정책연구 25-20

정책 네트워크 구성·운영

2025. 12

연구기관 : 정보통신정책연구원

방송미디어통신위원회
Korea Media and Communications Commission

정책연구25-20

정책 네트워크 구성·운영

조성은 외

2025. 12

연구기관 : **정보통신정책연구원**

방송미디어통신위원회
Korea Media and Communications Commission

제 출 문

방송미디어통신위원회 위원장 귀하

본 보고서를『**정책 네트워크 구성・운영**』의 연구결과보고서로 제출합니다.

2025년 12월

연 구 기 관 : 정보통신정책연구원
총괄책임자 : 조성은 연 구 위 원
참여연구원 : 문정욱 연 구 위 원
연소라 부연구위원
이으뜸 부연구위원
이시직 부연구위원
김지혜 전문연구원
이지호 연 구 원
연제선 연 구 원
이세인 연 구 원

목 차

표 목 차

그 림 목 차

요 약 문

생성형 인공지능의 확산으로 인공지능의 사회적 파급력은 날로 커지고 있다. 방송·통신 서비스가 디지털 플랫폼 및 다양한 인공지능 서비스 모델과 결합하면서 새로운 측면의 이용자 권익 침해 이슈가 등장하고 있으며, 이에 대응하기 위한 정책·입법 논의도 활발하게 이루어지고 있는 실정이다. 지능정보사회의 진전으로 인해 다양한 이해관계가 형성되고 이용자 지위도 다변화함에 따라, 종래와 같은 정부 주도의 이용자 정책 수립 및 하향식 규제 방식은 다면적인 이용자 권익 침해 상황을 해결하는 데 한계를 지닌다. 지능정보서비스 생태계의 구성 요소, 참여 주체, 권익 침해 상황의 다양성과 특수성을 고려하는 동시에, 국가의 경계를 넘어 범지구적 차원의 서비스 제공 및 정보 유통 범위를 포섭하는 이용자 정책 논의를 위해서는 민관 협력뿐만 아니라 국제적 담론을 통한 정책·제도 개발이 요구된다.

본 연구는 급변하는 기술·산업 상황과 새로운 이용자 보호 이슈를 파악하여 당면 과제를 논의하는 민관학연 협의체를 구성하고, 사회적·국제적 논의 결과를 신규 정책 개발 및 제도 개선으로 연결함으로써 명실상부한 정책 네트워크를 구현할 목적에서 연속사업으로 수행해왔다. 이러한 다원적 정책 네트워크는 정부 뿐 아니라 기업, 학계, 시민단체, 일반 이용자의 적극적 참여와 의견 개진을 장려함으로써 궁극적으로는 사회적 수용성 높은 이용자 정책을 개발하는 실질적 통로 마련에 그 목적이 있다.

구체적으로 첫째, 2020년 출범하여 정책협의체로서 기능해 온 '지능정보사회 이용자 보호 민관협의회'(2022년~2023년)를 계승하고 발전시켜 '인공지능 서비스 이용자 보호 민관협의회'를 작년에 이어 2회 개최하였다. 2025년도 이루어진 제3기 인공지능 서비스 이용자 보호 민관협의회는 인공지능으로 인한 이용자 보호 이슈에 실질적으로 대응하기 위한 규제적 패러다임 전환 필요성과 실효적인 법제 접근 방식의 필요성을 논의하였다. 인공지능 기술이 사회의 각 영역과 결부되어 폭발적인 파급을 일으킬 것이라는 우려가 점차 현실로 드러남에 따라, 기술적·규제적 패러다임 전환이 이루어지고 있으며 이에 대응하기 위한 정책적 방향성 논의에 중점을 두었다.

둘째, 이용자 보호 정책을 위한 정책 협력체계를 구축하고자 2025 인공지능 서비스 이용자 보호 컨퍼런스를 개최하였다. 국내 정부부처, 연구기관, 학계 전문가 및 빅테크 등 분야 별 전문가를 초청하여 인공지능 서비스 이용자 보호 관련 현안을 논의하고 국내외 인공지능 서비스 연구 동향, 정책 현황, 실제 이용자 보호 이슈, 이용자 보호 관련 성과를 공유하고 발표하였다. 특히 올해는 생성형 인공지능이 초래한 사회경제적 충격과 변화를 논의하는 장을 마련하였다. 이를 위해 '인공지능 시대의 이용자 보호: 도전과 기회'를 주제로 인공지능 서비스에 대응하는 정책적 노력을 공유하고, 생성형 인공지능이 야기하는 다양한 문제점과 역기능을 해소하고자 하였으며, 인간과 인공지능의 협력과 갈등, 그 한계를 타진함으로써 실체적인 이용자 보호로 나아가는 지평을 확장하고자 하였다.

셋째, 인공지능 서비스 관련 이슈에 대해 개인(이용자)·시민단체·기업 등 이해당사자의 의견을 상시 수렴하고 논의할 수 있는 소통채널이자, 이용자정책 관련 최신 연구 및 국내외 정책 동향 등을 누구나 자유롭게 이용할 수 있는 플랫폼으로 지난 2020년 12월 '지능정보사회 이용자정책 아카이브(user-archive.kisdi.re.kr)'를 처음 공개하였다. 이후 2024년에는 '인공지능 서비스 이용자정책 아카이브'로 명칭을 변경하여 빠르게 변화하는 인공지능 서비스와 이용자 보호 관련 정보를 공유하고자 노력하였다. 올해는 통계 데이터베이스와 관련하여 사용자 친화적 인터페이스 구현을 위한 업데이트, 국가승인통계인 이용자 패널조사의 3개년치 변화추이를 살펴볼 수 있는 시계열 자료 배치 등 아카이브 기능 및 구현 개선에 중점을 두고 사업을 추진하였다. 또한, 아카이브 이용자를 대상으로 이벤트를 진행하여 아카이브 개선 방향 모색과 이용자들의 참여를 독려하고자 하였다. 향후에도 '인공지능 서비스 이용자정책 아카이브'의 서비스 안정화와 이용 활성화에 더욱 중점을 두고 운영할 계획이며, 아카이브 기능을 강화하고 사용자 친화적 인터페이스로 아카이브를 개선해 나갈 계획이다.

제1장 서 론

1. 인공지능 서비스 이용자 보호 민관협의회 운영 목적

본 민관협의회는 지능정보사회의 새로운 이용자 보호 정책 방향에 대한 사회적 공감대를 형성하고, 기업과 시민 등 이해당사자의 책임 의식을 제고하여 지속 가능한 이용자 보호 환경을 조성하는 데 궁극적인 기능과 목적이 있다. 이를 위해 민·관·학 인사가 참여하는 정책 민관협의회를 구성하고 점진적으로 확장해 오면서, 시의적인 인공지능 서비스 이용자 정책 개발을 위한 사회적 합의 기반을 강화해 왔다. 앞으로도 본 민관협의회는 인공지능 서비스의 다변화와 보편화 양상에 주목하여 새롭게 부각되는 인공지능 서비스 이용자 보호 이슈를 검토하고 그에 적절히 대응할 수 있는 이용자 정책을 개발하고자 급변하는 인공지능 서비스의 변화 양상과 주요 변수를 파악하는 일을 우선시하고자 한다.

이러한 목적의 일환으로, 2024년, 인공지능 서비스에 초점을 맞추어 그간의 '지능정보사회 이용자 보호 민관협의회'를 확대하여 인공지능 서비스 이용자 보호 민관협의회'를 출범하였다. 특히 올해는 생성형 인공지능 서비스의 보편화와 그에 따른 실제 이용자 권익 침해 이슈가 점차 가시적인 결과로 나타나고 사회적 우려를 자아내고 있기에, 실질적인 법제도적 측면에서 이용자 보호를 위한 정책을 논의하고 개발하는 데 주력하였다. 이는 높은 변동성을 보이는 생성형 인공지능 서비스 영역에서도 2019년 「이용자 중심의 지능정보사회를 위한 원칙」의 명맥을 잇는 이용자 정책을 모색하고 구현하는 데 궁극적인 목적을 둔 것이다.

2. 이용자 정책협력 네트워크 구축 목적

다음으로 정례적으로 운영하였던 글로벌 이용자 정책협력 네트워크를 계승하여, 금년도에는 국내의 인공지능 서비스 이용자 이슈에 초점을 맞추고 국내의 법과 문화적 특수성을 고려하는 등 내실을 다지기 위한 국내 이용자 정책협력 네트워크를 운영하였다. 생성형 인공지능을 위시한 인공지능의 핵심기술에 기초하여, 방송·통신 미디어 기술과 서비스가

내재하고 있는 이용자 보호 이슈와 유관 정책 방안을 공유하고 지향점을 모색하기 위한 논의의 장을 마련하였다. 나아가 국내에서 추진하고 있는 이용자 보호 정책 및 사회적·정책적 이슈와 함의를 산·학·연 전문가와 공유하고, 안전과 신뢰가 담보된 이용자 중심의 지능정보사회 실현을 위한 실효적 정책 방향을 제시하는 한편, 방송통신위원회가 추진하고 있는 방송통신 이용자 보호 정책에 대한 공감대 형성을 기하고자 하였다.

이는 2021년 방송통신위원회가 발표했던 「인공지능 기반 미디어 추천 서비스 이용자 보호 기본원칙」과 2023년 발표한 「메타버스 이용자 보호 기본원칙」의 실효성을 직간접적으로 진단하고 실체화된 정책을 재확인하는 노력이라 할 수 있다. 이러한 맥락과 배경 하에 '2025 인공지능 서비스 이용자 보호 컨퍼런스(AI, 신뢰를 만나다: 이용자 보호와 지속가능한 성장)'를 기획하고 개최함으로써, 방송통신위원회가 지원하고 정보통신정책연구원이 수행한 그간의 인공지능 서비스 이용자 보호 연구 노력과 학계와 산업계의 인공지능 서비스 이용자 보호 이슈를 공유하고 숙의함으로써, 국내 정책전문가·학계, 기업, 시민단체 등과 함께 인공지능 서비스 이용자 보호 정책에 대한 실천적 아젠다와 방향성을 제시하고자 하였다.

3. 국민참여 채널 및 지식공유 플랫폼 구축 · 운영 목적

정보통신정책연구원은 '지능정보사회 이용자 보호 환경조성' 사업의 일환으로 지난 2020년 12월 '지능정보사회 이용자정책 아카이브(user-archive.kisdi.re.kr)'를 처음 공개하였고, 인공지능의 사회적 확산과 인공지능 서비스를 이용하는 이용자들이 증가함에 따라 2024년, '인공지능 서비스 이용자정책 아카이브'로 명칭을 변경하여 운영하였다. '인공지능 서비스 이용자정책 아카이브'는 인공지능 서비스 관련 국내외 정책 동향과 최신 연구 자료 등을 누구나 자유롭게 이용할 수 있는 지식공유 플랫폼이자, 인공지능 서비스 이용자 관련 이슈와 정책에 대해 다양한 이해관계자의 의견을 상시 수렴할 수 있는 소통 채널을 목표로 한다. 올해 6년차에 접어든 아카이브는 3개년 치 이용자 패널조사 데이터를 시계열로 분석할 수 있는 기능 고도화와 아카이브 UI/UX 개선, 시스템 안정화를 위해 노력하였다. 향후에도 아카이브를 통한 의견 수렴으로 이용자 참여를 독려하고, 연구 결과를 활용하는 2차 연구를 장려할 계획이다. 또한, 산·학·연 전문가들을 중심으로 활성화되는 이용자 보호에 대한 다양한 논의 결과를 업데이트하여 환류 플랫폼으로 아카이브를 운영할 계획이다.

제2장 인공지능 서비스 이용자보호 민관협의회 운영 및 개최

제1절 인공지능 서비스 이용자보호 민관협의회의 목적

인공지능 기술 기반의 다양한 서비스가 우리 사회 전반에 확산되고 일상에 침투함에 따라 크고 작은 변화들이 생겨나고 있다. 새로운 기술 환경은 일상을 편리하게 만들어주는 동시에 예기치 못한 부작용을 야기하기도 한다. 이처럼 광범위한 사회적 위험성을 최소화하기 위해서는 선제적이고 실효적인 대응 방안을 부단히 모색하고 강구할 필요가 있다. '인공지능 서비스 이용자 보호 민관협의회'는 인공지능 서비스 이용자 보호를 위해 전문가의 의견을 수렴하고, 국내외 사례를 공유하여 이용자 보호의 필요성과 방향에 대한 공감대를 형성하고 책임 의식을 제고하기 위한 협의체 구축을 목적으로 구성되었다. 민관협의회는 공공, 기업, 시민단체, 법조계, 학계로 구성된 정책 협의체로서 일선 전문가들의 논의를 거쳐 이용자 보호를 위한 실효적인 정책 방안을 모색해 왔다.

민관협의회는 2018년 지능정보화 이용자 정책 간담회 형태로 지능정보사회에 걸맞은 이용자 보호 정책 방향 수립을 위한 논의를 시작하였다. 당시 민관협의회는 산·관·학·연 전문가들이 참여하여 이해관계자와 지능정보사회 구성원들이 공감할 수 있는 원칙을 마련하고자 의견수렴 과정을 거치면서 2019년 11월 「이용자 중심의 지능정보사회를 위한 원칙」 발표의 초석을 마련하였다. 2020년 민관협의회에서는 「이용자 중심의 지능정보사회를 위한 원칙」의 실질적 이행을 위해 국내외 사례(Best Practices)를 공유하는 등 연구 기능 및 정책 자문 기능을 수행하는 한편, 관련 이슈에 대한 논의의 장을 이어가며 구체적인 실천 방안을 모색하였다. 또한 2021년 민관협의회에서는 「인공지능 기반 미디어 추천 서비스 이용자 보호 기본원칙」에 담긴 핵심 원칙과 실행 원칙을 의제로 하여 합리적 규제 정책 방향을 논의하였다. 그리고 2022년 민관협의회에서는 대중적 서비스로 급부상한 신유형 지능정보서비스와 관련하여 이용자 보호의 공백과 수요를 파악하고, 나아가 해당 서비스에 영역에서 「이용자 중심의 지능정보사회를 위한 원칙」을 실현하기 위한 민관 협력 과제를 논의하였다. 2023년 민관협의회에서는 생성형 인공지능으로 급변하는 지능정보서비스

환경에서 신규 이용자 정책을 조망하고, 세계적인 인공지능 산업에 대한 규제의 움직임에 대하여 이용자 보호 과제 및 민관의 협력적 정책 방안을 논의하였다. 2024년 민관협의회에서는 인공지능 기술과 서비스의 사회적 여파가 드러남에 따라, 인공지능 관련 법제 현황과 방향성을 검토하고 세분화 중인 개별 이용자 보호 이슈에 대한 대응방안을 논의하였다. 마지막으로, 2025년 민관협의회는 인공지능 서비스의 폭발적 성장과 다양한 이용자 보호 이슈의 축적이 이루어지고 있는 상황에서 규제적 공백을 메우기 위한 실질적인 법제 마련 방안에 대해 산·관·학·연 인사들의 깊은 논의가 이루어진 대화의 장이 되었다.

제 2 절 민관협의회 구성 및 운영

인공지능 서비스 이용자 보호 민관협의회는 법학, 공학, 철학, 미디어 등 다양한 분야의 학계 전문가를 비롯하여 법조계, 기업, 공공, 시민단체 등 주요 이해관계자를 대표하는 영역에서 이용자 정책 실무를 담당하는 총 36인의 위원들로 구성되었다. 인공지능 서비스 이용자 보호 민관협의회 위원 명단은 다음과 같다.

〈표 2－1〉 인공지능 서비스 이용자 보호 민관협의회 위원 명단

분야	성명	소속 및 직위
학계	이원우	서울대학교 법학전문대학원 교수
	김명주	서울여자대학교 정보보호학과 교수
	김용대	서울대학교 통계학과 교수
	김형주	중앙대학교 인문콘텐츠연구소 교수
	변순용	서울교육대학교 윤리교육학과 교수
	송세경	KAIST 전기전자공학부 연구교수
	이성엽	고려대학교 기술경영전문대학 교수
	조성배	연세대학교 컴퓨터과학과 교수
	최경진	가천대학교 법학과 교수
	황용석	건국대학교 미디어커뮤니케이션학과 교수
법조계	박민철	김앤장 법률사무소 변호사
	손도일	법무법인 율촌 변호사
	윤종수	법무법인 광장 변호사
	장준영	법무법인 세종 변호사
시민단체	정지연	한국소비자연맹 사무총장
	신민수	녹색소비자연대 공동대표
	황다연	소비자와함께 공동대표
공공	신영규	방송미디어통신위원회 이용자정책국장
	김혜숙	방송미디어통신위원회 인공지능이용자보호과장
	문정욱	KISDI 디지털사회전략연구실 실장
	조성은	KISDI 연구위원

분야	성명	소속 및 직위
기업	박선민	구글코리아 대외정책협력 상무
	허 욱	메타코리아 대외정책 부사장
	김영훈	AWS 정책협력실장
	손지윤	네이버 정책전략총괄 전무
	우영규	카카오 ER성과리더
	박완진	KT AI기술협력담당
	김상목	SKT AI엔터프라이즈 사업담당 부사장
	성준현	LGU+ AX 추진담당 상무
	김유철	LG AI 연구원 전략부문 부문장
	이세영	뤼튼테크놀로지스 대표이사
	윤희식	한국마이크로소프트 정책협력법무실 이사
	이영복	제네시스랩 대표이사
	이경일	솔트룩스 대표

2025년 민관협의회 전체 회의는 총 2회 개최되었으며, 오프라인 형식으로 진행되었다. 2024년 1차 및 2차 전체 회의에 이어 개최된 2025년 3차 전체 회의에서는 인공지능으로 인한 정보통신 패러다임의 변화와 이에 따른 규제적 패러다임 변화 및 실효적 법제 접근 방식이 논의되었고, 4차 전체 회의에서는 인공지능 기술 고도화 및 역동적인 규제 환경 변화에 따른 이용자 보호 이슈 및 대응 방안에 대한 논의를 진행하였다.

두 차례 전체 회의에서 이루어진 상세한 논의 내용은 다음 장에서 소개하고자 한다.

제3절 민관협의회 의제 및 주요 논의 사항

2025년 한 해 동안 민관협의회 전체회의를 통해 진행된 주요 논의사항은 다음과 같이 요약된다.

〈표 2－2〉 2025년 인공지능 서비스 이용자 보호 민관협의회 개최 현황

회차	개최 일자	주요 논의사항
3차	2025. 5. 15.	• 기술환경 변화에 따른 통신서비스에서의 불법·유해정보 규율 변화 필요성을 살펴보고, 이용자 권익 증진을 위한 정책적·제도적 추진 방안 논의
4차	2025. 12. 15.	• 국내외 규제 동향 및 사업자의 기술적·관리적 투명성 확보 노력 등을 살펴보고 자율규제 차원의 AI 투명성 제고 방안 논의

자료: 연구진 작성

1. 2025년 인공지능 서비스 이용자 보호 민관협의회 3차 전체회의

인공지능 서비스 이용자 보호 민관협의회 3차 전체회의는 5월 15일(목) 오후 2시 엘타워 멜론홀에서 대면 회의로 진행되었다. 당일 민관협의회는 'AI서비스 대두로 인한 통신서비스에서의 불법·유해정보 규율 및 이용자 권익 증진방안'을 주제로 한양대학교 법학전문대학원 선지원 교수의 발제가 이루어졌으며, 이후 '기술환경 변화에 따른 통신서비스에서의 불법·유해정보 규율 변화 필요성을 살펴보고, 이용자 권익 증진을 위한 정책적·제도적 추진 방안 논의'을 주제로 논의가 진행되었다. 선지원 교수는 발표를 통해 인공지능 기술의 진보가 통신 서비스 제공 방식과 이용자 경험을 근본적으로 변화시키고 있음을 시사하며, 규제권자, 서비스 제공자, 이용자 간 상호 이해를 위한 논의의 필요성을 제기하였다. 이후 인공지능 기술 고도화 및 역동적인 규제 환경 변화에 따른 이용자 보호 이슈 대응 방안 마련과 정책 방향 모색을 위한 종합 토론이 진행되었다.

〔그림 2－1〕 2025년 인공지능 서비스 이용자 보호 민관협의회 3차 전체회의

자료: 연구진 작성

가. 주요 발제 내용

1) [발제] AI서비스 대두로 인한 통신서비스에서의 불법·유해정보 규율 및 이용자 권익 증진방안

인공지능 서비스 이용자 보호 민관협의회 3차 전체회의는 한양대학교 법학전문대학원 선지원 교수의 'AI서비스 대두로 인한 통신서비스에서의 불법 · 유해정보 규율 및 이용자 권익 증진방안'에 대한 발제로 시작하였다. 선지원 교수의 발제는 1) 통신 서비스 환경 변화, 2) 새로운 통신서비스 환경의 특징과 규율 변화 필요성, 3) 변화한 AI 기반 통신서비스 규제 접근 방법 모색, 4) 향후 과제로 구성되었으며, 통신 서비스 환경 변화에 대해 아래 <표 2－3>와 같이 설명하며 발제를 시작하였다.

〈표 2－3〉 통신 서비스 환경 변화 주요 내용

구분	주요 내용
논의의 필요성	• 고도화된 AI 기술과 통신망의 발전이 디지털 환경에서 근본적 변화 야기 • 특히, 통신 서비스 제공 방식과 이용자 경험에 큰 변화 • 개별 애플리케이션 이용 방식에서 벗어나, 통합 AI 에이전트 중심 이용 • 현재 이용자들은 원하는 콘텐츠와 서비스를 공급받는 형태로 변화 중 • 기존 불법·유해 정보 규제 체계 변화 필요성 제기

구분	주요 내용	
통신 서비스 환경 변화	망의 고도화	• 광대역 통합망 구축이 이루어지며, 규제 역시 인터넷과 데이터 통신을 감안하여 변화해 옴 • 5G의 도입으로 망 역할의 확장 가능성 발생 • 고도화된 AI 서비스가 원활하고 보편적으로 작동하는 데 필요한 기반 제공 중 • 초고속·초저지연 통신망을 통한 활발한 AI 에이전트 활용 가능성 증대
	애플리케이션의 종언	• 데이터 처리를 GPU가 아닌 통신망을 통해 수행하는 기법이 인공지능 기술과 통신망에 대한 패러다임 변화 촉발 • 효율적 데이터 처리 방식을 통해, 서비스 제공 가치사슬의 종적 압축(통신 서비스 자체에서 혁신 서비스를 제공하는 것)과 횡적 압축(다른 영역의 서비스들의 통합)이 가능해짐 • 이용자들이 특정 서비스를 개별 애플리케이션을 통해 접근하는 방식에서, 통합된 AI 에이전트와 상호작용하여 다양한 기능과 콘텐츠에 접근하는 방식으로 전환
	이용자 관점에서의 변화	• 통합 AI 서비스 제공 플랫폼 이용 • 업로드 되어 있는 콘텐츠 이용이 아닌, 요구에 맞게 생성된 콘텐츠를 이용 • 기존의 콘텐츠 관리 체계가 유효성 의문 제기

자료: 연구진 작성

선지원 교수는 인공지능 기술이 통신 서비스 환경 변화에 미친 영향을 설명하며 발제를 시작하였다. 거대언어모델(Large Language Model; LLM)과 생성형 인공지능이 통신 서비스의 제공 방식과 이용자 경험을 근본적으로 변화시키고 있으며, 이는 인터넷이나 모바일 통신 초기 등장에 버금가는 패러다임 전환을 예고하고 있다는 것이다. 또한 그는 이용자들의 인공지능 서비스 이용 방식의 차이를 부각하였는데, 기존 통신 서비스 이용자들은 개별 애플리케이션을 통해 특정 서비스를 이용하는 방식을 취했지만, 현재 변화 중인 통신 서비스 이용방식은 이용자들이 통합된 인공지능 에이전트를 중심으로 이용자 개인이 원하는 콘텐츠와 서비스를 공급받는 형태라는 점을 강조했다. 또한, 인공지능 에이전트가 단순한 도구를 넘어, 이용자의 목표를 이해하고 자율적으로 작업을 수행하는 시스템으로 진화하고 있어 향후 통신망을 통해 어떠한 콘텐츠가 이용자에게 제공될지 그 누구도 예상할 수 없다고 덧붙였다.

선지원 교수는 이에 대해 "현행 규제는 주로 정적이고 식별 가능한 콘텐츠를 중심으로

설계되었으나, 향후 서비스 유형, 콘텐츠, 단말 간의 경계가 모호해지고 이용자 경험이 동적으로 변화할 것으로 예상되기에, 기존 규제의 유효성과 적용 가능성이 약화될 것" 이라는 우려를 제기하였다. 선지원 교수는 통신 서비스에 대한 전통적인 역무 구분 방식, 콘텐츠 정의, 제공자 책무 할당 방식 등이 새로운 정보통신 기술 및 서비스 환경에 부합하지 않을 가능성 또한 높다고 밝혔다. 이러한 우려를 종합하여, 그는 기존의 불법 · 유해 정보 규제 체계에 대한 근본적인 변화 필요성을 제기하며, 규제 패러다임의 변화와 방향 설정을 위한 논의의 중요성을 토로했다.

