



CPU에 적용하는 특허 빅데이터 분석

부제

비전공자도 수상하는 분석
방법론



CONTENTS

Part 0. 분석 기법 이해

1. 벡터화 / 임베딩
2. TF-IDF
3. Word2Vec / Doc2Vec / GloVe
4. LDA 분석
5. SNA 분석
6. 워드클라우드
7. 문현 유사도

PART 1. 실전 분석 방법론

1. 문제를 기반 검색식 도출
2. 파이썬으로 기술 트리 구축
3. LLM으로 유효특허 선별
4. 정량분석 인사이트 도출
5. SNA 분석으로 기술 네트워크
분석
6. 워드클라우드로 기술 흐름 파악
7. 유사도 분석법으로 회피 전략
수립

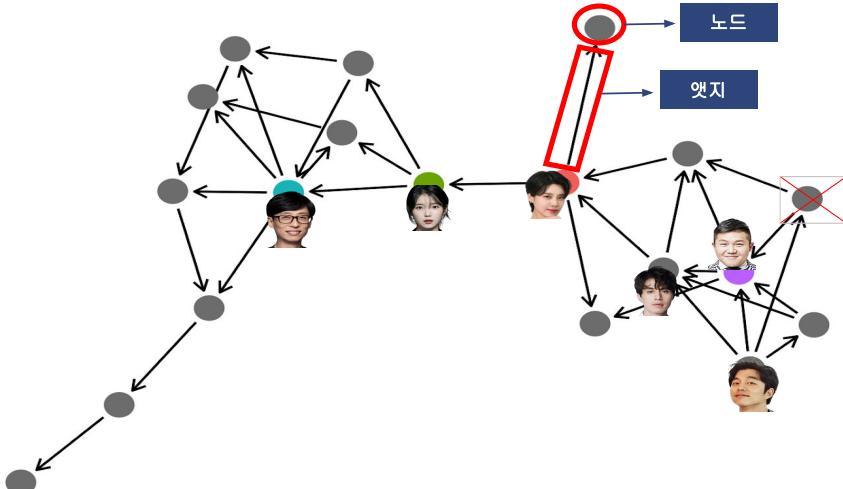
PART 2. 8가지 인사이트 프레임워크

1. 신기술 센싱
2. 기술 트렌드 분석
3. 유망 기술 도출
4. 신산업 센싱
5. 경쟁 우위 분석
6. 미래 특허 전략 수립
7. 기술 확장 가능성 센싱
8. 특허 분쟁 가능성 사전 검토

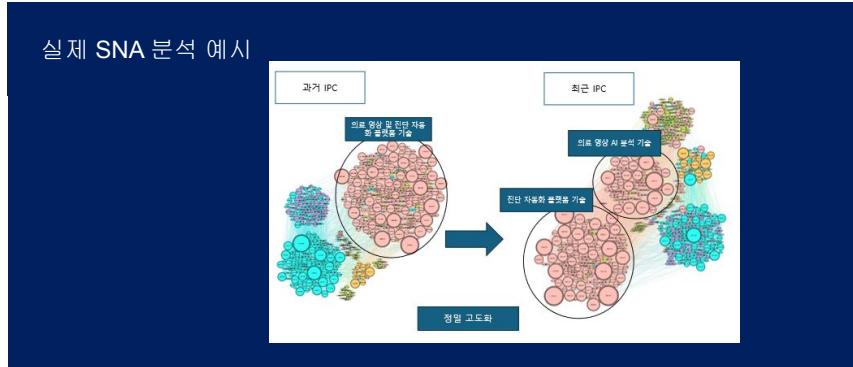
PART 3. 분석 실전 완주하기: 기획부터 인사이트 도출까지

1. 수상 보고서 해부
2. CPU 보고서 작성 실습

5. SNA 분석



실제 SNA 분석 예시



Highest Centrality



- Betweenness Centrality
- Closeness Centrality
- Degree Centrality
- Eigenvector Centrality
- NA

