

# 인권보장을 위한 인공지능 입법방향\*

## - 디지털 권리장전에 비춰본 영향분석의 필요성 -

심 우 민\*\*

- I. 서론
- II. AI 규범체계와 인권
- III. 한국의 디지털 권리장전
- IV. AI시대 인권보장을 위한 입법방향: 영향분석의 선제적 제도화의 요청
- V. 결론을 대신하여

### 국문초록

새로운 AI 규범체계 정립 시도가 전 세계적으로 이루어지고 있다. 경우에 따라서는 직접적인 AI 규제입법을 시도하는 사례들이 존재하는 것이 사실이지만, AI 기반 사회가 나아가야할 규범적 기준 및 원칙을 제시하고자 하는 것이 주류적인 경향이라고 할 수 있다. 결국 이는 향후 더욱 본격화할 것으로 보이는 AI 관련 입법의 구체적인 지향점으로 활용될 것으로 보인다. 따라서 현재 단계에서는 AI 규범체계의 목적을 명확하게 정립하는 것이 필요하다. 이러한 규범체계 목적은 궁극적으로 실질적인 인권의 보장이라고 할 수 있다. 한국도 이러한 맥락에서 최근 「디지털 권리헌장」을 공표하였다. 이는 당초 새로운 인권 및 가치질서를 제시하겠다는 목표를 가지는 것이었다. 그러나 인권의 도구적 역할에만 주목한 나머지, 이미 국제사회 규범이나 헌법을 통해 보장되는 권리들을 용어만 바꾸어 제시하는 수준에 머무르고 말았다. 실질적인 인권 보장에 기여할 수 있는 AI시대 입법에 기여하기 위해서는 단순한 규범적 선언에만 천착할 것이 아니라, 그

\* 이 논문은 2023년 11월 18일 서강대학교 법학연구소와 ICT법경제연구소의 추계 공동학술대회에서 발표한 내용을 논문의 체계로 재정리한 것임을 밝힙니다.

\*\* 경인교육대학교 부교수, 입법학센터장(법학박사)

에 앞서 면밀한 영향분석과 이에 입각한 가치적 공감대를 형성해 나가는 절차를 선제적으로 추구할 필요가 있다. 특히 AI 기술 및 그 활용의 맥락이 불확정성을 불가피하게 내포하고 있고, 그런 측면에서 일의적인 법적 규제가 아닌 위험 기반 접근방식이 필요하다는 측면에서 영향분석 제도화의 필요성은 아무리 강조해도 지나치지 않다.

• 핵심어 인공지능, 인권, 디지털 권리장전, 인공지능 입법, 영향분석

## I. 서론

IT법학 연구에 있어 인공지능(Artificial Intelligence, AI) 기술 도입과 발전은 중요한 전환의 계기를 마련해 주고 있다. 특히 최근 급격하게 관심이 집중되고 있는 대규모 언어모델에 입각한 생성형AI의 확산은 다시금 AI를 둘러싼 규범정립 또는 입법에 관한 논란을 불러일으키고 있다. 대체적으로는 AI를 통해 확보하게 된 유용성에 집중하면서, 이를 중장기적인 공동체 발전과 연계시키려는 입법 실무적 시도가 이어지고 있다.

최근 국내에서 진행되고 있는 AI 관련 입법에 관한 내용들을 살펴보면, 대부분 산업적 또는 경제적 측면을 강조하고 있는 양상을 보인다. 물론 산업의 안정적이고 지속 가능한 발전을 위해서는 AI로 인해 발생할 수 있는 역기능에도 주목할 필요가 있지만, 오히려 이러한 부분에 관한 명확한 입법정책적 지향점을 제공해주는 입법대안을 확인하기는 어렵다. 아마도 그 이유는 AI 기술의 발전 가능성과 방향성을 사전에 예측하기 어렵기 때문인 것으로 판단된다. 즉 AI 관련 입법에서 역기능 대응의 문제는 다소 선언적인 수준에서만 다루어진다. 이 논문은 바로 이런 현재 입법행태의 문제점에 관해 논의하고자 한다.

사실 입법은 그 속성상 매우 형식적인 개념과 요건을 전제로 할 수밖에 없다. 따라서 그 발생여부가 매우 유동적인 역기능에 대한 대응책을 중심으로 사전적인 규제입법이 이루어지는 경우에는, 포괄적이고 획일적인 규제 설정으로 인해 비례성을 상실한 입법이 이루어질 가능성이 높다. 더 나아가서는 그러한 입법으로 인해 기술의 창의적이고 자생적인 발전을 저해할 수 있다는 점 또한 고려되어야 한다. 그렇다고 AI에 관한 입법 고민과 시도를 그저 방기할 수만은 없다. AI 기술 자체는 현재의 수준만으로도 인간의 일상생활 방식은 물론이고 사회적인 구조의 변화까지도 충분히 추동할 수 있을 것으로 보이고, 이에 따라 과거에는 상상하지 못했던 다양한 역기능을 야기할 가능성이 높기 때문이다.

우리나라 정부는 「인공지능(AI) 윤리기준」<sup>1)</sup>, 「교육분야 인공지능 윤리원칙」<sup>2)</sup> 등을 내놓음으로써, 역기능에 관한 초기 대응을 수행하고 있는 것으로 판단된다. 그러나 문제는 이러한 윤리 기준 및 원칙의 국가적인 선언에도 불구하고, 이를 활용하여 관련 정책이나 입법의 방향성을 구체화하는 데에는 현실적으로 한계가 있다는 점이다. 특히 이러한 기준 및 원칙의 내용은 종래 우리사회가 추구하던 전통적인 규범적 가치를 재차 반복적으로 선언하는 수준에 머무르고 있어 입법정책적 견지에서 특수성도 발견하기도 어렵다.

상황이 이러하다보니, 사회적으로 AI를 둘러싼 사회적 문제가 불거지게 되면 중국에는 기존의 ‘법률적’인 기준이나 제도에 입각하여 사안을 해소하는 데 급급한 양상을 보여준다. 예를 들어, 2021년 우리나라에서 발생한 AI챗봇 ‘이루다’ 사건에 대한 정부의 대응에 있어서도 법률적 차원에서 「개인정보 보호법」 등 현행법 규정을 준수한 것인지 여부 판단에만 집중되었다.<sup>3)</sup> 그러나 실제 ‘이루다’ 사건의 본질적인 쟁점이었던 차별 및 혐오 표현 등의 쟁점들은 기존 법률 규정상 요건 등에 관한 판단만으로는 다루기 어려운 인권 쟁점들이었다. 결국 법(률)을 넘어선 보다 광범위한 인권에 대한 고려가 현 시점에서는 필요하고, 이를 바탕으로 과거와는 다른 새로운 법질서를 구축해 나가기 위한 고민을 수행해야 할 필요가 있다.

이와 같은 배경을 바탕으로, AI시대의 인권보장을 위한 입법은 어떠한 방향성을 가져야 하는지를 분석해 보고, 이에 입각하여 최근 정부가 제시한 「디지털 권리장전」의 정책추진 맥락을 비판적으로 고찰해 보고자 한다. 「디지털 권리장전」이 논하고자 하는 ‘권리’라는 것이 결국에는 법적 권리가 아닌 인권 관념을 전제로 한 것이었으리라 추론해볼 수 있기 때문이다. 또한 중국에는 AI에 관한 인권보장 입법에 있어 중요하다고 판단되는 영향분석 등의 필요성과 방향을 제언하도록 한다.

1) 관계부처 합동, 사람이 중심이 되는 「인공지능(AI) 윤리기준」, 2020.12.

2) 교육부, 사람의 성장을 지원하는 「교육분야 인공지능 윤리원칙」, 2022.8.

3) 개인정보보호위원회는 ‘이루다’의 개발사인 (주)스캐터랩에 대해 「개인정보 보호법」상 개인정보 처리 규정 위반을 이유로 1억330만원의 과징금(5,550만원) 및 과태료(4,780만원), 그리고 이에 따른 시정조치를 명령하였다.

## II. AI 규범체계와 인권

### 1. 새로운 규범체계의 요청

AI와 같은 혁신적 기술의 도입 및 일상화는 과거와는 다른 규범체계를 요청한다. 그 이유는 규범적 판단에 전제된 사실 자체가 변화하는 경우는 물론이고, 사회 구조적 측면에서 과거와는 다른 규범적 판단이 요청되는 경우도 있기 때문이다. 따라서 최근 국내외적으로 논의되고 있는 디지털 전환(digital transformation)에 있어 빈번히 논의되는 것이 법 패러다임 전환이며, 이는 현재 상황을 설명하는 데 있어 당연한 것으로 받아들여지고 있다. 이제 본격적으로 고민이 필요한 것은, 현 시점에서 상정해 두어야 하는 규범체계의 전환이 과연 어떠한 방향성을 가져야 하는지에 관한 쟁점이다. 이를 위해서는 디지털 전환의 배경 기술의 속성을 먼저 살펴보아야 할 것이다.

AI 기술은 이미 오랜시간을 거치면서 발전해 왔다고 할 수 있다. 그럼에도 불구하고 최근에 들어 본격적으로 AI를 둘러싼 규범체계 쟁점을 논의하기 시작한 것은, 이제 AI 기술 자체가 단지 유용한 객체 수준으로서만 존재하는 것이 아니라, 사회 전 영역에 영향을 미칠 수 있는 기술 수단으로 발전했기 때문이다. 그런데 이러한 발전으로 인한 영향은 쉽사리 예측하기 어려운 측면이 있다. 그 원인은 두 가지로 언급해 볼 수 있을 것이다. 우선 AI 기술 자체가 기계의 자율적인 판단과 의사결정을 추구하는 것이기에, 그것이 산출해낸 결과를 외부자인 인간의 입장에서 정확하게 예견하기 어려운 상황들이 빈번하게 발생한다. 또한 비약적인 발전을 거듭하고 있는 AI 기술은 기술 방식과 활용 분야 등에 있어서도 지속적으로 변화할 것으로 보인다. 따라서 AI는 과거 규범적 판단의 대상 또는 객체들과는 달리 매우 유동적인 속성, 즉 불확정성(indeterminacy)<sup>4)</sup>을 가지고, 이것이 바로 규범체계를 논하기 위한 대전제이다.

판단의 객체 또는 대상이 유동성을 가진다는 것은, 그에 대한 규범적 판단 또한 매우 유동성을 가질 수밖에 없다는 점을 의미한다. 그런데 이는 현재 우리의 규범체계가 전제로 하고 있는 것과는 다르다. 통상 이제까지의 법적이거나 규범적인 판단은 대상과 사안이 어느 정도 정형화 또는 고정되어 있는 상황을 상정하고 있다. 그래야 법적 효과를 부과하기 위한 개념과 요건을 형식적인 법규범에 실정화할 수 있기 때문이다. 이는 특히 형식

4) 심우민, “디지털 전환과 법교육의 미래: 법사회학적 접근 방식의 시사점”, 『법교육연구』 17(3), 2022, 109면 이하.

