

工信洞察系列之

生成式人工智能应用 发展报告（2024）

中国互联网络信息中心

2024年11月

工信洞察系列报告编委会

编委会主任 刘郁林

编委会副主任 张 晓

编 委 李 强 吕 坚 汪立东

编 写 组 王常青 郝丽阳 郭 悦

谢思娴 武志敏

前 言

2024 年是中华人民共和国成立 75 周年，是习近平总书记提出网络强国战略目标 10 周年，也是我国全功能接入国际互联网 30 周年。党的十八大以来，以习近平同志为核心的党中央高瞻远瞩，立足党和国家事业发展全局，着眼经济社会长远发展需求，从信息化发展大势和国内国际大局出发，重视互联网、发展互联网、治理互联网，作出了加快建设网络强国的重大决策部署。我国信息领域关键核心技术加速突破，数字经济成为稳增长促转型的重要引擎，互联网深度融入居民消费、教育、医疗、养老等多个民生领域，亿万人民共享互联网发展成果。

作为新一轮科技革命和产业变革的重要驱动力量，人工智能将对全球经济社会发展和人类文明进步产生深远影响。习近平总书记指出，“中国高度重视人工智能发展，积极推动互联网、大数据、人工智能和实体经济深度融合，培育壮大智能产业，加快发展新质生产力，为高质量发展提供新动能。”长期以来，我国高度重视人工智能技术、产业发展治理，出台《生成式人工智能服务管理暂行办法》《国家人工智能产业综合标准化体系建设指南》《人工智能安全治理框架》等一系列政策文件，为包括生成式人工智能在内的人工智能技术、产业健康有序发展奠定了坚实的制度基础，推动相关领域实现高质量发展。

在此背景下，中心组织编写《生成式人工智能应用发展报告》（以下简称《报告》），基于“发展历程-政策环境-技术生态-应用探索”的分析框架，对生成式人工智能的产业发展情况、投融资情况、用户使用情况进行研究分析，展现社会各界对新兴技术的探索实践，力图为政府部门、国内外行业机构、专家学者和广大人民群众了解我国生成式人工智能的发展现状提供参考。

感谢各参与机构对《报告》研究工作的支持，希望广大读者提出宝贵的意见建议，帮助我们不断提升重点问题研究能力，形成更多具有前瞻性、战略性的研究成果。

工信洞察系列报告编委会

2024 年 11 月

目 录

第一章 生成式人工智能发展综述	1
一、 生成式人工智能在创新探索中快速发展	1
二、 生成式人工智能产业持续拓展加速演进	3
三、 发达地区成为生成式人工智能产业领头羊	4
四、 良好融资环境为生成式人工智能插上翅膀	5
第二章 生成式人工智能政策环境	9
一、 国内政策：中央与地方政府积极引导，协同发力	9
二、 国际政策：欧美亚各国政策差异明显，各有所长	12
第三章 生成式人工智能技术生态	15
一、 体系化支撑促进大模型技术加速演进	15
二、 多模态大模型助力迈向通用人工智能	17
三、 智能体繁荣推动构建人工智能大生态	19
第四章 生成式人工智能应用情况	21
一、 用户端应用持续涌现，带来智能化便捷生活体验	21
二、 企业端应用日新月异，赋能千行百业智能化升级	27
第五章 生成式人工智能发展的难点与建议	35
一、 发展难点	35
二、 相关建议	37

第一章 生成式人工智能发展综述

一、生成式人工智能在创新探索中快速发展

生成式人工智能（Generative Artificial Intelligence，简称 GAI）¹是基于算法、模型、规则生成文本、图片、声音、视频、代码等内容的人工智能技术分支。

早在二十世纪 50 至 90 年代，研究人员就已经开始利用算法来模拟人类进行创作，比如利用隐马尔可夫模型、高斯混合模型、贝叶斯网络来生成语音或时间序列数据。但这些尝试受到当时计算能力和数据资源的限制，成果较为有限。人工智能技术在内容创作方面的发展也因此遇到了瓶颈。

进入二十一世纪，深度学习的出现给人工智能的内容生成能力带来了革命性变化。2014 年，生成对抗网络的概念被提出，成为人工智能技术发展的重要里程碑，并在图像生成等多个领域取得了显著成果，为日后生成式人工智能的出现打下基础。随后，变分自编码器和扩散生成模型等技术进一步推动了图像生成技术的发展。此外，随着长短期记忆网络和 Transformer 等模型的出现，自然语言生成领域也取得了显著进展。

2022 年 11 月，OpenAI 推出了 ChatGPT 以及后续一系列模型，使得生成式人工智能开始拥有特定的产品形态，并引起了全世界的关注。这类产品不但能够与用户进行自然语言交流，还能生成高质量的文本、音频、视频内容，执行翻译、摘要和问答等复杂任务，展现了十分广阔的应用前景。此后，全球各行各业开始投入对生成式人工智能的研究和应用探索。尤其在我国的百度、阿里巴巴、腾讯、华为等企业积极推动生成式人工智能技术的研发创新与应用落地，在经济发展、民生服务、科学发现等多个领域取得了积极成果。

参考材料：自然语言处理（NLP）模型发展历程²

在 AI 生成内容的早期，不同领域，如自然语言处理（natural language processing, NLP）和图像生成领域之间，没有太多的重叠。在 NLP 领域，最初是使用 N-gram 朴素语言模型，学习单词分布，通过前一个字符来预测下一个字符。因为该模型记忆能力有限，所以

¹ 尽管生成式人工智能（Generative Artificial Intelligence）在日常生活中经常与大语言模型（Large Language Models）被混淆使用，但二者在概念上存在一定差异。一般认为，大语言模型是生成式人工智能的一个特定类别，专注于自然语言处理。

² 来源：车璐,张志强,周金佳,等.生成式人工智能的研究现状和发展趋势[J].科技导报, 2024(12).

无法生成超过一定长度的连贯文本。相比之下，基于神经网络语言模型能够生成较长的连贯文本。用于建模语言的第 1 类神经网络是循环神经网络（recurrent neural networks, RNN），RNN 逐个阅读单词，同时更新思维状态，使得该模型具备短期记忆。由于 RNN 存在着短期依赖瓶颈问题，长短期记忆网络（long-short term memory, LSTM）被挖掘出来并用于长文本生成任务之中。在理论层面，LSTM 可以实现长时间记忆。然而，在具体的实践中，经过几十到 100 个词后，该模型就开始偏离主题。为了解决这一问题，一种基于注意力机制的新型神经网络结构 Transformer 在 2017 年被提出，同时受到了广泛关注。该架构的并行化处理使其能够充分利用图形处理单元（graphics processing unit, GPU）。此外，该结构在设计层面上允许不断地堆叠编码器或解码器结构，使得整个网络结构能够变得更为复杂，这为后续大语言模型的出现奠定了基础。

自 2018 年以来，基于 Transformer 架构，大语言模型开始逐步涌现，其中最著名的当属 Google 的来自 Transformer 的双向编码器表示（bidirectional encoder representations from transformer, BERT）模型和 OpenAI 的生成式预训练（generative pre-training, GPT）系列模型，其包括 OpenAI 于 2018 年率先提出的 GPT-1 模型，约有 1.2 亿个参数。紧接着 Google 于 2019 年提出了 BERT 模型，约有 3.4 亿个参数，其整体性能优于 GPT-1。随后，OpenAI 迅速提出了 GPT-2 模型，拥有的参数量高达 15 亿，并在 40 GB 的文本上进行了训练，实现了性能的进一步提升。之后，为实现更为优异的性能，OpenAI 于 2020 年提出 GPT-3 模型，该模型具有 1750 亿个参数，其性能足以碾压之前的 GPT-1、BERT 和 GPT-2 模型。然而，由于缺乏有效的引导，GPT-3 模型在生成文本内容时常常会出现一些不令人满意的结果。

为解决这一问题，OpenAI 提出了 Instruct GPT 模型，具体通过人类反馈强化学习（reinforcement learning from human feedback, RLHF）机制引导模型生成符合预期的内容结果。基于 Instruct GPT 模型，OpenAI 于 2022 年推出了 ChatGPT，带来了 AIGC 面向大模型时代的浪潮。紧接着，OpenAI 在 2023 年提出了 GPT-4，其参数量高达到 1.8 万亿，整体性能令人惊叹。这些模型擅长文本理解，在文本分类、实体检测和问题回答等能力上具有卓越的表现。同时，其他新兴的大模型如 Sora，也为 AIGC 领域带来了新的视角，Sora 模型通过其独特的架构和进阶的多模态处理能力，进一步拓宽了自然语言处理的应用范围。2024 年，一些新的突破性研究工作进一步推动了 NLP 领域的发展。

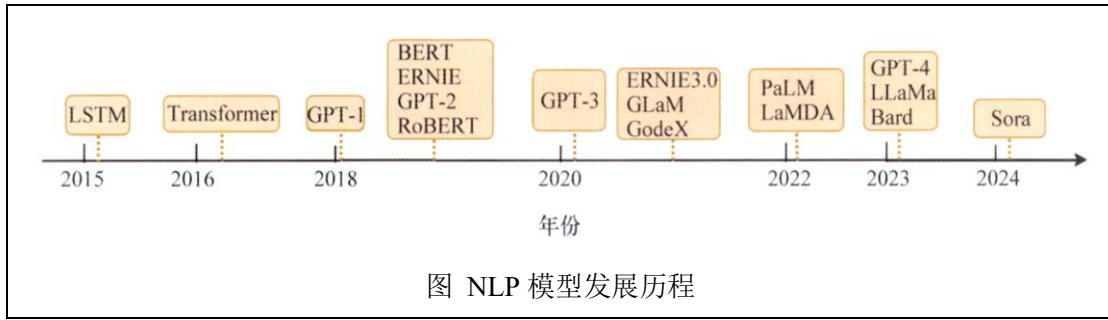


图 NLP 模型发展历程

二、生成式人工智能产业持续拓展加速演进

人工智能相关产业的发展在我国一直备受关注。早在2017年国务院印发的《新一代人工智能发展规划》就提出，“到2025年，人工智能基础理论实现重大突破，部分技术与应用达到世界领先水平，人工智能成为我国产业升级和经济转型的主要动力，智能社会建设取得积极进展”³。尤其自生成式人工智能问世以来，大型互联网厂商、中小型初创企业纷纷投入对此类产品和技术的研发，推动我国相关产业发展水平持续进步。

产业规模方面，我国人工智能产业体系更加全面。我国对人工智能技术发展一直予以高度关注，尤其自2023年生成式人工智能概念兴起至今，国产生成式人工智能大模型如雨后春笋般涌现。截至目前，我国已初步构建了较为全面的人工智能产业体系，相关企业超过4500家，核心产业规模已接近6000亿元人民币，产业链覆盖芯片、算法、数据、平台、应用等上下游关键环节⁴。

产品数量方面，生成式人工智能产品在我国百花齐放。截至2024年7月，我国完成备案并上线、能为公众提供服务的生成式人工智能服务大模型已达190多个⁵，我国以大模型为代表的人工智能普及率达16.4%⁶。百度的文心大模型、阿里的通义大模型、腾讯的混元大模型、字节跳动的豆包大模型、华为的盘古大模型、月之暗面的“Kimi智能助手”等生成式人工智能产品纷纷上线，为用户提供了丰富的选择空间和差异化的用户体验。

产业融合方面，生成式人工智能与各行各业的融合正在我国加速落地。生成式人工智能与制造业、农业、医疗、教育等传统行业深度融合，推动产业转型升级，促进新业态、新模式的不涌现。尤其在2024年，随着生成式人工智能技术的日趋成熟，各大科技企业的模

³ 来源：中国政府网，https://www.gov.cn/xinwen/2017-07/20/content_5212064.htm，2017年7月20日。

⁴ 来源：环球网，<https://3w.huanqiu.com/a/53309b/4KKnBYYjkLC>，2024年11月20日。

⁵ 来源：新华网，<http://www.xinhuanet.com/tech/20240814/57ef02dfa313413980511429a8252469/c.html>，2024年8月14日。