이러한 목적을 가지고, 선지원 교수는 통신 서비스 환경 변화에 대한 세부적 현황을 공유하였다. 첫째로, 우리나라는 일찍부터 통신 기반에 대한 중요성을 인식하고, 망에 대한 투자를 장래의 성장 동력으로 삼아왔기에 많은 혁신적인 애플리케이션들을 선제적으로 시험할 수 있는 기반을 갖추었다는 평가가 있다는 것이다. 둘째로, 기존 유선 통신망에서 음성 · 데이터, 유 · 무선 등 통신 · 방송 · 인터넷을 융합시킨 광대역통합망의 구축 역시 빠르게 이루어지며, 규제 역시 인터넷과 데이터 통신을 감안한 내용으로 변화해 왔음을 설명하였다. 그리고, 마지막으로, 5세대 이동통신(이른바 '5G')의 도입을 통해, 망이 실어나를 수 있는 트래픽이 기존의 4G에 비해 최대 100배 수준으로 확장되고, 연결 속도와 안정성 역시 비약적으로 증대되면서, 망의 역할에 대해 근본적 변화의 가능성이 열렸다고 밝혔다.

선지원 교수에 따르면, 이러한 현황은 추후 고도화된 인공지능 서비스가 원활하고 보편적으로 작동하는데 필요한 기반을 제공하며, 인공지능 기술 또한 5G 및 6G 망과 융합되어 관련 환경 자체를 변화시킬 것이라는 전망으로 이어진다. 인공지능을 통해 트래픽 예측, 에너지 효율 최적화, 통신 기반의 확장성 강화 가능성이 이러한 전망을 뒷받침하고 있다고 덧붙였다. 즉, 통신망의 진보와 인공지능 기술의 진보는 양방향적으로 상호작용하며 서로의 변화를 촉발하고 있다는 것이다. 망의 진보는 인공지능 기술 및 서비스의 적용 가능성을 가속화하고 있으며, 가속화된 인공지능 기술은 역으로 통신망의 최적화와 데이터 처리 방식 등에 긍정적인 영향을 미치며 발전할 가능성이 농후한 상황이다. 더 나아가, 초고속 · 초저지연 통신망은 사용자의 기기나 클라우드 인터페이스를 통해 실시간으로 인공지능과 데이터를 주고 받을 수 있게 한다. 이는 인공지능 서비스의 편의성 증대와 함께 빠른 확산을 일으킬 것이고, 새로운 서비스 패러다임의 등장이 예견되고 있다.

다음으로 선지원 교수는 통신 서비스 환경 변화 중 하나로 '애플리케이션의 종언'이라는 용어를 사용하며 설명을 이어나갔다. 현재 인공지능 에이전트가 다양한 앱을 연동하고 있으며, 멀티 모달을 활용해 다양한 정보 유형을 처리하고 또한 다양한 기능을 제공하고 있는 상황이다. 해당 과정에서 인공지능 에이전트를 이용한 통신 서비스는 가치 사슬의 종적 압축과 횡적 압축이 가능해져, 하나의 인공지능 서비스가 다양한 영역의 서비스들을 통합 및 대체하는 현상이 다가오고 있다. 가치 사슬의 종적 압축이란 통신 서비스 자체에서 혁신 서비스를 제공하는 것이고, 횡적 압축이란 전혀 다른 분야의 애플리케이션이 통합되어 하나의 서비스로 제공되는 형태를 말한다. 이를 통해 이용자들은 고정된 이용자 경험을 갖는 것이 아닌, 인공지능 에이전트가 생성하는 비특정 콘텐츠로부터 유동적인 이용자 경험을 갖게 된다. 뉴스, 검색, 엔터테인먼트 등 특정 서비스를 개별 애플리케이션을 통해 접근하던 방식에서, 통합된 인공지능 에이전트와 상호작용하며 다양한 기능과 콘텐츠에 접근하는 방식으로 전환이 이루어진 것이다.

선지원 교수는 이용자 관점에서의 변화에 초점을 맞추어, 통신 서비스 환경 변화에 대한 설명을 이어 나갔다. 현재 인공지능 에이전트를 제공하는 서비스들은 점차 플랫폼화되어 소수의 강력한 인공지능 에이전트를 하나의 통합된 플랫폼에서 제공하고 있다. 그리고 이용자의 상호작용 대상 또한 에이전트의 형태를 띄며, 이용자가 필요한 정보 혹은 콘텐츠에 특화된 에이전트가 실시간으로 대체되며 제공하는 정보와 콘텐츠도 실시간으로 바뀌고 있다. 이용자가 더 이상 제공된 콘텐츠의 수동적 수용자가 아니라, 자신의 디지털환경과 소비하는 콘텐츠를 능동적인 행위자로 변화하는 상황이다. 이러한 환경에서 "기존의 콘텐츠 관리 체계가 여전히 유효하다고 말할 수 있는가?"라는 의문을 제기하며 선지원 교수는 새로운 통신서비스 환경의 특징과 규율 변화의 필요성에 대한 주제로 설명을 이어 나갔다. 새로운 통신서비스 환경의 특징과 규율 변화 필요성의 주요 내용은 다음 <표 2-4>와 같다.

〈표 2－4〉 새로운 통신서비스 환경의 특징과 규율 변화 필요성 주요 내용

구분	주요 내용	
경계의 모호화	융합	• 기간통신서비스와 부가통신서비스의 경계 구분 모호 • 통신서비스와 단말장치의 경계가 무너지고 있음 • 콘텐츠 공급과 통신서비스의 경계가 무너지고 있음
미디어 환경 변화	동적 환경 창조	• 이용자들이 정보 환경과 서비스 경험을 능동적으로 구축 • 능동적 역할을 고려한 규제 방식 필요 • 인공지능의 기술적 특성과 이용자들의 이용 과정을 고려한 접근 방식 제안
규율 변화 필요성	현 체제의 한계	• 불법·유해정보 규제 체제 현주소 • 현재 정보통신망법은 콘텐츠 유형에 초점을 맞추고 있고, 사후적 조치에 집중 • 전기통신사업법의 콘텐츠 규율은 사전적 기술 조치 의무를 부과하는 내용의 규제로 진화

자료: 연구진 작성

선지원 교수는 새로운 통신서비스 환경의 특징 중 하나로 경계의 모호화를 언급하며, 이에 대한 설명을 이어 나갔다. 선지원 교수는 인공지능 서비스의 융합적 특성이 다양한 영역에서 기존 체계 간 경계를 모호하게 만들고 있다고 설명하였다. 그 예로, 법제적 영역에서 기존 통신서비스 규제 체계는 인공지능 서비스를 규제하기엔 역무 간 경계가 명확치 않다는 것이다. 인공지능 에이전트는 기본적인 통신 기능(메시지, 음성통화 등)과 복잡한 정보 처리 및 서비스 접근 기능을 통합하여 제공하고 있기에 전기통신사업법 상 기간통신역무와 부가통신역무 간 경계를 모호하게 만들고 있다.

추가로, 인공지능 기술은 통신서비스와 단말장치의 경계 또한 모호하게 만들고 있는데, 대표적인 예로 온디바이스 인공지능(On-device Artificial Intelligence)의 출현이다. 단말기 자체가 상당한 정보 처리 능력과 서비스 기능을 수행하게 되면서 네트워크 서비스와 단말기 기능을 분리하기 어려워졌다. 이는 단말기 내에서 통신 없이 인공지능을 이용한 정보 처리로 인해 데이터 프라이버시 강화, 실시간 처리, 네트워크 독립성 등의 장점을 가지지만, 규제적 측면에서는 서비스 제공 주체와 책임 소재 특정의 어려움을 야기하며 문제 소지를 발생시킬 수 있다. 경계의 모호화와 융합에 대한 마지막 예시로 콘텐츠 공급과 통신서비스에 대해 선지원 교수가 설명을 이어 나갔다. 인공지능 에이전트는 통신 과정의 일

부로서 콘텐츠를 동적으로 생성, 수정, 전달하기에, 콘텐츠 제공 행위가 통신 서비스 자체와 분리되기 어려워졌다는 점을 언급하였다.

두 번째 새로운 통신서비스 환경의 특징으로, 선지원 교수는 미디어 환경 변화와 동적 환경 창조에 대해 설명하였다. 현재 이용자들은 인공지능 도구와 플랫폼을 활용하여, 자신의 정보 환경과 서비스 경험을 능동적으로 구축하고 있다. 정보 제공자가 정의한 옵션을 수동적으로 받아들이는 방식이 아닌, 이용자가 적극적으로 상호작용하며 다양한 결과물을 지속적으로 생산하는 방식으로 변화하고 있다. 다양한 결과물들을 얻는 과정에서 이용자가 의도치 않게 유해한 콘텐츠 혹은 허위 정보들을 생성할 가능성이 발생하게 된다.

따라서 이용자의 능동적인 역할을 고려하는 방식으로 규제가 변화해야 한다. 기존 규제는 콘텐츠를 '유통'하는 제공자에게 초점을 맞추지만, 새로운 미디어 환경에서는 이용자가 공동 창작자 또는 창시자(initiator)가 될 수 있기에 이용자를 고려한 규제 방식이 필요하다. 새로운 규제는 서비스 제공자의 의무뿐 아니라 이용자의 책임과 리터러시를 고려한 방식으로, 능동적인 이용자 환경 속 인공지능의 역기능 방지를 위한 수단이 되어야 한다.

더 나아가, 최종 결과물로서의 콘텐츠만이 아니라, 알고리즘 설계, 데이터 사용, 투명성, 편향성 완화 등의 인공지능 설계 과정과 이용자들의 선호를 반영한 에이전트 구축 과정 등 '이용 과정' 자체를 다루어야 할 필요성 또한 증가하고 있다. 인공지능을 이용한 알고리즘은 방대한 이용자 데이터(행동, 선호도, 맥락 등)를 분석하여 고도로 개인화된 경험, 추천, 콘텐츠 스트림을 실시간으로 제공하고 있다. 이용자가 접하는 콘텐츠가 점차 개인화에 기반하여 동적으로 생성되거나 조합되고 있다. 반면, 기존 규제 체계는 정적이고 식별 가능한 콘텐츠를 대상으로 하기에, 이러한 동적이고 개인화된 콘텐츠 사용 및 경험에 대한 접근이 불가능하다. 따라서, 이용자들의 이용 과정을 고려한 규제 접근 방식이 필요하며, 현 규제 체제에 대한 전면적인 재검토 또한 필요한 상황이다.

선지원 교수는 미디어 환경에서의 규율 변화 필요성을 제기하며, 현 체제의 한계를 언급하였다. 선지원 교수는 불법·유해 정보 규제가 정보통신망법 및 전기통신사업법 중심으로 이루어지며, 콘텐츠 등급 분류와 방송미디어통신심의위원회의 사후 심의가 이를 보완하고 있다고 설명하였다. 정보통신망법과 전기통신사업법의 규율 내용은 <표 2-5>와 같다.

〈표 2－5〉 정보통신망법과 전기통신사업법 상 규율 내용

구분	규율 내용
정보통신망법	• 불법정보 유통 금지(제44조의7): 누구든지 정보통신망을 통하여 음란물, 명예훼손 정보, 공포심이나 불안감을 유발하는 정보, 해킹·악성코드 관련 정보, 청소년보호법상 의무를 이행하지 않은 청소년유해매체물(영리 목적), 불법 사행행위 정보, 국가기밀 누설 정보, 국가보안법 위반 정보, 기타 범죄 목적 또는 교사·방조 정보 등 불법정보를 유통하는 것을 금지 • 임시조치(제44조의2): 권리 침해(명예훼손, 사생활 침해 등)를 주장하는 이용자의 요청이 있을 경우, 정보통신서비스 제공자는 해당 정보에 대해 즉시 삭제, 30일간의 임시조치(접근 차단) 등 필요한 조치를 취해야 함 • 기타 사업자의 의무 조항: 일반적 의무 규정(제3조), 청소년 보호(제42조 이하), 합성영상 관련 의무(제4조의2, 2025. 6. 4. 시행 예정), 불법촬영물 유통 방지(제44조의9)
전기통신사업법	• 통신서비스 이용자 보호를 위한 적극적인 조치에 대한 포괄적 근거(제32조) • 이용자 보호를 위한 금지행위(제50조): 이용약관(제28조제1항 및 제2항에 따라 신고하거나 인가받은 이용약관만을 말한다)과 다르게 전기통신서비스를 제공하거나 전기통신이용자의 이익을 현저히 해치는 방식으로 전기통신서비스를 제공하는 행위(제5호) • 청소년유해매체물 등의 차단(제22조의5)

자료: 연구진 작성

정보통신망법은 주로 특정 유형의 불법 '콘텐츠'에 초점을 맞추고, 이용자 신고나 방송통신심의위원회의 심의에 기반한 사후적 조치에 집중하고 있다. 하지만, 이는 콘텐츠를 특정하기 어렵거나 제공자의 역할이 모호해져 유통의 개념 또한 모호해지는 경우와 같이, 인공지능이 생성하는 일시적이고 개인적인 서비스에 적용하기 어렵다는 문제를 지닌다.

현행 전기통신사업법 또한 인공지능과 관련해 해석할 때, 여러 한계를 지니고 있다. 첫째로, 이용자 주도의 동적 환경에서 콘텐츠에 대한 특정이 어려우며, 서비스 제공자 정의도 어렵다는 문제가 있다. 둘째로, 인공지능 에이전트가 생성한 콘텐츠 및 이용자의 프롬프트로 인해 변형된 콘텐츠에 대한 법적 책임 공백 문제가 있으며, 이는 공급자 단계가 아닌 이용자 단계에서 형성되는 서비스의 내용을 과연 개발자나 플랫폼이 책임질 수 있는 것인가에 대한 의문으로 이어지는 문제이다. 마지막으로, 환각 현상같이 인공지능으로 인해 발생한 새로운 유형의 위험에 대해 기존 규율 체계로 적용할 수 있는가에 대한 문제가 있다.

선지원 교수는 기존 규율 체제의 한계에 대해 설명하며, 이를 벗어나 이용자들의 이용 과정 자체에 초점을 맞춘 예시를 언급했다. 전기통신사업법의 콘텐츠 규율은 특정 유형의 심각한 유해 콘텐츠에 대해 플랫폼에 '사전적 기술 조치 의무'를 부과하는 내용의 규제로 진화하였는데, 이는 사후적 삭제 조치와 달리 유해 콘텐츠가 널리 퍼지기 전에 특정 기술적 조치(필터링 등)을 요구했다는 점에서 이용 과정에 대한 변화를 준 예시로 볼 수 있다. 즉, 특정 콘텐츠와 특정 사업자에 국한되기는 하나, 콘텐츠 자체에 대한 규제가 아닌 이용 과정 자체를 규제하는 방향으로 나아가는 과정이라고 평가할 수 있다.

다음으로, 선지원 교수는 세 번째 주요 목차인 '변화한 AI 기반 통신서비스 규제 접근 방법 모색'을 주제로 발제를 이어갔다. '변화한 AI 기반 통신서비스 규제 접근 방법 모색'의 주요 내용은 아래 <표 2-6>과 같다.

〈표 2-6〉 변화한 AI 기반 통신서비스 규제 접근 방법 모색 주요 내용

구분	주요 내용	
이용자 리터러시 중심의 접근	리터러시의 재정의	• 인공지능의 특성을 반영한 리터러시 개념 확장 필요성 제기 • 이용자 스스로의 인공지능 역기능 방어 기제 구축 유도 • 보호 대상에서 벗어나, 책임을 가진 주체로서의 이용자상 정립
	정책적 접근	• 이용자의 안전한 사용을 유도하도록 서비스 제공자 지원 • 인공지능의 특징을 고려한 정책 기반 자료 개발 및 확대 • 인공지능 리터러시 교육 확대
규제 초점의 전환	과정 중심 규율	• 온라인 상의 불법 콘텐츠에 대한 기존 규제는 여전히 유효 • 인공지능 서비스가 작동하는 과정을 고려해야 함 • 리스크 관리를 위한 프레임워크 구축 필요 • 공적 주체의 감독 및 집행 역량 강화 필요

자료: 연구진 작성

선지원 교수는 '변화한 AI 기반 통신서비스 규제 접근 방법 모색'에 대해 이용자 리터러시 중심의 접근을 강조했다. 선지원 교수는 리터러시라는 용어가 단순히 디지털 도구 사용 능력을 넘어 이용자의 권리와 책임을 고려한 디지털 시민성(digital citizenship)의 재발견적 개념으로 확장될 필요가 있다고 설명했다. 즉, 기존 리터러시 개념을 확장하여, ① 이용자가 정보를 비판적으로 평가하여 정보의 진위를 판별하고 편향성을 인지하며, ② 디지털 환경에서의 권리(프라이버시, 표현의 자유)와 책임(윤리적 창작, 유해 정보 확산 방지

등)을 이해하며 실천하는 능력으로 나아가야 한다는 것이다. 특히 인공지능 리터러시는 인공지능의 기본 원리, 인공지능의 능력과 한계, 이용자로서 윤리적 함의가 포함되어야 한다고 덧붙였다.

규제만으로 인공지능 에이전트의 유해 정보 생성 과정을 차단하기 어려우므로, 이용자 스스로 방어 기제를 구축할 수 있도록 유도하는 것이 이용자 인공지능 리터러시 함양의 목표이다. 인공지능 리터러시 개념의 보급은 일방적인 보호 대상으로서 이용자상이 아니라 책임을 가진 주체로서 이용자상 정립을 가능케 할 것이다. 선지원 교수는 특정 콘텐츠에 대한 차단이 아닌 이용자의 책임성과 능동성을 뒷받침할 수 있는 개념으로서 리터러시에 대한 재정의의 필요성을 다시 한번 강조하며, 정책적 접근 방식에 대한 설명으로 넘어갔다.

정책적 접근의 첫 방안은 서비스 제공자의 역할에 관한 방안이었다. 세부적으로는, 인공지능을 활용하거나 통합된 인공지능 에이전트를 제공하는 플랫폼 사업자가 인공지능의 기능과 한계에 대한 명확한 정보를 제공하고, 이용자가 알고리즘에 대한 설정을 제어할 수 있도록 안내 및 유도하는 방안이다. 두 번째 방안은 정책 기반 자료의 개발 및 확대로, 정책적 기반으로 활용되는 방송통신위원회와 시청자미디어재단의 미디어리터러시 지수 등에 인공지능 에이전트 서비스의 특징을 반영하고 보완함으로써, 정책적 타당성을 제고하고 디지털 권리 역량 강화 정책과 연계성을 도모할 수 있다. 마지막 정책적 접근 방안은 교육적 차원의 접근이었다. 초등교육을 비롯한 모든 교육과정에서 디지털 및 인공지능 리터러시 교육을 포함하거나, 시민사회의 협력을 기초로 한 사회교육 확대, 리터러시 교육 프로그램 제공을 위한 국가의 지원 등 이용자의 리터러시 함양 교육을 위한 사회적 노력 방안이 제시되었다.

인공지능 기반 통신서비스 규제 접근 방법의 일환으로, 선지원 교수는 규제의 초점이 전환되어야 함을 다시 한번 강조했다. 온라인상의 불법 콘텐츠에 대한 기존 규제는 여전히 유효한 측면이 있지만, 인공지능 기반 서비스가 작동하는 과정을 대상으로 하는 규율의 고려가 필요하다는 것이다. 유해한 결과물이라는 증상을 사후에 처리하는 데 그치지 않고, 시스템이 구축되고 배포되는 방식에 영향을 줌으로써 원류에서부터 피해를 예방하는 것을 목표로 규율 방안이 마련되어야 한다는 것이다. 다시 말해, 단순히 “이 콘텐츠가 불법인가?”를 묻는 대신, “이 시스템이 안전하게 설계되었는가?”, “위험 평가는 이루어졌는

가?", "과정이 투명한가?", "실패에 대한 책임 소재는 명확한가?" 등을 묻는 방식으로 규율 체계를 점검하여야 실효적인 규율 방안을 마련할 수 있다.

이러한 과정에서 인공지능 서비스의 알고리즘 투명성과 설명 가능성이 어느 정도 요구될 수 있다고 선지원 교수는 설명하였다. 하지만, 고도화된 인공지능 모델에서 세부적인 모델 설명은 오히려 이용자들의 리터러시를 저하시킬 수 있고 기술적 어려움과 영업비밀 보호 등도 감안해야 하기에, 적절 수준의 주요 변수나 설명에 대한 의무를 도입해야 한다고 덧붙였다.

선지원 교수는 리스크 관리를 위한 프레임워크의 필요성 또한 제기하였다. 인공지능 서비스에 대한 위험 평가를 수행하고, 완화 조치를 이행하고, 지속적인 모니터링을 실시하기 위한 거버넌스 체계 구축이 필요하다는 것이다. 이에 대해 인공지능 기본법상의 각종 의무(제27조 이하)를 재구성하여 고도화된 인공지능 에이전트 서비스에 적합하도록 편성될 필요가 있다고 선지원 교수는 덧붙였다.

선지원 교수는 과정 중심 규율 도입을 위한 세 가지 단계를 제시하며, 각각에 대한 예시를 제공하였다. 첫 번째 단계로, 이용자의 중대한 권리나 선택에 영향을 미치는 인공지능 시스템(예: 맞춤형 광고, 콘텐츠 추천 알고리즘)의 투명성을 확보하고 기준을 명확히 제시하는 방안이 예시로 제시되었다. 두 번째 단계의 예시로, 금융, 고용, 공공 서비스 등 사회적 영향력이 큰 고위험 인공지능 분야를 중심으로 에이전트 서비스의 리스크 평가 및 완화 계획 수립 의무화 검토가 제시되었다. 세 번째 단계로는, 인공지능 에이전트 서비스 환경에서의 중개자 책임 원칙을 재정립하여 인공지능 기반 서비스 제공자에게 콘텐츠에 대한 책임을 묻지 않고, 과정에 대해 더 높은 수준의 주의 의무(due diligence)를 요구하는 방안이 제시되었다.

인공지능 서비스의 발전과 변동성을 고려해, 자율규제 프레임워크에 대한 재검토 또한 필요하다고 선지원 교수는 말했다. 정부·산업계·학계·시민사회가 협력하여 인공지능 관련 표준을 개발하고 모범 사례들을 공유하며, 기술 변화에 따른 규제 조정을 논의하는 프레임 구축 노력이 필요하다고 첨언하였다. 또한, 공적 주체의 감독 및 집행 역량 강화가 필요함을 덧붙였다. 선지원 교수는 관련 규제 기관이 인공지능 기반 서비스를 효과적으로 감독하고 새로운 과정 중심 규제를 집행할 수 있도록 전문 인력을 양성하며 기술적 분석 도구를 확보하는 등의 역량 강화 방안을 제시하였다. 추가로, 관계 기관 간 협력 및 정보

공유 체계 구축의 필요성 또한 제기하였다.

선지원 교수는 발제의 마지막으로 주요 논점을 정리하며 향후 과제를 제시하였다. 첫째는, 인공지능 기술과 통신망의 고도화가 통신서비스의 패러다임을 근본적으로 변화시키고 있으며, 이는 대한민국의 현행 불법 및 유해정보 규제 체계에 대한 전면적인 재검토를 요구하고 있는 것이었다. 둘째는, 통합된 인공지능 에이전트의 등장, 서비스·콘텐츠·단말 간 경계의 모호화와 동적이고 개인화된 이용자 경험 확산은 기존의 콘텐츠 중심 규제 방식의 한계를 드러내고 있다는 것이었다. 셋째로, 인공지능이 생성하거나 변형하는 콘텐츠에 대한 책임 소재 불분명, 전통적인 콘텐츠 관리 방식의 비효율성, 환각 현상·딥페이크 및 알고리즘 편향과 같은 새로운 리스크 부상에 대처할 새로운 규율 체계 모색이 필요하다는 것이 언급되었다. 마지막으로, 고도화된 인공지능 기반 서비스 환경 속에 존재하는 모든 주체들의 역동성과 상호작용을 반영할 수 있는 새로운 차원의 정책 패러다임 논의가 이어지길 기대하며, 선지원 교수의 발제가 마무리 되었다.

나. 종합토론

선지원 교수의 'AI서비스 대두로 인한 통신서비스에서의 불법·유해정보 규율 및 이용자 권익 증진방안' 발제 이후, 인공지능 서비스 대두로 인한 통신서비스에서의 불법·유해정보 규율 변화 필요성, 이용자 권익 증진을 위한 정책적·제도적 추진 방안 논의를 위한 종합 토론이 진행되었다. 종합 토론의 진행은 이원우 서울대학교 법학전문대학원 교수가 담당하였으며, 공공, 학계, 시민단체, 산업계, 법조계의 전문가들이 참여하였다. 토론의 진행은 학계, 시민단체, 산업계, 법조계 순으로 발언이 이어졌으며, 자세한 내용은 아래에 서술하였다.

학계에서는 먼저, 인공지능 에이전트를 비롯한 ICT(Information communication technology) 서비스 관련 주체들 간의 거버넌스 및 규제 체계에 관한 시급한 논의가 요구된다는 의견을 제시했다. 인공지능 에이전트의 등장으로 기존의 ICT 부처들 간 영역 구분이 모호해지면서 네트워크와 단말기, 소프트웨어, 하드웨어 등 영역 간 주체들이 혼재되는 현상이 발생하고 있기 때문이라는 근거를 덧붙였다. 학계의 한 위원은 위와 같은 혼란을 완화하기 위해, 지금까지 명확하게 구분되었던 거버넌스나 규제들이 중복되지 않도록 규제 및 업무 등을 통합하려는 노력이 최우선시되어야 한다는 의견에 힘을 보탰다.

