

## 설명 가능한 인공 지능(XAI)



데이터분석본부 산업시장분석연구팀 책임연구원 **김한국** Tel: 02-3299-6065 e-mail: hgkim712@kisti.re.kr

### KEY FINDING

1. 설명 가능한 인공 지능(XAI)은 기계 학습 모델이 만들어낸 결정에 대한 신뢰성과 투명성을 확보하는 기술이다. 이는 의료, 금융, 보험, 군사 등 다양한 분야의 AI 기술 사용자로부터 신뢰를 얻기 위해 필요한 중요한 요소다.
2. XAI의 세계 시장 규모는 2021년에 44.1억 달러를 기록하였으며, 연평균 19.8 %의 높은 성장률을 유지하며 2026년에는 109.0억 달러를 기록할 것으로 전망되는데, 이는 XAI 기술의 가치와 중요성을 직관적으로 보여준다.
3. 국내 XAI 시장 규모는 2021년에 2,425억 원으로 나타났으며, 빠른 기술 발전과 함께 연평균 20.5 %의 성장률을 기록하며, 2026년에는 6,165억 원의 시장을 형성할 것으로 전망된다.
4. 국내에서는 네이버, LG, 카카오, 이글루코퍼레이션 등이 XAI 시장의 주요 업체로 활약하고 있으며, 전 세계적으로는 구글(Google)이 주요 연구개발을 주도하고 있다.
5. XAI 기술의 발전은 인공 지능의 '블랙박스' 문제를 해결하고, 인공 지능의 투명성과 신뢰성을 높이는 데 기여할 것이고, 또한 다양한 산업에서 인공 지능을 도입하고 활용하는 데 있어 핵심 요소가 될 것이다.

### 1) 시장의 개요

설명 가능한 인공 지능(Explainable AI, XAI)은 인공 지능(AI) 모

델이 입력값(영상, 음성, 문자 등)에서 출력을 생성할 때 주요 근거가 되는 부분을 해당 입력 값에서 추출하여 시각화하거나 문장으로 표현하는 기술이다. 최근의 인공지능 모델들은 대량의 퍼셉트론<sup>1)</sup>을 사

1) 퍼셉트론(perceptron) : 초기 신경망 분야의 연구에 커다란 영향을 미친 퍼셉트론은 본래 1950년대에 Rosenblatt가 제안한 것으로 수용층, 연합층, 반응층의 세 부분으로 구성되어 있다. 퍼셉트론 학습법은 특정 목적에 맞는 정확한 연결 강도에 수렴한다고 입증되었다. 초기 퍼셉트론은 인지가 가능하다는 관점에서 상당한 관심을 모았으나, AND, OR 연산 등 선형 분리가 가능한 문제의 해결에만 사용될 수 있었다.

용해 입력과 출력의 관계를 최대한 근사화할 수 있도록 다양한 계수를 학습한다.

하지만 학습에 사용된 퍼셉트론의 수가 많고, AI 모델을 구성하는 Perceptron들 간의 관계가 복잡하기 때문에 AI 모델의 판단 근거를 수학적으로 판단하기는 어렵다. 이에 대한 해결책으로 설명 가능한 AI 모델은 AI 모델이 학습시 특정 계수 혹은 특정 구조체에 모델의 판단 근거를 투영할 수 있도록 수학적으로 설계하는 것을 목표로 한다.

영상, 음성, 문자 등의 데이터를 다루는 AI 모델에 대해서는 특정 계수에 판단 근거를 투영할 수 있는 주의 기법(Attention Mechanism)을 반영해 모델을 설계한다. 반면에 화학식, 소셜 네트워크 등의 사회 관계망을 다루는 AI 모델에 대해서는 특정 구조체(SubGraph)에 모델의 판단 근거를 투영할 수 있도록 모델을 설계한다.

설명 가능한 AI는 일반적으로 분야별 적용 기법에 따라 4 개로 대별되며, 각 시스템 기술은 다른 종류의 사용자 요구를 처리한다. 이

는 데이터의 특성에 따라 사용되는 AI 모델의 대표적인 구조가 달라지며, 그에 따라 AI의 판단 근거를 추론할 수 있는 알고리즘이 변화하기 때문이다.

