

03

동영상 플랫폼의 선거보도 콘텐츠 규제 방안

김익현 지디넷코리아 미디어연구소장



1. 들어가며

미국 뉴햄프셔(New Hampshire) 주(州) 프라이머리(primary election, 예비선거)를 앞둔 지난 1월 말 유권자들에게 조 바이든(Joe Biden) 대통령의 전화가 걸려왔다. ‘경선 불참’을 권유하는 로보콜(자동응답 전화)이었다. 물론 ‘진짜’ 바이든의 전화는 아니었다. 인공지능(AI) 기술을 이용해 바이든 대통령의 음성을 정교하게 조작한 ‘딥페이크(deepfake)’였다. 이 사건은 선거를 앞둔 민감한 시기에 최첨단 AI 기술을 활용해 여론 조작을 시도했다는 점에서 많은 사람들에게 큰 충격을 안겨줬다.

‘딥러닝(deep learning)’과 가짜(fake)의 합성어인 ‘딥페이크’는 AI를 활용한 합성 기술을 의미한다. 뉴햄프셔 프라이머리에서 등장한 것은 음성 파일이었지만 딥페이크가 더 많이 활용되는 것은 영상과 이미지다. 최근 미국에선 조 바이든 대통령이 트랜스젠더 혐오발언을 하는 딥페이크 영상이 등장했으며, 국내에서도 윤석열 대통령 고백 영상을 가장한 딥페이크 때문에 한바탕 논란이 벌어졌다. 전 세계 76개국에 대통령 선거나 총선에 예정돼 있는 ‘슈퍼 선거의 해’를 맞아 각국들이 딥페이크를 비롯한 허위조작 영상을 효과적으로 규제하기 위해 다양한 노력을 하고 있다. 허위 영상이 여론을 왜곡하고 민주주의를 오염시키는 것을 두고 볼 수만은 없기 때문이다. 한국에서도 22대 국회의원 선거를 앞두고 중앙선거관리위원회가 ‘딥페이크 영상’을 이용하는 선거운동을 전면 금지시켰다. 딥페이크 같은 허위조작 영상은 주로 대형 동영상 플랫폼을 통해 유포되기 때문에 관련 법규 역시 플랫폼 사업자 규제에 초점이 맞춰져 있다. 더 정확하게는, 플랫폼 사업자들이 자신들의 플랫폼에서 유통되는 허위영상에 대해 어디까지 책임을 질 것이냐는 것이 핵심 논쟁거리다.

이 글에서는 선거 국면을 맞아 각국의 관심사로 떠오르고 있는 허위조작 영상 규제 및 심의현황과 함께 대형 온라인 동영상 플랫폼들의 자율 규제 움직임에 대해 살펴본다. 각국 정부나 입법기관의 규제 동향과 함께 동영상 유포의 핵심 채널인 플랫폼 사업자들의 허위조작 영상 심의 대책에 대해서도 함께 살펴본다.

2. 플랫폼에 강력한 책임 요구하는 유럽

전 세계 인터넷 산업을 주도하고 있는 미국이 플랫폼에 대해 느슨한 규제 정책을 유지해 온 반면, 유럽연합(EU)은 강력한 책임을 요구하고 있다. 미국이 플랫폼에 강력한 책임을 묻지 않는 것은 초기부터 자국 기업들이 인터넷 시장을 주도했던 것과 관련이 있다. 1990년대 후반 인터넷 혁명을 주도하면서 새로운 산업 육성을 위해 규제보다는 진흥 쪽에 좀 더 무게를 뒀던 전통

이 지금까지 이어지고 있는 것이다. 반면 전통적으로 나치를 비롯한 극우 혐오 발언에 대해 민감한 역사적 배경을 갖고 있는 EU는 플랫폼들에게 좀 더 적극적인 책임을 요구하고 있다. EU 인터넷 시장을 미국 기업들이 장악하고 있는 점도 강력한 규제와 관련이 있다.

