

2020-12호

AI TREND WATCH

2020. 8. 15.

클라우드 기반 AI에 대한 엣지
AI의 도전과 영향

AI전략센터 김민진 전문연구원



정보통신정책연구원
KOREA INFORMATION SOCIETY DEVELOPMENT INSTITUTE

클라우드 기반 AI에 대한 엣지 AI의 도전과 영향

정보통신정책연구원 김민진 전문연구원

개요

- ◆ AI 알고리즘이 대량의 데이터를 처리하고 다량의 컴퓨팅 리소스를 소비함에 따라 소프트웨어 공급자 중심으로 클라우드 기반 AI가 발전했으나(Cloud-based AI), 최근에는 하드웨어 공급자 중심으로 디바이스에서의 인공지능인 엣지 AI(Edge AI)에 대한 관심이 증가
- ◆ 클라우드 기반 AI와 엣지 AI의 특징을 살펴보고 엣지 AI의 영향과 역할을 전망하고자 함

주요 내용

- ◆ (엣지 컴퓨팅 기반의 AI 시장 부상) 전통적으로 빅데이터 · 서비스로 대표되는 클라우드 기반 AI가 인공지능 시장 성장을 주도하였으나, 최근 디바이스에 대한 의존도가 증가하는 사회 · 문화적인 환경 변화에 따라 하드웨어 공급자 중심의 엣지 AI(Edge AI)가 주목 받고 있음
 - ▶ (컴퓨팅 기술) 클라우드 컴퓨팅, 포그(Fog) 컴퓨팅, 엣지 컴퓨팅으로 구분할 수 있음
 - (클라우드 컴퓨팅) 네트워크로 연결된 다수의 컴퓨팅 자원을 이용하여 데이터를 저장 및 처리하는 기술¹⁾
 - (장점) 방대한 양의 데이터를 관리할 수 있는 효율적인 솔루션 제공
 - (한계) 정보 보안 및 개인 정보 보호 이슈 발생 가능, 대량의 데이터를 전송하는 데 비용이 많이 들고 대기 시간 문제 및 네트워크 병목 현상 발생 가능
 - (포그 컴퓨팅) 데이터를 멀리 있는 클라우드까지 보낼 필요 없이 가까운 포그에서 처리 하는 것으로, 중앙 집중식 클라우드 방식의 한계를 보완하는 기술²⁾
 - 안전하고 빠른 링크를 제공하여 데이터 양을 줄이면서 의미있는 정보를 클라우드로 전송
 - 클라우드 기술을 따르지만 데이터 분석, 의사 결정 및 제어를 분산하여 로컬의 자율성 향상

1) 아이티데일리(2017. 7. 31.), “[전문가 강좌] 4차 산업혁명과 지능정보기술의 미래방향(7)”

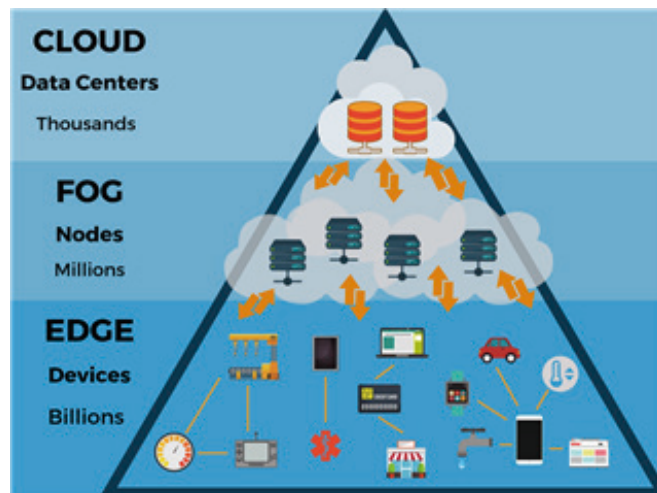
2) Queiroz, J. et. al.(2019), “Distributing Intelligence among Cloud, Fog and Edge in Industrial Cyber-physical Systems”.

