

# 우리나라 AI 연구(2016~2023)의 글로벌 경쟁력과 토픽 진화분석

김문구, 안지현\*, 박종현  
한국전자통신연구원, \*동국대학교

mkkim@etri.re.kr, anjh@dgu.ac.kr, stephanos@etri.re.kr

## Global competitiveness and topic evolution in Korean AI research: 2016~2023

Moon-Koo Kim, Jihyun An\*, Jong-Hyun Park  
ETRI, \*Dongguk Univ.

### 요약

인공지능(AI) 기술을 선점하기 위해 글로벌 주요 국가간 치열한 경쟁이 펼쳐지고 있으며, 우리나라도 정책적 우선순위를 부여하고 기술개발과 전문 인력양성, 산업생태계 구축에 투자를 확대하고 있다. 본 연구에서는 우리나라 AI 연구(2016년~2023년)에 초점을 집중하여 글로벌 차원에서 지식생산과 영향의 경쟁력을 파악하고 주제의 변화를 체계적으로 분석하였다. 우리나라는 논문생산과 영향력, 협업비율에서 글로벌 10 위권을 조금 하회하였지만 최근 전반적으로 순위와 지표가 상승하여 10위 이내의 경쟁력을 확보하였다. LDA에 의한 토픽모델링 분석결과, AI 기술발전에 따른 기술적 적용이 우리나라 AI 연구의 핵심 주제를 형성하였으며 시기에 따라 핵심 주제에서 상당한 변화가 나타났다.

### I. 서론

혁신적 기술진화와 빠른 사회적 확산을 바탕으로 인공지능(AI)이 인류의 전 영역을 변화/전환시킬 기술로 부각됨에 따라, 글로벌 주요국가에서는 선제적 연구개발 투자와 생태계 육성에 국가적 자원과 역량을 집중하고 있다[1]. 우리나라 역시 2010년대 중반 이후 국가필수 전략기술로 AI를 선정하고 기술개발과 산업/공공 활용, 전문인력 양성에 막대한 투자를 확대하고 있다[2]. AI 기술과 산업 선점이 글로벌 핵심 이슈로 부각되면서 AI 국가 지수/순위에 대해 다양한 발표가 있으나 국내에 초점을 집중하고 지식생산과 영향, 주제 진화를 분석한 연구는 드문 실정이다. 이에 본 연구에서는 2016년 이후 2023년까지 우리나라의 AI 연구의 글로벌 경쟁력을 체계화하고 핵심주제의 변화를 파악하고자 하였다.

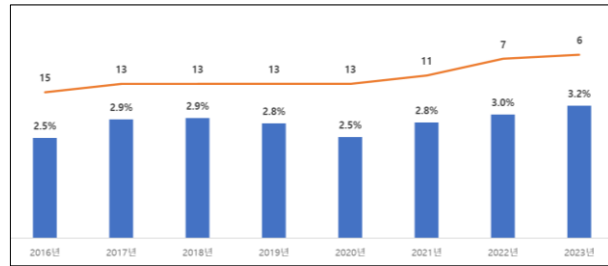
### II. 분석 절차

우선 전문 학술 DB인 Web of Science에서 AI에 해당하는 Research categories 검색(Computer Science, Artificial Intelligence)을 수행하였다. 2016년 이후 2023년까지 영문으로 발표된 논문(article)을 한정하여 검색한 결과, 총 175,694 논문이 검색되었다. 이 가운데 국내 저자의 논문은 5,033 편(2.9%)으로 나타났다. 저자 키워드를 대상으로 오타자 정정, 특수문자 제거, 용어 표준화, 소문자 변환 등으로 전처리 과정을 거쳤으며 LDA(Latent Dirichlet allocation)를 Python 패키지(spacy, Gensim)를 통해 수행하였다.

### III. 연구결과

#### (1) 우리나라의 글로벌 AI 연구 경쟁력 분석

우리나라는 2016년부터 2023년까지 AI 연구에서 논문 점유율이 글로벌 12위였으나(그림 1)과 같이 2022년 이후 최근 3%를 상회하면서 6위로 상승하였다. AI에서 우리나라 논문의 영향력은 <표 1>과 같이 나타났다. 피인용에서 상위 1%에 해당하는 논문들 가운데 국내 논문은 3.7%를 점유하여 전체 10위였으며 상위 비율이 낮아질수록 우리나라 논문의 점유율과 순위도 조금 낮아졌다.

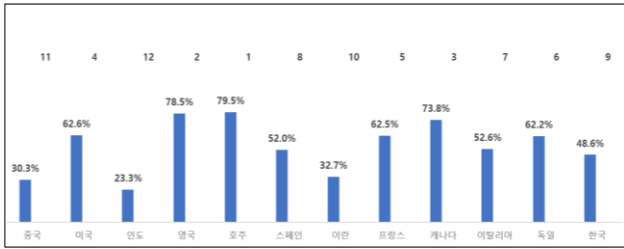


(그림 1) AI 연구에서 우리나라 논문 점유율과 순위

<표 1> AI 연구에서 우리나라 영향력 점유율과 순위

구분	글로벌 논문 수	우리나라 논문 수	우리나라 비율	우리나라 순위
상위 1%	1771	65	3.7%	10
상위 0.5%	885	29	3.3%	12
상위 0.1%	176	5	2.8%	13

우리나라는(그림 2)와 같이 AI 연구에서 다른 국가와의 협업비율은 48.6%(9위)였으며 글로벌 논문 생산 상위 12개국 평균(43.6%)과 비교하면 상대적으로 높았다. 2016년에는 38.6%였으나 2021년에는 56.1%를 거쳐 2023년에는 48.8%로 글로벌 협업이 전반적으로 활성화되는 추세를 보였다.



