



Transwarp Infinity 白皮书

星环无涯金融大语言模型

星环信息科技(上海)股份有限公司

Transwarp Infinity 星环无涯金融大语言模型白皮书

一、产品介绍

■ 产品背景

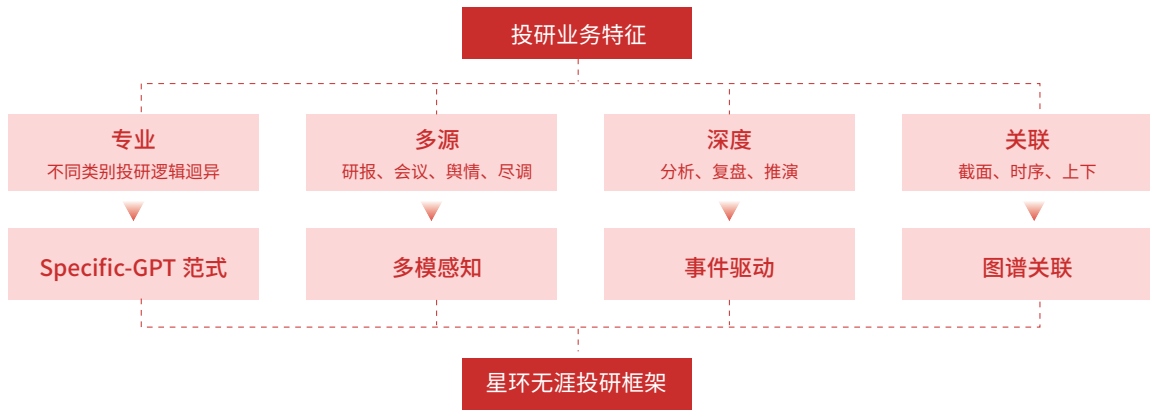
随着通用大语言模型（LLM）的正式落地问世，人类与数据的交互正在发生重大变革，数据获取、处理与分析已逐步朝向智能化、多模态和平民化趋势发展。在多模态大数据越来越备受关注的金融领域，大语言模型的落地实践更是当前行业的关注焦点。大模型的多模态涌现能力，是实现高质量泛化推理的基础，可为金融领域中投研溯源、量化投资、智能推理带来更多可能性，并可为整体行业的降本增效与科技创新提供显著成效。而在近年出台的各项金融科技政策与纲要，如由央行发布的《金融科技发展规划（2022-2025）》和党的二十大报告中，均将人工智能领域中多模态数据应用、自然语言处理与生成等技术放在重点发展方向中，更为推动金融行业挖掘创新实践带来确定性利好。

虽然大语言模型已在飞速发展，但通用大语言模型仍存在缺乏领域知识和知识推演能力等问题，在专业领域为企业提供服务仍有较大的能力缺陷，因而专业领域的大语言模型开发与实践已成为行业共识。星环科技作为企业级的大数据基础软件服务商，长期耕耘于数据服务与创新应用实践，在金融行业积累了丰富的业务经验与公开语料数据，同时不断开拓新型技术、研发了兼容多模态大数据的向量数据库，且通过为多家龙头金融机构服务积累了成熟的领域专业经验，为构建符合金融市场逻辑、贴合专业人士使用习惯的金融领域大模型与相应的成套解决方案奠定了夯实基础。



产品简介

星环科技无涯大模型 Transwarp Infinity，寓意学海无涯，既代表了投资领域终身学习的精神，也蕴含了大模型本身在参数架构方面持续迭代的内涵。针对量化投研领域特定的业务逻辑，星环科技通过预训、提示、增强、推导范式的构建，实现了 Financial-Specific-LLM 的训练；且基于大模型的事件驱动与深度图引擎，实现了对事件语义刻画、定价因子挖掘、时序编码、异构关系图卷积传播等功能，进一步丰富了应用场景、提高了应用上限；多种功能有机融合，最终构建成了可应用于金融行业量化投研的大模型星环无涯 Transwarp Infinity。



