

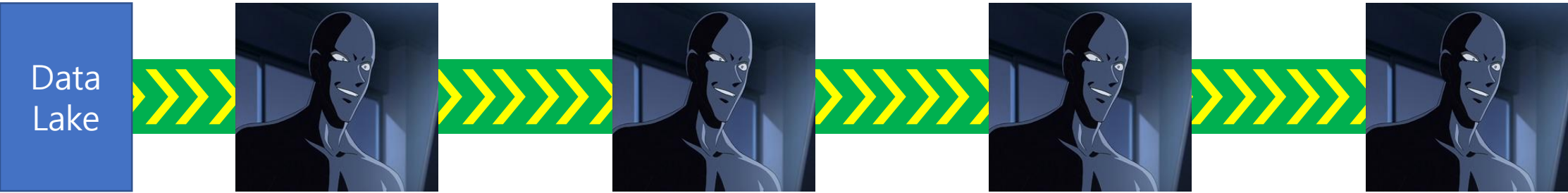
Machine Learning Workflow 구조 살펴보기

Study group <Awesome MLops>

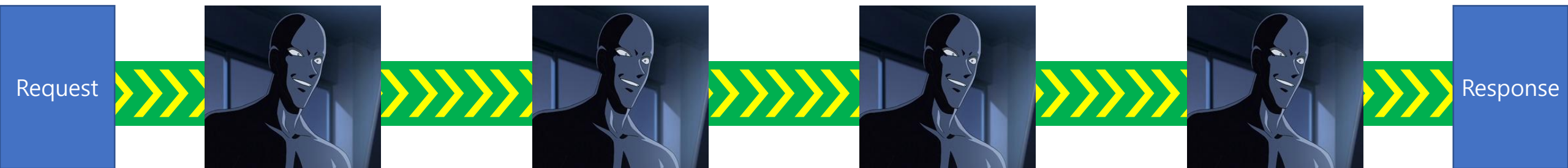
윤성국

Machine Learning 서비스를 하려면
뭐가 필요할까?

Learning workflow



Service workflow



Machine Learning

머신 러닝

학습

일단은.... 학습하는 놈이 있어야 하지 않을까?

학습하는 놈

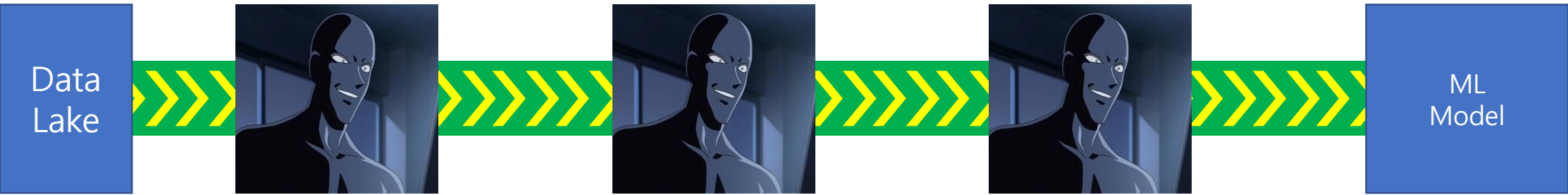
그거시 바로 **Model**

어떤 객체나 시스템, 현상, 개념을 설명하는 수단
잘 정의된 개념을 잘 정의된 구조로 잘 설명

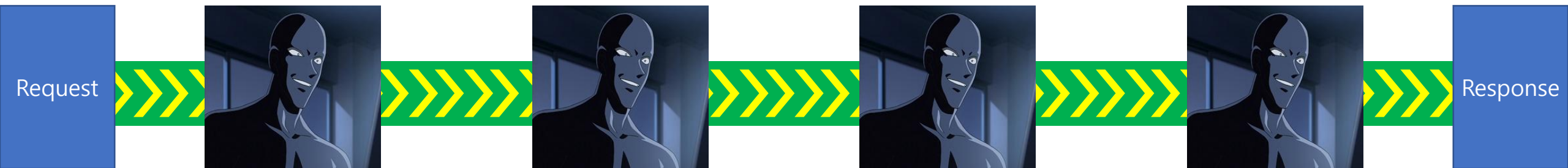
(대충 잘 만들라는 내용)

학습모델
학습을 통해 구조를 정의하여 데이터의 관계,
데이터를 통해 얻을 수 있는 정보를 설명

Learning workflow



Service workflow



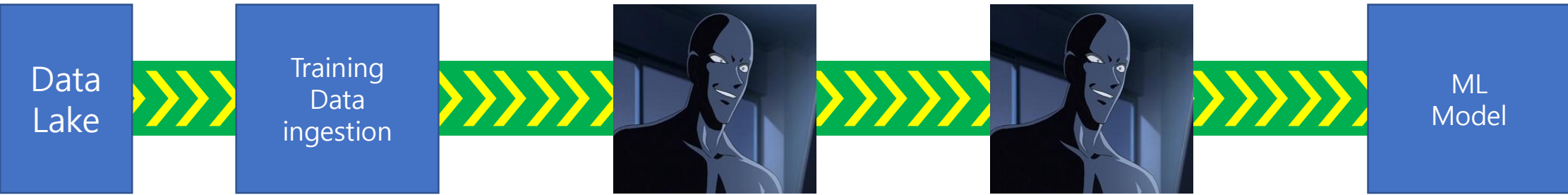
학습하려면 필요한건?

DATA



일단 먹여 놓고 갈귀라!

