

ChatGPT与人工智能 技术发展报告

——关键词：ChatGPT 自然语言处理 AI数据 AI芯片

广闻广识，全球行业专家网络



成为广闻专家，有偿提供行业知识与观察

——请关注公众号填写表单

企业需要咨询行业专家

——请邮件info@guangwen.online

目录

前言：人工智能市场现状与发展回顾	6
(一) 人工智能整体回顾	6
(二) 全球及中国人工智能市场规模	6
(三) 人工智能投融资情况与企业上市情况	7
(四) 人工智能专利技术情况	9
4.1 全球人工智能专利公开与申请情况	9
4.2 全球人工智能专利来源国分布情况	10
(五) 全球人工智能部分政策及规划	11
(六) 人工智能部分重要进展更新	12
6.1 国外最新进展	12
6.2 国内最新进展	12
第一章：爆红的 chatGPT	14
(一) ChatGPT 基本概念	14
(二) 发展历程	14
(三) 市场现状与发展概况	14
3.1 市场规模	15

3.2 投融资情况	16
3.3 国内外产业布局情况	17
3.4 产业链图谱	17
(四) 相关重要技术	18
4.1 大型语言模型 (LLM)	19
4.2 人类反馈强化学习 (RLHF)	19
(五) 重点应用领域	20
(六) 未来发展趋势	21
(七) 龙头企业发展动向	22
7.1 国内外龙头企业布局情况	22
7.2 国际龙头企业——OpenAI	23
7.3 国内龙头企业——百度	24
第二章：相关行业发展之自然语言处理	26
(一) 基本概念及作用	26
(二) 发展阶段	27
(三) 市场现状与发展概况	28
3.1 市场规模	28

3.2 投融资情况	30
3.3 专利技术.....	31
3.4 产业链	32
(四) 重要技术进展	32
(五) 未来发展趋势	33
(六) 龙头企业发展动向	34
6.1 国际龙头企业——Google	34
6.2 国内龙头企业——科大讯飞	35
第三章：相关行业发展之 AI 数据	37
(一) 基本概念及作用	37
(二) 发展阶段	38
(三) 市场现状与发展概况.....	38
3.1 市场规模	38
3.2 投融资情况	40
3.3 专利技术.....	40
3.4 产业链	41
(四) 重要技术进展	42

(五) AI 数据发展趋势	43
(六) 龙头企业发展动向	43
6.1 国际龙头企业——Appen	43
6.2 国内龙头企业——海天瑞声	45
第四章：相关行业发展之 AI 芯片	46
(一) 基本概念及作用	46
(二) AI 芯片分类	47
(三) 发展阶段	48
(四) 市场现状与发展概况	49
4.1 市场规模	49
4.2 投融资情况	51
4.3 产业链图谱	51
(五) 芯片类型对比	52
(六) 未来发展趋势	53
(七) 龙头企业发展动向	54
7.1 国际龙头企业——英伟达	54
7.2 国内龙头企业——寒武纪	55

文献参考..... 57

广闻广识公司简介 59

前言：人工智能市场现状与发展回顾

（一）人工智能整体回顾

2021-2022 年，全球人工智能市场呈现出蓬勃发展的态势：

应用领域不断扩大：人工智能应用领域已经不再局限于传统的图像和语音识别、自然语言处理等领域，而是涵盖了金融、医疗、制造业、物流、零售等各个行业，并且在这些领域中得到了广泛应用。

技术创新不断推进：随着深度学习、自适应系统、自主学习技术的不断发展，人工智能技术不断向前推进，越来越多的企业开始尝试将人工智能技术应用到其核心业务中。

产业生态加速形成：人工智能已经成为一个拥有完整产业链的领域，从芯片、算法、平台、应用等方面都有相应的企业存在，并且这些企业之间形成了一定的合作与竞争格局。

政策不断推动：各国政府都将人工智能视为未来的重要战略领域，出台了一系列支持政策和计划，以促进人工智能技术的发展和應用。

未来随着技术不断提升和应用广泛，人工智能市场将会变得更加成熟和完善。

（二）全球及中国人工智能市场规模

近年来，随着计算机技术的不断进步和应用领域的扩大，全球人工智能市场正在快速发展。据预测，到 2022 年，该市场规模已达 1290 亿美元，其中中国

市场规模预计达到 130.3 亿美元,超过全球市场规模的十分之一。未来 2022-2026 年,中国人工智能市场预计将以 19.6%的年复合增长率稳步发展,2026 年将有望超过 266.9 亿美元。

