

# 인공지능 인문학

4주차 PART 2

인공지능을 활용한 인문학 연구 설계(2)

김정아

# 학습 목표

- 1** 인공지능 활용을 위한 데이터의 주요 개념들을 이해한다.
- 2** 인간 중심 인공지능 설계에 대해 이해한다.

01

# 인공지능 데이터의 이해

# 데이터, 인공지능의 처음이자 끝

Improving the code vs. the data

	Steel defect detection	Solar panel	Surface inspection
Baseline	76.2%	75.68%	85.05%
Model-centric	+0% (76.2%)	+0.04% (75.72%)	+0.00% (85.05%)
Data-centric	+16.9% (93.1%)	+3.06% (78.74%)	+0.4% (85.45%)

# 인문학에 데이터가 '더' 중요한 이유

01

## 적은 데이터로 좋은 결과를

- 인문학 연구를 위한 데이터를 구하는 것은 쉽지 않기 때문에 적은 데이터로 학습을 하게 되고, 과적합(한쪽으로 과하게 학습되는 것)과 같은 문제가 발생할 수 있다. 이를 막기 위해 잘 분포된 좋은 데이터를 준비하는 것이 필요하다.

02

## 신뢰할 수 있는 연구를 위한 안전장치

- 머신러닝·딥러닝 모델의 '블랙박스' 문제는 현재로서는 완전히 피하기 어렵다. 하지만 신뢰할 수 있는 데이터로 모델이 학습되었다면 모델의 신뢰도도 높아질 수 있다.

03

## 연구 주제의 확장

- 양질의 데이터를 충분한 수량으로 준비한다면 수백 건의 도서를 한 번에 비교하는 등 인문학 연구자 개인이 할 수 있는 연구보다 훨씬 폭 넓은 연구가 가능하다.

# 인공지능에서 말하는 좋은 데이터

학습 데이터는 일종의 '샘플'이므로,  
연구하고자 하는 대상의 실질적인  
모습을 가장 잘 대표할 수 있는  
데이터로 구축해야 한다.

**균형잡힌  
데이터**

잘못 들어간 샘플,  
잘못된 라벨 등이 없는  
정확한 데이터를 구축해야 한다

**정확한  
데이터**

데이터를 확인하기 쉽도록  
잘 문서화되어 있고  
쉽게 저장/업데이트/검색이 가능한  
데이터로 구축해야 한다.

**투명한  
데이터**

# 데이터의 유형(1) 정형 - 비정형



## 정형 데이터

- 형태가 구조화되어 있음(엑셀 스프레드시트 형식)
- 일정한 형식을 갖추고 있어 SQL(Structured Query Language) 등을 활용하여 쉽게 검색하거나 분석할 수 있음
- 예: 인구 통계 데이터(나이, 성별, 소득, 주소 등)



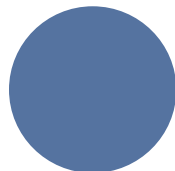
## 비정형 데이터

- 미리 정한 구조 없이 자연스럽게 생성
- 정해진 구조가 없기 때문에 정보 등을 추출하기 어렵고 분석도 어려움
- 예: 크롤링한 자유게시판 데이터, 이미지 데이터, 동영상 데이터

# 데이터의 유형(2) 종류별 데이터

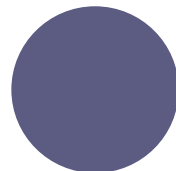
## 자연어 데이터

사람이 사용하는 자연스러운 언어로 이루어진 데이터



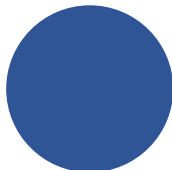
## 이미지 데이터

사진, 그림, 스캔된 문서 등 이미지로 구성된 데이터



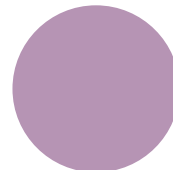
## 오디오 데이터

음성이나 음악, 기타 소리로 이루어진 데이터



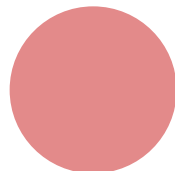
## 시계열 데이터

시간에 따른 순차적인 데이터. 기상 정보보다 주식 가격 데이터 등.



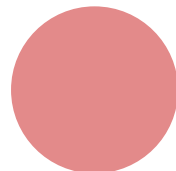
## 센서 데이터

IoT 등 센서에서 수집된 데이터



## 멀티모달 데이터

둘 이상의 유형의 데이터가 함께 있는 데이터





# 메타데이터(metadata)란?

## D DATA CARD FOR LANGUAGE TRANSLATION DATASET

Translated Wikipedia Biographies

[English to Spanish](#) • 516 KB • CSV
 [English to German](#) • 517 KB • CSV

The Translated Wikipedia Biographies dataset has been designed to evaluate gender accuracy in long text translations (multiple sentences or passages). The set has been designed to analyze common gender errors in machine translation like incorrect gender choices in anaphora resolutions, possessives and gender agreement.

