

게시판 내 논문 정보 크롤링 구현 프로젝트

컴퓨터공학전공 2학년

김수정

정적 웹 페이지

- 정적 웹 페이지
 - 특정 url만으로 항상 같은 화면 출력
- 동적 웹 페이지
 - 사용자의 조작에 의해 실시간으로 화면 전환
 - ex) 네이버 메일, 네이버 지도, 인스타그램

정적 수집

- 정적 페이지의 html을 파싱하여 정보 수집
- 과정
 1. 목표로 하는 웹 페이지의 HTML 다운로드
 2. 가져온 HTML 문서 전체 파싱
 3. 필요한 정보만 추출하여 리스트 생성
 4. 리스트 출력

정적 수집

1. 목표로 하는 웹 페이지의 HTML 다운로드 – requests
2. 가져온 HTML 문서 전체 파싱 – BeautifulSoup

```
# 필요한 패키지 가져오기
import requests
from bs4 import BeautifulSoup as bs

url = 'https://www.naver.com' # 가져올 url 문자열로 입력
response = requests.get(url) # requests의 get함수를 이용해 해당 url의 html 문서가 담긴 자료 받기
html_text = response.text    # html 문서만 빼기

# html을 잘 정리된 형태로 변환 ==> 파싱(parsing)
html = bs(html_text, 'html.parser')
```

BeautifulSoup 함수

- find() / find_all()
 - HTML에서 원하는 태그 검색
 - HTML의 태그 및 속성 이용

The diagram illustrates the structure of an HTML tag: `<p class="para">TCPSchool.com</p>`. Labels with arrows and brackets identify the parts: '태그 이름' (tag name) points to 'p'; '속성명' (attribute name) points to 'class'; '속성값' (attribute value) points to '"para"'; '내용' (content) points to 'TCPSchool.com'; '시작 태그' (start tag) points to the entire opening tag; and '종료 태그' (end tag) points to the closing tag '</p>'.

BeautifulSoup 함수

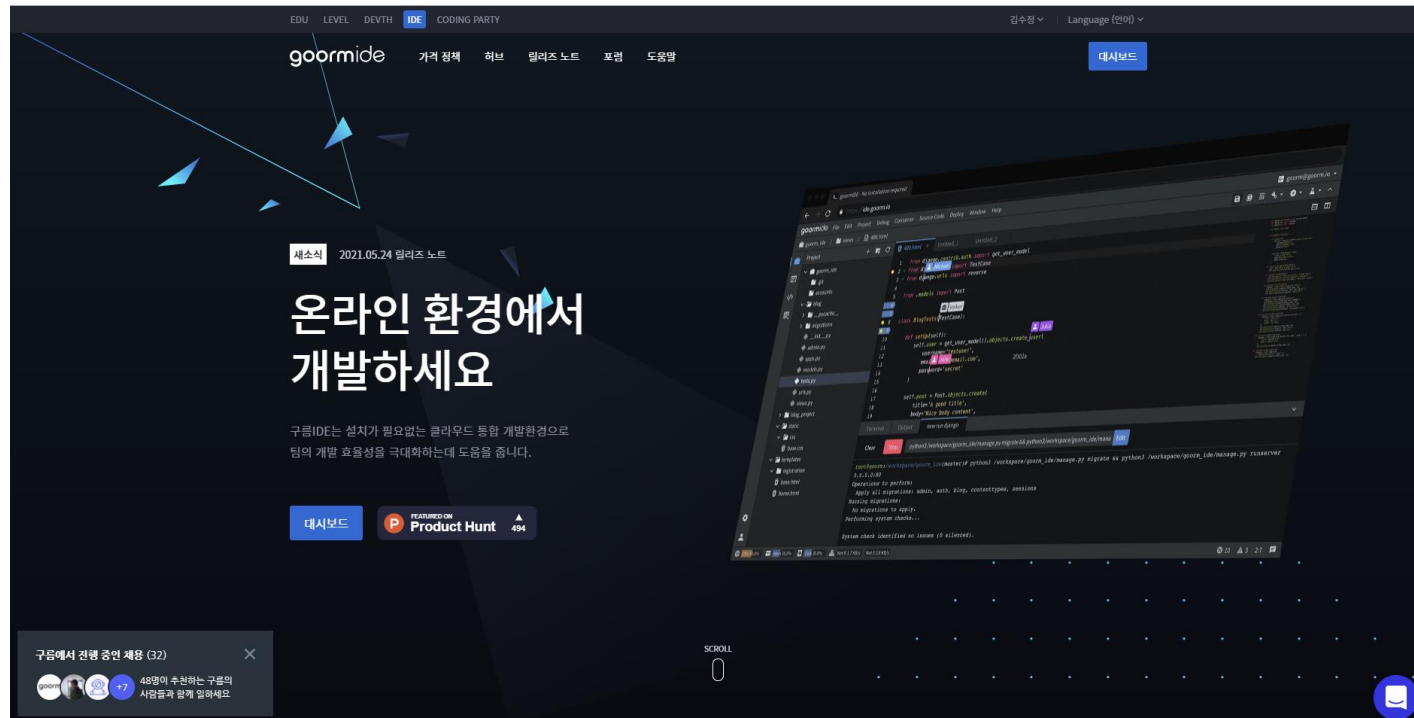
- select_one() / select()
 - CSS에서 HTML 태그 검색
 - CSS 선택자 이용

The diagram illustrates the components of a CSS selector and declaration block. The selector 'a' is labeled '선택자' (Selector) with a downward arrow. The opening curly brace '{' is labeled '선언 시작' (Declaration start) with an upward arrow. The declaration 'background-color: yellow;' is grouped by a blue bracket labeled '속성명' (Property name). The value 'yellow' is labeled '속성값' (Property value) with a yellow bracket and an upward arrow labeled '선언 구분자' (Declaration separator). The declaration 'font-size: 16px;' is grouped by a blue bracket labeled '속성명' (Property name). The value '16px' is labeled '속성값' (Property value) with a yellow bracket and an upward arrow labeled '선언 끝' (Declaration end). The closing curly brace '}' is labeled '선언 끝' (Declaration end) with an upward arrow.

```
a { background-color: yellow; font-size: 16px; }
```

Python 웹 크롤링 예제 연습

- 개발 환경
 - 구름 IDE - 웹 기반 클라우드 통합 개발 환경



1) 네이버 뉴스 기사 제목 크롤링



웹 크롤링 예제

1) 네이버 뉴스 기사 제목 크롤링

```
news_search_covid19.py ×
1 from bs4 import BeautifulSoup
2 import requests
3
4 url = 'https://search.naver.com/search.naver?where=news&sm=tab_jum&query=코로나'
5
6 response = requests.get(url)
7 html_text = response.text
8
9 html = BeautifulSoup(html_text, 'html.parser')
10
11 titles = html.select('a.news_tit')
12
13 for i in titles:
14     title = i.get_text()
15     print(title)
```

```
root@goorm:/workspace/Python_Crawling# python3 news_search_covid19.py
코로나 신규확진 587명...유흥업소·변이바이러스 '촉각' [종합]
코로나19 확진 인도 주재원 등 14명 특별기편으로 귀국(종합)
앤씨-넷마블, 코로나19 백신 유급휴가 시행
"시간이 많지 않다"...본격 코로나 기원 찾기 나서나
[속보] 코로나19 신규확진 587명...지역발생 571명
티몬, 코로나19 백신휴가 시행...접종 회차별 유급휴가 지급
[속보] 국내 코로나19 항체 보유율 0.27%...2천248명 중 6명서 확인
현대건설, 코로나 시국에 특화된 새 'H 시리즈' 개발
'코로나 격무' 간호직 공무원 극단 선택...부산시, 인력 총원 대책 발표
바이든, '90일 재조사' 코로나19 보고서 "공개하겠다"
```

웹 크롤링 예제

2) 네이버 뉴스 검색 결과 크롤링

```
news_search.py ×
1 from bs4 import BeautifulSoup
2 import requests
3
4 query = input('검색할 키워드를 입력하세요: ')
5 url = 'https://search.naver.com/search.naver?where=news&sm=tab_jum&query=%s' % query
6
7 response = requests.get(url)
8 html_text = response.text
9
10 html = BeautifulSoup(html_text, 'html.parser')
11
12 titles = html.select('a.news_tit')
13
14 for i in titles:
15     title = i.get_text()
16     print(title)
17
```

```
root@goorm:/workspace/Python_Crawling# python3 news_search.py
검색할 키워드를 입력하세요: 대구가톨릭대학교
대구가톨릭대 병원 박기혁 교수 우수논문상 수상
대구가톨릭대, 신명고서 '전공 탐색의 날' 운영
대구시의회, 대구가톨릭대와 상호협력 MOU
대구가톨릭대, 고교생 대상 '전공 탐색의 날' 행사 실시
[대구소식]화성자원봉사단, 적십자와 '든든도시락 나눔 봉사' 등
대구시의회·대구가톨릭대, MOU 체결
[포토] 대구가톨릭대, '코비 트리' 전시
대구 4개 대학병원, 경북권 감염병전문병원 유치전 본격화
대구 달성교육지원청, 돌봄교실 운영, 지역대학과 손잡다.
대구시의회-대구가톨릭대, 지역사회 발전 상호협력
```

웹 크롤링 예제

3) 구글 검색 결과 크롤링

The image shows a Google search for '대구가톨릭대학교' (Daegu Catholic University). The search results page is visible, showing the university's website link and a Wikipedia entry. A developer tools overlay is present on the right side, displaying the DOM structure of the search results. The DOM tree shows the following structure:

```
<div id="center_col">
  <style>...</style>
  <div id="taw">...</div>
  <div class="eqAnXb" id="res" role="main">
    <div id="topstuff">...</div>
    <div id="search">
      <div data-hveid="CAEQNA" data-ved="2ahUKEwiTw0Oo2uvvAhXJa94KHSpbBugQGnoECAEQNA">
        <h1 class="Uo8X3b">검색결과</h1>
        <div id="3IewYJP1WnX-QaqtPnADg" data-async-context="query:%EB%8C%80%EA%B5%AC%EA%B8%80%ED%86%A8%EA%B6%AD%EB%8C%80%ED%95%99%EA%B5%90" id="rso">
          <div class="g" data-hveid="CAGQAA">...</div>
          <div class="g">
            <h2 class="Uo8X3b">웹 검색결과</h2>
            <div data-hveid="CAYQAA" data-ved="2ahUKEwiTw0Oo2uvvAhXJa94KHSpbBugQF5gAMAd6BAGGEAA">
              <div class="tF2Cxc">
                <div class="yuRbf">
                  <a href="https://namu.wiki/w/%EB%8C%80%EA%B5%AC%EA%B8%80%ED%86%A8%EA%B6%AD%EB%8C%80%ED%95%99%EA%B5%90" data-ved="2ahUKEwiTw0Oo2uvvAhXJa94KHSpbBugQFjAHegQI8hAD" ping="/url?sa=t&source=web&rc=t&jurl=https://namu.wiki/w/%25EB%258C%2580%25EA%B5%25AC%25EA%B8%2580%25ED%2586%25A8%25E8%25A6%25AD%25EB%258C%2580%25ED%2595%2599%25EA%B5%2590&ved=2ahUKEwiTw0Oo2uvvAhXJa94KHSpbBugQFjAHegQI8hAD&csid=1622181878416275">
                    <br>
                    <h3 class="LC201b DKV0Md">대구가톨릭대학교 - 나무위키</h3> == $0
                  </a>
                  <div class="TbwUpd Njxre">...</div>
                  </div>
                  <div class="B6fmyf">...</div>
                  </div>
                  <div class="IsZvec">...</div>
                  <div jscontroller="m6a01" id="eob_15" jsdata="fxg5tf;_j;AMdnNk" jsaction="rcuQ6b:npT2md" data-ved="2ahUKEwiTw0Oo2uvvAhXJa94KHSpbBugQZ28BMA6BAGGEAs">...</div>
                </div>
              </div>
            </div>
          </div>
        </div>
      </div>
    </div>
  </div>
  <div id="bottomads">...</div>
  <div id="hntstuff">...</div>
```

웹 크롤링 예제

3) 구글 검색 결과 크롤링

```
google_search_result.py ×  
1 from bs4 import BeautifulSoup  
2 import requests  
3 from urllib.parse import quote_plus  
4  
5 baseUrl = 'https://www.google.com/search?q='  
6 query = input('무엇을 검색할까요? :')  
7 searchUrl = baseUrl + quote_plus(query)  
8 url = baseUrl + '%s'%query  
9  
10 response = requests.get(url)  
11 html_text = response.text  
12  
13 html = BeautifulSoup(html_text, 'html.parser')  
14  
15 results = html.find_all('h3')  
16  
17 print(searchUrl)  
18 for i in results:  
19     result = i.get_text()  
20     print(result)  
21
```

```
root@goorm:/workspace/Python_Crawling# python3 google_search_result.py  
무엇을 검색할까요? :대구가톨릭대학교  
https://www.google.com/search?q=%EB%8C%80%EA%B5%AC%EA%B0%80%ED%86%A8%EB%A6%AD%E3%84%B7%EB%8C%80%ED%95%99%EA%B5%90  
대구가톨릭대학교  
대구가톨릭대학교 효성캠퍼스  
대구가톨릭대학교 입학종합정보  
공지사항 > 공지사항 > 광장 - 대구가톨릭대학교  
학생지원통합시스템  
강의지원시스템 - 대구가톨릭대학교 강의지원시스템 System  
대표 지도 > 지도다운로드 > 캠퍼스안내 > 소개 - 대구가톨릭대학교  
대구가톨릭대학교 의과대학  
대구가톨릭대학교 - 나무위키  
대구가톨릭대학교 - 위키백과, 우리 모두의 백과사전  
대구가톨릭대학교 - Home | Facebook
```

4) 멜론 음원차트 top100 크롤링



웹 크롤링 예제



4) 멜론 음원차트 top100 크롤링

```
melon_chart.py x
1 from bs4 import BeautifulSoup
2 import requests
3
4 header = {'User-Agent': 'Mozilla/5.0 (Windows NT 6.3; Trident/7.0; rv:11.0) like Gecko'} #헤더정보
5 response = requests.get('https://www.melon.com/chart/week/index.htm', headers=header)
6 html_text = response.text
7
8 html = BeautifulSoup(html_text, 'html.parser')
9
10 songlist = html.find_all("div", {"class": "ellipsis rank01"})
11 artistlist = html.find_all("div", {"class": "ellipsis rank02"})
12
13 song = []
14 artist = []
15
16 for s in songlist:
17     song.append(s.find('a').text)
18 for a in artistlist:
19     artist.append(a.find('a').text)
20
21 rank = 100
22 for i in range(rank):
23     print(' %3d위 : %s - %s'%(i+1, song[i], artist[i]))
24
```

```
root@goorm:/workspace/Python_Crawling# python3 melon_chart.py
1위 : Peaches (Feat. Daniel Caesar & Giveon) - Justin Bieber
2위 : 롤린 (Rollin') - 브레이브걸스
3위 : Dun Dun Dance - 오마이걸 (OH MY GIRL)
4위 : 라일락 - 아이유
5위 : 운전만해 (We Ride) - 브레이브걸스
6위 : Celebrity - 아이유
7위 : Timeless - SG 워너비
8위 : Dynamite - 방탄소년단
9위 : 상상더하기 - 라붐 (LABOUM)
10위 : Next Level - aespa
11위 : 라라라 - SG 워너비
12위 : 마.피.아. In the morning - ITZY (있지)
13위 : 내 손을 잡아 - 아이유
14위 : 내사람 - SG 워너비
15위 : Butter - 방탄소년단
16위 : 밝게 빛나는 별이 되어 비취줄게 - 송이한
17위 : Coin - 아이유
18위 : ASAP - STAYC(스테이씨)
19위 : 멜로디 - ASH ISLAND
20위 : LOVE DAY (2021) (바른연애 길잡이 x 양요섭, 정은지) - 양요섭
21위 : 사이렌 Remix (Feat. UNEDUCATED KID, Paul Blanco) - 호미들
22위 : On The Ground - 로제 (ROSÉ)
23위 : 밤하늘의 별을 (2020) - 경서
24위 : 잠이 오질 않네요 - 장범준
25위 : Blueming - 아이유
26위 : Lovesick Girls - BLACKPINK
27위 : 취기를 빌려 (취향저격 그녀 x 산들) - 산들
28위 : 그냥 안아달란 말야 - 다비치
29위 : 에잇 (Prod.&Feat. SUGA of BTS) - 아이유
30위 : 어푸 (Ah puh) - 아이유
31위 : 헤븐 우연 - 헤이즈 (Heize)
32위 : VVS (Feat. JUSTHIS) (Prod. GroovyRoom) - 미란이 (Mirani)
33위 : Savage Love (Laxed - Siren Beat) (BTS Remix) - Jawsh 685
```

