

챗지피티, 인공지능(AI)과 민주주의

이광석

서울과학기술대 교수

이광석은 서울과학기술대학교 교수이며, 비판적 문화이론지 『문화/과학』의 편집인이다. 주요 연구 분야는 기술문화연구, 커먼즈, 플랫폼, 기술 생태정치학, 자동화사회 등이다. 지은 책으로는 「디지털 폭식 사회」, 「피지털 커먼즈」, 「포스트디지털」, 「디지털의 배신」, 「데이터 사회 비판」, 「디지털 야만」 등이 있다.

챗지피티, 인공지능(AI)과 민주주의

이광석
서울과학기술대 교수

이 글은 인공지능이 우리 사회 민주주의에 미치는 영향을 크게 두 측면에서 우려한다. 하나는, 인간 사회관계를 '조회수' 알고리즘을 통해 '소셜(미디어)' 관계로 흡수하고 대체하는 경향이다. 다른 하나는, 생성형 AI의 빠른 확산과 범용화로 인해 인간의 사유 과정과 대상 세계의 탐구 행위가 자동화 프로세스에 위탁되는 경향이다. 인공지능은 기술 효율과 편리에도 불구하고, 한편으로 인간과 인간이 서로 물리적으로 소통하고 관계 맺는 과정을 전산학적 알고리즘 연결로 대체하고, 다른 한편으로 온라인 밖 대상 세계로부터 무엇인가 자극받고 탐색하는 인간의 공감과 성찰 능력을 탈취하는 경향이 있다. 민주주의적 실천이 나와 연결된 '나들'과 함께 공존하고 연결되어 있다는 의식에서 비롯한다면, 동시대 인공지능이 디자인하려는 사회 설계 방향은 크게 어긋나고 있다고 볼 수 있다. 궁극적으로, 이 글은 인공지능을 물신화하거나 지능기계를 혐오하는 양극단을 경계하면서도, 인공지능의 민주주의적 전환을 위한 사회적 숙의 과정의 중요성을 강조한다.

인간을 닮은 지능기계를 보는 기대와 불안

인공지능(AI)이 야기할 삶의 변화에 전 세계가 들썩인다. 무엇보다 챗지피티(ChatGPT) 등 '생성형(generative)' 인공지능의 출현이 마치 블랙홀처럼 세간의 거의 모든 애깃거리를 집어삼키고 있다. 인류가 배태한 텍스트와 이미지를 대량 기계 학습해 이로부터 생성된 글과 이미지를 쏟아내는 생성형 AI 기술에 다들 탄복하고 있다.

정확히 얘기하면, 이 신기술에 대한 대중의 과열된 흥분에는 불안과 기대감이 혼재해 있다. 우선 불안의 측면에서 보자면, 어떤 이에게 생성형 AI는 인간 그 이상의 능력을 지닌 인공지능체의 출현을 상징한다. 다른 이에게는 이 새로운 인공지능이란 인간의 인지(정신) 노동의 자동화로 사라질 수많은 일자리를 뜻하기도 하다. 또 다른 이에게는 더욱 정교해져 그 진위조차 판독 불가능해진 그럴듯한 '가짜' 창작과 지식이 창궐하는 새 국면으로 비춰지기도 한다. 이처럼 여러 연유로 생긴 불안에도 불구하고 인공지능의 가속화에 많은 이들이 대체로 낙관적인 듯싶다. 불안의 꼭 정반대의 이유로, 누군가의 물음에 답하고, 번역하고, 시와 소설을 쓰고, 글과 기사를 작성하고, 원하는 이미지를 생성하고, 컴퓨터 코드를 짜고, 파워포인트를 만

들고, 텍스트를 요약하는 등 생성형 AI 기술이 주는 새로움이 과할 정도의 대중적 열광으로 증폭되고 있다.