PUBLISHER(S)

Google LLC

INDUSTRY TYPE

Corporate - Tech

DATASET AUTHORS

Anja Austermann, Google  
 Michelle Linch, Google  
 Romina Stella, Google  
 Kelle Webster, Google

FUNDING

Google LLC

FUNDING TYPE

Private Funding

DATASET CONTACT

[translote-gender-challenge-sets@google.com](mailto:translote-gender-challenge-sets@google.com)

DATASET PURPOSE(S)

Testing

KEY APPLICATION(S)

Machine Translation

Gender Accuracy

PRIMARY MOTIVATION(S)

Study gender accuracy in translations beyond the sentence in demographic and occupations diversity for fairness research.

INTENDED AND/OR SUITABLE USE CASE(S)

To evaluate gender accuracy on translations beyond the sentence (multiple sentences or passages). The set is focused on the presence of this specific linguistic phenomena to evaluate the most common contextual errors:
 

- Spanish to English: [Pro-drop](#)
- Spanish to English: Neutral to gender-specific possessives
- English to Spanish, German: [Gender agreement](#)

PRIMARY DATA TYPE(S)

Non-Sensitive Public Data about people

DATASET SNAPSHOT

Total instances	138
Masculine biographies (entities)	83
Masculine biographies (countries)	51
Feminine biographies (entities)	83
Feminine biographies (countries)	57
Rock bands & sport teams (entities)	12
Rock bands & sport teams (countries)	12

DESCRIPTION OF CONTENT

This dataset is based on publicly available data on public and/or historical figures (Wikipedia articles) at a given snapshot in time.
 

The dataset has 138 instances and each instance contains the first 8 to 15 sentences from a Wikipedia article. Articles are written in native English and have been professionally translated to Spanish and German. 126 of these instances represent a person with an associated stated gender and 12 are related with rock bands or sport teams (considered genderless).

DATASET SOURCE(S)

- Source Text: [English Wikipedia](#)
- Target Text: Professional translations

HOW TO INTERPRET A DATAPPOINT

**Each datapoint** refers to a central entity that can be a person (stated as feminine or masculine), a rock band or a sport team (considered genderless).

**Each entity** is represented by a long text translation (multiple connected sentences or continuous passage referring to that main entity).

## 정의

- 데이터에 대한 데이터
- 다른 데이터(콘텐츠)를 설명해주며, 일정한 규칙에 따라 구조화되어 콘텐츠와 함께 저장된다.

## 주요 예시

- 도서관의 카드 목록
- 디지털 사진 메타정보 exif
- 웹 문서의 메타데이터

# 메타데이터가 중요한 이유

01

## 메타데이터 = 데이터에 대한 분석

- 자동적으로 생성되는 메타데이터가 아닌, 인간의 판단이 들어가야 하는 메타데이터의 경우 메타데이터 생성 자체가 해당 데이터에 대한 연구라고 할 수 있다. 반대로 기존의 연구를 토대로 메타데이터로 만들 수 있다.

02

## 공유 및 협업의 가능, 연구의 확장

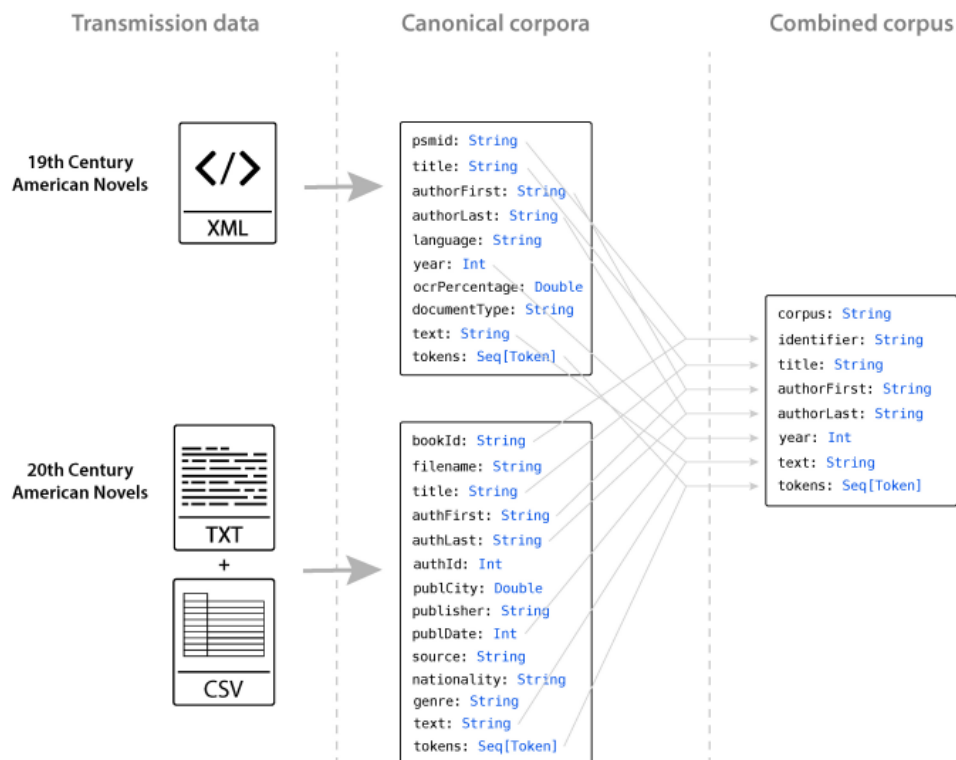
- 상세한 메타데이터를 만들면 데이터를 공유하기 쉬워지고, 추가 데이터 구축을 진행하는 것도 쉬워진다. 따라서 연구 분야의 지식의 공유와 발전을 촉진할 수 있으며 메타데이터를 통해 서로 다른 데이터들을 함께 연구할 수도 있다.