二、产品架构与解决方案

● 星环无涯整体框架

星环无涯的整体框架不仅包含了基于 prompt 进行预训微调好的金融大语言模型，且结合金融大模型分析与生成能力，通过搭建事件驱动引擎、叠加深度图计算、耦合星环量化投研框架，共同实现了对事件语义刻画、定价因子挖掘、时序编码、异构关系图卷积传播等功能，进而构建了包含事件冲击、时序变化、截面联动和决策博弈等多个维度的量化投研新范式。同时，结合 LangChain 模型，为开发人员提供更加多样的应用程序支持。





星环无涯构建了包括政策、舆情、ESG、风险、量价、产业链等六类大模型基础因子集，覆盖了固收、权益、商品等不同类别标的资产，擅长处理金融量化领域的各类问题；结合事件驱动引擎共同在政策和研报分析、新闻解读、事件总结和演绎推理上发挥强大的理解和生成能力；且能够对股票、债券、基金、商品等各类市场事件进行全面的复盘、传播和推演；还能够生成另类的策略因子集合，构建立体的归因解释体系。星环无涯大模型通过多模感知 + 事件驱动 + 深度图计算，从时间和空间、深度和广度等多个方面扩展投资研究的视角。



● 星环无涯：结合星环完整数据生命周期技术的金融行业解决方案

不同的投资领域、专家研究框架和投研习惯，都会造就业内客户对大语言模型的不同需求，因而一个可以更好地向全行业企业赋能的金融大语言模型不仅需要灵活的框架结构设计，还需要成体系化的技术解决方案。星环科技不仅自研了金融行业大语言模型，同时也提供了一站式的企业自建大语言模型工具链，该工具链不仅包含了与大语言模型应用落地紧密相连的向量数据库 Hippo，还包含了一系列针对各类数据存、算的分布式图数据库 StellarDB、时序数据库 TimeLyre、数据治理平台 TDS 等多款大数据、数据库产品，并耦合了可深度挖掘数据信息进行研、用、管的事件驱动引擎、深度图计算、投研管理等技术与框架功能，在星环无涯大模型的基础之上为不同企业的不同特需量身打造全套解决方案，助力企业迈向科技创新新征程。



