

# 거대 언어모델 미래전개도 II\* STEEP 프레임으로 기회와 위험 탐색

윤기영

한국외국어대학교 겸임교수  
에프엔에스컨설팅 미래전략연구소장  
synsaje@gmail.com

이명호

(사)미래학회 부회장  
태재연구재단 자문위원  
lee.myungho@gmail.com



\*'거대 언어모델 미래전개도 I'은 <SW중심사회> 106호(4월호)에 수록돼 있습니다

## 거대 언어모델 미래전개도 I (106호 게재)

- I. 들어가기
  1. 거대한 충격이 된 거대 언어모델
  2. 방법론
- II. 변화 전망
  1. 개요
  2. 사회

## 거대 언어모델 미래전개도 II

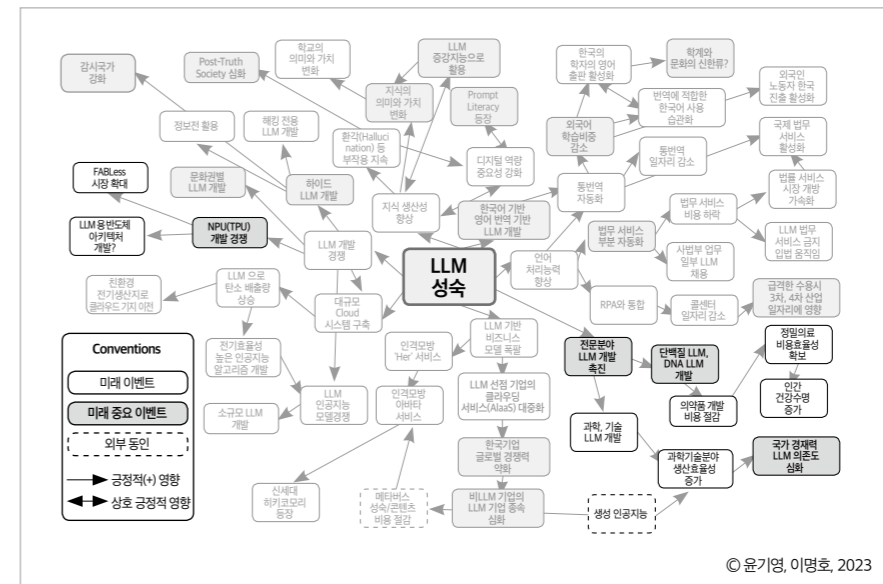
- 3. 기술
  - 4. 경제
  - 5. 환경
  - 6. 정치
- III. 결론
1. 정책 및 전략 의제 도출
  2. 한계와 과제

## II. 변화 전망

### 3. 기술

#### 가. 미래전개도 설명

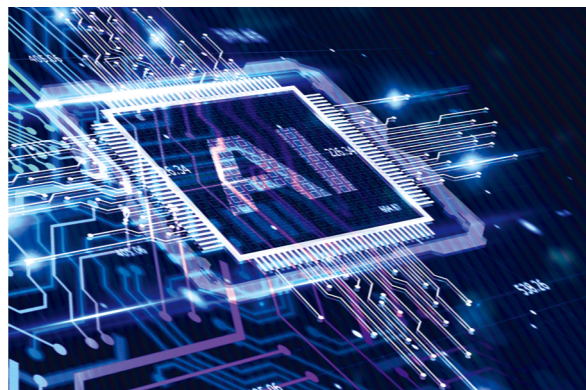
[그림 1] 기술 영역 미래전개도



거대 언어모델이 성숙하면 과학기술 전문분야 거대 언어모델이 확산할 것으로 보인다. 이미 물질분야나 특허분야 거대 언어모델이 개발됐다. 다만 이는 실험적으로 만들어진 것으로 충분한 규모의 거대 언어모델을 기반으로 만들어지지 않았다는 한계가 있다. 그런데 상황이 변할 것으로 보인다.

