

被低估的搜索型数据库

2022年11月21日

► **人工智能领军企业，大数据、安全等领域多点布局取得成效。** 1) **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，处于行业领先地位。2) **大数据领域**：公司是最早从事中文全文检索技术研发的企业，具有全球领先的智能检索技术，现已将领先优势拓展到数据采集、治理、分析挖掘等大数据核心技术的全生命周期，所有底层技术均自主可控。3) **数据安全领域**：公司子公司天行网安是国内最早从事网络安全和数据交换的企业，在数据视频交换、单向导入等方面具有雄厚的技术实力。4) **业绩稳步提升**：2012-2021 年公司整体收入复合增速近 20%。

► **百亿信创数据库市场蓄势待发，搜索引擎细分赛道蕴藏巨大机遇。** 1) **数据库市场**：2021 年国内数据库市场规模超 200 亿元，2021-2025 年复合增速有望达到近 30%。信创数据库市场到 2027 年有望超过 250 亿元，2022-2027 年复合增速有望达到约 35%。2) **搜索型数据库**：自主化势在必行。目前市场上开放源代码的领先的全文搜索引擎主要是 Solr 和 ElasticSearch，而一系列数据泄露、大国科技博弈等事件说明国产替代已经势在必行。3) **拓尔思：国内搜索型数据库龙头，是目前国内少有的从底层分词算法到全文搜索引擎，以及上层系统都完全自研的纯国产搜索型数据库，核心代码自主率高达 100%，适配国内主流信创环境，解决安全相关问题，并且已经具有多个标杆案例。**

► **依托核心能力多点开花，大数据、政务等业务乘风而起。** 1) **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，大型或头部企业级客户超过 8000 家，其中金融领域包括 5 大国有银行、3 大政策性银行、50% 的股份制商业银行、主要交易所及国家金融监管机构和约 300 家地方金融监管部门等用户；政府领域覆盖 80% 的中央和国务院机构，60% 的省级政府，50% 的地市政府用户。中央人民政府门户网站也采用公司软件构建。** 2) **着眼长远布局未来**：云和数据服务保持稳健增长，产品服务的 SaaS 化比例不断提升；积极布局新兴技术，“数字人”已从概念走向落地，并在北京冬奥会正式使用。

► **投资建议**：公司是人工智能领域龙头企业，依托自身能力布局政企大数据及安全领域，在重要细分领域有较高的客户覆盖水平。同时，在信创大趋势下搜索型数据库的国产替代具有重要潜力，公司作为细分领域龙头拥有全自主核心技术且树立多个标杆案例，未来有望长期受益于信创趋势。预计公司 2022-2024 年净利润分别为 3.24/4.20/5.44 亿元，对应 PE 分别为 30X、23X、18X，首次覆盖，给予“推荐”评级。

► **风险提示**：行业竞争加剧的风险；新业务开拓进展不及预期；其他业务存在波动的可能。

盈利预测与财务指标

项目/年度	2021A	2022E	2023E	2024E
营业收入 (百万元)	1029	1308	1646	2077
增长率 (%)	-21.4	27.2	25.8	26.2
归属母公司股东净利润 (百万元)	246	324	420	544
增长率 (%)	-23.0	31.9	29.6	29.6
每股收益 (元)	0.34	0.45	0.59	0.76
PE	39	30	23	18
PB	3.8	3.4	3.0	2.6

资料来源：Wind，民生证券研究院预测；(注：股价为 2022 年 11 月 18 日收盘价)

推荐
首次评级
当前价格：
13.53 元

分析师 吕伟

执业证书：S0100521110003

电话：021-80508288

邮箱：lvwei_yj@mszq.com

分析师 郭新宇

执业证书：S0100518120001

电话：010-85127654

邮箱：guoxinyu@mszq.com

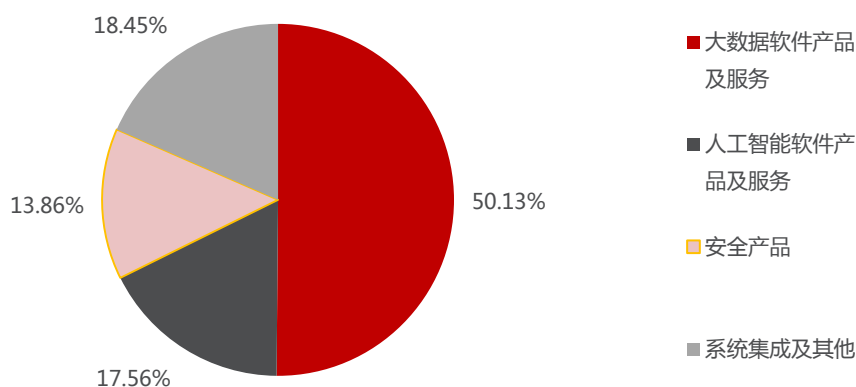
目录

1 人工智能领军企业，大数据、安全等领域多点布局取得成效	3
2 百亿信创数据库市场蓄势待发，搜索引擎细分赛道蕴藏巨大机遇	5
2.1 数据库市场：信创百亿市场蓄势待发，国产厂商大有可为	5
2.2 搜索型数据库：自主化势在必行，国产厂商勇担重任	6
2.3 拓尔思：国内搜索型数据库龙头，国产替代下长期机遇可期	7
3 依托核心能力多点开花，大数据、政务等业务乘风而起	10
4 盈利预测与投资建议	13
4.1 盈利预测假设与业务拆分	13
4.2 估值分析	14
4.3 投资建议	15
5 风险提示	16
插图目录	18
表格目录	18