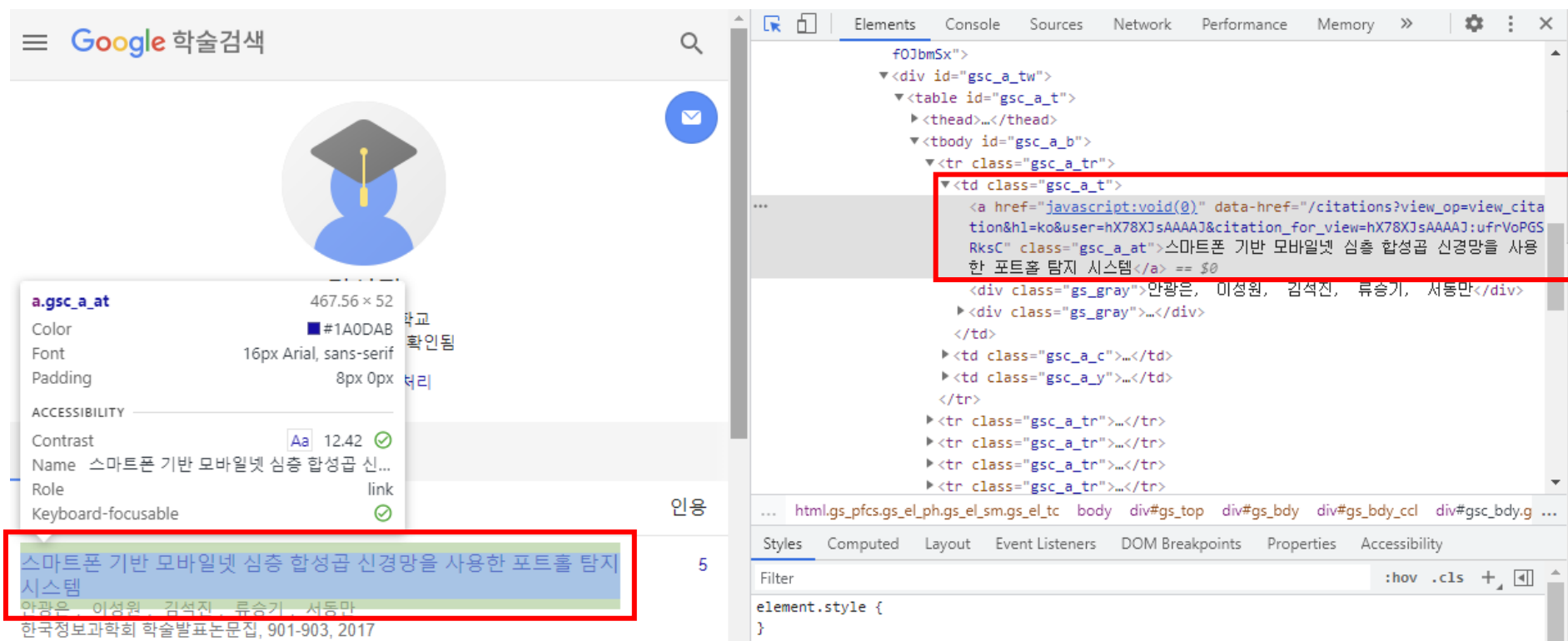
Google Scholar 개인 프로필의 논문 정보 크롤링

- Google Scholar 개인 프로필

Google 학술검색		
		
딤퍼닝 영상처리		
제목	인용	연도
스마트폰 기반 모바일넷 심층 합성곱 신경망을 사용한 포트홀 탐지 시스템 안광은, 이성원, 김석진, 류승기, 서동만 한국정보과학회 학술발표논문집, 901-903	5	2017
웹탐용 실시간 포트홀 탐지 시스템 이성원, 안광은, 김석진, 류승기, 서동만 한국정보과학회 학술발표논문집, 2113-2115	2	2017
모바일 기반 실시간 포트홀 탐지 어플리케이션 김석진, 이성원, 안광은, 류승기, 서동만 한국정보과학회 학술발표논문집, 1766-1768	2	2017
탐침 기반 토양 습도센서 데이터 수집 안정화를 위한 연구 김재훈, 김석진, 이동욱, 전수빈, 서동만 한국정보과학회 학술발표논문집, 1244-1246		2020
비대면 발열 검사를 위한 열화상 카메라 방식의 저비용 출입관리 시스템 구현 이재혁, 김수종, 김재훈, 김석진, 전수빈, 서동만 한국정보과학회 학술발표논문집, 1195-1197		2020

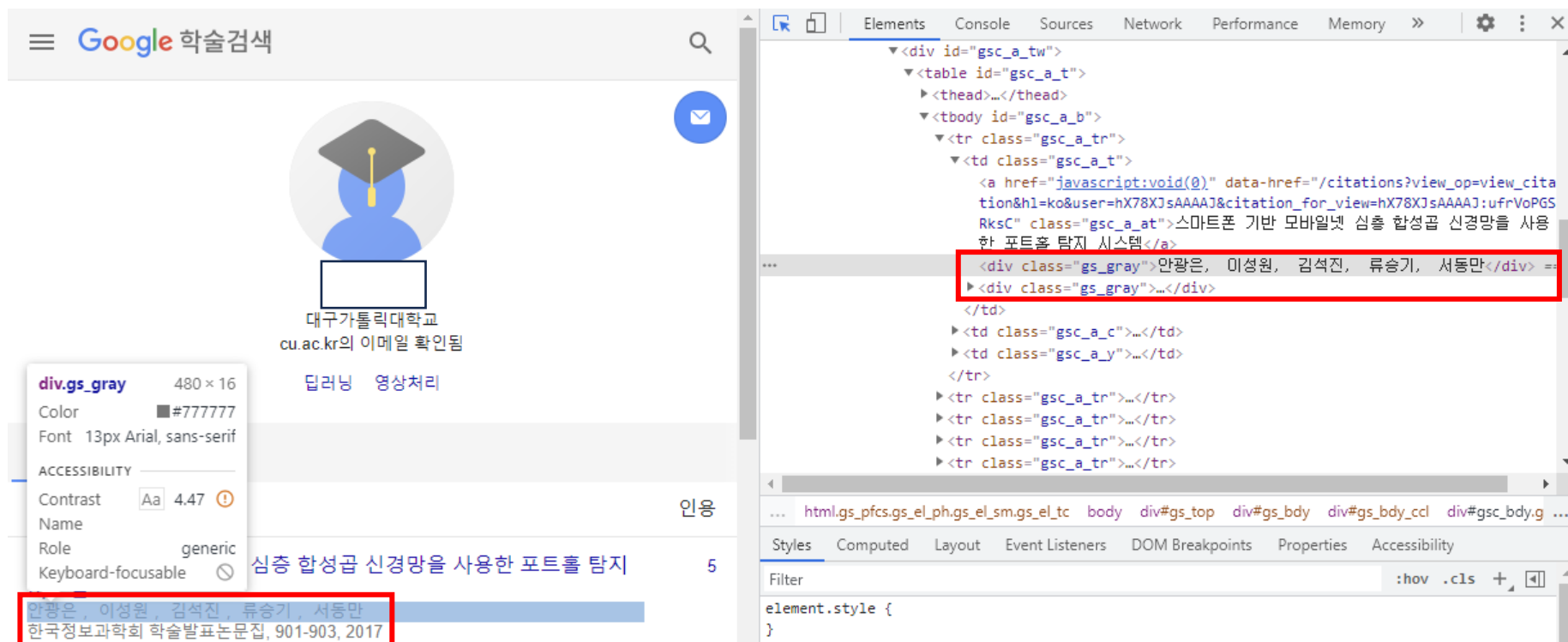
Google Scholar 개인 프로필의 논문 정보 크롤링

- 개인 프로필 논문 정보의 제목 정보



Google Scholar 개인 프로필의 논문 정보 크롤링

- 개인 프로필 논문 정보의 저자 및 발행기관 정보



Google Scholar 개인 프로필의 논문 정보 크롤링

- 크롤링 코드 및 실행 결과

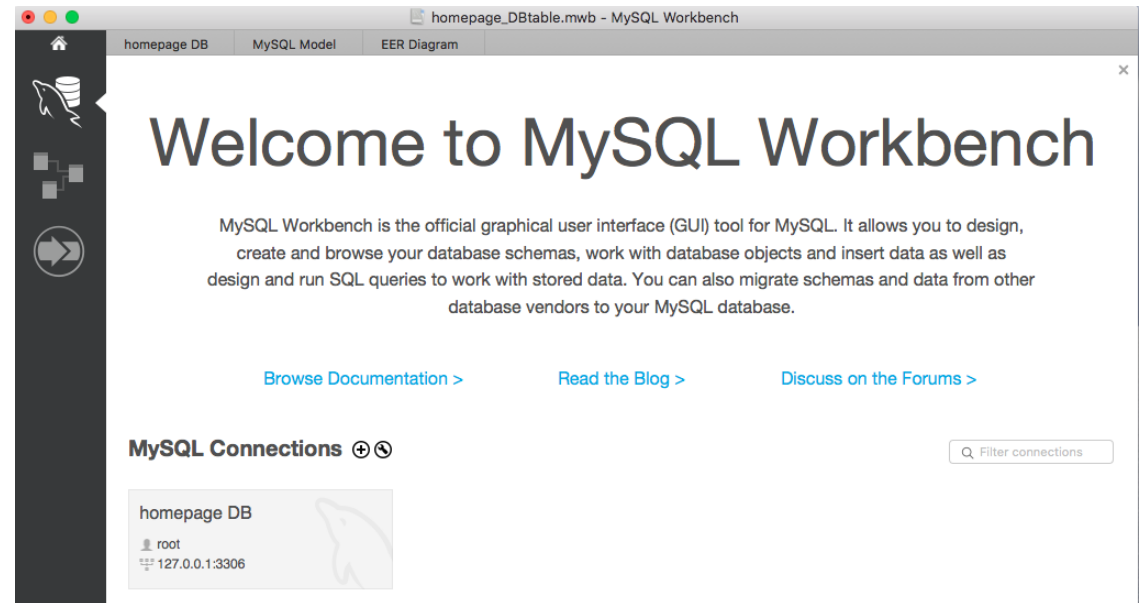
```
scholar.py x
1 import requests
2 from bs4 import BeautifulSoup
3
4 url = 'https://scholar.google.com/citations?user=hX78XJsAAAAJ&hl=ko&oi=ao'
5 response = requests.get(url)
6 html_text = response.text
7
8 html = BeautifulSoup(html_text, 'html.parser')
9
10 title_list = html.find_all("td", {"class": "gsc_a_t"})
11 info_list = html.find_all("div", {"class": "gs_gray"})
12
13 title = []
14 info = []
15
16 for t in title_list:
17     title.append(t.find('a').text)
18 for a in info_list:
19     info.append(a.text)
20
21 print("[논문 제목 리스트]")
22 print("\n".join(title))
23 print("\n[논문 저자 및 발행기관 리스트]")
24 print("\n".join(info))
25
```

```
root@goorm:/workspace/scholar_crawling# python3 scholar.py
[논문 제목 리스트]
스마트폰 기반 모바일넷 심층 합성곱 신경망을 사용한 포트홀 탐지 시스템
랩탑용 실시간 포트홀 탐지 시스템
모바일 기반 실시간 포트홀 탐지 어플리케이션
탐침 기반 토양 습도센서 데이터 수집 안정화를 위한 연구
비대면 발열 검사를 위한 열화상 카메라 방식의 저비용 출입관리 시스템 구현
항공 영상 내 오브젝트의 정확한 GPS 좌표 예측 방법 및 정복 표현
R-CNN 기반 포트홀 탐지 웹 서비스 시스템 구현
항공영상 기반 포트홀 위치 가시화 웹 서비스 구현 웹 어플리케이션
온라인 강의 상황에서의 프로그래밍 실습을 위한 온라인 저지 시스템 개발
항공 이미지에서의 Mask R-CNN 기반 포트홀 검출 연구
딥러닝 탐지용 임베디드 장치 간 처리속도 비교
다중 서버 기반 웹 서비스에서 하트비트를 이용한 서비스 안정화 방안 연구
MOSAICING INSIDE TUNNEL IMAGES FOR TUNNEL MONITORING
사용자 경험정보를 고려한 결정트리 기반 음악 추천 시스템
집단지성 기반 딥러닝용 데이터 셋 구축을 위한 이미지 라벨링 게임 어플리케이션
터널 내부의 파노라마 이미지 생성을 위한 이미지 특징 추출 알고리즘 비교 분석
교량 구조물 안전점검 자동화를 위한 정합 이미지 생성 시스템
양방향 통신 기반 침입자탐지 가정용 CCTV 시스템
데이터셋 생성을 위한 이미지 URI 및 메타데이터 수집 크롤러
딥러닝 기반 포트홀 탐지 시스템에서의 학습 모델 비교 분석
노인 복지 시설 CCTV 를 이용한 웹 기반 노인 행동 관리 서비스
모바일 기반 포트홀 탐지 시스템 효율화를 위한 영상 처리 및 딥러닝 기반 알고리즘 연구
IoT 기반 드론 확장 모듈 연구
기계학습과 이미지 처리 기반 실시간 포트홀 탐지 시스템 제안
반려 식물을 위한 스마트 화분

[논문 저자 및 발행기관 리스트]
안광은 , 이성원 , 김석진 , 류승기 , 서동만
한국정보과학회 학술발표논문집, 901-903, 2017
이성원 , 안광은 , 김석진 , 류승기 , 서동만
한국정보과학회 학술발표논문집, 2113-2115, 2017
김석진 , 이성원 , 안광은 , 류승기 , 서동만
```

