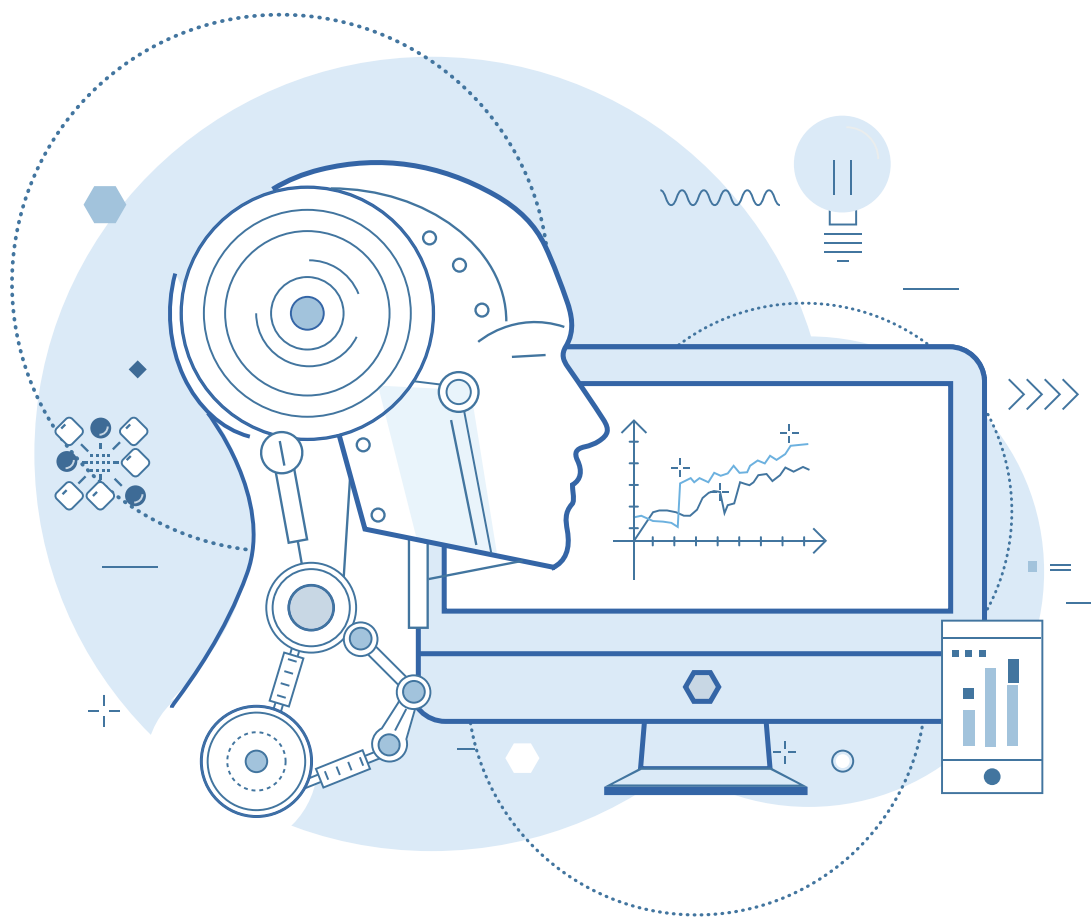


Transwarp Sophon

智能分析工具白皮书



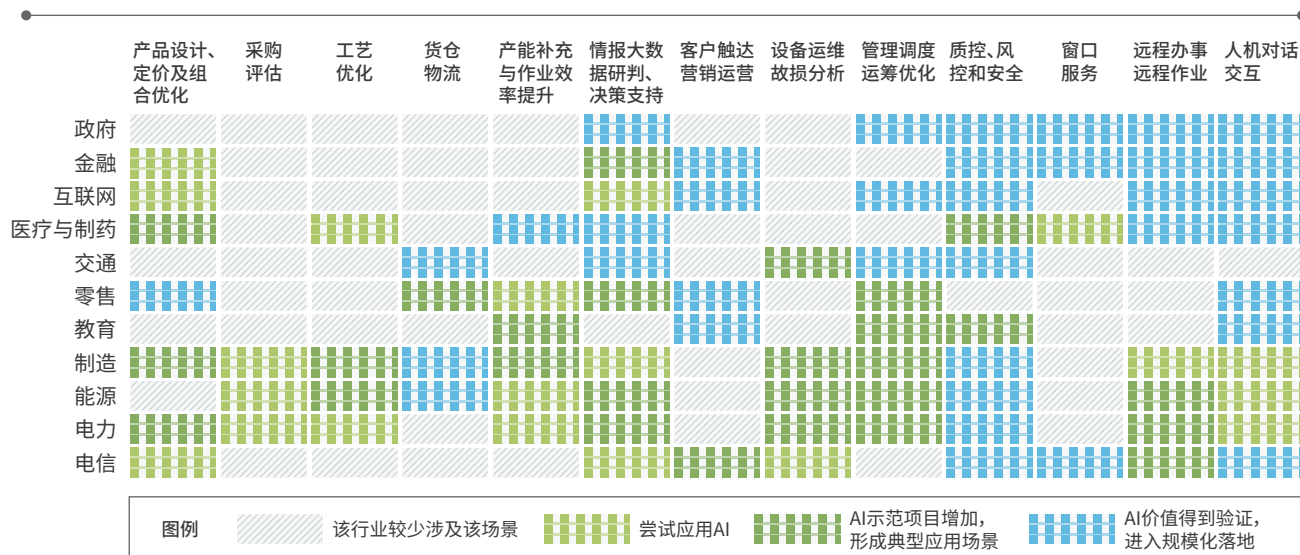
目录

DIRECTORY

1	产品背景 PRODUCT BACKGROUND	02
2	产品架构 PRODUCT ARCHITECTURE	03
3	产品特点 PRODUCT CHARACTERISTICS	04
4	Sophon Base 数据科学平台 DATA SCIENCE PLATFORM	05
5	Sophon KG 知识图谱平台 KNOWLEDGE GRAPH PLATFORM	09
6	Sophon Edge 边缘计算平台 EDGE COMPUTING PLATFORM	13
7	Sophon P²C 隐私计算平台 PRIVACY-PRESERVING COMPUTING PLATFORM	17

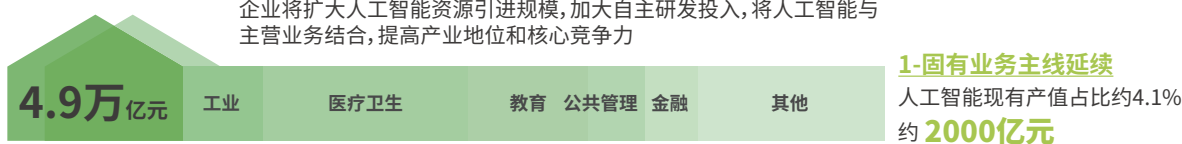


在“十四五”规划和2035年远景目标纲要提出的“打造数字经济新优势”建设方针及“加快推动数字产业化，大力推进产业数字化转型”发展主线影响下，各行业转型和升级的过程迸发出的智能化需求也逐步增多。人工智能作为数字经济新时代的引擎，能够在各行业企业设计、生产、管理、营销、销售多个环节中赋能其智能化需求。而据经弗若斯特沙利文公司整理的《中国全栈 AI 市场研究报告》显示“购买人工智能服务是传统企业应用人工智能技术的重要途径”。

