이달의 초점

사회보장영역에서 인공지능 기술 적용 현황과 과제

국내 사회보장 영역에서 인공지능 기술 적용 현황과 함의

|김은하|

국외 사회보장 영역의 인공지능 기술 적용 현황과 규제

|김기태·신영규|

한국 인공지능기본법의 내용과 사회보장제도 내 한계

|양승엽|

사회보장 영역 인공지능 기술 적용의 순기능 및 위험성에 대한 검토

|김기태|

한국보건사회연구원

KOREA INSTITUTE FOR HEALTH AND SOCIAL AFFAIRS

사회보장 영역 인공지능 기술 적용의 순기능 및 위험성에 대한 검토[1)]

Assessment of the Positive Impacts and Potential Risks of AI Applications in Social Security

김기태 한국보건사회연구원 사회보장정책연구실 연구위원

이 글에서는 사회보장 영역에서 인공지능 기술이 갖는 순기능과 위험성을 종합적으로 검토하고, 이에 대한 정책적 대응 방향을 제시한다. 먼저 국내외에서 인공지능의 활용 분야를 살펴봤다. 인공지능은 본인 인증, 자격심사, 복지급여 산정, 부정수급 탐지, 위험 분류, 맞춤형 서비스, 사례관리, 행정 효율화, 정책 평가 등 아홉 가지 영역에서 활용되고 있다. 인공지능 기술의 순기능으로는 행정 효율성 제고, 급여의 적시성과 정확성 향상, 사각지대 해소, 개인 맞춤형 서비스 제공, 정책 평가 기반 마련 등이 있다. 반면 개인정보 침해, 데이터 편향, 알고리즘 결정의 불투명성, 영리적 정보 활용 가능성, 설명 불가능성 등은 중요한 위험 요소로 지적된다. 이에 따라 데이터 품질 향상, 표준화, 통합적 거버넌스 구축, 사회보장 영역을 포함한 법제 정비, 전문조직의 신설과 국제 동향에 대한 모니터링이 필요하다는 정책 제언을 담았다.

1 들어가며

인공지능 기술 적용 과정은 필연적으로 윤리적인 쟁점을 낳는다. 김명주(2024)는 인공지능의 특성에 근거한 윤리적 원칙으로 공공성, 공정성, 책임성과 책무성, 안전성과 보안성을 제시했다. 더불어 인공지능의 차별화한 특성에 근거한 추가적인 윤리 원칙으로 통제 가능성, 투명성과 설명 가능성, 개인정보 및 사생활 보호를 들었다. 이와 같은 원칙들은 기술 적용 현장에서 서로 충돌되고 모순을 낳을 가

1) 이 글은 김기태, 신영규, 김명주, 김은하, 변소연. (2024). 사회보장행정에서 인공지능 적용 동향과 함의(한국보건사회연구원)의 내용 일부를 수정·보완하여 작성한 것이다.

능성이 높다. 사회적 공론화 과정을 통해 원칙들의 선순위를 정하고, 원칙들을 투명하고 공정하게 적용할 수 있는 민주적인 체계의 구축이 필요하다.

국내의 경우 조달정보개방포털과 나라장터에서 연구 시점 기준으로 최근 5년 동안(2019년 10월~2024년 9월) 확인되는 사회보장 분야의 복지기술 활용 현황을 조사해 36건의 사업을 확인했다(김기태 외, 2024). 이들 사업의 특징을 종합하면 다음과 같다. 첫째, 사회보장 분야의 복지기술은 노인에 대한 돌봄 수요에 다수 활용됐다. 둘째, 복지기술은 업무 담당자의 효율적 업무 처리를 지원하기 위해 활용됐다. 셋째, 빅데이터의 수집과 축적을 기반으로 하는 정보 제공 서비스가 다수 발견됐다. 넷째, 개인에게 특화된 맞춤형 서비스가 다수 차지했다. 마지막으로 정보 활용의 효율성과 효과성을 높이기 위한 데이터베이스 시스템의 활용 양상이 확인됐다. 종합해 보면 국내의 복지기술은 고차원적인 인공지능 기술이라기보다는 낮은 수준의 인공지능 기술 활용 수준에 머물고 있음을 확인할 수 있었다.

해외의 사례도 마찬가지다. 챗봇 등 상대적으로 초보적인 수준의 인공지능 기술이 활용되고 있다. 참고로, 해외 사회보장 영역에서 나타나는 양상은 국내와 비교했을 때 두 가지 특징이 있다. 첫째, 챗봇에 대한 활용도가 상대적으로 높았다. 특히 브라질, 아르헨티나, 파나마 등 남미 국가에서 활용도가 높다는 점이 특징적이다. 둘째, 사회보장 행정의 인공지능 적용 과정에서 적지 않은 사회적 논란을 낳고 있다. 특히 네덜란드와 덴마크 등에서는 개인정보 유출, 알고리즘의 차별적 관리, 국가 기관의 과도한 개입 등을 둘러싼 논란이 지속되고 있다. 한국에서 개인정보를 활용한 복지행정이 대규모로 이뤄지고 있지만, 이와 같은 논란이 드물었던 점을 고려하면 주목할 만한 대목이다. 한국에서 빅데이터와 인공지능을 활용한 행정 과정에서 서구와 같은 논란이 앞으로 불붙을 가능성도 배제할 수 없다. 불필요한 사회적 갈등을 막기 위해서는 관련 제도와 규제를 점검할 필요가 있다. 이 글에서는 국내외의 문헌 분석을 종합하면서 ① 사회보장 영역에서 인공지능 기술이 적용되는 아홉 영역을 살펴보고, ② 인공지능 기술 적용에 따른 순기능과 위험성을 각각 일곱 가지씩을 점검한 뒤 ③ 순기능을 극대화하고 위험성을 예방하기 위한 여덟 가지의 정책 제언을 시도한다.

2 사회보장 영역의 인공지능 기술 활용

국내외의 인공지능 활용 양태를 보면 사회보장 영역에서 인공지능이 활용되는 영역을 확인할 수 있다. 과거 복지국가에서 디지털 및 인공지능 기술이 활용되는 영역을 제시했던 Alston(2019), 김기태(2024), Zaber et al.(2024)의 논의를 종합하면 아래 아홉 가지 영역이 제시될 수 있을 것이다.[2)]

2) OECD(2024)도 'Modernizing Access to Social Protection' 보고서에서 인공지능의 활용 영역을 다음과 같은 네 가지로 제시했다. ① 공무원의 문서 작업 자동화 및 효율화, ② 개인이나 지역사회의 위험 예측 혹은 예방적 접근, ③ 급여 자격 기준 등에서 결정 자

첫째, 본인 인증(identity verification)이다. 본인 인증은 급여 신청, 자격심사, 급여 지급 과정에서 반드시 필요하다. 물론 한국에서는 지문 및 얼굴 정보가 포함된 주민등록 데이터베이스를 기반으로 매우 높은 수준의 본인 인증 체계가 구축돼 있음을 확인할 수 있다. 한국에서는 인공지능 기술까지 필요하지 않을 수도 있다.

