

ChatGPT: 百度文心一言畅想

AIGC行业深度报告(4)

华西计算机团队

2023年3月2日

分析师: 刘泽晶

SAC NO: S1120520020002

邮箱: liuzj1@hx168.com.cn

核心逻辑:

- ◆ **国产“ChatGPT”扬帆启航。** OpenAI的商业模式为API接口收费。我们认为此种商业模式具有“卡脖子”的风险，因此我国需要发展自主可控的“ChatGPT”。国产生态正在逐步繁荣，百度打响国产ChatGPT领域“第一枪”，其在算法、算力、数据、生态、平台五方面皆有储备；ChatGPT的竞争本质即大模型储备竞赛，大模型是人工智能发展的必然趋势，也是辅助式人工智能向通用性人工智能转变的坚实底座。大模型分为NLP(自然语言处理)、CV(计算机视觉)、多模态和科学计算四类。此外，中美科技巨头已经开启大模型储备“军备赛”。
- ◆ **百度文心一言，开启国产ChatGPT新征程。** 百度是少有大模型语言训练能力的公司，模型储备方面，百度实现了全生态布局。1、**NLP(自然语言处理)**，已经具备智能创作、摘要生成、问答、语义检索、情感分析、信息抽取等能力，且可以让机器人像人一样具有逻辑且自由对话；2、**CV(计算机视觉)**，可用于应用于图像分类、目标检测、语义分割等场景，此外还可以应用于文档、卡证、票据等图像文字识别和结构化理解；3、**跨境大模型**，可实现AI作画、场景融合视觉常识推理、跨模态图像检索、跨模态文本检索等多场景；4、**生物计算**，应用场景为蛋白结构预测和小分子药物研发等领域。
- ◆ **百度为国产ChatGPT“领军企业”，具有算力积累和生态优势。** **平台方面：**拥有自主生态的百度百舸·AI异构计算平台，具备高效率、多密度、高易用性、多场景部署、乐高式拼接等能力。**算力方面：**百度自身具有建设智能算力中心的实力，技术领先且自主可控，已有典型落地案例；服务器方面拥有自研的昆仑芯云服务器；**芯片方面**，昆仑芯AI芯片是百度自主研发的芯片，2代芯片已量产，具备算力支撑强、高速互联等多重优势。**生态：**百度大模型赋能千行百业，已有落地应用，合作厂商分别覆盖科技、金融、航天、影视、汽车、电子制造等诸多产业。此外，我们推测ChatGPT有望成为搜索引擎的流量入口，百度搜索引擎有望借助文心一言大模型的能力重回巅峰。此外，目前国产科技巨头已经开启大模型的“军备竞赛”，因此，我们判断，未来AI+有望赋能千行百业，具有AI+能力的厂商有望呈现“百花齐放”的态势。
- ✓ **投资建议：**关注三条投资主线 1)具备算力基础的厂商，重点推荐**中科曙光**，其他受益标的为**寒武纪、商汤、海光信息、浪潮信息、景嘉微、拓维信息、神州数码、龙芯中科**；2)具备AI算法商业落地的厂商，重点推荐**科大讯飞、拓尔思**，其他受益标的为**海天瑞声**；3) AIGC相关技术储备的应用厂商，受益标的为：**百度、同花顺、三六零、金山办公**。
- ◆ **风险提示：**核心技术水平升级不及预期的风险、AI伦理风险、政策推进不及预期的风险、中美贸易摩擦升级的风险。



目录

01 国产ChatGPT扬帆启航

02 百度文心一言，开启国产ChatGPT新征程

03 投资建议：梳理AIGC相关受益厂商

04 风险提示



01 国产“ChatGPT”扬帆启航

1.1 ChatGPT为API接口收费，我国需自主可控

- ◆ **ChatGPT, AI的旷世之作, 持续引爆市场:** 是OpenAI于2022年11月推出的聊天机器人, 由于其功能强大, 例如实现文章创作、代码创作、回答问题等功能, 我们认为其具有跨时代的意义, 例如实现从“模糊搜索”到“精准推送”的跨越, 因此持续引爆市场。
- ◆ **ChatGPT为AI接口收费, 我国需自主可控:** OpenAI (ChatGPT母公司) 的商业模式为API接口收费, 客户可以通过接入OpenAI的API接口获取获取相关图像、语言、代码调整服务, 我们认为此种商业模式具有“卡脖子”的风险, 一旦海外禁止API接口权限, 会对我国AIGC生态造成严重影响, 因此我国需要发展自主可控的“ChatGPT”。
- ◆ **国产生态正在逐步繁荣, 百度打响“ChatGPT”领域“第一枪”:** 百度是少有预训练模型(大模型)语言训练能力的公司, 其在算法、算力、数据、生态、平台五方面皆有储备, 根据百度官方公众号, 百度计划于2023年3月16日在北京总部召开新闻发布会, 围绕国产版ChatGPT文心一言, 我们认为此举正式意味着我国自主可控的“ChatGPT”即将问世。

ChatGPT在回答“费马小定理”

样品

在以下示例中, ChatGPT能够理解对上一个问题(“费马小定理”)主题的引用(“它”)。

示例 / 43个 上一个 下一个

用户 什么是费马小定理

聊天GPT 费马小定理指出, 如果 p 是质数并且一种是一个不能被整除的整数 p , 然后一种(一种权力 p)是全等的一种取模 p . 换句话说, 如果我们划分一种经过 p , 余数总是一种.

例如, 如果 $p = 5$ 个和一种 $= 2$ 个, 然后 $2^5 \uparrow = 32$, 和 32 除以 5 个有剩余的 2 个. 所以, $2^5 \uparrow$ 是一致的 2 个取模 5 个, 费马小定理适用于这种情况.

费马小定理常用于密码学和其他需要快速高效地执行模算术运算的应用程序. 它也是证明数论中其他定理的有用工具

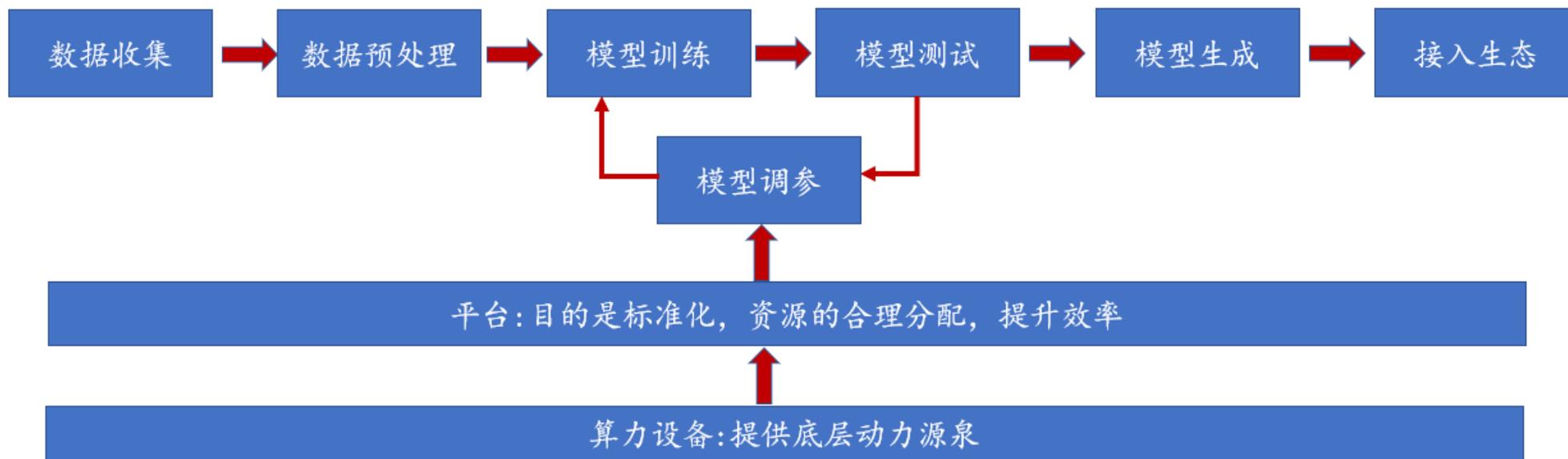
百度官方预告



1.2 ChatGPT的竞争本质即大模型储备竞赛

- ◆ **大模型的是人工智能发展的必然趋势**：大模型即“大算力+强算法”结合的产物。大模型通常是在大规模无标注数据上进行训练，学习出一种特征和规则。基于大模型进行应用开发时，将大模型进行微调，如在下游特定任务上的小规模有标注数据进行二次训练，或者不进行微调，就可以完成多个应用场景的任务。
- ◆ **大模型是辅助式人工智能向通用性人工智能转变的坚实底座**：大模型增强了人工智能的泛化性、通用性，生产水平得到质的飞跃，过去分散化模型研发下，单一AI应用场景需要多个模型支撑，每个模型需要算法开发、数据处理、模型训练、参数调优等过程。大模型实现了标准化AI研发范式，即简单方式规模化生产，具有“预训练+精调”等功能，显著降低AI开发门槛，即“低成本”和“高效率”。
- ◆ **数据、平台、算力是打造大模型生态的必备基础**：**数据**是训练大模型参数的必备，我们可以理解成“燃料”；**算力**是训练大模型的底层动力源泉，一个优秀的算力底座在大模型(AI算法)的训练和推理具备效率优势；**平台**是大模型和算力之间的“桥梁”，可针对不同的模型和硬件，实现资源的合理分配，达到软硬件的最优组合，从而大幅提升训练模型的效率。

数据、平台、算力、算法关系示意图



1.3 大模型带来的AI技术与应用变革潜能被广泛验证

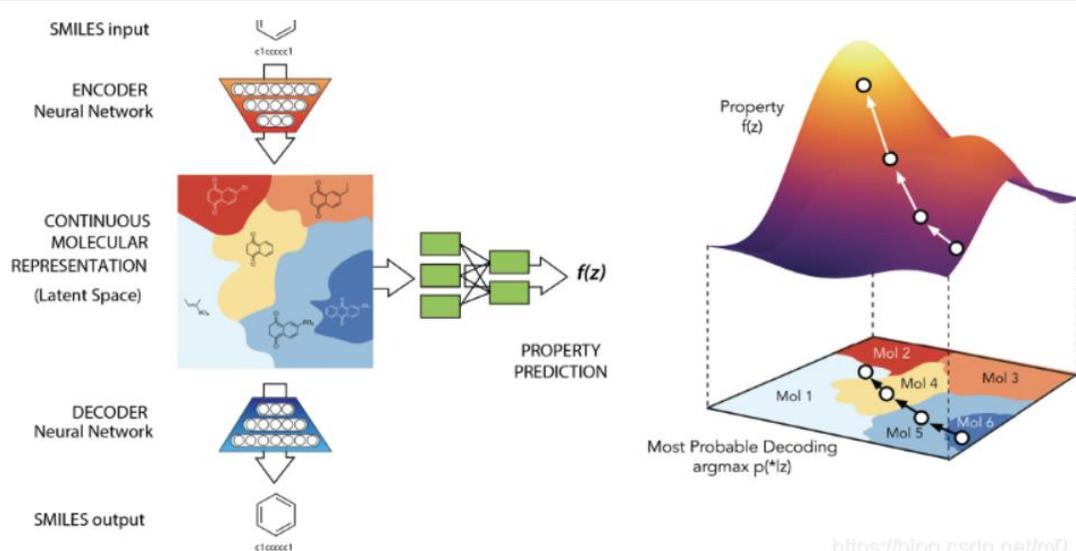
- ◆ **大模型带来的AI技术与应用变革潜能被广泛验证，可以分为四类，分别是NLP(自然语言处理)、CV(计算机视觉)、多模态和科学计算。**
- ✓ **NLP:** 近年来，随着预训练技术(大模型)、算力提升以及NLP海量数据和任务特性，大模型预训练在该领域取得显著突破，2018年，随着BERT(谷歌)的诞生，是利用海量无标注文本的监督学习，已经刷新多个AI权威技术榜单，3亿参数的BERT模型在权威GLUE的11任务刷新纪录，基准值推至80.4%，绝对值提升7.6个点。
- ✓ **CV:** 目前，主要以卷积神经网络(CNN)和Transformer为支撑的计算机大模型快速发展，比如，2021年的150亿参数的V-MoE推出，在ImageNET准确率高达90.35%，此外，国内厂商也逐渐发力，盘古CV大模型可提供OCR文字识别，并在多个场景落地。
- ✓ **多模态:** 大模型技术推动多模态模型不断迭代升级，比如阿里大模型M6，模型参数10万亿级别，持续拓宽大模型应用广度，覆盖电商、智能交互等场景，此外，OpenAI的CLIP(文本图像匹配)以及DALL·E2(文生图)持续落地。
- ✓ **科学计算:** AI+科学计算持续引发大变革，比如生物制药、气象预报、地震探查、材料等领域，例如Deep-Mind推出的AlphaFold2能够覆盖98.5%的人类蛋白质组，并对20种其他生物蛋白结构进行预测。

