16621370084 | super xiang@126.com | 男 | 上海 基本信息

贯: 宁 出生年月: 1994.04.19 现居地: 上海

信: lx81615836 民族: 汉族 博客: code.superxiang.com

技能标签: Java、Python、Golang、AI、运维 参加工作: 2016.6 最新项目更新至: https://me.superxiang.com

个人技能

熟练使用 Spring Boot、Spring Cloud、Alibaba Cloud、SpringAI 等相关框架以及 Spring Cloud 生态。

熟练使用 MySQL、Oracle、H2、PostgresSQL 等关系型数据库,MongoDB 与 Redis 等非关系型数据库。

熟练使用 Docker、Podman、Docker-Compose、Docker Swarm 与 Kubernetes 等容器化技术。

熟练使用 WebSocket、SSE 通信协议与 Netty 通信层框架,熟悉即时通讯系统及相关技术。

熟练使用 MyBatis-Plus、Spring Data、JPA、SQLAlchemy、Django ORM、Gorm 等 ORM 映射框架。

熟悉使用 ActiveMQ、RocketMQ、Kafka 和 RabbitMQ 等相关消息队列。

熟练使用 Python 语言开发,QT、爬虫、Selenium、Flask、FastAPI、Django 框架等 Web 框架。

熟练使用部分机器学习算法、yolo深度学习框架,GAN 对抗生成网络、RNN 神经网络等相关算法。

熟练使用大语言模型开发、RAG知识库开发,API开发以及文生图大语言模型SD与MJ的开发与部署。

熟练使用 Golang 语言开发,Gin、Beego 等 Web 框架,了解 Go-zero 框架。

熟练使用 Vue、React、HTML、JS 等前端开发语言,了解 Next.js 框架与 TypeScript 语言。 熟练使用 Linux、Unix 系统与各类发行版开发与部署,以及 Nginx 配置、CloudFlare 域名管理等 DNS 相关技术。

熟悉 PyTorch 开发包、Transformers 大模型架构以及大语言模型部署与微调。

了解 PHP 语言开发,包括 ThinkPHP 框架以及 PbootCMS 开源内容管理系统。

开发经验

宁夏如云网络科技(创业经历)

2023年7月——至今

该公司为在家乡创业时创办的公司,所有项目皆为本人搭建或开发,域名皆可正常访问。

使用 2H4G 的 Centos7.9 服务器, 通过 Swap 分区扩展内存、以满足服务的需求。

网站全域开启 Https 证书访问, Docker Compose 容器化部署, 使用 Python CertBot 证书机器人申请 Let's Encrypt 的 SSL 证书,通过 Aliyun-DNS 插件实现自动化托管。采用 Cloudflare DNS 托管实现 CDN 服务,并通过 Nginx 负载均衡配置。

自媒体数字人平台(主要负责人): Golang + Python + RuoYi (SpringBoot) + Vue

该项目为接到的外包项目,后端使用 Golang gin 框架开发,通过用户选择的性格标签,使用通义干问 API 生成用户数字人 人设介绍,同时使用用户上传的视频通过 gptsovits 克隆用户声音,将生成好的音频调用飞影 API 生成数字人视频,并将生成后 的视频, 转存到平台供用户下载。

职责:负责整个后端架构设计以及开发,并负责前端开发的指导

亮点:对于用户登录状态,通过 Jwt 生成 Token 并使用 Redis 管理在后台,针对前端使用 SeessionId 管理用户登录控制, 完成了 token 对于前端的透明化以及防篡改性,提高了平台的健壮性以及安全性。

对于平台耗时较久的操作使用协程进行异步处理,提升了用户体验,并节约了性能。

深度学习识别交通标志: Python + yolo5 (https://github.com/loks666/yolov5-5.0)

该项目为使用 yolo 深度学习框架训练的识别交通标志的项目,使用的数据集为 LISA,成功识别标志会标识标志的作用以 及识比赛别成功率,并可接入树莓派进行实时识别。

大模型知识库: 智谱知识库、DifyAI、FastAPI

使用知识库数据在智谱 AI 创建知识库,使用 Java+SpringAI 框架调用创建的知识库,共开发了三个接口,分别为:单次对 话,http 多轮对话,SSE 协议流式对话,DifyAI 和 FastAPI 为使用开源项目+中转 API 的方式进行数据的拆分并向量化,以此 实现知识库的建立与数据的索引。