1) EU <디지털서비스법>의 강력한 플랫폼 규제

EU의 플랫폼 규제 정책을 대표하는 것은 <디지털서비스법>(Digital Services Act, DSA)이다. 2023년 8월부터 본격 적용되고 있는 DSA는 개인의 권리가 보호되는, 안전하고 예측 가능하며 신뢰할 수 있는 온라인 환경을 확보한다는 입법 목적을 갖고 있다. 오프라인 공간에서 불법인 행위는 온라인에서도 통용되지 않는다는 것이 DSA의 기본 출발점이다. 특히 DSA는 온라인 플랫폼을 통해 허위정보나 혐오발언들이 무차별 확대 재생산되는 것을 막는 데 초점을 맞추고 있다. 이런 입법 목적을 달성하기 위해 플랫폼 사업자들은 인종차별을 조장하는 글이나 허위 조작 정보에 대해선 좀 더 강력한 관리 및 법적 책임을 지도록 했다. 위반할 경우엔 전 세계 매출의 6%까지 벌금을 부과할 수 있도록 하는 등 처벌 규정도 강력하다.

DSA의 핵심은 크게 세 가지로 요약할 수 있다.¹⁾

① 불법 콘텐츠 삭제

불법 콘텐츠를 적극적으로 찾아야 할 의무는 없지만, 사법 기관이나 행정 당국으로부터 조치를 명령 받으면 즉각 시행해야 한다. 이를 위해 이용자들이 불법 콘텐츠를 수월하게 신고할 수 있도록 해야 한다. 하지만 어떤 것이 '불법 정보'에 해당되는지는 구체적으로 규정하지 않았다. 이 부분은 개별 국가의 법체계에 따라 적용하도록 했다.

② 투명성 강화

분쟁 관련 정보나 콘텐츠 추천 알고리즘을 투명하게 공개해야 한다. 특히 대규모 플랫폼사업자는 추천 시스템의 주요 변수와 이용자가 조작할 수 있

¹⁾ 김현수·전성호 (2020). 유럽연합 디지털서비스법안(Digital Service Acts)의 주요 내용 및 시사점. <KISDI Premium Report>, 충북: 정보통신정책연구원.

는 방법을 약관에 명기하도록 했다. 또 매년 1회 콘텐츠 정책, 절차, 조치, 수단 등을 담은 보고서를 의무적으로 발간해야 한다.

③ 비례성 원칙 적용

규모가 큰 플랫폼 사업자일수록 더 큰 책임을 부여했다. 특히 월간 활성 이용자 4,500만 명 이상인 대규모 플랫폼 사업자는 리스크 평가·완화·외부 감사, 규정준수 책임자 지정, 연구자에 대한 데이터 접근 권한 제공 의무를 지도록 했다.

DSA는 시민들의 안전을 위협할 수 있는 범죄와 관련되었다고 의심할 만한 정보를 정부에게 알리거나(제21조), 불법정보가 아니더라도 정보 왜곡, 혐오, 차별이 될 수 있는 정보에 대해 대형 플랫폼들이 스스로의 판단 하에 조치(제27조)를 취하도록 하는 등 강력한 책임성을 부여하고 있다.²⁾ 따라서 DSA는 그 동안의 여러 법률들과는 다른 차원의 플랫폼 책임성을 요구한다는 점에서는 새로운 전범이 될 가능성이 많다. 하지만 개인 사업자인 플랫폼에게 과도한 권한을 부여한다는 비판도 만만치 않다.

DSA가 추천 알고리즘을 투명하게 공개하도록 한 것은 대형 온라인 플랫폼의 ‘필터버블(filter bubble)’ 차단을 겨냥한 것이다. 필터버블이란 정보 제공자가 알고리즘을 기반으로 이용자 맞춤형 정보를 제공함으로써 이용자들이 필터링 된 정보만을 접하게 되는 현상을 의미한다. ‘보고 싶은 것만 보여 주는’ 필터버블은 민주적 토론의 결립돌인 ‘확증편향’으로 이어질 가능성이 많다. 실제로 유튜브(YouTube)나 페이스북(Facebook) 같은 대형 플랫폼들은 비슷한 생각을 가진 사람들의 소통을 극단적으로 강화함으로써 오히려 진영 대립을 더 강화시키고 있다는 비판을 받고 있다. DSA는 유튜브나 인스타그램 이용자들이 추천 알고리즘 대신 ‘최신 게시물’이 최상단에 노출되도록 하는 표출 방식을 선택할 수 있도록 했다.

EU는 좀 더 효율적인 알고리즘 규제를 위해 2023년 4월 알고리즘투명성 센터(European Centre for Algorithmic Transparency, ECTA)를 발족했다. ECTA는 AI를 활용한 추천 시스템이나 각종 검색결과에 인증, 젠더

2) 박찬경 (2022). 유럽연합 디지털서비스법(Digital Service Acts): 규제나 자율이나를 넘어서. <언론중재>, 2022년 가을호(통권 164호), pp.62-73.