三、产品核心优势

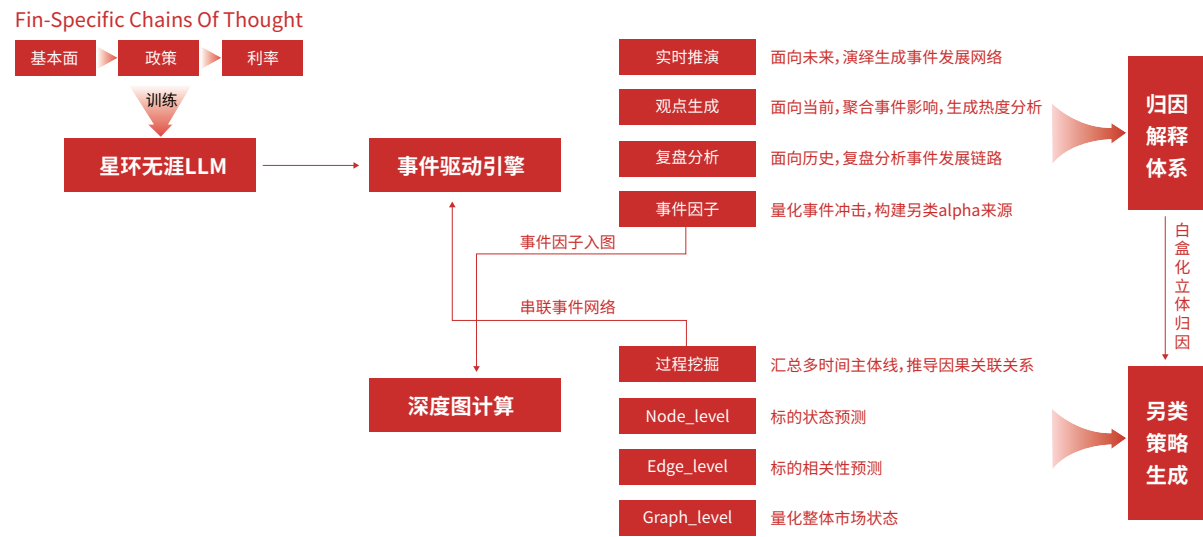
◆ 非结构化、非标准化信息转化为可溯因的标准化因子

基于强金融业务逻辑，构建六类大模型基础因子集、支撑复合因子策略体系，能够生成策略因子集合、构建立体的归因解释体系。



高精准、强逻辑的事理分析与推断力

有机耦合事件驱动引擎与深度图算法，实现事件复盘分析与推演，贯通宏观行业和大类资产分析逻辑。能够对股票、债券、基金、商品等各类市场事件进行全面的复盘、传播和推演。



专注于金融领域的大语言模型架构

针对金融行业，通过高质量金融语料训练、结合基于业务框架构建的事件驱动引擎与图计算功能，使星环科技的无涯大模型解决方案具备了准确理解和合理分析金融领域的专业能力。星环无涯擅长处理金融量化领域的各类问题，诸如在政策和研报分析、新闻解读、事件总结和演绎推理上都具备强大的理解和生成能力。



◆ 背靠大数据全生命周期技术栈的高灵活体系

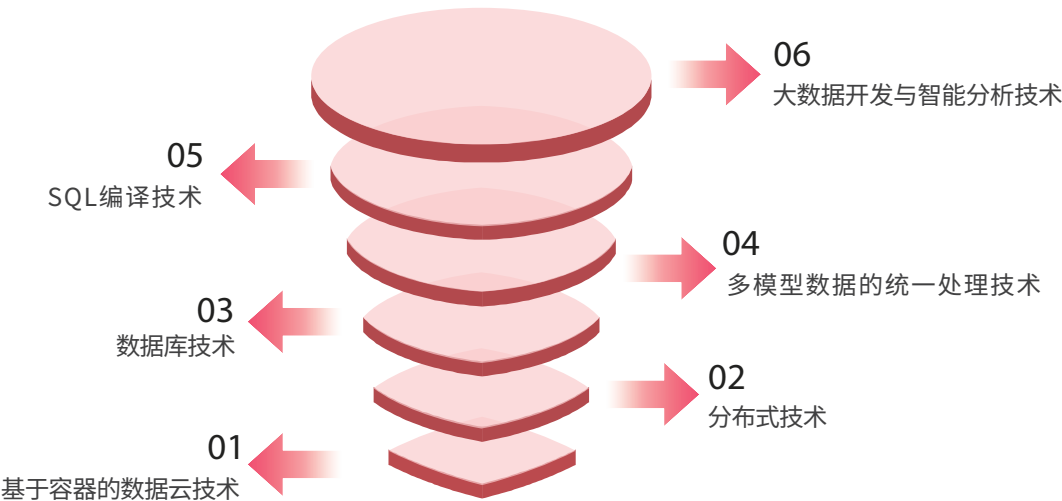
星环科技拥有多模态大数据全生命周期存算研用管的一系列技术积累，并不断基于整体框架进行创新性研发，可为不同需求的金融机构提供更加个性化、全面化的创新型服务与整套解决方案。