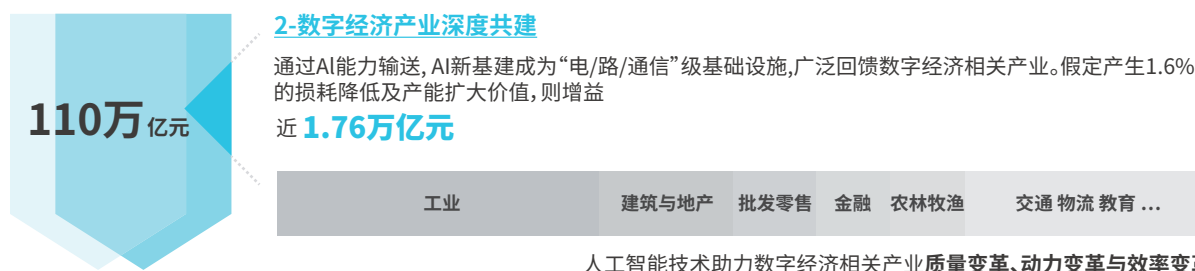


人工智能技术广泛渗透进经济生产活动主要环节

企业将扩大人工智能资源引进规模, 加大自主研发投入, 将人工智能与主营业务结合, 提高产业地位和核心竞争力



2021年中国各行业数字化、信息化与专业技术服务投入+Top 投入行业



2021年中国GDP+Top贡献行业

人工智能成为数字经济时代核心生产力

图源: 艾瑞咨询研究院2021年中国人工智能产业研究报告 (IV)

而星环科技则是通过Sophon为各企业用户提供其行业应用场景所需的人工智能服务。Sophon是一款包含数据分析与机器学习建模的一系列智能分析工具软件，能够一体化地完成从数据集成、模型构建、知识推理、辅助决策，支撑业务的数据分析、探索、服务流程，从而可以帮助开发者将智能化落地到企业应用中，帮助企业进行以数据驱动为核心的开发、运营和产业升级。

依托于底层大数据平台，Sophon主要覆盖数据分析中的计算智能（读取、计算）、感知智能（看、读、认）、认知智能（理解、认知、思考、推理）、安全智能（AI+安全——可信隐私计算）四个主要方向。同时其采用分布式计算模式，可以进行表格等结构化数据、标注文件等半结构化数据和图像文本等非结构化数据的接入、处理、分析与模型发布及管理工作。

○ 计算智能-Sophon Base数据科学平台

一款一站式数据科学平台工具，包括数据探索管理、可视化模型开发、编程式模型开发和AI模型运营等核心组件，帮助客户实现智能数据分析的全生命周期管理工作。

○ 认知智能-Sophon KG知识图谱平台

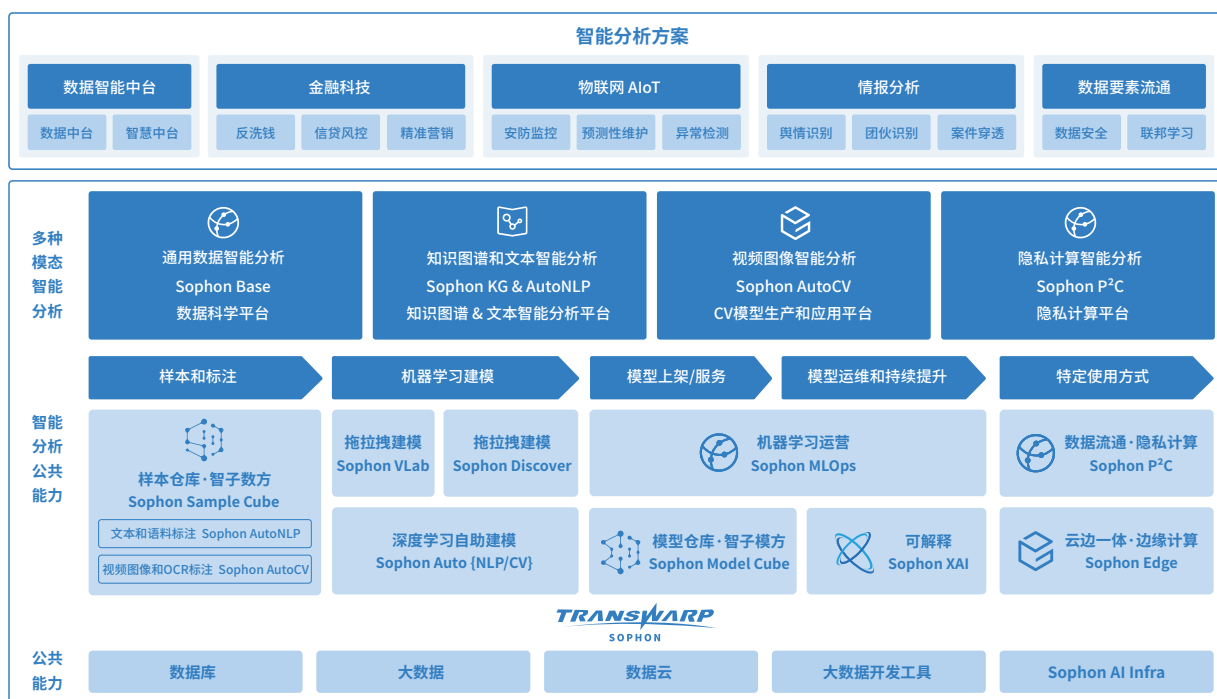
一站式知识生命周期管理平台工具，集知识的获取、融合、建模、存储、计算以及应用为一体。

○ 感知智能-Sophon Edge边缘计算平台

一款智能应用的构建平台工具，提供全面的设备连接、快速的模型异构部署和稳定安全的云边协同能力。

○ 安全智能——Sophon P²C隐私计算平台

一款数据的可信安全计算平台工具，基于密码学、联邦学习、多方安全计算和安全通信架构，在保障数据隐私安全的前提下，完成跨组织的数据协作。



Sophon作为数据要素的智能底座，可实现从数据到智能的全链路分析，具有以下八大特点：

01

覆盖数据分析建模全流程

支持一站式可视化流程，从数据接入、处理、标注，到模型训练、服务部署、线上监控，让整个流程更加高效、简单易用。此外，还支持容器化部署，弹性资源管理，提供跨平台、边缘设备等模型部署能力，以满足用户的不同需求。

02

提供数百种分布式机器学习算法并支持前沿的智能分析技术

平台提供超过200种分布式机器学习算法，并支持前沿的智能分析技术，包括高性能时间序列分析SDK、高性能深度图引擎ZenGraph以及知识图谱模块。这些工具可用于从多个来源的数据中提取信息，并推理出复杂问题的解决方案。

