

주제	영어 자기주도학습 리포트
요약	<ul style="list-style-type: none">- DeepMind AI outdoes human mathematicians on unsolved problem- China's short-sighted AI regulation- The UK AI Safety Summit, combined with a G7 declaration and US executive order, shows action is happening on AI, but what happens next? <p>이렇게 3가지 기사를 준비했습니다. [자료4]에 “세 개의 글을 읽고 알게 된 점”을 먼저 읽어보시고 기사 원문을 읽어보시길 추천합니다. 또, “3학년에 공부하고 싶은 분야”는 민성 학생이 원하시는 대로 수정하시길 바랍니다.</p>

자료 1. DeepMind AI outdoes human mathematicians on unsolved problem

1. 글제목 및 출처

DeepMind AI outdoes human mathematicians on unsolved problem

-Large language model improves on efforts to solve combinatorics problems inspired by the card game Set.

(14 December 2023)

<https://www.nature.com/articles/d41586-023-04043-w>

2. 선정한 이유

AI 기술은 점점 발전해왔고, 최근 chatGPT 등을 통해 보다 우리 실생활에 가깝게 다가왔습니다. 그러나 이러한 기술들은 기존의 데이터들을 기반으로 한 요약과 정보 검색, 전달에 그 방향성이 있었습니다. 이번 구글의 DeepMind AI의 진화는 그와는 달리 수학 문제 해결에 도움을 줌으로써 AI의 역할과 그 방향성이 보다 확장되어갈 것임을 암시하는 사례라고 보여져 선정하게 되었습니다.

3. 기사 원문

The card game Set has long inspired mathematicians to create interesting problems.

Now, a technique based on large language models (LLMs) is showing that artificial intelligence (AI) can help mathematicians to generate new solutions.

The AI system, called FunSearch, made progress on Set-inspired problems in combinatorics, a field of mathematics that studies how to count the possible arrangements of sets containing finitely many objects. But its inventors say that the method, described in Nature on 14 December¹, could be applied to a variety of questions in maths and computer science.

“This is the first time anyone has shown that an LLM-based system can go beyond what was known by mathematicians and computer scientists,” says Pushmeet Kohli, a computer scientist who heads the AI for Science team at Google Deepmind in London. “It’s not just novel, it’s more effective than anything else that exists today.”

This is in contrast to previous experiments, in which researchers have used LLMs to solve maths problems with known solutions, says Kohli.

Mathematical chatbot

FunSearch automatically creates requests for a specially trained LLM, asking it to write short computer programs that can generate solutions to a particular mathematical problem. The system then checks quickly to see whether those solutions are better than known ones. If not, it provides feedback to the LLM so that it can improve at the next round.

“The way we use the LLM is as a creativity engine,” says DeepMind computer scientist Bernardino Romera-Paredes. Not all programs that the LLM generates are useful, and some are so incorrect that they wouldn’t even be able to run, he says. But another program can quickly toss the incorrect ones away and test the output of the correct ones.

The team tested FunSearch on the ‘cap set problem’. This evolved out of the game Set, which was invented in the 1970s by geneticist Marsha Falco. The Set deck contains 81 cards. Each card displays one, two or three symbols that are identical in colour, shape and shading – and, for each of these features, there are three possible options. Together, these possibilities add up to $3 \times 3 \times 3 \times 3 = 81$. Players have to turn over the cards and spot special combinations of three cards called sets.

Mathematicians have shown that players are guaranteed to find a set if the number of upturned cards is at least 21. They have also found solutions for more-complex versions of the game, in which abstract versions of the cards have five or more properties. But some mysteries remain. For example, if there are n

properties, where n is any whole number, then there are $3n$ possible cards – but the minimum number of cards that must be revealed to guarantee a solution is unknown.

This problem can be expressed in terms of discrete geometry. There, it is equivalent to finding certain arrangements of three points in an n -dimensional space. Mathematicians have been able to put bounds on the possible general solution – given n , they have found that the required number of ‘cards on the table’ must be greater than that given by a certain formula, but smaller than that given by another.

Human-machine collaboration

FunSearch was able to improve on the lower bound for $n = 8$ by generating sets of cards that satisfy all the requirements of the game. “We don’t prove that we cannot improve over that, but we do get a construction that goes beyond what was known before,” says DeepMind computer scientist Alhussein Fawzi.

One important feature of FunSearch is that people can see the successful programs created by the LLM and learn from them, says co-author Jordan Ellenberg, a mathematician at the University of Wisconsin-Madison. This sets the technique apart from other applications, in which the AI is a black box.

“What’s most exciting to me is modelling new modes of human-machine collaboration,” Ellenberg adds. “I don’t look to use these as a replacement for human mathematicians, but as a force multiplier.”

4. 해석

카드 게임은 오랫동안 수학자들에게 있어 흥미로운 문제들을 만들 수 있는 영감을 주었습니다.

이제, 거대 언어 모델(LLMs) 기반의 기술은 인공지능(AI)이 수학자들에게 새로운 해결책을 내는 데 도움을 줄 수 있음을 보여주고 있습니다.

인공지능 시스템, FunSearch는 무한히 많은 물체를 포함하는 집합이 만들어 낼 수 있는 배열들을 세는 방법을 연구하는 수학 분야인 조합(combinatorics)에서 집합과 관련된 문제들을 해결하는 데 있어 발전을 이루었습니다. FunSearch의 개발자들은 12월 14일 네이처(Nature)에 기술된 이 방법이 수학과 컴퓨터 과학의 다양한 문제들에도 적용될 수 있다고 말합니다.