- (엣지 컴퓨팅) 사용자 또는 데이터 소스의 물리적인 위치나 그 위치와 가까운 곳(Edge)에서 컴퓨팅을 수행(데이터 처리, 수집, 분석)하는 것

- (장점) 데이터 익명화에 도움을 주며, 신속한 데이터 처리 가능
- (한계) 컴퓨팅 파워에 한계가 있으며, 디바이스를 네트워크에 연결하려면 추가 보안 절차 필요

※ 포그가 클라우드와 디바이스 사이의 중간 계층을 구성하는 정의를 따르는 반면 엣지는 임베디드 처리 기능이 있는 IoT 기반 디바이스라고 볼 수 있음

[그림 1] Cloud, Fog, Edge 컴퓨팅



출처: thinkbiz.net(2019. 2. 12.)

▶ (컴퓨팅 기술 기준 AI 시장) 전통적으로 소프트웨어 공급자 중심의 클라우드 기반 AI(Cloud-based AI) 시장이 성장해왔으나 최근 하드웨어 공급자 주도의 엣지 AI(Edge AI) 시장이 부상

- (클라우드 기반 AI 중심 시장 성장) AI는 방대한 양의 데이터를 저장, 분석, 처리해야 하므로 클라우드와 같은 거대한 인프라는 필수로 인식되어 왔으며, 향후에도 클라우드 기반 AI 성장이 가속화될 것으로 전망

* AI를 선도하는 글로벌 IT기업은 대부분 클라우드 서비스 리딩 기업이라는 공통점 보유

* * '20년까지 AI 기술을 채택한 기업 중 87%는 클라우드 기반 소프트웨어를 통해 AI 역량 확보, 83%는 클라우드 기반 개발 서비스를 사용하여 AI 애플리케이션을 만들 것으로 전망(Deloitte, 2018)³⁾

- (엣지 AI 부상) 소비자들의 디바이스에 대한 의존도가 높아짐에 따라 기업은 더 많은 서비스를 제공하기 위해 디바이스에 필수 컴퓨팅을 가져와야 할 필요성이 대두되면서 주목 받고 있음

3) Jeff Loucks(2018. 12. 11.), "Artificial intelligence: From expert-only to everywhere".

* 현재 데이터의 91%는 중앙 집중식 데이터 센터에서 처리되고 있으나 '22년까지 전체 데이터의 약 74%가 엣지에서 분석 및 처리될 것이며, '22년까지 출하되는 전체 스마트폰의 80%에 온디바이스(On-device) AI가 탑재될 것으로 예상(Gartner, 2018)⁴⁾

