

科技创新下传媒行业的投资机会

传媒行业深度研究

分析师：倪爽 S0910523020003

2024年6月19日



本报告仅供华金证券客户中的专业投资者参考
请仔细阅读在本报告尾部的重要法律声明

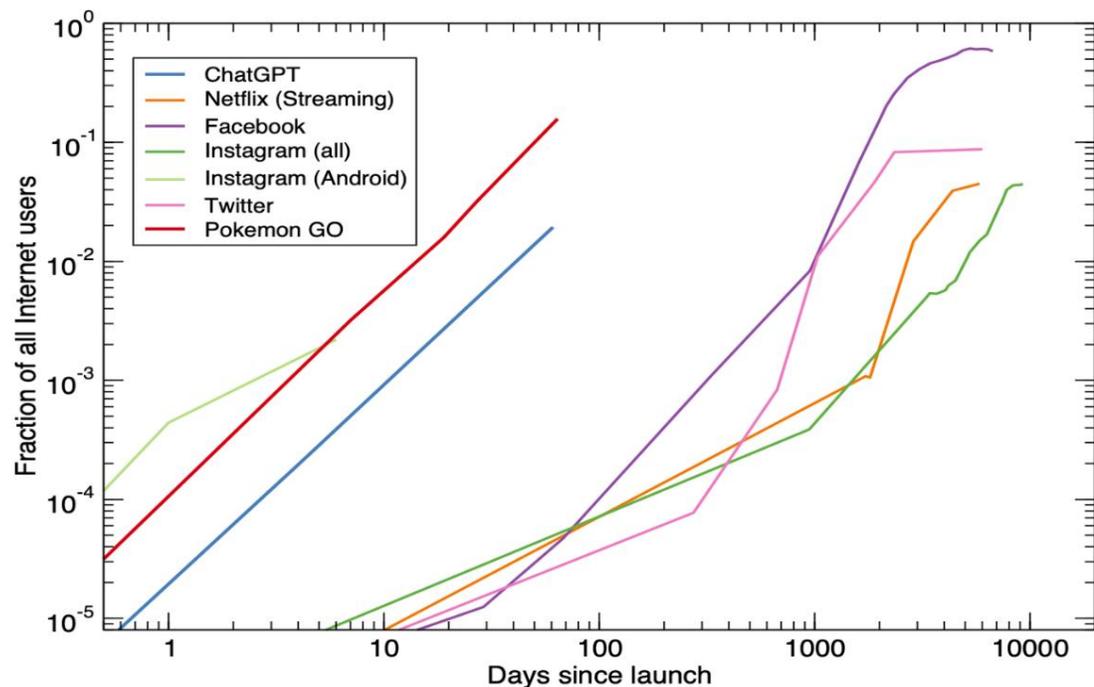
- ◆ 伴随着ChatGPT掀起的人工智能热潮，2023年迎来生成式AI元年。相比前代AI，具备高可用性、高拟人化的预训练大模型跨越技术奇点，国内外技术公司、高校、研究院的语言、图像、视频、音频大模型在2023年以极快的速度相继推出和迭代，基于预训练大模型的应用在全球范围内产生了爆炸式的影响，从社会群众到AI从业者，对人工智能技术能够带来的生产生活变革，都实现了颠覆性的再认识。
- ◆ 2022年1月，国务院发布“十四五”数字经济发展规划；2024年3月5日，李强总理代表国务院在十四届全国人大二次会议上作《政府工作报告》，《报告》提出，深入推进数字经济创新发展，开展“人工智能+”行动，打造具有国际竞争力的数字产业集群。AI作为建设数字中国，发展数字经济的核心一环，符合国家长远战略发展方向，受到了多项政策的支持。随着数字基础设施建设不断扩容提速，以及“人工智能+”及“数据要素×”等相关行动的不断推进，将持续推进数字经济与实体经济深度融合，赋能我国经济社会高质量发展。
- ◆ 因此我们长期看好AI的科技发展方向，以及国家战略方向的一致性。传媒行业作为AI与AIGC落地的核心应用端，同样拥有巨大的发展机会。当前，随着AI与AIGC的落地，降本增效在产业中的效果也陆续体现。目前应用最为广泛的是营销业务，其次包括游戏、影视、短视频等内容产业也将受益，出版教育等业务中AI技术的应用也在快速普及。
- ◆ 推荐关注：昆仑万维，恺英网络，姚记科技，天娱数科，视觉中国，捷成股份，天下秀，南方传媒，中文在线，蓝色光标，浙文互联等。
- ◆ 风险提示：政策不确定性；AI等技术发展低于预期；全球宏观形势与技术阻碍的风险等。

-  **01 数字经济赋予我国AI发展强动能**
-  **02 AIGC技术持续迭代**
-  **03 AIGC落地多场景，降本增效逐步显现**
-  **04 投资建议**
-  **05 风险提示**

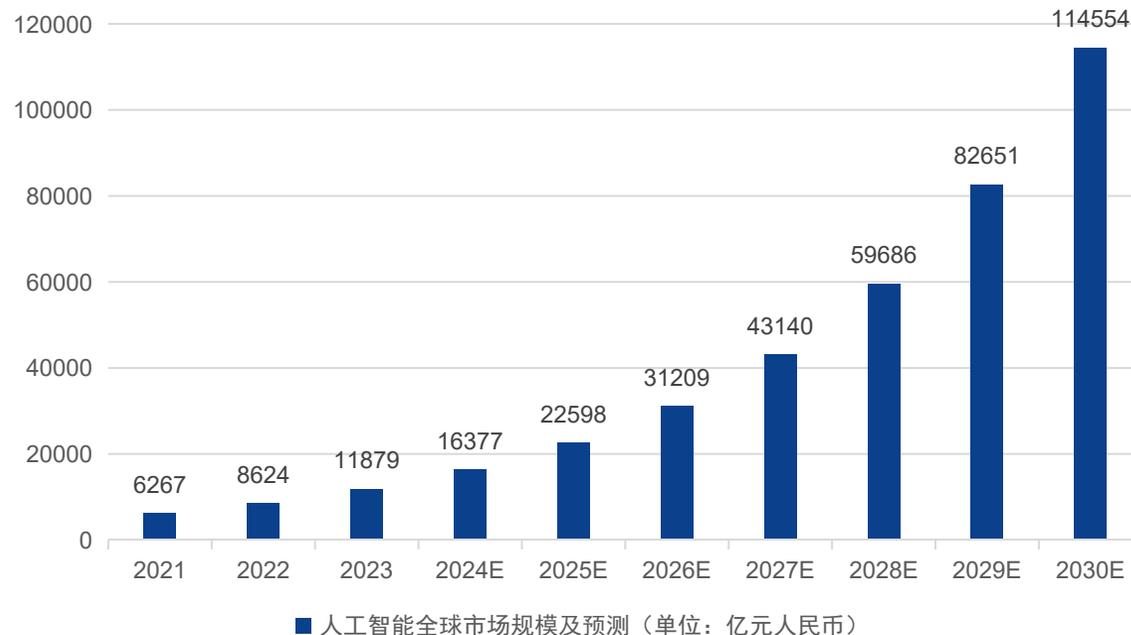
-  01 数字经济赋予我国AI发展强动能
-  02 AIGC技术持续迭代
-  03 AIGC落地多场景，降本增效逐步显现
-  04 投资建议
-  05 风险提示

- ◆ 2022年11月30日，ChatGPT-3.5横空出世，2023年3月15日，OpenAI正式发布GPT-4，让人类第一次体验了与AI的自然对话。ChatGPT达到1亿用户只用了2个月，此前即便是Apple APP store也花费了2年时间才达到1亿用户。人工智能以摧枯拉朽的势头席卷整个科技圈，英伟达CEO黄仁勋认为，ChatGPT就是AI的“iPhone时刻”。
- ◆ GPT-4发布后，众多巨头纷纷入局，与OpenAI展开合作。微软宣布，新必应搜索引擎正在运行GPT-4，同时推出Microsoft 365 Copilot将微软生产力工具与AI紧密联合，大幅提升生产力效率。此外，OpenAI官网披露的合作案例还包括摩根士丹利、KhanAcademy、冰岛政府、电子支付公司Stripe、语言学习应用Duolingo、应用程序 Be My Eyes等。ChatGPT是AI易用性革新的一大步，让用户可以通过自然语言便捷地使用。ChatGPT可以被视作是AI能力的集成，还可以内嵌入各类软件，实现对生产力的革命。
- ◆ 彼时，国内的大模型应用也蓬勃发展。百度于2023年3月16日发布文心一言，吹响了国内大模型应用的号角。

ChatGPT是发展最快的互联网产品之一



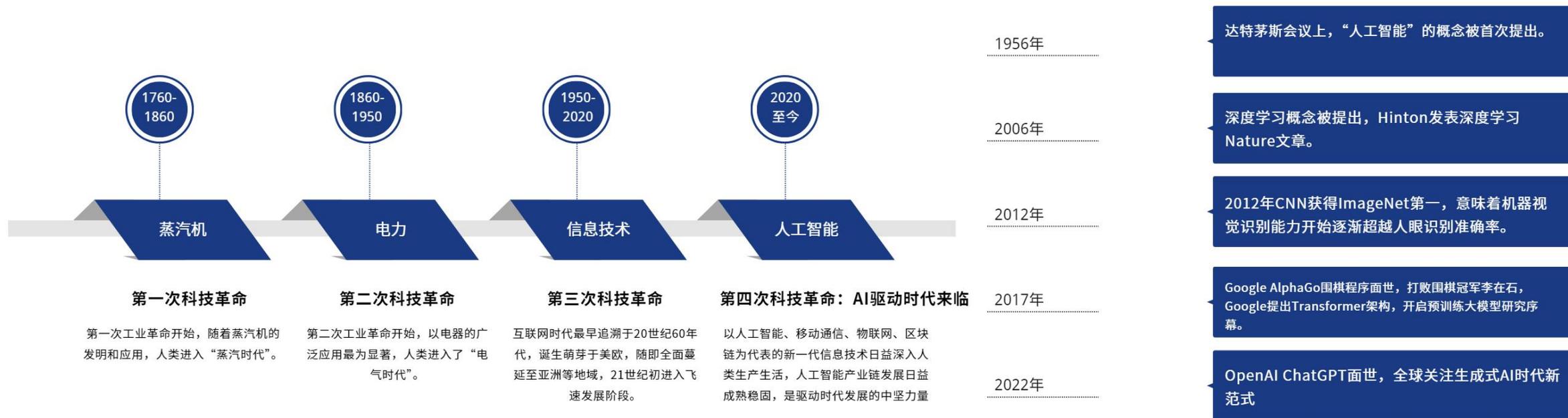
人工智能全球市场规模及预测



- ◆ AIGC (AI-Generated Content) 指利用人工智能技术 (生成式AI路径) 来生成内容的新型内容生产方式, 通常是文字、图像、音频或视频。这类内容可以通过自然语言处理, 机器学习和计算机视觉等技术生成, 即生成式AI。AI最初设立的目的是让机器像人类一样思考解决问题。目前AI的总体目的是通过各种算法解决问题提高生产效率。
- ◆ 人工智能 (AI) 伴生于信息技术时代, 经过数十年的研究积累及经验沉淀, 已逐步跨越科学与应用之间的技术鸿沟, 迎来新一轮的红利爆发与创新机遇。21世纪以来, 全球技术创新进入空前活跃期, 生成式AI技术的到来被誉为“最具革命性”的技术进步, 未来产业发展是抢占全球创新高地、重构全球创新版图、重塑全球经济结构的关键节点。

