计算机

证券研究报告
2023 年 04 月 11 日

以昇思为基，盘古生态引领中国 AI 未来

华为在 AI 的软硬件布局全面，全体系自主可控且水平国内领先

华为具备全栈全场景 AI 解决方案，为大模型提供坚实的软硬件平台，整体结构可以分为硬件层（昇腾芯片+服务器）、AI 赋能层与模型层。在硬件层华为拥有已经经过验证的训练端、推理端芯片和成体系的服务体系，达到了较好的自主可控；软件层面，公司在 AI 框架和 AI 一站式开发平台上都处于国内领先地位；在最上层的大模型，华为在 NLP、CV、多模态和科学计算上都有领先的布局。

MindSpore 昇思与 ModelArts 为华为 AI 软件生态打造了良好的基础

昇思自 2020 年至今已经成为国内领先的开源 AI 框架社区，在使用率上与百度飞桨并列国内双雄，针对大模型，昇思也在并行计算、内存复用等领域有针对性的改进；此外昇思拥有完善的生态伙伴体系与强大的“朋友圈”，众多科研机构与上市公司都基于 Mindspore 开发 AI 算法，也有包括盘古系列在内的多款大模型是基于 Mindspore 完成开发。ModelArts 作为基于华为云的一站式 AI 开发平台，也拥有较高的市占率和较完善的产品体系。由此可知华为在 AI 赋能层处于国内领先的水平，我们认为这会为未来 2B 大模型应用打下基础。

华为在大模型层布局完善，在 NLP 大语言模型积累已久，总体实力位列国内第一梯队

华为早在 2021 年就开展了大规模自回归中文预训练大语言模型的研究，最初与鹏程实验室、北京大学等联合推出了盘古 α，此模型与 GPT 系列结构相似，最高有 2000 亿参数，可见华为技术积累已久。随后在 2023 年 3 月份，公司发表的新论文描述了由华为完全自主开发的盘古 Σ，此模型是在盘古 α 基础上拓展而来的万亿参数稀疏模型，由 3290 亿 token 训练而成，从模型结果看，盘古 Σ 是国内第一梯队的中文大语言模型，无论在中文下游任务、中文对话生成、机器翻译上都处于领先地位，在英语自然语言理解上甚至不输 GPT3 13B 版本，我们认为长久的技术积淀造就了华为在大语言模型上的较强能力，未来持续看好基于盘古大语言模型的各类 2C 和 2B 应用。

风险提示：AI 应用落地不及预期、硬件领域发展不及预期、大模型进展不及预期

投资评级

行业评级	强于大市(维持评级)
上次评级	强于大市

作者

缪欣君 分析师
SAC 执业证书编号：S1110517080003
miaoxinjun@tfzq.com

行业走势图

资料来源：聚源数据

相关报告

1 《计算机-行业点评:从源头开始，AI 监管的落地在国家云》 2023-04-06
2 《计算机-行业深度研究:数据要素：数字经济发展核心引擎》 2023-04-05
3 《计算机-行业点评:当大模型遇见金融：海内外金融领域大模型对比》 2023-04-02

内容目录

图表目录

1. 华为在 AI 与大模型全栈布局完善，硬/软件皆国内领先

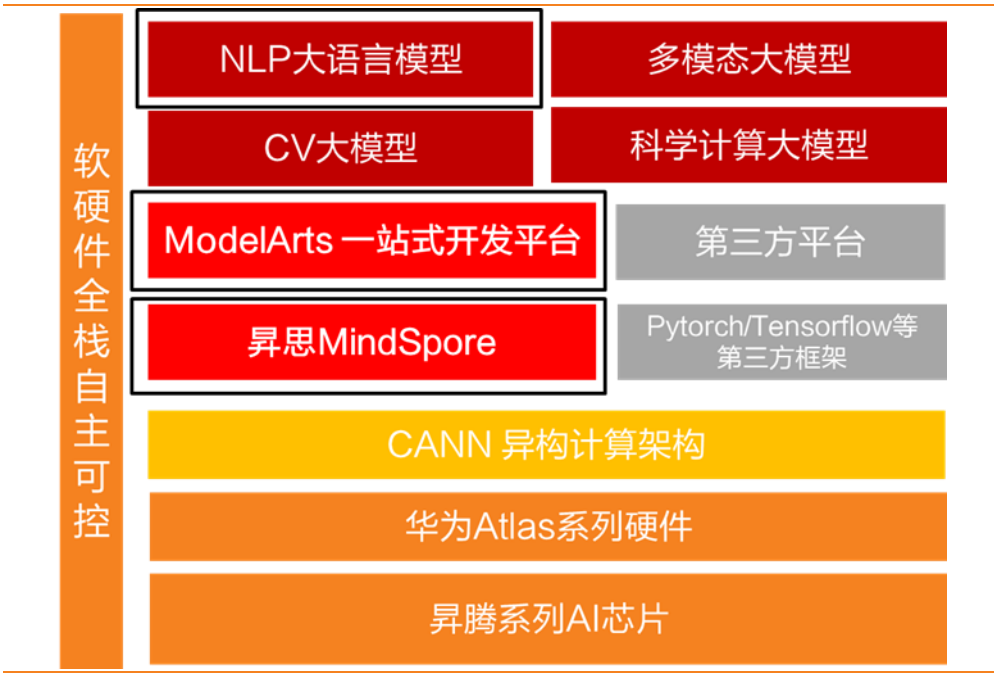
华为具备全栈全场景 AI 解决方案，为大模型提供坚实的软硬件平台。整体结构可以分为硬件层（昇腾芯片+服务器）、AI 赋能层与模型层。

1）在硬件层：华为拥有包括昇腾 910 与昇腾 310 在内的训练端与推理端芯片，在此之上有完善的 Atlas 系列硬件，随后有 CANN 异构计算架构，此架构对上支持多种 AI 框架，对下服务 AI 处理器与编程，发挥承上启下的关键作用，是提升昇腾 AI 处理器计算效率的关键平台；

