

# [ 연 수 보 고 서 ]

## AI 저널리즘, 팩트체크의 무기인가 새로운 가짜뉴스 위협인가

-AI 기사의 객관성과 한계, 법적·윤리적 이슈를 중심으로

### — 목 차 —

I. 서론 .....	1
II. 연구목적 및 의의 .....	1
III. 팩트체크 무기로서의 AI .....	4
IV. AI 저널리즘의 위험성 .....	6
V. 결론 .....	10

문화일보 차장 김성훈

연수기관: 미국 조지아대

## I. 서론

이른바 4차 산업혁명 시대의 도래와 함께 사회 전 분야에서 ‘디지털’과 ‘인공지능(AI)’이 유행어가 되고 있다. 필자가 출입하던 전통적 제조기업 중에서도 현대자동차가 AI 전담 조직을 설립해 자연어 처리기술(NLP)을 직접 연구하고 있을 정도입니다. 또 포스코는 ‘AI 제철소’를 도입하는 대전환 작업을 진행 중입니다. 이런 대격변의 현장을 취재하며 저도 AI에 관심을 갖게 됐습니다.

이와 같은 AI의 전방위 확산 흐름에서 언론도 예외가 될 수는 없습니다. 위에서 언급한 NLP는 기사 작성용 AI 알고리즘의 핵심 기술이기도 합니다. 이미 외국 우수 언론들은 뉴스 제작에 AI를 활용하고 있고, 국내에서도 AI가 작성한 기사가 노출되고 있습니다.

## II. 연구 목적 및 의의

### 1. 왜 인공지능(AI) 저널리즘에 주목해야 하는가.

로봇 저널리즘, 자동화(Automated) 저널리즘, 알고리즘 저널리즘, 컴퓨터 이용(Computational) 저널리즘 등 다양한 명칭으로 불리는 AI 저널리즘은 결국 소프트웨어를 활용해 인간의 자연어를 모방, 기계가 자동으로 기사를 작성하는 것을 말합니다.<sup>1)</sup>

AI는 인간보다 훨씬 방대한 양의 데이터를 훨씬 짧은 시간에 수집·분석할 수 있고, 알고리즘에 따라 순식간에 기사를 만들어낼 수 있습니다. 이에 온라인 포털과 SNS가 뉴스를 지배하는 시대에 AI 저널리즘은 정통 매체에게 가장 간단한 대응책이 될 수 있습니다. 하지만 AI가 잘못된 정보를 수집해 입력한다면 허위정보, 가짜뉴스가 급격히 증가할 위험성도 커지게 됩니다.<sup>2)</sup>

특히 AI가 가짜뉴스 생산 및 확대재생산에 악용될 경우, 정당한 출처를 가장해 진짜 같은 텍스트를 생성하거나, 정교한 가짜 영상을 만들어 없었던 일을 사실로 만들어버릴 수도 있습니다. 해외에서는 AI 시대, 당파성을 띤 가

---

1) M Tunez-Lopez, C Toural-Bran, C Valdiviezo-Abad, "Automation, bots and algorithms in newsmaking. Impact and quality of artificial journalism", *Revista Latina de Comunicacion Social*, 74, 2019, p. 1415.

2) Marju Himma-Kadakas, "Alternative facts and fake news entering journalistic content production cycle", *Cosmopolitan Civil Societies: An Interdisciplinary Journal*, 9(2), 2017, pp. 29-38.

짜 정치뉴스가 가져올 민주주의의 위기를 지적하는 연구도 나오고 있습니다.

그러나 AI가 정교한 가짜뉴스를 만들고 급속도로 퍼뜨릴 때 최대 피해자는 오히려 기업과 기업인이 될 수 있습니다. 정치인들은 맹목적 지지층에 기대어 자신에게 불리한 보도를 모두 가짜뉴스로 매도하는 등 오히려 가짜뉴스를 활용하는 존재입니다. 반면 기업과 기업인은 가짜뉴스가 한 번 퍼진 것만으로 이미지가 추락하고 막대한 손실을 입게 될 것입니다.

실제로 2016년 말 ‘힐러리 클린턴이 아동 성착취 조직에 연루돼 있고, 워싱턴 DC 기반 피자 브랜드 지하실에 근거지가 있다’는 가짜뉴스가 퍼진 후, 이를 믿은 한 남성이 피자가게에서 충기를 난사한 사건이 터졌습니다. 펍시콜라는 ‘CEO가 2016년 미국 대선에서 도널드 트럼프 대통령 당선을 비판했다’는 보수 온라인 매체들의 오보 때문에 SNS에서 불매운동이 일어, 1개월 새 주가가 5.5% 하락한 바 있습니다. 스타벅스는 2017년 8월 SNS에서 가짜 광고사진과 함께 ‘스타벅스가 이민서류 미비자들에게 공짜 프라푸치노를 준다’는 가짜뉴스가 퍼져 곤욕을 치렀습니다.