NLP和CV的发展现状与挑战对比

	NLP	CV
现状	分别在语言理解与生成、智能创作、机器翻译、智能对话、知识图谱和定制化语言解决方案落地应用，整体算法发展顺利，数据来源可获得性较强	2D数据工业质检、智慧城市落地完善，应用场景多、可商业化市场大，拥有最佳实践；人脸、OCR识别发展较为成熟
挑战	语言的歧义、文化差异及多样化、情感分析困难	3D/4D数据识别面临变形、光照、遮挡等可以依靠大规模预训练模型解决部分痛点的问题；数字人、数字孪生的数据获取困难，算法处理复杂
预期未来发展	以多个数据信息维度约束来验证情感分析及文本分析的准确性	打通数据融合以突破3D/4D获取瓶颈

资料来源： IDC, CSDN, 华西证券研究所

科学计算中药物发现和合成化合物示意图



https://blog.csdn.net/m0_46510245

1.4 中美科技巨头厂商开启大模型储备“军备赛”

国内外科技公司部分模型储备及应用场景

- ◆ **美国科技巨头公司开启AI大模型“军备赛”：**
- ◆ **谷歌：**谷歌推出聊天机器人Bard，底层代码为LaMDA，我们认为LaMDA与ChatGPT算法具备一战之力。此外，BERT算法具备库库时代的意义。
- ✓ **微软：**目前，根据智东西报道，微软推移动版Bing，语音接入、AI群聊等功能。此外，微软本身也有在NLP、CV相关大模型的技术储备。
- ✓ **其他：**例如FaceBook、亚马逊、DeepMind、英伟达等厂商已经加入大模型的“军备赛”，并分别在NLP、CV或多模态方面已有相应的技术储备，应用在语言生成、推理、代码生成、跨模态搜索等领域中。
- ◆ **我国需有自主可控AI大模型，生态正逐渐繁荣：**
- ✓ **百度：**是少有预训练模型(大模型)语言训练能力的公司，已经经历多次迭代，目前已覆盖众多方向，目前已有近百万开发者使用文心大模型，生态正在逐步繁荣，合作厂商覆盖科技、教育、工业、媒体、金融等诸多产业。
- ✓ **阿里：**根据钛媒体数据，M6模型是中文多模态模型，参数规模高达万亿，已在超40个场景中应用，可以实现剧本创作等功能。
- ✓ **其他：**例如腾讯、京东、科大讯飞、字节跳动、网易等纷纷加入“军备赛”，分别拥有独特技术架构，应用在智慧音效、AI视频创作、AI语音、AI作文、AI搜索等应用场景中。

厂商	AI大模型	参数	领域	应用场景
谷歌	BERT	4810亿	NLP	语言理解与生成
	LaMDA		NLP	对话系统
	PaLM	5400亿	NLP	语言理解与生成、推理、代码生成
	Imagen	110亿	多模态	语言理解与图像生成
微软	Parti	200亿	多模态	语言理解与图像生成
	Florence	6.4亿	CV	视觉识别
Facebook	Turing-NLG	170亿	NLP	语言理解、生成
	OPT-175B	1750亿	NLP	语言模型
Deep Mind	M2M-100	150亿	NLP	100种语言互译
	Gato	12亿	多模态	多面手的智能体
	Gopher	2800亿	NLP	语言理解与生成
OpenAI	AlphaCode	414亿	NLP	代码生成
	CLIP & DALL-E	120亿	NLP	图像生成、跨模态检索
	Codex	120亿	多模态	代码生成
百度	ChatGPT	-	NLP	语言理解与生成、推理等
	NLP大模型	千亿级别	NLP	语言理解、生成
	CV大模型		CV	图像识别
	跨模态计算大模型		多模态	语言理解与图像生成
生物计算大模型	CV		化合物表征学习、分子结构预测	
阿里巴巴	M6	万亿级别	多模态	语言理解与图像生成
腾讯	混元大模型	-	NLP	语言理解与生成
京东	K-PLUG	-	NLP	语言理解与生成、推理、代码生成
三六零	-	-	NLP	智能搜索
字节跳动	DA	-	NLP	语言理解
科大讯飞	中文预训练模型	-	NLP	语言理解与生成、语言互译

资料来源：公开资料整理，腾讯《AIGC发展趋势报告2023》，华西证券研究所



02 百度文心一言，开启国产ChatGPT新征程

2.1 百度文心一言开启国产ChatGPT新征程

- ◆ **百度是少有大模型语言训练能力的公司**：其文心大模型和Open AI的GPT模型类似，在2019年就已经推出，并且已经迭代了多代，从单一的自然语言理解延伸到多模态，包括视觉、文档、文图、语音等多模态多功能，因此“文心一言”所基于的ERNIE系列模型也已经具备较强泛化能力和性能。
- ◆ **我们认为百度作为国产ChatGPT“领军企业”，符合数据、平台、算力、大模型、生态多重要素需求**：**算力方面**，百度拥有自主生态的算力底座，比如百度智算中心；**AI芯片方面**，坐拥昆仑芯加速器，实现了核心自主可控，且性能优异；**平台方面**，百度具备百度百舸·AI异构计算平台，具有高性能、高弹性等优势；**大模型储备方面**，百度实现了全生态的布局，其中包括NLP、CV、大模型、生物计算等领域，且在不同场景中已有较多应用。**数据方面**，根据IDC数据，其拥有5500亿条知识，且已经应用于百度搜索、信息流、智能驾驶、百度地图、小度等产品。

百度文心全景图



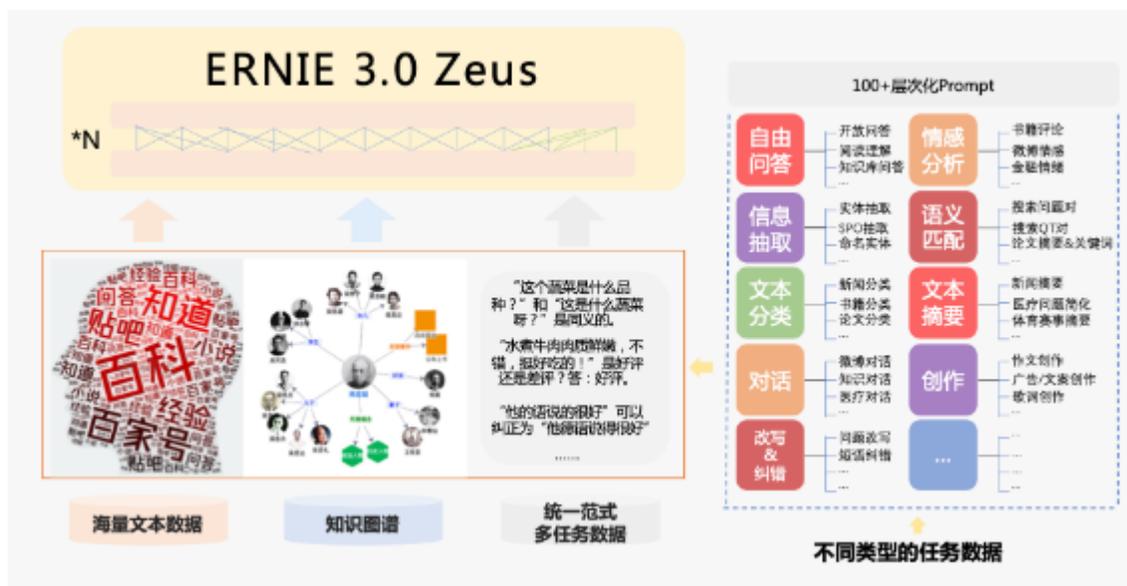
百度文心全景图



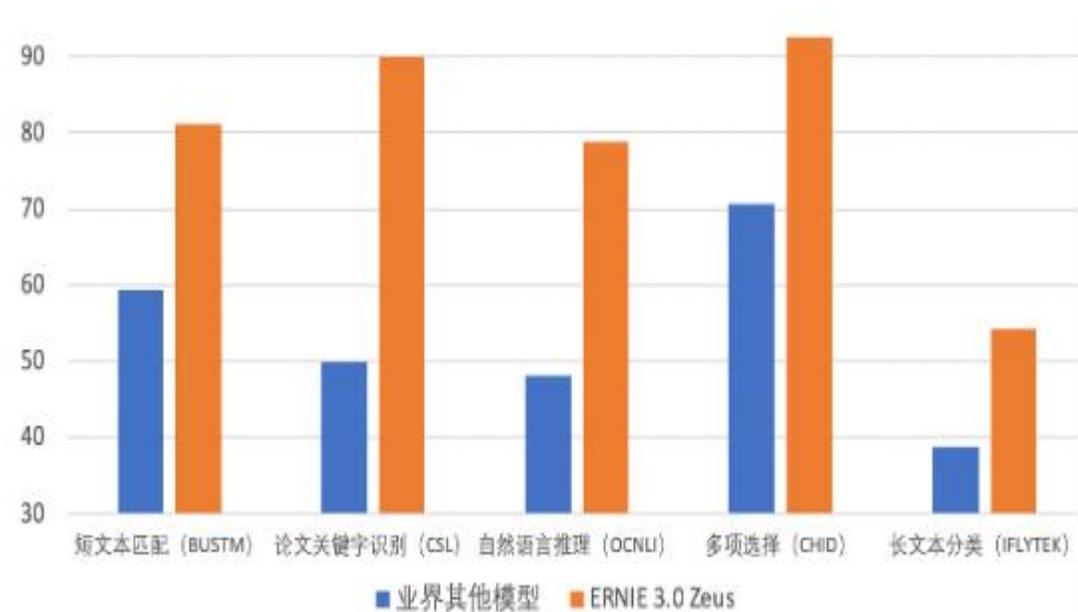
2.2.1 百度文心一言大模型储备齐全：NLP(自然语言处理)

- ◆ **百度文心一言NLP方向算法储备齐全：**其中著名的 ERNIE系列，是基于知识增强的千亿模型，用于智能创作、摘要生成、问答、语义检索、情感分析、信息抽取、文本匹配、文本纠错等各类自然语言理解和生成任务，并且模型已经可应用于医疗、金融、图语言、编程、跨模态、信息抽取等各个方面。此外，PLATO模型，是全球首个超百亿参数规模的中英文对话训练模型，可以让机器人像人一样具有逻辑且自由对话。
- ◆ **百度文心一言NLP功能强大：**例如ERNIE系列，ERNIE3.0基于知识增强的多范式统一预训练框架，深入融合的千亿级知识，具备强大的语言理解能力与小说、摘要、文案创意、歌词、诗歌等文学创作能力。其中与鹏城实验室合作发布了知识增强千亿大模型“鹏城-百度·文心”。目前文心ERNIE已经刷新93个中文NLP任务基准，并多次登顶SuperGLUE全球榜，已在机器阅读理解、文本分类、语义相似度计算等60多项任务中实际应用。

