

자동화된 의사결정 규제에 관한 비교법적 고찰 및 한국 법제에의 시사점

정 란*

서울대학교 법학연구소 선임연구원, 법학박사

국문 초록

디지털 미디어 환경에서 이루어지는 자동화된 의사결정(automated decision-making, ADM)은 표현의 자유, 인격권, 개인정보자기결정권 등 다양한 기본권에 중대한 영향을 미치는 과정으로, 특히 포털 뉴스 배열, 동영상 플랫폼 추천, AI 기반 콘텐츠 필터링 등 미디어 플랫폼의 알고리즘 운영 전반과 밀접하게 연관된다. 우리나라는 2023년 개인정보보호법 개정을 통해 자동화된 결정에 대한 정보주체의 권리 규정을 도입하였고, 시행령 및 고시를 통해 세부 절차를 마련하였으나, 개념 정의의 불명확성, 통지·설명 의무의 추상성, 재심 절차의 독립성 부족, 기록·감독 규정의 부재, 그리고 미디어 법제와의 구조적 연계 부족 등 여러 한계를 드러내고 있다.

본 연구는 이러한 한계를 극복하기 위하여 EU GDPR 제22조, EU 디지털서비스법(DSA), 캐나다 퀘벡주 「법률 25」, 브라질 LGPD, 영국 OSA, 호주 eSafety 체계 등을 비교법적으로 분석하고, 미디어 플랫폼에서 ADM이 작동하는 특수성을 고려하여 우리 법제의 개선 방향을 제시한다. 헌법상 명확성·비례성 및 국가의 기본권 보호의무에 비추어 현행 ADM 제도를 재검토하고, 특히 미디어 편집행위와 자동화된 결정의 경계가 흐려지는 현실에서 ADM을 '언론적 행위'로 평가할 수 있는 가능성을 검토하며, 기관 간 조정 부재로 인한 규제 공백 문제를 비판적으로 조명한다.

* flyran84@snu.ac.kr

연구는 핵심 개념들의 법정화와 표준화, 통지·설명 의무의 구체적 요소 확립, 독립적·실효적 재심 절차 마련, 기록의무 및 감독·감사 권한 도입, 미디어 법제와의 연동을 통한 플랫폼 규제 거버넌스 정비 등을 단계적 정책대안으로 제시한다. 이러한 제도 개편은 디지털 시대 인격권 및 개인 정보자기결정권 보호의 실효성을 강화하면서, 미디어 플랫폼 환경에서 공정하고 투명한 정보유통 구조를 확립하는 데 기여할 것이다.

주제어: 자동화된 의사결정, 개인정보보호, 거부권, 설명의무, 알고리즘 투명성, 비례원칙

목 차

- I. 서론
- II. 자동화된 의사결정 규제에 대한 현행 법제 검토
 1. 개인정보보호법 제37조의2의 도입 배경과 주요 내용
 2. 시행령, 고시, 안내서를 통한 세부 운영기준
 3. 자동화된 의사결정 규제에 대한 현행 법제의 한계
- III. 자동화된 의사결정에 대한 정보주체 권리 및 의무의 비교법적 검토
 1. EU GDPR 제22조
 2. 캐나다 퀘벡주 「Loi 25」
 3. 브라질 LGPD
 4. 시사점
- IV. 알고리즘 투명성과 플랫폼 책임에 대한 해외 규제 동향
 1. EU 디지털서비스법(DSA)
 2. 영국 「온라인안전법(Online Safety Act, OSA)」
 3. 호주 eSafety와 BOSE
 4. 시사점
- V. 자동화된 의사결정에 대한 규제의 헌법적 쟁점
 1. 국가의 기본권 보호의무 및 대사인효
 2. 기본권 충돌의 조화와 비례원칙-절차적 투명성을 통한 균형
 3. 명확성 원칙 및 예측가능성의 확보
 4. 디지털 공론장과 표현의 자유
- VI. 우리나라 법제의 개선 방향 및 정책 제언
 1. 주요 개념 정의의 법정화
 2. 통지 및 설명 의무 요소의 표준화
 3. 재심(인간 개입) 절차의 실효성 확보
 4. 기록 유지의무와 검증 인프라 구축
 5. 거버넌스 구축 및 기관 간 역할 조정
 6. 영업비밀 보호와 SME 차등규제
- VII. 결론

I. 서론

오늘날 디지털 플랫폼 사업자들은 인공지능과 빅데이터 기반 알고리즘을 활용하여 방대한 정보와 콘텐츠를 자동으로 분류·추천하거나 차단·필터링하고 있다. 미디어 플랫폼을 통한 콘텐츠 제공 및 유통이 활성화되면서, 플랫폼 사업자들이 제한된 디스플레이 공간에서 이용자들의 관심 및 행태정보를 활용하여 다양한 맞춤형 추천 콘텐츠 서비스를 운영하고 있다. 예컨대, 소셜미디어의 뉴스피드 추천, 동영상 플랫폼의 맞춤형 콘텐츠 추천, 검색엔진의 결과 정렬, 온라인 마켓플레이스의 상품노출 순위, 콘텐츠 모니터링 시스템의 자동 차단 결정 등은 모두 자동화된 의사결정(automated decision-making, ‘ADM’)의 대표적 사례로, 사전 설정된 기준 또는 알고리즘이 결과를 도출하는 방식이다.

이러한 알고리즘 결정은 단순한 사용자 경험의 문제를 넘어, 정보접근 기회와 여론 형성, 표현의 자유, 개인의 사회적 평가 등 핵심적 기본권에 직접적인 영향을 미치고, 더 나아가 고용, 신용, 의료, 교육 등 사회·경제적 권리에도 결정적인 영향을 미치는 사례가 증가하고 있다(Milano et al., 2022; Portes et al., 2020). 이에 따라 국제적으로도 자동화된 의사결정이 개인의 권리 또는 의무에 중대한 영향을 미치는 경우, 정보주체가 이를 인지하고 이의를 제기할 수 있는 권리의 필요성이 광범위하게 논의되어 왔다(Malgieri, G., 2019; Kaminski, M. E. et al., 2020; Future of Privacy Forum, 2022).

고도화된 인공지능 기반 추천 서비스는 이용자가 특정한 유형의 콘텐츠에 지나치게 편향되게 노출될 가능성을 증가시키며, 이용자들이 이러한 서비스의 작동 방식 자체를 정확히 인지하지 못하여 이를 시정하거나 대응할 기회를 갖지 못할 우려가 있다. 이로 인해 알고리즘 결정의 편향성이나 오류로 인한 차별·권리침해 문제, 예를 들어 AI 채용 시스템의 차별적 탈락 판단, 콘텐츠 추천 알고리즘의 편향된 정보노출 등 이른바 필터버블, 콘텐츠 자동차단으로 인한 부당한 표현 제한 등의 문제가

제기된다. 자동화된 의사결정의 주요 위험으로 의도하지 않은 편향(unintended bias), 데이터 불균형, 학습데이터의 부적절성 등으로 인해 개인에게 불이익을 줄 수 있음을 경고하며(Cathy O'Neil, 2017), 설명가능성의 확보 및 투명성 제고를 위한 입법 필요성을 강조하였다. 우리나라에서도 포털의 뉴스 배열 알고리즘 논란, 온라인 플랫폼의 검색 결과의 불공정성, 유튜브 추천 알고리즘의 편향성 등 알고리즘의 위험이 지속적으로 공론화되어 왔다(이성엽, 2024; 박노형, 2024). 구방송통신위원회는 생성형 인공지능서비스 이용자보호 가이드라인을 마련하고(방송통신위원회·정보통신정책연구원, 2025),¹⁾ 알고리즘 투명성위원회 설치 등을 추진하며 미디어 분야에서 알고리즘에 대한 규범화 논의를 본격화한 바 있다. 이러한 논의는 플랫폼의 추천·배열 알고리즘이 사실상 편집행위의 기능을 수행하며, 이는 여론 형성과 표현의 자유에 직접적인 영향을 미친다는 점에서 개인정보보호법 중심의 기존 규율만으로는 문제를 충분히 포착할 수 없음을 시사한다(Marta Beltrán, 2025).

이러한 맥락에서 자동화된 의사결정에 대한 대응권은 데이터 보호와 미디어 이용자 보호를 동시에 아우르는 새로운 권리보호 장치로 부각되고 있다. 이는 개인정보자기결정권과 디지털 시대 인격권 보호 측면에서 중요한 의의를 가지며, 동시에 건강한 정보유통 질서와 여론의 다양성 보장이라는 헌법적 가치와도 밀접하게 연결된다. 특히 2023년 3월 개정

1) 방송통신위원회는 OECD의 「인공지능에 대한 권고(Recommendation of the Council on Artificial Intelligence, '24)」와 유네스코(UNESCO)의 「인공지능 윤리에 대한 권고(Recommendation on the Ethics of Artificial Intelligence), '21)」를 참조하여, 생성형 인공지능 서비스의 전반에 걸쳐 추구해야 할 기본 방향성을 다음과 같이 제시하였다. 1) 생성형 인공지능 서비스는 인간의 존엄성을 보호하고 개인의 자유와 권리를 보장하며, 인간이 적절하게 통제하고 감독할 수 있는 방식으로 제공되어야 합니다. 2) 생성형 인공지능 서비스는 인공지능시스템 사용에 따른 작동 원리 및 결과, 이용자에게 미치는 영향 등에 대하여 이용자에게 알기 쉽게 설명되도록 제공되어야 합니다. 3) 생성형 인공지능 서비스는 안전하게 작동되어야 하며, 예상치 못한 피해를 최소화하고 악의적으로 이용되거나 변경되지 않도록 제공되어야 합니다. 4) 생성형 인공지능 서비스는 이용자에게 차별 또는 불공정한 결과가 발생하지 않도록 제공되어야 합니다.

되어 2024년부터 시행된 개인정보보호법 제37조의2는 자동화된 결정에 대한 정보주체의 거부권, 설명요구권, 검토 요구권 등을 명문화함으로써, 한국도 본격적인 알고리즘 규제와 정보주체 권리 보장 체계를 갖추게 되었다(이상용·이혜리, 2023).

이에 본 연구는 한국의 자동화 의사결정 통지·대응권 제도를 헌법적 관점에서 평가하고, 주요 해외 입법례 및 플랫폼 규제 동향을 비교 분석함으로써 향후 제도 개선 방향을 모색하는 것을 목적으로 한다. 특히 2023년 개정 개인정보보호법에 도입된 자동화된 결정에 대한 통지 및 대응권 규정을 중심으로, 그 실효성을 높이기 위한 입법 정책과제를 도출하고자 한다. 논문의 구성은 먼저 현행 우리나라 법제를 분석하고(II), 주요국의 개인정보 보호법상 자동화된 결정 규율(III)과 플랫폼 투명성 규제를 조망한 뒤(IV), 이러한 논의에 대한 헌법적 검토(V)를 수행한 후 제도 개선을 위한 구체적인 정책대안을 제안(VI)한다.

II. 자동화된 의사결정 규제에 대한 현행 법제 검토

1. 개인정보보호법 제37조의2의 도입 배경과 주요 내용

개인정보보호법 제37조의2는 2020년대 들어 부각된 인공지능 시대의 개인정보자기결정권을 강화하기 위한 장치로서, 2023년 3월 개정법에 신설되어 2024년 3월 15일부터 시행되었다.²⁾ 이 조항은 “완전히 자동화

2) 개인정보 보호법[시행 2025. 10. 2.] [법률 제20897호, 2025. 4. 1., 일부개정]

제37조의2(자동화된 결정에 대한 정보주체의 권리 등)

① 정보주체는 완전히 자동화된 시스템(인공지능 기술을 적용한 시스템을 포함한다)으로 개인정보를 처리하여 이루어지는 결정(「행정기본법」 제20조에 따른 행정청의 자동적 처분은 제외하며, 이하 이 조에서 “자동화된 결정”이라 한다)이 자신의 권리 또는 의무에 중대한 영향을 미치는 경우에는 해당 개인정보처리자에 대하여 해당 결정을 거부할 수 있는 권리를 가진다. 다만, 자동화된 결정이 제15조 제1항제1호·제2호 및 제4호에 따라 이루어지는 경우에는 그러하지 아니하다.

된 시스템으로 개인정보를 처리하여 이루어지는 결정”에 대해 정보주체가 일정한 권리를 가진다고 규정한다. 여기서 핵심이 되는 정보주체의 권리는 크게 해당 결정을 거부할 수 있는 권리와 해당 결정에 대한 설명 등을 요구할 수 있는 권리이다. 거부권은 그 결정이 정보주체의 권리 또는 의무에 중대한 영향을 미치는 경우에 인정되며, 설명요구권은 보다 일반적으로 자동화된 결정에 노출된 경우에 행사할 수 있는 것으로 해석된다. 또한 개인정보처리자가 자동화된 결정에 관하여 그 기준과 절차, 개인정보가 처리되는 방식 등을 공개할 의무를 진다고 규정하여, 알고리즘의 투명성에 대한 일반적 의무를 부과하고 있다. 이는 GDPR 제 22조의 구조적 영향을 받았지만, GDPR이 자동화된 의사결정 자체를 원칙적으로 제한하면서 예외적으로 허용하는 구조인 반면(Goodman & Flaxman, 2017), 우리 법은 자동화된 처리 자체를 제한하기보다 데이터 활용과 권리보호 간 균형을 도모하는 방식으로 입법설계를 하였다는 평가가 가능하다(최경진, 2024).³⁾ 아울러 법 제37조의2 제2항은 권리 행사의 예외 사유를 두어, 법률에 특별한 규정이 있거나 법령상 의무 준수를 위해 불가피한 경우, 정보주체와의 계약 체결·이행을 위해 필요한 경우로서 자동화 결정 사용을 알리고 동의를 받은 경우, 정보주체가 명시적

② 정보주체는 개인정보처리자가 자동화된 결정을 한 경우에는 그 결정에 대하여 설명 등을 요구할 수 있다.

③ 개인정보처리자는 제1항 또는 제2항에 따라 정보주체가 자동화된 결정을 거부하거나 이에 대한 설명 등을 요구한 경우에는 정당한 사유가 없는 한 자동화된 결정을 적용하지 아니하거나 인적 개입에 의한 재처리·설명 등 필요한 조치를 하여야 한다.

④ 개인정보처리자는 자동화된 결정의 기준과 절차, 개인정보가 처리되는 방식을 정보주체가 쉽게 확인할 수 있도록 공개하여야 한다.

⑤ 제1항부터 제4항까지에서 규정한 사항 외에 자동화된 결정의 거부·설명 등을 요구하는 절차 및 방법, 거부·설명 등의 요구에 따른 필요한 조치, 자동화된 결정의 기준·절차 및 개인정보가 처리되는 방식의 공개 등에 필요한 사항은 대통령령으로 정한다.

3) International Bar Association(2024)는 “Brazilian legal framework on automated decision-making”에서 브라질 LGPD 제20조와 유사하게, 한국 개인정보보호법도 자동화된 의사결정에 대한 권리를 인정하면서도 실무적 유연성을 확보하려 한 것으로 평가하였다.

으로 동의한 경우에는 예외를 인정하고 있다. 따라서 예외 사유에 해당하면 정보주체의 거부 요구를 받아들이지 않을 수 있다.

제37조의2의 도입 취지는, AI 기반 자동화기술 확산에 따른 개인정보 활용 증가 속에서 정보주체의 통제권을 강화하려는 것이다. 특히 채용, 신용평가, 금융거래 등 분야에서 AI 전형, 이상거래탐지, 자동차단 시스템이 도입되면서 편리성과 효율성 증대라는 이면에 차별 문제나 사생활 침해 우려가 제기되었다. 예컨대 AI 채용면접이 영상을 분석하여 합격 여부를 자동판단하거나, 자동화된 신용평가가 금융거래를 일괄 차단하는 사례가 보고되었다. 이러한 위험은 자동화된 결정이 인간의 사회적 기회에 심대한 영향을 미칠 수 있음에도, 정보주체가 그 내용을 알거나 다룰 절차가 없다는 점에서 본질적으로 문제된다(Citron & Pasquale, 2014).

여기서 주목할 점은, 이 규범의 의미가 단순한 개인정보 규제 기술이 아니라 헌법이론적 장치라는 점이다. Pollicino & De Gregorio는 플랫폼의 알고리즘이 여론 형성·정보노출·표현 통제까지 수행하면서 ‘준-공권력’으로 기능한다고 분석하며, 이러한 디지털 권력은 정보주체에게 이의 제기·설명·인간 개입 요청과 같은 절차적 권리를 부여하는 방식으로 헌법적으로 통제되어야 한다고 주장한다(Pollicino & De Gregorio, 2021). 또한 설명권과 재검토권이 단순한 정보 제공이 아니라, 차별 방지·투명성·인권 보호를 위한 ‘절차적 기본권’ 이라고 이해하기도 한다(Wachter & Mittelstadt, 2019; Kaminski et al., 2020).

정리하면, 개인정보보호법 제37조의2는 전적으로 자동화된 결정의 개념을 설정하고, 중대한 영향이 있는 경우 거부권과 설명요구권을 인정하며, 개인정보처리자에게 인적 개입 및 재처리 의무를 부과하고, 알고리즘 운용 정보의 공개 의무를 부과한다. 이를 통해 정보주체는 디지털 환경에서 ‘데이터의 대상’이 아니라 의사결정의 ‘참여 주체’가 될 수 있다. 즉, 이 조항은 디지털 시대의 개인정보자기결정권을 수평적 디지털 권력에 대응하는 새로운 거버넌스 참여권으로 확장하고 있다. 다만, 구체적인 적용을 위한 사항 상당수가 시행령 등 위임 규정에 맡겨져 있으며, ‘중

대한 영향' 등 핵심개념이 법률 자체에 정의되지 않았다는 한계가 지적된다(김연진, 2023).