(그림 2) AI 연구에서 글로벌 협업 비율과 순위

## (2) 우리나라의 AI 연구 핵심주제 진화

우리나라 AI 연구의 핵심주제를 분석하기 위해 <표 2, 3>과 같이 LDA 에 의한 토픽모델링을 수행하였다. 2016 년부터 2019 년까지 총 1,259 개의 키워드가 도출되었으며 10 개의 토픽이 최적인 것으로 나타났다[3]. AI 연구의 핵심 토픽은 Neural network (T1), Human robot interaction (T2), Generative algorithm (T3), Machine learning (T4), Fuzzy set (T5), Object detection (T6), Video code (T7), Image segmentation (T8), Social context analysis (T9), Support vector machine (T10)으로 종합되었다. AI 기술 자체, 이미지/텍스트 처리기법, AI 활용분야에 대한 키워드들이 핵심 연구 토픽이 되었다.

<표 2> 2016 년~2019 년 사이의 AI 연구의 핵심 토픽

토픽명	핵심 키워드	C	P
Neural network	neural_network, deep_learn, imag	0.331	11.14
Human robot interaction	system, human, model, robot, control, estim, mobil, interact, recommend	0.132	10.371
Generative algorithm	optim, model, data, genet_algorithm, cluster, estim, kernel	0.126	10.198
Machine learning	time, process, match, random, machin_learn, featur, pattern, line	0.111	10.399
Fuzzy set	fuzzi, set, filter, approxim, lattic, ideal, soft, oper, order, hesit	0.225	8.617
Object detection	object, track, detect, visual, learn, data, segment, model, base, method	0.258	9.336
Video code	video, comput, code, secur, base, effici, high, hev, softwar, decis	0.156	10.175
Image segmentation	imag, reconstruct, method, edg, comput, color, process, analysi, qualiti, filter	0.164	9.106
Social context analysis	network, detect, analysi, graph, servic, time, data, context, social, comput	0.071	10.722
SVM	featur, recognit, machin, local, face, support_vector, classif, structur	0.239	9.938

\* C: coherence, P: prevalence

2020 년부터 2023 년까지 총 1093 개의 키워드가 도출되었으며 역시 10 개의 토픽이 최적으로 나타났다. Particle Swarm Optimization (T1), Decision making (T2), Neural network (T3), Reinforcement learning (T4), Supervised learning (T5), Feature extraction (T6), Internet of things (T7), Convolutional neural network (T8), Generative adversary network (T9), Image analysis (T10)으로 핵심 토픽이 종합되었다. 이전 시기에 달리 새로운 AI 기술들이 부각되었으며 적용범위도 차별화되는 특징을 보였다.

<표 3> 2020 년~2023 년 사이의 AI 연구의 핵심 토픽

토픽명	핵심 키워드	C	P
PSO	optim, algorithm, comput, base, evolutionari, problem, analysi, process, design, search	0.168	9.852
Decision making	fuzzi, set, track, languag, cluster, system, soft, recommend, embed	0.210	9.378
Neural network	control, system, neural_network, time, fuzzi, delai, stabil, synchron	0.270	11.990
RL	deep_learn, machin_learn, classif, reinforc_learn	0.240	11.204
Supervised learning	detect, object, supervis_learn, memori, term, featur, recognit	0.189	9.182
Feature extraction	data, analysi, model, predict, featur, comput, learn, video, mine, extract	0.142	9.845
IoT	robot, system, function, vehicl, sensor, base, classif, real, internet, interact	0.190	9.041
CNN	convolut_neural_network, graph, deep_neural	0.375	11.507
GAN	learn, network, data, model, gener_adversari_network, train	0.211	9.287
Image analysis	imag, segment, resolut, analysi, detect, time, transform, reconstruct, vision, estim	0.184	8.713

\* C: coherence, P: prevalence

## IV. 결론: 우리나라 AI 연구 활성화 방향 제언

본 연구에서는 인공지능 분야를 중심으로 우리나라 연구 경쟁력을 파악하고 핵심 토픽의 변화를 분석하였다. 연구결과를 바탕으로 우리나라 AI 연구활성화 방향을 제언하면 다음과 같다.

첫째, 우리나라의 AI 연구의 지식생산이나 영향력, 글로벌 협업은 지속하여 증가하고 순위 역시 상승하는 추세가 나타났다. 우리나라가 글로벌 5 위 이내의 연구 경쟁력 확보를 위해 원천-상용 분야의 연구 기반을 강화하며 산학연 협력과 함께 글로벌 주요 국가들과 긴밀한 협업을 확대할 필요가 있다.

둘째, 우리나라 AI 연구의 핵심 토픽을 분석한 결과, AI 기술발전 추세가 적기에 반영되는 것으로 나타났다. 다만, 다양한 분야와의 융합연구나 응용연구의 토픽은 상대적으로 부족하였다. 이에 학제적 연구나 산업/공공 분야의 활용연구를 보다 활성화할 필요가 있다.

끝으로 본 연구에 이어 AI 연구의 핵심 토픽을 다른 국가들 (특히, 미국, 중국, EU)과 비교를 통해 AI 연구주제의 공통성과 함께 국가별 차별성을 분석하는 후속연구가 진행된다면 AI 연구를 심층적으로 이해하고 미래연구 방향을 설정하는데 도움이 될 것이다.

## 참 고 문 헌

- [1] 소프트웨어정책연구소, 미국, 영국, 캐나다의 인공지능(AI) 정책 동향, 2023.
- [2] 과기정통부, 인공지능·첨단바이오 전략로드맵, 2023.
- [3] Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. Multimedia Tools and Applications, 78, 15169-15211.