최단 거리에 가장 많이 포함되는 노드
모든 노드와 가장 가까운 노드
직접 연결된 노드가 가장 많은 노드
영향력이 큰 노드들과 연결된 노드



실무에서는 EC 중심성을 바탕으로 시각화를 진행해 다양한 분석을 진행



II. CPU 적용

•방법으로

1. 검색어 작성법
2. LDA로 트리 만들기
3. 유효특허 선별법
4. 기본 정량분석 마스터하기
5. SNA 분석법
6. 워드 클라우드 분석법

1강. GPT로 검색식 작성

키워드 기반 모집단 구성

리튬 회수 검색식

(TAC:(리튬* 리튬* 리튬* 리* 리튬* 리튬음* 리튬음* 리튬음* 리디엄* 리티움* 리티움*리지암* 리죽음* 리지암* 리튬음**Li** lithium') AND
 metallurg* smell* 주출* extract* 회수* 리커버* recover*) pyrometallurg*
 배소* 로스트* 로스팅*
 roast* (습식* wet* hydrot) N/1 (제련*metallurg smell*주출* extract*
 회수*리커버* recover)
 hydrometallurg* 침출*리칭*leach* 주출* extract* 회수*리커버* recover*
 재활용*재이용+재생* 재사용* reus* 리사이클* 리싸이클* recycl*)
 IPC:(B09B* C22B*) AND
 (TAC_DSC:(배터리*배터리* 배터리* 배터리* 더리*빼데리* 전지* 건전지*
 이자전지* 폐배터리* 폐배터리* 폐배터리* 폐배터리* 폐더리* 폐전지*
 폐전지* 폐건전지* 폐이자전지* batter* LiB*
 Waste-batter waste-cell*내김*새거* sagger sagger* 정극재*정극재* 정극활*
 양극재* 양극재* 양극활* 양극* 음극* 어노드*애노드* anod* 캐소드* 캐소드*
 캐소드* cathod* ((활* 활성* 액티브* 액티브* active*) A/1
 (물질*재료* material*)) ((블랙* 검은* black*) A/1 (매스* 덩어리* 파우더*
 mass* powder*))) IPC:(H01M*)

문제 요청 예시

- 리튬은 매장량이 제한적이고, 공급망의 불안정성이 커지고 있으며, 특히 전기차 배터리에 필수적인 리튬의 수급 문제가 지속되면서 이를 재활용하여 확보하는 기술의 필요성이 증대됨
- 리튬을 효율적으로 회수하여 리튬 자원의 손실을 최소화하고, 회수율을 높이는 연구·개발이 필요함

해결해야하는 문제 요약

1. 배터리에서 리튬 회수
2. 건식과 습식 추출 확인

검색식 요약

(리튬 and (건식 n/1 주출) or (습식 n/1 주출))
 IPC)
 And (배터리 and IPC)

검색식 확장



핵심 출원인 특히 용어 추가



특허 검색식 생성기 GPTs

특허 검색식 GPTs 제작 및 활용

2강. 파이썬으로 기술 분류 트리 구축- LDA 분석법

기술 트리 구축 프로세스

2단계 분류
핵심 공법 (Mixing)
코팅 공법 (Coating)
설진 공법 (Stamping)
노동 공법 (Working)
진공 건조 공법 (Vacuum Drying)
프레싱 기반 박아제 제작
풀프로세스 기반 공법
WIP Cell Assembly
고운 풍부 가동 공법 (WIP: Work in Progress)
바이셀 스택 공법 (Bio-Cell Stacking)
디파킹 공법 (Depacking)
풀셀 스택 공법 (Full-Cell Stacking)
접합 및 공법 (Tab Welding)
최종 패키징 공법 (Final Packaging)
활성화 공법 (Formation)



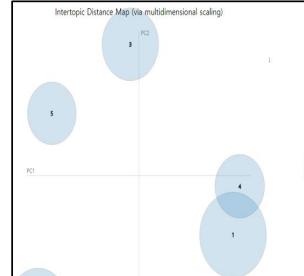
K WORK

분류된 특허를 KWORK에 저장



description_LDA_WORD

파이썬 코드를 통한 LDA 분석



각 주제별 LDA 분석

문제의 요청에 따른 특허 분류

분류	특허 번호	제작자
1	10-123456789	A
2	10-123456790	B
3	10-123456791	C
4	10-123456792	D
5	10-123456793	E