⁶ 来源：《第54次中国互联网络发展状况统计报告》。本数据调查了生成式人工智能产品的用户使用情况，包括ChatGPT、文心一言、通义千问等。

型调用价格显著下降⁷，从而明显降低了其他行业对生成式人工智能技术的应用成本。从智能语音助手到自动驾驶汽车，从机器翻译到智能医疗诊断，从智能制造到智慧城市，各类人工智能产品正逐步走进人们的生活，极大地提高了用户的生活质量和工作效率。

三、发达地区成为生成式人工智能产业领头羊

生成式人工智能作为新兴产业，其发展的地域分布与当地经济水平、产业结构存在明显相关性。北京、上海、广东等地充分发挥在融资机会、专业人才和政策支持等方面的优势，推动当地生成式人工智能产业蓬勃发展，形成了具有国际竞争力的产业集群。通过整理国家互联网信息办公室发布的《生成式人工智能服务已备案信息》⁸中的产品备案属地可以发现，截至 2024 年 11 月，我国共有 309 个生成式人工智能产品完成备案，北京、上海、广东三省的生成式人工智能备案产品数量占比分别达到 31.1%、27.2%和 11.7%。

表 1 《生成式人工智能已备案信息》的产品属地与数量分布

属地	数量	属地	数量
北京	96	上海	84
广东	36	浙江	25
江苏	18	四川	9
贵州	5	湖南	4
山东	4	天津	4
河北	3	重庆	3
海南	2	安徽	1
福建	1	河南	1
黑龙江	1	湖北	1
江西	1	辽宁	1
宁夏	1	陕西	1
云南	1	国资委	6

⁷ 来源：新华网，<http://www.xinhuanet.com/science/20240617/1cc3e5e73aa749268c45eec07b5d5bad/c.html>，2024 年 6 月 17 日。

⁸ 来源：国家互联网信息办公室，https://www.cac.gov.cn/2024-04/02/c_1713729983803145.htm，2024 年 4 月 2 日。

四、良好融资环境为生成式人工智能插上翅膀

生成式人工智能产品自问世以来，迅速成为全球投资热点领域，并一直延续至今。以该领域最典型的企业 OpenAI 为例，其在 2022 年底的估值仅为 29 亿美元，而在 2024 年 10 月初的估值已经高达 1570 亿美元，在不到两年的时间里增长了超过五十倍。以 OpenAI 为代表的相关企业受到资本市场的追捧，其根源在于生成式人工智能产品良好的市场前景。根据研究机构 IDC 的预测，2027 年全球生成式人工智能的市场规模将接近 1500 亿美元，复合增长率可达 85.7%⁹。我国在 2024 年初发布的《政府工作报告》也提出，要深化大数据、人工智能等研发应用，开展“人工智能+”行动，打造具有国际竞争力的数字产业集群。在此号召下，政府引导基金、民间资本、大型企业纷纷对生成式人工智能领域投入大量资金，推动该领域迅速发展。

首先，国有企业投资起到了领头羊作用。2024 年 2 月，国务院国有资产监督管理委员会召开“AI 赋能 产业焕新”中央企业人工智能专题推进会，提出“推动中央企业在人工智能领域实现更好发展、发挥更大作用”，要求中央企业加快布局和发展智能产业，加快建设一批智能算力中心等¹⁰。在智能制造、智慧物流、智能服务等方面，国有企业正通过实际行动，将人工智能技术与实体经济深度融合，以提高生产效率和创新能力。国有企业对人工智能领域的投资不仅起到带头示范作用，还带动了基础设施等产业环境的建设，为产业健康发展提供了有利条件。

其次，民间资本为产业腾飞插上翅膀。民间投资占我国全社会投资一半以上¹¹，对推动我国人工智能产业发展具有不可忽视的重要作用。随着全球金融领域对生成式人工智能的重视程度显著提高，我国民间资本对人工智能领域的投资热情也空前高涨。数据显示，2024 年前三个季度，我国共发生 504 起与人工智能相关的投融资事件，合计金额约 812 亿元¹²。在民间资本的推动下，人工智能初创企业如雨后春笋般涌现。据全国组织机构代码统一社会信用代码数据服务中心统计，2024 年上半年我国人工智能企业数量同比增长 35.65%。

最后，大型科技企业的投资力度持续加大。大型科技企业对生成式人工智能的投资与金融机构不同，往往直接将资金投入于企业内的研发、产品等环节，并已经开始产生实际的业

⁹ 来源：IDC，<https://www.idc.com/getdoc.jsp?containerId=prCHC51997124>，2024 年 3 月 28 日。

¹⁰ 来源：国务院国有资产监督管理委员会，<http://www.sasac.gov.cn/n2588025/n2588139/c30132964/content.html>，2024 年 2 月 26 日。

¹¹ 来源：国家发展和改革委员会，https://www.ndrc.gov.cn/xxgk/zcfb/tz/202211/t20221107_1340900.html?stae=123，2022 年 11 月 7 日。

¹² CNNIC 根据网络披露的投融资事件（不含定增、IPO 上市、退市等）公开资料整理并推算。

务收入。有数据显示，2024 年上半年国内三家大型互联网企业的支出较 2023 年同期提升 117%，支出增长的主要原因就是购买处理器和基础设施，为人工智能的大型语言模型训练提供支持¹³。

参考材料：2024 年 1-9 月我国人工智能相关投融资案例分布情况¹⁴

2024 年前三个季度，我国共发生 504 起与人工智能相关的投融资事件，合计金额约 812 亿元。

通过对 2024 年 1-9 月人工智能投融资事件的细分领域进行整理可以发现，从投融资事件的数量来看，投资于人工智能自身应用、投资于其他传统行业应用的投融资事件数量相同，均为 138 笔，占比均为 27.4%；投资于机器人相关领域的投融资事件数为 119 笔，占比为 23.6%；投资于人工智能硬件与技术的投融资事件数为 79 笔，占比为 15.7%。

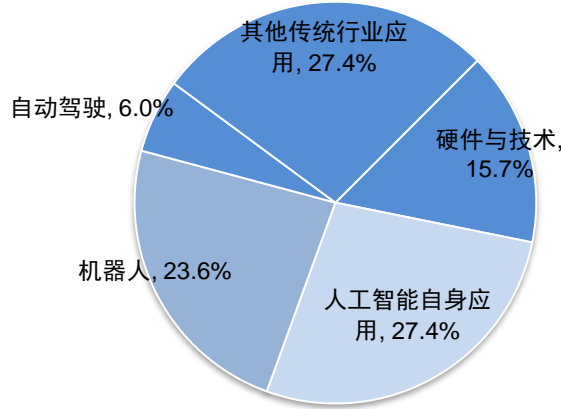


图 2024 年 1-9 月我国人工智能相关领域的投融资事件数量占比

从 2024 年 1-9 月人工智能相关领域的投融资事件金额来看，投资于人工智能自身应用的投融资金额为 371 亿元，占比高达 45.7%；其次为投资于其他传统行业应用，为 173.2 亿元，占比为 21.3%；投资于机器人相关领域的投融资金额为 111 亿元，占比为 13.7%；投资于人工智能硬件与技术的投融资金额为 96.8 亿元，占比为 11.9%。

¹³ 来源：海外网，https://m.haiwainet.cn/middle/3545022/2024/0827/content_32785586_1.html，2024 年 8 月 27 日。

¹⁴ 来源：CNNIC 根据网络披露的投融资事件（不含定增、IPO 上市、退市等）公开资料整理并推算。

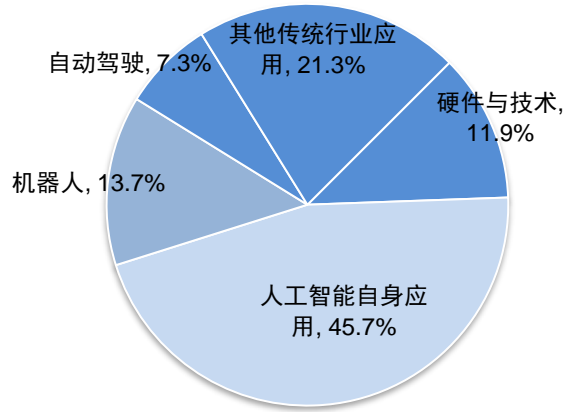


图 2024 年 1-9 月我国人工智能相关领域的投融资金额占比

从 2024 年 1-9 月人工智能相关领域投融资事件的地域分布情况来看，北京的人工智能相关投资金额高达 436 亿元，占比为 53.7%；其次为上海，金额为 112 亿元，占比为 13.8%；广东、深圳、浙江的金额均在 50-80 亿元之间。

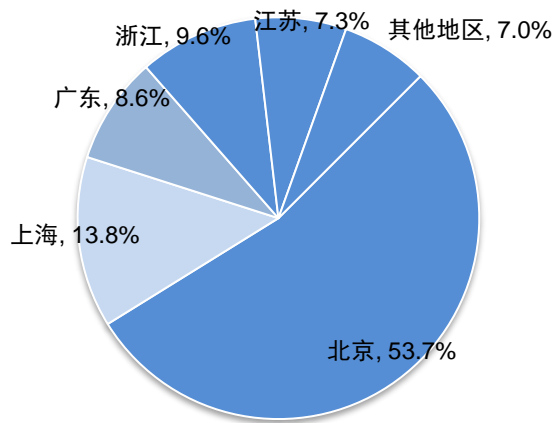


图 2024 年 1-9 月我国人工智能相关领域的投融资金额地域分布

第二章 生成式人工智能政策环境

一、国内政策：中央与地方政府积极引导，协同发力

早在生成式人工智能问世以前，我国政府就已经高度重视相关领域，并出台一系列政策推动人工智能产业健康有序发展（见表 2）。从人才培养到技术创新，从税收优惠到知识产权保护，一系列政策的密集出台，为我国人工智能相关产业的发展提供了有力支持。

整理自 2016 年至今的人工智能领域相关政策可以发现，2016 至 2020 年，我国政府出台的人工智能相关政策以发展规划、人才培养为主要目的；而在 2020 年以后，尤其从 2023 年至今，人工智能领域的政策逐渐侧重标准建设、行业规范，《生成式人工智能服务管理暂行办法》《国家人工智能产业综合标准化体系建设指南》《人工智能安全治理框架》陆续出台。

人工智能政策制定方向的变化直接反映了产业发展不同阶段。在 2022 年底 ChatGPT 问世以前，人工智能在我国产业界主要作为一项功能或技术，被应用到各行业的具体业务环节中；而 2022 年底 ChatGPT 问世之后，生成式人工智能开始作为一个独立产品的形式，在市场上大量涌现。以此为分界线，我国人工智能领域相关政策的制定也与时俱进，体现了动态调整的特点。

表 2 2016 年至今我国人工智能领域主要政策

时间	政策名称	类别
2016 年 5 月	《“互联网+”人工智能三年行动实施方案》	发展规划类
2017 年 7 月	《新一代人工智能发展规划》	发展规划类
2017 年 12 月	《促进新一代人工智能产业发展三年行动计划（2018-2020 年）》	发展规划类
2018 年 4 月	《高等学校人工智能创新行动计划》	人才培养类
2018 年 11 月	《新一代人工智能产业创新重点任务揭榜工作方案》	发展规划类
2019 年 4 月	《关于发布人工智能工程技术人员等职业信息的通知》	人才培养类
2019 年 8 月	《国家新一代人工智能创新发展试验区建设工作指引》	发展规划类
2019 年 11 月	《关于推动先进制造业和现代服务业深度融合发展的实施意见》	发展规划类

2020年1月	《关于“双一流”建设高校促进学科融合 加快人工智能领域研究生培养的若干意见》	人才培养类
2020年7月	《国家新一代人工智能标准体系建设指南》	标准规范类
2021年9月	《新一代人工智能伦理规范》	标准规范类
2022年7月	《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》	发展规划类
2022年8月	《关于支持建设新一代人工智能的示范应用场景的通知》	发展规划类
2023年7月	《生成式人工智能服务管理暂行办法》	标准规范类
2024年6月	《国家人工智能产业综合标准化体系建设指南（2024版）》	标准规范类
2024年9月	《人工智能安全治理框架》	标准规范类