2017년 3월 현대경제연구원은 가짜뉴스로 인한 경제적 피해가 연간 30조 900억 원에 달한다는 보고서를 내기도 했습니다. 매사추세츠공대(MIT) 연구팀이 2006~2007년 트위터에서 300만 명 이상이 총 450만 차례 이상 트윗한 뉴스 12만6000건을 선별, 뉴스가 1500명에게 퍼지는 속도를 비교한 결과 허위정보가 6배나 빨리 확산했습니다.<sup>3)</sup>

또 AI 저널리즘은 빅데이터를 활용해 더욱 개인화된 뉴스를 공급할 수 있는 만큼, 자신이 보고 싶은 것만 보는 트렌드와 결합한다면 AI가 생산한 가짜뉴스의 폐해는 더욱 심각해질 것입니다.

## 2. 현장에 적용된 AI 저널리즘 사례

뉴욕타임스는 ‘에디터(editor)’라는 알고리즘을 활용해 기자들의 기사 작성을 돕고 있고, BBC는 ‘주서(Juicer)’라는 AI 데이터 마이닝(대용량 데이터 속에서 유용한 정보를 발견하는 것) 시스템을 사용하고 있습니다. 블룸버그에서는 웹사이트 게재 기사의 약 3분의 1을 ‘사이보그’라는 AI 시스템이 작성합니다.<sup>4)</sup> 워싱턴포스트는 AI ‘헬리오그래프(Heliograf)’를 통해 고교 풋볼

3) Flores Vivar, "Artificial intelligence and journalism: diluting the impact of disinformation and fake news through bots", *Doxa Communication*, 29, 2019, p. 201.

4) Saad Saad and Talat A. Issa, "Integration or Replacement: Journalism in the Era of Artificial Intelligence and Robot Journalism", *International Journal of Media, Journalism and Mass Communication*, Vol. 6, Issue 3, 2020, pp. 7-8.

게임까지 다룰 수 있게 됐습니다.

LA타임스의 '퀘이크봇(Quakebot)'은 진도 3.0 이상 지진이 발생했을 때 자동으로 기사를 작성하며, 포브스는 데이터를 기사로 바꾸는 소프트웨어 '퀸(Quill)'을 활용해 증권 시황 기사를 만들고 있습니다.<sup>5)</sup> 퀸을 개발한 업체 '내러티브 사이언스'는 오는 2025년이 되면 로봇이 90%의 기사를 작성할 것이라고 주장했습니다.

중국 신화통신은 2018년 세계 최초 AI 앵커를, 2019년에는 최초의 AI 여성 앵커를 선보였습니다.



<그림 1> 신화통신이 2019년 공개한 AI 여성 앵커. 유튜브 캡처

제가 연수한 지역인 미국 조지아주의 유력 일간지 '애틀랜타 저널 컨스티튜션(AJC)'은 머신러닝(Machine Learning) 기술을 활용해 10만 건 이상의 징계 문건을 살살이 뒤져서 의사들의 성적(性的) 학대에 관한 탐사보도를 진행, 2017년 폴리처상 최종 후보까지 오르기도 했습니다.

국내에서는 서울대 언론정보학과 이준환 교수 팀이 개발한 로봇 기자 '나리'가 2017년 제19대 대선 당일 11시부터 다음날 개표 종료 때까지 250여 개의 기사를 작성했고, 170여 개는 실제 SBS에서 보도됐습니다.<sup>6)</sup> 이 교수 팀은 파이낸셜뉴스를 통해 증시 시황 기사도 선보였습니다. 연합뉴스는 AI NLP 기술을 적용해 날씨 보도를 하고 있습니다.<sup>7)</sup>

### 3. 팩트체크와 연계한 국내 AI 저널리즘 연구는 미비

이제는 AI가 자동으로 생산한 기사가 더 높은 호응을 얻는 경우까지 생겨나고 있습니다. 로이터재단과 옥스퍼드대 연구에 따르면, 35세 미만 응답자

5) 이창형, 「AI와 미디어」, 이창형 외 저, 『AI 시대의 미디어』, 북스타, 2020, 122-124쪽.

6) 이윤영-안종묵, 「로봇저널리즘의 사례와 이슈에 관한 탐색적 고찰」, 『커뮤니케이션학 연구』 제26권 2호(2018년 2호), 168쪽.

7) 이창형, 앞의 책, 120-121쪽.

의 65%가 인간이 쓴 기사보다 로봇이 쓴 기사를 더 마음에 들어 했습니다.<sup>8)</sup>

국내에서도 이준환 교수 팀이 기사작성 알고리즘을 상용화했을 정도로 기술적 발전을 이뤘으나, 학문적 탐구의 주제는 아직 언론매체에서 AI가 적용되는 현황이나 ‘AI가 인간의 일자리를 빼앗을 것인가’ 같은 철학적 테마에 집중된 경향을 보입니다. 반면 해외 연구는 AI를 활용한 가짜뉴스와의 전쟁, 정반대로 가짜뉴스 생산에 AI가 악용될 위험성, AI 저널리즘의 윤리 등으로 영역을 넓혀가고 있습니다.

