

拓尔思 (300229.SZ)

融媒体 AI 领军，加速落地进行时

积极拥抱 AI 产业趋势，加码布局垂直行业大模型。通用大模型使用的大规模训练语料来自互联网公开的通用数据，提供的是通用能力，适用于通用场景应用，在行业应用的专业度、精度、深度方面存在局限。公司积极拥抱大模型发展，启动了 trsGPT 的研发，可分为 trsGPT-G（政务专业模型）、trsGPT-F（金融专业模型）、trsGPT-M（媒体专业模型），在政务、金融、新闻媒体等垂直领域做应用落地。

多年深耕融媒体行业，拥有大量主流央媒客户资源，行业卡位优秀。融媒体是公司的成熟业务版块，主要面向各级新闻媒体单位围绕新闻生产策、采、编、发、评、运、屏全流程提供融媒体内容生产、用户行为资产、媒体大数据等综合服务。目前公司的融媒体业务已经合作了人民日报、新华通讯社、中央广播电视总台等主流央媒及其他新闻出版单位。截至 2022 年，公司融媒体用户已经覆盖了 72% 的中央媒体，61% 的省级媒体，40% 的行业媒体。

具备深厚新闻媒体领域高质量数据积累，为垂直领域大模型效果打下基础。1) 部分 ToB 场景中，用户对安全性、一致性、规范性、意识形态敏感性的要求更高，需要更多的定制和额外的工作等等，对训练数据质量提出更高要求。2) 数据层面，拓尔思从 2010 年自建数据中心以来，已采集了超过 10 年的互联网公开数据，拥有规模及质量均位列业界前茅的另类数据资产。目前，拓尔思拥有来自境内外、各行各业的公开数据规模超 1300 亿条，数据总量达 100TB 以上。拓尔思长年深耕政务、媒体、金融、专利等行业的信息化建设，已积累了 30+ 专业知识库，覆盖媒体、舆情、金融风控、产业投研、智能消保、开源情报、政务应用等场景；公司在新闻媒体等行业的深厚数据积累，将有利于公司将大模型在垂直领域 fine-tuning 后做落地，为公司垂直领域大模型的效果打下基础。

AIGC 赋能新闻媒体行业，融合大模型技术后公司未来有望实现产品化快速落地。近年来，公司积极与客户携手在 AIGC 相关领域，事实上，在 ChatGPT 出现之前，公司在 AIGC 领域已有一定的业务布局。根据公司 2022 年年报披露，2022 年公司在 AIGC 领域（机器写作、自动报告生成、对话式 AI 机器人、虚拟人播报等）的用户数已达 103 家，业务收入达 852 万元。近来，公司尝试在自研的自动写作产品基础上，基于大模型进行融合迭代，进一步提升现有 AIGC 产品的内容生成质量，取得一定效果。我们预计，伴随着公司未来 trsGPT 大模型的推出，公司有望快速实现 AIGC 产品化的落地，为客户带来更优秀的 AI 辅助创作效果。

维持“买入”评级。我们预计公司 2023-2025 年将实现收入 11.77/14.61/18.05 亿元，实现归母净利润 2.03/2.25/3.24 亿元，维持“买入”评级。

风险提示：AI 技术迭代不及预期风险；经济下行超预期风险；行业竞争加剧风险。

财务指标	2021A	2022A	2023E	2024E	2025E
营业收入（百万元）	1,029	907	1,177	1,461	1,805
增长率 yoy (%)	-21.4	-11.8	29.7	24.1	23.6
归母净利润（百万元）	246	128	203	225	324
增长率 yoy (%)	-23.0	-48.0	58.7	11.2	43.7
EPS 最新摊薄（元/股）	0.31	0.16	0.25	0.28	0.41
净资产收益率 (%)	9.6	4.6	7.1	7.4	9.6
P/E (倍)	87.0	167.4	105.5	94.8	66.0
P/B (倍)	9.1	8.6	8.0	7.5	6.8

