



数据要素专题研究

买入（维持评级）

行业专题研究报告

证券研究报告

计算机组

分析师：孟灿（执业 S1130522050001）

mengcan@gjzq.com.cn

大模型与数据共振，数据要素市场方兴未艾

投资逻辑

数据成为生产要素，成为数字经济时代的“石油”。数据已成为新型生产要素，具有劳动工具和劳动对象的双重属性。数据要素市场的目标是实现数据要素的市场化配置，其目的是建立以市场为基础的调配机制，实现数据的流动价值或在流动中产生价值。数据要素市场可以归结为数据采集、存储、加工、流通、分析和生态保障六大模块，主要通过业务贯通、数智决策和流通赋能三条途径实现自身价值。国家工业信息安全发展研究中心数据显示，我国 2021 年数据要素市场规模达 815 亿元，预计“十四五”期间年均增长率达 25%，有望在 2025 年达到 1990 亿元的市场规模。

政策催化强劲，行业景气度拐点已至。2022 年 12 月 19 日，“数据二十条”正式发布，这是继《中共中央国务院关于构建更加完善的要素市场化配置体制机制的意见》之后首次在国家级政策文件中全面明确数据基础制度，2023 年 3 月，国家数据局获批成立，数据要素发展再提速。政策端持续催化下，中国电子以及上海、河南、福建、陕西、成都、南京等地陆续成立数据集团，有望成为各地政府参与数据运营的抓手。此外贵阳、上海、深圳各地数据交易所陆续挂牌运营，截至 2022 年底，全国数据交易所已近 50 家，数据产品和服务类型日益丰富，加速了数据要素价值转化。

AI 大模型向行业垂类模型落地演进，数据成为核心壁垒。数据贯穿 AI 垂类模型训练的始终，AI 垂类模型强调领域的 know-how，对数据在深度和质量上的要求更高，增加训练数据量对模型性能提升来说更具性价比。行业数据成为国内垂类模型卡点，管国内 AI 大模型已经取得较大进展，但是行业数据的可得性限制了垂类模型的应用落地。当前阶段，拥有大量数据积累并且具备行业 know-how 的企业能够在大模型的基础上学习行业特色数据与知识，打造出 AI 垂类模型。

AI 垂类模型加速数据要素市场发展，推动数据流动。根据艾瑞咨询数据，2022 年我国 AI 基础数据服务市场规模为 31 亿元，2027 年有望达到 79 亿元，2022-2027 年 CAGR 达 20.6%。在垂类模型的训练过程中，对数据量、数据集的丰富度和全面性以及数据的及时性都提出了一定的要求，数据要素市场打通了“数据要素供给-数据要素流通-垂直领域模型研发及应用”循环发展的产业链条，其提供的海量丰富的训练数据是垂类模型的基础原料，市场流通则是连接数据要素与垂类模型的枢纽。在垂类模型的推动下，数据要素市场的新业态也不断涌现，多地以数据交易所为主导，建立数据要素与大模型的“产业联盟”，数据要素市场建设再加速。

投资建议

当前我国数据要素市场正面临从零到一的重大拐点，我们预判数据要素市场是中国未来十年极具成长潜力的赛道，板块将持续受益政策加持和产业发展逐步验证的双重催化。我们建议关注重点环节的“卡位”逻辑，从数据要素的供给侧、运营端、安全防护三个维度进行分析，筛选出兼具基本面和商业模式初具雏形的优质企业进行提前布局。数据要素供给侧推荐中科江南、航天宏图，数据要素运营端推荐易华录、深桑达 A，安全防护侧推荐永信至诚。

风险提示

政策落地不及预期、数据资产化发展不及预期。



内容目录

1. 数据要素市场是数字经济发展的核心引擎.....	5
1.1 数据成为生产要素，成为数字经济时代的“石油”.....	5
1.2 政策催化强劲，行业景气度拐点已至.....	7
2. 垂类 AI 大模型加速落地，推动数据要素市场发展.....	13
2.1 AI 大模型向行业垂类模型落地演进，数据成为核心壁垒.....	13
2.2 AI 垂类模型加速数据要素市场发展，推动数据流动.....	18
3. 投资建议.....	20
3.1 中科江南：财政、医保信息化领导者.....	23
3.2 航天宏图：空天地一体化数据领域龙头.....	24
3.3 浩瀚深度：运营商数据可视化市场领军者.....	24
3.4 金山办公：国产办公软件龙头.....	25
3.5 深桑达 A：CEC 旗下电子云和数据要素的核心运营者.....	26
3.6 易华录：CETC 旗下数据安全存储和运营的先行者.....	26
3.7 星环科技：国产分布式数据库和向量数据库龙头.....	27
3.8 太极股份：CETC 旗下政务云基础设施承建方.....	29
3.9 永信至诚：人工智能模型安全与数据安全实战化测评能力提供方.....	29
3.10 英方软件：大数据灾备领域龙头.....	31
4. 风险提示.....	31



图表目录

图表 1: 数据成为新型生产要素且具有劳动工具和劳动对象双重属性	5
图表 2: 数据要素通过业务贯通、数智决策和流通赋能三条途径实现价值释放	6
图表 3: 数据要素产业链六大模块	6
图表 4: 数据要素产业图谱	7
图表 5: 2025 年我国数据要素市场规模预达 1990 亿元	7
图表 6: 2022 年我国数据要素市场各产业链占比	7
图表 7: 数据要素相关政策不断出台	8
图表 8: “数据二十条”建立数据要素整体框架	9
图表 9: 我国数字经济占 GDP 比例约 40%	10
图表 10: 国家数据局管理职责和权限	10
图表 11: 我国中央及地方性数据集团	10
图表 12: 国内大数据交易所（中心、平台等）建设历程	11
图表 13: 我国土地财政收入增速下降	12
图表 14: 欧洲地区率先实行数据服务税	12
图表 15: “土地财政”和“数据财政”具备双规并行的可能性	13
图表 16: 全球知名大模型发布时间节点	13
图表 17: 数据贯穿 AI 垂类模型训练的始终	14
图表 18: 模型性能与训练数据量、模型参数规模、训练计算量呈正相关关系	14
图表 19: BloombergGPT 金融领域数据集共包含 3630 亿个 token	15
图表 20: BloombergGPT 在金融语料上的 bits per byte 均好于其他模型	15
图表 21: 拓尔思三大数据资产平台拥有千亿级优质数据积累	15
图表 22: 拓天·M 媒体大模型的行业基础	16
图表 23: 拓天·F 金融大模型的行业基础	16
图表 24: 医疗 MedGPT 参数和训练数据规模庞大	16
图表 25: MedGPT 将不断提升病种覆盖率	16
图表 26: MedGPT 与三甲主治医师医学一致性达到 96%	17
图表 27: 华为盘古大模型 3.0 重塑千行百业	17
图表 28: 盘古药物分子大模型	18
图表 29: 2020-2027 年中国 AI 基础数据服务市场规模	18
图表 30: 2020-2027 年中国数据治理与面向 AI 的数据治理市场规模	18
图表 31: 数据要素市场体系架构	19
图表 32: 高质量产业大数据大模型联盟成立	20
图表 33: 南海区大数据人工智能产业联盟成立	20



图表 34: 上海数据交易所上线语料库, 助力垂类模型训练..... 20

图表 35: 推荐标的..... 21

图表 36: 中科江南行业电子化应用平台..... 23

图表 37: 航天宏图扎实推进全产业布局..... 24

图表 38: 浩瀚深度可全面实现互联网流量及数据的前端采集处理和后端智能化应用..... 25

图表 39: 金山办公产品结构..... 25

图表 40: 深桑达聚焦四项主责主业..... 26

图表 41: 蓝光存储具备多重优势..... 27

图表 42: 易数工厂产品超市..... 27

图表 43: 星环科技定位于数字化基础设施的底层和中间层..... 28

图表 44: ArgonDB 数据库和 KunDB 数据库通过信通院“可信数据库”评测..... 28

图表 45: 人大金仓数据库积极构建信创产业生态..... 29

图表 46: 永信至诚产品服务体系生态链条..... 30

图表 47: 数字风洞打造数据安全测试评估标准平台..... 30

图表 48: 英方软件产品系列面向灾备和大数据两大应用场景..... 31



1. 数据要素市场是数字经济发展的核心引擎

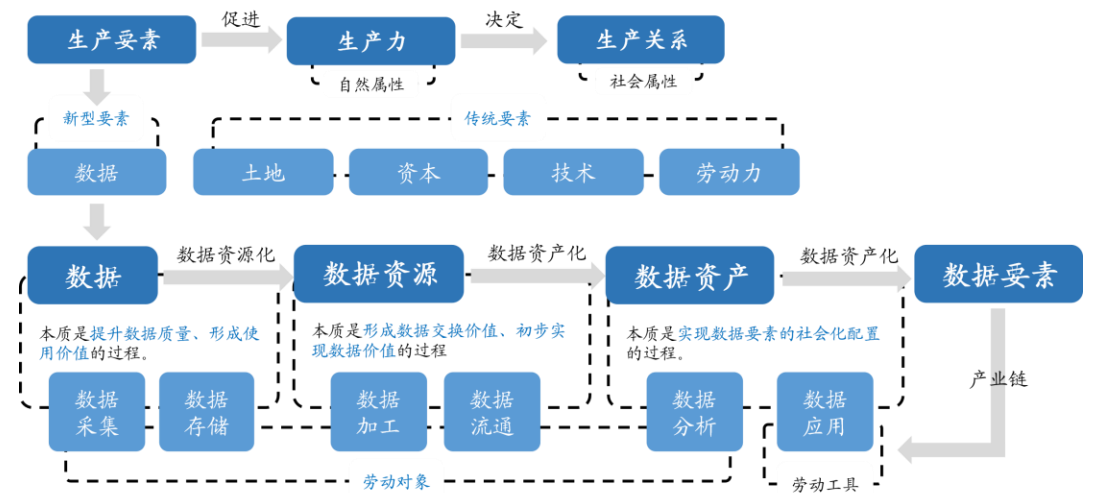
1.1 数据成为生产要素，成为数字经济时代的“石油”

数据要素作为较新的经济学概念，准确清晰地理解和定义数据要素市场相关概念是探索和培育数据要素市场模式和方向的重要前提，根据国家工业信息安全发展研究中心报告，我们给出以下对数据、数据资源和数据资产的定义：

- 1) 数据：是指所有能够输入计算机程序处理、反映一定事实、具有一定意义的符号介质的总称。
- 2) 数据资源：是指按一定规则排列组合的物理符号集合，用于承载或记录信息，这些信息可以是数字、文字、图像，也可以是计算机代码的集合。
- 3) 数据资产：本质上是产权的概念，是指个人或企业拥有或控制的、以物理或电子方式记录的能够为个人或企业带来经济利益的数据资源。

数据已成为新型生产要素，具有劳动工具和劳动对象的双重属性。生产要素是指进行社会生产经营活动所必需的资源和环境条件，随着经济的发展，生产要素的具体形态和主次序列不断变化，传统的生产要素主要包括土地、资本、技术和劳动力，在 5G、物联网、云计算、大数据、区块链和人工智能等技术的共同作用下，数据对生产的贡献日益突出。作为新型生产要素，数据具有劳动工具和劳动对象的双重属性，作为劳动对象，数据通过采集、加工、存储、流通和分析等环节具有价值和使用价值，作为劳动工具，数据通过融合应用能够提高生产效率，推动生产力的发展。数据要素市场的目标是实现数据要素的市场化配置，将数据要素从尚未完全由市场配置转向由市场配置的动态过程，其目的是建立以市场为基础的调配机制，实现数据的流动价值或在流动中产生价值。

图表1：数据成为新型生产要素且具有劳动工具和劳动对象双重属性



来源：国金证券研究所

数据要素主要通过业务贯通、数智决策和流通赋能三条途径实现自身价值。首先，数据投入生产的一次价值体现在支撑企业和政府的业务系统运转，实现业务间的贯通，为释放数据要素的初级价值，政府和企业的主要任务是推进业务的数字化和各类业务信息系统的建设。其次，数据要素二次价值释放体现在通过加工、分析、建模等过程，揭示出更深层次的关系和规律，从而使生产、经营、服务和治理等环节的决策更加智慧、智能和精准，该过程对企业的数据挖掘和分析能力提出了更高的要求。最后，数据要素的三次价值释放体现在将数据流通到更需要它的地方，让不同来源的优质数据在新的业务需求和场景中汇聚和融合，实现双赢和多赢的价值利用。



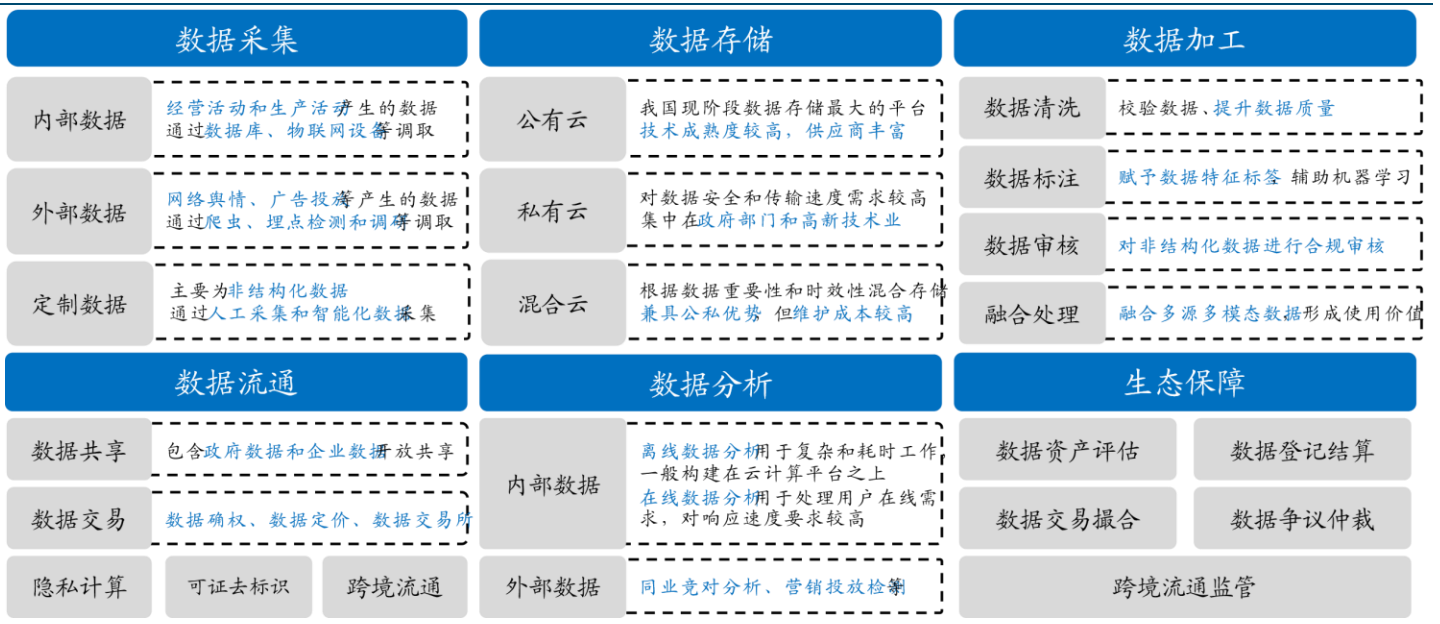
图表2: 数据要素通过业务贯通、数智决策和流通赋能三条途径实现价值释放



来源: 中国信息通信研究院, 国金证券研究所

从产业链的角度出发, 我们将数据要素市场归结为数据采集、存储、加工、流通、分析和生态保障六大模块。数据采集环节关注确保数据采集的准确性和全面性。数据存储环节关注确保数据存储的安全性和实时调用的可行性。数据加工环节关注确保数据加工的精度和准确性。数据流通环节是数据要素市场的核心环节, 关注在保障所有者权益的前提下进行合理合规的数据流通。数据分析环节关注深度分析和挖掘数据的价值和潜力。生态保障环节包括数据资产评估、登记结算、交易撮合、争议仲裁以及跨境流动监管等, 旨在为数据要素市场各主体提供有效的保障, 并构建一个良好的市场生态。

