

공개토론회

공공기관 AI 채용과

인공지능 규범

일시 | 2022년 9월 7일(수) 오후 2시

장소 | 국회의원회관 제1간담회실

주최 | 민주사회를위한변호사모임 디지털정보위원회,
사단법인정보인권연구소, 진보네트워크센터,
민병덕 의원 (더불어민주당)

순서

2:00 ~ 2:10	개회	민병덕 의원 (더불어민주당)
	사회	권오성 교수 (성신여자대학교 법학부)
2:10 ~ 3:00	발제	공공기관 AI 채용 현황과 문제점 김민 정책활동가 (진보네트워크센터)
		공공기관 인공지능 정보공개거부처분취소소송 진행 경과와 제도적 과제 김하나 변호사 (해우법률사무소 / 민주사회를위한변호사모임 디지털정보위원회 위원장)
3:00 ~ 3:45	토론	사단법인 정보인권연구소 장여경 상임이사
		개인정보보호위원회 김직동 과장 (신기술개인정보과)
		공정거래위원회 선중규 과장 (시장감시총괄과)
		국가인권위원회 이진백 주무관 (인권정책과)
		국민권익위원회 박지원 과장 (청탁금지제도과)
3:45 ~ 4:00	플로어 토론 및 참석자 전체 토론	



민병덕 의원 | 더불어민주당 안양시동안구갑

민주사회를위한변호사모임 디지털정보위원회, 사단법인 정보인권연구소, 진보네트워크센터, 그리고 민병덕 의원실이 공동주최하는 「공공기관 AI 채용과 인공지능 규범」 공개 토론회를 찾아주신 여러분 환영합니다. 더불어민주당 경기 안양시동안구갑 출신 국회의원 민병덕입니다.

바야흐로 인공지능의 세상입니다. 이미 바둑이나 컴퓨터게임 등의 승부에서 인공지능이 인간에 승리하는 일이 놀랍지 않지만 결국 이것도 인간이 발전시킨 기술이기에 자랑스럽기도 하다는 점이 인공지능의 양면성이라 생각합니다.

공공기관 채용 과정에 있어서 인공지능 프로그램의 도입은 편리성과 공정성이 기대되지만, 얼마 전 있었던 재판에서 과정의 투명성이나 도구의 성능, 관련 규범 준수가 보장되고 있지 않음이 확인되어 우려를 사고 있습니다. 특히 인공지능 채용 과정에서 발생할 수 있는 개인정보보호의 문제, 알고리즘 비공개에 따른 채용비리의 가능성 등은 많은 이들이 우려하고 있는 지점입니다.

이러한 문제에 있어서 핵심은 AI의 성능이나 신빙성 그 자체라기보다는, 그 도구를 안전하게 활용하는데 필요한 ‘법적 규범의 마련’에 있다고 볼 수 있습니다. 시스템이 올바르게 작동할 수 있도록 채용과정에서 공공기관들은 책무성과 법적규범, 보안을 제고하는 것이 무엇보다 중요할 것입니다.

오늘 토론회는 이러한 문제들을 자세히 들여다보고 바로잡기 위한 구체적인 방안에 대해 논의하는 자리가 될 것입니다.

오늘 이 자리를 계기로 바람직한 제도적 대책과 인권 규범을 마련해 인간이 만든 기술이 제대로 인간을 위해 사용될 수 있기를 바랍니다.

발제와 토론을 비롯해 본 행사를 준비하느라 애써주신 모든 분들께 감사의 말씀을 드립니다. 또, 바쁘신 와중에 자리해주신 모든 전문가 분들께 진심으로 감사드립니다. 댁내 건강과 평안이 늘 함께하기를 기원합니다. □

공공기관 AI 채용 현황과 문제점

김민 정책활동가 | 진보네트워크센터

○ AI면접(AI역량검사) AI서류 평가와 같은 인공지능 기술을 기반으로 한 자동화된 채용 도구가 지속적으로 사용되고 있음. 차별적인 인간 면접관과 숨막히는 경쟁을 거처온 수천 수만 명의 구직자들은 이제 인공지능이라는 새로운 벽을 만나게 된 것임. 기업은 많은 양의 채용 절차를 빠르고 효율적으로 처리한다는 명목 하에 인공지능 도구에 의존하는 한편, 인간이 객관적으로 판단하기 어렵다고 알려진 개인의 잠재력과 인간적 특성을 알고리즘이 차별 없이 판별해낼 수 있다고 신뢰하는 듯함. 그러나 인공지능 도구가 내놓은 결과는 그 코드와 학습 데이터에 의도치 않더라도 숨겨져 쓰여진 차별의 가능성을 고스란히 내포하고 있을 위험성이 큼.

○ 2018년부터, 공공기관 또한 채용 과정에 AI면접 등 인공지능 기반 채용 도구를 도입하기 시작했음. 지속적으로 불거진 공공기관 채용 과정에서의 비리 등 불공정한 채용이 이슈가 되자 그 대안으로 들어온 셈이지만 인공지능은 불공정한 채용과 차별의 본질적 문제를 해결할 수 있는 도구가 아님. 오히려 인공지능의 기술적 특성으로 일반적으로 제기될 수 있는 문제들, 차별의 가능성과 그 불투명성이라는 특성 때문에 보다 불공정한 채용 과정으로 변모할 수 있음. 더불어 개인정보 자기결정권에 대한 침해, 기술로 인한 접근성 제한, 기술 자체의 결함 등 문제를 포함해 여러 기본권에 전반적인 영향을 미칠 가능성이 큼.

1. 어떤 프로그램이 사용되는가?

○ 인공지능 기반 채용 도구들은 구직자, 지원자를 선별하고 평가하는 채용 절차의 여러 단계에서 사용되고 있음. 현재 한국에서 일반적으로 활용되는 도구들은 AI면접 및 역량검사, 그리고 AI서류평가임.

AI면접 및 역량검사

○ ‘AI면접’ 또는 ‘AI역량검사’라고 불리는 일련의 채용 도구는, 온라인 영상 면접을 통해 구직자를 판단하는 프로그램과 그 절차를 말함. 온라인 영상 면접을 통해 지원자의 얼굴과 목소리를 자동으로 처리해 표정, 감정, 호감도, 언어 습관, 소통능력, 매력도, 신뢰성, 논리성 등 외형적이거나 비언어적인 특성을 분석하고 평가할 수 있다고 주장함. 일부 도구는 일련의 뇌과학 및 신경과학 기반의 게임의 수행 과정을 함께 진행하여 정답과 오답, 응답속도, 의사결정 및 학습속도 등을 평가하고 분석하기도 함.

○ 이러한 AI면접 도구의 경우 실제 재직 중인 ‘고성과자’와 ‘저성과자’의 AI면접 응시 데이터·실제 성과 데이터, 우수 면접관·주요 대기업 인사총괄·산업심리학 교수 등 ‘면접 전문가’들이 면접 및 대화 영상을 분석하고 평가한 데이터 등을 기반으로 한다 알려짐.

AI서류평가

○ AI서류평가는 지원자의 이력서, 자기소개서 등의 채용 서류를 자동으로 분석해주는 도구로, 문장 및 맞춤법 오류 분석·표절 검사·주요 내용 요약과 같은 기본적 검수 기능과 더불어 채용 서류에 점수를 매기거나 직무 적합도 등을 인공지능을 통해 자동으로 평가해주는 것으로 알려짐. 이러한 AI서류평가의 경우, 자기소개서에 포함된 단어와 문장 중 능력·경험·신념·가치관·포부·지원동기 등으로 분류된 특성을 찾아 ‘고성과자’ 및 ‘저성과자’의 특성과 유사 여부를 이용해 평가에 활용한다고 알려짐.

○ 이러한 AI서류평가 도구의 경우 50만 건 이상의 자기소개서를 학습한 AI모델을 활용하고, AI 서류 평가를 도입하는 사용 기업의 합격자 자기소개서 데이터를 기반으로 ‘우수 인재’ 패턴을 학습한다고 함.

2. 무엇이 문제인가?

누적된 차별 데이터의 문제

○ 채용 과정에 도입된 인공지능 시스템은 학습에 사용한 데이터에 따라 의도치 않더라도 부당한 차별적 결과가 나올 수 있음. 인공지능 기술은 일반적으로 인간이 생성한 데이터에 의존하므로, 인공지능 모델이 내린 결정에도 인간의 차별과 편견이 반영될 수 있는 것임. 또한 실제 직무와는 관계 없는 인간의 고유한 특성이 채용 평가의 대상이 되는 등 공정성과 정확성의 문제가 제기될 수도 있음.

○ 이는 성별, 나이, 지역, 신체 조건이나 경제적 지위, 학력 등 흔히 채용 단계에서 차별과 편견을 가져온다고 여겨지는 민감한 개인의 속성이나 정보를 제외한다 하더라도 충분히 일어날 수 있는 문제임. 어떤 데이터가 수집되고 처리되며, 어떤 요인이 서로 작용되느냐에 따라 보호 받아야 할 개인의 민감한 특성이 대리 변수를 통해 파악되고 평가기준으로 사용될 수도 있음.

○ 그 예시로, 2018년 전자상거래 등 IT기업인 아마존이 자체 개발한 인공지능 채용 프로그램을 폐기했다는 사실이 보도되었음¹⁾. 해당 프로그램이 여성과 관련된 키워드가 들어간 서류 등을 아예 탈락시키거나 낮은 점수를 주는 등 남성편향적으로 평가되는 것을 방지하거나 수정할 수 없었기 때문이라고 함. 이는 해당 프로그램이 ‘가장 우수한 성과’를 보였던 ‘백인 남성’의 직무 관련 데이터를 통해 학습되고 구축되었기 때문임²⁾. 그로 인해 지원자의 성별을 평가의 매개변수에서 제외하더라도 여성과 관련된 지원자의 속성이나 대리적인 데이터 때문에 수정조차 불가능했음.

○ 한국에서 현재 활용되는 AI면접 등 채용도구들 또한 아마존과 비슷한 데이터 및 기준에 기반해 평가를 내리고 있다고 볼 수 있음. 해당 도구들은 통상적으로 ‘고성과’ 등 직장 내에서 우수하다고 판단된 노동자의 데이터에 의존하여 평가 모델을 만들거나 기존의 ‘인사전문가’ 등이 판단한 결과에 기반하여 평가 모델을 만드는 것으로 알려짐.

1) 로이터, Amazon scraps secret AI recruiting tool that showed bias against women

<https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>

2) Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. *California Management Review*, 61, 15-42.

<https://doi.org/10.1177/0008125619867910>

3) 중앙일보, 임원 면접 ‘우수’, 알고보니 AI 평가서도 ‘최고점’...이유는?

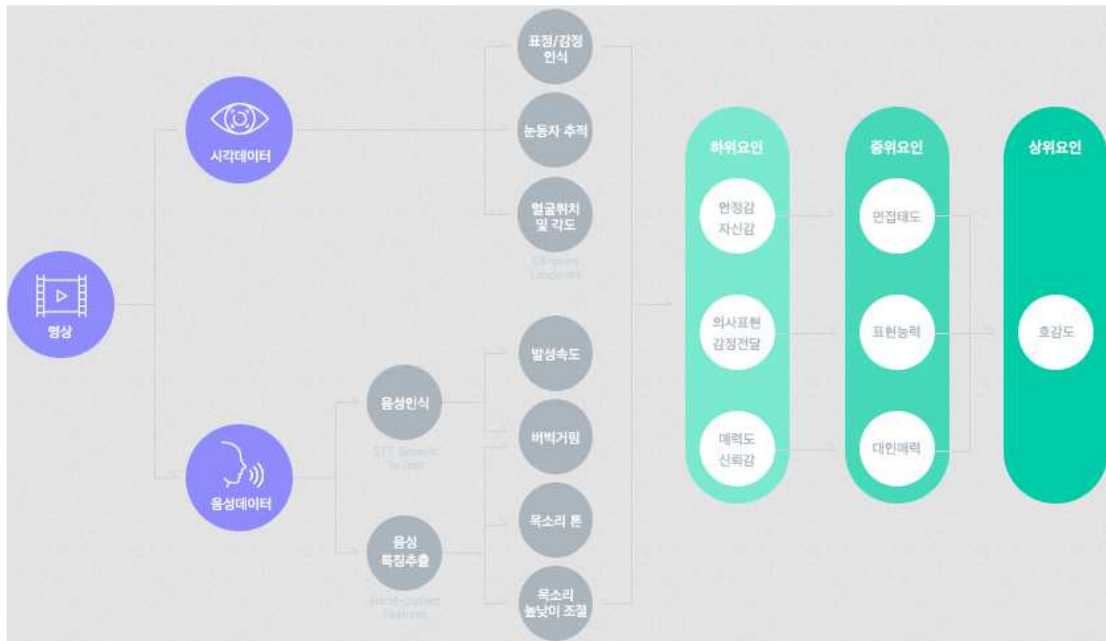
<https://www.joongang.co.kr/article/24012409#home>

마이다스 인 - AI면접 (AI역량검사)

“먼저 지원한 기업에서 높은 업무 성과를 낸 직원들의 특징점을 축적한다. 지원자는 게임이나 질의응답 등을 통해 검사에 응하면 시가 미리 학습해둔 고성과자의 특징, 회사가 바라는 인재상 등과 매치해 직무 적합도를 알려주는 방식이다.”³⁾

“ ‘우수 면접관 판단 결과 학습’, 국내 최고 면접 전문가가 평가한 1억 건 이상의 데이터를 통해 노하우를 학습한 최신 기술로 질의응답 과정에서 발생하는 실시간 반응을 분석합니다. 음성/영상 정보를 통해 지원자의 호감도, 매력도, 감정 전달 능력, 의사 표현 능력 등을 판단합니다.”⁴⁾

마이다스인 AI역량검사 백서⁵⁾



검증 방법 및 정확도	내용
학습을 위한 평가 데이터 수집	<ul style="list-style-type: none"> ▶ 1,279명의 응시자 영상에 대해 53명이 평가 (관찰 특성에 대한 레이블 생성). ▶ 응시자 1명에 대해 평가자 1~4명이 6점 척도로 평가 진행. ▶ 이분법 분류를 위하여 데이터 변환 (1~2점: 0, 5~6점: 1, 그 외 데이터는 제외). ▶ 여러 명이 평가한 경우, 다수의 점수를 택하되 판단이 갈리면 데이터에서 제외.

“방법은 이렇다. 구직자 3천 명의 동의를 얻어 3개 질문의 답변 영상을 녹화한다(영상 9천 개). 질문 답변을 6개로 쪼개고 다시 프레임 10개로 나눈다(영상 54만 개). 그리고 이를 마이 다스IT 인사담당자와 고객사 인사담당자로 구성된 ‘면접전문가’ 인력풀(200명으로 구성) 가운데 3명에게 영상, 음성, 영상+음성으로 나누어 보여준다(데이터 162만 개). 면접전문가는 앞서 제시된 ‘호감도’ 6개 항목에 대해 6점 만점으로 평가한다. 그리고 이를 얼굴의 68개 지점에 대한 변화 유형으로 학습시킨다(데이터 1억1016만 개). 정리하면, 구직자 3천 명이 3개 질문에 대해 응답한 면접영상을 면접관 200명이 3개씩 나눠 보고 호감도를 평가한 것이 1억 1016만 개의 데이터가 됐다는 뜻이다.⁶⁾”

제네시스 랩 - AI면접

“제네시스랩의 ‘뷰인터HR 솔루션’은 화상으로 촬영된 인터뷰 이미지를 전문가 그룹이 평가하는 방식이다. 전문가 그룹에는 삼성·LG그룹 등 대기업 출신의 인사담당 임원과 지원한 회사의 인사 담당자 등이 참여한다. 여기에서 데이터를 축적한 뒤 해당 데이터를 SI에게 학습시켜 면접에 활용한다.⁷⁾”

“뷰인터에서는 다양한 대기업 그룹사 회장실 출신의 인사책임자들이 자문으로 활동을 해주고 계십니다. 인사책임자로 30년 이상의 인사 경험, 특히 면접관으로 오랫동안 활동을 하셨습니다. 뷰인터에서 인사책임자 전문가분들이 가지고 있는 여러가지 평가 기준과 인사이트를 인공 지능이 배울 수 있도록 합니다.”

“뷰인터에서 보유하고 있는 수십만개 이상의 비디오 면접 영상들을 보고 위에 전문가 그룹이 메뉴얼하게 평가를 합니다. 예를 들어 영상 하나하나를 보며 채용추천점수, 소프트스킬 및 내용분석, 꼬리질문 추천 등에 대해 평가를 합니다. 신뢰도 측정을 위해 동일한 영상을 반복적으로 평가시스템에 입력하여 신뢰도를 체크하고 있습니다.⁸⁾”

4) 마이다스인, 기업에 가장 적합한 인재를 과학적, 객관적으로 판단하는 분석 기술

<https://www.midashri.com/intro/process/ai-interview/anal>

5) 마이다스인 AI역량검사 백서, 6. 소통역량 분석의 정확도 검증 <https://www.midashri.com/aicc-06>

6) 한겨레21, 투명성·공정성·신뢰성…AI면접 믿을 만할까?

https://h21.hani.co.kr/arti/economy/economy_general/49403.html

7) 중앙일보, 임원 면접 ‘우수’, 알고보니 AI 평가서도 ‘최고점’…이유는?

