

# Data Collection & Processing

## 데이터 수집

- 의료 이미지 데이터셋 확보
- 메타데이터 수집 및 라벨링
- 데이터 품질 검증

## 데이터 처리

- 이미지 정규화 및 리사이징
- 데이터 증강 (Augmentation)
- 배치 처리 최적화

## 데이터 플로우

- 실시간 데이터 파이프라인
- 배치 처리 스케줄링
- 데이터 버전 관리

## 구현 도구

- Pandas, NumPy
- OpenCV, PIL
- Apache Airflow