

불법콘텐츠 확산 방지를 위한 매개자 규제의 정당성과 유럽 법제 검토

정 애 령*

숙명여자대학교 법학연구소 연구교수, 법학박사

국문초록

본 연구는 인공지능 기술의 확산과 지능 정보화 사회의 진전으로 온라인 플랫폼이 사회적 소통의 장에서 핵심적인 역할을 수행하면서 발생하는 새로운 법익 침해 문제와 그에 따른 규제의 필요성을 탐구한다. 인공지능 기술의 활용이 증대됨에 따라 딥페이크, 가짜뉴스, 혐오 표현 등의 콘텐츠가 대중에게 빠르게 확산하고, 이는 여론을 왜곡하고 민주주의적 의사결정을 방해할 우려를 낳고 있다. 특히 정보의 소비자에서 능동적 주체로 전환된 사용자들이 플랫폼을 통해 만들어내는 콘텐츠가 언론의 자유와 인격권을 동시에 위협하며, 이는 플랫폼이 미디어 매개자로서 법적 책임을 가져야 한다는 주장의 근거가 된다. 의사소통의 장으로서 폭발적인 파급력과 영향력을 가진 동시에 구조적인 우월성을 지닌 온라인 플랫폼의 규제는 불법콘텐츠의 확산을 막기 위한 가장 효과적인 방법이기도 하다. 이에 본 연구는 이러한 새로운 위협에 대응하기 위한 플랫폼 규제의 정당성을 고찰하며, 유럽을 중심으로 한 규제 법제의 발전 양상을 검토하였다.

특히, 유럽의 대표적인 매개자 규제 사례로 독일의 네크워크집행법(NetzDG)과 프랑스의 정보조작대처법 및 혐오·불법콘텐츠 대응을 위한 아비아법, EU 디지털서비스법을 소개하며, 유럽의 법제에서 플랫폼 운영자에게 요구하는 법적 책임의 범위와 주요 규제 내용을 검토하였다. 또한

* arjung@sookmyung.ac.kr

EU 디지털서비스법의 제정으로 독일의 네트워크집행법이 일부 폐지되고, 독일의 디지털서비스법을 제정함으로써 유럽연합 전체 영역에서 매개자 규제에 대한 기준을 마련하였다. 그러나 이러한 규제들은 표현의 자유를 사전에 제한할 수 있는 문제를 내포하고 있으며, 사적 검열의 문제와 과잉 차단 위험성에 대한 우려가 존재한다. 그럼에도 불구하고 사회 구성원들의 자유로운 의견 수렴의 결과가 아닌 조작된 허위 정보의 경우, 이에 대한 보호 가치보다 공익이 현저히 높다고 판단된다.

또한 의사 표현 행위로 타인의 개인적 법익을 침해할 때 이에 대응할 법제가 마련되어 있다 하더라도, 정보통신기술과 접목된 폭발적 전파성과 파급효과를 고려할 때 매개자 규제를 통하여 무분별한 확산을 방지하고 법익 침해를 줄일 필요성이 인정된다. 다만 매개자 규제의 효율성 및 필요성을 인지하면서도 불법콘텐츠에 대한 조치 시 이를 판단하는 주체를 넓히거나, 의견을 청취하는 등의 절차적 측면을 보완하고, 표현의 자유 침해 가능성을 최소화할 다각적 노력이 필요할 것이다.

주제어: 매개자 규제, 온라인 플랫폼, 불법콘텐츠, EU 디지털서비스법

목 차

- I. 서론
- II. 지능정보사회 표현의 자유 변화와 매개자 규제의 정당성
 - 1. 첨단정보통신기술과 접목된 개인적·사회적 법익 침해
 - 2. 매개자 규제를 통한 해결 방안 모색
- III. 유럽 매개자 규제의 구체적 법제 검토
 - 1. 독일 네트워크집행법(NetzDG)과 디지털서비스법(DDG)
 - 2. 프랑스 정보조작대처법 및 혐오·불법콘텐츠 대응법
 - 3. EU 디지털서비스법(DSA)
- IV. 결론

I. 서론

정보통신기술의 발전은 특히 언론매체와 표현의 자유 실현에 많은 변화를 가져왔다. 표현의 자유와 조화를 이루어야 할 인격권 보호를 위한 법적 규제로 명예훼손 또는 모욕죄를 검토하는 데 중요한 요소는 바로 ‘전파성’이다. 정보의 소비자에 머물렀던 일반 국민은 현재 누구나 손쉽게 강렬한 미디어 콘텐츠를 생성하고, 이를 온라인 플랫폼을 통하여 더욱 빠르게 확산시킨다. 최근에는 딥페이크 기술을 이용하여, 실제와 구분이 힘든 사실적인 영상을 구현하여 중·고등학교 친구, 선생님 사진을 이용한 음란물을 생성하는 등 개인의 권리 침해가 심각하며,¹⁾ 가짜뉴스 생산을 통해 대중을 오도하고 사회적 혼란을 초래하는 사례들을 수없이 경험하였다. 그럼에도 불구하고 표현의 자유에 대한 규제는 늘 사전검열과 언론탄압으로 변질 위험성을 내포하고 있다.²⁾

1) MBC 뉴스(2024.10.21.), “올해 딥페이크 피해 학생 27명 더 늘어 841명...교직원 은 36명”

2) 방통심의위는 언론중재위원회의 중재 기능과 별도로 인터넷 언론사의 콘텐츠에 대해 정보통신망법에 따라 가짜뉴스 관련 불법, 유해 정보를 심의한다고 밝혔다. 이에 따라 정보통신에 관한 심의 규정을 위반한 인터넷 언론사의 온라인 콘텐츠

다양한 사상을 존중하고 이를 자유롭게 표현할 수 있도록 보장하는 것은 민주주의의 핵심 요소이다.³⁾ 방송과 신문 등 기존 언론매체를 통해 정보를 얻고 의사를 형성하던 과거와 달리 인터넷과 소셜미디어, 검색엔진, 그 외 플랫폼의 등장은 민주적 의사 형성의 새로운 공간을 마련하였다.⁴⁾ 특히 온라인 플랫폼은 이용자 간 상호작용을 촉진하는 것을 목적으로 개인의 자유 및 권리를 실현할 뿐만 아니라 사회적 가치를 실현하는 데 기여하였다. 온라인 플랫폼은 정치적 이데올로기가 아닌 개인적 표현이나 스토리텔링을 바탕으로 정치적 저항의 새로운 논리를 가능하게 하고, 모든 이용자가 온라인 콘텐츠의 생산자이자 소비자로 평등한 입장에서 참여할 수 있는 환경을 만들어 주었다.⁵⁾ 자아실현 및 인격 발현의 기본적 요소이자, 사회 구성원으로서 의사결정에 참여하는 수단인 표현의 자유의 중요성은 아무리 강조해도 지나치지 않는다. 그러나 허위 사실들이 무분별하게 양산되고 배포되어 무고한 피해를 낳는 결과 역시 두고만 볼 수는 없다. 디지털 환경의 변화가 우리 사회의 근간을 이루는 다원적 가치의 자유로운 소통과 민주주의를 지탱하는 신뢰에 위협을 가하고 있기 때문이다.⁶⁾

이에 따라 본 글에서는 정보통신기술의 발전이 가져온 표현의 자유의 변화와 침해되는 법익의 유형과 그 정도에 대해 언급하고, 발화자 이외의 매개자 규제의 정당성을 검토한다. (II) 불법콘텐츠의 확산을 막을 수

에 대해서 심의결과와 함께, 행정기관과 플랫폼사업자에게 불법, 유통사실을 통보하고, 경찰 수사의뢰 등 조치를 요청할 계획이다. 이에 대하여 내용에 대한 심의는 엄격한 요건하에 명확한 요건과 절차를 갖추어야 할 것인데, 이는 언론탄압으로 번질 위험성을 제기하며 비판이 일고 있다. (미디어스 2023.10.06., 방심위팀장 11인, 가짜뉴스 심의에 언론검열우려)

3) 박용상(2002), 〈표현의 자유〉, 헌암사, 59면.

4) Antje von Ungern-Sternberg(2019), 민주적 의사형성과 인공지능, in : *Demokratie und künstliche Intelligenz*, 한국법제연구원(역), 6면.

5) 송해엽(2021), 인공지능 기반 콘텐츠 필터링의 사회적 합의와 규제된 자율규제, 〈방송통신연구〉, 겨울호, 10면.

6) 권형돈(2020), 인터넷에서 표현의 자유와 그 규제법리-혐오표현·거짓정보에 대한 독일의 대응을 중심으로, 〈사법〉 제1권 제51호, 7면.

있도록 매개자를 규제한 독일 및 프랑스 법제에 표현의 자유 침해 가능성에 대한 비판과 헌법재판소의 결정 등을 함께 고찰하고, 올해 시행된 EU의 디지털서비스법의 내용과 한계를 살펴본다. (Ⅲ) 이로써 새로운 위협에 대응 방안을 모색하며 첨단과학기술을 이용하여 확장된 표현의 자유가 다른 법익과 조화를 이룰 수 있는 법제 마련에 도움이 되고자 한다.

Ⅱ. 지능정보사회 표현의 자유 변화와 매개자 규제의 정당성

1. 첨단정보통신기술과 접목된 개인적·사회적 법익 침해

가. 인공지능기술을 이용한 인격권 침해의 중대성

인터넷은 가장 표현 촉진적인 매체로서 기능하고, 온라인 플랫폼은 이용자들이 생산한 정보를 실시간으로 공유하고 전파함으로써 표현의 자유 확장에 기여했다. 그러나 이는 사회적으로 해로운 행동을 쉽게 하는 기회를 제공한 측면도 있다. 플랫폼은 개인적 상호작용에서 나타나는 무례함과 공격성, 공적 영역의 허위 정보나 조작, 정치 영역에서의 양극화와 착취 및 불법이 쉽게 전파되고 극대화되는 장을 마련한 것도 사실이다.⁷⁾ 여기에 인공지능 기술을 이용한 영상물 제작과 배포는 인격권 침해에 심각성을 더하고 있다. 딥러닝(deep-learning)과 거짓(fake)의 합성어인 딥페이크(Deepfake)는 특정 영상에 특정 인물의 얼굴과 음성 등을 합성하여 허위 정보를 만들어 내는 기술로 인공지능에서 주로 활용되는 딥러닝 기술을 기반으로 하고 있다.⁸⁾ 진실이 아닌 영상을 생성하는 AI 알고리즘은 사람의 표정과 몸동작, 음성뿐 아니라 기타 복합적 정보를

7) 송해엽(2021), 앞의 글, 10면.

8) 배상균(2021), 딥페이크(Deepfake)영상물에 관한 법적 대응조치 검토, <4차산업혁명 법과 정책> 제3호, 184-185면.

학습하여 모방하기에 사실과 구분하기 어렵다. 딥페이크 기술을 이용하여 합성한 선정적인 영상물이 유통되거나, 각종 정치인과 관련하여 교묘하게 만들어진 영상이 가짜뉴스로 활용되고, 이것이 선거에까지 영향을 미친 것이 사실이다.⁹⁾ 현재는 정치인이나 공인만이 타깃이 되는 것을 넘어 같은 학교, 같은 반 친구들이 딥페이크를 이용하여 불법 영상을 만들어 배포하는 상황에서, 일반 국민은 언제 어디에서 어떻게 조심해야 할지조차 모를 정도의 공포심을 갖게 되었다.