편견을 드러내지 않는지 감시하고 평가하는 역할을 하게 된다. 특히 ECTA는 알고리즘 감시를 위해 컴퓨터, 데이터 과학자 뿐 아니라 사회과학, 인지 과학 분야 전문 연구자들까지 채용해 ‘다학제적인 접근(multi-disciplinary approach)’을 하고 있다.

2) 독일의 <네트워크집행법>

EU가 DSA를 제정하기 오래전부터 독일은 <네트워크집행법>(Netzwerkdurchsetzungsgesetz, NetzDG)을 통해 플랫폼 사업자들을 규제해 왔다. 2017년 9월 발효된 <네트워크집행법>은 독일 내 이용자 200만 명이상이 플랫폼이 규제 대상이다. 규제 대상 사업자들은 자신들의 플랫폼에 올라온 위법 게시물을 24시간 이내 삭제해야 한다. 또 6개월에 한번씩 <네트워크집행법> 시행 결과를 담은 ‘투명성보고서’를 독일어로 발행해야 한다.

<네트워크집행법>은 발효된 지 3년이 채 안 된 2020년 6월에는 개정안이 의결되면서 이용자 보호 규정이 좀 더 강화됐다. 개정안에서는 특히 이용자 권리 강화 부문이 많이 보강됐다. 이에 따라 신고 게시물이 삭제 대상인지 여부에 대한 이의신청 절차를 도입하도록 했다. 또 이용자들이 위법 게시물에 대한 소송 절차를 좀 더 수월하게 진행할 수 있도록 독일 내에 소송대리인을 두도록 했으며, 이용자와 플랫폼 사업자 간의 분쟁을 조율할 수 있는 중립적인 민간중재기구를 의무적으로 두도록 했다.³⁾

독일 내 극우 발언이나 혐오 표현 등에 대한 규제에 초점을 맞춘 <네트워크집행법>은 DSA의 기반이 됐다는 평가를 받고 있다. 특히 개정안을 통해 이용자들의 신고 및 이의신청 절차를 명확하게 한 부분이나, 소송대리인제 도입을 통해 글로벌 사업자를 상대로 한 문제 제기를 좀 더 수월하게 할 수 있는 점은 이용자 보호 측면에서 진일보한 것으로 평가되고 있다.

하지만 이 법은 국가 기구가 아니라 사적 기업에게 콘텐츠에 대한 검열을 상당 부분 위임함으로써 사적 규제를 행사하도록 했다는 비판이 적지 않게 제기됐다. 특히 플랫폼 사업자들이 벌금을 피하려는 동기에서 법적으로 허

3) 이유진 (2020). 독일 네트워크 법집행법 개정안 발의: SNS 사업자 책임은 더 강화하고, 이용자 보호는 더 철저히. <언론중재>. 2020년 여름호(통권 155호), pp.106-117.



용될 가능성이 있는 콘텐츠까지 삭제함으로써 부적절한 사적 검열로 이어질 수 있다는 주장도 제기되고 있다. 이런 비판에 대해선 플랫폼 사업자들이 제 3자가 올린 콘텐츠에 대한 관리 책임을 좀 더 강하게 가지도록 해, 공익적 관점에서 허용 가능한 자율 규제를 할 수도 있을 것이란 반론도 만만치 않다.⁴⁾

3) 프랑스의 <정보조작대처법>

프랑스의 <정보조작대처법>(LOI n° 2018-1202 du 22 décembre 2018 relative à la lutte contre la manipulation de l'information)은 2017년 대통령 선거 당시 과도한 허위정보 난립이 문제가 되자 보완 차원에서 마련한 법이다. 마크롱 정부는 출범과 동시에 선거기간 동안 허위정보 확산을 막기 위해 관련 내용을 즉각 삭제하도록 하는 것을 골자로 하는 법안을 발의했다. 이후 몇 차례 수정 작업 끝에 2018년 11월 20일 <정보조작대처법>이 공식 발효됐다.