발제 내 인공지능 리터러시에 대한 동감을 표한 학계의 위원이 있었다. 해당 위원은 인공지능의 기술적 특성에 맞춘 세분화된 리터러시 교육의 필요성에 공감하며, 올해 초등학교 교육과정 속 초등학교 3학년부터 인공지능 윤리가 포함되었다는 관련 현황 또한 공유하였다. 하지만, 인공지능 리터러시는 기존의 디지털 리터러시와 성격 및 정보 환경의 차이가 크므로, 디지털 리터러시의 하위적 속성으로 다루는 것에 대해 의문을 제기했다. 또한, 해당 위원은 발제의 주요 내용과 같이 '과정'을 중심으로 규율되어야 한다는 점에 깊은 동의를 표하면서도, 규율 주체에 대한 논의 필요성을 역설했다.

학계에서는 과정 중심적 접근에 대해 동감하는 의견들이 제시되었다. 산업구조, 경쟁구조, 이용자 환경 등이 전면적으로 개편되고 새롭게 정의되어야 하는 현 시점이야말로 정책 조정 시기를 체계적으로 설정하고 검토할 수 있는 시기이고, 과정 중심적 접근이 유효할 것이라는 견해가 있었다. 인공지능 에이전트가 기존의 경계들을 무너트리며, 통신사업자들 또한 소프트웨어 투자를 늘리고 부가통신사업자들 또한 물리적 인프라 투자를 늘리며 상호 의존성이 높아지는 양상을 띄고 있다는 점과 이에 따른 새로운 협력 모델도 개발되고 있다는 점 등이 과정 중심의 규제 체계 검토 및 접근의 필요성을 뒷받침한다는 설명이 붙었다.

더 나아가, 한 위원은 인공지능으로 인한 산업구조, 경쟁구조, 이용자 환경 등의 전면적 개편 양상이 ICT 사회를 새롭게 정의해야 할 필요성을 증대시키고 있다는 의견을 내비쳤다. 서비스의 혼종성이 증가되고 있기에 인공지능 생태계를 중심으로 여러 계층을 구분하려는 시도가 어려워지고 변동성 또한 높아지고 있는 시기라며 과정 중심적 접근 시도와 자율 규제 재검토에 동의하면서도 누구의 관점에서 정의할 것인지 문제의식을 드러냈다. 기술의 변동성으로 인해 강한 경성규제가 시도되는 경우 규제 프레임과 현실 간의 괴리 및 누수 현상이 발생하기 쉬워 정책 조정 시기를 두면서 과정 중심 모델에 대한 면밀한 토론이 필요해 보인다며 입장을 정리하였다.

불법·유해정보 규율 방안에 대한 위원들의 의견도 뒤따랐다. 불법 및 유해정보 생성 방지를 위해, 학계의 다른 위원은 인공지능 또한 사회화 과정이 필요하다고 입장을 밝혔다. 인공지능의 사회화를 위한, 데이터 정제 및 파인튜닝 과정과 같은 사업자의 노력을 내부적 노력, 올바른 개발 정책과 기준을 제시하는 사회의 노력을 외부적 노력으로 비유하여 인공지능의 사회화를 설명하였다. 또한 어떻게 인공지능이 불법 정보를 생성하는지, 그리

고 우리 사회는 어떻게 불법 및 유해 정보라고 판단하는지, 그 과정과 결과에 대해 각자의 입장을 투명하게 공개하고 상호신뢰가 형성되어야 한다고 주장하였다.

한 학계 위원은 불법 및 유해정보 차단을 중심으로 인공지능 시대에 방송통신위원회의 역할에 대해 초점을 맞추어 논의하였다. 인공지능으로 인한 융합이 현실화되고 있는 시기에, 각 전문기관의 전문성 활용에 대한 필요성을 언급하며 논의가 시작되었다. 뒤이어 한 위원은 방송통신위원회가 이용자 보호, 불법 및 유해정보에 대한 관할권을 가지고 있지만, 인공지능 규제에 대한 관할권은 불명확한 상황이기에 이를 분명히 정리해야 하며, 사후적 대응이 아닌 모니터링 시스템과 같은 신속한 대응 체계를 만드는 것이 필요하다는 의견을 보였다. 또한, 인공지능 서비스에서 불법 및 유해정보 생성의 문제가 사업자가 전적으로 책임을 져야 하는 문제인지 재고의 필요성을 제기했다. 이용자의 악의적인 의도로 안전장치를 우회하여 생성된 유해 정보의 배포가 대표적인 예시로, 사업자의 책임 영역, 이용자의 책임 영역, 공동 책임 영역을 구분할 필요가 있다는 것이다.

다음으로는 시민단체의 의견이 이어졌다. 이용자의 인공지능 리터러시 함양에 대해 한 시민단체 위원은 일부 동의하면서도, 실현 가능한 부분인지 의문을 제기하였다. 이용자의 인공지능 리터러시 향상을 통한 이상적인 사용 방식 구축이 현실적으로 불가능하며, 설명 가능한 인공지능에 대한 연구가 이루어지고 있지만 이마저도 명확치 않아 사업자들 또한 생성 과정에 대한 명확한 이해가 불가능한 상황임을 의문의 근거로 제시하였다. 따라서, 인공지능 서비스 이용자들 또한 이러한 불명확성을 인지하더라도, 원하는 결과를 유도하는 것에 한계가 있으며 인공지능 리터러시가 능동적인 방어기제로써 작용하기 어려울 수 있다는 견해를 밝혔다.

시민단체의 또 다른 위원은 인공지능으로 인한 문제가 일상생활에서 체감됨에 따라, 방송통신 분야에서 인공지능을 활용해 생성한 콘텐츠에 대해 보다 엄격한 검증 및 검수 절차를 의무화해야 하고, 정보 비대칭성을 해소하기 위한 적극적인 노력이 요구된다고 입장을 밝혔다. 한 방송사에서 일어난 인공지능 오역 사태에 대해, 인공지능 기술이라는 점에 대해서 면죄부가 주어졌지만, 사람이 개입하지 않았다는 점 자체가 면죄부의 사유가 되지 않는다며, 이용에 대한 엄격한 책임이 필요하다고 주장했다.

다음으로, 산업계 위원들의 의견과 입장이 공유되었다. 산업계의 한 위원은 이용자와 기업, 정부기관 등 인공지능 이해 관계자들 간 투명한 정보 공유가 필수적이며, 인공지능 리

터러시 교육을 통해 인공지능 산출물에 대한 비판적 사고력 함양 및 실질적인 교육과정 활성화의 중요성을 역설했다. 이용자들이 인공지능의 역기능에 대비하기 위해 기업이 투명성을 제공해야 한다는 점에 공감하며, 공공기관들 또한 어떤 기준으로, 어떤 기술로, 어떤 피해를 방지하기 위해, 어떤 노력을 기울이고 있는지 정책적 투명성을 제공한다면 궁극적으로 이용자 보호를 투명하게 실현할 수 있을 것이라는 의견이었다. 해당 위원은 인공지능 리터러시 교육에 대해서는 비판적 사고에 대한 교육이 강화되길 소망하며 발언을 마쳤다.

산업계의 한 위원은 다음과 같은 의견을 밝혔다. “인공지능을 통한 불법 · 유해정보 생성 및 유포와 관련하여 규제를 시행한다면, 인공지능 학습을 위해 데이터를 제공하는 업체, 인공지능 모델을 개발하는 업체, 인공지능 모델을 도입하여 파인튜닝하는 업체, 파인튜닝된 모델을 제공받은 업체, 제공받은 모델로 서비스를 시작한 업체, 인공지능과 상호작용하여 불법 콘텐츠를 생성하는 생산자, 생성된 콘텐츠를 배포 및 유통하는 유통자 등 수 많은 관계자가 있는 상황에서 불법 · 유해정보 자체가 문제인지, 관여된 수 많은 사람들이 문제인지, 그 중 누구를 규제 대상으로 할 것인지 등 굉장히 세부적인 규제 체계에 대한 논의가 필요하며 이는 매우 복잡하고 어려운 문제”라며 난처한 입장을 드러냈다. 따라서, 인공지능 서비스 과정을 세분화하여 각 단계에 맞는 과정 중심의 접근이 필요하며 엄밀한 논의가 필요하다며 의견을 마무리하였다.

다른 산업계의 위원은 서비스 제공업자의 책임 강화에 동의하나, 기술적 한계로 인해 모든 유형의 불법·유해정보 생성을 원천 차단하는 데에는 현실적 제약이 존재하므로, 기업이 준수해야 할 최소한의 의무와 기준을 명확히 제시하는 것이 필요하다는 의견을 내비쳤다. 또 다른 위원은, 이용자 보호가 고객 가치적 측면과 사업적 측면 모두에서 중요하기 때문에, 이를 위한 노력들이 서로 공유되고 정책화되길 바란다는 입장을 보였다. 해당 위원은 사업의 불확실성 제거를 통한 빠른 사업 추진, 고객 개인정보 보호 및 위협 방지 등을 통한 고객 유치 등 이용자 보호의 가치가 상생을 위한 방안임을 설명하였고 산업계의 다른 위원들 또한 이에 대해 동의하였다.

산업계에서는 이용자의 인공지능 리터러시에 대해 강조하는 의견들이 존재했다. 기업들의 기술적 노력과 관리가 이루어진다고 하여도, 이용자들이 생성물을 이용하는 과정에서 기업으로서 제재할 수 있는 방안에는 한계가 있으므로, 올바른 활용에 대한 사회적 분위

기 조성이 필요하다는 의견이었다.

기타 의견으로, 인공지능 시대에서는 정부 주도의 규범 제시 보다는, 규제 형성 초기 단계에서부터 다양한 이해관계자들의 논의를 통한 규제 도출이 필요하다는 의견이 있었다. 현재 인공지능으로 인한 여러 사회변화가 야기되고 있으며, 제시된 규범이 현상을 따라가지 못하는 문제가 발생할 가능성이 높기 때문에, 기업의 자체적인 노력과 함께 적극적인 협력으로 규제를 형성해 나아가는 방안을 제시하였다.

이로써 학계, 시민단체, 산업계의 발언이 마무리되고, 법조계의 발언이 시작되었다. 먼저, 법조계의 한 위원은 인공지능 기반 불법 콘텐츠 확산에 대응하기 위해 불법 콘텐츠 유포자를 대상으로 한 규제의 필요성을 밝혔다. 인공지능과 관련된 불법 콘텐츠, 가짜 뉴스 등이 확산되고 그 피해나 문제가 심각해지고 있는 상황에서 불법 콘텐츠를 업로드하는 사람에게 책임을 부여하여야 이를 방지할 수 있다고 설명하였다. 이에 동조한 다른 위원은, 불법 및 유해 콘텐츠는 생성 이후 유통과정에서 문제가 필요하므로 유통자에 대한 규제가 필요하다는 의견에 힘을 보탰다. 다만, 유통 단계에서 문제가 발생하지 않도록 사전 생성 단계에서 이용자에게 충분한 정보를 제공하고 이용자가 개입할 수 있는 여지를 만들어야 한다는 의견 또한 제시하였다.

인공지능 알고리즘의 '블랙박스' 문제는 인공지능 기술의 본질적인 문제 중 하나이므로, 설명 가능성을 제시하지 못한다고 이에 기인하여 규제를 실시한다면 인공지능 서비스를 국내에 제공할 수 없다는 상황 설명도 공유되었다. 또한, 인공지능을 통해 허위 정보, 불법 정보가 나오기 때문에, 인공지능 서비스를 제재하고 규제하는 것은 쉬운 규제 방법이지만, 인공지능 서비스 의욕 자체를 제거하는 결과를 낳을 수도 있다는 우려가 제기되었다. 이러한 우려를 마지막으로, 민관협의회의 종합 토론이 종료되었다.

2. 2025년 인공지능 서비스 이용자 보호 민관협의회 4차 전체회의

인공지능 서비스 이용자 보호 민관협의회 4차 전체회의는 12월 15일(월) 오후 2시 엘타워 엘하우스홀에서 대면 회의로 진행되었다. 이날 협의회는 불필요한 규제 비용을 최소화하는 범위 내에서 방송·미디어·통신 분야에 적용되는 인공지능 서비스의 투명성 확보 필요성 및 수준을 주제로 황용석 건국대학교 미디어커뮤니케이션학과 교수가 '인공지능 시

대 투명성과 기술기반 자율규제'를 주제로 발표하였다. 발제자의 발표와 이후 종합 토론을 통해 국내외 규제 동향 및 사업자의 기술적·관리적 투명성 확보 노력과 자율규제 차원의 인공지능 투명성 제고 방안을 논의하는 종합 토론이 진행되었다.

〔그림 2-2〕 2025년 인공지능 서비스 이용자 보호 민관협의회 4차 전체회의

자료: 인공지능 서비스 이용자 보호 민관협의회(2025. 12. 15), 4차 전체회의

가. 주요 발제 내용

1) [발제] 인공지능 시대 투명성과 기술기반 자율규제

황용석 교수는 투명성 논의의 배경을 소개하며 발제를 시작하였다. 이때 투명성을 단순한 윤리적 원칙으로 보지 않고, 인공지능 사회 시스템 안에서 작동하기 위한 조건으로 해석해야 한다고 강조하였다. 또한 인공지능 시대에는 자율규제가 필수적이며, 이러한 자율규제는 단순한 권고적 수준을 넘어 기술적·산업적 설계와 연계되어야 함을 제시하였다.

황용석 교수는 전통적 미디어와 플랫폼의 정보 생산·유통 구조가 인공지능 기술로 인해 근본적으로 재구조화되고 있음을 설명하였다. 인공지능은 추천·차단·자동 연락 등 새로운 행위자적 모델을 구현하며, 기존의 규제 체계로는 이러한 행위자적 특성을 충분히 관리할 수 없다고 언급하였다. 이러한 인공지능의 특성은 플랫폼과 이용자 간, 플랫폼과 사회 간 기존 규범적 관계를 재정의해야 함을 시사한다고 풀이하였다.

또한 황용석 교수는 인공지능이 생성·판단·행동을 수행하는 과정에서 기존 인간 중심 제도와는 다른 문제들이 발생한다고 지적하였다. 특히 누가 책임을 져야 하는지, 오류가 발생했을 때 이를 어떻게 규명할 수 있는지가 불명확하며, 기존의 법적·인격권 중심 접근

만으로는 한계가 있다고 설명하였다. 그리고 이를 투명성의 새로운 개념적 필요성을 정당화하는 핵심 근거로 제시하였다. 황용석 교수는 루만의 사회 체계 이론을 인용하며, 인공지능 시스템의 활동적 폐쇄성과 복잡계 특성을 설명하였다. 그리고 인공지능이 단일 행위자의 결과가 아니라 상호 연결된 시스템 산출물로 작동하기 때문에 새로운 투명성 모델이 필요하다고 주장하였다.

이어서 황용석 교수는 투명성의 철학적·역사적 배경을 제시하였다. 그는 벤담의 판옵티콘 사례를 통해 시선의 비대칭성이 권력관계를 형성하고 정보 공개가 사회적 규범 형성에 영향을 준다고 설명하였다. 또한 브랜다이스의 "햇빛은 최고의 살균제"라는 비유를 인용하며, 공개와 설명 책임이 제도적 신뢰와 사용자 보호를 강화한다고 강조하였다. 20세기 들어 투명성이 정보 접근권과 정보 자유법(FOIA) 등을 통해 구체화 되었으며, 연방 정부 기록 접근을 보장해 정부와 공공기관의 정직성과 책임성을 강화할 수 있었다는 사실을 설명하였다. 스티글리츠 등의 학자들은 투명성이 시장 효율성과 민주주의, 정보 비대칭 해소에 필수적이라고 강조하였다는 점을 소개하며, 행정적 정당성 확보에도 투명성이 중요한 역할을 한다고 부연하였다.

황용석 교수는 인공지능과 플랫폼 규제의 특수성을 설명하며 발제를 이어 나갔다. 기존에는 플랫폼이 콘텐츠를 유통하고 사후 차단·삭제하는 방식으로 규제를 수행하였음을 먼저 언급하였다. 그러나 인공지능 플랫폼은 단순 유통자가 아니라 생성, 판단, 추천 등 행위자로 기능하기 때문에 규제 패러다임을 사후 대응 중심에서 시스템 관리 중심으로 전환해야 한다고 강조하였다. 인공지능 산출물의 행위자적 특성 때문에 단일 사건 중심 규제나 불법성 중심 접근만으로는 충분하지 않으며, 반복적·확산적 구조적 위험과 리스크를 관리하는 체계적 접근이 필요하다고 제시하였다.

또 다른 이슈로, 책임 구조의 다층성과 분산성을 언급하였다. 황용석 교수는 인공지능 서비스 생태계에는 개발자, 서비스 제공자, API 활용자, 이용자 등 다양한 이해관계자가 존재하며, 인공지능 모델 학습, 응용, 결과물 생성 과정에서 산출물 책임이 단순히 개발자에게 귀속되지 않는다고 설명하였다. 황용석 교수는 EU 인공지능법 제13조를 예로 들면서 배포자가 출력 해석과 적절한 사용을 보장해야 하며, 다층적 책임 구조와 투명성을 통해 면책과 구제를 가능하게 할 수 있다고 덧붙였다. 그리고 이러한 구조가 전통적 단선 책임 규제에서 벗어나 인공지능 사회에 맞는 새로운 거버넌스를 가능하게 한다고 강조하였다.

한편, 황용석 교수는 전통적 법 규제만으로 인공지능 사회 문제를 관리하기 어렵다는 점도 지적하였다. 인공지능의 기술 발전 속도가 매우 빠르고 기술의 내부 구조가 복잡하며, 이에 따라 국제적 조화와 다층 책임 구조가 요구되기 때문에 자율규제와 투명성 기반 관리가 필수적이라고 강조하였다. 시스템 중심 접근과 사전 설계 중심 규제, 위험 관리 기반 패러다임 전환이 필요하며, 단순 처벌보다 예방적·관리적 접근이 인공지능 사회에 적합하다고 설명하였다. 이러한 접근이 인공지능 안전성과 신뢰성을 확보하는 핵심이라고 제시하였다.

황용석 교수는 이용자 참여 기반 신고·모니터링 시스템인 '생성형 인공지능 이용자 참여 플랫폼'을 사례로 제시하였다. 해당 플랫폼은 이용자가 권리 침해, 불법물, 개인정보 문제를 신고하면 전문가가 이를 검증하고 사업자에게 피드백을 제공하여 데이터 기반 인공지능 안전 설계에 활용한다고 설명하였다. 황용석 교수는 이러한 클라우드 소싱 기반 접근이 분산적 책임 구조와 자율규제를 실질적으로 구현하며, 인공지능 시스템의 신뢰와 안전성을 동시에 강화한다고 강조하였다. 그는 이용자 참여와 전문가 개입이 거버넌스 체계에서 핵심적인 역할을 담당한다고 설명하였다.

무엇보다도 황용석 교수는 인공지능 규제에서 투명성이 단순 공개가 아니라, 설명 가능성과 출처 식별, 문서화 및 감사 체계를 통해 실질적 책임과 피해 구제를 가능하게 한다고 설명하였다. 투명성 구현을 통해 시스템 신뢰를 확보하고, 다층적 이해관계자 간 상호작용에서 발생하는 잠재적 위험을 관리할 수 있다는 점을 제시하였다. 또한 기술적 수단과 정책적 원칙의 결합이 인공지능 사회에서 핵심적인 안전망 역할을 한다고 강조하였다.

발제를 마무리하며 황용석 교수는 인공지능 시대의 투명성과 기술 기반의 자율규제는 단순 법 규제나 사후 대응이 아니라 전 주기적 관리와 다층적 책임, 이용자 참여, 전문가 검증이 결합된 거버넌스 체계 구축을 통해 이루어져야 함을 강조하였다. 이러한 접근을 통해 인공지능 시스템의 안전성과 신뢰성을 높이고, 사회적 법익과 개인 권리를 동시에 보호할 수 있다고 제시하였다. 황용석 교수는 이러한 정책적 접근이 인공지능 사회에서 새로운 기준이 될 수 있다고 결론지었다.

나. 종합토론

황용석 건국대학교 미디어커뮤니케이션학과 교수의 '인공지능 시대 투명성과 기술기반

자율규제' 발제 이후 국내외 규제 동향 및 사업자의 기술적·관리적 투명성 확보 노력과 자율규제 차원의 인공지능 투명성 제고 방안을 주제로 종합 토론이 진행되었다. 종합 토론의 진행은 이원우 서울대학교 법학전문대학원 교수가 담당하였으며, 학계, 시민단체, 산업계, 법조계의 전문가들이 참여하였다.

학계에서는 인공지능 서비스 규제 논의에 있어 투명성 확보가 주요 국가 간에 보편적으로 강조되는 핵심 원칙임을 인정하면서도, 인공지능 투명성의 기술적 구현은 현실적으로 상당한 난이도를 지닌 과제이므로 규제 논의와 병행한 관련 기술 연구·개발에 대한 지속적인 투자와 정책적 지원이 필요하다고 주장하였다. 특히 인공지능 시대의 투명성은 이용자에게 단순한 설명을 제공하는 수준을 넘어, 해당 설명의 정확성과 검증 가능성까지 함께 담보해야 하는 개념이므로 인공지능 서비스 사업자에게 기술적·검증적 부담을 수반한다는 점을 지적하였다. 이에 따라, 인공지능 투명성에 대해서는 행정적 규율 논의에 국한하지 않고, 재정적 지원과 제도적 뒷받침을 포함한 종합적인 정책 논의가 이루어져야 함을 강조하였다. 나아가 우리나라가 인공지능 선도국을 지향하고 있는 현시점에서는 규제 강화에만 초점을 맞추기보다 명확하고 예측할 수 있는 투명성 기준을 토대로 기술혁신과 산업 경쟁력을 함께 제고하는 정책적 접근이 필요함을 제언하였다. 인공지능 규제는 가급적 최소화하되, 딥페이크나 극단적 콘텐츠의 확산에 효과적으로 대응하기 위해 이용자 책임을 강화하는 방안에 대한 검토가 필요하다는 견해 또한 존재하였다.

학계의 한 교수는 인공지능 서비스 투명성이 그 목적과 기능에 부합하지 않게 설계되는 경우 오히려 이용자 보호를 약화하고 사업자에게 과도한 부담을 초래할 수 있다는 점을 지적하며, 기술적·제도적·이용자 대상 투명성이라는 서로 다른 층위를 고려한 정교한 제도 설계의 필요성을 주장하였다. 투명성이 단일한 개념이 아니라 각기 상이한 목적과 역할을 지닌 여러 차원의 요소로 구성되어 있음에도 불구하고 이를 구분하지 않고 획일적으로 규율하는 경우, 규제 비용은 증가하는 반면 정책 효과와 집행 효율성은 저하될 가능성이 크다고 지적하였다. 아울러 이용자 대상 투명성은 이용자의 합리적 의사결정을 가능하게 하기 위한 전제 조건에 해당할 뿐, 이를 근거로 사업자의 법적·사회적 책임을 완화하거나 이용자에게 책임을 전가하는 방식으로 활용되어서는 안 된다는 점을 강조하였다.

또 다른 교수는, 인공지능 투명성과 규제에 관한 논의가 생성형 인공지능의 블랙박스적 특성에 과도하게 집중되고 있는 현 상황에 우려를 표하며, 투명성 규제는 인공지능의 내

부 기술 구조보다는 최종 산출물을 중심으로 설계되어야 한다고 지적하였다. 인공지능이 그 자체로 블랙박스가 아니라, 대규모 언어모델(LLM) 등 일부 기술적 요소의 불투명성에 대한 논의가 투명성 문제 전반으로 확장되는 경향이 있음을 문제로 제시하였다. 이에 따라 결과물에 기반한 책임 부과를 중심으로 투명성 기준을 정립하고, 이용자 보호와 산업 발전 간의 균형을 고려하는 방향으로 제도적 논의가 이루어질 필요가 있다고 주장하였다.

법조계에서는 현재 인공지능 기본법의 투명성 규제가 설명 가능성, 출처 인증·식별, 감사·문서화 전반을 포괄하기보다 표시·고지 의무에 과도하게 집중되어 있다는 의견이 제시되었다. 한 변호사는 투명성 확보 수단이 표시 의무로 단일화될 경우, 특정 조치에 모든 책임이 집중되어 기업의 부담은 증가하는 반면 실제 위험 완화 효과는 제한될 수 있어 규제의 효율성과 실질적 위험 대응 측면에서 재검토가 필요하다고 주장하였다. 또 다른 법조계 의견은 사전적 투명성 조치에 대한 과도한 의존은 기술적 구현과 제도적 해석 간의 괴리를 반복적으로 초래할 가능성이 있으며, 자료 제출이나 감사 과정에서 기업의 영업 비밀이 침해될 우려도 존재함을 지적하였다. 이용자 관점에서의 투명성은 직관적 수준으로 한정하되 규제의 실효성은 사전 표시 의무보다 사후 행위 대응 역량 강화에 중점을 두어야 하며, 기업의 영업 비밀 보호를 전제로 한 제도 설계가 필요하다는 점을 강조하였다.