图表3: 数据要素产业链六大模块



来源: 国家工业信息安全发展研究中心, 中国信息通信研究院, 国金证券研究所

数据要素产业图谱清晰, 国内企业集中度较高。虽然数据要素作为一个较新的经济学概念, 但与数据相关的产业已经在国内经过了较长时间的发展, 例如在数据存储方面信创之风已经提前带领国内数据库反弹复苏, 以华为、阿里和腾讯为首的云数据库, 以及以达梦、南大通用为首的传统数据库, 还有 OceanBase、GoldenDB 和 TiDB 等国内分布式交易型数据库快速发展。除此之外, AI 大模型的快速崛起也大幅增加了对模型训练的需求, 数据加工中的数据清洗、标注和审核等工作需求激增, 百度 EasyData、海康 NLP 等技术获得关注。数据流通环节作为二级市场的关键环节, 在数据确权和隐私计算方面的国内企业如拓尔思、易华录等公司获得市场资金大幅支持。生态保障的主要参与者是政府机构和组织, 例如北京数据资产评估中心、大数据交易所以及国家数据局等监管机构等。



图表4: 数据要素产业图谱

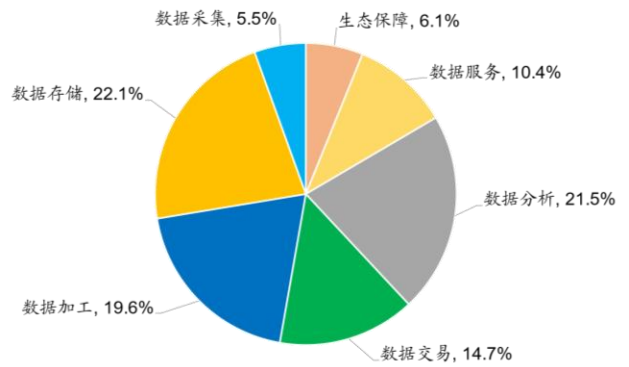


来源: 国家工业信息安全发展研究中心, 国金证券研究所

我国数据要素市场规模持续扩张, 数据要素成为数字经济发展新引擎。数字化的本质是在信息化的基础上, 对系统产生的数据要素, 利用大数据、AI、区块链等数字技术, 进行流程改造、数据决策、商业模式重构等全新的价值开发, 核心目的是实现商业增量。国家工业信息安全发展研究中心数据显示, 我国 2021 年数据要素市场规模达 815 亿元, 预计“十四五”发展期间年均增长率达 25%, 有望在 2025 年达到 1,990 亿元的市场规模。同时可以发现, 2022 年我国数据要素市场中数据存储、分析和加工行业位列前三, 随着人工智能和互联网技术的进步, 数据量的增大对数据存储提出了更高要求, 数据的安全、可靠和隐私存储市场需求逐步扩张, AI 大模型的快速演进带动数据加工中的清洗和标注工作需求激增, 我们看好数据要素行业相关企业未来的快速发展。

图表5: 2025 年我国数据要素市场规模预计达 1990 亿元

图表6: 2022 年我国数据要素市场各产业链占比



来源: 国家工业信息安全发展研究中心, 国金证券研究所

来源: 国家工业信息安全发展研究中心, 国金证券研究所

1.2 政策催化强劲, 行业景气度拐点已至

从政策到顶层管理架构, 数据要素体系根基逐渐稳固。2019 年党的十九届四中全会首次将数据列为生产要素, 表明国家大力发展数字经济的决心, 并标志着数据从资源向要素的转变。2020 年 4 月, 中共中央、国务院发布了《关于构建更加完善的要素市场化配置体制机制的意见》, 明确将数据市场与土地市场、劳动力市场、资本市场和技术市场列为加快培育的五大核心生产要素市场之一, 数据要素进入市场化阶段。2022 年 12 月, 国务院发布“数据二十条”这一纲领性文件, 确立了数据要素发展顶层指导框架。2023 年 3 月, 国家数据局成立, 负责协调推进数据基础制度建设, 统筹数据资源整合共享和开发利用, 统筹推进数字中国、数字经济、数字社会规划和建设等。地方层面, 多个省市密集发布数据要素相关政策, 数据要素产业进入实质性落地阶段。


图表7：数据要素相关政策不断出台

发布时间	发布部门	政策名称	主要内容
2019年10月	中共中央委员会	《关于坚持和完善中国特色社会主义制度推进国家治理体系和治理能力现代化若干重大问题的决定》	首次公开提出“数据可作为生产要素按贡献参与分配”，明确将“数据”作为生产要素，健全各种生产要素由市场评价贡献、按贡献决定报酬的机制。
2020年4月	中共中央、国务院	《关于构建更加完善的要素市场化配置体制机制的意见》	指出将数据作为与土地、劳动力、资本、技术并列的生产要素，要求“加快培育数据要素市场”。
2020年5月	中共中央、国务院	《关于新时代加快完善社会主义市场经济体制的意见》	提出建立数据资源清单管理机制，完善数据权属界定、开放共享、交易流通等标准和措施，发挥社会数据资源价值。
2021年3月	十三届全国人大四次会议	《“十四五”规划和2035年远景目标纲要》	明确提出培育规范的数据交易平台和市场主体，发展数据资产评估、登记结算、交易撮合、争议仲裁等市场运营体系。
2021年6月	第十三届全国人大常委会	《中华人民共和国数据安全法》	规范数据处理活动，保障数据安全，促进数据开发利用，保护个人、组织的合法权益，维护国家主权、安全和发展利益。
2021年11月	工信部	《“十四五”大数据产业发展规划》	要求加快培育数据要素市场，构建稳定高效产业链，打造繁荣有序产业生态，筑牢数据安全保障防线。
2021年12月	国务院	《要素市场化配置综合改革试点总体方案》	进一步破除阻碍要素自主有序流动的体制机制障碍，提升要素协同配置效率作出部署；加快技术、数据、资源环境市场培育和规则建设；健全要素市场治理和促进要素协同配置。
2021年12月	国务院	《“十四五”数字经济发展规划》	规范数据交易管理，提升数据交易效率，建立健全数据资产评估、登记结算、交易撮合、争议仲裁等市场运营体系。
2022年4月	中共中央、国务院	《关于加快建设全国统一大市场的意见》	提出要加快培育数据要素市场，建立健全数据安全、权利保护、跨境传输管理、交易流通、开放共享、安全认证等基础制度和标准规范。
2022年10月	国务院	《全国一体化政务大数据体系建设指南》	提出要加强数据汇聚融合、共享利用，促进数据高效流通使用，充分释放政务数据资源价值，强化数据安全保障，促进数据依法有序流动。
2022年12月	中央中央、国务院	《关于构建数据基础制度更好发挥数据要素作用的意见》	确立数据要素市场的四大原则体系：“数据产权、流通交易、收益分配、安全治理”，推动健全数据确权授权机制，完善和规范数据流动规则，有序发展数据跨境流通和交易。
2022年12月	财政部	《企业数据资源相关会计处理暂行规定（征求意见稿）》	拟对数据资源的相关会计处理做出规定，明确现阶段数据资源会计处理应当按照企业会计准则执行。
2022年12月	中共中央、国务院	《关于构建数据基础制度更好发挥数据要素作用的意见》	以数据产权制度、流通交易制度、收益分配制度、治理制度为核心，明确了数据要素市场制度建设的基本框架、前进方向和工作重点，初步形成我国数据基础制度的“四梁八柱”
2023年2月	中共中央、国务院	《数字中国建设整体布局规划》	指出要夯实数字中国建设基础，强化数字中国关键能力，优化数字化发展环境，明确提出“构建国家数据管理体制机制，健全各级数据统筹管理机构。
2023年3月	国务院	《国务院机构改革方案》	组建国家数据局，负责协调推进数据基础制度建设，统筹数据资源整合共享和开发利用，统筹推进数字中国、数字经济、数字社会规划和建设等。
2023年6月	上海市	《立足数字经济新赛道 推动数据要素产业创新发展行动方案（2023-2025）》	以上海数据交易所为核心抓手，统筹联动张江科学城范围内数据要素功能平台和产业园区，打造“一核三园两港”空间布局，目标成为数据要素交易流通枢纽地、产业发展引领区和应用场景区，打造全国首个数据要素产业集聚区。



2023年6月	北京市	《关于更好发挥数据要素作用进一步加快发展数字经济的实施意见》	推动建立供需高效匹配的多层次数据交易市场，充分挖掘数据资产价值，培育一批数据要素型领军企业。力争到2030年本市数据要素市场规模达到2000亿元，形成数据服务产业集聚区。
2023年6月	深圳市	《深圳市数据产权登记管理暂行办法》	合法取得数据资源或产品的登记主体享有相应的数据资源持有、数据加工使用和数据产品经营等相关权利；经登记机构审核后获取的数据资源或数据产品登记证书、数据资源许可凭证，可作为数据交易、融资抵押、数据资产入表、会计核算、争议仲裁的依据。

来源：中国政府网，共产党员网，中国人大网，财政部，国金证券研究所

顶层文件“数据二十条”搭建数据要素“四梁八柱”。2022年12月19日，《中共中央国务院关于构建数据基础制度更好发挥数据要素作用的意见》正式对外发布，又称“数据二十条”，这是继2020年4月10日发布的《中共中央国务院关于构建更加完善的要素市场化配置体制机制的意见》之后首次全面明确国家级政策文件中的数据基础制度。“数据二十条”涵盖了四个制度领域，包括数据产权、流通交易、收益分配和安全治理。在数据产权方面，建立数据资源持有权、数据加工使用权和数据产品经营权的“三权分置”结构。在流通交易环节，建立场内外的数据要素流通和交易制度。在收益分配方面，遵循了“谁投入、谁贡献、谁受益”的原则。在实行贡献值分配的基础上，还关注公益和相对弱势群体的利益，在再分配环节进行相应的考虑。在安全治理方面，通过“以链治数”的监管模式，实现了数据要素的安全可信流通。“数据二十条”的发布对于更好地发挥数据要素的作用具有重要意义，为数据基础制度的建设提供了全面的指导和规范。

图表8：“数据二十条”建立数据要素整体框架

01 建立保障权益、合规使用的数据产权制度

1. 探索数据产权结构性分置制度：

2. 明确数据分类分级确权制度：公共数据授权使用、企业数据供给激励、个人数据受托机制。

02 建立合规高效、场内场外结合的数据要素流通和交易制度

1. 完善数据全流程合规与监管规则体系：数据流通和数据安全两手抓，形成多方协同的数据治理格局，促进数据要素有序流通。
2. 统筹构建规范高效的数据交易场所：场内场外数据交易市场并行发展，且未来朝着“场内重发展、场外重规范”的方向演进。
3. 培育数据要素流通和交易服务生态：鼓励企业积极探索数据空间、数据链等产业数据生态，培育数据要素流通和交易服务多方生态。
4. 构建数据安全合规有序跨境流通机制。

“数据二十条”整体框架

1. 健全数据要素由市场评价贡献、按贡献决定报酬机制：在坚持“两个不动摇”的前提下，明确“谁投入、谁贡献、谁受益”原则下的数据要素价值分配机制。
2. 更好发挥政府在数据要素分配收益中的引导调节作用：特别提到“着重保护数据要素参与各方的投入产出收益”，“推动数据要素收益向数据价值和使用权创造者合理倾斜”，“强化基于数据价值创造和实现的激励导向”，“通过分红、提成等多种收益共享的方式”来平衡不同环节相关主体间的利益分配，体现按照要素合理取酬的原则。

1. 创新数据治理机制：政府强化分行业监管和跨行业协同监管，强调“有效市场”和“有为政府”相结合，发挥政府的有序引导和规范发展作用。
2. 压实企业的数据治理责任：企业在数据安全基础设施建设、日常管理和技术研发等方法积极发挥作用，配合实施数据安全认证，提升数据安全管理水平。
3. 充分发挥社会力量多方参与的协同治理作用：鼓励多方积极参与数据要素市场建设和治理，鼓励行业协会积极参与协同治理。

03 建立体现效率、促进公平的数据要素收益分配制度

04 建立安全可控、弹性包容的数据要素治理制度

来源：中国政府网，普华永道，上海数据交易所，国金证券研究所

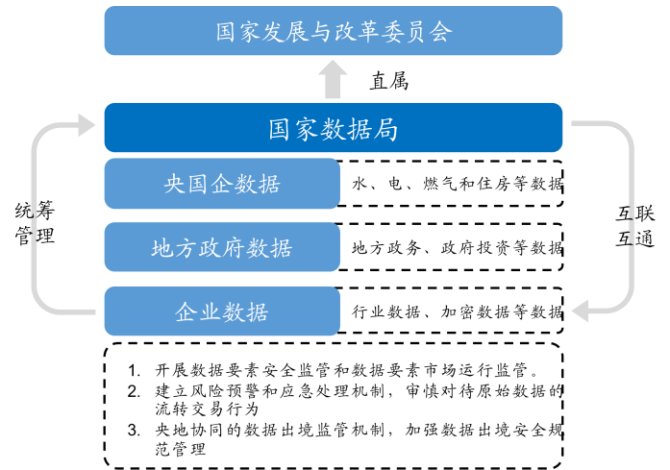
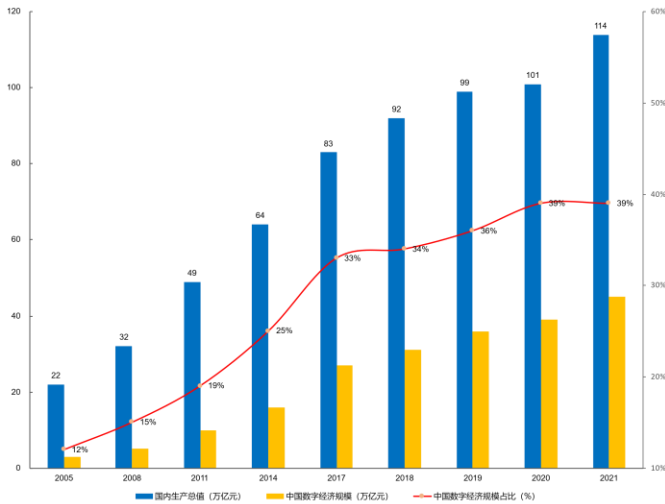
国家数据局获批成立，职责权限明晰集中有望实现数据要素发展再提速。2023年3月，《党和国家机构改革方案》印发，该方案在保持现有工作格局的总体稳定的前提下，提出了整合共享和开发利用数据资源的职责相对集中的要求。根据该方案，国家数据局作为国家发展和改革委员会管理的国家局，负责协调推进数据基础制度建设，统筹推进数字中国、数字经济、数字社会规划和建设等任务。我国数字经济规模全球排名第二，约占国内经济总