全球科技革命浪潮

全球人工智能产业发展关键节点回溯



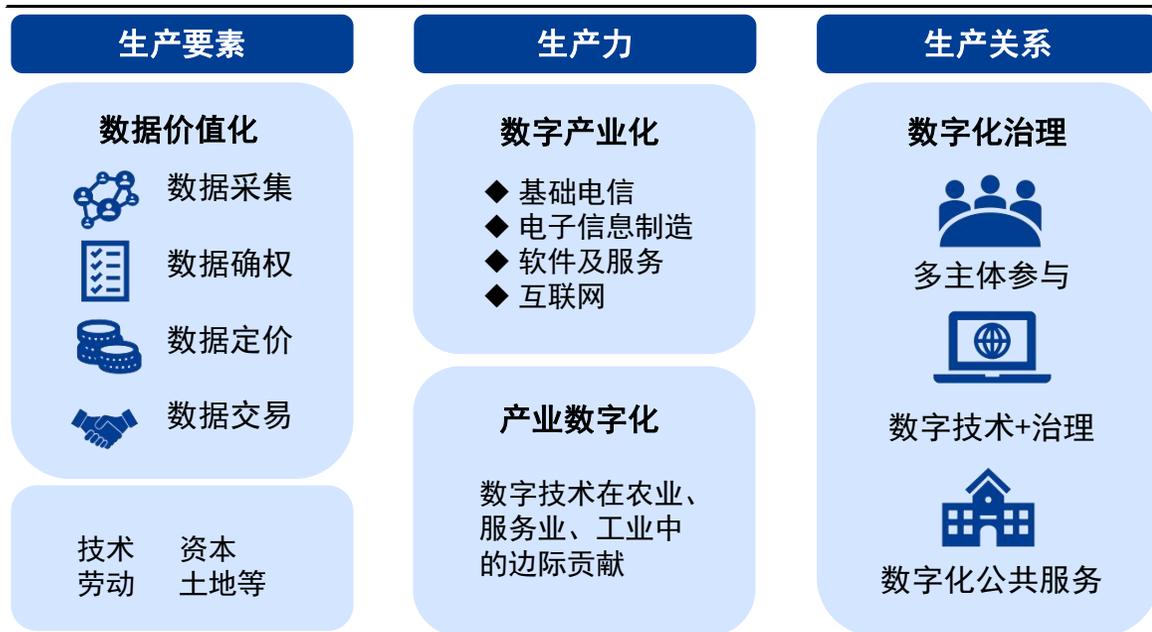
◆ 2023年迎来生成式AI元年。相比前代AI，具备高可用性、高拟人化的预训练大模型跨越技术奇点，国内外技术公司、高校、研究院的语言、图像、视频、音频大模型在2023年以极快的速度相继推出和迭代，基于预训练大模型的应用在全球范围内产生了爆炸式的影响，从社会群众到AI从业者，对人工智能技术能够带来的生产生活变革，都实现了颠覆性的再认识。

<h3>大模型小型化</h3> <p>2月25日，Meta AI在官网公开发布了 LLaMA大型语言模型，包括 7B、138、33B和65B 四种参数规模，后被国内外多款展垂类、端侧大模型作为技术底座。</p>	<h3>垂直行业化</h3> <p>3月30日，彭博社发布 BloombergGPT论文，该模型专门针对各类金融数据进行训练。</p>	<h3>端侧算力准备</h3> <p>10月25日，高通推出新一代移动芯片骁龙8Gen 3，其AI引擎支持多达100亿个参数的生成式AI模型。Hexagon NPU的性能提升了98%，持续的AI推理的每瓦性能提高了40%。</p>	<h3>合成数据再度引发关注</h3> <p>6月20日，国家互联网信息办公室发布关于发布深度合成服务备案信息的公告，公开发布境内深度合成服务算法备案信息。</p>	<h3>网络安全对抗升级</h3> <p>3月2日，美国拜登政府正式发布新版《国家网络安全战略》，提出中国是“对美国政府及私营部门网络最广泛、最活跃、最持久的威胁”。</p>
<h2>大模型技术进程</h2>	<h2>大模型落地应用进程</h2>	<h2>智算实力加强</h2>	<h2>国家动态监管</h2>	<h2>中美关系限制</h2>
<h3>多模态模型下场</h3> <p>3月15日，OpenAI发布了多模态预训练大模型 GPT-4，直接开放API。</p>	<h3>现有应用融合</h3> <p>2月7日，微软发布ChatGPT版搜索引擎New Bing，上线48/小时内获100万账户申请。</p>	<h3>英伟达持续扩大领先优势</h3> <p>11月13日，英伟达发布了目前世界最强的AI芯片H200，性能较H100提升了60%到90%，还能和H100兼容。</p>	<h3>生成式AI进入监管体系</h3> <p>7月13日，国家网信办等七部门联合公布的《生成式人工智能服务管理暂行办法》即将自2023年8月15日起施行。</p>	<h3>AI芯片出口再禁</h3> <p>10月17日，美国更新出口管制标准，英伟达针对中国市场的特供版H800、A800两款芯片也面临禁售。</p>
<h3>文生图模型快速进化</h3> <p>7月27日，Stability AI正式发布文生图模型SDXL 1.0。</p>	<h3>Agent崛起</h3> <p>11月7日，OpenAI发布了AI Agent初期形态产品GPTs，并推出了相应的制作工具GPT Builder。用户仅仅通过跟GPT Builder聊天，把想要的GPT功能描述一遍，就能生成专属GPT。</p>		<h3>L3自动驾驶迎来曙光</h3> <p>11月17日，工信部、公安部、住房和城乡建设部，交通运输部共同发布《关于开展智能网联汽车准入和上路通行试点工作的通知》。</p>	

数字经济核心产业增加值占GDP比重10%

- ◆ 2022年1月，国务院发布的“十四五”数字经济发展规划提出，到2025年数字经济核心产业增加值占GDP的比重要达到10%（2020年占比为7.8%）。2023年2月27日，国务院印发《数字中国建设整体布局规划》，提出建设数字中国是数字时代推进中国式现代化的重要引擎，是构筑国家竞争新优势的有力支撑；到2025年，数字中国建设取得重要进展；到2035年，数字化发展水平进入世界前列，数字中国建设取得重大成就。将数字经济的发展提上了一个新高度。
- ◆ 据2024年5月24日第七届数字中国建设峰会上发布的《数字中国发展报告（2023年）》，2023年我国数字经济保持稳健增长，数字经济核心产业增加值占GDP比重达到10%，提前实现“十四五”数字经济发展规划的预期目标；数据要素市场日趋活跃，2023年数据生产总量达32.85ZB，同比增长22.44%；数字基础设施不断扩容提速，算力总规模达到230EFlops，居全球第二位。
- ◆ 去年至今，产业数字化加速推进，人工智能、数字孪生等技术在90%以上的示范工厂得到应用；数据要素价值加快释放，数字产业化动能强劲。随着数字基础设施建设不断扩容提速，以及“人工智能+”、“数据要素×”相关行动的不断推进，将持续推进数字经济与实体经济深度融合，赋能我国经济社会高质量发展。

数字经济框架



数字经济相关政策

时间	政策	内容
2023年2月	《数字中国建设整体布局规划》	提出数字中国按照“2522”整体框架进行布局。
2022年1月	《“十四五”数字经济发展规划》	加快构建算力、数据等资源协同的全国一体化大数据中心体系，进一步推动算力基础设施的建设，加速打造多层次算力设施体系，实现算力水平的持续显著提升。
2021年11月	《“十四五”大数据产业发展规划》	提出，“十四五”时期，大数据产业发展要以推动高质量发展为主题，以供给侧结构性改革为主线。
2021年3月	《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》	迎接数字时代，激活数据要素潜能，推进网络强国建设，加快建设数字经济、数字社会、数字政府，以数字化转型整体驱动生产方式、生活方式和治理方式变革。
2020年3月	《工业和信息化部办公厅关于推动工业互联网加快发展的通知》	在加快新型基础设施建设方面，提出改造升级工业互联网内外网络

- ◆ 数字经济实现总量的快速增长。根据工信部公布的数据显示，中国数字经济规模已从2012年的11万亿元增长到2022年的50.2万亿元，预计2023年将达56.1万亿元，持续保持高位增长。在此背景下，2023年我国数字经济同比名义增长11.8%，高于GDP名义增速6.6个百分点，自2012年以来，连续12年显著高于GDP增速，数字经济持续发挥了经济“稳定器”与“加速器”的作用。
- ◆ 数字经济核心关键是“数实融合”。随着数字技术加速创新、快速迭代、群体突破，信息化和工业化融合从起步建设、到制造业与互联网深度融合、再到新一代信息技术与制造业融合发展，“由点成线、由线及面”向更大范围、更广领域和更深层次迈进，逐步进入以数字经济和实体经济深度融合为核心特征和重要模式的新阶段。在此阶段，Web3.0、区块链、元宇宙、AIGC等技术均在数实融合演进的道路上扮演重要角色。