적인 (실정)법을 규범체계의 중핵으로 삼고 있는 자유주의 규범체계에서 요구하는 바이기도 하다. 결과적으로 이제까지의 자유주의 규범체계와는 달리 규율 대상과 상황이 가지는 유동성을 반영할 수 있는 새로운 규범체계 정립방안의 모색이 필요하다.

## 2. AI에 관한 규범적 대응방식

### (1) 규제 입법적 대응

AI의 영향력과 위험 담론이 지속적으로 심화·확산되면서, 당연히 이의 역기능을 직접적으로 규율하기 위한 입법적 대응을 먼저 상정하는 실무적 행태가 증가하고 있다. 그 이유는 전통 규범체계 속에서 대응책을 고민하는 것이 가장 간명할 수 있기 때문일 것이다. 물론 아주 현실적으로만 보자면 직접적인 규제가 가능하다면 그러한 방안을 거부할 이유가 없을 것이다.

일반적으로 규제 입법의 측면에서 고려될 수 있는 가장 직접적인 것은 바로 ‘기술적·관리적 조치’에 관한 규정을 법제화하는 방식이다. 이에 대한 대표적인 사례는 EU의 「인공지능법안」에서 발견할 수 있다. 동 법안 제16조<sup>5)</sup>에서는 고위험 인공지능 시스템 제공자에게 품질관리시스템 마련 의무, 기술문서 작성의무 및 시스템 생성 로그 보관의무 등을 부과하고 있으며, 제52조<sup>6)</sup>에서는 “제공자는 자연인과 상호작용하는 인공지능 시스템을 설계하고 개발하는 경우, 자연인에게 인공지능 시스템과 교류하고 있음을 알리는 기능을 삽입해야 한다”는 등의 직접적인 기술적·관리적 조치를 의무화하고 있다.

이러한 직접적인 규제방식은 일응 의미가 있어 보이는 것은 맞지만, 실제 법제화되어 운영함에 있어서는 다양한 어려움에 봉착할 수 있는 한계가 있다. 예를 들어, 시스템 제공자에게 품질관리시스템 마련 의무, 기술문서 작성의무 및 시스템 생성 로그 보관의무 등을 부과하고 있는데, 매우 광범위한 다양성을 가지는 AI 기술과 서비스의 특성을 포괄할 수 있는 구체적·실무적 기준을 사전에 마련하는 데에는 상당한 논란이 발생할 가능성이 높다.

무엇보다 중요한 것은 동 규정의 규제 대상인 ‘고위험 인공지능’의 범주가 유동적이라는

5) European Commission, Proposal for a regulation laying down harmonised rules on Artificial Intelligence(Artificial Intelligence Act) and amending certain Union legislative acts(COM/2021/206 final), 2021.4.21, Article 16 (Obligations of providers of high-risk AI systems).

6) Artificial Intelligence Act, Article 52 (Transparency obligations for certain AI systems).

점이다. EU의 「인공지능법안」은 고위험 인공지능에 관한 기준을 제6조에서 제시하면서,<sup>7)</sup> 제7조에서는 EU 집행위원회가 부속서(annex) III 업데이트를 통해 고위험 인공지능의 범주를 넓힐 수 있도록 권한을 부여하고 있다.<sup>8)</sup> 사실상 「인공지능법안」의 주된 규제 목적이 고위험 인공지능의 관리에 초점을 맞추고 있다는 점을 고려한다면, 이는 고위험 인공지능의 개념적 범주가 유동적일 수밖에 없다는 점, 즉 불확정성을 나타내는 것이다. 따라서 비교적 간명해 보이는 기술적·관리적 조치와 같은 직접적인 규제방식의 경우에도, 중국에는 유동적 대응방식을 본질적으로 내포하고 있다는 점은 반드시 짚고 넘어가야 하는 사항이다.

위에서 언급한 가장 직접적인 규제방식 이외에도, ‘인증(유사)체계 활용방식’과 ‘권리 및 의무 설정 방식’이 모종의 규제 입법 유형으로 활용될 가능성도 있다. 권리 및 의무 설정 방식으로는 EU의 「개인정보보호규칙」 제13조~제15조에서 규정하고 있는 소위 ‘설명요청권(right to explanation)’ 부여방식<sup>9)</sup>이 대표적인 예시이다. 그러나 이 경우에도 AI 서비스 특성에 따라 어떤 내용을 어느 정도 수준까지 설명할 것인지 여부는 여전히 유동적으로 판단할 수밖에 없다. 인증(유사)체계 활용방식으로는 EU의 「인공지능법안」에 ‘적합성평가(conformity assessment)’라는 이름으로 규정되어 있다.<sup>10)</sup> 사실 이러한 인증체계 방식은 인공지능 담론 활성화 초기에 일본이 제안했던 「AI 개발가이드라인」에도 포함되어 있었다.<sup>11)</sup> 그러나 이러한 인증체계 활용의 경우에도 다양한 AI 활용방식을 모두 일괄적으로 포섭하는 데 있어서는 한계가 있을 것으로 보이고, 결국 상황변화에 대응할 수 있는 유동적인 방식으로 변모해 나갈 수밖에 없을 것이다.

결과적으로 세부적인 기준 설정이나 규제 집행에 있어서는, 구체적이고 유동적인 행정적 재량판단이 불가피하게 요청된다는 점을 확인할 수 있다. 따라서 현재 수준에서 이루어지는 규제 지향적 접근은 그 자체로 완결된 것이라기보다는, 다소 포괄적인 규제 설정을 위한 초기 단계 시도로 이해될 수 있을 것이다.

7) Artificial Intelligence Act, Article 6 (Classification rules for high-risk AI systems).

8) Artificial Intelligence Act, Article 7 (Amendments to Annex III).

9) “프로파일링을 포함한 자동화된 의사결정의 존재, 그리고 적어도 그러한 경우 이에 사용되는 로직에 관한 의미 있는 정보와 해당 처리가 정보 주체에 대해 갖는 중요성과 예상 결과”를 정보주체에게 정보 관리자가 설명토록 하는 의무를 규정하고 있다. 이와 유사한 맥락으로, EU 「온라인 플랫폼 규칙」 제5조도 온라인 플랫폼 사업자(검색엔진 포함)는 자사 웹사이트 화면에 배열되는 업체·상품 등의 우선순위를 결정짓는 주요 변수(main parameters) 및 고려되는 각 변수 간의 상대적 중요도를 약관에 명시해야 한다고 규정하고 있다. Regulation (EU) 2019/1150 of the European Parliament and of the Council of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services.

10) Artificial Intelligence Act, Article 19 (Conformity assessment)

11) AIネットワーク社会推進会議事務局(総務省情報通信政策研究所調査研究部), 「AI 開発ガイドライン」(仮称)の策定に向けた国際的議論の用に供する素案の作成に関する論点, 2016.12.28.

## (2) AI 기반 사회의 규범적 기준 및 원칙 제시

규범체계나 법 자체가 유동성을 가진다는 것은 법규범적 판단이 자의적으로 이루어질 수 있다는 점을 의미하기에, 이에 부응할 수 있는 새로운 규범질서를 상정하고 구체화해 나가는 데에는 어려움이 있다. 이런 의미에서 현재 세계 각국이 속속 제시하고 있는 다양한 AI 기반 사회 운영에 관한 규범적 기준이나 원칙들은, 규범적 판단의 자의성을 통제할 수 있는 규범 설정 ‘목적’을 명확히 한다는 점에서 나름 중요한 의미를 가지는 것이라고 할 수 있다. 물론 현실적으로 보자면, 이러한 원칙이나 기준들은 직접적으로 규범체계의 전환을 상정한 것이라기보다는, 즉자적이고 직접적인 입법적 규제의 어려움을 나타내는 것이라고 보는 것이 타당하다. 즉 현재 상황에서 직접적인 법적 규제를 설정하기에는 규율 대상이 명확하지 않거나 사안이 성숙하지 않았기 때문이다. 그러나 중장기적인 규범체계의 전환이 불가피하다는 점을 고려한다면, 결국 AI 관련 기준이나 원칙들은 향후 정립된 규범체계의 가치적 지향점을 명확히 해나가는 과정 중 한 단계를 나타내는 것이라고 평가할 수 있을 것이다.

그런데 현재 대체적으로 제시되어 있는 AI 기반 사회 운영에 관한 규범적 기준이나 원칙들은 매우 포괄적이고, 전통적인 규범적 기준을 사실상 거의 그대로 제시하고 있는 양상을 보인다. 예를 들어, 우리나라의 「인공지능(AI) 윤리기준」은 3대 기본원칙과 10대 핵심요건을 제시하고 있다. 3대 기본원칙으로는 ‘인간 존엄성 원칙’, ‘사회의 공공선 원칙’, ‘기술의 합목적성 원칙’을 제시하고 있으며, 이를 실천하고 이행할 수 있도록 AI 전체 생명주기에 걸쳐 충족되어야 하는 10대 핵심요건을 제시하고 있다. 이러한 10대 핵심요건으로는 ‘인권보장’, ‘프라이버시 보호’, ‘다양성 존중’, ‘침해금지’, ‘공공성’, ‘연대성’, ‘데이터 관리’, ‘책임성’, ‘안전성’, ‘투명성’이 제시되어 있다.<sup>12)</sup>

이러한 기본 원칙과 요건에서 확인할 수 있는 바와 같이, 비단 AI 기술 및 서비스에 한정하는 것이 아니라, 일반적으로 기술을 사회적으로 활용한다고 했을 때 당연히 고려해야 하는 전통적 사회 가치를 나열하고 있다. 이런 측면에서 이러한 국가 차원의 AI 윤리기준을 별도로 제시하는 것이 과연 어떠한 상징적인 의미를 가지는 것인지에 대해 의구심이 있다. 다만 동 윤리기준에서는 “각각 원칙들 사이에 고정된 형태의 우선순위를 제시하지는 않으며, 직간접적으로 영향을 받는 이해관계자가 지속적인 토론과 숙의 과정에 참여하여 절충점과 해결 방안을 모색하도록 권유한다”고 언급함으로써, 결국 사회적 담론 과정을 거

12) 관계부처 합동, 사람이 중심이 되는 「인공지능(AI) 윤리기준」, 2020.12.23.

처 구체화되어야 한다는 점을 간단히 제시하고 있다는 점에 주목해볼 필요는 있을 것이다.