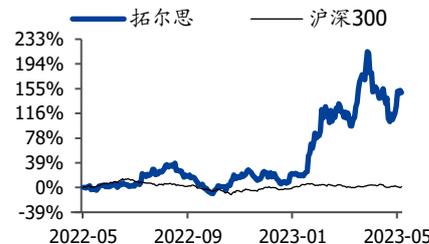
资料来源：Wind，国盛证券研究所 注：股价为 2023 年 5 月 15 日收盘价

买入（维持）

股票信息

行业	软件开发
前次评级	买入
5月15日收盘价(元)	26.88
总市值(百万元)	21,377.45
总股本(百万股)	795.29
其中自由流通股(%)	99.92
30日日均成交量(百万股)	109.04

股价走势



作者

分析师 刘高畅

执业证书编号：S0680518090001

邮箱：liugaochang@gszq.com

相关研究

- 《拓尔思 (300229.SZ)：2023 一季报利润增速亮眼，加码布局大模型及 AIGC》2023-04-15
- 《拓尔思 (300229.SZ)：业绩符合预期，疫情压力下维持稳健增长》2022-08-15
- 《拓尔思 (300229.SZ)：三十年磨一剑，政府 NLP 领军的稳健之路》2022-05-16



内容目录

1.泛政府领域 NLP 龙头，积极拥抱大模型产业趋势	4
2.深度卡位融媒体行业，高质量数据赋能 AIGC	6
3.投资建议	8
4.风险提示	8

图表目录

图表 1: 拓尔思是国内泛政府领域 NLP 领军企业	4
图表 2: 拓尔思 trsGPT 技术栈	5
图表 3: 公司优质央媒客户资源丰富，坐拥优质行业卡位	6
图表 4: 公司的海贝数据库可用于安全、政务、媒体等领域	7
图表 5: 拓尔思 AIGC 实践案例	7
图表 6: 公司盈利预测	8

1.泛政府领域 NLP 龙头，积极拥抱大模型产业趋势

公司是政府政务领域 NLP 龙头。1) 拓尔思成立于 1993 年，以“语义智能+”为发展战略深耕近 30 年，业务主要面向三大重点板块：内容安全和互联网空间治理、数字政府和数据智能，覆盖从党政机关到企事业单位的众多行业市场，提供大数据检索、智能风控营销、舆情监控等服务。公司以深耕多年的大数据积累与行业领先的 NLP 技术成为国内政府政务领域 NLP 龙头，截至 2022 年，公司融媒体用户已经覆盖了 72% 的中央媒体，61% 的省级媒体，40% 的行业媒体；政府网站集约化用户已经覆盖 80% 的中央和国务院机构，60% 的省级政府，50% 的地市政府。金融风险防控全链条云产品用户已经覆盖了 55% 的全国省市两级金融监管机构。政务舆情用户覆盖了全国近 50% 的省级政府。

图表 1: 拓尔思是国内泛政府领域 NLP 领军企业

公司介绍

拓尔思信息技术股份有限公司(简称“拓尔思”)成立于1993年，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。公司2011年在深圳证券交易所创业板上市，股票代码300229。公司净资产近30亿元，拥有员工2000余人，在全国设有4大区域中心31家分公司。

自成立以来，拓尔思始终坚持核心技术自主研发，拥有40+发明专利、1000+软件著作权，一直引领全文检索和搜索引擎数据库、自然语言处理(NLP)技术的原始创新和应用场景落地，成功服务了一批国家级重要信息化项目的建设，成为国家企业信用信息公示系统、信用中国、国家知识产权检索系统、GA云搜索等国家级基础设施的数据管理和搜索引擎。公司产品曾获得国家科技进步二等奖、上海市科技进步一等奖、中国电子学会一等奖、北京市科技进步二等奖、王选新闻科技特等奖等多项重要荣誉。

拓尔思坚持“数智+赛道”的发展战略，以优秀的科技能力打造丰富的行业解决方案，为企业、政府实现数字化升级赋能。公司自主研发的TRS大数据平台、TRS人工智能平台和安全一体化平台，已成为行业领先的数字化基座。同时，公司已建立了大规模的数据和知识资产，正在全面构建云和数据服务生态(SaaS&DaaS)。目前，公司的数智产品与服务已被国内外10000余家企业级用户广泛使用，尤其在企业数字化转型、数字政府、网络内容安全、融媒体、金融科技和开源情报等领域实现了深度应用。