시청각 콘텐츠는 단순한 루머보다 더 높은 신뢰성을 가지며, 특히 거짓이더라도 주관적으로 옳다고 판단되는 뉴스에 특별한 신빙성을 부여하는 확증 편향이 존재하는 상황에서 딥페이크와 같은 소프트웨어 발전은 인격권 침해를 더욱 심화시킨다.¹⁰⁾ 더구나 현재는 이미지 조작을 위해서 전문적 장비나 설비가 필요하지 않고, 가정용 컴퓨터의 성능만으로 충분히 정교한 품질의 영상 합성이 가능하며, 무료나 소액으로 이러한 기술의 이용이 가능해졌다.¹¹⁾ 이러한 딥페이크 기술은 복제 및 유포의 용이성과 익명성을 갖기 때문에 너무도 쉽고 그럴싸한 허위조작정보가 생산되고, 플랫폼을 통해 무서운 속도로 확산할 수 있는 현재의 기술은 이전과는 비교할 수 없을 정도의 중대한 인격권 침해를 만들어 내고 있으므로, 이에 대한 사회적 대응 노력이 필요하다.¹²⁾

나. 다양한 의견 형성 저해

필터버블(Filter Bubble)은 2011년 정치운동가 Eli Pariser가 2009년부터 도입되었던 추천 알고리즘으로 인하여 벌어지는 개인화 현상을 설명하

9) 이민영(2020), 딥페이크와 여론형성 -알고리즘의 권력화와 탈진실의 규제담론-, 〈미국헌법연구〉 제31권 제1호, 204면.

10) Kumkar/Rapp(2022), Deepfakes, *ZfDR* 2022, 199, 202-203.

11) 박주일(2020), 딥페이크 기술의 발전과 저널리즘의 새로운 위협-이미지 공정성 확보를 위한 영상 팩트체크의 필요성, 〈언론중재〉, 가을호, 56면.

12) 배상균(2021), 앞의 글, 188-189면.

며 사용한 개념이다.¹³⁾ 필터버블은 사용자의 위치나, 검색 이력 등의 정보에 기반하여 추천 알고리즘이 선별적으로 사용자가 보고 싶어 할 정보를 추측하고, 사용자들이 효율적으로 자신이 필요한 정보와 부합하지 않는 정보로부터 분리되는 과정에서 의도치 않게 자신만의 문화적·이념적 거품에 갇히게 되는 현상을 의미한다.¹⁴⁾ 구글 검색 결과와 페이스북의 뉴스피드를 통하여 이용자의 관심에 기반한 정보를 전달할 때 의사형성에 있어 중요한 콘텐츠가 개인맞춤형 또는 편향적 성격을 갖게 되는 것이다. 추천 알고리즘에 정치적 상업적 논리가 개입되면 정보 이용자는 정보 편식과 선택적 노출로 인해 가치관 왜곡이 일어날 수 있다.¹⁵⁾ 정보전달의 개인화는 특정한 정치적 의견 또는 주제로 국한하여 소통하는 결과를 초래한다. 소위 필터버블 혹은 에코챔버라 불리는 이 같은 현상에는 국민이 더 이상 다른 의견과 씨름하지 않고, 공개 토론과 분위기를 왜곡하여 받아들이고, 같은 의견을 가진 사람들의 무리에서 그들 자신의 정치적 견해를 과격화할 위험이 있다.¹⁶⁾

알고리즘은 문제 해결을 위한 개별 단계들의 세분된 순서이다. 구글 검색과 페이스북 뉴스피드의 관건은 다량의 사용 가능한 인터넷 사이트와 기타 정보로부터 모든 중요한 정보가 유의미한 순서로 이용자에게 제시되는 것이다. 이러한 시스템은 무엇이 중요하고 유의미한지를 플랫

13) Pariser(2012), *Filter Bubble*, 9 f.

14) 전준영·황소윤·윤영미(2018), 개인화 알고리즘으로 필터버블이 형성되는 과정에 대한 검증, <멀티미디어학회 논문지>, 제21권 제3호, 369-381면.

15) 사이언스에 실린 연구(Bakshy et al., 2015)는 페이스북상에서 이용자들이 실제로 어떤 정치 성향의 포스트에 노출되며, 어떤 포스트를 클릭하는지 분석했다. 이 연구는 필터버블을 낳을 수 있는 요인으로 친구들의 정치 성향, 추천 알고리즘, 이용자 개인의 정치 성향 등 세 가지로 설정하고 어떤 요인이 가장 반대의견에 대한 노출을 방해하는지 측정한 결과, 친구들의 정치 성향과 이용자 개인의 정치 성향이 페이스북 추천 알고리즘보다 더 크게 작용한다는 결과가 나왔다. 알고리즘보다는 사람의 성향이 필터버블을 낳는다는 것이다. (조진형·김규정(2022), 소셜미디어에서 에코챔버에 의한 필터버블 현상 개선방안 연구, <한국콘텐츠학회 논문지> 제22권 제5호, 60면.)

16) Dörr&Natt(2014), Suchmaschinen und Meinungs Vielfalt, *ZUM* 2014, 829, 837.

품이 규정하는 것으로 볼 수 있다.¹⁷⁾

원래 검색엔진은 정보를 쉽게 찾을 수 있게 접근할 수 있도록 하고, 특정 콘텐츠를 멀리하거나 억제하는 기능은 전혀 없다. 검색엔진 운영자의 관심은 오로지 수조 개의 서로 다른 인터넷 주소 중에서 각 사용자의 정보 요구와 가장 관련성이 높은 웹사이트를 찾는 데 있다. 검색엔진의 성공 여부는 오로지 이 작업을 얼마나 잘 수행하느냐에 따라 측정된다.¹⁸⁾ 이런 의미로 검색엔진은 다양한 정보 소스에 대한 접근을 제공함으로써 다양성을 촉진하기 때문에 필터버블의 반대라고 언급하기도 한다.¹⁹⁾ 그러나 플랫폼 사업자는 상업적 이익 내지 이해관계에 따른 요인을 중요시할 수밖에 없고, 의사의 다양성 측면보다 조회수나 머무는 시간을 연장할 수 있도록 사람들이 보거나 듣고 싶어 하는 콘텐츠를 추천하는 방식을 취할 수 밖에 없을 것이다.²⁰⁾ 즉, 검색엔진의 알고리즘은 “(다수) 선호”에 맞춰 인지 가능한 의견 스펙트럼의 축소”를 일으켜 “덜 개방 된 사회의 형성을 촉진”하고 마침내 “(다수) 의견을 영속화”함으로써 반다원주의 경향을 촉진한다.²¹⁾

이러한 이유로 포털 또는 소셜네트워크 시스템은 특정 이슈에 대한 의견의 스펙트럼이 구조적으로 축소될 수 있고 다양한 의사의 자유로운 전달을 추구하는 자유민주주의의 걸림돌로 작용할 수 있다. 우리가 생체 기능이 어떻게 작동하는지 정확히 잘 모르듯이, 이 모든 것은 우리가 의식하지 못한 사이에 일어난다. 사회환경이 사람들 사이의 상호 기대로

17) Antje von Ungern-Sternberg(2019), 민주적 의사형성과 인공지능, in : *Demokratie und künstliche Intelligenz*, 한국법제연구원(역), 7면.

18) Nolte(2017), Hate-Speech, Fake-News, das Netzwerkdurchsetzungsgesetz und vielfaltsicherung durch Suchmaschinen, *ZUM* 2017, 552, 563.

19) Nolte(2017), Hate-Speech, Fake-News, das Netzwerkdurchsetzungsgesetz und vielfaltsicherung durch Suchmaschinen, *ZUM* 2017, 552, 563.

20) 이부하(2018), 디지털 시대에 의사표현의 왜곡과 표현의 자유의 한계, <법학논문집> 제42집 제3호, 51면.

21) www.kek-online.de/fileadmin/Download_KEK/Gutachten/Gutachten_Die_Bedeutung_des_Internets_im_Rahmen_der_Vielfaltssicherung.pdf.

형성되지 않고 오로지 데이터에 기반한 조작만이 목표가 된다면 자율성으로서의 자유와 사회 전체에 문제가 될 수 있다.²²⁾ 따라서 반드시 조작의 의도로 사용되지는 않으나, 여론 형성에 상당한 영향을 미치는 여타의 정보 매개자와 구글 및 페이스북의 검색 및 정렬 알고리즘과 관련하여 기존법 혹은 신규 법이 원치 않는 왜곡의 경향을 어떻게 막을 수 있고, 여론의 다양성을 보호할 수 있는지에 대한 문제를 고민할 필요가 있다.²³⁾

다. 정치적 의사결정과정과 여론의 왜곡

인터넷을 시작으로 촉진적인 표현의 장이 마련되었고, 인공지능 기술을 이용하여 누구든 쉽게 영상물을 제작하고, 자신의 생각과 의사를 표현할 수 있는 세계가 열렸다. 그러나 앞서 언급한 바와 같이 이는 다른 개인의 중대한 인격권 침해의 가능성을 높일 뿐만 아니라 정치활동을 수동적으로 만들고, 딥페이크 기술 또는 마이크로 마케팅을 통해 탈진실을 가속화하며, 여론 형성과 선거의 과정을 왜곡시키고 결국 민주주의를 위협할 위험성이 있다.²⁴⁾ 정치권은 유권자들이 더 많이 투표하도록 대중매체를 통해 진행하던 선거운동 방식에 머무르지 않고, 데이터마이닝과

22) 마크 코겔버그지음/ 배현석 옮김(2022), 〈인공지능은 왜 정치적인 수밖에 없는가〉, 생각이음, 60면.

23) Antje von Ungern-Sternberg(2019), 민주적 의사형성과 인공지능, in : *Demokratie und künstliche Intelligenz*, 한국법제연구원(역), 19면.

24) 인공지능 기술이 민주주의의 위협이 되리라는 우려는 표현의 자유의 변화와 관련되지 않더라도, 개인 상호 간의 의사소통 및 토론·토의 같은 직접적이고 대면적 정치참여를 필요로 하지 않기 때문에, 축적된 데이터 분석에 의한 온라인 모형이 권력 획득의 정당성을 부여하여 민주주의 핵심인 다양한 의견의 숙의 과정을 제외시켜 민주적 절차가 사라질 수 있다고 지적하거나(손현주(2019), 인공지능 거버넌스와 민주주의의 미래, 〈사회사상과 문화〉 제22권 제2호, 동양사회사상학회, 337- 338면.), 정치적 의사결정은 사회문제의 모호성과 민주적 절차를 모두 고려해서 다양한 가치와 형량이 필요한데, 기계학습에 의한 결정은 다소 객관적 필요성의 구현으로 나타날 것이라는 우려를 드러내기도 한다. (Koster(2022), Das Ende des Politischen? Demokratische Politik und Künstliche Intelligenz, *Zeitschrift für Politikwissenschaft(ZPol)*, 32.).

마이크로 타게팅²⁵⁾을 활용하여 하이퍼미디어캠페인에 주력하는 추세다.²⁶⁾ 쇼핑을 할 때이나 취미와 관련된 사안을 선택할 때 효율적인 알고리즘 기반 개인맞춤형 추천이 결국 정치적 의사결정과정에도 영향을 끼치고 민주주의에 위협이 될 수 있다는 우려는 현실로 나타났다.²⁷⁾ 가치중립적일거라 생각하는 챗 GPT나 메타의 라마와 같은 생성형 인공지능의 경우에 있어서도 특정 정치적 성향을 드러냄을 밝혔다.²⁸⁾ 아주 객관적인 정보를 제공하리라 기대하는 생성형 인공지능의 경우에도 특정 성향을 보이고 있다면, 어떤 애플리케이션을 사용하느냐에 따라 다른 선택을 가지게 될 가능성도 존재한다는 것을 고려해야 한다.

또한 소위 소셜봇(bot)은 사람의 통제 없이 운영되는 소셜 미디어 계정으로, 조직적인 방식으로 메시지를 동시에 대량으로 전송할 수 있다. 이를 사용하여 단순 업무를 반복하고, 지정된 뉴스 또는 링크를 유포하거나 단순한 키워드로 식별되는 여타의 정보에 좋아요나 리트윗으로 정보를 강화하거나 심지어 가짜뉴스의 유포를 자동화할 수 있게 한다. 나아가 지능형 소셜봇은 복잡한 정보를 수집 및 분석할 수 있고, 이에 반응하여 활동한다.²⁹⁾ 이러한 기술들을 이용하여 러시아 가짜뉴스를 담당하며 수백 명의 직원을 보유한 상트페테르부르크의 인터넷 연구기관(Internet Research Agency)과 같은 소위 트롤 공장도 적은 노력과 저렴한 비용으

25) 마이크로 타게팅은 개인의 인구통계학적 특성과 소비패턴 및 생활 습관으로부터 수집된 데이터를 취합하고·분석하여 좁은 범주의 유권자로 분리한 후, 이들을 대상으로 정교하게 필요한 메시지를 도출하는 기법을 의미한다. (이민영(2020), 앞의 글, 204면.)