둘째, 자격 심사(eligibility assessment)다. 캐나다 온타리오주에서는 2014년부터 사회부조 운영시스템을 통해 급여 자격을 심사하고 있다. 빅데이터 및 인공지능을 통해 급여 자격 심사를 빠르고 정확하게 처리할 여지가 생긴다.

셋째, 복지급여액 산정 및 지급(welfare benefit calculation and payments)이다. 다수의 국가에서 점점 더 많은 복지급여액이 사람의 개입 없이 자동적으로 산정되고 지급되고 있다. 영국은 실시간 소득정보시스템(Real Time Information System)을 활용해 복지급여를 지급하고 있다.

넷째, 부정·오류 수급 예방 및 탐색(fraud prevention and detection)이다. 많은 복지국가에서 디지털 자료를 활용하는 주요한 이유 가운데 하나가 부정·오류 예방 및 탐색에 있다. 네덜란드에서 논란을 낳았던 SyRI(Systeem Risico Indicatie)가 여기에 해당한다. 참고로, 네덜란드의 SyRI는 중앙정부와 지방정부가 사회보장 영역 등에서 수급자의 부정행위를 방지하고 대처하기 위해 도입한 인공지능 시스템을 가리킨다. 2020년 네덜란드의 법원은 SyRI가 개인정보 처리 과정에서 투명성 및 안정성에 하자가 있다는 이유로 인공지능 시스템이 유럽인권협약을 위반한다고 판결했고, 시스템 작동은 종료됐다(Appelman et al., 2021, p. 263).

다섯째, 위험의 점수화 및 범주화(risk scoring and classification)다. 이는 데이터 자료의 수집과 빅데이터 분석 결과를 바탕으로 개인에게 일정한 점수를 부여하거나 일정한 범주를 부여하는 과정을 가리킨다. 여기에서 고위험 대상자도 도출된다. 한국의 사각지대 발굴관리 시스템이 여기에 해당한다. 이러한 접근은 세계 최초의 인공지능 관련 초국가적 규제인 유럽연합 인공지능법(Artificial Intelligence Act)에서 '수용할 수 없는 위험성(unacceptable risk)'으로 분류한 '사회적 평점(social scoring)'[3]과 흡사하다. 앞으로도 논란을 낳을 수 있는 대목이다.[4]

여섯째, 개인 맞춤형 정보 서비스(personalized information service)이다. 남미 지역에서 특히

동화, ④ 챗봇, 오피스 프로세스 자동화, 부정 수급 포착 영역 등이다. OECD(2024)가 전 세계적으로 사회보장 영역에서 인공지능은 '간헐적(infrequently)'(p. 9)으로 쓰인다고 논평한 점을 보면 OECD는 인공지능이 활용되는 영역을 지나치게 협소하게 파악한 것으로 보인다.

3) 개인의 데이터를 기반으로 개인을 평가하는 것을 의미하며, 여기에는 신용 행태, 교통 위반, 사회적 참여 등이 포함된다. 이러한 평가는 특정 서비스나 특권에 대한 접근을 규제하기 위해 사용된다.(Mosene, 2024)

4) Alston(2019)은 위험의 점수화 및 범주화와 관련해 세 가지 위험을 예시했다. 첫째, 전체 인구 집단의 데이터를 근거로 한 예측 모델에 따라 개인의 위험 수준을 파악하는 과정에서 나타날 수 있는 오류의 가능성, 둘째, 점수화 및 범주화의 근거가 되는 기술이 공개되지 않으면서 나타날 수 있는 권리 침해의 가능성, 셋째, 점수화 및 범주화가 현재의 불평등과 차별을 강화하거나 유지할 가능성이다.

활성화한 챗봇이 대표적 사례가 될 것이다. 다양한 영역에서 개인에게 최적화된 서비스를 제공하기 위해 빅데이터 분석 기반 기술을 활용하여 서비스 이용자에게 맞춤형 서비스를 제공할 수 있다. 한국의 인공지능 활용 초기상담시스템도 여기에 해당될 것이다. 인공지능 활용 초기상담시스템도 콜 기반 대화 시스템을 활용하여 잠재적 위기 대상자와 초기 상담을 진행한다. 이러한 기능은 복지수급자의 사례관리에까지 확장될 수 있다. 사회보장 기관이 인공지능을 사용하여 사례 처리 방식을 자동화하고 개인의 서비스 후속 조치를 위한 고충 대응 지원을 제공할 수 있다. 오스트리아의 사회보험연합은 청구 자동 처리를 지원하고 의사와 매칭하는 인공지능 기반 시스템을 구현했다.

일곱째, 온라인을 활용한 소통을 넘어 실제 돌봄 영역에서도 활용되고 있다.[5] 한국에서도 이미 폭넓게 활용되고 있다. 대부분 노인을 위한 것으로 인공지능과 사물인터넷을 활용한 건강 서비스, 생체 건강 셀프체크 서비스, 안전감지 센서를 활용한 안전 확인 서비스, 실종 방지를 위한 위치 기반 모니터링 서비스 등이다. 여가 활용에 도움이 되는 다양한 정보 제공 서비스 등이 제공되고 있다.

여덟째, 사회보장 행정 기관 내부적인 용도로 업무 담당자의 효율적 업무 처리를 돕고, 내부 교육 등의 용도로도 활용될 수 있다. 한국에서 다수의 복지기술은 업무 담당자의 효율적 업무 처리를 지원하기 위해 활용되고 있었다. 이를테면 빅데이터를 기반으로 의사결정을 효율적으로 지원하는 서비스도 확인되었는데, 수급자 상담을 위한 기초 자료를 제공한다.

아홉째, 사회정책의 효율성과 효과성을 평가하는 데 활용될 수 있다. 데이터의 가용성이 실시간화한다면 "적절한 성과지표의 주기적인 활용은 사업의 효과성과 효율성을 평가하는 데 필수적인 것으로, 증거기반정책(evidence-based policy)의 기초"(유종성, 2023, p. 8)가 될 수 있다. 또한 평가 과정에서 정책의 성과, 실패, 한계를 낳은 원인을 분석해 정책을 조정, 갱신, 폐기하는 근거가 될 수 있다.

