

陈磊

人工智能硕士研究生 · Rust 开发者 · 实用 AI 系统构建者

dev@somnus.wiki · GitHub · 个人博客 · 中国 · 南京

个人简介

中国科学院计算技术研究所人工智能专业硕士研究生（推免录取），本科毕业于西南交通大学智能制造工程专业。关注检索增强生成、Embedding 与 Reranking、参数高效微调、模型轻量化，以及面向机器学习场景的 Rust 工具开发；具备计算机视觉研究和电商数据工程实习经历。

教育经历

中国科学院计算技术研究所 / 中国科学院大学

2026 — 至今

人工智能专业 · 工程硕士（推免录取）

研究方向：高速视觉、多模态信息处理与实用人工智能系统。

西南交通大学

2022 — 2026

智能制造工程 · 工学学士

毕业设计：基于深度学习的双目立体匹配优化与轻量化方法研究。

实习经历

AI / 数据工程实习生 · 某电商公司

2026

- 搭建多店铺经营数据聚合与运营分析平台，统一管理销售、商品与经营指标。
- 开发每日销售额报数微信小程序，减少人工汇总和重复沟通。
- 设计电商后台数据采集、清洗、结构化存储与定时处理流程。
- 探索使用大语言模型完成商品名称、型号和 SKU 的归一化处理。

代表项目

Model Weight Parser

Rust · Ratatui · Safetensors

用 Rust 解析 Safetensors 模型权重，展示张量名称、数据类型和形状，并构建交互式终端界面，用于快速检查和理解模型权重结构。

RAFT-Stereo 优化与轻量化

PyTorch · Stereo Matching · FP16

围绕 Middlebury 2014 数据集评估细化模块、边缘损失和局部解冻策略，完成 FP16 推理部署；在保持精度基本稳定的情况下，将平均推理时间从 209.80 ms 降至 173.59 ms。

Superlabel

Python · 数据聚合 · 电商分析

面向电商商家的经营数据聚合与运营分析平台，整合多来源数据并提供统一查询、指标整理和报表能力。

技术能力

编程语言
人工智能
工程工具
开发环境

Rust、Python、TypeScript、JavaScript
PyTorch、RAG、Embedding、Reranking、LoRA、Safetensors
Git、Jujutsu、Docker、PostgreSQL、GitHub Actions、Cloudflare
CachyOS、Linux、Zed、Bash