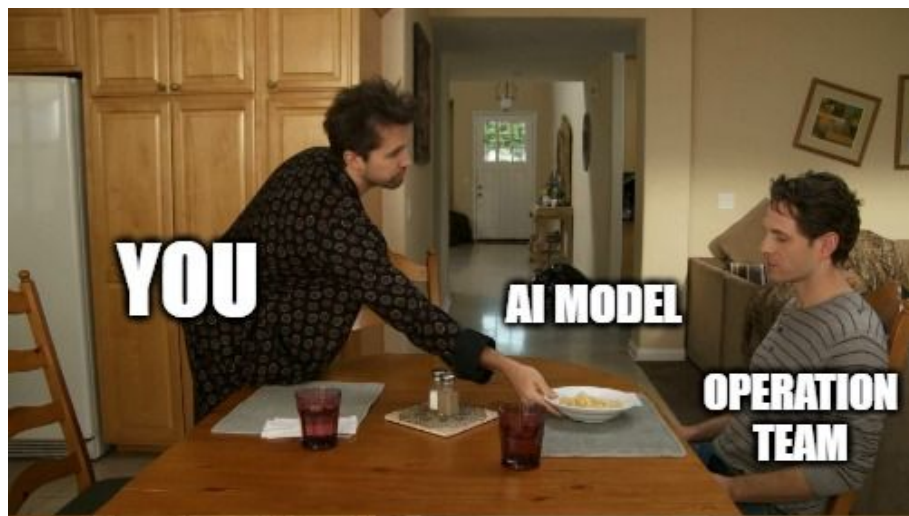


AI프로젝트의 빠른 도입을 위한 기술전략

김상우, 쏘카 데이터그룹/모빌리티그룹 그룹장

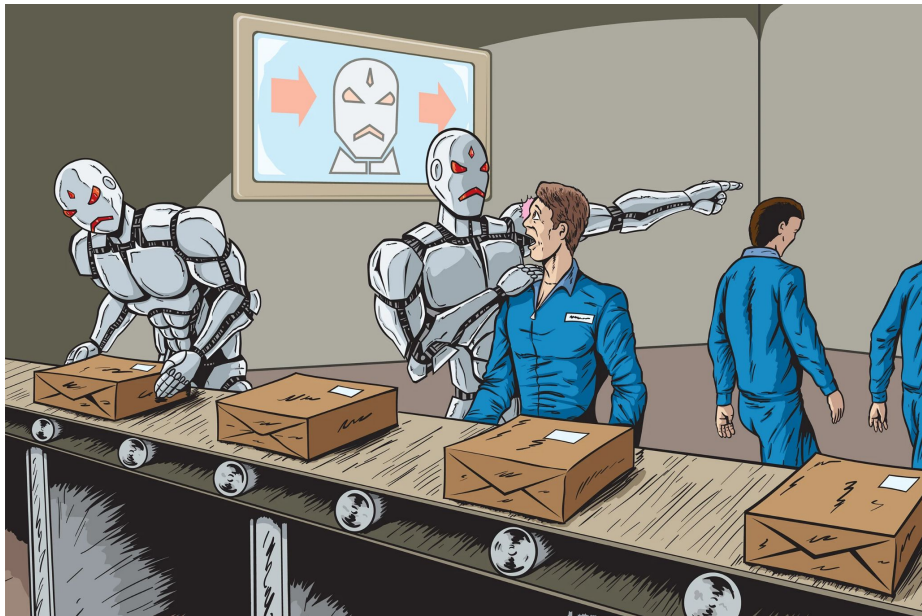
1

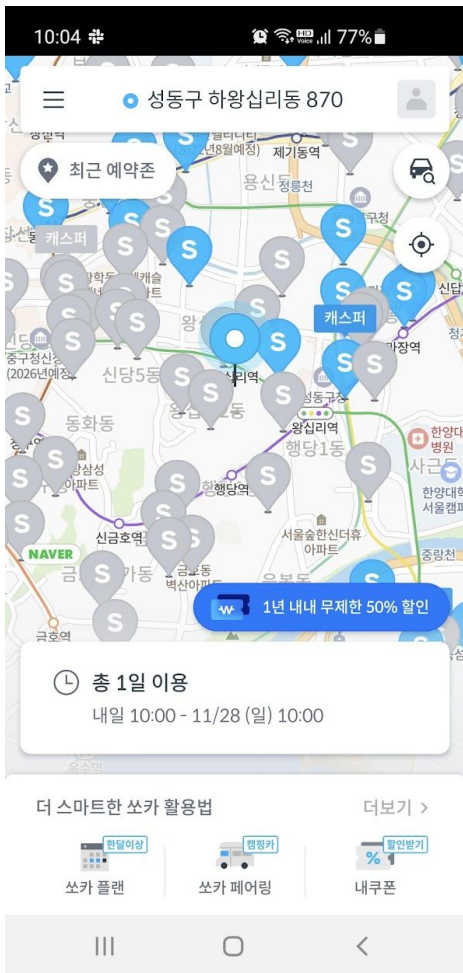


WHY?

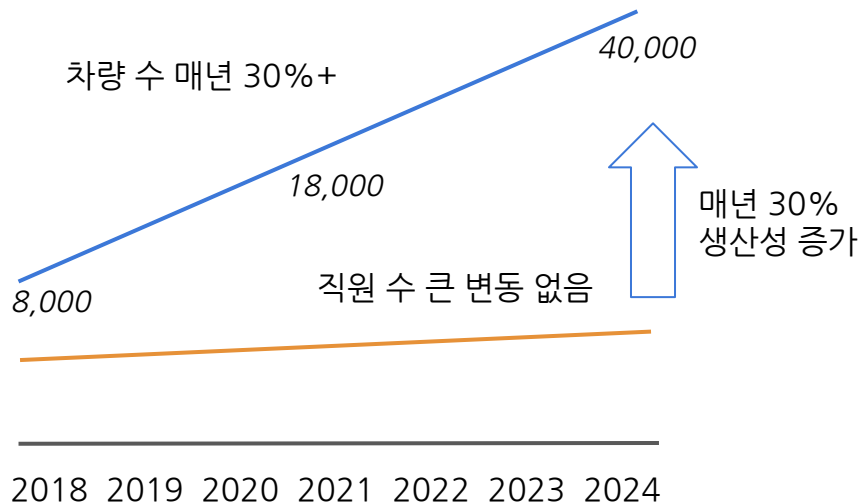
- 고용 안정성이 위협받는 경우들
 - 추천 알고리즘 vs MD (Merchandise)
- 부서간 갈등 (사내정치)
- 그냥 하기 싫어서 (원래 변화는 힘든것)
- 제대로된 AI모델이 아니어서
- 사실은 AI모델이 필요없어서 (안정적인 조직)

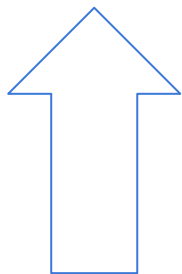
기존 기업들, 특히 대기업의 AI 도입이
실패하는 이유!





SOCAR





매년 30%
생산성 증가

- 운영, 사업 조직의 자동화, 무인화
- 더 나은 톨과 프로세스에 대한 강력한 동기

"Operational Excellence driven by AI"



?

