



# 직장 내 AI의 신뢰성 보장과 근로자 보호

## — 주요국 규제 정책 사례 및 시사점

손혜원

직장 내 AI를 통한 혁신의 지속가능성을 높이기 위해서는 기술적 진보와 직장 내 AI의 신뢰성 확보 조치가 조화를 이루어야 할 것이다. 국제적인 협력·조정과 함께, 신뢰할 수 있는 AI를 개발·사용하고 AI로 인한 근로자 권리 침해, 노동환경 악화 등의 위험 요소를 최소화하며 근로자를 보호하는 포용적이고 균형적인 정책 대응이 필요하다.

### 1 직장 내 AI 신뢰성 보장의 필요성

AI 기술이 빠르게 진보·확산되고 있고, 머지않은 시기에 AI가 직장에서도 다양하고 광범위하게 사용될 것으로 예상된다. 직장 내의 AI 사용은 생산성 향상, 근로환경 개선, 새로운 직업 창출 등의 긍정적 측면이 있으나, AI의 인력 대체에 따른 일자리 감소와 AI 알고리즘을 활용한 자동 인력관리로 일자리의 질 저하 등 부정적 영향이 우려되기도 한다.

직장 내에서 AI를 통한 혁신이 지속되기 위해서는 기술적 진보도 중요하지만, 직장 내 AI 사용을 신뢰할 수 있는 기반이 마련되어야 할 것이다.

2019년 5월, 채택된 OECD 권고안<sup>1)</sup>의 AI 원칙에 따르면, “신뢰할 수 있는 AI(trustworthy AI)”란 AI의 개발·사용이 안전하고 개인정보 보호·공정성·노동권 등 기본권을 존중하며, 고용 관련 결정 방식이 투명하고 이해할 수 있어야 한다. 또한, 사용자·근로자·구직자가 업무에 AI가 활용됨을 인지할 수 있어야 하고, 문제 발생 시 책임소재가 분명해야 함을 의미한다.<sup>2)</sup>

1) OECD, “Recommendation of the Council on Artificial Intelligence”, OECD/LEGAL/0449, 2019.05.

2) OECD, 『OECD Employment Outlook 2023 - ARTIFICIAL

Intelligence and Labour Market』(2023) 보고서에서 직장에서 사용되는 AI의 신뢰성을 보장하는 정책이 단기적으로는 안전하고 책임감 있는 AI 개발·사용을 촉진하고, 장기적으로는 혁신의 장애물을 피할 수 있게 할 것으로 전망한 바 있다.<sup>3)</sup>

이 글은 OECD의 보고서를 토대로 직장 내에서 사용되는 AI의 신뢰성을 보장하기 위한 세계 각국의 AI 관련 규제 논의를 근로자 권리 보호, 직장 내 AI의 투명성·설명가능성 및 책임성 보장 사례 위주로 살펴보고, 향후 관련 입법 및 정책 논의에 참고하고자 한다.

### 2 근로자 권리의 보호

#### (1) 인간의 결정과 상호작용

실제로 구인·업무 할당·보너스·승진 등 고용 관련 결정이 AI로 완전 자동화된 사례는 아직까지는 거의 없다. 이는 기술적 어려움, 비용 부담 또는 AI 활용 기술 부족 때문일 수 있으나, 고용에 관한 중요한 결정이 인간 개입 없이 AI를 통해 이루어지는 것을 금지하는 규제가 그 바탕이 되고 있다.<sup>4)</sup>

INTELLIGENCE AND LABOUR MARKET』, 2023.07, p.183.

3) OECD, 앞의 글, p.185.



이러한 규제 기초가 된 OECD의 AI 원칙은 AI 시스템에 인간의 개입과 감독을 보장하는 보호장치를 구현할 것을 권고하고 있으며, 각국의 AI 관련 규제도 이와 맥락을 같이 하고 있다.<sup>5)</sup>

또한, EU의 「일반 데이터 보호규정」(The EU General Data Protection Regulation; GDPR)은 고용 관련 데이터의 보호를 강조하며, 완전 자동화된 알고리즘 인력관리를 사실상 금지하고 있다.<sup>6)</sup>

최근 합의된 EU의 「AI ACT」도 고용 관련 결정에 사용되는 AI를 “고위험 AI 시스템”으로 정하여 적합성 평가와 기본권 영향평가 의무를 부과하며, 직장에서 근로자 감정을 추론하는 AI 시스템 사용을 금지한다.<sup>7)</sup>

