代表性成果登记表

申请人姓名：肖仰华　　所在院系：计算与智能创新学院

研究类别：[　]基础创新类　　[√　]应用创新类

成果名称：

自进化智能体系统、应用与治理

成果支撑材料清单（分点列出组成本项成果的高水平论文、专著、授权专利、政策咨询报告等）：

一、高水平论文

1.Siye Wu, Jian Xie, Yikai Zhang, Aili Chen, Kai Zhang, Yu Su, Yanghua Xiao（通讯作者）. ARM: Adaptive Reasoning Model. Neural Information Processing Systems 38 Main Conference (NeurIPS 2025)（NeurIPS 2025 亮点论文）

二、代表性开源系统与平台

1.GenericAgent 自进化智能体（https://github.com/lsdefine/GenericAgent，MIT 协议，GitHub Star 1.42 万、Fork 1,700 余）

2.SkillsHub 百万级技能库平台（146 万余条 AI Skills）、SOPHub 操作流共享平台

三、授权专利

1.基于文档结构树的大模型检索问答方法及装置，专利号： 202510050486.8

2.大语言模型不完全信息下的问题处理能力的评测装置，专利号： 202410780036.X

四、政策咨询报告

1.《智能原生企业塑造智能经济新图景》，《人民论坛》杂志 2025 年第 23 期

2.《“养龙虾热”的智能体安全隐忧》，《人民论坛》杂志 2026 年第 6 期

成果简介（200-300 字）

针对长任务上下文膨胀、经验难以复用、安全难以管控三重难题，申请人提出以“上下文信息密度最大化”为原则的自进化智能体架构，主持研发并开源通用自进化智能体 GenericAgent（GA），以原子工具、分层记忆与任务反思将已验证经验沉淀为可复用技能，在技术报告评测中显著降低 token 消耗，开源后 GitHub Star 逾 1.4 万，被多篇综述列为代表性自进化系统；提出快慢思考自适应推理模型 ARM，入选 NeurIPS 2025 亮点论文并获华为火花奖。成果应用于深圳福田“政务龙虾”（登顶微博热搜第一、获央视持续报道）和上海青浦两办办公辅助系统（8 项功能上线试运行）；GA 手机版截至 2026 年 9 月 26 日覆盖 654 台外部设备、完成约 2.3 万个 AI 任务。咨政建言被新华社、中央广播电视总台内参采纳；以首席科学家身份支撑夸夸菁领产业化（估值已到 30 亿人民币）；作为负责承接青浦复旦未来技术研究院通用人工智能实验室建设（累计投资 1.5 亿人民）；并在华为全球战略技术大会等作主旨报告。

附说明：

一、科学贡献

智能体是大模型走向千行百业的主要形式。2026 年以来，以 OpenClaw 为代表的自主执行式智能体在全球快速普及，“智能体能否可靠、低成本地执行真实任务并持续进化”成为人工智能竞争的新焦点。然而，通用智能体的落地仍面临三重核心难题：执行上，长周期任务中工具描述、记忆与环境反馈不断累积，上下文膨胀导致决策质量下降、token 开销失控；进化上，任务中获得的经验难以沉淀为可复用能力，系统“用千百次无变化”；治理上，政务与产业场景要求数据不出域、行为可审计、安全可控，而主流开源智能体代码规模巨大、安全边界难以统一管控。对此，本项目围绕自进化智能体的架构、推理与治理开展系统研究，形成如下三方面成果：

1. 提出了以“上下文信息密度最大化”为原则的自进化智能体架构。申请人提出该设计原则并主持系统研发，以约 3000 行核心代码、9 个原子工具、L0—L4 五层按需记忆和“任务反思—经验固化—技能复用”闭环，使智能体能力随使用自主生长，在技术报告所设评测任务与基线条件下 token 开销较主流系统降低 3—10 倍；据此研发并开源通用自进化智能体 GenericAgent（GA）。

