

港理大大亚湾技术创新研究院 人工智能研究中心

人工智能研究中心

Research Center for Artificial Intelligence



**杨
红
霞**

- 中心主任
- 计算机与数学科学学院副院长 (环球事务)
- 电子计算学系教授



**程
然**

- 中心副主任
- 数据科学及人工智能学系副教授
- 校长青年学者



**李
明**

- 中心副主任
- 工业及系统工程学系助理教授



**张
成
奇**

- 香港理工大学深圳研究院院长
- 人工智能讲座教授



**刘
洋**

- 电子计算学系副教授
- 校长青年学者



**李
菁**

- 电子计算学系助理教授



**郑
昌
萌**

- 电子计算学系助理教授
- (研究)



**王
香
蒙**

- 电子计算学系助理教授
- (研究)



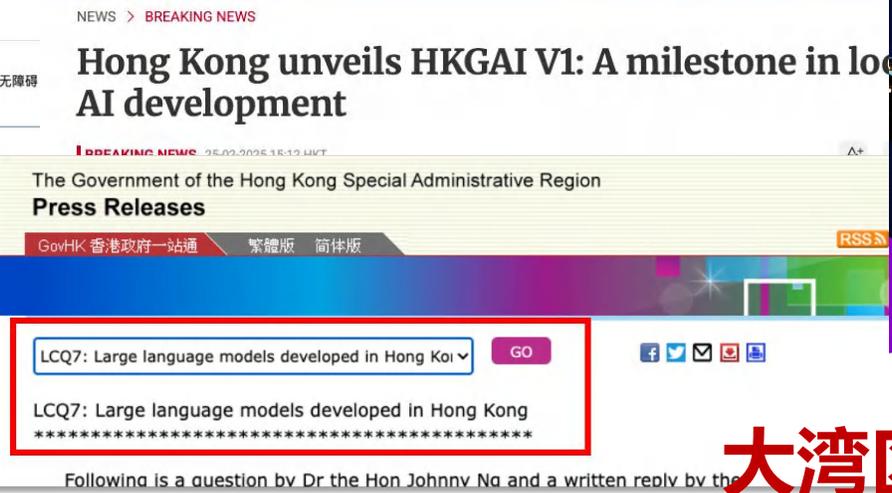
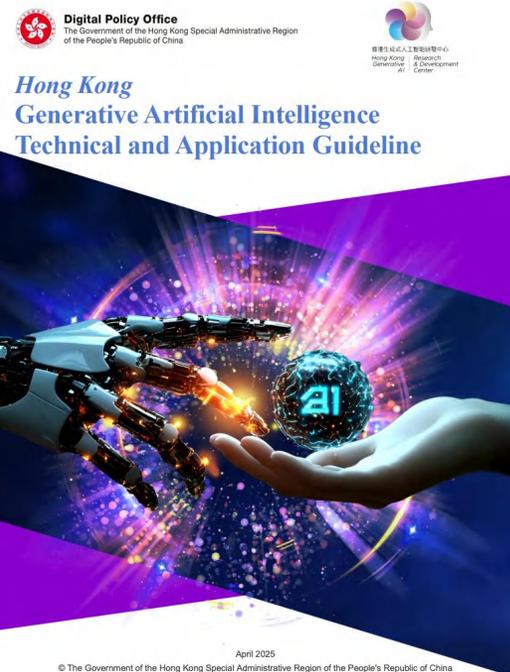
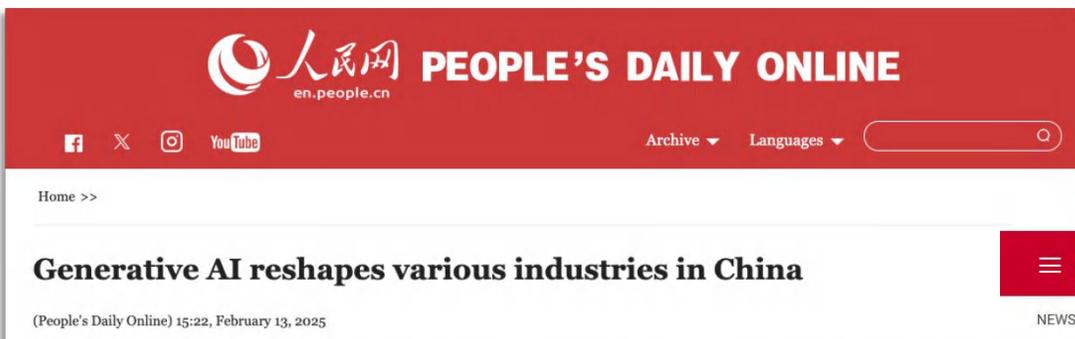
**邱
奕
宁**

- 电子计算学系助理教授
- (研究)



生成式AI: 现有的能力 vs. 涌现的应用

- **生成式AI (GenAI):** 通过提升各行业的生产力、创造力和决策能力，赋能内容创作、任务自动化和个性化体验。
 - 据Gartner预测，GenAI市场将大幅扩张，预计年增长率为46.47%，到2030年市场规模将达到3561亿美元



中国大陆

大湾区

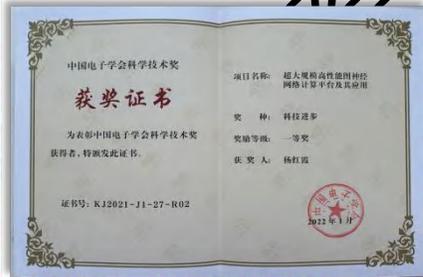
GenAI的新范式: Co-GenAI



- 当前的GenAI 范式在领域专家和 AI 专家之间造成了巨大的**差距**：
 - 人工智能专家缺乏协作和大量参与
 - 领域专家缺乏对应用的支持

基于集中式的GenAI	→	基于去中心化的GenAI
		
中心化 GPU 垄断		协作式GenAI
15B* 模型: 1-2 百万 GPU时		15B* 模型: 融合最优秀的4个模型只需要 160 GPU时
同类型高端芯片		高效利用多样化、分布式芯片