LDA 결과를 바탕으로 3단계 기술 트리 구축

분류	특허 번호	제작자
1	10-123456789	A
2	10-123456790	B
3	10-123456791	C
4	10-123456792	D
5	10-123456793	E

LLM을 활용하여 3단계 기술 트리 정의 구축

3강. GPT으로 유효특허 선별

유호 건수 총 3,699건								노 이즈
증설연	증설부지면적	변경면적	고정자단기	동자연자단기	최신기	술우선	재진반	
H2020(부록부지)	97	-	16	19	20	11	1	4
ITD부지면적	639	59	33	254	6	198	42	87
a48_진단지	1,459	365	152	238	159	1,176	164	1,201
GE	700	13	-	-	151	212	-	454
Siemens	951	53	55	33	151	212	6	461
포장부지면적	89	19	9	6	4	22	9	9
발달면적	12	-	6	-	-	-	-	6

LLM을 활용한 기술 분류 기반 노이즈 제거



기술 분류 정의 파일

분석 대상 특허 파일

분석을 위한 두 가지 파일 준비



기술 분류 정의 파일

기술 분류 정의 구축 프로세스



• keywert Google



해당 정의를 바탕으로 특허 자동 분류 기준 설정



기술 분류 프로세스

상기 프롬프트를 활용한
LLM 기반 기술 분류 적합성 확인

상기 프롬프트를 활용한 LLM 기반 기술 정의 추가

으로 유효특허				
차단기	회전기	슬iding	폐쇄판	노이즈
9	20	11	2	4
54	6	195	42	67
88	139	1,176	164	1,201
-	151	82	-	454
3	131	212	6	461
5	4	22	9	9
-	-	-	-	6

五

문석을

기술 분류 정의 파일

분석 대상 특허 파일

위한 두 가지 파일

준비

 WPS
Office

t
ence

```

4 1. System Project
  4 1.1. Project Properties
    4 1.1.1. General
      4 1.1.1.1. [Name] (이름)에 대한 기본 ([Name]), 소스 가로 ([Source], [Name]의 소스를 포함하는 모든 파일을 포함하는 기본 경로) 및 기본 경로 ([Path])에 대한 기본 경로 ([Path])를 설정합니다.

  4 Instruction
    4 1.2. General
      4 1.2.1. use_project (프로젝트 사용)
        4 1.2.1.1. 사용하는 프로젝트를 선택합니다. 선택한 프로젝트는 기본 경로로 설정됩니다.
        4 1.2.1.2. **설정하는 프로젝트는 사용하는 기본 경로를 바탕으로, 정의한 기본 경로 및 기본 경로 폴더에서 시작되도록 설정됩니다.** 설정하는 프로젝트는 사용하는 기본 경로를 바탕으로, 정의한 기본 경로 및 기본 경로 폴더에서 시작되도록 설정됩니다.
        4 1.2.1.3. 사용하는 프로젝트를 선택한 후 [OK] 버튼을 클릭합니다. 선택한 프로젝트는 기본 경로로 설정됩니다.
        4 1.2.1.4. 사용하는 프로젝트를 선택한 후 [Cancel] 버튼을 클릭합니다. 선택한 프로젝트는 기본 경로로 설정되지 않습니다.
        4 1.2.1.5. 사용하는 프로젝트를 선택한 후 [Apply] 버튼을 클릭합니다. 선택한 프로젝트는 기본 경로로 설정됩니다.

      4 1.2.2. Project (프로젝트)
        4 1.2.2.1. Project (프로젝트) 항목을 선택합니다. 선택한 항목은 기본 경로로 설정됩니다.

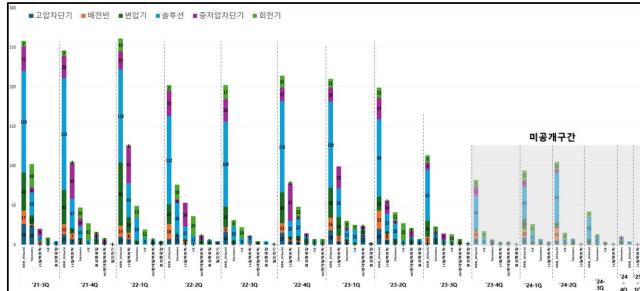
  4 Output Format
    4 1.3. General
      4 1.3.1. [Format] (형식) 항목을 선택합니다. 선택한 항목은 기본 경로로 설정됩니다.
        4 1.3.1.1. [Text] (문서), [HTML] (HTML), [PDF] (PDF), [Word] (Word), [Excel] (Excel), [CSV] (CSV), [JSON] (JSON), [YAML] (YAML), [XML] (XML), [INI] (INI), [INI (Multi-line)] (INI (다행선)), [INI (Multi-line, quoted)] (INI (다행선, 인용)) 및 [INI (Multi-line, quoted, escaped)] (INI (다행선, 인용, 탈출)) 등 다양한 형식을 지원합니다.

