



# 科研智能案例集（2025 年）

中国人工智能产业发展联盟

科学智能(AI4S)工作组

2026年1月

---

## 版权声明

---

本报告版权属于中国人工智能产业发展联盟科学智能工作组，并受法律保护。转载、摘编或利用其它方式使用本报告文字或者观点的，应注明“来源：中国人工智能产业发展联盟科学智能工作组”。违反上述声明者，本工作组将追究其相关法律责任。

# 目录

一、 科研智能基础平台 .....	3
(一) 星河启智科学智能开放平台 .....	3
(二) 石景山区 AI for Science 平台项目 .....	12
二、 材料研发 .....	20
(一) 基于科研智能的高分子材料研发 .....	20
(二) AI 自驱动大压铸工艺闭环系统 .....	26
(三) AI 赋能特钢研发 .....	1
(四) AI 驱动的碳捕集材料研发与验证平台 .....	8
(五) 材料显微图像智能分析及应用 .....	15
(六) 统一 NMR 智能解析在化学与材料科学的应用 .....	23
(七) 基于科研智能的锂电电解液研发 .....	31
(八) 深势科技 Piloteye® 电池设计智能研发平台 .....	37
(九) 基于专用算力的电解液 AI 数据平台建设 .....	47
(十) 基于数理双驱模型组合的二维材料发现 .....	56
(十一) 多模态技术赋能有机分子谱图解析模型研发 .....	60
三、 工业仿真 .....	64
(一) 高速动车组空气动力学智能化仿真大模型-斫轮·风驰 .....	64
(二) 民用客机超临界翼型函数生成和编辑模型 .....	72
(三) 风扇叶轮智能优化设计 .....	77
(四) 代理模型在轮胎设计仿真中的应用及仿真自动化 .....	84
(五) ChatCAD 生成式辅助工业设计 .....	93
(六) 面向高端制造业的多模态交互式预测性智能运维决策支持关键技术 .....	100
(七) 基于大数据的色漆漆料优化系统 .....	110
(八) 基于 MWORKS 的具身空间仿真平台 .....	115
四、 地球科学 .....	130
(一) 基于 OneScience 框架的气象 AI 大模型并行训练示范案例 .....	130
(二) 伏羲系列气象大模型研发项目 .....	136
(三) 基于深海基础模型的深海海山数字化智能系统 .....	142
(四) 大模型驱动灾害防御智能体 .....	150
(五) 一站式遥感变化监测智能体 .....	156
(六) 时空可信验证智能体 .....	161
(七) 滑坡隐患关联要素智能识别系统 .....	168
五、 生物医药 .....	179
(一) 人工智能辅助极端耐碱抗体的设计 .....	179
(二) 百度 PaddleHelix 平台针对 GPCR 难成药靶点的多肽设计系统的构建 .....	185
(三) 华为云盘古药物分子大模型助力华中科大快速研制出全球首个利什曼病抑制剂 .....	195
(四) 智慧芽 Eureka 平台：垂直领域大模型赋能生物医药研发全流程智能化 .....	199
六、 智慧农业 .....	208
(一) 华为云×中国农业科学院联合研发农科大模型 .....	208
七、 可控核聚变 .....	215
(一) 基于深度强化学习的球形托卡马克等离子体高精度磁控制 .....	215

## 一、科研智能基础平台

### （一）星河启智科学智能开放平台

#### 1.案例实施单位：

上海科学智能研究院、复旦大学、无限光年（上海）技术有限公司

#### 2.案例背景：

在科学智能的 1.0 时代，人工智能主要作为科研的“加速器”，科学家定义重大科学问题并积累领域数据，由 AI 专业人才针对性设计并优化算法模型，在特定领域取得突破，充分验证了 AI 赋能科学研究的巨大潜力。随着科学问题复杂度的指数级提升、多模态数据的爆发式增长以及大模型技术的成熟，科学智能正加速迈向 2.0 时代。2.0 时代的核心使命是让 AI 成为贯穿科学家科研全链条的核心伙伴，直接助力科学家去回答深邃的“科学之问”。科学领域迫切需要功能更强大、使用更便捷的基础设施和协作平台，让科学家能够直接利用平台化的 AI 工具，低门槛地处理高维异构数据、构建领域大模型、进行智能推理与实验规划，从而更高效地探索未知、验证猜想、发现规律。

#### 3.案例概述：

星河启智是智能体原生的全链路科学智能开放平台，致力于加速科学发现，为全球科学家提供覆盖高价值科学数据、开源科学智能模型、面向科研领域的高效智算、干湿实验闭环、多智能体推理规划及多学科众研的全栈基础设施，驱动科学智能进入新纪元。面向科学家、AI 工程师及广大开发者提供全链路服务，全面释放人工智能的生产

力，推进跨学科协作、攻坚关键科学问题、大幅加速科学发现，从而服务各个领域和行业。

#### 4.行业痛点：

**科学智能基础平台面临如下核心挑战：**一是高价值科学数据分散成孤岛，整合加工难度大，海量科研数据难以实现标准化沉淀与跨学科复用；二是智算基础设施适配不足，科学智能对智算基础设施提出了有别于大语言模型的新要求，需要满足异构融合、训推一体及多精度计算等多元需求；三是流程割裂化，数据管理、模型训练、实验模拟等环节相互孤立，跨学科协作缺乏支撑，AI难以贯穿科研全链，严重制约科学发现效率；四是技术门槛高，多数平台由技术驱动，界面复杂、流程繁琐，缺乏对领域科学家使用习惯的深度理解，导致“科学家想用不会用”；五是生态封闭性，缺乏开放共享的模型、数据和工具社区，跨机构、跨学科的协作创新受阻，难以形成“众研共创”的创新生态。

#### 5.案例详述：

星河启智科学智能开放平台以“AI驱动科研范式重构”为核心理念，围绕科学家科研全流程，构建覆盖数据、模型、算力、实验、推理与协作的全链路基础设施。平台致力于解决科学智能2.0时代中科学家主导科研过程中面临的数据割裂、智算新需求、模型门槛高、实验闭环难、协作效率低等核心痛点，推动科研从“工具辅助”走向“智能伙伴”的新范式。

##### (1) 总体思路

星河启智平台以“AI 原生、全链路贯通、开放协同”为设计理念，构建六大核心能力模块，形成“数据—模型—算力—实验—智能体—社区”一体化的科研智能闭环，旨在实现三大目标。**一是**降低科研门槛，通过开箱即用的模型广场、低代码实验编排与智能体辅助，使非 AI 背景的科学家也能高效开展智能科研。**二是**贯通科研流程，打通“数据采集—模型训练—干湿实验—结果验证”全链路，实现从假设生成到实验验证的智能闭环。**三是**构建开放生态，通过众研社区与开源机制，促进跨学科、跨机构协同，推动科学智能成果的快速传播与复用。

## (2) 软件平台

星河启智科学智能开放平台是面向科研全场景的一站式基础软件体系，其核心功能模块包括数据广场、模型广场、启智 Lab、AI 费米、自主科学探索引擎与众研社区。平台底层构建了统一的异构算力调度引擎与数据基座，实现 CPU/GPU 资源的弹性管理与海量多模态科学数据的可信治理。模型广场汇聚并管理来自多个学科的预训练模型，支持开箱即用与在线微调。自主科学探索引擎作为平台的智能中枢，内置多智能体框架与 workflow 编排系统，可自动调度数据、模型与计算资源，驱动“假设—仿真—实验—分析”的科研闭环。众研社区则构建了覆盖项目匹配、协同开发与成果共享的开放科学生态。该平台通过将分散的科研工具链系统性地整合，为科学家提供了贯穿研发全流程的智能伙伴。



来源：上海科学智能研究院

图 1 星河启智科学智能开放平台

### (3) 技术方案

星河启智平台的技术方案围绕构建“数据—模型—计算—实验”的智能科研闭环展开。方案实施始于数据与模型资源化，通过区块链技术实现科学数据的可信确权与标准化治理，并将涵盖气象、生物、材料等 12+学科的 200 余个前沿模型资源化，形成开箱即用的模型服务。



来源：上海科学智能研究院

图 2 星河启智智能平台科学数据集

核心流程为智能驱动的研发工作流：首先用户可通过自然语言或

可视化方式定义科研问题与约束条件，平台随即调用自主科学探索引擎，自动编排数据处理、模型调用与仿真计算任务，形成可执行的科研流水线。进而实现干湿实验的精准闭环，通过“AI 费米”模块连接虚拟仿真与高校、机构的实体实验设备，利用实验灵巧手等自动化装置执行物理实验，并将结果数据实时回传，用于验证 AI 预测并迭代优化假设。全过程由多智能体系统协同支撑：平台内嵌的专用智能体在统一编排下自主协同，并通过众研社区机制融汇专家知识，最终形成从问题提出到解决方案验证的全链路、自进化科研范式。



来源：上海科学智能研究院

图 3 多模态、多学科融合的科学智能模型

#### (4) 关键指标参数

建成超过 4 万个高质量科学数据集，数据总量达 12PB，平台打通 CPU/GPU 计算能力，大规模加速科学智能场景的强化学习，科研任务按照计算类型分层处理，小时级加速至分钟级；低成本 Serverless 动态部署，打通 GPU 显存、CPU 内存、SSD 闪存三级资源，同等计算资源下，可部署模型数量翻倍；子链路级别问题识别和任务恢复，异常中断秒级恢复，自动续训率 99.9%。

## 6.创新情况:

### (1) 技术创新情况

**数据层面:** 高质量数据集支持用户下载至本地使用, 或直接在平台的其他模块中调用。零样本预测与多任务学习的蛋白质语言模型平台汇聚多个高质量特色科学数据。生命科学领域, 以 **siRNAOD<sup>3</sup>** 数据集为例: 是目前唯一一个集多项关键特征于一体的公开数据库, 汇聚了超过 11 万条 **siRNA** 记录, 具备前所未有的数据规模。它不仅具备科研价值, 更对 **RNA** 药物开发、临床转化、**AI** 靶点筛选等领域具有高度适应性, 是构建未来 **siRNA** 研究与应用生态的重要基础设施之一。

**模型层面:** 学科前沿模型支持直接在平台中使用, 并支持用户自行部署模型, 在“先验知识+模型微调”的范式下, 模型无需从零学习, 即可在新场景中产生具有原创价值的科学发现。平台聚集了高校、科研机构和企业优质科学模型, 前沿成果包括打破气候科学知识碎片化与学科壁垒, 实现跨生态、经济、社会系统的智能认知与推理支撑的气候科学大语言模型 **PI@Climate**; 经 ECMWF 和谷歌认证精度领先的“伏羲气象大模型”; 涵盖分子、细胞、组织、器官等多尺度、多模态的女娲生命系列大模型; 预测精度达到世界一流商用软件薛定谔 **FEP+** 水平的物质大模型等。

**算力层面:** 平台打通 **CPU/GPU** 计算能力, 大规模加速科学智能场景的强化学习, 科研任务按照计算类型分层处理, 小时级加速至分钟级; 低成本 **Serverless** 动态部署, 打通 **GPU** 显存、**CPU** 内存、**SSD**

闪存三级资源，同等计算资源下，可部署模型数量翻倍；子链路级别问题识别和任务恢复，异常中断秒级恢复，自动续训率 99.9%。

**实验层面：**除了高校/产业丰富的湿实验资源，融合自动化的实验智能体以及通用的实验灵巧手。灵巧手具备长程动作思维链 CoAT 能力、泛化能力，在小样本数据下可以做到超高复杂、灵巧的操作。

## (2) 模式创新情况

**Agent 广场**汇聚了丰富的科研智能体资源，涵盖模型、数据处理、实验控制等多个方向。不仅提供资源集合，还通过“组合—执行—记录—反馈”的闭环机制，实现科研模式由分散化工具使用走向系统化、平台化的智能科研范式。同时，过往科研合作多靠学术关系或线下网络推动，存在匹配效率低、跨界沟通成本高等问题，众研社区构建科研任务拆解与合作对象匹配机制，将不同背景的科学家、AI 研究人员和工程师快速聚合，显著提升科研合作的组织效率与契合度。实现从被动合作到主动匹配、从封闭小组到开放生态的模式创新。

## 7.应用实效：

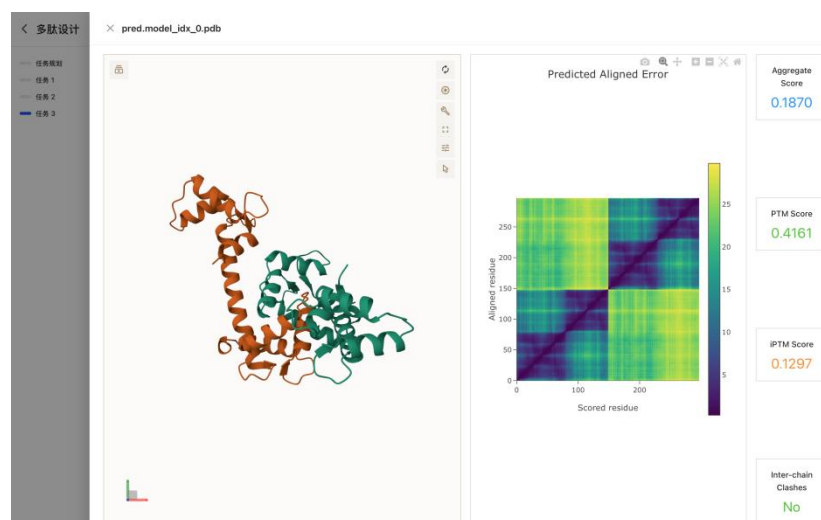
星河启智平台已于 2025 年 7 月 25 日在 2025 世界人工智能大会（WAIC 2025）上发布，开放产品试用，进入正式运营阶段。目前用户主要为：复旦大学在校师生、其他合作高校科研用户、上智院科研用户。同时，通过与魔搭社区的深度合作，共建 AI4S（AI for Science）专区面向广大开发者、C 端用户，目前已实现热门模型的代码共享、权重共享，全面落实开源开放理念。未来通过精细化的产品运营，将进一步拓展 B/C 端用户。

星河启智自主科学探索引擎已实际服务科研人员进行科学研究。基于用户问题自动检索科研知识库，提取关键信息并形成科研方案，并自动化编排科研流程。用户一方面可以在 Agent 广场中查看、编辑、迭代流程，一方面可以通过平台集成的计算资源与算法库，自动运行实验、模拟或分析任务，实时返回结果，并支持反复重跑。最终以图表、代码片段、文献引用等形式可视化输出，支持一键导出报告。

考虑科研问题的深度和流程的复杂性，当前在生命科学和物质科学领域的部分细分话题中，任务规划和推理能力已受到科学家的高度认可，特别是在“逻辑编排能力”“可视化与流程监控”“路径迭代与反馈驱动优化”层面。展示案例验证包括但不限于：

### (1) 多肽设计

胰高血糖素受体在葡萄糖稳态中扮演着关键角色。靶向它的多肽可以作为 2 型糖尿病的治疗方法，通过调节胰高血糖素的作用来发挥效用。设计一个肽序列(15-30 个氨基酸)，该序列能作为人胰高血糖素受体(GCGR,Uniprot ID:P47871)的选择性拮抗剂。



来源：上海科学智能研究院

图 4 多肽设计应用

### (2) 靶向药小分子结构预测

为治疗银屑病，寻找能够作用于 Tyk2 靶点的小分子结构。

### (3) 新材料结构发现

为设计医疗诊断应用和精密加工具有应用场景的深紫外光电器件，需要在禁带连续可调 AlGaIn 上实现能带值大于 4.5eV 的配比，帮用户设计一个配比方案。

## （二）石景山区 AI for Science 平台项目

### 1.案例实施单位：

北京火山引擎科技有限公司、北京枫清科技有限公司、中化数智科技有限公司

### 2.案例背景：

从国家层面，科技部会同国家自然科学基金委启动了“人工智能驱动的科学研究”（AI for Science）专项部署工作，布局“人工智能驱动的科学研究”前沿科技研发体系。石景山区作为全国首个 AI 大模型产业集聚区，近年重点布局人工智能产业，当前区域内尚未形成覆盖“数据—算法—算力—场景”的 AI for Science 一体化平台。与此同时，聚焦未来制造、未来信息、未来材料、未来能源、未来空间、未来健康六个方向，以 AI for Science 平台助力产业聚集，探索 AI+新材料等新领域，带动未来产业发展。

### 3.案例概述：

打造国内领先的区级 AI for Science 平台，赋能科研院所与企业在新材料合成、化工研发、生物医药等领域的创新实践，推动石景山区产业数字化转型。紧密围绕发展趋势，进一步加强石景山区智能科研服务平台建设，为新材料、生物医药等领域科研项目提供支撑：

一是提供 AI for Science 门户，将整体能力统一至平台之上，并且提供基础设施服务进行平台运行。二是打造 AI for Science 全流程的支撑平台和工具链，通过行业智能体平台、通用工具链和场景工具链，实现对各 AI for Science 业务场景的串联，助力新材料发现、药

物分析发现、生物信息分析等场景的智能实现。**三是**构建 AI for Science 领域大模型，优先根据客户需求实现化工材料大模型，统一框架和科学规律建模，打破学科边界，实现数据、算法与实验的深度协同。**四是**积极参与 AI for Science 数据集标准制定，支持构建 AI for Science 高质量数据集。

#### 4.行业痛点：

在 AI for Science 的工作中，有如下挑战：**一是**高质量科学数据稀缺；**二是**跨学科人才（AI+领域科学）缺口大，培养周期长；**三是**AI 模型训练对算力需求激增，需平衡硬件成本与能耗；**四是**缺乏支持 AI for Science 全生命周期的统一平台；**五是**需针对科学场景定制专用算法与大模型（如 AlphaFold），而非依赖通用模型；**六是**需从科研向产业延伸，看重实际落地。

#### 5.案例详述：

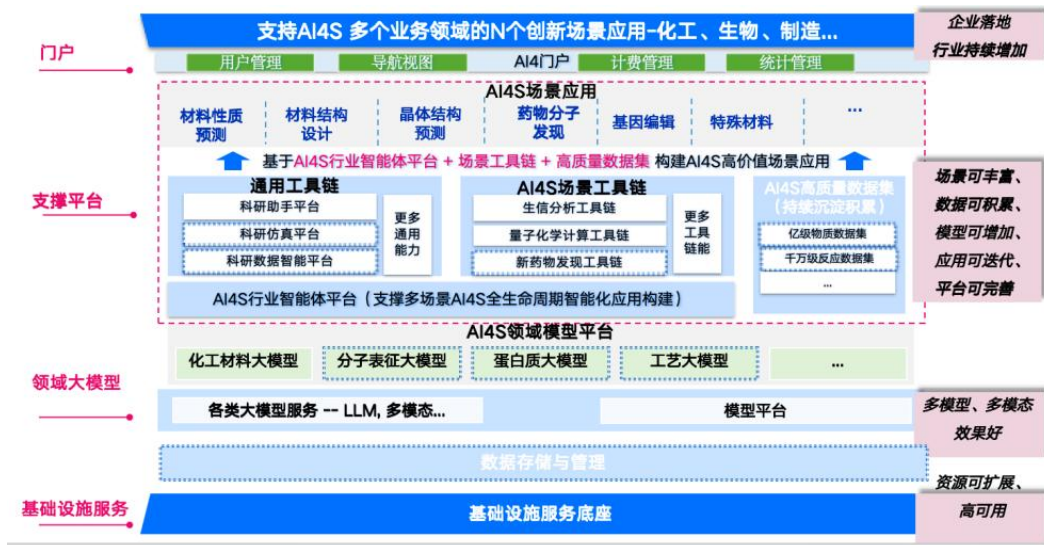
##### （1）总体思路

本项目以“平台化、场景化、生态化”为核心建设理念，打造覆盖“数据—算法—算力—场景”的 AI for Science 一体化平台。**总体思路围绕“四个一”展开：**一套基础设施，整合通用计算与智能算力，实现资源的弹性调度与高效利用；一个领域模型工厂，基于通用大模型底座，训练如化工材料、生物分子等垂直领域大模型，注入行业知识；一条智能支撑工具链，通过低代码智能体平台与可视化 workflow 引擎，将通用工具与场景化工具（如生物信息分析、量子化学计算）模块化封装，支撑科研全流程；一个统一门户，集成用户管理、资源调度与

场景应用，为区域创新主体提供一站式智能科研服务入口，最终构建产学研用深度融合的开放创新生态。

## (2) 软件平台

AI for Science 平台采用分层架构设计，构建了四大核心子系统。一是平台以 AI for Science 基础设施服务为底层支撑，提供通用计算与智能算力资源；二是在其上构建 AI for Science 领域模型平台，专注于化工材料、分子表征等垂直领域大模型的训练与调优；三是 AI for Science 支撑平台通过行业智能体平台、通用工具链和场景工具链，为科研全流程提供低代码开发与自动化支撑；四是最终通过统一的 AI for Science 门户，集成用户管理、业务导航与多场景应用，为科研人员提供一站式智能科研服务入口，形成完整闭环。



来源：北京火山引擎科技有限公司

图 5 石景山区 AI for Science 平台架构

## (3) 技术方案

平台采用"通用大模型+领域大模型+领域小模型"的协同技术路线。基于通用大模型的广泛知识理解能力，通过领域知识注入与专业

训练，构建具备行业深度的领域大模型；同时针对特定科研任务开发轻量级领域小模型，确保专业场景下的精度与效率。技术实施层面，通过智能体平台实现 workflow 可视化编排，将数据预处理、模型调用、仿真计算等环节串联成自动化科研流水线，并与信通院合作推进数据集标准制定，构建从数据管理、模型训练到智能应用的完整技术闭环。

#### (4) 关键指标参数

平台已覆盖化工、材料、生物医药等 5 大重点领域，集成超过 10 套科研工具链，通过低代码智能体平台，将科研智能应用的开发周期从传统的 4—6 周缩短至 3—5 天，效率提升 8 倍以上，极大降低了 AI 技术的应用门槛。

### 6. 创新情况：

#### (1) 技术创新

针对 AI for Science 平台，构建通用智能体平台，并且提出通用工具链和场景工具链的概念，从通用工具和科研场景角度对 AI for Science 提供支持。

在 AI for Science 科研支撑平台建设中，搭建低代码的智能体平台具有迫切的现实需求。随着科研数据多模态化（文本、图像、数值等）和工具平台异构化加剧，科研人员需频繁整合 NLP 模型、数据分析工具、实验模拟系统等资源，但技术实现门槛导致智能体开发效率低下（平均周期达 4—6 周）。科研智能体平台分系统通过低代码架构，提供可视化 workflow 编排、组件化 AI 能力封装和跨系统 API 集成，支持科研人员以拖拽方式快速构建诸如实验数据智能处理、文献

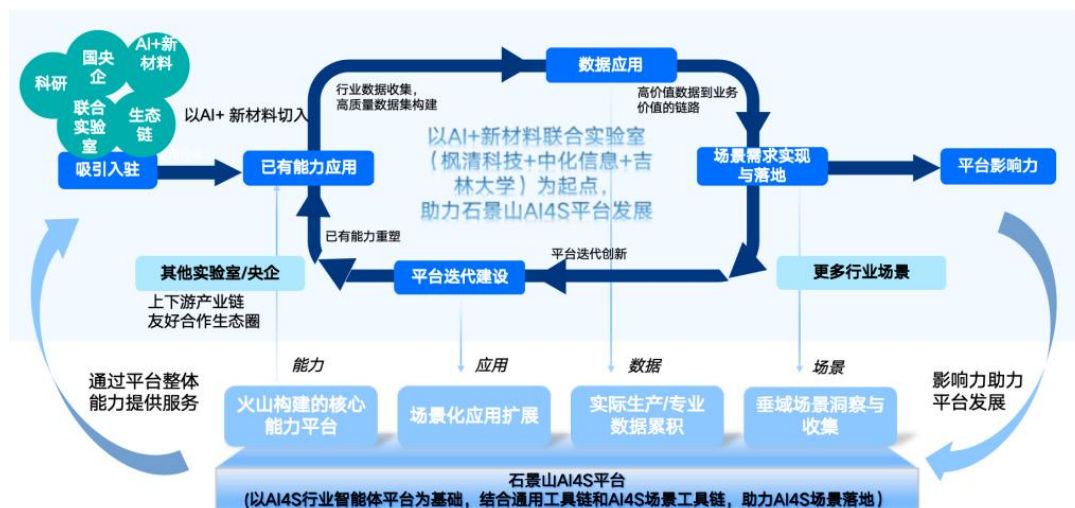
知识图谱分析、模型自动化调参等智能体应用，将开发周期缩短至3-5天，显著降低技术转化成本。

平台建设顺应 AI 与科研深度融合趋势，旨在突破技术开发瓶颈，赋能生物信息学、材料科学等领域科研人员自主构建智能化工具，推动科研范式从“试错驱动”向“智能设计”转型。作为区域科技创新的核心载体，石景山区聚焦新材料、生物医药、智慧农业等战略领域，但传统科研模式存在跨学科协作低效、数据资产利用率低、算力资源调度粗放等问题。例如，化工领域重复计算材料性能的项目占比达 42%，农业智能化场景中数据标注成本占研发周期 60% 以上。构建 AI for Science 科研智能体平台，是破解区域科研资源碎片化、加速科技成果转化转化的关键抓手。

提供科研所需智能体平台，用户在智能体平台之上可以进行以下工作：**一是**利用通用工具链辅助论文阅读、对论文进行翻译以及提供论文摘要，提供科研效率；**二是**利用场景工具链进行科研任务，比如利用生物医药平台进行蛋白质预测，利用量子化学工具链加速量子化学计算；**三是**用户也可以基于平台进行工具链扩展，助力自身的科研任务；**四是**在化工材料大模型方面，为新材料设计等提供助力，提供物质信息检索、逆合成、反应条件查询等功能。

## （2）模式创新

通过产学研用一体化的方式实现运营模式的创新，具体流程为：



来源：北京火山引擎科技有限公司

图 6 产业研用具体流程

以 AI+新材料联合实验室为起点，逐步实现从能力供给→数据应用→场景需求落地→平台影响力提升→更多行业场景拓展→反哺平台建设的循环闭环。其中关键环节包括：**一是**通过科研院所、央企、联合实验室和生态链伙伴形成创新共同体，以 AI+新材料为切入口；**二是**以“火山构建的核心能力平台”为基础，逐步沉淀通用能力；**三是**数据与场景驱动，通过行业数据采集、应用实践和生产数据积累，不断形成高质量数据和场景应用；**四是**将平台能力迁移到更多垂直行业和跨领域场景，带动影响力；**五是**平台在实际应用中不断迭代升级，实现能力重塑，形成长期可持续的创新循环。

## 7.应用实效：

平台构建全链条智能体开发与管理体系，覆盖工具链支撑、多模型兼容、协作管理、生态运营、数据监测及安全防护等核心维度，形成功能完备的 AI for Science 科研智能体开发与运行环境，具备如下能力且对外提供服务：

### (1) 智能体全生命周期工具链

提供完整的智能体搭建工具集合，支持提示词工程、插件开发、知识库构建、 workflow 设计、评测对比及内容审核等核心功能模块，实现从智能体设计、训练到部署的全流程可视化操作，降低技术门槛，提升开发效率。

### (2) 多模态模型适配与接入

支持开源模型、商业模型、本地部署模型、领域大模型等多元类型，具备灵活的模型适配框架，可快速集成不同技术路线的 AI 模型，满足科研场景对算法多样性和定制化的需求，实现“一键接入、统一管理”。

### (3) 协作空间与权限管理

构建层级化协作管理体系，支持工作空间创建、成员角色管理（添加/删除）及资源隔离共享，实现智能体、插件、知识库等资产在工作空间维度的安全管控。通过精细化权限配置，保障跨团队协作中的数据安全与知识产权保护。

### (4) 智能体与插件生态市场

搭建开放的智能体与插件交易中心，支持用户将自研智能体及插件发布至平台生态市场，实现科研工具与成果的沉淀、共享与复用。通过生态运营机制，促进跨机构技术交流，加速 AI for Science 应用创新与产业化进程。

### (5) 智能体数据分析与监控

开发多维度数据分析看板，实时跟踪智能体运行状态，涵盖消息

总量、互动用户画像、平均会话时长、Token 输出效率、用户满意度等关键指标。配套完善的日志系统与标注工具，全面记录智能体运行轨迹，为优化迭代提供数据支撑。

#### (6) 安全合规与内容管控

建立多层次安全防护体系，通过敏感词库管理、内容实时审核、数据加密传输存储等机制，确保模型输出合规可控。支持违规会话追溯与审计，可快速定位命中敏感词的交互记录，满足科研数据安全与合规要求。

#### (7) 安全合规与内容管控

提供开箱即用的问答会话页面组件及完备的 OpenAPI 接口，支持第三方应用快速接入平台能力，实现智能体功能与科研业务系统的深度融合，降低技术集成成本，提升平台服务的可扩展性与兼容性。平台提供的化工材料大模型，以中化信息的天枢大模型为基础版本，对中化各二级公司以及化工类企业可以以对话交互的形式提供针对分子检索、逆合成设计、反应条件推荐等研发任务的问答，为用户提供高效、便捷、精准的化工新材料合成提供助手服务，具有重要的应用价值。

## 二、材料研发

### （一）基于科研智能的高分子材料研发

#### 1. 案例实施单位：

厦门大学、嘉庚创新实验室、厦门智储大装置研究院有限公司、  
厦门依华智慧科技有限公司

#### 2. 案例背景：

高分子材料在世界科技前沿应用广泛，涵盖从高性能工程塑料到新能源材料、柔性电子材料等多个领域。传统依赖人工试错的研发模式存在周期长、成本高等瓶颈，严重制约高分子材料发展进程。国际科研机构加速推进人工智能驱动的科技研发，已在材料性能预测、合成路径设计等领域取得突破性进展，展现出颠覆传统研发范式的潜力。因此，亟需通过人工智能和机器人融合方式提升高分子材料研发效率，服务国家重大需求和新质生产力发展。



来源：厦门大学

图 7 AI for science 是高分子材料研发革命性的突破口

#### 3. 案例概述：

基于科研智能的高分子材料研发是以大数据和人工智能算法为基础，将高通量样品制备（手）、多维度样品表征系统（眼）、智能

决策迭代（脑）技术深度融合，以国家战略领域关键材料研发需求为牵引建设的无人智能材料研发实验系统。系统通过机器人技术和人工智能决策的结合，突破了无人干预条件下材料制备、表征和智能迭代的材料研发实验系统关键技术，实现材料的“决策—制备—表征”的干湿闭环研发。

#### 4.行业痛点：

高分子材料单体种类繁多且配方复杂，反应参数空间巨大，传统材料研发高度依赖人工试错，效率低、成本高，难以适应快速迭代需求；智能化转型中，国内缺乏自主可控的核心技术工具，实验数据碎片化且标准化不足，制约模型训练与验证；实验室研发与产业需求脱节，导致新材料从设计到量产验证周期长、风险高。人工智能与高分子材料研发深度融合不足，缺乏成功落地应用案例，亟需高效智能的新范式突破效率瓶颈。

#### 5.项目获奖、专利授权及获得投融资情况

- (1) 研发范式获批多个项目支持。
- (2) 与行业龙头企业联合开展技术研发，合同金额达数千万元。

#### 6.案例详述：

##### (1) 总体思路

针对高分子材料高效研发的技术挑战，项目组自主研发一系列结合机械手、精密位移运动控制系统等多智能体协同为基础的智慧材料研发技术。针对高分子材料高通量制备和多维度表征的技术挑战，建立对所制备材料的核磁、质谱、色谱、红外、拉曼等谱学表征和力

学、热学、电磁学等性能表征的多维度数据采集和自动化分析系统。针对高分子材料研发对象复杂、依赖试错和经验积累、难以理性设计的技术挑战，建立以科学智能(AI4S)为核心的智慧材料研发新范式，推动高分子材料开发进入人工智能驱动的新阶段。

## (2) 软件平台

项目组自主建设了基于具身科学智能的高分子材料研发系统，突破复杂配方设计及巨量参数空间的智能化寻优瓶颈，实现高分子复杂体系的高通量合成-多尺度结构表征-性能预测算法的闭环研发，并完成实验设备全链条国产化替代。通过联合国家实验室构建“人工智能模型-机器人实验-产业验证”三位一体的智能研发底座，并与相关企业共建垂直领域研发子平台，为多场景高分子材料提供定制化智能研发解决方案，构建自主可控的高分子材料研发体系，保障国家战略产业供应链安全。

## (3) 技术方案

**基于科研智能的高分子材料研发由三个子系统组成：**

一是高通量样品制备系统（手），以基于智能决策和机器人自动高通量材料—器件制备为目标，建立自动化、高通量的材料—器件制备平台，形成材料智能制备的自主可控仪器装备能力；二是多维度样品表征系统（眼）：建立面向材料研发的涵盖核磁、质谱、色谱、红外、拉曼等谱学表征和力学、热学、电磁学等性能表征的多维度数据采集和自动化分析平台，形成满足材料智能研发的多维度综合表征能力，三是智能决策迭代系统（脑）：建立结合标准化数据库、人工智

能算法与大模型的智能决策技术,实现材料研发全流程的决策、调度、控制、监测和反馈,形成“决策—制备—表征”的干湿闭环迭代研发能力。

#### (4) 关键指标参数

一是实现了材料的自动化高通量制备与表征,24小时内无人干预情况下完成超过100种制备条件的材料制备和表征;二是无人干预情况下自动运行连续工作时长大于240小时;三是针对特定材料需求,形成包括超过10000种制备条件及对应材料性能的数据库以及具有预测能力的智能预测算法。

### 7.创新情况:

#### (1) 技术创新

针对材料研发多尺度耦合、数据孤岛严重、依赖进口设备等行业痛点,项目在数据治理、模型与算法等层面实现多项关键突破,形成可落地的技术体系。

在数据层面,通过“感知—学习—交互”的智能体协作,将材料研发的“表征-决策-制备”流程智能化、高效化,联合苏州实验室、上海人工智能实验室、北京科学智能研究院等平台,中国科学技术大学、中国科学院高能物理研究所等一流科研院所,以及深势科技等企业,初步构建了高分子材料的百万级实验数据集、计算数据集和文献数据集,以及相关高分子材料大模型语料库。

模型与算法层面,突破跨尺度关联技术,首次构建了特种高分子材料微观分子结构—分子间作用力—宏观流变性能的跨尺度关联模

型，为材料性能的精准优化与高效迭代奠定了坚实基础。项目阶段性产品性能已得到应用单位的验证；开发了聚酰亚胺等高分子材料大模型，建立了结构化的知识库体系。基于材料性能需求，大模型智能推荐合成方案，并指导自动化实验平台完成全流程验证，成功开发出玻璃化转变温度突破 550°C 的高性能聚酰亚胺材料。

## (2) 模式创新

依托技术突破，实验室构建“跨领域协同+全链条智能化+场景化落地”的创新模式，形成高分子材料研发领域的新产品与新业态，推动行业范式升级。

**在产学研协同模式上**，打造“高校-科研院所-企业”跨领域创新网络，联合国家级平台、一流科研院所、以及行业龙头企业，构建智能实验室“算法驱动—材料创新—场景验证”闭环攻关体系。

**在服务模式上**，推出“定制化智能研发解决方案”，针对新能源、电子电器、航空航天等不同领域需求，提供从数据生产到模型验证的全链条服务。面向新能源电池领域，为相关企业定制基于 AI 算法的电池关键原料智能研发系统，实现原料的智能制备、自动表征、配方的智能筛选和迭代；面向芯片制造领域，与恒坤新材料共同构建高通量智能有机合成系统和开发专用智能算法，实现研发全流程闭环优化。

**在成果转化模式上**，基于科研智能的高分子材料研发，孵化厦门智储大装置研究院有限公司、厦门依华智慧科技有限公司等企业，推动技术向产业化延伸。同时，积极参与智能实验室和智能平台相关标准制定工作，形成“技术研发-标准制定-产业应用”的良性循环，为材

料领域智能化转型提供可复制的“产学研用”范本。

## 8.应用实效：

### (1) 新能源电池高分子材料领域

与新能源企业密切合作，建立基于 AI 算法的电池关键原料智能研发系统，实现原料的智能制备、自动表征、配方的智能筛选和迭代，具备每日 100 组配方的智能研发能力。初步对氟代羧酸酯、环状氟代羧酸酯以及氟代芳香类化合物组成的巨大配方空间进行搜索，实现锂电池电解液“干湿闭环”研发范式——非常规温度电解液的智能研发，通过自动化实验完成 40 组配方内筛选四元电解液配方的最优值。

### (2) 电子电器高分子材料领域

与光刻胶企业开展多项科研项目合作，共建校企联合研发中心，合作领域涵盖先进半导体材料、电子化学品，已实现对标 ArF 光刻胶树脂合成工艺路线开发；共建联合研发中心开展多项技术联合攻关。

### (3) 航空航天高分子材料领域

与航天企业合作，突破了无人干预条件下航空航天特种高分子材料合成、表征和智能迭代的材料研发实验系统关键技术，实现了实验系统全链条软硬件的国产化自主可控。

## （二）AI 自驱动大压铸工艺闭环系统

### 1. 案例实施单位：

小米技术委 AI 实验室：核心算法开发

小米汽车质量部：核心数据支持

小米汽车工厂压铸车间：生产实验测试支持

小米汽车智能智造装备部：压铸设备 plc 控制打通

小米汽车智能智造制造工艺部：现场工程实施

小米汽车数字化团队：MES 系统打通院

### 2. 案例背景：

当前，以 Deepseek 为代表的前沿大模型技术持续迭代，在全球多行业形成技术赋能趋势，展现出广泛的应用潜力。从行业发展现状来看，AI 与制造业的融合正处于加速渗透期，但传统智能制造领域在核心生产环节的技术转化效率仍有提升空间，大模型的赋能价值尚未完全释放。

2024 年小米 SU7 上市后获得市场广泛关注，面对规模化交付需求，小米将 AI 大模型技术深度应用于汽车生产制造全流程，重点聚焦一体化压铸这一关键工艺环节，构建了一套赋能压铸全链路的工艺自驱动闭环系统。该系统以大模型为核心支撑，分别在材料研发、工艺优化、质量检测三大核心场景实现技术落地——通过 AI 配方寻优缩短新材料研发周期，依托智能算法优化生产参数，借助视觉大模型实现毫米级缺陷精准检测，为高效交付提供了技术保障。



来源：小米汽车

图 8 小米工艺自驱动闭环系统

### 3. 案例概述：

本案例聚焦“小米在汽车制造领域推出的工艺自驱动闭环系统”，**面向一体化压铸全链路首创小米擎天柱工业大模型作为系统核心引擎**。该系统精准瞄准行业大压铸生产中的核心痛点，具体解决三大关键问题：材料研发层面研发周期长成本高、工艺优化层面生产稳定性欠缺与参数协同不畅、质量检测层面全流程质量管控难。

在技术层面，系统通过三大核心路径，分别突破材料研发、工艺优化、质量检测领域瓶颈，构建全链路价值闭环：一是深度挖掘多源异构材料知识，涵盖材料成分、力学性能、材料机理等多个维度，完成高维材料特征的统一提取、精准表征与深度融合，通过自研的多元材料 AI 仿真系统，降低材料研发阶段的试验成本与时间成本；二是搭建高度集成的中控系统，结合半监督学习、多模态信息融合和数据生成等技术，构建“机械性能全量预测—工艺优化—参数自动调控”的完整闭环链路，打通设备与系统协同的“最后一公里”，实现一体化压铸全链路的工艺闭环优化；三是深度集成 X 光射线等高性能工业检

测硬件设备，创新设计全自动提示与多阶段精细化微调技术方案，构建面向一体化压铸场景的工业视觉大模型，仅需少量标注数据即可实现毫米级微小缺陷的高精度检测与分割，打破传统人工质检与半自动化检测的效率和精度瓶颈，实现铸件的高可靠性、自动化闭环。

#### 4.行业痛点：

一体化大压铸技术是极高的工艺控制壁垒——虽发展多年却迟迟未能应用于汽车结构件的核心瓶颈。其关键挑战在于模具热场的纳米级稳定性控制：仅几度的温度波动，就可能致整批次后车身铸件产生蜂窝状气孔等致命缺陷，凸显了该工艺在热力学平衡与过程精度控制上的极端复杂性。工艺控制不好，不仅造成大量铝合金原料报废，严重的更会引发产线停摆、模具返修及交付违约的连锁损失，会严重影响整个产线的车辆交付力，综合以上，存在以下三项痛点：一是质量检测主要依赖破坏性检测和抽检，难以全面掌控每一件产品的质量全生命周期，导致缺陷品易流入下游环节，增加风险。二是生产过程复杂且动态，缺乏有效的实时监控和异常反馈，导致问题一经发现已造成大量不合格品，错失及时调整机会。三是 AI 模型难以与底层设备系统高效对接，缺乏自动化闭环调节能力，对模型可解释性与鲁棒性要求极高。

#### 5.创新情况：

工艺自驱动闭环系统已率先在多条压铸产线落地应用，推动质量管理实现从“事后救火”到“事前防火”的关键跃迁。在量产场景中，该系统不仅实现全流程质量可控，更达成废品率降低、用人成本最小化

的多重效益，成功斩获国际领先技术鉴定 2 项，为行业树立起智能制造的新标杆。

### (1) AI 驱动材料开发模式变革，推动材料研发迈入大模型时代

本系统突破传统材料研发瓶颈，实现从“实验驱动”到“大模型驱动”的创新跃迁：将传统依赖专家经验摸索配方的研发范式，升级为“大模型驱动”的研发范式，并创新性提出 MEE AI 模型，通过“性能预测 + 配方寻优”的双重能力，精准锁定性能最优的材料配方。依托大模型研发范式，材料研发周期缩短约 50%，其自主研发的小米泰坦合金已成功应用于汽车工厂的大压铸件生产，实现技术成果向实际产能的高效转化。

### (2) “超级工业巨兽”智慧大脑，引领压铸工厂迈向行业领先

压铸自驱动工艺闭环系统构建“感知—决策—执行”三位一体的智慧架构：感知系统方面，通过开发高度集成的中控系统，实现全流程数据实时采集，为后续决策提供精准数据支撑；决策系统方面，依托 AI 技术实现最优工艺参数推荐、非破坏性力学性能预测、事前质量预警与故障报警三大核心能力；执行系统方面，高度集成终端数百台设备，确保决策指令高效落地。这一架构彻底变革传统力学性能检测模式，从“破坏性抽检”升级为“100%AI 全检”；同时，AI 推荐的工艺参数不仅提升压铸件力学性能指标，还大幅缩短压铸岛设备调试时间，双管齐下保障工厂的质量与产能双达标。

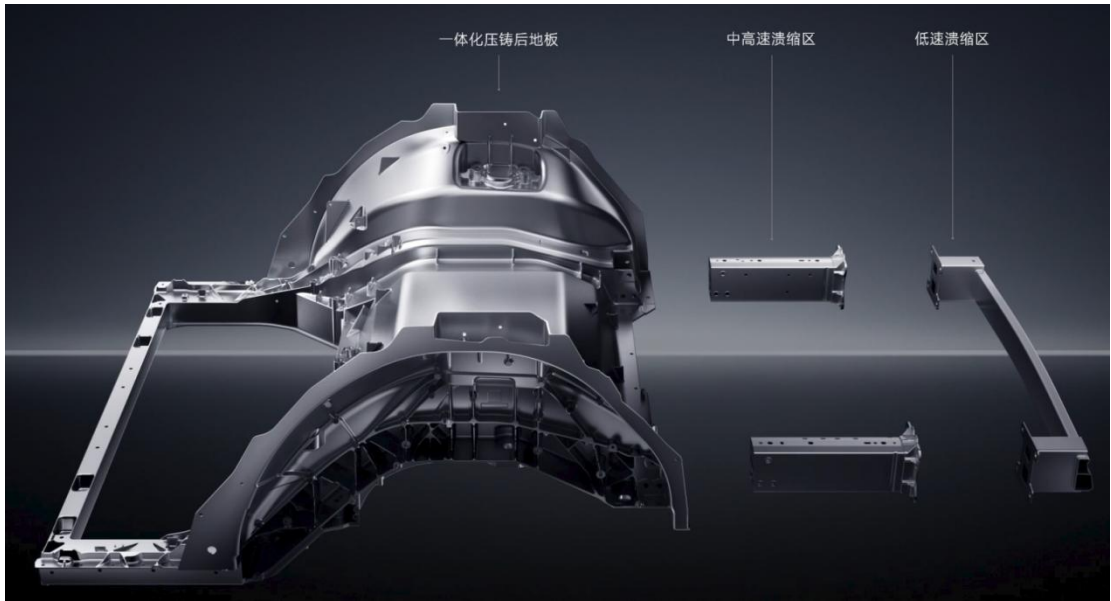
### (3) 创新大压铸件质检大模型技术，实现缺陷精准识别与闭环管控

针对大压铸件内部缺陷“样本量小、尺寸微小、形态多样”的行业难题，自研视觉大模型，质检准确率突破 99%，该技术能够实现毫米级缺陷的像素级分割，精准还原缺陷的位置与尺寸，检测效果远超人工。通过搭载该大模型大幅降低了人工成本，保障产品高质量交付。此外，系统通过缺陷分析反向辅助工艺优化，形成“检测—分析—提升”的质量闭环流程，并已实现平台化落地，可实时开展检测与质量改进。目前，该质检大模型已成功应用于量产的 SU7 与 YU7 两款车型。

## 6.应用实效：

### (1) 质量智能检测系统

当前，涵盖“内部缺陷检测”与“力学性能检测”的质量检测系统，已在小米汽车工厂压铸车间的各个压铸岛全面落地应用。这一系统彻底改变了传统质检模式：此前大压铸产品的机械性能质检采用“3%人工破坏性检测”，如今已升级为基于 AI 技术的“大压铸后地板 100%全检”。从检测效果来看，机械性能预测准确率高达 95%以上，其中 UTS（抗拉强度）、YS（屈服强度）、EL（延伸率）等关键力学性能指标的平均误差分别控制在 4.13MPa、4.01MPa、0.91%的优异水平。



来源：小米汽车

图9 小米一体化压铸集群

## (2) 智能参数优化系统

智能参数优化系统已全面覆盖小米汽车工厂压铸车间的各个压铸岛，为现场生产工艺参数调优带来全方位提效。借助该系统，工程师无需手动逐一监测生产测量参数，可通过智能监控系统直接识别设备异常，并依托参数推荐系统实现生产参数的自动调优。实际应用中，各压铸岛铸件的机械性能显著提升，其中机械性能平均合格率提高3.6%，有效夯实了量产质量基础。

### （三）AI 赋能特钢研发

#### 1.案例实施单位：

之江实验室、中信泰富特钢集团股份有限公司

#### 2.案例背景：

传统特钢企业在合金钢材材料研发与生产过程中，长期依赖专家经验与反复试验，面临研发周期长、成本高、知识难以沉淀等问题。尽管掌握丰富的工业数据资源，受限于技术门槛与数据壁垒，智能化工具难以落地。之江实验室与中信泰富特钢联合开展 AI4S 探索，构建以大语言模型为核心的智能研发系统，融合企业真实产线数据和专业知识，实现合金钢材设计、工艺优化、故障诊断等场景的一体化智能赋能，推动特钢行业从经验驱动迈向智能驱动。

#### 3.案例概述：

本案例聚焦“AI 赋能特种钢材研发”，通过“大模型+小模型/工具（1+X）”混合架构，构建以自然语言为接口的智能合金材料设计与工艺优化系统。系统基于 32B 大语言模型，融合冶金领域知识与企业私有数据，融合多种性能预测和视觉识别小模型，覆盖从文档问答、参数推荐到逆向设计、智能巡检等多类业务场景，实现特钢企业领域大语言模型落地应用。平台系统已在多个产线试点运行，显著提升研发效率与生产决策质量，形成“可推广、可自持、可信赖”的智能化转型范式。

#### 4.行业痛点：

钢铁行业研发依赖经验、试错成本高，传统信息化系统与实际业

务耦合难度大。同时，AI 模型部署门槛高、模型泛化能力差、缺乏行业语义理解能力，难以在生产一线广泛应用。亟需一种融合专业知识与通用智能的 AI 系统，打通从知识获取、数据分析到决策优化的全链条，赋能特钢企业高质量发展。

## 5.案例详述：

### (1) 总体思路

面对特钢行业材料设计周期长、试错成本高、专家经验难以复用等痛点，之江实验室联合中信泰富特钢，提出以“大语言模型为中枢、小模型为工具”的 AI4S 智能研发新范式。该范式通过将通用模型与企业私有数据、合金材料领域知识深度融合，构建对话式研发平台系统。平台系统支持自然语言交互，融合文本、结构化数据和图像信息，实现从知识问答、性能预测到合金材料逆向设计的多任务协同，同时推动从“AI 辅助专家”向“专家增强 AI”的范式转变。

### (2) 软件平台

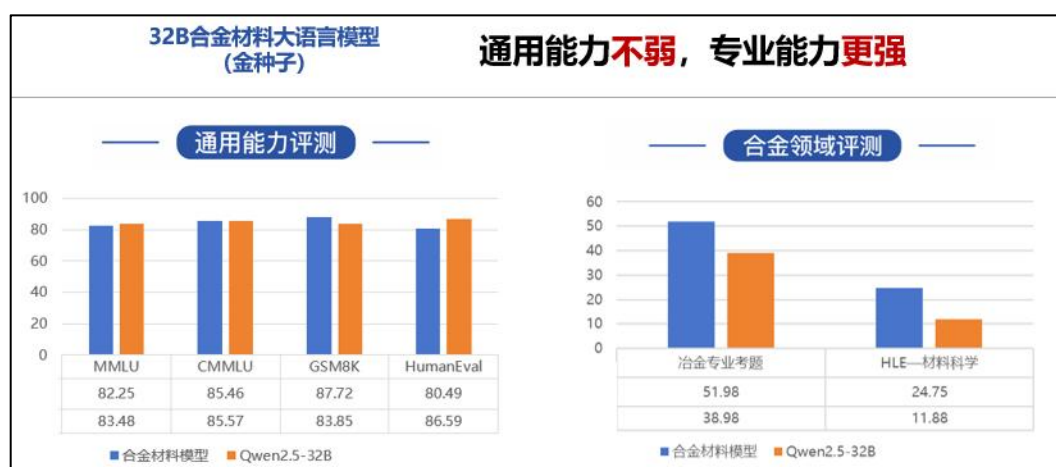
平台系统采用“1+X”混合智能架构。



来源：之江实验室

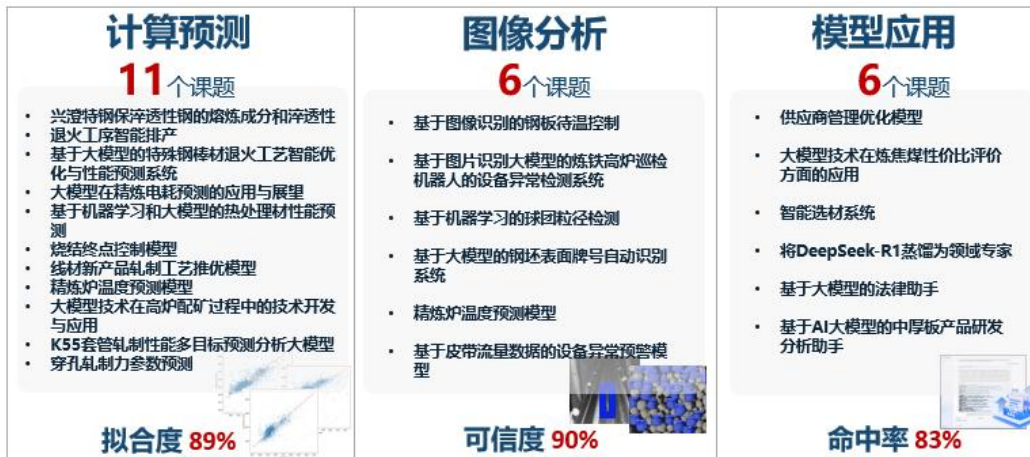
图 10 特钢企业领域大语言模型平台架构

包括以下核心模块：一是 32B 合金材料领域大语言模型。基于通用模型微调训练，具备冶金工艺问答、研发流程推理、文档解析等能力。二是 X 个专业领域模型/工具。专家模型包含炉温预测、淬透性计算、碳化层厚度预测、轧制工艺优化、合金材料逆向设计、图像识别等模块，支持高精度预测与模拟。RAG 知识库整合企业供应商信息、选材资料等文本，支持多轮语义检索与文档级对话。三是可视化 workflow 编排平台。用户通过自然语言描述需求，系统自动调度模型和工具执行任务，支持可视化流程监控与结果回溯。整个系统已在特钢多个产线试点，适配炼钢、轧制、热处理等典型流程场景。