“이것은 LLM 기반의 시스템이 기존에 수학자들과 컴퓨터 과학자들에 의해 알려진 것 이상

을 할 수 있음을 첫 사례입니다."라고 런던 구글 딥마인드에서 과학을 위한 인공지능을 연구하는 팀을 이끄는 컴퓨터 과학자 Pushmeet Kohli는 말합니다. "이는 새로운 정도가 아니라 현존하는 모든 것보다 효과적인 방식입니다."

이 방법은 기존의 연구자들이 LLM을 기존에 알려진 해결책을 통해 수학 문제들을 해결하는데 사용한 것과 대조적인 방식이라고 Kohli는 말합니다.

수학 챗봇

FunSearch는 특별하게 훈련된 LLM에게 요청을 자동적으로 생성하여, LLM이 특정 수학 문제에 대한 해답을 만들 수 있는 짧은 컴퓨터 프로그램을 작성하도록 요청합니다. 이후 시스템은 그렇게 만들어진 해답들이 기존의 해답들보다 더 나은지를 빠르게 확인합니다. 그렇지 않다면, LLM에 피드백을 제공해 다음 라운드에서 개선할 수 있습니다.

"우리는 LLM을 창의력 엔진으로 사용합니다."라고 딥마인드의 컴퓨터 과학자 Bernardino Romera-Paredes는 말합니다. 그는 LLM가 만들어내는 모든 프로그램이 유용한 것은 아니며, 일부 프로그램은 너무 부정확하여 실행되지도 않는다고 말합니다. 하지만 또다른 프로그램이 부정확한 프로그램은 빠르게 넘기고 올바른 프로그램의 출력을 테스트하게 합니다.

연구팀은 FunSearch를 '모자 세트 문제'에 대해 테스트했습니다. 이 문제는 1970년대에 유전학자 Marsha Falco에 의해 발명된 게임 세트에서 발전했습니다. 덱은 81장의 카드를 포함합니다. 각 카드는 색상, 모양, 음영이 동일한 한 개, 두 개 또는 세 개의 상징을 가지고, 이 각각의 특징에는 세 가지 가능한 옵션이 있습니다. 종합하면, 총 $3 \times 3 \times 3 = 81$ 가지의 경우가 존재합니다. 경기자들은 카드를 뒤집으며 세 개의 카드의 특별한 조합을 찾아야 합니다.

수학자들은 경기자가 적어도 21장 이상의 카드를 뒤집으면 한 세트를 찾을 수 있음을 확인했습니다. 또한 그들은 카드가 5개 이상의 속성을 가지고 있는 더 복잡한 버전의 게임에 대한 해결책도 찾아냈습니다. 그러나 여전히 의문점은 남아 있었습니다. 예를 들어, 임의의 정수 n 에 대해, 만약 n 개의 속성이 있을 경우, 총 3^n 의 경우의 수가 있을 것입니다. 그러나 경기자가 한 세트를 찾을 수 있음을 보장하는 최소 카드 수는 알려져 있지 않습니다.

이 문제는 이산기하학 측면에서 표현될 수 있습니다. 결국 이는 n 차원 공간에서 특정한 세 점의 배열을 찾는 것과 같습니다. 수학자들은 일반해로 해의 범위를 특정할 수 있었습니다. n 이 주어지면, 그들은 필요한 '표 위의 카드'의 수는 특정한 공식에 의한 값보다 커야 하지만 다른 공식에 의한 값보다 작아야 한다는 것을 발견했습니다.

인간-기계의 협업

FunSearch는 해답이 되는 모든 카드 세트를 생성함으로써 $n=8$ 의 하한선을 개선할 수 있었습니다. "우리는 그 이상으로 개선할 수 없다는 것을 증명하지 못하지만, 이전에 알려진

것 이상을 얻었습니다."라고 DeepMind 컴퓨터 과학자 Alhusein Fawzi는 말합니다.

FunSearch의 한 가지 중요한 특징은 사람들이 LLM에 의해 만들어진 성공적인 프로그램들을 보고 그것들로부터 배울 수 있는 것이라고 Wisconsin-Madison 대학의 수학자 Jordan Ellenberg는 말합니다. 이 특징은 해당 기술을 AI가 black box인 다른 기술들과 구별되게 합니다.

"저를 가장 흥분시키는 것은 인간-기계 협력의 새로운 방식을 모델링하는 것입니다"라고 Ellenberg는 덧붙입니다. "저는 이것이 인간 수학자들의 대체품이 아닌, 그들의 능력을 배로 시키는 데 사용될 것이라고 봅니다."

5. 단어 및 속어

arrangements : 배열

be guaranteed to : ~가 보장되다.

property : 성질

discrete geometry : 이산 기하학

the lower bound : 하한

application : 프로그램

force multiplier : 전력 승수

be equivalent to : 동등하다. 맞먹다.

in terms of : ~측면에서

set A apart from B : A를 B와 구별하다.

not A, but B : A가 아니라 B로

black box : 기능은 알지만 작동 원리를 이해할 수 없는 복잡한 기계 장치나 시스템, 물체

자료 2. China's short-sighted AI regulation

1. 글제목 및 출처

China's short-sighted AI regulation

(December 12, 2023)

TheKoreaTimes

https://www.koreatimes.co.kr/www/opinion/2023/12/137_364911.html?utm_source=KK

2. 선정한 이유

기술과 규제는 떼 놓을 수 없는 관계로, 정책에 따라 기술이 인간에게 어떻게 영향을 끼칠 수 있는지에 관한 의견을 잘 보여줍니다. 기사의 전체적인 내용은 AI 분야를 선도하고자 생성형 AI로 만든 콘텐츠의 저작권을 보장하며 AI 훈련에서 저작권 있는 자료의 학습을 쉽게 만들려는 중국의 의도를 AI 콘텐츠 저작권을 인정하지 않는 미국의 입장과 비교하며 중국 내부에 혼란을 일으킬 근시안적 결정이라고 비판하는 내용입니다.