与此同时，各地方政府也积极抢抓发展机遇，营造良好营商环境，积极支持当地人工智能产业发展。北京、上海、广州等地政府积极打造人工智能产业生态集聚区，引导资源投入人工智能产业创新，向相关技术人才提供优惠政策，为当地营造良好人工智能产业发展环境。自2023年1月至2024年9月，全国有16个省级行政区（或下属主要城市）发布了标题中含有“人工智能”的产业行动计划或措施文件（见表3）¹⁵；此外，还有5个省级行政区将人工智能作为重点领域涵盖在其他产业行动计划或措施文件中（见表4）。

表3 2023年1月-2024年9月各省级行政区发布的人工智能相关政策

发布时间	政策名称	发布省份
2023年7月	《关于印发重庆市以场景驱动人工智能产业高质量发展行动计划（2023—2025年）的通知》	重庆
2023年8月	《自治区人民政府办公厅关于印发促进人工智能创新发展政策措施的通知》	宁夏
2023年9月	《福建省促进人工智能产业发展十条措施的通知》	福建
2023年10月	《安徽省通用人工智能创新发展三年行动计划（2023—2025年）》	安徽

¹⁵ 来源：通过各地方政府网站搜索标题含有“人工智能”的政策文件，后经整理得到，统计时间为2023年1月1日至2024年9月30日。

2023年11月	《上海市推动人工智能大模型创新发展若干措施（2023-2025年）》	上海
2023年12月	《浙江省人民政府办公厅关于加快人工智能产业发展的指导意见》	浙江
2023年12月	《昆明市加快人工智能产业发展若干措施(试行)》	云南
2024年2月	《内蒙古自治区促进通用人工智能发展若干措施》	内蒙古
2024年4月	《陕西省培育千亿级人工智能产业创新集群行动计划》	陕西
2024年5月	《河北省人民政府办公厅关于进一步优化算力布局推动人工智能产业发展的意见》	河北
2024年5月	《南京市进一步促进人工智能创新发展行动计划（2024—2026年）》	江苏
2024年5月	《郑州市支持人工智能创新发展若干政策措施》	河南
2024年5月	《成都市人工智能产业高质量发展三年行动计划（2024—2026年）》	四川
2024年6月	《关于印发山西省促进先进算力与人工智能融合发展若干措施的通知》	山西
2024年6月	《广东省关于通用人工智能赋能千行百业的若干措施》	广东
2024年7月	《北京市推动“人工智能+”行动计划（2024-2025年）》	北京

表 4 2023 年 1 月-2024 年 9 月各省级行政区发布的将人工智能作为重点领域的政策

发布时间	政策名	发布省份
2023年1月	《江西省未来产业发展中长期规划（2023-2035年）》	江西
2023年12月	《加快推进吉林省数字经济高质量发展实施方案（2023-2025年）》	吉林
2023年12月	《山东省人民政府关于加快实施“十大工程”推动新一代信息技术产业高质量发展的指导意见》	山东
2024年6月	《湖北省加快未来产业发展实施方案（2024—2026年）》	湖北
2024年9月	《辽宁省人民政府办公厅关于科技引领未来产业创新发展的实施意见》	辽宁

二、 国际政策：欧美亚各国政策差异明显，各有所长

从国际视角来看，当前全球尚未形成国际上普遍认可的人工智能监管规则或标准框架，也缺乏专门的国际机构对人工智能相关领域的发展进行监管。欧美各国基于自身利益，在人工智能监管上总体呈现“各自为战”态势，表现出较强的差异化特征。此外，日本、韩国、新加坡等亚洲国家也正在积极制定相关政策，构建监管体系，从而应对高速发展变化的人工智能产业。

欧盟在人工智能监管方面采取了全面且严格的立法策略，旨在通过统一的法律框架来规范人工智能技术的应用及其潜在风险。自 2016 年以来，欧盟从伦理倡议开始，逐步升级到制定具有法律约束力的法案，强调基于风险的全过程监管方法。2023 年 12 月，欧洲议会、欧盟成员国和欧盟委员会三方就《人工智能法案》达成协议。该法案于 2024 年 8 月正式生效，是全球首部全面监管人工智能的法规，标志着欧盟在规范人工智能应用方面迈出重要一步。欧盟还针对特定数字场景制定监管方法，如《数字服务法案》《数字市场法案》《人工智能责任指令》《产品责任指令》修订版等，致力于率先构建人工智能领域的全球标准。

美国人工智能战略政策在 2023 年至 2024 年间展现了全面性和连贯性，通过密集出台的法规和政策（见图 1），强调了安全治理、基础创新、试点应用和国际合作的重要性。美国政府不仅在顶层设计上进行了战略布局，还在基础设施建设、人才培养、伦理法律问题处理以及跨部门合作等方面进行了深入探索和实践，同时积极推进国际合作，以确保在全球人工智能领域中的领导地位。

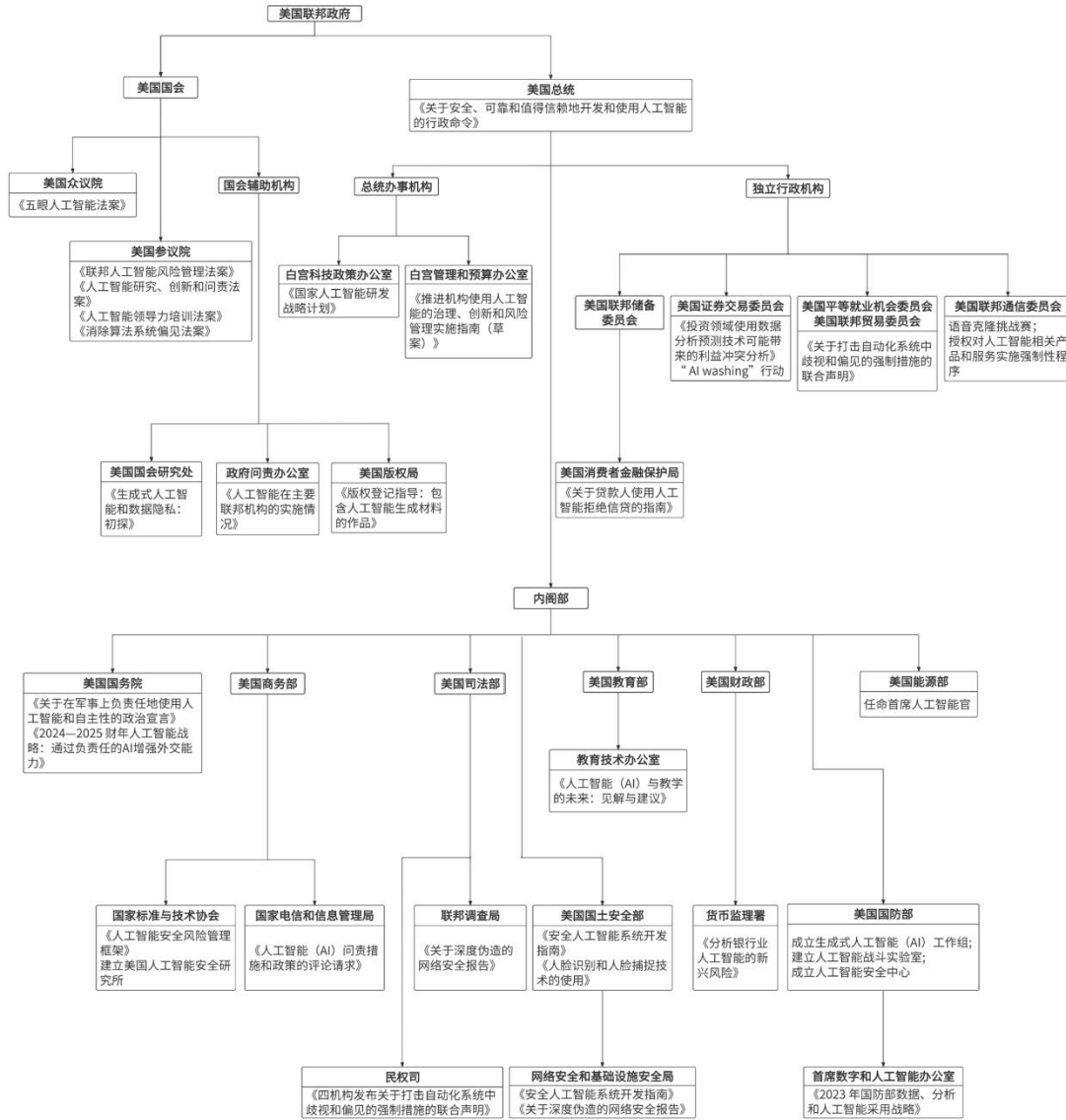


图 1 美国人工智能战略举措¹⁶

英国政府也在过去几年来连续推出人工智能发展战略，从 2018 年至 2024 年陆续发布了《人工智能行业政策》《国家人工智能战略》《英国科学技术框架》《人工智能监管的创新方法》《英国政府生成人工智能框架》《人工智能在政府中的应用》等多项政策。英国政府在其推出的人工智能相关发展战略中，多次强调人工智能在政府治理中的关键作用，希望更好地利用人工智能来提高公共服务供给，推动各部门制定人工智能实施计划。

日本在 2022 年发布《人工智能战略 2022》，涵盖人才培养、技术创新、基础设施、伦理规则等多方面目标和举措。2024 年 2 月，日本政府宣布设立“人工智能安全研究机构”，

¹⁶ 来源：宋艳飞,张瑶. 美国人工智能战略政策新动向及特点分析[J]. 人工智能,2024(2):70-78. DOI:10.16453/j.2096-5036.202410.

旨在促进日本开展可信赖的人工智能技术研发和应用。此外，日本政府内阁会议还在 2024 年 6 月通过了新版《科学技术创新白皮书》，指出日本需持续在人工智能领域进行投资，并预计 2023 年至 2028 年间，日本国内人工智能相关市场的规模将以年均 30% 的增长率扩大。

韩国由于其政治环境和产业结构，对人工智能的推动和应用较其他国家有所不同。2023 年 9 月，韩国政府宣布将投入 9090 亿韩元（约合 49.8 亿元人民币）用于推动人工智能技术的全民日常化，提升超大型 AI 的竞争力。2024 年 5 月，韩国公布《人工智能自主制造战略 1.0》，并发起了一个公私合作的人工智能自主制造联盟，计划开发一个标准模型，在 2028 年之前交付各行业使用。尤其值得注意的是，韩国军方在 2024 年提出“AI 科技强军”目标，加快人工智能军事化应用进程，并在 2028 年前建设约 90 个人工智能作战试点单位¹⁷。

新加坡一直走在全球人工智能管理框架开发和实施的前列。2023 年 12 月，新加坡发布《国家人工智能战略 2.0》，阐述了新加坡计划在未来三到五年内如何利用人工智能造福公共利益。2024 年 5 月，新加坡发布《用于生成式人工智能的人工智能模型管理框架》。该框架的制定参与者包括微软、谷歌等大型科技公司，美国商务部等政府机构大约 70 个组织和部门，对人工智能的新风险和挑战做出了回应。

¹⁷ 来源：光明网，https://mil.gmw.cn/2024-09/29/content_37590434.htm，2024 年 9 月 29 日。

第三章 生成式人工智能技术生态

一、体系化支撑促进大模型技术加速演进

(一) 技术架构变革为生成式人工智能发展打下基础

相比 PC、移动互联网时代，人工智能时代的 IT 技术栈¹⁸发生了重大变革。原来的芯片、操作系统、应用三层架构，演变为芯片层、框架层、模型层、应用层四层架构。新一代的四层架构技术体系不仅可以实现各层次间的有效沟通与协同工作，还能令整个系统得到持续优化。这种架构对生成式人工智能技术的不断迭代起到了重要支撑作用。

芯片层作为算力的物质基础，其性能直接影响了生成式人工智能模型的效能表现，因而推动芯片架构朝着专业化、定制化的方向发展。例如，英伟达的 Hooper 架构利用 Transformer 引擎，显著加快了模型的训练速度。此外，生成式人工智能模型参数规模持续增大，对芯片计算能力的需求也随之增长。高性能芯片不仅推动了框架和算法的优化，还对特定应用场景提供支持，加速了生成式人工智能技术的进步。

