

附件 1

《深入实施“人工智能+”行动——GPT-5 与企业级可控智能体系统构建与应用》高级实训讲座简章

一、实训时间和方式

时间：2025 年 9 月 26 日至 9 月 28 日（周五、周六、周日）

方式：线上直播

二、实训对象

涉及人工智能及大模型技术厂商、企业级 AI 解决方案与服务提供商、云计算与大数据平台、智能体框架与协议开发商、分布式系统与计算技术企业、互联网企业、电信/广电运营商及学术研究、科研院所、高等院校技术团队、AI 实验室成员、政务及事业单位。来自重点行业：

企业 AI 工程师 / 系统架构师 / 智能体开发者 —— 需要掌握 GPT-5 架构原理与 LangGraph、MCP、A2A 等核心协议，构建可控、可演化的企业级智能体系统。

人工智能+行业落地团队（金融、制造、运营、教育、政务、能源、医疗等） —— 希望将大模型推理与多 Agent 编排快速应用到业务自动化、决策优化、生产力提升。

投资机构技术负责人 / 战略顾问 / 创新孵化经理 —— 需要理解 GPT-5 及智能体闭环技术的商业化路径，评估项目可行性与投资回报。

政策落地推动方 / 智能化转型项目负责人 —— 希望在“人工智能+”国家战略下快速构建符合数据治理、安全合规与产业升级要求的 AI 系统。

博士生 / 博士后 / 海归研发人员 —— 聚焦 RLHF、DPO、GRPO、AZR 等前沿训练路径，准备申请欧盟 Horizon Europe、国家自然科学基金、重大专项等科研项目。

CIO / CTO / 创业团队核心技术合伙人 —— 需要构建 0-1 的全栈智能体产品原型并具备企业级部署与交付能力。

AI 安全与合规专家 / 数据治理负责人 —— 希望掌握多源上下文调度、策略边界控制（Policy Bounding）与全链路可观测性，实现安全可控的 GPT-5 智能体系统。

三、研修实训大纲

- ▶ 模块 一： "人工智能+" 国家战略与生产力变革
- ▶ 模块 二： 技术突围——政策驱动 AGI 路径创新
- ▶ 模块 三： 产业竞合——分层赋能抢占战略新赛道
- ▶ 模块 四： 场景革命——“AI+” 产业落地路径实践及案例
- ▶ 模块 五： 未来攻坚——AGI 融合与治理协同
- ▶ 模块 六： GPT-5 统一推理架构设计与高效开发核心技术详解
- ▶ 模块 七： LangGraph 多状态流程控制与 Agent 编排引擎设计
- ▶ 模块 八： MCP 协议全生命周期机制与分布式智能体上下文统一架构设计
- ▶ 模块 九： A2A 多智能体协同通信机制与 Agent 路由调度策略
- ▶ 模块 十： AG-UI 可视化建模系统与图形化 Agent 流程编排
- ▶ 模块十一： Agent 任务建模语言 DSL 设计原则与解释器构建
- ▶ 模块十二： Human-in-the-loop 人类反馈闭环与人工审核接入设计
- ▶ 模块十三： RLHF 完整对齐流程与偏好建模实战
- ▶ 模块十四： Constitutional AI 原则驱动的价值对齐方法
- ▶ 模块十五： PPO 强化学习算法与技术实现
- ▶ 模块十六： 基于 GRPO 驱动的下一代推理调优范式技术
- ▶ 模块十七： Clip-Higher 策略、动态样本和 Token-Level 策略 loss 结构
- ▶ 模块十八： TTRL - 测试时强化学习与多解投票奖励机制
- ▶ 模块十九： AZR - 零监督奖励 × 自演化推理智能体技术
- ▶ 模块二十： RL-AZR 推理评估与端到端测试策略全流程实现

四、实训内容

模块名称	主要内容
模块 一	<ul style="list-style-type: none"> ▶ 1. 工业革命视角下的 AI 定位。 2. 全球 AI 竞争格局（中/美/欧/日政策对比）。 3. AI 作为生产要素的变革价值。 4. 中国 AI 战略规划（2017-2025 年政策演进）。 5. AI+行动加速数据-模型-应用闭环（可规模化/可治理/可商用）。