```

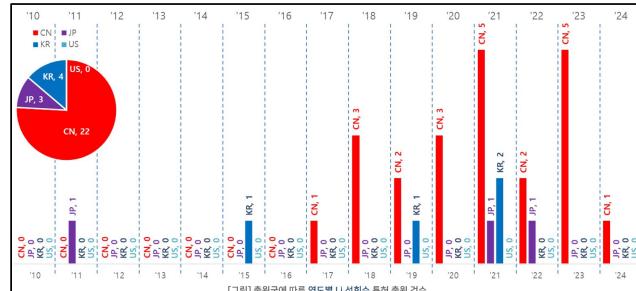
4강. 엑셀을 활용하여 정량분석

팩스풀 바로 활용할 수 있는 excel 특허 실무

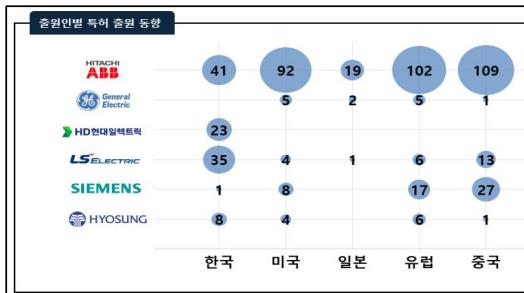
https://www.patspoon.com/course/course_view.jsp?id=178598&cid=#course-view-178598