### III. 팩트체크 무기로서의 AI

#### 1. 가짜뉴스 필터로써 AI의 유용성

구글, 페이스북 등 SNS 플랫폼들은 가짜뉴스를 걸러내는 데 AI를 적극 활용하고 있습니다.<sup>9)</sup> 실제로 AI는 가짜뉴스 감지에 효율성을 입증하고 있습니다. 미시건대 연구팀의 알고리즘은 76% 정확도(인간은 70%)로,<sup>10)</sup> 워싱턴대 앨런 AI 연구소가 개발한 ‘그로버(GROVER)’는 최고 92% 정확도로 가짜뉴스를 가려냈습니다.<sup>11)</sup> 듀크대 리포터스 랩(Reporters' Lab)은 자동화되고 빠른 팩트체크 기술을 연구 중입니다.<sup>12)</sup> 시러큐스대에서는 미 국방고등연구계획국(DARPA) 지원을 받아 다량의 진짜·가짜뉴스 데이터를 바탕으로 가짜뉴스를 식별해내도록 AI를 ‘훈련’시키는 프로그램을 개발 중입니다.

머신러닝 기법이 적용된 AI는 몇 가지 기준에 따라 가짜뉴스를 구분해냅니다. 제목에 감정적·극단적 어휘를 사용한 글, 가짜뉴스를 많이 올린 사람의 이름이 달린 글, 취재원이 불분명하거나 허위 출처를 인용한 글, 똑같은 내용이 여러 플랫폼을 통해 게재된 글, 극단적인 정치적 관점을 지닌 글 등은 가짜뉴스일 가능성이 크다고 판단합니다.<sup>13)</sup>

8) Tunez-Lopez et al., 2019, *op. cit.*, p. 1419.

9) Saad and Issa, 2020, *op. cit.*, p. 9.

10) Vivar, 2019, *op. cit.*, p. 206.

11) Rowan Zellers, Ari Holtzman, Hannah Rashkin, Yonatan Bisk, Ali Farhadi, Franziska Roesner and Yejin Choi, "Defending Against Neural Fake News", *arXiv: 1905.12616v2 [cs.CL]*, 29 Oct 2019, p. 2.

12) Jamie Gavin, "Using AI to fight fake news: Duke Reporters' Lab", 22 Aug 2018, <https://whatsnewinpublishing.com/using-ai-to-fight-fake-news-duke-reporters-lab/>

13) William Fleck, Nicholas Snell, Terry Traylor and Jeremy Straub, "Development of a 'Fake News' Machine Learning Classifier and a Dataset for its Testing", Proceeding volume 11013, Conference Paper, Disruptive Technologies in Information Science II, 110130A,

## 2. 가짜뉴스를 걸러내는 AI 사례

구글이 후원하는 ‘팩트마타(Factmata)’는 팩트체크가 필요한 문장을 선별한 후, 과거 텍스트, 통계수치 등과 비교해 사실인지 판단합니다. 수많은 웹사이트 링크에서 자동으로 문서를 수집, NLP 기술로 분석하고 색인을 달아 저장하는 구글 ‘지식 금고’도 자동 팩트체크에 활용될 수 있습니다. 텍사스대 연구팀의 ‘클레임버스터(Claimbuster)’는 어떤 문장이 사실인지를 0~1 사이 숫자로 나타내는데, 1에 가까울수록 팩트체크가 필요하다는 뜻입니다.<sup>14)</sup>

MIT 연구팀 알고리즘은 정보 출처의 정확성을 검증하고, 영국 스타트업 ‘파블라(Fabula)’는 가짜뉴스 특유의 공유·확산 양식을 토대로 선별합니다.<sup>15)</sup> 유럽연합(EU)도 AI를 이용해 가짜뉴스에 맞서는 ‘판당고(Fandango)’ 프로젝트를 진행하고 있습니다. 판당고는 AI가 조작한 이미지나 영상을 원래대로 되돌리고, 알고리즘을 이용해 기자들이 조작된 부분을 찾아낼 수 있게 도와줍니다.<sup>16)</sup>

그밖에 아래 <그림 2>에 소개된 프랑스 르몽드의 ‘데코되르(Les Decodeurs)’, 가짜뉴스를 골라내도록 돕는 웹사이트 ‘혹시(Hoaxy)’, 스놉스(Snopes)’ 등이 가짜뉴스와의 전쟁에 앞장서고 있습니다.<sup>17)</sup>

---

2019, <https://doi.org/10.1117/12.2520131>

14) 오세욱, 「팩트체크와 테크놀로지」, 오세욱 외 저, 『팩트체크 저널리즘』, 2019, 155-178쪽.

15) Vivar, 2019. *op. cit.*, pp. 206-209.