현재 도입기의 신생 산업에 속하는 인공지능은 고도의 기술이 필요한 기술 집약적이고, 진입 장벽이 높지만 성장 잠재력이 매우 높아 사회적 및 경제적 파급 효과를 끼칠 수 있는 기술로서 전도유망하다고 할 수 있다.



표 1 설명 가능한 인공지능(AI) 분류

분야	대표적인 인공지능 모델	주로 사용되는 분야	대표적인 설명 알고리즘	설명 알고리즘에 대한 객관적인 평가 지표	설명 알고리즘의 산업적 응용 분야
영상에 대한 설명 가능한 인공지능	Convolutional Neural Network, Faster-RCNN	객체 분류, 객체 검출, 미세 분류 등	CAM, GRAD-CAM	ROAR <sup>2)</sup>	의료 영상 분야
문자열에 대한 설명 가능한 인공지능	Long Short-Term Memory, Gated Recurrent Unit	문자열 번역, Natural Language Inference <sup>3)</sup> 등	Attention Mechanism	미확인	미확인
사회 관계망 데이터에 대한 설명 가능한 인공지능	Graph Neural Network	Graph Classification, Node Classification 등	GNN Explainer, PGM Explainer	미확인	온라인 시장에서 블랙 유저 검거 등
멀티 모달에 대한 설명 가능한 인공지능	Bottom-Up and Top-Down Attention Network	영상 캡셔닝, Visual Question Answering <sup>4)</sup> , Visual CommonSense Reasoning <sup>5)</sup> 등	Constrained Beam Search	VCR만 확인됨	미확인

출처: 한국과학기술정보연구원 작성

2) ROAR(RemOve And Retrain) : 원래 모드의 해석 값이 높은 순서대로 train data의 픽셀 일부를 제거하고 새로운 모델 재교육을 의미함.

3) Natural Language Inference : 자연어 추론, 하나의 문장(premise)이 주어졌을 때 해당 문장이 다른 문장(hypothesis)에 의미적으로 수반(entail)되는지 결정하는 문제임.

4) Visual Question Answering : 이미지(영상으로도 확장 가능)와 그 이미지에 대한 질문이 주어졌을 때, 해당 질문에 맞는 올바른 답변을 만들어내는 것임.

5) CommonSense Reasoning : 특정 상황에 대한 이미지와 이에 대한 질문이 주어지면, 정확한 대답과 함께 정당한 근거를 제시해야 하는 것임.

표 2 산업적 측면에서 설명 가능한 인공지능의 특징

특징	내용
도입기 산업	• 설명 가능한 인공지능은 기존 인공지능 산업의 빠른 확산에 따른 예상치 못한 사회적 이슈·우려가 대두되어 인공지능의 채택은 극대화하면서 위험과 부작용은 최소화하기 위해 대두된 것으로 도입기 산업임.
기술 집약적 산업	• 설명 가능한 인공지능을 구현하기 위해서는 데이터의 추상화, 고성능 컴퓨팅 능력, 빅데이터 기술, 설명 가능한 모델 구현 등 고도의 기술이 필요한 기술 집약적 산업임.
진입 장벽이 높은 산업	• 인공지능의 성장 잠재력은 높으나 고도의 기술력과 많은 자본력이 필요한 진입장벽이 높은 산업임.
파급 효과가 큰 산업	• 인공지능 산업은 법률, 금융, 의료, 자율주행차, 스마트 팩토리, 빅데이터 분석 등 산업 다방면에 적용이 가능하며 이로 인한 사회적/경제적 파급효과가 무척 큰 산업임.
대체재로부터의 위험이 낮은 산업	• 인공지능 기술을 직접 또는 간접적으로 대체할 수 있는 대체재는 없으므로 대체재의 위험은 낮은 산업임.

\*출처: 한국과학기술정보연구원 작성

## 2) 정책 및 규제 현황

최근 각국에서 인공지능 기술에 대한 투명성과 책임성을 높이기 위한 다양한 정책과 제도를 마련하고 있는데, 그 대안으로 설명 가능한 인공지능을 통해 사용자들이 AI의 결정 과정을 이해하고, 필요한 경우 적절한 조치를 취할 수 있도록 하는 방안을 모색하고 있다.