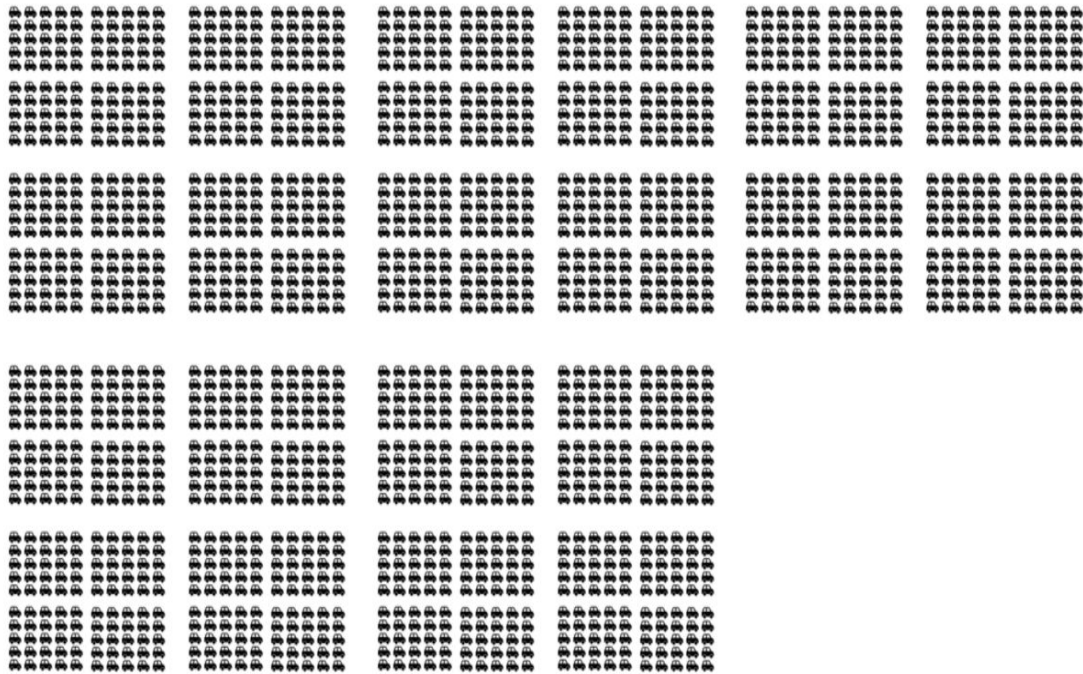
2018 - 200:1



2021 - 1000:1



2024 - 2000:1



SOCAR - Organizational Motivation

- 쏘카의 사업 100여명, 운영 100여명
- 기존 사업을 운영하는 조직
 - Job Security 이슈가 아니더라도..
 - 자발적인 AI Transformation은 쉽지 않음
- 성장과 효율화에 대한 압박감은 자연스럽게 AI Transformation에 대한
모티베이션이 됨
- Cooperation / Politics / KPI / Time



우리도 AI 한다

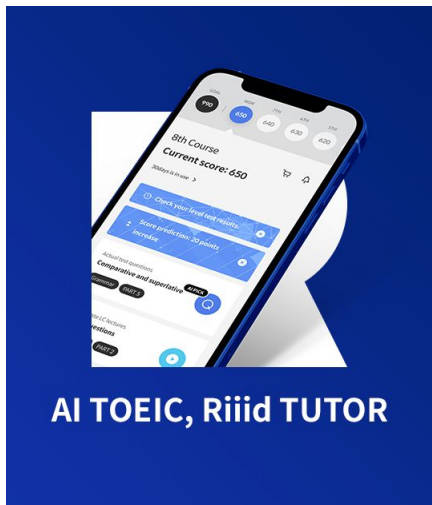


AI 활용하지 않으면 살아남을 수 없다!

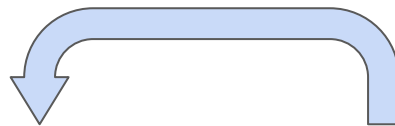
2

SOCAR - (Service) Domain + Operation

- 서비스를 이미 운영하고 있는 회사의 입장이므로
- 모든 문제를 AI 로 풀겠다! / Better, Faster, Cheaper
- 기존 오퍼레이션의 AI Transformation 을 위한 준비들이 필수적
 - 예) 정비 내역 확인 프로세스를 AI 화 하기 좋은 사진으로 남기고 체크한다던지



이용 데이터
(토익 문제 풀이)



더 좋은 서비스
(점수 예측,
족집게 문제풀이)



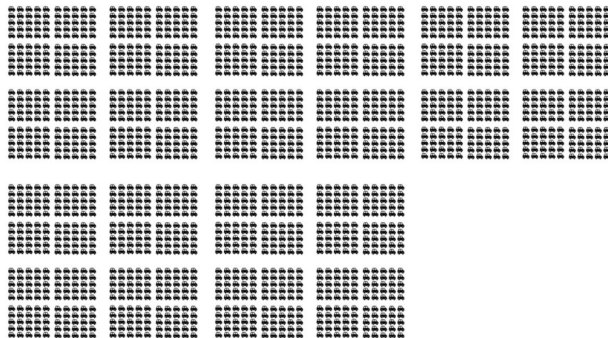
이용 데이터
(입사 지원, 합격 여부)



더 좋은 서비스
(지원 추천, 합격 예측)



- 자동화된 사고차량 확인
- 자동화된 세차 검수
- 세차 대상 차량 자동 선정
- 차량 위치 확인 자동화
- 블랙박스 장착 상태 확인 자동화
- .. many more!



