

计算机行业跟踪周报

华为盘古全栈自主，国产 toB 模型领军

增持（维持）

2024 年 06 月 11 日

证券分析师 王紫敬

执业证书：S0600521080005
021-60199781

wangzj@dwzq.com.cn

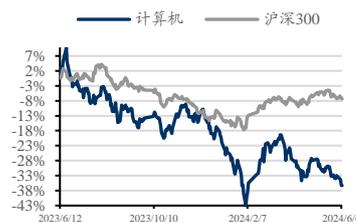
证券分析师 王世杰

执业证书：S0600523080004
wangshijie@dwzq.com.cn

投资要点

- **华为盘古是国产 toB 模型首选**：2024 年 3 月，工信部表示，要开展“人工智能+”行动，促进人工智能与实体经济深度融合，推动人工智能赋能新型工业化。华为盘古大模型坚持 AI for Industries，布局政务、金融、制造、矿山、气象等多个行业，依托华为强大的 toB 业务实现行业落地。
- **盘古大模型全栈自主，安全可靠**：华为盘古大模型是中国首个全栈自主的 AI 大模型。盘古大模型基于华为自主的昇腾人工智能生态，包括底层的鲲鹏和昇腾 AI 算力云平台，中上层的昇腾计算引擎 CANN、AI 开发框架 MindSpore，以及 AI 开发平台 ModelArts。
- **昇腾算力提供底层支持，新品值得期待**：华为昇腾在国产算力中一马当先。昇腾 AI 云服务可以提供更长稳的 AI 算力服务，千卡训练 30 天长稳率达到 90%，断点恢复时长不超过 10 分钟。2024 年 5 月，中移动华为昇腾 191 亿元大单落地。昇腾新一代新品有望推出，为盘古大模型发展提供稳定算力保障。
- **端侧 AI 风起，华为软硬件一体盘古最为受益**：微软、英伟达等全球科技巨头纷纷布局 AIPC 等端侧 AI，苹果公司也与 OpenAI 达成合作，将 ChatGPT 集成到 iPhone 的操作系统中，提供生成式 AI 服务。硬件是 AI 行业落地的最终载体。华为作为全球稀缺的同时在大模型、算力、硬件终端、操作系统等 AI 全产业链配置的公司，有望成为 AI 端侧落地的最大受益者。2024 年 4 月，华为发布首款接入华为盘古大模型的 AIPC MateBook X Pro。
- **投资建议**：政策强调 AI 大模型产业落地，华为盘古大模型聚焦多个行业，是中国首个全栈自主的 AI 大模型，拥有从 AI 模型到算力全栈自主产品，有望成为中国 AI 发展的 toB 主力军。同时，全球 AI 产业逐步向端侧落地，华为作为全球稀缺的同时在大模型、算力、硬件终端、操作系统等 AI 全产业链配置的公司，有望成为 AI 端侧落地的最大受益者。
- **相关标的**：**盘古大模型合作伙伴**：赛意信息、能科科技、云鼎科技、梅安森、北路智控、润达医疗、拓维信息、智洋创新、人民网。**昇腾算力产业链标的**：神州数码、高新发展、华丰科技、申菱环境、英维克等。**华为盘古端侧标的**：软通动力、九联科技、润和软件、证通电子等。
- **风险提示**：技术研发进度不及预期；行业竞争加剧。

行业走势



相关研究

《低空经济周总结：区域超预期，低空经济迎来密集政策催化》

2024-05-19

《低空经济基础设施之【规划设计】：规划设计先行，助力低空经济腾飞》

2024-05-14

内容目录

1. 聚焦行业，国产 2B 大模型领军.....	4
2. 全栈自主，安全可靠	5
3. 盘古大模型生态繁荣，与行业头部企业合作	6
4. 昇腾大单频出，910C 新品可期	6
5. 端侧 AI 风起，华为软硬件一体最为受益.....	7
6. 投资建议	7
7. 风险提示	8

图表目录

图 1: 盘古大模型 5+N+X 三层架构	4
图 2: 盘古大模型全栈技术生态.....	5
图 3: 盘古大模型部分行业合作伙伴.....	6

1. 聚焦行业，国产 2B 大模型领军

政策要求加快大模型行业落地。2024 年 3 月，工业和信息化部新闻发言人、总工程师赵志国在发布会上表示，要着力提升产业科技创新能力。加快推动以大模型为代表的人工智能赋能制造业发展。今年将开展“人工智能+”行动，促进人工智能与实体经济深度融合，推动人工智能赋能新型工业化。