数字经济规模快速增长

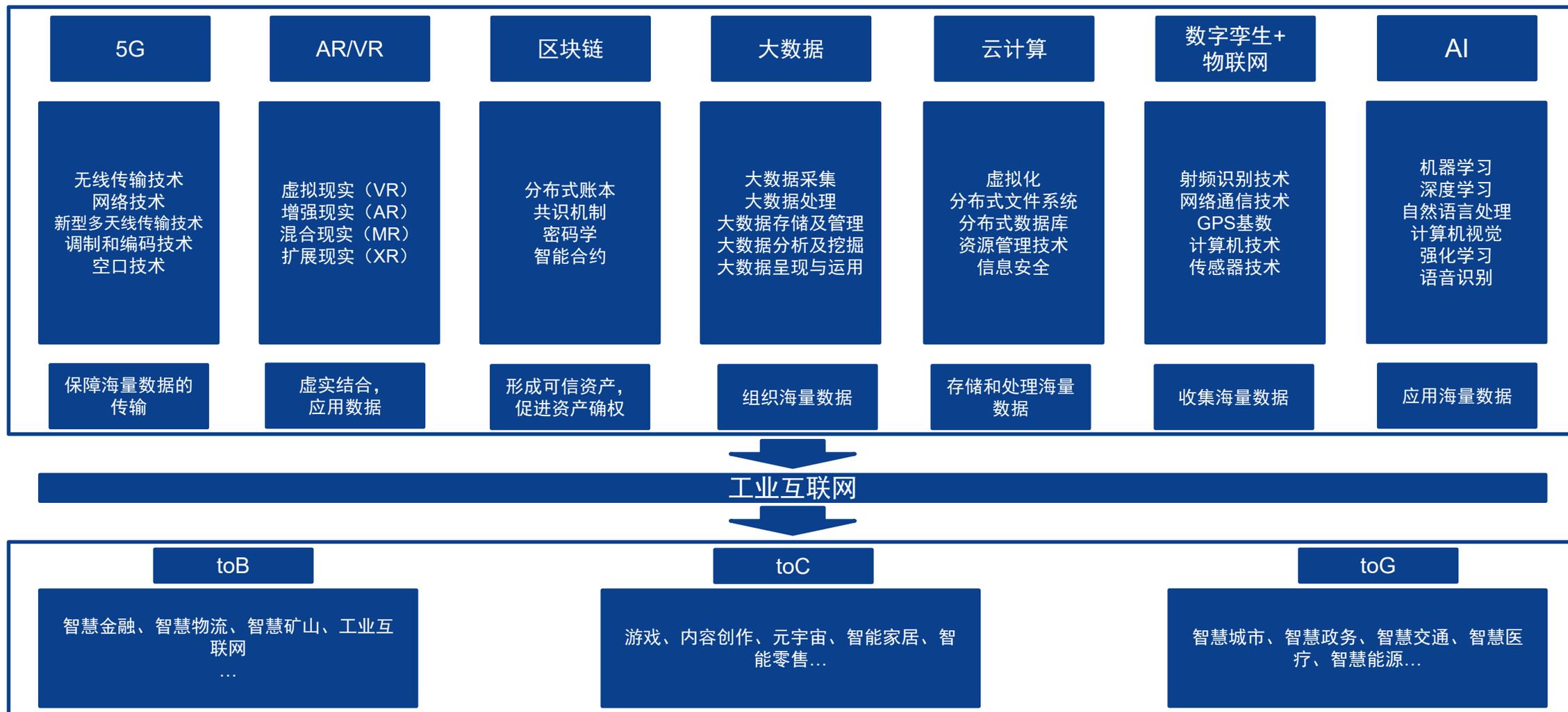


数字经济与GDP名义增速对比



数字产业化，为人工智能发展厚植土壤

数字产业化图谱



“新质生产力”深度赋能，政策助力AI发展

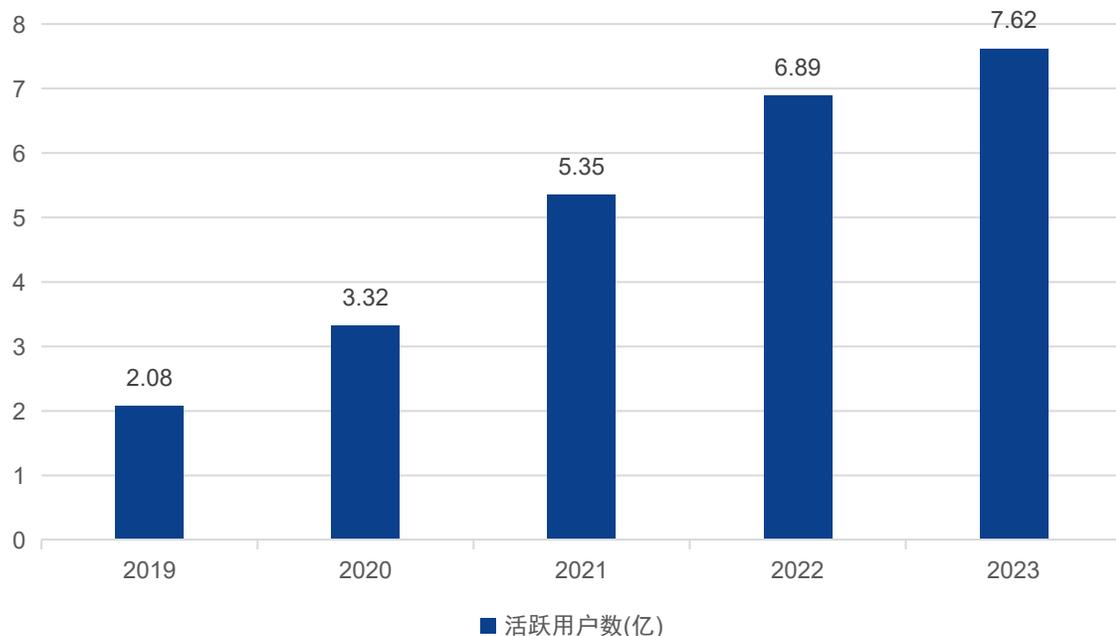
- ◆ 近年来，AI产业受到我国各级政府的高度重视和国家产业政策的重点支持。国家陆续出台了多项政策，鼓励行业发展与创新。
- ◆ 2024年3月5日，李强总理代表国务院在十四届全国人大二次会议上作《政府工作报告》，首次将“大力推进现代化产业体系建设，加快发展新质生产力”列为首要任务。《报告》提出，深入推进数字经济创新发展，开展“人工智能+”行动，打造具有国际竞争力的数字产业集群。近年来，传媒产业在AI浪潮下持续迎迭代催化。不仅在内容端获不同程度的赋能，多赛道或将持续迎来生产方式的变革。恰逢2024开年AI赛道爆发，通用模型百舸争流，AI从早期相对高度集中的toB端到toC端的下放，有望让传媒产业多赛道迎来爆发，开启AI赋能元年。

时间	发布单位	政策名称	主要内容
2023年7月	国家网信办、发改委等七部门	《生成式人工智能服务管理暂行办法》	明确了AIGC服务提供者应当依法开展的活动、遵守的规定，以及具体的服务规范，落实了监督检查和法律责任等。
2023年4月	国家互联网信息办公室	《生成式人工智能服务管理办法(征求意见稿)》	首次明确了AIGC“提供者”内容生产、数据保护、隐私安全等方面的法定责任及法律依据，确立了人工智能产品的安全评估规定及管理办法。
2022年8月	科技部	《关于支持建设新一代人工智能示范应用场景的通知》	坚持面向世界科技前沿、面向经济主战场、面向国家重大需求、面向人民生命健康，充分发挥AI赋能经济社会发展的作用，围绕构建全链条、全过程的AI行业应用生态，支持一批基础较好的AI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
2022年7月	科技部、教育部、工信部等	《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》	场景创新成为AI技术升级、产业增长的新路径，场景创新成果持续涌现，推动新一代AI发展上水平。鼓励在制造、农业、物流、金融、商务、家居等重点行业深入挖掘AI技术应用场景，促进智能经济高端高效发展。
2021年12月	工信部等	《“十四五”智能制造发展规划》	加强关键核心技术攻关。聚焦设计、生产、管理、服务等制造全过程，突破设计仿真、混合建模、协同优化等基础技术，开发应用增材制造、超精密加工等先进工艺技术，攻克智能感知、人机协作、供应链协同等共性技术，研发AI、5G、大数据、边缘计算等在工业领域的适用性技术。
2021年11月	工信部	《“十四五”软件和信息技术服务业发展规划》	支持AI算法库、工具集等研发。加快发展新型机器学习、生物特征识别自然语言理解、新型人机交互、智能控制与决策等产品和服务。推动AI开放平台建设。

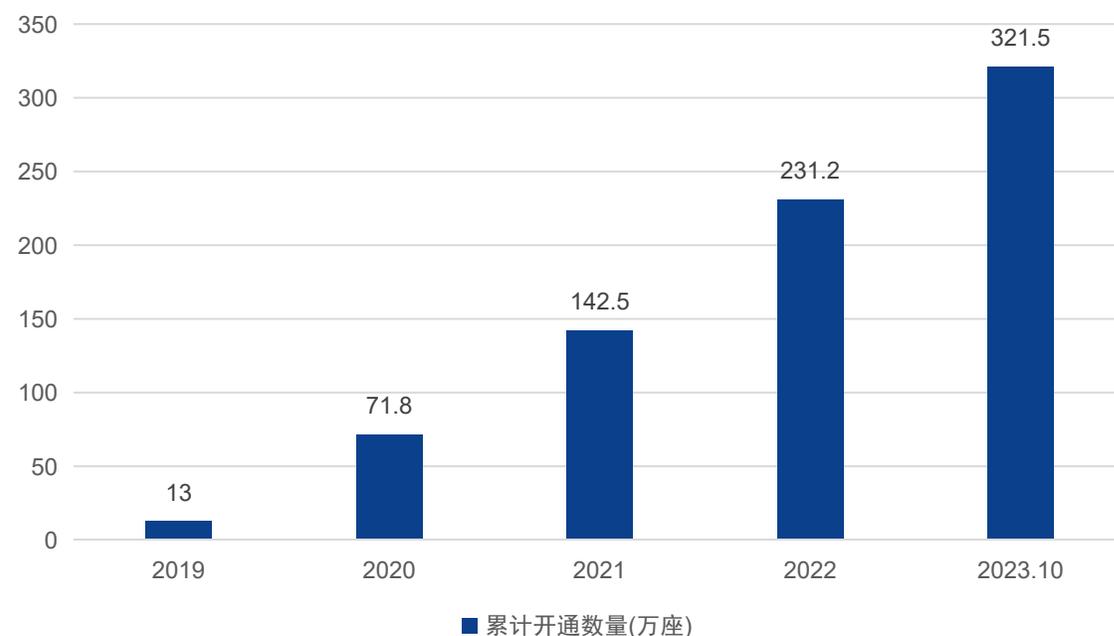
互联网生态多元化发展，多线并进催化AI应用落地

- ◆ 2023年，我国数字化、网络化、智能化发展日新月异，不断夯实数字底座，持续提升服务质量，有力推动互联网普及率增长。截至2023年12月，我国IPv6活跃用户数达7.62亿，同比增长10.60%。
- ◆ 5G方面，据工信部，截至2023年10月，中国已建成5G基站321.5万座，较2022年底增幅高达39.1%。5G网络已覆盖所有地级市城区、县城城区，超90%的5G基站实现共建共享，5G网络正加快向集约高效、绿色低碳发展。
- ◆ 终端方面，据TDIA统计，截至2023年9月，全球5G终端达到2916款，非手机终端1449款，占比超49.6%，5G终端呈现款型多样化发展趋势。
- ◆ 在网民规模持续提升、网络接入环境日益多元、企业数字化进程不断加速的宏观环境下，AIGC技术作为新型内容生产方式，有望渗透人类生产生活，为千行百业带来颠覆变革，开辟人类生产交互新纪元。

2019年-2023年IPv6互联网活跃用户数



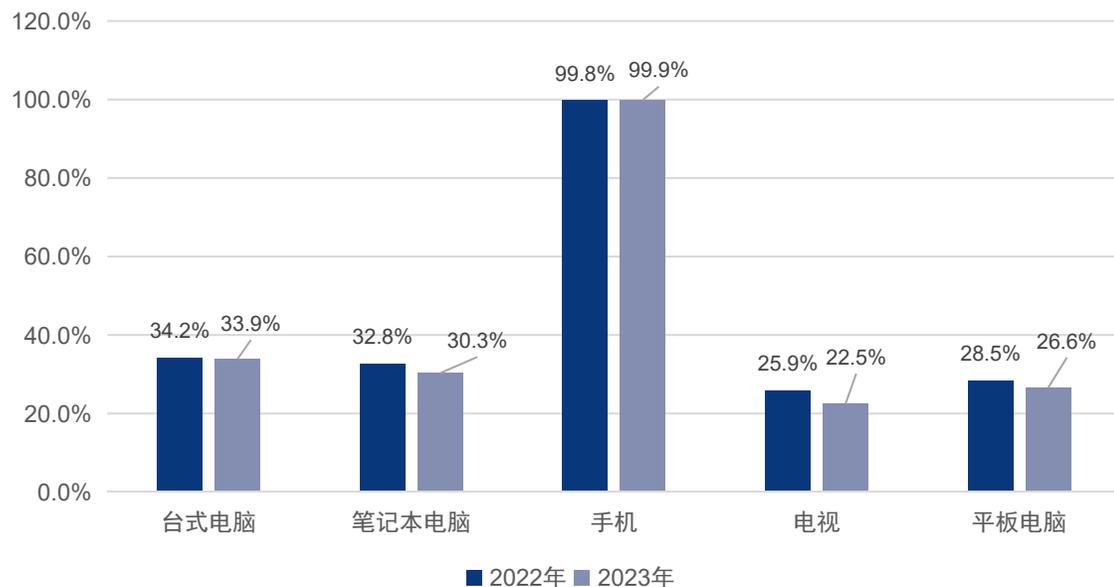
2019年-2023年10月中国5G基站累计开通数量(万座)