2. 시행령, 고시, 안내서를 통한 세부 운영기준

개인정보보호법 제37조의2가 자동화된 결정에 대한 기본권적 권리와 원칙을 선언적으로 규정하고 있는 데 반하여, 실제 제도의 이행과 운영에 관한 구체적 기준은 시행령 및 개인정보보호위원회 고시와 안내서(가이드라인)를 통해 구체화되고 있다. 2023년 개정 개인정보보호법 제37조의2를 규정하고, 2024년 3월 시행령 제44조의2 이하를 제정하였으며,⁴⁾ 개인정보처리자가 자동화된 결정의 기준과 절차를 사전에 공개할 의무를 추가로 명문화하였다. 즉, 개인정보처리자는 자동화된 결정이 이루어진다는 사실과 그 목적, 주요 개인정보의 유형, 처리 절차, 민감정보 또는 아동정보의 처리 여부, 거부나 설명, 검토 요구 방법을 인터넷 홈페이지

4) 개인정보 보호법 시행령 [시행 2024. 3. 15.] [대통령령 제34309호, 2024. 3. 12., 일부개정] (1) 개정이유 - 개인정보 보호위원회가 공공기관 등의 개인정보 보호 수준을 평가할 수 있도록 하고, 정보주체가 완전히 자동화된 시스템으로 개인정보를 처리하여 이루어지는 자동화된 결정에 대해 거부하거나 설명 등을 요구할 수 있도록 하는 등의 내용으로 「개인정보 보호법」이 개정(법률 제19234호, 2023. 3. 14. 공포, 2024. 3. 15. 시행)됨에 따라, 개인정보 보호수준 평가의 대상 및 방법 등을 정하고, 정보주체가 자동화된 결정에 대해 거부하거나 설명 등을 요구하는 방법·절차 등을 정하며, 개인정보처리자가 자동화된 결정을 하려는 경우 정보주체에게 알릴야 하는 사항을 정하는 등 법률에서 위임된 사항과 그 시행에 필요한 사항을 정하는 한편, 개인정보 보호수준 평가를 받았거나 개인정보 보호 인증을 받아 고유식별정보의 안전성 확보 조치에 대한 점검이 이루어진 경우에는 고유식별정보 안전성 확보 조치에 대한 정기적인 조사를 실시한 것으로 보도록 하는 등 현행 제도의 운영상 나타난 일부 미비점을 개선·보완하려는 것임. (2) 주요 내용 - 라. 자동화된 결정에 대한 정보주체의 거부·설명 등 요구의 방법·절차 등 (제44조의2부터 제44조의4까지 신설) 1) 정보주체는 자동화된 결정에 대해 거부하거나 설명 등을 요구하려는 경우에는 개인정보처리자가 마련하여 공개한 방법·절차에 따르도록 하고, 개인정보처리자는 정보주체의 거부 및 설명 등의 요구에 대하여 필요한 조치를 하고 그 결과를 정보주체에게 알리도록 함. 2) 개인정보처리자는 자동화된 결정에 사용되는 주요 개인정보의 유형과 자동화된 결정 간의 관계 등에 관한 사항을 인터넷 홈페이지 등을 통해 공개하도록 함.

이지 등을 통해 쉽게 확인 할 수 있도록 공개해야 한다. 시행령에 대한 실제 대응 기준은 2024년 9월 개인정보보호위원회 고시를 통해 구체화되었다. 고시는 「개인정보 보호법」 제37조의2 및 동법 시행령 제44조의3에 근거하여,⁵⁾ 정보주체가 자동화된 의사결정에 대해 거부·설명·검토를 요구할 때 개인정보처리자가 취해야 할 구체적 조치 기준을 정한 것

5) 개인정보 보호법 시행령[시행 2025. 10. 2.] [대통령령 제35780호, 2025. 9. 23., 일부개정]

제44조의3(거부·설명등요구에 따른 조치) ① 개인정보처리자는 정보주체가 제44조의2제1항에 따라 자동화된 결정에 대해 거부하는 경우에는 정당한 사유가 없으면 다음 각 호의 어느 하나에 해당하는 조치를 하고 그 결과를 정보주체에게 알려야 한다.

1. 해당 자동화된 결정을 적용하지 않는 조치

2. 인적 개입에 의한 재처리

② 개인정보처리자는 정보주체가 제44조의2제2항에 따라 자동화된 결정에 대해 같은 항 제1호에 따른 설명을 요구하는 경우 정당한 사유가 없으면 다음 각 호의 사항을 포함한 간결하고 의미있는 설명을 정보주체에게 제공해야 한다. 다만, 개인정보처리자는 해당 자동화된 결정이 정보주체의 권리 또는 의무에 중대한 영향을 미치지 않는 경우에는 정보주체에게 제44조의4제1항제2호 및 제3호의 사항을 알릴 수 있다.

1. 해당 자동화된 결정의 결과

2. 해당 자동화된 결정에 사용된 주요 개인정보의 유형

3. 제2호에 따른 개인정보의 유형이 자동화된 결정에 미친 영향 등 자동화된 결정의 주요 기준

4. 해당 자동화된 결정에 사용된 주요 개인정보의 처리 과정 등 자동화된 결정이 이루어지는 절차

③ 개인정보처리자는 정보주체가 제44조의2제2항에 따라 같은 항 제2호에 따른 검토를 요구하는 경우에는 정당한 사유가 없으면 정보주체가 제출한 의견의 반영 여부를 검토하고 정보주체에게 반영 여부 및 반영 결과를 알려야 한다.

④ 개인정보처리자는 다른 사람의 생명·신체·재산과 그 밖의 이익을 부당하게 침해할 우려가 있는 등 정당한 사유가 있어 법 제38조제5항에 따라 거부·설명등요구를 거절하는 경우에는 그 사유를 정보주체에게 지체 없이 서면등의 방법으로 알려야 한다.

⑤ 개인정보처리자는 제1항부터 제3항까지의 규정에 따라 정보주체의 거부·설명등요구에 따른 조치를 하는 경우에는 정보주체의 거부·설명등요구를 받은 날부터 30일 이내에 서면등의 방법으로 해야 한다. 다만, 30일 이내에 처리하기 어려운 정당한 사유가 있는 경우에는 정보주체에게 그 사유를 알리고 2회에 한정하여 각각 30일의 범위에서 그 기간을 연장할 수 있다.

⑥ 제1항부터 제5항까지의 규정에 따른 정보주체의 거부·설명등요구에 따른 조치에 관한 세부 사항은 보호위원회가 정하여 고시한다.

이다. 이로써 자동화된 결정 제도의 법률, 시행령, 고시를 통해 체계적으로 구체화하였다. 또한 2024년 9월 개인정보보호위원회는 고시 시행과 함께 「자동화된 결정에 대한 정보주체 권리 안내서」를 배포하여, 거부·설명·검토권의 행사 절차와 각 조치의 예시를 현장 실무자가 이해할 수 있도록 제공하였다.

「개인정보 보호법」 제37조의2, 「시행령」 제44조의2~4, 그리고 개인정보보호위원회 「고시」 제2024-9호는 자동화된 결정에 대한 3단계 규율 구조(법률-시행령-고시)를 형성하였다. 법률은 기본권적 권리와 원칙(거부·설명·검토 요구권)을, 시행령은 자동화된 결정의 공개·절차 의무를, 고시는 구체적 판단기준과 개인정보처리자의 조치 절차를 규정함으로써, AI·빅데이터 기반 자동화 의사결정에서의 투명성·공정성·책임성을 실질적으로 구현하는 단계적 규범체계를 일응 마련하였다.

3. 자동화된 의사결정 규제에 대한 현행 법제의 한계

가. 개념 조항들의 불명확성

고시 제2조에서 ‘자동화된 결정’을 정의하고⁶⁾ 제6조에서 ‘중대한 영향’의 5가지 판단기준을 제시했으나, 실제 적용 시 기업 간 해석 편차가 발생할 수 있다(이상용·이혜리, 2023; 김연진, 2023). 특히 ‘완전히 자동화’라는 개념의 범위가 모호한데, 제6조 제1호에서 ‘정당한 권한을 가진

6) 자동화된 결정에 대한 개인정보처리자의 조치 기준
[시행 2024. 9. 26.] [개인정보보호위원회고시 제2024-9호, 2024. 9. 26., 제정]

제2조(정의) 이 고시에서 사용하는 용어의 뜻은 다음과 같다.

1. “자동화된 결정”이란 법 제37조의2제1항에 따라 완전히 자동화된 시스템(인공지능 기술을 적용한 시스템을 포함한다. 이하 같다)으로 개인정보를 처리하여 이루어지는 결정(「행정기본법」 제20조에 따른 행정청의 자동적 처분은 제외한다)을 말한다.
2. “거부·설명등요구”란 영 제44조의2제1항 및 제2항에 따른 정보주체의 자동화된 결정에 대한 거부, 설명 및 검토 요구를 말한다.

사람의 실질적이고 의미 있는 개입 없이’라고 규정했으나, ‘실질적 개입’의 구체적 범위가 법령에 명시되지 않음으로써 해석의 여지를 남기고 있다.⁷⁾ 예를 들어, 사람이 AI의 결정 결과를 형식적으로 승인만 하는 경우를 ‘자동화된 결정’에 포함할 것인가 하는 문제가 명확하지 않다. EU GDPR은 ‘인간의 의미 있는 개입(meaningful human involvement)’ 여부를 기준으로 하며(Goodman & Flaxman, 2017), 우리 고시도 유사한 기준을 제시했으나 ‘의미 있음’의 수준이 구체적으로 규정되지 않음으로써 기업의 자의적 해석을 초래할 수 있다.

더욱이 ‘중대한 영향’의 판단 기준도 분명하지 않다. 고시 제6조 제2항의 5가지 기준이⁸⁾ 동시에 충족되어야 하는가, 아니면 하나라도 충족

7) 자동화된 결정의 의미에 대하여 다음과 같은 개정안이 제출되는 등, 법제화되었음에도 구체적인 내용에 대해서는 여전히 논쟁이 있다.

개인정보 보호법 일부개정법률안(백혜련의원 대표발의)(의안번호 5918/ 발의연월일 : 2024. 11. 27.)은 “완전히 자동화된”을 “자동화된”으로 개정하는 내용을 담고 있다. 제안이유 및 주요내용은 다음과 같다. ‘현행법은 완전히 자동화된 시스템으로 개인정보를 처리하여 이루어지는 결정(이하 “자동화된 결정”이라 함)이 정보주체의 권리 또는 의무에 중대한 영향을 미치는 경우에 해당 개인정보처리자에 대하여 해당 결정을 거부할 수 있는 권리와 자동화된 결정에 대하여 설명 등을 요구할 수 있는 권리를 규정하고 있음. 이는 사람의 개입 없이 인공지능(AI) 등을 활용하여 개인정보를 처리하는 과정에서 발생하는 정보주체인 국민과 개인정보처리자 간의 정보 비대칭 문제를 해소하기 위하여 도입된 제도임. 그런데 인공지능을 활용하여 개인정보를 처리하는 경우가 급격하게 증가함에도 불구하고 현행법은 사람의 개입이 전혀 없는 자동화된 결정에 대하여만 정보주체의 거부권 등을 허용하고 있어 실질적으로 인공지능을 활용하여 처리되는 개인정보의 보호에 한계가 있다는 지적이 있음. 이에 완전히 자동화된 시스템이 아닌 어느 정도 사람의 개입이 이루어지면서 자동화된 결정이 발생하는 경우에도 정보주체가 거부권 등을 행사할 수 있도록 하여 정보주체의 개인정보자기결정권을 강화하고, 개인정보를 두텁게 보호하고자 함(안 제4조제6호 및 제37조의2 제1항).

8) 자동화된 결정에 대한 개인정보처리자의 조치 기준

제6조(거부·설명등요구에 대한 조치시 고려사항) ② 개인정보처리자는 정보주체의 거부·설명등요구에 따른 조치를 위하여 자동화된 결정이 정보주체의 권리 또는 의무에 중대한 영향을 미치는 경우에 해당하는지 여부를 판단할 때에는 다음 각 호의 사항 및 [별표]의 예시를 종합적으로 고려해야 한다.

1. 사람의 생명, 신체의 안전과 관련된 정보주체의 권리 또는 의무인지 여부
2. 정보주체의 권리가 박탈되거나 권리 행사가 불가능하게 되는지 여부
3. 정보주체가 통상적으로 받아들일 수 있는 한도를 넘어서는 의무가 발생하는지 여부

되면 되는가가 불명확하기 때문이다. 예를 들어, 신용평가 AI 결정의 거절은 권리 박탈(대출 불가)로 볼 것인가, 의무 부과(높은 이자율 적용)로 볼 것인가 하는 판단이 명확하지 않다. 이는 명확성 원칙 측면에서 문제가 될 소지가 있는데, 기본권 제한적 규정이므로 그 내용이 분명해야 한다는 헌법적 요청을 고려할 때 현행 규정만으로는 기업과 이용자 간 분쟁이 증가할 가능성이 있다.

통지·설명 의무 내용과 절차도 보다 구체화될 필요가 있다. 시행령 제 44조의3 제2항⁹⁾ 및 그에 따른 고시 제4조에서는 사용된 개인정보 항목, 주요 기준, 처리 절차를 설명하도록 규정했지만, 주요 기준이 정확히 무엇을 의미하는가가 불분명하다. 상위의 요인만 설명하면 충분한가, 전체 요인을 설명해야 하는가 하는 기준이 법령에 명시되지 않은 것이다. GDPR 제13조, 제14조에서는 ‘자동화 처리의 논리, 그 의미 및 예상되는 결과’에 관한 ‘유의미한 정보’를 명시하고 있으며, 이에 비추어보면 우리 제도는 ‘간결하고 의미 있는’ 설명을 요구하면서도 기업이 제공 범위를 최소화하여 형식적 설명에 그칠 여지를 남기고 있다. 더욱이 시행령 제 44조의4는 기업의 ‘사전 공시’ 의무만 규정하고 있고, 결정 후 개별 통지의 시점(즉시인가, ‘지체 없이’인가, 다음 접속 시인가)이 명확하지 않

-
- 4. 정보주체의 권리 또는 의무에 지속적인 제한이 발생하는지 여부
 - 5. 정보주체에게 영향을 미치기 전의 상태로 회복하거나 해당 영향을 회피할 수 있는 가능성이 있는지 여부

9) 개인정보 보호법 시행령 제44조의3 (거부·설명등요구에 따른 조치) ② 개인정보 처리자는 정보주체가 제44조의2제2항에 따라 자동화된 결정에 대해 같은 항 제1호에 따른 설명을 요구하는 경우 정당한 사유가 없으면 다음 각 호의 사항을 포함하여 간결하고 의미있는 설명을 정보주체에게 제공해야 한다. 다만, 개인정보처리자는 해당 자동화된 결정이 정보주체의 권리 또는 의무에 중대한 영향을 미치지 않는 경우에는 정보주체에게 제44조의4제1항제2호 및 제3호의 사항을 알릴 수 있다.

- 1. 해당 자동화된 결정의 결과
- 2. 해당 자동화된 결정에 사용된 주요 개인정보의 유형
- 3. 제2호에 따른 개인정보의 유형이 자동화된 결정에 미친 영향 등 자동화된 결정의 주요 기준
- 4. 해당 자동화된 결정에 사용된 주요 개인정보의 처리 과정 등 자동화된 결정이 이루어지는 절차

아 기업에게 해석의 여지를 남기고 있다.

나. 조치의 체계적 절차 보장 미흡

절차 보장의 측면에서도 몇 가지 중요한 한계가 드러난다. 우선, 검토 절차의 독립성이 명시적으로 확보되어 있지 않다. 고시 제5조는 검토의무를 규정하지만, 초기 결정을 내린 담당자와 다른 주체가 검토해야 한다는 규정은 존재하지 않아, 동일한 담당자가 스스로의 결정을 다시 확인하는 형식적 재검토에 그칠 위험이 크다.

기한 측면에서, 시행령 제44조의3 제5항에서 정하는 30일의 기한(최대 90일)은 단순한 사건에는 충분하지만, 복잡한 신용평가나 장기간의 데이터 분석이 필요한 사건의 경우 불충분할 수 있으며, 특수한 상황에서는 충분한 검토를 방해할 수 있다. 절차적 정당성 관점에서 더 중요한 문제는 내부 검토 이후에도 외부적 구제수단이 제도적으로 마련되어 있지 않다는 점이다(이상용·이혜리, 2023). 정보 비대칭·소송 비용·증거 확보의 어려움 때문에 실질적 접근성이 크게 저하될 수밖에 없다.

또 하나의 중요한 한계는 기록의무의 부재이다. 자동화된 결정의 적정성이나 차별 여부를 사후에 검증하려면 당시의 결정 과정 기록이 필수적이나 현행 법제에는 이러한 명시적 기록의무 규정이 전혀 없다. 시행령·고시에서 기록의무를 규정하지 않고 있으며, 개인정보보호법 제29조(안전조치의무)에서 접속기록 보관 등에 대한 일반적 규정만 존재할 뿐이고, 위험평가 및 상시관리, 서비스 운영 기록 보관을 명시하지 않고 있다. 기록의무가 없다는 것은 차별 발생 시 이를 입증할 수 없고, 감독기관의 조사가 어려우며, 기업의 사후 책임회피가 용이해진다는 문제를 낳는다.