16) Tom Cassauwers, "Can artificial intelligence help end fake news?", *Horizon: EU Research & Innovation Magazine*, 15 Apr 2019, <https://horizon-magazine.eu/article/can-artificial-intelligence-help-end-fake-news.html>

17) <https://marutitech.com/artificial-intelligence-fake-news/>



<그림 2> 인도 마루티테크랩스가 소개한 가짜뉴스 선별 도구. 홈페이지 캡처

### 3. AI 팩트체크의 한계

하지만 AI가 만들어낸 가짜뉴스를 구분하는 데 AI 기술이 큰 효과를 발휘하지 못한다는 반론도 있습니다. 인간이 쓴 문장과 로봇이 쓴 문장을 구분해 내는 정확도는 뛰어나지만 어느 한 문장이 진실인지, 악의적인지를 검증하는 능력은 부족하다는 분석입니다.<sup>18)</sup> AI는 인간이 쉽게 알아챌 수 있는 함의를 파악하지 못하기도 하고, ‘과포괄성(Overinclusiveness)’ 때문에 진실인 글까지 가짜뉴스로 판별하기도 합니다.<sup>19)</sup> 오세욱 한국언론진흥재단 선임연구위원에 따르면 AI 팩트체크 혹은 자동 팩트체크는 지식기반 방식(기존 정보와 비교 검증), 맥락적 방식(정보 확산유형 기반 판단), 형식기반 방식(사실임을 확인할 수 있는 특정 형식을 기준으로 판단), 머신러닝 방식(인간이 알고 있는 바를 수학적 모델로 구현해 기계에게 학습시킴) 등 4가지 유형 모두 확실한 한계가 있습니다.<sup>20)</sup> 오 선임연구위원의 연구결과를 간단히 정리하면 아래 표와 같습니다.

18) Tal Schuster, Roei Schuster, Darsh J. Shah and Regina Barzilay, "The Limitations of Stylometry for Detecting Machine-generated Fake News", *Computational Linguistics*, Vol. 46, No. 2, Jun 2020, pp. 499-510.

19) Katarina Kertysova, "Artificial Intelligence and Disinformation-How AI Changes the Way Disinformation is Produced, Disseminated, and Can Be Counterd", *Security and Human rights* 29(2018), pp. 60-61.

20) 오세욱, 앞의 책, 160-179쪽.

지식기반 방식	기록되지 않은 데이터 확인 불가
맥락적 방식	사실이라도 빠르게 확산하면 가짜로 판단
형식기반 방식	형식을 제대로 갖춘 가짜뉴스 식별 불가
머신러닝 방식	학습데이터 부족, 모델링의 어려움

<표 1> 자동 팩트체크 기술의 방식별 한계

## IV. AI 저널리즘의 위험성

### 1. AI는 절대적 중립성·객관성을 보장할 수 없다

자연어 생성(NLG) 기술을 적용해 자동으로 작성된 기사와 인간이 쓴 기사의 신뢰도 및 객관성·중립성을 비교하기 위한 실험 결과, AI 작성 기사가 더 신뢰도가 높고 객관적인 것으로 평가됐다는 연구결과가 있습니다.<sup>21)</sup>

그러나 머신러닝 기술을 기반으로 한 AI는 결국 알고리즘 설계자의 의도대로 ‘훈련될’ 수 있다는 한계가 있습니다.<sup>22)</sup> AI가 학습하는 훈련용 데이터가 조작된다면 AI가 생성한 결과물도 왜곡될 수밖에 없습니다.<sup>23)</sup> 이른바 알고리즘 편향이 모델링 과정에서의 편향, 측정상의 편향, 입력데이터의 편향 등 3가지 방향에서 발생할 수 있다는 것입니다. 황용석 건국대 미디어커뮤니케이션학과 교수는 2019년 7월 18일 한국인공지능법학회 주최 제5차 AI 정책포럼에서 “AI에게는 편견이 없을 것이라는 생각이야말로 편견”이라고 역설하기도 했습니다.<sup>24)</sup> 저널리즘 분야는 아니지만, 국내에서도 AI 챗봇 ‘이루다’가 2021년 장애인·동성애에 대한 혐오를 학습해 논란이 된 바 있습니다.

### 2. AI가 생성한 가짜뉴스의 위험성

2017년 LA타임스에서는 AI가 ‘오보’를 낸 적이 있습니다. 웨이크봇이 캘

21) Yanfang Wu, "Is Automated Journalistic Writing Less Biased? An Experimental Test of Auto-Written and Human-Written News Stories", *Journalism Practice*, Vol. 14, Issue 8, 2020, pp. 1008-1028.

22) Stefan Rass, "Judging the quality of (fake) news on the internet", *Mind & Society*, Aug 2020, <https://doi.org/10.1007/s11299-020-00249-x>

23) Samuel Wooley, "We're fighting fake news AI bots by using more AI. That's a mistake", <https://www.technologyreview.com/2020/01/08/130983/were-fighting-fake-news-ai-bots-by-using-more-ai-thats-a-mistake/>

24) 노성열, 『AI 시대, 내 일의 내일』, 동아시야, 2020, 265-266쪽.