03

支持多种机器学习和深度学习框架

平台支持多种机器学习和深度学习框架，例如Spark、LightGBM、XGBoost、Tensorflow、Torch、Python和R等，能够满足不同的数据科学需求，并提供高性能深度学习推理框架DLIE，从数据处理、模型训练到模型推断等全流程都有覆盖。

04

提供具备“六个统一”的机器学习运维能力

提供机器学习模型运维（MLOps）的“六个统一”能力：统一纳管、统一运维、统一应用、统一监控、统一评估和统一解释。

05

提供分析即服务的能力

支持私有云、公有云和混合云部署模式下分析即服务的能力。

06

提供隐私计算技术为核心的数据要素流通平台

保障大数据交换时的信息安全、保护终端数据和个人数据隐私，在多参与方或多计算节点之间开展高效率的数据建模分析。

07

提供边缘计算推理能力

实现边缘端视频和物联网标准协议接入互通;可使用低代码方式定义视频处理、机器学习/深度学习模型分析、处理规则等;支持云端、边缘端统一管理部署;通过算法市场,提供面向如安全生产、工业质检、轨交运营和OCR图文识别等垂直场景的基准算法。

08

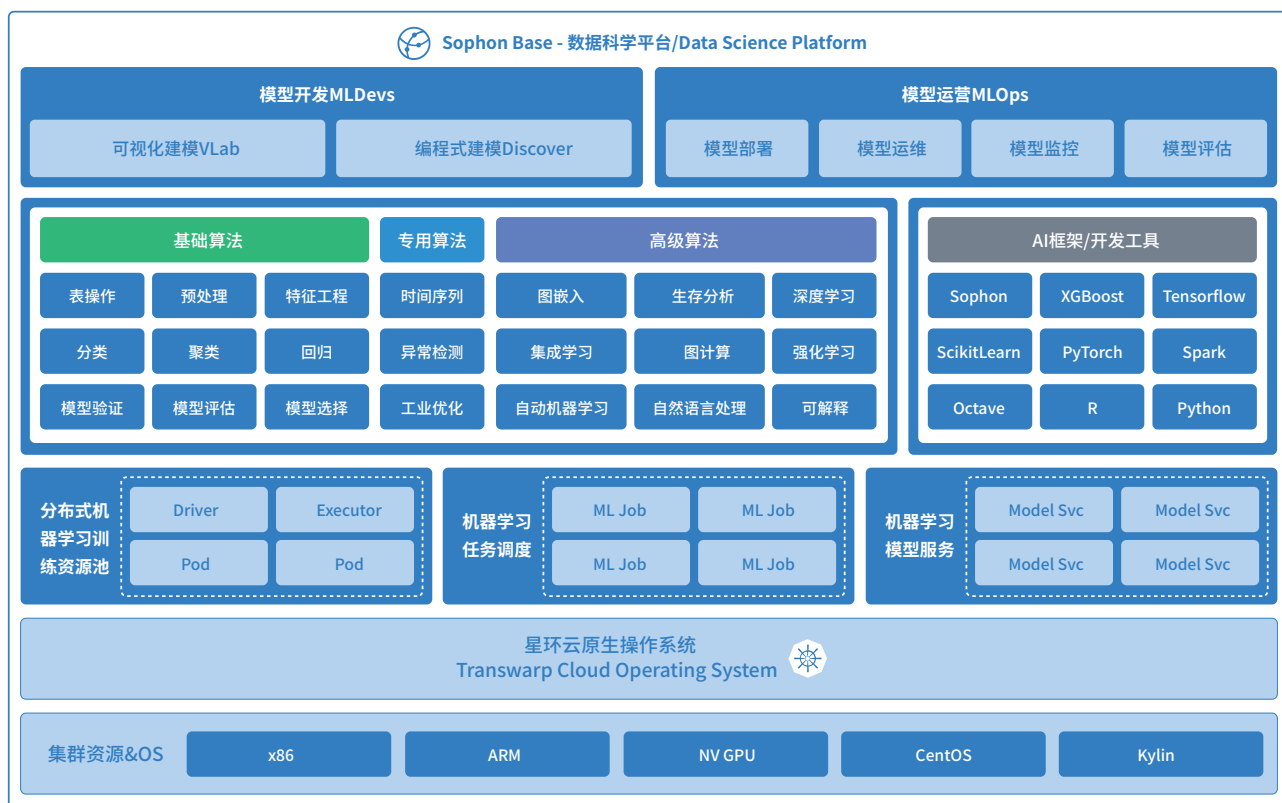
提供“三仓”实现以数据为中心的智能分析能力

提供样本仓、模型仓、场景仓:样本仓能够对多模型数据进行纳管、智能标注和探索分析,从而形成高质量的样本数据;模型仓则提供方便的训练、组装和评估,形成高质量的模型数据;场景仓能够规模化部署应用场景模版,并通过应用数据反馈支持高效率的业务迭代。

4 Sophon Base数据科学平台

平台介绍

Sophon Base是具备统计、机器学习、深度学习等完备算法和丰富算子的企业级分布式机器学习平台,支持从数据接入、数据处理,到模型训练、服务部署、线上监控的一站式可视化流程。



Sophon Base的平台优势可总结为以下四点：

○ 计算

提供覆盖数据分析全流程的可视化建模和推荐式建模服务，让用户无需编写代码即可完成机器学习建模，有效地降低了使用门槛；同时提供了支持多种编程语言和交互式代码输出的程式化建模服务，极简的操作界面、pipeline式的任务管理和灵活的框架支持帮助用户实现高效建模。

○ 运管

通过统一纳管、统一运维、统一应用、统一监控，赋予企业客户易用、高效且安全可靠的AI能力运营服务，协助客户规模化管理日益增长的机器学习模型，提升模型使用效率，降低模型集成管理成本，控制模型生产环境风险。

○ 可信

使用多种解释算法，可输出各特征参数对于模型预测结果的影响权重，如Kernel SHAP、LIME、LRP、DeepLift等，从而大大增强用户对AI模型的信任；另外在保证合规的同时，Sophon Base提供的可信AI服务还能推动优化模型，为迭代提供参考。

○ 安全

Sophon Base结合隐私计算和加密通信服务，可为多方安全建模提供完整的解决方案，解决跨组织协作时无法安全利用各方数据的困境。同时可提供多种联邦学习算法适用于各类垂直业务场景，为跨企业AI协作提供安全可靠的平台支持。

主要功能

01 //

多源数据接入及探索



可支持接入JDBC、Hive、HBase、HDFS及星环自研数据存储等多种数据源，也可直接接入本地CSV、TXT、PDF等数据文件；同时对系统接入数据，支持按统一视图及规范进行数据探索及数据清洗，同时提供文件管理系统等多种功能。

02 //

低门槛的可视化建模



具备完整的数据探索、多数据源接入、实验调度、智能分析、用户资产以及平台管理等功能，用户可以快速完成从数据读取、ETL、特征工程、模型训练、模型应用到模型评估的全流程拖拽式建模的强大能力，无需编写代码即可完成建模。同时，还提供了多场景的实验加工模板，协助用户快速构建特定场景的解决方案。