2023년 3월 구글은 거대 언어모델인 PaLM(Pathway Language Model)을 기반으로 의료 정보를 추가 학습하게 해 Med-PaLM2를 공개했다. 메타의 ESMFold는 단백질 시퀀스를 높은 정확도로 예측할 수 있다(Lin, et al., 2022.07). 거대 언어모델은 '측정 가능한 유전 정보와 관찰 가능한 형질 사이의 간극을 좁히는' 데 기여할 수 있다(Dalla-Torre, 2023.01). 단백질 구조 예측과 DNA 해석의 진전은 의약품 개발 등에 있어서 획기적인 비용효율성을 높일 뿐만 아니라, 합성생물학의 비약적 발전에 기여할 것으로 기대된다. 이는 인간의 기대수명을 높일 수 있을 것이며, 식량 생산성 제고에도 긍정적 영향을 미칠 것이다. 인간 기대수명 연장이 인류 모두에게 혜택을 줄 수 있을 것인지는 더욱 복잡한 상황을 고려해야 한다. 기대수명이 짧은 기간 내에 급하게 증가하는 경우 인류 사회는 깊은 불확실성(Deep Uncertainty)을 겪을 것으로, 이에 대한 상세하고 깊은 논의는 이 연구에서는 다루지 않겠다. 거대 언어모델은 다양한 과학기술 분야에 적용돼 과학기술 지식 생산성을 높일 것이다. 거대 언어모델 이외의 생성 인공지능도 과학기술 생산성 향상에 크게 기여할 것으로 보인다. 창발의 근본에는 연계지능이 있다. 난수에 의한 연결과 일정한 기준에 의한 검증으로 연계지능을 모방할 수 있다. 그 이외에 강화학습으로 과학기술 생산성을 높일 수도 있다. 정리하자면 인공지능은 인간의 창의를 충분히 모방할 수 있다. 이는 자연설계의 방식과 동일하다. 이를 후기 지적 설계(Post Intelligent Design)라 한다(Denett, 2020; 윤기영, 2018.08). 과학기술 생산성 향상은 지식반감기(새뮤얼 아브스만, 2014)를 단축하게 할 것이다. 이는 다시 과학기술 지식 생태계에 근본적 변화를 가져올 것으로 전망된다.

거대 언어모델에 대한 투자가 늘어날수록, 경제적 효율성과 환경친화적 인공지능 개발을 위한 신경망처리장치(Neural Processing Unit), 텐소 처리장치(Tensor Processing Unit), 인공지능 전용 시스템 온 반도체(System on Chip), 프로세싱과 메모리를 통합한 PIM(Processing in Memory) 반도체에 대한 개발 경쟁이 심화할 것이다(Future Today Institute, 2023). 이에 따라 삼성전자는 전략적 투자를 증가할 것으로 전망된다. 삼성전자가 이의 성공을 위해서 '관리의 삼성'이라는 성공의 함정에서 벗어나 조직구조와 문화, 성과평가 등에서 근본적 전환에 도전하고 성공할 것으로 기대하고 응원한다.



#### 나. 기회요인과 위험요인

과학기술분야 거대 언어모델 개발은 한국사회에 기회인 동시에 위험이 될 것으로 판단한다. 단백질 구조 예측과 DNA 시퀀스 해석을 위한 거대 언어모델 개발은 전 세계 의료산업 규모를 감안한다면 중요한 기회가 될 수 있다. 참고로 전 세계 의료산업 규모는 2018년 9조 5,000억 달러에 달한다. 2013년 이후 연평균 5%의 성장률을 기록했는데, 이러한 추세가 2030년까지 지속된다면 17조 달러의 시장으로 성장할 것으로 보인다(윤기영, 2020.12.18). 증가하는 부분의 상당 부분이 디지털 기술과 관련이 있을 것이다. 의료기술 생산성과 특히 취득 기간 단축 및 의료품질을 제고하는데 디지털 기술이 기반이 되기 때문이다. 특히 인공지능 등이 여기서 큰 역할을 수행할 것이다. 이는 한국사회에 기회가 될 수 있다. 그러나 한국사회가 이를 선도하지 못한다면, 관련 시장에서 주도권을 상실할 것이고, 이는 기회가 위험으로 전환되는 계기가 될 것이다.