2）AI 赋能层：华为拥有自主的 AI 框架昇思 MindSpore 和一站式开发平台 ModelArts。Mindspore 是国内第一梯队的 AI 框架，生态体系成熟；ModelArts 也是基于华为云的头部 AI 开发平台，助力企业提高开发效率降低开发成本；

3）盘古系列大模型层：华为开发的模型分别为 NLP 大模型、CV 大模型、科学计算大模和多模态大模型。

图 1：华为在 AI 与大模型的全栈布局完善



资料来源：华为昇腾官网，IT 之家公众号、天风证券研究所

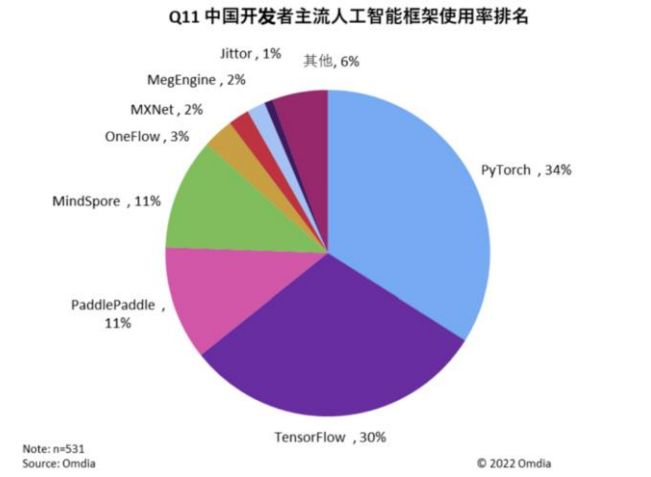
2. 昇思 MindSpore：国内领先的 AI 框架

AI 框架是智能经济时代的操作系统。根据中国信息通信研究院编写的《AI 框架发展白皮书》，AI 框架是 AI 算法模型设计、训练和验证的一套标准接口、特性库和工具包，集成了算法的封装、数据的调用以及计算资源的使用，同时面向开发者提供了开发界面和高效的执行平台，是现阶段 AI 算法开发的必备工具。

2.1. 国产 AI 框架呈现 MindSpore 与 PaddlePaddle 双寡头态势

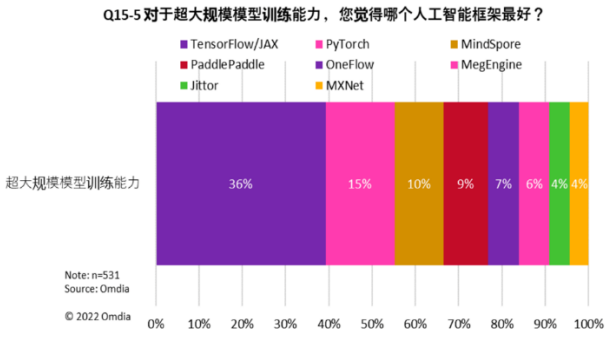
Mindspore 是使用率最高的国产 AI 框架之一。在 2023 年 2 月 Omdia 发布的《中国人工智能框架市场调研报告》中，分别针对中国开发者人工智能使用率、大模型使用意愿等进行调研，结果显示昇思 MindSpore 与 PyTorch、TensorFlow、PaddlePaddle 等人工智能框架在知名度与使用率市场份额上处于第一梯队，其中国产 AI 框架中昇思与百度飞浆并驾齐驱；在大模型开发上， Omidia 的调查显示，10%的调研用户认为昇思在超大规模模型训练能力上更强。

图 2：中国开发者主流人工智能框架使用率 11%，仅次于 Tensorflow



资料来源：昇思官网、天风证券研究所，注：样本=531

图 3：超大规模模型训练能力位列国产 AI 框架第一



资料来源：昇思官网、天风证券研究所，注：样本=531

2.2. Mindspore 昇思功能全面，针对大模型特殊优化

昇思 MindSpore 是一个全场景深度学习框架，旨在实现易开发、高效执行、全场景覆盖三大目标。其易开发表现为 API 友好、调试难度低；高效执行包括计算效率、数据预处理效率和分布式训练效率；全场景则指框架同时支持云、边缘以及端侧场景。昇思 MindSpore 包含的主要功能如下：

（1）ModelZoo（模型库）：ModelZoo 提供可用的深度学习算法网络，也欢迎更多开发者贡献新的网络(ModelZoo 地址)。

（2）Expression（全场景统一 API）：基于 Python 的前端表达与编程接口，支持两个融合（函数/OOP 编程范式融合、AI+数值计算表达融合）以及两个统一（动静表达统一、单机分布式表达统一）。

（3）第三方前端：支持第三方多语言前端表达，未来计划陆续提供 C/C++、华为自研编程语言前端-仓颉（目前还处于预研阶段）等第三方前端的对接工作，引入更多的第三方生态。

（4）Data（数据处理层）：提供高效的数据处理、常用数据集加载等功能和编程接口，支持用户灵活地定义处理注册和 pipeline 并行优化。

（5）Compiler（AI 编译器）：图层的核心编译器，主要基于端云统一的 MindIR 实现三大功能，包括硬件无关的优化（类型推导、自动微分、表达式化简等）、硬件相关优化（自动并行、内存优化、图算融合、流水线执行等）、部署推理相关的优化（量化、剪枝等）。