3. 기사 원문

PARIS - The Beijing Internet Court's ruling that content generated by artificial intelligence can be covered by copyright has caused a stir in the AI community, not least because it clashes with the stances adopted in other major jurisdictions, including the United States. In fact, that is partly the point: the ruling advances a wider effort in China to surpass the U.S. to become a global leader in AI.

Not everyone views the ruling as all that consequential. Some commentators point out that the Beijing Internet Court is a relatively low-level institution, operating within a legal system where courts are not obligated to follow precedents. But, while technically true, this interpretation misses the point, because it focuses narrowly on Chinese law, as written. In the Chinese legal context, decisions like this one both reflect and shape policy.

In 2017, China's leaders set the ambitious goal of achieving global AI supremacy by 2030. But the barriers to success have proved substantial - and they continue to multiply. Over the last year or so, the U.S. has made it increasingly difficult for China to acquire the chips it needs to develop advanced AI technologies, such as large language models, that can compete with those coming out of the U.S. President Joe Biden's administration further tightened those regulations in October.

In response to this campaign, China's government has mobilized a whole-of-society effort to accelerate AI development, channeling vast investment

toward the sector and limiting regulatory hurdles. In its Interim Measures for the Management of Generative Artificial Intelligence Services - which entered into effect in August - the government urged administrative authorities and courts at all levels to adopt a cautious and tolerant regulatory stance toward AI.

If the Beijing Internet Court's recent ruling is any indication, the judiciary has taken that guidance to heart. After all, making it possible to copyright some AI-generated content not only directly strengthens the incentive to use AI, but also boosts the commercial value of AI products and services.

Conversely, denying copyrights to AI-generated content could inadvertently encourage deceptive practices, with digital artists being tempted to misrepresent the origins of their creations. By blurring the lines between AI-generated and human-crafted works, this would pose a threat to the future development of AI foundational models, which rely heavily on training with high-quality data sourced from human-generated content.

For the U.S., the benefits of prohibiting copyright protection for AI-generated content seem to outweigh the risks. The U.S. Copyright Office has refused to recognize such copyrights in three cases, even if the content reflects a substantial human creative or intellectual contribution. In one case, an artist tried over 600 prompts - a considerable investment of effort and creativity - to create an AI-generated image that eventually won an award in an art competition, only to be told that the copyright would not be recognized.

This reluctance is hardly unfounded. While the Beijing Internet Court ruling might align with China's AI ambitions today, it also opens a Pandora's box of legal and ethical challenges. For instance, as creators of similar AI artworks dispute copyright infringement, Chinese courts could be burdened by a surge of litigation at a time when Chinese courts must face the contentious issue of whether copyright holders can obtain compensation for the use of their AI-generated works in AI training. This makes a revision of existing copyright laws and doctrines by Chinese courts and the legislature all but inevitable.

Questions about copyrights and AI training are already fueling heated debates in a number of jurisdictions. In the US, artists, writers, and others have launched a raft of lawsuits accusing major AI firms like OpenAI, Meta, and Stability AI of using their copyrighted work to train AI systems without permission. In Europe, the proposed AI Act by the European Parliament requires that companies disclose any copyrighted materials they use for training generative-AI systems - a rule

that would make AI firms vulnerable to copyright-infringement suits, while increasing the leverage of copyright-holders in compensation negotiations.

For China, addressing such questions might prove to be particularly complicated. Chinese law permits the free use of copyrighted materials only in very limited circumstances. But with Chinese courts increasingly aligning their rulings with directives from Beijing, it seems likely that, to facilitate the use of copyrighted materials by AI firms, they will soon start taking a laxer approach and approve a growing number of exceptions.

The price, however, could be steep. The adoption of a more lenient approach toward the use of copyrighted materials for AI training - as well as the likely flood of AI-generated content on the Chinese market - may end up discouraging human creativity in the long term.

From the government to the courts, Chinese authorities seem fixated on ensuring that the country can lead on AI. But the consequences of their approach could be profound and far-reaching. It is not inconceivable that this legal trend could lead to social crises such as massive job losses in creative industries and widespread public discontent. For now, however, China can be expected to continue nurturing its AI industry - at all costs.

4. 해석

중국의 근시안적인 AI 규제

파리 - 인공지능에 의해 생성된 콘텐츠가 저작권으로 보호될 수 있다는 베이징 인터넷 법원의 판결이 미국을 포함한 다른 주요 국가에서 채택된 입장과 충돌하기 때문에 AI계에 파문을 일으켰습니다. 사실, 부분적으로 이것이 핵심입니다: 이 판결은 중국에서 미국을 능가하여 AI 분야의 세계적인 리더가 되고자 하는 더 광범위한 노력을 발전시킵니다.

모두가 그 판결이 그렇게 중대하다고 보는 것은 아닙니다. 일부 논평가들은 베이징 인터넷 법원이 선례를 따를 의무가 없는 법적 시스템 내에서 운영되는 비교적 낮은 수준의 기관이라고 지적합니다. 그러나 엄밀히 말하면 사실이지만, 이러한 해석은 중국 법에 적혀 있는 대로 좁게 초점을 맞추고 있기 때문에 핵심을 놓치고 있습니다. 중국의 법적 맥락에서, 이와 같은 결정은 정책을 반영하고 구체화합니다.