1. 技术发展路径：机器学习→深度学习→生成式智能。
2. Scaling Law 挑战与计算范式转变。
3. 多模态大模型架构演进（图文音→全模态→智能体）。
4. 自监督学习机制与预测创新。
5. 涌现能力：自我反思/多步验证/长程思考。
6. 硬件支撑体系（算力-存力-运力基础设施）。

模块 二



1. 海外格局：OpenAI-DeepMind 双龙头 vs Meta 开源生态。
2. 中国模型梯队（互联网企业/创业公司/科研机构）。
3. DeepSeek 关键突破（V2 价格战/V3 开源/Janus-Pro 多模态）。
4. 紫东太初技术路线（1.0 单模态→3.0 智能体大模型）。
5. 模型性能对标（语言/视觉/推理能力对比 GPT-4o）。
6. 低成本训练范式革新。

模块 三



1. AI+场景的工业化路径。
2. AI+场景落地关键问题。
3. AI+场景的系统架构。
4. 办公革命：ChatGPT+Office 自动化/Copilot 全流程设计。
5. 工业制造：焊接质检大模型/机床预测性维护。
3. 智慧交通：BEV-Transformer 智驾系统/无人机巡检。
6. 医疗健康：心血管多模态诊断/手术器械识别。
7. 社会治理：AI 政务助手/反诈模型/联勤指挥。
8. 低空经济：空域管理/立体交通/应急消防。
9. 科学研究：AlphaFold 范式/材料发现/芯片设计。

模块 四



1. 生成式 AI 缺陷：事实性错误/价值观偏见/数据安全。
2. 具身智能新范式（机器人任务分解与环境交互）。
3. 第五科研范式：AI 驱动的科学猜想（A4Science）。
4. AGI 融合路径：类脑智能+博弈智能+信息智能。
5. 认知表征突破方向。

模块 五

▶ **GPT-5 统一推理架构：剖析 GPT-5 的推理链路、插件机制、上下文扩展策略，实操部署到企业级智能体架构。**

1. 统一架构设计：融合轻量响应、深度推理、实时路由，兼顾速度与精度。
2. 自适应思考强度：基于任务复杂度与 think hard 信号动态分配推理资源。
3. 多模态推理整合：统一处理文本、图像、视频、图表与空间理解任务。
4. 可控推理力度：通过 reasoning_effort 与 verbosity 精准控制质速与信息密度。
5. 端到端决策闭环：规划 → 工具调用（串/并行）→ 校验 → 回退 → 交付全链路。
6. 稳健工具编排：支持串并行、超时重试、回退与错误分级，提升稳定性。
7. 长上下文处理：272k 输入 + 128k 输出，切分、滑窗与持久上下文复用。
8. 结构化输出生成：支持 JSON / 文法约束，保证可解析与易集成。
9. 长文检索优化：召回 → 去重 → 排序 → 证据拼接与冲突消解。
10. 安全与事实校验：分级回答、幻觉抑制与二次验证机制。
11. RLT 优先对齐策略：以解释质量为奖励信号，减少大模型与重标注依赖。
12. 代码闭环执行：计划 → 生成/编辑 → 测试 → diff/patch 提交的自动化流程。
13. 性能与成本优化：Prompt 缓存、批处理与分层路由提升吞吐效率。

模块 六

14. 自定义工具扩展: plaintext + regex/CFG 实现稳健解析, 规避 JSON 脆弱性。

► **LangGraph 智能体编排引擎: 掌握多状态图、Agent 节点流转、任务跳转与异常控制核心机制。**

1. 掌握 LangGraph 状态机的有向图建模方式与节点控制流机制。
2. 构建支持分支、循环与异常处理的复杂任务状态图。
3. 实现状态之间数据依赖的参数绑定与上下文传递机制。
4. 集成 LLM 工具节点与条件判断节点, 实现 Agent 能力组合。
5. 利用 LCEL 表达式自定义流程逻辑与状态跳转规则。
6. 支持异步执行、多路径调度与并行任务处理。
7. 实现状态执行日志记录、故障重试与容错回滚机制。
8. 构建可复用的子流程模板与嵌套子图结构。
9. 实现用户输入到输出全链路的状态可视化与追踪。

