



F R O S T & S U L L I V A N

60 Years of Growth, Innovation and Leadership

中国AI基础软件
市场研究报告（2023）

A Frost & Sullivan
White Paper

目录

执行概要2

关键发现点.....3

章节一：中国 AI 基础软件发展背景4-5

 AI2.0 时代到来开启巨大的平台型机会..... 4

 庞大的数据基础为国内 AI 大模型发展提供根基 5

章节二：中国 AI 基础软件市场概览6-10

 AI 基础软件市场定义与解读..... 6

 蓬勃的 AI 大模型市场为 AI 基础软件发展注入动力 7

 利好的政策为 AI 基础软件市场提供制度保障 8

 开源演进为 AI 基础软件市场提供良好的生态土壤 9

 位于产业链中游的 AI 基础软件具有极高的附加值 10

章节三：AI 基础软件厂商竞争格局.....11-12

 AI 基础软件厂商竞争力图谱..... 11

 评估标准及主要成果 12

附录.....13

简介

过去十年，是人工智能从实验室走向产业化的十年。AI技术革命掀起的产业浪潮，势不可挡地席卷了全球的各行各业。作为现今最为活跃的创新领域，人工智能在迅速融入细分场景的同时也在不断地重塑传统行业模式，为经济社会输出独一无二的未来价值。

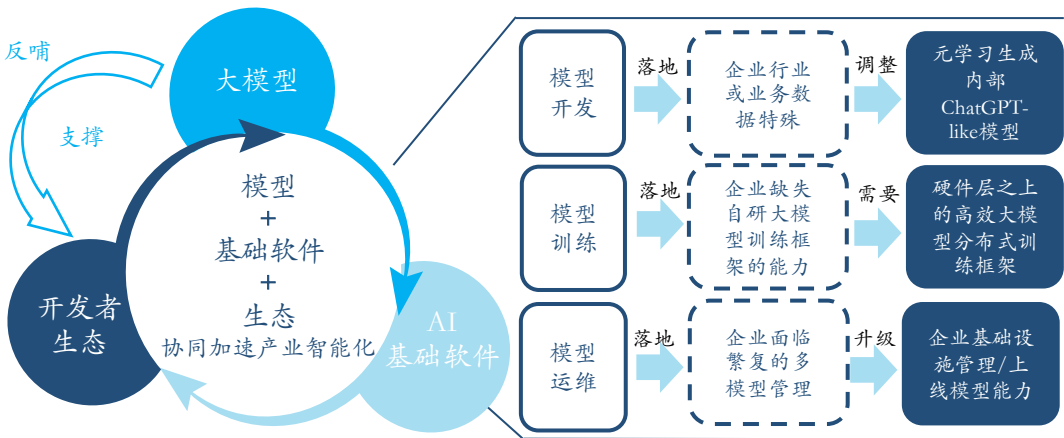
大模型是人工智能迈向通用智能的里程碑技术。业内人士认为，在基于数据的互联网时代和基于算力的云计算时代之后，将进入基于大模型的AI时代。大规模数据集和模型训练算力等领域取得的突破性进展，离不开AI基础软件的底层支撑。尽管大模型当前表现优异，但对于各行业使用者来说，实际应用于业务场景仍然存在较高的技术和成本门槛。当前大模型的参数标准并不统一，相对于参数级，模型的效果且是否能够支持快速迭代对于用户实际应用来说更为重要。用户能够在一个白盒大模型基础上快速地、低成本地微调和迭代出客制化

AI基础软件产业协同发展图

的小模型，才能高效地实现丰富场景的大模型应用。这就再次点明了AI基础软件工具链的重要性。

随着AI基础软件不断发展，近年来大模型落地应用呈现加速态势。在开发层面，AI基础软件可帮助提供大模型能力萃取；在训练层面，其提供了硬件层之上的高效大模型分布式训练框架；在运维层面，其作为企业的基础设施，可增强企业管理、上线模型的能力。

本报告旨在明晰AI基础软件框架的基本概念和分类，通过对行业发展历程及产业链的梳理探究行业核心驱动因素，剖析当下AI基础软件行业市场趋势，并基于应用、产品及生态三大关键维度构建厂商竞争力体系，评估主流玩家的核心竞争优势及综合壁垒，形成对中国市场AI基础软件框架发展情况的客观评价，并为行业未来发展提供参照建议与关键启发。