AI 프로젝트 소개 몇가지

☰

Accida

차량 번호 ▾

Q

2021-05-13 ~ 2021-05-14

📅

정렬 기준

피손 확률 순 ▾

차량 번호
12히7760

차량 id
8379

예약 번호
36502411

명차 id
4476763

지역
충청남도

모달 판정 시간
2021년 3월 3일 14시

<

43

44

45

46

47

48

49

50

51

52

53

>

AI 프로젝트 소개 몇가지



B1



B2



B3



B4



B5



B6



B7

AI 프로젝트 소개 몇가지



(인증 성공) 외부촬영 거품



(인증 성공) 내부 촬영 기계 세차



(인증 보류) 일반 외부 이미지



(인증 보류) 내부 이미지

Image Augmentation: Resize + Center Crop

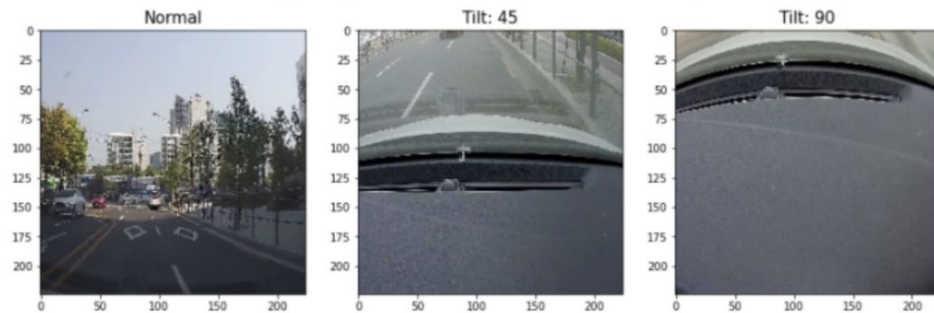
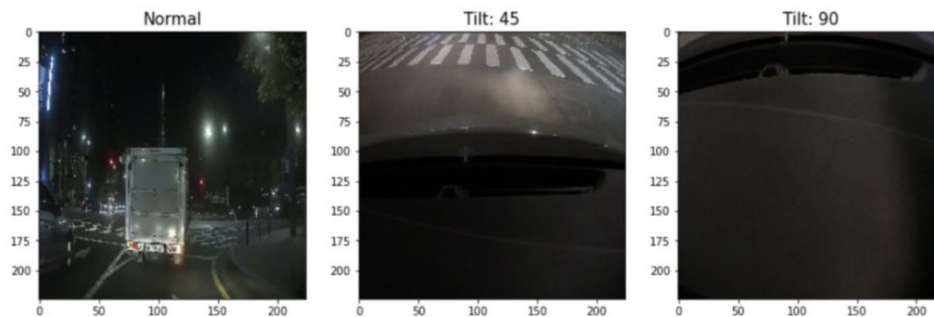
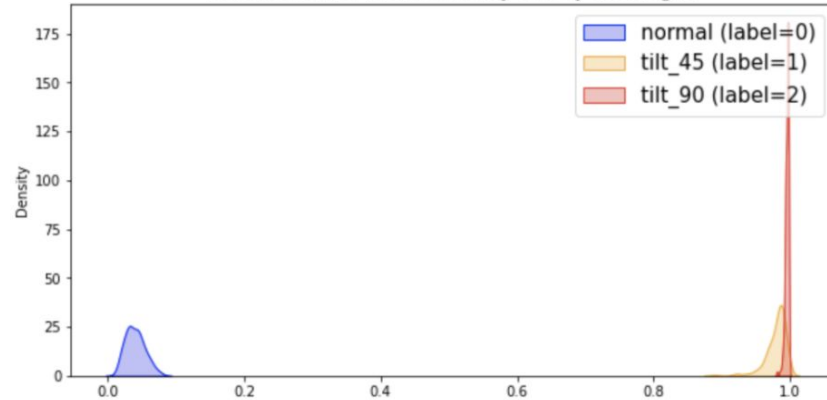