이 법은 대통령 선거 기간 인터넷으로 허위 정보가 대량 유통되면서 정상

4) 황용석 (2023. 12. 21). 온라인 혐오표현 규제는 왜 어려운가? <KISO Journal>, 서울: 한국인터넷자율정책기구. URL: <https://journal.kiso.or.kr/?p=12464>

적인 민의 수렴 절차가 심각하게 훼손됐다는 문제의식에서 출발했다. 선거 기간 동안 인터넷을 통해 허위정보 공세를 펼칠 경우 형법 같은 기존 법률만으로는 신속하게 대처하기 힘든 점을 보완하기 위해 ‘가짜뉴스 조항’ 등을 마련했다. 따라서 정보조작대처법은 신속한 대처가 필요한 선거 기간 동안 허위정보 유포를 막는 데 초점을 맞추고 있는 법이란 점에서 EU의 DSA나 독일의 <네트워크집행법>과는 입법 목표가 조금 다른 편이다.

국내에서 ‘가짜뉴스 대처법’으로 소개된 <정보조작대처법>은 선거 운동 기간과 선거 운동 이외 기간에 허위정보 공세에 대처하기 위한 규정을 담고 있다. 일단 이 법은 국가적 차원의 선거일 이전 3개월 동안 선거에 불법적인 영향을 미칠 목적으로 허위정보를 의도적으로 대량 배포하는 행위를 중단시키기 위해 민사 가처분 제도를 도입했다.

<정보조작대처법>은 선거일 3개월 전부터 선거일까지 정부기관, 후보자, 정당, 시민단체 등이 인터넷을 통해 유포되는 가짜뉴스에 대한 단독 판사의 개입을 요구할 수 있다고 규정하고 있다. 단독 판사는 4가지 조건을 충족할 경우에 유포 금지 가처분 명령을 내릴 수 있다.

- ① 해당 정보 내지 뉴스가 명백히 허위일 것
- ② 인위적일 것
- ③ 대량으로 유포될 것
- ④ 선거의 진정성에 영향을 미칠 수 있는 부정확하거나 허위 사실의 주장 내지 비판일 것

유포 금지 조건 못지않게 중요한 것은 신속 처리 규정이다. <정보조작대처법>은, 가짜뉴스에 대해 제소가 접수될 경우 단독 판사는 48시간 이내에 판결하도록 했다. 또 단독 판사의 결정에 대해 항소할 경우 항소법원도 48시간 이내에 선고해야만 한다. 선거라는 특수 상황을 감안해 신속한 처리에 초점을 맞추고 있는 셈이다. <정보조작대처법>은 선거기간 허위정보 유포 방지를 위해 플랫폼 사업자들에게도 몇 가지 의무를 부여하고 있다.

5) 이정년 (2021), 인터넷 가짜뉴스의 규율을 위한 형사입법모델 연구: 프랑스의 입법례를 분석하여, <서강법률논총>, 제10권 제1호, pp.129-152.
6) 이정년, 앞의 논문.

- ① 신고 체계 마련: 누구든지 가짜뉴스를 신고할 수 있도록 접근 가능하고 쉽게 인식할 수 있도록 할 것
- ② 알고리즘 투명성 확보: 이용자들이 정보(콘텐츠)에 대해 비판적인 판단을 내릴 수 있도록 할 것
- ③ 계정 탐지 시스템 구축: 플랫폼 상에서 방대한 양의 정보(콘텐츠)를 유포하는 계정 탐지할 수 있도록 할 것

〈정보조작대처법〉은 선거라는 특수한 상황을 감안해 판사들이 유포금지 가처분 명령 같은 비상조치를 발령할 수 있도록 한 것이 특징이다. 하지만 대다수 정보는 48시간 이내에 위법 여부를 판단하기 쉽지 않을 가능성이 많다. 그러다보니 〈정보조작대처법〉은 처음 발의될 당시부터 ‘표현의 자유’를 침해할 우려가 있다는 비판이 제기됐다. 실제로 프랑스는 2019년 유럽 선거와 2022년 대통령 선거 등 두 차례 선거에 〈정보조작대처법〉을 적용했지만 이렇다 할 효력을 발휘하지 못했다.

3. 미국의 느슨한 플랫폼 규제

미국은 거대 플랫폼 사업자들에 대해선 최소한의 관리자 역할만을 요구해 왔다. 이런 정책 기조는 인터넷 초창기부터 지금까지 20년 가까이 계속 유지돼 왔다. 이런 정책적 기조의 근간이 된 것이 ‘선한 사마리아인’ 면책 조항으로 불리는 〈통신품위법〉(Communications Decency Act, CDA) 230조다. 미국 뿐 아니라 세계 인터넷 발전의 원동력이 됐다는 평가를 받고 있는 이 조항은 최근 들어선 ‘가짜뉴스 현상’의 주범 중 하나라는 지적과 함께 강한 비판에 직면해 있다.