리포니아주 이슬라 비스타에서 리히터규모 6.8의 강진이 발생했다는 기사를 내보내고 트위터로도 알렸는데, 실제로는 1925년에 일어났던 지진이었습니다. 미 지질조사국(USGS)이 과거 지진 데이터베이스를 업데이트하는 과정에서 잘못된 지진경고를 내보냈는데, 이를 감지한 웨이크봇이 1분 만에 알고리즘에 따라 기사를 만들어 띄워버린 것입니다.<sup>25)</sup> 확인 취재를 거치지 않는 AI가 허위정보를 퍼뜨릴 수 있는 위험성을 잘 보여주는 사례입니다.

AI를 개발하는 비영리단체 ‘오픈AI(OpenAI)’는 2019년 ‘GPT-2’라는 NLP 모델을 제작했는데, 이 모델은 10번 만에 ‘영어를 할 줄 아는 유니콘’에 대한 논리 정연한 이야기를 만들어냈습니다. 오픈AI는 당시 “기만적이거나 가짜인 정보를 생성하는 데 악용될 우려가 있다”며 GPT-2의 축소 버전만 공개했습니다.<sup>26)</sup>

유엔의 빅데이터 혁신 계획 ‘글로벌 펄스(Global Pulse)’ 연구자들은 AI 문서생성 기술의 위험성을 보여주기 위해 위키피디아 문서와 유엔총회 연설 데이터베이스를 이용해 AI 언어모델을 훈련시킨 바 있습니다. 실험 결과 AI는 13시간 만에 기후변화, 이민자, 핵 군축 등 다양한 논제에 대해 실제 연설문 같은 글을 써냈습니다.<sup>27)</sup>

워싱턴대 앨런 AI 연구소의 그로버는 가짜뉴스를 식별하는 능력이 뛰어나지만, 역설적으로 ‘진짜 같은 가짜뉴스’를 만들어내는 능력도 탁월합니다. 그로버는 폴 크루그먼의 문체를 모방한 가짜 뉴욕타임스 칼럼, 워싱턴포스트 스타일의 가짜 정치뉴스, LA타임스 영화평론가 케네스 튜란 스타일의 가짜 영화평 등을 써냈습니다. 수돗물 불소 첨가 위험성을 주제로 한 그로버의 가짜 기사는 사람이 쓴 것보다 높은 평점을 받았습니다.<sup>28)</sup>

### 3. 갈수록 정교해지는 딥페이크의 위협

2020년 크리스마스를 맞아 영국 BBC 방송이 엘리자베스 2세 여왕의 연설을 방송한 뒤, 채널4 방송은 엘리자베스 여왕이 “(동영상 SNS) 틱톡에 올리겠다”며 춤을 추는 동영상을 내보냈습니다. 이 동영상은 AI로 정교하게 합성한 가짜, 즉 딥페이크(Deepfake) 영상이었습니다. 채널 4는 ‘허위정보와 가

25) "Will AI Save Journalism - or Kill it?", *The Wharton School online business journal*, 9 Apr 2020, <https://knowledge.wharton.upenn.edu/article/ai-in-journalism/>

26) Tessa Knight, "Who's afraid of robots? Fake news in the age of AI", 5 Mar 2019, <https://www.dailymaverick.co.za/article/2019-03-05-whos-afraid-of-robots-fake-news-in-the-age-of-ai/>

27) Kertysova, 2018, *op. cit.*, p. 68.

28) Zellers et al., 2019, *op. cit.*, pp. 5-21.

짜뉴스 확산을 가능케 하는 기술의 진화에 대한 경고'의 의미로 이 영상을 방송했다고 밝혔습니다.



<그림 3> 채널 4가 만든 춤추는 엘리자베스 여왕 딥페이크 영상. 유튜브 캡처

2017년 워싱턴대 연구진은 버락 오바마 전 미국 대통령의 딥페이크 영상을 만들어보였습니다. 2019년엔 낸시 펠로시 하원의장이 술에 취해 횡설수설하는 것처럼 만든 가짜영상이 나돌기도 했습니다. 2030년이 되면 딥페이크 기술은 진짜와 구분할 수 없을 정도로 발전하고, 가짜영상을 만들어내는 더욱 쉬워질 것이라고 합니다. 실제로는 한 적이 없는 말이나 행동을 정교하게 조작·합성해 만들어낸 딥페이크 영상이 퍼진다면 리더나 기관의 신뢰도는 크게 훼손될 것입니다. 가짜영상·녹음의 제작·유포가 쉬워지면 정치나 국제관계를 넘어 전방위로 협박범죄가 증가할 것이란 전망이 나옵니다.<sup>29)</sup>

사이버보안 그룹 '기록된 미래(Recorded Future)'의 CEO 크리스토퍼 알버그는 파이낸셜타임스 인터뷰에서 불과 3~4년 사이에 딥페이크 제작용 AI 프로그램을 구매하기가 너무 쉬워졌다고 지적했습니다.<sup>30)</sup>

