

검색엔진 구축 프로젝트

1. 연구배경 및 목적

2. 연구자 데이터베이스 검색시스템의 제한점

- 2-1. 키워드 검색 시 정확한 일치 (exact match) 결과 반환
- 2-2. 복합명사 토큰화 기능의 부재
- 2-3. 학술용어의 동의어와 약어 인식 불가
- 2-4. 데이터베이스 내 추상적인 연구정보

3. 검색시스템의 개선

4. 검색시스템의 확장

- 4-1. 유사도 검색 (semantic search)
 - 연구상세정보의 추가
 - 연구정보 추가 과정
 - 유사도 검색
- 4-2. 연구자 관계도

1. 연구배경 및 목적

본교의 가치 변화

- ❖ 융복합 학문 신설
 - ❖ 2019년: 스마트시티학과, 도시빅데이터융합학과
 - ❖ 2021년: 융합응용화학과, 인공지능학과
- ❖ 단일 학문의 성장보다는 다양한 학문 연계를 강조

산학협력단

- ❖ (설립이념) 산관학 연계협력 활동 증진
- ❖ (목표) 융복합 학문의 연구력 강화
- ❖ 연구자데이터베이스 구축 [그림 1]
 - 교내 연구자 및 연구현황 파악

연구자 DB

홈 > 연구자 DB

통합검색

통합검색

검색

☐ 반도체/디스플레이 ☐ 사이버보안 ☐ 수소 ☐ 양자 ☐ 이차전지 ☐ 인공지능 ☐ 차세대통신 ☐ 첨단로봇/제조

☐ 첨단모빌리티 ☐ 첨단바이오 ☐ 기타

전체 226 건 현재 페이지 1/20

학과명	교수명	기술분류	연구영역	전화번호	이메일
건축학부(건축공학)	강동화	기타	건축환경설비(실내공기환경, 공기조화설비)	02-6490-2766	dhkang@uos.ac.kr
도시공학과	강명구	기타	도시 및 지역경제, 도시정보시스템	02-6490-2799	mk@uos.ac.kr
전자전기컴퓨터공학부	강상혁	차세대통신	컴퓨터 네트워크, 유비쿼터스 센서 네트워크	02-6490-2333	shkang@uos.ac.kr
기계정보공학과	강인혜	인공지능	정형 기법, 소프트웨어 공학, 인공 지능	02-6490-2385	inhye@uos.ac.kr
환경원예학과	곽유리나	기타	채소재배생리, 시설원예	02-6490-2688	yurina911@uos.ac.kr
건축학부(건축공학)	구교진	기타	건설관리 및 전산	02-6490-2760	kook@uos.ac.kr

[그림 1] 서울시립대학교 산학협력단 내 연구자 [데이터베이스](#)

1. 연구배경 및 목적

통합검색

지능형

검색

☐ 반도체/디스플레이

☐ 사이버보안

☐ 수소

☐ 열차

☐ 이차전지

☐ 인공지능

☐ 차세대통신

☐ 첨단로봇/제조

☐ 첨단모빌리티

☐ 첨단바이오

☐ 기타

전체 6 건 현재 페이지 1/20

1. 검색 키워드와 일치하는 연구자 리스트 출력

	전화번호	이메일
건축학부(건축공학) 강동화 기타 건축환경설비(실내공기환경, 공기조화설비)	02-6490-2766	dhkang@uos.ac.kr
기계정보공학과 김태현 기타 임베디드 시스템, On-device AI, 지능형 IoT 시스템	02-6490-2388	thkim@uos.ac.kr
전자전기컴퓨터공학부 박경훈 장단로봇/제조 제어이론 및 응용 로봇학	02-6490-2322	gyunghoon.park@uos.ac.kr
교통공학과 이동민 장단모빌리티 교통시설, 교통안전, 인간공학, 지능형교통체계	02-6490-2827	dmlee@uos.ac.kr
건축학부(건축공학) 이상훈 기타 지능형 빌딩 정보 모델링 및 유지관리 (BIM) 빌딩 정보 네트워크 및 표준화된 정보 교환, 증강현실/가상현실 (AR/VR), 디지털 트윈	02-6490-2767	sanghoon.lee@uos.ac.kr
전자전기컴퓨터공학부 장석진 차세대통신 이동통신	02-6490-2342	schang213@uos.ac.kr

강동화

건축학부(건축공학)

연구실

건축환경설비연구실

홈페이지

-

연구영역

건축환경설비(실내공기환경, 공기조화설비)

이메일

dhkang@uos.ac.kr

전화번호

02-6490-2766

R&D 과제

No.	연구분야	직물분야	과제명(기간)	지원기관
1	건축환경/설비기술	건설업	건물 실내외 먼지부하 기반 미세먼지 통합제어알고리즘 개발()	교육부
2	건축환경/설비기술	건설	고위험 다중이용시설의 감염위험도 및 에너지소비 저감을 위한 지능형 실내환경제어시스템 개발()	과학기술정보통신부

2. 특정 연구자의 과제/논문/특허 내 필드값과 정확한 매치결과 반환

논문

No.	논문명	기준년도
1	Estimation of Outdoor PM2.5 Infiltration into Multifamily Homes Depending on Building Characteristics Using Regression Models	2021

[그림 2] 산학협력단 연구자 데이터베이스의 검색 과정

연구자 데이터베이스 현황

- ❖ 검색 과정: [그림 2]
- ❖ 검색 방법: 키워드 검색 [그림 3]

지능형

연구자	논문	과제	특허
연번	연번	연번	연번
기술분류 (12대)	단과대학	단과대학	단과대학
기술분류 (기존)	학과명	학과명	학과명
단과대학	교수명	교수명	교수명
학과명	논문명	과제명	특허명
교수명	기준년도	기준년도	국가코드
연구실	주관부처	주관부처	등록번호
연구영역	연구분야	연구분야	
전화번호	적용분야	적용분야	
이메일			
홈페이지			
비고			

[그림 3] 연구자 데이터베이스 테이블과 필드 상세. 실제 검색에 사용되는 필드는 밑줄로 표기함.