1) 플랫폼 정책의 기초가 된 〈통신품위법〉 230조

인터넷 초창기 미국의 최대 고민은 ‘음란물’이었다. 인터넷상에서 무차별 유포되는 음란물 공세로부터 미성년자들을 보호하는 것이 중요한 정책적 과제였다. 그런 문제의식을 담은 것이 1996년 제정된 〈통신품위법〉이었다.

목적부터 예사롭지 않았던 이 법은 과잉 규제 때문에 일부 조항이 위헌 판결을 받는 등 논란이 끊이지 않았다. 하지만 나중에 추가된 230조 때문에 초기 인터넷 발전을 이끈 중요한 법으로 꼽힌다. 〈통신품위법〉 230조는 “선



한 의도를 갖고 콘텐츠 중재 작업을 하는 플랫폼 사업자에 대해선 발행자에 준하는 의무를 부과하지 않아야 한다”는 내용을 담고 있다.

미국이 초기부터 플랫폼 사업자들에게 최소한의 관리자 의무만 부여했던 것은 크게 두 가지 이유 때문이다. 무엇보다 당시 인터넷은 막 떠오르기 시작한 신흥 산업이었다. 뉴미디어로 막 각광받기 시작한 인터넷 비즈니스를 제대로 키워내기 위해선 규제보다는 진흥이 훨씬 효과적이라는 판단에 따른 조치였다. 당시 미국이 전 세계 인터넷 산업을 주도적으로 이끌고 있었다는 점도 이런 정책 기조를 만들어낸 중요한 요인으로 꼽힌다. 이런 정책적 기조 덕분에 야후를 비롯한 초기 인터넷 사업자들이 빠른 속도로 세계 인터넷 시장을 주도할 수 있었다.

하지만 최근 몇 년 사이에 미국에서도 플랫폼 규제에 대한 입장이 조금씩 달라지고 있다. 페이스북, 유튜브 같은 대형 플랫폼에 허위정보가 넘쳐나면서 사회문제로 떠오르고 있기 때문이다. 이런 상황에서 2021년 1월 발생한 의사당 난입 사건은 미국의 플랫폼 규제 정책에 중요한 전기가 됐다. 페이스북, X(당시 트위터), 유튜브를 비롯한 소셜 플랫폼을 통한 허위조작 정보 유통이 ‘대선 불복 난동’의 원인 중 하나로 꼽힌 때문이다. 이런 상황에서 페이스북이 안전보다는 자사 이익을 우선하는 알고리즘을 유지하고 있다는 내부 고발까지 터져 나오면서 ‘플랫폼 책임성’ 공방이 뜨겁게 달아오르고 있다. 규제보다는 진흥에 초점을 맞췄던 정책 기조가 흔들리면서, 플랫폼 사업자들

에게도 강력한 책임을 부여해야 한다는 주장이 힘을 얻기 시작했다.

플랫폼 규제 강화에는 공감하지만 그 이유는 조금 다르다. 공화당 쪽은 소셜 플랫폼들이 보수 의견을 지나치게 제재하고 있다고 주장하고 있다. 따라서 <통신품위법> 230조가 플랫폼 사업자들에게 부여한 면책 특권을 없애야 한다는 입장이다. 반면 민주당은 <통신품위법> 230조가 보장한 면책 특권 때문에 소셜 플랫폼들이 허위정보에 좀 더 책임있게 대처하지 않고 있다고 비판하고 있다.

2) <통신품위법> 230조 개정 움직임

최근 몇 년 사이에 미국 의회는 <통신품위법> 230조 개정을 위해 많은 노력을 해왔다. 상하원에 발의된 법안도 꽤 많이 있다. 대표적인 것이 미국 상원에서 추진하고 있는 <세이프테크법>(SAFE TECH Act)이다. 마크 워너(Mark Warner), 매지 히로노(Mazie Hirono), 에이미 클로버샤(Amy Klobuchar) 등 민주당 상원 의원 세 명이 추진하고 있는 <세이프테크법>은 “양방향 컴퓨터 서비스 제공자나 이용자는 다른 정보 콘텐츠 제공자가 제공한 어떠한 정보의 발행자나 발화자로 간주해서는 안 된다”는 <통신품위법> 230조 규정 중 ‘정보(Information)’란 단어를 ‘진술(Speech)’로 바꿨다.⁷⁾ 이렇게 함으로써 잘못된 정보에 대해서는 일정 부분 책임을 부여하겠다는 것이다.