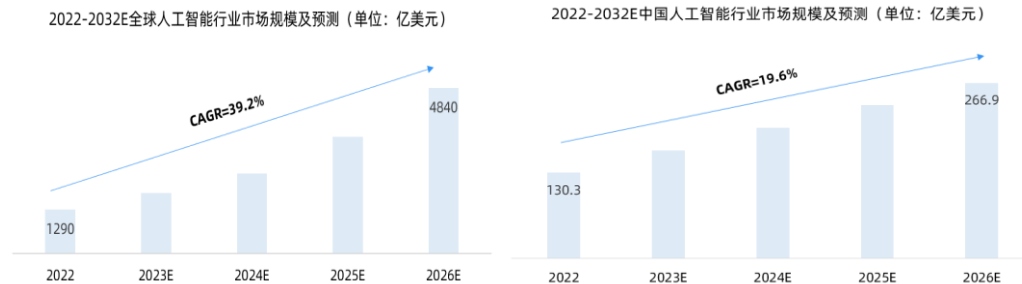


图 0-1 : 2022-2026E 全球及中国人工智能的市场规模 (数据来源: market.us, IDC)

(三) 人工智能投融资情况与企业上市情况

在 2014 年以前,多数重要模型都是由学术界发布。然而,随着先进人工智能系统的构建越来越依赖于大量数据、计算机能力和资金资源,相较于非营利组织和学术界,行业参与者拥有更丰富的资金资源,因此成为研发重要模型的主要力量。到 2022 年,已经有 32 个重要的机器学习模型诞生在工业界,而学术界仅有 3 个。

2022 年,全球人工智能领域的私人投资经历了十年来的首次下降。相比于 2021 年,2022 年 919 亿美元的投资金额下降了 26.7%。中国的人工智能领域的投资情况与全球趋势类似。在 2020 年至 2021 年期间,中国的人工智能领域一级市场总融资金额从 991.79 亿元增长至 2187.92 亿元,实现了 120%的增长。然而,截至 2022 年 11 月 10 日,总体呈现下降趋势。

2019-2022年11月10日中国人工智能领域一级市场股权融资情况

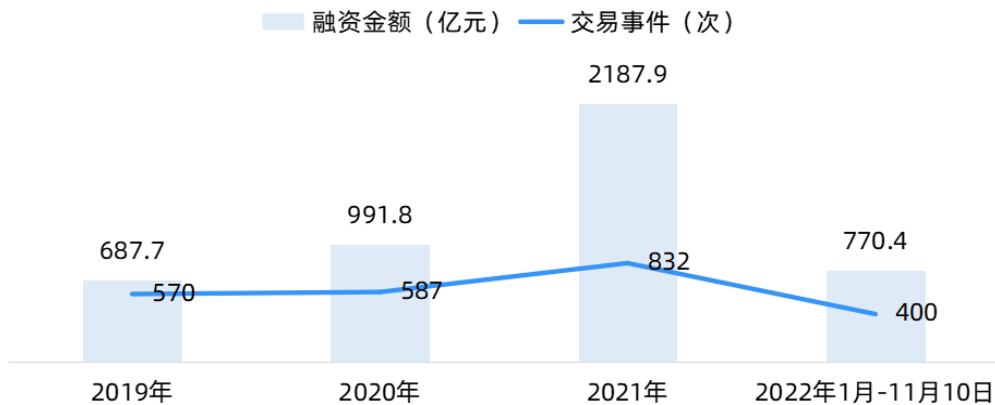


图 0-2: 2019-2022 年 11 月 10 日中国人工智能领域一级市场股权融资情况 (数据来源: IT 桔子)

2015-2022 年 11 月 10 日, 中国人工智能领域一级市场的平均单笔融资金额是呈波动增长的趋势, 在 2017 年后, 单笔融资金额平均不低于 1 亿元, 在一定程度上表明人工智能领域平均估值有所提升, 投资门槛越来越高。

2015-2022年11月10日中国人工智能领域一级市场平均单笔融资金额 (单位: 千万元)

