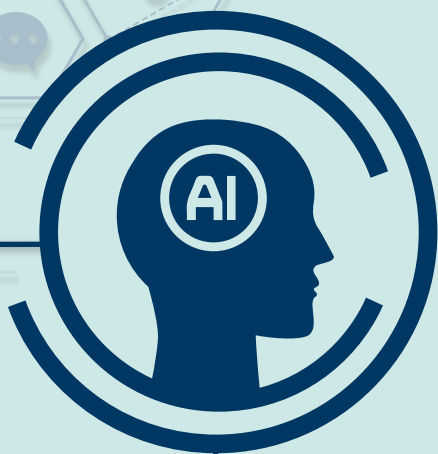


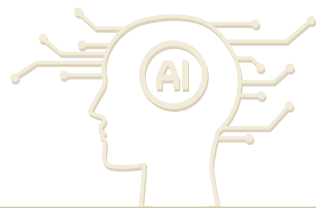


인공지능 기반 미디어 추천 서비스 이용자 보호 기본원칙 해설서



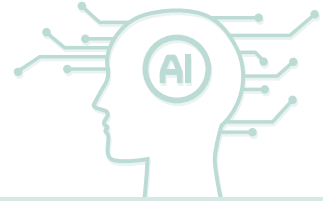
방송통신위원회

KISDI 정보통신정책연구원
KOREA INFORMATION SOCIETY DEVELOPMENT INSTITUTE



이 해설서는 방송통신위원회가 2021년 6월 30일 발표한 「인공지능 기반 미디어 추천 서비스 이용자 보호 기본원칙」에 제시된 개별 항목 및 원칙의 제정 취지와 주요 내용에 대한 이해를 돕기 위하여 마련되었습니다. 본 기본원칙 제정 전후의 의견 수렴 과정에서 확인되었던 각계각층의 인식과 수용도의 격차를 좁히고 디지털 미디어 영역에서 자율규제 원칙이 타당하게 정립되고 이행되도록 간략한 해석지침을 제공하는 데 주안점을 두었습니다. 기본원칙의 내용 중 5대 실행 원칙과 관련하여 구체적인 이행 방안, 적용 사례, 유의사항 등을 담은 실행지침(가칭)은 향후 별도로 마련될 계획입니다.

Contents



I	목적	1
II	용어의 정의	3
	* 참고 관련 법령(안) 규정	7
III	핵심 원칙	10
	1. 투명성	10
	2. 공정성	13
	3. 책무성	15

Contents

IV 실행 원칙.....17

1. 이용자를 위한 정보 공개.....	17
* 참고 관련 법령(안) 규정.....	19
2. 이용자의 선택권 보장.....	23
*참고 관련 법령(안) 규정.....	25
3. 자율검증 실행.....	28
*참고 국내외 사례.....	33
4. 불만 처리 및 분쟁 해결.....	37
*참고 국내외 사례.....	39
5. 내부 규칙 제정.....	42
*참고 국내외 사례.....	44

V 기본원칙 이행을 위한 정부의 역할.....47

1. 서비스 제공자 지원.....	47
2. 이용자 역량 강화.....	49
* 참고 국내외 사례.....	51

I 목적

목적

- 가. 본 기본원칙은 디지털 미디어 플랫폼에서 상용되는 인공지능 기반 추천 서비스의 투명성과 공정성을 제고하기 위하여 추천 서비스 제공자에게 권고되는 자율적인 실천규범이다.
- 나. 본 기본원칙은 디지털 미디어 콘텐츠의 편향적·차별적 제공을 방지하고 이용자의 권익을 적극적으로 보장함으로써 지능정보사회의 자유민주적 기반을 형성하는 미디어의 가치를 수호하는 것을 목적으로 한다.

1. 규정 취지

「인공지능 기반 미디어 추천 서비스 이용자 보호 기본원칙」은 디지털 플랫폼에서 제공되는 ‘미디어 추천 서비스’의 투명성과 공정성을 확보함으로써 미디어 이용자의 권익을 보장하고 미디어 본연의 가치와 기능이 디지털 미디어 영역에서도 유지되도록 하는 데 목적을 둔다. 이를 달성하기 위한 규범적 토대로서 본 기본원칙이 제정되었으며, 법적 강제 수단보다 수범자에 의한 자율규제 방식을 우선하여 추천 서비스 규율 원칙을 수립하는 것에 근본 취지가 있다.

2. 주요 내용

‘목적’ 규정에서는 본 기본원칙의 성격과 이행 목표를 밝히고 있다. 디지털 플랫폼을 중심으로 미디어 영역이 재편되는 현시점에도 미디어의 공익적·사회적 기능이 훼손되어서는 안 된다. 디지털 플랫폼이 지능정보사회의 ‘대중 미디어’로 온전히 기능할 수 있도록 추천 서비스에서 문제시되는 편향성, 차별성 등 잠재적 위험을 관리하는 자율규제를 체계화할 필요가 있다. 본 기본원칙은 추천 서비스 제공자가 직접 행하는 자율규제의 토대가 되는 권고적 규범에 해당한다. 이는 2019년 11월

방통위가 발표한 ‘이용자 중심의 지능정보사회를 위한 원칙’의 기초를 디지털 미디어 플랫폼의 추천 서비스를 대상으로 구체화한 특칙이라는 의미도 함께 지닌다.

방송·통신 분야의 디지털 전환 이후로 국내외 유수의 플랫폼 기업은 TV, 신문 등 언론사 중심의 전통적인 미디어와 정보 영역을 상당 부분 공유하고 있다. 디지털 플랫폼이 대중에게 광범위한 미디어 콘텐츠를 전달함으로써 정보 제공, 여론 형성 등 미디어 고유의 기능을 분담하게 된 오늘날에는 디지털 플랫폼 서비스가 ‘미디어 이용자’에 대한 불이익, 차별, 기타 권익 침해를 유발하지 않도록 경계할 필요가 있다. 본 기본원칙의 중심을 이루는 3대 핵심 원칙과 5대 실행 원칙은 주요한 미디어 서비스로 자리잡은 ‘인공지능 기반 추천 서비스’를 제공하는 자가 사회적·법적 책임을 다하는 차원에서 스스로 준수해야 할 목표와 수단을 제시한 것이다.

Ⅱ 용어의 정의

용어의 정의

본 기본원칙에서 사용되는 용어의 정의는 다음과 같다.

- 가. “디지털 미디어 플랫폼”이란 디지털 미디어 콘텐츠의 제공 또는 매개를 목적으로 하는 인터넷 홈페이지, 모바일 응용프로그램, 그 밖에 이에 준하는 전자적 시스템을 말한다.
- 나. “디지털 미디어 콘텐츠”란 정보제공·오락·교육·광고 등을 목적으로 일반 대중에게 제공하는 문자·이미지·영상 등의 자료 또는 정보로서 그 보존 및 이용의 효용을 높일 수 있도록 디지털 형태로 제작하거나 처리한 것을 말한다.
- 다. “인공지능 기반 추천 서비스”란 인공지능 알고리즘을 적용하여 완전히 또는 부분적으로 자동화된 콘텐츠 배열 시스템(이하 “추천 시스템”이라 한다)을 통하여 이용자에게 미디어 콘텐츠를 선별적으로 노출시키는 서비스(이하 “추천 서비스”라 한다)를 말한다.
- 라. “추천 서비스 제공자”란 디지털 미디어 플랫폼의 추천 시스템을 활용하여 미디어 콘텐츠를 제공하거나 매개하는 사업자를 말한다.
- 마. “이용자”란 디지털 미디어 플랫폼에서 제공되는 추천 서비스를 이용하거나 이용하려는 자를 말하며 해당 플랫폼을 통하여 미디어 콘텐츠를 소비하는 자 또는 플랫폼 서비스를 이용하여 미디어 콘텐츠를 공급하는 자를 모두 포함한다.

1. 규정 취지

‘용어의 정의’에서는 본 기본원칙에 쓰인 주된 용어의 뜻을 밝힌다. 본 기본원칙에서 채택한 용어의 일반적·추상적 의미를 규정함으로써 기본원칙의 수범자, 적용 범위, 보호 대상 등을 특정하여 규율의 근거를 명확히 하는 것에 용어 정의의 목적이 있다. 아직 규범적 용례가 없거나 미디어 분야에 부분적으로 도입된 신기술을 규율 대상으로 개념화하는 것이므로, 용어의 제한적 의미를 명시함으로써 본 기본원칙을 해석·적용하면서 발생하는 개념의 왜곡·확대나 적용상 오류를 방지하려는 것이다.

2. 주요 내용

가. 디지털 미디어 플랫폼

본 기본원칙에서 말하는 ‘디지털 미디어 플랫폼’은 인터넷 홈페이지, 모바일 응용프로그램, 그 밖에 이에 준하는 전자적 시스템을 통해 디지털 미디어 콘텐츠를 제공 또는 매개하는 자를 가리킨다. 종래의 미디어 영역이 콘텐츠 생산 주체(방송사, 신문사 등)를 중심으로 분류되었던 것과 달리 디지털 미디어 영역에서는 인터넷상의 ‘플랫폼’이 미디어 콘텐츠가 생산되고 소비되는 주된 장소적 개념이 되고 있다.

디지털 미디어는 기술적 의미에서 디지털 코드를 기반으로 동작하는 전자 매체로 정의되며, 인쇄된 책, 신문, 사진 필름, 오디오 테이프, 비디오 테이프와 같은 아날로그 미디어와는 형태 면에서 확연한 대조를 이룬다. 특히 본 기본원칙에서는 방송사, 언론사, 인터넷뉴스서비스사업자, 온라인 비디오 플랫폼 등 기존 또는 새로운 유형의 미디어 서비스를 제공하는 ‘플랫폼’을 대상으로 한다. 이러한 플랫폼 서비스 기반의 디지털 미디어가 본 기본원칙의 적용 영역에 해당한다.

나. 디지털 미디어 콘텐츠

디지털 플랫폼을 통해 유통되는 디지털 콘텐츠는 텍스트·오디오·비디오 파일에서 그래픽, 애니메이션, 기타 이미지에 이르기까지 다양한 형태로 제공된다. 그중에서도 디지털 미디어 콘텐츠는 ‘미디어’ 요소가 체화된 디지털 콘텐츠로 정의할 수 있다. 이와 달리, 디지털 콘텐츠 중에서도 전자상거래 플랫폼의 현물 거래 대상이 되는 재화

대한 일체의 정보가 자동 추천 대상이 되는 경우는 본 기본원칙의 적용 범위에서 제외한다.

디지털 콘텐츠, 미디어 서비스 등의 개념을 정의한 국내외 법 규정 사례를 종합하면, 디지털 미디어 콘텐츠는 ‘정보제공·오락·교육·광고 등을 목적으로 일반 대중에게 제공하는 문자·이미지·영상 등의 자료 또는 정보에 해당하는 디지털 정보’라는 광의의 의미로 정의될 수 있다. 다만 본 기본원칙이 인공지능 기반 추천 서비스에 한하여 적용되는 것인 이상, 여기서 말하는 디지털 미디어 콘텐츠는 ‘추천(자동 배열)의 대상’이 되는 콘텐츠로 한정된다.

다. 인공지능 기반 추천 서비스

‘인공지능 기반 추천 서비스’란 인공지능 알고리즘을 적용하여 완전히 또는 부분적으로 자동화된 콘텐츠 배열 시스템, 즉 ‘추천 시스템’을 통하여 이용자에게 미디어 콘텐츠를 선별적으로 노출시키는 서비스를 뜻한다. 추천 시스템은 알고리즘 시스템의 일종으로 이용자로부터 수집한 개인정보, 행태정보 등을 활용하여 이용자별로 맞춤형 추천 서비스를 지원하는 것이 일반적이다.

디지털 미디어 플랫폼에서 통상적으로 운용되는 추천 서비스는 이용자의 행동 패턴을 학습하여 특정 콘텐츠에 대한 개별 이용자 또는 유사 이용자 집단의 선호도를 예측하는 방식으로 콘텐츠 노출 순서가 결정된다. 요컨대 본 기본원칙의 적용 대상이 되는 추천 시스템은 인공지능 시스템 중에서도 머신러닝 기법을 비롯한 각종 추천 알고리즘이 적용된 콘텐츠 자동 배열 시스템으로 이해할 수 있다.

라. 추천 서비스 제공자

추천 서비스 제공자는 본 기본원칙의 직접적인 수범자로서 디지털 미디어 플랫폼의 추천 시스템을 활용하여 미디어 콘텐츠를 제공하거나 매개하는 사업자를 말한다. 플랫폼 사업자가 직접 추천 시스템을 관리·운영하는 경우에는 플랫폼 사업자가 곧 추천 서비스 제공자가 된다.

본 기본원칙은 자율규범으로서 개별 원칙의 불이행에 따른 법적 강제가 예정되지 않은 권고적 성격을 지니므로, 법률상 규제와 달리 수범자의 범위를 엄격히 제한할 필요는 없다. 물론 해당 플랫폼의 성격과 추천 서비스의 유형에 따라 잠재적 위험의 내용과 고저(高低)가 달라지므로 그에 따라 각 원칙을 이행하는 수준과 방식에

차이를 두어야 할 것이다.¹⁾

마. 이용자

본 기본원칙은 ‘이용자’를 디지털 미디어 플랫폼을 통해 제공되는 추천 서비스를 이용하거나 이용하려는 자로 정의한다. 또한 이용자의 범주에 플랫폼을 통하여 미디어 콘텐츠를 소비하는 자와 플랫폼 서비스를 이용하여 미디어 콘텐츠를 공급하는 자를 모두 포섭함으로써 본 기본원칙의 이용자 보호 범위를 두텁게 설정하고 있다. 즉, 추천 서비스의 투명성과 공정성을 제고하여 보호하려는 이용자는 ‘미디어 콘텐츠 소비자’와 ‘이용사업자’ 모두를 가리킨다.