1. 연구배경 및 목적

현 검색시스템의 문제점

- 1 [검색 기능의 한계]
 - ❖ 키워드 검색 시 정확한 일치 (exact match) 결과 반환
 - ❖ 복합명사 토큰화 기능의 부재
 - ❖ 학술용어의 동의어와 약어 인식 불가
 - ❖ 데이터베이스 내 추상적인 연구정보

- 2 [검색 방법의 한계]
 - ❖ 키워드 검색 방식

기대효과

- ❖ 교외로부터의 학내 접근성 향상
- ❖ 산학 협력 기회 증대

목표: 검색 시스템의 유연화

- 1 [검색 기능의 개선]
 - ❖ 불용어 제거
 - ❖ 복합명사 토큰화
 - ❖ 사용자 사전 도입 (약어/동의어 인식)

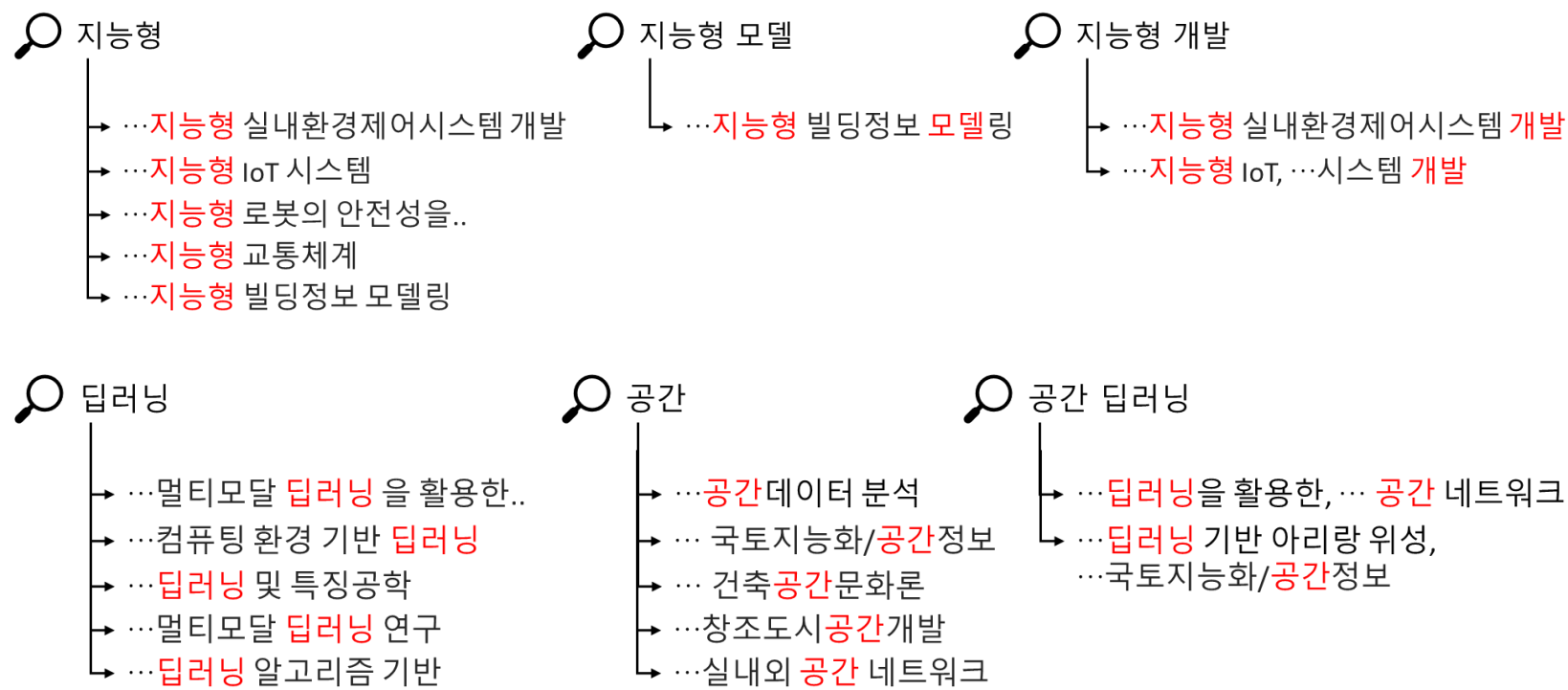
- 2 [검색 방법의 개선]
 - ❖ 연구세부정보 추가
 - ❖ 유사도 검색 방식

- 3 [추가 기능]
 - ❖ 연구자 관계도 시각화

2. 연구자 데이터베이스 검색시스템의 제한점

2-1. 키워드 검색 시 정확한 일치 (exact match) 결과 반환

- ❖ (현 상황) 검색어를 띄어쓰기 기반으로 분해, 분해된 모든 키워드를 정확히 포함한 결과를 반환
- ❖ (문제점) 여러 개의 검색 키워드를 입력 시, 모든 키워드를 가진 데이터가 없다면 검색 결과가 나오지 않음 [그림 4]
- ❖ (개선 필요성) 검색의 유연성이 떨어짐



[그림 4] 검색 키워드와 정확히 일치하는 결과 예시
(위) '지능형', '지능형 모델', '지능형 개발'의 키워드 검색 결과
(아래) '딥러닝', '공간', '딥러닝 공간'의 키워드 검색 결과