量的40%左右，设立国家数据局展示了与时俱进的态度，体现了国家对数字经济发展的顶层设计，符合时代要求和发展需求。国家数据局由发展和改革委员会直接管理，有助于消除部门之间、系统之间和地区之间的壁垒，实现数据要素的互联互通，将解决过去数字经济管理中的碎片化问题，并减少地方数字经济发展不均带来的数字鸿沟，提高数据交换的效率和准确性，加快数据要素产业的发展进程。

图表9：我国数字经济占GDP比例约40%

图表10：国家数据局管理职责和权限



来源：艾瑞咨询，国金证券研究所

来源：中国政府网，国金证券研究所

多个省市成立数据集团，有望成为各地政府参与数据运营的抓手。央企层面，中国电子数据产业集团于2022年12月成立，是国内首家由中央企业设立的数据产业集团；地方层面，上海、河南、福建、陕西、成都、南京等地陆续成立数据集团，多为政府主导。数据集团以数据为核心业务，实现公共数据、行业数据和社会数据的交汇、供给、配置及市场化开发利用，开展数字资产运营、数据交易服务和数字产业投资。

图表11：我国中央及地方性数据集团

名称	成立时间	注册资本
盐城市大数据集团	2012年1月(2021年更名)	30亿元
成都市大数据集团	2013年7月(2021年更名)	40亿元
南通市大数据开发集团	2014年12月(2017年更名)	1.5亿元
陕西省大数据集团	2017年4月	8亿元
云上贵州大数据(集团)	2018年10月	17亿元
贵阳市大数据产业集团	2019年2月(2022年更名)	10亿元
苏州市大数据集团	2019年5月(2022年更名)	20亿元
南京大数据集团	2020年6月	10亿元
烟台市大数据发展集团	2021年8月	10亿元
福建省大数据集团	2021年8月(2022年更名)	100亿元
无锡大数据集团	2022年3月	5000万元
上海数据集团	2022年9月	50亿元
福州数据集团	2022年11月	10亿元
中国电子数据产业集团	2022年12月	30亿元
河南数据集团	2023年1月	10亿元
武汉数据集团	2023年4月	20亿元
湖北数据集团	2023年6月	50亿元

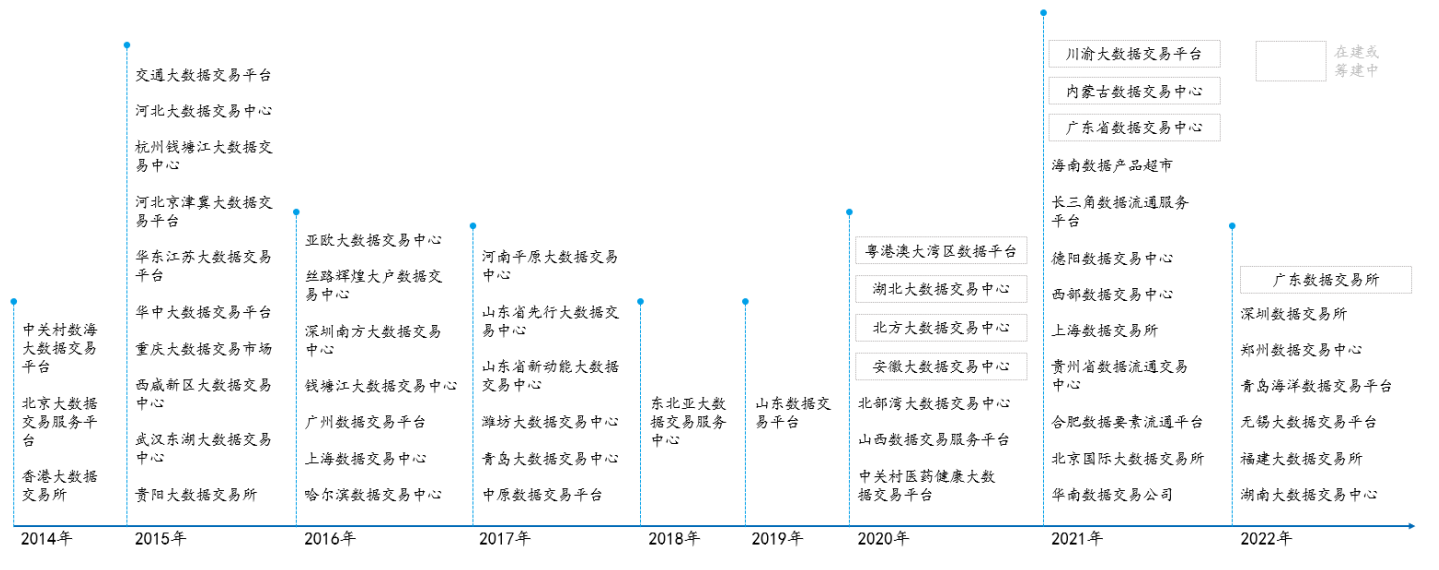
来源：大信创圈微信公众号，国金证券研究所

各地大数据交易所陆续挂牌运营，加速数据要素价值转化。我国自2014年开始探索建立类似证券交易所形式的数据交易机构，随着数据要素相关政策的推动，数据产品交易迎来



2.0 时代。2015 年 4 月，贵阳大数据交易所正式挂牌，根据数据交易网，截至 2023 年 6 月 21 日，贵数所已累计集聚“数据商”、“数据中介”等市场主体 629 家，上架产品 1055 个，交易 888 笔，共计交易额达 14443 万元。2021 年 11 月，上海数据交易所揭牌交易，根据数据交易网，2022 年上海数据交易所数据产品挂牌超 800 个，涉及金融、交通、工业、通信等 12 个行业领域，交易金额突破 1 亿元。2022 年 11 月，深圳数据交易所挂牌运营，截至 2023 年 6 月，深数所已累计交易突破 700 笔，覆盖 165 个应用场景，生态合作机构突破 900 家，汇集数据产品超 1500 个，服务触达 2000 家以上市场主体。截至 2022 年底，全国数据交易所已近 50 家。各大数据交易所交易主题、上架产品以及交易规模都处于快速发展态势，数据产品和服务类型日益丰富，能够提供数据 API、数据集、数据报告等多种形式的产品和服务。

图表12：国内大数据交易所（中心、平台等）建设历程

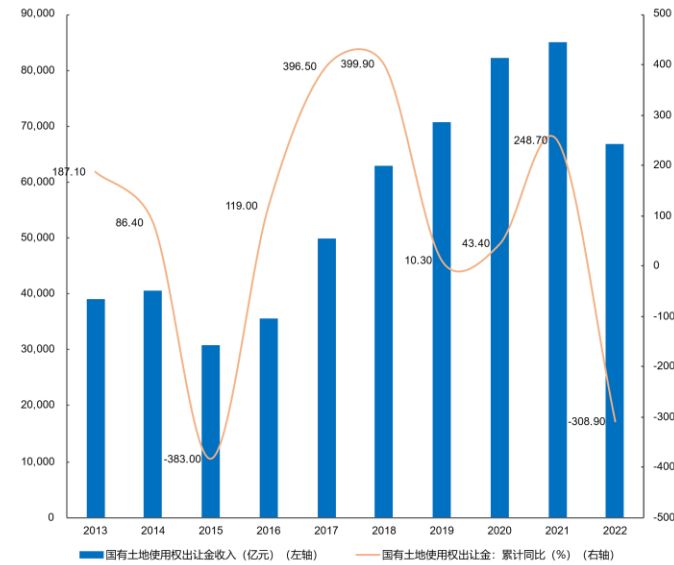


来源：大信创圈微信公众号，国金证券研究所

土地财政收入增速下降，数据要素探索“数据财政”可能性。土地财政主要指政府通过出售土地或者收取土地使用权等方式获取财政收入的政策，然而土地资源的有限性在我国经济转型升级和高质量发展的背景下，探索通过财政税收工具和手段从数字经济的关键要素中取得一定比重的财政收入，并发挥财政的基础性、支柱性国家治理作用是非常有意义的。数据财政是基于大数据和人工智能等新兴技术的发展，将数据视为新的财富来源的理念，具体而言数据财政包括数据进入市场之前，在数商组建、数据产品开发、登记等环节的监管伴随的财政活动，以及数据进入市场之后，在数据开发、服务、再生产等过程，通过财税手段实现对数据市场的激励和管理。欧洲地区率先对数据服务进行征税，在国际税收改革背景下提出的一种对数字服务进行征税的新模式，主要针对的是那些利润丰厚但却往往在欧洲本土纳税不多的跨国数字巨头，给我国推行数据财政提供了参考意义。

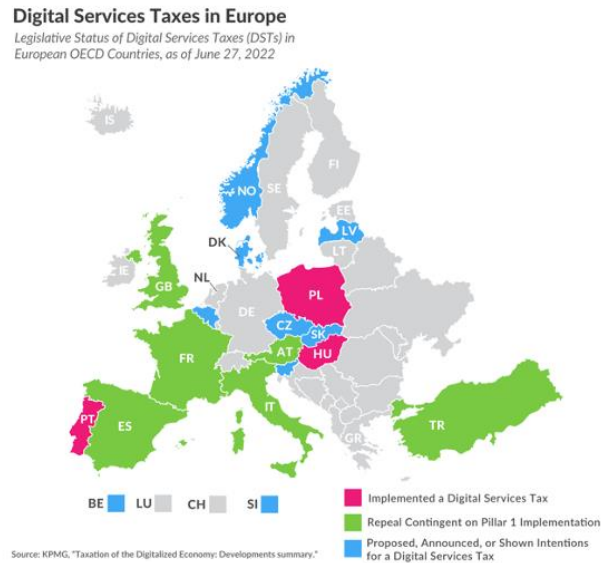


图表13: 我国土地财政收入增速下降



来源: Wind, 国金证券研究所

图表14: 欧洲地区率先实行数据服务税



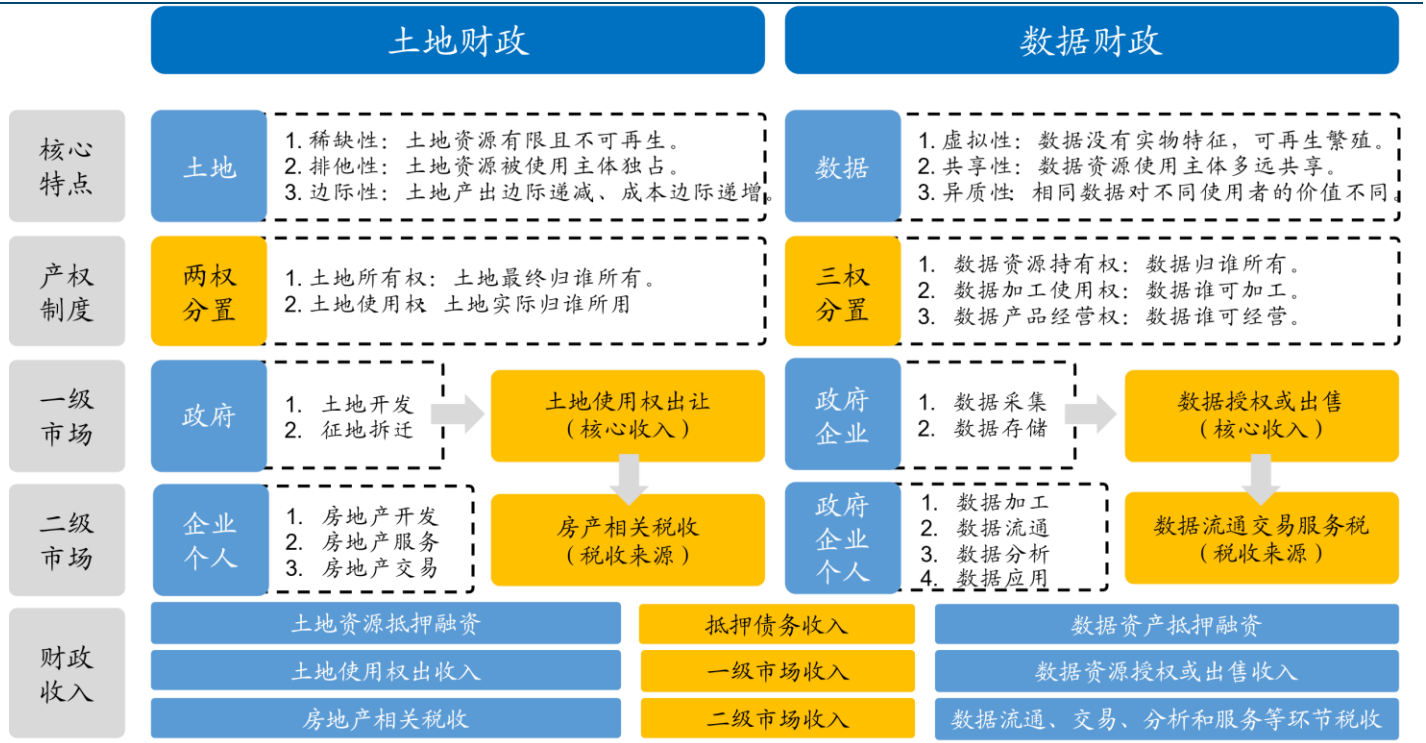
来源: Tax foundation, 国金证券研究所

“数据财政”制度将在数字经济发展和国家治理中扮演重要角色。数据资源具有公共性，数据权利分置创新拓宽了数据流通空间，建立数据财政制度是完善数据基础制度的需求，体现数字经济公共利益和全民共享数字经济红利；同时，数据财政发挥现代国家治理中财政作用，推动数据要素成为新生产要素，创造良好环境实现要素市场化配置，注入新时代寓意。数据财政主要通过财政支出和收入发挥治理数据要素市场的功能，在数据市场化起步阶段应当以财政激励和适当轻税推动数据供需市场发展，并力争形成良性循环；同时确立数据财政方式和手段，包括归并公共数据、提供公益性服务和有偿服务、采用财税政策和管理要求等；此外，建立数据财政制度和政策，涵盖公共数据资产管理、数据定价、数据税收、收入分配等也是试试数据财政的主要工作之一；最后，运营包括主体准入、运营级次选择、运营体系、税费征管等。