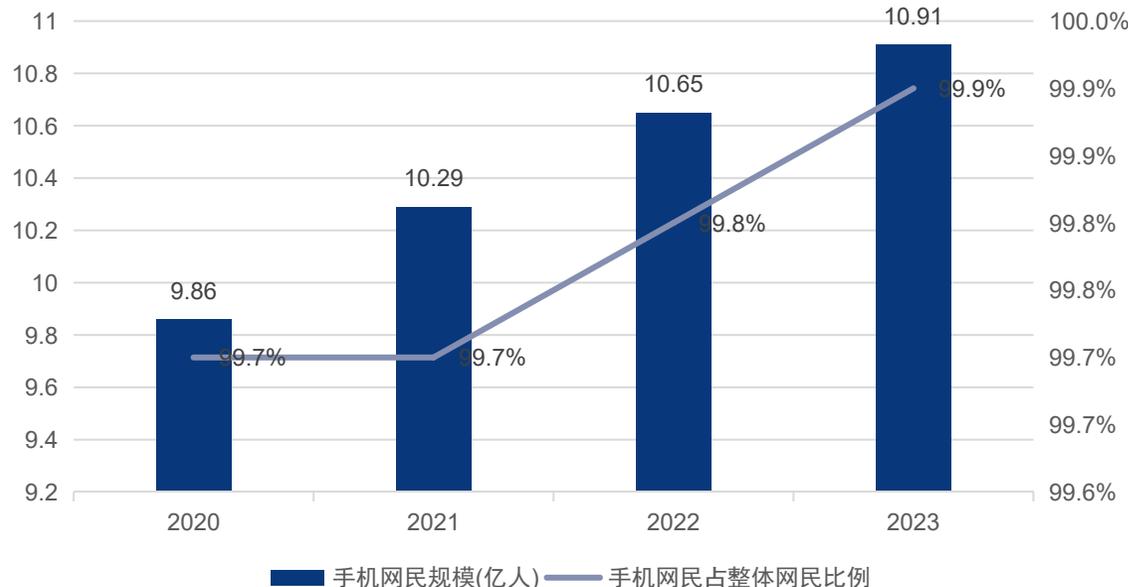
移动互联日渐普及，端侧AI蓄势待发

- ◆ 2023年，我国网络基础设施建设持续加强，通信服务质量深度优化，智慧生活提升了网民上网幸福感，数字适老普惠范围亦得到不断拓展。
- ◆ 上网设备方面，截至2023年12月，我国网民使用手机上网的比例高达99.9%；使用台式电脑、笔记本电脑、电视和平板电脑上网的比例分别为33.9%、30.3%、22.5%和26.6%。
- ◆ 2023年全年，国内市场手机总体出货量为2.89亿部，同比增长6.5%。其中，5G手机出货量为2.4亿部，同比增长11.9%，占同期手机出货量的82.8%。
- ◆ 截至2023年12月，我国手机网民规模达10.91亿人，较2022年12月增长2562万人，网民中使用手机上网的比例为99.9%。
- ◆ 端侧是AI模型应用不容忽视的重要场景。越来越多的AI计算和推理工作负载在手机、笔记本电脑、XR头显、汽车和其他边缘终端上运行，已知三星、Vivo、华为等手机厂商已在端侧布局AI；苹果亦在WWDC2024宣布推出端侧Apple Intelligence，携手OpenAI接入GPT-4o，以多模态能力赋能复杂任务处理。移动终端日渐普及，手机厂商先后入局端侧AI，有望双轮驱动长效赋能端侧“生产力”，加速AI应用落地。

2022-2023互联网络接入设备使用情况



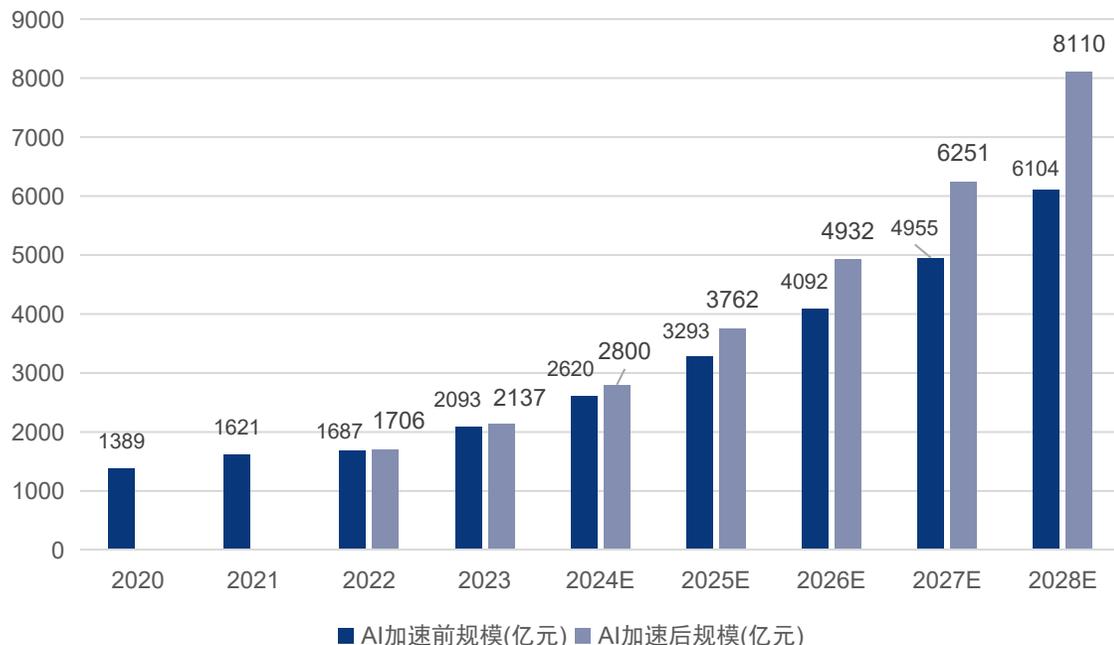
2020-2023手机网民规模及其占网民比例



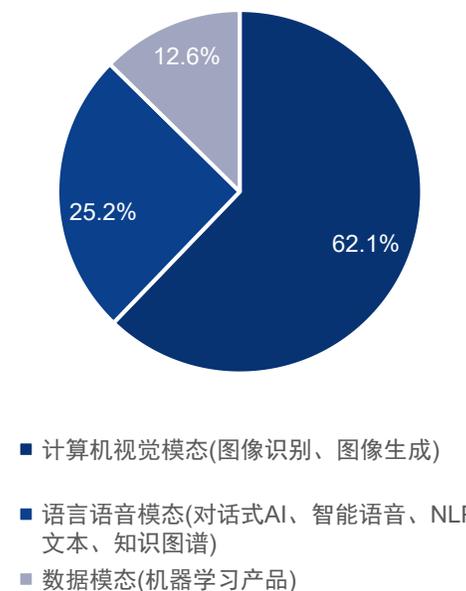
大模型涌现，赋能AI市场规模加速扩充

- ◆ 据艾瑞咨询研究院，2023年中国人工智能产业规模已达到2137亿元，大模型带来的底层技术革新将为中国人工智能产业的规模增长带来更多存量扩张与增量空间。2028年，中国人工智能产业规模将达到8110亿元。对比原本大模型未出现涌现能力的人工智能产业规模值，大模型带来的产业加成比例在2028年或达到32.9%，在语言语音模态规模加成最为显著。
- ◆ 未来大语言模型、语音大模型的产品门槛与应用成本或将逐步降低，带来更多API能力调用与产品解决方案的AI能力融合发展，尤其在2024年以后，更多AI产品逐步变现、AI能力下放至边缘侧与端侧之后的影响将更为明显；原本以图像识别为主的计算机视觉市场增长变缓，受政策及政府预算影响，泛安防类的业务增长更多被医疗、工业等CV产品取代，且图像生成市场将在未来3-5年迎来更多商业变现机会，进一步填充计算机视觉模态的市场空间驱动力。

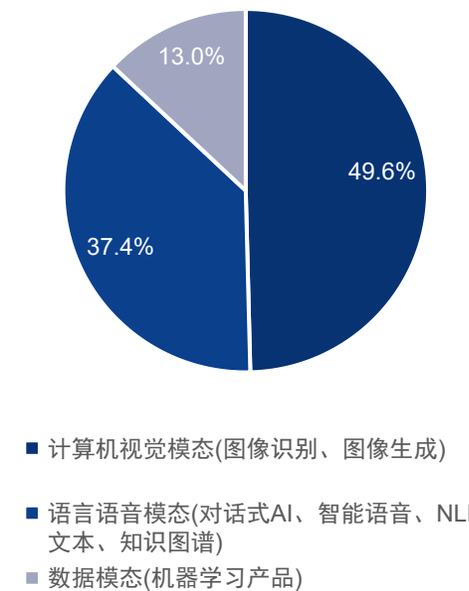
2020-2028年中国人工智能产业规模



2023中国人工智能产业模态分布



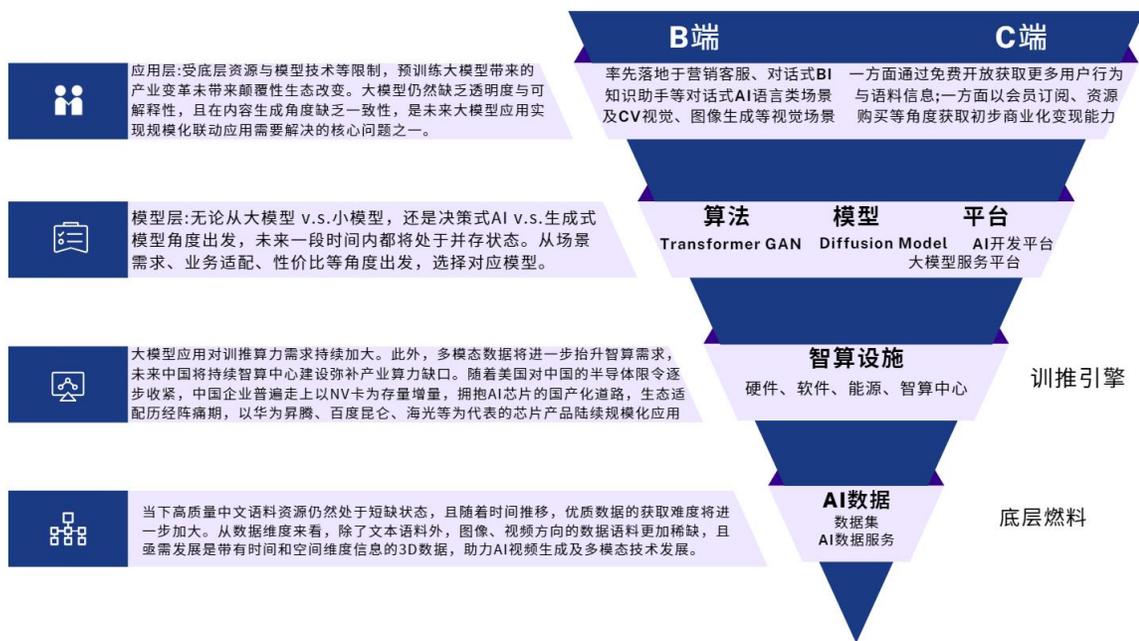
2028中国人工智能产业模态分布



2023年中国AI产业发展及演进形态

- ◆ B端热门产品特征：企业路径一般为内建产品应用，通常先打磨好服务方案，再对外输出。
- ◆ C端热门产品特征：收集用户行为及市场偏好，探索进一步产品演进及变现空间。
- ◆ 基于扩散模型的AI电商设计图、基于大语言模型的对话式AI解决方案等热门B端产品率先出圈变现，C端应用也涌现了一批AI搜索对话、图像生成等APP产品。2023年7月，妙鸭相机小程序横空出世，成为AIGC时代下第一个C端爆款。
- ◆ 中国人工智能产业发展呈现“倒三角”特征，数据层的规模质量、算力层的规模性能制约着模型层及应用层的飞速发展。2023年3月，中共中央、国务院印发的《党和国家机构改革方案》对外公布，组建国家数据局。2023年10月25日，国家数据局挂牌成立。国家从政策监管角度积极引导数据要素市场建设，鼓励开展数据确权授权，构建数据流通体系，旨在为大模型训推提供高质量数据市场资源。
- ◆ 目前，从B端和C端实际落脚点出发，AI产品方案落地仍需严谨方法论及场景结合。