来源：弗若斯特沙利文

关键发现点

在人工智能技术迭代的过程中，AI基础软件行业整体格局将逐步成型。科技巨头生态体系建设叠加中国肥沃的政策和技术土壤，将为业内厂商提供前所未有的发展机遇。本报告将从行业宏观环境洞察市场需求，从技术规律把握未来趋势，进而识别出在AI浪潮之下领先的AI基础软件厂商，并对其综合竞争实力进行分析：

- **驱动因素：**随着上游硬件、算力升级以及数据量增加，AI基础软件行业下一阶段驱动因素主要包含三个层面。**其一，在全球AI政策环境持续优化的趋势下，中国对大模型技术的宏观战略规划将为行业提供肥沃的成长土壤。**北京市政府已明确出台政策，加快建设算力中心，支撑多模态大模型、科学计算大模型等研发。**其二，下游大模型加速渗透应用场景和规模化应用将为行业带来曙光。**从2023年AI预训练大模型的应用来看，大模型已开始工业质检、智慧交通、时尚设计等多个行业中落地，其距离实现预训练大模型规模化商业化应用仍有一定的距离。**其三，头部科技企业加速布局和产业生态体系垂直整合，将为不同厂商提供多元化的战略切入点。**目前，业内已形成算力、基础软件、平台和服务纵向一体化的共识，从而为云服务企业、AI头部企业及AI芯片巨头提供了多环节的战略布局思路。相关头部AI企业，有望引领本轮生态体系建设。
- **市场趋势：**在核心驱动因素的催化下，中国AI基础软件行业展现出三大主要趋势，并在这个过程中逐步明晰竞争格局。**其一，中国原创技术创新突破加之人工智能政策导向，将推动市场**

产学研用融合创新共同体的成型。中国具备全球领先的视觉、自然语言处理、语音识别等智能任务的工程实现水平，人工智能原创技术创新正处于活跃期。大模型的技术突破结合2023年各省发布的支持政策，将引导行业形成连通企业、高校、研究院所、政府等创新主体的共同体生态。**其二，大模型应用逐步落地，将带来行业发展新需求。**作为模型生态系统的中坚力量，AI基础软件将会成为大模型应用落地的最主要的效率支撑，并通过大模型+小模型的方式，形成模型训练新范式。基础软件的重要性具体表现有三点：一是基础软件工程化的易用性、完备性；二是AI基础软件要与场景结合，能够进行完备的运维过程；三是需要具备安全可靠保障条件。

- **竞争格局：**在竞争格局渐趋明晰的过程中，相关厂商需要跨越技术、人才、资金和品牌四方壁垒，在技术实力、生态规模和功能体验三大成功要素上展开角逐。

来源：弗若斯特沙利文

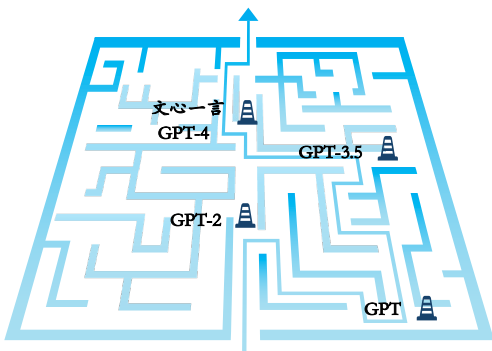
章节一 中国AI基础软件发展背景

AI2.0时代到来开启巨大的平台型机会

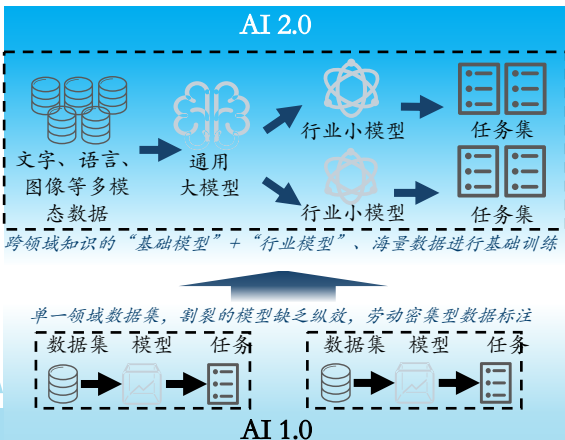
关键发现

- AI2.0是由大数据、云计算和人工智能技术等多种技术的融合所构成的，是一个巨大的平台式机会，将会在未来为各行业提供更加深入和全面的解决方案。随着AI 2.0 时代的到来，基础模型不再需要人工标注数据，也可以自行学习和阅读海量的文本。而且模型通过微调就能以低成本训练适应不同领域的任务。
- 其中，AI2.0相关的应用、平台和基础设施，将会成为包括产业及投资在内的热点。AI基础软件作为AI大模型训练、管理、应用的重要平台将会在AI2.0的长风中顺势演化出巨大的产业机会。