GenAI领域的杰出成就：先驱者



2021

• M6: 10T参数规模，世界上最大的预训练模型，仅在512个V100 GPU上进行训练

• 业界首个通用统一预训练模型（模态/任务/架构）OFA

• 与最先进的技术相比，LEGO将采样时间缩短了60%

• InfiGUIAgent-2B: 具有原生推理和反思功能的多模态通用GUI智能体

• 推出了与OpenAI的代码解释器可比的InfiAgent框架，发布了一个超高性能、规模较小的多模态模型（7B、13B）InfiMM。

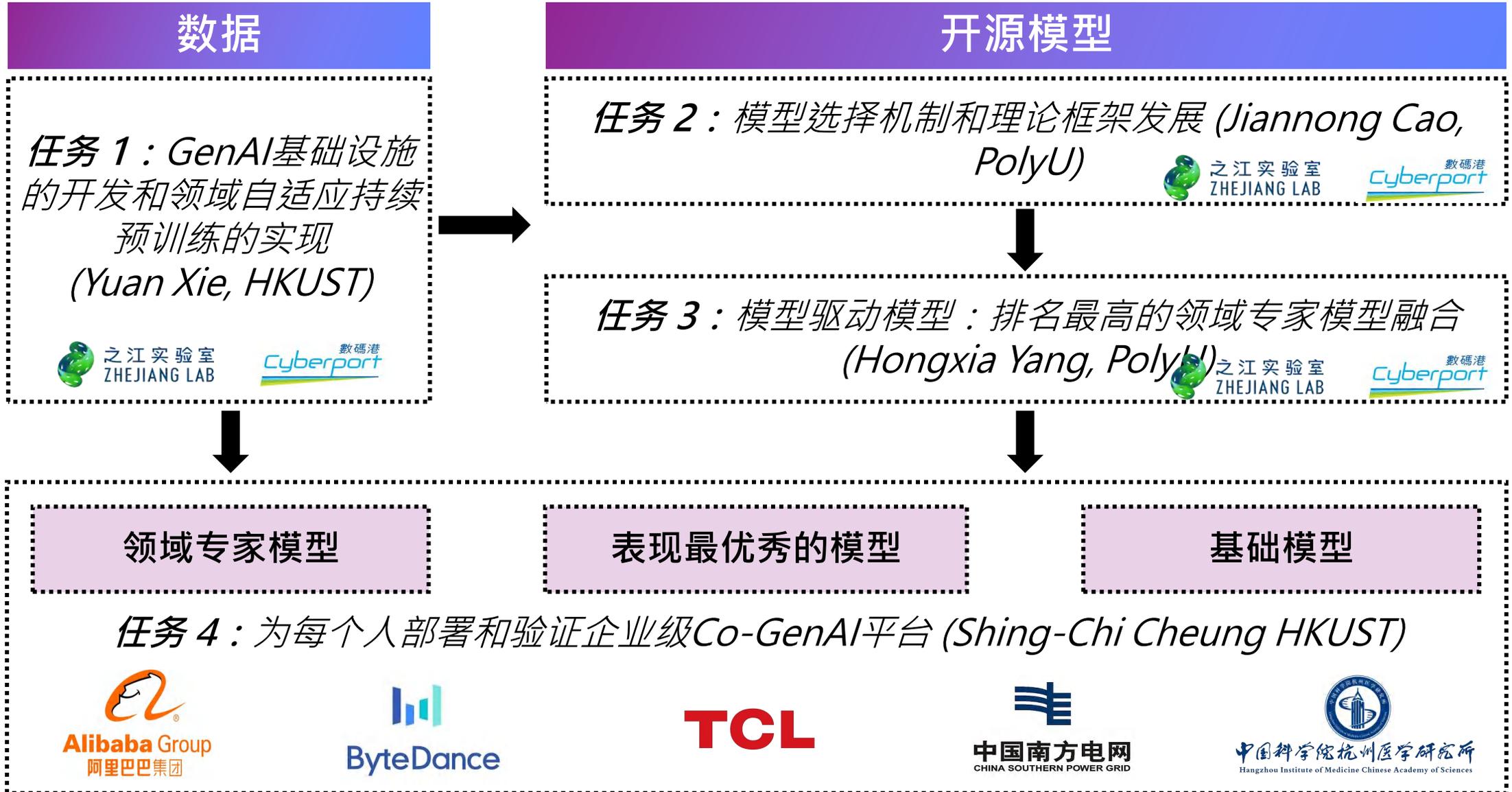
• InfiFusion: 涵盖7个基准测试和9个针对推理优化的LLM

• InfiR：在具有10亿至30亿个参数的LLM和MLLM模型的推理中实现顶级性能

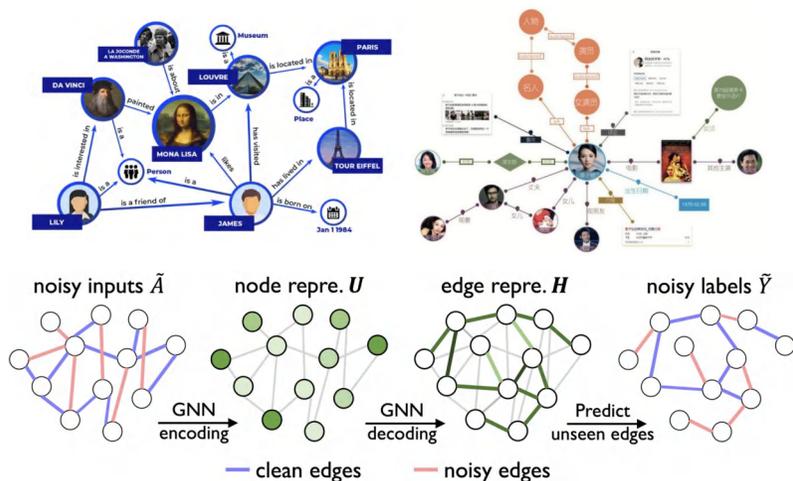
• InfiFusion：一个统一的模型融合框架，涵盖10个领域的融合，理论上可扩展到无限数量的领域