감독기관의 검증 권한도 제한적이다. 현행 법제는 자동화된 결정의 공정성 검증을 위한 특화된 감사·검증 권한을 명시하고 있지 않으며, 기업이 알고리즘은 영업비밀이라고 주장할 경우 감독기관이 비공개 환경에

서 소스나 모델을 검사하도록 하는 규정도 없다.

다. 기업의 영업비밀 보호 등과의 비교형량

알고리즘 설명 의무와 기업의 영업비밀 보호 간의 충돌이 제도의 실효성을 크게 저해하고 있다. 고시 제7조에서는 기업이 영업비밀·지식재산권 침해 우려를 이유로 거절할 수 있도록 규정했지만, 어디까지를 영업비밀로 볼지 하는 기준이 불명확하다.¹⁰⁾ 구체적 경계가 애매하기 때문에 이를 악용할 여지가 있다. 실제로 기업이 알고리즘의 핵심 요소를 전부 영업비밀로 주장할 수 있다면, 실질적 설명 제공은 이루어지지 않을 가능성이 높다.

현행 시행령·고시에는 기업 규모별 차등 규정이 없어 모든 기업에 동일한 의무가 적용되고 있다. 직원이 5명인 스타트업도 직원이 5,000명인 대기업과 동일한 기록·검토·설명 의무를 부담해야 하며, 이러한 동일한 규제는 소기업으로 하여금 자동화 의사결정 시스템 도입을 포기하거나 기술 혁신을 미루도록 만들 수 있다. EU의 디지털 서비스법(Digital Services Act, DSA)은 월간활동사용자 4,500만명 이상의 대형 플랫폼에만 강화된 의무를 적용하고 있다.¹¹⁾

10) 제7조(거부·설명등요구의 거절) ① 개인정보처리자는 정보주체가 자동화된 결정에 대하여 거부·설명등요구를 한 경우라고 하더라도 정보주체의 요구에 따른 조치를 거절할 정당한 사유가 있는 때에는 해당 조치를 하지 않을 수 있다. 이 경우 정당한 사유가 있는지 여부는 정보주체의 자동화된 결정에 대한 거부·설명등요구를 제한하여 정보주체가 입게 될 불이익과 개인정보처리자 또는 제3자의 이익을 비교·형량(衡量)하여 개별적으로 판단해야 한다.

③ 개인정보처리자는 정보주체가 자동화된 결정에 대하여 설명을 요구한 경우 제4조에 따른 조치를 거절할 정당한 사유에 해당하는지를 판단할 때에는 다음 각 호의 사항을 종합적으로 고려해야 한다.

4. 자동화된 결정에 대한 설명으로 인해 개인정보처리자의 재산과 그 밖의 이익을 부당하게 침해할 우려가 있는 경우로서 정보주체의 자동화된 결정에 대한 권리보다 우선하는지 여부

11) Digital Services Act(2022)(DSA)는 “매월 4,500만 명 이상의 EU 내 활성 이용자” (월간 활성 사용자 10% 이상, 2023년 기준)인 플랫폼을 ‘초대형 온라인 플랫폼(very large online platforms, VLOPs)’ 및 ‘초대형 온라인 검색엔진(very large online

라. 기관 간 조율 부재로 인한 규제 공백

자동화된 의사결정 통지·대응 제도는 개인정보보호법에 근거하지만, 실제 플랫폼 이용자들이 겪는 알고리즘 문제는 콘텐츠 노출, 표현의 자유·경쟁 질서 등 복합적인 영역과 얽혀있다. 현재 개인정보보호위원회는 자동화 결정의 개인정보보호 측면만을 규제하고, 방송미디어통신위원회는 정보통신망법상 임시조치, 자율규제 등 플랫폼 이용자 보호 영역을 담당하며, 공정거래위원회는 알고리즘 부당성 문제를 다루고 있다. 그러나 이들 기관 간에 명확한 역할 분담이 이루어지지 않고 있다. 구체적인 예를 들면, 유튜브의 추천 알고리즘이 특정 정치 진영의 콘텐츠를 체계적으로 배제하는 경우를 살펴보면, 개인정보보호위원회는 ‘개인정보와 무관’이라고 판단할 가능성이 있고, 방송미디어통신위원회는 ‘자율규제 영역’으로 분류할 가능성이 있으며, 결국 어느 기관도 책임지지 않을 수 있는 상황이 발생한다. 규제의 공백 문제를 인식한 구방송통신위원회는 2023년 ‘알고리즘 투명성위원회’ 설립을 추진하여 포털 뉴스 알고리즘 공개 등을 논의했으나, 현재까지 법제화되지 않은 상태로 남아 있다. 이는 기관 간 조율의 부재와 함께, 새로운 영역에 대한 법적 근거의 미흡을 보여준다.

마. 검토

제37조의2를 통해 개인정보보호 법제로서 데이터 기반 이용자 권익을 도모했다는 점에서 진일보한 것이며, 거부·설명·검토 권리를 체계적으로 규정한 점도 긍정적이다. 더욱이 2024년 고시·시행령으로 기본적인 틀을 구축했다는 것은, 구체적 운영을 가능하게 했다는 의미에서 중요하

search engines, VLOSEs)’으로 지정하고, VLOP/ VLOSE로 지정된 플랫폼(예: Facebook, YouTube, Amazon 등)은 훨씬 더 엄격한 투명성, 리스크 관리, 사외 독립감사, 신고의무 등 차등 규제를 받고 있다.

다. 그러나 입법 불비 부분을 향후 법 개정을 통해 보완하지 않는다면 실효성 확보가 어려울 것이다. 용어 정의의 경계 명확화, 통지·설명 의무의 구체적 내용 및 절차의 법정화, 기록·감독·차등규제의 신설, 분쟁 조정 체계 구축, 그리고 기관 간 조율 메커니즘이 모두 필요하다. 정보주체와 기업 간의 실질적 이익형량을 통해, 투명성과 혁신 간의 균형을 모색해야 하며, 우리 사회의 맥락에 맞는 균형 잡힌 제도 설계가 중요하다. 다음 목차에서는 주요 해외 입법례를 살펴보고, 한국 제도 개선에 시사점을 줄 수 있는 요소들을 찾아보고자 한다.

Ⅲ. 자동화된 의사결정에 대한 정보주체 권리 및 의무의 비교법적 검토

유럽연합(EU), 캐나다 퀘벡주, 브라질의 사례를 통해 자동화된 의사결정에 대한 정보주체 권리 및 의무가 어떻게 규정되어 있는지 비교한다. 이들 지역은 모두 최근 개인정보 보호체계에 ADM 관련 조항을 도입·정비하여, 우리나라 2023년 개정「개인정보보호법」 제37조의2와 유사한 규범 경로를 보인다. EU의 일반개인정보보호규정(GDPR) 제22조는 “전적으로 자동화된 처리에 기초한 결정으로부터의 보호(법적 효과 또는 이에 준하는 중대한 영향)”를 핵심 권리로 규정한 기준 모델이며, 우리 입법에도 많은 부분 반영되었다. 퀘벡주는 이른바 「Loi 25(구 Bill 64)」를 통해 「민간부문 개인정보보호법(Private Sector Act, CQLR c. P-39.1)」을 전면 개정하면서, 자동화결정의 통지·설명 요구권과 인간에 의한 재검토 기회를 명문화하여 GDPR을 상회하는 권리 보강으로 평가된다. 브라질의 「일반개인정보보호법(LGPD)」 제20조는 전적으로 자동화된 처리에 기초한 결정에 대해 재검토 청구권과 기준·절차에 관한 명확한 정보 제공의무를 규정하여 GDPR과 유사하되 권리 행사 방식과 통지 범위에

서 규율 기법의 차이를 보인다.

1. EU GDPR 제22조 - 자동화된 의사결정에 대한 규제와 권리

EU GDPR 제22조는 전 세계적으로 ADM 규제의 준거가 되는 조항이다. 동조 제1항은 정보주체가 프로파일링을 포함하여 “자신에 관한 법적 효과(legal effects)를 초래하거나 이와 유사하게 자신에게 중대한 영향(similarly significantly affects)을 미치는 전적으로 자동화된 처리(solely automated processing)에 근거한 결정의 대상이 되지 않을 권리”를 가진다고 규정한다. 이는 정보주체에게 적극적인 청구권을 부여하는 형식이 아니라, 특정 요건 하의 자동화된 결정을 원칙적으로 금지하는 ‘면책권(right not to be subject to)’의 성격을 갖는다.

다만 제22조 제2항은 ① 정보주체와 데이터 컨트롤러 간의 계약 체결 또는 이행에 필수적인 경우, ② 컨트롤러에 적용되고 정보주체의 권리와 자유 및 정당한 이익을 보호하는 적합한 조치를 규정하는 EU 또는 회원국 법이 허용하는 경우, ③ 정보주체의 명시적 동의(explicit consent)에 근거하는 경우의 세 가지 예외를 인정한다(박노형 외, 2017). 이러한 예외 사유에 해당하여 자동화된 결정이 허용되더라도, 컨트롤러는 제22조 제3항에 따라 정보주체의 권리와 자유 및 정당한 이익을 보호하기 위한 적절한 조치를 취해야 한다. 여기에는 최소한 컨트롤러 측의 사람의 개입(human intervention)을 요구할 권리, 자신의 견해를 표명할 권리(express point of view), 결정에 대해 이의를 제기할 권리(contest the decision)가 포함된다.

아울러 GDPR은 제13조(2)(f) 및 제14조(2)(g)를 통해 투명성을 보완한다. 자동화된 의사결정이 이루어지는 경우, 컨트롤러는 정보주체에게 ‘관련된 로직에 대한 유의미한 정보(meaningful information about the logic involved)’와 ‘해당 처리가 정보주체에게 미칠 중요성 및 예상되는 결과’를 제공해야 한다. 또한 제15조(1)(h)는 정보주체의 열람권 행사 시 자동

화된 의사결정 및 프로파일링에 관한 “유의미한 정보” 제공을 재차 확인한다.

EU GDPR의 자동화된 의사결정 규율 체계는 단순한 권리 선언을 넘어, 규제적 완결성을 갖춘 4단계 구조로 요약될 수 있다. 첫째, 자동결정의 원칙적 배제(Prohibition)이다. GDPR 제22조 제1항은 법적 효과 또는 중대한 영향을 미치는 전적으로 자동화된 결정을 원칙적으로 금지하는 효과를 갖는다. 둘째, 예외적 허용 요건이다. 제22조 제2항은 계약의 체결 및 이행에 필수적인 경우, EU 또는 회원국 법률에 근거가 있는 경우, 정보주체의 명시적 동의가 있는 경우에 한하여 예외적으로 자동화된 결정을 허용한다. 셋째, 예외 적용 시 보호조치이다. 예외적으로 허용되는 경우라도 제22조 제3항에 따라 데이터 통제자는 정보주체의 권리와 자유 및 정당한 이익을 보호하기 위한 적절한 조치를 취해야 하며, 여기에는 최소한 인간의 개입을 요구할 권리, 자신의 견해를 표명할 권리, 결정에 대해 이의를 제기할 권리가 포함된다. 넷째, 정보제공 의무이다. 제13조, 제14조, 제15조는 자동화된 결정이 이루어지는 경우, 그 논리에 관한 유의미한 정보와 예상되는 결과를 정보주체에게 제공하도록 의무화하여 절차적 투명성을 담보한다. 이러한 규정이 시행된 이후 유럽 학계와 실무에서는 “전적으로 자동화된 처리”의 엄격한 해석, “유사하게 중대한 영향”의 적용 범위, 그리고 알고리즘의 복잡성을 고려한 “유의미한 정보” 제공의 구체적 수준 등을 둘러싸고 심도 있는 논의가 지속되고 있다(Sandra Wachter et al., 2017).

우리 「개인정보보호법」 제37조의2는 GDPR 제22조와 그 관련 규정의 직접적 영향을 받아 설계된 것으로 평가되지만(이상용·이혜리, 2024), 양자는 규율의 출발점과 입법 기술 측면에서 뚜렷한 차이를 보인다. GDPR은 자동화된 결정을 원칙적으로 허용하지 않는다는 ‘금지-예외’ 구조를 취한다. 즉, 사전적 통제를 원칙으로 하되 엄격한 요건 하에서만 허용하는 방식이다. 반면, 우리 법은 자동화된 시스템을 활용한 결정을 원칙적으로 허용하면서, 정보주체에게 거부권, 설명 요구권, 재검토 요구

권 등의 권리를 부여하는 ‘권리 부여’ 형식을 취한다. 이는 정보주체의 적극적인 권리 행사를 전제로 하는 사후적 권리 구제 중심의 구조라 할 수 있다. 또한, GDPR은 원칙적으로 사전 동의 없는 자동결정을 금지하는 것에서 출발하지만, 우리 법은 자동결정의 허용을 전제로 하되 사후적인 거부나 설명 요구를 통해 정보주체를 보호한다는 점에서 규범적 접근 방식에 근본적인 차이가 존재한다.

2. 캐나다 퀘벡주 「Loi 25」- 광범위한 통지·설명 의무와 인간개입

캐나다 퀘벡주는 2021년 「개인정보 보호 입법 현대화에 관한 법률 (Act to modernize legislative provisions as regards the protection of personal information)」(일명 ‘Loi 25’, 구 Bill 64)을 제정하여 공·사적 부문 개인정보보호 법제를 전면 개편하였다. 이 개정은 EU GDPR의 구조적 영향을 받았으나, 자동화된 결정(ADM) 영역에서는 일부 측면에서 GDPR보다 더 강한 소비자 권리 보호를 도입한 것으로 평가된다.

Loi 25는 사적 부문 개인정보보호법(Private Sector Act)에 제12.1조를 신설하여 자동화된 결정에 대한 투명성을 강화하였다. 동조 제1항에 따르면, “개인정보에 근거한 전적으로 자동화된 처리(exclusively automated processing)에 기초하여 개인에 관한 결정을 내리는 조직(enterprise)은, 그 결정을 당사자에게 알릴 때 늦어도 그 시점까지(not later than at the time it informs the person of the decision)” 해당 결정이 전적으로 자동화된 처리에 기초한다는 사실을 통지해야 한다. 즉, 조직은 자동화된 결정이 내려지는 시점 또는 그 직전에 “이 결정은 자동화된 처리에 의해 내려졌음”을 개별적으로 고지해야 한다. 이는 GDPR이 정보수집 단계(제13조 및 제14조)에서 자동화된 의사결정의 존재 여부를 일반적·포괄적으로 설명하도록 요구하는 것과 달리, 개별 결정의 시점에 맞춘 실시간 통지 의무를 부과한다는 점에서 한층 강화된 모델이라 할 수 있다.

제12.1조 제1항 및 제3항은 정보주체의 요청이 있는 경우, 조직이 ①

해당 결정에 사용된 개인정보, ② 결정을 내리게 된 이유(reasons)와 주요 요인·매개변수(principal factors and parameters), ③ 관련 개인정보의 정정(correction)을 요구할 권리를 제공해야 한다고 규정한다.¹²⁾ 이는 GDPR이 “처리의 논리에 관한 유의미한 정보” 제공을 요구하는 것에 비해, 설명의 범위를 ‘주요 매개변수’로 명시하여 구체화함으로써 정보주체의 이해가능성과 예측가능성을 높인 것이다. 또한, 동조 제2항은 조직이 정보주체가 “해당 결정에 관하여 의견을 제출하거나 질문 또는 불만을 제기할 수 있는 담당자(contact person)”를 지정해야 함을 규정한다. 비록 조문이 ‘인간 개입(human intervention)’이라는 용어를 직접 사용하지는 않지만, 정보주체가 사람과 직접 소통하며 설명을 요구하고 조직 내 담당 인력이 이를 재검토하게 함으로써, 실질적인 인간에 의한 재심사 채널을 제도화한 것으로 해석된다.

제12.1조는 적용 범위에 관해 GDPR과 달리 “법적 효과 또는 유사하게 중대한 영향”이라는 제한을 두지 않는다. 즉, “개인정보를 기반으로 배타적으로 자동 처리된 모든 결정”에 적용되며, 조문 자체에서 별도의 영향력 기준(threshold)을 설정하지 않는다. 이로 인해 기업은 상대적으로 경미한 결정에 대해서도 자동화된 처리가 이루어진다면 통지 및 설명 의무를 부담하게 된다. 한편, 자료 보존과 관련하여 같은 법 제11조

12) Act respecting the protection of personal information in the private sector
 12.1. Any person carrying on an enterprise who uses personal information to render a decision based exclusively on an automated processing of such information must inform the person concerned accordingly not later than at the time it informs the person of the decision.
 He must also inform the person concerned, at the latter’s request,
 1° of the personal information used to render the decision;
 2° of the reasons and the principal factors and parameters that led to the decision; and
 3° of the right of the person concerned to have the personal information used to render the decision corrected.
 The person concerned must be given the opportunity to submit observations to a member of the personnel of the enterprise who is in a position to review the decision.