来源：之江实验室

图 11 模型通用能力和合金领域评测



来源：之江实验室

图 12 专业领域工具应用

### (3) 技术方案

一是大模型训练与行业适配。以 Qwen2.5-32B 为基础模型，利用收集的 TB 级合金材料领域语料进行微调训练，增强模型的行业专业性与本地化表达能力。二是领域模型/工具集成与异构桥接。围绕特钢场景构建多个轻量化预测模型与工具，包括产线参数预测模型、合金材料性能预测模型、图像识别模型、企业文档和数据库查询。通过设计规范的专用中间件模块，实现模型与工具之间的数据格式自动转化与传递（如自然语言到结构化数据、数据库查询结果到自然语言解释），确保流程无缝衔接。三是安全隔离与可控部署。平台系统支持本地化部署，所有模型推理均可在企业内网环境完成，知识库与模型权重完全可控，符合工业级数据安全要求。平台系统具备持续扩展与版本迭代能力，便于后续适配其他产线与应用场景。四是关键指标参数。32B 在通用评测集中表现与千问模型相当，在冶金专业考题和合金材料科学专家测试题中超越千问；炉温、碳化层、轧制工艺等产线参数预测拟合度均达到 90%以上；合金材料力学性能预测拟合度达到

89%；图像识别模型准确率 92%。

## 6.创新情况：

### (1) 技术创新

一方面行业专属语料与多模态数据融合，助力模型在专业测试中表现领先。平台构建了覆盖特钢全流程的专属行业语料库，并融合金属材料领域知识和通用开源语料，通过统一的预处理与标准化流程，实现不同领域、多模态数据的高质量集成。训练后的大模型在专业测试集中表现优异，专业能力显著超越同类模型，充分验证了“通用+行业”融合语料在提升模型准确性、稳定性与可信度方面的有效性和可行性。另一方面异构模型智能路由与统一接口，支撑复杂任务的高效调度。平台创新性构建“通用语言模型+垂直专家模型/工具链”的异构协同架构，具备高精度的自然语言意图识别与动态任务路由能力，能够根据用户输入自动识别任务类型并匹配最优的处理路径。通过统一接口机制打通大模型与各类结构化子模型之间的调用链路，实现知识问答、工艺推荐、预测计算等任务的自动化流转，显著提升系统的任务调度效率和跨场景泛化能力。

### (2) 模式创新

一方面从专家辅助到专家增强的合作开发新范式，重塑 AI 系统建设流程。本案例打破传统“AI 工程师单向服务业务专家”的建模范式，提出一种以大语言模型为中介的人机共创新模式。通过自然语言交互机制，业务专家可专注于知识体系构建、关键问题梳理与评测目标设定，AI 工程师则聚焦于技术实现与数据流程管理。借助 LLM 对

复杂语言指令的理解与转译能力链接知识与工具，降低双方协作门槛，实现跨专业高效协同开发，推动从“AI 辅助专家”向“专家增强 AI”的转型。另一方面面向企业多场景部署的智能中台结构，支撑制造多环节协同升级。平台采用高度模块化与解耦的架构设计，具备跨产线、跨流程、跨场景部署能力。在同一企业内部，不同产线或业务部门可按需部署平台子节点，并通过统一的接口协议接入中央智能中台，实现异地协同、数据互通、模型共享。该模式不仅满足制造企业在炼钢、轧制、热处理、质检等多环节的智能化升级需求，也具备良好的横向可扩展性，为行业内构建“可迁移、可组装、可控可管”的企业级 AI 基础设施提供了范式示范。

## 7.应用实效：

### (1) 应用场景

平台现已在特钢企业的炼钢、热处理、设备巡检等环节实现落地应用，典型用户包括生产调度员、设备巡检员、材料研发工程师等。系统通过自然语言交互接口，降低非 AI 背景员工使用门槛，广泛赋能一线操作人员与工艺专家，实现了从“辅助决策”到“协同优化”的能力升级。

### (2) 解决行业痛点情况

一是合金材料材料设计与成分预测模块支持从目标性能出发进行逆向设计，有效减少试验轮次，提升新材料开发速度与成功率。二是在炼钢测温场景中，通过部署炉温预测小模型，中包过热度降低 $2^{\circ}\text{C}$ ，避免钢水过热或温度波动导致的质量不稳定问题，可提升钢种开发和

生产过程的稳定性与可控性。三是在设备巡检环节，通过图像识别模型实现高炉主皮带异物检测，识别准确率达 92%，故障从无法及时发现缩短至毫秒识别，报警提示生产人员做出相应操作，降现场人力作业强度与环境高危问题。



来源：之江实验室

图 13 多场景领域检测与应用

### (3) 经济社会效益

以精炼炉温度预测场景为例，平台系统通过智能预测替代部分人工测温操作，并优化加热策略，使测温设备使用频次减少 50%，年均节省设备辅材消耗与能耗成本约百万元。平台部署后，产线事故发生率降低约 40%，产品合格率提升 2%，有效保障高端特钢的稳定生产与准时交付，显著提升企业整体运营效率与产品竞争力。

## （四）AI 驱动的碳捕集材料研发与验证平台

### 1. 案例实施单位：

中国华能集团清洁能源技术研究院有限公司

### 2. 案例背景：

在碳达峰碳中和目标推动下，碳捕获、利用与封存（CCUS）技术成为高排放行业减排的重要路径。然而，传统吸收剂开发周期长、效率低，阻碍了 CCUS 的大规模部署。为此，本项目聚焦吸收剂研发环节，探索 AI 赋能材料开发的路径，推动形成智能化、自动化、快速响应的碳捕集研发体系。

### 3. 案例概述：

项目构建了一体化智慧 CCUS 材料研发平台，融合分子建模、机器学习、自动化实验与高通量筛选技术，实现吸收剂从虚拟筛选、性能预测到实验验证的闭环流程。平台已应用于 CO<sub>2</sub> 吸收剂的结构-性能预测与新材料筛选，显著缩短研发周期，提高开发成功率。

### 4. 行业痛点：

碳捕集吸收剂研发普遍存在试错周期长、数据积累难、成本高、依赖专家经验等问题，亟需构建高效、智能、可扩展的研发范式。同时，缺乏统一平台协同推进算法、实验与产业验证，导致技术难以规模化转化。

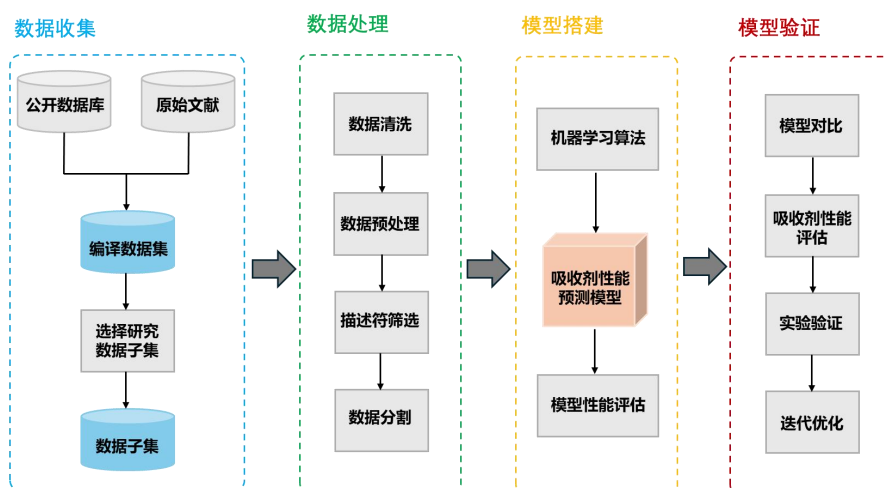
### 5. 案例详述：

碳捕集吸收剂是 CCUS 技术的核心之一，目前吸收剂性能和成本仍不满足大规模应用需求，开发更高效、更稳定、更经济的吸收剂配

方,可有效降低碳捕集成本,促进 CCUS 产业化进程。本项目围绕“AI 赋能碳捕集吸收剂研发”核心目标,构建了一套集数据整合、性能预测、高通量验证于一体的智能研发平台,实现碳捕集材料从分子结构设计到工艺适配的全流程数字化与自动化支撑。平台系统性整合机器学习算法、分子模拟工具与自动实验装置,具备预测-验证-反馈的闭环优化能力,显著提升研发效率和吸收剂性能表现。

### (1) 总体思路

项目设计遵循“数据驱动、模型预测、实验验证、持续迭代”机制,致力于破解传统吸收剂开发周期长、成本高、成功率低的技术瓶颈。研究从分子结构出发,构建高维描述符特征空间,利用机器学习模型预测其在特定工况下的 CO<sub>2</sub>吸收能力,再通过自动化实验平台进行快速验证与模型反馈迭代,实现精准筛选与性能优化。



来源: 中国华能集团清洁能源技术研究院有限公司

图 14 项目设计流程图

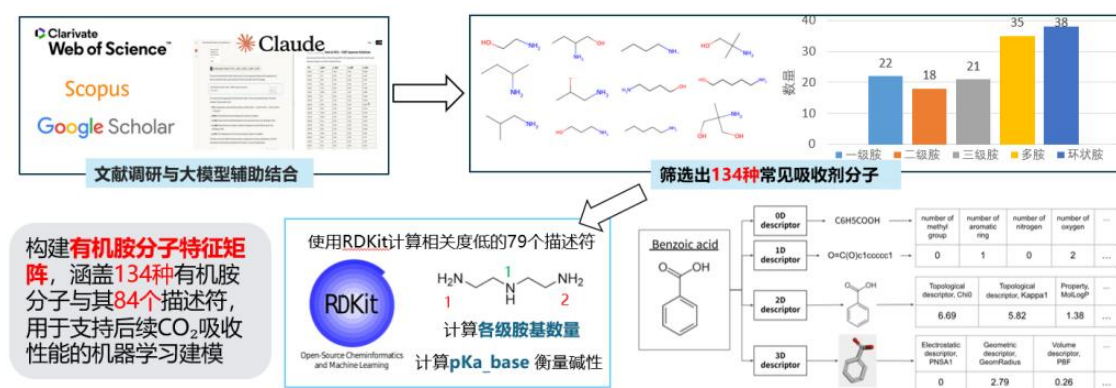
### (2) 软件平台

本平台构建了覆盖“数据-模型-实验-设计”全流程的碳捕集吸收剂智能研发体系。平台集成由 134 种有机胺构成的分子数据库与多维

描述符库，并融合 XGBoost、神经网络等 AI 算法，精准预测 CO<sub>2</sub>溶解度与吸收能力，模型具备 SHAP 可解释性。同时，自主研发的高通量自动化实验装置实现吸收性能的快速验证，效率提升 10 倍以上。通过反馈迭代与虚拟筛选系统，平台可根据目标性能反向推荐最优分子配方，形成“预测-验证-优化”的智能研发闭环，显著加速高性能吸收剂的发现与应用。

### (3) 技术方案

一是吸收剂分子数据库与描述符库。基于化学信息学工具，对 134 种主流有机胺提取分子结构特征，构建描述符矩阵；引入自定义特征，增强预测精度。通过大模型分析收集碳捕集领域吸收剂文献理论数据、实验数据。

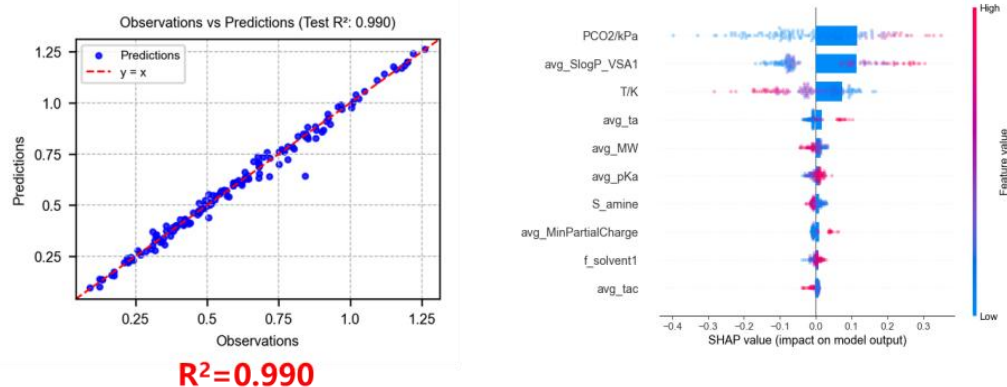


来源：中国华能集团清洁能源技术研究院有限公司

图 15 有机胺分子特征提取图

二是 AI 预测模型模块。采用 XGBoost、全连接神经网络(FCNN)、随机森林(RF)等算法，分别构建回归模型预测 CO<sub>2</sub>溶解度与吸收能力 ( $\alpha$ CO<sub>2</sub>)。模型通过十折交叉验证进行稳健评估，融合 SHAP 解

释框架识别关键影响因素，实现预测的可解释性。



来源：中国华能集团清洁能源技术研究院有限公司

图 16 回归模型预测图和 SHAP 框架图

三是自动化实验模块。自主开发一套碳捕集吸收剂高通量自动化开发系统，集成精密配液系统、物性参数检测模块、吸收反应模块、在线气体分析与自动记录系统、自动化仪器检测，支持温压可调、吸收反应实时监控、反应瓶重复利用，较人为实验研发效率提升约 10 倍以上。



来源：中国华能集团清洁能源技术研究院有限公司

图 17 碳捕集吸收剂高通量自动化开发系统

四是反馈迭代与虚拟筛选系统。系统以分子结构为输入，结合已有吸收性能数据库与实验回流数据，基于物化特征构建多维特征矩阵，

并集成吸收容量、吸收速率等指标的历史数据，形成初步结构筛选库。在此基础上，系统可根据目标工况反向推导出满足性能要求的结构组合，优先输出具有可测性与可工艺适应性的候选配方，为实验阶段提供明确方向，形成反馈迭代。

#### (4) 关键指标参数

一是吸收剂数据库参数量达到 10 万以上；二是吸收剂关键性能预测模型在验证集中达到  $R^2=0.95$  以上；三是筛选出的 1 种新型胺类吸收剂经实测性能优于现有 MDEA/PZ(3mol/L/1mol/L) 基准体系；四是开发 AI 驱动的碳捕集材料研发与验证平台，实现吸收剂自动化配制与性能检测，新配方材料实验开发验证周期由原来的 6 个月缩短至 3 周，兼容其他液体配方材料开发。

### 6.创新情况：

#### (1) 技术创新

本项目紧扣碳达峰碳中和背景下的技术需求，聚焦 CCUS 关键环节中捕集端吸收剂的开发瓶颈，创新构建了面向工程应用的研发与测试一体化平台，并在机制、流程和服务体系上形成了一套具有推广价值的研发组织模式。传统碳捕集吸收剂开发依赖人工配制和逐批测试，不仅周期长、效率低，且易受操作误差影响。针对这一问题，项目以性能可控、效率可复制为目标，搭建了具备制备、吸收性能评价、稳定性测试等功能的高通量自动化开发系统，贯通了吸收剂研发从实验室设计到工程选型的全流程。

该系统具备多配方并行处理、数据自动记录与批次对比分析能力，

在维持实验精度的前提下大幅提升了实验通量，适用于碳捕集工程中常见胺类、混胺类、添加剂配方等类型的快速筛选和工况适配，为大规模结构组合提供了支撑。围绕这一系统，项目制定了统一的配方试验流程模板、数据记录模板、评估标准、数据标准，使得研发过程具备可追溯性，提升协同研发效率。

## (2) 模式创新

在服务机制上，平台可灵活部署于 CCUS 示范工程配套实验室、高校研究平台或企业研发中心，支持内部使用和对外服务双路径运行，具备结构筛选、性能验证、指标对比、方案推荐等多项功能。目前已在华能正宁 150 万吨/年 CCUS 示范工程中投入使用，为捕集系统的吸收剂选型、能耗优化和运行稳定性提升提供了关键支撑。项目整套模式为 CCUS 行业在吸收剂开发阶段提供了自动化、系统性、可落地的新的研发范式，具备良好的推广基础和应用前景。

## 7.应用实效：

本项目构建的 AI 驱动的碳捕集材料研发与验证平台，已在真实应用场景中完成部署和实用验证，显著提升了碳捕集材料的研发效率和筛选质量，助力关键技术工程化突破。目前该平台已开展与高校科研单位的推广意向交流，如中国科学技术大学、湖南大学、怀柔实验室等。当前，正加快推进一体化、智慧化 CCUS 实验室建设，系统整合吸收剂等关键材料研发、高通量实验验证、工艺验证与系统优化等关键环节，构建覆盖 CCUS 全研发链条的智能化、集成化研发体系。

从社会效益层面来看，该系统显著提升了碳捕集材料开发的规范

化水平和转化效率,为我国 CCUS 产业的加速落地提供了有力的支撑手段。上述成果为 CCUS 技术持续创新和产业化应用提供了具有示范意义的支撑平台,对推动行业技术进步和产业高质量发展发挥了重要引领作用。项目部分组件和结构参数模块已完成开源,推动科研成果共享,降低行业应用门槛;已申请国家发明专利 3 项,形成了具备可复制推广条件的核心技术成果,相关成果相关项目已获得国家发展改革委“两重”专项和北京市科委项支持,并获得人民日报报道。

展望未来,随着全国 CCUS 产业链加速构建和规模化示范项目密集启动,碳捕集环节对新型吸收剂的需求将持续增长,对研发速度、筛选准确性与实验验证能力也提出更高要求。本项目所形成的研发机制与平台体系,顺应了 CCUS 产业化对材料端和工艺端一体化协同的趋势,不仅可服务于现有以电厂烟气为主的应用场景,也可拓展至钢铁、氢能、化工、天然气净化等新兴领域,为我国构建技术自主、安全可控、经济可行的 CCUS 产业体系提供基础支撑和工具保障,具备广阔的推广前景和长远的战略价值。

## （五）材料显微图像智能分析及应用

### 1. 案例实施单位：

北京低碳清洁能源研究院

### 2. 案例背景：

在材料科学研究中，颗粒尺寸及其分布对材料性能具有重要影响。扫描电子显微镜（SEM）作为微观结构分析的重要工具，广泛用于微米和纳米尺度的颗粒表征，提供直观、真实的数据支持，帮助科研人员深入理解和优化材料结构。然而，传统分析方式依赖人工标注与统计，操作繁琐、效率低，尤其在处理不规则颗粒时易引入误差。为提升分析效率与准确性，有必要引入智能化手段对 SEM 图像进行高效和精准的分析。

### 3. 案例概述：

本案例开发了基于人工智能的 SEM 图像分析模型，采用 SAM 架构，基于低碳院积累的上万张 SEM 图像，针对多组参数进行调优和适配，识别准确率达到 95%，处理速度较人工提升 200 倍（以 500 个颗粒物计算）。同时，构建了端到端的智能分析平台，支持批量图像上传、一键处理和多维统计等功能。平台上线后已应用于超过 20 个研发项目，覆盖多个材料研究领域，累计处理图像超 2000 张，节省约 170 小时分析时间，显著提升了 SEM 图像分析的效率与准确性。

### 4. 行业痛点：

在材料研究中，SEM 图像分析是不可或缺的重要微观表征和分析手段。当前该分析仍普遍依赖人工标注和统计，面临效率低、主观

误差大等问题。尤其在处理大量、不规则形状颗粒时，人工方法耗时长、重复性差，影响数据分析的准确性与一致性。此外，传统流程缺乏标准化，难以满足科研和工业应用对高通量、高精度图像处理的需求。因此，有必要引入高效、智能的图像分析手段，以提升分析质量与处理效率，支撑材料研究的深入发展。

### 5. 案例详述：

微观颗粒尺寸及其分布的精确表征是评估材料性能的核心环节。SEM 凭借其高分辨率和直观成像能力，成为微观尺度材料分析的常用工具。然而，传统 SEM 图像分析依赖人工标注和统计，不仅效率低下，且易受操作者经验影响，尤其在处理大量图像或复杂颗粒形态时，误差率显著增加。例如，在催化剂研究中，颗粒的不规则形状和密集分布会导致人工标注耗时长达数小时，且难以保证数据的一致性。



来源：北京低碳清洁能源研究院

图 18 传统人工 SEM 图像分析流程

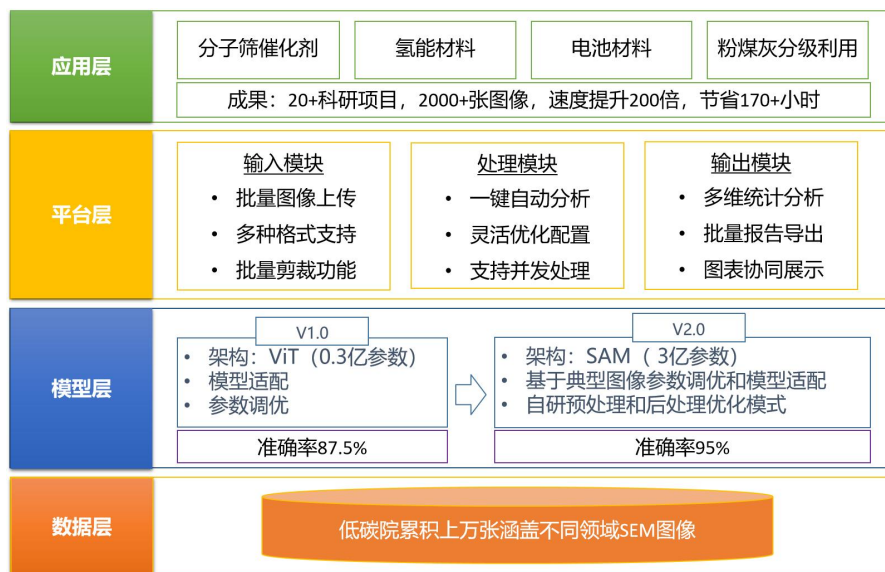
#### (1) 总体思路

针对上述问题，开发了一套显微图像智能分析模型及平台系统。总体思路围绕“算法优化+平台集成”展开，旨在通过人工智能技术实现 SEM 图像的高效、精准分析。技术路线分为三个阶段：模型定制化开发、平台功能构建与多场景应用。首先，针对 SEM 图像的特殊性（如对比度较低、颗粒边界模糊、背景噪声干扰等），从基础模型

架构入手，逐步完成从 V1.0 到 V2.0 的迭代升级。V1.0 版本基于 ViT (Vision Transformer) 架构初步实现了颗粒识别功能，但在复杂场景下存在边界识别不准确、小颗粒和模糊颗粒漏检等问题。为解决这一瓶颈，V2.0 版本引入 SAM (Segment Anything Model) 架构，经过定制化适配调优，并结合自研的预处理与后处理模块，显著提升了模型的泛化能力和分割精度。在适配调优过程中，基于低碳院积累的上万张图像，筛选了涵盖不同材料类型的典型 SEM 图像作为基础数据，进行模型参数调优，最终使识别准确率从 V1.0 的 87.5% 提升至 95% 以上。

## (2) 软件平台

基于上述模型构建了一套端到端的材料显微图像智能分析平台。平台采用模块化设计，核心功能包括：批量图像上传（支持多格式文件导入）、一键自动处理（无需人工干预即可完成多张图像分割与统计）、参数优化配置（可根据不同材料颗粒特征调整分割能力）、多维度结果输出（颗粒数量、平均尺寸、尺寸分布直方图等可视化图表）。平台界面设计简洁直观，用户仅需上传图像后，系统最快可在数秒内生成包含颗粒统计信息的完整报告。例如，在粉煤灰研究中，平台可在 10 秒内处理单张包含约 500 个颗粒的 SEM 图像，并输出直径和面积等分布统计结果和曲线，而传统人工标注需耗时 20 分钟以上。

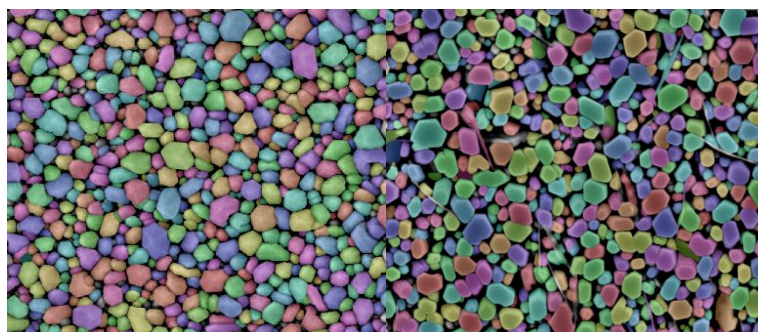


来源：北京低碳清洁能源研究院

图 19 材料显微图像智能分析模型及系统架构图

### (3) 技术方案

在颗粒识别任务中，模型能够对不同尺寸的颗粒均能实现精准识别，且对形状不规则颗粒（如椭圆形、多边形、立方体）的识别准确率超过 95%。在处理速度上。支持并发处理，可同时处理多张图像而无性能衰减。此外，平台通过引入多维度统计功能，能够自动生成粒径分布直方图、D50/D90 计算值及颗粒直径和面积分布图，为材料性能评估提供直观数据支持。例如，在工业分子筛催化剂研发中，根据颗粒尺寸进行分级利用具有重要意义。其中小颗粒分子筛，特别是纳米级别的颗粒，由于其较短的孔道，可以显著提升反应物和产物的扩散速率，从而大幅提高催化剂的性能和寿命。借助于本案例的平台，可以实现不同尺寸催化剂颗粒尺寸的快速统计和分类，助力分子筛催化剂的分级利用。



来源：北京低碳清洁能源研究院

图 20 v2.0 模型不规则颗粒识别可视化渲染结果

#### (4) 关键指标参数

目前该模型已为 20 个科研项目提供技术支持，涵盖包括电池材料、煤化工催化剂、氢能材料、碳纤维材料等多个研究领域，累计处理图像超 2156 张，模型调用次数达 2634 次，累计节省约 170 小时人工成本。

### 6. 创新情况：

#### (1) 技术创新

在模型研发过程中，术框架从 v1.0 的 ViT 过渡到 v2.0 的 SAM，同时结合上万张 SEM 图像的数据积累，实现了针对材料科学场景的定制化优化。原始数据涵盖电池材料、煤化工催化剂、粉煤灰等多种材料类型图像，颗粒形态包括规则球形、针状、多边形以及不规则形状等典型特征，为模型调优提供了丰富的样本基础。通过构建覆盖多尺度、多噪声水平的基础数据集，使模型具备适应材料颗粒边缘模糊、背景干扰等复杂场景的识别能力。在此基础上，参数组合优化框架针对多个关键指标进行了多参数精细化调整，最终使模型在识别准确率、响应速度及鲁棒性方面实现显著提升。这种基于行业数据的定制化适配，使模型能够精准捕捉材料微观结构特征，为材料性能预测提供可

靠数据支撑。例如，在粉煤灰利用研究中，模型通过分析颗粒尺寸分布，快速筛选出不同层级的粒径区间，显著缩短了实验周期，助力粉煤灰分级利用的相关研发。

**在技术实现层面**，预处理与后处理模块的设计进一步增强了模型的实用性。预处理阶段包含多种不同的图像增强功能，包括动态对比度增强（通过 CLAHE 算法调整局部亮度分布）、噪声抑制（结合多尺度高斯滤波与频域降噪）、多尺度边缘锐化（保留颗粒边界细节）、尺寸剪裁（提高处理效率和精度）以及图像归一化（消除成像条件差异）等。这些增强处理有效降低了图像质量波动对识别结果的影响。后处理环节则采用人机交互式操作模式，针对模型难以自动分割的粘连颗粒或漏检区域，用户可通过标注误识别区域优化结果，仅需少量人工干预即可完成修正。例如在催化材料分析中，通过对比度增强和少量后处理操作，解决了因材料表面反光导致的局部过曝问题，以及由于样本杂质噪声导致的背景干扰问题，使颗粒边界更清晰可辨，结果更可靠。

## （2）模式创新

**在应用创新层面**，该模型通过技术适配实现了多领域的拓展与优化。**一方面**，模型通过行业数据积累与参数优化，逐步覆盖电池材料、氢能催化剂、碳纤维、粉煤灰等多种材料类型，支持从纳米级颗粒到微米级纤维的全尺度分析。例如，在碳纤维材料研究中，通过统计纤维直径分布，为工艺参数调整提供数据依据；在氢能催化剂研发中，则通过分析颗粒尺寸分布，辅助研究人员优化催化性能。**另一方面**，

通过端到端应用模式，支持科研团队按需调整分析参数，并提供参数模板库以适配不同材料场景，降低了技术应用门槛。

### 7.应用实效：

材料显微图像智能分析模型已在多个材料科学研究领域实现应用，覆盖电池材料、煤化工催化剂、氢能催化剂、碳纤维及粉煤灰利用等方向，覆盖科研机构及化工企业实验室，验证了其在材料分析中的实际价值。平台上线后，已处理超过 2156 张 SEM 图像，模型调用次数达 2634 次，支持超 20 个科研项目的开展，累计节省约 170 小时的人工标注与统计时间。

一方面智能模型替代传统人工操作，解决了 SEM 图像分析中颗粒密集分布、形态多样性及团聚噪声干扰等问题。例如，在研发 MTO/ZSM-5 分子筛催化剂过程中，模型解决了传统方法在测量分子筛 SEM 图像中颗粒尺寸时因团聚或光学干扰产生的误差问题，达到对分子筛的孔道结构、晶体形貌及颗粒分布进行高精度表征，更全面地了解其催化活性位点分布和反应路径，从而为优化催化剂的制备工艺，提高其在甲醇制烯烃等反应中的选择性和稳定性提供深层次的洞察，显著提升科研工作效率的效果。

另一方面在粉煤灰分级利用研究中，基于模型的高效准确智能分割能力，通过分析颗粒尺寸分布，快速筛选出适合制备高性能材料的粒径区间，解决了传统方法在测量粉煤灰颗粒尺寸时因团聚或光学干扰产生的误差问题，达到高分辨表征粉煤灰的微观形貌、颗粒分布、孔隙结构等特性，准确率达到 95% 以上，远高于人工标注和统计。另

外，基于模型的应用平台支持批量处理（单次可分析超 100 张图像）及灵活的参数配置（如阈值调整、粒径范围筛选），通过多图批量处理、分析和筛选，可在 10 秒内处理单张包含约 500 个颗粒的 SEM 图像，并输出直径和面积等分布统计结果和曲线，而传统人工标注需耗时超过 20 分钟，分析速度提升至人工的 200 倍，帮助科研人员全面快速地了解其不同利用途径中的适用性，为实现粉煤灰的高效分级利用，开发高附加值产品提供更详细、准确的材料特性信息。

模型上线以后，集团准能科创中心、中科院物理所、山西煤化所、广西大学、德国弗劳恩霍夫材料回收与资源策略研究所（安卡院士团队）、ThermoFisher（原 FEI）、蔡司、日立、国仪量子等国内外多家单位应邀试用，样品涉及催化剂、电池材料、多孔材料、粉煤灰、金属材料、合金、陶瓷、细菌等多种类别，分析结果得到用户的肯定和好评。

上述应用表明显微图像智能分析模型能够满足不同材料场景的实际需求，显著降低了 SEM 图像自动和高通量分析的技术门槛，提升了科研工作效率。

## (六) 统一 NMR 智能解析在化学与材料科学的应用

### 1. 案例实施单位：

厦门大学、深势科技（宜宾）有限公司

### 2. 案例背景：

核磁共振（NMR）是化学、生物学和材料科学中用来解析分子结构和动力学的重要工具。准确预测 NMR 化学位移对光谱解析、结构修订以及构型确定具有重要意义。因此，研究结构与光谱之间的关系至关重要，它不仅影响光谱解析的精确度，还直接关系到分子结构的正确解析及材料设计的优化。

### 3. 案例概述：

本项目提出了 NMRNet，一个基于 SE(3)Transformer 架构的深度学习框架，采用预训练与微调的范式，旨在精准预测液相和固相 NMR 的化学位移。通过对原子环境的建模，NMRNet 能够统一处理不同状态下的 NMR 数据。此外，项目还构建了一个全面的基准数据集，并通过多个液相与固相 NMR 数据集验证了 NMRNet 的优越性，展示了其在实际应用中的鲁棒性和有效性。

### 4. 行业痛点：

传统的 NMR 化学位移预测方法，在面对复杂分子结构时，往往难以兼顾准确性与效率。同时，现有的深度学习框架多针对液相或固相 NMR 进行单独设计，缺乏通用性。另一方面，实验性 NMR 数据较为稀缺，这对模型的泛化能力构成了挑战。结构与光谱之间的复杂关系使得从分子结构预测光谱成为一项极具挑战性的任务。

## 5.项目获奖、专利授权及获得投融资情况

发表学术文章一篇：《Toward a unified benchmark and framework for deep learning-based prediction of nuclear magnetic resonance chemical shifts》

## 6.案例详述：

### (1) 总体思路

本项目旨在打破传统的液相与固相核磁共振（NMR）化学位移预测各自为政的局面，建立一个统一的深度学习预测框架。总体思路是采用“预训练-微调”的范式：首先利用海量的分子和晶体三维结构数据进行自监督预训练，使模型学习到通用的原子环境表征，从而缓解实验 NMR 数据稀缺的问题并提高模型的泛化能力。随后，针对具体的液相或固相 NMR 数据进行监督微调，以适应不同的物理状态和化学环境，最终实现对多样化化学系统（从有机小分子到复杂固体材料）的高精度、高效率光谱预测。

### (2) 软件平台

本项目开发的 NMRNet 是一个集成了数据准备、预训练、微调和推理四个模块的综合性深度学习工具，现已集成至人工智能应用电化学实验室(AI4EC)在线 app 平台。

核心能力方面，能够统一处理液相、固相甚至气相状态下的原子级任务，进行精确的 NMR 化学位移预测。应用场景方面，不仅提供数值预测，还支持峰归属 (Peak Assignment)、构型确定 (Configuration Determination) 和结构分析等关键应用，帮助研究人员深入理解结构

与光谱的相关性。用户服务方面，为提高易用性，平台包含一个基于 Web 的工具，使研究社区能够简便地进行化学位移预测。数据资源方面，平台内置了经过清洗和验证的标准化基准数据集（如 nmrshiftdb2-2024），为模型评估提供了可靠资源。

### （3）技术方案

一是模型架构方面，采用基于 SE(3)Transformer 的模型架构作为核心，该架构能够有效捕捉原子的三维空间信息和局部化学环境。二是预训练机制方面，利用超过 480 万个结构进行大规模自监督预训练。三是采用掩码原子预测（Masked Atom Prediction）和三维坐标恢复（3D Position Recovery）作为预训练任务。引入对数加权重采样（Log-weighted Resampling）技术解决数据集中元素分布不平衡的问题。四是差异化微调策略，液相 NMR 基于单分子结构进行微调，输入原子类型列表和成对原子距离矩阵。固相 NMR 引入周期性边界条件（PBC），采用 6Å 的截断半径来定义局部原子环境，以准确描述晶体中的长程相互作用。五是推理与应用，在推理阶段冻结模型参数，将其应用于各类下游任务，如利用预测的化学位移与实验值的比对来进行立体化学异构体的鉴定。

### （4）关键指标参数

在液相 NMR 预测精度方面，在实验场景中验证，NMRNet 的  $^1\text{H}$  NMR（氢谱）平均绝对误差（MAE）为 0.18 ppm， $^{13}\text{C}$  NMR（碳谱）平均绝对误差（MAE）为 1.09 ppm。在固相 NMR 预测精度方面，在 P2 型  $\text{Na}_{2/3}(\text{Mg}_{1/3}\text{Mn}_{2/3})\text{O}_2$  材料中， $^{23}\text{Na}$  的化学位移预测均方根误差

(RMSE) 为 48 ppm; 对于分子晶体的<sup>1</sup>H,<sup>13</sup>C,<sup>15</sup>N,<sup>17</sup>O 固体 NMR 数据, 预测结果与密度泛函理论 (DFT) 计算值表现出高度相关性。

## 7. 创新情况:

### (1) 技术创新

**在数据层面**, 一方面项目建立了一个全新的、标准化的基准数据集 nmrshiftdb2-2024。该数据集通过对现有数据库 (nmrshiftdb2) 进行手动筛选、纠正错误条目和广泛的清洗验证而创建。新的数据集相较于旧版 (nmrshiftdb2-2018) 包含更多的原子、更广泛的元素范围和更复杂的结构, 为评测模型在更具挑战性、更接近真实场景的任务中的表现提供了可靠基准。**另一方面**为预训练收集了包含超过 480 万个结构的广泛结构数据库。大规模的预训练旨在缓解实验性 NMR 数据稀缺的问题, 并提高模型的泛化能力。

**在模型层面**, 提出了一个名为 NMRNet 的统一深度学习框架。该框架的核心是采用了 SE(3)Transformer 架构, 使其能适应原子级的表示, 并扩展到固相、液相和气态体系。现有的深度学习框架通常只针对液相或固相 NMR, 限制了其通用性。NMRNet 通过共享的架构和统一的框架, 首次实现了对液相和固相系统原子环境的有效建模。

**在算法层面**, 一是采用了“预训练与微调”范式。预训练阶段使用大型结构数据库进行自监督学习, 任务包括掩蔽原子类型预测和 3D 坐标重建。预训练使模型能学习到原子环境的鲁棒表示。对比实验证明, 使用预训练权重可以显著降低预测误差, 尤其是在训练数据量有限时 (如<sup>19</sup>F)。二是针对固相系统, 开发了基于截止半径的局部环境

表示策略，并结合了针对晶体数据库的专门预训练。该策略被证明在固相 NMR 预测中表现最佳，优于仅使用晶胞或未针对晶体优化的预训练策略。**三**是在晶体预训练阶段，实施了对数加权重采样方法。此举解决了预训练数据中元素分布极度不均的问题，确保模型能有效学习稀有元素的化学环境。

**在价值对齐（可解释性）层面**，利用 Transformer 模型的多头注意力机制来可视化和分析原子间的相互作用。该技术提供了一种无需先验化学知识即可分析原子级相互作用的工具，增强了光谱解析和对复杂材料结构-性质关系的理解。

## (2) 模式创新

**在解决方案创新层面**，NMRNet 提供了一个集成的四模块框架（数据准备、预训练、微调、推理）。该框架不仅能进行高精度预测，其推理模块还被设计用于解决实际的化学问题。该框架超越了数值预测，将深度学习模型的能力从单纯的光谱预测扩展到更复杂的下游化学应用，包括：光谱峰归属、构型确定与结构修正与手性异构体鉴定。

**在服务模式创新层面**，项目成果以基于 Web 的工具和在线服务的形式提供。这种模式增强了模型的可及性，为广大的研究社区（包括非 AI 专业的化学家）提供了简化的化学位移预测服务。

**在应用范围拓展层面**，NMRNet 是统一框架，其应用范围同时覆盖了液相和固相 NMR。该模型被证实可广泛应用于多样化的化学系统，包括天然产物、大型有机分子（多达 150 个原子以上）和固体材料，展示了强大的泛化能力和外推鲁棒性。

## 8.应用实效:

### (1) 案例应用场景

NMRNet 框架已在多个关键科研与工业场景中实际落地应用。其核心应用是高精度的核磁共振 (NMR) 化学位移预测, 并在此基础上拓展至多种复杂化学任务: 一是光谱解析与结构阐明。作为化学、生物学和材料科学中的基础工具, 辅助研究人员解析光谱, 修正或确定分子结构。二是构型与立体化学确定。解决传统光谱解析中极具挑战性的任务, 包括结构修正 (如 TIC10、nevirapine) 和手性 R/S 异构体的鉴定。三是光谱峰归属。将实验测得的 NMR 信号自动匹配到分子结构中的特定原子。四是材料设计与分析。应用于复杂固体材料的分析, 例如 P2 型阴极电池材料, 以及分析原子级相互作用以理解结构—性质关系。五是公共安全。应用于关键分子的识别, 如对五种神经毒剂的化学位移进行精准预测, 为应对新型化学威胁提供支持。

该案例的目标用户是化学、生物化学、材料科学以及药物研发等领域的科研人员、工程师和学生。目前, 该项目已通过在线服务实现了广泛部署。截至目前, 该应用已吸引超过 5,000 名用户订阅, 累计处理了超过 1,700 次预测任务提交, 显示出研究社区的高度活跃度和对该工具的实际需求。

### (2) 解决行业痛点情况

一是解决了准确性与效率的平衡问题。传统方法在面对复杂分子时难以兼顾准确与高效。NMRNet 在保持高精度的同时, 展现了对大型复杂分子的强大外推鲁棒性。二是解决了模型的泛化性限制。现有

深度学习框架通常局限于液相或固相之一。NMRNet 通过统一的 SE(3)Transformer 架构和预训练范式，首次实现了对液相和固相 NMR 预测的统一建模，并均达到顶尖水平。三是缓解了实验数据稀缺的瓶颈。实验 NMR 数据稀缺限制了模型的泛化能力。NMRNet 通过在超过 480 万个无标签结构数据上进行大规模预训练，成功学习了原子环境的鲁棒表示，显著提升了模型在数据稀缺任务上的性能。

### (3) 经济社会效益

社会效益方面，该框架为分析化学和结构化学领域提供了强大的 AI 辅助工具，推动了该领域的发展。它通过提供开放的工具和服务，降低了高精度计算化学的门槛，使全球研究者受益。同时，其在神经毒剂识别等领域的应用也展现了潜在的公共安全价值。经济效益方面，NMRNet 能够显著加速新材料（如电池材料）和新药物的研发周期。通过提供比肩甚至超越传统 DFT 方法的预测精度，它极大降低了研发过程中对昂贵算力和时间的依赖，为企业和研究机构节省了大量的研发成本。

### (4) 示范推广价值及现状

NMRNet 作为统一 NMR 预测框架的成功范例，具有极高的示范推广价值。其“大规模结构预训练+领域数据微调”的策略可被推广至其他光谱（如拉曼、红外）或材料性质预测任务。项目推广现状良好，获得了学术界和产业界的广泛认可。从学术影响力来看，相关论文发表后，已迅速获得超过 10 次引用，表明其学术价值和创新性得到了同行的认可。从产业与科研合作来看，项目已与多家企业、高校和研

究院所达成合作，将 NMRNet 框架集成到不同的科研和工业体系中，有效助理了科研进度。

#### (5) 开源情况

该项目秉持开放科学精神，成果已全面开放。NMRNet 框架的源代码已在 GitHub 和 Zenodo 上开源。目前，GitHub 仓库已获得超过 45 个星标 (star)，吸引了业界的关注。一方面是模型开源，所有预训练的模型参数均在 Zenodo 上公开，便于复现和二次开发。另一方面是数据开源，所有用于微调的处理后 NMR 数据集，包括新建立的基准，均已在 Zenodo 上发布，已获得超过 400 次下载和 900 次访问。

#### (6) 相关标准成果产出

本项目最重要的成果之一是建立了一个新的行业标准基准：一方面是数据集，项目通过对现有数据进行广泛的清洗、验证和手动纠错，建立了一个更全面、更可靠的 NMR 化学位移基准数据集 nmrshiftdb2-2024。另一方面是统一基准框架，NMRNet 框架本身也为评估跨越液相和固相的 NMR 预测模型提供了一个统一的基准平台。

## （七）基于科研智能的锂电电解液研发

### 1.案例实施单位：

嘉庚创新实验室、厦门智储大装置研究院有限公司、厦门依华智慧科技有限公司

### 2.案例背景：

本案例构建了一套人工智能驱动的锂电电解液智能研发系统。该系统通过集成高通量自动化制备平台、多模态在线表征技术和大模型，实现了电解液配方从“设计-配制-测试-优化”的全流程闭环研发。有效将研发人员从重复性劳动中解放，极大提升了实验通量和数据一致性，显著缩短了研发周期，为快速获得高性能、高适配性的电解液配方提供了全新的智能化解决方案，并首次在低温电解液领域实现了“干湿闭环”。

### 3.案例概述：

本案例构建了一套人工智能驱动的锂电电解液智能研发系统。该系统通过集成高通量自动化制备平台、多模态在线表征技术和大模型，实现了电解液配方从“设计-配制-测试-优化”的全流程闭环研发。有效将研发人员从重复性劳动中解放，极大提升了实验通量和数据一致性，显著缩短了研发周期，为快速获得高性能、高适配性的电解液配方提供了全新的智能化解决方案，并首次在低温电解液领域实现了“干湿闭环”。

### 4.行业痛点：

当前锂电池行业面临三大难题：一是效率瓶颈，电解液研发依赖

人工试错，配方空间巨大，导致研发周期长达数年，无法满足市场快速迭代需求。二是数据瓶颈，人工操作易引入误差，数据质量差、一致性低，难以构建高质量数据库以支持有效的构效关系分析和模型训练。三是能力瓶颈，传统手段仅能监测电压、温度等宏观参数，对析锂、副反应等微观安全隐患缺乏有效的原位、实时监测与预警能力，制约了电池安全性能的提升。

## 5.案例详述：

### (1) 总体思路

本系统围绕锂离子电池电解液材料研发的迫切需求，瞄准人工智能和高通量自动化实验技术对缩短其材料筛选与优化周期的显著优势，通过人工智能筛选、高通量自动化制备技术、电池自动化装配技术和材料电化学性能高通量表征深度融合，解决电解液材料选择和配方设计决策等关键问题，通过人工智能预测及高通量制备系统所获取实验数据，助力电解液材料筛选效率的提升。

### (2) 软件平台

项目组自主建设了基于科研智能的锂电电解液研发系统，实现了智能化自动化制备、能源电池工况表征和数据驱动算法的闭环研发，及实验室仪器设备的全链条国产自主可控，建立了面向产业的锂电电解液研发新范式。通过与国家能源实验室和龙头企业开展深度战略协同，共同构建了“基础理论—技术攻关—标准制定”三位一体的创新生态，为引领新能源材料研发智能化变革、服务国家能源安全战略提供核心支撑。

### (3) 技术方案

基于科研智能的锂电电解液研发系统由“手、眼、脑”三个模块组成。一是“手”代表电解液智能制备模块，配备智能机器人，可执行编程控制、长时间不间断工作，自动高通量完成电解液配置、表征检测定制实验。通过自动化集成和 AI 学习决策指导，有效缩短材料研发周期、节约人力成本，有效结合领域学科积累、数据分析挖掘、机械自动化，将直觉经验数据化，减少无效试错，实现电解液材料的理性设计和精准配置。二是“眼”代表能源电池工况表征模块，能够实时监测并反馈电池制造和运行过程中的关键参数，显著提高电池器件的制造良率，确保生产过程的精确性和一致性以及电池使用过程中的安全监测。三是“脑”代表闭环调控模块，通过引入“工况表征—精准解析—实时调控”智能闭环系统，实现储能锂离子电池异常早期预警和精准调控，革新电池的研究范式，推动电池技术向更高效、更安全的方向发展。

### (4) 关键指标参数

一是智能扣式电池组装设备单机通量大于 2500 枚/周，较传统人工效率提升 50 倍；二是无人干预情况下自动运行连续工作时长大于 240 小时；三是数据生产质量显著提升，在无数据筛选模式下克容量波动率（COV）小于 0.67%、库伦效率波动率小于 0.45%，而在启用 80%数据分析后，克容量波动率进一步降至 0.40%以下，库伦效率波动率更优化至 0.05%以内。

## 6.创新情况：

## （1）技术创新

创新性地实现了锂电池电解液“干湿闭环”研发范式，针对非常规温度条件下的电解液研发需求，通过自动化实验完成 40 组配方筛选，精准锁定四元电解液配方最优值，为极端工况下电解液的性能优化提供了高效解决方案，推动了电解液研发从传统试错向数据驱动的范式转变。

数据采集层面，本系统深度融合 AI 技术与自动化实验平台，实现了多源异构数据的统一采集与实时处理。依托自动化电解液检测平台（包括电导率、拉曼光谱等多维度表征设备），系统集成电压、电流、温度、压力及气体排放等关键参数，并融合电化学工作站、质谱仪等设备输出，构建覆盖“配方—结构—性能”的电池电解液多模态数据库。AI 驱动数据清洗模块自动剔除噪声与异常值，实现高通量、高一致性数据预处理，显著提升数据质量与可用性，为智能推荐模型提供可靠输入。

算法优化层面，融合单目标与多目标参数协同策略。基于贝叶斯优化框架，首阶段（第 1—3 轮配方）聚焦 4.4V 库伦效率单目标寻优，快速锁定局部最优解；次阶段（第 4—7 轮配方）进行常压容量、高压库伦效率、高压容量以及常压库伦效率的多目标参数协同优化，实现在广泛参数空间内进行配方性能搜索。材料推荐模型通过持续学习实验数据，实现了更高的推荐精度，突破传统试错研发的盲目性。

性能与安全监测方面，发展多模态谱学工况表征和传感方法，集成电化学阻抗谱（EIS）、在线电化学质谱（OEMS）和光纤传感技

术。EIS 精准解析电池内部电阻/电容动态变化,捕捉离子扩散异常等; OEMS 监测氢气、一氧化碳等副产气体浓度,预警热失控风险; 光纤传感器技术能够实时监测电池内部的温度、压力和应力等关键参数,防止安全事故的发生。这种“电化学—气体—力学”多模态联用方案,实现对电池从微观反应到宏观热失控的全链路监控。

## (2) 模式创新

创建研发流程革命的无人实验室新范式。系统通过构建具身智能研发平台,集成机械臂智能化配置、多模态表征设备监测及人工智能决策,形成“配置—表征—优化”闭环,显著缩短研发周期,推动电解液研发从“试错法”转向“理性设计”,形成可复制的智能化研发框架。

**一方面建立产学研深度协同的产业化加速机制。**联合国家实验室协同开展国家级战略研究,参与筹建能源领域国家实验室体系创新联盟,开发适用应用场景的智能优化算法,赋能从假设提出到实验优化的干湿实验闭环;与龙头企业共建联合研究院,开展千万级科研项目合作。

**另一方面拓展跨领域的智能研发服务生态。**将电解液智能研发平台复用至机械加工切削液配方开发、氢能膜材料工艺等领域,形成覆盖新能源、高端制造的智能材料研发服务体系。

## 7.应用实效:

### (1) 目标用户及应用情况

与新能源龙头企业密切合作,共同搭建了基于 AI 算法的电池关键原料智能研发系统,实现原料的智能制备、自动表征、配方的智能筛选和迭代。目前,系统的无人化配置及在线表征硬件已通过企业转