유럽연합(EU) 각료 의사회는 2023년 6월 유럽연합 의회가 채택한 AI 법안(EU AIA)의 수정안을 검토하는 중이며 2023년 말까지 승인할 예정이다. 수정안은 인공지능을 '허용할 수 없는 위험', '고위험', '제한된 위험', '저위험 또는 최소위험'의 4단계 위험도로 분류하고, 이 중 '저위험 또는 최소 위험'을 제외한 나머지 3단계에 대해 단계별 규제를 부과한다. 특히, '범용 인공지능에 대한 의무' 항목에 챗지피티(ChatGPT)와 같은 모델을 기반으로 하는 생성형 인공지능 시스템의 투명성 요구 사항과 학습에 사용된 저작권이 있는 데이터의 자세한 요약도 공개해야 한다는 내용이 포함되어 있는데, 설명 가능한 인공지능의 특성이 이러한 규제를 준수할 방안으로 기대되고 있다.

우리나라는 인공지능 산업 육성과 관련 생태계 진흥에 초점을 맞춰 제도 정비에 나서고 있는데, 미국이나 중국 등이 인공지능 분야에서 기술 우위를 점하고 있는 상황에서 자칫 규제 중심으로 법·제도를 정비하면 '인공지능 주권'이 위협받을 수 있을 것이라는 인식이 존재한다. '인공지능 산업 육성 및 신뢰 기반 조성 등에 관한 법률안'

에서 인공지능 기술 연구 개발, 인공지능 서비스 출시를 우선 허용하고 있으며, 문제가 생기는 경우에 한해 규제를 하고 있는 것이 예이다. 그러나 미국과 유럽 연합에서 인공지능에 대한 규제 필요성이 대두되고 있는 만큼, 우리나라도 인공지능 윤리 및 신뢰성 강화, 고위험 인공지능에 대한 규제 필요성 등에 대한 논의가 이어지고 있다.

## 3) 시장동향

### | 시장 규모 및 전망

설명 가능한 인공지능의 세계 시장 규모는 2021년에 44.1억 달러의 규모로 나타났다. 관련 시장은 연평균 19.8 %의 성장률을 보이며 2026년에는 109.0억 달러의 시장 규모를 형성할 것으로 전망된다.

설명 가능한 인공지능의 국내 시장 규모는 2021년에 2,425억 원의 규모로 나타났다. 또한 관련 시장은 연평균 20.5 %의 지속 성장으로 2026년에는 6,165억 원의 시장 규모를 형성할 것으로 전망된다.

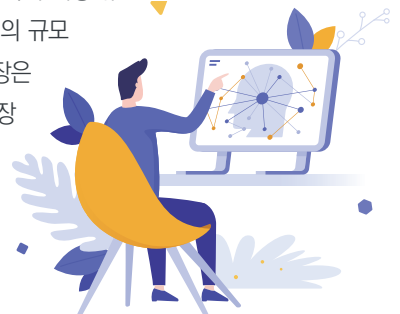
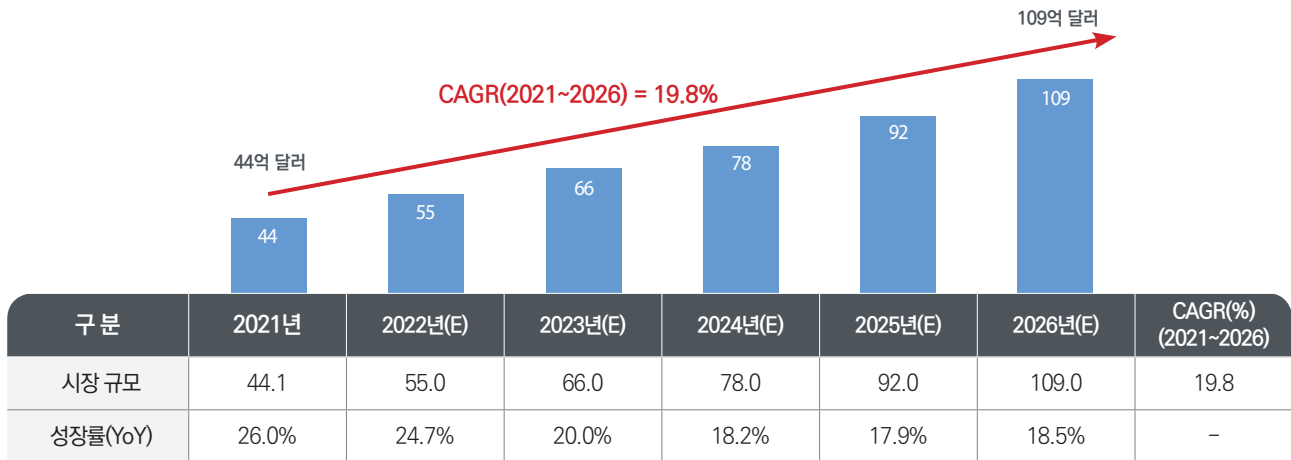