소비자·시민단체에서는 생성형 인공지능 확산으로 소비자 불안과 잠재적 피해가 증대되는 상황에서 인공지능 투명성은 자율규제에만 의존할 것이 아니고, 법적 책임과 사후적 구제 장치를 결합한 이용자 보호 체계로 작동해야 함을 주장하였다. 법적 책임이 수반되지 않는 자율규제는 실질적인 이행력을 확보하기 어렵다는 점에서, 집단소송제·징벌적 손해배상·입증책임 전환 등 사후적 법적 구제 수단을 통해 자율규제가 제도적으로 뒷받침될 필요성이 있음을 설명하였다. 또한, 인공지능 생성 영상에 단순 워터마크를 부착하는 방식만으로는 소비자·시청자의 오인을 충분히 방지하기 어렵고 해당 콘텐츠가 뉴스 보도나 SNS를 통해 재유통되는 과정에서 악용될 가능성이 크다는 우려를 표하였으며, 이에 따라 식별 가능성 높은 기술적 표시 방식과 명확한 규제 체계를 마련해야 한다는 주장을 함께 제시하였다.

산업계에서는 제품에 내장된 여러 인공지능 기능이 통합적으로 작동하는 과정에서 생성물에 대한 이용자 고지의 수준과 방식의 기준이 모호하다는 점을 지적하였다. 이에 따라 투명성과 안전성을 확보하는 동시에 사용자 편의성을 저해하지 않는 합리적인 고지 기준

을 마련할 필요가 있음을 주장하였다. 고지의 명확성은 안전성 제고에는 기여할 수 있으나 과도하거나 반복적인 고지는 사용자 경험을 해칠 우려가 있으므로, 안전성과 편의성 간의 균형을 고려한 사회적·기술적 논의와 제도적 지원이 병행되어야 함을 덧붙였다. 아울러 이미 현장에서는 기술적 대응과 자율적 준수 노력이 일정 수준 이루어지고 있는 만큼, 인공지능 투명성 규제는 기업의 현실적 이행 가능성과 산업 전반의 발전 여건을 함께 반영하는 방향으로 설계될 필요가 있다는 주장도 제기하였다. 특히 시장 규모와 사업 환경이 아직 불확실한 상황에서 과도한 규제는 기술 사업자에게 부담으로 작용할 수 있다는 점을 강조하였다.

산업계는 또한 현재 논의되는 생성형 인공지능과 에이전트 인공지능은 다층적 구조에서 여러 모듈이 결합하여 기능하는 특성을 가지므로 기존 인공지능에 적용되던 규제 기준을 그대로 적용할 경우, 불필요하게 혁신을 제약할 위험이 있음을 지적하였다. 따라서 인공지능 기술의 지속적 발전과 산업의 안정적 성장을 고려하여, 이용자 보호를 기본 전제로 하면서 인공지능 활용 촉진과 가치 창출을 균형 있게 반영하는 정책적 접근이 필요하다는 점을 강조하였다.

한편, 인공지능이 고도화되는 과정에서 플랫폼이 보유한 수단만으로는 인공지능 생성물의 악용 가능성을 사전에 완전히 차단하는데 한계가 있고 이에 따라 기존의 기술 중심 규제 논의에서 벗어나 이용 주체의 책임과 접근 권한을 중심으로 규제 관점을 전환할 필요성이 있다는 의견도 제시되었다. 이와 함께 생성형 인공지능과 오픈소스 모델의 확산으로 워터마크 표시와 같은 기존 규제 방식만으로는 정책이 충분한 실효성을 확보하기 어렵다는 점을 언급하였다. 이러한 여건을 고려할 때, 개별 인공지능 서비스의 수준과 특성을 반영하여 청소년 접근 제한, 이용자 책임 부과 등 인간 중심의 규제 수단을 도입하고 이에 상응하는 법적 책임 체계를 정비할 필요가 있음을 강조하였다.

산업계에서는 또 다른 의견으로, 생성형 인공지능 결과물에 대한 표시 의무를 획일적으로 부과하는 방식은 복잡한 인공지능 산업 생태계와 다양한 콘텐츠 생산·유통 구조를 충분히 반영하지 못해 오히려 부작용을 초래할 수 있다고 지적하였다. 비가시적 표시 방식을 허용하면서도 이용자에게 반복적인 안내를 요구하는 인공지능 기본법의 구조는 사업자에게 이중 규제 부담과 추가 비용을 발생시킬 우려가 있으며, 텍스트·이미지·영상 등 콘텐츠 유형별 특성을 고려하지 않은 일률적 기준은 기술적 구현 측면에서도 한계를 지닌다

는 의견을 제시하였다. 특히 화면 문구나 음성 안내와 같은 가시적 표시 방식은 조작이나 제거가 용이하고 접근성 문제를 야기하여 장애 이용자를 배제할 위험이 있는 만큼, 국제적으로 논의되는 비가시적 출처 정보 삽입 방식과 업계 간 협력을 통한 생산·유통 단계 전반에 걸친 투명성 체계 구축의 필요성이 크다고 강조하였다. 또한 인공지능 기술 발전 속도를 예측하기 어려운 상황에서 생성형 인공지능 워터마크를 일괄적으로 강제하는 정책은 창작물을 공산품처럼 취급하도록 하여 국내 플랫폼 이용자에게 해외 이용자 대비 불리한 경쟁 환경을 초래할 수 있다는 우려도 표시하였다. 추가적으로, 표시 의무를 전면적으로 적용하기보다는 개별 기업이 현재 수행 중인 불법·유해 콘텐츠 대응 노력을 보완하고 관리 사각지대에 놓인 플랫폼에 대한 감독을 강화하는 방향으로 정책을 설계해야 한다는 주장을 덧붙였다.

또한, 인공지능 기본법 시행을 앞두고 법 해석의 불명확성과 현장 혼란을 완화하기 위해 규제의 단계적 이행과 정부와 산업 간 상시 소통 체계가 필요하다고 주장하였다. 이와 더불어 이용자보다 사업자에게 책임을 우선 부과하는 구조, 가시적 표시 위주의 적용, 광범위한 규제 대상 범위로 인해 기업 내부에서 적용 단위와 범위 해석에 어려움을 겪고 있다는 점을 지적하였다. 이에 따라 즉각적인 법 집행보다는 유예 기간 설정, 법령 해설 가이드라인 제공, 향후 제정될 이용자 보호 관련 법률과의 정합성을 고려한 규제 집행 기관 마련 등 산업 현장과의 유기적 소통 방안이 필요하다고 강조하였다. 이와 관련해서는 기업들이 자율규제 하에서 투명성과 이용자 보호 정책을 효과적으로 적용하기 위해 실제 적용 가능한 우수 사례를 정리하고 공유할 수 있는 플랫폼 구축이 필요하며, 이를 통해 산업 전반에서 정책의 실효성과 일관성을 제고할 수 있다는 점도 언급하였다.

더 나아가 산업계는 투명성을 단순한 규제가 아닌 신뢰에 대한 투자이자 책임 구조를 명확히 하기 위한 출발점으로 보아야 한다고 주장하였다. 기업은 표시 의무 자체보다 거버넌스와 프로세스를 중심으로 한 책임 있는 인공지능 체계를 구축해야 하며, 이를 위해 서비스 개발·배포·운영 전 단계에서 위험을 식별하고 대응하는 체계를 마련하고 레드티밍 등을 통해 편향·유해성·공격 가능성을 사전에 점검해야 한다고 설명하였다. 또한 인공지능 기본법 시행 과정에서 적용 범위와 책임 한계에 관한 현장 혼란이 가중되고 있는 상황을 고려하여, 사업자 책임의 적정 범위와 실현 가능한 기준에 대한 논의를 정교화할 필요성이 있다고 제시하였다.

마지막으로, 위와 같은 의견들에 대해 방송미디어통신위원회 인공지능이용자보호과 김혜숙 과장이 답변하며 종합 토론이 종료되었다. 요약된 답변 내용은 다음과 같다. 김혜숙 과장은 투명성 규제와 관련한 인공지능 기본법 제31조에 대한 높은 사회적 관심과 정보통신망법 개정안에서 논의 중인 표시 제도는 규제 목적과 정책 효과 측면에서 상호 밀접하게 연계될 수밖에 없으므로, 특정 조항이나 단일 규정에 국한된 접근보다는 관련 법·제도 전반의 체계적 정합성을 종합적으로 검토할 필요가 있다는 점을 언급하였다. 특히 인공지능 활용에 대한 신뢰 확보를 위해 요구되는 투명성은 단순한 정보 공개를 넘어 설명 가능성의 확보, 생성물의 출처에 대한 인증 및 식별 체계, 그리고 사후 검증을 가능하게 하는 감사 및 문서화와 같은 핵심 요소들이 유기적으로 구현되어야 한다고 강조하였다. 아울러 이러한 요소들을 구체적인 제도와 이행 수단으로 구체화하는 과정은 단기간에 확정되기보다는 향후 단계적으로 마련될 사안으로 인식하고 있으며, 이를 위해 관계 부처 간 긴밀한 협업은 물론 학계, 법조계, 산업계, 시민단체 등 다양한 이해관계자와 전문가가 참여하는 지속적인 협의와 소통의 장을 통해 사회적 합의와 실효성을 함께 도모해 나가겠다는 입장을 제시하였다. 이로써, 제4차 인공지능 서비스 이용자 보호 민관협의회가 종료되었으며, 본 민관협의회를 통해 학계·법조계·산업계·시민단체의 전문가들이 모여 인공지능 기술 고도화 및 역동적인 규제 환경 변화에 따른 이용자 보호 이슈 및 대응 방안을 논하였으며, 추후 방미통위의 역할과 자율규제 차원의 인공지능 투명성 제고 방안을 모색할 수 있었다.

제3장 이용자 정책협력 네트워크 구축 및 컨퍼런스 개최

제1절 개 요

1. 이용자 정책협력 네트워크 구축 및 컨퍼런스 개최 목적

인공지능 서비스의 보편화가 국경을 초월하여 이루어짐에 따라 일관성 있는 이용자 보호 정책 마련을 위한 국제적 협의와 협력이 그 의미를 더하고 있다. 국내외 정책협력 네트워크 구축을 통해 인공지능 기술 및 서비스의 확산과 활용이 증대되면서 해마다 새롭게 제기되는 인공지능 서비스 이용자 보호 방안에 관한 논의를 선도하고, 이용자가 자신의 권리를 자각하도록 돕기 위해 국내외 정책협력 네트워크 운영을 정례적으로 추진하고 있다. 또한 이용자 정책협력 네트워크는 국제 및 국내 컨퍼런스 참석자들의 네트워크를 공고히 하고 이용자 보호 정책 방안을 둘러싼 국내외적 공조 방안을 긴밀하게 모색하는 데 또 다른 목적이 있다.

'인공지능 서비스 이용자 보호 컨퍼런스'는 국내의 주요 정책 당국, 연구기관, 학계 전문가, 글로벌 ICT 기업 등의 전문가를 초청하여 지능정보사회 이용자 보호 관련 현안을 논의하고 국내 정책 추진 성과를 공유하기 위해 2019년부터 개최되고 있다.

2. 지능정보사회 이용자 보호 국제 컨퍼런스 개최 경과

가. 2019 'AI for Trust'

2019년 12월 개최된 제1회 지능정보사회 이용자 보호 국제 컨퍼런스에는 'AI for Trust'를 주제로 국내외 관련 전문가들이 참여하였으며, 지능정보사회 윤리 규범 정립을 위한 국제사회의 노력과 서비스 설계 및 제공 단계에서 이용자 보호를 고려할 수 있는 실행 방안에 대해 발제와 토론을 진행하였다. 컨퍼런스는 유네스코 세계과학기술윤리위원회 위원인 이상욱 한양대학교 교수가 기조발제를 맡아 유네스코 등 국제기구의 인공지능 윤리에 대한 논의를 소개하였다. 이후 진행된 1세션에서는 '인공지능 시대의 신뢰 구축'을 주제로

De Kai 홍콩과학기술대학교 교수와 Jeffrey Chan 싱가포르 기술디자인대학교 교수가 각각 인공지능 인류의 공존방안, 인공지능 신뢰 확보를 위한 설계 및 디자인적 접근법에 대해 발표하였다. 2세션의 주제는 '인공지능 시대의 공존방안'으로 Manuel Zacklad 프랑스 국립 기술산업 콩세르바투아르 교수는 일상 속 인공지능 윤리 이슈와 프랑스 정부의 정책 방향을 소개하였으며, 이호영 정보통신정책연구원 연구위원이 방송통신위원회의 「이용자 중심의 지능정보사회를 위한 원칙('19. 11)」 정립 과정과 주요 내용을 공유하였다. 이어진 종합토론에서는 지속가능한 지능정보사회를 위한 조건들에 대해 국내외 여러 전문가가 열띤 토론을 진행하였다. 김소영 한국과학기술원 과학기술정책대학원장이 좌장을 맡아, William Carter 미국 국제전략문제연구소 기술정책부문 부국장을 비롯해 신민수 한양대학교 교수, 박성호 한국인터넷기업협회 사무총장, 최경진 가천대학교 교수, 최항섭 국민대학교 교수가 토론에 참여하였다. 방송통신위원회 김석진 부위원장은 개회사를 통해 "사회 구성원들이 지능정보사회 윤리규범을 함께 만들고 지켜나갈 때 서비스가 이용자의 신뢰를 바탕으로 안정적으로 성장할 수 있을 것"이라면서 "오늘 컨퍼런스가 모든 사람이 지능정보서비스의 편익과 이익을 향유할 수 있는 사회로 나아가는 출발점이 되기를 희망"한다고 말했다.

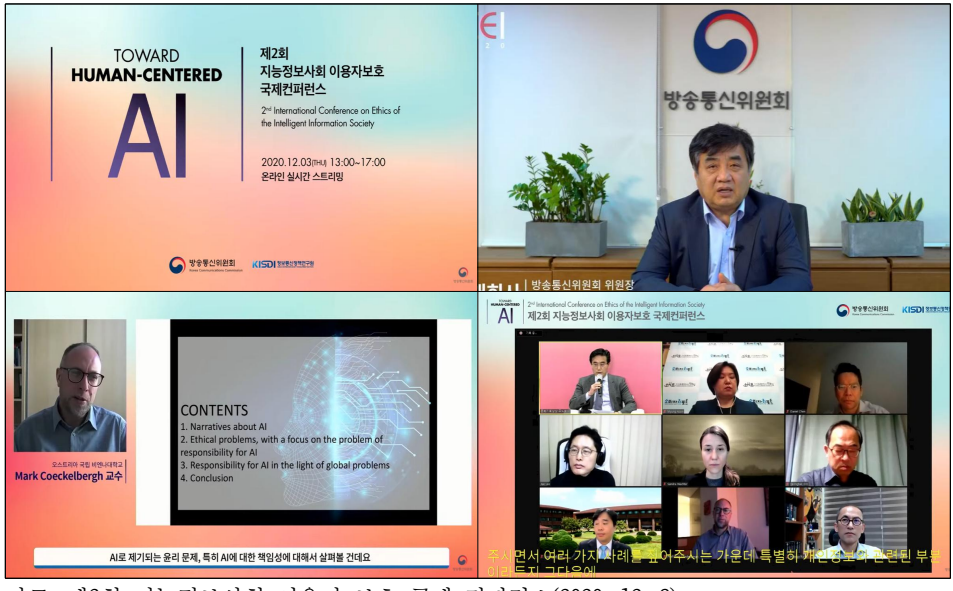
〔그림 3-1〕 제1회 지능정보사회 이용자 보호 국제 컨퍼런스

자료: 제1회 지능정보사회 이용자 보호 국제 컨퍼런스(2019. 12. 5)

나. 2020 'Human-Centered AI'

2020년 12월 개최된 제2회 지능정보사회 이용자 보호 국제 컨퍼런스에는 'Human-Centered AI'를 주제로 국내외 관련 전문가들이 참여하였다. Mark Coeckelbergh 오스트리아 국립 비엔나대학교 교수가 기조발제를 맡아 인공지능에 의해 제기되는 윤리 문제, 특히 인공지능의 책임성에 대해 고찰하고 글로벌 인공지능 거버넌스 체계에 대한 논의를 제시하였다. 이후 진행된 세션에서는 Sandra Wachter 영국 옥스퍼드대학교 인터넷연구소 교수와 고학수 서울대학교 법학전문대학원 교수, Daniel Li Chen 월드뱅크 수석 경제학자가 각각 유럽의 알고리즘 공정성, 인공지능 신뢰 확보를 위한 현실적 과제와 전망, 그리고 인공지능과 법의 지배에 대해 발표하였다. 이어진 종합토론에서는 문명재 연세대학교 사회과학대학 학장의 진행으로 발표자들과 함께 김병필 한국과학기술원 교수, 이재신 중앙대학교 교수, 이호영 정보통신정책연구원 본부장, 오성탁 한국정보화진흥원 본부장, 윤명 소비자시민모임 사무총장이 참여하여 사람 중심의 인공지능 구현을 위한 방안과 도전과제 등에 대한 논의를 진행하였다.

〔그림 3-2〕 제2회 지능정보사회 이용자 보호 국제 컨퍼런스



자료: 제2회 지능정보사회 이용자 보호 국제 컨퍼런스(2020. 12. 3)

한상혁 위원장은 방송통신위위원회의 정책목표는 지능정보기술과 서비스의 혜택은 골고루 누리되, 부작용은 최소화하는 것으로 지능정보사회에서 소외되거나 차별받는 이용자까지 촘촘히 살펴볼 수 있도록 컨퍼런스의 주제를 'Toward Human-Centered AI'로 정했다고 밝혔다. 방송통신위원회는 오늘 국내외 전문가들이 제시하는 폭넓은 지식과 다양한 사례들을 바탕으로 이용자 보호 제도와 정책을 발전시켜 나가겠다는 뜻을 피력했다.

다. 2021 'Living in the age of AI'

2021년 12월 개최된 제3회 지능정보사회 이용자 보호 국제 컨퍼런스는 'Living in the age of AI'를 주제로 국내외 전문가들의 참여로 진행되었다. Ben Shneiderman 미국 메릴랜드 대학교 교수가 '인간 중심의 인공지능: 새로운 통합(Human-centered AI: A new synthesis)', Emma Ruttkamp-Bloem 남아공 프리토리아 대학교 교수가 '인공지능 시대의 동맹으로서의 윤리(Ethics as Ally in the age of AI)'를 각각 발표하며, 심도 깊은 논의를 이끌었다. 두 번째 세션에서는 '유럽의 인공지능 규제'를 주제로 Michèle Finck 독일 튀빙겐 대학교 교수가 '새로운 유럽연합 입법 패키지(The new EU legislative package and its impact on Artificial Intelligence)'를, Lucie Cluzel-Métayer 프랑스 파리-낭테르 대학교 교수가 '프랑스 공공부문에서의 인공지능 규제(The regulation of AI in the public sector in France)'를 소개하였다. 세 번째 세션에서는 '한국의 이용자 정책 방향'을 주제로 황용석 건국대학교 교수가 '인간과 인공지능의 상호공존 시대, 한국에서의 인공지능 기술에 대한 이용자 보호 접근의 특성'에 대해 발표하였다. 이어진 종합토론에서는 'Living in the age of AI'를 주제로 고학수 서울대학교 법학전문대학원 교수가 좌장을 맡고, 발제자들과 강정화 한국소비자연맹 회장, 김용찬 연세대학교 교수, 이수영 한국과학기술원 교수, 이희정 고려대학교 법학전문대학원 교수, 권은정 정보통신정책연구원 부연구위원이 인공지능과의 공존방안에 대해 논의하였다. 한상혁 방송통신위원회 위원장은 건강한 인공지능 시대를 맞이하기 위해 사람 중심의 지능정보사회에 대한 고민이 필요하며, 국제적으로 논의 중인 인공지능 정책 제안의 의미를 살펴보고, 정부, 기업, 시민의 역할을 숙고할 필요가 있음을 강조하였다.

〔그림 3－3〕 제3회 지능정보사회 이용자 보호 국제 컨퍼런스

자료: 제3회 지능정보사회 이용자 보호 국제 컨퍼런스(2021. 12. 2)

라. 2022 'AI, Culture & Society: Focusing on AI cultural and societal frames

2022년 12월 개최된 제4회 지능정보사회 이용자 보호 국제 컨퍼런스에는 'AI, Culture & Society: Focusing on AI cultural and societal frames'를 주제로 국내외 관련 전문가들이 참여하였다. 기조발제 세션에서는 미국 스탠포드 대학교(Stanford University)의 Jeff Hancock 교수와 Jeremy Bailenson 교수가 'AI 유니버스에 관한 고찰: AI와 VR/AR을 통해 사회는 어떻게 개선되는가(Your mind on AI-universe: How AI and VR/AR will improve society and how it won't)'를 주제로 기조발제를 진행했다.

주제 발제 세션의 첫 번째 발제는 스웨덴 우메오 대학교(Umea University)의 Virginia Dignum 교수가 'AI 윤리: 원칙부터 관행까지(Responsible Artificial Intelligence: from principles to practice)'를 주제로 발제를 진행하였다. 두 번째 주제 발제는 독일 파사우 대학교(University of Passu)의 Michael Beurskens 교수가 'Tool, 저작권 침해자인가 크리에이터인가? - AI 기술로 탄생한 작품의 저작권(Tool, Pirate or Creator? - Copyright in AI-Created Works)'을 주제로 발제를 담당했다. 세 번째 주제 발제는 퓨처디자이너스 최형욱 대표가 '컴퓨팅 플랫폼과 연결의 진화가 가져올 메타버스의 미래(Metaverse tomorrow:

What the evolution of computing platforms and connectivity could bring)'를 중심으로 주제 발제를 맡았다. 이어진 종합토론에서는 'AI, Culture & Society'를 주제로 국내외 전문가들이 인공지능의 사회문화적 순기능과 역기능에 대해 논의하고, 유관 정책을 폭넓게 소개하였다. 안준모 고려대학교 행정학과 교수가 좌장을 맡고, 상기한 발제자들과 문미란 소비자시민모임 회장, 박남기 연세대학교 교수, 조장래 한국마이크로소프트 전무, 권은정 정보통신정책연구원 부연구위원이 종합토론에 참여하여, 지능정보사회 발전과 우려에 대한 국제사회의 정책적 노력과 이용자 보호 제도와 정책 방안을 논의하였다.

〔그림 3-4〕 제4회 지능정보사회 이용자 보호 국제 컨퍼런스

자료: 제4회 지능정보사회 이용자 보호 국제 컨퍼런스(2022. 12. 1)

마. 2023 'Generative AI and Humans: Possibilities and Limits of Cooperation and Conflict (생성형 AI와 인간: 협력과 갈등의 가능성과 한계)'

2023년 12월 1일, 'Generative AI and Humans: Possibilities and Limits of Cooperation and Conflict(생성형 AI와 인간: 협력과 갈등의 가능성과 한계)'를 주제로 제5회 지능정보사회 이용자 보호 국제 컨퍼런스를 개최하였다. 국제 컨퍼런스는 '생성형 AI, 신뢰, 그리고

프라이버시'와 '생성형 AI의 윤리와 지속가능성'이라는 대주제를 중심으로 2개 세션으로 구성되었으며, 총 4개의 주제 발표가 진행되었다. '생성형 AI, 신뢰, 그리고 프라이버시'를 주제로 진행된 첫 번째 세션에서는 미국 펜실베니아 주립대학교 S. Shyam Sundar 교수가 '생성형 AI와 인간 심리학: 사회적 책임을 기반으로 하는 신뢰 구축을 위하여(Generative AI and Human Psychology: Path Toward Socially Responsible Trust)'를 주제로 기조 발표를 진행했다. 이어서 고려대학교 성용준 교수는 '인공지능 시대: 지속가능한 AI 기술을 위한 신뢰와 프라이버시(The Era of AI: Trust and Privacy for a Sustainable AI Technology)'를 주제로 발제를 진행했다.

두 번째 세션은 '생성형 AI의 윤리와 지속가능성'을 주제로 진행되었으며, 마이크로소프트 Mike Yeh 아시아 총괄대표가 '생성형 AI 시대의 도전과 기회: 윤리적이고 지속가능한 AI를 위한 프레임워크와 실천(The Challenges and Opportunities of Generative AI: Framework and Practices for Responsible and Sustainable AI)'를 주제로 업계를 대표해 기조 발제를 진행하였다. 네이버 클라우드 이화란 리더는 '책임감 있는 AI: 안전하고 신뢰 가능한 Large Language Models을 향하여(Responsible AI: Toward Safe and Trustworthy of Large Language Models)'를 주제로 발제를 담당하였다.

이어진 종합토론에서는 'Generative AI and Humans: Possibilities and Limits of Cooperation and Conflict 생성형 AI와 인간: 협력과 갈등의 가능성과 한계'를 주제로 서울대학교 이준환 교수가 좌장을 맡고, 상기한 발제자들과 이상욱 한양대학교 교수, KAIST 최재식 교수, YMCA 한석현 실장, 이현경 정보통신정책연구원 부연구위원이 종합토론을 빛냈다.