AI 2.0时代：通用大模型崛起的AI平台时代



AI 1.0时代：割裂小模型、人工数据标注的时代



“自1980年首次看到图形用户界面 (graphical user interface) 以来，OpenAI的GPT人工智能模型是我所见过的最具革命性的技术进步。”

—— 比尔·盖茨



“AI1.0就像是发明电，AI 2.0就是电网。AI 2.0时代的来临，首先它是巨大的平台式机会，这个机会将比移动互联网大十倍，而且它是中国的第一次机会。”

—— 创新工场 李开复

李开复博士看好三类未来的创业机会：



AI 2.0智能应用

(各行各业的垂类应用，提高社会生产力)



AI 2.0 平台

(基础大模型 + 平台能力，中间层工具)



AI 基础设施

(基础软件、算法、框架等，压缩计算量，更好地进行模型运维、管理、训练。)

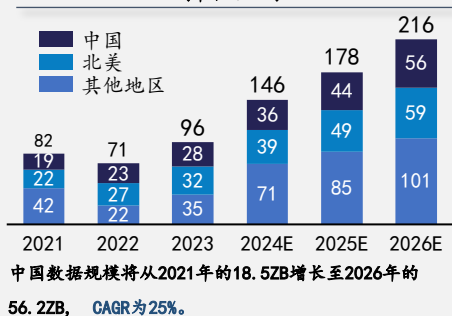
来源：弗若斯特沙利文

庞大的数据基础为国内AI大模型发展提供根基

数据量的爆发为人工智能的发展注入燃料

- 数据量的爆发推动了基础数据服务行业的发展和拓展基础数据服务提供的服务范围。中国的数据量在2017年到2021年，从2.3ZB(十万亿亿字节)增长到23.88ZB，预计在2026年达到76.6ZB，将位居全球第一，且未来依然保持爆发式增长。随着大数据的爆炸式增长，企业和组织对于处理、管理和分析这些数据的需求也在急剧增加。这为基础数据服务行业带来了巨大的市场需求，推动了行业的发展。其次，数据爆炸也推动了数据服务商提供更多元化、更专业化的服务。例如，针对不同类型的数据（如结构化数据、非结构化数据等），数据服务商可能需要提供不同的数据分析和处理服务。

全球数据量分布图-按地区划分 2021-2026
(单位: ZB)