框架层是连接硬件与上层应用的纽带，封装了底层芯片的计算能力。它为开发者提供了易于使用的程序接口和高效的模型构建工具，有效降低了模型开发的复杂程度。与此同时，框架层还集成了自动混合精度训练、分布式训练等先进特性，进一步降低了模型开发的门槛。这使得研究人员不需要在繁琐的底层实现细节花费过多精力，让他们更加聚焦于算法创新方面，从而加速了新模型的研发进程。

模型层是生成式人工智能的核心架构，涵盖了大语言模型、视觉大模型、多模态大模型等多种类型。得益于芯片层的算力支持和框架层的优化能力，这些模型的生成效率和泛化能力得到了显著增强，为创意内容的生成、智能对话等前沿应用带来可能。

应用层是将上述技术创新转化为实际经济效益的关键。在教育、娱乐、设计、医疗等领域，生成式人工智能为各行各业提供了广阔的应用前景，催生了诸如智能辅导系统、定制化内容推送、创作辅助工具等一系列创新应用。这些多样化的应用不但丰富了用户的日常生活，

¹⁸ IT 技术栈：指在信息技术领域中，为实现特定软件应用或系统功能而组合使用的一系列技术和工具的集合。通常包括多个层次，从底层的操作系统、数据库，到编程语言、框架、前端和后端技术，以及各种开发工具和服务。

而且还刺激了对更高级别芯片性能、更优化框架设计以及更智能化模型的追求，进而形成了充满活力的发展闭环，持续推动技术革新和产业转型升级。

(二) 深度学习框架成为生成式人工智能的迭代关键

深度学习框架作为生成式人工智能的“操作系统”，上承各种业务模型和行业应用，下接智能计算芯片，具备带动产业应用和硬件适配的双向导主权，已成为推动人工智能软硬件快速迭代升级的核心焦点。

从模型支持的角度看，深度学习框架极大简化了复杂算法的开发、优化和实际应用过程，让研究人员和开发者能够更加高效地设计、训练和部署深度学习模型。深度学习框架通过提供简单易用的接口，使得非专业人员也能构建复杂的神经网络模型，降低了算法研究到实际应用的门槛。此外，框架在加速算法创新的同时，还促进了这些算法在实际场景中的应用。通过标准化的环境和流程，让深度学习技术能够快速从实验室走向市场，解决实际问题，从而推动了人工智能技术的商业化和普及化。

从硬件的设计和推广来看，在人工智能领域，芯片设计与深度学习框架的发展是相辅相成的。头部框架厂商通过其在大模型训练和推理上的实践经验，能够提供宝贵的反馈给芯片制造商。这些反馈包括对算力需求、能效比、内存访问模式等的具体要求，帮助芯片厂商设计出更符合实际应用需求的人工智能芯片。这种反馈机制使得芯片研发不再孤立进行，而是根据实际应用中的算法需求进行优化，确保芯片能够高效执行特定的深度学习任务。与此同时，深度学习框架作为人工智能应用开发的基石，它的发展趋势和特性也会影响人工智能芯片的设计。框架可能会引入新的运算类型或优化策略，这要求芯片设计时需要考虑到这些变化，以实现更加紧密的软硬件协同。通过深度学习框架与人工智能芯片的适配工作，可以加速芯片的验证流程，确保其在实际应用中的性能表现，从而更快地将新技术推向市场并得到广泛应用。这种交叉指导和合作机制，促进了生成式人工智能技术的整体进步，使得算法和硬件能够更好地协同工作，提升了整体的人工智能系统效能。

从人工智能的产业应用来看，人工智能发展的早期阶段更加重视开发先进、复杂的算法模型，而随着产业的日趋成熟，业界开始更加关注如何将这些模型整合到实际的产品和服务中，实现真正的价值创造。因此，芯片厂商与深度学习框架的适配工作变得至关重要。通过这种合作，芯片厂商能够对其硬件产品进行持续优化，确保与主流框架的高效兼容。这些软件基础设施的完善，不仅提升了大模型的运行效率，还降低了应用开发的门槛，促进了整个

产业链的协同发展。

（三）数据飞轮让生成式人工智能系统形成闭环反馈

数据飞轮在生成式人工智能系统的持续优化和升级过程中扮演着动态加速器的角色。数据飞轮建立了一个闭环反馈系统，确保从芯片硬件到深度学习框架，再到模型训练和最终的应用场景，每一层都能基于实际使用中的数据反馈进行调整和优化。这一过程促进了技术的快速迭代，使得大模型能够更好地适应各种应用场景，不仅提升了性能，还提高了效率。随着数据的不断循环利用和模型的持续学习，生成式人工智能系统的智能水平和响应速度得到显著提升，从而推动整个产业向着更高效、更智能的方向发展。

数据飞轮通过一系列高效的数据处理和学习过程，不断优化和提升大模型的性能。首先，它从广泛的、多样化的数据源中筛选出高质量的数据，这些数据经过处理和标注，成为训练大模型的基础。高性能的芯片和先进的深度学习框架为模型训练提供了强大的计算能力和算法支持，使得模型能够快速学习并形成初步能力。当模型应用于实际场景后，与用户的每一次互动都成为反馈数据，这些数据被迅速收集并反馈回系统。这种即时的反馈机制使得模型能够了解到其在真实环境中的表现。数据飞轮利用这些反馈数据，通过智能策略增强训练集，使得模型根据实际应用中的表现进行针对性的再训练，不断调整和优化，以更精准地满足用户需求。

通过上述过程，数据飞轮对生成式人工智能系统进行持续的优化循环，每次迭代都使模型更加智能，更适应复杂多变的应用场景。随着循环的进行，生成式人工智能系统的能力得到显著提升，不仅在技术指标上进步，更重要的是在理解和响应用户需求上更加贴近人类水平，从而在各种应用场景中提供更加个性化和高效的解决方案。

二、多模态大模型助力迈向通用人工智能

（一）多模态能力大幅拓展生成式人工智能应用场景

大语言模型与多模态技术的融合正引领人工智能进入一个新阶段，通过整合处理文本、图像、音视频等多种数据类型，提升深度语义理解和跨模态处理能力，以实现更广泛场景下的智能决策与交互，展现出向通用人工智能迈进的巨大潜力。在为通用人工智能奠定基础的

同时，多模态大模型也寻求在计算效率与泛化能力上的平衡，以适应从医疗到娱乐等各领域的复杂需求。

在自动驾驶领域，多模态大模型充分展示了其作为实现通用人工智能重要路径的潜力。自动驾驶是视觉大模型的重要应用场景。通过重构自动驾驶技术栈，人工智能大模型具备了检测、跟踪、理解、建图四大能力，可以突破特定场景或预先规划路线的限制，更好处理不同道路条件、天气状况等场景需求。搭载了这类大模型的自动驾驶汽车能通过摄像头、雷达和超声波传感器等多种传感器采集道路信息，并利用人工智能算法实时分析、判断和决策，识别和分析各种道路和交通场景，适应不同的交通环境，包括城市、高速公路和乡村道路等。例如，早晚高峰的人车抢行环境下，自动驾驶汽车可以做到路权优先、合理避让、兼顾通行效率，从而提升了安全性和智能性。从更广泛的角度来看，多模态大模型在处理和理解现实世界中的复杂问题时具有天然优势。现实世界中的数据和知识通常是复杂交织在一起的，而多模态大模型通过整合来自不同来源的信息，能够提供更为丰富和准确的分析结果。物理世界中复杂场景的脱敏高价值数据回传，也将反哺大模型，推动人工智能技术跨越式发展。

(二) 四项基础能力成为迈向通用型人工智能的关键

理解、生成、逻辑、记忆是生成式人工智能大模型的四项基础能力。这四项基础能力的综合运用，形成了多模态大模型在复杂任务场景下，对多类型数据输入进行处理、创作、规划、决策的能力。这四项能力越强，多模态大模型就越接近通用人工智能的目标。

在理解能力方面，大模型通过深度学习和算法优化，能够高效处理文本、图像、音频等多种数据类型，精准捕捉信息的深层含义。这使得它们在复杂环境中能准确理解用户需求和环境动态。**在生成能力方面**，利用先进的生成技术，大模型能够创造高质量的内容，从文字到图像乃至音视频，展现出强大的创意潜力，为用户提供前所未有的互动体验。**在逻辑与记忆能力方面**，通过神经网络的复杂设计，大模型能够模拟人类的逻辑思维和长期记忆，进行推理和决策时更加连贯和准确，从而让大模型能够解决需要理解上下文的任务。

综合上述能力，多模态大模型正成为解决跨领域复杂问题的强有力工具，特别是在自动驾驶、教育、医疗等需要高度智能化和多维度理解的领域。这不仅提升了生成式人工智能技术的实用性，也预示着人工智能向更高级别智能体的转变。

三、智能体繁荣推动构建人工智能大生态

(一) 智能体的繁荣推动多领域智能化水平明显提升

智能体是指能够感知环境并采取行动以实现特定目标的实体，其运作机制包括理解、规划、反思和进化，能让机器像人一样思考和行动，自主调用工具完成复杂任务。智能体的最大特点是自主性，即在无需人类干预的情况下，根据外部传感器或数据输入自主做出决策并执行相应动作。这一特点使得智能体能够适应多变的环境，更高效、更智能地执行任务。

在人工智能的发展历程中，从最初的规则系统到后来的机器学习模型，再到具备自主性的智能体，技术演进不断推动着人工智能应用拓展边界。自主性智能体的出现，让人工智能从被动响应向主动决策转变，是人工智能应用落地的重要一步。智能体也将作为推动人工智能的重要应用形态，帮助生产制造和社会生活，向更加智能、自动、高效的方向迈进。例如：**在工业生产领域**，智能体将改变传统生产模式，显著提升自动化生产水平。它们能够自主监控生产线，实时调整生产参数，优化生产流程，甚至在检测到异常时自动停机进行故障诊断和修复。这种自主性不仅提高了生产效率和产品质量，还降低了人工成本和生产风险。**在科技研发领域**，智能体能够自主进行实验设计、数据分析和结果验证，甚至自主改进实验方案。这将极大缩短研发周期，降低研发成本，提高研发效率。**在公共安全领域**，智能体通过感知城市监控摄像头和传感器网络，实时监测安全状况，快速响应火灾等紧急情况，提高公共安全事件的响应速度和效率。**在交通管理领域**，智能体可以实时监控交通流量，自主分析交通数据、优化信控策略、减少拥堵，支持规划决策，提高城市交通整体效率和安全性。**在环境监测领域**，智能体可以收集和分析空气质量、水质等环境数据，自主判定污染情况、识别污染源，提出解决方案，快速应对突发污染事件。

(二) 智能体将为凝聚用户构建生态提供重要驱动力

随着生成式人工智能技术的不断发展，无代码编程逐渐成为现实，使得普通用户利用自然语言就能进行编程。借助生成式人工智能大模型的能力，用户能够以简单、高效的方式通过自然语言创建出具有特定功能的智能体。这一变革不但降低了用户的使用门槛，使得生成式人工智能变得触手可及，极大地提升了技术的易用性，也让智能体能与具体场景应用更加契合。无论是普通人、开发者还是中小企业，都可以低成本地快速制作属于自己的智能体。

尽管无代码编程得到的智能体在复杂场景下仍有所局限，但对于许多日常应用场景，这些智能体已经能够发挥很大作用。此外，智能体之间的协作能力也在不断提升，进一步增强了其应用潜力，提升了应用上限，能够解决越来越多的实际问题。智能体的出现不仅让生活更加智能化，还为各行各业带来了巨大的创新空间。随着技术的不断进步和开发门槛的逐渐降低，围绕智能体构建的人工智能生态将进一步发展壮大，成为推动生成式人工智能广泛应用的重要力量。

第四章 生成式人工智能应用情况

自 2023 年至今，生成式人工智能相关产业在我国快速发展，新产品新模型不断迭代，新应用新业态持续涌现，为经济社会发展注入了强劲动能。以应用场景进行区分，生成式人工智能在用户端和企业端的应用情况存在明显差异。

一、用户端应用持续涌现，带来智能化便捷生活体验

在用户端，生成式人工智能以“AI 助手”“智能助手”作为主要产品形态，并与即时通信、办公软件、线上客服、创作工具等传统互联网产品进行融合，为用户提供智能化的交流、办公、创作体验。