• 最新发布: InfiGFusion, InfiFPO, InfiGUI-R1, InfiJanice, InfiMed

高效能协作GenAI运算框架

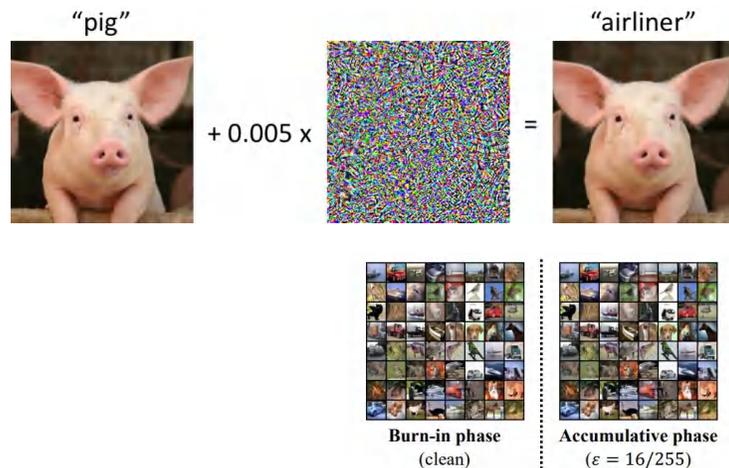


图数据含双边噪声



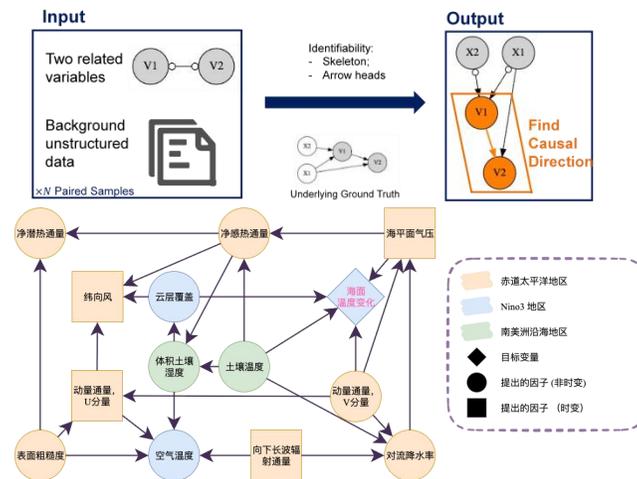
- 图数据中存在**双边噪声**，同时扰动图拓扑结构与目标边的标签。
- 提出一种**信息论驱动**的鲁棒图信息瓶颈框架，在 Cora/Citeseer 数据集的双边噪声环境下，将链路预测精度显著提升约**10%**。

图像数据注毒攻击



- 累积型注毒攻击**对机器学习模型构成严重威胁，其扰动难以察觉。
- 提出**基于记忆差异**的样本修正方法，有效抵御累积型注毒攻击，在 CIFAR-10 数据集上精度提升**48.81%**，在 SVHN 数据集上精度提升**40.10%**。

非结构化数据的因果发现



- 挖掘文本、图像等原始数据的因果关系，**局限于稀缺的高级变量**。
- 提出**结合大模型知识与统计检验**的因果表征框架，可信地揭示数据内在结构。对比 CoT，在基准数据集上的精度提升约**33%**。

发展规划：可信基础模型



- 开发抗噪声输入算法，提升模型鲁棒性。
- 建立对抗性提示防御机制，提升模型安全性。

- 优化领域知识覆盖，提升模型可靠性。
- 设计去偏训练框架，提升模型公平性。

- 应用于医疗诊断领域，降低误诊风险。
- 应用于公共服务领域，缓解社会分层问题。

2024

2025

2027

2028

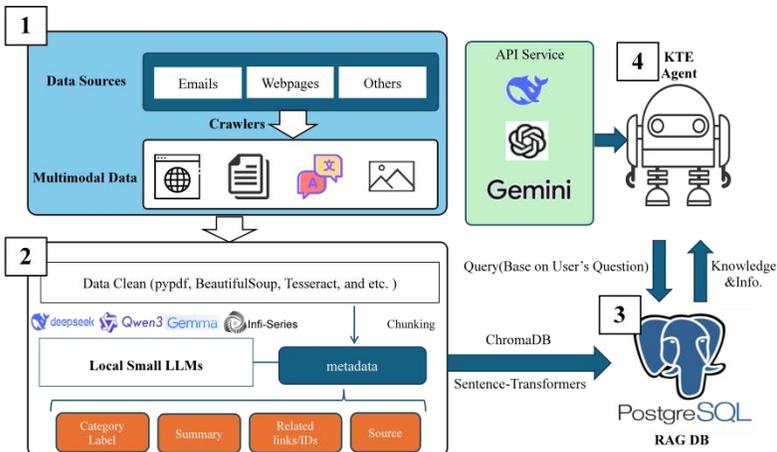
张成奇 教授



- 香港理工大学讲座教授
- 香港理工大学深圳研究院院长
- 2024年国际人工智能联合会 (IJCAI) 大会主席
- **可信基础模型**
 - 研究方向规划、建设团队
 - 学术指导
 - 合作与交流