ERNIE 3.0 Zeus算法框架



ERNIE 3.0 Zeus在公开数据集上零样本/小样本学习的效果



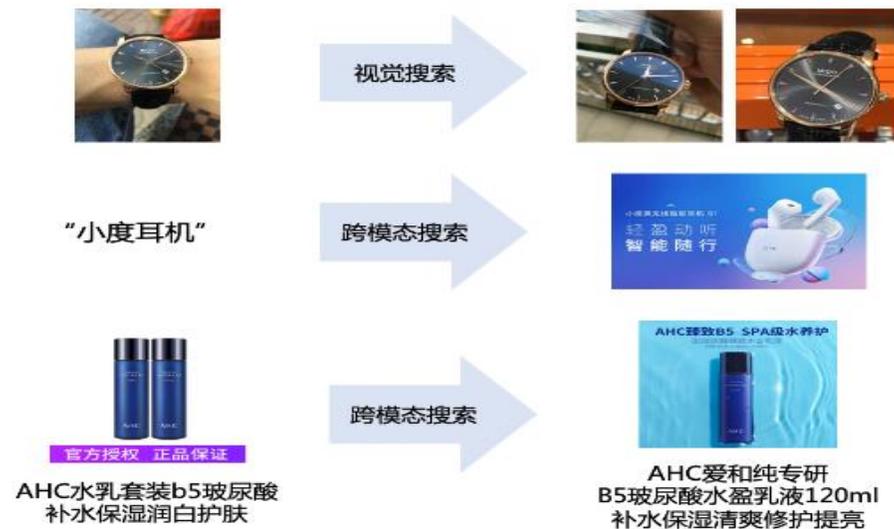
2.2.2 百度文心一言大模型储备齐全：CV(计算机视觉)

- ◆ **百度文心一言CV具有颠覆性**：VIMER-CAE：为视觉自监督预训练大模型，创新性地提出“在隐含的编码表征空间完成掩码预测任务”的预训练框架，在图像分类、目标检测、语义分割等经典下游任务上刷新SOTA结果。
- ✓ VIMER-UFO 2.0：多任务学习模型，行业最大170亿参数视觉多任务模型，覆盖人脸、人体、车辆、商品、食物细粒度分类等 20+ CV 基础任，具备支持各类任务、各类硬件的灵活部署等优势，可以有效解决大模型参数量大，推理性能差等问题。
- ✓ OCR -VIMER-StrucTexT 2.0：为表征学习预训练模型解决了训练数据匮乏和传统 OCR + NLP 链路过长导致的模型表达能力不足、优化效率偏低等问题，能够广泛应用于文档、卡证、票据等图像文字识别和结构化理解，例如泛卡证票据信息抽取应用、政务办公文档还原应用等场景。
- ✓ VIMER-UMS：是行业首个统一视觉单模态与多源图文模态表征的商品多模态预训练模型，可实现统一图文表征预训练同时覆盖商品视觉单模态、多模态识别与检索任务，可以显著改善商品视觉检索和商品多模态检索体验。

OCR -VIMER-StrucTexT 2.0典型应用场景



百度CV大模型VIMER-UMS功能介绍统一商品视觉与图文表征



2.2.3 百度文心一言大模型储备齐全：跨模态、生物计算

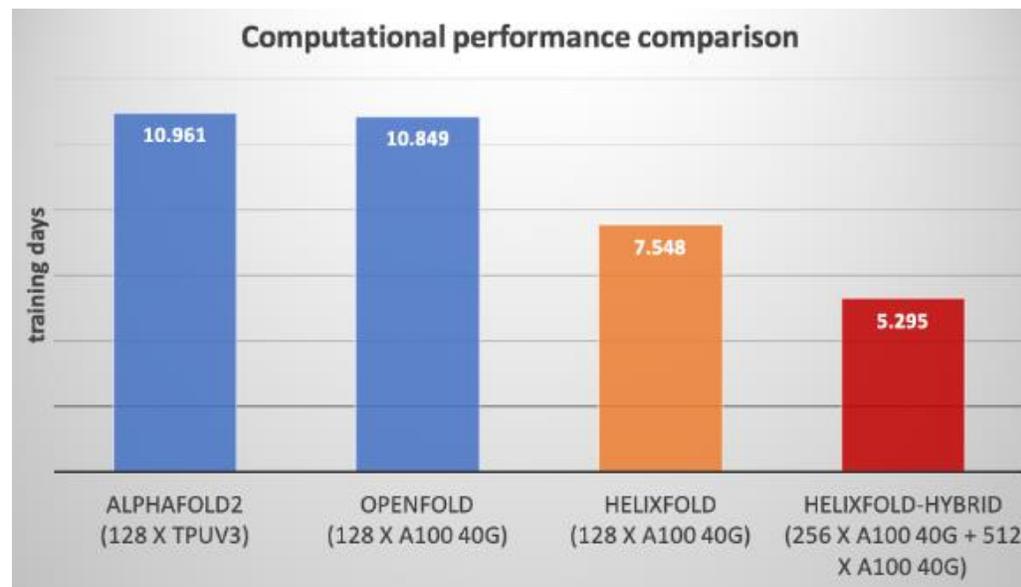
- ◆ **文心跨境大模型优势显著**：ERNIE-ViLG2.0是知识增强的 AI 作画大模型，在公开权威评测集MS-COCO上取得了当前该领域的领先效果，在语义可控性、图像清晰度、中国文化理解等方面均展现出了显著优势；跨模态文档智能大模型ERNIE-Layout，基于布局知识增强技术，融合文本、图像、布局等信息进行联合建模，在文档抽取、布局理解等5类11项任务刷新业界SOTA；ERNIE-ViL是首个只是业界首个融合场景图知识的多模态预训练模型。在视觉常识推理、跨模态图像检索、跨模态文本检索等 典型多模态任务中刷新了世界记录。
- ◆ **文心生物计算大模型在全球具有领先优势**：HelixFold-Single是秒级别的蛋白结构预测大模型，从近3亿的无标注蛋白质数据中隐式的学习MSA信息，在90% 的单体蛋白场景上预测效果持平AF2，在抗体结构预测场景下，比AlphaFold2预测结果更优；HelixGEM-2**为小分子药物研发模型**，在量子化学属性预测和虚拟筛选双场景上达到领先效果，其上个版本是业界首个基于几何构象增强的化合物表征模型，引入化合物的三维空间信息，在14个药物属性预测相关的benchmarks上效果达到业界最优；HelixFold**模型可以端到端地学习蛋白质结构**，在国产DCU环境下训练千万级别蛋白仅需到2.6天，在 GPU 上相同硬件环境配置下，训练性能和部分场景效果显著优 AlphaFold2。

文心 ERNIE-ViLG 2.0自动生成创意图片



资料来源：公开资料整理，腾讯《AIGC发展趋势报告2023》，华西证券研究所

百度HelixFold蛋白质结构模型与世界经典模型训练性能对比



2.3.1 百度底层算力技术实力强劲: 百度智算中心

- ◆ **百度智算中心是数字经济的重要底座**: 百度自身具有建设智能算力中心的实力, 百度智算中心面向人工智能应用场景, 为政府和行业客户提供普惠算力、算法模型和数据服务的人工智能基础设施。可面对人工智能应用场景, 支持大规模部署的同时, 满足高并发、高弹性、高精度等不同计算需求, 可支撑城市大脑、产业金融、自动驾驶等各个垂直化行业。
- ◆ **方案优势强大, 具备多重领先优势**: 1、**技术领先**, 现有千卡并行环境下实现加速比90%, 可支持大规模模型训练场景; 2、**绿色低碳**, 数据中PUE可以低至1.08, 左到低能耗、高性能运行; 3、**自主可控**, AI芯片、AI框架、开发平台、大模型层面、可以做到全栈自研自主可控; 4、**开放生态**, 飞桨已累计凝聚477万开发者等。
- ✓ 目前已在江苏盐城、湖北宜昌等地落地, 建设内容包括智算中心、城市级视频中台、数据中台系统、核心物联网通信系统、大规模图谱分析体系, 以及超过270类不同人工智能算法模型。

百度智算中心架构



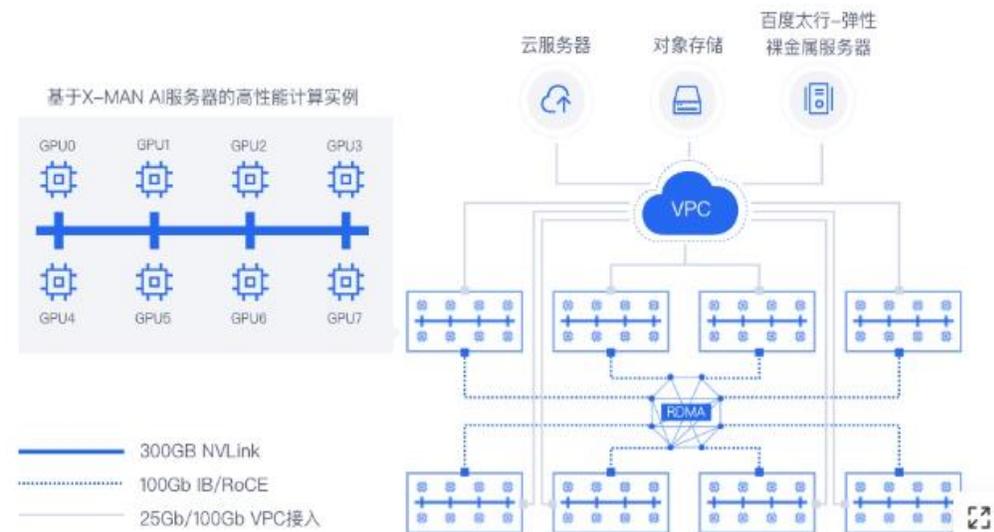
2.3.2 百度底层算力技术实力强劲:百度百舸·AI异构计算平台

- ◆ **百度AI异构计算平台技术积累浓厚:**包含AI计算、AI存储、AI加速、AI容器四大核心套件，具有高性能、高弹性、高速互联、高性价比等特性，其中平台已经充分汲取了百度多年技术积累，具备深度融合推荐、无人驾驶、生命科学、NLP等场景的实践经验，能为AI场景提供软硬一体解决方案，加速AI工程化落地。
- ◆ **百度智能计算平台具备多重优势:** 1、**高效率**，提供AI超级服务器、RDMA网络、大吞吐低延时的并行文件系统，可大幅提升训练速度和推理效率；2、**高密度**，支持GPU资源共享与隔离、架构感知调度，提升异构资源的容器化部署密度，提升资源利用率；3、**良好易用性**，支持PaddlePaddle、TensorFlow、Pytorch等多种主流深度学习框架；4、**多场景部署**，可根据业务需求在不同场景部署落地，公有云、IDC等场景；5、**乐高式拼接**，AI计算、AI存储、AI容器三大核心产品均可各自独立提供服务。
- ◆ **应用场景广泛:**可用于营销广告、无人驾驶、生物科技、语音语义、计算机视觉等场景中的训练及推理。

