

디지털인문학입문

2025.05.27.

07. AI 시대, 콘텐츠를 읽고 만드는 힘

강혜원

2025-01: 디지털 인문학입문

- **디지털 인문학**에 대한 기본적인 이해와 함께,
미디어 위키와 **시맨틱 데이터** 관련 기술을 배우고
이를 토대로 자신만의 **디지털 큐레이션 콘텐츠**를 제작해보는 수업
- 단순히 새로운 기술을 배우는 것을 넘어서,
왜 우리가 이런 기술을 인문학적 관점에서 이해하고 활용해야 하는지
함께 생각해보기

2025-01: 디지털인문학입문

- **Digital Curation** – Compilation

- ✓ **선별**: 가치 있는 정보 또는 콘텐츠를 선택
- ✓ **맥락화**: 자료를 새로운 시각이나 기준으로 해석
- ✓ **구조화**: 정보를 일정한 논리에 따라 정리
- ✓ **전달**: 다른 사용자 또는 독자가 쉽게 탐색하고 이해할 수 있게 재구성

- **Digital Literacy** - Reading & Writing

- ✓ 디지털 환경에서 정보를 탐색하고 해석하는 읽기 능력
- ✓ 그리고 그에 기반해 창의적으로 표현하고 연결하는 쓰기 능력
- ✓ 텍스트, 이미지, 하이퍼링크 등을 조합한 비선형적 텍스트

- **Semantic Data** – Designing

- ✓ 단순한 정보의 나열 $x \rightarrow$ 인물, 사건, 장소, 주제 등 다양한 요소들을 의미 기반으로 연결
- ✓ 내가 만든 콘텐츠가 누구에게, 어떤 방식으로 도달할지를 고려하며 설계

왜 필요할까

- 디지털 환경에서는 콘텐츠가 넘쳐남. 뉴스, 영상, 글, 댓글, 이미지, 채팅까지 모든 것이 정보이고 콘텐츠.
- 이제 단순히 정보를 '많이 아는 사람'보다, **정보를 잘 연결하고 해석하는 사람**이 더 중요한 시대
- 그 안에서 눈에 띄는 사람은?
 - ✓ 정보를 '잘 읽는 사람',
 - ✓ 의미를 '새롭게 엮는 사람',
 - ✓ 그리고 그것을 '다른 사람이 이해하기 쉽게 보여주는 사람'



Digital Curation

왜 필요할까

- 의미를 잘 읽고 맥락을 해석하는 것은 인문학의 영역
- 디지털 환경에서는 그 능력을 **디지털 방식으로 재현할 줄도 알아야**
- 단순히 글을 잘 쓰는 것만으로는 부족
- 카드뉴스, 인터랙티브 콘텐츠, 영상 캡션, 스토리텔링 구조 등 **멀티 모달 콘텐츠**에 대한 이해 필요
- 내가 만드는 콘텐츠가 어떤 플랫폼에서, 어떤 형태로, 어떤 사용자에게 다가갈지를 생각하면서 만들어야 함



Digital Literacy

왜 필요할까

- 내가 전달하고자 하는 맥락을 구조화하는 전략 필요
- 단지 콘텐츠나 정보를 나열하는 데 그치지 않고, 의미 단위로 나누고, 연결하고, 분류하는 작업
- 예를 들어 '김환기는 유명한 화가이다'라고 하기 보다, 그의 주요 작품, 수상 경력, 관련 인물 등과의 관계로 정보를 구성, 독자가 더 쉽게 탐색하고 활용할 수 있도록
- 즉, 정보의 밀도를 높이고 맥락과 관계성을 강화하는 작업