阶段验收，具备每日 100 组配方及每种电解液配方 5 种表征数据的智能研发能力，集成拉曼光谱、电导、饱和蒸汽压和黏度等多维自动化表征手段，实现了数据的智能化处理与分析，有效支撑了高性价比电解液的快速研发。

## (2) 标准编制与成果输出

团队积极参与智能实验室国标体系建设任务。由本项目团队牵头编制的《智能实验室行业发展报告》也将于近期发布。

## （八）深势科技 Piloteye® 电池设计智能研发平台

### 1. 案例实施单位：

北京深势科技有限公司、深势科技（宜宾）有限公司

### 2. 案例背景：

电池作为现代社会的核心储能载体，已深度融入电动汽车、移动电子等战略领域。随着“双碳”目标的推进与新能源产业的爆发，市场对电池安全性、快充效率、循环寿命等提出了更高要求。然而，电池研发往往涉及从微观反应机理到宏观电芯结构设计的跨尺度科学难题，且传统研发模式受限于单一尺度研究手段，难以实现多物理场耦合的系统性突破。在此背景下，深势科技推出 Piloteye® 电池设计智能研发平台，助力电池材料开发与电芯设计的智能化升级。

### 3. 案例概述：

Piloteye® 电池设计智能研发平台是深势科技自主研发且全球领先的电池设计自动化平台。平台基于机器学习、跨尺度建模等先进算法，为电池企业提供从 **Read**-文献调研、**Design**-实验设计、**Make**-合成制备、**Test**-表征测试到 **Analysis**-分析优化的一站式智能解决方案。平台通过构建材料微观尺度到电芯宏观性能的跨尺度预测体系，打破传统研发的信息壁垒，实现研发效率与创新能力的双重提升。自推出以来，已与多家头部电池企业展开合作，成功应用于电极材料开发、电芯性能优化等项目。

### 4. 行业痛点：

当前，电池行业面临诸多痛点。**一方面**，传统试错法研发周期长、

成本高、创新难，难以快速响应市场对电池高速迭代的需求，尤其是近年来备受关注的电池的安全性、稳定性、快充效率、循环寿命及能否大规模储能等。**另一方面**，电池研发涵盖从微观反应机理到宏观电芯结构设计的跨尺度复杂过程，涉及材料相变、离子输运、物化稳定性、析锂检测等诸多难题，传统手段难以高效解决。此外，负极新材料体系开发等关键领域也存在技术瓶颈，亟待智能化工具加速研发进程。

### 5.项目获奖、专利授权及获得投融资情况：

奖项方面，Piloteye 电池设计自动化计算平台入选全球数字经济大会《2024 北京数字经济标杆企业成果集》、第七届数字中国建设峰会《2024 中小企业数字化转型典型应用案例》、36 氪《AI Partner·2024 年 AI 应用标杆案例》；专利授权方面，包括一种用于锂电池电解液模拟的工作流任务处理系统；一种用于构建极化力场模型的处理方法和装置。

#### (1) 总体思路

随着国内外新能源产业的高速发展，电池材料的设计研发成为制约产业进一步突破的关键环节。传统的研发模式依赖试错法，存在周期长、成本高、效率低等痛点问题。深势科技围绕这类电池领域企业研发的关键需求，主要针对电极材料、电解液以及电芯这三个方向的应用场景，推出 Piloteye® 电池设计智能研发平台，从创新算法、工程化及“最后一公里”落地等方面系统的赋能电池材料开发和设计，为电池材料开发者提供创新算法与行业 know-how 相结合、操作便捷的

平台类工具。

## (2) 软件平台

深势科技 Piloteye® 电池设计智能研发平台是基于 AI for Science 新范式，利用 Deep Potential 系列多尺度建模算法、Uni-Mol 分子大模型、DPA 原子大模型、Uni-ELF 配方大模型、具有 GPU 加速与自动微分功能的新一代电化学计算模型、电芯老化模式定量分析算法等一系列人工智能新技术，致力于突破电池研发难点的电池设计自动化平台。平台从材料性质参数出发，预测材料颗粒的物化性质，进一步模拟电极与电芯尺度的性能；同时对加工工艺进行理性建模，同步模拟加工工艺对电极、电芯性能的影响，提高了电池设计的精确度和可靠性，并大幅缩短创新到量产间的周期，推动电池材料“设计理性化”、“开发平台化”和“制造智能化”的实现。

## (3) 技术方案

深势科技 Piloteye® 电池设计智能研发平台整体架构从底层到顶层依次为数据库、模型库和应用平台，三层紧密相连、协同运作，聚焦电池材料从理论设计到实验验证的核心环节，为电池研发设计提供全流程的智能化升级服务。



来源：北京深势科技有限公司

图 21 深势科技 Piloteye® 电池设计智能研发平台整体架构

数据库层：由深势科技玻尔科研空间站的数据库提供服务，其公共数据库涵盖电解液分子、正极材料等多领域的丰富数据资源，包括但不限于文献、专利、表征、测试、工艺等数据，且在实时更新拓展中，面向电池研发领域提供数据检索、结构保存、分子 SMILES 导出等功能，为电池研发提供全面、专业的数据支持。

模型库层：利用 Deep Potential 系列多尺度建模算法，融合电芯老化模式定量分析算法、Uni-Mol 分子大模型、DPA 原子大模型、Uni-ELF 配方大模型等多种能力，突破电池研发设计研发难点。

应用平台层：由 AI 读文献、AI 做计算、AI 做实验三个模块支撑组成应用平台层，实现电池研发设计全流程的系统打通、加速闭环，直接解决传统电池研发依赖“试错法”、成本高、效率低、创新难等痛点问题。AI 读文献方面，由 Uni-SMART 科学文献大模型作为数据底

座，Uni-Finder 文献阅读和信息提取系统提供知识库检索、数据整合、交互问答等服务，对电池材料研发设计涉及的分子结构、关键信息等进行分析总结。AI 做计算方面，由原子尺度的 DFT 计算、分子尺度的分子动力学模拟等多尺度模拟手段和 Uni-Mol 分子构象大模型、DPA 原子大模型、Uni-ELF 配方大模型驱动，可实现第一性原理正极掺杂筛选评估、“百万级”电解液配方设计和筛选、高精度固态电解质预训练大模型性质预测、高精度高效的电芯短期性能预测等多项关键场景的计算实践。AI 做实验方面，由 SimpFine 数据建模与分析平台和基于 Uni-AIMS 表征大模型的智能化电镜图像分析软件支撑，涵盖数据建模、数据预测、采样优化，表征分析等多种功能并可根据电池研发设计需求自由组合。

#### (4) 关键指标参数

一是平台预测电解液分子、配方的输运性质、电化学性质可达到高达 95% 的预测精度，可将研发周期缩短至传统实验周期的三分之一；二是平台预测固态电解质不同掺杂方案的离子电导率，计算结果和实验数据吻合度较高，超过 90%，加速固态电解质研发周期；三是平台通过电化学参数辨识以及利用电化学模型进行电芯短期性能预测，计算效率提升 2-3 倍，替代 60% 实验测试；四是基于平台工艺优化工具，通过数据搜集—算法建模—采样优化—迭代反馈，使用主动学习+实验强耦合，经过少量的优化迭代提高材料的性能，将开发成本减少至一半；五是平台基于高精度的电镜图像识别算法，颗粒识别准确率 >

90%，边界形状准确度>95%，单张图片分析时间<1分钟，与人工分析相比效率提升约30倍。

## 6.创新情况：

### (1) 技术创新

一是基于 Uni-Mol 电解液分子模型、Uin-ELF 电解液配方设计模型以及分子动力学方法对电解液分子和配方物性进行高通量高精度的快速预测，实现配方的设计和筛选。目前已支持将电解液探索空间扩展至  $10^{20}$ ，性质预测精度高达 95%，配方设计研发周期缩短至 1/3。

二是在第一性原理精度下对电池正极材料进行全面性质计算和掺杂评估。基于 DeePMD、主动学习策略等先进手段，建立可描述广泛正极掺杂空间的量子力学精度的势函数模型，基于该势函数模型能够实现快速精确的材料性质探索，计算效率比主流 DFT 方法快 1000 倍。涵盖 27 种元素，12w+样本数的 DPA 固态电解质原子大模型，具有第一性原理预测精度。基于该模型能模拟复杂组分固态电解质的性质，筛选最优固态电解质配方。

三是通过高效的参数优化和并行计算算法，对电化学模型参数进行辨识，利用精度更高的电化学模型预测电芯短期性能，计算结果与输入实验测试结果具有定量级的可比性。

四是通过建立基于多模态数据的正极前驱体工艺性能预测模型并对前驱体工艺进行优化，生产成本降低 30%，研发效率提高 4 倍。除此之外还广泛应用到电池原材料配方设计、溶液性能调优、金属成分工艺设计、陶瓷工艺性能设计等多个领域。

五是利用深度学习算法准确识别 SEM 图像的形态、尺寸、组成和分布，自动化批量处理电镜图像，与实验形成有效迭代，提升识别分析效率。

## (2) 模式创新

### 一是解决方案创新，即多尺度模拟与 AI 融合的创新算法体系。

平台创新性地整合了从微观到宏观尺度的建模与仿真方法，将 DPA、Uni-Mol、Uni-AIMS 和电化学模型 AI 自动调参等一系列基于 AI for Science 原理和数据驱动的创新算法融入电池研发过程。弥补了传统通用性仿真软件中经验参数过多导致模型精度不确定的缺陷，实现了从微观材料性质计算到介观材料颗粒物化性质，再到电极电芯尺度电化学性能的“全链条”全局优化，提高了计算模拟研究电池的精度和可靠性。

二是服务模式创新，即端到端的一站式智能解决方案。平台聚焦电池研发设计的正极材料、CEI、负极材料、电解液、SEI、电极、电芯、工艺等多个关键环节，通过多场景端到端的 APP 应用，实现文献解析、智能计算与实验仿真三轮驱动，为电池全生命周期的开发设计提供一站式智能解决方案，并根据个性化需求提供定制化服务，将传统试错式的电池研发模式转变为基于理性设计的智能研发模式，助力产业升级。

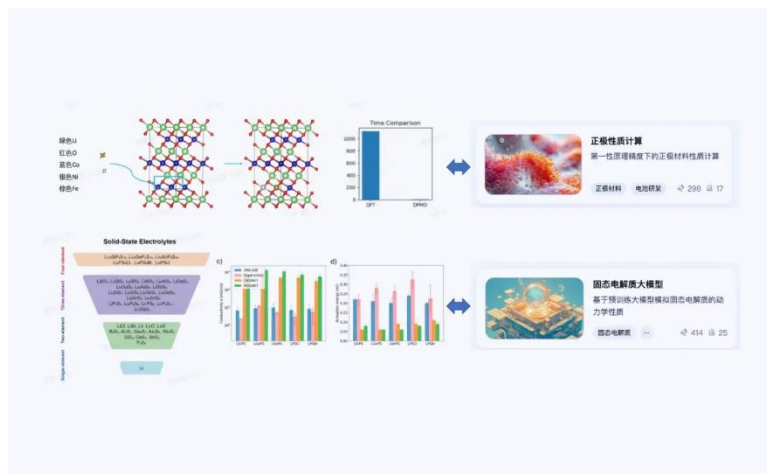
三是用户友好的交互设计与协同服务。平台采用了用户友好的设计理念，通过清晰的工作流、简单的输入和可视化交互的方式实现高通量、数据管理等一系列能力，让不同应用场景的研发人员都能轻松

上手，降低用户使用门槛，高效进行电池研发。

## 7.应用实效：

### (1) 正极掺杂计算

在电极材料的预测分析环节，平台基于 DeePMD 方法和 DPA 原子间势函数预训练模型，评估预测正极材料和固态电解质全掺杂空间的关键性质，准确度与实验可比，效率比主流方法快至少 1000 倍。平台支持产品级正极材料和固态电解质掺杂评估，交互式完成万级别的材料筛选，加速新材料研发周期缩短至原来 1/3。



来源：北京深势科技有限公司

图 22 正极材料和固态电解质掺杂评估

电池研发企业多氟多目标通过掺杂改性开发高性能钠电正极材料。深势利用“正极性质计算”APP 功能计算不同掺杂方案下正极材料的基本性质，快速筛选出性能表现最优的掺杂方案，经客户实验验证 50 周容量保持率提升 5%，加速正极材料开发。

### (2) 电解液配方筛选

基于 Uni-Mol 分子三维结果预训练模型，准确预测电解液全分子空间的多种关键物性，快速推荐候选目标分子。基于 Uni-ELF 配方模

型，几分钟内准确预测上千个电解液配方关键性质结果，快速设计和优化电解液配方。产品级电解液分子和配方性质预测和高通量筛选，一键式获得电解液配方性质报告，加速筛选流程，将百万级配方筛选缩短至天。



来源：北京深势科技有限公司

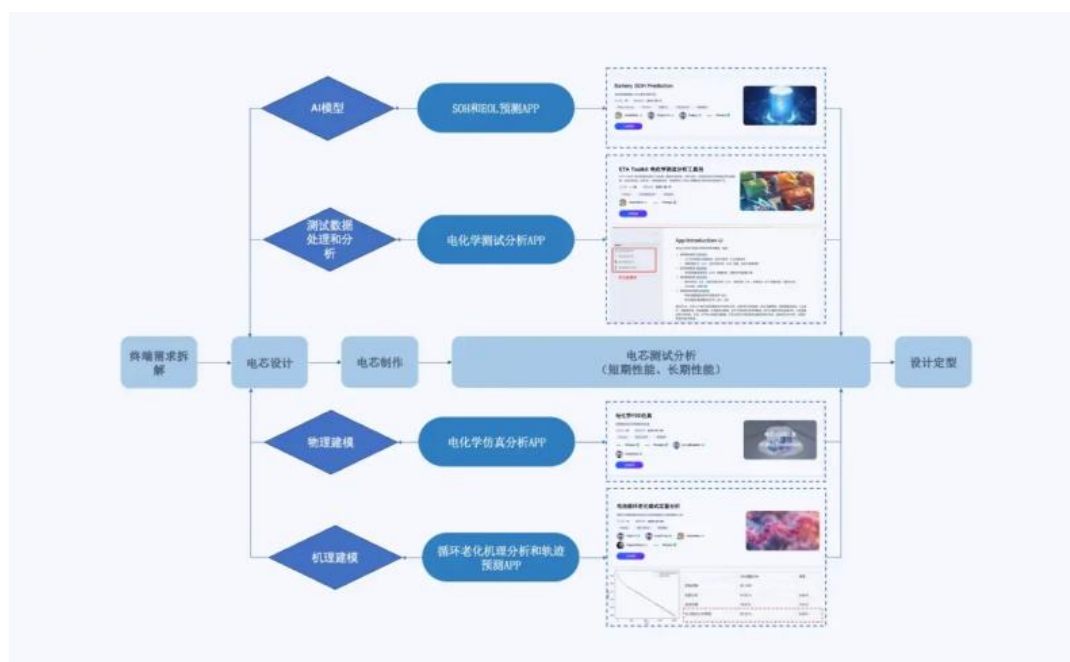
图 23 电解液分子及配方预测和筛选

头部新能源车企广汽埃安目标筛选耐高压的电解液配方。深势结合客户的实验经验，利用 Uni-ELF APP 分子生成以及物性预测功能，帮助客户从千万级空间中筛选十个符合目标需求的候选分子，经过客户实验迭代和验证，理论耐氧化电位大幅提升，电导率提升近 1 倍。

### (3) 电芯性能预测

平台基于电化学模型预测电芯性能，与实验测试结果进行详细对比分析。结合电芯机理模型，全面、准确、定量分析电池循环过程中各种衰减模式。利用 AI 数据驱动方式，基于少量电芯早期测试数据精确预测电芯健康状态。产品级电芯设计工具，涵盖“电芯仿真设计

->测试->分析->优化”全流程，一键式完成电芯性能的综合评估，将过去几个月的实验测试缩短至天级别。



来源：北京深势科技有限公司

图 24 电芯设计流程图

电池研发企业多氟多目标在已有的电芯数据基础上，快速评估电芯的长期性能，加速电芯开发迭代。深势引入 AI 机理模型有效处理已有的测试数据，对客户设计的电芯容量衰减情况进行全面的评测，预测结果与实验结果相比，误差不足 1%。

## （九）基于专用算力的电解液 AI 数据平台建设

### 1. 案例实施单位：

上海思朗科技股份有限公司、长江（孝感）3D 科学计算中心、上海科学智能研究院

### 2. 案例背景：

电解液是电池中离子传输的关键介质，作为新能源电池的核心材料之一，直接影响电池的能量密度、循环寿命、安全性及适用温度范围，这些关乎新能源电池的性能。由于电解液的结构复杂性以及配方成分的多样性，传统实验手段需投入大量的人力物力成本，且研发效率低下，难以满足市场对新型电解液材料的开发需求，亟需新的研发模式。

人工智能技术的兴起带来了新的思路，借助电解液数据集训练领域专用模型，成为加速电解液材料设计的重要破局手段之一。然而，当前既有的电解液数据集普遍存在配方覆盖不足、体系规模受限等局限性，难以支撑起 AI4S 在电解液研发领域落地，行业在呼吁更高质量的 AI-ready 的电解液数据集建设。

### 3. 案例概述：

本案例借助“天穹”材料计算专用科学计算机以及专为专用硬件架构设计的 Logic MD 软件，对不低于 10000 组电解液体系进行长时间大尺度的分子动力学模拟，其中单体系模拟时间至少 25 ns，模拟尺寸达到 10 万原子级别，确保模拟计算数据结果具备统计稳定性与物理可靠性；针对每组模拟轨迹，提取并计算其密度、黏度、扩散系

数、电导率、介电常数、结合能等关键物理化学性质；针对上述得到的电解液性质数据，进行智能化标注管理，并构建多模态检索功能、元数据管理功能、数据资产管理功能，形成便捷的对外服务接口。同时，利用上述电解液数据训练先进的机器学习模型，建立从电解质的分子结构（或配方组成）到其宏观物理化学性能的精准、快速的定量构效关系模型，实现近乎实时地预测新分子或新配方的性能。

#### 4.行业痛点：

**电解液设计面对巨大的挑战：**一是化学空间庞大，潜在的溶剂、盐、添加剂组近乎无限，传统试错法效率极低；二是多目标很难同时优化，单一模型仅预测单性能（如电导率），缺乏统一框架协同优化多指标，现有模型仅适用特定电解液类别，泛化性差；三是实验数据分散且规模小，覆盖的化学空间有限，同时缺乏微观结构信息，难以预测配方设计所需的多种物化性质。与此同时，国内外现有电解液数据库存在配方覆盖不足、微观信息不足、体系规模较小等局限性，这些缺陷共同导致无法满足令人满意的电解液 AI 模型训练需求，电解液研发陷入低效循环，亟需构建覆盖完整化学空间的高精度数据库以突破瓶颈。

#### 5.案例详述：

##### (1) 总体思路

借助“天穹”材料计算专用系统，对不低于 10000 组电解液体系进行长时间大尺度的分子动力学模拟，其中单体系模拟时间至少 25 ns，模拟尺寸达到 10 万原子级别，打造面向 AI for Science 领域的高精度、

高可靠性电解液分子动力学数据库，显著提升电解液分子动力学模拟的精度与效率，以此为电解液 AI 大模型训练提供高质量数据支撑，突破传统电解液研发的瓶颈局限，实现从经验驱动向数据驱动的根本性跃迁，大幅缩短新材料研发周期，有效降低研发成本。

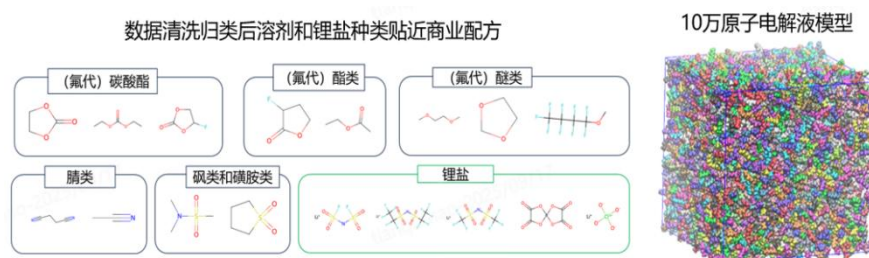
## (2) 软件平台

平台以构建“算力-数据-模型”三位一体的科研创新底座为核心战略目标，依托思朗科技全球领先的专用科学计算机“天穹”与上智院数据治理平台的深度融合，实现从科学计算、数据治理、质量控制到模型训练的全链条能力升级。其中，思朗科技“天穹”专用科学计算机基于 512 颗 MaPU 架构芯片组成的  $8 \times 8 \times 8$  立方体计算网络，辅以自研的 SL-link 通讯协议，实现分子动力学模拟效率较传统超算快 2 个数量级以上(达 5,000-10,000ns/天)，为海量电解液体系计算提供高效算力支撑。上智院提供的数据治理模块涵盖数据采集、清洗、标注、存储、验证的全流程治理体系，形成“计算加速-数据治理-模型训练”的完整技术闭环。这一创新平台通过算力与数据的深度协同，显著提升数据库精度与可用性，为行业构建可复用的科学智能基础设施奠定核心基础。

## (3) 技术方案

配方清洗和数据生产方面，针对行业应用需求与文献数据库，系统收集并清洗典型商用或学术关注度高的溶剂分子及锂盐配方组合，形成配方清单。利用“天穹”专用科学计算机对不低于 10000 组电解液体系进行长时间大尺度的分子动力学模拟，单体系模拟时间为至少

25 ns，模拟尺寸达到 10 万原子级别。为了确保模拟结果具备统计稳定性与物理可靠性，需对部分代表性电解液体系性质数据进行实验测量，基于实验数据对现有力场模型的精度进行评估并进行调优。



来源：上海思朗科技股份有限公司

图 25 高精度锂电池电解液配方分子动力学数据库构建示意图

关键物化性质提取方面，针对每组模拟轨迹，提取并计算其密度、黏度、扩散系数、电导率、介电常数、结合能等关键物理化学性质，并对得到的模拟计算数据进行集成标注、系统治理，构建结构化数据库。

电解液机器学习模型训练方面，利用上述高质量电解液数据集可训练先进的机器学习模型，建立从电解质的分子结构（或配方组成）到其宏观物理化学性能的精准、快速的定量构效关系模型，实现接近实时预测新分子或新配方的性能，为研发人员的快速迭代和配方优化提供强大且可靠的智能预测工具。

#### (4) 关键指标参数

一是电解液体系配方数量不少于 10000 组，每组模拟体系原子数量不低于 10 万个，每组体系的分子动力学轨迹时间长度不低于 25 ns；二是确保 95% 以上的数据资产在完整性、规范性、可追溯性、可用性方面保持一致；三是完成标注平台产品能力部署，支持解析电解质数

据领域核心文件格式，包含电化学测试数据（CSV、Excel）、分子式 smiles 等，确保支持解析的文件格式不少于 3 种；

## 6. 创新情况：

### (1) 技术创新

一是全新的材料计算架构：形成面向材料计算的专用算力架构，支持 10 万原子体系规模的超快速模拟计算，突破传统计算平台的规模限制，为电解液研发提供前所未有的计算效率。

二是全新的电解液数据库：实现 10 万级电解液大体系模型仿真计算，可提供更多高价值的微观溶剂化结构信息，达到配方级完整电解液性质功能信息覆盖；仿真力场基于实验数据优化、验证，电解液数据精度高，数据质量更加可靠；覆盖的配方种类数多，AI 成熟度高，可有效支撑电解液 AI 模型训练。

三是一体化全链路设计：突破传统数据平台单一功能的局限，创新性地 将数据资产管理&治理、对外服务体系、集成标注能力三大核心模块深度融合，构建从数据采集、处理、标注到服务输出的完整闭环体系，实现科学数据全生命周期的统一管理。

四是多模态智能检索：针对科学数据的复杂性和多样性，创新性地融合文本、图像、数值、结构化数据等多种模态的检索算法，支持化学分子结构相似性检索、实验图像内容检索等专业化检索方式，实现跨模态语义关联和智能聚合展示。

五是任务驱动的智能标注调度：创新性地采用任务拆分算法和智能调度机制，将复杂标注任务高效分解为细粒度单元，支持基于标注

人员技能和专业领域的智能分配，实现标注工作的并行化处理和动态优化管理。

## (2) 模式创新

一是构建行业级数据基础设施，重塑新材料研发范式：项目依托具有完全自主知识产权的科学计算核心技术和软硬件产品体系，在分子动力学等关键领域实现 2-4 个数量级的计算效率提升；基于此构建覆盖 10,000 组电解液体系的标准化数据库，提供高精度、大规模、长时间的分子动力学模拟数据，形成面向 AI 模型训练的高质量数据基础设施。这不仅为电解液研发、AI 辅助材料设计等提供核心数据支撑，更通过构建“算力-数据-模型”三位一体的科研创新底座，有效支撑我国在新型电池设计乃至新材料研发等战略领域形成技术代差优势，是夯实我国科技创新基础能力的关键举措。

二是推动 AI for Science 的产业化落地：项目成果将有效解决 AI 在材料科学领域应用的数据瓶颈问题，为 AI 大模型提供关键训练数据，解决 AI 在材料科学领域应用的数据瓶颈问题，推动 AI 在电池材料研发、电池管理系统优化等领域的应用，将传统“设计-合成-测试”流程重构为“数据驱动-AI 预测-实验验证”智能闭环，实现研发效率的大幅提升。

三是推动科学数据要素化和生态化：项目成果将推动材料学科学数据要素化和生态化。一方面，通过大规模建设物质科学领域的多模态语料库，显著增强科研数据供给能力；另一方面，通过加快科学语料的分类和有序开放共享，培育和壮大科学数据的种子用户群体，形

成广泛应用与价值转化的基础，进一步加速科研范式由算力驱动向数据价值驱动跃迁。

## 7.应用实效：

### (1) 案例应用场景

**电池研发与优化方面**，一是通过分析锂离子电池电解液成分（如溶剂、锂盐、添加剂）与性能的关系，优化配方设计。二是辅助开发宽温域（-40°C~80°C）极端环境电池电解液。**材料科学研究方面**，一是通过分子动力学模拟数据，辅助探讨验证电解液离子传输机制；二是筛选高稳定性、低成本的溶剂或功能性添加剂。**机器学习与AI应用方面**，训练模型预测电解液性能（如离子电导率），辅助特定性质的电解液设计开发，加速新材料发现。目前，该电解液数据集已上架国家新材料大数据中心，面向公众服务全国范围内的电解液乃至电池材料开发科研工作者。

### (2) 目标用户及当前用户情况

目标客户为国内大型动力电池生产企业，以及国内诸多涉及电池材料研发的高校/科研院所。

### (3) 解决行业痛点情况

想要开发一款性能优异的电解液十分困难：化学空间庞大，潜在的溶剂、盐、添加剂组合近乎无限；加之现有可利用的电解液数据集往往覆盖的化学空间有限，缺乏足够体现微观结构的信息，导致难以预测配方设计所需的多种物化性质，也无法满足AI模型训练需求；同时，传统试错法效率极低，需要耗费大量的时间成本和经济投入，

难以满足现代更高性能电池技术迭代对于电解液的研发需求。行业亟需构建覆盖完整化学空间的高精度数据库，以及基于仿真模拟+AI技术手段的电解液研发新模式。

基于此背景，本项目通过强大的 3D 科学计算机对众多电解液体系进行大体系、长时间的分子动力学仿真模拟计算，从而得到高精准确性、覆盖足够广度的高质量电解液数据集。一方面，通过进一步揭示了电解液微观结构层面关键性“构效关系”信息，另一方面能够有效支撑电解液 AI 大模型的训练开发，为 AI4S 在电池材料领域的落地实施打通了道路，最终助力实现新型电解液材料的高效开发。

#### (4) 经济社会效益

电解液作为新能源电池的核心材料，传统试错方法的研发模式存在研发周期长、成本高的弊端，且需依赖大量实验验证，难以满足新能源产业快速发展的需求。本案例构建面向 AI 的高质量电解液分子动力学数据库，并推动 AI 赋能新型电解液材料研发落地应用，将显著缩短研发周期与成本，是推动我国电池技术革新和产业升级的关键路径。

与此同时，当前核心数据的竞争已上升为世界各国安全博弈的核心领域，国际上常用的材料数据库约 90% 由欧美国家建设，一旦这些材料数据库的访问对中国实施封锁，会极大阻碍中国材料研发的发展。本案例基于国产的软硬件基础设施，构建自主、安全、可控的面向 AI 的高质量电解液分子动力学数据库，是有效抵御国外相关材料数据库“卡脖子”风险、确保新材料产业链创新链安全的必要手段。

### （5）示范推广价值及现状

本项目深刻践行国务院印发的《关于深入实施“人工智能+”行动的意见》要求，高度契合国家支持的“AI+科学技术”以及“人工智能+”产业发展，将助力打造国内 AI4S 在电解液材料领域落地应用的标杆样板，并成为 AI 在其他新材料，如橡胶材料、玻璃材料、水泥材料等研发领域推动新范式落地的典型示范。

### （6）开源情况

一方面，本案例涉及的电解液数据库一期上架国家新材料大数据中心，直接面向公众用户，服务全国范围内的电解液乃至电池材料开发科研工作者与企业研发人员。另一方面，案例中涉及基于“天穹”科学计算机构建的“算力—数据—模型”三位一体的科研创新“底座”平台，并将融入国家“1+N”架构体系的材料数据资源网络，成为重要的材料计算数据节点，通过对外提供算力服务接口，为广大的新材料科研工作者与企业研发人员提供高质量的材料仿真模拟算力服务；同时，平台也提供便捷完备的软件编译开发环境，能够让材料计算科学家围绕特定的计算任务需求灵活自主地设计相应的软件功能，满足各类计算场景需求。

### （7）相关标准成果产出

本案例服务于 3 家以上科研机构或企业，同时，围绕力场模型发表 1 篇文章和申请 1 篇专利。

## (十) 基于数理双驱模型组合的二维材料发现

### 1. 案例实施单位：

北京百度网讯科技有限公司、苏州实验室

### 2. 案例背景：

二维材料因其独特的量子限域效应与高比表面积，在柔性电子、光电、催化与储能等前沿领域具有重大应用前景。然而，新材料研发长期受限于传统“试错法”的高成本与长周期，且现有理论数据库规模有限、化学多样性不足，严重阻碍了“按需设计”的产业化进程。为突破此瓶颈，本项目旨在融合人工智能的数据驱动能力与第一性原理计算的物理机理，构建数理双驱的智能发现新范式，以实现二维材料宇宙的系统性探索与高效定向设计，为前沿科技产业提供颠覆性的材料解决方案。

### 3. 案例概述：

二维材料因量子限域、高比表面积等特性，在柔性电子、光电、催化与储能等领域具有关键应用价值，但现有数据库规模与成分多样性有限，难以支撑“按需设计”的产业化需求。为突破瓶颈，项目组提出以“数据驱动+物理机理”数理双驱为核心的闭环主动学习流程，融合深度神经网络与 DFT 高通量验证，系统扩展 2D 材料版图并面向目标性能开展条件生成。

### 4. 行业痛点：

当前，二维材料的探索严重依赖传统“试错法”与有限的计算模拟，面临三大核心瓶颈：一是现有数据库规模小、化学多样性不足，难以

支撑功能导向的材料定向设计；二是第一性原理计算虽精确但算力成本高昂，无法实现亿级候选空间的高通量筛选；三是材料生成与稳定性预测脱节，常产生大量不合理结构，导致研发周期长、成功率低。这些因素共同制约了新型二维材料的快速发现与产业应用。

## 5. 案例详述：

### (1) 总体思路

构建“数据收集—物理引导元素替换—MLIP（机器学习势）快速筛选—DFT 验证—条件扩散生成”的五段式闭环框架：以原型库与元素替换生成亿级候选相空间，轮次化抽样与 MLIP 预测凸包能  $E_{\text{hull}}^{\text{MLIP}}$  筛选，阈值通过 DFT 精确校核  $E_{\text{hull}}^{\text{MLIP}}$ ，合格结构入库并反哺模型迭代；在稳定库基础上训练等变分数网络，实现按化学组成与对称性等条件的结构生成。

### (2) 软件平台

平台分层包括：数据层（二维材料结构/能量/力数据与凸包参照）、模型层（CHGNetPaddle-MLIP、等变扩散生成器）、应用层（相图构建与定制生成）。流水线以 JAMIP 执行 DFT 校核，并以 MLIP 快速预测与筛选；在应用层可按给定体系构建相图并开展空间群约束生成，支撑从“筛—验—生”的一体化研发。

### (3) 技术方案

一是元素替换+主动学习，从已知 2D 原型出发，物理启发元素替换生成候选；每轮随机采样百万级结构，用 MLIP 筛选  $E_{\text{hull}}^{\text{MLIP}} < 50 \text{ meV/atom}$  者进入 DFT 验证；迭代更新后，MLIP 稳定

判别准确率由 19.01% 跃升至 92.36%。二是条件扩散生成，联合原子种类、坐标与晶格参数的联合扩散，并引入厚度与目标性质嵌入与适配器，实现按元素/空间群等条件的 S.U.N. (Stable/Unique/New) 结构生成；在 1,024 个生成结构的 DFT 复核中，74.8% 满足  $E_{\text{hullMLIP}} < 100 \text{ meV/atom}$ ，59.6% 满足严格的  $E_{\text{hullMLIP}} < 0$  标准。三是相图与快速验证，结合高精度 MLIP 对给定化学体系快速构建相图（如 V-Se-O），与 DFT 相图一致，MAE 仅 54 meV/atom。

#### (4) 关键指标参数

一是数据库规模与覆盖，构建 ML2DDB，收录 242,546 个经 DFT 验证的稳定单层结构，元素覆盖达 81 种，显著优于既有 2D 数据库。二是筛选精度，主动学习 5 轮后，MLIP 稳定性判别准确率达 92.36%，能量/力 MAE 降至 41 meV/atom 与 74 meV/Å。三是生成质量：生成结构与 DFT 弛豫后结构的  $\text{RMSD} < 0.26 \text{ \AA}$ ；千级生成保持 100% 唯一性，万级时唯一性仅下降 4.4%。

### 6. 创新情况：

#### (1) 技术创新

一方面提出“数理双驱”的 2D 材料发现范式，以 MLIP 加速与 DFT 校核构成内环精化；以条件扩散在稳定域上进行外延探索，贯通“发现—验证—再发现”的闭环。另一方面在生成侧实现 A/X/L 三变量联合扩散与性质约束适配，显式处理二维晶格周期与对称，提升 S.U.N. 材料产出与可控性。

#### (2) 模式创新

以闭环主动学习替代“一次性高通量”，在相同算力下实现更高的有效样本率与更快的“从候选到新材”转化。

### 7.应用实效：

一是库级扩容，相对现有 2D 库实现数量级提升，并显著增加三元/四元复合体系，为产业端的复合功能材料设计提供更大搜索空间。二是从筛到生一体化，面向具体化学体系（如 V-Se-O）快速构建相图并定向生成，缩短从“构想→候选→验证”的周期。三是面向多场景可迁移：可直接支撑柔性电子、光电与催化等方向的组合优化与参数倒推，提升研发效能与成功率。

## (十一) 多模态技术赋能有机分子谱图解析模型研发

### 1. 案例实施单位：

北京百度网讯科技有限公司、苏州实验室

### 2. 案例背景：

有机分子结构鉴定是药物发现、代谢物分析及环境检测等领域的关键环节，其核心任务是从核磁共振（NMR）等谱图数据中精准反推出分子结构。传统方法完全依赖化学专家的人工经验，通过对比谱图表征与已知结构数据库进行手动解析。这一过程不仅耗时冗长、成本高昂，且严重受制于分析人员的主观经验与知识储备，已成为制约新药研发与化学品创新的重要瓶颈。为突破此限制，本项目旨在利用多模态人工智能技术，开发能够自动、精准解析有机分子谱图的人工智能系统，推动化学分析从“经验驱动”迈向“数据驱动”的智能化新范式。

### 3. 案例概述：

有机分子结构解析长期依赖专家对 NMR 谱的手工判读，流程慢、依赖经验且可扩展性弱。该案例提出以“光谱—分子”多模态对齐为核心、以条件离散扩散为生成骨干的 DiffNMR 框架，旨在实现从  $1\text{H}/^{13}\text{C}$  NMR 到分子图的自动化反演。

### 4. 行业痛点：

当前，有机化合物的谱图解析面临三大核心挑战：一是高度依赖专家经验，解析结果主观性强，且资深专家资源稀缺，难以满足高通量研发需求；二是传统计算方法局限性大，基于规则的系统难以处理复杂或未知结构，而早期的机器学习方法存在准确率低、泛化能力弱

的问题；三是效率与精度难以兼得，人工解析耗时数小时甚至数天，且随着分子复杂性增加，正确率显著下降。这些痛点严重制约了科研与产业创新的速度，行业亟需一种能够兼顾自动化、高精度与强泛化能力的智能解析新方法。

## 5. 案例详述：

### (1) 总体思路

通过“两阶段预训练+端到端微调”实现谱-分子表征对齐与条件扩散生成：第一阶段以扩散自编码器建立分子图与潜表征的双向映射；第二阶段以对比学习将 NMR 编码对齐到分子潜空间；最终以 NMR 编码作为条件引导扩散解码器生成分子图。

### (2) 软件平台

一方面提供 NMR 编码器，采用 RBF 连续嵌入+双向跨模态注意力融合  $^1\text{H}$  与  $^{13}\text{C}$  信息，保留化学位移的连续性并刻画氢-碳对应等关键关系。另一方面提供图扩散解码器，以图 Transformer 为去噪网络，结合 FiLM 与 PNA 算子，迭代生成满足谱图约束的分子结构。

### (3) 技术方案

一方面支持对比表征对齐，通过谱-分子对的对比学习，获得具有判别性的公共潜空间，为相似度检索初始化与相似度过滤提供基础。另一方面支持推理增强：在无外部分子式先验时，引入相似度过滤将候选按谱相似性重排；当可用数据库时，以检索分子作为扩散初始状态，减少步数并提高准确率。

### (4) 关键指标参数

一是多模态协同增益，在提供分子式的设置下， $1H+13C$  联合输入 Top-1 达 68.26% ( $\leq 15$  个重原子)，优于单一模态；该趋势在  $\leq 20$  与  $\leq 25$  组别保持一致。二是无分子式场景的鲁棒性，仅用谱信息时(无分子式)，相似度过滤将  $\leq 25$  组别 Top-1 由 35.76%  $\rightarrow$  53.10%，引入检索初始化后进一步至 58.47%；平均 Tanimoto 相似度同步提升。三是效率提升，检索初始化可将去噪步数由 500 步降至 200 步并给出正确结构示例。

## 6. 创新情况：

### (1) 技术创新

提出条件离散扩散+多模态对齐的谱图反演架构，避免自回归式顺序误差累积，整体保证结构一致性；RBF 连续嵌入显著增强对近似位移差异的分辨能力。设计“检索初始化+相似度过滤”的推理增强策略，在大分子复杂场景下有效缩小准确率劣化。

### (2) 模式创新

将谱图解析从“专家规则+手工匹配”升级为“数据驱动+知识增强”的自动化流水线，既可离线训练，也可在线接入实验室数据库与生产 QC 系统。

## 7. 应用实效：

### (1) 降本增效

联合模态与推理增强在无分子式场景显著提升 Top-1 与 Tanimoto 指标，显著减少人工试错与谱图解读时间。

### (2) 场景落地

可用于药物先导筛选、代谢组学鉴定、精细化工质检等高通量场景；在有数据库的企业内网条件下进一步缩短推理时间与提升命中率。

### (3) 生态融合

基于 PaddlePaddle 与 PaddleMaterials 实现，便于与国产算力与材料研发工具链集成升级。

### 三、工业仿真

#### (一) 高速动车组空气动力学智能化仿真大模型-斫轮·风驰

##### 1. 案例实施单位：

中车青岛四方机车车辆股份有限公司、国家高速列车青岛技术创新中心、北京百度网讯科技有限公司

##### 2. 案例背景：

轨道交通装备已成为世界各国推动经济发展的战略性竞争新高地，速度是轨道交通装备研发的永恒追求。速度提升引发的空气动力学效应直接影响列车运行安全和节能环保，是整车设计的关键问题。气动性能的精准评估和优化是列车研发的核心技术，仿真是评估列车气动性能的核心手段。列车研发过程中，需开展多方案、多轮次、多车速、多场景的仿真评估优化，重点型号气动外形设计往往耗时数月甚至数年时间，严重制约研发效率。

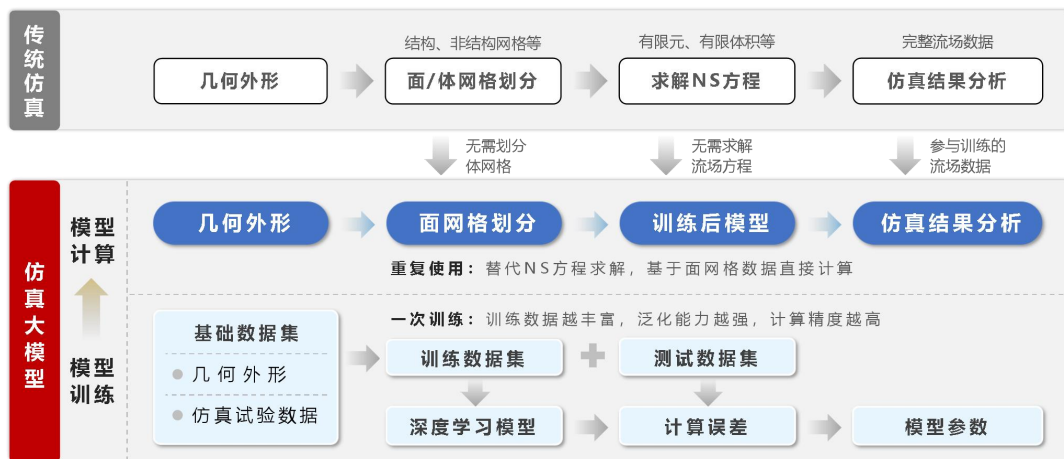
##### 3. 案例详述：

传统空气动力学仿真存在“精度-效率-成本”难以协调的突出问题，高精度的仿真结果依赖精细化的仿真建模策略和高精度的物理模型算法，导致计算量大幅增加。为了缓解仿真精度与效率之间的矛盾，需要采用大规模超算资源，仿真成本显著提升，且仿真效率仍然难以完全满足研发需求。同时，跨尺度的列车复杂几何模型前处理困难，耗时长；仿真软件的使用门槛高，需要具备专业的知识；海量历史数据的使用率低，未发挥潜在价值。依托中车斫轮大模型框架，基于科学计算大模型，以我国高速动车组二十余年自主研发过程中积累的

TB 级数据为基础，融合几何信息神经算子和傅里叶神经算子，开展深度神经网络模型训练和优化，开发高速动车组空气动力学智能化仿真大模型-砺轮·风驰，替代传统空气动力学仿真流程，使用单机显卡实现高速动车组气动性能的高精度快速仿真，解决传统空气动力学大计算量、长周期仿真“精度-效率-成本”难以协调的基本矛盾问题。

### (1) 总体思路

与传统 CFD 计算的技术路线不同，仿真大模型采用数据驱动的方法，以我国高速动车组二十余年自主研发过程中积累的 TB 级数据为基础，基于科学计算大模型，开展深度神经网络模型训练和优化，开发高速动车组空气动力学智能化仿真大模型-砺轮·风驰，替代传统的流场仿真求解，简化仿真建模过程，省略网格划分及流场求解时间，仿真效率大幅提升。



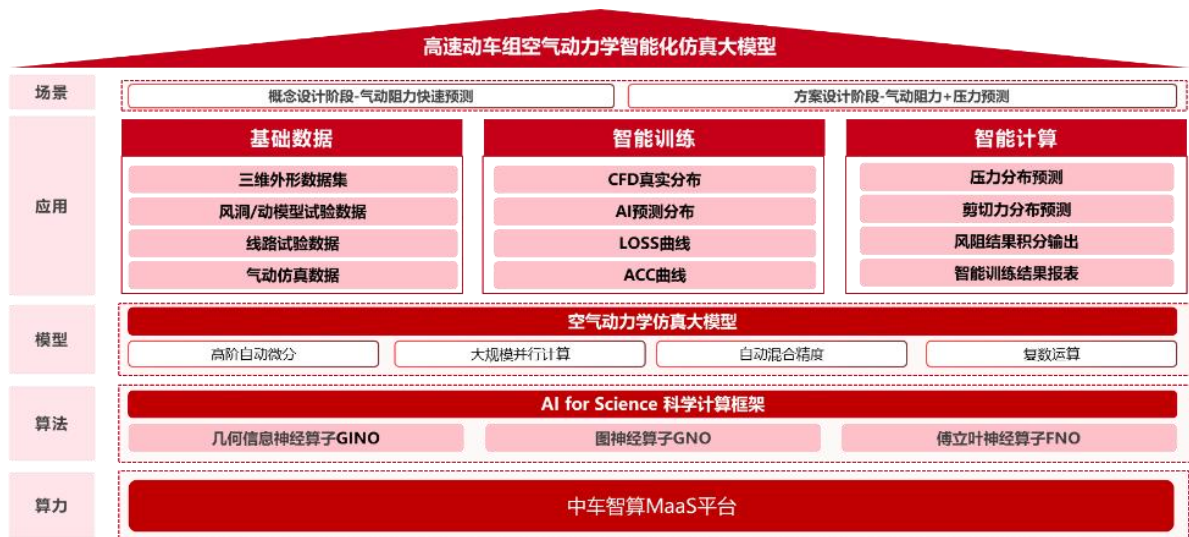
来源：中车青岛四方机车车辆股份有限公司

图 26 砺轮·风驰总体思路

### (2) 平台介绍

砺轮·风驰仿真大模型构架主要包括硬件层、算法层、模型层、应用层。硬件层为中车砺轮大模型平台，涵盖算力、MaaS 平台、基

模、应用、服务及生态的全链条内容，应用场景覆盖设计、制造、运营、维护、安全等各环节。**算法层**包括基于百度飞桨框架的几何信息神经算子、图神经算子、傅里叶神经算子等算子和模型预测精度修正算法、高速动车组复杂外形网格自动加密等模块。**模型层**利用海量高速动车组流场数据对深度学习神经网络进行训练和优化，形成空气动力学智能化仿真大模型。**应用层**利用仿真大模型进行高速动车组空气动力学性能的高精度快速预测。

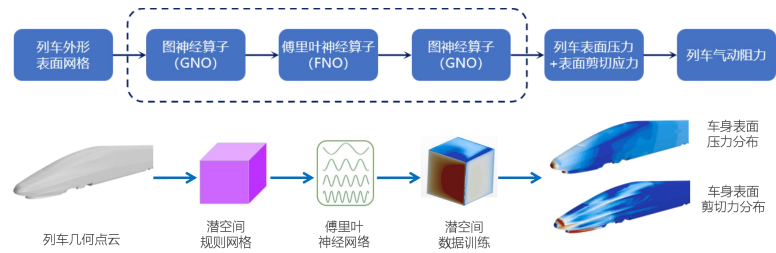


来源：中车青岛四方机车车辆股份有限公司

图 27 研轮·风驰大模型架构

### (3) 技术方案

研轮·风驰仿真大模型，融合了几何信息神经算子和傅里叶神经算子，建立了空气动力学智能化仿真框架。利用几何信息神经算子将不规则的复杂几何外形高保真映射到规则的潜空间，利用傅里叶神经算子将流场特征变换至频域空间，通过全局线性积分算子和局部非线性激活函数，进行高维强非线性神经算子学习，再利用傅里叶反变换和几何反映射，实现列车表面压力、切应力的高精度快速预测。



来源：中车青岛四方机车车辆股份有限公司

图 28 研轮·风驰大模型技术方案

研轮·风驰大模型优化了动静统一的自动并行编程范式，发展了基于字节码的动静转换技术，减少 80% 分布式代码开发；创新了神经网络张量的复数表示与运算技术，支持不同类型的算子学习模型。开发了异构多芯适配技术，全面支持昆仑 XPU、海光 DCU 等国产芯片及麒麟操作系统等国产系统，通用性高，可扩展性好，实现了关键技术全栈自主可控。

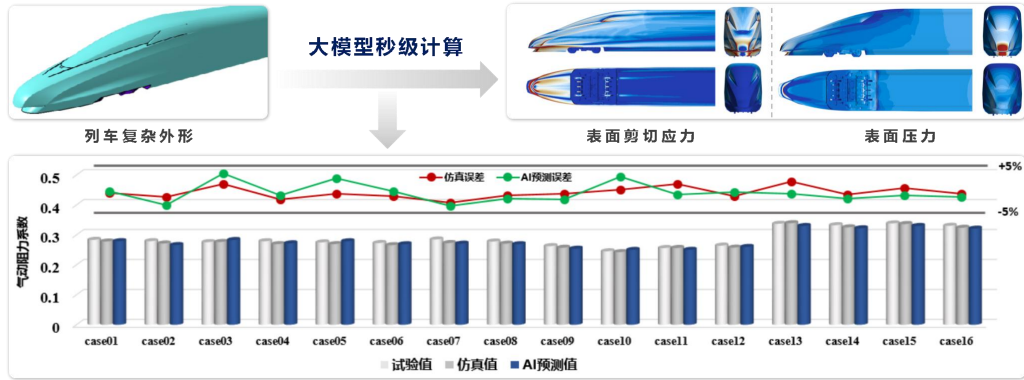


来源：中车青岛四方机车车辆股份有限公司

图 29 动静统一自动并行技术

#### (4) 关键指标参数

研轮·风驰大模型相当于研发人员的“超级计算器”，直接将列车几何文件导入仿真大模型，在 10 秒内获取列车表面压力、切应力及气动阻力等关键数据。与传统仿真相比，单工况的列车气动性能仿真周期由超算平台上的天级缩短至单机显卡上的秒级；与风洞试验结果相比，整车气动阻力的预测误差小于 5%。



来源：中车青岛四方机车车辆股份有限公司

图 30 研轮·风驰大模型仿真应用

#### 4. 创新情况：

##### (1) 技术创新情况

创新三维复杂几何外形气动性能的精准预测技术，攻克空气动力学智能化仿真大模型的几何泛化难题。研轮·风驰大模型创新了融合几何信息神经算子和傅里叶神经算子的空气动力学智能化仿真框架。在几何拓扑空间，利用几何信息神经算子解析列车三维几何特征，并将其高保真映射到规则化潜空间；在物理流场空间，利用傅里叶神经算子将列车流场特征变换至频域空间，通过全局线性积分算子和局部非线性激活函数，进行高维强非线性神经算子学习，实现复杂三维曲面表面压力、切应力的预测，在保证计算精度和计算效率的前提下，大幅提升大模型的几何泛化能力。

开发动静统一自动并行、复数运算机制和异构多芯适配技术，通用性高，可扩展性好，实现关键技术的全栈自主可控。研轮·风驰大模型优化了动静统一的自动并行编程范式，发展了基于字节码的动静转换技术，开发者无需关注分布式通信和调度逻辑，减少 80% 的分布式代码开发。创新了神经网络张量的复数表示与运算技术，支持不同

类型的算子学习模型，实现了傅里叶神经算子的快速训练与推理。构建了多硬件统一适配方案，降低了异构芯片与框架适配的复杂度，全面支持昆仑 XPU、海光 DCU 等国产芯片，及麒麟操作系统等国产系统，通用性高，可扩展性好，实现关键技术的全栈自主可控。