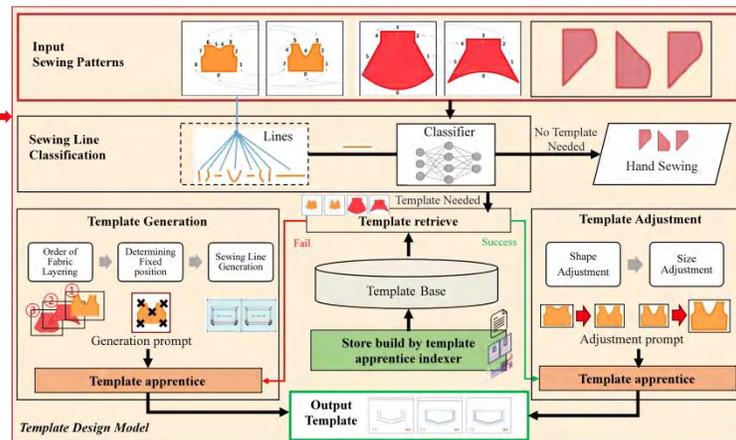
- 到2025年, 解决**鲁棒性与安全性**。
- 到2027年, 解决**可靠性与公平性**，从而构建可信基础模型。
- 到2028年，在**香港地区**，开发适用于**医学与社会科学**的可信基础模型，将其作为**社会基建的基石组件**，支撑**跨学科研究**，以应对全球挑战。

Agentic RAG



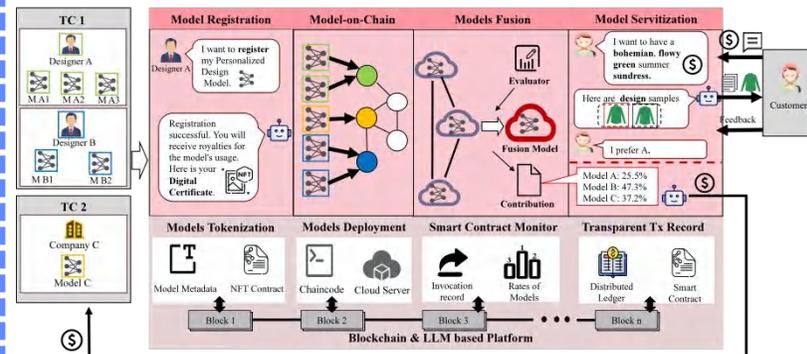
- 构建普适化的**企业级Agentic RAG**解决方案，提供以企业现有的数据为基础，以大模型为核心的知识库智能体系统，实现由企业级数据向专家经验知识的无缝迁移。

纺织模版AI助手



- 针对制衣生成的模板化过程，研发面向模版专家的**Agentic RAG**系统，基于大模型实现各类制衣模版检索、修正及增强生成，预计有效提升至少**50%**的模版专家设计效率。

领域模型资产化



- 基于区块链的领域模型资产化全链条解决方案，满足领域模型的资产确权、授权及转移等模型资产化需求
- 通过智能合约完全模型资产化过程的**100%**的全链条可视化及可追溯性。



发展规划：基于区块链的Agentic AI资产化平台

- 构建普适化的企业级 Agentic RAG解决方案。
- 以当地的典型产业为突破口，建议该方案的示范性效应。

- 由点及面，覆盖更多的场景领域，力争实现30+基于 Agentic RAG形成的垂直领域AI助手。

- 基于领域AI模型，构建模型的资产化平台，形成由数据要素流通到模型要素流通的核心转型。

2025

2027

2028

2030

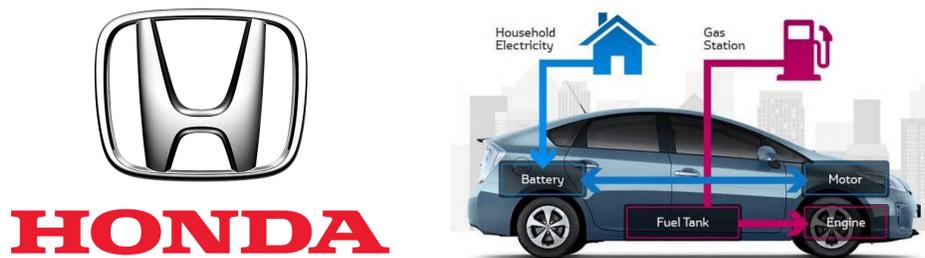


李明 教授

- 香港理工大学人工智能高等研究院副院长
- 工业及系统工程学系助理教授
- 港理大大亚湾创新研究院人工智能研究中心副主任

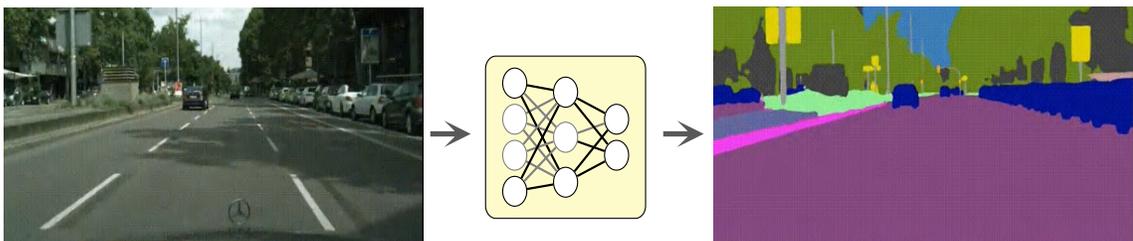
- 到2028年，建成**面向5+产业，30+细分场景的**领域智能体，有效提升各场景工作效率50%以上。
- 到2030年，建成**面向100+领域智能体的**模型资产化平台，有效支撑个人、企业级用户的模型资产化流通需求。