#### 4. AI에게 명예훼손 책임을 물을 수 있을까? 새로운 법적·윤리적 과제

2016년 8월 페이스북은 이용자들 사이에 가장 인기 있는 뉴스를 보여주는 '트렌딩 토픽스'를 담당하던 팀을 해고하고 AI 알고리즘이 자동으로 인기 주제를 인지해 알리도록 했습니다. 그런데 불과 이틀 뒤 AI는 당시 폭스뉴스 앵커였던 메건 켈리가 힐러리 클린턴을 지지해 해고됐다는 가짜뉴스를 8시간 이상이나 트렌딩 토픽스로 소개했고, 아무 근거 없는 허위정보가 20만

29) Kertysova, 2018, *op. cit.*, pp. 67-68.

30) Hannah Murphy, "The new AI tools spreading fake news in politics and business", 10 May 2020, <https://www.ft.com/content/55a39e92-8357-11ea-b872-8db45d5f6714>

개 이상의 ‘좋아요’를 받았습니다.<sup>31)</sup>

AI발 가짜뉴스 확산 위험성을 잘 보여준 사례인 동시에, 이 사건은 AI의 ‘책임성’에 대한 논쟁을 불러일으켰다는 점에서도 유명합니다. 로봇이 명예훼손을 저지를 수 있는가, 알고리즘이 저지른 잘못에 대해 페이스북에 법적 책임을 물을 수 있는가 등이 논란거리가 됐습니다. 페이스북 트렌딩 토픽스가 아니라 대형 언론사에서 AI가 가짜뉴스를 내보내 확산시켰다면 문제는 더욱 심각했을 것입니다.<sup>32)</sup>

AI 명예훼손 논쟁은 AI가 언론자유 보호의 대상이 될 수 있는냐는 질문으로까지 확대되고 있습니다. AI가 만든 가짜뉴스가 보도됐을 경우 언론사의 편집장이나 팩트체크 담당자가 책임을 져야 하는가, 프로그래머가 책임져야 하는가도 결론이 나지 않은 논쟁거리입니다. 더 나아가 AI에게도 ‘취재원 보호’ 윤리가 적용되는지, 아니면 AI 저널리즘의 생명은 데이터의 정확성인 만큼 독자에게 모든 출처를 투명하게 밝히는 것이 더 윤리적인지와 같은 보다 근본적인 문제제기도 이뤄지고 있습니다.<sup>33)</sup>

## V. 결론

지금까지 살펴봤듯이 저널리즘 영역에서 AI의 활용은 결국 ‘양날의 검’이라고 할 수 있습니다. 인간과는 비교할 수 없는 속도와 방대한 데이터를 처리할 수 있는 AI의 능력은 가짜뉴스를 걸러내고 보도의 정확성과 중립성을 높이는 데 유용하게 활용될 수 있습니다.

그러나 AI는 대면 또는 전화 등을 통해 직접 확인취재를 거칠 수 없다는 점에서 되레 허위정보를 퍼뜨릴 위험이 있습니다. 기술이 더욱 발전해 AI가 이메일이나 메신저 등을 통해 인터뷰까지 하는 시대가 올 수도 있겠지만, 이런 미래가 오더라도 인간 기자가 상대방을 직접 보면서 상호작용하고 발언의 맥락을 읽어내는 것과는 차이가 있을 것입니다.

또 학습하는 데이터나 알고리즘을 짜는 프로그래머의 의도 등에 따라 AI

---

31) Matt Carlson, "Automatic Judgment? Algorithmic judgment, news knowledge, and journalistic professionalism", *new media & society*, Vol. 20, Issue 5, 2017, pp. 1755-1772. 1-18, 2017.

32) Seth C. Lewis, Amy Kristin Sanders and Casey Carmody, "Libel by Algorithm? Automated Journalism and the Threat of Legal Liability", *Journalism & Communication Quarterly*, Vol. 96(1), 2019, pp. 60-69.

33) Matteo Monti, "Automated journalism and freedom of information: ethical and juridical problems related to AI in the press field", *Studies in Comparative and National Law*, Vol. 1, n.1/2018, pp. 146-151.

가 작성한 기사 내용이 왜곡될 가능성도 항상 존재합니다.

저는 1년간의 조지아대 그래디 저널리즘 및 매스커뮤니케이션 칼리지(Grady College of Journalism and Mass Communication) BPC(Business and Public Communications) 펠로우 프로그램 세미나 및 연구를 통해 이런 생각을 뒷받침해주는 전문가들의 의견을 충분히 접할 수 있었습니다.

