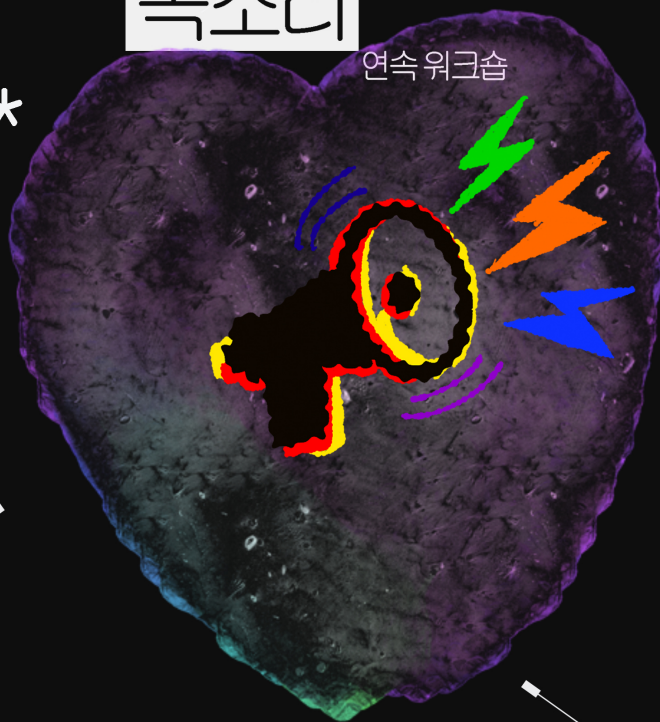


사회복지 분야 인공지능

도입 현황과 과제

영향 받는 사람의 목소리

연속 워크숍



일시 2026년 9월 2일(수) 오후 2시

장소 서울 청파동 아랫마을(용산구 청파로 320-28) - 온라인 생중계 병행

사회 회우 (디지털정의네트워크)

발제 윤영(빈곤사회연대)
초록(전국장애인차별철폐연대)

토론 장욱(홀리스행동)
박영아 변호사(공익인권법재단 공감)
조인영 변호사(공익인권법재단 공감)
권진 교수(예명대학원대학교 사회복지학과)

후원 아름다운재단

주최 인공지능 책임성과
공공성 강화를 위한
시민사회 공동행동

주관

공익인권법재단 공감, 디지털정의네트워크,
빈곤사회연대, 전국장애인차별철폐연대,
정보인권연구소, 홀리스행동

사회복지분야 인공지능 도입현황과 과제

_행복이음과 위기가구 발굴제도를 중심으로¹⁾

빈곤사회연대 김윤영
2026.08.28.

1. 들어가며

호주 정부는 2016년 로보데트 시스템(Robodebt Scheme)을 도입했다.²⁾ 이는 복지수급자 개인이 신고한 소득과 고용주가 국세청에 신고한 소득자료가 일치하는지 확인하고, 과다 지급된 복지수당이 있을 때 이를 환수하는 조치를 통고하는 자동 시스템이었다.

로보데트 시스템이 도입된 배경에는 복지제도가 낭비적으로 이용되고 있다는 정치적 선전이 있었다. 2013년 9월 총선에서 자유국민연합(Liberal-National Coalition)은 노동당 정부 아래 방만한 지출과 재정낭비가 이뤄졌다고 주장하면서 예산 통제와 정부부채 감소를 핵심 공약으로 걸고 승리했다. 2014년 7월 복지수당을 과도하게 지급받은 사람들을 대상으로 이를 환수하는 전략을 수립하고, 2015년 2월 사회서비스부 장관에 임명된 스콧 모리스 장관은 “호주인들은 사회보장시스템을 악용하려는 사람들을 용납하지 않을 것”이라고 선언했다.

로보데트 시스템이 복지수당 과다 수급자를 찾기 위해 이용한 정보는 복지부의 센터링크(Centerlink)에 복지수급자들이 직접 신고하는 2주간의 소득데이터이고, 다른 하나는 국세청이 가진 납세자의 연간 개인소득이었다. 로보데트는 데이터매칭 알고리즘을 통해 둘을 매칭하고, 과다 수급자를 가려내 이들에게 부채환수를 통지했다. 2016년 도입된 로보데트 시스템은 2010년의 복지수당 지급액부터 검토를 시작했다. 이를 통해 2010년부터 2012년까지 3년간 지급한 복지수당 지급 중 본인이 센터링크에 신고한 수입과 국세청이 파악한 소득이 불일치하는 866,857건을 찾아내 이 불일치를 설명할 것을 요구했다. 이를 설명하지 못하는 경우 로보데트는 국세청의 연소득 데이터를 26으로 나누어 2주간 소득 데이터의 평균값을 구한 후 이를 센터링크에 보고한 2주 소득의 데이터와 비교해 각 복지수급자가 정부에 진 부채를 구했다.

이러한 시스템에는 오류가 있다. 개인 소득이 연중 고르게 발생한다면 소득 평균화를 적용하는 것이 옳지만, 비정규직이나 일용직에 종사해 수입이 불규칙한 복지수급자의 경우 소득이 고르게 발생하지 않기 때문이다.³⁾ 결국 호주정부는 47만 건, 약 10억 호주달러에 이르는 부

1) 이 글은 월간 복지동향 326호(2025년 12월)에 실린 빈곤사회연대 김윤영의 글 <사회보장제도와 디지털 전환: 빈곤층 권리 보장 측면에서의 경향성과 대응과제>에 썼던 글이 재인용되었습니다.

2) 로보데트 시스템에 대한 이하 문단은 홍승헌, 황하, 「누구를 위한 디지털 전환인가? 자동화된 복지행정의 위험성」, 정부학연구 제30권 제2호(2024)를 참조

3) “로보데트 시스템에 대한 가장 중요한 비판은 이렇게 계산된 부채환수통보(debt notice)가 잘못된 계산일 수 있다는 점이다. 실업수당의 경우를 예로 들어보자. 실업수당 대상자 A씨는 1년간 4주를 일하였고, 일하는 동안 시급 25달러를 받아 총 4,000달러의 임금소득이 발생(25달러×주40시간×4주)하였다고 가정해보자. A씨는 소득이 발생하지 않은 48주간은 2주에 1,500달러의 실업수당을 지급받았으며, 임금소득이 발생한 4주간은 삭감된 금액으로 실업수당을 지급받았다. 로보데트 시스템이 도입되기 이전에는 소득이 발생한 4주간 50%가 삭감된 1,500달러만을 수령하게 되어 1년동안 받아야 할

채통보가 잘못 고지되었다는 점을 인정하고, 시스템을 폐지했다. 그러나 이는 수백에서 수천 달러에 이르는 부채환수통고를 받은 수많은 수급자들이 심리적 불안과 스트레스, 트라우마, 수치심을 얻고 심지어 자살에 이른 이후에야 일어났다.

호주의 로보데트 시스템뿐만 아니라 자동화된 사회보장제도의 위험성은 다양하게 고발되어 왔다. 2013년 네덜란드 국세청이 보육급여 신청자의 '위험 프로파일'을 만들어 부정수급 여부를 판정하게끔 했던 시스템은 사소한 서류상 오류를 부정수급의 증거로 취급하거나, 이중국적자를 부정수급 가능성이 높은 집단으로 분류하고, 약 2만 6천명의 부모를 부정수급자로 몰았다. 상환압박으로 인해 많은 사람들이 파산, 이혼, 가정해체를 경험하게 됐고 이는 2021년 네덜란드 정부 총사퇴로 이어졌다. 2013년 미국 미시간주에서는 실업급여 부정수급 자동판정시스템(MiDAS)를 도입을 통해 약 4만 8천 건의 부정수급 판정을 내렸는데 이는 이전에 비해 5배 증가한 수치였다. 시스템은 원금과 이자, 벌금 상황을 통보했고, 수급자들은 은행 계좌와 심지어 임금을 압류당하는 가혹한 징수를 경험했다. 이는 퇴거와 이혼, 신용등급의 파괴, 파산, 심지어 자살로 이어졌는데, 이후 주정부의 자체 검토에 따르면 93%가 오판이었던 것으로 드러났다. 부당하게 부정수급자로 간주된 3천명의 미시간 주민 원고는 미시간주 실업보험국(UIA)을 대상으로 소송을 했고, 미시간주는 2천만 달러의 합의금을 냈다. 유뱅크스는 <자동화된 불평등>을 통해 알고리즘 시스템이 가난한 이들을 감시하고 통제하는 '디지털 구빈원'을 낳는다고 비판했다. 가난한 이들의 정보를 수집하고, 이를 바탕으로 행동과 지출을 추적하고, 위험과 부정수급을 탐지하는 데이터베이스와 알고리즘, 위험예측 모델은 가난한 이들에 대한 감시와 통제를 손쉽게 강화하기 때문이다.

가난한 이들을 대상으로 하는 한국의 공공부조는 까다로운 소득재산 계산 방식과 선정기준, 급여 신청 단위로서의 '가구'와 부양의무자 기준 등 가족관계에 대한 제도적 개입에 이르기까지 복잡한 선정기준을 통해 매우 강한 선별/잔여주의적 성격을 강화해왔다. 빈곤사회연대는 이러한 제도 환경아래 추진되는 디지털화와 AI 도입은 다음과 같은 영향으로 이어질 수 있리라 우려해 왔다.

- ① 오류로 인한 피해와 수급자의 입증책임 강화. 디지털 처리의 복잡성과 일방성은 수급자의 이의신청이나 구명활동을 어렵게 만들고, 공적 자료와 실제 삶 사이의 차이를 해명할 책임이 수급자에게 돌아옴.
- ② 제도 밖의 예외 사례를 구제할 동기의 상실. 공적 자료를 중심으로 적격 여부를 판단하는 것으로 전담공무원과 일선 사회복지사의 역할이 제한되고, 적극적으로 제도의 목표(사회적 권리의 보장)를 실현할 필요가 약해짐.
- ③ 과도한 정보공개와 정보노출을 피하려는 사람들을 복지제도 이용에서 배제. 사회적 권리를 위해 데이터를 보호받을 권리를 포기해야하거나, 사생활을 방어하려 할 때 복지대상으로 '발굴되지 않는 책임' 또한 수급자의 몫이 됨.
- ④ 부정수급, 급여 삭감을 중심으로 한 제도 운영 강화. 일제조사와 같은 대량 감시가 잦아지

실업수당 중 1,500달러를 덜 받았다. 그러나, 로보데트 시스템에 의하면 4,000달러의 임금소득이 1년 평균소득으로 인정되므로, A씨는 매주 38.5달러의 실업수당을 삭감받아 연간 2,000달러의 실업수당을 덜 받아야 한다(38.5달러×52주). 이 경우, 로보데트 시스템은 A씨에게 500달러의 부채를 정부에 상환해야 함을 자동으로 통보하게 된다.” 홍승헌, 황하, 「누구를 위한 디지털 전환인가? 자동화된 복지행정의 위험성」, 정부학연구 제30권 제2호(2024), p69

고, 이는 수급자들로 하여금 급여를 받는 상태가 적정하지 않다는 인식을 강제함.

⑤ 정보유출 가능성. 제정한 정보가 사회복지급여 책정 외 다른 방식으로 활용될 수 있고, 특히 유출되거나 판매될 경우 저소득층에게 저질의 상품, 고리의 대출이 찾아가는 등 심각한 피해가 발생.

한국 사회보장제도의 디지털화와 AI도입은 크게 예산 효율화(부정수급 감시)와 복지 사각지대 발굴이라는 두 가지 목표아래 추진되어 왔다. 현재까지 도입된 제도를 가난한 이들의 권리 보장 관점에서 평가하고, 여기에 '영향 받는 사람들'의 얼굴을 떠올려보고자 한다.

2. 한국 사회보장제도의 디지털화, AI 도입 현황과 쟁점

한국 사회보장제도의 디지털화와 AI도입의 역사에서 중요한 전환의 장면을 꼽자면 2010년 행복이음의 도입과 2014년 사회보장급여법 제정, 최근 빠르게 넓어지는 AI활용이 있을 것이다. 행복이음은 사회보장제도 이용자들의 소득과 재산, 급여 이력 등 정보를 한 번에 관리할 수 있도록 정보화된 데이터를 통합하였고, 급여법은 사각지대 해소를 목표로 단전, 단수, 체납기록을 비롯한 위기 지표를 모두 데이터로 모을 수 있는 길을 열었다. 행복이음이 정해진 로직에 따라 데이터를 처리하는 전산화 시스템인 '규칙 기반 자동화(Rule-based Automation)'라면, 급여법에 따른 사각지대 발굴 시스템은 위기 가구 학습을 통해 위기도를 측정하는 '학습기반 AI(Learning-based AI)'로 빅데이터와 변수의 추가를 통해 알고리즘을 고도화시켜 왔다. 최근에는 화재나 활동을 감지해 응급 상황을 알리는 응급안전안심서비스 기기보급이나 대화를 나누는 AI인형 '효돌이'와 같은 서비스 제공형 AI 도입도 추진되어 왔다.

1)행복이음

행복이음은 사회보장정보시스템으로, 각종 사회복지 급여 및 서비스 지원 대상자의 자격과 이력에 대한 정보를 통합 관리하고, 지자체의 복지업무 처리를 지원하기 위해 시·군·구별 서울행정시스템의 31개 업무 지원 시스템 중 복지 분야를 분리, 개인·가구별 데이터베이스로 통합 구축한 정보시스템이다. 여기엔 7개 부처 170개 복지서비스 이력과, 국세청·노동부·건강보험공단 등에 등록된 소득과 재산자료, 인적정보가 개인·가구별로 모두 담긴다. 수급권자들이 읍·면·동 주민센터에서 복지 상담을 진행하면 초기상담과 접수 내용이 입력되고, 시·군·구청의 통합조사팀이 소득·재산 조사 자료를 바탕으로 지원 여부와 급여를 결정한다. 급여 지급이 결정된 수급자의 급여 지급과 관리를 위한 수급자의 소득·재산 변동사항 또한 행복이음에서 통합 관리 된다.



①수합한 정보가 흐르는 방향: 수급권자의 권리구제·편의보다 엄정한 관리와 효율화에 집중

2010년 이후 확인조사를 통해 절감한 사회보장 재정 금액은 보장 중지된 복지대상자에게 급여가 1년간 지급되었을 경우에 발생할 금액을 추계하는 것을 전제로 2013년 말까지 약 2조 812억원이며, 이 중 기초생활보장 재정 절감액은 약 1조 4,395억 원이다. (중략)

사회보장정보시스템(행복e음)은 정부 및 공공기관이 보유한 소득·재산 및 인적자료에 근거하여 자산조사에 적극 활용한 결과 2009년 평균 60일이 걸리던 자산조사 시간을 2010년 14일로 신속하게 수급을 받을 수 있게 되었다. 한편, 구축 이후 복지담당 공무원이 임의적 수정을 할 수 없도록 하였다.

- <국가정보화 20년의 성과>, 미래창조과학부 (2014. 9.1)

자산조사에 걸리는 시간이 14일로 단축되었다는 2010년 보도 자료의 언급이 무색하게도 행복이음 도입 당시 법률상 한 달로 정해져있던 기초생활수급자 적격조사 기간은 2026년 현재 두 달로 늘어났다. 2015년 맞춤형 개별급여 도입과 함께 기초생활보장법을 개정해 조사 기간을 연장했기 때문이다.

한국의 사람들은 어려움에 빠졌을 때 이용할 수 있는 여러 가지 제도가 마련되어있다는 홍보를 쉽게 접하지만, 막상 실제 사회보장제도를 신청하려고 하면 소득과 재산, 혹은 가구 구성이 신청에 적합하지 않아 이용할 수 없다는 판정을 쉽게 받는다. 더불어 이렇게 많은 정보를 수합하고 있는 정부가 왜 내가 필요한 제도를 제때 안내하지 않는지, 왜 직접 신청하지 않으면 아무런 서비스가 개시되지 않는지 의문을 갖는다.

실제 행복이음이 있더라도 수급권자들이 신청과정에서 요구받는 서류는 여전히 복잡하고 다양하다. 금융정보제공동의서를 제출하지만 6개월 치 통장 사본을 제출하라거나, 건강보험료 납입증명서를 떼어오라는 요청을 받기도 한다. 기초생활보장제도를 신청하면 수급자격을 갖추게 되었을 때 전기세 할인감면은 자동으로 신청되지만, 에너지바우처 신청 대상일 때도 그 사실을 스스로 알고 신청하지 않으면 이용하지 못한다. 이미 소득과 자산조사를 마친 주거급여 수급자가 생계, 의료급여 선정기준 이하로 소득이 하락해 신청 자격을 갖추게 되어도 신청을 안내하는 것은 각 지자체 담당자의 재량에 달렸을 뿐 통합적으로 이뤄지지 않는다.

