

# 디지털 인문학 입문

트와이닝 린지  
강화여자고등학교  
2019-05-12

# 목차

- ▶ 디지털 인문학의 정의
- ▶ 인문학에 대한 이해
- ▶ 아날로그와 디지털인문학의 관계
- ▶ 디지털인문학의 역사
- ▶ 디지털 인문학에서 주로 활용하는 기술
  - ▶ 마크업
  - ▶ 데이터베이스
- ▶ **AKS** 디지털인문학연구소 연구
  - ▶ 연구 방법
  - ▶ 프로그램
  - ▶ 언어
- ▶ 디지털 인문학 국내외 사례

# ‘디지털인문학’이란

- ▶ 디지털 기술을 활용하여 인문학 연구와 교육하는 것

# ‘인문학’이란

- ▶ 인간(의 문화)를 이해하려고 하는 연구
- ▶ 선사시대부터 현재까지
- ▶ 인간을 이해할 수 있는 통로
  - ▶ 문헌(기록, 일기, 시, 소설, 언론, 지도), 그림, 풍습, 습관, 음악, 무용, 음식, 종교, 유물, 언어, 건축, 의복 등
- ▶ 철학, 문학, 사학, 인류학, 언어학, 음악학, 종교학, 고고학, 지리학
  - ▶ 사회학, 정치학, 경제·경영학, 의학, 심리학

# 인간의 행위와 생각에 대한 정보 모집

- ▶ 현대인물
  - ▶ 인터넷 + IOT : SNS, GPS, 쇼핑 결제, RFID 센서, 스마트와치
  - ▶ 분석이 바로 가능하다
- ▶ 고대인물
  - ▶ 글, 유물, 유적
  - ▶ 옛날 사람을 이해하려면 아날로그 된 정보를 데이터로 전환
  - ▶ 정보가 없거나 찾기 어려운 경우
  - ▶ 언어, 개념 문제

# 인문학 연구 방법

- ▶ 인간을 이해하기 위한 자료
  - ▶ 도서관, 아카이브, 박물관, 유적
  - ▶ 온라인으로 제공되 데이터
- ▶ 한 학자 또는 연구팀
- ▶ 해당 주제에 대한 자료 모집
- ▶ 자료 검토
- ▶ 팩트 정리
- ▶ 정리한 팩트로 논쟁, 스토리 만든다

왜 이 주제, 왜 이 자료, 왜 이 팩트, 왜 지금, 왜 이 이야기?

- 중요성 및 객관성

# 인문학 2.0

## ~~전통~~ 아날로그 인문학

- ▶ 인원 : 홀로
- ▶ 연구 범위 : 소규모
- ▶ 생각 정리 : 글
- ▶ 결론 전달 : 글
- ▶ 지식 : 소유, 언급, 저작권

## 디지털 인문학

- ▶ 인원 : 협업
- ▶ 연구 범위 : 대규모
- ▶ 생각 정리 : 데이터
- ▶ 결론 전달 : 시각화, 지도 등
- ▶ 지식 : 공유, 재사용, 출처표시

# 해외 디지털 인문학 연혁

- ▶ Roberto Busa (1913-2011) 이탈리아 예수회 신부
  - ▶ 1946 - *Index Thomisticus*
  - ▶ IBM와 협력하여 토마스 아퀴나스(13세기 기독교 사상사)의 작품을 부명제화
  - ▶ ~1970 완료
- ▶ Text Encoding Initiative (TEI)
  - ▶ 텍스트 암호화(인코딩) 사업
  - ▶ 1980년대~
  - ▶ XML 포맷 지침
- ▶ Humanities Computing → Digital Humanities

# 한국과 디지털 인문학 연혁

- ▶ Edward Wagner (1924-2001) - Harvard 교수님
  - ▶ 조선 양반에 대한 연구
  - ▶ 문과 프로젝트 1967~ / 문과 방목, 족보
- ▶ 조선왕조실록 DB
  - ▶ 디지털화 및 마크업
  - ▶ 1990년대 초반
- ▶ IT New Deal 사업
  - ▶ 1997 IMF 위기
  - ▶ 역사 자료 디지털화
- ▶ 학회구성 - 2000년대 중반
- ▶ 다양한 주제로의 콘텐츠
  - ▶ 족보, 일기, 실록, 3.1 운동, 국채보상운동, 70년대 연극, 중인, 교육 정책
- ▶ 한문 고문헌 AI 번역