조지아대 그래디 칼리지의 콕스 저널리즘 혁신·관리·리더십연구소장(Director of James M. Cox Jr. Institute for Journalism Innovation, Management, and Leadership)을 맡고 있는 키스 헨던(Keith Herdon) 교수는 “AI가 단신 뉴스 생산을 넘어 이제는 딥페이크 기술의 경우 비디오 에디팅에도 쓰이고 있다”며 “이를 통해 인간 기자의 역할이 줄어드는 것뿐 아니라, 무엇보다 딥페이크를 통해 허위정보나 거짓정보가 생산돼 유통될 수 있다는 점이 무서운(scary) 부분”이라고 지적했습니다. 헨던 교수가 이끄는 콕스 연구소는 매년 조사를 실시해 뉴스 미디어 산업의 미래에 대한 함의를 담은 데이터를 제공하는 것으로 유명합니다.

또 커뮤니케이션법 전문가인 조너선 피터스(Jonathan Peters) 교수는 “시장 이론(Marketplace theory)에 따르면 좋은 견해와 나쁜 견해가 시장에서 경쟁, 더 좋은 견해가 살아남는다. 이에 더 보편적으로 수용되는 견해가 더 좋은 견해라는 주장이 언론의 자유를 다루는 학계의 주류였다”며 “하지만 소셜미디어 시대에는 훈련받지 않은 크리에이터(Creator), AI에 의해 여론이 쉽게 조작될 수 있다”고 강조했습니다. 미디어 분석학자로서 특히 소셜미디어 분석 권위자인 이타이 히멜보임(Itai Himelboim) 교수 역시 “정보의 출처가 어디인지, 즉 신뢰할 만한 기관에서 나온 것인지 검증하는 것이 더욱 중요해졌다”고 설명했습니다.

이에 대한 보완책과 관련, 위기관리 커뮤니케이션 전문가이자 조지아대 보건 및 위험 커뮤니케이션 센터의 부센터장(Associate Director of Center for Health & Risk Communication)을 맡고 있는 얀 진(Yan Jin) 교수는 “예를 들어 보건 분야에서 가짜뉴스 확산은 큰 재난을 초래할 수 있다”며 “AI가 만들어낸 잘못된 정보 혹은 거짓정보가 소셜미디어를 통해 퍼질 때, 이를 바로잡을 수 있는 것은 인간지능(Human Intelligence)”이라고 역설했습니다. 그는 “AI는 어떤 정보가 트렌드가 되고 있는지 감지하는 역할을 하고, 정보의 출처를 확인하고 진짜와 가짜를 가려내는 역할은 사람이 해야 한다”며 “인공지능과 인간지능이 결합할 때 비로소 ‘초지능(Super Intelligence)’이 될 수 있다”고 강조했습니다.

결국 AI 저널리즘 시대에도 오보나 가짜뉴스를 막기 위해 콘텐츠를 검증

하는 역할은 인간이 담당해야 하며<sup>34)</sup>, 끊임없는 사실 확인과 진실의 추구라는 기자 및 언론의 사명은 그 중요성이 더욱 커질 것입니다.

---

34) Patrick White, "How Artificial Intelligence Can Save Journalism?", 23 May 2020, <https://theconversation.com/how-artificial-intelligence-can-save-journalism-137544>

## <참고문헌>

### 1. 국내문헌

김동환·이준환, 「로봇저널리즘: 알고리즘을 통한 스포츠 기사 자동생성에 관한 연구」, 『한국언론학보』 59(5), 64-95쪽.

노성열, 『AI 시대, 내 일의 내일』, 동아시아, 2020.

박진아, 「인공지능 저널리즘의 법적 쟁점」, 『과학기술과 법』, 충북대 법학연구소, 2019, vol. 10, no. 2, 통권 19호, 119-154쪽.

오세욱, 「팩트체크와 테크놀로지」, 오세욱 외 저, 『팩트체크 저널리즘』, 2019, 149-194쪽.

이운영·안종목, 「로봇저널리즘의 사례와 이슈에 관한 탐색적 고찰」, 『커뮤니케이션학 연구』 제26권 2호(2018년 2호), 165-185쪽.

이창형, 「AI와 미디어」, 이창형 외 저, 『AI 시대의 미디어』, 북스타, 2020, 77-142쪽.

### 2. 외국문헌

Anderson, Samuel; Sulistryani, Hapsari, "Detecting and combating fake news on web 2.0 technology in the 2019 political season Indonesia", *Journal of Social Studies*, Vol. 15, NO. 2(2019), pp. 103-116.

Amali D, Geraldine; Poddar, Karishnu; Umadevi, K S, "Comparison of Various Machine Learning Models for Accurate Detection of Fake News", 2019 Innovation in Power and Advanced Computing Technology, Vellore, India, 2019, DOI: 10.1109/i-PACT 44901.2019.8960044.

Carlson, Matt, "Automatic Judgment? Algorithmic judgment, news knowledge, and journalistic professionalism", *new media & society*, Vol. 20, Issue 5, 2017, pp. 1755-1772.

Cassauwers, Tom, "Can artificial intelligence help end fake news?", *Horizon: EU Research & Innovation Magazine*, Apr 2019.