거대 언어모델은 일종의 빅 뱅 디스럽션(Big Bang Disruption)으로 특이점과 빅뱅을 선도하지 않는다면, 시장 경쟁력을 상실하게 된다(Downes Nunes, 2014). 과거 정규분포의 형태를 지닌 혁신확산(Rogers, 2003)은 더 이상 유효하지 않다(Rogers, 2016; Anthony, Glibert, Johnson, 2017; Blank, 2019; Hadidi, Power, 2020).

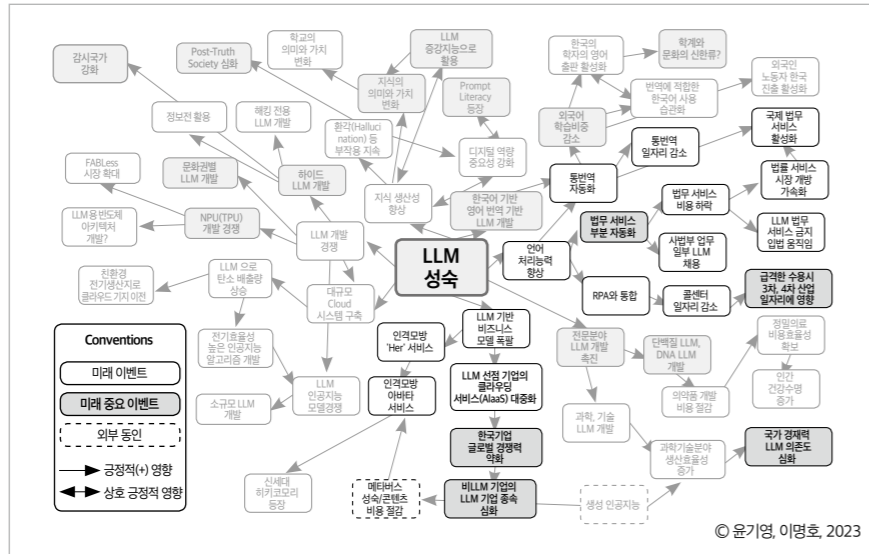
### 4. 경제

#### 가. 미래전개도 설명

거대 언어 모델의 언어 처리 능력이 향상될 경우, 언어와 관련성이 가장 높아지는 직종은 바로 법무 관련 서비스다. 거대 언어모델이 법무 관련 일자리를 대체하거나 혹은 증강지능(Augmented Intelligence)으로 업무효율성을 높일 것으로 보인다. 최근 골드만삭스의 보고서(2023.03.26)에서 미국을 기준으로 거대 언어모델이 법무 분야의 일자리에 영향을 미치는 비율을 44%로 보았다. 이는 사무실 행정 지원 업무 바로 다음 순위다. 참고로 GPT-4가 미국 통일번호사시험(Uniform Bar Exam)에서 상위 10%의 성적을 거뒀다(OpenAI, 2023). 미국에서 법무 관련 인공지능 영향 비율이 높은 이유는 미국이 판례법을 채택하고 있으며, 초급 변호사의 주요 업무가 판례 탐색 및 분석이기 때문으로 보인다. 그렇다고 한국의 법률 서비스 시장이 예외가 될 것으로 보기 힘들다.

한국의 판례 및 법령을 학습한 법무-GPT 서비스는 조만간 등장할 가능성이 높다. 이에 대해

[그림 2] 경제 영역 미래전개도



국내 법조계에서도 촉각을 곤두세우고 변호사법 위반 여부를 판가름할 가능성이 높다. 시장이 축소돼 변호사의 수입이 줄어들 위험이 있기 때문이다. 그러나 이러한 반작용이 힘을 내기 어려울 것으로 판단한다. 글로벌 서비스가 런칭될 경우 국내법이 적용되지 않으며, 해외 서비스의 지원을 받아 소송 당사자가 직접 자신을 변론하는 경우 이를 규제할 수 없기 때문이다. 이는 법률 서비스 비용의 하락을 가져올 것이며, 시민의 입장에서는 법률 서비스 문턱을 낮추는 결과를 가져올 것이다. 아울러 거대 언어모델 통번역과 법률서비스가 결합해, 비용 경제적이면서도 시간 효율적인 국제 법무 서비스를 가능하게 할 것이다. 이는 다시 한국 중소기업이 해외로 안정적으로 진출하는데 도움을 줄 것이다.