이처럼 이용자 개념을 양분하여 규정한 것은 5대 실행 원칙 중 ‘선택권 보장’이 미디어 콘텐츠 소비자만을 대상으로 한다는 점에서 의미가 있다.

1) 적용 대상에 따라 각 실행 원칙의 이행 수단을 차등화하는 방안에 대해서는 본 기본원칙의 실행지침을 마련하는 과정에서 논의를 진행할 예정이다.

i 참고 | 관련 법령(안) 규정

1) '디지털 미디어 플랫폼' 관련 규정

법령(안)명	조문
「온라인 플랫폼 기본법안」 (21. 11. 윤두현의원 대표발의)	“온라인플랫폼”이란 재화, 용역(일정한 시설을 이용하거나 용역을 제공할 수 있는 권리를 포함한다. 이하 같다) 또는 디지털 콘텐츠 제공에 관한 정보 교환 또는 계약 체결이 가능한 인터넷 홈페이지, 모바일 응용프로그램 및 이에 준하는 전자적 시스템을 말한다(제2조제1호).
「온라인 플랫폼 이용자 보호에 관한 법률안」 (20. 12. 전해숙의원 대표발의)	“온라인플랫폼”이란 이용자 간에 의사 소통 및 정보 교환, 재화·용역 또는 디지털콘텐츠(이하 “콘텐츠 등”이라 한다)의 거래 등 상호작용을 매개하기 위해 전기통신설비와 컴퓨터 및 컴퓨터의 이용기술을 활용하여 정보를 수집·가공·저장·검색·송신 또는 수신하는 정보통신체제를 말한다(제2조제1호).
「온라인 플랫폼 중개거래의 공정화에 관한 법률안」 (21. 1. 정부 제출)	“온라인플랫폼”이란 재화 또는 용역(일정한 시설을 이용하거나 용역을 제공할 수 있는 권리를 포함한다. 이하 같다)의 거래와 관련된 둘 이상 이용자 간의 상호작용을 위하여 정보통신설비를 이용하여 설정된 전자적 시스템으로서 대통령령으로 정하는 것을 말한다(제2조제1호).
일본 「특정 디지털플랫폼 투명성 및 공정성 향상에 관한 법률」 (特定デジタルプラットフォーム の透明性及び公正性の向上に 関する法律)	“디지털플랫폼”이란 다수의 이용을 예정하고 컴퓨터를 이용해 정보를 처리함으로써 구축한 장(場)으로 상품, 서비스 또는 권리를 제공하고자 하는 자의 해당 상품 등과 관련된 정보를 표현한 것을 다수에게 인터넷 및 그 밖의 고도정보통신 네트워크를 통해 제공하는 서비스를 말한다(제2조제1항).
유럽연합 「디지털서비스법안」 (Digital Services Act)	서비스 이용자의 요청에 따라 정보를 저장하고 대중에 배포하는 매개 서비스 제공자를 말한다(art. 2(h)).

2) '디지털 미디어 콘텐츠' 관련 규정

법령(안)명	조문
「온라인 디지털콘텐츠산업 발전법」	“디지털콘텐츠”란 부호·문자·음성·음향·이미지 또는 영상 등으로 표현된 자료 또는 정보로서 그 보존 및 이용에 있어서 효용을 높일 수 있도록 전자적 형태로 제작 또는 처리된 것을 말한다(제2조제1호).

법령(안)명	조문
「문화산업진흥기본법」	“디지털콘텐츠”란 부호·문자·도형·색채·음성·음향·이미지 및 영상 등(이들의 복합체를 포함한다)의 자료 또는 정보로서 그 보존 및 이용의 효용을 높일 수 있도록 디지털 형태로 제작하거나 처리한 것을 말한다(제2조제5호). “디지털문화콘텐츠”란 문화적 요소가 체화된 디지털콘텐츠를 말한다(제2조제6호).
「콘텐츠산업진흥법」	“콘텐츠”란 부호·문자·도형·색채·음성·음향·이미지 및 영상 등(이들의 복합체를 포함한다)의 자료 또는 정보를 말한다(제2조제1항제1호).
영국 「소비자권리법」 (Consumer Rights Act 2015)	“디지털콘텐츠”는 디지털 형태로 생산, 제공되는 데이터를 말한다(art.2 (9) “Digital content” means data which are produced and supplied in digital form).

3) ‘인공지능 기반 추천 서비스’ 관련 규정

법령(안)명	조문
유럽연합 「디지털서비스법안」 (Digital Services Act)	“추천 시스템”이란 온라인 플랫폼이 온라인 인터페이스에서 수신자가 시작한 검색의 결과 또는 표시된 정보의 상대적 순서 또는 중요도를 결정하는 특정 정보를 수신자에게 제안하기 위해 사용하는 전부 또는 일부 자동화된 시스템을 의미한다(art. 2 (o) ‘recommender system’ means a fully or partially automated system used by an online platform to suggest in its online interface specific information to recipients of the service, including as a result of a search initiated by the recipient or otherwise determining the relative order or prominence of information displayed).
미국 「알고리즘 책임성 법안」 (Algorithmic Accountability Act of 2019)	“자동화된 의사결정 시스템”이란 소비자에게 영향을 미치는 결정을 내리거나 관련된 인간의 의사결정을 용이하게 하는 기계학습, 통계 또는 데이터 처리 및 기타 인공지능 기술에서 파생된 계산 절차를 말한다(Sec. 2. (1) The term “automated decision system” means a computational process, including one derived from machine learning, statistics, or other data processing or artificial intelligence techniques, that makes a decision or facilitates human decision making, that impacts consumers).

4) '추천 서비스 제공자' 관련 규정

법령(안)명	조문
「온라인 플랫폼 기본법안」 (21. 11. 윤두현의원 대표발의)	“온라인 플랫폼 운영자”란 온라인 플랫폼 서비스 제공을 업으로 하는 사업자를 말한다(제2조제3호).
「온라인 플랫폼 이용자 보호에 관한 법률안」 (20. 12. 전해숙의원 대표발의)	“온라인 플랫폼 사업자”란 온라인 플랫폼 서비스를 제공하는 사업자를 말한다(제2조제3호).

5) '이용자' 관련 규정

법령(안)명	조문
「온라인 플랫폼 기본법안」 (21. 11. 윤두현의원 대표발의)	“이용자”란 온라인 플랫폼을 이용하는 고객 또는 공급자를 말한다(제2조제10호).
「온라인 플랫폼 이용자 보호에 관한 법률안」 (20. 12. 전해숙의원 대표발의)	“이용자”란 온라인 플랫폼 사업자가 제공하는 온라인 플랫폼 서비스를 이용하거나 이용하려는 자를 말한다(제2조제5호). “온라인 플랫폼 이용사업자”란 제5호의 이용자 중에 콘텐츠 등에 관한 정보를 제공하거나 콘텐츠 등의 거래를 위하여 온라인 플랫폼 서비스를 이용하거나 이용하려는 사업자를 말한다(제2조제6호).
유럽연합 「디지털서비스법안」 (Digital Services Act)	“서비스 수신자”란 관련 중개서비스를 이용하는 모든 자연인 또는 법인을 말한다(art. 2 (b) ‘recipient of the service’ means any natural or legal person who uses the relevant intermediary service). “소비자”란 거래, 사업 또는 직업적 목적 외의 활동을 수행하는 모든 자연인을 말한다(art. 2 (c) ‘consumer’ means any natural person who is acting for purposes which are outside his or her trade, business or profession).

Ⅲ 핵심 원칙

본 기본원칙은 ▲투명성 ▲공정성 ▲책임성이라는 세 가지 핵심 원칙을 지향한다. 이는 ‘인공지능 윤리’ 차원에서 제시되는 주요 원칙들 가운데 인공지능 기술을 활용한 디지털 미디어의 구조와 기능에 비추어 특별히 강조되어야 하는 원칙에 해당한다. 특히 투명성과 공정성은 인공지능 기술을 적용한 디지털 미디어가 반드시 확보·추구해야 하는 기본 가치인 동시에 인공지능 기술·시스템의 구조적 특성상 훼손되기 쉬운 규범적 가치이기도 하다.

1. 투명성

가. 추천 서비스 제공자는 이용자가 이용 시작 시점에 그 제공 사실을 인지하고 서비스의 내용에 영향을 미치는 주된 요인과 효과를 이해할 수 있도록 필요한 정보를 공개한다.

나. 추천 서비스 제공자는 추천 결과에 대하여 부정적인 요소가 발견되거나 이용자가 불만을 제기하는 경우에는 해당 결과에 도달한 과정을 충실히 설명한다.

1. 규정 취지

첫 번째 핵심원칙인 ‘투명성’은 인공지능 기반 추천 서비스의 신뢰성을 확보하기 위한 기본 전제로서, 인공지능 특유의 기술적 한계(시스템의 불투명성, 우연성 등)를 극복하고 보완하는 가치이다. 인공지능 서비스의 투명성은 인공지능 시스템에 의해 제공받는 서비스의 핵심적인 내용을 이해하도록 정보를 제공하는 것에 초점이 있다. 제1의 핵심 원칙으로 제시된 투명성은 추천 서비스에 관한 ‘정보 접근성’과 추천 결과에 이르는 과정에 대한 ‘설명 가능성’을 포괄하는 광의의 개념으로 규정하고 있다.

‘정보 접근성’ 차원의 투명성은 추천 서비스 이용자가 정보 활용 사실 자체, 추천 기준 및 원리 인지 등을 통해 자기정보통제권과 알 권리를 향상시켜 주체적·합리적 미디어 이용을 가능하게 하는 첫 단추가 된다는 점에서 의미를 지닌다. 또한 ‘정보 접근성’ 차원의 투명성 원칙에 기반을 두어야 권익 보장에 필요한 범위에서 추천

기준을 선택·배제함으로써 맞춤형 서비스를 최적화하는 것도 가능하다.

다른 한편, ‘설명 가능성’ 차원의 투명성은 인공지능 시스템의 추적·검토 가능성을 확보함으로써 제고할 수 있겠으나 아직은 기술적으로 온전한 설명 가능성을 담보하기 어려운 상황이다. 본 기본원칙의 투명성 원칙이 포괄하는 ‘설명 가능성’은 추천 시스템 사용에 따르는 설명책임을 이행하여 불투명성을 사후적으로 보완하는 맥락으로 이해된다.

요컨대, 미디어 추천 서비스가 지향해야 할 투명성 원칙은 이용자에게 서비스에 관한 주요 정보를 제공하여 합리적이고 주체적인 미디어 선택·이용을 지원하는 데 주된 목표가 있으며, 이용자가 서비스 이용 계약의 당사자로서 의사결정, 권리행사, 이익제기 등 전반의 절차에 대등하게 참여할 수 있게 하는 규범적 토대로서 중요한 의의를 지닌다.

2. 주요 내용

위 투명성 원칙 규정에 따르면, 미디어 추천 서비스의 ‘투명성’은 추천 서비스 제공 사실, 추천 서비스의 내용에 영향을 미치는 주된 요인과 효과, 추천 결과에 이르는 과정 등을 알리는 방식으로 실현 가능하다. 추천 시스템의 투명성(설명 가능성)은 기술적으로 완벽히 구현되기는 어렵기 때문에 관련 정보를 제공하거나 설명하는 적극적인 조치를 통해 이를 극복할 수밖에 없다.

먼저, 추천 서비스 제공자는 이용자가 서비스를 이용하는 시점에 인공지능 기반의 추천 서비스가 제공된다는 점을 인식할 수 있도록 적절한 방식으로 정보를 제공할 필요가 있다. 그와 동시에 투명성을 충족하는 첫 번째 방안으로 추천 서비스의 내용을 결정짓는 주된 요인과 효과, 즉 콘텐츠 추천(자동 배열)의 과정과 결과를 이해하는 데 필요한 정보를 공개하는 것을 제시하고 있다. 이는 이용자의 주체적·합리적 미디어 이용을 지원할 수 있는 ‘정보 접근성’ 차원의 투명성과 관련이 깊다.

또 다른 투명성 확보 방안으로 “추천 결과에 대하여 부정적인 요소가 발견되거나 이용자가 불만을 제기하는 경우에는 해당 결과에 도달한 과정을 충실히 설명”할 것을 규정하고 있다. 이는 ‘설명 가능성’ 차원의 투명성과 관련 있으며, 설명책임을 이행함으로써 투명성을 보완하는 경우에 해당한다. 추천 서비스 제공자가 설명책임을

항상 부담하는 것은 아니며, 추천 결과에 대하여 (자율검증 등을 통해) 이용자 권익 등을 침해하는 부정적 요소가 확인되거나 구체적인 불만 제기가 있는 경우에 추천 결과에 도달한 과정을 설명하도록 권고하는 내용을 담고 있다.

2. 공정성

가. 추천 서비스 제공자는 추천 시스템의 편향성으로 인하여 이용자의 권익 또는 미디어의 다양성이 훼손되지 않도록 콘텐츠 자동 배열의 기준 및 결과의 공정성 확보를 위한 조치를 강구한다.

나. 위에 따른 공정성 확보를 위한 조치는 이용자의 선택권 보장 및 불만 처리, 추천 서비스의 사전적·사후적 영향 평가, 지속적 시스템 개선 등 다양한 기술적·관리적 조치를 포괄한다.