1 人工智能领军企业 ,大数据、安全等领域多点布局 取得成效

公司是国内人工智能、大数据、数据安全等领域的领军企业。1) 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，处于行业领先地位。2) 大数据领域：公司是最早从事中文全文检索技术研发的企业，具有全球领先的智能检索技术，现已将领先优势拓展到数据采集、治理、分析挖掘等大数据核心技术的全生命周期，所有底层技术均自主可控。3) 数据安全领域：公司子公司天行网安是国内最早从事网络安全和数据交换的企业，发明了国内第一台安全隔离网闸，在数据视频交换、单向导入等方面具有雄厚的技术实力。

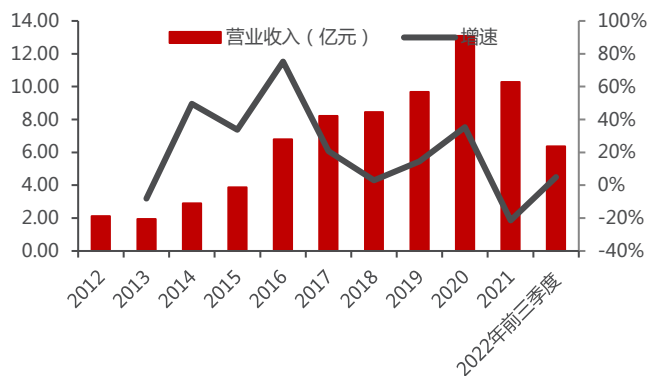
图1：2021 年公司各项细分业务收入占比



资料来源：公司 2021 年年报，民生证券研究院

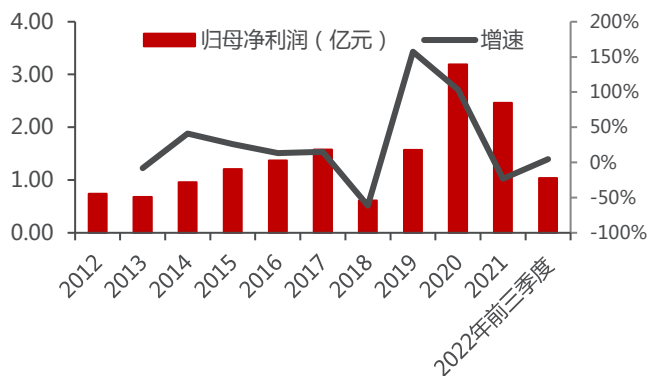
公司整体业绩稳步提升，2012-2021 年收入复合增速近 20%。公司整体业绩稳步提升，收入端 2012-2021 年收入复合增速近 20%，归母净利润 2012-2021 年收入复合增速近 15%。2021 年因为公司西部区域总部和大数据研发和运营基地项目已建成投入使用，其配套商业地产全部在上年度确认收入，因此无房地产业务收入，导致 2021 年收入、利润出现一定波动，但整体看成长趋势不改，2021 年人工智能软件产品及服务同比增速达到 52%。

图2：公司营业收入及同比增速变化



资料来源：wind，民生证券研究院

图3：公司归母净利润及同比增速变化



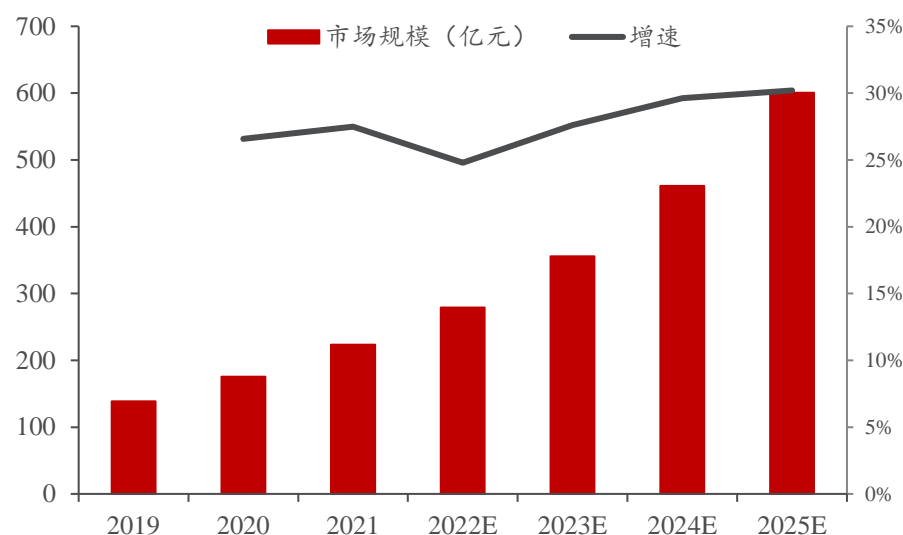
资料来源：wind，民生证券研究院

2 百亿信创数据库市场蓄势待发，搜索引擎细分赛道蕴藏巨大机遇

2.1 数据库市场：信创百亿市场蓄势待发，国产厂商大有可为

2021 年国内数据库市场规模超 200 亿元，2021-2025 年复合增速有望达到近 30%。根据赛迪顾问数据，2021 年，中国数据库市场保持快速增长，规模达到 223.5 亿元，2019-2021 年复合增速达到 27%，预计 2025 年市场规模达到约 600 亿元，2021-2025 年复合增速有望达到 28%。

图4：国内数据库市场规模及增速



资料来源：赛迪顾问，民生证券研究院

对信创数据库市场的测算：

数据库市场空间：2021-2025 年数据根据赛迪顾问数据。参照 2021-2025 年复合增速（28%），保守估计假设 2026、2027 年增速为 20%、20%。