실제 우리나라의 「인공지능(AI) 윤리기준」은 EU 집행위원회(European Commission)의 「신뢰할 만한 인공지능 윤리 가이드라인(Ethics Guidelines for Trustworthy AI)」<sup>13)</sup>의 영향을 받은 것이다. 그런데 EU의 「신뢰할 만한 인공지능 윤리 가이드라인」의 경우 포괄적인 원칙이나 기준 선언에 그치는 것이 아니라, 이것을 실제 다양한 개별 AI 개발자 및 서비스 제공자들이 활용할 수 있는 실천적인 자체평가(self assessment) 체크리스트까지 제공하기 위한 시범 운용절차(pilot process)를 포함하는 것이었다는 차별성이 있다.<sup>14)</sup> 이는 결국 보다 세부적인 기준 정립을 목표로 하는 분석 절차를 수행하고자 하는 것으로, 포괄적으로 제시한 기준의 해석 및 운용상의 유동성을 좁혀나가는 절차로 이해해볼 수 있을 것이다.

### 3. AI 규범체계 정립 근간으로서의 인권보장

종래 규범적 대응방식과 체계의 한계로 인하여, 불가피하게 가장 일반적으로 채택되는 방식은 AI 기반 사회를 위한 규범적 원칙이나 기준을 제시하는 것이다. 그런데 실상 문서 상으로 제시되어 있는 내용 자체로는 AI 기반 사회를 전제로 한다는 내용상의 특수성을 확인하기 어렵다. 다만 제시되어 있는 내용들을 종합해 보면, 결국 인권보장이라는 논제로 귀결되고 있다는 점을 확인할 수는 있다.

우리나라의 「인공지능(AI) 윤리기준」이 제시하고 있는 3대 기본원칙과 10대 핵심요건의 경우에도, 관련 내용을 종합해 보면 결국 인권보장을 내용으로 하는 것이다. 이는 「인공지능(AI) 윤리기준」이 “인간성을 위한 인공지능(AI for Humanity)”을 추구한다고 명시하고 있다는 점에서도 확인할 수 있다. 결과적으로 앞서 언급한 바 있는 유동적 성격을 가지는 규범적 판단이 불가피한 상황에서 그러한 판단의 자의성을 통제할 수 있는 규범 설정 ‘목적’은 바로 ‘인권보장’이다.

여기서 문제는 인권보장이 의미하는 개념적 범주가 비단 헌법이나 법률에 의해 뒷받침되는 법적 권리 영역에만 한정되는 것이 아니라는 점이다. 달리 말하자면, 기존의 법에 의

13) High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, European Commission, 2019.4.8.

14) High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, European Commission, 2019.4.8, 24면 이하. 이에 기반하여 High-Level Expert Group on Artificial Intelligence은 2020년 7월 17일 신뢰할 수 있는 AI 평가목록(체크리스트)을 제시하였다. High-Level Expert Group on Artificial Intelligence, The Assessment List for Trustworthy Artificial Intelligence(ALTAI) for self assessment, 2020.7.17.



해 보장되지 않을 수도 있는 인권의 가치나 내용이 존재한다는 가정 하에, 윤리 기준이나 원칙을 제시하는 것이라고 보는 것이 타당하다. 만일 법적으로 명확하게 보장할 수 있는 권리라고 한다면, 윤리기준 정립보다는 법률 등의 새로운 입법을 통해 충분히 대응할 수 있을 것이다. 결국에는 아직까지 법적 권리 보호나 직접적인 입법적 규제를 설정하기 어려운 상황이기 때문에, 윤리 규범이나 기준에 천착하는 양상을 보이게 된 것이라고 현 상황을 이해할 수 있다.

이와 관련하여, 국가인권위원회는 「인공지능 개발과 활용에 관한 인권 가이드라인」<sup>15)</sup>을 정부에 권고한 바 있다. 이 가이드라인은 각 정부 부처에서 발표하고 있는 다양한 인공지능 관련 윤리 기준들이 역사적 배경과 문화적 차이에 따라 자의적으로 해석될 수 있다는 점을 지적하고 있으며, 또한 이러한 우리나라의 윤리기준들이 국제적인 인권기준과는 의미와 규범성이 다르다고 평가하는 입장을 취하고 있다. 그리고 보다 직접적으로 “인공지능의 개발로 인해 예상치 못한 사회적 문제점이 점차 증가할 것을 고려할 때 인공지능의 개발과 활용에 있어 인권적 가치를 훼손하지 않고, 인간의 존엄성과 기본권 보장을 위한 인권 가이드라인의 필요성이 커지고 있다”고 밝히고 있다.<sup>16)</sup>

이 가이드라인은 ‘인간의 존엄성’, ‘투명성과 설명의무’, ‘자기결정권 보장’, ‘차별 금지’, ‘인공지능 인권영향평가 시행’, ‘위험도 등급 및 관련 법·제도 마련’ 등에 관한 주요 내용들을 권고하고 있다. 사실 여기에 나타난 주요 내용만 보자면, 실제 「인공지능(AI) 윤리기준」 등과 같이 우리나라가 제시하고 있는 윤리기준들과 현격한 내용상의 차별성이 존재하는 것이라고 보기는 어려울 정도로 포괄적이고 전통적인 인권규범을 나열하고 있는 것으로 보인다. 다만 특징적인 것은 ‘인공지능 인권영향평가 시행’과 ‘위험도 등급 및 관련 법제도 마련’을 제시하고 있다는 점에 주목할 필요가 있다.

인공지능 인권영향평가의 시행에 관해, 이 가이드라인은 “국가는 인공지능의 개발과 활용에 있어서 인권침해와 차별의 가능성 및 정도, 영향을 받는 당사자의 수, 사용된 데이터의 양 등을 고려하여 공공기관 및 민간기업을 대상으로 인권영향평가를 실시하여야” 한다고 권고하고 있다. 이는 결국 사전에 인권보장을 위한 일률적인 규제기준을 마련하기 어렵다는 지점을 고려한 것으로 보인다. 이를 위하여 이 가이드라인은 “인권영향평가 내용에는 인공지능의 특성, 상황, 범위 및 목적을 감안하여 인권 가이드라인이 제시한 원칙 및 내용, 국제 인권 기준, 관련 법률에서 정한 의무 등이 포함되어야 하며, 인권침해 위험요

15) 국가인권위원회 결정, 「인공지능 개발과 활용에 관한 인권 가이드라인」, 2022.4.

16) 국가인권위원회 결정, 「인공지능 개발과 활용에 관한 인권 가이드라인」, 2022.4, 4면.

인의 분석, 개선 사항 등을 도출해야” 한다고 부연한다.<sup>17)</sup>

또한 이 가이드라인은 “국가는 개인의 인권과 안전에 미치는 위험성이 매우 높아 인공지능이 금지되는 영역, 상당한 제한이 필요한 인공지능 고위험 영역, 일정 정도의 제한이 필요한 제한된 위험을 가진 영역, 위험성이 거의 없는 영역 등 적절하게 인공지능의 위험성 단계를 구분하고, 그에 맞는 규제 수준과 인적 개입이 이루어지도록 법과 제도를 마련하여야” 한다고 권고하고 있다. 이는 최근 EU의 「인공지능법안」 등 관련 규제 대안들이 견지하고 있는 위험 기반 접근방식(risk-based approach)을 보여주는 것이라고 할 수 있는데, 결국 인권보장이라는 가치를 중시여기면서도 비례성을 확보한 규제 대안 마련을 현실적으로 제안하고 있는 것으로 이해할 수 있다. 여기에서 주요하게 주목해야 하는 점은, 이러한 위험 기반 접근방식이 가능하기 위해서는 AI 기술의 위험도를 측정할 수 있는 구체적인 평가 및 분석이 선행되어야 한다는 것이다.

결과적으로, 인권보장을 궁극적인 규범 목적으로 한다고 하더라도, AI 기술 환경에서 요구되는 ‘인권보장의 특수성을 평가하고 분석할 수 있는 절차’가 요구된다. 따라서 현 시점에서는 즉각적인 법규범적 실체 조항(입법대안)들을 강구하기보다는, 인권보장이라는 보편적 가치를 추구하기 위한 목적 하에, 현 시점에 필요한 종합적 규범 체계를 단계적으로 정립해 가는 과정이 필요함을 인식할 수 있다.

### III. 한국의 디지털 권리장전

#### 1. 한국의 디지털 권리장전 논의 배경

OECD는 2022년 12월 ‘디지털 시대의 권리’(Rights in the digital age: Challenges and ways forward) 보고서<sup>18)</sup>를 발표한 바 있다. 이 보고서는 디지털 전환의 상황이 국제적으로 인정된 인권, 법적·헌법적 권리는 물론이고, 국내적으로 보호되는 이익에 미치는 영향을 검토한 것이다. 이를 토대로 표현의 자유, 개인정보 보호, 인터넷 접근권 등을 중심으로 디지털 권리를 보호하기 위한 국제사회는 물론이고 개별 국가들까지 포괄하는 이니셔티브를 제시했다. 즉 국제사회는 디지털 전환에 대응하여 국제적 차원의 인권 규범을 재

17) 국가인권위원회 결정, 「인공지능 개발과 활용에 관한 인권 가이드라인」, 2022.4, 7면.

18) OECD, Rights in the digital age: Challenges and ways forward, OECD DIGITAL ECONOMY PAPERS No. 347, 2022.12.

해석하여 인권의 적용 방법과 범위를 구체화하고자 하는 노력을 진행하고 있는 것이다. 이런 맥락에서 대부분의 국가들은 다양한 형태의 디지털 인권 문제에 대해 대응책을 마련할 수 있는 종합적이고 포괄적인 틀을 마련하고자 노력하고 있고, 한국도 예외는 아니었다.

유사한 견지에서, 한국 정부는 2022년 9월 「대한민국 디지털 전략」을 발표하였다.<sup>19)</sup> 이 전략에서는 디지털 활용을 보편적 권리로 보장하겠다는 취지를 가지는 (가칭)「디지털 권리장전」을 수립하겠다는 계획을 제시하였다. 변화하고 있는 디지털 환경에 부합하는 새로운 가치의 구체화 방향성을 제시하겠다는 정책방향을 밝힌 것이다. 또한 「디지털 권리장전」의 연장선에서 「대한민국 디지털 전략」은 새로운 디지털 사회 질서 속 국민의 기본 권리를 규정한 「디지털사회 기본법」 제정하겠다고 밝히고 있는데, 이는 「디지털 권리장전」의 구체적 실천을 위한 제도적 수단으로 기본법<sup>20)</sup> 유형을 상정하고 있는 것이라고 볼 수 있다.

정부는 「디지털사회 기본법」 제정에 관한 구체적인 내용으로, 기존의 「지능정보화 기본법」과 「정보통신 진흥 및 융합 활성화 등에 관한 특별법」을 통합하여, 새로운 디지털 사회 질서 속에서 국민의 기본 권리를 규정한 기본법을 제정하겠다는 전략을 제시했다. 이러한 정부의 법제 전략은 「데이터 산업진흥 및 이용촉진에 관한 기본법」(기 제정 완료), 「인공지능기본법」, 「메타버스특별법」, 「사이버안보기본법」, 「디지털포용법」 등 「디지털 경제」 5대 기반법의 제정을 넘어서서, 사회적 가치지향을 보여주는 기본법을 추가로 제정하겠다는 취지로 판단된다. 이러한 정부의 입법 구상은 궁극적으로 입법 분야를 「디지털경제 기본법」과 「디지털사회 기본법」으로 구분하여 접근하고 있는 특성이 있다.