数据与算力支撑AI产业顶层发展高度



2023年AI热门产品及演进形态

	toC			toB		
热门产品	AI对话助手 / AI Agent 平台	图像创作/音乐创作网站 App	AI手机 AI PC	营销/客服 数字员工/数字人	知识库助手 / 对话式BI	办公助手/ Copilot
内容及原理	拥有广泛的知识覆盖，提供知识问答、内容创作、情感陪伴等功能，还可以基于Agent平台打造个性化功能助理	提供基于图像、音频的生成创作，为专业工作者提供效率化生产平台，同时也降低创作门槛，让更多人参与	将大模型剪裁后置于端侧，基于端侧大模型，云端结合带来软硬件生态、用户交互体验的系统性变革	将B端对话式AI产品及解决方案融入大模型能力，大幅提升交互质量及轮次，拓宽营销、人力、财务等场景	结合企业文档数据实现知识信息的大规模解析、检索、问答，同时提供数据处理分析方面的智能问答与BI呈现	提供文档生成编辑、数据分析可视化、代码生成辅助、智能办公助理等功能，大幅提升员工的办公效率及职场体验

- 01 数字经济赋予我国AI发展强动能
- 02 AIGC技术持续迭代
- 03 AIGC落地多场景，降本增效逐步显现
- 04 投资建议
- 05 风险提示

- ◆ AI技术是现阶段互联网及各行业迈向新阶段的源动力。人工智能或成为推动第四次工业革命的核心力量；近年来，数据要素正成为经济增长新的主导性要素，而人工智能能够有效实现该要素的功能倍增，为生产力带来更为显著的效能提升。
- ◆ 各种关键技术的不断积累和突破，为人工智能的发展添砖加瓦，最终促成了大语言模型的巨大成功和AIGC领域的蓬勃兴起。AIGC多样化的内容生成能力使其覆盖各类内容形式，各类应用场景正随技术进步逐渐落地。
- ◆ AIGC不仅可覆盖文本、音频、图像、视频等基本内容模态，还可综合图像、视频、文本进行跨模态生成，并应用于各类细分行业成为具体的生产力要素，例如游戏行业中的AI、NPC、虚拟人的视频制作与生成等。

四次工业革命



蒸汽革命
核心：蒸汽动力

蒸汽动力替代重复劳动，生产方式机械化，开启城市化进程

电气革命
核心：电力 内燃机

加速工业化进程，带来大规模的工业生产和社会生产

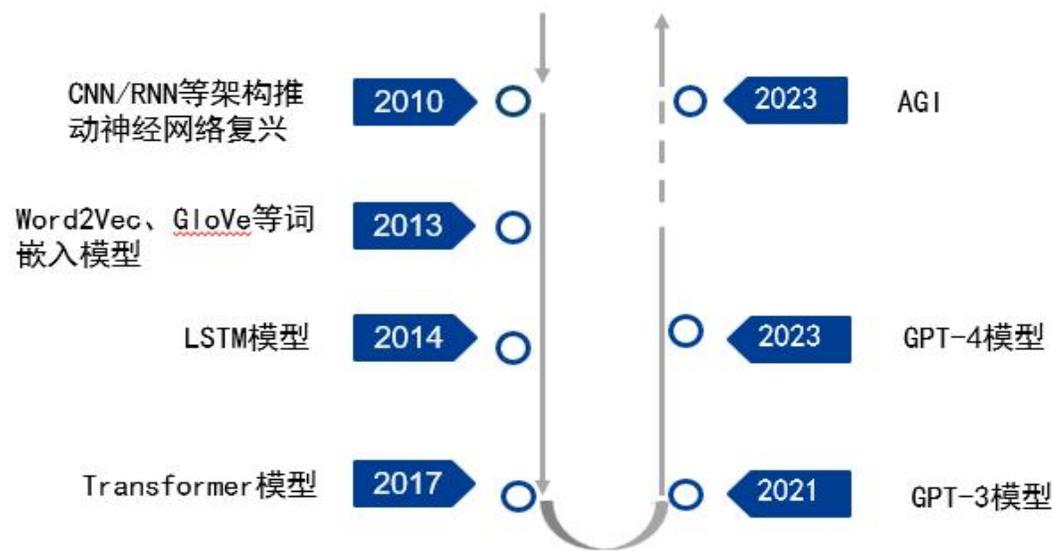
信息技术革命
核心：信息技术 电子计算机

信息传播速度加快，全球经济一体化，传统工业向信息经济转变

人工智能革命
核心：人工智能 互联网 云计算 大数据

进一步解放人类生产力和创造力，智能机器人进一步提高生产效率

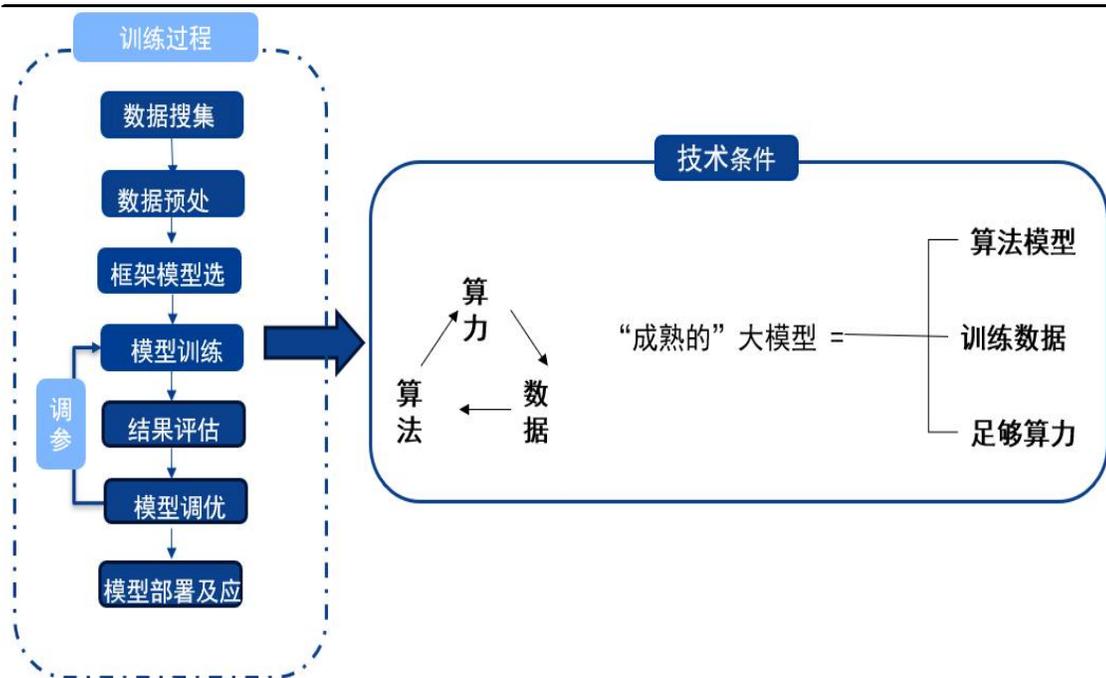
人工智能相关技术发展



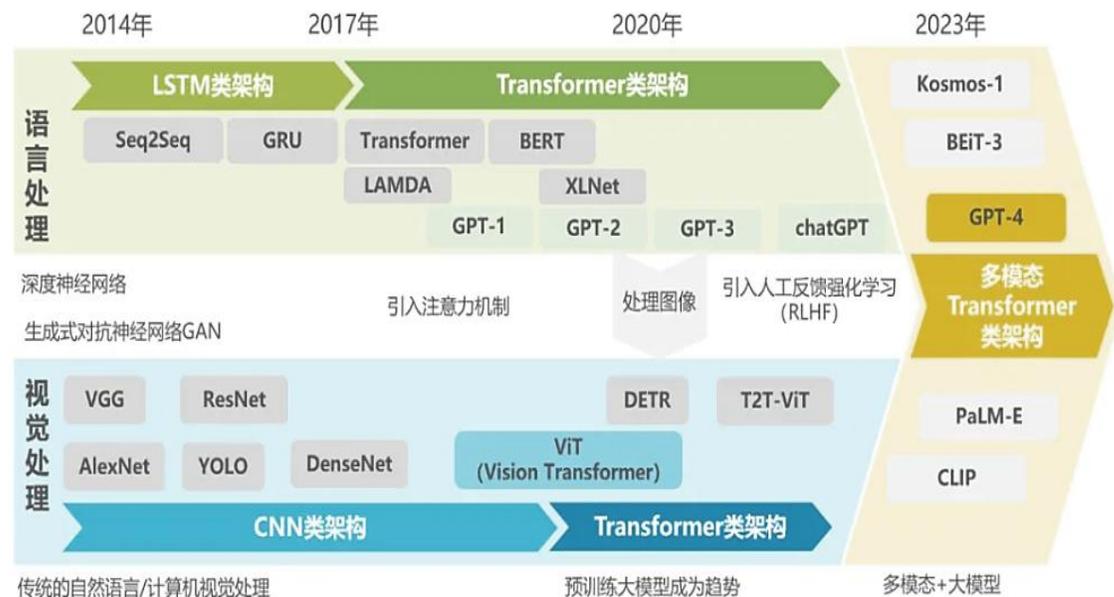
模型技术范式逐步确立

- ◆ 多模态架构可以建立统一的、跨场景/任务的模型，将不同类型的数据（如图像、文本、音频等）融合起来，极大地拓展了人工智能认知并理解世界的的能力边界。多模态学习在不同模态间搭建了桥梁，使得基础模型通过迁移学习和规模涌现达到能力跃迁成为可能，极大加速了通用模型的演进。
- ◆ Transformer类架构的发展带动多模态融合，为范式转变奠定基础。多模态架构逐渐成为图像生成领域的研究热点。
 - 在自然语言处理领域的机器翻译任务中，已经成为主流的模型架构。
 - 图像生成领域逐渐被应用，该架构可以有效地捕捉图像的全局结构和语义信息，同时可以生成高分辨率、逼真的图像。