通过对普通用户使用生成式人工智能产品的情况进行调查，可以发现以下四点情况：一是生成式人工智能产品在网民中已经形成一定用户规模，且未来仍有明显增长空间。二是与海外产品相比，国产生成式人工智能产品在网民中应用更加广泛。三是用户对生成式人工智能产品的主要使用诉求是回答问题、日常办公、休闲娱乐、创作内容。四是未成年人对人工智能相关知识存在明显兴趣爱好，并积极开展了自主学习实践。

（一）用户规模与属性

截至 2024 年 6 月，我国有 3.5 亿人表示自己听说过生成式人工智能产品¹⁹，占整体人口的 24.8%；有 2.3 亿人表示自己使用过生成式人工智能产品，占整体人口的 16.4%。对用户属性进行分析可以发现，年龄和学历明显影响了网民对生成式人工智能产品的使用意愿，中青年、高学历网民对生成式人工智能产品的使用更普遍。

从各年龄段网民的使用情况来看，20-29 岁网民使用生成式人工智能产品的比例最高，达到 40.5%；其次为 19 岁及以下网民，比例为 29.1%。40 岁及以上网民使用生成式人工智能产品开始明显下降，仅为 11.4%。

¹⁹ 在调查中，以具体的产品名称对被访者举例，如 ChatGPT、文心一言、通义千问等，而非询问被访者“是否听说过/使用过生成式人工智能产品”。

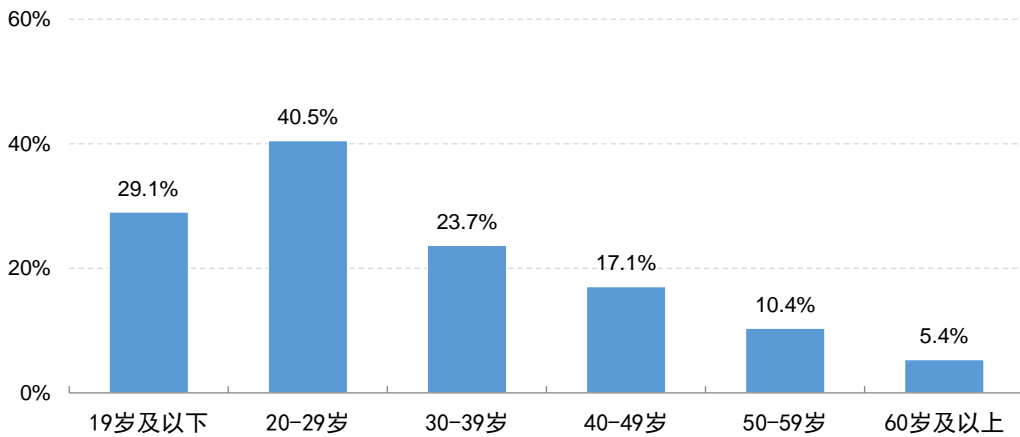


图 2 各年龄段网民使用生成式人工智能产品的比例

从不同学历网民的使用情况来看，大专及以上学历网民对生成式人工智能产品的使用率明显更高，达到 44.0%；大专以下学历网民对生成式人工智能产品的使用率则为 14.1%。

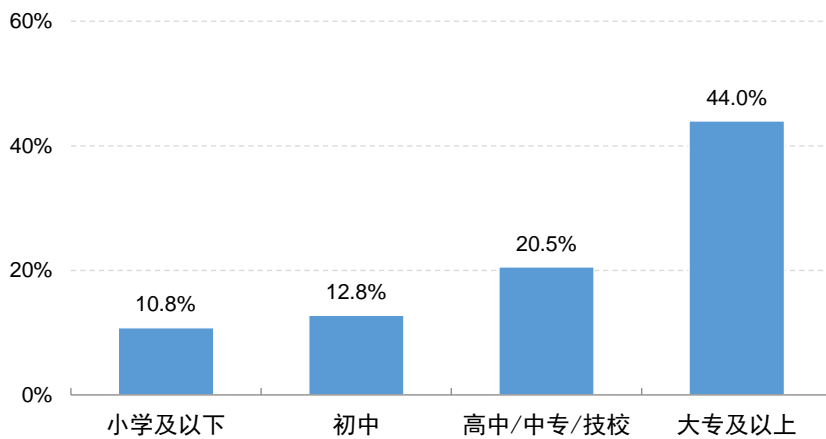


图 3 不同学历网民使用生成式人工智能产品的比例

(二) 典型产品使用情况

针对网民具体使用的生成式人工智能产品进行调查发现，截至 2024 年 6 月，国产生成式人工智能产品在网民中更加广泛；而海外产品，比如 OpenAI、谷歌等大型科技企业的生成式人工智能产品在国内也拥有一定规模的用户。

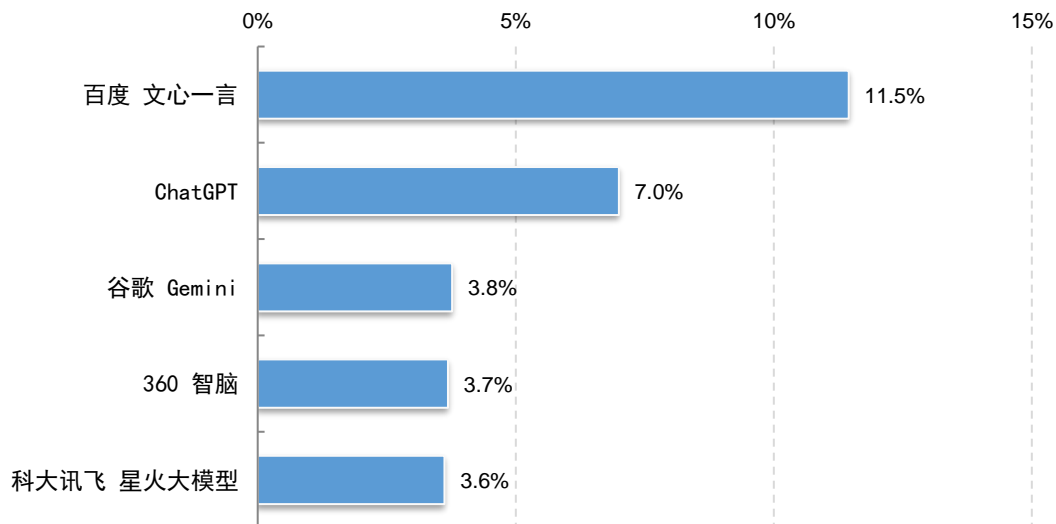


图 4 生成式人工智能典型产品在网民中的使用率

(三) 用户使用目的

生成式人工智能产品在普通用户群体中的主要应用场景包括：回答问题、日常办公、休闲娱乐、创作内容（包括文字、图片、音频、视频、代码）等。首先，利用生成式人工智能产品回答问题的用户最为广泛。62.2%的用户会对生成式人工智能产品进行提问。其次，约三分之一用户会将生成式人工智能产品作为自己的办公助手。以微软、金山办公为代表的办公软件企业已经将生成式人工智能融入到各自产品中，并以此作为新的业绩增长点。此外，生成式人工智能产品在休闲娱乐和内容创作领域的应用值得关注。尽管仅有约三成用户会将生成式人工智能产品用于休闲娱乐和创作内容，但这些作者创作的内容却影响到上亿用户。根据视频网站 Bilibili 的数据，每个月有超过 8000 万的用户在该平台观看生成式人工智能相关的内容²⁰。

²⁰ 来源：2024 年中国国际智能传播论坛。

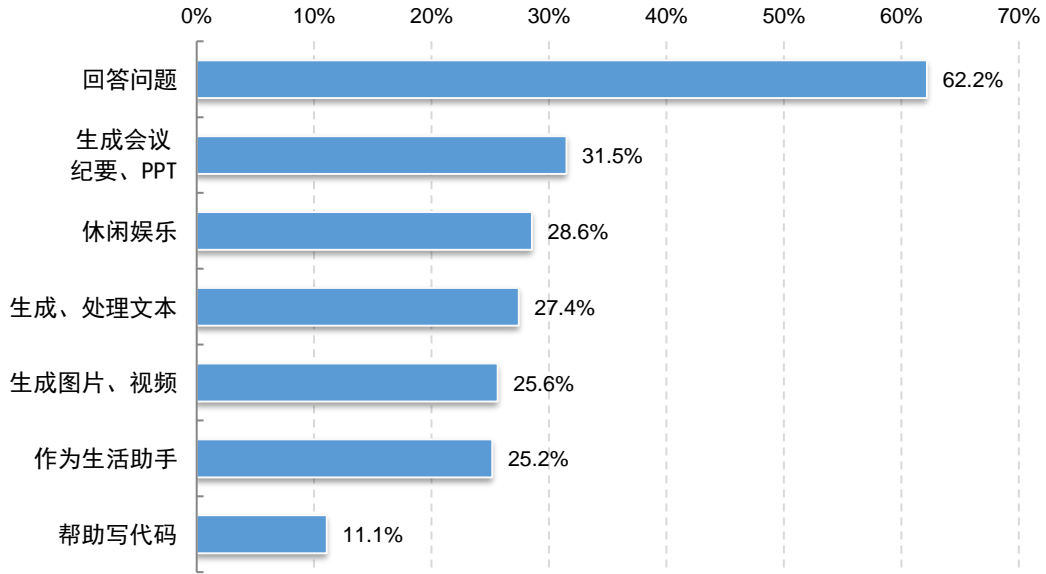


图 5 用户使用生成式人工智能产品的目的

专栏 1: 内容创作领域——大模型赋能内容获取和创作

我国内容创作市场规模庞大，文本、音频、视频等领域均对内容创作存在旺盛需求。内容创作过程中，用户往往需要进行大量的文档检索工作以支持自身创作，但由于技术水平和相关产品限制，用户不得不将检索与创作过程分离，创作效率十分有限，无法满足时间紧、质量高的应用场景要求。

生成式人工智能大模型为提升内容获取和创作效率提供了新途径，引发了智能办公领域的革新。一是实现内容获取和创作的智能化。生成式人工智能应用通过反思模型和对思维链、思维树的应用，参与到内容创作的每一个阶段，从构思、规划到最终的组装和编辑，为用户提供从创意到成品的无缝体验，让用户真正实现一键获取。二是实现内容获取和创作的个性化。生成式人工智能应用在执行环节借助混合专家模型（MoE）和检索增强生成（RAG），合理分配和调用多个不同尺寸模型，实现跨模态、跨格式内容生成，用户创作具有更高的自由度与选择性。三是实现内容获取和创作的专业化。生成式人工智能应用针对性地引入大量专业资源，通过专业知识内容的辅助来提升创作质量、拓宽创作服务范围。

例如，百度文库基于文心大模型重构为“一站式 AI 内容获取和创作平台”，发布智能 PPT、智能写作、思维导图、研究报告、拍图生文、智能画本、智能漫画、AI 搜索等多模态 AI 能力，满足了用户内容生成创作的多样性需求。百度文库的智能 PPT 功能具备辅助构思、丰富生成交互、内容生成与理解等能力，从 PPT 创作全过程为用户赋能提效。在

长文写作上，百度文库推出的 AI 产品“橙篇”，让用户可以轻松实现对超大量、超多格式、超长内容的文件进行深入理解、精准总结与即时问答，还支持超长篇幅的长文生成、深度编辑以及多模态的自由创作，极大地丰富了用户的内容创作手段²¹。百度文库、百度网盘联合推出的“自由画布”，打通了公域与个人授权的私域内容，通过“一拖一圈”的极简操作，实现对多格式、全模态文件的混合理解、生成、创作，并支持对富媒体文档的一键分享和存储，为用户提供“创作、编辑、存储、管理、查找、观看、使用、分享”的一站式服务，覆盖从内容生产起点到内容消费终点的全流程。

(四) 在未成年群体中的认知

未成年人是国家的未来，也是生成式人工智能产品的重要使用者，因此其对人工智能相关知识的认知和学习，具有非常重要的意义。针对老师（主要为班主任）的调查发现²²，94.4%的老师认为有必要让学生了解或学习人工智能相关知识。针对在校学生调查发现，40.6%的未成年网民表示听说过人工智能技术且对此非常感兴趣；16.4%的未成年网民表示没听说过，而这部分群体主要为小学生。