（6）Runtime（全场景运行时）：昇思 MindSpore 的运行时系统，包含云侧主机侧运行时系统、端侧以及更小 IoT 的轻量化运行时系统。

（7）Insight（可视化调试调优工具）：昇思 MindSpore 的可视化调试调优工具，能够可视化地查看训练过程、优化模型性能、调试精度问题、解释推理结果。

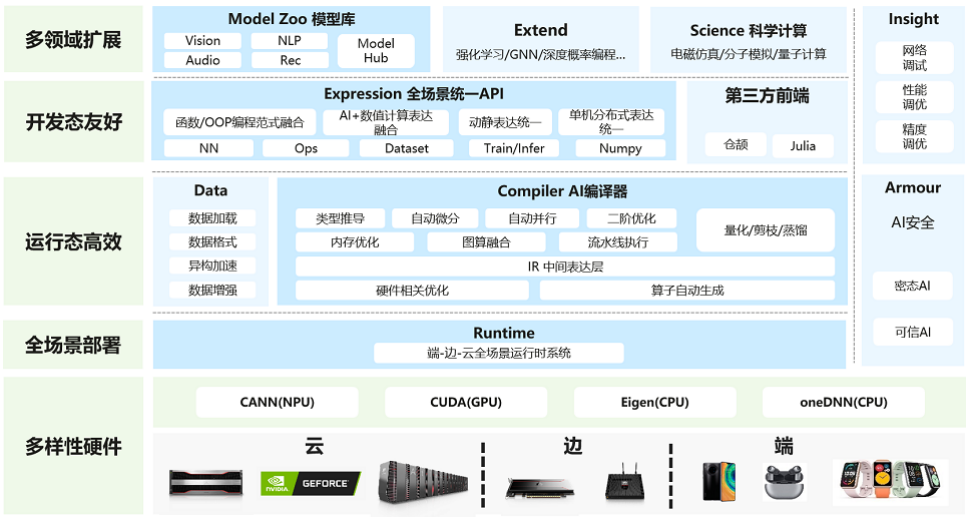
在支持超大规模模型训练开发方面，昇思 MindSpore 在进行架构设计时就考虑了大模型开发时遇到的内存占用、通信瓶颈、调试复杂、部署难等问题，针对性的技术研究与创新包括以下几点。

业界领先的全自动并行能力，提供 6 维混合并行算法，即数据并行、模型并行、流水并行、优化器并行等，一行代码实现模型自动切分、分布式并行计算，开发并行代码量降低 80%、系统调优时间下降 60%；

极致的全局内存复用能力，在开发者无感知的情况下，自动实现 NPU 内存 / CPU 内存 / NVMe 硬盘存储的多级存储优化，512 卡就可训练 10 万亿规模的参数模型，极大降低大模型训练成本；

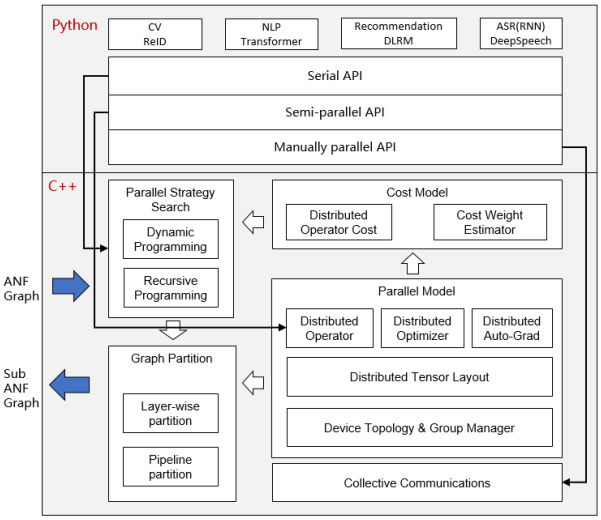
极简的断点续训能力，可解决大集群训练故障导致的任务中断问题，实现自动恢复、继承性

图 4：MindSpore 总体架构



资料来源：昇思官网、天风证券研究所

图 5：MindSpore 自动并行原理



资料来源：昇思官网，天风证券研究所

2.3. Mindspore 生态体系与应用场景广泛

截止 2023 年 4 月 8 日，昇思的社区用户超过 430 万，总 Star 数超过 22.1K。昇思自 2020 年开发至今走过了 3 个年头，到 2023 年 3 月底，昇思已经拥有了核心贡献者 9700+，下载量超过 386 万，服务企业 5500+，MSG 足迹遍布国内外 30 个城市，在码云（Gitee）千万级开源项目中活跃度排名第一。社区伙伴企业遍布各大高校和各类企业，长亮科技、华宇信息、润和软件、多伦科技、新开普等上市公司亦在其中。