1. 규정 취지

두 번째 핵심 원칙인 ‘공정성’은 ‘투명성’과 함께 미디어 추천 서비스가 반드시 충족해야 할 합목적적 가치에 해당한다. 인공지능 시스템에 내재된 편향성으로 인하여 현실에서 나타날 수 있는 다양한 형태의 부작용이 공정성 관점에서 논의되어 왔다. 특히 편향이 존재하는 데이터·알고리즘에 기초하여 도출된 추천 결과가 직간접적으로 불공정행위, 차별행위 등에 준하는 효과를 가져오는 경우가 대표적인 역기능 사례로 다루어지고 있다.

그런데 인공지능 서비스에서 추구하는 ‘공정성’은 분야별로 함축하는 가치에 차이가 있으며, 시스템적·결과적 공정성, 내용적·절차적 공정성 등 다양한 관점에서 양가적 또는 다층적 개념으로 파악할 수 있다. 따라서 공정성 목표는 분야에 따라 결과적 다양성, 통계적 정확성 등 상이하게 설정될 수 있으며 그와 결부된 수단도 다채롭게 제안될 수 있다.

본 기본원칙에서 제시하는 공정성 원칙은 기본적으로 인공지능의 편향에 대응하는 공정성 개념에 따른 것이다. 미디어 추천 서비스의 공정성 가치는 인공지능 시스템의 편향성에 의해 축소될 우려가 있는 미디어 콘텐츠 영역의 ‘다원성’ 및 ‘다양성’을 보장하는 관점에서 강조된다. 이와 더불어 디지털 플랫폼이 미디어 본연의 공적 기능을 분담하는 범위에서는 미디어의 객관성·균형성을 추구하는 차원의 공정성도 일부 고려될 수 있다.

2. 주요 내용

디지털 미디어 영역에서 추구되는 공정성 가치가 다층적인 만큼, 미디어 추천 서비스의 공정성을 효과적으로 구현할 수 있는 단일한 수단을 꼽기는 어렵다. 이에 본 기본원칙은 편향성 방지, 이용자 권익 보호, 미디어 다양성 추구 등이 곧 미디어 추천 서비스의 공정성 목표와 연결된다는 점을 전제로, 추천(자동 배열)의 기준과 결과 양 측면에서 추천 서비스가 공정성을 기할 수 있도록 다양한 합목적적 조치를 강구하는 것을 공정성 원칙으로 제시하고 있다.

본 기본원칙의 주된 수범자인 추천 서비스 제공자는 “추천 시스템의 편향성으로 인하여 이용자의 권익 또는 미디어의 다양성이 훼손되지 않도록 콘텐츠 추천(자동 배열)의 기준 및 결과의 공정성 확보를 위한 조치를 강구”할 것이 권고된다. 이때 공정성 확보를 위한 조치는 “이용자의 선택권 보장 및 불만 처리, 추천 서비스의 사전적·사후적 영향 평가, 지속적 시스템 개선 등 다양한 기술적·관리적 조치를 포괄”하는 것으로 폭넓게 예시하고 있다. 이러한 조치들은 본 기본원칙의 5대 실행 원칙 가운데 이용자의 선택권 보장, 자율검증 실행, 불만 처리 및 분쟁 해결 등과 직접 연결되는 기술적·관리적 수단에 해당한다.

3. 책무성

가. 추천 서비스 제공자는 투명성과 공정성을 제고하기 위하여 본 기본원칙을 준수할 사회적 책무를 지닌다.

나. 추천 서비스 제공자는 추천 서비스 운영 과정에서 기능적 오류·오작동, 현행 법령 위반 등 부정적인 결과가 발생한 경우에는 이를 제거·시정할 책임을 부담하며, 그와 관련한 이용자의 불만 또는 분쟁을 해결하기 위하여 노력하여야 한다.

1. 규정 취지

세 번째 핵심 원칙으로 - 권리·권한에 대응하는 책임(responsibility) 개념과 구분하여 - 윤리적·내적 규율과 신뢰를 중심으로 하는 자율규제 영역에서 사업자에게 부여되는 책무성(accountability)을 제시하고 있다.

현재 상용되는 추천 시스템의 콘텐츠 자동 편집(automated content curation) 기능을 기존 언론사 주도의 편집행위(editing)와 동일시하기는 어렵지만, 추천 서비스가 이용자의 미디어 소비에 유력한 영향을 미친다는 점에서는 추천 서비스 제공자의 사회적·법적 책임에 대한 재고가 필요한 상황이다.

추천 서비스 제공자가 오늘날의 미디어 기능을 분담하는 지위에 있는 이상 그 위상과 수익구조에 상응하는 책임을 부담해야 한다는 의견이 상당하다. 추천 서비스 제공자는 인공지능 시스템을 매개로 하여 미디어 콘텐츠를 추천(자동 배열)함으로써 디지털 미디어 영역을 재구성하므로, 전체 미디어 영역에서 콘텐츠 편집 기능 일부를 대체·보조하는 것으로 파악할 수 있다. 따라서 추천 서비스 제공자는 본 기본원칙의 수범자 지위에서 이용자 권익에 반하는 역기능을 최소화하고 추천 시스템 사용에 따른 사회적 책임을 다할 것이 요구된다.

2. 주요 내용

추천 서비스 제공자는 추천 서비스의 투명성과 공정성이라는 핵심 원칙을 지향하고 이를 충족시키기 위한 규범적 수단들, 즉 5대 실행 원칙을 준수할 사회적 책무를 진다. 또한 추천 서비스 결과에 대하여 광범위한 책임을 다하는 것도 책무성의 내용으로 명시하고 있다.

“추천 서비스 운영 과정에서 기능적 오류·오작동, 현행 법령 위반 등 부정적인 결과가 발생한 경우에는 이를 제거·시정할 책임”을 부담하며, 부정적 결과와 관련한 “이용자의 불만 또는 분쟁을 해결하기 위하여 노력”할 것을 권고한다. 이는 본 기본원칙이 제시하는 실행 원칙들 가운데 ‘자율검증 실행’과 ‘불만 처리 및 분쟁 해결’에 해당하는 내용으로, 본 기본원칙을 준수할 사회적 책무의 내용 중에서도 사후적인 책임 이행에 해당하는 두 가지 항목을 특별히 적시하여 책무성의 범주를 구체화하고 있다.

IV

실행 원칙

세 가지 핵심 원칙을 구현하기 위한 실천적 방안으로서 ▲이용자를 위한 정보 공개 ▲이용자의 선택권 보장 ▲자율검증 실행 ▲불만 처리 및 분쟁 해결 ▲내부 규칙 제정이라는 다섯 가지 실행 원칙을 제시하고 있다.

1. 이용자를 위한 정보 공개

- 가. 추천 서비스 제공자는 인공지능 기반 추천 시스템을 사용하여 해당 서비스를 제공한다는 점을 이용자에게 실시간 문자 전송, 홈화면 알림 등의 방법으로 통지 또는 표시한다.
- 나. 추천 서비스 제공자는 이용자의 소비·검색 이력, 콘텐츠별 조회 수, 연령, 성별 등 콘텐츠 자동 배열에 적용되는 주요 사항을 홈화면 알림, 팝업창 고지, 약관 명시 등 이용자가 이해하기 쉽고 명확한 방법으로 공개한다.

1. 규정 취지

첫 번째 실행 원칙으로 제시된 ‘이용자를 위한 정보 공개’는 인공지능 기반 추천 서비스의 제공 사실 및 적용 기준에 대하여 이용자에게 공개함으로써 이용자의 알 권리, 미디어 접근권, 자기정보통제권 등 주요한 기본권을 보장하는 데 목적이 있다. 이는 핵심 원칙 중 ‘투명성’을 확보하는 것과 직결되며, 그 중에서도 ‘정보 접근성 차원의 투명성’이 정보 공개 원칙의 주축이 된다. 즉, 미디어 이용 주체이자 정보주체의 지위에 있는 이용자가 추천(자동 배열)된 미디어 콘텐츠 영역의 성격을 이해하고 미디어 콘텐츠 소비·공급 활동을 주체적으로 영위할 수 있게 만드는 가장 기본적인 조치가 ‘정보 공개’이다.

추천 서비스 제공자가 콘텐츠 추천 기준 또는 방식에 관한 정보를 공개하는 것은 서비스 이용 계약 당사자인 서비스 제공자와 이용자 간의 정보 격차를 좁혀 양자의 지위 불균형을 완화하고 나아가 추천 서비스의 신뢰성을 제고하기 위한 규범적 수단으로 이해할 수 있다. 미디어 추천 서비스를 제공받는 이용자로 하여금 미디어

콘텐츠 추천 기준 및 방식에 관한 정보에 접근할 수 있도록 함으로써 정보 불균형에 따른 불이익을 방지하고 이용자의 권익을 보장하는 것에 정보 공개의 주된 취지가 있다.

2. 주요 내용

추천 서비스의 투명성을 충족시키기 위한 정보 공개 대상으로 추천 서비스 제공 사실과 콘텐츠 추천(자동 배열) 기준에 해당하는 데이터라는 점을 명확히 밝히고 있다. 즉, 이용자에게 인공지능 기반 추천 시스템을 사용하여 해당 서비스를 제공한다는 점을 통지·표시하거나 이용자의 소비·검색 이력, 콘텐츠별 조회 수, 연령, 성별 등 콘텐츠의 추천(자동 배열)에 적용되는 주요 사항을 쉽게 확인 가능한 상태로 두는 것이 ‘정보 공개’의 방식이다. 구체적인 정보 공개 방법으로 실시간 문자 전송, 홈페이지 알림, 팝업창 고지, 약관 명시 등을 언급하고 있으나 이는 예시에 지나지 않는다. 이용자가 관련 정보를 쉽게 접하고 이해할 수 있는 방법이라면 그 기술적 형태를 불문한다.

이러한 정보 공개 조치는 공개 대상이 되는 정보가 무엇인지에 따라 현행법상 보호 수단과 상충하거나 중첩될 여지도 있다. 예컨대 ① 알고리즘 자체를 공개하는 것은 추천 서비스 제공자의 영업비밀을 보호하는 법률상 목적·수단과 정면으로 충돌하므로, 알고리즘은 본 기본원칙에 기초한 정보 공개 대상이 아니다. 그리고 ② 추천 서비스 제공자는 회원가입 단계에서 수집한 이용자의 개인정보를 콘텐츠 추천(자동 배열) 기준으로 반영하기도 하는데, 이 경우에는 정보주체인 이용자가 자기정보통제권을 실현하는 차원에서 개인정보 수집·활용 사실을 명확히 인지할 수 있어야 하므로 정보 공개가 강력히 요구된다. 물론 특정인을 식별할 수 있는 개인정보를 활용하지 않은 경우라고 하더라도, 추천 서비스가 이용자의 권익에 직접적인 영향을 미치는 경우라면 마찬가지로 추천 서비스 제공 사실을 알릴 필요가 있다.

i 참고 | 관련 법령(안) 규정

1) 투명성 확보(정보 제공, 설명 가능성 등) 관련 사례

구분	내용
유럽연합 「디지털서비스법안」 (Digital Services Act)	• ‘대규모 온라인 플랫폼 사업자’는 추천 시스템에 사용되는 주요 매개변수와 함께 해당 매개변수를 수정하거나 변수에 영향을 줄 수 있는 방법 등을 쉽고 명확하게 이해하고 접근할 수 있는 형태로 약관에 명시하여야 함(제29조) ※ 대규모 온라인플랫폼 사업자: 서비스 이용자의 요청에 따라 정보를 저장하고 일반 대중에게 정보를 배포하는 호스팅 서비스 제공자 중 EU 내 월평균 4,500만 이상의 활성 서비스 수신자(active recipients of the service)에게 서비스 제공하는 온라인플랫폼
유럽연합 「인공지능법안」 (Artificial Intelligence Act)	• 사용자에 대한 정보 제공 및 투명성 보장(제13조) - 시스템이 투명하게 설계·개발되어야 하며, 사용자가 이해할 수 있는 간결, 완전, 명확한 정보를 포함하는 지침을 제공해야 함 ※ 정보 제공 내용: ① 공급업체의 신원 및 연락처 등, ② 고위험 AI 시스템 성능의 특성, 기능 및 한계, ③ 초기 적합성 평가 시점에 공급업체가 설정했던 고위험 AI 시스템 및 성능에 변경이 생긴 경우 이에 대한 정보, ④ 인간의 감독(통제)에 관한 사항, ⑤ 유지보수 및 관리조치 등 • 특정 AI 시스템에 대한 투명성 의무(제52조) - AI 시스템이 개인과 상호작용하도록 구성되어 있는 경우 공급업체는 AI 시스템과 상호작용하고 있음을 알리는 방식으로 설계·개발하여야 함 - 감정인식시스템 또는 생체분류시스템(biometric categorisation system) 사용자는 범죄수사나 법률에 의한 활용을 제외하고 시스템에 노출된 자연인에게 시스템 운영에 대하여 알려야 함 - AI 시스템이 딥페이크(deep fake)로 보이는 이미지, 오디오, 비디오 콘텐츠를 생산하거나 조작하는 AI 시스템의 사용자는 콘텐츠가 인위적으로 생성되거나 조작되었음을 공개하여야 함