장기적으로 보면 전 국민 대상 실시간 데이터에 근거한 인공지능의 활용은 사회보장제도 전체를 재편하는 방향으로 나갈 잠재력을 가지고 있다. 노대명(2024)은 소득보장제도 재구조화를 준비하자고 제안하면서 "소득 기반 사회보험의 실험이 중단됐지만 계속 추진할 필요"(p. 14)가 있으며, "현행 85개 제도를 4~5개 정도로 단순화하는 재정 기반 소득보장제도를 준비"(p. 14)해야 한다고 제안한다. 시민들의 다양한 욕구에 부응하면서 자리잡은 복지제도들이 제도 운영 주체 및 전달 체계에 따라 너무 복잡하게 얽혀 있는 점을 고려할 필요가 있다. 이에 따라 사회보장제도들의 재구조화에 대한 요구가 끊임없이 있었던 점을 고려하면(강혜규 외,

5) 돌봄 영역에서의 인공지능 기술 적용 속도가 빠르고, 분야도 다양하게 나뉘고 있다. 이 글에서는 이러한 상황을 깊게 다루지 못하는 한계를 노출하고 있다.

2022) 변화의 단초가 빅데이터와 인공지능에서 마련될 여지가 있다. 노대명(2024)은 현재의 급여 지급 정보관리 체계에서 복지 제공 부처별로 급여 자격 조건 심사를 따로 추진하면서 급여 지급이 지연되고 있다며 디지털 사회보장 허브 구축을 통해 급여 자격 심사와 지급을 효율화하자고 제안했다.

3 사회보장 영역 인공지능 기술 적용의 순기능

사회보장 영역에서 점점 폭넓게 활용되는 인공지능 기술 적용이 불러올 효과는 양면적이다. 먼저 인공지능이 공공의 민주적인 통제 아래 작동할 경우를 전제로 할 때, 기대되는 순기능은 다음과 같다.

첫째, 효율성이다. 앞서 살펴보았듯이 본인 인증, 자격 심사, 복지급여액 산정 및 지급 등의 일선 서류 행정이 더욱 자동화하게 된다. 이럴 경우 더 적은 인력으로 더 많은 행정 업무가 가능해진다. 일선 공무원은 민원인을 대상으로 하는 대면 서비스나 사례관리에 집중할 수 있다. 물론 인공지능의 적용에 따른 부가적인 업무도 생길 수 있다. 이를테면 미국의 행정명령 14110 7조 2항 (b)호에 따르면 급여 거부에 대해 급여 신청자가 사람인 심사자에게 이의를 제기할 수 있도록 하는 절차를 규정하고 있다.

둘째, 적시성이다. 빅데이터를 활용해 관리되는 인공지능 기술은 급여 신청과 심사, 지급에 이르는 과정을 단순화할 수 있다. 영국의 경우 통합급여를 도입하는 과정에서 실시간 소득파악 시스템을 활용해 급여 지급 절차를 간소화하고자 했다.

셋째, 정확성이다. 빅데이터에 근거한 인공지능의 판단, 범주화, 예측은 인간의 오류와 편견이 개입할 가능성을 줄일 수 있다. 영국의 경우 연금을 제외한 복지급여액 가운데 오류 또는 부정으로 인해 초과 지급된 액수가 2021~2022년 회계연도 기준으로 85억 파운드(약 12조 7000억 원)으로 추정됐다(Davies, 2022). 같은 기간 연금을 제외한 전체 복지급여 지급액의 7.6%를 차지하는 액수다. 특히 서구의 복지국가에서 디지털 기술을 통해 급여의 오류 및 부정 수급을 시정하고자 하는 의지가 강해 보인다.

넷째, 개인 맞춤형 급여 및 서비스 제공이다. 인공지능의 활용을 통한 행정 집행은 개인당 위험의 점수화 및 범주화까지 맞춤형 수준으로 정교화할 수 있다. 개인의 여건과 수요에 맞춘 현금 및 서비스, 일자리가 개인 맞춤형으로 제공될 여지가 커진다. 한국에서 인공지능 기반 일자리 매칭 서비스가 대표적인 예다. 인공지능 기반 일자리 매칭 서비스는 구직자의 이력서 정보와 구인 기업의 구인 정보를 활용하여 인공지능 알고리즘에 기반한 추천 서비스를 제공하고, 기업과 구직자 간 일자리 미스매치를 해소한다. 이를 통해 개인의 제도에 대한 체감도가 향상될 수 있다.

다섯째, 범용성이다. 챗봇을 통한 상담은 국민 입장에서는 시공간의 제약을 뛰어넘어 제도에 접근할 수 있는 기반을 제시해 준다. 시민들은 주민센터

를 방문하지 않고도 언제, 어디서나 급여와 관련한 상담과 더불어 자격을 확인할 수 있고, 급여를 신청할 수 있게 된다. 이를 통해 지역별 접근성에서의 격차를 해소할 수 있는 가능성도 커진다.

여섯째, 정책 평가의 용이성이다. 빅데이터를 활용한 급여 집행 내용은 정책의 효과를 평가하는 데 용이한 기반을 제공한다. 평가를 위한 데이터의 구축이 용이해지면서 근거 기반 정책 평가, 형성 및 집행의 토대가 마련된다.

일곱째, 사각지대 해소다. 복지 사각지대 발굴관리 시스템이나 인공지능 활용 초기상담시스템은 다중 위험의 시대에서 잠재적 수급 대상자들의 욕구를 빠르게 파악하는 토대가 된다. 한국 복지국가에서 고질적으로 지적되는 사각지대의 문제를 빠르게 해소할 수 있는 제도적 기반이 될 수 있다. 물론 사각지대 문제의 핵심은 위기 가구의 발굴에 있다기보다는 발굴된 위기 가구를 지원할 급여가 없거나 부족하다는 지적(함영진 외, 2023)도 함께 고려할 필요가 있다.

4 사회보장 영역 인공지능 기술 적용의 위험성

빅데이터 활용을 수반하는 인공지능 기술이 사회보장 영역에서 초래할 위험성도 함께 보겠다. 위험의 내용은 김수영(2016)과 김기태(2024)를 참조하면서 다음과 같이 제시한다.

첫째, 프라이버시의 문제다. 한국의 경우 국민기초생활보장제도의 수급자는 자신과 부양의무자의 소득 및 재산 정보를 국가에 제공하는 조건 아래 급여를 받을 수 있다. 복지급여 수급자들의 소득, 재산, 가족 정보뿐 아니라 일부 사례관리 정보까지 이미 방대하게 집적돼 있다. 공공기관을 넘나드는 개인정보의 유출 및 남용 가능성도 완전히 배제할 수는 없다. 또한 국가 권력에 의한 데이터 남용 가능성도 고려해야 한다(홍승헌, 황하, 2024).

