

陈嘉蔚

felix_chen@sjtu.edu.cn | academic.snowflake.zone | github.com/glimmeringlight | Tel: 159-0162-7247

教育经历

上海交通大学 | 网络空间安全，计算机学院 | 学术型硕士研究生 2025.09 – 2028.03 (预计)

- 导师：王烁，研究方向：可信人工智能与人工智能安全。
- **GPA**: 3.83/4.0。
- 荣誉：一等学业奖学金 (前 25%，2025 年)。

上海交通大学 | 信息安全，电子信息与电气工程学院 | 工学学士，辅修经济学学士 2021.09 – 2025.06

- 主修课程：程序设计思想与方法 (91)，数据结构 (96)，操作系统原理 (90)，编译原理 (93)，数据库原理 (96)。
- 学积分：91.35/100 (前 10%)。
- 荣誉：上海交通大学三好学生 (2021 年)、校友奖学金 (2022 年)。

专业技能

- 编程语言与工具：Python/C，熟悉 Linux；掌握 Git 进行版本控制。
- **AI 能力**：熟悉 PyTorch、Hugging Face Transformers 等主流 AI 开发框架，掌握 Transformer、视觉语言模型 (VLM) 等模型架构，具备模型训练、微调与评估实践经验。
- 语言：CET-6 (601/710)，CET-4 (668/710)。

实习经历

字节跳动 | Agent 应用算法 2026.08 – 至今

项目经历

在线求职平台 | Python Django / 业务逻辑设计与实现 / 课程项目 2023.09 – 2023.10

- 设计用户注册、登录等功能，并基于 Cookie 维持登录状态。
- 参与表设计工作，使用 MySQL 数据库进行持久化存储。
- 支持发布职位、上传简历、投递简历等基础功能；实现小型论坛，并利用 Python-Markdown 支持富文本编辑。
- 自 2024 春季学期起，该项目作为 NIS3368 课程的迭代基础项目。
- github.com/glimmeringlight/NIS3368_JobRecruitment

科研经历

大语言模型可答性检测与信息不足识别研究 | 研究负责人 2026.05 – 至今

- 围绕大语言模型在信息不足、问题歧义等场景下的可答性判断展开研究，构建“可回答性判断、是否拒答、澄清式提问”的任务框架，并对不同不可回答类型进行细粒度划分。
- 基于大语言模型中间层隐藏状态训练轻量级探针，研究模型内部表征对信息缺失的感知能力，并开展跨数据集、跨领域及跨模型的泛化实验。
- 针对跨域测试中的阈值偏移与数据集偏差问题，引入少样本阈值校准、对抗瓶颈和数据增强方法，在不微调大语言模型参数的情况下提升可答性检测的域外泛化性能。

“明岐”医学影像大模型矩阵 | 项目参与者 2025.07 – 2026.05

- 构建涵盖肠道疾病中息肉识别、息肉分类和息肉定位的测评体系，并对现有医学专用模型进行评测。
- 使用监督微调与强化学习方法对基座模型进行微调，提升模型在肠道息肉疾病上的精准识别与描述能力。
- 项目入选第二届“一带一路”科技交流大会十大人工智能应用场景，获评 2025 全国颠覆性技术创新大赛优胜奖。

“佑安”全灾程应急响应多智能体大模型协同智链 | 子课题负责人 2025.06 – 2025.10

- 项目提出“多模态融合编码器 + 专家智能体集群”架构，覆盖灾前预警、灾中评估和灾后响应全链条。
- 负责受灾评估模块的开发以及 Agent Tool 的实现，作为子智能体供主智能体调用，进行灾中受灾程度评估功能。
- 项目荣获 2025 年度“挑战杯”中国青年科技创新“揭榜挂帅”擂台赛人工智能赛道擂主 (该赛道全国唯一最高荣誉)。

基于 SM2 白盒数字签名算法的非接触式交互认证系统 | 毕业设计 2024.09 – 2025.05

- 对国密 SM2 数字签名算法有深入了解；熟悉其白盒化实现方案。
- 使用 C 语言实现 SM2 白盒数字签名算法，包括 256 位大整数运算、有限域运算、椭圆曲线点运算和算法层面实现。
- 掌握 NFC 技术原理；基于 LPC55S69 嵌入式平台和 PN7150 NFC 模块，完成交互验证系统设计与实现。