2017년 중국의 지도자들은 2030년까지 글로벌 AI 패권을 달성하겠다는 야심 찬 목표를 세웠습니다. 하지만 성공의 장벽은 상당한 것으로 입증되었고, 그 장벽은 계속해서 커지고 있습니다. 지난 1년여 동안, 미국은 중국이 미국에서 나오는 AI와 경쟁할 수 있는 대형 언어 모델과 같은 첨단 AI 기술을 개발하는 데 필요한 칩을 획득하는 것을 점점 더 어렵게 만들었습니다. 조 바이든 대통령 행정부는 10월에 이러한 규제들을 더욱 강화했습니다.

이러한 캠페인에 대응하여 중국 정부는 AI 개발을 가속화하기 위해 사회 전체의 노력을 동원하여 그 부문으로 막대한 투자금을 전달하고 규제 장애물을 제한했습니다. 8월에 발효된 생성형 AI 서비스 관리를 위한 임시 조치에서 정부는 행정 당국과 각급 법원이 AI에 대해 신중하고 관대한 규제 입장을 채택할 것을 촉구했습니다.

베이징 인터넷 법원의 최근 판결이 시사하는 바가 있다면, 사법부는 그 지침을 명심해야 합니다. 결국, AI가 생성한 일부 콘텐츠에 저작권을 부여하는 것은 AI를 사용하려는 동기를 직접적으로 강화할 뿐만 아니라 AI 제품과 서비스의 상업적 가치를 향상시킵니다.

반대로, AI가 생성한 콘텐츠에 대한 저작권을 부정하는 것은 의도치 않게 기만적인 관행을 부추길 수 있으며, 디지털 아티스트는 창작물의 출처를 거짓되게 전하려는 유혹을 받습니다. AI가 생성한 작품과 인간이 만든 작품 사이의 경계를 모호하게 함으로써, 이것은 인간이 생성한 콘텐츠에서 추출한 고품질 데이터를 사용하여 훈련에 크게 의존하는 미래의 AI 기반 모델 개발에 위협이 될 것입니다.

미국의 경우, AI가 생성한 콘텐츠에 대한 저작권 보호를 금지하는 것의 이점이 위험보다 더 큰 것으로 보입니다. 미국 저작권청은 콘텐츠가 상당한 인간의 창의적 또는 지적 기여를 반영한다고 하더라도 세 가지 경우에 그러한 저작권을 인정하는 것을 거부했습니다.

한 경우에, 한 예술가는 AI 생성 이미지를 만들기 위해 600개 이상의 프롬프트(노력과 창의성의 상당한 투자)를 시도하여 결국 예술 대회에서 상을 받았지만, 저작권이 인정되지 않을 것이라는 말을 들을 뿐이었습니다.

이러한 난색은 거의 근거가 없습니다. 베이징 인터넷 법원의 판결이 오늘날 중국의 AI 야망과 일치할 수도 있지만 법적, 윤리적 도전의 판도라의 상자를 열어줍니다. 예를 들어, 유사한 AI 예술품의 창작자들이 저작권 침해에 이의를 제기함에 따라 중국 법원은 저작권자가 AI 훈련에서 그들의 AI 생성 작품을 사용하는 것에 대한 보상을 받을 수 있는지에 관한 논쟁적인 문제에 직면해야 하는 시점에서 급증하는 소송에 대해 부담을 느낄 수 있습니다. 이로 인해 중국 법원과 입법부의 기존 저작권 법 및 정책 수정이 거의 불가피합니다.

저작권과 AI 훈련에 대한 질문은 이미 많은 국가에서 열띤 논쟁을 부채질하고 있습니다. 미국에서는 예술가, 작가 등이 OpenAI, Meta, Stability AI와 같은 주요 AI 회사들이 허가 없이 AI 시스템을 훈련시키기 위해 저작권이 있는 작업물을 사용했다고 고발하는 소송을 개시했습니다. 유럽에서는 유럽 의회가 제안한 AI법에 따라 회사가 생성형 AI 시스템 훈련에 사용하는 저작권이 있는 자료를 공개할 것을 요구하고 있는데, 이 규칙은 AI 회사가 저작권 침해 소송에 취약하도록 만드는 동시에 보상 협상에서 저작권자의 영향력을 높일 수 있는 규칙입니다.

중국에게 있어서, 그러한 질문들을 해결하는 것은 특히 복잡할 수 있습니다. 중국 법은 매우 제한된 상황에서만 저작권이 있는 자료의 자유로운 사용을 허용합니다. 그러나 중국 법원들이 점점 더 베이징의 지시에 그들의 판결들을 맞추면서, AI 회사들에 의한 저작권이 있는 자료의 사용을 용이하게 하기 위해, 그들은 곧 완화적인 접근법을 취하기 시작하고 점점 더 많은 예외들을 승인할 것으로 보입니다.

그러나 가격은 가파를 수 있습니다. AI 훈련을 위한 저작권이 있는 자료의 사용에 대한 더 관대한 접근법의 채택과 중국 시장에서 AI가 생성한 콘텐츠의 홍수 가능성은 결국 인간의 창의력을 저해할 수 있습니다.