구분	내용
영국 ICO ‘인공지능 기반 의사결정에 대한 설명’ (Explaining decisions made with AI)	• 인공지능의 설명 가능성 및 투명성을 확보하기 위하여 공개해야 하는 정보 ① 알고리즘 또는 모델이 사용되는 의사결정 과정의 전반적인 세부 사항 ② 의사결정 과정 내에서 알고리즘 또는 모델이 어떻게 사용되는지에 관한 사항 ③ 알고리즘 또는 모델 자체와 그 개발 방법에 대한 개요 ④ 관련 법률에 따라 수행된 데이터 보호, 평등, 인권에 대한 영향 평가를 포함하여 전반적인 의사결정 과정이 설계된 방식에 대한 근거 설명
일본 「특정 디지털플랫폼 투명성 및 공정성 향상에 관한 법률」 (特定デジタルプラットフォームの透明性及び公正性の向上に関する法律)	• 이용자에게 대한 정보 공개 의무 규정(제5조제2항제2호) ① 검색 표시 순위 결정에 이용되는 주요 사항(상품 가격이나 소비자 후기 등이 해당되며, 알고리즘의 공개는 제외) ② 해당 특정 디지털 플랫폼 제공자가 상품 등 구입 데이터(일반 이용자에 의한 상품 등과 관련된 정보의 검색이나 열람 또는 상품 등의 구입과 관련된 데이터)를 취득하거나 사용하는 경우의 해당 상품 등 구입 데이터의 내용 및 그 취득 또는 사용에 관한 조건 ③ ①, ② 외에 일반 이용자에 대한 해당 특정 디지털 플랫폼의 제공조건 중 개시하는 것이 특히 필요한 것으로서 경제산업성령으로 정하는 사항
미국 NIST ‘설명가능한 인공지능의 4가지 원칙’ (Four Principles for Explainable AI)	• (Explanation) 인공지능 시스템은 그들의 결과에 대해서 증거 또는 이유를 전달하여야 한다. • (Meaningful) 인공지능 시스템은 의미 있고 이해할 수 있는 설명을 개별 사용자에게 제공하여야 한다. • (Explanation Accuracy) 설명은 반드시 결과물을 생성하는 인공지능 시스템의 과정을 정확하게 반영하여야 한다. • (Knowledge Limits) 인공지능 시스템은 그 출력값에 대한 충분한 신뢰도를 충족하거나 그렇게 설계되었다는 조건하에서만 운영되어야 한다.

구분	내용
방송통신위원회 '온라인 맞춤형 광고 개인정보보호 가이드라인'	• 온라인 맞춤형 광고 개인정보보호 원칙 中 - (행태정보 수집·이용의 투명성) 광고 사업자 또는 매체 사업자는 이용자가 온라인상에서 자신의 행태정보가 수집·이용되는 사실 등을 쉽게 알 수 있도록 안내 방안을 마련하여야 한다.
개인정보보호위원회 '인공지능 개인정보보호 자율점검표'	• AI 관련 개인정보보호 6대 원칙 中 - (투명성) 개인정보 처리 내역을 정보주체가 알기 쉽게 공개한다. ※ 관련 자율점검사항 - AI 개발·운영 시 개인정보처리에 관한 구체적 사항을 개인정보처리방침에 포함·작성하여 홈페이지 등에 공개하는가? - AI 개발·운영 과정에서 자율적인 개인정보 보호활동을 적극적으로 수행하는가?

2) 국내외 정보 공개 사례 예시

기업명	서비스 유형	정보 공개 범위
네이버	맞춤형 광고	• 수집하는 행태정보의 종류 및 수집 방법 • 광고에서 행태정보를 활용하는 목적 • 행태정보 보유, 이용기간 및 파기 • 이용자 통제권 행사 방법 • 이용자 피해구제 방법
유튜브	동영상 콘텐츠 추천	• 이용자 계정 및 설정 관리를 통해 추천 서비스에 사용되는 본인의 시청 기록 및 검색 기록 확인 및 삭제 가능
클래스팅 (Classting)	맞춤형 교육	• 수집하는 정보의 항목 • 맞춤형 서비스 제공 및 광고 등 정보처리의 목적 • 개인정보의 파기 • 이용자 및 법정대리인의 권리행사 방법
넷플릭스	동영상 콘텐츠 추천	• 추천 시스템이 고려하는 데이터 항목 • 추천 결과에 따른 콘텐츠 배열 방법

기업명	서비스 유형	정보 공개 범위
뉴욕 타임즈	맞춤형 정보제공	• 개인화의 수행 목적 및 방법 • 개인화를 위해 사용하는 데이터 항목
BBC	맞춤형 정보제공, 맞춤형 영상 콘텐츠	• 미디어 추천 방법 • 개인화 설정 및 해제 방법 • 개인화의 수행 목적 및 방법 • 개인화를 위해 사용하는 데이터 항목
라쿠텐	맞춤형 정보제공	• 일본 ‘특정 디지털 플랫폼 투명성 및 공정성 향상에 관한 법률’에 따라 기본적 서비스와 검색순위 결정의 기본적 사항 등 공개
야후재팬	맞춤형 정보제공	• 일본 ‘특정 디지털 플랫폼 투명성 및 공정성 향상에 관한 법률’에 따라 추천 순서, 표시 순위 등 공개

2. 이용자의 선택권 보장

가. 추천 서비스 제공자는 미디어 콘텐츠를 소비하는 이용자가 본인의 선호와 필요에 따라 합리적으로 미디어 콘텐츠에 접근하고 이용할 수 있도록 선택권을 보장한다.

나. 추천 서비스 제공자는 위의 선택권을 보장하기 위하여 콘텐츠 자동 배열에 적용된 주요 사항에 대한 선택·변경 기능을 이용자가 이용하기 쉬운 방식으로 제공한다.

다. 선택·변경 기능은 콘텐츠 유형, 소요 시간, 비용, 가용 기술 등을 고려하여 합리적으로 실행 가능한 범위에서 제공할 수 있다.

1. 규정 취지

두 번째 실행 원칙으로 제시된 ‘이용자의 선택권 보장’은 미디어 콘텐츠를 소비하는 이용자가 서비스 이용 요건 설정에 실질적으로 동의·관여할 수 있도록 함으로써, 미디어 플랫폼과 이용자 간에 실질적인 상호 작용이 가능한 미디어 생산·소비 환경이 조성될 수 있도록 하는 데 목적을 둔다. 여기서 말하는 ‘선택권’은 미디어 소비자로서의 이용자가 본인이 원하는 콘텐츠에 자유롭게 접근하여 이용할 수 있는 권리에서 파생된 것으로 일종의 수단적 권리로 이해할 수 있다.

추천 서비스를 통해 미디어 콘텐츠를 소비하는 이용자는 추천 시스템을 직접 설계·운영·관리하는 사업자와 비교할 때 지극히 수동적인 지위에 놓이게 된다. 이용자가 쉽게 접근할 수 없는 시스템 내에서 이용자 데이터가 분석·활용되고 인공지능 알고리즘에 의하여 미디어 콘텐츠 영역이 자동적으로 구성됨에 따라 이용자의 합리적 선택 가능성 및 정보 통제 가능성은 대폭 축소될 수밖에 없기 때문이다. 더욱이 추천 시스템을 개발·운영하는 플랫폼 사업자 내지 추천 서비스 제공자와 이용자 간에 정보 불균형이 현저한 상황이라면 이용자 스스로 콘텐츠 추천(자동 배열)과 관련한 의도적 또는 결과적 편향성을 판단하기 어렵기도 하다.

따라서 이용자가 주체적으로 서비스 이용에 개입할 수 있는 권리를 보장하고, 이용자가 스스로 정보의 편향을 경계하고 콘텐츠 추천(자동 배열) 결과의 긍정적 영향과 부정적 영향을 구분하여 자신의 권익을 보호하기 위하여 - ‘이용자를 위한 정보 공개’라는 규범적 조치와 더불어 - 일정한 범위의 선택·변경 기능이 기술적으로 제공될 필요가 있다.

2. 주요 내용

본 실행 원칙에서는 앞서 정보 공개 대상이 되었던 콘텐츠 추천(자동 배열)에 적용된 주요 사항 — 이용자의 소비·검색 이력, 콘텐츠별 조회 수, 연령, 성별 등 추천 서비스 제공을 위해 수집·활용된 데이터 항목 — 을 이용자 본인의 선호와 필요에 따라 쉽게 선택·변경할 수 있는 기능을 제공할 것을 권고한다. 이때 선택권의 주체가 되는 이용자는 추천 서비스를 제공하는 플랫폼에서 ‘미디어 콘텐츠를 소비하는 이용자’에 한한다. 따라서 플랫폼 서비스를 이용하여 미디어 콘텐츠를 직접 공급하는 이용사업자는 선택권 보장 원칙의 보호 대상이 아니다.

선택 및 변경 기능을 제공하는 것은 서비스의 유형과 특성 등에 따라 각기 다른 결과를 발생시킬 수 있고, 추천 서비스 제공자에게 필요 이상의 책임을 부과하는 것으로 귀결될 우려도 있다. 예컨대, 추천 서비스 제공자에게 ‘추천 기준에 대한 선택·변경 기능’을 추가하는 것이 경우에 따라서는 영업활동에 과도한 부담으로 작용하거나 해당 추천 서비스의 독자성·특이성을 담보하기 위해 필수적으로 반영하는 데이터 항목에 대해서는 — 개인의 권익 침해가 발생하였거나 발생 개연성이 현저한 경우가 아니라면 — 이를 배제(옵트아웃, opt-out)할 수 있는 선택권을 보장하는 것은 오히려 비례성을 상실한 조치가 될 여지도 있다.

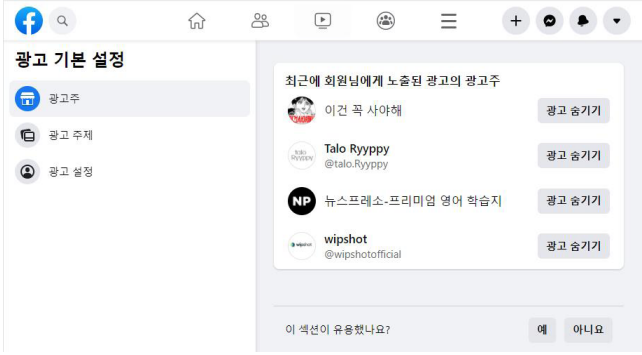
이처럼 미디어 이용자에게 ‘추천 기준에 대한 선택권’을 보장하는 것은 추천 서비스 제공자의 영업활동의 자유에 우선하는 이용자의 규범적 보호 가치가 인정되는 경우에 한하여 권고된다고 봄이 타당하다. 이러한 이유로 본 실행 원칙에서는 선택권의 주체를 제한하였을 뿐만 아니라, 콘텐츠의 유형, 선택·변경 기능을 제공하기 위해 드는 시간, 비용, 기술 등을 고려하여 합리적으로 실행 가능한 범위에서 제공하면 족하다는 점을 분명히 하고 있다.

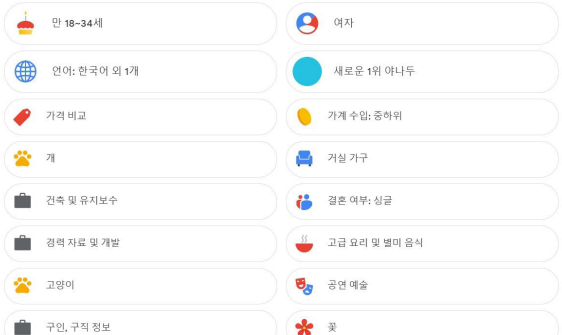
i 참고 | 관련 법령(안) 규정

1) 이용자의 선택권 보장 관련 사례

구분	내용
온라인 플랫폼 이용자 보호에 관한 법률안 (20. 12. 전혜숙의원 대표발의)	• 대규모 온라인 플랫폼 사업자에 대한 강화된 규제로서 정보공개의무 등 규정 - 제11조(노출 기준의 공개 등) ① 대규모 온라인 플랫폼 사업자는 검색결과, 추천 등을 결정하는 요소 등 콘텐츠 등의 노출 방식 및 노출 순서를 결정하는 기준을 공개하여야 한다. ② 대규모 온라인 플랫폼 사업자는 제1항의 기준 중 개인화된 기준의 적용 여부 및 내용과 범위 등을 이용자가 선택할 수 있도록 하여야 한다.
유럽연합 「디지털서비스법안」 (Digital Services Act)	• ‘대규모 온라인 플랫폼 사업자’는 추천 시스템에 사용되는 주요 매개변수와 함께 해당 매개변수를 수정하거나 변수에 영향을 줄 수 있는 방법 등을 쉽고 명확하게 이해하고 접근할 수 있는 형태로 약관에 명시하여야 함(제29조) • 추천 시스템의 주요 매개변수와 이를 수정(modify)하거나 영향을 줄 수 있는 옵션을 명확하고 이해하기 쉬운 방식으로 약관에 명시(GDPR상의 프로파일링에 따르지 않는 옵션을 하나 이상 포함해야 함)하여야 하며, 복수의 옵션이 부여된 경우, 표시된 정보의 순서를 결정하는 추천 시스템 각 매개변수에 대하여 언제든지 원하는 옵션을 간편하게 선택·수정하는 기능을 제공하여야 함
영국 CDEI ‘데이터 기반 온라인 타기팅’ (Data-driven online targeting)	• 본 보고서는 온라인 타기팅을 ‘사람에 대한 정보를 분석하고 그들의 온라인 경험을 맞춤화하는 데 사용되는 다양한 관행’이라고 정의하면서, 이러한 온라인 타기팅에 대한 정부의 규제 권고 사항을 제안 - 정부가 온라인 타기팅 시스템을 이용하는 플랫폼 사용자 등을 규제함에 있어 정보 제공 및 온라인 타기팅 시스템에 대한 제어를 핵심으로 할 것을 권고 - 온라인 플랫폼이 ‘fairness by design’ 의무를 이행하는 차원에서 정보에 입각한 소비자의 선택을 촉진하고, GDPR에 따른 설계 요구 사항에 따라 동의 및 개인정보보호 정책을 구상할 것을 제안