## 2. 인권 측면에서의 디지털 권리장전의 성격과 내용

### (1) 디지털 권리장전과 내재적·표출적 인권

「대한민국 디지털 전략」에 나타난 내용을 기반으로 본다면, 「디지털 권리장전」은 향후 디지털 사회가 지향해야 하는 가치 지향을 밝히기 위한 것이다. 그럼에도 「권리장전」이라 표현한 것은 ‘실정화된 법적 권리’가 아닌 ‘인권 규범’의 지향점을 제시하고자 하는 목적을 가지는 것이고, 법적 권리의 구체화는 이를 바탕으로 제정될 기본법을 통해 실현하고자 하는 것이다.

19) 관계부처 합동, 「대한민국 디지털 전략」, 2022.9.

20) 우리나라에서 활용되고 있는 ‘기본법’이라는 법제화 유형의 성격 등에 대해서는 다양한 논란이 있는 것이 사실이지만, 통상적으로 기본법은 정책적 차원의 가치 지향점을 현실화시키기 위한 입법 기술로 이해된다.

실제로 정부가 제시한 바와 같은, 권리장전을 시발점으로 하는 규범전략의 단계적 접근 방식은 타당한 측면이 있다. 따라서 제대로 된 「디지털 권리장전」을 구성하기 위해서는 이의 성격을 명확히 설정하는 것이 중요하다. 「디지털 권리장전」이 목적으로 하는 권리는 바로 ‘인권’을 의미하는 것이다. 이 지점에서 유의해야 하는 것은 ‘인권’이 가지는 의미이다. 이러한 인권은 크게 두 가지로 구분하여 접근이 가능하다. ‘내재적·표출적 인권’과 ‘도구적 인권’이 그것이다.

오늘날과 같이 법적 전문성과 자율성이 고양된 사회에서는 인권을 ‘도구적 인권’의 측면에서 접근하는 경향이 강하다. 이는 인권의 역할을 법과 제도의 형태로 구체화하여 인간의 존엄성을 보장하겠다는 데 방점을 두는 것을 의미한다. 즉 이러한 견지에서는 법제화되지 않은 인권은 의미가 없는 것이 된다.<sup>21)</sup> 물론 이러한 인권의 도구적 역할에 주목하는 접근방식은 근대 인권의 정립과정에서 중요한 역할을 해왔다는 점은 당연하다.

그러나 도구적 인권 관념은 인권의 역할을 매우 협소한 실정법적 틀 속에서 개별 권리 침해의 문제에만 집중하도록 하는 한계를 노정한다. 바로 이 지점에서 현대사회의 새로운 인권을 이야기하면서도 중국에는 인권 그 자체를 무의미한 추상적 구호로만 여기게 만드는 결과를 초래하는 것이다. 따라서 이러한 문제점을 극복하기 위해 ‘내재적·표출적 인권’의 관점을 재인식해야 할 필요가 있다. 인권의 내재적 역할은 인간의 존엄성을 최고의 가치로 승인함으로써 개인을 정치적 단위 주체로 정립하는 기능을 의미한다. 인권의 표출적 역할은 인간의 권리를 보편적으로 선포함으로써 개인의 자유와 자율을 인간 사회 발전과 민주정치를 위한 핵심 동력으로 선양하는 기능을 의미한다.<sup>22)</sup>

‘내재적·표출적 인권’과 ‘도구적 인권’이라는 양 관점 중, 사실 「디지털 권리장전」에서 의도하고 있는 인권은 현실적으로 ‘내재적·표출적 인권’에 해당한다. 그 이유는 이 권리장전을 통해 디지털 전환 또는 심화<sup>23)</sup> 사회의 새로운 보편적 권리에 관한 국내외 담론을 우리나라가 선도하겠다는 취지를 가지는 것이기 때문이다. 이와 달리, 권리장전 전략이 상정하고 있는 「디지털사회 기본법」 제정은 도구적 인권 측면에서의 전형적인 접근이라고 할 수 있다.

21) 조효제, 『인권의 지평: 새로운 인권 이론을 위한 밑그림』(후마니타스, 2016), 10면.

22) 조효제, 위의 책, 11면.

23) 「대한민국 디지털 전략」에서는 디지털 전환이라는 용어보다는 “디지털 심화”라는 용어를 사용한다. 이는 단순한 디지털화를 넘어선다는 점을 강조하기 위한 것으로 보인다. 관계부처 합동, 「대한민국 디지털 전략」, 2022.9, 12면.

## (2) 디지털 권리장전의 내용

그렇다면 「디지털 권리장전」의 내용은 무엇이 되어야 하는가. 이에 관해서는 다양한 논란이 있을 수 있다. 다만 현실적으로 생각해보면, 디지털 전환사회 또는 디지털 심화사회라고 하더라도 과거와는 다른 새로운 인권, 즉 실체적 권리를 추가적으로 언급할 만한 것들이 무엇이겠는가 하는 의문이 있는 것이 사실이다.

2023년 9월 25일 정부는 디지털 환경에서의 자유와 권리, 공정과 안전 등의 원칙을 담은 「디지털 권리장전」을 발표했다.<sup>24)</sup> 이는 국제사회가 공동으로 추구해 나갈 보편적인 미래상을 우리나라가 선도적으로 제시하고자 한 것으로, 일종의 가치지향에 관한 선언이라고 할 수 있다. 그러한 만큼 다양한 전문가 등의 자문 과정을 거쳤으며, 간담회 및 디지털 신질서 협의체를 운영해 가면서 이를 공론화 하는 작업을 진행한 바 있다. 「디지털 권리장전」은 총 6장, 28개조로 구성되어 있는데, 아래의 권리장전 체계에서 확인할 수 있는 바와 같이 구체적인 권리를 규정하기보다는 전반적인 미래 디지털 사회의 지향점을 선언한 것이라고 할 수 있다.

〈표 1〉 「디지털 권리장전」의 체계

장 구분	조문 목록
제1장 기본원칙	제1조 (자유와 권리의 보장)
	제2조 (공정한 접근과 기회의 균등)
	제3조 (안전과 신뢰의 확보)
	제4조 (디지털 혁신의 촉진)
	제5조 (인류 후생의 증진)
제2장 디지털 환경에서의 자유와 권리 보장	제6조 (디지털 접근의 보장)
	제7조 (디지털 표현의 자유)
	제8조 (디지털 다양성 존중)
	제9조 (개인정보의 접근·통제)
	제10조 (디지털 대체수단 요구)
	제11조 (디지털 근로·휴식의 보장)

24) 과학기술정보통신부, 「보도자료: 대한민국이 새로운 디지털 규범질서를 전 세계에 제시합니다!」, 2023.9.25.

제3장 디지털에 대한 공정한 접근과 기회의 균등	제12조 (공정경쟁의 촉진)
	제13조 (디지털 자산의 보호)
	제14조 (디지털 리더러시 향상)
	제15조 (데이터 접근 보장)
	제16조 (사회 안전망 강화)
제4장 안전하고 신뢰할 수 있는 디지털 사회	제17조 (디지털 기술의 윤리적 개발과 사용)
	제18조 (디지털 위험 대응)
	제19조 (디지털 프라이버시의 보호)
	제20조 (건강한 디지털 환경 조성)
	제21조 (아동·청소년의 보호)
제5장 자율과 창의 기반의 디지털 혁신의 촉진	제22조 (디지털 혁신활동의 자유)
	제23조 (디지털 규제 개선)
	제24조 (디지털 혁신 지원)
	제25조 (디지털 전환에 따른 갈등 조정)
제6장 인류 후생의 증진	제26조 (지속 가능한 디지털 사회)
	제27조 (국가 간 디지털 격차 해소)
	제28조 (디지털 국제규범 등을 위한 협력)

「디지털 권리장전」의 정식 명칭은 「디지털 공동번영사회의 가치와 원칙에 관한 헌장: 디지털 권리장전」이다. 사실 명칭상으로는 당초 「대한민국 디지털 전략」에서 밝힌 ‘권리장전’의 취지가 다소 후퇴한 느낌을 주는 것이 사실이다. 기존 전략상으로는 ‘권리’에 주안점을 두는 듯 하는 느낌을 주었으나, 명칭으로만 보더라도 디지털 사회의 가치와 원칙(지향점)을 제시하는 데 기본 목적이 있다는 점을 확인할 수 있다.

실제 「디지털 권리장전」의 내용도 그러하다. 직접적으로 인권 및 권리와 관계되어 있는 내용은 ‘제2장 디지털 환경에서의 자유와 권리 보장’에 나타나 있다. 최근 국제적으로 자주 언급되고 있는 ‘디지털 접근권’을 필두로 ‘디지털 표현의 자유’, ‘디지털 다양성’, ‘개인 정보 접근·통제’, ‘디지털 대체수단 요구’, ‘디지털 근로·휴식의 보장’ 등의 내용이 여기에 포함된다. 물론 이 밖에도 권리나 인권 보장과 연계되어 있는 내용들도 권리장전의 다른 부분들<sup>25)</sup>에 존재한다고 볼 수도 있을 것이다. 다만 권리장전의 편제 및 조문 표현 등

25) 제13조 (디지털 자산의 보호), 제15조 (데이터 접근 보장), 제16조 (사회 안전망 강화), 제19조 (디지털 프라이버시의 보호), 제22조 (디지털 혁신활동의 자유) 등.

을 고려할 때 직접적으로 권리로서의 성격을 부여한 것은 제2장 부분이라고 할 수 있다.

### (3) 디지털 권리장전 비판적 고찰

그렇다면, 이러한 「디지털 권리장전」이 과연 당초 의도한 전략적 목적을 달성할 수 있는 체계와 내용을 가진 것인지에 대해 검토할 필요가 있다. 실제 디지털 전환 또는 심화 상황 속에서 새로운 권리 개념이 충분히 생성될 여지가 있고, 또한 세계 각국은 과거와는 다른 권리 보장 환경에 대해 고민하고 있는 것이 사실이다. 그런 의미에서 우리나라의 이번 「디지털 권리장전」은 우리나라가 국제사회의 일원으로서 변화를 선도하겠다는 포부를 표명하고 있을 뿐만 아니라, 실제 권리장전의 제6장에서는 ‘인류 후생의 증진’이라는 다소 국제적 차원의 지향점을 제시하고 있기도 하다.