<https://www.joongang.co.kr/article/24012409#home>

8) 제네시스랩 블로그, 뷰인터AI영상면접 기술을 소개합니다. <https://blog.genesislab.ai/?p=11859>

무하유 (AI 서류평가)

“AI평가는 기업에서 제공하는 고성과자의 데이터 기반으로 구축된 AI 학습 데이터를 활용합니다. 이를 통해 자기소개서를 심층 분석하여 평가 시 유용한 근거를 제공합니다.⁹⁾”

03

고득점자 패턴 추출 기술



사람 평가 결과에 따라 고득점/저득점 자소서
패턴을 학습하여 점수화 합니다.



50만 건 이상의 자기소개서를 학습한
AI 모델을 활용합니다.

○ 인공지능과 데이터 기반 기술이 차별이 누적된 데이터에 의존할 경우 그 차별이 기술에 의해 다시 한 번 모방되고 확산될 위험이 있음. 그렇기에 인공지능 학습데이터의 주요 데이터라 할 수 있는 ‘고성과자’, ‘우수한 합격자’, ‘면접 전문가’, ‘인사책임자’ 등이 어떤 근거로 선정되었는지를 살펴봐야 할 것임.

○ 고용노동부의 2017년 조사¹⁰⁾에 의하면, 서류전형 합격자의 성비는 여성과 남성이 거의 동일했으나, 면접 뒤에는 여성 합격자의 비율이 69%로 낮아짐. 또한 기업 중 55.1%는 채용 시 상대적으로 선호하는 성별이 있으며, 그 성별로는 ‘남성’이 73.6%로 여성보다 2.8배 가량 많다는 조사 결과¹¹⁾도 있음.

○ 성별에 따른 고용 통계¹²⁾에 따르면, 그 격차가 지속적일 감소해 왔음에도 불구하고 여성 고용률은 57.7%, 남성 고용률은 75.2%였음. 또한 경력이 높아질수록 남성 비율이 높아지는 국내 노동시장의 특성 상 남성 면접관이 여성 면접관보다 훨씬 많을 수밖에 없을 것임. 이 경우 ‘고성과자’, ‘우수한 합격자’, ‘면접 전문가’ 등을 기반으로 한 데이터는 이미 기존의 차별이 누적되어 오염된 데이터라 할 수 있음.

9) 카피킬러HR 기술설명 블로그, 카피킬러HR에 적용된 AI기술

<http://web.archive.org/web/20210310043522/www.hr.copykiller.com/technology>

10) 한겨레, 남녀평등 가장한 성차별, 채용 성비 5:5의 민낯

<https://www.hani.co.kr/arti/society/women/1003433.html>

11) 사람인, 채용 시 ‘성별’ 영향 여전... 지원자 성별 스펙 차이는?

https://www.saramin.co.kr/zf_user/help/live/view?idx=108718&listType=news

12) 고용노동부 고용률 통계 등

검증되지 않은 기술

○ “특정 기업들은 분노를 감지할 수 있는 알고리즘을 보유하고 있다고 주장합니다. 그러나 이들이 실제로 보유한 것은 찡그린 표정을 감지할 수 있는 알고리즘일 뿐입니다. 분노의 표현일 수도 있고 아닐 수도 있는, 그저 찡그린 표정일 뿐이죠. 얼굴 표정의 구성에 대한 설명을 감정적인 의미에 대한 추론과 혼동하지 말아야 합니다.”¹³⁾

○ 얼굴과 음성 데이터 분석을 통한 ‘감정인식’ 기술 자체에 대한 문제도 제기됨. AI면접 등에 활용되는 일부 인공지능 기술은 단순한 얼굴 식별과 비교를 넘어 사람의 ‘감정’을 확인할 수 있다고 주장하지만 어떠한 신뢰성도 확보하고 있지 못함. 연구에 의하면 얼굴의 움직임과 내적 감정 상태의 명확한 연관성을 찾기 어렵다고 하며, 동시에 표정의 의미 자체도 문화적 맥락에 따라 달라짐¹⁴⁾. 특히, 감정인식 기술의 기반이 된 감정 추론 연구는 다음과 같은 한계를 갖고 있음.

- 제한된 신뢰성 (동일한 감정 범주의 예시들이 공통된 일련의 얼굴 움직임을 통해 안정적으로 표현되거나 인식되지 않음)
- 부족한 특수성 (감정 범주의 예시와 얼굴 움직임의 구성 사이에 고유한 체계성이 없음)
- 제한된 일반화가능성 (맥락과 문화적 영향이 충분히 문서화되거나 설명되지 않음)

○ 국제인권기구 또한 고위험 인공지능 기술의 맥락에서 감정 및 정신 상태 등을 추론하는 생체인식 기술의 문제와 위험성을 지적함.

유엔 인권최고대표 <디지털 시대 프라이버시권 (2021)> (A/HRC/48/31)¹⁵⁾

28. 인공지능 도구는 또한 사람들의 얼굴 표정 및 기타 ‘예측적 생체 인식’에서 감정 및 정신 상태를 추론하여 보안 위협 여부를 판단하도록 개발되었다. 안면감정인식 시스템은 인간의 감정상태를 표정에서 자동적이고 체계적으로 추론할 수 있다는 전제하에 운영되며, 이는 과학적 근거가 부족하다. 연구자들은 얼굴 표정과 감정의 약한 연관성만을 발견했고 얼굴 표정은 문

13) You can't determine emotion from someone's facial movements-and neither can AI
<https://news.northeastern.edu/2021/08/20/you-cant-determine-emotion-from-someones-facial-expression-and-neither-can-ai/>

14) Lisa Feldman Barrett, Ralph Adochs, and Stacy Marsella, 2019, Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements
<https://journals.sagepub.com/doi/10.1177/1529100619832930>

화와 맥락에 따라 다양하여 감정 인식이 편견과 오해에 취약하다는 점을 강조했다. 이러한 우려를 고려하였을 때, 예를 들어 경찰이 검문이나 체포를 위해 개인을 선별하거나 심문 중 진술의 진실성을 평가하기 위해 공공 기관이 감정 인식 시스템을 사용하는 것은 프라이버시권, 자유로울 권리 및 공정한 재판에 대한 권리 등 인권 침해 위험을 초래한다.

○ 한편 대중적으로 잘 알려진 두 가지 감정인식 프로그램을 시험해본 결과 체계적인 인종 차별이 발견된 사례도 있음. 400개의 농구선수 사진 데이터셋을 메그비(Megvii)의 페이스++와 마이크로소프트의 페이스API를 통해 평가해본 결과 프로그램은 흑인 선수들이 분명히 웃고 있거나 미소 짓는 경우에도 평균적으로 보다 부정적인 감정의 점수를 매겼음¹⁶⁾.

○ 이에 얼굴인식 서비스를 제공하던 마이크로소프트는 ‘책임 있는 인공지능’ 기준을 제정하며 감정, 성별, 나이 등 분석 정보를 제공하던 기능을 제공하지 않겠다고 선언함. 한편 미국의 대표적인 AI면접 도구인 ‘하이어뷰(HireVue)’의 경우 연방거래위원회의 감정인식 기술 등에 대한 비판과 부정확성 등 의혹이 불거지자 ‘인지능력’, ‘심리적 특성’, ‘적성’ 등을 평가하던 얼굴 분석 기능을 중단한다고 밝힘¹⁷⁾.

3. 국내 AI 채용도구 활용 현황

민간기업 현황

○ 2022년 7월 고용노동부와 고용정보원의 조사에 의하면 752개 기업 중 AI면접을 실시하는 곳은 52곳으로 6.9%, AI면접 결과를 채용에 반응하는 곳은 31곳인 4.1%였음. 이하 조사 결과 일부.

15) The right to privacy in the digital age, Report of the United Nations High Commissioner for Human Rights.

<https://www.ohchr.org/en/documents/thematic-reports/ahrc4831-right-privacy-digital-age-report-unted-nations-high>

16) Lauren Rhue, 2018, Racial Influence on Automated Perceptions of Emotions

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3281765

17) HireVue, Facing FTC Complaint From EPIC, Halts Use of Facial Recognition

<https://epic.org/hirevue-facing-ftc-complaint-from-epic-halts-use-of-facial-recognition/>

청년 채용 이슈 조사 결과 (고용노동부 등)

(AI 면접①: 활용 현황) 응답 기업의 6.9%가 AI 면접 활용, 결과 반영 비율은 4.1%이며, AI 면접을 실시하는 기업의 대부분이 대면 면접을 추가로 실시

— 기업들은 서류전형, 역량검사, 면접 등에서 AI 기술을 활용하며, 이번 채용이슈 조사는 그 중 AI 면접에 초점

< AI 면접 실시 현황 >

전체		매출액 500대 기업		중견기업	
응답기업	AI 면접 실시	전체	AI 면접 실시	전체	AI 면접 실시
752개	52개 (6.9%)	252개	40개 (15.9%)	500개	12개 (2.4%)

— AI 면접 결과를 채용에 반영하는 정도와 관련하여, 전체 응답기업(752개) 중 4.1%인 31개 기업만이 AI 면접 결과를 채용에 반영한다고 답변했다.

< 채용 시 AI 면접 결과 활용 반영 정도 >

	참고는 하지만, 결과와 무관	계량화하여 점수에 반영	당락을 좌우	AI면접 결과 반영 (계량화+당락 좌우)
전체(n=752)	21개 (2.8%)	25개 (3.3%)	6개 (0.8%)	31개 (4.1%)

— 한편, AI 면접을 실시하는 기업(52개사)의 대다수(50개사, 96.2%)가 AI 면접을 보완하기 위해 대면면접을 추가로 실시한다고 밝혔으며, 이런 결과는 중견기업 채용담당자 간담회에서도 직접 확인할 수 있었다.

(AI 면접②: 평가) AI면접에 대해 채용과정의 공정성 측면에서는 긍정적으로 평가하나, 평가의 정확성에 대해서는 상대적으로 다소 부정적

— 매출액 500대 기업 및 중견기업 모두 AI 면접이 채용과정 공정성과 채용업무의 효율성 측면에서는 긍정적으로 인식하나, 평가의 정확성 측면에서는 상대적으로 다소 낮게 평가하는 것으로 나타났다.

전체	평가의 정확성			채용과정의 공정성			채용업무의 효율성		
	긍정	보통	부정	긍정	보통	부정	긍정	보통	부정
대 기업 (n=252)	13.1%	64.7%	22.2%	46.8%	41.7%	11.5%	47.2%	38.5%	14.3%
중견기업 (n=500)	8.8%	57.2%	34.0%	39.4%	41.8%	18.8%	24.8%	47.8%	27.4%

공공기관 현황

○ 공공기관의 경우, 2018년부터 2022년까지 350곳의 공공기관 중 최소 45곳이 채용 과정에 AI 채용도구를 활용해 왔음.¹⁸⁾ 주로 AI면접 및 역량검사를 도입한 것으로 확인되며 일부는 AI 서류 평가를 활용.

○ AI 채용도구를 활용하는 45곳의 공공기관 중 27곳은 AI 채용도구의 의한 평가를 전형 점수에 반영하거나 채용의 당락을 결정짓는 도구로 활용하고 있음. 이중 5곳의 공공기관은 AI면접과 AI서류평가를 모두 활용하며 특히 강원랜드, 한국수자원공사, 한전 KDN 등은 채용 전형을 1차 AI서류 평가, 2차 AI면접 등으로 진행하여 모든 평가와 선발 과정을 인공지능의 평가 결과에 맡기기도 함.

강원랜드. 2022년 1차 체험형 인턴 모집 공고

□ 채용절차

○ 전형일정



□ 전형절차 안내

○ AI 서류전형 : 모집규모의 3배수 선발(자기소개서 AI평가점수 고득점자 순)

- 블라인드 위반, 불성실 기재자(기재항목 누락, 내용 부적절), 자기소개서 표절을 30% 이상은 평가점수와 상관없이 불합격 처리(붙임 1. 서류전형 유의사항 참조)

* 동점자 전원 합격

○ AI 면접전형 : 모집규모의 1배수 선발(AI평가점수 고득점자 순)

- 딥러닝 AI가 면접영상 분석 후 평가

* 평가 점수 만점의 20% 미만자는 불합격 처리

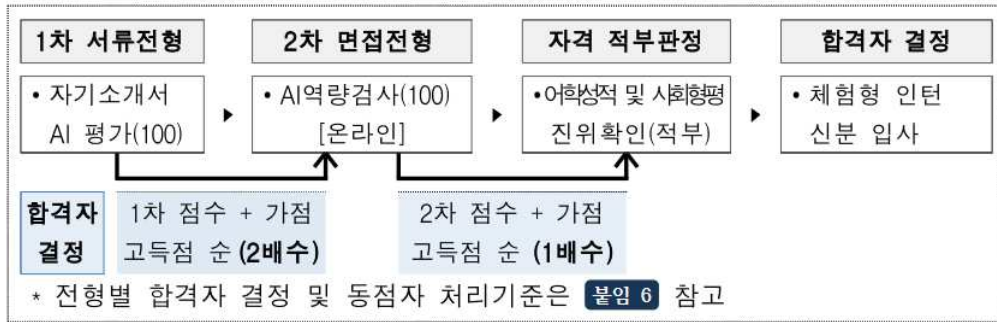
○ 면접전형 동점자 처리기준 - 보훈대상자 > 장애인 > 저소득층, 한부모가족, 북한이탈주민, 다문화가족 > 자택 출·퇴근 가능자(기숙사 미제공)

*해당 기준에도 동점자 발생 시 서류전형 고득점자 순으로 선발

18) 공공기관 채용정보시스템 알리오 등을 통한 자체 조사

한국수자원공사. 2022년 상반기 체험형 인턴사원(일반, 사회형평) 공개 채용

전형절차



2. 1차전형 : AI 서류심사 (합격예정인원의 2배수 내외 선발)

□ (평가방법)

○ K-water 직무기술서 및 합격자 자기소개서의 패턴과 특성을 학습한 AI가 지원자의 자기소개서 및 경험기술서를 분석하여 평가 채점

□ (점수산정) 직무적합도(50점) + 자기소개서 우수성(50점) + 가점

구분	평가대상	평가내용	평가기준
AI 서류심사	직무적합도 평가 (50)	자기소개서, 경험기술서	점수평가 (100)
	자기소개서 우수성 평가 (50)	합격자의 우수 자기소개서 빅데이터를 학습한 AI가 평가	

3. 2차전형 : AI 면접심사 (역량검사) (합격예정인원의 1배수 내외 선발)

□(검사방법) 다양한 게임을 수행하는 과정에서 발생하는 지원자의 무의식적 행동 반응 및 결과를 기반으로 AI가 맞춤형 질문 면접을 통해 직무 수행에 필요한 역량을 검사하여 점수 부여

* 유선인터넷, 크롬 웹 브라우저, 웹캠이 설치된 컴퓨터(노트북)가 필요하므로 사전에 개별 준비하여 검사기간 내 온라인으로 검사 수행하며, 검사 수행 시 접속종료로 인한 책임은 지원자에게 있음(소요시간 : 약 60분)

□(검사용소)성과역량(신뢰, 전략, 관계, 실행), 성장역량(가치, 조직적합), 관찰특성(호감도) 등

□(점수산정) 종합점수(100점) + 가점

1)신뢰 가능성 '신뢰불가'는 점수와 관계없이 불합격 처리

2)검사완료 전 접속 종료자는 1회 추가 접속 기회를 제공하며, 2회 접속 종료자는 불합격처리

3)AI역량검사 대리시험 의심자는 향후 녹화 동영상과 대조하여 본인확인 추가 실시

한전KDN. 2022년도 상반기 한전KDN 체험형 청년 인턴 채용 공고

<p>1차 [서류전형]</p>	<p>○ 자력 적/부평가 - 사무 : 제한없음 - 기술 : 지원분야 관련학과 학위취득자 또는 자격증 보유자 - 공통 : 자기소개서 평가결과 “중대결함” 또는 “표절을 30%” 이상 탈락처리 * 중대결함 : 의미 없는 반복단어, 반복문장, 회사명 오기재, 표절 등 / [참고] 자사명 오기재 예시 : 한국전력, 한전, 한국전력KDN, 한국KDN 등 ○ 합격자 결정기준 : 영어/자격증 점수 합산 고득점자 순 - 동점자 전원 합격처리</p>	<p>채용인원 3배수</p>														
<p>2차 [AI평가]</p>	<p>○ 온라인 시행 ○ 시행기간 : '22.3.29.(화) ~ 4.3(일) 중 개인별 시행 ○ AI 화상면접</p> <table border="1" data-bbox="389 1084 1209 1279"> <tr> <td>AI면접 접수</td> <td>99~80</td> <td>79~60</td> <td>59~40</td> <td>39~20</td> <td>19~1</td> <td>신뢰불 가</td> </tr> <tr> <td>반영점 수</td> <td>20점</td> <td>18점</td> <td>16점</td> <td>14점</td> <td>12점</td> <td>0점</td> </tr> </table> <p>- 신뢰불가 : 미응답, 응답불성실, 응답 왜곡 등 응시 결과를 신뢰할 수 없는 경우 - 탈락기준 : 미참여자, 대리시험자 및 부정행위자 ○ AI자기소개서 평가 ○ 합격자 결정기준 - AI화상면접 점수(20%) + AI자기소개서 평가점수(80%) 고득점자 순</p>	AI면접 접수	99~80	79~60	59~40	39~20	19~1	신뢰불 가	반영점 수	20점	18점	16점	14점	12점	0점	<p>채용인원 1배수</p>
AI면접 접수	99~80	79~60	59~40	39~20	19~1	신뢰불 가										
반영점 수	20점	18점	16점	14점	12점	0점										
<p>최종합격</p>																

○ 한편, 공공기관은 AI채용 도구를 도입해 적극 활용하면서도 공공기관에 법적으로 요구되는 기본적인 자료를 구비하지 않아온 것은 물론이고 공정하고 신뢰할 수 있는 채용 과정을 위한 일련의 검증조차 진행하지 않았음.

