

魏晓鹏

☎ (+86) 177-2460-1381 ✉ wxp_sampson@163.com 🏠 <https://simonwei97.github.io> in [linkedin.com/in/xiaopeng](https://www.linkedin.com/in/xiaopeng)

教育背景

南方科技大学

2015.9 - 2019.6

- GPA: 3.41/4.00。通信工程，工学学士。

工作经历

商汤科技 | 软件开发工程师

2020.7 - 至今

- 负责基于 LLMs 的本地知识库问答系统**检索服务**开发工作，从 0 到 1 的技术方案设计 & 工程落地。
- 主导 SenseFoundry 产品的视图设备管理服务、视图接入服务的开发工作，包括技术方案、技术文档撰写，代码开发，组件 Helm Chart 修改。完成了 7 个版本的开发工作，并且在职期间申请 2 个行业相关专利。

商汤科技 | 测试开发工程师

2019.7 - 2020.6

- 负责公司 SenseFoundry 产品设备管理服务、视图/视频接入服务的版本测试工作，包括测试方案制定、自动化脚本开发及测试报告输出。
- 使用 Python 开发自动化测试脚本，完成产品服务的功能、业务流、准确性及性能测试，共计完成了 3 个版本的测试工作。

项目经历

SenseNova mini - 知识库检索服务 | Python, FastAPI, Minio, Milvus, TiDB

- 知识库检索服务 (KnowledgeBase Retrieve worker)，接收用户的问题，Embedding 用户问题后，在文本向量中匹配出与问句向量最相似的 top_k 个，再对 top_k 结果进行上下文召回，最后将匹配出的文本作为上下文 (context) 和问题一起添加到 Prompt 中，发送给 LLM 生成回答。其中文本向量是由前序的解析组件解析切分后的 chunk 文本，为一个文本段落中的字符。
- 服务基于 FastAPI 框架进行开发，项目支持离线私有化部署。采用 Helm Chart 部署在 k8s 集群。
- 针对不同解析类型 (PDF/Doc/Docx, QA pairs)，对 Milvus 相似度检索结果进行上下文召回，不同 chunk 文本类型有不同的召回策略，chunk 文本类型包括标题、正文、表格等。

SenseFoundry - 视图接入服务 | Go, gRPC, Zookeeper, OSS, Redis, Kafka, Prometheus+Grafana

- 视图接入服务 (IIS)，接收 SDK 相机及其他视图协议的消息，将数据结构化不同的数据归一化 (清洗) 为指定内部结构，最后写入到消息队列，由视图解析服务消费后进行特征等细节分析。
- 服务基于 gRPC，使用了 Zookeeper, OSS, Redis, Kafka/MQTT，使用 Helm Chart 部署在 k8s 集群。
- 服务采用自研 OSS 存储部分对象大图，然后将消息写入 Kafka (带小图或大图)。后续因 Kafka 大消息传输压力过大，使用 Redis 缓存较大的图片，解析服务消费后拉取缓存的图片，降低 Kafka 压力，整体服务接入性能提升 20%。
- 服务提供链路数据治理，从输入到输出各个环节进行 metrics 埋点，最终使用 Prometheus+Grafana 实现可观测性监控。
- 开发 CLI 工具，支持服务的性能测试，Mock 上下游服务及数据，运维问题定位、异常检测等。

SenseFoundry - 视图设备管理服务 | Go, gRPC, MySQL (TiDB), Prometheus+Grafana

- 视图设备管理服务 (DMS)，将相机、NVR 设备、IoT 等设备纳入系统，统一进行管理。主要功能包括设备接入管理，状态管理，设备任务管理，根据业务流程与下游组件进行交互。支持多租户 (MySQL 分表)。
- 服务基于 gRPC，使用 TiDB (MariaDB 迁移到 TiDB) 进行数据存储，采用 Helm Chart 部署在 k8s 集群。

专业技能

- 熟悉 Go, Python 开发，熟悉 gRPC 框架，GC 垃圾回收策略，GMP 调度，熟练使用 pprof 分析定位问题。
- 熟悉 Kafka 的高性能、高可用，负载均衡策略，熟悉 Kafka partitions/replicas 分配策略。
- 了解 Redis 内存淘汰策略，Redis 分布式锁，Redis 持久化。
- 熟悉 Docker 及 Kubernetes 生态，有使用 Helm Chart 部署产品服务经验；了解 k8s pod 调度策略，熟悉 k8s 服务发现；有使用 Prometheus+Grafana 完成服务可观测性的项目经验。
- 大学英语四/六级 (CET-4/6)，熟悉英文读写。本科实验室期间已在阅读外文科研文献，具备英语日常交流的能力。