百度百舸 · AI异构计算平台架构



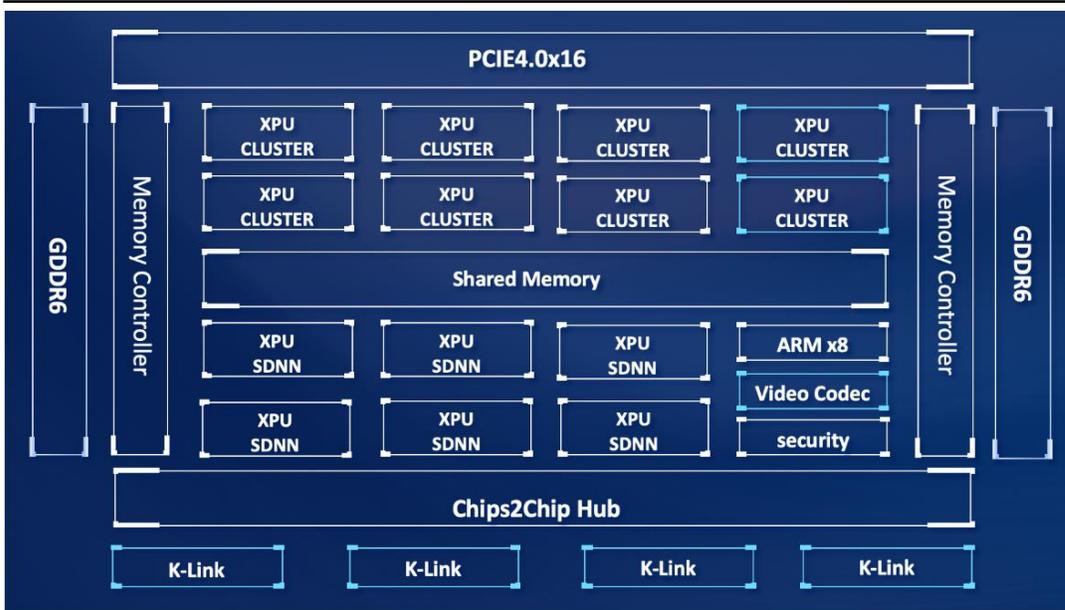
百度百舸 · AI异构计算平台AI计算示意图



2.3.3 百度底层算力技术实力强劲: 昆仑芯云服务器

- ◆ **百度昆仑云服务器服务器专为AI算力而生:** 昆仑芯云服务器是一种弹性按需、提供高性能通用AI算力的云服务器，应用于AI推理和AI训练加速。昆仑芯云服务器是搭载昆仑芯的云服务器，支持K100和R200型号。其中昆仑芯为自主研发的 AI 通用处理器芯片。其中R200 AI芯片采用7nm制成，广泛应用于计算机视觉、自然语言处理、大规模语音识别、大规模推荐等场景。
- ◆ **昆仑芯云服务器技术壁垒浓厚:** 1、**领先的AI计算能力**，基于7nm制程，单卡算力达到128TFLOPS，支持推理和训练；2、**生态兼容**，支持主流深度学习开发框架，例如Tensorflow、Pytorch、PaddlePaddle等；3、**具备硬件虚拟化能力**，优化了加速芯片的利用率，在保证延时和吞吐量的情况下支持推理和训练等混合工作负载；4、**易开发**，支持多种开发组件工具，编译引擎支持C和C++编程。
- ◆ **典型案例:** 1、**互联网搜索**，可支持搜索系统对高并发要求，提高NLP推理的效率；2、**超算中心**，实现人工智能算力大规模、高密度部署；3、**工业质检**，利用深度学习技术替代人工质检，大大节省了人力成本，提高整厂的智能化水平，帮助企业降低损耗，提升约10%的良品率；4、**智慧城市**，全方位支撑智慧政务、综合治理等智慧城市核心板块。

百度昆仑芯云服务器架构



资料来源：百度智能云官网，华西证券研究所

百度昆仑芯云服务器架构互联网搜索应用场景



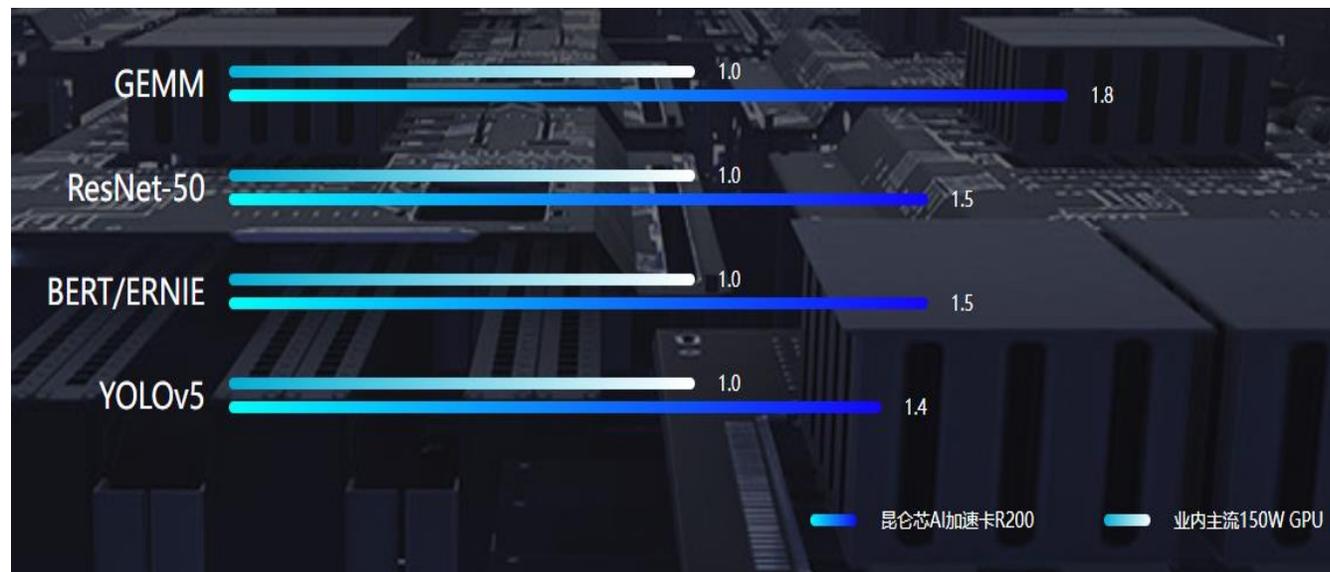
2.3.4 百度底层算力技术实力强劲: 昆仑芯AI芯片

- ◆ **百度自身AI芯片技术实力浓厚:** 昆仑芯AI芯片是百度自主研发的芯片, 昆仑芯科技前身是百度智能芯片及架构部昆仑芯科技深耕AI加速领域已十余年, 是一家在体系结构、芯片实现、软件系统和场景应用均有积累的AI芯片企业。昆仑芯1代是百度自研的第一一代昆仑芯片, 2020年底实现量产; 昆仑芯2代也已于2021年8月量产; 根据百度集团执行副总裁沈抖透露, 昆仑芯3代将于2024年初量产。
- ◆ **昆仑芯2代AI芯片可提供庞大算力支撑, 具有性能优势:** 采用7nm架构, 对比上一代, R200全面提升AI负载的运行效率, 巅峰算力可达256 TOPS@INT8, 昆仑芯2代具备多重优势, 1、**创新架构**, 通用和专用指令集融合, 支持训练、推理、虚拟化; 2、**算力支撑强大**, 通用计算能力明显增强, 可灵活支持AI算法的演进, 3、**高速互联**, 高性能分布式AI系统, 加速AI数据和模型并行中的数据交换。此外, 昆仑芯在著名算法的功耗比已有显著优势。

昆仑芯1-2代产品差异

型号	型号: K100 加速卡	型号: R200 加速卡
精度	INT4/8/16 XFP16/32	INT8/16/32 FP16/32
算力	INT8: 128 TOPS FP16: 32 TOPs FP32: 8 TOPS	INT8: 256TOPS FP16: 128TOPS FP32: 32TOPS
显存	8GB	16GB
访存带宽	256 GB/s	512 GB/s
系统互联	PCI-E Gen4 x 8, 兼容 3.0/2.0/1.0	PCI-E Gen4 x 8, 兼容 3.0/2.0/1.0
功耗	75W	150W

昆仑芯2代芯片与主流芯片对比



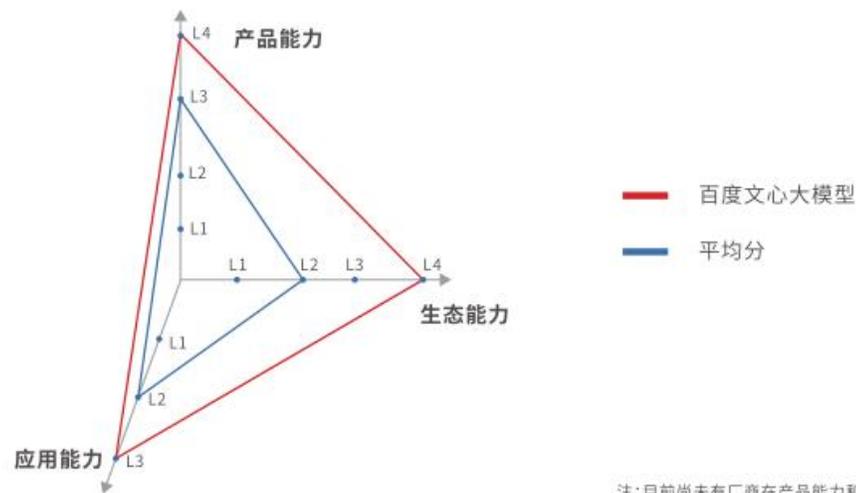
2.4 百度文心一言生态愈发繁荣

- ◆ **百度模型评估结果属于第一梯队，彰显其强大技术实力：**根据IDC的数据评估先实，百度文心大模型在市场格局中处于第一梯队，产品能力、生态能力达到L4水平，应用能力达到L3水平。产品能力呈现出较强技术实力和平台积累，“文心大模型+深度学习平台”创新了人工智能研发应用范式达到行业前端水平；**应用能力**方面，百度已在金融、能源、制造、城市、传媒、互联网等行业拥有实际落地的标杆案例，截止目前文心已累计发布11个行业大模型；在**生态能力**方面，百度文心大模型在社区用户的基础上，可以实现与开发者、行业用户、上下游产业的正向互动，在评估厂商中处于行业领先地位。
- ◆ **百度大模型赋能千行百业，已有落地应用：**目前百度文心在能源、金融、航天、制造、传媒、城市、社科以及影视等领域与国网、浦发、吉利、TCL、人民网、哈尔滨、上海辞书出版社等均有案例应用的行业大模型。这些行业大模型作为重要AI底座，在各行业的技术效果突破、产品创新、生产流程变革、降本增效等维度产生价值。

百度文心行业大模型全景图



百度文心大模型评估结果



来源：IDC

注：目前尚未有厂商在产品能力和生态能力方面达到L5，应用能力尚未有厂商达到L4，因此不在本次评估中凸显。

2.4 百度文心一言生态愈发繁荣

- ◆ 百度文心一言场景愈发繁荣，目前已有诸多厂商开展合作，我们认为其商业模式同样为API接口收费，属于SAAS商业模式，合作厂商分别覆盖科技、金融、航天、影视、汽车、电子制造等诸多产业。