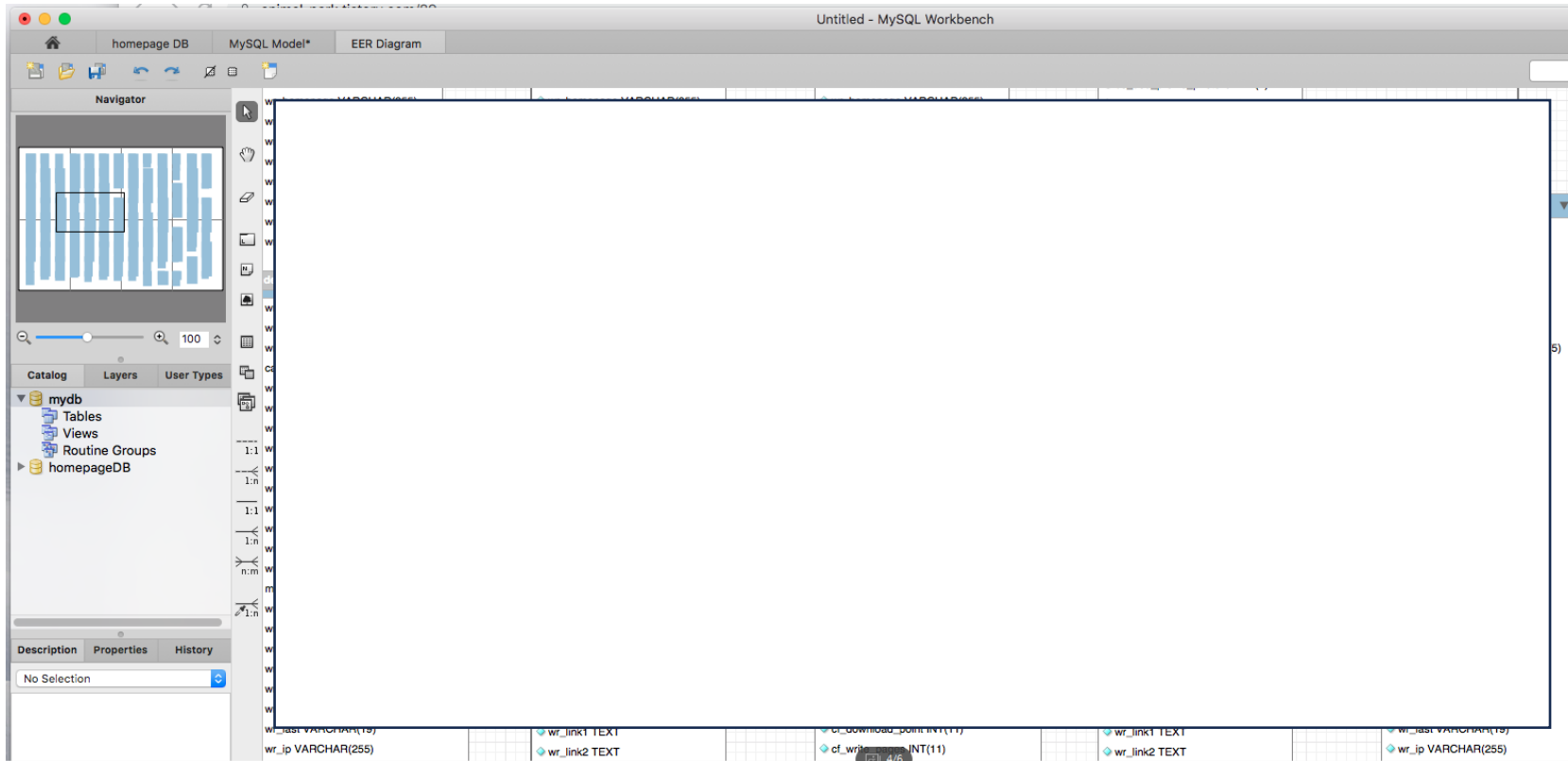
MySQL Workbench 설치

- MySQL
 - 오픈 소스 관계형 데이터베이스 관리 시스템
- MySQL Workbench
 - DB 설계 및 구현을 위한 도구



연구실 홈페이지 DB 테이블 구조

- 테이블 구조도 생성

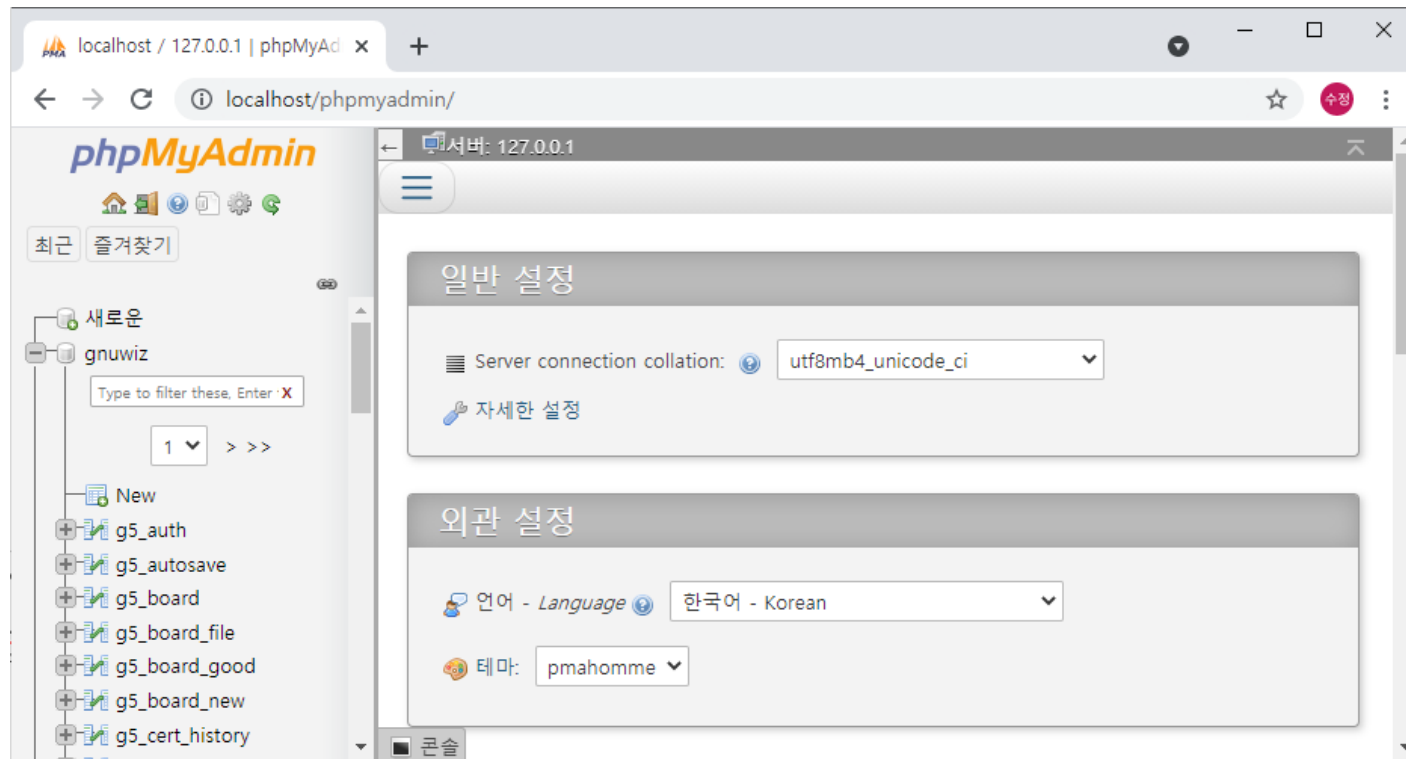


로컬 서버 구동

- XAMPP 설치
 - APM 환경을 한번에 묶어서 배포하는 소프트웨어 프로그램
 - APM (Apache & PHP & MySQL)
 - Apache : 웹 서버
 - PHP : 웹 프로그래밍 언어
 - MySQL : 데이터베이스

XAMPP 서버 동작 확인

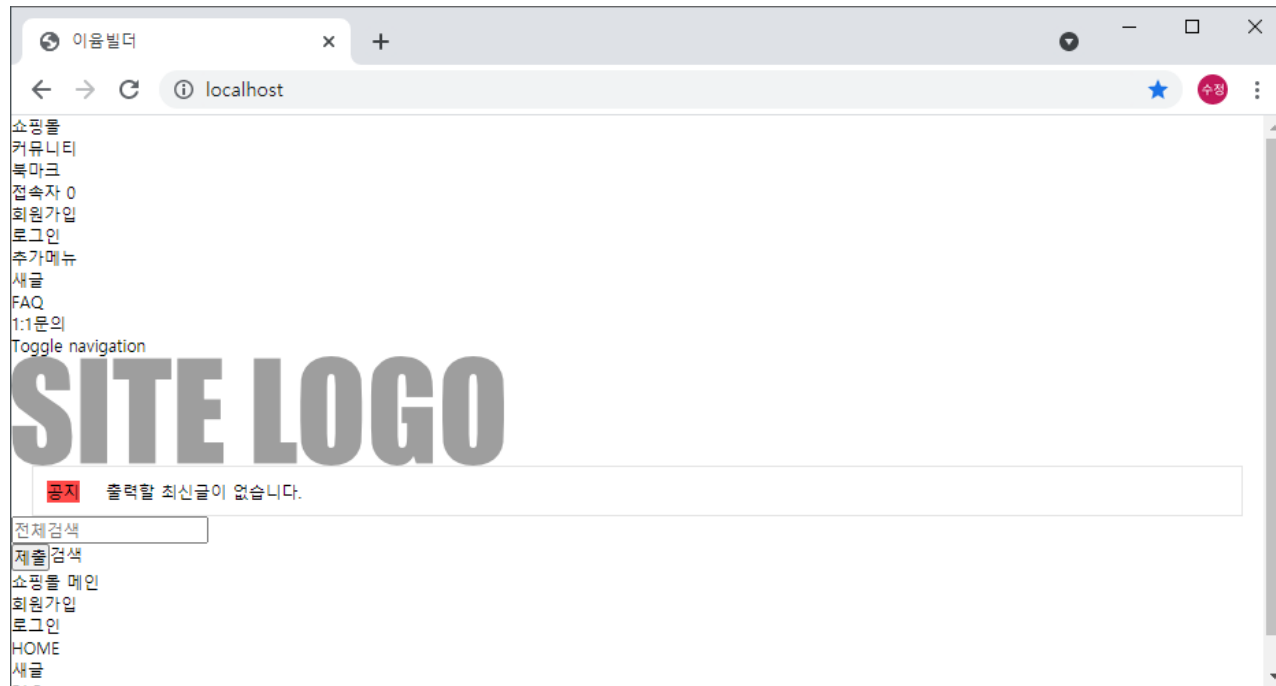
- localhost에서 서버 동작 확인



그누보드5 & 이음빌더 설치

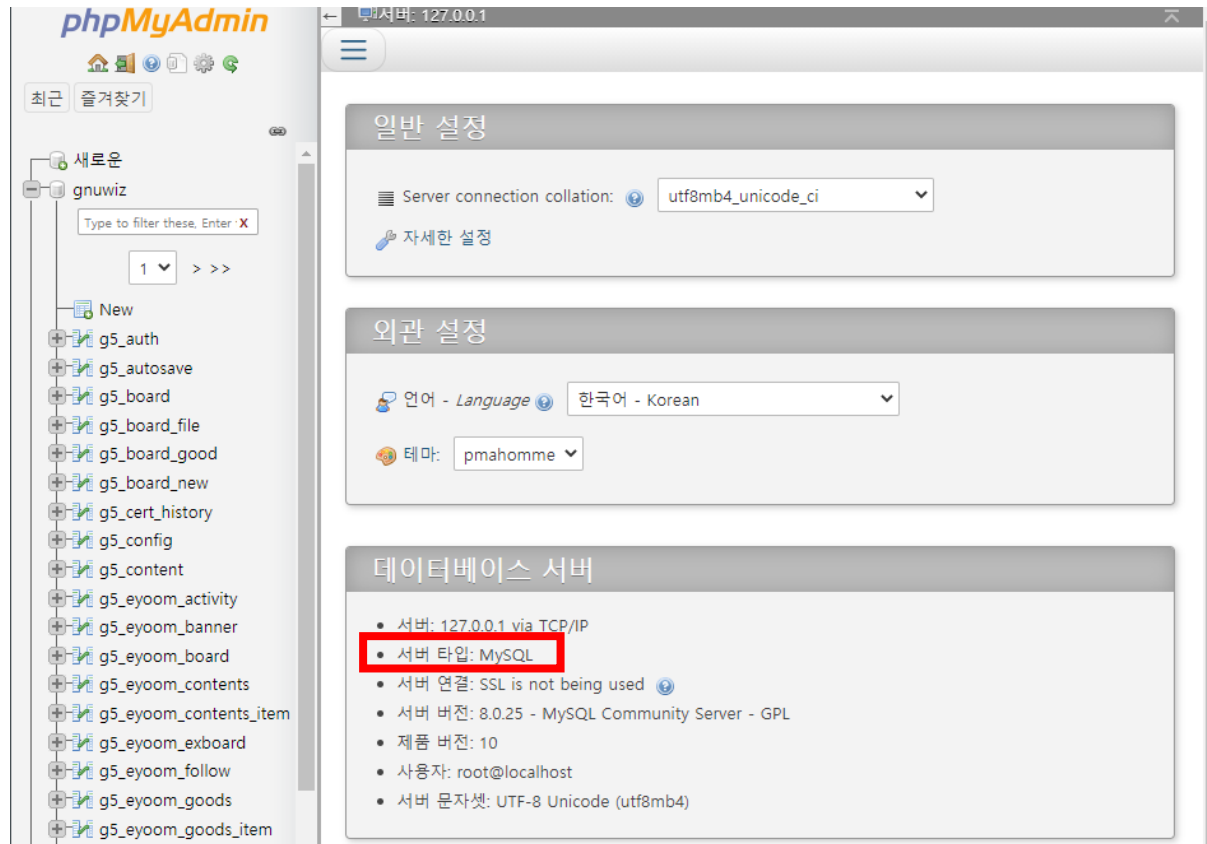
- 이슈 발생

- 설치 후, localhost/g5 에서 모든 버튼과 레이블들이 왼쪽으로 쏠림



그누보드 재설치 및 MySQL 변경

- MariaDB -> MySQL 변경



게시판 글 등록 관련 테이블 조사

- 자유게시판(free) 글 등록 테이블

	wr_id	wr_num	wr_reply	wr_parent	wr_is_comment	wr_comment	wr_subject	wr_content	wr_seo_title	ca_name	wr_comment_reply	wr_option
<input type="checkbox"/> 수정 <input type="checkbox"/> 복사 <input type="checkbox"/> 삭제	1	-1		1	0	1	안녕하세요	안녕하십니까 하하하 하하하하	안녕하세요			
<input type="checkbox"/> 수정 <input type="checkbox"/> 복사 <input type="checkbox"/> 삭제	2	-2		2	0	0	두번째 글입 니다	세번째아니고 두번째요	두번째-글입 니다			
<input type="checkbox"/> 수정 <input type="checkbox"/> 복사 <input type="checkbox"/> 삭제	3	-1	A	3	0	0	Re: 안녕하세 요	아름하세은	re-안녕하세요			
<input type="checkbox"/> 수정 <input type="checkbox"/> 복사 <input type="checkbox"/> 삭제	4	-2	A	4	0	0	Re: 두번째 글입니다	예예 두번째 H 답글이예 요	re-두번째-글 입니다			
<input type="checkbox"/> 수정 <input type="checkbox"/> 복사 <input type="checkbox"/> 삭제	5	-1		1	1	1		푸하하하				

게시판 글 등록 관련 테이블 조사

- 자유게시판(free) 글 등록 테이블

베이스: gnuwiz > 테이블: g5_write_free

검색 삽입 내보내기 가져오기 권한 테이블 작업 SQL 명령어 트래킹 트리거

wr_seo_title	ca_name	wr_comment_reply	wr_option	wr_link1	wr_link2	wr_link1_hit	wr_link2_hit	wr_hit	wr_good	wr_nogood	mb_id	wr_password	wr_name	wr_email	wr_homepage	wr_datetime	wr_file	wr_last	wr_ip
안녕하세요		1				0	0	1	0	0	admin		최고관리자	crystar2402@naver.com		2021-07-12 17:24:38	0	2021-07-12 18:05:32	::1
두번째-글입니다						0	0	1	0	0	admin		최고관리자	crystar2402@naver.com		2021-07-12 18:00:20	0	2021-07-12 18:00:20	::1
re-안녕하세요						0	0	1	0	0	admin		최고관리자	crystar2402@naver.com		2021-07-12 18:03:53	0	2021-07-12 18:03:53	::1
re-두번째-글입니다						0	0	1	0	0	admin		최고관리자	crystar2402@naver.com		2021-07-12 18:04:45	0	2021-07-12 18:04:45	::1
				a:0:{}		0	0	0	0	0	admin		최고관리자	crystar2402@naver.com		2021-07-12 18:05:32	0		::1

게시판 글 등록 관련 테이블 조사

- 자유게시판(free) 글 등록 테이블

📍 자유게시판

전체 4 건 - 1 페이지 ☐ 관심게시판 ☐ 글쓰기

번호	<input type="checkbox"/>	제목	글쓴이	날짜	뷰
4	<input type="checkbox"/>	세번째 글인데유 ~?	최고관리자	2021.07.12	2
3	<input type="checkbox"/>	두번째 글입니다	최고관리자	2021.07.12	1
2	<input type="checkbox"/>	Re: 두번째 글입니다	최고관리자	2021.07.12	1
1	<input type="checkbox"/>	+1 안녕하세요	최고관리자	2021.07.12	1
0	<input type="checkbox"/>	Re: 안녕하세요	최고관리자	2021.07.12	1

게시판 글 등록 관련 테이블 조사

- 자유게시판(free) 글 등록 테이블 (1/3)
 - wr_id : 등록 순 각 번호
 - wr_num : 원글 번호 (음수 / -1, -2, -3, ...)
 - wr_reply : 답글 여부 (대문자 알파벳 순 / A, B, C, D...)
 - wr_parent : 원글, 답글, 댓글의 원글 번호
 - wr_is_comment : 댓글 여부 (댓글(1), 아니면(0))
 - wr_comment : 댓글 갯수