中国AI预训练大模型, 2023年

收集大量数据, 并训练超大规模神经网络

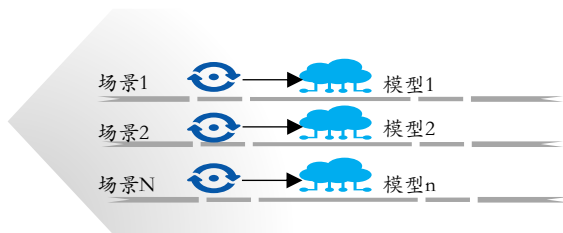


预训练

微调&部署

迭代

对模型微调以达到特定的场景应用/行业应用



来源: 弗若斯特沙利文

章节二 中国AI基础软件市场概览

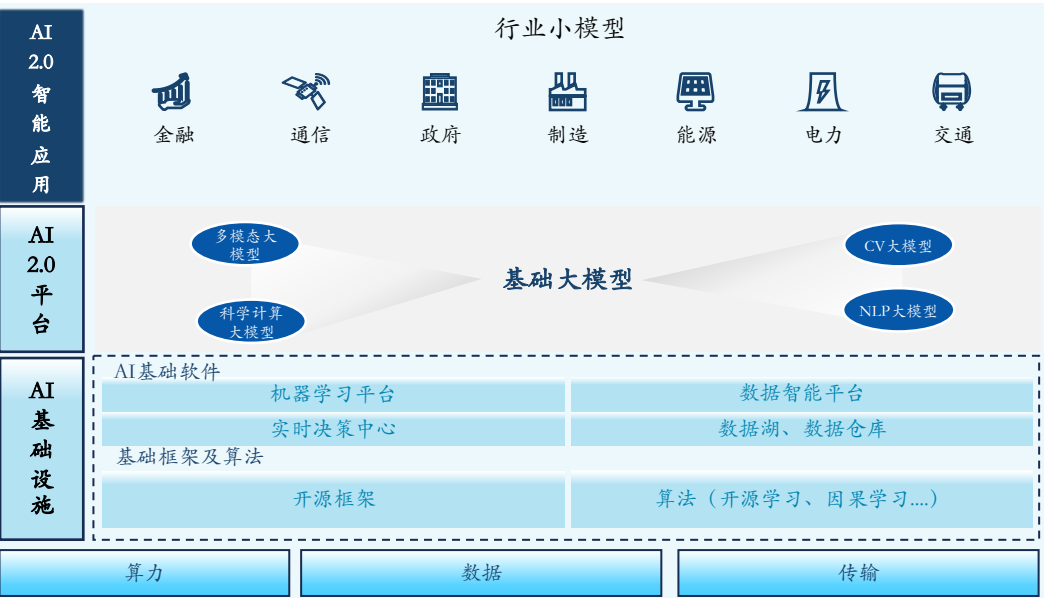
AI基础软件市场定义与解读

关键发现

- AI基础软件包含一系列企业级AI应用所需的平台软件产品及解决方案，是为大模型应用落地的最主要的效率支撑，AI基础软件的发展决定了人工智能发展的深度、高度、广度，其催化大模型应用快速发展，推动政府和企业AI规模化应用。

AI基础软件定义与解读

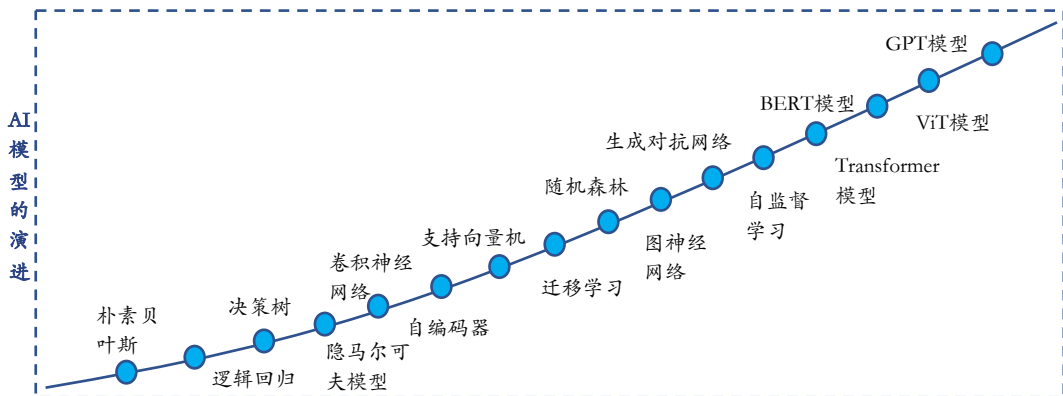
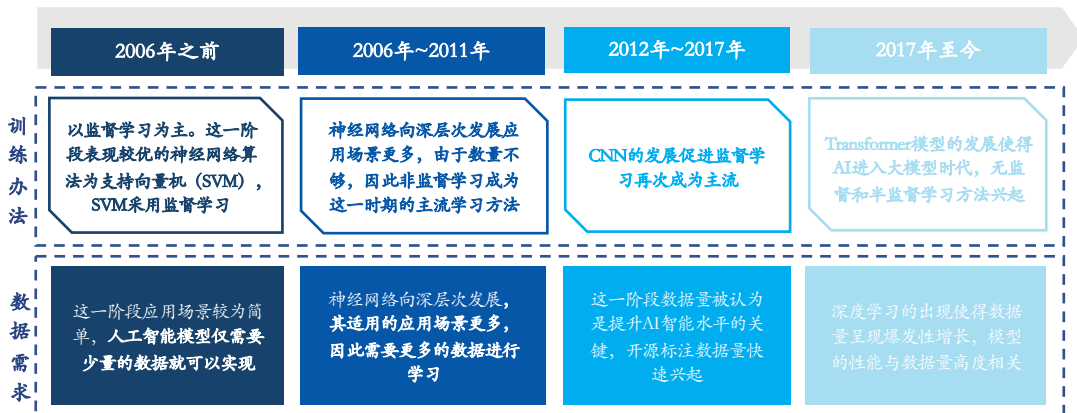
- AI2.0时代的到来，让各行业对大模型应用的需求日渐提升，对大模型对于业务的支撑力也更为重视，但大多数企业都会工程、技术等能力不足的问题。AI基础软件作为AI基础设施的重中之重，为企业客户提供全方位的AI调度以及模型服务，包含机器学习平台等一站式模型平台，以及数据智能平台、实时决策中心、数据湖、数据仓库等服务于AI的数据平台，从而降低各行业客户训练自己人工智能模型的门槛，实现降本增效。