Learning workflow



Service workflow





근데 어떤 데이터?

~~~뿐

DATA

Yee~~~뿐 데이터를 얻기위한

preprocessing

Yee~~쁜 데이터를 얻기위한

# preprocessing

## Data Cleaning

쓸데없는 부분 제거  
노이즈 제거  
이상한 부분 수정



그렇다고 프로토스 행성정화 마냥 싹 태우면 안됨

## Feature engineering

모델이 좋아하는 형태로 변환 : Dataset  
특징 선택(feature selection)  
차원 축소(dimension reduction)

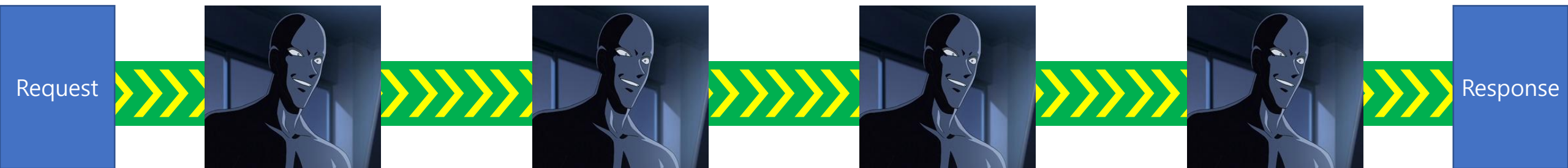