표 3 설명 가능한 인공지능의 세계 시장 규모

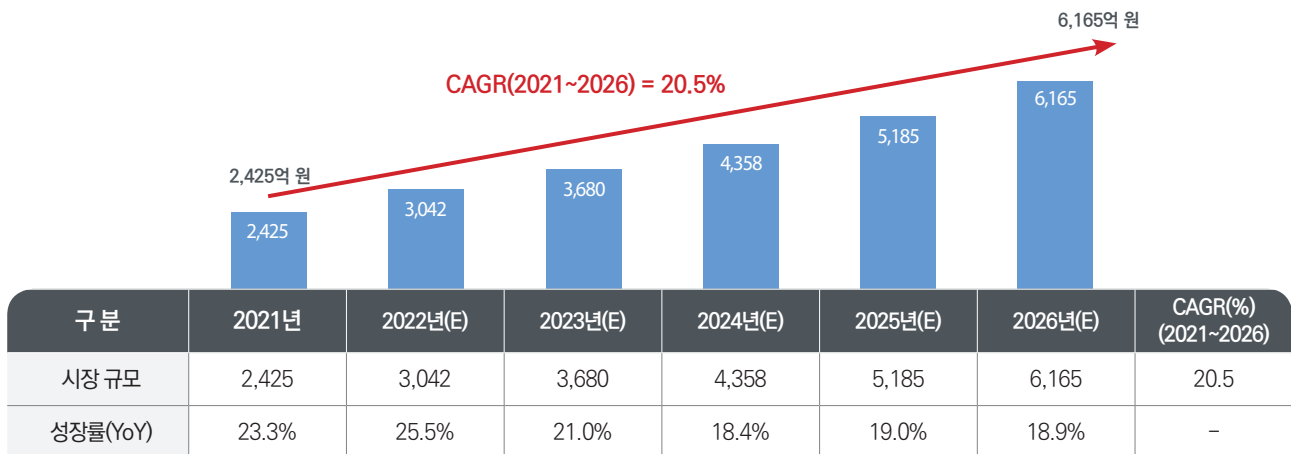
(단위 : 억 달러)



출처: "Explainable AI Market, Forecast, 2021 - 2030", Next Move Strategy Consulting, 2022, "Artificial Intelligence Market - Global forecast to 2026", MarketsandMarkets, 2021

표 4 설명 가능한 인공지능의 국내 시장 규모

(단위 : 억 원)



출처: "Explainable AI Market, Forecast, 2021 - 2030", Next Move Strategy Consulting, 2022, "Artificial Intelligence Market - Global forecast to 2026", MarketsandMarkets, 2021

\* 산출근거: 세계 시장규모 × 한국 연도별 인공지능 시장 점유율(MarketsandMarkets, 2021)

\* 환율(KRW/USD): ('21~)1,144.42

## 경쟁 현황

해외의 주요 설명 가능한 인공지능 업체는 구글(Google), 마이크로소프트(Microsoft), 애플(Apple), 아마존(Amazon), 아이비엠(IBM), 그리고 엔비디아(Nvidia) 등이며, 모두 미국에 위치하고 있다.

