

证券代码：300229

证券简称：拓尔思

拓尔思信息技术股份有限公司 投资者关系活动记录表

编号：2023-002

投资者关系活动类别	<input checked="" type="checkbox"/> 特定对象调研 <input type="checkbox"/> 分析师会议 <input type="checkbox"/> 媒体采访 <input type="checkbox"/> 业绩说明会 <input type="checkbox"/> 新闻发布会 <input type="checkbox"/> 路演活动 <input type="checkbox"/> 现场参观 <input type="checkbox"/> 其他
参与单位名称及人员姓名	兴业证券：吴鸣远； 嘉实基金：何鸣晓、顾晶菁、罗丹、刘晔； 景顺长城基金：张雪薇。
时间	2023年2月20日 10:00-11:00、15:00-16:00。
地点	拓尔思信息技术股份有限公司会议室
上市公司接待人员姓名	副总经理、董事会秘书李党生先生。
投资者关系活动主要内容介绍	<p>一、副总经理、董事会秘书李党生先生就公司主营业务、核心产品及优势、未来战略规划进行介绍。</p> <p>二、参会机构提问及回答。</p> <p>（一）公司在 AIGC 领域的应用案例有哪些？</p> <p>回答：知识图谱是 AIGC 利用 AI 学习实现自动生成的基础之一。早在 2015 年，拓尔思参与了一项基于大数据的类人智能关键技术与系统的 863 课题《面向基础教育的知识能力智能测评与类人答题验证-高考机器人》，主要负责十亿级条目规模知识图谱以及大数据知识管理和检索平台的构建，助力该“高考机器人”能实现对类人答题原型系统的试题理解、常识记忆、语言运用、知识关联、知识推理、初等数学求解等各个智力维度进行科学评价。该项目于 2017 年顺利通过验收。</p>

基于 NLP 技术的文本生成可以算作是 AIGC 中发展最早的一部分技术。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中（简称“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。

2019 年，拓尔思提出以“语义智能+行业/领域”的营销策略。公司深耕自然语言处理，致力于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，并积极与客户携手，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已相继打造出一批优秀实践案例。

（二）公司在 AIGC 领域有哪些发展规划？

回答：在发展路径上，拓尔思重点发展垂直行业的专业大模型，以“专业大模型+领域知识数据”为核，以 NLP 技术突破来推动更多 AIGC 商业落地。公司将基于通用 AIGC 大模型，专注优势行业进行专业大模型的研发，融合学习行业特有的大数据和知识，提升大模型对行业应用的适配性。公司也将以预训练大模型、In-Context Learning、instruction tuning 等技术为基础，将公司自主研发的文本生成、交互式生成、跨模态生成、小样本学习、大模型与外部知识库的融合等功能作为研发重点，突破基于大模型的 AIGC 关键技术，在问答式 AI、智能创作、搜索引擎等领域实现成功应用。

同时，公司将加快推进“智创”（一款专注文字生成类的内容自动生产平台）AIGC 平台的研发，该平台依托公司长期积累的自然语言处理技术和人工智能平台产品为基础，融合开源大模型，将专注在辅助型、应用型、创作型等文本内容的自动生成研发，以个性化、专业性的内容自动生成为壁垒，保持在 AI 领域的技术竞争力。公司将结合优势行业和客户群体，优先关注新闻出版、政务服务、金融领域、通用行业、元宇宙和云服务等 AIGC 细分市场，从降本增效走向额外价值转移。

（三）公司在数据要素领域有哪些布局？

回答：数据要素作为重要的生产要素，由数据采集、存储、流通交

易、分析应用等环节构成。拓尔思将重点关注数据整理、数据质量的把控，以及信创背景下数据要素如何释放价值。

公司正在积极推进自身数据采集、标注、存储、传输、管理、应用等全生命周期价值管理，加强公司在政府、媒体、金融、产业大数据等重点行业的数据应用模式探索，深化数据驱动的全流程应用。目前，公司基于深度学习的自训练模型和专业领域，打造了面向产业、媒体资讯、网络舆情三大领域的数星资产平台，不仅可以面向内部数星风控、海策政策大脑、冒烟指数、天目开源情报等云和数据服务产品进行数据赋能，也可为政府、媒体、金融、企业、安全等行业用户提供云和数据服务。

公司将积极推动数据产品与数据服务融合，以让数据创造价值作为重要的业务拓展方向，助力我国数字经济建设。

（四）公司云和数据服务的收费模式和业务拓展方式是怎样的？

回答：公司云和数据服务的签约均以年为单位，付款模式主要分为按季度、半年度、年度支付，按年付费的客户居多。产品的收费标准均根据数据量、用户数制定，比如公司在金融行业的产品“TRS 数星金融智能风控平台”的收费标准是根据客户监测发放贷款或投资并购的交易对手数量而制定。

公司云和数据服务的市场拓展主要以渠道代理的形式进行，一是将合作伙伴作为公司拓展客户的渠道，通过把公司的标准化平台产品嵌入到合作伙伴的产品中，进而提高客户数量；二是与垂直行业的专业服务商合作，由其进行渠道代理销售，进而拓展产品销售路径。目前已在金融领域有成功案例。

