발행일 2019. 04. 23
발행처 융합연구정책센터

2019 **04** APRIL vol. 144

융합 FOCUS

융합연구정책센터
Convergence Research Policy Center

융합 FOCUS

융합연구정책센터
Convergence Research Policy Center

발행일 2019년 04월 23일
발행처 한국과학기술연구원 융합연구정책센터
02792 서울특별시 성북구 화랑로 14길 5
Tel. 02-958-4980 http://crpc.kist. re. kr
펴낸곳 (주)동진문화사 Tel. 02-2269-4783

국제특허분류(IPC) 코드를 활용한 기술 간 연관 분석
- 인공지능 사례 중심으로-

박혜경 한국과학기술연구원 융합연구정책센터

01
선정배경

- 특허 데이터는 기술의 명칭, 출원 국가, 출원인, 출원일, 등록일, 청구항, 국제특허분류(IPC) 코드 등 다수 정보 포함
- 특허 데이터의 국제특허분류(IPC) 코드*를 활용하여 기술 간 연관성 분석 가능
  * 국제적으로 통일한 특허분류체계로 모든 특허 문헌을 8개 기술 분야로 분류
- 국제특허분류(IPC) 코드에 대해 구체적으로 알아보고 이를 활용하여 기술 간 관계를 알아보기 위한 방법을 소개하고, 인공지능 기술을 사례로 제시하고자 함

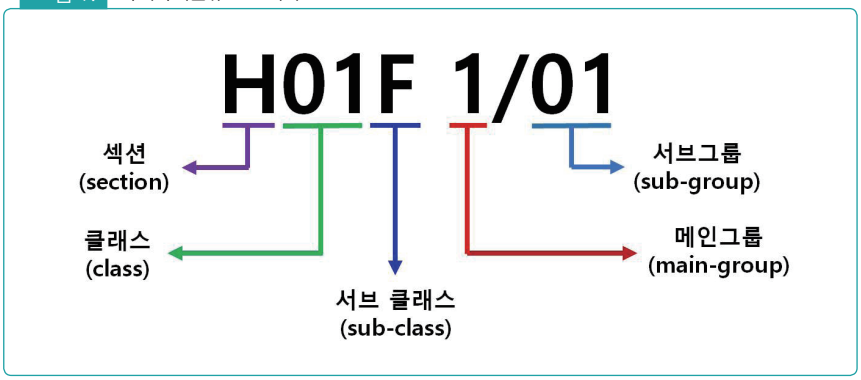
02
국제특허분류(IPC)의 개요

- **(국제특허분류의 정의)** 국제특허분류(IPC: International Patent Classification)는 각국마다 상이하게 사용되던 특허분류체계를 국제적으로 통일한 특허분류체계
  - 미국(USPC: US Patent Classification), 일본(JPC: Japanese Patent Classification), 유럽(ECLA: European Classification system)
- **(국제특허분류의 발생 배경)** 국가 간 기술교류 및 타국가의 특허문헌조사 요구를 반영하여 국제협약을 통해 통일된 하나의 분류체계 확립
  - 국제특허분류에 관한 내용이 담긴 IPC 제 1판이 유럽조약에서 처음으로 도입('68년)
  - 이후, '71년 스트라스부르 협정(Strasbourg Agreement concerning the International Patent Classification)에서 새롭게 개정되어 '75년 10월 발효
  - '98년까지 미국, 러시아, 독일, 영국, 일본 등 전세계의 많은 국가들이 가입
  - 우리나라는 '81년 7월, 국제특허분류를 공식적으로 도입, '99년 10월 8일 스트라스부르 협정에 가입

- 기술발전에 따라 특허분류도 수정이 필요하여 '75년부터 5년마다 가맹국, 국제사무국 등의 제안에 따라 전문가위원회를 통해 정기적으로 개정

(국제특허분류 코드의 구조) 섹션(section), 클래스(class), 서브클래스(sub-class), 메인그룹(main-group), 서브그룹(sub-group)으로 구성된 계층적 구조로 이루어짐

그림 1. 국제특허분류 코드의 구조



- 모든 기술을 8개의 분야로 분류, 첫 단계인 섹션은 가장 큰 기술 분류를 나타내며 A부터 H까지의 알파벳 기호로 표시

표 1. 국제특허분류 코드의 각 섹션 설명

섹션	내용
A	Human Necessities (생활필수품)
B	Performing Operations; Transporting (처리조작; 운수 등)
C	Chemistry, Metallurgy (화학, 야금 등)
D	Textiles, Paper (섬유, 지류 등)
E	Fixed Constructions (고정구조물 등)
F	Mechanical Engineering, Lighting, Heating, Weapons(기계공학, 조명, 가열, 무기, 폭파 등)
G	Physics (물리학)
H	Electricity (전기)