来源: 弗若斯特沙利文

蓬勃的AI大模型市场为AI基础软件发展注入动力

中国AI模型的演进历程，2023年



- 在数据层面，模型训练数据的质量和规模对模型迭代的效果至关重要，尤其是目前内外数据共享和数据共创、数据类别不均衡、极端场景数据缺失等问题，呼唤业内对AIGC结构化数据合成领域的技术探索
- 在技术层面，生成式AI模型的推理能力愈发重要，同时对AI的可信度、可解释度的需求也在提升，呼唤自动机器学习、深度学习、因果学习等领域的基础软件性能提升
- 在商业模式层面，随着大模型的逐渐成熟，通用大模型+行业小模型会成为越来越多企业采用的落地模式，因此，帮助企业自建AI模型的AI基础软件成为大势所趋

来源: 弗若斯特沙利文

利好的政策为AI基础软件市场提供制度保障

2022年国家新一代人工智能开放
创新平台

15个

国家积极布局人工智能产业

在国家监管层面，中国积极布局人工智能产业，竞跑“未来赛道”。随着各行业、各领域对人工智能需求的日益增长，与实体经济深度融合的新模式不断涌现，形成了具有中国特色的研发体系和应用生态，引领着经济社会各领域从数字化、网络化向智能化跃升。以“大模型”为代表的技术爆发，加速了人工智能产业的发展。

如何抓住这一轮技术变革的浪潮，促进区域以及产业发展，北京、上海、深圳等地纷纷出台相关政策举措。北京更是闻风而动，一天之内发布两项政策，冲刺“人工智能第一城”。

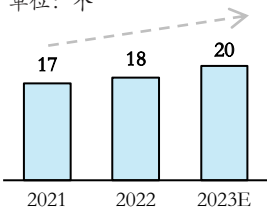
最新地方政策文件中均重点提到了算力端发展，加大算力基础设施的投资力度，同时强调了人工智能的高质量发展，拓展AI创新应用场景的深度与广度。国家及地方出台的多项AI产业支持政策将给产业发展带来助推力，更进一步推动数字经济与实体经济的融合发展。

2022年
国家人工
智能创
新应用
先导区

8个

国家新一代人工智能创新发 展的试验区数量

单位：个



来源：中关村数智人工智能产业联盟

国家指导政策推动AI基础软件发展

数据
优势

加快数字化发展，建设**数字中国**，同时打造数字经济新优势，充分发挥海量**数据**和丰富应用场景优势。同时指出要加强关键数字技术创新应用：聚焦高端芯片、操作系统、人工智能关键算法、传感器等关键领域；建设重点行业**人工智能数据集**，发展**算法推理训练场景**

——《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》

来源：第十三届全国人大四次会议

算法
推理

智能
服务

支持人工智能算法库、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**人工智能开放平台**建设。推进面向行业企业智能服务应用。面向金融、电信、能源等行业，推动智能信息技术服务平台的建设应用，构建**智能服务体系**。

——《“十四五”软件和信息技术服务业发展规划》

来源：中华人民共和国工业和信息化部

平台
建设

北京市出台政策促进大模型研发

算力
中心

将新增算力建设项目纳入算力伙伴计划，加快推动海淀区、朝阳区建设北京人工智能**公共算力中心**、北京数字经济算力中心，形成规模化先进算力供给能力，支撑千亿级参数量的大型语言模型、大型视觉模型、**多模态大模型**、**科学计算大模型**等研发。

——《北京市促进通用人工智能创新发展的若干措施》

来源：北京市人民政府

大模型
赋能

各省份激励AI基础软件产业发展

各省份相关政策名称

来源单位名

《数字经济全产业链开放发展行动方案》	北京市
《建设杭州国家人工智能创新应用先导区行动计划(2020-2024年)》	浙江省
《数字经济发展规划(2020-2025)》	湖南省
《广东省数字经济发展指引1.0》	广东省
《智慧海南总体方案(2020-2025年)》	海南省
《上海市促进人工智能发展条例》	上海市
《“十四五”新型基础设施建设规划》	四川省