03 //

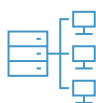
简单高效的编程建模



支持多种编程语言和交互式代码输出，提供灵活可复用的自定义开发环境，便于在有相似建模需求的场景中快速复用，后续也能更好地进行模型上架上线；构建为自动化运行的pipeline任务，避免训练过程中出现因实例资源关闭，导致任务中断的情况，保障模型训练的稳定性。

04 //

统一的Dashboard展示



支持统一查看和管理系统内建模项目、项目中的建模实验、模型文件和模型镜像等的数量统计，项目运行状态随时间变化情况，以及各集群节点上CPU/内存的资源利用率变化情况。同时，支持查看系统内用户的操作审计信息，若遇到功能使用问题，还可迅速跳转至对应模块的帮助手册。

05 //

规模化的集成管理



统一纳管多源异构的模型文件，通过规范化集成管理，实现统一部署大规模机器学习模型，从而帮助用户大幅降低模型管理成本。

06 //

高效模型推理服务



使用图形交互的方式，帮助用户流程化快速构建服务推理图，并支持自定义配置包含多模型的复杂推理逻辑，及可视化快速部署模型应用，简化了模型配置和部署的复杂流程。

07 //

多维度监控及评估



实时监控模型服务的运行状态，帮助用户控制模型生产环境风险；通过多维度评估模型预测效果，为用户后续迭代优化模型提供重要参考。

08 //

联邦任务与建模支持



支持自定义联邦任务及角色，并支持让联邦任务的参与方根据任务向导，基于加密协议服务以及内置的多种联邦特征工程及联邦模型算子完成联邦任务的数据预处理、特征工程、模型参数配置等流程；同时在建模前后可分别从数据质量评估、模型贡献评估两方面衡量参与方的数据质量以及模型贡献度，为用户筛选联邦任务参与方及量化模型收益提供了有力的数据支撑。



应用案例

银行AI平台建设解决方案

模型统一集成 与管理

使用Sophon Base搭建全行统一的AI平台，快速接入行内积累的由Sophon Base或其他框架/平台训练生成的大量模型文件，解决了异构模型的管理问题，集成了全行多种算法框架生成的数百个机器学习模型。

模型统一运维 与应用

使用Base中的标准化流程，统一构建模型推理逻辑，一键部署模型应用，并实现对模型应用的统一运维，让部署模型应用的平均时间由1.5天降至0.5小时，同时配置成本降低近80%。

模型统一监控 与评估

持续监控所有已上线的模型应用，并评估模型预测性能。最终同时在线上百个模型预测服务，单条数据实现毫秒级响应；同时打通了模型全生命周期流程，使模型的平均迭代周期由1月降至1周。

期货市场AI智能解决方案

基于机器学习 进行风控建模

基于Sophon Base的协作开发，采用专家规则与机器学习相结合的方式的规则发现，搭建风控措施辅助决策系统；实现了专家规则+AI模型的双轨制决策模式，增强了AI模型在业务场景中的应用可解释性。

强化学习辅助 异常交易识别

运用Sophon Base构建逆向强化学习模型，采用对敲交易识别分析、场外配资账户识别分析和异常交易指标计算等手段，结合市场行情分析交易特征，进而识别异常交易模式，提前预测市场风险。

○ 智能配用电分析解决方案

居民用电行为分析

根据居民长期用电数据，利用Sophon Base对居民进行聚类分析，并将居民按照用电行为分为控制用户、季节敏感用户、周期性用户等不同类型，深入挖掘居民用电特征，准确分析了用户用电分析规律，实现了异常用电行为检测。

配电台区降损分析

基于配电台区供电量、峰谷比、功率因数等特征，对某一行政区域所有台区进行多维聚类，并基于随机森林算法建立线损分析模型，估计所有台区合理线损，帮助电力公司实现了台区精细化管理。

用电预测应用分析

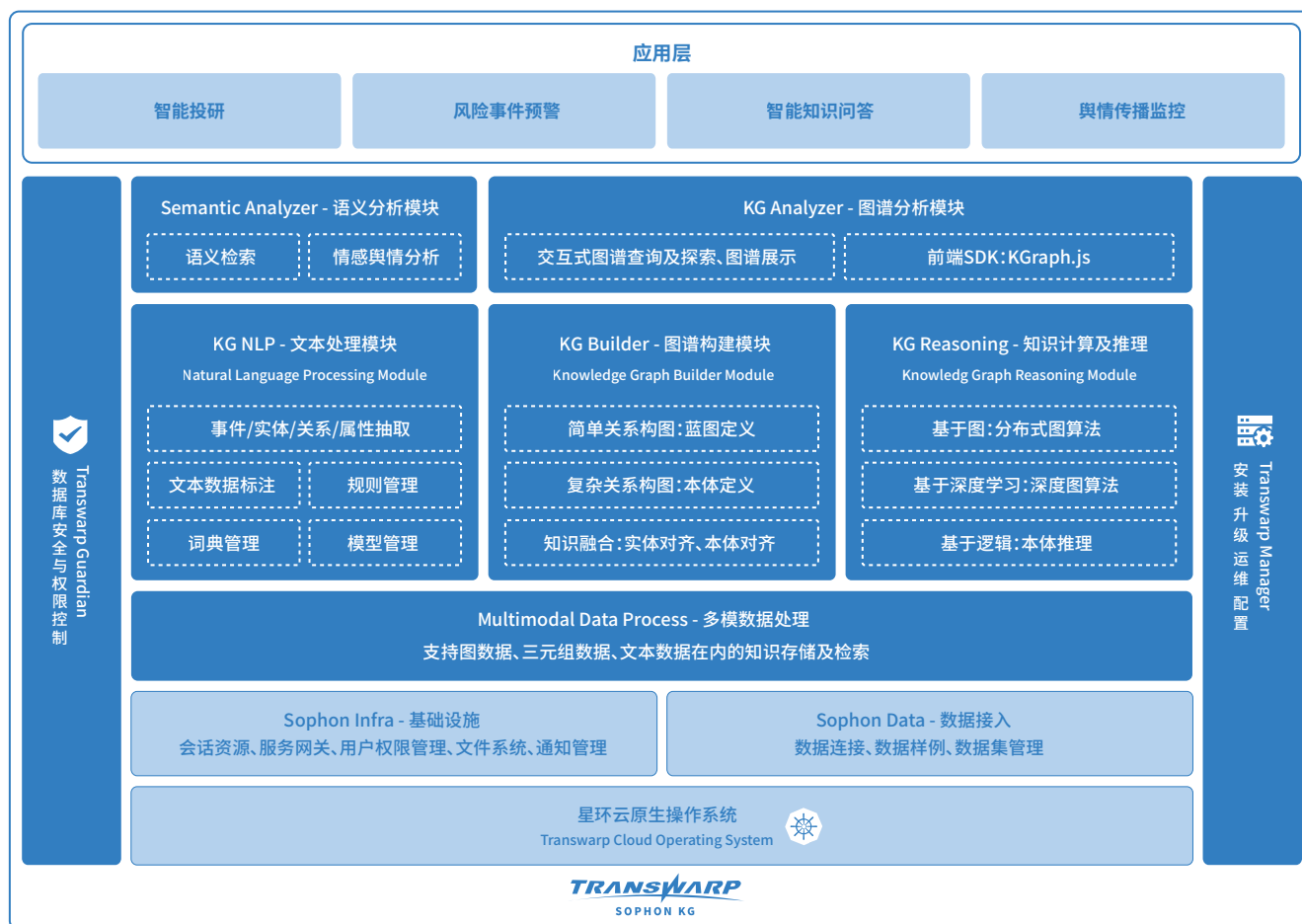
综合考虑气温、湿度、气压等因素，建立样本本次适应的负荷预测模型和聚类预测模型，实现了对台区每日用电的最大/最小值、以及未来七天用电负荷曲线的较为准确预测。