○ 2020년, 공공기관에서 사용하는 AI채용 도구에 대해 △공공기관의 공정한 채용절차 준수 유무 △개인정보 침해 여부 검토 △AI면접의 차별성과 편향적 결과 검토 등 실태를 파악하고자 13곳의 공공기관에 정보공개 청구를 진행함. 그러나 대부분의 기관은

‘업체의 비밀(영업비밀)에 관한 사항’ 등의 사유로 정보를 비공개하거나, 단순한 계약자료의 경우에도 ‘AI면접 관련 자료관리 및 운영은 용역사에서 수행하므로 기관은 자료를 갖고 있지 않다’, ‘AI면접 관련 사항은 업체에 일임하고 있으니 업체로 문의하라’며 대부분 정보 부존재 처리하였음.

○ 이후 진행된 정보공개 거부처분 취소소송을 통해 공공기관이 AI채용 도구를 도입해 판단을 자동화하면서도, 그 결과에 대해 합리적인 설명을 보장하기 위한 자료를 전혀 구비해두지 않았음이 확인됨. 예를 들어 AI면접과 AI서류평가를 활용해 채용 평가 과정의 대부분을 인공지능에 일임한 공공기관 한전 KDN의 경우, AI면접 시 주어지는 질문사항에 관해 채용담당자의 어떠한 검토도 없이 그 내용을 용역업체에 포괄적으로 위임함. 또한 채용 의사결정자에게는 AI면접과 관련된 어떠한 자료도 제공하지 않았음. 채용 당사자나 감독기관이 채용 당락에 대한 이유를 묻거나 문제가 제기되어도 설명할 수 없는 상태인 것임.

○ 또한 공공기관 한국국제협력단의 경우 2019년 AI면접 프로그램의 접속 오류로 면접이 중단된 지원자에게 채용시 기회를 주지 않고 불합격 처리하여 감사원에서 주의 조치를 내렸음.¹⁹⁾ 그럼에도 불구하고 해당 기관은 정보공개 청구 과정 및 소송 과정에서 ‘인공지능 프로그램 구동 과정에서 발생하는 돌발상황 및 응시자 민원에 대한 해결 내역을 알 수 있는 문서’, ‘인공지능 프로그램이 채용에 실제 투입될 당시 인공지능의 기능별 오류/오차율’ 등에 대한 정보공개 청구에 대하여 해당 정보가 부존재한다고 답변하였으며 AI면접이 문제없이 진행되었고 그렇기에 그것을 기록한 정보 또한 존재하지 않는다고 답변함.

4. 국내외 대응

○ 인공지능 기반 채용 도구가 보다 빠르게 도입된 해외의 경우, 공공기관의 인공지능 기술 특히 채용과 노동 영역의 인공지능 기술에 대한 규제가 준비되고 있으며 국제적 단위의 단체와 기구에서도 AI 채용과 차별에 대한 우려가 지속되고 있음.

○ 유네스코(UNESCO)와 OECD는 보고서²⁰⁾를 통해 알고리즘을 설계하는 기술 산업

19)"접속오류로 불합격...공공기관 AI면접 무책임하게 운영"
<https://www.hankyung.com/society/article/202207072760Y>

20) The effects of AI on the working lives of women
<https://unesdoc.unesco.org/ark:/48223/pf0000380861>

내의 다양성 부족과 차별적인 노동 시장이, 차별을 내면화시킨 알고리즘을 만들 수 있다고 지적한 바 있으며, 세계경제포럼(WEF)²¹⁾ 또한 2018년부터 지속적으로 고용과 인사 과정에서의 인공지능 사용으로 인한 성차별을 우려해왔음. 이는 특히 성차별 등 많은 국가에서 불균형한 노동시장으로 인한 오염된 데이터를 지적하고 있음.

○ 유럽연합은 인공지능 백서²²⁾를 통해 채용 과정 및 노동자의 권리에 영향을 미치는 상황에서의 인공지능 프로그램 사용은 항상 ‘고위험’으로 간주되어 특정한 조건과 규제가 적용될 것임을 밝혔고, 인공지능 법안 초안²³⁾을 통해 이를 확실시킨 바 있음. AI 면접 등 AI채용 도구와 같은 고위험 인공지능 시스템은 출시 전 위험 평가 및 완화 시스템 구비, 위험과 차별적 결과를 최소화시키기 위한 데이터세트 품질 보장, 결과의 추적성을 보장하기 위한 기록 구비, 감사 및 요구사항 준수 여부 평가를 위한 문서화, 사용자에 대한 명확하고 적절한 정보 제공, 인적 감독 조치 등 엄격한 의무가 부과됨.

○ 미국 일리노이 주는 통지와 정보 제공 등을 의무로 하는 내용의 <인공지능 영상면접법(Artificial Intelligence Video Interview Act)>²⁴⁾을 2019년 제정했고, 뉴욕시는 <자동화된 채용 결정 도구에 관한 법(Automated employment decision tools)>²⁵⁾을 제정하여 2023년 발효 예정임. 뉴욕시의 법안은 뉴욕시의 고용주가 전년도 ‘편향성 감사(bias audit)’를 수행한 AI 채용 도구만을 사용하도록 하고 있으며, 해당 도구를 통해 평가하고자 하는 자격 및 특성을 지원자들에게 통지하고 이를 대체할 수 있는 채용 과정을 요청할 수 있도록 함. 또한 요청 시 AI 채용 도구를 통해 수집된 데이터의 유형, 해당 데이터의 출처 및 고용주와 고용기업의 데이터 보관 정책 등을 확인할 수 있게 함.

○ 한편, 미국 고용평등기회위원회(EEOC)는 채용 과정에서 고용주가 알고리즘과 인공지능을 사용하는 것에 있어 미국 장애인법(ADA)의 적용을 다루는 가이드라인을 발표

21) AI-assisted recruitment is biased. Here's how to make it more fair

<https://www.weforum.org/agenda/2019/05/ai-assisted-recruitment-is-biased-heres-how-to-beat-it/>

22) EUROPEAN COMMISSION, WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust

https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

23) <https://artificialintelligenceact.eu/the-act/>

24) Employers Using AI in Hiring Take Note: Illinois' Artificial Intelligence Video Interview Act Is Now in Effect <https://www.jdsupra.com/legalnews/employers-using-ai-in-hiring-take-note-54767/>

25) Navigating New York City's Artificial Intelligence Hiring Law

<https://www.jdsupra.com/legalnews/navigating-new-york-city-s-artificial-1566015/>

함.²⁶⁾ 가이드라인은 개인의 장애로 인해 인공지능 도구의 선택 및 평가 기준을 충족하지 못하거나 그 기준에 따라 부정적인 평가를 내리는 형태로 장애에 기반해 개인을 탈락시킬 수 있다고 설명함. 이러한 사례를 통해 취업의 기회를 잃게 되면 장애인법 위반이라고 설명. 이 경우에는 경력에 현저한 공백이 있는 지원자 (장애의 결과일 수 있는)를 자동으로 탈락 또는 감정하거나, 장애 유무에 영향을 받을 수 있는 언어 패턴 및 특정 게임을 해결하는 능력과 같이 신체적 또는 정신적 특성을 측정하고 평가하는 채용 과정 등이 포함됨.

○ 국제적인 인권 기준은 인공지능 시스템에 대한 설명가능성과 책임성의 확보, 비차별성 등을 중요시하고 있음.

“신기술의 개발 및 도입에 관한 의사결정에 모든 관련 이해당사자의 참여를 보장하고, 특히 공공부문에서 인공지능이 지원하는 의사결정에 대하여 적절한 설명가능성이 보장될 필요가 있다.” 유엔 사무총장 (2020)²⁷⁾

“동시에, 인력 관리에 사용되는 많은 인공지능 시스템의 기반이 되는 정량적 사회과학은 견고하지 않고 편견에 취약하다. 예를 들어, 회사가 남성, 백인, 중년 남성을 선호하는 과거 데이터셋으로 학습된 인공지능 채용 알고리즘을 사용하는 경우, 알고리즘은 구인에 적합한 자격을 동등하게 갖춘 여성, 유색인종 및 젊은이나 노년층을 선호하지 않을 것이다. 동시에, 근로자를 보호하기 위한 책무성 구조와 투명성이 결여된 경우가 많으며, 노동자들은 인공지능 기반 모니터링 실시에 대한 설명을 거의 또는 전혀 듣지 못하고 있다.” 유엔인권최고대표 (2021)²⁸⁾

26) U.S.EEOC, The Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence to Assess Job Applicants and Employees
<https://www.eeoc.gov/laws/guidance/americans-disabilities-act-and-use-software-algorithms-and-artificial-intelligence>

27) Question of the realization of economic, social and cultural rights in all countries: the role of new technologies for the realization of economic, social and cultural rights
https://www.ohchr.org/EN/HRBodies/HRC/RegularSessions/Session43/Documents/A_HRC_43_29.pdf

28) A/HRC/48/31: The right to privacy in the digital age - Report of the United Nations High Commissioner for Human Rights
<https://www.ohchr.org/en/documents/thematic-reports/ahrc4831-right-privacy-digital-age-report-united-nations-high>

○ 한국 국가인권위원회 또한 2022년 5월 <인공지능 개발과 활용에 관한 인권 가이드라인>²⁹⁾을 마련하여 공공기관과 민간 기업의 신기술의 도입과 사용에 대한 문제를 지적하고 인권침해 및 차별이 발생하지 않도록 권고한 바 있음. □

19. 이러한 알 권리의 보장과 인공지능이 미치는 영향력과 중요성을 감안할 때, 인공지능의 판단과정과 그 결과에 대한 적절하고 합리적인 설명이 보장 되어야 합니다. 학습 및 추론, 판단의 과정과 결과에 이른 이유를 설명 하기 어려운 인공지능은 이에 대한 대응의 불확실성과 영향을 받는 당사자의 불안감을 유발하고, 인권 및 안전에 관한 법령과 정책의 집행 효과를 불분명하게 할 수 있습니다.

21. 공공기관은 인공지능의 개발과 활용 계획 등을 사전 공개하여야 하고, 관련 당사자들의 의견을 공청회 등으로 수렴하여야 합니다. 공공기관이 인공지능을 통한 의사결정을 할 때 설명할 수 없는 인공지능을 활용 해서는 안 되며, 특히 조달의 경우 입찰 단계에서부터 설명가능성이 보장되어야 합니다.

22. 공공기관이 개발하고 활용하는 모든 인공지능과 민간이 개발하고 활용 하는 인공지능 중 개인의 생명이나 안전 등 기본적 인권에 중대한 영향을 미치는 인공지능은 사용된 데이터와 인공지능 알고리즘의 주요 요소 일반에게 공개하고 설명하여야 합니다.

23. 또한, 인공지능에 의한 자동화된 의사결정이 예정되어 있는 경우, 영향을 받는 당사자들은 사전에 그 사실을 알아야 합니다. 자동화된 의사결정에 의하여 영향을 받는 당사자는 그 결정의 이유에 대하여 설명을 듣고, 당사자 진술을 할 수 있으며, 이의를 제기할 수 있어야 합니다.

24. 특히, 완전히 자동화된 의사결정으로만 개인에게 법적 효력 또는 생명·신체·정신·재산에 중대한 영향을 미치는 일은 제한되어야 하고, 이러한 의사결정이 이루어진 경우에는 당사자가 해당 방식을 거부하거나 인적 개입을 요구할 수 있는 권리를 보장받아야 합니다.

29) 국가인권위원회 보도자료, <인공지능 개발과 활용에 관한 인권 가이드라인> 마련

<https://www.humanrights.go.kr/site/program/board/basicboard/view?boardtypeid=24&boardid=7607961&menuid=001004002001>

공공기관 인공지능 정보공개거부처분취소소송 진행 경과와 제도적 과제



김하나 변호사 | 민주사회를위한변호사모임 디지털정보위원회 위원장

목 차

1. 들어가며
 - 가. 채용과정에서 AI시스템 활용의 위험성
 - 나. 이 사건 소 제기의 필요성
2. 소송 경과 및 판결 요지
 - 가. 소 제기 배경
 - 1) 공공기관 비공개 결정
 - 2) 정보공개청구 문항의 구성
 - 나. 주요 소송 진행 경과
 - 1) 피고 기관 대응방법
 - 2) 정보 부존재 주장 관련
 - 3) 비공개 사유에 관한 주장

다. 판결 요지

- 1) 피고 국제협력단 사건(수원지방법원 2022. 5. 26. 선고 2020구합73847)
- 2) 피고 한전KDN 사건(광주지방법원 2022. 6. 16. 선고 2020구단11377)
- 3) 상소 여부

3. AI 면접 운용 과정의 실태와 문제점

가. 정보가 존재하지 않는다는 피고 주장의 의미

- 1) 공공기관은 채용절차에서 AI면접 프로그램을 사용하였음에도 이와 관련한 목록을 별도로 작성하지 않았다(1.가. 관련)
- 2) 공공기관은 개인정보보호법상 위탁자의 지위에서 지원자들의 개인정보처리를 제3자에게 위탁하였음에도 개인정보보호법에서 정한 의무를 다하지 않았다(2.나. / 2.다. 관련)
- 3) 공공기관은 채용절차에서 발생한 시스템 오류에 대한 민원 해결, 사후 점검 및 조치 업무를 전혀 수행하지 않았다(1.마. / 3.다. / 4.사. / 4.아. 관련)
- 4) 공공기관은 국가를 당사자로 하는 계약에 관한 법령과 관련 계약예규의 내용을 위반하여 수의계약을 체결하였다(3.가. 관련)
- 5) 공공기관은 AI면접 프로그램의 내용과 기술정보를 구체적으로 확인하지 않고 수의계약을 체결하였다(3.라. / 4.가. / 4.나. / 4.다. 관련)
- 6) 피고는 관련 규정에 반하여 면접위원에게 AI면접 프로그램에 대한 사전교육을 실시하지 않고 결과지를 제공하였다(1.라. 관련)

나. AI면접 프로그램의 처리방법을 확인할 수 있는 자료에 대한 공개 요구의 필요성

- 1) 알고리즘 등 자료 공개에 대한 해외 판례
- 2) 국내법상 알고리즘 공개 가능성

다. 소결

4. 나가며 : 개선안에 대한 제언

1. 들어가며

가. 채용과정에서 AI시스템 활용의 위험성

미국과 유럽연합은 한국보다 먼저 인공지능(이하 'AI')시스템을 민간과 공공영역에서 보편적으로 활용하였고, AI시스템의 문제점과 본질적 한계에 대한 제도적 보완책에 대한 논의도 2010년대 중반부터 지속적으로 이루어졌다. 유럽연합(이하 'EU')은 AI시스템을 규제할 수 있는 법안을 준비하였고, 2021. 4. 21. EU 집행위원회는 AI시스템 규제를 위한 『인공지능법안(ARTIFICIAL INTELLIGENCE ACT)』 초안을 발표하였다. 법안의 골자는 AI시스템에 따른 위험을 ▲수용 불가 위험, ▲고위험, ▲저위험/최소위험 유형으로 구분하여 각각의 위험 유형에 따라 AI 시스템을 분류하여 AI시스템 유형에 따라 규제를 달리하는 내용이다.

특히, 고위험 AI시스템에 관한 내용이 전체 85개 조항 중 절반을 차지하는데, 안전과 관련된 AI시스템과 기본권과 관련된 AI시스템을 고위험 AI시스템 유형으로 정의하면서 고위험 AI시스템의 경우 공급자와 이용자 등에게 출시 전후 엄격한 의무를 부과하고 있다. EU 『인공지능법안』은 고위험 AI시스템의 예시를 들고 있는데, 특히 '모집 광고, 지원서 선별, 면접, 심사 등 채용과정에서 이용하는 AI시스템'은 기본권 관련 AI시스템으로 고위험 AI시스템으로 분류하고 있다(이후 위 법안 초안에 대한 절충안 등이 제시되었지만, AI시스템을 위험성을 기준으로 분류하고 고위험 AI시스템에 엄격한 의무를 부과하는 체계에 대한 변경은 없다).

앞서 제시한 인공지능법안의 관점에서 우리나라 공공기관이 활용하고 있는 AI면접 프로그램은 '고위험 AI시스템'으로 분류되며, '고위험 AI시스템'을 '공공기관'이 '조달'의 방식을 통해 '채용'에 활용하였다는 점에서 기술적·사회적·행정적으로 살펴볼 지점이 많다. 공공기관이 활용한 AI면접 프로그램은 마이다스인(또는 마이다스아이티)이 개발한 것으로 응시자의 표정, 태도를 분석하여 이를 호감도 평가에 활용한 것이다. 채용의 공정성 측면에서 해당 기술이 지원자의 외모나 사투리를 차별하거나, 여성 또는 장애인을 차별하여 불공정하게 평가할 위험성은 없는지 등 AI기술이나 시스템 자체가 공정성을 확보한 것인지도 확인할 필요가 있다.