정부에서 법정에 이르기까지, 중국 당국은 중국이 AI를 선도할 수 있도록 보장하는 데 집착(fixated)하는 것처럼 보입니다. 그러나 그들의 접근 방식의 결과는 심오하고 광범위할 수 있습니다. 이러한 법적 추세가 창조 산업의 대규모 실직과 광범위한 대중 불만과 같은 사회적 위기로 이어질 수 있다는 것은 상상할 수 없는 일이 아닙니다. 그러나 현재로서 중국은 어떤 대가를 치르더라도 AI 산업을 계속 육성할 것으로 예상될 수 있습니다.

5. 단어 및 속어

copyright: 저작권, 판권

jurisdiction: 사법[재판]권; 지배(권); 관할, 권한; (경찰 등의) 관할권, 관할구, 관구 (기사에 서는 major jurisdiction: 주요국이라는 맥락으로 쓰임.)

surpass: ...보다 낫다, 능가하다, 뛰어나다

consequential: 중대한

supremacy: 주권, 대권; 패권, 지배권, 지상권

substantial: 상당한 (양적), 많은

campaign: (사회적, 정치적) 운동, 캠페인

administration: 행정부

mobilize: 동원하다, <산업, 자원 등을> 전시 체제로 하다, <지지, 힘 등을> 결집하다

interim: 당분간[임시]의, 중간의; 가..., 잠정의

administrative authority: 행정 당국

inadvertent: 고의가 아닌, 우연의; 의도하지 않은

deceptive: <사람을> 속이는, 현혹시키는; 믿을 수 없는; 오해를 사는

blur: <광경, 의식, 눈 등을> 흐리게 하다, (잉크 따위로) 더럽히다, ...에 얼룩이 지게 하다

pose: 자세[포즈]를 취하다, 꾸민 태도를 취하다, (...처럼) 보이려고 애쓰다, (...인) 체하다

competition: 경기, 시합, 경쟁 시험

reluctance: 싫음, 마지못해 함, 꺼림, 마음 내키지 않음; [드물게] 반항

infringement: (법규) 위반, 위배; (특허권, 판권 등의) 침해; 위반[침해] 행위

dispute: 논쟁하다, (...에 대해) 논의하다, 반론하다, 이의를 말하다

surge: <군중, 감정 등이> 파도처럼 밀려오다, 쇄도하다, 밀어닥치다, 들끓다

litigation: 소송, 기소

contentious: 다투기 좋아하는, 논쟁하기 좋아하는

copyright holder: 저작권자
doctrin: (종교의) 교의, 교리, (정치, 학문상의) 주의; [미] 공식 (외교) 정책; 원칙, 학설
legislature: 입법부
raft: 뗏목, 뗏목배, 고무 보트
launch a raft of lawsuit: 소송을 개시하다
parliament: 의회, 국회; 하원
vulnerable: 상처 입기 쉬운, 공격받기 쉬운, 비난받기 쉬운, 취약성[약점]이 있는, (유혹 등
에) 넘어가기 쉬운
leverage: (목적 달성을 위한) 수단, 효력; 권력, 세력
compensation: 배상, 보상; 보충; 보상[배상]금; 보수, 봉급
directive: 지령, 명령; [컴퓨터] 지시어, 지시문
facilitate: <사정이><일을>용이하게 하다, 쉽게 하다, <행동, 과정 등을> 촉진[조장]하다
lax: 느슨한, 헐거운, 완만한, <규율, 사람 등이> 엄격하지 못한; 단정치 않은, 트릿한, 방종
한
fixated: (...에) 집착한
profound: 깊은, <책, 사상 등이>뜻 깊은, 난해한
inconceivable: 상상할 수도 없는, 생각조차 할 수 없는, 터무니없는, 믿을 수 없는
nurture: <아이를> 양육하다, 기르다, 양성하다, 가르쳐 길들이다; 영양물을 공급하다

자료 3. The UK AI Safety Summit, combined with a G7 declaration and US executive order, shows action is happening on AI, but what happens next?

1. 글제목 및 출처

The UK AI Safety Summit, combined with a G7 declaration and US executive order, shows action is happening on AI, but what happens next?

<https://www.bbc.com/future/article/20231107-why-global-regulation-of-artificial-intelligence-is-still-a-long-way-off>

2. 선정한 이유

인공지능이 최근 화두로 떠오른 이래로, 인공지능의 개발이 점차 가속화되면서 무분별하게 개발되고 있는 인공지능을 보며 AI 안정성과 관련된 윤리적 문제와 같은 문제들에 대해 생각해보았고, 그리고 그에 대해 각국에서 어떻게 제도적으로 규제를 하려 하는지에 대해 관심이 생겨 다음 기사를 선정하게 되었다.

3. 기사 원문

As more than 100 attendees from civil society, the world's leading tech companies, and governments gathered in an English stately home, there was some tension.

This was the world's first AI Safety Summit convened by a major international country, and was intended to help chart the future direction of the technology. And it was held at Bletchley Park, where more than 80 years ago, a group of British scientists broke the German Enigma code.

The term "artificial intelligence" has been around since 1956, with the concept of machine learning existing for centuries before that - albeit often more in theory than in practice. In November, Collins Dictionary named the term as its "word of the year" as it had been the talking point of 2023.

AI already dictates large parts of our lives without us knowing it, and has for years. But the October 2022 release of ChatGPT by OpenAI, a company still less than a decade old, changed the paradigm.

Generative AI - where artificial intelligence takes its collective knowledge and an input, often in the form of a question or conversation starter, and produces new text, images or other content - has changed the game. And the speed of its development has alarmed the public and politicians who legislate for them.

Consensus has been agreed on the need for regulation: 28 countries including

the US, UK and China, alongside the European Union, signed the Bletchley Declaration, a world-first global agreement at the UK's AI Safety Summit, saying as much.