来源：弗若斯特沙利文

开源演进为AI基础软件市场提供良好的生态土壤

世界开源发展史

开源概念最早可以追溯到1970年，这一年也是“UNIX元年”。

1985年，Richard Stallman发表《GUN宣言》，吸收“前UNIX”和BSD的开源成果，开发推出Emacs编译器自由软件（Free Software）。同年，自由软件基金会（FSF）正式成立。

1991年，Linus Torvalds在Minix和GUN成果的基础上开发了Linux操作系统（Linux 0.01）。

从中国开源发展阶段的维度看开源发展的变化：

围绕产品操作系统及其生态，解决“缺芯少魂”短板；围绕基于开源的深度信息技术（大数据、云原生、区块链、人工智能等）发展与应用，为第四次世界工业革命做准备；围绕开发/编制开源软件供应链以保障其安全，并保障其上游的网络安全，国家关键基础设施的安全。

中国开源发展史

1991年，中国与AT&T Bell labs USL/USG合作，引进UNIXSVR 4.2版本源代码，并发布了中文版本，合作组建了中国UNIX公司。

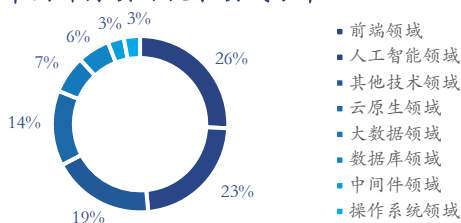
1999年，中科红旗、中软网络在引进Red Hat公司Linux发行版的基础上，分别推出了最早的Linux中文版本。

1999年至今，中国开源已经发展了32年。这32年的开源发展大致可以分为三个阶段：

第一阶段，主要围绕企业产品的操作系统及其生态建设；
第二阶段，主要结合研发基于开源的深度信息技术（如大数据、云原生、区块链、人工智能等）及其应用；
第三阶段，在经济双循环基础上规范建设或改造我国的供应链并主要在促进产业链、供应链数字化的基础上，采用取物料表格格式样的开源代码，保障其安全。

中国开源发展现状

中国开源项目技术领域分布



数据说明：中国开源项目技术领域分布共统计931个项目，代码托管平台包括GitLab/Gitee/GitLab/openl启智社区等

中国开源项目贡献者和星量统计

中国开源生态图谱中项目共有

58872

位贡献者 注：此数据未进行去重

中国开源生态图谱中项目共获得

2934777

颗星（Star）

数据说明：中国开源项目贡献者和星量统计数据来源于托管在GitLab上的870个项目3月数据统计

开源对AI基础软件的意义

□ 降低AI基础软件的开发门槛

开源代码通常是免费提供的，企业和个人可以不用支付高昂的许可费，降低了获取成本。开源的文档使得开发者可以更容易地学习和理解AI算法和模型，降低了开发门槛，使更多的人能够参与到开发中。

□ 促进开发者的知识共享和协作

在AI基础软件领域，开源项目允许研究人员、工程师和开发者共享他们的算法、模型和技术，鼓励全球志愿者协助开发、修复其源代码，凝聚人才，集结大众智慧，促进了全球范围内的协作和创新以及集体智慧的发挥。

□ 扩大AI基础软件的应用领域

开源协同建设生态和供应链，加速AI开发创新。开源已成为全球流行的创新和协同模式。开源为广大开发者提供了高度灵活和可定制的基础，使他们能够根据具体应用场景来定制和优化解决方案，满足各种不同行业和需求的需求。

来源：弗若斯特沙利文

位于产业链中游的AI基础软件具有极高的附加值

AI基础软件主要位于产业链的中游。其产业链由上游基础设施与资源提供商、中游AI基础软件平台与一站式AI开发平台和下游应用领域企业构成。上游是人工智能软件部署的基础，为AI基础软件提供算力支撑。中

游是产业链的核心，以模拟人的智能相关特征为出发点，构建应用技术路径，主要包括以基础开源框架、技术开放平台为代表的AI基础软件。下游主要为AI基础软件在各细分场景的应用，主要包括制造、安防、金融、医疗、零售、交通、等领域。