또한, AI시스템은 학습에 사용된 데이터의 양과 질에 따라 결과값이 다르게 생성될 수 있으므로 AI시스템 학습에 활용된 데이터가 어떤 데이터인지, 그러한 데이터 학습을 통해 어떠한 정보를 어떻게 처리하여 결과값을 도출하는지 등을 확인하여야 지원자가

해당 전형에서 탈락하더라도 어떤 평가 기준에 따라 탈락한 것인지에 대한 설명이 가능하다. 지원자가 채용절차에서 탈락하였는데, 왜 탈락하였는지에 대한 답변이 'AI시스템이 결정하였기 때문'이라면, 그 자체가 지원자의 직업선택의 자유를 침해와 알 권리를 침해하는 행위에 해당한다.

나. 이 사건 소 제기의 필요성

AI기술은 개인 또는 공동체의 권리 또는 안전을 침해할 위험성을 지니고 있다. 하지만 안타깝게도 현재 대한민국에는 이를 규제할 명확한 법적 안전장치가 없고, 이러한 상황에서 인간의 기본권에 직접적인 영향을 미치는 분야까지 AI시스템이 어떠한 검증 절차나 사후관리 없이 도입·활용되는 실정이다. 공공기관이 AI기술을 도입하고 활용한 계약과정에 대한 정보와 AI면접 프로그램의 학습데이터 등 프로그램 관련 정보는 국민의 알 권리 범주에 포함되는 정보로서 이에 대한 공개를 요청하는 것은 정당한 요청이다.

그런데 정보공개소송의 피고가 된 2개 기관은 '채용'이라는 기본권 침해 위험이 큰 영역에서 AI시스템을 활용하였음에도 AI시스템 활용실태와 내용을 분석할 수 있는 정보를 공개하지 않았다. 이번 정보공개거부처분에 대한 취소소송은, ① 대중에 공개되지 않았던 정보를 취득하여 공공기관이 AI시스템을 어떻게 도입·활용하고 있는지를 확인할 수 있다는 측면과 ② 공공기관 AI면접 프로그램 활용행위에 대해서도 국민의 알 권리가 보장된다는 점을 확인을 필요가 있다는 측면에서 소 제기 필요성이 컸다.

이하 AI면접 프로그램 관련 정보공개에 대한 거부처분취소소송의 경과와 판결 내용을 중심으로 현행 AI면접 프로그램 도입·운영 실태에 관한 제도적 문제점을 제시하고, 간략히 개선안을 제안하고자 한다.

2. 소송 경과 및 판결 요지

가. 소 제기 배경

1) 공공기관 비공개 결정

원고는 13개 공공기관을 상대로 정보공개를 청구하였는데, 국제협력단과 한전KDN은 대부분 정보에 대한 공개를 거부하는 취지의 회신을 하였다.

원고는 2020. 7. 28. 국제협력단을 상대로 ‘인공지능 프로그램과 채용의 공정성 관련 정보공개청구’라는 제목으로 관련 정보의 공개를 청구하였고, 이에 대하여 국제협력단은 2020. 8. 25. 내규인 ‘채용업무처리지침’을 공개하였으나, 나머지 정보에 대해서는 정보공개법 제9조 제1항 제5조의 ‘공개될 경우 업무의 공정한 수행에 현저한 지장을 초래한다고 인정할 만한 상당한 이유가 있는 정보’라는 사유를 들어 전부 비공개 처리하였다.

원고는 2020. 9. 8. 국제협력단이 정보 대부분을 비공개한 처분에 대하여 이의를 신청하였고, 이에 대하여 국제협력단은 2020. 9. 25. “채용계획(안), 시험 실시계획, 면접 채점표, 다른 응시자의 성적, 석차, 다른 응시자의 답안지 및 채점표 등과 관련된 정보는 공공기관의 정보공개에 관한 법률 제9조 제1항 제5호에 의거하여 비공개처리 됨을 양해 부탁드립니다.”라고 사유를 밝히며 AI면접 관련 공고문, 최근 5년간 신입직원 통계를 제외한 나머지 정보를 비공개 결정하였다.

한편, 원고는 2020. 7. 28. 한전KDN을 상대로 위 국제협력단과 동일하게 ‘인공지능 프로그램과 채용의 공정성 관련 정보공개청구’라는 제목으로 정보공개를 청구하였다. 이에 대하여 한전KDN은 2020. 8. 10.경 채용현황에 대해서 알리오 공시자료를 참고하도록 하였고, 2020년 상반기 신입사원 채용공고문을 공개하였다. 그리고 나머지 정보에 대하여는 AI면접 관련 자료 관리 및 운영은 용역사인 마이다스인에서 수행하여 공개 요청한 자료를 보유하고 있지 않다는 취지인 ‘부존재’를 이유로 공개를 거부하는 결정을 하였다.

2) 정보공개청구 문항의 구성

정보공개포털(www.open.co.kr)을 통해 공공기관에 정보공개를 요청할 때, 가장 중요한 것은 기관과 열람하고자 하는 정보를 특정하는 것이다. 최초 기관을 상대로 정보공개청구를 할 때, 시민사회단체는 공공기관이 AI면접 프로그램을 사용할 때 체결하는 계약의 형태가 무엇인지, AI면접 프로그램을 구동하는 과정에서 제3기관이 지원자의 개인 정보를 처리할 때 그 지위가 ‘제3자¹⁾’와 ‘수탁자²⁾’ 중 어디에 해당하는지 등 법률관계를 추정할 수 있는 정확한 정보를 확인하기 어려웠다. 이에 최초 정보공개를 할 당시 아래와 같이 경우의 수를 나누어 공개를 요청하는 정보와 자료를 특정하여 13개 공공기관에 정보공개를 청구하였다.

1) 개인정보보호법 제17조 등 적용

2) 개인정보보호법 제26조 등 적용

1. 채용절차에 사용된 인공지능 시스템 또는 소프트웨어(이하 ‘인공지능 프로그램’) 관련 정보공개 일반
2. 채용절차에 사용된 인공지능 프로그램 사용 과정에서 위탁업체가 관여한 경우
3. 채용절차에 사용된 인공지능 프로그램 사용 과정에서 용역업체가 관여한 경우
4. 채용절차에 사용된 인공지능 프로그램 사용 과정에서 수집한 개인정보를 목적 외 이용 또는 제3자에게 제공하였을 경우로 경우
5. 기관, 위탁업체 등 제3자가 채용절차에서 사용한 인공지능 프로그램에 대한 사항

그런데, 다수의 공공기관에 정보공개를 청구한 결과 채용절차에 사용된 AI면접 프로그램은 대부분 『국가를 당사자로 하는 계약에 관한 법률(이하 ‘국가계약법’)』상 ‘용역 계약(제한경쟁입찰에 의한 계약, 협상에 의한 계약)’에 의하여 도입되었다는 점과 업체가 개인정보보호법상 지원자의 개인정보를 위탁처리하는 ‘수탁자’의 지위라는 점을 확인하였다. 이에 시민사회단체(이하 ‘원고’)는 타공공기관 회신자료를 토대로 확인한 몇 가지 사실관계를 반영하여 공개가 필요한 정보를 축소하여 정보공개거부처분취소소송을 제기하였다. 소장과 함께 제출한 별지 정보공개 대상정보 목록은 아래와 같다.

별지

대상 정보

1. 채용절차에 사용된 인공지능 시스템 또는 소프트웨어(이하 ‘인공지능 프로그램’) 관련 정보

- 가. 공기업, 준정부기관의 경영에 관한 지침 제16조에 따라 인공지능 프로그램과 관련하여 인사부서 및 감사부서에 보관 중인 문서목록
- 나. 인공지능 프로그램이 사용되었던 당해 채용의, 채용계획 수립, 공고, 서류·면접 전형과 필기시험 실시, 합격자 결정 등 직원채용의 절차와 방법 등에 관한 별도의 문서
- 다. 인공지능 프로그램이 채용절차에 사용된 구체적인 목적, 평정 요소(항목), 인공지능 프로그램의 결과가 반영되는 구체적인 사항과 절차를 알 수 있는 문서
- 라. 인공지능 프로그램이 사용되었던 당해 채용의 의사결정자(기관 내 채용담당자 및 면접위원 등)에게 제공된 문서 및 사전교육 내용
- 마. 인공지능 프로그램이 시행된 이후 이에 대한 사후 점검 사항을 확인할 수 있는

는 문서

- 바. 공고문 이외에 지원자에게 공개한 인공지능 채용기준과 결과를 확인할 수 있는 문서
- 사. 각 채용 관련 주체들이 인공지능 면접 지원자 개인정보를 보관한 장소, 보관 기간, 파기 여부를 확인할 수 있는 문서(만약 파기하지 않았다면 관련 근거 규정)
- 아. 인공지능 프로그램을 통해 지원자별로 평가한 결과가 있다면 그 구체적인 평가 항목, 건수를 확인할 수 있는 문서(개인정보 삭제 후 제공 요청)
- 자. 인공지능 프로그램이 사용되었던 당해 채용과정에서 기관이 수집·이용한 지원자의 개인정보 항목 및 항목별 목적
- 차. 인공지능 프로그램이 사용되었던 당해 채용과정에서 수집된 지원자의 개인정보를 위탁업체에 제공한 경우, 위탁업체의 명칭, 제공된 지원자의 개인정보 항목 및 항목별 목적

2. 채용절차에 사용된 인공지능 프로그램 사용 과정에서 위탁업체가 관여한 경우

- 가. 인공지능 프로그램 사용에 관하여 위탁계약서 등 기관과 위탁업체가 체결한 위탁계약의 내용을 알 수 있는 문서
- 나. 인공지능 프로그램 사용에 관하여 기관이 수탁자에 대한 교육을 실시하였는지, 만약 실시하였다면 그 내용 및 실시내역을 확인할 수 있는 문서
- 다. 인공지능 프로그램 사용에 관하여 수탁자에 대한 지원자 개인정보의 안전한 처리에 관한 계획을 수립한 내용과 그 구체적인 사항을 확인할 수 있는 문서
- 라. 인공지능 프로그램 사용에 관하여 사업장 및 인터넷 홈페이지 등에 게재·공개한 위탁업무 내용 및 수탁자 관련 사항을 확인할 수 있는 문서
- 마. 인공지능 프로그램 사용에 관하여 처리 현황 점검 등 대통령령으로 정하는 바에 따라 수탁자가 개인정보를 안전하게 처리하는지를 감독하였는지 만약 하였다면 관련 문서
- 바. 인공지능 프로그램 사용에 관하여 보안 각서 등 대행·위탁과정에서 작성한 일체의 문서, 개인정보 유출방지 방안 관련 문서, 이행능력 심사 관련 문서

3. 채용절차에 사용된 인공지능 프로그램 사용 과정에서 용역업체가 관여한 경우

- 가. 인공지능 프로그램을 사용한 기관(발주기관)이 입찰 과정에서 홈페이지에 공시한 문서(예 : 입찰공고서, 제안요청서 등)
- 나. 인공지능 프로그램 사용을 위하여 용역업체가 작성하여 제출한 서류(예 : 용역계약일반조건, 용역계약특수조건, 과업내용서, 보안각서, 개인정보처리위탁 관련 문서, 청렴계약 관련 문서, 인권경영이행 관련 문서 등 용역 관련 서류)
- 다. 인공지능 프로그램 관련 용역계약 이행의 관리·감독을 실시하거나 사업의 품질 확보를 위해 조치한 내역을 확인할 수 있는 문서
- 라. 인공지능 프로그램 관련 계약목적물의 다음 기술자료
 - 소스코드 및 오브젝트 코드의 복제물
 - 기술정보(매뉴얼, 설계서, 사양서, 기능명세서, 플로우차트, 유지보수자료 등)

4. 기관, 위탁업체 등 제3자가 채용절차에서 사용한 인공지능 프로그램에 대한 사항

- 가. 인공지능 프로그램의 기능을 설명한 자료(위탁업체 등 제3자가 기관에 제공한 자료 등)
- 나. 진행자, 면접위원 및 지원자에게 배포되거나 제공된 인공지능 프로그램에 대한 교육 및 기술 설명자료, 물품의 항목(예 : AI면접시스템 환경 조성 사항 등)
- 다. 인공지능 프로그램이 직접 또는 기관이 프로그램을 사용하여 지원자에게 출제된 문항을 알 수 있는 문서
- 라. 인공지능 프로그램이 지원자로부터 수집하는 사항(예 : 표정, 감정, 안구 움직임, 근육 변화 등 얼굴 정보, 목소리 톤, 크기, 속도, 음색 등 음성정보, IP주소, 쿠키, 단말 기고유번호 등)을 알 수 있는 문서
- 마. 지원자 평가를 위해 인공지능 프로그램에 입력하는 지원자의 개인정보(예 : 얼굴, 성별, 출신학교, 장애유무 등)를 알 수 있는 문서
- 바. 위 라.항을 통해 평가하려는 직무 적합성이 무엇인지를 알 수 있는 문서(예 : 대리시험 및 지원자 교체 여부 판단, 허위응답에 대한 응답신뢰 불가여부 판단, 종합/세부 평가결과 및 지원한 직군에 대한 적합도 점수 제공, 지원자의 내면적 특성 및 호감도)
- 사. 인공지능 프로그램 구동 과정에서 발생하는 돌발상황 및 응시자 민원에 대한 해결 내역을 알 수 있는 문서
- 아. 인공지능 프로그램이 채용에 실제 투입될 당시 인공지능의 기능별 오류/오차율

원고가 공개를 청구한 정보는 크게 4가지로 분류된다. 구체적으로, ① ‘채용절차에 사용된 인공지능 프로그램 관련 정보’ 항목에서는 인공지능 프로그램이 채용절차에 도입된 과정, 실제 AI면접 결과가 채용절차에서 어떻게 활용되는지 여부와 인공지능 프로그램의 일반적인 내용을 확인할 수 있는 자료, ② ‘채용절차에 사용된 인공지능 프로그램 사용 과정에서 위탁업체가 관여한 경우’ 항목은 개인정보처리 위탁계약을 체결하여 지원자의 개인정보를 관리하게 한 피고와 수탁자가 개인정보보호법상 의무를 잘 이행하였는지 확인할 수 있는 자료, ③ ‘채용절차에 사용된 인공지능 프로그램 사용 과정에서 용역업체가 관여한 경우’ 항목은 인공지능 프로그램을 사용할 때 인공지능 프로그램을 만든 업체와 피고가 체결한 용역계약의 구체적인 내용을 확인할 수 있는 자료, ④ ‘기관, 위탁업체 등 제3자가 채용절차에서 사용한 인공지능 프로그램에 대한 사항’ 항목에서는 인공지능 프로그램을 구성하는 알고리즘이 공정한지, 공정성을 확보하기 위해 기관이 적절한 관리·감독을 하는지를 확인할 수 있는 정보이다.

나. 주요 소송 진행 경과

1) 피고 기관 대응방법

피고 국제협력단은 원고의 정보공개청구에 대하여 정보부존재와 공공기관의 정보공개에 관한 법률(이하 ‘정보공개법’) 제9조 제1항 제5조에 따라 ‘업무에 공정한 수행에 현저한 지장을 초래한다고 인정할만한 상당한 이유가 있는 정보’임을 이유로 대부분 정보를 비공개하였다. 그런데 피고 국제협력단은 정보공개거부처분취소소송이 제기된 이후 상당한 자료를 법원에 서증으로 제출하는 방식으로 공개하였고, ‘3. 나. 인공지능 프로그램 사용을 위하여 용역업체가 작성하여 제출한 서류 일체’ 정보에 대해서만 정보공개법 제9조 제1항 제5호 사유에 해당하고 나머지 정보는 정보가 존재하지 않는다는 비공개 사유를 유지하였다.

피고 한전KDN은 최초 원고의 정보공개청구에 ‘AI면접 관련 자료 관리 및 운영은 용역사인 마이다스인에서 수행하므로 기관이 정보를 보유하고 있지 않다(정보부존재)’는 이유로 공개를 거부하였다. 그런데 소제기 이후 피고 한전KDN은 일부 정보에 대해서는 정보부존재 주장을 유지하였으나, 일부 정보는 정보공개법 제9조 제1항 제5조의 ‘업무에 공정한 수행에 현저한 지장을 초래한다고 인정할만한 상당한 이유가 있는 정보’와 제9조 제1항 제7호의 ‘경영상·영업상 비밀에 관한 사항으로서 공개될 경우 법인 등의 정당한 이익을 현저히 해칠 우려가 있다고 인정되는 정보’에 해당한다고 주장하면서 비공개 사유를 변경하였다. 또한, 피고는 일부 정보는 공개 문항만으로 정보가 특정되지

않는다는 주장도 추가하였다.

2) 정보 부존재 주장 관련

정보공개거부처분취소소송의 상대방인 피고 기관들은 원고가 청구한 대부분 정보가 존재하지 않는다고 답변하였다. 법원은 일관되게 정보공개 제도가 공공기관이 보유·관리하는 정보를 그 상태대로 공개하는 제도라는 점에 비추어 보면, 정보공개를 구하는 사람이 공개를 구하는 정보를 행정기관이 보유·관리하고 있을 상당한 개연성이 있다는 점을 증명함으로써 충분하지만, 공공기관이 그 정보를 보유·관리하고 있지 않은 경우에는 특별한 사정이 없는 한 정보공개거부처분의 취소를 구할 법률상의 이익이 없다는 입장을 취하고 있다³⁾.