이러한 사례는 행복이음이 '수요자 중심 서비스'를 표방했으나 제도에 이미 포착되어 있는 기수급자들의 과소수급 상태를 해소(권리 구제)하는 방향으로는 잘 작동하지 않는 반면, 행정의 오류나 정보 수합의 시차로 발생한 과오지급에 대해서는 더 잘게 조사할 수 있도록 도움을 주고 있다는 점을 보여준다.

행복이음은 이명박 정부의 국정과제 중 하나로, 복지예산 부정수급을 막겠다는 목표 아래 추진됐다. 보건복지부는 관리망을 통해 2010년과 2011년 복지급여 대상자 확인조사를 실시, 44만 8천 9백명이 수급권을 박탈했다. 이 중 기초생활수급자는 11만 6천명에 이른다. 정부는 이를 통해 3천 962억 원의 비용을 절감했다고 홍보⁴⁾했다. 행복이음 도입 이후 수급자 숫자는 2009년 말 157만 명에서 2010년 5월 152만 명으로, 같은 해 8월 148만 명으로 줄어들었다.

〈표 41〉 부양불능, 부양거부·기피 현황

(단위: 가구)

시 도	부 양 불 능									부양거부·기피			시 도
	계	행방불명	사실상이혼	해외이주	복역	군복무	시설수급자	사망후 호적미정리	기타 부양불능	계	양자·양부· 서모·의부	기타가족 관계해제	
2006년	112,762	46,772	12,165	9,696	3,329	3,435	1,508	723	35,134	115,554	28,019	87,535	2006년
2007년	120,909	52,198	12,794	11,638	3,759	3,701	1,812	731	34,276	129,917	29,803	100,114	2007년
2008년	111,533	50,342	11,014	12,350	3,641	3,237	1,877	454	28,618	137,455	26,918	110,537	2008년
2009년	102,340	48,594	10,226	13,511	3,653	3,040	1,950	376	20,990	148,300	25,814	122,486	2009년
2010년	9,163	738	230	690	1,947	2,641	2,104	193	620	119,254	19,063	100,191	2010년
2011년	20,915	8,148	136	988	1,750	4,840	3,673	1,042	338	168,756	19,806	148,950	2011년
2012년	16,352	6,695	110	957	1,400	2,903	3,101	932	254	193,457	18,534	174,923	2012년
2013년	13,250	5,829	68	1,240	1,109	1,826	2,625	355	198	213,651	15,142	198,509	2013년
2014년	12,859	6,284	1	1,806	1,028	1,383	2,111	239	7	258,562	14,293	244,269	2014년
2015년	21,503	9,819	12	7,251	1,229	1,201	1,791	165	35	282,609	13,340	269,269	2015년
2016년	23,822	11,244	9	8,219	1,368	1,133	1,607	217	25	295,998	12,093	283,905	2016년
2017년	25,791	12,436	7	9,147	1,385	979	1,505	312	20	318,690	11,353	307,337	2017년
2018년	26,540	12,868	7	9,665	1,379	910	1,390	303	18	340,311	10,521	329,790	2018년
2019년	27,440	13,381	7	10,212	1,407	867	1,300	252	14	363,019	9,719	353,300	2019년
2020년	27,918	13,581	7	10,478	1,505	866	1,277	191	13	383,472	9,049	374,423	2020년
2021년	28,863	13,972	6	11,098	1,550	823	1,230	172	12	401,725	8,496	393,229	2021년
2022년	26,324	12,180	528	11,150	1,078	5	147	1,224	12	399,533	7,484	392,049	2022년
2023년	28,937	13,297	562	11,710	1,231	5	692	1,427	13	424,347	7,460	416,887	2023년
2024년	29,208	12,967	571	11,903	1,267	6	903	1,580	11	430,543	6,813	423,730	2024년

같은 기간 빈곤 현장에서는 일제조사가 일어날 때마다 연락도 되지 않는 부양의무자의 소득과 재산을 근거로 한 일방적인 수급 삭감·박탈 통보가 이어졌다. 요양병원의 노인들은 창밖으로 몸을 던졌고, 끊어진 가족관계를 다시 소명해야 하는 사람들은 수급비가 우선 정지된 상황에서 막막한 생계에 고통을 겪었다. 2011년 하반기에는 고용주가 신고한 일용소득자료를 입수해 1만 2천명의 수급을 중지시켰다. 동기간 수급이 끊긴 수급자 3만 9천 명 중 31%에 달하는 수치였다. 이들 중에는 고용주의 신고소득이 실제 근로여부 및 급여액수와 차이가 있거나, 이미 해당 소득이 사라진 사람들도 있지만 이를 개인이 소명하기란 까다롭거나 불가능했다. 사회복지통합관리망 속 정보는 담당자가 임의로 수정할 수 없도록 되어 있었기 때문에 공적자료를 수정해야만 했는데, 이는 수급권자가 자력으로 할 수 있는 일이 아니기 때문이다.

특히 부양의무자기준은 심각한 문제를 초래했다. 기존에는 부양의무자가 있어도 부양받지 못하는 경우의 판별에 담당자의 현장조사가 중요한 역할을 했는데, 일제조사에 따라 기존에 인정했던 부양 불능을 무더기 취소했기 때문이다. 2010년 부양 불능에 따른 인정은 2009년에 비해 10분의 1수준으로 낮아졌고, 거부기피 사유에 대한 인정 건수 역시 낮아졌다. 둘을 합해 부양의무자가 있지만 부양받지 못하는 것으로 인정받은 건수는 2009년 250,640건에서 2010년 128,417건으로 절반가까이 줄었다.

이들이 부적정, 혹은 부정수급 상태였다고 보기는 어렵다. 이미 연락이 끊긴 가족이나 생사가 불분명했던 가족의 소득 때문에 수급에서 탈락한 여러 사례⁵⁾도 이를 뒷받침하지만, 2024년 부양불능, 거부기피 인정건수의 합이 다시 459,751건으로 늘어났다는 사실 또한 이를 방증한다. 일반적으로 결정한 배제의 총량이 오랜 시간을 거쳐 다시 제자리로 돌아온 셈이기 때문이다. 그러나 이것이 원상태로의 완전한 '회복'을 의미할 수는 없다. 일제히 탈락을 통보받은 이들 중 몇몇은 일정 시간을 보낸 뒤 다시 제도에 진입했겠지만 그 사이 어떤 이들은 건강을 잃고, 어떤 이는 가족관계가 한번 더 악화하거나 죽음을 맞았기 때문이다.

②권리구제 유인은 하락, 빈곤당사자의 입증부담은 증가

디지털화는 눈앞의 빈곤보다 공적전산망 속 자료를 중요하게 여기게끔 현실을 조직한다. 서울 역에서 지내는 한 홈리스는 사회보장제도 신청을 시도했으나 한해 전 해약한 주택청약통장을 어디에 썼는지 소명하지 못해 수급신청을 거절당했다. 현행 기초생활보장제도는 현재 소득과 자산 뿐만 아니라 그간 자산 처분 흐름에 대해서도 조사하는데, 수급자가 되기 위해 의도적으로 자산을 줄이는 것을 방지한다는 목적이다. 그런데 사라진 자산의 내역을 소명하는 항목은 매우 제한적이다. 공적자료로 확인되는 부채를 상환했거나, 교육비 혹은 의료비, 장제비와 같은 비용으로 사용했다는 것을 '공적자료'로 입증할 때만 재산을 실제 사용했다고 간주하기 때문이다. 청약통장을 해지하기 전에도 어려운 생활을 전전했던 그는 그간 친구에게 빌린 돈도 갚고, 신세진 이웃들에게 식사도 대접했다고 말했으나 담당 공무원은 1인가구 기준중위소득 50%에 해당하는 '자연감소분'만을 인정할 수 있다며, 내년에 신청하라고 말했다.

그의 통장에는 5백 원의 출금기록이 가득하다. 하루에도 대여섯 번씩 찍힌 5백 원은 커피 자판기 결제 내역이다. 추위에 언 몸을 녹이기 위해 아침에 두어 잔, 잠들기 전 몸을 데우기 위해 두어 잔씩 마신 출금 기록은 명백한 빈곤의 기록이었다. 수십 줄 이어진 5백원 출금 기록의 내용이 무엇인지 물어본 담당 공무원은 '추워서 마신 커피'라는 답변을 듣고 잠시 머뭇거리는 듯 했지만 '그래도 어쩔 수 없으니 내년에 다시 오라'고 말했다.

행복이음 도입은 최일선 현장의 권리구제 인식에도 영향을 주었다. 유뱅크는 디지털화가 초래하는 광범위한 정보수집과 자동화된 결정구조가 자격 선별 및 급여 지급과정에서 발생하는 문제에 대한 항의나 호소를 어렵게 만든다고 지적했는데, 이는 단지 복지 신청자 뿐만 아니라 업무를 담당하는 노동자에게도 영향을 미친다. 결정 근거나 오류발생 여부를 알기 어렵고, 수급권자가 겪는 공적자료와 실제 삶 사이의 차이를 노동자 본인이 입증하고 책임지기보다 수급권자와 관계기관에 청구하는 것이 '안전한' 선택이 되기 때문이다.

중요한 문제는, 취약한 상황에 내몰린 사람일수록 전산망 속 공적 자료와 실제 삶 사이의 차

5) 7살에 시설에 들어가 23년 동안 생활하다가 2010년 자립생활을 시작한 한 장애인은 2011년 6월 부양의무자기준으로 수급에서 탈락했다. 시설에 사는 동안 그는 부모님에 대한 이야기를 한 번도 들은 적이 없다. IMF 외환위기로 가정불화를 겪은 이후 가족을 한 번도 만나지 못했던 한 남성은 자녀에게 소득이 있다며 수급권을 박탈당했다. 신장장애를 가진 30대 수급자는 3년 전 아버지가 돌아가셨지만, 20년 전 아버지와 이혼한 어머니의 소득이 확인되어 수급에서 탈락했다. (2011 [비마이네](#))

명백한 현실 앞에서도 공적 자료를 1차 자료로 활용하라는 지침 안에서 사회복지 노동자들은 수급자들의 호소에 반응하기보다 전산망 속 숫자를 더 중요한 판단 기준으로 삼게 된다. 이런 모든 상황은 빈곤층이 급여를 포기하기를 종용하는 장치가 된다. 한국의 사회보장제도 선정기준이 매우 까다롭고 디테일하다는 점에서 자동화는 이러한 문제를 증폭시킨다.

2014년 12월, ‘송파 세 모녀법’이라는 이름으로 기초생활보장법과 긴급복지지원법이 개정되고, 「사회보장급여의 이용·제공 및 수급권자 발굴에 관한 법률」(이하: 급여법)이 제정되었다. 이를 통해 정부는 건보료, 수도·광열비 등의 체납정보를 수집해 위기 가구를 선정하고 있다. 급여법은 제도 시행 초기 13개 기관으로부터 23종의 위기정보를 수집했으나, 2026년 현재 26개 기관 47개 정보를 수집한다.

복지사각지대 발굴대상자 처리 업무구성도

발굴대상자 이관/선정/분배

발굴대상자 처리

발굴대상자 실적

중앙발굴 (보건복지부)

중앙발굴대상자 선정

발굴보조시스템을 통한 대상자 선정

중앙발굴대상자 이관

발굴대상자 전입출입 경우, 특별 소개발굴대상으로 자동 이관

중앙발굴대상자 분배

한도이월 추정정보의 소개 발원정보로 대상자 분배

신규대상자 등록
추가발굴대상자 등록

발굴대상자 알림
읍면동(읍) 담당자에게
본래 산하 발굴대상자 알림

발굴대상자 조회

(시군구)산하 읍면동발굴대상자 조회
(읍면동)소속 읍면동발굴대상자 조회
안내문 출력(※그린 연계)

발굴대상자 현황

주제정보, 발굴변수
본인을 포함한 가구구성원 일괄이력,
상당이력 등 상세정보 조회

비대상자 등록

공가, 취업, 소득초과등 비대상 조건
에 해당하는 비대상자 등록
중앙발굴대상자 비대상자 관리

**지자체
상시관리대상자**

**지도서비스
자재관리대상자**

상담

초기상담 등록

발굴대상자
정보일괄접수-경로
복지사각지대발굴 선택

연계요청

공공
서비스
신청

민간
서비스
연계

근로
복지
요청

사례
관리
요청

신청

통합
신청

건강
복지
신청

보육
신청

바우
처신청

실적등록

(중앙/지자체발굴)

 (시도)발굴년월 별, 시군구 별 발굴
대상자 지원실적 조회
 (시군구)발굴년월 별, 읍면동 별 발
굴대상자 지원실적조회
 (읍면동)발굴년월 별 발굴대상자
지원실적 조회

보건복지부, 지자체

지자체

보건복지부, 지자체

2024년 11월부터 발굴시스템은 AI를 활용한 초기상담 사업을 진행하고 있는데, 발굴된 위기 가구에게 사전 SMS를 발송해 안내하고, AI가 초기상담을 요청, 행복이음을 통해 AI 초기상담 결과를 확인하고, 이에 기반해 지자체는 심층 상담과 복지지원을 이어가게 된다.

[AI 초기상담 결과(예시)]

AI 초기상담 시스템		복지대상자 ※ (예시) 대상자: 김 OO / 위기정보(유형분류): 통신비체납(체납위기), 건보료 체납(체납위기), 일용근로대상자(고용위기)
①인사 및 본인확인	“김OO 님, OO시군구 OO읍면동 인공지능 상담 전화입니다. 어려움이 있어 복지도움이 필요하신지 저와 상담 하시겠어요?” 김OO 님, 본인 맞으신가요?	네. 상담할게요 맞아요.
②위기확인	(행복이음에서 대상자 위기정보 확인)	-
③복지요구 확인	네, 5개의 질문을 드릴게요. 끝까지 참여 부탁드립니다	-
	(건강) 본인이나 가족 중 신체나 심리, 정신 등 건강을 위한 도움이 필요하신가요?	네, 저는 교통사고 후유증으로 허리도 아프고, 어머니께서 치매 증상이 있으세요.
	(경제) 생계비 지출이나, 체납, 채무 등으로 도움이 필요하신가요?	네, 내야할 돈은 많고 너무 힘드네요
	(고용) 일자리 등 취업이 필요하신가요?	네, 몸도 아프고 무슨 일을 해야할지 모르겠네요
	(돌봄) 본인이나 가족 중 돌봄이 필요한 분이 계신가요?	네, 어머니께서 치매증상이 있으셔서 돌보는 데 어려움이 있습니다.
	(주거) 혹시 주거 지원이나 주거환경 개선에 도움이 필요하신가요?	아니요, 괜찮습니다.
④상담종료	김OO 님, 말씀하신 내용은 주민센터에 전달하여 지원이 가능한지 확인하여 연락드리겠습니다. 오늘도 좋은 하루 보내세요!	돈이 없어서 경제적으로 너무 어려워요. 휴대폰 요금도 밀려있고, 도움 주시면 좋을 것 같아요. (통화 종료)

①‘위기가구 발굴’을 이유로 한 광범위한 정보 수집

복지 사각지대에서 새로운 죽음이 발생할 때마다 발굴시스템에 수합되는 정보의 개수는 무비판적으로 확대되어 왔다. 정보가 넓게 파악되고 각 정보가 연결될수록 개인정보 침해 위험의 수준도 더 커지지만, 이에 대한 우려는 찾아보기 어렵다. ‘위기가구 지원율이 낮다’는 비판은 ‘더 많은 데이터를 수합하면 적중률이 높아진다’는 기술주의적 해법으로만 흡수되기를 반복하고 있기 때문이다.

개인정보 보호법은 최소 수집과 목적 제한, 정보 주체의 통제 원칙이 있으나 사회보장급여법은 이를 넘어서는 수준으로 정보를 넓게 수집하고, 이를 모두 연계해 관리하는 문제가 있다. 법 12조(자료 또는 정보의 처리)는 단전, 단가스를 비롯한 임대료, 관리비, 대출금등의 연체정보와 부채 정보 등을 모을 수 있도록 하고 있는데, 시행령 개정을 통해 그 내용과 개수를 계속 늘릴 수 있는 구조다. 제23조(사회보장정보의 처리 등)는 사회보장정보시스템에 이 모든 정보를 합해 관리하도록 하고, 제24조(사회보장정보시스템의 이용 등)⁶⁾와 제24조의 2(사회서비스정보시스템의 구축·운영 등)⁷⁾는 각 시스템의 정보를 연계·통합하고, 민간 사회서비스 제공기관의 정보까지 통합하도록 하고 있다. 수합된 정보를 결합하는 것은 생활에 대한 훨씬 더 정밀한 프로파일링을 가능하게 한다. 발굴 시스템이 채무와 신용등급, 심지어 건강 정보와 같은 아주 민감한 정보를 다룬다는 점에서 이는 언제든지 심각한 문제로 이어질 수 있다. 현재 수준의 정보 수합과 결합이 사각지대 발굴이라는 이용 목적을 고려할 때 최소한의 수준에서 수집되고 있는 것이 맞는지, 목적과 효과에 맞는 최선의 방법인지에 대한 평가가 필요하다.