26) 이민영(2020), 앞의 글, 204면

27) 영국이 EU의 탈퇴 여부를 결정하는 ‘브렉시트(Brexit) 투표’ 과정에서 특정 정치적 성향을 가진 사람들을 대상으로 정치적 캠페인이 집중되는 형태로 현실화되거나, 우리나라에서도 플랫폼 뉴스배열 알고리즘의 편향성 여부가 중요한 정치적 쟁점으로 부상하기도 하였다. (문의빈(2023), 알고리즘 기반 추천 플랫폼의 기본권 제한에 관한 비교법적 연구, <헌법학연구> 제29권 제2호, 405-406면.)

28) 김광석·박세익·박정호·오태민(2024), <트럼프 2.0>, 이든하우스.

29) Antje von Ungern-Sternberg(2019), 민주적 의사형성과 인공지능, in :*Demokratie und künstliche Intelligenz*, 한국법제연구원(역), 11면.

로 표적화된 허위 정보 캠페인을 시작할 수 있었다.³⁰⁾ 이러한 방식으로 확산한 뉴스는 많은 대중에게 도달하여 공개 토론에서 주제를 빠르게 설정하거나 강화할 수 있다.³¹⁾

마케팅 소셜봇 등의 이용을 막을 수는 없겠으나 인터넷상 소통에 있어 주변 의견이 갑자기 주류로 인식될 수도 있는 영향력을 가진 소셜봇을 이용한 의견 표현은 여론은 왜곡시키는 효과를 가져온다.³²⁾ 특히 생성형 AI모델은 사람이 만든 콘텐츠와 구별하기 어려운 가짜뉴스 등 합성 콘텐츠 역시 만들 수 있고, 이는 민주적 담론을 크게 왜곡시키는 데 일조한다.³³⁾

2. 매개자 규제를 통한 해결 방안 모색

가. 매개자 규제에 소극적인 미국

워싱턴 국회의사당 습격은 가짜뉴스와 혐오 표현의 확산이 민주주의의 근간을 흔들 수도 있다는 것을 분명하게 보여준 사건이다. 여기에서 SNS는 트럼프 대통령이 대중을 동원하는 수단으로 사용되었고, 의사 표현 매개체들의 책임에 대하여 생각해 보지 않을 수 없다. 미국은 정보의 유통 및 배포를 담당하는 플랫폼 사업자에게 가능한 책임을 부과하지 않고 최소한의 부담을 주고자 하는 전통적 입장을 고수한다. 다시 말해 통신품위법(Communications Decency Act)의 주요 원칙과 더불어 선한 사마리아인 법칙을 적용하여 명확한 불법성이 드러나지 않을 경우 매개자인 플랫폼 사업자 규제는 거의 하지 않는다. 사법기관도 아닌 민간사

30) „How the world was trolled“, *The Economist* v. 4.11.2017, S. 21 f.

31) Holznagel(2018), Phänomen “Fake News”- Was ist zu tun?, *MMR* 2018, 18, 20.

32) 정애령(2021), 코로나 시대 공정한 선거를 위한 자유로운 선거운동 방안모색, <유럽헌법연구> 제37호, 134면.

33) Berz&Engel&Hacker(2023), Generative KI, Datenschutz, Hassrede und Desinformation- zur Regulierung von KI-Meinungen, *ZUM* 2023, 586, 586.

업자가 뉴스의 진실성 및 허위 여부를 판단하는 것은 어렵기도 하지만, 표현의 자유 보장에 해가 될 수 있다는 이유에서다.³⁴⁾ 통신품위법 제230조는 제3자에 의해 만들어진 콘텐츠를 단순히 저장하거나 전파하는 역할을 하는 양방향 컴퓨터 서비스 사업자(interactive computer service providers)에 대해서 민사상 책임을 면제하는 면책조항이다. 이에 따라 양방향 컴퓨터 서비스 제공자나 이용자는 제3가 제공한 정보에 대해 발행자 또는 발화자로서 간주하지 않는다고 명시적으로 규정하고 있다.³⁵⁾ 그러나 러시아가 가짜뉴스를 통해서 선거 캠페인을 조작하려는 시도에도 불구하고 어떠한 액션도 취하지 않은 페이스북에 대한 비난 이후, 미국은 분위기가 바뀌고 있다.³⁶⁾

페이스북, 인스타그램, 틱톡, 링크트인과 같은 소셜네트워크는 다수의 수신자와 소통할 수 있는 플랫폼으로, 이 공간은 사이버 괴롭힘이나 증오 선동의 범죄 현장으로도 악용되는 것이 사실이다. 그 범위는 낮은 수위의 모욕부터 7만 명의 좋아요 지지를 받은 린치 요구까지 다양하다.³⁷⁾ 소셜네트워크는 정기적으로 텍스트, 이미지 및 비디오와 같은 모든 종류의 정보를 소위 “친구”에게 전송할 기회를 제공하고,³⁸⁾ 극단주의 테러 집단의 콘텐츠 역시 전송 또는 배포된다. 이와 관련하여 테러 사건의 피해자 유족이 온라인 플랫폼 기업을 상대로 소송을 제기한 바 있었으나, 2023년 2월 미국 연방대법원은 통신품위법 제230조를 통한 유튜브, 트위터 등 온라인 플랫폼의 면책특권을 인정하는 판결을 내렸다.³⁹⁾ 연방대

34) 유의선(2018), 가짜뉴스의 법적 규제- 사회적 법익 보호를 중심으로-, <언론과 법> 제17권 제2호, 61면.

35) 47 U.S.C Art. 230 Abs. c (1).

36) Wardle&Derakhshan, Information Disorder: Toward an interdisciplinary framework for research and policy making, (https://firstdraftnews.com/wp-content/uploads/2017/10/Information_Disorder_FirstDraft-CoE_2018.pdf?x40896, S. 61.)

37) Ostendorf&Frahm&Doege(2012), Internetaufrufe zur Lynchjustiz und organisiertes Mobbing, *NStZ* 2012, 529, 529.

38) Eckel&Rottmeier(2021), “Liken als Haten”: Strafverfolgung von Hatespeech in Sozialen Netzwerken, *NStZ* 2021, 1, 1.

39) 2015년 발생한 파리 테러사건의 피해자의 유족들이 유튜브가 극단주의 콘텐츠 배

법원의 판결로 테러와 관련한 알고리즘 기반 추천에 대하여도 플랫폼 사업자의 책임이 없다는 것을 확인하였지만, 켄타지 브라운 잭슨 대법관은 보충 의견에서 통신품위법 제230조가 알고리즘을 기반으로 추천하는 모든 경우의 면책 근거로 일반화할 수 없다고 지적하였다. 또한 이 판결 이후 통신품위법 제230조에 대한 비판 여론 역시 커지고 있어, 플랫폼 사업자의 면책특권에 대하여 제한적인 해석을 가하고자 하는 움직임이 크다.⁴⁰⁾

나. 매개자 규제의 정당성

의사 표현의 매개자, 즉 기술적 유포자는 의사 표현의 행위자가 즉각적 교란자(unmittelbarer Störer)와 다르게, 간접교란자(mittelbarer Störer)로 표현할 수 있다.⁴¹⁾ 소셜네트워크가 위법한 목적으로 개설된 것도 아니고, 운영자에게는 위법한 표현물이 유포되는 것에 대한 고의도 없으므로 플랫폼을 제공한 것 자체에 대하여 형법상 책임을 물을 여지는 없다.⁴²⁾ 그러나 소셜네트워크가 누구나, 어디서나 많은 사람들이 접근할 가능성을 제공하여 의사소통을 촉진하고, 이는 잠재적으로 무한한 청취자를 대상으로 순식간에 유포시키는 데 큰 역할을 하는 것은 명확한 사

포의 책임을 저야 한다며 소송을 제기하였다. 또한 2017년 이스탄불에서 발생한 IS 총기난사 사건의 피해자들 역시 트위터가 테러집단의 콘텐츠를 방조하고 있다며 트위터를 상대로 소송을 제기하였다. (2023.5.22. 이코리아, 미국당국 통신품위법 제23조 논쟁에 빅테크 손들여준 대법원, <https://www.ekoreanews.co.kr/news/articleView.html?idxno=67072> 2024.11.24. 최종확인)

- 40) 악성 알고리즘 방지법안(Justice Against Malicious Algorithms Act), 세이프 테크법안(SAFE TECH Act), 플랫폼 책임 및 소비자 투명성 법안(Platform Accountability and Consumer Transparency Act) 등 플랫폼 사업자와 같은 매개자에 의무를 부과하는 법안이 발의되었다. 자세한 내용은 문의빈(2023), 앞의 글, 420면 이하 참조.
- 41) Kastl-Riemann(2023), Algorithmen und Künstliche Intelligenz im Äußerungsrecht ZUM 2023, 578, 583.
- 42) 신상현(2017), 위법한 명예훼손 게시물의 삭제 및 차단 의무에 관한 소셜네트워크 운영자의 형법적 책임- NetzDG에 대한 비판적 검토, 인하대학교 <법학연구> 제20권 제3호, 64면.

실이다.⁴³⁾ 온라인 플랫폼 사업자가 사람들이 자유롭게 정보를 제공받고 소통할 수 있는 커뮤니케이션의 장을 제공하고, 이를 통해 가짜뉴스 및 혐오 표현을 포함한 타인의 인격권 침해는 물론 민주주의를 훼손할 위험에 기여했다는 점에서, 이러한 위험을 억제하는 데도 그 역할이 필수적이라고 볼 수 있다.⁴⁴⁾

(1) 의사소통에 기여하는 매개자로서 언론매체 유사 책임

언론사나 방송사와 다르게 검색엔진은 사용자의 정보수집 및 활동에 중요한 역할을 하지만, 그 자체로 여론 형성 과정에 큰 영향을 미치지 않는다는 이유로 검색엔진에 대한 규제는 적절하지도 필요하지도 않은 추가규제를 비판하는 의견이 존재한다.⁴⁵⁾ 이와 반대로 헌법상 방송 개념을 광범위하게 설정하여 뉴스피드와 같은 플랫폼도 방송에 포함함으로써 입법자는 통제 의무를 지게 되고, 방송 질서 내에서 규제를 위한 해결 방안을 모색하는 주장도 공존한다.⁴⁶⁾ 물론 온라인 플랫폼의 종류와 기능이 다양하므로 모든 플랫폼에 언론매체 유사 책임을 적용할 수는 없으나 뉴스피드의 포스팅은 프로필 설정에 따라 다르지만, 불특정 다수의 집단을 목표로 하며, 전파를 이용한 배포 및 독일연방헌법재판소가 요구하는 최신성과 시청자들에 대한 광범위한 영향력 요건을 충족한

43) Kühling(2021), »Fake News« und »Hate Speech« - Die Verantwortung der Medienintermediäre zwischen neuen NetzDG, MStV und Digital Services Act., ZUM, 2021, 461, 463.

44) Mondal&Silva&Benevenuto(2017), A Measurement Study of Hate Speech in Social Media, HT '17: Proceedings of the 28th ACM Conference on Hypertext and Social Media, 86 ff.

45) Nolte(2017), Hate-Speech, Fake-News, das Netzwerkdurchsetzungsgesetz und vielfaltsicherung durch Suchmaschinen, ZUM 2017, 563 ff.

46) Gerdorf(2019), Hate Speech in sozialen Netzwerken, MMR, 2017, 439, 444f.; Kaiser&IReiling, 소셜미디어를 통한 의견의 필터링- 그리고 의견다양성이라는 민주주의적 이상?, in : Demokratie und künstliche Intelligenz, 한국법제연구원(역), 88-98면.