5 Sophon KG知识图谱平台

平台介绍

星环知识图谱软件（Sophon KG）是一款覆盖知识全生命周期，集知识的建模、抽取、融合、存储、计算、推理以及应用为一体的知识图谱产品。平台支持低代码图谱构建、智能化知识抽取、多模态知识存储与融合、多形式知识计算和推理以及多维度的图谱分析。除了具备上述的链路完备性，平台还从业务场景出发，沉淀了几个场景的图数据模型、规则模型和算法模型，可以帮助客户快速解决同场景下的业务问题。





Sophon KG的平台优势可总结为以下六点：

○ 链路完备性

从图谱定义、知识抽取、知识融合、知识存储、知识推理到知识应用和发布，Sophon KG对于知识的全生命周期管理进行了全链路的支撑与管理。

○ 多模型数据对接

Sophon KG支持多源、多方式、多类型的数据接入，支持对接星环分布式图数据库StellarDB和开源图数据库Neo4j、分布式本体存储推理框架SANSa、支持对接开源图查询语言openCypher和本体查询语言SPARQL，以实现基于多种模型数据的分析与检索。

○ 全流程可视化

无论是简单图谱构建还是复杂图谱构建，用户均只需通过简单的交互式配置即可完成。解决行业知识图谱构建过程中更新维护成本高，操作复杂繁琐等效率性不高的痛点。

○ 多模态数据融合

通过NLP技术和图技术的结合，实现结构化、半结构化、非结构化数据的半自动化融合构图。解决行业知识图谱构建过程中的数据多样性的痛点。

深度图计算引擎

Sophon KG提供了一套自研的深度图计算引擎ZenGraph，可以帮助用户从图中挖掘出更多有价值的信息。在技术上，ZenGraph支持和多种数据进行通信，有效降低了图上分布式计算额外的数据传输消耗。

NLP能力迁移泛化

平台即内置了金融场景的NLP规则模型、词典、算法模型，也支持用户自定义上传模型、字典和规则，并支持一键和NLP任务绑定，以帮助客户快速解决同场景下的业务问题。此外，用户还可以通过利用平台的预训练模块，标注少量数据即可训练出一个高精度的抽取模型，以实现NLP模型的规模化应用。

多维度图谱分析

Sophon KG不仅支持所见即所得的交互式图谱查询与分析，而且支持基于自然语言的检索与推荐、智能问答、情感分析，还支持基于图算法的模式发现功能，上述功能可辅助用户从多维度理解大数据，并为达到业务目标提供决策支持。

主要功能

简单图谱构建

Sophon KG蓝图定义模块的产品设计能让用户通过拖拉拽结构化数据表，直接完成实体、关系及属性的抽取，暨通过图模式和数据映射的一体化设计，解决图模式在数据出现之后才能确定的痛点。同时，对特定场景，系统内置了蓝图和蓝图模板，用户只需将业务数据与蓝图模板做映射即可生成图谱。

复杂图谱构建

针对概念层级多、概念之间的关系复杂的图谱，一般要花费的时间成本是非常高的。而使用Sophon KG的本体定义模块，用户通过交互式配置即可完成概念层级以及实体和关系属性的继承定义。平台支持知识库的可视化展示，能够使用户一目了然的获得知识结构以及知识之间的复杂关系。

智能知识抽取

平台内置金融场景的知识抽取模型，可以实现事件、实体、关系、属性的自动化抽取。针对其他业务场景，用户也可以上传自己训练好的抽取模型，只要将模型和对应的抽取任务绑定即可实现知识的自动化抽取。此外，平台支持用户将抽取结果与已有的图谱进行融合，实现非结构化数据的半自动化构图。

知识计算与推理

平台既支持基于逻辑或者规则的演绎推理也支持基于图和深度学习的归纳推理。

智能语义检索

平台内置金融保险业务场景的智能问答能力。用户可以不改变业务使用习惯，在平台上直接输入关键词或者语句，平台就能识别出用户意图，给出用户所关心的答案，并以概念-实体-关系构成的图谱形式进行结果呈现。



应用案例

保险智能问答解决方案

支持将产品、类型、条款等非结构化关系网数据通过图谱建模、知识抽取和实体融合等技术，构建成一个专业知识库。在知识库的基础上，实现产品条款等问题的自动语义检索、问答。

01

通过自然语言处理技术，理解人类真实意图，结合QA问答对、知识图谱等能力提供一个面向内部员工、保险代理人的保险智能问答客服。有效减少了人工客服的接入，同时最终问答准确率达到95%，也使获客效率提升了50%。

02

交易反欺诈解决方案

建立商户交易分析图谱，运用商户群体的聚类、分类模型，基于商户群的行为数据结合业务规则进行分析并挖掘出存在信用卡套利套现行为的商户个体及商户团伙。

01

基于商户的知识图谱，运用LPA社区聚类算法，配置商户间关联的共同消费者与历史交易数据这两种边的权重，挖掘出存在异常交易的商户个体及商户团伙。最终实现了输出异常名单的TOP 50%覆盖到当月上报案例的95%。

02

舆情传播分析解决方案

通过星环提供的Sophon KG+StellarDB解决方案，可以从海量新闻资讯文本中完成实体-关系-事件的智能抽取，并支持主流的图查询语言openCypher。

