

邓雅琪

yaqidengtravail@gmail.com | <https://dengyaqi.github.io/about/>

教育背景

法国索邦大学(Sorbonne Université)	2017.09 - 2019.10
语言与计算机, 全日制, 硕士, 主修 NLP 和 Java。中国留学网证书编号: 120210282897。	
法国巴黎第七大学(Université de Paris 7 Denis Diderot)	2015.09 - 2017.09
计算机语言科学, 全日制, 本科, 主修 NLP 和 Java。中国留学网证书编号: 120210287879。	
ESGCI(École de Commerce International et Marketing à Paris)	2014.09 - 2015.09
金融语言预科, 全日制。	
深圳职业技术大学	2010.09 - 2013.06
应用法语, 全日制, 专科。	

工作经历

高级 Java 开发工程师	北京汉克时代科技有限公司 - 杭州, 中国	2023.06 - 2024.10
工作内容: 在蚂蚁集团网商项目组, 负责一站式策略研发、行业风险管理及风险策略等平台的后端开发。通过深入的需求分析、系统架构设计、编码实现及测试, 提升系统整体效率, 同时在团队内部提供技术指导, 与项目组保持高效沟通并优化工作流程。		
安卓软件开发工程师	浙江清华柔性电子技术研究院 - 杭州, 中国	2021.11 - 2022.12
工作内容: 负责多个医疗方向安卓软件的全栈开发, 涵盖软件设计、编码调试及发布后支持, 确保软件高效迭代和稳定运行。		
自由职业者	巴黎, 法国	2019.10 - 2021.08
工作内容: 负责 NLP 算法开发、Java Web 应用设计及安卓软件开发, 通过高效的技术解决方案推动了项目交付和提高客户满意度。		
自然语言处理工程师	Yseop - 巴黎, 法国	2019.04 - 2019.09
工作内容: 负责知识图谱项目研究, 包括知识本体融合算法的研发与验证, 并完成 Lexicon 网站的全栈开发, 显著提升检索准确度。		

后端项目经历

一站式策略研发平台 I 期(bkpolycystudio)、行业风险管理平台(bkindustryassert)等项目

项目简介: 为简化研发过程和提高效率, 整合多平台功能并形成闭环, 实现了高价值研发方案的沉淀和业务流程的优化。

责任模块: 负责文件编译器和目录树的功能开发与重构、性能优化与遗留问题修复。

工作业绩: 模块重构: 通过重新设计架构、并行或异步处理、安全性改进等方向重构代码, 提高代码质量和运行效率;
性能优化: 基于数据库表优化、合并网络请求、性能监控和分析等措施确保系统稳定性和用户体验;
漏洞修复: 设计并实施修复方案, 处理系统权限问题导致的网站崩溃和登录故障, 确保修复措施不引入新问题;
多平台开发: 采用外观模式对接新老平台功能, 开发多版本 Facade 的功能, 并确保接口兼容性与可持续性开发。

高性能商城项目

项目简介: 基于 Java SpringBoot 开发了包括商品与订单管理、用户权限、搜索等功能的商城系统, 并实现了高并发秒杀模块。

项目内容: 秒杀设计: 基于 RocketMQ 控制订单提交, 通过 Redis 做热点缓存与分库分表以实现降低数据库压力;

搜索优化: 基于 ES 与 Redis 实现热门搜索词的二级缓存, 并通过 Tag-Based Recommendation 智能推荐商品;

存储冷热分离: redis 缓存热门图片, 实现智能置换淘汰。图片压缩及 lazy 加载, 降低客户流量冲击;

适配普通机型: JVM 调优限制内存占用、避免进程 OOM, 调优 GC 策略在性能和内存占用取 tradeoff;

部署运维: 通过 Kubernetes 实现自动化部署和运维, 提升了系统的可伸缩性和管理效率。

RAFT 分布式系统项目

项目简介: 完成一款具有高可用、强一致性的分布式 Key-Value 存储系统。

项目内容: 论文复现: 根据 Raft 论文, 实现 Raft 一致性共识算法, 具备 Leader 选举, 日志复制, 快照更新等功能;

数据迁移: 基于 Multi-Raft 架构, 对数据进行 shared 分区, 并且可以在多组 Raft Group 中迁移;

优化存储接口: 线性一致性的 key-value 存储接口, 通过异步 Apply、FollowerRead 和 Prevote 等进行性能优化;

高性能分布式事务: 实现了 MVCC 多版本并发控制, 并基于 Precolator 模型, 支持高效分布式事务。

后端技术栈

编程语言: Java、C++。

相关技术：系统架构设计、分布式开发、JVM 调优、ES 优化、MySQL 性能调优、库表设计、并发编程、Redis、RocketMQ、Git。
相关工具：IntelliJ IDEA、Kibana、Visual VM、Postman、Visual studio、Xampp、StarUML。