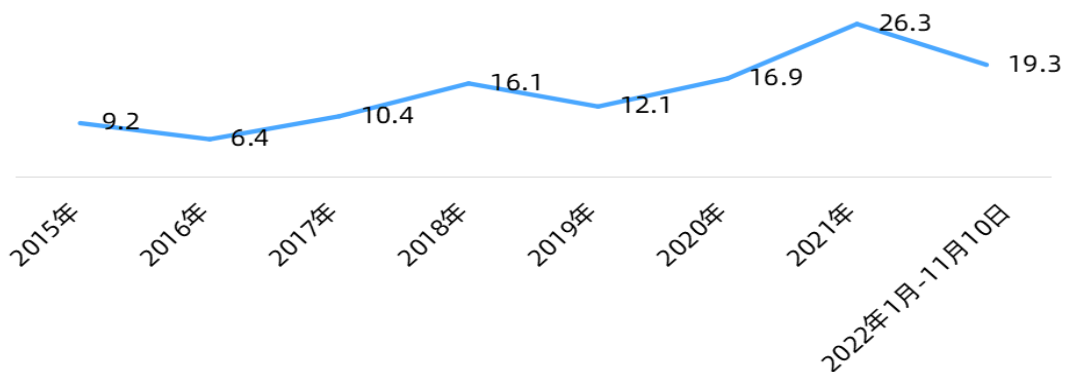


图 0-3: 2015-2022 年 11 月 10 日中国人工智能领域一级市场平均单笔融资金额

(数据来源: IT 桔子)

2015 年至 2022 年 11 月 10 日, 中国人工智能领域早期投资呈现显著下降趋势, 波动范围最大, 从占比 60%下降至 25%。相比之下, 成长期投资占比变化

较小，基本保持稳定；而战略投资和中后期投资则有一定程度的增加。

值得注意的是，从 2019 年开始，成长期投资的占比开始高于早期投资，且一直保持这种趋势至 2022 年 11 月 10 日。在一定程度上表明国内人工智能领域的风险投资偏好正在逐渐走向成熟的态势。

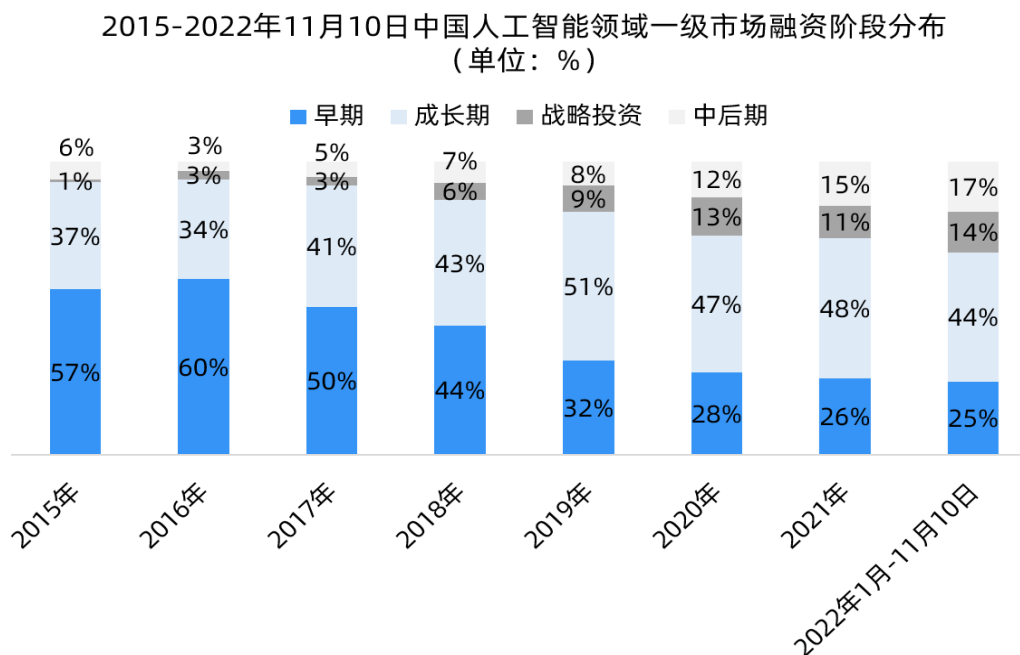


图 0-4: 2015-2022 年 11 月 10 日中国人工智能领域一级市场融资阶段分布 (数据来源: IT 桔子)

(四) 人工智能专利技术情况

4.1 全球人工智能专利公开与申请情况

全球人工智能领域每年专利公开数量呈逐年上升趋势,2022 年达到 3 万项以上。从专利数量上看,人工智能在不断的发展进步。

2017-2023年4月7日全球人工智能行业专利公开与申请情况(单位：项)