► **MCP 协议上下文调度: 实现分布式多智能体间上下文统一、权限控制、资源管理。**

1. 理解 MCP 的七阶段上下文生命周期流程与全链路控制路径。
2. 拆解 MCP 三大组件 Host、Client、Server 的角色与职责关系。
3. 掌握模型上下文协议在 LangGraph、Langflow 等系统中的深度集成方式。
4. 分析 Model-controlled Tools 与 AI 托管资源的功能调用流程。
5. 理解 MCP 中的 tool annotations 标注机制与函数参数约定规范。
6. 掌握 Client 与 Server 之间长连接、短连接与 Streaming 的通信机制演进。
7. 理解 Namespacing 如何构建逻辑分组与上下文隔离作用域。
8. 掌握 MCP Inspector 工具的使用, 追踪智能体上下文生命周期变化。
9. 理解 MCP 如何实现用户控制 Prompt、应用控制 Resource 的权限模型。

► **A2A 智能体协同机制: 构建 Planner、Coder、Critic 多角色 Agent 体系, 实现任务级协作。**

1. 设计多角色 Agent 协作模型: Planner、Researcher、Coder、Critic 等。
2. 实现 Agent 之间通过消息传递队列或共享状态图通信。
3. 配置基于意图、任务类型的 AgentRouter 路由机制。
4. 构建串行、并行、条件分支等多种协同执行策略。
5. 支持输出投票、融合策略、异步等待与优先级重调机制。
6. Agent 之间共享上下文缓存并保持语义一致性 (基于 MCP)。
7. 跨 Agent 调用工具并进行信息回传与中间结果协商。
8. 支持异构模型 Agent 协作 (GPT、Claude、CommandR 等)。
9. 构建 Agent 健康检查、心跳检测与任务容错恢复逻辑。
10. 结合 LangGraph 构建多 Agent 协同的任务流水线图。

► **AG-UI 图形建模系统: 零代码可视化任务建模, 自动生成 Agent 流程结构。**

1. 使用 AG-UI 创建流程节点、Agent 配置、工具绑定的可视化界面。
2. 实现流程图拖拽式建模, 节点属性可编辑与参数注入。
3. 构建任务状态图与实际 LangGraph 节点之间的映射机制。
4. 设计运行日志可视化与执行轨迹热区高亮回放功能。
5. 实现用户权限区分、协作建模与操作审计功能。
6. 支持基于 MCP 上下文结构的输入输出字段可视化。
7. 集成多 Agent 协同流程并提供 YAML/JSON 导入导出。
8. 支持从 AG-UI 动态热更新 LangGraph 执行逻辑。
9. 提供模型调用结果的可视化展示与工具输出渲染视图。

模块 七

模块 八

模块 九

模块 十

10. 接入 HITL 审核流程，支持图形化人工干预接入点设定。

► **Agent DSL 任务语言设计：基于业务场景自定义任务语言，映射到 LangGraph 执行。**

1. 理解 DSL 的定义、应用场景与业务价值（保险、客服、审批等）。
2. 使用 ANTLR/Lark 定义语法规则并生成解析器。
3. 构建基于 AST 的解释器：包括状态控制器、条件判别器与工具调用器。
4. 支持流程编排语言中的变量传递、上下文引用与嵌套执行。
5. 将 DSL 与 LangChain/CrewAI 执行引擎集成，实现任务调用。
6. 实现 DSL 热更新、运行时参数注入与版本控制机制。
7. 提供 DSL → JSON/YAML 转换支持图形化呈现。
8. 支持 Agent 路由、任务描述、执行权限等结构建模。
9. 将 DSL 映射到 LangGraph 节点，实现低代码任务控制。
10. 支持 DSL 的可视化开发工具联动（AG-UI）。