大模型训练过程及部分条件

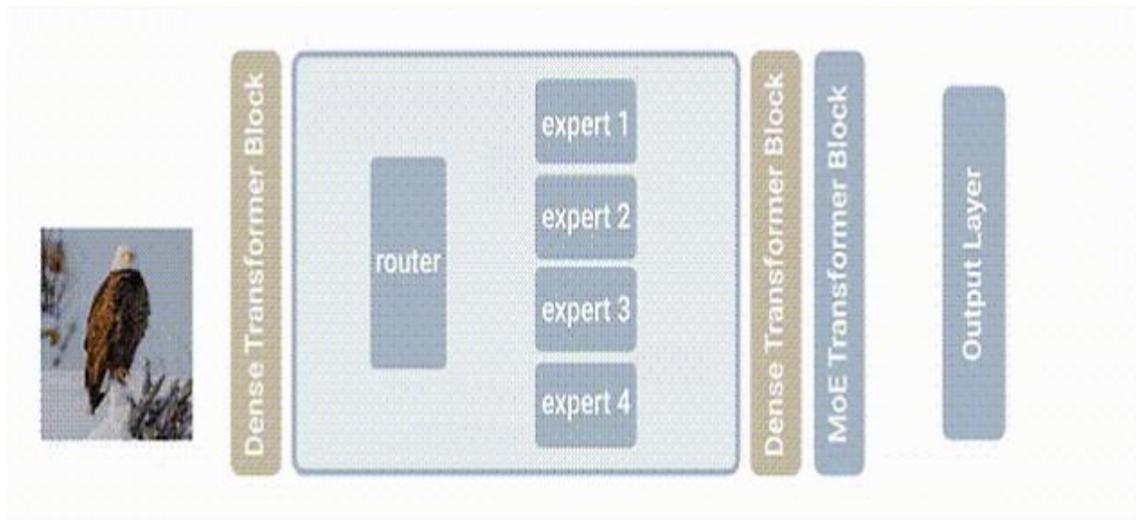


Transformer类结构发展

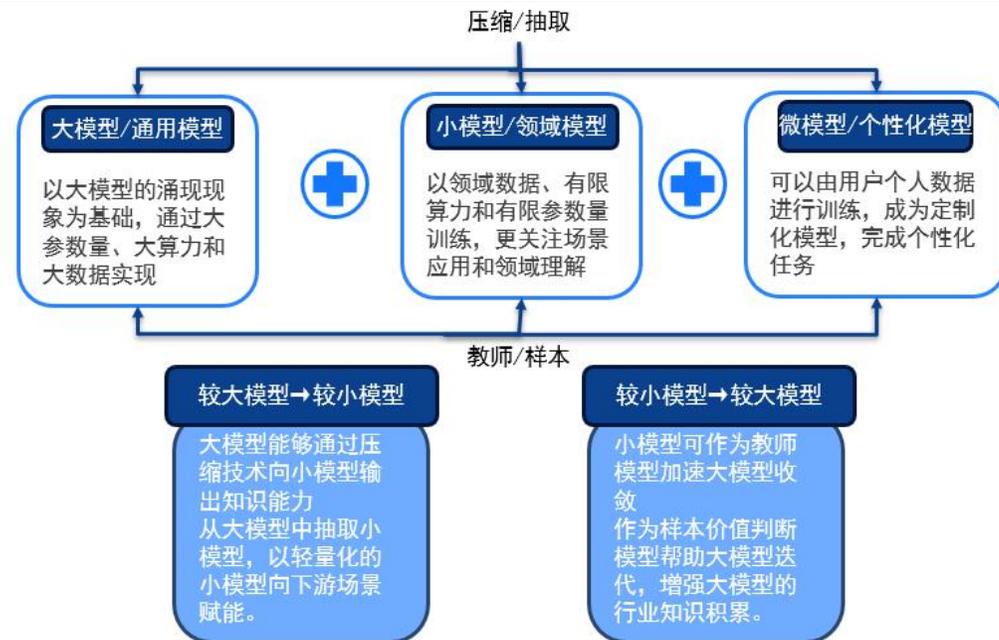


- ◆ 大模型发展路径多元化，正向推动生态建立。MoE模型（专家混合）由多个子模型（即专家）组成，每个子模型为一个局部模型，处理输入空间特定子集。MoE使用门控网络基于数据适配特定模型降低不同类型样本之间干扰。
- ◆ Transformer模型相比MoE模型不具备循环结构，使得训练AI大模型需要大量算力及时间。以GPT为例，其Transformer解码器训练过程中每天约消耗超50万度电力，训练成本以万亿美元为单位计数。
- ◆ 短期内通用类大模型市场热度较高，长期内多种模型组合是未来发展方向。根据模型体量，模型可分为大模型、小模型、微模型。
 - 小模型：多专注于垂类领域，垂直场景具备充足的数据以及问题处理能力。
 - 微模型：更加个性化，由用户个人数据训练而成。

MoE模型特点



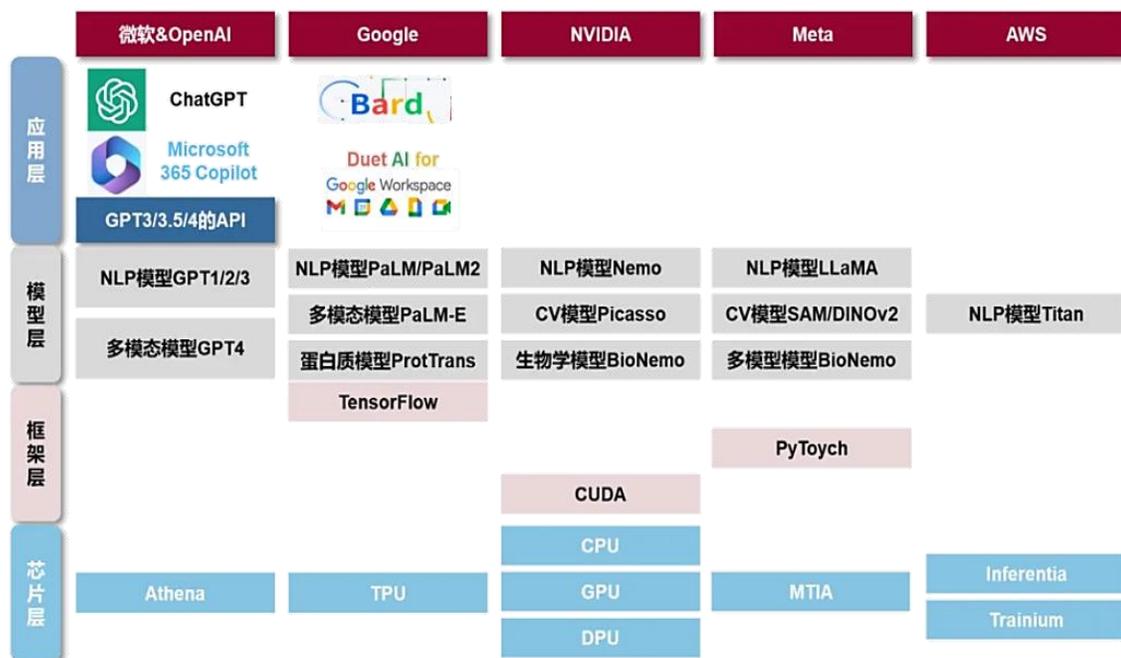
可能的AIGC模型格局及各方关系



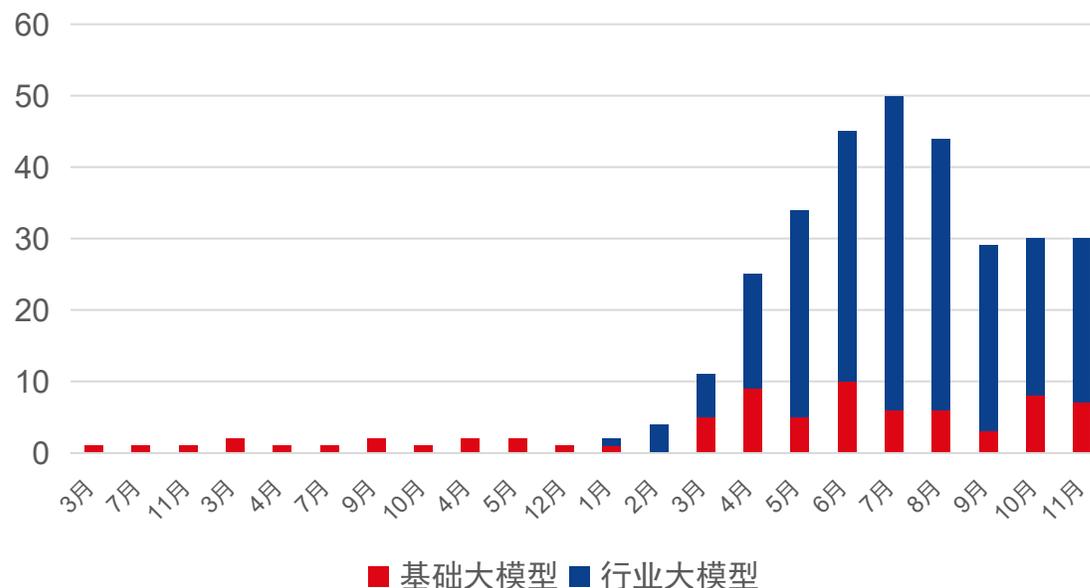
模型厂商建立生态框架

- ◆ AI生态加速发展，大模型技术逐步收敛，生态走向聚合，模型更收敛、框架更归一。
- ◆ 根据AI PORT数据显示，全球共有11家公司开发不止一种类型的AI模型。Stability AI为五种不同的AI模型类型（图像、视频、音频、3D 和代码）领先，其次为 OpenAI（聊天机器人、音频、视频和多模态）和 Google（文本、图像、音频和多模态），两者均有有四种模型类型。微软、Meta、腾讯、百度和Yandex等公司开发两到三种不同AI模型。
- ◆ 中国自2023年5月起，行业大模型发布数量显著增加，互联网巨头达成进一步行业分化及产业伙伴合作，同样受益于开源生态的建设，更多垂类厂商结合开源模型研发契合自身业务的行业大模型产品。根据艾瑞不完全统计，截止2023年底，中国行业模型的个数占比已经超过8成。