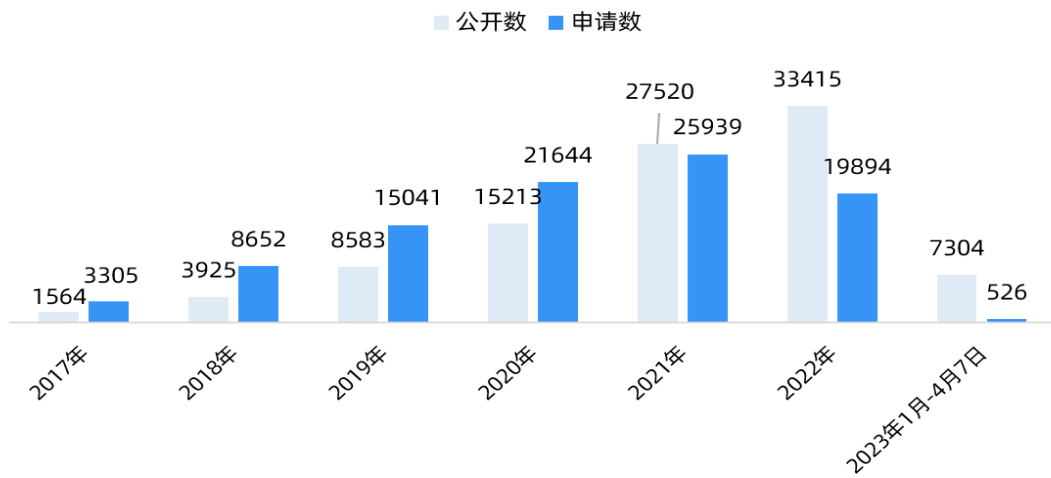


图 0-5: 2017-2023 年 4 月 7 日全球人工智能行业专利公开与申请情况 (数据来源: 智慧芽)

2017-2023年4月10日
全球人工智能前十大热门技术词： 电子设备、深度学习、神经网络、服务器、语音、机器人、人工智能系统、大数据、数据处理、数据集。

图 0-6: 2017-2023 年 4 月 10 日全球人工智能行业前十大热门技术词 (数据来源: 智慧芽)

4.2 全球人工智能专利来源国分布情况

2018-2023 年 4 月 10 日, 我国在人工智能专利公开数量上处于绝对领先地位, 远超美、日韩等其他国家。

2017-2023年4月10日人工智能领域专利技术来源国分布

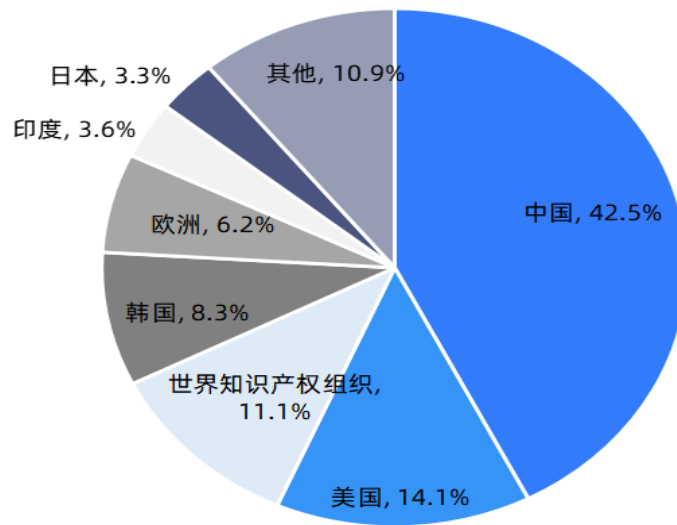


图 0-7: 2017-2023 年 4 月 10 日全球人工智能国家技术分布 (数据来源: 智慧芽)

(五) 全球人工智能部分政策及规划

国家	年份	人工智能领域相关内容
美国	2022	在人工智能咨询委员会的第一次会议上, 提出了建立五个工作组, 旨在为美国人工智能领域的发展提供决策建议
	2022	《国防部负责任人工智能战略和实施路径》
	2021	《人工智能/机器学习战略规划》
	2021	《AI国家安全委员会最终建议报告》
英国	2022	《国防人工智能战略》
俄罗斯	2021	政府通过为AI项目提供补贴等支持AI发展的两份政府决议, 推动AI与各产业深度融合
澳大利亚	2021	发布“AI行动计划”
日本	2022	《AI战略2022》
法国	2021	《AI国家战略》新计划
中国	2022-2023	全国各省市积极出台相关政策以支持和推进人工智能产业发展
	2022	《扩大内需战略规划纲要(2022-2035年)》
	2022	《最高人民法院关于规范和加强人工智能司法应用的意见》

图 0-8：2021-2023 年 4 月 10 日全球人工智能行业相关政策及规划部分汇总

（六）人工智能部分重要进展更新

6.1 国外最新进展

2023 年 1 月 3 日，Meta 发布自监督算法 data2vec 2.0：通过整合语音、视觉和文本三种数据模态，使用一个框架进行训练，可以在保持相同精度的情况下，将训练效率提高 16 倍。