Fleck, William; Snell, Nicholas; Traylor, Terry; Straub, Jeremy, "Development of a 'Fake News' Machine Learning Classifier and a Dataset for its Testing", Proceeding volume 11013, Conference: Disruptive Technologies in Information Science II, 2019, DOI: 10.1117/12.2520131.

Himma-Kadakas, Marju, "Alternative facts and fake news entering

journalistic content production cycle", *Cosmopolitan Civil Societies: An Interdisciplinary Journal*, 9(2), 2017, pp. 25–41.

Ireton, Cherilyn; Posetti, Julie, Journalism, *'Fake News' & Disinformation, Handbook for Journalism Education and Training*, UNSESCO, 2018.

Kertysova, Katarina, "Artificial Intelligence and Disinformation—How AI Changes the Way Disinformation is Produced, Disseminated, and Can Be Counterd", *Security and Human rights* 29(2018), pp. 55–81.

Lewis, Seth C.; Sanders, Amy Kristin; Carmody, Casey, "Libel by Algorithm? Automated Journalism and the Threat of Legal Liability", *Journalism & Communication Quarterly*, Vol. 96(1), 2019, pp. 60–81.

Martens, Bertin; Aguiar, Luis; Gomez–Herrera, Estrella; Muller–Langer, Frank, "The digital transformation of news media and the rise of disinformation and fake news", *JRC Digital Economy Working Paper*, No. 2018–02, European Commission, Joint Research Centre, Seville.

Monti, Matteo, "Automated journalism and freedom of information: ethical and juridicial problems related to AI in the press field", *Studies in Comparative and National Law*, Vol. 1, n.1/2018, pp. 139–155.

Pathak, Archita; Srihari, Rohini, "BREAKING! Presenting Fake News Corpus For Automated Fact Checking", *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop*, Florence, Italy, 28 Jul–2 Aug, 2019, pp. 357–362.

Rass, Stefan, "Judging the quality of (fake) news on the internet", *Mind & Society*, Aug 2020, <https://doi.org/10.1007/s11299-020-00249-x>

Saad, Saad; Issa, Talat A., "Integration or Replacement: Journalism in the Era of Artificial Intelligence and Robot Journalism", *International Journal of Media, Journalism and Mass Communication*, Vol. 6, Issue 3, 2020, pp. 1–13.

Schuster, Tal; Schuster, Roei; Shah, Darsh J.; Barzilay, Regina, "The Limitations of Stylometry for Detecting Machine–generated Fake News", *Computational Linguistics*, Vol. 46, No. 2, Jun 2020, pp. 499–510.

Tunez–Lopez M; Toural–Bran, C; Valdiviezo–Abad, C, "Automation, bots and algorithms in newsmaking. Impact and quality of artificial journalism", *Revista Latina de Comunicacion Social*, 74, 2019, pp.

1411–1433.

Vivar, Flores, "Artificial intelligence and journalism: diluting the impact of disinformation and fake news through bots", *Doxa Communication*, 29, 2019, pp. 197–212.

Wang, Patrick; Angarita, Rafael; Renna, Ilaria, "Is this the Era of Misinformation yet? Combining Social Bots and Fake News to Deceive the Masses", *The 2019 Web Conference Companion*, Mar 2018, Lyon, France, 10.1145/3184558.3191610.

Wu, Yanfang, "Is Automated Journalistic Writing Less Biased? An Experimental Test of Auto-Written and Human-Written News Stories", *Journalism Practice*, Vol. 14, Issue 8, 2020, pp. 1008–1028.

<https://channels.theinnovationenterprise.com/articles/how-can-artificial-intelligence-combat-fake-news>

<https://knowledge.wharton.upenn.edu/article/ai-in-journalism/>

<https://marutitech.com/artificial-intelligence-fake-news/>

<https://theconversation.com/how-artificial-intelligence-can-save-journalism-137544>

<https://whatsnewinpublishing.com/turning-fake-news-against-itself-ai-tool-can-detect-disinformation-with-92-accuracy/>

<https://whatsnewinpublishing.com/using-ai-to-fight-fake-news-duke-reporters-lab/>

<https://www.bbc.com/news/technology-49446729>

<https://www.dailymaverick.co.za/article/2019-03-05-whos-afraid-of-ro-bots-fake-news-in-the-age-of-ai/>

<https://www.ft.com/content/55a39e92-8357-11ea-b872-8db45d5f6714>

<https://www.globenewswire.com/news-release/2020/08/28/2085316/0/en/Syracuse-University-Researchers-Using-Artificial-Intelligence-to-Combat-Fake-News.html>

<https://www.poynter.org/fact-checking/2019/these-projects-are-using-ai-to-fight-misinformation>

<https://www.technologyreview.com/2020/01/08/130983/were-fighting-fake-news-ai-bots-by-using-more-ai-thats-a-mistake/>

<https://www.vice.com/en/article/vvjzvd/how-robo-journalists-could-flood-the-internet-with-fake-news>