챗지피티와 같은 생성형 AI의 이전과 다른 가장 큰 특징이라면, 전문가뿐만 아니라 이제 일반인의 지식 생산과 창의 활동에 결합되고 부착되는 인공지능의 '범용성'일 것이다. 검색 기능을 대신해 인공지능에 묻거나, 플러그인된 앱과 소프트웨어 프로그램을 쓰거나, 글을 쓰면서도 코파일럿 형태로 작동하는 생성형 AI를 늘 곁에 두면서 그것이 만들어낸 결과 값을 어디든 덧대고 응용할 수 있는 일상이 곧 열릴 것이다.

편리나 효율성의 견지에서 보면 신생 인공지능이 바야흐로 지식 생산과 창작의 범용 기술이 되는 것이 당연해 보인다. 하지만, 우리가 마주하는 인공지능의 세례와 축복에 비해 아직은 그것이 향후 일상 삶과 우리의 의식에 미칠 영향에 대해 그리 쉽게 낙관하기에 선부르다. 앞으로 챗지피티류의 인공지능을 마치 거의 전 국민이 가입한 카톡처럼 쓰게 된다면 인간에 어떤 변화를 야기할까?

이 글은 최근 인공지능이 민주주의에 미치는 영향을 크게 두 측면에서 우려한다. 1) 인간의 사회관계를 '조회수' 알고리즘을 통해 '소셜(미디어)' 관계로 흡수하고 대체하는 경향이다. 2) 생성형 AI의 빠른 확산과 범용화로 인해 인간의 사유 과정과 대상 세계의 탐구 행위가 자동화 프로세스에 위탁되는 경향이다. 다시 말해, 인공지능은 그 수많은 효율과 편리에도 불구하고, 한편으로 인간과 인간이 서로 물리적으로 소통하고 관계 맺는 과정을 전산학적 알고리즘 연결로 대체하고, 다른 한편으로 온라인 밖 대상 세계(사물)로부터 무엇인가 자극받고 탐색하는 인간의 공감과 성찰 능력을 탈취하는 경향이 있다. '똑똑한' 기술이 인간의 편리를 돕겠다고 나서고 있지만, SNS 알고리즘 자동 추천과 '필터버블' 효과가 사회 갈등과 극단의 정치를 낳고 있는 한편으로, 생성형 AI는 다른 한편에서 대상 세계에 대한 인간의 무관심과 '사물의 소멸'¹을 야기할 모양새다.

1 한병철, 『사물의 소멸 - 우리는 오늘 어떤 세계에 살고 있나』, 전대호 옮김, 김영사, 2022.

인공지능의 야만성과 '조회수' 알고리즘 정치

사실상 인공지능 기술은 챗지피티 출현 이전에도 이미 다른 모습으로 우리 곁에 스며들고 있었다. 그렇지만 이제까지의 응용 사례들은 그리 유쾌하지만은 않았다. 오늘날 자본주의 사회에서 가동되는 인공지능의 민낯은 생각보다 비릿하다. 가령, 플랫폼 배달 노동자는 인공지능 알고리즘에 마치 아바타처럼 속박된 채 노동을 수행한다. 주문 앱을 통한 이용자 손끝의 별점들은 영세업자에게 마치 비수처럼 꽂힌다. 소셜미디어 인증 샷과 누군가의 댓글에 부동산 프

2 에이미 추아, 『정치적 부족주의』, 김승진 옮김, 부키, 2020, 17-18쪽.

리미엄 가격이 널뛰기하고 젠트리피케이션이 야기되는 현실은 또 어떠한가. 자본주의의 인공지능 기술로 매개된 알고리즘 논리가 우리 사회를 뒤덮고 있다.

인공지능 알고리즘 추천 기술에 의해 편향된 콘텐츠 소비가 이뤄지면서 우리는 상호 극단적 배제와 증오로 뿔뿔 뭉친 ‘정치적 부족주의’를 또한 마주하고 있다.² 정치적 부족주의는 전통의 계급과 정체성의 정치를 증오와 혐오의 정치로 환원하는 우익 포퓰리즘의 양상으로 나타났다. 우익 포퓰리즘의 논리는, 비정규직 노동자, 여성, 장애인 등 사회 약자와 타인을 표적 삼고 그들을 배제하는 혐오 발언과 폭력 심리를 내(內)집단의 동일시 논리로 삼게 하고 이를 부족 응집력으로 고착화한다는 점에서 퇴행적이다.