이 같은 최소한의 안전장치가 있기는 하나, 직장 내에서 AI가 활용되는 것만으로도 근로자의 자율성과 자유의지를 저해하고, 근로자 간의 교류를 감소시켜 사회적 고립과 복지 저하를 초래하며, 더 나아가 근로자의 업무 관련 존엄성에 악영향을 미칠 가능성이 있다.<sup>8)</sup> 산업안전보건 규제가 근로자의 정신건강을 보호하는 토대가 될 수는 있으나, AI로 인한 심리·사회적 위험까지 규율할 수 있을지는 확실하지 않고, 직장 내 AI 사용으로 인한 인간 교류 감소의 위험성을 방지하기 위한 규제 논의는 드문 실정이다. 예외적으로, 독일은 AI 조사위원회가 작성한 보고서를 통해 “인간은 사회적 존재로서 직장에서 다른 인간과 사회적으로 상호작용하고, 인간의 피드백을 받으며, 스스로를 노동력의 일부로 인식하도록 보장해야 한다”고 명확하게 적시하고 있다.<sup>9)</sup>

## (2) 개인정보 침해

AI 시스템은 광범위한 개인정보를 처리하는 경우가 많기 때문에 개인정보 침해 피해가 더 커질 수 있다. 특히, 민감하지 않은 대량 데이터를 활용하여 개인 민감정보까지 추론할 수 있으므로 위험성이 더 크다.

AI로 처리되는 개인정보 보호에 관한 규제로는 EU의 GDPR이 있다. 정보시스템에서 개인정보가 사용될 때, 정보주체의 명시적 동의를 구하도록 규정하는데, 많은 국가에서 유사한 체계를 취하고 있다.

그러나, 고용관계의 정보·권력 비대칭성과 AI 시스템에 대한 근로자의 이해가 부족한 현실을 감안할 때, 사전동의의 효과 여부에 대해서는 논란이 있다. 유럽 데이터 보호 위원회(European Data Protection Board)는 근로자의 동의가 자유롭게 이루어지지 않았을 가능성이 있기 때문에 사용자가 동의를 근거로 현재나 미래의 직원 개인정보를 처리하는 것은 문제가 있다고 보았다.<sup>10)</sup>

또한, GDPR은 일반적인 내용이라 ‘고용’ 측면에서 개인정보 처리를 규율하기에는 한계가 있으므로 각국의 규제가 GDPR 조항을 단순 반복할 것이 아니라 보다 구체적으로 규정해야 한다는 견해도 있다. 이를테면, 특정한 종류·목적에 위한 데이터 수집의 금지, 직원의 개인정보 접근권 확대를 통한 정보 비대칭성 개선, 자동해고 같은 고위험 결정의 완전 자동화 금지, 인간 개입에 대한 명확한 기준 확립, 이의제기, 자동화된 결정에 대한 인간 검토의 권리, 알고리즘 관리의 급속한 진보로부터의 보호 등을 고용의 맥락에서 명시할 필요가 있다는 것이다.<sup>11)</sup>

## (3) 직장에서의 편견과 차별

AI는 인간이 만든 데이터로 학습하기 때문에 인간

4) OECD, 앞의 글, p.191.

5) OECD (2019), Recommendation of the Council on Artificial Intelligence, OECD Paris,

6) Official Journal of the European Union (2016), Regulation (EU) 2016/679 of the European Parliament and of the Council

7) 심소연, 「규제중심의 유럽연합 인공지능법(EU AI ACT)」, 『최신 외국 입법정보』, 국회도서관, 2024.04. pp.3~4.

8) Briône(2020), “My boss the algorithm: an ethical look at algorithms in the workplace”, <<https://www.acas.org.uk/research-and-commentary/my-boss-the-algorithm-an-ethical-look-at-algorithms-in-the-workplace/paper>> (최종검색일 2024.04.23.)

9) Deutscher Bundestag Enquete-Kommission (2020), Artificial Intelligence – Social Responsibility and Economic, Social and Ecological Potential

10) EDPB(2020) Guidelines 05/2020 on consent under Regulation 2016/679 European Data protection Board <<http://edpb.europa.eu>> (최종검색일 2024.04.23.)

11) Abrahams, H., M. Silberman and J. Adams-Prassl(2022), The Need for Employee-specific Data Protection Law: Potential Lessons from Germany for the EU, European Law Blog, <<https://europeanlawblog.eu/2022/09/30/the-need-for-employee-specific-data-protection-law-potential-lessons-from-germany-for-the-eu>> (최종검색일 2024.05.10.)

의 편견이 여과 없이 들어간 결과물을 도출할 가능성이 있다. 예를 들어, 데이터 알고리즘에 따라 채용 공고 볼 수 있는 지원자의 범위가 고학력 남성 위주로 맞춰진다거나, 면접 후보자 선택이나 직원 성과평가 등에서 유사한 차별이 일어나는 경우 등이다. 차별금지법 같은 기존의 법률은 AI로 인한 차별을 규제하기에는 한계가 있을 수 있으므로 데이터에 내재된 편향성을 완화하는 방안 마련이 필요하다.