**创新多源异构流场数据标准化处理及融合技术，构建满足空气动力学智能化仿真大模型训练的数据集。**以我国高速动车组二十余年自主研发过程中积累的 TB 级数据为基础，首次系统整合了数值仿真、模型试验、线路试验等多源异构流场数据及列车三维外形数据，提出了流场数据和几何数据的标准化处理框架。发展了多源异构数据融合技术，结合张量分解技术实现非结构化流场数据与结构化试验数据的维度对齐，形成了具有严格一致性的列车气动载荷标准数据库，构建了满足仿真大模型训练要求的数据集。

## (2) 模式创新情况

该案例开发的高易用低功耗的仿真大模型软件已应用于列车气动外形设计及优化，并可拓展至风电、航空、汽车等装备制造领域。软件架构方面，采用三层架构，底层为数据层，包括几何数据、仿真及试验数据；核心层为模型层，包括核心算法调优、大模型训练；顶层为应用层，进行列车气动性能的高精度快速预测。仿真前处理方面，只需导入列车三维几何文件，无需进行缝隙缝合和破面修补，人工建模时间由 6 小时缩短至 1 小时。气动载荷计算方面，利用仿真大模型替代传统的流场仿真，计算时间由超算平台上的天级缩短至单机显卡上的秒级。

## 5.应用实效:

**天级飞跃至秒级的仿真效率提升。** 斫轮·风驰大模型首次实现人工智能技术与轨道交通装备仿真领域的深度融合。以列车真实几何外形为输入，利用空气动力学智能化仿真大模型，将单工况的列车气动性能仿真周期由超算平台上的天级缩短至单机显卡上的秒级，实现了高速动车组空气动力学仿真效能的飞跃，为列车流线型头型、车体断面、转向架舱等气动外形设计研发提供秒级响应的决策支持，加速新一代高速动车组产品的迭代创新。

**单点突破到全域升级的效能提升。** 斫轮·风驰大模型具有强大的复杂几何泛化能力，通过单次模型训练即可适用多种车型，形成“一次训练、多车型适用”的创新模式。随着训练样本的持续扩充，大模型还将向“多场景适用”方向发展，形成“全车型、全场景”气动性能评估能力。结合智能优化算法，实现整车气动外形的智能优选和快速优化，为高端装备制造领域提供“AI+工业制造”融合的示范样本。

**全流程融入产品研发运维与跨行业革新。** 斫轮·风驰大模型深度融合产品研发与运维全流程。在研发端，大模型相当于研发人员的“超级计算器”，支撑列车外形快速迭代设计；在运维端，大模型未来可接入线路气象数据，进行强风下列车气动性能实时预测，支撑列车安全运营。该大模型技术可拓展推广至高速磁浮列车、城际动车组、高速地铁列车等轨道交通装备，可在风电、航空、汽车等装备制造领域实现跨行业迁移应用。

**助力高速动车组降本增效与节能优化。** 斫轮·风驰大模型可实现

全链条的降本增效。在研发过程中，利用大模型可显著降低列车气动性能评估验证成本，大幅缩短气动外形研发周期，节省大量的人力物力。利用大模型开展列车气动外形减阻优化设计，降低列车运行能耗。以时速 350 公里高速动车组推算，8 车编组高速动车组的阻力每降低 1%，每百公里可节省 14 度电，京沪高铁单程运行可节省 168 度电，按每日 1 个往返，每年运行 330 天计算，一列 8 辆编组的高速动车组每年可节省 11 万度电，经济效益显著。

**赋能自主仿真软件突破与产业生态重构。** 斫轮·风驰大模型具备完全自主知识产权，其技术框架及核心算法可衍生为自主可控的商业仿真软件，后续将开发轨道交通装备智能化仿真软件，并通过与国产工业软件平台的深度适配，推动轨道交通装备研发从“经验驱动”向“数据驱动”转型升级。以 AI+仿真为代表的“技术突破-软件替代-生态重构”递进式发展路径，为我国高端装备制造业智能化升级提供可复制范式。

## (二) 民用客机超临界翼型函数生成和编辑模型

### 1. 案例实施单位：

上海人工智能创新中心、中国商飞上海飞机设计研究院

### 2. 案例背景：

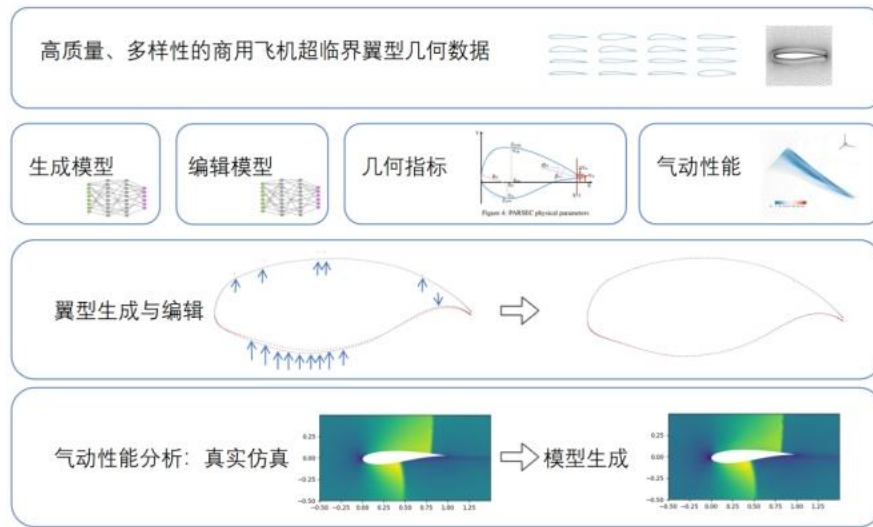
翼型设计在飞机设计制造中至关重要，但传统设计方法耗时且效率不高。工程师常采用迭代试错流程，从有限数据集中手动筛选和优化翼型，往往需数月才能定稿。尽管机器学习被视为潜在替代方案，现有翼型表征方法却存在根本性权衡。参数化模型（如贝塞尔曲线、NURBS）保证翼型光滑且支持任意分辨率采样，但将设计空间限制于预设函数族，制约了设计灵活性。离散点模型虽具备更大灵活性，但缺乏光滑性，无法适应制造对分辨率的需求。这些限制严重影响了翼型设计的效率、成本效益和创新潜力。因此，亟需一种融合参数化模型连续光滑性、离散点模型分辨率灵活性，并能最大化设计空间的统一方法，以推进翼型设计发展。

### 3. 案例详述：

FuncGenFoil 将翼型建模为连续函数，通过神经算子实现光滑几何结构与分辨率无关的控制点，利用流匹配技术合成超越参数化约束的新型翼型，拓展设计空间。用户可通过拖拽固定点交互式优化生成翼型，同时保持几何完整性。在 AFBench 数据集上，该方法相比前沿方法降低了 74.4% 的设计误差，设计多样性提升 23.2%。此外，FuncGenFoil 的翼型编辑功能十分自由灵活，与约束条件实现了接近零误差的效果。

## (1) 总体思路

FuncGenFoil 通过神经算子将翼型表征为连续函数，弥合参数化与离散点方法鸿沟，在保持光滑性和设计灵活性的同时，实现分辨率无关生成。关键创新包括：通过神经算子（FNO）和流匹配技术，利用常微分方程（ODE）将隐函数映射至翼型几何；以任意分辨率生成高质量翼型；通过函数空间优化精确修改翼型，保持几何完整性。



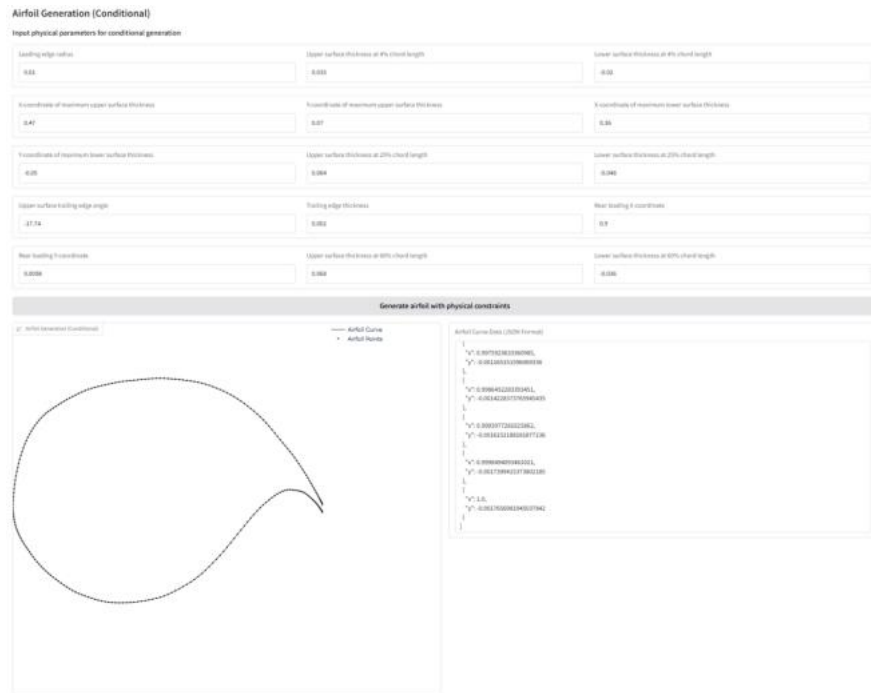
来源：上海人工智能创新中心、中国商飞上海飞机设计研究院

图 31 FuncGenFoil 算法框架组建与功能示例

## (2) 模型技术方案

FuncGenFoil 以神经算子为骨干神经网络，采用连续时间生成模型，包含两大核心模块。生成模型融合神经算子与流匹配，通过线性 ODE 路径将翼型扰动为隐码，并利用学习的速度算子反向求解 ODE 重构翼型。翼型参数化通过角度参数编码翼型的环形拓扑，并变换坐标系。训练过程使用流匹配算法进行模型训练，最小化流匹配损失以对齐生成与目标翼型，推理时从高斯过程隐码出发求解 ODE 生成翼型。编辑模块使用最大后验（MAP）估计优化隐码，确保编辑区域

满足用户约束同时保持生成先验。编辑流程包括将原始翼型反演至隐空间，通过微调模型最小化编辑损失。



来源：上海人工智能创新中心、中国商飞上海飞机设计研究院

图 32 FuncGenFoil 算法框架操作界面示例

### (3) 关键指标参数

在 AF-200K 数据集上的生成性能表现为：标签误差降低 74.4%，设计多样性提升 23.2%，光滑度达到  $1.16 \times 10^{-2}$ ，优于参数化模型。编辑精度方面，点调整误差小于  $2.75 \times 10^{-7}$ ，编辑耗时为 10~50 步微调即可完成。分辨率灵活性表现为以 4 倍训练分辨率生成翼型时差异极小（误差增幅小于  $9.1 \times 10^{-3}$ ）。气动验证显示，生成的机翼与原始数据集升阻比（L/D）分布一致。

## 4. 创新情况：

### (1) 技术创新情况

FuncGenFoil 融合前沿 AI 技术与工程需求，在 AI 驱动气动设计

**领域实现重大突破。**关键技术进展包括：**一是**通过神经算子和流匹配将翼型建模为连续函数，规避离散方法分辨率限制与参数化模型局限；**二是**在函数空间求解 ODE 生成翼型，实现光滑、无限分辨率的模型输出，满足行业对高保真翼型生成的需求，为制造环节提供直接可用输出；**三是**引入贝叶斯优化框架精准编辑翼型，通过微调隐码实现近零误差，赋能工程师交互式优化 AI 生成设计，同时保持气动完整性。以流匹配替代传统扩散模型，降低训练复杂度并提升收敛速度，显著降低生成误差并提升数据多样性。

## (2) 模式创新情况

**FuncGenFoil 开创航空航天工程领域 AI 驱动新范式，重塑传统设计流程。**通过生成任意分辨率翼型，无缝对接 CAD/CAM 系统，支持高分辨率机翼制造，消除人工后处理误差。工程师通过拖拽界面与 AI 协同优化设计，融合领域知识与生成先验，CFD 验证轮次减少 50%，降低算力成本。该框架可扩展至 3D 涡轮/压气机叶片、汽车气动外形等参数化形状设计，并通过代码和预训练模型发布，推动跨行业应用。隐空间可解释性确保符合航空安全标准，生成翼型经物理定律验证（如光滑度达标），避免缺陷设计。FuncGenFoil 作为首款函数空间模型，开启连续、分辨率自由的翼型生成范式，为气动设计树立新标杆，能将开发周期由数月缩短至数日，支持高性能翼型快速探索，是气动外形优化领域的关键进展，契合工业界对效率、创新与适应性的核心诉求。

## 5.应用实效：

FuncGenFoil 模型自部署以来，已在实际工程场景中完成落地，覆盖了从二维翼型初步设计到工程优化的多个环节，并展现出显著的经济与社会效益。

### (1) 典型应用场景及用户

FuncGenFoil 已在中国商飞上海飞机设计研究院二维翼型气动总体设计与优化环节实现应用。模型将服务于新一代民用飞机预研，支撑从方案探索、性能对比到快速优化的自动化设计。

### (2) 行业痛点解决成效

FuncGenFoil **显著缩短翼型设计迭代周期**。原需数月的设计流程现可在数天内完成大批量高质量备选方案，极大提速工程研发。据统计，仅在上飞院二维翼型工作中，2 周内高效生成 500 余组候选方案。通过函数空间建模带来的无限分辨率的模型输出，工程师无需反复修改模型以兼容不同制造分辨率，极大降低了人工后处理与 CAD/CAM 对接的复杂度。自动化高保真生成能力减少了传统 CFD 验证的重复轮次，部分场景下验证工作量减少超过 50%，显著降低算力与人力成本。设计空间的拓展带来多样性提升，为新型翼型创新提供更多可能。

### （三）风扇叶轮智能优化设计

#### 1. 案例实施单位：

南京天泱软件有限公司

#### 2. 案例背景：

广东某风扇研制企业面临产品性能升级需求，希望在相同转速下将风扇风量提升 5%。传统设计依赖工程师经验，需经历参数调整、打样测试、反复优化的冗长流程，研发周期长达 4 个月，成本高昂。企业缺乏空气动力学领域的顶级专家，难以通过常规手段突破技术瓶颈。天泱软件基于 AI 与工业软件融合的技术优势，为其提供智能化解决方案。

#### 3. 案例详述：

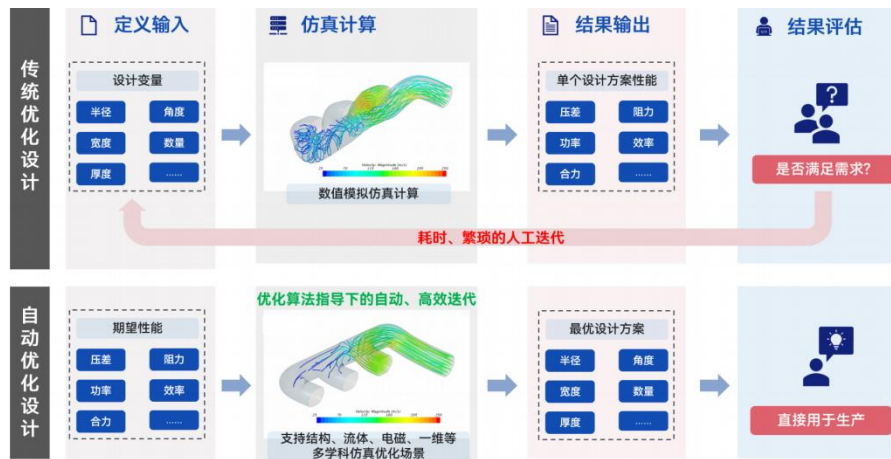
叶片机械设计高度依赖工程师经验，参数调整缺乏系统性，试错成本高。传统流程需多次打样测试，迭代效率低，难以适应快速变化的市场需求。同时叶片性能优化涉及流体力学、材料科学等多学科知识，传统方法难以高效融合。现有商业软件操作复杂，需专业背景，非专家难以直接应用。天泱软件通过智能优化设计软件 AIPOD，对风扇叶片结构进行参数化建模（涉及 7 个关键变量，如弦长、扭转角、天花板距离等），结合仿真模拟与智能算法，在 2 天内自动生成 61 种优化方案。最终方案将风量提升 20%，远超客户预期。AIPOD 还分析出“天花板距离”是影响风量的核心因素，验证了 AI 技术在复杂工程问题中的高效决策能力。该案例突破了传统设计依赖经验的局限，为中小企业提供了技术平权工具。

## (1) 总体思路

天泱软件以“AI 驱动工业设计革命”为核心理念，通过将人工智能技术深度融入传统工业软件，构建“参数化建模-智能仿真-多目标优化-知识沉淀”的全流程解决方案。针对风扇叶片设计的复杂性，提出解决方案。一是参数化建模，将叶片几何结构转化为可量化的设计变量，打破经验依赖。二是 AI+仿真融合，通过智能算法自动生成优化方案，结合高精度仿真验证，加速迭代。三是知识图谱赋能，整合流体力学、材料科学等跨学科知识，形成可计算的数字资产。四是实现技术平权，通过“傻瓜式”操作界面，降低中小企业技术门槛，实现行业普惠。

## (2) 软件平台

AIPOD 是天泱软件自主研发的工业级智能优化平台，核心功能包括多学科优化算法、参数化建模工具、智能仿真驱动、知识图谱系统和可视化分析。多学科优化算法集成遗传算法、粒子群算法、神经网络等 AI 技术，支持多目标、多约束优化。参数化建模工具内置几何参数化模块，支持用户自定义设计变量（如弦长、扭转角、安装角度等）。智能仿真驱动可自动调用 CFD（计算流体力学）、FEM（有限元分析）等仿真工具，实现“方案生成-仿真-优化”闭环。知识图谱系统将领域专家经验编码为可计算的规则，指导优化方向并沉淀设计知识。可视化与分析实现输出优化结果对比、参数敏感性分析及 3D 模型，辅助用户决策。

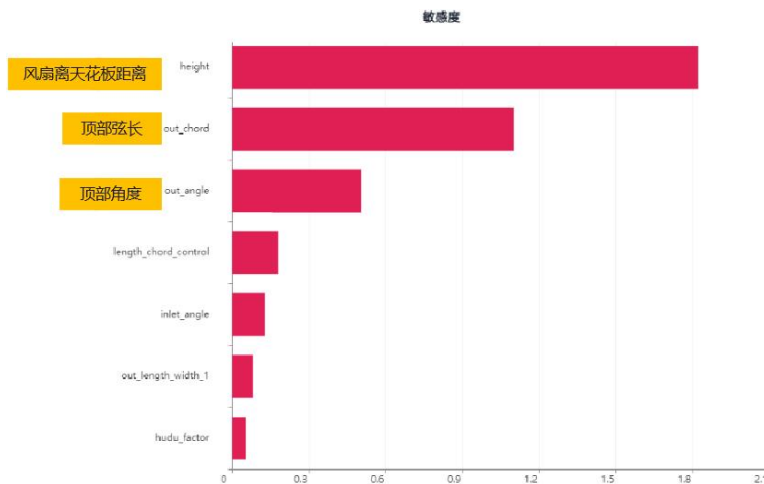


来源：南京天湫软件有限公司

图 33 AIPOD 智能优化平台

### (3) 技术方案

技术实施方案包括参数化建模与变量定义、智能优化流程。关键设计变量共 7 个，包括设计变量有天花板距离、叶根弦长、叶顶弦长、叶根角度、叶顶角度等。约束条件包括叶片强度、加工工艺可行性、成本限制等。

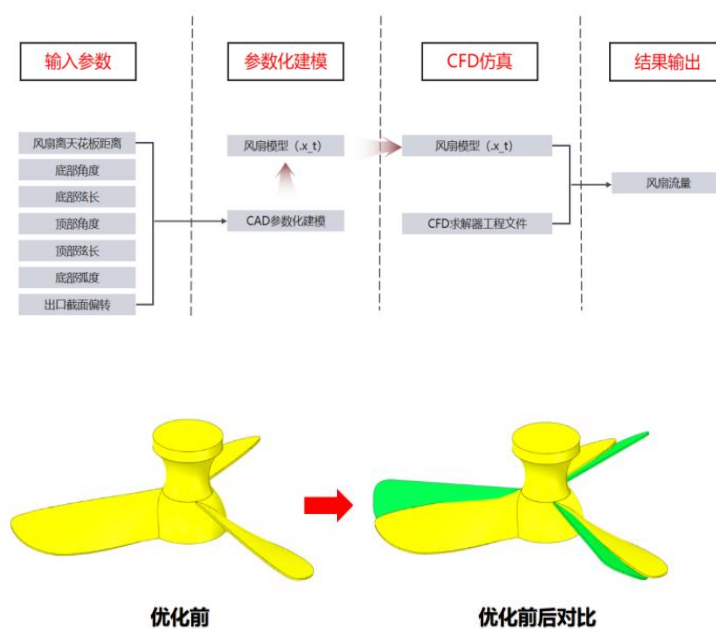


来源：南京天湫软件有限公司

图 34 风扇优化中的关键设计变量

AIPOD 智能优化流程通过算法驱动与自动化工具实现高效设计优化，减少 70%人工操作，3-5 倍效率提升，支持复杂工程场景下的

**多学科协同优化。**一是参数化建模与变量定义，提取 CAD 模型参数或通过非参数化变形技术定义设计变量，集成 CAE 工具链(如 ANSYS、Fluent) 搭建自动化仿真流程。**二是**智能采样与代理模型构建，利用 AI Agent 进行智能采样，生成初始训练数据集，构建高精度代理模型(如 RBF、Polynomial)，通过主动学习追加样本提升模型可靠性。**三是**多目标优化求解，采用改进 Pareto 算法(如 NSGA3、SilverBullet) 进行多目标寻优，结合分布式计算加速收敛，生成最优解集(Pareto 前沿)。**四是**稳健性设计与验证，区分控制因子与噪声因子，评估设计稳健性；通过敏捷后处理工具筛选方案，自动回传至 CAD 系统验证工程可行性。**五是**知识融合与迭代，集成专家规则库约束搜索方向，挖掘设计规律复用至后续优化，形成闭环迭代优化体系。



来源：南京天湫软件有限公司

图 35 风扇设计优化前后对比

#### (4) 关键指标参数

关键指标参数如下表所示。

表 1 AIPOD 风扇优化设计前后指标对比

指标	优化前	优化后	提升幅度
风量 (m <sup>3</sup> /h)	5,000	6,000	+20%
研发周期 (天)	120	2	-98%
仿真计算成本 (万元)	15	3	-80%
参数敏感性分析精度	依赖经验	±5%误差	—
跨学科知识复用率	0%	70%	—

来源：南京天湫软件有限公司

#### 4. 创新情况：

##### (1) 技术创新情况

**跨学科知识图谱与智能决策。**数据整合与模型构建方面，天湫软件通过整合空气动力学、材料科学等多学科知识，构建工业设计领域知识图谱，将叶片几何参数（如弦长、扭转角）与流体力学性能（风量、效率）的关联规则编码为可计算的数字资产。**算法突破方面**，采用遗传算法、粒子群算法等，还有分布式资源功能的升级，在 2 天内完成 61 种方案的自动优化，风量提升 20%，参数敏感性分析误差控制在±5%以内。

**智能仿真与算力优化。**仿真加速技术方面，通过 GPU 并行计算和模型降阶，将单方案 CFD 仿真时间从 8 小时缩短至 2 小时，支撑全流程自动化迭代。**闭环自动化方面**，实现“参数化建模→AI 优化→仿真验证”全流程无人干预，工程师仅需定义目标与边界条件，操作效率提升 98%。

**工业 AI 底座与价值对齐。**平台化能力方面，AIPOD 内置工业

AI 底座，支持跨行业场景快速迁移。风扇案例中的参数化建模与优化逻辑，已复用至船舶减阻（阻力降低 3%）、新能源汽车热管理等场景。价值验证方面，广东风扇企业通过 AIPOD 实现风量提升 20%，研发成本降低 80%，技术投资回报率达 5 倍。

**具身智能与软件工程。**具身智能应用方面，系统自动解析用户需求（如“风量提升”目标），生成几何模型并调用仿真工具，完成物理闭环。软件工程架构方面，采用微服务架构，支持模块灵活扩展。例如，知识图谱模块与优化算法模块可独立升级，保障系统迭代效率。

## （2）模式创新情况

**技术平权解决方案。**“傻瓜式”操作界面，通过引导式问答（如“风扇安装位置是否受限？”）替代复杂仿真参数设置，非专业工程师也能完成高精度优化。

**产学研协同创新生态。**人才培养模式，与清华大学共建博士生实践基地，每年输送 20 名博士生参与工业软件研发，实现“高校理论—企业实践—技术转化”闭环。与江苏省人工智能学会共建科技创新平台，聚焦智能优化算法与工业机理模型研发。跨行业场景拓展，新能源领域突破，基于风扇优化技术积累，天泱软件将 AI+仿真能力延伸至新能源汽车电耗分析。通过分析 2000 条车主数据，精准定位气温、电池老化等影响续航的关键因素，为车企优化设计提供依据。

## 5.应用实效：

### （1）应用场景与目标用户

**场景方面**，可应用于工业风扇、通风设备、散热系统等旋转机械

的叶片设计与性能优化。**用户方面**，已覆盖全国 300+制造企业，包括广东某风扇研制企业、华北储能企业、新能源车企等，其中中小型企业占比 70%，重点解决其缺乏高端人才、研发效率低的痛点。

## (2) 解决行业痛点

**经验依赖突破方面**，通过 AIPOD 智能优化，将传统 4 个月的研究周期缩短至 2 天，参数敏感性分析误差控制在 $\pm 5\%$ 以内。**技术平权普惠方面**，为中小企业提供与头部企业对等的 AI+仿真工具，如广东风扇企业通过 AIPOD 实现风量提升 20%，研发成本降低 80%。**跨学科整合方面**，整合空气动力学、材料科学知识，构建“参数-性能”映射关系库，支持多物理场耦合优化。

## (3) 经济社会效益

**直接经济效益方面**，广东某风扇研制企业缩短研发周期、无需引进顶级专家研发成本减少近百万，风量提升 20%，销售额提升百万元。某散货船企业应用 AIPOD 后，航行阻力降低 3%，年燃油费用节约近千万元。累计优化产品设计超千项，节约研发成本超亿元。**社会效益方面**，推动制造业绿色转型，如新能源汽车电耗分析项目，助力电车续航提升 5%。

## (4) 示范推广价值

软件已成功应用于船舶结构减重、电机电磁性能优化、旋转机械性能提升等领域，验证了其在复杂工业场景中的通用性。

#### (四) 代理模型在轮胎设计仿真中的应用及仿真自动化

##### 1. 案例实施单位：

中化信息技术有限公司

##### 2. 案例背景：

随着业务水平的提升，传统轮胎设计仿真面临几个困境，参数化仿真重复工作量大，人工操作易出错；流程自动化程度低，效率受限；传统试验设计方法难以实现多参数多目标耦合优化。针对这些问题，本项目采用代理模型技术，通过机器学习建立设计参数与性能指标的映射关系，替代耗时的高精度仿真；同时构建自动化仿真流程，集成参数化建模，结果分析，形成闭环优化系统。该方案有望提升轮胎性能优化的效率，提供了智能化的研发新路径。

##### 3. 案例详述：

案例聚焦轮胎胎冠设计的关键参数对印痕图特征的影响规律。通过构建参数化模型与代理模型，实现了设计参数到印痕图的高精度预测：**一方面**基于参数化设计变量建立数值回归模型，预测误差控制在5%以内；**另一方面**结合原始点云坐标数据，成功复现了与仿真结果高度吻合的印痕图形状。该研究验证了代理模型在轮胎接地特性预测中的有效性，为快速评估设计参数对轮胎接触性能的影响提供了可靠的技术手段，显著提升了设计迭代效率。

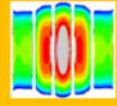
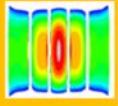
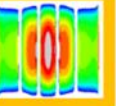
##### (1) 总体思路

传统轮胎设计流程中，工程师需要反复调整设计关键参数，通过耗时冗长的有限元分析来获取印痕图特征数据。单次仿真通常需要

4-6 小时的计算时间，而完整的设计迭代往往需要进行数十次这样的仿真。更关键的是，设计参数（如胎冠中心弧（R1）、胎冠总弧高（H）和行驶面宽（W））与印痕性能指标（包括左右胎肩长轴 L1/L2、中心长轴 L0 和 F 值）之间存在着复杂的非线性关系，使得传统的正交试验设计方法越来越难以满足现代轮胎多目标优化的需求。

针对这些行业痛点，案例创新性地构建了一套完整的解决方案。技术路线的核心在于将参数化建模、自动化仿真和代理模型预测三个关键环节有机整合，形成一个高效的设计优化闭环。在自动化仿真方面，本项目先针对仿真人员业务负担最大的网格前处理工作展开了试点，通过对 HyperMesh 的脚本开发，大幅提升了网格划分的整体效率，将原需要 1 小时的前处理流程压缩到 1 分钟，将单个方案网格划分总时间从 2.5 小时压缩到 1.5 小时，将工作效率提升了 60%。同时针对已进行标准化的仿真流程进行深度的二次开发，预期将实现，方案上传，自动化仿真，报告生成的闭环流程，大幅减少由于仿真过程中大量重复人为操作所产生的的误差和误差。实现高效可靠的仿真并未后续的代理模型自动寻优奠定良好的数据基础。

**代理模型的自动化构建是整个项目的核心技术突破。**研究团队针对数值指标和三维网格图像两种模型预测目标创造性地采用了参数-数值回归和参数+点云-点云的两种技术路线实现了技术目标。对参数-数值回归将三个设计参数和性能指标构建了基于采用径向基函数（RBF）的 SVR 回归模型，并通过自动化流程完成了调参等复杂工作，并成功的准确输出 5 项关键印痕性能指标预测误差小于 2%。

		训练数据范围外		训练数据范围内	
胎冠中心弧 $R_n$		1200	600	1000	800
胎冠总弧高 $h$		3	7	4	5
行驶面宽 $b$		265	220	255	230
印痕图(FEA) (以1为准, 1.22061)					
左胎肩长轴 (mm)	模型预测值	214.7	207.1	211.9	212.0
	仿真值	213.0	203.2	212	213
中心长轴 (mm)	模型预测值	214.6	247.1	222.2	235.4
	仿真值	217.0	249.2	222.4	236
右胎肩长轴 (mm)	模型预测值	214.4	216.5	212.1	212.4
	仿真值	213.0	203.2	212	213
F值	模型直接预测值	2.1	12.6	4.4	11.9
	模型预测间接计算值	0.0	17.6	5.1	11.6
	仿真值	2.0	23.0	5.2	11.5

不准

准确

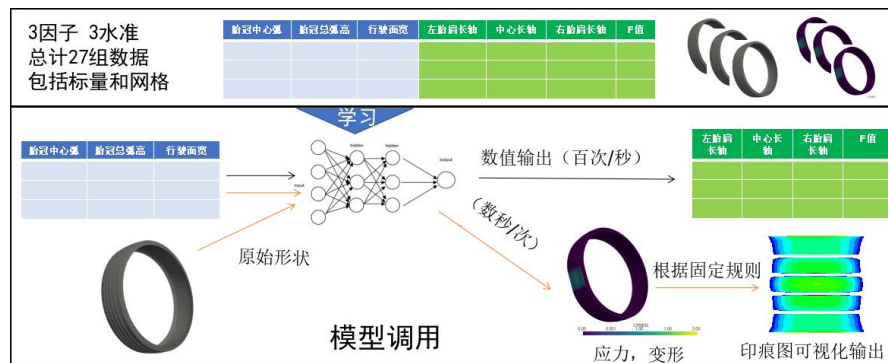
来源：中化信息技术有限公司

图 36 数值回归验证结果

## (2) 技术方案

对于参数+点云-点云预测任务，本任务的核心为基于原始点云和设计参数预测出最终的轮胎印痕图。本任务中，研究团队采用了双输入架构，将 3 个关键设计参数与原始点云坐标特征有机结合。同时也考虑了印痕图形成的复杂过程和实际物理过程。若采用从设计参数直接生成印痕图的端到端的模型构建方法，会导致机器学习模型所需要学习的物理知识过于复杂，需要数据量过大。且印痕图中存在大量的边缘信息（印痕图内侧压应力为某一数值，外侧压应力为 0，这主要是由于仿真流程上生成印痕图会使用某一压应力阈值对所形成的压应力分布图进行剪裁），难以通过简单的机器学习模型进行拟合和处理。且印痕图的生成过程中，实际是根据网格连通性和节点数值进行大量的插值处理后最终得到图像，这也给端到端的模型设计带来了很大困难。所以本方法考虑到实际印痕图仿真中采用的流程。不采用

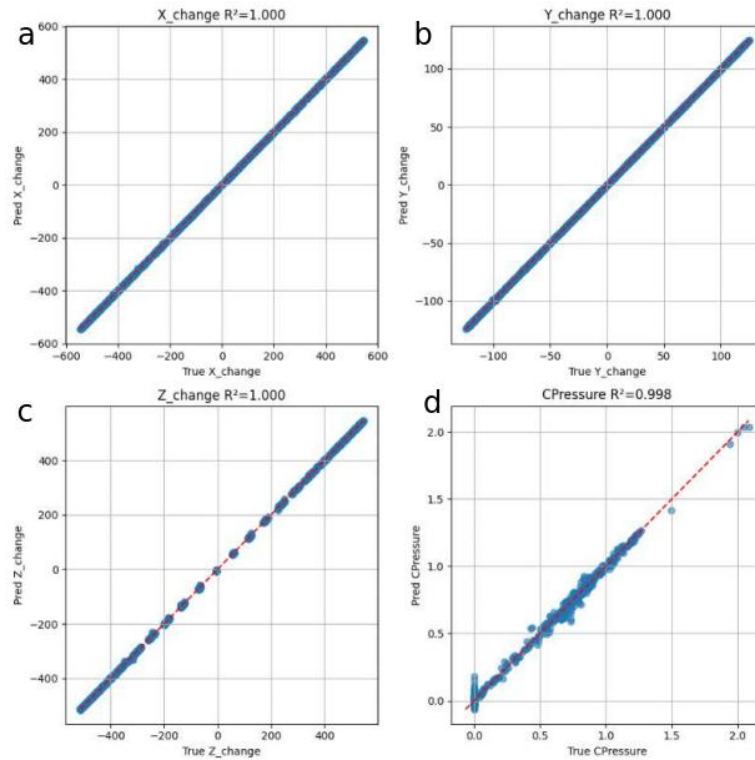
直接从设计参数生成印痕图的端到端的模型构建方法，而是通过 XGBOOST 算法构建了以原始点云+设计参数参数为输入，以变形后点云坐标和压应力值为输出的预测模型。预测生成印痕图所需要的变形和压应力两个物理量，后根据生成印痕图的固定规则生成印痕图。



来源：中化信息技术有限公司

图 37 参数-数值回归和参数+点云-点云重建

对于变形后点云的预测，在测试数据上表现良好，R2 指标均接近于 1（大于 0.9999）。相对的对于压应力的预测相对表现较差 R2=0.998。

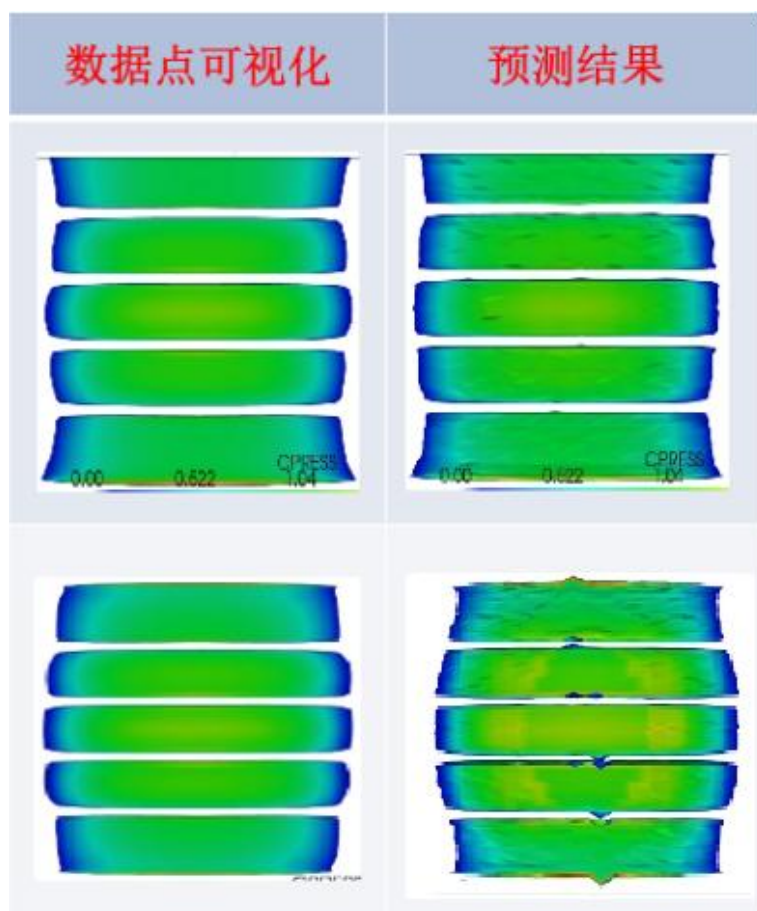


注：a, b,c: 变形后坐标(X,Y,Z)的预测值和仿真值；d: 压应力预测值和仿真值。

来源：中化信息技术有限公司

图 38 预测仿真对比

基于上述的预测结果，基于原始的网格连通关系，预测的变形后点云和预测的压应力对网格进行了重建，后对于印痕图进行了可视化。这种创新的建模方法既保留了参数化设计的灵活性，又充分考虑了轮胎接地的微观特征，实现了预测精度的大幅提升。



来源：中化信息技术有限公司

图 39 CAE 仿真结果数据与代理模型预测结果

效率提升方面单次预测时间小于 2 秒以内，较传统仿真方法提速达 9000 倍。大幅降低了计算资源消耗降低了。这些突破性的改进使得多目标优化和容差性能（考虑设计参数的统计学分布的性能预测）预测成为了可能。代理模型预测结果充分证明了代理模型技术在轮胎设计领域的巨大价值。

#### 4.创新情况：

##### (1) 技术创新情况

**数据层面**，传统仿真流程高度依赖人工操作，网格划分和参数调整耗时巨大。案例通过 HyperMesh 脚本自动化开发，将网格前处理时间从 1 小时压缩至 1 分钟，整体效率提升 60%，同时通过标准化仿

真流程，确保数据的高质量和可复用性，为代理模型训练奠定基础。

**模型架构层面**，案例创新性地提出双模态预测体系。一是参数-数值回归模型，基于径向基函数（RBF）的 SVR 模型，精准预测 5 项关键印痕指标（如 L1/L2、L0、F 值），误差控制在 2% 以内。二是参数+点云-点云预测模型：采用 XGBOOST 算法构建双输入架构，结合设计参数（R1、H、W）和原始点云数据，实现变形后点云坐标及应力分布的同步预测。测试显示，变形预测  $R^2 > 0.9999$ ，应力预测  $R^2 = 0.998$ ，精度远超传统方法。

**算法设计层面**，案例突破纯数据驱动的限制，嵌入轮胎接地的物理机理。例如，通过点云变形与应力分布的双阶段建模，更真实地反映轮胎受力过程；同时利用网格连通关系进行后处理，确保预测结果符合工程实际，避免“黑箱”模型的非物理解释。

**算力层面**，代理模型将单次预测时间从 4-6 小时缩短至 2 秒，效率提升 9000 倍，使得多目标优化和容差性能分析（如参数公差对性能的影响）成为可能。此外，通过自动化闭环流程（参数化建模→仿真→报告生成），大幅降低人为误差，提升工程适用性。

## （2）模式创新情况

传统轮胎设计依赖“试错法”，效率低下且成本高昂。案例构建了“仿真-预测-优化”全流程智能解决方案，实现设计范式和服务模式两大突破。一是设计范式变革，工程师输入参数后，2 秒内即可获取印痕性能预测，支持胎肩磨损、接地压力等多目标协同优化。二是服务模式升级，通过标准化，将高门槛的仿真技术转化为“仿真即服务”

(Simulation-as-a-Service)，计划支持 GUI 操作和自动报告生成，显著降低技术门槛。

**在应用范围上**，技术已从轮胎设计延伸至智能制造与质量控制。例如，通过容差性能预测，可量化量产轮胎的性能波动，优化生产工艺；同时，该框架可迁移至密封件、减震器等橡胶制品的力学性能优化，拓展性极强。

**在安全与可信度方面**，可视化对比（如印痕图叠加分析）增强模型可解释性，确保结果符合工程直觉。

### 5.应用实效：

**在自动化仿真方面**，案例团队针对网格前处理这一耗时环节进行了重点优化。通过对 HyperMesh 软件的脚本开发，实现了胎面网格的自动化划分。具体实施中，将原本需要 1 小时的手动前处理流程压缩到 1 分钟完成，单个方案的网格划分总时间从 2.5 小时减少到 1.5 小时，整体工作效率提升了 60%。这一改进显著减轻了仿真工程师的重复性工作负担。经实践，平均每月节约人天投入 8 个人天，大幅减轻了业务人员的业务负担，使其能够投入到价值更高的工作中。同时将开发的相关代码汇总于公司内部的知识库内，通过大模型增强检索的功能帮助仿真工程师快速的根据以往开发的脚本，实现业务上所需的功能，掌握相关能力，提升整体的公司业务水平。

**代理模型的应用是本案的核心技术突破**。研究团队构建了基于径向基函数（RBF）的支持向量回归（SVR）模型，将胎冠中心弧、胎冠总弧高和行驶面宽三个关键设计参数与五项印痕性能指标建立

映射关系。通过自动化调参流程，模型的预测误差控制在 2%以内。更突出的是，该模型将原本需要 4—6 小时的仿真流程缩短至 2 秒，效率提升达 9000 倍，同时大幅降低了计算资源消耗。这一突破使得多目标优化和容差性能预测成为可能，为轮胎设计提供了新的技术手段。

**从行业视角来看，这项研究为轮胎设计领域带来了革命性的改变。**它突破了传统试错法的效率瓶颈，推动整个行业向“仿真驱动设计”的新范式转型。展望未来，研究团队计划将这项技术扩展至全工况预测，集成前后端和优化算法，可以完成设计生成—AI 代理模型寻优-实际仿真验证的高效流程，大幅减少 CAE 仿真的仿真个数。这套解决方案的技术框架不仅适用于印痕性能预测，还构建了一个可以广泛应用于 CAE 代理模型的标准流程和方法，可以与滚阻，形变，压力分布等其他关键性能指标的仿真相结合，并配合多目的遗传算法进行优化，快速的求得帕累托前沿，帮助设计人员快速的确定实际设计方案。展现出广阔的产业化应用前景。

## （五）ChatCAD 生成式辅助工业设计

### 1. 案例实施单位：

创新奇智科技集团股份有限公司

### 2. 案例背景：

传统的 CAD 计算机辅助设计软件，操作界面复杂且使用门槛高，对操作人员具有严苛的经验要求。创新奇智率先将工业大模型技术引入工业设计领域，推出“ChatCAD 生成式辅助工业设计”。通过简单的对话问答形式，便能迅速理解设计师的创意意图，自动生成符合要求的工业设计图，还支持导出到传统工业设计软件进行手动微调。这不仅极大地提升了工业设计的效率和质量，降低使用门槛，更推动了大模型技术在工业生产上游的广泛应用。该产品目前可以应用于机械元器件设计、工程建筑图纸哑图读取、工业品外观设计、电气设计、建筑图纸设计等，将显著降低工业设计门槛。

### 3. 案例详述：

CAD 的一个重要应用方向是 PID(Piping and Instrument Diagram)，即管道与仪表流程图，是石油、化工、天然气等工业领域中不可或缺的工程设计图纸，它详细描绘了工艺流程中的管道布局、连接关系以及仪表位置等核心信息。然而，在过往长期的历史项目中，积累了海量的 PID 图纸，它们仅以 PDF 格式的哑图形式保存，若需转换为可编辑的 CAD 电子工业设计图，则必须通过繁琐且易出错的手工录入过程。这一过程不仅要求深厚的行业知识，还极为耗时费力。尽管国内外众多企业已尝试采用传统的深度学习方法进行识别与转化，但这

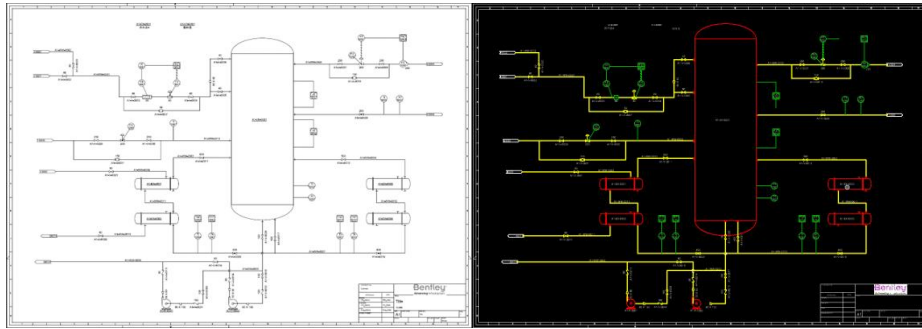
些方法在面对 PID 图纸中复杂多样的组件关系及管线识别时，往往力不从心，远未达到工业应用的实际需求。

在工程基础设施领域，创新奇智与基础设施工程软件公司 Bentley 合作，研发 PID 工艺流程图，由样图自动生成 CAD 项目。该项目主要应用 ChatCAD 的 Image-to-CAD 的能力，基于多模态工业大模型，将历史项目中的 PDF 图纸逆向转化为标准的 CAD 设计图，解决当前手工录入效率低、错误率高的问题。

iPID（Intelligent Process Piping and Instrument Diagram 智能工艺管道和仪表流程图）基于创新奇智自研的工业大模型 AIInno-75B 的“CAD 设计”创新生成技术，其主要功能可将传统图片格式的工艺和仪表流程图转化为可交互、可分析、可扩展的智能 PID，实现从静态图纸到动态智能 PID 的跨越，助力流程工业数字化转型。iPID 是一款真正实现了开箱即用，支持图像复原、属性复原和逻辑关系复原的新一代智能 PID 产品。

### （1）核心能力

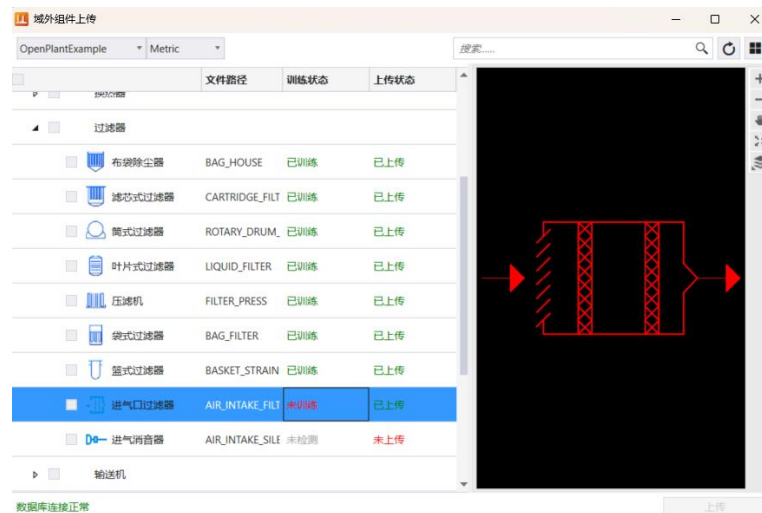
**多种格式和尺寸图纸的智能解析与生成。**iPID 利用端到端大模型技术可以直接将 pdf、jpg、jpeg、png、bmp、webp 等格式的图纸导入 Bentley 的 OpenPlant PID 中，生成可编辑、可扩展的矢量图。该功能支持 A0-A4 多尺寸图纸，解决传统 PDF 及图片图纸不可编辑、信息割裂等痛点，为后续处理提供高质量基础数据。



来源：创新奇智科技集团股份有限公司

图 40 自动生成 CAD 图

识别各种图形元素并理解图纸中的逻辑。大模型技术让 iPID 并不是简单的读图，而是可以真正理解图纸中的逻辑，可以识别并理解设备、管线、管件、阀门、仪表等图例，结合大模型预训练过的工艺图例库，精准提取图形信息，为智能 PID 生成提供关键元素。对于大模型未训练过图元图例，可通过“域外组件上传”提交给大模型，iPID 提供简单的整库上传、分类上传和单个图元上传功能，并能确保通过大模型分析后的图纸图例符号的规范性与准确性。

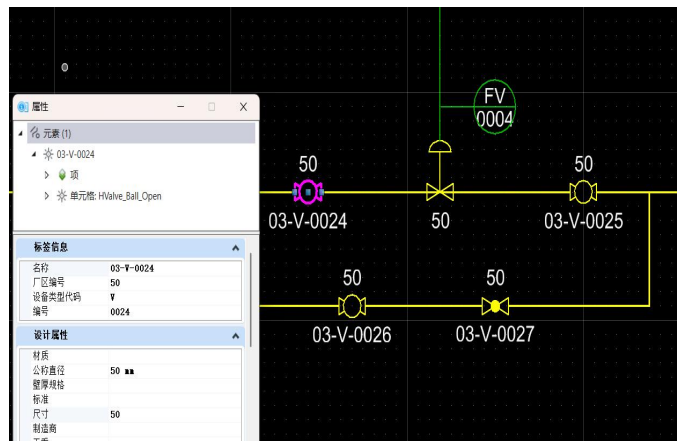


来源：创新奇智科技集团股份有限公司

图 41 ChatCAD 识别图形元素

深度集成 OpenPlant PID 设计平台。在 iPID 生成与图纸一致的矢量图后，可以进一步对图的内容进行智能编辑，如属性的修改、注释

的添加等功能。该产品不仅支持导出 dgn 或 dwg 等通用格式，同时也支持导出为多种其他格式应用与不同场景。还支持基于同一项目的多用户协同设计，属性标记，注释批改等，保障团队协作顺畅有序。



来源：创新奇智科技集团股份有限公司

图 42 内容智能编辑

**BOM 清单实时统计。**iPID 基于 OpenPlant PID 还可实现基于图纸内部的报表统计，并实现报表内容的实时刷新，同时也支持对于 BOM 清单的外部报表统计，并支持多种格式的导出再编辑，报表模版也可以基于项目要求灵活定制。

## (2) 应用场景

**设计场景，**用户可以直接将 PDF、图片等 PID 图纸内容快速转化为智能 PID，将整个设计过程加速。再基于 OpenPlant PID 内置的结构化数据和行业规范校验，实时检验设计错误，确保设计方案的合理性和可行性，提高设计质量降低项目风险。

