

남궁찬

분산 시스템, 데이터 엔지니어링, 웹 백엔드 분야에 관심을 갖고 있습니다.

화려함보다는 기본이 탄탄하고 깊이 있는 컴퓨터 엔지니어가 되고자 합니다.

✉ 9511chn@gmail.com

🏠 <https://getchan.github.io/>

🐙 <https://github.com/getChan>

Experience

카카오엔터프라이즈

2020년 11월 - 현재 (1년 8개월)

웹 크롤러 개발 및 운영 (검색 도메인)

- 신규 Hadoop/HBase/Spark 운영 클러스터 구성 및 데이터 이관.
(33개 서버 클러스터. On-Prem)
- 레거시 MapReduce 잡을 Spark 잡으로 전환하여 처리 속도를 약 4배 개선
- 신규 크롤러 시스템을 독립적으로 설계, 개발, 운영 및 개선.
- 타겟팅 웹 크롤러 운영 및 유지보수. (데이터 : 1.5TB, 130만개 URL)
- 레거시 코드를 신뢰할 수 있는 코드로 리팩토링하였고 모니터링 자동화하여 시스템과 데이터 안정성을 높임.

데이터 플랫폼 / 파이프라인 개발 및 운영 (지식그래프 도메인)

- Spring Cloud Data Flow / k8s 기반의 데이터 처리 플랫폼 개발 및 운영.
- 소스 데이터로부터 지식그래프로 변환 및 적재하는 Streaming 데이터 파이프라인 개발.
- 파이프라인 내의 각 어플리케이션을 auto scaling 하는 시스템 개발

카카오

2020년 7월 - 2020년 8월 (2개월)

Data Architect (Intern).

- 파일(csv, tsv...)형태의 데이터를 RDB 테이블에 적재하는 웹 운영툴 개발
(Python, Flask)

Mondrian AI

2020년 3월 - 2020년 6월 (4개월)

Backend / Platform

- FTP, REST API 로부터 수집한 csv, json 데이터를 가공하여 RDB에 적재하는 데이터 파이프라인 개발. 데이터를 제공하는 Backend API 개발.
(Python, PostgreSQL, Airflow, Nodejs)

Skill

Intermediate

유연하게 다룰 수 있습니다.

- Java
- Spring Boot
- Kubernetes / Helm
- Apache HBase
- Apache Kafka

Basic

일반적인 유스케이스에 적용 및 운영한 경험이 있습니다.

- Apache Spark
- HDFS / MapReduce
- RDB
- ElasticSearch
- Prometheus
- Python 3

Education

인천대학교, — 학사

2014년 - 2020년

컴퓨터공학 전공