盘古大模型坚持 AI for Industries。华为云盘古大模型自 2021 年发布以来，就聚焦行业价值场景，思考如何用 AI 来优化产品研发、生产供应、市场营销等行业客户所关注的问题。

2023 年 7 月，华为盘古大模型 3.0 发布。盘古大模型 3.0 是面向行业的大模型，包含 L0 中 5 类基础大模型、L1 行业大模型及 L2 场景模型三层架构。

盘古大模型 3.0 是一个完全面向行业的大模型系列包括 5+N+X 三层架构。

图1: 盘古大模型 5+N+X 三层架构



数据来源：华为云公众号，东吴证券研究所

L0 层包括自然语言、视觉、多模态、预测、科学计算 5 大基础模型，提供满足行业场景的多种技能。盘古大模型 3.0 可为客户提供 100 亿参数、380 亿参数、710 亿参数和 1000 亿参数的系列化基础大模型，匹配客户不同需求。无论多大参数规模的大模型，盘古提供一致的能力集，包括 NLP 大模型的知识问答、文案生成、代码生成、NL2SQL，

插件调用等能力以及多模态大模型的图像生成、图像理解等能力，供客户和伙伴企业直接调用。

L1 层是 N 个行业大模型。盘古大模型型首要的就是要为行业客户提供行业大模型能力。华为云既可以提供使用行业公开数据训练的行业通用大模型，包括政务、金融、制造、矿山、气象等，也可以基于行业客户的自有数据，在盘古的 L0 和 L1 上，为客户训练自己的专有大模型。在 L0 和 L1 大模型的基础上，华为云还为客户提供了大模型行业开发套件，通过对客户自有数据的二次训练，客户就可以拥有自己的专属行业大模型。

L2 层为客户提供了更多细化场景的模型，更加专注于某个具体的应用场景或特定业务，为客户提供“开箱即用”的模型服务。

2. 全栈自主，安全可靠

华为盘古大模型是中国首个全栈自主的 AI 大模型。

华为盘古大模型基于华为昇腾人工智能生态。华为在最底层构建了以鲲鹏和昇腾为基础的 AI 算力云平台，构建了昇腾计算引擎 CANN、AI 开发框架 MindSpore，以及 AI 开发平台 ModelArts 为大模型开发和运行提供分布式并行加速、算子和编译优化、集群级通信优化等关键能力。基于华为的 AI 根技术，大模型训练效能可以调优到业界主流 GPU 的 1.1 倍。