6) 제24조 ①보장기관의 장은 제5조부터 제22조까지에 따른 업무를 효율적으로 수행하기 위하여 사회보장정보시스템을 이용하거나 관할 업무시스템과 사회보장정보시스템을 연계하여 이용할 수 있다.

7) 제24조의2 ①보건복지부장관은 보장기관이 다음 각 호의 법인·단체·시설·기관(이하 “사회서비스제공기관”이라 한다)의 업무를 전자화하고 업무 수행에 필요한 정보를 통합·연계하여 처리·기록 및 관리하는 정보시스템(이하 “사회서비스정보시스템”이라 한다)을 구축·운영할 수 있다.

○ 발굴시스템 연계 위기정보 현황 (21개 기관 47종, '26.6월 기준)

연번	데이터 항목명	수집 주기	출처 기관명	
1	단전	매월	1	한국전력공사
2	전기료 체납	매월		
3	단수	격월	2	각 지자체 상수도사업본부(사업소)
4	수도요금 체납	격월		
5	단가스	격월	3	도시가스사(33개소)
6	가스요금 체납	격월		
7	초중고 교육비 지원사업 대상자	격월	4	교육부
8	건보료 체납	매월	5	건강보험공단
9	국민연금보험료 체납	매월		
10	피부양자 장기요양상태	격월		
11	의료비과다지출대상자	격월		
12	중증질환 산정특례	격월		
13	의료기관 장기 미이용	격월		
14	노인 장기요양 등급	격월		
15	건강보험료 부과내역	격월		
16	재난적 의료비 지원대상	매월		
17	전세취약가구(확정일자, 주택조사)	격월	6	국토교통부, 한국토지주택공사
18	월세취약가구(확정일자, 주택조사)	격월		
19	공공임대주택 임대료 체납	매월/격월	7	한국토지주택공사, 12개 지방도시공사(격월)
20	개별연장급여대상	매월	8	고용노동부
21	실업급여 수급자	매월		
22	실업급여 미수급자	매월		
23	고용위기(단절, 실업)	매월		
24	일용근로대상자	매월		
25	산재요양종결후 근로단절자	매월	9	근로복지공단
26	자살고위험관리대상자	매월	10	자살예방센터
27	세대주가 사망한 가구	격월	11	행정안전부
28	주민등록세대원	격월		
29	금융연체대상자	매월	12	한국신용정보원
30	채무조정 중지(실효)자	매월		
31	공동주택 관리비 체납	격월	13	아파트 관리사무소(관리업체 5개소)
32	집합건물(오피스텔) 관리비 체납	격월	14	집합건물 관리단
33	휴폐업자	격월	15	국세청
34	통신비체납	매월	16	한국정보통신진흥협회
35	노후긴급자금 대부(실버론)	매월	17	국민연금공단
36	정책서민금융신청반려자	격월	18	서민금융진흥원
37	기초·긴급 신청탈락·중지	격월	-	보건복지부
38	시설 퇴소자	격월		
39	보건소 집중관리군	매월		
40	저소득층 기저귀조제분유 지원 대상자	격월		
41	영양플러스 미지원 대상자	매월		
42	맞춤형급여 신청	격월		
43	범죄피해	격월	19	경찰청
44	화재피해	격월	20	소방청
45	내원사유 자해/자살 대상자	매월	21	중앙응급의료센터
46	재난피해가구	격월	-	행정안전부
47	신생아 난청 확진자	매월	-	보건복지부

②‘발굴’하지만 ‘지원’하지 않는다

위기가구 발굴 시스템을 통해 신청을 기다리는 것이 아니라 ‘찾아가는 복지’를 실현한다고 설명하지만, 발굴하고도 지원하지 않는 상황은 계속되고 있다. 보건복지부에 따르면 2025년 위기가구 발굴대상자 중 복지서비스로 연결된 경우는 63.4%이고, 이 중 민간서비스로 연결된 경우는 66%다. 전체 발굴대상자 137만 명 가운데 공적 서비스로 연결된 경우는 29만 명으로 21.7 4%이고, 상대적으로 안정적인 지원에 해당하는 기초생활보장제도로 연결된 경우는 58만 명으로 전체 발굴 대상자 중 4.28%에 불과하다. 민간서비스와 기타 공공서비스가 상대적으로 충분하거나 지속적이지 않다는 점을 고려할 때 발굴된 위기가구 대다수는 필요한 지원을 받지 못하는 상황이 반복될 것이라는 점을 예상할 수 있다.

(단위: 명, %)

연도	발굴 대상 (a)	복지서비스 지원(b) (b/a x 100)	공공 서비스					민간서비스 연계(d)** (d/b*100)
			소계(c) (c/b x 100)	기초 생활보장	차상위	긴급 복지	기타 공공 서비스*	
'24년	1,423,466	831,660 (58.4%)	252,425 (30.4%)	48,451 (5.8%)	9,888 (1.2%)	23,796 (2.9%)	170,290 (20.5%)	579,235 (69.6%)
'25년	1,371,307	876,613 (63.9%)	298,060 (34.0%)	58,703 (6.7%)	13,253 (1.5%)	25,355 (2.9%)	200,749 (22.9%)	578,553 (66.0%)

* 기초연금, 장애인연금, 사회서비스이용권(노인돌봄, 장애인 활동지원) 등

** (민간서비스) 공동모금회, 푸드뱅크, 민간기관 결연후원금 등

이는 위기의 확인에도 불구하고 지원할 방안이 없는 제도적 공백으로 인해 발생한다. 2023년 복지사각지대 발굴시스템을 통해 각 지자체에 분배된 발굴 대상자는 141만 4천명, 이 중 지원을 연결하지 않는 비대상자로 처리 된 사유 중 가장 높은 것은 소득재산 초과(31.2%), 기수급자(11.3%) 등 지원기준 초과자 비중이 64.2%⁸⁾였다. 이렇게 발굴되더라도 연계할 자원이 없는 경우 행정기관에 대한 당사자의 불신을 초래하거나 민원에 직면하기도 한다. 이를 무마하기 위해 간단한 후원물품을 건네면서 ‘달래고 넘어가는’ 상황이 벌어지기도 한다.⁹⁾

이런 상황은 발굴제도의 효과성에 의문을 갖게 한다. 보건복지부는 발굴시스템의 정확도가 높아지고 있다고 자찬하지만, 발굴 대상이 사회보장제도에 진입하는 것은 단지 발굴 기술의 발달이 아니라 선정기준의 개선이나 빈곤의 확대와 같은 요인이 영향을 미치기 때문에 보장 비율만으로 개선을 파악하기 어렵다. 게다가 ‘발굴’은 인공지능에 의해 완성되는 것이 아니라 곳곳에 노동과 비용이 투입되는 작업이기 때문에, 투입된 역량 대비 실제 효과가 있는가에 대한 검증이 필요하다. 1년에 여섯 차례 대규모의 위기가구 발굴 작업이 진행되는데 적은 숫자의

8) <복지사각지대발굴시스템 비대상자 실태분석과 정책제안>, 한국보건사회연구원, 2024

9) “차량이나 소득이 일용직 같은 경우에는 좀 넘어가잖아요. 그래서 안 되는 경우가 더 많은데 계속 안 부만 확인하는게 그래서 그분들이 좀 짜증내는 부분도 있는 것 같아요. 어려워해서 막상 상담 한번 해보면 기초생활수급자 안에는 안 들어와요. 그러면 이제 저희가 기껏 할 수 있는건 후원 물품 들어오는 거 드리면서 우선 라면이나 이런거 드리면서 한번 달래고 넘어가는 경우가 많아요.” 2025, 최정은, 민소영 「접근이 힘든 복지사각지대 발굴 개선방안 연구: 지자체 업무 담당자의 경험을 중심으로」, 한국사회복지행정학 제27권 제3호, p114

사회복지 전담공무원들이 다양한 업무를 처리하는 동시에 위기가구 최대 두 번의 현장 방문을 기한 내 달성하고, 이를 다시 데이터로 입력하는 추가 업무를 처리하고 있기 때문이다. 일선 담당자들은 내려오는 정보가 너무 방대하다거나, 얼마 전 비대상 처리한 대상자가 재발굴되어 명단에 오르거나, ‘이미 알고 있는 사각지대’인 경우가 빈번하다는 것을 경험할 때 효과적이지 않은 것 같다고 느낀다.

‘발굴’하더라도 지원하지 않는 황망한 현실은 2025년 3월 주검으로 발견된 A씨를 통해서도 잘 드러난다. 서울 강남구 반지하에 살던 세입자 50대 남성 A씨는 6개월 이상 무직상태였고 월세와 전기요금이 수개월 체납되었으며, 도시가스도 중단된 상태였다. 강남구청은 2024년 A씨를 위기가구로 선정하고 자택을 방문했으나 만나지 못했다. 이후 A씨가 강남구청을 찾아가 긴급복지지원제도 신청을 문의했으나 강남구는 예산 소진을 이유로 1월에 방문하라는 안내를 했고, 2025년 1월 다시 주민센터를 찾았지만 새해 예산 집행이 개시되지 않았으며 5일 뒤 다시 방문하라고 반려를 받았다. 그 이후 다시 주민센터를 찾지 않은 A씨는 3월 주검으로 발견됐다. 긴급복지지원제도는 48시간 내 선지원을 표방하는 제도지만 위기가구로 발굴된 A씨의 경우 3개월이 넘도록 아무런 지원을 받지 못했다.

이렇듯 실제 해결해야 하는 문제는 ‘비대상’으로 밀려나는 이들의 권리를 보장하기 위한 제도 개선과 예산확보, 적극적인 현장 개입이지만 2014년 이래 사회보장제도의 정책결정과 자원 흐름은 발굴 시스템의 강화로 집중됐다.

③귀납적 추론의 한계 벗어날 수 없는 AI, 빈곤 프로파일링의 위험

위기가구 발굴시스템이 위기도를 산정하는 알고리즘은 공적지원을 받은 대상자 정보를 훈련 데이터로 활용한 ‘머신러닝 기반 예측 모델’이다. 이는 사전에 정의된 산정 알고리즘을 통해 계산되는 것이 아니라 위험도 예측 모델을 자동으로 생성, 갱신하는 방식으로 작동한다. 위기 정보가 확대되고, 새로운 사각지대 유형의 문제가 드러날 때마다 복수의 모형을 만들어낼 것으로 추정되나 구체적인 내용은 공개하지 않는다.

AI, 인공지능이라는 단어는 마치 컴퓨터가 스스로 언어나 이미지를 구상하고 창조하는 것처럼 보이게 하지만 이는 사실과 다르다. 인공지능은 기계학습을 통해 데이터의 패턴을 찾고, 연관 관계를 귀납적으로 추론하는 것이기 때문이다. 위기가구 발굴시스템 또한 기존 복지대상자의 조건을 학습해 이와 비슷한 대상자를 추리는 것이기 때문에, 현실에 등장하는 새로운 형태의 빈곤과 예외상황에 놓인 사람들의 위기를 알아서 찾아내지 못한다.

또, 이러한 발굴시스템은 개인의 공적 데이터를 수집하고 빅데이터 분석을 통해 고위험도 발굴 대상자를 추출한다는 점에서 위험의 점수화와 범주화를 수행하는 방식이다. 이는 사회적 평점을 부과하는 것으로, 유럽 인공지능법이 정하는 4단계 위험¹⁰⁾ 중 ‘수용할 수 없는 위험(Unacceptable risk)’인 ‘사회적 평점(Social scoring)’에 해당한다. 유럽 인공지능법은 이를 엄격히 금지한다. 필립 알스톤은 위험의 점수화, 범주화가 오류의 가능성, 점수화 및 범주화의

10) ①불용인 위험(Unacceptable Risk): 사회적 신용 평가, 공공장소에서의 실시간 원격 생체인식 등은 원천적으로 금지 됨.

②고위험(High Risk): 실수적인 공공서비스 및 혜택에 대한 접근을 포함해 교육, 노동, 출입국 관리 등에 관한 시스템. 엄격한 적합성평가(Conformity Assessment)가 필요함.

③제한적 위험(Limited Risk): 챗 봇이나 딥페이크, 감정인식 시스템 및 생체 분류시스템은 사용자가 AI와 상호작용하고 있음을 인지할 수 있도록 투명성 의무가 부과 됨.

④최소위험(Minimal Risk): 비디오 게임 AI나 스팸 필터링 시스템 등 위험성이 매우 낮거나 거의 없는 시스템. 별도의 법적 의무나 규제 없음.

근거를 공개하지 않으면서 생기는 권리침해 가능성, 점수화 및 범주화가 현재의 불평등과 차별을 강화하거나 합리화 할 가능성이 있다는 점을 지적했다.

3. ‘영향 받는 사람들’의 얼굴을 떠올리며

초기상담 전화가 AI로 대체되는 것을 어떻게 생각하는지 질문을 던졌을 때 수급권자들은 ‘차라리 좋을 것 같다’는 대답을 했다. 수화기 너머에 있는 사람의 기계적이거나 차별적인 반응을 이미 경험했기 때문에 ‘차라리 진짜 기계라면 이해라도 되니까’ 그렇게 답한 사람도 있었고, ‘기분을 맞춰줘야 한다는 압박’에서 자유로울 것 같다고 기대하기도 하며, ‘오히려 더 공정하지 않을까’ 기대하기도 했다. 담당 공무원 역시 정보가 부족한 상황에서 발굴된 위기가구를 무조건 만나야 한다는 압력에 시달리고, 막상 얼굴을 마주친 사람에게 아무것도 주지 못한 채 상담을 종료할 때의 민망함을 마주하느니 AI초기 상담을 이용하는게 효과적이라고 느낀다. 이는 AI를 선호한다기보다 기존 시스템에 대한 불신이나 실패가 AI에 대한 기대로 드러나는 반응이라고 보아야 할 것이다. 그러나 차별하는 제도 안에서 작동하는 AI는 같은 실패를 피하기 어렵다.

행복이음 도입과 급여법을 통해 진행된 대량의 데이터 수합과 디지털화는 이미 사회복지 현장에 상당한 변화와 영향을 미쳤다. 이 두 가지 제도에 대한 평가는 더 빠른 속도로 점증할 AI 도입에 앞서 선행되어야 할 질문이 무엇인지 묻고 있다. 최근 사회복지 학계와 현장에서는 AI 도입에 대한 다양한 관심과 논의가 이어지고 있다. 최대한 AI를 활용해 사회적 이익과 효율성을 강화해야 한다는 입장, 실시간 소득을 사회보장제도와 연계하고 신청 자동화로 나아가자는 입장 등 여러 입장이 토론되고 있지만, 도입과 활용에 대한 우려의 목소리는 드물다. AI에 대한 몇 가지 위험에도 불구하고 인간의 결정과 개입을 남긴다면 이를 통제할 수 있다고 보는 것이 대체의 합의인 것으로 보인다.

2019년 UN 극빈과 인권에 관한 특별보고관 필립 알스톤은 디지털 복지국가의 위험성을 경고하는 보고서¹¹⁾를 통해 디지털화가 매우 가차없는 속도로 진행되기 때문에 복지 디스토피아를 막기 위해서는 진로를 크게, 신속하게 바꿀 필요가 있다는 점을 경고했다. 그는 AI에 의해 영향을 받는 프라이버시는 단지 사적인 행동이 다른 사람들에게 보여지는 수준이 아니라 개인의 자율성과 선택을 훼손하는 영역에 걸친다는 것을 강조하며 더 자주 빈곤층이 감시당하고 이로 인해 처벌받을 가능성을 염두할 것을 경고했다.

보고서는 결론을 대신해 다음과 같은 것들을 지적했다. ①디지털 사회복지기술이 과학적 진보의 필연적 결과물이 아니라 인간의 정치적 선택을 반영한다는 점 ②시장 논리가 우세하도록 방치한다면 인권이 무시될 것이라는 점 ③신기술을 뒷받침하는 가치가 가난한 이들이 아니라 백인, 남성, 부유층, 북반구에 맞춰져있다는 점을 인지하고 이를 조율하기 위한 세심한 노력을 기울여야 한다는 점 ④예측 분석과 알고리즘은 기존 데이터에 반영된 편향을 재현하고 악화시킬 가능성이 높다는 점 ⑤빅테크는 사회보장이 아니라 부유층을 위한 상품 개발에 관심을 쏟는다는 점 ⑥비용절감이 아니라 뒤쳐진 사람들을 돌보는 새로운 방법, 취약계층의 생활수준을 보장하기 위한 기술에 대한 관심이 증진되어야 하며, 이것이야말로 디지털 사회복지 혁명이

11)

http://antipoverty.kr/xe/index.php?_filter=search&mid=publish&search_target=title_content&search_keyword=%EB%94%94%EC%A7%80%ED%84%B8&document_srl=1265364

될 것이라는 점이다. 한국에서 AI는 아직 사회보장제도 적격 판정의 도구로 사용되거나 고지능 인공지능을 도입하지 않았으나, AI도입에 대한 범정부적 의지와 산업 육성을 최우선 가치로 고려하는 인공지능법의 등장은 머지않아 새로운 국면이 도래할 것을 예고한다.