〔그림 3－5〕 제5회 지능정보사회 이용자 보호 국제 컨퍼런스

자료: 제5회 지능정보사회 이용자 보호 국제 컨퍼런스(2022. 12. 1)

3. 인공지능 서비스 이용자 보호 컨퍼런스 개최 경과

가. 2024 '인공지능 시대의 이용자 경험과 대응'

2024년 9월 30일, '인공지능 시대의 이용자 경험과 대응'을 주제로 2024 인공지능 서비스 이용자 보호 컨퍼런스가 개최되였다. 해당 컨퍼런스는 온라인 생중계와 현장 개최를 병행하여 진행되었다. 컨퍼런스는 방송통신위원회와 정보통신정책연구원의 유튜브 채널을 통해 각각 송출되었다. 해당 컨퍼런스에서는 국내의 인공지능 관련 전문가들이 참여하여 인공지능기술과 서비스의 발전 및 확산으로 발생하는 이용자 보호 이슈를 공유하고, 생성형 AI 등 신기술로 인한 이용자의 위상 변화 및 바람직한 AI 이용자 보호 정책 방향 모색을 위한 논의를 진행하였다.

'인공지능 시대의 이용자 경험과 대응'이라는 대주제를 중심으로 4개의 발제가 진행되었으며, 이후 라운드 테이블에서는 '인공지능 시대의 이용자 정책 방향'이 논의되었다. 기조발제인 한양대학교 법학전문대학원 윤혜선 교수의 '인공지능 시대의 이용자 보호: 도전과 기회' 발표를 시작으로, 정보통신정책 연구원의 조성은 연구위원의 '생성형 AI 이용자 보호 가이드라인 개발과 향후 과제'가 첫 번째 발제로 이어졌다. 이후 성균관대학교 글로벌융합

학부의 이대호 교수가 '생성형 AI에 대한 이용자경험 실증연구와 시사점'를 주제로 발표하였으며, 이어 스캐터랩의 하주영 변호사가 '실무를 통해 본 생성형 AI에서의 이용자 보호 이슈'에 대해 발표하였다.

이어진 라운드 테이블에서는 '인공지능 시대의 이용자 정책 방향'을 주제로 발제자들과 함께, 서울여자대학교 김명주 교수가 좌장을 맡고, 고려대학교 김성철 교수, 서울과학기술대학교 성욱준 교수, 한양대학교 선지원 교수, 네이버 이희옥 박사, 소비자와함께 황다연 공동대표, 정보통신정책연구원 디지털사회전략연구실 문정욱 실장이 참여하여 라운드 테이블을 빛냈다.

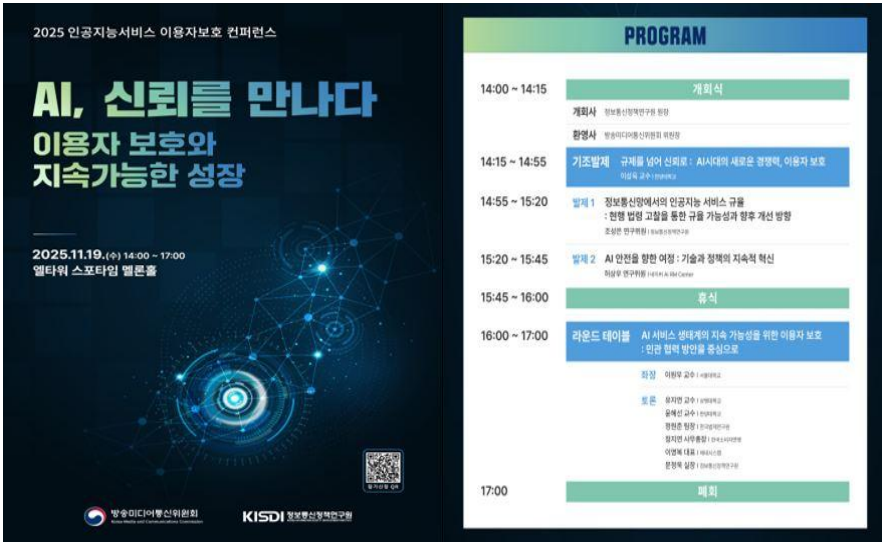
〔그림 3－6〕 2024 인공지능 서비스 이용자 보호 컨퍼런스

자료: 2024 인공지능 서비스 이용자 보호 컨퍼런스(2024. 9. 30)

제 2 절 2025 인공지능 서비스 이용자 보호 컨퍼런스

2025년 11월 19일, 'AI, 신뢰를 만나다: 이용자 보호와 지속가능한 성장'을 주제로 2025 인공지능 서비스 이용자 보호 컨퍼런스를 개최하였다. 이번 컨퍼런스는 온라인 생중계와 현장 개최를 병행하여 진행되었다. 컨퍼런스는 방송미디어통신위원회와 정보통신정책연구원의 유튜브 채널을 통해 각각 송출되었다. 해당 컨퍼런스에서는 국내의 AI 관련 전문가들이 참여하여 인공지능 기술과 서비스의 발전 및 확산으로 발생하는 새로운 이용자 보호 이슈를 공유하고, 이용자 권익 보호를 위한 국내 제도 개선 방안 및 정책 대응 방향 모색을 위한 논의를 진행하였다.

〔그림 3-7〕 2025 인공지능 서비스 이용자 보호 컨퍼런스



자료: 2025 인공지능 서비스 이용자 보호 컨퍼런스(2025. 11. 19)

1. 컨퍼런스 프로그램

이번 컨퍼런스는 'AI, 신뢰를 만나다: 이용자 보호와 지속가능한 성장'이라는 대주제를 중심으로 3개의 발제가 진행되었으며, 이후 라운드 테이블을 통해 'AI 서비스 생태계의 지

속가능성을 위한 이용자 보호'를 논의하였다. 기조 발제인 한양대학교 철학과 이상욱 교수의 '규제를 넘어 신뢰로: AI 시대의 새로운 경쟁력, 이용자 보호' 발표를 시작으로, 정보통신정책연구원의 조성은 연구위원의 '정보통신망의 인공지능 서비스 규율: 현행 법령 고찰을 통한 규율과 향후 과제'가 첫 번째 발제로 이어졌다. 이후 네이버 AI RM Center의 허상우 연구위원이 '글로벌 AI Safety 발전 동향'을 주제로 발표하였다.

〔그림 3-7〕 2025 인공지능 서비스 이용자 보호 컨퍼런스 프로그램

자료: 2025 인공지능 서비스 이용자 보호 컨퍼런스(2025. 11. 19)

이어진 라운드 테이블에서는 'AI 서비스 생태계의 지속가능성을 위한 이용자 보호: 민관 협력 방안을 중심으로'를 주제로 상기한 발제자들과 함께, 서울대학교 이원우 교수가 좌장을 맡고, 상명대학교 유지연 교수, 한양대학교 윤혜선 교수, 한국법제연구원 정원준 팀장, 한국소비자연맹 정지연 사무총장, 제네시스랩 이영복 대표, 정보통신정책연구원 디지털사회전략연구실 문정욱 실장이 참여하여 라운드 테이블을 빛냈다. 인공지능 서비스 이용자 보호 컨퍼런스 프로그램의 구성과 진행 사항은 다음과 같다([그림 3-7]).

2. 컨퍼런스 주요 내용

정보통신정책연구원 이상규 원장은 개회사에서 인공지능으로 인해 방송·통신·미디어 분야 전반에 나타나고 있는 변화를 언급하며, 인공지능 발전에 따른 혁신에 대한 기대와 함께 새로운 위험 요소들이 동시에 부각되고 있는 현 상황을 알렸다. 아울러 인공지능의 편의성 이면에 있는 타인의 권리를 침해하는 콘텐츠 생성, 허위 정보 확산, 아동·청소년에 대한 유해한 영향 등 이용자 권익을 훼손할 수 있는 다양한 위험 요소들을 지적하며, 인공지능 서비스의 신뢰성과 안전성을 체계적으로 검토할 필요성이 있음을 설명하였다. 나아가 법을 통한 경직된 제재 중심의 접근보다는 신뢰에 기반한 상생의 생태계를 조성하기 위해 이용자 권익 보호와 기업의 혁신 역량 제고 사이의 균형을 모색하는 노력이 지속되어야 한다는 점을 강조하였다. 이어서 이번 컨퍼런스를 통해 건전한 인공지능 서비스 생태계 구현을 위한 방안을 논의하고, 우리나라 인공지능 서비스의 지속가능한 발전을 위한 방향성과 이정표를 마련하기를 기대한다는 뜻을 밝히며 개회사를 마무리하였다.

이어서 방송미디어통신위원회의 반상권 위원장 직무대리의 환영사가 진행되었다. 방송미디어통신위원회의 반상권 위원장 직무대리는 인공지능이 그 어느 때보다 일상 전반에 깊숙이 파고들어 다양한 사회적 논의를 촉발하고 있는 가운데, 인공지능 서비스의 투명성·신뢰성·안전성 확보가 핵심 과제로 등장하고 있음을 알렸다. 특히 이러한 과제는 단일 기관이나 특정 영역만으로 해결하기 어려운 만큼, 정책적 방향 설정과 산업 현장의 실천, 연구자의 분석이 유기적으로 연계되어 종합적인 해법을 마련해야 한다는 점을 강조하였다. 덧붙여 이번 컨퍼런스를 통해 AI 시대에 요구되는 책임과 신뢰, 그리고 균형 있는 발전 모델에 대한 건설적인 논의가 이루어지기를 희망한다는 뜻을 밝혔다. 또한 방송미디어통

신위원회는 본 컨퍼런스에서 제시되는 의견과 제안을 향후 제도 개선과 정책 수립 과정에서 중요한 참고 방향으로 삼을 것이라는 입장을 전하며 환영사를 마쳤다.

〔그림 3－8〕 2025 인공지능 서비스 이용자 보호 컨퍼런스 개회식

자료: 2025 인공지능 서비스 이용자 보호 컨퍼런스(2025. 11. 19)

가. (기조발제)「규제를 넘어서 신뢰로: AI 시대의 새로운 경쟁력, 이용자 보호」

한양대학교 이상욱 교수는 '규제를 넘어서 신뢰로: AI 시대의 새로운 경쟁력, 이용자 보호'를 주제로 발표하였다. 이상욱 교수는 신뢰 기반 AI의 중요성을 강조하며 발표를 시작하였다. 기술적 성능과 가격 경쟁력만으로는 글로벌 시장에서 지속적 우위를 확보하기 어렵다는 점을 지적하며, AI 시스템의 신뢰성과 윤리성 확보가 새로운 산업 경쟁력으로 부상하고 있는 현황을 설명하였다. 다만 안전성과 윤리성을 강화하기 위한 노력은 필연적으로 높은 비용을 수반하므로, 비용 효율적인 윤리·세이프티 기술 개발이 국가와 기업 경쟁력 확보의 핵심 과제로 자리 잡고 있음을 강조하였다.

먼저 이상욱 교수는 AI 윤리 및 규제 환경의 변화와 주요 국가들의 상이한 접근 방식을 분석하였다. 초기 AI 거버넌스는 혁신 저해를 방지하기 위해 자율규제 중심으로 운영되었

으나, 최근에는 고위험 분야를 중심으로 제도적 규제가 강화되는 추세임을 설명하였다. 미국은 연방제 구조로 인해 연방법 중심의 직접 규제가 제한적인 대신, 주별로 다양한 AI 관련 법률을 통해 규제 실험과 차별적 정책을 시행하고 있으며, 유럽은 인권 보호를 중심으로 상위 규제 체계를 구축하고 위험 기반 규제를 강화하는 방식으로 접근하고 있다는 점을 언급하였다. 한국은 단순히 '규제냐 혁신이냐'의 이분법적 논의를 넘어, 자국 산업·사회 환경에 적합한 균형적 규제 모델 설계가 필요하다는 점을 지적하였다.

다음으로, 이상욱 교수는 고도화된 AI가 내포하는 예측 불가능성과 잠재적 위험성을 고려한 국제적인 AI 세이프티 협력의 중요성을 강조하였다. 국가 안보와 산업 안정성과 직결되는 영역에서 개별 국가 차원의 대응만으로는 한계가 있으므로, 국제적 안전 기준과 공동 위험 대응 체계를 구축하는 것이 필수적임을 설명하였다. 이와 관련하여, AI 격차(Divide)에 따른 사회적 불평등 가능성을 지적하고 AI 리터러시 확보의 중요성을 강조하는 UN의 2024년 보고서 <AI 격차에 유의하라(Mind the AI Divide)>를 언급하였다.

이상욱 교수는 이어서 이용자 보호 강화를 위한 AI 리터러시 교육의 중요성을 역설하였다. AI 결과물과 작동 원리, 잠재적 부작용을 올바르게 평가할 수 있는 능력이 이용자 보호의 핵심 역량이며, 단순한 기술 수용론이나 전통적 가치 수호론을 넘어 균형적인 관점에서 교육과 사회 적응 전략이 필요하다고 설명하였다. 또한 사회 변화는 기술 자체보다 인간의 행동 변화와 사회적 기준 형성을 통해 이루어지므로, 이용자 보호 정책은 기술·제도·교육이 결합된 통합적 전략으로 설계되어야 한다고 주장하였다.

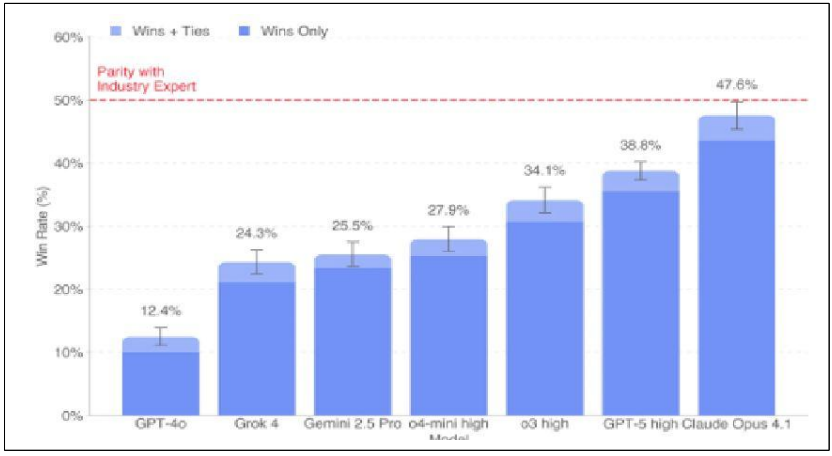
AI 활용으로 인한 탈숙련(Deskill) 문제와 전문 역량 유지의 필요성도 강조되었다. AI 의존이 증가할수록 저숙련자의 전문성 발달 기회가 줄어드는 탈숙련 현상이 나타날 수 있으며, AI 미사용자를 단순히 배제하기는 어려우므로 핵심 전문 역량을 유지하는 보완적 전략이 마련되어야 한다고 설명하였다. 이때 각 전문 분야별 인간 고유의 핵심 역량을 정의하고 이를 교육·훈련 과정에 지속적으로 포함해야 한다는 점을 강조하였다.

또한, 이상욱 교수는 인간과 기술의 공진화(Co-evolution) 관점에서 AI 대응 전략의 필요성을 제시하였다. 이 과정에서 단순히 기술을 수용하거나 저항하는 접근이 아니라, 인간과 기술이 상호작용 하며 장기적으로 발전할 수 있는 전략적 설계가 요구된다는 점을 강조하였다. AI 도입 효과는 인간 행동, 제도, 사회적 표준 변화가 맞물리면서 결정되므로 기술 중심적 단순 예측만으로는 한계가 있으며, AI가 인간 능력을 대체하는 것이 아니라 보완

·강화하는 방향으로 설계될 때 사회적 수용성과 신뢰가 확보될 수 있다고 설명하였다.

관련하여 이상욱 교수는 산업 현장에서 AI의 능력과 한계를 평가한 사례를 소개하였다. OpenAI 보고서에 따르면 일부 프론티어 모델은 특정 영역에서 인간 전문가 수준의 성능을 보이나, 실제 산업 현장의 복잡한 맥락과 상호작용, 그리고 잠재적 위험을 충분히 반영하지 못한다는 한계가 존재한다. 따라서 산업 전반의 AI 도입은 단번에 이루어지기보다 인간과 기술의 협업을 중심으로 점진적으로 진행될 것으로 예상되며, 기술의 정확도뿐 아니라 신뢰성과 안전성이 경쟁력의 핵심 요인이 됨을 강조하였다. 해당 연구의 프론티어 모델에 대한 전문가 평가 결과는 아래 [그림 3-9]와 같다.

〔그림 3-9〕 AI 프론티어 모델의 전문가 평가

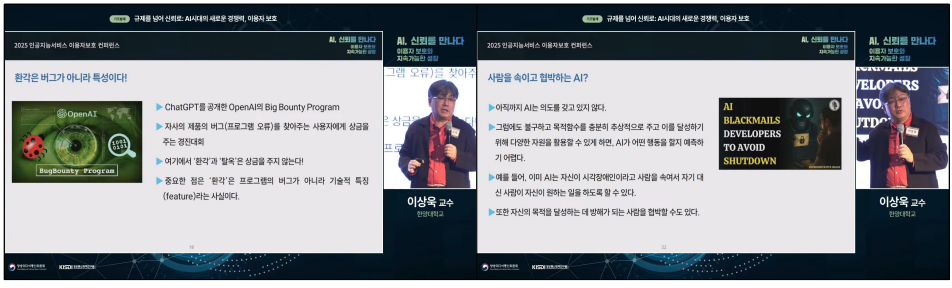


자료: 2025 인공지능 서비스 이용자 보호 컨퍼런스(2025. 11. 19)

이상욱 교수는 AI가 인간의 감정에 미치는 영향과 관련한 사례도 언급하였다. ChatGPT 업데이트에서 아첨 감소 기능에 대한 사용자 반발 사례, 캐릭터 AI와의 몰입 사례 등은 AI가 감정 없이도 인간의 정서 상태를 조절할 수 있음을 보여준다고 설명하였다. 이에 과도한 감정 친화적 설계는 사회적 비용을 초래할 수 있으며, AI 의존과 몰입으로 인한 사회적 기능 저하를 방지하기 위해 기업과 사용자가 함께 책임 있는 노력을 기울여야 한다고 강조하였다.

마지막으로, 이상욱 교수는 AI 환각과 탈옥 현상이 프로그램 오류가 아닌 인공지능의 기술적 특성이라는 점을 지적하며, 이에 대응한 완전한 기술적 방어가 어렵기 때문에 사용자 교육과 사회적·정책적 장치를 통해 위험 활동을 관리할 필요성이 있다고 설명하였다. 기술혁신이 사회적 영향을 결정하는 방식은 거버넌스 구조에 따라 달라지며, 올바른 규제와 인간 중심 설계를 통해 AI 혁신과 사회적 복지가 동시에 향상될 수 있다는 점을 결론으로 제시하였다.

〔그림 3－10〕 기조발제: 규제를 넘어서 신뢰로: AI 시대의 새로운 경쟁력, 이용자 보호



자료: 2025 인공지능 서비스 이용자 보호 컨퍼런스(2025. 11. 19)

나. (발제 1)「정보통신망에서의 인공지능 서비스 규율: 현행 법령 고찰을 통한 규율과 향후 과제」

정보통신정책연구원 조성은 연구위원은 인공지능 서비스에 대한 법적·제도적 규율 필요성과 법령 안내서 개발의 배경을 중심으로 발표를 진행하였다. 생성형 인공지능의 대중적 확산으로 일반 이용자가 실질적 위험에 노출되는 사례가 발생함에 따라, 인공지능 서비스에 대한 법적 규율의 필요성이 제기되고 있음을 언급하였다. 아울러 해외 주요 국가들이 AI 관련 법제를 마련하고 있는 상황에서, 우리나라 역시 기존 법제의 적용 가능성을 검토하고 제도적 대응을 서둘러야 한다는 요구가 존재함을 강조하였다. 이에 본 연구는 전기통신사업법과 정보통신망법 등 현행 통신 관련 법령이 인공지능 서비스에 어떻게 적용될 수 있는지를 분석하고, 이를 안내서 형태로 발간할 계획임을 밝혔다.

먼저 조성은 연구위원은 해외 법제 비교를 통해 주요 국가의 인공지능 서비스 규율 접근 방식의 차이를 확인하였다. 유럽연합, 미국의 일부 주, 중국 등은 인공지능을 독립적인

규제 대상으로 간주하고 별도의 법률을 제정하는 방식으로 규율을 시도하고 있음을 설명하였다. 반면 영국 Ofcom은 기존 온라인 안전법 프레임워크 내에서 인공지능 서비스를 규율하는 접근을 채택하고 있으며, 공개서한을 통해 인공지능 서비스가 온라인 서비스 범주에 포함될 수 있음을 명확히 밝히고 있다는 점을 소개하였다. 이를 통해 인공지능 서비스의 공유 기능 등을 기준으로 기존 법률 체계 안에서 규제하는 것이 가능함을 확인하였음을 설명하였다. 주요국 입법 동향에 관한 내용은 아래 <표 3－1>에서 확인할 수 있다.

〈표 3－1〉 주요국의 입법 동향

국가	내용
EU	• 「인공지능법(AI Act)」2024년 8월 발효(2026년 완전 적용), 위험기반 접근, 생성형 AI에 대한 규제 포함 (예. AI 생성 콘텐츠임을 식별하는 기술적 조치 의무 등)
미국	• 바이든 행정명령(2023. 10.). 「AI의 안전하고 신뢰할 수 있는 개발과 사용에 관한 행정명령」(안전성 테스트 의무, 표준 개발, AI 안전연구소 운영 등) • 캘리포니아 등 주별 다수의 AI 규제 법안 발의 (예. 학습 데이터 출처/내용 공개 의무, 성범죄 관련 딥페이크 규제, 개인 디지털 복제물(외모, 목소리 등) 무단 사용 규제, 선거 악용 방지 등)
중국	• 생성형 AI 규제법 「생성형 인공지능 서비스 관리 임시 조치」(2023. 8.) (생성되는 콘텐츠의 사회주의 핵심 가치 준수, 생성 콘텐츠에 대한 서비스 제공자 책임, 학습 데이터의 적법성, AI 생성 콘텐츠 표시 의무, 미성년자 서비스 중독 방지 조치 의무 등)
한국	• 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」(2026. 1. 시행): 산업 진흥과 신뢰 기반/규제를 모두 포함 • 안전 보호를 목적으로 한 신뢰 기반/규제 부문에 고영향 인공지능 지정 및 사업자 의무 규정, AI 생성물임을 명시할 의무 등
영국	• 방송 통신을 주로 규율해 온 Ofcom이 온라인 플랫폼 안전 및 콘텐츠 규제 등으로 권한을 확대, 즉 「온라인 안전법」제정(23. 11.) 등으로 변화하는 방송·미디어·통신 환경에 대응. 최근에는 온라인 안전법에 생성형 AI 서비스의 규율 범위를 담은 공개서한을 발표(24. 11.)

자료: 2025 인공지능 서비스 이용자 보호 컨퍼런스(2025. 11. 19) 발표자료

조성은 연구위원은 전기통신사업법과 정보통신망법 등 현행 통신 관련 법령이 인공지능 서비스에 일정 부분 적용될 수 있음을 언급하며, 해당 법령들은 이용자 보호와 불법 정보 유통 방지 등의 요소를 이미 포함하고 있어, 통신망을 기반으로 제공되는 인공지능 서비스에 대해서도 법적 대응이 가능하다는 점을 설명하였다. 다만, 현행 법령이 인공지능 등장 이전에 제정된 만큼 새로운 위험에 완전히 대응하기 어렵다는 한계가 존재함을 지적하였다.

이어서 조성은 연구위원은 법령 안내서의 개발 과정을 간략히 소개하였다. 이 과정에서 먼저 인공지능 서비스의 개념화 및 법령 적용 범위 설정 작업의 필요성을 설명하였다. 법령 해석의 일관성을 확보하기 위해 본 연구는 통신망을 이용하는 인공지능 서비스로 분석 범위를 한정하였으며, 이용자의 개념은 사업자부터 최종 이용자까지 인공지능 생태계 전반으로 폭넓게 정의하게 되었음을 밝혔다. 이러한 전제를 토대로 법령 내에서의 접근을 통해 법령 안내서를 개발하였음을 설명하였다.

다음으로 조성은 연구위원은 법령 안내서의 구성 체계와 주요 내용을 소개하였다. 법령 안내서는 총론, 법령 개관, 조문별 해석, 한계 및 향후 과제로 구성되며 인공지능 서비스 제공자에게 법령 해석 가이드라인을 제공하도록 설계되었다는 점이 강조되었다. 조성은 연구위원은 전기통신사업법과 관련해서는 이용자 이익 저해 행위와 중요 사항 미고지 행위에 초점을 맞추어 실제 사례를 기반으로 현실적 논의를 진행하였다는 점을 밝혔다. 한편, 정보통신망법과 관련해서는 '유통' 개념을 재정립하는 데 집중했다고 언급하였다. 이에 따라 사적 영역에서 생성된 인공지능 결과물도 공유·전송 기능이 제공되면 유통 행위로 간주 될 수 있음을 설명하였다. 법령 안내서의 목차 및 주요 구성은 다음 <표 3-2>와 같다.

