

# PERSONAL RESUME

## 何龙

电话：13544175837

邮箱：1248551752@qq.com

性别：男

年龄：25



### 求职意向

意向岗位: 大模型AI数据标注（搜索评测、模型训练） 求职类型: 社招

### 工作经历

2025.07 - 至今

阿里巴巴夸克通义组

AI训练师

- 负责单轮及多轮对话数据的改写与质检，运用NLP技术优化对话逻辑与语义连贯性，制定并推动标准化质检规则落地，提升对话数据集质量。
- 建立并迭代数据标注规则体系，针对歧义场景、上下文关联等问题提供解决方案；通过实时答疑与规则优化，团队标注效率提升50%。
- 能快速检阅数据并判定保留或舍弃，确保数据合格率；主导标注规则的收集、制定与迭代，降低团队理解成本及标注误差。
- 精通提示词工程（Prompt Engineering），熟练通过PE技术进行高质量训练数据的生产与过滤，更擅长通过优化指令赋能团队，曾独立撰写模板显著降低新人上手门槛。
- 针对现有标注工具提出改进建议（如rag验真工具、自动化校验），显著提升标注效率。
- 运用提示词工程（PE）赋能提效：针对项目动态需求，持续优化并撰写核心标注环节的Prompt指令。
- 持续关注AI数据领域前沿技术，主动探索更高效的标注方法或工具改进方案，有运营个人抖音自媒体尝试AI知识科普，积累内容创作与用户沟通经验。（抖音号60733657275）
- 独立撰写全套SOP文档，覆盖数据预处理、标注规范到质检流程，并担任内部讲师培训新员工，缩短新人上手周期，提升团队整体产能。
- 通过对用户问题与模型回答进行质量评分，系统性标注并分类错误类型，精细化改写轻微问题回答，最终整合优化后的对话数据构建高质量数据集，为模型迭代提供精准反馈与数据支撑。

### 实习经历

2024.09 - 2025.06

阿里巴巴智能信息

大模型标注助理

- 负责文本数据的筛选、清洗及改写，通过文本对齐分析优化逻辑与相关性，累计处理数据超3万条，数据集可用性达98%以上，支撑模型迭代。
- “Bad Case分析与标注标准迭代：建立错误归类体系（语病性/逻辑性/安全性），针对不同场景诊断典型错误，输出标注规则优化建议，建立误判归类体系。
- “RAG多模态场景优化：主导构建RAG参考资料训练集，针对数据信息不一致等问题进行标注与改写，提升模型调用数据库信息的能力。

- “高风险多模态语料处理：批量生成图文/视频高质量泛化样本，训练覆盖率提升10%；专项处理数千条涉台、涉港图文/音视频文本，政策贴合度100%；按月维护同义词库，持续迭代风险词库。
- 模型评测体系：参与设定文本训练集，完成模型输出评分与排序，为优化提供定量依据。
- “多模态智能体对话标注：拆解图文/音视频多轮对话中的用户意图与上下文依赖，构建逻辑链支撑智能体决策。独立建立并标注视觉-文本多轮对话数据集，为模型在复杂多模态场景下的应用提供核心数据支持。”
- 运用提示词工程（PE）赋能提效：针对项目动态需求，持续优化并撰写核心标注环节的Prompt指令。成功开发的格式改写模板，显著降低了新成员对Markdown等工具的学习门槛，使新人无需深入学习即可快速上手试标任务，极大加速了团队的整体培训与上手效率。

**2024.03 - 2024.08**

**金山办公**

### 数据标注实习生

语音数据校对：5个月内完成1万条语音数据校对，通过精细化标注将指令文本准确率从82%提升至93%。

- “多模态标注规则迭代：参与撰写与优化视觉理解SFT方向AI训练项目的标注规则，通过汇总图文/音视频错误案例、补充边界规则，推动规则覆盖度与团队执行效率提升。
- 标注规则迭代：参与撰写与优化多项AI训练项目的标注规则通过汇总错误案例、补充边界规则，推动规则覆盖度与团队执行效率提升。