다.⁴⁷⁾ 이는 제3자의 진술에 대한 책임 있는 매개자의 법적 의무를 인터넷의 특별한 특징이 아닌 언론의 차별화된 책임으로 접근하고자 하는 것이다.⁴⁸⁾ 오늘날 소셜미디어는 방송과 라디오가 그랬던 것처럼 여론 형성의 매체이자 요소이고, 언론의 자유를 ‘봉사하는’ 담당자로서 그 안에서 만들어지는 의사소통의 자유를 확보할 특별한 책임을 부여하는 것이다.⁴⁹⁾

사회적 플랫폼 운영자를 간접적으로 구속하는 것을 정당화하기 위한 다양한 관점 중 가장 핵심적인 관점은 바로 플랫폼이 일반적으로 접근 가능한 의사소통의 공간을 자발적으로 제공한다는 점이다. 이때 이 의사소통의 공간은 기타 맥락에서 이용자의 커뮤니케이션 자유의 행사에 대한 권리까지 포함한다.⁵⁰⁾

(2) 구조적 우월성

독일 연방헌법재판소는 경기장 방문 금지와 관련된 사안에서, 경기장 운영주는 “신분의 구별 없이 공평하게 다수의 관중에게 축구 경기를 관람할 수 있도록 개방”하였고, 그 기회로부터의 배제는 “당사자의 사회생활 참여에 있어 결정적”이므로 경기장의 잠재적 방문객을 자의적으로 불평등하게 대해서는 안 된다고 밝혔다.⁵¹⁾ 독점 또는 구조적 우월성에서 발생하는 결정권의 경우에도 이와 유사하게 적용될 수 있으므로 경기장 방문이 사회생활의 참여로 간주한다면, 광범위한 인구집단이 의사소통을

47) Kaiser& Reiling(2019), 소셜미디어를 통한 의견의 필터링- 그리고 의견다양성이라는 민주주의적 이상?, in : *Demokratie und künstliche Intelligenz*, 한국법제연구원(역), 88면.

48) Nolte(2017), Hate-Speech, Fake-News, das Netzwerkdurchsetzungsgesetz und vielfaltsicherung durch Suchmaschinen, *ZUM* 2017, 552, 553.

49) Wagner(2022), Kommunikation in sozialen Medien und Grundrechte - Zur Drittwirkung der Meinungsfreiheit und Ausgestaltung einer dienenden Medienfreiheit, *Zum* 2022, 861, 861.

50) BVerfGE 128, 226 (252f.)

51) 독일연방헌법재판소는 경기장 입장금지에 대한 판결에서 사법(私法)관계에서 간접적인 기본권의 제3자효의 의미를 확정하였다. (BVerfGE 148, 267, 283f.)

위하여 포괄적이고 집중적으로 사용하는 커뮤니케이션 플랫폼의 게시물
의 수신과 발신의 영역에도 적용되어야 한다는 것이다.⁵²⁾

온라인 플랫폼은 다른 제공자와 연결 내지 진입 가능성을 높여주거나
제한하는 프레임워크의 구조이다.⁵³⁾ 인터넷 포털사업자에 의해 의견 형
성의 선호 또는 의사에의 제어 가능성이 존재하고, 검색 및 선택 알고리즘
이 다수 이용자의 선호에 맞추게 됨으로써 자강 효과는 증가하고 결국
인기 있는 사이트나 콘텐츠에 유리한 구조적 불균형이 발생한다.⁵⁴⁾ 사적
플랫폼 운영자를 간접적으로 구속하는 것을 정당화하기 위한 또 다른
이유는 플랫폼 운영자와 이용자 간의 세력 불균형으로, 이는 주요 플랫
폼의 독점과 유사한 지위에 근거한다.⁵⁵⁾

(3) 관리와 통제의 효과적 지위

또한 영향력의 측면에서 가짜뉴스 또는 혐오 표현이 소셜네트워크에
도달하여 확산한 이후 사후적으로 효과적인 개입을 하기는 어렵다. 정보
는 쏠살같이 퍼지고 통제는 사실상 불가능하기 때문이다.⁵⁶⁾ 명백히 불법
적인 콘텐츠가 일반 사람들이 많이 접근하는 사이트에 나타나지 않도록
광고주가 가하는 압력은 경제적으로 훨씬 더 효과적인 것으로 판단된
다.⁵⁷⁾ 이와 같은 맥락에서 “Stop Hate for Profit(이익을 위한 증오를 멈

⁵²⁾ Antje von Ungern-Sternberg(2019), 민주적 의사형성과 인공지능, in: *Demokratie und künstliche Intelligenz*, 한국법제연구원(역), 24면.

⁵³⁾ 이부하(2018), 디지털 시대에 의사표현의 왜곡과 표현의 자유의 한계, 법학논문집 제42집 제3호, 50면.

⁵⁴⁾ Eric Goldman(2006), Search Engine Bias and the Demise of Search Engine Utopianism, *Yale J. L. & Tech.* p. 118.

⁵⁵⁾ Antje von Ungern-Sternberg(2019), 민주적 의사형성과 인공지능, in: *Demokratie und künstliche Intelligenz*, 한국법제연구원(역), 24면.

⁵⁶⁾ Berz&Engel&Hacker(2023), Generative KI, Datenschutz, Hassrede und Desinformation- zur Regulierung von KI-Meinungen, *ZUM* 2023, 586, 591.

⁵⁷⁾ Kühling(2021) »Fake News« und »Hate Speech« - Die Verantwortung der Medienintermediäre zwischen neuen NetzDG, MStV und Digital Services Act.,

취라.)”⁵⁸⁾는 광고 보이콧 캠페인도 존재한 바 있다.

플랫폼 사업자는 정보를 게시하고 사람들이 의사소통하는 공간을 제공하고 관리하는데 다양한 데이터를 취득하고 경제적 이익을 얻는 지위에 있으며, 이는 동시에 기술적으로 혐오 표현과 조작된 정보의 무분별한 확산을 방지할 수 있는 위치에 있다. 또한 플랫폼 사업자의 경우 명백한 불법콘텐츠가 타인의 사생활 및 인격권 등에 대한 침해 가능성을 충분히 인지할 수 있으므로, 기술적으로 그의 관리와 통제가 미치는 범위 안에서 확산 방지를 위한 주의의무를 다하여야 하는 것이다.⁵⁹⁾ 의사소통의 장인 인터넷 플랫폼은 혐오 표현이나 거짓 정보 등을 사전에 삭제할 수 있는 자동화된 시스템을 갖추고 신속히 대응할 수 있기 때문에 그에게 주어진 사회적 책임을 다함으로써,⁶⁰⁾ 불법적 콘텐츠의 유포를 막아 피해를 줄일 수 있다.

Ⅲ. 유럽 매개자 규제의 구체적 법제 검토

1. 독일 네트워크집행법(NetzDG)과 디지털서비스법(DDG)

가. 네트워크집행법

(1) 네트워크 집행법의 제정 및 내용

독일의 경우, 2015년 시리아 난민 등에 대한 모욕적 발언이 문제 되

ZUM, 2021, 461, 464.

⁵⁸⁾ <https://www.stophateforprofit.org>

⁵⁹⁾ 양형우(2009), “명예훼손에 대한 인터넷 종합 정보제공 사업자의 손해배상책임 - 대법원 2009. 4. 6. 선고 2008다53812 전원합의체 판결”, <재산법연구> 26권 2호, 206-207면.

⁶⁰⁾ 권형돈(2020), 앞의 글, 25면.

면서 SNS상 위법한 콘텐츠를 규제하는 법률(Netzwerkdurchsetzungsgesetz: NetzDG)이 만들어졌다.⁶¹⁾ 이 법은 유튜브나 페이스북 또는 트위터와 같이 SNS상에서 확산하는 혐오 표현과 가짜뉴스 등의 유포를 억제하고자 매개자를 규제하는 방식을 취하고 있으며,⁶²⁾ 플랫폼 운영자는 권리를 침해하는 콘텐츠의 신속한 삭제와 딥페이크의 불신으로 영향받는 데이터를 효과적으로 보장한다.⁶³⁾ 2017년 6개의 조항으로 만들어 네트워크집행법은 제1조에서 본 법의 적용 대상 범위를 규정하고, 제2조 보고의무(Berichtspflicht)와 제3조 불법콘텐츠에 대한 문제 제기 처리, 제4조 의무 위반에 따른 벌금에 관한 조항 및 제5조 서비스를 위한 국내 공인 대리인 제6조 경과규정을 두었다.⁶⁴⁾

이 법의 적용 대상은 사용자가 임의의 콘텐츠를 다른 사용자와 공유하거나 일반에 공개하는 데 사용되는 플랫폼을 영리 목적으로 운영하는 텔레미디어 서비스 제공자다.⁶⁵⁾ 플랫폼 사업자에게 의무를 부과하는 방식으로 불법콘텐츠의 확산 방지를 위한 입법 목적은 정당하지만, 사실 독일 네트워크집행법은 제정 당시 플랫폼 서비스제공자에게 사적 검열을 허용하는 결과를 초래하는 것이라는 주장과 함께 표현의 자유 침해 위험성을 이유로 여야 할 것 없이 반대가 심하였다.⁶⁶⁾ 네트워크집행법상 “의심스러운 경우 차단(Im Zweifel Sperren)”이라는 슬로건을 실현하기 위해 표현의 자유를 지나치게 제한하지 않도록, 명백히 불법적인 콘텐츠의 범위 설정이 중요한데, 콘텐츠의 불법성을 평가하기 위해서 사실 관계 및 법적 관계 모두 매우 면밀하고 복잡한 검토가 필요하다.

61) Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (BGBI, I S. 3352).

62) 권형돈(2020), 앞의 글, 17면.

63) Kumkar&Rapp(2022), Deepfakes, *ZfDR* 2022, 199, 219.

64) Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (BGBI, I S. 3352).

65) NetzDG Art, 1 Abs .1.

66) 권형돈(2020), 앞의 글, 17면.

SNS 사업자가 삭제하여야 할 불법콘텐츠는 테러 조직 선전자료의 제작 및 배포 등과 같이 독일 형법상 금지하는 행위를 대상으로, 동법에 구체적으로 언급된 내용은 위헌적·테러 조직의 선전자료 제작 및 배포(형법 제86조), 위헌적·테러 조직(예컨대 나치 등) 상징의 반포(형법 제86a조), 국가를 위태롭게 하는 중대한 폭력행위의 준비(형법 제89a조) 및 이를 문서로써 교사하는 것(형법 제91조), 국가의 반역적인 사실의 왜곡(형법 제100조a)과 범죄의 선동(형법 제111조), 범죄행위를 실행하려는 협박에 의하여 공공의 평온을 해하는 것(형법 제126조), 국내외 테러 및 범죄 조직의 결성(형법 제129조-129b조) 등과 관련한 표현은 삭제되어야 할 콘텐츠로 본다. 또한 불법적 콘텐츠로 분류되는 내용을 가진 형법상 범죄는 형법 제130조 대중선동죄로서 공공의 평온을 해치는 편파적 발언, 아우슈비츠의 거짓말, 나치의 폭력적 지배와 자의적 지배에 대한 찬양 및 경시를 금지하고, 형법 제131조의 폭력적 표현 그리고 형법 제166조에 따라 공공의 평화를 어지럽히는 방식으로 종교적 철학적 신념을 모욕하는 행위를 금지한다. 나아가 형법 제184b조에 해당하는 아동 포르노 소지 및 배포 금지, 형법 제184조부터 제184c조에 규정된 포르노 방송과 인터넷에 의한 공공연한 진열과 청소년에 대한 제공을 금지하며 제185조 모욕죄, 제186조 악평유포죄, 제187조 허위 사실 비방죄 등 명예훼손적 표현을 금지한다. 또한 사진 촬영으로 내밀한 삶의 영역과 개인의 권리를 침해하는 행위(형법 제201a조) 협박죄(형법 제241조) 또는 증거 데이터의 위조(형법 제269조)의 구성요건에 해당하는 내용을 말한다. 즉, 공공질서나 공적 평온을 교란하는 혐오범죄나 거짓 정보 그리고 내밀한 사적 영역을 침해하는 형법상 구성요건에 해당하는 불법 게시물을 규율 대상으로 한다.⁶⁷⁾ 그러나 이러한 규정에도 불구하고, 법원에서조차 범죄행위로서의 모욕 또는 명예훼손과 보호받아야 하는 표현의 자유 한계 내에서 진술을 구별하기 까다로운 경우가 많다. 하

⁶⁷⁾ Liesching(2018), *NetzDG Kommentar*, Nomos, § 1, Rn. 12

물며 “의심스러운 경우 차단(Im Zweifel Sperren)”이라는 슬로건을 실현하기 위해 짧은 삭제 기간 내 불법적 콘텐츠 판단의 어려움은, 신중한 검토 없이 우선 차단 조치를 하게 할 것이고, 이는 표현의 자유를 위축시킬 것이라는 우려가 있다는 점을 경청할 필요가 있다.⁶⁸⁾

(2) 네트워크집행법의 평가 및 개정

이렇듯 사적 검열의 문제와 표현의 자유 위축 효과 등에 대한 비판의 목소리가 끊이지 않자, 독일 네트워크집행법은 몇 차례 개정을 통하여, 신고당한 자와 신고자 모두의 이익제기 절차 등을 규정하고 제3의 기관에 의하여 재심을 받을 수 있는 규정을 통해 개선 방안을 마련하였다. 또한 동법 개정을 통하여 비디오(동영상) 공유플랫폼 서비스 제공 업체까지 통지 및 조치 의무 적용 대상자를 확대하는 등의 방향으로 개정하였지만,⁶⁹⁾ 후술하는 EU 디지털서비스법이 제정됨에 따라 네트워크집행법이 사실상 폐지되고, 관련 국내법을 제정비하였다.