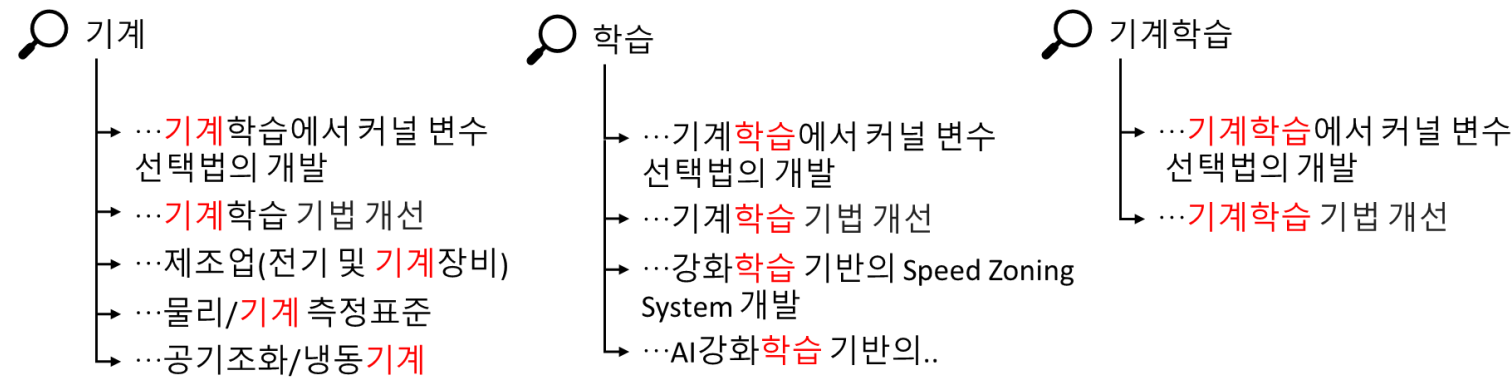
2. 연구자 데이터베이스 검색시스템의 제한점

2-2. 복합명사 토큰화 기능의 부재

- ❖ (현 상황) 연구분야에서 사용하는 용어들은 여러 개의 명사가 결합되어 만들어진 단어가 많음 [표 1]
- ❖ (문제점) 토큰화하지 못하고 하나의 단일 키워드로 인식함
 - '기계학습'은 '기계'나 '학습'이라는 개별 키워드로 인식 불가 & '기계학습'을 정확하게 포함하는 데이터만 검색 [그림 5]
- ❖ (개선 필요성) 복합명사를 이루는 단어들도 복합명사가 갖는 의미의 일부를 내포하기에 토큰화를 통해 개개의 단어를 검색 키워드로 인식하는 것이 필요함

기계정보공학과	기계+정보+공학과
도시빅데이터융합학과	도시+빅+데이터+융합+학과
응집물질물리이론	응집+물질+물리+이론
채소재배생리	채소+재배+생리
창조도시공간개발	창조+도시+공간+개발
기계학습	기계+학습

[표 1] 연구자 데이터베이스 내 복합명사의 예시



[그림 5] 복합명사 토큰화 기능의 부재 예시: '기계학습'

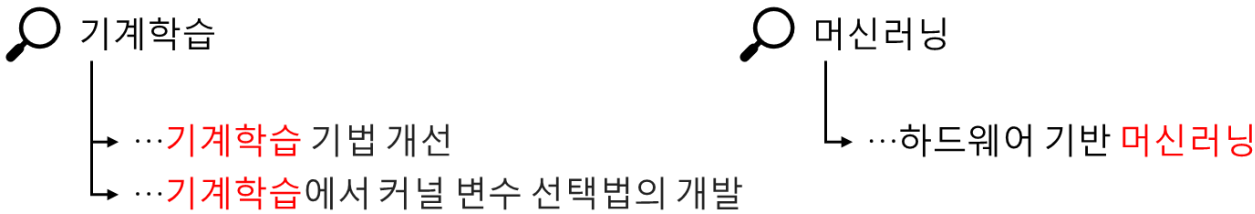
2. 연구자 데이터베이스 검색시스템의 제한점

2-3. 학술용어의 동의어와 약어 인식 불가

- ❖ (현 상황) 데이터의 대다수는 학술용어. 영어 기반의 학술용어가 한국어로 번역되는 과정에서 많은 동의어가 발생함 [표 2]
- ❖ (문제점) 이러한 동의어와 약어를 동일한 단어로 인식하지 못함 [그림 6]
- ❖ (개선 필요성) 연구 데이터 검색을 위한 기능적 향상이 필요함

동의어	
기계학습, 머신러닝	강화학습, 보상학습, 보상기반학습
AI, 인공지능	의사결정나무, 의사결정트리, 트리모델
DNN, 딥러닝, 심층 학습, 심층 신경망, 딥네트워크	클러스터링, 군집화, 군집분석, 클러스터분석
ANN, 인공신경망, 뉴럴네트워크	Computer Vision, CV, 컴퓨터비전
NLP, 자연어처리, 언어처리	과학기술정보통신부, 과기부, 과기정통부, 정통부, MSIT, Misistry of Science and ICT

[표 2] 제한점 3: 학술용어의 동의어 예시



[그림 6] 제한점 3: '기계학습'과 '머신러닝'은 동의어지만 개별 단어로 인식됨

2. 연구자 데이터베이스 검색시스템의 제한점

2-4. 데이터베이스 내 추상적인 연구정보

- ❖ (현 상황) 연구 분야를 검색하는데 실질적으로 필요한 정보는 연구영역, 논문명, 과제명, 특허명 뿐임 [표 3]
- ❖ (문제점) 추상적인 정보임
- ❖ (개선 필요성) 각 연구 산출물마다 구체적인 정보로 데이터를 구축하는 것이 바람직함

연구영역	데이터베이스, 정보검색, 자연어 처리
논문명	- Improving Complex Scene Generation by Enhancing Multi-scale Representations of GAN Discriminators - Retrieval-Augmented Response Generation for Knowledge-Grounded Conversation in the Wild - Selectively Connected Self-Attentions for Semantic Role Labeling - Exploiting Text Matching Techniques for Knowledge-Grounded Conversation
과제명	외부 지식베이스 전이를 통한 멀티모달 딥러닝 연구
특허명	-

[표 3] 제한점 4: 통계학과 박재휘 교수의 연구정보

3. 검색시스템의 개선

- ❖ 사용 프로그램: 엘라스틱서치
 - 검색엔진이자 데이터베이스, 형태소분석 인덱스 지원
- ❖ 시스템 구축
 - 작업환경: Window 10



❖ 데모 화면

The screenshot shows a web application titled '연구자 검색시스템 demo'. It has a menu with '키워드 검색' selected. A dropdown menu shows '키워드 검색' and '유사도 검색'. The '키워드 검색' section shows a search bar with '기계학습' and a '검색' button. Below it is a table with 10 rows of search results. The '유사도 검색' section shows a search bar with '기계학습' and a '검색' button. Below it is a table with 10 rows of search results.