AI基础软件产业链



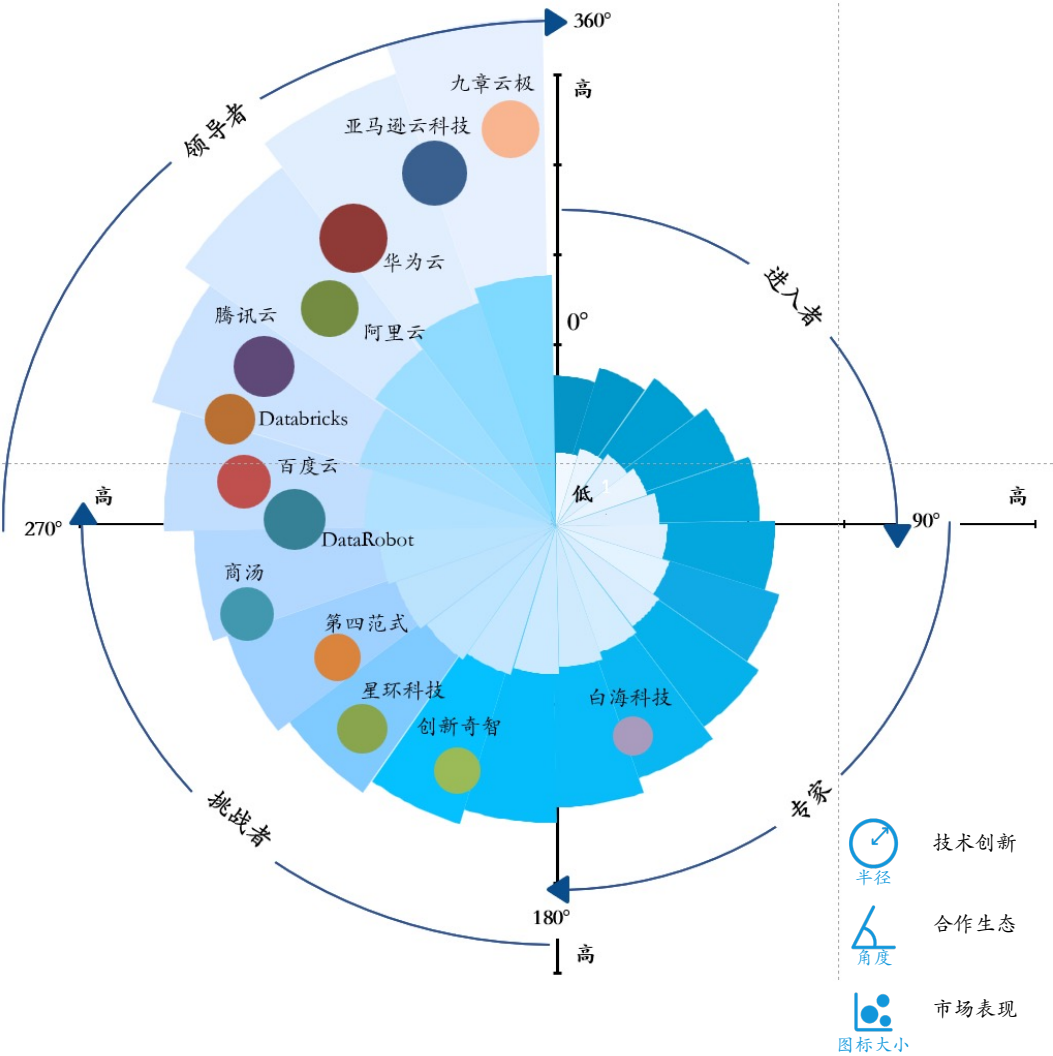
上游基础运算加速升级，赋能AI基础软件发展。基础运算作为AI基础软件的产业上游，主要为人工智能提供算力和数据的支持。中游应用技术趋于成熟，行业格局基本形成，主要包括AI基础软件平台与一站式AI开发平台。云厂商倾向于向客户提供端到端的整体解决方案，**期望整体捆绑客户**。相比之下，目前领先的AI基础软件平台能够以模块化的方式嵌入客户云的体系中，能够充分满足客户的自主性和可控性的要求，例如，九章云极采用“云中云”的战略部署，将自身软件