2023 年 1 月 27 日，谷歌公布文字-图像生成模型 Muse（缪斯）、MusicLM 等：文字可以直接生成高保真的音乐。

2023 年 1 月 28 日，斯坦福团队推出 DetectGPT：旨在帮助教育工作者识别出 ChatGPT 生成的论文。

2023 年 1 月 28 日，马斯克合伙人研制大脑“第七层”：瘫痪病人可用意念操作电子设备。

2023 年 1 月底，GPT-3 核心成员打造 ChatGPT 竞品 Claude：目标成为一款安全、符合人类价值观且遵循道德规范的人工智能系统。

2023 年 2 月 24 日，Meta 公布一款新的人工智能大型语言模型 LLaMA：能支持研究人员在多项复杂任务中取得进展，如生成文本、进行对话、撰写总结性文件、证明数学定理或预测蛋白质结构等。

6.2 国内最新进展

2022 年 12 月 28 日，新型智能皮肤可实现手部任务快速识别：未来或许可

实现在隐形键盘上打字，实现沉浸式触摸。

2023年1月25日，中山大学联合团队实现机器人液化越狱：在未来，医疗保健、精准货物运输以及电子维修等领域可能会受益于该新技术的发展。

2023年1月底，中山大学研发“数字流式”系统 DeepFundus：促进人工智能在临床方面的应用。

2023年3月16日，百度推出中国式的 ChatGPT 服务：为中国提供了一个更加便捷和高效的自然语言处理平台，有助于推动中国在人工智能领域的发展和应用。

2023年4月7日，阿里云自研大模型“通义千问”开始邀请用户测试体验：每一家企业未来都能够在阿里云上调用通义千问的全部功能，同时利用自身的行业知识和应用场景训练自己的企业大模型。

2023年4月10日，商汤科技宣布推出自研的中文语言大模型应用平台“商量”：它具备自然语言生成、文本生成图、感知模型标注和模型研发功能，能够理解多轮对话和超长文本。

第一章：爆红的 chatGPT

（一）ChatGPT 基本概念

ChatGPT 是由 OpenAI 开发的一种 AI 语言模型，它可以根据给定的输入生成类似于人类的文本。该模型经过在大量文本数据集上的训练，具备生成问题回答、长文本总结、故事创作等能力。通常情况下，它被用于对话式 AI 应用中，以模拟与用户进行类人对话。

（二）发展历程



图 1-1：GPT 模型迭代历程

（三）市场现状与发展概况

3.1 市场规模

ChatGPT 所涉及的行业属于 AI 生成内容行业（Artificial Intelligence Generated Content, AIGC）。这个行业整合了人工智能、大数据、云计算、5G 等多个技术领域，可以在创意、表现力、迭代、传播、个性化等方面发挥技术的优势。在 2021 年之前，AIGC 主要生成文字内容，但是新一代的模型 ChatGPT-4 已经可以处理文字、语音、代码、图像、视频、机器人动作等多种格式的内容。随着深度学习模型的不断完善、开源模式的推动以及大型模型商业化的可能，AIGC 行业未来发展前景广阔。

如需查看全文，请直接购买报告！

点击下方链接：

<https://www.djyanbao.com/paidReport/detail?id=7858>

第二章：相关行业发展之自然语言处理

（一）基本概念及作用

自然语言处理（Natural Language Processing，简称 NLP），是指利用人类交流所使用的自然语言与机器进行交互通讯的技术，通过人工智能技术对自然

语言进行处理，使得计算机对其能够可读并理解，也是 chatGPT 类的 AI 文本生成功能的核心技术之一。

自然语言处理涉及细分领域很多，其核心任务涉及两部分，自然语言理解 (Natural-language Understanding, 简称 NLU) 和自然语言生成 (Natural-language generation, 简称 NLG) 。