自主开源 get jobs(https://github.com/loks666/get jobs) Java + Selenium

基于 Selenium 实现的简历分析投递项目,可以定时执行,接入 AI 分析等功能,目前 900+ star,目前该项目已成为 Github 同类型仓库排行第一的项目,该项目所有功能皆由本人开发,并同时维护相关的社区。

图书管理系统(https://github.com/loks666/ruyun library): Python + Django + MySQL

该项目使用 Django 框架开发的图书管理系统,前端使用 Bootstrap,后端使用 EasyUI,主要负责图书信息的添加、修改、 组合查询和删除以及后台管理等功能。

■ 微博與情分析(https://github.com/loks666/weibo-analyse): Python + FLask + MySQL

该项目通过 selenium 爬虫获取微博评论的信息,并对数据进行分析,检测舆论与热点情况,并进行图表与词云图的绘制。

■ 运动会报名管理系统(https://github.com/loks666/sports meet): Java + SpringBoot + Mysql

该项目使用的 SpringBoot 与 Vue 框架开发的运动会报名系统,分别有报名管理,用户管理,比赛管理以及自动分组等功能。

■ 机器学习识别视频语义分割(https://gitee.com/lok666/video): Python + FastAPI + NLP + MongoDB + FFmpeg 该项目为使用深度学习模型 DeepSpeech 的中文模型将视频转为文本,同时使用 NLP 进行语义分析,并使用 MongoDB 保 存 NLP 识别到的实体与语义关系,同时调用科大讯飞语音转写 API 获取视频的时间戳,使用 Python 调用 FFmpeg 命令,对 原视频根据 NLP 分析的结果进行分割子视频,将分割后的视频保存并将路径返回页面,进行展示,完成整个流程的实现。

■ Fly 助手【ai.ruyun.fun】: Golang & Gin + Vue + Mysql + Redis

该项目为购买的 Golang 编译成的二进制程序包,通过 API 调用聚合市面上主流的大模型,如 GPT3.5、GPT4、讯飞星火大模型、阿里通义干问、清华智普等。其中 GPT4 模型为节省成本采用的是逆向 Copilot 接口,即通过抓包的方式获取 Github Copilot 的 Device ID,使用 OneAPI 转换为 API Key 调用的方式接入。尝试过逆向其他平台例如:Claude2、ChatGPT、Qwen 等,但目前主流的方式为获取 Access Token 后调用,考虑到稳定性与时效性放弃了类似方案。同时网站遭遇过三次黑客攻击,积累了安全相关的部分知识与经验。

自主接入功能: QQ 登录、微信公众号登录、Github 登录、支付宝当面付与网站支付、邮件发送、腾讯云短信验证码接入等。

■ 如云科技官网【www.ruyun.fun】PHP7.5 + Mysql + Docker

使用 PbootCMS 的网站模板,二次开发后打包为 Docker 镜像推送到 DockerHub 中央仓库使用。并通过域名授权方式进行访问控制。

■ 技术博客与 Fly 助手资料站【code.superxiang.com | | aiblog.ruyun.fun】

该博客使用开源博客技术框架 Halo,使用 JDK17 & SpringBoot + Mybatis + Mysql + Redis + Thymeleaf 架构实现,通过插件驱动的方式实现各类模板,附件管理,博客管理等功能。

■ One API [one.ruyun.fun | | fly.ruyun.fun | | new.ruyun.fun]

该项目采用 Golang & Gin + React + Mysql + Redis 技术栈。主要目的是将各类大模型 API 调用的方式聚合到同一系统内,并实现密钥的分发与多渠道管理。同时,配置支付模块可以自定义渠道价格与速率。其中,OneAPI 源自 songquanpeng/one-api,经过二次开发后改为了 FlyAPI 版本 loks666/fly-api。后发现 Calcium-lon/new-api 的 fork 版本更为优雅,故目前主要使用 New api 版本。三版本同时运行,数据互通,仅前端不同。目前聚合了阿里通义干问、讯飞星火 v3、ChatGPT、Copilot 等主流大模型。以供 Fly 助手以及分发下层 API 渠道使用。

◆ 如云发卡平台【pay.ruyun.fun】Python + VUE3.0 + FLASK + Mysql

此项目为基于 Baiyuetribe/kamiFaka 构建的卡密发卡系统,通过集成支付宝当面付,管理自定义兑换码以及账号类商品的自动发货卡密发卡系统。

■ mall-swarm 在线商城系统【JDK8 & AlibabaCloud + Elasticsearch + RabbitMq + Mysql + Kubernetes】