百度文心一言合作部分场景

合作场景	合作公司	合作目的	模型效果	应用场景
电网	国网	国家电网与百度联合发布知识增强的电力行业大模型，建设更适配电力行业场景的AI基础设施，降低数据标注成本，提升细分场景模型效果	使用更少算力在更短的训练时间内即可取得良好的效果，模型综合指标达到91.18%，相比较传统方式提升11.38%。在小样本场景下的模型效果优势明显	设备技术标准知识库、电力知识问答系统、电力文档智能分析等多个能源行业应用场景
金融	浦发银行	浦发银行与百度在行业数据、大算力和AI算法上实现优势互补，联合研发面向金融行业的大模型，将在金融行业各类智能场景落地，降低金融AI应用落地的门槛	金融文本命名实体识别模型的F1值提升约3%，有效提升授信报告的处理效率；在金融语义相似度，金融事件主体抽取等一系列公开数据集上相对通用大模型取得显著提升	智能对话、智能文档分析、智能双录等多个金融行业应用场景
航天	航天工程	中国探月与航天工程和百度合作，让大模型从航天科研场景的数据和知识中融合学习，进而对航天数据进行智能采集、分析和理解，助力航天科研工作效率提升	在航天文献领域分类与航天技术文档匹配检索等下游任务上，效果综合提升约3%	航天故障部位信息抽取、航天工程文献情报分类、航天技术文献检索与摘要生成等多个航天领域应用场景
新闻媒体	人民网	对于新闻行业内容繁杂、信息提取难度极大等痛点，人民网与百度在传媒行业、大数据和AI算法上结合，引入人民网舆情数据中心积淀的行业知识，顺应内容与技术发展，联合研发知识增强的传媒行业大模型，打造全媒体时代内容科技创新引擎，加速传媒行业的智能化升级	大幅度提升传媒行业自然语言处理任务效果，在诸多媒体行业下游任务中得到了验证。其中，新闻摘要生成任务人工测评可用率提升约7%；在新闻内容审核分类、舆情分析等一系列媒体行业任务上相对通用大模型取得4-6%的提升	新闻内容分类、新闻情感分析、新闻摘要生成等多个媒体场景
影视	电影频道	电影频道与百度深度合作，通过发挥双方在视频修复数据、AI技术与大算力上的优势，全方位提升视频修复效率和观感体验的效果，以AI技术加速超高清视频产业的智能化发展进程	测试数据显示，使用电影频道-百度·文心大模型每天可修复视频28.5万帧，解决了绝大部分画面的修复问题。即便是需要进一步精修，修复速度也能提升3-4倍	划痕、噪点、抖动等常见画质损坏因素的修复及清晰度提升
汽车	吉利	立足于汽车行业知识庞杂、客服人力需求大等业务痛点，吉利与百度融合了双方在通用人工智能科研与汽车领域AI应用的优势，联合研发了聚焦于汽车行业的大模型，共同推进汽车产业通用智能化，在降低AI应用落地成本的同时，进一步提升汽车行业AI应用的效果与价值	共建模型的智能客服知识库扩充任务人工评估可利用率相对基线提升24.37%，能大幅提高问答的泛化水平，提高客服系统的智能化体验；除此以外，在车载语音系统短答案生成、汽车领域知识库构建等生成任务上的效果均有20%-35%的提升，可辐射至多个汽车行业场景	智能客服系统、车载语音系统、领域知识库构建等汽车行业应用场景
电子制造	TCL	面对电子制造行业产线繁多、质检工艺复杂且精度要求高的痛点，TCL和百度在行业数据、AI算法及大算力上实现优势互补，共同构建TCL-百度·文心电子制造行业大模型，不仅变革了传统质检算法流程，同时大幅提升了模型研发与运营效率，实现了成本的降低和业务效率的提升	使用TCL-百度·文心大模型，在TCL两个产线检测mAP指标平均提升10%+；训练样本减少到原有训练样本30%~40%，产线指标即可达到原有产线效果，新产线冷启动效率可提升3倍，产线上线开发周期降低30%	电子制造行业各种不同产线工艺场景的缺陷检测

资料来源：百度文心大模型官网，华西证券研究所

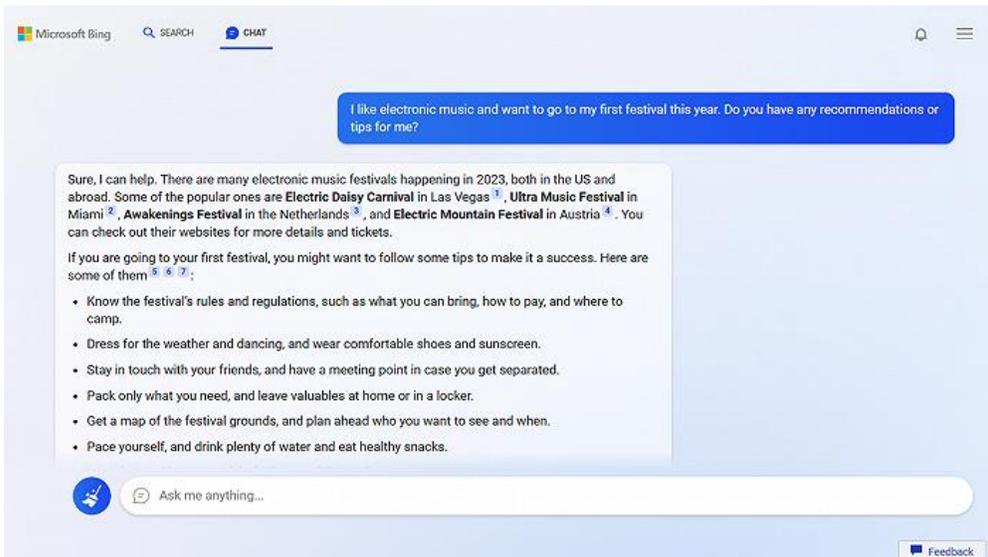
2.5 百度文心一言五大要素齐全，开启新篇章

- ◆ **我们认为大模型成功必备的五要素，分别是数据、算力、平台、模型和生态，而百度作为国产ChatGPT的领头羊，五大要素皆已配备。**
- ◆ **海量数据：**即训练大参数模型的“燃料”，根据IDC数据，其拥有5500亿条知识，且已经应用于百度搜索、信息流、智能驾驶、百度地图、小度等产品。
- ◆ **模型储备：百度实现了全生态的布局，**其中包括NLP、CV、跨境大模型、生物计算等领域。
 - ✓ **NLP领域：**ERNIE系列，是基于知识增强的千亿模型，用于智能创作、摘要生成、问答、语义检索、情感分析、信息抽取等能力。PLATO模型，是全球首个超百亿参数规模的中英文对话训练模型，可以让机器人像人一样具有逻辑且自由对话的功能。
 - ✓ **CV领域：**VIMER-CAE，应用于图像分类、目标检测、语义分割等场景；VIMER-UFO，拥有170亿参数，覆盖人脸、人体、车辆、商品、食物细粒度分类等 20+的视觉多任务模型；OCR -VIMER-StrucTexT广泛应用于文档、卡证、票据等图像文字识别和结构化理解；VIMER-UMS业内首发多源图文模态表征的商品多模态预训练模型；
 - ✓ **跨境大模型：**可实现AI作画、场景融合视觉常识推理、跨模态图像检索、跨模态文本检索等多场景。
 - ✓ **生物计算：**应用场景为蛋白结构预测和小分子药物研发等领域。
- ◆ **平台方面：**拥有自主生态的百度百舸 · AI异构计算平台，具备高效率、多密度、高易用性、多场景部署、乐高式拼接等能力。
- ◆ **算力底座：**百度自身具有建设智能算力中心的实力，技术领先且自主可控，已有典型落地案例；服务器方面拥有自研的昆仑芯云服务器，具有AI计算能力领先、生态兼容、兼具硬件虚拟化、易开发等能力；芯片方面，昆仑芯AI芯片是百度自主研发的芯片，2代芯片已量产，具备算力支撑强、高速互联等多重优势。
- ◆ **生态：**百度大模型赋能千行百业，已有落地应用，合作厂商分别覆盖科技、金融、航天、影视、汽车、电子制造等诸多产业。

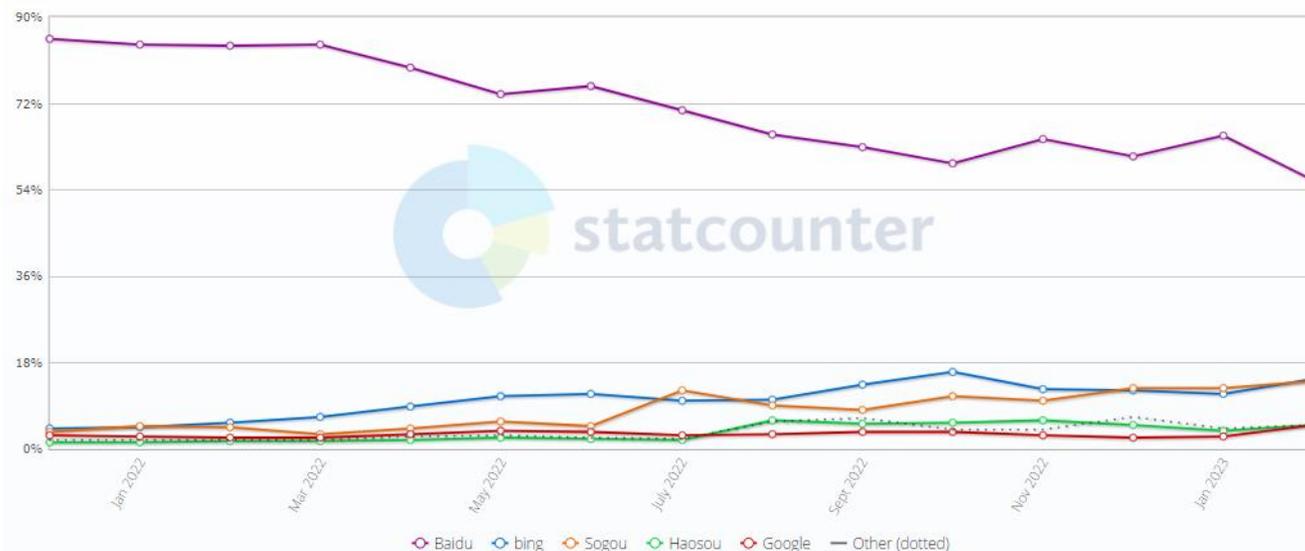
2.6 百度文心一言畅想，搜索引擎的下一章

- ◆ **ChatGPT有望重新定义搜索“入口”**：我们已经在《ChatGPT: 重新定义搜索“入口”》论证过ChatGPT有望成为下一代搜索引擎的催化剂，原因有二，第一，ChatGPT实现了从“模糊搜索”到“精准推送”跨越；第二，ChatGPT符合方便、快捷、便于交互、安全且生态完善四大关键因素。
- ◆ **商业巨头开启搜索引擎争夺战**：科技巨头微软和谷歌开启的下一时代搜索引擎争夺战也论证了我们的结论，新版Bing浏览器将有ChatGPT支持，可实现在线客户交流，此外，微软已经推出移动版Bing，包括安卓和IOS系统；谷歌浏览器已经开启新版谷歌发布会，并推出聊天机器人Bard，具有对话、多重搜索等功能，底层算法为LaMDA算法。
- ◆ **我们认为文心一言对百度搜索引擎而言“如虎添翼”**：根据Statcounter数据，2021年12月，百度搜索引擎国内市占率为85.48%，2023年百度市占率55.92%，虽依旧市占率国内第一，但是占比大幅下降。目前文心一言大模型已经具备语义检索、情感分析、逻辑思考、自然对话等功能，我们认为其具有新版Bing类似功能，参考新版Bing用户增值数量（根据腾讯网消息，微软公布ChatGPT版BING不到48小时，申请用户数量已经超过百万），我们推测百度搜索引擎市占率有望重回巅峰。

用户可以和新版BING搜索引擎交流



中国搜索引擎市场份额(2021年12月至2023年2月)