2) 국내외 이행 사례 예시

기업명	내용
페이스북 (現 메타)	• 페이스북은 개인 설정 내 광고 기본 설정에서 개인에게 특정 광고가 도달하기 위해 사용된 카테고리(키워드)를 확인할 수 있도록 하여 광고 배열 기준을 이용자가 확인할 수 있도록 함 • 해당 키워드 여부를 사용하지 않을 수 있도록 설정할 수 있도록 하여 배열 기준의 변경을 가능하게 함 • 특정 개인에게 노출된 광고의 광고주를 확인하여 특정 광고를 더 이상 보지 않을 수 있도록 설정하거나 특정 광고 주제(주류, 육아, 반려동물)의 표시 빈도를 조절할 수 있도록 하여 개인의 선호도에 따른 접근 선택권을 보장하고 있음  〈그림 1〉 페이스북 내 노출된 광고에 대한 설정창

구분	내용
구글	•구글 대시보드 내 광고 설정에서 ‘광고 개인 최적화 방법’을 통하여 어떤 키워드를 기준으로 광고가 제공되는지에 대하여 표시하여 광고 배열 기준을 명시 광고 개인 최적화 방법 광고는 Google 계정에 추가한 개인정보, Google과 파트너 관계를 맺은 광고주의 데이터, Google이 추정한 관심분야를 바탕으로 표시됩니다. 자세히 알아보거나 환경설정을 업데이트하려면 요소를 선택하세요. 표시되는 광고 제어하는 방법 알아보기  〈그림 2〉 구글 광고 설정 내 개인 최적화 방법 • 각 키워드를 클릭하면 광고 추천 시 해당 키워드 사용 여부를 확인 및 변경할 수 있는 옵션을 제공함으로써 이용자 선호도(선택)에 따른 광고 콘텐츠 접근을 보장

3. 자율검증 실행

가. 추천 서비스 제공자는 추천 서비스가 이용자에게 미치는 영향을 평가하고 조정함으로써 위험성을 상시 관리할 수 있도록 자율검증 체계를 마련한다. 자율검증 실행의 예시는 다음과 같다.

- ① 추천 서비스가 유발할 수 있는 위험을 사전에 예측·평가하여 위험성의 정도에 비례하는 효과적인 통제수단을 갖춘다.
- ② 서비스 운영 과정에서 추천 서비스가 이용자에게 미치는 영향에 대하여 지속적으로 확인 및 관리한다.
- ③ 추천 시스템의 개발·적용 과정을 접근성, 가독성, 검증가능성 등을 충족하는 방법으로 기록하고, 이를 일정 기간 보관하여 추천 서비스에 대한 자율검증 시 활용할 수 있도록 한다.

나. 자율검증은 위험 수준에 따라 차등적인 체계를 구성하여 실행하며, 소요 시간, 비용, 가용 기술 등을 고려하여 합리적으로 실행 가능한 범위에서 운영한다.

1. 규정 취지

‘자율검증’ 방식의 자율규제는 인공지능 시스템의 위험성을 관리하는 최선의 방법으로 제안되고 있다. 이는 인공지능 기반 미디어 추천 서비스를 개발·운영하여 영리활동을 하는 사업자의 자율성을 존중하면서도 이용자에 미치는 영향, 특히 부정적 영향을 제어하는 데 적합한 수단이 될 수 있다. 본 실행 원칙은 추천 서비스 제공자의 미디어 콘텐츠 공급 방식에 따라 이용자의 미디어 콘텐츠 소비 영역이 양적·질적으로 달라질 수 있다는 점을 고려하여, 콘텐츠 공급 시스템을 상시 관리함으로써 미디어 콘텐츠를 적정하고 균형 있게 제공하도록 하는 데 목적이 있다.

법적 규제를 통해 콘텐츠 공급 방식을 제한하는 것은 추천 서비스 제공자의

영업활동의 자유를 지나치게 침해할 우려가 있으며, 혁신적이고 이용자 친화적인 서비스를 개발하는 유인을 해칠 수 있다. 더욱이 인공지능 알고리즘 기반의 추천 서비스는 사업자의 독자적 기술 및 설계에 기초한 기술집약적 서비스로서 매우 복잡한 기술 구조를 갖추고 있기 때문에 외부에 의한 통제가 사실상 어려운 측면이 있다.

따라서 본 기본원칙이 제시하는 원칙들을 관철하는 효과적인 방식은 추천 서비스 제공자 스스로가 추천 서비스의 내용이 이용자에게 미치는 영향을 평가하고, 이를 지속적으로 관리할 수 있는 구조를 갖추는 일이다. 추천 서비스 제공자가 자율검증 체계를 마련하고 운영하는 일은 미디어 서비스의 자율성을 지키면서도 상시적인 위험성 관리를 통해 이용자에게 미치는 영향을 효과적으로 통제할 수 있는 가장 실효적인 방안에 해당한다.

2. 주요 내용

본 기본원칙은 기본적으로 자율검증 체계의 구성 방식과 내용 역시 추천 서비스 제공자의 자율에 맡기고 있으나, 일정한 준거를 제시한다는 측면에서 자율검증 실행 방법의 예시를 제공하고 있다. 먼저 미디어 분야에서의 추천 서비스 이용이 야기할 수 있는 위험을 사전에 예측·평가함으로써, 적절한 통제 수단을 갖추 수 있도록 하는 기반을 마련하도록 권고하고 있다. 자율검증은 어디까지나 ‘위험(성) 기반의 차등적 규율’에 기초한 자율규제 수단이다. 서비스 분야, 추천 콘텐츠의 유형, 사용되는 데이터 등에 따라 이용자 권익 침해의 위험성에 차이가 있기 때문에 자율검증에 수반하는 기술적·관리적 통제 수단도 위험성 수준에 따라 차등을 두는 것이 합리적이다. 이러한 관점에서 자율검증 실행체계를 아래와 같이 예시하고 있다.

① 위험성 사전 예측·평가 및 예방적 통제

“추천 서비스가 유발할 수 있는 위험을 사전에 예측·평가하여 위험성의 정도에 비례하는 효과적인 통제수단을 갖춘다.”

이 항목은 추천 서비스 개발 단계에서 미리 위험성을 평가할 필요가 있음을 강조한 것으로, 사전예방적 성격을 갖는 통제 수단 역시 위험성에 비례하는 적합한 수단이면

충분하다. 분야를 막론하고 인공지능 시스템을 사용하는 서비스 분야의 자율규제 내용으로 시스템 개발 단계의 '위험성 예측·평가'는 공통된 위험 관리 내지 위험 통제 수단으로서 중요한 의미를 갖는다.

위험에 대한 예측과 평가는 다음과 같은 절차를 거쳐 수행할 수 있다.

- 1단계: 사전 시뮬레이션을 통해 추천 서비스가 제공하는 미디어 콘텐츠 예측
- 2단계: 예측 결과에 따라 사용 알고리즘의 위험성 정도를 분류
- 3단계: 위험성의 정도에 비례하는 적합한 통제수단 제시

미디어 분야에서 특히 주안점을 두어야 할 위험에 대한 평가 요소로 다음과 같은 것들을 예시할 수 있다.

- 알고리즘(콘텐츠 추천(자동 배열) 기준)에 편향성이 존재하는지의 여부
- 추천 서비스의 제공을 위해 사용하는 데이터를 이용자에 맞게 적절하게 구성하고 있으며, 양적으로 충분한 데이터를 사용하는지 여부
- 뉴스 콘텐츠 제공 시 자동 배열의 주요 기준 및 사용 데이터의 정확성·최신성 등
- 알고리즘의 적용을 위한 준비 절차가 적법하게 실행되었는지 여부
- 알고리즘에 의한 자동추천 서비스의 작동이 소비자에게 충분히 설명되었는지 여부
- 알고리즘의 사용이 서비스의 품질을 향상시키는지 여부
- 알고리즘 오용에 대비한 안전장치를 갖추고 있는지 여부
- 자동 추천 알고리즘의 사용이 기타 법적 요건과 부합하는지 여부

② 서비스 운영 중 지속적인 위험성 관리

“서비스 운영 과정에서 추천 서비스가 이용자에게 미치는 영향에 대하여 지속적으로 확인 및 관리한다.”

두 번째 항목으로, 서비스 운영 과정에서 지속적으로 자율검증을 실행할 것을 권고한다. 미디어 추천 서비스 결과에 대한 모니터링, 피드백, 영향 분석 등을 거쳐 발생 가능한 위험 요소가 누적·결합되는 것을 막을 수 있다. 서비스 운영 과정에서 지속적으로 확인 및 관리하는 방안은 다음 세 가지로 요약된다.

- 자동 추천 시스템 사용에 따른 결과 주시 및 영향 분석
- 명백한 부작용이 확인되거나 불만사항으로 반복 접수되는 경우 즉각 시정 또는 서비스 중단
- 알고리즘 시스템 자체 또는 사용 결과의 위험성에 부합하는 기술적·관리적 조치 이행

③ 검증 가능한 형태의 기록 및 그 보관·활용

“추천 시스템의 개발·적용 과정을 접근성, 가독성, 검증가능성 등을 충족하는 방법으로 기록하고, 이를 일정 기간 보관하여 추천 서비스에 대한 자율검증 시 활용할 수 있도록 한다.”

자율검증 대상이 되는 추천 시스템의 위험성을 예측·평가·관리하기 위해서는 추천 시스템의 구성 요소, 구현 과정 등 제반 사항이 접근·독해·검증 가능한 형태로 보관되어야 한다. 따라서 자율검증 체계의 중요한 항목으로 추천 시스템의 개발·적용 과정을 기술적으로 검토 가능한 형태로 기록하고 이를 보관하여 자율검증 시 활용하도록 하는 방안을 제시하고 있다.

추천 서비스 제공자는 자율검증 실행의 방법으로 내부 전담조직을 운영하거나 사업자단체 및 외부 전문기구에 검증을 위탁하는 등 다양한 방식을 선택할 수 있다. 내부 전담기구를 두는 경우에는 그 책임자에게 해당 업무를 독립적으로 처리할 수 있는 지위와 권한을 부여하여 자율규제의 완결성을 확보할 수 있다. 이와 달리 외부 전문기구에 검증을 위탁하는 경우에는 전문기구가 추천 시스템의 작동 원리 및 결과를 조사, 이해, 검토할 수 있도록 관련 정보 및 기록을 전달하는 것이 불가피하므로, 이때 기업의 기밀 사항 유출이 발생하지 않도록 충분한 안전장치를 마련해 두어야 할 것이다.

또한 추천 서비스 제공자에게 과도한 부담을 지우지 않고 실행할 수 있는 한도 내에서 자율검증 체계가 마련되고 이행되면 충분하다고 보는 것이 자율규제의 본질에 비추어 타당하다. 자율검증을 통해 기대할 수 있는 피해 예방·완화 효과에 비하여 시간, 비용, 기술 등의 소요 정도가 과도한 경우 자율검증이 경제적·합리적 수단이라고 보기 어렵다. 특히 미디어 영역에서 콘텐츠 추천(자동 배열)으로 인해 발생할 수

있는 위험(성)은 - 추천 알고리즘 설계·적용 행위 자체가 위법한 것이 아닌 이상 - 주관적이고 불확실한 경우가 많다. 그러므로 미디어 추천 서비스에 대한 자율검증은 사전적·상시적으로 위험(성)을 예방·관리함으로써 추천 시스템의 위법성을 자발적으로 통제하는 것에 주된 목적이 있다고 볼 수 있다.

i 참고 | 국내외 사례

1) 자율검증 관련 법규·지침

구분	내용
유럽연합 「디지털서비스법안」 (Digital Services Act)	• 대규모 온라인 플랫폼의 리스크 평가 의무를 규정(제26조) - 대규모 온라인 플랫폼은 EU 내 서비스 기능 및 이용에 따른 중요한 시스템 리스크에 대하여 연 1회 이상 식별·분석·평가해야 함 - 리스크 평가 시, 자사의 추천 시스템이 불법 콘텐츠 확산 등 시스템 리스크에 어떤 영향을 미치는지 고려해야 함
독일 「미디어 국가 협약」 (Staatsvertrag zur Modernisierung der Medienordnung in Deutschland, Medienstaatsvertrag)	• 미디어 서비스 제공자에게 콘텐츠와 정보 제공에 관하여 일정한 자율검증 의무를 규정(제19조 제1항) - 완전히 또는 부분적으로 사진 또는 문서 형태의 인쇄물이 정기적으로 제작되는, 언론형식으로 편집된 텔레미디어 서비스”(제1문)와 “영업상 정기적으로 보도 또는 정치적 정보를 제공하는 언론형식으로 편집된 텔레미디어”(제2문)에 대하여 공인된 저널리즘 원칙, 예컨대 언론윤리강령 규정을 준수할 것을 의무화함 - 인터넷 언론보도 사이트, 특정 인터넷 블로그 등은 저널리즘 원칙의 준수 하에 보도의 유포 전 그 내용과 출처, 사실 여부를 주의 깊게 검토해야 함
유럽연합 ‘신뢰 가능한 인공지능을 위한 자체평가목록’ (The Assessment List for Trustworthy Artificial Intelligence for self assessment)	• 인공지능 시스템과 관계되는 조직의 내외부에서 인공지능 시스템에 대해 평가할 수 있도록 하는 항목을 제시 • 인공지능 시스템의 설계자와 개발자, 데이터 기술자, 조달 책임자 또는 전문가, 인공지능 시스템을 활용하는 직원, 법무담당직원 및 경영관리자 등의 자율검증 항목 제시