**旧改项目场景，**用户只需将老旧 PDF 图纸导入 iPID 即可智能识别与转换，快速完成方案的改造，提高改造项目的规划和实施效率。产品内置的行业规范和企业标准，可以大幅减少因改造不当导致的安

全隐患和生产事故，保障企业改造过程的安全和顺利进行。

#### 4.创新情况：

##### (1) 技术创新情况

**自然语言驱动的工业设计生成。**突破传统 CAD 软件交互模式，基于生成式大模型技术，用户通过直观的文本输入描述设计需求，系统即可快速解析并自动生成参数化工业模型，摆脱传统软件复杂操作限制，显著降低使用门槛，实现设计流程的革命性提效。

**交互式设计修改与优化。**ChatCAD 提供强大的交互式修改功能，使用户能够对已生成的设计模型进行精细化调整，修改指令效果实时可视化呈现，确保用户能够即时查看修改效果，并根据需要进行多次迭代优化，直至达到行业交付标准。

**专业设计要求的精准解析。**依托 AIInno-75B 工业大模型对工业领域知识和专业设计标准的深度认知，ChatCAD 能够准确识别工业领域中的关键设计参数，尤其在处理含多重约束的复杂指令时，仍可精准捕捉用户意图并按标准执行设计操作。

**主流 CAD 格式无缝兼容。**输出文件全面兼容行业通用格式，实现与主流工业设计软件的无障碍交互，确保设计成果在现有生态体系中的流通应用，促进了设计成果在更广泛领域的应用与共享。

##### (2) 模式创新情况

**高效设计相应市场变化。**传统工业设计流程中，设计师需要耗费大量时间和精力在图纸的绘制和修改上。然而，ChatCAD 通过自然语言处理技术的运用，实现了设计的快速生成和优化，从而将设计周

期大大缩短。这意味着企业能够更快地响应市场变化，抢占先机。

**自然语言完成工业设计降低设计门槛。**传统的工业设计软件操作复杂，往往需要经过专业培训才能上手。然而，ChatCAD 的易用性设计使得即使是非专业人士也能通过简单的自然语言输入轻松参与到设计过程中来。这无疑降低了设计的门槛，让更多人能够发挥自己的创意和想象力。

**灵活优化满足设计多样化需求。**传统的工业设计流程往往较为固定，一旦设计初稿确定，后续的修改和优化工作可能会变得繁琐且耗时。基于 ChatCAD 的辅助，设计师可以通过自然语言随时调整设计参数，无论是尺寸、形状还是参数特性，都能迅速作出相应的调整，并实时展示修改后的设计效果。这种灵活性使得设计师能够更好地满足客户多样化的需求，从容应对设计变更的挑战。

**智能化设计提升设计创新性。**作为大模型应用，ChatCAD 具有强大的学习能力，能够持续地从设计师的反馈中进行学习，逐渐理解其设计风格和偏好，自主完成复杂的设计任务，不仅减轻了设计师的工作负担，同时提高了设计的整体质量和创新性。

## 5.应用实效：

工作效率提升 85%，绘制一张图的时间从平均 8 小时，降至平均 1 小时。基于更高效的架构设计，可以在消耗更低算力的情况下实现更全面的信息提取。同时具备标准用户认证与鉴权机制，严格把控数据访问权限，确保信息的机密性、完整性和可用性。产品实现了易部署、规模易扩展、服务能力高可靠。

未来，随着技术的不断进步和市场需求的不断增长，AI 加持的工业软件 ChatCAD 有望在多个领域发挥重要作用，推动整体设计行业的创新和发展。比如，在机械设计领域，ChatCAD 通过自然语言驱动的生成式设计，能够高效创建复杂的机械部件和整体设计方案。在电气设计领域，ChatCAD 将支持通过自然语言交互，生成准确的电路图和电气布局设计方案。此外，与高级电气仿真软件的结合使得生成的电气设计能够进行实时仿真和分析，帮助工程师在设计阶段发现并解决潜在问题，降低后期修改和调试的成本和风险。在建筑设计中，ChatCAD 能够快速生成建筑模型和平面布局图，提供设计师更直观的设计参考。在电子设计中，ChatCAD 可以帮助工程师快速创建电路板和电子元件布局。ChatCAD 还可以在教育培训领域发挥重要作用，帮助学生和初学者更轻松地学习工业设计软件操作，并提升他们的设计技能。

## （六）面向高端制造业的多模态交互式预测性智能运维决策支持关键技术

### 1.案例实施单位：

北京中科汇联科技股份有限公司

### 2.案例背景：

为推动创新和经济增长，优化决策和公共服务，开放公共数据有助于公众更好地了解政府行为，从而加强政府的透明度和公信力。在保证开放数据的同时也面临许多挑战，如确保个人隐私和敏感信息需要得到充分保护。确保开放的数据具有高质量和一致性，避免误导公众或造成不必要的混乱。为促进数据的广泛应用和共享，需要建立统一的数据标准和确保数据的互操作性。以及有效的治理机制和解决方案是确保公共数据开放可持续性和可访问性的关键，也是构建可信赖的数据开放生态系统的基石。针对以上内容中科汇联自主研发了AiMAX数智资产管理决策平台以及中科汇联BIM云平台，以BIM为基础，立足服务于实际项目应用，提升企业的BIM项目应用能力和管理能力的目的。

### 3.案例概述：

旨在帮助组织有效地管理和利用数据资源。提供了一系列的功能和工具，通过这些功能，数据中台能够帮助组织建立统一的数语义和格式，确保数据的一致性和准确性，优化数据的存储和利用方式，提高数的可信度和可用性，保护数据的安全和合规性，以及管理数的整个生命周期。数据中台的基础构成为组织提供了一个全面而可靠的数

据管理基础，帮助组织实现数据驱动的决策和创新，提升业务效率和竞争力。

#### 4.案例详述：

##### (1) 总体思路

该项目基于中科汇联自主研发的 AiMAX 智资产管理决策平台的数据中台是一个综合性的数据管理系统，旨在帮助组织有效地管理和利用数据资源。它提供了一系列的功能和工具，通过这些功能，数据中台能够帮助组织建立统一的数据语义和格式，确保数据的一致性和准确性，优化数据的存储和利用方式，提高数据的可信度和可用性，保护数据的安全和合规性，以及管理数据的整个生命周期。数据中台的基础构成为组织提供了一个全面而可靠的数据管理基础，帮助组织实现数据驱动的决策和创新，提升业务效率和竞争力。

##### (2) 平台介绍

**AiMAX 数智资产管理决策平台主要由以下基础内容构成：**

**数据集成：**数据集成是指将不同来源的数据整合到一个统一的数据仓库中，以便进行更加全面和准确的数据分析。数据集成可以帮助企业更好地了解市场和客户需求，制定更加精准的业务决策和策略。当前数据集成的技术包括：数据源管理、数据调度、任务编排、注册中心、任务模板配置、执行器管、调度日志工具等。

**数据治理：**数据治理是指通过规范化、标准化等方式，对数据进行管理和维护，确保数据的一致性、准确性和可靠性。数据治理可以帮助企业更好地利用数据，提高数据分析的精度和效率。当前数据治

理的技术趋势包括：数据标准化、数据质量监控、数据治理工具等。

**数据质量：**数据质量是指数据的准确性、完整性、一致性、可靠性等方面的表现。数据质量对于数据分析和业务决策具有至关重要的作用。当前数据质量的技术趋势包括：数据清洗、数据标准化、数据质量监控等。

**数据服务：**数据服务是指将数据提供给用户进行业务分析和决策。数据服务可以帮助企业更好地理解市场和客户需求，制定更加精准的业务决策和策略。当前数据服务的技术趋势包括：API 设计、微服务架构、数据安全保障等。

**运维管理：**数据仓库和数据分析系统需要不断进行运维管理，以确保其正常运行和数据安全。运维管理包括备份和恢复、监控和故障排除、性能优化等方面。当前运维管理的技术趋势包括：自动化运维、DevOps 思想、数据安全保障等。

**在 BIM 云平台方面包括以下几个内容：**

**统筹平台应用管理：**统筹行业、股份公司的 BIM 云平台应用，对整个股份公司的 BIM 相关工作进行全量管理，解决分子公司平台建设的各自为战情况，避免重复性建设。

**促进资源共享：**生产、管理数据统一管理、全局协调，协调了部分地区资源闲置、部分地区资源缺失的情况；行业、企业规范标准自由上传、预览、下载。

**打造 BIM 人才聚集地：**成为行业或企业的 BIM 人才聚集地，避免部分项目、处级公司的 BIM 人才孤军奋战的情况，加强人才之间

的技术、经验交流，企业之间的人才培养、培养体系建设的经验交流。

### (3) 技术方案

为推动创新和经济增长，优化决策和公共服务，开放公共数据有助于公众更好地了解政府行为，从而加强政府的透明度和公信力。在保证在开放数据的同时也面临许多挑战，如确保个人隐私和敏感信息需要得到充分保护。确保开放的数据具有高质量和一致性，避免误导公众或造成不必要的混乱。为促进数据的广泛应用和共享，需要建立统一的数据标准和确保数据的互操作性。以及有效的治理机制和解决方案是确保公共数据开放可持续性和可访问性的关键，也是构建可信赖的数据开放生态系统的基石。针对以上内容中科汇联自主研发了 AiMAX 数智资产管理决策平台及 BIM 云平台，在大数据技术方面探索了以下几个重点内容。

**可视化建仓和数据建模：** 中科汇联致力于开发可视化建仓工具，使数据仓库的搭建和管理更加简化和高效。同时研究数据建模技术，以支持数据的结构化和语义化描述，为数据的组织、查询和分析提供更好的支持。

**多种数据集成模式：** 提供多种数据集成模式，包括向导模式、脚本模式、API 模式和文件集成。这些模式可以满足不同用户的需求，使数据集成过程更加灵活和便捷。

**多种数据服务方式：** 提供多种数据服务形式，包括 API、数据集、数据表、主题域和数据编目。这些服务可以帮助用户更好地管理和利用数据资源，实现数据的高效共享和价值挖掘。

**字段级别的数据血缘支持：**支持字段级别的数据血缘追踪，可以跟踪数据从源头到目标的流动路径，帮助用户了解数据的来源和变化过程。

**元数据分类的值传递：**通过数据血缘实现元数据分类值的传播，确保元数据的分类值在数据变化时能够及时更新，保持准确性和一致性。

**数据标准映射：**支持手动映射和自动映射两种数据标准映射模式，帮助用户实现数据的一致性和标准化。

**数据运维：**同时支持对主机集群、服务以及告警的监测，确保数据平台的稳定运行和及时响应。

**可插拔的数据运维组件：**提供可插拔的数据运维组件，可以根据用户需求进行定制和扩展，满足不同业务场景的需求。

**数据治理业务一体化：**支持数据集成、元数据管理、数据质量检核、数据对外服务的提供以及数据运维一体化业务服务。这种一体化的业务服务可以帮助组织实现数据的全面管理和有效利用，提升数据治理的效率和效果。

## 5.创新情况：

### (1) 可视化建仓、数据建模

采用先进的可视化工具，使用户能够以图形化界面进行数据仓库的建模和设计，大大降低了技术门槛，提高了建模效率。与其他项目相比，可视化建仓工具更加直观易用，且支持更丰富的数据建模功能，使用户能够更轻松地进行数据仓库的构建和管理。

## (2) 多种数据集成模式

持多种数据集成模式，包括向导模式、脚本模式、API 模式和文件集成，以满足不同用户的需求和习惯。与竞争对手相比，集成模式会更加灵活多样，覆盖了更广泛的集成方式，使用户能够更灵活地处理不同来源的数据。

## (3) 多种数据服务方式

通过提供多种数据服务方式，包括 API、数据集、数据表、主题域和数据编目，满足了不同业务场景下的数据访问需求。与其他类似项目相比，数据服务方式更加多样化，覆盖了更广泛的业务需求，提供了更丰富的数据服务功能。

## (4) 支持字段级别的数据血缘

支持字段级别的数据血缘追溯，能够精确记录数据的来源和变化过程，帮助用户更好地理解数据的真实性和可信度。相较于其他项目，数据血缘功能更加细致和精确，能够提供更全面的数据追溯和分析能力。

## (5) 元数据分类的值传递

实现了元数据分类值通过数据血缘传播的功能，使得元数据的分类信息能够自动传递和更新，减少了人工干预的成本和错误率。这一功能相对于其他项目来说，提高了数据管理的效率和准确性，使得数据的元数据更加准确和完整。

## (6) 数据标准映射

支持手动映射和自动映射两种模式，用户可以根据实际需求选择

合适的映射方式，提高了数据标准化的灵活性和效率。相比于其他项目，映射功能更加智能化和灵活，能够满足不同数据标准化需求的场景。

### （7）数据运维

同时支持对主机集群、服务以及告警的监测，实现了全方位的数据运维管理，保障了系统的稳定性和可靠性。与竞争对手相比，数据运维功能更加全面，覆盖了更广泛的监控对象，提供了更强大的运维管理能力。

### （8）可插拔的数据运维组件

提供可插拔的数据运维组件，用户可以根据实际需求灵活选择和扩展运维功能，满足不同场景下的运维需求。相对于其他项目，可插拔组件更加开放和灵活，能够更好地适应不同用户的运维需求和环境。

### （9）数据治理业务一体化

支持数据集成、元数据管理、数据质量检核、数据对外服务等业务的一体化管理，实现了数据治理全流程的闭环管理。与其他项目相比，数据治理业务一体化能力更加强大，能够提供更全面和高效的数据治理服务。

### （10）DAG 有向无环图与数据血缘结合研发

将 DAG 有向无环图与数据血缘结合，实现了数据流程的可视化展示和数据血缘的动态追踪，帮助用户更好地理解数据流程和关系。这一功能在同类项目中较为独特，提升了数据管理和分析的可视化程度，增强了用户对数据的理解和信任。

### (11) 支持 flinkx 和 datax 采集二次开发

支持对 flinkx 和 datax 采集工具进行二次开发，扩展了数据采集的灵活性和适用性，满足了更多数据源的接入需求。相对于其他项目，二次开发能力更加强大，能够满足更复杂和多样化的数据采集需求，提供更全面的数据接入解决方案。

### (12) hdp 集群及各组件适配国产系统

对 hdp 集群及各组件进行了适配，实现了与国产系统的无缝集成，提高了系统的兼容性和适用性。相较于其他项目，适配能力更加全面，能够满足国内用户对国产系统的需求，降低了系统集成和部署的难度。

### (13) 通过绘图实现可视化建设数仓分层、分区和表的建设

通过绘图工具实现了数仓分层、分区和表的可视化建设，使用户能够清晰地了解数仓结构和数据布局，提高了数据管理的效率和可靠性。相对于其他项目，可视化建设功能更加直观和易用，帮助用户更快速地构建和管理数据仓库。

### (14) 信创适配、国产化主流数据库、中间件支持

支持信创适配、国产化主流数据库和中间件，满足了国内用户对国产化软件的采用需求，提高了系统的适用性和可靠性。这一特点相较于其他项目来说，增强了系统在国内市场的竞争力，使得更多国内用户能够选择我们的解决方案

## 6.应用实效：

### (1) 社会效应

**提升政府透明度与公信力：**通过开放公共数据，政府行为能够更

透明地被公众了解，从而增强公众对政府的信任。例如，中科汇联开发的 AiMAX 数智资产管理决策平台和 BIM 云平台，为政府数据的开放提供了技术支持，确保数据的高质量和一致性，避免误导公众，同时加强数据隐私保护。

**促进数据共享与协同治理：**数据中台的建设打破了信息孤岛，促进了跨部门、跨区域、跨层级的数据共享与协同管理。例如，政务数据中台的建设实现了政务数据的统一管理，提升了政府服务的效率和质量。此外，通过数据共享，政府能够更好地支持公共服务和社会治理，提高民众的满意度。

**推动行业创新与人才培养：**中科汇联 BIM 云平台的建设不仅提升了企业的 BIM 应用能力，还成为行业人才聚集地，促进了技术交流和经验分享，避免了人才孤岛现象。

**助力智慧城市建设：**数据中台在智慧城市中的应用，如城市规划、公共安全、交通管理、建筑施工等，为城市治理提供了科学依据，提升了城市管理的智能化水平。

## (2) 经济效益

**提高企业运营效率：**数据中台通过整合企业内外部数据资源，打破数据孤岛，实现数据的标准化和一致性，从而提升数据利用率，降低运营成本，优化企业决策流程。例如，通过数据中台，企业能够更快速地响应市场变化，优化资源配置，提高业务效率和竞争力。

**促进业务创新与增长：**数据中台为企业提供了丰富的数据资源和分析工具，支持企业进行业务创新和战略规划。例如，通过数据挖掘

和机器学习，企业可以预测市场需求变化，优化生产计划，发现新的业务机会。

**降低数据处理成本与风险：**数据中台通过集中管理和自动化处理，降低了数据存储、处理和分析的成本，同时通过数据加密和权限管理，保障了数据的安全性和隐私性。

**推动产业链协同创新：**数据中台通过数据共享和开放平台，促进了企业与外部合作伙伴的协同创新，推动了产业链上下游的协同运作，从而提升了整体产业的竞争力。

**支持中小企业融资与数字化转型：**数据中台通过提供安全的数据共享机制，帮助中小企业获取融资支持，同时助力中小企业实现数字化转型，提升其市场竞争力。

**推动高端制造业智能化升级：**中科汇联 BIM 云平台通过数据建模、可视化建仓和数据血缘追踪等技术，提升了制造业企业的项目管理能力和生产效率，降低了生产成本，推动了制造业向智能化转型。

综上所述该案例通过构建数据中台和智能运维平台，实现了数据资源的高效利用、业务流程的优化、创新能力的提升以及社会透明度的增强。这些成果不仅提升了政府和社会的治理能力，也为企业的数字化转型和产业升级提供了有力支撑。同时，通过数据安全和隐私保护机制的完善，进一步保障了数据开放和应用的可持续性。

## （七）基于大数据的色漆漆料优化系统

### 1.案例实施单位：

上海波士内智能科技有限公司

### 2.案例概述：

作为国民经济的支柱产业之一，汽车对于保持经济的平稳增长、稳定社会发展全局具有重要的战略意义。国家明确提出了到 2025 年，智能制造能力成熟度水平明显提升的转型升级目标；作为汽车生产的四大工艺之一，涂装在汽车生产中扮演着非常重要的角色，它不仅影响着汽车外观的美观度和质感，还具有保护车身金属、提高耐久性和抗腐蚀能力的功能。涂装过程并不是简单的喷涂，而是一个复杂的工艺流程，涉及到底漆、面漆、清漆等多道工序。基于大数据的色漆漆料优化系统旨在基于色漆各维度的参数进行大数据分析，给出满足工艺质量要求下的最优生产参数推荐，并且推荐的生产参数相较于现场的生产参数有油漆耗量的减少。同时系统能给出单车油漆耗量统计和分析结果，并有生产参数变化前后耗量对比功能。目前，喷涂参数设置往往基于工程师多年的经验积累或者喷涂机器人预设的基础数据，需要通过人工优化来完成，耗费大量的时间和精力。这种传统的人工调整方法虽然有效，但也存在一定局限性，对工程师的要求极高。

### 3.案例详述：

#### （1）总体思路

本项目《基于大数据的色漆漆料优化项目》在技术上具有创新性，采用了大数据分析、人工智能和深度学习等先进技术。通过使用这些

技术，项目显著缩短了喷涂参数设置的时间，减少了对工艺工程师经验的依赖；并且确保推荐的喷涂参数符合工艺质量要求，同时降低油漆消耗，提高生产效率。通过项目的优化，实现了更高效的生产流程，提升了产品质量，并为企业节约成本，同时也积极响应环保理念。



来源：上海波士内智能科技有限公司

图 43 色漆工艺智能管理系统

## (2) 技术方案

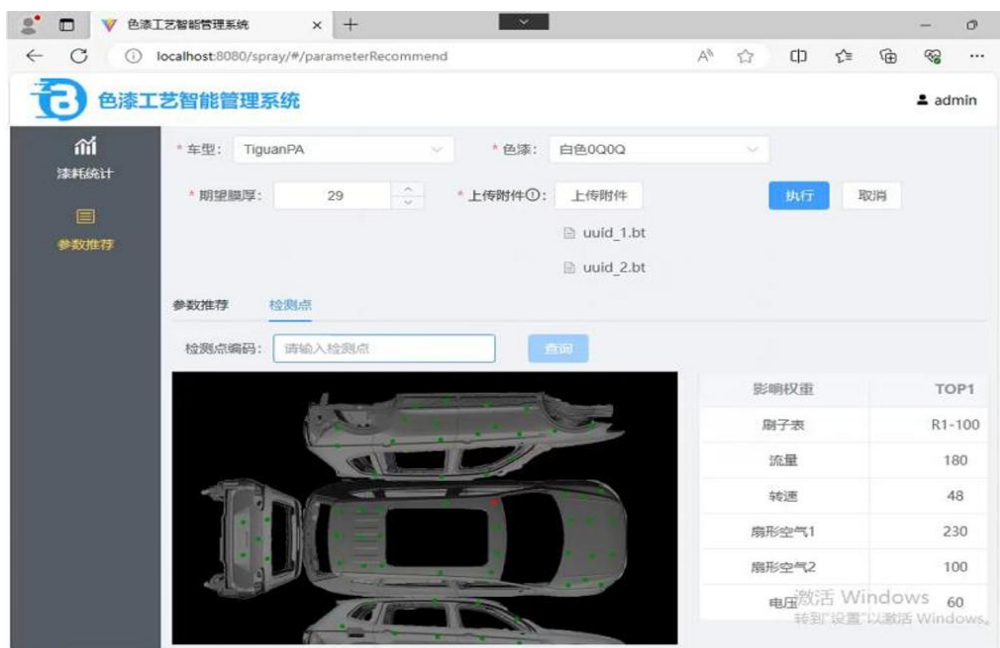
一是参数推荐根据用户输入的车型、色漆、期望膜厚以及上传的喷涂参数文件，系统应该能够自动给出相对最优的推荐参数。这样可以减少工艺工程师手动调整参数的时间，同时不依赖工程师的经验，新手工程师也能够获得相对最优的推荐参数。二是根据设定的喷涂参数和车辆信息预测膜厚根据当前设置的喷涂参数，系统应该能够预测出每个检测点的膜厚。这样可以帮助工程师更好地了解当前参数设定下的膜厚情况，并通过热力图的形式直观的展现出来，有助于优化生产流程。三是针对单测点的刷子号调整推荐针对单个检测点膜厚较厚

的情况，系统提供了针对单测点的刷子号调整推荐功能。通过这一功能，可以根据特定检测点的情况，实现针对性、细粒度的参数调整，进一步优化推荐的喷涂参数。这种精细调节能够有效提升产品品质，并确保每个检测点的膜厚达到最佳状态。通过系统的智能推荐和调整功能，工程师可以更加高效地进行优化，实现更好的生产效果。

### (3) 关键指标参数

本项目《基于大数据的色漆漆料优化项目》的预期技术指标如下：

- 1、膜厚预测准确率达到 90%以上，膜厚预测值和实际值的偏差不高于 0.5 微米；
- 2、推荐参数可以达到不小于 5%的色漆耗量优化结果，并能根据期望膜厚进行调整。



来源：上海波士内智能科技有限公司

图 44 膜厚预测技术指标

## 4. 创新情况：

本项目《基于大数据的色漆漆料优化项目》利用大数据分析、人工智能和深度学习等先进技术，实现了基于大数据分析的单测点刷子

号调整推荐，基于深度学习的参数推荐和膜厚预测功能。

**基于大数据分析的单测点刷子号推荐调整技术系统**对采集到的多维度数据，包括喷涂工艺参数、材料参数和环境参数等进行大数据分析并建立分析模型，实现针对性、细粒度的参数调整，进一步优化推荐的喷涂参数。

**基于深度学习的参数推荐和膜厚预测技术系统**能够根据用户的输入自动给出相对最优的推荐参数，并根据当前设置的喷涂参数，预测出每个检测点的膜厚，帮助用户更好地了解 and 掌握生产过程中的关键指标。实现生产过程中的智能化和高效化，为企业提供更可靠、精准的生产决策支持本项目《基于大数据的色漆漆料优化项目》将显著缩短了喷涂参数设置的时间，减少了对工艺工程师经验的依赖；降低油漆消耗，实现更高效的生产流程，提升了产品质量，并为企业节约成本，同时也积极响应环保理念。根据期望膜厚以及上传的喷涂参数文件，自动给出相对最优的推荐参数；根据设定的喷涂参数和车辆信息预测膜厚；针对单测点的刷子号调整推荐。

### 5.应用实效：

目前的喷涂参数设置往往基于工程师多年的经验积累或者喷涂机器人预设的基础数据，需要通过人工优化来完成，耗费大量的时间和精力。这种传统的人工调整方法虽然有效，但也存在一定局限性，对工程师的要求极高。通过监测不同喷涂参数设置前后的油漆消耗量以及检测点膜厚是否符合工艺质量要求，以便现场工艺工程师能够及时调整工艺。同时，借助知识图谱、大数据分析和算法建模，综合分

析多个影响因素，以创建膜厚预测模型和喷涂参数推荐模型。膜厚预测模型可根据设定的喷涂参数预测检测点的膜厚值，为现场工程师的优化调整提供支持；而喷涂参数推荐模型则可根据所需的工艺膜厚质量要求，提供最佳喷涂参数，从而减少工程师的工作量、提高工作效率，并达到节省油漆消耗的目的，同时可以有效提高生产效率、降低成本，并确保生产质量的稳定性与一致性。通过采集现场数据并结合知识图谱、大数据分析和算法建模，新型的膜厚预测模型和喷涂参数推荐模型将为企业带来更大效益，具有巨大的市场潜力和前景。本项目方案通过对海量数据和实时生产情况的监控与分析，能够实现精准的参数推荐和膜厚预测，提高了生产过程的稳定性和可靠性。综合来看，本项目喷涂工艺优化方案在技术上经过充分验证与实践，具有较高的成熟度和可靠性，能够为客户提供稳定、高效的解决方案，帮助客户实现生产效率的提升和产品质量的保障。相关技术正在申请核心自主知识产权（计算机软件著作权）。

## （八）基于 MWORKS 的具身空间仿真平台

### 1. 案例实施单位：

牵头单位：北京理工大学

参与单位：苏州同元软控信息技术有限公司（系统建模与仿真平台支持）

### 2. 案例背景：

针对智能无人系统在复杂现实场景下面临的训练成本高、跨场景适应性弱及安全验证难等瓶颈，北京理工大学周治国团队与同元软控联合研发基于 MWORKS 的具身空间仿真平台。该平台以创新提出的“具身域”为核心，融合信息域（算法与数据处理）、物理域（真实环境交互）与虚拟空间（高精度建模），构建涵盖感知、决策、执行闭环验证的仿真训练环境。通过高保真动态场景建模（如海洋扰动、多天气模拟）与数字孪生技术，实现无人系统（如无人船、无人机）从虚拟训练到物理样机的低成本高效迁移，提升其在复杂水域、城市环境中的自主导航与多任务协同能力，推动具身智能技术在工业自动化、智慧城市等领域的产业化落地。

### 3. 案例概述：

该平台基于 MWORKS 系统建模仿真技术，融合具身域、信息域与物理域，集成三维虚拟引擎、强化学习框架及多模态感知算法库，构建高保真动态仿真环境（支持多天气、海风扰动等复杂场景建模）。通过数字孪生监控与自定义算法接口，实现无人系统（如无人船、无人机）的感知决策优化与“数字样机训练-物理样机部署”全链路闭环

验证，应用于工业自动化、智慧城市等领域，显著提升智能体在复杂水域、城市环境中的自主导航、集群协同与任务执行能力，并以具身智能无人船为例验证跨场景迁移应用价值。

#### 4.行业痛点：

智能无人系统研发面临真实环境训练成本高、极端场景数据获取难、跨场景算法泛化弱、物理样机安全验证风险大等多重瓶颈。具体表现为：**一是**依赖实体硬件测试导致硬件损耗与周期成本攀升；**二是**传统仿真建模对动态物理规则（如流体扰动、多智能体协作）还原度不足，制约算法适应性；**三是**物理样机直接部署易引发失控隐患，且缺乏虚实联动的闭环验证机制；**四是**感知、决策、控制等多模态技术开发分散，亟需统一平台整合数字孪生建模、强化学习训练与硬件部署全流程。本案例通过具身域技术融合高精度物理建模与虚拟仿真，构建低成本、高保真训练环境，并依托数字孪生与跨场景迁移能力，破解安全验证与产业化落地难题。

#### 5.项目获奖、专利授权及获得投融资情况：

2022 年获中国造船工程学会科技进步三等奖；

2024 年获江苏省电力科学技术进步一等奖。

#### 6.案例详述：

##### （1）总体思路

北京理工大学周治国教授团队与苏州同元软控联合研发的基于 MWORKS 的具身空间仿真平台，以“虚实共生”为核心理念，旨在解决智能无人系统在复杂物理环境中训练成本高、场景适配性差的核心

痛点。团队提出“具身信息物理系统（ECPS）”理论框架，将智能体的物理实体、环境交互与数字模型深度融合，构建了一个支持感知、学习、推理和行动的全链路仿真生态。平台通过信息域、物理域、具身域的三域协同，实现从虚拟训练到实体部署的无缝衔接，为无人船、无人机、无人车等系统提供安全、高效且可扩展的算法验证环境。



来源：北京理工大学

图 45 具身域、信息域、物理域

**信息域（Cyber System Domain）**指现代产品中负责数据的处理、传输、存储和管理的虚拟系统，包括控制算法设计、数据分析、人工智能算法等。**物理域（Physical System Domain）**指现代产品系统中与现实世界直接交互的部分，包括物理实体、设备和过程，其囊括机械、电气、流体、热、磁等各个专业学科。**具身域（Embodied Domain）**指机器人或智能体所处的物理环境以及其自身物理结构所构成的一个整体范畴，这个范畴对智能体的行为、感知、学习和决策有着根本性的影响。具身域可视为一个由物理实体（智能体及其所处的物理环

境)、信息流(传感器-处理器-执行器的信息交互)和功能规则(身体的物理规律、环境约束以及智能体的行为逻辑)所构成的复杂系统,包括具身信息物理系统和具身空间。

**具身信息物理系统 (Embodied Cyber Physical System)** 是一种将具身智能体深度集成到信息物理系统框架中的新型系统架构。它不仅包含了传统 CPS 中的物理系统、网络系统和信息系统,还特别强调具身智能体在其中的核心作用。**具身空间 (Embodied Space)** 是一个综合性的智能环境,有机结合了物理空间、虚拟空间和数据空间,融合了虚拟现实、计算通信、感知融合、决策推理和机器人控制等多种先进技术。旨在支持无人系统(如机器人)进行感知、学习、推理和行动,从而提升其具身智能。具身空间的核心目标是提供一个高度仿真的训练和测试环境,使机器人能够在其中学习并适应各种复杂场景,进而将这些能力迁移到真实物理世界中。通过这种方式,具身空间赋予了无人系统自主学习和适应复杂环境的能力,使其能够执行各种高级任务。

## (2) 平台介绍

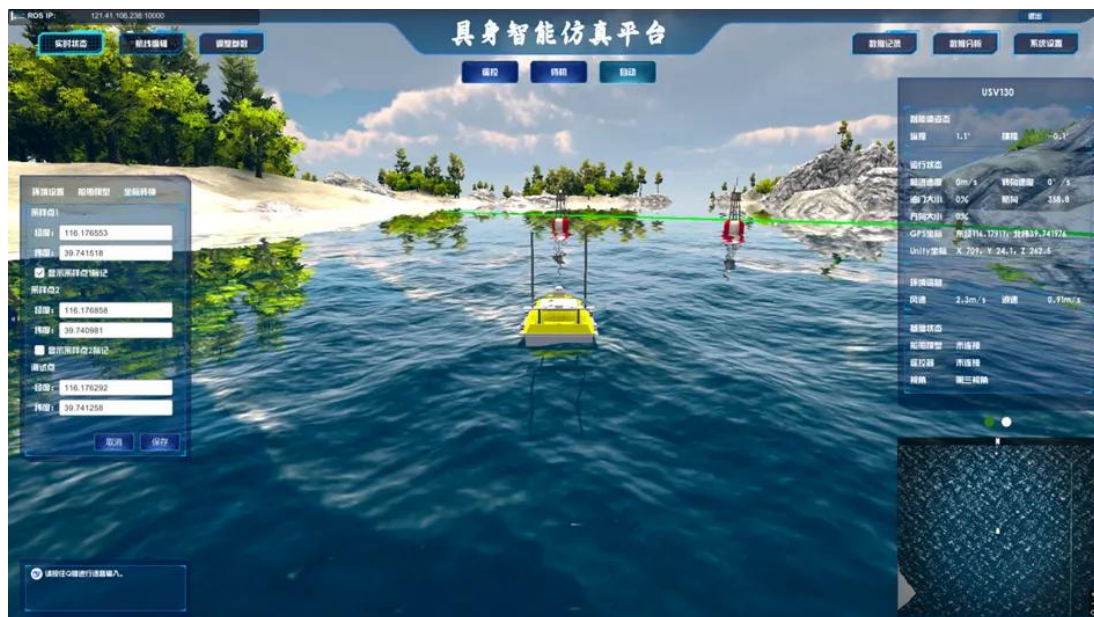
基于 MWORKS 的具身空间仿真平台是一个高度综合的人工智能平台,专注于具身空间建模、无人系统建模和仿真测试。它通过集成物理空间、虚拟空间和数据空间,支持无人系统的感知、学习、推理和行动,旨在提升无人系统的具身智能。在具身空间仿真平台框架中,具身空间是一个核心概念,它结合了物理空间、虚拟空间和数据空间,构建了一个支持无人系统具身智能感知、学习、推理和行动的

综合性环境。具身空间仿真平台框架通过其关键模块（如虚拟仿真引擎、生成智能引擎和推理进化引擎等）来实现具身空间的构建和功能支持。



来源：北京理工大学

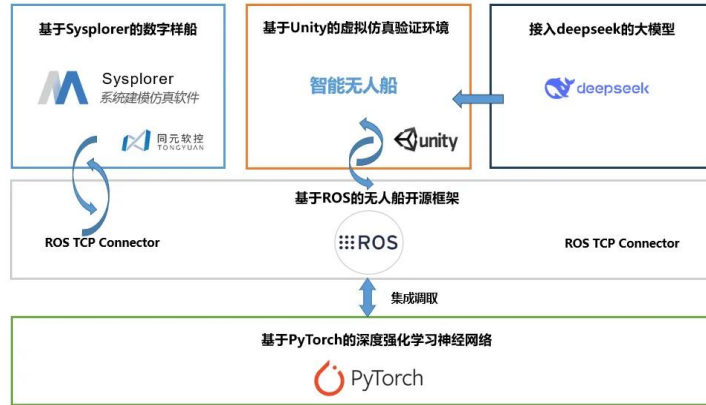
图 46 具身空间仿真平台社区



来源：北京理工大学

图 47 具身空间仿真平台

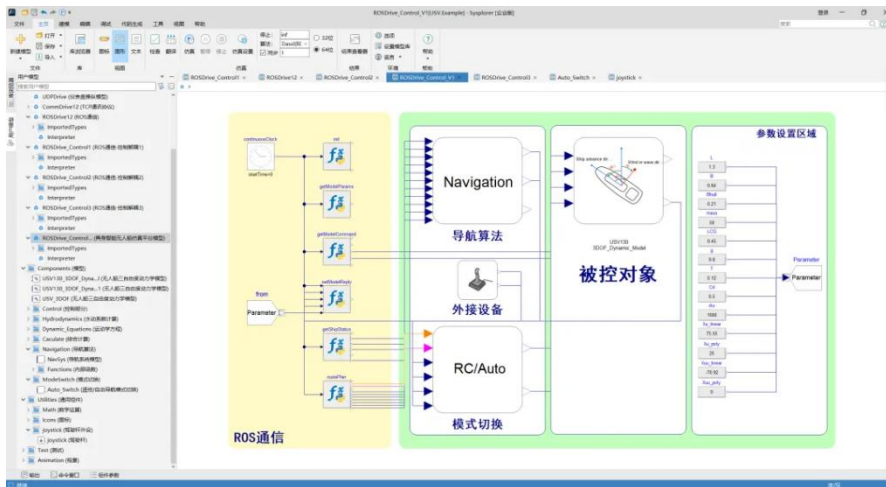
该平台以高精度三维虚拟仿真引擎为基底，结合 MWORKS 的系统建模能力，构建了多层级技术架构：



来源：北京理工大学

图 48 平台架构与核心功能

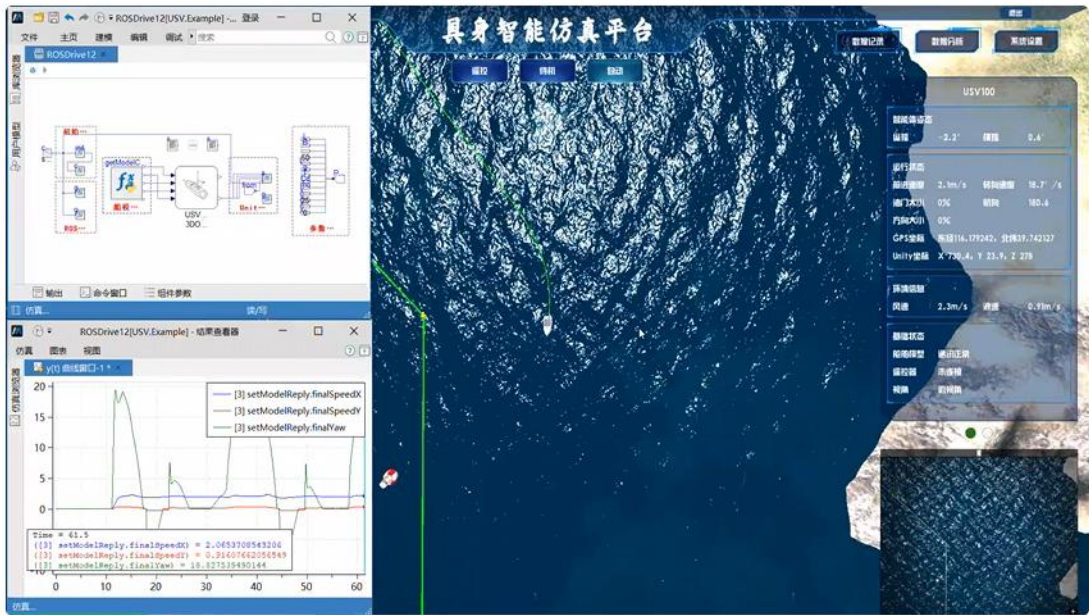
底层物理引擎集成流体力学、机械动力学等真实物理规则，支持海况风浪、多天气条件（雨雪、昼夜）的动态模拟，确保虚拟环境与真实世界的物理一致性。例如，在无人船案例中，平台可模拟六级海况下的波浪扰动对船体运动的非线性影响。



来源：北京理工大学

图 49 底层物理引擎

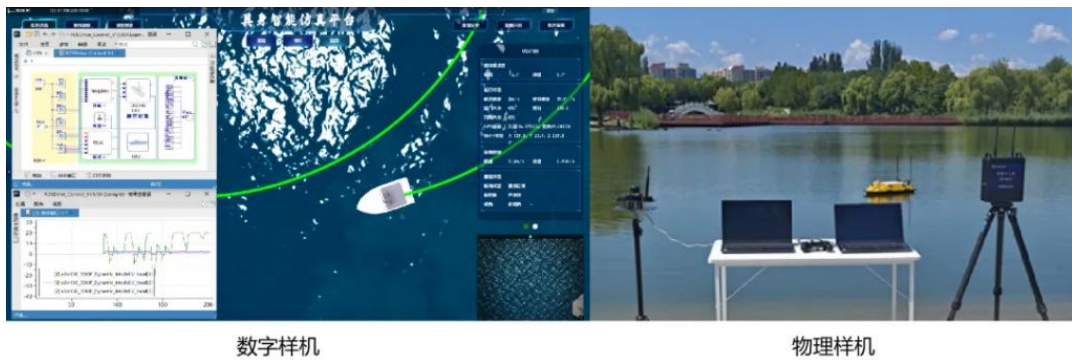
中台智能训练模块内置 SLAM 建图、路径规划、多智能体协同等算法库，兼容 ROS 通信协议与 Python 自定义接口。通过一键式强化学习训练，用户无需关注 Unity 引擎与 ROS 的底层适配，可直接调用预置环境进行算法迭代，训练效率提升 60%以上。



来源：北京理工大学

图 50 中台智能训练模块

顶层数字孪生系统实时同步物理样机与虚拟环境数据，支持虚实双向映射。例如，无人船实船航行数据可动态反馈至虚拟场景，用于优化导航策略；反之，虚拟训练的避障模型也能直接部署至实体船体实现“训练-验证-部署”闭环。



来源：北京理工大学

图 51 中台智能训练模块

### (3) 技术方案和创新突破

平台的核心技术突破体现在具身域建模与多模态智能进化两大方向。具身域建模技术，首创“智能体-环境耦合”模型，将机器人的物理结构（如无人船的船体动力学参数）与任务场景（如港口航道约

束) 统一编码为可计算的数学表达。通过 MWORKS 的 Modelica 多学科建模语言, 平台可自动生成包含机械、电气、流体力学的联合仿真模型, 支持毫秒级实时解算; **多模态智能训练框架**, 融合视觉、激光雷达、惯导等多传感器数据流, 结合 DeepSeek 大模型的自然语言指令解析能力, 实现跨模态语义对齐。例如, 无人船可通过语音指令接收任务目标, 同时结合视觉感知自主规划航行路径, 并在训练中动态优化决策逻辑。

#### (4) 关键指标参数

关键指标参数如下表:

表 2 关键指标参数

指标类别	参数详情
物理仿真精度	流体动力学模拟误差<3%, 机械系统响应延迟<10ms
多智能体协同规模	支持 50+智能体集群训练, 任务冲突消解成功率≥95%
算法训练效率	强化学习收敛速度较传统方案提升 2.5 倍, GPU 资源利用率达 90%
硬件兼容性	支持 ROS1/ROS2、NVIDIA Jetson、STM32 等主流控制器, 协议适配种类超 20 种

### 7. 创新情况:

#### (1) 技术创新情况

**数据层面: 多模态感知融合与异构数据治理。**研发多源异构数据实时融合算法, 解决传感器数据(激光雷达、视觉、IMU)时空异步、精度差异问题, 实现厘米级定位与毫米级环境重建。在具身智能无人

船案例中，融合声呐、光学与气象数据，降低了海洋环境建模误差率，支撑复杂水域避障与路径规划。

**模型层面：动态物理规则建模与高保真仿真。**基于 MWORKS 平台构建动态物理模型库，支持流体力学、多体动力学等复杂规则仿真，实现海浪、风速等扰动场景的毫秒级响应。缩短了船舶领域仿真训练周期，提升极端天气（如台风）场景覆盖率，助力无人系统快速适配真实环境。

**算法层面：强化学习与多智能体协同优化。**提出“分层强化学习+博弈论”协同决策框架，解决多无人系统（如无人机集群）任务冲突与资源竞争问题。在智慧城市物流场景中，实现无人机协同配送，路径冲突率下降，能耗降低。

**算力层面：分布式计算与云边端协同优化。**开发轻量化模型压缩算法与边缘端推理引擎，算力需求降低，支持低功耗硬件（如嵌入式芯片）实时运行。无人船边缘计算延迟控制在 50ms 以内，单设备日均训练成本下降 40%。

**安全层面：虚实联动安全验证与风险预测。**构建“数字孪生-物理样机”双向校验体系，通过异常行为检测算法提前识别失控风险，并生成安全策略库。在工业巡检机器人部署中，事故率降低，保障高危场景下的安全运行。

**价值对齐：伦理约束与场景自适应规则嵌入。**设计基于规则引擎的伦理边界模型，动态约束智能体行为（如避让优先级、隐私保护），通过对抗训练提升合规性。在城市安防机器人中实现隐私数据脱敏，

提升公共区域任务执行合规率。

**具身智能：具身域框架与跨场景迁移能力。**提出“具身域”理论框架，统一物理空间（环境交互）、虚拟空间（仿真训练）与数据空间（算法迭代），支持智能体跨场景策略迁移。无人船训练模型迁移至其它场景，适配周期大为缩短，任务执行效率提升。

**软件工程：模块化平台架构与低代码开发。**基于 MWORKS 打造模块化仿真平台，提供多协议接口，支持用户通过拖拽组件与脚本自定义算法，开发效率提升。平台已服务多家企业客户，在能源、交通领域形成解决方案，代码复用率达 80%。

**应用必要性及行业实效来看，**一方面实现降本增效，虚拟训练替代大部分实体测试，硬件损耗减少，企业研发成本降低。另一方面实现技术壁垒突破，填补国内高精度动态仿真技术空白，突破国外仿真软件（如 ANSYS、Gazebo）在复杂规则建模与多智能体协同领域的垄断。

**产业化落地情况来看，**在船舶、物流、工业检测等领域形成标杆案例，推动国产智能无人系统全球竞争力提升。北理工通过跨学科技术融合与产学研协同，在数据治理、算法优化、安全验证等八大层面实现突破，构建“仿真-训练-部署”全链路能力，解决行业训练成本高、跨场景弱、安全风险大等痛点，为智能无人系统规模化落地提供技术底座。

## （2）模式创新情况

**解决方案创新：跨行业标准化智能训练平台。**推出基于 MWORKS

的“具身空间仿真平台”，集成高精度物理建模、多模态感知算法与强化学习框架，形成“仿真训练-数字孪生-实地部署”全链路解决方案，覆盖船舶、物流、工业检测等场景。为某海洋科技企业提供无人船集群训练服务，替代海上实测，研发周期缩短。

**服务模式创新：平台化“仿真即服务”（SimaaS）。**构建“技术订阅+定制开发”服务体系，面向中小型企业提供低代码开发接口与预训练模型库，支持按需调用算力资源与场景模块，降低 AI 技术使用门槛。与智慧城市服务商合作，为城市部署无人机物流网络，通过平台订阅模式实现算法动态更新，配送效率提升，客户技术投入减少。

**应用范围拓展：从单一场景到多域协同生态。**突破传统无人系统单一领域限制，依托“具身域”技术实现跨场景（水域、陆地、空域）策略迁移，衍生出水域巡检-陆地安防-空中物流协同服务。利用 AI 生成技术，自动创建设备故障等低概率高风险的训练场景，提升算法鲁棒性。江苏国家电网借助该功能，在虚拟环境中模拟地下输电管廊巡检场景，训练机器人快速定位与应急，故障处理效率提升。

**安全可信治理：全生命周期风险管控体系。**建立“伦理规则嵌入+动态安全评估”双轨制，提供从算法训练到物理部署的合规性验证服务，形成行业安全标准。应用成效：为工业机器人企业提供安全认证服务，事故率下降。

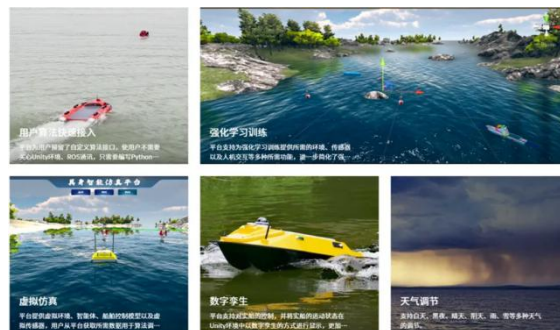
**生态共建模式：产学研用协同创新网络。**联合地方政府、头部企业共建智能无人系统联合实验室，形成“技术研发—场景验证—商业化推广”闭环生态。通过 MOHUB 模型共享平台，开放无人船、无人

机等案例源码及模型库，吸引 500+开发者参与共建，形成“产学研用”协同生态。例如，江苏科技大学基于平台二次开发了某海域航行算法，已应用于某型科考船虚拟测试。

通过“平台化服务+跨场景协同+安全治理”模式创新，打破传统 AI 应用碎片化瓶颈，构建覆盖技术、场景与生态的全价值链服务体系，在船舶、物流、城市治理等领域形成规模化落地，拉动行业增效，推动智能无人系统从单点突破向生态赋能升级。

### 8.应用实效：

以具身智能无人船为例，平台已实现从虚拟训练到近海测试的全流程验证。虚拟训练阶段在 Unity 引擎中构建高精度港口航道模型，模拟不同能见度、潮汐周期下的航行场景。通过强化学习算法，无人船在 10 万次迭代中掌握复杂障碍物规避策略，训练耗时仅 72 小时。数字孪生验证将训练模型部署至实体船体，实时对比虚拟与真实传感器的数据偏差（如激光雷达点云匹配误差 $<0.1m$ ），动态校准算法参数。实地部署表现：在某实测中，无人船自主导航成功率从传统方法的 78%提升至 98%，响应延迟降低至 200ms 以内。



来源：北京理工大学

图 52 具身智能无人船

极地水下航行器破冰探测仿真验证。海军工程大学针对极地水下

航行器破冰探测任务中冰层环境复杂、实地测试成本高、周期长的难题，基于 MWORKS 具身空间仿真平台构建“冰-水-航行器”多学科耦合模型，融合卫星遥感与实测数据，高精度还原冰层结构及洋流扰动环境，开发强化学习驱动的破冰策略优化算法，实现多物理场联合仿真与虚拟测试。该方案突破传统仿真精度不足的瓶颈，显著降低破冰能耗并提升任务成功率，大幅缩减研发成本与周期，支撑极地装备研制及冰区通信中继测试。相关技术纳入行业标准，核心模型通过开源平台共享，推动极地装备研发从“经验驱动”向“模型驱动”转型，为我国极地战略提供高效、低风险的“虚拟试验场”范式，兼具军事应用与科研推广价值。

**港口无人船集群协同导航与避碰仿真应用。**中船重工 711 所针对国际集装箱港口复杂环境下无人船自主导航效率低、多船避碰风险高的难题，基于 MWORKS 具身空间仿真平台构建港口数字孪生系统，融合航道三维建模与实时动态数据，模拟潮汐变化、极端天气及高密度船舶交互场景，通过多智能体强化学习算法优化集群协同避障策略。方案突破动态障碍物预测与高精度感知定位技术瓶颈，显著提升恶劣天气下的自主导航成功率与作业效率，有效降低碰撞风险。相关算法模型通过开源平台共享，形成“仿真训练—实船验证—国际复制”的全链条赋能模式，为全球智慧港口无人化升级提供标准化解决方案。

**地下输电管廊智能巡检机器人仿真应用。**国家电网江苏省公司针对跨江地下高压输电管廊环境复杂、人工巡检效率低且安全隐患大的问题，基于 MWORKS 具身空间仿真平台构建高精度三维虚拟环境，

模拟电缆过热、气体泄漏等多类故障场景，训练搭载多模态传感器的智能机器人。通过强化学习优化动态路径规划与故障定位能力，融合数字孪生技术实现机器人运动姿态实时校准，攻克复杂地形下步态稳定性与精准识别技术瓶颈。实际部署后，机器人巡检效率大幅提升，显著节省人力成本并成功预警潜在事故，相关成果参与制定行业技术标准，为城市地下空间智能化运维提供可复制的技术范式。

**变电站“空天一体”无人机智能巡检体系。**国家电网江苏省公司针对变电站人工巡检效率低、高危作业风险大的问题，依托数字电网平台构建“空天一体”立体巡检体系，通过高精度三维建模技术对变电站及输电设备建立数字孪生体，部署固定与移动无人机基站实现集群自主巡检。无人机搭载双模检测设备，结合边端 AI 模型实时识别设备缺陷与异常，显著提升巡检精度与效率，降低高危区域人工介入风险。方案实现变电站全业务无人化作业，大幅降低运维成本并减少非计划停电，成果覆盖多省市输电网络，参与制定多项行业标准，为新型电力系统打造智能化、无人化运维示范样板。

**船舶智能无人系统虚拟仿真教学体系。**江苏科技大学针对船舶与海洋工程教学实训高成本、高风险及技术滞后问题，构建“虚拟—实船—数字孪生”一体化教学平台，开设智能无人系统系列课程，涵盖动力学建模、感知融合与算法开发。基于高精度船舶模型及典型水域场景，支持多场景虚拟实验与实船验证，学生通过开源算法库自主开发导航模型并应用于企业项目。教学成果多次获国家级奖项，研发算法被行业采纳，校企联合输送专业人才，推动产学研融合创新。课程