2.7 国产大模型“军备竞赛”已开启，AI+生态将百花齐放

- ◆ 目前国产科技巨头已经开启大模型的“军备竞赛”，且此类科技巨头已有部分AI+的落地，我们认为未来AI+将赋能千行百业，且具备AI+能力的厂商将处于百花齐放的状态。
- ✓ **阿里巴巴**：根据钛媒体表示，阿里巴巴达摩院在2020年初启动中文多模态预训练模型M6项目，2021年10月，M6的参数规模扩展到10万亿，成为当时全球最大的 AI 预训练模型。M6已在超40个场景中应用，日调用量上亿。在阿里云内部，M6大模型的应用包括但不限于在犀牛智造为品牌设计的服饰已在淘宝上线、为天猫虚拟主播创作剧本，以及增进淘宝、支付宝等平台的搜索及内容认知精度等场景。
- ✓ **京东将推出产业版ChatGPT，名为ChatJD**：据中国科技网2月10日消息，ChatJD旨在打造优势、高频、刚需的产业版通用ChatGPT。未来，ChatJD智能人机对话平台有望落实在零售、金融等领域，内容包括内容生成、人机对话、用户意图理解、信息抽取、情感分类等，在客户咨询与服务、营销文案生成、商品摘要生成、电商直播、数字人、研报生成、金融分析等领域将发挥广泛的落地价值。

京东ChatJD125计划图



资料来源：公开资料整理，中国科技，澎湃新闻，华西证券研究所

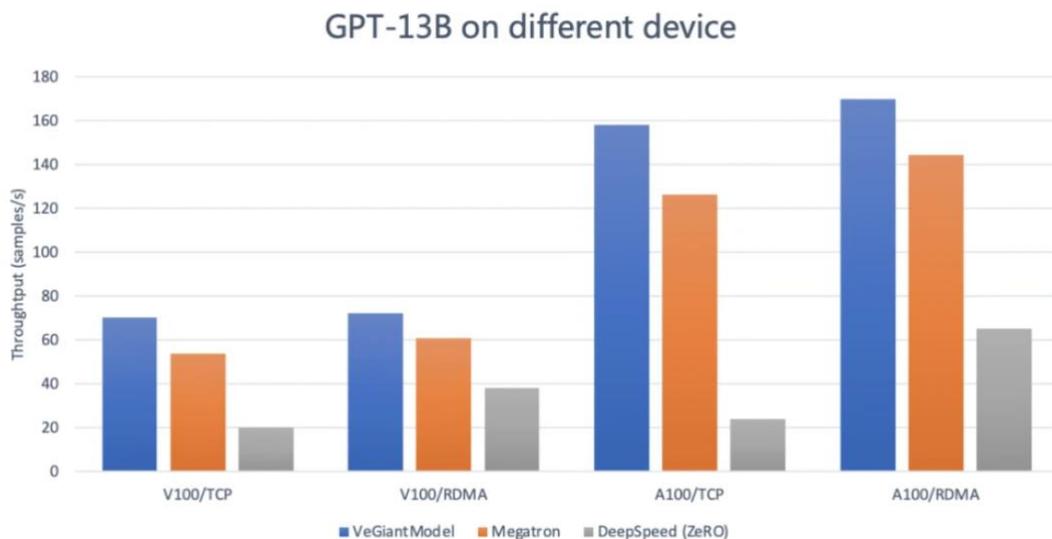
阿里版ChatGPT



2.7 国产大模型“军备竞赛”已开启，AI+生态将百花齐放

- ✓ **腾讯**:根据《AIGC发展趋势报告2023》在广告领域，腾讯混元AI大模型能够支持广告智能制作，即利用AIGC将广告文案自动生成广告视频，大大降低了广告视频制作成本。巨大的应用前景将带来市场规模的快速增长。
- ✓ **字节跳动**:字节跳动已经拥有强大的推送算法，此外根据腾讯网消息，为应对大模型对算力带来的显著压力(例如BERT、GPT-3等)，字节跳动 AML 团队内部开发了火山引擎大模型训练框架 veGiantModel，为应用于自然语言处理领域的大模型训练，最高可将大模型训练的性能提升6.9倍，大幅降低训练系统的压力。目前火山引擎已在机器学习平台上支持该算法，目前该平台正在公测中。
- ✓ **华为**:根据MindSpore介绍，早在2021年，华为诺亚方舟实验室、华为中央软件院、鹏城实验室以及北京大学相关技术团队组建的盘古α首次基于“鹏城云脑II”训练，该模型属于中文大规模预训练语言模型，主要用于中文理解与生成，参数数量超过2000亿。
- ✓ **其他**:科大讯飞、网易、三六零等国内科技巨头纷纷加入大模型技术战中，分别拥有独特技术架构，应用在智慧音效、AI视频创作、AI语音、AI作文、AI搜索等应用场景中。

大模型训练框架veGiantModel与世界一流算法对比



华为云脑成功训练千亿级开源中文大模型

云脑II成功训练千亿级开源中文大模型——盘古α

模型创新 | 工程创新 | 开源应用

首创随机顺序自回归预训练语言模型ALM

MindSpore软硬件协同优化、大规模分布式自动并行

训练了全球首个2千亿参数中文全开源预训练语言模型

模型	数据开源	算法开源	模型开放	服务开放
英文GPT-3 175B	X	X	X	√
开源清漆CPM 2.0B	X	X	√	X
盘古α 200B+13B	√	√	√	√

首次实现代码全开源

HDC.Cloud



03 投资建议：梳理AIGC相关受益厂商

3.1 投资建议: 梳理AIGC的受益厂商

- ◆ 我们认为AIGC的出世会产生革命性的影响，同时有望赋能千行百业。我们梳理了三条路径图，积极的推荐以下三条投资主线：
 - ✓ 1) 具备算力基础的厂商，重点推荐**中科曙光**，其他受益标的为**寒武纪、商汤、海光信息、浪潮信息、景嘉微、拓维信息、神州数码、龙芯中科**；
 - ✓ 2) 具备AI算法商业落地的厂商，重点推荐**科大讯飞、拓尔思**，其他受益标的为**海天瑞声**；
 - ✓ 3) AIGC相关技术储备的应用厂商，受益标的为：**百度、同花顺、三六零、金山办公**。

AIGC的A股受益标的

公司名称	股票代码	收盘价	市值(亿元)	EPS(元)			PE(倍)		
		2023/3/1	2023/3/1	2021	2022E	2023E	2021	2022E	2023E
寒武纪*	688256.SH	90.60	363.14	-2.06	-2.79	-1.78	-	-	-
拓维信息*	002261.SZ	10.96	137.64	0.07	-0.04	0.15	156.6	-	72.3
神州数码*	000034.SZ	27.74	185.54	0.37	1.56	1.90	75.2	17.8	14.6
龙芯中科*	688047.SH	119.65	479.80	0.66	0.43	0.72	181.3	278.0	166.3
浪潮信息*	000977.SZ	42.88	627.64	1.38	1.67	2.01	31.1	25.7	21.3
景嘉微*	300474.SZ	75.02	341.39	0.97	0.64	0.93	77.3	118.0	80.6
中科曙光	603019.SH	34.33	502.60	0.80	1.03	1.47	42.9	33.3	23.4
海光信息	688041.SH	53.50	1243.52	0.16	0.35	0.64	331.1	152.9	83.6
科大讯飞	002230.SZ	49.88	1158.75	0.70	0.25	0.94	71.3	199.5	53.1
拓尔思	300229.SZ	23.72	183.03	0.34	0.42	0.60	69.1	56.5	39.5
海天瑞声*	688787.SH	230.30	98.57	0.89	0.48	1.23	258.8	476.9	188.0
同花顺	300033.SZ	125.11	672.59	3.56	3.15	4.40	35.1	39.7	28.4
三六零*	601360.SH	11.31	808.14	0.13	-0.35	0.07	87.0	-	163.7
金山办公	688111.SH	300.00	1383.79	2.26	2.42	3.88	132.8	124.0	77.3

3.2.1 浪潮信息：中国服务器/AI服务器市占率稳居榜首

- ◆ **浪潮信息是全球领先的新型IT基础架构产品、方案及服务提供商：**公司是全球领先的 AI 基础设施供应商，拥有业内最全的人工智能计算全堆栈解决方案，涉及训练、推理、边缘等全栈 AI 场景，构建起领先的 AI 算法模型、AI 框架优化、AI 开发管理和应用优化等全栈 AI 能力，为智慧时代提供坚实的基础设施支撑。
- ◆ **公司算力技术壁垒浓厚：**生产算力方面，公司拥有业内最强最全的 AI 计算产品阵列，业界性能最好的Transformer 训练服务器 NF5488、全球首个 AI 开放加速计算系统 MX1、自研 AI 大模型计算框架 LMS。聚合算力层面，公司针对高并发训练推理集群进行架构优化，构建了高性能的NVMe 存储池，深度优化了软件栈，性能提升 3.5 倍以上。调度算力层面，浪潮信息 AIstation 计算资源平台可支持 AI 训练和推理，是业界功能最全的 AI 管理平台；同时，浪潮信息还有自动机器学习平台 AutoML Suite，可实现自动建模，加速产业化应用。

浪潮信息智算中心



浪潮信息智算中心



3.2.2 中科曙光：我国高性能计算、智能计算领军企业

- ◆ **中科曙光作我国核心信息基础设施领军企业：**在高端计算、存储、安全、数据中心等领域拥有深厚的技术积淀和领先的市场份额，并充分发挥高端计算优势，布局智能计算、云计算、大数据等领域的技术研发，打造计算产业生态，为科研探索创新、行业信息化建设、产业转型升级、数字经济发展提供了坚实可信的支撑。
- ◆ **依托先进计算领域的先发优势和技术细节，中科曙光全面布局智能计算：**完成了包括AI核心组件、人工智能服务器、人工智能管理平台、软件等多项创新，构建了完整的AI计算服务体系。并积极响应时代需求，在智能计算中心建设浪潮下，形成了5A级智能计算中心整体方案。目前，曙光5A智能计算中心已在广东、安徽、浙江等地建成，江苏、湖北、湖南等地已进入建设阶段，其他地区也在紧张筹备和规划中。

中科曙光主要产品

<p>通用服务器</p> <ul style="list-style-type: none"> 机架式服务器 高密度服务器 刀片服务器 核心应用服务器 	<p>智能计算服务器</p> <ul style="list-style-type: none"> 深度学习训练 智能应用推理 	<p>终端&工作站</p> <ul style="list-style-type: none"> 微型计算机 工作站 	<p>高性能计算机</p> <ul style="list-style-type: none"> 通用高性能计算机 高性能计算机系统组件 高性能计算机的服务支撑
<p>机房冷却设施</p> <ul style="list-style-type: none"> 微模块产品 液冷基础设施产品 	<p>存储产品</p> <ul style="list-style-type: none"> 分布式统一存储 多控统一存储 高密度存储服务器 备份一体机 	<p>网络安全产品</p> <ul style="list-style-type: none"> 数据中心安全产品 汇聚分流设备 智能加速卡 网络内容识别分析系统 网络态势感知系统 	<p>大数据平台软件</p> <ul style="list-style-type: none"> 大数据智能引擎系列 数据工程服务系列 视频智能分析系列 大数据与人工智能实训平台
<p>云计算平台软件</p> <ul style="list-style-type: none"> 云计算操作系统 超融合一体机 云桌面 云容灾 	<p>计算服务</p> <ul style="list-style-type: none"> 弹性计算服务 混合计算服务 专有计算服务 API 托管、运营 	<p>云计算服务</p> <ul style="list-style-type: none"> 云服务器 ECS 裸金属 BMS 对象存储 OSS 云容器实例 CCI 人工智能服务 数据开发 DDS 数据治理中心 DGS 数据服务 DSS 数据可视化 DAV 数据集成 Data Integration 	<p>城市云</p> <ul style="list-style-type: none"> 智慧城市 国资云 交通云 医疗云
<p>5A级智算中心</p>			