Score	Document
13.8777	기계학습 및 지식처리
13.8777	컴퓨터비전, 기계학습
12.949	의료 빅데이터, 통계적 기계학습
12.5298	기계학습에서 커널 변수 선택법의 개발
12.1368	데이터마이닝, 데이터베이스, 정보추출, 기계학습
12.1368	공간 자료의 특성을 반영한 기계학습 기법 개선
11.7678	고에너지 물리에서 유온 트래거링을 위한 하드웨어 기반 머신러닝
11.7678	기계학습 및 앙상블 기법을 이용하여 데이터를 분류하는 장치 및 방법
5.7885	딥러닝 모델 학습 방법 및 학습기
5.3963	그래프 임베딩을 위한 적응적 자기지도학습 프레임워크

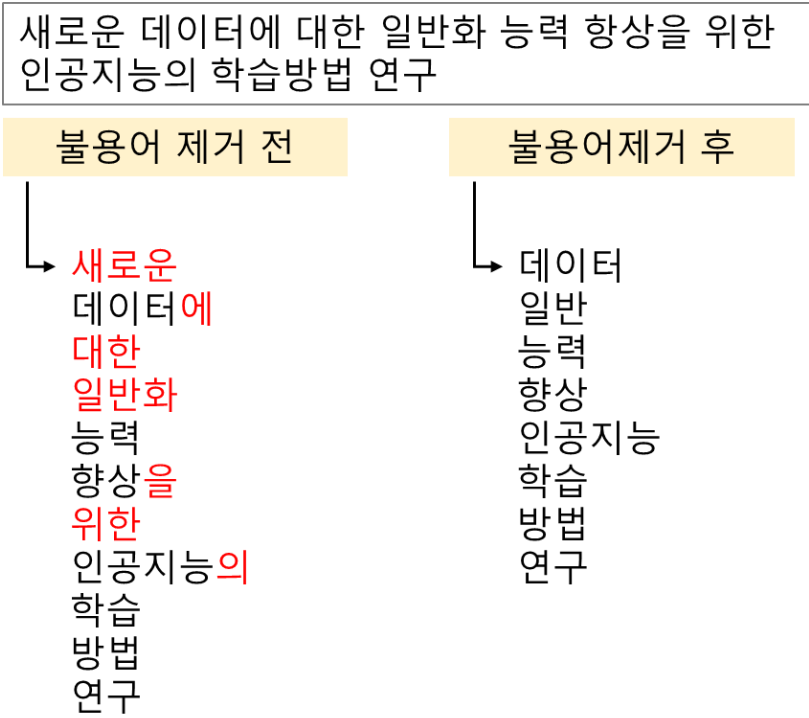
Row	Score	연번	학과명	교수명	논문명	
0	1	0.7407	51	인공지능학과	김한준	A Tensor Space Model-Based Deep Neural Network for
1	2	0.7336	733	통계학과	박창이	Multiclass Laplacian support vector machine with func
2	3	0.7267	733	통계학과	박창이	Negative binomial graphical model with excess zero
3	4	0.7201	53	인공지능학과	유하진	Capturing Discriminative Information Using a Deep Arc
4	5	0.7195	733	통계학과	박창이	Variable selection for nonlinear support vector machin
5	6	0.7159	61	인공지능학과	최호식	Primal path algorithm for compositional data analysis
6	7	0.7112	58	인공지능학과	이병정	Utilizing Topic-Based Similar Commit Information and
7	8	0.7073	733	통계학과	박창이	Kernel variable selection for multicategory support vec
8	9	0.7011	56	인공지능학과	유하진	Regularized Within-Class Precision Matrix Based PLDA
9	10	0.7002	324	공간정보공학과	정형섭	Ship Detection from X-Band SAR Images Using M2Det C

- ❖ 개선작업 후 더 유연한 검색 결과를 제공하지만 검색 결과의 양이 상당히 증가하는 문제가 발생
 - 상위 n개의 검색 결과만을 출력하도록 제한함

3. 검색시스템의 개선

❖ 불용어 제거

- ❖ 불용어: 조사, 접속사 등과 같이 문장을 구성하는 필수 요소지만 문맥적으로 큰 의미가 없는 단어
- ❖ (개선 이유) 불필요한 정보가 늘어나거나 중요한 단어로 인지할 수 있음
- ❖ (개선 후) '대한'의 검색 X & '데이터'에 대한 검색이 늘어남 [그림 8]



[그림 7] 불용어 제거: 불용어 제거 후 인덱스의 변화. 조사, 수식어 등의 제거되었다.

통합검색

데이터에 대한

산단 검색 시스템

☐ 반도체/디스플레이 ☐ 사이버보안 ☐ 수소 ☐ 양자 ☐ 이차전지 ☐ 인공지능 ☐ 차세대통신 ☐ 첨단로봇/제조

☐ 첨단모빌리티 ☐ 첨단바이오 ☐ 기타

전체 2 건 현재 페이지 1/20

학과명	교수명	기술분류	연구영역	전화번호	이메일
전자전기컴퓨터공학부	김한준	인공지능	딥러닝 및 특징공학 AI 데이터 품질 개선	02-6490-2339	khj@uos.ac.kr
인공지능학과	노영민	인공지능	인공지능, 컴퓨터비전	02-6490-2476	youngmin.ro@uos.ac.kr

데이터에 대한

개선 후

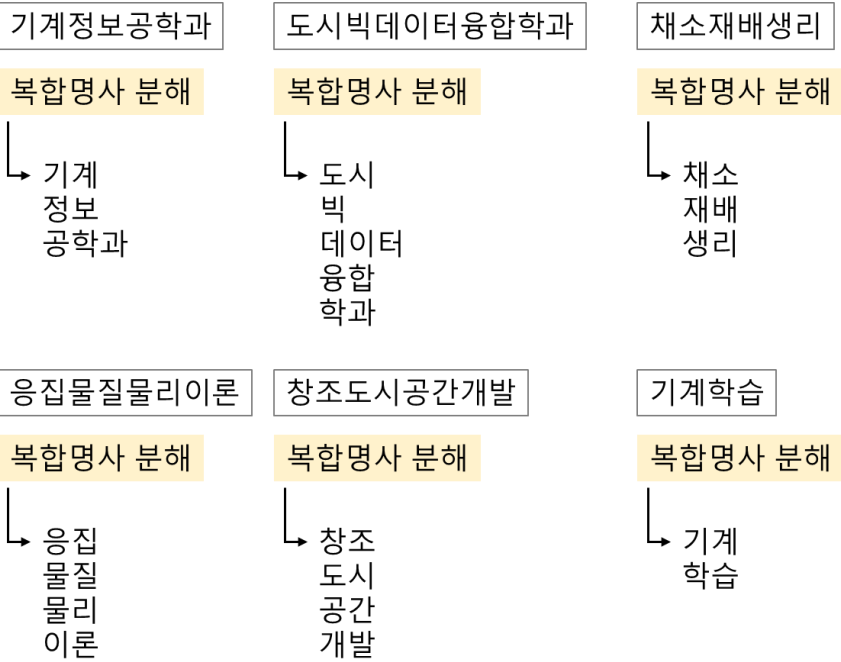
	Score	Document
0	4.7053	데이터 처리 장치 및 데이터 처리 방법
1	4.391	영상데이터와 라이다데이터의 기하학적 정합방법 및 그 장치
2	4.1161	정형 테이블 데이터셋에 대한 딥러닝 기반 데이터 융합 기술 개발
3	4.0079	공간데이터 분석 및 시각화
4	4.0079	3DGIS, 데이터 모델
5	4.0079	환경 데이터 분석 및 모델링
6	4.0079	데이터를 시각화하는 장치 및 방법
7	4.0079	데이터 기록 및 판독장치
8	4.0079	데이터 기록 및 판독장치
9	4.0079	데이터 연산 방법 및 시스템