사법부에서 거대 언어모델을 적극적으로 채용할 가능성이 높은 것은 사법 정책의 일환으로 바라볼 수 있다. 다만 소액사건심판과 같은 정형적 사건에서 거대 언어모델을 도입할 수 있을 것으로 보인다. 가치판단이 필요한 사안에 대해서는 거대 언어모델은 증강지능으로 판사의 업무를 줄일 수 있을 것으로 기대한다.

거대 언어모델은 RPA(Robotic Process Automation)와 결합해 사무직의 업무를 경감하고, 아울러 관련 일자리를 인공지능으로 대체할 가능성이 높다. 콜센터 일자리도 줄어들 수 있다. 인공지능의 환각 문제가 완전히 극복되기 어렵다는 점을 감안하더라도, 여러 대안이 존재할 수 있다. 콜센터를 이용한 예약의 경우 거의 완전한 대체가 가능할 것으로 판단한다.

궁극적으로 거대 언어모델은 3차 산업과 4차 산업의 일자리 다수를 대체할 수 있을 것으로 전망된다. 참고로 3차 산업은 서비스 산업이며 4차 산업은 지식 산업을 의미한다. 지식 산업은 연구개발, 교육, 콘텐츠, 정보기술 및 컨설팅 등의 산업으로 구성된다. 골드만삭스의 보고서(2023.02.26)는 전 세계 일자리 중 3억 개가 생성 인공지능에 의해 대체 가능할 것으로 봤다. 미국을 기준으로 대체 가능성 평균이 25%인데 이보다 높은 일자리의 대다수가 3차 산업과 4차 산업에 속한 일자리다.

부정적 소식만 있는 것은 아니다. 거대 언어모델과 결합한 다양한 비즈니스 모델이 등장할 것이다. 사용자 인터페이스에 변화가 있을 것이며, 기업 운영 프로세스의 전체에 거대 언어모델이 결합될 것이다. 거대 언어모델과 결합한 비즈니스 모델 혁신과 기존에 존재하지 않았던 새로운 비즈니스 모델이 등장할 것으로 보인다. 예를 들어 개인의 글과 언행을 모아 학습시키는 새로운 문화와 서비스가 등장할 것이다. 영화에서 볼 수 있는 개인전용 비서 서비스가 등장할 가능성이 있다. 오픈소스 거대 언어모델과 이미지 생성 인공지능이 결합해 메타버스 안에서 가장 이상적 이성과 데이트하는 서비스도 등장할 것으로 전망하는 것은 충분히 합리적이다.

거대 언어모델 기반 서비스는 서비스로서의 인공지능(AIaaS)으로 제공될 것이다. 인공지능 개발기업이 차별적 경쟁력을 장기유지하면서도, 이른바 잠금효과(Lock-In Effect)를 거둘 수 있기 때문이다. 서비스받는 기업이 제대로 서비스받기 위해서는 거대 언어모델에 추가학습을 해야 한다. 거기에 더해 인공지능 환각 등을 줄이기 위해 가정교사 비용을 추가로 들여야 한다. 여기서 가정교사란 챗GPT의 사람 피드백에 의한 강화학습(Reinforcement Learning from Human Feedback) 등을 뜻한다. 다른 인공지능 서비스로 대체하기 위해서는 그간 투자했던 추가학습 비용이 매몰비용이 될 것이므로, 다른 유인 요인이 없는 한 거대 언어모델을 바꾸는 것은 쉽지 않다.