는 “개인에 관한 결정을 내리기 위하여 사용된 정보는 그 결정이 내려진 후 최소 1년간 보존되어야 한다”고 규정한다. 이 규정은 자동·비자동 결정을 불문하고 적용되나, 특히 자동화된 결정에 있어 정보주체의 이의 제기나 감독기관의 사후 조사 시 핵심적인 근거자료를 확보하게 하는 법적 장치로 기능한다. 요컨대 퀘벡주의 접근은 기업에게 상당한 준수 비용을 발생시키지만, 결정 시점 통지의 명확화, 설명 내용의 구체화, 기록 보존 의무화 등을 통해 AI 시대의 투명성과 책임성을 확보하기 위한 강력한 규범 모델을 제시하고 있다.

3. 브라질 LGPD - 권리는 인정하나 인간개입 명시 없는 절충 모델

브라질의 「일반개인정보보호법(Lei Geral de Proteção de Dados Pessoais, 이하 ‘LGPD’)」은 2018년 제정(Lei n^o 13.709/2018) 이후 2019년 개정(Lei n^o 13.853/2019)을 거치며 정비되었다. 자동화된 결정에 대한 규율은 제20조에 집중되어 있는데, 동조 본문(caput)은 “데이터 주체는 전적으로 자동화된 개인정보 처리에만 의존하여 이루어진 결정(decisões tomadas unicamente com base em tratamento automatizado de dados pessoais)으로부터 발생하는 결과에 대하여, 자신의 이익에 영향을 미치는 결정에 대해 재검토(revisão)를 요청할 권리를 가진다”고 규정한다.¹³⁾

13) Article 20. The data subject is entitled to request the review of decisions made solely based on automated processing of personal data that affect his/her interests, including decisions intended to define his/her personal, professional, consumer and credit profile or aspects of his/her personality. (New wording given by Law No. 13,853/2019)

Paragraph 1. The controller shall provide, upon request, clear and adequate information on the criteria and procedures used for an automated decision, complying with trade and industrial secrets.

Paragraph 2. In the event of failure to offer the information set forth in paragraph 1 of this article based on the compliance with trade and industrial secrets, the national authority may carry out an audit to verify discriminatory aspects in the automated processing of personal data.

Paragraph 3. (vetoed). (Included by Law No. 13,853/2019)

이는 개인정보 처리에 관한 동의, 계약 이행, 정당한 이익 등 일반적인 적법 근거가 존재하면 자동화된 결정을 허용하되, 사후적으로 정보주체에게 재검토 요청권을 부여하는 방식이다. 즉, 사전적 금지와 예외적 허용 구조를 취하는 GDPR과 달리, LGPD는 자동화된 결정의 허용을 전제로 한 사후적 구제 중심의 모델을 채택하고 있다(IBA, 2024).

LGPD 제정 과정에서 가장 첨예한 쟁점은 재검토의 주체를 ‘사람’으로 명시할 것인가의 문제였다. 당초 의회를 통과한 법안에는 해당 재검토가 “자연인에 의한 재심(revisão por pessoa natural)”이어야 한다는 문구가 포함되어 있었다. 그러나 대통령 거부권 행사와 후속 입법 과정을 거치며 ‘자연인에 의한’이라는 문구가 삭제되었다. 그 결과 현행 LGPD 제20조는 인간 개입(human intervention)을 명시적으로 요구하지 않으며, 재검토 절차가 반드시 사람에게 의해 수행되어야 하는지, 아니면 다른 알고리즘에 의한 산출로도 족한지에 대해 범문상 모호성을 남기게 되었다. 이에 대해 인간에 의한 실질적 재평가 절차가 제도적으로 보장되지 않는다면 알고리즘 결정의 내재적 불투명성을 해소하기 어렵다는 비판이 지속적으로 제기되고 있다.

정보제공의무와 관련하여 LGPD 제20조 제1항은 데이터 통제자(controlador)가 정보주체의 요청 시 자동화된 결정에 사용된 “기준(criterion)과 절차(process)에 관한 명확하고 적절한 정보”를 제공해야 한다고 규정한다. 단, 이는 ‘영업비밀(trade secrets) 및 산업기밀’을 침해하지 않는 범위 내로 제한된다. 주목할 점은 제20조 제2항의 견제 장치이다. LGPD는 통제자가 영업비밀을 이유로 정보 제공을 거부하는 경우, 브라질 데이터 보호당국(ANPD)이 해당 결정에 차별적 요소가 포함되어 있는지 확인하기 위해 직접 감사(audit)를 수행할 수 있는 권한을 부여하였다. 이는 알고리즘의 투명성과 기업의 이익이 충돌할 때, 규제 기관이 개입하여 그 균형을 맞추도록 설계된 제도로서, 사법적 판단 이전에 행정적 감독을 통해 실효성을 확보하려는 시도로 평가된다.

LGPD는 GDPR과 달리 자동화된 결정 자체의 원칙적 허용·금지 여부

를 별도로 규정하지 않는다. GDPR이 ‘법적 근거가 없으면 금지’라는 금지-예외 구조를 취하는 반면, LGPD는 동의, 정당한 이익 등 일반적인 법적 근거가 존재하면 자동결정을 허용한다. 대신 재검토 요청권, 정보청구권 등 사후적 권리를 통해 통제장치를 두는 방식이다(IBA, 2024). 이러한 절충적 모델은 기업 입장에서는 유연성과 실무적 예측가능성을 제공하지만, 정보주체의 관점에서는 인간 개입 부재로 인한 권리보호의 미흡이라는 한계를 가진다. 브라질 사례가 시사하는 바는, 입법과정에서 인간 개입 조항이 누락되면 제도의 실효성이 떨어질 수 있다는 점, 영업비밀 예외에 대한 감독기관 통제장치의 중요성을 함께 고려해야 한다는 것이다.

4. 시사점

EU GDPR, 캐나다 퀘벡주 Loi 25, 브라질 LGPD의 입법례를 종합하면, 자동화된 의사결정을 규율하는 방식은 국가별로 상이한 규범적 지향점을 보인다. 퀘벡 모델은 통지·설명 의무와 재검토 절차, 자료 보존 의무를 강력하게 부과하는 ‘적극적 권리보호 모델’로, 브라질 모델은 자동결정을 폭넓게 허용하되 사후적 권리와 감독당국의 개입을 통해 균형을 맞추는 ‘완화된 절충 모델’로, GDPR과 우리 법은 그 중간 지점에서 금지 또는 권리와 예외를 교차시키는 ‘중간적 모델’로 유형화할 수 있다. 이러한 비교법적 분석을 토대로 우리 「개인정보보호법」 제37조의2 및 관련 제도에 주는 시사점을 정리하면 다음과 같다.

첫째, 정의 및 적용 범위의 구체화가 시급하다. 퀘벡주는 “배타적으로 자동화된 처리”라면 영향의 경중을 불문하고 적용하며, GDPR은 “법적 효과 또는 중대한 영향”이 있는 경우로 한정한다. 우리 법은 “권리 또는 의무에 중대한 영향을 미치는 경우”를 거부권의 요건으로 하고 있으나, 법률 차원에서 ‘중대한 영향’의 정의가 부재하여 법적 불확실성이 크다. 따라서 플랫폼의 추천·배열이 이용자의 정보 접근권이나 여론 형성에

미치는 영향이 ‘중대한 영향’에 포섭될 수 있는지, 구체적인 판단 기준과 예시를 제시하여 규범력을 확보해야 한다(이상용·이혜리, 2024).

둘째, 통지 시점의 실시간성 확보이다. GDPR이 정보 수집 단계에서 일반적인 사전 고지를 요구하는 것과 달리, 퀘벡 모델은 결정이 내려지는 시점(at the time of decision)에 “이 결정이 자동화된 것임”을 통지하도록 하여 정보주체의 ‘실시간 인지 가능성(real-time awareness)’을 보장한다. 한국 역시 개인정보 처리방침을 통한 포괄적 고지만으로는 부족하며, 신용평가, 채용, 대출 등 침익적 성격이 강한 결정에 대해서는 결과 통보 시 자동화 여부를 개별 안내하도록 의무화하는 방안을 검토해야 한다.

셋째, 설명 의무의 실질화이다. 설명의 내용에 있어 GDPR의 “논리에 관한 유의미한 정보”라는 표현은 다소 추상적이다. 반면 퀘벡법은 “결정에 사용된 데이터, 주된 이유, 주요 매개변수”를 명시하여 ‘사건 특유의 설명(case-specific explanation)’을 지향한다. 이는 정보주체가 자신의 구체적 상황을 이해하고 예측가능성을 확보하는 데 필수적이다(Wachter et al., 2017). 우리 법도 단순히 ‘처리 기준과 절차’를 나열하는 것을 넘어, 해당 결정이 왜 그렇게 도출되었는지에 대한 개별적 사유를 설명하도록 설명 요구권의 내용을 구체화할 필요가 있다.

넷째, 인간 개입의 제도적 담보이다. 자동화된 결정 규율의 성패는 형식적 재처리가 아닌 실질적 심사의 보장에 달려 있다. 브라질의 입법 과정에서 ‘자연인에 의한’ 문구가 삭제됨으로써 발생한 논란은 타산지석이 된다(IBA, 2024). 우리 법 제37조의2가 규정한 ‘재처리 요구권’이 실효성을 갖기 위해서는, 단순한 알고리즘 재구동이 아니라 ‘권한 있는 사람에 의한 검토, 사유 제시, 기록화’가 필수적 요건임을 하위 법령과 가이드라인에 명문화해야 한다(이상용·이혜리, 2024).

다섯째, 영업비밀과 투명성의 조화를 위한 감독기구의 역할 강화이다. 기업의 알고리즘 공개 거부와 이용자의 알 권리 충돌은 필연적이다. 브라질 LGPD는 영업비밀로 인해 정보 제공이 제한될 경우, 데이터보호당

국이 차별 여부를 감사할 수 있도록 하여 균형을 모색한다. 한국의 개인정보보호위원회 역시 기업이 영업비밀을 이유로 설명을 거부할 때, 비공개 절차를 통해 알고리즘의 적정성을 검증할 수 있는 조사 권한을 명시적으로 확보할 필요가 있다.

여섯째, 사후적 통제를 위한 기록 보존 의무화이다. 퀘백주가 의사결정 관련 정보를 최소 1년 이상 보존하도록 법제화한 것은 사후적 권리구제를 위한 물적 토대를 마련한 것이다. 알고리즘의 휘발성을 고려할 때, 결정 당시의 입력 데이터와 로직 상태가 보존되지 않으면 사후 검증은 불가능하다. 한국도 고시 수준의 권고를 넘어, 자동화된 결정 기록의 보존을 법적 의무로 격상시켜 책임성을 강화해야 한다.

요컨대, 비교법적 검토 결과는 한국의 자동화된 의사결정 규제가 적용 범위의 명확화, 실시간 통지 및 구체적 설명, 실질적 인간 개입, 기록 보존을 통한 책임성 확보라는 네 가지 축을 중심으로 보완되어야 함을 시사한다. 이는 단순한 개인정보 보호를 넘어, 디지털 플랫폼 시대의 정보인격권을 실질적으로 보장하기 위한 최소한의 법적 과제이다.

IV. 알고리즘 투명성과 플랫폼 책임에 대한 해외 규제 동향

개인정보 보호를 중심으로 자동화된 의사결정에 대한 권리 규율을 살펴해보았다. 앞서 논의가 자동화된 의사결정으로부터 개인을 보호하기 위한 권리 중심의 규범을 검토한 것이라면, 디지털 플랫폼 환경에서 알고리즘의 영향은 개개인뿐 아니라 사회 전체의 여론 형성, 정보 유통, 이용자 안전에 걸쳐 있다. 이에 최근 각국에서는 대규모 온라인 플랫폼을 겨냥한 알고리즘 투명성 제고와 콘텐츠 관리(content moderation) 절차의 공정성을 확보하기 위한 규범들이 등장하고 있다. 그중에서도 EU의 디지털서비스법(Digital Services Act, DSA), 영국의 온라인안전법(Online Safety Act, OSA), 호주의 eSafety 제도는 플랫폼의 투명성 및 보고의무

제도 측면에서 대표적인 모델로 평가된다. 이러한 법제들은 개인정보보호법과 직접적인 영역은 다르지만, “자동화된 결정으로부터 이용자를 보호한다”는 공통의 목적을 공유한다는 점에서 ADM 통지제도의 발전에 중요한 비교 기준이 된다.

1. EU 디지털서비스법(DSA)

EU의 디지털서비스법(Digital Services Act, DSA)은 불법·유해 콘텐츠의 확산 방지와 이용자 권리 보호를 목적으로 하는 포괄적 온라인 플랫폼 규제법이다. DSA는 플랫폼의 면책특권(safe harbor)을 전제로 하되, 알고리즘의 투명성과 책임성(accountability)을 강화하는 다양한 의무를 부과함으로써 ‘블랙박스’로 남아 있던 플랫폼의 의사결정 구조를 법적 규율 영역으로 포섭하였다. 특히 DSA는 콘텐츠 조치 통지, 내부 불복 절차, 추천 시스템의 매개변수 공개, 위험 평가 및 독립 감사 등을 통해 플랫폼의 콘텐츠 모더레이션(content moderation)과 알고리즘 시스템을 다층적으로 통제한다. 이는 개인정보보호법이 다루는 개인적 차원의 데이터 처리를 넘어, 미디어 플랫폼의 시스템적 위험을 관리하는 모델로서 중요한 참조점이 된다.

DSA 제17조는 플랫폼이 이용자의 게시물이나 계정에 대해 삭제, 차단, 노출 제한(demotion), 수익화 중지 등 ‘제한 조치(restrictive measures)’를 취하는 경우, 해당 이용자에게 “명확하고 구체적인 사유(clear and specific statement of reasons)”를 지체 없이 제공해야 한다고 규정한다.¹⁴⁾ 이 사

¹⁴⁾ Article 17 Statement of reasons

1. Providers of hosting services shall provide a clear and specific statement of reasons to any affected recipients of the service for any of the following restrictions imposed on the ground that the information provided by the recipient of the service is illegal content or incompatible with their terms and conditions:

(a) any restrictions of the visibility of specific items of information provided by the recipient of the service, including removal of content, disabling access to

유서에는 조치의 법적 근거, 조치의 구체적 내용과 범위, 의사결정 과정에 자동화된 수단이 사용되었는지 여부, 이용자의 이의제기 및 구제 수단이 필수적으로 포함되어야 한다. 이러한 정보는 동시에 EU 집행위원회가 운영하는 ‘DSA 투명성 데이터베이스(Transparency Database)’에 제출되어 공개된다. 이는 사실상 전 세계 플랫폼의 자동화된 콘텐츠 삭제·차단 결정에 대한 거대한 공적 통지 시스템으로 기능하며, 이용자의 알 권리를 보장함과 동시에 연구자와 규제 기관이 플랫폼의 자의적 검열이나 과도한 차단을 감시할 수 있게 한다.

DSA 제20조는 온라인 플랫폼에게 효과적인 내부 불만 처리 시스템(internal complaint-handling system)을 구축할 의무를 부과한다.¹⁵⁾ 이용자는 콘텐츠 조치 결정에 불복할 경우 이 절차를 통해 이의를 제기할 수 있으며, 플랫폼은 이를 차별 없이, 적시에, 자의적이지 않게 처리해야 한다. 주목할 점은 제20조 제4항이 “적절한 자격을 갖춘 인력의 감독 하에(under the supervision of appropriately qualified staff)” 결정을 처리하

content, or demoting content:

- (b) suspension, termination or other restriction of monetary payments;
- (c) suspension or termination of the provision of the service in whole or in part;
- (d) suspension or termination of the recipient of the service's account.

¹⁵⁾ Article 20 Internal complaint-handling system

1. Providers of online platforms shall provide recipients of the service, including individuals or entities that have submitted a notice, for a period of at least six months following the decision referred to in this paragraph, with access to an effective internal complaint-handling system that enables them to lodge complaints, electronically and free of charge, against the decision taken by the provider of the online platform upon the receipt of a notice or against the following decisions taken by the provider of the online platform on the grounds that the information provided by the recipients constitutes illegal content or is incompatible with its terms and conditions:
 - (a) decisions whether or not to remove or disable access to or restrict visibility of the information;
 - (b) decisions whether or not to suspend or terminate the provision of the service, in whole or in part, to the recipients;
 - (c) decisions whether or not to suspend or terminate the recipients' account;
 - (d) decisions whether or not to suspend, terminate or otherwise restrict the ability to monetise information provided by the recipients.

도록 규정한 것이다. 이는 완전 자동화된 처리에 의한 기각을 방지하고, 실질적인 ‘인간에 의한 검토(Human Review)’를 제도적으로 보장하는 조항으로 해석된다. 나아가 내부 절차로 해결되지 않을 경우, 이용자는 제 21조에 따라 독립적인 법외 분쟁해결기구(out-of-court dispute settlement body)를 선택하여 구제를 받을 수 있다. 이러한 다층적 구제 구조는 헌법상 적법절차(due process) 원칙을 사적 플랫폼 영역에 구현한 것으로 평가된다.