信创目标市场：考虑到 2021 年国内服务器出货量中，30%左右为互联网行业（根据 IDC、华经产业研究院 2021 年数据），但信创主要集中在政府、电信、金融、能源等行业，参考政府、电信、金融市场占比（共约 25%）同时考虑还有医疗、教育等细分市场也存在国产替代机遇，因此假设整个市场中 40%左右为信创市场。由于服务器和数据库往往成一定比例，因此我们也可以假设整个数据库市场中约 40%为信创的目标市场。

海外厂商份额：根据前文数据，2021 年下半年，在国内关系型数据库市场中，Oracle、微软、SAP、IBM 份额总和仍达到近 45%。考虑到国外大厂在国内长期占据市场优势、有部分海外厂商份额在“其他”里，我们假设 2021 年海外厂商总

份额仍达到 60%。后续随着国产替代推进，2022-2027 年海外厂商份额逐步下滑至 50%/45%/40%/35%/30%/25% (上述假设，基于在未来五年内，在党政和重要行业基本实现国产化目标；同时考虑到在部分关键环节中国产替代可能需要一定时间，因此未将国产化率设定为 100%)。

测算结果：2022-2027 年信创数据库市场 (国产厂商在信创市场的可得市场规模) 到 2027 年有望超过 250 亿元，2022-2027 年复合增速有望达到约 35%。

表1：信创数据库市场规模预测

	2021	2022E	2023E	2024E	2025E	2026E	2027E
数据库市场规模 (亿元)	224	279	356	461	601	721	865
yoy		25%	28%	30%	30%	20%	20%
信创目标市场占比	40%	40%	40%	40%	40%	40%	40%
信创目标市场规模 (亿元)	89	112	142	185	240	288	346
海外厂商份额	60%	50%	45%	40%	35%	30%	25%
国产厂商份额	40%	50%	55%	60%	65%	70%	75%
国产厂商可得市场规模 (亿元)	36	56	78	111	156	202	259

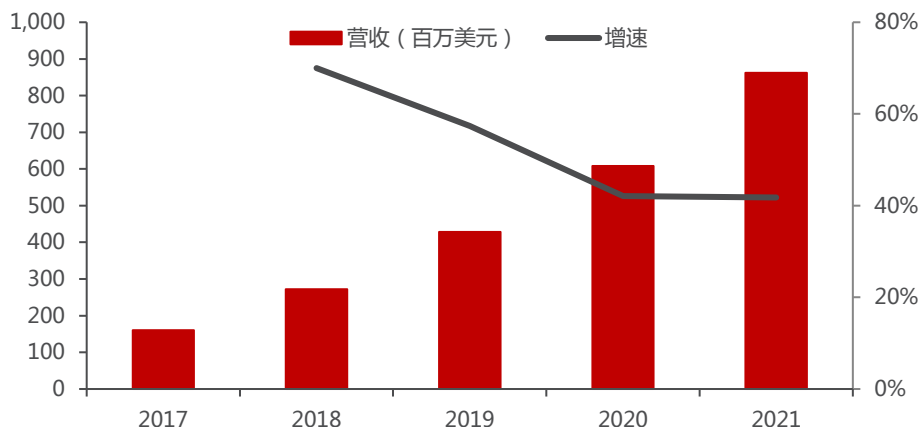
资料来源：赛迪顾问、华经产业研究院、IDC，民生证券研究院预测

2.2 搜索型数据库：自主化势在必行，国产厂商勇担重任

国产数据库浪潮已起，大多数组织都通过关系型数据库很好地管理与利用了结构化数据，但是很多业务的重要见解都隐藏在非结构化数据中。非关系型数据库又称为 NoSQL 数据库，意味着其将不仅仅限于传统的 SQL 二元模式，而是进行更广泛的拓展。随着互联网的兴起，传统的关系数据库在处理电商、社交等超大规模和高并发的 SNS 类型动态网站表现出了功能上的欠缺，而非关系型的数据库则由于其本身的特点得到了非常迅速的发展。

随着我国数字化转型的蓬勃发展，非结构化数据的处理需要依赖基于全文检索的搜索引擎技术，而目前市场上开放源代码的领先的全文搜索引擎主要是 Solr 和 ElasticSearch(以下简称 ES)。ES 自发布以来迅速占领了全文搜索引擎市场，目前国内外很多企业已放弃自主研发，逐渐转投 ES 阵营。根据 ES 财报显示，公司在 2021 年的营业总收入达到 8.62 亿美金，2017-2021 年收入复合增速达到 52%。

图5：ElasticSearch 营收及变化情况



资料来源：wind，民生证券研究院

国外危机频发，搜索型数据库自主势在必行。1) ES 数据泄露事件频发给国内各行业用户敲响了数据安全的警钟。根据 Group-IB 报告显示，2021 年网络上暴露的 ES 实例超过 10 万个，约占 2021 年暴露数据库总数的 30%。2019 年底发生的一起 ES 数据库泄露，包括 27 亿个电子邮件地址，其中 10 亿个密码是以简单的明文存储，涉及国内多家互联网公司。**2) 开源软件和安全工具受到国外出口法律限制。**2021 年美国商务部和安全局发布控制措施，禁止美国公司向中国和俄罗斯等国家出口和转售网络安全产品；2022 年美国商务部和安全局又发布了一项针对网络安全领域的出口管制规定，未经审批禁止向中国分享安全漏洞。**3) 开源协议变更也带来商业风险。**2021 年初，Elastic 公司决定将 Server Side Public License 和 Elastic License 两款开源软件的 Apache License 2.0 变更为双授权许可。其核心条款是“如果将程序的功能或修改后的版本作为服务提供给第三方，那么必须免费公开提供服务源代码”。这意味着不法分子可以获得其源代码并研究其漏洞，给企业用户带来巨大的安全风险