- ▶ 클라우드 기반 AI와 엣지 AI 각각의 개념, 장점 및 한계점을 알아보고 AI 시장의 구도 변화를 전망하고자 함

◆ (클라우드 기반 AI) 효율적인 솔루션을 제공하나 보안 위험이 있고 시간과 비용이 많이 소요

- ▶ (개요) 개발자가 AI 또는 기계학습을 응용프로그램에 통합할 수 있도록 지원하는 모든 유형의 클라우드 서비스로 Alaas가 대표적 사례
 - AI 알고리즘이 대량의 데이터를 처리하고 다량의 컴퓨팅 리소스를 소비하므로 소프트웨어 공급자 중심으로 클라우드 기반 AI가 발전
 - 클라우드 개발자는 이미지 인식, 빅데이터 분석에 이르기까지 다양한 AI 도구를 제공
 - (주요 기업) Alphabet(Google), Apple, Amazon, IBM, CognitiveScale, Microsoft, Intel, SAP, Salesforce, Fair Isaac 등
 - (대표적 사례) 아마존웹서비스, 마이크로소프트 애저, 구글 클라우드와 같은 클라우드 기반 Alaas(서비스형 AI)
- ▶ (장점) 방대한 양의 데이터를 관리할 수 있는 효율적인 솔루션을 제공하여 사용자가 AI를 쉽게 적용할 수 있도록 지원
 - (서비스 보편성 확보) 중앙 서버에서 서비스를 제공하므로 사용자는 디바이스의 제약에서 벗어나 클라우드 서비스를 이용할 수 있으며, 이러한 특징은 강력한 컴퓨팅 파워를 필요로 하는 AI 서비스에서 중요
 - (서비스 접근성 향상) 네트워크만 연결되어 있다면 해당 서비스에 접근 가능
 - (AI 구현 가능성 제고) 클라우드로부터 AI 개발에 필요한 서비스를 받아 구현의 난이도를 낮추어 주어 기업 간 AI 격차 해소
 - (운영 효율성 향상) 사용한 만큼 비용을 지불(Pay as you go)하게 되므로 이용자 관점에서 낭비 요소 제거
- ▶ (한계점) 정보 보안, 개인 정보 보호, 시간 및 비용 증가, 통신 인프라에의 높은 의존성 등이 비즈니스 적용에의 한계로 작용
 - (보안 및 개인 정보 보호 이슈) 애플리케이션 개발자는 사실상 데이터의 수집이 용이하지 않은 경우 존재

4) Gartner(2018. 10. 3.), "What Edge Computing Means for Infrastructure and Operations Leaders".

- 클라우드로 전송하는 정보가 개인 식별 가능한 데이터를 포함하고 있는지, 개인 정보를 저장하는 것이 개인 정보 보호법을 위반하는지 여부를 고려해야 함
- (고비용·긴 대기시간 소요) 대량의 데이터 전송 시 많은 비용이 들고 긴 대기 시간이 발생하며 데이터 양이 급격히 증가함에 따라 이 문제는 더 심화될 전망
- (높은 통신 인프라 의존성) 통신 인프라가 부족하거나 관련 작업 및 정보가 민감하여 네트워크 연결이 제한되는 환경일 경우 AI 애플리케이션 접근이 제한될 위험

◆ (엣지 AI) 소량의 데이터로 디바이스에서 독립적으로 AI를 실행할 수 있어, 보안이나 정보 유출 위험 없이 즉각적으로 AI 서비스를 제공할 수 있다는 점에서 클라우드 기반 AI의 한계점을 보완할 수 있는 기술로 주목 받기 시작

▶ (개요) 원격 서버가 아닌 로컬 컴퓨터 또는 임베디드 시스템에서 기계학습 알고리즘을 실행하는 것으로, 디바이스 자체에서 인공지능 서비스가 구현되는 온디바이스 AI(On-device AI)가 대표적 사례

- 알고리즘이 강력하지는 않지만 디바이스에서 네트워크 연결 없이 데이터를 처리하고 독립적으로 결정을 내릴 수 있음
- 인공지능을 엣지에서 구현하려는 흐름은 스마트워치, 스마트폰과 같은 스마트기기, 센서 등의 빠른 성장에 기인

* IBM은 현재 약 150억 개의 지능형 디바이스가 있다고 추정하였으며, IDC는 2025년까지 이러한 지능형 디바이스가 약 1,500억 개로 늘어날 것으로 예측⁵⁾

- 최근 코로나19 대응에 있어서도 엣지 AI 소프트웨어의 활약 예상

* 엣지 AI를 활용하여 공공장소를 모니터링⁶⁾하고 접촉 추적 프로그램을 구현, 잠재적으로 코로나 바이러스의 확산을 늦추는 데 도움을 주는 비디오 감시 애플리케이션이 향후 2년 동안 가장 큰 성장을 보일 것으로 예상(Markets and Markets, 2020. 5)⁷⁾

- (주요 기업) 소프트웨어 분야와 하드웨어 분야에서의 시장 주도 기업은 다음과 같음

- (소프트웨어) IBM, Google, Microsoft, Amazon, Invision.AI 등
- (하드웨어) Microsoft, 삼성전자, 화웨이, NVIDIA 등

- (대표적 사례) 스마트폰, 자율주행차, 지능형 스피커 등에서 자체적으로 AI 서비스가 구현되는 온디바이스 AI

5) ciokorea(2020. 8. 6.), “기대반 우려반... 클라우드를 능가할 엣지 시대는 어떠한가”.