But what should be done, and by whom, is still up for debate.

The UK government announced on the first day of the summit it was launching a UK AI Safety Institute. But this was followed immediately after by the US unveiling its own version. Gina Raimondo, the US commerce secretary, acknowledged the competition. "Even as nations compete vigorously, we can and must search for global solutions to global problems," she said in a speech at the UK summit.

The day before the Bletchley Park summit began, the US unveiled an executive order outlining how it planned to regulate AI. It was a deliberate decision, believe experts watching the developments play out, for the US to grab control. And around the same time in Japan, the G7 group of industrialised countries issued a joint statement on the importance of regulating AI that seemed timed to remind the world that they, too, had a stake in the debate.

Plenty of processes are running in parallel across the world to try and rein in AI companies from unsafely developing their tools. But they're not always working together.

I think there's pretty divergent views across various countries around what exactly to do - Deb Raji

That much was clear at the UK summit itself, according to Deb Raji, a fellow at the Mozilla Foundation, who attended many of the meetings at Bletchley Park. "There was a global consensus on the thought that AI was something to reflect on and to regulate," she says. "But I think there's pretty divergent views across various countries around what exactly to do."

Raji points out that in some of the sessions she attended, even member states from within the European Union were advocating for different, competing and contradictory approaches to regulating AI that were also in opposition to the act the countries within the trading bloc agreed earlier this year. "Even within that more coordinated discourse around AI, there was a shocking amount of diversity in the perspectives involved," she says.

You might also like:

The employees secretly using AI at work
Why humans will never understand AI
Have we evolved to trust machines?

Such diversity and such competition is natural, says Margaret Mitchell, a researcher and chief ethics scientist at Hugging Face, an AI company lobbying for a more thoughtful, humanity-focused approach to AI. "Governments will seek to protect their national interests, and many of them will seek to establish themselves as leaders," she says. "That kind of competition to each be the leader falls out, I think, from the tendencies of humankind and the kinds of personalities that are promoted and empowered."

Mitchell says that much of the jostling to dictate global AI rules is down to "alpha manoeuvres" that are a function of who is in power in different countries worldwide. However, both Mitchell, Raji and Mike Katell, an ethics fellow at the Alan Turing Institute, highlighted the issue that voices within the Global North and economically more developed countries were given more prominence at the table this week, rather than from elsewhere in the world.

"You can look at maps where the AI regulation is happening around the world," says Katell. "There are big gaps in the Global South. There's very little happening in Africa."

Katell says that "if there's a single word that correctly describes the AI landscape, it's competition". That brings challenges with duplication and contradiction, believes Raji. "I think it's going to be really hard to come up with a consensus," she says.

Some forward steps are happening despite that competition. "The actions across countries this week are important steps in what is a difficult regulatory process on a global level," says David Haber, co-founder and chief executive of Lakera, an AI safety firm, who was an adviser to the European Union as it passed its own AI Act.

I am a little bit more optimistic than I was before this summit and the move by the Americans - Mike Katell

However, Haber acknowledges that it's progress, rather than action - and there's plenty of and duplication (and occasionally contradiction) depending on the jurisdiction. Until some baseline global consensus is reached, companies will have

to self-police, he says.

"While governments triangulate their policies, the real responsibility will continue to remain with industrial players," he says. "AI risks will continue to evolve quickly. We cannot wait for the technology to fully mature before applying standard security and safety practices."

As for who will win out, and what it means, the experts aren't sure. "I'm not really holding my breath that it will amount to much, but I am a little bit more optimistic than I was before this summit and the move by the Americans," says Katell.

Raji believes the Americans are leading the way. Their intervention this week "was a much more meaty regulatory intervention, and I think them putting that out was an important signal," she says. However, she hopes that the US alone doesn't have the final say. "I hope there is an opportunity for various countries to enter the conversation and influence each other," she says.

4. 해석

시민사회와 세계를 선도하는 기술 기업들, 정부들 등에서 온 100여 명의 참석자들이 영국의 대저택에 모여들면서 다소 긴장감이 감돌았습니다.

이것은 주요 국제 국가에 의해 소집된 세계 최초의 인공지능 안정성 정상 회담이었으며, AI 기술의 미래 방향을 설정하는 것을 돕기 위한 의도였습니다. 그리고 이는 블레츨리 공원에서 열렸는데, 그곳에서 80년도 더 전에 한 영국 과학자 그룹이 독일 에니그마 암호를 해독했습니다.

"인공 지능"이라는 용어는 1956년부터 존재해 왔으며, 기계 학습의 개념은 그 이전 수세기 동안 존재해 왔습니다 - 종종 실제보다 이론적인 논의에서 더 많았음에도 불구하고 말입니다. 11월에, 콜린스 사전은 2023년의 화제였던 이 용어를 "올해의 단어"로 명명했습니다.

AI는 이미 우리도 모르는 사이에 우리 삶의 많은 부분을 결정하고 수년 동안 그래왔습니다. 하지만 아직 10년도 채 되지 않은 회사인 OpenAI의 2022년 10월 ChatGPT 출시는 패러다임을 바꿨습니다.

생성형 인공 지능은 - 인공 지능이 종종 질문이나 대화 시작의 형태로 집단 지식과 입력을 취하고 새로운 텍스트, 이미지 또는 기타 콘텐츠를 생산하는 - 판도를 변화시켰습니다. 그리고 그것의 발전 속도는 대중과 그들을 위해 입법을 하는 정치인들을 놀라게 했습니다.