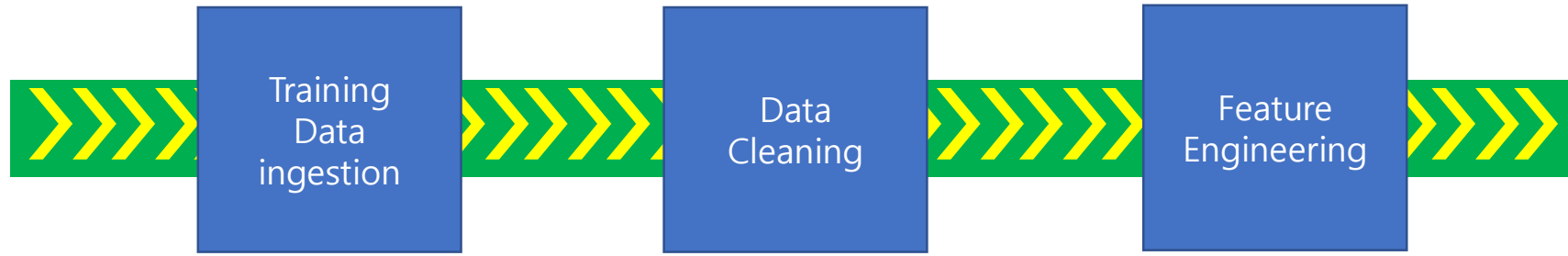
그 피쳐말고 임파

# Learning workflow



# Service workflow





Data는 처리될 때마다 계속 바뀝니다.  
모델에 따라 처리되는 형태가 바뀌기도 합니다.

이럴 때 필요한 것은 무엇인가

# Versioning

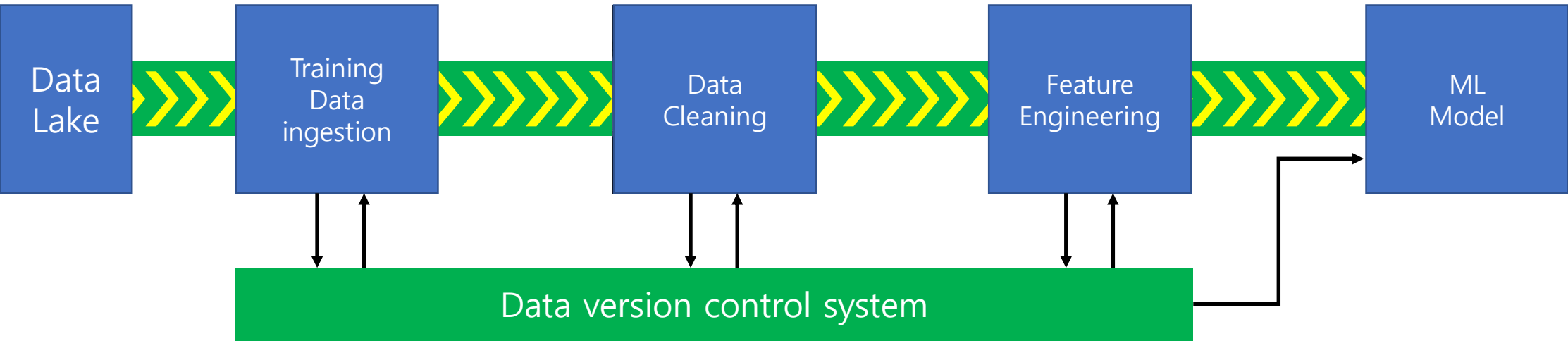


데이터도 CI/  가 필요하다~ 이



이야

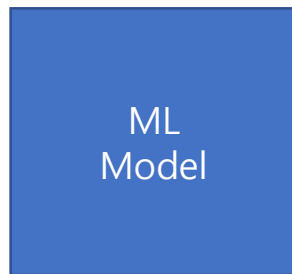
# Learning workflow



# Service workflow



우리

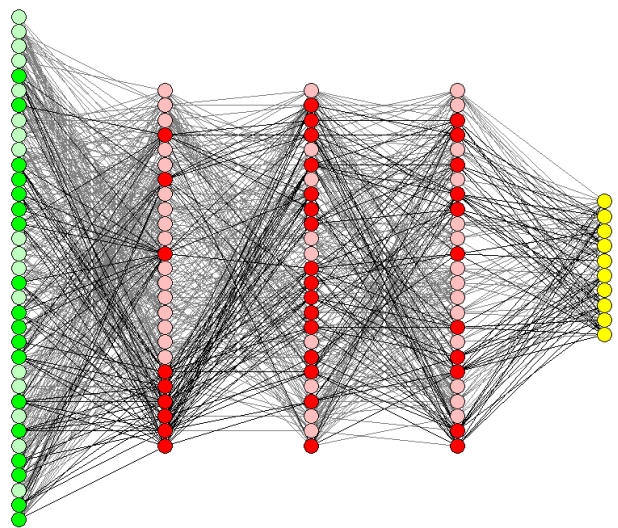