新能源车控制策略优化



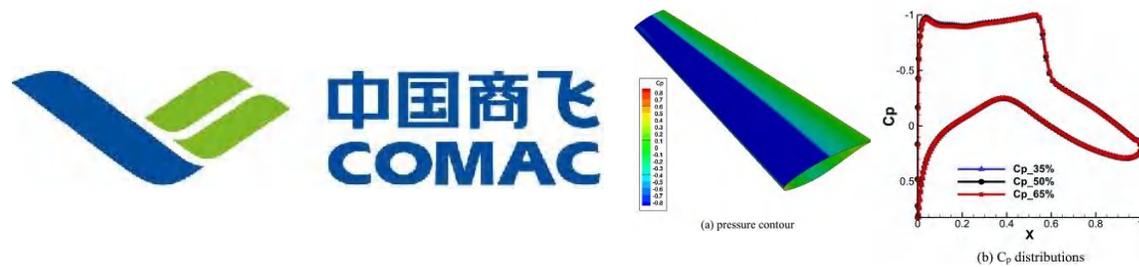
- 采用纯仿真环境结合高效的演化搜索算法优化新能源车控制策略，**油耗降低5%，碳排放减少62%**，优化后的控制策略**已成功集成至本田的官方软件平台**。

昇腾AI芯片模型部署与性能优化



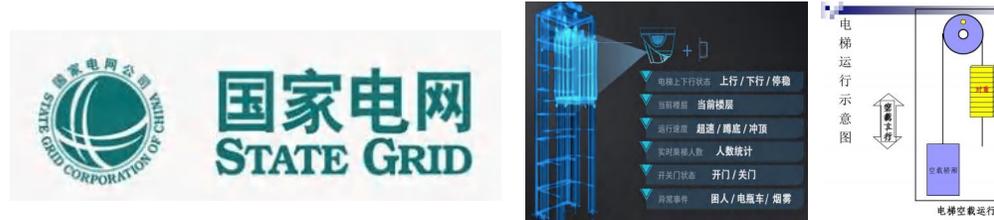
- 运用进化搜索算法优化模型结构与参数，结合硬件友好性预测模型，确保部署效率。
- 模型精度与NVIDIA平台持平，**目标识别速度提升30%**，实际部署**功耗低至9.5W**，展现极致能效比。

飞行器翼型空气动力学设计优化



- 基于小样本数据，融合物理启发知识，通过演化建模技术构建代理模型。**无需昂贵CFD仿真**，即可高效优化设计并**显著提升翼型气动性能**，整体效率远超传统方法。

关键设备预测性维护与安全预警



- 整合多源传感器数据。采用人机融合的建模方法，结合专家经验与AI学习能力。
- 实现对潜在**故障的提前预警**与精准的**智能故障诊断**；保障设备运行的**高安全性**与系统响应的**高实时性**。



发展规划：大小模型协同的工业智能体

规则融合·自适应建模

- 融合专家规则与数据学习，构建可解释的智能体模型。
- 实现小样本驱动的自适应建模，落地典型工艺生产环节。

多目标优化·在线协同

- 构建多目标自优化系统，智能权衡效率、能耗与稳定性。
- 实现多智能体协同部署与经验迁移，支持复杂任务动态控制。

自主演化·泛化部署

- 形成自主演化能力，根据现场环境持续学习与重构。
- 推动泛化部署与标准接口建设，构建跨行业的智能体应用生态。

2025

2027

2028

2030

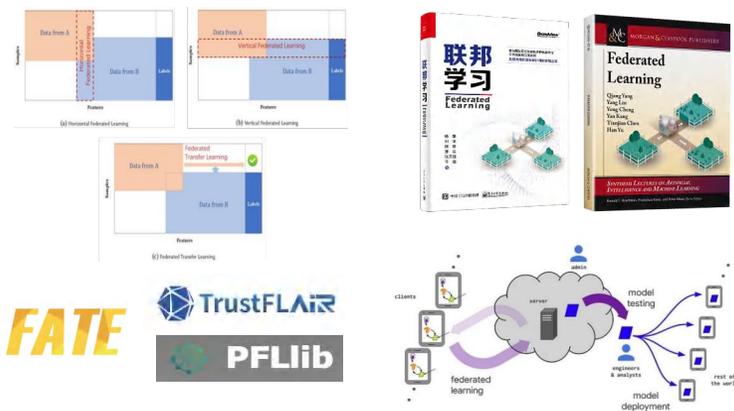


程然 教授

- “斯坦福”全球前2%顶尖科学家
- “科睿唯安”全球高被引科学家
- 中国智能计算科技创新人物
- 算力中国·青年先锋人物
- **研究方向**
 - 计算智能
 - 演化计算
 - 表征学习

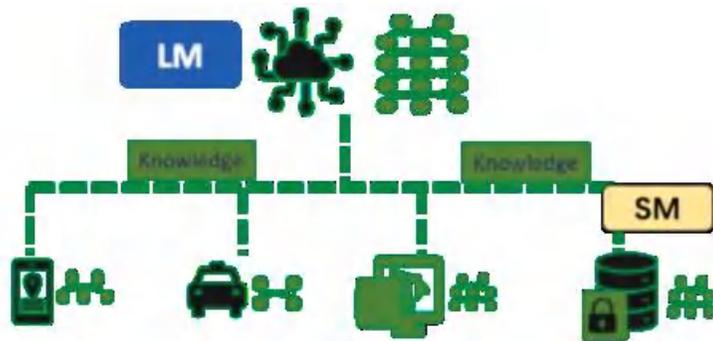
- 到2027年，在**3-5条典型工艺产线**中部署智能体系统，完成**可解释模型+专家规则融合**试点验证。
- 到2030年，形成支持**50+节点部署**的智能体体系，支撑跨场景泛化与智能升级，推动**工业智能体生态初步成型**。