图 6：昇思的部分生态伙伴企业一览

资料来源：昇思官网、天风证券研究所

此外，基于 MindSpore，华为和生态伙伴目前已开发出多款面向多模态、遥感、生物医药等领域的大模型：

紫东太初：业界首个三模态千亿参数大模型，支持文本、视觉、语音不同模态间的高效协同，可支撑影视创作、工业质检、智能驾驶等产业应用；

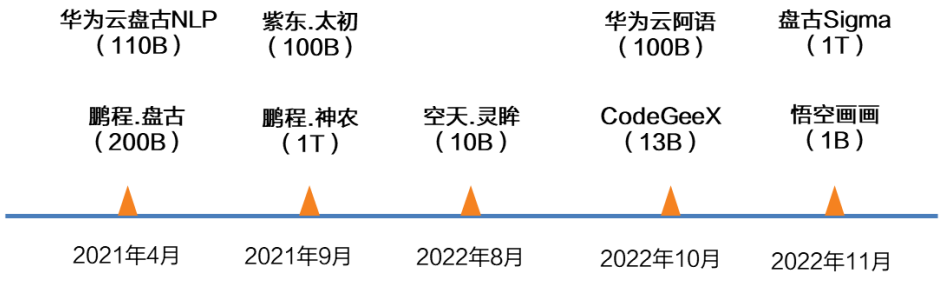
武汉 Luojia：遥感领域首个国产化自主可控的遥感专用机器学习框架；

鹏程盘古：业界首个千亿级参数中文自然语言处理大模型，可支持知识问答、知识检索、知识推理、阅读理解等丰富的下游应用；

鹏程神农：面向生物医学领域的人工智能平台，包含蛋白质结构预测等多个模块，为制药企业和医学研究机构提供平台能力，加速新型药物的筛选与创制；

空天灵眸：首个面向大规模跨模态数据的遥感智能解译生成式大模型，共享学习遥感多模态多任务的通用特征，加速 AI 应用于遥感领域。

图 7：基于昇思 MindSpore 开发的大模型一览



资料来源：量子位公众号、天风证券研究所

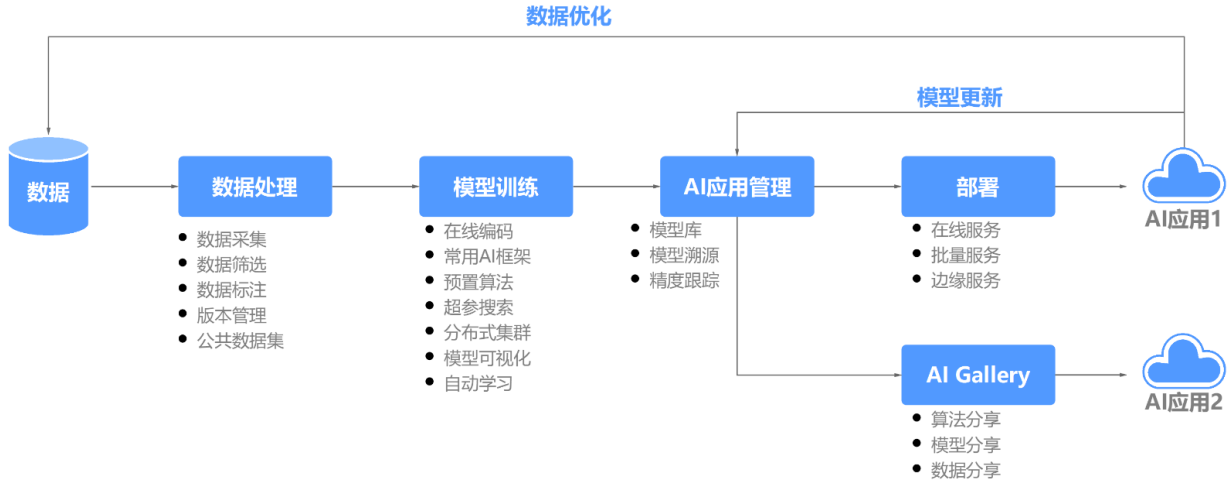
3. AI 框架之上，更有业内领先的一站式开发平台

ModelArts 是面向开发者的一站式 AI 开发平台，为机器学习与深度学习提供海量数据预处理及半自动化标注、大规模分布式 Training、自动化模型生成，及端-边-云模型按需部署能力，帮助用户快速创建和部署模型，管理全周期 AI 工作流。

ModelArts 的理念就是让 AI 开发变得更简单、更方便。从技术上看，ModelArts 底层支持各种异构计算资源，开发者可以根据需要灵活选择使用，而不需要关心底层的技术。同时，ModelArts 支持 Tensorflow、PyTorch、MindSpore 等主流开源的 AI 开发框架，也支持开发

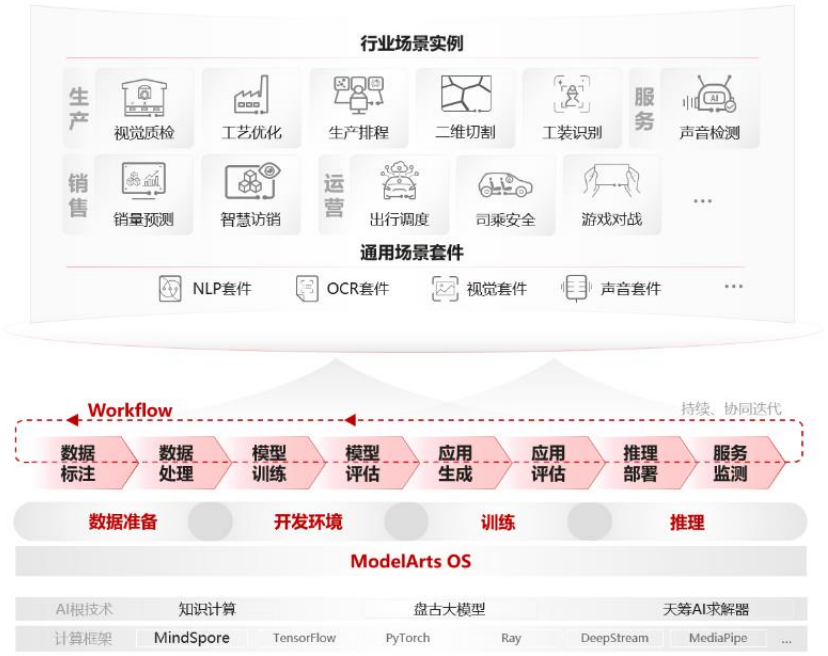
者使用自研的算法框架，匹配您的使用习惯。在产品架构上，能够支撑开发者从数据到 AI 应用的全流程开发过程。包含数据处理、模型训练、模型管理、模型部署等操作，并且提供 AI Gallery 功能，能够在市场内与其他开发者分享模型。