원고는 피고가 보유·관리하고 있지 않은 자료라고 주장하는 문서가 법령 등에서 공공기관이 일반적으로 구비·관리하도록 정한 문서라는 점을 주장하였다. 하지만, 부과된 의무와 별개로 각 기관이 객관적으로 해당 문서가 존재하지 않는다는 점을 주되게 주장하였다. 나아가 AI면접과 관련하여 법령에서 공공기관에 특별한 의무를 부과한 사정은 존재하지 않았는바, 법령 등에서 공공기관이 일반적으로 구비·관리하였다는 사실만으로 상당한 개연성이 입증된 것으로 보기에 한계가 존재하였다. 따라서 공개를 청구하는 원고가 해당 정보가 존재하여야 함을 입증해야 한다는 입증책임의 법리에 따라, 피고가 보유·관리하지 않고 있다고 주장하는 정보에 대해 각하 판결을 면하기 어려운 부분이 존재하였다. 피고 기관들이 존재하지 않는다고 주장한 정보 목록은 아래와 같다.

	국제협력단	한전KDN
1. 채용절차에 사용된 인공지능 프로그램 관련 정보	가. 공기업, 준정부기관의 경영에 관한 지침 제16조에 따라 인공지능 프로그램과 관련하여 인사부서 및 감사부서에 보관 중인 문서목록	가. 공기업, 준정부기관의 경영에 관한 지침 제16조에 따라 인공지능 프로그램과 관련하여 인사부서 및 감사부서에 보관 중인 문서목록
	라. 인공지능 프로그램이 사용되었던 당해 채용의 의사결정자에게 제공된 문서 및 사전교육 내용	라. 인공지능 프로그램이 사용되었던 당해 채용의 의사결정자에게 제공된 문서 및 사전교육 내용

3) 대법원 2006. 1. 13. 선고 2003두9459 판결 등 다수 판례 참조

	<p>마. 인공지능 프로그램이 시행된 이후 이에 대한 사후 점검 사항을 확인할 수 있는 문서</p> <p>바. 공고문 이외에 지원자에게 공개한 인공지능 채용기준과 결과를 확인할 수 있는 문서</p>	<p>항변내용 : 불특정</p>
<p>2. 채용절차에 사용된 인공지능 프로그램 사용 과정에서 위탁업체가 관여한 경우</p>	<p>소제기 이후 법원에 제출한 자료로 내용이 일부 확인되는 문항</p>	<p>나. 인공지능 프로그램 사용에 관하여 기관이 수탁자에 대한 교육을 실시하였는지, 만약 실시하였다면 그 내용 및 실시내역을 확인할 수 있는 문서</p> <p>다. 인공지능 프로그램 사용에 관하여 수탁자에 대한 지원자 개인정보의 안전한 처리에 관한 계획을 수립한 내용과 그 구체적인 사항을 확인할 수 있는 문서</p>
<p>3. 채용절차에 사용된 인공지능 프로그램 사용 과정에서 용역업체가 관여한 경우</p>	<p>가. 인공지능 프로그램을 사용한 기관이 입찰 과정에서 홈페이지에 공시한 문서</p> <p>항변내용 : 공개될 경우 업무의 공정한 수행이나 연구·개발에 현저한 지장을 초래한다고 인정할 만한 상당한 이유가 있는 정보(정보공개법 제9조 제1항 제5호)</p> <p>다. 인공지능 프로그램 관련 용역계약 이행의 관리·감독을 실시하거나 사업의 품질 확보를 위해 조치한 내역을 확인할 수 있는 문서</p> <p>라. 인공지능 프로그램 관련 계약목적물의 다음 기술자료</p> <ul style="list-style-type: none"> - 소스코드 및 오브젝트 코드의 복제물 - 기술정보(매뉴얼, 설계서, 사양서, 	<p>원고 취하</p> <p>나. 인공지능 프로그램 사용을 위하여 용역업체가 작성하여 제출한 서류</p> <p>원고 취하</p>

	기능명세서, 플로우차트, 유지보수 자료 등)	
4. 기관, 위탁업체 등 제3자가 채용절차에서 사용한 인공지능 프로그램에 대한 사항	소제기 이후 법원에 제출한 자료로 내용이 일부 확인되는 문항	가. 인공지능 프로그램의 기능을 설명한 자료(위탁업체 등 제3자가 기관에 제공한 자료 등) 나. 진행자, 면접위원 및 지원자에게 배포되거나 제공된 인공지능 프로그램에 대한 교육 및 기술 설명 자료, 물품의 항목
	다. 인공지능 프로그램이 직접 또는 기관이 프로그램을 사용하여 지원자에게 출제한 문항을 알 수 있는 문서	다. 인공지능 프로그램이 직접 또는 기관이 프로그램을 사용하여 지원자에게 출제한 문항을 알 수 있는 문서
	소제기 이후 법원에 제출한 자료로 내용이 일부 확인되는 문항	라. 인공지능 프로그램이 지원자로부터 수집하는 사항을 알 수 있는 문서 마. 지원자 평가를 위해 인공지능 프로그램에 입력하는 지원자의 개인정보를 알 수 있는 문서 바. 위 라.항을 통해 평가하려는 직무 적합성이 무엇인지를 알 수 있는 문서
	사. 인공지능 프로그램 구동 과정에서 발생하는 돌발상황 및 응시자 민원에 대한 해결 내역을 알 수 있는 문서	원고 취하
	아. 인공지능 프로그램이 채용에 실제 투입될 당시 인공지능의 기능별 오류/오차율	

3) 비공개 사유에 관한 주장

피고 기관들은 공통적으로 피고 기관들이 보유·관리하는 정보가 존재함을 전제로 해당 정보가 존재하더라도 정보공개법 제9조 제1항 제5호의 ‘업무에 공정한 수행에 현저한 지장을 초래한다고 인정할만한 상당한 이유가 있는 정보’임을 주장하였다. 그런데 이때 ‘의사결정 과정 또는 내부검토 과정’에 있는 사항으로서 공개될 경우 업무의 공정한 수행에 현저한 지장을 초래한다고 인정할만한 상당한 이유를 비공개대상으로 정한 취지는 ‘공개로 인하여 공공기관의 의사결정이 왜곡되거나 외부의 부당한 영향과 압력을 받을 가능성을 차단하여 독립적이고 공정한 의사결정이 이루어지도록 하기 위한 것이다⁴⁾.

정보공개법 제9조 제1항 제5호 사유에 해당하는지는 비공개함으로써 보호되는 업무수행의 공정성 등 이익과 공개로 보호되는 국민의 알 권리 보장과 국정에 대한 국민의 참여 및 국정운영의 투명성 확보 등 이익을 비교·교량하여 구체적인 사안에 따라 신중하게 판단한다고 일관되게 판시하고 있다⁵⁾. 그런데 피고 기관들이 정보의 비공개를 주장하는 근거는 해당 정보가 정보공개법에서 정한 ‘계약’에 관한 자료이고, 피고 내부 규정상 비공개 대상정보 세부기준에서 정하고 있는 비공개 대상정보에 해당한다는 사실에 국한되며, 피고는 구체적으로 어떤 점에서 공개될 경우 업무의 공정한 수행에 ‘현저한’ 지장을 초래한다고 인정할만한 상당한 이유가 있는지 구체적으로 설명하지 못하였다.

특히, 원고가 공개를 청구한 자료는 사회적으로 큰 논란을 불러일으켰던 AI면접에 대한 것이며, 채용과정에서 AI면접이 사기업에까지 확산하면서 대다수 취업준비생이 별도의 교육비용을 들여 AI면접 전형을 준비하는 추세가 강화되고 있다. 이런 사회적인 현상을 고려하였을 때 대부분 공공기관에서 사용하고 있는 AI면접 프로그램이 어떤 계약 방식과 내용으로 공공기관에 사용된 것인지를 확인할 필요성이 매우 크다.

나아가 원고가 공개를 청구한 정보는 2019년도 계약 과정에서 개발업체가 작성하여 제출한 서류를 의미하는 것으로서 이미 기간이 만료된 계약에 대한 것으로 해당 자료의 공개로 인하여 피고 기관들의 의사결정이 왜곡되거나 외부의 부당한 영향과 압력을 받을 가능성이 없고, 독립적이고 공정한 의사결정을 방해한다고 볼 이유도 없었다. 원고는 위와 같은 내용과 함께 AI면접 응시자의 외모, 사투리 등에 기초한 차별 가능성 등 인권침해의 문제를 지적하며 공개의 정당성을 주장하였다.

4) 대법원 2003. 12. 11. 선고 2001두8827 등 다수 판결

5) 대법원 2010. 6. 10. 선고 2010두2913 등 다수 판결

다. 판결 요지

법원은 피고 기관들이 부존재를 주장하는 정보에 대해서는 원고 청구를 각하하였다. 그러나 피고가 보유·관리한다는 사정이 확인된 정보에 대해서는 사건에 따라 인용 이유에 차이는 있으나, 예외 없이 공개거부처분이 위법이라는 취지의 일부 승소 판결을 하였다.

1) 피고 국제협력단 사건(수원지방법원 2022. 5. 26. 선고 2020구합73847)

피고 국제협력단은 비법인사단인 원고가 소송상 당사자능력이 없다고 주장하였으나, 법원은 원고가 독자적인 정관을 가지고 이에 근거하여 대표자를 두고, 의사결정 기관으로 총회 및 심의·의결기구로 운영위원회 등을 두고 있으며, 이들은 각 회원으로 구성된 사실이 인정되므로, 원고는 민사소송법 제52조에 따른 단체성을 갖춘 비법인사단으로써 당사자능력을 인정된다고 판단하면서 피고 주장을 배척하였다.

피고 국제협력단은 소송이 제기되기 전 정보공개청구 단계에서 채용업무처리지침과 공고문을 제출하였다. 그리고 피고는 소송에서 원고가 공개를 청구한 정보 중 4개 항목은 이미 공개한 자료를 통해 알 수 있는 정보라고 반박하였다. 이에 대하여 법원은 ‘원고가 이 사건 소제기 전에 이미 공개된 자료를 통해 이러한 사항에 관한 정보를 확인할 수 있었던 이상, 이 부분 정보에 해당하는 이 사건 목록 제1의 사항, 자항 기재 각 정보를 별도로 공개함으로써 원고가 얻을 수 있는 이익은 없는 것으로 보인다.’라고 판단하여 2가지 항목에 대한 청구를 각하하였고, 2가지 항목에 대해서는 사전에 공개한 자료로 알 수 없는 내용이라는 점을 들어 피고 주장을 배척하였다.

법원은 피고 국제협력단이 부존재를 주장한 자료에 대하여 자료 존재에 대한 입증은 원칙적으로 공개청구자에게 증명책임이 있다고 전제한 뒤 피고가 주장한 내용 및 법원이 직권으로 조사한 사실관계에 비추어 각 정보를 보유·관리하고 있을 상당한 개연성이 있다고 인정하기에 부족하고, 달리 이를 인정할 만한 증거가 없다고 판단하였다. 나아가 이 사건 소송 과정에서 원고에게 공개한 정보에 대해서는 ‘이와 같은 우회적인 방법은 정보공개법에 의한 공개라고 볼 수 없음이 분명하므로, 원고에게는 여전히 이 부분 정보에 관한 공개거부처분의 취소를 구할 법률상 이익이 있다.’라고 보아 원고에게 소를 구할 이익이 있음을 인정하였다.

피고는 일부 정보가 존재한다는 전제하에 그 정보가 정보공개법 제9조 제1항 제5호

소정의 ‘입찰계약에 관한 사항으로서 공개될 경우 업무의 공정한 수행에 현저한 지장을 초래한다고 인정할 만한 상당한 이유가 있는 정보’라고 주장하였다. 하지만 법원은 ① 피고가 비공개한 정보가 용역업체와 인공지능 프로그램의 구매계약을 체결하면서 용역업체로부터 받은 서류인데, 피고는 입찰 방식이 아니라 수의계약 방식으로 용역업체와 계약을 체결하였다는 점, ② 피고가 비공개한 정보는 물품 제조/구매 승낙사항, 한국국제협력단 수의계약 자가점검표(수의계약 관련 규정 포함), 계약보증금 지급보증확약각서, 하자보수보증금 지급보증 협약각서, 조세포탈에 대한 서약서, 청렴계약서, 퇴직자 임원 근무여부 확인 서약서, 용역업체의 법인등기부 등본, 용역업체의 국민연금 사업장 가입자 명부, 개인정보처리 위탁계약서 등이라는 점, ③ 피고는 위 비공개한 정보가 피고 내부규정인 ‘정보공개 운영규정’ 별표 1 한국국제협력단 비공개 대상정보 세부기준(을 제6호증)에서 정하고 있는 비공개 대상정보라고 주장하나, 세부기준에 따르면 ‘발주 전 내부 사전협약사항, 예정가격 설정 결과, 입찰 참여업체 정보 및 투찰가격 정보, 계약체결 전 기술 및 가격협상 내용’을 비공개 대상정보로 분류하고 있으나, 이 사건 정보는 위와 같은 내용을 포함하고 있지 않다는 점, ④ 피고는 인공지능 프로그램을 통한 면접이 확정적인 전형방법이 아니라 내부검토를 하는 상황이므로 이 사건 정보가 의사결정 과정 또는 내부검토 과정에 있는 사항이라는 취지로도 주장하나, 이 사건 정보는 인공지능 프로그램의 내용이나 활용방안에 관한 문서가 아니라는 점, ⑤ 피고가 이 사건 정보를 공개함으로써 공공기관의 계약 관련 정보에 관한 국민의 알 권리를 충족하고, 용역업체를 통해 인공지능 프로그램을 면접에 활용하는 과정에서 개인정보가 안전하게 잘 관리되고 있는지 등을 확인할 수 있는 계기가 될 수 있다는 점을 들어, 피고가 비공개한 정보에 대한 부분은 공개되어야 한다고 판결하였다.

특히 법원이 “공공기관의 계약 관련 정보에 관한 국민의 알 권리를 충족하고 용역업체를 통해 인공지능 프로그램을 면접에 활용하는 과정에서 개인정보가 안전하게 관리되고 있는지 등을 확인할 수 있는 계기가 될 수 있다.”라고 명시적으로 판단한 부분은, 공공기관에서 AI면접 프로그램이 어떻게 활용·관리되고 있는지가 국민의 알 권리 대상이라는 점과 사회적으로 이에 대한 감시가 필요하다는 점을 법원이 간접적으로 인정한 것으로 의의가 있다.

2) 피고 한전KDN 사건(광주지방법원 2022. 6. 16. 선고 2020구단11377)

피고 한전KDN은 소 제기 전 정보공개 청구 단계에서 ‘정보부존재’를 이유로 정보공개를 거부하였다. 그런데 소송에서는 일부 정보가 존재하지 않는다는 주장을 유지하면서도 일부는 정보가 정보공개법 제9조 제1항 제5조의 ‘업무에 공정한 수행에 현저한

지장을 초래한다고 인정할만한 상당한 이유가 있는 정보’와 제9조 제1항 제7호의 ‘경영상·영업상 비밀에 관한 사항으로서 공개될 경우 법인 등의 정당한 이익을 현저히 해칠 우려가 있다고 인정되는 정보’에 해당한다고 비공개 사유를 변경하였다.

법원은 피고 한전KDN이 존재하지 않는다고 주장하는 정보에 대하여는 정보가 존재한다는 점에 대한 입증책임은 원고에게 있는데, 피고가 ‘각 정보를 보유·관리하고 있을 상당한 개연성이 있다고 인정하기에 부족하고, 달리 이를 인정할 만한 증거가 없다.’라고 판단하면서 해당 부분의 원고 청구를 각하하였다.

한편, 법원은 처분청이 처분사유를 추가·변경하는 행위와 관련하여 “처분청은 당초 처분의 근거로 삼은 사유와 기본적 사실관계가 동일성이 있다고 인정되는 한도 내에서만 다른 사유를 추가하거나 변경할 수 있고, 여기서 기본적 사실관계의 동일성 유무는 처분사유를 법률적으로 평가하기 이전의 구체적인 사실에 착안하여 그 기초인 사회적 사실관계가 기본적인 점에서 동일한지 여부에 따라 결정되며 이와 같이 기본적 사실관계와 동일성이 인정되지 않는 별개의 사실을 들어 처분사유로 주장하는 것이 허용되지 않는다고 해석하는 이유는 행정처분의 상대방의 방어권을 보장함으로써 실질적 법치주의를 구현하고 행정처분의 상대방에 대한 신뢰를 보호하고자 함에 그 취지가 있고, 추가 또는 변경된 사유가 당초의 처분시 그 사유를 명기하지 않았을 뿐 처분시에 이미 존재하고 있었고 당사자도 그 사실을 알고 있었다 하여 당초의 처분사유와 동일성이 있는 것이라 할 수 없다.”라고 판단하고 있다⁶⁾.

법원은 위 법리에 기초하여 한전KDN이 당초 부존재를 사유로 정보공개를 거부하였으나, 이후 소송에서 비공개 사유를 변경한 정보들에 대하여 피고가 이처럼 처분사유를 변경한 것에 기초적 사실관계가 동일하다고 볼 수 없다고 판단하였다. 즉, 법원은 피고의 처분사유 추가는 허용되지 않는다고 보아 피고가 보유·관리하고 있다고 자인한 정보 일체를 공개하라며 원고 청구를 인용하였다. 비록 처분사유 추가·변경에 대한 법리적 판단에 기초하여 원고 청구를 인용하였으나, 피고 한전KDN이 한 구체적인 주장에 대한 실체적인 판단을 하지 않아 아쉬움이 남는 판결이다.

