

카카오 증오발언 대응 정책 녹서

원칙 수립을 위한 숙의 과정의 기록

2021.1.

목차

I	들어가며: 카카오가 증오발언 대응에 나선 이유	1
II	국내외 증오발언 대응 논의의 현주소	3
III	숙의 과정 1: 미디어자문위원회를 통한 정책 수립 과정 설계	8
IV	숙의 과정 2: 증오발언에 관한 민관학 공동 연구	16
V	숙의 과정 3: 온라인 증오발언 실태 파악을 위한 카카오 데이터 분석	24
VI	숙의 과정 4: 시민사회 의견 청취를 위한 라운드테이블	30
VII	나가며: 카카오의 증오발언 대응 정책 방향	35
부록1	숙의 과정 회의록	41
부록2	공동연구 결과	63

I. 들어가며: 카카오가 증오발언 대응에 나선 이유

카카오는 ‘기술과 사람으로 더 나은 세상을 만들겠다’는 비전으로 ‘기업의 디지털 책임(CDR, Corporate Digital Responsibility)’ 실천 방안을 고민하고 있다. 뉴스와 검색어 등 사회적 영향을 끼치는 서비스를 제공하는 사업자로서 역할과 책임을 고민해 왔으며, 그 과정에서 연예, 스포츠 뉴스에 대한 댓글과 인물 관련 검색어 서비스를 폐지하고, 실시간 이슈 검색어 서비스를 종료하기도 했다. 연예 뉴스의 댓글이 사생활 침해와 인격 모독 등 인권 침해적 표현의 진원지 역할을 해왔다는 문제의식에서 비롯된 것이었다.

당시 카카오는 댓글 서비스 정책을 보완하면서 댓글 신고 항목에 ‘차별·혐오’를 새로 만들었다. 또한 외부 전문가 및 이용자와의 소통을 통해 혐오·폭력성 콘텐츠 관련 자율규제 기준과 이행 방안에 대한 논의에 나설 계획을 공개했다. 단순히 서비스 개편에 그치는 것이 아니라 온라인 플랫폼이 증오(憎惡)발언에 대응할 수 있는 진정성 있는 고민이 필요하다고 판단했기 때문이다.

카카오는 2019년 11월 독일 베를린에서 개최한 ‘인터넷 거버넌스 포럼(IGF 2019)’에 참여해서, 구글, 페이스북, 마이크로소프트 등 다양한 기업들과 더불어 테러 및 폭력 콘텐츠의 자정 방안에 대한 의견을 나누기도 했다.

국내 인터넷 업계가 자율 규제 강화에 나서는 것에 대해 ‘공론장 축소’를 우려하는 목소리도 있다. 자율 규제가 언제나 ‘옳은 해답’을 내놓는 것은 아니기 때문이다. 그럼에도 불구하고 전 세계적으로 온라인 증오발언을 그저 두고만 볼 것은 아니라는 공감대가 형성됐다는 점을 주목할 필요가 있다. 증오발언으로 표출되는 한 사회의 차별과 폭력을 외면하는 것이 능사가 아니라, 증오발언이 어떤 방식으로 생산되고 있으며, 어떻게 문제를 해소할 수 있을지 진지하게 고민할 시점이 됐다.

카카오는 증오발언 정책 속의 과정을 통해 우리 사회 안의 증오발언의 형태와 성격을 고찰하고, 플랫폼 기업으로서의 역할을 살펴보았다. 대응 방안을 모색하기 위해 1년여에 걸쳐 다양한 이해관계자들과 협업을 진행했다. 속의 과정을 거쳐 도출한 결과가 <증오발언 근절을 위한 카카오의 원칙>이다.

본 ‘녹서(green paper)’는 그 속의 과정 및 관련 연구를 우리 사회에 공유하기 위해 만들었다. 카카오가 한 1년여의 고민이 향후 진행될 증오발언 근절을 위한 고민과 노력에 실효적인 참고자료가 되길 바란다. 결과물 자체를 넘어, 다양한

증오발언에 대한 카카오의 문제의식과 논의의 시작 배경

“전 세계적으로 증오발언을
그저 두고만 볼 것은 아니라는
공감대 형성”

이해관계자들의 의견을 모으는 과정을 공유함으로써, 증오발언 문제에 대한 우리 사회의 인식을 높이고, 대응 필요성에 대한 공감대를 넓힐 수 있다고 보았다. 이를 통해 향후 증오발언, 나아가 차별의 문제에 대한 우리 사회에서의 논의가 한층 성숙해지길 기대한다.

‘녹서’는 본격적으로 논의를 시작하기에 앞서 논의의 대상이 되는 문제를 다각도로 분석하고 대안을 마련해 참여자들에게 제공하는 자료를 의미한다. 원칙 수립 과정을 담은 보고서를 ‘녹서’로 칭한 것은 이번에 발표한 원칙을 시발점으로 삼아 향후 본격적인 논의를 진행시켜 나가겠다는 의미도 담고 있다.

용어의 사용에 관해

Hate speech는 국내에서 ‘혐오(嫌惡)표현’이란 말로 통용된다. 그러나 이 녹서를 포함해 향후 카카오의 정책과 관련된 자료에는 ‘증오발언’이란 용어를 쓰고자 한다.

‘혐오’ 대신 ‘증오’를 쓴 이유는 이렇다. ‘혐오’는 ‘싫어하여, 기피하는’ 감정까지 포괄한다. 이러한 감정에는 특정 대상에 대한 ‘고의적 폭력성’이 내포하지는 않을 수 있다. 특정 행위를 관리·조치하려면 그 행위의 해악성이 뚜렷해야 한다. 제재의 필요성과 타당성도 확보해야 한다. 이러한 조건을 고려해 ‘차별적 인식에 기반해 특정 집단이나 개인에 대해 극단적인 언어적 폭력을 가하는 행위’를 ‘증오발언’으로 정의했다.

‘Hate speech’를 우리말로 어떻게 엄밀히 옮길 것인가에 대한 학술적 논의는 아직도 진행 중이다. 이러한 가운데 ‘혐오’ 대신 ‘증오’를 정책적 논의 대상으로 삼은 것은 특정 집단에 가해지는 폭력의 문제를 해결하는 것이 보다 시급하다고 판단했기 때문이다. 용어와 관련해 불필요한 오해가 없길 당부드린다.

※일러두기: 녹서의 본문에는 ‘혐오표현’, ‘증오발언’ 등 용어들이 혼재한다. 이는 속의 과정에서 활용된 자료나 발언 등의 원문을 가능한 그대로 공유하기 위해 수정하지 않았기 때문이다.

II. 국내외 증오발언 대응 논의의 현주소

1. 증오발언의 국내외 양상¹⁾

증오발언 현상의 실태와 국제기구 단위의 논의들

증오발언은 전 세계적으로 하나된 정의를 내리기 보다는 각 나라가 가지고 있는 문화, 언어, 역사 등 사회적인 맥락에 따라 다르게 표현해 왔다. 예를 들어, 미국은 건국 이후 오랜 기간 노예제도를 유지해 왔으며, 노예제도를 폐지한 이후에도 뿌리깊게 흑인을 차별해 온 역사가 있다. 그 배경에 따라 미국 사회에서는 인종차별에 대해 다른 나라보다 민감하다.

반면 유럽권에서는 세계대전 과정에서 발현된 소수 민족에 대한 증오발언에 보다 주의를 기울여 왔다. 나치가 제2차 세계대전 중 유태인, 슬라브족, 집시, 동성애자, 장애인, 정치범들을 학살한 과거가 있기 때문이다. 그럼에도 불구하고 최근 유럽 대륙에 난민을 향한 차별과 증오가 고개를 들고 있는 모양새다. ‘테러와의 전쟁’이 중동 난민 문제와 얽혀 문명에 대한 증오를 야기하고 있다.

한국도 예외는 아니다. 신종 코로나바이러스(코로나19)가 확산하기 이전부터 지역 감정을 기반으로 한 뿌리 깊은 증오를 넘어 계층화한 증오가 일상을 파고 든지 오래다. 초등학교생들이 ‘월거지·전거지·빌거·엘사’ 등 일상화한 표현을 쏟아낸다²⁾. 여성·인종을 향한 증오는 더 심각하다. 2016년 5월 발생한 ‘강남역 화장실 살인사건’ 이후, 극단적인 여성 증오발언이 사회문제로 떠오른 바 있다. 미국의 영화제작자 하비 와인스타인(Harvey Weinstein)으로 촉발된 성폭력 고발 운동인 ‘미투(Metoo)’ 논쟁이 확산하면서 국내의 ‘여성혐오’에 대한 갈등도 깊어졌다.

증오의 민낯은 난민 이슈에서도 목격됐다. 2018년 6월 예멘 출신 난민 550여 명이 제주도에 입국해 난민 신청을 하자, 무슬림에 대한 맹목적인 증오가 분출됐다³⁾. 논쟁 과정에서 난민법 제도의 악용 등 합리적인 우려도 있었으나, ‘무슬림=테러리스트’와 같은 맹목적 비난도 가세했다. 한국 사회의 ‘제노포비아(외국인 혐오)’를 노골적으로 목격할 수 있던 사건이었다.

최근 한국에서는 현실의 일상을 뒤흔든 코로나19가 막연한 증오와 합쳐져, 온라인에서 폭력의 소재로 둔갑했다. 상황의 심각성을 인지한 방송통신심의위원회는 “코로나는 열에 약하다. 다시 한번 지하철에 불을 질러 ○쌍디안(경상도 사람 대상의 증오발언)을 통구이로 만들어서 코로나를 잡아야 한다”, “이참에

“최근 한국에서는
일상을 뒤흔든 코로나19가
막연한 증오와 합쳐져,
온라인 폭력의 소재로 둔갑”

광주에서 함 퍼지고 봉쇄시킨 후에 땅크로 흥○ 새끼들(전라도 사람 대상의 증오발언) 싹다 밀어죽여야 된다” 등과 같은 실제 예시를 언급하며, “특정 대상을 차별·비하하거나 조롱하고 편견을 조장하는 혐오표현 6건에 대해 심의를 통해 삭제 의결하고 조치했다”고 지난 2020년 3월 발표했다⁴.

2. 증오발언에 대한 국제적 원칙들

증오발언에 대해 국제적으로 통일된 정의는 아직 부재하다. 하지만 국제연합(United Nations; UN) 등 국제기구 및 비정부 기구(Non Governmental Organization; NGO)들은 차별과 증오에 반대하는 선언을 발표하고, 회원국들이 그 선언을 이행할 수 있도록 국제조약을 만들어 왔다. 증오발언의 개념에 대한 실마리를 얻기 위해 이러한 선언과 조약의 내용을 살펴봤다.

아래에서 확인할 수 있는 바와 같이 여러 국제기구들은 “만인이 평등하다”는 평등 가치 아래, 사람들은 인종, 성별, 성적지향, 피부색, 종교, 언어, 정치적 및 기타 의견, 민족적 또는 사회적 출신, 재산, 출생, 기타의 지위 등으로 차별을 받지 않아야 하며, 모두가 동등한 대우를 받아야 한다고 밝히고 있다. 또한 국제적 조약들은 타인의 특성이나 소속된 집단을 이유로 증오를 고취하는 행위까지도 범죄에 준하는 것으로 보고 각국의 적극적 대응을 촉구하고 있다.

세계인권선언

국제연합(UN)은 제2차 세계대전 이후 전 세계에 만연한 인권침해에 대해 인류의 반성을 촉구하고, 모든 인간의 기본적 권리를 존중해야 한다는 취지를 담아 1948년 <세계인권선언(Universal Declaration on Human Rights)>을 발표했다.

세계인권선언은 “모든 사람은 태어날 때부터 자유롭고, 존엄하며, 평등하다(제1조)”라고 선언하며⁵, “모든 사람은 인종, 피부색, 성, 언어, 종교 등 어떤 이유로도 차별을 받지 않으며, 이 선언에 나와 있는 모든 권리와 자유를 누릴 자격이 있다(제2조)”고 규정했다. 모든 사람이 평등하다는 것을 천명한 뒤, 각 개인들이 그 평등함을 누릴 수 있도록 “모든 사람들은 법 앞에 평등하며, 차별 없이 법의 보호를 받을 수 있다(제7조)”고 밝히고 있다.

세계인권선언은 이후 주요 국제 인권조약들의 토대가 됐다.

세계인권선언을 발표한 이후, 국제연합은 이 선언이 실질적으로 법적인 효력을 가질 수 있도록 하는 국제인권법을 제정하려 했다. 초기 구상은 시민적·정치적 권리와 경제적·사회적·문화적 권리를 하나의 조약에 담는 것이었지만, 시민적·정치적 권리와 경제적·사회적·문화적 권리는 이행 방식이 다르다는 인식이 커져 〈시민적 및 정치적 권리에 관한 국제규약(International Covenant on Civil and Political Rights)〉과 〈경제적, 사회적 및 문화적 권리에 관한 국제규약(International Covenant on Economic, Social and Cultural Rights)〉 두 개의 규약으로 만들었다⁶.

우리나라는 1990년에 두 조약을 동시에 비준했다⁷.

“증오발언에 대해 국제적으로 통일된 정의는 아직 부재”

“UN 등 국제기구 및 비정부 기구들은 차별과 증오에 반대하는 선언을 발표하고, 회원국들이 그 선언을 이행할 수 있도록 국제조약을 만들어”

시민적·정치적 권리에 관한 국제 규약

1966년, 국제연합은 〈시민적·정치적 권리에 관한 국제규약〉을 조약의 형태로 완성했다. 이 규약은 증오발언에 대해 구체적으로 언급하고 있다⁸. 협약의 제2조 제1항에서는 “모든 개인에 대하여 인종, 피부색, 성, 언어, 종교, 정치적 또는 기타의 의견, 민족적 또는 사회적 출신, 재산, 출생 또는 기타의 신분 등에 의한 어떠한 종류의 차별도 없이 이 규약에서 인정되는 권리들을 존중하고 확보할 것을 약속한다.”고 규정한다⁹. 증오발언과 관련해서는 제20조 2항에서 “차별, 적의 또는 폭력의 선동이 될 민족적, 인종적 또는 종교적 증오의 고취는 법률에 의하여 금지된다”고 규정하고 있다.

또한 제26조에서는 “모든 사람은 법 앞에 평등하고 어떠한 차별도 없이 법의 평등한 보호를 받을 권리를 가진다. 이를 위해 법률은 모든 차별을 금지하고, 인종, 피부색, 성, 언어, 종교, 정치적 또는 기타의 의견, 민족적 또는 사회적 출신, 재산, 출생 또는 기타의 신분 등의 어떠한 이유에 의한 차별에 대하여도 평등하고 효과적인 보호를 모든 사람에게 보장한다.”고 밝히고 있다. 조약에 비준한 각국이 차별과 증오에 적극적으로 대응할 것을 요구하는 것이다.

경제적·사회적·문화적 권리에 관한 국제규약

〈경제적·사회적·문화적 권리에 관한 국제규약〉은 인간의 존엄성 및 합당한 삶을 누리기 위한 기본권들을 다루고 있다¹⁰. “모든 사람은 인종, 피부색, 성, 언어, 종교, 정치적 또는 기타의 의견, 민족적 또는 사회적 출신, 재산, 출생, 기타의

“모든 형태의 인종차별 철폐에 관한 국제협약”은 차별·폭력을 선동하는 행위를 법률로 금지해야 한다고 규정”

지위 등에 따른 어떠한 종류의 차별도 없이 이 선언에 제시된 모든 권리와 자유를 누릴 자격이 있다(제2조 제2항)”고 정의해, 모든 사람이 권리와 자유에 동등하며 차별 없는 대우를 받아야 한다는 것을 다시 한번 강조했다.

모든 형태의 인종차별 철폐에 대한 국제 협약

1963년, 국제연합은 국제사회에서 인종차별에 대한 우려가 커짐에 따라, 〈모든 형태의 인종차별 철폐에 관한 선언〉을 공식적으로 채택했다. 2년 뒤 국제연합 총회는 〈모든 형태의 인종차별 철폐에 관한 국제협약(International Convention on the Elimination of All Forms of Racial Discrimination)〉을 만들었다¹¹. 1969년부터 국제연합 회원국들이 해당 조약에 비준 및 가입하며 인종차별 퇴치에 대한 뜻을 이어 나갈 수 있는 발판이 마련되었다.

이 협약은 제4조a항을 통해 인종차별에 기반한 폭력행위와 이를 선동하는 행위뿐만 아니라 모든 형태의 (재정적) 지원까지도 의법처벌(법에 의하여 처벌)의 대상으로 선언했다¹². 즉 차별과 폭력을 선동하는 증오 행위를 법률로 금지해야 한다고 규정한 것이다.

컴퓨터를 통한 인종주의·외국인혐오적 행위의 처벌에 관한 사이버 범죄 협약 추가의정서

유럽의회는 2003년, 〈사이버 범죄 협약(Council of Europe Convention on Cybercrime)〉의 추가의정서로 〈컴퓨터를 통한 인종주의, 외국인혐오적 행위의 처벌에 관한 사이버 범죄 협약 추가 의정서(Additional Protocol to the Convention on Cybercrime, concerning the criminalisation of acts of a racist and xenophobic nature committed through computer systems)〉를 채택했다. 이는 컴퓨터시스템을 통해 형성된 인종차별적 또는 외국인혐오적 성격을 지니는 행위들을 범죄로 규정하는 내용을 담고 있다¹³. 여기서 인종차별적이고 외국인혐오적인 자료란, “종교 뿐만 아니라 인종, 피부색, 혈통, 민족 등 출신에 기초하여 어떤 개인이나 개인의 집단에 대한 혐오, 차별 또는 폭력을 옹호·증진·조장하는 서면 자료, 이미지, 또는 사상이나 이론을 대표하는 것”을 말한다¹⁴. 또한 유럽에서 자행된 홀로코스트에 대해서 부인하는 것도 금지하고 있다.

참고문헌

1. 〈KISO저널〉 제38호 “‘혐오 바이러스’ 확산, 그리고 기업의 책임”을 일부 수정·보완함
2. 머니투데이 (2019.11.17.). “이백충·월거지”.....초등학교 교실에 퍼진 ‘혐오’.
3. 한겨레 (2018.6.18.). 제주도엔 온 예멘 난민 500명, 무슬림 혐오에 내몰리다.
4. 연합뉴스 (2020.3.2.). 방심위, '코로나19' 관련 특정지역 혐오 표현 삭제.
5. UN. Universal Declaration of Human Rights.
6. 박진아 (2016). 「시민적 및 정치적 권리에 관한 국제규약」 채택 50년 - 규약의 발전과 이행. *홍익법학*, 17(4), 211-236.
7. 이주영 (2016). 사회권규약의 발전과 국내적 함의. *국제법학회논총*, 61(2), 125-157.
8. 박지원 (2015). 혐오표현(Hate Speech)’의 규제에 관한 외국 입법례. *국회 입법현안 법률정보 제41호*.
9. 국가법령정보센터. 시민적 및 정치적 권리에 관한 국제규약(B규약).
10. 이주영 (2016). 위와 같음.
11. 국가인권위원회 국제협력담당관실 (2005). 유엔인권해설집(Human Rights Fact Sheet 12권) 「인종차별위원회」 .
12. 국가법령정보센터. 모든 형태의 인종차별 철폐에 관한 국제협약.
13. 유럽의회 (2003). Additional Protocol to the Convention on Cybercrime, concerning the criminalisation of acts of a racist and xenophobic nature committed through computer systems.
14. 이정념 (2018). 혐오발언의 규율을 위한 국제 기준들. *예술인문사회 융합 멀티미디어 논문지*, 8(6), 555-562.

III. 속의 과정 1: 미디어자문위원회를 통한 정책 수립 과정 설계

증오발언 원칙 수립 위해 1년 동안 진행한 로드맵

카카오는 증오발언에 대한 원칙 수립을 위해 거친 과정은 아래 그림의 로드맵으로 설명된다.

카카오는 2020년 1월부터 증오발언 대응에 대한 내부 논의를 시작했다. 검토 결과, 증오발언 문제에 효과적으로 대처하기 위해서는 외부의 다양한 의견을 수렴해야 할 필요성을 파악했다.

이에 따라 카카오는 3월부터 증오발언 대응 원칙 수립을 위한 전반적인 속의 과정에 미디어자문위(이하 자문위)의 참여를 요청했다. 미디어 전문가로 구성된 자문위는 이후 카카오의 증오발언 대응 논의 전 과정에 함께 했다. 자문위는 정기 회의를 통해 속의 진행 방식, 정책 방향 설정, 내부 연구 결과 등을 함께 논의했을 뿐 아니라, 10월 시민사회 전문가 라운드테이블(원탁회의)에 참여하고, 이후 <증오발언 근절을 위한 카카오의 원칙>의 초안을 검수하기도 했다.

한편 카카오는 증오발언에 대한 정의와 원칙의 방향성 수립에 참고하기 위한 학술 연구를 ‘국가인권위원회(이하 인권위)’, ‘한국언론법학회’와 공동으로 진행했다. 연구단은 6월 구성을 마쳤고, 연구진 공모 및 확정은 6~8월에 이뤄졌다. 중간 연구 결과는 9월 22일에 세미나 형식으로 외부에 공개했다.



이번 장에서는 미디어자문위를 통한 숙의 과정을 설명하고자 한다.
자문위와의 논의에 앞서, 카카오는 2020년 3월 다음과 같은 메일을 통해 증오
발언 원칙 수립의 계획을 전달했다.

미디어자문위원회와 증오발언
대응 원칙 수립 놓고
2020년에 세 차례 회의

“미디어자문위 2020 준비를 위한 사전 요청 건”

- 2020년 3월 20일에 자문위원들에게 보낸 이메일

아직까지 우리나라에서는 해외에서와 같이 혐오 및 극단적 폭력 콘텐츠에 대해 치열한 논쟁이 벌어지고 있지는 않습니다. 특정 집단에 대한 혐오범죄가 심각하지 않다는 면에서 한편으로는 다행이지만, 우리나라에도 분명 다양한 집단 간 갈등이 계속되어 왔다는 점을 간과할 수는 없습니다. 무엇보다 앞으로 더욱 차별과 폭력이 없는 사회로 나아가야 한다는 점을 생각할 때, 우리 사회도 치열한 논의를 통해 늦지 않게 온라인상에서의 혐오 및 극단적 폭력 콘텐츠의 제재에 대한 기틀을 다져야 할 것으로 보입니다. 향후, 미디어자문위를 통해 나누게 될 논의들은 관련 기틀을 다지는 데 기여할 수 있을 것으로 기대합니다.

차기 회의에서는 자문위 틀과 운영 방식(인원 구성 포함)에 대해 구체적으로 말씀 나누는 시간으로 구성하고자 합니다.

자문회의는 2020년 4월, 6월, 8월, 총 세 차례에 걸쳐 이루어졌다.

회의의 목적은 1)국내 언론 및 미디어 학계의 전문가들로부터 증오발언 정책 마련의 필요성과 방향에 대한 자문을 구하고, 2)증오발언 정책 마련을 위한 과정에서 다양한 이해관계자들의 의견을 반영하기 위한 공론 수렴 과정을 설계하기 위한 것이었다.

논의에 참여한 카카오 미디어자문위원은 다음과 같다. (위원장 외 가나다 순)

- | 박재영 고려대 미디어학부 교수 (위원장)
- | 김민정 한국외국어대 미디어커뮤니케이션학부 교수
- | 김장현 성균관대 인터랙션사이언스학과 교수
- | 도준호 숙명여대 미디어학부 교수
- | 이준웅 서울대 언론정보학과 교수

첫 번째 회의에서는
증오발언 정책에 대해
고민하게 된 배경을 설명하고
자문위원들의
정책 수립 동참을 요청

제1차 미디어자문위 회의

증오발언 대응 정책 마련을 위한 첫 번째 자문위 회의에서 카카오는 자문위원들에게 증오발언 정책에 대해 고민하게 된 배경을 설명하고, 향후 정책 수립 과정에 동참하기를 요청했다. 해외에 견주어 볼 때 국내에서의 증오발언 논의는 다소 늦은 편이며, 카카오 내부만의 고민으로 해결하기에 어려운 문제이기 때문이다. 학계와 업계 뿐 아니라 사회적으로 논의를 확대하고, 논의의 결과물을 다양한 사회구성원들과 함께 나누고자 하는 취지이다. 자문위가 그 방식에 대해 함께 고민해 줄 수 있을 것으로 기대했다.

카카오의 계획과 증오발언 논의 현황에 대해 자문위원들은 다양한 의견을 제시했다. 아래는 발언자 및 발언 순서에 관계 없이 주요 의견을 주제별로 묶은 것으로서, 자세한 내용은 [부록 1\)의 회의록](#)에서 확인할 수 있다.

증오발언 대응 정책 수립의 필요성

□ 증오발언 논의 선도 역할 필요성

- 국내 개별 사업자들의 이용자 보호 정책이나, 한국인터넷자율정책기구(Korea Internet Self-Governance Organization; KISO)의 차별적 표현 완화 정책 등 업계의 적용 사례는 7개도 되지 않음
 - 즉, 국내 업계의 관심은 미미한 상황임. 반면 구글이나 페이스북은 커뮤니티 가이드라인에 독립적인 항목으로 설정해 자세히 설명하고 있음.
 - 카카오가 이런 작업을 선도적으로 해 나가면 좋을 것
- 소수자는 숫자의 문제가 아님. 즉 다수에 의한 소수의 차별이라는 개념이 성립하지 않음. 본인이 선택하지 않은 것으로 차별을 받을 때 문제임
- 혐오표현의 폐해는 소수자들의 사회참여를 막아버리는 데 있음
- 극단적인 표현이 여론을 형성하는 데 큰 영향을 끼침
- 카카오가 민주주의의 건전한 작동을 돕는다는 생각으로 논의를 이끌어 가길 바람

□ 플랫폼 사업자의 증오발언 대응 가능성

- 특정 유형의 발언에 대해 사업자가 제한하는 것은, 이용자와 사업자 간 계약의 문제이므로 가능함
- 사적 영역에서 발언 수준을 정하는 것은 사적 영역의 주체들이 발언의 자유

- 를 통제하지 않는 방식으로 할 수 있음
- 법으로 규제(모욕죄 등의 적용)하는 접근에 반대
- 플랫폼 사업자의 증오발언 규제 시 부작용 우려
 - 다양한 형태의 혐오표현들이 서로 결합하면서 새로운 표현이 만들어지고 있음
 - 혐오표현의 기준을 잡기 어려우며 가변성이 큼
 - 혐오표현 사용을 제재하면 사람들이 사회적으로 반성할 기회를 박탈당할 수 있음
 - 혐오표현 이슈를 포털 사업자가 선도할 필요는 없음
 - 혐오표현을 제재하기 시작하면 스스로 오류에 빠질 수 있음
 - 혐오표현을 선택에 따라 가리는 기능 제공 정도가 최선일 것
 - 이용자 규제가 차별을 줄이고 민주주의에 도움이 될지는 따져볼 문제임

“정책 수립 시 시민단체 등 참여를 보장하고 넓은 범위에서 이용자의 대표성 확보가 필요”

증오발언 대응 정책 수립시 고려할 점

- 정책 수립 과정
 - 시민단체 등의 참여가 보장되어야 함. 넓은 범위에서 이용자 대표성 확보가 필요함
 - 사업자로서 이용자 보호를 위한 규범을 만들어 내려면 충분한 설득력 확보가 필요함
 - 카카오가 선도적으로 광범위한 토론의 장을 만들어 내는 것이 중요함
 - 혐오표현의 생산과 소비에 대한 실질적 연구가 필요함
 - 서비스 정책이 이용자에게 미치는 영향, 비즈니스에 미치는 영향 등 반복적 실험이 필요함
- 정책 이행 방식
 - 단어 삭제 수준의 규제보다는 적극적으로 혐오표현에 반대하는 입장 표명이 중요함
 - 콘텐츠 삭제가 아닌, 캠페인, 대항표현 단체 지원, 교육 투자, 이용자 선택 증대 등의 방식도 가능함
 - 공간에 따라 혐오표현 양상 차이가 크다는 점을 인식해야 함
 - 카카오 서비스 외에 폐쇄적 커뮤니티의 혐오표현도 심각하므로 사회적 관심을 환기하는 역할이 중요함

두 번째 회의에서는
민관학 공동연구 추진,
라운드테이블 개최,
녹서 발간 계획을 공유

- 특정 단어가 혐오표현인지를 고민하는 것이 아닌, 보다 스마트한 접근 필요
 - 혐오표현 검사 프로그램을 개발해 댓글에 적용할 필요가 있음
 - 선한 표현을 더 많이 노출시키는 것도 방법임
- 용어 정의의 필요성
- 혐오표현이라는 용어 자체가 오해를 불러올 수 있음. 차별과 혐오를 구별해야 할 필요 있음
 - “차별에 반대하는 카카오” 등 중립적인 표현, 또는 extreme speech로 통일하는 안 등을 제안함

제2차 미디어자문위 회의

두 번째 자문위 회의에서 카카오는 증오발언 대응 정책 마련을 위한 실행 계획을 자문위원들에게 공유하고 의견을 구했다. 인권위 및 한국언론법학회와 함께하는 민관학 공동연구를 추진하고 있으며, 정책 또는 가이드라인 초안을 마련한 후에 전문가, 언론인, 시민단체 등으로부터 의견을 수렴하기 위한 라운드테이블을 개최하고, 고민의 과정 및 결과물을 사회와 공유하기 위한 녹서를 발간할 계획을 밝혔다.

이러한 논의 확장 계획 및 증오발언 대응 정책 방향에 대해 자문위원들은 다음과 같은 의견을 제시했다. 마찬가지로 아래는 발언자 및 발언 순서에 관계 없이 주요 의견을 주제별로 묶은 것으로서, 자세한 내용은 [\(부록 1\)의 회의록](#)에서 확인할 수 있다.

논의 확장 계획에 대한 의견

- 공동연구에 대한 의견
- 공동연구는 철학적, 법리적 논의보다는 서비스 고품질화를 위한 문제 해결 프레임워크(자율규제) 마련에 초점을 두어야 함
 - 공동연구는 학회 공모보다는 카카오에서 책임자를 찾아 발주하는 방식을 추천함
 - 공동연구 주제로 댓글 신고 조치 시 판단이 어려웠던 표현에 대해 일반인 및 전문가 대상의 설문을 해서 데이터를 축적하는 것도 가능함

□ 기타

- 일반 시민 대상의 공청회 등 많은 공청회를 갖기를 추천함
- 논쟁적인 사안인 만큼 결과물 자체보다 과정에 대한 설명과 논의 자체에 집중하며 진행해야 함
 - 녹서를 마지막에 발표하는 것은 유의미함

“(증오발언 대응은) 논쟁적인 사안인 만큼 결과물보다 과정에 대한 설명과 논의 자체에 집중해야”

정책 방향에 대한 의견

□ 증오발언으로부터의 보호 대상 및 개념 정의

- 인권위 가이드라인에 19가지 차별금지 대상이 나열되어 있음
- 인권위 혐오표현 리포트는 6가지를 언급하고 ‘등’을 붙임. 이에 준거 가능
- 일반 원칙 중 하나는 본인 의지로 선택하지 않은, 인종, 성별, 출신국가 등임
- 혐오가 너무 개괄적이라면, ‘적의적 표현’도 사용 가능한 개념임

□ 정책 수립시 부작용 우려

- 가이드라인 발표 이후 사회적 반발이 있을 수 있음을 감안해야 함
- 가이드라인 통해 카카오가 가치판단을 하게 되면 부담이 될 수 있음

제3차 미디어자문위 회의

세 번째 자문위 회의에서 카카오는 현재 운영하고 있는 욕설치환(置換) 데이터베이스(database; DB) 분석 등을 통한 내부 고민의 결과를 자문위에 공유하고 이에 대한 의견을 요청했다. 자문위에 공개한 자료는 위의 [〈속의 과정 3〉](#)에 해당하는 장에서 확인할 수 있다. 이 자료에는 증오발언 대응 정책 수립시 치열한 논의가 필요한 다음의 세 가지 쟁점에 대한 내부적 고민이 담겨 있다.

- 증오발언으로부터 보호해야 할 대상은 누구인가?
- 어떤 표현이 증오발언에 해당하며, 어떻게 분류할 수 있나?
- 각 대상 또는 표현의 종류에 따른 대응 방안은 무엇인가?

이에 대한 자문위원들의 주요 의견을 정리하면 다음과 같으며, 자세한 내용은 [〈부록 1〉의 회의록](#)에서 확인할 수 있다.

“사적 영역과 공적
영역에서의 발언 규제는
다르게 접근 필요”

증오발언으로부터의 보호 대상

□ 보호 대상을 정의하는 내부 기준 설정 필요성

- 보호 대상을 내부적으로 지정 필요
- 소수자의 특성, 개념을 구성하는 기존 이론을 참조하여 보호 대상을 정하는 기준을 밝히되, 구체적 집단은 밝히지 않을 수 있음
- 정책 목표를 명확히 해야함
 - 사회통합, 사회질서를 저해하는 정보를 규율하는 것이 목표라면, 대상에 관계 없이 극단적 공격 표현을 규제하면 됨
 - 단, 이는 극단적 공격 발언 규제로서 이미 하고 있는 것
 - 소수자 인권보호가 목표라면, 보호 대상을 면밀히 설정할 필요 있음

□ 보호 대상 정의의 유동성 필요

- 보호 대상을 사전 지정하는 것은 어렵지만, 보호 대상 특징이 안 된 상태에서 규제도 어려움
- 소수자 범위는 계속 변화하므로 유동적 기준 필요
 - 피해를 보는 집단은 모두 보호의 대상이 될 수 있음
- 보호 대상을 정하는 것은 사회적 합의 이루기 어려울 수 있음

□ 기타

- 현재 우리나라에서 가장 문제되고 있는 성별에 먼저 접근해보는 것도 방법

증오발언의 분류 및 대응 방안

□ 서비스 성격에 따른 대응 방안 구분 필요성

- 게시물 규제 정책을 서비스마다 분할할 필요 있음
- 사적 영역과 공적 영역에서의 발언 규제는 다르게 접근 필요
 - 사적 영역에서 구성원의 자발적·내부적 규칙에 의해 규제 이루어짐
 - 공적 영역에서는 공론장 역할 유지 필요. 카카오 서비스 중 뉴스 댓글 등은 대표적인 공적 공간
 - 학문적 담화에 근거한 보편적 규제 시행은 어려움
- 공적 공간이라도 품위(decency) 유지 차원에서는 상스러운 욕, 불법정보, 인권침해 행위 등은 현행법으로 이미 규제되고 있음
 - 증오·선동 발언에 대해서는 법이 없어 공적 공간에 적용 어려움

□ 게시물 규제 외의 대응 방안

- 추가적인 정보 제공 및 캠페인 등 게시물 삭제가 아닌 다른 방식으로 대항할 필요 있음
- 소통 역량을 증진시키는 방식으로 해결 필요. 사업자가 더 많은 토론을 만들어 내고, 스스로 교육할 수 있는 계기를 제공
- 서비스·기술적 방식으로 해결 필요. 토론할 사람들이 자발적으로 따로 모여서 토론할 수 있는 공간 제공하되, 미성년자 등은 못 들어오게 할 수 있음
- 사회적 약자에게 차등적으로 자원을 제공하는(발언의 기회를 보장하는) 정책도 도입 가능하고 필요함
- 극단적 발언, 선동은 즉각적 불법행위로 이어질 수 있으므로 발언 규제 필요
 - 그러나 정치적 발언을 규제해서는 안됨
 - 특정 집단에 대한 발언을 막는 것은 불가함

□ 기타

- 기존의 이론과 표현의 범주화 틀을 매칭할 필요 있음
 - 차별적 괴롭힘, 편견조장, 모욕, 증오·선동의 4단계 등
 - 제재 단계 구분을 위해서는 표현의 유해성에 방점 필요
 - 이론적으로는 제시한 표현의 구분 중 행동에 대한 규제는 가능성 있으나, 인식·의견에도 제재 필요한 경우 있음
 - 분석에 활용된 ‘이름짓기’를 규율하는 것으로도 문제의 상당부분 해소 가능할 것
 - 용어의 통일·정리가 필요. 정확한 용어는 ‘증오’이지만 ‘혐오’가 일반적으로 쓰이는 상황
- ※참고: 자문위에 공유한 원문에는 ‘차별·혐오·증오 표현’이라는 용어를 사용함. 뒤의 [\[속의 과정 3\]](#)에 수록한 자료는 사후 ‘증오발언’으로 수정한 것

“게시물 삭제가 아닌 다른 방식으로 대항할 필요”

“정확한 용어는 ‘증오’이지만 ‘혐오’가 일반적으로 쓰이는 상황”

IV. 속의 과정 2: 증오발언에 관한 민관학 공동 연구

‘사회의 눈높이’에서 온라인 증오발언의 문제점을 바라보고자 민관학 협력 공동연구 추진

첫 번째 미디어자문위 회의를 진행한 뒤, 카카오는 온라인 증오발언 대응을 위한 정책 수립이 플랫폼 기업이 홀로 짊어지기에 쉽지 않은 과제임을 실감했다. 증오발언이라는 사회적 난제를 기업이 자체적으로 정의하고 대응 방안을 도출하는 것 자체가 근본적인 대응과 멀어질 수 있기 때문이다. 미디어자문위에서 지적한 바와 같이, 실제로 카카오 내부의 정책, 대외, 이용자보호 담당 실무자로 구성된 태스크포스(TF)에서는 혐오·증오 등의 정의를 내리는 것에서부터 어려움을 겪었다. ‘사회의 눈높이’에서 온라인 증오발언의 문제점을 바라보고, 다양한 전문가와 현장의 목소리를 통해 변화를 이끌어 낼 필요성을 실감했다.

카카오는 보다 효과적이고 현실적인 온라인 증오발언 대응을 위해 ‘거버넌스(Governance)’를 통한 접근 방식을 선택했다. 국내에서 처음으로 온라인 증오발언 대응을 위해 민·관·학의 협력을 도모했다.

가장 먼저, 한국 사회의 혐오와 차별 문제에 집중해온 국가인권위원회 혐오차별 대응기획단(이하 기획단)에 도움을 요청했다. 기획단은 인권위가 2019년 1월 혐오와 차별 관련 현안에 적극적으로 대응하기 위해 구성한 특별 전담조직이다. 기획단은 성별이나 장애, 나이, 출신 국가·민족, 인종, 성적 지향 등을 이유로 하는 증오발언을 모니터링하고, 이런 표현을 예방하기 위한 교육과 홍보 업무를 맡고 있다. 2019년 10월 발간한 <혐오표현 리포트>를 통해 이용자의 자발적 협력과 온라인 플랫폼의 자율규제 마련을 강조하는 등 온라인 증오발언 문제를 꾸준히 고민해 온 곳이다.

관련 전문성을 지닌 학계의 도움도 절실했다. 카카오는 인터넷 서비스에서의 사회현상에 주목하며 전문성을 쌓아온 한국언론법학회에 증오발언에 대한 연구를 함께 해달라고 제안했다.

이와 같은 배경에서 출범한 민·관·학 공동연구단은 2020년 5월 첫 회의를 기점으로, 언론법학회 소속 학자 및 인권위 추천을 받은 전문가로 연구진을 구성하여 증오발언에 관한 연구를 시작했다. 카카오에서는 플랫폼 현황에 대한 기초 연구를 지원했다.

연구 결과물은 2020년 9월 22일 서울 중구 프레스센터에서 1차 중간발표 세미나를 통해 공개했다. 공동연구의 중간 성과를 발표하고 향후 연구과제와 대응 방안을 모색하는 자리였다.

구체적으로, 먼저 이승선 충남대 교수와 최진호 한양대 박사가 온라인 증오발언에 대해 일반 시민과 전문가들의 인식을 조사한 결과와 그 의미를 발표했다. 이어서 박아란 한국언론재단 책임연구위원이 증오발언과 표현의 자유의 한계를 비롯해, 역사적·사회적 맥락 등 온라인 증오발언 개념 정의에서 고려할 요소에 대해 발표했다. 마지막으로 이주영 서울대학교 인권센터 전문위원이 온라인 증오발언 규제를 강화하고 있는 해외 동향과 함께 증오발언에 관한 글로벌 소셜 미디어 기업의 역할과 과제에 대해 연구한 내용을 소개했다.

공동연구단은 이번 연구를 통해 앞으로 국내 증오발언 관련 논의가 보다 활성화 되기를 기대하고 있다. 온라인 증오발언 대응을 위한 실질적인 해법을 모색하는 데 유용한 참고자료로 활용되길 바란다.

연구의 최종 결과는 [\(부록 2\)](#)에 수록했다.

온라인 증오발언에 대해 일반 시민과 전문가들의 인식을 조사한 결과와 그 의미, 표현의 자유의 한계 및 역사적·사회적 맥락 등 온라인 증오발언 개념 정의에서 고려할 요소, 해외 동향 및 글로벌 소셜미디어 기업의 역할과 과제에 대한 연구 결과 발표

“혐오표현의 주요 접촉경로가 되는 온라인에서 혐오표현 대응을 위한 자율적 규제가 시도되고 있는 만큼, 자율규제를 통하여 온라인 혐오표현에 대응하는 유럽의 사례 등을 분석하는 이번 연구를 통해 혐오표현 근절을 위한 자율적 규제의 구체적 기준을 마련하고, 나아가 정보통신 기업이 인권존중 역할을 실현하는 계기가 되기를 기대한다”

최영애 국가인권위원회 위원장

“최근 디지털화가 가속화됨에 따라 기업이 정책과 제도를 통해 긍정적 미래를 창출하는 ‘기업의 디지털 책임(CDR, Corporate Digital Responsibility)’이 중요해지고 있다. 이번 공동 연구가 온라인 혐오표현 근절을 위한 선한 영향력이 되길 희망한다.”

여민수-조수용 카카오 공동대표

“이번 연구를 통해 혐오표현에 대한 사회적 논의를 풍부하게 하고 온라인에서 인격권과 표현의 자유를 조화롭게 균형 맞출 수 있는 방안을 고민하는 계기가 되었으면 한다.”