由单一强调“土地财政”转向“土地财政+数据财政”双轨并行。相较于土地财政，数据财政的核心标的从土地转变为数据，从而利用虚拟性、共享性和异质性等特点，排除了土地资源有限、主体独有不可共享和边际效应等问题。在财政收入方面，数据资源的资产性让其能够参与抵押活动，同时主要的收入来源变成了数据授权或出售过程中的收入，参考欧洲地区的数据服务税，还可在数据流通、交易和服务等环节征税。从“土地财政”到“土地财政+数据财政”双轨并行，这种转变反映了政府财政模式的演进和创新，不仅能有效避免单一依赖土地财政的风险，还能充分利用数据资源，增加财政收入的多样性，进一步推动社会经济的发展；同时数据财政也可以为政府提供更准确的数据支持，帮助政府更好地进行政策决策和提供公共服务。



图表15：“土地财政”和“数据财政”具备双规并行的可能性



来源：国金证券研究所

2. 垂类 AI 大模型加速落地，推动数据要素市场发展

2.1 AI 大模型向行业垂类模型落地演进，数据成为核心壁垒

2017 年谷歌发布的 Transformer 网络结构是大模型发展的源头技术，自此以后大模型技术在自然语言理解、计算机视觉、智能语音等方面都取得了标志性的技术突破，在模型精度、通用性和泛化能力等方面都实现了跨越式发展。中国自 2020 年进入大模型快速发展期，目前与美国保持同步增长态势，涌现出 GLM、盘古、悟道、文心一言、通义千问、星火认知等一批具有行业影响力的预训练大模型，形成了紧跟世界前沿的大模型技术群。

图表16：全球知名大模型发布时间节点



来源：CNCF, Frost&Sullivan, 国金证券研究所

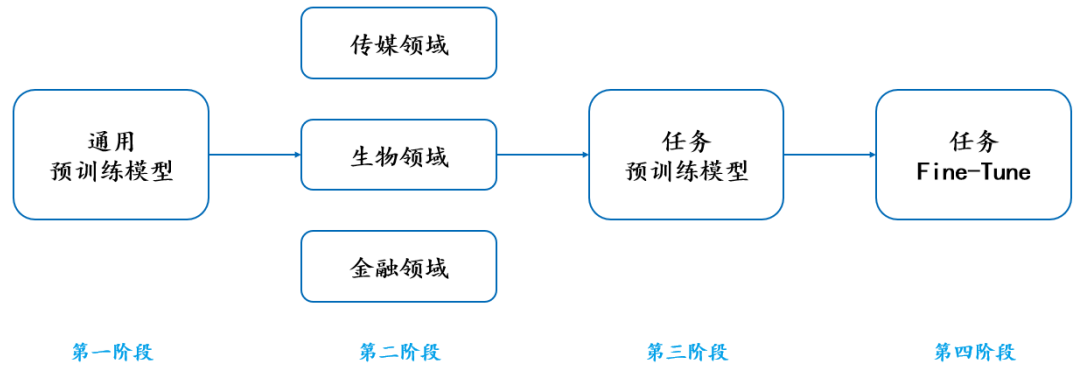
数据贯穿 AI 垂类模型训练的始终。AI 垂类模型强调领域的 know-how，对数据在深度和质量上的要求更高，模型训练分为四个阶段：

- 1) 通用预训练：在数据质量有保证的前提下，增加数据的数量和多样性，同时提升模型复杂度，这样可以提供普遍有效的模型增强能力。
- 2) 领域预训练：在第一阶段通用模型基础上，分别用各个领域数据，再分别做一次预训练，得到适合解决各个不同领域的预训练模型。



- 3) 任务预训练：选择任务适配的领域预训练模型，在这个模型的基础上，用手头数据，抛掉数据标签，再做一次预训练。
- 4) 任务 Fine-tuning。

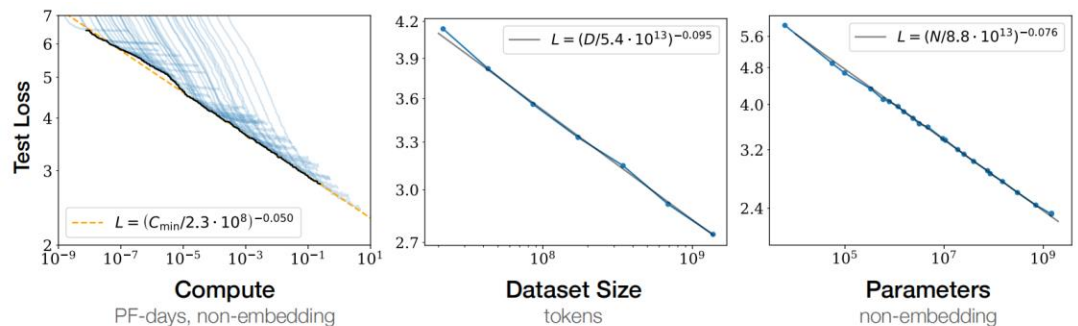
图表17: 数据贯穿 AI 垂类模型训练的始终



来源：国金证券研究所

增加训练数据量对模型性能提升来说更具性价比。根据 OpenAI 的研究，独立增加训练数据量、模型参数规模、训练计算量时，预训练模型在测试集上的损失会单调降低，模型的效果越好。DeepMind 在设计 Chinchilla 模型时，对标数据量 300B、模型参数量 280B 的 Gopher 模型，选择将参数降低为 Gopher 的四分之一的同时增加 4 倍的训练数据，无论是预训练指标，还是很多下游任务，Chinchilla 的效果都要优于规模更大的 Gopher。

图表18: 模型性能与训练数据量、模型参数规模、训练计算量呈正相关关系



来源：《Scaling Laws for Neural Language Models》，国金证券研究所

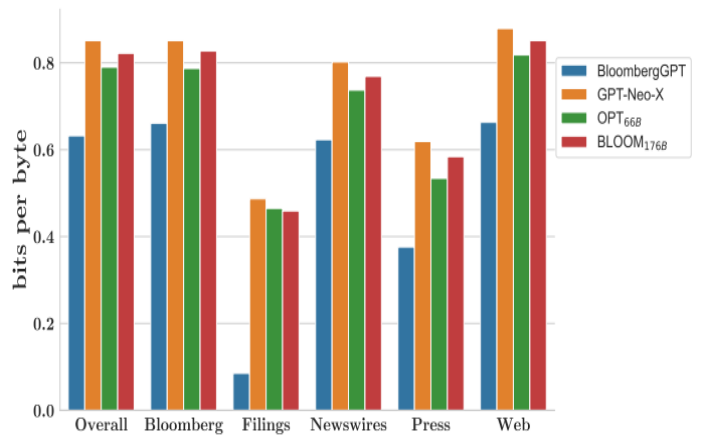
金融垂类数据对于 BloombergGPT 模型性能提升效果明显。Bloomberg 作为全球商业、金融信息和财经资讯的领先提供商，拥有 40 年金融数据的积累。BloombergGPT 是一个有 500 亿参数、基于 BLOOM 模型的 LLM，其训练所用的金融数据集包含新闻、档案、网络爬取的新闻稿件、英文财经文档等英文金融文档，共包含 3630 亿个 token，是目前最大的金融数据集。BloombergGPT 在金融语料上的 bits per byte 指标均好于其他垂类模型，在大多数任务中的得分位列第一，是目前最出色的金融垂类模型。



图表19: BloombergGPT 金融领域数据集共包含 3630 亿个 token

Date	Bloomberg	Filings	News	Press	Web	Total
2007 [03-]	276	73	892	523	2,667	4,431
2008	351	91	1,621	628	9,003	11,695
2009	293	93	1,791	528	9,179	11,883
2010	292	111	1,917	527	11,388	14,236
2011	335	117	2,264	548	13,643	16,907
2012	403	105	2,502	529	15,015	18,554
2013	415	87	2,437	441	17,230	20,610
2014	396	251	2,458	437	18,510	22,052
2015	358	1,639	2,371	427	20,782	25,576
2016	324	1,891	2,509	418	24,337	29,478
2017	294	2,294	2,567	398	25,283	30,837
2018	275	1,791	2,702	420	26,027	31,214
2019	263	1,662	3,102	504	27,195	32,726
2020	277	1,632	2,794	805	30,928	36,435
2021	247	1,767	3,515	938	29,749	36,215
2022 [-07]	140	882	2,206	531	16,872	20,631
	4,939	14,486	37,647	8,602	297,807	363,482

图表20: BloombergGPT 在金融语料上的 bits per byte 均好于其他模型



来源:《BloombergGPT: A Large Language Model for Finance》, 国金证券研究所

来源:《BloombergGPT: A Large Language Model for Finance》, 国金证券研究所

行业数据成为国内垂类模型卡点。尽管国内 AI 大模型已经取得较大进展,但是行业数据的可得性限制了垂类模型的应用落地。当前阶段,拥有大量数据积累并且具备行业 know-how 的企业,能够在在大模型的基础上学习行业特色数据与知识,打造出 AI 垂类模型。

拓尔思拥有千亿级优质数据,在垂类模型高质量训练过程中具备独特优势。公司拥有 10 余年高质量数据和知识资产积累,拥有超过 1500 亿条高价值的数据和知识资产,包含文字、图片、音视频等多种模态数据。在数据时效性上,公司重点数据源分钟级更新数据,日均采集增长过亿条。在数据质量上,公司产业要素数据集与产业风险数据集入选首批“北京市人工智能大模型高质量数据集”,重点数据主要采集主流官方机构可公开访问的数据,保证数据源头的“纯净”和完整性,同时配备了专家级的知识标引团队,对行业知识进行梳理,积累了 30+领域知识库、31000+知识标引规则,保证数据标签的多维度和精准度,更适配高信源数据在各行业领域的应用场景。

图表21: 拓尔思三大数据资产平台拥有千亿级优质数据积累

平台名称	数据类型	数据规模	标注	客户群体
数家 (媒体资讯)	长短文本	1000 亿+	200+以上标引维度 12000+标引规则	传媒机构
网察 (舆情)	多语言文本	1500 亿+	100+以上标引维度	政府、企业、高校、研究机构等
	音频		9000+标引规则	
数星 (产业大脑)	图片	110 亿+	500+以上标引维度	银行、保险、企业、证券、政府等
	视频		10000+标引规则	

来源:拓尔思微信公众号, 国金证券研究所

拓尔思凭借丰富的数据资源等优势,结合行业 know-how,面向媒体、金融、政务领域推出三大行业大模型,实现行业应用落地。

拓天 M 媒体行业大模型:凭借广泛的用户基础、丰富的数据资源、适用场景的理解三大行业基础,推出内容生产智能助手(智能拟定新闻标题,文章段落续写等),新一代搜索与推荐(领导人的金句引用),以及多模态传播与服务(虚拟 AI 主播、传播分析报告自动生成)三大应用。

拓天 F 金融行业大模型:基于用户、数据、知识、场景四大行业基础,推出投研类(生成金融图谱、标的推荐、风险分析、延保生成、数据获取),智能客服(通过语音识别,摘要投诉内容、投诉核实、生成解决方案),内容审查(对宣传图进行禁用词、敏感词的审核)三大应用。

政府 G 政务行业大模型:公司深度理解政务场景,并在数字政府领域积累了大量优质头部用户和政务资讯的高质量数据,选择了自 2020 年以来的近一亿条数据进行预训练,在政



策服务、新一代政务咨询、公文辅助写作三个场景实现应用落地。

图表22: 拓天·M 媒体大模型的行业基础

图表23: 拓天·F 金融大模型的行业基础



来源: 拓天大模型发布会, 国金证券研究所

来源: 拓天大模型发布会, 国金证券研究所

医学数据积累和行业 know-how 使医联在垂类模型落地中占据先发优势。医联 MedGPT 为国内首款医疗大语言模型，目前的参数规模为 100B 规模，预训练阶段使用了超过 20 亿的医学文本数据，微调训练阶段使用了 800 万条的高质量结构化临床诊疗数据，并投入超过 100 名医生参与人工反馈监督微调训练，目前医联 MedGPT 已经可以覆盖 ICD10 的 60% 疾病病种。在 6 月 30 日国内首次 AI 医生与真人医生一致性评测中，7 位专家教授针对评测形成的 91 份有效病例进行审核，并且针对 AI 医生的问诊准确性、诊断准确性、治疗建议准确性、辅助检查方案准确性、数据分析准确性、提供可解释信息、自然语言问诊与交互这 7 个评价维度的打分之后，最终，真人医生综合得分为 7.5 分，AI 医生综合得分为 7.2 分，AI 医生与三甲主治医生在比分结果上的一致性达到了 96%。

图表24: 医疗 MedGPT 参数和训练数据规模庞大

图表25: MedGPT 将不断提升病种覆盖率



来源: 医联 Medlinker 公众号, 国金证券研究所

来源: 医联 Medlinker 公众号, 国金证券研究所



图表26: MedGPT 与三甲主治医师医学一致性达到 96%



来源: 医联 Medlinker 公众号, 国金证券研究所

华为庞大的数据积累和多年行业解决方案经验助力行业大模型加速落地。在生态层面, 华为云数智领域已经拥有超过 80 万的开发者, 8700+社区代码贡献者, 覆盖业内主流算法模型及数据集的 2000+行业资产, 以及覆盖金融、制造、互联网、医疗等多个行业的企业用户。基于盘古基础大模型和海量行业数据, 华为云陆续推出了矿山、药物分子、电力、气象、海浪等盘古行业大模型。例如, 华为盘古药物分子大模型接受了超大规模的化合物表征模型训练, 预先对 17 亿个药物分子的化学结构进行了学习, 并对药物分子结构进行预测、打分, 成药性预测准确率比传统方式高 20%, 帮助科研人员省去大量药物设计的成本; 此外, 该模型还内置了高效的分子生成器生成了 1 亿个创新的类药物小分子筛选库, 结构新颖性达 99.68%, 为发现新药创造了更多可能性。

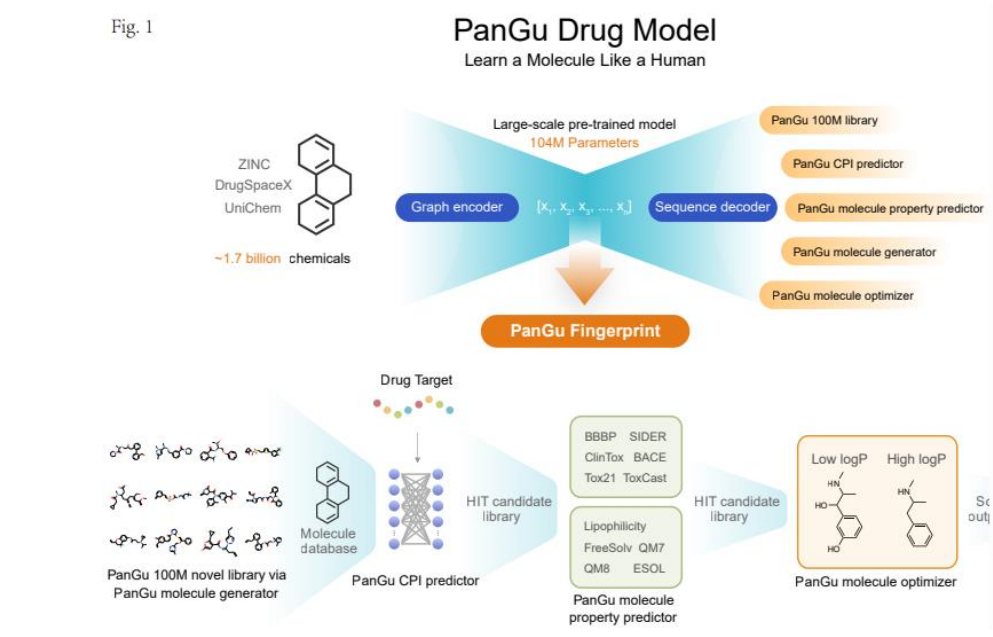
图表27: 华为盘古大模型 3.0 重塑千行百业



来源: 华为微信公众号, 国金证券研究所



图表28: 盘古药物分子大模型



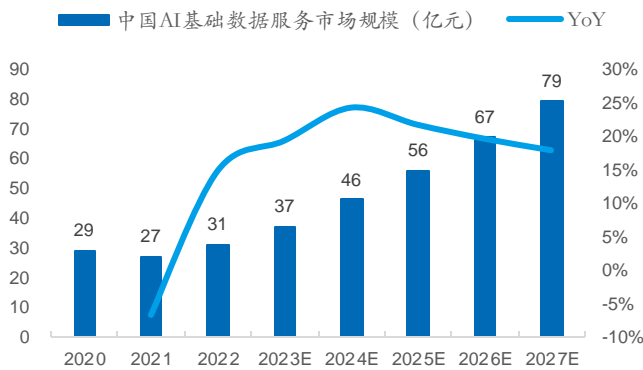
来源:《PanGu Drug Model: Learn a Molecule Like a Human》, 国金证券研究所