[그림 8] 불용어 제거 적용 예: '데이터에 대한'의 검색 결과 비교

3. 검색시스템의 개선

❖ 복합명사 분해

❖ (개선 후) '응집물질물리', '응집물질'과 같은 데이터들도 검색됨 [그림 10]



통합검색

응집물질물리

산단 검색 시스템

☐ 반도체/디스플레이 ☐ 사이버보안 ☐ 수소 ☐ 양자 ☐ 이차전지 ☐ 인공지능 ☐ 차세대통신 ☐ 첨단로봇/제조

☐ 첨단모빌리티 ☐ 첨단바이오 ☐ 기타

전체 2 건 현재 페이지 1/20

학과명	교수명	기술분류	연구영역	전화번호	이메일
물리학과	정석범	기타	응집물질물리 이론 (위상학적 물질 상태 및 초전도체)	02-6490-2646	sbchung0@uos.ac.kr
물리학과	정재일	반도체/디스플레이	응집물질물리이론	02-6490-2655	jeiljung@uos.ac.kr

응집물질물리이론

개선 후

	Score	Document
0	23.5015	응집물질물리이론
1	18.3568	응집물질물리 이론(위상학적 물질 상태 및 초전도체)
2	16.1015	응집물질물리실험(신물질박막연구)
3	15.3649	응집물질이론(전자구조계산)
4	10.8266	응집물질의 전자구조연구
5	10.3265	응집물질실험 (고온초전도체, 자성물질)
6	9.4776	이론천체물리, 은하 및 블랙홀 진화 수치실험
7	6.8236	통계물리학
8	6.4489	물리화학적 수처리
9	6.4489	입자실험 및 전산물리

[그림 9] 복합명사 분해: 복합명사 분해 후 인덱스의 변화

[그림 10] 복합명사 분해 적용 예: '응집물질물리이론'의 검색 결과 비교

3. 검색시스템의 개선

❖ 사용자 사전 도입 (약어/동의어 인식)

❖ (개선 후) 머신러닝 → 기계학습, 딥네트워크 → 딥러닝/심층학습 검색 가능 [그림 11]

통합검색

머신러닝

산단 검색 시스템

☐ 반도체/디스플레이

☐ 사이버보안

☐ 수소

☐ 양자

☐ 이차전지

☐ 인공지능

☐ 차세대통신

☐ 첨단로봇/제조

☐ 첨단모빌리티

☐ 첨단바이오

☐ 기타

전체 1 건 현재 페이지 1/20

머신러닝

개선 후

통합검색

심층 학습

산단 검색 시스템

☐ 반도체/디스플레이

☐ 사이버보안

☐ 수소

☐ 양자

☐ 이차전지

☐ 인공지능

☐ 차세대통신

☐ 첨단로봇/제조

☐ 첨단모빌리티

☐ 첨단바이오

☐ 기타

전체 2 건 현재 페이지 1/20

딥네트워크

개선 후

[그림 11] 동의어 인식 예: '머신러닝'과 '딥네트워크'의 검색

4. 검색시스템의 확장

4-1. 유사도 검색 (semantic search)

- ❖ (배경) 한정된 정보를 보완하고자 연구내용을 잘 나타내는 추가 정보를 이용하고자 함 [그림 12]
- ❖ (방법) 문장 단위의 연구정보를 추가, 검색어와의 문맥 상 유사성을 기준으로 검색

연구자	논문	과제	특허
연번	연번	연번	연번
기술분류 (12대)	단과대학	단과대학	단과대학
기술분류 (기존)	학과명	학과명	학과명
단과대학	교수명	교수명	교수명
학과명	논문명	과제명	발명의 명칭
교수명	기준년도	기준년도	국가코드
연구실	논문내용상세	주관부처	등록번호
연구영역		연구분야	특허설명상세
전화번호		적용분야	
이메일		과제내용상세	
홈페이지			
비고			
연구분야상세			

[그림 12] 연구정보를 추가한 최종 데이터

4. 검색시스템의 확장

4-1. 유사도 검색 (semantic search)

❖ '논문상세정보' 구축과정

초록 전문 발췌
The electrochemical behaviour of direct methanol fuel cells (DMFCs) is sensitive to methanol concentration; thus, to avoid external sensors, it is a promising candidate to monitor the concentration of methanol in the fuel circulation loop, which is central to the efficient operation of direct methanol fuel cell systems. We address this issue and report on an extremely robust electrochemical methanol sensing technique that is not sensitive to temperature, cell degradation and membrane electrode assembly (MEA) type. We develop a temperature independent empirical correlation of the dynamic response of cell voltage to step changes in current with methanol concentration. This equation is successfully validated under various operating scenarios at both the single cell and stack levels. Our sensing method achieves an impressive accuracy of ± 0.1 M and this is expected to increase the reliability of methanol sensing and simplify the control logic of DMFC systems.



LLM을 사용하여 원하는 기준으로 요약함

나는 현재 {학과명}분야의 논문의 초록을 '논문이 어떤 연구를 했는지 가장 잘 나타내는 문장'으로 요약할 것이다.

아래의 초록을 요약하는데 다음의 다섯가지 기준을 반영해줘.

