

# 남궁찬

부산 시스템, 데이터 엔지니어링, 웹 백엔드 관련 기술에 관심을 갖고 있습니다.

화려함보다는 기본이 탄탄하고 깊이 있는 컴퓨터 엔지니어가 되고자 합니다.

5년 뒤에는 다양한 문제 상황에서 최선의 결과를 내는 데이터 처리 전문가로,

10년 뒤에는 많은 사람들의 문제 해결에 도움이 되는 데이터 플랫폼을 개발하고자 합니다.

✉ [9511chn@gmail.com](mailto:9511chn@gmail.com)

🏠 <https://getchan.github.io/>

🔗 <https://github.com/getChan>

## Experience

### 카카오엔터프라이즈

2020년 11월 - 현재 (1년 3개월)

Data Collection Platform

- 신규 Hadoop/HBase/Spark 클러스터 구성하고 데이터 이관.  
(33개 서버 클러스터. On-Premise)
- MapReduce 잡을 Spark 잡으로 변경하여 처리 속도를 약 4배 빠르게 개선하였고, 실시간 데이터 집계를 가능하게 함.
- Sitemap 크롤러 시스템을 신규 설계, 개발, 운영 및 유지보수.
- 타겟팅 웹 크롤러 운영 및 유지보수 (1.5TB, 130만개 URL)
- 레거시 코드를 캡슐화 등 신뢰할 수 있는 코드로 리팩토링하였고 모니터링 자동화하여 시스템과 데이터 안정성을 높임.

### 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, Web API 로부터 수집한 csv, json 데이터를 가공하여 RDB에 적재하는 데이터 파이프라인 개발. 데이터를 제공하는 Web API 개발.  
(Python, PostgreSQL, Airflow, Nodejs)

## Skill

### Intermediate

고급 개념 및 아키텍처를 알고 있으며, 유연하게 다룰 수 있습니다.

- Web Crawling
- Java 8
- Python 3
- Apache HBase
- Spring Boot

### Basic

핵심 개념을 알고 있으며, 일반적인 유스케이스에 적용할 수 있습니다.

- Apache Spark
- HDFS
- Kubernetes
- RDB
- Apache Airflow
- ElasticSearch

## Education

### 인천대학교, — 학사

2014년 - 2020년

컴퓨터공학 전공