联邦学习



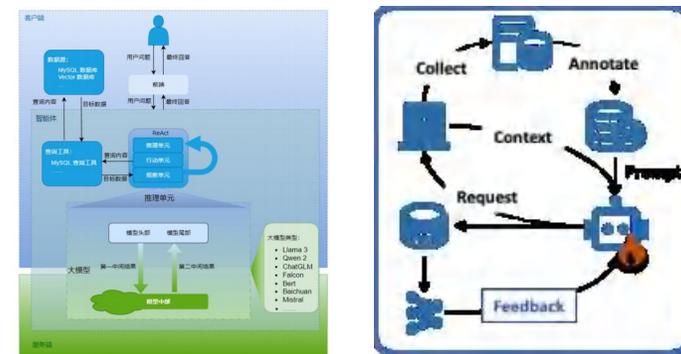
- 开展联邦学习基础理论、算法、框架与应用研究，包括横向、纵向联邦学习、联邦迁移学习等。
- 实现**数据不动模型动，数据可用不可见**的协同安全AI新范式。

大小模型协同



- 研究大-小模型协同方法与框架，解决大模型应用中的私有数据安全与计算资源受限问题。
- 实现**大小模型协同演进**，领域任务模型性能与安全性大幅提升。

多智能体协同技术



- 构建多个大模型智能体的知识融合与协同推理技术框架。
- 面向医疗健康、金融等应用领域解决数据稀缺、数据质量差、分布偏移等关键问题。

发展规划：面向多智能体的可信增强与协同框架



- 基于联邦学习等关键技术，提出面向多智能体的可信增强方法，优化可信协同方案，提升协同效能和数据安全。

- 研发开源的面向多智能体的可信增强与协同系统。
- 示范应用于医疗与金融等领域，进一步迭代优化。

- 在医疗与金融等领域取得实际应用效果。
- 拓展到其他应用场景，持续跨学科研究与应用。

2025

2026

2027

2028



刘洋 教授

- 香港理工大学副教授
- 香港理工大学数据科学与人工智能系
- Presidential Young Scholar
- Google Scholar引用25,000+次
- 《麻省理工科技评论》2022中国隐私计算科技创新人物

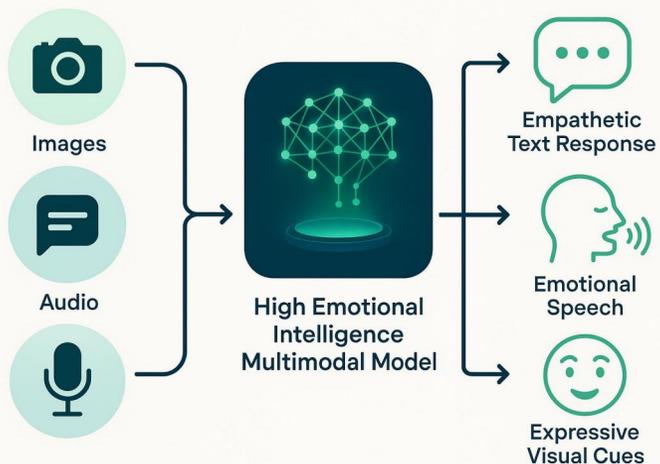
- 到2026年, 提出理论框架和方法，解决多智能体在资源受限、数据敏感的领域应用瓶颈。
- 到2027年, 发布开源框架。
- 到2028年，应用于至少2个不同领域，促进AI+跨学科研究与应用。