2.3 拓尔思：国内搜索型数据库龙头，国产替代下长期机遇可期

拓尔思作为中文全文检索技术的开创者和引领者，早在上世纪 90 年代就开始深耕搜索型数据库领域。目前公司的 TRS 海贝大数据管理系统（以下简称海贝）是拓尔思自主研发的一款搜索型数据库，适用于数字、文本、地理位置、结构化数据、非结构化数据等所有数据类型，为大数据应用提供高效的数据存储、全文检索、分析统计等数据管理服务。

公司产品具有多个优势，核心代码自主率达 100%：

1) 自主可控：海贝是目前国内少有的从底层分词算法到全文搜索型数据库，以及上层系统都完全自研的纯国产搜索型数据库，核心代码自主率高达 100%。

2) 生态完善：已完成与龙芯、海光、飞腾、鲲鹏等国产芯片以及中标麒麟、

统信 UOS 等国产操作系统的适配工作，同时与达梦、人大金仓、神州通用、南大通用等主流国产数据库完成互认，可为企业搭建完整的信创数据库体系提供支撑。

3 数据安全 采用多副本机制解决数据的可用性问题 通过数据校验以及 WAL 技术解决数据的完整性问题。

4)具有多个标杆案例 :公司目前已经有中国知识产权大数据与智慧服务系统、信用中国等典型标杆案例，为后续拓展市场奠定了重要的基础。

表2：海贝引擎数据库的典型案例分析及具体成果

应用项目	具体成果
中国知识产权大数据与智慧服务系统	TRS Hybase 单一分词器既支持中文、日文、韩文等方块文字，也支持英文、法文、德文等拉丁语系，还支持藏文、蒙文、维文等少数民族语言，实现全语种数据统一处理，即使各种语种都混在一起也能轻松处理，大大降低系统在多语种情况下的数据处理难度，英文词根准确率 99%以上
信用中国	“信用中国”系统，自检索服务迁移至 TRS Hybase 平台后，利用 12 台服务器中的 6 个节点可支撑供每秒 5000+次的高并发检索，且运行稳定，国家公共信用信息中心给予了高度认可和赞赏
第二次全国污染源普查数据清洗项目	通过 TRS ETL 处理原始数据总量为 75,956,218 条数据，交叉对比处理后 56,344,815 条数据。最终交付企业名录库按照国民经济行业分类的“B类-采矿业”、“C类-制造业”等六大门类提供，同时六大门类包括其各自大类、中类、小类，最终交付企业名录库 17,314,608 条数据
工商总局 12315 大数据分析系统	在实际大数据场景下，运行良好，较原有系统有较大性能提升。这也展示了 TRS ETL 与主流大数据计算框架的集成能力，可以将复杂的业务逻辑提交到计算平台进行处理
海南省政府网站集约化项目	新版网站突出集约化这个重点，坚持省级带头、基层跟进，按照“基本栏目+特色栏目”的模式，统一规划，稳步推进，实现了集约整合后的网站技术平台、页面风格和栏目框架三统一
南宁市政府网站集约化项目	全面调动了政府和社会公共服务资源向市民提供服务，充分展示服务型政府对“以人为本，民生关怀”的执政理念，提高市民办事便利性和满意度
襄阳市政府网站集约化项目	实现了全市 5 套班子政务门户网站的集约化管理，规范站群栏目设置、信息采集、信息审核、信息发布等业务流程，进一步突出信息服务、网上办事、互动交流等门户网站的核心功能，全新的襄阳市政府门户网站已成为统一管理、统一形象、统一服务的“一站式”政务服务门户

资料来源：拓尔思官网，民生证券研究院整理

能力已经可以与国外龙头相对比，细节布局更显优势。在搜索能力方面，相较于以 Lucene (Java 语言) 为引擎的 ES，海贝基于一个纯 C 内核的 TRS 引擎，系统资源的使用更加合理可控，复杂搜索响应性能更优。同时，公司提供了更加专业的段、句、位检索、XML 检索，支持度量衡数据的数值区间检索等高级功能。

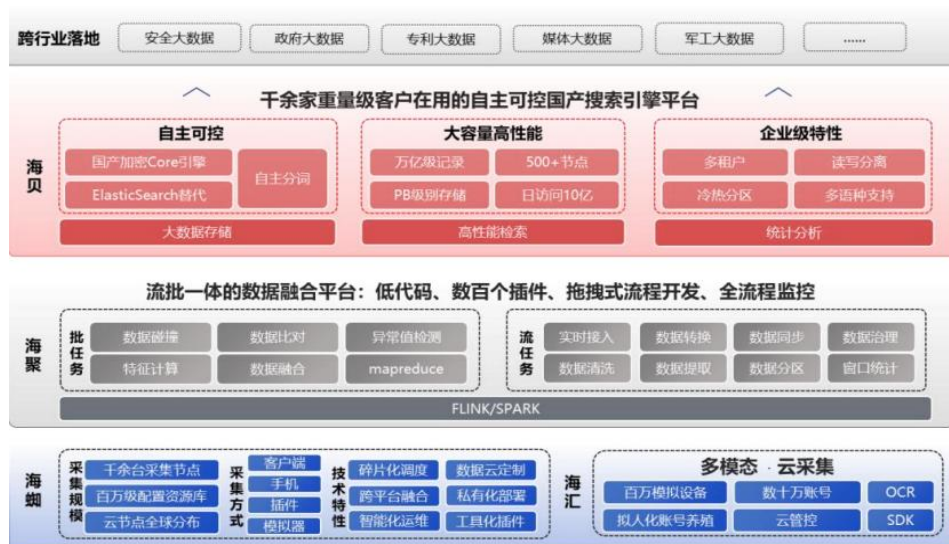
表3：海贝与国外龙头 ES、SolrCloud 在技术等方面的对比

	海贝	ES	SolrCloud
分词器	单一分词器处理全语种	多语种支持（分别处理）	多语种支持（分别处理）
查询语法	Lucene 语法/trs 语法	JSON	Lucene 语法
多引擎	Lucene 引擎（Java）TRS 引擎（C）	Lucene（Java）	Lucene（Java）
易用性	全中文可视化数据管理后台	Kibana（英文）	较弱（英文）
冷热分区	支持	不支持	不支持
读写分离	支持	不支持	不支持
访问隔离	支持	不支持	不支持
索引加密	支持	不支持	不支持
安全控制	支持	不支持	不支持
国密算法	支持	不支持	不支持