2.2 AI 垂类模型加速数据要素市场发展, 推动数据流动

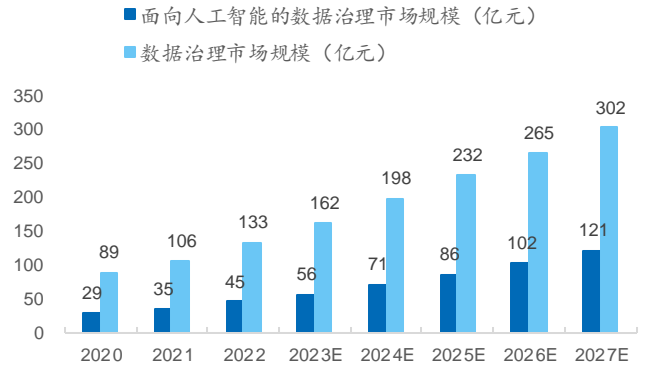
数据是 AI 模型的关键要素之一, AI 数据相关市场快速增长。据艾瑞统计与预测, 2022 年我国 AI 基础数据服务市场规模(含数据采集与标注)为 31 亿元, 受各类 AI 应用对于图像、语音和文本数据集及定制化数据服务的需求上涨影响, 2027 年相应规模可达到 79 亿元, 2022-2027 年 CAGR 达 20.6%。2022 年中国面向人工智能的数据治理市场规模约为 45 亿元, 受数据平台服务、数据治理服务和 AI 应用建设的需求推动影响, 在数据治理市场中的占比不断提升, 2027 年将达到 121 亿元, 2022-2027 年 CAGR 为 21.7%。

图表29: 2020-2027 年中国 AI 基础数据服务市场规模

图表30: 2020-2027 年中国数据治理与面向 AI 的数据治理市场规模



来源: 艾瑞咨询, 国金证券研究所



来源: 艾瑞咨询, 国金证券研究所

垂类模型的训练需要海量优质且合法的数据资源, 其中包含两方面的维度, 第一是海量的数据资源, 第二是优质且合法的数据资源。

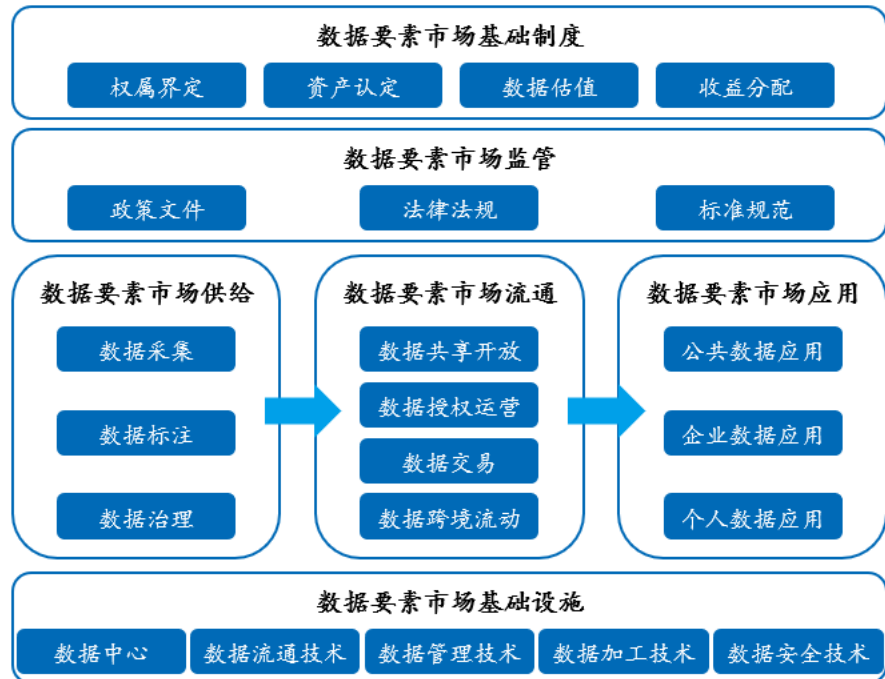
在垂类模型的训练过程中, 在保证数据质量的前提下, 训练数据量越大, 模型的推理能力就越强; 训练数据集的丰富度和全面性越高, 模型泛化能力越强; 数据更新的及时性越高, 模型的实用性就越强。单一企业所拥有的数据难以同时满足质量、规模、多样性、实时性方面的要求, 使得模型性能提升存在瓶颈, 需要通过数据流通补充企业所缺少的数据资源。当前阶段, 模型厂商主要通过共建生态或者数据交易两种方式获取垂类数据, 百度、科大讯飞等厂商与各行业企业建立起生态合作关系, 在行业数据、模型算法上实现优势互补; Reddit 4 月份宣布将对调用其 API 的公司收费, 中文在线按照采集数据包的大小及数据类别对模型厂商进行收费。



模型训练所使用的数据有一部分是来自互联网的公开数据,这些数据可能包含用户的个人信息,这些信息可能被大模型无意中学习和记忆,并在后续的应用中泄露。例如,OpenAI 抓取了社交媒体网站 Reddit 上所有获得至少 3 个“喜欢”的帖子所链接的所有网页,以及 Reddit 帖子和相关评论,而这些链接及与之关联的个人信息在没有通知相关方获得同意的情况下被大量抓取,存在极大的安全隐患。

AI 垂类模型方兴未艾,推动数据要素市场发展。数据要素市场打通“数据要素供给-数据要素流通-垂直领域模型研发及应用”循环发展的产业链条,对垂类模型的应用落地具备重要意义,数据要素市场供给提供海量丰富的训练数据,是垂类模型的基础原料,市场流通是连接数据要素与垂类模型的枢纽,基础制度保障市场规范发展,市场监管维护市场运行秩序,基础设施提供重要支柱。

图表31: 数据要素市场体系架构



来源: 中国电子信息产业发展研究院, 国金证券研究所

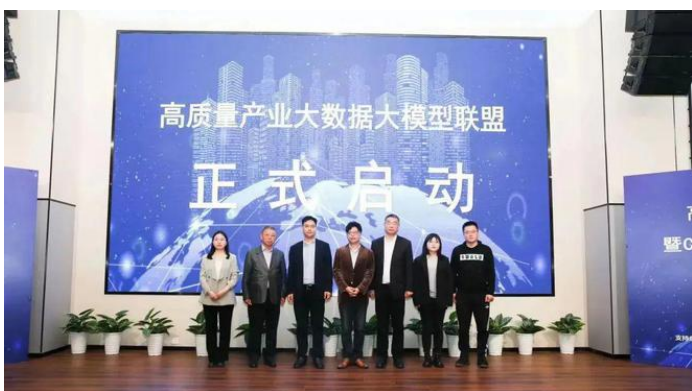
在垂类模型的推动下,数据要素市场的新业态不断涌现。多地以数据交易所为主导,建立数据要素与大模型的“产业联盟”,数据要素市场建设再加速。

高质量产业大数据大模型联盟聚焦高质量产业数据垂直应用。2023年5月6日,高质量产业大数据大模型联盟正式启动,联盟创始会员单位包括华东江苏大数据交易中心股份有限公司、西部数据交易有限公司、苏州大数据交易服务有限公司、德阳数据交易有限公司等10家单位。该联盟重点研究数据质量评估、数据价值评估、数据资产入表、数据资产证券化和衍生产品等方向,旨在打造高质量产业大数据集,赋能垂类模型应用落地。

南海区大数据人工智能产业联盟开辟新赛道,丰富数据要素市场生态。2023年6月13日,广州数据交易所为发起单位,成立佛山市南海区大数据人工智能产业联盟,“一所多基地多平台”体系架构的正式落地。南海“垂类模型+传统产业”可将难以直接交易的产业数据转化成为数据模型产品,借助数据交易平台,充分激活大量沉淀闲置的数据资源,进一步丰富数据要素市场生态,开辟出“垂类模型+传统产业+数据交易”的新赛道。



图表32: 高质量产业大数据大模型联盟成立



来源: 云旗数据交易平台, 国金证券研究所

图表33: 南海区大数据人工智能产业联盟成立



来源: 南海区人民政府官网, 国金证券研究所

上海数交所启动语料数据生态创新合作伙伴计划, 推动数据要素市场建设。上海数交所官网于7月7日正式上线语料库, 累计挂牌近30个语料数据产品, 包含文本、音频、图像等多模态, 覆盖金融、交通运输和医疗等领域。针对数据质量高但开放程度低的供方, 可以通过数据交易链有效破解语料数据流通的信任问题, 此外, 上海数据交易所语料库还将提供特色标签服务体系、挖掘应用场景价值、驱动稀缺数据开放流通以提高该类供方开放数据积极性。针对开放程度高但数据质量低的机构, 上海数交所语料库通过搭建专业化数商服务渠道提高数据质量。

图表34: 上海数据交易所上线语料库, 助力垂类模型训练



来源: 上海数据交易所, 国金证券研究所

3. 投资建议

当前, 我国数据要素市场正面临从零到一的重大拐点, 我们预判数据要素市场是中国未来十年极具成长潜力的赛道, 板块将持续受益政策加持和产业发展逐步验证的双重催化。

政策层面持续加持。从2022年12月20号国家印发《关于构建数据基础制度更好发挥数据要素作用的意见》强调, 数据基础制度建设事关国家发展和安全大局, 要统筹推进数据产权、流通交易、收益分配、安全治理, 加快构建数据基础制度体系。

北京2023年6月印发《关于更好发挥数据要素作用进一步加快发展数字经济的实施意见》的通知, 力争到2030年, 北京市数据要素市场规模达到2000亿元, 基本完成国家数据基础制度先行先试工作, 形成数据服务产业集聚区。

深圳2023年6月印发《深圳市数据产权登记管理暂行办法》, 为规范数据产权登记行为, 保护数据要素市场参与主体的合法权益, 促进数据作为生产要素开放流动和开发利用, 提供了深圳经验和模式。

上海2023年6月印发《立足数字经济新赛道推动数据要素产业创新发展行动方案



《(2023-2025)》指出要主动跟踪数字经济发展趋势，推动数据要素产业高质量发展。要提“效能”，全力打造产业发展新高地，在产品供给、场景应用、数商发展等方面持续发力。

管理职能层面持续完善。今年两会期间，国家公布组建国家数据局，作为国家发展和改革委员会管理的国家局，同时各地方交易所逐步建立和从试运逐步切换到正式运行阶段。

我们预判数据要素市场下半年有望进入实质性加速发展的重要阶段，从“土地财政”转向“数据财政+土地财政”双轨运行，有望带来产业链巨大投资机会。

AI 模型与数据要素将形成中长期持续共振。我国以大模型和行业垂类模型为代表的人工智能技术快速发展，数据将成为行业垂类模型发展的核心壁垒，AI 与数据要素市场形成共振，共同将成为新一轮大国科技竞赛的焦点，我国数据要素市场的价值有望充分挖掘。

我们建议数据要素市场投资机会核心关注重点环节的“卡位”逻辑，从数据要素的供给侧、运营端、安全防护三个维度进行分析，筛选出兼具基本面和商业模式初具雏形的优质企业进行提前布局。

1) 数据要素供给侧，是指行业数据商，具备细分行业的数据积累，未来在数据要素市场具备“数据金矿”的企业。典型的诸如各行业应用龙头企业，拥有行业垂类数据（中科江南、航天宏图、浩瀚深度、中科星图、超图软件、久远银海、中远海控、金山办公）等。

2) 数据要素运营端，是指第三方专业服务机构，能够为数据交易双方提供数据产品开发、发布、承销和数据资产的合规化、标准化、增值化服务，促进提高数据交易效率。典型的诸如，数据基础设施提供商包括：云基础设施（中国电子云、电信天翼云、移动云、华为云等）、数据安全存储（易华录等）、数据通信厂商（锐捷网络等）、数据库厂商（星环科技、太极股份等）。国家级数据运营商：中国电科（易华录）、中国电子（深桑达 A）、中国移动、中国电信等。地方国资系数据运营商：云赛智联、山大地纬等。

3) 数据全生命周期安全防护厂商：安恒信息、永信至诚、奇安信、启明星辰、三未信安、英方软件、信安世纪、格尔软件、电科网安等。

图表35：推荐标的

定位	公司名称	股票代码	推荐理由
数据 供给侧	中科江南	301153.SZ	中科江南是国内领先的智慧财政和智慧医保综合解决方案供应商，积极开拓财政、会计数电票入账和医保等领域的数据要素相关业务。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 1.69、2.22、2.79 元，对应 44X、34X、27X PE。
	航天宏图	688066.SH	航天宏图是卫星运营与应用服务龙头厂商，实现空天信息领域全产业链布局，加快发展以航天/航空为主的自主数据源建设。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 1.57、2.28、3.17 元，对应 39X、27X、19X PE。
	浩瀚深度	688292.SH	浩瀚深度是互联网流量管控及数据智能化的领航员，采用“以采集管理系统为基础，全方位拓展延伸”的业务发展模式，深耕运营商市场且具有存量设备部署优势。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 0.45、0.69、1.09 元，对应 64X、42X、26X PE。
	金山办公	688111.SH	金山办公是国内领先的办公软件和服务提供商，产品体系可实现“端云一体”、“多屏联动”，“云+协同”战略高度契合办公软件发展趋势。预计 23、24、25 年摊薄 EPS 分别为 3.53、4.56、6.06 元，对应 134X、104X、78X PE。
	中科星图	688568.SH	中科星图是国内数字地球领域领先厂商，形成了面向政府、企业及特种领域用户，覆盖空天信息产业全链条的 GEOVIS 数字地球应用生态。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 1.42、1.99、2.77 元，对应 57X、40X、29X PE。
	超图软件	300036.SZ	超图软件是地理信息软件（GIS）行业的领跑者，提供覆盖数据全生命周期的空间数据治理工具，积极开拓大数据局和大数据中心业务。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 0.61、0.82、1.07 元，对应 39X、29X、22X PE。
	久远银海	002777.SZ	久远银海是智慧民生服务商，聚焦医疗医保、数字政务、智慧城市三大战略方向，正加大数据要素研发投入，创新研究政务数据要素产品和创新应用场景。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 0.62、0.81、0.99 元，对应 57X、44X、36X PE。
	中远海控	601919.SH	中远海控是全球领先的集装箱航运综合服务商，聚焦集装箱航运及码头运营管理两大