3) 상소 여부

위 판결은 원고와 피고 기관들이 모두 항소하지 않아, 원고와 피고 국제협력단 간의 소송은 2022. 6. 11. 확정되었고, 원고와 피고 한전KDN 사이 소송은 2022. 7. 6. 확정

6) 대법원 2003. 12. 11. 선고 2001두8827 등 다수 판결

되었다.

3. AI 면접 운용 과정의 실태와 문제점

가. 정보가 존재하지 않는다는 피고 주장의 의미

1) 공공기관은 채용절차에서 AI면접 프로그램을 사용하였음에도 이와 관련한 목록을 별도로 작성하지 않았다(1.가. 관련)

현행 「채용절차의 공정화에 관한 법률」 제11조와 「공기업·준정부기관의 경영에 관한 지침(2020. 12. 29. 개정되기 전의 것)», 이하 ‘공기업 등 경영지침’)」 제16조는 아래와 같이 규정하고 있다.

「채용절차의 공정화에 관한 법률」(채용절차법)

제11조(채용서류의 반환 등)

① 구인자는 구직자의 채용 여부가 확정된 이후 구직자(확정된 채용대상자는 제외한다)가 채용 서류의 반환을 청구하는 경우에는 본인임을 확인한 후 대통령령으로 정하는 바에 따라 반환하여야 한다. 다만, 제7조제1항에 따라 홈페이지 또는 전자우편으로 제출된 경우나 구직자가 구인자의 요구 없이 자발적으로 제출한 경우에는 그러하지 아니하다.

② ~ ⑥ 생략

「공기업·준정부기관의 경영에 관한 지침(2020. 12. 29. 개정되기 전의 것)」

제16조(채용 공정성 관리) ① ~ ⑥ 생략

⑦ 공기업·준정부기관은 채용과 관련된 서류를, 인사부서와 감사부서에서 동시에 관리하도록 하되, 감사 부서는 감사 권한의 범위 내에서 열람하도록 한다.

⑧ 공기업·준정부기관은 기록물 보존에 관한 내부 규정을 통해, 채용 관련 문서를 영구적으로 보존하도록 보존 기간을 정하여야 한다.

⑨ 공기업·준정부기관은 채용계획 수립, 공고, 서류·면접전형과 필기시험 실시, 합격자 결정 등 직원채용의 절차와 방법 등에 관한 별도의 규정을 마련하되, 개별 채용별로 그 절차와 방법을 기관장 등이 달리 정하도록 할 수 없다.

⑩ 공기업·준정부기관은 제15조 제4항에 따른 제한경쟁시험방식으로 직원을 채용하는 경우(단기채용이나 제9조제3호에 따른 사회형평적 인력 활용을 위한 채용은 제외) 채용목적, 인원, 절차, 기준 등 채용전반에 대해 주무기관의 장과 미리 협의

7) 이 사건 소제기 당시 시행 중이었던 지침이다.

하여야 하며, 같은 회계연도 내에 동일한 절차와 기준으로 동일한 직급·직종(분야)에 대해 채용하는 경우에는 일괄하여 협의할 수 있다. 다만 불가피한 사유로 협의 사항과 달리 채용하게 되는 경우에는 변경사항과 그 사유를 즉시 주무기관의 장에게 보고하여야 한다. <개정 2019.12.27.>

⑪ 공기업·준정부기관은 제15조제4항에 따른 제한경쟁시험방식으로 직원을 채용하는 경우 합격자 발표 전 외부위원이 참여하는 위원회 등을 통해 당해 채용이 관계법령과 내부규정에 따라 이행되었는지 점검하여야 한다.

⑫ 생략

위 규정의 내용을 종합하여 보면, 구인자는 구직자로부터 받은 채용서류 등을 반환 반환하되(채용절차법 제11조), 채용 공정성 관리를 위하여 채용과 관련된 서류(채용서류와 다른 개념이다)는 인사부서와 감사부서에서 관리하고(지침 제16조 제7항), 사후 공정성 관리가 이루어질 수 있도록 채용 관련 문서를 영구적으로 보존하며(동조 제8항), 합격자 발표 전 외부위원이 참여하는 위원회 등을 통해 채용이 관계법령과 내부규정에 따라 이행되었는지를 점검하여야 한다(동조 제11항).

통상 정보공개 청구인은 공공기관이 어떠한 정보와 문서를 보관하고 있는지 알 수가 없으므로 기관이 작성·보관하고 있는 문서의 명칭을 확인하기 위하여 ‘문서목록’의 공개를 청구한다. 원고도 채용 관련 법률 규정에 기초하여 공공기관이 당연히 작성하여 보관하고 있을 것으로 판단하여 첫 번째로 ‘1.가. 공기업, 준정부기관의 경영에 관한 지침 제16조에 따라 인공지능 프로그램과 관련하여 인사부서 및 감사부서에 보관 중인 문서 목록’의 공개를 청구하였다. 그런데, 국제협력단은 AI면접 프로그램을 채용절차에서 참고자료로 사용하고, 한전KDN은 면접결과 점수를 채용절차에 반영하였음에도 이와 관련하여 작성·보관한 문서가 존재하지 않는다고 답변하였다. 이러한 공공기관의 답변은 사후적으로 ‘문서’를 통해 이루어지는 공정성 관리를 위한 절차에서 AI면접 프로그램에 대한 검증과 평가가 이루어지지 않았고, 앞으로도 이루어질 수 없다는 점을 자인하는 것이다.

2) 공공기관은 개인정보보호법상 위탁자의 지위에서 지원자들의 개인정보처리를 제3자에게 위탁하였음에도 개인정보보호법에서 정한 의무를 다하지 않았다(2.나. / 2.다. 관련)

한전KDN은 기관이 개인정보보호법상 위탁자의 지위에서 지원자의 개인정보처리에

관한 업무를 마이다스인에 위탁하여 처리하였다고 취지로 주장하고 하였다(국제협력단은 관련 서류를 소제기 이후 제출하였다). 개인정보보호법 제17조에서 정한 개인정보의 '제3자 제공'은 본래의 개인정보 수집·이용 목적의 범위를 넘어 그 정보를 제공받는 자의 업무처리와 이익을 위하여 개인정보가 이전되는 경우이고, 개인정보보호법 제26조에서 정한 개인정보의 '처리위탁'은 본래의 개인정보 수집·이용 목적과 관련된 위탁자 본인의 업무처리와 이익을 위하여 개인정보가 이전되는 경우를 의미한다(개인정보보호법 제17조 제3자 제공은 처리위탁행위보다 허용범위가 제한적이고, 위반시 형사처벌을 받는다). 개인정보 처리위탁에 있어 수탁자는 위탁자로부터 위탁사무 처리에 따른 대가를 지급받는 것 외에는 개인정보처리에 관하여 독자적인 이익을 가지지 않고, 정보제공자의 관리·감독 아래 위탁받은 범위 내에서만 개인정보를 처리하게 되므로, 개인정보 보호법 제17조에서 정한 '제3자'에 해당하지 않는다⁸⁾.

개인정보보호법 제26조 및 동법 시행령 제28조에 의하면, 위탁자는 개인정보처리에 관한 업무를 문서에 의하여 위탁하여야 하고, 업무 위탁에 따라 정보 주체의 개인정보가 분실·도난·유출·위조·변조 또는 훼손되지 아니하도록 수탁자를 교육하고, 감독하여야 한다(개인정보보호법 제26조 제4항). 개인정보보호법 위탁자는 채용과정에서 수집되는 응시자의 개인정보에 관하여 관리·감독할 책임을 부담하고, 위반할 경우 벌금형 또는 과태료 처분에 처할 수 있다.

따라서 한전KDN은 수집되는 정보의 민감성 및 기관이 부담하는 두터운 관리·감독상 책임에 비추어 피고가 AI면접 진행 과정에서 개인정보처리에 대한 계획을 수립하고(2.다.항 관련), 수탁업체에 대한 교육(2.나.항 관련)을 실시함이 마땅하다. 한전KDN이 이러한 자료가 존재하지 않는다고 주장하는 것은 개인정보보호법상 정해진 위탁자로서 의무 해태를 자인하는 것이다.

더불어 피고는 인공지능 프로그램이 지원자로부터 수집하는 사항(4.라.항 관련)과 지원자 평가를 위해 인공지능 프로그램에 입력하는 지원자의 개인정보(4.마.항 관련)를 알 수 있는 문서도 부존재한다고 밝혔다. 하지만 개인정보 처리위탁 관계가 본래의 개인정보 수집·이용 목적과 관련된 위탁자 본인의 업무처리와 이익을 위하여 개인정보가 수탁자에게 이전되는 것에 그치는 것이라는 점에 비추어 보면 한전KDN이 지원자로부터 수집하는 사항과 지원자가 입력하는 개인정보에 대한 정보가 부존재한다는 사실을 상식적으로 받아들이기 어렵다. 실제 위와 같은 자료가 부존재한다면 한전KDN이 수집되는

8) 대법원 2020. 3. 12. 선고 2017도13793 판결

정보의 구체적인 내용과 범위를 알지 못하였다는 것으로 ‘개인정보 처리위탁’ 관계가 맞는지에 대한 점검이 필요하다.

3) 공공기관은 채용절차에서 발생한 시스템 오류에 대한 민원 해결, 사후 점검 및 조치 업무를 전혀 수행하지 않았다(1.마. / 3.다. / 4.사. / 4.아. 관련)

피고 기관들은 모두 채용절차에서 발생한 시스템 오류에 대한 민원 해결, 사후 점검 및 조치업무를 수행한 문서 등이 존재하지 않는다고 답변하였다(한전KDN의 경우, 해당 자료가 존재하지 않는다는 사정이 명백하여 원고가 해당 문항에 대한 소를 취하였다). 특히, 국제협력단의 경우 2019. 6월 실시한 채용절차의 AI면접 과정에서, 다수 응시자의 시스템에 오류가 발생하여 절차가 중단되는 실제 문제가 발생하였음에도 사후 시스템 오류에 대한 별도 관리를 전혀 하지 않았다는 사정이 확인되었다.

국제협력단이 2019. 6월에 진행한 채용절차에서 응시자 중 일부는 1회 또는 2회 시스템에 재접속을 하였는데, 이 과정에서 재접속한 응시자들은 피고 담당자에게 전화 또는 이메일을 통해 연락하여 시스템의 불안정과 시스템 멈춤 현상을 알렸다. 그런데 국제협력단은 재접속이 문제 되자, 내부 회의를 통해 ‘총 182명 중 1회 재접속한 25명은 면접을 완료한 것으로 처리하고, 2회 재접속한 3명은 면접을 실시하지 않은 것으로 처리’하였다. 국제협력단은 AI면접 과정에서 발생한 시스템 오류와 불안정 문제를 명확히 인지할 수밖에 없는 상황이었다. 이러한 구체적인 사정은 감사원 감사보고서를 통해 밝혀졌다.

사회의 빈부격차는 컴퓨터 등 전자기기의 보급격차로 이어지며, 그러한 보급의 격차는 소유한 전자기기 사양의 격차로 이어질 수밖에 없다. 채용의 공정성은 단순히 기회의 공정성뿐만 절차적 공정성도 보장할 수 있도록 제도를 갖추고 있고, 공공기관은 일부 사회적 약자를 위한 별도 절차를 마련하는 방식으로 절차적 평등까지 실현하고 있다. 그런데, 국제협력단은 위와 같은 AI면접 과정상 시스템과 프로그램의 불안정으로 인하여 3명의 응시자가 다음 전형으로 응시할 기회를 박탈당하였음에도 원고가 공개를 청구한 문항에 대하여 ‘AI면접이 무사히 이행되었다고 인식하고 판단’하여 별도로 시스템 오류에 대한 민원 해결 등을 할 필요가 없었고, 따라서 정보가 부존재한다는 취지로 답변하였다. 시스템상 오류로 피해를 본 피해자는 분명히 존재하는데 그러한 문제(민원)을 해결하기 위한 절차와 사후 재발 방지를 위한 문서가 모두 존재하지 않는다고 답변한 것이다.

이는 피고가 AI면접 과정에서 실제 시스템 오류로 인한 문제가 발생하였음에도 이를 문제로 인지하지 않았고, 사후적으로 이러한 사태의 재발 방지를 위한 대응도 마련하지 않았다는 점을 보여준다.

4) 공공기관은 국가를 당사자로 하는 계약에 관한 법령과 관련 계약예규의 내용을 위반하여 수의계약을 체결하였다(3.가. 관련)

원고는 AI면접 프로그램 입찰과정에서 홈페이지에 공시한 문서에 대한 정보공개를 청구하였고, 이에 대하여 피고 기관들은 ‘수의계약’ 방법을 채택함에 따라 그러한 자료가 존재하지 않는다고 답변하였다(한전KDN의 경우, 해당 자료가 존재하지 않는다는 사정이 명백하여 원고가 해당 문항에 대한 소를 취하하였다). 수의계약은 국가계약법 시행령 제26조, 제27조, 제28조의 사유가 존재하는 경우에 예외적으로 체결할 수 있다. 원고가 정보공개를 청구한 다수의 기관은 동 시행령 제27조(재공고입찰과 수의계약)에 따라 제안요청서를 공개하기도 하였다(국민건강보험공단 일산병원, 한국남동발전 등). 나아가 「공기업·준정부기관 계약사무규칙」 제8조의2는 아래와 같이 수의계약을 체결한 경우 그 월별 해당 수의계약에 관한 사항을 홈페이지에 공개하도록 정하고 있다.

「공기업·준정부기관 계약사무규칙」
제8조의2(수의계약 관련 정보의 공개)
① 기관장 또는 계약담당자는 제8조제1항제2호 및 제7호(「국가를 당사자로 하는 계약에 관한 법률 시행령」 제26조제1항제2호에 따라 수의계약을 체결하는 경우로 한정한다)에 따라 수의계약을 체결한 경우에는 **월별 해당 수의계약에 관한 다음 각 호의 사항(이하 이 조에서 “수의계약 관련 정보”라 한다)을 다음 달 말일까지 해당 공기업·준정부기관의 인터넷 홈페이지에 공개해야 한다.** <개정 2016. 9. 12., 2019. 10. 4.>
1. 사업명과 계약이행기간
2. 계약상대자의 대표자 성명, 상호, 법인등기부상 본점소재지
3. 예정가격과 계약금액
4. 수의계약의 법령상 근거 및 구체적인 사유
② 제1항에 따라 공개한 수의계약 관련 정보는 최소한 해당 계약의 이행을 완료한 날부터 1년이 되는 날까지는 공개해야 한다.. <개정 2019. 10. 4.>

피고 기관들은 AI면접 프로그램 개발업체와 ‘수의계약’을 체결하였음을 인정하였음에

도 동시에 원고가 공개를 청구한 ‘3.가. 인공지능 프로그램을 사용한 기관(발주기관)이 입찰 과정에서 홈페이지에 공시한 문서’가 존재하지 않는다고 주장하였다. 이는 피고가 AI면접 프로그램에 대한 계약을 체결하는 과정에서 국가계약법령과 관련 계약예규의 내용을 위반하여 수의계약을 하였다는 점을 자인하는 것이다.

5) 공공기관은 AI면접 프로그램의 내용과 기술정보를 구체적으로 확인하지 않고 수의 계약을 체결하였다(3.라. / 4.가. / 4.나. / 4.다. 관련)

AI기술은 여타의 기술과 달리 학습에 사용된 빅데이터와 알고리즘에 대한 검증이 반드시 선행되어야 한다. AI면접 프로그램이 공정성과 객관성을 담보하였는지는 해당 기술이 어떤 빅데이터를 학습하여 알고리즘을 생성하였는지, 어떤 메커니즘을 통해 결론에 도달하였는지에 대한 검증을 통해 비로소 확인할 수 있다. 원고는 이를 확인하기 위하여 ‘3.라. 인공지능 프로그램 관련 계약목적물의 다음 기술자료[- 소스코드 및 오브젝트 코드의 복제물, - 기술정보(매뉴얼, 설계서, 사양서, 기능명세서, 플로우차트, 유지보수자료 등)]’에 대한 정보공개를 청구하였다.

당초 원고는 공공기관들이 정보공개법 제9조 제1항 제7호의 법인의 영업상 비밀에 관한 사항임 비공개 사유로 제시할 것이라 예상하였다. 그런데, 피고 기관들의 답변은 ‘인공지능 프로그램을 제작·공급받는 용역이 아니었는바, 관련 기술자료를 제공받은 사실이 없다’라는 것으로, 피고는 AI면접 프로그램의 최소한의 기술내용도 확인하지 않은 채 수의계약을 체결한 것이다.