图 8：ModelArts 的产品架构



资料来源：华为云官网、天风证券研究所

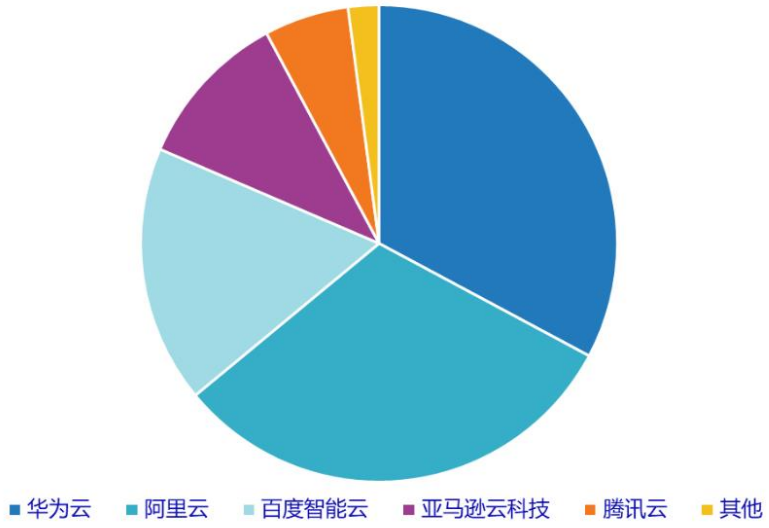
图 9： ModelArts 架构与 AI 开发的工作流相契合



资料来源：华为云公众号、天风证券研究所

华为云的 ModelArts 在中国机器学习公有云服务市场的份额位居前列。华为云 AI 开发生产线 ModelArts 在 AI 云服务方面的竞争优势明显。历经多年的技术创新，ModelArts 已成功在十多个领域进行商业化落地，持续领跑机器学习公有云市场，为 AI 开发带来变革。据 IDC 发布的《2022 H1 中国 AI 云服务市场研究报告》统计，华为云在中国机器学习公有云服务市场份额排名第一，迄今为止华为云 AI 开发生产线 ModelArts 已经连续五次登上榜首。

图 10：2022H1 华为云在中国机器学习公有云服务市场中位居前列



资料来源：IDC、天风证券研究所

在互联网领域，华为云 ModelArts 基于算法优化、语音质检等途径，有效提升了 T3 出行司乘安全检测模型的准确率和召回率，使危险驾驶事件率下降 **38.6%**，同时大幅降低模型开发和交付周期。

在自动驾驶领域，针对 AI 算法训练，华为云 ModelArts 支撑端到端训练效率提升；分布式多级缓存技术可以将训练时长缩短 **50%**；针对大规模集群训练，拓扑感知调度和动态软路由技术可以提升训练性能 **30%**。

4. 盘古系列大语言模型，从 2021 年-2023 年不新突破，凤凰终将涅槃

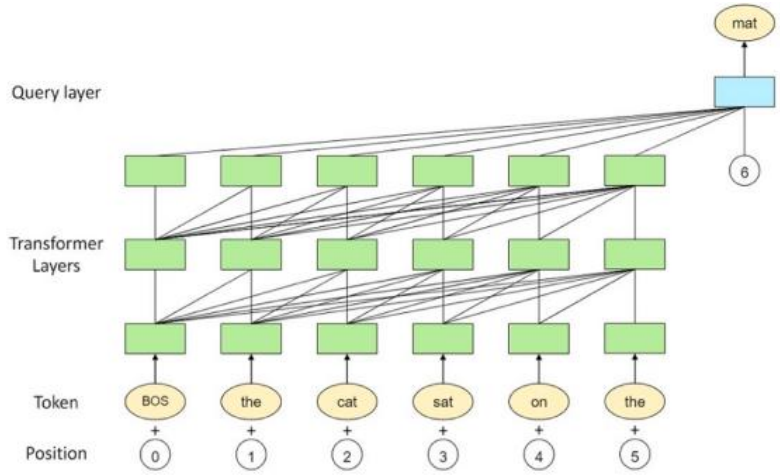
4.1. 鹏程.盘古α：业界首个 2000 亿参数中文自然语言处理大模型

华为基于 MindSpore 框架训练出业界首个 2000 亿参数以中文为核心的预训练生成语言模型。鹏程·盘古α基于“鹏城云脑Ⅱ”和国产 MindSpore 框架的自动混合并行模式，实现了在 2048 卡算力集群上的大规模分布式训练，并完成业界首个 2000 亿参数以中文为核心的预训练生成语言模型。鹏程·盘古α预训练模型支持丰富的场景应用，在知识问答、知识检索、知识推理、阅读理解等文本生成领域表现突出，具备很强的小样本学习能力。

该模型是基于单向的 Transformer decoder 发展而来。query 层堆叠在 transformer 层之上。query 层的基本结构与 transformer 层相似，只是引入了一个额外的 Query layer，来预测生成下一个 query Q 的位置。

盘古α从 80T 原始数据中清洗除了 1.1TB 高质量中文语料数据集投喂训练，总计 250B 规模 Tokens。海量语料是预训练模型研究的基础，联合团队从开源开放数据集、common crawl 网页数据、电子书等收集了近 80TB 原始数据，搭建了面向大型语料库预处理的分布式集群，通过数据清洗过滤、去重、质量评估等处理流程，构建了一个约 1.1TB 的高质量中文语料数据集，经统计 Token 数量约为 250B 规模。通过对不同的开源数据集独立进行处理，完全清除了跟下游任务相关的标签信息，以保证源数据的无偏性。