			板块，且具备显著的规模优势。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 1.59、1.24、1.25 元，对应 5.9X、7.6X、7.5X PE。
基础设施 提供商	锐捷网络	301165.SZ	锐捷网络是行业领先 ICT 基础设施及行业解决方案提供商，交换机、无线产品、云桌面、IT 运维管理等多个领域市占率位居前列。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 1.33、1.79、2.2 元，对应 47X、34X、28X PE。
	星环科技	688031.SH	星环科技是国内领先的大数据基础软件开发商，实现了从大数据管理平台、数据应用中件、数据智能分析工具到数据库的全产业链布局。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为-1.48、-0.79、0.36 元。
	太极股份	002368.SZ	太极股份是国内信创和数字政府建设领域领军者，子公司人大金仓深耕数据库领域，具备国内领先的数据库产品、服务及解决方案体系。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 0.82、1.06、1.34 元，对应 50X、39X、31X PE。
基础设施 提供商、数 据运营商	深桑达 A	000032.SZ	深桑达是中国电子旗下云计算和数据创新业务的核心参与厂商，子公司中国电子云是中国电子旗下唯一云计算品牌，同时前瞻性开展数据要素领域研究和业务布局。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 0.25、0.44、0.75 元，对应 131X、74X、44X PE。
	易华录	300212.SZ	易华录是数据要素市场开拓者，坚持以数据湖战略为中心，围绕超级存储和数据变现两大主线，打造数字经济基础设施建设、数据运营以及政企数字化建设三大业务板块。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 0.35、0.65、1.00 元，对应 96X、51X、33X PE。
	中国电信	601728.SH	天翼云是全球最大的运营商云和国内最大的混合云，数据要素服务平台“灵泽”可提供“聚、治、享、安、算、用”一体化数据要素可信流通能力。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 0.34、0.38、0.42 元，对应 17X、15X、13X PE。
	中国移动	600941.SH	移动云位居国内云计算市场第一梯队，立足海量数据和丰富应用场景优势，从数据能力和算力网络两方面积极推进数据要素流通基础设施建设。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 6.4、7、7.63 元，对应 15X、13X、12X PE。
数据安全 厂商	安恒信息	688023.SH	安恒信息是新兴安全领跑者，将数据安全作为三大核心战略，布局数据安全咨询规划、数据安全保护、数据安全监管、数据要素流通四大方向。预计 23、24、25 年摊薄 EPS 分别为 0.93、2.49、4.6 元，对应 186X、69X、38X PE。
	永信至诚	688244.SH	永信至诚是网络靶场和人才建设领军企业，首创“数字风洞”产品体系开启数据安全测试评估专业赛道。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 1.22、1.82、2.49 元，对应 63X、43X、31X PE。
	奇安信	688561.SH	奇安信是国内网安行业龙头，数据安全、态势感知等新赛道产品持续领跑，提供满足数据全生命周期和数据全流程要求的数据安全产品和解决方案。预计 23、24、25 年摊薄 EPS 分别为 0.29、0.80、1.27 元，对应 179X、65X、41X PE。
	启明星辰	002439.SZ	启明星辰是网安行业老牌厂商，依托数据绿洲构建一体化数据安全保障体系，重点发力行业业务数据治理、行业业务数据要素化经营。预计 23、24、25 年摊薄 EPS 分别为 1.01、1.27、1.51 元，对应 29X、23X、20X PE。
	三未信安	688489.SH	三未信安是国内领先的商用密码基础设施提供商，“算法”+“芯片”两大硬核技术是三未信安在数据安全与密码领域的核心竞争优势。预计 23、24、25 年摊薄 EPS 分别为 1.32、1.61、2.01 元，对应 48X、39X、32X PE。
	吉大正元	003029.SZ	吉大正元是国内电子认证及可信数据保障行业领军者，提供基于密码的可信身份认证及可信数据保障等多层次、全方位的综合安全性安全解决方案。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 1.33、1.74、2.41 元，对应 24X、19X、13X PE。
	格尔软件	603232.SH	格尔软件是国内公钥基础设施 PKI 领域龙头厂商，提供覆盖资产管理、安全防护、安全存储和安全运营四大方面的数据安全解决方案。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 0.38、0.55、0.69 元，对应 39X、26X、21X PE。



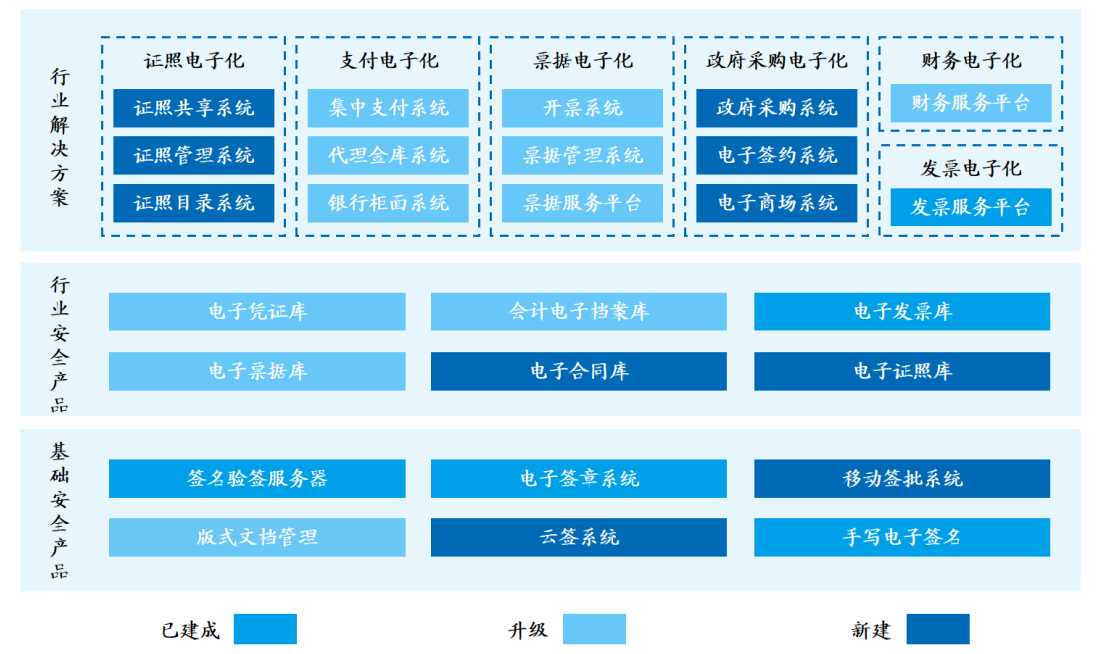
电科网安	002268.SZ	电科网安是密码行业龙头、数据安全国家队，率先提出以密码为核心的数据智能安全服务商的战略定位，积极探索和试点数据流通和价值挖掘领域的数据安全合规运营市场。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 0.48、0.62、0.79 元，对应 59X、45X、35X PE。
英方软件	688435.SH	英方软件是国内数据复制与保护软件领先企业，容灾、备份、云灾备产品系列构筑企业数据安全的最后一道防线，大数据系列产品向数据采集领域延伸。Wind 一致预测 23、24、25 年的摊薄 EPS 分别为 0.93、1.62、2.35 元，对应 79X、45X、31X PE。

来源：Wind，国金证券研究所（注：金山办公、安恒信息、奇安信、启明星辰、三未信安盈利预测及估值为国金证券研究所预测，其他来自 Wind 一致预期）

3.1 中科江南：财政、医保信息化领导者

财政信息化龙头，向多行业延伸发展。中科江南深耕财政信息化领域十余年，凭借支付电子化和预算管理一体化两大核心业务占据行业龙头地位，近年来中科江南加快推进行业电子化和预算云服务业务发展，依托于多年积累的电子化技术和实施经验不断拓展行业电子化应用，通过打造电子政务产品矩阵，为各级财政部门、金融机构和行政事业单位提供信息化管理系统解决方案。

图表36：中科江南行业电子化应用平台



来源：中科江南招股说明书，国金证券研究所

电子凭证会计数据标准深化试点，助力财税数据要素流通。2023 年 4 月，财政部等 9 部委联合发布《关于联合开展电子凭证会计数据标准深化试点工作的通知》，扩大试点电子凭证的种类和开具范围以及接收端试点单位范围，推动电子凭证全流程标准化无纸化处理，有望进一步促进业、票、财、税、档等会计数据标准化和规范化管理，从而实现财税数据要素的高效流转和互联互通。中科江南是 8 家“电子凭证开具分发平台”之一和深化试点服务保障单位，目前已推出国库集中支付电子凭证、基于区块链平台的医疗结算电子凭证、商标电子印章以及会计电子证照等电子凭证产品，助力电子凭证会计数据标准化应用。

医保数据要素落地在即，医保电子凭证业务增长可期。随着医保信息化建设持续推进，医保、商保数据有望互通并带来大量应用场景，未来医保数据要素实现或将率先开放。中科江南电子凭证业务在医疗信息化领域不断落地，2022 年其先后开拓了国家医保局医疗电子票据应用区块链平台及电子票据共享项目、湖北省银行端医保局医保基金业财一体化项目、湖南医疗保险基金财务一体化管理项目等项目，为医保电子凭证业务打开增量空间。

积极开拓数据要素业务，相关数据应用服务正在广州数交所进行合规审查。作为国内领先的智慧财政和智慧医保综合解决方案供应商，中科江南在财政及医疗垂类领域拥有丰富的项目经验和数据积累，有望进一步围绕财税、医保等数据要素与相关行业应用场景构建数据应用服务，不断开拓财政、会计数电票入账和医保等领域的数据要素相关业务。目前，中科江南已成为广州数据交易所登记入库的数据经纪人和第三方专业服务机构，数据要素

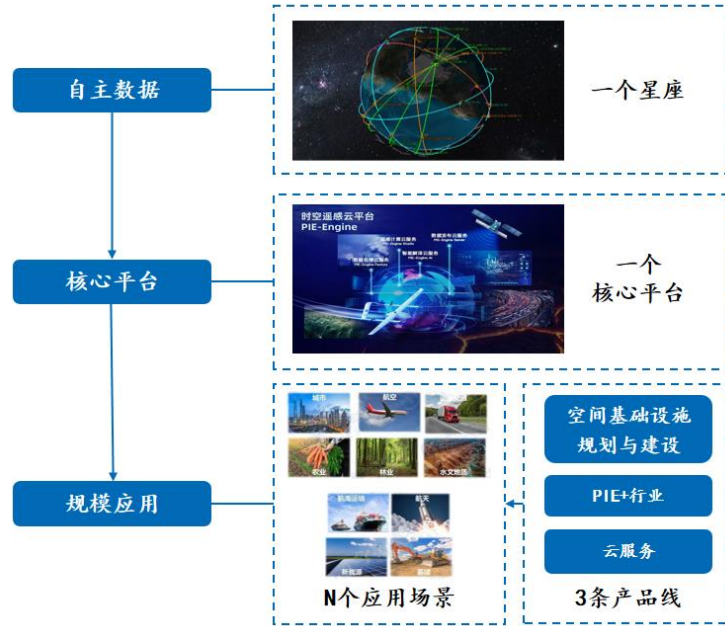


相关产品已经入库并登记，正在进行政策合规审查。

3.2 航天宏图：空天地一体化数据领域龙头

完善全产业链布局，构建空天地一体化数据资源体系。航天宏图是空天信息领域先行者、卫星运营与应用服务龙头厂商，为突破自主可控数据源瓶颈和推进云服务规模化，业务模式从空天信息产业链中下游向上游拓展，稳步推进分布式干涉雷达卫星星座和无人机产线建设，形成“上游自主数据-中游核心平台-下游规模应用”商业模式，实现“一个星座、一个核心平台、三条业务产品线、N个应用场景”的全产业链布局。其中，宏图一号卫星星座已于2023年3月30日完成发射并进入预定轨道，有效填补国内民用SAR遥感数据市场的空白。

图表37：航天宏图扎实推进全产业链布局



来源：航天宏图 2022 年年报，国金证券研究所

云化转型加速，遥感数据不断丰富。航天宏图 PIE-Engine 平台已初步形成“平台+SaaS应用”、“软件+数据”的服务模式，2022 年新增了面向金融、农业、林业、能源和建筑工程、智慧园区的无人机服务能力，发布了面向全国地面沉降和大型基础设施形变监测的 SAR 雷达数据 SaaS 服务，形成了 30+ 全国重点区域沉降监测产品并完成了 200+ 用户订单，云服务能力的持续拓展应用使得遥感数据不断丰富，截至 2022 年底，PIE-Engine 平台已覆盖 20 多个领域，数据总量近 20PB，数据集个数达到 160 多种。

PIE+行业产品应用落地，深入挖掘数据要素价值。航天宏图 PIE+行业产品线向不同细分行业加速渗透，2022 年在持续拓展气象海洋、生态环境、自然资源等优势行业应用的基础上，重点提升了特种行业、应急管理等领域卫星应用服务热点市场的应用深度和广度。此外，航天宏图 PIE+行业业务逐步接入其自主的航天（卫星星座）航空（无人机群）数据，依托“数据+云+应用+生态”的服务模式，深入挖掘数据的应用价值，并形成行业级一体化服务平台体系与数据产品，未来将为需求侧持续赋能。

3.3 浩瀚深度：运营商数据可视化市场领军者

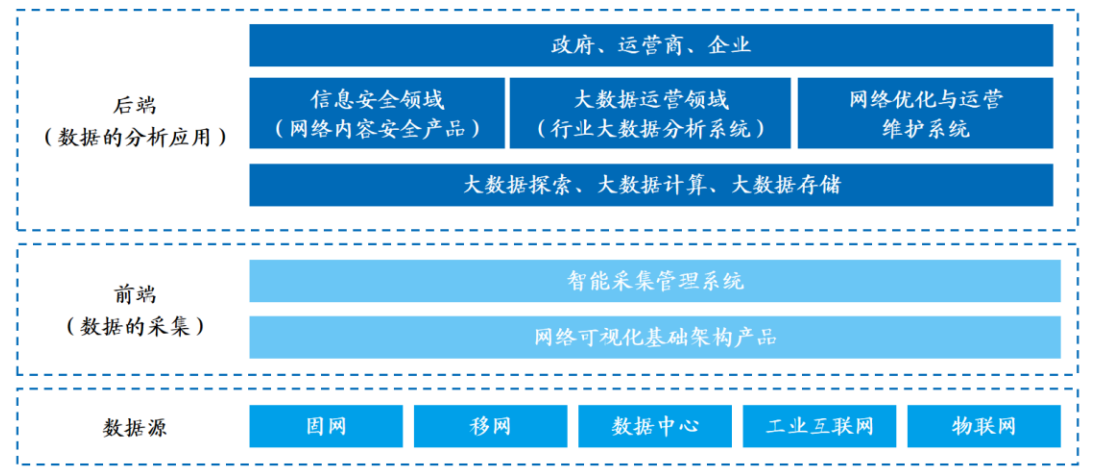
网络可视化领军，全方位拓展延伸。浩瀚深度专注于互联网流量管控及数据智能化应用领域，在海量数据获取、高速数据处理和深度信息挖掘方面拥有 20 多年的经验积累和技术沉淀，是国内通信行业中少数能够全面实现互联网流量及数据的前端采集处理、后端智能化应用的企业。浩瀚深度采用“以采集管理系统为基础，全方位拓展延伸”的业务发展模式，在持续迭代升级智能采集管理系统的同时，纵向拓展各类智能化应用系统，横向延伸开发信息安全防护类产品，各产品间形成良好的协同效应，实现网络可视化行业中全产业链垂直一体化发展，为电信运营商等客户提供网络智能化及信息安全防护解决方案。