未来，拓尔思将不忘初心与守正创新，在大数据、人工智能和数据安全领域坚持自主可控的核心技术研发，积极探索创新场景，加强建设合作生态，全面深化行业应用，通过不断满足用户的信息化和数字化需求，助力数字中国高质量发展，为实现“两个一百年”奋斗目标和中华民族伟大复兴的中国梦不断作出应有的贡献！

30年 技术积累	三大核心技术领域 大数据 人工智能 数据安全	A股第一家 上市的大数据技术企业	市场占有率第一 网络舆情监测 政府门户网站 融媒体 边界安全 金融风控
-------------	-------------------------------	---------------------	---

资料来源：拓尔思官网，国盛证券研究所

积极拥抱 AI 产业趋势，加码布局垂行业大模型。通用大模型使用的大规模训练语料来自互联网公开的通用数据，提供的是通用能力，适用于通用场景应用，在行业应用的专业度、精度、深度方面存在局限。公司积极拥抱大模型发展，启动了 trsGPT 的研发，其技术栈可以分为 5 层：

- 应用层：将 trsGPT 生成的 trsGPT-G（政务专业模型）、trsGPT-F（金融专业模型）、trsGPT-M（媒体专业模型）。三大行业模型（可通过接口）集成到自研云服务或面向用户的应用程序，运行自有模型或通过第三方接口运行模型，帮助用户使用 trsGPT 带来的人工智能协作能力和生成能力，形成各类行业应用，如公文辅助写作、投研自动报告生成、智能投研问答、新闻资讯知识型搜索、以文生图配稿等。
- 接口层：将应用层和模型层衔接，方便应用层调用，使得开发者和用户能够以编程方式与模型进行交互。这可以简化 trsGPT 在实际应用中的部署和调用，从而降低使用门槛。
- 模型层：trsGPT 支持接入各类主流大模型，包括各类开源或非开源模型，以及各种模型的共享平台。大模型负责向 trsGPT 模型输出模型能力，trsGPT 模型更精确地处理自己“擅长”的任务。这一层提供了不同的模型数据和功能，通过接口层为应

用层提供专业模型的功能支持，包括各种调用 API 和数据中心调用工具，同时提供对应的提示工程接口和模型精调接口等。

- 框架层：提供训练或云部署的深度学习框架和中间件等，包括 PyTorch、TensorFlow 等知名深度学习框架和中间件。
- 计算层：为模型层提供模型计算和调度的各种算力支持，为训练 AI 模型运行训练和运行推理任务提供基础设施。计算层包括了各种云计算平台和计算芯片。

图表 2: 拓尔思 trsGPT 技术栈



图： trsGPT 技术栈

资料来源：公司公告，国盛证券研究所

2. 深度卡位融媒体行业，高质量数据赋能 AIGC

多年深耕融媒体行业，拥有大量主流央媒客户资源，行业卡位优秀。融媒体是公司的成熟业务版块，主要面向各级新闻媒体单位围绕新闻生产策、采、编、发、评、运、屏全流程提供融媒体内容生产、用户行为资产、媒体大数据等综合服务。融媒体内容生产服务平台是以内容资产为核心的新一代数据型媒体业务平台，涵盖了报、网、端、微、视和自媒体平台的全流程内容生产发布管理，该平台由数据资源聚合、融合生产创作、选题策划分析、传播效果分析、协调指挥调度、用户资产运营、大屏综合管理等多个平台组成。目前，公司的融媒体业务主要以“软件产品+大数据服务”的融合模式为客户提供服务，已经合作了人民日报、新华通讯社、中央广播电视总台、解放军报、光明日报、经济日报、中国日报、科技日报、中国纪检监察报、工人日报、农民日报、中国新闻社等各类主流央媒及其他新闻出版单位。截至 2022 年，公司融媒体用户已经覆盖了 72% 的中央媒体，61% 的省级媒体，40% 的行业媒体。