中科曙光硅立方液体相变冷却计算机



3.2.3 海光信息：支持全精度，GPU实现规模量产

- ◆ **海光信息主要从事高端处理器、加速器等计算芯片产品和系统的研究、开发，主要产品包括海光CPU和海光DCU:**2018年10月，公司启动深算一号DCU产品设计，海光8100采用先进的FinFET工艺，典型应用场景下性能指标可以达到国际同类型高端产品的同期水平。2020年1月，公司启动DCU深算二号的产品研发。
- ◆ **海光DCU性能强大:**海光DCU基于大规模并行计算微结构进行设计，不但具备强大的双精度浮点计算能力，同时在单精度、半精度、整型计算方面表现同样优异，是一款计算性能强大、能效比较高的通用协处理器。海光DCU集成片上高带宽内存芯片，可以在大规模数据计算过程中提供优异的数据处理能力。

海光信息主要产品



系列	7000系列CPU	5000系列CPU	3000系列CPU	系列	8000系列DCU
核心规格	最大32个物理核心	最大16个物理核心	最大8个物理核心	核心规格	60-64个深度计算单元
应用领域	高端通用服务器、先进计算系统	通用服务器	个人工作站、工控设备等终端产品	应用领域	先进计算系统、人工智能

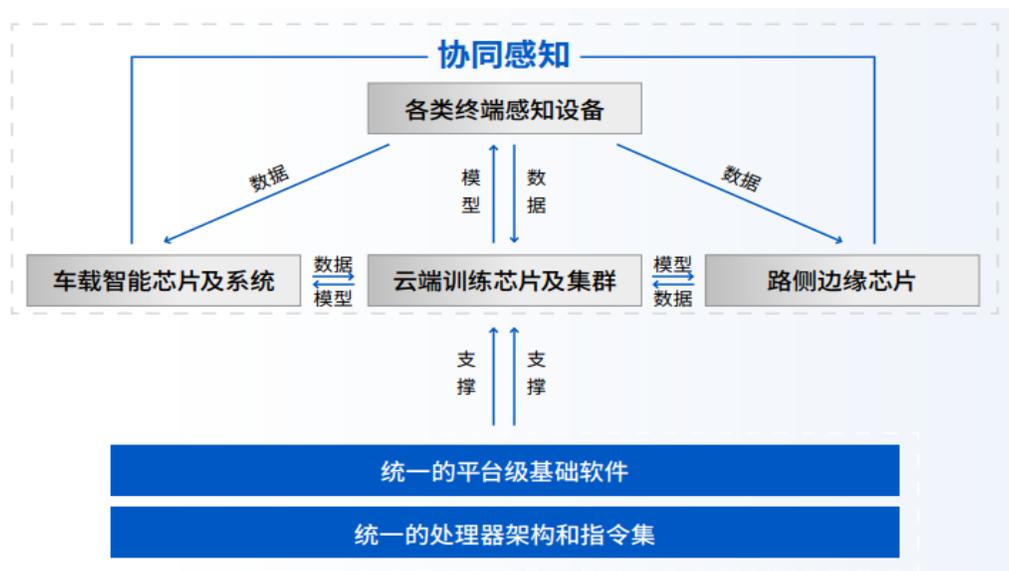
海光深算一号性能达到国际同类产品水平

项目	海光	NVIDIA	AMD
品牌	深算一号	Ampere 100	M1100
生产工艺	7nm FinFET	7nm FinFET	7nm FinFET
核心数量	4096 (64CUs)	2560 CUDA processors 640 Tensor processors	120CUs
内核频率	Up to 1.5GHz (FP64) Up to 1.7Ghz (FP32)	Up to 1.53Ghz	Up to 1.5GHz (FP64) Up to 1.7Ghz (FP32)
显存容量	32GB HBM2	80GB HBM2e	32GB HBM2
显存位宽	4096 bit	5120 bit	4096bit
显存频率	2.0 GHz	3.2 GHz	2.4 GHz
显存带宽	1024 GB/s	2039 GB/s	1228 GB/s
TDP	350 W	400 W	300 W
CPU to GPU 互联	PCIe Gen4 x 16	PCIe Gen4 x 16	PCIe GEN4 x 16
GPU to GPU 互联	xGMI x 2, Up to 184 GB/s	NVLink up to 600 GB/s	Infinity Fabric x 3, up to 276 GB/s

3.2.4 寒武纪：少数全面掌握AI芯片技术的企业之一

- ◆ **寒武纪是目前国际上少数几家全面系统掌握了通用型智能芯片及其基础系统软件研发和产品化核心技术的企业之一**：寒武纪主营业务是应用于各类云服务器、边缘计算设备、终端设备中人工智能核心芯片的研发和销售。公司的主要产品包括终端智能处理器IP、云端智能芯片及加速卡、边缘智能芯片及加速卡以及与上述产品配套的基础系统软件平台。
- ◆ **公司AI技术积累浓厚**：能提供云边端一体、软硬件协同、训练推理融合、具备统一生态的系列化智能芯片产品和平台化基础系统软件。2022年3月，寒武纪正式发布了新款训练加速卡“MLU370-X8”，搭载双芯片四芯粒封装的思元370，集成寒武纪MLU-Link多芯互联技术，主要面向AI训练任务。

寒武纪“云边端车”协同



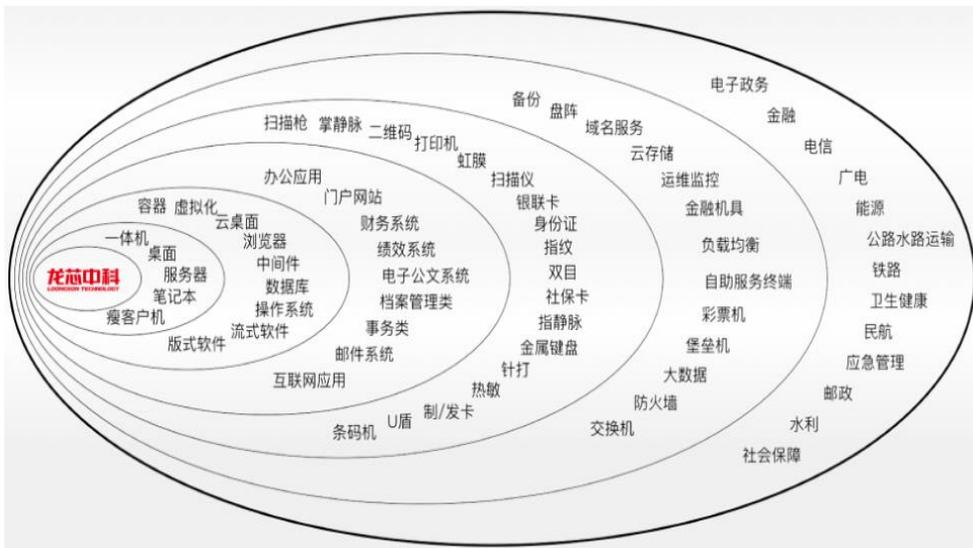
寒武纪产品技术图谱



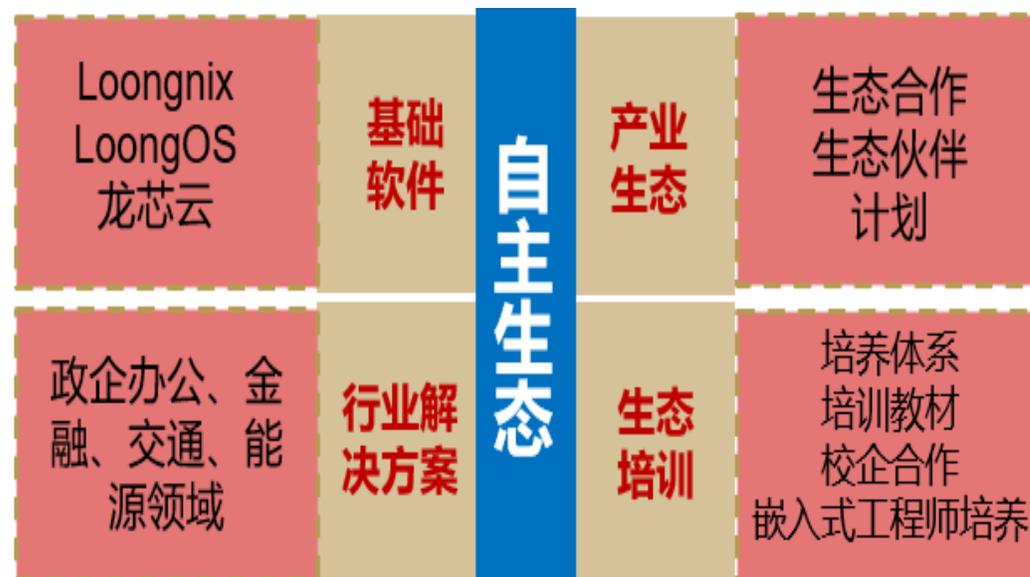
3.2.5 龙芯中科：2K2000系列集成自主GPU

- ◆ **龙芯中科主营业务为处理器及配套芯片的研制、销售及服务：**主要产品与服务包括处理器及配套芯片产品与基础软硬件解决方案业务。公司基于信息系统和工控系统两条主线开展产业生态建设，面向网络安全、办公与业务信息化、工控及物联网等领域与合作伙伴保持全面合作，产品在电子政务、能源、交通、金融、电信、教育等行业领域已获得广泛应用。
- ◆ **公司自主研发2K200系列GPU：**2022年12月，龙芯2K2000完成了初步功能调试及性能测试，达到其设计目标，2023年将推出试用。龙芯2K2000集成了两个LA364处理器核，典型工作频率为1.5GHz，共享2MB的L2缓存，SPEC2006INT (base) 单核定/浮点分值达到13.5/14.9分。龙芯2K2000芯片集成了龙芯自主研发的GPU，并优化了图形算法和性能。

龙芯中科生态合作示意图



龙芯中科自主生态



3.2.6 景嘉微：新一代JM9系列有望打开商用市场

- ◆ **国产GPU龙头企业:** 公司成立于2006年，主要从事军用电子产品的研发、生产、销售，目前形成了三大业务板块分别是图形线控模块、小型专用雷达和芯片业务。GPU方面，2014年首推JM5400实现了军用GPU的国产替代；第二款芯片JM7200于2018年研发成功，具备了PC端的功能；日前，公司9系列芯片研发成功，具备高性能计算能力。
- ◆ **新一代JM9系列有望打开商用市场:** 日前，公司JM9系列图形处理芯片已顺利发布，应用领域涵盖地理信息系统、媒体处理、CAD辅助设计、游戏、虚拟化等高性能显示和人工智能计算领域。目前，信创市场为公司提供了新的业务增长点，JM9系列图形处理芯片的成功发布将为公司未来进一步拓展通用市场提供强有力的产品支撑。

景嘉微GPU系列产品



景嘉微7系列GPU示意图



3.3.1 科大讯飞: 自然语言处理的全球龙头厂商

- ◆ **科大讯飞于2022年初正式发布了“讯飞超脑 2030 计划”，其目的是向“全球人工智能产业领导者”的长期愿景迈进。**该计划是公司的核心战略，目的是构建基于认知的人机协作、自我进化的复杂系统，即让机器人感官超越人类，具备自主进化的能力，**打造可持续自主进化的复杂智能系统，助力机器人走进千家万户。**其计划分为三个阶段性里程碑：
 - ✓ **第一阶段** (2022-2023) : 推出可养成的宠物玩具、仿生动物等软硬件一体机器人，同期推出专业数字虚拟人家族，担当老师、医生等角色；
 - ✓ **第二阶段** (2023-2025) : 推出自适应行走的外骨骼机器人和陪伴数字虚拟人家族，老人通过外骨骼机器人能够实现正常行走和运动，同期推出面向青少年的抑郁症筛查平台；
 - ✓ **第三阶段** (2025-2030) : 最终推出懂知识、会学习的陪伴机器人和自主学习虚拟人家族，全面进入家庭。