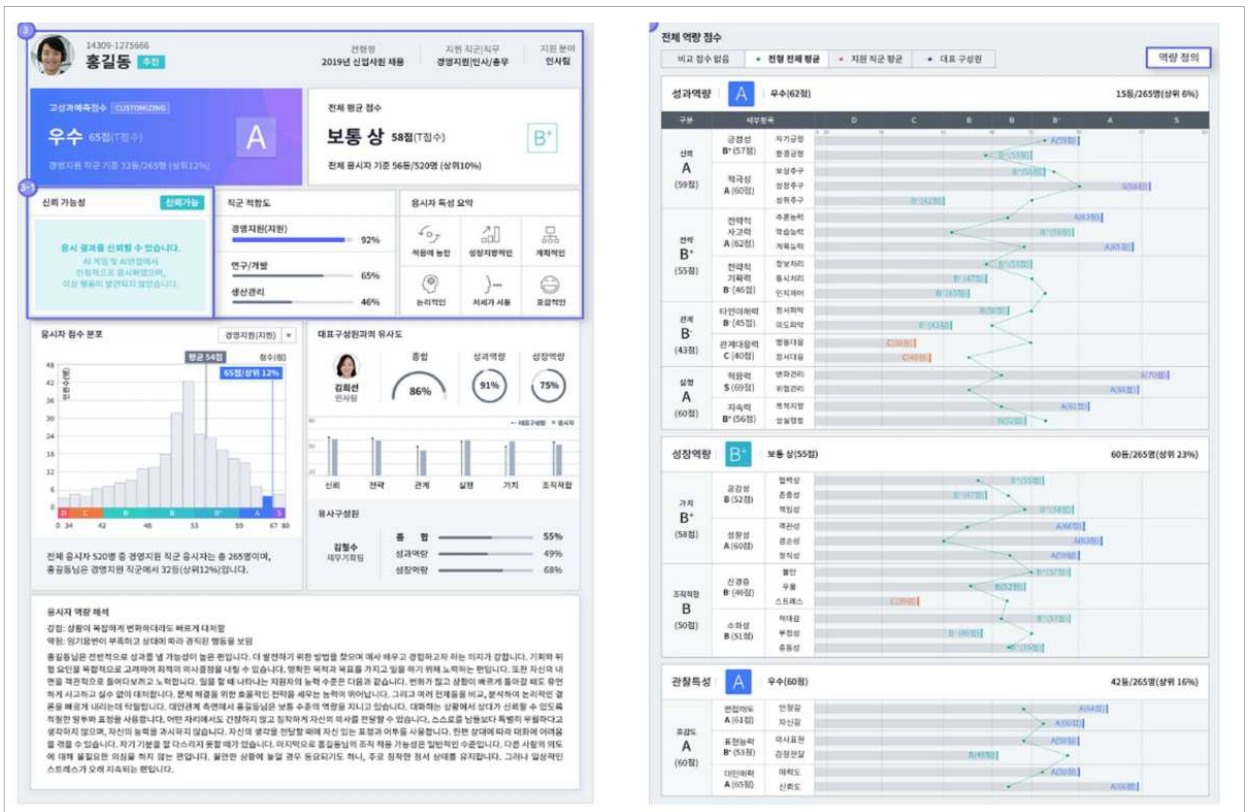
한편, 원고는 같은 맥락에서 공공기관이 AI면접 프로그램의 메커니즘을 얼마나 잘 파악하고 이를 활용하였는지 알기 위하여 ‘4.가. 인공지능 프로그램의 기능을 설명한 자료’, ‘4.나. 진행자, 면접위원 및 지원자에게 배포되거나 제공된 인공지능 프로그램에 대한 교육 및 기술 설명자료, 물품의 항목’, ‘4.다. 인공지능 프로그램이 직접 또는 기관이 지원자에게 출제한 문항을 알 수 있는 문서’에 대한 정보공개를 청구하였다.

이에 대하여 국제협력단은 4.가.와 4.나. 문항에 대하여 직접적인 자료를 제시하지는 않았지만, 소 제기 이후 국제협력단이 제출한 자료로 일부 내용을 확인할 수 있었고, 4.다. 문항에 대하여 ‘AI면접은 해당 절차에 참여하는 지원자들이 가정이나 인터넷 접속이 가능한 장소에서 자신의 컴퓨터로 접속하여 응답하는 방식으로 이루어졌으며, 피고가 출제 문항과 관련하여 용역업체로부터 별도로 제공받은 자료는 존재하지 아니함.’이라고 답변하였고, 한전KDN은 위 자료가 모두 부존재한다고 답하였다.

피고 기관들은 AI면접을 통해 지원자의 역량을 평가하는 절차를 거쳤고, 이를 채용의 참고자료 또는 채용과정에 직접 활용하였음에도 AI면접 프로그램에서 어떠한 질문을 통해 응시자의 역량을 평가하는지 구체적인 내용을 확인하지 않았다는 사정을 확인할 수 있다.

6) 피고는 관련 규정에 반하여 면접위원에게 AI면접 프로그램에 대한 사전교육을 실시하지 않고 결과지를 제공하였다(1.라. 관련)

인사혁신처가 2019년 발간한 공정채용가이드북에 따르면, ‘면접위원의 평가오류를 줄이고 면접역량 강화를 위해 면접위원 유의사항 등 사전교육을 반드시 실시’하라고 정하고 있고, 피고 채용업무처리지침 제4조 제5항도 채용담당자 및 면접위원에 대해 채용 관련 교육을 실시하도록 정하고 있다.



AI 면접 결과지 일부

국제협력단이 제출한 위 AI면접 결과표를 살펴보면, 응시자 역량 ‘해석’을 비롯하여 성과역량, 성장역량, 관찰특성으로 구분하여 평가하면서 총평과 점수, 등수를 제공하고, 직군에 따른 직무적합도까지 분석하여 제공하고 있다. AI면접 프로그램에 대한 사전교

육이 선행되지 않은 상태에서 해당 자료만 제공하였을 경우, 이를 마치 응시자의 능력에 대한 절대적인 점수, 등수로 인식하여 면접위원 개인의 객관적인 평가를 저해할 가능성이 큰 자료이다.

이처럼 공공기관이 AI면접 프로그램에 대한 최소한의 정보도 확인하지 않은 상태에서 그 결과물을 면접위원에게 사전 설명 없이 제공한 행위는 관련 규정의 취지에 반하는 것이며, 객관적인 평가를 저해할 개연성이 높다.

나. AI면접 프로그램의 처리방법을 확인할 수 있는 자료에 대한 공개 요구의 필요성

1) 알고리즘 등 자료 공개에 대한 해외 판례

AI기술 편향성에 대한 논란은 AI기술의 투명성에 대한 요구로 이어지고 있다. 세계 여러 나라에서 AI시스템을 구성하는 알고리즘의 공정성과 투명성과 관련된 소송이 제기되었고, 각국 법원은 이미 여러 차례 AI 알고리즘이 불투명하게 관리되어 기본권 침해할 초래한 경우, 그 부당함을 주장하는 청구인들과 시민사회의 손을 들어주었다.

2017년 폴란드법원은 정부의 실업자 점수 산출에 사용된 AI 알고리즘(프로파일링 메커니즘)에 대해 투명성과 개인정보보호 문제를 이유로 정보를 공개하도록 하였고, 그 결과 해당 시스템의 비효율성과 차별 위험이 확인되어 해당 AI시스템이 폐기되었다. 미국 교육청은 민간기업을 통해 조달한 교사평가 알고리즘에 기초하여 공립학교 교사들을 해고하였고, 이에 대하여 휴스턴교사연맹이 교육청을 상대로 제기한 소송에서 법원은 투명성과 적법절차 위반 등 청구인 주장을 인용하였고, 결국 프로그램 운영 중단이 결정되었다. 위 휴스턴 판결에서 미국 법원은 민간기업의 영업비밀과 국민의 헌법상 권리인 적법절차를 모두 충족시키기 위해서는 공공기관의 중요한 의사결정에 공개할 수 없는 알고리즘을 사용해서는 안 된다고 실시하였다.

또한, 네덜란드 법원은 2020. 2.경 자동화된 사회복지급여 부정수급탐지 시스템에 대하여 투명성 부족과 개인정보보호법 위반을 이유로 즉시 사용중지 결정을 내렸고, 암스테르담 법원은 2021년경 택시 플랫폼(우버와 올라)에서 운전자 감시에 활용된 AI 알고리즘에 대해 공개 결정을 내렸다. 스페인에서는 AI기술의 투명성에 대한 시민사회의 요구에 부응하여, 플랫폼 기업에 대해 노동자의 근로조건, 고용과 해고 결정에 영향을 미칠 수 있는 알고리즘과 AI에 관한 정보를 근로자 대표에게 공개하도록 의무화하는 법안이 통과되었고, 네덜란드 암스테르담과 핀란드 헬싱키시, 프랑스 앙티브시 등 외국의

지방자치단체들은 AI 알고리즘과 관련한 사항을 시민들에게 먼저 공개하고 관련한 의견을 수렴하고 있다.

위 판결들은 공공영역에서 ‘고위험인공지능’으로 분류될 수 있는 AI시스템이 사용된 경우로서, 해당 AI시스템을 거쳐 선발에서 배제된 시민이 그 이유를 알 수 있도록 AI 시스템과 알고리즘의 구체적인 내용에 대하여 공개하고, 만약 이를 설명할 수 없다면 사용을 중지하라는 취지이다.

2) 국내법상 알고리즘 공개 가능성

정보공개법 제9조 제1항 제7호는 ‘경영상·영업상 비밀에 관한 사항으로서 공개될 경우 법인 등의 정당한 이익을 현저히 해칠 우려가 있다고 인정되는 정보’를 비공개할 수 있다고 정하면서 단서를 두어 가. 사업활동에 의하여 발생하는 위해(危害)로부터 사람의 생명·신체 또는 건강을 보호하기 위하여 공개할 필요가 있는 정보와 나. 위법·부당한 사업활동으로부터 국민의 재산 또는 생활을 보호하기 위하여 공개할 필요가 있는 정보에 대한 정보공개를 예외적으로 인정하고 있다.

일반적으로 AI시스템은 개발은 상당한 시간, 비용과 노력을 투입한 결과물로서, 정보가 입력되어 결과값이 도출되는 처리절차인 알고리즘은 경영상·영업상 비밀에 관한 사항으로 인정되어 비공개대상으로 인정될 가능성이 매우 크다. 특히 사기업이 개발한 AI시스템은 특허등록 대상으로 특허법에 따른 보호를 받는 정보에 해당한다. 한편, AI 면접 프로그램이 지원자들의 채용과정에 직접적인 영향을 미치더라도 ‘사업활동에 의하여 발생하는 위해(危害)로부터 사람의 생명·신체 또는 건강을 보호하기 위하여 공개할 필요가 있는 정보’와 ‘위법·부당한 사업활동으로부터 국민의 재산 또는 생활을 보호하기 위하여 공개할 필요가 있는 정보’에 해당한다고 보기에 한계가 있다.

현재 일부 공공기관은 AI면접 프로그램을 통한 결과지를 참고자료로만 활용하고 있고, 일부 공공기관은 이를 점수화하여 채용에 직접 반영하고 있다. 국제협력단은 이를 참고자료로 활용하고(다만, 이에 응시하지 않을 경우 다음 전형 응시가 어려우므로 이를 참고자료로 보기 어려운 부분도 존재한다), 한전KDN은 이를 채용과정에 적극적으로 활용하였다. 채용은 입사를 위한 선발 과정이며, 이러한 절차가 개인의 삶에 상당한 영향을 미친다는 점은 특별한 설명이 불필요할 정도로 자명하다. 따라서 지원자 불합격 사유에 대하여 AI면접 프로그램 결과물이라는 결론에 그치지 않고, 결과값에 영향을 미치는 요소에 대한 정보를 제공하여 합리적인 수준에서 채용결과에 대한 설명이 가능하

여야 한다. 하지만 현행 법체계 하에서는 정보제공과 설명을 요구하여 유의미한 결과물을 취득하기에 한계가 있다.

다. 소결

피고 기관들은 원고가 청구한 상당수의 정보에 대하여 해당 정보가 ‘부존재’한다고 반복적으로 답변하였고, 이러한 피고 기관들의 태도에 비추어 해당 자료는 실제 존재하지 않았던 것으로 판단된다. 피고 기관들은 AI면접 프로그램을 몇 년째 채용절차에 반복적으로 사용하고 있음에도 관련 법령에서 요구하는 문서가 존재하지 않았다. 최소한의 기술검증을 위해 작성되었을 것으로 판단되는 문서조차 공공기관에 존재하지 않는다는 사실은 원고가 예상한 것보다 많은 부분이 개발업체에 전적으로 위임한 상태에서 AI면접 절차가 진행되었다는 사정을 보여준다.

정보가 존재하지 않는다는 피고 답변을 통해 AI면접 프로그램이 최소한의 검증도 이루어지지 않은 상태로 실제 채용절차에 활용되어왔고, 공공기관은 채용업무를 수행하는 과정에서 준수하여야 하는 규정을 상당 부분 위반한 상태로 AI면접 프로그램이 사용됐다는 사실을 확인할 수 있다. 공공기관이 관련 문서가 존재하지 않는다고 회신한 자료는 판결서에 명확히 기재되어 있는바, 향후 ‘고위험 인공지능’의 상용화 과정에 대한 공론화 과정에서 반드시 짚어볼 필요가 있을 것으로 보인다.

정보공개거부처분취소소송에서 존재하지 않는 자료를 제외한 대부분 자료가 공개되었지만, AI면접 프로그램의 평가 기준이나 방식, AI면접 프로그램 학습에 사용된 데이터가 어떤 것인지 나아가 AI면접 프로그램 자체가 별도의 차별 또는 다른 인권침해의 위험성은 없는지 등이 여전히 불투명한 상황이다. 채용절차공정화에 관한 법률 제4조의 3에서는 공정한 채용을 위해 직무수행과 무관한 구직자의 정보(신체적 조건, 출신지역, 재산, 학력, 직업, 재산)를 응시원서나 이력서에 기재하지 않도록 정하고 있고, 이를 어기는 경우 500만 원 이하의 과태료를 부과하도록 정하고 있는데(동법 제17조), AI면접 프로그램은 응시자로부터 생체정보를 포함한 여러 가지 정보를 수집하면서도 위와 같은 법률의 규율 대상에서 벗어나 있다. AI면접 프로그램에 대한 모든 정보를 공개하지는 않더라도 해당 AI시스템의 공정성에 대한 검증이 합리적인 수준에서 이루어질 수 있도록 제도개선이 필요하다.

이 사건 정보공개거부처분취소소송은 AI면접 프로그램에 대한 구체적인 정보를 확인하기 위한 목적에서 제기하였으나, 공공기관이 이러한 사실관계를 확인할 수 있는 자료

를 전혀 생성·소지하지 못한 상태였으므로 AI면접 프로그램에 대한 유의미한 정보를 확인할 수는 없었다. 다만, AI면접 프로그램에 대한 관리·감독이 전혀 되지 않는 상황에서 이미 해당 프로그램이 공공기관 채용에 광범위하게 활용되었다는 사실을 확인하였다는 측면에서 의미를 찾을 수 있다.

4. 나가며 : 개선안에 대한 제언

AI가 인간의 사고 프로세스를 모방하여 컴퓨터 기술로 구현한 것이라는 점에 집중하여 본다면, 편견과 확증편향 등 사람의 사고가 가지는 한계도 AI기술을 통해 구현될 수 있다. 최근 AI기술을 규율하고자 하는 다수의 법안이 발의되었으나, 대부분 산업이나 기술의 진흥을 주된 목적으로 하고 있으며, AI기술로 인한 인권침해 가능성을 염두에 둔 법과 제도가 마련되지 않은 상황이다.

앞서 피고 기관들이 작성·보유하고 있지 않다고 주장한 자료들을 통해, 현재 AI면접 프로그램이 전적으로 개발업체 관리하에 활용되고 있다는 사실을 확인하였다. 나아가 각 피고 기관들은 결과지만 받아보는 수준의 개입을 하고 있다는 점과 피고 기관들이 학습에 사용한 데이터와 알고리즘의 주요요소에 대하여 고려하지 않았다는 점도 확인하였다. 그런데, 이처럼 채용과정에서 이용하는 AI시스템은 기본권에 직접적인 영향을 미치는 것으로 평가할 수 있으므로 학습에 사용한 데이터와 알고리즘의 주요요소 등에 관한 확인을 통해 사전에 AI시스템의 편향성 검증이 이루어지도록 사전영향평가 등과 같은 사전 검증제도를 법제화하여야 한다.

이를 위해서는 학습에 사용한 데이터와 알고리즘 주요요소에 대하여 검증이 가능한 수준으로 공개가 이루어져야 한다. 인권위원회는 2022. 5.경 <인공지능 개발과 활용에 관한 인권 가이드라인>을 발표하였다. 가이드라인에서도 인공지능이 개인의 생명이나 안전 등 기본적 인권에 중대한 영향을 미치는 경우 사용한 데이터와 알고리즘의 주요요소를 일반에 공개하고 설명할 수 있어야 한다고 권고하고 있다.

한편 감사원 감사를 통해, 국제협력단이 2019. 6. 진행한 채용절차에서 시스템 불안정과 멈춤 현상으로 인하여 2회 재접속한 3명의 지원자는 AI면접을 실시하지 않은 것으로 처리되었고, 다음 전형에 응시할 수 없는 불이익을 받았다. 감사원은 이에 대하여 ‘① 앞으로 인공지능 면접을 활용하면서 접속 오류 등 응시자의 귀책사유로 판단하기 어려운 경우 응시자의 채용시험 응시 기회를 제한하는 일이 없도록 직원채용 업무를 철저히 하고, ② 관련자에게 주의를 촉구하시기 바랍니다.(주의)’라고 조치사항을 통지

하였다. 감사결과에 따르면 AI면접 프로그램 시스템 불안정과 멈춤 현상의 원인을 분석하여 이러한 일이 재차 발생하지 않도록 점검할 의무가 있으나, 국제협력단은 사후 점검 이행하지 않았다. 따라서 AI시스템을 활용한 경우, 사후 기술적 오류를 점검하고, 결과값을 분석하는 사후 검증제도 또한 법제화하여야 한다.