► 关于我们

星环科技(股票代码:688031)致力于打造企业级大数据基础软件,围绕数据的集成、存储、治理、建模、分析、挖掘和流通等数据全生命周期提供基础软件与服务,构建明日数据世界。公司以上海为总部,以北京、南京、广州、新加坡为区域总部,在郑州、成都、重庆、济南设有支持中心,同时在深圳、西安等地设有办事机构,并在加拿大设有海外分支机构。经过多年自主研发,星环科技建立了多个产品系列:一站式大数据基础平台TDH、分布式分析型数据库ArgoDB及交易型数据库KunDB、基于容器的智能数据云平台TDC、大数据开发工具TDS、智能分析工具Sophon和超融合大数据一体机TxData Appliance等,并拥有多项专利技术。目前公司产品已经在十几个行业应用落地,拥有超过一千家终端用户。2016年公司成为中国首个进入Gartner数据仓库及数据管理解决方案魔力象限的厂商,且被评为最具前瞻性的远见者;2017年被IDC评为中国大数据市场领导者;2018年星环科技成为12年来全球首个完成TPC-DS测试并通过官方审计的数据库厂商;2020年再次被IDC评为中国大数据管理平台领导者。自2021年起,星环科技蝉联Gartner增强分析技术中国推荐供应商;2022年,公司入选Gartner数据中台领域全球推荐供应商;同年6月,入选Gartner中国数据库管理系统产品品类最多的厂商之一。2022年10月,成功登陆上交所科创板。

► 核心技术



► 应用行业

产品已经在金融、政府、能源、交通、制造、公共安全、电信运营商、零售、媒体、教育、医疗等细分领域落地。

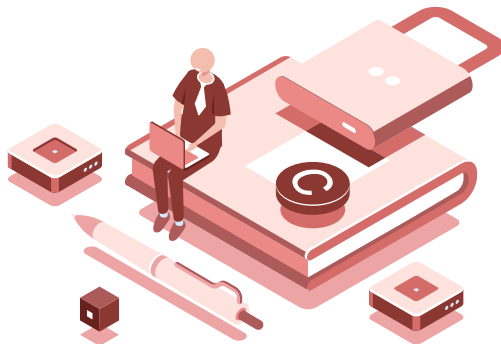
► 公司部分用户



► 产品废弃处理

本产品的废弃处理应根据国家的法律和规章进行





版权声明 © 2023 星环信息科技(上海)股份有限公司保留一切权利

任何单位或个人未经星环科技书面许可,不得擅自摘抄、复制本文件中的内容,不得以任何形式传播。

商标声明

本文件展示、提及或使用的所有商标归星环科技或者其他商标持有人所有。本文件内容不视为以明示、暗示、默许或者其他形式授予任何单位或个人商标使用权。未经星环科技书面许可,任何单位或个人不得以任何形式使用星环科技的商标或标记。

安全港声明

您购买的产品、服务或功能等受您与星环科技所签订的商业合同约束,本文件所描述的产品、服务或功能可能不在您购买或使用范围之内。由于产品版本升级或其他原因,本文件内容会不定期进行更新,对此不会另行通知。除非另有约定,本文件仅作参考、指导作用,所有陈述不构成对合同相对方的任何担保、承诺,不视为合同的组成部分或者附件,星环科技对此保留最终解释权。



TRANSWARP

- ☎ 电话: 021-60932577 北京分公司: 010-68278990
- 💻 网址: www.transwarp.io 社区: <https://community.transwarp.cn>
- @ 合作: mkt@transwarp.io 技术支持: service@transwarp.io 400-7676-098
- 📍 上海: 徐汇区虹漕路88号A座9F, B座11F&12F&15F
- 📍 北京: 海淀区复兴路69号院 华熙LIVE中心B座5层
- 📍 南京: 雨花台区宁双路19号 云密城J栋10层、建邺区贤坤路60号05幢10层01-06单元
- 📍 郑州: 郑东新区龙子湖崇德街17号 星联创科中心13层
- 📍 广州: 天河区林和西路161号中泰国际广场B座3601
- 📍 成都: 高新区益州大道北段555号 创新时代广场3号楼1701
- 📍 重庆: 重庆市大学城景阳路35号光谷智创园A座3楼
- 📍 济南: 历城区经十东路28666号济南超算主楼5楼538-542
- 📍 深圳: 深圳市福田区深南大道6031号杭钢富春大厦1523



公众号



服务号



B站



视频号

版本 V1.0

本手册中的图形或图片, 主要辅助描述本公司的产品功能和特点, 本公司保留最终解释权。