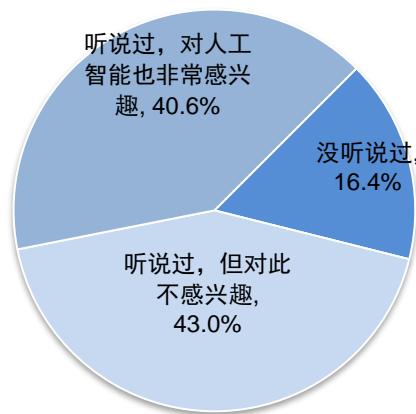


图 6 未成年网民对人工智能技术的认知

通过对未成年网民了解人工智能技术的信息渠道进行调查发现，大部分未成年网民主要是在上网时了解到这一技术，占比达到 55.5%；老师介绍和课外阅读也是未成年人了解人工智能技术的重要渠道，占比分别为 37.0%和 36.4%。

²¹ 来源：环球网，<https://tech.huanqiu.com/article/4I0CRcarCBa>，2024 年 5 月 30 日。

²² 来源：《第 5 次全国未成年人互联网使用情况调查报告》。本段数据调查的是未成年人对“人工智能”的认知情况，而非仅限于“生成式人工智能”。

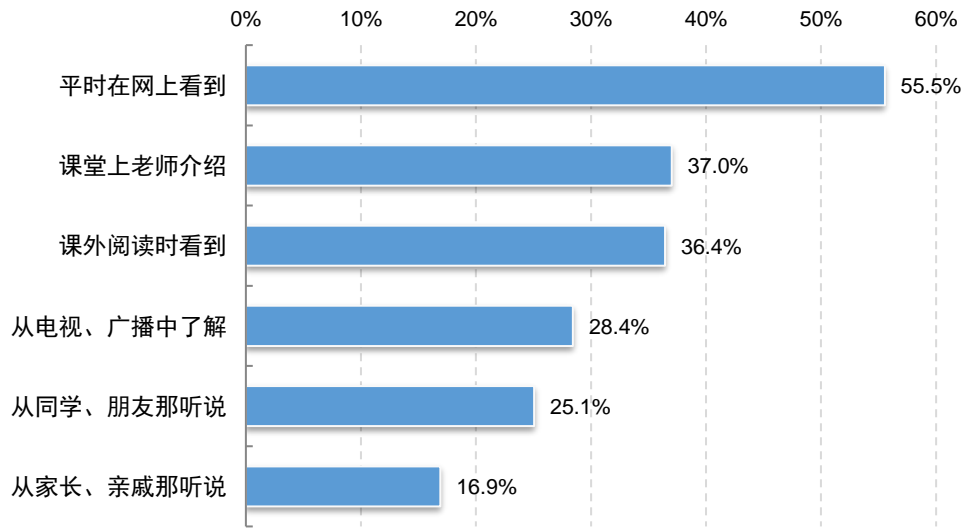


图 7 未成年网民了解人工智能技术的渠道

二、企业端应用日新月异，赋能千行百业智能化升级

在企业端，各行各业都在积极拥抱生成式人工智能带来的智能化升级浪潮。在交通、能源、制造、化工等多个领域，高科技企业与传统行业积极进行合作，投入大量资源，共同研发各行业专用的生成式人工智能大模型，探索如何利用这项新兴技术赋能实体经济创新发展。在 2023 年底开展的一项针对全球 16 个国家的企业管理人员调查中，79% 的受访者预计生成式人工智能将在三年内为组织和行业带来实质性变革，其中近三分之一的受访者预计实质性变革将在当下（14%）或不到一年内（17%）发生²³。

（一）应用效果：各领域争相进行智能化改造实现降本增效

随着“人工智能+”行动等政策深入推进，生成式人工智能大模型在各行业的应用成效开始显现。降本和增效是各行业选择应用大模型的核心驱动力。大模型的应用不仅能够自动化完成重复性任务，还可以通过深入数据分析提供更为精准的洞察，从而显著提升机构的工作效率。例如，海淀区政府利用政务大模型，将查找数据、指标计算等场景所需的 3 天工作降低到 1 分钟，把图标绘制、报告撰写等场景所需的 5 天工作降低到 30 分钟²⁴。此外，通过提升客户体验和加速产品开发周期，大模型还为企业创造了新的价值增长点，使其能够在不断变化的市场中保持灵活性、发现新机遇。

当前，各行各业都在尝试利用生成式人工智能技术实现智能化升级。从农业生产中的智能化管理，到交通领域的出行效率提升，再到科学研究中的创新实验，各个领域将大模型与实际的应用场景结合后，都产生了较好效果。

专栏 2：智慧交通领域——贵阳智能信控驱动城市交通精细化管理

生成式人工智能大模型正在推动交通行业的转型升级，为建设交通强国创造了战略性的新机遇。一是**感知能力升级**，大模型能够提升交通路侧设备的精准感知能力，更加清晰地描绘出交通运行图；二是**研判能力升级**，大模型能够强化对图像、文本和地图数据等的研判能力，准确识别出道路隐患信息；三是**优化能力升级**，大模型进一步提升对道路路况的研判效率，优化交通信号控制效果；四是**服务能力升级**，大模型驱动的智能问答，为管理者提供高效准确的问答服务，实现便捷交通管理。

²³ 来源：德勤《企业生成式人工智能应用现状》2024 年第一季度报告，2024 年 3 月。

²⁴ 来源：环球网，<https://tech.huanqiu.com/article/4KIwjPcebsa>，2024 年 11 月 18 日。

在此基础上，大模型能对各路口、各流向的交通流量进行精准感知和预测，根据城市路口、干线等区域的差异，高效应对恶劣天气、施工占道等交通环境快速变化，因时因地制定信号控制优化策略，动态调节路口信号灯时长，实现城市全域交通通行效率的提升。

以贵阳市为例，近年来，贵阳市汽车保有量持续攀升，如何解决交通拥堵这一“城市病”，成为摆在城市管理者面前的共同问题。此前，贵阳市委市政府开展了一系列治理交通拥堵举措，取得了阶段性成效。但由于地形的限制和道路条件的约束，距离彻底破解贵阳城市交通拥堵，仍面临着一定难题。

为此，贵阳市公安局交通管理局联合百度智能云，选取观山湖核心区域的 31 个重要交通节点落地智能信控，通过充分融合互联网和贵阳本地交警数据，对路口交通流量进行感知与预测，能够预测出 5 分钟内的交通流量分布，准确率在 80% 以上。随后，系统会动态调节每个路口的红绿灯时长，在高峰期最大化绿灯利用率。同时，与“百度地图”APP 的信息进行同步，实现红绿灯读秒、绿波提示、交通事故、道路施工等多重信息实时推送更新。通过多项举措联合实施，该区域工作日车均延误平均下降 15%，全天拥堵指数从 1.7 降至 1.58，早高峰拥堵时长缩短 30 分钟，显著提升了市民交通出行体验。

未来，贵阳市公安局交通管理局还将进一步依托百度智能云交通大模型，以大语言模型、视觉大模型、跨模态大模型支撑全域研判、全域信控、安全防控、数字应用四大类场景，对全市所有路口进行一盘棋优化，以创新技术手段助力贵阳市道路交通管理向全域智能控制体系迈进，促进交通系统的现代化升级。

专栏 3：智慧农业领域——中国农业大学发布“神农大模型 2.0”²⁵

7 月 15 日消息，中国农业大学近期正式发布了“神农大模型 2.0”。相较于“神农大模型 1.0”，“神农大模型 2.0”实现了技术上的突破，不仅包含农业知识问答、农业文本语义理解、文本摘要生成、农业生产决策推理等核心功能，更在支持图像、声音、视频、文件等多模态交互、智能化推理等方面取得了显著进展。

从技术上看，“神农大模型 2.0”在知识图谱、向量数据库、检索增强生成等常规大模型技术基础上，在多模态编码器、多模态投影等方面进行了创新。

技术的更新迭代促进了“神农大模型 2.0”功能的全面提升，同时，中国农业大学还发布了“神农·固芯”育种大模型、“神农·筑基”种植大模型、“神农·强牧”养殖大模

²⁵ 来源：环球网，<https://tech.huanqiu.com/article/4IcS99uYdM0>，2024 年 7 月 15 日。

型及“神农·问穹”遥感气象大模型，能够服务于育种、种植、养殖、遥感和气象等农业多场景的实践应用。在应用中引入了多智能体的设计思路，实现农业物联网、传感器、农业智能装备与大模型的互联融合，进一步提升了现代农业智能化控制与决策的效率。

(二) 应用环节：研发与服务环节最先尝试进行智能化升级

整体来看，生成式人工智能大模型在产业链各环节的落地应用进展并不完全一致，呈现出“研发与服务先行，中间生成制造等环节后发”的特征。在产品研发方面，大模型可以迅速提高数据分析和预测能力，加快产品研发和创新步伐。在市场应用方面，特别是在营销、运营等服务场景下，智能化的解决方案能够快速提升用户体验和服务效率。目前，在法律咨询、智慧诊疗、线上客服和智能机器人等领域，基于生成式人工智能技术的“智能助手”已经十分常见。在生产制造和供应链管理等中间环节，大模型的整合和优化往往面临技术复杂性高、数据质量和积累不足、传统流程改造等挑战，需要在积累实践中持续突破。从当前应用进展看，信息提炼、智能交互、辅助决策是生产制造类企业与大模型结合的共性需求。例如，中天钢铁与百度合作，对钢铁行业的数据库和知识库进行建设，在此基础上部署中天南通大模型，落地在生产、财务、人事等多个应用场景。在企业管理方面，依托基于文心大模型的自助数据分析工具辅助决策，提升了企业的经营决策能力；在设备运维方面，依托智能运维助手实现了设备运维端到端的业务闭环，强化了对一线运维人员的业务支持能力，将设备维护效率提升了10%。

专栏 4：科研探索领域——国家天文台人工智能工作组发布大模型“星语 3.0”，已接入望远镜阵列²⁶

4月14日，中国科学院国家天文台人工智能工作组发布新一代天文大模型——“星语 3.0”。“星语 3.0”基于阿里云通义千问开源模型打造，目前已成功接入国家天文台兴隆观测站望远镜阵列——Mini“司天”。

这是大模型在科学领域落地的经典案例，也是大模型在天文观测领域的首次应用。

天文望远镜是人类探索宇宙的“照相机”。相比单体望远镜，大型望远镜阵列能更有效地整合高性能望远镜资源，成本更低，观测效率更高。随着望远镜数量的不断增加，如何控制大型望远镜阵列已成为当今天文界共同面临的挑战。

以往天文观测主要依赖观测助手和科研人员的配合。科研人员往往需要根据观测所在

²⁶ 来源：央广网，https://tech.cnr.cn/techgd/20240415/t20240415_526665674.shtml，2024年4月15日。

地气象情况修改观测计划，在观测室等待数据返回并实时分析数据，再根据结果修改观测计划。重要观测目标的每个环节都需要人工参与，效率低且难以同时控制多个望远镜。

“星语 3.0”尝试解决这一难题。依托阿里云通义千问底模和百炼平台，“星语 3.0”基于超 30 万专家订正数据完成训练，在天文物理等专业能力上表现突出。

目前，“星语 3.0”已成功接入国家天文台兴隆观测站望远镜阵列——Mini “司天”，可实现自主控制望远镜进行观测、分析观测结果，智能地给出下一步观测建议。

例如，当收到“观测某星体”的任务时，“星语 3.0”首先查询某星体的坐标信息，反馈给观测人员；当得到确认后，“星语 3.0”将按照观测人员输入的计划进行自动观测；单次曝光完成后，大模型将回收并处理数据，根据结果判断目标源的观测价值，推荐接下来的观测计划。

未来，星语大模型将接入国家天文台大型望远镜阵列“司天”（Mini “司天”即为“司天”的一部分）。“司天”是我国天文学家面向时域天文学提出的国家级重大天文基础设施项目，至 2030 年米级望远镜数量预期达到 72 架，每晚产生约 140TB 处理后数据，成为全球巡天效率最高的项目。星语大模型的接入将科研人员从繁琐的观测中解放出来，更加专注于研究。