①디지털화와 AI도입을 촉발하는 발굴주의(신청주의 해소)라는 트리거

지금 필요한 또 하나의 질문은 행정절차 개시에 있어 수급권자의 필요나 요구보다 AI의 모니터링 결과가 더 많은 영향을 미치고 있지 않은가에 관한 것이다. 더 많은 데이터가 수합될수록 양질의 데이터를 관리하는데 더 많은 노동시간을 할애하고, 실제 수급권자를 만나는 것이 아니라 변경된 소득·자산정보에 관한 행복이음 시스템의 알람이 사회복지 현장 노동자의 행동을 불러낼 때 현실에 대한 개입 역량은 축소하게 될 것이다.

이재명 대통령은 복지 분야에서 주요하게 두 가지 쟁점을 던지고 있는데, 하나는 자동 지급을 통한 신청주의의 해소¹²⁾이고, 또 다른 하나는 AI 적용 확대다. 이는 기존 복지 사각지대가 ‘미신청’에 의해 발생한다는 생각에서 기인한다. 그러나 사각지대는 단지 미신청이나 거부에 의해서만 발생하는 것이 아니라, 까다로운 선정기준에 따른 제도적 배제, 급여를 받고 있으나 실제 필요에 미치지 못하는 과소 수급에 의해서도 발생한다. 2014년 송파 세모녀에서 2020년 방배동 모자, 2022년 창신동 모자와 2025년 강남 남성에 이르기까지 사각지대의 죽음이 발생할 때마다 일제조사와 발굴이 해법으로 제시되었으나 이들은 보이지 않는 사람들이 아니었다. 여러 번 주민센터에 찾아가 사정을 호소했으나 선정기준을 벗어난다는 답변만을 들었거나, 부양의무자에게 연락 가는 것을 원하지 않아 수급신청을 포기했거나, ‘발굴’되었음에도 예산 부족으로 밀려난 사람들이다.

‘신청주의’를 문제로 지목하는 일련의 경향은 신청의 단계가 사라지면 사각지대와 신청 과정에서의 낙인감이 해소될 것으로 기대하지만, 자동신청과 지급이 실제 이뤄지더라도 제도개선이 없는 한 사각지대는 그대로 유지될 것이다. 오히려 발굴을 강조하는 흐름이 낳는 부작용도 크다. AI와 공적 데이터에 의존한 발굴 강조는 현실의 복잡함을 외면하게 만들기 때문이다. 시스템에 존재하는 ‘사실’ 뒤에는 복잡한 복지제도와 그보다 더 난해한 빈곤 현실이 있다. 이 사이를 헤치고 나올 수 있도록 당사자를 지원하는 것이야말로 사회보장제도의 역할이라는 사실을 잊어서는 안 된다. 디지털화와 AI활용이 광범위한 데이터 수집으로 빈곤층의 정보를 추출하고, 컴퓨터에 의해 발굴된 위기가구 방문에 힘을 쏟도록 유인하지만, 눈앞에 있는 사람의 호소는 놓치도록 만들지 않는지에 관한 평가가 필요하다.

②‘AI기본사회’에서 벌어지는 복지수급권자의 데이터 추출과 공공서비스 민영화 위험

지난 7월 21일 김용범 청와대 정책실장은 본인의 페이스북에 “AI에 대한 접근은 이제 하나의 기본권”이라고 썼다. 이를 구체화할 계획은 ‘모두의 AI’에서 시작한다. 국민에게 국산 AI를 무료로 제공해 수요와 데이터를 꾸준히 쌓아가고, 이를 통해 자국 AI시장을 키운다는 계획이다. 두 번째 단계는 공공 에이전트다. 그는 “공공서비스를 국민 대신 찾아 신청까지 도와주는 AI 에이전트, 그리고 공공데이터를 열어 민간이 가져다 쓰게 하는 일”이며 “개별 병원이나 기업이 홀로 감당하기 어려웠던 첫 도입의 위험을 국가가 먼저 떠안아, 공공 영역에서 사용 사례

12) 보건복지부는 위기가구 생계급여 직권신청과 관련해 지난 5월 공무원 면책 규정을 마련하였는데, 긴급복지 지원 이력이 있는 위기가구 가운데 미성년자, 발달장애인 등 스스로 동의할 수 없는 가구원이 있고, 친권자등으로부터 동의를 받기 어려운 경우 담당 공무원이 직권으로 신청할 수 있도록 하는 사회보장급여법의 규정을 확대 적용하는 내용이다. 그러나 이는 3개월 안에 금융정보제공동의서를 보완해 다시 조사하는 방식이라 효과가 제한적일 것으로 예상된다.

와 데이터와 신뢰를 만들어”낼 것이라고 했다.

이러한 정부의 구상은 AI산업육성에 초점을 두고 있는 인공지능기본법을 통해서도 드러난다. 과학기술정보통신부는 지난 8월 28일 SK텔레콤·카카오·KT 컨소시엄을 모두의 AI 사업수행기관으로 선정했는데, 이는 비용과 데이터, 위험은 사회가 부담하고 핵심 인프라와 수익은 민간 기업이 독점하게 될 수 있다. 특히 이 과정에서 공공서비스를 이용하기 위해 신청자들은 자신의 일상생활을 데이터로 공개해야 하는 압력에 놓이게 된다. 민감 정보가 통합되어 있는 사회보장정보시스템과 결합 될 때의 위험은 더 커진다.

‘모두의 AI’는 과연 무료일까? 참여 기업은 대규모의 이용자와 데이터를 확보하고, 안정적으로 시장을 독점해 플랫폼 생태계를 선점할 수 있는 반면, 사회 전체는 함께 비용을 치른다. GPU·데이터센터 구축과 전력, 통신망, 데이터와 행정기관 인력을 비롯해 차별적 판단이나 오류나 실패와 같은 모든 비용이 사회 전체에 전가되기 때문이다. 전형적인 위험의 사회화, 이윤의 사유화다. 이는 중국에 AI기업의 기술과 데이터, 인프라에 행정이 종속되도록 만들고 공공서비스의 디지털 인프라를 민영화하는 결과로 드러날 수 있다.

지난해 한국을 방문한 필립 알스톤은 한국 사회운동과의 대담에서 인공지능이 빈곤층, 홈리스, 그리고 사회 정책 전반의 상황에 활용될 수 있으나 이것이 개선에 이를 가능성은 거의 없다는 점을 강조했다. 이는 인공지능이 정부의 역할을 축소하고, 효율성을 중시하는 완전히 다른 가치 체계에 기반을 두고 있기 때문이며, 본질적으로 자유주의적 가치는 개인이 스스로를 돌봐야 한다고 가정하기 때문이다.

복지수급자들의 행동과 지출, 정보에 대한 대량의 데이터 수집은 유출이나 오류가 발생할 때만 문제를 일으킨다는 관점이나 ‘인간의 최종적 개입이 있으면 괜찮다’는 관점에 재고가 필요하다. 빅데이터에 의한 행동과 생활의 추적은 단지 프라이버시를 보호하지 못하는 문제가 아니라 일상을 데이터화 하고, 패턴을 추적당하고, 감시와 통제를 손쉽게 강화하는 도구를 만들어내는 것이기 때문이다.

③누가 결정하는가, 책임지는가

복지의 디지털화는 사회보장정보의 대규모 데이터 수집과 결합으로 이어졌고, 이제는 AI 에이전트에 의한 공공서비스 제공이 예보되고 있다. 복지 사각지대를 해소한다는 명분 아래 대규모의 데이터가 통합된 것처럼 사회보장제도 이용의 효율화와 사용자 편의를 추구한다는 명분은 민간 AI기업들에게 데이터와 산업 설비를 지원하고 있다.

로보데트 스캔들에 관한 호주 왕립조사위원회 보고서는 다음과 같은 반성을 담았다.

“사회복지정책과 관련하여 취할 수 있는 사고방식에는 여러 가지가 있다. 하나는 많은 시민이 인생의 다양한 시기에 소득지원이 필요하다는 점을 인정하는 것이다. 학업이나 구직 과정에서 일시적으로 필요한 경우도 있으며, 연령이나 불우함 또는 장애 등의 이유로 말미암아 소득지원이 보다 장기적으로 필요한 이들도 있다. 이러한 입장에서는 이들에 대한 지원을 기꺼이, 적절하게, 그리고 존중하는 마음으로 제공해야 한다고 믿는다. 다른 접근 방식은 사회보장 혜택을 받는 사람들을 국가경제의 발목을 잡는 존재로 간주하고, 가능한 모든 방법을 동원해 이에 대한 예산을 줄여야 한다고 주장한다. 이러한 입장에서는 복지수급자를 납세자의 부담이라 지칭하고, 복지 혜택을 청구한 사람들에게 부담스러운 조건을 요구하고, 공무원의 복지지원을 최소화하고, 정당하든 아니든 혜택을 회수하고, 사회보장수급자의 상태를 불쾌하고 바람직하지 않게 만들고자 한다. 로보데트 시스템은 후자의 대표적인 사례다.”¹³⁾

‘영향 받는 사람들’은 누구인가? 공적 데이터가 나의 현실을 설명하는데 실패해 복지제도에서 미끄러진 사람의 낙담, 눈앞의 빈곤 현실을 보면서도 권리를 구제하는 대신 퇴근을 선택하는 모멸감, 외로움을 견디지 못해 친밀함을 가장한 저질의 상품과 관계에 돈을 지불하는 찻잔함, 모두가 스크린 속 정보만을 바라보는 사이 ‘말할 곳’을 찾지 못한 사람들의 텅 빈 얼굴. 복지사각지대 해소를 명분으로 데이터 긁어모으기가 반복되어 왔지만, 실제 중요한 것은 누구에게, 얼마를 지급할 것인가를 에 관한 정치사회적 합의다. 그렇기 때문에 지금 우리가 궁금 해하고 질문해야 하는 것은 ‘AI도입이 우리의 삶을 어떻게 바꿀 것인가’가 아니라 ‘우리의 오늘을 변화시키기 위해 필요한 것’이다. 캐시오닐은 “빅데이터는 과거를 암호화 할 뿐, 미래를 창조하지 않는다. 즉, 기존의 질서를 자동화할 뿐”이라고 했다. AI는 우리의 삶에 개입한다. 그러나 새로운 미래를 결정하는 것은 언제나 우리 앞에 놓인 책임이었다. 빈곤과 불평등을 해소하기 위한 노력 또한 그렇다.

13) Commonwealth of Australia, 2023. 「정부학연구」제30권 제2호 <누구를 위한 디지털 전환인가? 자동화된 복지행정의 위험성>, 홍승헌·황하, 2024에서 재인용

AI, 장애해방의 도구인가, 관리 통제의 기술인가?

전국장애인차별철폐연대 초록

2025년 11월, 이재명 대통령은 국회 본회의장에서 2026년도 정부 예산안에 관한 시정연설을 진행했다. 해당 연설에서 “AI”와 “인공지능”은 총 스물여덟 번이나 언급되었다. 같은 연설에서 “장애인”과 “민생”은 각각 두 번, “복지”는 단 한 번 언급되었다.

장애운동의 현장에서 체감하는 AI에 대한 관심 역시 뜨겁다. 장애인 당사자를 위한 AI 활용교육이 요청되는가 하면, 2025년 12월 장애와 인공지능을 다룬 심포지엄에서 한 관계자는 “Agentic AI 기술이 결합된다면, 업무 환경에서의 장애는 마법처럼 소멸할 것”¹⁾이라고 확인하기까지 했다.

과연 AI는 ‘장애를 소멸’시킬 것인가? 그렇다면 그것은 온당한가?

장애인 당사자의 AI 활용

옳고 그름을 따져보기도 전에 이미 AI는 장애인의 삶의 한 장면이 되었다. 한 중증장애여성은 처음 전동휠체어를 갖게 되었던 때를 회고하며 “지금 생각해 보면 전동휠체어는 단순한 이동보조 기기가 아니라 내 삶을 제대로 살아가게 해 준 기적의 날개였다”고²⁾ 말했다. 타인의 지원 없이 스스로 이동이 가능해지면서 자신감이 생겼고, 자신감이 곧 삶의 변화에 대한 욕망으로 변졌다는 것이다. AI 역시 많은 장애인 당사자에게 ‘누군가에게 지원을 요청하지 않아도 되는 일상’의 범위를 넓혀주고 있다.

발달장애인 당사자들이 ChatGPT에게 가장 자주 하는 말은 “쉽게 설명해줘”였다. 이해하기 어려운 문서 내용이나 뜻 모를 단어를 붙여넣고 쉽게 설명해달라는 문구를 덧붙이면 즉시 알기 쉬운 해설을 받을 수 있다.

장애우권인문제연구소에서 실시한 심층인터뷰³⁾에 참여한 청각장애인 당사자는 수업을 들을 때 실시간 자막서비스 ‘소보로SOVORO’를 적극적으로 활용하고 있다고 답했다. 해당 서비스를 이용하면 말소리를 태블릿 화면에서 실시간 자막으로 확인할 수 있어 내용을 이해하는 데 도움이 되고, 전보다 적극적으로 수업에 참여할 수 있다는 것이다. 화재 경보를 진동과 시각 정보로 전달하는 서비스도 개발되었다. 사용자는 베개의 진동이나 조명의 반짝임으로 위험을 감지할 수 있다. 불완전하지만 음성 언어를 인식해 수어로 통역하는 서비스도 개발되어 있다.

같은 인터뷰에서 시각장애인 당사자는 대화로 인터넷 쇼핑을 할 수 있도록 지원하는 AI 브라우저나 이미지를 음성으로 해설해 주는 서비스를 활용한다고 밝혔다. 삼성전자가 만든 저시력 보

1) 박민준, 심포지엄 <장애와 인공지능: 기회와 위기를 논하다> 자료집, 2025. 12. 12.

2) 수동휠체어에서 전동휠체어로, 나의 삶 ‘보조기기’. 김상희, 2021. 07. 18., 비마이너

3) 장애 관련 인공지능의 활용과 접근에서의 차별, 김영연, 심포지엄 <장애와 인공지능: 기회와 위기를 논하다> 자료집, 2025. 12. 12.

조기기 ‘릴루미노’는 웨어러블 기기 ‘릴루미노 글래스’로 사물을 촬영하고 윤곽선 강조, 확대/축소, 색반전/대비 등의 후처리를 통해 인식률을 높일 수 있는 형태로 변환한다.

신체장애로 인해 타이핑이나 클릭이 불편한 장애인 당사자들에게 같은 동작을 무수히 반복해야 하는 리서치는 매우 수고롭고 오랜 시간이 드는 작업이었으나 정보량을 압축해주는 AI 기술은 훨씬 더 수준 높은 과제를 단시간에 완성할 수 있게 해주었다. 음성으로 집 안의 기기를 작동하거나 제어하는 장치 역시 지체장애인들의 자유도를 높이는 도구가 된다. 장애우권익문제연구소의 인터뷰에서 정신장애인 당사자는 대화형 AI를 이용해 정서적 지지를 얻는다고 답하기도 했다.

언어장애를 가진 이들이 대체 보완 의사소통 수단으로 사용하는 AAC 역시 AI 기술이 도입되면서 좀 더 자연스러운 음성으로 하고자 하는 말을 전달할 수 있게 되었다. 그만큼 장애인 당사자의 공적 발화 경험도 늘어났다.

지자체별, 기관별로 흩어져 있는 장애인 복지서비스 가운데 특정 개인이 신청할 수 있는 서비스를 선별해 신청 절차를 안내받거나, 국가법령정보센터 데이터를 활용해 차별 상황에서 침해받은 법적 권리를 확인하고 구제 절차를 안내받는 것도 가능하다.

AI 사용이 늘면서 기획의도를 알 수 없는 영상이나 음악을 종종 받게 된다. 앱으로 작곡과 영상 편집이 가능해지자 자기표현의 수단으로 음악과 영상을 활용하여 배포하는 동지들이 있어서다. 직접 손으로 그림을 그리지 않더라도 원하는 그림체로 원하는 이미지를 생성할 수 있고 어느 정도 완성도 높은 결과물이 나오니 재미있어하는 동료들이 많은 것 같다. 아직은 앱 화면이나 브라우저에만 존재하는 AI를 넘어 피지컬 AI가 확대 보급되면 활동지원 시간이 부족한 당사자들이 온갖 가정용 보조기기를 한데 합쳐둔 것 같은 피지컬 AI를 활용할 수 있게 될까? AI 시대 완전 자율주행 저상버스는 버스정류장에서 기다리는 휠체어 이용자를 “못 봤다”며 지나쳐 가지 않고 램프를 내릴 수 있는 위치에 한 번에 정확하게 정차해 승차를 지원할까? AI 기술은 정말로 ‘장애를 마법처럼 소멸’시킬까?