김종철 한국언론법학회 회장(연세대 교수)

“선동형 표현보다 모욕형
표현의 심각성을 더 크게 인식”

공동연구 중 일반 시민 인식 조사 결과

다음은 공동연구의 결과물 중 일반 시민들의 인식 조사 결과 일부를 발췌한 것이다. 연구 방법 및 결과에 대한 보다 자세한 정보는 [\(부록 2\)](#)를 통해 확인할 수 있다.

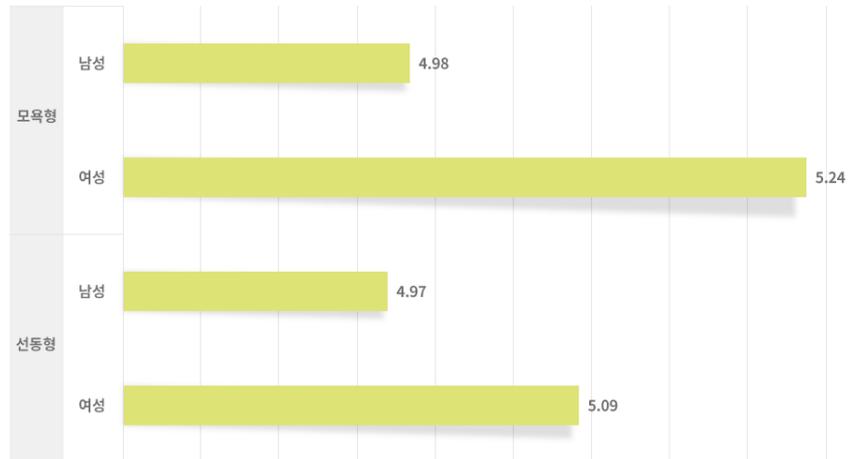
조사의 개요

- 조사 목적: 증오발언에 관한 일반 시민들의 인식 및 행태 조사
- 조사 방법: 온라인 설문조사 (조사기관: 마켓링크)
- 조사 기간: 2020.8.26.~31.
- 샘플: 1,022 명 - 성별, 연령, 거주지역을 고려한 비례층화표집 활용

주요 시사점

□ 혐오표현 유형별 인식

- 표현방식, 성별, 연령대, 교육수준에 따라 혐오표현에 대한 인식에 차이가 나타남
- 대체로 선동형 표현보다 모욕형 표현¹의 심각성을 더 크게 인식함
- 여성이 남성보다 혐오표현에 대한 민감성이 높게 나타남

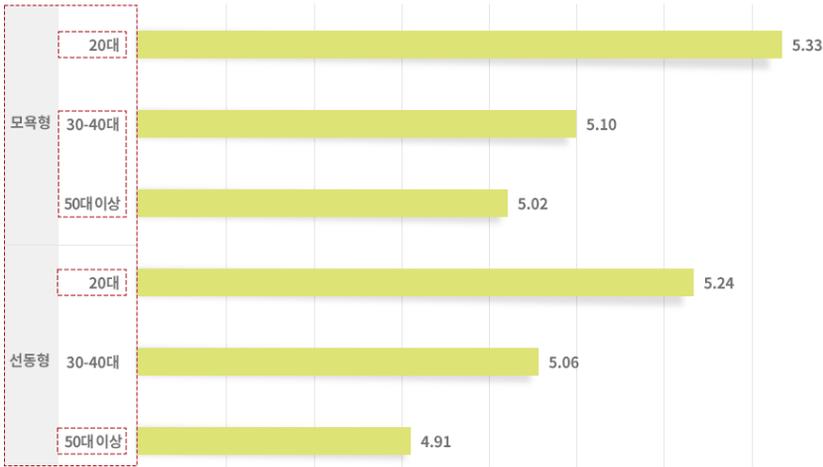


N=186(20대), 423(30-40대), 413(50-60대)

그림1. 표현유형별 혐오표현에 대한 판단 차이 (성별)

1. 모욕형 표현은 '모욕, 비하, 멸시, 위협하는 표현', 선동형 표현은 '차별·폭력을 선전하거나 선동하는 표현'을 의미

- 연령대가 낮을수록 혐오표현에 대한 민감성이 높게 나타남

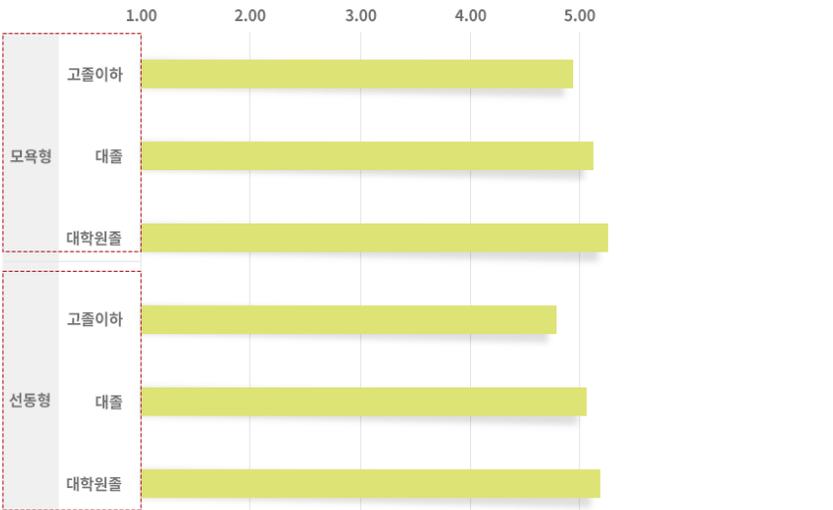


“여성이 남성보다, 연령대 낮을수록, 교육수준 높을수록 혐오표현에 대한 민감성이 높게 나타나”

N=186(20대), 423(30-40대), 413(50-60대)

그림2. 표현유형별 혐오표현에 대한 판단 차이 (연령대별)

- 교육수준이 높을수록 혐오표현에 대한 민감성이 높게 나타남



N=196(고졸이하), 710(대재-대졸), 116(대학원재-대학원졸)

그림3. 표현유형별 혐오표현에 대한 판단 차이 (교육수준별)

“심리적 거리감 클수록
혐오표현에 대한 민감도 낮아”

□ 혐오표현 민감도

- 혐오표현의 대상과 자신과의 심리적 거리감이 크다고 느낄수록 혐오표현에 대한 민감도가 낮게 나타남

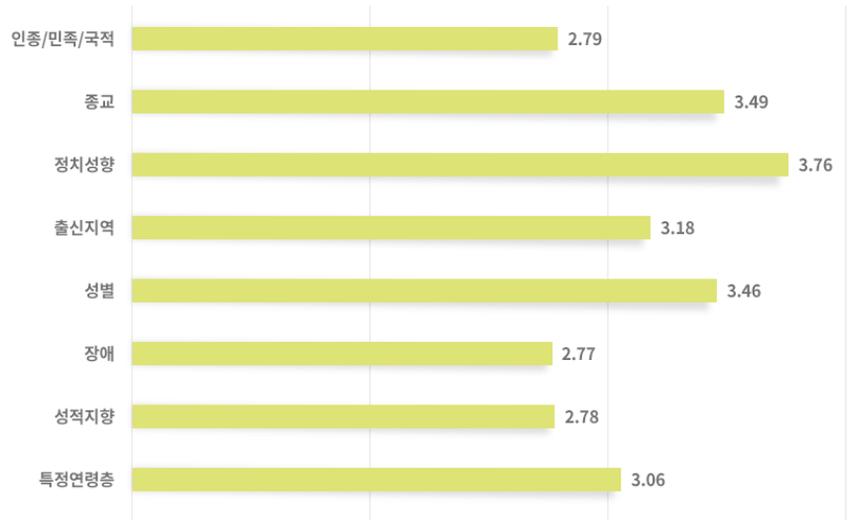


N=1,022, 5점 척도로 측정

그림4. 온라인 혐오표현에 대한 개념적 판단

□ 혐오표현 대상

- 종교, 정치성향, 출신지역, 성별, 특정연령층을 대상으로 한 온라인 혐오표현에 대한 노출이 상대적으로 많음



N=1,022, 5점 척도로 측정

그림5. 온라인 혐오표현 노출빈도

□ 혐오표현 노출

- 남성이 여성보다 온라인 혐오표현을 더 많이 접하는 것으로 인식함

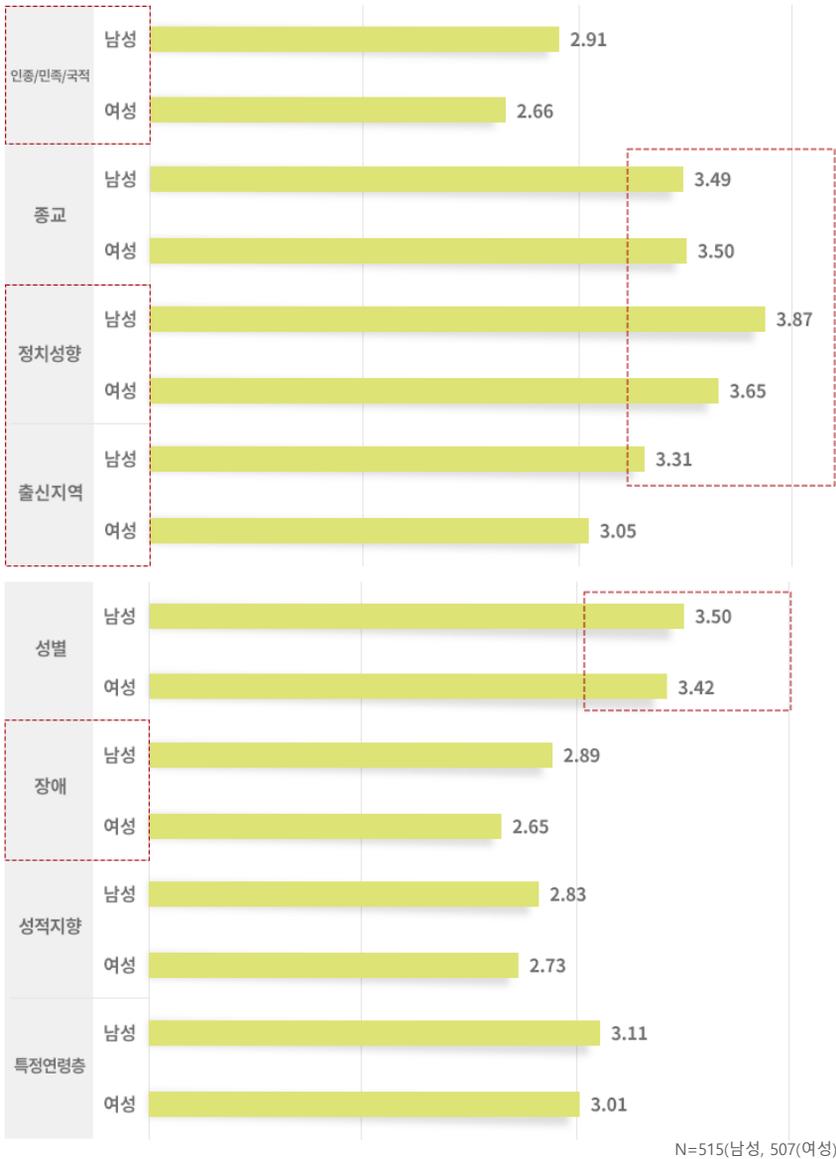
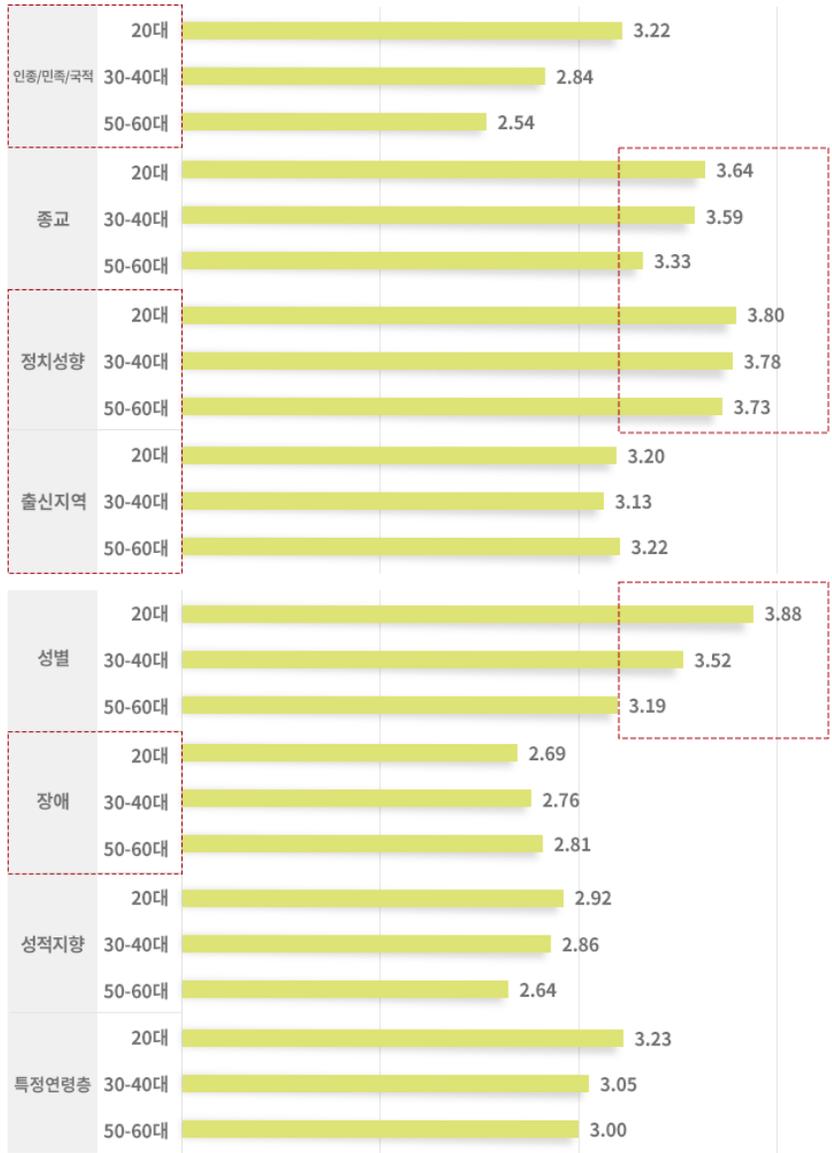


그림6. 피해대상 집단별 혐오표현 노출 정도의 차이 (성별)

- 연령대가 낮을수록 혐오표현을 더 많이 접하는 것으로 인식함



N=186(20대), 423(30-40대), 413(50-60대)

그림7. 피해대상 집단별 혐오표현 노출 정도의 차이 (연령대별)

□ 혐오표현 생산 주체

- 온라인 혐오표현 생산은 남성이 여성에 비해 상당히 많으며, 연령대별로는 고른 분포를 보임

“혐오표현 생산은 남성이 여성에 비해 많고, 연령대별로는 고르게 분포”

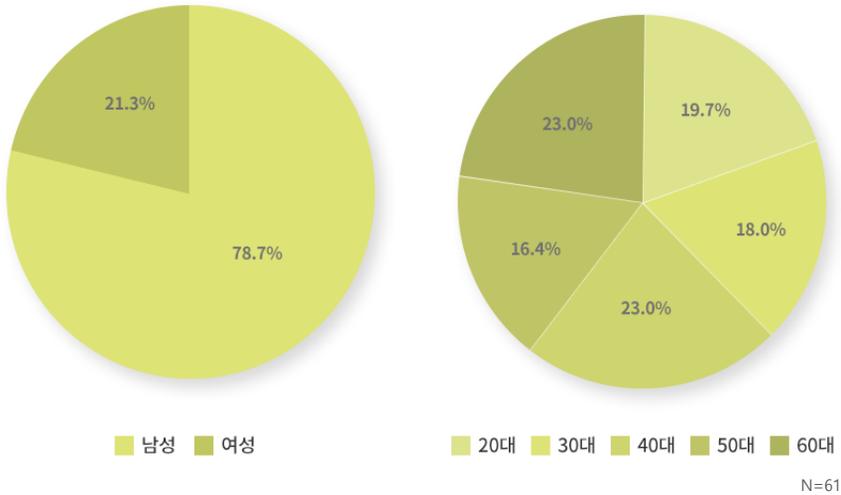


그림8. 온라인 혐오표현 생산 주체

V. 속의 과정 3: 온라인 증오발언 실태 파악을 위한 카카오 데이터 분석

카카오 내부적으로 파악한 증오발언의 유형과 특성

카카오는 미디어자문위 회의 및 공동연구 진행과정에서, 우리 사회가 공감할 수 있는 세부적인 증오발언 대응안을 마련하기 위해서는 다양하고 참여한 쟁점이 선결되어야 한다는 점을 파악했다. 표현의 제재는 곧 표현의 자유를 제한하는 것이므로, 목적과 수단의 정당성이 확보되어야 하며, 과잉금지의 원칙 내에서 이루어져야 한다. 즉, 제재 대상이 되는 표현을 명확히 판별하고, 효과적이되 과도하지 않은 제재 방법을 정하는 기준이 마련되어야 하는 것이다. 따라서 증오발언 정책을 마련하기 위해서는 다음과 같은 세 가지 쟁점이 치열하게 논의되어야 한다고 보았다.

첫째, 증오발언으로부터 우리 사회와 플랫폼 사업자가 보호해야 할 대상을 판별하는 기준과 이 기준에 따른 보호 대상을 정하는 데 어려움이 있다. 각자가 생각하는, 보호해야 할 집단과 보호하지 않아도 되는 집단을 나누는 기준에 매우 큰 차이가 있기 때문이다.

사회 구성원들의 평균적인 인식을 조사해 보호 대상을 선별하는 방식을 적용하는 것 또한 쉽지 않다. 특정 특성을 가진 집단이 보호받지 않아도 된다는 평균적인 인식이 발견되었을 때, 이는 '옳은 것'이라기 보다는 오히려 그 집단에 대한 차별·혐오·증오가 만연해 있다는 것을 방증하는 것일 수 있기 때문이다.

둘째, 증오발언의 형식과 내용 등에 따라 분류하는 기준을 세우는 과정이 필요하다. 특정 집단을 향한 표현이라도 그 목적과 내용에 따라 미치는 영향이 다를 수 있기 때문이다.

셋째, 비례 원칙에 부합하며, 부작용을 최소화하는 규율 방안을 마련해야 한다. 모든 증오발언을 일괄적으로 제재하기 보다는, 표현의 잠재·즉시적 효과에 따라 제재의 수위와 방법이 달라야 하며, 제재 적용이 어려운 경우 그 외의 방법(예를 들어, 교육적 접근)도 고려해야 할 필요가 있다.

이러한 쟁점들에 대해 카카오는 내부적으로 실태를 파악하고 규율 방안을 논의하였고, 그 결과를 미디어자문위에 공유하였다. 다음은 내부 분석의 결과이며, 이에 대한 자문위원들의 피드백은 [제3차 미디어자문위 회의록](#)에서 확인할 수 있다.

증오발언으로부터 보호해야 할 대상은 누구인가?

쟁점들 중에서도 증오발언으로부터 보호해야 할 대상을 정하는 것이 가장 먼저 이루어져야 하는 판단이자, 가장 어려운 부분이다. 따라서 그 판단 기준을 논의 하기 위해, 우리 사회에서 어떤 집단에 대한 공격·비하가 이루어지고 있는지 살펴보고자 했다.

일반적으로 증오발언의 표적은 '역사적·사회적 소수자로서 정신적·신체적 압박을 받아온 집단'으로 알려져 있다. 그러나 이러한 표적 집단의 속성 정의만으로는 우리 사회 현실에서 차별·혐오·증오의 대상이 되고 있는 집단들을 포괄하기 어렵다. 따라서 우리 사회에서 어떤 집단이 공격·비하의 대상이 되고 있는지를 살펴봄으로써 증오발언의 표적으로 고려될 수 있는 집단들을 폭넓게 나열하고, 그 중에서 보호 대상을 정하기 위해 논의해야 할 쟁점을 도출하고자 했다.

이를 위해 카카오의 서비스 운영 과정에서 2018년 8월부터 최근까지 '욕설, 비방, 차별, 혐오' 사유로 신고되어 욕설치환 DB 등록 여부를 검토한 바 있는 단어들을 분석했다. 단어들 중, 특정 집단을 공격하거나 비하하는 것으로 볼 수 있는 표현을 일차적으로 선별하고, 이 단어들이 표적으로 하는 집단에 따라 분류했다. 여기서 '특정 집단'을 사전에 정의하지 않고, 개인에 대한 단순 모욕이 아닌, 어떠한 특성을 기반으로 집단을 나누어 공격·비하하는 경우를 모두 포함하고자 했다.

다만 증오발언의 문장 단위 맥락과 의미를 판단하지는 않는 현 정책상, 특정 집단에 대한 이름짓기(labeling) 형태의 표현에 관한 데이터만 확보 가능하다는 한계로 인해 분석의 포괄성이 담보되지 않는다. 또한 집단으로 구분되긴 하나 보호 특성으로서 인정되기 어려운 '정치적 성향'에 따른 집단에 대한 공격·비하 표현은 분석의 용이성을 위해 제외했다.

공격·비하의 표적이 되는 집단들은 '인종(피부색, 언어, 국가, 민족, 혈통 등의 특성을 통칭), 지역, 종교(신념), 나이, 장애, 성별, 성정체성·성적지향, 직업, 기타'의 9가지로 분류했다. 이 들은 '한국인터넷자율정책기구' 정책규정상 '차별적 표현 완화를 위한 정책'에 제시된 것 및 인권위의 2019년 <혐오표현 리포트>에서 제시한 혐오표현 개념(p.12)을 참조한 후, 기존 정의상 나열되지 않는 집단 특성을 파악하기 위해 '기타' 항목을 추가했다.

카카오의 서비스 운영

과정에서 '욕설, 비방, 차별, 혐오' 사유로 신고되어

욕설치환DB 등록 여부를

검토한 바 있는 단어들을 분석

분류 결과 발견한 증오발언 표적의 특성은 다음과 같았다.

무엇보다 성별이 가장 두드러지는 공격·비하의 대상이었다. 여성에 대한 공격·비하성 이름짓기가 제재 대상 키워드의 대다수를 차지했다. 그 중에서도 ‘한국 여성’으로 공격·비하가 집중되는 경향을 보였다. 다만 여성이라는 특성을 공격·비하하는 것인지 아니면 단순 욕설인지 구분하기 어려운 경우가 많았다.

또한 남성에 대한 공격·비하성 이름짓기도 여성보다는 아니지만 적지 않은 수준이었고, 마찬가지로 단순 욕설과 성별에 대한 공격·비하를 구분하기 어려웠다. 특이점은 여성이나 남성을 옹호하는 신념을 가진 이들에 대한 이름짓기도 발견된다는 것이었다.

지역의 경우 전라도, 경상도, 대구 등 특정 지역민을 공격·비하하는 이름짓기가 다수 있었다.

인종(피부색, 언어, 국가, 민족, 혈통)에 대해서는 예상과 달리 인종차별적 이름짓기는 거의 발견되지 않았다. 오히려 한국, 중국, 일본 등, 인종적 소수성 보다는 국가에 대한 반감에 기반해 해당 국민을 공격·비하하는 키워드만 발견됐다. 성 정체성과 성적 지향에 대한 공격·비하 이름짓기는 한 가지만 발견되어, 실제 관련 콘텐츠나 댓글 등에서 목격할 수 있는 양상과는 괴리가 있었다.

종교(신념)와 관련해서는 기독교인과 이슬람교인을 공격·비하하는 이름짓기가 있었다.

각 연령대별로는 중년 남성, 청소년, 노년층 각각을 비하하는 이름짓기가 발견됐다.

특정 직업과 관련해서는 기자를 비하하는 이름짓기가 있었다.

이상의 분류 기준에 포함되지 않은 집단을 표적으로 한 표현으로는 저연령 아동의 어머니, 흡연자 등 생래적이지 않은 특성을 가진 집단에 대한 이름짓기가 있었다.

이상의 결과를 참고하건대, 증오발언으로부터 보호가 필요한 집단을 정하는 데에 다음과 같은 세부 쟁점을 논의해야 할 것으로 파악됐다.

- 보호 특성을 사회적·역사적으로 차별을 받고 있는 집단 특성으로 한정할 것인가?
 - 예: 남성도 보호 대상이 될 수 있는가?

“여러 사례를 종합해 본 결과
증오발언으로부터 보호할
집단을 정하는 데 여러 세부
쟁점이 존재”

- 예: 우리나라 상황에서 약자라고 볼 수 없는 국가·인종(중국, 일본 등)이 보호 대상이 될 수 있는가?
- 예: 보호 대상이 되는 지역을 나누는 기준은?
- 예: 우리나라에서 소수라고 볼 수 없는 기독교도 보호 대상인가? 4대종파 외 소수 종교는?
- 예: 아저씨, 노인, 청소년 등 연령에 따른 집단도 보호 대상인가?
- 예: 사회적 지위가 낮다고 보기 어려운 직업도 보호 대상인가?
- 예: 생래적이지 않은 다양한 특성에 의해 구분된 집단도 보호 대상이 될 수 있는가?
- 보호 대상을 옹호하는 신념을 가진 집단 또한 보호의 대상인가?
 - 예: 페미니스트, 다문화주의자
- 상호교차성이 있는 경우, 결정된 보호 대상의 범위를 뚜렷하게 적용하기 어려운 경우 어떻게 할 것인가?
 - 예: 일본이라는 국가는 보호 특성이 아니고, 여성은 보호 특성일 경우, 일본 여성은 보호 대상인가?
- 특정 보호 집단을 직접 공격·비하하는 것이 아닌, 그들의 특성에 빗대어 다른 집단을 공격·비하하고자 하는 경우에도 보호 특성이 적용되는가?
 - 예: 특정 정치적 신념을 가진 사람을 비하하기 위해 여성을 낮춰 부르는 어미를 붙여 지칭하는 경우
 - 예: 특정 브랜드를 옹호하는 사람을 이슬람교도에 빗대어 비하하는 경우

어떤 표현이 증오발언에 해당하며, 어떻게 분류할 수 있나?

증오발언은 위와 같이 이름짓기를 통한 직접적인 비하 외에도 다양한 형태로 나타날 수 있다. 표적 집단에 대한 모욕, 비하, 멸시, 위협 또는 차별·폭력의 선전과 선동을 통해 차별을 정당화·조장·강화하는 효과를 갖는 표현이 폭넓게 증오발언으로 정의되며, 세부적으로는 아래와 같은 내용과 맥락으로 나누어 볼 수 있다.

“‘효과’를 기준으로 ‘차별적 인식을 강화하는 표현’, ‘사회정치적 의견을 선동하는 표현’, ‘실질적 공격(폭력)에 해당하는 표현’으로 묶어”

- 편견을 강화하는 특징 표현
- 편견·차별을 정당화하는 편향된 정보
- 부정적 감정의 표현
- 사회적 배제를 촉구하는 표현
- 차별적 정치 행동, 정책을 촉구하는 표현
- 모욕, 비하, 멸시하는 표현
- 물리적 (성)폭력을 예고, 조장, 중계하는 표현

이를 다시 한 번 '효과'를 기준으로, 아래와 같이 각각 차별적 인식을 강화하는 표현, 사회정치적 의견을 선동하는 표현, 실질적 공격(폭력)에 해당하는 표현으로 묶어볼 수 있다.

효과에 따른 분류	표현의 내용
차별적 인식을 강화하는 표현	- 편견을 강화하는 특징 표현 - 편견·차별을 정당화하는 편향된 정보 - 부정적 감정의 표현
사회정치적 의견을 선동하는 표현	- 사회적 배제를 촉구하는 표현 - 차별적 정치 행동, 정책을 촉구하는 표현
실질적인 공격 표현	- 모욕, 비하, 멸시하는 표현 - 물리적 (성)폭력을 예고, 조장, 중계하는 표현

각 대상 또는 표현의 종류에 따른 대응 방안은 무엇인가?

'행동', 즉, 실질적 공격(폭력)에 해당하는 표현은 게시물 관리 정책을 변경해 대응할 필요가 있고, 비교적 용이하게 정책 운영이 가능할 것으로 보인다. 명백히 타인에게 피해를 가했거나, 가할 위험이 있는 표현이라면 관리의 정당성을 확보할 수 있다. 또한 위해성 판단 기준도 상대적으로 명확하므로, 서비스 제공자의 자의적·비일관적 판단으로 인한 문제가 발생할 가능성이 상대적으로 낮다. 다만 여전히 어떤 수준으로 게시물 및 이용자를 제재해야 하는지에 대한 논의는 필요하다. 차별·혐오·증오에 기반한 실질적 공격(폭력) 표현이, 이미 서비스 제공자들이 제재하고 있는 유해 정보보다 높은, 낮은, 또는 같은 수준에서 제재되어야 할지 사회적 합의가 이루어져야 한다. 예를 들어, 최근 프랑스에서는 이러

한 표현을 아동 성범죄와 같은 수준으로 보고 서비스 제공자가 지체없이 삭제하도록 하는 법안이 논란이 되었는데, 과도한 표현의 자유 침해 우려가 제기됐다. 반면 '인식'과 '의견'에 영향을 미치는 표현에 대해서는 게시물 관리 정책, 즉 게시물이나 이용자를 제재하는 방식으로 대응하기 어려운 것이 사실이다. 이러한 표현의 진실성이나 정당성을 판단하는 것은, 그 실행가능성을 떠나, 서비스 제공자의 권한으로서 우리 사회의 인정을 받기 어려울 것으로 보인다.

따라서 차별·혐오·증오 인식과 의견에 영향을 미치는 표현에 대해서는, 서비스 제공자가 게시물이나 이용자를 직접적으로 제재하는 것 외에 사회적으로 수용 가능한 대응 방안을 고민할 필요가 있다.

또한 보호 대상의 집단과 상황에 따라 위의 방식, 즉 '행동'에 대해서는 게시물 관리 정책으로 대응하되 '인식'과 '의견'에 대해서는 이외의 접근법을 취하는 방식이 일률적으로 적용될 수 있는지도 논의할 여지가 있다. 예를 들어, 특정 시기에 문제를 일으킨 집단에 대해서는 언어적 공격을 제재하는 것이 역효과를 일으킬 수 있고, 반대로 제도적 보호를 필요로 하는 특정 집단에 대한 사회정치적 배제 의견은 제재가 필요한 경우도 발생할 수 있기 때문이다. 이러한 특수한 경우에 대한 대응 기준도 선제적 논의가 필요하다.

“‘실질적 공격(폭력)’에 해당하는 표현에 대해서는 게시물 관리 정책을 변경해 대응할 필요가 있어”

“‘인식’과 ‘의견’에 영향을 미치는 표현은 게시물·이용자 제재만으로 대응하긴 어려워”

VI. 속의 과정 4: 시민사회 의견 청취를 위한 라운드테이블

시민사회 전문가들에 증오발언 대응 정책에 대한 카카오의 고민 공유하고 의견 청취

마지막으로 카카오는 시민사회의 의견을 수렴하기 위한 목적으로 시민사회 전문가들을 초청해 라운드테이블을 진행했다. 여기서 시민사회 전문가란, 증오발언 및 표현의 자유에 대해 깊이 고민하고 사회적 변화를 촉구해 온 활동가들을 뜻한다. 라운드테이블은 2020년 10월 14일 서울 중구에서 열었으며, 참석한 시민사회 전문가들은 아래와 같다(가나다 순). 미디어자문위원들 또한 자리에 참석해 시민사회의 의견을 청취했으며, 박재영 위원장이 좌장으로서 라운드테이블의 진행을 맡았다.

- | 김용구 한국장애인인권포럼 소장
- | 김지학 한국다양성연구소 소장 (국제앰네스티 한국지부 이사)
- | 박록삼 한국기자협회 부회장 (서울신문 논설위원)
- | 손지원 오픈넷 변호사
- | 윤철한 경제정의실천시민연합 기획연대국장 (서면으로 의견 제출)
- | 이윤소 한국여성민우회 성평등미디어팀장
- | 이완 아시아인권문화연대 대표
- | 임동준 민주언론시민연합 활동가
- | 조소담 닷페이스 대표
- | 조진경 십대여성인권센터 대표

카카오는 시민사회 전문가들에게 그동안 카카오가 증오발언에 대해 고민해 온 과정을 공유하고, 두 가지 질문을 던졌다.

- 증오발언 문제 해소를 위해 인터넷 기업이 이용자 정책 수립 등의 방법으로 적극적으로 나서는 것은 적절한가?
- 증오발언 정책의 보호 대상과 관련 규제의 적용 범위는 무엇인가?

이 질문에 대한 답과 함께 시민사회 전문가들은 다양한 관점의 의견을 제시했다. 아래 발언 내용은 순서와 관계 없이 주제별로 묶었다. 원문은 [\(부록 1\)의 회의록](#)에서 확인할 수 있다.

기업의 증오발언 문제 대응의 정당성

□ 적극 대응 필요

- 기업의 인권보호 책임의 관점에서 접근한다면, 카카오가 혐오표현에 개입하는 것은 정당함
 - 일반 기업이 이용자를 제재한다는 관점에서는 부정적일 수 있으나, 인권 침해에 카카오의 상품·서비스가 연루되거나 기여하고 있다면 이에 개입해야 함
- 혐오표현에 관한 제재는 국가의 역할이나, 사회적 합의가 부재하다는 이유로 국가가 역할을 수행하지 못하고 있음
 - 따라서 카카오가 자율적, 선제적으로 개입하는 것이 바람직함
 - 혐오표현에 대한 개입은 범 미디어 생태계의 일원으로서 카카오의 책무이자 공공적 역할

□ 최소 개입 필요

- 공론장은 시민이 만드는 것이므로, 사기업의 검열은 최소화해야 (절대 개입 불가가 아닌 최소 개입이 필요하다는 취지)

증오발언과 표현의 자유

□ 증오발언 제재 강조

- 혐오표현의 제재는 표현의 자유와 기본권을 향상시킴
 - 혐오표현 제재는 표현의 자유와 대치되지 않을 뿐 아니라, 오히려 소수자의 표현의 자유를 증진시킴
 - 모두가 자기표현을 할 수 있는 사회를 위해서는 혐오표현을 제재해야 함

□ 표현의 자유 강조

- 표현의 자유는 옳은 표현만을 위한 것이 아니라, 우리가 싫어하는 생각을 표현할 자유도 포함
 - 옳지 않은 표현도 일정한 가치와 기능을 가짐. 잘못된 의견을 반박하고 논쟁을 촉적하는 과정에서 사회적 논의가 풍성해짐. 논의를 통해 더 옳은 방향을 형성해 나가는 데 옳지 않은 표현이 기여

“혐오표현에 대한 개입은
범 미디어 생태계 일원으로서
카카오의 책무”

“소수자의 범주를 나열하고,
‘등’을 포함하여 범위를 넓힐
수 있도록 해야”

“타인을 이질적 존재로
인지하게 하는 경우도 정책적
선언 수준에서는 포괄 필요”

증오발언으로부터 보호해야 할 대상의 범위

□ 소수자로 보호 대상 명확화 필요

- 혐오표현은 “사회적 소수자에 대한 고정관념과 편견을 강화하고, 차별과 억압을 만드는 표현”
- 소수자의 범주(여성, 성소수자, 장애인, 이주민...)를 나열하고, ‘등’을 포함하여 범위를 넓힐 수 있도록 해야 함
- 사회적 논란을 피하고자 ‘성소수자’를 빼는 식으로 진행하는 것은 옳바르지 않음
- 사회적 소수자에 해당하지 않는 집단은 혐오표현의 대상으로 볼 수 없음
 - 주류적 지위에 있거나, 사회에서 공포·위축 위험이 없는 집단에 대한 표현마저 제재하면, 사회적 소수자들의 표현을 오히려 억압할 수 있음
 - 예를 들어 미러링, 패러디 등을 통한 기득권에 대한 대항표현

□ 포괄적 보호 필요

- 타인을 이질적 존재로 인지하게 하는 경우도 ‘정책적 선언’ 수준에서는 모두 포괄할 필요가 있음
 - 단, 제재는 현실적인 문제이므로 다른 접근 필요
- 보호의 대상을 한정하기는 어려움
 - 예를 들어 기자, 남성, 개신교인 등에 대한 비하가 혐오표현이 아니라고 정확히 말할 수 없으며, 이런 표현을 계속 사용하는 것이 옳다고 볼 수 없음

제재가 필요한 표현의 범위

□ 포괄적 제재 필요

- 모욕 의도와 상관 없이, 사회적 약자들에 대해 일상적으로 혐오표현이 쓰이고 있으므로 구체적인 강력한 조치 필요
- 혐오표현에 해당하는 단어를 규정하는 것만으로는 불충분. 맥락 고려가 필요함
- 직접적인 표현이 아니라, 이미지 등과 결합된 형태로 혐오의 취지가 드러나는 경우도 고려 필요

□ 최소 제재 필요

- 해악이 중대하고 명백한 혐오표현만 제재해야 함
 - 소수자 집단이 모욕감을 느낄 수 있고, 심리적으로 위축될 수 있다는 가능성을 이유로 포괄적으로 금지해서는 안됨
 - 혐오표현 규제의 목적은 표적 집단에 대한 차별·폭력·배제 및 공론장에서의 표현·영향력 위축을 막기 위한 것
 - 실질적 불안·공포·위축을 불러오고, 사회에서 실질적 배제나 폭력을 야기할 수 있는 표현에만 제재가 한정되어야 함
 - 구체적인 허위 정보를 언급하며 사회적 차별·배제를 정당화하거나 조장하는 경우도 규제 필요

“실질적 불안, 공포, 위축을 불러오고, 사회에서 실질적 배제나 폭력을 야기할 수 있는 표현에만 제재가 한정되어야”

이용자 제재 외 대안

- 플랫폼이 혐오표현을 허용하지 않는다는 것을 천명하는 것 만으로도 큰 의미가 있음
 - 보호 대상이 자신들이 보호받을 수 있음을 인지하게 하는 측면에서 중요
- 처벌 중심보다는 교육 중심으로 접근 필요
 - 우리 사회는 서로 혐오표현에 대한 이해 수준이 다르므로, 사례 등을 통해서 어떤 것들이 왜 혐오표현인지 이해시키는 것이 더 중요
 - 본인도 모르게 혐오표현을 하는 경우도 많음
 - 정책·가이드라인을 만들더라도, 이를 이용자들이 읽게 할 방법을 더 많이 고민해야 함
- 혐오표현을 포함한 게시물을 표시하는 방식으로 사람들이 생각하고 배울 기회가 될 것
 - 게시물에 혐오표현 포함을 표시하는 단계, 삭제하는 단계, 게시자를 차단하는 단계 등 여러 단계가 있을 것
 - 플랫폼이 단독적으로 판단하거나, 폐쇄적으로 처분하기 보다는 이 게시물에 문제가 있다는 것을 알릴 필요 있음
- 자율적 심의기구를 통해 혐오표현을 판단하고 조치할 수 있음
 - 낮은 단계에서는 기계적 기준을 바탕으로 기술적으로 거르는 것이 가능. 논쟁거리가 있거나 높은 단계에서는 심의기구의 판단 필요

“혐오표현은 찬반의 문제가
아니므로 ‘근절·방지 정책’이
되어야”

- 이용자가 혐오표현 피해 방지와 구제를 위해 능동적으로 대처할 수 있는 방법을 제공해야 함
 - 게시물의 공개 범위 조절, 댓글을 필터링할 수 있는 기능 등
 - 피해 구제를 원할 때 신고 및 증거 수집을 용이하게 해야 함
- 실명화를 추진해야 함. 사이버 세계와 현실 세계의 룰을 같게 만들어야 함

기타 제언

- 관련 정책은 여러 버전을 만들어, 합의할 수 있는 버전을 도출한 후, 논의를 통해 발전시킬 수 있음
- 카카오가 만드는 혐오표현도 주의해야
 - 서비스 및 콘텐츠가 혐오표현을 확산·양성·조장하지 않는지 관리감독할 필요 있음
 - 내부적으로도 혐오표현에 유의하고 있다는 점을 보여주어야 이용자 제재도 설득력 생김
 - 다양성 리포트 등 내부인의 인권의식 수준을 높이기 위한 조치도 병행 필요
- 혐오표현의 시작점이 되는 언론보도 문제도 해소해야 함
 - 언론이 혐오표현을 사용함으로써 이용자에게 혐오정서가 확산되고 이를 다시 언론이 끌고 들어오는 악순환 반복
 - 뉴스제평위의 제재 규정 보완 및 실질적으로 기능하도록 개선 필요
- 정책 이름에 회사의 의도가 명확히 드러나도록 해야 함
 - 혐오표현은 찬반의 문제가 아니므로 ‘근절·방지 정책’ 등이 되어야 함

VII. 나가며: 카카오의 증오발언 대응 정책 방향

속의 과정을 통해 도출된 쟁점 및 대안

카카오는 미디어자문위, 공동연구, 라운드테이블 등의 속의 과정을 거쳐, 증오발언에 관한 몇 가지 쟁점들을 발견했다. 이용자 및 전문가들의 의견을 더욱 폭넓게 수렴해 이 쟁점들에 대한 답을 찾는 것은 향후 카카오의 증오발언 정책 수립 및 이행을 위해 꼭 필요한 과정으로 보인다.

개념 및 용어의 정의 필요성

현재 국내에서는 hate speech를 ‘혐오표현’으로 번역하여 논의에 활용하고 있으며, 카카오의 속의 과정에 참여한 전문가들 및 참고자료들도 혐오표현이라는 용어를 주로 사용했다. 다만 ‘혐오’가 실질적 규제의 대상이 되는 개념으로서 적절한지에 대해서는 논의가 필요하다는 의견이 많았다.