얘 이야기 좀 더 해봅시다

# Modeler

재는 (Researcher) 가 만듭니다.

그 사람들은



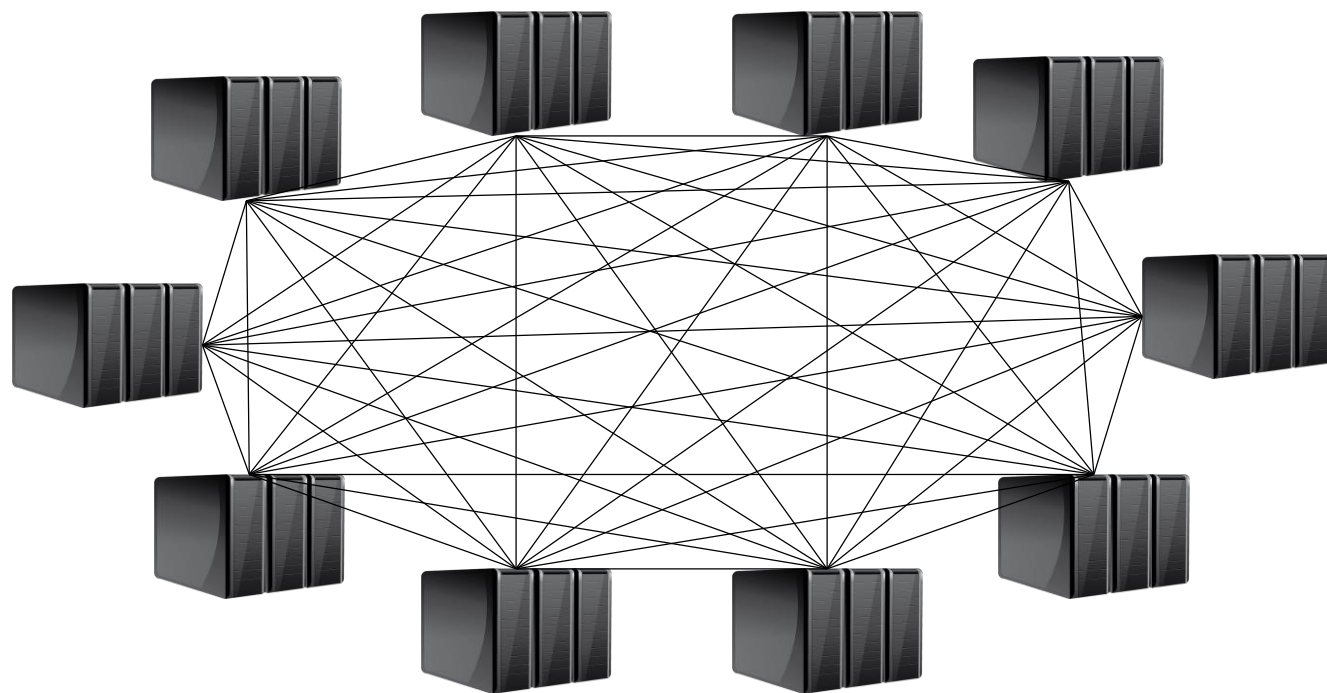
---

이런 거 설계하기도 바쁩니다.  
논문도 보고 말이죠

그런데 학습모델을 돌리려면



넘나 많은 GPU를



넘나 분산된 환경에서

운영하(고 싶어하)게 됩니다.

행복한 고민



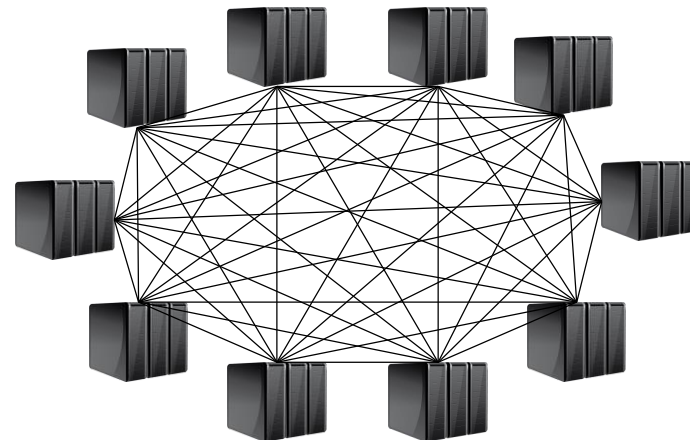
# 자 모델 알아서 돌려봐

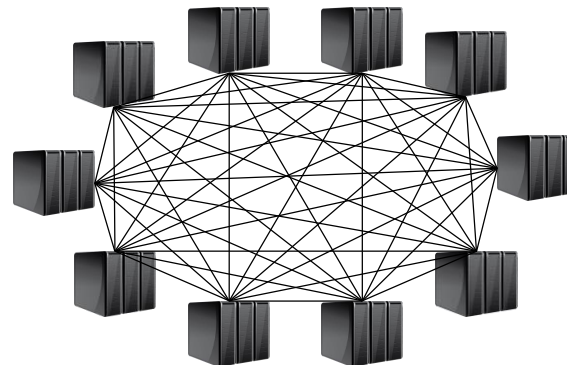


선생님 자비 좀....

나에게 주피터를 달라!

그런 건 없어