네트워크집행법의 위헌성과 관련하여 독일 연방헌법재판소는 공익 보호를 위해 제정된 일반법에 의해 제한되는 기본권은 비례원칙에 반하지 않는 한 헌법합치로 해석될 필요가 있다는 입장이다.⁷⁰⁾ 연방헌법재판소가 시정해야 하는 헌법위반은 기본권의 의미와 범위, 특히 보호영역의 범위에 대한 근본적으로 잘못된 이해에 기반한 해석의 오류를 드러내는 경우를 의미하는 것이고, 법의 집행에 따라 일어날 수 있는 특정 개별인의 표현의 자유 위배 여부가 법 조항 자체의 위헌 여부를 판단하는 기초가 되는 것은 아니라고 밝혔다.⁷¹⁾ 플랫폼 사업자의 표현물 삭제에

68) Nolte(2017), Hate-Speech, Fake-News, das Netzwerkdurchsetzungsgesetz und vielfaltsicherung durch Suchmaschinen, ZUM 2017, 552, 556 이하.

69) 2022년 7월 21일 개정(https://world.moleg.go.kr/web/wli/lgsInfoReadPage.do?A=A&searchType=all&searchPageRowCnt=10&CTS_SEQ=48407&AST_SEQ=69&ETC=1)

70) BVerfGE 7, 198 (208).

대한 우려에 대해서, 플랫폼 사업자 역시 사용자의 기대에 부응하여 시장에서 살아남아야 하므로 자신의 이익을 위해 콘텐츠를 과도하게 삭제하지 않을 것이라고 밝히며, 대규모의 부당한 삭제의 가능성은 높지 않을 것이라 예견했고, 실제로 우려했던 ‘과잉 차단’은 발생하지 않았다고 확인하였다.⁷²⁾ 독일연방의회는 자유롭고 개방적인 민주적인 사회에서 조화롭게 함께 살아가기 위하여 증오범죄를 촉발할 수 있는 소셜네트워크 플랫폼 내의 혐오 표현과 기타 불법콘텐츠에 대항하고자 하는 법의 목표를 달성하는 것으로 평가하였다.⁷³⁾

그러나 네트워크집행법에 따라 독일에 등록된 이용자 수가 2백만 명 이상인 SNS 제공자에게 보고의무와 이의제기에 대한 처리 의무를 부과하던 규정은 폐지되었다.⁷⁴⁾ 다만 이는 독일 연방헌법재판소의 판단에서도 알 수 있듯이 플랫폼 사업자 등에게 책임을 부과하는 매개자에 대한 규제 방법에 대한 부정적 평가에서 기인한 것은 아니며, EU 디지털서비스법이 제정됨에 따라 독일도 관련된 국내법률을 정비하는 차원에서, 2024년 5월 6일 독일 네트워크집행법 규정을 폐지하고 독일 디지털서비스법을 제정하였다.

나. 독일 디지털서비스법

독일 디지털서비스법은 언론매체법에서 규정하는 방송을 제외하고, 동

71) BVerfGE 7, 198 (209).

72) 소셜 네트워크 Facebook이 투명성 보고서에서 발표한 수치를 확인하면, 2018년 1월 1일부터 6월30일까지 서비스 제공 업체는 886건의 불만 사항 중 362개의 콘텐츠만 삭제하였다. Facebook의 후속 보고서에 따르면 네트워크집행법으로 말미암아 “오버 블로킹”의 위험이 생성되지 않았으며, 2020년 상반기에 2,308건의 콘텐츠가 삭제된 반면, 2020년 하반기에는 그 수가 1,276건으로 거의 절반으로 감소한 것으로 나타났다. (https://about.fb.com/wp-content/uploads/2020/07/facebook_netzdg_July_2020_German.pdf)

73) BT-Drucksache 18/12356, S. 13.

74) <https://www.gesetze-im-internet.de/netzdg/BJNR335210017.html>. (최종확인 2024. 12.4.)

법의 다른 규정이 없는 경우 모든 디지털 서비스 제공자에게 적용된다.⁷⁵⁾ 여기에서 말하는 디지털 서비스는 정보사회 서비스, 다시 말해 일반적으로 보수를 받고 원거리에서 전자적 수단을 통하여 수신자의 개별적 요청에 따라 제공하는 모든 서비스를 의미한다.⁷⁶⁾

독일 디지털서비스법의 적용 대상에서 동영상 공유플랫폼 서비스도 구체적으로 명시하고 있는데, 이는 서비스 제공자가 주체적으로 편성이나 편집 등에 대한 책임을 지지 않는 방송물이나 사용자가 제작한 영상물을 일반 대중에게 제공하는 것을 주로 하는 디지털 서비스를 말한다.⁷⁷⁾ 또한 디지털 서비스를 제공하는 기업의 법적 관할권에 대한 구체적인 규정을 두며, 정보제공의무를 규정하고 있다. 서비스 제공자는 평소 유료로 제공하는 디지털 서비스에 대하여 쉽게 알아볼 수 있도록 일반 정보를 게시하여야 하며, 이름 및 개업지와 법적 형태는 물론이고 이메일 주소와 같은 신속하게 연락할 수 있는 정보와 관할감독청에 대한 정보 역시 게시할 것을 규정하고 있다.⁷⁸⁾ 본 법에 따르면 주법에 따른 관할관청은 독일이 소재국이거나 소재국으로 간주하는 시청각미디어서비스제공자와 동영상공유플랫폼제공자의 목록을 작성하여야 한다고 규정하고, 이러한 목록의 모든 갱신 사항을 언론매체를 관할하는 연방최고관청에게 전달할 것과 이를 다시 유럽연합집행위원회에 전달할 것을 규정하고 있다.⁷⁹⁾ 또한 권리침해 시 차단청구권을 규정하여, 통신망에서 이용자가 다른 사람의 지식재산권을 침해하였고 권리자가 다른 방법으로 는 침해를 시정할 수 없는 경우 해당 서비스제공자에게 정보 이용 차단을 요구할 수 있는 권리를 규정하고 있다. 다만 이러한 차단은 과도하지

75) Art. 1 Abs. 1 DDG.

76) Richtlinie (EU) 2015/1535 des Europäischen Parlaments und des Rates vom 9. September 2015 über ein Informationsverfahren auf dem Gebiet der technischen Vorschriften und der Vorschriften für die Dienste der Informationsgesellschaft, Art. 1. Abs. 1 (b)

77) Art.3 Abs. 8 (a) DDG.

78) Art. 5. DDG.

79) Art. 9. DDG.

않아야 하고 비례원칙에 부합할 것을 요구한다.⁸⁰⁾

또한 이용자의 위법행위와 관련하여 EU의 디지털서비스법의 제2장 중개서비스제공자의 책임의 제4조에서 제8조까지의 규정을 적용하도록 하여,⁸¹⁾ 실질적인 매개자에 대한 의무 내용을 EU 디지털서비스법(DSA)으로 대체하였다.⁸²⁾ 관련 사항은 EU 디지털서비스법에 대한 내용을 언급하면서 매개자 규제의 내용을 검토하도록 하겠다.

2. 프랑스 정보조작대처법 및 혐오·불법콘텐츠 대응법

가. 정보조작대처법

프랑스는 허위 정보 유포자인 행위자에 대한 규제(언론법 제27조, 선거법 제97조)와 동시에, 정치적 결정 과정을 왜곡하여 선거에 영향을 줄 수 있는 허위 정보의 확산 방지를 위하여 행위자뿐 아니라 매개자 역시 규제하는 방식을 취하고 있다. 프랑스는 일반적인 허위 정보 규제가 아닌, 선거기간에 한하여 선거에 영향을 미칠 것이 명백한 허위 정보 유포에 대응하기 위하여 매개자를 규제하는 법을 제정하였다. 선거법 제L97조를 통하여 새로운 위조물, 비방 또는 그 외의 기만적 술책으로 1인 또는 수인의 선거인들로 하여금 기권하도록 만드는 자는 1년 이하의 징역이나 15,000유로의 벌금에 처하고 있었다.⁸³⁾ 이후 프랑스는 2017년 가짜뉴스가 대선 기간에 심각한 정치, 사회적 문제로 등장하면서 ‘정보 조작 방지에 관한 법’ n° 2018-1202’을 발표하였다. 이 법은 선거 후보자가 선거기간 동안 가짜뉴스 및 허위 정보가 확산하는 것을 제한할 수

80) Art. 8 Abs. 1-2 DDG.

81) Art. 7 Abs. 1 DDG.

82) 이유진(2023), 네트워크집행법 대체하는 디지털서비스법, 더 이상 책임 피할 수 없는 글로벌 플랫폼, <신문과 방송>, 2023년 8월호, 한국언론진흥재단, 103면.

83) 한갑운·윤종민(2018), 가짜뉴스의 규율방법에 대한 법적 고찰, <과학기술과 법> 제8권 제1호, 72면.

있도록 하는 법으로, 후보자는 소셜미디어상의 거짓 정보를 선거 전 3개월 기간 동안 삭제요청을 할 수 있다는 내용을 담고 있다. 판사가 이에 동의하면, 해당 웹사이트를 차단하거나 폐쇄할 수 있게 된다. 본 법에 따르면 온라인 서비스 제공자에게 공적 질서를 혼란스럽게 할 수 있거나, 선거의 진실성을 변질시킬 수 있는 허위조작정보 유포에 대처하기 위한 조치를 실시할 의무를 부가한다.⁸⁴⁾ 법은 판사가 정보의 허위성과 관련하여 이의가 접수된 후 48시간 이내에 허위여부를 판단하여 삭제 등 배포 중지를 결정할 수 있도록 하고 있다. 판사의 개입이 정당화되기 위해서는 문제 된 정보가 ‘의도적으로’ 생성된 명백한 허위정보여야 하고, 대량으로 유포된 것이어야 한다.⁸⁵⁾

이에 대하여 판사가 단독으로 가짜뉴스의 개념 또는 선거의 진실성을 변질시키는 것에 대해 판단을 하는 것이 판사 개인에게 너무 과도한 부담을 줄 수 있다는 비판과 프랑스 학계의 부정적 평가가 존재한다.⁸⁶⁾ 표현의 자유에 대한 침해 가능성이 염려되지만, 프랑스 헌법위원회는 표현의 자유가 정치적 토론 및 선거운동에서 특별한 중요성을 가진다는 점을 전제하며, “일정한 정보 내용의 배포를 중단할 수 있게 하는 절차의 결과를 고려할 때, 문제가 되는 주장이나 비난은 부정확하거나 거짓된 특성이 명백한 경우에만 표현 및 통신의 자유를 침해하지 않는 것이고, 이와 같은 어떤 조치를 정당화할 수 있다. 선거의 진실성을 변질시킬 위협도 명백해야 함”을 판시하며, 문제 되는 주장 또는 비판에 대한 가처분이 허용되는 경우에 대한 한정적 해석을 전개하였다.⁸⁷⁾

84) 한동훈(2019), 프랑스 정보조작대처법의 합헌성에 관한 판단, <언론중재>, 108면 이하.