정치적 부족주의를 키우는 데는 이른바 ‘조희수’ 알고리즘이 깊게 자리한다. 우리의 온라인 대화와 행위는 닷컴의 SI 알고리즘 계산식에 갈수록 크게 좌우되고 있다. 즉 일상 사회관계가 지난 십여 년 사이 ‘소셜(미디어)’ 관계로 크게 대체되어왔다. 사적 소통과 연결은 물론이고 사회적, 공적 의사결정이나 연결망으로 빅테크 소유의 소셜미디어를 사회관계적 말길로 채택해왔던 까닭이다. 이는 우리에게 소통의 편리와 효율을 제공하기도 했지만, 궁극에는 수많은 왜곡된 데이터의 흐름을 야기해왔다. 특히 주목을 끌어 조희수를 모으고 구독을 높이는 일이 곧 돈과 직결되면서, 관계적 흐름을 왜곡하는 현상이 크게 증가했다. 가령, 조희수와 클릭을 유도하기 위해, 자극적 소재의 콘텐츠는 물론이고 혐오, 편견, 극단의 정서까지도 불러일으킬 정도로 인플루언서라 불리는 이들은 자신의 동조 세력을 ‘구독’과 ‘좋아요’로 끝없이 끌어모으려 한다.

3 김병권, 『진보의 상상력』, 이상북스, 2021, 227쪽.

‘주목’과 ‘조희수’에 의해 구축된 왜곡된 ‘소셜(미디어)’ 관계는 궁극에 우리의 비판적 사유와 판단을 후퇴시킨다. 일종의 “시장경제를 닮아가는 정치”를 만들어낸다.³ 가령, ‘알고리즘 정치’는 내 동지를 팔로워와 폐북 친구로, 적과 ‘그들’의 구분은 언팔과 폐절된 이들로, 정치 저항은 인증 ‘셀카’와 해시태크로, 연대는 좋아요와 구독으로, 사유와 판단은 알고리즘 추천 등으로 쉽게 치환된다.⁴ 정치는 취향으로 쪼그라들고, 그런 취향조차 알고리즘 채널로 굳어지고, 경합과 저항의 미디어 공론장은 관중과 ‘어그로꾼’(남들의 시선을 붙들어두는 것에 집착하면서 상징 폭력의 극적 희생양을 찾기 위해 항시 몰두하는 악질의 인플루언서)에 의해 무력화된다. 그렇게 우리의 물리적 사회관계를 대체해 SI 알고리즘이 구성하는 배열과 가중치에 의해 축조된 전산 공학적 관계가 지배적으로 되면서, 사실상 시민들 스스로 민주적 숙의나 사회적 공동체 감각을 배양하기 어려워진다.

4 김곡, 『과잉존재』, 한겨레출판, 2021, 154쪽.

생성형 AI의 범용화가 미칠 민주주의의 새로운 위협

5 이광석, 『챗봇 '이루다'가
우리 사회에 남긴 문제:
인공지능에 인권 메뉴얼
탑재하기』 『문화/과학』
105호, 2021, 183-198쪽.

6 니콜라 부리오,
『포스트프로덕션』
정연삼·손부경 옮김,
그레파이트온핑크, 2016.

기술적으로 보면, 챗지피티와 같은 인공지능은 거대언어모델(LLM)을 기반으로 초거대 데이터를 기계학습(머신러닝)해 맞춤형 문장들을 자동 생성하는 언어 예측 알고리즘 모델이다. 다시 말해, 지난 수십 년간 축적된 인간의 디지털 지식과 창작 데이터를 무차별적으로 대거 습득한 후 사용자가 원하는 특성의 프롬프트형 질문에 맞춤형 답을 제시하는 자동화 장치이다. 이전 시점의 데이터를 기반으로 해 특정 단어의 다음에 올 단어를 확률적으로 추론해 문장들을 자동 생성해내는 패턴을 갖추고 있다.