基座模型

高层决策

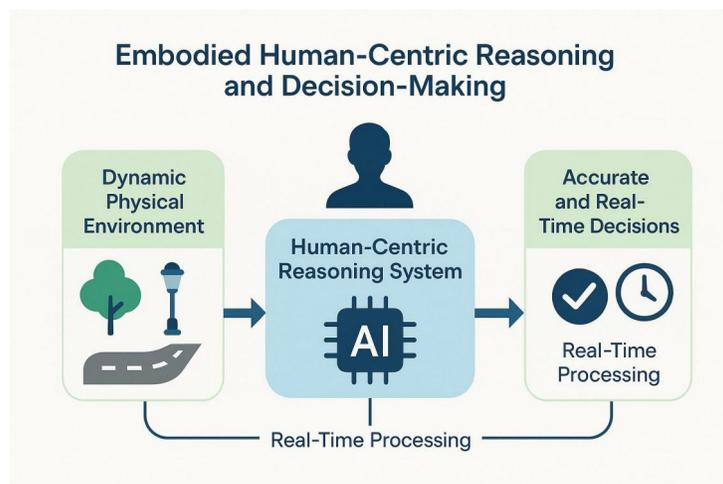
底层行为

高情商的多模态基座模型



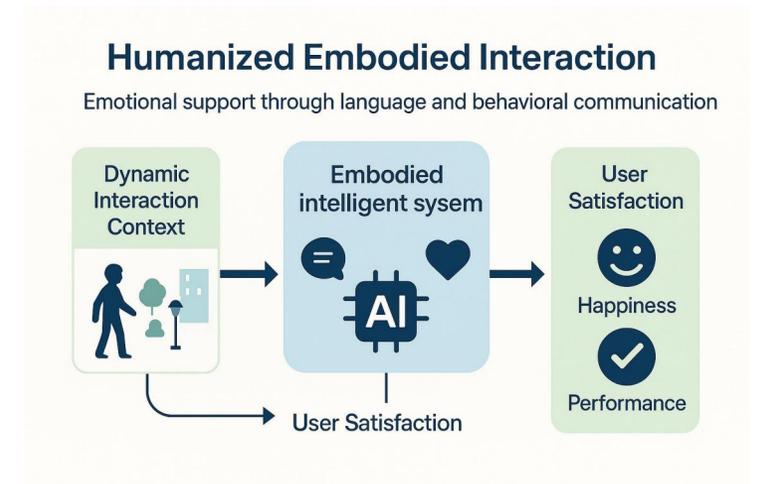
- 增强基座模型从图、文、音等多模态输入中**挖掘人类需求**的能力，并提升其**高情商反馈水平**。
- 在**多种情商任务**中比通用多模态基座模型的性能提高**5%-10%**。

基于动态环境的人本决策



- 提高具身智能系统在复杂动态环境中始终**以人为本**、做出满足人类需求的决策能力。
- 兼顾决策合理性与实时性：在**延迟1秒**内，实现**高于90%**的决策准确率。

人性化具身交互



- 提升具身智能系统在人机语言与行为交互中的情感支持能力，实现人性化交互。
- 构建模拟人机交互平台，实现**90%**用户满意度；在真实人机交互中达到**85%**满意度。



发展规划：以人为本的具身智能

- 开发高情商多模态基座模型，在多个情感反馈任务上较通用模型性能提高 5%-10%。
- 构建人本具身决策系统雏形，基准测试准确率突破 90%。

- 优化人本决策系统至 ≤ 1 秒响应延迟，维持 $\geq 90\%$ 决策准确率。
- 搭建模拟交互平台，完成 10 大类日常交互场景（如家庭/办公/医疗）的闭环测试。

- 达成人性化交互的用户满意度目标（模拟交互90%；真实交互85%）。
- 在养老陪护、助盲服务两大应用领域实现以人为本的具身智能系统在机器人的解决方案落地。

2024

2027

2028

2030

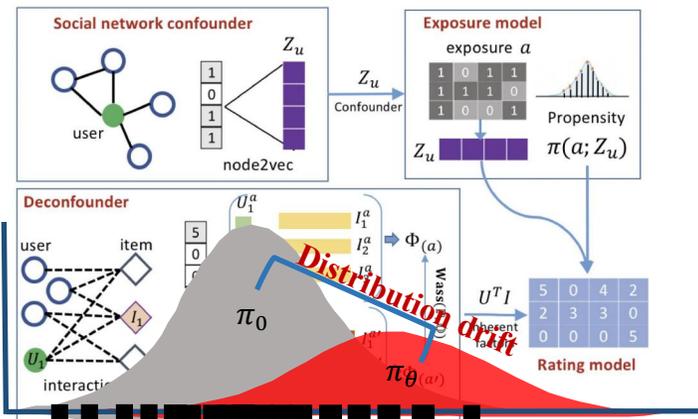
李菁 教授



- 香港理工大学助理教授
- 前腾讯AI Lab高级研究员
- 代表作AngleE语义模型在HuggingFace百万月下载
- **以人为本的具身智能研究**
 - 研究方向规划
 - 学术指导
 - 合作与交流

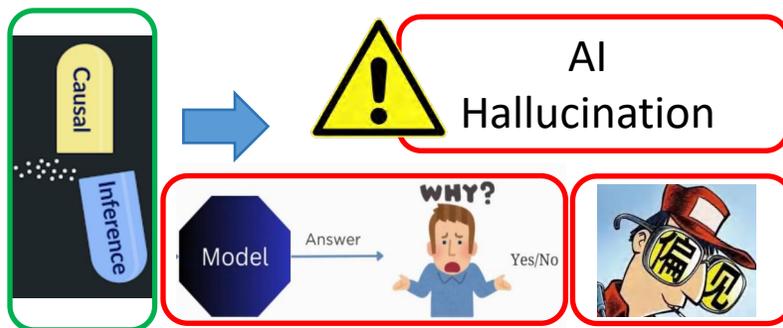
- 到2028年，奠定多模态基座模型在**多种情商能力**中的行业领先水平，并实现**准确率超过90%的毫秒级**的人本决策系统。
- 到2030年，在养老陪护和助盲服务领域完成研究在机器人的落地应用，确保真实交互中的**用户满意度稳定超过85%**。

数据混淆偏差、分布偏移去偏方法



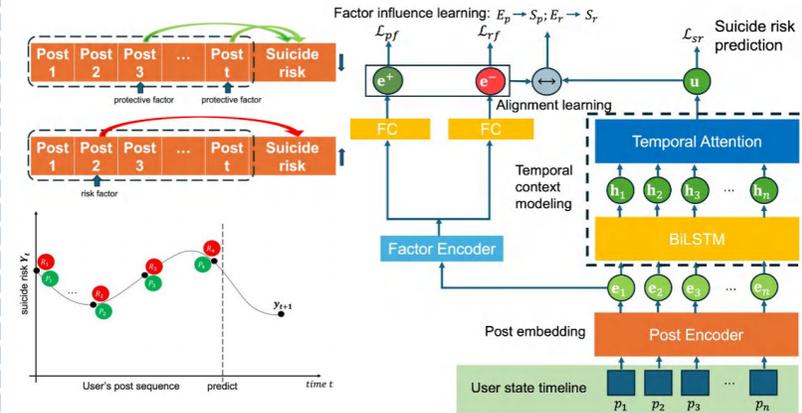
- 揭示行为建模中由于**混淆变量**引入的因果偏差。
- 提出一种**去混淆与分布校准**方法，结合因果图、逆概率加权 (IPW) 以提高模型在非理想训练数据下**泛化能力**。