- 하위로 갈수록 기술 내용이 더욱 구체화 됨

표 2. 국제특허분류 코드의 예

IPC 코드 계층	설명
H	전기
H01	기본적 전기소자
H01F	자석; 인덕턴스(Inductance); 변성기; 자기특성을 위한 재료의 선택
H01F 1/01	무기재료로 된 것

(국제특허분류 코드 사용의 장점) 해당 기술의 국제특허분류 코드 확인을 통해 관련 특허 검색 용이

03
국제특허분류 코드를 이용한 기술 간 연관 분석 예시: 인공지능

(분석의 배경 및 목적) 인공지능은 4차 산업혁명의 기반기술로, 국가 성장 동력의 중요한 수단으로 각광받게 되면서 세계 각국에서는 이를 위한 연구개발 투자를 증대

- 인공지능은 여러 학문이 연계된 융합분야이며, 기술적 관점에서 인간의 인지, 학습, 추론 등 지적능력의 일부 또는 전체를 컴퓨터로 구현하는 기술로 정의

표 3. 인공지능의 다양한 정의

구분	개념
존 매카시(1956)	지능적 기계를 만드는 과학 및 엔지니어링
마쓰오 유타카(2015)	기계(컴퓨터)가 인간수준의 인지, 이해, 추론, 학습 등의 사고 능력을 모방하는 기술
ETRI(2017)	기계(컴퓨터)가 인간수준의 인지, 이해, 추론, 학습 등의 사고능력을 모방하는 기술
Gartner(2018)	사람과 자연스러운 대화를 나누고, 인간의 인지 능력을 향상시키거나, 반복적 작업 수행 시 사람들을 대체함으로써 인간을 모방하는 기술
과학기술정보통신부(2018)	인지, 학습 등 인간의 지적능력(지능)의 일부 또는 전체를 컴퓨터를 이용해 구현하는 지능

(출처 : KISTEP 기술동향브리프(2018))

- 여러 학문이 연계된 기술인 인공지능에 대한 특허 데이터의 국제특허분류 코드 분석을 통해 기술 간 연관관계를 제시

(인공지능의 기술분류) 학습 및 추론, 상황이해, 언어이해, 시각이해, 인지인식 및 인지기술로 구분

- 인공지능은 연구자에 따라 분류방식이 다양하며, 본 사례분석은 한국지식재산연구원(2016)에서 활용된 분류를 적용하여 분석

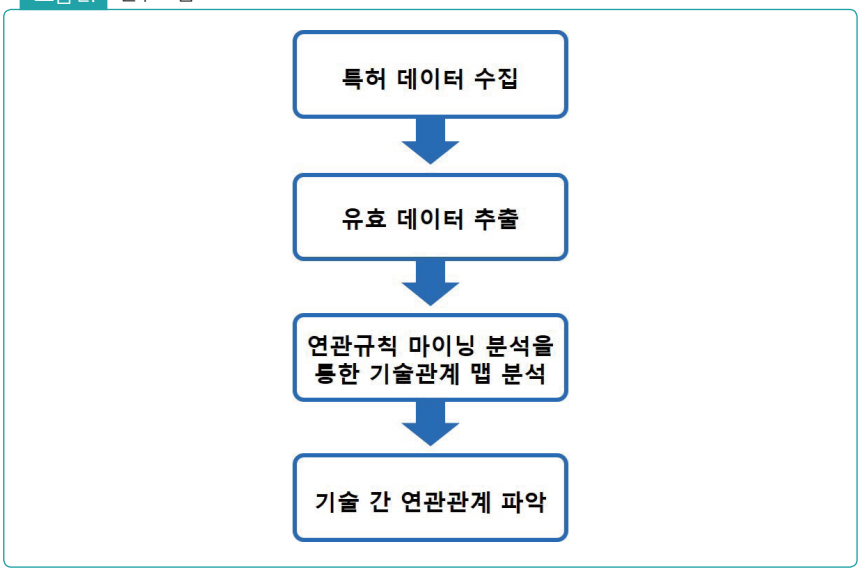
표 4. 인공지능 핵심 기술 및 세부기술 분류

구분	세부기술
학습 및 추론	지식표현, 지식베이스
상황이해 기술	감정이해, 공간이해, 협력지능, 자가이해
언어이해기술	자연어 처리(형태소 분석, 개체명 인식, 구문분석, 의미분석) 질의응답, 음성처리, 자동통번역
시각이해 기술	내용 기반 영상 검색, 행동인식, 시각지식
인식 및 인지기술	휴먼라이프 이해, 인지 아키텍처