图 11：鹏程·盘古α网络结构



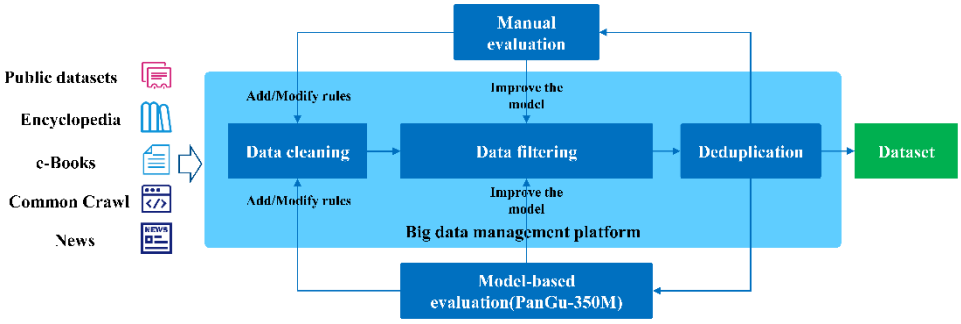
资料来源：昇思官网、天风证券研究所

表 1：1.1TB 中文语料数据组成

数据来源	大小/GB	数据源	数据处理步骤
开放数据集	27.9	15 个开放数据集，如 DuReader、BaiDuQA、CAIL2018、Sougou-CA 等	数据格式转换、文本去重
百科数据	22.0	百度百科、搜狗百科等百科类数据	文本去重
电子书籍	299.0	不同主题的电子书籍，如小说、历史、诗歌、古文等	敏感词过滤、基于模型的文本过滤
Common Crawl	714.9	2018 年 1 约—2020 年 12 月的 Common Crawl 网页数据	数据清洗、过滤、去重等所有数据处理步骤
新闻数据	35.5	1992—2011 年的新闻数据	文本去重

资料来源：《鹏程·盘古：大规模自回归中文预训练语言模型及应用》,作者曾炜、苏腾等、天风证券研究所

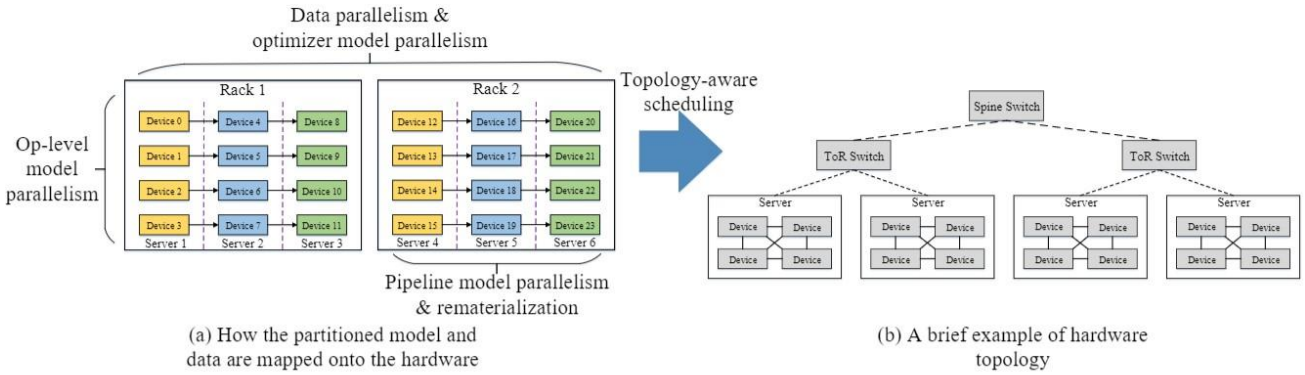
图 12：鹏程·盘古α数据集处理流程



资料来源：启智社区、天风证券研究所

盘古大模型基于 MindSpore 超大规模自动并行完成训练，同时采取全国产硬件。大集群下高效训练千亿至万亿参数模型，用户需要综合考虑参数量、计算量、计算类型、集群带宽拓扑和样本数量等才能设计出性能较优的并行切分策略，模型编码除了考虑算法以外，还需要编写大量并行切分和通信代码。鹏程·盘古α大模型基于国产全栈式软硬件协同生态(MindSpore+CANN+昇腾 910+ModelArts）完成开发。

图 13：模型并行&集群处理架构



资料来源：昇思官网、天风证券研究所

图 14：盘古 α 基于昇腾硬件平台构成的 AI 集群

硬件平台	设备数量	操作系统	集群管理	框架
Ascend 910	2048	EulerOS-aarch64	ModelArts	MindSpore

资料来源：昇思官网、天风证券研究所

4.2. 盘古 Σ：探索万亿参数稀疏模型，最终效果位列中文大模型第一梯队

盘古 Σ 是拥有万亿参数的系数模型，以盘古 α 为基础改进而来。盘古 Σ 基于 Ascend 910 AI 处理器和 MindSpore 框架开发而来，训练完成 1.085T 参数的语言模型。其从盘古 α 继承了的参数，使用随机路由专家（RRE）将密集 Transformer 模型扩展为稀疏模型，并通过专家计算和存储分离（ECSS）高效地训练了 329B 个 Tokens 的模型，最终异构计算的训练吞吐量增加了 6.3 倍。

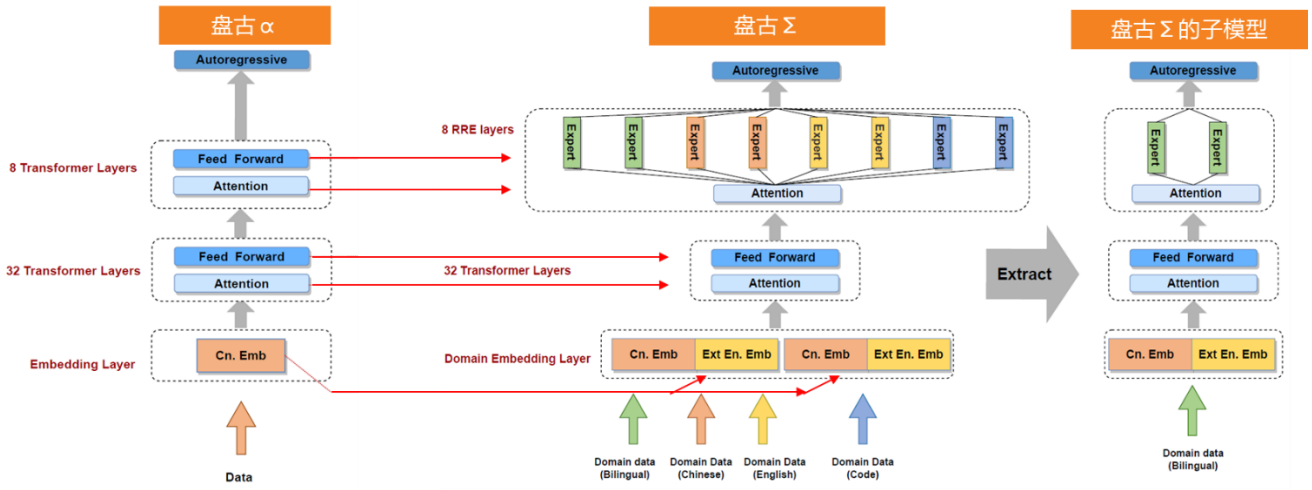
盘古 Σ 架构由密集 Transformer 层和稀疏 Transformer 层混合组成，同样采取自回归语言建模。较低的 M 层是在不同领域之间共享的密集层，上面的 N 个 Transformer 层的前馈部分通过随机路由专家（RRE）进行稀疏激活。