장애인 지원 체계에서의 AI 활용

정부가 발표한 ‘3대 메가프로젝트’[반도체(AI의 두뇌), 피지컬 AI(AI의 신체), AI 데이터센터(AI의 심장)에 1500조 원 이상을 투자하겠다는 초대형 프로젝트.-1500조 원은 2026년 예산을 기준으로 492년 동안 장애인 활동지원과 교통약자 이동지원을 할 수 있는 금액이다]를 보면 정말로 장애 여부가 개인의 일상에 아무런 영향을 미치지 못하게 되어야만 할 것 같다. 정부가 AI에 막대한 중요성을 부여한 만큼 장애인 지원 체계에서도 다양한 방식으로 AI를 활용하고 있다.

서울, 경기, 대전은 특별교통수단에, 경남은 바우처택시에 AI에 배차 시스템을 도입했다. 아이디어는 지역마다 조금씩 다른데, 서울시는 실시간 교통흐름과 차량의 위치, 목적지 등을 종합적으로 분석해 가장 효율적인 차량을 매칭해 특별교통수단의 고질적인 문제로 지적되는 대기시간

을 줄이고자 한다. 대전시는 AI 데이터 분석을 기반으로 차고지와 차량 대기 장소를 결정해 배차 시 이동 거리와 대기시간을 감축했다고 한다. 경기도와 경남은 AI 상담원을 도입했다. 이용자의 최근 이용 내역을 파악해 출발지와 목적지를 예측해 배차 신청에 드는 시간을 줄인다는 계산이다.

SK텔레콤이 개발한 AI행동분석시스템 케어비야는 한국정보과학진흥협회가 주관한 ‘ICT어워드 코리아 2025’에서 AI 미래혁신 분과 대상(과학기술정보통신부 장관상)을 수상했다. 해당 기술은 발달장애인 평생교육 영역에 도입되었다. 발달장애인의 어려운 행동들을 촬영, 개인별 행동패턴을 분석해 자신 또는 타인에게 위험이 될 행동을 인식하고 사전에 예측하여 지원 인력에게 적절한 대응 방안을 제안한다.

교육부는 2025년 발달장애인 학생들의 학습을 지원하는 AI 교육자료를 19개 시범학교에 보급했다. 학생용 교육자료는 시각자료를 제공하고 버튼을 누르는 간단한 행동으로도 수업에 참여할 수 있으며 반복 학습이 가능한 앱 형태로 개발되었으며 교사용 플랫폼은 개별 학생의 흥미, 추가 학습이 필요한 영역을 분석해 개별 학생의 학습 속도를 세밀하게 관리할 수 있도록 정보를 제공한다.

한편 한국사회보장정보원은 AI를 이용해 사회서비스 사각지대를 탐지하고 부정수급을 탐지하는 시스템을 도입했다. 기존의 샘플링 기반 검사에서 AI 기반 이상거래탐지시스템(FDS)을 구축해 실시간으로 부정의심 결제를 모니터링하고 사후뿐 아니라 사전에도 부정수급 의심 사례 심사를 강화하겠다는 것이다.

효율적 관리와 비용 절감의 기술

자립생활을 확장하기 위한 장애인 당사자들의 개인적 AI 활용과 달리 지원 체계의 AI 적용은 장애인 지원에 드는 비용을 절감하고 장애인을 효율적으로 관리하기 위한 기술 적용으로 보인다. 전국장애인이동권연대가 2026년 6월 서울시 특별교통수단 API를 분석한 결과에 따르면 분석 기간 중 서울시에서 운행되는 840대의 특별교통수단이 모두 가동된 시간은 단 한 시간도 없다. 해당 기간 최대 대기시간은 약 3시간 20분에 이른다. 일상적 이동을 위해 차량 한 대를 신청하고 탑승할 때까지 3시간 20분이 걸린 것이다. 그런 와중에도 해당 시간 운행 차량은 422대로, 절반 가까운 차량이 차고지에 방치되어 있었다. 이는 특별교통수단을 운전할 운전원이 부족해서 일어난 일이다. AI 배차 시스템이 아니라 운전원 고용으로 해결했어야 하는 문제다. 경기도는 이번 추가경정예산에 약속한 특별교통수단 예산을 반영하지 않아 연말 특별교통수단 서비스 축소와 대기시간 대란이 예상된다. 대전시는 전국 특광역시자치시 중 유일하게, 특별교통수단의 법정 대수를 충족하지 못한 지역이다. AI가 아니라 지방자치단체의 예산 반영과 채용으로 해결해야 하는 절대적 서비스 부족 문제를 ‘AI를 통한 효율성 제고’라는 허울로 가리고 있다.

또한 AI 배차의 알고리즘을 담당자조차 알기 어렵기 때문에 이용자는 자기보다 늦게 신청한 사람이 먼저 배차를 받더라도 사유에 관한 설명을 들을 수 없다. 사후에 문제제기를 해서 오류를

밝혀내기 어렵고 밝혀진다 할지라도 침해는 이미 저질러진 뒤다. 사건이 발생하는 순간 당사자는 사정을 요구할 수 없고, 책임을 묻기도 어렵다.

AI행동분석시스템에서 이러한 문제는 더 큰 위협으로 작용할 수 있다. 먼저 센터 내에 설치된 카메라를 통해 이용자의 행동 정보가 실시간으로 촬영, 기록된다는 점에서 개인정보 침해 문제가 제기된다. 또한 기록된 정보는 알지 못하는 알고리즘에 의해 ‘위험할 수도 있는 문제 행동’ 또는 ‘문제 행동을 하기 직전의 행동’으로 분류된다. ‘전문가’에 의해 제안되는 대응 방안이 구금, 격리, 결박은 아니겠으나, 얼마나 위험한 행동으로 분류되느냐에 따라, 누가 프로그래밍하느냐에 따라 얼마든지 인권침해적 대응이 제안될 수 있다. 특히 행동을 하기 ‘전에’ 예측한다는 구상은 발달장애인이 하지 않은 행동을 이유로 행동을 제약받을 수 있게 된다는 의미에서 머리뼈의 형태로 범죄 가능성을 점쳤던 골상학을 떠올리게 한다.

마찬가지로 AI의 오류 가능성은 인권침해 발생 당시 확인할 수 없고 ‘위험행동 징후’로 분류된 행동이 왜 ‘위험행동 징후’로 분류되었는지 입증하기도 어렵다. 알고리즘이 불투명한 탓에 책임 소재도 불분명하다. 문제가 발생했을 때 책임은 ‘AI’에게 돌아간다.

충분한 시간을 들여 상호작용하고 당사자의 욕구를 파악해 행동의 원인을 확인하고 대체 행동을 연습할 시간과 비용을 절감하고, 적은 인원으로 효율적 관리를 달성하기에는 적절한 방안으로 보인다.

부정수급 탐지 시스템 역시 같은 위협을 내포하고 있다. 결제 패턴을 분석하는 시스템에 의해 부정수급을 의심받은 이용자는 부정수급이 아님을 증명해야 하나 증명을 요구하는 쪽은 어떤 사유로 부정수급을 의심하게 되었는지 설명하지 않고 “AI에 의한 탐지”만을 근거로 내세울 수 있다.

공통적으로 지원 체계에 도입된 AI는 무인서비스와 대량 정보 처리를 가능케 하고 AI 내부에서 일어나는 일이 밖으로 드러나지 않는다. 이는 동시에 차별이 발생했을 때 입증하거나 당사자의 사정권 보장을 어렵게 하고 개인정보를 마구잡이로 취득 기록 학습한다는 문제점으로 작용한다.

불평등한 해방의 도구로는 장애를 해방할 수 없다

그렇다면 기관의 AI 사용을 제한, 감시하고 장애인 당사자의 개인적 활용만을 권장하는 것이 해답이 될 수 있을까? 그간 새로이 등장한 많은 기술이 장애해방을 약속했고 그것들은 하나같이 불평등했다. 다양한 기능을 갖춘 개인 맞춤형 전동휠체어는 너무 비싸 필요한 사람 누구나 가질 수 없고, 전동휠체어로 독립적 이동이 가능해졌어도 가게 앞의 턱을 없앨 수는 없었다. AI 역시 앞선 많은 기술이 그랬듯 ‘장애를 소멸시킬’ 수는 없어 보인다.

노들장애인자립생활센터 김상희 소장은 2026년 장애학회 춘계학술대회에서 언어장애를 가진 사람들이 표준 발음을 구사할 수 있도록 지원하는 AAC가 말하는 사람보다는 듣는 사람을 위해

개발된 기술 같다고 언급했다. 발화자는 즉시 대화에 참여하지 못하고 따로 입력을 해야 하는 불편이 있음에도 듣는 사람에게만 편리하다는 것이다. 그는 또한 AAC 사용을 요구하면서 입력할 시간을 기다려주지 않는 문화도 문제라고 지적했다.

말하는 사람을 위한 AI 기술이라면 적어도 언어장애인의 말을 듣고 이용자의 말하기 습관을 분석해 입력을 용이하게 해주는 서비스 정도는 상상할 수 있을 것 같다. 그 경우에도 표준발음과 비표준발음 간의 위계를 강화한다는 혐의를 벗을 수 없겠으나, 준비할 시간이 없는 일상 대화에서도 소통의 편의를 강화할 수는 있을 것이다. 한국어를 서비스하지는 않지만, 표준발음을 사용하지 않는 영어 사용자의 말을 표준발음으로 바꾸어 들려주는 Voiceitt라는 앱이 실제 개발되어 서비스 중이다. 불편한 신체로 오랜 시간을 들여 직접 글을 입력하면 표준발음만 알아들을 수 있는 사람들이 이해할 수 있는 표준발음으로 읽어주는 앱은 진작 개발된 데 반해 여전히 언어장애가 있는 한국어 사용자의 말을 알아듣는 AI가 없는 현실은 어떤 기술을 개발할지 자체가 불평등하게 결정된다는 사실을 보여준다.

김상희 소장은 이어 기술이 비장애중심주의를 강화한다는 섬뜩한 지적을 내놓았다. AAC가 정확한 의사소통을 가능하게 해주는 한편 AAC를 사용할수록 사람들은 표준발음에 익숙해지고 그럴수록 언어장애를 가진 장애인이 직접 말할 기회가 줄어들지 않을까 우려된다는 것이다. AI는 많은 장애인 당사자가 독립적으로 할 수 있는 활동을 늘려주었지만, 이는 사회적 환경을 바꾸는 대신 AI를 이용해 장애인 당사자가 ‘비장애인처럼 기능할 것’을 기대하는 분위기를 형성할 수 있다. 장애차별적 사회 환경이 장애로 인한 불편을 만든다는 장애의 사회모델은 사라지고 장애를 개인이 치료하고 극복해야 할 것으로 보는 장애의 의료모델이 회귀한다. AI가 도입되면 장애가 소멸하리라는 확언은 AI를 이용해 장애를 소멸시키라는 말에 다름 아니다.

가격 역시 장벽으로 작동한다. 생성형 AI 구독료는 물론이고, Voiceitt, 릴루미노, SOVORO 모두 비용이 든다. 장애인이 비장애인과 같은 기능을 수행할 것을 요구받으면서 그러기 위해 유료 서비스를 이용해야 한다면 이 비용을 ‘장애비용’으로 부를 수도 있을 것 같다. 피지컬 AI가 도입된다면 ‘AI의 신체’는 이미 고가인 전동휠체어보다 훨씬 더 비싸, 최저임금 적용도 받지 못하는 장애인들에게는 없는 집을 팔아도 못 살 물건이지 않을까.

어떤 장애인은 돈이 있다 해도 AI를 활용하기 어렵다. 한국지능정보사회진흥원이 조사한 2025 디지털정보격차실태조사에 따르면 인터넷 사용자 중 심화 디지털 서비스[키오스크, OTT, 생성형 AI, 사물인터넷, 가상/증강현실, 블록체인 기반 서비스 사용을 포함한다]를 활용하는 비율은 일반 국민 84.9%, 장애인 65.1%로 장애인 심화 디지털 서비스 활동률이 19.8%p 더 낮게 나타났다. 장애인의 인터넷 이용률은 87%로, 일반국민의 인터넷 이용률 95%에 비해 7%p 이상 낮다는 점에서 장애인과 비장애인 간 실제 심화 디지털 서비스 활동률 격차는 드러난 통계보다 크다. 생성형 AI의 경우, 인터넷 사용자 일반국민 45.3%가 사용하고 있다고 응답한 반면 장애인은 19.6%만이 사용한다고 응답해 두 배 이상 차이가 났다.

즉 장애인을 위한 기술은 잘 개발되지 않고, 개발된다 해도 비싸고, 접근하기 위한 정보가 충분히 주어지지 않는 것으로 보인다. 따라서 AI로 ‘장애를 소멸’시키기 위해서는 활동지원용 피지컬 AI를 개발하라고 요구해야 하고, 피지컬 AI의 보조기기 급여화를 요구해야 하며, 급여량을 늘릴 것을 요구해야 할 것 같다. 마찬가지로 AI 완전자율주행 저상버스로 장애인 이동권을 보장하기 위해서는 AI 완전자율주행 차량을 개발할 때 휠체어 탑승설비를 포함하라고 요구해야 하고, 차량을 도입하려는 지자체에 휠체어 탑승설비가 포함된 모델로 도입하라고 요구해야 하며, 비용 지원을 함께 요구해야 할 것 같다. 그게 다 될 때쯤이면 비장애인들은 물통 속에 뇌를 넣고 전극만 꽂아둔 신인류로 탈바꿈해 버리는 게 아닐까.

AI, 해방도 통제도 아닌

AI는 장애를 마법처럼 소멸시킬 수 없겠지만, 우리 사회는 AI의 장애차별을 철폐해야 한다. AI가 제공하는 편리함이 장애인에게도 평등하게 제공될 수 있도록 기술 접근성을 강화해야 한다. 사용 과정에서 인권침해가 발생하지 않도록 살피고 공공의 책임하에 활용될 수 있도록 요구하며 당사자의 사정 권한을 보장해야 한다. 새로운 현장에 AI가 도입될 때마다 영향을 예측하고 우려사항을 점검해야 한다. AI가 학습할 데이터가 장애차별적인 내용을 포함하지 않도록 필터링하고 AI의 장애차별을 적발하고 수정해야 한다. 그리고 그 모든 과정에서 장애인 당사자가 참여할 수 있어야 한다. 이것은 장애운동이 새로운 기술이 도입될 때마다 해왔던 요구다.

AI는 당연하게도 마법이 아니고, AI 세계에서 장애인들은 장애인일 것이다. 그리고 그것이야말로 AI가 가장 온당하게 사용된 결과일 테다. 장애평등한 사회란 장애를 AI 마법으로 소멸시킨 사회가 아니라 장애를 가진 채로 AI를 잘 활용하며 살아갈 수 있는 사회일 것이기 때문이다.

연속 워크숍 6차: 사회복지 분야 인공지능 도입 현황과 과제 토론문

주장욱(홈리스행동)

1. 들어가며: 이미 ‘지원’이 아닌 ‘통제’가 기본값

2025년도 서울특별시의회 보건복지위원회 회의록(2025. 11. 10.) 일부 발췌(아래)

○강석주 위원 그리고 이건 경각심 차원에서 이야기합니다. 왜 그러냐면 본 위원이 생각할 때는 외국은 특히 일본 같은 데는 굉장히 잘 되어 있습니다. 그리고 지금 아시는 분은 아시겠지만 일본 같은 경우는 지방에 있는 노숙인이 동경으로 올라왔을 때 안 받아줘요, 자기 동네로 가라 그런다고. 그래서 또 노숙인들끼리도 마찬가지예요. 어디서 왔냐고 물어봐서 저기 하면 너네들 여기 못 들어와 이래서 그 노숙인들은 또 집단 노숙인 그런 시설이나 그다음에 자기네들 홈리스 그런 주거지로 가는 게 아니고 떠돌아 다녀요. 그러다가 결국은 안 돼서 자기 동네로 가거든요. 그런데 우리는 지금 일단 열차만 탔다 하면 다 서울이 종착역이라서 여기 오면 안 받아줄 수 없이 계속 오고, 지금 쪽방촌 같은 경우도 실질적으로 이 시설에 안 들어가지만 쪽방촌에 들어오는 노숙인들 월세를 사실 서울시에서 다 내주는 거 아닙니까? 내주고 있죠?

○복지실장 윤종장 네. 그건 이제 주거 급여인 경우에…….

○강석주 위원 그러니까 내주고 있는데 그것도 마찬가지예요. 그게 소문이 나니까 전반기 때 현장 방문 가서 이야기 들어보면 사유가 아주 조금 불분명한, 특히 여성 노숙인 같은 경우는 불분명해. 그런 것들이 호불호가 확실하게 구분이 안 되더라. 이런 것도 냉정하게 해 볼 필요성이 있지 않느냐 이렇게 생각하는데 어떻게 생각하세요?