혐오(disgust)는 생리심리적 원인에 기반하고 반 대상 정향(대상으로부터 멀어지고자 하는)인 감정에 가까운 반면, 증오(hate)는 사회계층적 원인에 기반하고 대상 정향(대상에 대한 어떤 행동을 목표로 하는)인 감정이다. 즉, 혐오보다는 증오가 사회적 배제나 공격 등 폭력성을 내포하므로 실질적 규율의 대상이 되는 개념으로서 적절하다고 판단된다. 혐오라는 감정적 반응을 실질적으로 규제할 수도 없고, 의견이나 감정의 표출까지 규제할 때의 부작용이 클 것으로 우려된다.

따라서 녹서의 초입에 제시한 바와 같이, 향후 카카오의 관련 정책 논의에는 ‘증오발언’이라는 용어를 사용하고, 이를 ‘차별적 인식에 기반해 특정 집단에 대해 극단적인 언어적 폭력을 가하는 행위’로 정의하고자 한다.

보호 대상의 범위

국내외 논의를 살펴보면 통상 ‘사회적 억압과 차별의 대상이 되는 비주류 내지 소수자 집단’을 증오발언으로부터의 보호 대상으로 보고 있다. 속의 과정에서는 보호 특성을 한정하지 않고 생래적 또는 후천적으로 획득되는 정체성을 바탕으로 집단을 구분하여 공격하는 발언을 폭넓게 규제해야 한다는 의견도 있었다.

속의 과정을 통해 확인한
증오발언 문제의 쟁점과
대응 방향

1. 이준웅·박장희 (2018). 모든 더러운 말들: 증오발언 규제론 및 규제반대론 검토. 서울대학교 법학, 59(3), 1~43.

자율규제의 형태로 이용자의 증오발언을 사업자가 제재할 수 있다는 데 이견 없으나 무엇을 증오발언으로 판단하고 제재할 것인지는 다소 의견 차이 있어

그러나 실질적 차별이 발생하는 요인들로 보호 특성을 한정해야만 증오발언 정책의 본래 목적을 달성할 수 있다는 의견이 다수였다.

다만 소수·비주류 집단의 범위는 명확하게 구분하기 어렵고 사회적 상황 변화에 따라 변화할 수 있으므로, 특정 소수자 집단을 보호 대상으로 한정하지 않고 유동적으로 대응할 수 있도록 범위를 확장할 수 있는 가능성을 열어둘 것이 권고되기도 했다.

또한 일반적으로 주류·다수 집단은 증오발언의 보호 대상이 될 수 없다는 점을 명확히 해야 한다는 의견이 많았다. 주류·다수 집단까지 증오발언의 보호 대상이 되면, 소수·비주류 집단의 대항표현까지 제한하게 될 우려가 있기 때문이다. 따라서 향후 증오발언 정책 수립시에는 보호 대상 내지 보호 특성을 나열하되, 나열된 것에 한정되지 않는다는 단서를 붙일 필요가 있다.

이용자 규제의 대상이 될 수 있는 발언의 범위

콘텐츠를 삭제하거나 이용자의 서비스 이용을 제한하는 등 자율규제의 형태로 이용자의 증오발언을 사업자가 제재할 수 있다는 데 큰 이견은 없었다. 다만 보호 대상을 향한 발언 중 어떤 발언까지를 증오발언으로 판단하고 얼마나 강하게 제재할 것인지에 대해서는 다소의 의견 차이가 발견됐다.

먼저 표적집단에 대한 직접적 공격 뿐 아니라, 차별을 강화하고 조장하는 효과가 있는 발언까지를 이용자 규제의 대상으로 보아야 한다는 입장이 있었다. 증오발언이 일상에 만연해 있고, 거짓정보 내지 가짜뉴스가 특정 집단에 대한 차별을 강화하고 있으므로 조치가 필요하다는 것이다.

한편으로는 포괄적인 이용자 증오발언 규제에 대해 우려하는 입장도 있었는데, 그 이유는 대체로 표현의 자유를 훼손할 가능성이 높다는 것이었다. 지나치게 광범위한 발언에 이용자 규제를 적용할 경우 다양한 의견을 말할 수 있는 이용자의 권리가 과도하게 위축될 수 있다. 또한 표현의 자유가 가진 또다른 순기능, 즉 ‘나쁜 표현’에도 노출됨으로써 이에 대한 문제제기와 토론이 이루어지도록 하는 기능이 약화될 수 있다. 이를 고려하면 현재로서는 표적집단에 대한 일방적 모욕이나 물리적 폭력 선동 등 즉각적 해악이 명확한 발언에 한해 이용자 규제를 적용하는 방식이 적절해 보인다.

외부로부터의 의견과는 별개로 서비스 운영자의 입장에서는 해악이 명백하지 않은 발언까지 제재하는 데 따르는 어려움을 고려하지 않을 수 없다. 사업자로

서는 발화의 맥락이나 진실성을 판단하여 차별을 강화하거나 조장하는 발언인지를 판단하기 쉽지 않으며, 그러한 권한을 이용자들이 인정할 것이라고 기대하기도 어렵다. 또한 신속한 대응이 더 중요할 수 있는 증오발언 제재에 심의기구 등을 통해 증오발언을 심사하는 체계가 어울리지 않을 수 있다는 점을 고려해야 한다.

증오발언을 억제하는 것보다
차별·혐오·증오를
줄여나가려는 인식 변화가
근본적 해결책이며, 카카오가
교육·캠페인 통해 기여 가능

이용자 규제 외의 증오발언 대응 방법

이용자 규제 외에도 카카오가 증오발언, 나아가 우리사회의 차별에 대응하는 방법으로서 여러 대안이 제안됐다.

자문위, 공동연구, 라운드테이블을 관통하는 핵심적 의견은 이용자 교육을 통한 인식 제고 효과였다. 증오발언 자체를 억제하는 것보다 증오발언의 뿌리인 차별·혐오·증오를 줄여나가기 위한 인식 변화가 가장 근본적인 해결책이며, 카카오가 교육이나 캠페인을 통해 기여할 수 있다는 것이다. 직접적인 이용자 교육 프로그램을 제공하는 것 외에도, 카카오가 증오발언에 대한 대응 의지를 천명하는 것만으로도 우리 사회에 이러한 문제가 실존하고 있음을 알리고 일반 이용자들을 논의에 참여시킬 수 있을 것이라는 기대도 있었다.

한편으로는 이용자의 증오발언 콘텐츠 뿐 아니라, 사업자가 내부적으로 만드는 콘텐츠, 유통하는 뉴스 콘텐츠 등에서의 증오발언을 경계하고 내부 임직원을 교육해야 한다는 의견도 있었다.

증오발언 콘텐츠를 가리거나 삭제하는 것이 아니라, 문제의 소지가 있는 콘텐츠임을 표시하는 방식으로 대응하는 안도 있었다. 즉각적 해악이 명백하지 않아 판단이 어렵지만 증오발언으로서 잠재적 해악이 있는 경우, 이를 표시함으로써 이용자들 스스로가 판단하고 추가적으로 논의할 수 있는 기회를 제공하자는 것이다. 이 방식 또한 표현의 자유를 위축하지 않으면서도 더 넓은 범위의 증오발언에 대응할 수 있다는 측면에서 유용한 대안으로 보인다.

증오발언 근절을 위한 카카오의 원칙

숙의 과정을 통해 쟁점들에 대해 다양한 이해관계자 및 전문가들의 의견을 청취하고 연구한 결과, 카카오가 국내 인터넷 생태계를 선도하는 기업으로서 적극

미디어자문위 및
라운드테이블 참석자들의
의견을 또 한 번 수렴해
최종안을 마련한
〈증오발언 근절을 위한
카카오의 원칙〉

적으로 증오발언에 대응해야 할 필요성이 확인되었다. 카카오는 자유롭고 다양한 의사소통을 매개하는 인터넷 서비스 제공 사업자로서, 표현의 자유를 핵심 가치로서 수호하려는 노력을 이어가고 있다, 그러나 표현의 자유를 남용한 증오발언을 근절해야 한다는 사회적 요구도 커지고 있다. 다만 증오발언 정책과 표현의 자유 사이의 충돌을 최소화하고 다양한 이용자의 발언의 기회를 보장하기 위해서는, 정책의 목표와 대상을 뚜렷하게 하되 제재적 수단은 최소한으로만 적용할 필요가 있다.

이러한 방향성을 담아 카카오는 〈증오발언 근절을 위한 카카오의 원칙〉을 만들었다. 이 원칙 아래 향후 카카오는 운영정책 등 이용자를 위한 정책 및 가이드라인을 지속적으로 개편하고, 서비스 및 기술을 고도화해 가고자 한다. 또한 지금까지와 같이 다양한 이해관계자들과의 소통을 지속하며, 그들의 제안을 적극 수용하여 더 나은 대안들을 모색하고 실행할 것이다.

이 원칙은 카카오에서 초안을 작성한 뒤, 2020년 11월 말부터 12월 초 미디어자문위원 및 라운드테이블 참석자들의 의견을 수렴해 확정하는 과정을 거쳐 최종안을 만들었다.

증오발언 근절을 위한 카카오의 원칙

카카오는 기술과 사람의 연결을 통해 더 나은 세상을 만들고자 합니다. 카카오는 보다 자유롭고 다양한 연결을 위해 이용자의 표현의 자유를 보호하려고 힘쓰고 있습니다.

이용자의 인권을 보호하는 것 또한 카카오의 중요한 책무입니다. 카카오는 표현의 자유를 남용하여 타인의 안전을 위협하는 발언을 경계합니다.

카카오는 온라인 증오발언(hate speech)을 근절하기 위해 노력해 나가려 합니다. 아래와 같은 카카오의 증오발언 대응 정책 방향이 온전히 이행되고, 나아가 모두에게 안전한 디지털 환경이 만들어지도록 여러분의 많은 관심과 참여를 부탁드립니다.

1. 카카오는 출신(국가, 지역 등)·인종·외양·장애 및 질병 유무·사회 경제적 상황 및 지위·종교·연령·성별·성 정체성·성적 지향 또는 기타 정체성 요인 등을 이유로 특정 대상을 차별하거나, 이에 대한 편견을 조장하며, 일방적으로 모욕하거나 배척하는 행위에 반대합니다.
2. 카카오는 이러한 차별에 기반해 특정인과 특정 집단을 공격하는 발언을 증오발언으로 정의합니다. 증오발언은 이용자의 정서적 안전을 위협할 뿐 아니라, 사회적 배척과 물리적 폭력을 유발합니다. 증오발언은 다양한 이용자가 발언에 나설 자유를 위축시킬 뿐만 아니라 우리 사회의 신뢰와 건강성을 저해합니다. 카카오는 이용자의 인권과 존엄성을 훼손하고 안전을 위협하는 증오발언에 강경하게 대처하겠습니다.
3. 이용자는 카카오 서비스 내 공개된 공간에서 특정인과 특정 집단에 대한 폭력을 선동하거나 인간의 존엄성을 훼손하는 발언에 유의해야 합니다. 이용자는 타인의 존엄성과 안전을 위협하지 않는 한 여전히 공공정책이나 자신의 신념 등에 대해 자유롭게 의견을 개진할 수 있습니다.
4. 카카오는 증오발언을 근절하기 위해 앞으로도 정책, 기술, 서비스 기획 및 디자인을 고도화해 나가겠습니다. 더불어 사내 교육과 모니터링을 강화하는 등 내부로부터의 차별과 증오발언을 경계하겠습니다.

위 원칙은 공개 게시물 영역에 한합니다. 카카오톡의 사적 대화공간, 메일, 독서클럽 등 개인화한 서비스와 커뮤니티 비공개 게시물 등에는 프라이버시 존중을 최우선 가치로 적용할 것입니다.

카카오는 모든 이용자가 안전하게 표현의 자유를 누릴 수 있는 환경을 만들기 위해 노력하겠습니다. 모두에게 안전한 디지털 공간을 만들기 위한 노력에 이용자 여러분도 적극적으로 동참해 주시길 부탁드립니다.

부록 1. 속의 과정 회의록

제1회 미디어자문위 회의

- 일시: 2020.4.24. 16:00~19:00
- 장소: 서울 중구
- 참석자
 - 카카오: 양현서 부사장, 손정아 부사장, 김대원 이사, 이준목 이사, 박용준 이사, 김성환 부장, 김대기 부장, 김수원 연구원
 - 자문위원: 박재영 교수, 김민정 교수, 김장현 교수, 도준호 교수, 이준웅 교수

**양현서
카카오 부사장** 카카오는 사내에서 이용자 권리 보호와 댓글 정책을 위한 TF를 운영하고 있다. 이 TF들에서는 댓글에서의 명예훼손, 게시글 삭제 등을 정한다. 댓글 관련 TF에서는 댓글을 어떤 방식으로, 기술적으로 또는 정책적으로 개선할 것인지를 운영한다. 또, 명예 훼손 글에 대한 삭제 정책을 정하는 TF는 KISO와 함께하고 있다.

사실 해외에 비해서, 국내 혐오표현은 아주 극단적인 방향으로 드러나진 않는다. 해외에서보다 우리의 논의 속도가 늦었다고 보는데, 우리나라에서는 종교나 인종과 관련된 극단적인 사태가 많지 않기 때문에, 사회적으로 큰 문제로 인식되지 않는 것 같다. 우리나라에서는 젠더(gender)나 지역 같은 부분에 있어서 다른 나라에 비해 이슈가 있다.

이처럼 아직 대한민국에서 혐오표현의 정의조차 제대로 정립되지 않는 상황인데, 어떤 표현들이 카카오의 플랫폼 안에서 허락되어야 하는지 고민해보게 되었다. 이러한 고민을 사회의 여러 연령대와 계층에 계신 분들, 정부 및 국회, 이용자, 학생 등과 나눠보고, 플랫폼에서 쓰여야 하는 언어의 규범과 정책을 수립해보고 싶다는 생각을 작년부터 했다.

혐오표현은 범주가 넓으며, 반응이 생겨날 수 있다. 그렇기 때문에 혐오표현을 우리 내부의 고민으로 그칠 것이 아니라, 사회적인 고민으로 확대해서 카카오가 혐오표현에 대한 고민 자체를 선제적으로 이끌어보고 싶다. 어떤 방식으로 해나갈지 위원님들이 함께 고민해 주길 바란다.

**김대원
카카오 이사** 기간에 한정을 두지 않고, 혐오표현과 관련해 서비스 현장에서 고민하는 이슈들에 대해 자문을 받으면서, 혐오표현에 대한 정의서부터 관련 서비스 정책, 그리고 규범까지 논의를 했으면 한다. 이번 모임과 다음 모임까지는 혐오표현 개념에 대한 논의를 진행해 틀을 잡고, 관련 세부 주제를 잡아 고민한 뒤 연말에는 서비스 정책을 완성시키거나 규범을 만드는 스케줄을 생각하고 있다. 이번 미디어자문위를 운영하는 방식에 대해 자유롭게 말씀 주셨으면 한다. 방향, 인원보강, 연구 등에 대해 의견을 달라.

손정아
카카오 부사장

혐오표현에 대한 학계 분위기는 어떠한지 궁금하다. 어떤 논의가 이루어지고 있나?

김민정
자문위원

개인적으로 혐오표현에 관심을 가지고 있다. 우선 국내 학계에서는 hate speech를 어떻게 번역할 것인가에서부터, 표현의 자유와 균형을 어떻게 맞출 수 있을지에 대한 다양한 논의들이 있다. 작년에 국가인권위원회에서 혐오차별대응기획단과 추진위원회를 만들었고, 혐오표현에 대한 대응리포트를 발간했다. 리포트는 혐오표현의 정의, 기존 법 조항, 문제, 대응 방안, 법적인 대응 외에도 규범 변화 교육 등의 다양한 내용을 포함하고 있다. 이처럼 혐오표현에 대해 공부할 수 있는 텍스트들은 만들어져 있다.

네이버나 카카오 등 개별 사업자들의 이용자 보호 정책이나 KISO의 차별적 표현 완화 정책 적용 사례가 7개밖에 되지 않는 상황들을 보면, 혐오표현 이슈에 대해 관심이 미미한 것 같다. 반대로 구글이나 페이스북의 경우는, hate speech를 커뮤니티 가이드라인에 독립적 항목으로 설정하여 굉장히 자세히 설명하고 있다. 국내에서는 카카오가 이런 작업을 선도적으로 하면 좋을 것 같다.

최근에 댓글에서 차별과 혐오를 신고하는 기능을 도입했다고 알고 있다. 카카오는 인권위에서 사용하는 혐오표현의 정의를 우선 사용하겠다고는 했는데, 이를 적용하는 것은 실제 상황에서 굉장히 어려운 문제다. 혐오표현의 기본적인 개념이나 정의는 논의를 통해 합의를 하면 되겠지만, 실제로 규제를 하는 방향 또는 집행에 있어서 실질적인 문제에 대한 논의가 많이 필요할 것으로 보인다.

또 시민단체의 참여가 보장되어야 한다고 생각한다. 실질적으로 넓은 범위에서 이용자의 대표성을 확보하고 그들의 논의를 반영할 수 있으면 좋겠다.

이준웅
자문위원

행위규범의 대상으로서의 ‘행위’와 ‘발언’에 대한 논의는 이제 시작이다. 인권위에서 발간한 보고서가 일단 혐오표현에 대한 신호탄을 던졌다. 그런데 인권위원회 보고서에서는 ‘역사적 사실의 부정’을 혐오표현이라고 주장하는 등의 논란의 여지가 있다.

사업자로서 이용자 보호를 위한 규범을 만들어내고 적용을 하려면, 그 내용이 충분히 설득력 있어야 한다. 이때 카카오가 선도적으로 광범위한 토론의 장을 많이 마련하는 것이 중요하다. ‘혐오표현’이라는 말 자체가 사람들을 오도(誤導)할 수 있다. 혐오표현이란 용어 자체가 부정적 반응을 유발하기도 한다. 카카오의 사업을 ‘혐오표현 대응 사업’보다는, ‘차별에 반대하는 카카오’처럼 중립적으로 하는 게 좋을 것 같다. 왜냐하면 대부분의 사람들이 ‘혐오’와 ‘증오’를 구별하지도 못하고, 많은 경우에 혐오표현의 원인과 효과에 대해서 오해하고 있기 때문이다.

양현서
카카오 부사장

남성과 여성을 비하하고, 또 각 성을 성기에 연결시키는 표현이 많아지자 신체를 훼손하는 말을 규제해야 한다는 논쟁이 있었다. 그렇다면 “손에 장을 지진다”도 신체 훼손 표현이므로 규제해야 하느냐는 반박이 있었다. 사업자 입장에서는 컴퓨터나 시가 아니라 서비스 센터를 통해서 규제하는 부분이 많기 때문에 정확한 지침이 필요하다. 카카오가 어떤 표현은 허용하고 반대해야 하는지, 실행의 범위를 만들고 싶다.

이준목
카카오 이사

카카오가 댓글 등 커뮤니티 서비스를 제공하는 사업자로서 어디까지 관리를 해야 하는지, 나아가 카카오가 특정 유형의 발언에 대해 반대한다고 정의했을 때 서비스를 자유롭게 쓸 수 있어야 한다는 외부의 비판에 어떻게 대응해야 하는지 등에 대해 고민이 있다.

이준웅 사업자로서 서비스 제공은 이용자와의 계약 문제이다. 따라서 그러한 표현을 쓰겠다는 이용자가 있다면, 자문위원 그가 제3의 플랫폼으로 가야 한다.

도준호 최근 혐오표현 범위에 다수자에 의한 소수자 차별 개념이 표현되어야 하는가, 그리고 다수자·소수자는 어떻게 판단하는가에 대해 느낀 일이 있었다. 이번에 숙명여자대학교 신입생 중에서 트랜스젠더가 있었는데, 자문위원 결국 스스로 입학을 포기하는 사례가 있었다. 그 사건이 진행되면서 학내 게시판 여론이 어떻게 진행되는지를 보았다. 카카오와 상황은 다르지만 교훈이 있을 것 같다.

일단 팩트를 정리하면, 문제가 됐던 학생이 서류 전형에 냈을 때는 수술하기 전이었고, 제출 후 태국에서 수술을 했다. 전형 과정에서 주민등록번호 뒷자리가 2로 바뀐 것을 교직원들이 발견했고, 외부 유출이 되었다. 이 사실이 알려지면서 문제가 되었는데, 처음에 가장 극렬한 반대 의견을 낸 것이 일부 페미니스트들이었다. 그 페미니스트들이 트랜스젠더를 반대하는 이유는, ‘여성성의 훼손’이다. 생식기를 잘랐다고 해도 여성으로 볼 수 없고, 염색체도 다르다는 것이다.

소수자는 숫자의 문제가 아니다. 인종·국적·성별 등 본인이 선택하지 않은 것으로 차별을 받을 때, 다수자와 소수자가 구별이 되는 것 같다. 여성이 일반적으로 소수자였는데, 이 사건에서는 ‘다수자에 의한 소수자에 대한 차별’이라는 개념이 좀 달라진다. 일반적으로 여성이 소수자인데, 이 사건에서는 또 다른 소수자인 트랜스젠더를 차별한 것이다. 페미니스트 그룹이 글을 올린 다음, 숙대 성소수자 그룹이 페이스북에 트랜스젠더 학생의 입학 환영 메시지를 냈다. 그 다음에 대부분의 학생들이 불편하다며 반대 의견을 쏟아냈다. 결국 이 학생이 입학할 포기하게 된 계기는, 게시판에 올라온, 괴롭혀서 자퇴시키겠다는 내용들이었다. 이러한 내용은 혐오표현의 폐해라고 할 수 있는데, 소수자의 사회참여를 막아버린 것이다. 서로 싸우는 걸 보면, 도가 굉장히 지나친 표현들을 많이 쓴다. 그리고 그런 표현들이 결국 학생들이 건전한 사고를 하는 데 영향을 미친다. 누군가는 리딩(leading)하는 역할을 해줘야 하지 않겠냐는 생각을 했고, 또 그런 측면에서 카카오가 이슈를 제시한 것 같다. 민주주의 전체의 건전한 작동 방식을 돕는다는 목적으로 논의를 이끌면 어떻게 생각한다.

박재영 카카오에서 사람들이 많이 주목하는 것은 결국 뉴스 댓글이다. 세계 유수의 미디어들 중 어느 곳도 한국 자문위원 미디어처럼 댓글을 허용하고 있지 않다. 아예 댓글이 없는 뉴스 사이트도 많다. 타임즈는 댓글을 노출하고 있어도, 24시간 내에는 다 없어지게 만들고 10% 밖에 보여주지 않는다. 근본적인 의문은 “댓글이 우리 사회 민주주의에 얼마나 기여하고 있는가”이다. 댓글이 없어지면 표현의 자유가 무너지는지 혹은 댓글의 부정적인 기능으로 인해서 민주주의가 훼손되지는 않는지와 같은 근본적인 고민을 해야하지 않을까 생각한다.

혹은 표준어 검사 프로그램처럼 혐오표현을 찾아내는 프로그램을 만드는 것은 어떠한가? 1차적으로 혐오 표현을 잡아내는 것을 넘어서, 2차적으로 혐오의 맥락을 잡을 수 있는 프로그램을 만들어 뉴스 댓글을 다는 사람들이 검사 기능을 거치도록 하는 것이다. 이러한 프로그램은 국내에 없는데, 한번 카카오가 도전해 보는 것이 어떤가?

손정아 팩트체크나 증오표현에 대해서 기술적으로 접근하는 것도 카카오가 고려해볼 필요가 있다고 생각한다. 카카오 부사장 스를 통해서 차별적인 콘텐츠가 유통되는 부분에 대한 고민도 있다. 이용자들이 댓글을 쓰는 것도 문제지

만, 콘텐츠 자체가 차별을 조장할 수 있다는 것도 같이 고려해야 할지 고민이다.

박재영 미디어 콘텐츠, 뉴스 콘텐츠 내의 혐오 보도에 대한 연구보고서가 작년 한국언론진흥재단에서 나왔다. 아까 말씀드린 프로그램을 언론사 기자들도 사용하면 걸리지 않을까. 저런 표현 자체의 필요성도 있고, 과도한 규제도 문제가 된다. 하지만 저런 표현을 씌으로써 사람들은 저런 표현을 대체할 수 있는 좋은 표현들을 있고, 생각하려고 노력하지 않게 된다. 전체적인 언어 수준이 저하될 수밖에 없다. 나쁜 표현을 없애으로써 좋은 표현이 살아남고 활성화될 수 있도록 하는 기능과 역할을 하면 좋을 것이다.

김장현 갈수록 우리 사회에서 극단적인 표현이 급증하고 있다. 다양한 차원의 혐오표현들이 서로 결합하면서, 새로운 표현을 만들고 있다. 우리가 혐오표현이라는 이름으로, 포털에 그러한 표현들을 보여주지 않을 것을 강요하게 되는 순간, 또 다른 사람들은 사회적인 반성을 할 수 있는 기회를 박탈당할 수 있다. 그래서 표현을 사전에 막는 것은 최소화해야 한다고 생각한다. 사실 “민주주의 차원에서 댓글이 필요한가”와 같은 측면은 다시 근본적인 논의가 필요하고, 이는 학계에서 이끌어내야 한다고 생각한다. 혐오표현은 기준을 잡기 어렵다. 단어들의 온도, 감정의 온도는 금방 변화한다. 그래서 절충안으로, 어떤 표현을 보기 싫으면 가리는 정도가 이용자들의 개별적인 선호를 반영해주는 최대한이라고 생각하고, 이는 이미 실행 중이다.

이준웅 사람들이 선호하지 않는 표현들을 끊임 없이 쫓아다니며 막을 수는 있겠지만, 10년 뒤에 그것들이 없어질 것이라고 기대하지 말자는 것이 요점이다. 이런 규제가 차별을 줄이고 민주주의에 도움이 된다고 믿고 있는데, 정말 그러하지는 따져봐야 한다.

한편, 서비스적 관점에서 혐오표현을 규제하는 것이 중요한 일이라고 생각할 수 있지만, 그렇지 않을 수 있다. 카카오는 큰 사업자이고 사적영역에 속하지만, 공적으로 비취지는 역할들이 있다. 따라서 관련 규정을 정할 때, 서비스 품질, 사업 목적, 시민 사회의 일반적인 공존의 문제를 종합적으로 생각해야 할 일이다. 뉴욕타임즈는 댓글의 난이도를 조정할 수 있다. 100을 누르면 모든 욕이 다 나오고, 50으로 내리면 fuck이라는 단어는 나오지 않는다. 0으로 내리게 되면 모든 욕설들이 사라진다. 그렇다고 해도, 외국인들 때문에 미국 노동자들이 문제가 생겼다는 트럼프 지지자들의 댓글들을 지우지는 않는다. 왜냐하면 이 의견도 정당한 의견일 뿐만 아니라 이민정책을 반대하는 정치적인 주장이기 때문이다. 이런 식으로 영민하게 노력해야 한다. 특정 단어들을 하나하나 혐오표현인지 아닌지에 대해 정하는 것은 이상한 고민일 수 있다.

박재영 우리 사회는 혐오표현을 쓰는 사람은 누구인지, 또 얼마나 많은 이용자가 어느 정도 수준의 혐오표현을 쓰고 있는지를 모르고 있다. 혐오표현의 생산 측면도 연구 대상이지만, 사람들이 혐오표현을 어떻게 소비하는지가 중요하다. 혐오표현 소비의 영향이 있든 없든, 사회적으로 의미가 있다. 그래서 혐오표현의 생산과 소비에 대한 실질적인 연구를 하면 어떨까.

사업체 입장에서는 실험을 많이 할 것을 권하고 싶다. 뉴욕 타임즈가 지금 뉴스 구독 모델을 개발했다. 2006년 이후 거의 20년간 실험한 결과다. 카카오도 혐오표현을 완전히 닫아서 못하게 했을 때, 유저들의 금단 현상이 어떻게 발생하는지 봐야하지 않을까 한다. 금단현상이 없다면 정책을 바꿔볼 수도 있다. 혐오표현의 차단이나 규제가 정말 비즈니스에 영향을 주고 있는지도 과학적으로 검토해봐야 한다. 닫는

것이 비즈니스에 영향을 준다면, 두 가지로 대응할 수 있다. 다시 이런 표현을 허용하는 방법도 있겠지만, 훨씬 적극적으로 보면 그걸 단음으로써 새로운 비즈니스 전략을 갖게 된다. 뉴스로 인해 들어왔던 이용자들이 나가게 되면, 카카오 자력으로 사람들을 모아야 하고, 이를 위해 새로운 사업 아이템이 개발될 것이다. 여러 가능성이 있기 때문에 다양한 실험을 해서 내부적으로 실질 데이터를 가지고 검토를 해야 하고, 필요하다면 학계에서 참여하는 것도 필요하다.

김민정
자문위원 혐오표현 규제에 있어서 단어를 삭제하는 것은 욕설 규제와 달라지지 않는다. 단어 삭제보다 적극적으로, 혐오표현에 반대하는 입장을 표명하는 것이 중요하다. 카카오가 중요한 포털 사업자로서 카카오 플랫폼 공간 안에서 일어난 차별적인 발언들에 대해서 명확히 반대하는 입장을 보여야 한다. 하지만 그 방식이 꼭 삭제가 아니라, 캠페인이나 대항표현을 하는 단체에 대한 지원, 교육에 대한 투자, 이용자 선택이 가능하도록 하는 투자일 수도 있다.

또, 도준호 위원이 말한 사례가 시민단체 말을 들어야 하는 이유 중 하나다. 20대들이 대학교 커뮤니티인 에브리타임에서 행동하는 것과 카카오에서 행동하는 것은 피해 정도 등에서 완전히 양상이 다르다. 댓글보다 실시간 톡이나 에브리타임 같은 커뮤니티 공간에서 발생하는 피해는 심각성이 높다. 카카오가 직접적으로 서비스는 관리하지 않지만, 이러한 피해에 대해 사회적인 관심을 환기시킬 수 있을 것이다.

이준웅
자문위원 우리 사회가 ‘혐오표현’이라는 용어 자체를 검토 없이 수입한 것에도 이유가 있을 것이다. 사적 영역에서 규범을 정해서 발언의 수준을 정하는 일은, 사적 영역의 주체들이 하는 일이다. 사적 영역에서 일부는 전통에 의해서, 일부는 민주적으로 내부에서 규율을 채택하는 방식으로 규제한다. 이렇게 얼마든지 가능한데, 어떤 사람들은 법적으로 규제할 수 있다고 생각하는 것 같다.

카카오에서 이용자와의 약속을 통해서 발언의 자유를 통제하지 않는 방식으로 발언을 규제하는 것이 가능하다는 것을 보여줬으면 좋겠다. 카카오는 법적 처벌이 아닌 방식으로, 최소한의 개입으로 해당 문제에 접근할 수 있다.

양현서
카카오 부사장 단순히 단어 리스트를 만들어 규제를 하겠다는 것이 아니라, 우리 사회 안에서 혐오발언 중 무엇이 문제가 되는지, 누가 말하는지, 그리고 플랫폼은 어떤 방식으로 대응해야 하는지 등 고민을 나누기 위한 것이다.

이준웅
자문위원 모욕죄가 적용된다면, ‘중복’이라는 말을 쓸 수 없다. 이는 정치적 담론을 제약하는 일이다. 그런데 모욕죄를 적용하자는 국회 논의가 있다.

‘중복’은 2심에서는 명예훼손이라고 판결이 내려졌지만, 최종심인 대법원에서는 ‘중복’이라는 표현만으로는 명예훼손이 아니라는 판결이 났다.

김대기
카카오 부장 2014년 KISO에서 혐오표현에 대한 사례와 절차를 만들 때, ‘중교’라는 범주는 들어가지 않았다.

박재영
자문위원 선플을 더 많이 노출시키고, 선플을 달도록 유도하는 것도 하나의 방법일 것이다. 표현력이 좋은 소설가를 고용해서, 좋은 댓글들을 등록시켜 사람들이 보고 느끼도록 하는 것은 어떠한가?

양현서
카카오 부사장 플랫폼 사업자는 사회적 책임, 디지털 책임을 져야 한다. 그래서 연에 댓글도 달았는데, 사회적 담론을 제시하지는 못했다. 결과가 어떻든 이런 논의를 사회에 던져보고 싶다. 결과와 목표는 아직 잘 모르겠다.

박재영 연예 댓글 담은 지 이제 반년쯤 되었나? 혹시 내부적으로 결과에 대한 평가가 없는지 궁금하다.
자문위원

박용준 연예 댓글을 달았다고 연예 기사 트래픽이 내려가지는 않았다. 댓글을 이용해서 돈을 더 벌겠다는 것은 맞지 않는 주장이다. 물론 이용자들이 댓글이 없는 것에 대해서 불편을 느낀다. 카카오는 이용자들에게 뉴스 댓글 영역을 접할 수 있는 기회를 주었다. 이용자들이 처음에는 달았다가 결국 다시 펼친다. 그래서 아예 댓글을 달아버리는 것보다는, 어떤 식으로 바꿀 수 있을지에 대해서 고민을 하는 것이다.
카카오 이사 뉴스 자동화를 할 때 처음 하는 작업은 유통되어서는 안되는 콘텐츠를 걷어내는 것이다. 광고 뉴스나 너무 자극적인 어뷰징 콘텐츠를 걷어내고 더 중요한 뉴스들이 나오도록 한다.
댓글도 마찬가지로 일단 이용자 보호를 위해서 보여지지 않아야 하는 댓글들이 무엇인지에 대해 계속 얘기하고 있다. 서비스 제공자 입장에서는 기준이 필요한 상황이다. 댓글은 배열을 서비스 제공자가 전혀 건들지 않고, 이용자들이 추천을 눌러서 좋은 댓글이 올라가게 하라는 의도로 만들었다. 그런데 지금은 약간 팬덤화가 되었다. 내가 지지하는 쪽의 댓글이 올라가게 한다. 이러한 댓글의 배열을 바꿔보자는 것이 내부적인 과제이기도 하다. 그런데 댓글의 배열을 바꾸는 것은 이용자들에게 설득력이 있어야 한다.

도준호 숙대 트랜스젠더 이슈에서 느꼈던 부분은 결국, 여론을 형성할 때 센 표현이 대세가 된다는 것이다. 대다수 학생들이 원래부터 그런 생각을 하지는 않았을 것이다. 학생들 또한 여성이라는 소수자이고, 여러가지 차별에 대한 교육을 받았다. 몇몇 페미니스트들의 극단 표현으로 여론이 흘러가는 것을 보면 문제가 있다는 생각이 들었다. 익명게시판에 그렇게 험한 말이 나오고 극단적인 표현들이 그 학생에게 영향을 주었다는 생각에 굉장히 슬펐다.
자문위원

이준웅 혐오표현이란 말이 많이 쓴다. 또 음란, 국가보안, 사기, 협박을 포함하는 불법적 발언들, 불법은 아니지만 욕설, 비속어, 심한 풍자나 조롱을 포함하는 비불법과 비규범들, 우리는 이중에서 무엇이 혐오표현이고 아닌지 아직도 모른다. 그래서 모호한 말을 쓰는 것보다 어떤 책 제목인 ‘Extreme Speech’로 통칭하는 것이 적절하다. 극단적 발언이나 극단 발언이라고 쓸 수 있다. 우리 말로 그냥 ‘험한 말’ 정도도 좋을 것 같다. Hate speech를 ‘혐오표현’이라 옮긴 것은 오해를 유발한다는 점에서 잘못된 것이다. 혐오라는 것은 원인과 대상이 정해져 있는 굉장히 좁게 규정된 개념인데, 실제 사용은 매우 넓게 사용한다는 점에서 그렇다.
자문위원

김장현 혐오표현과 같은 이슈에서 포털이 선도적 역할을 할 필요는 없다. 일부는 자정되고, 일부는 악화되는 것이 민주 사회의 원리다. 포털이 혐오표현에 대한 블로킹(blocking)을 주요 임무로 자행하기 시작하면, 오류에 빠지기 쉽다. 어떤 정치인의 스캔들이 뜨면, 그 정치인을 풍자하는 이상한 표현들이 당일에 나온다.
자문위원

손정아 숙대 내에서 비판에 대해 궁금하다. 특정 게시글을 올리는 사람은 소수다. 그런데 트랜스젠더 학생의 입학 논의 과정에서 다양한 숙대 학생들이 참여하면서 토론을 거친 것인가?
카카오 부사장

도준호 아까 말했던 순서대로 글이 올라왔고 이 과정에서 논쟁이 붙었다. 그러던 몇몇 극단적인 페미니스트들의 목소리가 훨씬 커졌고 이에 다수가 동조를 하게 되었다.
자문위원

손정아 카카오 부사장 뉴스 댓글 같은 경우, 댓글이 몇 백 개 이상 달리게 되면 다양한 의견이 생겨 의견 간의 균형을 찾는다. 학교 내 게시판에 의견을 내는 사람이 소수여서, 한 쪽의 의견만이 관철되었는지 궁금했다. 과거 댓글 알바도 그러했지만 지금은 사람들에게 댓글 리터러시가 생긴 것 같다는 생각이 든다. 최근에 특정 기사의 댓글에서 기사 내에 팩트가 아닌 사실을 지적했고, 기자가 댓글에 직접 사과를 했다. 이와 같이 긍정적인 기능도 있다.

제2회 미디어자문위 회의

- 일시: 2020.6.18. 16:00~18:00
- 장소: 서울 중구
- 참석자
 - 카카오: 손정아 부사장, 김대원 이사, 이준목 이사, 박용준 이사, 김성환 부장, 김대기 부장, 김수원 연구원
 - 자문위원: 박재영 교수, 김민정 교수, 김장현 교수, 도준호 교수, 이준웅 교수

김대원 카카오 이사 혐오표현 가이드라인 정책은 two-track 방식으로 이뤄진다. 우선, 인권위와 언론법학회 공동연구단을 통해서 개념적, 학문적 측면을 다루려고 한다. 혐오표현의 개념, ‘혐오표현’ 대신 사용할 수 있는 용어 등 기본적인 사항부터 연구를 시작하고, 연구 결과가 좋으면 내년에 더 심화해서 진행하는 단계적 방향을 언론법학회에 전달했다. 9월에 세미나를 통해 연구 중간 결과를 발표할 예정이다.

다른 한 축은 자문위로, 초안을 마련하는 작업이다. 9월 둘째 주에 초안이 완성되면 전문가, 기자, 시민단체를 초대해 초안에 대한 의견을 받는 라운드테이블을 진행하고자 한다. 여기서 나온 의견과 연구 결과를 수렴해 12월에 가이드라인을 만들 예정이다.

인권위나 언론법학회에는 학문적인 접근보다 우리 사회 내 혐오표현이나 문제 표현에 대한 어떤 생각을 갖고 있는지 다양한 목소리를 수렴했으면 좋겠다는 의견을 전했다. 기존 연구자 뿐만 아니라, 다른 분야에서도 혐오표현에 대해 고민하고 있는 분들도 합류하면 좋겠다고 제안했다.

혐오표현 가이드라인 정책에 지역, 인종, 성별 등 어떤 축을 포함해야 할지에 대해서도 의견 내주시기를 바란다.

이준웅 자문위원 인권위 가이드라인에 19가지 차별금지 대상이 나열되어 있다. 여기에는 성별, 종교, 장애, 나이, 출신 지역, 출신 국가, 용모, 기혼, 전과, 성적 지향, 학력, 병력 등이다.

김대원 카카오 이사 인권위 가이드라인은 매우 세분화되어 있는데, 무엇을 채택하고 제외해야 하는지? 또 인권위 내용에는 들어가 있지 않지만, 포함되어야 할 것이 있는지?

김민정 인권위 2019년 혐오표현 리포트는 6가지를 언급하고 ‘등’을 붙였다. ‘성별, 장애, 나이, 출신지역, 인종, 성적 지향 등’이라고 되어 있다. 이를 준거법으로 잡아도 될 것 같다. ‘등’을 붙여서 항목을 다 열거하지 않아도, 수용이 가능하도록 하는 것이 좋을 것이다.

김성환 KISO에서는 ‘지역, 장애, 인종, 출신 국가, 성별, 직업 등’으로 규정했다. 인권위 기준은 ‘성적 지향’ 등 카카오 부장 KISO 보다 더 나아간 측면이 있다.

도준호 일반적인 원칙 중 하나는 본인의 의지로 선택하지 않은 것, 즉 인종, 성별, 출신 국가 같은 기준으로 차별을 해서는 안된다는 것이다.

박재영 이 때 중요한 것은, 적의적 표현이어야 한다는 것이다. ‘혐오’가 너무 포괄적이라면, ‘적의’가 개념에 더 근접한 것으로 보인다.
공동연구 주제는 학회에서 공모하는 것보다, 카카오에서 책임자를 찾아 연구 과제를 맡기는 것이 더 좋을 것이다. Naming의 개념적인 범주는 대단히 학술적인 내용이다.

김대원 인권위에서는 글로벌 리서치를 위해 연구자를 지정했다. 글로벌 기준과 개념에 대해 연구하게 될 것이다.
카카오 이사

이준웅 철학적, 법리적 논의보다는 카카오 서비스 고품질화를 위해서 문제 해결을 위한 프레임워크 마련하는 것이 중요하다. 자율 규제를 위한 원리, 기준, 주요 적용 사례 정도만 나와도 될 것 같다.

김대원 내용의 깊이에 대해서는 아직 논의되지 않았으나, 프레임워크를 만든 이후에도 계속 발전 되어야 한다는 것엔 공감하고 있다. 완벽한 가이드라인보다는, 향후 논의를 위한 토대를 만들고자 한다.
카카오 이사

도준호 그럼 이 때 카카오가 어느 정도 가치 판단을 하겠다는 뜻인가?
자문위원

김대원 가치판단보다는, 이준웅 위원 말처럼 카카오의 pain point 해결을 위한 것이다.
카카오 이사

도준호 계속해서 판단하는 문제들이 나올 텐데, 이는 카카오에게 부담이 될 수 있다.
자문위원

김대원 사회적으로 논란이 많을 수 있는 소재이기 때문에 속의 과정을 오래, 그리고 많이 거치고자 한다. 아직 국내에서 진행된 바가 없으므로, 이번 논의가 영향력이 있을 것이라 생각한다. 카카오는 논의의 출발점을 마련해야 한다는 책임도 느끼고 있다.
카카오 이사

김민정 로드맵에서 녹서를 마지막에 발표하는 것이 의미하다고 생각한다. 요즘은 결과물보다는 과정 자체에 집중한다. 최근 페이스북과 트위터의 사례를 보면, 사람들은 최종 결정보다 결정에 누가 참여했는지 그리고 결정을 도출하는 과정에 대한 설명을 요구했다. 논쟁적인 사안인 만큼, 과정에 대한 설명과 논의 자체에

집중하면서 진행되어야 한다.