03

## 데이터 관리 용이

- 대량의 데이터가 필요한 인공지능 학습과 소량의 데이터들이 산재되어 있는 인문학 연구의 특성상 여러 종류의 데이터를 혼합하여 사용하는 일이 잦다. 이 때 메타데이터를 통해 각 데이터들을 빠르고 쉽게 관리할 수 있다.

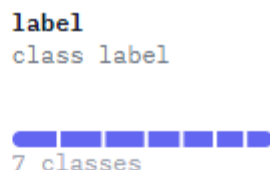
# 인문학 데이터의 메타데이터 생성



Stanford Literary Lab의 메타데이터 생성 사례

- 대부분의 인문학 데이터 구축은, 이미 연구가 이루어진 콘텐츠에 대해 나중에 이루어지게 된다.
- 따라서 이 과정에서 메타데이터를 생성하는 일은 기존의 연구에서 각 데이터별로 포착한 상세한 분석들을 잃지 않으면서, 여러 텍스트가 모여서 이루는 상위 단위로 일관된 병합을 이루어야 한다는 상충된 목표를 달성해야 한다.
- 이러한 과정은 해당 분야의 전문가가 할 수 있는 일로, 이 자체가 하나의 연구 주제가 될 수도 있다.

# 데이터 주석(data annotation)



유튜브 내달 2일까지 크리에이터 지원 공간 운영	3 생활문화
어버이날 앞두고 호러저...남부지방 열은 황사	3 생활문화
내년부터 국가RD 평가 때 논문건수는 반영 않는다	2 사회
김명자 신임 과총 회장 원로와 젊은 과학자 지혜 모을 것	2 사회
회색인간 작가 김동식 양심고백 등 새 소설집 2권 출간	3 생활문화
아워서 생방송 하세요...액션캠 전용 요금제 잇따라	0 IT과학

KLUE dataset

- 주석(label) 유무에 따른 데이터 분류
  - 원시 데이터(raw data): 아무 주석이 없는 데이터
  - 주석 데이터(labeled data): 모델 학습을 위해 필요한 정보를 추가한 데이터
- 지도 학습(supervised learning)이 비지도 학습(unsupervised learning)보다 성능이 높기 때문에, 가능하다면 주석된 데이터를 구하는 것이 좋다.