Image Augmentation: Resize + Center Crop



Distribution of Tilt Probability at Day and Night

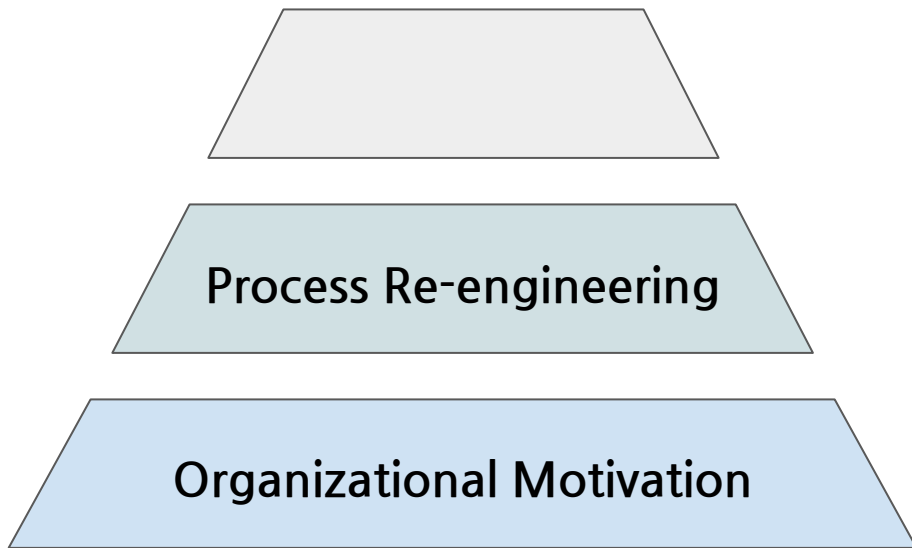




- 기존 정비 프로세스
 - 엄밀하지 못한 정비 카테고리
 - 사진으로 검수
- AI 도입을 위한 새로운 정비 프로세스
 - 정비 프로세스를 정규화, 규격화
 - 데이터 취득하기 좋아짐
 - 검수용 사진의 각도를 규격화
 - 데이터 품질
 - 학습기반의 알고리즘을 도입하고, 자동화하기 좋아지는 방향으로