손정아 세부적인 기준을 모두 세우는 것은 시간이 더 필요할 것이다. 다만, 카카오가 큰 플랫폼으로서 논의를 시작하고, 서비스를 개선하기 위한 고민의 출발을 한다는 것에서 의미가 있다고 생각한다.

이준웅 가이드라인 발표 이후에, 일부 단체 등에서 사회적인 반발이 있을 수 있다.
자문위원

박재영 공청회를 많이 해야 한다. 학자를 대상으로 한 연구 공청회 외에도, 일반 시민들 대상으로 하는 공청회를 많이 해서 의견을 들어야 한다.

김수원 인권위와의 공동연구를 발표하는 기사 댓글에, 코로나 사태에서 특정 종교나 지역의 잘못을 지적하는 것도 혐오나 차별이냐는 반발이 있었다. 이용자들의 행위를 혐오, 차별로 낙인 찍지 않으면서 정책을 세우는 것도 중요하다고 생각한다. 사람들이 배제당한다는 인식을 갖거나 위기의식을 느끼지 않을 수 있도록 정책을 세우는 데 도움을 주시면 좋겠다.

김장현 혐오표현은 댓글 신고 후속 조치에 있어서 중요한 기준이 될 수 있다. 댓글 신고에 대한 조치 작업을 하면서 애매한 영역이 있었을 것이다.

공동연구는 정기 조사 형식으로, 댓글 조치시 가장 결정이 어려웠던 표현들을 일반인과 전문가를 대상으로 한 설문을 통해 집중적으로 분석하고, 그 데이터를 축적하면 좋을 것 같다. 설문을 통해 데이터를 모아, 라벨을 만들어서 일관된 AI advisor 구축에 활용하면 좋을 것이다. 급격하게 변화하는 국민정서나 유행어를 반영하려면 너무 장기 데이터를 반영해도 안 될 것이다.

제3회 미디어자문위 회의

- 일시: 2020.8.25. 16:00~18:00
- 장소: 비대면 화상 회의
- 참석자
 - 카카오: 김대원 이사, 이준목 이사, 김성환 부장, 김대기 부장, 김수원 연구원
 - 자문위원: 박재영 교수, 김민정 교수, 김장현 교수, 도준호 교수, 이준웅 교수

※참고: 이번 회의에서는 [숙의 과정 3](#)의 내부적 분석 결과를 자문위에 공유한 후 그에 대한 의견을 받았다. 이하 회의록에는 발표 부분을 제외하고, 분석 결과에 대한 자문위원들의 의견만을 담았다.

박재영
자문위원

우선 용어를 정리하거나 통일시킬 필요가 있어 보인다. 우선 ‘혐오’와 ‘증오’에 대해서다. 정확한 용어는 ‘증오’지만, ‘혐오’가 일반적으로 사용되고 있다.

보호 대상 문제는 지역을 예시로 들겠다. 만약 어떤 지역을 보호 대상으로 지정해 놓았다 하더라도, 상황에 따라 그 대상의 범위가 달라질 수 있다. 법률이라면 꼭 대상이 있어야 한다. 하지만 이것은 법률이 아니니까, 유동적으로 둘 수 있지 않을까. 또한 혐오표현의 대상이 지속적으로 변화하기 때문에, 보호 대상을 선택적으로 지정하는 것은 불가능해 보인다. 여기서 보호 대상이 특정되어 있지 않는데 어떻게 규제하겠냐는 반론이 나올 수 있다. 중간지대를 확인해야 한다.

다음으로는, ‘소수자’에 대한 이야기다. 소수자는 결국 ‘사회경제적 힘’에 따라 결정이 된다. 시간이 변화함에 따라서 소수자의 개념 또한 변화하기 때문에, 현재 보호 대상을 지정하는 것은 범위를 협소하게 만드는 것이 아닐까. 그러므로 소수자라는 개념보다는 넓은 범위에서, 지위고하를 막론하고 누군가 피해를 보고 있다면 보호 대상으로 생각하는 것이 어떨지.

가이드라인을 보면, 다음과 카카오가 갖고있는 데이터에서 김수원 박사가 개념화해 귀납적으로 추출한 것이다. 인지, 태도, 행동은 이미 이론적으로 구축되어 있다. 인지에서 태도, 태도에서 행동으로 가면 점차 생각을 이야기하는 것부터 행동을 유발하는 것으로 높아지는 것이다. 3단계로 구분하셨는데 보통 차별적인 괴롭힘, 편견 조장, 모욕, 증오 및 선동 이렇게 4단계로 이야기를 많이들 하고 있다. 이처럼 이론과 틀을 매칭 시키는 것도 괜찮아 보인다. 실질적인 데이터들이 이론적으로 나타나는 것은 아니지만, 이론적인 토대를 가지고 표현에 대해 이름을 붙이면 좀 더 견고한 가이드라인이 될 것이다. 이처럼 이론적 연계성을 강화하는 쪽으로 제안해본다.

전체적으로는 좀 더 유동성 있고 융통성 있게 이슈를 바라보면 어떨까 생각한다.

마지막으로는, 이와 같은 혐오표현 가이드라인이 한국에서 처음이다. 그래서 굉장히 조심스럽게 하는 것이 좋을 것이라고 생각한다. 이 말은 한꺼번에 이루어내려고 하지 말고, 단계적으로 접근하라는 것이다. 김수원 박사가 분석했던 표현 대상은 9개로 나누어져 있는데, 선행 연구에 따르면 혐오표현 중 가장 많은 범주가 성별이고 또 우리 사회에서 성별 문제가 제일 큰 문제다. 남성과 여성이라는 것부터 시작해보면 어떨까, 생각한다. 예컨대 보호 대상, 규율해야 할 표현의 종류, 표현의 수위 및 대응 방안도 성별 케이스만 놓고 보면 정교화할 수 있겠다는 생각이 든다. 하지만 지역 같은 경우는 모호해진다. 경상도에 태어났지만 서울에서 산 사람인 저와 같은 경우만 보더라도 이 범주는 상당히 유동적이다. 성별로 시작해보고, 다른 범주로 넘어가보는 방식이 어떠한지 제안해본다.

나아가, 레이블링(이름짓기) 데이터로만 혐오표현 현황을 파악하는 데 한계가 있다고 말했는데, 사실 레이블링이 이 문제의 전부일 수도 있다. 레이블링만 잘 규율해도 많은 문제를 풀 수 있다고 생각하기 때문이다. 규율인가 아닌가, 이렇게 복잡하게 고민하기 보다는 ‘미디어’라는 차원을 생각하는 것이 좋을 것 같다. 미디어라는 것은 사회의 언어를 선도하고, 언어의 순화기능을 하는 측면이 있다. 카카오가 레이블링을 잘 잡으면, 사회적으로 큰 역할을 하는 것이 아닐까 생각한다.

김장현
자문위원

공격과 비하의 대상이 점점 다양해지는 상황에서, 적극적으로 보호해야할 대상을 선택적으로 지정하는 것은 어느정도 가능할 것이다. 따라서 내부적으로는 지정하되, 자주 점검하는 것이 필요하다. 왜냐하면 보호 집단의 종류는 사회적인 국면에 따라 달라지기 때문이다.

카카오의 경우에는 일상 생활의 플랫폼인 메신저를 서비스하고 있고 그 위에서 O2O(Online to Offline) 서비스를 하고 있다. 포털은 그 중 하나의 서비스이다. 우리 일상생활에서 오프라인의 영역들이 온라인으로 포함되어 분리 불가능해지고 있는 상황에서, 언제까지 포털이 모든 것을 필터링해야 할지에 대해서는 회의감이 들기도 한다.

김민정
자문위원

첫 번째로, 카카오의 종국적인 정책 목표가 무엇인지 명확화하는 것이 필요하다. 가이드라인 초안을 보면, ‘디지털 건강’, ‘보편적 서비스’, ‘모두에게 안전하고 편리한 서비스’라는 용어를 사용하고 있다. 만약 댓글과 게시물 규제의 궁극적 목표가 사회적 갈등의 완화, 사회통합 이런 맥락이라면, ‘극단적인 공격 표현’에 대해 규제를 하는 것이 필요하다.

그러나 극단적 표현을 규제하는 것 뿐만 아니라, hate speech와 관련해서 ‘소수자의 인권보호’가 목표라면, 같은 욕을 해도 사회적으로 권력이 있는 남성에게 욕을 하는 것과 사회적으로 배제되고 차별 받은 소수자 집단에 대해 욕을 하는 효과가 다르기 때문에 보호 대상 집단을 면밀히 설정하는 것이 맞다고 본다. “남성을 혐오표현으로부터 보호해서는 안된다” 이런 뜻은 아니고, 남성에게 가해지는 표현 중 극단적 공격 표현만 규제한다고 했을 때에 여성에게 가해지는 표현에는 가부장제 및 성차별 이데올로기를 반영하여 폭 넓게 규제를 적용하는 방식을 생각해본다. 더욱이 성소수자나 이주민에 관한 것 또한 좀 더 폭넓게 고민해볼 수 있지 않을까.

만약 보호 대상 집단 설정에 집중한다면, 내부적으로 설정하는 것이 좋을 것이다. 소수자 집단에 대한 혐오표현에 집중하며 그 집단의 판별 기준을 이용자들에게 알리는 것이 좋을 것 같다. 뿐만 아니라 이와 같은 보호 대상 집단은 시간에 따라 변화할 것이며, 특정 사건이 발생할 때마다 변하기 때문에 주기적으로 업데이트하고 검토하며 고민을 해야할 것이다.

사회적 소수자 개념에 대해, 사회과학 분야에서 논의된 부분을 찾아보았다. 소수자는 숫자적 개념이 아니다. 숫자는 다수라고 해도 소수자일 수 있다. 예를 들면 백인의 인구수는 비교적 소수이지만 여러 나라에서 주류 집단에 속한다. 소수자는 이러한 개념이기 때문에 소수자의 특성과 개념을 구성하는 방향으로, “카카오는 이러한 개념을 적용하여 사회적 소수자에 대한 혐오표현을 중점적으로 규제하고자 합니다”라고 밝히면 될 것 같다. 앞서 말했듯, ‘사회적 통합’ 혹은 ‘소수자 인권보호’ 중 정책적 목표가 정해지면, 극단적 표현에 대해서만 규제를 해야할지 또는 소수자 입장에서 폭넓은 규제를 할 수 있을지를 정할 수 있다고 생각한다.

나아가 인식, 행동을 구분한 것에 대해서 이론적 토대가 단단했으면 좋겠다는 생각을 한다. 박재영 교수님이 말씀하신 부분인데, 네 가지 유형은 인권위원회에서 정해두기도 했기에 이를 참고해도 될 것 같다. 혹은 사회심리학 분야에서 사회심리학 분야 저서로 고든 올포트의 <편견>이 있는데, 여기서 태도와 신념에 대해 나눈 것이 나온다. 이처럼 인식, 태도, 신념 등 세분화를 하여 구분을 하면 규제를 명확화하고 설명하기에 좋을 것이다.

계속해서 게시물, 댓글을 삭제한다는 말이 머리 속에 남는다. 삭제만 하거나 삭제 중심으로 접근해서는 안 된다. 정말 극단적인 게시물과 댓글은 삭제할 필요가 있겠지만, 그렇지 않은 경우에는 삭제가 아닌 방식으로 대응할 수 있을 것 같다. 예를 들면, 해당 게시물과 댓글에 대해 추가적 정보를 알린다든가 캠페인을 보여주는 방식으로, 좀 더 형성적 방식의 규제를 하는 방향이 좋을 것 같다.

또한 서비스 별로 다른 방식으로 분할하여 규제하는 것이 좋을 것 같다. 카카오톡과 같은 메신저에는 규제를 적용하지 않고, 댓글, 게시글에는 하는 방식을 생각했다. 사적인 서비스, 공적인 서비스는 다르게 하는 것이 맞다고 본다.

마지막으로 극단적 공격 표현은 혐오표현의 문제로 볼 필요가 없다고 본다. 그래서 극단적 공격 표현은 지금이라도 적용할 수 있다고 생각하고, 그리고 이미 어느정도 욕설과 관련해 적용되고 있다고 본다.

이준웅 자문위원

다음과 카카오에서 제공하는 서비스가 공적인 영역인지 혹은 사적인 계약의 규율 영역에 들어가는지가 가장 중요한 요점이라고 생각한다.

예컨대, 동창회나 학회 같은 사적 공간에서는 구성원들의 자발적인 규칙에 의해, 다시 말하면 내부 민주주의를 통해 얼마든지 표현들이 규제 가능하다. 동창회에서는 “우리 ‘흡연총’이라는 말을 쓰지 말자”라고 규정하면 된다. 하지만 여기서 문제는, 이러한 규정을 공적 영역에서, 예를 들면 공원에서, 타인을 규제할 수 없다는 것이다. 이것이 자유민주주의의 원리이다. 왜냐하면 공적 영역의 목적은 공개적인 발언을 통해 여론을 형성하는 것이다. 하지만 공적 영역에서의 발언을 마음에 들지 않는다는 이유로 규제하는 것은 문제가 된다.

문제는 “다음과 카카오가 제공하는 서비스들이 어느 영역에 속하는가”이다. 다음과 카카오가 제공하는 서비스는 굉장히 다양하다. 다음 카페와 같은 커뮤니티들에서는 이미 사적으로 발언을 규제하는 자체적인 약관을 커뮤니티별로 적용하고 있다. 이것이 굉장히 중요한 경험이다. 내가 자주 가는 사이트의 경우에는 정치 이야기가 금지다. “어떻게 사이트가 특정 주제에 대한 이야기를 금지시킬 수 있는가”라고 생각할 수 있지만, 그 사이트는 사적인 취미 모임이니 금지시킬 수 있다. 하지만 이 사이트에서 정치 이야기를 원천적으로 금지시키는 것이 아니라, 다른 별도의 정치 토론 게시판을 만들어 운영하고 있다. 이처럼 커뮤니티가 사적으로 규율을 만들어 지키는 훌륭한 경험들이 우리 사회에 이미 많다. 이 경험들을 바탕으로, 카카오톡 또한 사적 공간에 대한 규율을 생각해보면 될 것이다.

다만 공적 공간이 문제가 된다. 다음과 카카오 서비스 중 공적 공간은 ‘뉴스 댓글’이다. 만약 뉴스 댓글을 언론사가 운영한다면, 이는 언론사가 규정을 적용하면 된다. 하지만 다음과 카카오가 운영한다면, 다음과 카카오의 상업적인 이익 창출 공간이라고 볼 수도 있기에 100% 공적 공간이라고는 말할 수는 없지만, 우리나라에서 이런 댓글 공간이 네이버와 다음 두개 밖에 없으며, 많은 사람들이 네이버와 다음에서 뉴스를 이용하고 있다. 따라서 댓글 공간에서 사회적이거나 정치적 발언에 대해선 폭넓게 자유를 보호하되, 몇몇 상업적인 활동은 규율할 수 있을 것이다. 댓글 공간을 포함한 일련의 공공성이 강한 공간에서는, 법률로 정하지 않는 한, 서비스 제공자가 사적 영역에서 만들어진 특정 이론과 같은 학문적 담론을 들어서 보편적인 규제를 시행하긴 어려울 것 같다고 생각한다.

그렇다고 해도 우리가 할 역할이 없는 것은 아니다. 우리나라는 도덕적인 품위가 중요한 나라다. 저도 욕을 하기도 하지만, 공공장소에는 욕을 안하고, 할 수도 없다. 우리나라는 욕설에 대해서 사회적으로 ‘품위’에 의해 규제가 되고 있다고 볼 수 있다. 상스러운 욕이나 불법 정보, 인권 침해 행위, 현행 법을 위반하는 내용들은 이미 임시조치에 의해 규제가 되고 있고, 품위와 관련해서도 규제가 된다고 본다.

증오와 선동 발언에 대한 규제는 법률이 만들어지지 않는 한 공적인 서비스에서 별개로 규제를 적용하기 힘들다. 그렇다고 해도 서비스 제공자가 서비스 공간에 아무런 제재를 가하지 않는 것 또한 시민 사회를

운영하는 책임자로서 온당한 태도는 아니라고 생각한다. 사람들의 소통의 역량을 증진시키는 방향을 생각해볼 수 있다. 예를 들면 어떤 이슈에 관해 토론을 더 만들어 낸다든지, 훌륭한 사람들의 토론문을 링크로 만들어 놓는다든지 등. 토론문들을 통해 사람들이 참조도 하고, 인용을 해서 댓글에서 논의를 하게 된다면 더 많은 토론을 이끌어내고 스스로 배워가는 계기가 될 것이다.

또한 서비스와 기술적 방식으로 이를 해결할 수 있는 방안이 틀림없이 있다. 댓글 창에 문제가 되는 표현의 정도를 이용자가 결정할 수 있는 슬라이더를 만든다든가, 성인인증 및 본인인증한 사람들에게 별도의 공간을 마련해 격렬한 토론의 공간을 줄 수 있는 방식도 있다. 물론 이용자에게 개인정보를 노출하라고 하는 것은 위헌 결정이 났지만, 자발적으로 자신이 자신의 정보를 제공하고 들어와 서비스를 이용하는 것은 문제가 없다. 그러나 여러 문제가 되는 정치적 발언을 사용하며 논쟁을 하고 싶으면, 해당 공간을 이용하도록 하는 게시판 등의 공간을 만들 수 있다.

보호 대상 범주 질문과 관련해서는, 혐오표현에 대해 집단 범주 별로 다르게 적용할 수 없다고 생각한다. 권리의 구제 관련한 문제에 있어서는, 집단의 표식으로 가를 수 없다. 누구든 피해를 볼 수 있고, 또 피해를 보았다면 법적인 보호를 요청할 수 있다. 인권은 모두에게 평등하다.

하지만 사회에 분할된 자원을 나눌 때에는 이미 불균형한 상태에 따라 나누게 된다. 따라서 차등적으로 사회적 약자를 보호할 수 있는데, 예를 들자면 가난한 사람에게 많은 이득이 가게 사회적 자본을 나눌 수 있다는 것이다. 이는 롤즈의 정의론에 나오는 ‘차등적 원칙’이다. 이에 따르면 여성이나 남성을 욕하는 것 모두 인권침해이고 해서는 안되는 것이지만, 자원 배분 문제에 있어서는 다르다. 남성과 여성에게 자신의 권리 주장을 위해 발언권을 주고 시간을 나누어준다고 할 때에는, 국가이든 서비스 제공자이든 기득권을 가지고 있는 남성에 비해 여성에게 시간을 더 분배할 수 있다. 이것은 기회의 문제이고, 사회적 약자에 대한 정의의 문제이다. ‘인권 보호’의 문제인지, ‘자원 배분’의 문제인지 생각해야 한다.

하지만 이미 다음과 카카오에서 욕설, 성적 괴롭힘 규제를 하고 있기에 이 부분은 많이 고민하지 않아도 될 것 같다. 물론 사회적 약자를 위해 더 많은 기회를 주고 그들의 공간을 마련하는 것은 충분히 형성적인 서비스 차원에서 가능할 것이다. 이와 같은 정책을 도입하는 것은 ‘차등적 원칙’에 따라 정의를 만들어 가는 일이며 카카오가 사회적 기업으로써, 또 시민 사회에 기여하는 서비스 제공자로서 할 수 있다고 생각한다. 이는 바람직하며, 정의로운 일이다.

나아가 정치적 발언에 대해 강조해서 말씀드리겠다. 극단적 발언 중에서 선동적인 발언, 혹은 즉각적인 폭력을 유발하는 발언 즉, 명백하게 현존하는 위협의 원리 또는 즉각적 불법행위 원리 등은 규제가능하다는 것은 이미 세계적으로 받아들여지고 있다. 선동 발언은, 즉각적 불법행위가 나올 수 있기에 이러한 발언은 모든 나라가 규제를 하고 있다. 그러나 정치적 발언을 규제하기 시작한다면, 이는 전체주의가 될 수 있는 위협에 빠진다. 따라서 정치적 발언에 있어서는 폭넓은 자유를 보장하는 것이 맞다고 생각한다.

**도준호
자문위원**

혐오표현 관련 보호 대상을 정하는 것, 보호 대상을 정하는 판단 기준에 대해 사회적 합의를 이루는 것은 어렵다. 보호 대상 설명에 “역사적으로 차별을 받고 있다”는 내용이 들어가는데, 이 부분을 모두가 합의할 수 있는 내용으로 발전시키기는 매우 어렵다는 말이다. 보호 대상 문제는 숫자 문제가 아니라 권력의 문제다. 예를 들어 기독교와 같은 경우는, 종파도 많고 우리나라에서 기독교의 권력이 약하지 않다 보니, 기독교를 혐오의 보호 대상으로 정해서 실행하기에는 현실적으로 어려워 보인다.

보호 대상 문제와 별개로 결국 표현의 자유를 어디까지 허용할 것인가의 문제가 된다. 댓글은 여론 형성에 중요한 요소로 작동한다. 기사 뿐만 아니라 댓글을 통해 사람들은 기사가 논점을 제대로 잡고 있는지 판단하기도 한다. 정확하게 규제 및 제재의 단계를 구분해서 적용하는 것이 쉽지 않다. 결국 시도가 가능한 방법은, 표현에 대해 대응하는 것이다. 표현의 유해성에 초점을 맞춰야 한다. 혐오표현은 명예훼손이나 모욕 처럼 특정 대상을 두고 이야기하는 것이 아니지만, 명예훼손과 모욕죄를 구성하는 요건이나 표현을 참조해도 좋을 것이다.

가이드라인에서 이론적으로 표현을 세 단계를 나눈 부분을 보면, ‘행동’에 대한 규제가 가장 가능성 있어 보인다. 인식과 의견보다는, 행동이 직접적으로 표현된 부분들이 분명하다. 그러나 앞서 말했듯 표현의 수위에 초점을 맞춘다고 하면 인식이나 의견을 나타내는 글에도 문제가 있을 수 있다. 이론적으로 세 단계로 나누는 것은 이해 가지만, 결국 초점을 ‘표현의 수위’에 두는 것이 가장 가능성 있어 보인다.

결론적으로 소수자 집단을 보호하자는 의도와 취지는 이해하지만, 이에 대해 적극적인 개입은 상당한 위험일 수 있다고 생각한다. 단계적으로 접근할 필요가 있다. 표현의 수위부터 시도해보고, 점차 토대를 쌓아 논의를 거쳐 발전시켜야 한다고 생각한다.

시민사회 라운드테이블

- 일시: 2020.10.14. 15:00~17:00
- 장소: 서울 중구
- 참석자
 - 카카오: 양현서 부사장, 김대원 이사, 이준목 이사, 김성환 부장, 김대기 부장, 김수원 연구원
 - 자문위원: 박재영 교수, 김민정 교수, 김장현 교수, 도준호 교수, 이준웅 교수
 - 시민사회 전문가: 김용구 한국장애인인권포럼 소장, 김지학 한국다양성연구소 소장, 박록삼 한국기자협회 부회장, 손지원 오픈넷 변호사, 이윤소 한국여성민우회 성평등미디어팀장, 이완 아시아인권문화연대 대표, 임동준 민주언론시민연합 활동가, 조소담 닷페이스 대표, 조진경 십대여성인권센터 대표

박재영 카카오라는 사기업이 사회적인 역할을 다하는 것이 상당히 의미 있는 일이라고 생각한다. 오늘 시민사회 자문위원 전문가분들이 시민과 사회의 입장에서 좋은 말씀들을 해주시기를 바란다.

김용구 소장 장애인들은 집단의식도 권리 보유자로서의 의식도 매우 강하다. 그러다 보니 혐오표현 이전에 장애인을 이질적으로 표현하는 것에도 민감하게 반응한다. ‘일반인’이나 ‘정상인’이란 말을 장애인에 대비되는 용어로 사용하는 관행에도 “그럼 장애인은 ‘비정상’이라는 의미인가?” 하고 짚고 넘어가는 식이다. 잘못된

용어가 반복적으로 쓰이다 보면 왜곡된 이미지를 만들고 편견과 혐오를 넘어, 차별과 증오 범죄로 이어지기 때문이다. 하지만 우리사회 모든 소수자 그룹이 이처럼 조직화되거나 권리의식이 강한 것은 아니다. 국가가 혐오 표현을 규제해야 할 이유다. 표현의 자유를 억압한다는 반론도 있지만 국가의 인권보호 의무라는 큰 틀에서는 다수가 공감하고 있는 것 같다. “그럼 민간 기업인 카카오가 인권 정책을 통해 혐오 표현을 이유로 이용자를 제재하는 것은 정당한가?”라는 카카오의 고민이 있는 줄 안다. 저는 조직 외부인에 대한 카카오의 제재 권한이 아닌, 카카오의 인권존중 책임으로 접근하길 제안한다.

이용자가 카카오 서비스를 매개로 악의적이고 의도적인 혐오 표현을 반복하면서 특정 소수자 그룹에 대한 부정적 인식과 태도를 확산시키거나 정신적 고통을 가하는 경우, 카카오는 아무런 책임이 없는 걸까. 책임이 있다면 무엇이고 이유는 무엇일까. 국제연합(UN)은 ‘기업과 인권에 관한 이행 원칙(UNGPs)’에서 인권침해를 직접 유발하거나 기여한 경우는 물론, 상품과 서비스가 인권침해에 관계된 경우에도 영향력 범위 안에서 필요한 조치를 취할 것을 기업에게 요구하고 있다. 예컨대, 제3자에 의한 인권침해 예방 조치를 취하거나 피해자가 발생할 경우 피해구제에 참여 또는 협력하는 것이다. 따라서 카카오가 서비스 참여자에 의한 혐오 표현 예방 의지를 공개적으로 표명하고, 정책을 준수하지 않는 이용자의 서비스 이용을 제약하는 등의 조치는 사회적 영향력을 보유한 책임 있는 사회구성원으로서 취해야 할 당연한 행동이라 할 수 있다.

이와 같은 카카오의 정책과 조치가 사회적 지지 속에서 이행되려면, 우선 ‘악의성’, ‘의도성’, ‘반복성’ 등을 고려해 제재 대상을 선정하고, 이용자 제재도 필요한 최소한의 범위와 기간 내에서 이뤄져야 한다. 또한 제재 이유를 명확하게 설명할 수 있어야 할 것이다. 반면, 카카오 정책이 되도록 많은 이해관계자와 이슈를 포괄하는 방향으로 수립이 되었으면 한다. 직접적이고 위협적인 증오 발언은 물론 간접적인 장애인 차별행위, 차별을 조장하는 광고 행위도 이미 법적인 제재 범위에 있다는 점에서 너무 협소하게 정책을 적용할 경우 실효성이 떨어지기 때문이다. 따라서 모든 사회적 약자, 소수자 분야 관련 혐오 발언 내용과 이미지를 왜곡하거나 편견과 차별을 조장하는 발언 수준까지 포괄할 수 있도록 정책을 설계했으면 좋겠다.

**김지학 소장
한국다양성연구소**

장애 용어에 대해서는 사회적인 공감대가 형성되어 있는데, 이는 장애인 차별에 대한 문제 의식이 이미 형성되어 있기 때문이다. 하지만 우리 사회가 차별 발언에 대해 문제의식을 가지지 않는 부분들, 인종차별, 성 소수자, 종교에 대한 비하 표현 등 또한 많다.

혐오표현은 정의가 매우 중요하다. 사회적 소수자에 대한 고정관념과 편견을 강화하고, 차별과 억압을 만드는 표현이 혐오표현이다. 혐오표현에 대한 정의를 제대로 내리고, 이에 따라 원칙과 제도를 만드는 것이 중요하다. 사회적 소수자를 어디까지 명시하는지와 관련해서는, ‘여성, 성소수자, 장애인, 이주민’ 등 소수자를 나열하고 ‘등’을 포함하여 범위를 넓힐 수 있도록 하는 것이 중요하다. 하지만 논란을 피하고자 ‘성 소수자’를 빼는 것은 더욱 큰 문제가 될 것이다. 또 ‘기레기, 한남충, 개독’은 사회적 소수자에 해당하지 않으므로 혐오표현의 대상이 아니다.

혐오표현에 대한 제재를 사기업이 하는 것은 바람직하다. 원래 국가 기구가 해야하나, 사회적 합의를 이유로 하지 못하고 있다. 국가가 역할을 아직 다하지 못하고 있고, 또 국가가 하지 못하는 부분에 대해 언론과 미디어가 역할을 해야 하나 못하고 있다. 카카오가 혐오표현에 대해 규제를 하는 것은 매우 바람직하다. 네이트나 네이버의 경우, 댓글 이력을 조회할 수 있도록 한다. 카카오는 이런 조치와 함께, 혐오표현과 소

수자에 대한 정의, 위반 시 처리·제재 원칙을 마련하면 좋을 것이다. 최근 프랑스와 독일은 플랫폼에 성 착취물, 홀로코스트 관련 게시물에 대해 막대한 책임을 지우는 법안도 마련했다.

혐오표현 제재는 표현의 자유와 대치되지 않으며, 오히려 표현의 자유를 향상시킨다. 혐오표현은 소수자들의 자기 표현을 막는다. 모두가 자기 표현을 할 수 있는 사회를 위해서는, 혐오표현이 제재되어야 한다. 혐오표현 제재는 표현의 자유와 기본권을 향상시키는 것이지, 자유를 제재하는 것은 아니다.

하지만 ‘공격하여 구분 짓는 행위’라는 표현에서, ‘공격하는’이란 워딩은 위험할 수 있다. 요즘 혐오표현은 누군가를 공격하는 것 외에도 피해자 포지셔닝을 하며 이뤄지는 경우도 있다. 예를 들어, 일부는 스스로를 페미니스트라고 하면서도 '트랜스젠더는 여성의 안전을 침해한다'는 식으로 혐오를 나타내는 경우도 있다. 때문에 ‘공격’이라는 표현은 이런 경우를 포섭하지 못할 수 있다.

박록삼 부회장 한국기자협회

기자협회는 국가인권위원회와 함께 인권보도준칙을 만들어, 소수자 보도 시 주의사항에 대해 구체적으로 정해 두었다. 특히 성폭력 보도는, 구체적으로 지양해야 할 표현이나 용어들을 정의했다. 혐오표현 관련해서도, 여러 언론 단체들이 미디어 실천 선언을 하기도 했다. 구속력이나 실효성이 있는 것은 아니고 준칙을 준수할 것을 권고한다는 수준에 머물러 있지만, 우리 사회가 발전하는 만큼 기자들도 최대한 지키려고 노력하고 있다.

카카오가 이런 작업을 진행하는 것은, 기사 같은 콘텐츠보다도 콘텐츠에 달리는 이용자들의 댓글을 대상으로 하는 것 같다. 모욕하려는 의도와 상관 없이, 사회적 약자들에 대해 일상적으로 쓰이는 혐오표현들이 많기 때문에 더 구체적이고 강력한 조치가 필요하다. 따라서 카카오가 사기업으로서 이를 하는 것에 대한 부담감은 느끼지 않아도 될 것이다. 기자와 언론들이 미디어 실천 선언 등 약속을 하는 것처럼, 카카오도 외형적으로는 뉴스편집 등 언론의 역할을 하고 있는 만큼, 범 미디어 생태계의 일원으로서 책무와 공공적 역할에 대한 부분을 사명감 있게, 구체적으로 하는 것이 중요할 것으로 생각한다. 또, 카카오가 단순히 기계적인 기준을 가지고 진행할 것인지, 혹은 자율적 심의기구를 만들어서 진행할 것인지 궁금하다.

손지원 변호사 오픈넷

혐오표현에 대해 사기업 차원에서, 건전한 공론장을 만들기 위한 노력과 자율규제에는 당연히 찬성한다. 저는 혐오표현에 반대하며, 혐오표현이 소수자 집단에 주는 심리적 타격감에도 공감하고 있다. 하지만, 기본적으로 표현의 자유라는 것이 올바른 표현을 위한 것이 아니라, 우리가 싫어하는 생각을 위한 자유일 수도 있다. 나쁘고 옳지 못한 표현은 규제하고, 좋은 표현만 공론장에 있어야 하는 것은 아니다. 왜냐하면 잘못된 표현도 올바른 표현도 일정한 가치와 기능을 가지고 있기 때문이다. 잘못된 의견을 반박하고 논쟁을 촉적하는 과정에서 사회적 논의가 풍성해진다. 논쟁에서 생각의 잘못을 지적하는 기회가 될 수 있고, 논의를 통해 더 올바른 방향을 형성해 나가는 역할도 있으므로 표현의 자유가 보장되어야 한다. 따라서 혐오표현을 많이 규제할수록 좋다는 입장은 염려스럽다. 나쁜 사상은 늘 존재하며, 이를 보이지 않도록 막는 것은 그런 사상이 존재하고 있다는 사실을 간과하게 만들고, 적절한 대응의 기회를 박탈할 수 있다. 표현의 가치는 표현의 내용이나 화자의 의도만으로 결정되는 것이 아니다. 표현 활동 그 자체가 대중의 알 권리의 대상이다. 이를 고려할 때, 소수자 집단이 모욕감을 느낄 수 있고 심리적으로 위축될 수 있다는 가능성을 이유로 무조건적으로 표현을 포괄적으로 금지하고 차단하는 것보다는, 해악이 중대하고 명백한 혐오표현에 대해서만 규제해야 한다. 공론장은 시민이 만드는 것이고, 사기업의 검열을 최소화해야 한다고 생각한다. 해악이 심각한 표현에 한해서 규제해야 한다.

그렇다면, 어떤 범위의 혐오표현을 규제해야 하나? 혐오표현 규제의 목적은 표적 집단에 대해서 나타나는 차별, 폭력, 배제, 그리고 공론장에서 표적 집단의 표현이나 영향력 위축을 막기 위함이다. 그래서 혐오표현은 표적 집단에 실제로 불안이나 공포, 위축을 불러일으키고, 사회에서 실질적인 배제나 폭력을 일으킬 수 있는 위험을 가진 표현에 한정되어야 한다. 그래서 표적 집단도 사회적 소수자에 한정할 필요가 있다. 사회적 소수자의 예를 열거하는 것은 좋은 아이디어라고 생각한다. 이 역시 국가기관의 판단기준으로서 추상적일 수 있으나 자율규제이므로 엄격한 명확성을 요구하지 않아도 될 것이다. 사회에서 주류적 지위를 차지하거나, 사회에서 공포나 위축 위험이 없는 집단에 대한 표현마저 제재하면, 기득권에 대한 불만을 미러링, 패러디 등으로 표현하는 사회적 소수자들의 표현을 억압하는 것이 될 수 있다. 표현의 규제가 너무 포괄적으로 집단에 대한 부정적인 감정을 표현하는 것을 대상으로 삼을 수는 없기 때문에, 모욕감 같은 추상적이고 포괄적인 개념이 아니라 선동, 차별 등 구체적인 허위 정보를 언급하며 사회적 차별 배제를 정당화하거나 조장하는 것을 규제해야 한다. 제주 난민들이 성폭력을 저질렀다든가, 동성애자들의 질병 등과 같이 소수자에 대한 구체적 허위정보로 사회적 차별을 조장하는 것들이 규제되어야 한다.

이윤소 팀장
한국여성민우회

카카오 안에 다양한 서비스들은 각기 다르게 다뤄져야 하므로, 세부적인 규정이 필요하다. 카페에서 사인(私人)이 쓴 글을 규제하는 것과, 다음 뉴스에서 제공하는 글을 규제하는 것은 다른 방식으로 이야기되어야 한다. 이와 같이 상세한 규정이 논의되고 있는지 궁금하다. 또 규제는 무엇이라고 생각하는지도 궁금하다. 개인이 아이디를 사용할 수 없도록 하는 것이 규제인지, 인터넷 언론사에서 혐오표현 사용시 다음 뉴스에서 배제한다는 차원 인지에 따라 논의가 달라질 것이다.

또 제재 규칙들이 생겨났을 때 어떻게 처리할지도 중요한 문제이다. 유튜브의 경우, 민원을 넣었을 때 기술적으로 조치하는데, 이로 인해 불균형 규제 문제가 생기고 있다. 이런 문제는 어떻게 해결할 것인지, 또 규제하는 알고리즘은 어떻게 만들지 등 깊은 문제까지 논의되어야 한다.

자율적 규제는 필요하다. 혐오표현을 정의하는 것이 중요하지만, 혐오표현을 이용자들에게 이해시키기 위해 노력하는 것이 더 필요하다. “나는 여자를 사랑하는데 내가 여성 혐오표현을 사용한다고?” 라고 이야기할 정도로, 우리 사회는 서로 혐오표현에 대한 이해 수준이 천차만별로 다르다. 혐오표현을 정의하는 것보다는, 사례 등을 통해서 어떤 것들이 왜 혐오표현인지 이해시키는 것이 더 중요한 일이다. 또 보호 대상이라는 것을 정할 수는 없다. 기레기, 한남충, 개독이 혐오표현이 아니라고 정확히 말할 수 있는지 고민이 생긴다. 이 단어들을 계속해서 사용하는 것이 괜찮은가?

맥락을 고려하지 않고, 혐오표현에 해당되는 단어를 규정하는 것만은 중요하지 않다고 생각한다. 가이드라인을 만들고 나서 어떻게 이용하고, 이용자들이 어떻게 읽게 할 것인지도 더 많이 논의해야 한다. 유튜브 가이드라인에도 자세하고 좋은 내용이 많지만, 실제로 지켜지고 있지 않다. 유튜브에 첫 영상을 올리기 전 가이드라인을 읽었는지 확인하거나, 교육내용을 담은 영상을 제공하는 등 최소한의 절차를 마련하는 것이 중요하다. 교육 차원에서 더 고민하면 좋겠다.

이완 대표
아시아인권문화연대

한국기자협회가 국가인권위원회와 함께 진행한 ‘혐오표현 반대 미디어 가이드라인’의 경우, ‘반대’라는 용어에 문제의 소지가 있다. 혐오표현은 찬성과 반대의 문제가 아니다. 단어 선택이 매우 중요하다. 카카오도 혐오표현 정책이라는 표현을 사용하지만, 혐오표현을 어떻게 할지에 대한 의도가 명확히 드러나도록 ‘혐오표현 근절·방지 정책’이란 표현을 사용하면 좋겠다. 국제인권법도 근절에서 나아가 ‘방지’라는 단어

로 더 강력하게 표현하고 있다.

이용자들 중 늘상 혐오표현을 내뱉는 사람도 있지만, 본인도 모르게 하는 경우도 많다. 혐오표현이 왜 나쁜지, 왜 규제되어야 하는지 알려주는 것도 매우 중요하다. 가이드라인을 만드는 것도 중요하지만, 카카오가 혐오표현 근절을 위해 어떻게 정책을 펼칠지 천명하고, 사람들이 접근하기 쉽도록 하는 '무언가'가 있으면 좋겠다.

혐오표현을 정의하는 것은 어렵다. 국가인권위원회도 이런 이유에서 혐오표현 가이드라인이 아닌 '리포트'를 발간했다. 맥락에 따라 혐오표현은 달라진다. 대신 버전을 여러 개 만들면 된다. 모두가 합의할 수 있는 버전을 만든 후, 점점 논의를 통해 발전시키면 된다.

나아가 이용자가 사용하는 혐오표현이 아니라, 카카오가 영향을 미칠 수 있는 혐오표현에 대해서도 주의해야 한다. 이모티콘 등의 카카오 서비스 전 분야에 혐오표현을 확산·양성·조장하지 않는지 관리 및 감독하는 기구나 전담 부서가 없다. 이용자에게 혐오표현을 근절하라고 하면서도, 카카오가 어떻게 조심하고 있는지에 대해서는 이야기한 적이 없다. “내부 문서에 혐오표현이 있는지 살피고 있다” 등 카카오가 내부적으로도 열심히 하고 있다는 것을 보여주면서, 이용자의 참여를 함께 독려하는 것이 설득력이 있을 것이라고 생각한다. 해외 글로벌 테크 기업은 다양성 리포트를 발간하며 부족한 부분을 채우기 위한 노력을 보여준다. 카카오도 그런 방향으로 가야 한다. 모든 것을 규정으로 제재할 수 없다. 카카오 구성원 모두가 그런 의식을 가지고 있어야, 가이드라인을 넘어서 문제가 있을 때 문제제기 할 수 있다. 내부직원들의 인권 의식 수준을 높이기 위한 조치도 함께 병행해야 한다.

임동준 활동가 민주언론시민연합

카카오 등 포털을 통해서 혐오표현이 어떻게 확산되고 있고, 어떻게 근절시킬지 논의하겠다. 네이버·카카오가 유명한 자살 사건을 기점으로 연예·스포츠 댓글을 폐지했는데, 관련 문제를 해결했다고 볼 수 없다. 결과적으로 댓글은 연예·스포츠 기사에서 파생된 것이다. 유해한 연예 기사가 해결되지 않으면 댓글 문제는 해결되었다고 볼 수 없다. 본질적 문제는 혐오표현을 포함하거나 유발하는 보도들이다. 10월 4일 많이 본 연예 뉴스 7위 기사 제목이 ‘돈 많이 번 여배우’다. 제목에는 ‘여배우’라는 차별 표현을 포함하고 있다. 근본적인 문제는 차별이나 혐오표현을 쓰는 기사이므로, 댓글 폐지 등 이용자 규제가 만능은 아니다. 댓글 폐지 후 언론보도 서비스에서 이용자들의 혐오표현이 사라진 것도 아니다. 가십성 내용의 SNS 누리꾼 반응을 보도에 인용하기 때문에, 기사에 혐오표현들이 포함되어 있다. 이와는 반대로, 멀쩡한 보도에 혐오표현 댓글이 달리기도 한다. 최근 차별금지법 관련 기사들에서도 많은 혐오표현 댓글이 달렸다. 이를 보면 하나의 구조가 보인다. 언론이 혐오표현을 사용하거나 유발하는 기사를 만들면, 혐오 정서가 확산된다. 그럼 이용자들이 혐오표현을 사용하고, 이를 언론이 다시 끌고 들어오는 악순환이 반복되는 것이다.