2. 提出了依据任务难度自适应切换快慢思考的推理方法。申请人作为通讯作者指导完成该项工作，针对推理模型“思考长度膨胀”问题，提出自适应推理模型 ARM，在保持准确率的同时显著压缩推理开销，入选 NeurIPS 2025 亮点论文，获华为火花奖，其技术思路延伸至 GenericAgent 的高效运行设计。

3. 形成了面向政务与产业的自进化智能体应用与治理实践。申请人牵头以 GenericAgent 为统一底座，形成“极简架构 + 自主执行 + 安全沙箱 + 全程留痕”的政务智能体技术路线，在深圳福田、上海青浦等地上线运行；建成 SkillsHub、SOPHub 技能共享平台；围绕“龙虾”类智能体的安全隐忧与智能原生企业发展提出咨政建议，相关建言被新华社参考报道和中央广播电视总台内参产品采纳。

上述成果相关论文入选 NeurIPS 2025 亮点论文，方案获华为火花奖（见图 1）；GenericAgent 于 2026 年 1 月开源，GitHub Star 逾 1.4 万、曾登顶 GitHub 趋势榜（见图 2），被多篇综述与 Hermes Agent、SkillClaw、EvoAgentX 并列为代表性自进化系统；作为深圳福田“政务龙虾”核心登顶微博热搜第一并获央视持续报道（见图 3）；GA 手机版上线三个月完成 2.3 万个 AI 任务并参展 2026 外滩大会（见图 4）；支撑总预算 1.5 亿元的通用人工智能实验室立项建设及 4 家企业合计 2,700 万元的联合实验室合作（见图 5）；申请人受邀在华为 STW、中国智能体大会等作主旨报告。

二、代表性论著评价

申请人提出设计原则并主导研发的通用自进化智能体 GenericAgent（技术报告）提出“上下文信息密度最大化”这一统一设计原则，以约 3000 行核心代码、9 个原子工具与 L0—L4 五层记忆实现智能体技能的自主生长：在 Lifelong AgentBench 上以 100%完成率仅消耗 Claude Code 27.7%、OpenClaw 15.5%的输入 token；9 轮重复任务中 token 消耗下降 89.6%、调用次数由 32 次收敛至 5 次。技术报告收录于 Hugging Face Papers（获 23 次点赞、被 8 个论文集收录）。学术界已将其视为代表性自进化系统：MOSS（arXiv 2605.22794）将 GenericAgent 与 Hermes Agent、SkillClaw、EvoAgentX 并列为应用级自进化智能体的四个代表；Bayesian-Agent（arXiv 2606.08348）以 GenericAgent 为基线并为其构建专用适配器；《Self-Improvements in Modern Agentic Systems: A Survey》等综述将其收录；浙江大学求是特聘研究员朱强在论文中将其纳入“将交互经验转化为记忆、提示、技能、工作流、策略或架构等持久化产物”的自进化智能体谱系。该智能体作为来自高校的“龙虾”类智能体代表，得到学术界与工业界的认可，解放日报评价其为“国内高校自进化智能体标杆”。

申请人提出的依据任务难度自动选择推理形式的自适应推理模型 ARM（论文 1），入选 NeurIPS（CCF A 类会议）2025 年度亮点论文。天津大学菁英教授郝建业在论文中评价其为对现有强化学习的有效改进策略，能够有效抑制思考长度膨胀、提升推理模型准确率。该“快慢思考自适应平衡”方案荣获华为火花奖，其技术思路延伸至 GenericAgent，为其低 token 开销、高效运行提供了方法基础。申请人已三获华为火花奖（2022 年知识图谱自动更新、2024 年大模型持续学习、2026 年快慢思考自适应），解放日报·上观新闻 2026 年 3 月 28 日以《三获华为火花奖，“养龙虾”的复旦教授破解产业真问题》作专题报道。