► **HITL 人类反馈闭环系统：接入审核机制，实现偏好采集、强化反馈与策略调优。**

1. 构建人工审核界面与任务触发机制。
2. 支持模型输出异常检测自动触发 HITL 审核节点。
3. 构建人工反馈 → 奖励建模 → RL 更新的闭环机制。
4. 提供标注质量控制工具：审核、重审、仲裁、共识。
5. 将人工反馈作为偏好信号用于 RLHF 强化对齐。
6. 多人协同审核支持角色分工与任务流调度。
7. 提供审核数据的结构化输出并进入知识库。
8. 支持用户行为数据作为奖励建模的弱监督信号。
9. 构建人机共创机制：人工修改模型生成结果。
10. 支持 HITL 在 LangGraph 中的异步挂起与流程中断恢复。

► **RLHF 三阶段训练系统：涵盖 SFT、RewardModel、PPO，完整构建偏好驱动对齐流程。**

1. 掌握 RLHF 三阶段流程：SFT → RM → PPO。
2. 构建人工标注的偏好对比数据集（pairwise, ranking）。
3. 训练 RewardModel 对输出结果进行评分建模。
4. PPO 阶段使用 KL 惩罚避免过度偏移初始策略。
5. 构建策略迭代训练流程与经验回放机制。
6. 支持多批次更新、Early Stopping 与奖励归一化。
7. 提供打分数据的可视化与 Reward 变化趋势监控。
8. 接入 LangGraph 实现 RLHF 在流程图中的闭环强化。
9. 实现 Preference Mining：从用户行为中挖掘偏好信号。
10. 比较 RLHF 与 DPO、GRPO 等对齐算法的异同与适用场景。

► **Constitutional AI 自我修正系统：引入宪法原则，实现 Self-Critique 与价值对齐。**

1. 构建自然语言宪法（Constitution）驱动的大语言模型训练流程。
2. 基于模型自评的对齐反馈机制（Self-Critique & Revision）设计。
3. 构建三阶段训练流水线：SFT → AI Preference Comparison → Fine-tuning。
4. 定义用于偏好对比的宪法原则集合详解。
5. 训练 Preference Model 时使用 AI 代替人类打分，构建更可扩展的偏好数据集。
6. 实现 Harmlessness - Helpfulness 的 Pareto 优化（双提升）目标。
7. 引入自监督对齐（RLAIF）策略：AI 打分 + 强化学习优化策略生成器。
8. 在流程中编码透明性（Interpretability）：显式生成行为理由与拒答解释

模块十一

模块十二

模块十三

模块十四

9. 支持多版本宪法实验：民主型宪法生成流程 + 多文化偏好注入机制探索

► PPO 强化学习算法实现：从策略迭代到收敛曲线，掌握核心参数与调试路径。

1. 理解 PPO 的核心目标函数与 Clip Trick 推导。
2. 构建 Actor-Critic 架构下的 LLM 策略网络。
3. 使用 KL penalty 控制策略更新距离。
4. 配置 Rollout、Reward Normalization 与 Value Baseline。
5. 使用奖励模型评分作为训练信号并更新策略。
6. 实现 Multi-Batch Update 与 GAE (优势估计) 机制。
7. 监控策略变化、Entropy、KL 值等收敛曲线。
8. PPO 参数调优实践 (学习率、Batch size、Clip Range)。