1. 영어약자와 특수문자는 반드시 한국어 용어로 바꿔줘.
2. simulation과 result이후 내용은 제외해줘.
3. propose가 들어간 문장을 중점으로 연구 내용과 목적이 자세하게 드러나게 해줘 ('-다'체 구사).
4. 한국어로 답변해줘.
5. 답변은 최대 200자 이내로 해줘.

초록: '{input}'



요약 결과 = '논문내용상세' 필드
직접 메탄올 연료전지(DMFCs)의 전기화학적 행동이 메탄올의 농도에 민감하기 때문에, 연료 순환 루프 내의 메탄올 농도를 모니터링하는 것이 DMFC 시스템의 효율적인 작동에 중요하다. 이 문제를 해결하려고 현재 온도, 셀의 변질 및 막 전극 조립(MEA) 유형에 민감하지 않은 매우 견고한 전기화학적 메탄올 감지 기법을 제안한다.

4. 검색시스템의 확장

4-1. 유사도 검색 (semantic search)

- ❖ '논문상세정보'의 상위 5개 검색 결과를 살펴보자.
- ❖ 시나리오는 1) 검색어가 키워드인 경우와 2) 키워드 검색으로 검색이 불가능한 경우로 구성함

시나리오 1. 검색어가 키워드인 경우 (기계학습)

산단 검색 시스템

학과명	교수명	기술분류	연구영역	전화번호	이메일
공간정보공학과	구형모	첨단로봇/제조	공간데이터 분석 및 시각화	02-6490-2885	hmokoo@uos.ac.kr
인공지능학과	김한준	인공지능	딥러닝 및 특징공학 AI 데이터 품질 개선	02-6490-2339	khj@uos.ac.kr
전자전기컴퓨터공학부	김한준	인공지능	딥러닝 및 특징공학 AI 데이터 품질 개선	02-6490-2339	khj@uos.ac.kr
통계학과	박창이	인공지능	통계적 학습, 비모수적 추론	02-6490-2634	park463@uos.ac.kr

개선 후							
	Row	Score	연번	학과명	교수명	논문명	초록_요약
0	1	0.7407	51	인공지	김한준	A Tensc	텍스트 분류에 일반적으로 사용되는 머신러닝 알고리즘 중 나이브 베이즈 및 서포트 벡터 머신은 합리적인 성능을 제공하지만, 딥 러닝을 활용한 텍스트 분류는 이전
1	2	0.7336	733	통계학	박창이	Multick	이 논문은 다중 클래스 분류 문제에서 레이블이 지정되지 않은 샘플을 활용하여 더 나은 분류 함수를 학습하는 라플라시안 서포트 벡터 머신(LapSVM)을 제안합니다. 드
2	3	0.7267	733	통계학	박창이	Negativ	과분산 계수 데이터를 위한 새로운 그래픽 모델인 영과잉 지역 음이항 그래픽 모델을 제시합니다. 음이항 분포의 두 가지 매개변수화를 이용하여 목적 함수를 최적화하
3	4	0.7201	53	인공지	유하진	Capturi	이 연구에서는 음향 장면 분류에서 자주 오분류되는 클래스 쌍을 구별하기 위해 구체적인 세부 정보를 효과적으로 활용하는 방법을 탐구하였습니다. 딥 뉴럴 네트워크
4	5	0.7195	733	통계학	박창이	Variabl	지지벡터기계(SVM)는 분류문제에서 주로 사용되는 학습 방법으로, 고차원 데이터를 효과적으로 다룰 수 있습니다. 그러나 잡음 변수가 많은 경우 예측력이 저하되고,
5	6	0.7159	61	인공지	최호식	Primal	논문에서는 공분포 데이터에 대한 LASSO(Least Absolute Shrinkage and Selection Operator) 추정기를 개발하였습니다. 이는 회귀 계수에 대한 선형 제약 조건을 고려하
6	7	0.7112	58	인공지	이병정	Utilizin	소프트웨어 유지보수 비용을 줄이기 위해 주제 기반 유사 커밋 정보를 사용하여 버그 위치 지정을 개선하는 방법을 제안한다.버그 보고서의 주제를 결정하고 유사한 버

4. 검색시스템의 확장

4-1. 유사도 검색 (semantic search)

시나리오 1. 검색어가 키워드인 경우 (영상분석)

산단 검색 시스템

학과명	교수명	기술분류	연구영역	전화번호	이메일
기계정보공학과	김태현	기타	임베디드 시스템, On-device AI, 지능형 IoT 시스템	02-6490-2388	thkim@uos.ac.kr
교통공학과	박신형	첨단모빌리티	Smart Mobility, ITS, 교통안전, 도로공학	02-6490-2829	shinhpark@uos.ac.kr
토목공학과	조수진	기타	스마트구조기술	02-6490-2434	soojin@uos.ac.kr

개선 후

	Row	Score	연번	학과명	교수명	논문명	초록_요약
0	1	0.7261	63	전자전	김윤	NOR-Ty	이 연구에서는 직장암 환자의 신장 전 항암 방사선 요법(nCRT) 후 자기 공명 영상(MRI) 평가를 개선하기 위한 효과적인 pCR(병리학적 완전 반응) 예측 방법을 제안합니다.
1	2	0.7118	53	인공지능	유하진	Capturi	이 연구에서는 음향 장면 분류에서 자주 오분류되는 클래스 쌍을 구별하기 위해 구체적인 세부 정보를 효과적으로 활용하는 방법을 탐구하였습니다. 딥 뉴럴 네트워크
2	3	0.71	732	통계학	박재휘	Retriev	GAN 모델의 발전으로 물체 이미지 합성은 사실적이거나, 복잡한 장면 이미지 모델링에 어려움이 남아있음. 논문에서는 자기 지도 보조 작업 및 다중 스케일 특징 앙상블
3	4	0.709	732	통계학	박재휘	Improv	논문에서는 자기 지도 보조 작업 및 다중 스케일 특징 앙상블과 같은 시각적 표현 학습의 패러다임을 활용하여 판별자의 구별 능력을 향상시켜, 도전적인 장면 데이터
4	5	0.6978	62	인공지능	황혜수	Pathol	이 연구에서는 다중 매개 변수 MRI를 활용한 직장암 nCRT 후 pCR 예측에 효과적인 방법을 제안합니다. 여러 MRI 시퀀스를 결합하여 얻은 체적 특징을 활용하는데, 제
5	6	0.6967	20	신소재	김상일	Effect c	이 연구는 브롬 치환이 n-형 층상 SnSe2의 열전 속성을 향상시키는 데 미치는 영향을 조사했다. 브롬 도핑으로 인해 전기 전도도 및 세벡 계수가 동시에 증가하면서, 상
6	7	0.6907	733	통계학	박창이	Variabl	지지벡터기계(SVM)는 분류문제에서 주로 사용되는 학습 방법으로, 고차원 데이터를 효과적으로 다룰 수 있습니다. 그러나 잡음 변수가 많은 경우 예측력이 저하되고,