게시판 글 등록 관련 테이블 조사

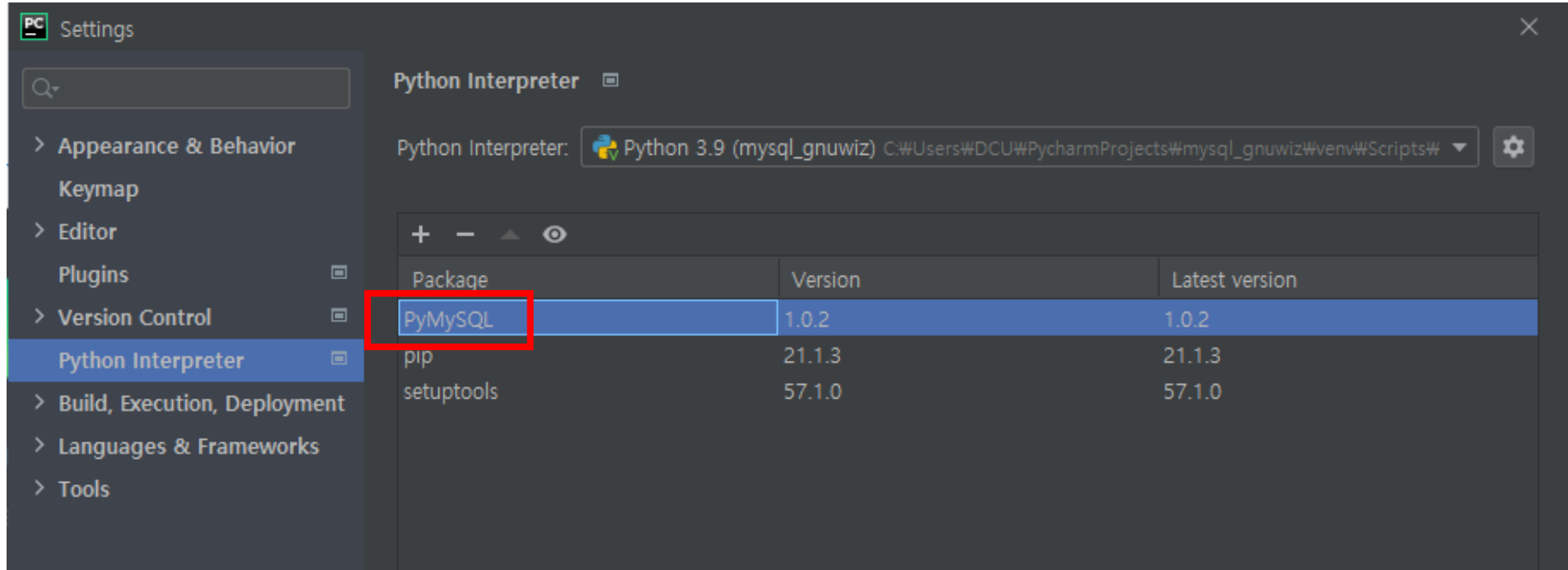
- 자유게시판(free) 글 등록 테이블 (2/3)
 - wr_subject : 제목
 - wr_content : 내용
 - wr_seo_title : 목록 표시용 제목 (대문자, 공백 X)
 - wr_hit : 조회수
 - wr_good : 추천수 / wr_nogood : 비추천수

게시판 글 등록 관련 테이블 조사

- 자유게시판(free) 글 등록 테이블 (3/3)
 - mb_id : 작성자 회원 아이디
 - wr_name : 작성자 이름
 - wr_email : 작성자 이메일
 - wr_datetime : 작성일
 - wr_last : 최근 수정일 (댓글 반영 최근일)
 - wr_ip : 작성자 아이피

Pycharm & MySQL 연동

- Pymysql 라이브러리 설치

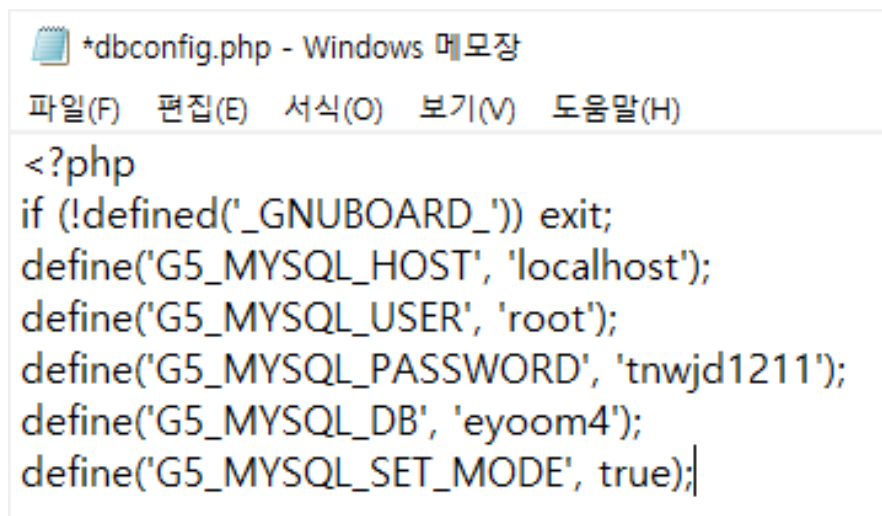


SQL Struct dump

- MySQL 테이블 구조 백업
 - phpmyadmin에 연구실 홈페이지 DB 덤프파일 업로드
 - 로컬 경로(g5)에 연구실 홈페이지 전체 백업 파일 업로드
 - 발생한 에러
 - 문제 : 정의되지 않은 함수 존재 → php 버전이 다르기 때문에 발생
 - 해결방법 : PHP 재설치 (PHP 7.1.3)

SQL Struct dump

- MySQL 테이블 구조 백업
 - dbconfig.php 파일 수정 (user & password 설정 변경)



```
*dbconfig.php - Windows 메모장
파일(F)  편집(E)  서식(O)  보기(V)  도움말(H)

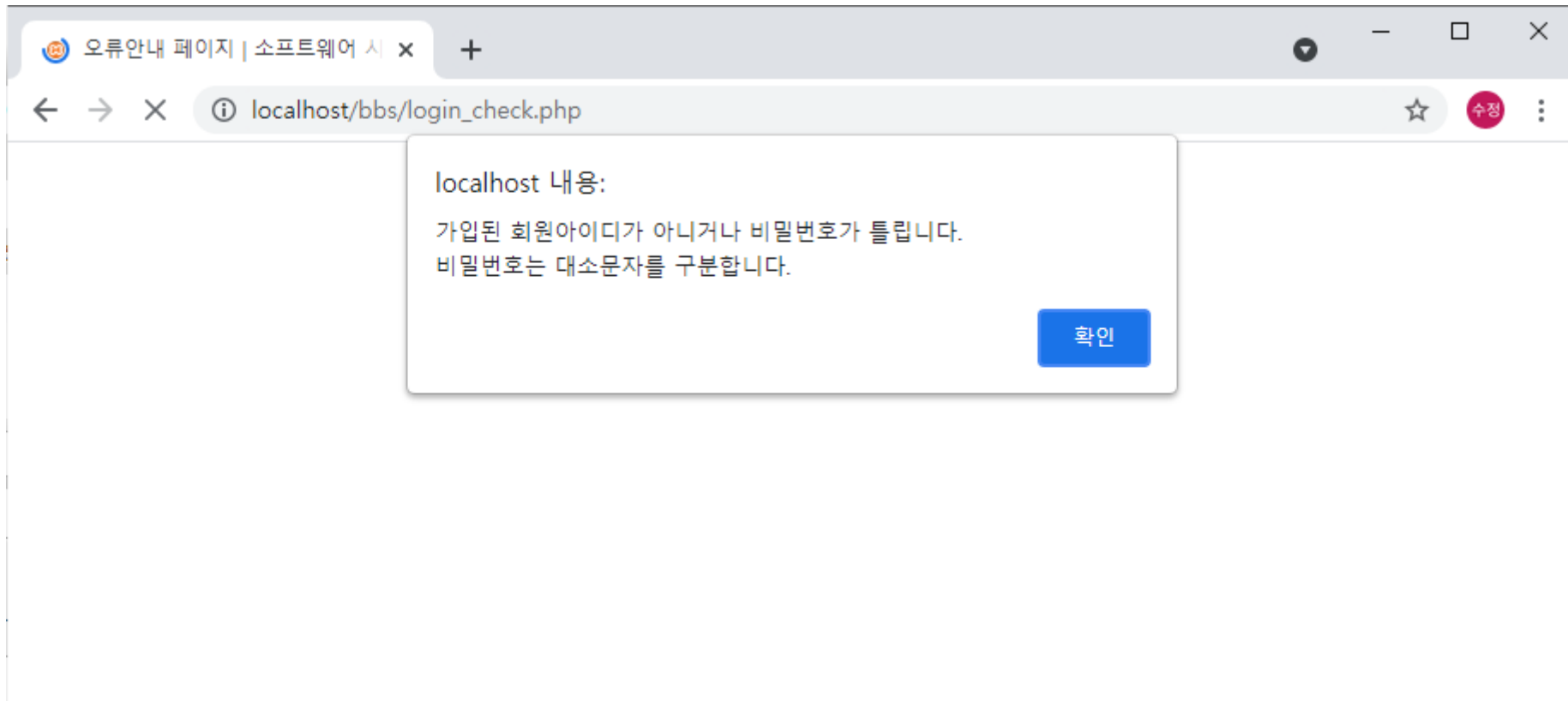
<?php
if (!defined('_GNUBOARD_')) exit;
define('G5_MYSQL_HOST', 'localhost');
define('G5_MYSQL_USER', 'root');
define('G5_MYSQL_PASSWORD', 'tnwjd1211');
define('G5_MYSQL_DB', 'eyoom4');
define('G5_MYSQL_SET_MODE', true);
```

SQL Struct dump

- MySQL 테이블 구조 백업
 - dbconfig.php 파일 수정 (user & password 설정 변경)
 - 발생한 에러
 - 문제 : 파일 수정 후 localhost/dashboard 실행 시 xampp 서버 확인 페이지만 등장
 - 해결 방법 : dashboard 상위 폴더에 g5 내 파일 (연구실 홈페이지 전체 백업 파일) 복사

이슈 발생

- localhost로 접속한 연구실 홈페이지에서 로그인 불가



이슈 발생

- localhost로 접속한 연구실 홈페이지에서 로그인 불가

서버: 127.0.0.1 > 데이터베이스: homepage > 테이블: g5_member

보기 구조 SQL 검색 삽입 내보내기 가져오기 권한 테이블 작업 SQL 명령어 트레이스

✓ 25 - 29행 표시 중 (합계 30, 질의 실행시간 0.0018 초.)

SELECT * FROM `g5_member`

☐ 프로파일링 [인라인 편집] [수정] [SQL 해석] [PHP 코드 보기] [다시 보기]

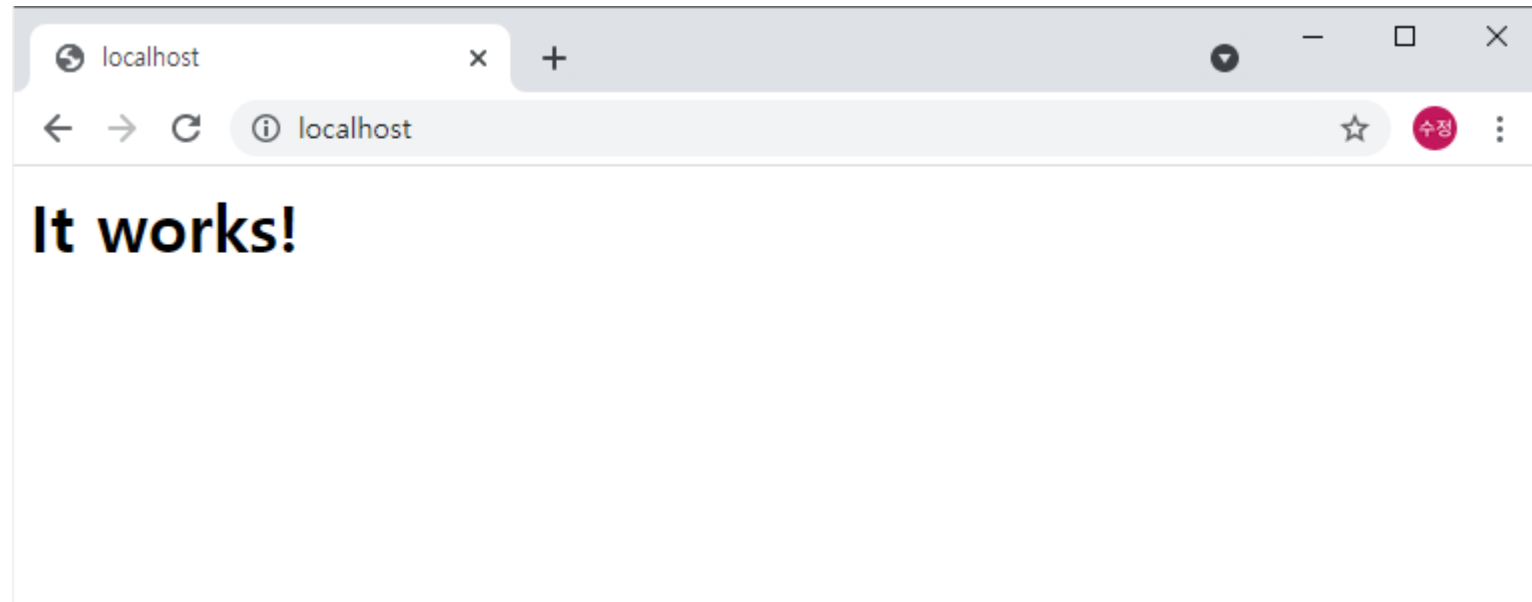
<< < 2 > >> ☐ 모두 보기 | 행 갯수: 25 | 행 필터링: 현재 테이블 검색 | Sort by key: 없음

+ 옵션

	mb_no	mb_id	mb_password	mb_name	mb_nick	mb_nick_date
<input type="checkbox"/> 수정 복사 삭제	26					2021-03-03
<input type="checkbox"/> 수정 복사 삭제	27					2021-02-25
<input type="checkbox"/> 수정 복사 삭제	28					2021-03-15
<input type="checkbox"/> 수정 복사 삭제	29					2021-03-29
<input type="checkbox"/> 수정 복사 삭제	30					2021-07-16

APM 재구축

- Apache 설치
 - 설정 완료 및 정상 실행



APM 재구축

- PHP 설치

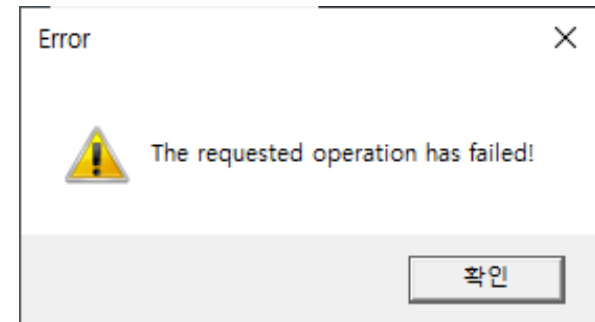
- 설정 완료 후 Apache 재실행 시 에러 발생

- 에러 메시지 창

- <The requested operation has failed>

- 명령 프롬프트

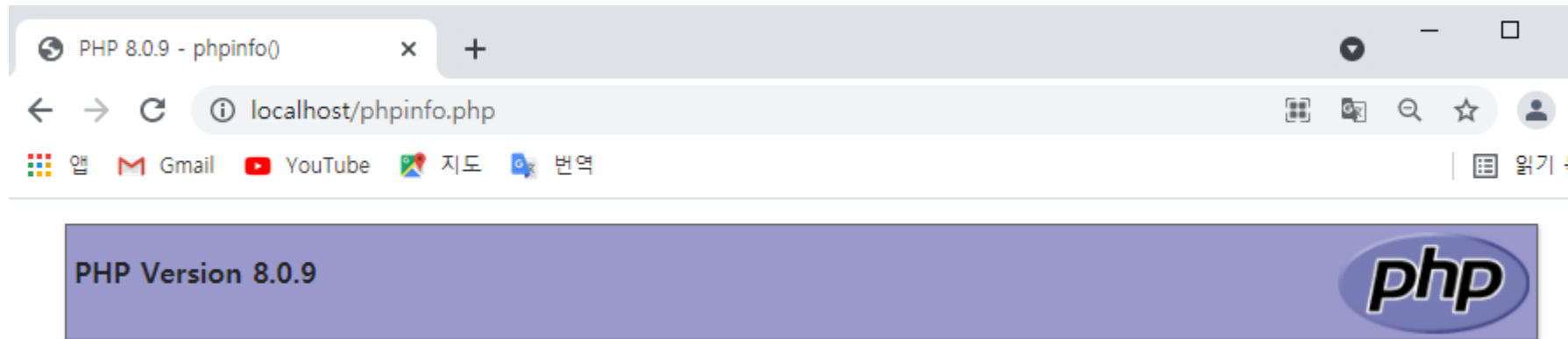
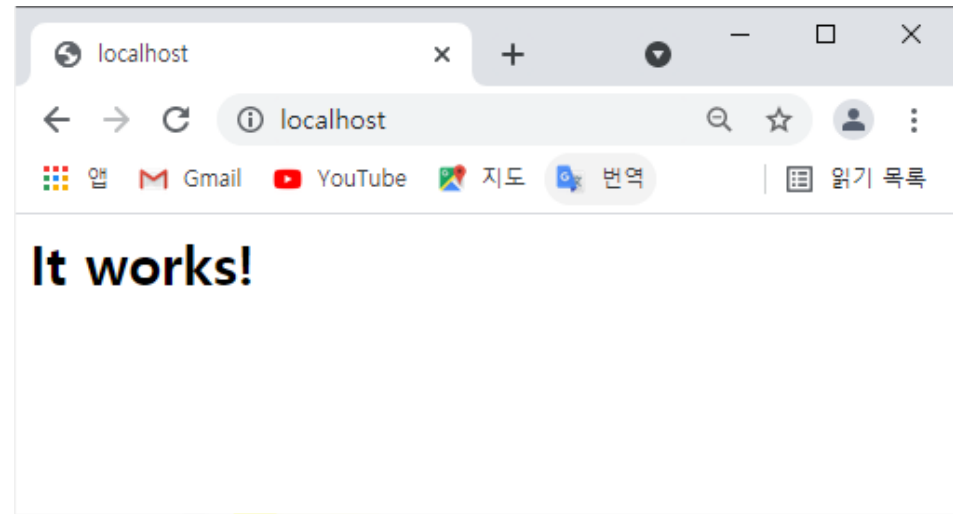
- Apache/conf/httpd.conf에서 설정한 경로에 파일(php7apache2_4.dll)이 없음



```
c:\Apache24\bin>httpd.exe -k start
httpd.exe: Syntax error on line 539 of C:/Apache24/conf/httpd.conf: Cannot load C:/
php7/php7apache2_4.dll into server: \xc1\xf6\xc1\xa4\xb5\xc8 \xb8\xf0\xb5\xe2\xc0\x
bb \xc3\xa3\xc0\xbb \xbc\xf6 \xbe\xf8\xbd\xc0\xb4\xcf\xb4\xd9.
```