01

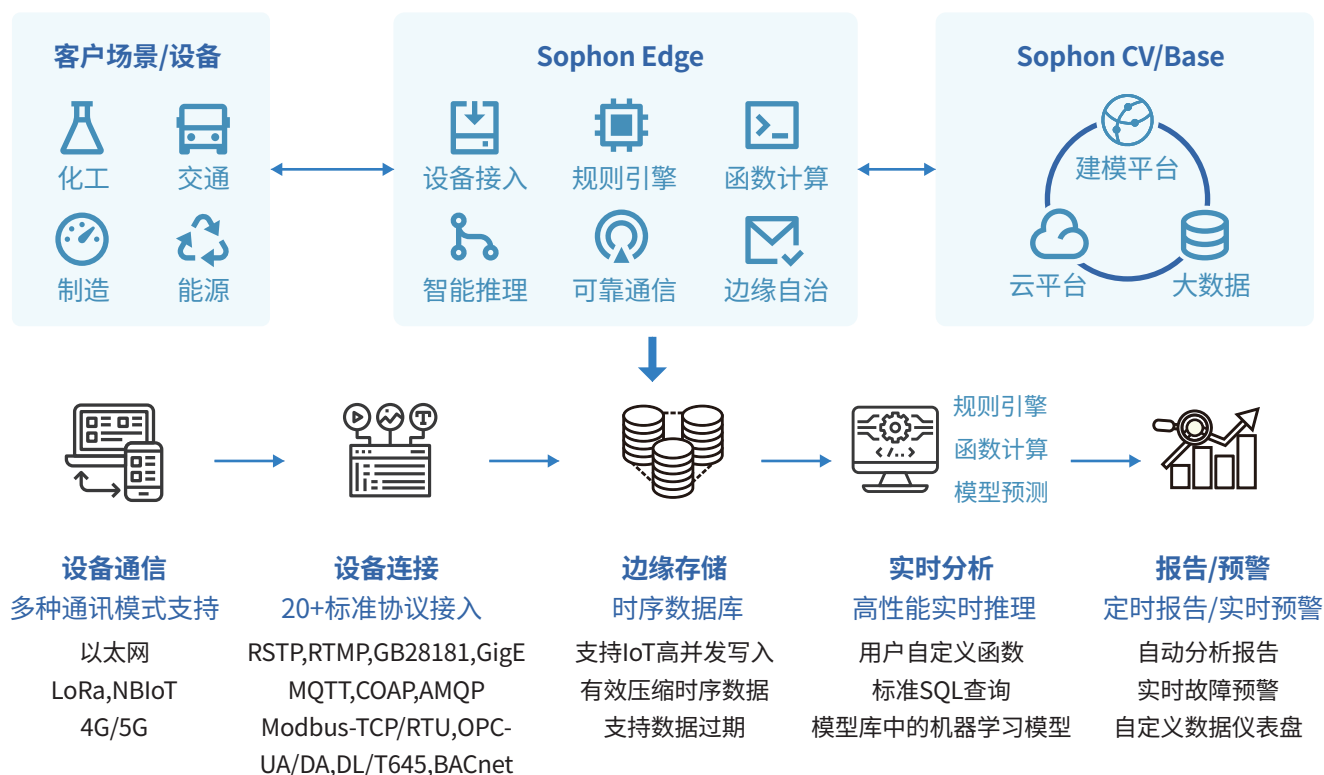
以指定事件为中心提取相关舆情信息，融合星环自研舆情传播算法并基于图结构完成计算。最终训练得到的舆情传播算法模型精度达0.84，召回率达0.63，帮助用户实现了高效的舆情传导可视化分析，最终获得的舆情传播数值也能辅助用户业务决策输出。

02

6 Sophon Edge边缘计算平台

I 平台介绍

星环边缘计算平台Sophon Edge，作为一款边缘AI应用构建平台，结合CV模型训练模块和MW模型仓库模块，提供了一整套完整的视觉模型训练、应用部署和管理能力，帮助企业有效提升模型开发部署效率，同时实现过程数据资产的积累。



Sophon Edge的平台优势可总结为以下五点：

○ 高效率：一站式模型训练、部署和应用平台

平台覆盖了从样本管理、数据标注、模型训练到模型部署应用的全流程模块，能够基本满足用户在CV模型落地过程中的各种需求。流程紧凑，具有较强的碎片化场景适应能力，能够实现任意场景的快速建模和部署应用，很好地解决建模流程效率低的问题。

○ 零门槛:引导式训练模型、低代码构建应用

支持全程可视化引导式、零代码模型训练，结合预置的业内优秀算法及预训练模型，无需编程即可训练出高精度的CV模型。在模型部署方面，能够基于平台预置的标准化算子通过可视化拖拉拽的方式，快速搭建业务应用规则。

○ 高性能:高性能模型推理和硬件适配

在现场部署的模型，使用了推理加速工具对模型进行推理优化，以实现模型高性能推理。此外平台针对多种异构硬件进行推理加速，可支持国内外多款主流加速卡。

○ 易管理:过程数据资产积累和管理

平台支持数据、模型和应用的版本管理，以应对使用过程中业务需求的动态变更。对于过程中的数据、模型或者相关应用的过程资产，能够通过系统有效的沉淀和积累。

○ 很安全:用户数据安全得以保障

在保障数据安全方面，通过设样本水印防止标注人员泄露数据，全面记录审计日志防范恶意操作，生成的模型文件加密处理保证模型安全。从数据通信安全、模型文件安全、业务代码安全等不同维度，保障了业务在边缘侧能够稳定且安全的运行。

■ 主要功能

○ 数据管理模块

1. 文件系统：样本管理用于对用户上传的原始图像数据进行存储和管理，支持用户自定义文件存储路径，支持样本集的预览，支持样本集增删改查操作，支持样本集的导出功能。支持应用落地后实际业务回流数据的接入。
2. 数据标注：支持图片分类、目标检测和图像分割等常见图像类标注任务。提供模型智能预标注和协同标注功能，大大提升了标注效率。
3. 数据共享：支持平台内样本数据和标注数据的多账户共享，同时支持通过导出样本数据和标注数据的方式在平台外进行数据共享。

○ 模型生产模块

1. 作业管理：支持对训练作业进行管理，支持训练作业的增删查操作。支持查看训练作业详细信息，包括模型的训练数据信息、训练参数信息等。
2. 引导式建模：针对分类和目标检测任务均预置了业内优秀算法框架及预训练模型，同时提供引导式模型训练服务，帮助用户快速训练好模型。
3. 分布式训练：模型训练支持分布式训练，模型训练效率高。且支持GPU资源可视化，用户能够手动选择训练资源进行作业训练，从而实现资源弹性分配。

应用部署模块

1. 设备管理：平台支持如Onvif、GB28181、RTSP、RTMP、UVC等超过20种的标准协议，用户只需要进行简单配置便可快速将物联网设备或视频设备接入平台，并进行设备数据实时预览和统一管理。
2. 应用构建：平台提供拖拽的交互方式辅助用户快速搭建智能化场景方案，同时平台提供了面向如强监管、工地安防、电网运维和市政监管等不同行业领域的标准场景模板，可在相似场景下快速复制落地，能够有效降低用户开发成本，快速实现业务流程的对接。
3. 云边协同：平台能够实现云端对边缘节点或集群的统一管理，实现云边数据协同、状态协同、控制协同，极大降低了边缘端的管理以及运维成本。

模型持续管理模块

1. 模型管理：支持文件模型、镜像模型和DLIE模型（星环自研深度推理引擎，支持多种推理加速硬件）的一键上架，以及多版本管理。支持模型加密保护，支持模型一键部署到服务器进行落地使用。
2. 模型评估：基于与模型关联独立测试集，对模型进行一键测试评估，对评估指标进行可视化展示，并支持版本间的评估指标对比，帮助用户快速捕捉版本间性能变化。
3. 持续迭代：针对上架模型，用户可设置模型持续迭代策略，对模型的线上请求数据进行抽样，扩充至指定的样本集，为模型线上数据分析以及下一轮模型迭代做好数据积累。