(출처 : 한국지식재산연구원 이슈페이퍼(2016) 발췌 및 정리)

(연구방법) 인공지능에 대한 특허 데이터를 수집 후 유효 데이터를 추출하여 연관규칙 마이닝 분석을 통해 기술관계 맵으로 연관된 기술 파악

그림 2. 연구 흐름도



※ **연관규칙 마이닝(Association rule mining)**

- 아이템(item)과 트랜잭션(transaction)을 이용하여 기술 연관성 및 기술 간 관계 분석을 제시하는 데이터 마이닝 분석 방법
  - 아이템: IPC코드, 트랜잭션: 특허
- 개별 트랜잭션은 번호와 이에 해당하는 아이템들로 구성됨

표 5. 트랜잭션 데이터 집합

특허번호	포함된 IPC코드
US09998219	G10D, F06B
US12472680	F06B, H04C
US12719217	G10D, F06B, H04C

- 연관 규칙 발견을 위한 3가지의 평가기준

기준	설명
지지도(Support)	· 아이템 X와 아이템 Y가 동시에 발생할 확률 의미 · 1에 가까울수록 X와 Y는 연관성 높음
신뢰도(Confidence)	· 아이템 X가 발생할 때 아이템 Y가 발생할 가능성 나타냄
향상도(Lift)	· X와 Y의 신뢰도를 Y의 지지도로 나눈 값 · 1보다 크면 아이템 X와 아이템 Y가 상호보완 관계,1이면 X와 Y는 독립관계, 1보다 작으면 X와 Y가 상호대체 관계(X가 증가하면 Y는 감소, X가 감소하면 Y는 증가)

(데이터 수집) 특허 데이터베이스인 WIPSON을 이용하여 인공지능 관련 특허 데이터 수집
- '18년 12월 31일까지 한국, 일본, 중국, 미국, 유럽에 출원된 총 693건의 특허
- 인공지능 관련 기술* 분류 키워드를 포함하여 검색식 작성

  * 학습 및 추론, 상황, 이해, 인식 및 인지

특허데이터 도출 검색식
((((인공* OR artificial*) ADJ (지능* OR intelligen*)) OR AI).KEY) AND ((자율* OR 자가* OR autonom*) AND ((학습 OR 러닝 OR learn*) AND (추론 OR 추정 OR 예측 OR infer* OR reasoning OR presum* OR conject* OR predict*) AND (이해 OR understand* OR comprehend* OR comprehension) AND (인식 OR 인지 OR 지각 OR 감지 OR 식별 OR recognit* OR recogniz* OR cognit* OR cogniz* OR awar* OR dicern* OR percept* OR perceiv*)).DSC)

(유효 데이터 추출) 인공지능과 관련이 없는 특허 및 중복 데이터를 제거한 총 612건의 특허 데이터 중 국제특허분류 코드가 2개 이상인 유효 특허 데이터 403건 추출