최근엔 <아동성착취물방지법>(Stop CSAM Act)과 <온라인아동안전법>(Kids Online Safety Act·KOSA)이 미국 전역의 큰 관심을 끌고 있다. 미국 연방 상원은 지난 1월 말 이 법 제정을 위한 청문회에 메타(Meta), 스냅챗(Snapchat), 틱톡(TikTok), X 등의 최고경영자(CEO)들을 불러 플랫폼 책임성 문제를 따지기도 했다.

이 법들이 통과될 경우 페이스북, 유튜브 등 플랫폼 사업자들에게 아동 성착취물 등 콘텐츠 관리 책임을 물을 수 있게 된다. <통신품위법> 230조의 보호를 받던 플랫폼 사업자들의 방어막이 균열되는 셈이다. KOSA는 상원에서 60명 이상 의원의 지지를 받고 있어 통과 가능성이 상당히 높은 편이다.⁸⁾

7) 김익현 (2021). 힘 실리는 '빅테크 견제법' 면책 특권 줄이고 규제는 대폭 강화. <신문과방송> 12월호. 서울: 한국언론진흥재단

8) Feiner, L. (2024, 2, 16). Kids Online Safety Act gains enough supporters to pass the Senate, <The Verge>. URL: <https://www.theverge.com/2024/2/15/24073878/kids-online-safety-act-new-senate-support>

4. 거대 플랫폼 사업자들의 두 가지 상반된 움직임

거대 플랫폼에서 유통되는 선거 조작 영상을 가려내는 작업은 생각처럼 간단하지 않다. 특히 신속한 판단이 요구되는 선거 관련 영상은 난이도가 더 높다. 플랫폼 사업자에 대해 단순 중개자 이상의 책임을 요구하지 않고 있는 미국 뿐 아니라 강경 입장을 보이고 있는 EU도 마찬가지다. 프랑스 <정보조작대처법> 사례에서 볼 수 있는 것처럼 자칫하면 민주주의의 핵심 가치인 표현의 자유와 충돌할 우려가 있기 때문이다.

따라서 올해 같은 ‘슈퍼 선거의 해’엔 허위조작 영상 퇴치를 각국 정부의 규제 정책에만 의존할 수는 없는 상황이다. 거대 플랫폼 사업자의 자율규제가 중요한 역할을 할 수밖에 없다. 이 문제와 관련해 거대 플랫폼 사업자들은 두 가지 상반된 행보를 보이고 있어 눈길을 끈다.

1) 20개 빅테크의 자율규제 선언

마이크로소프트(Microsoft), 메타, 구글 등 20개 빅테크 기업들이 유권자를 속이는 AI 거짓 정보와의 전쟁을 선포했다. 이들은 지난 2월 독일에서 열린 뮌헨안보회의(Munich Security Conference)에서 AI 기술을 활용해 선거에 영향을 끼치는 행위에 공동 대응하겠다고 선언했다.⁹⁾

협약에 참여한 기업들은 초거대 AI 모델의 위험성을 산정하고, 자신들의 플랫폼에서 유포되는 각종 허위정보를 탐지하는 등의 대책을 마련하기로 했다. 또 자신들의 플랫폼에서 허위정보를 탐지해 ‘꼬리표’를 붙이는 등의 방법을 동원하기로 했다. 허위정보 퇴치 모범 사례는 협약에 참여한 다른 기업들과 적극 공유하기로 했다. 한 마디로 AI를 활용한 선거 허위정보로 신뢰를 무너뜨리는 행위에 함께 대응하겠다고 선언한 셈이다.

이들이 특히 초점을 맞춘 것은 ‘딥페이크’를 비롯한 허위 조작 영상이다. 20개 빅테크 기업들이 ‘딥페이크와의 전쟁’에 관심이 쏠리는 것은 2024년이 ‘슈퍼 선거의 해’이기 때문이다. 올해는 전 세계 76개국에서 대통령 선거나 총선이 치러질 예정이다. 사실상 전 세계 인구의 절반 이상이 선거를 경험할 것으로 예상된다. 그러다보니 그 어느 때보다 ‘딥페이크’ 영상이 여론을 왜곡하고 민주주의를 오염시킬지 모른다는 우려가 커지고 있다. ‘뮌헨안보회의’에 참

9) O'Brien, M & Swenson, A. (2024, 1. 17). Tech companies sign accord to combat AI-generated election trickery. (AP). URL: <https://apnews.com/article/ai-generated-election-deepfakes-munich-accord-meta-google-microsoft-tiktok-x-c40924ffc68c94fac74fa994c520fc06>

석한 20개 빅테크 기업들의 협약 역시 이런 문제의식에서 나온 것이다.