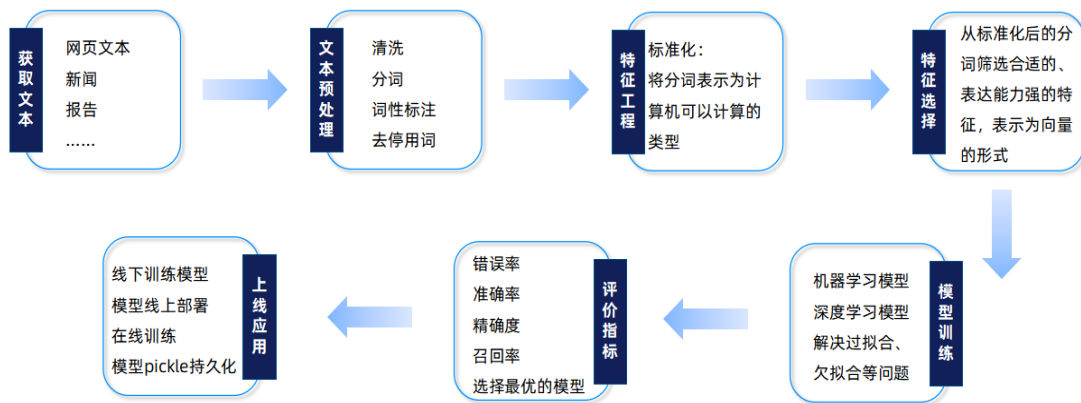


图 2-1：自然语言处理的工作流程

ChatGPT 和自然语言处理之间是相辅相成的关系：

首先，自然语言处理作为 chatGPT 的核心技术之一，它的研究启发了大规模语言模型的构建，使得 ChatGPT 能够更好地理解自然语言，生成更加优质的文本。其次，自然语言处理任务中的数据集和评估指标也为 ChatGPT 的训练和评估提供了基础和参考。此外，自然语言处理还促进了 ChatGPT 在多任务学习、迁移学习等方面的应用，进一步拓展了 ChatGPT 的应用范围和性能。

反之，ChatGPT 的出现也深刻地影响了自然语言处理领域。一方面，大规模语言模型中蕴含的信息使得传统自然语言处理任务（如分词、词性标注、句法分析等）不再需要专门的研究。另一方面，ChatGPT 将各种自然语言处理任务转化

为序列到序列生成问题，打破了任务间的界限，促进了多任务学习和迁移学习的发展。此外，工业界拥有丰富的计算资源、用户信息和反馈数据，这与学术界进行系统级研究形成互补，产生了相互促进的作用。

（二）发展阶段

自然语言处理的发展过程对人类具有重要意义。自然语言处理最早可以追溯到 20 世纪 50 年代，随着自然语言处理技术的不断进步，我们可以更好地利用自然语言与计算机进行交互，从而实现更高效、更便捷的信息处理和沟通方式。

ChatGPT 就是自然语言处理技术应用的优秀代表。

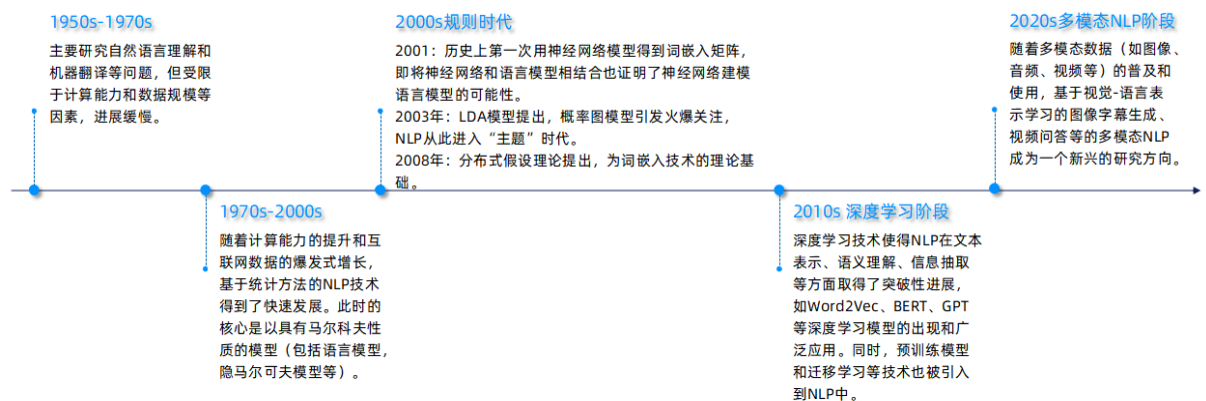


图 2-2：自然语言处理发展历程

（三）市场现状与发展概况

3.1 市场规模

如需查看全文，请直接购买报告！

点击下方链接：

<https://www.djyanbao.com/paidReport/detail?id=7858>

第三章：相关行业发展之 AI 数据

（一）基本概念及作用

数据是 AI 的三大基石（算法、算力）之一，ChatGPT 在对语言做出反应时，需要的不仅仅是在不同行业数据的采集和利用，还包括数据准备、模型选择、训练、评估、参数调整、预测等流程。