资料来源：拓尔思官方公众号，民生证券研究院

从搜索安全拓展到数据管理，TRS 大数据平台是公司的主要数字底座。TRS 大数据平台汇集了海蜘蛛（数据采集）、海汇（采集管控）、海聚（流批一体的融合平台）、海贝（自主可控的国产搜索型平台）四大板块，已经成为公司重要的数据底座。同时，自 2010 年起，公司自主投资建设的大数据中心。目前，TRS 数据中心已具备数千亿数据量的数据索引、标记、查询、挖掘分析能力，万亿级数据总量的秒级检索能力，日均亿级数据获取能力。

图6：ElasticSearch 营收及变化情况

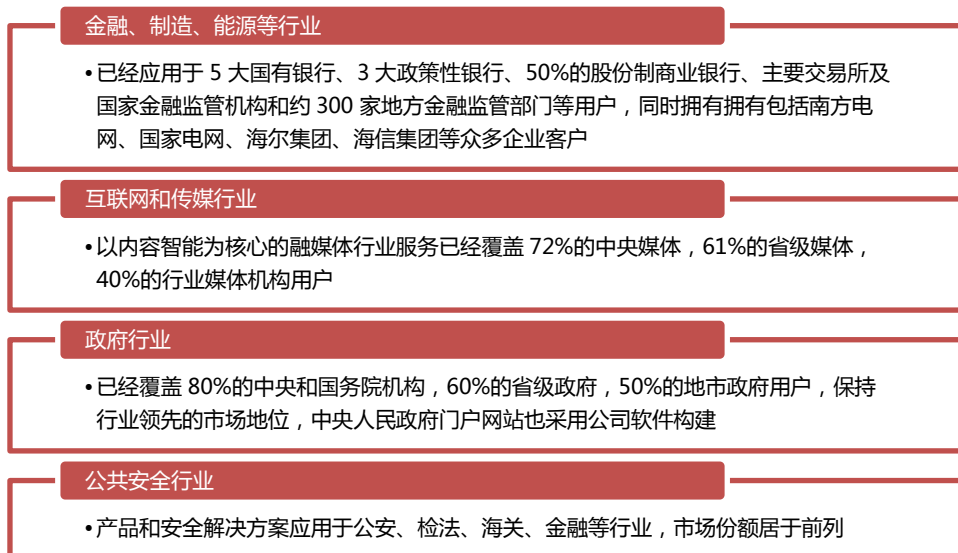


资料来源：公司 2022 年半年报，民生证券研究院

3 依托核心能力多点开花 ,大数据、政务等业务乘风而起

AI 技术为基础，服务不同行业重点客户。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，处于行业领先地位。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。公司“语义智能+行业/领域”覆盖的主要行业包括金融、互联网、传媒、政府、公共安全等，并拥有大量行业重要客户。

图7：公司布局的重点行业及重点客户情况



资料来源：公司 2022 年半年报，民生证券研究院

国家对数据治理高度重视，多个政策和重要事件催化数据治理的发展。1) 中央全面深化改革委员会第二十六次会议审议通过了《关于构建数据基础制度更好发挥数据要素作用的意见》，在下一步任务中明确加快出台数据要素基础制度及配套政策，推进公共数据、企业数据、个人数据分类分级确权授权使用。2) 2022 年 9 月，上海数据集团有限公司正式揭牌成立，是上海加快打造数据要素市场化配置核心载体的重要举措，模式有望全国推广。3) 国务院办公厅印发了《全国一体化政务大数据体系建设指南》，该文件的印发将加速数据横向打通，提升政务数据与社会数据融合应用水平。

表4：近期关于数据治理方面的重要政策

政策名称	核心内容或意义
关于构建数据基础制度更好发挥数据要素作用的意见》	在下一步任务中明确加快出台数据要素基础制度及配套政策，推进公共数据、企业数据、个人数据分类分级确权授权使用，构建数据产权、流通交易、收益分配、安全治理制度规则，统筹推进全国数据要素市场体系。
上海数据集团有限公司正式揭牌成立	组建上海数据集团，是上海加快打造数据要素市场化配置核心载体的重要举措，将加快公共数据和国企数据的要素化进程，激活数据资源，同时带动社会数据的汇聚融合，以市场化运作的方式对外输出数据价值。这一模式有望全国推广
《全国一体化政务大数据体系建设指南》	2023 年底前，全国一体化政务大数据体系初步形成；到 2025 年，全国一体化政务大数据体系更加完备，政务数据管理更加高效，政务数据资源全部纳入目录管理

资料来源：上海国资委官网、中国政府网、光明网，民生证券研究院整理

数字政府领域龙头，政务大数据持续发力。公司在数字政府和公共服务领域形成了产品+云服务+解决方案的业务体系，已经覆盖 80%的中央和国务院机构，60%的省级政府，50%的地市政府用户，保持行业领先的市场地位，中央人民政府门户网站也采用公司软件构建。公司在信用、专利、公安、政务服务和环保等领域，已经帮助各级政府部门建立了多个政务大数据应用；同时积极探索创新领域，先后推出政策大脑、产业大脑等典型应用，为政府部门提供决策支持与治理工具。