资源包通过开源平台推广至多所高校，参与智能船舶教学标准制定，年均培养数百名复合型人才，助力船舶工业智能化升级与技术创新。

## 四、地球科学

### （一）基于 OneScience 框架的气象 AI 大模型并行训练示范案例

#### 1. 案例实施单位：

曙光信息产业股份有限公司

#### 2. 案例背景：

气象行业在数值预报模式主导下取得长足进展，但仍面临多重挑战。传统物理模型依赖高分辨率、复杂参数化与昂贵的计算资源，难以在有限预算下实现快速迭代。现有深度学习预报模型虽展现出初步优势，但受限于模型规模与计算框架，难以进一步提升对多尺度天气过程的刻画能力。如何在保持高分辨率物理一致性的基础上，实现高效、可扩展的深度预报模型，成为当前气象人工智能领域的核心技术瓶颈。

提升模型参数规模是增强预报泛化能力、提升空间-时间分辨率的有效路径。现有训练方案采用数据并行，在扩展模型时面临显存受限问题，通信开销随模型规模上升而剧增。本案例采用张量并行、流水线并行对模型进行规模扩展，以算子内维度切分为核心机制，提升通信与计算重叠效率，具备更优的可扩展性与资源适配性，是实现大规模高效训练的合理技术路径。

#### 3. 案例详述：

##### （1）总体思路

本项目在现有气象深度学习预测模型原有结构的基础上，引入张

量并行、流水线并行机制，对模型内部关键模块进行结构化划分与分布式部署，实现参数级别的多异构加速卡协同训练。对注意力机制中多头结构进行并行拆解与映射，按需将计算分配至多个计算单元，降低单卡显存占用并提升并行效率，有效突破了单卡显存瓶颈。该策略使模型在保持数值稳定性的同时具备更强的预测能力，为提升中长期天气预测精度提供了结构性支撑，为下一步开展大参数量预报模型的系统训练奠定了技术基础。

为突破现有气象深度学习模型在参数规模与表达能力方面的瓶颈，本项目在复现 Pangu-weather 预测模型的基础上，探索基于张量并行（Tensor Parallel, TP）、流水线并行（Pipeline Parallel, PP）的模型重构与分布式训练路径。Pangu-weather 原始模型采用数据并行（Distributed Data Parallel, DDP）方式进行训练，DDP 在每张加速卡上存放一份完整模型进行训练，存在显存利用受限、通信效率低下等问题，难以进一步扩展模型参数。为提升模型对多尺度气象过程的建模能力、增强中长期预报精度，本项目引入张量级别的算子划分机制，对 Pangu-weather 内参数量较大的 Transformer 模块进行结构级解耦，在多加速卡之间完成大规模模型的高效训练。同时，以原 DDP 训练方式为基线进行性能测试。

本方案旨在不改变 Pangu-weather 已有物理建模与气象变量结构的基础上，通过 TP、PP 等并行训练策略支持更大模型容量，从而提升模型精度、增强对极端天气事件的响应能力，并为未来 AI+NWP 混合模式打下技术基础。

## (2) 平台介绍

本项目部署于基于曙光新一代国产异构加速卡机器的高性能训练平台。新一代国产异构加速卡拥有较高的 AI 异构加速性能和较大的显存和带宽，适合构建原型级 AI 并行训练方案。新一代国产异构加速卡配套的软件栈及 AI 框架能够支持主流深度学习框架及相关的并行训练框架（如 PyTorch、Megatron-LM 等）。凭借良好的能效比与灵活部署能力，为本次 TP、PP 改造提供了理想的验证平台。

## (3) 技术方案

本项目采用 TP 训练模式，以 Transformer 内部的注意力机制（Self-Attention）与前馈网络（Feed Forward Network）为主要张量切分对象。在多异构加速卡架构下，本案例将每一层中包含的大矩阵张量按列或行进行划分，并分别部署至多个异构加速卡中同步计算。例如，对于一个含有 16 个注意力头的结构，在 TP 模式下，当异构加速数量为 4 时，按照每个异构加速卡计算 4 个注意力头的方式，实现头部级别的均匀分配，降低单卡显存峰值压力；在 PP 模式下，若 Transformer 包含 16 个模块且异构加速卡数量为 4 时，按照每个异构加速卡按照顺序计算 4 个模块的方式，实现流水线并行，突破同参数下显存受限的瓶颈。在实际部署中，所有张量切分/通信操作通过 NCCL 实现异步通信，实现计算与数据同步的最大重叠。

## (4) 关键指标参数

本案例所复现的 Pangu-weather 原始模型参数量为 64M，通过 TP 和 PP 并行训练策略分别提升为 2 倍、5 倍、10 倍参数量，并评估其

原始模型在 DDP 训练策略下，同训练轮次下达到的预测准确性或同预测准确性时所需的训练轮次。

#### 4.创新情况：

##### (1) 技术创新情况

本项目聚焦于气象大模型的结构扩展与并行化优化，以 Pangu-weather 为具体案例，首先进行模型复现，随后在模型结构设计、训练并行架构、算力利用、参数规模扩展等关键技术方向实现多项创新性突破。

在复现模型的基础上，本项目对 Pangu-weather 内含的经典 Transformer 结构进行了并行可重构设计，引入算子级别计算切分，通过 TP 或 PP 机制，将模型内部的注意力与前馈计算划分至多个异构加速卡运行，突破了单卡显存限制，显著提升模型对大气过程的建模能力与可扩展性。通过对经典结构的切分，为其他案例提供可借鉴的技术路线及切分思路。

在算法与计算架构层面，本项目摒弃传统 DDP 训练的全量梯度同步方式，采用参数级细粒度并行策略，结合异步通信优化与算子级内存调度，实现通信与计算的高效重叠，大幅降低通信瓶颈。

在算力利用与软件工程方面，本项目使用 Megatron-LM 构建张量并行训练流水线，完整实现了多异构加速卡训练与推理过程的自动图调度，具备良好的复用性与扩展性。本项目提供的 TP 和 PP 改造案例，支持同类模型结构的快速适配与迁移。

综上，本项目探索并实现了人工智能模型并行化的关键技术突破，

针对应用最广泛的 **Transformer** 结构在进行参数扩展时遇到的显存受限问题，提供 **TP** 和 **PP** 方式的解决方案，该方案具备向 **10B** 级参数气象模型扩展的技术基础，对推动 **AI+**气象智能计算融合具有显著示范意义。

## (2) 模式创新情况

本项目以“大模型并行化”为核心路径，探索了新型人工智能气象模型训练模式与部署机制，在产品形态、服务结构与应用边界方面形成了如下创新：

**轻部署产品形态：**项目通过张量并行技术，在显存受限平静下，将现有气象大模型进行参数量提升，降低了大模型研发门槛，具备在科研机构、高校实验室等轻型平台落地的可行性，推动气象 **AI** 从“中心化训练”向“广域分布式训练”转变。

**服务模式创新：**构建支持大模型扩展的气象预报开发框架，提供常见深度学习模块的并行划分策略模板，具备较强的迁移性。

**应用边界拓展：**通过对模型参数容量的扩大与结构细化，提升了模型在极端天气预测、气象驱动灾害模拟、智能农业与能源调度等场景中的精度与适应性，拓宽了 **AI** 气象模型的应用边界。

该模式为新一代 **AI** 气象模型提供了具备工程可实施性与产业可落地性的范式，具有良好的复制推广价值。

## 5.应用实效

本项目以气象大模型中的 **Pangu-weather** 大模型为例，在复现模型的基础上，针对传统 **Transformer** 的结构扩展与并行化训练，在训

练框架、模型结构、并行策略及工程实现等方面完成了关键技术准备，具备广泛的潜在应用场景。主要应用方向包括：气象部门的中长期天气与灾害性天气预报、能源行业中的风电/光伏功率预测、农业气象服务中的气候风险预警、交通与物流行业的天气调度辅助决策、高校与科研机构的 AI 气象建模研究等。该方案也可作为国家气象大模型体系构建的边缘试验平台或技术支撑模块，推动气象 AI 模型能力普及化。用户涉及气象业务单位（国家/省级气象中心）、能源企业的预测调度部门、涉农涉灾政务平台、智能交通平台、高校大气科学与人工智能相关科研团队，以及致力于 AI 模型工程化开发的开源社区成员与创业企业。由于该项目显著降低了大模型训练的硬件门槛，使得具备中等异构加速卡资源的机构（如高校实验室、地方气象台）也具备构建与部署大模型的能力，从而显著拓展了高质量模型服务的可及性。

本项目核心组件依托 OneScience 项目发布，除本案例包含的 DDP、TP、PP 模式下的 Pangu-weather 以外，OneScience 项目还包含常见气象海洋预测大模型，如 FourCastNet、GraphCast、CorrDiff 以及 NowCastNet 等。针对气象海洋领域，OneScience 项目包含常见的 ERA5、CMEMS 气象海洋数据集、数据加载器、张量并行训练框架、模型配置文件、推理测试脚本及基本文档等，遵循 Apache 2.0 开源协议。后续计划在模型优化、多变量耦合预测能力及数据接口方面进一步开放，鼓励更多用户参与共建，并逐步推进构建气象 AI 开源社区生态体系。

## (二) 伏羲系列气象大模型研发项目

### 1. 案例实施单位：

上海科学智能研究院、复旦大学

### 2. 案例背景：

数值模式是现代天气预报的基石，近 20 年，我国也建立起较为完整的数值天气预报（CMA）模式体系，然而与欧洲中期天气预报中心（ECMWF）等世界一流预报中心相比仍有一定差距。本项目通过多学科交叉研究，基于海量大气海洋数据驱动，研发“伏羲”系列气象大模型，不仅在科技自立自强、业务自主可控等方面具有重要意义，促进大气海洋预报领域的科技发展，还能提升灾害性天气预报能力，服务于防灾减灾需求和产业发展。

### 3. 案例详述：

#### (1) 总体思路

数值模式天气预报虽然符合物理规律、无需训练即可应用于任意区域，但是计算复杂度高，依赖大量超算资源，且无法通过海量的观测数据系统性提升预报精度。人工智能气象大模型带来了天气预报发展新机遇，在解决数值模式因向更复杂系统和更高分辨率发展带来的计算问题、天气气候无缝隙模拟、计算效率和迭代更新方面具有优势。伏羲团队自主研发基于人工智能的“伏羲”系列气象大模型，能够预报涵盖未来数小时、数天、甚至次季节时间尺度的天气过程，精准刻画从 30 米微气象空间尺度到全球数千公里行星尺度天气形势，预测精度超越传统最优物理模式，计算速度相对物理模型实现千倍提升。

## (2) 平台介绍

伏羲团队相继推出一系列基于人工智能的伏羲气象大模型，包括了中期天气预报大模型、极端天气预报大模型、次季节预报大模型、集合预报大模型、端到端大模型等。伏羲气象大模型能够预报涵盖未来数小时、数天、甚至次季节（60天）时间尺度的天气过程，精准刻画从30米微气象空间尺度到全球数千公里行星尺度天气形势，具备天气、气候一体化，确定性、概率预报协同能力。

## (3) 技术方案和关键指标参数

伏羲中期气象大模型（FuXi），提供全球未来15天逐小时的天气预测，该模型在国际权威机构ECMWF和谷歌WeatherBench2业务评测中整体精度领先，表现超过ECMWF数值模式IFS及其大模型AIFS，也优于谷歌DeepMind的大模型GraphCast；在中国气象局组织的人工智能天气预报大模型示范计划中，总分排名第一；该模型还获得2025年国际基础科学大会前沿科学奖。相关成果发表于npj Climate and Atmospheric Science。

通过融合扩散模型，并与伏羲模型进行端到端结合，开发出了针对极端天气预报优化的首个气象大模型FuXi-Extreme，极大提升极端天气事件预测的准确性和及时性，相关成果发表于SCIENCE CHINA Earth Sciences。

伏羲次季节气象大模型是首个基于人工智能的次季节预报大模型，将MJO（次季节最关键预报因子）的可预报天数由30天提升到了36天，相关工作发表于Nature Communications。在此基础上，联合

国家气候中心等单位发布了人工智能全球次季节-季节预测大模型系统“风顺”，该模型为中国气象局发布的三大模型之一，荣获“2024 年度中国十大气象科技进展”和“2024 年度中国气候研究重大进展”。

伏羲中短期集合预报大模型基于变分自编码器（VAE）构建扰动生成机制，引入连续排名概率分数（CRPS）与预测分布和目标分布间的 KL 散度作为联合优化目标，替代传统 VAE 中的 L1 与 KL 损失，有效提升了集合预报的统计一致性与预报技巧。系统评估表明，FuXi-ENS 在 360 个变量与预报时长组合中，有 98.1%的情况在 CRPS 指标上优于 ECMWF 集合预报系统。相关成果发表于 Science Advances。

研发首个预报精度超过 ECMWF 的端到端气象大模型，实现了“多源观测数据-同化初始场-预报”的全链路自主可控和系统性优化，能够完全摆脱初始场数据被国外“卡脖子”风险，相关工作发表于 Nature Communications。

#### （4）获奖、专利授权及获得投融资情况

“FuXi: A cascade machine learning forecasting system for 15-day global weather forecast”论文获 2025 国际基础科学大会（International Conference of Basic Science）科学前沿奖（2025 Frontiers of Science Award）。

“伏羲”系列气象大模型研发项目入选世界人工智能大会 2025 SAIL AWARD TOP 30 榜单。

“人工智能大模型首次实现全球气候预测”入选全国气象科教融

合创新联盟 2024 年度“中国十大气象科技进展”。

“人工智能全球次季节-季节气候预测系统的研制与应用”入选中国气候研究委员会 2024 年度气候研究重大进展。

“伏羲”模型在中国气象局组织的“2024—2025 年度人工智能天气预报模型示范计划”中，获得综合评估总分第一名。

“自主可控的远洋船舶气象导航”项目获 2025 年“数据要素×”大赛全国总决赛气象服务赛道二等奖、上海分赛气象服务赛道一等奖。

“伏羲气象大模型”入选中国气象服务协会 2024 年度首批优秀科技成果。

“精准气候预测智慧应用平台”获第十三届中国创新创业大赛河北雄安赛区暨第五届雄安新区“智绘未来”创新创业大赛团队组亚军。

“智能气候预测大模型团队”获第十三届中国创新创业大赛河北赛区暨第十二届河北省创新创业大赛行业赛一等奖。

#### 4. 创新情况：

##### (1) 技术创新情况

伏羲中短期气象大模型（FuXi）采用级联模型架构，首次在预测性能上达到了与国际权威的欧洲中期天气预报中心（ECMWF）的集合平均性能相当的水平。FuXi-Extreme 首次将扩散模型与伏羲模型端到端结合，极大提升极端天气事件预测能力。伏羲次季节大模型（FuXi-S2S）将 MJO 预测时间从 30 天延长到了 36 天，超过公认权威的 ECMWF 次季节模式。伏羲集合预报大模型（FuXi-ENS）利用变分自编码器（VAE）构建扰动模型，实现大气不确定性的量化估计。

端到端大模型（FuXi-Weather）是首个预报精度超过 ECMWF 的端到端气象大模型，完全摆脱对数值模式依赖。

## （2）模式创新情况

伏羲系列气象大模型通过人工智能与大气海洋科学的深度融合，成功构建了“技术研发-产业赋能-开源共享”的新范式。技术研发方面，以级联架构、扩散模型、变分自编码器等突破传统数值模式性能瓶颈，多项指标超越国际权威机构 ECMWF；产业赋能方面，伏羲模式实现了从技术突破到产业应用的跨越，通过精准气象预测，显著提升防灾减灾效能，推动新能源高效利用，驱动交通、航运、农业等行业领域的高质量发展；在开源共享方面，构建了协同创新的生态基础，实现了从专业工具向普惠服务的转变，创造了巨大的经济与社会价值。

## 5.应用实效

伏羲气象大模型作为人工智能与传统气象科学深度融合的代表性成果，成功构建了气象大数据与深度学习技术协同发展的新范式。项目实施以来，在提升气象预报精度、推动气象服务智能化、带动相关产业发展等方面取得了显著成效：

### （1）防灾减灾效能显著提升

伏羲模型具备高精度的中短期天气预报能力，能提前准确预报台风、暴雨等极端天气事件。该能力显著提升了各级政府和应急管理部的响应效率，为灾前预警、灾中调度和灾后评估提供了有力支撑。通过提前部署防范措施，有效减少了因灾害天气造成的人员伤亡和财产损失，切实增强了社会的防灾减灾能力。

## (2) 推动新能源高效利用

项目在风电和光伏等新能源领域应用广泛，通过对风速、太阳辐射等关键气象要素的精准预测，有助于评估可再生能源潜力，建模灾害事件所引发的发电能力波动，协助企业优化发电调度，提升能源利用效率。伏羲模型的高时空分辨率预报能力，为新能源站点的选址、系统设计和运维决策提供科学依据，助力实现电网调度智能化，有效推动了新能源规模化发展。

## (3) 赋能交通、航运与农业领域

在交通运输与航运管理方面，伏羲模型提供的高精度气象服务，帮助航空、航运及地面交通部门制定安全高效的运行方案，减少因恶劣天气导致的延误和事故，提升整体运行效率和安全保障水平。在农业方面，模型提供的精准气象预报帮助农户合理安排农事活动，规避气象风险，提升农作物产量与品质，促进农业稳产增收。

## (4) 开源开放，推动领域协同创新

2025年2月在hugging face开源FuXi-S2S数据集和FuXi-ENS数据集，单月下载量超24万次，总下载量超70万次。基于已业务运行的伏羲中期气象大模型，实现全球任意地区的未来15天逐小时预报的可视化。预报一天更新两次，包括温度、降水、湿度、风和光辐照五大关键气象要素。在魔搭社区开源伏羲超分模型、中期模型和污染物模型，开源的伏羲模型已被应用于学术研究。

### （三）基于深海基础模型的深海海山数字化智能系统

#### 1.案例实施单位：

之江实验室、中国大洋事务管理局、自然资源部第二海洋研究所、国家深海基地管理中心、厦门大学

#### 2.案例背景：

深海作为地球最大的生态系统，覆盖 65% 的地表面积。根据《联合国海洋法公约》及《生物多样性公约》定义，生境指生物或种群的自然生长场所及类型。开展深海生境研究，不仅能深化对深海矿区生态系统服务功能的认知，为科学划定开采与保护区域提供依据，更可避免重蹈近海“过度利用—认知突破—治理完善”的被动发展老路。当前深海探测长期依赖人工判读声呐与视频数据，存在效率低下（单次科考需 5000 余小时人工分析）、漏检率高（小型生物识别率不足 40%）、易对脆弱底栖生态造成二次破坏等问题。因此我国深海矿区勘探迫切需要智能化技术突破极端环境探测瓶颈，为国家资源战略提供支撑。

#### 3.行业痛点：

首先是数据稀缺性问题，极端深海环境导致高质量样本匮乏，传统 AI 模型训练所需标注数据不足，生物识别漏检率较高，加之缺乏统一的生物-地质联合标签体系，跨航次数据复用难度大；其次是边缘算力瓶颈，科考船载设备算力普遍偏低，难以支撑 AI 实时推理，导致数据不得不回传至陆地处理；最后是技术自主化不足的风险，深海探测核心技术长期依赖进口，不仅存在战略安全隐患，更制约着我国深海资源开发的主权行使。

## 4. 案例详述：

### (1) 总体思路

本项目聚焦解决深海探测中的三大核心瓶颈——效率低下、数据稀缺和算力不足。依托之江智算平台及 Qwen2.5-VL 多模态大模型架构，成功开发作为深海海山数字化智能系统核心智能底座的深海基础模型。该模型通过构建“多模态数据融合—云端大模型训练—轻量化部署”三级技术闭环，显著提升从数据采集到决策支持的响应速度，突破传统科考分析动辄数周的时效瓶颈，实现分钟级响应。

### (2) 平台介绍

本项目基于之江智算平台，开发了在线交互式智能化检测平台，旨在为生物与地质图像的智能识别提供高效解决方案。用户只需上传图片并提出相关问题，平台可借助先进 AI 算法快速识别图像内容，自动绘制识别框，准确标识生物种类或地质类型。平台支持生物图像和地质图像的多种格式上传，用户通过简单直观的交互进行提问，平台则依托深度学习技术开展图像分析和数据处理，提供即时反馈。通过之江智算平台的强大计算能力，平台可高效处理大规模数据，大幅提升用户工作效率。

### (3) 技术方案

**模型基于先进的深度学习技术和神经网络架构，以开源的多模态大模型为基础进行开发。**为支持多源异构数据有效融合，对模型输入层进行扩展，强化了数据处理能力。在此基础上，通过模型压缩和蒸馏技术，将模型推理速度提升 60%，显著降低计算资源占用，同时保

持模型的准确性与高效性。

**在深海生物物种检测方面**，系统整合多源异构公开数据集与经专家严格标注校准的自建高质量数据，有效弥补当前深海生态研究中高可信度数据资源的匮乏问题。基于此构建起覆盖 30 万余条标注样本的深海生物识别数据集，训练出具备强泛化能力与实用性的深度学习识别模型。该模型在真实科考应用中显著提高深海物种识别的准确率与处理效率，为深海生态监测与生物多样性研究提供关键技术支撑。

**在深海地质特征识别方面**，建立了多样化、丰富的地质数据集，涵盖多平台多元数据源，包括维基百科、谷歌搜图和百度搜图等。此外，利用 BenthicNet 等开源数据集，搜集整理出 11.7 万条图文数据。这些数据为 AI 大模型训练提供了扎实支撑，确保模型能准确识别和分类不同地质特征。

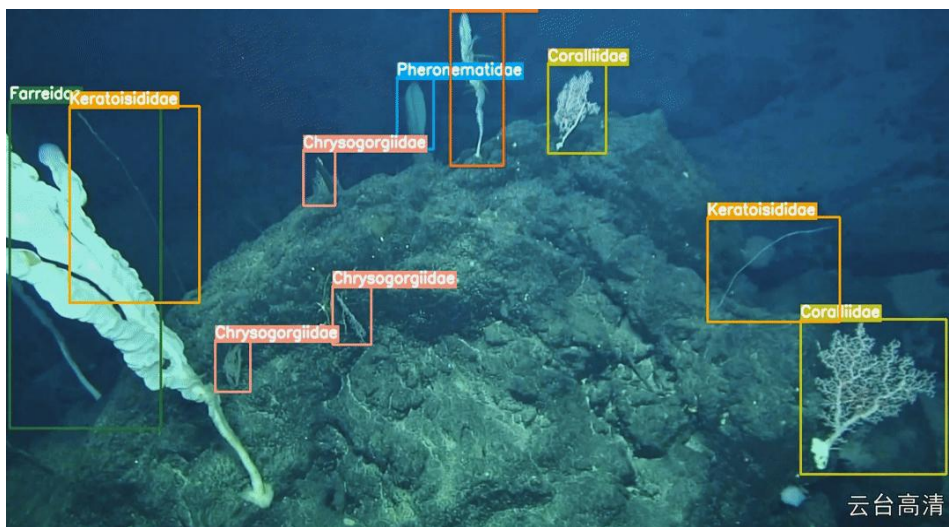
#### (4) 相关成果

四项发明专利聚焦多模态技术在大模型推理、深海生物时空轨迹分析、大模型融合的生态监测及训练数据集构建等方向，目前均处于申请阶段；一篇题为《基于多模态大模型的深海生境理解与应用》的论文已向《自然资源学报》投稿，正处于审稿流程中。



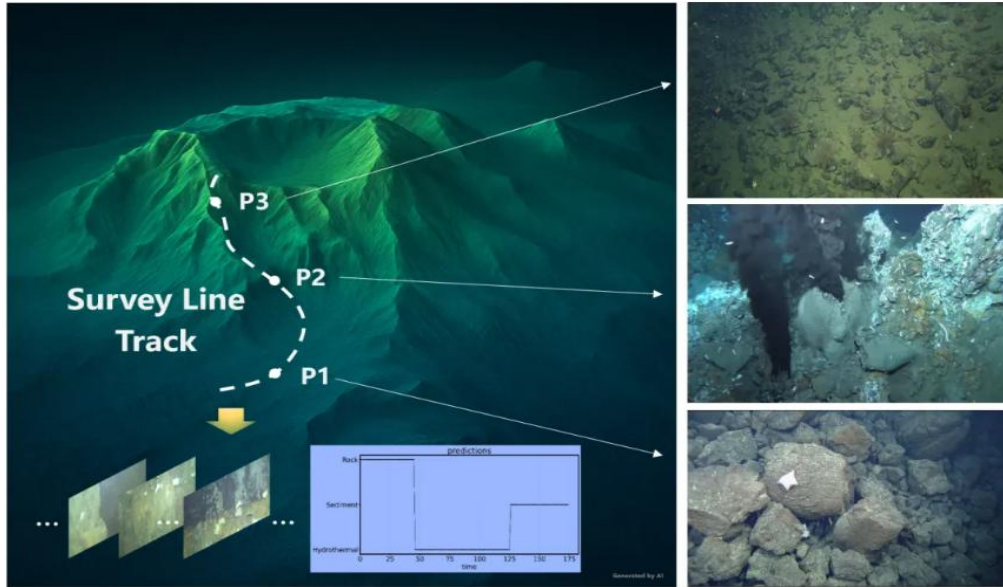
来源：之江实验室

图 53 深海海山数字化智能系统



来源：之江实验室

图 54 深海生物物种检测示例



来源：之江实验室

图 55 深海地质特征识别示例

## 5. 创新情况：

### (1) 技术创新情况

本项目在人工智能关键技术多个层面实现突破：**一是**数据融合与模型训练方面，采用“多模态数据融合—云端大模型训练—轻量化部署”闭环，有效解决深海探测效率低和数据稀缺问题。融合多平台数据，提升采集准确性和时效性。基于 Qwen2.5-VL 多模态大模型架构，优化输入层设计，成功支持多源异构数据高效融合，确保处理精准度；**二是**深度学习与算力优化方面，基于大模型深度开发优化，扩展输入层并增强处理能力，显著提升对复杂多模态数据的处理效率。实现模型推理速度提升 60%，大幅降低计算资源消耗。依托之江智算平台强大算力，结合轻量化部署、模型压缩与蒸馏技术，实现分钟级全流程响应，突破传统科考分析需数周的限制，有效节省算力的同时拓展应用场景；**三是**安全与保护方面，采用先进数据加密和多层次安全防护，

保障用户上传的生物与地质图像数据安全及隐私，符合相关行业安全标准。

## (2) 模式创新情况

本项目在服务模式和应用范围上实现多维度创新：**首先是交互式智能检测服务**，开发基于 AI 的在线交互式检测平台，用户可通过直观对话框上传图片并提问，平台依托深度学习技术即时分析并反馈结果。相比传统人工干预或批量处理，该模式极大简化操作流程，提升用户体验与操作便捷性，推动 AI 技术更广泛服务于生物与地质领域的实际需求；**其次是精准识别与应用拓展**，通过对生物与地质图像的精准识别分析，成功应用于深海探测、地质勘察、生物多样性研究等多个领域。平台能准确识别生物种类、描述地质类型，为科研、环保、资源勘探等工作提供高效辅助工具，尤其在深海实时分析识别海底地质特征方面提供有力决策支持，在高精度、快响应场景表现卓越；**进一步是智能化解决方案**，提供智能化自动标注、分类与识别服务，显著提升生物图像识别和地质图像分析效率。将深海探测等传统需数周的分析过程缩短至分钟级，对资源开发、环境保护等领域具有重大现实应用价值；**最后实现跨领域融合与生态构建**，融合深海探测与生物、地质图像分析技术，形成创新技术产品和服务模式，开辟了 AI 应用新领域。为相关行业提供全面的智能化服务，促进产业链深度协同，构建可操作、可拓展、可持续发展的产业生态系统。平台的开放性和智能化特性具备广阔应用前景，推动相关领域技术与产业升级。

## 6.应用实效

本项目开发的交互式深海生物物种检测和地质特征识别检测平台已经成功应用于多个领域，特别是在生物多样性研究与海洋资源勘探方面，取得了显著的社会效益。平台目前已面向大众开放公测，并与自然资源部第二海洋研究所展开深度合作，成为其生物识别项目的重要技术支撑。

**在实际应用中**，自然资源部第二海洋研究所的研究人员通过深海海山数字化智能系统进行深海生物的智能识别，极大提升了研究效率。平台支持多种图像格式上传，用户仅需提供深海图像，系统便能迅速分析图像，自动识别深海生物种类及定位框，减少了人工标注和分析时间。通过深度学习算法，平台不仅能够快速准确地识别出多样生物种类，还能绘制精准识别框，保障识别结果的高准确率。深海海山数字化智能系统的投入使用，使得自然资源部第二海洋研究所的研究人员进行大规模生物调查时，节省了大量时间和人力成本。依托平台的高效数据处理能力，研究人员可在短时间内完成深海生物样本的处理和数据分析，加快了研究进度，提升了科研成果的转化效率。本系统应用价值不仅局限于科研领域，还广泛适用于环境保护、海洋资源勘探等行业，具有广阔推广潜力。

**社会效益方面**，平台的智能化识别技术推动了海洋生物资源的可持续开发，为相关环保政策提供数据支持。同时，平台的高效运作提升了海洋行业的科技水平，有助于促进区域经济的可持续发展，成为业内科技创新的重要标杆。2025年世界海洋日及联合国海洋大会期间，深海海山数字化智能系统作为全球首个深海领域数字化公共科技

产品正式开放公测。这一突破不仅填补了人工智能技术在深海应用的空白，更标志着深海发展迈入数智化时代。本成果在法国尼斯举办的联合国海洋大会“启智海洋”边会上亮相后，迅速引发国际社会广泛关注，《人民日报》《新华网》等国内主流媒体也对其进行了报道。基于开放共享原则，本平台即将开源开放，支持第三方开发者及科研机构开展技术二次开发，进一步拓展其在深海探测、生物多样性、生态保护等领域的创新应用潜力。

## （四）大模型驱动的灾害防御智能体

### 1.案例实施单位：

土豆数据科技集团有限公司、中国地质环境监测院

### 2.案例背景：

地质灾害防治是国家防灾减灾体系的重要组成部分，但传统管理模式依赖人工巡查和经验判断，存在数据割裂、响应滞后、决策低效等问题。随着极端天气增多，地灾风险加剧，亟需智能化手段提升防控能力。在自然资源部指导下，中国地质环境监测院率先启动“地灾风险双控智能化”改革，依托大模型、GIS、物联网等技术，构建全国首个“地灾双控智能体”，实现从“人防”到“智防”的跨越式升级。

### 3.行业痛点

传统地灾防控面临三大痛点：**一是数据沉睡**：地质勘探、监测数据分散在纸质报表或孤立系统中，无法联动分析；**二是响应迟滞**：预警依赖人工层层传达，平均耗时6小时以上，错失避险黄金期；**三是决策粗放**：风险研判依赖个人经验，缺乏量化依据，易出现“过度防控”或“防控漏洞”。以某次滑坡灾害为例，因未能及时关联气象数据和隐患点历史位移，导致XX人伤亡。智能体通过大模型实时关联多源数据，可从根本上解决这些问题。

### 4.案例详述：

#### （1）总体思路

地灾双控智能体是一款面向区县级地质灾害防治的AI决策中枢系统，以“隐患点+风险区”双控理念为核心，构建了集“智能预警—预

案生成—防御响应—灾害评估—科普宣传”于一体的全链路智能防控平台。产品深度融合大模型技术和知识图谱技术，实现了从传统“人防”到现代“技防”的转型升级，是地质灾害防治领域的革命性解决方案。



土豆数据科技集团有限公司

图 56 地灾双控智能体

## (2) 平台介绍

具备六大核心能力：

**智能风险识别。**隐患点一张图：聚合展示全县隐患点分布，支持按灾害类型、险情等级等多维度筛选，如查询“XX镇威胁50人以上的滑坡隐患点”，地图自动高亮并生成统计报表。风险区动态评估：结合降雨量、地质构造等数据，实时计算风险区等级变化。

**智能监测预警。**多源数据融合：接入北斗位移监测、雨量站、InSAR遥感等数据，实现分钟级风险感知。

**智能应急响应。**一键启动会商：自动关联隐患点数据、气象预报、应急预案，生成处置建议。责任人联动：地图显示各级责任人位置与

联系方式，支持任务派发与跟踪。

**智能辅助决策。**自然语言交互：通过语音或文本提问（如“XX项目是否符合避让要求？”），大模型自动解析并调取相关法规与案例。



土豆数据科技集团有限公司

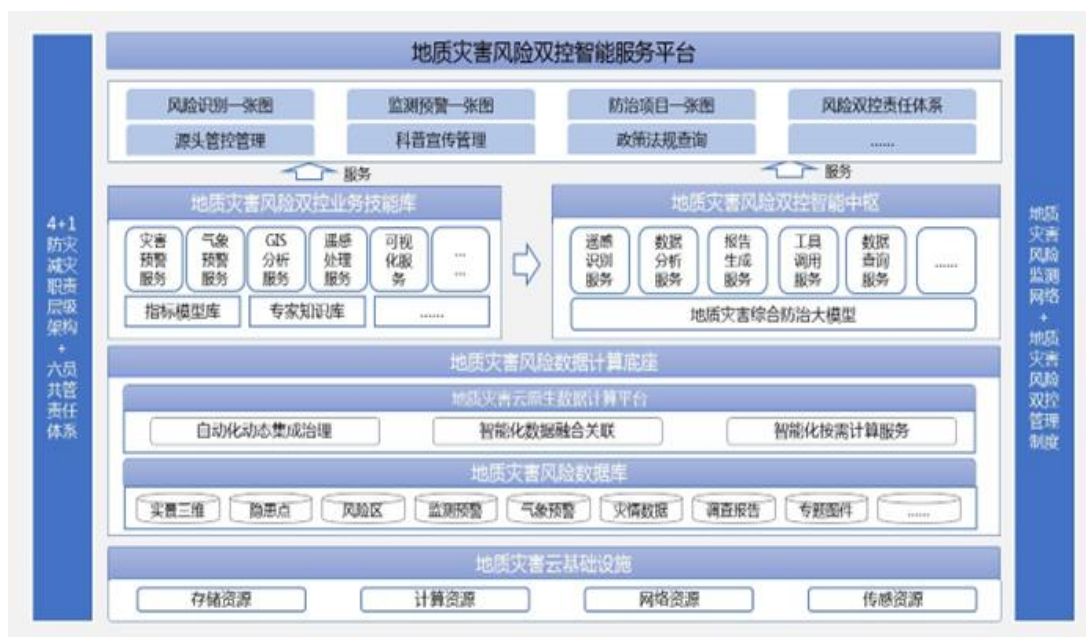
图 57 地灾双控智能体六大核心能力

### (3) 技术方案

**精准监测与预警：**一是实时处理分析大量传感器数据。二是综合多因素判断，提高预警准确性。

**高效评估与管理：**一是全面分析风险，识别高风险区域和隐患点。二是优化防治策略，分析成本效益。

**智能决策支持：**一是模拟预测灾害，规划应对措施。二是促进快速响应与协同决策。三是灵活查询与应用：基于自然语言交互，开放式数据查询。四是灵活化的数据可视化展示。



土豆数据科技集团有限公司

图 58 地灾双控智能体平台技术服务

## 5. 创新情况：

### (1) 技术创新情况

**自然语言大模型：**将自然语言大模型与地灾行业紧密结合，融合卫星影像、监测数据与专业文献，实现"语音提问→地图定位→报告生成"全流程智能化，风险研判准确率提升 40%。

**动态知识图谱：**构建包含 8 万+隐患点、200+地质参数的时空知识库，支持"降雨-位移-灾情"多维度关联分析，数据利用率从 30%提升至 90%。

### (2) 模式创新情况

**业务效率提升：**一是传统人工分析需 2-3 天的工作系统。二是可在 5 分钟内完成并生成报告。三是基层人员工作效率提升 80%。

**决策质量优化：**一是融合多源数据的智能分析。二是风险研判准确率提升 40%。

**管理模式革新：**一是实现"数据驱动"的精准防控。二是构建"县-乡-村"三级协同机制。三是应急响应时间缩短 70%。

## 6.应用实效

### (1) 试点成果

在龙川县试点县成功应用；形成可复制的技术方案，正在 6 个省推广应用，并获评行业标杆案例。

### (2) 经济社会效益

依托遥感大模型核心技术，智能体对卫星影像进行全自动解析，智能识别地表新建构筑物、新增线性地物等变化特征，精准提取疑似违法图斑。针对每个识别结果，深度运用空间分析引擎，自动套合耕地保护红线、生态保护区等管控数据，进行合规性智能研判。借助分布式计算架构，系统可实现每日数万平方公里影像的实时处理，构建 7×24 小时无人值守监测体系。无论是年度国土变更调查，还是日常执法巡查，均能实现分钟级响应，大幅提升违法用地发现效率，将传统人工巡查效率提升 10 倍以上。

### (3) 工作流程优化

智能体通过洛书多模态大模型，自动对比图斑举证信息，智能化输出审核结论，大幅提升审核效率；智能体审核结果的准确率达到 98%以上，完全满足变更调查业务规范要求，确保了审核结果的可用性，大幅度减少了人工核查工作量。

同时优化了原有监测、分析、报告环节割裂，流程不连贯的现象；打造了统一的信息共享工作平台，快速获取全面、准确的违法态势分

析；并自动胜场可视化、智能化的决策支持工具，全面替代了原有专业门槛高、数据凌乱、决策链条长而慢的低效率的人工时代，将传统人工巡查效率提升 10 倍以上。

## （五）一站式遥感变化监测智能体

### 1.案例实施单位：

土豆数据科技集团有限公司、莫干山地信实验室

### 2.案例背景：

为推进遥感 AI 技术在国土执法领域的深度应用，莫干山地信实验室联合土豆数据科技集团有限公司，共同研发大模型驱动的遥感变化检测智能体。项目依托实验室的“地理信息+AI”技术积累，重点解决城乡违法用地监测难题。通过融合多源卫星影像和地理信息数据，构建了智能化监管平台，实现从“人防”到“技防”的执法模式转型。

### 3.案例概述：

本智能体创新性地集成了莫干山实验室的多时相遥感分析算法与空间知识图谱技术，具备三大突破：一是首创“变化检测—合规分析—执法建议”全链路 AI 处理，处理效率达 2000 平方公里/天；二是支持耕地“非粮化”等复杂场景识别，准确率提升至 91%；三是研发自然语言交互引擎，实现“查询莫干山占用基本农田的违建”等智能问答。

### 4.行业痛点：

一是**技术门槛高**，现有算法难以适配多变的地物特征，城乡结合部识别准确率不足 65%；二是**业务融合浅**，多数系统仅提供基础检测，缺乏与执法流程的深度结合；三是**数据壁垒强**，遥感数据与国土业务数据割裂，分析结果难以直接指导执法；四是**更新迭代慢**，传统模型需 3-6 个月重新训练，无法适应新出现的违建类型。

### 5.案例详述：

## (1) 总体思路

本项目由莫干山地信实验室主导研发，旨在打造新一代“智能感知—自主分析—决策输出”的遥感监测技术体系。系统基于多模态大模型+地理空间认知计算的创新架构，深度融合遥感影像解译与地信专业知识，实现从数据输入到业务决策的端到端智能化处理。



土豆数据科技集团有限公司

图 59 遥感智能监测技术体系

## (2) 平台介绍

融合遥感大模型与数字化能力的智能工具，以多维度创新功能构建高效运作体系：**在监测识别环节**，凭借亚米级变化识别精度，结合遥感大模型精准识别违规图斑，突破了传统识别的准确性局限；**在分析管理层面**，通过任务智能创建+全流程数字化留痕，搭建起闭环式的高效管理链路；**在决策支持维度**，依托跨年份数据对比、影像数据智能推荐，为任务建设提供精准的智能辅助；**在查询应用端**，以自然语言交互的开放式查询+多维度数据筛选，实现了更灵活便捷的使用体验。

### (3) 技术方案

以多模态融合为核心，构建了高效且兼具创新性的分层技术架构：  
**一是多模态输入层**，实现三类数据的全面接入——卫星影像模态处理 0.5-10m 分辨率光学/SAR 数据，矢量数据模态融合国土调查、生态红线等空间数据，文本规则模态内嵌自然资源领域 2000+条政策条款；  
**二是跨模态特征融合环节**，采用双流 Transformer 架构并创新性引入空间注意力机制，使特征融合精度较单模态基线提升 32%；  
**三是通过自适应训练框架**，将模型迭代周期从传统 3 个月压缩至 1 周，同时支持增量学习新地物类型，兼顾了训练效率与扩展灵活性。

## 6.创新情况：

### (1) 技术创新情况

**多模态数据融合方面**，项目突破传统遥感分析的单模态局限，创新性地构建了“影像—矢量—文本”三模态融合框架。通过研发空间对齐注意力机制，实现了 0.5 米分辨率卫星影像与国土调查数据的像素级匹配，解决了异构空间数据融合的行业难题，使变化检测精度提升至 93.1%（较传统方法提升 28%）。

**工程化实现方面**，开发了轻量化模型部署方案，通过动态量化技术和自适应计算框架，使模型在 8 卡 GPU 服务器上可实现 2000 平方公里/天的处理能力，同时支持 10 m<sup>2</sup>级小目标检测，较行业标准提升 5 倍灵敏度。创新设计的增量学习系统，可将模型迭代周期从传统 3 个月缩短至 1 周，大幅提升业务适应性。

### (2) 模式创新情况

**产品形态方面**，开创了“AI 智能体+业务中台”的新模式。将大模型能力封装为可插拔的微服务模块，支持与现有国土调查云、执法监管系统无缝对接。在德清县试点中，该模式使系统部署时间缩短 80%，用户无需改变原有工作流程即可享受 AI 赋能。**服务模式方面**，建立“持续学习即服务”（CLaaS）新机制。通过实时收集各地执法人员的反馈标注，形成模型优化闭环。**应用拓展方面**，突破传统卫片执法范畴，成功将技术复用于生态保护、应急监测等领域。**治理体系创新方面**，研发了行业首个“人机协同决策看板”。通过可视化展示 AI 判断依据、置信度及人工修正记录，构建了透明可信的决策机制。

## 7.应用实效

### （1）技术验证成果

目前系统已在浙江省莫干山地信实验室完成技术验证环境部署，成果如下：在技术验证环节，算法性能达到预期指标——实验室测试数据集上实现 91.2%的变化检测准确率、8 m<sup>2</sup>的最小识别面积，模拟环境下处理速度可达 1500 平方公里/天。德清县部分区域的概念验证（POC），成功识别测试区域内的模拟违建目标，实现与现有地理信息平台的初步对接，还完成了自然语言交互功能的原型测试。

### （2）应用价值

一方面提升效率，可缩减人工判读时间 70%以上，大幅降低人力成本与时间消耗；报告生成效率提升 10-15 倍，显著加快工作流程的周转速度。另一方面改善执法精度，复杂场景识别准确率达 90%，提升了复杂环境下目标识别的可靠性；小目标检出率突破 80%，强化了

对微小违规目标的捕捉能力。

### (3) 经济社会效益

减少人力成本可达 70%，大大降低人力成本。

### (4) 标准产出

目前正与莫干山地信实验室共建行业标准，并拟于年底左右进行开源发布。

## （六）时空可信验证智能体

### 1.案例实施单位：

土豆数据科技集团有限公司、四川省国土科学技术研究院

### 2.案例背景：

四川省位于中国西南，涵盖多个地形类型：高山、盆地、平原、丘陵等，地形复杂多样。全省共 183 个区县，年度变更核查图斑量达到 300 万，核查任务庞大，传统的人工核查方法已无法满足核查需求，为应对这一挑战，我们在四川省开展了基于智能审核的变更核查项目，旨在通过融合多模态大模型与视觉大模型，结合智能化审核流程，显著提升核查效率、节约人力成本并提高核查质量。

### 3.案例概述：

本案例展示了举证核查智能体在变更核查中的创新应用，智能体基于多模态大模型的视觉识别与智能推理技术，利用外业举证照片、专题数据和技术规程进行综合分析，自动化执行核查图斑的精准审核。通过智能化审核流程，系统实现了 50%以上人力节约，且核查结果的准确率达 95%以上。有效解决了传统核查业务依赖人工存在的耗时久、质量差、成本高等难题。

### 4.行业痛点：

国土变更核查工作通常集中在年底，面临时间紧任务重的问题；传统的变更核查依赖于人工逐图斑、逐县、逐市、逐省分析，效率低且容易出错，且对核查单位技术人员的专业水平、工作态度依赖较大，容易导致错判、漏判，影响核查质量；此外，当前软件智能化水平低，

缺乏自动化、智能化审核技术，造成人力资源、设备资源高峰期投放大，非高峰时段设备和人员闲置浪费。

## 5.案例详述：

### (1) 总体思路

举证核查智能体是基于洛书多模态大模型的智能核查平台，旨在利用图像识别、深度学习等前沿技术，对变更核查数据进行智能识别与分析，结合年度变更核查技术规程，自动生成审核结论，并且具备灵活的个性化处理能力，能够针对不同类型地类、作物、种植属性及作物的不同生长阶段进行精准判定。智能体突破了传统人工审核的瓶颈。提升了核查效率，提高了核查质量。

### (2) 平台介绍

该智能体基于云平台架构设计，支持大规模数据的快速处理，平台能够接入国土调查云平台的数据，实时审核，审核结果迅速回传。智能体支持广泛覆盖的地类识别，能够精准识别耕地、园地、林地等10种主要地类，覆盖90%以上变更调查图斑；智能体不仅能够实现众多作物的精准识别，如水稻、玉米、小麦、高粱等粮食作物，蔬菜、烟叶、野芋、药材等非粮食作物，同时能够精准区分旱生作物和水生作物；针对同一作物从幼苗期到成熟期及收割后的各个生长阶段，实现精准识别，确保对举证信息的准确判断。

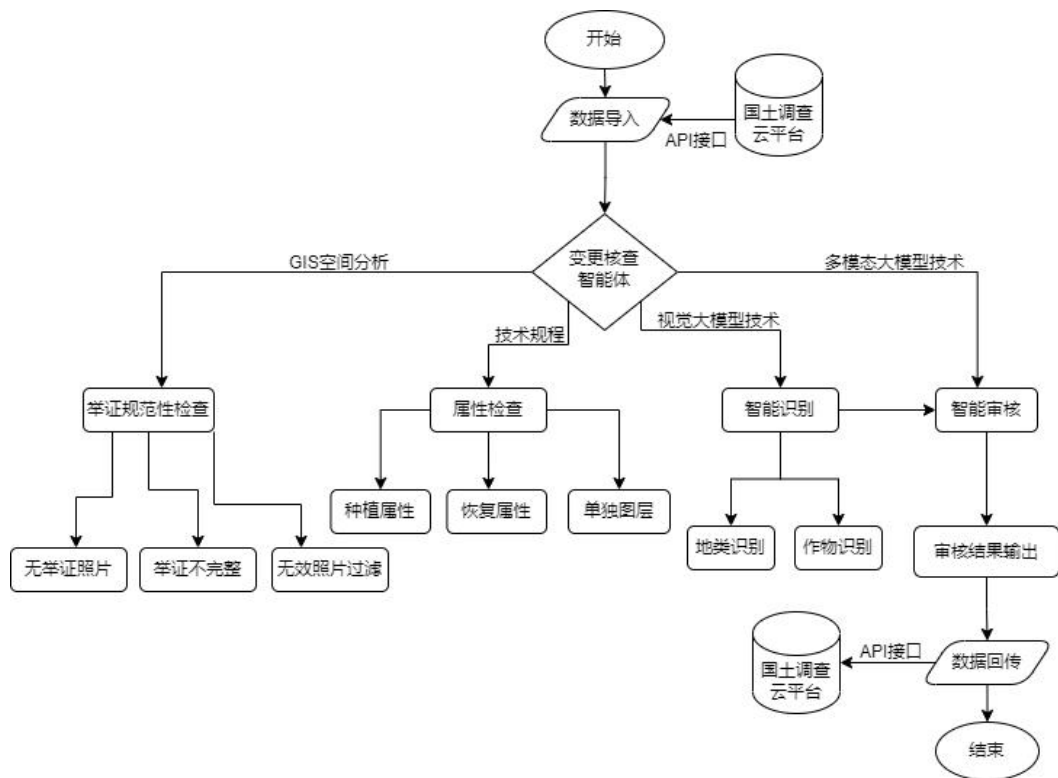
### (3) 技术方案

智能体依托于强大的云计算平台，采用全链路智能中枢架构，实现从举证数据上传——智能审核——结论回传一步到位，彻底打破

「县→市→省」的传统逐级核查体系，消除冗余流程，显著提升核查效率与响应速度；通过构建全自动智能审核流水线，支持夜间任务队列智能处理，确保审核工作全天候不间断运行，大幅提升业务处理时效性与连续性；弹性算力资源池，有效降低硬件采购成本；同时支持公有云、混合云等多种部署方式，在保障数据处理灵活性的同时，确保敏感数据在内网环境下的安全处理；基于技术规程的审核规则与推理大模型的深度思考能力，通过双重校验机制，对调查举证数据的真实性、合规性进行严格把控，确保输出结果真实可靠、准确无误。

#### （4）关键指标参数

在四川省 30 万图斑的测试中，智能体动态协调 200 台单卡 TRTX 4090 服务器共同参与并行计算，仅用 8 小时便完成了全部审核任务，充分展示了平台的高效处理和实时审核能力。智能体在平均人工替代率和准确率方面也展现了显著优势，尤其在处理复杂地理环境和大规模数据时表现尤为突出。具体来说，智能体的平均人工替代率达到了 50% 以上，且准确率超过 95%，确保了核查结论的精确性与可靠性，体现了深度学习模型的强大能力。



来源：土豆数据科技集团有限公司

图 60 年度变更核查技术实现流程图

## 6.创新情况：

### (1) 技术创新情况

举证核查智能体能深度融合多源异构数据（举证照片、调查监测数据、地理信息矢量数据）实现多模态数据融合、综合分析，从而提供全面精准的变更核查结果；基于深度学习和机器学习，能够自动提取地物特征进行精准分类识别，这一创新不仅提高了对地类、作物识别的精准度，还优化了传统人工判定过程中的主观偏差，尤其在复杂地理特征或大规模任务监测中，智能体能够适应不同的变化模式，显著减少人工操作。通过大数据分析与人工智能算法的深度融合，智能体实现了多维度的自动化决策机制，有效提升核查结果的准确性；在算力层面，采用云计算和分布式计算架构，支持海量数据快速处理与

实时分析，保证了变更核查的实时性，同时避免了单一计算节点的性能瓶颈；在安全方面，举证核查智能体支持公有云、私有云、混合云等多种部署方式，在保障数据处理灵活性的同时，确保敏感数据在内网环境下的安全流转与处理；举证核查智能体具备轻量化、可移植的特性，结合模块化设计，显著提升了系统的可扩展性和可维护性，能够灵活应对不断变化的行业需求及技术挑战；这一创新为政府及行业提供了革命性的解决方案，有利推动行业数字化转型。

## （2）模式创新情况

基于洛书多模态大模型的智能核查平台，突破了传统人工审核的瓶颈，实现了全链路自动化审核流程，通过图像识别、深度学习等技术，自动化完成海量数据的处理，生成精准的审核结论，尤其是对不同地类、作物及其生长阶段的准识别与判定。智能体采用了云平台架构与弹性算力资源池，支持大规模数据的快速处理。这一模式使得各级政府能够实时接入国土调查云平台的数据，自动化完成数据审核、回传。平台还支持夜间任务队列智能处理，确保全天候不间断的智能审核，提升了审核的连续性和时效性。此外，平台的灵活部署方式（公有云、混合云等）使得数据处理既具备灵活性，又确保了敏感数据的安全处理。智能体通过构建双重校验机制，结合技术规程和推理大模型的深度分析能力，对数据的真实性、合规性进行严格把控，确保了审核结论的准确性与可靠性。在实际应用中，核查智能体已经实现了显著的行业效益。以四川省 30 万图斑测试为例，平台在 8 小时内完成了全部审核任务，其人工替代率达到 50% 以上，审核准确率超过