DSA 제27조는 ‘추천 시스템의 투명성(Transparency of recommender systems)’을 명문화하였다.¹⁶⁾ 플랫폼은 추천 알고리즘이 콘텐츠 순위를 결정하는 데 사용하는 주요 매개변수(main parameters)와 이용자가 그 매개변수를 수정하거나 영향을 미칠 수 있는 옵션을 약관에 명확히 설명해야 한다. 예컨대, 유튜브나 페이스북은 이용자의 시청 기록, 친구 관계, 위치 정보 등이 추천에 어떻게 반영되는지 설명하고, 이용자가 이를 쉽게 변경할 수 있는 기능을 제공해야 한다. 특히 초대형 온라인 플랫폼(Very Large Online Platforms, VLOPs)과 초대형 검색엔진(Very Large Online Search Engines, VLOSEs)에 대해서는 제38조를 통해 ‘프로파일링

¹⁶⁾ Article 27 Recommender system transparency

1. Providers of online platforms that use recommender systems shall set out in their terms and conditions, in plain and intelligible language, the main parameters used in their recommender systems, as well as any options for the recipients of the service to modify or influence those main parameters.
2. The main parameters referred to in paragraph 1 shall explain why certain information is suggested to the recipient of the service. They shall include, at least:
 - (a) the criteria which are most significant in determining the information suggested to the recipient of the service;
 - (b) the reasons for the relative importance of those parameters.
3. Where several options are available pursuant to paragraph 1 for recommender systems that determine the relative order of information presented to recipients of the service, providers of online platforms shall also make available a functionality that allows the recipient of the service to select and to modify at any time their preferred option. That functionality shall be directly and easily accessible from the specific section of the online platform’s online interface where the information is being prioritised.

에 기반하지 않은(non-profiling based)’ 추천 시스템 옵션을 최소화한 하나 이상 제공하도록 의무화하였다.¹⁷⁾ 이는 이용자가 알고리즘에 의한 맞춤형 필터버블(Filter Bubble)이나 확증편향의 굴레에서 벗어날 수 있는 ‘적극적 선택권(Opt-out)’을 보장하는 핵심 조항이다.

DSA는 개별적 권리 구제를 넘어, 알고리즘이 초래하는 사회적 위험을 통제한다. 제34조에 따라 VLOPs는 자신의 서비스가 불법 콘텐츠 확산, 기본권 침해, 선거 개입, 공중보건 위협 등의 시스템적 위험(systemic risks)을 초래하는지 매년 평가하고, 이를 완화하기 위한 조치를 취해야 한다(제35조). 또한 제37조는 이러한 의무 이행 여부에 대해 독립적인 외부 감사(Independent Audit)를 받도록 규정하며, 제40조는 학계 연구자들에게 플랫폼 데이터에 접근할 수 있는 권한을 부여한다. 이는 기업의 영업비밀 보호 논리를 넘어, 알고리즘의 사회적 영향력에 대한 민주적 통제와 투명성을 확보하기 위한 강력한 수단이다.

DSA 모델의 핵심은 플랫폼이 더 이상 알고리즘을 ‘블랙박스’ 뒤에 숨길 수 없으며, 이용자와 사회에 대해 설명할 책임(Accountability)을 진다는 점이다. 이는 직접적인 콘텐츠 검열을 지양하면서도, ‘절차적 투명성(Procedural Transparency)’ 강화를 통해 이용자의 권익과 표현의 자유를 보호하려는 접근이다. 현행 우리 「개인정보보호법」상의 ADM 대응권이 ‘개인 단위’의 권리 보호에 초점을 맞춘다면, DSA와 같은 플랫폼 규제법은 ‘시스템 단위’의 투명성과 위험 관리를 담당함으로써 상호보완적으로 작동할 수 있다. 즉, 개인정보보호법으로는 해결하기 어려운 여론 왜곡이나 혐오 확산과 같은 사회적 문제는 DSA식의 위험 평가와 투명성 보고서 의무를 통해 해결하고, 개별적인 차단이나 추천 배제는 ADM 통지

¹⁷⁾ Article 38 Recommender systems

In addition to the obligations laid down in Article 27, the Digital Services Act (2022) requires providers of very large online platforms and very large online search engines that use recommender systems to offer at least one option for each recommender system that is not based on profiling (Article 38(1) DSA (2022)). The notion of “profiling” follows the definition provided in Article 4(4) GDPR (2016).

및 이의제기권을 통해 해결하는 이원적 거버넌스(Dual Governance) 체계의 구축이 필요하다. EU DSA는 포괄적 온라인 플랫폼 규제법으로 불법·유해 콘텐츠의 확산 방지와 이용자 권리 보호목적으로 제정되었으며, 그 수단으로 플랫폼 알고리즘의 투명성과 책임성을 다양한 조문에서 요구하고 있다. DSA의 투명성 요건이 플랫폼의 콘텐츠 모더레이션 및 알고리즘 시스템의 구조와 영향력을 평가할 수 있도록 설계되었다.

2. 영국 「온라인안전법(Online Safety Act, OSA)」

영국의 온라인안전법(Online Safety Act)은 EU 탈퇴 이후 영국이 자국 기준에 따라 제정한 종합 온라인 안전 규제법이다. OSA의 목적은 불법 또는 유해한 온라인 콘텐츠(특히 아동에게 해로운 콘텐츠)의 확산을 방지하고, 플랫폼이 이에 대해 책임을 지도록 하는 데 있다. 이 법은 기술중립적 언어로 작성되었지만, 알고리즘이 이용자에게 유해 콘텐츠를 노출시키는 구조적 역할을 인식하고 이에 대응하는 여러 조항을 포함한다(Online Safety Act, 2023; UK Department for Science, Innovation & Technology, 2025; Online Safety Act Network, 2025).

OSA는 모든 플랫폼 사업자에게 서비스 내 불법 콘텐츠 및 아동 유해 콘텐츠 위험을 정기적으로 평가하고, 그 위험을 완화하는 조치를 취할 것을 의무화한다. 이때 사업자는 서비스의 알고리즘이 이용자의 유해 콘텐츠 노출에 미치는 영향을 구체적으로 고려해야 한다. 제공자는 알고리즘이 이용자의 불법 콘텐츠 노출에 어떤 영향을 미칠 수 있는지, 또한 아동의 유해 콘텐츠 노출에 어떠한 영향을 미칠 수 있는지를 위험평가의 일부로 명시적으로 고려해야 한다. 플랫폼이 스스로의 추천 알고리즘·프로파일링 시스템이 아동에게 폭력적 콘텐츠를 추천하거나, 성인에게 불법 콘텐츠 확산을 유발하고 있는지 자가 점검(self-assessment) 하도록 강제하는 것이다. 이 평가 결과는 규제기관 Ofcom이 감독하며, 필요시 기술적 수정이나 시스템 변경을 명령할 수 있다. OSA가 플랫폼의

설계·기능·알고리즘이 불법콘텐츠 차단 및 아동 안전 의무 준수에 미치는 영향을 평가하도록 요구한다.

OSA의 특징 중 하나는 성인 이용자에게 ‘합법적이지만 해로운 콘텐츠(legal but harmful content)’의 노출을 줄일 수 있는 제어 도구를 제공하도록 요구한 점이다. 이는 DSA의 ‘비프로파일링 추천 옵션’보다 한층 세분화된 방식이다. 대형 플랫폼은 자살·자해 조장 콘텐츠, 혐오표현 등 합법적이지만 해로운 콘텐츠 유형에 대하여 이용자가 원치 않으면 노출을 최소화할 수 있는 선택옵션(opt-out mechanism)을 사전에 제공해야 하고 미인증 사용자 콘텐츠 차단 기능(anonymous-user filtering)을 제공해야 한다. OSA가 “Safer Algorithms 조항”을 통해 중간 또는 고위험 서비스에 특정 콘텐츠 유형별로 추가적 조치를 부과한다.

OSA는 플랫폼에게 연례 투명성 보고서(annual transparency report) 제출을 의무화한다. 보고서에는 서비스 운영 방식과 알고리즘 사용 현황이 포함되어야 한다. 범주화된 서비스들은 알고리즘 사용 및 이용자 경험에 미치는 영향을 포함한 온라인안전 관련 정보를 담은 연례 보고서를 발행해야 한다. 특히 플랫폼은 스스로 설정한 약관(community guidelines)상의 내용을 일관되게 집행할 법적 의무를 진다. 또한 모든 온라인 서비스에 이용자 신고채널 및 불만처리 절차(complaint mechanism) 설치를 의무화한다. 이 불만에는 “플랫폼이 내 콘텐츠 또는 계정을 조치한 행위” 뿐 아니라 “알고리즘이 내 게시물을 부당하게 축소노출(deprioritisation)했거나 노출 자체를 차단한 경우”도 포함된다. 즉, 이용자는 “AI 필터링으로 인해 내 게시글이 후순위 처리되었다”고 주장하며 불만을 제기할 수 있고, 플랫폼은 이에 대해 명시적 답변을 제공해야 한다. 이러한 절차는 DSA의 내부 불복절차와 구조적으로 유사하며, 내부구제-외부구제(사법절차·Ofcom 조사)로 이어지는 다층적 권리보호 체계를 형성한다.

OSA는 EU DSA보다 규제 대상이 명확히 ‘콘텐츠 안전’에 집중되어 있고, 국가기관(Ofcom)의 감독 권한이 더 강력한 적극적 행정개입 모델이다. 그러나 알고리즘 투명성과 책임 측면에서는 DSA와 궤를 같이한

다. 특히 위험평가 과정에서 알고리즘의 영향을 명시적으로 고려하도록 한 점은, 향후 각국이 AI 영향평가(AI Impact Assessment) 제도를 도입할 때 중요한 선례가 될 것이다. OSA는 알고리즘 프로파일링과 프라이버시 보호 간의 균형을 모색하며 이용자 선택권 제도는 알고리즘 통제권을 개인에게 되돌려준다는 점에서 헌법상 개인 자율성 보장의 구현으로 볼 수 있다. 한국의 경우 정보통신망법상 게시물 임시 조치 제도나 방송미디어통신위원회의 시정 요구 절차 등 유사한 표현행위 피해 구제 장치가 있으나, 이는 주로 타인의 권리 침해 콘텐츠에 국한된다. 반면 OSA는 플랫폼 알고리즘 자체로부터 발생하는 이용자 권리 침해를 다룬다는 점에서, 향후 한국의 플랫폼 규제 논의에서 주목할 만한 방향을 제시한다.

3. 호주 eSafety와 BOSE

호주는 「온라인안전법(Online Safety Act, OSA)」을 제정하고, 세계 최초의 온라인 안전 전담 규제기관인 eSafety Commissioner(온라인안전국장)를 중심으로 유해물 대응 체계를 구축하였다. 호주의 모델은 EU나 영국과 같이 포괄적인 안전 규제 모델을 취하면서도, 전담 규제기관에게 광범위한 정보 요구권과 제재 권한을 부여하여 실효적 집행을 강조한다는 점에서 차별성을 가진다.

OSA는 ‘기본 온라인 안전 기대사항(Basic Online Safety Expectations, BOSE)’이라는 독특한 규범 체계를 도입하였다. 이는 플랫폼 운영자가 이용자 안전을 위해 합리적 조치를 취하고, 불법 콘텐츠를 최소화하며, 투명성을 확보해야 한다는 일련의 ‘기대사항’을 법령(Determination)으로 명시한 것이다. BOSE 자체는 추상적인 기대치로 보일 수 있으나, eSafety Commissioner가 이를 근거로 사업자에게 비정기·정기 보고를 요구하는 ‘법적 통지(Legal Notice)’를 발부할 수 있다는 점에서 강력한 구속력을 갖는다. eSafety는 2022년 애플, 메타, 마이크로소프트 등 7개 빅테크 기

업에 아동성착취물(CSEA) 대응을 위한 알고리즘 탐지 현황 제출을 요구하는 통지를 발부하였고, 그 결과를 분석하여 ‘투명성 보고서’로 대외에 공개하였다. 이는 기업의 자율적 공개에 의존하지 않고, 규제 기관이 필요한 시점에 특정 알고리즘 정보를 강제로 수집할 수 있는 권한을 행사한 사례이다.

OSA의 가장 큰 특징은 규제 당국의 정보 요구권이 실질적인 제재와 사법적 판단에 의해 뒷받침된다는 점이다. eSafety Commissioner는 2023년 2월, 트위터(현 X)가 아동 성착취물 대응 관련 투명성 통지에 대해 부실하게 응답했다는 이유로 61만 500호주달러(약 5억 원)의 벌금을 부과하고 민사소송을 제기하였다. 이와 관련하여 호주 연방법원(Full Federal Court)은 2025년 7월 31일 선고한 판결에서 규제 당국의 손을 들어주었다(X Corp v eSafety Commissioner, 2025; Federal Court of Australia, 2025).¹⁸⁾ 법원은 OSA 제56조 제2항에 근거한 투명성 통지 의무는 단순한 행정적 절차가 아니라 기업이 승계해야 할 ‘규제적 의무(regulatory obligations)’이며, 이는 Twitter Inc.에서 X Corp.로의 기업의 합병 후에도 소멸하지 않고 포괄적인 ‘법적 책임(liabilities)’으로 승계된다고 판시하였다. 이 판결은 플랫폼 기업이 법인 구조 변경이나 인수합병을 통해 알고리즘 투명성 의무를 회피할 수 없음을 명확히 하였으며, eSafety Commissioner의 정보 수집 및 제재 권한이 사법적으로도 유효함을 확인해 준 중요한 선례로 평가된다.

eSafety Commissioner는 법적 강제 수단 외에도 ‘추천 시스템 및 알고리즘’에 관한 입장문(Position Statement)을 발표하여, 추천 알고리즘의

18) 호주 연방법원은 eSafety 국장이 Online Safety Act 2021 (Cth) s 56(2)에 따라 트위터(Twitter Inc.)에 발부한 비정기 투명성 보고요구(Reporting Notice)의 이행을 촉구한 조치가 합병 후 X Corp에게 그대로 승계된다고 판단하면서, X Corp의 항소를 기각하였다. 본 판결은 Nevada Revised Statutes § 92A.250(1)(d)에 따른 ‘liabilities’의 범위가 규제적 의무(regulatory obligations)까지 포함됨을 확인하고, 온라인 플랫폼의 알고리즘·콘텐츠 관리 투명성 의무에 대한 행정기관(eSafety Commissioner)의 집행권한이 실질적으로 관철될 수 있음을 명확히 한 중요한 선례로 평가된다.

위험성을 경고하고 ‘안전 설계(Safety by Design)’ 원칙을 준수할 것을 권고하고 있다. 이는 알고리즘이 이용자의 체류 시간을 늘리기 위해 유해 콘텐츠를 증폭(amplification)시키는 것을 방지하기 위해, 설계 단계에서부터 위험 완화 조치를 포함해야 한다는 규범적 가이드라인이다. 호주는 이처럼 규제 기관이 알고리즘의 작동 원리에 대해 직접적인 설명을 요구하고, 미이행 시 즉각적인 제재로 압박하는 ‘행정 규제 중심의 모델’을 취하고 있다. 이는 개별 이용자의 권리 구제에 초점을 맞춘 개인정보보호법 중심 모델과는 달리, 사회 전체의 온라인 안전을 확보하기 위해 플랫폼의 상시적 투명성을 강제한다는 점에서 한국의 규제 거버넌스 설계에 시사하는 바가 크다.

호주의 사례는 플랫폼 알고리즘의 불투명성을 해소하기 위해서는 규제 기관에 실질적인 ‘정보 제출 요구권’이 부여되어야 함을 보여준다. 우리나라는 방송미디어통신위원회나 공정거래위원회가 시정조치 권한을 보유하고 있으나, 평상시 알고리즘의 위험성을 모니터링하거나 덤페이크, 혐오표현 등 특정 이슈 발생 시 즉각적으로 알고리즘 데이터를 요구할 수 있는 법적 근거는 미약하다. 따라서 향후 한국의 플랫폼 규제는 개인정보보호법상의 ‘개인별 ADM 통지권’과 별도로, 방송통신 규제 기관이 플랫폼의 알고리즘 운영 실태를 파악할 수 있는 ‘직권 조사 및 자료 제출 요구권’을 입법화하는 방향으로 보완될 필요가 있다. 이는 개인의 권리 구제와 사회적 안전망 확보라는 두 가지 목표를 동시에 달성하기 위한 통합 거버넌스의 핵심 요건이다.

4. 시사점

EU DSA, 영국 OSA, 호주 eSafety/BOSE는 개인정보보호법 영역을 넘어 플랫폼의 알고리즘 책임성(algorithmic accountability)을 구현하려는 시도라는 공통점을 가진다. 이들 규범은 표현의 자유, 이용자 안전, 공정 경쟁 등 다층적 공익을 겨냥하면서도, 결과적으로 알고리즘의 투명성과

인간 통제 가능성을 확대한다는 점에서 개인정보보호법상 ADM 통지·설명·재검토권과 지향을 공유한다. 특히 새도우 배닝(shadow banning), 검색 알고리즘 편향처럼 개인정보 규제와 콘텐츠 규제가 경합하는 영역에서는 중복 규제나 규제 사각지대를 피하기 위한 제도적 구획과 기관 간 협력 거버넌스가 중요하다. 개인정보보호법은 개인 단위의 ADM 통지·대응권을 보장하는 반면, 플랫폼 투명성 규범은 시스템 단위의 위험 평가·보고·감사를 담당한다. 양자는 상호보완적으로 설계되어야 하며, 이는 우리나라에서도 개인정보보호위원회와 방송미디어통신위원회 등 관계기관 간 협업을 통해 구현 가능하다.

디지털 시대 알고리즘 규율은 개인정보보호법의 틀을 넘어 디지털 서비스 전반으로 확장되고 있으며, 이는 헌법상 표현의 자유 및 절차적 투명성 원칙과도 긴밀히 연결된다. 이러한 맥락에서 자동화 의사결정 통지 제도는 플랫폼 투명성 규범과 조화롭게 결합될 때 실효성을 극대화할 수 있다. 즉, 개인정보 보호와 콘텐츠 투명성이라는 이원적 접근을 통합함으로써, 디지털 환경에서 책임 있는 알고리즘 거버넌스(responsible algorithmic governance)의 완성이 가능하다.