移动端项目经历

运动心电衣项目

项目简介：基于安卓原生实现一款实验室人员使用的连接心电衣设备与云后端实时通信的安卓移动端软件。

责任模块：安卓原生全栈开发。

工作业绩：UI 设计：严格按照 UCD 与产品定义的需求，基于 Figma 进行界面的交互设计，输出终端界面视觉设计图；

软件设计：整体框架使用 Jetpack 组件构建 MVVM 架构，包含实时监测、缓存、用户登入、权限管理、蓝牙操作等模块

前端开发：根据交互设计图实现自定义 View 页面架构和布局，基于 MPAndroidChat 绘制心电图。

设备交互：基于 nordic 库与自定义蓝牙协议交互，基于 Rxjava 异步回调监测实时数据、基于 greenrobot 实现本地数据库。

云端交互：基于 Retrofit+OkHttpClient 请求云计算资源和上发设备实时数据。

人工关节术中与术后项目

项目简介：完成两款软件帮助医生与实验人员对于人工关节设备的研究。1) 基于 Matlab 实现术中实时检测关节移动轨迹与数据上传云端。2) 基于安卓原生实现一款 APP，可实时链接嵌入式设备，术后监测关节压力并具有报警系统。

责任模块：安卓原生全栈开发，在运动心电衣 APP 的基础上进行。

工作业绩：UI 设计、软件设计、设备交互、云端交互等模块与运动心电衣项目类似。

UI 优化：持续优化产品的用户体验；

性能优化：不同移动设备不同版本的兼容问题优化、基于 Service 保持 CPU 持续运行、防止系统回收软件；

内存优化：使用 Leakcanary 检测内存泄漏，并修复例如 viewbinding 导致内存泄漏等问题；

移动端技术栈

编程语言：Java、Kotlin、HTML5、Javascript、Typescript、CSS。

相关技术：网络编程、搜索引擎、缓存设计、应用上架、UI 设计、多媒体处理、蓝牙协议设计、内存优化、IOT 设备联调。

开发工具：Android studio、Postman、VUE、Figma、Sketch、Photoshop、Illustrator、微信开发者工具。

NLP 项目经历

多语言流行病检测系统项目

项目简介：传染病疫情的信息必须及时跟世界各国分享，以便及时围堵与消除，例如非典、Covid-19 等。

责任模块：通过信息抽取、字符细分化分析和依赖于文本体裁来实现《不受语言制约》的一个流行病检测系统。

工作业绩：数据采集：通过第三方合作与爬虫技术，采集各国第一线报道数据。

数据预处理：文章分割、模式提取、模式过滤、事件定位等获取文本体裁所需的元数据。

模型训练与优化：加入语言识别模块后，系统的准确率从 93% 提升至 96%；召回率从 94% 提升到 95%；错误率减少 1%。

日语语音识别与合成项目

项目简介：建立一个连续语音中的单词识别系统，并在 HTK(马尔科夫隐藏模型工具包)环境下对其进行验证。

责任模块：识别和定位给定的日语语音连续体中的关键词列表中的所有出现的词，并生成新的语音段。

工作业绩：语音分析：分析报道中的发言(发言顺序，识别不同人，发言现象的重叠，犹豫，连音，外来词，缩写等)。

数据预处理：识别音素在音段中为位置，使用 hmm 模型通过音素进行分割出 wav 文件，文件单位为单词。

模型训练：预处理完成的数据提供给 HTK 模型进行训练。

数据生成：生成词典(词和罗马音对照表)、生成语法(测试内容中不足的单词使用相似率最高的单词进行替换)。

语音生成：通过使用 MBROLA、Praat 软件和上述识别的模型，加上 Emofilt 添加情感，最终合成新句子。

Steam 法语评论自动摘要生成项目

项目简介：生成简约连贯的评论(好评和差评)摘要。

责任模块：识别和定位给定的日语语音连续体中的关键词列表中的所有出现的词，并生成新的语音段。

工作业绩：数据采集：通过 Scrapy 与 Sream API 采集百万条数据，并在 kaggle 下载相关数据，最终达到千万级数据量。

数据预处理：使用词向量、keyword、NER、KNN 等基本方法，对文本处理(不处理文本内的表情包，网络用语和缩写)。

机器学习：使用非监督学习进行情感、好评差评等自动分类。标签进行手动标注。

应用部署：使用 Flask、Bottle、Bootstrap 变成一个网络应用并且部署在服务器上。

NLP 技术栈

编程语言：Python、Matlab、R、PHP。

相关技术：语言学、知识图谱、统计调查与研究、数据采集与预处理、数据挖掘与分析、数据标注审核、智能问答系统开发与运维。

开发工具：Pycharm、Brat、Praat、Labelme、Power BI、精灵标准助手、Weka、Excel。

其他

项目管理：Scrum。

证书：人工智能训练师四级、DELF B2、C2 驾照。

语言：普通话(母语)、粤语(母语)、客家话(母语)、法语(商务)、英语(商务)、日语(流畅)。