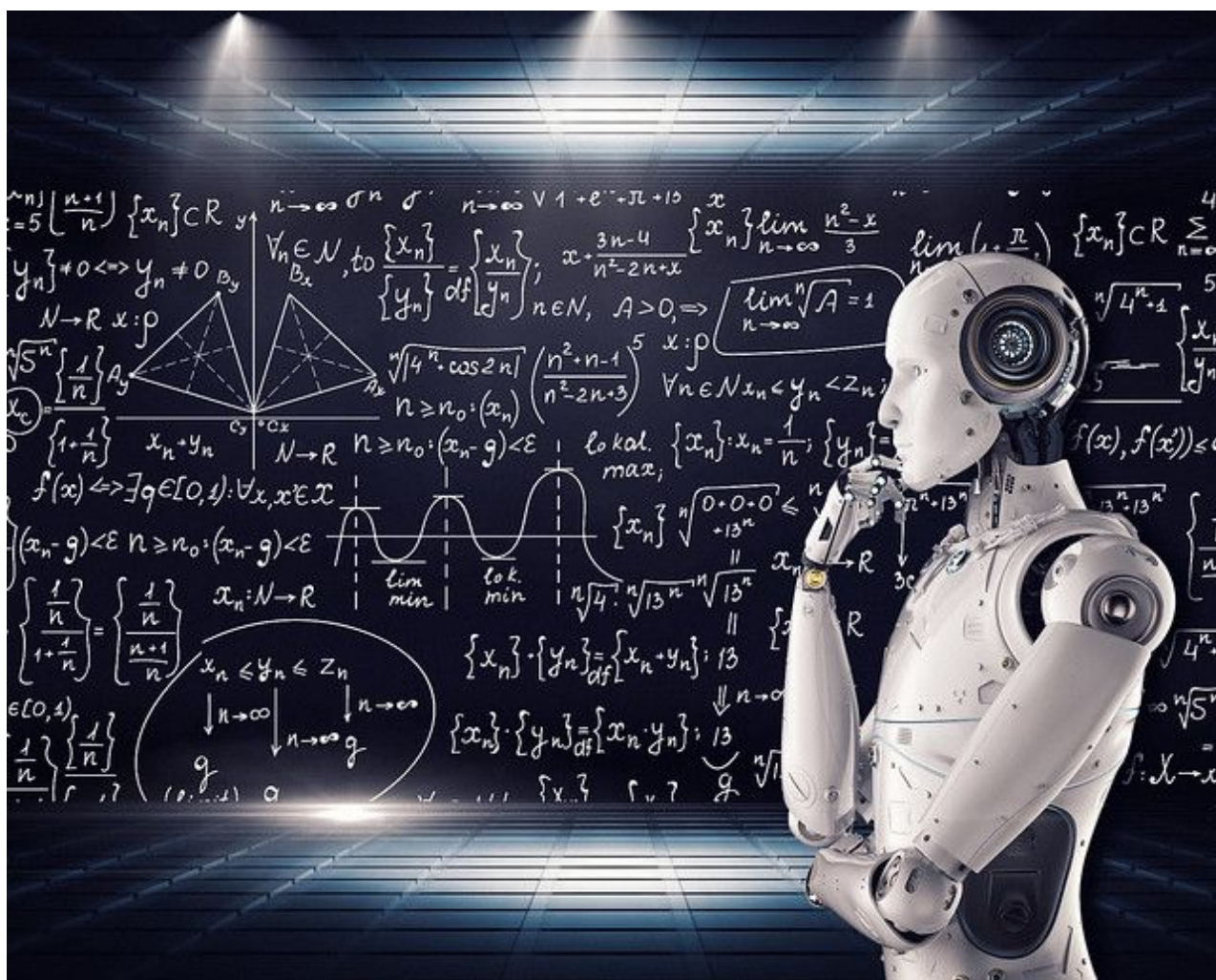


AI프로젝트를 통해 우리도 좋아지겠지



AI도입을 위한 프로세스/비즈니스모델 개선은 필수

3



Problems & Algorithms

- 회사에서는 각자의 영역(도메인)에서 매우 특수하고 섬세한 문제들을 해결하게 된다
- 이러한 문제들은 오픈도메인의 제너럴한 알고리즘으로 해결되지 않음
 - 예) ImageNet이 1,000개의 클래스를 분류해내지만, 쏘카에서 요구하는 차종이나 자동차 부위는 알수가 없다
 - 예) 초대형 자연어 모델이 회사에서 사용하는 용어들을 다 이해하기 어렵다
- 엄청나게 다양한 회사와 산업분야가 있고, 이에 맞는 모델들이 학습되고 운영되어야 함



- 도메인을 이해하는 AI Backbone
- 데이터 품질
- MLOps
- (도구 개발, 프로세스 자동화)

도메인을 이해하는 AI Backbone

- 도메인에서의 특수하고 디테일한 문제를 해결하는 모델들을 대량 양산해야 함
- AI Backbone 모델과 Transfer Learning (w. Self supervision)
- 차량 외관 사진 도메인
 - 사진에서 파손을 탐지하는 모델을 학습
 - 여기에서 파생해서 세차 오퍼레이션을 위해 오염을 탐지하는 모델을 쉽게 학습할 수 있다 (Transfer Learning)
 - 차종을 맞춘다던가, 부품 부위를 명명한다던가 등의 여러가지 모델 파생이 가능하다
 - 100여장의 매우 적은 수의 레이블로도 80%정도 성능의 모델이 제작 가능