○복지실장 윤종장 맞습니다. 지난 상임위에서 계속 위원님들도 노숙인 정책이 마냥 호혜적으로 가는 것은 맞지 않다, 과학적인 관리도 필요하다는 말씀에 동의를 저희들이 드렸고요. 그런 방향으로 가야 한다고 생각합니다. 지금 위원님 말씀하신 것처럼 저도 얼마 전에 일본을 다녀와서 오사카나 도쿄의 예를 같이 얘기도 하고 들어봤는데 그쪽은 철저하게 칸막이를 해서 더 늘어나지 않도록 그렇게 관리하고 있었고요. 저희들도 아마 그런 부분들은 관리를 해야, 그렇게 막아야 될 부분이 있고 전국에서 공통 기준을 통해서 같이 가야 할 부분도 있기 때문에 그런 것들 관리해 나가는 데 과학적인 방법이 될 수 있도록 그렇게 하겠습니다.

○강석주 위원 우리도 서울 보면 예를 들어서 AI 시대, 무슨 디지털 시대 해서 굉장히 시스템이 잘 되어 있는데 한 번은 손을 봐야 될 그런 때가 왔다고 생각합니다. 그래서 이건 말로만 할 게 아니라 실장님이 한번 기초작업을 해 놓는 게 좋겠다 이런 말씀을 드립니다.

2. 본론①: 복지의 디지털화에 대한 사례 중심 비판

1) 비대면 복지는 이미 ‘중첩된 취약성’을 배제하고 있다(참고: 홈리스의 실태)

- 보건복지부의 <2024년도 노숙인 등의 실태조사> 결과에 따르면, 시설노숙인¹⁾ 중

36.8%가 노인에 해당함. 전체 시설노숙인 중 등록장애인은 50.2%. 장애 유형 중에는 지적장애(20.0%), 정신장애(19.0%) 비중이 타 유형에 비해 압도적으로 높음. 학력은 ‘초졸 이하’가 26.6%, ‘중졸 이하’로 확대하면 47%. 오직 ‘일자리 복귀’를 통한 ‘사회 복귀’, 즉 ‘자활’을 정책 기조로 삼는 정부와 지자체의 홈리스 지원체계 안에서 연령, 장애, 질환, 문해력 등은 취약성으로 기능하고 이들이 ‘중첩’될수록 더욱 지원체계에서 멀어지게 되거나 시설에 장기 입소하게 됨. 실제로 홈리스행동과 밀접하게 활동하는 당사자 활동가와 야학 학생 중에는 ‘중첩된 취약성’을 가지고 있어 지원체계에서 배제당해 온 이들이 상당수임.

- 이들 당사자의 경우 이미 ARS(자동응답시스템) 이용 과정, 심지어 사람 상담사와의 통화에서조차도 어려움을 겪고 있음. ‘대면’하지 않는 상황에서 당사자가 겪는 ‘곤란함’은 유선으로 전달되기 어렵고, 상담사가 단 몇 초 안에 당사자의 특성을 파악해 그에 맞는 소통 방식을 장착하기란 더욱 어려움. **즉, 어떤 소통에는 ‘대면’과 ‘시간’이 필수적임.**²⁾

- 참고로, 홈리스가 겪는 ‘중첩된 취약성’에는 ‘지불 능력 부족’도 포함됨. 보건복지부 실태조사에 따르면, 거리노숙인 중 휴대폰(스마트폰, 피쳐폰) 소지자는 54.2%, 시설노숙인 중 소지자는 41.4%로 상당수가 비대면 소통 자체를 할 수 없는 조건에 놓여 있음.

2) 정보 수집에 동의하지 않으면 복지도 없다(참고: ‘따스한채움터’의 전자회원증 도입)

- 지난 2020년 코로나19 팬데믹 당시, 서울시는 “방역 관리”를 목적으로 시립 무료 급식장 ‘따스한채움터’에 전자 태그(RFID) 방식의 전자회원증을 도입하였음. 회원증을 만들기 위해서는 사진 촬영, 신분증 제출, 노숙 이력 조회에까지 응하여야 했음. 공공 급식시설의 태부족이 해당 조치가 발생한 근본적인 이유였으나, 서울시는 홈리스 당사자로 하여금 정보 인권과 식량에 대한 접근권을 저울질하게 만드는 것으로 ‘따스한채움터’에 이용자가 물리는 문제를 해결하고자 하였음.
- 실제 당시 당사자들은 “노숙했던 기록이 남을까 봐 노숙인 등록도 하지 않았는데, 전자회원증을 만들지 않으면 밥을 안 준다고 해서 노숙인 등록까지 하게 됐

- 1) 보건복지부는 노숙인자활·재활·요양시설에서 거주하는 ‘노숙인 등’을 ‘생활시설 노숙인’이라고 하고 있는데, 토론문에 한정하여 이들을 ‘시설노숙인’이라고 한다.
- 2) 국가인권위원회는 2024년 노인 인권에 관하여 과학기술정보통신부에 다음과 같은 권고를 내린 바 있다: **“현대 사회에서 정보접근권은 헌법에서 보장하는 다른 기본권 실현의 필수조건**이 되는 부분이 많으므로 단순 편의제공 또는 복지 차원이 아닌 기본적 권리 측면에서 정보접근권이 보장될 수 있도록 정부가 적극적 조치를 취하여야 한다. 이에 일상 생활에 필수적인 정보와 서비스 전달 방식에 있어서 디지털 방식(사회관계망서비스, 키오스크 등 온라인 방식)과 아날로그 방식(우편, 유선 안내, 대면 등 오프라인 방식)을 병행하여, 노인이 아날로그 방식을 통해서도 정보와 서비스에 접근할 수 있도록 선택의 기회를 부여함으로써, **디지털 환경에 익숙하지 않은 노인이 차별받거나 불이익을 겪지 않도록 ‘아날로그 접근권’을 보장할 필요**가 있다.”(출처: 국가인권위원회, <디지털 격차로 인한 노인의 인권 보호 제도개선 권고>, 2023. 12. 7.).

다”, “회원증을 가진 사람은 (회원증을) 찍고 나서 곧장 식당에 입장하지만, 없는 사람은 수기로 방역 명부를 작성한 뒤에도 한참을 대기실에서 기다려야 한다”라는 등의 반응을 보였음.³⁾ 디지털화를 위한 정보 수집 동의 여부와 관계없이 기본권은 모두가 차별 없이 보장받을 수 있어야 함.

3) 차별·혐오 해소 없이는 낙인에 기반한 기술만 나온다(참고: ICT콤플렉스 SW공모전)

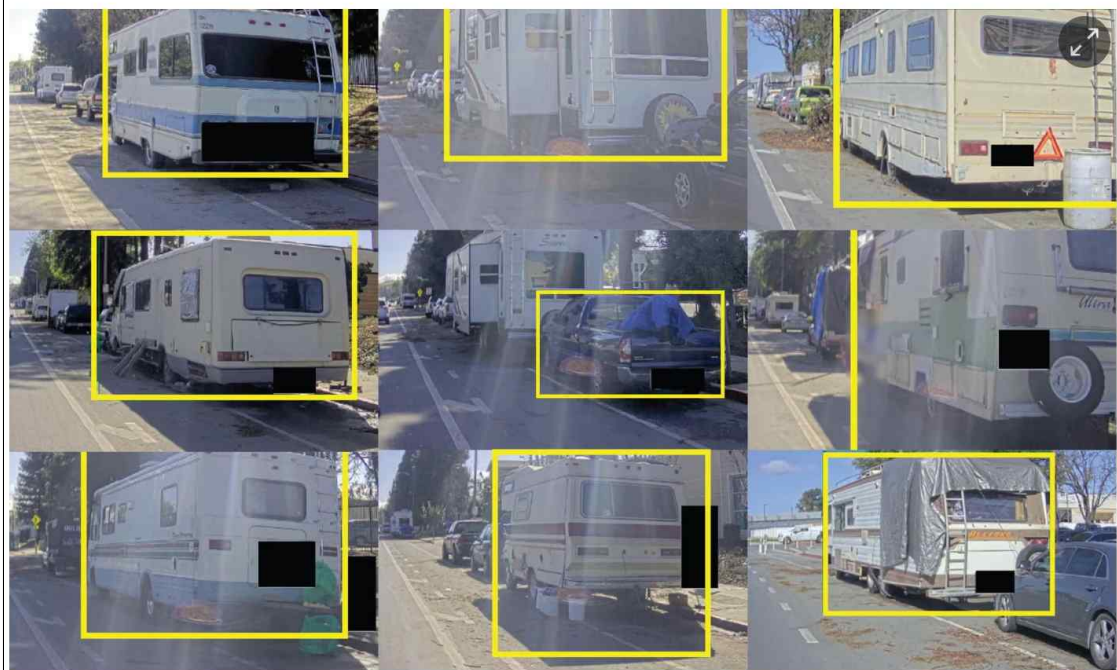
- 지난 2024년 과학기술정보통신부가 주최한 ‘2024 제12회 ICT콤플렉스 SW개발 공모전’에 홈리스에 대한 차별과 낙인에 기초한 소프트웨어들이 최우수상 등 수상하는 일이 발생함. “들꽃가드닝”은 시설 입소 홈리스의 외출, 외박 관리와 긴급 상황 대응을 위해 이들의 위치를 실시간으로 추적하는 기능이 있었고, 시설 입소자 관리를 위해 “센터 내 모든 노숙인의 상태를 한눈에 파악할 수 있는 대시보드”를 제공하였음. “R U OK?”는 홈리스 위치 파악 후 이를 시설 종사자 간에 공유하기 위해 당사자의 사진, 위치 정보 등을 당사자의 동의 여부와 상관없이 저장할 수 있는 소프트웨어였음. 이름 기록을 원하지 않는 홈리스의 정보는 “이름을 알 수 없는 노숙인 모임” 기능을 통해 저장하도록 만들어 둬. “노숙인 프로필 목록”, “노숙인 즐겨찾기” 등 기능도 존재. 단지 홈리스 상태에 처해 있다는 이유만으로 모든 행색과 행위가 기록되는 것.
- 홈리스 상태의 특성상 얼굴, 실명 등 특정 가능한 개인 정보가 아닌 정보라고 하더라도 여러 정보(예: 의복 색깔, 잠자리 위치 등)의 조합으로 충분히 개인이 특정 가능하다는 점에서 민감하게 다루어져야 함. 만약 수집된 정보가 홈리스 개인을 식별할 수준이 되지 않는다면 그 또한 특성에 맞는 지원을 하기 어려워지므로, 무엇보다 당사자의 상황과 동의에 따른 적절한 정보 수집과 신뢰가 중요함. 또한, 응급 상황 시 홈리스 개인의 위치를 빠르게 특정하여야 하는 주체는 화재 진압, 구급 등 전문 요원들임. 홈리스 개인에게 자신의 상황을 스스로 통제 및 결정할 수 있도록 연락 가능한 디지털 기기를 보급하는 등 개개인에게 자원을 마련하는 일이 더 필요함. 홈리스 상태를 자원과 권리 박탈의 상태로 이해하지 않고, 오히려 이를 ‘개인의 일탈’로 지속 이해하게 된다면 낙인에 기반하여 감시와 통제를 강화하는 방향으로 기술이 개발될 수밖에 없음.
- 한편, 이미 특정 정신질환과 그에 따른 행동이 기록되어 있는 홈리스의 경우, 해당 이력을 이유로 ‘임시주거지원’이 거부된 사례가 존재함. 이는 ‘주거’의 문제이기도 하지만, 홈리스 개인의 동의 여부와 무관하게 수집된 정보를 바탕으로 복지 연계 단계에서 종사자에 의한 임의적 판단이 이루어지고 있다는 것을 보여 줌. ‘기록’, ‘관점’이 모두 영향을 끼침.

3) 홈리스뉴스 82호, “방역 이유로 ‘따스한채움터’ 이용문턱 높은 서울시”(출처: 홈리스행동 홈페이지).

3. 본론②: 홈리스 지원체계 내 AI 도입에 대한 우려

- 디지털 기기 사용 등 기술에 대한 접근성 문제(2-1), 기술 사용 및 제도 이용 과정에서 배제되는 취약성 문제(2-1), 정보 수집이 복지 이용을 포함한 기본권 보장의 전제가 되는 문제(2-2), 결정권자들의 문제 인식과 관점이 바뀌지 않으면 결국 기술도 차별과 혐오를 답습할 것이라는 점(2-3) 등이 해소되지 않은 상태에서 AI를 도입하는 것은 오히려 홈리스를 더욱 소외시킬 것이 자명함.
- 특히 2-3과 관련하여, 아래 사진은 과거(2024) 주택 가격이 가장 비싼 도시 중 하나인 미국 캘리포니아주의 산호세시에서 시행한 시범사업의 하나로, 기업들로 하여금 거리, 공공장소에 주차된 노숙 차량을 포착해 내는 AI 기술을 개발하고 훈련시키도록 하였다고 함. 일반 차량과 주택으로 사용되는 차량을 구분하기 위해 “차량 내부 창문이 가려져 있나요?”, “사생활을 보호하기 위해 수건으로 내부를 가렸나요?”, “차량 외부에 이들이 음식을 먹거나 위생 관리가 어렵다는 것을 보여 주는 쓰레기가 있나요?”와 같은 기준을 바탕으로 노숙 차량을 감지해 내도록 하였다고 함. 지역 비영리단체는 이를 “집 없는 사람들을 쓰레기나 낙서처럼 흉물로 취급하는 것”이라고 평가했다고 함.⁴⁾

Revealed: a California city is training AI to spot homeless encampments



San Jose's foray into automated surveillance of homelessness is the first of its kind in the country, experts say. Illustration: Guardian Design

4. 나가며: 경험과 질문들

- 4) The Guardian, “Revealed: a California city is training AI to spot homeless encampments”, 2024. 3. 25.

- 경험적으로, 홈리스 개인에게 복지 서비스를 안내하고 신청과 이용을 돕기 위해서는 아웃리치(outreach)와 같은 현장지원 방식이 필수임. 이 과정에서 쌓인 신뢰와 서로에 대한 정보를 바탕으로 필요한 서비스를 요구하거나 개인 특성에 맞는 서비스를 제안하게 됨. 또한, 홈리스 상태는 매우 복잡한 상황이기 때문에, 장애, 중독 등 질환, 채무 등 각각의 문제를 다룰 수 있는 지원 인력이 함께 움직여야 함. 문제 파악이 됐다고 하더라도, 소득, 거처 등 자원이 부족하기 때문에 각각의 문제를 풀 수 있는 기관에 함께 밀착해 이동 및 방문하는 작업이 필요함. AI가 수많은 정보를 판단, 분석, 제공할 수 있다고 하더라도, ‘시간’, ‘신뢰’, ‘이동’과 같은 요소들 또한 가져갈 수 있을 것인가?

- 끝 -

토론문

박영아
공익인권법재단 공감

1. 들어가며

2025. 8. 보건복지부는 ‘AI 복지·돌봄 혁신 추진단’ 발족했다. 당시 보도자료¹⁾에 따르면 추진단은 ‘기획총괄반’, ‘복지행정·데이터혁신반’, ‘돌봄기술혁신반’과 사회적 논의의 장을 마련하는 ‘AI 혁신포럼반’의 4개 작업반으로 구성되었다. 이 중 ‘복지행정·데이터혁신반’에서는 “AI를 활용한 사각지대 발굴, 상담, 서비스 추천, 신청 지원 등 복지행정 전반의 혁신방안”과 “사회보장데이터 개방·활용 활성화 방안”을, ‘돌봄기술 혁신반’은 “돌봄기술 R&D 투자 확대, 기술 상용화를 위한 실증 및 제도적 지원 등 전주기 기술지원 방안”을 마련하기로 되어 있다.

보도자료에서 당면한 과제가 무엇이며, 인공지능이 과제 해결에 어떠한 기여를 할 수 있을지에 대한 방향성을 찾아볼 수 없다. 돌봄과 복지를, 인공지능을 활용하여 “혁신” 하겠다는 막연한 목표만 제시되어 있다. 주요 과제로 포함된 “사회보장데이터 개방·활용 활성화 방안” 마련은 정작 복지·돌봄 보장과 아무런 관련이 없다. 인공지능을 활용한 복지·돌봄 혁신의 목표가 복지와 돌봄의 보장이 아닌 “인공지능 활용” 그 자체가 아닌지, 다시 말해서 사회보장 분야에서의 인공지능 활용의 접근부터가 잘못된 게 아닌지 의구심을 가질 수밖에 없는 이유다.

2. 사회보장 분야에서 인공지능 활용 관련 국제적 움직임

지난 2021년 USP2030²⁾은 세계 각국 사회보장 체계의 전자화(digital transformation)와 interoperability(상호운용성) 지원 목적으로 디지털 융합 구상(Digital Convergence Initiative, DCI) 작업반(working group)을 발족하였다. 이후 2025년 DCI에 ‘AI Hub for Social Protection’이 신설되었다. 웹페이지에 따르면

1) https://www.mohw.go.kr/board.es?mid=a10503010100&bid=0027&act=view&list_no=1487161&tag=&nPage=1

2) 2030년까지 전 세계 모든 이들에게 빈곤, 질병, 재해 등 삶의 위험에 대응할 수 있는 '보편적 사회보장(Universal Social Protection)'을 제공하기 위해 결성된, 국제노동기구(ILO)와 세계은행(World Bank)이 공동 의장을 맡고 있으며, 각국 정부, 시민사회, 국제기구(UN Agencies, 개발은행 등) 참여하는 글로벌 파트너십

‘DCI AI Hub for Social Protection’의 역할은 사회보장 기관들이 인공지능을 관리, 설계 및 구현할 수 있는 역량을 키울 수 있도록 지원함으로써, 이러한 기관들이 인공지능의 혜택과 위험을 책임감 있고 혁신적으로 관리하고 디지털 주권을 수호할 수 있도록 돕는 데 있다.