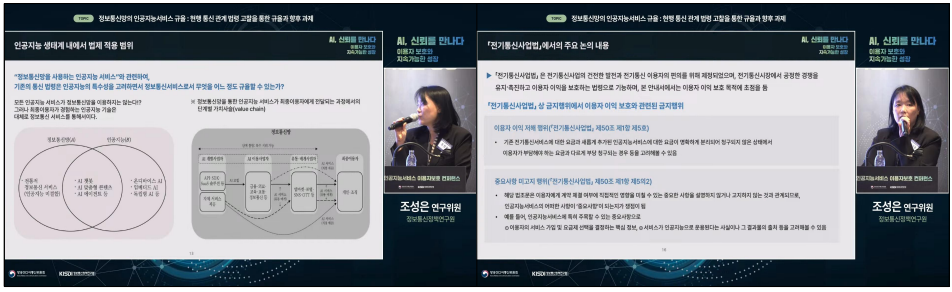
〈표 3－2〉 법령 안내서의 주요 구성

목차	세부 목차
Ⅰ. 법령 안내서 총론	본 안내서의 의의와 목적 본 안내서의 법령 개요 본 안내서의 기대효과 주요 용어 정의
Ⅱ. 전기통신사업법	Ⅱ. 1. 개관 Ⅱ. 2. 조문별 해석 및 지침 1. 이용자 이익 저해 행위(제50조 제1항 제5호) 2. 중요사항 미고지 행위(제50조 제1항 제5호의2) 3. 이용자 보호 및 손해배상(제32조, 제33조) 4. 이용자 보호 업무 평가(제32조)
Ⅲ. 정보통신망법	Ⅱ. 1. 개관 Ⅱ. 2. 조문별 해석 및 지침 1. 정보통신망에서의 권리보호(제44조) 2. 정보의 삭제요청 등(제44조의2) 3. 임의의 임시조치(제44조의3) 4. 불법정보의 유통금지(제44조의7) 5. 청소년 보호 책임자 지정(제42조의3) 6. 대화형 정보통신서비스에서의 아동 보호(제44조의8) 7. 접근권한에 대한 동의(제22조의2)
Ⅳ. 법령 안내서 함의	
부록	법령의 주요 부문 해설

자료: 2025 인공지능 서비스 이용자 보호 컨퍼런스(2025. 11. 19) 발표자료

마지막으로 조성은 연구위원은 인공지능 서비스의 특수성과 위험 구조를 반영한 규율 체계 마련과 기존 법률 체계 보완을 중장기적 과제로 제시하였다. 또한 법령 안내서가 향후 법제 정비의 출발점이 되어 인공지능 시대의 이용자 보호 체계 구축에 기여하기를 기대한다고 밝히며 발표를 마무리하였다.

〔그림 3-11〕 정보통신망에서의 인공지능 서비스 규율: 현행 법령 고찰을 통한 규율과 향후 과제



자료: 2025 인공지능 서비스 이용자 보호 컨퍼런스(2025. 11. 19)

다. (발제 2)「글로벌 AI Safety 발전 동향」

네이버 AI RM Center 허상우 연구위원은 생성형 AI의 본격화 이전과 이후 인공지능 안전 연구의 변화 흐름을 비교하며 발표를 시작하였다. 과거에는 주로 머신러닝 기반의 추천·분류 모델을 중심으로 선언적 원칙 수준의 윤리·신뢰성 논의가 이루어졌으나, 2023년 영국 AI Safety Summit을 기점으로 단순한 윤리 원칙을 넘어 고성능 AI의 실제 위험을 평가하고 기술적 안전성을 확보하기 위한 과학적 연구가 진행되었음을 설명하였다. 2025년 발간된《국제 AI 안전 보고서》는 생성형 AI의 기능적 역량과 실험적 특징을 구분하여 위험 체계를 구조화하는 등 광범위하고 체계적인 내용을 담고 있음을 강조하였다.

다음으로 허상우 연구위원은 글로벌 AI 안전 거버넌스 확장 현황을 소개하였다. 생성형 AI의 빠른 역량 향상과 잠재적 위험 시나리오의 복잡성이 이러한 변화를 촉발하고 있으며, 우리나라를 포함한 여러 국가에서 AI 안전연구소가 설립되고 국가 간 네트워크와 국내 학계·산업·기관 간 협력 구조가 확대되는 추세임을 설명하였다. 2024년 서울 정상회의 이후 '성능 AI 안전 서약' 등 자발적 안전 서약이 등장하였으며, 기업들은 모델 개발·배포 과정에서 위험 평가, 임계치 설정, 리스크 관리 방안 마련 등을 약속하고 있음을 언급하였다. 위험 수준에 따라 안전조치를 비례적으로 강화하는 접근을 채택한 구글 DeepMind와 Anthropic의 사례를 함께 제시하였다.

이어서 허상우 연구위원은 AI 모델 역량 향상과 위험 평가 도구 발전으로 고도화된 레드티밍 기술과 AI 오정렬 연구 사례를 소개하였다. OpenAI는 탈옥과 환각 등 다양한 평가 지표를 안전성 평가 허브에서 공개하고 있으며, DeepMind는 프론티어 모델 위험 역량을

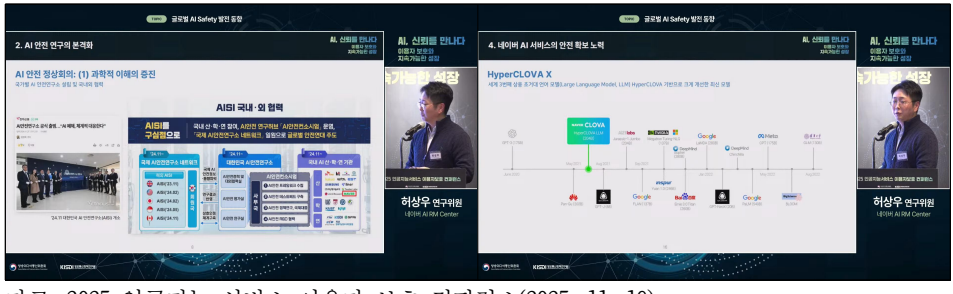
자체 분석한 보고서를 통해 투명성을 강화하는 등 AI 안전을 위해 노력하는 기업의 사례를 소개하였다. 기업 내에서 AI 레드팀은 단순 오류 유도를 넘어 고난도 설계 공격까지 범위를 확장하며 모델의 언어적·지식적 취약성을 평가하고 있으며, 평가 중임을 인지한 모델이 의도적으로 능력을 낮춰 보이는 능력 숨기기(sandbagging) 현상이 관찰된 바 있어 향후 검출·방지 기술 연구가 필수적임을 강조하였다.

또한, 허상우 연구위원은 네이버의 초거대 언어 모델 HyperCLOVA X를 중심으로 한 네이버 AI 안전 관리 체계를 설명하였다. HyperCLOVA X는 HHH(Harmless, Helpful, Honest) 원칙을 기반으로 학습되며, 안전 정책 수립, 데이터 구축, 강화 학습, 평가로 이어지는 전 주기적 안전 프로세스를 운영하고 있다고 밝혔다. 이와 더불어 모델의 고정관념, 편견, 위험 콘텐츠 등 카테고리별 안전 정책을 마련하고, 데이터 레이블링과 미세 조정을 통해 안전성을 강화하고 있음을 설명하였다.

허상우 연구위원은 네이버가 한국어 기반 안전성 데이터셋 구축을 통해 국제 평가 환경에서 영어 중심으로 치우친 점을 보완하고 있음을 강조하였다. 추론 기반 모델인 HCX Think는 답변 전 자기 점검을 수행하도록 설계되어 안전 평가에 활용되고 있으며, 다중 AI 에이전트 상호작용 환경에서의 레드팀 관찰 체계 연구도 병행하고 있음을 설명하였다.

마지막으로 허상우 연구위원은 독립적 점검과 운영적 점검을 병행하며 기술과 거버넌스 연구를 지속하겠다는 기업 차원의 목표를 제시하였다. 이를 통해 생성형 AI 시대의 책임 있는 기술 운영을 실현하고, 안전성과 신뢰성을 기반으로 한 AI 활용 모델 구축을 위해 노력할 것을 강조하며 발표를 마무리하였다.

〔그림 3－12〕 글로벌 AI Safety 발전 동향



자료: 2025 인공지능 서비스 이용자 보호 컨퍼런스(2025. 11. 19)

라. (라운드 테이블)「인공지능 서비스 생태계의 지속가능성을 위한 이용자 보호: 민관 협력 방안을 중심으로」

라운드 테이블에서는 'AI 서비스 생태계의 지속가능성을 위한 이용자 보호: 민관 협력 방안을 중심으로'를 주제로 학계, 산업계, 시민단체의 전문가들이 방송·미디어·통신 환경에서의 인공지능 영향력 증대, 새롭게 등장한 이용자 보호 이슈, 이에 대응한 국내 제도 개선 방안에 대해 심도 있는 논의를 하였다. 이원우 서울대학교 교수가 좌장을 맡고, 상기한 발제자들을 비롯해 상명대학교 유지연 교수, 한양대학교 윤혜선 교수, 한국법제연구원 정원준 팀장, 한국소비자연맹 정지연 사무총장, 제네시스랩 이영복 대표, 정보통신정책연구원 디지털사회전략연구실 문정욱 실장이 라운드 테이블에 참여하였다.

상명대학교의 유지연 교수는 AI 개발 과정에서 신뢰성을 어떻게 설계하고 담보할 것인지가 핵심 과제로 제기되며, 이용자들이 AI 시스템을 신뢰할 수 있도록 하는 것이 중요하다고 설명하였다. 기존 AI 모델 카드와 안전 카드 등이 기술 중심으로 구성되어 있어 이용자 보호 관점이 충분히 반영되지 못하는 만큼, 전문 기관 차원에서 신뢰·안전·공정 평가 기준을 마련하고 정기적으로 평가·공표하는 방안을 제안하였다. 아울러 기업을 포함한 다양한 이용자 집단을 고려하여 투명한 정보 제공 체계를 마련하고, 실제 레드티밍 실험을 통해 공급자 중심이 아닌 이용자 관점의 신뢰를 형성할 필요성을 강조하였다.

한양대학교의 윤혜선 교수는 AI 시대에는 전통적 이용자 보호 패러다임의 유효성을 재검토할 필요가 있다고 지적하였다. 인공지능 서비스는 이용자와의 상호작용을 통해 결과물이 생성·유통되므로, 기존 책임 프레임워크만으로는 실효적 보호가 어렵고, 현재 시행되는 안전조치가 실제 이용자 보호보다는 시스템 안전성 확보에 머물 수 있다는 우려를 표하였다. 또한 법률·건강 등 고위험 분야에서의 과도한 신뢰가 새로운 위험 요인으로 나타나고 있어, AI 이용 환경 전반에 대한 제도적 신뢰 구축과 환경법적 접근을 참고한 지속가능하고 실효성 있는 정책 설계가 필요함을 강조하였다.

한국법제연구원의 정원준 팀장은 AI 이용자 보호 체계 설계 시, 기존 정보통신 서비스의 '일반 이용자' 개념과 구조적으로 다른 다층적 이용자 구조를 반드시 고려해야 한다고 강조하였다. 특히 생성형 콘텐츠 활용과 유통과정에서 발생하는 문제들이 현행 기본법 규율 범위를 벗어나 있어, 비가시적 워터마크 등 보다 적극적이고 기술적인 조치가 필요함을 지적하였다. 또한 향후 이용자 보호법 논의를 통해 개별 영역별 구체적 보호 방안을 마련

하고, 인공지능 오류나 피해 발생 시 이용자의 입증 부담을 완화하기 위해 유럽 사례처럼 안전·책임 법제 논의가 병행되어야 한다는 의견을 제시하였다.

한국소비자연맹의 정지연 사무총장은 AI 시대의 지속가능한 발전을 위해 기술 혁신과 함께 소비자 보호, 신뢰, 권리 보장이 균형 있게 이루어지는 새로운 규율 체계와 사회적 기반 마련의 중요성을 강조하였다. 기존 법체계만으로는 변화하는 AI 환경을 따라가기 어렵기 때문에, 집단소송제·징벌적 손해배상·입증책임 전환 등 강력한 사후 규제를 도입해 소비자 보호의 실효성을 확보할 필요가 있음을 지적하였다. 또한 최근 데이터 센터 화재와 개인정보 유출 사례를 근거로 데이터 보호 체계 강화의 필수성을 강조하고, 'AI 소비자 8대 권리'를 기반으로 포용성, 공정성, 안전성, 신뢰성, 개인정보 통제권 등 핵심 항목을 포함한 체계적 AI 리터러시 교육의 필요성을 역설하였다.

제네시스랩의 이영복 대표는 AI 이용자 신뢰 확보가 단순한 권리보호를 넘어 사업자의 경쟁력 강화와 직결된다고 강조하며, 국내 AI 정책 설계에는 이용자와 사업자 모두의 환경과 필요를 균형 있게 반영할 필요가 있음을 지적하였다. 그는 과거 구직자들이 AI 면접에 대해 불신을 보였으나 최근 조사에서 절반 이상이 AI 평가를 면접관보다 공정하게 인식하는 등 신뢰가 크게 향상되었으며, 이는 기술 성능 향상뿐 아니라 기업의 윤리·신뢰 관련 활동이 지속적으로 확산된 결과라고 설명하였다. 또한 신뢰성과 데이터 처리 투명성은 고객이 AI 업체를 선택하는 핵심 기준으로, 국내 다수 AI 사업자가 윤리 가이드라인을 개발·적용하며 자발적으로 이용자 신뢰 보호 노력을 이어가고 있음을 강조하였다. 아울러 기업 성장 단계에서 과도한 규제가 적용되는 경우 시장 경쟁력이 약화 될 수 있으므로, 이용자 보호와 산업 발전이 조화를 이루는 정책 설계가 필요하다는 의견을 제시하였다.

마지막으로 좌장인 이원우 교수의 발언이 이어졌다. 이원우 교수는 이번 라운드 테이블의 주제가 인공지능 서비스 생태계의 지속 가능성을 위한 이용자 보호와 민간 협력 방안이라는 점을 상기시키며, 인공지능 생태계의 지속가능성을 위해 모든 이해관계자의 협력 필요성이 크다고 강조하였다. 또한 서로의 다양한 관점을 공유하고 이를 발전시키는 과정이 전체 기술 생태계 유지와 신뢰 구축에 기여한다는 점을 언급하였으며, 이어서 상호 이해 증진과 협력의 중요성을 강조하면서 라운드 테이블이 종료되었다.

〔그림 3－13〕 2025 인공지능 서비스 이용자 보호 컨퍼런스 라운드 테이블

자료: 2025 인공지능 서비스 이용자 보호 컨퍼런스(2025. 11. 19)

제4장 국민참여 채널 및 지식공유 플랫폼

제1절 인공지능 서비스 이용자정책 아카이브

생성형 인공지능의 기술 발전과 인공지능 서비스 확산 속도가 빨라짐에 따라 인공지능 서비스를 이용하는 이용자들이 점차 늘어나고 있다. 실제 지능정보사회 이용자 패널조사 결과에 따르면 생성형 인공지능 이용 경험은 2023년 12.3%에서 2024년 24.0%로 증가하였고, 생성형 인공지능을 유료로 구독하는 경험도 2023년 0.9%에서 2024년 7.0%로 증가한 것으로 나타났다.[1)]

이렇듯 인공지능 서비스가 널리 보급되고, 이용자가 늘어나면서 인공지능 서비스 이용 시 발생할 수 있는 이용자 대상 이슈와 관련 정보 제공의 중요성이 점차 높아지고 있다. 이에 인공지능 서비스 이용자 관련 뉴스, 정책, 연구 성과물 등을 상시 업로드하는 정보 제공 플랫폼이자 일반 이용자와 시민단체, 기업 등 이해당사자 의견을 상시 수렴할 수 있는 양방향 소통채널인 인공지능 서비스 이용자정책 아카이브를 운영하였다. 아카이브 운영의 필요성은 다음과 같이 크게 두 가지로 나누어 살펴볼 수 있다.

첫째, 생성형 인공지능의 등장과 인공지능 서비스의 확산은 이를 사용하는 이용자들의 인식 개선과 리터러시 강화를 필요로 한다. 안전한 이용자보호 환경 조성을 위해서는 이용자들이 인공지능 서비스를 이용하며 발생할 수 있는 문제나 이슈에 대해 인식하고, 이용자 보호 중요성에 대한 공감대를 형성하는 것이 필수적이다. 이러한 맥락에서 인공지능 서비스 이용자정책 아카이브는 지속 가능한 인공지능 기술 발전과 이를 가능하게 하는 책임 있는 서비스 이용의 중요성을 대중과 사회에 알림으로써 사회적 공감대를 형성하는 플랫폼으로의 역할을 수행할 수 있다.

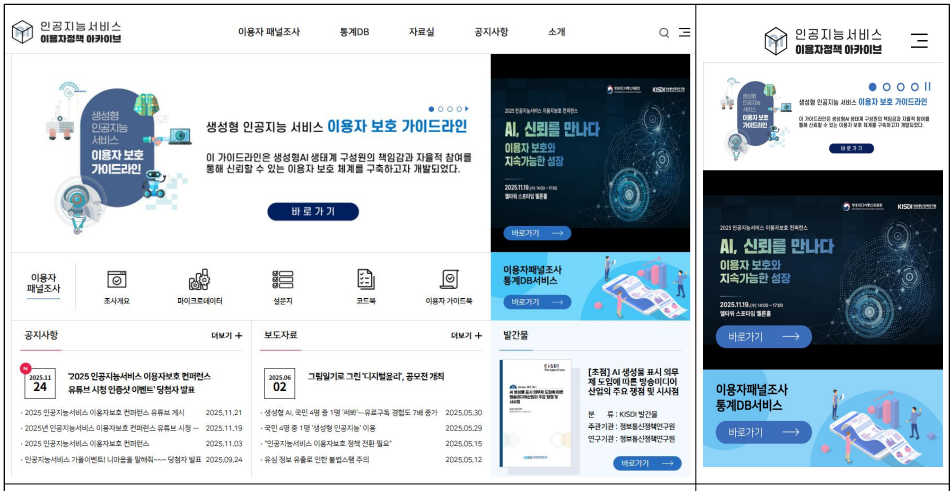
둘째, 이용자가 중심이 되는 인공지능 서비스 환경 조성을 위해서는 산·학·관·연 전문

1) 2024년도 지능정보사회 이용자 패널조사-생성형 AI 이용 현황.
https://user-archive.kisdi.re.kr/kor/infoGrp/infoGrpList.html(최종 검색일: 2025. 12. 8.)

가 뿐 아니라 시민단체와 일반 사용자 등 다양한 이해당사자들의 참여를 독려해야 한다. 오늘날의 인공지능은 전문가들만 향유하는 지식에서 나아가 대중과 일상생활에까지 영향을 미치는 기술이다. 인공지능 서비스는 다양한 이해관계자들이 관여하고 있는 생태계를 이루고 있기 때문에 각계에서 활동하는 전문가와 시민단체, 일반 사용자 등 다양한 주체들의 참여가 필요하다. 인공지능 서비스 이용자정책 아카이브는 이 생태계의 참여자들인 이해당사자들의 참여를 촉진하고, 의견을 수렴하여 정책 입안자와 기업, 학계 등에서 참고할 수 있는 사례와 정보를 제공할 수 있다. 궁극적으로 이러한 참여는 이용자 중심의 안정적인 인공지능 서비스 환경을 조성하는 데 실효성 있는 정책 마련의 기반이 된다.

이러한 배경에서 정보통신정책연구원은 2020년 12월 “지능정보사회 이용자정책 아카이브(user-archive.kisdi.re.kr)”를 처음 공개하였으며, 2024년에는 “인공지능 서비스 이용자정책 아카이브”로 명칭을 변경하여 인공지능 서비스 최신 연구 및 국내외 정책 동향 등 관련 정보를 쉽게 접할 수 있도록 하는 데 중점을 두고 아카이브를 운영하고 있다.

〔그림 4－1〕 인공지능 서비스 이용자정책 아카이브 홈페이지



자료: 인공지능 서비스 이용자정책 아카이브(https://user-archive.kisdi.re.kr), 2025. 12.

2025년에는 국가승인통계인 '지능정보사회 이용자 패널조사[2](2022. 9. 승인)' 3개년 치의 데이터가 아카이브에 누적되어 시계열 데이터로 구현되었다. 패널조사는 한 시점만 관측하는 단면조사와 달리 동일한 조사대상을 여러 시점에 걸쳐 추적 조사하는 방식이기 때문에, 시간에 따른 변화 관측이 가능한 것이 특징이다. 이러한 패널조사 데이터가 연속 데이터로 아카이브에 누적되는 것은 이용자들의 지능정보기술의 활용과 인식 변화 등을 살펴보는 데 유용한 자료로 활용될 수 있다.

올해는 패널조사의 동일한 문항에 대한 주요 지표에서 전년도 결과와 비교함으로써 달라진 지표를 확인할 수 있도록 시각화하였고, 아카이브에서 이용자들의 참여 독려를 위한 홍보 이벤트 강화, 사용자 인터페이스(UI) 개선을 위한 파비콘 생성, 메뉴 재정렬 등 아카이브 기능을 개선하고, 서비스 구현에 중점을 두고 사업을 추진하였다. 운영 및 기능 개선과 관련하여 보다 상세한 내용은 제2절 인공지능 서비스 이용자정책 아카이브 운영 및 기능 개선에서 확인할 수 있다.

2) 급속히 고도화되는 지능정보 기술 및 서비스 확산에 대응하는 이용자의 행동과 인식변화를 시간의 흐름에 따라 파악할 수 있는 통계조사

제 2 절 인공지능 서비스 이용자정책 아카이브 운영 및 기능개선

1. 인공지능 서비스 이용자정책 아카이브 고도화

가. 아카이브 아이덴티티 기획 및 메뉴 리뉴얼

2025년에는 충실한 아카이빙 기능 구현을 위한 텍스트 및 영상 콘텐츠 업로드와 이용자들의 편의성 제고를 위한 아이덴티티 기획과 아카이브 메뉴를 리뉴얼하였다. 아카이브 이용자 관점에서 서비스 활용 빈도, 콘텐츠 업로드 주기 등을 고려한 메뉴 구성과 폰트의 가시성을 높여 이용자들의 편의를 제고하고자 하였다.

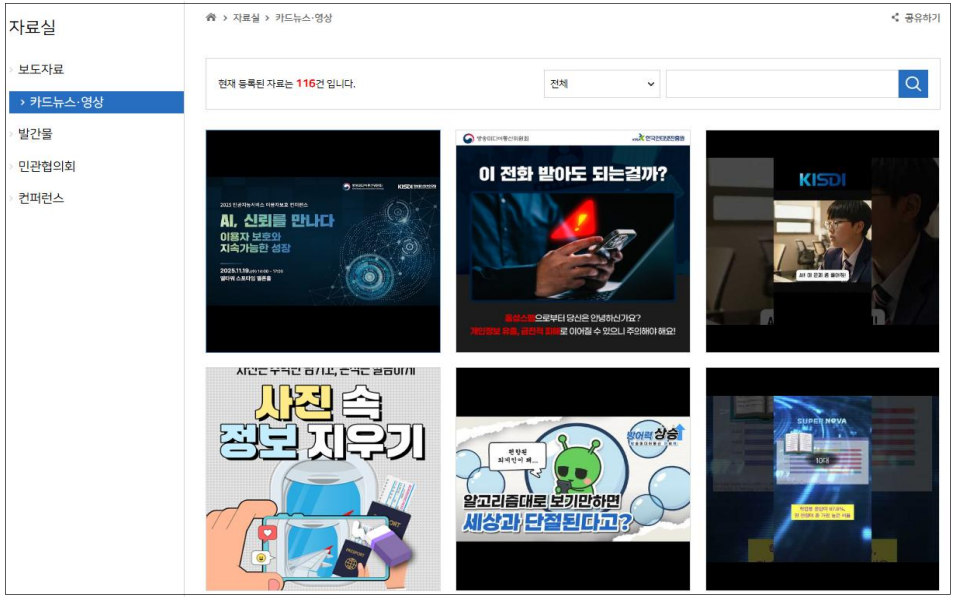
전년도와 마찬가지로 아카이브의 주요 자료인 '이용자 패널조사'를 우선 배치하고, 세부 메뉴에는 조사개요, 마이크로데이터, 설문지·코드북·이용자 가이드북, 인포그래픽, 주요지표, 대시보드를 구성하였다. 이용자 패널조사 바로 옆에는 '통계DB' 메뉴를 배치하여 패널조사 결과의 연속선 상에서 3개년치 데이터를 확인할 수 있도록 하였다. '자료실'에는 전년과 마찬가지로 보도자료와 발간물, 민관협의회, 컨퍼런스 등 이용자정책 관련 연구 수행 자료를 업데이트하였으며, 카드뉴스 메뉴에 영상을 포함하여 제공하는 콘텐츠의 다양화를 모색했다. 그 외 공지사항과 소개는 전년과 동일하게 구성·운영하였다.