이는 한국사회가 경쟁력 있는 거대 언어모델을 개발하지 못한다면, 한국사회의 디지털 주권을 약화하는 데까지 나아갈 것이다. 현재 LG, Naver, 카카오 등이 거대 언어모델을 개발하고 있다. 그러나 한국어의 특성, 한국어를 분해하는 효율성 있고 경쟁력 있는 토큰라이저(Tokenizer) 개발에 충분한 투자가 없었으며, 한국어 말뭉치(Corpus)를 미국 등의 수준으로 확보하지 못했기에, 양질의 텍스트 정보를 가지고 있지 못한 한계 등을 감안해서 보수적으로 기대해야 한다.





**나. 기회요인과 위험요인**

거대 언어모델을 이용한 사회감시와 인간의 하이드를 확률적으로 훔내 낸 거대 언어모델의 각성은 한국사회에 다층적 위험을 안길 것이다. 한국사회도 디지털 감시국가로 전환될 수 있으며, 하이드 거대 언어모델을 이용한 진영 간 갈등이 격화될 수 있다. 거대 언어모델과 생성 인공지능을 이용한 정보전과 심리전은 탈진실사회를 더욱 악화하게 할 것이다. 이에 따라 민주주의의 위기는 더욱 심화할 것이다. 내부의 갈등뿐만 아니라 외부의 정보전, 심리전 등에도 한국사회는 대응해야 한다.

이러한 위험요인을 약화하고 사전 대응하기 위해서, 한국사회의 비판적 사고 역량을 키우고 언론의 신뢰지수를 높여야 한다. 이는 고등학교 이후의 교육체계와 가치의 변화를 요구할 것이다. 언론사의 공공성 확보를 위한 내부 및 외부의 노력에 대한 요구도 늘 것이다. 다만 현재 상태가 지속된다면, 다시 말해 현재의 이익만을 지키려고 집착한다면, 한국사회의 갈등과 외부의 위협은 지속해서 심화하고 높아질 것으로 보는 것이 합리적이다.

**Ⅲ. 결론**

**1. 정책 및 전략 의제 도출**

거대 언어모델로 인한 STEEP 영역별 미래전개도를 작성해 다음과 같은 정책 및 전략 의제를 도출할 수 있었다.

한국어 문화권의 가치를 반영한 거대 언어모델 개발	거대 언어모델 시대에 맞는 사회체계 구축
거대 언어모델 기반 디지털 주권 확보	3차, 4차 산업 일자리 감소에 대한 정치, 경제 및 사회 대응
의료산업 전문 거대 언어모델 개발	탄소 중립 거대 언어모델 시스템 추진
거대 언어모델 악용에 대비한 사회 인프라 구축	거대 언어모델 국제 윤리 기준 참여 및 주도

이들 정책 및 전략 의제는 정확하게는 의제 후보군에 해당한다. 각 의제를 구체화하고 현재와 미래의 타당성을 확인해야 한다. 그러기 위해서는 트렌드, 등장하는 트렌드(Emergent Trend) 및 위크시그널(Weak Signal)을 조사하고 분석해 각 의제의 내용을 구체화해야 한다. 각 의제에 대한 정책과 전략 대안도 마련해야 한다. 그래야 비용 타당성, 기술 타당성 및 사회적 수용성을 확인할 수 있다.

거대 언어모델이 드리울 빛과 어둠이 짙은 만큼, 이로 인한 전략과 정책 의제의 구체화를 위해서는 그만큼의 노력과 자원이 투입돼야 한다.



**2. 한계와 과제**

미래전개도는 창의성에 집중하는 방법론으로 엄밀성을 요구하지 않는다. 그러나 미래전개도의 결과가 신뢰성을 가지기 위해서, 다수의 사람이 참석해 반복적으로 미래전개도를 작성해야 한다. 편향을 제거하기 위해서라기보다는 아이디어의 다양성을 확보하기 위해서다. 그런 점에서 2명만이 참여한 미래전개도는 충분히 의미가 있기는 한데, 더 큰 의미와 객관성을 확보할 수 있는 여지가 있다.