둘째, 부정확성의 문제다. 앞서 인공지능이 불러올 장점으로 정확성을 제시한 점을 고려하면 다소 모순적인 지적일 수 있다. 그럼에도 국내의 사회보장정보시스템에서 개인의 사망 및 출생 신고가 반영되지 않거나 과거 소득이 현재 소득과 합산돼 제시되는 등의 문제가 끊임없이 발생하고 있다. 현장 공무원들은 복잡한 데이터를 처리하고 오류를 수정하느라 대면 접촉을 통해 사각지대를 발굴할 기회를 오히려 놓치고 있다는 지적(임덕영, 남윤재, 2023)도 나온다. 물론 이러한 문제는 데이터 기반 인공지능 기술이 현장에 접목되는 과정에서 발생하는 일시적이고 과도기적인 문제라고 볼 수 있다. 그러나 인공지능 기술에 대한 맹신적인 태도 역시 경계할 필요는 있다. 이 대목은 사회보장 행정에서 관리하는 데이터와 알고리즘의 질 관리와도 직결되는 문제다.

셋째, 데이터 소유권의 문제다. 김수영(2016)은 급여 신청자는 소득, 재산, 가족정보, 신체 등에 대한 자기 정보 제공에 동의하게 되는데, 여기서 정보의 소유권에 대한 문제가 제기된다고 지적한다. “국

가는 정보 제공자의 의도와는 상관없이 이들이 제공한 정보를 보유할 뿐만 아니라 향후 해당 정보의 활용, 분배, 처분에 관한 권한을 얻게 되는 상황." (p. 211) 이 대목은 앞서 논의한 프라이버시와 함께 앞으로 논점으로 부상할 여지가 크다.

넷째, 개인정보의 영리 목적 활용 문제다. '디지털헬스'를 둘러싼 데이터 활용은 한국에서 이미 오랜 의제다. 건강보험의 개인정보에 대한 접근권을 민간기업은 지속적으로 요구하고 있다. 보건복지부는 2024년 11월 '보건의료 데이터 혁신포럼'을 전국 43개 의료 데이터 중심 병원과 진행한 바 있다. 복지부는 "정부와 의료계, 학계, 산업계가 데이터 협력 체계를 구축하여 미래 의료 혁신과 국민 건강 증진으로 나아가는 계기가 되기를 희망"(보건복지부, 2024. 11. 26. p. 4)한다고 밝혔다. 여기에서 개인 건강 정보의 영리 목적 활용의 여지는 잠재적으로 남아 있다.

다섯째, 알고리즘 결정에 근거한 개입의 문제다. 덴마크는 취약 아동을 포착하기 위해 실업 및 의료를 비롯한 다양한 사회경제적 영역의 데이터를 결합하여 200개 이상의 위험 지표를 분석하는 머신러닝 모델을 구축했다(Jørgensen, 2021). 그리고 모델이 위험 신호를 보내는 가구에 대해 부모의 동의 없이 개입이 가능하도록 했다. 이후 해당 모델은 신뢰성에 의문이 제기되면서 2018년 12월 중단된 바 있다(Algorithm Watch & Bertelsmann Stiftung, 2020). 한국에서도 2023년부터 44종의 위기 정보에 근거해 위기 가구에 대해서는 경찰·소방 협력을 통한 비상 개문(開門) 지침을 마련했다(정세정 외, 2023). 이 대목에서 빅데이터에 근거한 인공지능의 판단에 따라 국가는 어디까지 개인의 삶에 개입할 수 있는가라는 윤리적인 문제가 발생한다.

여섯째, 데이터와 알고리즘의 편향성의 문제다. 미국의 데이터 과학자인 Cathy O'Neil(2017)은 알고리즘이 현대 사회에서 대량살상무기(Weapons of Mass Destruction)에 준하는 위험성을 가졌다고 경고하기 위해 책 '대량살상수학무기(Weapons of Math Destruction)'를 출판한 바 있다. 그는 여기에서 편향된 데이터와 알고리즘이 취약계층에 차별적인 결과를 낳게 된다고 경고했다. 그는 알고리즘에 대한 기술 낙관론을 '테크노유토피아'로 명명하면서 "알고리즘과 기술이 가져다줄 혜택에 대한 무제한적이고 부적절한 희망에서 깨어나야 한다"(p. 343)고 강조한다.

일곱째, 인공지능의 '설명 불가능성' 문제다. 인공지능의 복잡성을 고려할 때, 인공지능이 어떤 기준과 원칙에 따라 작동하는지에 대해 인간은 이해할 수 있어야 한다. 이는 인공지능 관련 핵심 원칙 가운데 하나인 설명 가능성(explainability)이다. 문제는 인공지능이 점점 더 설명 가능하지 않다는 점이다. 인공지능이 고도화할수록 설명 가능성은 떨어진다. [그림 1]을 보면 인공지능 모델의 작동이 고도화할수록 모델이 산출하는 결과에 대한 근거를 인간이 파악하기 어렵다. 고성능 인공지능은 왜 그러한 결과값을 산출했는지에 대해 인간에게 근거나

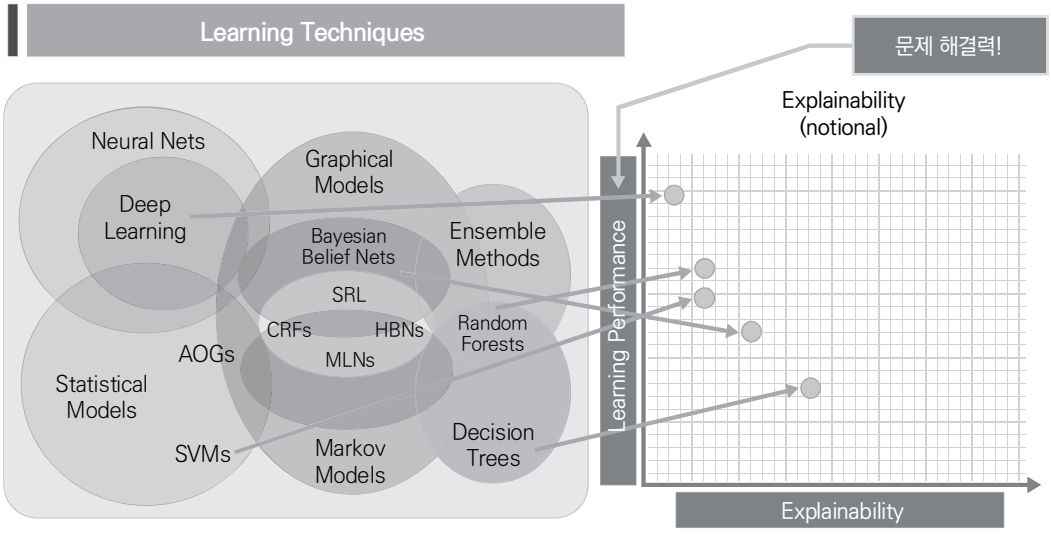
이유를 제시하지 않게 된다. 결국 인간은 근거가 없는 인공지능의 판단 혹은 설명을 어디까지, 무슨 근거로 신뢰해야 할지에 대한 딜레마에 빠질 수 있다.