85) 이향선(2018), <가짜뉴스 대응개선을 위한 정책방안연구>, 방송통신심의위원회, 61면.

86) 한동훈(2019), 앞의 글, 117면.

87) 관련된 자세한 내용은 한동훈(2019), 앞의 글, 114면 이하 참조.

나. 혐오·불법콘텐츠 대응을 위한 아비아법과 헌법위원회의 결정

또한 프랑스는 2020년 5월 독일 네트워크집행법과 유사한 인터넷상 불법콘텐츠 대응에 관한 법률⁸⁸⁾을 제정하였다. 아비아 의원이 발의해서 제정된 법률이라 일명 아비아법으로 불리는 이 법 역시 불법콘텐츠를 적극적으로 규정하지 않고 다른 법률에서 규제되는 표현을 규율 대상으로 한다.⁸⁹⁾ 네트워크집행법과 유사하게 플랫폼 사업자에게 신고가 들어오면 24시간 이내에 명백한 불법콘텐츠를 삭제할 의무를 부과한다. 이를 위반한 사업자는 1년 이하 구금 및 25만 유로 이하 벌금으로 처벌하도록 규정하였고, 테러를 미화하거나 선동하는 콘텐츠 또는 아동 음란물의 경우 행정당국의 요청이 있으면 1시간 이내 삭제해야 함을 규정하였다.⁹⁰⁾

프랑스 법률 역시 논의 과정에서 무분별한 삭제 위험에 대한 비판에 직면하였고, 신고 남용을 방지하고자 신고자 신상정보를 제출하도록 하거나 악의를 가지고 신고한 사람에 대한 처벌 조항이 신설되었다. 또한 온라인 플랫폼 사업자가 내리는 결정의 자의성을 보완하기 위하여 투명성에 관련된 내용을 보장하였다.⁹¹⁾ 그럼에도 불구하고 프랑스 헌법재판소에 사전적 위헌법률심판이 청구되었고 이 법의 대부분 조항에 대해 위헌결정이 내려졌다.⁹²⁾

프랑스 헌법위원회는 아비아 법률안의 입법 목적과 관련하여, 아동 포르노그래피의 배포와 테러활동 유발이나 이를 찬동하는 것은 공공의 질서와 제3자의 권리를 심히 침해하는 것으로 표현의 자유와 통신의 자유

88) loi n° 2020-766 du 24 juin 2020 visant à lutter contre les contenus haineux sur internet)

89) 임효준(2023), 인터넷 서비스 제공자를 통한 표현규제에 관한 절차적 보호장치의 탐색-최근 유럽의 사례를 중심으로-, 〈언론과 법〉 제22권 제1호, 263면.

90) 전학선(2020), 프랑스에서 인터넷상의 혐오표현에 대한 규제- 헌법재판소 2020.6. 18. 아비아법률결정으로 중심으로, 〈외법논집〉, 제44권 제3호, 52면.

91) 임효준(2023), 앞의 글, 266면

92) Décision n° 2020-801 DC du 18 juin 2020.

결정의 자세한 내용은 전학선(2020), 앞의 글, 51면 이하 참조.

를 남용하는 것으로 보고, 이에 대한 표현물의 확산을 막기 위해 행정청의 요청으로 서비스 운영자에게 삭제 의무를 부과하는 것은 정당하다고 보았다.⁹³⁾ 그러나 이용자가 신고하면 사업자가 불법콘텐츠⁹⁴⁾를 삭제해야 하는 의무를 지는데, 불법을 판단함에 있어 모호하기에 명확성 원칙에 위배되고, 이를 판단하기 위해 주어진 시한은 24시간으로 너무 짧아, 플랫폼 사업자는 처벌을 피하기 위한 목적으로 우선 삭제 결정을 하게 될 위험이 있고, 법관은 어떠한 개입도 할 수 없는 점을 지적하여 위헌임을 선언하였다.⁹⁵⁾

3. EU 디지털서비스법(DSA)

가. 적용 범위의 구분 및 책임의 차별화

2020년 12월, EU 집행위원회는 디지털 중개서비스에 대한 EU 차원의 통일된 의무를 이행하고, 시대에 뒤떨어지는 것으로 인식되는 온라인 서비스 책임에 관한 전자상거래 지침의 법적 체계를 현대화하기 위한 디지털서비스법 초안을 발표하였다.⁹⁶⁾ 그리고 2023년 8월 25일 온라인상 불법·허위 콘텐츠가 유포되는 것을 방지하는 글로벌 플랫폼 기업에 수 십억 유로의 벌금을 부과하는 유럽연합 디지털서비스(Gesetz über digitale Dienste, 이하 DSA)법이 발효되었고, 2024.2.17.부터 시행되었다.⁹⁷⁾

⁹³⁾ 임효준(2023), 앞의 글, 263면.

⁹⁴⁾ 본 법에 적용되는 불법인 콘텐츠는 출신지나 국가와 민족, 인종 및 종교, 성별, 성적 지향 등을 이유로 차별하거나 혐오감정으로 인하여 개인 또는 단체에 대한 폭력을 선동하는 행위, 인종학살과 노예착취행위, 전쟁범죄 등 반인류 범죄행위를 부정 또는 축소하는 행위와 성폭력, 테러리즘에 대한 찬양·선동행위, 미성년자가 보거나 인지할 수 있는 상황에서 음란 요소가 있는 메시지를 유포하는 행위 등을 의미한다.

⁹⁵⁾ 전학선(2020), 앞의 글, 59면.

⁹⁶⁾ Kumkar&Rapp(2022), Deepfakes, *ZfDR* 2022, 199, 220.

⁹⁷⁾ Verordnung (EU) 2022/2065 des Europäischen Parlaments und des Rates vom 19. Oktober 2022 über einen Binnenmarkt für digitale Dienste und zur Änderung

이 법은 알고리즘을 통한 권력의 사회적 영향력을 통제하고 시민의 권리와 민주주의에 영향을 규제하기 위하여 고안되었다.⁹⁸⁾ 디지털서비스법을 통한 중개서비스 제공자의 책임 있고 성실한 행동은 안전하고 예측 가능하며 신뢰할 수 있는 온라인 환경을 조성하는 데 기여할 것이다. 또한 이 법은 정보를 중개하고 유통하는 서비스 업체의 의무와 책임을 부여하는 법이지만, 이와 동시에 중개서비스 제공업체가 사용자들의 콘텐츠를 과도하게 규제하거나 합법적 사용자의 콘텐츠를 삭제하는 일이 발생하지 않도록 균형을 맞추어야 한다.⁹⁹⁾

또한 불법 정보의 유통에 대한 책임은 각 서비스의 유형에 따라 차등적으로 적용된다.¹⁰⁰⁾ 디지털서비스법 대부분의 조항은 정보화 서비스 일반이 아닌 중개서비스, 즉 정보를 직접 생산하지 않고 특정 방식에 따라 유통하는 사업자들의 의무와 책임을 다루고 있으며,¹⁰¹⁾ 중개서비스 내에서도 단순 전송서비스¹⁰²⁾, 캐싱 서비스¹⁰³⁾, 호스팅 서비스¹⁰⁴⁾ 그리고 온라인 플랫폼과 초대형 온라인 플랫폼 등 규모에 따라 분류하고 이를 바탕으로 차등적 규제를 명시하고 있다. 단순한 전송과 캐싱, 호스팅의 경

der Richtlinie 2000/31/EG (Gesetz über digitale Dienste) 디지털서비스법의 형식은 명령(Regulation)으로 회원국에 직접 적용된다.

98) 박찬경(2022), 유럽연합 디지털서비스법(Digital Services Act): 규제냐, 자율이냐를 넘어서, 〈언론중재〉 통권 제164호, 63면.

99) Wiedemann(2024), Transparenzberichtspflichten für Anbieter von Online-Plattformen nach dem DSA, ZUM 2024, 364, S. 364.

100) 이권일(2024), 독일에서의 규제 논의의 현황과 과제 - 독일 네트워크법(NetzDG)과 EU 디지털서비스법(DSA)을 중심으로 -, 〈공법학연구〉 제25권 제2호, 8면.

101) DSA Art. 1 Abs.1.

102) 단순 전송은 인터넷 전송서비스와 같이 서비스 이용자가 제공하는 통신네트워크 내의 정보전송 또는 통신네트워크에 대한 접근제공으로 구성된다.

103) 캐싱 서비스는 서비스 이용자가 제공한 정보의 통신망 내 전송으로 구성된다. 캐싱은 이용자의 요청에 따라 다른 이용자에게 보다 정보가 효율적으로 전송될 수 있도록 해당 정보의 임시적 저장을 포함한다. (예: 제한된 기간만 정보를 저장하는 정보 캐피 네트워크)

104) 호스팅은 서비스 이용자가 제공하거나 그 요청에 따라 정보를 저장하는 서비스를 말한다. (대표적 예: SNS플랫폼, 비디오, 이미지 또는 오디오 공유서비스, 파일 공유서비스, 클라우드 서비스 등)

우, 중개자가 제3자 콘텐츠에 대한 인지를 하지 않는 한, 원칙적으로 통지 및 조치의 책임이 없다.¹⁰⁵⁾ 여기에서 중개서비스의 개념은 정보의 매개라는 사실적·기술적 의미를 갖는 것으로, 우리 정보통신망법상 정보통신서비스 제공자의 개념 정의(제2조 제3호)에서 사용하는 ‘매개’ 개념과 일치하는 것으로 볼 수 있다.¹⁰⁶⁾

본 법에 따르면 중개서비스 제공자는 신의와 성실에 따라 불법콘텐츠를 탐지하고 식별하여 삭제 또는 접근차단조치를 목적으로 하는 기타 조치를 하거나, 본 규정에 명시된 요건을 포함하여 유럽연합법 및 국내법 요건을 준수하기 위해 필요한 조치를 했다면 단순 전송, 캐싱, 호스팅 서비스에 언급된 책임면제를 받을 자격이 있다고 봐야 한다.¹⁰⁷⁾ 이러한 중개서비스 제공자에게는 전송되거나 저장하는 정보를 감시하거나 불법행위를 나타내는 상황을 적극적으로 조사할 일반적인 모니터링 의무가 부과되지 않는다.¹⁰⁸⁾

반면 디지털서비스법은 서비스사용자가 역내 인구의 10%(약 4,500만 명)를 넘는 기업에 대해서는 초대형 온라인 플랫폼으로 분류하고 고강도 규제함으로써,¹⁰⁹⁾ 전 세계 국민들이 사용하는 페이스북, 틱톡, X(구 트위터), 유튜브, 인스타그램, 링크트인, 핀터레스트, 스냅챗 등 8개 소셜미디어 기업과 아마존, 알리바바 등 5개 전자상거래업체는 체계적 위험을 처리하기 위한 추가 의무를 진다. 구글플레이, 애플 앱스토어를 비롯한 모바일 앱스토어, 마이크로소프트의 Bing 검색엔진 등도 이에 포함된다.¹¹⁰⁾ 초대형 온라인 플랫폼과 검색엔진의 경우 특수한 역할과 범위를

¹⁰⁵⁾ vgl. Art. 3 ff. DSA

¹⁰⁶⁾ 이병준(2021), 법체계 측면에서 유럽연합 디지털서비스법과 플랫폼 규제 특징, 〈유통법연구〉 제8권 제1호, 50면.

¹⁰⁷⁾ Art. 7 DSA

¹⁰⁸⁾ Art. 8 DSA.

¹⁰⁹⁾ Art. 33-43 DSA.

¹¹⁰⁾ 한국경제(2023), 가짜뉴스 방치하면 대규모 벌금 EU, 디지털플랫폼 규제법 발효. 2023.08.27.