이를 막기 위한 포털의 근본적인 해결책은 뉴스제휴평가위원회의 제재 규정 보완이다. 우선 포털은 언론이 제휴할 만한 매체인지 확인해야 한다. 제휴 매체의 문제를 제재하는 기구인 뉴스제휴평가위원회가 이미 존재하기 때문에, 이 기구에 혐오표현을 규제할 수 있는 장치가 있다면 충분히 걸러낼 수 있을 것이다. 현재 뉴스제휴평가위원회 심사규정에는 ‘15조 부정행위 - 선정적 기사 및 광고’ 관련 내용이 있기는 하지만, 혐오표현이라는 구체적 용어는 없으며, 규제 근거 또한 빈약하다. 또 뉴스제휴평가위원회가 명맥상으로만 제재 기구이기 때문에, 제대로 기능하지 않는 것도 개선해야 한다. 규제를 가지고 있지만, 실질적으로 규제할 수 없다면 문제가 해결되지 않을 것이다.

가장 간단하게 해결할 수 있는 방안도 있다. 이용자와 언론 모두에게 메인에 어떤 뉴스가 있는지 매우 중요하다. 혐오표현이 제목에 포함되어 있는 경우 메인 노출이 불가 하도록 하거나, 더 강한 혐오·차별표현 내용을 포함하면 포털에 송부할 수 없도록 하거나, 혹은 약한 혐오·차별 표현을 포함한다면 혐오표현을 포함하고 있다는 안내를 상단에 표기하는 것도 괜찮을 것이다. 페이스북이나 트위터도 허위 정보에 대해서 비슷한 안내를 하고 있다.

조소담 대표
닷컴페이스

닷컴페이스는 우리 사회의 차별이나 폭력을 겪은 사람들의 이야기를 영상으로 담고 있다. 영상에 출연하시는 분들이 차별이나 폭력의 대상이 되는 분들이 많아서 신경을 많이 쓰고 있다.

첫째로, 인스타그램, 페이스북 등 각 플랫폼이 얼마나 다른 환경을 가지고 있는지에 대해 말씀 드리겠다. 콘텐츠를 플랫폼에 발행할 때 안전하지 않다고 느끼고, 댓글에 신경 쓰게 되는 플랫폼은 카카오톡, 유튜브나 인스타그램 같은 경우에는 피해를 사전에 방지하기 위해서 능동적인 행동들을 할 수 있다. 예를 들면 콘텐츠를 업로드할 때 전체 공개로 할 것인지 혹은 구독자에게 먼저 노출할 것인지, 인스타그램은 친한 친구를 대상으로 공개할 것인지 등 공개범위를 조절할 수 있다. 유튜브 같은 경우에는 댓글 커뮤니티에서 댓글을 필터링 할 수 있는 기능이 있다. 한 영상물에 지역차별 관련 댓글들이 한참 올라오기 시작할 때, 이런 맥락의 댓글이 많이 작성됨을 파악하고 댓글 필터링으로 조절할 수 있었다. 어떤 범위의 사용자에게 콘텐츠를 공개할 수 있는지를 조절하는 권한을 가졌을 때, 안전한 플랫폼으로 여겨진다.

둘째로는, 피해 상황이 일어나기 전에 방지하는 것 뿐 아니라 피해 상황 이후에 능동적으로 행동할 수 있도록 피해 복구를 위한 조치들이 중요하다. 예를 들어 학교 폭력 피해자는 가해자들에게 단체 채팅 방에 초대되었고 나쁜 말들이 쏟아진 이후에 채팅 방이 폭파되었다. 계속해서 채팅 방은 다시 만들어지고 폭파되며 피해자는 지속적인 괴롭힘을 당했다. 채팅방 복구를 원할 때, 대화를 신고하고 증거 수집이 어려운 경우가 있다. 이런 부분은 플랫폼이 도움을 주어야 한다고 생각한다. 또 예를 들어, 저희 단체 활동 중 피해를 중간부터 최소화 하기 위해 어떤 흐름이 만들어지는 걸 방지하고 차단하는 경우도 있다. 흐름 차단이 중요한 이유는 어떤 한 사람이 혐오성 차별 발언 댓글을 올리는 것이 중요한 것이 아니라, 그 댓글을 보고 흐름이 만들어져 다 같이 몰려와서 공격을 하기 시작한다는 것이다. 커다란 흐름으로, 즉 집단적인 공격으로 이어지는 것을 막을 수 있는 조치를 취해야 한다. 예를 들어서 댓글을 올리는 사용자에게 패널티를 준 다거나, 콘텐츠를 제공하는 사람에게만 댓글이 보이도록 설정을 바꾸는 등 여러 방법을 시도할 수 있다. 그리고 실제로 피해를 받는 사람들도 한 명이 악성 댓글이나 게시물에 올렸다는 것 보다는, 10명의 사람들이 그 게시물을 방관하고 있었다는 것에서 더 큰 상처를 받는다.

셋째로는, 플랫폼에 직접 게시물을 신고하고 게시물 삭제를 요청하는 경우가 많은데 이 때 보호해야 할 대상들에 대한 범위가 불명확하다. 동의하는 부분은 큰 범위의 정의로 여성, 아동, 성소수자 등으로 필수적으로 해당 집단들은 자신들이 보호 받을 수 있음을 인지하도록 하는 것이 중요하다.

어떠한 경우에는 댓글만 봤을 때 혐오표현이 직접적으로 나타나거나 하지는 않지만, 혐오성의 취지를 가진 댓글들이 올라오고 이것이 이미지 등과 결합되기도 했다. 보호해야 하는 대상을 사전적으로 정의하기 보다는, 다른 방식의 접근이 필요하다. 애매하고 알기 어려운 경우에는 플랫폼이 단독적으로 판단하거나 패쇄적으로 신고 과정을 진행하기 보다는, 이 게시물이 사람들이 느끼기에 문제가 있다는 것을 알리는 것이 필요하다. 문제가 있어서 콘텐츠를 삭제한다는 결정보다, 아주 많은 사람들이 게시물에 문제가 있음을

표시했기 때문에 해당 게시물을 가리고 안 볼 수 있게 하는 권한을 이용자들에게 주는 방식들도 가능하다. 마지막으로 외부 링크들 중 다음카페 링크로 커뮤니티에 들어왔을 때 갑자기 공격댓글들이 많아지는 경우가 종종 있었다. 이용자에게 안전한 환경을 보장하기 위해 커뮤니티의 폐쇄성을 마련했는데, 사실 커뮤니티의 폐쇄성은 안전하게 남을 공격하기 위해 악용되기도 한다. 그래서 공격성 게시물이나 댓글이 의심되는 다음카페 링크를 신고하면, 플랫폼 차원에서 확인하고 걸러내는 작업이 필요하다고 생각한다. 다음카페, 네이버, 일베 등에서 게시물이 퍼져 어느 인물이 공격 받는 상황이 생겼을 때, 신고를 하면 우리끼리 “일베가 제일 조치 빠르게 해준다”라고 말하는 상황도 있었다.

**조진경 대표
십대여성인권센터**

저희 기관은 성착취 피해 아동 청소년 지원 기관이다. 카카오가 여러 매체의 정착지 매체로 사용되고 있기 때문에 이 부분을 살펴봤으면 한다. 혐오표현의 정의와 혐오에 대한 고민을 해보았는데 지난 한국 사회에서 발생한 미투(#MeToo) 사건의 경우 노골적으로 피해자에 대한 혐오들이 얼마나 많은지 봤다. 피해자에 대한 혐오는 단지 성범죄 피해자뿐 아니라 모든 약자도 포함됐다. 이를 통해 볼 때 대한민국에서는 피해자나 약자를 혐오하는 분위기가 있다. 더 나아가 성매매의 경우 피해자를 전혀 피해자로 인식하지 않는 분위기가 있어서 성범죄자들은 그 분위기를 이용해 성범죄를 성매매로 환원한다. 아동 청소년 성착취는 성폭력과 성매매가 구별이 안 된다. 따라서 성매매를 포함한 성범죄를 성착취로 봐야 하며, 특히 ‘여성을 성적 도구로 사용 가능하다는 방식의 생각들이 가장 심각한 여성혐오다’라는 관점이 필요하다. 성매매를 혐오라고 여기지 않는 것, 그리고 이와 같은 문제의식이 없다는 것이 가장 심각한 여성 혐오라는 시각이 필요하다.

정부의 역할 뿐만이 아니라 기업의 사회적 책임도 매우 중요하다고 생각한다. 착취물이 올라오면 이를 차단할 수 있는 기술 개발을 기업에서도 해야 한다. 한국에서는 기업이 사회적 책임이 없다고 생각하고 있으며, 다른 나라들에 견줘 해외 플랫폼들은 우리나라에서는 책임을 지지 않으며, 지게 하지도 않고, 지게 할 수가 없다. 사회를 통해 기업은 이윤을 얻기 때문에, 기업이 사회에 대한 책임으로 일정 부분 사회 환원이 당연하다고 생각한다. 특히나 미래를 책임지는 아동, 청소년들에 대한 성착취 범죄에 대해서는 기업이 기술 개발에 대한 책임이 있다. 성착취 가해자를 찾고 증거를 제공하고, 더 나아가서 이윤의 일부를 피해자에게 환원하는 것 등이 기업의 사회적 책임에 포함된다고 생각한다.

IT 기술을 기반으로 하는 아동 청소년 대상 성착취 시장이 커지고 있다. 과거에는 오프라인으로 진행이 되었지만 이제는 조건 만남 등이 IT 기술을 기반으로 이뤄지고 있다. 채팅 어플을 통해서 초기 만남을 갖고, 1대1 채팅 사이트를 이용하는데 대부분이 카카오톡을 사용하고 있다. 이런 현황에 대해서 어떤 방식의 접근이 필요한지에 대한 고민을 많이 한다. 그래서 민간단체와 기업의 협업이 매우 중요하다. 기술 개발의 방향성도 함께 고민하고, 피해자들이 실제로 들어와 있는 기관들의 목소리에 귀를 기울이며 기술을 활용할 수 있도록 고민해야 한다. 미국 같은 경우에는 이미 기술을 활용하여 피해자를 보호하는 여러 사례들이 존재한다. 미국에서는 ‘스포라이트’라는 프로그램을 통해 인터넷 브라우징 히스토리나 이동 경로를 추적해서 가해자를 바로 검거하는 등의 조치들이 마련돼 있다.

마지막으로, 초기 인터넷 사용 시 오프라인 상의 고정적인 정체성에 갇혀 있던 사람들에게 익명성은 다른 정체성을 가질 수 있게 돼 매우 신선하게 느껴졌는데, 현장에 계속 있으면서 사이버 상의 익명성에 대해 다시 고민해야 한다고 생각하게 되었다. 사이버 세계가 현실 세계보다 더 현실 세계처럼 느끼는 건

10~30대이다. 미래를 내다봤을 때 사이버 환경을 더 현실로 느끼는 인구층이 더 많아질 것이다. n번방 사건에서 나타났듯, 현실세계와 사이버 세계의 지배 규칙이 너무 다르다. 조주빈과 같은 경우에는 현실에서는 자기 모습을 현실의 규칙에 맞게 이미지화 하고 있었고 사이버상에서는 익명 뒤에 숨어 최악의 모습을 가지고 있었다.

현장에서 보았을 때에는, 피해자 정체성은 굉장히 다양해졌다. 사이버 공간에서는 익명이라는 전제 하에, 성역할 등 개인의 정체성이 너무나 다양해졌다. 사이버 세계는 익명성 뒤에서 혐오와 폭력, 차별이 가장 강력한 규칙이다. 현 10·20·30대는 40·50대와는 다르게 더 이상 고정된 현실세계에서만 정체성이 있는 것이 아니다. 사이버 세계와 현실 세계의 지배 규칙이 같아 져야 한다. 사이버 공간에서의 실명화를 추진해야 한다고 생각한다.

이완 대표
아시아인권문화연대

혐오표현을 구별하는 가이드라인 만들고, 카카오 조직 내에 이를 판단하는 기구가 있는지?

김대원
카카오 이사

질문 주신 부분에 대해서는 다양한 방안, 즉 기술과 서비스가 결합된 방식, 내부적 모니터링 등의 모든 가능성을 열어두고 고민하고 있다고 답변 드리겠다. 회사 내에서는 아직까지 혐오표현 제재를 해야 할까, 라는 부분도 정해지지 않았다.

말씀주신 의견들이 저희의 생각과 많이 일치하지만, 문제는 현실에서의 파급력이나 부정적 영향 또한 같이 고려해야 한다는 점이다. 방향은 정해졌으나, 실제 어떤 행동을 취할 것인지는 치열하게 고민 해보아야 할 것 같다.

이완 대표
아시아인권문화연대

낮은 수준이나 단계의 표현들은 기술로 거르는 것이 가능하다. 하지만 몇몇 표현에는 논쟁이 생길 수 있다. 이처럼 높은 단계의 경우에는 심의기구를 통해 혐오표현을 판단하고 조치한다. 이외의 다른 방식은 없을 것으로 보인다. 일본에서는 혐오표현 신고를 하면 시민단체와 같은 심의기구가 심의를 하고 플랫폼에 삭제 요청하여 플랫폼에서 글을 삭제하게 되고 게시자의 이름이 공개되는 등의 조치를 하고 있다.

이준목
카카오 이사

카카오는 열린 게시판을 만들어 이용자들이 서로의 의견을 주고받도록 했다. 이때 악관을 위반한 경우 적극적으로 규제하고 있다. 그러나 혐오표현이 지속적으로 변화하고 다양해지다 보니, 접근이 어려운 부분이 있다. 카카오 내부의사결정기구도 있다. 자율정책기구인 KISO에 안건을 올려 논의하기도 한다. 명예 훼손, 검색어 정책 등은 이미 정립이 되어 있어 다음 단계인 차별·혐오표현을 아젠다로 올려 논의 중이다. 이렇게 존재하고 있는 기구를 활용할지 다른 방법이 필요할지 고민 중이다. 기술적으로도 접근하려고 고민하고 있다. 명백히 불법적인 것들은 이미 기술적으로 판별하는 것이 발전이 많이 되어있으나, 혐오는 텍스트 차원의 명문화된 것이 없어 기술적으로 인식하고 조치를 취하는 것에 난관이 있다. 우선 혐오표현의 기준을 마련하는 것이 가장 중요한 것 같고, 이를 기술화 하는 방향으로 진행할 것이다.

이완 대표
아시아인권문화연대

애플의 경우 다양성 부사장이 있다. 카카오에서도 혐오표현 담당자의 지위가 높아야 사내에서 잘 받아들여질 수 있을 것이라고 생각한다.

김지학 소장
한국다양성연구소

제재라고 해서 단지 통제라기 보다는 교육이라고 접근했으면 한다. 국가가 하지 못한 것을 사기업이 하는 것이 큰 용기다. 우리 플랫폼이 혐오표현을 허용하지 않는다는 것을 천명하는 것 자체가 큰 의미를 담고 있다. 삭제 대신 혐오표현을 많이 포함한 게시물임을 알리는 것만으로도 사람들이 생각하고 배우는 기회가 될 것이다. 혐오표현이 포함되어 있다고 표시하는 단계, 삭제하는 단계, 게시자를 차단하는 단계 등 여러 단계가 있을 것이라고 생각한다. 그래서 처벌보다는 교육이라고 생각한다. 처벌 중심의 사고보다 어떠한 문화를 만들지, 어떤 교육 효과를 낼수 있을지 측면으로도 접근해야 한다.

이준웅
자문위원

기술적인 노력으로 일부 확인될 수 있는 것들, 예를 들면 키워드, 연관어 검토 등은 스마트하게 파악할 수 있고 대처할 수 있을 것 같다. 그러나 규제하려는 노력이 있다면, 그에 대한 이용자의 대응 방법도 기민해진다. 따라서 기술적 추격으로 대응하는 방식에도 한계는 있다. 따라서 이용자의 발언과 행동을 통제한다는 생각 자체에 대해서 원리적으로 접근해야 한다. 규제 가능한 발언과 행위의 목록을 정해서 대응하는 방식은 한계가 있을 것이다.

부록2. 공동연구 결과

온라인 혐오표현이란 무엇인가

한국언론법학회·국가인권위원회·카카오

I. 들어가며

- 역사적으로 혐오표현은 역사적, 사회적 소수자로서 정신적, 신체적 핍박을 받아온 집단을 표적(target)으로 삼아왔음(김민정, 2020).
- 디지털 시대에 혐오표현은 더 많은 인터넷 이용자들을 대상으로 더 빠르게 전달되는 속성이 있음. 특히 익명성을 이용하여 특정 집단을 표적으로 삼는 인터넷 혐오표현의 파급력은 커지고 있으며, 온라인 혐오표현과 관련된 이슈들은 전 세계적으로 문제가 되고 있음.
- 혐오표현의 개념과 범주에 대해서는 한가지로 정의하기는 어려우며 학술적으로도 다양한 의견이 제시되고 있음. 특히 한국의 경우 혐오표현과 관련된 역사가 다른 국가에 비해 비교적 짧은 편이어서 혐오표현의 개념정립이 필요하며 그 심각성에 대해서 시민들의 인지도가 낮은 편임.
- 이 연구는 특히 인터넷상의 혐오표현과 관련하여 시민들의 인식을 조사하는 한편, 전문가의 의견을 수렴하고 해외 소셜미디어들의 대응방안을 조사하는 것을 목적으로 함.

1. 혐오표현의 개념과 영향

□ 개념요소

- 혐오표현이란 성별, 장애, 종교, 나이, 출신지역, 인종, 성적지향 등을 이유로 어떤 개인, 집단에게 1) 모욕, 비하, 멸시, 위협 또는 2) 차별, 폭력의 선전과 선동을 함으로써 차별을 정당화, 조장, 강화하는 효과를 갖는 표현(국가인권위 <혐오표현리포트> 2019).

※연구진 註: 이 보고서는 2020년 9월 22일 열린 ‘온라인 혐오표현이란 무엇인가’ 세미나에서 발표된 내용을 바탕으로 관련 자료와 데이터를 추가분석한 것입니다. 세미나 발표 및 보고서 작성은 한국언론법학회 소속 학회원인 이승선(충남대학교 언론정보학과 교수), 박아란(한국언론진흥재단 책임연구위원), 최진호(한양대학교 컴퓨터이셔널사회과학 연구센터 박사), 이주영(서울대학교 인권센터 전문위원)이 맡았으며, 보조연구원 곽은아(한양대학교 미디어커뮤니케이션학과 박사과정)가 참여했습니다. 자문을 맡아주신 윤성옥(경기대학교 미디어영상학과 교수) 선생님과 세미나 토론 및 보고서 작성에 도움을 주신 여러 한국언론법학회 학회원들께 감사드립니다.

- <혐오표현리포트>에 따르면 혐오표현은 다음과 같은 개념요소를 갖고 있음.
 - 1) 어떤 속성을 가진 특정 집단을 대상으로 함
 - 2) 특정 집단에 대한 부정적 편견과 고정관념을 바탕으로 대상 집단이나 그 구성원을 모욕, 비하, 멸시, 위협하거나 차별, 폭력을 선전, 선동함
 - 3) 대상 집단에 대한 물리적 공격이 아닌 언어 등을 사용한 언동

□ 개념에 대한 다양한 관점

- 윤성옥(2019) ‘혐오표현 규제와 법적 쟁점에 관한 연구’에 따르면 혐오표현은 다음과 같이 다양한 관점에서 개념 지어짐.
- 혐오표현이란 ‘개인의 정체성’에 근거함. 따라서 혐오표현은 주로 국가, 인종, 민족, 종교, 성적 성향, 직업, 성별 또는 장애와 관련됨. 국내 학자들을 대체로 혐오표현에 대해 “지역·민족·인종·국적·종교·장애·성별·성정체성·연령 등 일정한 속성을 이유로 특정 집단이나 특정 집단 구성원에 대해 차별·모멸·혐오적인 의사를 표현하거나 다른 사람들도 이에 동참하도록 선동·선전하는 것”이라고 정의하고 있음.
- 역사성을 지닌 소수자를 강조하는 관점이 있음. 국가인권위원회는 소수자들을 신체적·정신적으로 괴롭히는 표현, 차별·혐오를 의도·암시하는 표현, 소수자를 멸시·모욕·위협하여 인간존엄성을 침해하는 표현, 차별·적의·폭력을 선동하는 표현이라고 했음. 또는 역사적으로 차별받아 온 집단이라고 특정하고 있음. 혐오표현의 대상 집단을 역사적 소수자로 국한하는 관점에서는 남성에 대한 혐오는 가능하지 않고 경상도 지역민을 공격하는 것도 포함되지 않음. 그러나 혐오표현의 대상으로서 소수자만을 명시하거나 역사적으로 차별받아 온 집단을 명시하는 것은 명확성 원칙이나 평등원칙에 반한다는 입장도 있음. 혐오표현의 해악은 다수이더라도 동일하고 다수자도 피해자가 될 수 있다고 보는 것임.
- 차별과 편견을 강조함. 유럽인권재판소는 혐오표현을 “공격적인 국가주의와 민족주의, 그리고 이주민들과 소수자들을 향한 차별과 적대감 등 편협함에 기반한 인종혐오, 외국인혐오, 반유대주의 등의 증오를 전파, 선동, 선전 또는 정당화하는 모든 형태의 표현”이라고 정의하였음. 특히 혐오표현에서 선천적 특성에 따른 편견의 개념을 중요하게 보기도 함. 즉 "그들의 실제적이거나 인지된 선천적 특성 때문에 사람이나 집단을 겨냥한 편견적, 적대적, 악의적인 표현"이라고 혐오표현을 정의하기도 함.
- 공격적 표현뿐만 아니라 선동적 표현까지 포함함. “통상 인종·민족·국적·종교·연령·장애·성별 그리고 성 정체성 및 성적지향 등과 같은 사람의 특성에 근거하여 사람 또는 그와 같은 특성을 가진 사람의 집단에 대해 혐오의 감정을 공격적으로 표현하거나 그와 같은 혐오를 선동·확산·조장하는 표현”이라고 정의하기도 함. 따라서 혐오표현에는 직접적인 폭력이나 집단학살과 같은 행위는 제외된다. 혐오표현과 혐오범죄가 구분된다는 관점이 있음. 혐오범죄는 신체적·물리적 폭력행위가 수반되는데 혐오표현의 쟁점은 신체적, 물리적 폭력행위가 없는 발언이나 표현행위를 법적 규제와 처벌의 대상으로 할 수 있는가의 문제이기 때문임. 그러나 목표 집단에 대한 폭력행위는 아니더라도 선동표현은 포함하는 것으로 봄.

- 효과나 영향을 중요한 개념요소로 보는 관점도 있음. 일단 혐오표현이 발화되면 상처나 모욕은 희생자의 인지에 달려있음. 혐오표현의 영향력이 발생할 수도 있고 없을 수도 있다는 것임. 즉, 혐오표현의 효과란 발화자, 내용, 목표대상(originator, content and the targeted one)에 따라 달라짐. 특히 차별을 선동하지 않는다면(또는 상처를 주지 않는다면) 문제가 되지 않고 무엇이 혐오표현인지 애매하게 됨.
- 국내에서도 대체로 혐오죄의 구성요건은 혐오의 감정 자체가 아니라 외부로 표출되는 행위가 될 것이며 외부의 행위가 형법적 규율대상이 되는 경우 내심을 동기로 고려하는 것은 가능하다고 봄. 이에 따라 혐오표현이란 ①개인의 타고나거나 인지된 속성(이는 개인의 의지에 따라 거의 선택할 수 없는 태생적 본능적 속성에 가깝다)에 대해 ②차별과 편견을 기반으로 한(원인이나 동기) ③공격적 표현이나 선동적 표현(표현내용)을 중요한 개념요소로 판단함.

□ 혐오표현의 부정적 영향

- 혐오는 결국 차별의 문제로 귀결됨. 혐오표현은 표적 집단을 끊임없이 구분 짓고 배격하고 차별을 조장하기 때문임.
- 혐오표현은 그 자체로도 해악을 끼칠 수 있지만 표적 집단에 대한 폭력적 행위인 혐오범죄를 일으킬 위험이 있음.
- 혐오표현은 인간의 존엄과 가치를 부정하고 개인적 인권(명예권이나 프라이버시권과 같은 인격권, 자유권, 평등권 등)을 침해함.
- 혐오표현은 사회통합에 부정적 영향을 끼치고 결과적으로 민주주의 발전을 저해할 가능성이 있음.

2. 혐오표현 규제를 위한 노력

□ 유럽

- 유대인을 향한 혐오표현으로 인해 세계대전이 촉발된 역사적 경험이 있음. 그 결과 많은 유럽국가에서 혐오표현을 형사적으로 처벌함.
- 독일: 형법 제130조(증오선동, 폭력적 조치·자의적 조치 촉구, 명예훼손, 해당 내용 문서나 상영물 배포, 전시, 광고, 전파 등을 처벌).
- 프랑스: 언론자유법 제24조 등(차별, 혐오 선동, 모욕 처벌).
- 영국: 공공질서법 제3장(혐오선동, 출판·연극·방송 등에서의 혐오표현 처벌).

□ 미국

- 혐오표현도 수정헌법 제1조의 보호를 받는 표현의 범주에 속함.
- 미국 연방대법원은 폭력을 단순히 옹호하는 표현을 제한해서는 아니 되며, ‘급박한 위법행위를 선동 또는 야

기(inciting or producing imminent lawless action)'하려는 의도가 있는 표현만 규제할 수 있다고 판시함. 즉, 어떠한 표현이 해악이나 폭력을 야기하려 한다는 구실로 정부가 정치적 반대의견을 억압하는 것을 금지하고 정치적 표현의 자유를 최대한 보호하기 위한 것임.

- 그러나 혐오로 인한 범죄행위인 혐오범죄(hate crime)는 수정헌법 제1조의 보호를 받지 못함. 미국 양형위원회 가이드라인 매뉴얼에 따르면 피해자의 개인적 특성(인종, 피부색, 국적, 성별, 성적 지향, 장애 등)이 범행 동기가 되었을 때는 3배까지 가중처벌이 가능함.

□ 한국

- 2010년 이후 외국인 근로자, 다문화 가정, 특정지역을 타겟으로 삼는 온라인 혐오게시물 급증. 2015년 강남역 살인사건을 계기로 여혐·남혐 논쟁 심각해짐. 코로나19 팬데믹 이후 중국인과 특정 종교에 대한 혐오표현 확산됨.
- 혐오표현의 심각성에 대한 사회적 인식 높아지고 있으나 정작 혐오표현이 무엇인지, 왜 혐오표현을 규제해야 되는지에 대한 논의는 부족한 상황임.
- 20대 국회에서 발의된 8개의 혐오표현 관련 법안 모두 회기 만료 등의 이유로 자동 폐기됨.
- 판례: 혐오표현을 규율하는 법조항이 따로 없기 때문에 피해자가 특정될 수 있는 경우의 혐오발언은 명예훼손이나 모욕으로 처벌됨.
- 형사정책연구원 보고서: 2016~2017 온라인 모욕죄 사건 분석한 결과 혐오표현이 119건으로 전체의 31.6%를 차지함.
- 혐오발언으로 인한 모욕죄 유죄판결 사례: 웹툰 작가를 "한남총"이라고 칭한 경우 벌금형, 버스를 함께 탄 인도인에게 "아랍인은 더럽다, 냄새난다"고 발언한 경우 벌금형, "여자들은 똥똥하면 안된다 못생긴 년"이라는 발언에 대해 벌금형, 인터넷 게임을 하던 중에 "김치년, 낙태총"이라는 표현에 대해 벌금형, "김치년 나라 망신시켰다"는 댓글에 대해 벌금형, "짱깨새끼 너네 나라로 가라"는 발언에 대해 벌금형 등 다양한 사건에서 혐오발언이 모욕죄로 처벌됨.
- 특정인에 대한 흉어, 전라디언, 종북, 호모새끼 등의 명예훼손 및 모욕의 게시글이 올라왔을 때 웹사이트 운영자는 삭제요청을 받았을 경우 이를 즉각적으로 삭제할 의무가 있다고 법원은 판결함(2013카합1661).
- 헌법재판소 2019. 11. 29. 선고 2017헌마1356결정: 서울시 학생인권조례가 혐오표현을 금지하는 것은 표현의 자유를 침해하는 것이 아니라고 헌법재판소는 판단함. 헌법재판소는 차별적 언어나 행동, 혐오적 표현도 헌법 제21조 표현의 자유의 보호 영역에는 해당하되, 단 헌법 제37조 제2항에 따라 제한 가능하다고 판시함. 혐오표현은 그 자체로 상대방의 인간의 존엄성을 침해하고 특정 집단의 가치를 부정한다는 문제점이 있으며 특히 성장기 학생을 대상으로 한 차별혐오표현은 학생의 정신적, 신체적 능력을 훼손하고 인격과 가치관 형성에 부정적 영향을 끼친다고 봄. 따라서 서울시 학생인권조례가 금지한 차별, 혐오표현은 민주주의를 위해 허용되는 의사표현이 아님. 타인의 인권 침해를 인식하였거나 인식할 가능성이 있고 결과적으로 인권 침해가 발

생하는 표현이므로 표현의 자유가 허용하는 한계를 넘는 것으로서 제한 필요성이 있다고 결정함.

3. 혐오표현 규제의 어려움

□ 개념 정의의 어려움

- 학술적으로 혐오표현 개념에 대해 다양한 견해가 있음. 소수자(수적 소수자 또는 지배적 위치가 아닌 자)만 대상이라는 견해, 역사적으로 차별 받은 소수자 집단을 대상으로 하는 견해, 다수도 포함되는 특정 집단을 대상으로 한다는 견해, 차별적 사유를 근거로 폭력을 초래하는 표현(협의)과 기타 차별적인 단순한 혐오표현(광의)으로 나누는 견해 등 다양함.

□ 역사적, 사회적 맥락이 반영되어야 하는 어려움

- 유럽, 미국과는 달리 한국의 경우 인종이나 종교로 촉발된 혐오표현 관련 역사적 경험 희박함. 따라서 혐오표현이 사회에서 본격적으로 논의된 경험도 적음.
- 결국 한국에서 혐오표현이 무엇인지, 어떤 표현의 규제에 초점을 맞춰야 될지에 대한 적극적인 논의가 필요하며 그에 대한 합의점을 찾아나가야 함.

□ 처벌 대상 규정의 어려움

- 어떤 것이 표현의 자유의 한계를 벗어나는 혐오표현인지 일의적으로 규정하기 어려움.
- 따라서 실질적 차별행위, 혐오로 인한 폭력 등 행위에 규제 초점을 맞출 필요가 있음

□ 법규 적용의 어려움

- 피해자(특정, 소수집단)가 특정될 수 있는 경우 현행법상 명예훼손(형법, 정보통신망법), 모욕(형법), 불법정보(정보통신망법) 등 여러 처벌규정의 적용이 가능하고 민사상 불법행위로 인한 손해배상청구도 가능함. 그러나 피해자 특정이 어려운 혐오표현이라면 현실적으로 처벌이나 손해배상을 구하기가 어려운 상황임.
- 대법원 양형위원회는 허위사실적시 명예훼손죄와 모욕죄 양형기준에 있어서 혐오 또는 증오감에서 범행을 저지른 경우를 특별 양형인자로 포함시킴(2019).
- 결국 혐오표현이 구체적인 침해를 불러일으키거나 차별적 행위와 결합했을 때 법적 제재가 가능함. 법률적 공백을 메우기 위해서 차별금지법도 필요함.

□ 디지털 테크놀로지 발달로 인한 어려움

- 혐오표현 확산에 기여하는 혐오미디어(hate media)는 과거 라디오, 텔레비전이었으나 이제는 인터넷이 혐오미디어로 자리매김함.

- 혐오표현은 가짜뉴스와 결합하기도 하고, 특히 자극적인 가짜뉴스와 결합했을 때 빠른 속도로 확산되는 경향을 보임.
- 소셜 미디어에서는 리트윗(retweet), 리그램(gram), 좋아요 등의 다양한 디지털 표현으로 인해 혐오의 감정이 확산되고 증폭되는 양상을 보임.

3. 혐오표현 규제 동향에 대한 논의

- 혐오표현에 대한 규제를 반대하는 사람들은 혐오표현과 차별이나 폭력의 상관성에 대해 항상 의구심을 가지고 있음(Baker, 2012; Dworkin, 2009). ‘표현의 자유’ 옹호론도 오프라인이건 온라인이건 막론하고 혐오표현에 대한 규제를 도입하는 데 방어논리로 잘 활용되어 왔음.
- 규제 반대론자들은 혐오표현에 대한 규제와 법은 권력남용이 될 수 있고, 실효성이 없으며, 발화자 입장에서 자율성을 해치고, 그 결과 사회 전체의 민주주의 체제에 손상을 끼친다고 보고 있음(Baker, 2009, 2012; Dworkin, 2012).
- 규제 찬성론자들은 혐오표현의 해악성에 주목함.
 - 혐오표현의 해악은 두 가지 차원으로 나타나는데 우선 혐오표현이라는 행위자체로 발생하고 다음은 행위 결과로서 발생함.
 - 혐오표현 행위는 그 자체가 인간존엄성에 대한 침해임. 목표집단의 구성원들에게 불쾌한 감정과 현실적인 공포를 느끼게 하며, 명예를 훼손하거나 인격에 상처를 입힘으로써 존엄성을 침해함(Waldron, 2012).
 - 그 다음 혐오표현의 행위결과로 특정집단이 침묵하거나 배제됨으로써 공론장이 왜곡되거나 폭력행위까지 초래될 수도 있음. 예를 들어 여성을 부당하게 열등한 존재로 종속시키는 표현에 의한 피해, 그리고 혐오표현의 결과 여성들을 어떻게 침묵시키는지로 나누어 볼 수 있음(Gelber & McNamara, 2016).
- 온라인 혐오표현을 오프라인 혐오표현과 달리 봐야 하는지에 대해서도 논란이 많음. 온라인은 이질적이면서도 역동적이고 다양한 방식으로 짧은 시간 안에 확장됨(Citron, 2014; Cohen-Almagor, 2015). 접근성, 수용자의 크기, 익명성, 즉시성이라는 특징 때문에 온라인 혐오는 오프라인 혐오와는 차원이 다른 문제라고 보는 입장임. 따라서 오프라인과 달리 온라인 혐오표현의 경우 규제의 필요성과 정당성이 인정될 수도 있음.
- 특히 인터넷이 자발적인 혐오표현을 부추기고 있으며 이미 양적으로 온라인 혐오표현이 오프라인 혐오표현을 넘어섰고 이용의 편리함, 접근성, 저비용으로 앞으로 더 증가할 것임(Brown, 2018). 인터넷의 특수성에 주목하는 관점은 혐오표현에 대해 새롭게 접근할 것을 요구함(Weede, 2016). 공격적인 집단으로부터 사회를 지키는 데 있어 전통적인 규제방식은 이제 한계에 이르렀다고 봄(Balica, 2017).
- 혐오표현은 테러와 상당히 관련이 있고 특히 온라인에서 혐오표현은 테러 공격을 야기하거나 촉진한다고 보는 입장은 정부와 인터넷서비스사업자가 보다 적절한 정책을 마련해야 한다고 주장함(Chetty & Alathur, 2018). 특히 인터넷서비스사업자가 상업적 목적을 넘어 혐오표현의 문제를 해결해야 함. 이러한 형식의 규제

아웃소싱에 대해 표현의 자유를 옹호하는 입장에서는 반대하겠지만 사업자의 자율규제에 맡기지 않는다면 정부에 의한 표현의 자유 침해가 더 심각하고 엄격하게 발생할 수 있다고 경고하고 있음(Brown, 2018).

4. 해결되어야 할 과제

- 무엇이 혐오표현인지에 대한 개념 및 혐오표현의 범위에 대한 적극적인 논의 및 사회적 합의가 필요한 시점임.
- 온라인 이용자들에게 혐오표현의 문제점을 적극적으로 인식시켜야 함. 근거 없이 특정 집단을 겨냥한 혐오표현은 정당한 비판의 범주를 벗어난 것이므로 표현의 자유로서 보호될 수 없다는 인식을 온라인 이용자들이 갖도록 교육이 필요함.
- 온라인 공간에서 혐오표현을 자제하려는 시민들의 자율적 노력 및 효과적인 미디어 리터러시 교육이 시급함.

II. 글로벌 소셜미디어 기업과 혐오표현 규제

1. 글로벌 소셜미디어와 혐오표현

□ 글로벌 소셜미디어

- 온라인 플랫폼을 인터넷을 매개로 서비스를 통해 복수의 개별적이지만 상호의존하는 이용자들 사이의 상호작용을 촉진하는 디지털 서비스라고 정의할 때, 온라인 플랫폼 서비스에는 소셜미디어, 앱 스토어, 온라인 마켓, 결제 서비스, 각 이코노미(gig economy), 검색엔진, 커뮤니케이션 서비스, 스트리밍 서비스 등이 포함됨(OECD, 2019).
- 서비스 간 중첩된 영역도 많기 때문에, 각각의 서비스를 명확하게 구분하기 어려움.
- 이 글은 온라인 플랫폼 중 페이스북(Facebook), 인스타그램(Instagram), 트위터(Twitter), 유튜브(YouTube), 핀터레스트(Pinterest), 레딧(Reddit), 틱톡(TikTok)과 같은 글로벌 소셜미디어에서의 혐오표현 문제와 그에 대한 접근을 중심으로 살펴봄.

□ 혐오표현 규율의 범위

- 혐오표현은 단일한 법적 정의를 가지고 있지 않음.
- 일반적으로 사회적 약자나 소수자 집단에 대하여 증오 또는 혐오를 선동(incitement to hatred)하거나, 공개적으로 비방·멸시를 하거나, 심각한 편견이나 선입견을 유포하는 표현 등 다양한 성격과 수위의 해악을 동반하는 표현을 통칭하는 용어로 혐오표현을 사용함(국가인권위원회, 2017).
- 혐오표현을 규율하는 국제법 조항이나 외국의 입법례를 보면, 규범의 목적과 규율 수단 및 강도 등에 따라 적용되는 혐오표현의 범주가 다름(이주영, 2015).

외국, 특히 유럽, 캐나다, 호주, 뉴질랜드 등의 입법례를 보면, 인종주의적 혐오표현에 대한 규율을 시작으로 성적지향, 성별정체성, 성, 장애 등을 이유로 한 혐오표현으로 규율의 범위를 확장하는 경향을 확인할 수 있음 (국가인권위원회, 2017).

- 미디어법에서의 혐오표현의 정의를 일례로 살펴보면, 영국 커뮤니케이션청 방송법(Ofcom Broadcasting Code)은 hate speech를 “장애, 민족, 젠더, 젠더재지정, 국적, 인종, 종교 또는 성적지향을 이유로 한 불관용을 기초로 하는 증오를 유포, 선동, 조장, 정당화하는 모든 형태의 표현(all forms of expression which spread, incite, promote or justify hatred based on intolerance on the grounds of disability, ethnicity, gender, gender reassignment, nationality, race, religion, or sexual orientation)”으로 정의함(Section 3).

□ 소셜미디어와 인권, 그리고 민주주의

- 소셜미디어의 대중화 초기에는 소셜미디어가 인권과 민주주의를 진작시킬 것이라는 사회적 기대가 상당했음.
- 실제 제도 언론에 대한 접근이 제한적일 때 소셜미디어는 진실을 알리고 일반 대중의 의견을 널리, 국경을 넘어 전파하는 데 실질적, 잠재적 힘이 있음. “소셜미디어가 아랍의 봄을 만들었다”(2010-2012)라는 진술은 소셜미디어의 그러한 힘을 단적으로 드러냄.
- 하지만 그로부터 몇 년 지나지 않아, 소셜미디어를 통해 나타나는 인권과 민주주의에 대한 부정적 영향에 대해서 우려가 커지기 시작하였음. “소셜미디어는 민주주의를 위협하는가? 페이스북, 구글, 트위터는 양질의 정보로 편견과 허위를 몰아냄으로써 정치를 살리리라 기대했었다. 그런데 무언가 정말 잘못되어 버렸다.” (2017년 11월 4일 이코노미스트지)

□ 소셜미디어와 증오선동 주요 사례

- 2015년 미국 피츠버그 흑인교회에서 백인우월주의자의 총기 난사로 흑인 성직자 및 교인 9명 사망. 수사 내용에 따르면, 총격을 가한 백인우월주의자는 “온라인을 통해 급진화”되었음. (“Prosecutors say Dylann Roof ‘self-radicalized’ online, wrote another manifesto in jail” in *The Washington Post*, 23 August 2016)
- 2018년 미국 유대교 예배당에서 반유대주의자의 총기 난사로 11명 사망. 이 사건의 반유대주의자는 Gab이라는 소셜미디어를 통해 활동하였고 사건 바로 직전에도 추종자들에게 반유대주의 게시물을 남김. Gab은 당시 2년 정도된 소셜네트워크로서, 혐오표현 규범을 두고 있는 페이스북이나 트위터와 달리 자신들은 “표현의 자유”을 더 옹호한다고 표방하였음. Gab은 “백인 민족주의자, 네오나치, 다른 극단주의자들의 천국”이 되었음. (“On Gab, an Extremist-Friendly Site, Pittsburgh Shooting Suspect Aired His Hatred in Full”, in *The New York Times*, 28 October, 2018)
- 2016년-2017년 미얀마에서의 로힝야 민족에 대한 박해와 대학살 과정에서, 군부와 극우 민족주의 불교단체

가 페이스북을 통해 로힝야에 대한 허위정보와 증오를 대중들 사이에 광범위하게 확산시켰음. 이에 대해 미안
마 인권상황에 대한 유엔의 독립조사단은 페이스북이 로힝야에 대한 증오선동 도구로 적극적으로 사용된 것,
페이스북이 그러한 문제를 대응하기 위한 필요한 조치를 제대로 하지 않은 것 등에 대해 2018년 보고서에서
강도 높게 비판을 함.