云和数据服务保持稳健增长，产品服务的 SaaS 化比例不断提升。公司的“数家-资讯大数据平台”“数星-智能风控平台”“金融风险监测平台”“网察-舆情大数据分析平台”“开源情报分析平台”等均已经实现了产品的全面 SaaS 化，新签约了中邮人寿保险、善智互联、南威软件、中国能源建设集团、河北报业、南京市信息中心等多家重要企事业单位客户。2021 年度营收占比达 36.84%，产品和服务的 SaaS 化比例也快速提升。2022 年前三季度，公司云和数据服务收入占比稳中有升，其中在金融、政府领域增速较快。

积极布局新兴技术，“数字人”已从概念走向落地。依托自主可控的自然语言处理(NLP)技术和行业知识图谱沉淀，拓尔思通过自身积累以及与腾讯等国内顶级数字人形象团队签订战略合作协议，虚拟数字人 SaaS 平台已开发上线，并已经成为公司客户订购虚拟数字人的全链服务平台，在互联网领域、农业、政务等落地了一批服务型虚拟人项目。2022 年北京冬奥会期间，公司发布具有元宇宙自然语言处理技术的虚拟人“小思”，实现自动采编、智能写稿、虚拟播报等全自动功能。

表5：公司“数字人”业务具体布局情况

涉及领域	“数字人”布局具体情况
互联网	公司定制网红主播的真人数字分身，并探索在直播带货领域的场景服务
农业	与广东省共建“南方乡村振兴新农人 AI 数智赋能平台”，定位在农产品直播内容智能创作的 AIGC 在线服务平台，主要面向农产品电商主播提供直播脚本智能创作、带货虚拟人全链租用等知识赋能服务
政务、媒体和出版等	为知识产权出版社和中国体育报等多家用户提供智能播报为主要方向的虚拟主播服务，并围绕知识图谱为核心的智能问答服务建设自主互动的服务型虚拟人建设
重大赛事	2022 年北京冬奥会期间，公司发布具有元宇宙自然语言处理技术的虚拟人“小思”，实现自动采编、智能写稿、虚拟播报等全自动功能

资料来源：公司 2022 年半年报，民生证券研究院整理

图8：北京冬奥会期间，公司的虚拟人“小思”播报北京冬奥会



资料来源：公司官方公众号，民生证券研究院

4 盈利预测与投资建议

4.1 盈利预测假设与业务拆分

大数据软件产品及服务：1) 收入：公司为金融、制造、能源等企业客户，公共安全，政府部门，互联网和传媒等主要行业提供大数据产品及服务。随着政策在数据治理，特别是政务数据治理方面的不断推进，相关领域景气度有望持续提升。同时，公司是最早从事中文全文检索技术研发的企业，具有全球领先的智能检索技术，有望在信创中发挥重要作用。我们认为，随着信创的推进，搜索型数据库的国产替代需求有望持续释放，因此相关业务有望实现持续的高增长。综上，我们预计 2022-2024 年收入增速分别为 30%、30%、30%；**2) 毛利率：**商业模式相对稳定，公司在行业内地位相对稳定且业务模式相对成熟，因此综合考虑 2022-2024 年毛利率保持在 71%的水平。

人工智能软件产品及服务：1) 收入：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，处于行业领先地位。人工智能发展是大势所趋，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，涵盖多个行业/领域，企业级客户超过 8000 家，在人工智能发展的大趋势下有望持续成长，整体看相关业务正处于快速发展期。我们预计 2022-2024 年收入增速分别为 35%、30%、30%。**2) 毛利率：**商业模式相对稳定，公司在行业内地位相对稳定且业务模式相对成熟，因此综合考虑 2022-2024 年毛利率分别为 74%、74%、75%。

安全产品：1) 收入：公司子公司天行网安是国内最早从事网络安全和数据交换的企业，在数据视频交换、单向导入等方面具有雄厚的技术实力。公司面向政府及公共安全市场，政府端已经覆盖 80%的中央和国务院机构，60%的省级政府，50%的地市政府用户，公共安全端面向公安、防务等涉密和安全部门市场，产品和安全解决方案应用于公安、检法、海关、金融等行业，市场份额居于前列，同时在网络舆情技术和数据服务领域处于国内领导地位。考虑到国内对安全的需求不断提升，同时公司具有领先的市场地位，因此假设相关收入保持稳定增长，我们预计 2022-2024 年收入增速分别为 10%、10%、10%。**2) 毛利率：**商业模式相对稳定，公司在行业内地位相对稳定且业务模式相对成熟，下游客户类型也相对稳定，综合考虑 2022-2024 年毛利率分别为 74%、74%、74%。

系统集成及其他：1) 收入：主要是配合公司上述业务进行的业务，因此规模预计随上述业务整体增长而增长 我们预计 2022-2024 年收入增速分别为 25%、20%、20%。**2) 毛利率：**商业模式相对稳定，因此综合考虑 2022-2024 年毛利率分别为 19%、19%、19%。

表6：公司细分业务及毛利率预测

百万元	2021	2022E	2023E	2024E
大数据软件产品及服务	515.7	670.5	871.6	1133.1
增速	-6%	30%	30%	30%
毛利率	71%	71%	71%	71%
人工智能软件产品及服务	180.6	243.8	317.0	412.0
增速	52%	35%	30%	30%
毛利率	74%	74%	74%	75%
安全产品	142.5	156.8	172.5	189.7
增速	8%	10%	10%	10%
毛利率	74%	74%	74%	74%
系统集成及其他	189.8	237.3	284.7	341.7
增速	211%	25%	20%	20%
毛利率	19%	19%	19%	19%
收入总计	1028.7	1308.4	1645.8	2076.5
增速	-21%	27%	26%	26%
毛利率	62%	62%	63%	63%