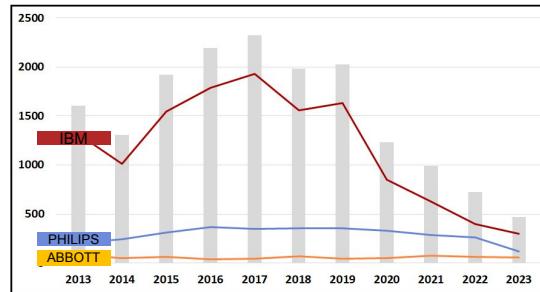
출원인·기술 분류별 특허 출원 동향 분석 시각화



국가별 유효특허 출원 건수 시계열 분석 시각화



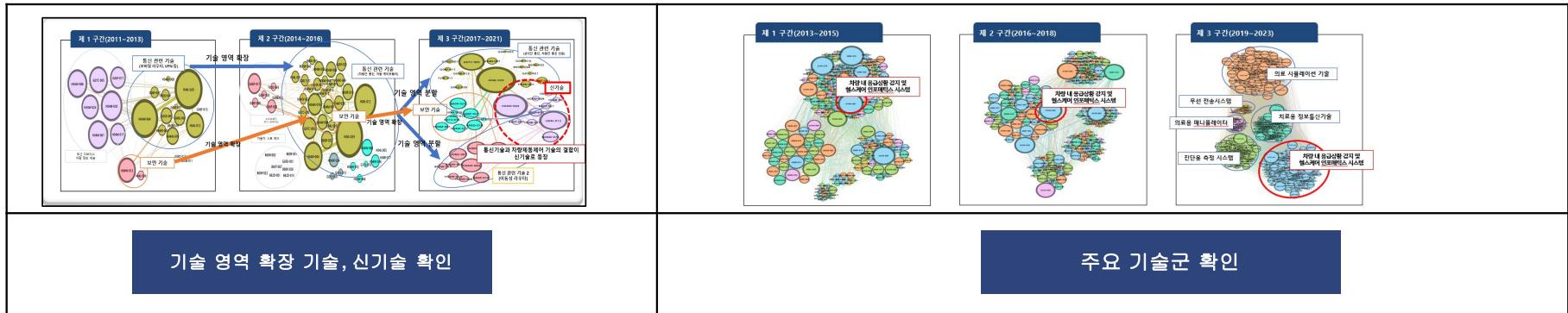
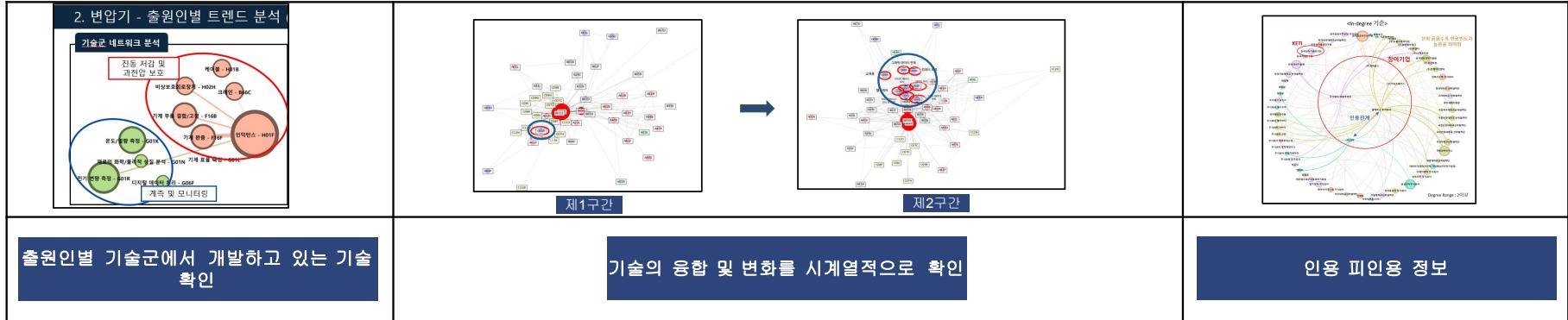
국가별·출원인별 특허 출원 동향 분석
시각화