“讯飞超脑”计划和里程碑



3.3.2 拓尔思: 全文检索技术的始创者

- ◆ **拓尔思是中文全文检索技术的始创者**：领先的大数据、人工智能和数据安全产品及服务提供商，成立于1993年。其主要核心业务包括软件平台产品研发，行业应用系统解决方案和大数据SaaS/DaaS云服务三大板块，涵盖大数据、人工智能、互联网内容管理、网络信息安全和互联网营销等领域方向。
- ◆ **公司技术实力雄厚**：人工智能领域：公司是国内最早从事自然语言处理（NLP）研发的企业之一，在 NLP、知识图谱、OCR、图像视频结构化领域都具备自主可控的底层技术；在大数据方面，数据采集、挖掘、分析等底层技术均实现技术可控。
- ◆ **公司行业领域覆盖广**：由于公司人工智能和大数据核心技术具有跨行业/领域的通用普适性，因此公司技术产品和服务所支持的业务应用落地场景众多，涵盖多个行业/领域，大型或头部企业级客户超过 8000 家。广泛覆盖于金融、制造、能源、互联网、传媒、政府、公共安全等领域。

拓尔思数据应用相关产品

数据应用 (SaaS 服务)

- 网察 (舆情监测)
- 数家 (资讯数据云服务)
- 海策 (政策大脑)
- 数星产业大脑
- 数星风控
- 数星电报
- 冒烟指数 (金融风险监测)
- 网脉
- 网巡
- 网搜
- 天目 (开源情报)
- 全球基因专利数据云服务

数据采集

- 海蜘蛛 (分布式采集系统)
- 数据治理
- 海聚 (数据整合系统)
- TRS海贝搜索数据库
- 数据中台
- 数据分析
- 水晶球分析师

数据可视化

- 数据可视化大屏云服务平台

拓尔思网络安全相关产品

边界安全

- 天行安全单向导入系统UIS
- 天行安全隔离与信息交换系统GAP
- 天行安全请求服务系统RSS
- 天行安全视频交换系统MTP
- 天行安全数据交换系统DTP
- 天行部门间信息共享与服务平台ISSP
- 天行单向隔离光闸FGAP
- 天行防火墙FW
- 天行集中监管控制系统3AS
- 天行数据交换与接口服务系统DTS

大数据安全

- 天行边界安全运维管理平台BOP
- 天行数据库安全审计及性能监测审计系统DAS
- 天行应用日志安全审计系统ALSA

物联网安全

- 天行应用系统安全审计系统VAS

3.3.3 海天瑞声: AI训练数据服务领军者

- ◆ **海天瑞声是我国领先的训练数据专业提供商，发行人始终致力于为AI产业链上的各类机构提供算法模型开发训练所需的专业数据集。**发行人所提供的训练数据覆盖智能语音、计算机视觉、自然语言等多个核心领域，全面服务于人机交互、智能驾驶、智慧城市等多种创新应用场景。
- ◆ **公司产品倍受认可：**公司的产品和服务已获得阿里巴巴、腾讯、百度、科大讯飞、海康威视、微软、亚马逊、三星等国内外客户的认可，应用于其研发的个人助手、智能音箱、语音导航、搜索服务、智能驾驶、机器翻译等多种人工智能产品相关算法模型的训练过程中。
- ◆ **公司具备多重技术领先性：**公司核心技术先进性包括算法与数据处理技术并用、工具和平台共建、在语音语言学基础研究方面有深厚积累等，公司可提供数据的一站式解决方案，广泛支持复杂的算法数据标注需求。

云从科技产品服务体系



3.4.1 百度: 国内搜索引擎龙头公司，深度AI技术储备厂商

- ◆ **百度是国内信息和知识为核心的互联网综合服务龙头公司:** 百度于 2000 年 1 月 1 日在中关村创立，2019 年百度用户规模突破10 亿，百度 App 日活跃用户规模达 2 亿，信息流位居中国第一。作为 全球领先的人工智能平台型公司，百度大脑 集成了百度超过240项的 AI 能力，是 中国唯一的“软硬一体AI大生产平台”。
- ◆ **百度着力AI核心技术引擎发展，全面布局人工智能:** 基于百度内部开发的 AI 百度大脑平台，百度人工智能新业务主要包括 DuerOS (语音助手和智能设备)和 Apollo自动驾驶和智能交通)。百度大脑是百度 AI 核心技术引擎，包括视觉、语音、自然语言处理、知识图谱、深度学习等 AI 核心技术和 AI 开放平台。
- ◆ **百度为中国第一大市场搜索引擎:** 百度提供形式丰富的搜索服务，包括文本、语音、视觉、图片、视频、文献等，百度可提供多种搜索模式，包括文本输入、基于 AI 的语音搜索、视觉搜索和垂直搜索服务。

百度部分产品



百度DuerOS内容生态



3.4.2 同花顺: 金融IT龙头公司，领先金融AI服务商

- ◆ **公司为国内领先的互联网金融信息服务提供商**：公司成立于2001年,公司在业内拥有完整的产业链，产品及服务覆盖产业链上下游。主要客户涵盖金融市场的各层次参与主体，包括券商、基金、私募、银行、保险、政府、科研院所、上市公司等机构客户和广大的个人投资者用户。公司主要业务是为各类机构客户提供软件产品和系统维护服务、金融数据服务、智能推广服务，为个人投资者提供金融资讯、投资理财分析工具、理财产品投资交易服务等。
- ◆ **公司继续加大对人工智能技术的应用研究**：公司已构建同花顺 AI开放平台，可面向客户提供智能语音、智能客服、智能金融问答、智能投顾、智能质检机、会议转写系统、智能医疗辅助系统等多项 AI产品及服务，可为银行、证券、保险、基金、私募、高校、政府等行业提供智能化解决方案。目前，公司AI产品正在积极拓展至生活、医疗、教育等领域。

同花顺部分人工智能产品



文字识别



智能语音



自然语言基础服务



知识图谱



智能金融问答



数据智能



智能投顾

智能投顾

投顾AI辅助系统

人机结合的服务模式，针对服务客户的真实场景打造的投顾助手，帮助投顾提高服务效率和服务质量

智能机器人

运用 AI 技术打造更懂财经的个性化智能助理，开启股市人工智能新纪元

资产配置

基于用户画像、产品问卷，构建用户风险测评系统，提供个性化资产配置方案，持续跟踪监控服务用户

同花顺智能金融问答产品

问句输入操作简便，解析识别度高，支持各类形式问法

1. 自然语言解析，识别度高

- 连续三年净利润增长的股票
- 净利润和营收下跌大于10%
- 台州制药净利润（台州制药是002728.SZ特一药业的原文中文简称）

2. 支持四则运算

- $(每股红股+每股转增股本) \geq 0.5$

3. 多chunk、多条件组合选股

- $(每股红股+每股转增股本) \geq 0.5$
- 增发解禁股票
- 成交量放大，5日内涨幅小于10%

4. 支持时间推理逻辑

- 首发上市日期大于5年
- 1月20日创今年新高

来源多样、种类齐全、结构化非结构化数据应有尽有

1. 结构化数据：数据源于同花顺FinD数据，数据及时准确齐全

- 概念指数
- 昨日小单总额，近三天小单总额占比大于10小于20
- 总额占比排名

2. 技术指标数据，包含多种均线形态、指标形态、术语

- 非停牌，非Est，macd低位金叉，5月10日均线粘合，中小盘
- 高位二次死叉

3. 非结构化数据，包含公告、研报、新闻、微博、股吧等

- 午间利好公告组合（公告）
- 或将戴帽（新闻）
- 中报点评（研报和新闻）

4. 数据精确度高，一般保留原始数据精确度，便于精确计算

- 2015/12/4人民币汇率升值中间价



04 风险提示

风险提示

- ◆ **核心技术水平升级不及预期的风险:** AIGC相关产业技术壁垒较高，公司核心技术难以突破，进程低于预期，影响整体进度。
- ◆ **AI伦理风险:** AI可能会生产违反道德、常规、法律等内容。
- ◆ **政策推进不及预期的风险:** 受到宏观经济、财政、疫情影响，政策推进节奏不及预期。
- ◆ **中美贸易摩擦升级的风险:** 供应链存在部分海外提供商，容易受到美国“卡脖子”制裁，导致产品研发不及预期。

分析师与研究助理简介

刘泽晶（首席分析师）2014-2015年新财富计算机行业团队第三、第五名，水晶球第三名，10年证券从业经验。

分析师承诺

作者具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，保证报告所采用的数据均来自合规渠道，分析逻辑基于作者的职业理解，通过合理判断并得出结论，力求客观、公正，结论不受任何第三方的授意、影响，特此声明。

评级说明

公司评级标准	投资评级	说明
以报告发布日后的6个月内公司股价相对上证指数的涨跌幅为基准。	买入	分析师预测在此期间股价相对强于上证指数达到或超过15%
	增持	分析师预测在此期间股价相对强于上证指数在5%—15%之间
	中性	分析师预测在此期间股价相对上证指数在-5%—5%之间
	减持	分析师预测在此期间股价相对弱于上证指数5%—15%之间
	卖出	分析师预测在此期间股价相对弱于上证指数达到或超过15%
行业评级标准		
以报告发布日后的6个月内行业指数的涨跌幅为基准。	推荐	分析师预测在此期间行业指数相对强于上证指数达到或超过10%
	中性	分析师预测在此期间行业指数相对上证指数在-10%—10%之间
	回避	分析师预测在此期间行业指数相对弱于上证指数达到或超过10%

华西证券研究所：

地址：北京市西城区太平桥大街丰汇园11号丰汇时代大厦南座5层

网址：<http://www.hx168.com.cn/hxqz/hxindex.html>

华西证券股份有限公司（以下简称“本公司”）具备证券投资咨询业务资格。本报告仅供本公司签约客户使用。本公司不会因接收人收到或者经由其他渠道转发收到本报告而直接视其为本公司客户。

本报告基于本公司研究所及其研究人员认为的已经公开的资料或者研究人员的实地调研资料，但本公司对该等信息的准确性、完整性或可靠性不作任何保证。本报告所载资料、意见以及推测仅于本报告发布当日的判断，且这种判断受到研究方法、研究依据等多方面的制约。在不同时期，本公司可发出与本报告所载资料、意见及预测不一致的报告。本公司不保证本报告所含信息始终保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者需自行关注相应更新或修改。

在任何情况下，本报告仅提供给签约客户参考使用，任何信息或所表述的意见绝不构成对任何人的投资建议。市场有风险，投资需谨慎。投资者不应将本报告视为做出投资决策的惟一参考因素，亦不应认为本报告可以取代自己的判断。在任何情况下，本报告均未考虑到个别客户的特殊投资目标、财务状况或需求，不能作为客户进行客户买卖、认购证券或者其他金融工具的保证或邀请。在任何情况下，本公司、本公司员工或者其他关联方均不承诺投资者一定获利，不与投资者分享投资收益，也不对任何人因使用本报告而导致的任何可能损失负有任何责任。投资者因使用本公司研究报告做出的任何投资决策均是独立行为，与本公司、本公司员工及其他关联方无关。

本公司建立起信息隔离墙制度、跨墙制度来规范管理跨部门、跨关联机构之间的信息流动。务请投资者注意，在法律许可的前提下，本公司及其所属关联机构可能会持有报告中提到的公司所发行的证券或期权并进行证券或期权交易，也可能为这些公司提供或者争取提供投资银行、财务顾问或者金融产品等相关服务。在法律许可的前提下，本公司的董事、高级职员或员工可能担任本报告所提到的公司的董事。

所有报告版权均归本公司所有。未经本公司事先书面授权，任何机构或个人不得以任何形式复制、转发或公开传播本报告的全部或部分内容，如需引用、刊发或转载本报告，需注明出处为华西证券研究所，且不得对本报告进行任何有悖原意的引用、删节和修改。

THANKS