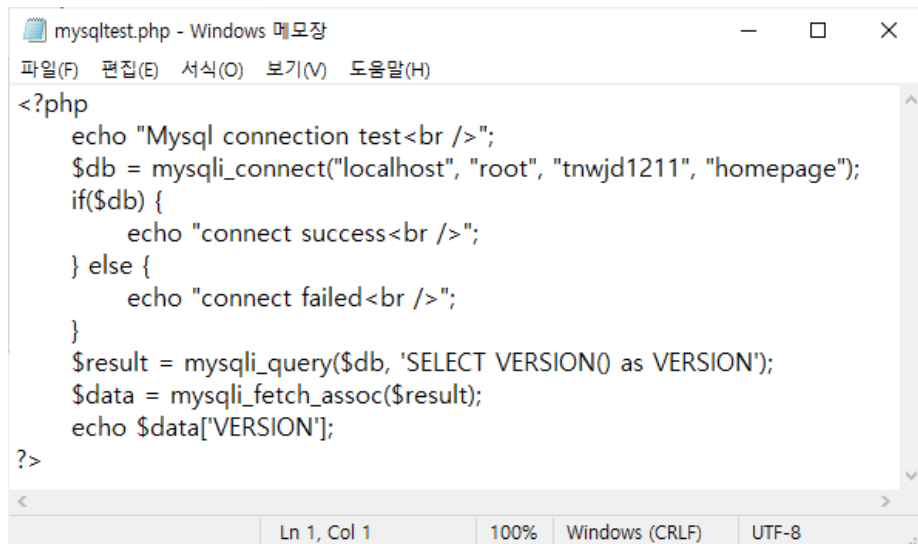
APM 재구축

- Apache & PHP 설치
 - 설정 후 정상 작동 확인

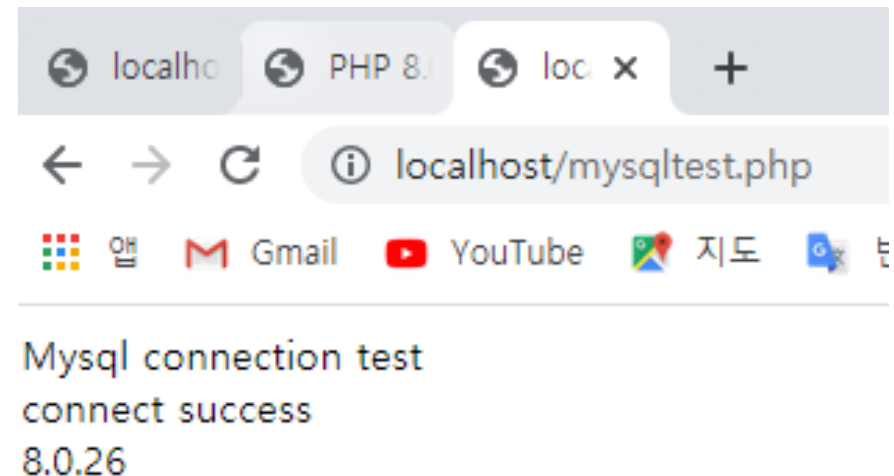


APM 재구축

- MySQL 설치
 - 설정 및 DB 생성 후 정상 작동 확인
 - APM 전체 연동 테스트

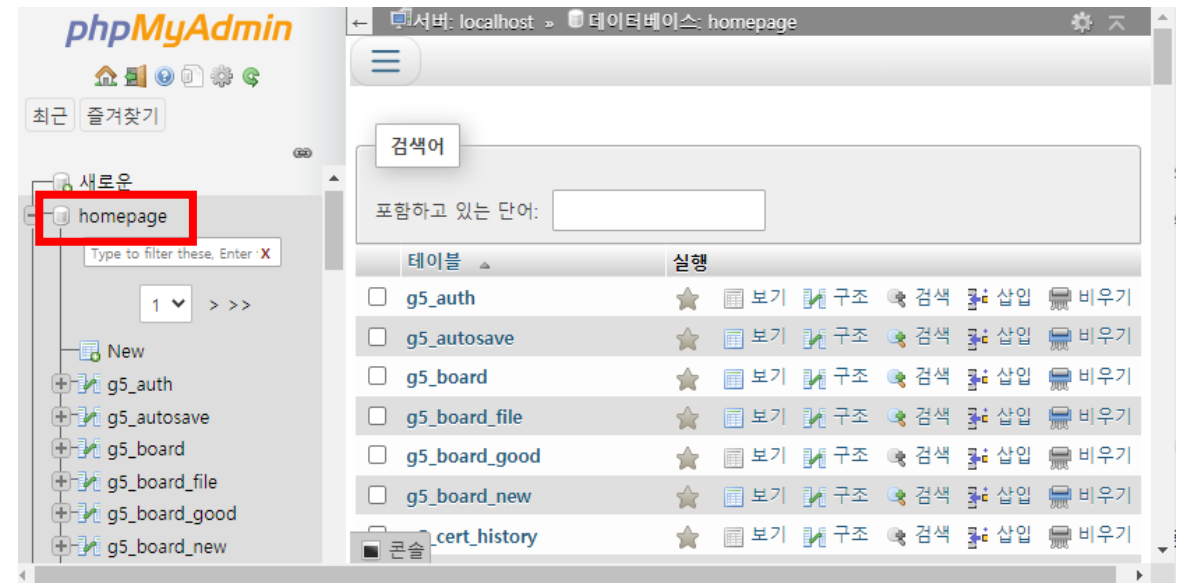
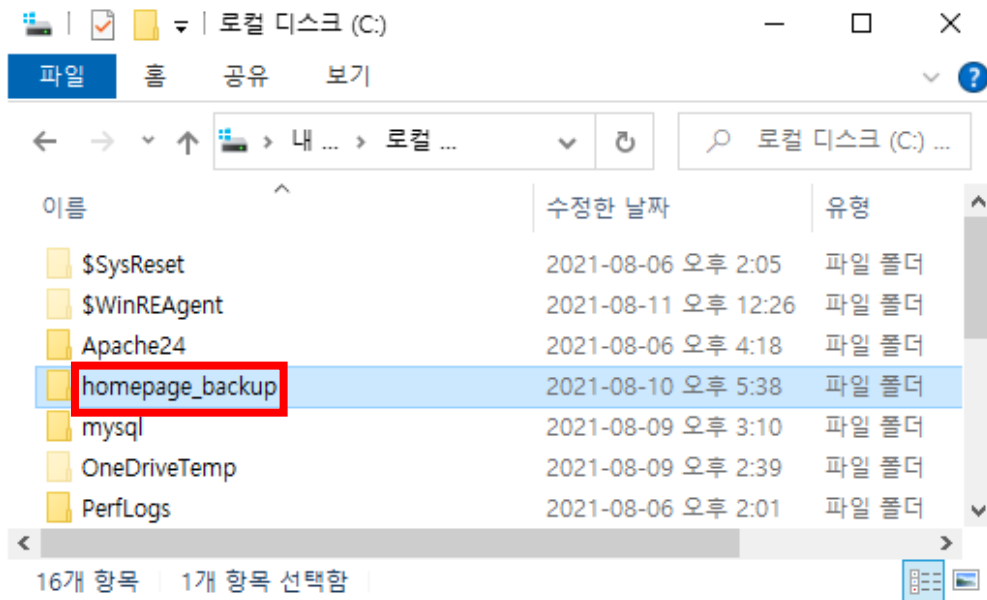


```
mysqltest.php - Windows 메모장
파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)
<?php
    echo "Mysql connection test<br />";
    $db = mysqli_connect("localhost", "root", "tnwjd1211", "homepage");
    if($db) {
        echo "connect success<br />";
    } else {
        echo "connect failed<br />";
    }
    $result = mysqli_query($db, 'SELECT VERSION() as VERSION');
    $data = mysqli_fetch_assoc($result);
    echo $data['VERSION'];
?>
```



연구실 홈페이지 환경 세팅

- 로컬 경로에 백업 파일 업로드
- MySQL Struct dump



localhost 에러 사항 (1/4)

- Fatal error: Uncaught Error: Call to undefined function get_magic_quotes_gpc() in C:\Homepage_backup\W\common.php:96 Stack trace: #0 C:\Homepage_backup\W\common.php(2): include_once() #1 C:\Homepage_backup\W\index.php(2): include_once('C:\W\Homepage_bac...') #2 {main} thrown in C:\Homepage_backup\W\common.php on line 96

➤ 발생 원인

- get_magic_quotes_gpc()는 php8에 존재하지 않는 함수

➤ 해결 방안

- !(function_exists('magic_quotes_gpc') && get_magic_quotes_gpc()) 추가

localhost 에러 사항 (2/4)

- Fatal error: Array and string offset access syntax with curly braces is no longer supported in C:\w\homepage_backup\w\lib\w\common.lib.php on line 3514

- `if($path[0]=='/')`

➤ 발생 원인

- php7.4부터 배열 인덱스 호출 시 중괄호 대신 대괄호 사용

➤ 해결 방안

- 배열 인덱스 괄호 {}를 []로 변경

localhost 에러 사항 (3/4)

- Warning: mysqli_connect(): (HY000/1045): Access denied for user 'kairos0'@'localhost' (using password: YES) in C:\homepage_backup\lib\wcommon.lib.php on line 1443
Connect Error: Access denied for user 'kairos0'@'localhost' (using password: YES)

➤ 발생 원인

- MySQL 정보가 올바르지 않아서 DB 연결 불가

➤ 해결 방안

- DB 정보 변경

localhost 에러 사항 (4/4)

- Fatal error: Uncaught Error: Undefined constant "G5_SHOP_DIR" in C:\Homepage_backup\Weyoom\Wconfig.php:44 Stack trace:
#0 C:\Homepage_backup\Weyoom\Wcommon.php(10): include_once() #1
C:\Homepage_backup\Wextend\Wxeyoom.extend.php(53): include_once('C:\W\Homepage_bac...') #2
C:\Homepage_backup\Wcommon.php(646): include_once('C:\W\5Homepage_bac...') #3
C:\Homepage_backup\Wcommon.php(2): include_once('C:\W\Homepage_bac...') #4
C:\Homepage_backup\Windex.php(2): include_once('C:\W\Homepage_bac...') #5
{main} thrown in C:\Homepage_backup\Weyoom\Wconfig.php on line 44

➤ 발생 원인

- 정확한 원인 파악 실패
- 그누보드는 php8에서 테스트되지 않았다는 의견

➤ 해결 방안

- php7으로 재설치

PHP7 재설치

- 발생한 에러
 - Apache 작동 실패
 - The requested operation has failed
 - Can't locate API module structure `php_module' in file C:/php7/php7apache2_4.dll: No error
- 해결 방안
 - php7module 설정 부분 수정

localhost 접속

- Warning: Use of undefined constant G5_SHOP_DIR - assumed 'G5_SHOP_DIR' (this will throw an Error in a future version of PHP)
in C:\homepage_backup\weyoom\class\theme.class.php on line 85

➤ 발생 원인

- PHP5.5 이상 버전부터 배열 사용 시 인용부호 적용

➤ 해결 방안

- 상수에 인용부호 적용하여 에러 사항 부분 모두 수정

localhost 접속

- Warning: count(): Parameter must be an array or an object that implements Countable in C:\homepage_backup\themeWeb4_basic\head.html.php on line 302