资料来源：wind，民生证券研究院预测

费用率预测：1) 销售费用率：由于公司数据库、AI等业务正处于拓展期，因此预计将保持一定销售投入强度，预计22-24年销售费用率为11%、11%、11%。
 2) 管理费用率：公司业务扩张预计伴随着合理的人员扩张，因此在业务快速发展的同时，人员预计将持续扩张，预计22-24年管理费用率为13%、13%、13%。
 3) 研发费用率：公司数据库等新兴业务正处于快速拓展期，随着下游客户提出不同类型要求，公司相关研发投入预计将持续，因此预计22-24年研发费用率为10%、10%、10%。

表7：公司主要费用率预测

	2020	2021	2022E	2023E	2024E
销售费用率	11.0%	11.5%	11.0%	11.0%	11.0%
管理费用率	8.1%	14.8%	13.0%	13.0%	13.0%
研发费用率	6.3%	9.5%	10.0%	10.0%	10.0%

资料来源：wind，民生证券研究院预测

4.2 估值分析

根据公司三大主要业务，选取业务中涉及 AI、数据库或大数据、安全相关业务的公司，包括科大讯飞（涉及上述所有业务）、太极股份（涉及上述所有业务）、海天瑞声（涉及上述 AI、大数据等业务），且相关公司均为细分领域领军企业，与公司的市场地位具有一定可比性。相关公司2022/2023年平均PE分别为66X、37X，公司PE为30X、23X，具有一定估值优势。

表8：公司与可比公司估值对比表

股票代码	公司简称	收盘价 (元)	EPS (元)			PE (倍)		
			2021A	2022E	2023E	2021A	2022E	2023E
002230	科大讯飞	34.85	0.67	0.80	1.11	52	44	31
688787	海天瑞声	63.70	0.74	0.56	1.31	86	114	49
002368	太极股份	30.89	0.64	0.78	0.96	48	40	32
	平均					62	66	37
300229	拓尔思	13.53	0.34	0.45	0.59	39	30	23

资料来源：wind，民生证券研究院预测（收盘价为2022年11月18日收盘价；可比公司eps来自于wind一致预期）

4.3 投资建议

公司是人工智能领域龙头企业，依托自身能力布局政企大数据及安全领域，在重要细分领域有较高的客户覆盖水平。同时，在信创大趋势下搜索型数据库的国产替代具有重要潜力，公司作为细分领域龙头拥有全自主核心技术且树立多个标杆案例，未来有望长期受益于信创趋势。预计公司2022-2024年净利润分别为3.24/4.20/5.44亿元，对应PE分别为30X、23X、18X，首次覆盖，给予“推荐”评级。

5 风险提示

1) 行业竞争加剧的风险。政务、金融等信息化分为多个细分领域，细分领域竞争有加剧的可能；同时，公司还向其他行业渗透，同样将面对对应细分领域已有竞争者的竞争。

2) 新业务开拓进展不及预期。公司数据库等业务属于创新业务，而且在行业内也属于率先开拓相关业务的探索者。未来业务拓展中可能遇到困难，导致业务拓展进度存在不确定性。

3) 其他业务存在波动的可能。公司大部分业务近年来业绩存在波动，相关领域由于自身财力等方面影响，下游景气度提升缓慢，未来仍存在波动的可能。

公司财务报表数据预测汇总

利润表 (百万元)	2021A	2022E	2023E	2024E
营业总收入	1029	1308	1646	2077
营业成本	390	494	613	759
营业税金及附加	11	13	17	21
销售费用	118	144	181	228
管理费用	153	170	214	270
研发费用	98	131	165	208
EBIT	284	393	503	649
财务费用	20	36	39	35
资产减值损失	-13	-5	-5	-5
投资收益	19	13	16	21
营业利润	274	363	465	616
营业外收支	4	1	1	1
利润总额	279	364	466	617
所得税	31	36	42	62
净利润	248	327	424	556
归属于母公司净利润	246	324	420	544
EBITDA	371	511	698	967

资产负债表 (百万元)	2021A	2022E	2023E	2024E
货币资金	491	327	528	881
应收账款及票据	326	384	482	609
预付款项	3	2	2	3
存货	94	103	113	120
其他流动资产	669	727	804	902
流动资产合计	1583	1542	1929	2514
长期股权投资	77	91	107	128
固定资产	298	223	237	693
无形资产	302	348	349	357
非流动资产合计	2143	2558	2657	2696
资产合计	3726	4101	4587	5210
短期借款	10	5	2	2
应付账款及票据	214	271	319	374
其他流动负债	260	299	363	432
流动负债合计	484	574	684	808
长期借款	0	0	0	0
其他长期负债	664	656	654	659
非流动负债合计	664	656	654	659
负债合计	1148	1230	1339	1467
股本	715	715	715	715
少数股东权益	14	18	22	33
股东权益合计	2578	2871	3248	3742
负债和股东权益合计	3726	4101	4587	5210

资料来源：公司公告、民生证券研究院预测

主要财务指标	2021A	2022E	2023E	2024E
成长能力 (%)				
营业收入增长率	-21.44	27.18	25.79	26.17
EBIT 增长率	-24.24	38.47	27.83	29.14
净利润增长率	-23.00	31.91	29.61	29.57
盈利能力 (%)				
毛利率	62.13	62.27	62.75	63.47
净利润率	23.89	24.78	25.53	26.22
总资产收益率 ROA	6.60	7.91	9.16	10.45
净资产收益率 ROE	9.59	11.36	13.02	14.68
偿债能力				
流动比率	3.27	2.69	2.82	3.11
速动比率	2.61	2.02	2.15	2.43
现金比率	1.02	0.57	0.77	1.09
资产负债率 (%)	30.82	29.99	29.18	28.17
经营效率				
应收账款周转天数	104.72	100.00	100.00	100.00
存货周转天数	87.77	80.00	70.00	60.00
总资产周转率	0.31	0.33	0.38	0.42
每股指标 (元)				
每股收益	0.34	0.45	0.59	0.76
每股净资产	3.59	3.99	4.51	5.19
每股经营现金流	0.26	0.60	0.79	1.07
每股股利	0.05	0.07	0.09	0.11
估值分析				
PE	39	30	23	18
PB	3.8	3.4	3.0	2.6
EV/EBITDA	26.50	19.53	13.99	9.75
股息收益率 (%)	0.37	0.49	0.63	0.82