图表 3: 公司优质央媒客户资源丰富，坐拥优质行业卡位



资料来源：公司官网，国盛证券研究所

具备深厚新闻媒体领域高质量数据积累，为垂直领域大模型效果打下基础。

1) 在部分 ToB 场景中，可用的数据集是有限的，需要更多工程化和特定方法的干预；部分 ToB 场景中，用户对安全性、一致性、规范性、意识形态敏感性的要求更高，需要更多的定制和额外的相关工作等等。而这些都是拓尔思不断在思考、研究、开发和应用实践去解决的问题。

2) 数据层面，拓尔思从 2010 年自建数据中心以来，已采集了超过 10 年的互联网公开数据，拥有规模及质量均位列业界前茅的另类数据资产。目前，拓尔思拥有来自境内外、各行各业的公开数据规模超 1300 亿条，数据类型涵盖新闻、资讯、政策、视频、图片、百科、社交等多模态，数据总量达 100TB 以上。公司的海贝数据库有千余家重量级客户在使用，可用于公共安全、政务公开与媒体融合等领域。算法层面，拓尔思长年深耕政务、媒体、金融、专利等行业的信息化建设，已积累了 30+ 专业知识库，涵盖通用语义分析、人物/机构、行业分类、专利、媒体、金融、科技情报、乡村振兴等领域；30000+ 标签模型，覆盖媒体、舆情、金融风控、产业投研、智能消保、开源情报、政务应用等场景；350+ 深度学习算法模型，包括 NLP、金融监管、风控征信、公共安全、产业服务、传播分析、事件研判、舆情态势等通用模型、指数模型、领域模型等。公司在新闻媒体等行业的深厚数据积累，将有利于公司将大模型在垂直领域 fine-tuning 后做落地，为公司垂直领域大模型的效果打下基础。

图表 4: 公司的海贝数据库可用于安全、政务、媒体等领域



资料来源: 公司官网, 国盛证券研究所

AIGC 有望赋能新闻媒体行业，辅助智能写稿与内容生成。近年来，公司以数据智能应用为核心赋能不同行业的数字化转型与降本增效的场景应用，积极与客户携手在 AIGC 相关领域，如机器写作、对话式 AI、内容人机协同和自动报告生成等应用场景相继打造出一批实践案例，在经济日报、浙江日报等新闻媒体出版单位已有落地应用。早在 2016 年，拓尔思在浙江日报报业集团的融媒体智能传播服务平台（简称浙报媒立方）率先推出了机器写作，主要根据拓尔思的全网资讯大数据自动生成体育、气象和财经等领域的主题稿件，推送到浙报媒立方。

图表 5: 拓尔思 AIGC 实践案例



图: 拓尔思 AIGC 实践案例

资料来源: 公司公告, 国盛证券研究所

融合大模型技术后，公司未来有望实现产品化快速落地，为客户带来更好的 AIGC 使用效果。事实上，在 ChatGPT 出现之前，公司在 AIGC 领域已有一定的业务布局。根据公司 2022 年年报披露，2022 年公司在 AIGC 领域（机器写作、自动报告生成、对话式 AI 机器人、虚拟人播报等）的用户数已达 103 家，业务收入达 852 万元。2022 年发布的 9 款 SaaS 服务产品中，有 8 款是融合了大数据和人工智能技术，如麦文智创云服务作为一款农产品电商直播脚本的自动写作产品，在第七届广东省农村电子商务峰会暨数美汕尾助农电商消费节上正式向社会各界发布，得到业界一致好评。近来，公司尝试在自研的自动写作产品基础上，基于大模型进行融合迭代，进一步提升现有 AIGC 产品的内容生成质量，取得一定效果。我们预计，伴随着公司未来 trsGPT 大模型的推出，公司有望快速实现 AIGC 产品化的落地，为客户带来更优秀的 AI 辅助创作效果。

3.投资建议

维持“买入”评级。根据以下盈利预测，我们预计公司 2023-2025 年将实现收入 11.77/14.61/18.05 亿元，实现归母净利润 2.03/2.25/3.24 亿元，维持“买入”评级。