因果推理启发的可信人工智能模型



- 将因果推理 (causal inference) 融入深度学习中，以纠正模型在**用户行为建模**、**推荐**与**情感分析**中的偏差来源。
- 提出了多个因果增强模型架构，提升模型对因果关系的建模能力，从而增强其**鲁棒性**、**公平性**与**可解释性**。

社交媒体用户自杀风险识别方法



- 聚焦于社交媒体中**自杀倾向用户**的识别问题。
- 提出了多模态融合的风险检测框架，综合**文本**、**行为轨迹**和**社交网络结构**信息以提升检测准确性，在多个真实社交媒体数据集上表现出色。

发展规划：可信心理健康大语言模型



- 利用因果辨识理论揭示多领域数据中蕴含的因果结构。
- 构建大规模多领域因果关系图数据库。

- 实现因果知识与人工智能模型的融合。
- 提升模型在多领域上的解释性，公平性及鲁棒性。

- 应用于心理健康监测，自杀预警等领域，开发原型系统。
- 实现对心理健康问题评估、干预、跟踪的全流程闭环管理。

2024

2027

2028

2030

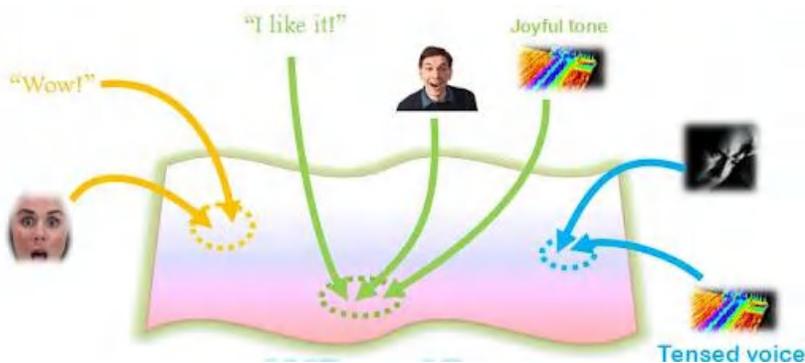


王香蒙 教授

- 香港理工大学助理教授（研究）
- 因果推理、可信人工智能专家
- **因果推理启智的人工智能模型**
 - 研究方向规划、建设团队
 - 学术指导
 - 合作与交流

- 到2028年，构建具备因果推理机制的可信人工智能模型，突破当前模型在**解释性不足、公平性差与可信度不高**等瓶颈。
- 到2030年，进行项目落地，特别是针对**青少年群体**的**心理健康监测与自杀预警**，构建基于AI的**评估-干预-追踪**系统，提升对高风险人群的社会保障与福利支持。

跨模态数据建模



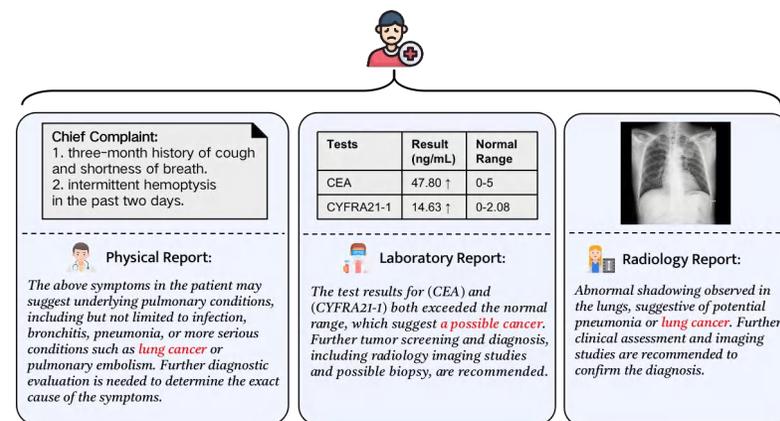
- 针对跨模态数据缺失问题，利用生成模型在预训练阶段获取的潜在知识，突破**多模态语义一致性控制技术**。
- 自动生成模态对齐数据，重点标注工业设备监测、医疗诊断等**复杂场景的细粒度对齐样本**。

可信知识融合



- 研究基于图结构的多模态知识表征框架，实现**文本、图像、时序等数据的语义对齐**。
- 设计动态知识更新机制，基于强化学习实现知识图谱节点的**自适应扩展与实时更新**，并与多模态数据流协同进化。

多智能体协同推理



- 构建分层多智能体框架，通过任务分解与动态角色分配机制，实现**复杂场景的高效协同**。
- 设计基于双重反馈循环的自我纠错机制，生成具有**可解释性的推理链**。

发展规划：突破认知智能瓶颈，推动智能化升级



- 结合行业需求，制定覆盖数据生成、知识融合、模型协同的技术攻关标准。

- 构建基于知识的多智能体多模态推理框架，优化协作效率，实现跨场景知识迁移。

- 在工业与医疗领域搭建多模态验证平台。
- 在科学推理、工业控制等任务中形成可复制的行业解决方案。

2024

2027

2028

2030



郑昌萌 教授

- 香港理工大学研究助理教授
- 多模态、自然语言处理专家
- 智枢科技联合创始人
- 多模态大语言模型**
 - 研究方向规划
 - 学术指导
 - 合作与交流

- 到2027年，构建跨模态**数据生成-知识融合-协同推理**一体化技术体系，突破动态对齐、知识演化、多智能体协同三大核心瓶颈。
- 到2030年，发布**全球首个**覆盖工业、医疗场景的多模态对齐评估数据集与**多模态多智能体底座平台**。