95%。这一实际应用成果，标志着核查智能体在土地变更核查领域的技术创新已取得可观的实效，进一步促进了政府对智能化技术的应用和推广，如自然资源调查与监测领域，林地和草监测，永久基本农田核查等。

## 7.应用实效：

举证核查智能体广泛应用于自然资源调查监测各类业务中，涵盖国土变更核查、森林资源监测、农田保护等多个领域，该智能体不仅提升了各类资源的监管效率和精准度，还为自然资源的可持续管理与保护提供强大技术支持。通过精确的数据分析与智能决策，它促进了资源管理的精细化与智能化。

举证核查智能体的目标用户包括政府部门、相关企业集科研机构。目前，主要用户群体为地方政府及其下属的自然资源管理机构、测绘局、规划院等单位，主要用于年度变更核查和日常变更核查。该智能体适用于需要高频率、全覆盖且大规模资源监控的场景，满足了从日常管理到年度变更的多元化需求。

与传统的人工核查方式相比，举证核查智能体极大的提高了工作效率，显著降低了运营成本，传统人工核查不仅效率低下、成本高昂，而且容易存在遗漏和误差。通过融合大数据技术与人工智能（AI）算法，举证核查智能体能够处理海量核查数据，且依托多模态大模型，显著提高了核查过程的精确性与智能化水平。无论是复杂地形的资源监测，还是数据量庞大的土地利用变更分析，智能体都能快速、准确地完成任务，确保了数据的可靠性与精确性。

核查变更智能体的应用，不仅降低了政府部门在广泛区域内开展资源监测与管理的成本，还有效减少了资源浪费。通过实时获取和分析核查数据，政府能够即时发现问题并做出相应调整，优化政策与监管措施，从而提升社会治理的科学性和响应效率。此外，智能体的应用还为资源管理决策提供了更加全面和精细的数据支持，帮助各级政府制定更加精准和有针对性的资源保护政策。

目前，举证核查智能体已经获得多个省级自然资源管理机构的认可，并与多个省市变更核查主管单位达成了合作意向。通过不断的技术创新和应用实践，智能体可在全国范围内的推广和应用，逐步成为推动自然资源智能化监管的重要工具，展现了其在实际应用中的巨大潜力和市场前景。

## （七）滑坡隐患关联要素智能识别系统

### 1.案例实施单位：

中国自然资源航空物探遥感中心

### 2.案例背景：

我国西部山区大型滑坡灾害易发程度高，孕灾机制复杂，隐蔽性强，严重制约和威胁新型城镇化以及重大基础设施的建设与安全运行。近年来，以“综合遥感普查识别+地面核查”的工作模式逐渐成为重大崩滑隐患早期识别，动态掌控风险隐患“底数”的有效手段。然而，面对广域尺度动态遥感识别的需求以及日益积累的海量遥感数据，如何提高滑坡隐患识别技术的智能化和工程化水平成为当前面临的主要问题。

### 3.案例概述：

中国自然资源航空物探遥感中心依托国家重点研发计划项目《广域重大地质灾害隐患综合遥感识别技术研发》下属课题《时空信息协同的崩滑隐患关联要素智能识别技术研发》，建立了滑坡隐患要素智能识别技术框架，联合多个知名高校和企业研发了地质灾害隐患关联要素智能识别系统。该系统整合了多源遥感数据，结合前沿 AI 技术，实现了广域尺度滑坡隐患关联要素的遥感智能识别，有效改善了以人工解译为主的遥感调查方式，提高了广域动态识别的准确性与时效性，为我国西部地区地质灾害隐患遥感调查提供了有效支撑。

### 4.行业痛点：

滑坡隐患遥感智能识别研究尚处于起步阶段，缺少工程化应用实

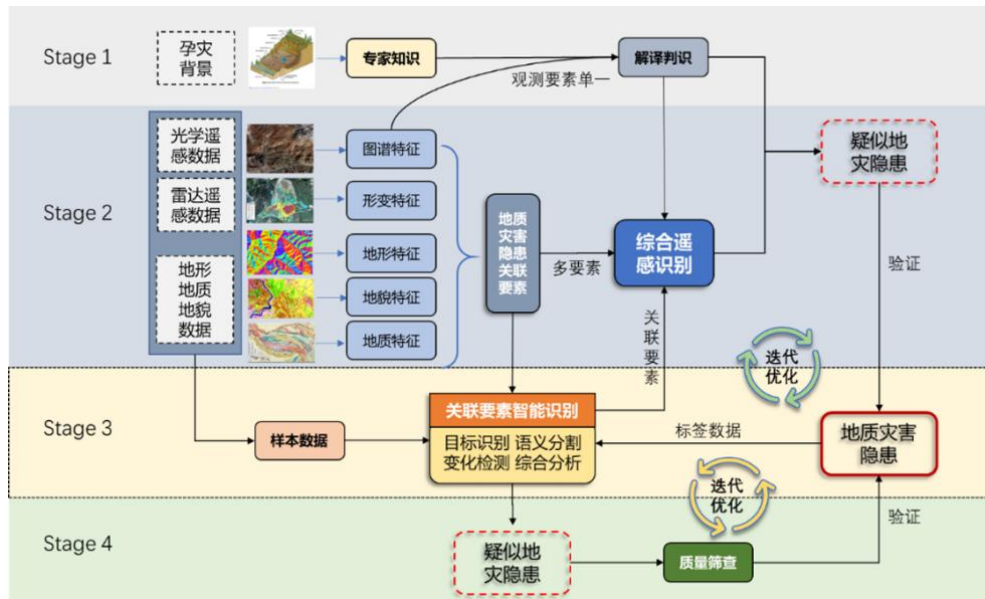
践。除了面临如何理清人工智能在隐患识别任务中的技术定位、找准识别对象、与传统业务实现技术优势互补的难题外，还面临着多源异构遥感数据分析、智能模型高效训练与优化、业务化应用部署等及技术难题。面向地质灾害隐患识别应用场景，研发融合前沿 AI 技术的新一代 AI4S 遥感综合分析平台，构建符合业务场景需求的知识库、模型库、样本库，已成为推进识别业务工程化应用的核心挑战。

### 5.项目获奖、专利授权及获得投融资情况：

系统研发过程中，在滑坡图谱特征智能识别和变化检测，滑坡隐患关联的广域 InSAR 形变区自动识别，智能模型可泛化研究与工程化应用策略研究等方面取得技术突破，发表 SCI 论文 12 篇，EI 论文 2 篇，授权发明专利 12 项，授权软件著作权 11 项，培养研究生 11 名，自然资源部青年科技人才 1 人。系统相关成果入选 2023 爱分析·人工智能最佳实践案例，2024“算力中关村”算力技术创新与应用服务优秀案例，2025“智算大会”算力产业十大创新实践案例。

### 6.案例详述：

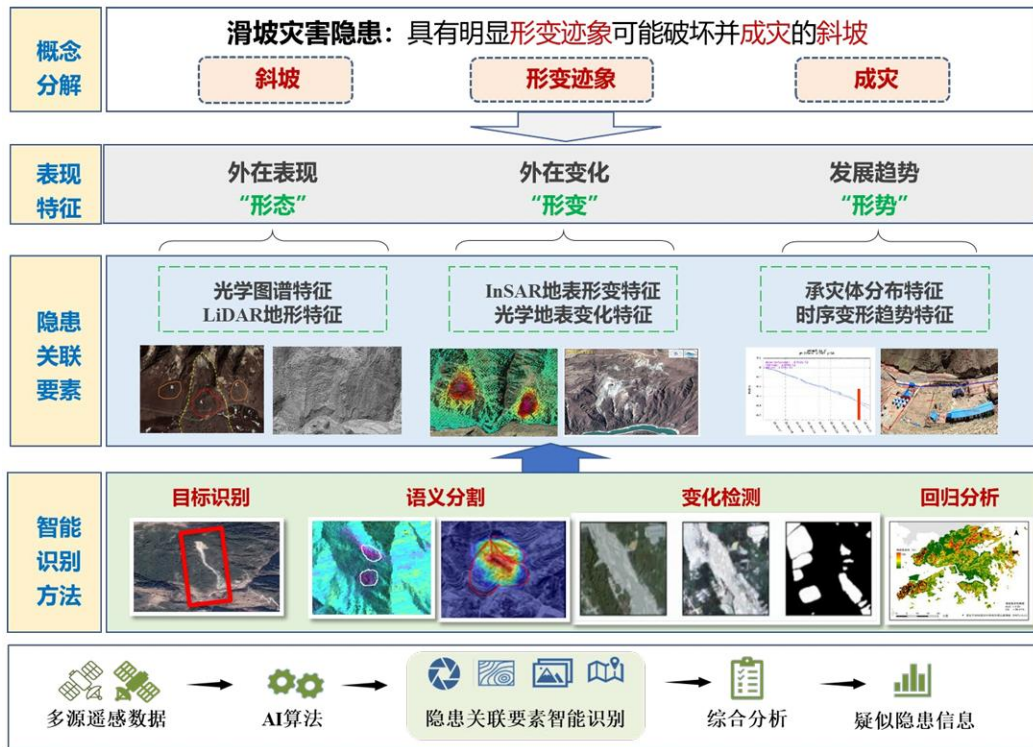
滑坡隐患的智能识别是一个极其前沿的课题，相关概念和技术都处在探索阶段。因此，首要问题是明确智能识别的对象和 AI 技术在传统业务链条中的技术定位。滑坡隐患遥感识别技术发展可概括为四个阶段，“AI+遥感”技术的核心定位是突破滑坡隐患关联要素的智能识别技术。当前研究的主要挑战是创新智能识别技术，构建相关系统平台并为智能识别算法研发、工程化应用实践提供可靠的技术支持与平台保障。



来源：中国自然资源航空物探遥感中心

图 61 滑坡灾害隐患遥感识别发展阶段

明确“AI+遥感”在滑坡隐患识别场景下的对象是重中之重。虽然滑坡灾害的形成过程极为复杂、表现形式多样。但通过光学遥感、InSAR、激光雷达等综合遥感技术可以对滑坡隐患表现出来的“形态、形变、形势”特征进行探测。因此，项目从遥感观测要素-隐患表现特征的关联关系出发，客观的将智能识别的对象限定为“地质灾害隐患关联要素”，如滑坡体、拉裂缝等光学特征、InSAR 地表形变特征以及承灾体等，这些要素是隐患综合判识的重要判断依据。



来源：中国自然资源航空物探遥感中心

图 62 滑坡隐患关联要素识别技术路线图

### (1) 平台研发与工程化应用

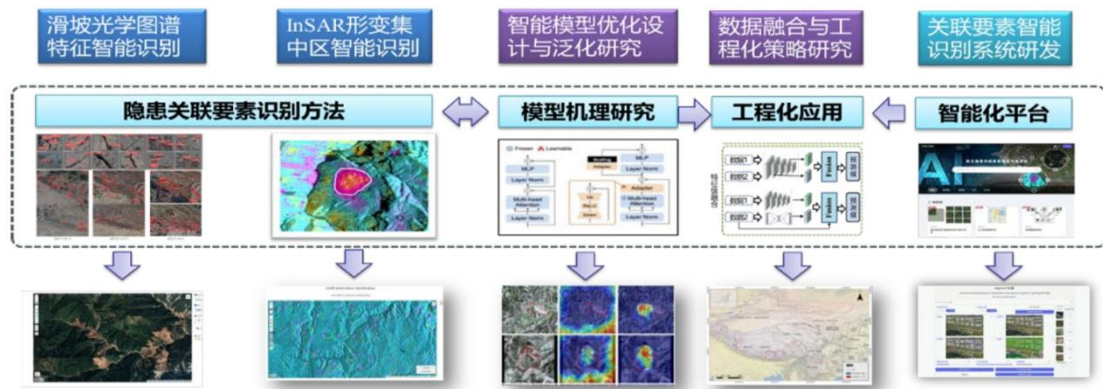
项目组充分结合传统遥感分析平台、人工智能平台与云计算平台的优势，结合地质灾害隐患识别业务需求，以 AI for Geoscience 为核心理念，研发了可支撑开展全国尺度遥感识别应用的综合分析平台。一是实现了 AI 模型从开发到部署的全链条功能，二是研发了面向应用场景的样本库、知识库和模型库，三是提出“1+N”的新型服务模式满足并行式、交互式、流程式等不同场景应用需求。同时，还具备支持大模型等新型 AI 技术的能力。

应用方面，以县为单位开展应用，通过“样本-模型”持续迭代，逐步省级至全国尺度业务化应用推广。目前已成了西部山区滑坡关联地表形变聚集区智能识别工作，共 377 景数据，覆盖面积超 500 万平

方公里，平均推理效率达到 5min/景，进而实现从 AI 模型研发到业务化应用闭环。

## (2) 技术方案

项目组通过开展关键技术攻关，先后在小样本识别、智能模型可解释与可泛化分析、视觉大模型自适应迁移、智能云计算与实时推理、智能交互解译、样本-模型协同增效、模型剪枝等方面取得技术突破。为隐患关联要素识别奠定了扎实的应用基础。



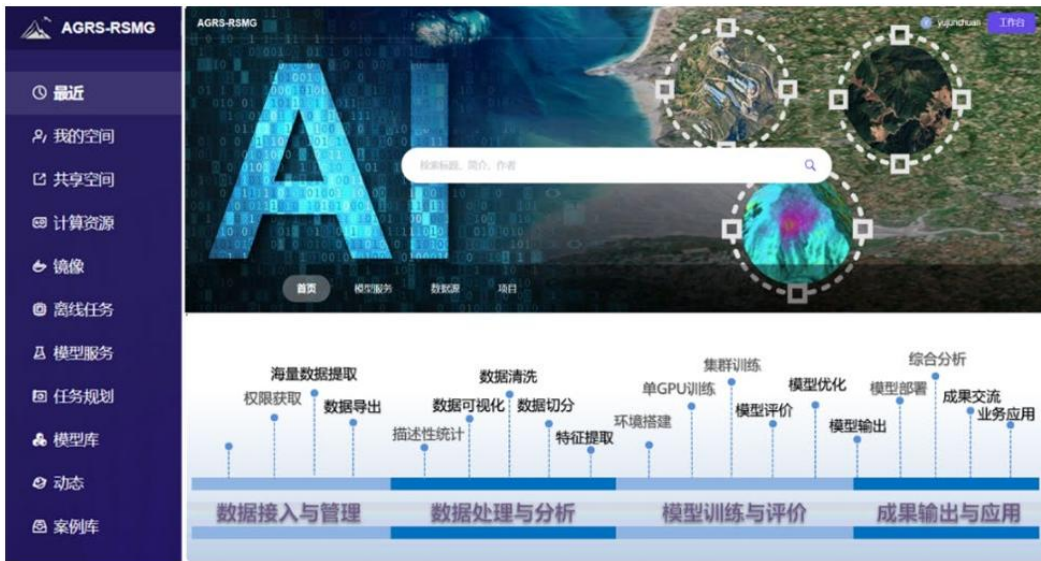
来源：中国自然资源航空物探遥感中心

图 63 项目基本技术路线

## 7. 创新情况：

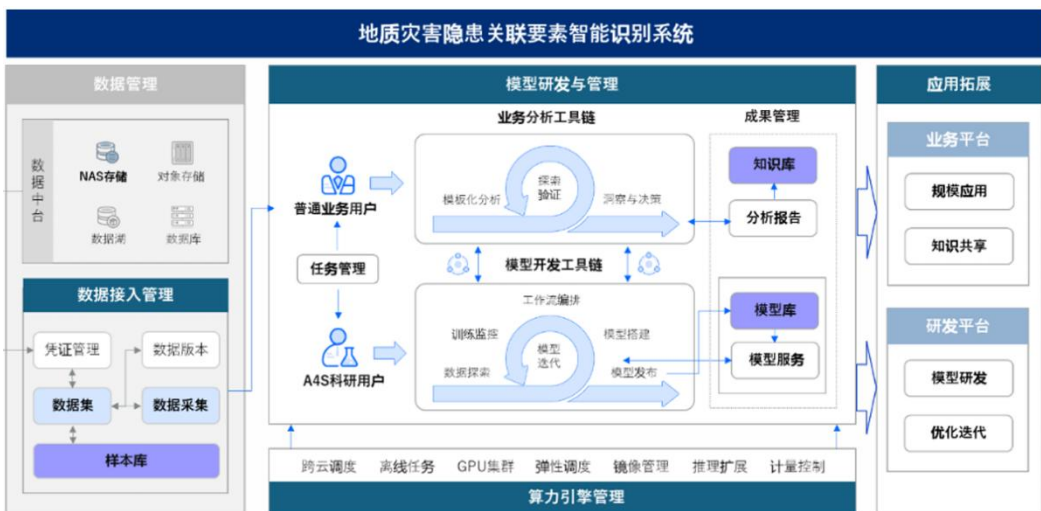
### (1) 系统架构创新

基于微服务架构和容器技术构建集“样本分析—模型迭代—应用推理”于一体的滑坡隐患关联要素智能识别系统。相较于传统遥感分析系统，在多源遥感数据处理，智能模型全流程管理，并行推理等方面优势明显。系统采用组件插件化、多层架构、多模块结构的设计方式。综合利用插遥感智能分析引擎、高性能并行计算引擎和插件 workflow 技术，实现多模态、多场景、多要素、多框架、多机并行的智能遥感智能识别应用能力。



来源：中国自然资源航空物探遥感中心

图 64 系统界面



来源：中国自然资源航空物探遥感中心

图 65 系统基本框架

## (2) 系统功能创新

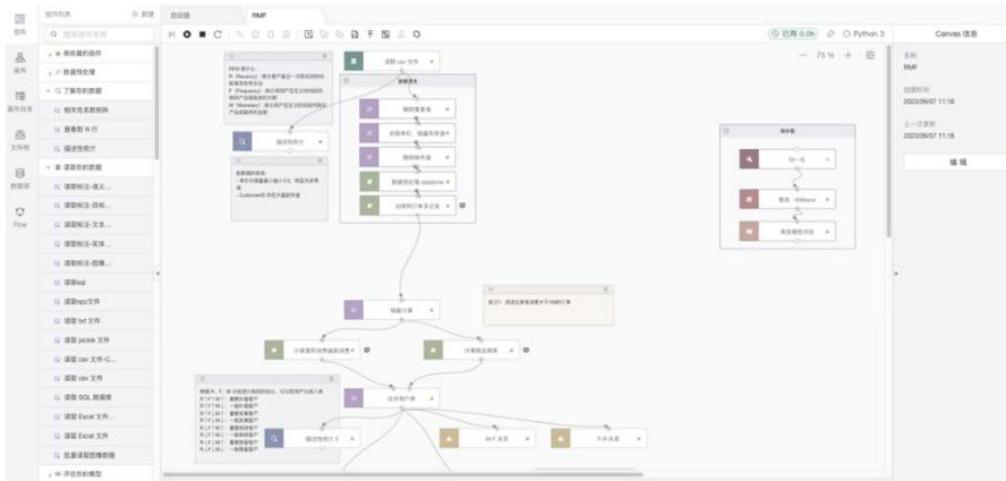
系统包含三个子系统，共集成了 40 余个 AI 和遥感功能算子。功能上的特点可以概括为“六个动态”，即“多源数据动态存取，计算资源动态平衡，软件环境动态适配，代码多版本动态管理，智能模型动态发布，知识动态分享”通过集群存储、有限开源、环境定制、算力调度、多模推理和知识库功能，实现遥感数据高效处理与智能分析的

一体化协同。同时，系统还具备对于语言大模型、多模态大模型和智能体的支持。



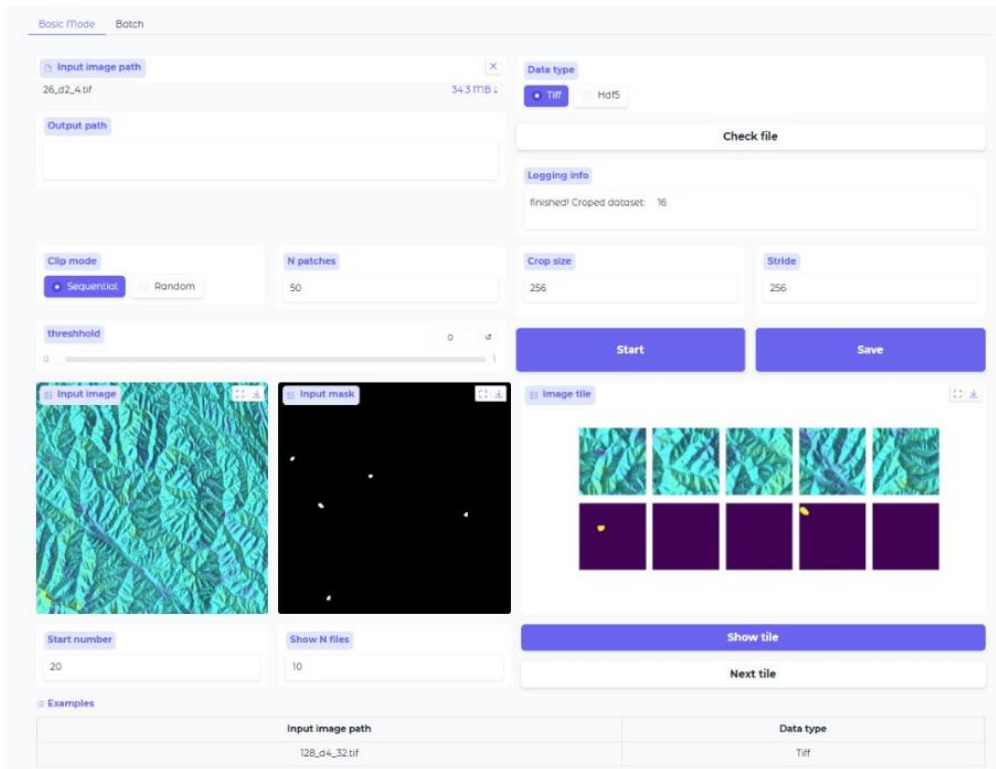
来源：中国自然资源航空物探遥感中心

图 66 滑坡隐患识别功能插件



来源：中国自然资源航空物探遥感中心

图 67 Canvas 低代码模型建模



来源：中国自然资源航空物探遥感中心

图 68 智能交互式 APP 发布

### (3) 智能识别技术创新

系统研发过程中突破了一系列滑坡隐患关联要素智能识别的关键技术。具体包括：突破了基于多源遥感滑坡隐患关联地表形变异常智能识别与广域泛化技术，以视觉大模型交互式解译和“样本—模型”迭代为算法核心，识别精度提升至 85%；提出了滑坡隐患关联图谱特征智能识别及变化检测技术，利用困难样本挖掘技术将隐蔽型滑坡识别精度提升至 70%；建立了基于协同增效的智能模型优化方法，通过特征可视化与数据贡献度定量分析技术，提升复杂场景下关联要素识别可解释性和可泛化性；发展了多模态遥感信息融合与智能模型剪枝技术，实现参数缩减 50%情况下且精度衰减<1%。

### (4) 模式创新情况

系统采用“数据—模型—服务”一体化设计，依托强大的 GPU 计算集群，不仅具备 274 万亿次浮点运算/秒的运算能力和 PB 级存储，还支持离线模型训练和在线推理，显著提高了工程化应用的效率和效果。创新性的研发出“1+N”多场景应用发布能力，为不同需求场景的隐患识别提供服务，打破了传统遥感系统的应用局限性。该系统以 AI for Geoscience 作为研发理念，为新型地学数据智能分析平台的研发提供了有价值的参考。



来源：感识别发展阶中国自然资源航空物探遥感中心

图 69 相对传统系统的优势明显

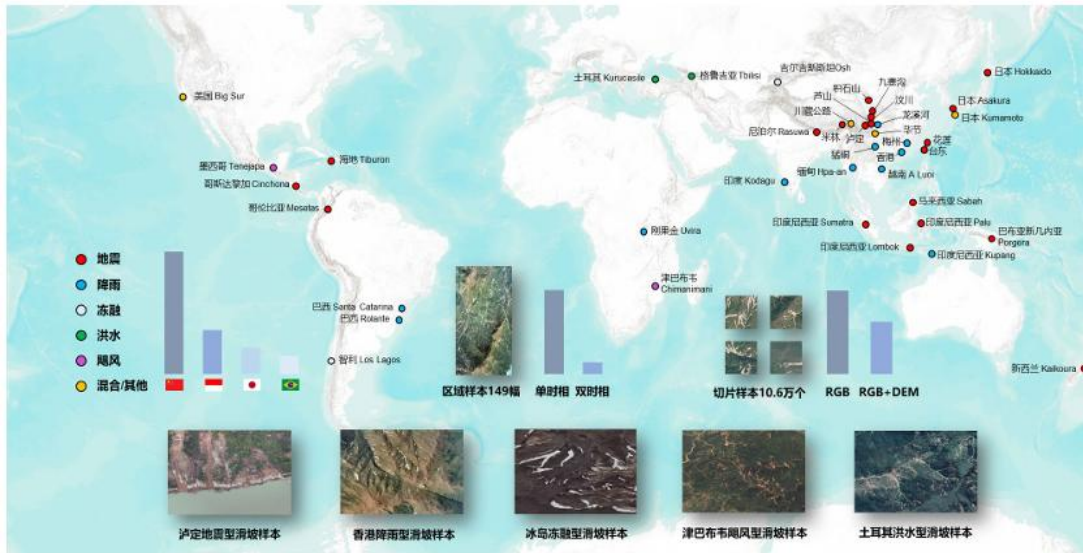
## 8.应用实效：

依托地质灾害隐患关联要素智能识别系统完成了西部山区滑坡关联地表形变聚集区智能识别工作，应用覆盖面积超 500 万平方公里，共识别潜在滑坡目标数万处。该数据直接服务于全国高中易发区崩滑隐患综合遥感识别工作，有效服务于地质灾害防治管理。

地质灾害隐患关联要素智能识别系统已为多个行业用户及实验

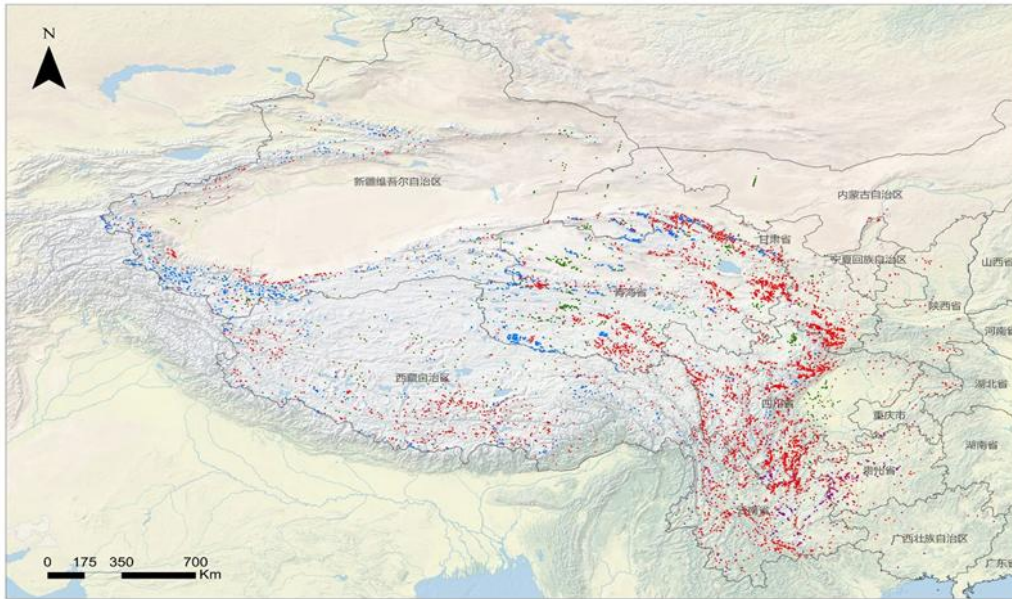
室机构提供超 3 万计算时的智能模型训练与应用服务，受到行业用户的一致好评。

系统研发过程中，项目组构建了两个关联要素智能识别综合遥感数据集，为推动该领域的智能化研究提供了基础数据。一是全球滑坡多源遥感综合样本数据集，该数据集空间范围达全球尺度，总样本量超 10 万。二是全国 InSAR 地表形变聚集区智能识别样本数据集，空间范围覆盖中国西部地区，共包含仅 1 万个样本。



来源：中国自然资源航空物探遥感中心

图 70 全球滑坡多源遥感综合样本数据集



来源：中国自然资源航空物探遥感中心

图 71 全国 InSAR 地表形变聚集区智能识别样本数据集

从遥感观测要素与地灾隐患表现特征的关联关系出发提出了隐患关联要素智能识别的基本理论框架，逐步建立了数据知识联合驱动的隐患关联要素智能识别技术方法体系。从“概念提出—方法研究—生产应用”全方位的展现了基于人工智能和遥感技术在滑坡隐患关联要素识别方面的技术细节，并编制了相关技术规范，对于推动地质灾害隐患识别智能化水平，拓展人工智能在该领域的应用能力具有参考意义。

## 五、生物医药

### (一) 人工智能辅助极端耐碱抗体的设计

#### 1. 案例实施单位：

牵头单位：上海人工智能创新中心

参与单位：上海交通大学、长春金赛药业有限公司

#### 2. 案例背景：

抗体亲和层析技术是生物制药纯化的核心工艺，需耐受强碱清洗（0.5 M NaOH）以重复使用。然而，天然 VHH 抗体在此极端环境下易降解，导致纯化成本激增。传统理性设计依赖专家经验筛选突变，需大量实验验证，效率低且难以捕捉多点突变的复杂上位效应。此外，工业蛋白质需适应非自然极端环境（如 pH>13），而自然蛋白质数据稀缺，AI 模型训练缺乏参考，成为技术瓶颈。

#### 3. 案例概述：

本案例基于上海人工智能创新中心和上海交通大学联合研发的 Pro-PRIME 蛋白质大语言模型，通过两轮 AI 驱动设计实现 VHH 抗体的耐极端环境改造。通过 65 次实验验证即获得工业级突变体，突破传统方法需上千次实验的局限，研发周期缩短至 2 个月。优化后的 VHH 抗体已在长春金赛药业生长激素生产线应用，层析柱复用次数提升 3 倍，年节约生产成本超百万美元。

#### 4. 行业痛点：

极端环境蛋白质自然样本稀缺，AI 模型缺乏训练数据，难以预测非自然适应突变。传统理性设计依赖试错法，实验验证量巨大，单

点突变组合易陷局部最优。抗体纯化占生物制药生产成本的 30%-50%，耐碱性不足导致层析填料频繁更换。

## 5.项目获奖、专利授权及获得投融资情况：

全球首次实现 AI 驱动极端耐碱抗体设计，改造设计的蛋白质产品落地金赛药业生产线，年节约成本数百万美元。

## 6.案例详述：

### (1) 总体思路

本项目以 AI 大模型为核心驱动，构建“AI 预测—实验验证—工业迭代”闭环体系，突破传统蛋白质设计的效率瓶颈。具体流程如下：一是 AI 零样本预测，基于 Pro-PRIME 模型对 VHH 抗体单点突变进行全序列饱和预测，筛选潜在耐碱位点；二是小样本实验调优，利用首轮实验数据（EC50、Tm 值）微调模型，预测多点突变组合；三是工业级验证，综合评估突变体耐碱性、热稳定性及亲和力，优化生产参数并规模化应用。

### (2) 平台介绍

**Pro-PRIME 模型：**多任务预训练框架：基于自主收集的带极端环境标注的序列数据集进行掩码语言建模(MLM)和最适生长温度 OGT 预测，以及相关性 loss 多任务联合训练，学习蛋白质序列语法与其温度属性。

**微调策略：**采用小样本迁移学习技术，仅需 10~20 组实验数据即可适配特定蛋白质的耐性优化目标。

### (3) 技术方案及关键指标参数

### 单点突变设计：

零样本预测：输入 VHH 抗体序列，Pro-PRIME 对 130 个位点进行 20 种氨基酸替换评分，筛选 Top 45 突变体。

### 实验验证：

耐碱性测试：0.3 M/0.5 M NaOH 处理 24 小时，ELISA 法检测 EC50。

热稳定性：差示扫描量热法（DSC）测定 T<sub>m</sub> 值。

结构分析：SDS-PAGE 检测降解程度，质谱定位断裂位点（Q4、G10 等）。

验证结果：15 个突变体在 0.5 M NaOH 下 EC50 降低 30%~60%，35 个突变体 T<sub>m</sub> 值有所提升。

### 多点突变组合优化：

上位效应建模：将单点实验数据输入微调后的 Pro-PRIME，预测 20 个多突变组合，优先选择耐碱性评分和 T<sub>m</sub> 全部大于最佳单点突变体的候选序列。

关键发现：“双负得正”组合：A57D (EC50=1.3) 与 P29T (EC50=1.2) 单独表现不佳，但组合后 EC50 降至 0.4，降解带比例从 61.1% 降至 18.7%。

热稳定协同：A15P；R20T 双突变体 T<sub>m</sub> 值达 75°C (Δ+10°C)，耐碱性 (0.5 M NaOH) EC50 降低 55%。

## 7. 创新情况：

### (1) 技术创新情况

模型层面，实现零样本预测与多任务学习的蛋白质语言模型。工

业场景要求抗体在极端碱性和高温下保持功能，但自然蛋白质数据中此类样本稀缺。基于 Pro-PRIME 大语言模型，首次实现极端环境蛋白质的零样本预测。模型通过自监督学习自然蛋白质序列规律，和生长环境标签数据的联合训练，即可高效预测抗体的耐碱突变；引入多任务学习框架，同时优化耐碱性、热稳定性( $T_m$  值)和亲和力( $EC_{50}$ )，在单点突变筛选中兼顾多重工业指标。从效果来看，Pro-PRIME 的零样本能力填补了数据空白，避免传统方法依赖大量实验的局限性。首轮设计筛选 45 个单点突变体，实验验证 15 个耐碱性提升，其中 A15P 突变体在 0.5 M NaOH 处理 24 小时后活性保留率提升至野生型的 3 倍， $T_m$  值提高 6°C。

**算法层面，实现上位效应感知的多点突变组合优化。**传统方法依赖逐步叠加“最优”单点突变，无法发现“双负得正”组合。通过开发多点突变协同评分算法，捕捉单点突变间的非线性相互作用(上位效应)。例如，A57D 和 P29T 单点突变耐碱性评分较低，但组合后  $EC_{50}$  降低 60%，突破传统贪心算法的局部最优限制。从实效来看，本算法通过实验数据微调模型，精准预测 20 个多点突变，其中 5 个组合耐碱性超越所有单点突变。多点突变体 A57D;G134R;D114Y 在 0.5 M NaOH 处理 24 小时后动态结合容量 (DBC) 保留 52.3%，较野生型 (15.2%) 提升 3.4 倍，直接应用于金赛药业生长激素生产线。

**数据层面，小样本驱动模型快速调优。**工业蛋白质设计需快速响应产线需求，传统深度学习模型依赖海量数据。本方法将实验验证量从数百次压缩至 65 次，显著降低成本。仅用 45 个单点突变实验数

据（耐碱性和  $T_m$  值），通过迁移学习对 Pro-PRIME 模型微调，使其适配极端环境设计需求，大幅提高预测多点突变的准确率。经验证，两轮设计耗时仅 2 个月，筛选出 20 个多点突变体，其中 A15P; R20T 等多个突变体已规模化生产，年节约成本超百万美元。

## （2）模式创新情况

**一方面是解决方案创新**，即实现 AI 驱动的蛋白质快速设计流程。构建“零样本预测→实验验证→工业反馈”全流程解决方案，突破传统试错法。首轮 AI 筛选候选突变，次轮结合实验数据优化模型，形成闭环迭代。设计效率提升 5 倍，金赛药业生产线应用后，层析柱使用寿命从 5 次提升至 15 次。

**另一方面是服务模式创新**，即实现产学研协同的工业适配机制。上海人工智能创新中心（模型开发）、上海交通大学（实验验证）、金赛药业（工业试产）三方协同，实现“AI 设计—实验室—生产线”无缝衔接。突变体从实验室到规模化生产周期缩短至 3 个月，推动生长激素纯化工艺升级，大幅提高产品合格率。

## 8.应用实效

### （1）应用场景与落地规模

核心场景为 VHH 抗体作为亲和层析填料，用于生长激素的工业化纯化。传统抗体在 0.5 M NaOH 清洗后失活，需频繁更换层析柱，成本高昂。优化后的耐碱突变体已在长春金赛药业规模化应用，已经完成 5000L 的大规模放大生产，覆盖其生长激素的多条生产线，累计运行超 12 个月。

## (2) 目标用户与当前使用情况

目标用户为生物制药企业（抗体/酶依赖型纯化工艺）、工业酶生产企业。当前用户包括：长春金赛药业，全面替换野生型 VHH 抗体，层析柱复用次数从 5 次提升至 15 次，填料采购成本降低 65%。

## (3) 解决行业痛点

一是极端环境稳定性不足，野生型 VHH 在 0.5 M NaOH 处理 24 小时后活性仅存 15.2%，突变体 A57D;G134R;D114Y 保留 52.3%，解决层析柱寿命短的核心瓶颈；二是实验成本高，传统设计需上千次实验，AI 模型仅用 65 次实验筛选出工业级突变体，研发成本降低 80%；三是多点突变协同效应，发现“双负得正”组合（如 A57D;P29T），突破理性设计经验局限，上位效应预测准确率超 85%。

## (4) 经济社会效益

一方面实现直接降本，金赛药业年节约层析填料采购、清洗试剂及废液处理成本近百万美元；另一方面促进产能提升，层析柱复用周期缩短 30%，生产线年增产生长激素数十万支，增收数百万美元。

## (5) 开源与生态贡献

技术开源方面，Pro-PRIME 模型部分代码已开源，相关文章已经发表在 *Science Advances*.10,eadr2641(2024)。本项目的抗体改造设计的结果也已经发表在 *eLife*13:RP102788,2024；产学研生态方面，联合上海交通大学、金赛药业打通产学研闭环，共享实验数据与工业场景，加速技术迭代。

## （二）百度 PaddleHelix 平台针对 GPCR 难成药靶点的多肽设计系统的构建

### 1. 案例实施单位：

牵头单位：北京百度网讯科技有限公司

参与单位：中晟全肽

### 2. 案例背景：

多肽作为一类兼具天然活性与良好生物相容性的分子，在新药研发、抗菌材料等领域具有广阔的应用前景。传统的多肽药物设计主要依赖于理性设计、定点修饰或基于天然序列的优化，这些方法往往存在筛选效率低、结构与功能关联认知不足、难以覆盖庞大序列空间等局限性。近年来，随着人工智能技术的发展，基于深度学习的蛋白质设计方法逐渐兴起。相比传统方法，AI 驱动的方法在创新性、新颖性上具有显著优势，为高效开发新型多肽药物和功能材料提供了全新的技术路径。

GPCR 作为一类重要的跨膜蛋白靶点，在药物发现中占有核心地位，是很多疾病的关键靶标。但其结构复杂性和动态特征又给药物设计带来了巨大挑战，GPCR 家族成员众多，亚型间高度相似，如何在提高结合亲和力的同时实现亚型选择性，是多肽药物开发的关键难题。因此，想要针对 GPCR 难成药靶点设计多肽药物分子，就成为了业界的一个难题。百度螺旋桨 PaddleHelix 生物计算团队联合湖南中晟全肽生物科技公司，针对产业界的这一难题，进行联合攻坚研发，开发出业界首个研发出针对 GPCR 跨膜蛋白的多肽药物设计方法，不仅在

技术层面取得突破，并且已在中晟全肽公司内实际应用落地，帮助其将原本需要 3-6 个月的实验周期，压缩到仅需 14 天，大大提升了研发效率。

### 3.案例详述：

#### (1) 总体思路

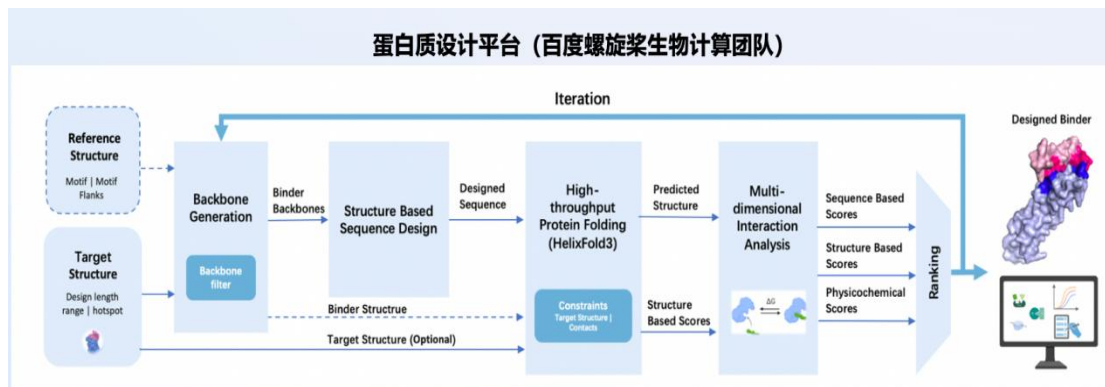
多肽作为一类兼具天然活性与良好生物相容性的分子，在新药研发、抗菌材料等领域具有广阔的应用前景。与小分子药物相比，多肽在与靶点结合时通常具有更大的接触面积和更高的结合特异性，从而能显著降低脱靶效应和毒副作用；同时，多肽分子量相对较小，组织穿透性优于大分子抗体，且不易引发免疫排斥反应。与抗体类药物相比，多肽合成更为简便、成本更低，并能通过化学修饰、环化或非天然氨基酸引入等手段进一步改善其稳定性和药代动力学特性。正因如此，多肽在兼具小分子和抗体优势的同时，能够在多种“难成药”靶点上展现独特潜力。

然而，传统的多肽设计主要依赖于理性设计、定点修饰或基于天然序列的优化，这些方法往往存在筛选效率低、对结构与功能关系的认知不足、难以覆盖庞大序列空间等局限性。近年来，随着人工智能的发展，基于深度学习的多肽设计方法逐渐兴起。AI 能够在大规模多肽数据中学习，实现对多肽性质（如稳定性、亲和力、膜穿透性等）的快速预测与优化。相比传统方法，AI 驱动的设计不仅在搜索效率和创新性上具备突破性优势，还能在功能优化精度上实现大幅提升，为高效开发新型多肽药物和功能材料提供了全新的技术路径。

## (2) 平台介绍

蛋白设计平台 HelixDesign 以自研蛋白结构预测模型 HelixFold3 为核心，其使用同源序列搜索 MSA 特征，参考 Template 特征及输入蛋白序列特征作为输入，利用了特征处理模块 (PairFormer) 和构象预测模块 (Diffusion) 实现了全原子级别的高精度的蛋白复合体构象建模。作为全球首个研发出并开源对标 AlphaFold3 技术的团队，早在 2024 年 8 月开源 HelixFold3 模型的同时，便开始了蛋白质设计全流程的探索工作。在此基础上，为了充分利用先验信息提升预测效果，HelixFold3 在功能上进一步扩展，支持以约束为条件的构象预测：包括残基距离约束为条件的构象预测，参考结构为条件的构象预测等。这些技术和功能的创新为高精度的蛋白设计的评估奠定了坚实基础。

基于 HelixFold3 搭建的 HelixDesign 蛋白质设计全流程系统，通过融合多种蛋白设计模型和工具，实现了一体化端到端蛋白设计。全流程的蛋白质设计，通过高精度的骨架设计模块，序列设计模块，构象预测模块和互作用分析模块的协同，让整体的 pipeline 流程发挥出最大效果。通过引入“RAG”等大语言模型中已经被验证有效的技术，基于输入的靶蛋白构象信息进行蛋白构象数据库的检索用于骨架设计。流程中构象预测模块充分利用了输入及“RAG”得到的靶蛋白信息做为条件，利用 HelixFold3 实现高精度的构象预测。最后通过多维度的评估最终筛选得到候选序列。



来源：北京百度网讯科技有限公司

图 72 蛋白质设计平台流程示意图

HelixDesign 蛋白设计系统支持包括抗体设计，多肽设计，酶设计，小蛋白设计等多场景的蛋白设计能力。设计流程针对不同场景做了针对性优化，如在抗体设计场景，HelixDesign 实现了自动化识别抗体从而使用高精度特征处理流程以保证设计质量。在小蛋白设计场景中，HelixDesign 支持保留关键功能区域实现了设计成功率的提升。在酶设计场景中实现了关键活性位点固定功能，以提升酶设计的成功率。这些针对场景的功能适配和优化，在提升设计精度的同时也增强了平台的易用性。

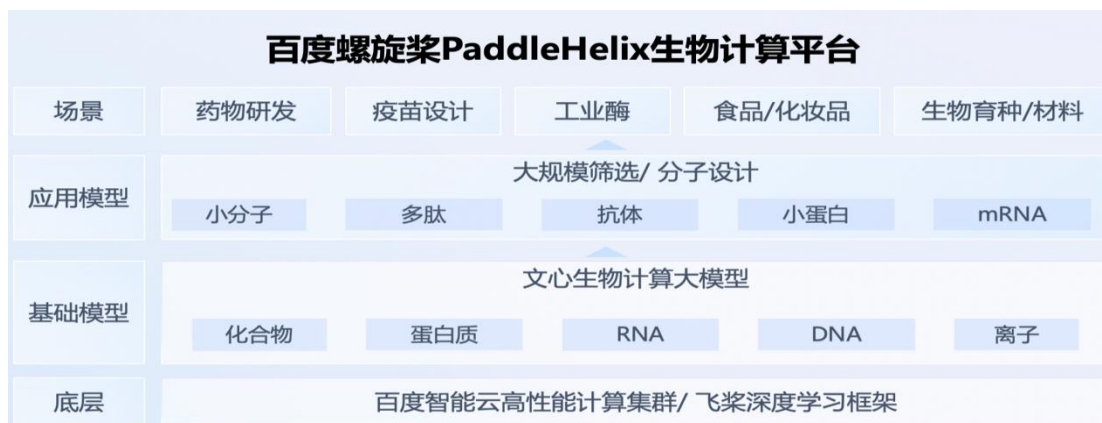
当前，HelixDesign 设计系统，已成功在 20+个靶点上，完成数十万大分子蛋白药物的设计和验证，湿实验的成功率达到 30%，设计出的序列相比野生型有更强的结合能力，在序列的新颖性和多样性上也表现更佳。其设计效果在中晟全肽等公司合作项目中得到验证。

### (3) 技术方案

GPCR 作为一类重要的跨膜蛋白靶点，在药物研发中占有核心地位，但其结构复杂性和动态特征给多肽设计带来了巨大挑战。一方面，GPCR 的配体结合位点多位于跨膜区或膜内腔，空间狭窄且富含疏水

环境，使得体积较大的多肽分子难以有效进入并形成稳定结合；另一方面，多肽自身普遍存在稳定性差、易被蛋白酶降解以及膜渗透性不足的问题，限制了其在体内的药效维持。此外，GPCR 家族成员众多，亚型间高度相似，如何在提高结合亲和力的同时实现亚型选择性和信号偏向性调控，也是多肽药物开发的关键难题。再加之 GPCR 实验结构解析与功能验证通量有限，进一步加剧了传统多肽设计的困难。

百度螺旋桨 PaddleHelix 生物计算团队，依托自研的高精度 HelixFold 系列结构预测模型，构建了一体化的多肽药物设计平台，涵盖骨架搜索→序列设计→结构评估→能量打分排序的自动化流程。同时，针对 GPCR 靶点的特性，创新性的提出：**一是**状态特异性的多肽设计新策略：首次在多肽设计过程中显式引入 GPCR 构象状态，能够定向获得激动剂或拮抗剂多肽，实现对受体功能的精准调控。**二是**建立优化的结构预测与筛选体系：开发 HelixFold 系列模型，针对 GPCR 靶蛋白进行微调训练，显著提升 GPCR-多肽复合物在不同状态下的预测准确性和筛选效率。**三是**实现多靶点实验验证与突破：在 APJR、GLP-1R 和 GHSR 三个 GPCR 靶点上成功设计并验证高亲和力的激动剂和拮抗剂，部分拮抗剂优于临床候选药物，展现了方法的普适性和应用前景。



来源：北京百度网讯科技有限公司

图 73 百度螺旋桨 PaddleHelix 生物计算平台

#### 4.创新情况：

G 蛋白偶联受体（GPCRs）是一个庞大且多样的膜受体家族，广泛参与细胞信号转导，因而成为众多药物研发的重要靶点。然而，现有的药物开发方法，如小分子药物和生物制剂，通常面临脱靶效应等局限，基于多肽的治疗策略因其卓越的高亲和力、选择性和效力，逐渐成为一种有前景的替代方案。尽管近年来基于人工智能（AI）的多肽设计方法，如利用 AlphaFold 等结构预测模型进行高亲和力多肽的设计，已取得了一定进展，但针对 GPCR 的多肽设计依然面临特殊挑战。由于 GPCR 能够在活跃和非活跃状态之间发生动态转变，且其与不同多肽配体的结合方式取决于这些状态，因此，设计既能选择性结合特定受体构象的多肽至关重要。这要求精准预测 GPCR 与多肽结合时的构象状态，才能设计出具备明确功能（如激活或抑制）的多肽。

为了解决这一挑战，百度螺旋桨生物计算团队提出了一种基于结构的 GPCR-多肽设计方法，针对 GPCR 难成药靶点这一难题，在算法层面、流程层面、系统层面进行了一系列的创新优化：

**算法层面**，提出了状态特异性的多肽设计策略&开发了针对 GPCR-多肽复合物优化的预测模型，首次将 GPCR 的动态构象状态（活化与失活）显式引入多肽设计流程，实现了激动剂与拮抗剂的定向设计。相比以往仅基于单一静态结构的设计方法，该策略能够更精确地模拟配体与受体间的构象依赖性相互作用，从而在功能层面实现对 GPCR 的精准调控。这一创新为解决 GPCR 药物开发中普遍存在的“状态依赖性难题”提供了突破口。以往方法多忽略 GPCR 受体的动态构象，本研究首次在设计方法上显式引入受体活化/失活状态，针对性地产生激动剂或拮抗剂多肽，从而提高功能调控的精确性。在 HelixFold 系列基础上，结合大规模 GPCR-多肽复合物数据进行精调，提出并研发了 HelixFold-Multistate 模型。该模型不仅在 GPCR-多肽界面预测上表现优于现有方法，还能更准确地区分 GPCR 关键跨膜螺旋（如 TM3、TM6、TM7）的状态构象，显著提升了在多肽设计中对受体功能状态的捕捉能力。这为功能特异性肽类药物设计提供了坚实的计算基础。

**流程层面**，引入多源多样化肽骨架与逆折叠策略，拓展候选库多样性创新性地使用骨架检索等工具从结构数据库中高效检索多样化的肽骨架，并结合逆折叠模型（Inverse Folding）生成大量候选序列，突破了以往方法依赖天然配体或少数已知结构的局限。这种“结构驱动+序列生成”的设计思路，不仅扩展了可探索的序列空间，还提升了发现新型、非天然多肽的可能性。