V. 자동화된 의사결정에 대한 규제의 헌법적 쟁점

자동화된 의사결정은 개인 데이터 처리의 영역을 넘어, 플랫폼이 이용자의 정보 접근·표현 기회·사회적 평판을 구조적으로 재편하는 핵심 메커니즘으로 작동하고 있다. 알고리즘에 의해 추천·배열·축소·차단되는 과정은 단순한 기술적 선택이 아니라, 공론장의 형성과 민주적 의사소통 환경에 실질적 영향을 미치는 사적 권력 행사에 해당한다. 이러한 맥락에서 ADM은 더 이상 개인정보 보호만의 문제로 이해될 수 없으며, 표현의 자유·정보접근권·평등권·절차적 기본권 등 복합적 영역과 교차되는 규범의 대상이 된다. 그럼에도 국내 법제는 ADM을 개인정보보호법

의 좁은 범위 안에서만 규율하고 있어, 실제 쟁점이 발생하는 플랫폼 편집권, 디지털 공론장 보호, 미디어 규제와의 결합 문제를 거의 반영하지 못하고 있다. 따라서 자동화된 의사결정 통지제도의 헌법적 평가 역시 단순한 개인정보자기결정권 보호 여부를 넘어, 사적 정보 권력에 대한 국가의 보호의무, 기본권 충돌 조정, 규제의 명확성과 비례성, 플랫폼 편집권과 표현의 자유 보장이라는 다층적 관점에서 종합적으로 검토될 필요가 있다.

1. 국가의 기본권 보호의무 및 대사인호

오늘날 플랫폼 알고리즘은 이용자의 정보 선택, 사회적 여론 형성, 정치·경제 활동에 광범위한 영향력을 행사하며, 이용자의 행위 데이터를 기반으로 정보의 흐름을 결정하고 특정 콘텐츠를 배제하거나 노출을 차단할 수 있는 실질적 권력을 형성한다. 알고리즘에 의해 추천·배열·축소·차단되는 과정은 정보 접근·선별·노출 순위를 조정하는 구조적 작용으로, 이는 단순한 기술적 선택이 아니라 사실상 디지털 공론장의 접근을 통제하는 ‘사적 권력(Private Power)’으로 기능한다(박진완, 2025). 이러한 상황에서 플랫폼이 상업적 이익이나 불투명한 기준에 따라 특정 콘텐츠의 노출을 배제(de-ranking)하거나 차별적으로 필터링할 경우, 이용자의 표현 기회와 정보접근권은 실질적으로 제약될 수 있다.

따라서 플랫폼의 편향적 추천·배열이나 불투명한 차단·노출 제한(Shadow Banning)으로 인해 이용자의 표현의 자유 및 정보접근권이 침해되는 상황을 국가가 방치할 경우, 이는 헌법상 ‘국가의 기본권 보호의무’를 위반할 우려가 있다. 자동화된 의사결정에 대한 법제는 단순한 행정 규제가 아니라, 알고리즘에 의한 구조적 기본권 침해로부터 디지털 공론장의 건전성을 방어하기 위한 헌법적 보호의무의 이행으로 이해되어야 한다. 즉, 자동화된 의사결정 규율은 공론장 참여권과 여론의 다양성을 보장하기 위한 헌법적 장치로 기능해야 하며, 정보주체의 개인정보

자기결정권을 보호하는 소극적 조치를 넘어, 국가가 실제로 취해야 할 ‘적극적 보호 조치’로서 정당성을 확보한다.

이처럼 자동화된 의사결정의 통지·설명·재검토 제도는 개인정보자기결정권의 보장을 넘어서, 디지털 공론장의 질차적 투명성을 확보하고 이용자의 표현의 자유를 실질화하는 헌법적 기능을 수행한다. 이는 디지털 환경에서 표현권 보장의 의미가 단순히 국가 권력 개입의 부정(소극적 자유)을 넘어, 사적 정보 권력의 통제에 대한 적극적 규범 개입으로 발전하고 있음을 시사한다. 따라서 ADM에 대한 제도의 도입과 준비는 디지털 공론장의 보호와 여론 다양성의 유지라는 관점에서 국가의 기본권 보호의무 실현을 위한 핵심 제도적 기초가 된다.

플랫폼 알고리즘의 편향성·차별성은 이용자의 정보접근권, 알 권리 및 여론 다양성에 직접적인 영향을 미치며, 특히 특정 정치·사회적 콘텐츠의 비노출이나 순위 하락은 사실상의 검열 효과(chilling effect)를 초래할 위험이 있다. 이 지점에서 자동화된 의사결정 통지제도는 단지 개인정보 보호 조치가 아니라, 국가가 플랫폼 기반의 정보 유통구조로 인해 침해되는 국민의 개인정보자기결정권·표현의 자유를 보호하기 위한 적극적 보호의무로 정당화된다.

이와 같은 문제는 알고리즘 편향의 성격에서도 확인된다. 알고리즘 편향은 단순한 기술적 오류가 아니라, 국가와 사적 플랫폼이 기본권 보장에 관하여 어떠한 적극적 보호의무를 부담하는가를 직접적으로 드러내는 문제이다. 기계학습 시스템은 역사적으로 축적된 데이터 속 편견을 그대로 학습하여 특정 성별·직업군·사회경제적 지위를 차별적 위험 요인으로 구성할 수 있다. 예컨대 언어 데이터 기반 성역할 분석이 남성은 ‘프로그래머’, 여성은 ‘가정 내 돌봄 및 보조 직업’으로 자동 분류된 사례는, 데이터 편향이 곧 헌법 제11조의 실질적 평등 보장에 반하는 결과로 이어질 수 있음을 보여준다(Bolukbasi et al., 2016). 또한 범죄 예측 모델이 ‘가난한 사람’일수록 범죄 위험 점수가 높게 산출된다는 연구는, 알고리즘이 사회적 낙인을 자동화하며 형사 절차 전반에 평등권 및 적

법절차의 위기를 초래할 수 있음을 시사한다(Oswald et al., 2018). 이러한 편향은 소비자 가격·추천 알고리즘 영역에서도 동일하게 나타난다. 이커머스 플랫폼이 이용자의 정보에 따라 가격을 차별적으로 설정한다는 연구는, 사인 간 불평등이 알고리즘을 매개로 구조화됨으로써 개인정보자기결정권과 평등권이 결합된 새로운 헌법적 문제영역을 구성함을 보여준다(Hannak et al., 2014). 나아가 이른바 ‘공정 알고리즘’ 조차 차별을 지속시킬 수 있다는 비판은, 공정성 정의 자체가 다원적이고 일률적 기준 설정이 불가능하다는 연구결과와 맞물려 기술적 공정성만으로는 헌법적 평등 보장의 기준이 될 수 없음을 시사한다.

결국 표현의 자유 영역에서도 알고리즘이 콘텐츠 노출·추천·선별을 통제함으로써 특정 직업·성별의 사회적 이미지와 명예를 왜곡할 위험이 존재한다. 이는 표현의 자유가 단순히 국가 검열을 부정하는 소극적 자유권 차원을 넘어, 사적 권력에 의해 왜곡되는 정보 구조에 대한 적극적 보호의무가 요구되는 단계로 이행하고 있음을 나타낸다. 따라서 알고리즘 편향에 대한 규범적 통제는 형식적 중립성을 넘어, 헌법적 평등·절차적 기본권·개인정보자기결정권·사적 권력 통제가 결합된 구조적 접근이 필요하다.

2. 기본권 충돌의 조화와 비례원칙-절차적 투명성을 통한 균형

자동화된 의사결정에 대한 규제를 둘러싸고 다양한 기본권들이 충돌한다. 정보주체의 권리로는 인격권, 개인정보자기결정권, 평등권, 알 권리, 행복추구권 등이 포함되며, 특히 개인정보자기결정권은 개인정보의 수집·이용 여부뿐 아니라 자동처리 과정 자체에 대한 통제권을 포함한다(박상돈, 2017; 김희정, 2022). 반면 개인정보처리자인 플랫폼 사업자는 표현의 자유, 직업수행의 자유, 재산권, 기업경영의 자유 등을 주장할 수 있다. 알고리즘은 기업 경쟁력의 핵심 자산이므로 그 강제 공개는 영업비밀 침해에 해당할 수 있으며, 추천 여부와 배열 방식은 플랫폼 서비

스 설계의 본질적 영역에 해당하므로 국가 개입은 기업 활동의 자유를 제한할 수 있다. 또한 이용자 중 콘텐츠 제작자의 경우, 알고리즘이 특정 콘텐츠를 축소하거나 숨기는 방식으로 노출을 통제하면 표현의 자유가 침해될 수 있으므로, 알고리즘 통제는 창작자 권리 보호와도 연결된다.

이와 같은 충돌을 헌법적으로 조정할 때, 핵심 축은 개인정보자기결정권 등 정보주체의 인격권 보호와 사업자의 경제적 자유 간의 조화를 이루는 것이다. 개인정보자기결정권은 헌법 제10조·제17조에 기반한 중요한 기본권이며, 반면 재산권(제23조)과 직업의 자유(제15조)는 공공복리에 따라 제한 가능한 자유권이다. 나아가 플랫폼 사업자의 표현의 자유(제21조) 역시 고려해야 하는데, 플랫폼의 배열과 추천이 직접적 표현 행위인지, 이용자 표현에 대한 매개 행위인지에 따라 자유의 범위가 달라질 수 있다. 표현의 자유는 민주주의 핵심권리이므로 플랫폼 자체의 자유라 하더라도 공론장에서의 정보접근 구조와 불가분적으로 연결된다.

결국 입법자는 필요 최소한의 규제를 통해 개인정보자기결정권을 실질적으로 보호하면서 기업의 자유 제한을 최소화해야 하며, 이를 위해 영업비밀 예외 인정, 스타트업·중소기업 대상 완화, 기업에게도 절차적 권리 부여 등이 고려될 수 있다. 이와 같은 기본권 조정은 과잉금지원칙에 따라 목적의 정당성, 수단의 적정성, 피해 최소화성, 법익 균형성을 충족해야 한다. 이때 중요한 점은 자동화 통지제도가 알고리즘의 소스코드 공개나 추천 순서 변경을 강제하는 ‘내용 규제’가 아니라, 결정에 이르는 논리와 주요 요인에 대한 설명권 및 이의제기 절차를 보장하는 ‘절차적 투명성(Procedural Transparency)’에 기반한 규제라는 점이다. EU 디지털 서비스법(DSA)이나 영국 Online Safety Act가 알고리즘 구성 자체 공개 대신 설명책임·위험 평가 의무를 부과하는 것도 이와 같은 균형을 도모하기 위함이다.

자동화된 의사결정 통지·설명·재검토 제도는 개인정보자기결정권 보호를 넘어서, 디지털 공론장의 절차적 투명성과 이용자의 표현의 자유 실현을 위한 헌법적 장치를 제공한다. 플랫폼 사업자의 재산권과 영업

자유는 전면적으로 제한되는 것이 아니라, 국가의 보호의무 이행과 공정한 정보질서 확립이라는 공익을 위해 비례적으로 조정되는 것이다. 정책적으로도 규제 대상과 범위를 명확히 하여 과도한 부담을 방지하고, 특히 스타트업·중소기업에 대한 차등 규제가 필요하다. 또한 영업비밀 침해 우려가 있는 경우, 설명 의무의 범위를 소스코드가 아닌 데이터 항목·요인 수준으로 한정하는 방식 등이 고려될 수 있다. 즉 알고리즘 투명성과 영업비밀 보호 간 균형 조정은 AI 규제의 핵심 과제이다.

3. 명확성 원칙 및 예측가능성의 확보

법치국가의 원리로부터 도출되는 명확성 원칙은 규제의 수범자가 자신의 행위를 결정할 수 있도록 법규범의 내용을 명확히 규정할 것을 요구한다. 국민의 권리·의무에 영향을 미치는 법규범은 누구라도 이해할 수 있을 정도로 예측 가능하게 규정되어야 하며, 이는 법치주의의 핵심 요소로서 특히 형벌법규나 표현의 자유 제한 법규에서 더욱 엄격히 요구된다(성낙인, 2024). 기본권 제한 법규범의 내용이 불명확한 경우 국민은 자신의 행위 기준을 예측하지 못해 위축 효과를 경험할 수 있으므로, 법률 조항으로부터 무엇이 금지되고 무엇이 허용되는지를 수범자가 구체적으로 인식할 수 있어야 한다(장영수, 2022).

자동화된 의사결정에 대한 규율은 플랫폼의 추천·배열·비노출 등의 정보 유통 구조에 영향을 미치므로, 이는 간접적으로 표현의 자유 영역에까지 미치는 규제로 평가될 수 있다. 따라서 이러한 규율에 대한 명확성 요구 수준은 일반적인 개인정보 규제보다 더욱 엄격하게 적용될 필요가 있다. 특히 알고리즘 운영에 관한 의무 부과는 플랫폼의 편집적 기능과 직접 결부될 수 있으므로, 표현의 자유와의 충돌 가능성을 고려할 때 예측 가능성이 확보되어야 한다.

명확성 원칙은 플랫폼 사업자 측면에서도 규제 정당성 및 집행 가능성 확보를 위해 중요한 의미를 가진다. 규제 내용이 불명확하면 기업들

은 어느 정도의 준수 의무가 요구되는지 판단하지 못해 과도하게 위축되거나, 반대로 규제를 무시하는 방식으로 대응할 수 있다. 이는 법 집행의 예측가능성을 떨어뜨릴 뿐 아니라, 자의적 집행 가능성을 높인다. 따라서 자동화 의사결정 규제의 내용은 의무의 구체성을 법률 차원에서 명확히 규정함으로써 입법의 정당성을 확보해야 한다.

다만 명확성 원칙을 지나치게 엄격하게 적용하면 입법의 기술적 유연성이 저해될 우려 또한 존재한다. AI 분야는 기술 변화 속도가 매우 빠르며, 알고리즘 설계 방식이나 데이터 처리 유형이 지속적으로 변화하므로, 일정 범위의 기술중립적 표현과 포괄적 규율은 불가피하다(이원우 외, 2019). 그러나 현재 우리 법 체계는 고시·가이드라인 등 하위규범에 과도하게 의존하고 있으며, 이는 법률유보 원칙 및 예측가능성의 관점에서 보완이 요구된다. 따라서 자동화 의사결정 규제는 기술 중립성과 명확성 간의 균형을 달성하기 위해, 법률에 기본권 제한 요건과 절차적 보장 원칙을 분명히 하고 세부 사항은 하위규범에서 발전시켜 나가는 방식으로 설계될 필요가 있다.

4. 디지털 공론장과 표현의 자유

알고리즘에 의한 추천·배열·축소 노출(deprioritization) 및 차단은 오늘날 사실상의 미디어 편집행위로 기능한다. 통제받지 않는 플랫폼 알고리즘은 콘텐츠 접근 구조를 자의적으로 통제할 수 있으므로, 이는 국가에 의한 검열이 아닌 사적 검열(Private Censorship)의 위험을 초래한다. 만약 국가가 이러한 ADM 작용을 규제 영역 밖에 두게 되면, 오히려 거대 플랫폼의 편집권 행사로 인해 개별 이용자의 표현의 자유가 침해되는 역설적 상황이 발생한다. 따라서 플랫폼의 ‘편집의 자유’를 절대시할 것이 아니라, 공론장 형성에 중대한 영향력을 행사하는 정보 권력에 대해 절차적 규제를 정당화하는 헌법적 근거가 마련되어야 한다. 이는 헌법 제21조가 금지하는 사전검열과 충돌하지 않는다. 자동화 통지제도는

콘텐츠 내용 심사가 아니라, 차단·축소·배열의 절차를 투명화하고 이에 대해 다룰 수 있는 권리를 보장하는 장치이기 때문이다. 이러한 절차 보장은 오히려 표현의 자유 실현을 위해 필수적이다.

자동화된 의사결정 통제는 단순한 기술 규제가 아니라 표현의 자유와 정보접근권을 실질적으로 보장하기 위한 새로운 형태의 헌법적 장치로 이해되어야 한다(Diana Treščáková, 2024). 자동화된 추천·배열·노출·차단은 개인정보 처리의 영역을 넘어, 포털 뉴스 배열, 유튜브·틱톡의 추천 구조, AI 기반 콘텐츠 필터링 등과 같이 사실상 언론의 편집 기능(editing function)을 수행한다. 이러한 변화에 따라 플랫폼 알고리즘의 편집적 영향은 이용자의 정보환경 및 여론 형성에 직접적 효과를 미치기 때문에, 국제적으로도 알고리즘 편집성을 인정하는 논의가 확대되고 있다(Balkin, 2021).

헌법 제21조가 보장하는 표현의 자유는 단지 개인의 표현뿐 아니라, 언론의 자유까지 포괄하는 개념이다. 전통적으로 편집행위는 신문·방송 사업자 등 언론기관에 귀속된 자유로 이해되어 왔으나, 오늘날 뉴스·여론의 상당 부분이 플랫폼 알고리즘에 의해 자동적·구조적으로 재편되면서 플랫폼 역시 사실상 언론 기능을 수행하고 있다는 평가가 제기된다(Balkin, 2018; Keller, 2021). 따라서 알고리즘에 대한 규제가 편집의 자유 침해라는 비판을 받을 수 있지만, 플랫폼이 스스로를 언론기관으로 인정하지 않으면서도 자동화된 편집을 통해 공론장을 사실상 지배하는 현실에서는 기존의 표현의 자유 이론을 디지털 맥락에 맞게 재구성할 필요가 있다.