그런데 개별 조문의 내용들을 살펴보면, 과연 그것이 우리가 과거에 상정하지 못했던 권리 개념, 더욱 정확하게는 인권 달성의 목표를 제시하고 있는 것인지에 대해서는 의문이 든다. 특히 제2장에서 자유와 권리의 보장이라는 표제 하에 제시하고 있는 권리 또는 인권 개념들 그 자체는 이미 현재의 국제 인권규범들이나 우리나라의 헌법 체제 속에서도 충분히 수용 및 논의되어 온 것들이라고 할 수 있다. 물론 그러한 권리 및 인권의 가치 지향을 국가 공동체가 선언함으로써 보다 명확한 정책적 지향점을 제시하고자 한다는 점에서는 전혀 의미가 없는 것은 아니지만, 그렇다고 하더라도 단순 조문식으로 이를 나열하는 것이 과연 어떠한 의미를 가질 것인지에 대해서는 의문이 있다. 실제 우리나라 정부는 이미 과거에도 ‘종합계획’, ‘전략’ 등의 명목으로 유사한 내용들에 대해 정책적 의지를 표명해 왔다는 점에서 보더라도 「디지털 권리장전」의 의미가 다소 퇴색된 측면이 있다.

특히 조문 형식으로 권리를 나열하는 방식은, 앞서 언급한 바 있는 ‘내재적·표출적 인권’과 ‘도구적 인권’의 경계를 모호하게 한다는 점에 유의할 필요가 있다. 연혁적으로 볼 때, 소위 ‘권리장전’의 형식에 관해서는 일의적인 기준 등이 존재하지 않는다. 따라서 실제 권리장전의 그것은 매우 다양할 수 있는 것이 사실이다. 그러나 실제 「디지털 권리장전」은 우리나라 법령 형식과 동일한 조문체계의 입법기술을 활용함으로써, ‘내재적·표출적 인권’의 관점이 강조되어야 할 권리의 내용이 결과적으로는 권리 침해여부를 법적 논의의 틀 속으로 협소화시키는 듯한 인상을 준다. 또한 앞서 소개한 바와 같이 정부가 당초 이러한 권리장전의 내용을 곧바로 「디지털사회 기본법」에 국민들의 기본적 권리로 구체화하겠다고 한 점에서 권리장전의 형식은 더욱 문제점을 가지는 것이다.

추측건대, 정부가 「디지털 권리장전」 정립 전략을 계획하면서, ‘내재적·표출적 인권’과 ‘도구적 인권’이라는 두 가지 관점을 정확하게 인식하지 못했기 때문에, 위에서 언급했던 문제점이 발생할 것이 아닌가 한다. 현재 국내 인권담론의 지배적 경향은 도구적 인권 관념에 입각해 있다고 볼 수 있는데, 이는 인권의 문제를 법적 권리 침해의 문제와 등치시키는 경우가 빈번하다는 사실로 확인할 수 있는 부분이다. 실무적으로는 이러한 관념적 입장에서 권리장전을 고민하면서, 결국 도구적 활용을 전제로 한 실체적 권리 개념을 도출하고자 노력했을 것이다. 그러나 현실적으로 새로운 권리 실체나 가치를 도출하기는 당연히 어려웠을 것이고, 그러다보니 기존의 인권 관념을 유지하는 선에서 디지털 전환 또는 심화에 어울리는 용어들로 변경하는 방식을 활용했을 것이다.

물론 현행 법제에서 수용하고 있는 근대 초기의 인권 관념과 가치는 디지털 전환사회에서 상황적 변화에 직면하고, 그에 따라 과거에는 상정하지 못했던 권리 및 인권 관념이 충분히 탄생할 수 있다. 또한 인간의 존엄, 자유, 평등 등 기본적인 인권 가치들의 구체적인 달성 방식(권리 구체화를 위한 입법이나 정책 추진 내용 등)도 과거와는 충분히 달라질 여지가 높다. 그러나 현 시점에서 논해지는 디지털 전환 또는 심화의 맥락은 이제 초입단계로 접어들고 있는 것이어서, 아직까지 새로운 인권 및 권리 관념의 탄생이나, 법을 통한 권리 구체화의 방식에 대해서 즉자적으로 언급하기 어려운 상황이다.

결국 향후 인권이나 권리 관념의 변화를 수긍한다고 할지라도, 현재 국가 공동체적 수준에서 필요한 조치는 변화하는 기술 환경이 다소 본원적인 인권의 가치에는 어떠한 영향을 미치는 것인지, 그리고 그러한 인권적 가치를 보호하기 위한 최소한의 법적 조치는 어떻게 변화해야 하는지에 대해 면밀한 견지에서 분석하는 것이 필요하다. 그렇다고 하여 이러한 분석 그 자체로 과업이 완성되는 것은 아니다. 인권 및 권리의 종국적인 구체화는 정치적으로 구성되어질 수밖에 없는 성격을 가지고, 그래야만이 수용성과 타당성을 가지는 권리 관념이 생성될 수 있다. 따라서 시민적 담론 절차는 매우 필수적인 것이다. 즉 통상 권리장전에서 선언되어 있는 권리는 엄밀하게 말하자면 법적 권리가 아닌 ‘정치적·도덕적 권리’라고 할 수 있다. 이러한 정치적·도덕적 권리는 당대의 시대적 상황에 부합하는 방향으로 구체화되기 마련이다. 그리고 그러한 구체화는 적극적인 시민참여와 담론과정을 거치면서 현실적인 법적 권리로 변모하게 된다.

그런데 정부는 2023년 9월 25일 「디지털 권리장전」을 공표하면서, 「디지털 권리장전」을 기준으로 삼아 디지털 심화시대의 쟁점들을 해소하고 구체적인 법·제도를 정비하기 위한 실질적인 노력들을 이어나갈 계획이라고 밝히면서, ‘디지털 심화대응 실태조사’를 통



해 관계부처의 정책·제도들을 「디지털 권리장전」에 따라 정비될 수 있도록 지원할 예정이라고 밝혔다.<sup>26)</sup> 이러한 정부의 계획에 대해서는 구체적으로 실무 수준에서 다양한 문제점이 제시될 수 있겠지만, 무엇보다도 ‘디지털 심화대응 실태조사’를 권리장전 도출 과정에서 수행하지 않았다는 점을 명시적으로 보여주고 있다.<sup>27)</sup> 즉 구체적인 분석이나 담론절차 없이 실제적인 권리나 가치지향을 먼저 제시하고, 이를 구체화하기 위한 실태분석을 사후에 진행하겠다는 것이다. 이렇게 되면 실상 「디지털 권리장전」은 당초 의도한 목적에 부합하는 역할을 수행하기 어렵고, 단순한 정부의 일상적 정책 표명에 그치게 된다. 결과적으로 바람직하게는 다소 시일이 걸리더라도 「디지털 권리장전」 제정 과정에서 디지털 전환 및 심화 상황의 인권(권리)에 대한 영향을 분석하여 시민들에게 먼저 제시하고, 이를 기반으로 시민들이 자율적으로 참여할 수 있는 담론절차 및 정책반영 절차를 운영하는 것이 선행되었어야 했다.

#### IV. AI시대 인권보장을 위한 입법방향:

##### 영향분석의 선제적 제도화의 요청

##### 1. AI와 영향분석

###### (1) 초기 AI 윤리기준 정립과정

앞서 AI 등에 기반한 디지털 전환 또는 심화 상황에서 기술 환경의 현격한 변화가 미치는 인권에 대한 영향을 선제적으로 분석해야한다고 언급했다. 이와 관련하여 다양한 사례들이 있을 수 있지만, 여기에서는 AI와 관련한 EU의 초기대응 과정에 대해 살펴보고자 한다. EU의 경우에는 최근 「인공지능법안」이라는 구체적인 법제화 논의까지 이어지고 있는데, 여기에서는 「신뢰 가능한 AI 윤리지침」(Ethics Guidelines for Trustworthy AI) 정립 및 후속절차와 연계되어 있는 초기 상황을 확인해 보도록 한다.

26) 과학기술정보통신부, 「보도자료: 대한민국이 새로운 디지털 규범질서를 전 세계에 제시합니다」, 2023. 9.25.

27) 물론 정부가 수차례 밝힌 바와 같이, 권리장전 수립 과정에서는 다양한 관계 전문가 등의 의견을 수렴했다. 그러나 이러한 의견수렴이 인권 및 가치 구성을 위한 담론절차에 근접하는 성격을 가지는지에 대해서는 의문이 있다. 특히 정부는 일반 시민들의 의견수렴을 위한 ‘디지털 공론장’(<http://beingdigital.kr>, 2023.11.17. 방문)을 운영하기도 하였는데, 실제로는 사후적이고 형식적인 의견수렴의 과정이었다는 평가가 지배적이다.

AI 관련 입법정책에 관해, 초기 EU의 담론형성을 주도한 유럽 의회는 유럽집행위원회에 AI에 관한 영향분석(impact assessment)<sup>28)</sup> 실시 촉구를 핵심 내용으로 하는 ‘Civil Law Rules on Robotics 권고안’을 2017년 1월 결의한 바 있었다. 이는 의회 결의로서의 속성을 가지는 것이기 때문에 전형적이고 정식화된 집행위원회의 영향분석 절차를 거친 것은 아니지만, 의회 자체의 다양한 영향분석 절차 및 이와 병행된 의견수렴(consultation) 절차를 거쳤다는 점을 확인할 수 있다.<sup>29)</sup> 특히 집행위원회 권고(촉구)를 위한 근거를 마련하기 위하여, EU의회는 결의 이후 대대적인 공개 의견수렴 절차를 추가적으로 진행하기도 하였으며,<sup>30)</sup> 결의안 검토 과정에서 EU의 입법지원기구, 특히 STOA와 같은 기술영향분석 조직의 보고서들이 광범위하게 검토 및 수렴되었다는 사실을 확인할 수 있다.

EU의회의 Civil Law Rules on Robotics 보고서는, 이후 진행된 집행위원회의 AI 정책 수립 방향 설정에 결정적인 기여를 한 것으로 보인다. 이러한 배경에서 집행위원회는 2018년 유럽에서의 ‘AI 개발을 진흥하기 위한 커뮤니케이션’을 채택했으며,<sup>31)</sup> 2019년에는 EU 회원국의 국가들의 AI 국가전략을 조화시키기 위한 조정계획을 발표하였다(EU 이사회 승인).

이러한 과정들을 거쳐 중국에는, 2019년 4월 집행위원회가 지원하고 결과를 공표한 Ethics guidelines for trustworthy AI로 이어졌다. 이는 집행위원회가 앞서 EU의회 권고(Civil Law Rules on Robotics)에 포함되어 있었던 윤리지침(부록)을 구체화한 것이라고 할 수 있다. 집행위원회는 이 과정에서 52명의 독립적인 전문가들로 구성된 High-Level Expert Group on AI를 출범시켰다. 이 전문가 그룹이 작성한 문서는 EU에서 윤리적 AI 시스템 개발을 촉진하고 확보하는 방법에 대한 지침을 제공하는 것을 목표로 했다. 전문가 그룹의 운영은 실질적으로 전문가 의견수렴 과정의 일환이라고 평가할 수 있을 것이다. Ethics guidelines for trustworthy AI에는 동 가이드라인 취지에 부합하는 AI 및 그 운용에 대해 평가할 수 있는 요건들을 포함한 평가 목록을 제시하고 있다. 이를 기반으로 파일럿

28) ‘영향분석’이라는 용어는 국내에서 ‘영향평가’로도 표현된다. 다만 이 글에서는 논의의 일관성을 위해 ‘영향평가’로 용어가 정착된 일부 사례를 제외하고는 ‘영향분석’이라는 용어를 사용하기로 한다.