직장 내 편견·차별방지 규제 사례로는 미국의 「알고리즘·AI 사용으로 인한 장애인 채용 차별 방지 지침」(Algorithms, Artificial Intelligence, and Disability Discrimination in Hiring)<sup>12)</sup>과 프라이버시 침해와 여성 및 소수 민족 차별 우려에서 기인한 안면인식 기술 관련 규제를 참고할 수 있다.<sup>13)</sup>

미국 메릴랜드 주는 채용 면접에서 안면인식 사용을 금지하는 법안을 통과시켰고,<sup>14)</sup> 유엔인권이사회 보고서는 안면인식 사용의 일시금지를 요구했으며<sup>15)</sup>, 유럽평의회(Council of Europe)는 얼굴 이미지에서 성격 특성, 내면의 감정, 정신건강을 감지하는 얼굴 분석 도구의 금지를 촉구한 바 있다.<sup>16)</sup>

#### (4) 결사의 자유와 단체교섭권

AI 사용은 근로자·근로자대표 및 사용자 간의 권력 불균형과 정보 비대칭으로 근로자의 교섭력을 약화시킬 수 있으며, AI 기반 모니터링으로 노조 활동을 제약할 우려가 있다.

스페인에 이에 대응하여 「라이더법」(Lei Rider)을 통해 디지털 플랫폼 기업으로 하여금 근로조건이나 고

용 결정에 사용하는 알고리즘 정보를 근로자대표에게 제공하도록 의무화하였다. 이는 AI 알고리즘에 관한 정보 및 협의권을 확대하고, 단체교섭을 통한 세부 협상이 가능하게 되었다는 점에서 의미를 가진다.<sup>17)</sup>

#### (5) 산업안전보건

직장 내 AI 오류로 인한 기계 오작동이 근로자의 안전을 위협할 수 있고, 단순·반복 업무가 감소되는 대신 복잡한 업무가 증가하거나 작업강도가 세지면서 근로자의 건강과 복지를 해칠 수 있으며, AI 도입에 따른 새로운 기술 습득의 필요성과 모니터링 강화로 스트레스가 증가될 수도 있다.

노동법 및 산업안전보건 규제가 직장 내 AI 사용에 따른 근로자의 보호에 일반적으로 적용되나, 독립적으로 인식하고 결정할 수 있는 AI의 특성과 작업환경 변화를 고려할 때, AI 사용으로 제기되는 문제를 효과적으로 해결할 수 있는 보완책을 논의해야 할 것이다.<sup>18)</sup>

### 3 직장내 AI의 투명성·설명가능성·책임성

#### (1) 직장 내 AI의 투명성

“투명성”이란 AI가 어떤 알고리즘으로 결과를 도출했는지를 투명하게 공개해서 공적 검토가 가능하도록 해야 함을 의미한다.<sup>19)</sup> 직장 AI의 사용에 관한 투명성 보장은 근로자·사용자가 채용, 모니터링, 승진 등의 고용 관련 결정에 대한 AI의 영향 여부와 결과가 도출된 근거 및 방식을 확인할 수 있을 때 가능해진다.

투명성 보장을 위하여 EU의 「AI ACT」는 제한적 위험 AI 시스템에 AI 생성 콘텐츠에 대한 표시의무와 저작권이 적용되는 학습데이터 공개 의무를 규정하고 있다.<sup>20)</sup> 또한, EU의 「플랫폼노동의 근로조건개선지침

12) 국회도서관, 『초거대 AI 한눈에 보기』, FACT BOOK 2023-5호, p.96.

13) Buolamwini, J. and T. Gebru (2018), “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification”, Proceedings of Machine Learning Research, Vol. 81, pp.1-15.

14) Fisher et al.(2020), Bill Text:MD HB1202, <https://legiscan.com/MD/text/HB1202/id/2169556> (최종검색일 2024.05.10.)

15) UN Human Rights Council (2021), A/HRC/48/31: The right to privacy in the digital age – Report of the United Nations High Commissioner for Human Rights.

16) Council of Europe (2021), Guidelines on Facial Recognition, pp.19-26.

17) Pérez del Prado, D. (2021), “The legal framework of platform work in Spain: The new Spanish “Riders’ Law””, Comparative Labor Law and Policy Journal, p.4.

18) Jarota, M.(2021), “Artificial intelligence and robotisation in the EU—should we change OHS law?”, Journal of Occupational Medicine and Toxicology, Vol.16/1, pp.5-6.

19) 이상욱, 「AI 윤리의 쟁점과 과제」, 『AI 윤리의 쟁점과 거버넌스 연구』, 2021.12. p.26.