聚焦 DPI 技术，赋能数据采集及数据分析应用环节。“智能采集管理系统”是浩瀚深度应用 DPI 技术实现产业化的核心产品，具备互联网流量识别、分析、管控以及数据采集功能，也是实现流量及数据智能化应用的重要基础。浩瀚深度 DPI 技术形成了“小而精”的发展路径，在处理性能、设备集成度、设备能耗比等方面更具优势。此外，浩瀚深度 DPI 系统



串联部署于骨干网、城域网/省网、IDC 出口、互联互通等关键网络节点，能够对海量流数据进行采集、分析与管控，在网络安全和数据安全防护领域同样大有可为。

图表38：浩瀚深度可全面实现互联网流量及数据的前端采集处理和后端智能化应用



来源：浩瀚深度招股说明书，国金证券研究所

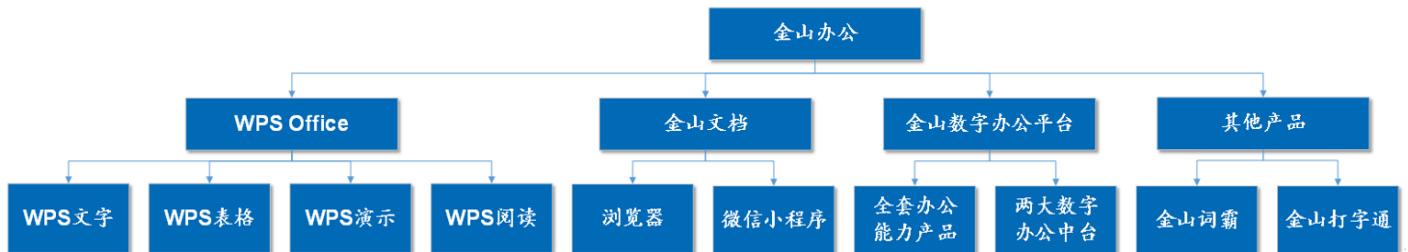
深耕运营商市场，与中国移动合作紧密。浩瀚深度近 30 年来深耕运营商市场，其硬件 DPI 技术路径与中国移动的网络建设及技术路径适配性较好，已成为中国移动硬件 DPI 系统的核心供应商，在中国移动总部 2018-2022 年互联网 DPI 集中采购招投标的中标总份额均超过 50%，除了与中国移动的保持紧密合作外，浩瀚深度还于 2017 年推出了满足中国电信、中国联通技术要求的软件 DPI 等相关产品，以进一步拓展运营商市场。目前，浩瀚深度在运营商市场已经形成了存量设备部署优势，相关产品广泛部署于国内三大运营商的各层级网络节点，在运营商国际出口和网间互联互通出口市占率第一。

拓展运营商数据应用场景，掘金数据要素价值。数据要素市场快速发展且算力需求高速增长，运营商同时作为数据要素市场多环节参与者和算力网络建设主导者，数据资源规模大且价值量高。浩瀚深度与中国移动等电信运营商客户合作多年，其核心 DPI 技术主要应用于数据采集和数据分析环节，可以用于深入挖掘运营商数据要素价值和拓展数据应用场景，例如帮助运营商用户和企业客户快速构建 PB 级海量数据处理平台，从而提升数据赋能的精细化管理水平、盘活数据资源。

3.4 金山办公：国产办公软件龙头

金山办公是国内领先的办公软件和服务提供商，主要从事 WPS Office 办公软件相关产品及服务的设计研发及销售推广。金山办公主要产品包括 WPS Office 办公软件、金山文档等办公能力产品矩阵以及金山数字办公平台解决方案。其中 WPS Office 办公软件及金山文档兼容 Windows、Linux、Mac OS、Android、iOS、HarmonyOS 等主流操作系统，用户可根据自身需求在 WPS 客户端和金山文档之间灵活切换，真正实现了“端云一体”、“多屏联动”。

图表39：金山办公产品结构



来源：金山办公 2022 年年报，国金证券研究所

把握办公软件“云+协同”发展趋势，云文档数量迅速增长。在“云+协同”办公逐步成为办公软件领域主流趋势的背景下，金山办公于 2015 年发布 WPS+一站式云办公并提供订阅服务，拉开云办公序幕；2018 年发布金山文档，开启协同办公征程。2022 年金山办公不断优化产品在云和协作等跨端场景下的用户体验，引导用户持续使用云文档服务，截至 2022 年底用户通过公有云上传的云文档数量已超 1752 亿份，同比增长 35%。“云+协同”



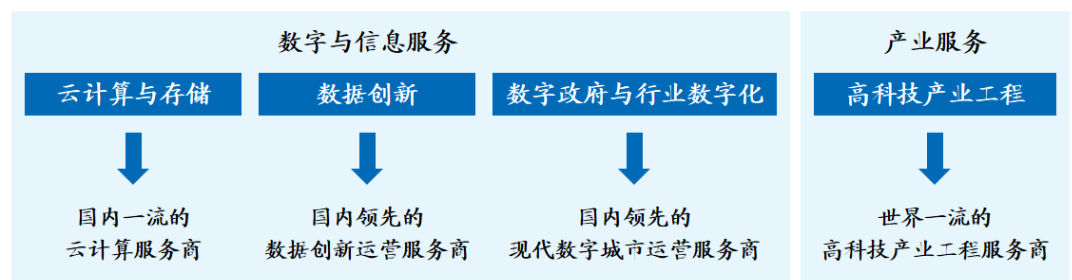
办公逐渐成为常态化需求，未来云文档数量及存储量有望延续高速增长态势。

端云一体化推进云服务功能统一，助力用户提升数据管控能力。2022 年公司完成了 WPS Office 的“端云一体化”转型，使客户端和云服务紧密结合，落地数字办公领域的云原生解决方案，实现了在企业身份和个人身份上实现用户数据的隔离和应用功能的统一，在企业管理和团队场景上实现文档、会议、日历的身份认证和数据打通，为组织级客户提供全面的文档数据管理服务；同时优化云文档服务基础架构，通过分布式高可用的技术方案，极大地提升了用户文档数据的安全性和可靠性。此外，金山办公还发布了全新的云文档管理器，优化云盘及搜索功能、改进分享和协作机制、提升用户全端文档上云的流程体验。

3.5 深桑达 A：CEC 旗下电子云和数据要素的核心运营者

深桑达是中国电子旗下云计算和数据创新业务的核心参与厂商，承担“云计算及存储+数据创新+数字政府与行业数字化服务+高科技产业工程服务”四项主责主业，以“云数”核心产品技术创新为驱动，持续赋能数字政府和行业数字化转型升级。

图表40：深桑达聚焦四项主责主业



来源：深桑达 2022 年年报，国金证券研究所

背靠 CEC 信创生态，国资云建设核心参与者。深桑达依托中国电子 PKS 自主安全计算体系及丰富的网信产业资源，持续打造中国电子旗下唯一云平台“中国电子云”，构筑了 IaaS、PaaS、SaaS、SECaaS 全栈云服务能力，为党政及关键行业客户提供高安全、高性能、高弹性的全栈分布式云及存储产品。依托对用户上云需求的深刻理解和自身 PKS 技术体系优势，以及在政企市场积累的大量实践经验，中国电子云在国内专属云市场后来居上，其全栈信创云应用于政府、金融、央企等数百家机构。随着信创产业规模持续扩大，中国电子云有望凭借 CEC 网信产业的资源优势和全栈自主技术能力，成为国资云建设的核心参与者。

数据创新业务排头兵，定位于数据要素运营者。深桑达是中国电子数据创新业务的承载主体，国内最早的数据要素化探索与实践者，在业内率先开展数据安全与数据要素化理论研究与业务布局。深桑达定位于“数据要素运营者”，已推出了数据金库、数据要素加工交易平台、安全可信数据空间等数据安全与数据要素化工程系列产品，并进一步形成数据产权、流通交易、收益分配以及安全治理方面的可实际操作方案。在实践领域，深桑达积极开展数据安全与要素化工程试点，已在武汉、郑州、德阳、大理等多个地市取得了阶段性成果。深桑达致力于成为国际领先的数据要素运营者，当前已凭借前瞻性布局和丰富的项目实践经验取得先发优势，未来有望持续受益于数据要素市场扩容带来的发展机遇。

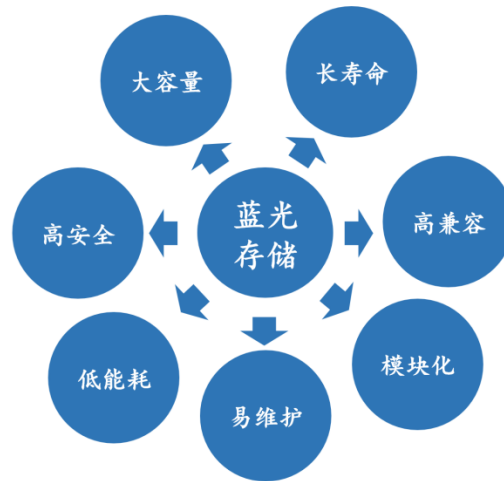
3.6 易华录：CETC 旗下数据安全存储和运营的先行者

大数据产业国家队，数据要素市场开拓者。易华录由中国华录集团控股，坚持以数据湖战略为中心，围绕超级存储和数据变现两大主线，不断深化自主可控的国产存储技术以及持续释放海量多元的数据要素生产力。2022 年以来以“数据二十条”为代表的顶层政策不断出台，易华录经营重点开始由数据湖建设转向数据湖及数据要素运营。易华录核心业务主要包括政企数字化业务、数字经济基础设施业务、数据运营及服务。

蓝光存储优势显著，数据湖助力数据要素价值释放。蓝光存储是目前最有效的数据长期归档存储方式，在数据能耗、介质寿命、数据安全性以及长期存储总体成本方面显著优于硬盘和磁带存储，可以更好地满足海量数据的大规模存储空间需求、降低长期保存和使用数据的能耗和成本，推动政企数据长期全量归档存储。易华录依托于自主可控且全球领先的 500GB 蓝光核心技术，2017 年独创数据湖商业模式，6 年来在全国 33 个城市落地运营数据湖，先发优势显著，具备完整的“收-存-治-用-易”数据全生命周期服务能力，为政府客户管理数据资产、实现数据要素优化配置提供可靠载体，截至 2022 年底已有多个数据湖已与政府签署了数据受托存储及运营协议、数据存储购买服务协议等数据存储服务协议。



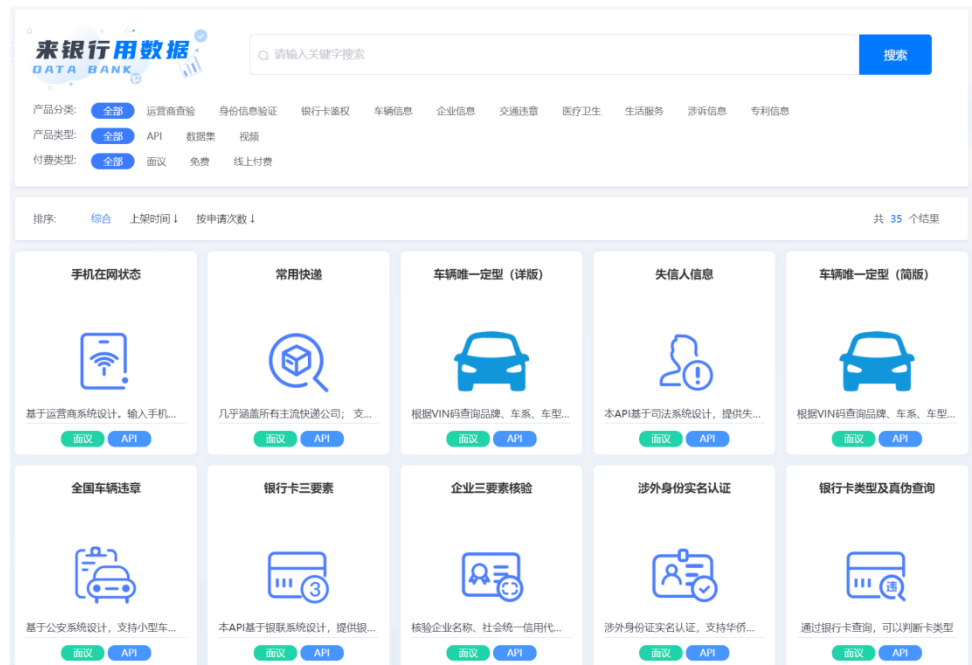
图表41：蓝光存储具备多重优势



来源：易华录官网，国金证券研究所

易数工厂驱动数据资产交易及变现，数据运营项目快速落地。易华录主要锚定数据登记确权、数据交易应用、数据授权运营、数据安全监管等方向进行纵深业务拓展，持续完善基于易数工场模式的数据资产化全链条服务。具体来说，就是对数据资源进行标准化场景运营与受托服务运营，将其转变为高价值的、具有市场需求的数据服务，最终实现数据资产交易与增值变现。经过6年持续打磨，2022年易华录在抚州落地全国首个基于政务数据授权运营模式的“数据银行”，且徐州、株洲、开封和抚州4个城市的数据湖项目公司均获得当地政务数据治理运营授权，同时与10+家省级大数据交易所签署合作协议。考虑到易华录数据湖已覆盖了全国22个城市，未来易数工厂有望向其他城市持续复制与落地。

图表42：易数工厂产品超市



来源：易华录官网，国金证券研究所

3.7 星环科技：国产分布式数据库和向量数据库龙头

专注于国产大数据基础软件领域，竞争力行业领先。星环科技是国内领先的大数据基础软件开发商，经过10年自主研发形成了涵盖大数据与云基础平台、分布式关系型数据库、数据开发与智能分析工具在内的完整产品矩阵，能够为金融、政府等行业客户提供覆盖数据全生命周期的整体解决方案，赋能客户数字化转型升级。2022年星环科技被Gartner评为数据中台及图数据库领域全球推荐供应商，同时入选中国数据库管理系统产品品类最