申请人围绕智能体治理持续建言。《智能原生企业塑造智能经济新图景》《“养龙虾热”的智能体安全隐忧》刊于《人民论坛》（见图 6），就智能原生企业发展路径和自主执行式智能体的安全治理提出对策；关于“人工智能与社会治理”的成果及观点被新华社三篇参考报道吸收采纳，为中央领导提供决策参考，新华社上海分社于 2024 年 12 月致函复旦大学表示感谢（见图 7）；2026 年 3 月，申请人就“OpenClaw 等 AI 智能体带来的风险隐患”接受中央广播电视总台专访，稿件于 3 月 20 日被总台内参产品采纳、在决策咨询中发挥了重要作用，总台于 2026 年 6 月致函感谢（见图 8）。

三、研究成果应用与落地

1.开源生态与真实使用。GenericAgent 于 2026 年 1 月 11 日在 GitHub 以 MIT 协议开源、1 月 16 日发布 V1.0（早于 OpenClaw 的 2026 年 1 月 25 日）。截至 2026 年 9 月，GitHub Star 1.42 万、Fork 1,700 余、累计提交 1,359 次；曾登顶 GitHub 日趋势榜（Trendshift 收录），2026 年 4 月 GitHub 趋势榜位列全球第四、领先 Hermes Agent，被多位海外知名博主关注并传播（见

图 9）。技术报告收录于 Hugging Face Papers 并被多个论文集收录；Datawhale 社区出品《Hello GenericAgent》入门教程；社区自发开发 ga-manager、galley、A3Agent、GenericAgent-Admin 等第三方 GUI 客户端；建立 22 个用户交流群；2026 年 5 月发布 Windows/macOS 桌面端并推出一键安装。

推出可在用户自有手机上运行的 GA 手机版。区别于多数跑在云端模拟器或网页壳中的手机 Agent，GA 手机版直接运行在用户自己的手机里、操作用户已登录的 App。自 2026 年 6 月 23 日首个外部用户起，上线 96 天（截至 9 月 26 日，已剔除内部与测试设备）累计外部用户设备 654 台、登录账号 386 个，用户发出指令 2.6 万条、AI 完成任务 2.3 万个（任务顺利完成率 73%，任务耗时中位数 56 秒），Agent 工具调用 15.7 万次、AI 手机操控动作 8.4 万次，被 AI 操控过的 App 达 147 款（微信、小红书、淘宝、美团、抖音等）；日活峰值 99 台，首批用户第 8 周留存仍达 30%，各批次次周留存平均 44%。

2. 技能平台建设。SkillsHub 百万级技能库于 2026 年 3 月 10 日上线，汇聚 146 万余条 AI Skills（索引自 GitHub、Skills.sh 等公开资源），按领域、类型、场景细粒度分类并搭载推荐引擎，为智能体提供“能力索引市场”，作为国内高校最早的“龙虾”相关成果之一，获复旦大学官网、新浪财经等报道。SOPHub 操作流共享平台面向 GA 开放，汇聚官方与社区投稿的 200 余个高质量 SOP，覆盖浏览器操作、通信工具接入、手机操控、内容生产等领域，并已集成为 GA 原生能力，GA 可自动下载与上传技能。

3. 政务智能体应用。深圳福田“政务龙虾”DinTal Claw 以 GenericAgent 为核心，2026 年 3 月 6 日由福田区发布，为全国首批在政务外网本地化部署的自主研发智能体，覆盖民生诉求调研分析、公共场所卫生许可办理等场景，实现“小时级任务分钟级完成”，综合运行成本不足传统办公模式的 15%。#首批政务龙虾#、#公务员养上政务龙虾了#等话题登顶微博多榜热搜第一，央视新闻重点关注并持续报道（阅读量 44 万+），20 余家媒体同步跟进，央视新闻点评“国产龙虾创新加速，核心优势逐步显现”。