6) 실제로 중국에서는 산업용 엣지 컴퓨터를 탑재한 5G 경찰로봇이 중국 주요 공항과 쇼핑몰 현장에 투입돼 운영, AI 로봇이 수많은 유동인구를 대상으로 신속하게 마스크 착용 여부나 체온을 확인(출처: 전자신문, 2020. 4. 1.).

7) markets and markets(2020. 5), “Covid-19 Impact On Edge AI Software Market”.

- ▶ (장점) 충분한 데이터 수집이 어렵고 컴퓨팅 능력도 부족한 환경에서 개인정보 유출 걱정 없이 AI를 원활히 지원
 - (보안·개인정보 보호 강화) 데이터 수집과 활용에 따른 민감 정보 유출 논란, 빅브라더 출현 위험을 경감시킬 수 있음
 - 디바이스는 AI를 기반으로 한 맞춤형 기능을 제공하면서도 이용자 데이터가 어떻게 수집되는지 혹은 의사결정의 근거는 무엇인지 이해시킬 수 있음
 - 필요한 데이터만 선별적으로 데이터 센터로 전송 가능
 - (지연시간 감소) 클라우드 기반 AI에서 필연적으로 발생하는 데이터 전송 시간을 단축하여 효율성을 제고하고 보다 빠른 의사결정 가능
 - (오프라인 가용성 확보) 네트워크 연결 없이 안정적인 AI 서비스 제공 가능
 - (클라우드 이용 비용 및 네트워크 연결 비용 절감) 클라우드 서비스 비용 및 대역폭과 셀룰러 데이터 비용 절감
- ▶ (한계점) AI가 구현되는 디바이스의 사양에 따른 한계가 존재
 - (성능 제약) 클라우드 대비 처리 능력이 미흡
 - (모델 업데이트 시 연결 필요) 훈련(training)은 디바이스에서 온전히 수행하기에 한계가 있어 모델 개선을 위해서는 연결이 필요할 수 있음
 - (구동 제약) 디바이스에서의 배터리 수명과 열손실 문제 발생 가능

〈표 1〉 클라우드 기반 AI와 엣지 AI 비교

구분	클라우드 기반 AI	엣지 AI
개념 및 특징	• 개발자가 AI 또는 기계학습을 응용 프로그램에 통합할 수 있도록 지원하는 모든 유형의 클라우드 서비스 • 소프트웨어 공급자 중심	• 로컬 컴퓨터 또는 임베디드 시스템에서 기계학습 알고리즘을 실행 • 하드웨어 공급자 주도
장점	• 서비스 보편성 확보 • 서비스 접근성 향상 • AI 구현 가능성 제고 • 운영 효율성 향상	• 보안·개인정보 보호 강화 • 지연시간 감소 • 오프라인 가용성 확보 • 클라우드 이용 비용, 네트워크 연결 비용 절감
한계점	• 보안·개인정보 보호 이슈 • 고비용·긴 대기시간 소요 • 높은 통신 인프라 의존성	• 성능 제약 • 모델 업데이트 시 네트워크 연결 필요 • 구동 제약

◆ (AI 시장 구도 변화 전망) 엣지 AI 시장의 급속한 성장이 전망되고 있으나 이것이 인공지능 시장 발전을 주도할 것인지에 대해서는 엇갈린 전망 존재

▶ (엣지 AI 시장의 성장) 엣지 AI 소프트웨어 및 하드웨어 시장은 각각 연평균 25%, 21% 성장이 전망 됨

- (엣지 AI 소프트웨어 시장) 조사기관마다 차이가 있으나 '18년~'24년까지 엣지 AI Solution&Service 시장은 세계적으로 연평균 약 25% 이상의 성장률을 기록하여 11억 달러 수준의 시장 규모가 형성될 것으로 예상