应用案例

生物生长观测解决方案

使用Sophon Edge平台的模型生产模块，能够训练生物识别相关的模型。所得到的生物识别模型测试效果良好，平均精度可达96%以上。

01

使用平台中部署相关的模块对模型进行部署应用，在模型应用过程中，平台会对实时处理的数据进行监测，对有价值的内容进行回传。业务人员可自行判断回传的数据信息，从中选取可用于迭代训练的数据，从而实现模型的优化更新,最终使得模型在不同情况下都能有优秀的识别表现。

02



○ 炼钢生产优化解决方案

通过Edge AIoT平台能力，对接炉炼钢的生产过程数据，通过数据预处理，特征工程，尝试不同建模方案，解决客户吹炼过程中的实际问题，保证出钢的质量，提高炼钢的生产效率，降低成本，节能减排。

01

通过平台提升了原始数据的治理能力，通过模型沉淀了转炉炼钢的工业生产模型，找出了吹炼过程中供氧和造渣操作之间最佳的协调方式，做到了标准化的炼钢生产，提高了出钢终点温度和碳含量的命中率。

02

○ 地铁火花检测解决方案

使用Sophon Edge对地铁线路运营视频进行图像识别，将图像中的火花信息提取成结构化的时序信息。

01

通过Sophon Edge训练得到的时序智能分析模型，根据火花发生频次和位置来评价地铁线路健康状况，并结合现有的数据大盘对多条线路的健康状况在Sophon Edge提供的仪表盘中进行综合展示。

02

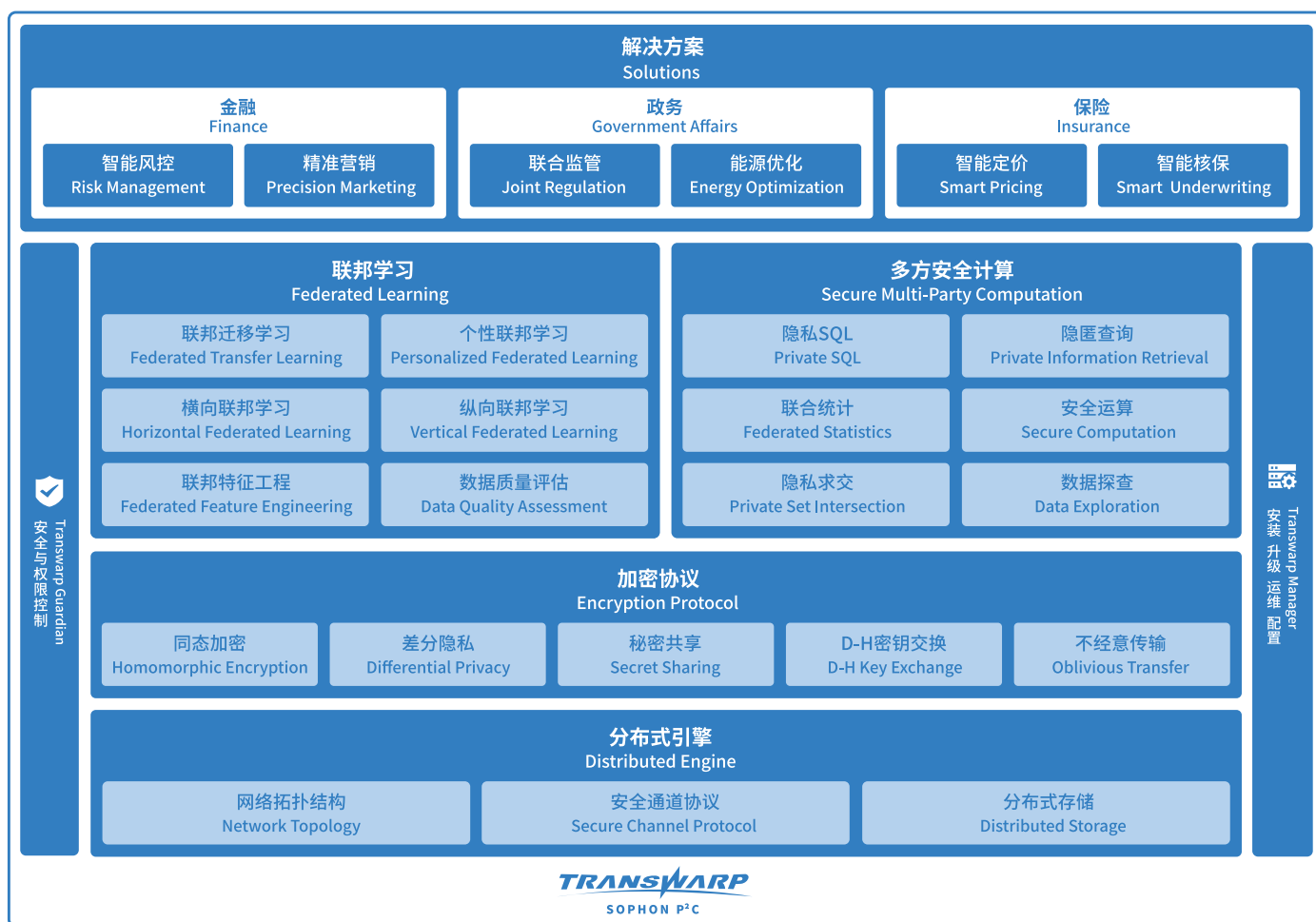
通过对整个线路打火的次数及火花大小进行聚类分析，为接触网检修提供了指导依据，通过视频识别和及时告警预防了设备故障的发生，相较于传统依靠人工监测的方法显著节约了人力成本，建立起了一套完整的燃弧质量健康评价体系。

03

7 Sophon P²C隐私计算平台

I 平台介绍

Sophon P²C作为一款企业级隐私计算平台，在保证明文数据不出本地的前提下，支持多参与方或多计算节点之间开展高效率的数据分析、建模和模型发布，达到数据不动模型动、数据可用不可见的效果。Sophon P²C支持不同场景的隐私计算需求，包括横纵向联邦学习、多方安全计算、基于差分隐私的数据发布、匿踪查询等，可为多方数据安全协作提供完整的平台底座。



Sophon P²C的平台优势可总结为以下四点：

○ 支持多种隐私计算框架, 平台易用易部署

1. 采用同态加密、差分隐私、秘密分享、不经意传输等隐私技术，覆盖联邦学习（FL）、多方安全计算（MPC）、匿踪查询（PIR）、隐私求交（PSI）等多种隐私计算功能。
2. 支持大数据规模的隐私计算场景，支持亿级数据进行联邦学习、多方安全计算和隐私求交。
3. 提供页面可视化安装部署，并支持实体部署、容器部署、云上部署等多种方式，轻松完成平台上线。
4. 低代码可视化隐私计算平台，提供算子托拉拽式本地数据清洗与特征工程能力，实现高效数据预处理。
5. 多行业AI落地经验，可提供专家级应用服务，如联邦风控、联邦反欺诈、联邦推荐等个性化AI场景建模服务，助力企业降本增效。