물론 사회보장 분야에서는 인공지능이 활용된다고 하더라도 딥러닝 수준까지 활용되는 경우는 많지 않다. 따라서 인공지능의 설명 가능성에 대한 문제가 적어도 한동안은 중요하지 않을 수도 있다. 그렇지만 기술의 발전 속도를 고려한다면 인공지능의 설명 불가능성이 조만간 난제로 등장할 가능성도 있다. 사회보장 행정에서는 특히 취약계층에 대한 차별적인 결정이 이뤄질 가능성도 염두에 둬야 한다.

지금까지 살펴본 인공지능의 잠재적 위험성을 살펴봤다. 다만 국외에서 제기되는 위험성을 고려할 때, 한국과 외국의 제도적 환경이 다르다는 점도 염두에 둘 필요가 있다. 정세정 외(2023)에서는 세 가지로 나누어서 설명하고 있다(pp. 213-214).

첫째, 한국에서는 지문 및 얼굴 정보가 포함된 전 국민 주민등록 데이터베이스 덕분에 개인 인증 영역에서 압도적인 인프라를 구축했다. 전 국민의 일률적인 국민 식별 번호를 운영하는 국가는 한국이 거의 유일하다(성준호, 2016). 둘째, 해외에서는 개인정보의 영리적 집적 및 활용에 대한 우려가 크다(Eubanks, 2018; Alston, 2019). 미국에서는 민간 보험회사에서 개인 데이터를 활용하는 빈도가 더 높다. 반면 한국은 공공이 전 국민 소득, 재산, 건강, 인적 데이터를 주도해 관리하고 있다. 셋째, 다른 복지국가에서는 복지 제도와 성숙화 과정에서 디지털 기술을 급여의 부정 및 오류 수급 포착에 활

[그림 1] 인공지능 윤리 관련 의제와 원칙



출처: "인공지능 윤리 핵심 가치 분석: 한국 사례를 중심으로"[발표자료], 김형주, 2024, 인공지능 발전과 공공의 과제 세미나, p. 5.

용하는 경향이 있다. 서구가 보편적 복지국가의 성숙기를 지난 이후에 나타나는 문제에 대응하는 과정에 있다면 한국은 보편적 복지국가와는 거리가 멀다. 한국에서는 디지털 기술에 대해서도 초점이 사각지대 해소 혹은 위기 가구 포착에 있다. 사회보장 영역에서 인공지능의 활용과 규제를 모색할 때, 이러한 한국적 특수성도 함께 고려할 필요가 있다.

5 정책 제언

지금까지 우리는 인공지능 기술이 사회보장 영역에 불러올 수 있는 순기능과 위험성을 두루 확인했다. 앞으로의 정책 방향은 자연스럽게 도출된다. 위험성을 최소한으로 관리하면서 동시에 순기능을 최대화하는 것이다. 리스크 가능성을 규제하고, 편의를 증진하는 방향으로의 지원 사이에서 적절한 정책적 배합이 중요하다는 의미다. 여기에서 신기술에 대한 규제와 지원을 길항 혹은 모순 관계로 볼 필요는 없다. 오히려 지원의 전제는 규제이고, 규제의 이유는 지원이다. 규제와 지원 사이의 동적인 관계에 대해 김병권(2024. 12. 26.)의 다소 긴 언급을 들어 보자.

(정부의 인공지능 지원 편향적 정책에 대해) "우리는 얼른 인공지능을 경쟁력 있게 개발해야 하니 당연한 거 아니냐는 분들도 많다. 그러나 누차 말하지만, 식품과 의약품은 정확한 안전성 규제를 해야 소비자들이 맘 놓고 구입하여 먹을 수 있고, 이럴 때 진짜 식품산업과 제약업이 발전한다. 지금처럼 우리가 항공기를 자유롭게 타는 것은 엄청나게 까다로운 항공기 운항 규제가 생겼기 때문이다. 그 이전에는 워낙 사고가 많아 비행기는 모험가들만의 이용에 국한되었다. 지금 지구상에 15억 대가 넘는 자동차가 굴러다니는 것은 오직 엄격한 '교통법규'와 '자동차 안전성 규제' 때문이다. 인공지능도 잘 사용하면 좋지만, 위험성도 함께 있기에 '제대로 안전성을 엄격히 규제'해야 번성할 수 있다."

국제사회에서 인공지능에 대한 대표적인 규제 사례에 해당하는 유럽연합과 미국 행정명령 14110도 인공지능에 대한 규제와 지원 사이에서 일정한 균형을 갖췄다. 유럽연합 인공지능법은 "이 규정은 공공행정이 준수되고 안전한 인공지능 시스템의 광범위한 사용을 통해 이익을 얻을 수 있도록 혁신적인 접근 방식의 개발과 사용을 저해해서는 안 된다"(EU, 2024)라고 명기하고 있다. 미국 행정명령도 7조 2항 (b)호에서 보건복지부는 급여 및 서비스를 시행할 때 자동화 또는 알고리즘 시스템의 사용을 촉진하는 계획을 행정명령 시점 기준으로 180일 이내에 발표해야 한다고 규정했다.

안타깝게도 한국에서는 사회보장 영역의 인공지능 적용에 관한 공적인 비전 혹은 계획이 제시된 바 없다. 물론 2024년 12월 26일 한국에서 인공지능기본법이 국회 본회의를 통과한 상황을 고려할 필요는 있다. 그러나 한국의 인공지능기본법에서는 사회보장 영역에서의 인공지능 기술 적용을 촉진하거나 규제하는 내용이 반영되지 않았다. 유럽연합의 인공지능법과 미국의 행정명령에서 사회보장 영

역을 비중 있게 다룬 점과 대조된다. 한국의 사회보장 영역에서도 인공지능 기술 적용이 폭넓게 활용되고 있다. 한국의 인공지능기본법이 사회보장 영역을 간과한 점은 의아한 대목이다. 법은 처음부터 이렇게 중요한 결함을 가지고 있다. 앞으로라도 사회보장 영역을 포괄하는 법 개정을 시도하거나, 사회보장 영역에서의 인공지능 적용 및 규제에 대한 별도의 입법을 시도하는 방안을 검토할 수 있을 것이다.