(분석방법) R*을 활용하여 연관규칙 마이닝 분석 실시 후 기술관계 지도(map)** 활용

* 오픈소스 프로그램으로 통계, 데이터 마이닝 및 그래프를 위한 언어

** 도출된 연관규칙을 한눈에 파악 가능하도록 시각화 한 지도

- 연관성 규칙을 추출하는 R 코드 apriori() 함수를 이용하여 지지도(support) 기준으로 상위 15개의 규칙 분석

(분석결과) 인공지능과 관련하여 기술 간 연관 관계를 제시

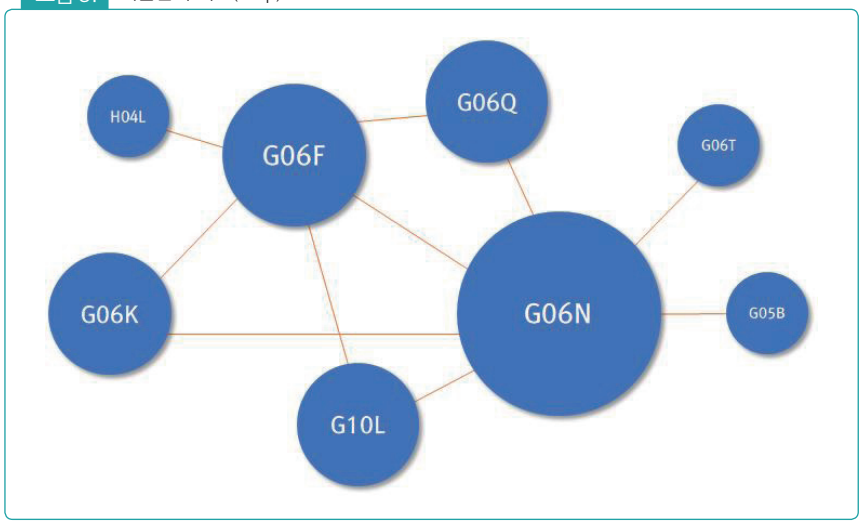
표 6. 지지도 상위 15개 규칙의 도출된 연관 기술

구분	개념
'전기에 의한 디지털 데이터처리'(G06F) ↔ (G06N)'특정 계산모델 방식의 컴퓨터시스템'	0.2123 (21.2%)
'예측용으로 특히 적합한 데이터 처리 시스템 또는 방법'(G06Q) ↔ (G06N)'특정 계산모델 방식의 컴퓨터시스템'	0.0864 (8.6%)
'데이터의 인식'(G06K) ↔ (G06N)'특정 계산모델 방식의 컴퓨터시스템'	0.0716 (7.2%)
'예측용으로 특히 적합한 데이터 처리 시스템 또는 방법'(G06Q) ↔ (G06F)'전기에 의한 디지털 데이터처리'	0.0518 (5.2%)
'음성인식'(G10L) ↔ (G06F)'전기에 의한 디지털 데이터처리'	0.0469 (4.7%)
'음성인식'(G10L) → (G06N)'특정 계산모델 방식의 컴퓨터시스템'	0.0419 (4.2%)
'데이터의 인식'(G06K) → (G06F)'전기에 의한 디지털 데이터처리'	0.0395 (4.0%)
'제어계'(G05B) → (G06N)'특정 계산모델 방식의 컴퓨터시스템'	0.0345 (3.5%)
'이미지 데이터 처리'(G06T) → (G06N)'특정 계산모델 방식의 컴퓨터시스템'	0.0320 (3.2%)
'디지털 정보의 전송'(H04L) → (G06F)'전기에 의한 디지털 데이터처리'	0.0271 (2.7%)

* R 분석결과 재구성

- 전체 데이터에서 '특정 계산모델 방식의 컴퓨터시스템'(G06N)과 '전기에 의한 디지털 데이터처리'(G06F) 기술이 같이 사용된 비율이 21.2%로 기술간 상관관계가 높은 것으로 나타남
- 즉, '특정 계산모델 방식의 컴퓨터시스템'(G06N) 기술과 '전기에 의한 디지털 데이터처리'(G06F) 기술이 가장 많이 연관되어 있는 것을 확인

그림 3. 기술관계 지도(map)



* R 분석결과 재구성

- '특정 계산모델 방식의 컴퓨터시스템'(G06N) 기술은 여러 기술들과 가장 많이 관련되어 있는 것으로 나타남
- '음성 인식'(G10L), '예측용으로 특히 적합한 데이터 처리 시스템 또는 방법'(G06Q), '데이터의 인식'(G06K), '제어계'(G05B), '이미지 데이터 처리'(G06T) 기술과 연관되어 있는 것으로 분석됨

04 결론 및 시사점

- 국제특허분류 코드를 활용하여 기술의 연관성뿐만 아니라 기술 간 선후 관계, 핵심기술 등 파악 가능
  - 국제특허분류 코드에 기반하여 분석 목적에 맞는 다양한 활용 가능
- 특허 데이터의 국제특허분류 코드를 이용하여 인공지능 관련 기술들의 연관성에 대한 기술관계 맵을 제시
  - 인공지능은 인간의 지능을 따라잡기 위해 컴퓨터 시스템을 활용하여 대량의 데이터를 기반으로 학습한다는 점 유추 가능

참고자료

1. Jun, S. (2011). IPC Code Analysis of Patent Documents Using Association Rules and Maps – Patent Analysis of Database Technology, 21–30, Springer-Verlag Berlin Heidelberg, 2011.
2. 전성해, 박상성, 장동식(2014). 특허분석과 기술예측. 서울: 교우사.
3. 특허청·한국발명진흥회 편저(2011). 지식재산의 이해. 서울: 박문각.
4. 곽현, 전성태, 박성혁, 석왕헌(2016). 인공지능(AI) 기술 및 정책 동향(2016 지식재산 이슈페이퍼). 한국지식재산연구원.
5. 나영식, 조재혁(2018). KISTEP 기술동향브리프 인공지능(SW). 한국과학기술기획평가원.

융합연구정책센터
Convergence Research Policy Center
02792 서울특별시 성북구 화랑로 14길 5
Tel. 02-958-4980 Fax. 02-958-4989