► DPO 直接偏好优化：构建高效免 RewardModel 的训练流程，兼容多样性控制。

1. 为什么 RFT (Reinforcement Fine-Tuning) 更适用于多任务推理与数据稀缺场景。
2. RLHF 与 RFT 的底层训练架构区别 (reward learning vs reward programming)。
3. GRPO vs PPO / DPO / RLO 的 loss 结构对比：监督来源、优势估计、KL 控制。
4. GRPO 的核心理念：直接优化可编程 reward function 而非训练 reward model。
5. GRPO 总 loss 分解：ratio loss、advantage loss、clip loss、KL penalty 四因子组合。
6. 可编程 reward function 示例：Python 结构模板 + task-specific 规则注入。
7. 使用 LLM-as-a-Judge 架构定义 reward pipeline (如 GPT-Eval, GPT-4 Voting)。
8. 构建主观性敏感 reward: correctness \times confidence 等 soft reward 设计。
9. Verifiable reward 的重要性及数据不一致导致的优化崩溃问题分析。
10. 多目标 reward 组合: correctness + novelty + efficiency 的线性加权与归一化。
11. Reward clipping 与 normalization 技术在稳定训练中的关键作用。
12. GRPO + LoRA 微调方法在 7B 以下模型的轻量部署路径。

► GRPO 生成式奖励调优系统：以 Python reward function 编程实现灵活策略训练。

1. DAPO 融合三大策略：Clip-Higher、动态采样、Token-Level loss。
2. 提出 Decoupled Clipping，正则化梯度提升训练稳定性。
3. DAPO 适配非均匀 reward 分布，优化多样训练样本表现。
4. Clip-Higher 通过双边剪裁控制策略偏移与探索程度。
5. ϵ_{high} 设定 reward 上限，防止模型偏移与异常奖励。
6. ϵ_{low} 设定学习下限，保留低分样本的训练可能性。
7. 联合 KL penalty 共同控制策略收敛与泛化能力。
8. Clip-Higher 增强 policy entropy，引导多样性输出。
9. 设定 ϵ 区间经验值，适配多任务与多模型训练。
10. Clip 机制在策略优化中实现探索与利用的平衡。
11. 动态采样 Top-K 筛选，提升高质量样本梯度密度。
12. reward normalization 保持采样分布均衡与稳定。
13. 训练中逐步剔除低 reward 样本，提升有效样本率。
14. 样本老化机制防止单批次样本主导训练方向。
15. 动态分阶段采样：预热宽松、收敛阶段精准控制。
16. 引入 Token-Level PG loss 细化到每个 token 级别奖励。

► DAPO 策略融合机制：整合 Clip-Higher、动态采样、Token-Level loss 等技术路径。

1. 定义 TTRL：在测试阶段使用强化学习提升模型推理对齐能力。
2. 构建 Majority Voting 奖励，无需人工标签或奖励模型参与。

本文件内容格式受中国软件行业协会版权保护，任何未经授权的格式模仿和抄袭行为我方将予以追责。

模块十五

模块十六

模块十七

模块十八

3. 使用 Top-k 与温度采样生成多样化解答样本作为学习基础。
4. 自动提取每轮输出中的最终答案，用于构建多数派奖励。
5. 设计 1/0 或 0~1 连续奖励向量用于梯度优化与策略学习。
6. 利用 GRPO 算法实现对多轮 reward 的策略稳定更新优化。
7. 无需 RewardModel 与 Preference 数据，支持轻量级无监督训练。
8. 提高奖励信号稳定性：多轮投票与解答一致性增强方式。

► TTRL 测试时强化学习方法：引入多解路径与 Majority Voting 策略完成无监督对齐。

1. 定义 AZR：无需标签和偏好数据实现从零开始的推理能力学习。
2. 使用 Zero-Signal Bootstrap 机制启动无监督推理行为优化过程。
3. 通过 Batch 级 Self-Play 并行生成多个推理路径进行自我博弈。
4. Latent Replay Buffer 用于存储隐空间信息与奖励信号对照数据。
5. 构建 Majority Voting Proxy 代替人工偏好用于策略评估与更新。
6. 应用温度退火策略逐步收敛探索行为形成稳定推理策略。
7. 构建完整自演化闭环系统：从任务输入到奖励输出再自我优化。
8. Token-Level 奖励归因将分数精确映射到关键推理 token 位置。
9. Prompt 重写机制根据模型反馈自动调整结构与提示格式。
10. 利用 Self-Distillation 方法模仿自身优解实现自我强化学习。