4. 검색시스템의 확장

4-1. 유사도 검색 (semantic search)

시나리오 1. 검색어가 키워드인 경우 (MRI)

산단 검색 시스템

학과명	교수명	기술분류	연구영역	전화번호	이메일
전자전기컴퓨터공학부	김윤	반도체/디스플레이	나노 스케일 반도체 트랜지스터	02-6490-2352	yoonkim82@uos.ac.kr
전자전기컴퓨터공학부	이주한	차세대통신	광섬유, 군사용, 산업용 레이저	02-6490-2344	jhl@uos.ac.kr
인공지능학과	황혜수	인공지능	빅데이터분석 및 정보검색	02-6490-2454	hwang@uos.ac.kr

개선 후

	Row	Score	연번	학과명	교수명	논문명	초록_요약
0	1	0.7686	63	전자전	김윤	NOR-Ty	이 연구에서는 직장암 환자의 신장 전 항암 방사선 요법(nCRT) 후 자기 공명 영상(MRI) 평가를 개선하기 위한 효과적인 pCR(병리학적 완전 반응) 예측 방법을 제안합니
1	2	0.7545	62	인공지	황혜수	Patholo	이 연구에서는 다중 매개 변수 MRI를 활용한 직장암 nCRT 후 pCR 예측에 효과적인 방법을 제안합니다. 여러 MRI 시퀀스를 결합하여 얻은 체적 특징을 활용하는데, 제인
2	3	0.6639	12	신소재	김상일	Electric	이 연구는 Se로 도핑된 Re2Te5의 합성과 전기 및 열 전도 특성에 관한 것이다. Se 도핑은 캐리어 농도 및 전기 전도도 증가를 통해 880 K에서 최대 전력 인자를 0.135 mW
3	4	0.6523	8	신소재	김상일	Thermoe	이 연구는 MoSe2-MoTe2 시스템의 열전 및 전기 전도 특성을 조사했다. 높은 Te 함량을 갖는 MoTe2 샘플이 높은 전력 인자를 나타내며, 음이온 치환에 따른 격자 열전도
4	5	0.6459	23	신소재	김상일	Enhanc	이 연구는 InSe에 Sn 치환 도핑을 통한 열전 속성의 향상을 보고한다. Sn 도핑에 의해 전기적 특성이 향상되어 전력 인자와 차원이 없는 zT가 증가했다.
5	6	0.6454	325	공간정	정형섭	Precise	이 연구에서는 기존 SAR 간섭측량 및 다중 아파처 SAR 간섭측량의 통합을 통해 국소 및 복잡한 지형에서의 정밀한 3D 변형 검색의 가능성을 평가하였습니다. 특히 201
6	7	0.6394	15	신소재	김상일	Phase f	이 연구에서는 FeSe2-FeS2 시스템의 열전 및 전기 전도 특성을 조사했다. S 도핑에 따른 구조적 변화로 인해 전력 인자의 증가와 효과적인 열전도도의 감소가 관찰되었

4. 검색시스템의 확장

4-1. 유사도 검색 (semantic search)

시나리오 2. 키워드 검색으로 검색이 불가능한 경우

1) 검색어: 희소한 해를 찾기 위한 회귀모형

	Row	Score	연번	학과명	교수명	논문명	초록_요약	
	0	1	0.737	733	통계학	박창이	Negative	과분산 계수 데이터를 위한 새로운 그래픽 모델인 영과잉 지역 음이항 그래픽 모델을 제시합니다. 음이항 분포의 두 가지 매개변수화를 이용하여 목적 함수를 최적화
	1	2	0.7071	733	통계학	박창이	Multicla	이 논문은 다중 클래스 분류 문제에서 레이블이 지정되지 않은 샘플을 활용하여 더 나은 분류 함수를 학습하는 라플라시안 서포트 벡터 머신(LapSVM)을 제안합니다. 다
	2	3	0.699	61	인공지능	최호식	Primal	논문에서는 공분포 데이터에 대한 LASSO(Least Absolute Shrinkage and Selection Operator) 추정기를 개발하였습니다. 이는 회귀 계수에 대한 선형 제약 조건을 고려하
	3	4	0.6965	51	인공지능	김한준	A Tensc	텍스트 분류에 일반적으로 사용되는 머신러닝 알고리즘 중 나이브 베이즈 및 서포트 벡터 머신은 합리적인 성능을 제공하지만, 딥 러닝을 활용한 텍스트 분류는 이전
	4	5	0.6917	321	공간정보	정형섭	Mappir	최근 기후 변화로 인한 온난화로 가뭄 빈도가 증가하며 국가적인 수위 관리의 중요성이 증대되고 있습니다. 본 연구는 대한민국 공주시의 다중 계절 광학 위성 이미지
	5	6	0.6836	41	신소재	김영욱	Electric	이 연구에서는 N 도핑된 SiC 세라믹의 전기 전도도에 영향을 미치는 주요 요인들을 검토하고, SiC 폴리타입, 입계 구조, 용해성 원자, 기공률 등이 전도도에 미치는 영향
	6	7	0.6832	56	인공지능	유하진	Regula	이 연구에서는 기존의 화자 확인 모델인 확률적 선형 판별 분석(PLDA)을 개선하기 위해 정규화된 클래스 내 정밀도 행렬을 도입한 방법을 제안합니다. 이를 통해 GLAS