이런 노력에도 불구하고 빅테크들의 자율규제 조치가 AI 기술을 이용한 허위 콘텐츠, 특히 강력한 파괴력을 자랑하는 허위 동영상을 제어하는 역할을 제대로 해낼 수 있을지 의문이다. 합의 내용이 모호할 뿐 아니라 구속력이 없기 때문이다. 그러다보니 빅테크들의 공동 행보가 의미는 있지만, 선거를 겨냥한 교묘한 조작 영상을 제대로 걸러낼 수 있을 지에 대해서는 진한 의문 부호가 붙어 있다.

2) 빅테크들의 콘텐츠 관리 인력 축소

플랫폼 사업자들은 겉으로는 대형 선거를 앞두고 허위정보 추방을 위해 만반의 준비를 하고 있다. 선거에 악영향을 끼치는 허위 조작 영상 추방을 위해 협력하는 것을 골자로 하는 협약을 발표하는 등 대책 마련에 적극 나서고 있는 것처럼 보인다. 하지만 실제 움직임은 그런 외부적인 선언과는 조금 다르다. 지난해부터 콘텐츠 심의 관련 인력을 감축하는 등 대외적으로 내세운 명분과는 상반된 행보를 보이고 있다. 게다가 콘텐츠 관리 정책도 일관성이 없는 것으로 나타났다.¹⁰⁾

그 중 대표적인 것이 세계 최대 동영상 유통 플랫폼인 유튜브다. 유튜브는 미국 대통령 선거를 1년 5개월가량 남겨 놓은 2023년 6월 ‘선거 허위정보’ 처리 관련 전략을 공개했다. 2020년 대통령 선거 관련 허위정보를 더 이상 삭제 처리하지 않겠다는 것이 달라진 콘텐츠 전략의 핵심이었다. 유튜브는 2020년 미국 대선 당시 엄격하게 적용했던 제재 정책이 허위정보를 추방하는 데 이렇다 할 기여를 하지 못하면서, 오히려 자유로운 토론을 저해하는 부작용만 가져왔다고 평가했다.

이런 판단에 따라 ‘공개적인 정치적 논쟁(open political debate)’을 진작하기 위해 콘텐츠 모더레이션(Content-Moderation) 정책을 수정하게 됐다고 설명했다. 콘텐츠 삭제 조치가 일부 허위정보를 제어하는 효과를 보긴 했지만, 오히려 폭력이나 기타 실제 세계에 피해를 입힐 위험을 의미 있을 정도로 줄여주지는 못하면서 정치적 발언을 억제하는 의도치 않은 결과로 이어졌다

10) YouTube Team (2023, 6. 2). An update on our approach to US election misinformation. (YouTube official blog). URL: <https://blog.youtube/inside-youtube/us-election-misinformation-update-2023/>

는 것이다.

콘텐츠 관리 정책을 느슨하게 바꾼 것은 유튜브 뿐만이 아니다. 2022년 일론 머스크(Elon Musk)로 주인이 바뀐 트위터 역시 콘텐츠 관리 인력을 대폭 감축했다. 팩트체크 전문기관인 폴리팩트(PolitiFact)에 따르면 트위터는 2022년 11월 신뢰 및 안전팀 인원 15%를 감축했으며, 이듬해 1월에도 추가로 허위정보 모니터링 업무를 담당하는 콘텐츠 모데레이션(content moderation) 직원 수십 명을 해고했다.¹¹⁾

메타는 플랫폼 내에서 유포되는 허위정보를 가려내기 위해 폴리팩트를 비롯한 11개 팩트체크 전문 기관들과 협력하고 있다. 허위정보로 판명될 경우 노출을 제한하는 등의 조치를 적용하고 있다. 다만 이 기준은 현재 활동 중인 정치인들에게는 적용하지 않는다. 하지만 메타 역시 허위정보를 가려내는 콘텐츠 관리 업무를 대폭 축소했다. 2013년 초 200명가량의 콘텐츠 모데레이터를 해고했을 뿐 아니라, 6개월가량 준비했던 팩트체크 관련 도구 개발 프로젝트도 중단시켜버렸다.¹²⁾