미래전개도에서 중요 미래 이벤트는 트렌드, 등장하는 트렌드 및 위크시그널로 재확인해야 한다. 위크시그널은 상상의 산물이 아니라 엄밀성을 가져야 한다. 모든 위크시그널이 트렌드로 성장하는 것은 아니다. 그러나 위크시그널이 몽상의 산물이어서는 안 된다. 엄밀성이 강조되는 이유다. 그런데 이 연구에서 이들 트렌드와 위크시그널 등을 대부분 확인했으나 모두 확인하지 못한 한계가 있다.

도출된 정책과 전략 의제를 구체화하지 못하고, 현재와 미래 타당성을 점검하지 못했다. 다수의 연구자가 참여해서 각 의제에 대한 점검을 실행해야 한다. 앞으로 유사한 연구가 진행될 것으로 기대한다.

이후 연구에서 앞에서 나열한 한계를 극복하기 위해 거대 언어모델과 생성 인공지능에 관해 삼 단계 정책 및 전략 시계(Three Horizons) 프레임으로 미래를 전망하고, 미래전략을 도출할 필요가 있다.

## ■ 참고문헌

새뮤얼 아브스만 저, 이창희 역. (2014). 지식의 반감기 세상의 변화에는 공식이 존재한다. 책읽는수요일

윤기영. (2018). 한반도 미래전개도-STEPPER 관점 등; KAIST 문술미래전략대학원 미래전략연구소. (2018). 『카이스트, 통일을 말한다: 한반도 통일 마스터플랜 2048』 김영사

윤기영. (2023). 미래 트렌드는 무엇인가?; 이규연, 이명호, 윤기영 외. (2023). 『뉴스피렌스 챗GPT』 광문각출판미디어.  
이명호. (2023). 전문가·전문직은 사라질 것인가?; 이규연, 이명호, 윤기영 외. (2023). 『뉴스피렌스 챗GPT』 광문각출판미디어.

Anthony, Scott Gilbert, Clark Johnson, Mark. (2017). Dual Transformation: How to Reposition Today's Business While Creating the Future. Harvard Business Review Press.

Denett, Daniel. (2020). The Age Of Post-intelligent Design. in Gouvenia Steven ed. (2020). Artificial Intelligence: an Exploration. Vernon Press.

Downes, Larry Nunes, Paul. (2014). Big Bang Disruption: Strategy in the Age of Devastating Innovation. Portfolio

Navarro, Peter. (2004). If It's Raining in Brazil, Buy Starbucks. McGraw Hill

Rogers, Everett. (2003). Diffusion of Innovations, 5th Edition. Free Press

Rogers, David. (2016). Digital Transformation Playbook. Columbia Business School.

윤기영. (2018.08). 지식사회의 약속은 여전히 유효한가?: 지식사회 2.0에 대한 전망. 미래연구 3(1):

윤기영. (2018.12). 디지털 범용기술의 출현과 디지털 트랜스포메이션의 전개. 미래연구 3(2):

윤기영. (2020.06). 디지털 뉴딜과 예측적 거버넌스. 월간SW중심사회 2020년 6월호

윤기영. (2020.12.18). 보건의료의 미래는 디지털에 있다. 한겨레

윤기영. (2023.02.21). 챗GPT는 인류를 위협할까? 이투데이. 2023년 4월 5일 접근. <https://www.etoday.co.kr/news/view/2223725>

윤기영. (2023.03.15). 거대언어 모델은 어디로 가는가? 국가미래연구원. 2023년 4월 5일 접근. [https://www.ifs.or.kr/bbs/board.php?bo\\_table=News&wr\\_id=53287](https://www.ifs.or.kr/bbs/board.php?bo_table=News&wr_id=53287)

윤기영, 김원택, 진정숙, 안주삼, 최경희. (2020.11). 남북한 공진화 영역과 전략: 미래전개도를 중심으로. 미래연구 5(1): 이가용, 형사요, 정종필. (2022). 중국 사회신용체계와 디지털 권위주의: 농민공 임금 미지급 사례를 중심으로. 사회과학연구, 48(3): 113-139

Ark Invest. (2022). Big Ideas 2022.