Experiment ENV platform



심심하면 커널 나가버리는 내가 싫어하는 주피터노트북 읊음

하지만 모델러가 늘어나면 어떨까?



프레임워크 고만 가져와 미친놈들아

이걸 그냥 싹 합치면?



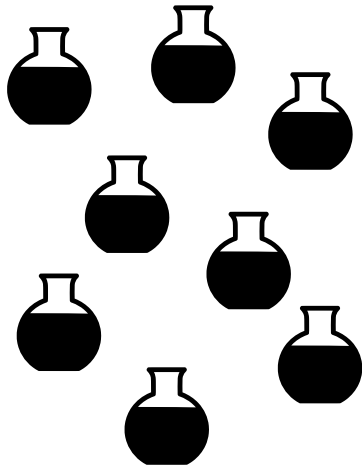
거기에 계속되는 실험을 한다면



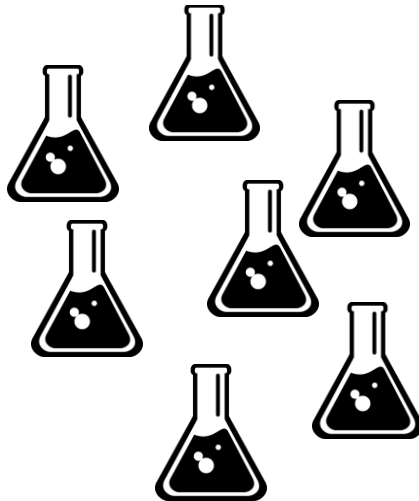
  
TensorFlow



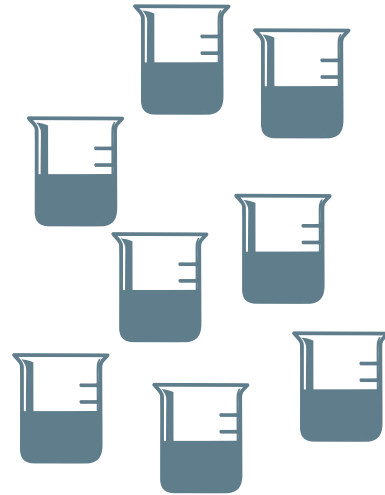
 Keras



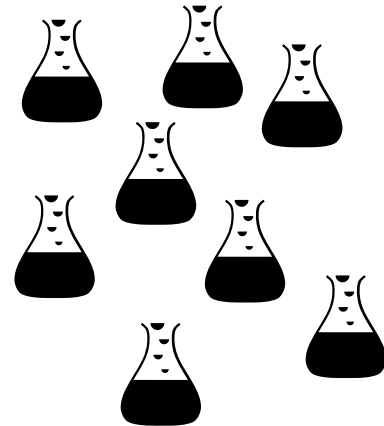
 Caffe2

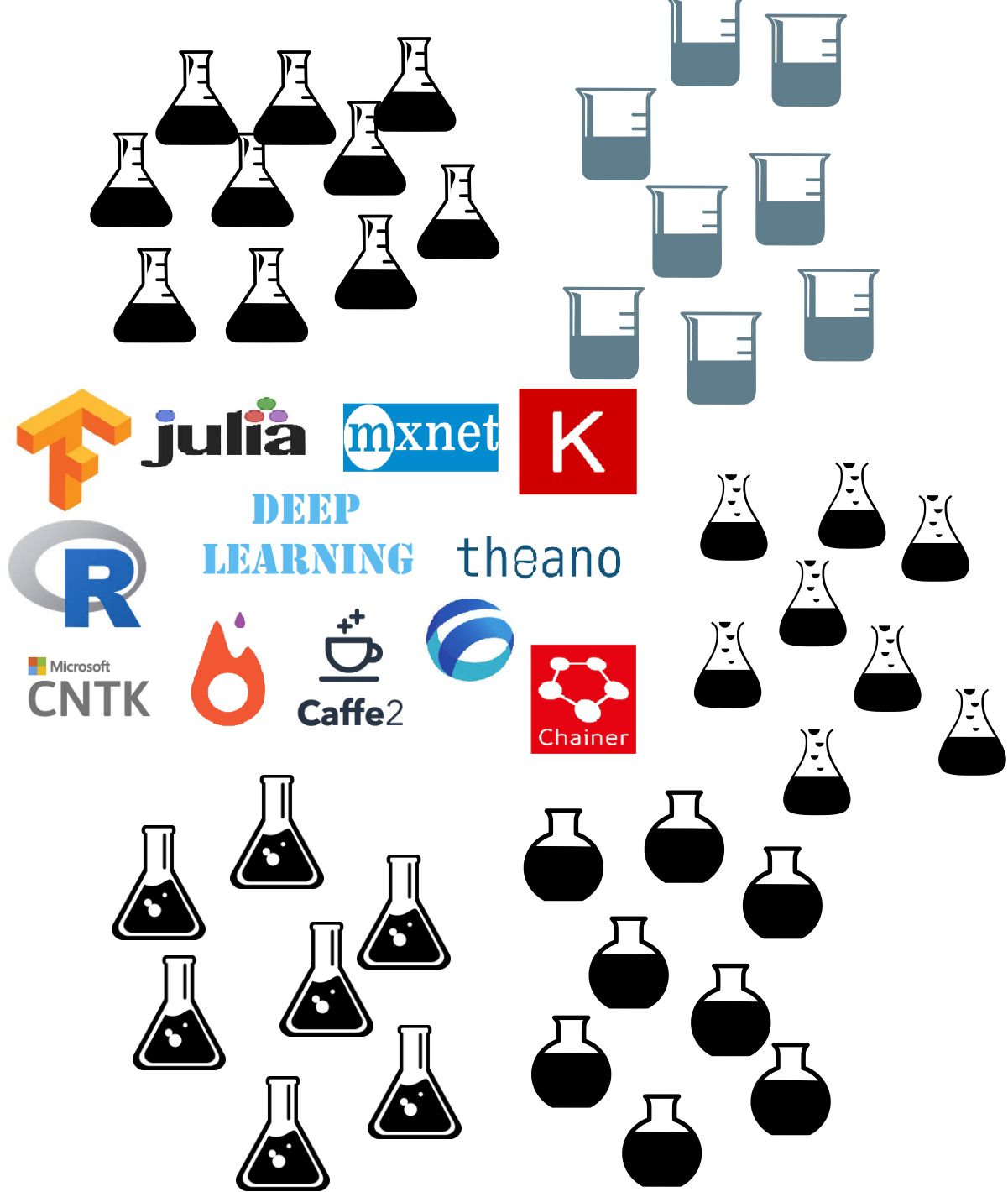