데이터 품질

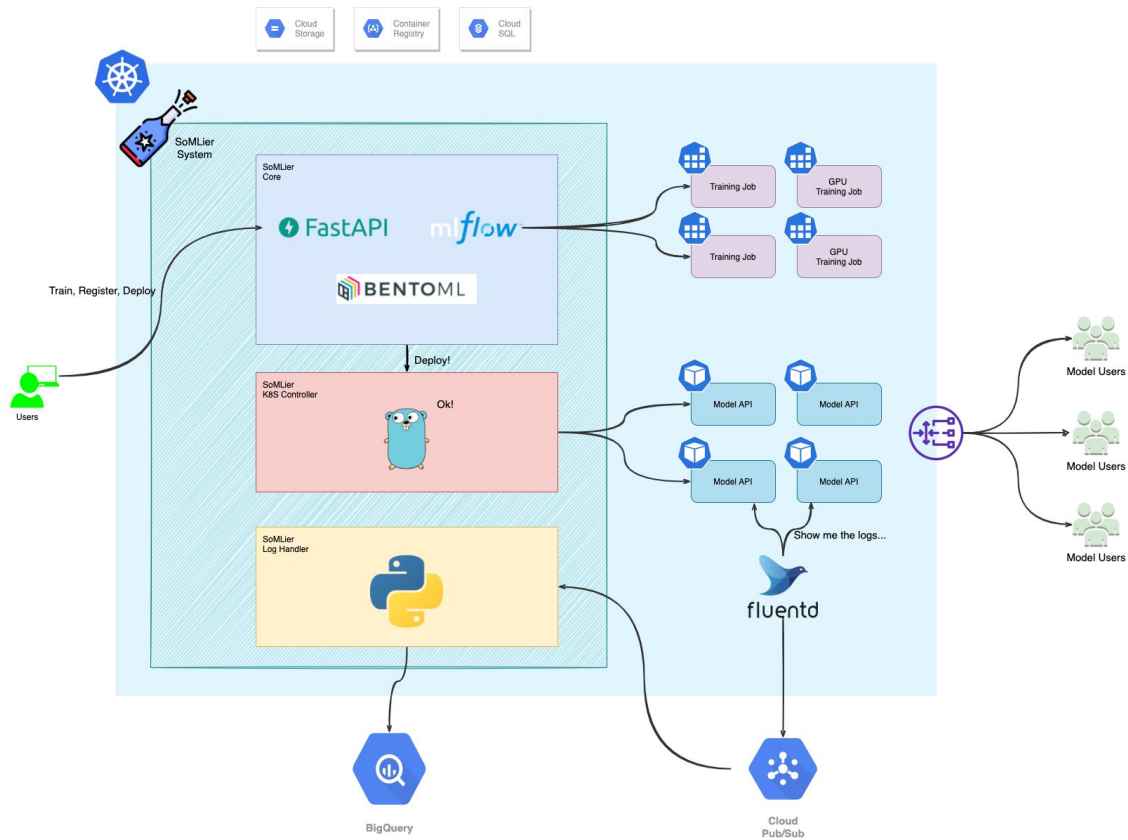
- 데이터
 - 데이터가 존재하는가?
 - 데이터가 충분한가?
 - 품질이 좋은가?
- HILL (Human In the Loop Learning)
 - 애매함을 정리하는 기술과 프로세스
- 데이터 품질에 대한 쏘카의 접근
 - 레이블링은 매우 까다롭게
 - 일정한 데이터를 얻을 수 있도록 (프로세스 개선을 통해)
 - 레이블이 애매한 경우는 극단적인 데이터부터, 단계적으로
 - 배포 이후의 데이터 품질 관리는 MLOps 에서 커버

MLOps - SoMLier

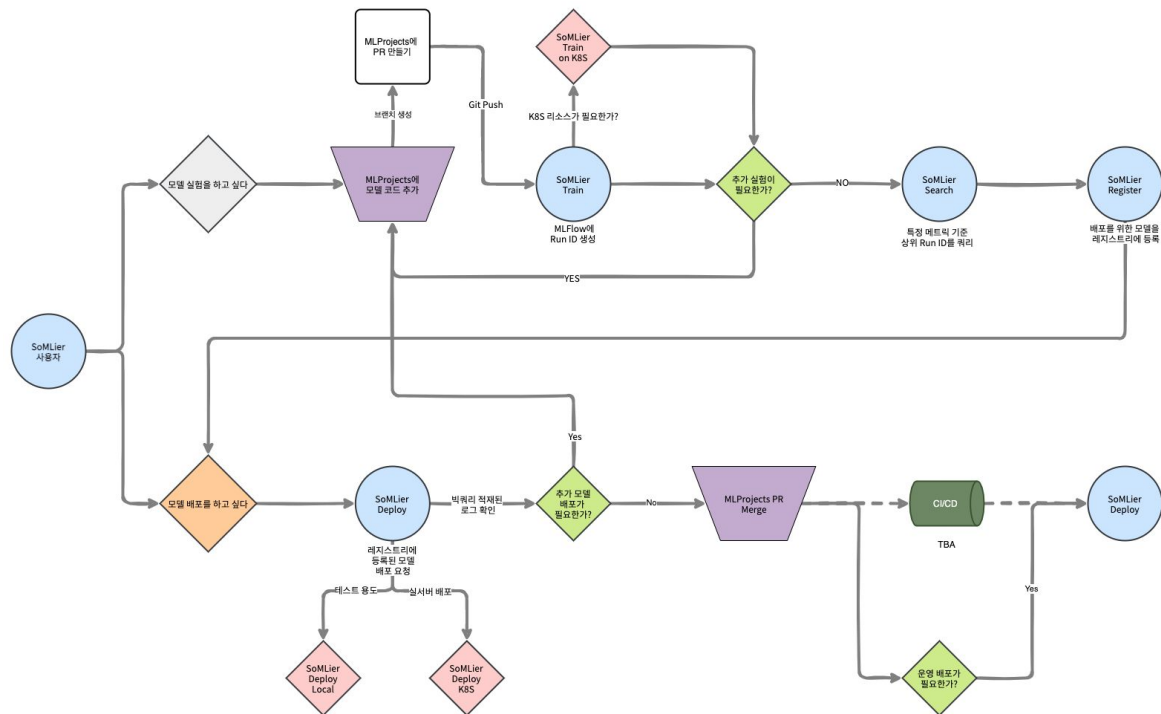
- 모델 개발 (노트북 도구)
- 모델 서빙 인프라, 배포 (CI/CD)
- 모델 분석, 모니터링, 테스트, 디버깅
- 피쳐스토어 (in progress)
- 데이터 퀄리티 (in progress)