구분	내용
유럽연합 '신뢰 가능한 인공지능을 위한 자체평가목록' (The Assessment List for Trustworthy Artificial Intelligence for self assessment)	• 인공지능 시스템 관련 결정 사항 - 자기학습 내지 자동화된 시스템인가 - 프로세스 도중에 인간이 핵심이 되어 검증이 가능한가 (Human-in-the-Loop) - 프로세스 전반에 대한 인간의 총괄적 검증이 가능한가 (Human-on-the-Loop) - 인간의 명령에 따라 검증이 가능한가(Human-in-Command) • 검증을 담당하는 인간에게 감시 집행 방법에 관한 전문적인 훈련이 제공되었는가? • 인공지능 시스템의 이용자나 대상에 대한 의도되지 않은 반대 효과에 대한 탐지 및 대응 메커니즘을 갖추고 있는가? • 필요한 경우 작동을 중지시킬 수 있는 "정지버튼" 또는 절차를 보장하고 있는가? • 인공지능 시스템의 자기학습 또는 자동화된 성격을 반영한 특수한 감시 및 통제 수단을 부여하고 있는가?
캐나다 「정책(의사) 결정에서의 머신러닝 사용에 관한 지침」 (Directive on the Use of Machine Learning for Decision-Making)	• 공공 부문에서의 인공지능 사용에 관한 지침에 해당 • 지침상 자율검증 책무 - 결정 절차 개시 전 사용된 데이터의 편향성 여부 등에 대해 검증 및 모니터링 - 사용하는 데이터가 정확한 것인지, 최신인지, 정보관리 정책과 프라이버시법에 부합하는지의 여부 검증 - 사전에 알고리즘의 영향성 정도에 대해 전문가 검증 - 자동 결정 시스템의 사용이 법적 요건들과 부합하는지 검증

2) 국내외 이행 사례 예시

구분	내용
독일 연방정부 「데이터윤리위원회 의견서」 (Gutachten der Datenethikkommission, 2019)	■ 알고리즘 시스템의 리스크에 대한 자율검증 방법과 위험 판단 척도를 제시 • 리스크 관리에 적합한 규제 방식 권장 - 애플리케이션 개발, 구현, 평가 및 수정 단계에서의 제반 요소(각 단계에 참여하는 모든 인간 행위자 포함)를 이루는 각종 사회·기술 시스템을 고려해야 함 - 알고리즘 시스템에 대한 규제 수단 및 요구 사항으로 시스템 수정 및 제어 톨, 투명성에 대한 지침, 결과의 설명 가능성 및 추적 가능성, 사용에 대한 책임 규정 등을 포함해야 함 - 알고리즘 시스템의 손해가능성에 대해서는 포괄적 모델을 적용하여 개인 또는 사회에 대하여 일괄적으로 결정하도록 권장
	◎ 손해가능성 정도에 따라 5단계로 구분 - 1단계: 손해가능성이 거의 또는 전혀 없는 시스템 ⇒ 알고리즘 요소가 없는 제품에도 적용되는 일반적인 품질 요구 이상의 별도의 조치 필요 없음 - 2단계: 손해가능성이 약간 있는 시스템 ⇒ 형식적·실질적 요건(투명성 의무, 리스크 영향평가 공개의무) 및 통제 절차(감독기관에게 공개의무, 사후 통제, 감사절차) 필요 - 3단계: 손해가능성이 일반적이거나 분명한 시스템 ⇒ 추가적인 사전 허가 절차 필요 - 4단계: 손해가능성이 상당한 시스템 ⇒ 알고리즘 시스템의 계산 과정에 영향을 주는 요소, 요소 간 중요도 등을 포함하는 더욱 강화된 투명성 의무 부과 및 라이브 인터페이스를 통한 감독기관의 지속적인 통제가능성 확보 필요 - 5단계: 손해가능성이 용인될 수 없는 시스템 ⇒ 알고리즘 시스템 전면적·부분적 금지

구분	내용
전기전자기술자협회(IEEE) ‘윤리적 연구 및 설계를 위한 방법론’ (Methods to Guide Ethical Research and Design, 2019)	■ ‘자동화된 지능형 시스템(Autonomous Intelligent System; A/IS)’의 개발에 수반되어야 할 자율검증 체계를 제안 * Autonomous Intelligent System(A/IS): 인공지능(AI)이라는 표현을 대체하여 IEEE가 사용하는 용어 • Section 3. 책임과 평가(Responsibility and Assessment) ① 알고리즘 관리감독(Oversight for algorithms) - (배경) A/IS의 기저에 있는 알고리즘은 지속적인 감독 및 평가에서 벗어나 있고 최종 사용자는 알고리즘이 어떻게 출력값을 산출하였는지 알 수 없는 문제가 있음 - (권고사항) 의료, 도시공학, 항공 분야 등과 같이 사유재산권과 관리감독의 균형을 찾을 필요 있음, 알고리즘의 세부사항은 일반대중에게 공개되지 않지만 윤리적 문제에 대한 공론화와 최종영향평가의 공개 필요 ② 독립적인 감사 기구(Independent review organization) - (배경) 대중들이 A/IS 기술을 충분히 활용하기 위해서는 독립적이고 전문적 의견을 제공할 수 있는 심의 기관이 필요함, 이 기관은 A/IS의 가장 적합한 사용법 및 위험성을 경고해야 하며 A/IS의 품질보증을 담당 - (권고사항) A/IS 상품이 디자인되고 개발되고 배치되는 과정에서 실제로 윤리적 기준을 만족하는지 체크할 수 있는 ISO와 같은 독립적인 국제기관이 구성되어야 함 ③ 기록 향상의 필요성(Need for better technical documentation) - (배경) 인간의 의사결정으로 인해 A/IS가 불가해한 특성을 나타내기도 함, 내/외부의 감사가 동반되는 향상된 기록이 필요하며, 이는 시스템의 윤리적 영향력을 파악하는 데 중요 - (권고사항) 기술자는 데이터의 흐름, 성능, 한계, 위험 등 전반적인 제품의 내용을 철저히 기록해야 하며 프로세스 과정에서 특이 사항, 규정 준수 여부, 방법론 또한 경험적 증거와 함께 명시적으로 기록해야 함(감사가능성, 접근성, 유의미성, 가독성 충족을 요함)

4. 불만 처리 및 분쟁 해결

가. 추천 서비스 제공자는 추천 시스템 적용 결과와 관련한 이용자의 불만을 신속히 처리하고 이용자 간의 분쟁을 적절하게 해결하기 위하여 필요한 조치를 취한다.

나. 불만 처리 및 분쟁 해결은 신규 혹은 기존 창구를 통해 시행하며, 창구 이용에 관한 안내는 홈페이지 알림 등 이용자가 쉽게 접근할 수 있도록 한다.

1. 규정 취지

디지털 플랫폼 내에서의 ‘불만 처리 및 분쟁 해결’은 이용자 보호 목적에서 권리침해, 이의신청 등을 처리하는 절차로서 서비스 이용 약관에 관련 내용이 규정된 경우가 많다. 추천 서비스의 내외부적 위험(성)은 사전에 전적으로 예방하는 것이 불가능하기 때문에 이용자의 불만 또는 분쟁 사항을 신속하게 접수하고 사후적으로 해결하는 것이 이용자의 권익 보장에 주효하다. 위 실행 원칙은 추천 서비스 이용자의 불만·분쟁 사항에 관한 절차적 권리를 보장함으로써 이용자의 능동적이고 종교적인 권리구제를 가능하게 하는 데 목적이 있다. 이는 앞서 제시된 실행 원칙들 - 이용자의 정보 공개 및 추천 기준 선택·변경에 대한 권리 등 - 이 실제로 담보될 수 있도록 이용자의 권리 행사·구제와 관련한 절차와 방법을 제시하는 것과도 무관하지 않다.

2. 주요 내용

추천 서비스 제공자는 이용자의 불만 처리와 분쟁 해결을 위하여 필요한 절차를 마련해야 한다. 추천 서비스 제공자는 추천 시스템 적용 결과에 대하여 결과 오류에 대한 사실확인, 불만 접수 등 이용자의 불만과 분쟁을 적절하게 해결하기 위한 절차와 방법 등을 마련하고 이를 이용자에게 안내할 필요가 있다. 불만 처리와 분쟁 해결은 별도의 전담 창구를 통해 실시하여야 하며, 해당 창구는 해당 플랫폼에서 이미 운영되고 있는 창구나 조직을 통해서도 이루어질 수 있으며, 신속하고 전문적인

해결을 위해 전담 조직을 별도로 설치·운영하는 것도 가능하다.

창구 이용에 대한 안내는 웹사이트·모바일앱 시작페이지 게시, 탭 메뉴 설정 등 이용자가 언제든지 쉽게 접근 가능한 방법으로 제공한다. 불만 처리 등 절차의 안내·진행 방식은 정형화될 수 없고 플랫폼 서비스의 형태에 따라 다양한 형태로 유연하게 구현할 수 있다. 다만, 절차의 유무보다 중요한 것은 창구 담당자의 전문가적 역량이다. 창구 담당자는 콘텐츠 추천 결과에 대한 불만 또는 분쟁 사안을 접수한 경우, 해당 사안을 유형별로 분류하여 추천 시스템의 작동 원리 및 효과에 대하여 일반 이용자 입장에서 이해할 수 있는 수준으로 충분히 설명할 수 있어야 한다. 인공지능 시스템의 투명성 내지 설명 가능성은 기술적으로 완벽히 확보하기 어렵고 이용자 스스로 충족시키는 것은 더욱 어려운 상황에서, 불만 처리 및 분쟁 해결 과정에서 이루어지는 ‘충분한 설명’은 결과적으로 추천 서비스의 투명성과 이용자의 정보 접근권 향상에 기여하는 측면도 있다.

i 참고 | 국내외 사례

1) 관련 법령(안) 및 지침

구분	내용
유럽연합 「디지털서비스법안」 (Digital Services Act)	■ 내부 불만 처리 시스템에 관한 의무조항을 둠 - (규모를 불문하고) 온라인 플랫폼은 이용자가 무상으로 이용할 수 있는 내부 불만 처리 시스템을 마련해야 함(제17조, Internal complaint-handling system) - 내부 불만 처리 시스템으로 문제를 해결할 수 없는 경우에 이용자는 법정 외 분쟁해결절차를 선택할 수 있으며, 온라인 플랫폼은 선정된 분쟁해결기구의 결정을 따라야 함(제18조, Out-of-court dispute settlement) ※ 조정기관은 해당 분쟁해결기구가 ① 온라인 플랫폼 및 해당 이용자와의 독립성, ② 해당 이슈에 대한 전문성, ③ 전자통신기술을 통해 분쟁 해결에 대한 접근 용이성, ④ 신속하고 효율적이며 비용 효과적인 방법으로 분쟁 해결 가능, ⑤ 명확하고 공정한 절차규정에 따라 분쟁 해결 가능한 여건 등을 증빙하여 요청하면 자격을 부여해야 함
독일 「소셜 네트워크에서의 법 집행 개선에 관한 법률」 (Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken: NetzDG)	■ 이의신청 처리 절차에 관한 의무조항을 둠 - 이용자 수가 2백만 명 이상인 대규모의 소셜 네트워크 서비스 제공자들에게는 불법 게시물에 대해 제기된 이의신청을 처리하고 그 결과를 보고해야 할 의무 부여(제2조 및 제3조) - 이용자가 불법 게시물을 쉽게 신고할 수 있도록 서비스 제공자는 쉽게 인지할 수 있고, 직접 접근이 가능하고, 상시로 이용할 수 있는 절차를 마련해야 함(제3조 제1항 제2문). • 게시물에 대한 신고가 들어온 경우 서비스 제공자는 즉시 해당 게시물의 삭제 또는 접근 차단 필요 여부를 검토해야 함(제3조 제2항 제1문 제1호) • 명확히 불법인 게시물에 대해서는 24시간 이내에 삭제 또는 접근을 차단해야 함(제3조 제2항 제1문 제2호) - 불법 게시물과 관련하여 이미 내려진 결정에 대해 2주 안에 다시 이의를 제기할 수 있는 절차를 마련할 것을 의무화(제3b조 제1항 제2문 및 제3문)

구분	내용
독일 「소셜 네트워크에서의 법 집행 개선에 관한 법률」 (Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken: NetzDG)	- 불법 게시물에 대해 1년간 100건 이상의 이의신청을 받은 경우 이의신청의 처리 및 결과와 관련해서 반년마다 보고서를 작성하여, 이를 연방관보(Bundesanzeiger) 및 플랫폼 홈페이지에 공개함(제2조 제1항 제1문)
일본 「특정 디지털플랫폼의 투명성 및 공정성 향상에 관한 법률」	■ 특정 디지털 플랫폼 제공자로 하여금 플랫폼 이용사업자와의 거래 관계에서의 공정성 확보를 위한 절차나 체제 정비, 고충 접수 또는 분쟁 처리 시스템 정비, 긴밀한 연락을 취하기 위한 국내 업무 관리자 선임 등을 하도록 함(제7조 제3항 제3호 및 제4호)
유럽연합 ‘신뢰 가능한 인공지능을 위한 자체평가목록’ (The Assessment List for Trustworthy Artificial Intelligence for self assessment)	■ 신뢰 가능한 인공지능 구현을 위하여 필요한 7대 요건을 제시하면서 이를 자체적으로 평가하기 위한 평가 목록을 제안 - 기술적 견고성과 안전성 관련 평가 항목으로 정확성을 제시하면서, 정확성 확보 여부를 확인하기 위한 질문 중 하나로 “이용자나 정보주체가 기대하는 AI 시스템의 정확성 수준에 대하여 소통될 수 있도록 보장하는 절차가 있는가”를 포함