29) EU Parliament(JURI), Report with recommendations to the Commission on Civil Law Rules on Robotics, 2017(<https://www.europarl.europa.eu/committees/en/report-with-recommendations-to-the-commi/product-details/20170202CDT01121>, 2023.11.17. 방문).

30) 공개 의견수렴 결과에 관한 개괄적인 내용은 EPRS, Public consultation on robotics and Artificial Intelligence - First (preliminary) results of public consultation, 2017을 참고할 것(<https://www.europarl.europa.eu/cmsdata/128500/juri-committee-eprs-consultation-robotics-results.pdf>, 2023.11.17. 방문).

31) European Commission, Communication Artificial Intelligence for Europe, 2018 (<https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe>, 2023.11.17. 방문).

평가 프로세스 및 의견수렴 등 피드백 과정을 진행했다. 특히 이러한 피드백을 보다 역동적으로 수행하기 위하여 The European AI Alliance<sup>32)</sup>라는 이해관계자 포럼을 운영했다는 점도 특징이었다.

이상과 같은 비구속적인 윤리 논의를 어느 정도 궤도에 올려놓은 뒤, EU 집행위원회는 2020년 2월 인공지능에 관한 백서<sup>33)</sup>를 발간하였다. 이후 보다 구체적인 법적 규제 및 정책 방향성을 결정하기 위한 영향분석 개시절차(inception impact assessment)를 7월 23일부터 9월 10일까지 실시하였다.<sup>34)</sup> 이러한 절차 운용은 향후 보다 구체적인 영향분석 논제들을 설정하기 위해 진행되고 있는 것으로, 전형적인 집행위원회 영향분석의 초기 절차에 해당한다고 평가할 수 있겠다. 특히 이러한 절차에서는 광범위한 의견수렴을 거쳤으며, 접수된 모든 의견들은 웹사이트를 통해 공개했다. 실제로 EU의 경우 의견수렴 과정을 사실상 모든 영향분석 과정과 병행하는 절차적 특징이 있다.

이상에서 살펴본 바와 같이, EU는 AI 관련 윤리기준 정립 및 입법정책적 대응 초기에 매우 광범위하고 면밀한 영향분석 절차를 진행했다는 사실을 확인할 수 있다. 이는 기본적으로 새로운 기술 환경의 변화가 미칠 영향들을 사전에 분석하여, 보다 명확한 AI 관련 정책적 가치의 방향성을 설정하기 위한 시도였다고 할 수 있을 것이다. 현실적으로도 EU가 현재, 그 내용상의 타당성 여부는 별론으로 하더라도, 「인공지능법안」을 주축으로 AI 입법에 관한 논의를 세계적으로 자신감 있게 주도하고 있는 상황은 이러한 초기 대응방식에 힘입은 바가 크다고 하겠다.

## (2) AI 입법과 영향분석의 필수적 견련성

영향분석 절차 진행의 필요성과 효과는 비단 AI 관련 규범정립 초기 과정에서만 필요한 것이 아니다. AI 기술의 속성상 직접적인 법적 규제 및 입법의 내용 중에도 영향분석이 필연적으로 내재될 수밖에 없다. AI 기술 및 서비스의 불확정적인 성격은 이에 대한 규제 목적 달성의 불확정성을 야기한다. 이러한 측면에서 대부분의 입법 시도 사례들은 전통적

32) The European AI Alliance(<https://ec.europa.eu/digital-single-market/en/european-ai-alliance>, 2023.11.17. 방문).

33) EUROPEAN COMMISSION, WHITE PAPER - On Artificial Intelligence - A European approach to excellence and trust, 2020. 2. 19.

34) <https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Requirements-for-Artificial-Intelligence>(2023.11.17. 방문) 참조. 의견수렴 결과는 <https://ec.europa.eu/digital-single-market/en/news/white-paper-artificial-intelligence-public-consultation-towards-european-approach-excellence>(2023.11.17. 방문)에 정리되어 소개되어 있으며, 제시된 의견 모두를 문서 형태로 확인할 수 있다.

인 법적 규제와 같이 포괄적이고 일의적인 규제 기준을 정립해 활용하기보다는, 위험 기반 접근방식(risk-based approach)을 취하고 있다.

EU의 「인공지능법안」도 ‘제안이유 (14)’<sup>35)</sup>에서 이를 명확하게 밝히고 있으며, 이러한 취지에서 동 법안은 AI의 위험을 금지된 위험(unacceptable risk), 고위험(high risk), 제한적 위험(limited risk), 최소 위험(minimal risk)으로 분류하고, 각 범주에 부합한다고 판단되는 규제방식을 설정하고 있다. 이러한 위험 기반 접근방식은 실제 기술적 불확정성이 전제된 상황에서 보편적으로 논의되는 규제방식이다. 결국 제기되는 위험이 무엇인지는 확인하고, 그러한 위험의 정도에 따라 규제 방식과 수준을 달리 선택하게 된다는 것을 의미한다. 그렇다면 비법적인 정책적 대응이 필요한지, 아니면 법규범적 규제방식 중 어떤 유형을 선택할 것인지의 문제는, 제기되는 위험이 무엇인지에 대한 규명을 요청한다. 바로 이 지점에서 등장하는 것이 ‘영향분석’이다.

EU 국가들은 이미 영향분석을 보편적으로 활용하고 있음에도, 최근 EU 의회가 의결한 「디지털 시대의 인공지능에 대한 결의안(Resolution on artificial intelligence in a digital age)」<sup>36)</sup>은 영향분석의 중요성을 재차 강조하고 있다. 동 결의안은 ‘입법조치 및 관련 로드맵’에 관한 내용에 관해 기술하면서, 제123항에서는 “AI와 같은 분야에서 새로운 디지털 법안을 발표하기 전에, EU 집행위원회가 적절한 예측 및 위험 분석을 통해 심층적인 사전 영향분석 수행을 촉구 한다”고 제시하고 있다. 이는 위험기반 접근 방식의 전제로서 영향분석의 중요성을 시사하는 것이다.

영향분석의 중요성은 AI 규제에 대해 그간 미온적인 태도를 취해온 미국에서도 나타나고 있다. 대표적인 것이 최근에 수정 발의된 「알고리즘 책무성 법안(Algorithmic Accountability Act)」<sup>37)</sup>이다. 이 법안은 인간의 개입이 없는 ‘자동화된 의사결정 시스템(Automated Decision System)’뿐만 아니라, 인간이 의사결정을 하더라도 AI 알고리즘 등의 도움을 받는 ‘증강된 중요 의사결정 프로세스(Augmented Critical Decision Process)’에도 자체적인 영향분석을 실시하고, 그 결과(연례) 보고서를 연방 FTC(Federal Trade

35) (14) AI 시스템에 대해 비례적이고 효과적인 규제 수단을 도입하려면 명확하게 정의된 위험 기반 접근방식을 준수해야 한다. 그러한 접근방식은 규칙의 유형과 내용을 AI 시스템이 생성할 수 있는 위험의 강도와 범위에 맞게 세분화해 나가야 한다. 따라서 모종의 AI 관행을 억제하기 위하여, 고위험 AI 시스템의 요건과 관련 운영자에 대한 의무를 규정하고, 특정 AI 시스템에 대한 투명성 의무를 규정하는 것이 필요하다.

36) European Parliament resolution of 3 May 2022 on artificial intelligence in a digital age(2020/2266(INI))

37) Wyden, R., Boone, C., & Clarke, Y., Algorithmic Accountability Act of 2022(S.3572, H.R.6580), 2022.2.3.

Commission)에 제출토록 하고 있다. 이러한 영향분석 체계의 구성은 기본적으로 AI 알고리즘을 활용하는 기업들의 책무성을 강화하기 위한 수단으로 기능하도록 의도한 것이며, 연방 FTC 차원에서는 AI 알고리즘 활용 상황과 그에 따른 위험에 대해 지속적으로 추적 및 평가할 수 있는 기반을 제공하겠다는 의도인 것이다.

## 2. AI 인권과 영향분석 제도화 방향

### (1) 인권의 실질화를 위한 AI 영향분석의 필요성

앞서 다른 국가들의 사례에서 확인한 바와 같이, AI와 관련한 입법정책은 초기 기준 정립 과정에서는 물론이고, 실제 법의 정립 및 집행 과정 전반이 영향분석과 긴밀한 연관성을 가진다는 점을 알 수 있다. 이는 불확정성이 전제된 AI의 속성과 위험 기반 접근방식의 상황적 속성을 고려할 때 불가피한 대응방식인 것이다.

AI 기술의 다양한 활용이 논해지고는 있지만, 실제적으로는 현시점은 AI 활용 보편화의 초기단계라고 할 수 있다. 그런 의미에서 인권 보장 또는 달성을 위하여 국가가 어떠한 입법정책적 조치를 수행해야 하는지를 명확하게 확정하기란 쉽지 않다. 그럼에도 손에 잡히는 실체적(법적)권리 개념을 고안해 내기 위한 시도들이 빈번하게 이루어지고 있다. 이는 앞서 언급한 바와 같이 ‘도구적 인권’ 관념이 정책 실무 현장에서 지배적인 관행으로 자리 잡았기 때문이다.

사실 이제까지 우리나라의 입법 관행에 있어 ‘도구적 인권’ 관념이 지배적일 수밖에 없었던 것은, 거시적인 인권의 목록과 그것을 달성하기 위한 가치적 지향점을 새롭게 설정하기보다는, 근대사회 발전 이래로 이미 정립된 인권적 가치 토대 위에서 그것을 어떻게 현실에서 법적 권리로 구체화할 것인가에 주력해 왔기 때문이다. 그러나 이제는 AI 등 신기술(emerging technologies)의 발전으로 인하여 과거와는 다른 차원의 인권 환경이 대두되고 있다. 이에 따라 단순히 확립된 인권 관념과 가치 지향만을 추구하기에는 어려움이 있고, 그러하기에 전통적 인권 관념 및 가치에 대해 다시금 숙고해야 하는 시점인 것이다.

물론 현실 변화로 인하여, 전통적으로 논해져 왔던 인권의 구체적 목록이 현격하게 변화하는 것은 아닐 것으로 보인다. 다만 중요한 것은 인권의 내재적 역할과 표출적 역할을 어떻게 현실에 구현할 것인지를 고민하고, 인권 보장을 위하여 기존에는 고민하지 못했던 가치 지향점을 명확히 해나가는 작업이 중요하다. 물론 이 과정에서 아직 정립되지 않은

새로운 인권 목록들이 발견될 수는 있을 것이지만, 그것은 어디까지나 새로운 가치지향이 어느 정도 현실화된 이후의 문제이다.