➤ 발생 원인

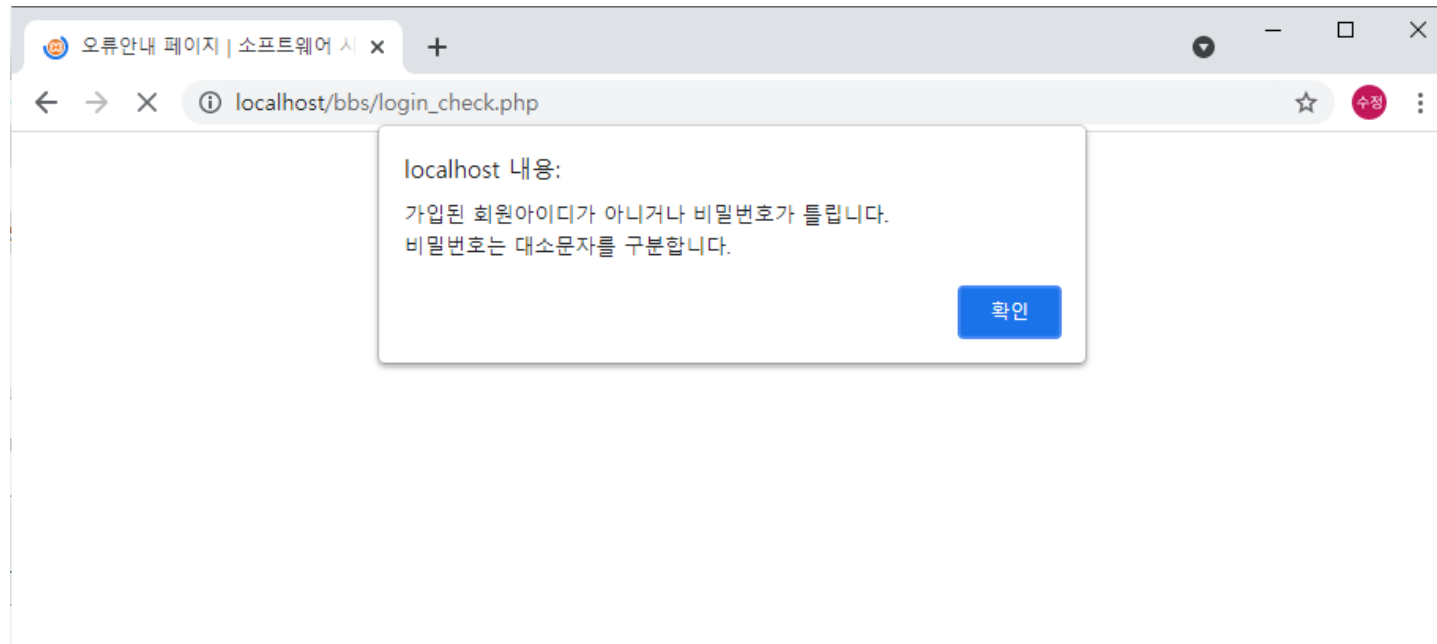
- php7 이상에서 발생하는 count() 함수와 관련된 문제

➤ 해결 방안

- php에서 count함수 사용 시 해결 소스코드 인용하여 html 코드 추가

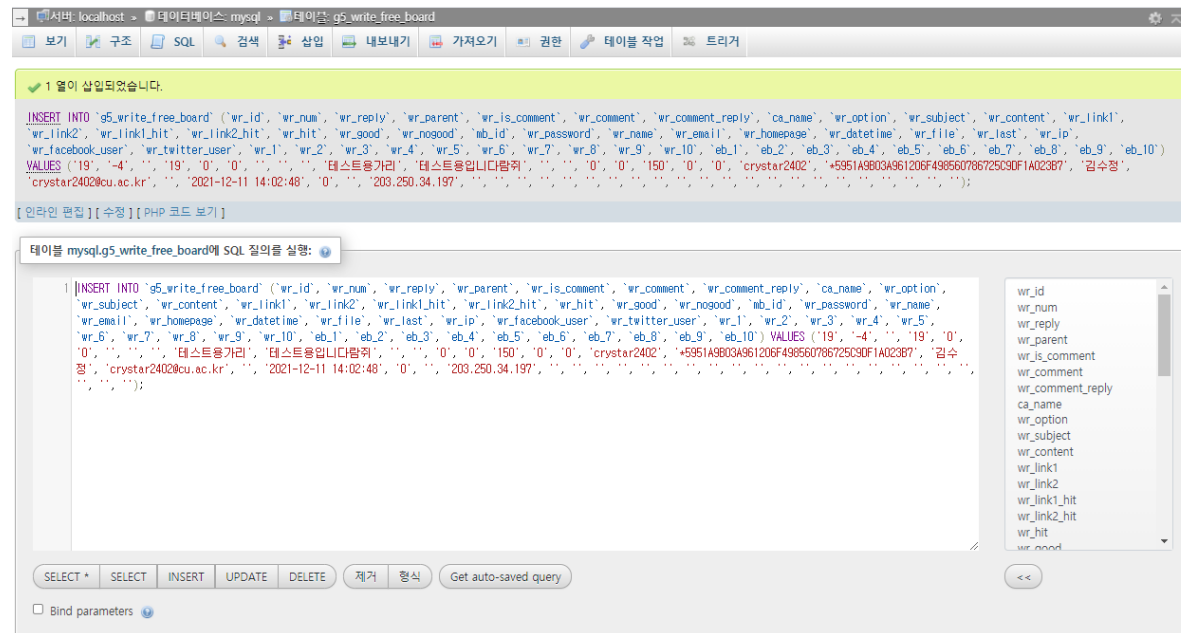
localhost 접속

- localhost로 접속한 연구실 홈페이지에서 로그인 불가



DB 상호작용 확인

- 게시물 업로드 통한 DB 상호작용 확인
 - 자유게시판 테이블(g5_write_free_board) 내 열 삽입




DB 상호작용 확인

- 게시물 업로드 통한 DB 상호작용 확인
 - 자유게시판 테이블(g5_write_free_board) 내 열 삽입

📍 자유게시판

Home / 게시판 / 자유게시판



김수정
학사과정

작성
0

내글반응
0

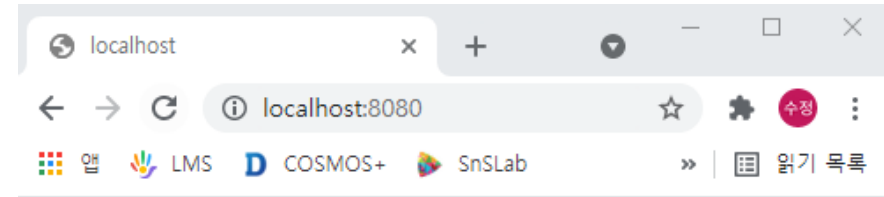
로그아웃

전체 3 건 - 1 페이지

번호	제목	글쓴이	날짜	뷰
3	● 테스트용가리	김수정	51분전	161
2	스	Ryan	2020.11.12	48
1		정인범	2020.03.24	69
0		안광은	2019.10.21	66

홈페이지 접속 에러 사항

- ERR_CONNECTION_RESET
- 예상 원인
 - ① 방화벽 인한 인터넷 연결 차단
 - ② data 디렉토리 권한 낮음
 - ③ 캐쉬 용량 초과
 - ④ 포트 연결 오류



사이트에 연결할 수 없음

연결이 재설정되었습니다.

다음 방법을 시도해 보세요.

- 연결 확인
- 프록시 및 방화벽 확인
- Windows 네트워크 진단 프로그램 실행

ERR_CONNECTION_RESET

새로고침

세부정보

홈페이지 접속 에러 사항

- 해결 시도 방안
 - VPN 설치 후 접속 (실패)
 - Data 디렉토리 권한 777 부여 (실패)
 - session.use_only_cookies을 1로 변경 (실패)
 - 인터넷 캐쉬 초기화 <ipconfig/flushdns> (실패)

홈페이지 접속 에러 사항

- 해결 시도 방안
 - 리눅스 내
 - 8080 포트 서버 닫혀 있음
 - 80 포트 리스닝
 - 윈도우 내
 - 8080 포트 발견 → 프로세스 종료

```
ksj@DESKTOP-H83D2JG:~$ nmap -p 8080 127.0.0.1

Starting Nmap 7.60 ( https://nmap.org ) at 2021-09-03 16:31 KST
Nmap scan report for localhost (127.0.0.1)
Host is up (0.000074s latency).

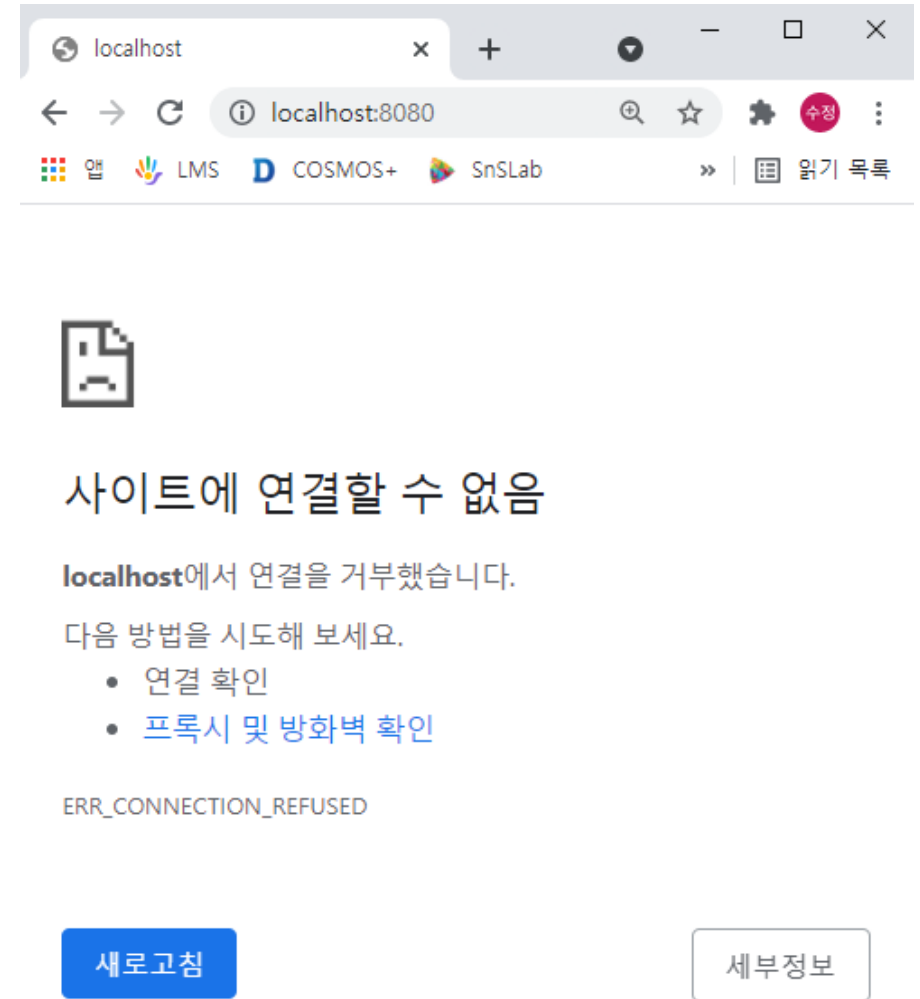
PORT      STATE SERVICE
8080/tcp   closed http-proxy

Nmap done: 1 IP address (1 host up) scanned in 0.03 seconds
```

```
root@DESKTOP-H83D2JG:/var/www/html# netstat -tnlp
Active Internet connections (only servers)
Proto Recv-Q Send-Q Local Address           Foreign Address         State       PID/Program name
tcp6      0      0 :::3306                 :::*                    LISTEN      -
tcp6      0      0 :::80                   :::*                    LISTEN      1690/apache2
```

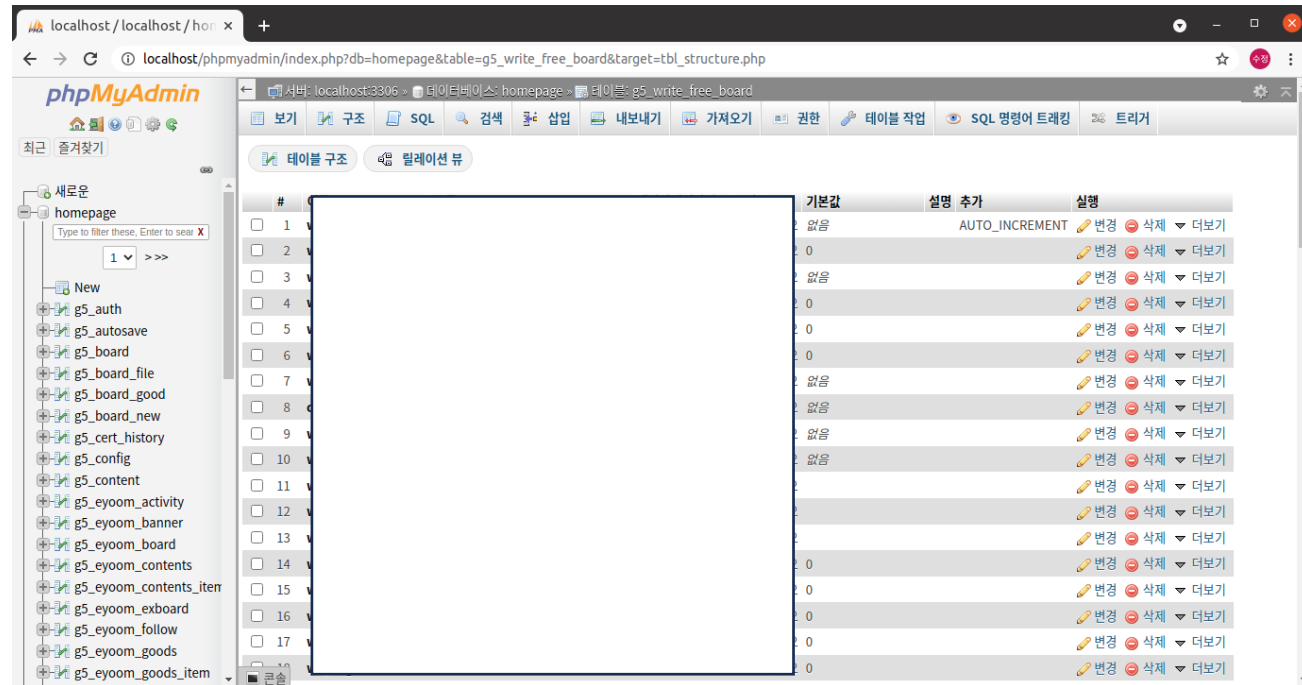
홈페이지 접속 에러 사항

- ERR_CONNECTION_REFUSED



연구실 홈페이지 구축

- APM(Apache & PHP & MySQL) 환경 구축
- 홈페이지 백업 파일 업로드 및 DB import



연구실 홈페이지 구축

- 이슈 발생
 - 문제 상황
 - 로그인 불가 & 캡차 이미지 표기 오류 문제
 - 원인
 - MySQL 버전 불일치로 인한 password() 함수 오류
 - MySQL 4.0 이하 → password() 함수 결과가 16bytes
 - MySQL 4.0 이하 → password() 함수 결과가 41bytes
- 해결 방법
 - sql_password() 함수 수정

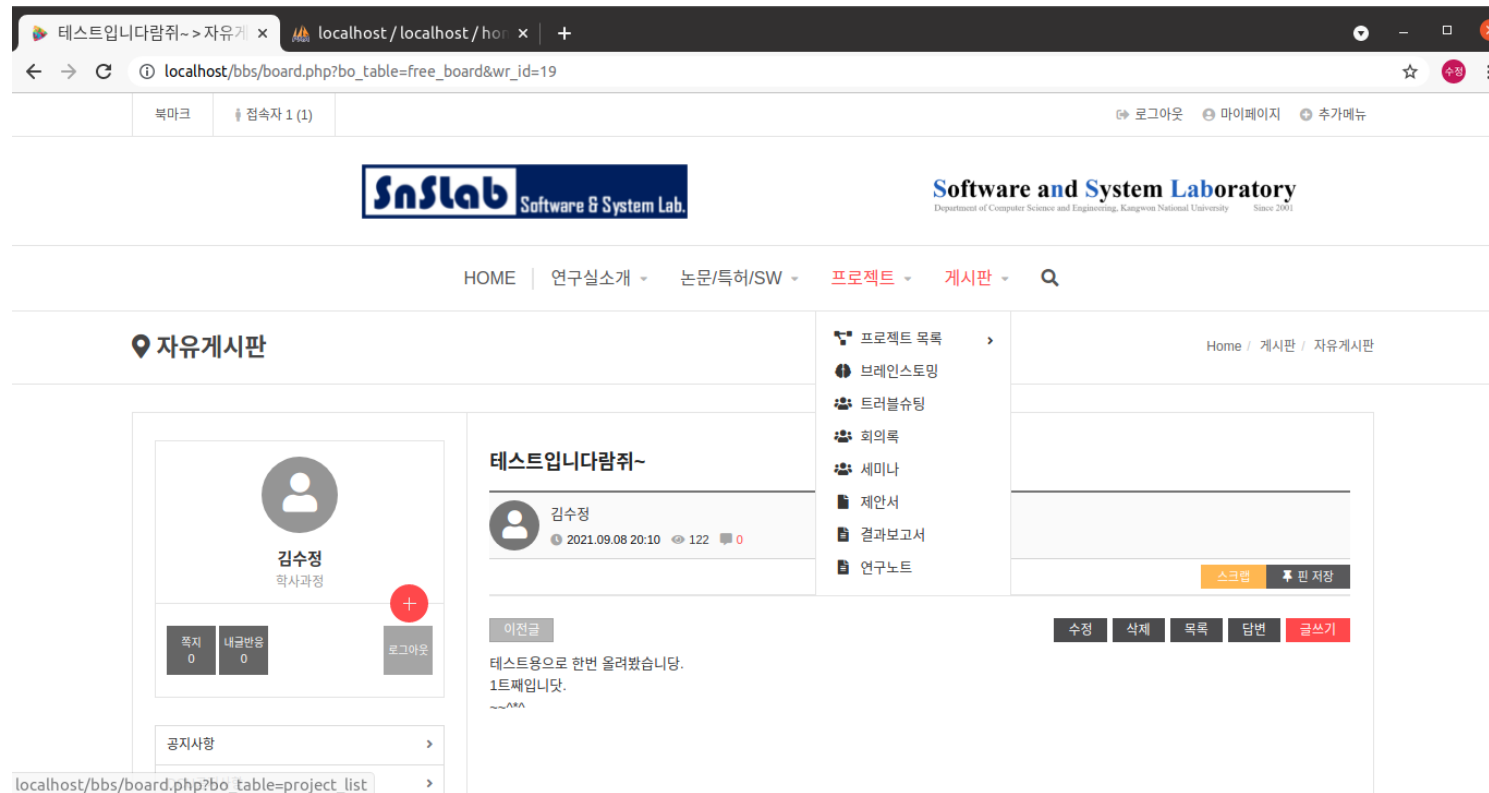
연구실 홈페이지 구축

- phpMyAdmin 이용한 게시물 업로드 확인



연구실 홈페이지 구축

- phpMyAdmin 이용한 게시물 업로드 확인

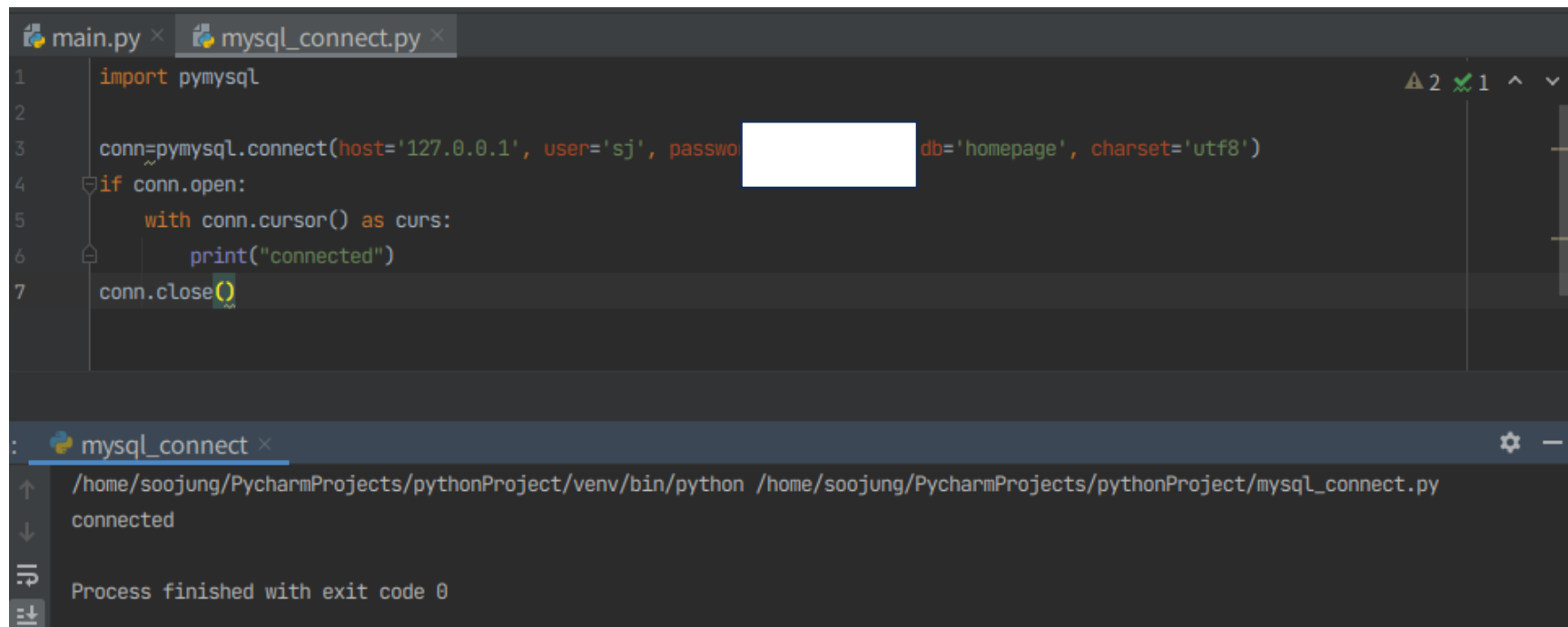


논문 정보 크롤링

- Python 웹 크롤링 환경 구축
 - Python 및 Pycharm 설치
 - 크롤링 라이브러리 설치
 - BeautifulSoup4, requests, PyMySQL