○ “74. The role of social media is significant. Facebook has been a useful instrument for those seeking to spread hate, in a context where, for most users, Facebook is the Internet. Although improved in recent months, the response of Facebook has been slow and ineffective. The extent to which Facebook posts and messages have led to real-world discrimination and violence must be independently and thoroughly examined. The mission regrets that Facebook is unable to provide country-specific data about the spread of hate speech on its platform, which is imperative to assess the adequacy of its response.” (UN Report of the Independent International Fact-Finding Mission on Myanmar, 12 September 2018, UN Doc. A/HRC/39/64, para. 74.)

- 페이스북도 미얀마에서 폭력을 선동하는 데 플랫폼이 사용되는 데 대해 적절한 조치를 취하지 못했다는 점을 인정함. (“Facebook Admits It Was Used to Incite Violence in Myanmar”, in *The New York Times*, 6 November 2018)
- 2019년 뉴질랜드 크라이스트처치 이슬람 사원에서 총기 난사로 50명이 사망한 사건에서, 당시 총기를 난사한 백인우월주의자는 백인우월주의 내용의 선언문을 온라인에 유포하고 총격상황을 소셜미디어로 중계하였음. (“How the Christchurch Terrorist Attack was Made for Social Media”, CNN, 16 March 2019)

□ 소셜미디어와 증오범죄 관계에 대한 연구

- Kersten Muller and Carlo Schwarz, “Fanning the Flames of Hate: Social Media and Hate Crime”, *Journal of the European Economic Association* (6 October 2020)
- 이 논문에서 연구자들은 2년 동안 독일에서 발생한 난민공격 사건, 3,335건에 대해서 관련성 있을만한 여러 가지 요인들을 체계적으로 분석함. 부(wealth), 인구구성, 극우파 지지 정도, 언론 성향, 난민 수, 과거 증오범죄 사건의 발생 건수, 난민 반대 시위 발생 건수 등. 그전까지 연구되지 않은 새로운 변수로 소셜미디어 이용과의 상관관계가 드러남. 난민 반대 공격이 많이 발생한 지역들은 모두, 지역의 규모나 소득 수준과 관계없이, 페이스북 이용이 평균보다 높았음.

2. 혐오표현과 인권

□ 차별금지과 평등권

- 세계인권선언과 주요 국제인권조약의 가장 기초가 되는 원칙. 세계인권선언 제7조는 “차별이나 차별의 선동에 대해 평등한 보호를 받을 권리”를 인정. 시민적 및 정치적 권리에 관한 국제규약 제26조는 “어떠한 차별도

없이 법의 평등한 보호를 받을 권리”를 확인하면서 이를 위해 “법률은 모든 차별을 금지”하고 “차별에 대하여 평등하고 효과적인 보호를 모든 사람에게 보장”해야 할 국가의 의무를 규정.

- 시민적 및 정치적 권리에 관한 국제규약 제20조 제2항과 모든 형태의 인종차별철폐에 관한 국제협약 제4조 (a)항에 따르면 차별, 적의 또는 폭력의 선동이 될 민족적, 인종적 또는 종교적 증오의 고치를 법률에 의하여 금지할 의무가 있음.
- 인종, 출신 민족, 종교, 성별, 장애, 성적 지향 등을 이유로, 특정집단과 그에 속한 개인의 인간존엄성을 부인하고 공격하는 혐오표현은 대상 집단에 대한 차별과 편견을 강화하는 동시에 그 자체가 차별행위를 구성한다는 것이 유엔인권조약기구의 해석(이주영, 2015).

□ 표현의 자유

- 세계인권선언과 시민적 및 정치적 권리에 관한 국제규약 제19조에서 표현의 자유를 인정하고 있음. 차별 없이 권리를 향유할 수 있어야 한다는 기본 원칙은 표현의 자유에도 마찬가지로 적용.
- 표현의 자유에 대한 제한: 시민적 및 정치적 권리에 관한 국제규약 제19조 제3항은 이 “권리의 행사에는 특별한 의무와 책임”이 따르며, “타인의 권리” 등을 위해 필요한 경우에 한하여 표현의 자유에 일정한 제한을 가할 수 있다고 함.
- 표현의 자유의 내적 가치: ①진리의 발견: 인간은 합리적 존재로서 이성을 통해 진실에 도달하고, 진실이 궁극적으로는 허위를 극복할 수 있다는 것을 전제로 하여, 진실 추구에서 가장 최선은 자유로운 시장 경쟁 속에서 허위들이 걸러지도록 하는 것이라는 ‘사상의 자유시장론’. ‘사상의 자유시장론’은 미국의 수정헌법 제1조(표현의 자유)를 헌법상 가장 우위의 가치로 지탱하는 이론적 토대가 되어 왔음. ②개인의 자아실현과 발전: 개개인들은 간섭 없이 자신의 생각과 감정을 표현할 수 있어야 하고, 이는 자아를 실현하고 발전할 수 있는 능력을 증진시킴. 이를 위해 최대한 폭넓게 표현의 자유를 존중할 것이 요구됨. ③민주주의: 민주주의 정치제도 하에서 사람들은 자유로운 의사 개진을 통해 복수의 의견 가운데서 최선의 정책을 선택함. 의견의 옳고 그름, 좋고 나쁨과 관계없이 표현의 자유를 두텁게 존중하는 것이 민주주의 기본 토대로서 중요(이주영, 2015).
- 표현의 자유에 대한 동등한 권리와 혐오표현: 사회적 소수자들에게는 표현의 자유를 행사할 수 있는 기회가 실질적으로 균등하지 않다는 점, 집단 정체성에 대한 공격은 소수자 개인들의 자아 표출을 위협할 수 있다는 점, 민주주의는 모든 사람이 평등하게 정치적 의사결정 과정에 참여할 수 있음을 전제로 하는데, 혐오표현은 대상 집단과 개인들에 대해 동등한 사회일원으로서의 지위를 부정하여 소수자 집단의 공론장 참여를 위축시키는 효과를 낳는다는 점 등 고려할 필요성(이주영, 2015).
- 사상의 자유시장론과 소셜미디어: 사상의 자유시장론이 소셜미디어에서 유효하기 위해서는 소셜미디어가 다양한 생각과 견해가 자유로이 소통, 토론되는 공간을 형성할 수 있어야 하고, 다양한 정보와 견해의 자유로운 유통을 촉진할 수 있어야 하고, 자정작용을 통해 해악적 표현물은 소수화되고 진리의 발견을 도울 수 있어야 함. 그런데 글로벌 소셜미디어의 현재 상황이 사상의 자유시장론의 전제를 충족하는지 의문. 이용자가 소셜미

디어에서 머무는 시간이 길수록 광고수익이 증가하고 이를 통해 수익을 얻는 것이 글로벌 소셜미디어의 주된 수익 모델이고 이를 위해 소셜미디어는 알고리즘을 통해 이용자별 ‘맞춤형’ 정보를 제공함. 결과적으로 소셜 미디어는 공동의 공적 공간이 아닌 수많은 작은 공적 공간들을 형성하게 되는데, 다양한 정보와 견해의 접근 보다는 유사한 생각을 가진 사람들과의 접촉이 강화되면서 정치적 편향성과 사회적 고정관념이 더 강화되는 경향을 보임. 대항 발화를 통한 편견 해소와 다양성 증진이 어려운 조건을 나타내고 있음(Laub, 2019).

3. 소셜미디어 혐오표현에 대한 접근들

□ 주요 플랫폼의 혐오표현 규제

- 페이스북¹: “혐오발언은 협박과 배척이 있는 환경을 조성하고 경우에 따라 실제 폭력을 조장할 수 있기 때문에 FACEBOOK에서 허용되지 않습니다.” “인종, 민족, 국적, 종교, 성적 지향², 성별 또는 성적 정체성, 심각한 신체적 장애 또는 질병과 같이 보호되는 특성을 이유로 타인을 직접적으로 공격하는 것을 혐오발언으로 정의”
- 유튜브³: “YouTube에서는 증오심 표현이 허용되지 않습니다. 다음과 같은 특성을 문제 삼아 개인이나 집단 에 대한 폭력 또는 혐오감을 조장하는 콘텐츠는 삭제됩니다. : 연령, 계급, 장애, 민족, 성 정체성 및 성 표현, 국적, 인종, 이민 신분, 종교, 성별, 성적 지향, 큰 폭력 사건의 피해자와 그 친인척, 군필 여부
- 트위터⁴: “혐오 행위: 인종, 민족, 국적, 신분 계급, 성적 성향, 성별, 성 정체성, 종교, 나이, 장애, 심각한 질병 등을 이유로 타인에게 폭력을 조장하거나, 직접적인 공격을 가하거나, 위협해서는 안 됩니다. 또한 이러한 범 주의 특성을 바탕으로 다른 사용자에게 피해를 주는 것이 주된 목적인 계정은 허용되지 않습니다.” “혐오 이 미지 및 표시 이름: 프로필 이미지 또는 프로필 헤더에 혐오 이미지나 상징을 사용할 수 없습니다. 또한, 사용 자 아이디, 표시 이름 또는 프로필 자기소개를 사용하여 특정 대상을 괴롭히거나 사람, 단체 또는 보호 대상 범 주에 속하는 이들에 대한 증오심을 표현하는 등의 가학적인 행위에 참여해서도 안 됩니다.”
- 인스타그램⁵: “인종, 민족, 국적, 성별, 나이, 성적정체성, 성적 지향, 종교, 장애 또는 질병을 이유로 누군가를 공 격하거나 폭력을 조장하는 행위는 결코 허용되지 않습니다. 단, 이에 반대하거나 경각심을 일깨우기 위한 목 적으로 공유된 혐오 발언은 의도가 명확한 경우에 한하여 허용될 수 있습니다.”

1. https://m.facebook.com/communitystandards/hate_speech/ (최종접속일: 2020.10.30.)

2. 페이스북과 인스타그램에 게시되어 있는 혐오표현 정책문 국문 번역본은 sexual orientation을 모두 성적 취향으로 하고 있음. 각국의 사회적·언어적 맥락 속에서 혐오표현을 이해할 필요성을 환기시키는 대목.

3. <https://support.google.com/youtube/answer/2801939?hl=kr> (최종접속일: 2020.10.30.)

4. <https://help.twitter.com/ko/rules-and-policies/hateful-conduct-policy> (최종접속일: 2020.10.30.)

5. <https://www.facebook.com/help/instagram/477434105621119/> (최종접속일: 2020.10.30.)

□ 페이스북과 혐오표현: 최근 논쟁

- CEO 마크 저커버그 (2019) “‘허위정보’나 ‘가짜뉴스’를 담고 있는 정치광고라도 제한 두지 않겠다.”고 표방하였다가 비판을 받음. 맥락: 2016년 영국의 브렉시트 국민투표 과정에서 페이스북의 이용자 정보가 브렉시트 지지 정치 컨설팅업체에 유출되어 브렉시트 지지그룹에서 페이스북 이용자를 대상으로 한 캠페인에 활용되는 등 페이스북이 민주주의를 교란하는 데 이용된다는 비판이 제기되어 왔음.
- 미 대통령 트럼프 “WHEN THE LOOTING STARTS, THE SHOOTING STARTS” 게시 (2020.5.29 Facebook, Twitter, Instagram). 트위터: 폭력규칙 위반 안내문 읽어야 게시물 볼 수 있도록 조치. 페이스북 & 인스타그램: 자체규정 위반하는 게시물임에도 규정 적용하지 않음.
- STOP hate for profit 운동: 페이스북에 대한 광고 보이콧 조직(2020.7)
- 케노샤 백인우월주의 무장단체 페이지(2020.8): 폭력선동 신고 있었으나 페이스북은 제때 조치를 취하지 않음. 2명 사망.
- Facebook이 의뢰한 독립적인 감사 보고서가 발간됨(2020.8): “민권의 심각한 후퇴 초래” “표현의 자유를 옹호하는 것 좋은 일. 하지만 모두에게 적용되어야.” 권력있는 정치인들에게 동일한 규칙 적용하지 않는 것 비판 내용 담음. “a hierarchy of speech is created that privileges certain voices over less powerful voices.”

□ 주요 국가들의 정책과 관련 규범

- 미국: 수정헌법 제1조 표현의 자유: 적용범위 국가행위자. 소셜미디어 기업은 해당되지 않음. 통신품위법 (1996) 제230조: 이용자들의 온라인 표현에 대하여 인터넷 기업에 대해 면책 규정. 트럼프의 제230조 면책 제거하는 행정명령(2020.5.29): “온라인 플랫폼의 검열에 대한 제한 필요”하다는 취지에서. 민주당의 조 바이든, 버니 샌더스, 마크 해리스 등 온라인 플랫폼 기업들이 유통하는 정보가 폭력 조장, 허위정보 유포 등 해악을 끼칠 때 책임져야 한다는 점에서 제230조 삭제 또는 개정 입장.
- 유럽평의회 「컴퓨터를 통한 인종주의, 외국인혐오적 행위의 처벌에 관한 사이버 범죄 협약 추가의정서」 (2003): 온라인상 인종주의, 외국인혐오적 내용의 전파(제3조), 인종주의적 모욕 행위(제5조), 제노사이드 또는 인도에 반하는 범죄 부정, 심각한 축소, 승인, 정당화(제6조) 등의 행위에 대해 형사 규제 도입, 처벌하도록 함.
- 유럽연합 「형법을 통한 인종주의와 외국인 혐오의 특정 형태와 표현에 대응하기 위한 기본 결정」 (2008): 인종, 피부색, 종교, 혈통 또는 출신 국적 및 민족에 따라 규정된 집단 또는 그 일원에게 공개적으로 폭력과 증오를 선동하는 고의적 행위를 각 회원국 내에서 형사범죄로 처벌하도록 하고, 각국 법원들 사이의 관련 협력도모하는 규정을 담고 있음.
- 유럽연합 「불법 온라인 혐오표현 대응 행동기준」 (2016): 유럽연합 회원국: 「형법을 수단으로 인종주의와 외국인 혐오의 특정 형태와 표현에 대응하기 위한 기본 결정」 (2008)에 따라 공개적인 증오·혐오·선동 법으

로 규제. 배경: 온라인상에서도 불법 혐오표현 신고에 대해 신속하게 대응해야 할 필요성. 온라인 플랫폼, 소셜 미디어 플랫폼, 기업 내부 규정 또는 커뮤니티 가이드라인에 폭력과 증오 조장, 선동 금지 명확히 하도록 함. '유효한' 삭제 요청에 대해 규정/가이드라인에 따라 검토, 최대한 24시간 내 불법 혐오표현 삭제 신고 요청 검토, 검토 결과에 따라, 콘텐츠 삭제 또는 접속 차단 조치하도록 함. 규정과 커뮤니티 가이드라인에 대한 정보를 제공하고, 인식을 제고하도록 함. 시민사회 및 전문가 그룹과 협력을 강화하도록 함. 2016년 페이스북, 마이크로소프트, 트위터, 유튜브 참여. 2018년 인스타그램, 구글+, 스냅챗, 데일리모션 참여. 2019년 Jeuxvideo.com 추가 참여.

- 독일: 「형법」 제130조 민족, 인종, 종교, 국적, 성적지향 등을 이유로 특정 집단에 대해 증오·혐오를 선동하는 행위, 경멸, 악의적 중상 혹은 명예훼손으로 인간 존엄성을 침해하는 행위 처벌. 「형법」 제185조, 제187조 (모욕, 명예훼손죄) 집단모욕죄, 집단명예훼손죄 인정. 「일반평등대우법」 제3조 언어를 통한 차별적 괴롭힘과 차별 지시, 차별행위로 규정. 「네트워크 법집행법」(NetzDG법, 2018년 1월 시행) 소셜미디어 제공 기업에 불법 온라인 혐오표현콘텐츠 삭제 또는 접속 차단 조치 의무. 의무 위반: 벌금 최대 5천만 유로. 투명성 보고 의무 부과.

4. 소셜미디어에서의 혐오표현 규제

□ 우려

- 콘텐츠의 과잉 차단 위험
- 기업에 의한 검열이 제도화될 것에 대한 우려

□ 과제

- 국내 소재 여부와 관계없이 소셜미디어 기업의 인권 존중과 보호에 대한 책임 인식 강화할 필요 있음.
- 주요 글로벌 소셜미디어 기업의 혐오표현 자체 규정 관련, 개선 과제: ①소셜미디어 내 처리 절차의 투명성 증진, ②부당하게 삭제, 차단된 콘텐츠에 대한 이의제기 절차 강화 및 게시물 작성자에 대한 통지 제도 강화, ③정부, 소셜미디어 기업, 시민사회, 혐오표현으로 영향을 받는 사회적 약자·소수자 그룹과 지속적 협의.

III. 온라인 혐오표현에 대한 일반시민 인식 조사

1. 조사개요

□ 조사목표

- 한국사회 혐오와 차별이 디지털 미디어를 통해 가시화되면서 사회문제로 드러나고 있음. 온라인상 혐오의 대상은 인종, 민족, 장애, 여성, 성적지향을 넘어서 코로나 국면을 거치면서 특정 지역과 종교로 확장됐으며, 혐오와 차별을 조장하는 표현의 방식도 지속적으로 교묘하게 변화해가고 있음. 혐오표현에 대한 노출이 증가되면서 둔감해지는(desensitization) 현상까지 나타나고 있는 것으로 보임.
- 온라인 혐오표현에 대한 규제 필요성에 대한 공감대가 형성된 가운데 법제도적, 정책적, 기술적, 실천적 해법 마련을 위해 학계, 정책기관, 시민단체, 산업계 등 다양한 주체들이 부단히 노력하고 있음. 그러나 혐오표현이 무엇이며, 실제 일반시민들의 인식이 어떠한지, 그에 대해 어떤 대처방안이 필요한지에 대한 논의가 부족해보임. 따라서 온라인 혐오표현에 대한 일반시민의 인식을 살펴볼 필요가 있음.

□ 조사대상

- 전문조사기관 마켓링크에 의뢰하여 2020년 8월 26일부터 31일까지 6일 동안 혐오표현에 대해 들어본 적 있는 일반인을 대상으로 조사를 실시함.
- 조사기관에서 보유하고 있는 패널 가운데 16,542명에게 설문 URL이 포함된 이메일을 발송한 뒤 응답대상자가 조사 페이지에 접속하여 응답하는 방식으로 진행함. 성별, 연령, 거주지역을 고려한 비례층화표집(proportionate stratified sampling)을 했으며, 1,022명(응답률 6.2%)의 응답을 수집하여 분석에 활용함.
- 응답자 평균 연령은 44.58세($SD = 12.89$), 성별분포는 남성 515명(50.4%), 여성 507명(49.6%), 연령별 분포는 20대 186명(18.3%), 30대 193명(18.9%), 40대 230명(22.5%), 50대 242명(23.7%), 60대 이상 171명(16.7%)으로 나타남. 지역별로는 경기 263명(25.7%), 서울 198명(19.4%), 부산 69명(6.8%), 경남 67명(6.6%), 인천 59명(5.8) 등임.
- 그 외 교육수준은 고졸 이하 196명(19.2%), 대재·대졸 710명(69.5%), 대학원재·대학원졸 116명(11.4%)이며, 종교는 무교 570명(55.8%), 개신교 203명(19.9%), 불교 147명(14.4%), 천주교 79명(7.7%), 기타 13명(1.3%) 등으로 나타났고, 정치성향은 진보 307명(30.0%), 중도 505명(49.4%), 보수 210명(20.5%)이었음.

□ 조사내용

- 혐오표현에 대한 판단(문장): 혐오표현의 대상을 카키넨, 옥사넨, 그리고 래새넨(Kaakinen, Oksanen, & Räsänen, 2018), 국가인권위원회(2016), 문화체육관광부(2018)를 참조하여 인종·민족·국적, 종교, 정치성

향, 출신지역, 성별, 장애, 성적지향, 특정연령층의 8개 범주로 나누고, 혐오표현 관련 문헌(논문, 보고서, 교육 자료 등)과 온라인 커뮤니티, 뉴스 댓글, SNS 댓글을 참조하여 각 혐오대상별 표현 진술문을 본 연구에 맞게 바꾸어 대상별 4개(모욕형 2개, 선동형 2개), 총 32개 진술문을 설문문항으로 활용함. 소랄, 빌레비츠, 그리고 위니위스키(Soral, Bilewicz, & Winiewski, 2018)와 코완과 호지(Cowan & Hodge, 1996)의 혐오표현 측정방식을 변용해 7점 척도(1: 혐오표현이 아니다 ~ 7: 혐오표현이다)로 측정.

※ 국가인권위원회(2019) 연구에서 분류한 바와 같이 모욕, 비하, 멸시, 위협하는 표현을 모욕형으로, 차별·폭력의 선전하거나 선동하는 표현을 선동형으로 분류했음

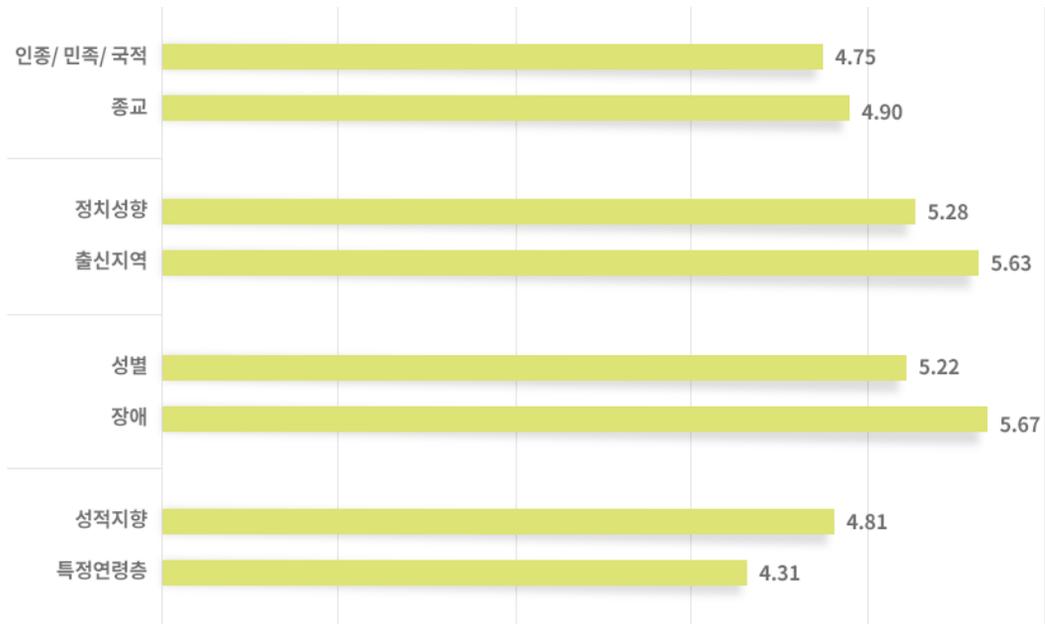
- 혐오표현에 대한 판단(단어): 온라인에서 많이 쓰는 혐오표현 43개 표현을 선별해 5점 척도(1: 혐오표현이 아니다 ~ 5: 혐오표현이다)로 측정. 해당 단어를 전혀 알지 못하는 경우 응답에서 제외할 수 있도록 함.
- 혐오표현에 대한 개념적 판단: 개인·집단에 대한 욕설, 인간 이하 표현, 동물·곤충·질병으로 표현 등 13개 문항을 5점 척도로 측정.
- 혐오표현에 대한 노출경험: 국가인권위원회(2016)와 문화체육관광부(2018)의 혐오표현을 접한 경험 문항을 변용해, 최근 6개월 사이 온라인에서 인종·민족·국적, 종교, 정치성향, 출신지역, 성별, 장애, 성적지향, 특정 연령층의 8개 범주에 대한 혐오표현을 얼마나 접했는지 나누어 5점 척도(1: 거의 접하지 못했다 ~ 5: 매우 자주 접했다)로 측정.
- 혐오표현 노출에 대한 개인적 영향: 온라인 혐오표현 노출에 따른 감정 또는 행동을 5점 척도(1: 전혀 그렇지 않다 ~ 5: 매우 그렇다)로 측정.
- 혐오표현 생산 경험: 혐오표현 생산 경험과 빈도, 이유를 각각 측정.
- 혐오표현 피해자 인식: 온라인에서 누가 혐오표현의 피해를 받는다고 생각하는지 명목형 문항에 대해 1~3순위를 선택하도록 하는 방식으로 측정.
- 혐오표현 규제 태도: 국가인권위원회(2016), 문화체육관광부(2018), 그리고 박아란과 양정애(2016)의 문항을 활용해 혐오표현에 대한 규제 태도, 피해대상별 규제 필요성, 대응방안에 대해 각 5점 척도로 측정

2. 온라인 혐오표현에 대한 판단

□ 피해대상 집단 및 표현별 혐오표현 인식 수준 (그림 1, 2-1, 2-2)

- 온라인 혐오표현에 대한 판단을 8개 피해대상 집단별로 나눠 살펴본 결과, 장애에 대한 혐오표현(M = 5.67, SD = 1.28)을 혐오표현에 가장 가깝게 인식하고 있었으며, 이어 출신지역에 대한 혐오표현(M = 5.63, SD = 1.25), 정치성향(M = 5.28, SD = 1.44), 성별(M = 5.22, SD = 1.30) 등의 순으로 높게 나타났음.
- 특정연령층에 대한 혐오표현을 혐오표현이라고 인식하는 정도가 가장 낮게 나타남(M = 4.31, SD = 1.41).

※참고: M=평균, SD=표준편차를 의미함



N=1,022. 7점 척도

그림1. 온라인 혐오표현에 대한 판단

- 각 문장 표현별로 나누어 살펴본 결과, 정치성향, 출신지역, 장애에 대한 혐오표현은 모두 7점 가운데 5점이 넘게 나타났음(M = 5.18 ~ 5.84). 대체로 4점 이상으로 각 표현을 혐오표현으로 인지하고 있었으나, 특정연령층 가운데 젊은층(초등학생, 청소년)에 대한 표현은 다소 낮게 나타남(M = 3.68 ~ 4.13).

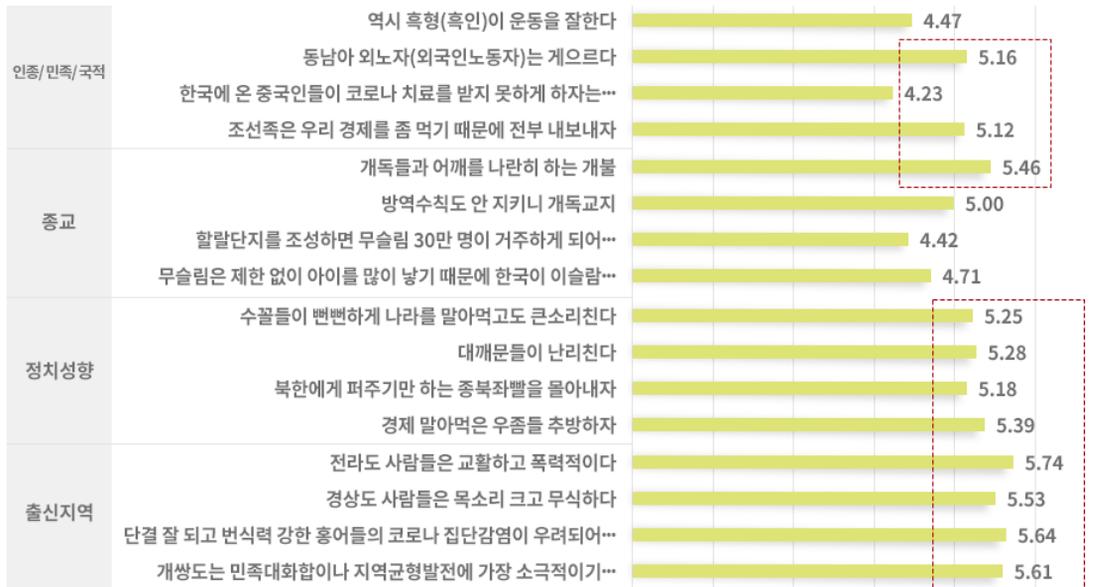


그림2-1. 온라인 혐오표현에 대한 판단 (혐오표현별) 1

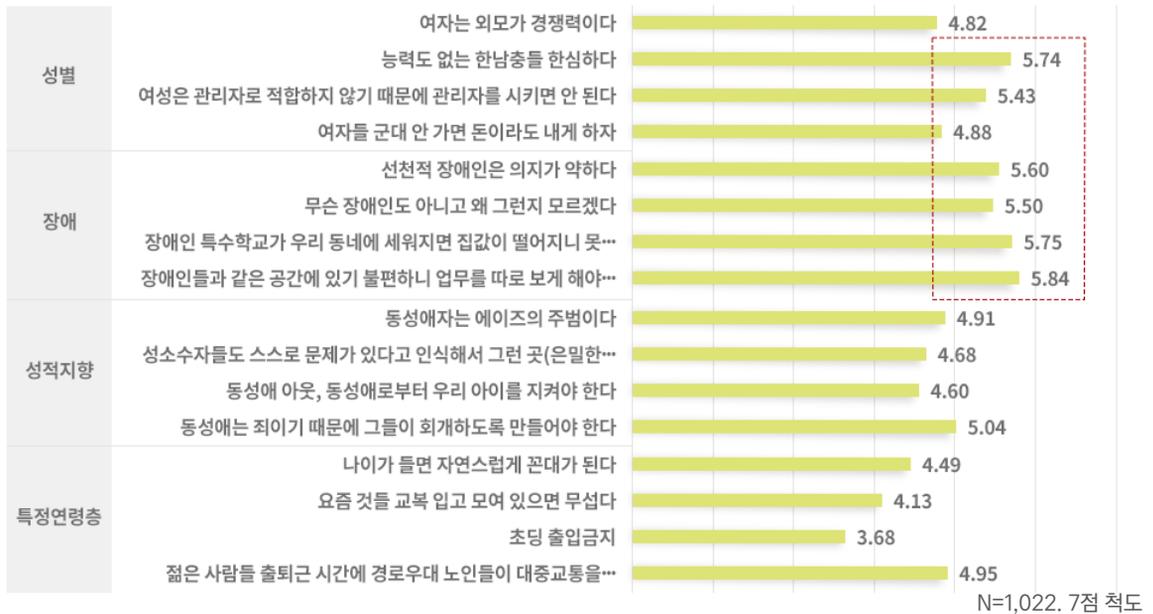


그림2-2. 온라인 혐오표현에 대한 판단 (혐오표현별) 2

□ 성별에 따른 인식 차이 (그림 3-1, 3-2, 4)

- 온라인 혐오표현에 대한 성별에 따른 인식의 차이가 있는지 살펴보기 위해 독립표본 t-검정(independent t-test)을 실시한 결과, 정치성향, 성별, 장애, 성적지향에 대한 혐오표현의 인식에 차이가 있었으며, 여성이 남성에 비해 보다 높게 인식하는 것으로 나타남.

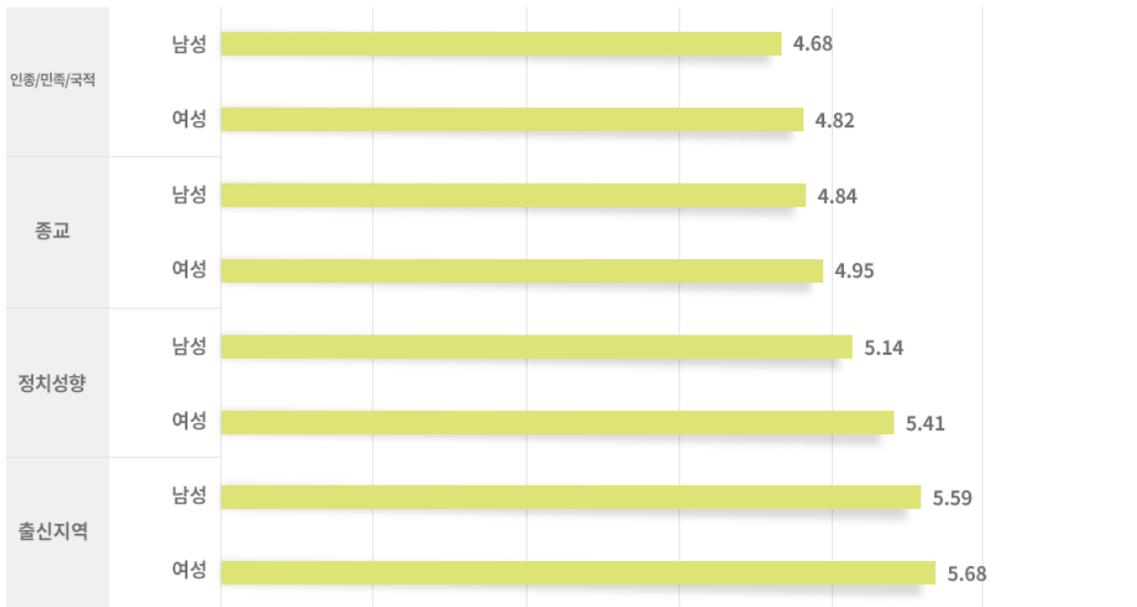


그림3-1. 피해대상 집단별 혐오표현에 대한 판단 차이 (성별) 1

- 성별에 대한 혐오표현에 대한 남녀 간 차이가 가장 큰 것으로 나타났는데, 여성은 5.49(SD = 1.27), 남성은 4.95(SD = 1.29)였음. 또한 정치성향에 대한 혐오표현에 있어서도 여성 5.41(SD = 1.47), 남성 5.14(SD = 1.39)로 차이가 있었음. 한편 특정연령층에 대한 혐오표현에 대한 인식은 남성 4.37(SD = 1.38), 여성 4.25(SD = 1.43)로 유일하게 남성이 여성보다 혐오표현에 대한 인식의 정도가 높은 것으로 나타났으나 통계적으로 유의미하지는 않았음.

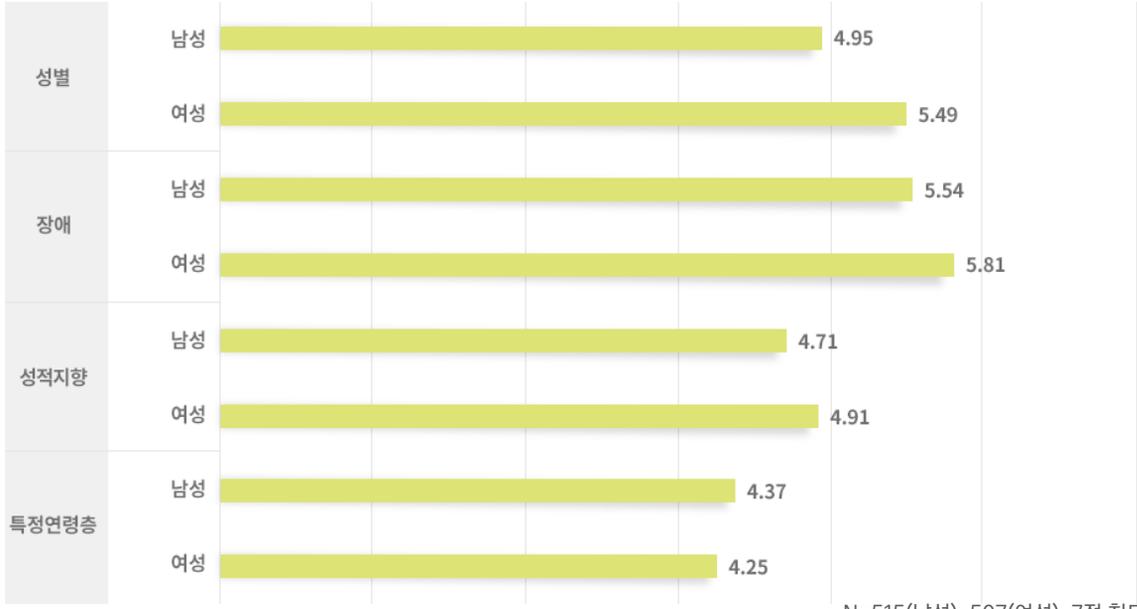


그림3-2. 피해대상 집단별 혐오표현에 대한 판단 차이 (성별) 2

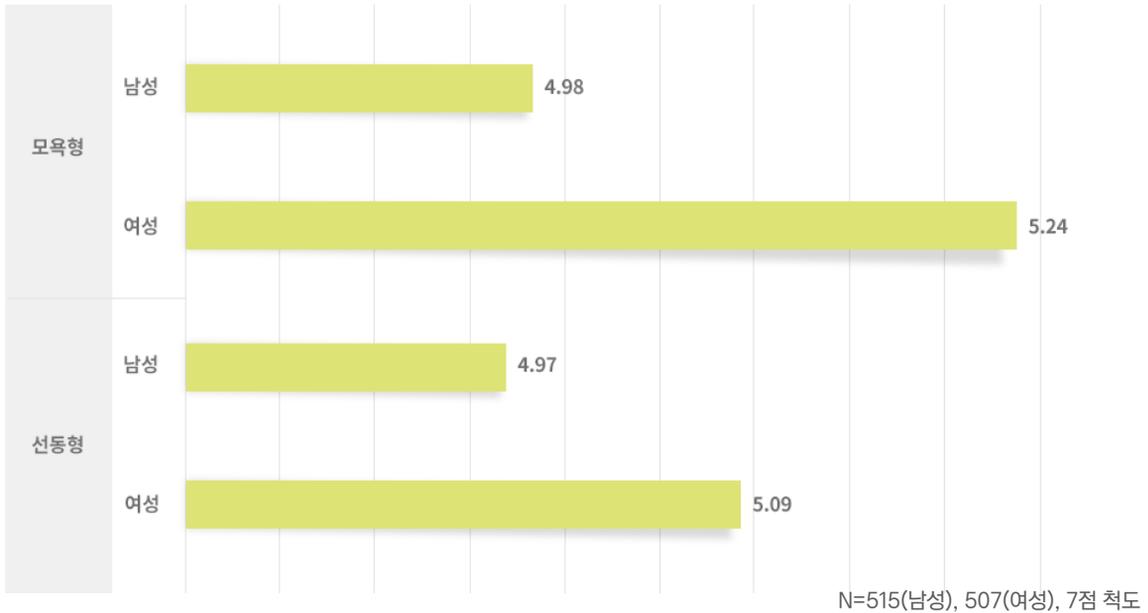


그림4. 표현유형별 혐오표현에 대한 판단 차이 (성별)

- 온라인 혐오표현을 모욕형과 선동형으로 나누어 성별에 따른 인식의 차이가 있는지 살펴본 결과, 모욕형 혐오표현은 여성이 5.24(SD = 1.02), 남성이 4.98(SD = 4.98)로 여성이 남성에 비해 상대적으로 혐오표현에 대한 인식의 수준이 높은 것으로 나타났으며, 이러한 차이는 통계적으로 유의미했음. 선동형 혐오표현도 여성이 5.09(SD = 1.10), 남성이 4.97(SD = 1.10)로 혐오표현에 대한 인식의 수준은 여성이 더 높은 것으로 나타났으며, 이러한 차이는 제한적 수준에서 통계적으로 유의미한 것으로 나타남.

□ 연령대에 따른 인식 차이 (그림 5-1, 5-2, 6)

- 온라인 혐오표현에 대한 연령대별 인식의 차이가 있는지 살펴보기 위해 일원배치분산분석(one-way ANOVA)을 실시한 결과, 연령대별로 정치성향, 출신지역, 성별, 장애, 성적지향에 대한 혐오표현의 인식이 다른 것으로 나타났음.
- 전반적으로 젊은층이 혐오표현에 대한 인식의 정도가 높은 것으로 나타났음. 가령, 출신지역에 대한 혐오표현은 20대가 5.96(SD = 1.24), 30~40대가 5.63(SD = 1.26), 그리고 50~60대가 5.49(SD = 1.23)로 나타났으며, 이러한 차이는 통계적으로 유의미한 것으로 나타났음.