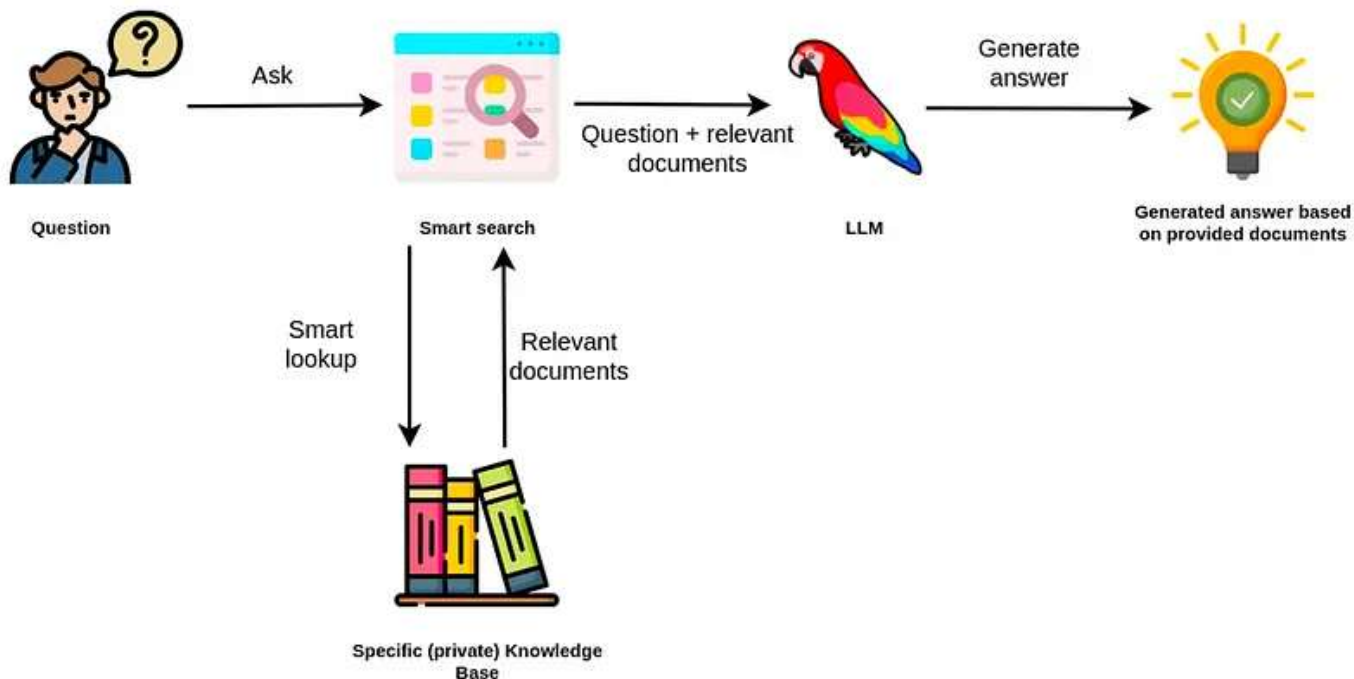
Semantic Data

AI 시대의 디지털 리터러시

- 이러한 정보의 구조화 방식은 일종의 'AI에게 말걸기'
- 시맨틱 데이터를 통해 구체적, 명시적으로 정리된 맥락은 AI가 더 정확하게 정보를 검색하고 이해하는 데 도움을 줄 수 있음
- 콘텐츠 산업, 출판, 미디어, 전시, 교육, 관광, 지역문화... 거의 모든 분야에서 **AI를 활용한 콘텐츠 생성, 요약, 추천, 큐레이션**을 고민하고 있음
- 그 AI가 제대로 작동하려면, AI가 이해할 수 있는 데이터가 필요
- 기획자, 편집자, 스토리텔러, 작가가 AI를 다루기 위해서는, **AI가 이해할 수 있는 정보 구조를 만드는 능력**이 필수

RAG Retrieval-Augmented Generation 검색-증강 생성

- 외부 데이터에서 정보를 검색해와서 LLM이 생성하는 답변을 보완하는 기법
- 데이터베이스, 문서, 웹사이트 등 특정 분야의 독점적이고 맥락적인 정보를 대상으로, LLM이 훈련받은 데이터 범위를 넘어 더 정확하고 구체적인 정보를 제공



AI Agent

- 사람이 일일이 지시하지 않아도 주어진 목표를 달성하기 위해 알아서 계획을 세우고 실행하는, 일종의 “자율 실행 AI 시스템”



Data

- RAG와 AI Agent 모두 데이터를 얼마나 구조화하고 의미 있게 제공하느냐가 중요
- 동일한 AI 모델이라도 특정 분야의 고품질 데이터가 풍부하면, 해당 영역에서 훨씬 뛰어난 성능을 보임
- "AI를 어떻게 우리 데이터 위에 잘 얹을 것인가"
- AI가 우리 지식과 콘텐츠를 잘 읽고 활용하게 하는 '훈련 환경'을 만드는 일이 중요해질 것