논문 정보 크롤링

- PyMySQL로 DB 연결
 - connect 함수 이용하여 DB 서버 연결

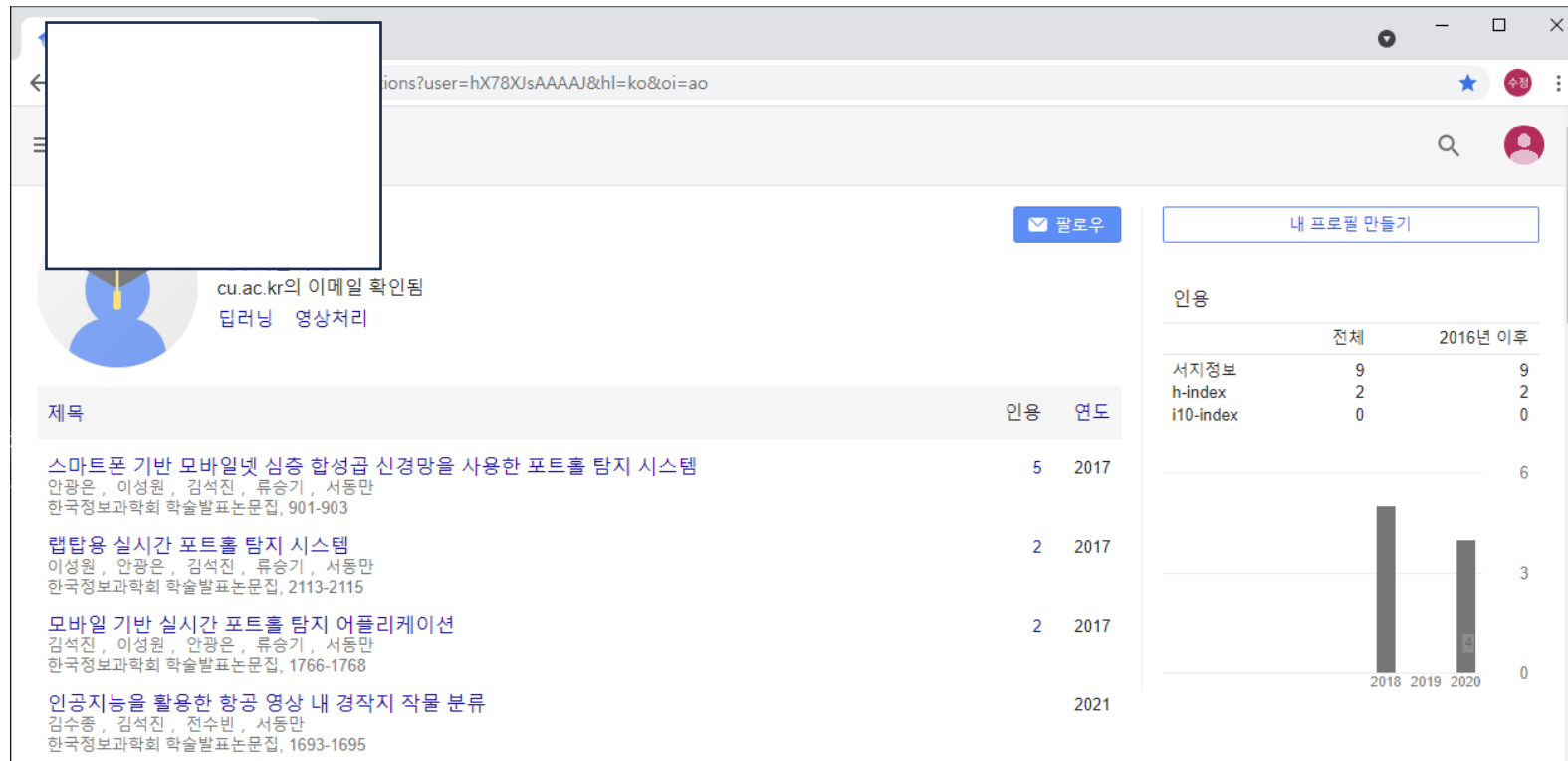


```
main.py x mysql_connect.py x
1 import pymysql
2
3 conn=pymysql.connect(host='127.0.0.1', user='sj', passwo db='homepage', charset='utf8')
4 if conn.open:
5     with conn.cursor() as curs:
6         print("connected")
7     conn.close()

mysql_connect x
/home/soojung/PycharmProjects/pythonProject/venv/bin/python /home/soojung/PycharmProjects/pythonProject/mysql_connect.py
connected
Process finished with exit code 0
```

논문 정보 크롤링

- 논문 정보 크롤링 구현



cu.ac.kr의 이메일 확인됨
딤러닝 영상처리

제목	인용	연도
스마트폰 기반 모바일넷 심층 합성곱 신경망을 사용한 포트를 탐지 시스템 안광은, 이성원, 김석진, 류승기, 서동만 한국정보과학회 학술발표논문집, 901-903	5	2017
웹탐용 실시간 포트를 탐지 시스템 이성원, 안광은, 김석진, 류승기, 서동만 한국정보과학회 학술발표논문집, 2113-2115	2	2017
모바일 기반 실시간 포트를 탐지 어플리케이션 김석진, 이성원, 안광은, 류승기, 서동만 한국정보과학회 학술발표논문집, 1766-1768	2	2017
인공지능을 활용한 항공 영상 내 경작지 작물 분류 김수중, 김석진, 전수빈, 서동만 한국정보과학회 학술발표논문집, 1693-1695		2021

필로우

내 프로필 만들기

인용

	전체	2016년 이후
서지정보	9	9
h-index	2	2
i10-index	0	0

6

3

0

2018 2019 2020

- 논문 정보 크롤링 구현

63

selenium 이용한 사이트 자동화

- selenium 사용 배경
 - 한 페이지 내에 모든 정보가 명시되지 않음
 - 더보기 버튼 클릭 필요



selenium 이용한 사이트 자동화

- selenium & google driver 설치
- 생성한 자동화 동작
 - 스크롤 내리기 동작
 - 버튼 클릭 동작

```
##### webdriver selenium 설정
##### 창 숨기는 옵션 추가
options = webdriver.ChromeOptions()
options.add_argument("headless")
# 구글 드라이버 경로 설정
driverPath = "/home/soojung/PycharmProjects/google_scholar_crawling/chromedriver"
driver = webdriver.Chrome(driverPath)
# chromedriver 실행 시 홈페이지 내 데이터 로딩 시간 기다리기 설정
driver.implicitly_wait(2)
driver.get(url=url)
# 스크롤 내리기
driver.execute_script("window.scrollTo(0, document.body.scrollHeight)")
driver.implicitly_wait(5)
# 버튼 클릭
driver.find_element_by_xpath("//*[@id='gsc_bpf_more']/span").click()
```

리스트 세분화

- 기존 리스트 정보

- title : 논문 제목

- info : 저자 및 발행 기관

전체 65 건 - 1 페이지

글쓰기

번호	제목	글쓴이	날짜	뷰
65	● 사용자 경험정보를 고려한 결정트리 기반 음악 추천 시스템	SJ	3시간 38분전	1
64	● MOSAICING INSIDE TUNNEL IMAGES FOR TUNNEL MONITORING	SJ	3시간 38분전	1
63	● 다중 서버 기반 웹 서비스에서 하트비트를 이용한 서비스 안정화 방안 연구	SJ	3시간 38분전	1
62	● 딥러닝 탐지용 임베디드 장치 간 처리속도 비교	SJ	3시간 38분전	1
61	● 항공 이미지에서의 Mask R-CNN 기반 포트홀 검출 연구	SJ	3시간 38분전	1
60	● 온라인 강의 상황에서의 프로그래밍 실습을 위한 온라인 저지 시스템 개발	SJ	3시간 38분전	1
59	● 항공영상 기반 포트홀 위치 가시화 웹 서비스 구현 웹 어플리케이션	SJ	3시간 38분전	1
58	● R-CNN 기반 포트홀 탐지 웹 서비스 시스템 구현	SJ	3시간 38분전	1
57	● 항공 영상 내 오브젝트의 정확한 GPS 좌표 예측 방법 및 정복 표현	SJ	3시간 38분전	1
56	● 비대면 발열 검사를 위한 열화상 카메라 방식의 저비용 출입관리 시스템 구현	SJ	3시간 38분전	1
55	● 탐침 기반 토양 습도센서 데이터 수집 안정화를 위한 연구	SJ	3시간 38분전	1
54	● 항공 영상 내 오브젝트의 정확한 면적 측정을 위한 외곽선 검출 방법	SJ	3시간 38분전	1
53	● 터널 내부 영상 정합 알고리즘 개선 연구	SJ	3시간 38분전	1
52	● 항공 영상 내 포트홀 탐지를 위한 합성곱 신경망 성능 비교 분석	SJ	3시간 38분전	1
51	● 잘못된 어노테이션 제거기법을 활용한 딥러닝 학습용 이미지 데이터 셋 구축 시스템...	SJ	3시간 38분전	1

리스트 세분화

• 변경 후 리스트 정보

- title : 논문 제목
- author : 저자
- publisher : 발행 기관
- year : 발행 연도

안광은,이성원,김석진,류승기,서동만
이성원,안광은,김석진,류승기,서동만
김석진,이성원,안광은,류승기,서동만
김수중,김석진,전수빈,서동만
김수정,정윤지,김석진,김미혜,서동만
이동욱,김석진,박준홍,전수빈,서동만
한윤성,전민건,김석진,전수빈,서동만
김석진,전민건,전수빈,서동만
한윤성,김수중,김석진,이동욱,전수빈,서동만

author

한국정보과학회 학술발표논문집, 901-903, 2017
한국정보과학회 학술발표논문집, 2113-2115, 2017
한국정보과학회 학술발표논문집, 1766-1768, 2017
한국정보과학회 학술발표논문집, 1693-1695, 2021
한국정보과학회 학술발표논문집, 2086-2088, 2021
한국정보과학회 학술발표논문집, 651-653, 2021
한국정보과학회 학술발표논문집, 274-276, 2021
한국정보과학회 학술발표논문집, 271-273, 2021
정보과학회 컴퓨팅의 실제 논문지 27 (6), 282-288, 2021

publisher

2017
2017
2017
2021
2021
2021
2021
2021
2021
2021
2020

year

스마트폰 기반 모바일넷 심층 합성곱 신경망을 사용한 포트홀 탐지 시스템
랩탑용 실시간 포트홀 탐지 시스템
모바일 기반 실시간 포트홀 탐지 어플리케이션
인공지능을 활용한 항공 영상 내 경작지 작물 분류
교육용 로봇 팔 기기의 춤 동작 구현
잘못된 어노테이션 제거기법을 활용한 딥러닝 학습용 이미지 데이터 셋 구축 시스템 연구

title

리스트 세분화

- 게시물 제목 설정
 - 각 리스트 원소 병합 후 저장
 - title_list : 게시물 제목 리스트

```
# 게시물 이름 생성
for i in range(list_len):
    title_list[i] = "[" + "".join(year[i]) + "]" + "".join(author[i]) + " \\" + "".join(title[i]) + "\\", " + "".join(publisher[i])
```

```
[2017] 안광은, 이성원, 김석진, 류승기, 서동만 "스마트폰 기반 모바일넷 심층 합성곱 신경망을 사용한 포트홀 탐지 시스템", 한국정보과학회 학술발표논문집, 901-903, 2017
[2017] 이성원, 안광은, 김석진, 류승기, 서동만 "랩탑용 실시간 포트홀 탐지 시스템", 한국정보과학회 학술발표논문집, 2113-2115, 2017
[2017] 김석진, 이성원, 안광은, 류승기, 서동만 "모바일 기반 실시간 포트홀 탐지 어플리케이션", 한국정보과학회 학술발표논문집, 1766-1768, 2017
[2021] 김수중, 김석진, 전수빈, 서동만 "인공지능을 활용한 항공 영상 내 경작지 작물 분류", 한국정보과학회 학술발표논문집, 1693-1695, 2021
[2021] 김수정, 정운지, 김석진, 김미혜, 서동만 "교육용 로봇 팔 기기의 춤 동작 구현", 한국정보과학회 학술발표논문집, 2086-2088, 2021
[2021] 이동욱, 김석진, 박준홍, 전수빈, 서동만 "잘못된 어노테이션 제거기법을 활용한 딥러닝 학습용 이미지 데이터 셋 구축 시스템 연구", 한국정보과학회 학술발표논문집, 651-653, 2021
[2021] 한운성, 전민건, 김석진, 전수빈, 서동만 "항공 영상 내 포트홀 탐지를 위한 합성곱 신경망 성능 비교 분석", 한국정보과학회 학술발표논문집, 274-276, 2021
```

발행 연도 순 정렬

- 발행 연도 리스트(year) 이용
- 오름차순 정렬

```
# 발행연도 순 게시글 정렬
for m in range(list_len):
    minimum = year[m]
    index_of_min = m
    for n in range(m, list_len):
        if year[n] < minimum:
            minimum = year[n]
            index_of_min = n
    year[m], year[index_of_min] = year[index_of_min], year[m]
    title_list[m], title_list[index_of_min] = title_list[index_of_min], title_list[m]
    print(title_list[m])
```

[2017] 안광은, 이성원, 김석진, 류승기, 서동만 "스마트폰 기반 모바일넷 심층 합성곱 신경망을 사용한 포트홀 탐지 시스템", 한국정보과학회 학술발표논문집, 901-903, 2017

[2017] 이성원, 안광은, 김석진, 류승기, 서동만 "랩탐용 실시간 포트홀 탐지 시스템", 한국정보과학회 학술발표논문집, 2113-2115, 2017

[2017] 김석진, 이성원, 안광은, 류승기, 서동만 "모바일 기반 실시간 포트홀 탐지 어플리케이션", 한국정보과학회 학술발표논문집, 1766-1768, 2017

[2017] 이민영, 정영주, 이성원, 한정, 김석진, 한현관, 서동만 "반려 식물을 위한 스마트 화분", 한국정보처리학회 학술대회논문집 24 (2), 1192-1194, 2017

[2018] 안광은, 이성원, 정영주, 김석진, 류승기, 서동만 "딥러닝 기반 포트홀 탐지 시스템에서의 학습 모델 비교 분석", 한국정보과학회 학술발표논문집, 731-733, 2018

[2018] 한정, 김석진, 박준홍, 이성원, 고희동, 서동만 "노인 복지 시설 CCTV 를 이용한 웹 기반 노인 행동 관리 서비스", 한국정보과학회 학술발표논문집, 1707-1709, 2018

[2018] 김석진, 안광은, 이성원, 류승기, 서동만 "모바일 기반 포트홀 탐지 시스템 효율화를 위한 영상 처리 및 딥러닝 기반 알고리즘 연구", 한국정보과학회 학술발표논문집, 1595-1597, 2018