该项目是一套微服务商城系统,基于 AlibabaCloud 构建、使用 Elasticsearch、Docker、Kubernetes 等核心技术,同时提供了基于 Vue 的管理后台搭建系统。集成了注册中心、配置中心、监控中心、网关等系统功能。

该项目由于服务器资源紧张,故只部署了本地环境,初期使用 Rancher 部署,但由于 Rancher 在 2.5 版本后受 bug 影响部署过程中问题较多,故使用 Kompose 将 Docker Compose 配置文件转换为 deployment 配置文件使用 kubectl 部署,后在终端使用 k9s 客户端管理,客户端使用 k8s lens 管理。

■ 大模型的本地部署与微调

电脑环境为 2080Tl 11G + 9700k,内存为 32G,原生部署了 Stable Diffusion、ChatGLM-6B 与 Qwen-VL-7B 大模型,其中 SD 使用 WebUI 整合包部署,civitai 的模型进行训练,ChatGLM 与 Qwen-VL 部署了 web 版本与 API 版本,其中 ChatGLM 额外部署了 Cil 终端命令行版本,而后恰逢 Gemma 模型发布,发现了 Ollama 框架与 open-webui,故使用 open-webui 为前端,Ollama 框架作为后端集成部署了 Gemma7B 模型与 LLaMA2 模型。

■ Web版 ChatGPT [LobeChat (React) | | NextChat (Vue)]

以上两项目皆为通过 Next.js 构建的 ChatGPT 应用,使用 API 调用各类大模型接口,同时搭配 One API 实现聚合调用。但由于上线后出现了盗刷问题,故停止了后台配置 Key 的方式,而采用用户页面配置方式。

■ 项目生成 ChatNio 【chatnio.ruyun.fun】 Golang & Gin + React + Mysql + Redis

该项目为一个 AI 聚合聊天平台,主要特色功能为 AI 项目生成器,通过 Prompt 提示词的配置实现文本生成 app 的 AI 化自动构建,具有订阅和 Token 弹性计费系统,Key 中转服务,多模型聚合支持等。

■ 开源的聊天系统【webchat.ruyun.fun】JDK17 & SrpingBoot + Mysql

此项目为之前开源的系统 niubaide im 的重构版本 webchat, 优化了界面, 新增了好友管理等功能。

■ 网站应用导航【guide.ruyun.fun】Golang & Gin + Vue + Mysql

该项目基于开源网站导航项目 WebStackPage.github.io 二次开发版本 webstack-go, 使用 Golang 作为管理端后台并构建为 Docker 镜像在 Docker Hub 使用。

■ 微信接入 ChatGPT

初期使用的是 chatgpt-on-wechat,该项目为 Python 开发,通过 Linux Wechat Web 的方式实现微信的接入,并使用 Chatgpt 回复应用消息,但由于该项目消息处理延时较高,且不支持 Dalle 模型,故与朋友合作开发了该项目的 Golang 版本 gpt-wechat-bot,使用 openwechat 包实现了 Wechat 的热登录方式,消息推送以及文生图功能,其中文本消息使用 ChatGPT 处理,图片功能使用逆向 Coze 应用(字节跳动大模型平台)托管的机器人,通过 Discord 平台的频道 Token,推送应用消息实现文 生图功能,以及通过链接回传,微信下载的方式回传生成的图片。

■ 服务监控管理与备份【mon.superxiang.com】

服务管理使用的为 uptime-kuma,通过轮询请求的方式监测各服务的存活状态,并将结果以看板形式展现。 使用 Python 脚本进行备份,使用 pyby 包上传备份文件,同时使用 cron 定时任务每天自动备份。

上海富数科技 (客户交付部-高级 Java 开发工程师)

2021年8月——2023年6月

该公司下主要分为两个平台,Avatar 为基础应用平台,主要分为用户管理、隐私数据管理、密钥管理、算法管理、组件管理等模块。Fmpc 为 Python 开发的联邦学习算法平台,对 Avatar 进行算法支撑。

主要职责: Avatar 平台的功能开发与迭代,对 Fmpc 平台的接口的支撑与接入提供支持。

■ Avatar 安全计算平台 (JDK8 & SpringBoot + Mybatis + Mysql + Redis + Docker)