PYTORCH

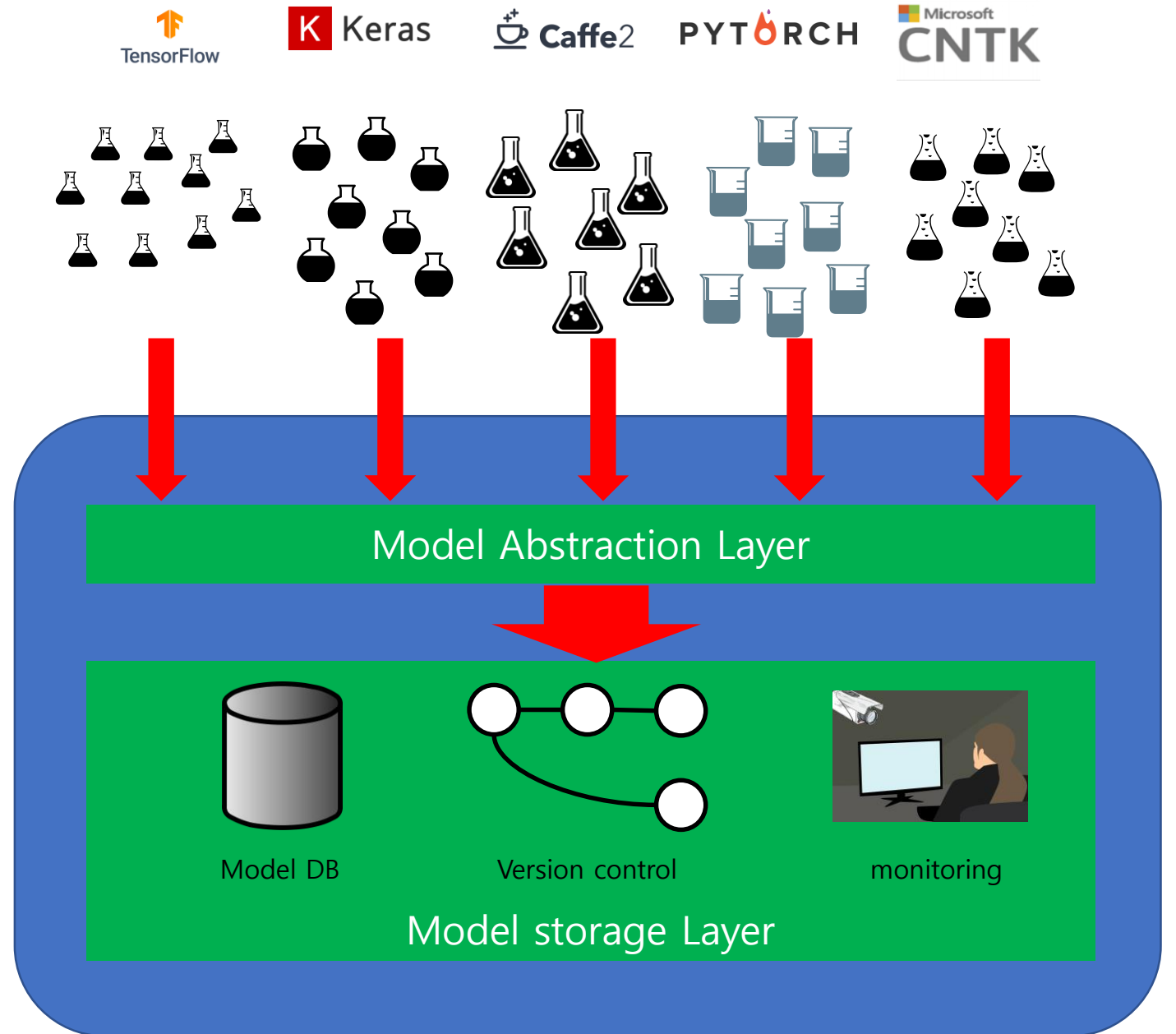


 Microsoft  
CNTK

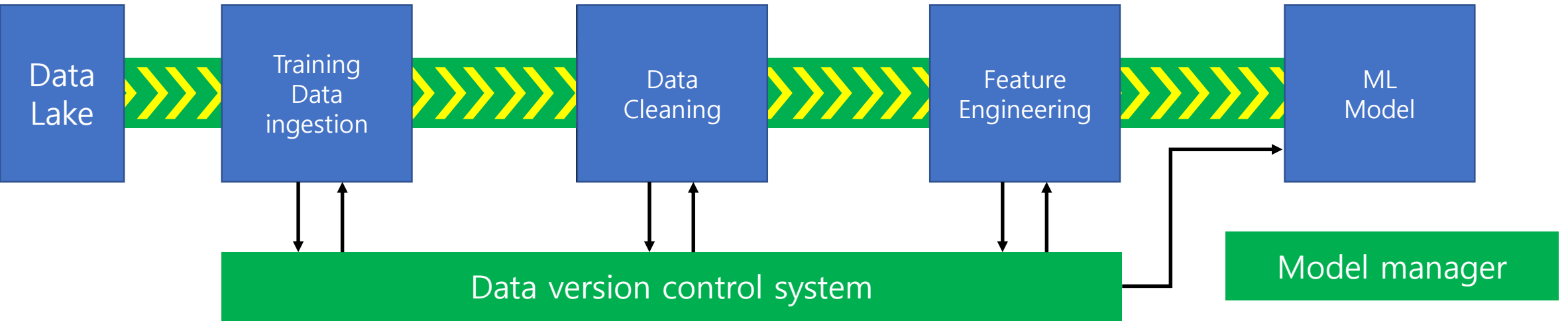




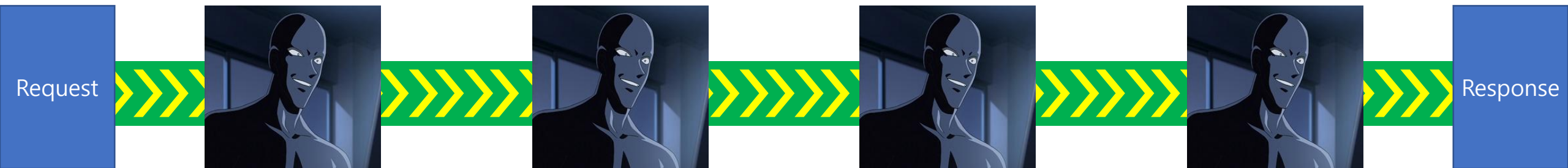




# Learning workflow

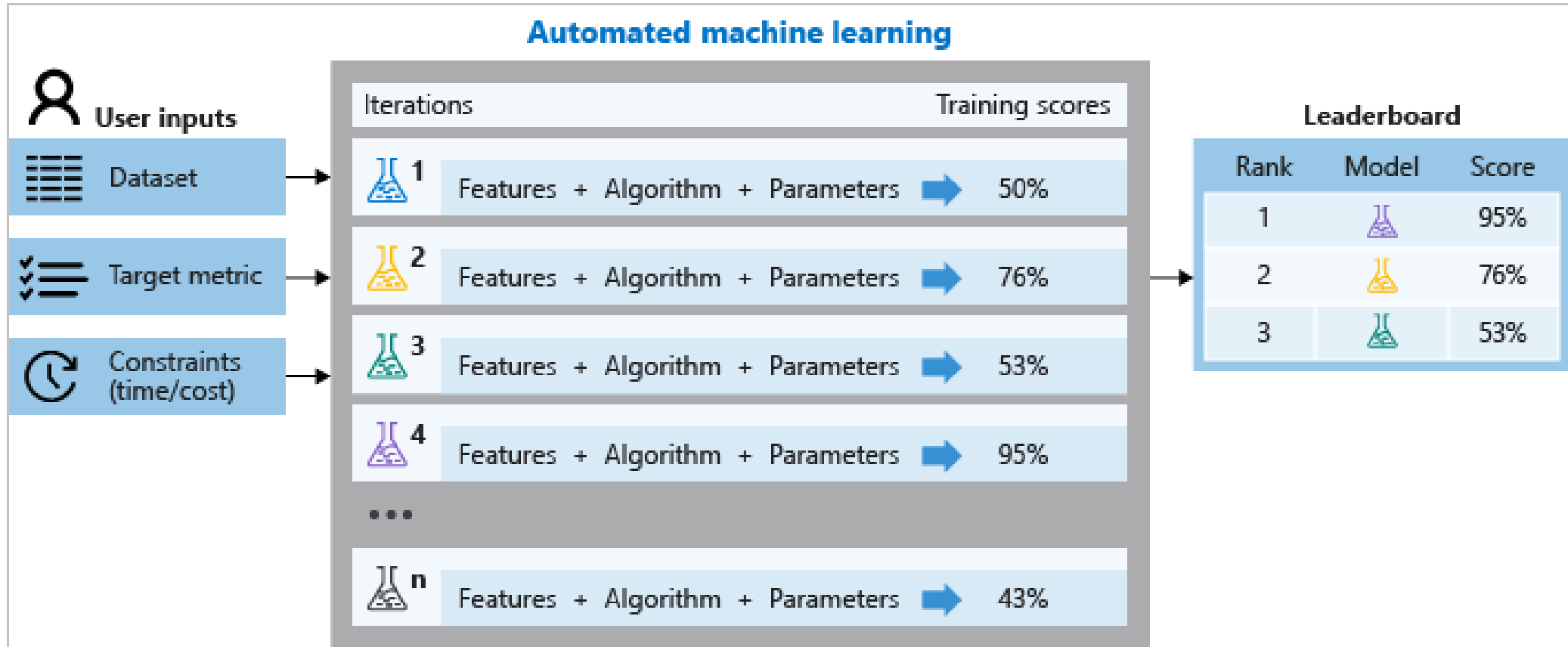


# Service workflow





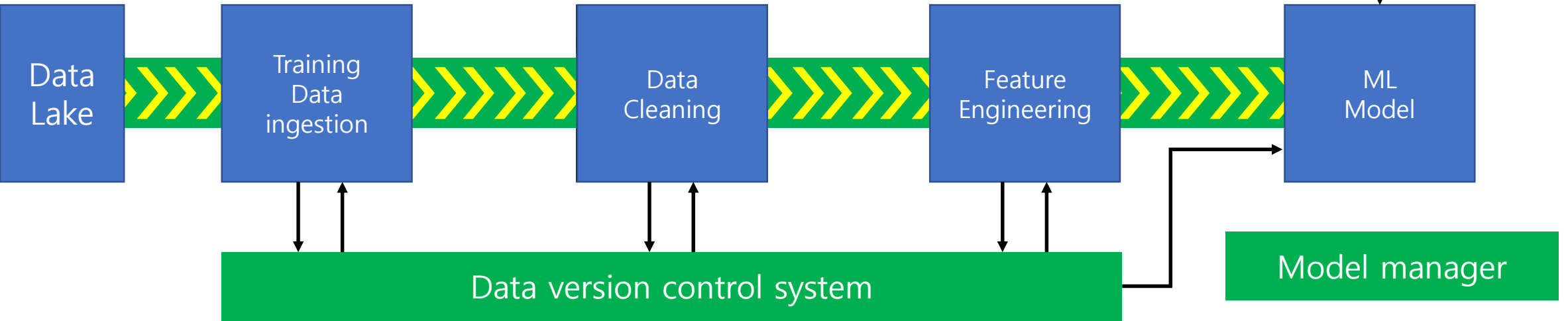
그런데 요즘은 자동으로 합니다.



# Auto ML

모델러 : AI 앞잡이가 되면 고용보장은 될 줄 알았는데 ㅎㅎ  