AI Hub은 이후 ‘사회보장 분야 위험 평가 체계’³⁾, ‘사회보장 분야에서의 인공지능 분류체계’⁴⁾, ‘사회보장 분야에서의 인공지능: 데이터 주권 관련 고려’⁵⁾ 등의 지침 및 안내서를 발간했다. 인공지능이 사회보장의 어떤 국면에서 어떤 방식으로 활용될 가능성이 있으며, 이때 고려해야 하는 사항과 위험을 구체적으로 제시하고 있는 점에서 사회보장 분야에서의 인공지능 활용에 관한 이해도를 높이는 데 참고할 만하다고 생각한다. 이 글에서는 ‘사회보장 분야에서의 인공지능 분류체계’ 문서(이하 “DCI분류체계문서”)를 잠깐 소개할까 한다.

DCI분류체계문서는 사회보장 분야에서 활용되는 인공지능 시스템을 완결성 있고 체계적으로 분류하는 방법을 제시하는 데 의의를 두고 있다. 우선 인공지능 기능을 분석적 작업과 생성형 작업으로 분류한다. 이러한 분류는 인공지능의 유형에 따른 분류라기보다 생산물에 따른 분류라고 한다. 최근의 대규모 언어 모델(Large Language Model, LLM)들이 양 업무 모두 수행할 수 있기 때문이다. 분석적 작업으로는 예측/추정, 분류, 집단화, 이상 및 변화 감지, 비정형 자료로부터의 인식 및 추출(비디오, 오디오 및 텍스트 분석), 순위 및 결정 체계(무엇을 할 것인지 또는 무엇을 우선적으로 할 것인지 제시) 등이 있고, 생성형 작업으로는 같은 형식 내의 생성 및 전환, 서로 다른 형식 간 전환, 합성 데이터셋 생성이 있다고 한다.

문서는 이어서 인공지능이 활용되는 방식을 10가지로 분류하고 각각의 항목마다 권리 보장/침해 관련성 정도에 관한 설명을 부가하고 있다.

- (1) 예측적 분석을 포함한 취약성, 필요 및 위험 평가: 개별적 결정이 아닌 정책, 설계 및 운영 관련 결정에 참고할 수 있도록 빈곤, 취약성, 위험수준을 예측하고 그에 따라 인구집단을 구분하는 것이므로 권리에 대한 영향이 간접적이라고 한다.

3) DCI AI Hub for Social Protection, AI Hub Risk Assessment Framework for Social Protection, Deutsche Gesellschaft für Internationale Zusammenarbeit, Germany, 2026

4) DCI AI Hub for Social Protection, A Taxonomy for AI in Social Protection: Defining Capabilities and Application Areas, Deutsche Gesellschaft für Internationale Zusammenarbeit, Germany, 2026.

5) DCI AI Hub for Social Protection, AI in Social Protection: Data Sovereignty Considerations, Deutsche Gesellschaft für Internationale Zusammenarbeit, Germany, 2026

다만 취약성 평가가 개별적 자격여부 결정에 반영되는 경우 거버넌스 민감도(통제 체계 하자가 권리에 미치는 영향)가 고위험에 해당하게 됨 >> 우선적 개입 대상을 선별하는 위기가구 발굴시스템이 여기에 해당한다고 볼 수 있음

- (2) 경향 및 충격(shock) 예측: 기후, 분쟁, 경제위기 등을 예측하여 대응할 수 있도록 함. 주로 정보 제공용
- (3) 식별, 검증 및 데이터 연계: 권리 및 수급권에 직접적으로 영향을 미칠 가능성 있음. 오류는 잘못된 배제, 급여 거부 또는 잘못된 부정수급 신호 감지로 이어질 수 있음
- (4) 결정 지원: AI가 직접 자격판정 및 급여수준 결정을 지원하는 경우. 권리 및 수급권에 직접적 영향을 미치기 때문에 가장 강력한 수준의 인적 감독, 요건에 대한 검증가능성과 접근가능한 불복 장치를 요함
- (5) 데이터 품질 및 이상 감지: 권리 및 수급권에 직접적으로 영향을 미칠 가능성 있음. 특히 이상 신호가 급여 중지 또는 부정수급 조사와 같은 후속조치의 근거가 되는 경우 거버넌스 민감도가 높음
- (6) 사용자 소통 및 상호작용: 수급권을 결정하는 근거가 되지 않으나 챗봇 안내 오류, 잘못된 번역이나 정보제공이 수급권자의 행동 및 제도에 대한 신뢰에 영향을 미칠 수 있음
- (7) 매칭과 권유(고객서비스와 사례관리 지원): 중간 정도의 민감도. 권유는 서비스, 일자리 및 기회에 대한 접근성에 영향을 미치나 통상 사람에 의한 확인을 전제로 함. 매칭 결과가 케이스워커의 검토를 거치지 않고 직접적으로 프로그램 연계, 일자리 배치 또는 서비스 접근성을 결정하는 경우 거버넌스 민감도가 상승함
- (8) 운영 및 절차 자동화(생산성 향상, 지원부서 효율성 향상과 특정 기능 자동화): 문서 처리 및 생성 지원은 데이터 품질 및 정확성 위험을 초래하는 반면(특히 추출되거나 생성된 콘텐츠가 후속 의사결정에 반영되는 경우), 업무량 예측은 직접적인 권리 영향이 미미하다. 거버넌스 민감도는 자동화된 업무흐름 경로가 처리 시간에 영향을 미쳐 특정 그룹이 혜택을 받는 데 실질적인 지연을 초래할 때 더욱 높아짐 >> 사회보장정보시스템이 복지전담공무원의 업무처리를 사실상 지배하는 구조는 문제 없는가?
- (9) 준법 및 윤리성: 권리와 수급권에 직접적 영향을 미칠 가능성 있음
- (10) 정책 분석, 학습 및 모니터링·평가(편향성과 공정성 감지/평가 포함): 주로 내부적이며 실무적.

DCI분류체계문서는 나아가 위 10가지 활용 방식이 사회보장 정책 수립(법적 및 정책

적 기틀, 자원 마련, 조정/관리 및 기술적/기능적 역량), 프로그램 설계(급여 및 서비스 최적화, 자격 범주 및 요건 최적화, 지원의 수준, 가치 및 빈도 최적화 관련 결정), 및 집행(아웃리치, 소통, 인식제고 활동, 등록, 기여/청구/철회 관리, 필요/상태 평가, 가입, 급여 및 서비스 제공, 권리구제, 사례관리 등)의 각 단계별로 어떻게 적용될 수 있는지 제시한다.

또한 인공지능 활용에 수반되는 위험은 특정 단계에서 나타나기보다 시간이 지남에 따라 누적됨을 전제로, 인공지능의 생애주기를 다음과 같이 분류한다:

- (1) 과제의 식별 및 설계: 처음부터 인공지능으로 해결하고자 하는 문제, 예정된 사용자 및 맥락, 성공 기준을 정의함과 동시에 실현가능성, 위험 및 윤리적 함의 고려
- (2) 데이터 수집 및 전처리: 모델을 훈련시키기 위한 데이터의 수집, 정리, 분류와 관리. 데이터의 품질, 대표성 및 법적/윤리적 기준 준수 보장
- (3) 모델 선정 및 훈련: 적절한 알고리즘과 모델 설계를 선택한 후 모델을 훈련시키고 검증하여 성능을 최적화하면서 편향성, 과적합과 설명가능성 등의 문제 대처
- (4) 편입과 활용: 훈련된 모델을 기존의 체계 또는 업무흐름에 편입하여 실제 사용. 신뢰성, 보안성 및 사용자 상호작용 검증
- (5) 모니터링, 유지관리 및 운용중단: 성능과 영향을 지속적으로 추적, 조건 변화에 맞춘 모델 갱신 또는 재훈련, 더 이상 효과적이거나 적절하지 않을 경우 안전한 운용중단.

DCI분류체계문서는 인공지능 관련 위험(risk) 범주를 구조적 차원과 인적/사회적 차원으로 구분해서 분류한다.

(1) 구조적 차원

- 가) 데이터 관련 위험: 인공지능을 훈련, 검증 및 운용하기 위한 데이터가 실제 상황을 반영하지 않는 위험
- 나) 모델 관련 위험: 알고리즘의 설계, 목적 및 작동 관련 위험
- 다) 운용 및 체계 편입 관련 위험: 인공지능 알고리즘이 실제 인프라, 업무흐름, 데이터 파이프라인에 활용될 때 발생하는 위험. 이러한 위험은 담당자가 비판적 검토 없이 알고리즘 결과물에 의존할 때 가중됨
- 라) 거버넌스 및 제도적 감독 위험: 인공지능의 책임성, 성능, 감독 확보장치, 법적 안전장치 또는 효과적 통제가 결여되어 있을 때 발생하는 위험
- 마) 시장, 주권 및 산업구조적 위험: 공급자, 기반시설 제공자 그리고 세계적 기술 생태계에 대한 의존성에서 비롯된 위험

(2) 인적/사회적 차원

- 가) 사생활 및 개인정보 보호에 관련 위험
- 나) 평등, 비차별, 공정성 및 포용성에 관련 위험
- 다) 자율성, 존엄성 및 적법절차에 관련 위험
- 라) 책임성, 투명성 및 권리구제에 관련 위험
- 마) 구조적 및 사회적 위험

DCI분류체계문서는 인공지능이 사회보장 분야에서 어떻게 활용될 가능성이 있는지, 그리고 어떤 위험이 어떤 국면에서 발생할 가능성 있는지를 이해하는 데 지침을 제공하고, 관련 논의에 활용할 수 있는 언어와 범주를 제공하는 점에서 의미가 있다고 생각한다.

문서의 결론 중 다음 대목이 특히 눈에 띈다.

“실무적으로, 인공지능의 채용은 기술을 선택하는 데서 시작되는 것이 아니라, 해결하고자 하는 문제에 대한 명확한 정의에서부터 시작한다.”

“사회보장 제도에서 인공지능의 가치는 기술의 정교함으로 결정되는 것이 아니라, 사람들에 대한 서비스 전달을 향상시키는 데 기여할 수 있는 능력에 의해 결정된다.”

3. 사회보장정보시스템과 기초생활보장제도의 경직되고 획일화된 수급요건 및 행정실무로 인한 문제와 인공지능

현행 기초생활보장제도의 경직되고 획일화된 수급요건, 행정실무와 사회보장정보시스템 운용실무에 비추어볼 때, DCI의 분류체계 문서에서 제시된 구조적 위험, 그리고 인적/사회적 위험도가 매우 높다고 할 수 있다.

발제문에서 언급된, 가공자산으로 인해 수급권이 배제된 서울역 홈리스 사례의 경우처럼, 수급요건의 경직성 및 획일성은 사람이 실제로 처한 상황의 반영이 처음부터 배제됨을 의미한다. 그리고 발제문도 지적하는 것처럼 이는 요건만의 문제가 아니다. 요건 충족의 근거가 되는 사실, 즉 소득 및 재산 또는 그 부재를 입증할 수 있는 방법 역시 사회보장정보시스템으로 수집되는 “공적 자료”로 한정된다. 공적 자료가 실제 상황과 다를 때에도 공적 자료가 우선되고, 요부조자는 공적 자료를 수정하라는 요구를 받는다. 예를 들면, 근로소득을 파악하는 ‘공작 자료’로 사용자가 근로복지공단 등에 신고한 전년도 보수를 총개월수로 나눈 월평균 보수가 활용된다. 시간제 대

학강사처럼 보수가 해마다, 그리고 연중에도 수시로 줄거나 느는 등 일정하지 않은 경우에도 마찬가지다. 수급권자가 보장기관에 소득변경을 신고해도, 보장기관은 공적 자료를 수정해야 반영할 수 있다고 한다. 그러나 근로복지공단 등에 대한 법적 신고 의무자는 사용자이기 때문에 수급권자는 결국 사용자에게 수정신고 해달라고 해야 하는데, 사용자 입장에서 연중 소득 변경은 이듬해 연초에 정산하면 되는 일이기 때문에 그러한 요청에 응할 이유도, 의무도 없다. 인공지능이 도입되기도 전에 이미 사회보장정보시스템상 데이터가 실제 상황에 우선하는 것이다. DCI분류체계문서에서 얘기하는 구조적 위험 중 데이터가 실제 상황을 반영하지 않는 위험이 인공지능 도입 전 전자화(디지털화) 단계에서 이미 발생하고 있다.

구조적 위험 중 운용 및 체계 편입 관련 위험의 발생 역시 예견된다. 수급자격 및 급여를 결정하는 전담공무원이 인공지능이 생성한 결과물에 절대적으로 의존할 것으로 예상되기 때문이다. 공무원의 입장에서 인공지능 결과물에 대한 의존은 책임을 최소화하는 방법이다. 이는 위에서 본 것처럼 사회보장정보시스템상 정보를 우선하는 행정처리, 그리고 예를 들어 국민연금공단에 대한 근로능력평가 위탁의 경우처럼 일정 기능을 외부에 위탁했을 때 그 결과를 재검토하지 않고 결정에 그대로 반영하는 행정 실무에 비추어 볼 때 명백히 예상되는 바이다.

거버넌스 및 제도적 감독 관련 위험 역시 크다. 위기가구 발굴을 위해 수집되는 개인민감정보는 21개 기관 47개종인데, 수집되는 정보를 늘리기 위한 조치가 발빠르게 취해지는 반면, 이와 관련된 거버넌스 및 제도적 감독을 위한 조치에 대한 검토 및 논의는 뒷전이다. 공공부조 수급권이 시혜라는 인식, 폭넓은 행정재량과 법원의 소극적 사법심사 경향, 수급권자의 경제적·사회적 취약성으로 인해 권리구제가 활성화되어 있지 않고, 구제를 시도하는 경우에도 성공률이 높지 않은 점 또한 인공지능 도입 시 거버넌스 및 제도적 감독 관련 위험을 높이는 요인이 된다.

정부 쪽에서 최근 자주 언급되고 있는 “복지급여 자동지급”은 인공지능과 연결될 때 가장 위험도가 높은 ‘인공지능을 활용한 결정’에 해당하게 되어 거버넌스가 제대로 작동하지 않을 때 직접적 권리침해로 이어진다. 그러나 경직된 요건으로부터 시작되는 공공부조 전달체계의 문제점을 모두 신청주의 하나로 수렴시키려는 정부의 논의에서는 이에 대한 관점을 찾아볼 수 없다. 게다가 행정기본법 제20조는 “행정청은 법률로 정하는 바에 따라 완전히 자동화된 시스템(인공지능 기술을 적용한 시스템을 포함한다)으로 처분을 할 수 있다”고 규정한다. 자동화된 시스템을 통한 결정의 책임성에 대

해서는 침묵하고 있다. 단서로 처분에 재량이 있는 경우를 제외하고 있지만, 요건 해당 여부에도 판단이 개입될 수밖에 없다. 따라서 이 조항은 사문화되어야 할 것이나, 사문화되지 않을 경우 인공지능을 활용한 결정에 대한 거버넌스 및 통제 관련 위험을 더욱 높이는 독소조항으로 작동할 소지가 다분하다.

가장 근본적 문제는 서두에서 언급한 것처럼 해결하고자 하는 문제를 명확히 정의하기도 전에 인공지능 도입을 기정사실화하는 정책 당국의 접근이다. 이러한 접근은 문제 해결에 다가서지 못함은 물론, 사회보장 급여 전달체계 개선 등 당면한 과제보다 예산 절감, 효율화 그리고 “사회보장데이터 개방·활용 활성화 방안”을 마련하겠다는 보건복지부의 보도자료처럼 사회보장과 무관한 목적 개입으로 인한 왜곡, 그리고 책임성 및 투명성 결여로 권리구제는 물론, 제도개선을 더욱 어렵게 만드는 결과를 가져올 것이라고 하면 기우일까.