（五）公司推出的“海贝”是一款搜索型数据库，请问和百度、360这些常见的搜索引擎有什么不同？

回答：根据服务对象、搜索场景、搜索内容以及商业模式的不同，搜索引擎分为互联网搜索引擎、垂直搜索引擎、企业级搜索引擎。

互联网搜索引擎是从互联网提取各个网站的信息（以网页的文字为主），建立起数据库，并能检索与用户查询条件相匹配的记录，按一定的

排列顺序返回结果。垂直搜索引擎适用于有明确搜索意图情况下进行搜索。例如，用户购买机票、火车票、汽车票时，或想要浏览网络视频资源时，都可以直接选用行业内专用搜索引擎，以准确、迅速获得相关信息。这两种搜索引擎都是面向互联网 To C 的应用。企业级搜索引擎是 To B 或 To G 的，需要满足企业多样化的需求，特别是各种非结构化数据的搜索使用更是重中之重，常见的应用场景包括站内搜索、企业搜索、智能问答、数据分析、数据库加速、运维监控等。

众所周知的百度和 360 属于互联网搜索引擎，公司 TRS 海贝搜索数据库属于企业级搜索引擎，是构建搜索引擎应用的核心支撑软件。虽然两种类型搜索引擎的最终目标都是从庞大的知识库里取得用户需要的答案，但各自索引的内容差异化却非常大，应用场景和侧重的技术领域也不同。

企业级搜索针对的目标是企业内部各个应用系统中的文本文件和各类业务数据，如内网新闻、工作资讯、工作报告、专题研报、项目文档、知识库、培训音视频等网页和各类文件，还有 ERP 系统、CRM 系统、HR 系统中的业务数据等。这些数据存储结构、文件格式和来源等各不相同，不同企业的内部数据千差万别。企业级搜索的核心是提高内部员工的知识共享与工作效率，因此，对索引数据治理的高效、灵活性以及智能化，还有检索结果返回的精准性、个性化和高响应速度提出了更高的要求。企业级搜索还非常注重检索权限和访问权限的问题，需要继承各个业务系统的原子权限，对安全性要求非常高。在这方面常用的互联网搜索一般无法做到。

从产品特点上讲，公司 TRS 海贝搜索数据库是一款从内核到系统完全国产自研的搜索型数据库，早已融入了自研的 NLP 技术，如文本分词、自动分类、相似文本检索、拼音检索提示等，也集成了深度学习引擎，支持以文搜图、以图搜图、图文融合搜索，可以对图像中的文字进行 OCR 识别，也可以提取图像或者文本的特征数据，通过基因编码存储到海贝数据库里，成功实现图像相似性检索功能。而将搜索引擎技术与人工智能技术结合，对搜索结果进行智能整理与归类，并以精准化、智能化、

个性化的方式将搜索结果呈现给用户是“海贝”的产品优势。

（六）TRS 海贝搜索数据库目前有哪些应用案例？

回答：海贝已经广泛服务于安全大数据、媒体大数据、政府大数据以及军队大数据等众多细分行业，公安部、新华社、国家市场监督管理总局、海关总署、专利局、商标局等一系列重量级客户都选择了公司的产品。在国家层面，几个重大的基础数据库，例如云搜系统、企业信用信息公示系统、专利检索系统等都构建在海贝之上。

（七）在信创的大背景下，TRS 海贝搜索数据库是否存在国产替代的市场机遇？

回答：目前国内搜索引擎数据库市场为美国上市公司开源软件 ElasticSearch 一家独大的局面，且开源软件数据泄露和涉及国家信息安全事件频出。信创潮下，搜索型数据库坚持自主研发是国产替代的必由之路。

国内搜索引擎数据库厂商竞争方面，十几年前国内厂商圈也曾百花齐放，但在 ElasticSearch 开放源代码后大部分公司退出了该领域，目前国内坚持在搜索引擎数据库领域自主研发并保持技术领先的企业已凤毛麟角。拓尔思作为中文全文检索技术的创始者与引领者，在相关领域研发已有近 30 年的历史。从自主可控角度讲，海贝搜索型数据库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引擎数据库，吸收了公司在自然语言处理和信息检索方面的长期技术经验积累，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，完全满足信创要求和国产化替换需求。

（八）在海贝搜索数据库的国产替代方面，是否有成功案例？

回答：在国产替代方面，海贝已拥有众多国家标杆项目成功经验以及本土化服务能力，积累了一套成熟的、标准化的 ElasticSearch 平滑迁移方法，全程保障数据与业务的平滑迁移。以信用中国项目为例，该项目由国家发展改革委、人民银行指导，国家信息中心主办。原来在某云平台上由 ElasticSearch 对外提供检索服务，2018 年迁移至公司“TRS

	<p>海贝大数据管理系统”上，目前海贝在该项目中每秒需要支持大约 1.2 万的并发检索，日访问量接近 10 亿，获得了用户的高度认可。</p> <p>目前公司正在与行业有关部门积极沟通，同时客户群体中已产生国产替代的实际需求，公司已在积极对接。</p>
附有清单(如有)	无
日期	2023 年 2 月 21 日