技术嵌入天翼云的生态当中，以独立第三方的身份专注提供AI服务，对于大型客户来说更可控更独立。下游行业与应用场景深度融合，赋能行业加速发展。目前，主流AI基础软件已逐步应用于各类行业场景中，模拟人类解决实践问题。ResNet、GPT-3等模型不断提升视觉处理、阅读理解等基础智能任务水平，语音识别、自然语言处理和计算机视觉等感知类任务上的应用技术成就可直接应用于实践产品，带来了广阔产业应用前景。

来源：弗若斯特沙利文

章节三 AI 基础软件厂商竞争格局

AI基础软件厂商竞争力图谱



来源: 弗若斯特沙利文

评估标准及主要成果

评估标准

- **AI基础软件厂商的“技术创新”**用圆心到坐标原点之间的距离（半径）来表示：距离衡量AI基础软件厂商在大数据管理、云安全治理、云原生架构、产品应用能力、可视化开发和底层框架等方面的技术能力，半径越大，AI基础软件厂商的技术能力越强。
- **AI基础软件厂商的“合作生态”**用圆心与坐标原点到Y正轴的夹角（角度）表示：角度衡量AI基础软件厂商在国内市场各应用领域（主要包括政务、金融、医疗、交通、制造等）方面的合作伙伴和头部厂商的数量，角度越大，AI基础软件厂商的合作生态力越强：360度即表示合作生态力满分。
- **AI基础软件厂商的“市场表现”**用图标大小来表示（图标大小）：图标大小衡量AI基础软件厂商在应用领域广度、易用性、客户服务和定价策略的市场表现力，每家厂商的竞争力都适用于代表不同服务能力的四个规模等级之一。



主要研究结果

九章云极、亚马逊云科技、华为云、阿里云、腾讯云、Databricks和百度云被评为国内AI基础软件市场的“领导者”。

- 这类厂商用大模型、云原生技术等提升应用数据安全、技术兼容性和AI基础软件开发部署能力，使应用体验更具安全性、先进性和易用性。此外，这些厂商还在各应用领域拥有强大的市场表现力，并始终如一地为客户提供完善的服务，以满足他们的需求。

DataRobot、商汤、第四范式、星环科技和创新奇智被评为国内AI基础软件市场的“挑战者”。

- 这类厂商通过增加产品应用价效比、区别于竞争对手以及提供更多的AI技术开放性来获得推动增长的力量。

白海科技

被评为国内AI基础软件市场的“专家”。

- 这类厂商都是利基型企业，具有开发优质的应用产品和技术创新的潜力。

来源: 弗若斯特沙利文

名词解释

ITU	<p>国际电信联盟：国际电信联盟是联合国负责信息通信技术（ICT）事务的专门机构，旨在促进国际上通信网络的互联互通并制定技术标准以确保实现网络和技术无缝互连。ITU从2016年开始开展人工智能标准化研究，目前是国际人工智能标准化领域的权威机构。</p>
ISO	<p>国际标准化组织：国际标准化组织的主要活动是制定国际标准，协调世界范围的标准化工作，以及与其他国际组织进行合作，共同研究有关标准化问题。ISO主要在工业机器人、智能金融、智能驾驶方面等开展了人工智能标准化研究，是国际标准化领域中一个十分重要的组织。</p>
GPU	<p>Graphics Processing Unit：GPU是图形处理器，又称显示核心、视觉处理器、显示芯片，是一种专门在个人电脑、工作站、游戏机和一些移动设备（如平板电脑、智能手机等）上做图像和图形相关运算工作的微处理器。GPU使显卡减少了对CPU的依赖，能很好地解决电影级图像质量需要解决的透明性、高质量反走样、运动模糊、景深和微多边形染色等问题。</p>
ASIC	<p>Application-Specific Integrated Circuit：ASIC是专用集成电路，是针对特定用户要求和特定电子系统设计、制造的专有应用程序芯片，其计算能力和计算效率可根据算法需要进行定制。ASIC芯片广泛应用于人工智能设备、虚拟货币挖矿设备、军事国防设备等智慧终端。</p>
IoT	<p>物联网：物联网是指互联网的延伸，其终端是嵌入式计算机系统及其配套的传感器，硬件或产品连上网所发生的数据交互称为物联网。</p>
CLIP	<p>Contrastive Language-Image Pre-training：CLIP是一种基于对比文本-图像对的预训练方法或者模型。CLIP的训练数据是文本-图像对，通过对比学习，模型能够学习到文本-图像对的匹配关系。CLIP包括两个模型：其中Text Encoder用来提取文本的特征，而Image Encoder用来提取图像的特征。</p>