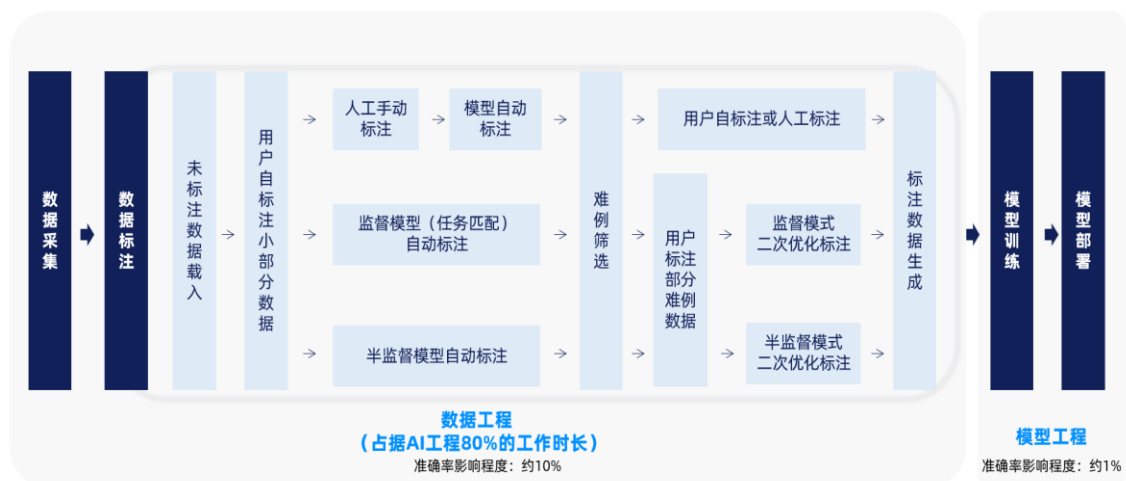


图 3-1：AI 数据工作流程

AI 数据对 ChatGPT 的重要性在于为 ChatGPT 提供了大量的训练数据和算法评估服务，从而使其在自然语言处理方面表现更出色。AI 数据平台可以提供海量的对话样本和其他类型的 NLP 数据，这些数据可以用来训练和优化 ChatGPT 模型，并使其能够更好地理解和回答用户的问题。此外，AI 数据平台还可以提供常规任务数据，如文本分类、实体识别等，这些任务可以帮助 ChatGPT 实现更复杂的应用场景。同时，AI 数据平台还可以提供精准的模型评估服务，以评估 ChatGPT 模型的准确度和效果。

（二）发展阶段

AI 数据的发展经历了四个阶段：数据收集阶段、数据标注和整理阶段、数据共享和开放阶段、AI 数据平台阶段。AI 数据为 AI 模型提供了用于训练、测试和优化的数据，使它们能够更快地学习和适应新的任务。AI 数据的广泛应用和使用，有利于提高 AI 模型的准确性和效率。这些数据比如包括从标准低级任务（如图像分类）到复杂高级任务（如对话生成）的各种数据来源，可以为 AI 模型提供足够的知识和内容，促进 AI 模型逐步提升到能够解决高级任务的水平。AI 数据行业正在迅速发展，AI 数据市场正在扩大，对于推动 AI 技术发展、提高 AI 模型能力、促进 AI 应用落地和促进整个 AI 行业的发展方面都具有重要的作用。



图 3-2: AI 数据发展历程

(三) 市场现状与发展概况

如需查看全文，请直接购买报告！

点击下方链接：

<https://www.djyanbao.com/paidReport/detail?id=7858>

第四章：相关行业发展之 AI 芯片

(一) 基本概念及作用

AI 芯片是 ChatGPT 实现的基础，更是承载人工智能行业发展的基础设施，为模型训练提供算力支持。

广义上讲，能够运行人工智能算法的芯片都叫做 AI 芯片，但是通常意义上的 AI 芯片指的是针对人工智能算法做了特殊加速设计的芯片。AI 芯片也被称为 AI 加速器或计算卡，即专门用于处理人工智能应用中的大量计算任务的模块。深度学习模型与推荐式系统模型复杂度进一步提升，对芯片算力提出了更高的要求，AI 芯片算力已步入蓬勃发展阶段。

根据机器学习算法步骤，可分为训练（training）芯片和推断（inference）芯片。从技术架构来看，AI 芯片主要分为图形处理器（GPU）、现场可编程门阵列（FPGA）、专用集成电路（ASIC）、类脑芯片四大类。

分类方式	具体类别	功能定义
根据机器学习算法步骤	训练（training）芯片	通过大量的数据输入，构建复杂的深度神经网络模型，运算能力较强。
	推断（inference）芯片	利用训练出来的模型加载数据，计算“推理”出各种结论，侧重考虑单位能耗算力、时延、成本等性能
根据技术架构	图形处理器（GPU）	通用型人工智能芯片
	现场可编程门阵列（FPGA）	半定制芯片
	专用集成电路（ASIC）	全定制芯片
	类脑芯片	模拟人脑神经元结构，处于起步阶段