专栏 5：工业生产领域——中化信息利用大模型创新新材料研发范式

中化信息是世界 500 强企业中国中化的直属数字科技公司，担负着支持中国中化数字化转型，助力提升研发效率和安全保障能力的重任。过去，在化工行业新材料研发领域，科研人员需要依赖人工经验，进行筛选目标、设计实验、报告分析等环节，整个研发过程要对分散在各网站文献、业务系统中的众多专业信息进行检索，严重影响研发效率。在安全生产领域，由于传统视频监控识别算法准确率低，带来企业运维管理难度大、安全隐患响应和处置不及时等难题。

中化信息与百度合作探索人工智能在化工行业的应用，基于千帆 AppBuilder 打造“化小易”知识助手，能够基于自然语言提问，快速检索并回答分子特性、分子合成路线等专业知识，替代了传统的人工查询方式，对特定分子的检索效率提升 5 倍以上，极大提高了研发效率。基于文心大模型提升在安全生产能力，建设工业视觉智能应用平台，在企业侧对生产区域内不安全行为进行实时视频分析，形成报警数据，将安全事件响应效率提升至分钟级，大大提升了企业安全生产的综合能力。

中化信息利用“化小易”提升新材料研发过程中的信息获取效率，辅助专家作出决策，是大模型在化工新材料研发领域的创新落地。未来，中化信息也将继续探索大模型在分子材料特性预测、辅助合成设计、自主学习迭代优化流程等方面应用，利用大模型进一步提升安全生产管理，帮助企业在技改中找到更高的生产力，推动数字科技与产业深度融合，引领化工行业智能化发展。

生成式人工智能大模型的初期应用侧重于解决实际操作中的效率问题，而随着技术的成熟和企业对其潜力的深入理解，其应用范围将逐渐扩大，影响企业的决策制定和战略方向，乃至整个行业的结构和运行逻辑。

从短期看，生成式人工智能大模型的应用聚焦于人机交互频繁且任务重复性高的场景。这些场景对自动化和效率提升的需求最为迫切，且相对容易验证模型的准确性和效率。例如，在客户服务、文档处理、数据分析等领域，大模型可以迅速提升处理速度，减少人力成本，同时通过不断优化，提供更加个性化的用户体验。这种应用策略考虑到了技术成熟度、企业接受度以及快速实现价值回报的需求。

从长期看，生成式人工智能大模型的应用将不仅仅局限于执行操作层面，而是会扩展到决策支持和战略管理层面。这意味着大模型将更深入地参与到企业的核心业务流程中，如市场预测、风险评估、策略规划等，这就对大模型的理解和分析能力提出了更高要求。随着模型的不断学习和优化，它们将能够提供更加精准的决策建议，甚至在某些情况下自主做出决策。这种变革不仅会改变工作方式，还会推动企业组织结构、运营模式的创新，以及产品和服务的创新设计，最终可能重构整个行业的运作模式。

专栏 6：智慧能源领域——深圳燃气利用视觉大模型提升安全生产能力

随着城市燃气管道网络的不断扩展，燃气设施的监管与维护工作日益繁重。为保障居民用气安全，深圳燃气已投入大量资源，建设了数千路的视频监控设备，并推动二级业务单位广泛采用移动布控、执法记录仪、操作记录仪等物联网设备，以实现更全面的监控与管理。但是由于智能化识别能力不足、系统集成度不够，现场巡查仍然以人工巡查为主，不仅难以满足当前对安全、效率的高要求，面对井盖未关、工程器械防护不足、燃气泄漏等危险情况时，人工响应的滞后性可能导致安全隐患的扩大，影响居民用气安全。

为此，深圳燃气与百度合作探索将大模型应用于燃气行业，首次在行业引入视觉大模型，建设智能视频中台，解决燃气企业运营场景繁杂、识别困难等难题。视频中台应用场景涵盖燃气场站、营业厅、光伏电站、燃气热电厂、工地、执法及物业管理等多个领域，

目前已应用在 68 个场站、58 个营业厅，接入监控设备超 2000 多路，每天为各类其他来源的 14 万张图像进行集中分析、识别、业务闭环处理，在安全管理落实、工程质量提升、操作规范保障等方面一线实践中取得显著效果。

集成大模型能力的智能视频中台，提升了深圳燃气安全管理的整体智能化和数字化水平。实现了视频采集、加工处理、管理的全流程能力提升，原本隔离不互通的各个监控系统成功实现互联互通，系统层级减少，大大提升管理人员效率。大模型提升了智能化分析水平，视频中台能够充分挖掘信息背后隐藏的数据价值，为管理人员提供辅助决策，提升“自动预警”能力。视觉大模型强化了视觉识别算法，比如烟火识别、穿戴识别、槽车作业区管控、静电放电监控、光伏板清洁度监控等等。过去，这些主要靠人力巡查来完成的识别任务，非常考验巡检员的“眼力”。如今，人工智能算法模型库里累积了各种识别算法，相当于把经验丰富的巡查员的所有知识进行数字化和沉淀。同时在一些危险区域还可以代替人力巡察，消减盲区之余还可以降低巡察人员工作风险。

未来，深圳燃气还将继续推进大模型落地应用，将通过建设统一视频数据中心，建立人/车/设施/事件等多维数据的关联图谱等举措，进一步挖掘视频数据价值，对视频 AI 感知的数据进行归档统计，让大模型不仅助力发现问题，并且通过挖掘视频数据价值，为企业生产运营提供科学有力的决策依据。

(三) 应用领域：多模态能力拓展生成式人工智能应用边界

多模态能力是显著拓宽生成式人工智能应用领域的重要因素。通过整合文本、图像、语音和视频等不同数据类型，生成式人工智能大模型可以实现对各类信息的深层理解和联动。多模态能力赋予大模型更高的灵活性和适应性，使其能够在复杂的现实场景中提供更精准的解决方案，从而推动行业的深刻变革。多模态大模型能够更好地处理噪声和数据缺失。如果一种模态不可靠或不可用，系统可以依靠其他模态来维持性能。例如，生成式人工智能的多模态能力显著提升了自动驾驶系统的安全和交互性能。通过对摄像头、激光雷达等多种传感器信息的理解与判断，可以实现对车辆周围环境的全面感知，支持车辆做出即时决策和精准控制。

专栏 7：自动驾驶领域——国内外自动驾驶出行服务（Robotaxi）提速发展

自动驾驶是视觉大模型重要应用场景，是新质生产力的典型代表。人类驾驶员会因疲劳、分心等原因导致驾驶失误，自动驾驶可以做出更为理性、准确的驾驶决策，降低事故

率。相关数据显示，2023年特斯拉全自动驾驶，平均每百万英里(约161万公里)，发生事故的次数为0.21次，显著低于同期美国道路交通对应数据1.49次。国内也有数据显示，自动驾驶车辆出险率仅为人类司机的1/14。

当前，加快推进自动驾驶出行服务已成为国内外共识。在一些国家，自动驾驶服务已经加速应用。美国加州公共事业委员会(CPUC)公布的数据显示，谷歌旗下自动驾驶出行服务公司Waymo在旧金山市8月的无人车日均服务单量约8800单，高于同期的出租车单量约6300单。此外，特斯拉也在10月发布Robotaxi车型，预期将在2026年大规模推开Robotaxi服务，并量产Cybercab无人驾驶出租车；通用旗下Cruise宣布，将在2024年晚些时候恢复全无人驾驶服务，计划2025年初重启收费服务。美国出行服务平台Lyft、Uber等也在加快布局自动驾驶。

在国内，北京、武汉和上海等地都在积极推进自动驾驶应用。以武汉为例，武汉市智能网联汽车开放测试道路里程累计突破3379公里，覆盖武汉12个行政区，辐射面积约3000平方公里，触达人口超770万，实现跨区通行、跨江通行、机场高速通行等多个自动驾驶应用场景的全国创新突破。以萝卜快跑为代表的国内企业正加快进入全面拓展阶段。目前，萝卜快跑已覆盖全国10余个城市，并在多个城市提供了自动驾驶出行服务，在超1亿公里实际道路测试里程中，实现无重大伤亡事故。

自动驾驶在提供产业增量空间、提升交通安全出行服务上，已经呈现了令人期待的应用前景。未来，随着技术的持续进步和自动驾驶出行服务模式的不断发展完善，将能够为公众提供更加便捷、高效的出行体验，提升交通安全性与效率，让科技更好服务民生。

(四) 应用方式：智能体成为连接用户与各类服务的新桥梁

智能体是目前生成式人工智能应用的主流形态之一，能够通过理解、规划、反思和进化，执行复杂任务并做出决策。在应用层面，智能体能够根据特定目标自主或半自主地完成任务。它们既可以作为人类的助手协同工作，也可以独立完成复杂任务。同时，智能体不局限于单一软件形式，可以调用工具，与其他实体协作，成为更强大的智能网络。例如，在制造行业，智能体能实时监控生产线的运行状态，通过对实时数据的深度分析，实现自主预防性维护，从而提高生产效率和设备可靠性。在零售业，智能体可以优化库存管理，通过分析消费趋势和预测需求调整库存水平，提高供应链效率，并为客户提供个性化的购物建议，提升客户体验。在能源管理方面，智能体在能源行业可以优化电网管理，通过分析供需数据，调整能源

分配，提高能源利用率，支持可再生能源的整合。当前，OpenAI、谷歌、百度等国内外企业相继推出智能体产品。如百度文心智能体平台已吸引 15 万家企业和 80 万名开发者²⁷。

专栏 8：智能体成为生成式人工智能应用的主流形态之一

从产业界实践来看，生成式人工智能的发展将经历三个阶段。第一个阶段是“AI 助手”，需要人类进行辅助、把关后，才能交付成果；第二阶段是“智能体”，具备自主使用工具、反思、自我进化等能力；第三阶段是“AI Worker”，能够像人一样独立完成各种脑力和体力劳动。当前生成式人工智能的发展正处于智能体阶段。随着技术的不断迭代与创新，智能体将持续扩大应用范围、扩展工作能力，在各行各业中发挥着日益重要的作用。

例如，在政务服务领域，中卫慧通基于百度文心大模型打造的基层政务智能体“村长 copilot”。“村长 copilot”服务基层政务工作人员，具备信息获取与整合、工单管理、方案定制、联系方式管理等功能，可以帮助工作人员快速整合各类信息、制定活动方案、自动生成工单等等。“村长 copilot”与面向村民群众的智能体“居民助理”配合，实现大模型对基层治理业务场景的重构，已经在 6000 多个村庄辅助基层工作人员提升效率。截至 2024 年 7 月，累计大模型调用量突破 1 亿，日均超过 200 万次。近 3 万名基层政务工作人员正在用大模型服务超千万居民，实现对居民需求的快速精准反馈处理。

再如，在企业运营领域，比亚迪、沙宣、洋河等企业积极打造企业官方智能体，既作为企业品牌顾问，也承担销售和客服的职责，能够为用户提供一对一专业服务，大幅提升互动营销效率。比亚迪官方智能体能够针对用户个性化问题，基于大模型的理解、生成、逻辑、记忆四大能力和检索增强能力，快速从网上获取最新、最准确的消息，给出正确答案，并通过多轮互动、沟通，挖掘满足用户购车需求，节省用户对比时间。数据显示，比亚迪官方智能体上线后，销售线索转化率提升了 119%。

²⁷ 来源：中国工信新闻网，https://www.cnii.com.cn/rmydb/202411/t20241114_615852.html，2024 年 11 月 14 日。

第五章 生成式人工智能发展的难点 与建议

一、发展难点

(一) 产业发展面临算力瓶颈

我国发展生成式人工智能相关产业面临的算力瓶颈，主要体现在算力供给、算力分布和算力管理调度等方面：**一是高性能芯片进口受限，国产人工智能芯片仍未达到顶尖水平。**随着全球科技竞争加剧，一些国家出于自身利益考虑，对我国实施严格的人工智能芯片出口管制措施，阻碍了人工智能相关产业的发展。近年，国内已经涌现出一些自主可控的算力芯片产品，但这些产品在计算性能、通用性方面仍较全球领先水平存在一定差距。**二是算力基础设施分布不均衡。**生成式人工智能发展所需的高端芯片价格昂贵，且海量算力需求使得数据中心能耗极高，使得产业发展主要集中于经济发达地区。从地域上看，东部沿海算力资源较为集中，中西部则相对匮乏；从企业规模上看，大型企业资金充足，因此往往可以购买优质的算力资源，而中小企业由于难以承担高昂的硬件成本，导致其算力资源相对短缺。**三是模型应用面临算力管理难题。**随着应用的爆发式增长，模型的训练和推理需求也在大幅提升。由于大模型需要强大的计算能力，芯片规模增加使得运维管理的复杂性显著提升。硬件在使用过程中不可避免地会发生故障，而随着规模的扩大，故障发生的概率也随之增加²⁸。如果无法有效解决算力管理问题，训练成本和效率等一系列挑战将随之而来，极大地减缓企业在生成式人工智能应用方面的进程。