► AZR 零监督自演化策略与 RL-AZR 评估框架：构建全流程自博弈智能体体系，实现自我演化与可控验证。

1. scripts/test/test_end_to_end.py 脚本支持从 prompt 到结果全流程测试。
2. 支持 CLI 指定 config、strategy、prompt 类型与 judge 模型。
3. 测试框架支持 math、code、logic 多任务类型同时运行。
4. EvaluationMetrics 统一评估指标：correctness、reward、trace_score。
5. 支持 reward_judger 模块配置多种 judge 模型用于奖励评分。
6. 测试用例支持 prompt_path、output_dir 等多参数控制任务结构。
7. judge 结果保存为 judge_output.jsonl，便于后续分析与评分校验。
8. 测试可输入多个 checkpoint 执行策略对比，输出 win-rate 曲线。
9. 支持生成 reward_curves.pdf 曲线分析策略收敛与评分变化趋势。
10. 推理日志输出 reasoning_trace.html 可视化展示推理结构过程。
11. reward_ensemble 模块融合多个 reward model 投票结果提高稳定性。
12. failed_cases/ 自动保存失败推理样本用于训练数据补充与复盘。
13. chain_metric.py 支持计算推理链长度、深度等行为指标。
14. 测试日志结构化保存 logs/test_run_{timestamp}.json 可追踪可回放。
15. strategy tester 支持扰动下策略鲁棒性测试，检测模型稳定性。
16. reward validator 注入行为链验证器，自动校验奖励合理性。
17. reward-quality 曲线与 token-sparsity 散点图支持策略差异分析。
18. 支持推理失败自动 fallback，避免长链任务执行失败中断流程。
19. 各轮推理记录 latency 与 token usage 用于性能优化分析。
20. 支持 push-to-hub 上传测试结果至 Huggingface Spaces 实现集中管理。

模块十九

模块二十

五、实训收益

► GPT-5 赋能新纪元：

► AI+产业落地：

本文件内容格式受中国软件行业协会版权保护，任何未经授权的格式模仿和抄袭行为我方将予以追责。

率先引入 GPT-5 的统一推理 (Unified Reasoning) 与原生 Agent OS 技术, 实现跨场景推理、自动工具编排、长上下文记忆与实时多智能体协作, 直接驱动企业智能化跃迁。

► 国家战略直通:

课程体系紧扣“人工智能+”政策, 聚焦可规模化、可治理、可商用的 AGI 路径, 助力企业抢占国家战略新赛道。

► 全栈实战导向:

涵盖 LangGraph 流程编排、MCP 上下文调度、A2A 多 Agent 协同、RLHF 对齐训练、自演化 AZR 等核心技术, 全流程源码级操作, 企业即学即用。

深度聚焦金融、制造、科研、运营、政务等“人工智能+”重点场景, 实现传统业务智能升级与 AI 原生业务场景共创。

► 构建可控可进化架构:

构建可持续演化的企业级智能体系统, 确保全链路可观测、可治理、可扩展, 为长期商业竞争力提供坚实技术底座。

► 核心价值:

打造可控的企业级智能体 AI 系统

► 第一阶段:

1. 掌握“三可”AGI 实施框架→ 深度理解可规模化 (场景开放)、可治理 (伦理嵌入)、可商用 (生态反哺) 的政策落地路径, 抢占国家战略新赛道。
2. 贯通国产化全栈技术链→ 实践昇腾芯片+紫东太初大模型的国产基座适配, 覆盖数据清洗、模型微调、安全部署全流程, 筑牢自主可控技术根基。
3. 解锁五大场景 AI 原生变革→ 聚焦 金融/制造/科研/政务/运营 领域, 通过认知工业范式、动态治理模型等方法论, 推动传统业务智能升级与 AI 原生场景共创。
4. 构建行业治理实战能力→ 演练金融可追溯决策、医疗伦理委员会、低空经济联防机制等治理工具, 破解 AI 落地合规难题、洞察 AGI 融合路径 (类脑/博弈/信息智能)、第五科研范式 (AI 驱动科学猜想) 及治理挑战, 布局长期竞争力。
5. 接入产业生态网络→ 联通政府开放场景、算力券资源池及模型共享社区 (如 OpenI 启智), 赋能企业轻量化转型与普惠 AI 落地。