고려할 때 이용약관에 관한 추가 요건을 부과하는 것이 합당하다고 언급한다.¹¹¹⁾ 이와 같은 초대형 온라인 플랫폼은 사전 예방적 위험 관리 차원에서 모든 위험에 대한 분석 및 평가가 필요하고, 적절하고 효율적인 위험 완화 조치를 취해야 하고, 독립적인 감사를 포함하는 포괄적인 규정을 준수하도록 시스템을 구축해야 한다.¹¹²⁾ 디지털서비스법은 중개사업자가 유통하는 정보에 대한 책임이 있음을 원칙으로 하되, 특수한 면책사유에 대하여 규정하는 방식으로 다양한 중개사업자의 차등적 책임을 부여한다.¹¹³⁾

나. 투명성 제고를 위한 장치

투명성 원칙은 공정성과 깊은 연관성을 가지는 개인정보보호를 위한 기본적 원칙으로, 알고리즘에 의한 의사결정 시 규범적 측면에서 그 기반이 되는 원칙 및 논리가 이용자에게 적절히 설명될 수 있어야 함을 의미한다.¹¹⁴⁾ 디지털서비스법 전문에서는 중개서비스 제공자의 계약의 자유는 원칙적으로 존중되어야 하지만, 투명성과 서비스 이용자의 보호, 불공정하거나 자의적 결과를 방지하기 위한 이용약관 내용과 적용 및 시행에 관한 특정 규칙을 설정하는 것이 적절하다 밝히고 있다.¹¹⁵⁾ 디지털서비스법 제15조는 중개서비스 제공자의 투명성 보고의무를 규정하고, 최소한 1년에 한 번 쉽게 접근할 수 있는 방식으로 관여된 모든 콘텐츠 조정에 대한 보고서를 공개적으로 제공할 것을 요구한다.¹¹⁶⁾ 또한 온라인 플랫폼 제공자는 제15조의 보고서 동조에서 정한 정보 이외에도 명

111) DSA 전문 (48).

112) Art. 33-37 DSA.

113) 박찬경(2022), 앞의 글, 66면.

114) 이대회(2024), 인공지능에 의한 개인정보의 자동 처리에 대한 투명성 확보에 관한 연구, <저스티스> 통권 제200호, 한국법학원, 304면.

115) DSA 전문 (45).

116) Art. 15. DSA.

백히 불법적 콘텐츠 제공한 경우와 명백한 근거 없이 신고에 따른 중단과 불만 제출에 따른 중단을 구분하는 등 추가적 정보를 포함하여 보고할 것을 제24조에서 규정하고 있다. 나아가 추천시스템의 투명성 제고를 위하여 동법 제27조 제1항은 “이용약관 등에 추천시스템에 사용되는 주요 매개변수(parameters)와 이용자가 주요 매개변수를 변경하거나 영향을 줄 수 있는 옵션을 최소 1개 이상 명확하고 이해하기 쉽게 명시하도록 규정하고 있다.” 그리고 동조 2항에서 온라인 플랫폼이 알고리즘을 기반으로 추천하는 결정적 기준이 무엇인지, 매개변수들이 상대적으로 중요한 이유에 대하여 이용자들에게 설명할 의무를 규정한다. 한편 온라인 플랫폼은 추천시스템에 영향을 줄 수 있는 옵션을 공개한 것을 기반으로 이용자가 이를 선택하거나 변경할 수 있도록 장치를 마련해야 하고 이러한 기능에 쉽게 접근할 수 있도록 동조 3항에서 규정하고 있다.¹¹⁷⁾

디지털서비스법은 알고리즘 자체가 아니라 그에 투입되는 매개변수로 어떤 것들이 있으며, 가장 중요한 기준이 무엇이고, 매개변수별 차이를 둔 이유가 무엇인지 설명하도록 규정하고 있다.¹¹⁸⁾ 호스팅 서비스 제공업체는 개인 또는 단체가 불법콘텐츠로 간주하는 개별 항목이 서비스에 존재한다는 사실을 해당 개인 또는 단체에 통지할 수 있는 절차를 수립해야 하고, 이러한 절차는 쉽게 접근할 수 있고 사용자 친화적이어야 하며 전자적 수단을 통해서만 통지를 전송할 수 있도록 해야 한다. 이 외에도 제39조 추가적 온라인 광고의 투명성 및 제42조 투명성 보고의무를 규정하여 대규모 온라인 플랫폼과 대규모 검색엔진이 감사보고서를 받은 후 3개월 이내 제34조에 따른 위험성 평가, 제35조 제1항에 따른 위험성 완화 조치, 제37조 제4항에 따른 감사보고서 및 제37조 제6항에 따른 감사 이행 보고서 등을 발간하고 디지털 서비스 조정기관 및 집행위원회에 송부하도록 규정하고 있다.¹¹⁹⁾ 다만 투명성 의무의 영향을 받

117) Art. 27 DSA.

118) 문의빈(2023), 앞의 글, 428면.

119) 문의빈(2023), 앞의 글, 430면.

는 온라인 플랫폼의 경우 예컨대 동법 제24조 제5항의 투명성 의무를 어떻게 이행해야 하는지 구체적인 방법이 명시되어 있지 않고, 특히 투명성 의무의 적용 범위와 유럽위원회에 대한 정보전송의 실질적 이행과 관련한 실무적 모호함이 존재하여 비판이 있다.¹²⁰⁾ 현재 상황에서는 입법 공백과 모호성에 대한 문제는 구체적 사건을 다루는 과정에서 구체화 될 것을 기대하며, 동법 제24조 제5항의 적용 범위는 광범위하게 해석되어야 한다고 볼 수 있다. 즉, 콘텐츠 제한의 결정 및 제한의 정당성을 전달할 의무의 범위가, 온라인 플랫폼 제공자가 실제로 배포한 콘텐츠뿐 아니라 플랫폼에서 제공되지 않았고 업로드 필터를 통해 사전에 삭제된 콘텐츠도 포함하여 보고할 것으로 해석하는 것이 적절하겠다.¹²¹⁾

다. 규제가능한 불법콘텐츠의 범위와 한계

한편 디지털서비스법에서 의미하는 ‘불법콘텐츠’란 해당 법률의 정확한 주제나 성격과 무관하게 제품의 판매나 서비스 제공을 포함하는 활동 자체 또는 이에 대한 언급이 유럽연합법 또는 회원국의 법률을 준수하지 않는 모든 정보를 의미한다.¹²²⁾ 다시 말해, EU 디지털서비스법은 불법 콘텐츠를 해석하는 데 독일 네트워크집행법상의 형법 내용만 다루는 것이 아니라 이미지, 저작권 또는 데이터 보호법의 관련 조항을 위반하는 것만으로도 네트워크집행법에 비해 광범위한 조치가 가능하다.¹²³⁾

그렇다고 이것이 가짜뉴스의 범람을 완벽히 막을 수는 없다. 왜냐하면 디지털서비스법에 따르더라도 유럽연합 또는 국내법에 따라 불법적인 콘텐츠에만 적용되는 것이고, 진실이 아닌 정보, 즉 가짜뉴스가 모든 경

¹²⁰⁾ Wiedemann(2024), Transparenzberichtspflichten für Anbieter von Online-Plattformen nach dem DSA, *ZUM* 2024, 364, S. 365.

¹²¹⁾ Wiedemann(2024), Transparenzberichtspflichten für Anbieter von Online-Plattformen nach dem DSA, *ZUM* 2024, 364, S. 368.

¹²²⁾ Art. 3 h) DSA

¹²³⁾ Kumkar&Rapp(2022), Deepfakes, *ZfDR* 2022, 199, 223.

우 불법적인 것은 아니기 때문이다.¹²⁴⁾

EU 디지털서비스법하에서 딥페이크 기술을 이용한 불법콘텐츠에 대한 규제는 가능할 수 있으나, 이 경우에도 플랫폼 사업자가 딥페이크의 신원 확인을 신뢰할 수 있는지에 대한 문제와 상충하는 기본권 입장 간의 복잡한 균형 조정 과정이 여전히 남아 있다. 또한 디지털서비스법상 구속력 있는 삭제 기한이 없다는 점을 고려할 때, 이 법 하에서 더욱 시급한 문제는 삭제 결정의 시간적 차원이다. 이러한 이유로 딥페이크 문제와 관련해서는 법적 규제보다 더 효과적인 것은 기술적 자체에 대한 규제라고 설명한다.¹²⁵⁾

딥페이크 기술 자체의 사용금지를 주장하는 경우도 있지만,¹²⁶⁾ 이보다는 기술의 발전을 통해 이를 감지하는 시스템을 함께 발전시켜,¹²⁷⁾ 디지털 플랫폼 사업자가 인공지능 영역에서 이루어지는 기술을 바탕으로 불법콘텐츠를 식별하고 필터링하는 방법이 바람직할 것이다. 인공지능 기반 필터링은 정책 입안자가 사업자의 기술적 역량에 의존하여 문제를 해결하는 방식으로, 이와 같은 규제 방식은 검열 시비 논란을 피하고 표현의 자유를 훼손하지 않은 채 규제 목적을 달성하기에 선호되는 방식이다.¹²⁸⁾ 중고차 판매자가 차량의 문제점을 미리 공개하도록 의무화하는 법과 같이, 소셜 미디어 플랫폼에 조작이 있는 동영상과 오디오를 확인하여 표시하도록 하는 법은 사용자가 자신이 시청한 영상이 조작없는

¹²⁴⁾ Berz&Engel&Hacker(2023), Generative KI, Datenschutz, Hassrede und Desinformation- zur Regulierung von KI-Meinungen, *ZUM* 2023, 586, 593.

¹²⁵⁾ Kumkar&Rapp(2022), Deepfakes, *ZfDR* 2022, 199, 223-224.

¹²⁶⁾ Linardatos(2022), „Das Ende der Verlässlichkeit“, LTO v. 27.7.2021, <https://www.lto.de/recht/hintergruende/h/deepfakes-regulierung-europa-eu-schaden-demokratie-manipulation/> [20.07.2022].

¹²⁷⁾ 정교한 딥페이크 탐지 기술은 아직 개발되지 않았지만, 국가 안보 차원에서 상당한 자원을 투자하여 신속하게 개발한다면, 누군가가 변조된 미디어를 게시했을 때 이를 적발할 수 있는 적절한 수단이 곧 개발될 수 있을 것이라는 전망도 존재한다. (Hasen(2019), Deep Fakes, Bots, and Siloed Justices: American Election Law in a Post-Truth World, *UC Irvine School of Law Research*, p. 549.)

¹²⁸⁾ 송해엽(2021), 앞의 글, 17-18면.

진실이라고 믿고 의사결정을 할 수 있도록 돕는다.¹²⁹⁾ 딥페이크에 대한 라벨링 의무를 부과함으로써 플랫폼과 네트워크 사업자의 삭제 및 차단 의무를 보완하여 딥페이크의 투명성을 높이고 피해자를 보호하는 동시에 딥페이크 기술을 합법적으로 안전하게 사용할 수 있는 방안을 계속 고찰해 가야 할 것이다.¹³⁰⁾ 이러한 기술적 수단은 일반적으로 언론의 자유에 대한 규제보다 덜 침익적으로 평가되며, 탐지 기술이 계속 발전함에 따라, 진실한 라벨링 제도는 위조라는 심각한 문제를 해결할 수 있는 최선의 방법이며, 적절하지 않은 정보로 인해 제대로 된 의사결정을 할 수 없는 문제를 해결해 준다.¹³¹⁾

디지털서비스법의 제정에도 불구하고 아직 여러 한계점이 존재하지만, 기술적 발전과 함께 법과 제도를 보완하여 새로운 위협에 대한 대응 방안을 마련해야 할 것이다.

IV. 결론

정보통신기술의 비약적 발전과 데이터 처리능력이 급성장하면서 어느 때보다 생각과 사상에 대한 표현물을 만들고 전파할 환경이 구축된 반

¹²⁹⁾ Hasen(2019), Deep Fakes, Bots, and Siloed Justices : American Election Law in a Post-Truth World, *UC Irvine School of Law Research*, p. 550.

¹³⁰⁾ Kumkar&Rapp(2022), Deepfakes, *ZfDR* 2022, 199, 226.