2) 현행 통신분쟁조정 절차와의 관계

- 「전기통신사업법」 제45조의5는 전기통신사업자와 이용자 사이의 소규모 분쟁을 처리하기 위한 통신분쟁조정위원회의 통신분쟁조정 절차를 규정함
 - 인공지능 기반 미디어 서비스 관련 분쟁은 일반적인 통신분쟁 사안과 달리 미디어 분야의 자율성, 추천 서비스의 특수한 성격 및 기능 등을 특별히 고려해야 하므로, 위 통신분쟁조정 절차와 별개로 추천 서비스 관련 분쟁을 전문적으로 해결할 수 있는 조직·절차를 마련할 필요 있음
 - 소규모 분쟁인 경우에도 이용자가 용이하게 접근하여 권리구제에 적극 임할 수 있는 경로가 필요하므로, 행정절차 외에 플랫폼의 내부 절차를 이원적으로 운영한다면 이용자 권익 보장에 더욱 유리할 것으로 예상됨

〈참고〉 전기통신사업법 제45조의5(분쟁조정 절차)

- ① 전기통신에 관한 분쟁의 조정을 원하는 자는 대통령령으로 정하는 사항을 기재하여 분쟁조정위원회에 조정을 신청할 수 있다.
- ② 분쟁조정위원회는 제1항에 따른 분쟁조정 신청을 받은 때에는 그 사실을 다른 당사자에게 통지하여야 한다.
- ③ 분쟁조정위원회는 당사자 또는 이해관계인이 의견을 진술하려는 경우에는 특별한 사유가 없으면 의견진술의 기회를 주어야 한다. 다만, 당사자가 정당한 사유 없이 이에 따르지 아니하는 때에는 그러하지 아니하다.
- ④ 분쟁조정위원회는 분쟁의 조정을 위하여 필요하다고 인정하는 경우에는 당사자, 이해관계인 등에 필요한 자료의 제출을 요구할 수 있다.
- ⑤ 분쟁조정위원회는 제1항에 따라 분쟁조정 신청을 받았을 때에는 당사자에게 그 내용을 제시하고 조정 전 합의를 권고할 수 있다.
- ⑥ 분쟁조정위원회는 분쟁조정 신청을 접수한 날부터 60일 이내에 이를 심사하여 조정안을 작성하여야 한다. 다만, 부득이한 사정이 있는 경우에는 한 차례만 30일의 범위에서 분쟁조정위원회의 의결로 처리기간을 연장할 수 있다.
- ⑦ 분쟁조정위원회는 제6항 단서에 따라 처리기간을 연장한 경우에는 기간연장의 사유와 기한을 명시하여 당사자에게 통지하여야 한다.
- ⑧ 그 밖에 분쟁조정의 절차와 방법 등에 관하여 필요한 사항은 대통령령으로 정한다.

5. 내부 규칙 제정

가. 추천 서비스 제공자는 현행 법령, 기업 윤리, 인공지능 윤리 등을 감안하여 추천 시스템 사용 및 관리에 관한 내부 규칙을 자율적으로 제정하고 이를 공개한다.

나. 내부 규칙에는 본 기본원칙 준수를 위한 세부 지침 및 기술적·관리적 수단을 비롯하여 자율검증을 위한 조직 및 체계, 분쟁해결절차 등의 내용을 포함할 수 있다.

1. 규정 취지

디지털 미디어 플랫폼이 우리 사회의 정보 흐름에 상당한 영향을 미침에 따라 허위정보, 극단주의, 사이버 폭력 및 다양한 형태의 괴롭힘 및 학대를 양산할 우려도 커지고 있다. 하지만 다양한 유형의 콘텐츠를 유통시키는 플랫폼 사업자를 엄격하게 규제하는 것은 미디어의 자율성과 독립성을 직접 침해하는 결과를 초래할 수 있기 때문에 디지털 미디어 영역에서는 ‘미디어 자율규제 원칙’이 더욱 중시된다. 아울러 자율규제의 내용을 이루는 ‘자율규범’을 제정하는 권한도 수범자인 미디어 서비스 제공자에게 속하게 하는 것이 자율규제의 본질에 비추어 이상적이다.

‘내부 규칙 제정’은 추천 서비스 제공자의 다양성과 자율성이 존중되는 자율규제 기반을 갖추기 위한 필요조건으로 볼 수 있다. 다만, 본 기본원칙은 자율규범의 성격을 지니므로 위 실행 원칙에 근거하여 추천 서비스 제공자에게 내부 규칙 제정이 강제되는 것은 아니다.

2. 주요 내용

‘내부 규칙 제정’ 원칙은 서비스 제공자의 자율성을 존중함으로써 미디어의 다원성을 추구하는 관점을 부각한 것이다. 종래 미디어 자율규제는 미디어 사업자의 ‘윤리강령’에 근거할 때가 많았다. 윤리강령 기반의 자율규제는 미디어의 진실성과 공정성을 확보하겠다는 취지로, 법률에 의한 강제보다 사업자의 자율적인 준수를 권고하는 데 무게를 둔 것이다.

추천 서비스 제공자는 추천 시스템 사용 및 관리를 목적으로 자율검증을 위한 조직·체계, 분쟁해결절차 등에 관한 사항을 내부 규칙으로 제정할 것이 권고된다. 각 실행 원칙을 이행하는 세부 지침 역시 수범자 스스로 수립 가능하며, 현실적인 여건에 맞는 실행 가능한 수준의 기술적·관리적 조치를 구체화하여 자율규제의 직접적인 실행 방안으로 삼을 수 있다. 윤리강령에 준하는 내부규칙이라고 하더라도 그 규범적 근거는 윤리에 한하지 않고 현행 법체계에 담긴 이념, 원칙, 규정, 제도 등에도 뿌리를 둘 수 있다. 따라서 내용상 투명성 및 공정성 확보와 관련된 현행 법령의 준수 의무를 내부 규칙에 포함시킬 수 있음은 물론이다.

① 참고 | 국내외 사례

1) 관련 법령(안) 및 지침

구분	내용
「온라인 플랫폼 이용자 보호에 관한 법률안」 (20. 12. 전해숙의원 대표발의)	제5조(온라인 플랫폼 사업자의 자율규약) ① 방송통신위원회는 온라인 플랫폼 사업자 및 사업자 단체로 하여금 공정하고 투명한 거래질서를 확립하고 이용자 권익을 보호하기 위하여 자율적으로 규약을 정하도록 권장할 수 있다. ② 온라인 플랫폼 사업자 또는 이를 대표하는 사업자 단체는 방송통신위원회에 제1항에 따른 규약이 이 법을 위반하는지 여부에 대한 심사를 요청할 수 있다. ③ 방송통신위원회는 제2항에 따른 심사를 받은 규약의 이행을 독려하기 위하여 지원시책을 마련하여 시행할 수 있다.
「정보통신망 이용촉진 및 정보보호 등에 관한 법률」	제44조의4(자율규제) ① 정보통신서비스 제공자단체는 이용자를 보호하고 안전하며 신뢰할 수 있는 정보통신서비스를 제공하기 위하여 정보통신서비스 제공자 행동강령을 정하여 시행할 수 있다. ② 정보통신서비스 제공자단체는 다음 각 호의 어느 하나에 해당하는 정보가 정보통신망에 유통되지 아니하도록 모니터링 등 자율규제 가이드라인을 정하여 시행할 수 있다. 1. 청소년유해정보 2. 제44조의7에 따른 불법정보 ③ 정부는 제1항 및 제2항에 따른 정보통신서비스 제공자단체의 자율규제를 위한 활동을 지원할 수 있다.
일본 ‘AI 활용 가이드라인’ - AI 활용을 위한 실천자료 (A 利活用ガイドライン ~ A 利活用のためのプラクティカルリファレンス~)	• AI 서비스 제공자 및 AI 서비스를 활용하는 사업자는 이용자 등이 AI의 활용에 대해 적절히 인식할 수 있도록 다음과 같이 AI에 관한 이용방침을 작성, 공개하고 통지해야 함 (i) AI의 판단이 소비자 및 제3자에게 직접 영향을 미치는 경우 소비자 및 제3자가 그러한 사실을 적절히 인식할 수 있도록 AI 이용방침을 작성, 공표하고 문기가 있는 경우 통지해야 함

구분	내용
일본 ‘AI 활용 가이드라인’ - AI 활용을 위한 실천 자료 (A I 利活用ガイドライン ~ A I 利活用のためのプラクティカルリファレンス~)	(ii) (i)에 관하여 소비자 및 제3자의 권리와 이익에 중대한 영향을 미칠 가능성이 있는 경우에는 관련 사항을 적극적으로 통지하여야 함. 또한 통지나 공표는 이용 개시 전뿐만 아니라 AI의 운용 및 작동에 변동사항이 생겼거나 운용 및 작동을 종료하는 경우 등에도 이루어져야 함
유럽연합 ‘허위 정보에 대한 실천 강령’ (Code of Practice on Disinformation)	• 유럽연합은 자율규제와 관련한 다양한 권고안을 마련해 오면서, 추천 서비스에 의한 허위정보와 혐오표현 증폭을 해결하기 위한 다양한 규제 수단을 논의해 옴 - 2018년 유럽 위원회는 실천 강령(2018 European Commission’s Code of Practice)의 일환으로 허위 정보에 대한 실천 강령(Code of Practice on Disinformation)을 마련

2) 국내외 기업 내부 정책 수립 동향

구분	내용
카카오	• 카카오는 국내 기업 중 최초로 사회적 책임에 걸맞은 AI 기술 개발·활용에 관한 윤리 규범으로서 2018년 1월 <알고리즘 윤리 헌장>을 발표 - 카카오 알고리즘의 기본원칙, 차별에 대한 경계, 학습 데이터 운영, 알고리즘의 독립성, 알고리즘에 대한 설명, 기술의 포용성, 아동과 청소년에 대한 보호에 관한 규정을 둠
네이버	• 네이버는 첨단 AI 기술을 누구나 쉽고 편리하게 활용할 수 있는 일상의 도구로 만든다는 목표 아래, 2021년 2월 <AI 윤리 준칙>을 발표 - 사람을 위한 AI 개발, 다양성 존중, 합리적인 설명과 편리성의 조화, 안전을 고려한 서비스 설계, 프라이버시 보호와 정보 보안 등을 핵심 내용으로 함

구분	내용
페이스북 (現 메타)	• 페이스북 감독위원회(The Facebook Oversight Board; FOB) - 일종의 자율규제 메커니즘을 도입한 것으로, 페이스북은 FOB 문서(결정/권고)에 동의하고 권장 사항을 준수하는 조치를 취해야 함 - 이러한 규제 메커니즘이 제대로 작동할 경우 콘텐츠 큐레이션과 관련된 문제를 실질적으로 개선할 수 있을 것으로 기대 • 메타 데이터 정책 - 광고, Facebook 뉴스피드, Instagram 피드 및 Instagram 스토리 등의 기능 및 콘텐츠를 맞춤화하고, 관심을 가질 만한 그룹 또는 이벤트나 팔로우할 만한 주제 등을 추천 - 맞춤형 서비스 제공을 위해 인맥, 기호, 관심사 및 활동, 서비스 이용 및 상호 작용 방법, 메타 안팎으로 연결되어 있거나 관심 있는 사람, 장소 또는 사물에 관한 정보 등 수집하거나 제공받은 데이터 이용
구글	• 구성원을 대상으로 책임 있는 AI 개발 및 사용에 대한 안내를 위해 <Recommendations for Regulating AI> 수립
마이크로소프트	• 마이크로소프트는 에테르 위원회(Aether Committee), 책임 있는 AI 사무소(ORA), 책임 있는 AI 엔지니어링 전략(RAISE) 등을 통해 작업 전반에 책임 있는 AI 원칙을 준수
BBC	• BBC 머신러닝 원칙(BBC Machine Learning Principles) - BBC의 가치 반영, 고객 경험 개선, 투명성, 책임성, 보안과 공정성 보장, 저널리즘의 가치 확장, 인간 참여와 검토의 10가지 원칙 및 뉴스 AI 개발 시 고려사항 제시

V

기본원칙 이행을 위한 정부의 역할

1. 서비스 제공자 지원

가. 정부는 추천 서비스 제공자가 본 기본원칙을 효과적으로 이행할 수 있도록 이해관계자와 전문가의 의견을 반영하여 실행 가이드와 지원 정책을 수립한다.

나. 정부는 본 기본원칙을 이행하는 데 필요한 기술적·관리적 조치에 관한 실행 가이드를 제공하고 지속적으로 추천 서비스의 투명성·공정성을 향상시키는 방안을 모색하기 위하여 정책자문기구를 둘 수 있다.