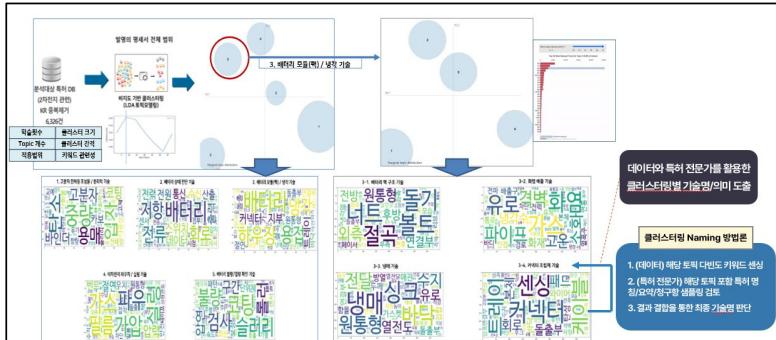
출원인별 연도별 출원 동향 시각화

5강. SNA 분석으로 기술 네트워크 분석

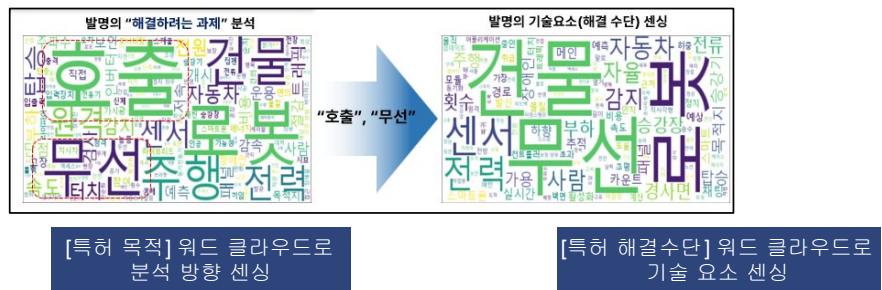
분석 적용 예시



6강. 워드클라우드로 기술 흐름 파악

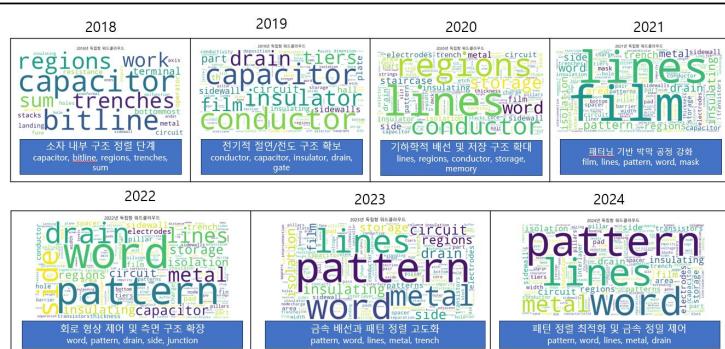


LDA로 도출한 클러스터 네이밍

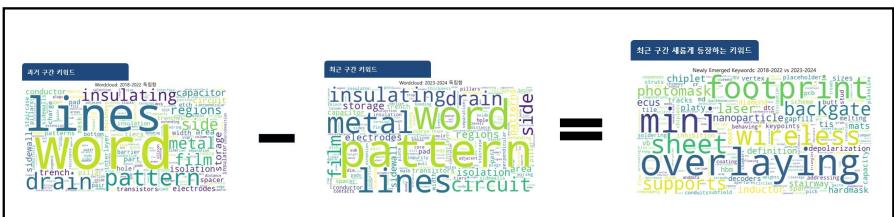


[특허 목적] 워드 클라우드로 분석 방향 센싱

[특허 해결수단] 워드 클라우드로 기술 요소 센싱



[특허 독립항] 시계열 적인 워드클라우드로 기술 흐름 확인



과거에 등장하지 않았던 새로운 키워드들을 바탕으로 신기술
세상

빅데이터 인사이트 프레임 워크 문제 적용 방안

빅데이터 분석 전략 목표	2025 CPU 특허전략 부분 적용 가능 문제
기술 트렌드 분석	A4,A6,A8, A10 ,A11,A13,A14,A15,B1,B3,B5, B6,C2
경쟁 우위 분석 (경쟁사 분석)	A1,A4,A7,A8,A9, A10 ,B3,B6,C1
유망 기술 도출 (성장 가능성이 높은 기술)	A3,A9,A13,A15,B2,B6,C1,C2
미래 전략 수립(기술 격차) -OSMATRIX	A2,A4,A5, A10 ,A11,A12,A14,B1,B2, B5,B6,C1
리스크 사전 검토	A1,A2,A5,A6,A7,A8, A10 ,B3,B4,B5,C1,C2
신기술 센싱 (새롭게 떠오르는 기술)	A6,A8,A9,A14,A16,B6,C2
신규 기술 요소 센싱 (고유명사의 기능성 확장 신호 확인)	A7,B4
신산업 구조 분석(새로운 시장 탐색)-대안	A8

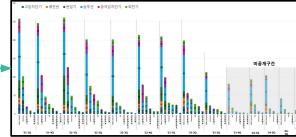
A10 번 문제 적용 예시

출제 문제	Display Sensor
-------	----------------

■ [기술 또는 문제에 대한 간단한 배경설명]

최근 디스플레이 기술이 발달하면서 다양한 센서가 외장형 또는 내장형으로 디스플레이에 통합되는 추세이다. 디스플레이에 화면을 통해 정보를 제공하는 동시에 사용자의 다양한 신호를 감지하여 상호 작용하는 중요한 인터페이스 역할을 한다. 디터셔서 앞에서 살펴본 디터셔에서, 디터셔뿐만 아니라, **광 센서**, **바이오 센서** 등 다양한 센서가 개발됨에 따라 디스플레이와 결합하여 스마트폰, 태블릿, 자동차, 웨어러블 기기 등 애플리케이션의 활용 범위도 넓어졌다.

이처럼 디스플레이 센서 기술이 점점 정교해지면서 높은 정확도, 빠른 응답속도, 내구성 확보 등의 기술적 과제가 부각되고 있다. 관련하여 다음과 같은 조사를 수행하시오.