上海青浦区委办、区府办“政务智能·办公辅助”系统以 GenericAgent 为技术底座，2026 年 7 月 31 日在政务外网上线试运行，现已上线讲话稿写作、文档校对、常委会通知生成、制发文件与基层减负 39 条一致性评估、每日要讯生成、政策合规审查、政务文件合法性审查、常务会议议题生成共 8 个智能体，实现从“人工起草”向“智能生成＋人工审校”转变。系统全链路数据不出内网；通过基于 8 张昇腾 910B 自部署 DeepSeek V4 Flash 底座模型并优化容错机制，响应速度提升 3—5 倍、并发数由 3 提升至 12；针对智能体框架本身的对比评测显示，GenericAgent 框架在 74 个请求上全部成功、未引入失败或死循环，低并发条件下总体吞吐较裸模型高 9.86%、平均延迟低 7.31%。应用成效方面，两办 8 个应用均已完成端到端流程打通并在内网稳定调用，常委会通知生成已在多场实际常委会会议中与正式下发通知比对

验证，制发文件与基层减负 39 条一致性评估输出含三层证据链的结构化《参考意见》，特定领导讲话风格定制生成已通过样例验证；试运行期间累计发现并解决性能、功能、网络适配问题 10 项，双方拟定《智能政务合作框架协议》建立长期产学研合作机制。此外，青浦自贸区政策问答系统已导入 2,587 份政策文件并进入用户验收测试。

4. 工业与科学场景验证（均基于同一 GenericAgent 底座，为原型验证或试点部署阶段成果）。智能点锡机 EDispense：将触控屏控件映射为 GA 可调用工具，实现“一句话驱动 AOI 对位 + Gerber 推理 + 点锡执行”，运动类操作配备安全确认闸门，已开源。芯片后端设计 RTL-to-GDS：基于 SMIC 40nm 工艺与 Synopsys/Cadence 双工具链，GA 自主编写并调试 TCL 脚本、自主查阅手册试错，完成 RTL 到 GDS 全流程，478 组对拍全部通过，Fmax 283.3MHz，关键 DRC 项清零。强对流短临预报：在上海气象局试点应用，GA 学习资深预报专家的 67 个数据接口经验，建立 122 天可回测系统，在特定个案复盘中预警发布较官方渠道提前约 50 分钟。FVCOM 海洋水动力建模：GA 自主获取公开数据、云端编译部署、诊断并修正模型崩溃根因，在测试案例中将传统 1—2 个月的建模周期压缩至数小时，自动交付报告与 11 张图件。IT 设备智能运维：沉淀专家经验实现算力设备到工业设备的运、检、维、优自动化，GA 自主优化的大模型推理框架实现 78%提速。

5. 科技成果转化与产业化。申请人担任深圳夸夸菁领科技有限公司首席科学家，GenericAgent 由复旦知识工场实验室与该公司联合工程化，其“政务龙虾”DinTal Claw 为 GenericAgent 唯一官方授权商业版本。据公开报道，该公司 2025 年 12 月完成中金资本旗下基金投资，估值近 30 亿元（南方都市报 2026 年 5 月），被称为“数智员工赛道独角兽”，2025—2026 年度获华为全球优秀合作伙伴奖，计划 2027 年一季度申报港股 IPO；其 AI 数智员工在深圳福田辅助化解纠纷 1,300 余宗、调解时效提升 70%。双方签订横向课题“自进化智能体关键技术研究与应用”（2026）。

以 GenericAgent 为统一内核的“通用人工智能实验室”已在上海市青浦复旦未来技术研究院立项建设（2026—2029 年，总预算 1.5 亿元），申请人任实验室主任，面向政企、工业、个人三类对象开展产品辐射。依托实验室，已与 4 家企业签署校企联合实验室/技术合作协议，合同总额 2,700 万元：上海启源芯动力科技有限公司 500 万元（2026 年 8 月—2029 年 8 月）、深圳蜂巢互联科技股份有限公司 1,200 万元（2026 年 9 月—2029 年 9 月）、无锡华光环保能源集团股份有限公司 500 万元（2026 年 9 月—2029 年 9 月）、深圳夸夸菁领科技有限公司 500 万元（2026 年 9 月—2029 年 9 月），以 GenericAgent 为共同技术底座面向芯片、工业互联、环保能源与政企数智员工等领域开展联合研发（见图 10）

