

Data Collection & Processing

Data Collection

- Securing medical image datasets
- Metadata collection and labeling
- Data quality verification

Data Processing

- Image normalization and resizing
- Data augmentation
- Batch processing optimization

Data Flow

- Real-time data pipeline
- Batch processing scheduling
- Data version management

Implementation Tools

- Pandas, NumPy
- OpenCV, PIL
- Apache Airflow