图表 6: 公司盈利预测

公司各业务增速预测	2021A	2022A	2023E	2024E	2025E
金融、制造、能源等企业客户	43.38%	0.88%	45%	35%	30%
公共安全	104.84%	-31.84%	40%	30%	0.30
政府行业	-29.00%	-13.69%	0%	0%	0%
互联网和传媒	-51.95%	33.60%	10%	0%	0%
合计增速	-21.44%	-11.81%	29.72%	24.14%	23.56%
营业收入（亿元）	10.29	9.07	11.77	14.61	18.05

资料来源：公司公告，国盛证券研究所

4.风险提示

AI 技术迭代不及预期风险：若 AI 技术迭代不及预期，NLP 技术理解人类意图水平未能取得突破，则对产业链相关公司会造成一定不利影响。

经济下行超预期风险：若宏观经济景气度下行，固定资产投资额放缓，影响企业再投资意愿，从而影响消费者消费意愿和产业链生产意愿，对整个行业将会造成不利影响，NLP 技术应用落地将会受限。

行业竞争加剧风险：若相关企业加快技术迭代和应用布局，整体行业竞争程度加剧，将会对目前行业内企业的增长产生威胁。

免责声明

国盛证券有限责任公司（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，本公司不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。

本报告的信息均来源于本公司认为可信的公开资料，但本公司及其研究人员对该等信息的准确性及完整性不作任何保证。本报告中的资料、意见及预测仅反映本公司于发布本报告当日的判断，可能会随时调整。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。本公司不保证本报告所含信息及资料保持在最新状态，对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本公司力求报告内容客观、公正，但本报告所载的资料、工具、意见、信息及推测只提供给客户作参考之用，不构成任何投资、法律、会计或税务的最终操作建议，本公司不就报告中的内容对最终操作建议做出任何担保。本报告中所指的投资及服务可能不适合个别客户，不构成客户私人咨询建议。投资者应当充分考虑自身特定状况，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。

投资者应注意，在法律许可的情况下，本公司及其本公司的关联机构可能会持有本报告中涉及的公司所发行的证券并进行交易，也可能为这些公司正在提供或争取提供投资银行、财务顾问和金融产品等各种金融服务。

本报告版权归“国盛证券有限责任公司”所有。未经事先本公司书面授权，任何机构或个人不得对本报告进行任何形式的发布、复制。任何机构或个人如引用、刊发本报告，需注明出处为“国盛证券研究所”，且不得对本报告进行有悖原意的删节或修改。

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，本报告所表述的任何观点均精准地反映了我们对标的证券和发行人的个人看法，结论不受任何第三方的授意或影响。我们所得报酬的任何部分无论是在过去、现在及将来均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。

投资评级说明

投资建议的评级标准		评级	说明
评级标准为报告发布日后的6个月内公司股价（或行业指数）相对同期基准指数的相对市场表现。其中A股市场以沪深300指数为基准；新三板市场以三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）为基准；香港市场以摩根士丹利中国指数为基准，美股市场以标普500指数或纳斯达克综合指数为基准。	股票评级	买入	相对同期基准指数涨幅在15%以上
		增持	相对同期基准指数涨幅在5%~15%之间
		持有	相对同期基准指数涨幅在-5%~+5%之间
		减持	相对同期基准指数跌幅在5%以上
	行业评级	增持	相对同期基准指数涨幅在10%以上
		中性	相对同期基准指数涨幅在-10%~+10%之间
减持		相对同期基准指数跌幅在10%以上	

国盛证券研究所

北京

地址：北京市西城区平安里西大街26号楼3层

邮编：100032

传真：010-57671718

邮箱：gsresearch@gszq.com

南昌

地址：南昌市红谷滩新区凤凰中大道1115号北京银行大厦

邮编：330038

传真：0791-86281485

邮箱：gsresearch@gszq.com

上海

地址：上海市浦明路868号保利One56 1号楼10层

邮编：200120

电话：021-38124100

邮箱：gsresearch@gszq.com

深圳

地址：深圳市福田区福华三路100号鼎和大厦24楼

邮编：518033

邮箱：gsresearch@gszq.com