발행 연도 순 정렬

번호	논문
32	● [2021] 김수종,김석진,전수빈,서동만 "인공지능을 활용한 항공 영상 내 경작지 작물 분류", 한국정보과학회 학술발표논문집, 1693-1695, 2021
31	● [2021] 한윤성,김수종,김석진,이동욱,전수빈,서동만 "항공 영상 내 오브젝트의 정확한 면적 측정을 위한 외곽선 검출 방법", 정보과학회 컴퓨팅의 실제 논문지 27 (6), 282-288, 2021
30	● [2021] 김석진,전민건,전수빈,서동만 "터널 내부 영상 정합 알고리즘 개선 연구", 한국정보과학회 학술발표논문집, 271-273, 2021
29	● [2021] 한윤성,전민건,김석진,전수빈,서동만 "항공 영상 내 포트홀 탐지를 위한 합성곱 신경망 성능 비교 분석", 한국정보과학회 학술발표 논문집, 274-276, 2021
28	● [2021] 이동욱,김석진,박준홍,전수빈,서동만 "잘못된 어노테이션 제거기법을 활용한 딥러닝 학습용 이미지 데이터 셋 구축 시스템 연구", 한국정보과학회 학술발표논문집, 651-653, 2021
27	● [2021] 김수정,정윤지,김석진,김미혜,서동만 "교육용 로봇 팔 기기의 춤 동작 구현", 한국정보과학회 학술발표논문집, 2086-2088, 2021
26	● [2020] 김재훈,김석진,전수빈,서동만 "항공영상 기반 포트홀 위치 가시화 웹 서비스 구현 웹 어플리케이션", 한국정보과학회 학술발표논문집, 1423-1425, 2020
25	● [2020] 전민건,한윤성,김석진,이동욱,전수빈,서동만 "R-CNN 기반 포트홀 탐지 웹 서비스 시스템 구현", 한국정보과학회 학술발표논문집, 1426-1428, 2020
24	● [2020] 김수종,한윤성,김석진,전수빈,서동만 "항공 영상 내 오브젝트의 정확한 GPS 좌표 예측 방법 및 정복 표현", 한국정보과학회 학술발표논문집, 1420-1422, 2020
23	● [2020] 이재혁,김수종,김재훈,김석진,전수빈,서동만 "비대면 발열 검사를 위한 열화상 카메라 방식의 저비용 출입관리 시스템 구현", 한국정보과학회 학술발표논문집, 1195-1197, 2020
22	● [2020] 김재훈,김석진,이동욱,전수빈,서동만 "탐침 기반 토양 습도센서 데이터 수집 안정화를 위한 연구", 한국정보과학회 학술발표논문집, 1244-1246, 2020
21	● [2020] 김유리,김성지,김정호,조재림,이동욱,김석진,전수빈,서동만 "사용자 경험정보를 고려한 결정트리 기반 음악 추천 시스템", 한국정보처리학회 학술대회논문집 27 (2), 655-658, 2020
20	● [2020] SJ, Kim, SW, Lee, D, Seo "MOSAICING INSIDE TUNNEL IMAGES FOR TUNNEL MONITORING", JP Journal of Heat and Mass Transfer, 71 - 76, 2020
19	● [2020] 김석진,박준홍,전수빈,서동만 "다중 서버 기반 웹 서비스에서 하트비트를 이용한 서비스 안정화 방안 연구", 한국정보과학회 학술발표논문집, 1171-1172, 2020
18	● [2020] 이재혁,전민건,한윤성,김석진,서동만,전수빈 "딥러닝 탐지용 임베디드 장치 간 처리속도 비교", 한국정보과학회 학술발표논문집, 1622-1624, 2020

번호	논문
17	● [2020] 전민건,한윤성,김석진,전수빈,서동만 "항공 이미지에서의 Mask R-CNN 기반 포트홀 검출 연구", 한국정보과학회 학술발표논문집, 1640-1642, 2020
16	● [2020] 박준홍,전민건,이재혁,김석진,전수빈,서동만 "온라인 강의 상황에서의 프로그래밍 실습을 위한 온라인 저지 시스템 개발", 한국정보과학회 학술발표논문집, 1667-1669, 2020
15	● [2019] 박준홍,김석진,정연욱,이동욱,정영주,서동만 "데이터셋 생성을 위한 이미지 URI 및 메타데이터 수집 크롤러", 한국정보처리학회 학술대회논문집 26 (2), 1082-1084, 2019
14	● [2019] 이재혁,김석진,정영주,한윤성,서동만 "양방향 통신 기반 침입자탐지 가정용 CCTV 시스템", 한국정보처리학회 학술대회논문집 26 (2), 319-321, 2019
13	● [2019] 김석진,이성원,전민건,김수종,서동만 "교량 구조물 안전점검 자동화를 위한 정합 이미지 생성 시스템", 한국정보처리학회 학술대회논문집 26 (2), 1105-1106, 2019
12	● [2019] 김석진,이성원,서동만 "터널 내부의 파노라마 이미지 생성을 위한 이미지 특징 추출 알고리즘 비교 분석", 한국정보과학회 학술발표논문집, 1409-1411, 2019
11	● [2019] 정연욱,김석진,유진희,박준홍,서동만 "집단지성 기반 딥러닝용 데이터 셋 구축을 위한 이미지 라벨링 게임 어플리케이션", 한국정보과학회 학술발표논문집, 1467-1469, 2019
10	● [2018] 이성원,안광은,정영주,김석진,류승기,서동만 "기계학습과 이미지 처리 기반 실시간 포트홀 탐지 시스템 제안", 한국정보과학회 학술발표논문집, 7-9, 2018
9	● [2018] 김석진,정영주,이성원,서동만 "IoT 기반 드론 확장 모듈 연구", 한국정보과학회 학술발표논문집, 1828-1830, 2018
8	● [2018] 김석진,안광은,이성원,류승기,서동만 "모바일 기반 포트홀 탐지 시스템 효율화를 위한 영상 처리 및 딥러닝 기반 알고리즘 연구", 한국정보과학회 학술발표논문집, 1595-1597, 2018
7	● [2018] 한정,김석진,박준홍,이성원,고희동,서동만 "노인 복지 시설 CCTV 를 이용한 웹 기반 노인 행동 관리 서비스", 한국정보과학회 학술발표논문집, 1707-1709, 2018
6	● [2018] 안광은,이성원,정영주,김석진,류승기,서동만 "딥러닝 기반 포트홀 탐지 시스템에서의 학습 모델 비교 분석", 한국정보과학회 학술발표논문집, 731-733, 2018
5	● [2017] 이민영,정영주,이성원,한정,김석진,한현관,서동만 "반려 식물을 위한 스마트 화분", 한국정보처리학회 학술대회논문집 24 (2), 1192-1194, 2017
4	● [2017] 김석진,이성원,안광은,류승기,서동만 "모바일 기반 실시간 포트홀 탐지 어플리케이션", 한국정보과학회 학술발표논문집, 1766-1768, 2017
3	● [2017] 이성원,안광은,김석진,류승기,서동만 "랩탑용 실시간 포트홀 탐지 시스템", 한국정보과학회 학술발표논문집, 2113-2115, 2017

발행 기관 리스트 세분화

- 기존 발행 기관 리스트
 - publisher
- 변경 후 발행 기관 리스트
 - publisher : 발행 기관
 - publication : 간행물
 - publication_page : 해당 페이지

```
한국정보과학회 학술발표논문집, 901-903, 2017
한국정보과학회 학술발표논문집, 2113-2115, 2017
한국정보과학회 학술발표논문집, 1766-1768, 2017
한국정보과학회 학술발표논문집, 1693-1695, 2021
한국정보과학회 학술발표논문집, 2086-2088, 2021
한국정보과학회 학술발표논문집, 651-653, 2021
한국정보과학회 학술발표논문집, 274-276, 2021
한국정보과학회 학술발표논문집, 271-273, 2021
정보과학회 컴퓨팅의 실제 논문지 27 (6), 282-288, 2021
```

publisher

```
list [0] : 한국정보과학회 / 학술발표논문집 / 901-903
list [1] : 한국정보과학회 / 학술발표논문집 / 2113-2115
list [2] : 한국정보과학회 / 학술발표논문집 / 1766-1768
list [3] : 한국정보과학회 / 학술발표논문집 / 1693-1695
list [4] : 한국정보과학회 / 학술발표논문집 / 2086-2088
list [5] : 한국정보과학회 / 학술발표논문집 / 651-653
list [6] : 한국정보과학회 / 학술발표논문집 / 274-276
list [7] : 한국정보과학회 / 학술발표논문집 / 271-273
list [8] : 정보과학회 / 컴퓨팅의 실제 논문지 27 (6) / 282-288
list [9] : 한국정보과학회 / 학술발표논문집 / 1244-1246
list [10] : 한국정보과학회 / 학술발표논문집 / 1195-1197
list [11] : 한국정보과학회 / 학술발표논문집 / 1420-1422
list [12] : 한국정보과학회 / 학술발표논문집 / 1426-1428
list [13] : 한국정보과학회 / 학술발표논문집 / 1423-1425
list [14] : 한국정보과학회 / 학술발표논문집 / 1667-1669
list [15] : 한국정보과학회 / 학술발표논문집 / 1640-1642
list [16] : 한국정보과학회 / 학술발표논문집 / 1622-1624
list [17] : 한국정보과학회 / 학술발표논문집 / 1171-1172
list [18] : JP / Journal of Heat and Mass Transfer / 71 - 76
list [19] : 한국정보처리학회 / 학술대회논문집 27 (2) / 655-658
list [20] : 한국정보과학회 / 학술발표논문집 / 1467-1469
list [21] : 한국정보과학회 / 학술발표논문집 / 1409-1411
list [22] : 한국정보처리학회 / 학술대회논문집 26 (2) / 1105-1106
list [23] : 한국정보처리학회 / 학술대회논문집 26 (2) / 319-321
list [24] : 한국정보처리학회 / 학술대회논문집 26 (2) / 1082-1084
list [25] : 한국정보과학회 / 학술발표논문집 / 731-733
list [26] : 한국정보과학회 / 학술발표논문집 / 1707-1709
list [27] : 한국정보과학회 / 학술발표논문집 / 1595-1597
list [28] : 한국정보과학회 / 학술발표논문집 / 1828-1830
list [29] : 한국정보과학회 / 학술발표논문집 / 7-9
list [30] : 한국정보처리학회 / 학술대회논문집 24 (2) / 1192-1194
```

publisher / publication
/ publication_page

DB 연도 칼럼 추가

- 논문 게시판(g5_write_thesis) 내 연도 칼럼 (g5_write_thesis) 추가

번호	논문
32	● 이민영,정영주,이성원,한정,김석진,한현관,서동만 "반려 식물을 위한 스마트 화분", 한국정보처리학회 학술대회는문집 24 (2), 1192-1194, 2017
31	● 이성원,안광은,정영주,김석진,류승기,서동만 "기계학습과 이미지 처리 기반 실시간 포트를 탐지 시스템 제안", 한국정보과학회 학술발표논문집, 7-9, 2018
30	● 김석진,정영주,이성원,서동만 "IoT 기반 드론 확장 모듈 연구", 한국정보과학회 학술발표논문집, 1828-1830, 2018
29	● 김석진,안광은,이성원,류승기,서동만 "모바일 기반 포트를 탐지 시스템 효율화를 위한 영상 처리 및 딥러닝 기반 알고리즘 연구", 한국정보과학회 학술발표논문집, 1595-1597, 2018
28	● 한정,김석진,박준홍,이성원,고희동,서동만 "노인 복지 시설 CCTV 를 이용한 웹 기반 노인 행동 관리 서비스", 한국정보과학회 학술발표논문집, 1707-1709, 2018
27	● 안광은,이성원,정영주,김석진,류승기,서동만 "딥러닝 기반 포트를 탐지 시스템에서의 학습 모델 비교 분석", 한국정보과학회 학술발표논문집, 731-733, 2018
26	● 박준홍,김석진,정영주,이동욱,정영주,서동만 "데이터셋 생성을 위한 이미지 URI 및 메타데이터 수집 크롤러", 한국정보처리학회 학술대회는문집 26 (2), 1082-1084, 2019
25	● 이재혁,김석진,정영주,한윤성,서동만 "양방향 통신 기반 침입탐지 가정용 CCTV 시스템", 한국정보처리학회 학술대회는문집 26 (2), 319-321, 2019
24	● 김석진,이성원,전민간,김수중,서동만 "교량 구조물 안전점검 자동화를 위한 정합 이미지 생성 시스템", 한국정보처리학회 학술대회는문집 26 (2), 1105-1106, 2019
23	● 김석진,이성원,서동만 "터널 내부의 파노라마 이미지 생성을 위한 이미지 특징 추출 알고리즘 비교 분석", 한국정보과학회 학술발표논문집, 1409-1411, 2019
22	● 정영주,김석진,유진희,박준홍,서동만 "집단지성 기반 딥러닝용 데이터 셋 구축을 위한 이미지 라벨링 게임 어플리케이션", 한국정보과학회 학술발표논문집, 1467-1469, 2019
21	● 김유리,김성지,김정호,조재림,이동욱,김석진,전수빈,서동만 "사용자 경험정보를 고려한 결정트리 기반 음악 추천 시스템", 한국정보처리학회 학술대회는문집 27 (2), 655-658, 2020

Server: localhost:3306 > Database: homepage > Table: g5_write_thesis

Options	wr_id	wr_subject	wr_content	wr_publication_year
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	69	안광은,이성원,김석진,류승기,서동만 "스마트폰 기반 모바일넷 심층 학습곱 신경망을 사용한 ...	안광은,이성원,김석진,류승기,서동만 "스마트폰 기반 모바일넷 심층 학습곱 신경망을 사용한 ...	2017
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	70	안광은,이성원,김석진,류승기,서동만 "스마트폰 기반 모바일넷 심층 학습곱 신경망을 사용한 ...	안광은,이성원,김석진,류승기,서동만 "스마트폰 기반 모바일넷 심층 학습곱 신경망을 사용한 ...	2017
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	71	이성원,안광은,김석진,류승기,서동만 "랩탐용 실시간 포트를 탐지 시스템", 한국정보과학회 ...	이성원,안광은,김석진,류승기,서동만 "랩탐용 실시간 포트를 탐지 시스템", 한국정보과학회 ...	2017
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	72	김석진,이성원,안광은,류승기,서동만 "모바일 기반 실시간 포트를 탐지 어플리케이션", 한국...	김석진,이성원,안광은,류승기,서동만 "모바일 기반 실시간 포트를 탐지 어플리케이션", 한국...	2017
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	73	김수중,김석진,전수빈,서동만 "인공지능을 활용한 항공 영상 내 경작지 작물 분류", 한국정...	김수중,김석진,전수빈,서동만 "인공지능을 활용한 항공 영상 내 경작지 작물 분류", 한국정...	2021
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	74	김수정,정윤지,김석진,김미혜,서동만 "교육용 로봇 팔 기기의 춤 동작 구현", 한국정보과학...	김수정,정윤지,김석진,김미혜,서동만 "교육용 로봇 팔 기기의 춤 동작 구현", 한국정보과학...	2021
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	75	이동욱,김석진,박준홍,전수빈,서동만 "잘못된 어노테이션 제거기법을 활용한 딥러닝 학습용 이...	이동욱,김석진,박준홍,전수빈,서동만 "잘못된 어노테이션 제거기법을 활용한 딥러닝 학습용 이...	2021
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	76	한윤성,전민간,김석진,전수빈,서동만 "항공 영상 내 포트를 탐지를 위한 합성곱 신경망 성능...	한윤성,전민간,김석진,전수빈,서동만 "항공 영상 내 포트를 탐지를 위한 합성곱 신경망 성능...	2021

논문 개조식 작성

- <웹 크롤링 기반 논문 자동 등록 시스템>
 1. 서론
 2. 관련 연구
 3. 시스템 설계
 4. 구현
 5. 결과
 6. 결론 및 향후계획

논문 개조식 작성 (1/4)

1. 서론

- 연구실 홈페이지 내 논문 게시판
- 논문 자동 등록 시스템의 필요성
- 크롤링 구현 목표

2. 관련 연구

- 웹 크롤링이란?
- 웹 크롤링 구현 시스템 관련 연구

논문 개조식 작성 (2/4)

3. 시스템 설계

- 데이터 수집
 - 목표
 - 경로
- 개발 환경
 - Python 크롤러 모듈
 - pymysql / bs4 / request / selenium

논문 개조식 작성 (3/4)

4. 구현

- 사이트 자동화
 - 생성한 자동화 동작 (스크롤 내리기 / 버튼 클릭)
- 데이터 추출
 - 논문 명, 저자, 발행기관, 간행물, 페이지, 연도 등
- DB 저장
 - 논문 게시판 테이블 내 데이터 저장

논문 개조식 작성 (4/4)

5. 결과

- 논문 게시판 내 게시물 자동 등록

6. 결론 및 향후 계획

- python 기반 웹 크롤러 구현
- 연도, 논문 명, 저널 명 순 정렬 기능 추가할 계획
- 각 인원 프로필 내 논문 등록 및 정렬 기능 추가할 계획