图2：盘古大模型全栈技术生态



数据来源：华为云公众号，东吴证券研究所

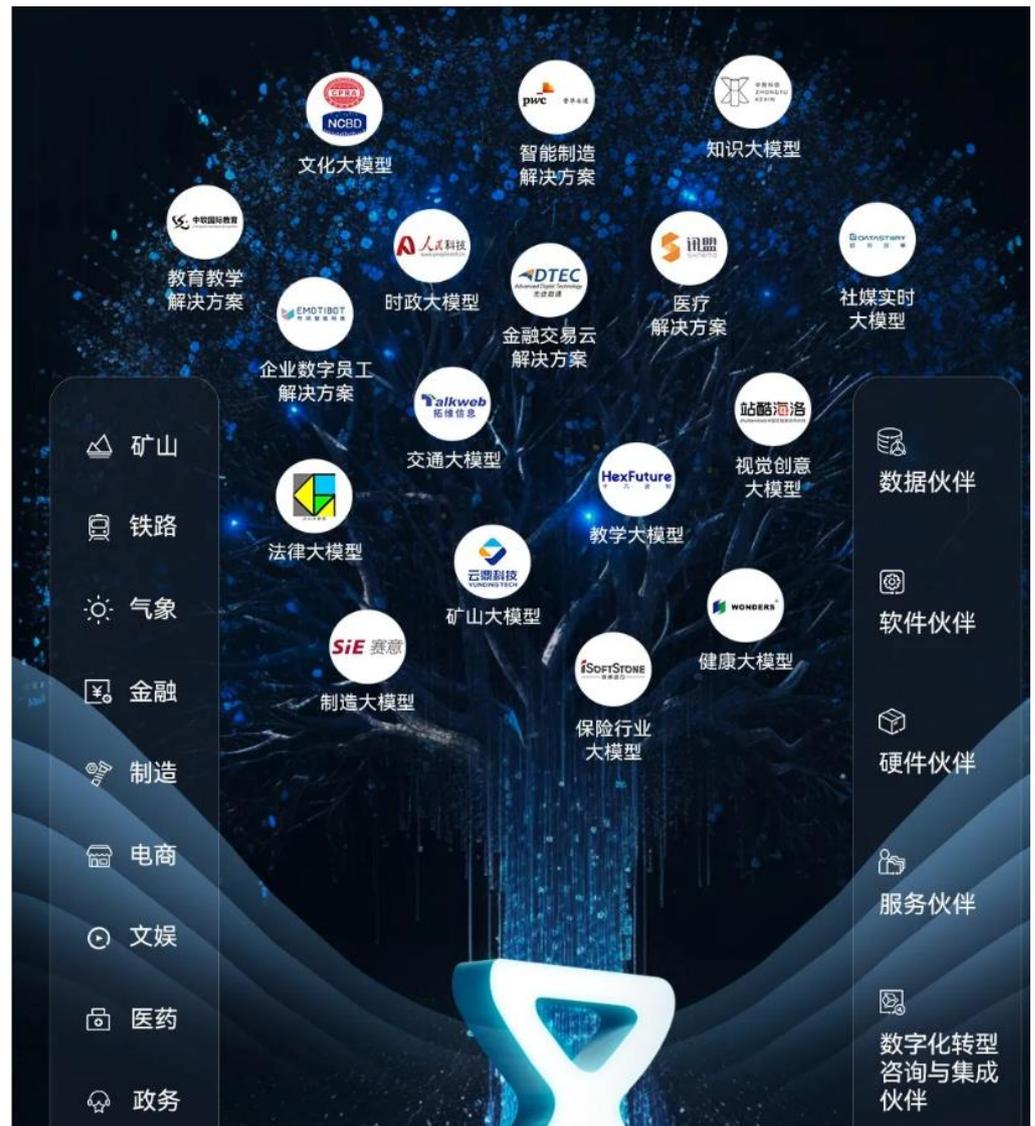
华为云盘古安全护栏获信通院满级评分，全链路筑牢企业大模型可信。3月15日，在中国人工智能产业发展联盟(AIIA)第十一次全会上，中国信通院发布大模型标准符合性验证结果，华为云盘古凭借大模型安全解决方案——盘古安全护栏完成软硬件、数据、

模型、应用等可信模块测试，成为业界首个在大模型可信领域获得卓越级（5级）评分的大模型产品。

3. 盘古大模型生态繁荣，与行业头部企业合作

盘古大模型正在重塑千行百业，与各行业头部企业达成合作。盘古大模型的生态系统已初具规模，众多行业头部客户和伙伴正在与华为云联合创新，基于盘古开发与各具特色的行业大模型，升级各行各业的解决方案。

图3: 盘古大模型部分行业合作伙伴



数据来源：华为云公众号，东吴证券研究所

4. 昇腾大单频出，910C新品可期

算力是AI发展的基石，华为昇腾是国产算力领军。稳定、充足的算力是AI大模型

能够快速迭代的基本要求。在海外制裁背景下，算力国产化迫在眉睫，华为昇腾在国产算力中一马当先。昇腾 AI 云服务可以提供更长稳的 AI 算力服务，千卡训练 30 天长稳率达到 90%，断点恢复时长不超过 10 分钟。

大单频出，910C 新品将至。2024 年 5 月，中移动华为昇腾 191 亿元大单落地。根据芯生代信息，昇腾 910C 芯片目前处于芯片级测试阶段，预计今年第四季度有样机，2025 年第一季度小批量量产。

5. 端侧 AI 风起，华为软硬件一体最为受益

微软发布 AIPC，2025 年 AIPC 占比有望达到 43%。2024 年 5 月 21 日，微软推出一款名为“Copilot+PC”的 AIPC。根据 Gartner 数据，预计到 2025 年 AIPC 出货量会占所有 PC 出货量的 43%。

2024 年 6 月 2 日 —NVIDIA 宣布推出全新 NVIDIA RTX™ 技术，用于支持在全新 GeForce RTX™ AI 笔记本电脑上运行的 AI 助手及数字人。