部分海外企业AI布局情况



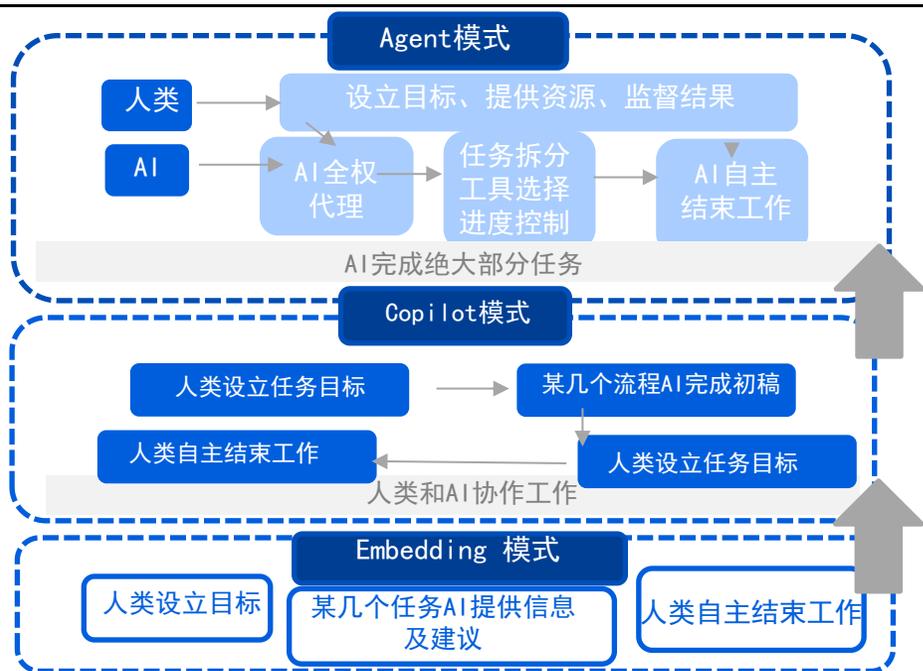
2019-2023年中国大模型发布情况



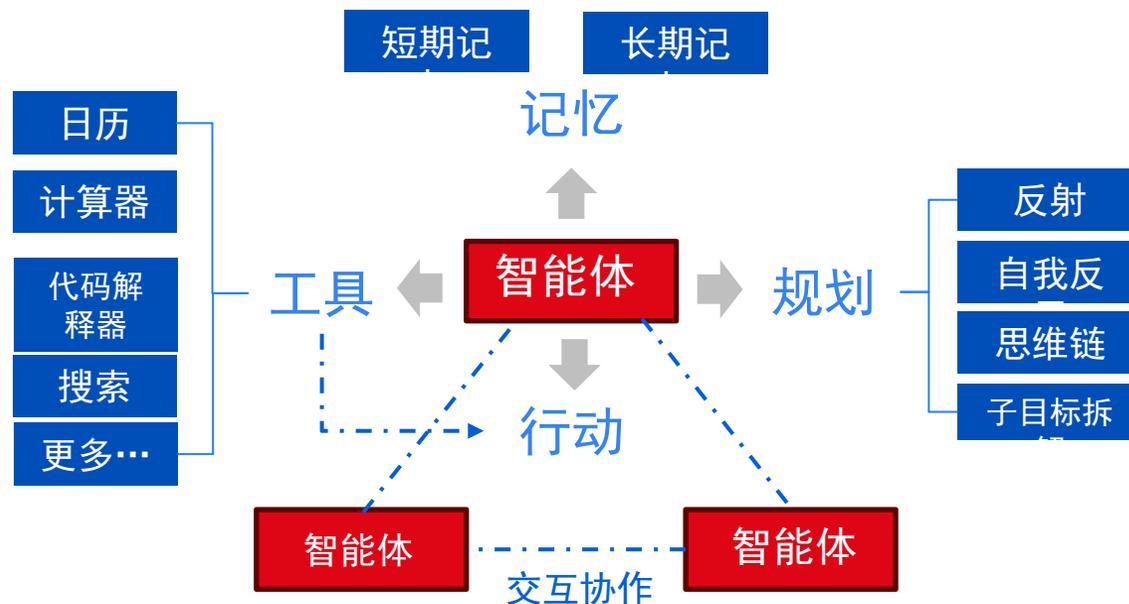
AI 产品形态发生演变

- ◆ AI产品形态向内容生成、智能助手以及数字代理等方向演进。随着大模型技术的成熟，规模增大，大模型为AI Agent提供强大能力。
- ◆ Agent在人工智能领域被赋予具有自主性、反应性、交互性等特征的智能“代理”。大型语言模型(LLMs)的出现为智能代理的进一步发展带来契机。
- ◆ AI Agent具有更广泛的应用范围可处理多个任务，并在不同领域中执行各种功能；具有更自然和灵活的交互方式，能够理解复杂的自然语言指令，与用户进行更智能对话。Agent+大模型将有望构建具备自主思考、决策和执行能力的智能体，进一步提升大模型的应用能力。

人类与AI 协作三种模式



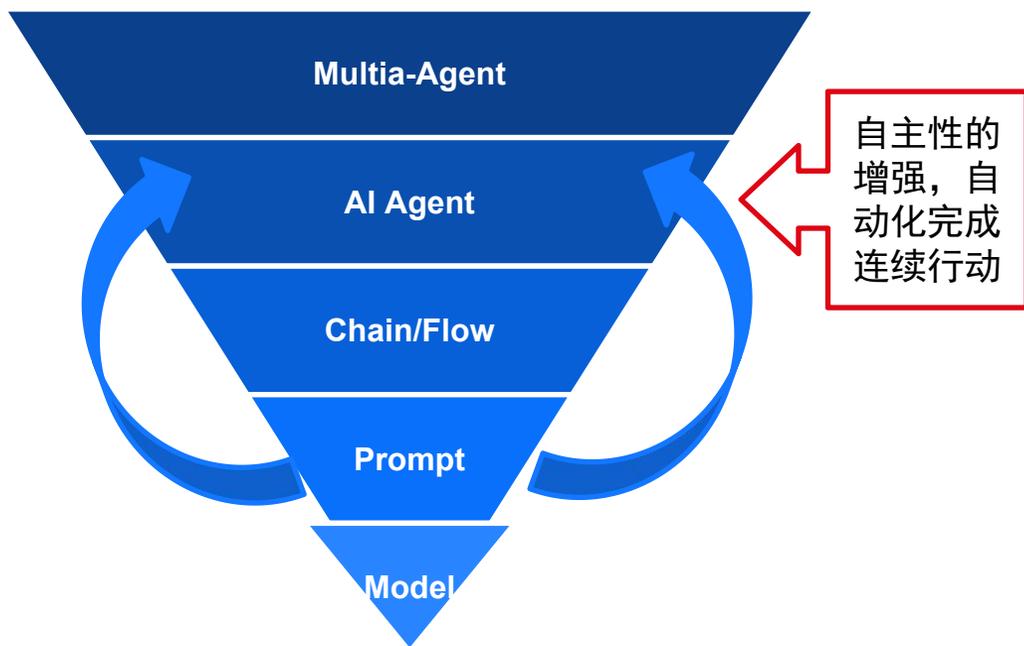
LLM驱动Agent基本框架



AI Agent或将引领产业变化

- ◆ 对比AI与人类的交互模式，AI Agent较目前广泛使用的Copilot模式更加的独立，能够自主调用资源完成任务，人类在其中起到督促和评估的作用。能否提供个性化、连贯的交互体验，为AI Agent的构建关键。
- ◆ 目前，AI Agent可以视为在LLM技术下Prompt工程的进化。AI Agent是Prompt工程的升级，Agent的核心在于自主性的增强，可有效完成某一个工作点或工作单元，尽量减少人的干预。
- ◆ 随着大模型参数量提升将提高AI Agent的理解力和泛化能力，使其能更好地处理多种任务和上下文信息。将逐步增强AI代理的自然语言处理能力。AI Agent将给应用程序的形态和业态带来颠覆性变化，成为垂直应用的超级入口。通过重构过往软件形态实现商业场景的落地。

AI Agent的升级



workflow 变化

LLM-based 代理

非代理 workflow

从头开始一次性输出一篇关于X主题的论文，不使用退格键。

开始 → 完成
文章

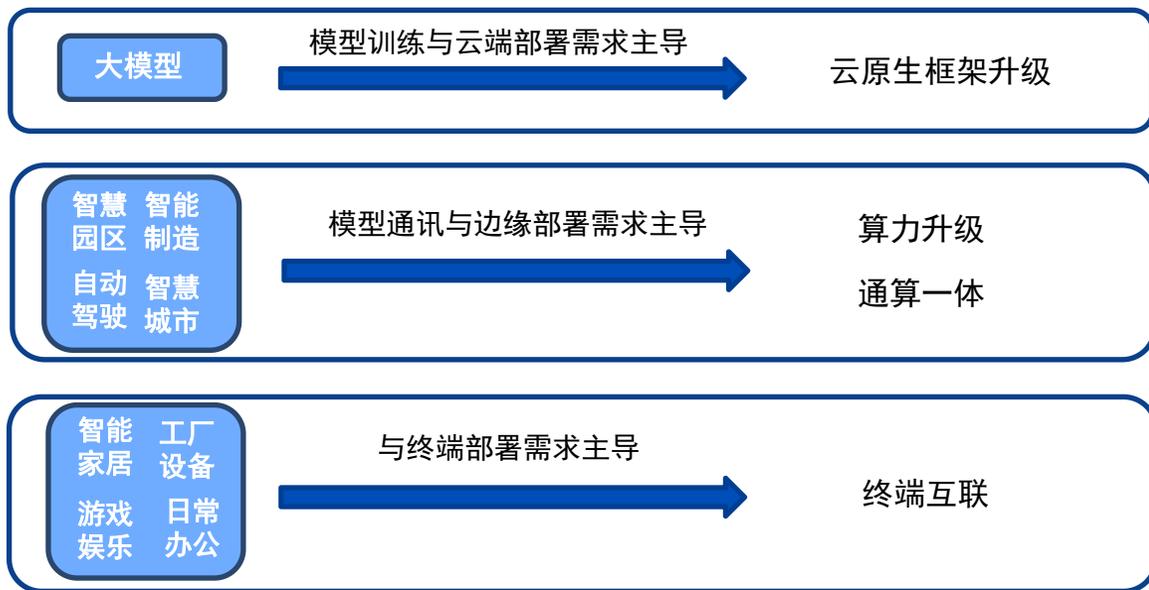
代理 workflow

- 给出关于X主题的大纲
- 你需要网站资源吗？
- 写一个初稿
- 考虑哪些部分需要修改或更多的研究
- 修改初稿

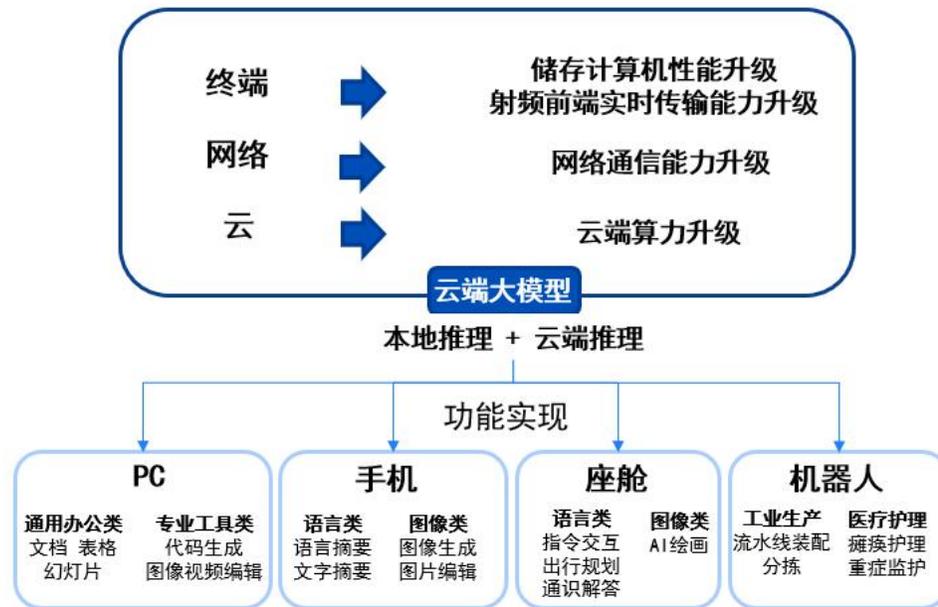


- ◆ 大模型正在从算力统管和场景优化两个维度在边缘侧进行落地尝试。
 - 算力：大模型能够部分替代和接管原有云端计算中心的算力调度权限与能力，减少云端传输所带来的时间损耗，对边缘侧算力使用效率带来改进。
 - 场景：大模型可取代原有边缘侧用于预测、决策、判别、生成等多类任务的小模型，提升场景泛化能力和使用效果。
- ◆ 终端模型与云端模型协同提供服务，存、算、网同步升级。大模型已经率先在手机和汽车座舱中得到初步应用。计算存储需求也在催化终端硬件和网络性能迭代。在大模型裁剪技术以及终端算力制约下，端侧部署大模型参数量小，功能相对有限，云端大模型相互协同为用户提供更丰富的场景体验。

