

# 데이터베이스 기말 프로젝트 본과제

2022년 5월

# 기말 프로젝트

- 제목: 네이버 영화 구축
- 내용: 네이버 영화 사이트(movies.naver.com)로부터 영화 데이터를 수집하여, mySQL에 DB화하고, 검색 인터페이스 제공하는 프로젝트
- 프로젝트 주요 구성요소
  - DB 설계
  - 네이버 영화 데이터 수집(crawling) 및 DB 구축
  - 검색 인터페이스 구현
- 팀구성: 1인 ~ 3인

# 프로젝트 specification (1)

- 네이버 영화 DB 구축
  - 네이버 영화디렉토리의 링크를 활용하여, 가능한 많은 영화를 수집하여, DB화 할 것.
    - 평가요소: 수집된 영화의 수
  - 네이버의 각 영화 정보를 나타내는 URL에서 보이는 가능한 모든 정보를 수집하여 DB화 할 것. 그대로 재현할 수 있으면 가장 이상적임.
    - 예: 장르, 주요정보, 배우/제작진, 포토, 동영상, 평점, 리뷰 등등
    - 평가요소: 영화 정보 속성의 갯수 및 충실성
  - DB설계에서 배운 데이터 모델링 기법을 적용하여, normalized 된 형태로 DB를 설계하고, 저장할 것.
    - 예: 범죄도시2 에서 "범죄, 액션"은 해당영화의 장르로서 다중값을 가짐.
    - 평가요소: 제대로된 설계가 되었는가?
- 데이터 수집 및 DB 구축 언어는 python이어야 함.

# 프로젝트 specification (2)

- 영화 검색 인터페이스 구현
  - 다양한 속성으로 영화를 검색하는 웹인터페이스를 구현할 것.
  - 검색 속성 예: 영화제목, 배우이름, 감독이름, 장르, 년도, 국가 등
  - 정렬 속성 예: 가나다순, 년도순, 평점순 등
  - 검색결과 리스트에서 영화를 클릭시, 최대한 네이버와 동일한 화면 구성으로 영화정보를 보여줄 것.
  - 검색의 속도 향상을 위해서 적절한 index를 구축할 것.
  - 검색 인터페이스 개발언어는 무관.
  - 평가요소: 검색 인터페이스의 작동성, 검색결과의 정확성, 검색 속도, 영화정보 화면의 충실성, 검색 속도 등

# 과제 일정

- 과제 Q&A : 조교를 통해서 질문내용을 미리 전달하면, 15주차 강의에서 질의응답을 할 예정임.
- 팀구성: 1인 ~ 3인
  - 팀원의 역할 분담 내용을 보고서에 쓰고, 동영상에서 각자가 맡은 주요 부분을 발표할 것.
- 팀단위 과제 제출: 과제 제출시 팀내 팀장만 과제를 제출하고, 구성원이 있을 경우, 해당학생들의 학과/학번/이름을 명시
- 제출물
  - 보고서
  - 소스코드
  - 유튜브 발표 동영상 (10분 내외)
- 과제 제출 마감: 6월 19일