규제의 필요성에 공감대가 형성된 것은 유럽연합과 함께 미국, 영국, 중국 등 28개국이 영

국의 AI 안전성 정상회의에서 세계 최초로 세계적 합의를 이룬 '블레츨리 선언'에 서명하면서 부터입니다.

그러나 무엇이 누구에 의해 규제되어야 하는지는 여전히 논쟁의 여지가 있습니다.

영국 정부는 정상회담 첫날 영국 인공지능 안정성 연구소를 설립한다고 발표했습니다. 그러나 미국이 자체 버전을 공개한 직후에 이것이 뒤따랐습니다. 지나 라이몬도(Gina Raimondo) 미국 상무장관은 경쟁을 인정했습니다. "국가들이 격렬하게 경쟁할 때에도, 우리는 세계적인 문제에 대한 세계적인 해결책을 찾을 수 있고, 찾아야 합니다."라고 그녀는 영국 정상회담 연설에서 말했습니다.

블레츨리 파크 정상회담이 시작되기 전날, 미국은 AI를 어떻게 규제할 것인지를 설명하는 행정명령을 발표했습니다. 전문가들은 미국이 통제권을 쥐기 위한 의도적인 결정이었다고 믿고 있습니다. 그리고 비슷한 시기에 일본에서도 G7 선진국 그룹이 AI 규제의 중요성에 대한 공동 성명을 발표했는데, 이는 자신들도 이 논쟁에 참여하고 있다는 것을 전 세계에 상기시키기 위한 때가 된 것 같습니다.

인공지능 회사들이 그들의 도구를 안전하게 개발하는 것을 시도하고 억제하기 위해 전 세계에서 많은 프로세스들이 병렬로 실행되고 있습니다. 그러나 그들이 항상 함께 일하는 것은 아닙니다.

나는 정확히 무엇을 해야 하는지에 대해서는 여러 국가에서 매우 다양한 견해가 있다고 생각합니다. - Deb Raji

블레츨리 파크(Bletchley Park)의 많은 회의에 참석한 모질라 재단(Mozilla Foundation)의 동료인 데브 라지(Deb Raji)에 따르면, 그러한 정도는 영국 정상회담 자체에서 분명했습니다. "인공지능은 반성하고 규제해야 할 어떤 것이라는 생각에 세계적인 합의가 있었습니다."라고 그녀는 말합니다. "하지만 저는 정확하게 무엇을 해야 하는지에 대해 여러 국가에서 꽤 다양한 견해가 있다고 생각합니다."

Raji는 자신이 참석한 일부 세션에서 유럽 연합 내 회원국조차도 올해 초 무역 블록 내 국가들이 합의한 법에 반대하는 AI 규제에 대해 서로 다른, 경쟁적이고 모순적인 접근법을 옹호하고 있었다고 지적합니다. "AI에 대한 보다 조정된 담론 내에서도 충격적인 다양한 관점들이 있었습니다."라고 그녀는 말합니다.

다음 기사들에도 관심을 가져주세요.:

직원들이 몰래 AI를 업무에 사용합니다.

왜 인간은 AI를 절대 이해하지 못할까요?

우리는 기계를 신뢰하도록 진화했을까요?

이러한 다양성과 논쟁들은 당연하다고 인공지능에 대한 좀 더 사려 깊고 인간 중심의 접근법을 위해 로비를 하고 있는 인공지능 회사 Hugging Face의 연구원이자 수석 윤리 과학자인 마가렛 미첼은 말합니다. "정부들은 그들의 국가 이익을 보호하기 위해 노력할 것이고, 대부분은 자신들을 지도자로 확립하고자 할 것입니다"라고 그녀는 말합니다. "저는 각자의 지도자가 되기 위한 경쟁은 인류의 성향과 권력을 가지는 사람들의 성격에서 발생한다고 생각합니다."

미첼은 전 세계적인 인공지능 규칙을 좌우하는 많은 부분은 세계 다른 나라에서 누가 권력을 잡는지에 대한 함수인 "알파 기동"에 달려있다고 말합니다. 하지만, 앨런 튜링 연구소의 윤리 연구원인 미첼, 라지, 마이크 카텔 모두 이번 주 논의에서 북부와 경제적으로 더 발전된 나라들의 목소리가 다른 나라들에 비해 더 부각되었다는 점을 강조했습니다.

"여러분은 전세계에서 AI 규제가 일어나고 곳을 지도에서 볼 수 있습니다"라고 카텔은 말합니다. "지구 남부에는 큰 격차가 있습니다. 아프리카에서는 거의 일어나지 않습니다."

카텔(Katell)은 "AI 지형을 정확하게 설명하는 단 하나의 단어가 있다면, 그것은 경쟁이다"라고 말합니다. 라지(Raji)는 그것이 복제와 모순이라는 난관을 가져온다고 믿습니다. "저는 합의를 도출하는 것이 정말 어려울 것이라고 생각합니다."라고 그녀는 말합니다.

그러한 경쟁에도 불구하고 몇몇 진전들이 일어나고 있습니다. "이번 주에 국가 간의 조치들은 글로벌 수준에서 어려운 규제 과정의 중요한 단계들입니다."라고 자체 인공지능 법을 통과시킨 유럽 연합의 고문이었던 인공지능 안전 회사 라케라의 공동 설립자이자 최고 경영자인 데이비드 하버가 말합니다.

저는 이번 정상회담과 미국의 움직임에 대해 이전보다 조금 더 낙관적입니다 - 마이크 카텔

하지만, Haber는 그것이 행동이 아니라 절차이며, 관할권에 따라 많은 혼선과 중복 (그리고 때때로 모순)이 있다는 것을 인정합니다. Haber는 기본적인 글로벌 합의가 이루어지기 전까지는, 기업들이 자체적으로 단속해야 한다고 말합니다.