정부의 행정 영역에서 사회보장 분야가 차지하는 비중을 고려하면 더욱 그러하다. 미국의 경우 연방정부에서 인공지능을 활용하는 사례 가운데 3분의 1 이상이 보건복지부 업무 영역에 속한다(Burt, 2024. 10. 17.). 미국의 행정명령에서 지속적으로 '보건복지부'가 호명되는 이유다. 한국의 정부 지출 가운데 보건·복지·고용 영역 비율이 가장 높은 점은 주지의 사실이다. 2024년 기준으로 보면 전체 정부 예산 656조 6000억 원 가운데 보건·복지·고용 분야가 237조 6000억 원으로, 총지출의 36.1%를 차지하는데, 비율은 앞으로도 증가할 것으로 예상된다(기획재정부, 2024). 이와 같은 점을 염두에 두고 빅데이터와 인공지능의 사회보장 영역 적용을 촉진하고 부작용을 선제적으로 예방하기 위한 정책 제언을 하도록 하겠다.

첫째, 사회보장 정보에 활용되는 데이터 품질을 개선하기 위한 정책적 노력이 필요하다. 인공지능의 데이터 누적 과정에서 쓰레기 데이터가 쓰레기 출력값을 산출하는(garbage in, garbage out) 문제가 제시된다(James, 2024). 데이터와 데이터에 근거한 인공지능의 결과값에 대한 신뢰도를 높이기 위해서는 양질의 데이터 확보 및 집적, 관리가 필요하다. "사회보장제도에 대한 공정한 평가와 이를 기반으로 한 사회보장제도의 설계를 위해서는 정확한 사회보장제도 관련 정보가 필요하다. 따라서 정보 포괄성을 바탕으로 하는 정확한 데이터 품질 관리가 요구된다."(이상원, 2023, p. 3) 공적 데이터를 관리하는 기관들에 대한 인적, 행정적 인프라 개선 및 확충이 필요함을 의미한다.

둘째, 사회보장 영역에서의 데이터 통합, 연계, 관리를 위한 노력이 필요하다. 사회보장 영역에서는 2022년부터 사회보장기본법 42조 1에 근거해 사회보장 정책의 심의·조정과 관련 연구를 위한 행정 데이터가 수집되고 있다. 한국에서는 모든 개인에게 부여되는 주민등록번호로 기관들의 데이터를 연계할 수 있는 강력한 여건을 구축하고 있다. 그렇지만 데이터 연계 과정에서 주민등록번호를 결합키 생성에 사용하지 못하도록 하는 법적인 규제가 발목을 잡고 있다. 더욱이 대부분의 정부 부처와 공공기관들도 데이터 공개에 소극적이다. 북유럽, 미국, 영국, 캐나다 등에서 행정 자료 구축과 공개, 연계에 적극적인 점(유종성, 2023)을 참고할 필요가 있다.

셋째, 다양한 기관이 집적한 정보들의 표준화 및 단순화가 필요하다(노대명, 2024). 이는 앞선 데이터 질 개선 및 연계와 직결되는 문제다. 노대명(2024)은 사회보장정보시스템의 많은 데이터가 제대로 활용되고 있지 않다며 사회보장정보원의 데이

터 테이블은 1만 1280개에 달하지만, 정작 자주 활용되는 정보는 500칼럼에 불과하다고 지적한다. 다른 부처와 기관이 생산하는 데이터들 역시 통합 은커녕 연계가 힘든 상황이다. 현재는 구슬은 서 말이지만, 이들을 꿰지는 못하고 있다.

넷째, 데이터 구축, 연계, 활용 과정에서 데이터 보안 및 안전에 대한 엄격한 기준 설정이 반드시 필요하다. 유종성(2023)은 행정 데이터 활용에서의 향후 과제를 제안하면서 다음과 같이 언급했다. "전수를 포괄하는 빅데이터인 경우가 많아 개인정보 보호에 보다 세심한 주의가 필요하다. 개인정보 보호에 대한 안전장치를 이중 삼중으로 설치할 필요가 있다. 과거 스웨덴 등 북유럽 국가들에서도 데이터 활용과 개인정보 보호를 둘러싼 논쟁이 있었다. 개인정보 보호를 철저히 하면서 데이터를 연계하고 가명 처리된 데이터를 연구자들이 안전하게 사용할 수 있는 방법들이 발전되어 왔다."(p. 15) 소득, 재산, 건강, 가족 등 개인정보가 결합될수록 데이터 유출에 따른 충격은 커질 수밖에 없다. 이 문제는 데이터 보안을 위한 데이터 거버넌스의 문제와도 직결된다.

다섯째, 데이터 관리를 넘어 이를 활용하는 알고리즘 및 인공지능 시스템의 편향성을 최소화하면서 정확성을 개선하기 위한 노력이 반드시 필요하다. 과거 네덜란드, 덴마크, 호주에서 인공지능 활용을 섣불리 시도하다가 사회적 논란을 낳았던 사례를 복기할 필요가 있다(정세정 외, 2023). 인공지능을 활용한 행정에서 가장 흔하게 지목되는 리스크가 편견과 불평등의 심화(O'neil, 2017; Eubanks, 2018)라는 점도 상기할 필요가 있다. 인공지능 기반 행정에 대한 신뢰는 과학성, 중립성에 대한 입증을 통해 조금씩 누적될 수 있다. 그리고 그 신뢰가 무너지는 것도 한순간이다.

여섯째, 한국의 정부 부처에서, 특히 사회보장 영역에 한정하자면 보건복지부에서 인공지능 관련 조직과 인력을 신설 및 배치하는 안을 검토할 필요가 있다. 유럽연합의 인공지능법에 따르면 고위험성 인공지능 시스템[6] 제공자는 17개의 의무를 지게 된다(윤혜선, 2024).[7] 물론 유럽연합의 다소 엄격한 기준이 그대로 한국에 적용될 것으로 확언하기는 어렵다. 미국에서조차 보건복지부는 다양한 이행 조치를 담은 행정명령을 강요받았다. 이를테면 'Office of the Chief Artificial Intelligence Officer(OCAIO)'를 임명하고, 'HHS Artificial

6) 2024년 제정된 유럽연합 인공지능기본법(EU AI Act)은 인간의 안전, 건강, 기본권, 생계 등에 중대한 영향을 미칠 수 있는 분야에서 사용하는 인공지능 시스템을 '고위험성(high-risk)'으로 분류한다. 여기에는 주로 생체인식, 교육, 노동시장 접근, 사회보장 급여 결정, 법 집행, 출입국 심사, 사법 분야 등에서 사용되는 시스템이 포함된다. 등록, 투명성, 위험관리, 데이터 품질, 인간 감독 등의 엄격한 의무가 부과된다.

