

## 6일차. 코디네이팅 서비스

아파치 에어플로우를 통해 코디네이팅 및 스케줄링 서비스를 실습합니다

- 목차
  - 에어플로우 기동 및 접속
  - 배시 오퍼레이터 실습
  - 파이썬 오퍼레이터 실습
  - [에어플로우 설치 가이드](#)
  - 참고 사이트
    - <https://github.com/puckel/docker-airflow>
    - <https://airflow.apache.org/docs/stable/cli-ref>
    - <https://airflow.apache.org/docs/stable/tutorial.html>
    - <https://airflow.apache.org/docs/stable/best-practices.html>
    - <https://airflow.apache.org/docs/stable/howto/operator/python.html>

### 1 에어플로우 서비스 기동

에어플로우 실습을 위한 도커 컨테이너를 기동합니다

- 최신 소스를 내려 받습니다

```
bash>
cd /home/ubuntu/work/data-engineer-intermediate-training
git pull
```

- 모든 컨테이너를 종료하고, 더 이상 사용하지 않는 도커 이미지 및 볼륨을 제거합니다

```
bash>
docker rm -f `docker ps -aq`
docker image prune
docker volume prune
```

- 워크스페이스로 이동하여 에어플로우 빌드 및 컨테이너를 기동합니다

```
bash>
cd /home/ubuntu/work/data-engineer-intermediate-training/day6
```

`docker-compose up -d` `docker-compose ps`

```
* 에어플로우 데이터베이스를 초기화합니다
```bash
bash>
./airflow.sh
./airflow.sh initdb
```

### 2. 에어플로우 아키텍처