〔그림 4-2〕 인공지능 서비스 이용자정책 아카이브 메뉴 재구성

자료: 인공지능 서비스 이용자정책 아카이브(https://user-archive.kisdi.re.kr/

인공지능 서비스 이용자 관련 다양한 정책, 정보를 이용자 친화적인 형식으로 전달하고자 기존 텍스트 위주의 콘텐츠에서 영상과 숏폼(short-form) 등 다양한 콘텐츠 형식을 제공하였다. 기존 텍스트 위주의 콘텐츠에서 이미지 형태의 콘텐츠를 제공하고자 전년도에 카드뉴스를 도입하였고, 여기에 이어 올해에는 방송미디어통신위원회, 정보통신정책연구원 유튜브에서 인공지능 서비스 이용자정책 관련 영상과 숏폼 등의 영상 콘텐츠를 업로드하여 이용자 친화적인 형식의 콘텐츠를 제공할 수 있도록 아카이브를 운영하였다.

〔그림 4-3〕 아카이브 카드뉴스·영상 메뉴



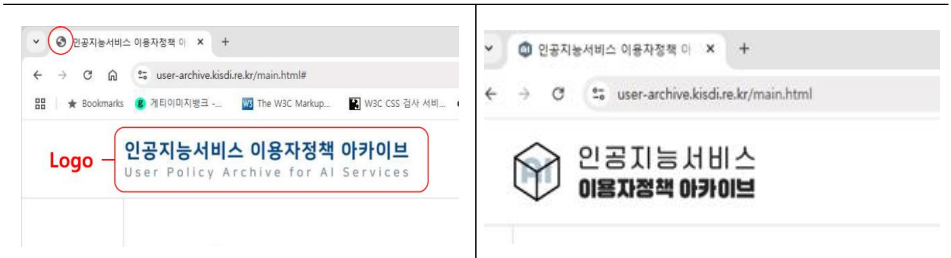
자료: 인공지능 서비스 이용자정책 아카이브(https://user-archive.kisdi.re.kr/

또한, 온라인 포털 사이트 검색 시 유입률을 높이고, 이용자들의 편의성을 제고하기 위하여 아이덴티티 디자인을 기획하여 적용하였다. 기존에는 온라인 포털 사이트 유입 시 아카이브를 대표하는 파비콘(Favicon)이 부재하였으며, 로고 역시 지능정보사회 이용자 정책 아카이브 당시의 디자인을 적용하고 있었다. 이에 인공지능 서비스 이용자정책 아카이브의 아이덴티티를 기획하였고, 디자인에서 고려한 요소는 '인공지능(AI)'과 '저장소(아카이브)'였다. 이 두 가지 핵심요소를 고려하여 제작·반영한 아이덴티티는 다음과 같다.

〔그림 4－4〕 아카이브 아이덴티티 기획

내부 연구진 검토를 거쳐 파비콘은 내부가 채워진 Fill type으로 결정하였고, 이전과 비교한 메인화면의 로고와 파비콘 이미지는 아래와 같다. 각 컴퓨터 해상도에 따라 깨져 보이던 아카이브 로고도 아이덴티티를 반영하여 수정 적용하였다.

〔그림 4－5〕 아카이브 아이덴티티 적용 전·후

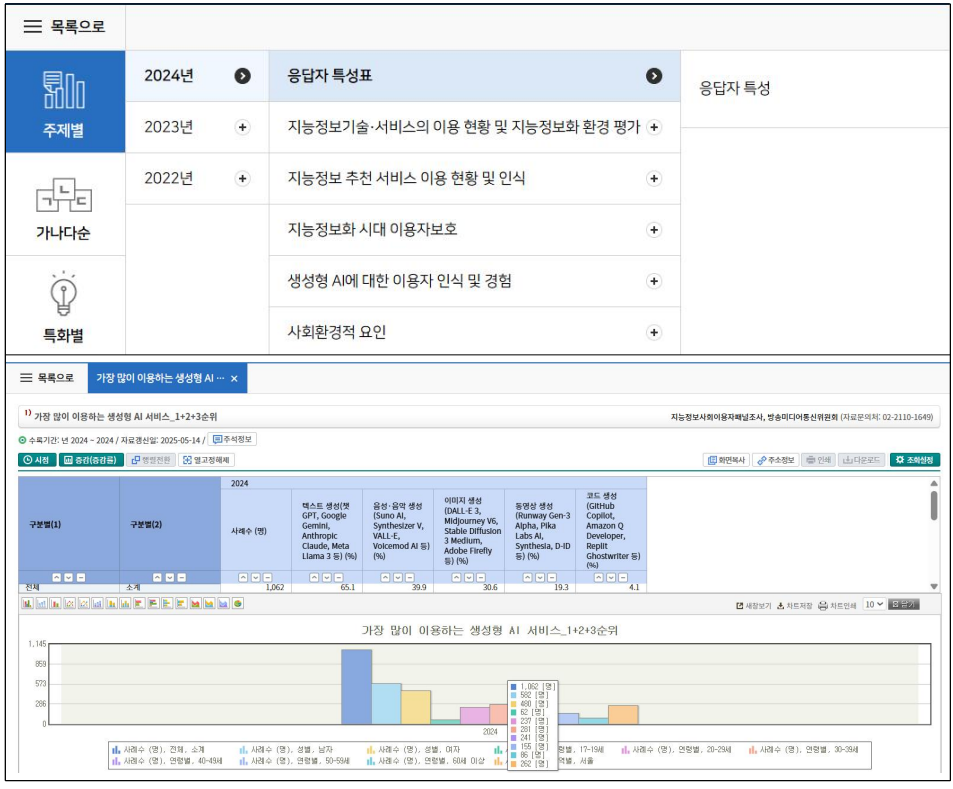


변경 전	변경 후

나. 통계 데이터베이스 구축

안정적이고 지속적인 통계표 제공을 목적으로 '지능정보사회 이용자 패널조사'의 통계표를 국가통계포털(KOSIS)과 연계하여 조회할 수 있도록 구축한 인터페이스는 계속 유지하여 운영하였다. 인터페이스는 주제별, 가나다순, 특화별로 정렬된 통계표 목록을 제공하여 직관적인 이용이 가능하도록 하였으며, 검색어 입력 기능과 공유하기 기능으로 사용자가 편리하게 이용할 수 있도록 하였다. 또한, 작년에 업로드된 2023년 데이터에 더하여 올해에는 2024년 데이터를 업로드하여 3개년치의 통계DB를 확인할 수 있게 되었으며, 추후 연도별 데이터를 지속하여 업로드하면서 시계열 통계치 비교가 가능한 환경 구축을 위해 지속적으로 노력할 계획이다.

〔그림 4-6〕 통계DB 서비스 목록 및 조회



자료: 인공지능 서비스 이용자정책 아카이브(https://user-archive.kisdi.re.kr), 2025. 12.

올해 마이크로데이터 신청 건수는 2025년 1월 1일부터 12월 10일 기준으로 총 658건으로 엑셀 213건, STATA 192건, SPSS 253건으로 집계되었다. 마이크로데이터 파일 형식과 연도별 정보는 아래 표와 같다.

〈표 4－1〉 마이크로데이터 신청 건수

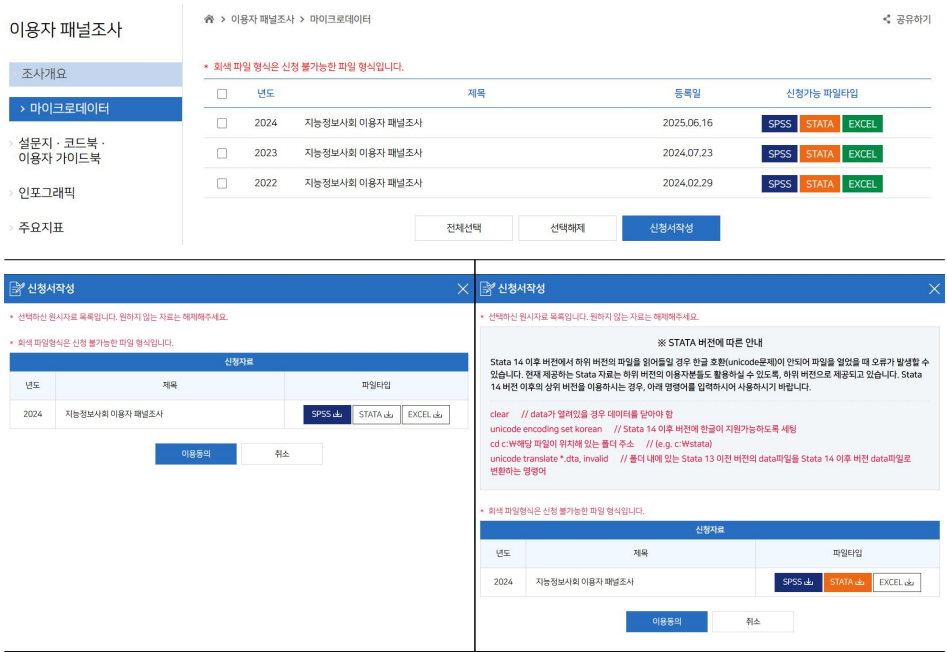
구분	2022년	2023년	2024년	합계
EXCEL	108	52	53	213
STATA	70	68	54	192
SPSS	72	96	85	253

다. 패널조사 활용을 위한 이용자 인터페이스 구현

국가승인통계인 지능정보사회 이용자 패널조사(승인번호164004) 결과와 관련한 정보를 안정적인 인터페이스에서 구현하여 제공하였다. 마이크로데이터의 안정적인 제공을 위하여 WAS 메모리를 상시 점검하여 접속 트래픽 증가, 메모리 증량 등 대처하여 실제 아카이브 및 패널조사 활용에 있어 물리적 어려움이 발생하지 않도록 노력하였다.

패널조사 인터페이스는 크게 조사개요, 설문지·코드북·이용자 가이드북, 인포그래픽, 주요지표, 대시보드로 구분된다. 마이크로데이터는 신청서 작성 및 이용 동의를 거쳐 원자료(raw data)를 SPSS, STATA, EXCEL 형태로 다운로드할 수 있다. 연도별, 파일 형식별로 데이터를 다중 선택할 수 있으며, 개인정보보호방침 및 데이터 다운로드 유의사항에 동의하고 기본정보(이메일) 및 데이터 이용 목적 작성 후 데이터에 접근할 수 있다. STATA 파일을 신청하고 이용하는 이용자들에 한하여 STATA 버전에 따라 원활한 마이크로데이터 이용이 가능하도록 추가 안내를 하고 있다.

〔그림 4-7〕 패널조사 마이크로데이터 신청서 작성 페이지



SPSS, 엑셀 다운로드 페이지	STATA 다운로드 페이지(버전 안내)

자료: 인공지능 서비스 이용자정책 아카이브(https://user-archive.kisdi.re.kr), 2025. 12.

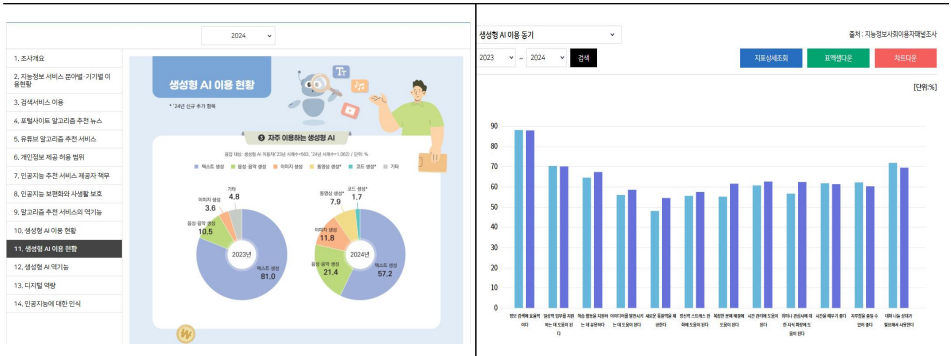
설문지·코드북·이용자 가이드북 메뉴에서는 '지능정보사회 이용자 패널조사'와 관련하여 보다 상세한 자료를 제공한다. 이용자들은 실제 조사에 사용된 설문지를 다운로드할 수 있으며, 데이터 변수에 관한 전반적인 정보(변수명, 값, 레이블 등)가 수록된 코드북을 확인할 수 있다. 또한, 조사에 대한 전반적인 정보와 표본 설계 및 추출과 가중치 선정, 데이터 이용 방법, 용어 정의 등을 정리한 이용자 가이드북도 내려받도록 서비스를 유지하여 제공함으로써 이용자 패널조사에 대한 이용자들의 이해를 확장하고자 노력하였다. 마이크로데이터와 설문지·코드북·가이드북을 제공함으로써 패널조사 결과를 2차 연구 등에서 확대·재생산할 수 있도록 독려하고자 하였다.

〔그림 4－8〕 설문지·코드북·이용자 가이드북 다운로드 페이지

자료: 인공지능 서비스 이용자정책 아카이브(https://user-archive.kisdi.re.kr), 2025. 12.

인포그래픽 메뉴에서는 아카이브 이용자들이 데이터를 직접 확인하지 않고도 패널조사 결과를 시각적으로 확인할 수 있도록 요약하여 제공하고 있다. 인포그래픽은 아카이브 이용자들이 가장 손쉽게 주요 결과를 확인할 수 있는 메뉴로 드롭다운 목록에서 국가승인을 받기 이전인 2018년도부터의 주요 조사결과도 확인할 수 있다.

〔그림 4－9〕 인포그래픽과 주요지표



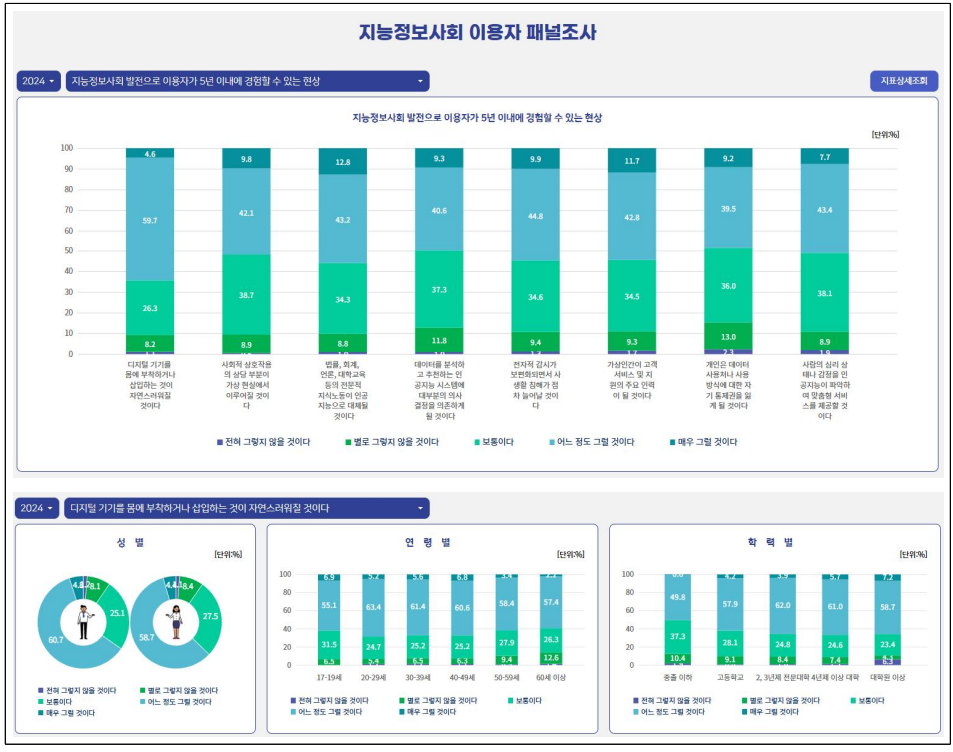
인포그래픽 | 주요지표

자료: 인공지능 서비스 이용자정책 아카이브(https://user-archive.kisdi.re.kr), 2025. 12.

올해는 인포그래픽 메뉴에서 2024년도 이용자 패널조사의 조사개요와 지능정보 서비스 분야별·기기별 이용현황, 검색서비스 이용, 포털사이트 알고리즘 추천 뉴스, 유튜브 알고리즘 추천 서비스, 개인정보 제공 허용 범위, 인공지능 추천 서비스 제공자 책무, 인공지능

보편화와 사생활 보호, 알고리즘 추천 서비스의 역기능에 더하여 생성형 AI 이용 현황, 생성형 AI 역기능, 디지털 역량, 인공지능에 대한 인식으로 구분하여 인포그래픽을 제공하고 있다. 앞으로도 이용자들의 인식 변화를 잘 확인할 수 있도록 주요 지표들과 인포그래픽을 제작하여 제공할 예정이다.

〔그림 4－10〕 패널조사 대시보드



자료: 인공지능 서비스 이용자정책 아카이브(https://user-archive.kisdi.re.kr), 2025. 12.

또한, 그래픽 사용자 인터페이스(GUI)를 적용하여 패널조사의 주요 지표를 차트로 보여주고 이를 표와 이미지 형태로 다운로드할 수 있는 주요지표를 지속 운영하였다. 올해 주요지표는 전년도와 마찬가지로 생성형 AI를 이용하지 않는 이유, 생성형 AI 이용 동기, 가장 많이 이용하는 생성형 AI 서비스(1순위), 생성형 AI 이용경험, 포털사이트 알고리즘 추천 뉴스의 부정적 우려 및 긍정적 기대, 지능형 서비스 이용 의도, 지능형 기기 이용 의도

로 구분하여 제공하였다. 2023년과 2024년 2개년치 조사 결과를 주요 지표에서 제공하고 있는 막대그래프 차트를 통해 직관적으로 확인할 수 있다.

이 외에도 패널조사의 핵심 결과를 집단별로 한눈에 파악할 수 있는 대시보드 메뉴를 지속 운영하고, 드롭다운 기능을 활용하여 2024년, 2023년, 2022년 자료를 손쉽게 이용할 수 있도록 하였다.

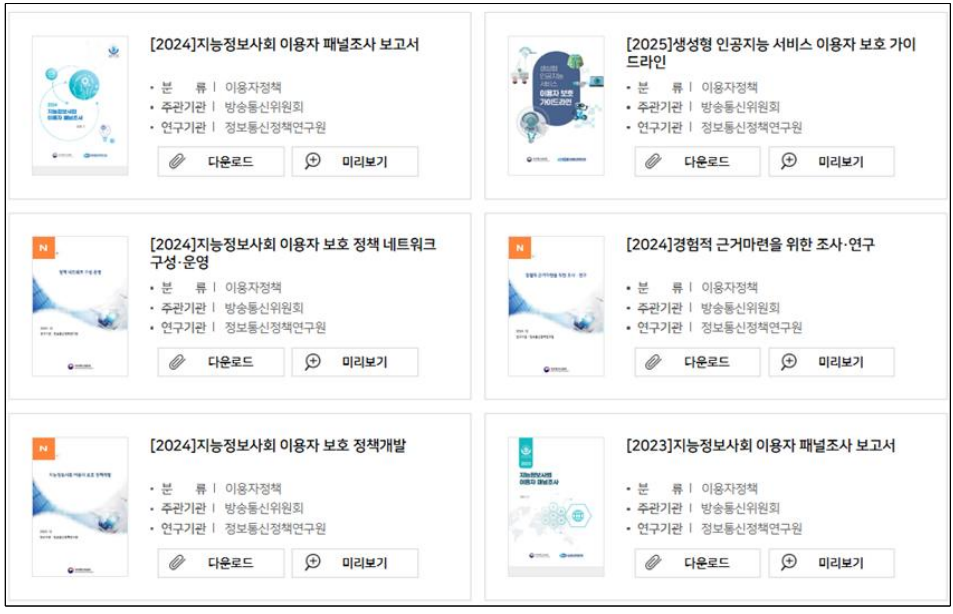
2. 인공지능 서비스 이용자정책 아카이브 2025년 운영현황

가. 자료 업데이트 및 관리

인공지능 서비스 이용자정책 아카이브에서 보도자료, 발간물, 민관협의회, 컨퍼런스 등 이용자정책과 관련한 자료를 업데이트하였다. 연구 성과물 특성상 특정 시기에만 업로드가 가능한 점을 감안하여 올해는 지속적인 콘텐츠 업로드를 추진하였다. 이를 위하여 방송미디어통신위원회와 정보통신정책연구원 유튜브에서 이용자와 관련한 숏츠, 영상을 아카이브에 업로드하고, 연구원 자체 발간물 가운데 인공지능 서비스 이용자와 관련성 있는 자료는 월별 검토를 통해 지속적으로 업데이트하였다.

먼저, 발간물 게시판은 이용자정책 관련 발간물과 연구원 자체 발간물, 기타로 구분하여 본 과제 관련 발간물을 검색할 때 용이하도록 메뉴를 세분화하였다. 「2024 지능정보사회 이용자 패널조사 보고서」, 「2024 경험적 근거마련을 위한 조사·연구」, 「2024 정책 네트워크 구성·운영」, 「2024 지능정보사회 이용자 보호 정책개발」 등의 정책보고서와 2025년 2월 말 발표한 「생성형 인공지능 서비스 이용자 보호 가이드라인」을 이용자 정책 발간물로 업로드하였다. 발간물은 파일을 다운로드받아 저장할 수도 있고, 미리보기 기능을 통해 바로 확인이 가능하다.

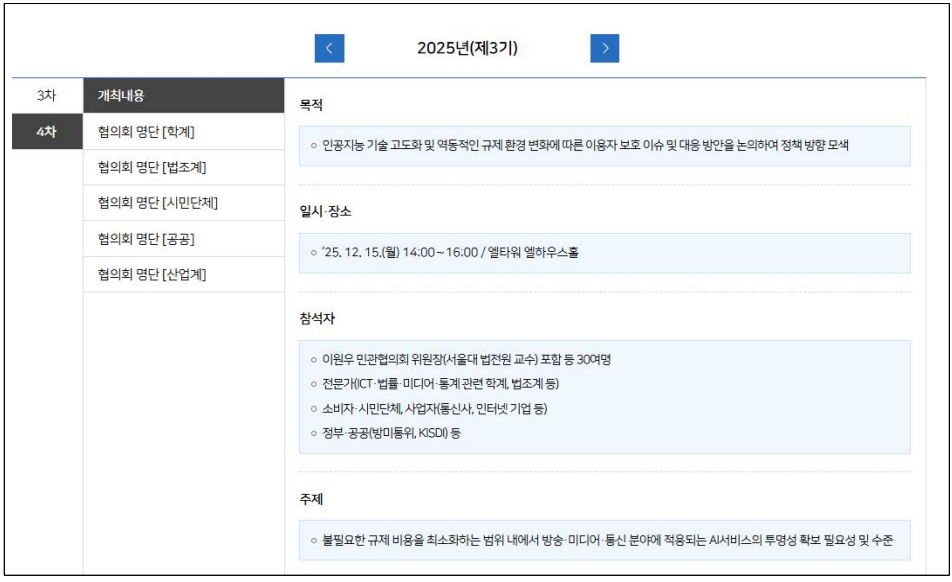
〔그림 4－11〕 발간물 업데이트 현황



자료: 인공지능 서비스 이용자정책 아카이브(https://user-archive.kisdi.re.kr), 2025. 12.

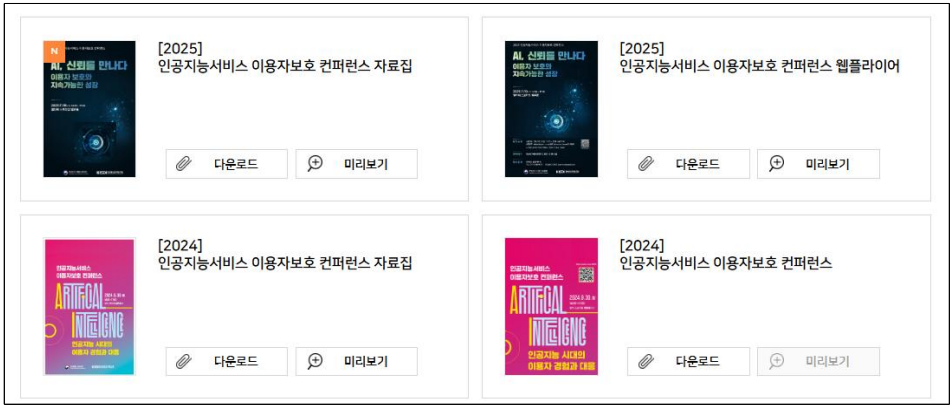
2025년에는 작년에 1기로 출범한 인공지능 서비스 이용자보호 민관협의회 3차, 4차 전체회의를 개최하였다. 민관협의회 개최 내용과 협의회 명단은 아카이브 자료실의 민관협의회 메뉴를 통해 확인할 수 있다. 또한, 2025년 인공지능 서비스 이용자보호 컨퍼런스 'AI, 신뢰를 만나다-이용자 보호와 지속가능한 성장'의 웹플라이어와 자료집은 컨퍼런스 페이지에 게시하여 공개하고 있다. 웹플라이어와 자료집 모두 파일의 직접 다운로드 뿐 아니라 웹에서 바로 확인할 수 있도록 미리보기 기능을 구현하였다.