现金流量表 (百万元)	2021A	2022E	2023E	2024E
净利润	248	327	424	556
折旧和摊销	87	118	195	317
营运资金变动	-161	-46	-84	-130
经营活动现金流	189	426	564	766
资本开支	-476	-630	-203	-275
投资	-307	-80	-50	-30
投资活动现金流	-754	-494	-253	-305
股权募资	0	0	0	0
债务募资	8	-4	-11	0
筹资活动现金流	756	-96	-110	-108
现金净流量	190	-164	201	353

插图目录

图 1：2021 年公司各项细分业务收入占比.....	3
图 2：公司营业收入及同比增速变化.....	4
图 3：公司归母净利润及同比增速变化.....	4
图 4：国内数据库市场规模及增速.....	5
图 5：ElasticSearch 营收及变化情况.....	7
图 6：ElasticSearch 营收及变化情况.....	9
图 7：公司布局的重点行业及重点客户情况.....	10
图 8：北京冬奥会期间，公司的虚拟人“小思”播报北京冬奥会.....	12

表格目录

盈利预测与财务指标.....	1
表 1：信创数据库市场规模预测.....	6
表 2：海贝引擎数据库的典型案例分析及具体成果.....	8
表 3：海贝与国外龙头 ES、SolrCloud 在技术等方面的对比.....	9
表 4：近期关于数据治理方面的重要政策.....	11
表 5：公司“数字人”业务具体布局情况.....	11
表 6：公司细分业务及毛利率预测.....	14
表 7：公司主要费用率预测.....	14
表 8：公司与可比公司估值对比表.....	15
公司财务报表数据预测汇总.....	17

分析师承诺

本报告署名分析师具有中国证券业协会授予的证券投资咨询执业资格并登记为注册分析师，基于认真审慎的工作态度、专业严谨的研究方法与分析逻辑得出研究结论，独立、客观地出具本报告，并对本报告的内容和观点负责。本报告清晰地反映了研究人员的研究观点，结论不受任何第三方的授意、影响，研究人员不曾因、不因、也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿。

评级说明

投资建议评级标准	评级	说明
以报告发布日后的 12 个月内公司股价（或行业指数）相对同期基准指数的涨跌幅为基准。其中：A 股以沪深 300 指数为基准；新三板以三板成指或三板做市指数为基准；港股以恒生指数为基准；美股以纳斯达克综合指数或标普 500 指数为基准。	推荐	相对基准指数涨幅 15%以上
	谨慎推荐	相对基准指数涨幅 5% ~ 15%之间
	中性	相对基准指数涨幅-5% ~ 5%之间
	回避	相对基准指数跌幅 5%以上
行业评级	推荐	相对基准指数涨幅 5%以上
	中性	相对基准指数涨幅-5% ~ 5%之间
	回避	相对基准指数跌幅 5%以上

免责声明

民生证券股份有限公司（以下简称“本公司”）具有中国证监会许可的证券投资咨询业务资格。

本报告仅供本公司境内客户使用。本公司不会因接收人收到本报告而视其为客户。本报告仅为参考之用，并不构成对客户的投资建议，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，客户应当充分考虑自身特定状况，不应单纯依靠本报告所载的内容而取代个人的独立判断。在任何情况下，本公司不对任何人因使用本报告中的任何内容而导致的任何可能的损失负任何责任。

本报告是基于已公开信息撰写，但本公司不保证该等信息的准确性或完整性。本报告所载的资料、意见及预测仅反映本公司于发布本报告当日的判断，且预测方法及结果存在一定程度局限性。在不同时期，本公司可发出与本报告所刊载的意见、预测不一致的报告，但本公司没有义务和责任及时更新本报告所涉及的内容并通知客户。

在法律允许的情况下，本公司及其附属机构可能持有报告中提及的公司所发行证券的头寸并进行交易，也可能为这些公司提供或正在争取提供投资银行、财务顾问、咨询服务等相关服务，本公司的员工可能担任本报告所提及的公司的董事。客户应充分考虑可能存在的利益冲突，勿将本报告作为投资决策的唯一参考依据。

若本公司以外的金融机构发送本报告，则由该金融机构独自为此发送行为负责。该机构的客户应联系该机构以交易本报告提及的证券或要求获悉更详细的信息。本报告不构成本公司向发送本报告金融机构之客户提供的投资建议。本公司不会因任何机构或个人从其他机构获得本报告而将其视为本公司客户。

本报告的版权仅归本公司所有，未经书面许可，任何机构或个人不得以任何形式、任何目的进行翻版、转载、发表、篡改或引用。所有在本报告中使用的商标、服务标识及标记，除非另有说明，均为本公司的商标、服务标识及标记。本公司版权所有并保留一切权利。

民生证券研究院：

上海：上海市浦东新区浦明路 8 号财富金融广场 1 幢 5F； 200120

北京：北京市东城区建国门内大街 28 号民生金融中心 A 座 18 层； 100005

深圳：广东省深圳市福田区益田路 6001 号太平金融大厦 32 层 05 单元； 518026