구글(Google)은 검색 엔진, 클라우드 컴퓨팅, 소프트웨어 등을 중심으로 설명 가능한 인공지능 기술을 개발하고 있으며, 여러 자

회사를 통해 '설명 가능한 인공지능(Explainable AI)'과 '모델 서치(Model Search)' 등의 플랫폼을 선보였다.

마이크로소프트(Microsoft)는 설명 가능한 인공지능 툴킷인 페어러른(Fairlearn)과 뉴앙스 커뮤니케이션즈(Nuance Communications)의 인수를 통해 다양한 산업에 대화형 인공지능 및 주변 인공지능을 제공하고 있다.

애플(Apple)은 AI 연구를 비공개로 진행하며, 다양한 AI 기업을

인수해 자체 제품의 인공지능 기능을 강화하고 있다.

아마존(Amazon)은 전자 상거래와 클라우드 컴퓨팅을 중심으로 활동하며, 머신 러닝 추론의 최종 검증을 가능케 하는 아마존 증강 인공지능 'Amazon Augmented Artificial Intelligence (A2I)'를 선보였다.

아이비엠(IBM)은 AI 페어니스 360(AI Fairness 360)을 통해 설

명 가능한 인공지능 개발에 주력하고 있으며, 머신 러닝 모델의 차별과 편향을 조사, 보고, 완화하는데 중점을 두고 있다.

엔비디아(Nvidia)는 그래픽 처리 장치, 인공지능 하드웨어 및 소프트웨어 공급에 주력하며, AI 처리를 위한 공격적인 투자를 통해 미국의 멜라노스(Mellanox)와 영국의 암(ARM)을 인수하였다.

표 5 해외 설명 가능한 인공지능 업체

업 체	현 황
구글(Google) (미국)	<ul style="list-style-type: none"> <li>구글(Google)은 1998년 설립되어 검색 엔진 기술, 온라인 광고, 클라우드 컴퓨팅, 컴퓨터 소프트웨어, 인공지능에 중점을 둔 미국의 다국적 기술 회사임.</li> <li>동사는 TensorFlow, X Development(구 구글(Google) X), Deepmind Technologies 및 Waymo와 같은 다양한 자회사를 통해 인공지능 개발에 중점을 둠.</li> <li>2019년 데이터와 모델 사이의 편향, 드리프트 등의 차이를 감지하고 어떤 특징이 모델에 예측 결과에 얼마나 영향을 미쳤는지 알 수 있는 'Explainable AI'를 발표함.</li> <li>2021년 딥 러닝 모델을 설계하기 위한 Automated Machine Learning 플랫폼 'Model Search'를 오픈 소스로 공개함.</li> </ul>
마이크로소프트 (Microsoft) (미국)	<ul style="list-style-type: none"> <li>마이크로소프트(Microsoft)는 대화형 시각화 대시보드와 불공평 완화 알고리즘으로 구성되어 개발자가 AI 시스템 공정성을 높일 수 있는 설명 가능한 인공지능 툴킷인 Fairlearn을 개발해 공개함.</li> <li>2022년 의료, 금융 등 산업 전반에 걸쳐 대화형 AI 및 주변 인공지능 분야의 선두업체인 Nuance Communications를 인수함.</li> </ul>
애플(Apple) (미국)	<ul style="list-style-type: none"> <li>애플(Apple)은 1976년 미국에 설립된 소비자 전자 제품, 소프트웨어 및 온라인 서비스를 제공하는 다국적 기술 회사임.</li> <li>구글(Google)이 AI 연구에 오픈 소스 기반의 공동 프로그램에 참여하고 선도하는 것에 반해 동사는 대부분의 연구를 비공개로 진행하고 있으며 아이폰, 아이패드 등에 머신 러닝을 추가하지만 상세 스펙을 공개하지 않음.</li> <li>영국의 시장조사기관 글로벌데이터에 따르면, 2016년부터 2020년까지 60 개의 AI 기술 기업 중 25 개사를 인수해 가장 많은 AI 기업을 인수한 것으로 밝혀짐.</li> </ul>
아마존(Amazon) (미국)	<ul style="list-style-type: none"> <li>아마존(Amazon)은 1994년 미국에 설립된 전자 상거래, 클라우드 컴퓨팅, 인공지능에 초점을 둔 미국의 다국적 기술 회사임.</li> <li>2020년 머신 러닝의 낮은 신뢰도 예측을 지속적으로 식별하고 개선해 머신 러닝 추론에 사람의 판단이 필요한 상황에서 확인을 통해 추론을 최종 검증할 수 있는 완전 관리형 서비스 아마존 증강 인공지능(Amazon Augmented Artificial Intelligence, A2I)을 발표함.</li> </ul>
아이비엠(IBM) (미국)	<ul style="list-style-type: none"> <li>아이비엠(IBM)은 1911년 미국에 설립된 컴퓨터 하드웨어, 미들웨어 및 소프트웨어를 생산 및 판매하는 글로벌 업체임.</li> <li>설명 가능한 인공지능을 위해 AI 애플리케이션 수명 주기 전반에 걸쳐 머신 러닝 모델의 차별과 편향을 조사, 보고 및 완화할 수 있는 AI Fairness 360을 개발함.</li> </ul>
엔비디아(Nvidia) (미국)	<ul style="list-style-type: none"> <li>엔비디아(Nvidia)는 1993년 미국에 설립된 고성능 컴퓨팅을 위한 그래픽 처리 장치(GPU), 응용 프로그래밍 인터페이스(API), 인공지능 하드웨어 및 소프트웨어 공급 업체임.</li> <li>2019년 고성능 네트워킹 기술을 가진 미국의 멜라노스(Mellanox)와 2020년 영국의 반도체 설계회사 암(ARM)을 각각 69억 달러와 400억 달러에 인수하여 AI 처리를 위해 공격적인 투자를 감행함.</li> </ul>