"정부들이 정책을 삼각으로 진행하는 동안, 실질적인 책임은 계속해서 산업 주체들에게 있을 것입니다"라고 그는 말합니다. "AI 위험은 계속해서 빠르게 진화할 것입니다. 우리는 표준 보안 및 안전 법들이 마련되기 전에 기술이 완전히 성숙하기를 기다릴 수 없습니다."

누가 이길 것인지, 그리고 그것이 무엇을 의미하는지에 대해 전문가들은 확신하지 못합니다. "저는 이 회담이 큰 영향을 줄 것이라고 기대하지는 않지만, 이번 정상회담과 미국의 움직임에 대해 이전보다 조금 더 낙관적입니다"라고 카텔은 말합니다.

라지는 미국이 주도권을 쥐고 있다고 믿고 있습니다. 이번 주 그들의 개입은 "훨씬 더 강

력한 규제 개입이었고, 저는 그들이 이를 발의하는 것이 중요한 신호였다고 생각합니다"라고 그녀는 말합니다. 하지만, 그녀는 미국만이 최종 결정권을 갖지 않기를 바랍니다. "저는 다양한 국가들이 대화에 참여하고 서로 영향을 미칠 수 있는 기회가 있기를 바랍니다"라고 그녀는 말합니다.

5. 단어 및 속어

stately home : (특히 일반인에게 공개되는) 대저택

summit : (산의) 정상, 산꼭대기; 절정, 정점, 정상 회담

albeit : 비록 ...일지라도

talking point : 논란거리, (회의에서 흔히 특정 주장을 지지하는) 연설 사항

alongside : ...옆에, 나란히 ...와 함께, ...와 동시에

divergent : 다양한

rein in : 말의 고삐를 당기다, ~을 억제하다

vigorously : 발랄하게, 힘차게.

executive : 행정적인

outline : 개요를 서술하다 (=sketch), 윤곽을 보여주다[나타내다]

triangulate : 삼각으로 만들다

자료 4. 정리

세 개의 글을 읽고 알게 된 점

첫 번째 기사에선 수학 분야에서의 AI의 활약을 다룹니다. Nature에 올라온 기사에선 거대 언어 모델에게 수학 문제를 학습시켜 이산기하학 문제를 풀어내는데 사용합니다. 비록 컴퓨터가 문제의 완전한 증명을 할 순 없지만 찾고자 하는 새로운 해 집합을 발견했다는 데 의의가 있으며 수학자들과의 협업에서 실용성을 입증하였습니다.

두 번째 기사는 중국의 AI 정책에 대한 전문가의 비평입니다. 최근 중국은 AI의 창작물에 대한 저작권을 인정했습니다. 이에 따라 AI를 사용하려는 동기를 유발시키고 중국에서 생성되는 AI 저작물이 많아질 것입니다. 하지만 Angela Huyue Zhang, professor of Law at New York University는 이런 정책이 근시안적이라고 비판하는데, 당장은 중국이 AI 시장에서 주도권을 잡는 데에 성공할지라도 AI의 남용이 인간의 창의력을 저해하기 때문입니다.

세 번째 기사는 AI기술에 대한 규제와 우려에 대해 다룹니다. AI에 대한 규제는 국가 간의 군비 경쟁과 비슷하게 진행되기 때문에 모든 국가들이 모여 규제를 만들지 않는 이상 AI의 위험은 계속해서 커져 나갈 것입니다. 현재 AI에 대한 정책 결정 과정에서 미국이 주도권을 잡고 있지만 국가간 합의가 이루어지기 전까진 산업 주체들에게 윤리적 책임이 있습니다.

종합하면 AI의 활용 방법은 점점 더 다양해지고 있고 그 만큼 위험도 증가하고 있다고 볼 수 있습니다. 과학 기술 연구, 기업 경영, 문화 활동 등등 소위 국력을 이루는 요소들 전반에 걸쳐 AI의 활용이 이루어지기 때문에 AI 기술 경쟁은 국가 간 경쟁으로 번진 상태

이고 각자 이 기술에서의 주도권을 잡기 위해 정책적, 외교적 노력을 하고 있습니다. 중국이 AI 창작물의 저작권을 인정한 것, 미국이 중국으로의 첨단 반도체 수출을 금지한 것은 이러한 경쟁들의 예시라고 볼 수 있습니다. 하지만 국가 간 기술 경쟁만 심화 되고 윤리적, 경제적 문제는 뒷전으로 하여 AI활용에 대한 적절한 규제를 하지 않으면 짧게는 대규모 실업, 길게는 인간의 사고력 발달 저해 등의 문제가 발생할 것이기 때문에 모든 국가들이 참여하는 인공지능 관련 윤리, 법 체계 회의가 필요하다고 생각합니다.

3학년 진학 후 공부하고 싶은 분야

이번 학습을 하면서 인공지능 중 거대언어모델을 접하게 되었습니다. ChatGPT같이 인간의 언어를 이해하고 생성하도록 훈련된 인공지능을 통칭하는 말인데, 대규모의 언어 데이터를 학습해 문장 구조나 문법, 의미, 단어 내에 내재된 다른 의미 등을 이해하고 생성할 수 있습니다. 거대 언어 모델은 딥러닝과 통계를 이용한 자연어 처리를 하는데 이 자연어 처리를 하는 방법에 대해 알아보고 싶습니다.