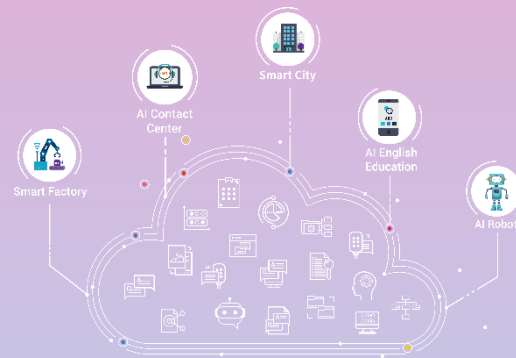


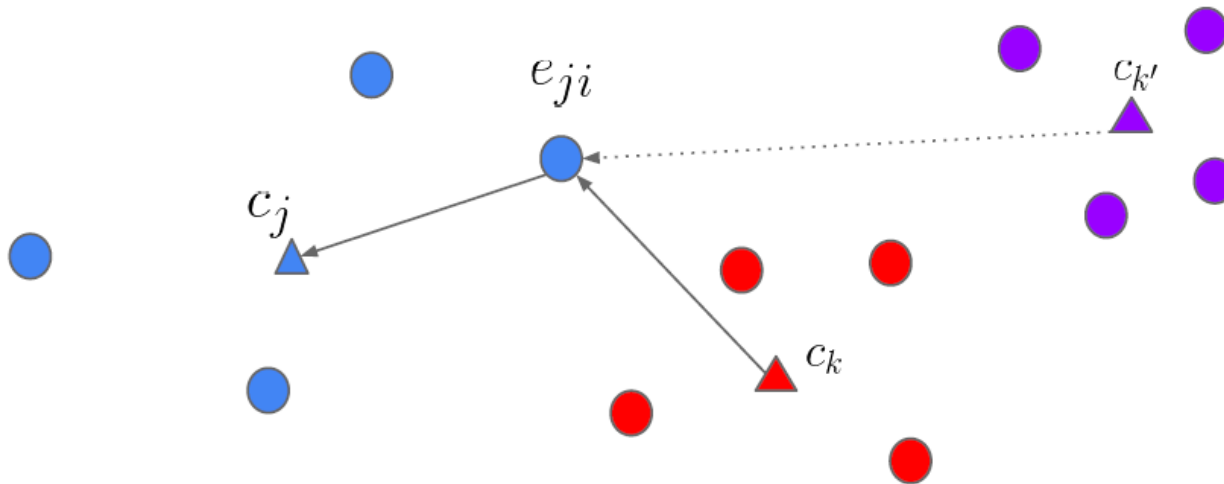
Level 2. How to customize AI engines



<http://mindslab.ai>

화자인증

- 화자인증(Speaker Verification):
 - 2개의 음성이 같은 사람의 것인지 판별해주는 엔진
 - 임의의 길이의 음성을 고정 차원 화자벡터(Speaker Vector)로 만들어줌
 - 벡터 사이의 거리를 통해 유사도를 판별



화자인증 학습 데이터

- N명의 사람들이 발화한 음성(5초 이상) M개
 - N, M 은 다다익선
- 동일한 사람에 대해 녹음 환경, 날짜, 감정, 발화 언어 등이 다양해야 함
 - VoxCeleb2 데이터 사용
 - Language agnostic (영어, 한국어 등 동시 사용 가능)



VoxCeleb is an audio-visual dataset consisting of short clips of human speech, extracted from interview videos uploaded to YouTube

7,000 +
speakers

1 million +
utterances

2,000 +
hours

화자인증 학습

- VoxCeleb2로 한번 학습해두고 나면 추가 학습이 필요 없음
- 화자인증이 아닌 화자인식으로 사용하려 한다면 transfer learning 가능

관련 자료



- <https://arxiv.org/abs/1710.10467> (화자인증)
- <https://arxiv.org/abs/1710.10468> (화자인증을 이용한 화자분리)
- <https://arxiv.org/abs/1810.04826> (화자인증을 이용한 음성분리 – VoiceFilter)
- <https://arxiv.org/abs/1902.10107> (화자인증 고도화)