유튜브, 메타, 트위터 같은 거대 플랫폼 사업자들은 최근 선거가 있을 때마다 허위정보 유통의 주범으로 엄청난 비난을 받았다. 이런 상황에서 '슈퍼 선거의 해'를 앞두고 콘텐츠 관리 역량을 대폭 축소하는 것은 '정치 사회적 고려' 보다는 '경제적 계산'이 강하게 작용한 때문이란 게 대체적인 분석이다. 허위정보는 사실을 전달하는 콘텐츠보다 유튜브 같은 동영상 플랫폼에서 훨씬 빠른 속도로 전파되는 경향이 있기 때문이다. 허위 콘텐츠를 적극적으로 관리하지 않는 것이 수익 측면에선 훨씬 이득이 된다는 것이다. 그러다보니 선거 관련 허위정보를 더 이상 삭제하지 않겠다는 유튜브의 결정에 대해선 '민주주의를 위협하는 행보'라는 비판까지 제기되고 있다.¹³⁾

11) Wenn, S.(2023. 7. 7). How Meta, Tiktok, Twitter and YouTube plan to address 2024 election misinformation. (Poynter). URL: <https://www.poynter.org/fact-checking/2023/how-meta-tiktok-twitter-and-youtube-plan-to-address-2024-election-misinformation/>

12) Field, H. (2023. 5. 26). Tech layoffs ravage the teams that fight online misinformation and hate speech. (CBNC). URL: https://www.cnbc.com/2023/05/26/tech-companies-are-laying-off-their-ethics-and-safety-teams-.html?utm_content=Main&utm_medium=Social&utm_source=Twitter

13) Wenn, S.(2023. 7. 7). op. cit.

5. 나가며

페이스북, 유튜브 등 소셜 플랫폼은 한때 참여 민주주의의 토대로 많은 각광을 받았다. 그런 역할은 지금도 변함이 없다. 하지만 전 세계적으로 정파적 대립이 극심해지면서 소셜 플랫폼의 부작용이 커지고 있다. 특히 올해처럼 전 세계적으로 대형 선거가 예정돼 있을 때면 허위정보와 혐오발언 공세가 더 거셀 것으로 우려된다. 4월 총선이 예정돼 있는 우리나라도 벌써부터 유튜브 같은 플랫폼에 다양한 정치 콘텐츠들이 올라오고 있다.

이 글에서는 선거 관련 허위 영상을 가려내려는 각국의 정책과 함께 유튜브를 비롯한 대형 플랫폼 사업자들의 자율규제 동향을 간단하게 살펴봤다. 그동안 플랫폼에 대해 느슨한 규제를 적용했던 미국은 최근 들어 <통신품위법> 230조 개정을 추진하면서 그 동안의 기조가 바뀔 조짐을 보이고 있다. 반면 지난해부터 <디지털서비스법>을 적용하고 있는 EU에선 유튜브, 페이스북 같은 대형 플랫폼들에게 강력한 관리 책임을 부여하고 있다. 하지만 강력한 책임을 요구하는 쪽이나, 느슨한 규제를 적용하고 있는 쪽이나, 짧은 선거기간 동안 치고 빠지는 각종 허위 콘텐츠를 효과적으로 제어할 수 있을지는 여전히 미지수다.

그렇기 때문에 플랫폼사업자들의 적극적인 행보가 중요한 상황이다. 전 세계 20개 빅테크 기업들이 플랫폼 내 허위정보 자율규제에 나서겠다고 선언한 부분이 관심을 끄는 것은 이런 배경 때문이다. 실제로 이들은 ‘딥페이크’ 같은 정교한 조작 영상을 제대로 가려내기 위해 긴밀하게 협력하겠다고 약속했다. 거대 플랫폼 사업자들이 한 목소리를 낸 점은 의미가 있지만, 거창한 약속이 제대로 지켜질지는 의문이다. 자율 규제와 표현의 자유의 경계선을 명쾌하게 규정하는 것이 생각처럼 쉽지는 않기 때문이다. 게다가 협약에 동참한 대다수 플랫폼 사업자들이 내부적으로는 콘텐츠 모더레이션 관련 인력과 투자를 대폭 감축하고 있다는 점도 이들의 약속에 깊은 신뢰를 보내기 힘들게 만든다. 🗣️