图 4-1：AI 芯片分类

（二）AI 芯片分类

不同的应用场景，按照“端-边-云”，对 AI 芯片有着不同的要求。其中端和边上进行的大部分是 AI “推理”；大部分的训练过程是在云和数据中心进行，云端训练和推断计算主要由 AI 服务器完成，底层算力芯片包括 CPU、GPU、FPGA、ASIC 等。

CPU 是 AI 计算的基础，负责控制和协调所有的计算操作。GPU、FPGA、ASIC 是 AI 计算的核心作为加速芯片处理大规模并行计算。

	端			边		云	
	耳机电话	智能手机	个人电脑	网络摄像机 IPC	边缘服务器	数据中心	
算力 (OPS)	20M	100G	1-10T	10-20T	10-20T	10-500T	200+T
功耗	1mW	10mW	1-2W	3-10W	3-10w	10-300W	200+W
模型大小	10KB	100KB	10MB	10-100MB	10-100MB	100+MB	300+MB
延时	<10ms	~10ms	10-100ms	10-500ms	10-500ms	ms~s	ms~s
推理	√	√	√	√	√	√	√
训练	X	X	√	√	√	√	√

图 4-2: AI 芯片“端-边-云”分类及对比 (数据来源: ofweek 电子工程网、华为全联接大会)

芯片类型	GPU	FPGA	ASIC
定制化程度	通用型	半定制化	全定制化
可编程性	不可编程	容易编程	难以编程
算力	中	高	高
价格	高	中	低
优点	通用性强, 适合大规模并行运算; 设计制造工艺成熟	可灵活配置架构适应算法迭代; 功耗较低, 开发时间较短 (6个月)	固化算法实现极致性能和能效、平均算法很强, 功耗低, 体积小; 量产后成本最低
缺点	并行运算能力在推理端无法完全发挥	量产单价高, 峰值计算能力较低, 硬件编程难	前期投入高, 研发时间长 (1年), 技术风险大
应用场景	高级复杂算法和通用型人工智能平台	适用于各种具体的行业	特殊场景独立设计智能算法软件

图 4-3: AI 芯片对比

(三) 发展阶段

如需查看全文，请直接购买报告！

点击下方链接：

<https://www.djyanbao.com/paidReport/detail?id=7858>

文献参考

- [1] 【Acumen Research and Consulting】 Generative AI Market and Region Forecast, 2022 - 2030
- [2] 【IDC】 2021-2022 中国人工智能算力发展评估报告
- [3] 【IT 桔子】 2022 年中国人工智能投融资分析报告
- [4] 【market.us】 Global Artificial Intelligence Market By Solution (Hardware, Software, and Services), By Technology (Deep Learning, Machine Learning, NLP), By End-Use, By Region, Segment Outlook and Forecast 2023-2032
- [5] 【艾媒网】 2023 年中国 AIGC 行业发展研究报告
- [6] 【国金证券】 传媒互联网产业行业深度研究：AIGC 步入快车道，游戏行业降本增效或超预期
- [7] 【国金证券】 计算机软件行业深度研究：新场景催化 AI “燃料” 数据采标需求加速释放
- [8] 【国盛证券】 百度集团-SW (09888)：类 Chat-GPT 产品将上线，AI 商业化探索行业领先
- [9] 【国盛证券】 金山办公 (688111)：AI 能力积淀已久，国产软件扛旗者迈入智能办公新纪元
- [10] 【华西证券】 “文心一言” 发布，国内厂商距离复现 ChatGPT 有多远？
- [11] 【申万宏源】 AIGC 系列之七：国产 AI 芯片的百倍算力需求

[12] 【斯坦福】2022 年人工智能指数报告

[13] 【浙商证券】寒武纪深度报告：GPT 算力系列深度·寒武纪（688256）：中国 AI 芯片先行者

[14] 【智谱 AI】人工智能发展月报

如需查看全文，请直接购买报告！

点击下方链接：

<https://www.djyanbao.com/paidReport/detail?id=7858>

广闻广识公司简介

上海广闻广识信息科技有限公司致力于打造行业领先的信息服务平台，以庞大的专家网络与专业合规的服务体系为依托，希望通过促成各种形式的沟通交流，从而拓展人脉，使得各方都能在沟通中获益。

广闻广识通过线上线下会议、专家访谈等形式，快速灵活为各类基金公司和投资机构提供高价值、高定制化的行业知识服务，分享与传递专业意见，赋能客户决策。

向世界分享你的声音

让您的观点与智慧**广闻广识!**