보통 대상 물질세계나 온라인 세계의 양계(兩界)로부터 이해와 통찰을 구하는 인간의 학습 방식과 달리, 생성형 AI의 기계학습은 굳이 외부 물질세계와 사물들에 대한 이해를 필요로 하지 않는다. 더불어, 생성형 AI는 확률 기반형 인간 언어·문장·이미지 등 확률 추론으로 생성된 작업물의 에러 비율을 낮추기 위해 인간의 수작업을 통해 오류를 잡고 피드백을 주는 '강화학습'을 추가한다. 강화학습이란 인간의 단순반복 '미세노동(microwork)'을 통한 생성된 결과의 텍스트 오류나 결함을 잡아내는 후작업에 해당한다.

01. 주류 세계관의 과잉 대표

챗지피티나 미드저니 등 텍스트·이미지 생성형 인공지능의 과열 조짐은, 국내에서 대화형 '챗봇 이루다'⁵가 극단의 혐오와 편견을 실어 나르며 개인 민감 정보를 여과 없이 노출해 사회 이슈가 됐던 불과 일 이년 전과 사뭇 달라 보인다. 물론 그 기능의 세련됨에도 불구하고 생성형 AI의 '기술적 결함'들은 또 다른 형태로 발견된다. 가령, 사용자가 인공지능을 일부러 속이는 할루시네이션(환각)에 취약하거나, 헛소리를 그럴듯하게 포장하거나, 몇 회 이상의 집요한 대화로 젠더 편향 및 편견 에러를 생성하거나, 누군가의 이용자 (결제) 정보가 흘러나오거나, 기업 보안사항이 유출되기도 했다. 이들 문제는 지속적인 '증강 데이터'나 '강화학습'의 보완을 통해 시간이 지남에 따라 부분적으로나 형식적으로 '중화(中和)'될 기술 결점들로 보는 편이 낫다. 오히려 우리가 향후 크게 주목해야 할 것은 이들 생성형 AI가 지닌 근원적 모순과 민주주의 사회의 작동에 미치는 부정적인 문제에 있다.

생성형 AI는 일종의 거대 '리믹스'(무수한 원본 이미지와 텍스트를 뒤섞어 변형해 새롭게 창작하는 행위) 자판기와 흡사하다. 이같은 리믹스 자동 제조기 시대의 도래로 말미암아, 기성의 텍스트와 이미지, 오디오 등을 재조합해 창의적 행위의 결과물을 얻는 '포스트-프로덕션'⁶ 하위문화 혹은 문화정치 또한 이제 인공지능에 의해 쉽게 거세될 운명이 되었다. 왜냐면 이제 단 몇 초면 근사한 문장과 생성 이미지가 똑딱 만들어지는 리믹스 과소비 세계에 우리 대부분이 곧 입문할 까닭이다. 리믹스가 자동화 연산 처리에 의해 점차 대체되는 현실에서는, 리믹

스 행위의 그 어떤 미학적·창의적 혁명성도 상실하게 된다. 오히려 대상 세계와 분리된 흐릿한 디지털 사본들 사이의 상호 참조와 변주로 생성된 자동 창작 지식의 온라인 세계 내 과포화 상태가 절정에 이를 것이다.

생성형 AI는 이로 인해 그것이 이미지이건 텍스트건 태생적으로 인간 사회의 주류 의식과 규범적 사고의 확률적 평균값을 대규모로 변주해 쏟아낼 것이다. 이들 대부분은 기성 세계의 주류 논리와 의미 구도를 빼담을 수밖에 없다. 인류 지식과 정보의 거대 빅데이터를 원료로 갈아 넣은 생성형 AI 자판기가 토해내는 전산 확률적 구성물이 차고 넘칠수록, 인간 사회의 표준화된 세계관이나 규범이 과잉 대표될 확률이 높다는 얘기다. 인간이 물질·비물질 대상 세계를 모사하고 탐색해 만드는 지식 콘텐츠보다 이 생성형 자동화 기계들이 마구 쏟아내는 리믹스 콘텐츠가 표상의 세계를 압도하는 때가 훨씬 빠르게 도래할지도 모른다.