1. 규정 취지

통상 플랫폼 사업자의 지위에 있는 추천 서비스 제공자는 ‘규제된 자율규제’ 방식에 기초한 각종 법률상 의무 규정의 수범자가 된다. 법률에 근거를 둔 자율규제의 경우에는 수범자인 추천 서비스 제공자가 근거규범의 제정(입법절차)에 참여하는 것이 한시적이고 제한적인 반면에, 본 기본원칙은 규범 제정뿐만 아니라 합리적인 이행 방안을 모색하는 후속 단계에서도 수범자를 비롯한 이해관계자의 실질적인 참여를 유도하고자 한다. 그 과정에서 정부의 후견적·중개적 역할이 불가결하다.

정부로서는 디지털 미디어 영역에서 보편화되고 있는 추천 서비스가 신규 미디어 서비스로서 혁신성과 부가가치를 생산하면서 이용자의 권익과 편의도 향상시키는 미디어 환경을 구현할 수 있도록 정책적 방안을 마련하는 노력을 기울여야 한다. 본 기본원칙은 추천 서비스에 대한 자율규제의 근거규범으로 마련된 것으로, 실제로 서비스 제공자가 이를 준수하기 위해서는 관련 기준이나 요건이 더욱 구체적으로 제시될 필요가 있다. 특히 실행 원칙의 세부 지침을 마련하고 수범자의 이행 여건을 조성하는 정책적 목표 아래 ‘서비스 제공자 지원’이 주요한 정부의 역할로 제시된 것이다.

2. 주요 내용

본 기본원칙과 관련한 정부의 후속 과제는 3대 핵심 원칙과 5대 실행 원칙에 대한 수범자의 이해도와 수용도를 높이고 자발적으로 준수할 수 있는 방침과 지원책을 마련하는 것이다. 본 기본원칙이 추천 서비스 규율의 원칙으로서 사실상 규범력을 발휘하고 효과적으로 이행되기 위해서는 무엇보다 서비스 제공자와 이용자를 포함하는 이해관계자의 입장을 광범위하게 청취하고 이를 반영할 필요가 있다. 특히 정보 공개, 선택권 보장, 자율검증 세 가지 실행 수단과 관련하여 현실적으로 실행 가능한 수단을 제시하기 위해서는 이해당사자 외에도 전문가들의 다양한 의견을 수렴하여 균형 있게 검토하는 과정 역시 필수적이다.

이와 같이 기본원칙에 대한 해석지침과 실행지침을 마련하고 지속적으로 현재 기술 수준과 산업 여건에 맞는 실효적인 이행 방안을 집중적으로 연구하고 도출하기 위하여 관련 정책 수립을 지원하는 정책자문기구를 두는 것이 필요하다. 정부는 정책자문기구를 통해 이해관계자 의견을 충분히 반영한 유연하고 합리적인 실행지침을 마련하고, 장기적으로 인공지능 기반의 미디어 서비스의 투명성·공정성을 제고하여 이용자 권익을 보장하는 정책을 개발·추진해 나가야 할 것이다.

2. 이용자 역량 강화

가. 정부는 추천 서비스 이용에 따른 불이익과 피해를 예방하고 합리적인 미디어 소비 환경을 조성하기 위하여 이용자의 인공지능 활용 역량을 강화하는 방안을 마련하여야 한다.

나. 정부는 추천 서비스 제공자가 이용자 대상의 인공지능 역량 교육, 인공지능 교육 콘텐츠 보급 등을 확대 및 강화할 수 있는 정책을 마련하여야 한다.

1. 규정 취지

여기서 말하는 이용자 역량은 ‘디지털 미디어 리터러시(digital media literacy)’로서 디지털 미디어를 사용하기 위해 요구되는 지식, 기술, 능력 등을 의미한다. 이는 스크린 및 방송이 전달하는 영상 언어를 이해하는 차원의 ‘미디어 리터러시’에서 발전한 개념으로, 각종 디지털 미디어의 기능을 이해하고 활용할 수 있는 역량을 뜻한다. 즉, 미디어가 전달하는 메시지를 단순히 이해하는 것뿐만 아니라 미디어를 능동적으로 선택하고 이용하며 콘텐츠 생산 과정에도 참여하는 능동적인 리터러시 개념에 해당한다.

국내외 미디어 플랫폼에서 추천 서비스가 널리 상용화되고 있음에도 불구하고 이에 대한 이용자의 지식과 이해도는 여전히 낮은 수준에 머물고 있다. 이에 본 기본원칙은 ‘이용자 역량 강화’를 정부의 중요한 임무로 명시한 것이다. 추천 서비스와 관련한 이용자 역량은 궁극적으로 이용자 스스로 서비스의 성격과 원리를 파악하여 합리적인 미디어 소비를 주도할 수 있는 능력으로 보아야 한다. 이러한 추천 서비스 이용자의 역량을 효과적으로 함양하기 위해서는 서비스 제공자의 관련 정보 제공을 확보하고 다양한 수준의 교육 프로그램을 지원하는 정부의 정책적 노력이 수반될 필요가 있다.

2. 주요 내용

추천 서비스 이용자의 역량 강화는 서비스 이용 과정에서 미디어 콘텐츠 자동 배열에 따른 불이익과 피해를 예방하는 데 일차적인 목적이 있다. 이때 불이익은 다양한 차원에서 ‘공정성’ 원칙을 위배 또는 훼손하는 이익 침해 유형으로 볼 수 있다. 미디어 영역에서 추천 서비스가 유발할 수 있는 불이익은 — 완전히 자동화된 의사결정이 이루는 경우와 달리 — 이용자의 선택행위에 의하여 비로소 구체화된다. 따라서 이용자가 추천 기준과 결과를 이해하고 본인의 목적과 선호에 맞게 미디어 콘텐츠를 최종 소비할 수 있다면 다른 불법적인 요인이 없는 이상 추천 서비스 이용에 따른 불이익이나 피해를 인정하기는 어렵다. 요컨대, 이용자의 ‘디지털 미디어 리터러시’를 강화하는 것은 추천 서비스 이용자의 권익을 보장하고 피해를 사전에 차단하기 위한 가장 근본적인 방안이다.

오늘날 디지털 미디어 리터러시가 국제적 관심사로 떠오르면서, UNESCO는 미디어·정보 리터러시 관련 커리큘럼 개발, 국제 공조, 정책 수립, 글로벌 프레임워크 완성 등의 포괄적인 전략을 통해 미디어·정보 리터러시가 높은 사회를 구현하는 것을 국가의 중요한 목표로 제시하고 있다. 위에서 언급한 ‘인공지능 역량(AI literacy)’은 디지털 미디어 리터러시 범주에 속하는 것으로, 추천 서비스에 사용되는 인공지능 기술 또는 인공지능 시스템에 대한 이해·이용·응용 능력을 포괄하는 의미를 지닌다. 인공지능 기반의 지능형 미디어 플랫폼의 성장세를 감안하면 인공지능 역량 교육은 향후 디지털 미디어 리터러시 교육에서 절대적인 비중을 차지하게 될 것이다.

정부는 이용자의 디지털 미디어 리터러시 증진을 위한 각종 교육프로그램과 정보의 제공을 통해 건전한 미디어 환경이 조성될 수 있도록 노력하여야 하고, 이를 실질적으로 달성하기 위하여 구체적인 정책과 재원을 마련하여야 한다. 이러한 정부의 후견적 역할은 이용자의 분별력 있는 미디어 소비를 유도할 수 있다. 또한 플랫폼 사업자로서는 향상된 이용자 역량에 상응하여 경쟁력 있는 양질의 미디어 서비스를 제공하기 위해 추천 서비스의 공정성과 투명성을 확보하는 노력을 기울이는 선순환을 기대할 수 있다.

i 참고 | 국내외 사례

1) 미디어·인공지능 리터러시 관련 법령·지침·정책

구분	내용
「방송통신발전 기본법」	• 방송통신발전 기금의 용도 중 하나로 시청자가 직접 제작한 방송프로그램 및 미디어 교육 지원을 명시(제26조제1항제7호)
「국가정보화 기본법」	• 국가기관 및 지방자치단체는 인터넷 중독의 예방 및 해소를 위하여 필요한 교육을 실시할 수 있음(제30조의8제1항)
유럽연합 「시청각 미디어 서비스 지침」 (Audiovisual Media Services Directive: AVMSD)	• 회원국은 미디어 리터러시 기술을 개발하는 조치를 촉진해야 함(제33a조) • 비디오 공유 플랫폼은 효과적인 미디어 리터러시 측정과 도구를 제공해야 함(제28b조 제2항)
핀란드	• 핀란드 정부는 2013년부터 미디어와 정보 리터러시(media and information literacy)에 중점을 둔 미디어 리터러시 교육 추진 - 교육을 통해 아동 및 청소년의 미래를 위한 지식과 경쟁력을 제고하는 동시에 어느 누구도 이러한 교육의 기회와 혜택에서 배제되지 않도록 하는 데 목표를 둠 - 학령기 아동의 오전 및 오후 활동(activity)에 미디어 리터러시 교과과정을 포함시킴 ※ 유럽 35개국을 대상으로 한 뉴미디어 리터러시 조사에서 핀란드는 미디어 리터러시 점수가 가장 높은 국가로 나타남
프랑스	• 2005년 「학교의 미래를 위한 기본법 및 계획법(Loi d'orientation et de programme pour l'avenir de l'école)」 개정하여 미디어 교육을 공교육에 포함시킴 • 이후 2013년 개정된 「프랑스 공화국의 학교 재건을 위한 기본법 및 계획법(Loi d'orientation et de programmation pour la refondation de l'école de la République)」에서 디지털 교육 및 원격교육을 공공 서비스로 도입하는 것을 명시하여 미디어 및 정보 리터러시 교육의 중요성 강조
영국 오프콤(Ofcom)	• 오프콤은 「커뮤니케이션법(Communications Act 2003)」 제11조에 따라 미디어 리터러시를 촉진시킬 의무가 있고, 동법 제14조(6)(a)에 따라 관련 연구를 수행해야 함

2) 인공지능 활용 역량 강화에 관한 지침

지침명	내용
일본 'AI Utilization Guidelines'	• 적정활용의 원칙(Principle of proper utilization) “사용자는 AI 시스템이나 AI 서비스를 적절한 범위와 방식으로, 인간과 AI 시스템 간 또는 사용자 간 적절한 역할 분담 하에 활용하도록 노력해야 한다” - 사용자 등이 AI의 활용 목적과 용도, 성격, 능력을 제대로 인식한 후 활용할 수 있도록 개발자 및 AI 서비스 제공자는 필요한 정보·설명을 제공할 필요 있음
일본 'Social Principles of Human-centric AI'	• 인간중심주의(The Human-Centric Principle) - AI 활용이 헌법과 국제 기준에 의해 보장된 기본적 인권을 침해해서는 안 됨 • 교육/문해력의 원칙(The Principle of Education/Literacy) - 문해력과 AI의 올바른 활용 촉진을 위해 적절한 메커니즘(공교육, 재교육 등)이 도입될 필요 있음 - AI 교육과 문해력을 키우는 교육환경이 모든 국민에게 평등하게 제공되어야 함
EU AI HLEG 'Ethics Guideline for Trustworthy AI'	• '비기술적 방법'으로서 리터러시 고려 - 윤리적 사고방식을 함양하기 위한 교육 및 인식
OECD 'Ethics Guideline for Trustworthy AI Recommendation of the Council on Artificial Intelligence'	• 인공지능 사용 확산을 위한 정부의 역할로 '인력 양성 및 노동시장 전환 준비' 명시 “정부는 이해관계자들과 긴밀히 협력하여 직업 세계와 사회의 변화에 대비해야 한다. 필요한 기술을 갖추는 것을 포함하여 다양한 응용 분야에서 사람들이 AI 시스템을 효과적으로 사용하고 상호작용할 수 있도록 해야 한다”
전기전자기술자협회(IEEE) 'Ethically Aligned Design'	• 윤리적 설계 요건 중 하나로 '리터러시' 제시 “자동화된 지능형 시스템(A/IS) 개발자는 안전하고 효과적인 운영에 필요한 지식과 기술을 명시하고 준수하여야 한다”

3) 국내외 민간·공공 이행 사례

지칭명	내용
구글	• 대중에게 머신러닝 교육을 무료로 제공하는 ‘구글 AI 학습’ 프로그램을 도입 - 본 프로그램은 초보자부터 전문가에 이르는 다양한 수준의 AI 학습 자료 및 커리큘럼을 제공
마이크로소프트	• AI 학습 트랙을 대중에게 공개 - ① AI 소개, ② 데이터 과학을 위한 파이썬, ③ 수학과 통계, ④ AI를 위한 윤리, ⑤ 딥 러닝 등 - 본 프로그램을 접하는 개발자들이 실무 실습과 일련의 온라인 전문 강좌를 통해 취업 준비 기술과 실제 경험을 쌓아 AI와 데이터 과학에 대한 기술을 개발·고도화하는 것을 지원
독일 ‘인공지능 전략 2020’	• 독일 연방정부는 2018년 11월 업무, 환경, 기후 등 다양한 분야에서 AI 개발·이용을 활성화하여 국가 경쟁력을 강화하는 것을 목표로 ‘인공지능 전략 2020(Strategie Künstliche Intelligenz der Bundesregierung Fortschreibung 2020)’ 수립 - 독일의 모든 시민이 인공지능에 대한 편견을 해소하고 이해도를 높일 수 있도록 AI 관련 교육을 충실하게 제공하는 것을 목표에 포함시킴 - AI 메커니즘과 응용 가능성에 대한 이해도를 높이기 위해 무료 온라인 강좌 “AI의 요소(Elements of AI)” 프로그램을 지원하는 등 시민의 AI 활용 역량 강화에 주력