* '18년 3억5,600만 달러에서 '23년 11억5,200만 달러로 성장('18년~'23년 CAGR: 26.5%, Markets and Markets, 2019. 2)⁸⁾

* '19년 3억4,650만 달러에서 '24년 10억8,770만 달러로 성장('19년~'24년 CAGR: 25.7%, 360 Research Reports, 2020. 5)⁹⁾

- (엣지 AI 하드웨어 시장) 연평균 약 21%대의 성장을 보일 것으로 예상

* '18년 4억2,334만 달러에서 '26년 19억2,921만 달러로 성장('18년~'26년 CAGR: 20.9%, Fior Markets, 2020. 1)¹⁰⁾

▶ (엣지 AI 중심으로의 이동 가능성) 지속적으로 클라우드 기반 AI 중심의 AI 시장 성장이 예상되나, 하드웨어 측면에서는 엣지 AI가 추론용 AI 칩 시장에서 큰 비중을 차지하게 될 것

* Fortune Business Insight는 클라우드 기반 AI 서비스를 채택하는 기업의 증가가 글로벌 AI 시장 성장의 주요 요인 중 하나라고 언급¹¹⁾

- (소프트웨어) AI 소프트웨어 시장에서 엣지 AI 비중은 5% 미만

* 세계 AI 소프트웨어 시장 규모: '18년 100억 1천만 달러에서 '25년 1,260억 달러 예상(tractica, 2020. 1)¹²⁾

- (하드웨어: AI Chip) 클라우드 기반 AI가 전반적인 AI 칩 시장을 선도할 것으로 예상됨. 다만 추론용 반도체 시장에서는 엣지 AI 시장 규모가 클라우드 기반 AI를 넘어설 것

• (학습용) 클라우드 기반 AI 칩 우위 유지

* 학습용 시장에서 클라우드 기반 AI 반도체가 꾸준히 주도권 유지(ABI research, 2019)¹³⁾

8) markets and markets(2019. 2), "Edge AI Software Market".

9) 360 Research Reports(2020. 5.), "Global Edge AI Software Market Growth (Status and Outlook) 2019-2024".

10) Fior Markets(2020. 1), "Global Edge AI Hardware Market".

11) Fortune Business Insights(2020. 7), "Artificial Intelligence (AI) Market size, Share & COVID-19 Impact Analysis".

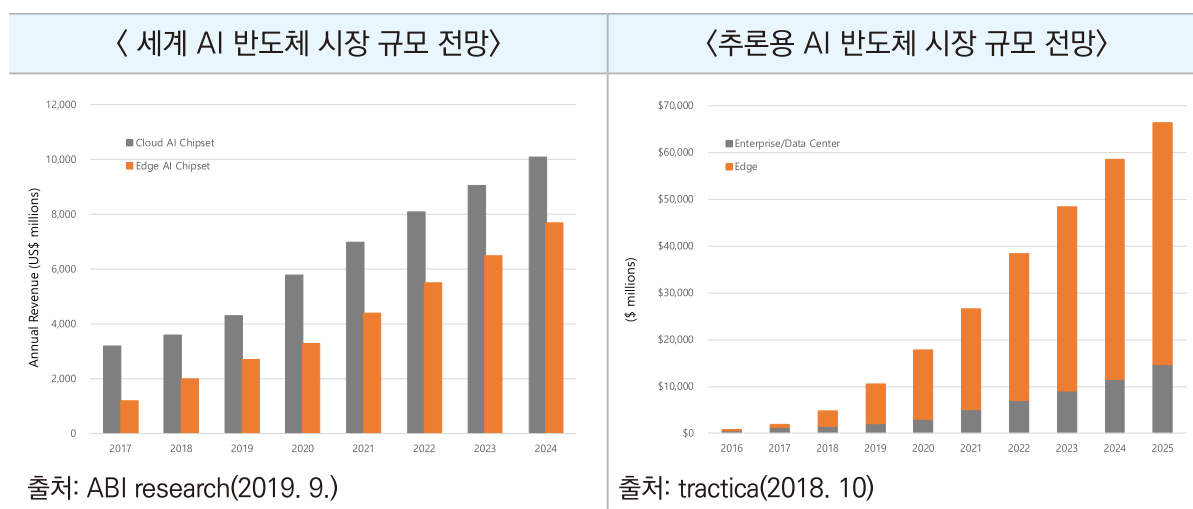
12) tractica(2020. 1), "Artificial Intelligence Market Forecasts".

