

# Báo cáo tháng 4.

## Những việc đã hoàn thành

- Thêm tính năng delimited string to key-value pair (Kafka-Redis Connector)
- Tìm hiểu công nghệ Spring, viết API demo theo mô hình SpringMVC và thư viện SpringWeb, đóng gói, triển khai thử
- Tài liệu báo cáo luồng dữ liệu kiểu mới
- Set up proxy để build dễ hơn với Maven, tìm hiểu qua quá trình build Maven và viết file pom.xml
- Hoàn thành nghiên cứu về SpringBoot cơ bản và code Spring Web cơ bản với request GET và POST
- Tiếp tục tìm hiểu công nghệ SpringBoot, công nghệ kết nối API với Redis: Spring Data Redis
- (Bên lề, 30m) hoàn thành tìm hiểu và sử dụng công nghệ JMeter để đo tải, độ trễ trung bình cho request
- Tiếp tục tìm hiểu công nghệ, fix bug cho API thử nghiệm deploy trên CCAI TEST
- Triển khai toàn bộ luồng lên CCAI TEST
- Triển khai toàn bộ luồng test lên CCAI TEST, config kết nối với cụm Kafka test
- Thử tải và trễ API từ PC riêng. Đánh giá chất lượng luồng.
- Viết test case cho luồng mới. Test sơ bộ ban đầu các case đã thiết kế (anh Lộc thiết kế)

## Những kiến thức mới học được

### Về lý thuyết ML/DS

- Similarity Searching với Product Quantization (có đính kèm báo cáo mở rộng **PQ.docx** và **PQ-Presentation.pptx**)
- 2 Chương đầu của Designing Data Intensive Application, đề cập tới 3 trụ cột của các good data apps, cùng với những Data Model cơ bản cùng lịch sử, usecase của chúng. (chưa có báo cáo mở rộng đính kèm)

### Về cài đặt non-ML/DS, engineering

- Spring, DI/IoC, viết RestAPI với SpringBoot (Có báo cáo mở rộng **Spring Notes.pdf**)
- Kết nối API với Redis và query từ Redis
- Luồng dữ liệu mới Nifi-Kafka-Redis (lý có luồng và mô tả ngắn gọn công nghệ mới trong báo cáo mở rộng **Báo cáo kỹ thuật.docx**)
- Java Unit Testing (đang tạm dừng ở dependency mocking với Mockito và asserting AssertJ)
- Java Logging với Log4J (có đính kèm báo cáo mở rộng **Logging with Log4J.docx**)
- Đo tải, thông lượng và độ trễ cơ bản sử dụng JMeter. (chưa có báo cáo đính kèm)

### *Kế hoạch tháng sau*

- Theo sát project MyViettel Recommender và nâng cấp cải tiến API nếu cần thiết.
- Tạm dừng và kết thúc nghiên cứu về Similarity Searching, thay vào đó bắt đầu nghiên cứu về những công nghệ mới có tính ứng dụng cao hướng NoSQL và Graph databases.
- Mở rộng củng cố kiến thức về thiết kế/cài đặt hệ thống dữ liệu với sách Designing Data Intensive Applications.