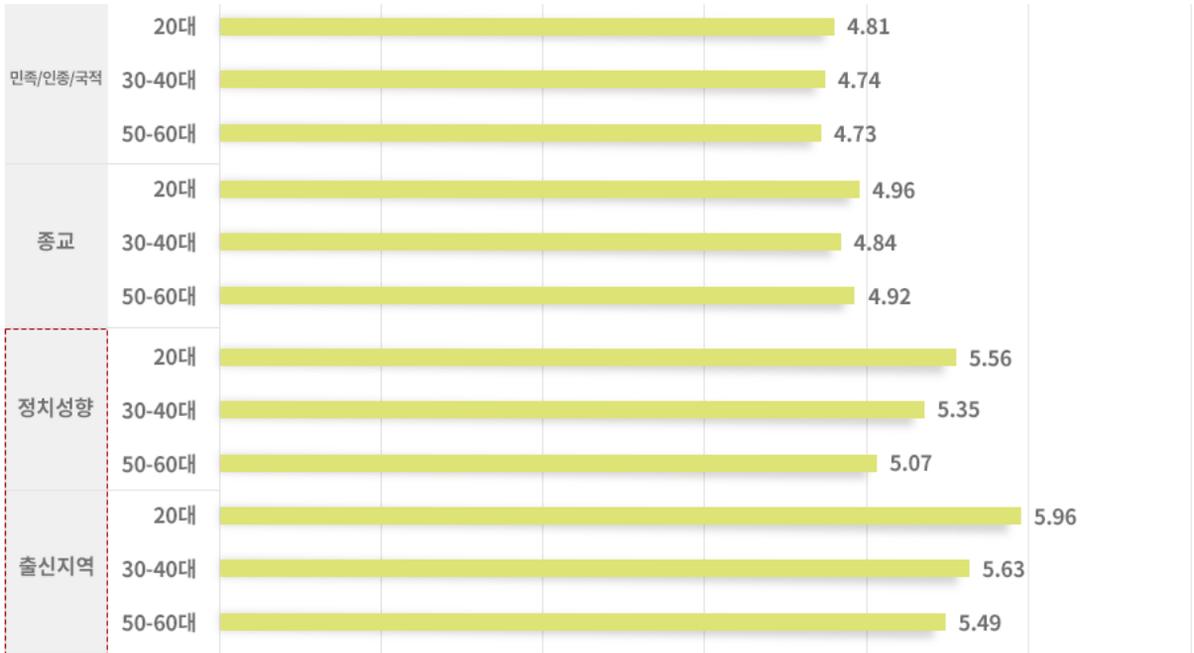
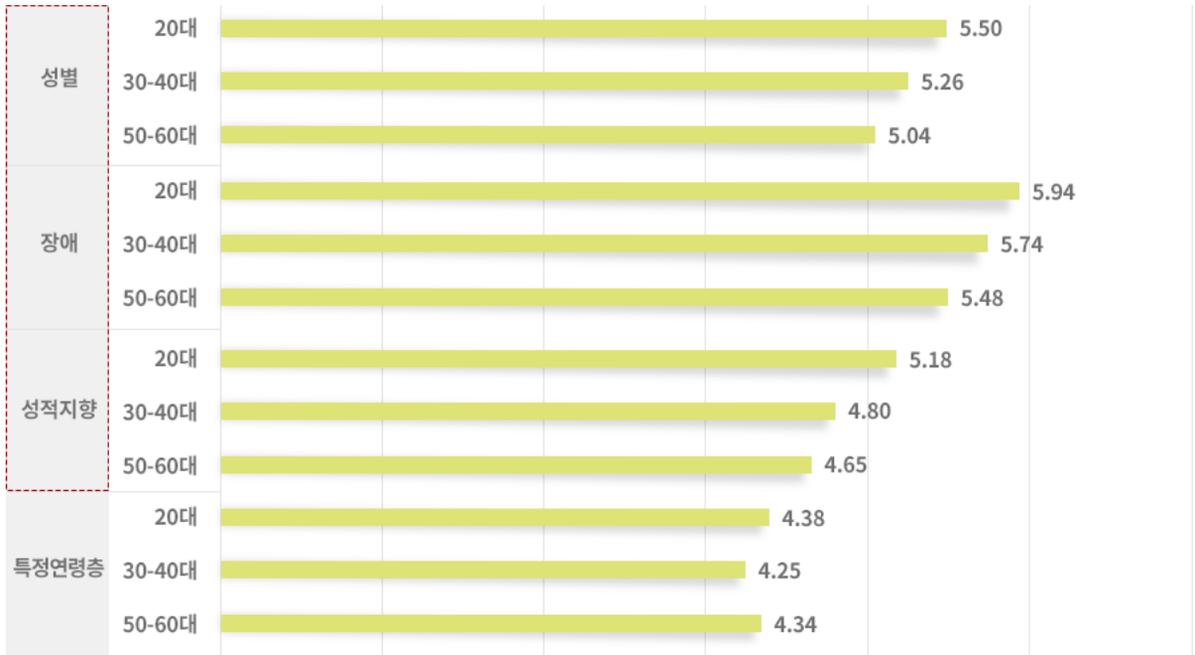


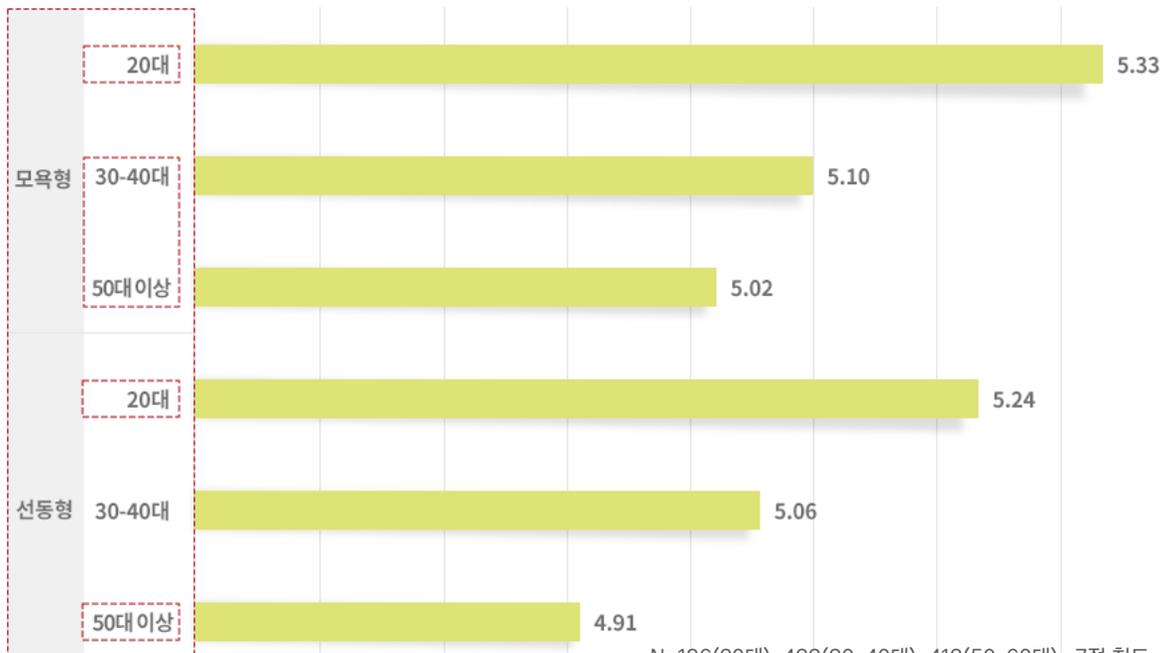
그림5-1. 피해대상 집단별 혐오표현에 대한 판단 차이 (연령대별) 1



N=186(20대), 423(30-40대), 413(50-60대). 7점 척도

그림5-2. 피해대상 집단별 혐오표현에 대한 판단 차이 (연령대별) 2

- 온라인 혐오표현을 모욕형과 선동형으로 나누어 연령대에 따른 인식의 차이가 있는지 살펴본 결과, 통계적으로 유의미한 차이가 있는 것으로 나타남.



N=186(20대), 423(30-40대), 413(50-60대). 7점 척도

그림6. 표현유형별 혐오표현에 대한 판단 차이 (연령대별)

- 모욕형 혐오표현에 대해서는 20대가 5.33(SD = 1.06), 30~40대가 5.10(SD = 1.07), 그리고 50~60대가 5.02(SD = 1.03)로 나타났으며, 사후검증결과 30~40대와 50~60대 간에는 유의미한 차이는 없었음. 즉 20대가 다른 두 연령대에 비해 모욕형 혐오표현에 대한 인식이 상대적으로 높은 것으로 나타남.
- 선동형 혐오표현에 대해서는 20대가 5.24(SD = 1.13), 30~40대가 5.06(SD = 1.09), 그리고 50~60대가 4.91(SD = 1.09)로 나타났으며, 사후검증결과 20대와 50~60대 간의 차이만 유의미한 것으로 나타났음. 즉 20대가 50~60대에 비해 선동형 혐오표현에 대한 인식이 상대적으로 높은 것으로 나타남.

□ 교육수준에 따른 인식 차이 (그림 7-1, 7-2, 8)

- 온라인 혐오표현에 대해 교육수준에 따른 인식의 차이가 있는지 살펴보기 위해 일원배치분산분석(one-way ANOVA)을 실시한 결과, 교육수준별로 인종·민족·국적, 종교, 정치성향, 출신지역, 장애, 특정연령층에 대한 혐오표현 인식에 차이가 있는 것으로 나타났음
- 전반적으로 교육수준이 높을 때 혐오표현에 대한 인식의 정도도 높은 것으로 나타났음. 인종·민족·국적의 경우 대학원졸(재학포함) 5.00(SD = 1.38), 대졸(재학포함) 4.76(SD = 1.41), 그리고 고졸이하 4.53(SD = 1.46) 순으로 나타났으며, 교육수준에 따른 차이는 통계적으로 유의미했음. 사후검증 결과, 대학원졸이 고졸 이하보다 통계적으로 높은 인식을 보이거나, 대학원졸과 대졸, 대졸과 고졸이하 간 차이는 유의미하지 않은 것으로 나타남.

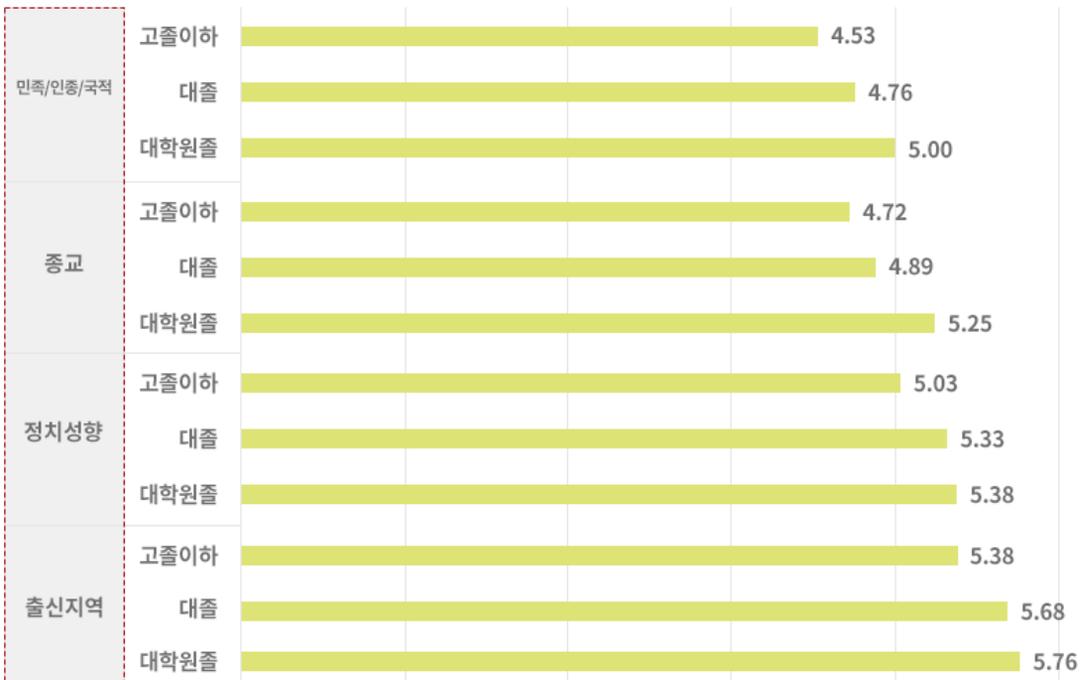
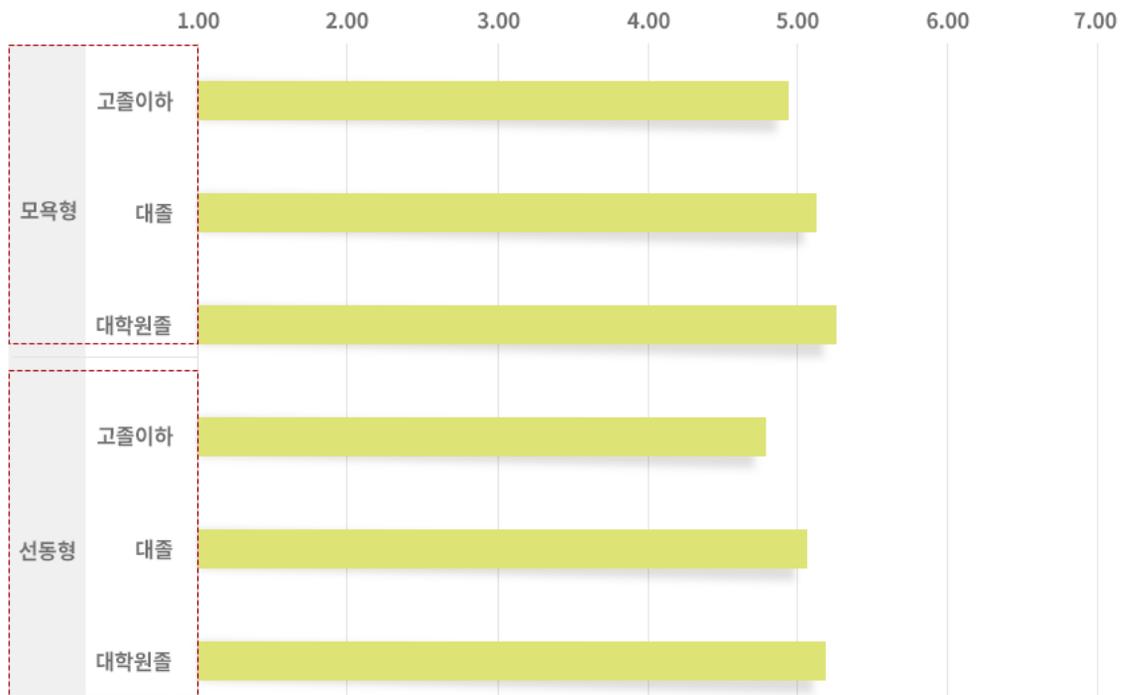


그림7-1. 피해대상 집단별 혐오표현에 대한 판단 차이 (교육수준) 1



N=196(고졸이하), 710(대재·대졸), 116(대학원재·대학원졸). 7점 척도

그림7-2. 피해대상 집단별 혐오표현에 대한 판단 차이 (교육수준) 2



N=196(고졸이하), 710(대재·대졸), 116(대학원재·대학원졸)

그림8. 표현유형별 혐오표현에 대한 판단 차이 (교육수준)

- 온라인 혐오표현을 모욕형과 선동형으로 나누어 교육수준에 따른 인식의 차이가 있는지 살펴본 결과, 통계적으로 유의미한 차이가 있는 것으로 나타남.
- 모욕형 혐오표현에 대해서는 대학원졸(재학포함) 5.26(SD = 1.10), 대졸(재학포함) 5.13(SD = 1.05), 그리고 고졸이하 4.94(SD = 1.03)로 나타났으며, 이들 간 차이에 대한 사후검증 결과, 대학원졸과 고졸이하 간 차이가 있었으며 그 외에는 통계적으로 유의미한 차이는 없었음.
- 선동형 혐오표현에 대해서는 대학원졸(재학포함) 5.19(SD = 1.09), 대졸(재학포함) 5.07(SD = 1.10), 그리고 고졸이하 4.80(SD = 1.10)로 나타났으며, 이들 간 차이에 대한 사후검증 결과, 대졸과 대학원졸이 고졸이하에 비해 인식이 높은 것으로 나타났으나 대졸과 대학원졸 간에서는 통계적으로 유의미한 차이가 없었음.

□ 온라인 혐오표현(단어)에 대한 판단 (그림 9)

- 온라인에서 많이 사용하는, 혐오의 의미를 담고 있는 단어들에 대해 응답자들이 어떻게 인식하고 있는지 살펴 보기 위해 온라인 혐오표현으로 많이 사용되고 있는 43개 단어를 선정하여 조사한 결과, 전반적으로 4점 이상으로 혐오표현에 가깝게 인식하고 있는 것으로 나타남.
- 지역혐오 단어인 개쌍도(M = 5.84, SD = 1.38), 전라디언(M = 5.58, SD = 1.60), 멍청도(M = 5.57, SD = 1.58)이 다른 단어에 비해 사람들이 혐오표현에 더욱 가깝게 인식하고 있는 반면, 민족·인종·국적에 대한 혐오 단어인 쫄꺨(M = 4.44, SD = 2.05), 흑형(M = 4.65, SD = 1.94), 쪽바리(M = 4.77, SD = 2.13), 짱깨(M = 4.65, SD = 1.94)는 혐오표현으로 인식하는 정도가 상대적으로 낮았음.
- 또한 개인이나 집단의 특성을 벌레로 표현한 일부 단어들에 대한 인식도 상대적으로 낮게 나타났음. 예를 들어, 급식충(M = 4.91, SD = 1.81), 학식충(M = 4.64, SD = 1.89), 설명충(M = 4.41, SD = 1.90), 진지충(M = 4.44, SD = 1.86), 라떼충(M = 4.70, SD = 1.84)을 혐오표현으로 인식하는 정도가 상대적으로 낮게 나타남. 다만, 성적지향(똥꼬충), 노년충(틀딱충), 남성(한남충)을 벌레로 빗대어 표현한 단어들에 대해서는 상대적으로 높게 나타남.

□ 온라인 혐오표현에 대한 개념적 판단 (그림 10)

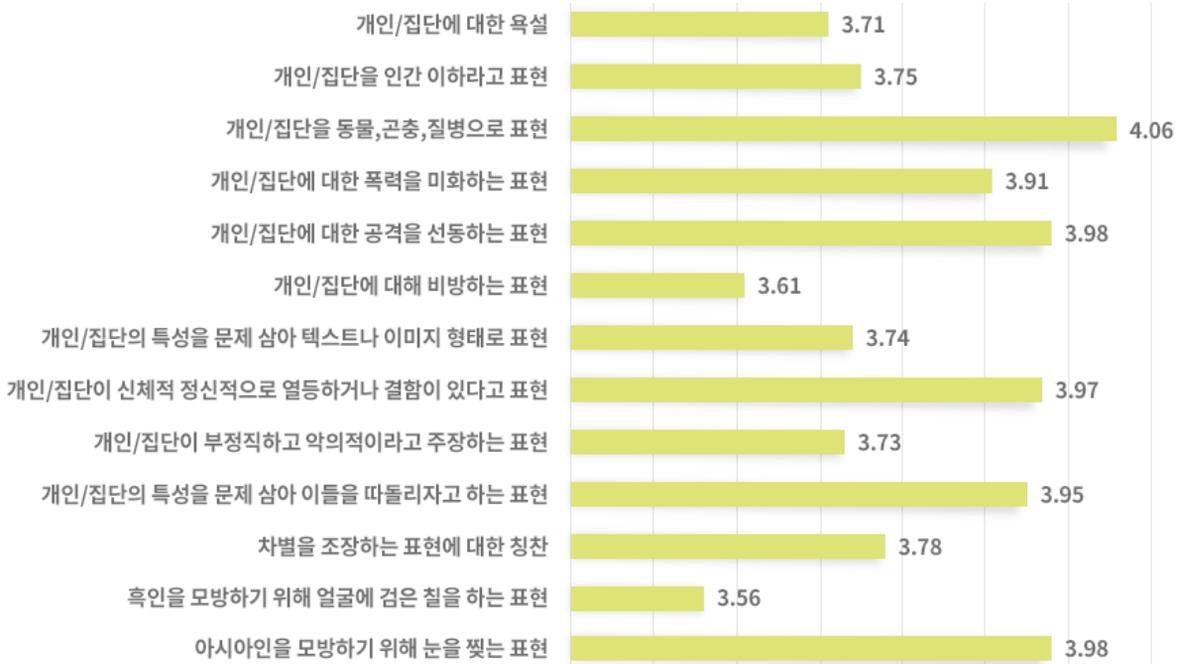
- 온라인 혐오표현에 대한 여러 기준들에 대해 응답자들이 얼마만큼 혐오표현이라고 인식하는지 살펴본 결과, 대체로 3점 이상 넘는 것으로 나타나 혐오표현에 가깝게 인식하고 있는 것으로 나타남.
- 여러 기준 가운데 특히 개인·집단을 동물, 곤충, 질병으로 표현하는 것(M = 4.06, SD = 1.01)을 혐오표현에 가깝게 인식하고 있었으며, 개인·집단에 대한 공격을 선동하는 표현(M = 3.98, SD = 0.96), 개인·집단이 신체적·정신적으로 열등하거나 결함이 있다고 표현(M = 3.97, SD = 1.01), 개인·집단의 특성을 문제 삼아 이들을 따돌리자고 하는 표현(M = 3.95, SD = 1.01) 등의 순으로 높게 인식하고 있었음.
- 한편 행위적 표현으로서 흑인을 모방하기 위해 얼굴에 검은 칠을 하는 표현(M = 3.56, SD = 1.19)을 혐오표현으로 인식하는 정도는 상대적으로 가장 낮게 나타났으며, 이에 비해 아시아인을 모방하기 위해 눈을 찡는



N=1,022. 7점 척도

그림9. 혐오 의미를 담고 있는 단어에 대한 혐오표현 판단

표현(M = 3.98, SD = 1.01)은 다소 높게 나타나, 혐오표현의 대상과 자신의 거리감에 따라 자신과 멀다고 느끼는 경우 혐오표현이라고 느끼는 정도가 낮아진다는 것을 탐색해볼 수 있음.



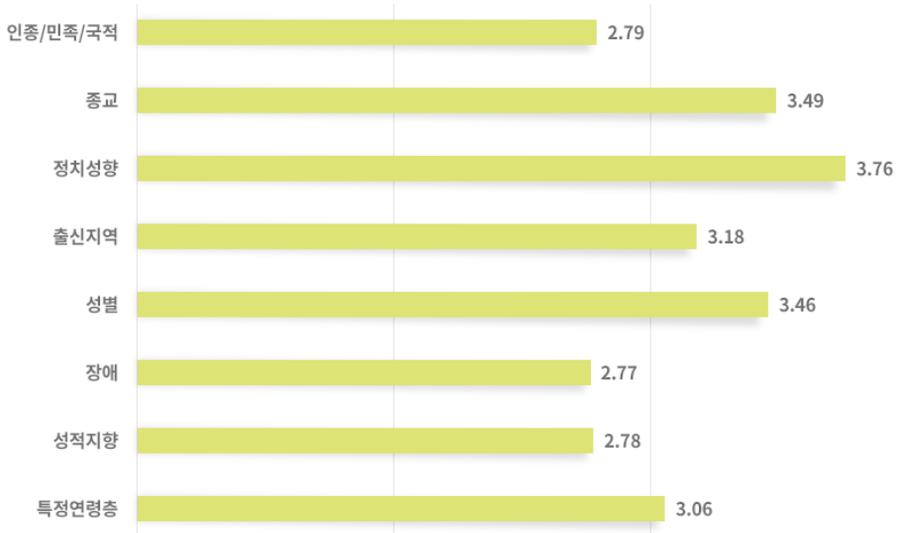
N=1,022. 5점 척도

그림10. 온라인 혐오표현에 대한 개념적 판단

3. 온라인 혐오표현에 대한 노출 및 생산

□ 온라인 혐오표현 노출빈도 (그림 11, 12-1, 12-2, 13-1, 13-2)

- 응답자들의 온라인 혐오표현 노출량을 살펴본 결과, 정치성향(M = 3.76, SD = 1.19), 종교(M = 3.49, SD = 1.29), 성별(M = 3.46, SD = 3.46), 출신지역(M = 3.18, SD = 1.29), 특정연령층(M = 3.06, SD = SD = 1.18)에 대한 혐오표현 등의 순으로 높게 나타났음. 그 외 인종·민족·국적, 장애, 성적지향에 대한 혐오표현은 3점 보다 낮게 나타나 노출빈도가 높지 않았음.
- 성별에 따라 온라인 혐오표현에 대한 노출량의 차이가 있는지 살펴보기 위해 독립표본 t 검증을 실시한 결과, 인종·민족·국적, 정치성향, 출신지역, 장애에 대한 혐오표현 노출 정도에 통계적으로 유의미한 차이가 있는 것으로 나타남. 즉 남성이 여성에 비해 위 네 개 피해대상 집단에 대한 혐오표현에 더 많이 노출됨



N=1,022. 5점 척도

그림11. 온라인 혐오표현 노출빈도

- 연령대에 따라 온라인 혐오표현에 대한 노출량의 차이가 있는지 살펴보기 위해 일원배치분산분석을 실시한 결과, 인종·민족·국적, 종교, 성별, 성적지향에 대한 혐오표현 노출의 정도에 통계적으로 유의미한 차이가 있는 것으로 나타남. 전반적으로 20대가 다른 연령대에 비해 노출의 정도가 높은 것으로 나타났으며, 인종·민족·국적, 그리고 성별에 대한 혐오표현에 대해 30~40대가 50~60대에 비해 노출의 정도가 높음.

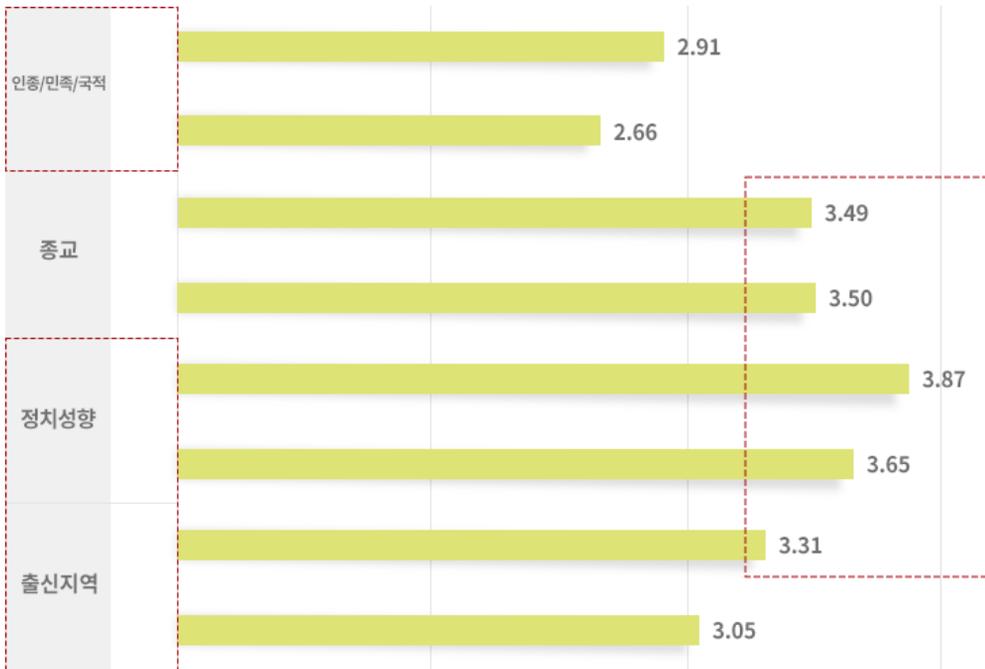


그림12-1. 피해대상 집단별 혐오표현 노출 정도의 차이 (성별) 1

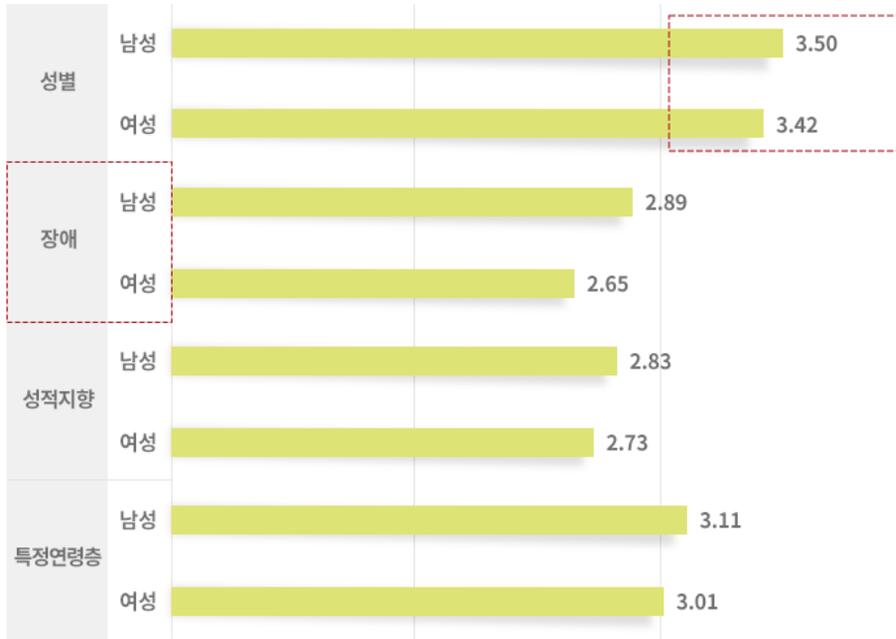


그림12-2. 피해대상 집단별 혐오표현 노출 정도의 차이 (성별) 2
 N=515(남성), 507(여성). 5점 척도

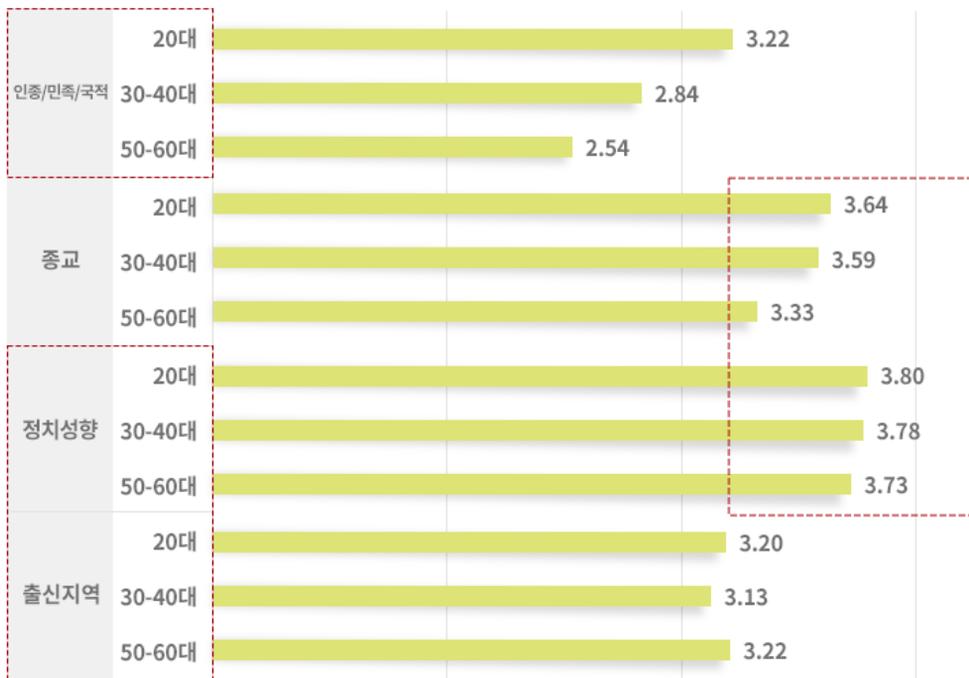
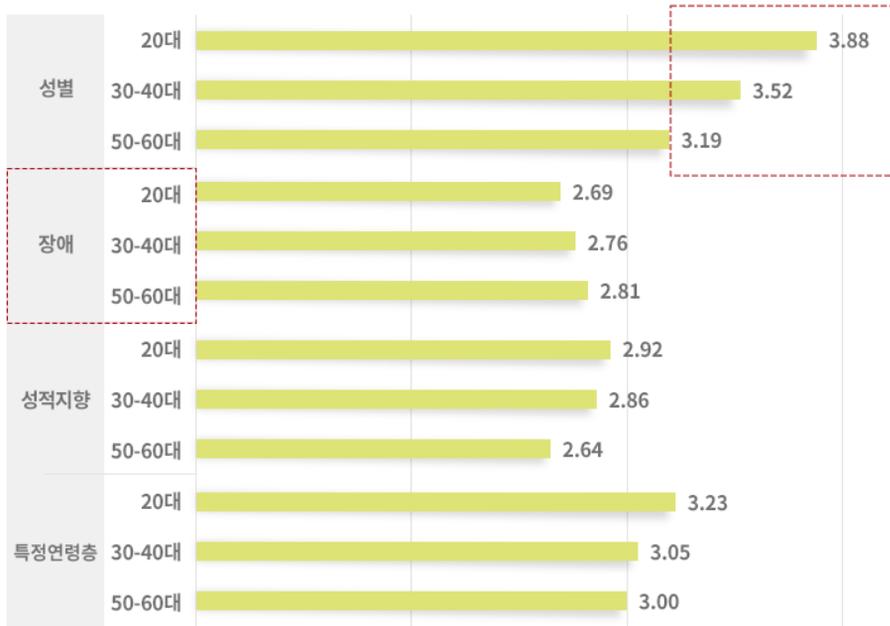


그림13-1. 피해대상 집단별 혐오표현 노출 정도의 차이 (연령대별) 1

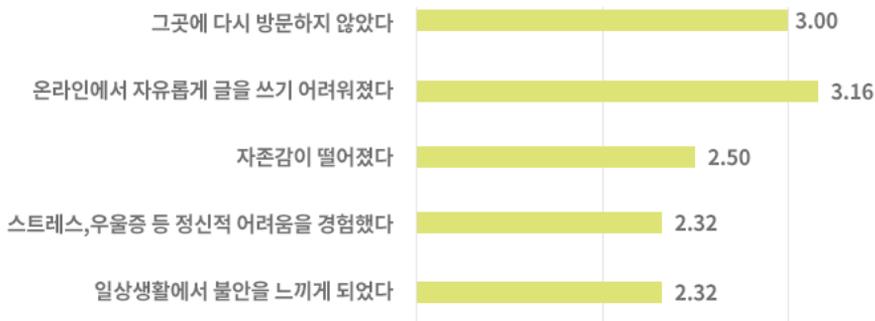


N=186(20대), 423(30-40대), 413(50-60대). 5점 척도

그림13-2. 피해대상 집단별 혐오표현 노출 정도의 차이 (연령대별) 2

□ 온라인 혐오표현 노출의 영향 (그림 14)

- 응답자들이 혐오표현에 노출되었을 때 어떠한 영향이 있는지 살펴본 결과, ‘온라인에서 자유롭게 글을 쓰기 어려워졌다’(M = 3.16, SD = 1.27), ‘그곳에 다시 방문하지 않았다’(M = 3.00, SD = 1.20)는 3점 이상으로 상대적으로 높게 나타났으나, ‘자존감이 떨어졌다’(M = 2.50, SD = 1.22), ‘스트레스, 우울증 등 정신적 어려움을 경험했다’(M = 2.32, SD = 1.22), ‘일상생활에서 불안을 느끼게 되었다’(M = 2.32, SD = 1.20)와 같은 문항은 3점 미만으로 나타났음.
- 즉 온라인 혐오표현을 접한 뒤 심리적 위축보다 행동적 위축이 상대적으로 더 큰 것으로 나타남.



N=1,022. 5점 척도

그림14. 온라인 혐오표현 노출의 영향

□ 온라인 혐오표현 생산 경험 (그림 15, 16, 17)

- 응답자 중 온라인에서 혐오표현을 올린 경험이 있는지 살펴본 결과, 응답자의 6%에 해당되는 61명만 그렇다고 응답함. 이들 가운데 남성(78.7%)이 여성(21.3%)에 비해 약 3.7배나 많은 것으로 나타났으며, 연령대별로는 20대에서 60대까지 고르게 나타남.
- 온라인에서 혐오표현을 얼마나 자주 올리는지 살펴본 결과, '1개월에 1회 이하'가 19명(31.1%), '6개월에 1회 이하'가 11명(18.0%), '1주일에 1회 이하'가 10명(16.4%) 등이 순으로 나타남.
- 온라인에서 혐오표현을 올리는 이유를 살펴본 결과, '실제로 그렇게 생각해서'가 30명(27.5%), '상대방에 먼저 내가 속한 집단을 비난하는 내용을 올려서'가 23명(21.1%), '다른 사람으로부터 인정받고 싶어서'가 16명(14.7%) 등의 순으로 나타남.

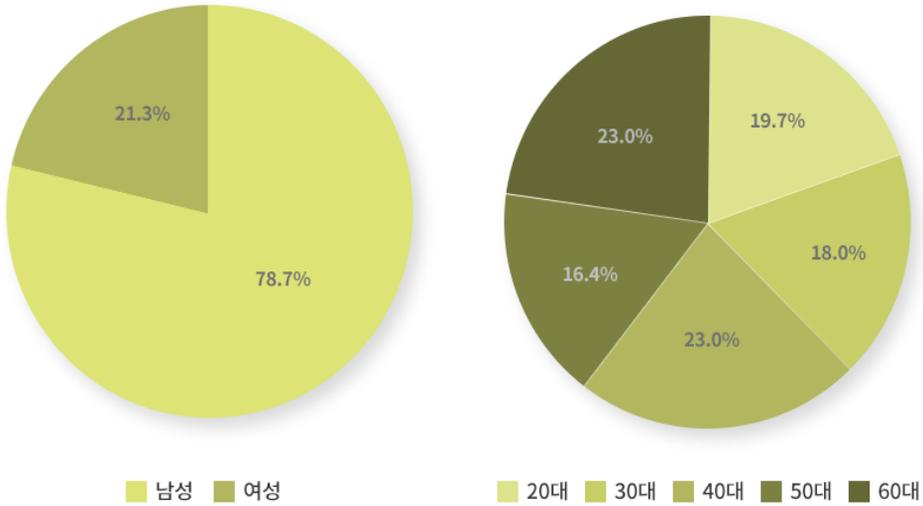


그림15. 온라인 혐오표현 생산 주체

N=61

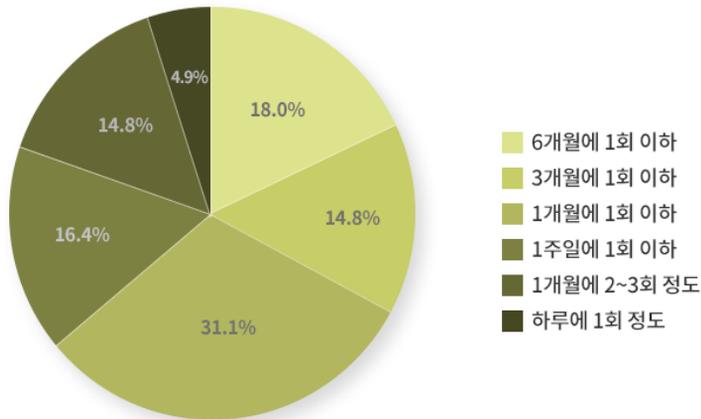
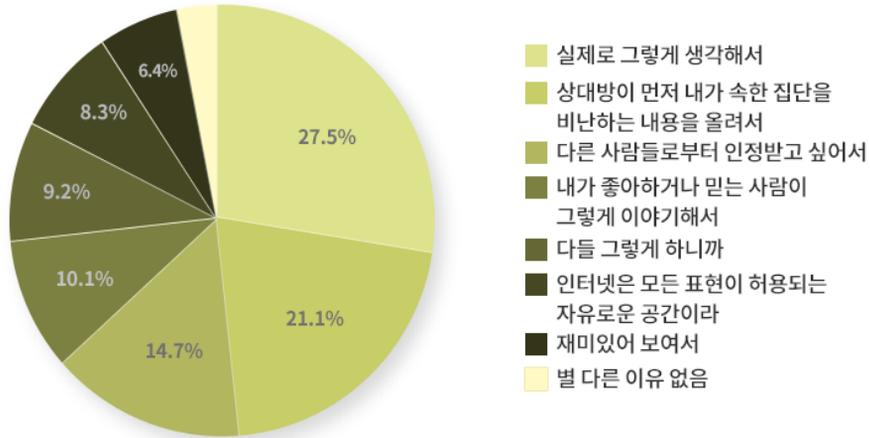


그림16. 온라인 혐오표현 생산 빈도

N=61



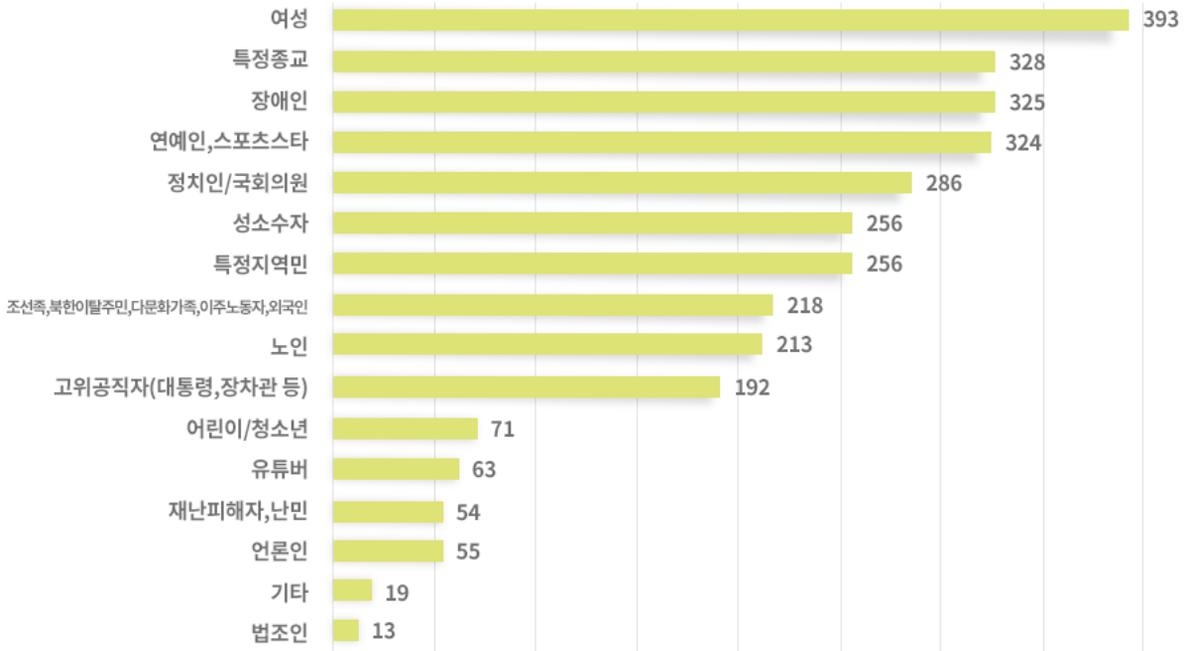
N=61

그림17. 온라인 혐오표현 생산 이유

4. 온라인 혐오표현의 피해대상 인식

□ 온라인 혐오표현의 피해대상 (그림 18, 19)

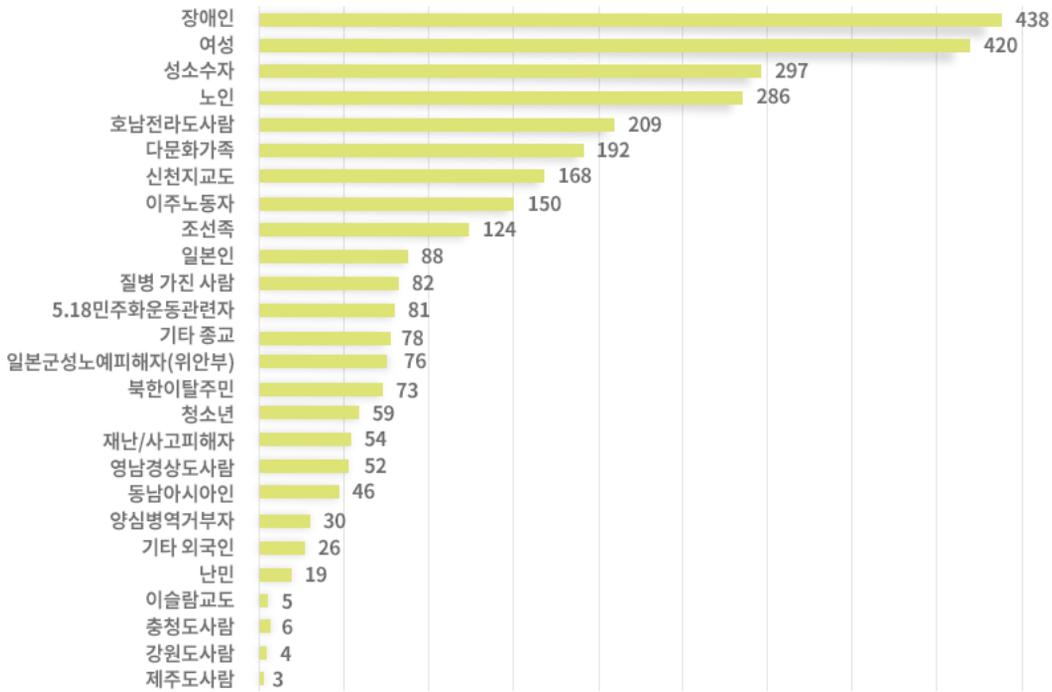
- 온라인에서 누가 혐오표현의 피해를 많이 받는다고 생각하는지 살펴본 결과, 여성 393명(12.8%), 특정종교 328명(10.7%), 장애인 325건(10.6%), 연예인·스포츠스타 324명(10.6%) 등의 순으로 나타남.