7) ① 고위험성 인공지능 시스템 준수 사항 준수 여부 확인, ② 고위험성 인공지능 시스템에 연락처 표시, ③ 품질 관리 시스템 준수, ④ 기술 문서, 품질관리 시스템 문서 등 필수 자료 보관, ⑤ 통제하에 있는 경우 자동으로 생성된 로그 보관, ⑥ 시장 출시·서비스에 투입되기 전 적합성 평가 절차 준수, ⑦ 유럽연합 적합성 선언 작성, ⑧ CE 마크 부착, ⑨ 등록 의무 준수, ⑩ 필요한 시정 조치(규정 위반 시 적절한 조치) 및 정보 제공, ⑪ 국가 관할 기관의 합리적인 요청이 있을 때, 고위험성 인공지능 시스템이 모든 준수 사항에 부합함을 입증, ⑫ 제품 및 서비스에 대한 접근성 의무 보장, ⑬ 관할 당국과의 협력, ⑭ 중요 내용 변경 시 적합성 평가, ⑮ 출시 후 모니터링 체계 설계·구축·운영, ⑯ 중대 사고 관련 정보 공유 및 신고 의무, ⑰ 인공지능 리터러시 보장(윤혜선, 2024).

Intelligence(AI) Strategy'를 공표해야 했다. 한국 보건복지부는 시스템 제공자 혹은 제공자의 관리감독 기관으로서 해당 조치들을 적극 검토할 필요가 있다. 지금까지 보건복지부에서는 보건 분야의 일부를 제외하고는 사회보장 영역의 인공지능 관련 규제 조치를 내놓은 바가 없다.

일곱째, 국제적인 인공지능 발전과 규제 동향에 대한 모니터링이 필요하다. 유럽연합의 인공지능법은 법 집행 과정을 거치면서 규제의 내용이 구체화할 가능성이 높다. 미국의 경우 트럼프 행정부 집권 이후 바이든 정부의 행정명령이 폐지되면서 또 다른 격변을 거치고 있다. 앞으로 인공지능 발전의 속도, 발전이 사회에 미치는 파장, 인공지능 패권을 둘러싼 미·중·유럽권의 경쟁, 국제기구의 개입 등이 복잡하고 역동적으로 작용할 것으로 예상된다. 무엇보다 중요한 변수는 정치인과 정책 전문가들마저도 인공지능이 정확히 무엇인지, 그래서 인공지능이 미칠 파장이 어떠할지에 대해 잘 모르고 있다는 점이다. 놀라운 일은 아니다. 심지어 기술 전문가들 사이에서도 인공지능의 영향에 대해 비관론자(doomer)와 낙관론자(boomer)로 엇갈리고 있는 점을 상기할 필요가 있다. 물론 미지(未知)보다 해소되기 어려운 문제는 상충하는 이해관계라는 점도 확인해 둔다. 한국의 사회보장 영역에서 인공지능의 적용을 지원하고 규제하고자 한다면 이러한 내외 환경에 대한 동적인 모니터링이 필요할 수밖에 없다.

여덟째, 지금까지 제시한 데이터 관리, 연계, 표준화 및 알고리즘 질 관리 등을 총괄하는 거버넌스 체계 구축이 필요하다. 한국에서 행정 데이터 활용에 대한 논의는 데이터 활용도를 높이는 방향을 중심으로 논의되는 경향이 있다. 데이터를 활용한 복지행정에서의 오류 혹은 사고는 한 번만 발생해도 전체 시스템을 폐쇄할 정도로 충격이 큰 점을 고려할 필요가 있다(홍승헌, 황하, 2024). 관련해서 2024년 국회 본회의를 통과한 인공지능기본법에서는 규제의 주체로 과학기술정보통신부를 호명하고, 독립적인 규제 및 감독 기구를 구체적으로 제시하지는 않은 것으로 보인다. 지금까지 공개된 초안에서는 인공지능 사업자나 대통령령으로 정하는 인공지능 기술 관련 기관이 '민간자율인공지능윤리위원회'를 '둘 수 있다'는 느슨한 규제를 제시하고 있다. 한국의 인공지능기본법이 진흥과 규제 사이에서 진흥 쪽으로 과하게 경도됐다고 보는 것이 적절할 것이다. 김병권(2024. 12. 26.)이 지적한 대로 새로운 법은 규제법은 아니고 그냥 '진흥법'이다.

앞서 언급한 대로 제대로 된 규제 없이 인공지능은 발전할 수 없다. 사회보장 영역은 사람을 대상으로, 특히 빈곤, 아동, 노인 등 취약계층을 대상으로 한다. 그리고 이들의 사적인 소득, 건강, 재산 자료에 근거해 정책을 편다. 유럽연합에서 금지하는 '사회적 평점 부여(social scoring)'에 가깝다. 그래서 시장 친화적인 인공지능 정책을 펴는 미국에서도 지난 정권 때 보건복지부에 대해서는 보다 상세한 지침을 제시했다.

6 나가며

한국 공공 영역의 디지털 기술, 인공지능 기술 도입은 빠른 반면 관련 도덕적, 법적 쟁점에 대한 사회적 논의 및 규제 형성 과정은 매우 더디다. 사회보장 영역의 인공지능 기반 기술 도입 속도는 다른 복지국가들에 견줘 전혀 늦지 않다. 이러한 점을 고려한 디지털 거버넌스 구축이 필요하다. 관련해서 유도진(2024. 12. 27.)은 인공지능의 민주적 활용을 위해 세 가지의 제도적 기구 설치를 제안했다. 첫째, 인공지능 활용 정책과 데이터 사용의 적법성, 투명성을 감사하는 독립기구다. 여기에는 인권, 기술, 법률 등 다양한 전문가가 참여한다. 둘째, 정부의 인공지능 활용에 대한 시민의 감시와 통제를 위한 옴부즈맨 프로그램이다. 신고 절차를 간소화해 시민 참여를 확대하는 식의 접근을 제안했다. 셋째, 인공지능 시스템의 알고리즘 공정성, 데이터 처리 투명성, 윤리적 설계를 평가하고 인증하는 독립 인증 기구다. 여기에서는 기술적·윤리적 위험 요소를 정기적으로 점검하고, 문제 발생 시 시정을 요구할 수 있는 권한도 가지게 된다.

종합하면 사회보장 행정에서 인공지능 기술을 적용할 때의 거버넌스는 '인간 중심적 활용'이라는 원칙 아래 ① 민주적 통제, ② 시민 참여, ③ 프라이버시 보장을 전제로 해야 하고, 동시에 시민의 적극적 사회권 보호라는 원칙 아래 ① 사각지대 해소, ② 행정 효율성 제고, ③ 급여 지급의 정확성과 적시성 보장 등을 추구하는 정책 방향을 확인할 필요가 있다.