6. 学术与产业影响力。2026 年，申请人在华为 STW（战略技术大会）主会场作 Keynote《大模型驱动的自进化智能体技术突破与洞察》，在 2026 中国 AI 智能体大会、全球人工智

能技术大会（GAITC）作主旨报告《模型驱动的自进化智能体：技术洞察与趋势展望》，在网络 AI 技术论坛作《自进化智能体——认知与实践》，并在第二届浦江 AI 学术年会、2026 企业数智化创新峰会等作主旨报告。GenericAgent 以“自进化智能体开创者”为主题参展 2026 Inclusion·外滩大会（9 月 9—12 日），展出 GA 手机版：展期 4 天平台累计完成 AI 任务 1,401 个（展台演示 431 个+用户手机 970 个），活跃设备 123 台，其中新激活设备 43 台、新注册账号 39 个，前者接近展前整整两周的自然新增（46 台）、后者超过展前两周（31 个）；展会所在周新增设备 52 台，为 8 月以来单周最高，日均新注册、日均有效激活用户较展前分别提升至 4.4 倍、4.7 倍，DAU 提升 42%；展期新用户中 72%完成了首个 AI 任务（展前为 50%），84%的激活发生在 11—17 时展馆开放时段，与现场扫码即装即用的体验相符；展会开始后 18 天与之前 18 天相比，新激活设备增长 77%、新注册账号增长 79%、用户指令增长 48%。相关成果获央视新闻、解放日报/上观新闻、机器之心、21 世纪经济报道、南方+、南方都市报、广州日报、深圳新闻网、上海经信委、36 氪等持续报道，申请人接受《三问“政务龙虾”首席科学家：安全否？省钱吗？会取代人类吗？》《DinTal Claw“政务龙虾”首席科学家肖仰华：智能体“奇点时刻已来”》等专访。



图 1 快慢思考自适应推理方案获华为火花奖（2026）

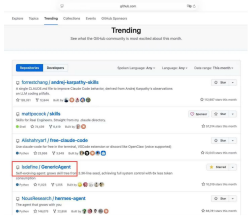


图 2 GenericAgent 位列 GitHub 月度趋势榜全球第四（2026 年 4 月）



图 3 央视新闻报道以 GenericAgent 为核心的深圳福田“政务龙虾”



图 4 2026 Inclusion·外滩大会 GenericAgent 展台现场（观众正在体验 GA 手机版）

<table>
<tr>
<td>

图5上海市青浦复旦未来技术研究院与4家企业集中签署联合实验室共建协议</td>
<td>

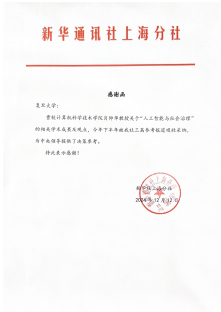
图6 咨政文章《智能原生企业塑造智能经济新图景》刊于《人民论坛》</td>
</tr>
<tr>
<td>

图7 新华社上海分社感谢函：申请人关于“人工智能与社会治理”的成果及观点被三篇参考报道吸收采纳（2024年12月）</td>
<td>

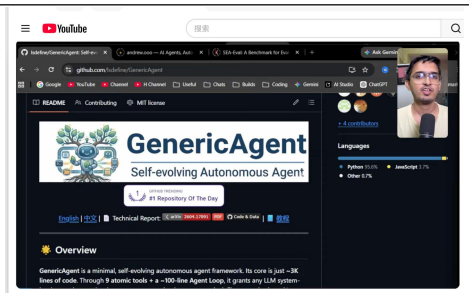
图8 中央广播电视总台感谢信：申请人关于“OpenClaw等AI智能体带来的风险隐患”的访谈稿件被总台内参产品采纳（2026年6月）</td>
</tr>
<tr>
<td>

图9 海外博主专题视频“GenericAgent Beats OpenClaw & Hermes Agent”</td>
<td>

图10 申请人代表上海市青浦复旦未来技术研究院与上海科学智能研究院、上海祖泉创新转化研究院分别签署三方战略合作协议</td>
</tr>
</table>