# Markup Language 마크업 언어

- ▶ 마크업은 '문서에 주석, 메모를 적는다'
- ▶ XML – eXtensible Markup Language
  - ▶ 확장 가능한
- ▶ HTML – Hypertext Markup Language
  - ▶ Hypertext는 다른 문자로 연결이 가능한 문자, 즉 순차적인 문자를 넘는다
  - ▶ 웹 기반 문서 연결, 이동을 위한 것
  - ▶ XML에 비해 한정된 용도
- ▶ 마크업은 태그(tag)활용한다
  - ▶ Hello, my name is <person>Lyndsey</person>.
- ▶ 정보 추출
  - ▶ 같은 기사에 언급한 인물
  - ▶ 인물과 장소
  - ▶ 인물 목록
- ▶ 다른 정보와 연결

# Database

- ▶ 여러 종류

## Relational

- ▶ SQL(Structured Query Language)라는 언어를 통해 DB와 소통

## NoSQL (Non-SQL, SQL외)

- ▶ Triple-store or RDF
  - ▶ 순이 --> 좋아하다 --> 돌이
  - ▶ SPARQL
- ▶ Graph
  - ▶ 네트워크망 형식
  - ▶ Neo4J - Cypher

# Markup + Database의 중요성

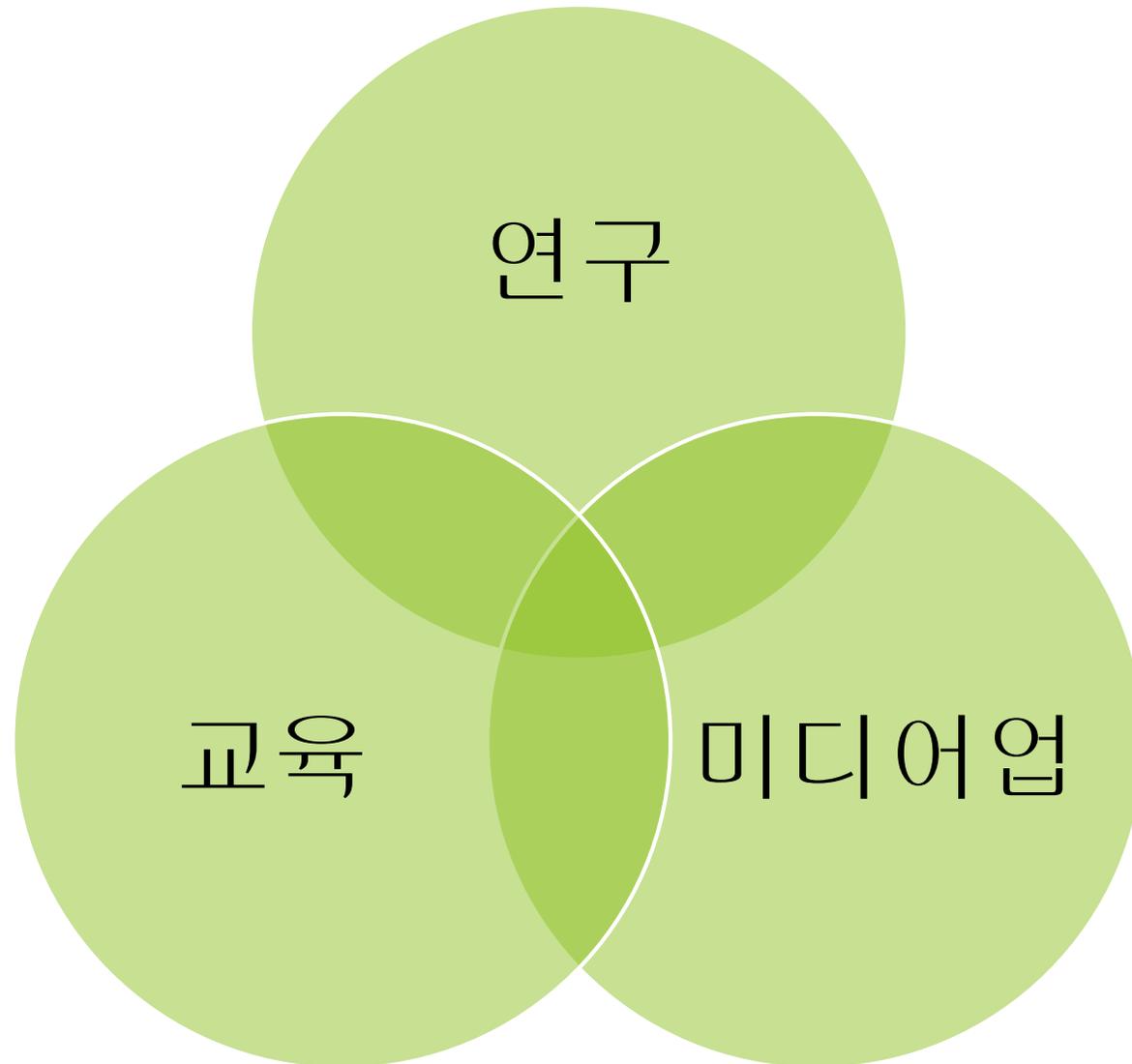
- ▶ 마크업을 통한 구조를 알아야 **CSS**를 통해 스타일(디자인)을 입힐 수 있고 데이터 추출이 가능하다
- ▶ **Python**이나 다른 프로그래밍언어로 **DB**에 있는 정보를 불러려고 할 때 **DB**의 구조를 이해하고 **SQL**과 같은 쿼리 언어도 알아야 가능하다