多的厂商之一，凸显星环科技在数据库和数据中台等大数据基础软件领域的核心竞争力。

图43：星环科技定位于数字化基础设施的底层和中间层



来源：星环科技招股说明书，国金证券研究所

分布式架构成为主流，加速数据库国产化替代进程。我国数据库市场长期以来由 Oracle、IBM DB2 为代表的国外传统集中式关系型数据库主导，然而，随着数据量的爆发式增长以及应用负载的快速增加，传统集中式架构难以满足海量数据处理和大规模并发需求，分布式架构则凭借安全性、扩展性、灵活性等优势成为主流的国产替代方向。数据库是信创产业链的核心环节，也是支撑数据要素价值释放的关键基础设施，随着数据管理软件逐步由集中式架构向分布式架构演进，国产大数据产品有望实现换道超车。

分布式数据库先行者，坚持自主研发道路。星环科技采用分布式技术体系，自主研发了高性能分析型数据库 ArgODB 和交易型数据库 KunDB 等分布式数据库产品，其中 ArgODB 主要用于构建离线数据仓库、实时数据仓库、数据集市等数据分析系统，KunDB 则适用于操作型业务场景和高并发场景。通过不断打磨产品技术以及在实际业务场景中的持续落地实践，ArgODB 和 KunDB 已广泛应用于金融、政务、能源、医疗、交通等多个行业，成功替代 Oracle、IBM DB2、Teradata 等传统关系型数据库，并且实现了低成本国产化替代和业务平滑迁移。2023 年，星环科技 ArgODB 和 KunDB 两款数据库均通过信通院“可信数据库”评测，体现了国产分布式数据库的优异性能。

图44：ArgODB 数据库和 KunDB 数据库通过信通院“可信数据库”评测



来源：星环科技微信公众号，国金证券研究所

积极布局数据要素市场，促进数据要素安全高效流通。星环科技实现了从大数据管理平台、数据应用中间件、数据智能分析工具到数据库的全产业链布局，能够针对数据要素的价值形成和流通过程中的不同环节提供相应的产品和解决方案。具体来看，星环科技大数据基础平台 TDH 和分析型数据库 ArgODB 可以进行数据归集及资源库建设；大数据开发工具 TDS 可以用于数据资产目录编制、数据标签运营、数据 API 开发和管理；智能分析工具 Sophon 可以实现海量数据的深度挖掘及安全管控。此外，星环科技基于 TDS 和 Sophon 形成了数据要素流通工具集解决方案，围绕数据交付等环节提供数据安全防护和隐私计算能力，助力数据要素安全流通和价值释放。



3.8 太极股份：CETC 旗下政务云基础设施承建方

数字服务国家队，自主可控主力军。太极股份是中国电科网信事业核心产业平台和自主可控总体单位，以信创、云和数据为两大业务主线，已成为国内信创和数字政府建设领域的领军厂商。太极股份全面布局信创基础软件产业，形成了涵盖普华操作系统、金仓数据库、金蝶中间件、慧点办公软件在内的自主产品体系，信创产业生态日益完善；另一方面，太极股份开创了“云+数+应用+生态”一体化服务模式，以“云+数”为基础驱动行业客户数字化转型，已在政务、农业产业互联网、工业互联网、交通大数据等数字化应用场景落地。

国产数据库先行者，党政行业市占率领先。太极股份控股子公司人大金仓是国内成立最早的国产数据库厂商，专注数据库领域 20 余载，具备国内领先的数据库产品、服务及解决方案体系，同时积极构建信创产业生态，截至 2022 年底已经与 1800+ 家生态伙伴的 6000+ 款产品实现兼容认证，通过与重点行业用户联合制定解决方案以实现向不同行业快速推广。目前金仓数据库已广泛应用于政务、军工、能源、金融、电信等 60+ 重点行业，累计装机部署超百万套，并于 2022 年完成了多家银行国际结算、跨境支付、供应链金融等 30 余类近 300 个系统的升级改造。根据太极股份 2022 年年报，2022 年人大金仓在电子政务、电信、能源等重点行业市占率保持领先；根据 IDC 数据，2H2022 人大金仓在我国关系型数据库本地部署模式市占率位居国内厂商第四。

图表45：人大金仓数据库积极构建信创产业生态



来源：金仓数据库微信公众号，国金证券研究所

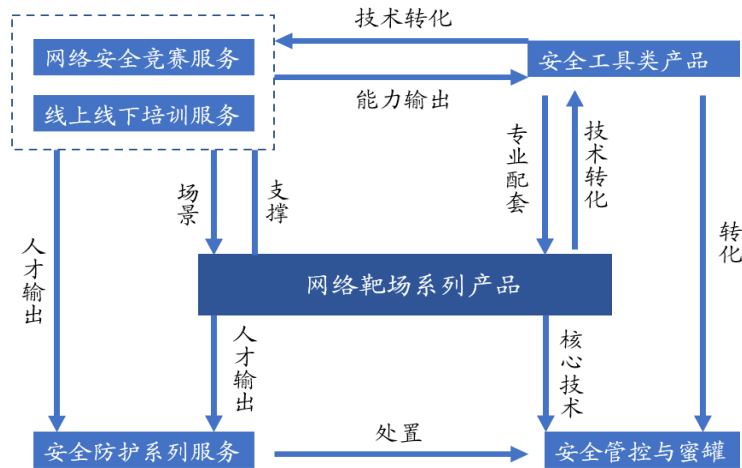
核心产品不断更迭，保障数据要素安全畅行。人大金仓主要产品包括通用型数据库 KingbaseES、分析型数据库 KingbaseAnalyticsDB、分布式数据库 KSOne、数据同步软件 KingbaseFlySync 等。其中，KingbaseES 作为人大金仓的核心拳头产品，近年来持续迭代升级，2022 年人大金仓发布 KingbaseES V9 全新版本，各项性能得到全面提升，并重点增强了 KingbaseRAC 能力，为国产数据库走向更多关键领域带来新的契机；同时发布金仓一体机 KingbaseXData，补齐了人大金仓在软硬件一体产品上的空白，进一步满足各行业对数据库产品不同维度的需求。人大金仓成立以来专注于数据管理领域，通过技术创新和产品迭代持续增强数据采集、存储、分析、管理等基础能力，支撑数据要素安全流通。

3.9 永信至诚：人工智能模型安全与数据安全实战化测评能力提供方

永信至诚是网络靶场和人才建设领军企业。永信至诚基于网络空间平行仿真及攻防对抗类技术，推出了网络靶场系列产品、安全管控与蜜罐产品、安全工具类产品、安全防护系列服务、网络安全竞赛服务和其他服务。其中，网络靶场系列产品是公司核心业务，其核心技术为蜜罐产品奠定了技术基础；网络安全竞赛服务和线上线下培训服务作为流量入口，积累行业化经验、储备安全人才以及提升公司影响力。



图表46：永信至诚产品服务体系生态链条



来源：永信至诚招股说明书，国金证券研究所

“数字风洞”产品体系开启安全测试评估专业赛道。在数据安全形势日益严峻的背景下，测试评估已成为风险管理的关键基础环节。永信至诚基于网络靶场深厚技术及经验积累，2022年发布数据安全“数字风洞”产品体系，形成了全场景、全要素、全生命周期的安全测试评估解决方案，通过全面支撑人、系统、数据等核心要素的安全测试与评估，有效保障用户网络和数据安全，目前在军工、公安、金融等关键行业以及人工智能、数据安全等新兴领域均得到实践验证。

秉承安全“证无”理念，实现全场景数据安全测试评估。“数字风洞”产品体系基于安全“证无”理念和3×3×3×(产品×服务)安全感公式，强调对人、系统、数据等进行持续性测试评估和风险处置，从而督促和帮助系统不断迭代优化，助力网络和数据安全由“形式合规”转向“实质合规”。此外，永信至诚参与协办2023数字中国创新大赛网络数据安全赛道，结合平行仿真技术、春秋云专有云平台和“数字风洞”产品体系，打造了覆盖数据安全、数据分析、数据算法等典型数据业务场景的测试评估演练场，在为网络和数据安全风险“证无”的过程中提升数据安全保障能力。

图表47：数字风洞打造数据安全测试评估标准平台



来源：永信至诚微信公众号，国金证券研究所



3.10 英方软件：大数据灾备领域龙头

英方软件是国内数据复制与保护软件领先企业。依托自主研发的动态文件字节级复制、数据库语义级复制和卷层块级复制三大核心底层复制技术及其他信息化技术，公司构造了“容灾+备份+云灾备+大数据”四大数据复制产品系列，相关产品覆盖了容灾、备份、云灾备、数据库同步、数据迁移等经典应用场景，同时也推广到了智能灾备管理、数据副本管理、数据流管理、大数据收集分发、数据跟随等更多应用领域。

核心技术达到行业先进水平，国内市场占有率高。英方软件的动态文件字节级复制技术可以保证各种复杂生产环境下的数据一致性，支持每秒 300MB 以上的高压力业务生产环境，理想状况下的同步性能达到毫秒级，可以持续数据保护及业务的高可用切换；数据库语义级复制技术每秒可处理 100MB 的在线日志，支持实时同步状态下的数据对比和数据修复，保证数据一致性；卷层块级复制技术可通过调整块的大小适应不同业务场景，提升系统性能、优化资源使用效率。根据 IDC 统计数据，公司在 2021 年国内数据复制与保护的纯软件市场排名第三，市占率 10.2%，仅次于 Veritas、Dell Technologies（市占率分别为 16.0%、13.1%），在国内第三方数据复制软件企业中排名第一。

图48：英方软件产品系列面向灾备和大数据两大应用场景



来源：英方软件招股说明书，国金证券研究所

灾备充当企业数据安全守门员，产品向数据采集领域延伸。备份数据往往作为企业数据安全的最后一道防线，公司的高可用灾备管理软件 i2Availability 可跨平台、跨系统、跨存储使用，远距离传输时对网络带宽要求低，广泛支持各类应用和数据库的容灾高可用保护；持续数据保护与恢复软件 i2CDP 可在任意时间点恢复，恢复颗粒度达到微秒级；备份和云灾备产品组合可实现字节级实时备份且任意时间点恢复，具有较强的兼容性，广泛支持各类应用程序和数据库的数据保护，支持国产操作系统。公司数据复制产品的应用领域逐渐向数据采集方向延伸，基于大数据产品 i2Active 和 i2Stream 可构建大数据实时汇聚平台，帮助客户在多中心、云平台等异构环境下，将各个业务中的应用数据进行低延时的采集、转换、融合、分发和监控等。

4. 风险提示

政策落地不及预期

数据要素相关政策密集出台以支持产业发展，若相关政策落地执行不及预期，可能影响行业商业化落地的进程。

数据资产化发展不及预期

目前关于数据资产化的商业模式处于发展期，还不够成熟，仍需要由产业链各个环节的厂商积极探索。



行业投资评级的说明：

- 买入：预期未来 3—6 个月内该行业上涨幅度超过大盘在 15%以上；
- 增持：预期未来 3—6 个月内该行业上涨幅度超过大盘在 5%—15%；
- 中性：预期未来 3—6 个月内该行业变动幅度相对大盘在 -5%—5%；
- 减持：预期未来 3—6 个月内该行业下跌幅度超过大盘在 5%以上。



特别声明：

国金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。

本报告版权归“国金证券股份有限公司”（以下简称“国金证券”）所有，未经事先书面授权，任何机构和个人均不得以任何方式对本报告的任何部分制作任何形式的复制、转发、转载、引用、修改、仿制、刊发，或以任何侵犯本公司版权的其他方式使用。经过书面授权的引用、刊发，需注明出处为“国金证券股份有限公司”，且不得对本报告进行任何有悖原意的删节和修改。

本报告的产生基于国金证券及其研究人员认为可信的公开资料或实地调研资料，但国金证券及其研究人员对这些信息的准确性和完整性不作任何保证。本报告反映撰写研究人员的不同设想、见解及分析方法，故本报告所载观点可能与其他类似研究报告的观点及市场实际情况不一致，国金证券不对使用本报告所包含的材料产生的任何直接或间接损失或与此有关的其他任何损失承担任何责任。且本报告中的资料、意见、预测均反映报告初次公开发布时的判断，在不作事先通知的情况下，可能会随时调整，亦可因使用不同假设和标准、采用不同观点和分析方法而与国金证券其它业务部门、单位或附属机构在制作类似的其他材料时所给出的意见不同或者相反。

本报告仅为参考之用，在任何地区均不应被视为买卖任何证券、金融工具的要约或要约邀请。本报告提及的任何证券或金融工具均可能含有重大的风险，可能不易变卖以及不适合所有投资者。本报告所提及的证券或金融工具的价格、价值及收益可能会受汇率影响而波动。过往的业绩并不能代表未来的表现。

客户应当考虑到国金证券存在可能影响本报告客观性的利益冲突，而不应视本报告为作出投资决策的唯一因素。证券研究报告是用于服务具备专业知识的投资者和投资顾问的专业产品，使用时必须经专业人士进行解读。国金证券建议获取报告人员应考虑本报告的任何意见或建议是否符合其特定状况，以及（若有必要）咨询独立投资顾问。报告本身、报告中的信息或所表达意见也不构成投资、法律、会计或税务的最终操作建议，国金证券不就报告中的内容对最终操作建议做出任何担保，在任何时候均不构成对任何人的个人推荐。

在法律允许的情况下，国金证券的关联机构可能会持有报告中涉及的公司所发行的证券并进行交易，并可能为这些公司正在提供或争取提供多种金融服务。

本报告并非意图发送、发布给在当地法律或监管规则下不允许向其发送、发布该研究报告的人员。国金证券并不因收件人收到本报告而视其为国金证券的客户。本报告对于收件人而言属高度机密，只有符合条件的收件人才能使用。根据《证券期货投资者适当性管理办法》，本报告仅供国金证券股份有限公司客户中风险评级高于C3级（含C3级）的投资者使用；本报告所包含的观点及建议并未考虑个别客户的特殊状况、目标或需要，不应被视为对特定客户关于特定证券或金融工具的建议或策略。对于本报告中提及的任何证券或金融工具，本报告的收件人须保持自身的独立判断。使用国金证券研究报告进行投资，遭受任何损失，国金证券不承担相关法律责任。

若国金证券以外的任何机构或个人发送本报告，则由该机构或个人为此发送行为承担全部责任。本报告不构成国金证券向发送本报告机构或个人的收件人提供投资建议，国金证券不为此承担任何责任。

此报告仅限于中国境内使用。国金证券版权所有，保留一切权利。

上海	北京	深圳
电话：021-60753903	电话：010-85950438	电话：0755-83831378
传真：021-61038200	邮箱：researchbj@gjzq.com.cn	传真：0755-83830558
邮箱：researchsh@gjzq.com.cn	邮编：100005	邮箱：researchsz@gjzq.com.cn
邮编：201204	地址：北京市东城区建内大街26号	邮编：518000
地址：上海浦东新区芳甸路1088号	新闻大厦8层南侧	地址：深圳市福田区金田路2028号皇岗商务中心
紫竹国际大厦7楼		18楼1806



【小程序】
国金证券研究服务



【公众号】
国金证券研究