在实际部署环节，盘古 sigma 可以根据实际需求和场景提取和部署子模型，而子模型可能仅包含百亿级参数。通过这样的设计，盘古 sigma 可以实现性能、效率、可用性和部署这四个目标。

在训练数据集上，盘古 sigma 一共收集了 40 个领域的数据集，其中主要来自于中文、英文、双语（中文和英文）和代码这四个领域，其余领域包括 26 种其他单语自然语言、6 种编程语言，以及分别来自金融、健康、法律和诗歌领域的文本数据。

中文数据集包含 200GB 的 WuDaoCorpora 2.0 和 100GB 的 CLUECorpus2020；英文文本包含 800GB 的 the Pile 数据集和包含 750GB 的 C4 数据集；代码部分使用了 147GB 的 Python 代码和来自 GHTorrent 的 161GB 的 Java 代码。以上数据根据文件大小（<1MB）、每行平均字符数（<200）、每行最大字符数（<1000）和它们的可编译性进行过滤，最终获取了超过 3000 亿的 tokens。

图 15：盘古Σ总体架构



资料来源：《PANGU-Σ: TOWARDS TRILLION PARAMETER LANGUAGE MODEL WITH SPARSE HETEROGENEOUS COMPUTING》(作者 Xiaozhe Ren、 Pingyi Zhou 等)、天风证券研究所

图 16：盘古Σ训练数据集中的 4 个主要领域数据来源

Domain ID	Domain	Tokens (Billion)	Data source
0	**Bilingual** (Chinese, English)	77.51 B Chinese (38.75) + English(38.76B)	CLUECorpus2020 , C4
1	**Chinese**	75.47 B	WuDaoCorpora 2.0
2	**English**	75.90 B	Pile , C4
3	**Code** (Python, Java)	75.24 B Python (50.24B) + Java (25B)	Python (PanGu-Coder) Java (GHTorrent)

资料来源：《PANGU-Σ: TOWARDS TRILLION PARAMETER LANGUAGE MODEL WITH SPARSE HETEROGENEOUS COMPUTING》(作者 Xiaozhe Ren、 Pingyi Zhou 等)、天风证券研究所

盘古Σ在中文 NLP 中展现强大性能，在英文语言理解、机器翻译问题上也表现不俗，表明华为在大语言模型上的长期积淀与技术实力。根据论文《PANGU-Σ: TOWARDS TRILLION PARAMETER LANGUAGE MODEL WITH SPARSE HETEROGENEOUS COMPUTING》(作者 Xiaozhe Ren、Pingyi Zhou 等)，研究员将盘古Σ和主流模型做对比，实验结果表明，盘古Σ在各种中文 NLP 下游任务的零样本学习中表现出较强性能。此外，模型在开放域对话、问答、机器翻译和代码生成中也展现出较强能力。

在机器阅读理解、自然语言推理、文本分类、语义相似度分析等中文下游任务下，研究人员选择了 16 种数据集来验证，并选择了 ERNIE 3.0 Titan 和盘古 α-13B 作为比较对象，结果如图 17 所示，**与拥有 2600 亿参数的 ERNIE 3.0 Titan 相比，PanGu-Σ在 16 个数据集中的 11 个上超过了百度，在所有数据集上的平均得分高出 3.96 分。**在中文对话能力测评中，研究人员比较了包括自我聊天、主题对话聊天、开放领域问答等领域，结果也是强于所选择的 baselines。

在英文自然语言理解下，研究人员在 SuperGLUE 作为测试基准，其中包含了 8 项自然语言理解任务。此测试中，研究人员从万亿模型中抽去了一个 380 亿参数的子模型与 GPT-3 13B 在对比了 zero-shot 能力。结果显示出，虽然盘古Σ仅训练了 1120 亿 tokens 的英文语料，但仍然表现出了不输于 GPT-3 13B 的能力。

在机器翻译上，研究员将盘古Σ与包括 CeMAT、ERNIE3.0 等模型进行了比较，评价指标为 SacreBLEU，测试集为 WMT17 和 WMT20，主要验证中英文的翻译。这里的盘古Σ模型

在翻译任务的数据集上进行了 fine-tune，最终结果表明，盘古 Σ 展现出了较强的机器翻译能力。

此外，在代码生成任务上，研究人员在 MBPP 任务上进行了测试，MBPP 是一个测量与训练模型从自然语言生成 Python 语言能力的任务集，最终结果也表明模型在代码生成上能力较强。