결과적으로 단순한 AI 위험에 대응하기 위한 규제입법이 아니라, 그 기저에 존재하는 인권 보장이라는 근원적 입법정책의 목적을 구체화하기 위한 전략이 필요하다. 이런 측면에서 현 단계의 세계 각국은 AI 등 디지털 신기술을 둘러싼 다양한 영향들을 분석하는 데 주력하고 있는 것이다.

## (2) 기존 영향분석 제도의 활용 가능성

그렇다면 AI와 관련하여 우리나라에서 활용할 수 있는 기존의 영향분석 제도들은 어떤 것들이 있는지 파악해 보아야 할 필요가 있다. 물론 현실적으로는 개별 정부부처들이 사안별로 수행하는 정책연구보고서들도 다분히 영향분석으로서의 기능을 가진다. 그러나 AI 사안의 중요성과 다양성을 고려한다면, 제도적으로 활용할 수 있는 정례화·체계화된 영향분석 제도는 물론이고, 인권 보장 및 달성에 실질적으로 기여할 수 있는 실효적인 영향분석 제도가 요청된다.

첫째,, AI와 관련한 영향분석에 있어 가장 먼저 언급할 수 있는 것은 바로 「과학기술기본법」 제14조에 규정되어 있는 ‘기술영향평가’ 제도가 존재한다. 이는 과학기술의 발전이 경제, 사회, 문화, 윤리, 환경 등에 미치는 영향을 사전에 평가하고 그 결과를 정책에 반영하기 위해 운영되는 제도이다. AI 기술도 과학기술의 일환이라는 점에서 당연히 기술영향평가의 대상이 될 수 있다. 실제로 2015년에 수행된 기술영향평가 주제가 ‘인공지능 기술’이었다.<sup>38)</sup> 그러나 기술영향평가는 새로운 기술에 대해 이벤트적이고 일회적으로 평가절차를 수행하는 성격을 가지고 있을 뿐만 아니라, 실제 영향평가 결과도 정책적으로 활용하기에는 구체성이 떨어지는 내용들을 담고 있어, 실제 AI에 관한 영향분석 제도로써 활용하는데 한계를 가진다.

둘째, 보다 직접적으로 AI에 한정한 영향분석에 관한 제도로는 「지능정보화 기본법」 제56조의 ‘지능정보서비스 등의 사회적 영향평가’가 있다. 이는 AI와 같은 지능정보서비스 등의 활용과 확산이 사회·경제·문화 및 국민의 일상생활 등에 미치는 영향에 대하여 조사·평가하는 제도를 의미한다. 그러나 이 영향평가는 서비스를 대상으로 하고 기술에 관해서는 앞서 언급한 기술영향평가에서 수행토록 하고 있어, 기술과 서비스 간 모호한 구분

38) 미래창조과학부·한국과학기술기획평가원, 『2015 기술영향평가-제2권 인공지능 기술』, 2016.

으로 인해 혼선을 유발하고 있다. 또한 구체적인 영향평가 기준 및 절차에 대해서도 위임 규정의 부재로 인하여 법령상 구체화할 수 없다는 한계도 존재한다. 결과적으로 이 영향 평가는 2020년 도입된 이후 아직까지 제대로 수행된 바가 없다.

셋째, AI 사안들을 모두 포괄하기에는 한계가 있지만 상당한 연관성을 가지는 영향분석 제도로 「개인정보 보호법」 제33조의 ‘개인정보 영향평가’ 제도가 있다. 이 영향평가는 개인정보파일의 운용으로 인하여 정보주체의 개인정보 침해가 우려되는 경우에 그 위험요인의 분석과 개선 사항을 도출하기 위해 활용되는 평가이다. 이는 앞서 언급한 영향평가들과는 달리 포괄적인 기술이나 서비스가 아니라, 개별적인 사안 및 대상의 위험성에 집중한다는 특성을 가진다. 현실적으로 개인정보 영향평가는 공공기관들에 대해서만 의무화되어 있는데, AI의 활용이 민간 영역에서 더욱 빈번하게 이루어진다는 점에서 AI와 관련한 영향 분석 수단으로 활용하기에는 대상 범주 측면의 한계가 있는 제도이다.

이상에서 살펴본 바와 같이, AI 일상화 시대에 인권의 새로운 전환과 실질화를 위한 정책 운영에 적실하게 활용할 수 있는 영향분석 제도는 부재한 상황이다. 물론 개별 영향분석 제도들이 나름의 의미를 가지지 않는 것은 아니지만, 보다 상시적이고 포괄적인 영향분석 제도로서 활용하는 데에는 분명 한계가 있다.

### (3) 영향분석의 제도화 방안

AI 인권의 새로운 실질화에 종국적인 목적을 둔 영향분석을 제도화함에 있어서는, AI가 활용되는 사회적 현실 상황을 고려할 필요가 있다. AI 기술은 단순히 특정 영역에 한정된 것이 아니라 일상생활의 거의 모든 영역에서 활용될 것으로 예상된다. AI 기술은 그 자체가 인간의 판단을 지원하는 기계적·자율적 의사결정을 추구한다는 측면에서, 인간의 판단이 개입될 수 있는 거의 모든 영역에서 활용될 여지가 높다. 이와 더불어, AI 기술은 민간의 주도적이고 창의적인 서비스 개발 경쟁의 맥락에서 그 활용 영역을 넓혀갈 것으로 예상된다. 과거 우리나라의 경우 대체적으로 과학기술 및 정보통신 기술은 공공부문의 주도로 확산되어 온 측면이 있지만, AI 기술의 경우에는 민간 사업자 간 경쟁이 더 나은 서비스를 확산 및 발전시키는 기반이 되고 있다. 이러한 상황을 고려할 때, 다음과 같은 영향 분석 제도화 방안을 고려해야할 것으로 보인다.

첫째, ‘입법 및 정책’을 대상으로 하는 포괄적인 영향분석의 제도화가 필요하다. 과거와 같은 기술 환경이었다고 한다면, 특정 기술에 직접적으로 연관된 입법정책적 영향을 분석

하는 데 초점을 두는 것만으로도 의미가 있을 것이다. 그러나 디지털 전환 또는 심화 상황에서 AI는 매우 미시적 수준의 변화는 물론이고 국가 공동체적 거버넌스에도 상당한 영향을 미칠 것이 예견되기 때문에, 국가 전반의 정책 영역을 포괄하는 영향분석을 제도화하는 것이 필요하리라고 본다. 이와 관련해서는 서구 주요 국가들의 경우 통상적인 입법 및 정책에 관한 영향분석을 제도화하고 있으며,<sup>39)</sup> 이는 실제 AI 관련 영향분석을 수행하는데 활용되어 왔다. 즉 주요 국가들은 AI에 한정한 AI 관련 영향평가 제도를 활용하기보다는, 앞서 살펴본 바와 같이 기존에 제도화되어 있는 포괄적 영향분석 제도를 활용하고 있다. 우리나라의 경우, 이러한 입법 및 정책에 관한 포괄적 영향분석의 필요성이 지속적으로 제기되어 왔지만, 제도화가 이루어지지는 못했다. 물론 정부입법 영역에 규제영향분석 제도가 존재하지만, 이는 말 그대로 국민의 권리를 제한하거나 의무를 부과하는 법령을 의미하는 ‘규제’에만 초점을 두는 것이고, 또한 정부입법에서만 활용되는 제도라는 점에서 한계가 있다. 따라서 향후 포괄적 영향분석 제도의 정립에 관심을 기울일 필요가 있다.<sup>40)</sup>

둘째, 직접적으로 ‘AI 기술 및 서비스’를 대상으로 하는 영향분석을 제도화할 필요가 있다. AI 기술 및 서비스는 현재도 마찬가지로 향후에도 민간의 경쟁 속에서 계속 신규화되어갈 것으로 예상된다. 그런 의미에서 새로운 방식의 AI 기술이나 서비스가 인권에 미치는 영향을 미치거나 위험을 유발할 가능성이 있다면, 이에 대해 상시적으로 평가할 수 있는 체계를 확보하는 것이 중요하다. EU의 「인공지능법안」이나 미국의 「알고리즘 책임성 법안」의 경우에도 위험의 측면에서 이러한 영향분석 체계를 상정하고 있다. 만일 우리나라에 이러한 영향분석 체계를 제도화하고자 하는 경우에는 앞서 언급한 ‘개인정보 영향평가’의 규정 사례를 일부 참조해볼 수 있을 것으로 판단된다. 또한 앞서 국가인권위원회의 「인공지능 개발과 활용에 관한 인권 가이드라인」 권고<sup>41)</sup>에도 나타난 바와 같이, 직접적으로 ‘인공지능 인권영향평가’<sup>42)</sup> 시행에 대한 논의도 진행되고 있다. 다만 이러한 인공지능 인권영향평가 제도화를 위해서는 법률 차원의 입법이 반드시 필요할 것으로 보이는

39) 영향분석 절차에 관한 가장 대표적인 사례는 지속적인 변화를 거듭하고 있는 EU라고 할 수 있다. 세부적인 절차 및 분석 방법에 관한 사항에 대해서는 European Commission, Better Regulation Guidelines, 2021.11.; European Commission, Better Regulation Guidelines Toolbox, 2023.7을 참고할 것.

40) 국회 차원에서 입법영향분석 및 규제영향분석을 제도화하기 위한 노력은 현재 진행중에 있다. 다만 규제영향분석의 경우에는 ‘규제’에 한정한 영향분석을 상정하고 있다는 특징이 있다. 따라서 다소 포괄적인 영향분석을 제시하고 있는 법안으로 김태년의원 대표발의, 「국회입법조사처법 일부개정법률안」(의안번호 제1888호), 2020.7.14.; 「국회법 일부개정법률안」(의안번호 제1891호), 2020.7.14을 참고해볼 수 있겠다.

41) 국가인권위원회 결정, 「인공지능 개발과 활용에 관한 인권 가이드라인」, 2022.4.

42) 이에 대한 개괄적인 사항은 유승익 외, 『인공지능 인권영향평가 도입 방안 연구』(국가인권위원회, 2022) 참고.



데, 그 이유는 특정 수준 이상의 위험을 유발하는 서비스 개발자 및 제공자들에게 영향평가를 의무화하는 것은 헌법상 기본권으로서의 성격을 가지기 때문이다.