► 第二阶段:

1. 理解 GPT-5 最新架构与推理机制内幕、全面解析其多阶段推理链路、工具调用机制、上下文扩展策略与训练优化方法;
2. 掌握 LangGraph、MCP、A2A 等主流智能体核心协议与源码架构;
3. 能够构建具备状态感知、任务规划、结果评估与自我演化能力的 Agent 系统;

4. 具备独立实现 RLHF、DPO、GRPO、AZR 等主流对齐路径的工程实践能力；

5. 企业级大模型智能体项目从任务建模 → Agent 编排 → RL 训练 → HITL 反馈的闭环实现；

6. 具备从 0-1 搭建“人工智能+”场景智能系统的全栈工程能力与项目交付能力；

▶ 第三阶段：（互动答疑及考核认证）

本期我们特别设置了互动答疑环节，为参会代表提供一个深度交流的平台，鼓励大家积极参与讨论和提问，将自身企业在业务发展中遇到的问题向专家咨询，参加实训并经过考核合格的学员，将颁发高级证书。证书将作为您在求职、升职加薪、招投标过程中的重要参考依据。

六、专家介绍

专家	简介
王老师（中国科学院）	<p>王老师：中国科学院自动化研究所副总工程师，武汉人工智能研究院院长，紫东太初大模型研究中心常务副主任，研究员，博士生导师，中国科学院大学人工智能学院岗位教授，多模态人工智能产业联盟秘书长。</p> <p>主要从事多模态大模型、视频分析与检索、大规模目标识别等方面的研究。其团队研发的“紫东太初”多模态大模型，成为全球首个千亿参数级多模态预训练模型，斩获年世界人工智能大会最高奖项卓越引领者（SAIL）奖。</p>
王老师（硅谷 CTO）	<p>王老师：人工智能专家，现任硅谷一家智能体企业 CTO，专注于可控大模型、以人为本的强化学习（Human-Centered RL）与智能体系统。毕业于斯坦福大学，长期在硅谷从事前沿 AI 研发与落地，带领团队成功交付 11 个大型 NLP 项目，服务覆盖字节跳动、Apple、PayPal、摩根大通、培生教育、LinkedIn、腾讯 等。拥有 10+ 年湾区经验，历任 CTO、执行副总裁、首席数据科学家 等关键岗位。</p>

技术专长：以 Controllable LLMs、Human-Centered Deep RL、Agentic AI 为核心；Conversational AI 与可控自然语言生成（Controllable NLG）以及面向对话的智能体系统上具备行业领导力。熟练基于 GPT、Llama、Claude、DeepSeek 等前沿模型交付高影响力解决方案，并结合 RLHF、PPO、DPO、GRPO, KTO, ORPO, RLAIIF, SPCT, MCTS, DAPO, TTRL, AZR 等对齐与策略优化算法，确保系统可控、可靠且可规模化。

七、证书及费用

项 目	说明
测 评	本次培训结束后，将进行专业测评考试，经考核合格，可申请以下证书：
证书类型	A类：由中国软件行业协会颁发《企业级智能体架构》、《生成式AI治理》、《大模型推理架构师》高级职业技术水平证书（三选一） B类：在A类基础上增加《数字技术应用（人工智能）》、《人工智能应用》高级技术证书（二选一）
费用标准	A类：3980元/人（含报名费、培训费、专家费、资料费、考核建档及申报证书费） B类：5380元/人（含两本证书费用）

八、报名方式

天津市软件行业协会：	
联 系 人：胡老师	手 机：13820397363（同微信）
电 话：022-23789297 转 810	报名邮箱：hn@tsia.com.cn
报名材料：	<ol style="list-style-type: none"> 1. 填写完整报名回执表（附件（二）） 2. 2寸电子版证件照 3. 汇款凭证（如选择电汇） 4. 以上发送到邮箱并联系教务老师