SoMLier 소개



SoMLier 소개



ML 모델 사용자 입장에서 모델을 만들고 배포하는 흐름에서 고려해야하는 복잡도를 최소화, 모델 개발에만 집중할 수 있도록 도와줍니다.

도구 개발, 프로세스 자동화

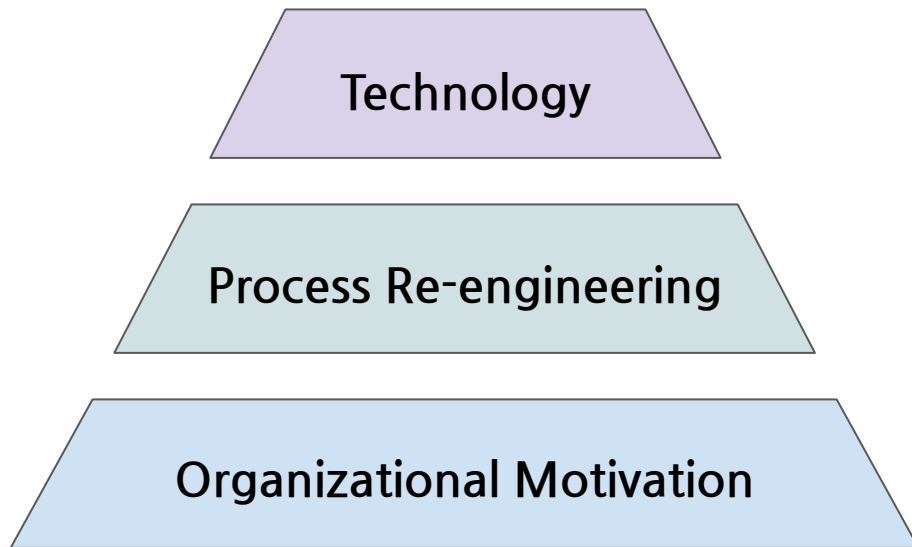
- 사례
 - 예) 파손차량검수 - 자동화된 파이프라인, 툴 개발
 - 예) 침수탐지 - 모니터링 에이전트, 툴 개발
 - 예) 세차 인증 - 자동화 에이전트
- 옵션
 - 슬랙봇 / 자동화 에이전트 / 시트 / 내부 어드민 연동 / 전용 프로세스 툴 개발 / API 제공 (MSA) / 외부 API 연동 등
- 다양한 옵션과 전문성을 가지고 현업부서의 니즈에 대응

AI 프로젝트 도입 과정

- 2019: 차량 파손 탐지 모델 및 검수 툴을 만들었으나 도입 실패
 - 커뮤니케이션 미스로, 운영부서의 니즈 파악을 잘못함
- 2020: 같은 기술로 Accida 프로젝트 성공
 - 운영부서의 니즈 정확히 파악 및 도입에 대한 정교한 조율
 - 차량 파손 탐지 태스크의 완전 자동화, 무인화
 - 기존 대비 x4 이상의 효율 향상
- 2021: Accida를 기반으로 여러 파생 모델 개발 성공
 - AI Backbone 프로젝트
 - 모델 개발 기간 3~6개월 => 1개월
- 2022~: 다양한 주제로 동시에 10여개의 AI 프로젝트 수행 예정



기술을 어디에 도입하지?



조직적 니즈와 프로세스 개선이 우선, 기술은 그 다음

정리

1. Organizational Motivation

- 모든것에 앞서서 조직 전체가 변화를 절실하게 필요로해야만 한다

2. Process Re-engineering

- AI 프로젝트가 성공적으로 녹아들기 위한 프로세스 개선

3. Technology

- 디테일한 문제들을 해결하고, 생산성을 높이기 위한 기술들

Q&A