○ 星环底座支持, 隐私保护全链路

1. 基于星环全链路数据隐私保护能力，可围绕数据的集成、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。
2. 将星环大数据分析能力与隐私计算相结合，多种加密安全手段和优异的通信架构，确保平台在大数据量下也能获得卓越的性能。
3. 可与星环数据安全平台Defensor、数据商城Datamall、数据云平台TDC等星环产品以松耦合的方式组成可信数据流通解决方案，提供安全、合规的数据跨域流通能力。

○ 权威机构认证, 数据安全有保障

1. Sophon P²C作为国内首批通过信通院联邦学习资质认证的隐私计算平台，支持多方在数据隐私保护前提下进行协作。
2. 深度参与隐私计算行业标准的编写，如中国支付清算协会领头编写的《多方安全计算金融应用评估规范》、北京金融科技产业联盟领头编写的《联邦学习技术金融应用白皮书》等。
3. 政务和营销隐私计算落地案例在荣获2021信通院《星河案例——隐私计算优秀案例》。

○ 企业级应用平台, 支持国产化适配

1. 作为企业级应用平台，能有效应对非恶意性错误和显式攻击，拥有良好的系统稳健性，保障企业间数据协作的安全。
2. 支持自动服务部署与调度，提供强大的服务编排调度和资源管理，方便进行集群运维和安全管控，并可实现应用的快速部署、急速扩展和无缝对接，优化硬件资源的使用。
3. 主流国产化生态厂商适配互认，可适配多种国产芯片、操作系统和中间件，满足国产化需求。

主要功能

分布式隐私计算技术支持

Sophon P²C使用了差分隐私、同态加密、掩码技术、不经意传输的隐私计算技术和可信计算技术，结合Sophon Base提供的高性能的分布式算法，可以在海量密文数据上运行分布式机器学习算法，实现“原始数据不动模型动”、“数据可用不可见”的效果。

多样化自定义任务角色配置

Sophon P²C支持用户根据自身业务场景数据情况选择横向或纵向联邦建模任务，并能自定义联邦任务中的主动方、参与方、协调方、模型使用方等，通过向导式及图形化的界面完成角色和任务配置，让用户能够更大效率地配置其数据资产。

丰富的联邦学习建模算法提供

Sophon P²C内置丰富的算子，如联邦缺失值填充、联邦字符串索引、联邦标准化与联邦归一化等联邦特征工程算子，XGBoost、逻辑回归、评分卡、线性回归、泊松回归、决策树、随机森林K均值等通用联邦学习算子，联邦图计算、联邦知识蒸馏、联邦知识融合等个性化联邦学习算子以及多种模型评估算子，在数据不出本地、保障数据安全的情况下，联合各方数据进行联邦学习建模，直观地从准确性、稳定性等方面对联邦模型进行评估，从而确定效果最优的模型。

强大的多方安全计算能力协助

Sophon P²C支持初筛方法和带有计算条件的隐私求交，保证大数据量下的卓越性能，并支持在各方数据不出库的前提下完成匿踪查询，在保证数据隐私安全的情况下完成数据探查；同时可基于秘密分享的安全运算，在各方不共享各自隐私数据明文的情况下，安全地完成基本的四则运算、比较运算、逻辑运算、三角函数、对数函数、指数函数等多种约定的数学函数和统计函数。

应用案例

智慧政务解决方案

在某市群租房排查中，电力企业（主动方、结果获取方）与水务部门（参与方、数据提供方）通过Sophon P²C纵向联邦学习功能服务构建了群租房预测模型，相比本地用电数据训练的模型，AUC提升10%以上，且全程明文数据不出本地，有效保护了居民用水用电的数据隐私信息。

01

在某省疫情防控中，通过Sophon P²C提供的隐私求交技术，卫健委和药监局在保护公民个人隐私数据的前提下，生成了未打疫苗且购买感冒药的重点人员名单，为精准防控提供了有力支撑。

02



联合智能风控解决方案

在Sophon P²C中基于金融企业及其子企业的用户交易数据、外部税务数据、工商信息等多源数据进行联合统计分析，选择需要引入的数据与特征，通过联邦学习构建风控模型，提升模型准确度，增强对高风险人群的辨别能力。

01

根据业务需求，金融企业可在Sophon P²C中基于生成的高风险人群名单对外提供其他隐私计算服务，如匿踪查询、隐私求交、联邦学习等，全面提升联合风控能力。

02

联合精准营销解决方案

车企可通过Sophon P²C提供的联邦学习服务，借助与支付、社交应用等企业合作，安全引入多方数据，丰富用户特征维度，构建诸如高潜客户预测、沉默用户促活等业务模型，实现目标用户挖掘与用户意向的精准触达。

01

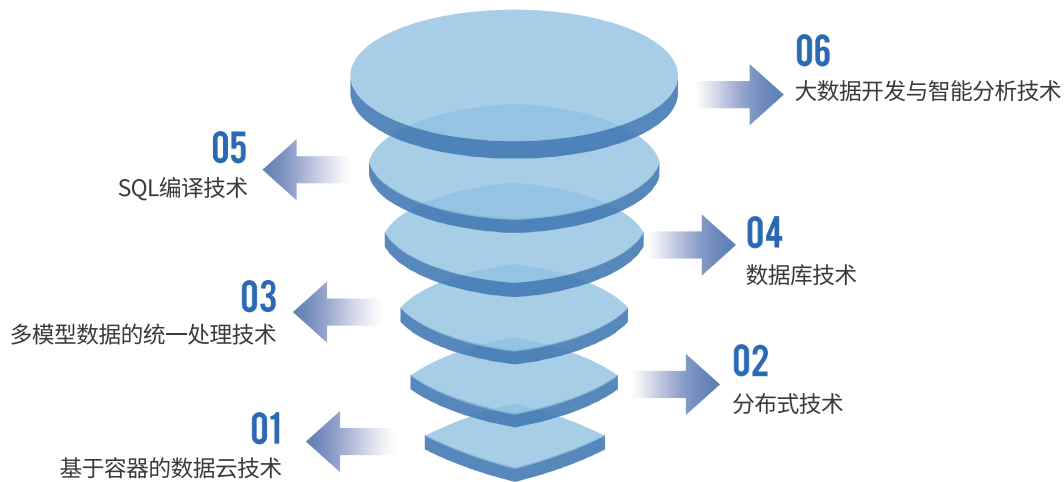
通过Sophon P²C提供的多方安全计算服务，企业可对用户个性行为进行全面统计分析，尤其在面对长尾客户时，能够有效开展精细化运营，进一步提升用户留存率与活跃度。

02

► 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

公司产品已经在金融、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





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。

- ☎ 电话:021-60932577 北京分公司:010-68278990
- 💻 网址:www.transwarp.io
- @ 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
 - 📍 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3012-3013
 - 📍 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

版本 V3.2

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。



公众号



服务号



B站



视频号