**系统层面**，构建了基于多指标的多肽筛选与排序体系。在筛选评

价环节引入 pLDDT、pAE、iPTM 等多维度置信度指标，并结合受体状态一致性过滤，实现了对候选多肽亲和力、稳定性和功能正确性的综合评估。该体系显著提高了从大规模候选库中筛选高质量多肽的命中率，其效果优于现有其他方法，为高效发现可实验验证的候选分子提供了可行路径。

从结果上看，该方法成功推动 GPCR 难成药靶点的多肽发现，特别是在拮抗剂设计方面取得突破（如 GLP-1R 拮抗肽优于临床候选药 Avexitide），展示了对传统难成药靶点提供新型候选分子的潜力。在 APJR、GLP-1R 和 GHSR 三个 GPCR 靶点上分别设计并验证了激动剂与拮抗剂，其中包括低至纳摩尔级的激动剂和功能明确的拮抗剂，证明了方法的普适性与可迁移性。

## 5.应用实效

湖南中晟全肽生物科技公司（以下简称“中晟全肽”），是国家级专精特新“小巨人”企业，凭借全球首创的多肽信息压缩技术（PICT），构建起包含近 5 亿种多肽的分子发现库，为肿瘤、自身免疫疾病等领域的 20 余条创新药管线提供源头创新支持。经过对多家技术供应商的系统评估，中晟全肽最终选择与百度螺旋桨团队达成战略合作。聚焦多肽药物研发的核心难题——跨膜 GPCR 蛋白靶点的多肽药物设计。依托百度的生物计算大模型能力，研发团队创新性地引入“激动与拮抗”两种状态的构象分析，通过指定 GPCR 的构象状态提升蛋白-多肽构象的预测精度，从而精准设计和筛选出指定功能（激动与拮抗）的多肽，研发出基于结构的靶向 GPCR 蛋白的多肽药物设计系统，攻

克了传统方法中无法指定状态的设计问题。该系统在 9 个 GPCR 靶点上均成功设计出多个具有活性的多肽激动剂或抑制剂，为 GPCR 药物设计提供了新范式。相关的研究成果也已在中晟全肽内部的多条管线上落地，帮助其将原本需要 3-6 个月的实验周期，压缩到仅需 14 天，大大提升了研发效率。基于一期良好的合作成果，中晟全肽公司又与百度螺旋桨团队启动了二期合作，探索更具挑战性的多肽修饰设计场景的技术新高地。

本案例提出并验证了一种基于结构预测的多肽设计方法，成功设计并鉴定了针对多种 G 蛋白偶联受体 (GPCR) 的肽类激动剂和拮抗剂。通过该方法，我们在 9 个 GPCR 靶点上筛选并发现了多个具有高活性的肽类激动剂与拮抗剂，重点展示了针对 APJR、GLP-1R 和 GHSR 的设计成果。

在 APJR 的激动剂设计中，经过多轮计算筛选，从中挑选出 11 个高排名肽，并在体外实验中验证其显著的结合亲和力。特别是 A6 和 A11 肽，其 EC50 值低于 10 nM，表现出卓越的活性。此外，也在 APJR 的拮抗剂设计中取得了显著成果，这些肽在抑制 APJR 活性方面表现出了强大的效果。对于 GLP-1R，设计了多种拮抗剂肽，尤其是 I3 肽，其亲和力超越了现有的临床药物 Avexitide，展示了更优的抑制效果。而在 GHSR 激动剂设计方面，筛选出多种具有低于 1 $\mu$ M EC50 值的肽，特别是 G4 肽，其效力接近 100 nM。在拮抗剂设计中，Peptide G8 成功抑制了 Ghrelin 与 GHSR 的结合，表现出 3.3 $\mu$ M 的 IC50 值。

最终验证了基于结构的多肽设计方法的有效性,证明了该方法能够成功应用于多个 **GPCR** 受体靶点,并为肽类药物的开发提供了坚实的理论依据与实验数据支持。

### （三）华为云盘古药物分子大模型助力华中科大快速研制出全球首个利什曼病抑制剂

#### 1.案例实施单位：

华为云计算技术有限公司

#### 2.案例背景：

利什曼病是一种由利什曼原虫感染引起、白蛉传播的人兽共患寄生虫病，主要有皮肤型、内脏型和黏膜皮肤型 3 种类型。其中内脏利什曼病又称为黑热病，死亡率超过 95%。利什曼病在全球 102 个国家和地区广泛流行，威胁超过 3.5 亿人口，每年新发病例为 70 万至 100 万例，但至今无有效疫苗和特异性药物。

#### 3.案例概述：

华中科技大学的李岩教授团队利用华为云盘古药物分子大模型，在短时间内完成了对利什曼原虫 SCARE1 毒力蛋白的发现、机制解析及新型小分子抑制剂的开发。该模型显著缩短了药物筛选和优化的时间，使得利什曼病的预防性药物研发周期缩短至数月，研发成本降低了 60%以上。这些新型抑制剂不仅在细胞和动物模型中显示出显著的抑制效果，还展示了一定广谱抗微生物感染的潜力，为解决全球范围内的利什曼病防控问题提供了新方案。

#### 4.行业痛点：

（1）利什曼病在全球多地流行，人畜共患，数亿人面临感染风险，却没有人用疫苗，临床治疗手段有限，存在耐药性严重、注射给药方式不便、毒副作用较大等问题，急需研发好用的特异性药物。

(2) 新药研发周期长、成本高，创新药有个不成文的定律，“双十定律”，一款创新药的研发需要经历 10 年花费 10 亿美元才有可能成功。

## 5. 案例详述：

### (1) 总体思路

利什曼病，这种被忽视的热带传染病，每年威胁全球数亿人口。然而，至今尚无有效的疫苗或特异性药物来预防这一疾病，给全球公共卫生带来了巨大挑战。近年来，随着病原微生物耐药性逐渐增加，开发新型广谱抗感染药物的需求变得更加迫切。

华中科技大学同济医学院李岩教授团队专注于研究利什曼原虫的致病机制，旨在揭示其免疫逃逸的分子机制。结合华为云盘古药物分子大模型进行蛋白-蛋白相互作用预测以及基于质谱的互作蛋白垂钓技术验证，团队首次发现了利什曼原虫的一种关键毒力蛋白——SCARE1，并揭示了其通过劫持宿主关键蛋白来阻碍巨噬细胞吞噬溶酶体合成的机制，从而使利什曼原虫能够成功逃避宿主免疫系统的攻击。这一发现为利什曼病的防治提供了全新的科学基础。

发现 SCARE1 蛋白只是第一步。团队还需要找到一种能够有效阻断这一机制的小分子抑制剂，以实现利对利什曼病的预防和治疗。这一过程往往涉及大量的小分子化合物筛选和结构优化工作，通常需要数年时间。然而，李岩教授团队意识到，传统的药物筛选方式在时间和成本上都难以满足这一急迫的公共卫生需求。

### (2) 技术方案

为了加速药物的发现和开发，李岩教授团队与华为云合作，利用其先进的盘古药物分子大模型。该模型具备强大的蛋白质-蛋白质相互作用（PPI）预测功能，通过分析 SCARE1 蛋白的氨基酸序列和结构信息，快速预测出与其互作的关键宿主蛋白。这一预测结果不仅通过实验得到了验证，还为团队提供了潜在药物靶点的明确方向。

接下来，团队利用华为云盘古大模型的分子对接和药效团筛选功能，从庞大的化合物库中精准筛选出多种具有显著宿主蛋白结合潜力的小分子抑制剂。通过实验验证，这些小分子抑制剂显示出了显著的抑制 SCARE1 与宿主蛋白互作的效果，有效阻断了利什曼原虫的感染过程。

在 AI 技术的强大助力下，李岩教授团队成功开发出了全球首个能够预防性给药的利什曼病抑制剂。这一新型药物不仅在细胞和动物模型中展示了优异的预防效果，还在其他病原微生物（包括李斯特菌等）感染中展现出广谱抗感染潜力。这一发现为利什曼病的防治以及药物耐药性等全球性公共卫生挑战提供了新的解决方案。

此外，华为云盘古大模型的应用极大缩短了药物筛选和优化的周期，使得从发现到药物候选体的研发时间从传统的数年缩短至数月，研发成本降低了 60% 以上。这一进展不仅使得新药研发更为高效和经济，还大大降低了药物研发的门槛，激发了更多研究团队和企业参与到药物开发的过程中。

## 6. 创新情况：

(1) 利用华为云盘古药物分子大模型进行高效蛋白质-蛋白质

相互作用（PPI）预测，首次发现利什曼原虫中的关键毒力蛋白 SCARE1。

(2) 揭示了利什曼原虫毒力蛋白 SCARE1 劫持宿主关键蛋白、防止吞噬溶酶体合成的免疫逃逸机制。

(3) 利用华为盘古药物分子大模型，快速筛选出能够靶向宿主蛋白的新型小分子抑制剂，并通过实验验证其有效性。

### 7.前景与未来应用：

当前，李岩教授团队正在积极推进该预防性药物的临床前研究，并计划在全球范围内开展临床试验。这一突破性成果不仅为利什曼病的防治带来了希望，还为其他感染性疾病的治疗提供了新的思路和方法。特别是在全球抗生素耐药性问题日益严峻的今天，这一研究为开发新型广谱抗感染药物奠定了坚实的基础。

未来，团队计划继续利用华为盘古药物分子大模型，探索更多病原微生物的免疫逃逸机制，并开发出更多具有广谱抗感染潜力的预防性药物。这一研究方向有望为全球公共卫生带来重大贡献，并在应对未来传染病爆发和抗生素耐药性挑战中发挥关键作用。

总之，AI 技术与生物医学研究的结合，正在加速推动新药研发的进程，为解决全球性公共卫生问题带来了前所未有的机遇。李岩教授团队的研究成果，正是这一变革的生动例证。

## （四）智慧芽 Eureka 平台：垂直领域大模型赋能生物医药研发全流程智能化

### 1. 案例实施单位：

智慧芽信息科技有限公司（苏州）有限公司

### 2. 案例背景：

在生物医药领域，新药研发具有成本高、研发周期长、成功率低三大高风险性质。据《Nature》报道，新药研发平均成本高达约 26 亿美元，研发周期耗时约 10 年，而成功率却不足十分之一。传统人工检索分析模式难以应对海量数据，60%的早期研发因专利布局疏漏导致商业化失败，亟需 AI 技术重构研发流程。

### 3. 案例概述：

智慧芽凭借其在知识产权领域十余年的深厚知识积累与深入研究，推出了“Eureka”平台——AI 驱动的生物医药数据分析平台。该平台依托智慧芽生物医药垂直大模型（70B 参数），构建靶点分析、专利布局、临床数据关联等 AI Agent，已服务药明康德等 500 余家生物医药企业，实现研发关键环节效率提升 3-5 倍。有助于全方位提升生物医药领域知识产权的创造、运用、保护、管理和服务能力，赋能新质生产力发展。

### 4. 行业痛点：

在信息爆炸时代，传统的情报模式难以助力药物创新，面临诸多瓶颈，包括：数据碎片化问题突出，分析能力薄弱，技术迭代风险难以预判。例如，靶点专利检索需人工选 2000+文献，耗时 1 周；临床数

据与专利布局脱节,40%药物因侵权风险被迫终止研发;研发方案生成依赖专家经验,创新效率低。

## 5.项目获奖、专利授权及获得投融资情况:

与本项目相关的知识产权共 20 项,包括授权发明专利 11 项,软件著作权 5 项,审理中发明专利 4 项。智慧芽垂直领域大模型已通过国家网信办双备案;入选“2024 中国 AI 大模型产业图谱”。相关成果获“数据要素×”大赛江苏分赛科技创新赛道一等奖、北京分赛医疗健康赛道二等奖、广东分赛科技创新赛道优秀奖。

## 6.案例详述:

### (1) 总体思路

本项目以“垂直大模型+生物医药研发全流程”为核心,通过 AI Agent 替代人工完成专利检索、靶点分析等重复性工作,构建“数据驱动—智能决策—成果转化”的闭环体系

### (2) 平台介绍

智慧芽自主研发的垂直领域大模型(70B 参数,基于 Transformer 的 Decoder-only 架构),以“AI+科研”为核心定位,构建了覆盖知识产权、生物医药、材料科学等领域的专业化 AI Agent 矩阵——“Eureka”平台。该平台首批上线近 20 个高精度智能体,包括查新检索、专利说明书撰写、技术方案探索、生物医药百科问答、材料性能分析等,通过行业级数据与业务逻辑的深度融合,显著提升科研任务的自动化水平。例如:一是专利查新检索, AI Agent 可一键生成专家级查新报告,降低 30%无效检索工作。二是技术方案探索,基于 TRIZ 创新理

论，AI可替代传统“头脑风暴+专家访谈”流程，将方案生成周期从2周缩短至2小时。三是生物医药与材料分析，AI Agent支持即时解答专业问题，并提供可视化分析报告，大幅提升研发效率。目前，Eureka平台已服务超5000家客户，并与多家行业头部企业达成战略合作，推动科研创新从“人工驱动”向“人机协同”转变，加速科研范式升级。

### (3) 技术方案

**一方面构建“多层次、全链路”架构。**在数据层，自建“AI技术+专家审核”数据生产方式，整合全球172个受理局专利、PubMed文献、ClinicalTrials.gov临床试验数据及上海本地药企研发数据，形成结构化数据库。运用大数据分析、机器学习及人工校验技术，实现数据的高效治理与智能应用。**在算法层**，部署NLP模型（处理专利文本、科研论文）、计算机视觉模型（解析化学结构式、基因序列图谱）、知识图谱（关联靶点-药物-专利企业关系）及垂域大模型（PharmaGPT），实现多模态智能分析。**在应用层**，开发标准化API接口，支撑高价值专利识别、侵权监测、价值评估等功能模块，通过Web端、移动端及政企对接平台实现服务输出。

**另一方面实现九项关键技术。**一是基于先进的预训练语言模型技术和实体识别、实体链接技术，实现多源异构数据的实体关系建模，建立生物医药实体与科技文献之间的关系，提高生物医药领域信息抽取的准确性与全面性。二是通过识别查询语句中的实体与属性，并对实体进行链接，对属性进行标准化，判断查询语句的意图，再得到搜索引擎的查询语句，最终得到查询问题的结果。三是改变DPO数据

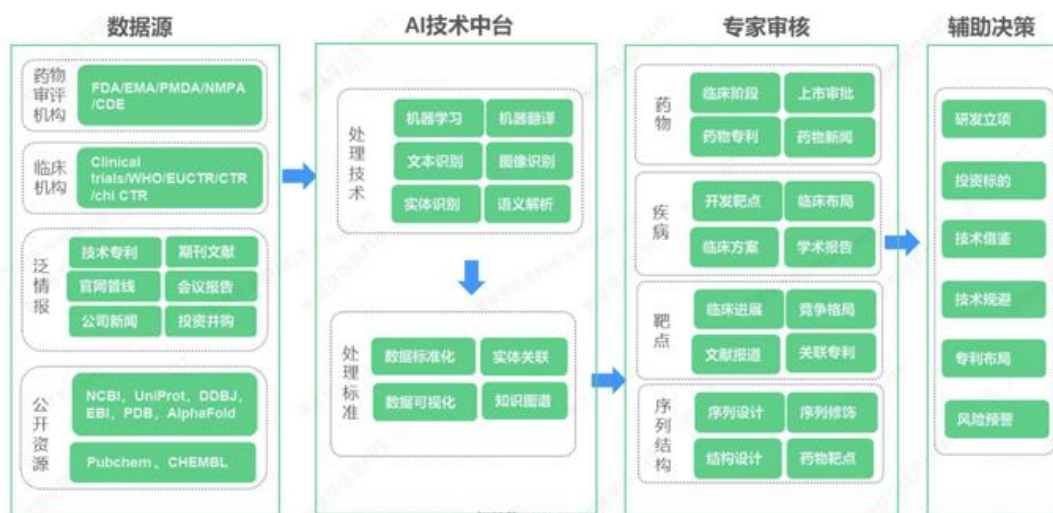
利用方式，改为“做加法”，针对典型的需要负反馈的业务场景合成少量 DPO 数据进行训练尝试，再逐步增加基于 nemotron 的合成数据。

四是构建知识增强的问答框架，融合知识图谱和大语言模型技术，解决生物学领域问答系统的准确性问题，提供基于可靠知识的精确答案，缓解 LLMs 可能产生的“幻觉”问题。五是构建技术思维框架，开发结构化决策支持模型，通过系统性分析问题、识别关键因素，制定解决方案，提高技术决策的可行性、严谨性和周密性。六是通过 GraphRAG 技术实现问题扩展能力，通过连接不同信息片段提供新的合成洞察力。七是设计跨文档知识聚合机制，通过预构建的语义关联网络实现分布式知识点的动态重组，支持全景式知识呈现。八是创新偏好建模方法，在模型中添加两个 MLP 层，使得模型能够输出一个表示人类偏好的分数，并使用二元排序损失来优化模型。九是研发领域自适应训练框架，利用知识产权领域的垂类数据，持续训练开源 Transformer 结构的模型，实现通用模型向专业领域的精准迁移。



来源：智慧芽信息科技（苏州）有限公司

图 74 总体架构图



来源：智慧芽信息科技（苏州）有限公司

图 75 技术架构图

#### (4) 关键指标参数

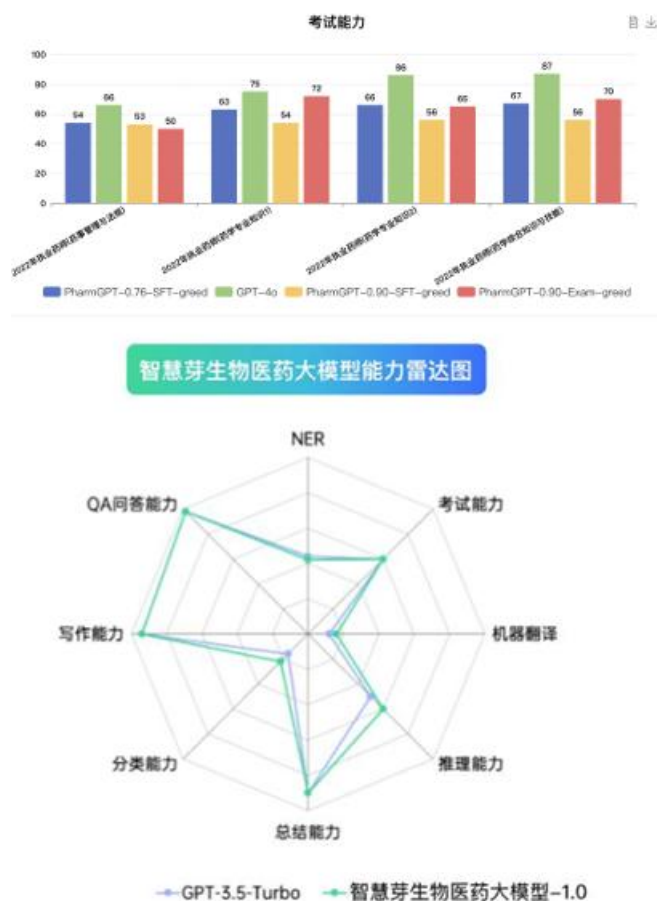
医药实体提取技术从专利、文献、临床实验、药物审批、新闻中提取药物、靶点、适应症信息，一般采用精确率（precision）来评价

最终效果。如下表所示。

表 3 关键指标参数

	Paprr	Patent	CT	DA	News
Drug	79.08	89.24	89.92	86.96	83.59
Target	79.46	92.09	88.35	81.58	72.50
Disease	82.06	91.75	91.40	80.61	80.28

自研垂域大模型方面，自研大模型的能力一般从考试能力、QA 问答能力、NER 能力、机器翻译能力、写作能力、推理能力、分类能力、总结能力等这几个维度来评价结果。

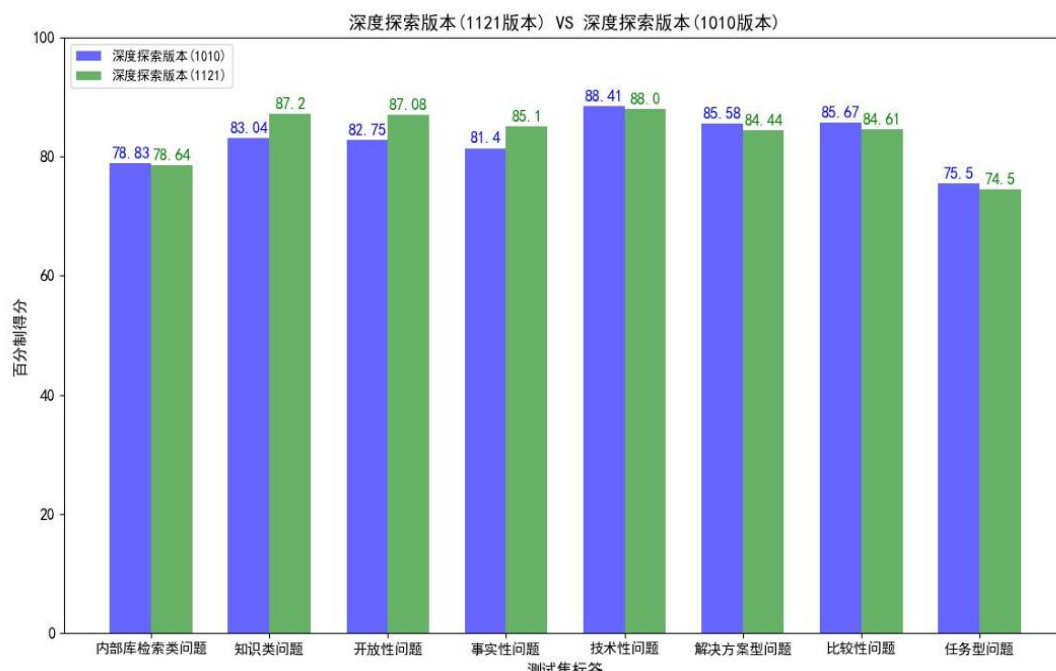


来源：智慧芽信息科技（苏州）有限公司

图 76 智慧芽生物医药大模型能力雷达图

深度探索技术方面，从内部库检索类问题、知识类问题、开放性 问题、事实性问题、技术性问题、解决方案型问题、比较性问题、任

务型问题这几个维度来评测加入深度探索技术后，模型回答的效果。



来源：智慧芽信息科技有限公司（苏州）有限公司

图 77 智慧芽生物医药大模型深度探索技术

## 7. 创新情况：

### (1) 技术创新情况

自研医药垂直领域大模型自主研发 PharmGPT 模型，融合命名实体识别、化学式图像识别与行业知识增强技术，构建覆盖 1600 万+ 专利、100 万+ 临床试验的医药语义理解框架。采用 200 张 A800 GPU 训练 200G 纯文本数据，突破性解决 AI 医药问答中的“幻觉”问题，实现分子结构-适应症-研发进展的多维度精准解析。

多模态数据协同与应用打通新药情报、生物序列与化学物质数据库，建立跨模态检索分析引擎。通过化学结构式与靶点蛋白序列的关联映射，支持从化合物筛选到适应症预测的全链条研发分析，帮助药企缩短先导化合物发现周期，降低临床前研究成本。

引入专家审核机制针对化学结构式、适应症描述等复杂数据，借

助强大药专家团队进行数据指导与审核，同时开发规则引擎与自适应标注工具，确保药品研发数据全生命周期可信度。

产品矩阵协同深度融合 **Pharmsnap** 与 **Bio/Chemical** 数据库，构建跨维度关联分析引擎，基于强化学习推荐新药研发路径。开发竞争情报热力图与管线碰撞预警系统，实时预测靶点扎堆风险，辅助企业避开同质化竞争。

## (2) 模式创新情况

“AI+专家”协同服务模式重构突破传统 AI 工具“纯技术输出”的局限，构建“AI 基础分析+专家深度优化”的分层服务体系。AI Agent 先完成专利文献检索、靶点数据整合等标准化工作，将原始数据转化为结构化分析报告——这一步骤较全人工模式效率提升 3 倍；随后由拥有 10 年以上经验的生物医药专家介入，结合行业经验对报告进行校验、补充产业洞察，并提供定制化建议。例如为某 mRNA 疫苗企业提供的服务中，AI 在 2 小时内完成全球 3000+ 相关专利的初步分析，专家用 1 个工作日完成方案采纳率大幅提升。

研发数据资产化与共享机制首创“生物医药研发知识模块”体系，将每次专利分析、靶点研究的成果拆解为可复用的标准化模块（如“PD-1 靶点专利布局模板”“单抗药物侵权风险核查流程”），存入企业专属知识库。中小企业无需重复开展基础研究，可直接调用模块并结合自身需求调整，技术调研成本降低 60%。

## 8. 应用实效

本项目已落地药明康德、某创新药企等场景，覆盖药物研发前期

靶点筛选、中期专利布局、后期商业化全阶段，服务生物医药企业超 500 家。

为某单抗药物企业缩短靶点专利检索时间从 1 周至 6 小时，漏检率从 18%降至 2%；帮助某药企规避 2 项核心专利侵权风险，避免研发投入损失 1.2 亿元。

服务企业平均缩短研发周期 1.5 年，单药研发成本降低 15%，累计带动产业增值超 3 亿元。应用实效（1000 字内）社会效益：提升生物医药研发的创新效率，加速新药研发和上市，为患者提供更多有效的治疗药物，改善民众健康水平。为国家及地方知识产权战略实施提供实践经验，形成可复制、可推广的“AI+生物医药”服务模式，推动全国生物医药产业升级。

本项目基于预训练语言模型技术训练得到全球领先的生物医药实体抽取模型，从情报提供上大幅提高新药研发立项成功率，降低新药研发风险，获得生物医药相关客户的广泛好评。相关知识产权保证了服务的技术优势。已形成标准化解决方案，在长三角生物医药产业园推广，与生物医药行业协会建立合作，共同开展“AI 赋能生物医药研发”培训，累计培训企业研发人员 1200 人次，推动行业整体智能化认知提升。

## 六、智慧农业

### （一）华为云×中国农业科学院联合研发农科大模型

#### 1.案例实施单位：

华为云计算技术有限公司、中国农业科学院生物技术研究所、中国农业科学院作物科学研究所、中国农业科学院北京畜牧兽医研究所

#### 2.案例背景：

强国必先强农，智慧农业是农业强国建设的战略制高点。生物育种作为农业核心领域，长期受限于经验依赖强、育种偶然性大、研发周期冗长等问题，试错成本高昂，严重影响现代农业发展进程。为破解这一困境，中国农业科学院生物技术研究所牵头，联合华为云，以技术创新推动育种范式变革。

#### 3.案例概述：

多方联合打造了农科大模型，依托华为云 AI 科学计算平台与盘古 LLM 模型，构建三大核心引擎，形成“海量知识综述生成—全维度数据挖掘—基因功能深度解码—性状智能设计模拟—种植系统智慧管理”全流程数智化闭环。该模型在助力水稻株型改良等领域取得突破，使水稻抗倒伏能力显著提升且产量稳定，同时将基因研究方案设计周期从数月缩短至数周，高效推动育种范式从“经验试错”迈向了“智能创制”。

#### 4.行业痛点：

传统育种高度依赖育种家经验，品种培育偶然性大，培育周期普遍较长，新品种迭代速度往往滞后于市场与生产需求。高昂的试错成

本与漫长的育种流程，导致种质创制效率低下，优良品种供给抗风险能力有限，不仅制约农业生产提质增效，更影响种业科技自立自强与国家粮食安全保障。

## 5.案例详述：

### (1) 总体思路

项目构建了三位一体的协同创新体系，合作组织架构方面，项目由中国农业科学院生物技术研究所、作物科学研究所、北京畜牧兽医研究所三大核心科研单位，与华为云 AI 科学计算团队组成联合攻关组。技术底座则深度融合了华为云 AI 科学计算平台、数仓和盘古基础 LLM 模型的能力，构建了“算力-模型-数据”三位一体的农业科研的新架构。该架构具备自主可控、可持续演进的特性，旨在打造国际领先的 AI 生物育种引擎。

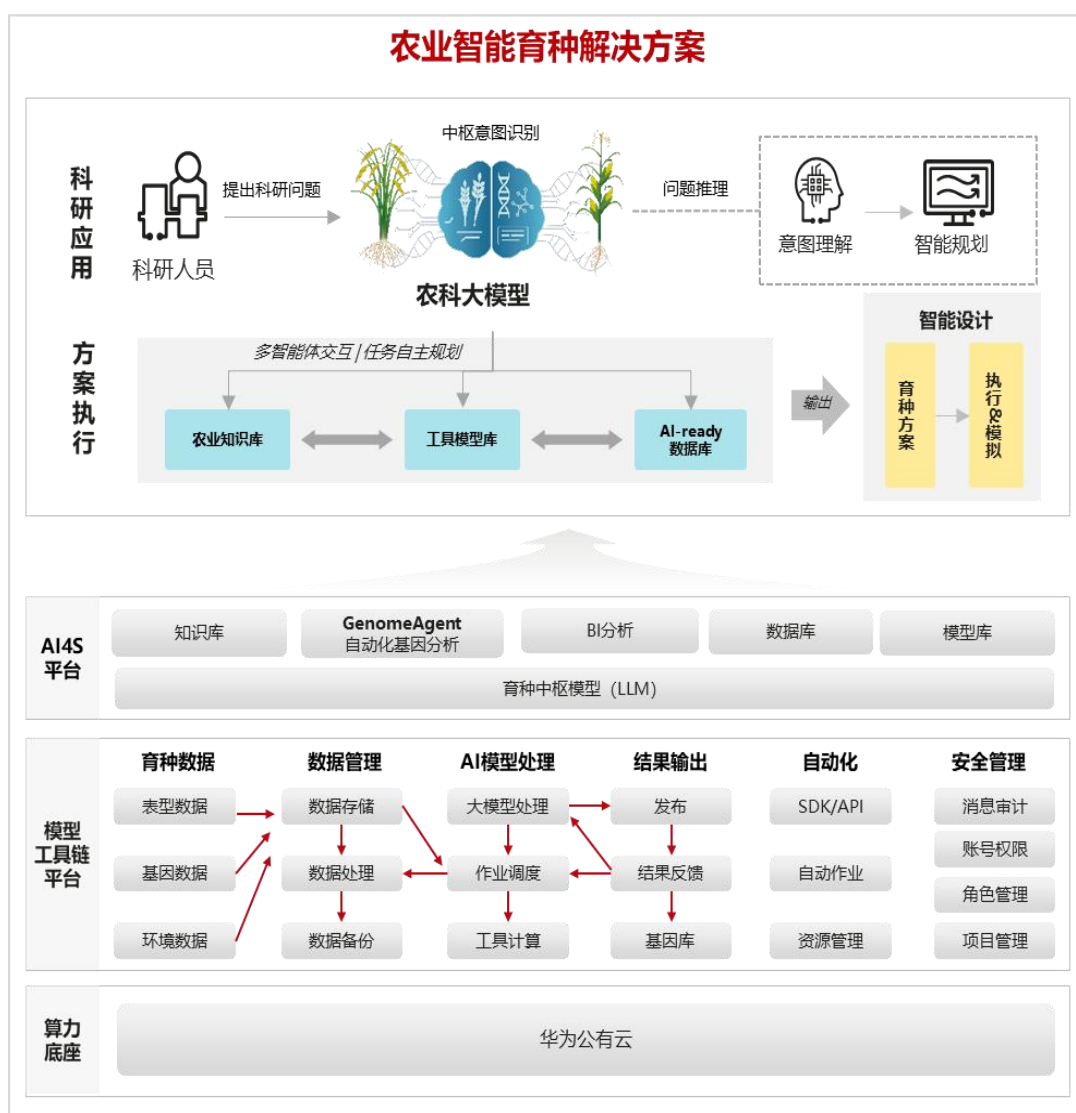
### (2) 平台介绍

平台采用“三引擎一系统”核心架构，打造生物育种全流程数智化支撑体系：**一是**知识计算引擎整合超 300 万篇农业文献及 1000 万+摘要，实现知识智能挖掘；**二是**数据分析引擎汇聚 50+个核心物种的百亿条多组学数据，支持跨维度深度分析；**三是**智能交互平台提供“零代码”对话式操作，可智能调用 100+种专业工具；**四是**智能体系统实现科研任务自主分解与并行计算，高效支撑育种任务全链路多步骤地执行。

### (3) 技术方案

以“协同攻关+技术融合+全链覆盖”为核心技术路径：合作层面组

建科研与技术联合团队，形成跨领域攻坚合力；技术层面构建“算力-模型-数据”一体化底座，支撑“三引擎一系统”高效运行；应用层面覆盖基因挖掘、性状设计、方案规划、智慧种植全链条，创新引入专业AI智能体，实现“实验室-温室-大田”闭环验证，推动育种范式从“经验试错”向“智能创制”转型。



来源：华为云计算技术有限公司

图 78 农业智能育种解决方案

#### (4) 关键指标参数

知识库规模方面，收录超 300 万篇农业全文文献、1000 万+文献

摘要，知识覆盖率较传统数据库提升 3 倍以上；数据资源方面，整合 50+个核心物种（覆盖中国 90%主要农作物），多组学数据总量超百亿条；方面工具支持，内置 100+种专业生物信息学工具，支持“零代码”自动化分析；效率提升方面，单个分析任务可完成超千步骤模拟推演，方案设计周期从数月缩短至数周。

## 6.创新情况：

### （1）技术创新情况

**全流程数智化闭环：**构建了从“AI 读文献筛选候选基因”到“AI 分析多组学数据验证基因功能”，再到“AI 辅助设计实验方案”的完整闭环，实现育种科研范式升级。

**五大核心能力跃升：**包括知识规模，构建包含 AI 农业数据库 300 万篇文献、1000 万摘要的农业 AI 知识库；**数据维度**，覆盖 50+个核心物种，数据量超百亿条；**工具集成**，融合 100+种专业分析工具与预测模型；**智能设计**，支持多维度基因功能预测与位点精准设计；**生产延伸**，将能力延伸至大田种植的智慧管理环节。

**世界一流且自主可控的 AI 技术底座：**一方面支持澎湃算力，基于昇腾的 AI 科学计算平台，为海量数据的处理、大模型的训练提供了稳定、高效的算力保障，夯实了项目的物理基础。另一方面支持强大模型，以盘古 LLM 模型作为领域模型的技术基座，其出色的通用能力大幅降低了领域适配成本，使科研团队能站在巨人的肩膀上，专注于解决农业领域的专业化难题。

**系统架构创新方面：**一是建立中国首个农业科研领域大模型系统

评测体系；二是开发全新框架农业智能系统；三是实现多智能体自主调用与编排支持 24 小时内完成超千步骤的计算与设计模拟。

**交互方式突破方面：**融合自然语言交互能力，科研人员可通过“零代码”方式完成复杂分析，关键研发环节效率提升约 10 倍。

## (2) 模式创新情况

**一方面构建了“结对攻关”的深度联创的协作模式创新。**从人才与智力投入来看，双方集结了各自领域的精锐力量，由华为的 AI 算法专家、架构师和工程师团队，与中国农业科学院的农业科学家、育种家和数据科学家组成联合攻关组，实行长期“驻场协同”办公。这种“AI+农业”的深度智力融合，打破了技术与学科的壁垒，是项目成功的关键组织保障。**从敏捷高效的项目管理来看，**采用“按周滚动、按日调度”的敏捷开发模式，确保了项目在面对复杂的技术难题时，能够快速响应、迭代优化，有力保障了研发进度。

**另一方面赋能了从“技术突破”到“行业引领”的产业生态。**从科研成果影响力来看，双方充分发挥各自领域的平台优势，构建了科技与农业深度融合的立体化推广。依托华为开发者大会（HDC）、华为全联接大会（HC）展示技术底座的先进性，同时借助中国农业科学院在中国国际农产品交易会（CIAPF）、经济信息联播（CCTV2）等权威渠道的影响力彰显行业应用的示范价值。这种跨界联动的传播模式，实现了科研成果在科技圈与农业领域的双重引爆，极大地提升了项目的行业话语权和国际知名度。**从共建生态来看，**基于华为提供的 AI4S（AI for Science）开放创新框架，中国农业科学院正在从单纯的技术

创新者，转型为农业 AI 生态的核心共建者与规则制定者。通过将自身的顶尖农业专业知识（Know-How）沉淀为标准化的模型能力，赋能更多行业伙伴，共同开拓智能育种的万亿蓝海。

## 7.应用实效：

### (1) 案例应用场景与目标用户

**水稻株型改良：**以中国农业科学院团队利用该系统对水稻种质资源进行株型改良为例，实现株高较常规品种降低约 25%，抗倒伏能力显著提升，同时保持产量不受影响。该案例验证了模型在单一性状定向改良上的可行性。

**基因编辑靶点智能设计：**模型可针对水稻等重要基因，预测编辑不同位点后基因功能的变化，智能推荐多个高潜力编辑靶点，将方案设计与评估工作周期从数月缩短至数周。

**目标用户定位：**明确面向农业研发和育种领域，特别是从事生物育种、基因功能研究的研究人员。

**内部用户：**核心用户为中国农业科学院生物技术研究所、作物科学研究所、北京畜牧兽医研究所等科研团队。

**外部用户：**项目目前处于推广阶段，正与各省市农业科学院、以及相关农牧业企业进行对接交流。

### (2) 解决行业痛点情况

**基因筛选准确性不足：**传统基因功能筛选依赖序列比对和人工设定参数阈值，易导致结果偏差。

**研发周期长：**从基因挖掘到功能验证、实验设计的完整链条依赖

人工经验，周期长达数月甚至数年。

**缺乏系统化评测体系：**农业科研领域此前缺乏对大模型性能的系统化评估标准。

### (3) 经济社会效益

通过本项目实施，能够显著缩短早期研发周期，降低研发试错成本，提升种质创新效率。

### (4) 示范推广价值及现状

构建了中国首个农业科研领域大模型系统评测体系，为行业提供参考框架。形成“AI 文献知识综述基因功能—AI 数据整合分析基因网络—AI 设计试验方案”的全流程数智化闭环，确立了智能育种的新范式。同时，依托双方共建的生态平台，将中国农业科学院的专业能力标准化输出，赋能全行业，加速推动我国种业科技能力的整体跃升。

## 七、可控核聚变

### （一）基于深度强化学习的球形托卡马克等离子体高精度磁控制

#### 1. 案例实施单位：

新奥能源研究院

#### 2. 案例背景：

磁约束可控核聚变是解决未来能源问题的关键路径之一。托卡马克装置中的等离子体是一个高温、高维、非线性的复杂系统，其稳定运行和高性能实现极度依赖于精确的磁控制。传统的控制方法，如传统比例-积分-微分（PID）控制方法，虽已广泛应用，但在处理多变量强耦合问题时面临设计复杂、调试周期长、难以适应新运行模式等挑战。这限制了探索新型、高性能等离子体位形（如球形托卡马克中的高参数位形）的效率，成为磁约束聚变研究领域亟待突破的技术瓶颈。

#### 3. 案例概述：

本案例提出并实现了针对球形托卡马克的等离子体智能控制系统。通过多保真数字孪生环境中进行离线训练，采用数据-机理融合建模的方式消减仿真环境与实验装置的误差，利用强化学习使智能体自主学习复杂的控制策略。在球形托卡马克 EXL-50U 中实现了对极向场线圈电源电压的毫秒级控制，实验成功达成了对等离子体电流和位置的稳定控制，在限制器和偏滤器位形下均实现了数百毫秒的稳定运行，是全球首个完成球形托卡马克智能控制的验证案例。

#### 4.行业痛点:

传统托卡马克磁控制面临三大痛点：**一是**设计调试极其复杂：传统 PID 控制器依赖线性模型，需要专家投入大量时间进行参数整定和前馈设计，每种新的等离子体位形都需重复此过程，迭代效率低下；**二是**控制性能存在瓶颈：等离子体是典型的多输入多输出（MIMO）强耦合系统，独立的单输入单输出（SISO）控制器难以实现最优协同控制，限制了对复杂位形的精确控制能力；**三是**探索新模式成本高：新物理现象和高性能运行模式的探索，受限于控制系统开发的滞后。开发一种能快速适应新目标、甚至自主发现更优控制策略的通用方法，是加速聚变研究的关键。

#### 5.案例详述:

##### (1) 总体思路

本案例结合多保真度数据-机理融合模型与深度强化学习控制策略，构建了一个端到端的等离子体智能控制系统。其核心思路是：首先，利用机理仿真和实验数据联合驱动，建立能够精确描述 EXL-50U 装置电磁响应和等离子体动力学行为的多保真数字孪生环境；其次，在该环境中，利用深度强化学习（DRL）算法训练一个神经网络控制器，使其学会从等离子体状态到线圈电压指令的非线性映射；然后，将训练好的控制器移植到真实的物理实验装置中，利用“少样本”（Few-shot）数据适应环境，执行实时控制任务；最后，在实际实验环境完成测试验证，并且对参数调优，助力于核聚变装置的科学实验。

##### (2) 平台介绍

**物理平台与控制目标：**实验平台为新奥自主研发的 EXL-50U 球形托卡马克。控制目标是实时调节 12 组独立的极向场 (PF) 线圈和中心螺线管 (CS) 线圈的电压, 以精确追踪预设的等离子体电流 ( $I_p$ ), 径向位置 (R) 和垂直位置 (Z) 以及等离子体形状 ( $a$ 、 $\kappa$ 、 $\delta$ 等) 的时间演化波形。

**“数字孪生”模拟环境：**为保证策略能够成功迁移至真实设备, 此外还需平衡与强化学习交互的仿真计算量, 我们构建了三种不同保真度的物理模拟器作为 DRL 的训练环境, 此外为了满足实验信号实时反馈 (亚毫秒级), 还发展了对传统反演方法的代理模型:

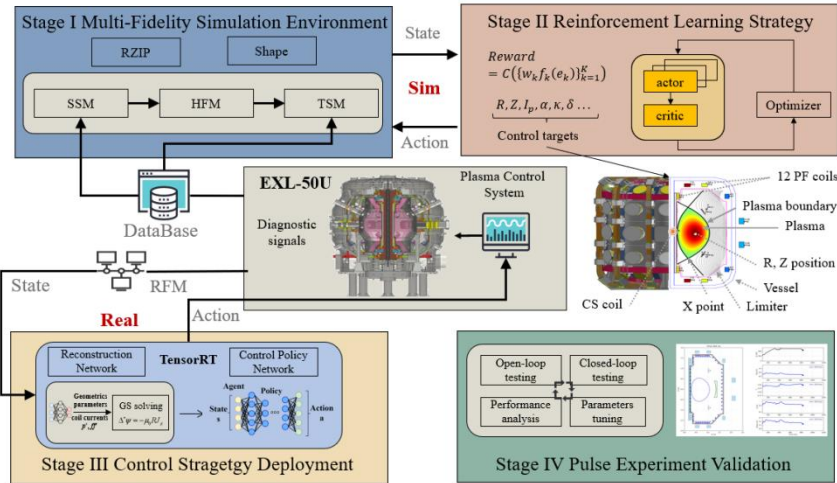
**状态空间模型 (SSM)：**一个基于线性化麦克斯韦方程组和刚性等离子体假设的快速响应模型。该模型计算速度快, 适用于大规模、并行的探索性训练, 帮助智能体快速学习基础的控制逻辑。

**高保真度模拟器 (HFM)：**一个求解非线性自由边界 Grad-Shafranov 平衡方程与电路方程耦合的模拟器。该模型能更精确地模拟等离子体的形变和动态演化, 用于对策略进行精细优化, 保证仿真稳定性的关键。

**时间序列预测模型 (TSM)：**一个基于完全数据驱动的深度学习模型, 进行, 可以通过与真实环境的自监督学习框架, 可以学习实验数据并修正高保真机理模型, 是成功实现“Sim-to-Real”迁移的关键。

**实时反演代理模型 (EFIT-mini)：**一种融合神经网络与物理模型的实时平衡反演算法。通过将神经网络嵌入传统 Grad-Shafranov 方程求解流程, 仅输出关键等离子体参数替代传统迭代计算, 在保持物理

自治性的同时实现 0.36 毫秒级高速反演。该算法兼具传统方法的物理精度与数据驱动方法的数值稳定性，为托卡马克等离子体实时位形控制提供可靠支撑。



来源：新奥能源研究院

图 79 整体技术方案示意图

### (3) 技术方案

**智能体 (Agent) 设计：状态 (Observation)：** 输入为等离子体的实时状态，包括等离子体电流  $I_p$ 、径向位置  $R$ 、垂直位置  $Z$ ，以及这些量的目标值。状态信息由装置的磁探针等诊断系统测量，并通过 EFIT-mini 快速计算得出。**算法：** 采用近端策略优化 (PPO) 算法。**动作 (Action)：** 输出为一个包含 13 个元素的向量，分别对应 12 组 PF 线圈和 CS 线圈电源的电压指令值。**奖励 (Reward)：** 设计了一个多目标奖励函数，主要惩罚当前状态 ( $I_p$ 、 $R$ 、 $Z$ 、 $a$ 、 $\kappa$ 、 $\delta$ 等) 与目标值之间的误差。通过映射加权聚合各分量，引导智能体优先改善表现最差的控制维度，从而实现全面、精确的追踪。

### (4) 关键指标参数

训练完成的神经网络控制器被编译为可在 EXL-50U 实时等离子体控制系统 (PCS) 中高效运行的程序。在实验中, 该控制器以 1kHz 的频率运行, 接收诊断数据并输出电压指令。

限制器位形放电 (#12204): 在 500kA 等离子体电流下实现了在平顶段 650 ms 的稳定闭环控制。与基准 PID 控制 (#12165) 相比, 径向位置 R 的控制标准差从 0.0052 米优化至 0.0036 米, 显示出更高的稳定性。

偏滤器位形放电 (#12662): 在更复杂的偏滤器位形下, 同样实现了超过 300 ms 的稳定控制, 验证了控制器对不同位形的适应能力。

控制器行为分析: 分析发现, DRL 控制器在不同实验中自主学习到了新颖的线圈电流分配策略, 例如在 #12204 中优先使用外侧 PF9/10 线圈进行控制, 而在 #12662 中则采用了不同的非均衡电流分布, 这展示了其超越人类直觉寻找最优解的潜力。

## 6. 创新情况:

### (1) 技术创新情况

基于机理-数据的物理模型+端到端深度强化学习控制: 本案例在国内首次将基于多保真度物理模拟器训练的 DRL 控制器成功迁移应用于真实托卡马克装置。与依赖历史数据的模型不同, 我们通过状态空间模型 SSM、高保真模型 HFM 等物理模型以及时序代理模型 TSM 和反演代理模型 EFIT-mini 构建数字孪生, 使智能体能够在符合物理原理的且能根据实验数据修正的仿真环境中探索和学习, 这极大地增强了控制策略的泛化能力和物理可解释性, 是连接人工智能算法

与聚变工程实践的关键技术突破。

**高维、非线性 MIMO 问题的智能解耦与协同控制：**传统方法通常将多变量控制问题分解为多个独立的单变量回路，难以处理线圈间的强耦合效应。本案例中的单一神经网络控制器，通过端到端的学习，隐式地学习到了 13 个执行单元之间的复杂耦合关系，实现了对整个系统的高效协同控制，无需人工设计解耦矩阵，代表了复杂工业和大型科学装置控制的一种新范式。

**成功的“Sim-to-Real”少样本迁移：**在经过少量真实装置数据微调的情况下，将在模拟器中训练的策略直接成功部署于 EXL-50U，是本案例的核心技术亮点。这得益于多保真模拟器的建立与自适应学习，包括对电源系统响应、诊断噪声等实际工程因素的建模，有效克服了模拟与现实之间的“鸿沟”，验证了该技术路线的可行性。

## (2) 模式创新情况

**聚变实验研发模式的变革与加速：**本案例展示了一种全新的“目标驱动”式控制器设计模式。研究人员只需定义高级的物理目标（如期望的等离子体形状和参数），而无需深入复杂的控制工程细节，人工智能模型即可自主学习形成一个高性能控制器。这极大地降低了新位形探索的门槛和时间成本，将传统以“月”为单位的控制器开发周期缩短至以“天”为单位的自动化训练周期，从而能够加速对先进聚变运行模式的探索和验证。

**发现超越人类经验的控制策略：**强化学习的探索机制使其能够发现非传统的、但可能更优的控制方案。如在实验中观察到的，DRL

控制器自主采用了与传统 PID 不同的线圈电流组合策略来实现控制目标。这种“黑盒”中的智慧为我们理解和优化等离子体控制提供了新的视角，有望突破基于人类先验知识的性能天花板。

**为未来聚变堆提供可扩展的智能控制解决方案：**随着国际热核聚变实验堆计划（ITER）和商业聚变堆（DEMO）的建设，其运行的复杂性和对可靠性的要求将远超现有装置。本案例所验证的 DRL 控制框架具备良好的可扩展性，能够通过增加状态输入和动作输出来适应更复杂的控制任务（如剖面控制、不稳定性抑制等），为未来聚变堆的“一键式”智能运行和自主控制奠定了技术基础。

## 7.应用实效：

### （1）案例应用场景与目标用户

本案例直接应用于 EXL-50U 球形托卡马克装置的等离子体磁控制实验。目标用户是托卡马克实验装置的物理研究人员和运行操作团队。该系统旨在替代或辅助传统的 PID 控制系统，用于日常的科学实验，特别是用于快速实现和稳定维持新型、高性能的等离子体位形。

### （2）解决行业痛点情况

**显著降低控制器设计复杂度：**通过自动化的训练流程取代了繁琐的人工参数整定，为 EXL-50U 快速实现多种实验目标（限制器、偏滤器）提供了敏捷的控制能力。

**提升复杂位形控制性能：**实验数据显示，在平顶段放电中，DRL 控制器相比 PID 在位置控制上实现了更高的稳定性（径向位置标准差降低了约 30%），证明了其在处理多变量耦合问题上的优势。

**加速新运行模式的探索：**该框架已成功用于控制 EXL-50U 的偏滤器位形，未来可快速拓展至 XPT、雪花偏滤器等更前沿的位形，大大缩短了从理论设计到实验验证的周期。

### (3) 经济社会效益

**经济效益：**直接效益体现在提升实验效率和装置运行的成功率上。通过缩短控制器开发时间，增加了有效放电炮次，加速了科研产出。长远来看，该技术是实现未来商业聚变电站低成本、高可靠性运行的关键技术之一，对于降低聚变能源的度电成本具有深远影响。

**社会效益：**核聚变被视为解决人类能源危机的“终极方案”。本案例作为 AI for Science 在这一领域的成功应用，展示了人工智能加速重大科学发现的巨大潜力。通过推动聚变研究的智能化进程，本项目为早日实现清洁、无限的聚变能源贡献了力量，对保障国家能源安全、应对气候变化具有重大的战略意义。

### (4) 示范推广价值及现状

本案例所建立的基于物理模型的 DRL 控制方法具有极强的通用性和可移植性。该技术框架不仅限于 EXL-50U，其核心思想和算法模块可推广至国内外其他托卡马克装置（如 EAST、HL-3、DIII-D 等）乃至未来的 ITER 和 DEMO。本项目在国内率先完成了从模拟训练到真实装置成功控制的完整闭环，为国内同行提供了宝贵的工程实践经验和技術示范。目前，我们正基于此框架，进一步开发针对更复杂形状（如大拉长、强非对称偏滤器）和剖面参数的集成控制，以期在更高参数下实现更全面的智能控制。

#### (5) 开源情况、相关标准成果产出等

本案例的核心算法、高保真度物理模型及训练数据集均可开源。相关研究成果已整理为学术论文，计划在国内外高水平聚变及人工智能顶级期刊发表，并积极参与国内外相关学术会议进行交流。