참고문헌

강혜규, 강지원, 강희정, 김기태, 김세진, 김태완, 주하나. (2022). **사회보장제도 실태분석 및 개선방안 연구**. 보건복지부, 한국보건사회연구원.

기획재정부. (2024). **2024~2028년 국가재정운용계획 주요내용**.

김기태. (2024). 디지털 복지국가의 개념과 논점. **비판과 대안을 위한 사회복지학회 학술대회 발표논문집** (pp. 133-145).

김기태, 김명주, 김은하, 신영규, 변소연. (2024). **사회보장행정에서의 인공지능 적용 동향 및 함의**. 한국보건사회연구원.

김명주. (2024). **AI 윤리와 규제** [발표자료]. 한국보건사회연구원 사회보장적용에서의 인공지능 적용 동향과 함의 세미나 서울.

김병권. (2024. 12. 26.). **한국의 처참한 AI법**. Facebook. https://www.facebook.com/byoungkweon.kim

김수영. (2016). 사회복지정보화의 윤리적 쟁점-사회보장정보시스템을 통한 데이터감시를 중심으로-. **한국사복지학. 68**(1), 193-224.

김형주. (2024). **인공지능 윤리 핵심 가치 분석: 한국 사례를 중심으로** [발표자료]. 인공지능 발전과 공공의 과제 세미나, 세종.

노대명. (2024). **디지털 전환과 한국 사회보장제도의 다가오는 과제** [발표자료]. 한국보건사회연구원 빈곤불평등연구실 세미나, 세종.

보건복지부. (2024. 11. 26.). **보건의료데이터와 인공지능이 열어가는 디지털 헬스케어 미래** [보도자료].

성준호. (2016). 주민등록번호에 의존한 본인확인제도의 문제점: 각국의 개인식별번호제도 및 관련 법률의 검토를 통한 시사점. **공공사회연구, 6**(2), 208-246.

유도진. (2024. 12. 27.). AI 기본법, 기술과 민주주의의 균형, 해결해야 할 과제들. **Ohmynews**. https://www.ohmynews.com/NWS_Web/View/at_pg.aspx?CNTN_CD=A0003092112&CMPT_CD=P0010&utm_source=naver&utm_medium=newsearch&utm_campaign=naver_news

유종성. (2023). 사회보장 행정데이터 활용사례와 향후 과제. **보건복지포럼, 325**, 6-19.

윤혜선. (2024). **EU 인공지능법의 주요 내용과 함의**. [발표 자료]. 사회보장적용에서의 인공지능 적용 동향과 함의 세미나. 한국보건사회연구원. 서울.

이상원. (2023). 사회보장 분야 행정데이터의 활용과 전망. **보건복지포럼, 325**, 2-4.

임덕영, 남윤재. (2023). **사회보장 현장 모니터링: 현장 전문가와 실무자 포럼**. 한국보건사회연구원.

정세정, 김기태, 곽윤경, 우선희, 최준영, 이영수. (2023). **한국 복지국가의 재구조화를 위한 연구 - Ⅰ. 디지털 복지국가의 딜레마**. 한국보건사회연구원.

함영진, 이현주, 어유경, 김가희, 박성준, 조용찬, 이영글, 문용필, 오민수. (2023). **복지 전달체계 혁신을 위한 대안적 고찰: 취약계층 발굴정책 개선을 중심으로**. 한국보건사회연구원.

홍승헌, 황하. (2024). 누구를 위한 디지털 전환인가?: 자동화된 복지행정의 위험성. **정부학연구, 30**(2), 61-84.

Algorithm Watch & Bertelsmann Stiftung. (2020) *Automating Society Report 2020*. https://automatingsociety.algorithmwatch.org

Alston, P. (2019). *Digital Welfare States and Human Rights; Report of the Special Rapporteur on Extreme Poverty and Human Rights*. https://undocs.org/A/74/493

Appelman, N., Fathaigh, R. O., & van Hoboken, J. (2021). Social welfare, risk profiling and fundamental rights: The case of SyRI in the Netherlands. *J. Intell. Prop. Info. Tech. & Elec. Com. L., 12*, 257-271.

Burt. C. (2024. 10. 17.). HHS Strategic Plan for AI Coming Soon. *ExecutiveGov*. https://executivegov.com/2024/10/department-of-health-and-human-services-strategic-ai-plan/

Davies, G. (2022). *Report on Accounts: Department for Work & Pensions*. National Audit Office.

EU. (2024). EU AI Act: first regulation on artificial intelligence. [online]. https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

James, G. (2024). Concerns over AI data quality gives new meaning to the phrase "garbage in, garbage out". *SC World*. https://www.scworld.com/perspective/concerns-over-ai-data-quality-gives-new-meaning-to-the-phrase-garbage-in-garbage-out

Jørgensen, R. F. (2021). Data and rights in the digital welfare state: the case of Denmark. *Information, Communication & Society*, 1-16.

Mosene, K. (2024). One step forward, two steps

back: Why artificial intelligence is currently mainly predicting the past. *Digital Society Blog, Humboldt Institut für Internet und Gesellschaft*. https://www.hiig.de/en/why-ai-is-currently-mainly-predicting-the-past/

OECD. (2024). *Modernizing Access to Social Protection*. https://doi.org/10.1787/af31746d-en

O'Neil, C. (2017). **대량살상수학무기**. (김정혜 옮김). 흐름출판. (Original work published 2016)

Zaber, M., Casu, O., Brodersohn, E. (2024). *Artificial Intelligence in social security organizations*. International Social Security Association.

Assessment of the Positive Impacts and Potential Risks of AI Applications in Social Security

Kim, Ki-tae

(Korea Institute for Health and Social Affairs)

In this article, I provide a comprehensive overview of the benefits and risks of AI applications in the area of social security and discuss their policy implications. AI has been applied across various dimensions of social security administration in Korea and elsewhere—notably in identity verification, eligibility screening, benefit calculation, fraud detection, risk classification, personalized services, case management, administrative process optimization, and policy evaluation. Using AI in these areas has proved beneficial in increasing administrative efficiency, improving timeliness and accuracy of benefit payments, reducing unmet needs, tailoring services to individual needs, and establishing databases for policy evaluation. However, AI applications also carry risks involving personal information breaches, data bias, the opacity of algorithmic decision-making, potential commercial exploitation of information, and the problem of explainability. The implications drawn point to the need to improve data quality, promote standardization, establish integrated governance, revise legal frameworks to include provisions on social security, create specialized organizations, and monitor international trends.