(제프 딘을 원망해라)

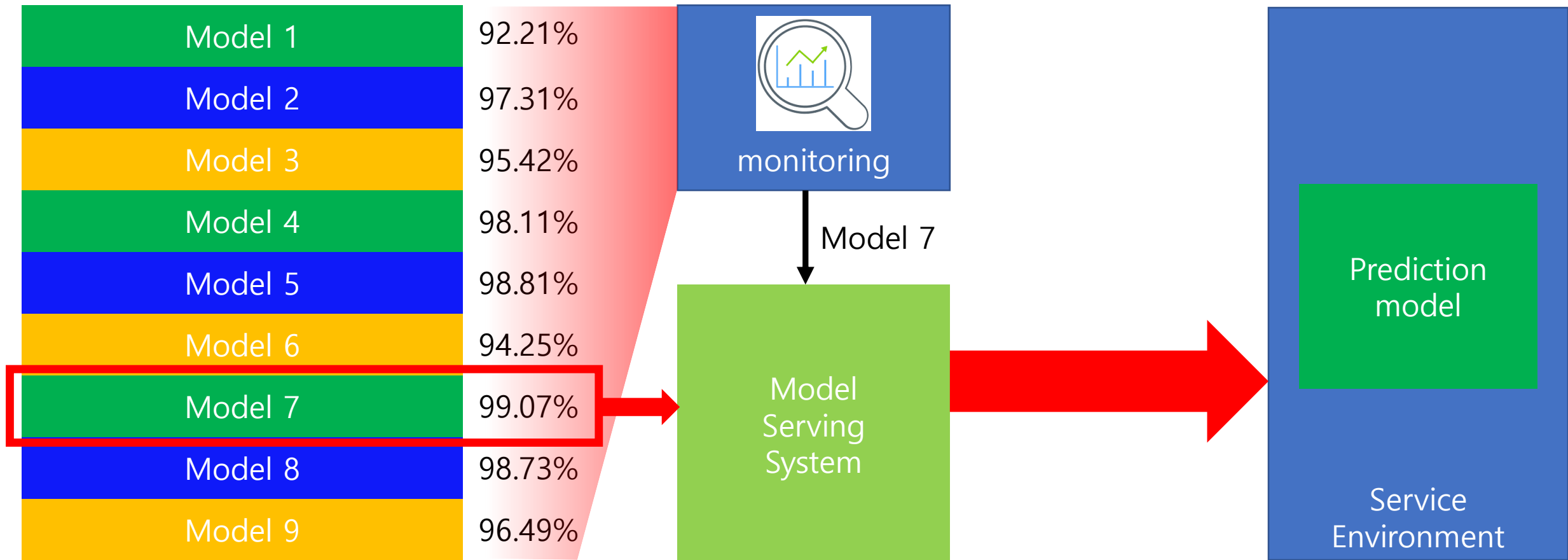
# Learning workflow



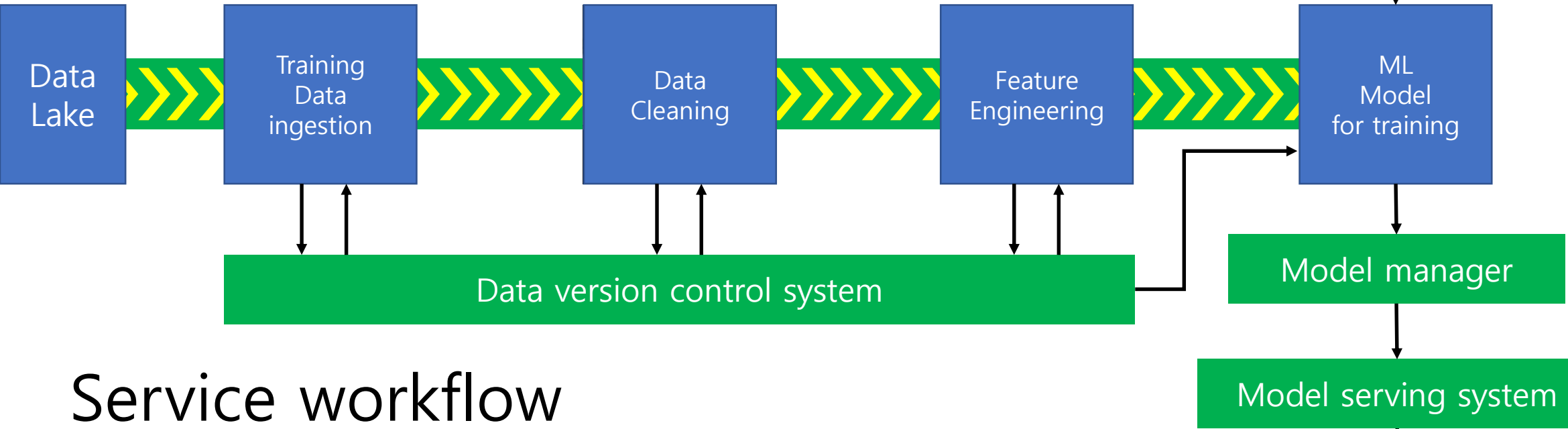
# Service workflow



모델 서빙(모델배포)도 자동으로 해야 합니다.



# Learning workflow



# Service workflow



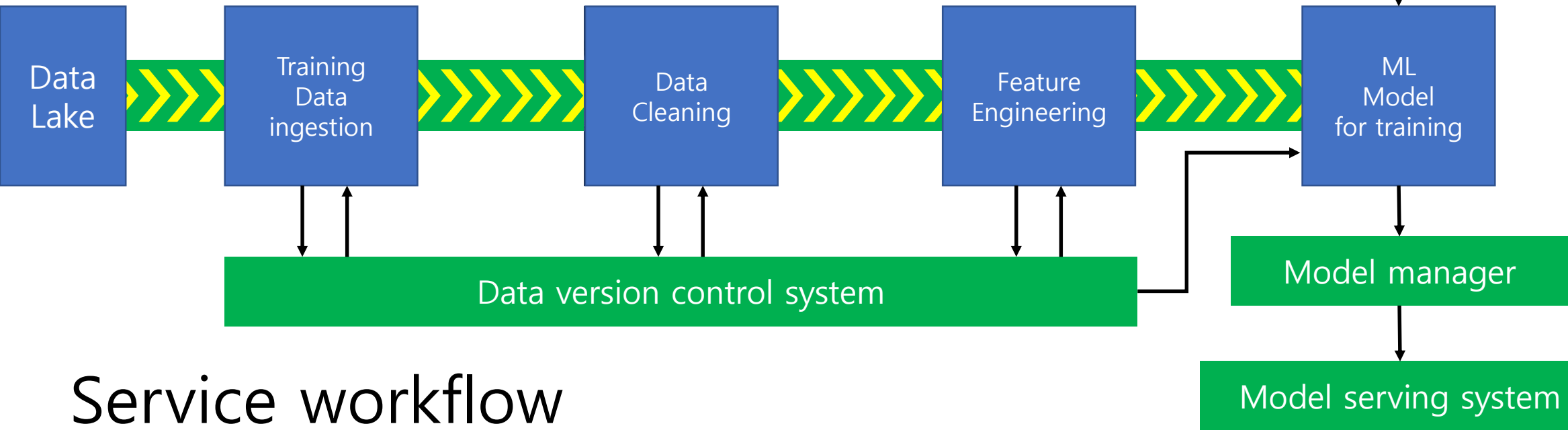
실제 서비스에서 모델을 사용할때도

필요한건? **DATA**



나는 유저가 준 데이터를 원한다

# Learning workflow



# Service workflow





근데 어떤 데이터?