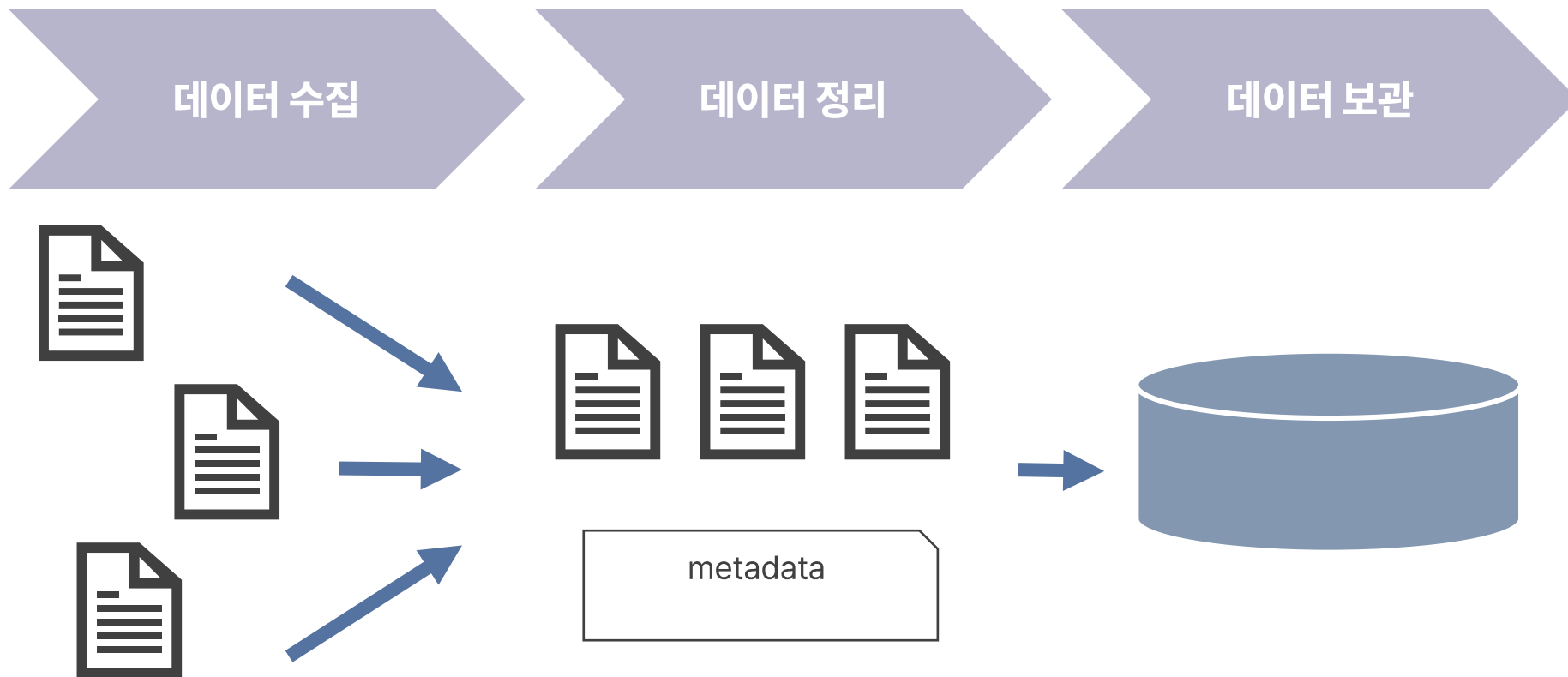
# 데이터 주석의 실제(1)

- 인명, 지명, 단체명 이 3가지 카테고리에 대해 고유명사(named entity)를 태깅한 데이터를 구축하고자 할 때 아래 문장을 어떻게 처리할 것인가?
  - 오늘 **서울시장**이 새로 당선되었다.
    - <서울>시장 / <서울시장>
  - **현기차** 독점 심화
    - <현><기>차, <현기차>
- 문장의 감성(긍정/부정/중립)을 나누는 데이터를 구축하고자 할 때 아래 문장을 어떻게 처리할 것인가?
  - 이 영화 지루하긴 한데 취향에 맞으면 진짜 인생영화입니다.

# 데이터 주석의 실제(2)

- 주석 체계 만들기
  - 고문헌의 연도 구분 모델을 만들기로 했다. 이런 경우 보통 분류 모델로 해결하기 때문에 연도를 일정 단위로 나누어야 하는데 어떻게 나누면 좋을까?
    - 균일하게 N년 단위 / 역사적인 큰 사건을 기준으로 나누기 등등
  - 18세기 편지 데이터를 주제별로 주석을 달기로 하였다. 이 때 어떤 기준으로 주석 체계를 만들 것인가?

# 데이터 관리



02

# 인간 중심 인공지능



# Human Centered AI란?

AI 시스템은 법치, 인권, 민주적 가치 및 다양성을 존중하는 방식으로 설계되어야 하며, 공정하고 정의로운 사회를 보장하기 위한 적절한 안전장치를 포함해야 합니다.

– OECD AI Principle 1.2

인공지능 성능 향상에  
무엇이 필요한가?



인간을 위해 인공지능이  
뭘 할 수 있는가?

# HCAI의 두 가지 측면

01

## 사회 문화적 관점에서 인간을 이해하는 AI 시스템

- 기계 중심적인 관점으로 작동되는 인공지능이 아니라 인간 중심적 관점으로 인간을 이해하고 그에 맞추어 작동하는 인공지능 시스템

우리 전공

02

## 인간이 AI를 이해하도록 돕는 AI 시스템

- 궁극적으로 블랙박스를 벗어나 사용자가 그 내부를 분명하게 인식할 수 있는 투명성을 갖춘 시스템

엔지니어링  
과 협력

Riedl, M. O. (2019). Human-centered artificial intelligence and machine learning. *Human behavior and emerging technologies*, 1(1), 33-36.

# Q&A

질의 응답은 커뮤니티에