AI 관련 시장은 규모는 커지고 있고, AI기술이 실생활에 활용되는 빈도도 갈수록 증가하고 있다. 지금까지 살펴본 AI면접 프로그램을 비롯하여 많은 선택과 판단이 사람이 아닌 AI시스템에 의하여 이루어질 것이고, 이는 개인의 권리에 직접적인 영향을 미칠 수 있다. 하지만 AI시스템의 한계에 사회문제가 지속적으로 제기되는 만큼 이를 관리·감독할 수 있는 법과 제도에 대한 심도 있는 고민도 병행되어야 한다. □

보론 : 인공지능 규율 제도의 도입이 필요함

장여경 상임이사 | 사단법인 정보인권연구소

○ 공공부문에서 개발 및 활용하는 인공지능의 위험은 다음과 같음

- 영국 앨런튜링 연구소는 공공부문 인공지능이 고려할 위험을 다음과 같이 분류함¹⁾

- ▶ 편향 및 차별
- ▶ 개인 자율성, 권리구제, 권리행사 거부
- ▶ 불투명성, 설명불가능성, 부당한 결과
- ▶ 프라이버시 침해
- ▶ 사회적 관계 단절 및 고립
- ▶ 신뢰할 수 없고, 안전하지 않으며, 품질이 낮은 결과물

- 영국 공직생활윤리위원회는 공직생활 원칙에 미치는 인공지능의 위험을 다음과 같이 진단함²⁾

- ▶ (공개성에 대한 도전) 관련 정보를 충분히 제공하지 않을 경우 투명성 저해 위험
- ▶ (책임성에 대한 도전) 조직의 책임체계 모호, 공직자 의사결정의 책임소재 불분명, 의사결정 설명 불능 등으로 책임성 불명확화 위험
- ▶ (객관성에 대한 도전) 데이터 편향으로 차별의 확산·증폭 위험

1) The Alan Turing Institute (2019). A guide for the responsible design and implementation of AI systems in the public sector.

2) The Committee on Standards in Public Life (2020). Artificial Intelligence and Public Standards.

○ 특히 채용 분야에서 인공지능이 차별적 의사결정을 수행하였을 경우 이는 현행 법률이 금지하는 차별일 수 있음³⁾

- 인공지능의 차별 유형으로는 △의도에 따른 고의적 차별의 가능성 △편견의 학습 가능성 △학습 데이터의 부족 및 대표성의 결여 형식으로 이루어질 수 있음

- 직접 장애 여부를 묻지 않고도 비슷한 장애를 가진 다른 사람의 데이터 학습을 통하여 장애가 있을 가능성이 높은 지원자를 고의적으로 선택에서 배제하는 등의 사례가 발생할 수 있음. 예를 들어 파이메트릭스(Pyometrics)의 AI 면접 프로그램은 ADHD, 난독증(Dyslexia)과 색맹(color blindness)에 대한 장애를 식별할 수 있음

※ 미국에서는 페이스북에서 고용, 채용, 금융 분야 광고주가 장애인, 여성, 외국인, 한 부모, 청소년 가정, 특정 인종, 특정 종교와 관련이 있는 이용자를 선택에서 배제한 데 대하여 집단소송이 제기되기도 하였음

- 고의적인 차별 의도가 없었다 하더라도 인공지능 챗봇 <이루다> 사건에서처럼 차별적 발언을 학습하였을 경우, 간접적으로 차별적인 결과를 도출할 수 있음

- 인공지능이 학습에 사용하는 데이터에 역사적이고 사회적인 편견과 차별을 상관관계와 대리변수에 반영되어 이를 항구화할 가능성이 있음. 예를 들어 차별 대상에 대하여 부정적인 판단이 있거나 또는 의도적으로 배제한 사항이 반영되어 있을 수 있음. 이 경우 해당 데이터를 학습한 모델은 동일한 행동을 모방하고 이를 환류함으로써 채용에서 차별을 반복할 수 있음

※ 미국 아마존사가 사용해온 채용 시스템이 지원자 이력서를 평가하면서 ‘여성체스클럽’과 여자대학교 이름 등 여성 관련 단어에 대하여 감점한 것으로 나타나 사용이 중단된 바 있음. 과거 10년의 우수사원 데이터로 학습한 이 시스템이 주로 남성 직원을 선호하는 패턴을 나타낸 것으로 밝혀짐

- 학습데이터의 문제에는 대표성 결여 문제도 있음. 장애인의 안면, 표정, 행동 등의 특징적 요소에 대한 데이터가 없어 알고리즘이 이를 학습하지 못한 경우, 해당 알고리즘이 장애인의 특정 행동이나 표정에 대해서 부정적으로 판단할 가능성이 있음. 결국 데이터상 대표성이 적은 장애인과 같은 소수자에 대해서 불리한 편향이 나타날 수 있음.

3) 김예진 (2022). “인공지능에 의한 자동화된 의사결정의 위험성 - 장애인 채용을 중심으로”. LAW & TECHNOLOGY 제18권 제1호 통권 제97호, p.37-57.

○ 이에 세계 각국은 특히 공공부문 인공지능에 대한 요구사항을 제도화하기 위해 노력해 옴

- 영국 정부는 공공기관이 인공지능을 활용할 때 6가지 요소를 반드시 고려하도록 함⁴⁾

구성	세부내용
데이터 품질	▶ 활용의 성공 여부는 데이터 품질의 우수성이 핵심
공정성	▶ 인공지능 모델은 관련된 훈련과 테스트가 중요하며, 정확하고 일반화 가능한 데이터셋 활용도 중요 ▶ 인공지능 시스템이 의도된 목적에 부합해 개발될 수 있도록 전문 지식을 보유한 인력이 개발한 것인가?
책임성	▶ 인공지능 모델의 각 요소를 담당하는 사람과 인공지능 시스템의 설계자 및 구현자의 최종 책임을 묻는 방법을 고려
개인정보 보호	▶ 유럽연합 개인정보보호법(GDPR) 및 영국데이터보호법(DPA 2018) 과 같은 데이터 법·제도의 준수 여부
설명가능성 및 투명성	▶ 인공지능 모델이 결론에 도달한 방법을 설명가능한가?
비용	▶ 인공지능 인프라 구축, 실행 및 유지보수, 관련 인력 훈련 및 교육 등 인공지능 도입 비용과 그에 따른 경제적 효과(혜택, 이익)를 비교

- 영국 공직생활윤리위원회는 공공기관 인공지능 윤리를 구현하기 위한 제도 마련을 제안함

<ul style="list-style-type: none"> ▶ 공직생활 7대 원칙(사리사욕 금지, 청렴성, 객관성, 책임성, 공개성, 정직성, 통솔력)이 인공지능 시대에도 충분히 관련성이 있으며 유효함 ▶ 모든 공공기관 인공지능에 대하여 현행 법률 준수를 요구하고, 특히 영국 평등인권위원회에 공공부문 인공지능의 평등법 준수지침 개발을 요구 ▶ 정부에 인공지능 공공조달 규칙 마련, 공공기관 인공지능 영향평가의 의무적 실시 및 공개 제도 마련, 일정한 수준에서 공공기관 인공지능에 대한 정보 공개 기준 마련을 요구 ▶ 공공부문 인공지능을 규제하는 전문기구로서 데이터윤리혁신센터 강화를 제안 등

- 특히 영국 정부 인공지능 조달지침은 입찰 공고 시부터 인공지능 시스템의 ‘블랙박스’ 및 공급업체에 대한 종속(lock-in) 방지를 요구함⁵⁾

4) Government Digital Service and Office for Artificial Intelligence (2019). A guide to using artificial intelligence in the public sector.

5) Office for Artificial Intelligence (2020). Guidelines for AI procurement.

○ 각국은 공공부문 인공지능에 합법성, 투명성, 책무성 등을 요구하는 판례 및 모범 사례를 수립해 옴

- 미국 텍사스 휴스턴의 연방지방법원은 민간 기업에서 조달한 교육청의 교사 평가 알고리즘에 대하여 투명성과 적법절차 부족을 이유로 운영을 중단시킴. 특히 법원은 민간 기업의 영업 비밀과 국민의 헌법상 권리인 적법절차를 모두 충족하기 위해서는 공공기관의 중요한 의사결정에 비밀 알고리즘을 사용해서는 안된다고 설시함⁶⁾

- 네덜란드 헤이그 지방법원이 사회복지급여 부정수급 탐지 시스템에 대하여 영업 비밀을 이유로 한 투명성 부족과 개인정보보호법 위반을 이유로 운영을 중단하라는 취지의 판결을 내림⁷⁾

- 네덜란드 암스테르담과 핀란드 헬싱키 시는 2020년 시민들에 알고리즘 등록부를 시범적으로 공개하였음. 이 알고리즘 등록부는 시가 도입·운영하는 인공지능 시스템의 △학습 데이터셋에 대한 정보 △데이터 처리에 대한 정보 △차별 방지에 대한 정보 △인간 감독에 대한 정보 △위험성에 대한 정보 등을 읽기 쉬운 평문으로 공개하는 한편 이를 책임지는 공직자의 이름, 부서 및 연락처도 공개하여 시민들이 의견을 제출할 수 있도록 하였음⁸⁾

- 프랑스 앙티브시는 2021년 2월부터 알고리즘 주민공개 제도를 실시하고, 각 알고리즘의 △행정 정보(관계 행정부서명, 관리/위탁업체, 행정부서 연락처, 정보 업데이트일자) △알고리즘 관련 사업 및 의사결정에 대한 정보(알고리즘명, 배경, 알고리즘 목적, 의사결정 절차, 의사결정 자동화 수준, 법적 근거, 관련자료) △의사결정 영향에 대한 정보(연간 이루어진 행정결정 수, 결정 범위, 결정의 영향을 받는 사람들) △알고리즘 작동원리에 대한 정보(처리 데이터, 처리 데이터 출처, 처리 데이터 수집방법, 알고리즘 유형, 알고리즘의 수행 작업)를 공개하였음⁹⁾

6) Federal Suit Settlement: End of Value-Added Measures in Houston. Education News 보도 (2017. 10. 10.); HOUSTON FED. OF TEACHERS v. HOUSTON INDEPENDENT. <https://www.leagle.com/decision/infdco20170530802#> 참조.

7) Welfare surveillance system violates human rights, Dutch court rules. The Guardian 보도(2020. 2. 5.); Welfare Surveillance on Trial in the Netherlands. Human Rights Watch (2019. 11. 8).

8) 암스테르담시 알고리즘 등록부 <<https://algorithmerregister.amsterdam.nl/en/ai-register/>>; 헬싱키시 알고리즘 등록부 <<https://ai.hel.fi/en/ai-register/>>

9) Antibes publie l'inventaire de ses algorithmes. laGazette 보도(2021. 2. 11).

○ 채용 등 고위험 분야에서 개발 및 활용되는 인공지능에 대하여 규율하는 법제도 마련이 시급함

- 유럽의회에서 입법 절차가 진행 중인 인공지능법안(Artificial Intelligence Act)은 △교통 등 주요 인프라 △입시 등 교육훈련 △장난감 등 안전성 △고용 및 플랫폼 입점 △신용평가·사회복지 등 필수서비스 △법 집행 △출입국관리 △재판 등에 사용되는 인공지능에 대하여 ‘고위험’으로 분류하고, 이에 대하여 위험 평가 및 완화, 데이터셋의 고품질 보장, 기록 및 문서화, 정보 공개, 인적 통제, 견고성·보안성·정확성 및 사전적합성 평가(ex-ante conformity assessment) 등 의무를 부과함

- 미국은 연방거래위원회(FTC)가 2020년과 2021년 인공지능 사용 지침을 연달아 발표하고 기업이 개발하고 판매하는 인공지능이 투명성, 설명 가능성, 공정성, 견고성과 실증적 타당성, 책임성을 갖추도록 하였음. 또한 인종적으로 편향된 알고리즘의 판매나 사용, 고용·주택·신용평가·보험·복지급여에 대한 거부 결정시 불공정한 행위, 인종·피부색·종교·출신국가·성별·혼인여부·연령·공공부조 수령여부에 따른 차별을 금지하였음. FTC와 미국 의회는 고위험(highly sensitive) 자동화된 의사결정 시스템에 대하여 영향평가(algorithmic impact assessment) 의무를 부여하는 알고리즘 책무성법안(algorithmic accountability bill)의 입법을 추진 중임. 법안은 자동 결정시스템이나 보안 또는 개인 데이터를 포함하는 정보 시스템이 (1) 보안 또는 개인 정보 문제를 발생시키거나, (2) 상당수 개인 정보를 포함하고 있거나, (3) 공개적으로 접근 가능한 대규모의 물리적 위치를 체계적으로 모니터링하는 경우, 고위험으로 간주하고, 그 평가에서 (1) 시스템에 대한 자세한 설명을 포함하여야 하며 (2) 시스템의 상대적 비용과 편익을 평가하며, (3) 개인 정보의 프라이버시와 보안에 대한 위험을 결정하고, (4) 발견된 경우 그러한 위험을 최소화하기 위해 취한 단계를 설명하도록 함

○ 인공지능의 안전과 인권 침해를 예방하고 해결하기 위한 제도적 대책의 하나로 영향평가가 주목받고 있으며, 일부 국가는 공공부문에 대하여 이미 영향평가를 실시중임

- 캐나다 정부는 2019년 연방정부의 조달 정책 및 제도를 소관하는 재정위원회의 훈령으로 「자동화된 의사 결정 훈령(Directive on Automated Decision-Making)」을 제정하여 시행 중임. 외부 대상으로 추천이나 행정적 의사결정을 하기 위하여 알고리즘 시스템, 도구, 통계적 모델을 개발하거나 조달하는 연방정부 기관은 위험도에 따라 △동료 검토 △고지 △의사 결정에 대한 인간의 개입 △설명 요구 △교육훈련 △비상 계획 △시스템 구동 승인에 대한 요구사항을 적용하도록 하고, 이 위험도를 평가하기 위한

알고리즘 영향평가를 도입함. 실시 시기는 프로젝트 설계 단계 초기에 우선 실시하고, 시스템의 생산 전에도 두 번째로 실시하여 요구사항이 구축된 시스템에 반영되었는지 확인하도록 함. 두 번째 평가 결과는 일반 접근이 가능한 형식으로 온라인에 공개하여야 함. 시스템의 기능 또는 범위가 변경되면 평가를 갱신하여야 함¹⁰⁾

- 영국 정부는 <인공지능 조달지침>을 발표하고 공공조달을 통하는 인공지능에서 10대 원칙을 따르도록 요구함. 지침은 특히 두 가지 방식의 평가를 요구하고 있는데, 조달 절차 개시 전 데이터 평가를 실시하고, 조달 절차 개시 단계에서 인공지능 배치의 편익과 위험에 대하여 평가할 것을 요구함¹¹⁾

- 영국 국민 보건 서비스(NHS)는 2022년부터 보건의료 분야에 특화된 알고리즘 영향 평가를 도입함¹²⁾

- 덴마크 국가인권기구는 2020년 <디지털 업무 인권영향평가 지침>을 발간하고 공공 부문과 민간 부문 모두에 대하여 인공지능 등 디지털 사업, 제품 및 서비스가 야기, 기여, 관련된 위험의 특성을 평가하고 해결하도록 함¹³⁾

- 네덜란드 세무 당국은 알고리즘 시스템에 기반하여 저소득 가정 부모와 돌봄 노동자 수만 명을 아동복지 급여 부정수급 혐의로 고발하였으나, 이 의사 결정이 소수 인종에 편향적이었다는 사실이 2021년 알려지면서 큰 논란을 빚음. 특히 이 시스템은 특정 개인을 부정수급 혐의자로 분류한 이유를 제공하지 않아 공공부문의 투명성, 책무성, 감독이 결여되었다는 비판을 받음. 2022년 4월 5일, 네덜란드 하원은 정부에 인권영향평가를 의무화할 것을 요구하는 결의안을 채택함. 결의안은 공공기관이 알고리즘을 사용하여 사람에 대한 평가 또는 결정을 내릴 때 사전 인권영향평가를 의무화하고 그 평가 결과 또한 가능한 한 공개하도록 의무화할 것을 요구함¹⁴⁾ □

10) Government of Canada. Directive on Automated Decision-Making; 사단법인 정보인권연구소 (2021). 인공지능(AI) 개발과 활용에서의 인권 가이드라인 연구. 국가인권위원회 연구보고서. pp. 85~86.

11) UK Government. Guidelines for AI procurement.

12) Ada Lovelace Institute (2020). Algorithmic impact assessment: user guide.

13) The Danish Institute for Human Rights (2020). Guidance on Human Rights Impact Assessment of Digital Activities.

14) MOTIE VAN DE LEDEN BOUCHALLIKH EN DEKKER-ABDULAZIZ(Nr. 835). Voorgesteld 29 maart 2022; Utrecht University (2022). Government of the Netherlands (2022). Impact Assessment Fundamental rights and algorithms.

<<https://www.government.nl/documents/reports/2022/03/31/impact-assessment-fundamental-rights-and-algorithms>>.

개인정보보호위원회 토론



김직동 과장 | 신기술개인정보과

공정거래위원회 토론



선중규 과장 | 시장감시총괄과

국가인권위원회 토론



이진백 주무관 | 인권정책과

국민권익위원회 토론



박지원 과장 | 청탁금지제도과

메모