13) eetasia(2019. 9. 11.), "Market for Cloud AI Chipsets to Double in Next 5yrs".

- (추론용) '25년 엣지에 활용되는 추론용 칩 시장이 클라우드 기반의 학습용 칩 시장보다 3.5배 클 것으로 예측

* '25년 클라우드 기반 AI 반도체는 145억 달러, 엣지 AI 반도체는 514억 달러를 차지할 것으로 전망(Tractica, 2018)¹⁴⁾

[그림 2] AI 반도체의 시장 규모 및 전망



시 사 점

◆ 엣지 AI는 클라우드 기반 AI의 단점을 보완할 수 있다는 점에서 각광받고 있음

- ▶ 이용자의 디바이스에 대한 의존도가 증가하고 스마트기기와 센서 등의 발전 및 빠른 성장에 힘입어 하드웨어 공급자 중심으로 엣지 AI 시장이 부상
 - 엣지 AI는 보안 및 정보보호, 지연시간 절감, 오프라인에서의 구동 가능성 확보, 컴퓨팅 파워 절약 등의 강점을 갖고 있으며 적은양의 데이터로도 AI 서비스 제공이 가능
 - 스마트기기의 사양이 높아지는 상황에서 엣지 AI는 클라우드 기반 AI의 단점을 보완하면서 더욱 성장할 것이라는 전망
- ▶ 메인프레임에서 개인용 컴퓨터로의 전환과 같이 기존 중앙 집중 성격의 클라우드 기반 AI에서 점차 분권적 성격의 엣지 AI로의 전환이 일어날 것이라는 전망이 존재
 - 엣지에서의 데이터 처리 비중 증가, AI 기능이 탑재된 디바이스의 증가가 엣지 시장으로의 패러다임 전환의 근거로 제시되고 있음

14) semiengineering.com(2018. 11. 28.), "AI Chip Architectures Race To The Edge".

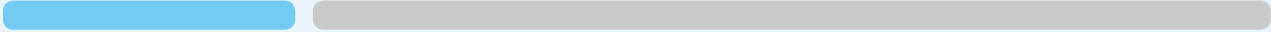
◆ 엣지 AI는 클라우드 기반 AI와 차별화된 영역에서 새로운 시장 기회를 창출할 수 있다는 점에서 의미가 있을 것

- ▶ 시장 규모와 기술적 장단점을 고려할 때 엣지 AI 시장이 주류가 될 것이라는 전망 보다는 활용 영역이 확장된다고 보는 것이 적절
 - 엣지 AI가 전체 시장에서 차지하는 비중이 미미하다는 점에서 패러다임의 전환으로 볼 수 있는 근거가 부족
 - 또한 클라우드 AI와 엣지 AI는 모두 장점과 한계점을 가지고 있다는 점에서 하나의 시장이 다른 시장을 대체하기 쉽지 않을 것
 - 클라우드 기반 AI는 방대한 양의 데이터와 연산 능력을 기반으로 AI 알고리즘을 구축하고 훈련하는 과정에서 보다 유용
 - 엣지 AI는 훈련된 알고리즘을 디바이스 단위에서 추론하여 판단하는 과정에서 보다 유용하게 적용 가능
- ▶ 따라서 엣지 AI는 하드웨어 공급자 뿐 아니라 소프트웨어 공급자에게도 새로운 시장 기회를 제공하면서 클라우드 기반 AI와 함께 성장할 가능성
 - (하드웨어 산업) AI 칩 공급자는 추론용 AI 반도체 시장에서 급격한 시장 확대 기회를, 디바이스 제조사는 AI 기능이 탑재된 기기를 통해 제품 차별화의 기회를 갖게될 것
 - (소프트웨어 산업) 엣지 AI와 클라우드 기반 AI의 소프트웨어 공급자가 대동소이하다는 점에서 엣지 AI의 시장 성장은 기존 AI 소프트웨어 공급사의 시장 확대 기회가 될 것