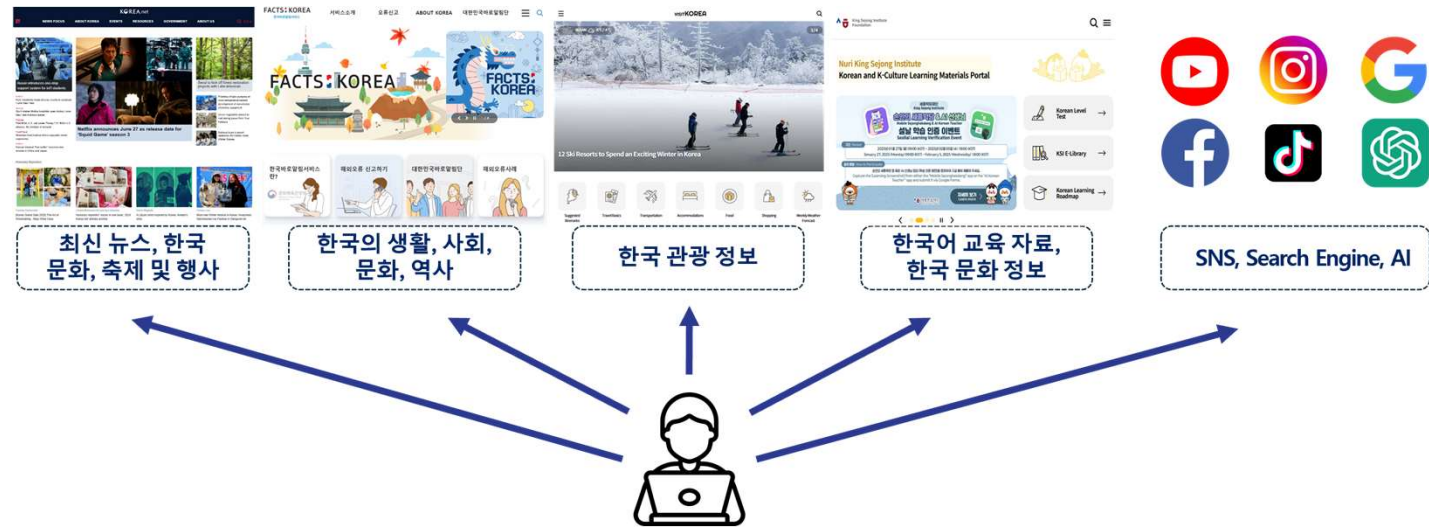
Semantic Data

- “시맨틱 기반 데이터 설계”는 현재 GPT 같은 대규모 언어모델(LLM)의 한계를 보완하고 응용 분야를 확장하기 위한 전략적 기반 중 하나로 주목받고 있음
- LLM은 통계 기반 확률 모델이라, 지식의 정확도·일관성·출처 추적성 등에서 한계가 있지만,
- 특정 분야에 대해 정확한 지식 정보를 담은 시맨틱 데이터(용어 정의, 관계망)가 LLM의 프롬프트 또는 RAG 소스로 활용되어,
- 개념 간 관계, 주제 간 맥락을 기반으로 사용자 맞춤형 응답을 가능하게 함

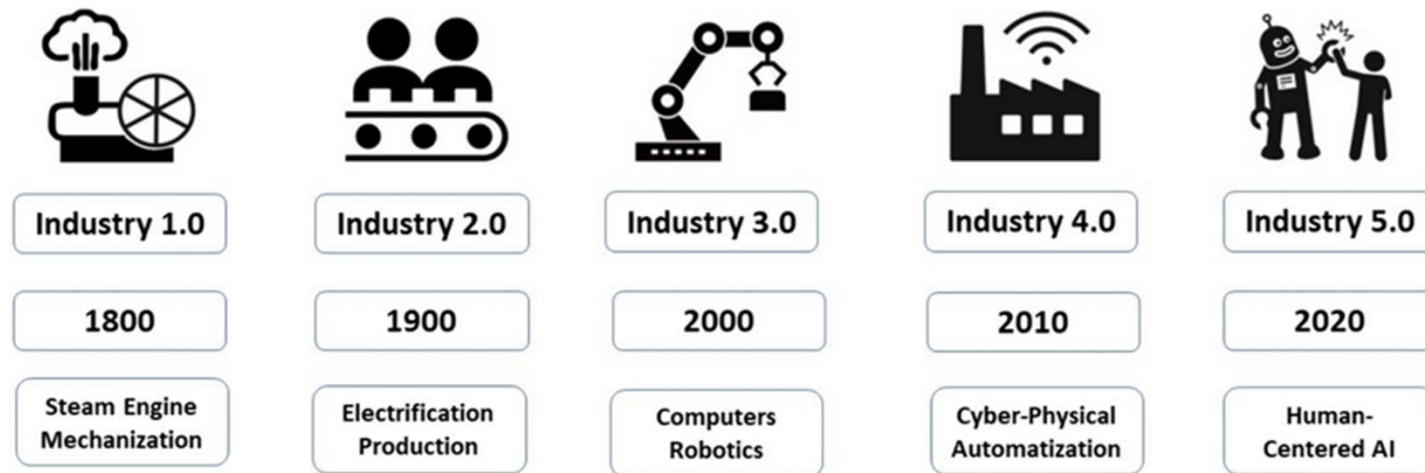
AS-IS



TO-BE



AI와 인간이 협업하는 시대



https://www.researchgate.net/figure/The-industrial-revolution-from-Steam-Engines-to-human-centered-AI_fig1_383950180

- AI는 인간을 대체할 것인가....?
→ 문과적 지식과 스토리텔링이 AI 시대의 미래에 어떤 기여를 할 수 있을까
- 우리가 디지털 리터러시와 디지털 큐레이션을 배우는 이유
→ 인문학적 감수성과 지식을 디지털 환경에서 잘 활용하면,
AI는 인간의 강력한 도구가 될 수 있음
- 인공지능에게 '말을 걸고', 인공지능이 '잘 이해할 수 있게끔 의미를 구조화'하는 능력이 있다면,
기술에 종속되는 사람이 아니라 그것을 적극적으로 활용하는 사람이 될 수 있음