2) 검색어: AI를 활용한 MRI영상분석

	Row	Score	연번	학과명	교수명	논문명	초록_요약
	0	1	0.7942	63	전자전	김윤	NOR-Ty 이 연구에서는 직장암 환자의 신장 전 항암 방사선 요법(nCRT) 후 자기 공명 영상(MRI) 평가를 개선하기 위한 효과적인 pCR(병리학적 완전 반응) 예측 방법을 제안합니
	1	2	0.7479	62	인공지능	황혜수	Patholo 이 연구에서는 다중 매개 변수 MRI를 활용한 직장암 nCRT 후 pCR 예측에 효과적인 방법을 제안합니다. 여러 MRI 시퀀스를 결합하여 얻은 체적 특징을 활용하는데, 제안
	2	3	0.7314	317	공간정보	정형섭	ForestV 항공 이미지 및 머신 러닝 기술을 활용한 연구에서는 무인 항공기를 통해 획득한 LiDAR 유도 DSM을 활용하여 숲 구조를 효과적으로 분류하고 성능을 향상시켰다. XGB
	3	4	0.7214	51	인공지능	김한준	A Tensc 텍스트 분류에 일반적으로 사용되는 머신러닝 알고리즘 중 나이브 베이즈 및 서포트 벡터 머신은 합리적인 성능을 제공하지만, 딥 러닝을 활용한 텍스트 분류는 이전
	4	5	0.7173	52	인공지능	유하진	Self-Su 본 연구는 딥 러닝을 활용한 자가-지도 학습 방법을 제안하여, 낮은 방사선량 CT(LDCT) 이미지와 정상적인 방사선량 CT(NDCT) 이미지의 수가 부족한 상황에서도 효과
	5	6	0.717	324	공간정보	정형섭	Ship De 해상 합성곱계 레이더 (SAR) 이미지를 활용한 선박 탐지는 시간과 날씨의 영향을 받지 않아 많은 연구에서 주목받고 있습니다. 본 연구는 딥러닝을 활용한 SAR 이미지
	6	7	0.7123	733	통계학	박창이	Variabl 지지벡터기계(SVM)는 분류문제에서 주로 사용되는 학습 방법으로, 고차원 데이터를 효과적으로 다룰 수 있습니다. 그러나 잡음 변수가 많은 경우 예측력이 저하되고

4. 검색시스템의 확장

4-2. 연구자 관계도

❖ 특정 연구 내용과 관련 있는 연구자를 찾는 것을 시각화 관점의 서비스 제공

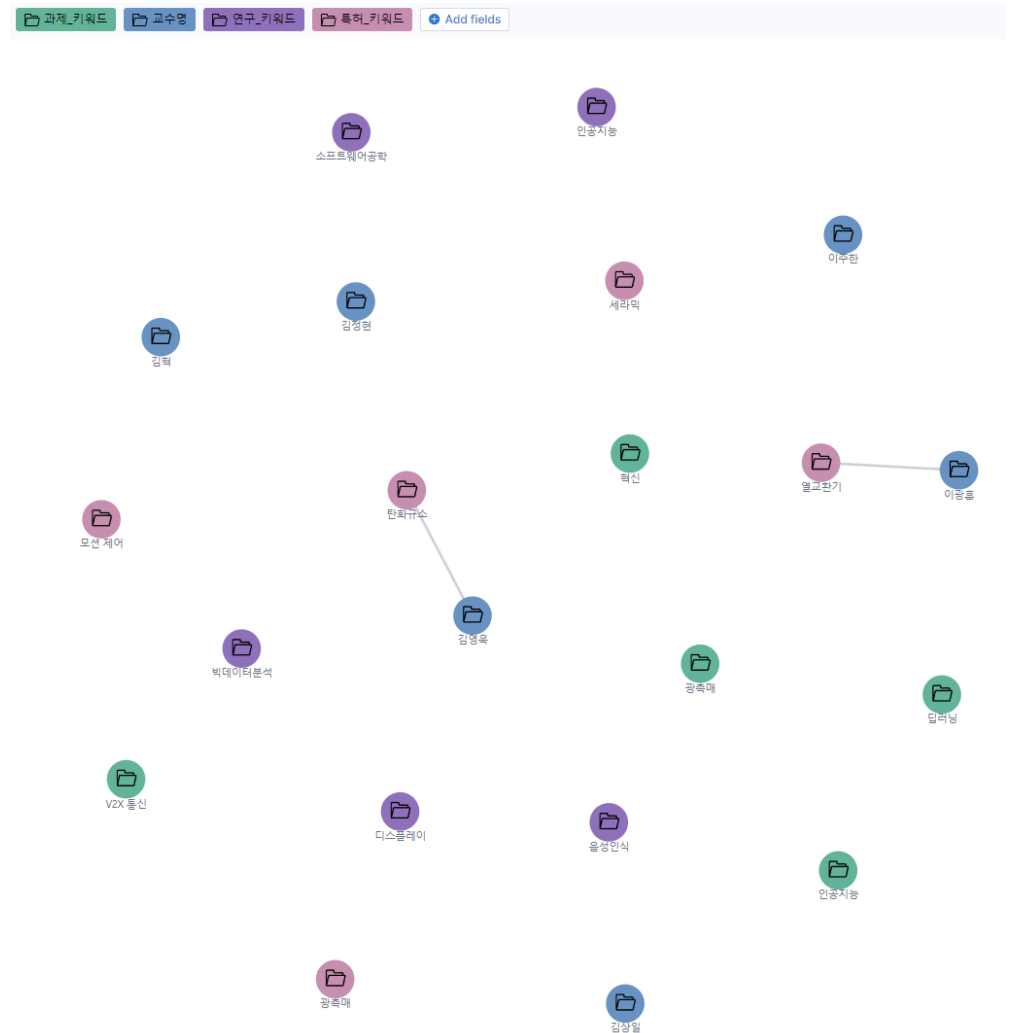
연구자	논문	과제	특허
교수명	교수명	교수명	교수명
연구영역	논문명	과제명	특허명

아래의 데이터에서 이 연구자의 연구영역을 가장 잘 나타내는 키워드를 '한글 명사형의 단어 한 개'를 추출해줘.

데이터: '{연구영역}'

교수명	연구_키워드	논문_키워드	과제_키워드	특허_키워드
강인혜	인공지능			
나영승	연료전지			
황면중	로봇운동계획			
나영승		에너지저장		
유하진		인공지능		
김태현			IoT	
황면중			로봇	
황면중				모션제어

[그림 13] 연구자 관계도를 위한 데이터 재구성



[그림 14] 연구자 관계도