多方安全计算平台,用于对用户数据加密进行安全计算,基于 Java 平台搭建,与 Fmpc 算法平台交互,采用 Redis 及 FastDfs 进行数据同步,使用环境变量传递的方式保证信息的安全性。实现安全多方计算、联邦学习、匿踪查询等实现分布式 隐私计算方案。

主要负责:平台的常规迭代,新功能的开发与改造等。

主要任务: Avatar 平台的计算流程的优化,算法可插拔方案的设计与实现,文件管理架构改造升级等。

■ Fmpc 算法支撑平台 (Python3.6 + Flask + Redis + FastDfs)

对 Avatar 平台提供算法支撑。支持逻辑回归、随机森林、神经网络、样本对齐等算法形成虚拟融合样本集,对各方中间结果交互计算,多轮迭代产生训练模型。

主要负责: Fmpc 平台接入 Avatar 平台的技术支持,对接接口的开发与联调测试等支持。确保两个平台之间的数据传输和功能交互的顺利进行,以及保证系统的稳定性与可靠性。

上海亿企赢网络科技有限公司(技术部-Java 后端工程师)

2020年5月——2021年6月

■ 上海电子税务局即时通讯系统 (SpringMVC + Bonde 框架(自研) + Netty 4.0.15 + Oracle)

该项目为在上海电子税务局系统中,根据管理人端的组织结构,开发的即时通讯系统,降低了税务人员内部沟通的成本,同时保证了信息的安全性,为税务人员提供了方便内部沟通的平台。

职责描述:负责评估和选择适合即时通讯系统的技术栈和框架。并对即时通讯系统的各项功能进行详细设计,并实现相应的技术方案。开发实时消息传递、群组聊天、消息推送等功能模块。后负责测试和生产环境的部署的支持工作。在应用上线后,着手设计和实现即时通讯系统基于 Zookeeper 管理 Netty 集群的高可用方案,但高可用架构在进入开发阶段时由于开发资源不足故未投入使用。以单体应用方式发布。

成果:在开发期间解决了 Netty 的死锁问题,设计了整个系统各个阶段的架构与功能,并在后期负责系统内 Oracle 数据库向国产数据库金牛座的架构升级。

■ 上海保税区税收预测平台 (SpringBoot + Mybatis-Plus + Oracle)

该项目针对上海保税区税收情况,分析并预测保税区各行业的税收情况,通过系统历史数据的整合与预测,生成分析报告, 经过实验测试,该系统的成功实施显著降低了税务人员的沟通成本,减少了不必要的资源投入,为客户方节省了宝贵的资源和时间。

职责描述:负责该系统中报告的生成与管理、各税种的统计与展示、系统内的数据处理以及框架组件的搭建。并负责系统上线后的需求变更和修改性开发工作。记录和维护开发文档和代码库,以便后续维护和团队知识共享。

上海讯联数据服务有限公司(研发部-Java 后端工程师)

2016年6月——2020年5月

■ 平安银行收单系统 GP 商户处理平台 (SpringMVC + Spring + Mybaits + Mysql)

该系统为面向客户及平安银行的核心系统,客户为使用公司开发的POS 机的商户,系统主要用于处理银行卡交易数据以及提供平台供应商入驻、审核、提交银行审核、放款清分和风控等业务功能。公司内其他系统皆为以此系统作为模板系统开发,并相应做出的一部分修改完成的。

职责描述: GP 商户处理平台的放款清分和风控相关功能模块的需求开发与改造: 参与了 GP 商户处理平台的开发和改进工作, 主要为放款、清分和风控相关的功能模块。负责需求分析、功能开发、代码改进等工作,以确保平台的正常运行和安全性。

平安银行审核机制处理流程的修改:负责了系统内平安银行的审核机制处理流程的修改工作。包括对审核流程的优化和改进,以满足平安银行的要求和标准。与平安银行的相关部门合作,确保审核机制的顺利运行。

公司泰国分公司商户处理平台的开发工作:负责了公司在泰国分公司的商户处理平台的开发工作。参与了泰国分公司商户处理 平台的构建和开发,以满足该分公司的业务需求。负责整个开发流程,包括需求分析、系统设计、编码实现等。

毕业院校:上海出版印刷高等专科学校(印刷包装管理) 2013.9~2016.6

自我评价:

具有较强的学习动力和自我提升欲望,对技术怀有浓厚兴趣并喜欢钻研;

良好的表达与沟通能力,积极主动,对工作尽心尽责;

具有有良好的工作态度和团队合作精神;

热爱软件开发行业,较强的学习能力,喜爱折腾技术;