6 月 10 日消息，苹果 WWDC 24（全球开发者大会）将于明日凌晨 1 点揭幕。据彭博社 6 月 5 日的报道，苹果预计很快宣布与 OpenAI 达成协议，将 OpenAI 的 ChatGPT 集成到 iPhone 的操作系统中，提供生成式 AI 服务。

行业落地是 AI 大模型发展的必然之路，硬件是 AI 行业落地的最终载体。苹果、dell 等硬件公司在 AI 上逐步布局，标志着 AI 在端侧落地的快速推进。华为作为全球稀缺的同时在大模型、算力、硬件终端、操作系统等 AI 全产业链配置的公司，有望成为 AI 端侧落地的最大受益者。

2024 年 4 月，华为发布首款接入华为盘古大模型的 AIPC MateBook X Pro。

6. 投资建议

政策强调 AI 大模型产业落地，华为盘古大模型聚焦多个行业，是中国首个全栈自主的 AI 大模型，拥有从 AI 模型到算力全栈自主产品，有望成为中国 AI 发展的 toB 主力军。同时，全球 AI 产业逐步向端侧落地，华为作为全球稀缺的同时在大模型、算力、硬件终端、操作系统等 AI 全产业链配置的公司，有望成为 AI 端侧落地的最大受益者。

相关标的：

盘古大模型合作伙伴：赛意信息（制造）、能科科技（制造）、云鼎科技（煤矿）、梅安森（煤矿）、北路智控（煤矿）、润达医疗（医疗）、拓维信息（交通）、智洋创新（电力）、人民网（时政）。

昇腾算力产业链标的：神州数码、高新发展、华丰科技、申菱环境、英维克等。

华为盘古端侧标的：软通动力、九联科技、润和软件、证通电子等。

7. 风险提示

技术研发进度不及预期。大模型属于新兴技术，国内与海外仍有一定差距，研发进度可能不及预期。

行业竞争加剧。大模型是未来技术趋势，中国多家巨头宣布进军大模型，未来竞争格局有可能恶化。

免责声明

东吴证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本研究报告仅供东吴证券股份有限公司（以下简称“本公司”）的客户使用。本公司不会因接收人收到本报告而视其为客户。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议，本公司及作者不对任何人因使用本报告中的内容所导致的任何后果负任何责任。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

在法律许可的情况下，东吴证券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并进行交易，还可能为这些公司提供投资银行服务或其他服务。

市场有风险，投资需谨慎。本报告是基于本公司分析师认为可靠且已公开的信息，本公司力求但不保证这些信息的准确性和完整性，也不保证文中观点或陈述不会发生任何变更，在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。

本报告的版权归本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制和发布。经授权刊载、转发本报告或者摘要的，应当注明出处为东吴证券研究所，并注明本报告发布人和发布日期，提示使用本报告的风险，且不得对本报告进行有悖原意的引用、删节和修改。未经授权或未按要求刊载、转发本报告的，应当承担相应的法律责任。本公司将保留向其追究法律责任的权利。

东吴证券投资评级标准

投资评级基于分析师对报告发布日后 6 至 12 个月内行业或公司回报潜力相对基准表现的预期（A 股市场基准为沪深 300 指数，香港市场基准为恒生指数，美国市场基准为标普 500 指数，新三板基准指数为三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的），北交所基准指数为北证 50 指数），具体如下：

公司投资评级：

- 买入：预期未来 6 个月个股涨跌幅相对基准在 15% 以上；
- 增持：预期未来 6 个月个股涨跌幅相对基准介于 5% 与 15% 之间；
- 中性：预期未来 6 个月个股涨跌幅相对基准介于 -5% 与 5% 之间；
- 减持：预期未来 6 个月个股涨跌幅相对基准介于 -15% 与 -5% 之间；
- 卖出：预期未来 6 个月个股涨跌幅相对基准在 -15% 以下。

行业投资评级：

- 增持：预期未来 6 个月内，行业指数相对强于基准 5% 以上；
- 中性：预期未来 6 个月内，行业指数相对基准 -5% 与 5%；
- 减持：预期未来 6 个月内，行业指数相对弱于基准 5% 以上。

我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重建议。投资者买入或者卖出证券的决定应当充分考虑自身特定状况，如具体投资目的、财务状况以及特定需求等，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。

东吴证券研究所
苏州工业园区星阳街 5 号
邮政编码：215021

传真：（0512）62938527

公司网址：<http://www.dwzq.com.cn>