(Directive of the European Parliament and of the Council on Improving working conditions in platform work)은 디지털 노동 플랫폼이 플랫폼 노동자의 근로조건에 관한 중요 결정에 관해 AI의 사용·기능 정보를 언제 제공해야 하는지 그 형식과 시점을 명시하고 있다.<sup>21)</sup>

## (2) 직장 내 AI 사용에 관한 설명가능성

직장 내 AI 사용에 관한 설명가능성은 AI 알고리즘에 의한 고용 관련 결정에 관한 설명이 근로자·근로자 대표 및 사용자가 이해할 수 있는 쉬운 언어로 제공되어야 함을 의미한다. AI 알고리즘에 기반한 고용 관련 결정의 논리가 불투명하고 이해하기 힘들면 AI의 부정적 영향이 의심되더라도 그 결정을 바로잡기 어렵다. 이는 근로자의 권리를 침해할 뿐 아니라 AI 기반 결정을 근로자가 수용할 가능성도 떨어뜨리고 AI의 직장 내 적용을 어렵게 할 수 있다.

캐나다 「인공지능 및 데이터법」(Canadian Artificial Intelligence and Data Act ;AIDA)과 「소비자 개인정보보호법」(Consumer Privacy Protection Act ; CPPA)은 자동화된 의사결정 시스템의 개발자·사용자에게 AI 시스템이 특정 결과에 도달한 방법에 관해 쉬운 언어로 된 설명을 제공하게 함으로써 투명성과 설명가능성을 규율하고 있다.<sup>22)</sup>

## (3) 직장 내 AI의 책임성 제고

직장에서 사용되는 AI의 책임성 보장은 신뢰성을 확보하기 위한 기반이 된다.<sup>23)</sup> 그러나, AI는 무형성, 복잡성, 불투명성 및 예측가능성 제한 등의 특성으로 인해 문제 발생 시 책임소재 규명과 피해구제가 어려울 수 있다.

직장 내 AI의 책임성을 제고하기 위해서는 근로자

가 AI 알고리즘으로 인한 결정에 이의를 제기하거나 수정할 수 있는 절차를 마련하고, 차별금지 원칙의 준수와 AI 시스템의 안전한 작동을 보장할 책임이 있는 AI 행위자를 특정할 수 있는 체계를 구축하는 방안을 고려할 필요가 있다.<sup>24)</sup>

이와 관련하여 유럽연합 집행위원회(European Commission)의 AI 책임지침안(AI Liability Directive : AILD)은 AI 피해자에 대한 입증책임을 완화하고, 고위험 AI 시스템 제공자에 대한 증거 공개를 규정하여 알고리즘 정보에 보다 쉽게 접근할 수 있도록 하고 있다.<sup>25)</sup>

## 5 시사점

이상에서 살펴본 직장 내 AI와 관련된 국제적 규제 사례는 신뢰할 수 있는 직장 내 AI의 개발·사용을 필요로 하고 직장 내 AI가 촉발할 수 있는 위험성을 직면한 여러 국가에 참고가 되고 있으며, AI 관련 규율의 국제적 격차를 방지하는 데에 기여하고 있다.

직장에서 사용되는 AI와 노동환경의 미래는 어떻게 설계되고 적용되는지에 따라 크게 달라질 수 있으므로 투명성·설명가능성·책임성의 확보를 통한 신뢰할 수 있는 AI 개발·사용과 함께, AI로 인한 근로자 권리 침해, 일자리 질 악화 등의 위험 요소를 최소화하고 근로자를 보호하는 포용성이 균형을 이루게 하는 면밀한 정책 대응이 필요하다. 또한, AI 규제 논의에 관한 국제적 협력 및 조정을 지속하면서, AI에 대한 우리 사회의 전반적 이해도를 높이는 데에도 정책적 역량과 자원을 투입하여 신뢰 구축 및 도약의 발판을 마련할 수 있도록 해야 할 것이다.

『이슈와 논점』은 국회의원의 입법활동을 지원하기 위해 최신 국내외 동향 및 현안에 대해 수시로 발간하는 보고서입니다.

20) 심소연, 「규제중심의 유럽연합 인공지능법(EU AI Act)」, 『최신 외국 입법정보』 2024-04호, 국회도서관, p.3.

21) European Commission (2021), Directive of the European Parliament and of the Council on Improving working conditions in platform work, pp.16-17.

22) OECD, 앞의 글, p.198-199.

23) 신봉근, 「인공지능(AI)의 민사책임-EU의 입법을 중심으로」, 『민사법학』 제105호, p.215.

24) Loi, M. (2020), People Analytics must benefit the people – An ethical analysis of data-driven algorithmic systems in human resources management, Algorithm Watch, pp.43-49.

25) European Commission (2022), New liability rules on products and AI to protect consumers and foster innovation, <[https://ec.europa.eu/commission/presscorner/detail/en/IP\\_22\\_5807](https://ec.europa.eu/commission/presscorner/detail/en/IP_22_5807)>