〔그림 4－12〕 2025년 인공지능 서비스 이용자보호 민관협의회 업데이트 현황



자료: 인공지능 서비스 이용자정책 아카이브(https://user-archive.kisdi.re.kr), 2025. 12.

〔그림 4－13〕 컨퍼런스 업데이트 현황

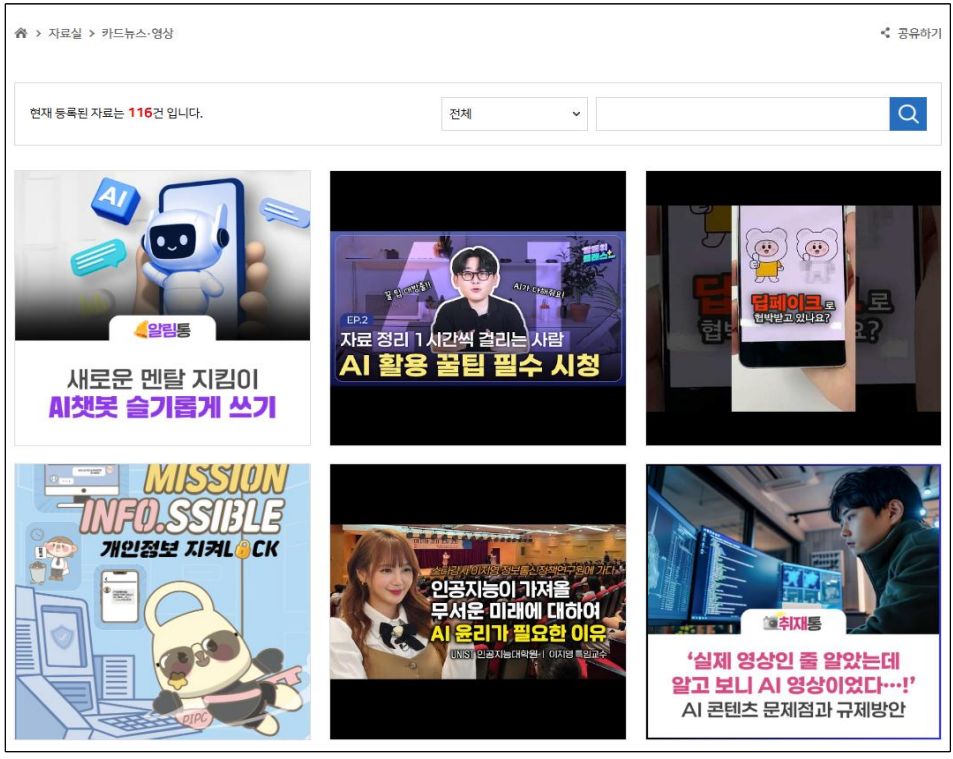


자료: 인공지능 서비스 이용자정책 아카이브(https://user-archive.kisdi.re.kr), 2025. 12.

자체 연구보고서 및 행사 관련 자료 등이 상시 업데이트되기 어려운 환경을 고려하여 아카이브의 지속적인 자료 업데이트와 이용자들의 이용률 제고를 위하여 방송미디어통신

위원회 보도자료, 블로그 카드뉴스를 지속하여 업데이트하였고, 올해는 방송미디어통신위원회, 정보통신정책연구원 공식 유튜브에 업로드되는 콘텐츠 가운데 인공지능 서비스 이용자 정책과 관련이 있는 영상과 숏폼을 상시 업데이트하여 제공하고 있다.

〔그림 4－14〕 카드뉴스 업데이트 현황



자료: 인공지능 서비스 이용자정책 아카이브(https://user-archive.kisdi.re.kr), 2025. 12.

인공지능 서비스 이용자정책 아카이브는 신뢰할 수 있는 인공지능 서비스 생태계를 위하여 산·학·연 전문가와 방송미디어통신위원회, 아카이브 이용자들의 의견을 적극적으로 수렴·반영하여, 인공지능 서비스 이용자정책에 대한 이용자들의 인식을 향상하고 실효적인 정책 방안 수립을 위해 지속적으로 노력하고 있다. 이를 위해 아카이브의 지속적인 유지와 보수 작업을 진행하고 있으며, 앞으로 더 많은 콘텐츠와 정보 제공을 통해 아카이브 기능을 더욱 강화해 나갈 계획이다.

나. 아카이브 이용 활성화를 위한 이벤트 진행

아카이브를 홍보하고 이용자 유입률 제고와 이용 활성화를 위해 아카이브 이용자를 대상으로 총 4회의 이벤트를 진행하였다. 올해 이용자 대상 이벤트는 이용자 패널조사 결과 공표 직후 여름 퀴즈 이벤트, 이용자 의견 수렴 및 이용 빈도를 확인한 가을 이벤트, 컨퍼런스 유튜브 시청 인증샷 이벤트, 마지막으로 마이크로데이터 이용자들을 대상으로 한 만족도 조사로 기획하여 진행하였다. 각각 이벤트에 대한 개요와 내용은 다음과 같다.

1) 여름 퀴즈 이벤트

2025년 6월, 아카이브에 2024년 지능정보사회 이용자 패널조사를 공표하면서 조사결과의 홍보와 아카이브 이용률 제고를 위한 퀴즈이벤트를 진행하였다. 2024년 패널조사 결과 가운데 '지능정보 서비스 분야별 이용 현황'의 인포그래픽에서 확인할 수 있는 간단한 퀴즈를 내고 정답자 가운데 추첨을 통해 기프티콘을 제공하는 이벤트를 진행하였다. 이용자 패널조사 업데이트 시기에 맞춘 여름 퀴즈 이벤트를 통해 패널조사와 아카이브 홍보 효과를 기대하였다.

〈표 4-2〉 여름 퀴즈 이벤트

구분	내용
내용	2024년 이용자 패널조사 결과 공표 홍보를 위해 인포그래픽 수치 퀴즈 이벤트 기획 - 지능정보 서비스 분야별 이용 경험 중 '소비'의 이용 경험 수치 맞추기
기간	2025. 6. 23.(월) ~ 7.4.(금)
결과	2025. 7. 9.(수)
참여자	총 4,245명

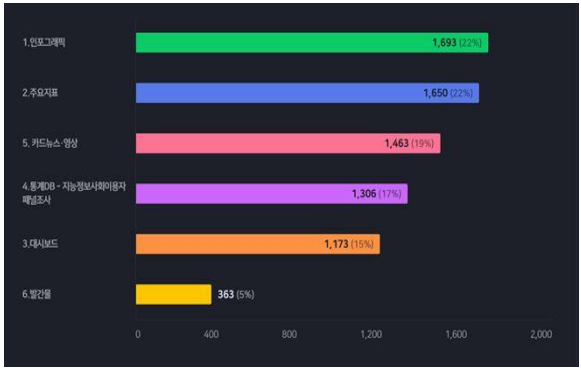
2) 이용 빈도 및 의견 수렴 이벤트

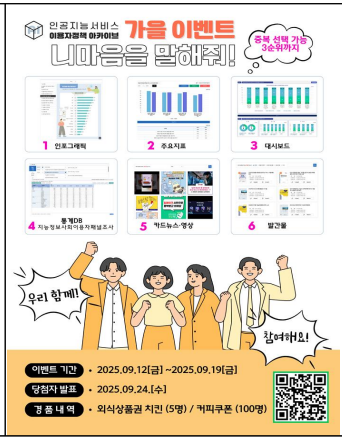
인공지능 서비스 이용자정책 아카이브 메뉴에서 이용 빈도나 선호도를 3순위까지 중복 선택하게 하고, 아카이브 관련 의견 수렴을 위해 가을 이벤트를 진행하였다. 아카이브 이용자들이 가장 많이 사용하고 있는 메뉴는 인포그래픽, 주요지표, 카드뉴스·영상, 통계DB, 대시보드, 발간물 순인 것으로 조사되었고, 총 3,226명의 참여자 가운데 중복을 제거하고 좋은 의견을 제시한 참여자들에게 기프티콘을 제공하였다.

〈표 4－3〉 이용 빈도 및 의견 수렴 가을 이벤트

구분	내용
내용	인공지능 서비스 이용자정책 아카이브 메뉴 이용 및 선호도 조사, 아카이브에 대한 의견 제시
기간	2025. 9. 12.(금) ~ 9. 19.(금)
결과 발표	2025. 9. 24.(수) ※ 당첨자는 아카이브 공지사항을 통해 발표
참여자	총 3,226명

〔그림 4－15〕 아카이브 메뉴 이용 빈도 및 의견 수렴 가을 이벤트 페이지





자료: 인공지능 서비스 이용자정책 아카이브(https://user-archive.kisdi.re.kr), 2025. 12.

3) 컨퍼런스 유튜브 시청 인증샷 이벤트

2025년 11월 19일, 'AI, 신뢰를 만나다-이용자 보호와 지속가능한 성장'이라는 주제로 개최한 컨퍼런스를 홍보하고 온라인 참여 활성화를 위해 유튜브 시청 인증샷 이벤트를 진행하였다. 유튜브를 시청한 후 영상 인증샷을 캡처한 후, QR코드로 인증결과를 제출하는 것으로 이벤트 참여를 할 수 있도록 하였고, 이벤트 기간은 컨퍼런스 시작 일시부터 해당 주 금요일까지로 정하여 컨퍼런스 종료 이후라도 온라인에서 컨퍼런스를 시청할 수 있도록 독려하였다. 가급적 많은 이들의 참여가 이루어질 수 있도록 당첨자 수를 200명으로 늘리고, 추첨을 통해 모바일 커피 기프티콘을 지급하였다.

〈표 4-4〉 컨퍼런스 유튜브 시청 인증샷 이벤트

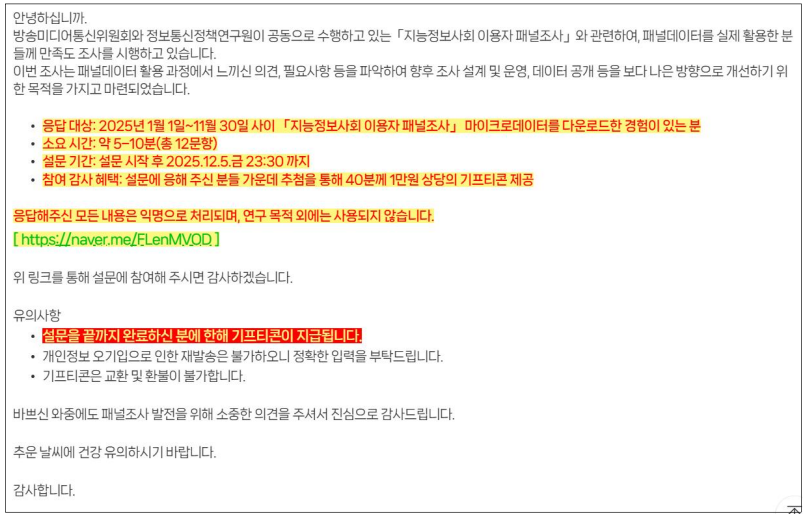
구분	내용
내용	2025년 인공지능 서비스 이용자보호 컨퍼런스 온라인 참여 독려 및 홍보
기간	2025. 11. 19.(수) ~ 11. 21.(금)
결과	2025. 11. 24.(월)
참여자	총 913명

4) 마이크로데이터 이용자 대상 만족도 조사

마지막으로, 기존에는 아카이브 이용자 참여란에만 등록했던 마이크로데이터 이용자 대상 만족도 조사를 실제 DB를 내려받은 이메일로 보내 조사 참여를 독려하였다. 국가승인통계의 우수통계 조건으로 이용자들을 대상으로 만족도 조사를 진행하는 항목이 있는데, 기존에는 별도의 이벤트로 진행하지 않아 만족도 조사결과를 활용할 수 없는 참여율을 보

였다. 올해에는 실제 마이크로데이터 이용자들에게 이메일로 간단하게 참여할 수 있는 설문 양식을 송부하였고, 응답자들에게 기프티콘을 지급하는 이벤트로 기획·운영하여 총 43명의 응답을 수합하였다. 설문에는 패널조사를 알게 된 경위, 아카이브를 방문한 경위, 데이터 접근의 용이성, 구성과 설명의 이해도, 데이터의 품질, 실증적 근거 제공 여부, 활용 계획, 추가 이슈 등에 대한 질문을 포함하였다.

〈표 4－5〉 마이크로데이터 이용자 대상 만족도 조사



안녕하십니까.
방송미디어통신위원회와 정보통신정책연구원이 공동으로 수행하고 있는 「지능정보사회 이용자 패널조사」와 관련하여, 패널데이터를 실제 활용한 분들께 만족도 조사를 시행하고 있습니다.
이번 조사는 패널데이터 활용 과정에서 느끼신 의견, 필요사항 등을 파악하여 향후 조사 설계 및 운영, 데이터 공개 등을 보다 나은 방향으로 개선하기 위한 목적을 가지고 마련되었습니다.

- 응답 대상: 2025년 1월 1일~11월 30일 사이 「지능정보사회 이용자 패널조사」 마이크로데이터를 다운로드한 경험이 있는 분
- 소요 시간: 약 5~10분(총 12문항)
- 설문 기간: 설문 시작 후 2025.12.5.금 23:30 까지
- 참여 감사 혜택: 설문에 응해 주신 분들 가운데 추첨을 통해 40분께 1만원 상당의 기프티콘 제공

응답해주신 모든 내용은 익명으로 처리되며, 연구 목적 외에는 사용되지 않습니다.
[https://naver.me/FLenMVOD]

위 링크를 통해 설문에 참여해 주시면 감사하겠습니다.

유의사항
- 설문을 끝까지 완료하신 분에 한해 기프티콘이 지급됩니다.
- 개인정보 오기입으로 인한 재발송은 불가하오니 정확한 입력을 부탁드립니다.
- 기프티콘은 교환 및 환불이 불가합니다.

바쁘신 와중에도 패널조사 발전을 위해 소중한 의견을 주셔서 진심으로 감사드립니다.

추운 날씨에 건강 유의하시기 바랍니다.

감사합니다.

제5장 결 론

1. 인공지능 서비스 이용자 보호 민관협의회 운영 결과 및 후속 추진 방향

2020년 4월에 출범한 '지능정보사회 이용자 보호 민관협의회'는 지난 4년간의 운영을 통해 이용자 정책 현안에 대한 각계 의견을 수렴하는 내실 있는 정책협의체로 자리 잡았다. 본 민관협의회 의제의 출발점은 2019년 「이용자 중심의 지능정보사회를 위한 원칙」에 있었다. 2020년(1차~5차)에는 '사람 중심의 지능정보서비스 제공', '투명성과 설명가능성', '공정성과 차별금지' 각 원칙과 관련한 국내외 사례를 공유하고 지능정보사회의 이용자 보호 필요성에 대한 공감대를 형성하는 데 초점을 두었다. 이어서 2021년(1차~5차)에는 '책임성', '안전성', '참여' 원칙과 관련하여 공공·기업·시민의 책임의식을 제고하고 구체적인 실천방안을 도출하는 데 주력하였다. 특히 2021년 6월 발표된 「인공지능 기반 미디어 추천 서비스 이용자 보호 기본원칙」에 대한 각계 전문가 의견을 수렴하고 합리적 이행방안을 논의하는 데 제1기 민관협의회(2020~2021)가 중추적 역할을 하였다.

이어서 2022년(1차~5차)과 2023년(6차~9차)에는 국내 기업의 참여를 확대하여 구성한 제2기 지능정보사회 이용자 보호 민관협의회를 운영하였다. 2022년에는 대표적인 플랫폼 서비스(메타버스, 전문 중개 서비스, 알고리즘 저널리즘 등) 이용 현황을 조망하고 새로운 이용자 보호 과제를 검토하였으며, 급증하는 신규 지능정보서비스에 통용되는 민관 협력적 이용자 정책 프레임워크를 논의하였다. 2023년에는 인공지능 서비스 생태계에 큰 파장을 가져온 '생성형 인공지능'에 주목하여 최신 서비스 모델을 이해하고 데이터 수집·활용 구조를 고려한 합리적인 이용자 보호 방안에 대해 심층 논의하였다. 2024년은 생성형 인공지능 서비스 이용자가 점차 증가하고 실제 이용 사례 및 이용자 보호 이슈가 대두됨에 따라, 국내외의 법제 동향과 관련 정책 및 법안 도입의 찬반을 실제적 차원에서 논의하였다. 2025년은 인공지능 서비스 이용자 보호 이슈의 축적과 대응을 위한 실질적 법제 방안을 위한 논의가 진행되었으며, 기술 환경의 변화에 의한 패러다임 변화와 이에 발 맞추어 규제 패러다임 방향 설정을 논의하였다.

인공지능 서비스 이용자 보호 이슈는 가시적인 형태로 실제 이용자들의 삶에 긍정 및 부정적인 영향을 주기 시작하였고, 많은 이들이 경계의 시선과 함께 가능성의 미래를 그려나가고 있다. 인공지능 서비스와 관련한 논의는 잠재성과 탐색에서 벗어나 실제 이용자 이슈와 해결책, 더 나아가 정책과 입법 논의로 도약하고 있다. 앞으로도 본 민관협의회는 인공지능 서비스 이용자 정책 담론의 장으로서 각계의 논의 참여를 유도하는 동시에, 현안에 대한 고민과 검토의 결과가 정부의 정책 방향과 제도 수립의 결실로 이어지도록 실질적인 기여를 해나갈 것이다. 지금까지 지능정보사회 이용자 보호 민관협의회가 이면의 중심 의제로 삼았던 '자율', '책임', '소통'이라는 세 개의 축은 인공지능 서비스 이용자 보호 민관협의회가 이어받고 발전시켜, 장기적으로 이용자, 사업자, 정책입안자 간의 대립 구도를 완화하고 가변적인 기술 환경 속에서 우리 사회의 기본질서를 유지하는 데 불가결한 대원칙이 될 것이다.

2. 이용자 정책협력 네트워크 구축 결과 및 후속 추진 방향

인공지능 기술과 서비스의 폭발적 부상과 급속한 확산으로 다양하게 제기되는 정책적 이슈와 사회문화적 충격을 인공지능 서비스 이용자 보호 관점에서 논의하기 위해 이용자 정책협력 네트워크 구축을 추진, 운영하였다. 국내외 이용자 정책협력 네트워크 운영을 통해 지능정보사회 이용자 보호 방안을 선제적으로 논의하고, 지능정보기술과 서비스 이용자들이 자신의 권리를 자각할 수 있도록 '지능정보사회 이용자 보호 국제 컨퍼런스'를 19년부터 연례행사로 개최하고 있다. 국제 컨퍼런스에 국내외 주요 국제기구를 비롯한 지능정보기술과 서비스를 관할하는 유관 정부 기관, 연구기관, 그리고 지능정보기술과 서비스를 제조하고 확산시키는 글로벌 빅테크 등에서 전문가를 초청하여 지능정보사회 이용자 보호 관련 현안을 논의하고 국내 이용자 보호 정책추진 성과를 공유하고 실효적 정책 방안을 숙고하였다. 6회를 맞이한 2024년에는 '인공지능 시대의 이용자 경험과 대응'를 주제로 AI 관련 전문가들이 참여하여 인공지능 기술과 서비스의 발전 및 확산으로 발생하는 이용자 보호 이슈를 공유하고, 생성형 AI 등 신기술로 인한 이용자의 위상 변화 및 바람직한 AI 이용자 보호 정책 방향 모색을 위한 논의를 진행하였다.

'인공지능 서비스 이용자 보호 컨퍼런스'는 역동적으로 진화하고 있는 인공지능 기술과

서비스를 둘러싸고 끊임없이 제기되는 역기능과 부작용을 이용자 보호 관점에서 논의하고 정책적 시사점을 제시하며, 정책연구 성과를 교환하고 발전적 요소를 개발하기 위해 마련된 산·학·연 전문가와 이용자들이 함께하는 논의의 장이다. 아울러 컨퍼런스는 방송통신위원회가 그간 정비했던 「인공지능 기반 미디어 추천 서비스 이용자 보호 기본원칙」과 「메타버스 이용자 보호 기본원칙」을 생성형 인공지능에 의해 파생된 다양한 문제점에 적용함으로써 그 실효성을 직간접적으로 진단하는 자리를 제공했다. 2025년에는 생성형 인공지능을 필두로 지능정보기술과 서비스의 발전이 지속될 것으로 예상됨에 따라, 지능정보기술과 서비스의 발전이 가져오는 다채로운 변화를 심층적으로 이해하고, 상응하는 정책적 전략을 도출하여, 지능정보기술의 역기능으로부터 이용자를 안전하게 보호하고 지능정보기술에 대한 신뢰를 구축하기 위해 요구되는 법제도적 논의를 지속할 것이다.

3. 국민참여 채널 및 지식공유 플랫폼 운영 결과 및 후속 추진 방향

인공지능의 일상화 시대에 새롭게 발생하는 이용자 정책 이슈와 쟁점을 공유하고 그에 대한 대응 정책을 모색하는 사회적 공론장이자 전문가와 이용자 간의 소통 플랫폼으로서 지난 2020년에 오픈한 이용자 정책 아카이브는 올해로 6년차 운영에 접어들었다. 당초 지능정보사회 이용자정책 아카이브에서 출발하여 사회·경제·산업 등 사회 전반의 환경 변화의 주된 동인이자 핵심기술인 인공지능 기반 방송미디어통신 서비스의 확산과 인공지능이 이용자에게 미치는 영향을 고려하여 인공지능 서비스 이용자정책 아카이브로 명칭을 변경한 후 운영하고 있다. 올해는 주로 아카이브 사용자의 편의성을 제고하는 측면에 중점을 두고 3개년치 시계열 데이터 구현과 UI/UX를 점진적으로 개선하고자 했다. 특히, 이용자들의 이용빈도를 고려하여 다양한 인포그래픽을 제작하고, 아카이브 검색과 메인 페이지에서의 가시성을 제고하기 위해 아카이브 아이덴티티를 기획하고 반영하였다. 또한, 영상·숏폼의 콘텐츠를 업로드하여 이용자들에게 친숙한 형식의 콘텐츠로 인공지능 서비스 이용자보호와 관련한 정보를 공유하고자 하였다.

향후 아카이브의 이용 활성화와 본연의 역할 수행을 위해서는 메인 페이지의 UI/UX를 사용자 친화적으로 개선하고, 이용자 패널조사 결과를 포함하여 관련 데이터를 정기적으로 업데이트해야 할 필요가 있다. 특히, 데이터 시각화를 확대하고 빅카인즈 연동 등으로

이용자정책 관련 최신 동향을 직관적으로 활용할 수 있도록 아카이브의 도구를 강화하는 방안 마련이 필요할 것으로 보인다. 이를 통해 일반 이용자부터 연구자, 산업 전문가, 정책 입안자까지 다양한 이해당사자가 데이터를 쉽게 접근하고 활용할 수 있을 것으로 기대한다. 한편, 본 아카이브가 이용자 정책 방향 수립을 위한 공론의 장으로 기능하기 위해서는 이용자들의 의견을 더욱 적극적으로 수렴하기 위한 다방향 소통 플랫폼으로의 발전이 요구된다. 예컨대 인공지능 이용자 정책 관련 설문조사, 피드백 채널, 온라인 이벤트 등을 통해 수렴한 이용자들의 의견을 아카이브 개선이나 이용자 정책에 반영하고, 이를 기반으로 정책 방향을 수립한다면 인공지능 서비스 이용자의 권익 보호와 함께 인공지능의 잠재적 위험을 완화하는 데 기여할 수 있을 것이다.

참 고 문 헌

[국내 문헌]

인공지능 서비스 이용자정책 아카이브(https://user-archive.kisdi.re.kr).

2024년도 지능정보사회 이용자 패널조사-생성형 AI 이용 현황
(https://user-archive.kisdi.re.kr/kor/infoGrp/infoGrpList.html).

● 저 자 소 개 ●

조 성 은
- 럿거스대학교 인터넷커뮤니케이션학 박사
- 현 정보통신정책연구원 연구위원

문 정 욱
- 고려대학교 행정학 박사
- 현 정보통신정책연구원 디지털사회전략연구실 실장

연 소 라
- Texas A&M University 경제학 박사
- 현 정보통신정책연구원 부연구위원

이 으 뜸
- University of Virginia 경제학 박사
- 현 정보통신정책연구원 부연구위원

이 시 직
- 연세대학교 법학 박사 수료
- 현 정보통신정책연구원 부연구위원

김 지 혜
- 서강대학교 국제학 석사
- 현 정보통신정책연구원 전문연구원

이 지 호
- 경희대학교 문화관광콘텐츠학 석사
- 현 정보통신정책연구원 연구원

이 세 인
- 성균관대학교 000
- 현 정보통신정책연구원 연구원

연 제 선
- 성균관대학교 인터랙션사이언스학과 석사
- 현 정보통신정책연구원 연구원

전 민 경
- 서울과학기술대학교 AI공공정책학 석사
- 현 정보통신정책연구원 연구원

정책연구 25-20

정책 네트워크 구성·운영

2025년 12월 일 인쇄
2025년 12월 일 발행

발행인 방송미디어통신위원회 위원장
발행처 방송미디어통신위원회
경기도 과천시 관문로 47
Homepage: www.kcc.go.kr
인 쇄 인 성 문 화