

# 王晓丽

算法工程师

♀ 女 📅 1993-05 🎓 中国科学院大学 · 硕士 🌐 英语: CET6  
☎ 18322494286 ✉ cocoder@qq.com

## 技能专长

- 技术栈: Python, PyTorch, Transformers, scikit-learn, BERT, GPT, ChatGLM, Flan-t5, Linux, Flask, Docker, Git, JavaScript, React, Nodejs, langchain 等。

## 项目经历

### 关系抽取框架 InstructRE

指令微调

Flan-t5

Flask

React

Docker

2024年1月 - 2024年9月

**项目概述:** 针对传统生成式关系抽取框架结构化输出能力受限导致抽取效果不佳的问题, 独立设计并开发了面向指令微调模型的关系抽取框架, 大幅提升多场景下关系抽取效果。

- 独立设计并实现了多任务学习+指令微调的关系抽取框架, 提出自然文本对齐的三元组输出模板。
- 独立设计包含4种辅助任务的多任务联合训练框架, 基于FLAN-T5模型实现指令微调。
- 主导系统集成和部署, 基于Flask将算法封装为API服务, 使用React与AntV构建前端可视化界面, 支持在线识别与知识图谱展示。

**效果:** 在 5 个公开基准数据集上与4个主流模型对比取得 SOTA 效果, 平均 F1 提升 **3.51%**。

### 面向关系抽取的数据增强方法 TripleToText

数据增强

GPT3.5

LangChain

2024年3月 - 2024年9月

**项目概述:** 针对关系抽取任务样本标注成本高的问题, 提出利用语言模型融合反向生成自动进行数据扩充方法, 减少标注成本。

- 设计多样化的三元组扩充方法(如实体置换等), 利用语言模型实现高质量的语料自动生成。
- 开发基于实体匹配原则的数据质量过滤机制, 确保增强数据的可靠性。

**效果:** 在3个公开数据集的低资源场景下显著性能提升, 增强后实体抽取 F1提升 4.77%, 关系抽取 F1 提升了 3.27%, 且适用于多领域任务, 已在通用领域和科研论文领域验证有效。

### 基于知识图谱的医疗问答系统

Neo4j

BiLSTM-CRF

BERT+TextCNN

BM25

ESIM

2023年6月 - 2023年12月

**项目概述:** 针对现有医疗问答准确率低问题, 建立医疗知识图谱, 开发面向诊疗场景的智能问答系统。

- 设计项目架构并实现, 采用意图识别→实体提取→知识检索→答案生成的四层架构。
- 主导知识图谱构建, 设计数据 schema, 从寻医问药网站爬取数据作为问答基础知识库。
- 主导算法训练优化部署, 设计并实现分层意图识别模块, 采用LR+GBDT区分粗粒度意图, 微调BERT+TextCNN实现细粒度诊疗意图, 训练 BiLSTM-CRF 从用户输入中抽取实体, 采用BM25算法+ESIM模型进行实体规范化。
- 设计模板检索及自动应答流程, 从用户意图确定查询模板, 基于数据库返回生成准确的, 可解释的答案。

**效果:** 构建包含2万+实体、4万+关系的医疗知识图谱, 实现可溯源的精准问答。

## 教育背景

中国科学院大学 人工智能

硕士

2021年9月 - 2025年1月

山西大同大学 计算机科学与技术

学士

2012年9月 - 2016年6月

天津卓朗科技有限公司等 前端工程师

2016.7 - 2021.9

- 5 年前端开发经验，参与并交付 15+ 个项目，独立完成多个项目，包括运营系统开发，官网开发以及参与多个移动端项目，包括 React Native 跨平台应用开发，移动端 H5 页面等。
- 独立负责核心运营管理后台前端开发，基于React+Ant Design Pro搭建完整管理系统框架，整合ECharts/AntV满足多维数据统计需求。
- 主导3D模型交互页面和可复用UI组件库开发，支持数据中心实时监控业务。

#### 完美工事运营管理后台

React

Ant Design Pro

ECharts

AntV

TypeScript

项目概述：针对手机考勤APP所设计的运营后台，实现对app端各类数据的可视化分析及运营管理功能。

- 独立负责运营平台前端开发，基于antd-pro构建完整的运营管理系统，整合echarts、antv等可视化组件实现用户注册分布、活跃度等核心数据展示。
- 开发权限管理、订单管理、可配置表单等核心模块，开发基于角色的精细化权限控制和运营数据的查询导出。

#### 日志生成工具

项目概述：将工作记录生成工作日报的命令行工具

- 使用 Git API 获取本地代码仓库的提交记录，并按时间维度智能归类整理。
- 填充个人信息将其导出为 Excel。

#### Bee APP

react-native

react-navigation

项目概述：一款为用户提供保险产品查询、投保、理赔等服务的APP

- 根据UI设计稿还原页面布局和交互效果；开发保险产品展示、用户注册登录、保单管理等核心业务页面；优化页面性能和用户体验，确保iOS和Android平台的界面一致性。

#### 看脸吃饭APP及运营管理后端

React

Redux

Less

Webpack

Babel

Hybrid

项目概述：主打直播和订餐的hybrid app，以及对应的运营管理后台。

- 基于React技术栈开发hybrid应用的H5页面，实现webview与原生代码的交互通信，支持直播和订餐核心功能。
- 负责公共展示组件和容器组件的开发，后期引入Redux进行状态管理重构，提升应用架构和可维护性。
- 基于jQuery和H-ui框架开发运营管理后台，实现用户数据的管理、导出和表格分页等核心功能。
- 集成Validform和Mustache模板引擎，完善表单验证机制和数据展示模板，提升后台操作体验。

#### 南开大学企业号及后台管理系统

Vue.js

Bootstrap

ECharts

响应式设计

项目概述：主要实现校内事务集中到微信端处理。

- 独立负责微信端企业号前端开发，基于Vue.js和WEUI实现校内事务处理页面，开发应用图标拖拽排序、tab切换等交互功能，并适配不同设备。
- 独立PC端运营后台系统开发，实现管理员权限可视化配置、移动端数据分析展示、应用上下架管理等核心功能，并集成用户反馈处理机制。