플랫폼 알고리즘의 편향성·차별성은 이용자의 정보접근권 및 알 권리, 여론 다양성에 직접적인 영향을 미치며, 특정 정치·사회적 콘텐츠의 비노출 또는 순위 하락은 사실상의 검열 효과(chilling effect)를 유발한다. 이 지점에서 자동화된 의사결정 통지제도는 개인정보보호 조치가 아니라, 국가가 디지털 기반 정보 유통구조로 인해 침해되는 국민의 표현의 자유와 개인정보자기결정권을 보호하기 위한 적극적 보호의무 이행으로

정당화될 수 있다.

비교법적으로도 이러한 방향성은 확인된다. EU 「디지털서비스법(DSA)」, 호주 「Online Safety Act」, 캐나다 Online Harms Safety 체계는 모두 알고리즘 투명성 공개, 프로파일링 기반 위험 평가, 이용자 대상 설명 및 이의제기 절차를 핵심적 의무로 부과한다. 이들 규범은 개인정보보호법적 권리 보호를 넘어 플랫폼 책임(platform accountability)을 통해 표현의 자유 및 정보접근권을 보호하려는 구조이며, 각국은 이를 헌법상 이용자 권리 보호 차원에서 정당화하고 있다. 이러한 비교법적 규범들은 한국 또한 플랫폼의 자동화된 편집행위에 대한 공적 통제를 “이용자 권리 보호”의 관점에서 구축해야 함을 시사한다.

VI. 우리나라 법제의 개선 방향 및 정책 제언

국내외 법제 비교와 헌법적 검토를 통해 드러난 쟁점들을 토대로, 한국의 자동화 의사결정 통지·대응권 제도를 보다 실효성 있고 헌법적으로 정합하게 만들기 위한 입법·정책 개선안을 제시한다. 핵심 개선 방향은 주요 개념 정의의 법정화, 통지·설명 의무요소의 표준화(법률 명문화), 재심(인간 개입) 절차의 요건화, 기록의무 및 검증 인프라 구축, 거버넌스 정비 및 기관 간 조정과 아울러 영업비밀 보호와 소규모 사업자 부담 경감을 위한 조치도 함께 제안한다.

1. 주요 개념 정의의 법정화

개인정보보호법 제37조의2는 여러 핵심 개념을 포괄적 문언으로 규정하거나 위임하고 있다. 예를 들어 ‘완전히 자동화된 결정’, ‘중대한 영향’, ‘의미 있는 설명’ 등의 용어가 법집행 단계에서 불명확성 논란을 일으킬 소지가 충분하다. 입법자는 이러한 개념들을 법률에 보다 구체적으

로 규정하거나 최소한 법률에 위임 범위를 명확히 적시할 필요가 있다. GDPR도 ‘프로파일링’ 등 용어 정의를 규정(제4조 제4항)에 두고, ‘법적 효과 또는 유사한 중대성’도 해석지침(Recital 71)에서 어느 정도 설명해 주고 있다(Wachter et al., 2017).

자동화된 결정이 어떻게 이루어져야 하는지 명확하게 알 수 없고(이상용·이혜리, 2024), 프로파일링에 대해서 명확하게 규정하고 있지 않다(박노형·김규동, 2024). 이는 집행과정에서 혼선을 불러올 수 있으므로, 법률 본문 또는 별도 정의조항에 개념을 규정해야 한다. 정당한 권한을 가진 사람의 실질적이고 의미 있는 개입 없이, 인공지능 기술 등 자동화된 방법으로 개인정보를 처리하여 개별 정보주체에 대해 행해지는 결정 등으로 의미를 구체화할 필요가 있고, 프로파일링류 처리에 대해서도 명확하게 규정할 필요가 있다(박노형·김효권, 2022). ‘정보주체의 권리 또는 의무에 중대한 영향’은 법률에 직접 열거하기 어렵더라도, 위임 규정으로 대통령령에 판단기준 위임을 명시하거나, 최소한 설명적 정의를 두어야 한다. ‘권리 또는 의무에 중대한 영향을 미치는’이라 함은 생명·신체, 자유 또는 재산 등 정보주체의 중대한 법익에 관한 권리·의무에 지속적이거나 돌이킬 수 없는 영향을 주는 것을 의미함을 포괄적으로라도 규정하며 해석 기준으로 활용될 수 있다(김연진, 2023). ‘설명’에 대하여 정보주체가 요구할 수 있는 설명의 내용을 법률에 명시하여 ‘해당 자동화된 결정에 사용된 개인정보 항목, 결정에 이르게 한 주요 요소 및 그 논리의 개요를 정보주체가 이해하기 쉬운 형태로 제공하는 것을 말한다’ 등으로 구체화할 필요가 있다. 현재는 고시에서만 ‘간결하고 의미 있는 설명’이라고 추상적으로 규정하고 있다. 재심 요구 시 수행되는 ‘검토’의 성격을 법률에 인간에 의한 재평가를 포함한 검토라고 명시하여 LGPD와 같이 인간이란 말이 없으면 기업이 AI로 재검토할 위험이 있으므로, 이를 방지할 필요가 있다.

이러한 정의 조항들은 법률의 명확성을 높여주며, 기업과 이용자에게도 예측가능성을 준다. 다소 기술적 내용이라 입법과정에서 충분한 숙의

가 필요하지만, GDPR 등 해외입법을 참고하면 충분히 도출 가능하다. 특히 AI 기본법 등과 정합성도 고려해서, 가능하면 AI 기본법의 정의와 상충되지 않게 해야 한다. 개인정보보호법에서 정의를 명확히 해두면, 다른 법으로 파생될 때도 기준점이 될 것이다.

2. 통지 및 설명 의무 요소의 표준화

현행 법령 체계에서 통지·설명 의무의 구체적 내용은 시행령과 고시에 위임되어 있으나 개인정보자기결정권의 본질적 내용을 형성하는 부분으므로, 법률유보의 원칙상 중요한 사항은 가급적 법률에 직접 규정할 필요가 있다. 특히, 정보주체에게 언제 통지해야 하는지(결정과 동시에, 또는 사전에 등), 어떤 방식(개별통지나 일반공개 등)을 허용하는지에 대해, 통지 시점과 방식을 명확하게 규정할 필요가 있다. 쉼백의 예와 같이 ‘결정이 행하여질 때 또는 통보될 때’라고 규정한다면, 보다 실효성있을 것이다. ‘통지’에는 해당 결정이 자동화 처리로 이루어졌다는 사실과 정보주체가 가지는 권리(거부·설명요구권 등)를 명시하여야 한다고 규정하며, 설명 요구 시 제공할 내용으로 해당 결정의 주요 원인과 고려된 정보, 그 결정의 의미와 예상 결과 등을 규정하도록 고려할 필요가 있다. 정보주체가 거부권을 행사한 때에는 개인정보처리자는 지체 없이 해당 자동화된 결정의 효력을 정지하거나 그 적용을 배제하기 위한 합리적 조치를 취하여야 한다는 규정을 두어, 기업이 아무 조치를 취하지 않고 시간을 끄는 상황을 방지할 필요가 있다. 또한 재심 절차에 있어, 개인정보처리자는 정보주체의 검토 요청을 받은 경우, 사람에 의해 해당 결정의 타당성을 재검토하고 그 결과를 통지해야 한다. 이 경우 검토에는 정보주체가 제출한 의견을 고려하여야 한다는 방식으로 절차의 핵심 사항을 법률에 규정해야 한다. 이러한 사항들은 시행령이나 고시에서 다루고 있으나 이를 법률 단계에서 그 중요사항을 규정하여 규범력을 갖게 할 필요가 있으며, 특히 기업들이 법률에서 명시한 의무로 준수를 제고

해야 한다. 통지·설명 의무위반에 대한 제재 규정도 명확하게 규정해야 한다.

표준양식 개발도 고려할 만하다. 법률에는 아니더라도, 개인정보보호 위원회가 가이드라인 형태로 ‘ADM 통지 표준양식’을 제공하면 기업들이 참고해 쉽게 적용할 수 있다. 안내서에서 대략의 내용을 안내하고 있으나 충분하지 않고, 이러한 표준 양식의 제공은 이용자의 실질적 이해도를 높이기 위해 필요하다(박미사, 2024). 이용자에 제공될 정보의 종류와 수준을 법률로 표준화하고, 그 이행을 확보할 장치를 만드는 것이 개선 과제이다. 이는 국내 기업 간, 외국 기업 간 규칙의 통일성을 가져와, 규제의 예측가능성과 공정경쟁 환경을 조성할 수 있다.

3. 재심(인간 개입) 절차의 실효성 확보

인간에 의한 재검토 절차는 ADM 규제의 핵심이라고 할 수 있다. 그러나 현행 우리 제도에서는 구체성이 떨어져 형식화될 우려가 있다. 이를 개선하기 위해 몇 가지 최소 요건을 명문화할 필요가 있다. 첫째, 독립성의 보장으로 재검토는 가급적 초기 결정에 관여하지 않은 사람이 맡도록 권장 또는 의무화한다. 예를 들어 ‘개인정보처리자는 정보주체의 검토 요구에 대하여 해당 자동화결정의 산출에 책임이 없는 사람으로 하여금 검토를 수행하게 해야 한다’라고 규정할 수 있다. 이는 내부 통제를 강화하는 것이며, 최소한 부서 간 분리라도 이루어져야 한다. 예를 들어 AI 채용의 경우 HR팀 외 별도 컴플라이언스팀이 재심하는 등으로 독립성이 보장되어야 한다. 재심 절차의 독립성 보장이 실질적 이용자 보호의 핵심이다. 둘째, 신속한 처리 기한을 마련할 필요가 있다. 개인정보 열람·정정 요청도 법정 처리기한이 있는 것처럼, ADM 재검토도 기한을 명시해야 기업의 지연을 방지할 수 있다. 셋째, 결과 통지 및 이유 기재와 관련하여, 재심 후 결과를 통지할 때 단순히 ‘유지함’ 통지뿐만 아니라 간략한 이유도 알려주도록 하여 이용자가 납득하거나 추가적인

분쟁조정 등 이의제기 절차에서 판단의 기초가 될 수 있도록 한다. 넷째, 구체 절차를 다층적, 체계적으로 설계할 필요가 있다. DSA는 내부 불복 후 외부 분쟁조정을 두었는데, 이를 고려할 수 있다. 예컨대 방송 미디어통신위원회 산하 ‘알고리즘 투명성심의위원회’를 만들어, 이용자, 플랫폼, 이의신청으로 단계적으로 심리하는 모델도 가능하다(시사저널, 2023). 개인정보보호위원회 소관으로, 개인정보 분쟁조정위원회 활용이 현실적일 수 있으나 ‘자동화 결정 관련 분쟁’도 대상으로 포함하여 이용자가 기업의 재심결과에 불복하면 분쟁조정위원회에 조정신청하고, 기업은 권고에 따르면 해결되고, 그렇지 않을 경우 소송이 가능하도록 한다. 또는 집단분쟁조정도 가능한 바, 만약 특정 AI 시스템 문제로 많은 피해자가 생기면, 집단분쟁조정으로 일괄적인 조치를 유도할 수 있고, 개인정보 피해 집단분쟁 사례도 활용가능하다. 다층적 구체체계를 염두에 두고, 1차적으로 기업 내부 절차를 체계적으로 마련하되, 2차로 분쟁조정, 관련 부처의 조사 등 외부 절차 마련이 필요하다. 재심 결과에 불복 시 다음 단계로 불만 제기 가능한지 여부, 해당 기관 등을 안내하는 고지 절차도 마련하여 실효적인 권리 구체절차가 이루어질 수 있도록 할 필요가 있다.

4. 기록 유지의무와 검증 인프라 구축

추후 논란이 될 때 사후 입증 및 검증을 위해, 중요한 과정을 로그 기록으로 남기는 것이 중요하다. 개인정보처리자는 정보주체에 관한 자동화된 결정의 주요 처리과정과 결과, 그에 사용된 개인정보 항목을 결정 시점으로부터 최소 기간 동안 안전하게 보존하여야 한다는 등의 명확한 조항을 둘 필요가 있다. 또한 시행령 등 하위법령으로 위임되어 구체화 되더라도, ‘기록에는 자동화 결정이 내려진 일시, 사용된 시스템 또는 모델명, 투입된 개인정보의 종류, 산출된 결과, 개입한 사람의 신원, 정보주체 통지·이의 여부 등을 포함한다’고 규정하여 사후에 입증이나 분쟁

시의 증거로 활용될 수 있도록 할 필요가 있다. 기록의무만 두면 기업의 준수가 어려울 수 있기에 감독기관이 필요시 그 로그를 열람·제출받을 권한을 부여하고, 조사권 조항에 ‘제37조의2 관련 자료 제출 요구권’을 명시하는 방안을 고려하여야 한다.

더 나아가, 제3자에 의한 알고리즘 감사나 투명성 인증제도를 도입을 고려할 수도 있다. 예컨대 개인정보보호위원회가 일정 규모 이상 기업의 ADM 시스템에 대해 정기적으로 점검하거나, 기업이 자체 컨설팅을 받아 공정성 인증마크를 받게 유도하는 방식이다. 이러한 소프트 인프라는 기업의 자발적 개선을 유인할 수 있다. 이미 EU는 AI 알고리즘 감사(Algorithmic Auditing) 시장을 키우는 중이며 독일의 경우도 개인정보중심설계(Privacy by Design) 원칙 기반 인증제도를 운영 중인 바, 이러한 사례들을 참고하여 보다 적극적으로 인증제도를 시행할 필요가 있다.

정부 차원에서 설명가능한 AI(Explainable AI) 기술 연구를 지원해, 기업이 보다 용이하게 설명정보를 생성하고 로그를 관리하도록 장려해야 한다. 기록 및 검증 인프라는 자동화된 의사결정 제도의 실질적 투명성, 책임성 담보의 토대이다. 통지와 설명은 정보주체와 직면하는 ‘프런트엔드(front-end)’ 요소로서 외형적 투명성 제공에 집중하고, 기록관리와 감사체계는 실질적 신뢰·책임 담보를 뒷받침하는 ‘백엔드(back-end)’로 기능한다. 실효적 개인정보 보호, 분쟁처리 및 감독을 위해서 반드시 기록관리와 검증 인프라에 대한 입법적·행정적 보완이 필수적이다. 특히 영업비밀, 지식재산권 등의 사유로 기업이 설명을 거부할 때 감독기관의 실질적 검증권은 신뢰와 수용성에 결정적 역할을 한다. 다만, 모든 기업에 동일한 기록 및 감사 의무를 둘 경우, 중소기업은 비용과 기술적 부담으로 인해 사업 포기 또는 혁신 지연 가능성이 있으므로, 중소기업자에게는 기록·감사 의무의 차등적용이나 지원책도 함께 고려될 필요가 있다.

5. 거버넌스 구축 및 기관 간 역할 조정

제도 운영 거버넌스 측면에서 현재 자동화 결정 통지·대응권은 개인정보보호위원회 소관으로 사실상 개인정보보호위원회가 규제·감독한다. 그러나 앞서 본 바와 같이 플랫폼 알고리즘 문제는 방송미디어통신위원회, 과학기술정보통신부, 공정거래위원회, 금융위원회 등 다양한 영역과 관계된다. 개인정보위원회와 방미통위 등이 정례 협의회를 구성해, 관련 정책을 조율하고 정보도 공유하며, ‘디지털 플랫폼 알고리즘 규제 협의회’ 같은 합동기구를 설치하거나 ‘디지털 서비스 위원회’ 같은 통합 규제기구 신설을 생각해볼 수 있다. 한국 행정 환경상 현실적이진 않을 수 있으므로, 대신 위원회 산하에 ‘AI·알고리즘 전문위원회’를 설치하고 방미통위·과기정통부 인사도 참여시켜 정책개발에 협업하도록 할 수 있다. AI 기본법에 대하여 AI 윤리원칙, 이용자 권리, 생성 AI 결과물 표시의무, 고위험 AI 사전인증 등이 논의되었는데, 이러한 법규들과 개인정보보호법 간 중복·충돌 없게 정비되어야 하며, 기본법이 기본원칙을 두고, 구체 권리는 개인정보보호법이 담당하는 식으로 역할 분담을 명시한다. 알고리즘 검증 등에는 기술 전문성이 필요하므로, 감독기관 내 AI 전문인력 양성과 영입을 통해 데이터사이언스 인력을 보강하고, 추후 분쟁해결절차에서의 판사, 변호사들도 AI에 대한 이해를 증진하기 위한 시스템 마련이 필요하다. 글로벌 플랫폼의 알고리즘은 한 국가 규제만으로 어려우므로, EU의 사례 공유, 국제공동조사, 표준 개발 등을 위한 국제기구의 논의에 적극 참여할 필요가 있다.

