竞争。智能芯片等硬件设施将首先受益于 AIGC 的发展，对强大算力的需求和美对华高端芯片的出口限制或将加速国产化芯片的

替代进程。建议集团抓住相关产业国产化机遇，加大对国产高算力芯片、模型算法等领域自主创新企业、科研团队及可转化专利成果的持续关注。

**三是围绕核心企业开展供应链金融业务。**ChatGPT 的火爆既为头部智能企业展示了新的展业方向及思路，其生态的建立也为人工智能中小企业提供了充足的市场空间。从全市角度出发，在人工语音智能领域，现阶段合肥相较其他城市发展优势明显，龙头企业科大讯飞在 AI 领域积累深厚，公司 AI 能力已在教育、医疗、政务等多个领域落地，行业语料数据量超 50TB，具备场景和数据优势，数据和模型正向反馈循环，模型的能力持续迭代优化。据公司公告显示，科大讯飞将于 2023 年 5 月发布生成式预训练大模型的产品级应用。科大讯飞在人工智能行业的快速突破，将为其上下游企业带来巨大的市场空间。建议相关单位围绕合肥市内人工智能产业链龙头企业，探索为其提供高效便捷的融资服务，重点加强与科大讯飞产业链上下游企业的挖掘和对接，拓展业务新领域。

# 免责声明

兴泰智库成立于 2016 年 8 月，是由合肥兴泰金融控股（集团）有限公司（以下简称“公司”）发起，并联合有关政府部门、高等院校、研究机构、金融机构和专家学者组成的，以“汇聚高端智慧，服务地方金融”为宗旨的非营利性、非法人学术团体。

《兴泰智库研究报告》是兴泰智库自主研究成果的输出平台，内容以宏观报告、政策解读、行业观察、专题研究为主，对内交流学习，对外寻求发声，致力于为合肥区域金融中心建设提供最贴近市场前沿的前瞻性、储备性、战略性智力支持。

本报告基于兴泰智库研究人员认为可信的公开资料或实地调研资料，但对这些信息的准确性和完整性不作任何保证。报告中的信息或所表达意见仅代表研究人员观点，不构成对任何人的投资建议，公司不对本报告任何运用产生的结果负责。

本报告版权归公司所有，未经事先书面授权，任何人不得对本报告进行任何形式发布、复制。如引用、转载、刊发，需注明出处，且不得对本报告进行有悖原意的删节和修改。