출처: 각사 홈페이지, 한국과학기술정보연구원 재구성

국내 설명 가능한 인공지능 시장의 주요 업체로는 네이버, LG, 카카오, 이글루코퍼레이션 등이 있다.

네이버는 한국 최대 검색 포털을 운영하며, '네이버 AI랩'을 설립해 인공지능 기술의 경쟁력을 강화하고 있다.

카카오는 다양한 모바일 및 인터넷 서비스를 제공하며, 자체 개발한 인공지능 엔진을 통해 다양한 비즈니스 활동을 진행하고 있다.

LG는 전자, 화학, 통신 제품 생산을 중심으로 활동하며, LG AI연구원을 통해 인공지능 전문 인력을 양성하고 설명 가능한 인공지능 기술 연구를 본격화하였다.

이글루코퍼레이션은 설명 가능한 인공지능에 기반한 AI 특허를 보유하고 있으며, 파이오링크 인수를 통해 네트워크 및 클라우드 보안 분야까지 확장하였다.

표 6 국내 업체 현황

업 체	현 황
네이버	<ul style="list-style-type: none"> <li>• 네이버는 1999년 설립되어 한국 최대 검색포털 네이버를 서비스하고 있으며, 인공지능, 로봇틱스, 모빌리티 등 미래 기술을 지속적으로 개발하고 있음.</li> <li>• 동사는 2020년 인공지능 기술 경쟁력 강화를 위해 중장기 AI 기술연구소 '네이버 AI랩'을 신설함.</li> <li>• 2022년 세계적으로 권위 있는 '표현학습국제학회(ICLR) 2022'에서 동사의 컨볼루션, 심층신경망, 강결합 연구 등 관련 논문 17개가 채택됨.</li> <li>• 현재 동사는 인공지능 전문회사 설립 추진을 검토 중에 있음.</li> </ul>
엘지(LG)	<ul style="list-style-type: none"> <li>• LG는 1947년 설립되어 전자, 화학, 통신 제품을 생산하는 글로벌 기업임.</li> <li>• 동사는 2020년 LG AI 연구원을 출범한 뒤 오는 2023년까지 2,000억원을 투자하여 그룹 AI 전문가 1,000명을 양성할 계획임.</li> <li>• LG AI 연구원은 토론토대 Konstantinos Plataniotis 교수팀과 공동으로 설명가능한 인공지능 기술 연구에 매진하여 인공지능학회(AAAI)에서 '설명가능한 AI'와 '연속학습' 연구에 대한 성과를 발표함.</li> </ul>
카카오	<ul style="list-style-type: none"> <li>• 카카오는 1995년 설립된 다양한 모바일 및 인터넷 서비스를 제공하는 IT 기업임.</li> <li>• 동사는 2017년 인공지능 연구전문회사인 카카오엔터프라이즈와 카카오브레인이 엔진들을 개발하고 비즈니스 서비스 플랫폼으로 카카오톡을 공개함.</li> <li>• 동사의 자회사 카카오브레인은 2022년 헬스케어 분야내 '초거대 인공지능 모델' 개발에 본격 착수함.</li> </ul>
이글루코퍼레이션	<ul style="list-style-type: none"> <li>• 이글루코퍼레이션은 2021년 자사의 AI 보안관계 솔루션에 적용할 설명가능한 인공지능에 기반한 AI특허 2건을 등록함.</li> <li>• 동사는 2021년 애플리케이션 전송 컨트롤러(ADC) 전문 업체인 파이오링크를 인수하여 네트워크·클라우드 보안 분야까지 사업을 확장함.</li> </ul>