마치 이는 이미지 생성 AI의 자동화된 ‘리믹스’ 공정으로 보자면, 특정 가중치 값이 과해 유명 작가, 화풍, 사조의 패턴화된 색감과 질감이 이미지 결과물에 두드러지는 것과 유사하다. 텍스트 생성형 AI로 보면, 특정 문체, 패턴, 서술 구조, 세계관이 여러 문장 속에서 건조하게 변주되며 무한 복제되는 현상과 같다. 문제는 프롬프트형 질문과 답변이 일상 속에 흔하게 기입될 때, 주류 세계관의 과잉 표상 속에서 과연 소수 의견과 타자성의 관점이 살아남을 수 있을까 하는 의문이다. 젠더 편향이나 인종·약자 혐오 등이 이와 같은 자동화된 세계관에서는 사회적 논쟁이나 주장 없이 강화학습에 의해 수면 아래 가라앉거나 뻘한 확률적 모범 답안이 제공될 공산이 크다.

02. 사유 과정의 자동화와 퇴화

적어도 검색의 시대는 온라인에서 무언가 찾고 추리하는 행위를 통해 스스로 정리하고 생각하는 힘을 인간 그 자신에 귀속시켰다. 이제 생성형 인공지능에 묻고 답하는 시대는 그 자신이 풀어야 할 물음을 자주 생략하고 이를 지능기계에게 쉽게 위탁해버린다. 이미 구축된 지식과 데이터의 생성형 지능 자판기로부터 맞춤형 답을 찾고 인간 자신의 사유 프로세스를 대체하는 ‘구매’ 습관에 익숙해지면, 일정 부분 사유하는 힘과 탐구 능력을 인공지능의 자동화 능력에 자주 위임하게 될 확률이 높다. 질문형 프롬프트조차 로봇 에이전트에 자동화하는 ‘오토-지피티’의 출현은, 인간의 물음조차 생략해 인공지능에 위임하는 가까운 미래를 가정하고 있다.

문장과 이미지 생성의 리믹스 기계들이 어디에든 산재하는 자동화 세상에서는, 직접적인 인간의 감각, 사물과 사건, 타자의 존재, 원본 사이트와 출처 등 대상 세계로부터 캐묻고 찾고 경

인공지능 기술의 민주주의적 전환

7 AI Index Steering Committee, The AI Index 2023 Annual Report, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2023.

8 Future of Life Institute에 의해 발의(2023. 3. 22.)된 공개 청원 페이지 (Pause Giant AI Experiments: An Open Letter)에 한 달도 안된 4월 12일 현재, 2만 여명이 넘는 숫자가 서명했다. 최근 지피티-4 모델 이후의 모든 AI 개발을 향후 6개월 동안 멈춰 세우고 사회적 합의와 규제기구 마련을 요청하는 한 공개서한에 대한 전 세계 수만 명에 달하는 IT 전문가들의 자발적인 지지 성명이 화제다. 다음 페이지 참고: <https://futureoflife.org/open-letter/pause-giant-ai-experiments>

힘하고 따지는 비판적 성찰 과정이 줄어들 공산이 크다. 궁극에 이는 대상 세계와 생명 타자와 함께 공존하고 있다는 사회 공동체 감각인 '공통감각'(sensus communis)을 퇴화시킨다. 이는 사회적으로 인공지능에 묻고 답하는 자동화에 대한 의존도를 높여 상대적으로 대상 세계와의 접촉을 멀리하면서, 현대인의 대상 세계에 대한 감응 능력의 퇴화와 사회화 과정 자체의 '탈속련화'를 이끈다.