AKS 디지털인문학연구소 연구 및 교육

# 주로 하는 연구

- ▶ **DB 기반 콘텐츠 제작**
  - ▶ 데이터베이스 구축
  - ▶ 데이터 생산 및 입력
  - ▶ 데이터 시각화 (네트워크 그래프, 지도)
  - ▶ 데이터 분석
- ▶ **가상현실 콘텐츠 제작**
  - ▶ 360도 사진 및 동영상, 드론 촬영
  - ▶ 가상 갤러리 만들기
- ▶ **미디어위키로 프로젝트 관리 및 콘텐츠 공유**
  - ▶ 기사 작성, 분류, 미디어 콘텐츠, 협업, 하이퍼링크 등
- ▶ **AI 번역**

# 사업 목적



# DB기반 프로젝트 진행 방법

1. 주제 선정
  2. 자료 조사
    1. 종류, 양
    2. 내용의 성격
  3. 온톨로지 구성
  4. **DB** 구축
  5. 데이터 생산
  6. 시각화 및 분석
- ▶ 위키로 정리, 진행 검토
  - ▶ **Iterative** - 작은 예시 만들고, 피드백 받고, 개선하고, 다시 하기

# 사용하는 프로그램

- ▶ Filezilla FTP
- ▶ MediaWiki
- ▶ Semantic MediaWiki
- ▶ MySQL
- ▶ Neo4J
- ▶ Gephi
- ▶ Protege
- ▶ PTGui
- ▶ KR pano
- ▶ Google NMT

# 사용하는 프로그래밍 언어

- ▶ SQL, Cypher
- ▶ MediaWiki Syntax
- ▶ XML
- ▶ HTML
- ▶ CSS
- ▶ Vis.js
- ▶ Python
- ▶ Command Line

# 연구 사례

## 국내

- ▶ [Encyves](#)
- ▶ [국립중앙박물관 미술품 갤러리](#)
- ▶ [Korean Ceramics](#)
- ▶ [VH2018 한국의 불교 사찰](#)
- ▶ [Korean Cultural Heritage VR Materials](#)

## 해외

- ▶ [Kindred Britain](#)
- ▶ [China Biographical Database](#)
- ▶ [Lviv Interactive](#)
- ▶ [Islandic Saga Map](#)
- ▶ [Flooding in the Low Countries Inundation Map](#)