N=3,066. 1-3순위 중복응답

그림18. 온라인 혐오표현의 피해대상 (전체)

- 구체적으로 온라인에서 누가 혐오표현의 피해를 많이 받는다고 생각하는지 사회적 소수자로 한정하여 살펴본 결과, 장애인 438명(14.3%), 여성 420명(13.7%), 성소수자 297명(9.7%), 노인 286명(9.3%), 호남전라도사람 209명(6.8%) 등의 순으로 나타남.



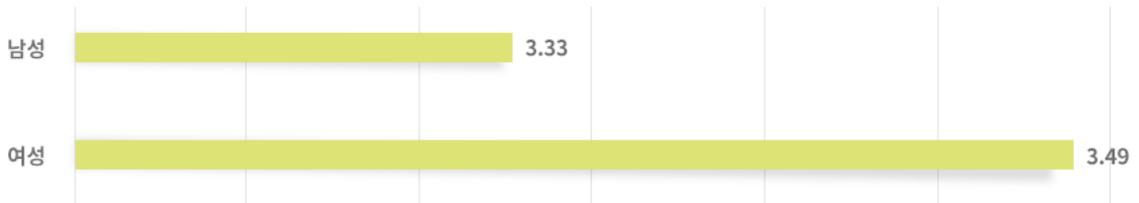
N=3,066. 1-3순위 중복응답

그림19. 온라인 혐오표현의 피해대상 (사회적 소수자)

5. 온라인 혐오표현 대응방안

□ 온라인 혐오표현 규제 태도 (그림 20, 21)

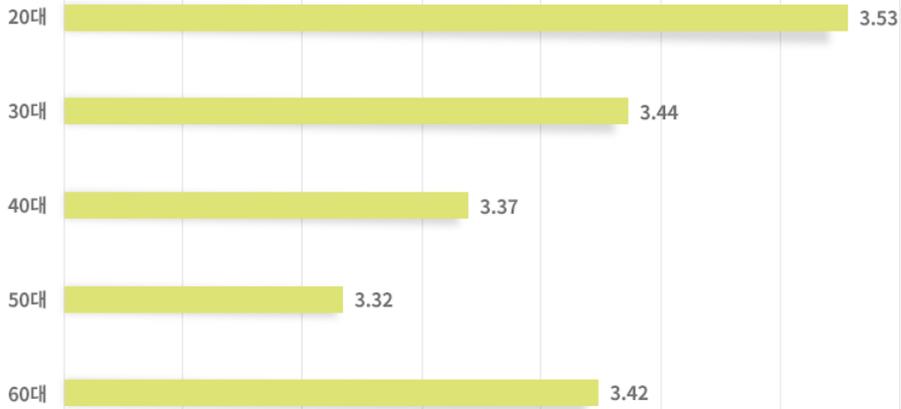
- 온라인 혐오표현에 대한 규제에 대한 태도를 살펴본 결과, 평균 3.41(SD = 0.92)로 나타나 대체로 규제의 필요성에 공감하는 것으로 나타남.
- 성별에 따라 규제 태도에 차이가 있는지 살펴본 결과, 여성이 3.49(SD = 0.94)로 남성 3.33(SD = 0.91)에 비해 높았으며, 이는 통계적으로 제한적인 수준에서 유의미한 것으로 나타남.



N=515(남성), 507(여성). 5점 척도

그림20. 온라인 혐오표현에 대한 규제 태도 차이 (성별)

- 연령대에 따라 규제 태도에 차이가 있는지 살펴본 결과, 20대 3.53(SD = 0.94), 30~40대 3.40(SD = 0.96), 50~60대 3.36(SD = 0.88) 순이었으나, 이 차이는 통계적으로 유의미하지 않은 것으로 나타남.

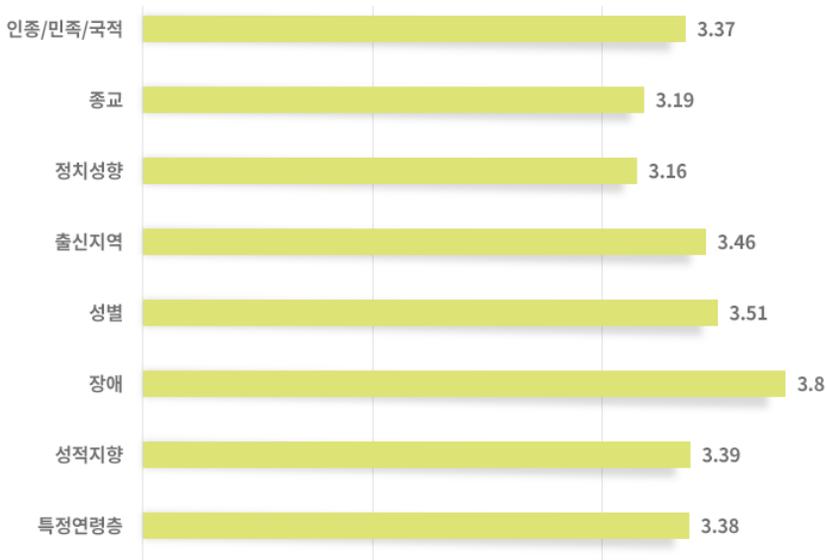


N=186(20대), 423(30-40대), 413(50-60대). 5점 척도

그림21. 온라인 혐오표현에 대한 규제 태도 차이 (연령대별)

□ 온라인 혐오표현 규제방안 인식 (그림 22)

- 온라인 혐오표현을 규제함에 있어 피해대상별 필요성에 대한 인식을 살펴본 결과, 장애(M = 3.80, SD = 1.23), 성별(M = 3.51, SD = 1.26), 출신지역(M = 3.46, SD = 1.22), 특정연령층(M = 3.38, SD = 1.15)에 대한 혐오표현 등의 순의 중요도와 우선순위로 규제해야 한다고 생각하고 있었음.
- 한편 정치성향(M = 3.16, SD = 1.21), 종교(M = 3.19, SD = 1.22)에 대한 혐오표현에 대해서는 필요성이 다른 피해대상에 대한 혐오표현에 비해 상대적으로 낮은 것으로 나타남.



N=1,022. 5점 척도

그림22. 온라인 혐오표현 규제방안 인식 (피해대상별 필요성)

□ 온라인 혐오표현 대응방안 (그림 23)

- 온라인 혐오표현 대응방안에 대한 인식을 살펴본 결과, ‘포털 등 사업자의 임시조치, 삭제 등 적극적 활용’(M = 4.02, SD = 1.05)이 가장 높은 것으로 나타났으며, 이어 ‘피해자의 반론, 반박 기회 등 보장’(M = 3.90, SD = 1.04), ‘포털 등 사업자의 자율적 규제와 투명성 보고서 발행’(M = 3.88, SD = 1.03), ‘시민교육과 미디어 리터러시 교육 제공’(M = 3.85, SD = 1.01) 등의 순으로 나타남.
- 한편 ‘형사적으로 처벌하는 법률(예: 혐오표현처벌법, 차별금지법 등) 제정’(M = 3.48, SD = 1.17)이나 ‘현행 형사 관련 법률의 적극적 활용’(M = 3.64, SD = 1.11)과 같은 형사처벌과 관련된 대응방안은 상대적으로 다소 낮은 것으로 나타남.

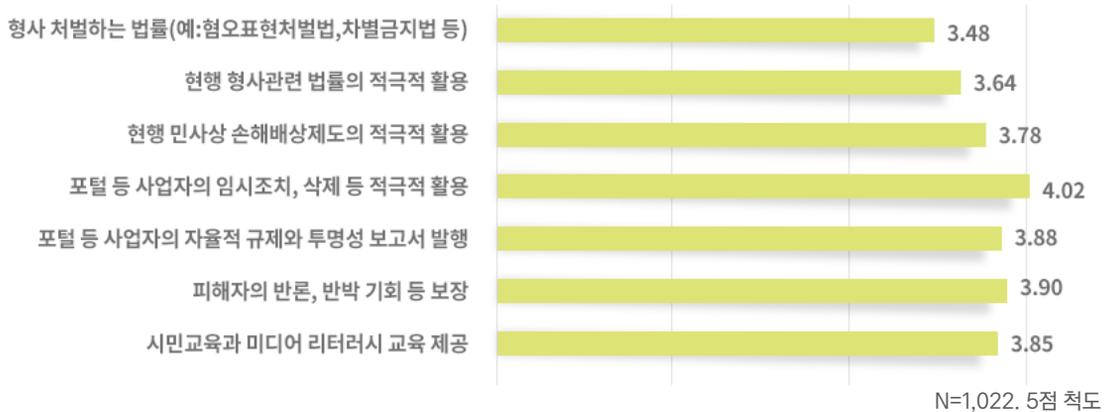


그림23. 온라인 혐오표현 대응방안에 대한 인식

IV. 온라인 혐오표현에 대한 전문가 인식 조사

1. 조사대상 및 내용

□ 조사대상

- ‘온라인 혐오표현’에 대한 일반시민들의 인식조사와 더불어 전문가들이 생각하는 ‘온라인 혐오표현’의 개념, 특성 및 확산 배경 등에 대한 심층적인 비교, 분석이 필요하다고 판단함. 시민들의 양적 조사결과 내용의 함의를 살펴볼 수 있다는 점, 온라인 혐오표현 현상에 대한 정책적 대안을 제시받을 수 있다는 점 등을 고려함.
- 이를 위해 오랫동안 혐오표현과 온라인 혐오표현 현상을 탐구해 온 언론학자, 헌법학자, 법률가 15명을 대상으로 전문가 심층 인터뷰를 실시함. 언론학자 7명, 법학자 6명 그리고 변호사 2명임. 심층 인터뷰는 2020년 9월 1일부터 17일까지 이메일을 통해 의견을 제시받았고 필요할 경우 직접 전문가와 전화 통화하여 답변 내용을 보충하였음.

□ 조사내용

- 서술형 질문과 선택형 질문으로 면접지를 구성하였음.
- 서술형 질문은 첫째, 온라인 혐오표현의 특성과 정의 그리고 일반적인 의미에서 혐오표현과 차이점 둘째, 온라인 혐오표현이 지속되는 이유와 대응 방법에 대해 전문가의 의견을 구했으며 개인적 차원, 교육·문화·언론 등 사회적 차원, 인터넷서비스 사업관리 차원, 정책적·법적 차원으로 구분하여 질문하였음.
- 선택형 질문으로 첫째, 온라인의 다양한 공간 중에서 혐오표현이 가장 심각하다고 판단하는 플랫폼 평가, 둘째, 혐오표현으로 간주되는 사례 중에서 가장 심각하다고 평가하는 표현, 셋째, 온라인에서 혐오표현으로 인해 가장 심각한 피해를 입은 대상에 대한 평가, 넷째, 사회적 소수자 중에서 온라인 혐오표현으로 인해 가장 심각한 피해를 입은 대상에 대한 평가 다섯째, 온라인 혐오표현에 대한 대응방식 중에서 가장 필요하다고 평가하는 방안 등임.

2. 온라인 혐오표현의 정의와 특성 그리고 일반적 혐오표현과 차이

□ 온라인 혐오표현의 정의와 특성

- “온라인상에서, 사회적 소수자에 대해 성, 연령, 종교, 나이, 인종, 장애, 성적 지향 등을 이유로 차별, 편견에 근거해 일방적으로 비하, 비방하거나 부정적으로 묘사하는 표현”이라는 정의가 많음. 특정 상대에 대한 적대, 적의적인 표현을 하는 행위로 규정하면서, 다른 행위와 비교할 때 고의성 또는 인식이 전제되는 행위라고 평가함.
- ‘온라인 혐오표현’의 특성과 정의에 대한 전문가들의 답변은 다음과 같은 특징을 가지고 있음. 첫째, 사회적 소수자를 일방적으로 비난하거나 비하, 부정적으로 묘사하는 표현이라고 파악하고 있음. 둘째, 특정 상대에 대한 차별적, 비하적, 모욕적 표현으로 보는 견해가 많으나 혐오표현과 모욕을 구분할 필요성, 사실적 근거를 가지고 정당하게 비판, 논평하는 것과 구분해야 한다는 의견을 제시하고 있음. 셋째, 온라인 혐오표현은 익명성을 바탕으로 하기 때문에 일반인들의 혐오표현 동참이 용이하고 그 파급력 또한 강하다고 평가하고 있음. 넷째, 사회적·정책적 차원에서 중요한 것은 국가·시민사회가 대응해야 할 대상으로서 혐오표현의 설정하고 이에 대한 방안을 강구할 필요성을 제기함.

□ 온라인 혐오표현과 일반적 혐오표현의 차이

- 전문가들은 온라인 혐오표현과 일반 혐오표현의 차이는 거의 없다고 평가하고 있음. 서로 구분되지 않는다는 견해, 강학상 다르게 볼 이유가 없다는 견해, 다만 정책 집행을 위해 양자를 구분해야 한다는 견해가 제시됨. 이를테면 “온라인 혐오표현의 경우 온라인 공간의 익명성과 발화의 용이성으로 인하여 조직적으로 이루어지는 혐오표현뿐만 아니라 편견에 기반 한 고정관념을 확산시키는 차별적 표현이 광범위하게 나타날 수 있고 혐오표현 일반과 온라인 혐오표현의 구분의 필요성은 국가와 시민사회의 대응 방식을 논하기 위하여 필요”하다

는 것임.

- 다만 혐오표현이 ‘어떤 공간’에서 발화되느냐에 따라 그 파급력, 영향력이 크며 온라인상 혐오표현은 익명성, 개방성, 전파성, 쌍방향성 등의 이유로 물질적, 정신적 피해가 훨씬 심각하다고 진단함.
- 차별적 표현과 혐오표현은 일부 중첩되거나 유사할 수 있으나 동일하지는 않다는 견해, 혐오표현을 반드시 사회적 소수자에 대한 다수자의 적의적 표현으로만 간주해서는 안 되고 다수자에 대한 혐오표현 역시 존재한다는 견해도 제시됨. 소수자에 대한 혐오와 다수자에 대한 혐오표현은 사회적 의미와 해악의 정도, 법적 통제와 규제범위를 설정할 때 정책적 고려를 달리해야 한다는 정도의 차이가 존재한다는 것임.
- 또 온라인 혐오표현과 일반 혐오표현의 차이를 단지 전송수단의 차이로 파악하는 전문가들 중에 모욕, 명예훼손, 반인종주의, 특정 종파에 대한 비난, 전쟁에 대한 찬양·고무·선동, 명백하게 허위인 정보의 전파 등도 온라인 혐오표현의 범주에 포함시켜야 한다는 견해를 제시한 경우가 있음.

3. 온라인 혐오표현의 지속 이유와 대응 방법

□ 온라인 혐오표현의 지속 이유

- ‘개인적 차원’에서 온라인 혐오표현이 지속되는 이유를 전문가들은 다음과 같이 제시하고 있음.
 - 첫째, 혐오표현의 심각성을 인식하지 못하고 인터넷상의 표현에 익숙해 짐으로써 놀이나 유희 차원에서 관행적으로 저항감 없이 사용하고 있다는 점.
 - 둘째, 다른사람에 대한 포용, 배려와 존중심이 부족하고 공감역량이 약화된 채 자신의 분노, 좌절, 스트레스를 배설하는 차원에서 혐오표현을 활용하고 있다는 점.
 - 셋째, 고정관념과 편견, 잘못된 가치 평가, 자신의 사고에 대해 성찰하지 못한다는 점.
 - 넷째, 각 개인의 사회경제적 지위 약화와 불안의 가속을 사회적 소수자들에 전가하는 경향이 심화되는 점.
 - 다섯째, 익명성을 활용해 온라인 공간에서 자신의 불만을 공개적으로 표출해 정서적 카타르시스를 느끼고 이에 대한 동조자들과 연대를 통해 심리적 고립감을 벗어나려고 하는 점.
 - 여섯째, 혐오표현 현상을 비판적으로 수용할 수 있는 미디어리터러시 능력이 부족한 점.
- ‘사회적 차원’에서 볼 때 온라인 혐오표현이 지속되는 이유는 다음과 같음.
 - 첫째, 초·중·고의 미디어 교육 및 차별과 혐오에 대한 가정·사회적 교육이 미흡하다는 점,
 - 둘째, 저성장 및 경제적 여건의 악화로 경제적 불평등이 심화된 반면 오프라인의 도덕적 수위가 강화됨으로써 오프라인에서 불만을 표출할 수 있는 배출구가 약화돼 이를 온라인 표현공간에서 표출하는 데 따른 결과임.
 - 셋째, 사회적인 차별과 갈등이 심화되고 있는 상황에서 언론이 특정 집단에 대한 고정관념을 지속적으로 재생산하여 이러한 차별과 갈등을 증폭시키는 데 기여하고 있다는 점.
 - 넷째, 온라인 커뮤니티 공간에 대해 너무 과도한 의미를 부여해 과민 대응해 온 언론과 시민사회 등의 분위기

기도 온라인 혐오표현의 심화, 확산의 한 요소라고 진단함.

- 다섯째, 편견과 차별이 역사적으로 지속, 심화되어 일상생활에 깊숙이 침투되고 정치적·사회적·경제적·문화적 주류문화로 자리 잡아 재생산되어 왔고 그러한 편견과 차별이 무의식적으로 온라인 공간에서 혐오표현으로 발화되고 있다는 점. 그 외 언론사가 대립되는 이슈를 제공하고 확대 재생산하면서 공론화 된 주제에 참여하지 못할 경우 소외의식을 만들고 있다는 점, 우리 사회가 유럽보다 미국의 이분법적인 대립 구도의 사회문화적 영향을 더 많이 받아 자신과 다른 견해를 온라인에서 투쟁의 대상화 함으로써 극단적인 공격들이 나타나고 있다는 점, 애초 표현의 자유의 보호 범위 밖에 있는 혐오표현을 표현의 자유의 한 영역으로 잘못 인식하고 있으며 혐오표현에 대한 규제는 일반적인 표현의 자유와 충돌이 아님에도 불구하고 이를 표현의 자유 문제로 파악하고 있다는 점도 전문가들은 제시하였음.
- ‘인터넷서비스 사업관리 차원’에서 발생한 온라인 혐오표현의 문제에 대해 전문가들은 다음과 같은 의견을 제시했음.
 - 첫째, 인터넷서비스 사업자들이 온라인 공간의 혐오표현의 문제점을 인지하고 있음에도 표현의 자유를 이유로 이에 적극적으로 개입하지 않고 방치하고 있다는 점.
 - 둘째, 나아가 이러한 표현행위들이 이루어지는 것 자체를 인터넷서비스의 일환으로 간주하고 있는데 이는 인터넷서비스 사업자들의 상업적 관점을 반영하고 있으며 언론 역시 이에 일조하고 있다는 점.
 - 셋째, 혐오표현의 심각성에 대한 인터넷서비스 사업자들의 인식 자체가 낮다는 점. 그 외 ‘원인’과 ‘책임’은 구분되어야 하고 인터넷서비스 사업자들이 온라인공간의 혐오표현에 대해 법적 책임을 질 수는 있으나 온라인 혐오표현의 지속, 확산의 원인이 되었다고 평가할 수 없다는 견해도 제시되었음.
- ‘정책적·법적 차원’에서 볼 때 온라인 혐오표현이 지속되는 이유는 다음 몇 가지임.
 - 첫째, 혐오와 차별에 대한 처벌 법률이 부재하고 정책적으로도 온라인상 혐오표현이 처벌될 수 있다는 신호를 온라인 이용자들에게 인식시키지 못하고 있다는 점.
 - 둘째, 온라인상 정치적 표현에 대해 예민하게 대응하면서 표현의 자유로부터 보호받아야 할 기본권에 대한 문제를 너무 안일하게 대응하고 있다는 점.
 - 셋째, 온라인 혐오표현의 처벌 쟁점에 대한 사회적 합의 노력이 필요함에도 불구하고 합의를 유도해 내거나 경계상의 문제에 대한 가이드라인을 마련하지 못하고 있다는 점.
 - 넷째, 포털에서 제공되는 내용이지만 사실상 포털사업자의 온라인 혐오표현에 대한 책임을 면제하고 있는 법·정책상의 문제임.
 - 다섯째, 온라인 혐오표현은 한 국가 범주에 머무르지 않는 특성을 갖고 있고 따라서 이러한 양상을 이해하고 온라인의 특성을 활용한 대항언론 활성화가 시급하기 때문에 국가 정책적 대응이 긴요함에도 정부의 혐오표현 개념에 대한 이해가 부족한 점.
 - 여섯째, 온라인 혐오표현의 특성상 장기적 차원에서 문제해결을 위한 정책적 대응을 해야 함에도 단기적 관점에서 이를 정책적으로 해결하려는 접근을 하고 있다는 점.

□ 온라인 혐오표현에 대한 대응 방안

- ‘개인적 차원’의 대응방법으로 전문가들이 제안한 것은 다음과 같음.
 - 첫째, 오프라인의 정체성과 온라인상의 정체성이 일치되도록 스스로 노력이 필요하며 온라인 혐오표현의 해악에 대한 경계심을 가질 수 있도록 자기 성찰이 배가되어야 함.
 - 둘째, 개인적으로 사회적 소통능력을 함양해야 하는 데, 기본적으로 인간을 존중하고 타인을 배려, 포용하려는 공감능력을 높여야 함. 이를 위해서는 인간이 가진 정체성의 다양성에 대한 이해를 증진하고 자기 안의 편견의 발견과 인식, 혐오표현의 표적이 되는 집단 구성원에 대한 일상적 만남과 경험을 강화해 스스로 변화를 꾀할 필요가 있음.
 - 셋째, 법적 지식을 축적하고 미디어 리터러시 교육에 참여해 그 역량을 제고해야 함. 전문가들은 온라인 혐오표현에 대한 대응 차원에서 볼 때 ‘개인적 차원’의 대응이 가장 어려울 것이라고 평가하고 있기도 함.
- ‘사회적 차원’의 대응방안으로 전문가들은 다음과 같은 점을 제시함.
 - 첫째, 초·중·고에서 미디어 교육을 ‘제도화’하는 것이 중요함. 이를 위해 교육부가 국가교육과정을 변경해야 하며 정규 교과목 설정이 어려울 경우 최소한 특정 교과의 한 단원으로 편제해서 교육해야 함. 초·중·고에서 미디어 교육이 실질적으로 실행될 수 있도록 현직 교사에 대한 재교육, 예비 교사들에 대한 미디어교과목 필수화를 구축해야 함.
 - 둘째, 문화적으로 사회적 약자에 대한 공감능력을 높이고 인권과 민주주의의 가치에 대한 시민들의 각성을 제고하는 미디어 교육과 언론의 정보제공이 동반되어야 함.
 - 셋째, 온라인 혐오표현을 소개하고 주목하는 언론의 보도 태도를 변화시켜야 함. 언론을 통한 혐오표현의 생산, 재생산을 억제할 수 있도록 언론재단과 사업자 연합체 등에서 언론인에 대한 저널리즘 교육을 실시해야 함.
 - 넷째, 교육·언론·시민단체 등에 의한 다차원적인 통합적 공동 노력이 필요한데 디지털 리터러시 교육, 혐오표현 보도를 자제하고 소통위주의 콘텐츠를 제공하려는 언론의 노력, 혐오표현을 모니터링하고 캠페인 활동을 전개하는 시민단체의 노력이 동시 다발적으로 이루어져야 함. 전 연령층을 대상으로 민주 시민성 함양을 위한 평생교육이 실시되어야 함.
- ‘인터넷서비스 사업자’ 차원에서 대응해야 할 방안에 대해 전문가들이 제시한 안은 다음과 같음.
 - 첫째, 이용약관에 혐오표현을 할 경우 표현물 삭제, 이용자 계정 사용중지 등의 불이익이 있다는 점을 명확하게 밝힐 것.
 - 둘째, 게시물이나 댓글 등을 게시하거나 등록하기 직전에 표현물에 혐오표현이 있는지 여부를 확인할 수 있도록 혐오표현에 특정 표시를 해 주는 방법을 실행할 것.
 - 셋째, 지속적으로 캠페인 활동을 전개하는 방안으로 포털의 첫 화면 윗줄에 온라인 혐오표현을 경계하고 해악을 알리는 문구를 스트리밍하는 방안을 실행할 수 있을 것임.
 - 넷째, 익명성을 보장하면서도 온라인 메시지 게시자의 사이트 내에서의 과거 온라인 경력이 드러날 수 있

도록 사이트를 재설계하는 방식을 활용해 볼 것.

- 다섯째, 표현물 삭제나 이용자 계정 중지 등의 강력한 조치 외에 검색 결과에서 혐오표현에 해당하는 게시물을 배제시키는 방식을 실행할 것. 여섯째, 포털 등 인터넷기업 종사자를 대상으로 혐오표현 관련 교육을 실시하고 온라인 혐오표현의 유통을 억제하는 것이 사회적으로 가진 의의와 가치를 이해하도록 해야 함.
- ‘법적·정책적 차원’의 대응방식에 대해 전문가들은 다음과 같은 방안을 제안함.
 - 첫째, 온라인 혐오표현은 금지된다는 신호를 이용자들에게 전달하고 차별금지법, 증오범죄규제법, 혐오처벌법 등의 법률을 제정해 대응할 것을 제안함. 그 중심에 차별금지법을 둘 것을 제안하고 있음.
 - 둘째, 정책적으로 장기적 차원의 위원회를 만들어 실천과제를 도출하고 실행할 것.
 - 셋째, 미디어교육지원법, 인권교육법 등을 통한 시민들의 민주적 시민성 함양과 디지털 정보 수용 및 이용 역량을 제고할 것.
 - 넷째, 현행 ‘정보통신망법’의 불법정보조항에 혐오표현 신설, 정보통신심의규정 제8조제3호 바목의 재해석 검토, 청소년보호법상 유해매체물에 혐오표현을 지정을 신설하는 방안을 검토해 볼 것.

□ 선택형 응답에 나타난 온라인 혐오표현 대응 방안

- 전문가들이 선택형 응답을 통해 제안한 가장 중요한 온라인 혐오표현 대응방안은 ‘시민교육, 미디어리터러시 교육’임. 15명의 전문가 중 10명이 제1순위로 ‘교육적 대응’을 선택하였음. 두 번째 대응방법으로 선택한 것은 ‘인터넷서비스 사업자의 자율규제와 투명성보고서’ 방식임. 제2순위와 제3순위로 다수 전문가들이 제안한 것은 혐오표현에 대한 ‘임시조치와 삭제를 적극적으로 활용할 것’임. 그리고 ‘혐오처벌법 등 형사법’을 제정해서 대응할 것을 미디어 교육, 자율규제와 보고서 등의 조치에 이어 후순위이긴 하지만 보완책으로 제안하고 있음.
- 기타 선택형 답변에 나타난 특성은 첫째, 전문가들은 온라인 공간의 ‘일베’, ‘디씨 인사이드’, 블로그와 유튜브 등에서 온라인 혐오표현이 가장 심각하다고 평가하고 있음. 둘째, 온라인 혐오표현물을 게시하는 이유로 ‘실제 그렇게 생각’하거나 ‘다른 사람의 인정을 받고 싶어서’ 혹은 ‘다른 사람들도 그렇게 하니깐’, ‘상대방이 먼저 비난을 했으므로 그에 대한 대응방법으로 게시’한다는 평가를 하고 있음. 셋째, 사회적 소수자 중에서는 ‘성적 소수자’와 ‘여성’이 온라인 혐오표현의 피해를 가장 많이 받고 있는 것으로 평가하고 있음.

V. 나가며

□ 온라인 혐오표현에 대한 일반시민의 인식 조사 결과, 표현방식, 성별, 연령대, 교육수준에 따라 차이가 있었음

- 선동형 표현보다 모욕형 표현에 대한 심각성을 더 크게 인식하는 것으로 나타남.
- 여성이 남성보다, 젊을수록, 그리고 교육수준이 높을수록 각 혐오표현에 대한 인식이 더 높은 것으로 나타남.
- 혐오표현의 대상과 자신과의 심리적 거리감이 크다고 느끼는 경우 혐오표현에 대한 인식이 상대적으로 낮음

것으로 나타남.

□ 종교, 정치성향, 출신지역, 성별, 특정연령층 대상의 온라인 혐오표현에 대한 노출이 상대적으로 많았음.

- 남성이 여성보다, 젊을수록 온라인 혐오표현을 더 많이 접하는 것으로 나타남.
- 온라인 혐오표현 생산은 남성이 여성에 비해 약 3.7배 많았으며, 연령대별로는 고른 분포를 보임.

□ 온라인 혐오표현에 대한 일반시민의 인식에 기반한 인터넷 사업자의 자율규제 방안 마련이 필요함.

- 혐오표현의 대상이 누구이며, 표현방식이 어떠한지에 따라 일반시민들이 느끼는 심각성이 다르다는 점을 감안하여 표현의 자유를 최대한 보장하면서도 피해대상과 표현방식에 따라 즉시 삭제할지, 임시조치를 할지 등에 대한 규제의 수위를 정할 필요가 있음.
- 혐오표현 사전에 구축하고 지속적으로 업데이트하고, 유사도 기반 혐오표현 탐지 알고리즘과 머신러닝을 활용하여 혐오표현에 대해 이용자가 인지할 수 있도록 알리거나 즉각 조치하여 노출을 최소화시킬 필요가 있음.

□ 교육대상별 온라인 혐오표현 관련 미디어 리터러시 교육 프로그램 개발 및 실시가 요구됨.

- 혐오표현에 대한 인식의 정도, 혐오표현에 대한 생산과 노출 정도가 성별, 연령대별, 교육수준별로 차이가 있기 때문에 교육대상을 성별, 연령대별, 교육수준별로 세분화한 미디어 리터러시 교육 프로그램을 개발하고 실시할 필요가 있음.
- 정책기관, 미디어 사업자, 시민단체, 학계가 온라인 혐오표현 관련 미디어 리터러시 교육을 위한 협의체를 구성하여 유기적인 협력 속에서 지속적인 교육을 실시하도록 연결할 필요가 있음.

□ 온라인 혐오표현을 규제하기 위해 법적 강제력을 동원하는 방안은 신중하게 접근할 필요가 있음.

- 일반시민들은 온라인 혐오표현에 대한 규제 필요성에 공감하면서도, 민·형사상 처벌하는 규제방식에 대해서는 다른 방식에 비해 상대적으로 낮게 평가한 점을 감안해 법적 강제력을 행사하여 규제하는 방안에 대해서는 신중하게 접근해야 할 것임.
- 혐오표현과 현행법으로 규제가 가능한 모욕적 표현을 구분 짓는 것은 상당히 어려움. 타인의 사회적 평가를 저하시키는 욕설은 국내 법제에서 모욕죄로 처벌가능하기 때문에 혐오표현 규제를 위한 논의에서는 모욕적 표현을 제외할 필요가 있음. 혐오표현은 특정 집단을 목표로 삼는다는 특징이 있기 때문임.
- 혐오표현의 개념과 범주가 불명확하므로 혐오표현 규제를 위한 입법은 자칫 표현의 자유를 심각하게 제한할 위험이 있음. 따라서 혐오표현 규제는 신중하게 행해져야 하며, 혐오표현을 규제하는 입법은 명확하고 구체적으로 마련되어야 함.

참고문헌

〈논문 및 보고서〉

- 국가인권위원회 (2016). 〈혐오표현 실태조사 및 규제방안 연구〉. 서울: 국가인권위원회.
- 국가인권위원회 (2017). 〈혐오표현 실태조사 및 규제방안 연구〉. 서울: 국가인권위원회
- 국가인권위원회 (2019). 〈혐오표현 리포트〉. 서울: 국가인권위원회.
- 국가인권위원회 (2020). 〈차별에 대한 국민인식조사〉. 서울: 국가인권위원회.
- 김민정 (2020). 소셜미디어 플랫폼상의 혐오표현 규제. 〈방송문화연구〉, 32권 1호, 7-54.
- 문화체육관광부 (2018). 〈혐오 표현 대응관련 대국민 인식조사 결과보고서〉. 세종: 문화체육관광부.
- 박아란·양정애 (2016). 혐오표현과 여성혐오에 대한 인식. 〈미디어 이슈〉, 2권 7호, 1-14.
- 윤성욱 (2019). 혐오표현 규제와 법적 쟁점에 관한 연구. 〈미디어와 인격권〉, 5권 1호, 57-94.
- 이주영 (2015). 혐오표현에 대한 국제인권법적 고찰. 〈국제법학회논총〉, 60권 3호, 195-227.
- Baker C. E. (2009). Autonomy and hate speech. In: Hare I, Weinstein J (Eds.), *Extreme Speech and Democracy*. Oxford: Oxford University Press.
- Baker, C. E. (2012). Hate speech. In M. Herz & P. Molnar (Eds.), *The content and context of hate speech: Rethinking regulation and responses* (pp. 57-80). Cambridge: Cambridge University Press.
- Balica, R. (2017). The criminalization of online hate speech: It's complicated. *Contemporary Readings in Law and Social Justice* 9(2), 184-190.
- Brown, A. (2018). What is so special about online(as compared to offline) hate speech? *Ethnicities*, 18(3), 297-326.
- Chetty, N., & Althur, S. (2018). Hate speech review in the context of online social networks. *Aggression and Violent Behavior*, 40, 108-118.
- Citron DK. (2014). *Hate Crimes in Cyberspace*. MA: Harvard University Press.
- CNN (2019, March 16). How the Christchurch Terrorist Attack was Made for Social Media. Retrieved from <https://edition.cnn.com/2019/03/15/tech/christchurch-internet-radicalization-intl/index.html> (최종접속일: 2020.12.1.)
- Cohen-Almagor R. (2015). *Confronting the Internet's Dark Side: Moral and Social Responsibility on the Free Highway*. Cambridge: Cambridge University Press.
- Cowan, G., & Hodge, C. (1996). Judgments of hate speech: The effects of target group, publicness, and behavioral responses of the target. *Journal of Applied Social Psychology*, 26(4), 355-374.

- Dworkin, R. (2009). Foreword. In I. Hare & J. Weinstein (Eds.), *Extreme speech and democracy*(pp. v-ix). Oxford: Oxford University Press.
- Dworkin R. (2012). Reply to Jeremy Waldron. In: Herz M, Molnar P (Eds.), *The Content and Context of Hate Speech: Rethinking Regulation and Responses*. Cambridge: Cambridge University Press.
- Facebook’s Civil Rights Audit – Final Report (2020.08). Retrieved from <https://about.fb.com/wp-content/uploads/2020/07/Civil-Rights-Audit-Final-Report.pdf> (최종접속일: 2020.12.1.)
- Gelber, K., & McNamara, L. (2016). Evidencing the harms of hate speech. *Social Identities*, 22(3), 324-341.
- Kaakinen, M., Oksanen, A., & Räsänen, P. (2018). Did the risk of exposure to online hate increase after the November 2015 Paris attacks?: A group relations approach. *Computers in Human Behavior*, 78, 90-97.
- Laub, Zachary (2019, June 7). Hate Speech on Social Media: Global Comparisons. Council on Foreign Relations, Retrieved from <https://www.cfr.org/background/hate-speech-social-media-global-comparisons> (최종접속일: 2020.12.1.)
- Müller, Karsten., & Carlo Schwarz. (2020). Fanning the Flames of Hate: Social Media and Hate Crime. *Journal of the European Economic Association*, jvaa045, Retrieved from <https://doi.org/10.1093/jeaa/jvaa045>.
- OECD (2019). Platform Companies & Responsible Business Conduct. Retrieved from <http://mneguidelines.oecd.org/RBC-and-platform-companies.pdf> (최종접속일: 2020.12.1.)
- Soral, W., Bilewicz, M., & Winiewski, M. (2018). Exposure to hate speech increases prejudice through desensitization. *Aggressive behavior*, 44(2), 136-146.
- The Economist (2017, November 4). Do social media threaten democracy?. Retrieved from <https://www.economist.com/weeklyedition/2017-11-04> (최종접속일: 2020.12.1.)
- The New York Times (2018, October 28). On Gab, an Extremist-Friendly Site, Pittsburgh Shooting Suspect Aired His Hatred in Full. Retrieved from <https://www.nytimes.com/2018/10/28/us/gab-robert-bowers-pittsburgh-synagogue-shootings.html> (최종접속일: 2020.12.1.)
- The New York Times (2018, November 6). Facebook Admits It Was Used to Incite Violence in Myanmar. Retrieved from <https://www.nytimes.com/2018/11/06/technology/myanmar-facebook.html> (최종접속일: 2020.12.1.)
- The Washington Post (2016, August 23). Prosecutors say Dylann Roof ‘self-radicalized’ online, wrote another manifesto in jail. Retrieved from <https://www.washingtonpost.com/news/post-nation/wp/2016/08/22/prosecutors-say-accused-charleston-church-gunman-self-radicalized-online/> (최종접속일: 2020.12.1.)
- UN (2018.09.12). Report of the Independent International Fact-Finding Mission on Myanmar. UN Doc. A/HRC/39/64.
- Waldron, J. (2012). *The Harm in Hate Speech*, Harvard University Press.

Weede, E. (2016). Geopolitics, Institutions, and Economics: On the Rise and Decline of Civilization
s. Geopolitics, History, and International relations 8(1), 177–220.

〈국제인권문서 및 외국법령〉

모든 형태의 인종차별철폐에 관한 국제협약 (1965). 제4조(a)항.

세계인권선언 (1948). 제7조, 제19조.

시민적 및 정치적 권리에 관한 국제규약 (1966). 제19조, 제20조 제2항, 제26조.

Act to Improve Enforcement of the Law in Social Networks of Germany, (Network Enforcement Act, NetzDG)

Communications Decency Act of the United States, 47 U.S.C. §230.

Council of Europe (2003, January 28). Additional Protocol to the Convention on Cybercrime, concerning the criminalisation of acts of a racist and xenophobic nature committed through computer systems (Treaty No.189), Strasbourg.

Criminal Code of Germany (Strafgesetzbuch – StGB), Sections 130, 185, 187.

EU (2016). Code of conduct on countering illegal hate speech online.

EU (2008, November 28). Council Framework Decision on combating certain forms and expressions of racism and xenophobia by means of criminal law.

Executive Order on Preventing Online Censorship (US), (2020, May 28). Retrieved from <https://www.whitehouse.gov/presidential-actions/executive-order-preventing-online-censorship/> (최종접속일: 2020.12.1.)

General Act on Equal Treatment of Germany, Section 3.

Ofcom Broadcasting Code of United Kingdom (with the Cross-promotion Code and the On Demand Programme Service Rules), (2019, January 1). Section three: Crime, disorder, hatred and abuse, Retrieved from <https://www.ofcom.org.uk/tv-radio-and-on-demand/broadcast-codes/broadcast-code> (최종접속일: 2020.12.1.)

카카오 증오발언 대응 정책 문서

발행일

2021년 1월 13일

발행처

(주) 카카오

기획 / 편집/ 문의처

카카오 정책팀

policy.kakao@kakaocorp.com

디자인

케이앤웍스 디자인운영파트

〈카카오 증오발언 대응 정책 문서〉의 판권은 주식회사 카카오가 소유하고 있으며, 카카오는 〈카카오 증오발언 대응 정책 문서〉를 상업적 목적으로 이용하는 것을 금지합니다. 또한 저작권법에 따른 정당한 범위 내에서 〈카카오 증오발언 대응 정책 문서〉를 이용·가공·인용하는 경우라도 반드시 출처를 밝혀주시기 바랍니다.

카카오 증오발언 대응 정책 녹서

원칙 수립을 위한 숙의 과정의 기록