图 17：盘古 Σ 在多种中文下游任务下表现出较强的性能

Task Type	Dataset	Split	Metric	PanGu-α 13B	ERNIE 3.0 Titan	PanGu-Σ
Reading comprehension	CMRC2018	dev	avg(EM/F1)	10.37(1.46/19.28)	30.41(16.62/44.20)	**31.23(15.97/46.49)**
	DRCD	dev	avg(EM/F1)	5.61(0.66/10.55)	29.46(21.08/37.83)	**37.78(27.70/47.86)**
	DuReader	dev	ROUGE-1	24.46	32.13	**32.20**
	C3	dev	Acc	54.47	54.85	**56.93**
Natural language inference	CMNLI	dev	Acc	48.44	**51.70**	51.14
	OCNLI	dev	Acc	41.53	44.61	**45.97**
Text classification	TNEWS	dev	Acc	60.26	**72.60**	69.19
	IFLYTEK	dev	Acc	73.80	**79.84**	75.72
Semantic similarity	AFQMC	dev	Acc	65.76	**68.99**	68.49
	CSL	dev	Acc	49.30	55.80	**56.93**
Winograd Schema Challenge	CLUEWSC2020	dev	Acc	75.00	81.08	**85.20**
Cloze and completion	CHID	dev	Acc	70.64	**86.21**	81.01
	PD	test	Acc	43.86	67.06	**77.80**
	CFT	test	Acc	46.60	66.14	**86.84**
	CMRC2017	test	Acc	38.90	74.63	**83.57**
	CMRC2019	dev	Acc	68.19	75.00	**93.87**
/	Overall	/	Average	48.57	60.66	**64.62**

资料来源：《PANGU-Σ: TOWARDS TRILLION PARAMETER LANGUAGE MODEL WITH SPARSE HETEROGENEOUS COMPUTING》（作者 Xiaozhe Ren、 Pingyi Zhou 等）、天风证券研究所

图 18：盘古 Σ 在英语自然语言理解上表现不俗

Dataset	Metric	GPT3 13B	PanGu-Σ
BoolQ	acc	**66.2**	65.54
CB	acc	19.6	**55.36**
Copa	acc	**84.0**	79.00
RTE	acc	**62.8**	59.21
WiC	acc	0.0	**50.78**
WSC	acc	**64.4**	63.46
MultiRC	$F1_a$	**71.4**	59.31
ReCoRD	acc	**89.0**	84.37
SuperGLUE	average	57.2	**64.62**

资料来源：《PANGU-Σ: TOWARDS TRILLION PARAMETER LANGUAGE MODEL WITH SPARSE HETEROGENEOUS COMPUTING》（作者 Xiaozhe Ren、 Pingyi Zhou 等）、天风证券研究所

图 19：盘古 Σ 在机器翻译上能力较强

Data	WMT20	
Lang	Corpus	BLEU
mT5-XXL	26.0M	24.0
CPM-2	26.0M	26.2
Ernie3.0	26.0M	26.8
CeMAT	26.0M	37.1
PanGu-Σ	26.0M	36.6
PanGu-Σ (Low-resource)	0.3M	31.0

资料来源：《PANGU-Σ: TOWARDS TRILLION PARAMETER LANGUAGE MODEL WITH SPARSE HETEROGENEOUS COMPUTING》（作者 Xiaozhe Ren、 Pingyi Zhou 等）、天风证券研究所

5. 建议关注

我们认为华为在 AI 软硬件全栈的布局较全面，除去硬件端的优势，华为在 AI 软件层也国内领先，基于此我们看好华为的底层算力伙伴企业与基于盘古大模型生态的软件应用。

1. 底层算力：东华软件、拓维信息、四川长虹、神州数码、广电运通、卓易信息
2. to C 入口重塑：四维图新、石基信息
3. to C 应用增效：金山办公、同花顺、科大讯飞、视源股份（与电子组联合覆盖）、万兴科技、光云科技
4. to B 应用增效：
1）企业服务：泛微网络、金蝶国际、致远互联、用友网络、软通动力、中国软件国际、汉得信息、东方国信
2）金融科技：恒生电子、长亮科技、顶点软件、中科软、宇信科技、金证股份

3）视频多模态：中科创达（与电子组联合覆盖）、海康威视、当虹科技、智洋创新、东方电子

5. AI 监管：深桑达、安恒信息、启明星辰、美亚柏科、深信服、信安世纪、三未信安、博汇科技

6. 风险提示

1）**AI 应用落地不及预期：**若 AI 相关应用的落地不及预期，相关公司或将受到影响；

2）**硬件领域发展不及预期：**华为受到美国制裁，在芯片端存在诸多限制，如未来制裁进一步加剧，公司在 AI 软硬件栈的布局或受到影响；

3）**大模型进展不及预期：**GPT 系受到国内外广泛关注，国内加速国产大模型研发，华为若后续研发进展及应用不及预期，相关公司或将受到影响。

分析师声明

本报告署名分析师在此声明：我们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，本报告所表述的所有观点均准确地反映了我们对标的证券和发行人的个人看法。我们所得报酬的任何部分不曾与，不与，也将不会与本报告中的具体投资建议或观点有直接或间接联系。

一般声明

特别声明

投资评级声明

类别	说明	评级	体系
股票投资评级	自报告日后的 6 个月内，相对同期沪深 300 指数的涨跌幅	买入	预期股价相对收益 20%以上
		增持	预期股价相对收益 10%-20%
		持有	预期股价相对收益-10%-10%
		卖出	预期股价相对收益-10%以下
行业投资评级	自报告日后的 6 个月内，相对同期沪深 300 指数的涨跌幅	强于大市	预期行业指数涨幅 5%以上
		中性	预期行业指数涨幅-5%-5%
		弱于大市	预期行业指数涨幅-5%以下

天风证券研究

北京	海口	上海	深圳
北京市西城区佟麟阁路 36 号	海南省海口市美兰区国兴大道 3 号互联网金融大厦 A 栋 23 层 2301 房	上海市虹口区北外滩国际客运中心 6 号楼 4 层	深圳市福田区益田路 5033 号平安金融中心 71 楼
邮编：100031	邮编：570102	邮编：200086	邮编：518000
邮箱：research@tfzq.com	电话：(0898)-65365390	电话：(8621)-65055515	电话：(86755)-23915663
	邮箱：research@tfzq.com	传真：(8621)-61069806	传真：(86755)-82571995
		邮箱：research@tfzq.com	邮箱：research@tfzq.com