출처: 각사 홈페이지 참조 한국과학기술정보연구원 재구성

## 4) 분석자 인사이트

XAI는 인공지능의 판단 근거와 신뢰성을 평가하는 핵심 기술로서 의료, 금융, 보험, 군사 등의 중요한 분야에서 강력한 도구로 인식되게 만들었다. 그러나 이러한 장점에도 불구하고, 기존의 XAI 모델은 서로 다른 기법 간의 객관적인 성능 비교가 어려웠다.

구글(Google)은 이 문제를 해결하기 위해 영상 데이터에 대한 인공지능 모델에 한하여 서로 다른 설명 기법이 생성해 내는 판단 근거가 실제 모델 출력값 생성에 미치는 영향을 파악하는 알고리즘을 제안하였다. 하지만 이 알고리즘은 메모리 구조를 이용하지 않는 경

우에만 사용 가능하며, 문자열, 멀티 모달 데이터, 시계열 데이터를 포함하는 사회 관계망 데이터에 대해서는 적용이 제한적이다.


그럼에도 불구하고 인공지능 모델은 점차 트랜스포머(Transformer)라는 특수한 구조로 대체되고 있다. 이는 인공지능 모델의 일련의 처리 과정을 단순화하며 성능을 향상시킨다. 트랜스포머(Transformer) 모델은 내부의 처리 과정이 더욱 복잡해졌기 때문에 이를 이해하고 설명할 수 있는 XAI의 연구가 필요하게 되었다. 이와 같은 변화와 수요는 XAI 기술의 발전을 촉진시키고 있다.

그리고 시장 규모, 성장성, 장단기 시장 전망, 수익성, 경쟁 강도, 수명 주기 및 시장 촉진·저해 요인을 종합적으로 검토한 결과 XAI의



시장 매력도는 '양호' 수준이다. XAI 기술의 필요성과 이에 따른 시장 수요는 지속적으로 증가하고 있으며, 이는 해당 기술의 중요성을 더욱 부각시킨다.

총체적으로 보면, XAI는 인공지능 기술의 발전과 사회적 수용 가능성을 동시에 추진하는 중요한 역할을 담당하고 있다. 이 기술의 진

화와 확산은 인공지능의 투명성과 신뢰성을 높이는 데 크게 기여하며, 이는 더욱 많은 분야에서 인공지능의 적용을 가능하게 한다. 따라서 XAI는 단순히 기술적인 발전을 넘어 사회적 변화와 발전에 기여하는 중요한 장치로서의 역할을 수행하고 있다. 



www.astinet.kr  
에서 원문을 다운로드  
받으실 수 있습니다.

# ASTI MARKET INSIGHT



**본원** (우)34141 대전광역시 유성구 대학로 245 한국과학기술정보연구원  
T. 042) 869-1004, 1234 F. 042) 869-1091

**분원** (우)02456 서울특별시 동대문구 회기로 66 한국과학기술정보연구원  
T. 02) 3299-6114 F. 02) 3299-6244