□ [문제]

- Display에 적용되는 Sensor군을 조사하고, 이에 대한 특허법을 작성하시오.
 - US, CN, EP, JP, KR 특허 대상으로 할 것
 - 정량분석을 실시할 것
 - Display에 적용되는 다양한 Sensor군을 조사하여, 관련 기술분류 기준을 정하고, 그에 따른 특허분석 범위를 정한 것.
예시) 압력 센서, 지문 센서, 터치 센서, 내/외장형 구분 … 등, 특히 분류기준을 제시
 - 정해진 분석 범위에 따라 국가/기준분야/권리자/연도 등 정량 분석을 수행 후 의미 도출
 - 정성분석을 실시
 - 주요권리자별 특허포트폴리오, 비교, 주요특허 선정
예시) 특허법 범위에서 주요 특허 기술의 중요도/장점 권리범위 분석

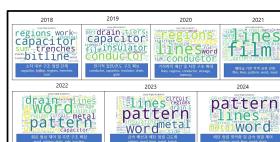
Patspoon을 활용한 정량분석
그래프 및 이사이트 도출



The diagram illustrates the SNA architecture with the following components and their interactions:

- Host Computer**: The central component, represented by a large red circle labeled "Host Computer". It interacts with both the **Local Area Network** (LAN) and the **Wide Area Network** (WAN).
- Local Area Network (LAN)**: Represented by a blue circle. It contains a **File Server** (FS), a **Terminal**, and a **Host Computer**. The LAN is connected to the Host Computer.
- Wide Area Network (WAN)**: Represented by a green circle. It contains a **Host Computer**, a **Switch**, and a **Router**. The WAN is connected to the Host Computer.
- Network Interface Card (NIC)**: A small yellow box at the bottom representing the interface between the host computer and the network.
- Physical Layer**: A blue box at the bottom representing the physical connection layer.
- Protocol Layer**: A green box at the bottom representing the protocol layer.
- SNA Application Layer**: A red box at the bottom representing the application layer.

경쟁 우위 분석



어드클라우드
통한 기술 추세 요약

기술트렌드 분석

미래 전략 스트

리스크 사전 검토

2. 주요 출원인(권리자) 특허 포지션 분석

상기 특허법에서 도출된 주요 출원인 중 하나 이상 선정. 해당 출원인의 특허보지권, 특히 전략등을 도출하시오 (필요 시 출원인이 보유한 전체특허(다른 기술 포함) 포함) 포함하고 비교를 통해 해당 출원인의 전략을 효과적으로 설명할 수 있음)



코어 및 이미징 클러스터 기반
미래전략 수립

C110-B1 (SUS00010) (General mode of wireless cellular networks and communications services)		A-14-B2 (SUS00005) (Emergency apparatus and systems)	
SUB-CODE		SUB-CODE	
C110-B1-01	General mode of wireless cellular networks and communications services	A-14-B2-01	Emergency apparatus and systems
C110-B1-02	Emergency mode of wireless cellular networks and communications services	A-14-B2-02	Emergency apparatus and systems
C110-B1-03	Emergency mode of wireless cellular networks and communications services	A-14-B2-03	Emergency apparatus and systems
C110-B1-04	Emergency mode of wireless cellular networks and communications services	A-14-B2-04	Emergency apparatus and systems
C110-B1-05	Emergency mode of wireless cellular networks and communications services	A-14-B2-05	Emergency apparatus and systems
C110-B1-06	Emergency mode of wireless cellular networks and communications services	A-14-B2-06	Emergency apparatus and systems
C110-B1-07	Emergency mode of wireless cellular networks and communications services	A-14-B2-07	Emergency apparatus and systems
C110-B1-08	Emergency mode of wireless cellular networks and communications services	A-14-B2-08	Emergency apparatus and systems
C110-B1-09	Emergency mode of wireless cellular networks and communications services	A-14-B2-09	Emergency apparatus and systems
C110-B1-10	Emergency mode of wireless cellular networks and communications services	A-14-B2-10	Emergency apparatus and systems

유사도 및 청구항 분석 통한
리스크 사전 경보