大模型在云、边、端落地对物联网技术体系的影响与塑造



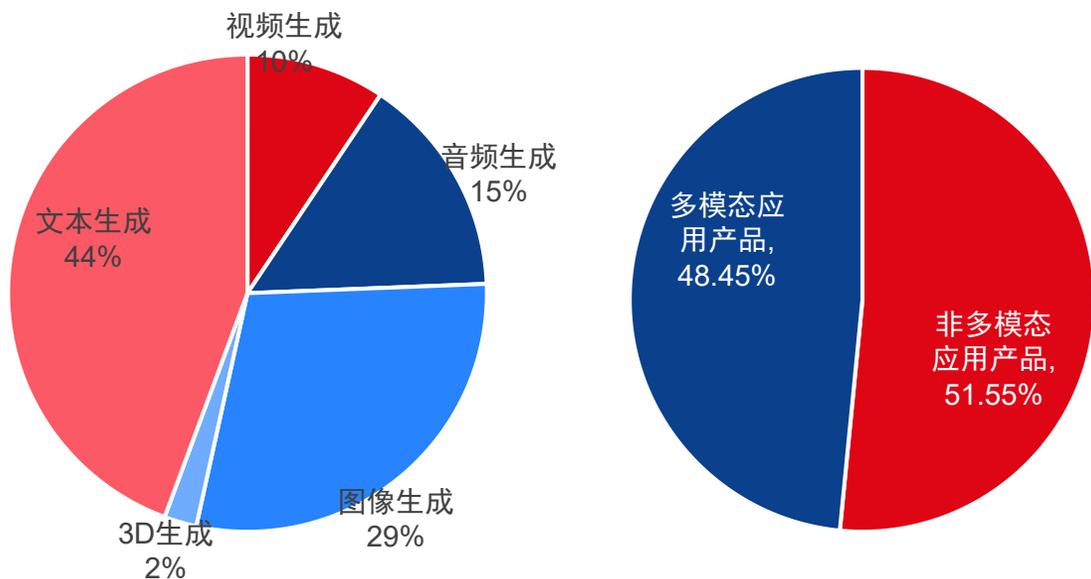
云端与终端示意图



各类形态产品格局逐步清晰

- ◆ AIGC应用基于所生成的模态，可分为文本生成、音频生成、图像生成、视频生成、3D生成的产品。在所统计的产品中，44%以上的应用专注于文本模态的生成，在AIGC应用中最为普及。图像生成产品约占29%，应用也相对成熟。音频生成应用数量占比约为15%，生成式AI为其带来新技术驱动。
- ◆ AI应用90%+为软件形态，AI硬件开始进行尝试。AI硬件主要分为两种：
 - AI+万物：即在传统产品上集成AI功能，主要包含AI手机、AIPC、XR、具身智能、学习机、AI电视、汽车等。
 - AI新产品：即利用大模型的特性，创新消费电子的产品形态与交互方式，代表性产品有AI Pin、Rabbit R1等。

产品模态分布（左）与多模态应用占比（右）

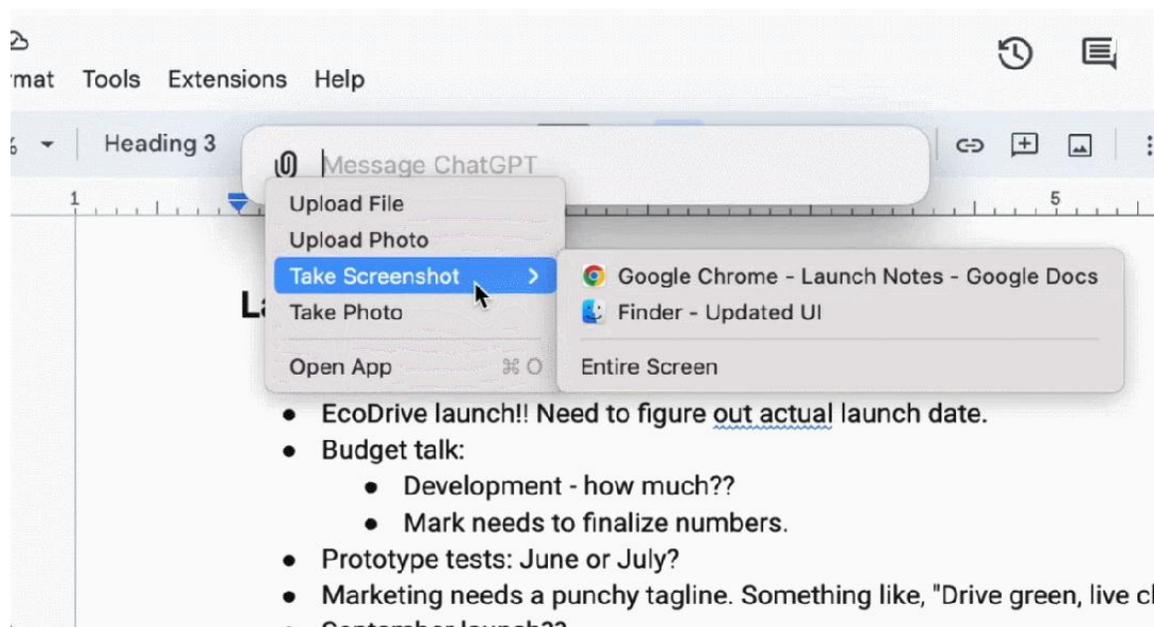


AI与硬件结合代表：Rabbit R1



- ◆ 手机AI PC应用从用户感知视角，多模态的人机交互将提高实用性。AI终端将从存储—应用—交互一体的娱乐/工具机，逐渐演化为用户随身携带的智能BOX。
- ◆ 从技术栈层面，操作系统作为全机能力调度的核心将发挥主体性作用，硬件厂商将以AI操作系统为核心重塑自身生态，原有软件厂商的用户数据与流量入口优势或将被削弱。
- ◆ 端侧应用主要功能：
 - 内容总结：借助浏览器AIGC插件总结文字、视频等内容，提升阅读效率等，并以标签形式提炼，助力工作提效。
 - 信息生成：围绕工作场景，用户可借助AI生成所需的文本、图片等内容信息，也可生成办公文档等，提高效率。

PC端调用示例



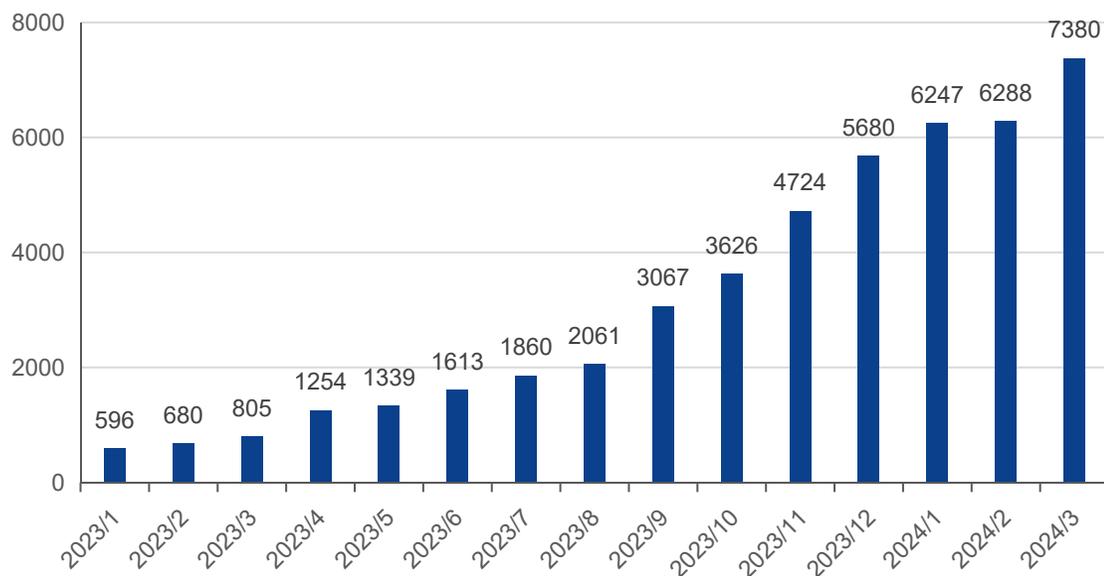
Meta与雷朋联名智能眼镜



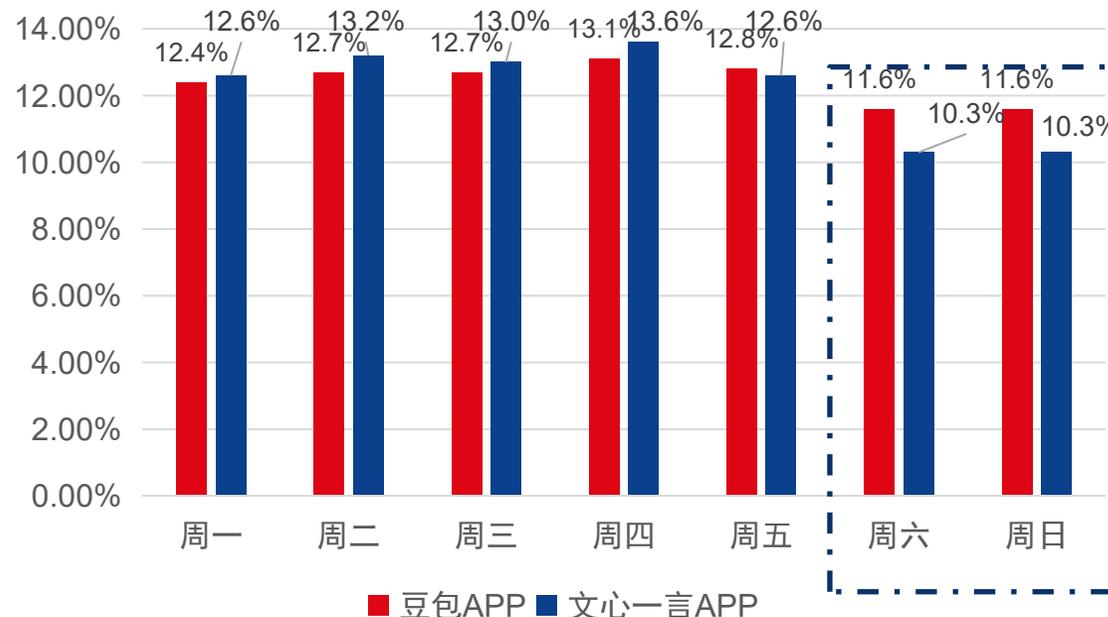
应用侧持续增长

- ◆ AIGC多模态应用涉及PC和移动端，头部APP拥有超5千万活跃用户。开发者在PC、移动端陆续以独立和嵌入的形式，完成AIGC产品在各领域布局。
- ◆ 开发者在PC、移动端陆续以独立和嵌入的形式，完成AIGC产品在各领域布局。
- ◆ 从市场表现来看，用户对AIGC应用热情度较高，独立APP的需求持续增长。根据QuestMobile数据显示，2024年3月，AIGC APP行业用户规模同比增长8倍以上。
- ◆ AIGC APP在工作日（周一至周五）期间的活跃度高于休息日（周六、周日），推测现阶段AIGC APP与办公场景的适配性更强。

AIGC APP行业月活跃用户规模趋势（单位：万）



2024年1月 千万级AIGC APP周均活跃占比分布



以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/387132101013006124>