(二) 高质量数据集相对匮乏

我国发展生成式人工智能相关产业面临的数据集质量问题，主要体现在以下几个方面：**一是高质量数据集相对缺乏。**尽管我国已有部分中文开源数据集，但数据完整性不足、清洗

²⁸ 美国科技公司 Meta 在一份研究报告中表示，其训练 Llama 3 405B 参数模型在 16384 个 Nvidia H100 GPU 集群上运行，54 天内发生了 419 次意外故障，平均每三个小时就会发生一次故障。

和标注不精确等问题普遍存在，导致数据质量相比国际英文公开数据集存在较大差距。由于国内数据集质量参差不齐、部分内容陈旧，导致国产大模型的训练效果并不十分理想。二是**数据产业生态有待健全**。我国数据要素市场发展还处于初级阶段，数据流通规则和供需对接机制尚未完全建立，因而未能形成高效完整的人工智能数据产品供应链，企业获取数据资源的成本高、难度大。三是**企业间数据孤岛问题依然存在**。数据作为新型生产要素虽然已得到国家层面的重视，但在企业实际运营过程中往往不易落实。很多企业出于自身利益的考虑，并不乐于将生产经营数据对社会开放，以免帮助竞争对手训练模型，导致专业领域的行业应用数据集很难获取，抑制了行业大模型的发展。

(三) 行业应用场景急需扩展

尽管生成式人工智能在各行业的应用前景被社会各界寄予厚望，但在实际应用过程中依然面临着各类难点。一是**部分前沿应用领域需进一步加快探索**。当前，生成式人工智能的技术发展与应用落地正处于关键时期，需要持续加大各方面支持力度，确保我国在全球竞争中的相对优势。如自动驾驶领域，作为人工智能的重要应用场景之一，正在以欧美为代表的海外市场高速发展。美国 Waymo、特斯拉、英国 Wayve 等企业都在各国政府的积极推动下落地自动驾驶技术应用。二是**部分企业对大模型技术本身和安全性存在顾虑**。很多企业对大模型应用存在“私有化部署才安全”的传统观念，对更易用更高效的公有云部署持有疑虑。而一味强调私有化、定制化，一方面会大幅增加企业的运维和服务成本；另一方面，“手工作坊”“施工队”的应用模式，会折损大模型应用的效率和质量，制约企业提质增效、发展新质生产力。三是**部分企业在利用大模型实现智能化转型方面经验尚浅**。目前，大模型并不总是适用于所有企业场景，需要结合行业特定数据，通过精细调优来提高输出的准确性。然而，技术掌握的不足、数据积累的不高、缺少大模型应用工具等因素，限制了其在一些企业中的有效应用。

(四) 专业人才培养任重道远

我国发展生产式人工智能相关产业面临的专业人才紧缺问题，主要体现在以下几个方面：一是**新兴技术应用导致人才缺口加大**。生成式人工智能自 2022 年底问世至今，产品持续高速迭代，使得相关专业技术人才需求不断提升。根据央视财经报道，生成式人工智能的

招聘市场在 2024 年十分火爆，第一季度相关岗位需求同比增长超 300%²⁹。二是**高校人才培养需要与时俱进**。高校在人工智能相关教育中缺乏高水平的课程建设经验，根据各学校的实际情况，可能在数据算力资源、教学实践资源、专业师资力量等方面存在不同程度的短板，从而导致学生学习的专业知识与实际就业存在一定差异。三是**未成年人的教育培养有所不足**。尽管很多未成年人对人工智能技术存在兴趣，但其获取信息的主要渠道是“自己在网上浏览相关知识”，而非来自学校的教育课程。调查数据显示，55.5%的未成年网民对人工智能技术的了解主要依靠网上自学，而通过学校老师了解人工智能技术的比例仅为 37.0%³⁰。

（五）安全伦理问题面临挑战

生成式人工智能产品的涌现，使得人工智能技术对人类的影响空前扩大，相关的安全和伦理问题已经引起社会重视。从目前看来，生成式人工智能的潜在安全和伦理风险主要在于以下三个方面：一是**虚假信息问题**。生成式人工智能自出现以来，其“幻觉”³¹问题就一直存在。尽管在过去两年时间里，生成式人工智能的“幻觉”已经通过技术手段明显减少，但考虑到信息甄别过程中的实际困难，纯粹通过技术手段消除生成式人工智能虚假信息问题的前景并不乐观。二是**版权纠纷问题**。“内容创作”是生成式人工智能的重要功能之一。但对于生成式人工智能产品而言，无论其训练过程，还是生成内容，都有可能引发知识产权的纠纷问题。尤其在文艺创作领域，已有相当多的案例涉及生成式人工智能的版权问题。三是**误用滥用问题**。生成式人工智能作为工具，不可避免地面临着误用滥用问题。这类产品不仅可能被不法分子用于造谣、欺诈，就连普通学生也可能利用这类产品抄袭作业、生成论文，从而影响正常的社会秩序。

二、相关建议

（一）推动算力芯片产业发展

高算力芯片是生成式人工智能发展的物质基础，因此要确保我国在该领域的研发和生产

²⁹ 来源：央视财经，https://content-static.cctvnews.cctv.com/snow-book/index.html?item_id=5936248638283909967&track_id=B1A8CEDD-4BED-4AD5-A781-6E0F44BA58D2_736815666270，2024 年 5 月 8 日。

³⁰ 来源：《第 5 次全国未成年人互联网使用情况调查报告》。

³¹ 生成式人工智能的“幻觉”：指生成式人工智能创造出看似合理、逻辑自治，但实际上不正确或不准确的信息。

自主可控，保障产业链供应链安全，进而实现国家安全。一是**加强技术创新与研发投入，推动国际交流合作**。从国家层面加大对国产人工智能芯片研发生产的支持力度，合理利用税收优惠、资金补贴等优惠政策激励措施，鼓励产学研用合作，推动国产芯片企业自主创新，促进产业生态不断完善。在保护国家安全和国内企业正当利益的前提下，积极寻求与国际伙伴的交流合作，积极学习、引进先进芯片生产制造技术，参与新一代人工智能芯片的研发与标准制定。二是**提升算力集群管理调度能力**。积极利用国内智能云计算企业在异构计算平台建设应用方面的优势，促进算力建设运营方与头部云计算厂商合作，支持通过“一云多芯”等方式保障不同型号芯片在统一框架下高效调度，支撑大规模算力集群建设和稳定运行。强化面向多芯片混用场景的算力保障技术，充分挖掘既有算力潜力。提升算力平台调度和机器管理能力，支持算力芯片高速互联、高效存储，解决大规模集群稳定和性能问题，给大模型训练和推理加速。提升算力系统鲁棒性³²，在少数异常事件发生时能迅速响应和恢复，快速、高效支撑大模型开发全流程的算力需求。

(二) 推动数据资源开放共享

数据资源是训练生成式人工智能产品的“养料”，因此加快国内数据生态建设、产出高质量的数据资源，是推动生成式人工智能发展的另一项重要工作。一是**建立公共数据利用制度**。要统筹发展和安全，兼顾效率与公平，提高我国公共数据资源的开发利用水平。通过建立国家级大数据平台，促进数据的合法合规共享，打破行业壁垒，实现公共数据资源的整合与开放。二是**激发产业活力**。鼓励产学研合作，共同构建标注详尽的专用数据集，特别是在医疗、金融、教育等专业领域，通过政府引导资金支持数据集建设项目，同时考虑利用众包和技术社区力量，发动技术社区用户参与数据标注，扩大数据覆盖范围。三是**保障数据安全**。完善数据隐私相关的法规制度，探索隐私计算等新兴技术的应用前景，确保数据资源在训练人工智能产品的过程中能够保护国家安全和个人隐私，增强数据收集和使用的合法性和可靠性。

(三) 推动产业引导扎实有效

为加快生成式人工智能在更多行业领域落地应用，需要政府、企业和社会各界的共同努

³² 鲁棒性：指系统在面对内部结构或外部环境的改变时，能够维持其功能稳定运行的能力。

力协作，实现科技创新、产业带动和社会价值的综合效应。一是对自动驾驶等大模型创新应用场景适度放宽准入。在车路云一体化应用试点等政策支持下，充分授权地方试点示范，引导营造包容创新的舆论环境，支持自动驾驶车辆在公共交通等领域规模化试点应用。二是鼓励公有云在生成式人工智能领域的规模化应用。推动央国企先行先试，带动更多企业参与，牵引人工智能产业发展。央国企是大模型应用的主力军，在人工智能产业发展中发挥着牵引作用。因此应通过对场景业务分类分级，精细化区隔适合采取公有云模式的大模型应用场景，让大模型在更多场景先落地应用，再逐步完善优化，带动行业企业更便捷高效应用大模型。三是打造行业标杆，提升示范效应。考虑优先推动大型企业与科技企业合力打造“龙头项目”，突出引领作用，树立业界标杆，形成更多标准化、模块化、可复制推广的先进案例和经验，让中小企业看到如何应用人工智能大模型，及应用能够产生的实际经济效益，从而主动拥抱智能化升级改造。

(四) 推动专业人才教育培养

“功以才成，业由才广”，人才是发展生成式人工智能的根本驱动力，因此需要持续完善“引才”“育才”“留才”机制，为数字时代夯实人才基础。一是在高校开展实用性教育。高校相关机构应紧密结合市场需求，开展实用性教育。学校应尝试开设与生成式人工智能相关的课程，加强理论与实践结合，并加强校企合作，为学生提供实习和实践的机会，使学生能够在真实的工作环境中应用所学知识，培养解决实际问题的能力。二是在中小学开展普及教育。在中小学阶段，应将人工智能的基础知识纳入素质教育体系，根据不同年龄段学生的认知特点，设计合适的课程内容，激发他们对人工智能的兴趣，培养他们对技术的初步理解和应用技能，为未来培养专业人才打下坚实基础。三是完善海外引进人才保障机制。要完善海外人才引进、回流的支持保障机制，包括提供有竞争力的薪酬待遇、科研环境、生活条件等，形成具有国际竞争力的人才制度体系。

(五) 推动安全体系持续完善

为应对生成式人工智能技术的快速发展和广泛应用，需要与时俱进，持续完善相关安全体系，并强化对公众的安全意识教育。一是动态完善安全体系。尽管我国已经出台了相关管理办法与治理框架，但生成式人工智能发展快、覆盖广、种类多，需要与时俱进，动态完善监管体系，确保技术发展符合社会伦理和价值观，不被误用滥用。二是加强安全技术研究探

索。从历史上看，技术的进步常常与安全风险相伴而生，为解决新的安全风险，总有相应的技术手段被创造出来。因此，需要持续加强对新技术的探索，开发更加先进的安全防护和检测手段，应对生成式人工智能可能出现的各种潜在威胁，提高模型算法的透明度，加强人工智能模型的可解释性，让用户和监管机构能够看到程序的决策过程，减少“黑箱”风险。三是注重公众宣传教育。对于防范生成式人工智能的潜在风险，不能仅依靠政策和平台的力量，对社会公众的宣传教育同样必不可少。要通过广泛的宣传，让用户对生成式人工智能技术形成正确的认知，帮助其树立正确的使用技能和安全防范意识，避免生成式人工智能产品的误用滥用。

本报告版权归中国互联网络信息中心（CNNIC）所有。

感谢百度发展研究中心对本报告技术等相关内容提供的素材支持。

如引用或转载，请注明来源。

如有合作需求，请致电 010-59116603 / 010-59116606。