◆ 나아가 향후 엣지 AI의 성장은 소비자에게 보다 밀접한 인공지능 서비스를 제공할 수 있으며 데이터가 저장 및 가공되는 방식에 있어서도 변화를 야기할 수 있어 주목할 필요

참고문헌

- 360 Research Reports(2020. 5.), “Global Edge AI Software Market Growth (Status and Outlook) 2019–2024”.
- ciokorea(2020. 8. 6.), “기대반 우려반... 클라우드를 능가할 엣지 시대는 어떠한가”.
- Duncan Stewart et. al.(2019. 12. 9.), “Bringing AI to the device: Edge AI chips come into their own”, Deloitte Insights.
- eetasia(2019. 9. 11.), “Market for Cloud AI Chipsets to Double in Next 5yrs”.
- Fortune Business Insights(2020. 7), “Artificail Intelligence (AI) Market size, Share & COVID-19 Impact Analysis”.
- Fior Markets(2020. 1), “Global Edge AI Hardware Market”.
- Gartner(2018. 10. 3.), “What Edge Computing Means for Infrastructure and Operations Leaders”.
- Jeff Loucks(2018. 12. 11.), “Artificial intelligence: From expert-only to everywhere”. Deloitte Insights.
- markets and markets(2019. 2), “Edge AI Software Market”.
- _____ (2019. 4), “Edge AI Hardware Market”.
- _____ (2020. 5), “Covid-19 Impact On Edge AI Software Market”.
- Nutenix(2020. 6. 19.), “AI in the Cloud: Questions Many Aren’t Asking”.
- Queiroz, J., Leitão, P., Barbosa, J. and Oliveira, E.(2019), “Distributing Intelligence among Cloud, Fog and Edge in Industrial Cyber-physical Systems”, ICINCO.
- semiengineering.com(2018. 11. 28.), “AI Chip Architectures Race To The Edge”.
- thinkbiz.net(2019. 2. 12.), “EDGE COMPUTING – MOVING INTELLIGENCE TO THE EDGE”.
- tractica(2020. 1), “Artificial Intelligence Market Forecasts”.
- 아시아경제(2019. 12. 4.), “인텔 2023년, AI 업무 43% 엣지에서 처리”.
- 아이티데일리(2017. 7. 31.), “[전문가 강좌] 4차 산업혁명과 지능정보기술의 미래방향(7)”.
- 전자신문(2020. 4. 1.), “어드밴텍, AI 에지컴퓨터 ‘MIC-770’ 5G 정찰로봇에 장착... 中 공항서 코로나 검사 활용”.
- 컴퓨터월드(2019. 12. 30), “[시장동향] AI에도 클라우드 바람, AlaaS로 코드 없는 AI 개발 가능”.
- 한국경제(2020. 7. 29.), “엣지 AI 시장, 10년 뒤 155조로 성장...비대면·스마트시티에 필수”.



KISDI AI TREND WATCH는 인공지능 관련 주요 이슈와 최신 동향 정보를 제공하는 온라인 정기간행물입니다.

KISDI AI전략센터 및 산학연 전문가들이 참여하여 매월 15일과 30일에 온라인으로 배포합니다.

본지에 게재된 내용은 본 연구원의 공식 견해와 다를 수 있습니다.

보고서와 관련된 문의는 AI전략센터 김민진 전문연구원(minjinkim@kisdi.re.kr, 043-531-4356)으로 연락주시기 바랍니다.