민주주의적 실천은 나와 연결된 '나들'과 함께 공존한다는 공통감각을 실천으로 삼는다. 그런데, 외부 세계로부터 유리된 채 지능 기계가 던져주는 인공 디지털 세계의 또 다른 모사로부터 의식의 출발점으로 삼는 일이 흔해지고, 대상 세계와 타자로부터 배우는 에너지와 시간을 마찰비용 정도로 간주한다면 어떠할까? 사물과 물질성에 대한 궁금증과 물음이 프롬프트 자동 명령어로, 아니 로봇 에이전트의 자동 응답으로 대부분 대체되는 세상은 과연 '멋진' 세상일까? 이제 인간에게 주어진 일이란 것이 물질세계와 유리된 '탈물질(탈신체)'의 데이터 질서 내 부유하는 표상들의 탐닉이라면?

인공지능 기술 영역은 지난 2014년부터 학계와 비영리 기관의 연구에서 급격하게 멀어져 이제는 빅테크의 사업으로 굳어졌다.⁷ 기계학습의 가장 근본이 되는 광범위한 데이터 수집 능력과 자본력에 있어서 학계와 비영리 기관이 빅테크에 밀렸다고 볼 수 있다. 개방성에 기반 인공지능 코드 개발 문화와 오픈AI의 개방 철학의 기초 또한 이제 무너졌다. 가령, 오픈AI는 지피티-4 모델부터 모든 데이터셋, 코드, 알고리즘을 전격 비공개하기로 결정했다. 최대 주주인 마이크로소프트(MS)의 욕망이 반영된 결과다.

인공지능의 이렇듯 시장 기술로 가속화 하는 현실에서 우리는 무엇을 할 수 있을 것인가? 기술의 진행 방향에 대한 어떤 사회적 합의 과정이나 사회 영향 평가도 없이 무분별하게 인공지능을 전방위로 도입하는 빅테크의 폭주하는 행보에 제동을 걸어야 한다.⁸ 인공지능이 우리의 일상과 사회 저변에 착근되는 방식과 효과에 대한 근본적인 규제안 마련 및 재검토가 시작돼야 한다.

근원적으로는, 청정의 비물질 혁신 기술로 행세하는 동시대 인공지능을 떠받치는 현실 인프라와 물질성을 드러내는 일이 중요하다. 이를테면, 인간 의식과 생체 데이터의 무차별 추출과 포획, 암흑상자화된 데이터 인클로저(사유화) 질서, 생성형 AI의 연산 처리에 소모되는 과도한 에너지 소비와 탄소 배출, 인공지능의 강화학습에 동원되는 산노동 등 기술 열광으로 쉽게

망각되는 인공지능의 사각지대를 드러내야 한다.

물론 인공지능을 물신화하거나 지능기계를 아예 혐오하는 양극단을 경계해야 한다. 도구적 기술주의를 벗어나 인공지능 기술을 좀 더 민주주의 기획에 합당하게 사회 공학적으로 튜닝할 필요가 있다. 무엇보다 생성형 인공지능이 표준화된 기성의 주류 세계관을 전산 확률 평균적으로 재생산하면서, 사회 소수자의 목소리를 텍스트와 이미지 생성에서 배제하고 왜곡하거나 침묵의 나선 속에 가두려는 민주주의의 위협 상황에 대한 적극적인 사회 개입이 필요하다. 앞으로는 생성형 인공지능의 자동화된 사회적 편향 기제를 면밀히 따지는 일 그 자체가 민주주의의 가치를 지키는 일과 직결될 것이다. 이를 위해서는 인공지능에 의해 왜곡된 사회 관계를 복구하고 사회 자동화의 폭주를 조절하는 시민사회 주도의 대안 논의 테이블을 마련하는 일이 시급하다.

챗지피티, 인공지능(AI)과 민주주의

발행일 2023년 4월 28일

발행처  민주회 운동기금사업회
Korea Democracy Foundation

 한국민주주의연구소
Institute for Korean Democracy

ikd@kdemo.or.kr

경기도 의왕시 내손순환로 132

이 리포트의 내용은 연구자의 견해로서 연구소의 공식 견해와 다를 수 있습니다.



9 772508 324001
ISSN 2508-3244