6. 영업비밀 보호와 SME 차등규제

알고리즘 투명성 요구와 기업의 영업비밀 보호 간 균형이 제도의 수용성에 중요한 바, 자동화 결정 통지·설명 의무를 과도하게 요구하면, 기업의 핵심 알고리즘이 노출되어 경쟁상 불이익을 입을 수 있다. 따라

서 법에 ‘상업상 기밀 또는 지식재산 침해를 초래하지 않는 범위에서’ 등의 단서를 두어 기업들이 안심하고 협조하도록 하고, 다만 남용을 막기 위해, ‘개인정보처리자가 영업비밀을 이유로 정보 제공을 거부하는 경우, 위원회는 전문기관에 해당 알고리즘의 차별성 여부 등에 대한 검사를 명할 수 있다’ 등의 대리 감사 규정을 둘 수 있다. 설명 요구에 응할 때, 소스코드 등 영업비밀 부분은 적절히 블랙박스 처리하여 넘길 수 있도록, 기밀유지 협약하에 제공하는 절차 마련도 고려할 수 있다.

중소사업자에 대한 차등규제를 고려할 수 있는 바, 모든 기업에게 동일 의무를 부과하면 영세한 곳은 감당 못하거나 비용과다로 사업 포기 위험이 있다. 그러므로 매출, 이용자수 등 일정 기준 이하 기업은 일부 의무를 면제 또는 완화해줄 수 있다. EU DSA도 중소 플랫폼은 위험평가 등을 면제하는 등 조치를 둔 것과 같이 대규모 AI 활용 사업자와 영세 사업자를 구분해 규제의 단계적 적용을 도모해야 한다. 다만, 예외가 남용될 수 있기에 연매출, 이용자수, 민감영역 여부 등을 종합적으로 고려해 합리적으로 적용 대상을 구분하는 것이 과제이다. 자동화 결정이 이뤄지는 분야가 다양한데, 분야별로 요구사항이 다를 수 있으므로 분야별 특성을 반영하는 유연한 방식도 고려해야 한다. 이를 위해 고시나 향후 가이드라인에서 분야별 세부 지침을 마련할 수 있다.

VII. 결론

디지털 미디어 시대에 자동화된 알고리즘 의사결정이 개인과 사회에 미치는 영향력은 점점 커지고 있다. 본 연구는 개인정보보호법상 자동화 의사결정 통지·대응권 제도를 중심으로 국내외 법제와 정책을 비교 분석하고, 이를 헌법적으로 검토하여 제도 개선 방향을 제시하였다. 특히 알고리즘이 미디어 환경에서 사실상 ‘편집행위’로 작동하면서 정보접근권·여론의 다양성·표현의 자유에 중대한 영향을 미친다는 점을 강조함

으로써, 단순한 개인정보 규제에 차원을 넘어서 입법적·헌법적 논의를 제기하였다.

우리나라는 2023년 통지·대응권 제도를 도입하여 이용자의 알고리즘 투명성 권리를 인정하는 중요한 진전을 이루었다. 그러나 현행 법령은 개념 정의의 불명확성, 절차의 구체성 부족, 실효성 확보장치 미흡 등의 문제를 안고 있어 제도의 목적이 충분히 달성되지 못할 우려가 있다. 또한 알고리즘 결정이 실제 이루어지는 영역은 미디어 플랫폼임에도 불구하고, 개인정보보호위원회·방미통위·공정위 간 역할 분절 등 구조적 규제 공백이 발생하고 있다.

본 연구는 GDPR, 캐나다 퀘벡주 Loi 25, 브라질 LGPD 등 정보주체 권리를 중심으로 한 외국 입법례와 EU DSA, 영국 OSA, 호주 eSafety 제도 등 플랫폼 책임 규범을 비교함으로써, 자동화된 의사결정 규제의 국제적 흐름을 체계적으로 검토하였다. 이를 바탕으로 다음과 같은 개선 방향을 제안하였다. ‘완전 자동화된 결정’, ‘중대한 영향’ 등 핵심 개념을 법률에 명확히 정의하여 예측가능성을 제고하고 자의적 집행을 방지할 것, 통지·설명 기준을 법정화하여 이용자가 언제, 어떤 방식으로 정보를 제공받을 수 있는지 명확히 할 것, 재검토 절차에 독립성·신속성·결과통지 의무를 포함하여 절차적 실효성을 확보할 것, 기록의무·사후검증 체계를 강화하여 알고리즘 결정의 책임추적과 투명성을 보장할 것, 관련 기관 간 협력 및 역할 분담 체계를 법제화하여 규제 공백을 해소할 것, 영업비밀 보호장치 마련과 차등규제를 도입하여 규제의 비례성과 현실성을 확보할 것 등이다.

이러한 개선 방향은 헌법이 요구하는 개인정보자기결정권, 인격권 및 국가의 기본권 보호의무와 기업의 표현의 자유 및 영업의 자유 사이에서 균형을 도모하려는 시도이며, 개인의 존엄 보호와 데이터 경제의 혁신이라는 두 축을 조화시키려는 것이다.

■ 참고 문헌

[국내문헌]

- 강나운 (2023, 2, 3). “뉴스 배치 공정하게”…방통위, ‘알고리즘투명성심의위원회’ 설치. <시사저널>. URL: <https://www.sisajournal.com/news/articleView.html?idxno=255549>
- 강현구 (2021). 인공지능의 법적정의와 인공지능기술 발달에 따른 국가의 기본권 보호의무. <헌법재판연구>, 8권 2호, 29-64.
- 김연진 (2023). 개정 개인정보 보호법에 관한 시론적 고찰. <성균관법학>, 35권 2호, 327-363.
- 김희정 (2020). 알고리즘 자동의사결정으로부터 개인의 보호 - 인간의 개입권(the right to human intervention)을 중심으로 -. <헌법학연구>, 26권 1호, 207-237.
- 박노형 외 8인 (2017). <EU 개인정보보호법-GDPR을 중심으로->. 박영사.
- 박노형 (2024). <개인정보보호법>. 박영사.
- 박노형, 김효권 (2022). 자동화된 결정에 관한 개인정보보호법 정부 개정안 신설 규정의 문제점 <사법> 61, 361-390.
- 박노형, 김규동 (2024). 자동화된 의사결정의 일반적 금지원칙에 대한 EU 사법법원의 첫 판결 분석, <저스티스> 204, 239-266.
- 박미사 (2024). 인공지능(AI) 시대 개인정보보호법의 한계와 개선방안에 관한 고찰. <서강법률논총>, 13권 3호, 129-160.
- 박상돈 (2017). 헌법상 자동의사결정 알고리즘 설명요구권에 관한 개괄적 고찰. <헌법학연구>, 23권 3호, 185-218.
- 박아란 (2016). 알고리즘 편집과 인터넷뉴스서비스사업자의 법적 책임. <언론중재>, 138, 18-27.
- 박진완 (2025). 인공지능을 통한 자동화된 처리에 의한 의사결정, GDPR 제22조 그리고 유럽연합법원의 Schufa 결정. <법학논고>, 90, 1-30.
- 방송통신위원회, 정보통신정책연구원 (2022). <인공지능 기반 미디어 추천 서비스 이용자 보호 기본원칙 해설서>.
- 방송통신위원회, 정보통신정책연구원 (2025). <생성형 인공지능 서비스 이용자 보호 가이드라인>.
- 성낙인 (2024). <헌법학>. 법문사.

- 신현아 (2025). 온라인 추천 시스템 알고리즘 투명성 확보와 이용자의 선택권 보장 - EU 디지털서비스법과의 비교를 중심으로 -. <소비자법연구>, 11권 2호, 9-39.
- 유연주 (2023). <온라인플랫폼 상품검색·추천의 투명성 제고 및 정보공개 방안 연구>. 서울대학교 소비자학과 박사학위논문.
- 이대희 (2024). 인공지능에 의한 개인정보의 자동 처리에 대한 투명성 확보에 관한 연구 - 알고리즘에 의한 자동 의사결정에 따른 '설명권' 도입을 중심으로 -. <저스티스>, 200호, 298-323.
- 이상용·이혜리 (2024). 개인정보보호법에 의한 자동화된 결정의 규율. <선진사법률연구>, 105호, 69-112.
- 오세욱. (2021). <기계가 드러내는 편향, '이루다'만 문제였을까?> KISO 저널, 42, 13-17.
- 이성엽 (2025). <데이터와 법>. 박영사.
- 이원우, 토마스 팻처, 카렌 영, 베르트랑 뒤 마레, 제임스 스페타, 크리스티 포드, 김태오, 이광호, 김용대, 장원철, 윤혜선, 김태호 (2019). <4차 산업혁명 시대의 기술혁신과 규제정책>. 홍문사.
- 장영수 (2022). <헌법학>. 홍문사.
- 주민호 (2023). 자동화된 의사결정에 대한 기본권의 실효적 보장- EU AI법 제54조와 GDPR 제22조의 관계를 중심으로 -. <법학논고>, 82호, 63-88.
- 최경진 외 12인 공저 (2024). <개인정보보호법>. 박영사.
- 허완중 (2014). 기본권의 대사인적 효력과 기본권보호의무 그리고 기본권충돌의 관계. <헌법논총>, 25권, 1-59.

[해외문헌]

- AlgorithmWatch (2022). A guide to the Digital Services Act: The EU's new law to rein in Big Tech. URL: <https://algorithmwatch.org/en/dsa-explained/>
- Australian Senate Economics References Committee (2023). Influence of international digital platforms. Canberra. URL: https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Economics/Digital_platforms/Report
- Balkin, J. M. (2018). Free Speech in the Algorithmic Society: Big Data,

- Private Governance, and New School Speech Regulation. *U.C. Davis Law Review*, 51, 1149–1210.
- Balkin, J. M. (2021). How to Regulate (and Not Regulate) Social Media. *Journal of Free Speech Law*, 1, 71–96.
- Beltran, M. (2025). AI algorithms under scrutiny: GDPR, DSA, AI Act and CRA as pillars for algorithmic security and privacy in the European Union. *Computers & Security*, 158.
- Bolukbasi, T., Chang, K.-W., Zou, J., Saligrama, V., & Kalai, A. (2016). Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. *arXiv*: 1607.06520.
- Bygrave, L. A. (2019). Article 22 Automated Individual Decision-Making, Including Profiling. In C. Kuner et al. (Eds.), *〈The EU General Data Protection Regulation (GDPR): A Commentary〉*, Oxford University Press.
- Bundesbeauftragte für Datenschutz und Informationsfreiheit. (2023). Privacy by Design Certification in Germany. URL: <https://www.bfdi.bund.de/DE/Fachthemen/Inhalte/Technik/DPbD.html>
- Citron, D. K., & Pasquale, F. (2014). The Scored Society: Due Process for Automated Predictions. *Washington Law Review*, 89(1).
- Crawford, K., & Schultz, J. (2014). Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms. *Boston College Law Review* 55(1).
- De Gregorio, G. (2019). Democratising Content Moderation. A Constitutional Framework. *Computer Law & Security Review*, 36.
- De Gregorio, G. (2021). The Rise of Digital Constitutionalism in the European Union. *International Journal of Constitutional Law*, 19(1), 41–57.
- Federal Court of Australia. (2025, July 31). *Online safety regulator's powers confirmed in X Corp ruling*.
- Future of Privacy Forum. (2022). Automated Decision-Making Under the GDPR. URL: <https://fpf.org/wp-content/uploads/2022/05/FPF-ADM-Report-R2-singles.pdf>
- Gillespie, T. (2018). Custodians of the Internet: Platforms, Content

- Moderation, and the Hidden Decisions that Shape Social Media. Yale University Press.
- Goodman, B., & Flaxman, S. (2017). European Union Regulations on Algorithmic Decision-Making and a “Right to Explanation”. *AI Magazine* 38(3), 50–57.
- Hannak, A., Soeller, G., Lazer, D., Mislove, A., & Wilson, C. (2014). Measuring Price Discrimination and Steering on E-commerce Web Sites. *Proceedings of the 14th ACM/USENIX Internet Measurement Conference (IMC'14)*, 305–318. URL: <https://conferences.sigcomm.org/imc/2014/papers/p305.pdf>
- International Bar Association. (2024). Brazilian legal framework on automated decision-making. URL: <https://www.ibanet.org/Brazilian-legal-framework-on-automated-decision-making>
- Information Technology and Innovation Foundation. (2025). EU should improve transparency in the Digital Services Act. URL: <https://itif.org/publications/2025/10/20/eu-should-improve-transparency-in-the-digital-services-act/>
- Kaminski, M. E., & Urban, J. M. (2021). The Right to Contest AI. *Columbia Law Review*, 121(7).
- Keller, D. (2022). The DSA's Industrial Model for Content Moderation. *Verfassungsblog*, February 23. <https://verfassungsblog.de/dsa-industrial-model/>
- Klonick, K. (2018). The New Governors: The People, Rules, and Processes Governing Online Speech. *Harvard Law Review*, 131(6), 1598–1670.
- Lyon, D. (2015). *Surveillance after Snowden*. Polity Press.
- Malgieri, G. (2019). Automated decision-making in the EU Member States: The right to explanation and other “suitable safeguards” in the national legislations. *Computer Law & Security Review*, 35(2).
- Milano, S., Taddeo, M., & Floridi, L. (2020). Recommender Systems and Their Ethical Challenges. *AI & Society*, 35.
- O’Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Penguin Books.
- Online Safety Act Network. (2025). OSA regime: Final documents. URL:

- <https://www.onlinesafetyact.net/resources/osa-regime-final-documents/>
- Oswald, M., Grace, J., Urwin, S., & Barnes, G. (2018). Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality. *Information & Communications Technology Law*, 27(2).
- Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.
- Pollicino, O. (2021). Judicial Protection of Fundamental Rights on the Internet: A Road Towards Digital Constitutionalism?. Hart Publishing.
- Pollicino, O., & De Gregorio, G. (2021). Constitutional Law in the Algorithmic Society. In O. Pollicino & G. De Gregorio (Eds.), *Constitutional Law and the EU Digital Agenda*. Cambridge University Press.
- Portes, A., N'Goala, G., & Cases, A. (2020). Digital Transparency: Dimensions, Antecedents and Consequences on the Quality of Customer Relationships. *Recherche et Applications en Marketing (English Edition)*, 35(4), 72-98.
- Soderlund, K., Engstrom, E., Haresamudram, K., Larsson, S., & Strimling, P. (2024). Regulating high-reach AI: On transparency directions in the Digital Services Act. *Internet Policy Review*, 13(1), 1-31.
- Suzor, N. (2018). Digital Constitutionalism: Using the Rule of Law to Evaluate the Legitimacy of Governance by Platforms. *Social Media Society*, 4(3), 1-10.
- Treščakova, D. (2023). Freedom of Expression on the Internet vs. Digital Services Act: Chosen Aspects. *ECLIC*, 9.
- UK Department for Science, Innovation & Techonology. (2025). Online safety act: Explainer. URL: <https://www.gov.uk/government/publications/online-safety-act-explainer/online-safety-act-explainer>
- Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law*, 7(2).
- Wachter, S., & Mittelstadt, B. (2019). A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI.

Columbia Business Law Review, 2019(2).

[법령·판례]

Online Safety (Basic Online Safety Expectations) Determination (2022).

URL: <https://www.legislation.gov.au/Details/F2022L00213>

Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act)(2022).

X Corp v eSafety Commissioner [2025] FCAFC 99.

■ ABSTRACT

A Comparative Legal Study on the Regulation of Automated Decision-Making and Its Implications for Korean Law

Jeong, Ran

Ph.D., Senior Research Fellow, Seoul National University Law Research Institute

Automated decision-making (ADM) in the digital media environment significantly affects fundamental rights, including privacy, freedom of expression, equality, informational self-determination, and the integrity of public discourse. In Korea, the 2023 amendment to the Personal Information Protection Act (PIPA) created a statutory framework for ADM notifications and user response rights, supported by enforcement decrees and administrative guidelines. However, significant limitations persist: key concepts lack statutory clarity, notification and explanation obligations are vague, review procedures have limited independence and effectiveness, record-keeping and audit duties are absent, and ADM rules are structurally disconnected from the broader media regulatory framework. To address these issues, this study uses a comparative framework to examine regulatory approaches in other jurisdictions. This study uses a comparative framework to examine the European Union's GDPR Article 22, the Digital Services Act (DSA), Quebec's Loi 25, Brazil's LGPD, the United Kingdom's Online Safety Act, and Australia's eSafety regime. The aim is to identify best practices and propose reforms to improve the regulation of ADM in Korea's media

environment. The analysis focuses on the specific risks of ADM in media platforms, such as news ranking, personalized content curation, algorithmic filtering, and automated moderation, and evaluates their constitutional implications under Korean constitutional principles of legality, clarity, proportionality, freedom from prior restraint, and the State's positive duty to protect fundamental rights. Based on this comparative and constitutional analysis, the study proposes a multi-layered reform model for Korea: statutory clarification of core ADM concepts; standardized, enforceable notification and explanation duties; independent, timely human review procedures; mandatory record-keeping and supervisory audit mechanisms; and an integrated governance structure linking ADM regulation under PIPA with media sector laws, such as the Broadcasting Act, Newspaper Act, and Information and Communications Network Act. These reforms aim to improve the effectiveness of rights to human intervention, explanation, and contestation, and to strengthen protection of personality rights and informational self-determination in the algorithm-driven media ecosystem. This study contributes to the literature by providing a comprehensive comparative analysis and a reform model specific to the Korean context.

Keywords: Automated Decision-Making (ADM), Personal Data Protection, LGPD
Algorithmic Transparency, Proportionality, GDPR, DSA, Loi 25

[논문투고일 2025. 10. 27. 논문수정일 2025. 11. 23. 게재확정일 2025. 11. 27.]