~~~뿐

DATA

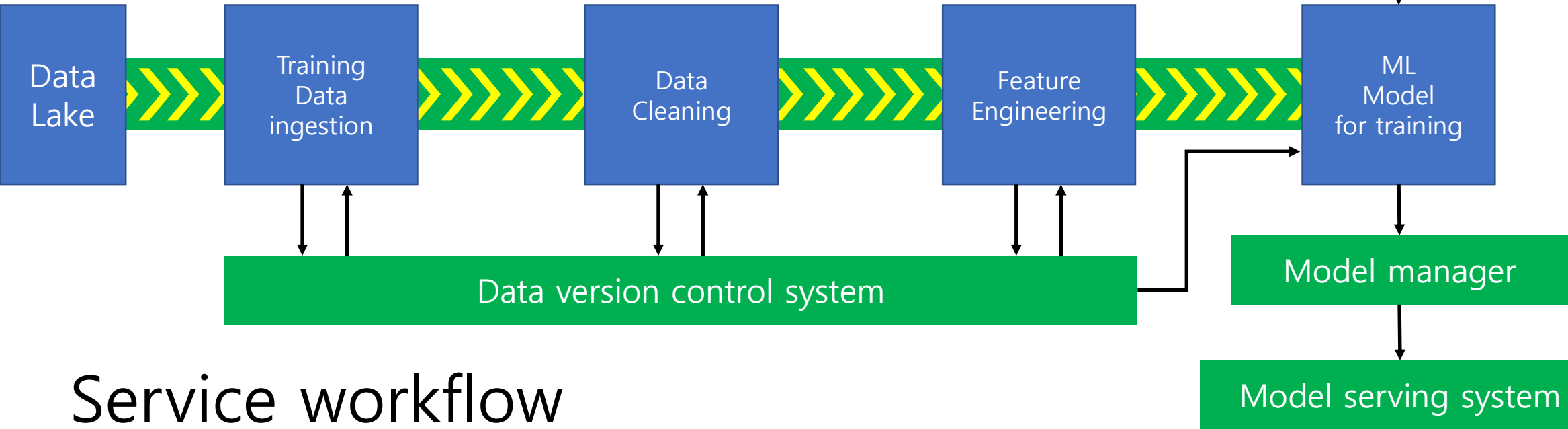
학습 과정에서 본 동일한 성능을C

재현성 을 위한

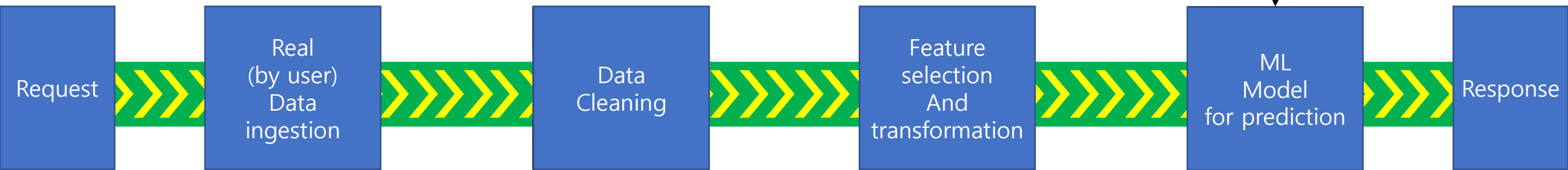
학습 과정에서 한 동일한 과정으로

preprocessing

Learning workflow



Service workflow



Data pipelining을 하려면



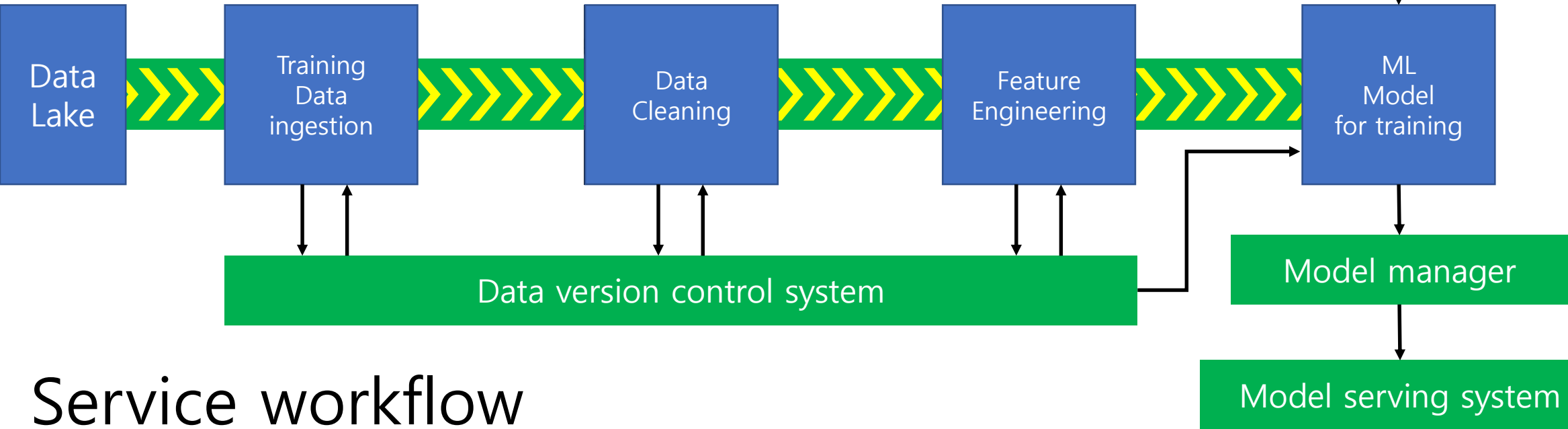
Orchestration 도 중요합니다.

참 중요한데
표현할 방법이 없네

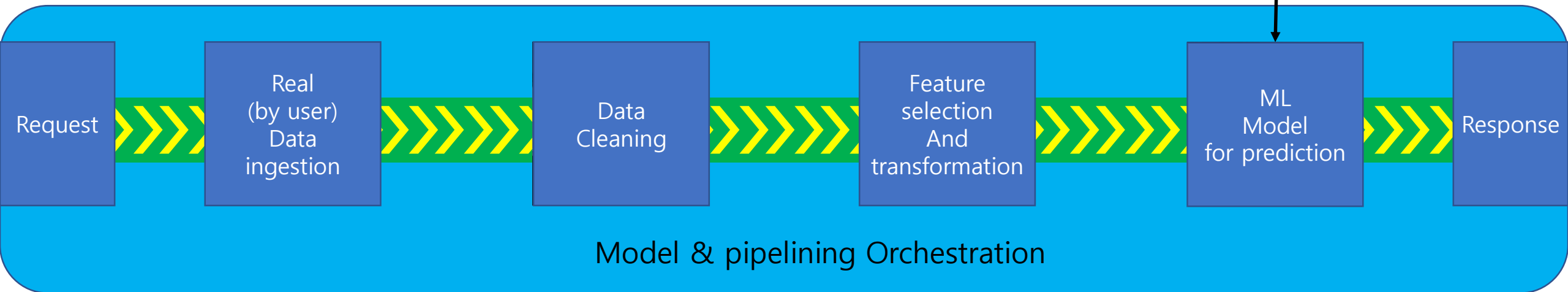


이것이 전체 workflow 입니다.

Learning workflow



Service workflow



다음 주는 각 파트별로 학습해와서
별표하는 형식으로 진행합니다.

아빠!
어디가?

춘천호 얼음캠핑
저녁 재료 장보기

MBC

화색
아 그냥 이게
끝이에요?

네... 끝이에요



거기에 계속되는 실험을 한다면