마지막으로 AI를 둘러싼 입법정책의 인권 친화적 방향성 확립을 위하여 영향분석을 위와 같은 방식으로 제도화한다고 할 때, 방법론적 측면에서도 특별한 고려가 필요하다. 일반적으로 영향분석은 사회과학 방법론을 활용한 실증적인 접근을 지향한다. 그러다보니, 실무적으로 이를 양적 연구방법론 일변도로 영향분석을 생각하는 경향이 존재한다. 그러나 최근에는 이러한 문제점을 인식하고 질적 연구방법론, 즉 이해관계자에 대한 공적 의견수렴 절차의 체계화에 주목하는 경향이 증가하고 있다. 특히나 AI 관련 사안의 경우 양적 통계나 데이터만으로는 관련 기술의 영향이나 위험을 사전에 예측하는 데 분명 한계가 있기 때문에, 결국 공적 의견수렴 및 피드백 과정이 매우 중요하다. 이는 앞서 살펴본 EU 대응사례에서도 확인할 수 있는 부분이다. 따라서 보다 나은 영향분석 절차 도입을 위해서는 최대한 공적 의견수렴 및 피드백 절차를 내포한 제도화가 가능토록 주의를 기울여야 한다.<sup>43)</sup>

## V. 결론을 대신하여

AI 시대의 새로운 기회와 위기에 대응하기 위한 다양한 논의들이 국내외적으로 활발하게 이루어지고 있다. 특히 최근 수년간 이루어진 대응은 주로 AI 윤리기준 및 사회적 대응의 원칙적 방향을 설정하는 데에 노력이 집중되어 있는 양상이다. 그 이유는 직접적인 대응(규제)입법을 추진하기에는 발생 가능한 사안들에 대한 판단 기준 및 요건이 명확하지 않기 때문이다. 즉 포괄적인 대응방향을 설정함과 동시에, 이에 기반한 새로운 가치질서를 공동체적으로 구성해 나가기 위한 공감대를 구축해 나가겠다는 의미로 이해될 수 있다. 이 과정에서 중요한 것은 결과적으로 인권 보장 및 달성이라는 공동체 운영의 궁극적 목적이라고 할 수 있다.

사실 ‘인권’이라는 관념은 그것을 실현하는 수단인 법적 권리에 비하여 상당히 추상적인 성격을 가지기 때문에, AI 시대의 새로운 입법에서 기준이 될 수 있는 구체적이고 명확한 지침을 제공해 주지 못하는 성격을 가진다. 예를 들어, 우리나라의 AI 윤리기준 등의 정립

43) 이러한 영향분석의 제도화에 관한 세부적인 사항은 심우민, 『국정과제 이행을 위한 과학기술 법체계 중장기 발전방향 연구』(국가과학기술자문회의, 2018), 111면 이하; “디지털 전환과 사회갈등: 입법학적 분석과 대안”, 『유럽헌법연구』 37, 2021, 509면 이하 참고.

사례 등을 살펴보면, 기존에도 당연한 것으로 여겨지던 추상적 원칙의 반복 나열 수준을 크게 넘어서지 못하고 있다. 그 결과 실제 AI기술 개발 및 활용 현장에서는 당해 기준들을 행위준칙으로 고려하지 못하는 현상이 발생한다. 따라서 향후에는 AI 윤리기준이나 그것의 배경을 이루는 인권적 가치를 바라보는 시각의 변화가 필요하다.

이러한 배경 속에서, 최근 한국 정부는 AI를 주축으로 하는 디지털 사회의 변화에 대응하는 원칙과 기준의 마련의 일환으로 「디지털 권리장전」을 공표하였다. 이번 시도 또한 그것이 가지는 긍정적 의미에도 불구하고, 실제적으로는 기존 AI 윤리기준 정립에서와 마찬가지로 오류를 반복하고 있는 것으로 판단된다. 디지털 전환 또는 심화 상황에서 도구적으로 활용할 수 있는 실체적 권리 도출에 주력하다보니, 결과적으로 당연한 권리들을 재차 나열하는 상황을 연출하고 있다. 이러한 교착 상태를 극복하기 위해서는 새로운 디지털 사회가 추구해야 하는 권리, 좀 더 구체적으로는 비법적인 인권이 가지는 본원적인 가치와 역할을 어떻게 달성토록 할 수 있는지에 주목할 필요가 있다. 즉 새로운 실체적 권리 관념의 도출보다는 인권의 가치를 공동체적으로 어떻게 실현시킬 것인지에 관한 구성적 절차의 제공에 초점을 둘 필요가 있다.

이를 위해서는 급변하는 디지털 환경이 이미 정립된 인권 및 그 가치에 어떠한 영향을 미치는지에 대한 분석이 선행되어야 한다. 즉 AI와 이를 둘러싼 사회 상황에 대한 엄밀한 영향분석을 수행할 필요가 있다. 또한 AI 기술 및 그것의 활용의 불확정성을 고려한다면, 모종의 기준과 가치를 정립한다고 할지라도 이는 재차 지속적으로 변화하는 속성을 가질 것이기 때문에, 영향분석 절차를 정례화·제도화할 필요가 있다. 또한 이렇게 영향분석을 통해 도출된 결과는 직접적으로 일의적인 사회적 대응방향을 제시해주지는 못한다. 결국 공동체 구성원들의 사회 변화에 대한 인식과 공감대를 바탕으로 새로운 인권의 가치와 그 실현 방안이 구체화될 때, 그것이 우리 공동체가 지향해야할 새로운 가치 기준으로 활용될 수 있다. 따라서 해외 주요 국가들의 영향분석 절차가 거의 필수적으로 공적 의견수렴 및 피드백 과정을 포함하고 있는 점에 착안하여, 우리의 경우에도 영향분석 제도화에 이러한 부분들이 포함될 수 있도록 해야 한다. 물론 그렇다고 하여 영향분석이 인권보장을 위한 만병통치 처방이라고 주장하는 것은 아니다. 다만 이는 AI를 둘러싼 인권보장 입법을 위한 매우 기초적이고 필수적인 기반 중 하나로서 역할을 할 것이다.

결국 중요한 것은 새로운 실체적인 권리나 인권 관념이 아니라, 인권의 가치와 목적을 달성하기 위한 포괄적이고 합리적인 절차의 제공이다. 따라서 향후 AI 시대의 인권보장을 위한 정책 추진에 있어서는 이러한 절차적인 측면에 비중을 두는 노력이 이루어지기를 기

대한다. 이를 통해 디지털 전환 및 심화 맥락에서 인권의 가치와 역할이 올바르게 정립될 수 있고, 또한 그러한 토대 위에서 의미 있는 AI 입법이 구체화될 수 있을 것이다.

## ■ 참고문헌

과학기술정보통신부, 「보도자료: 대한민국이 새로운 디지털 규범질서를 전 세계에 제시합니다」, 2023.9.25.

관계부처 합동, 「대한민국 디지털 전략」, 2022.9.

\_\_\_\_\_, 사람이 중심이 되는 「인공지능(AI) 윤리기준」, 2020.12.

교육부, 사람의 성장을 지원하는 「교육분야 인공지능 윤리원칙」, 2022.8.

국가인권위원회 결정, 「인공지능 개발과 활용에 관한 인권 가이드라인」, 2022.4.

김태년의원 대표발의, 「국회법 일부개정법률안」(의안번호 제1891호), 2020.7.14.

\_\_\_\_\_, 「국회입법조사처법 일부개정법률안」(의안번호 제1888호), 2020.7.14.

미래창조과학부·한국과학기술기획평가원, 『2015 기술영향평가-제2권 인공지능 기술』, 2016.

심우민, “디지털 전환과 법교육의 미래: 법사회학적 접근 방식의 시사점”, 『법교육연구』 17(3), 2022

심우민, “디지털 전환과 사회갈등: 입법학적 분석과 대안”, 『유럽헌법연구』 37, 2021.

심우민, 『국정과제 이행을 위한 과학기술 법체계 중장기 발전방향 연구』(국가과학기술자문회의, 2018)

유승익 외, 『인공지능 인권영향평가 도입 방안 연구』(국가인권위원회, 2022)

조효제, 『인권의 지평: 새로운 인권 이론을 위한 밑그림』(후마니타스, 2016)

AIネットワーク社会推進会議事務局(総務省情報通信政策研究所調査研究部), 「AI 開発ガイドライン」(仮称) の策定に向けた国際的議論の用に供する素案の作成に関する論点, 2016.12.28.

EPRS, Public consultation on robotics and Artificial Intelligence-First(preliminary) results of public consultation, 2017.

EU Parliament(JURI), Report with recommendations to the Commission on Civil Law Rules on Robotics, 2017.

European Commission, Better Regulation Guidelines Toolbox, 2023.7.

\_\_\_\_\_, Better Regulation Guidelines, 2021.11.

\_\_\_\_\_, Communication Artificial Intelligence for Europe, 2018.

\_\_\_\_\_, Proposal for a regulation laying down harmonised rules on Artificial Intelligence(Artificial Intelligence Act) and amending certain Union legislative acts(COM/2021/206 final), 2021.4.21.

\_\_\_\_\_, WHITE PAPER-On Artificial Intelligence-A European approach to excellence and trust, 2020. 2. 19.

European Parliament resolution of 3 May 2022 on artificial intelligence in a digital

age(2020/2266(INI))

High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, European Commission, 2019.4.8.

High-Level Expert Group on Artificial Intelligence, The Assessment List for Trustworthy Artificial Intelligence(ALTAI) for self assessment, 2020.7.17.

OECD, Rights in the digital age: Challenges and ways forward, OECD DIGITAL ECONOMY PAPERS No. 347, 2022.12.

Regulation (EU) 2019/1150 of the European Parliament and of the Council of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services.

Wyden, R., Booke, C., & Clarke, Y., Algorithmic Accountability Act of 2022(S.3572, H.R.6580), 2022.2.3.

## ■ Abstract

# The Direction of AI Legislation for Human Rights

- The Need for Impact Assessment in light of the Digital Bill of Rights -

Shim, Woomin\*

Attempts to establish new AI normative frameworks are being made around the world. While it is true that there are trials of direct AI regulatory lawmaking in some cases, it can be said that the mainstream is to present normative standards and principles for AI-based society. In the end, these are likely to be referred as specific directions for AI-related legislation, which is expected to become more serious in the future. Therefore, at this stage, it is necessary to concretely establish the purpose of the AI normative frameworks. Ultimately, the purpose of such normative frameworks is to ensure substantive human rights. In this context, South Korea recently promulgated the Digital Bill of Rights. However, by focusing only on the instrumental role of human rights, it merely rephrased the rights already guaranteed through international norms and constitutions. In order to contribute to legislation in the AI era that can actually guarantee human rights, it is necessary to proactively propose a process to analyze or assess the impact and build a value consensus based on it, rather than simply relying on normative declarations. In particular, the need to institutionalize impact assessment cannot be overemphasized in that the context of AI technology and its use inevitably contains indeterminacy, and requires a risk-based approach rather than standardized legal regulation.

• Key Words Artificial Intelligence, Human Rights, Digital Bill of Rights, AI Legislation, Impact Assessment

---

\* Associate Professor(Ph.D. in Law), Department of Social Studies Education, Gyeongin National University of Education