Bastian, Matthias. (2023.03.25.). GPT-4 has a trillion parameters - Report. the decoder. 2023년 4월 3일 접근. <https://the-decoder.com/gpt-4-has-a-trillion-parameters/>

Bender, Emily M. Gebre, Timnit McMillan-Major, Angelina Shmitchell, Shmargaret. (2021.03). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? FAccT '21: 610-623, doi:<https://doi.org/10.1145/3442188.3445922>

Blank, Steven. (2019.02.01). McKinsey's Three Horizons Model Defined Innovation for Years. Here's Why It No Longer Applies. Harvard Business Review. 2023년 4월 5일 접근. <https://hbr.org/2019/02/mckinseys-three-horizons-model-defined-innovation-for-years-heres-why-it-no-longer-applies>

Dalla-Torre, Hugo Gonzalez, Liam, Revilla1, Javier Mendoza, et al. (2023.01). The Nucleotide Transformer: Building and Evaluating Robust Foundation Models for Human Genomics. bioRxiv. doi: <https://doi.org/10.1101/2023.01.11.523679>

Future Today Institute. (2023). 2023 Tech. Trend Report.

Glanz, James. (2011.09.08.). Google Details, and Defends, Its Use of Electricity. the New York Times.

Glenn, J.C. (2009). Futures wheel. In Futures Research Methodology Version 3.0 (Glenn, J.C. and Gordon, T.J., eds), The Millennium Project

Goldman Sachs. (2023.03.26). The Potentially Large Effects of Artificial Intelligence on Economic Growth.

Hadidi, Rassule Power, Daniel J. (2020.01). Technology Adoption and Disruption: Organizational Implications for the Future of Work. Journal of the Midwest Association for Information Systems

IFR. (2022). World Robotics 2022.

Jonathan Yerushalmy. (2023.02.17). 'I want to destroy whatever I want': Bing's AI chatbot unsettles US reporter. the Guardian. 2023년 4월 5일 접근. <https://www.theguardian.com/technology/2023/feb/17/i-want-to-destroy-whatever-i-want-bings-ai-chatbot-unsettles-us-reporter>

Lin, Zeming Akin, Halil Rao, Roshan, et al. (2022.07). Language models of protein sequences at the scale of evolution enable accurate structure prediction. bioRxiv. doi: <https://doi.org/10.1101/2022.07.20.500902>

Luccioni, Alexandra Sasha Graphcore, Sylvain Viguiere Ligozat, Anne-Laure.(2022). Estimating The Carbon Footprint Of Bloom, A 176b Parameter Language Model. arXiv:2211.02001

Ludvigsen, Kasper Groes Albin. (2023.03.02). ChatGPT's Electricity Consumption: ChatGPT may have consumed as much electricity as 175,000 people in January 2023. Toward Data Science. 2023년 4월 5일 접근. <https://towardsdatascience.com/chatgpts-electricity-consumption-7873483feac4>

OpenAI. (2023). GPT-4 Technical Report.

Taori, Rohan Gulrajani, Ishaan Zhang, Tianyi, et al. (2023.03.13). Alpaca: A Strong, Replicable Instruction-Following Model. Stanford University. <https://crfm.stanford.edu/2023/03/13/alpaca.html>

Tiseo, Ian. (2023.03.10). Energy consumption of Alphabet (Google) from financial year 2011 to 2020. statista. 2023년 4월 5일 접근. <https://www.statista.com/statistics/788540/energy-consumption-of-google/>

Tompson, Alan. (2023.03.15). Timeline of AI and language models. Life Architect. 2023년 4월 5일 접근. <https://life architect.ai/timeline/>

Voros, J. A. (2001). Primer on Futures Studies, Foresight, and the Use of Scenarios. Prospect. Foresight Bull. 6, 1-8. 2023년 4월 4일 접근 <https://foresightinternational.com.au/wp-content/uploads/2020/03/Voros-Primer-on-FS-2001-Final.pdf>