¹³¹⁾ Hasen(2019), Deep Fakes, Bots, and Siloed Justices : American Election Law in a Post-Truth World, *UC Irvine School of Law Research*, 2019, p. 550. 그러나 알고리즘을 이용하여 현실과 허구를 구별하는 것에도 한계가 있다, 이는 소위 고양이와 쥐 게임에 비유되는데, 가짜 콘텐츠를 탐지하는 방법이 개선될 때마다 자동으로 생성되는 방법도 개선되기 때문이다. 또한 생성형 AI인 Chat GPT가 혐오표현이나 인종차별을 생성하거나 원자폭탄을 만드는 방법 등을 공개하지 못하도록 필터를 걸어두지만, 악의적 행위자는 콘텐츠 필터를 우회하는 방법에 대한 지침을 찾을 수 있고, 연구원들은 Chat GPT를 사용하여 재배포 코드를 포함하여 전체 혐오표현 캠페인을 조율할 수 있음을 보여주었다.(Beuth(2023.1.12.), Wie sich Chat GPT mit Worten hacken lässt, *Spiegel Online*)

면, 이로 인한 피해 역시 심각하다. 이에 자아실현의 발판이자 민주주의의 근간이 되는 표현의 자유에도 일정 제한이 불가피한데, 디지털 사회에서 쟁점이 되는 것은 표현 행위자 자체의 처벌을 넘어, 이를 매개하는 플랫폼 사업자에 대한 규제이다.

국가별로 인격권 및 다른 법익과의 형량 아래 표현의 자유를 제한하는 규정은 이미 두고 있으나, 인터넷의 특성상 저장의 영구성 및 유포의 신속성 등으로 말미암아 그로 인한 피해는 건잡을 수 없고 불법콘텐츠의 확산을 막기 위해서는 플랫폼사업자의 협조가 절대적이다. 혐오 표현 내지 허위 정보의 확산에 최첨단 기술이 접목되면서, 이를 규제하기 위한 기술의 발전과 함께 행위자 그리고 플랫폼 사업자와 같은 매개자에 대한 법적 규제 역시 불가피한 상황이다. 우리 법제 역시 유럽과 유사하게 각 법익 침해 유형별로 행위자를 규제하고, 정보통신망법 등을 통하여 매개자에게 조치 의무를 부과한다. 이는 이용자를 직접 규제하는 데 있어 발생하는 실질적 어려움을 극복함과 동시에, 표현의 자유를 제한한다는 비판을 누그러뜨리는 방식으로 매개자 규제 방식을 취하고 있다. 이는 국가기관에 의한 규제라는 비난을 피하고 사적 규제의 법제화로 헌법적 논란을 줄이려는 선택이지만, 사적 검열의 문제 등 제기되는 여러 비판점 역시 인지하여야 할 것이다.

불법콘텐츠의 무분별한 확산을 막기 위하여 매개자를 규제하면서도 이로써 발생할 수 있는 표현의 자유 위반 요소 검토와 적법절차의 준수 등의 면밀한 형량과 고민이 필요하다. 특히 EU 디지털서비스법 제정을 통해 유럽의 각 국가는 EU의 디지털서비스법을 준용하도록 함으로써 관련 법령을 대체하거나 유럽연합 법 규정의 준수를 위한 작업을 진행하였다. 디지털서비스법은 여러 비판에도 불구하고 불법콘텐츠 확산 방지의 기여가 인정되는 독일의 네트워크집행법과 비슷한 맥락의 법으로 볼 수 있다. 민주사회에 지대한 영향을 미치는 표현의 자유 및 인격권의 조화를 모색하되, 플랫폼 기업 영업의 자유 역시 과도하게 제한되어서는 안 된다. 이러한 맥락에서 EU의 디지털서비스법의 책임의 차별화 정책

은 의의가 있다. 또한 매개자 규제의 효율성 및 필요성을 인지하면서도 불법콘텐츠에 대한 조치 시 이를 판단하는 주체를 넓히거나, 의견을 청취하는 등의 절차적 측면을 보완하고, 표현의 자유 침해 가능성을 최소화할 다각적 노력이 필요하다. 또한 원론적으로는 표현의 자유를 억압할 위험성을 감수하면서 강제하는 법적 규제가 바람직하지 않기 때문에 미디어 리터러시 교육 등을 통하여 허위조작정보의 유포 등이 스스로 억제되는 사회환경을 만드는 노력 역시 필요할 것이다. 이런 면에서 팩트체크링이나 분쟁조정을 담당하는 기구 등이 제대로 작동할 수 있도록 합법적 절차와 기준을 마련하고, 재원과 전문 인력의 지원, 인증 시스템 부여 등이 사회적 차원에서 함께 강구되어야 할 것이다.

■ 참고 문헌

- 권형돈(2020), 인터넷에서 표현의 자유와 그 규제법리-혐오표현·거짓정보에 대한 독일의 대응을 중심으로, <사법> 제1권 제51호.
- 김성화(2019), 독일의 SNS규제와 표현의 자유, <인권법평론> 제22호.
- 마크 코겔버그지음/배현석 옮김(2022), <인공지능은 왜 정치적인 수밖에 없는가>, 생각이음.
- 문의빈(2023), 알고리즘 기반 추천과 플랫폼의 기본권 제한에 관한 비교법적 연구, <헌법학연구> 제29권 제2호.
- 박주일(2020), 딥페이크 기술의 발전과 저널리즘의 새로운 위협-이미지 공정성 확보를 위한 영상 팩트체크의 필요성, <언론중재>, 2020 가을호.
- 박찬경(2022), 유럽연합 디지털서비스법(Digital Services Act): 규제냐, 자율이나를 넘어서, <언론중재> 2022년 가을호(통권 제164호).
- 배상균(2021), 딥페이크(Deepfake)영상물에 관한 법적 대응조치 검토, <4차 산업혁명 법과 정책> 제3호.
- 송해엽(2021), 인공지능 기반 콘텐츠 필터링의 사회적 합의와 규제된 자율 규제, <방송통신연구>, 2021 겨울호.
- 신상현(2017), 위험한 명예훼손 게시물의 삭제 및 차단 의무에 관한 소셜네트워크 운영자의 형법적 책임- NetzDG에 대한 비판적 검토, 인하대 학교 <법학연구> 제20권 제3호.
- 양형우(2009), “명예훼손에 대한 인터넷 종합 정보제공 사업자의 손해배상 책임 -대법원 2009. 4. 6. 선고 2008다53812 전원합의체 판결”, <재산법연구> 26권 2호.
- 유의선(2018), 가짜뉴스의 법적 규제- 사회적 법익 보호를 중심으로-, <언론과 법> 제17권 제2호.
- 이권일(2024), 독일에서의 규제 논의의 현황과 과제 -독일 네트워크법(NetzDG)과 EU 디지털서비스법(DSA)을 중심으로-, <공법학연구> 제25권 제2호.
- 이민영(2020), 딥페이크와 여론형성 -알고리즘의 권력화와 탈진실의 규제담론-, <미국헌법연구> 제31권 제1호.
- 이병준(2021), 법체계 측면에서 유럽연합 디지털서비스법과 플랫폼 규제의 특징, <유통법연구> 제8권 제1호.

- 이부하(2018), 디지털 시대에 의사표현의 왜곡과 표현의 자유의 한계, 〈법학논문집〉 제42집 제3호.
- 이향선(2018), 〈가짜뉴스 대응개선을 위한 정책방안연구〉, 방송통신심의위원회
- 임효준(2023), 인터넷 서비스 제공자를 통한 표현규제에 관한 절차적 보호 장치의 탐색-최근 유럽의 사례를 중심으로-, 〈언론과 법〉 제22권 제1호.
- 전학선(2020), 프랑스에서 인터넷상의 혐오표현에 대한 규제- 헌법재판소 2020.6.18. 아바아법률결정으로 중심으로, 〈외법논집〉, 제44권 제3호.
- 정애령(2021), 코로나 시대 공정한 선거를 위한 자유로운 선거운동 방안 모색, 〈유럽헌법연구〉 제37호.
- 한갑윤·윤종민(2018), 가짜뉴스의 규율방법에 대한 법적 고찰, 〈과학기술과 법〉 제8권 제1호.
- 한동훈(2019), 프랑스 정보조작대처법의 합헌성에 관한 판단, 〈언론중재〉.
- Anna-Bettina Kaiser/Ines Reiling(2019), 소셜미디어를 통한 의견의 필터링- 그리고 의견다양성이라는 민주주의적 이상?, in :*Demokratie und künstliche Intelligenz*, 한국법제연구원(역).
- Antje von Ungern-Sternberg(2019), 민주적 의사형성과 인공지능, in :*Demokratie und künstliche Intelligenz*, 한국법제연구원(역).
- Berz, Amelie&Engel, Andreas&Hacker, Philipp(2023), Generative KI, Datenschutz, Hassrede und Desinformation- zur Regulierung von KI-Meinungen, *ZUM* 2023, 586.
- Eckel/Rottmeier(2021), "Liken als Haten": Strafverfolgung von Hatespeech in Sozialen Netzwerken, *NStZ* 2021, 1.
- Eric Goldman(2006), Search Engine Bias and the Demise of Search Engine Utopianism, *Yale J. L. & Tech.*
- Gerdorf(2017), Hate Speech in sozialen Netzwerken, *MMR*, 2017, 439.
- Holznagel(2018), Phänomen "Fake News"- Was ist zu tun?, *MMR* 2018, 18.
- Kastl-Riemann(2023), Graziana: Algorithmen und Künstliche Intelligenz im Äußerungsrecht *ZUM* 2023, 578.
- Kumkar&Rapp(2022), Deepfakes, *ZfDR* 2022, 199.

- Kühling, Jürgen(2021): »Fake News« und »Hate Speech« - Die Verantwortung der Medienintermediäre zwischen neuen NetzDG, MStV und Digital Services Act, *ZUM*, 2021, 461.
- Mondal&Silva&Benevenuto(2017), A Measurement Study of Hate Speech in Social Media, *HT '17: Proceedings of the 28th ACM Conference on Hypertext and Social Media*, 85-94.
- Nolte(2017), Hate-Speech, Fake-News, das Netzwerkdurchsetzungsgesetz und vielfaltsicherung durch Suchmaschinen, *ZUM* 2017, 552.
- Ostendorf&Frahm&Doege(2012), Internetaufrufe zur Lynchjustiz und organisiertes Mobbing *NStZ* 2012, 529.
- Peukert(2018) Gewährleistung der Meinungs- und Informationsfreiheit in sozialen Netzwerke, *MMR* 2018, 572.
- Wagner(2022), Kommunikation in sozialen Medien und Grundrechte - Zur Drittwirkung der Meinungsfreiheit und Ausgestaltung einer dienenden Medienfreiheit, *Zum* 2022, 861.

■ ABSTRACT

Justification for intermediary regulation and a review of European legislation

Jung, Aeryung
Research Professor(NRF), Institute for Legal Studies,
Sookmyung Women's University

This study explores the emerging issues of legal and moral infringement arising from the pivotal role of media platforms in social communication, particularly with the spread of artificial intelligence technology and the advancement of the information society, and particularly highlights the need for corresponding regulations. As the use of artificial intelligence increases, content such as deepfakes, fake news, and hate speech is spreading rapidly among the public, raising concerns that it will distort public opinion and hinder democratic decision-making.

Regulating digital platforms, which play a dominant role in communication and have explosive ripple effects and influence due to their structural advantages, is the most effective way to curb the spread of illegal content. Therefore, this study examines the legitimacy of platform regulation as a response to these new threats and evaluates the current state of regulatory framework development in Europe.

In particular, we introduce Germany's Network Enforcement Act (NetzDG), France's Countering Disinformation Act, and the EU Digital Services Act, which are representative examples of intermediary

regulation in Europe, and examine the scope of legal liability, as well as key regulatory details required of platform operators under European law. In addition, Germany's Network Enforcement Act, often regarded as the beginning of intermediary regulation, was partially repealed with the enactment of the EU Digital Services Act, and the enactment of the German Digital Services Act has set the standard for intermediary regulation across Europe. However, these regulations may restrict the freedom of expression preemptively, raising concerns about private censorship and overblocking. Therefore, while recognizing the effectiveness and necessity of intermediary regulation, it is essential to take measures to minimize the risk of infringing on freedom of expression by expanding the regulatory body's scope for addressing illegal content, and supplementing procedural aspects by listening to public opinion.

Keywords: Mediator Regulation, Online Platforms, Illegal Content, Artificial Intelligence, EU Digital Service Act

[논문투고일 2024. 10. 28. 논문수정일 2024. 11. 25. 게재확정일 2024. 12. 5.]