[토론문] 비장애중심으로 설계된 세계 위에 세운 AI는 무엇을 이어받는가

공익인권법재단 공감 조인영

1. 기술은 자기가 본 세계를 학습한다

AI는 무에서 판단을 만들어내지 않고 이미 존재하는 데이터에서 패턴을 찾아가 패턴에 가장 잘 맞는 답을 낼 뿐이므로, 어떤 세계가 그 데이터를 남겼는가가 결과를 결정한다. 미국 뉴욕시변호사회 대통령직속 AI·디지털기술 태스크포스가 2025년 6월 발표한 보고서 「AI 사용이 장애인에게 미치는 영향」은 그 세계를 하나의 숫자로 보여주는데, 2023년 기준 전 세계 상위 100만 개 웹사이트 홈페이지의 97%가 가장 기본적인 웹 접근성 기준조차 충족하지 못한다.¹⁾

접근성 없는 자료라도 AI에게 고쳐 달라고 하면 되지 않느냐는 반문이 나오지만, 보고서는 그렇게 되지 않는다고 말한다. AI가 사진 설명이나 입력창 표시를 만들어내는 해도 그것이 쓸모 있는지 판정할 기준이 없고, 좋은 설명이 무엇인지는 맥락이 정하는데 사람이 그 판단을 어떻게 해왔는지가 데이터에 거의 남아 있지 않기 때문이다. 그리하여 형식만 채워진 결과물이 쌓이고 그것이 다시 다음 AI의 학습 자료가 되는데, 더 나쁜 것은 아무것도 없을 때는 보이던 문제가 형식이 채워지면 보이지 않게 된다는 점이다. 같은 구조는 사람에 관한 판단에서도 반복된다. 채용 알고리즘은 장애인이 배제되어 온 노동시장의 역사를 정상 패턴으로 학습하고, 신용평가는 장애인이 축적할 기회를 갖지 못했던 금융 이력을 기준으로 삼게 된다.

2. 데이터를 더 넣으면 해결되는가

이에 대해 AI기술자들 사이에서는 장애에 대한 정보가 담긴 데이터를 더 넣으면 되지 않겠느냐는 처방들이 제시되기도 했다. 실제로 AI 편향에 관한 대부분의 논의가 이러한 방향으로 제시되기도 한다. 그런데 포용적 디자인 연구자 유타 트레비라누스(Jutta Treviranus, OCAD대학 포용적디자인연구센터장)는 장애에서는 이 처방이 원리적으로 작동하지 않는다고 말한다. 비례적 대표성이 완

1) New York City Bar Association, Presidential Task Force on Artificial Intelligence and Digital Technologies, The Impact of the Use of AI on People with Disabilities (2025. 6. 12.)

전히 확보되더라도 AI는 여전히 평균과 다른 것에 대해서는 불리하게 판단한다는 것이다.²⁾

트레비라누스가 수십 년간 "충분히 참여하고 살아가려면 무엇이 필요한가"를 물어 모은 데이터를 그래프에 찍으면, 80%가 중앙의 좁은 공간에 몰리고 나머지 20%가 주변의 넓은 공간에 흩뿌려지는 형태가 나타난다. 그가 '인간 성운'이라 부르는 모양이다. 장애인 사이의 공통점이라고는 평균에서 충분히 멀다는 것 하나뿐이라면, 그 집단은 애초에 하나의 집단이 아니다. 그래서 데이터를 늘리는 일은 중앙의 밀도만 높일 뿐 주변부를 중앙으로 옮기지 못한다.

트레비라누스의 결론은 "최적 패턴을 향한 편향은 곧 차이에 대한 편향이다. AI가 좋아질수록 차별도 정교해진다."는 것이다. 직관에 반하지만 논리는 단순하다. AI가 잘한다는 것은 평균적 패턴을 더 정확히 찾아낸다는 뜻이고, 패턴을 정확히 찾아낼수록 거기서 벗어난 것도 더 확실하게 걸러낸다. 기다리면 나아진다는 기대가 성립하지 않는다는 뜻이며, 이것이 이 논의의 가장 중요한 실천적 함의다.

이 구조는 발제문이 다른 세 시스템에 그대로 나타난다. 특별교통수단 AI 배차가 학습하는 효율적 배차는 평일 아침 집에서 복지관으로 향하는 것과 같은 다수의 이동 패턴이다. 야간에 친구를 만나러 가거나 주말에 처음 가보는 곳으로 향하는 이동, 즉 비장애인이 아무 설명 없이 하는 종류의 이동이 장애인에 대한 알고리즘에게는 예외로 취급된다. AI 행동분석 시스템에서는 어떤 발달장애인의 특정한 손동작이 그 사람에게 즐거움의 표현이라 해도, 유사 사례가 적으면 시스템은 그것을 기준으로 분류할 근거를 갖지 못한다. 개별성이 강한 사람일수록 위험하거나 예외적인 특성으로 분류될 확률이 높아지는 것이다. 부정수급 탐지 시스템에서도 마찬가지다. 활동지원사와 함께 장을 보느라 결제가 특정 시간대에 몰리거나 이동이 어려워 한 곳에서 몰아서 결제하는 사람은 그 자체로 패턴에서 벗어나므로, 부정수급을 하지 않았는데도 부정수급을 닮은 소비를 하게 된다.

세 경우 모두 평균에서 먼 사람일수록 불리해지는 구조이고, 데이터를 더 모으는 것으로는 교정되지 않는다. 그리고 이 구조는 다음 항에서 보듯 차별을 판정하는 단계에서 한 번 더 반복된다.

2) Jutta Treviranus, "The tyranny of the average", University Affairs (2026)

3. 차별을 판정하는 방식도 같은 함정에 있다

지금까지는 AI가 만드는 차별을 다뤘지만, 뉴욕시변호사회 보고서가 짚는 다음 문제는 한 단계 더 나아간다. AI의 차별을 막기 위해 만든 지침과 법제 역시 똑같이 통계에 의존하고 있어서, 차별을 만드는 도구와 그것을 감시하는 도구가 결국 같은 방식으로 사고한다는 것이다. 그렇다면 감시는 처음부터 특정한 종류의 차별을 볼 수 없게 되어 있는 셈이다.

어떤 채용 알고리즘이 차별적인지는 대개 이렇게 판정한다. 통과율이 남성은 60%인데 여성은 30%라면 격차가 크고 표본이 충분하므로 차별을 인정할 수 있다는 것인데, 뉴욕시가 채용 알고리즘에 의무화한 편향감사도 우리 법의 간접차별 판단도 기본 구조는 다르지 않다. 개인이 자기 사건에서 차별 의도를 입증하기란 사실상 불가능하니 집단 간 결과의 격차로 대신 증명하게 해주는 장치이고, 애초에 차별 피해자를 돕기 위해 고안된 방법이다.

이 장치가 장애에서는 작동하지 않는다는 지적은 새로운 것이 아니다. IBM의 접근성 연구자였던 샤리 트레윈(Shari Trewin)은 2018년 「장애인을 위한 AI 공정성」에서 그 근거를 장애인권리협약의 장애 정의에서 찾았다.³⁾ 협약은 장애를 사람의 속성이 아니라 손상을 가진 사람과 환경적·태도적 장벽 사이의 상호작용에서 발생하는 것으로 보는데, 그렇다면 장애는 몇 개의 값을 갖는 단순한 변수가 아니라 여러 차원을 가지며 강도가 다르고 시간에 따라 변하는 것이 된다. 자폐 당사자 사회에 "한 명의 자폐인을 만났다면 당신은 한 명의 자폐인을 만난 것"이라는 말이 있듯, 동질해 보이는 집단 안에서도 사정은 같다.

그 결과 차별이 실제로 있었고 그것이 알고리즘의 설계에서 비롯되었더라도, 판정 단계에서는 개별적 일화이거나 통계적으로 유의미하지 않은 것으로 정리된다.

4. 이것은 새로운 문제가 아니다

발제문은 그간 새로 등장한 많은 기술이 장애해방을 약속했고 그것들은 하나 같이 불평등했다고 썼다. 이 진단은 이미 학술 개념으로 정식화되어 있는데, 애슐리 슈(Ashley Shew, 버지니아공대)가 명명한 테크노에이블리즘

3) Shari Trewin, 'Considerations for AI fairness for people with disabilities'(2019)

(technoableism) 이 그것이다.⁴⁾ 기술이 장애의 해결책이라는 믿음, 장애인은 기술의 마법으로 고쳐지기를 기다리는 존재라는 믿음, 그리고 사회를 접근 가능하게 만드는 일은 그보다 부차적이라는 믿음이 하나의 세트로 작동한다는 것이다. 앞의 두 믿음이 세 번째를 정당화하므로, 테크노에이블리즘은 기술에 대한 과잉기대라기보다 책임을 옮기는 장치에 가깝다. 이 장치는 두 방향으로 작동한다.

첫째, 사회가 바뀌어야 할 문제가 기술을 쓰지 않는 개인의 문제로 환원된다. 슈가 드는 근거는 발제문의 장애비용 논의와 겹친다. 가장 단순한 무릎 아래 의족조차 8,000~16,000달러이고 평생에 걸친 조정과 교체가 따르는데, 보험이 전부를 부담하지 않으니 가난한 사람은 그냥 없이 산다. 기적적인 기술이 존재한다는 사실과 그것을 가질 수 있다는 사실은 전혀 다른 문제다. 그런데도 기술이 있다는 사실만으로 세계는 조정될 이유를 면제받고, 남은 불편은 그것을 갖추지 못했거나 잘 다루지 못하는 개인의 몫이 된다. 그래서 슈는 오히려 휠체어야말로 나은 이동 방식이라고 말하는데, 휠체어는 세계가 장애인에게 맞춰 조정되기를 요구하는 물건이기 때문이다. 발제문이 전동휠체어로 독립적 이동이 가능해졌어도 가게 앞의 턱을 없앨 수는 없었다고 쓴 것이 이 이야기다.

여기서 새로운 위계가 생긴다. 기술을 쓸 돈과 정보와 기기를 갖춘 장애인과 그렇지 못한 장애인이 나뉘고, 후자는 지원을 받지 못하는 이유조차 개인의 책임이 될 수 있다. 발제문이 인용한 디지털정보격차 통계, 즉 심화 디지털 서비스 활용률이 일반국민 84.9%인 데 비해 장애인은 65.1%에 그친다는 수치가 위험한 이유가 여기 있다. 이 격차는 흔히 교육과 보급을 늘리자는 결론으로 회수되지만, 그 결론은 격차의 원인을 개인의 역량에 두고 사회가 바뀌어야 한다는 요구를 다시 뒤로 미루게 된다.

둘째, 기술이 기존의 것을 대체할 수 있다는 믿음이 원래 있던 권리를 박탈할 수 있다. 마르티어 더뮐더(Maartje De Meulder)의 2025년 논문 「AI 속의 농인: AI 언어기술과 언어권의 침식」이 이 문제를 정면으로 다룬다.⁵⁾ 그는 수어 아바타와 자동 번역 인터페이스를 검토하면서 AI 도구가 혁신을 약속하는 한편 편향을 지속시키고 테크노에이블리즘을 강화하며 설계상의 결함을 통해 불

4) Ashley Shew, Against Technoableism: Rethinking Who Needs Improvement (W. W. Norton, 2023)

5) Maartje De Meulder, Deaf in AI: AI language technologies and the erosion of linguistic rights(2025)

평등을 심화시킨다고 지적하는데, 그가 든 사례 하나가 시사적이다. 영국수어 대규모언어모델을 만들어 세계 최대의 수어 데이터셋을 구축하겠다는 845만 파운드 규모의 사업이 시작되었는데 연구책임자가 전원 청인 학자였다.

핵심은 프레이밍이다. 이 기술들은 본래 인간 통역사와 경쟁하기 위한 것이 아니라 농인 사회의 다양한 언어 실천을 지원하고 반영하기 위한 것인데, 대체물로 규정되는 순간 접근성의 위계를 강화하고 기술에 대한 새로운 형태의 언어적 종속을 만들어낸다. 그래서 농인 이용자와 수어통역사와 농인 단체의 우려는 누가 이 기술을 만드느냐를 넘어, 그것이 쓸 만해졌을 때 어떻게 사용될 것이며 그로 인해 인간 통역 서비스에 뿌리를 둔 어렵게 얻어낸 언어권과 접근권이 침식되지 않겠는가에 놓인다.

우리가 AI기술로서 이루고자 하는 가치와 목적이 명확하지 않으면, 그토록 오랫동안 싸워왔던 장애에 대한 의료적 모델, 개별적 모델이 또 다시 등장하게 되며, 더 공고해질 수 있다.

문제는 기술 자체가 아니라 누가 무엇을 만들지 어떻게 활용할지를 정하는가이다. 새로운 기술이 등장할 때마다 장애운동이 같은 요구를 반복해 온 것은 매번 같은 일이 벌어졌기 때문이다.

5. 그렇다면 무엇을 요구할 것인가

첫째, 인간에 의한 결정을 요구할 권리. 누가 고용되고 입학하는지, 누가 서비스를 받는지, 누가 감사 대상이 되는지, 누가 위협으로 간주되는지, 어떤 의료적 도움을 받을 자격이 있는지에 대한 AI 결정, 그리고 자막 같은 AI 서비스 모두에 대해서 인간이 제공하는 서비스를 받을 권리에 대한 요구이다. 사람은 자신이 평균이 아니고 학습 데이터에 공정하게 대표되어 있지도 않으므로 AI가 자신에 대해 공정한 결정을 내릴 수 없다고 말하고 사람의 판단을 요구할 수 있어야 한다. 이때 인간 대안은 AI서비스와 편의성·비용·소요시간에서 동등해야 한다. 신청은 가능하지만 몇 주가 걸리고 별도 비용이 든다면 아무도 쓰지 못하는 권리가 되기 때문이다.

둘째, 도입 전 영향평가와 실질적 인적 검토의 법정 의무화. 사후 구제로는 해결하기 어렵고 수정하기 어려우며 책임을 무기도 어렵다. 뉴욕주는 2024년 12

월 LOADinG Act를 제정해, 주 기관이 공적부조 급여와 개인의 권리·자유에 영향을 미치는 업무에 자동화 의사결정 시스템을 쓰는 것을 법률상 근거 없이는 금지하고, 2년마다 목적·알고리즘·정확성·공정성을 평가하며, 차별적·편향적 결과가 확인되면 사용을 중단하도록 했다. 여기서 실질적 인적 검토는 시스템의 위험과 한계와 기능을 이해하고 훈련받았으며 결정을 승인·거부·수정할 실제 권한이 있는 사람의 검토로 정의된다. 발제문이 지적한 "담당자조차 알고리즘을 알기 어렵다"는 상태가 바로 이 법이 금지하려는 상태이기도 하다.

셋째, 장애 데이터의 가중치와 문서화, 그리고 법제 검토. 단순히 장애 데이터를 포함시키라는 것이 아니라 과소대표된 경험에 상응하는 가중치를 부여하고 장애인 고유의 사례에는 양의 가중치를 주어 다수 중심 모델에 의한 배제를 막으라는 요구이다. 이는 데이터의 단순 증량으로는 해결되지 않기 때문에 나온 처방이다. 그리고 이와 연결되어 제공자가 자기 시스템이 장애를 어떻게 다루는지에 관한 데이터 출처와 설계 결정을 문서화하고 접근 가능한 이의제기 절차를 갖추도록 해야 하고, 장애인에게 발생하는 고유하고 개별적인 형태의 피해가 통계적 척도에 의해 일화적이라며 기각되지 않도록 차별금지법제를 검토하고 강화할 필요가 있다.

넷째, 예측에 근거한 제약의 금지선. 발제문이 골상학을 떠올린 것은 정확한 직관이다. 골상학의 문제는 예측이 틀렸다는 것만이 아니라 예측만으로 사람을 다르게 대우해도 된다고 믿었다는 데 있다. 다만 예측 기술의 전면 금지를 요구하면 위기행동 지원의 현실적 필요라는 반론에 부딪히므로, 실무적으로 제시할 수 있는 것은 예측 결과가 무엇을 정당화할 수 있는가를 제한하는 것이다. 발생하지 않은 행동을 근거로 제약적 조치를 할 수 없어야 하고, 어떤 개입이든 사후 기록 열람과 분류 근거의 설명, 이의제기와 중단 청구가 보장되어야 한다.

다섯째, 개발 단계의 참여를 조달로 강제하기. 한국어 언어장애 음성인식이 없는 것은 기술의 한계가 아니라 무엇을 개발할지가 불평등하게 결정된 결과다. 이 불평등은 시장이 만든 것이지만 시장 논리로는 교정되지 않는다. 그러므로 요구는 개발을 기다리는 것이 아니라 개발을 발주하게 만드는 쪽이어야 하고, 그 지렛대가 공공조달이다.

6. 맺으며

우리가 지나온 세계를 바꾸지 않는다면 아무것도 바뀌지 않을 것이다. AI를 규율하는 일과 AI가 학습하는 세계를 바꾸는 일이 함께 가야 한다는 것이다. 접근할 수 없는 웹페이지, 장애인이 배제되어 온 채용 기록, 휠체어 이용자를 태우지 않고 지나간 버스들의 데이터 위에서 만들어진 기술은 그 역사를 학습하고 재생산한다. 기술이 중립적이지 않은 이유는 개발자가 악의적이어서가 아니라 기술이 배우는 세계가 이미 차별적이기 때문이며, 그래서 AI에 대한 요구는 결국 장애계의 오래된 요구와 같아질 수밖에 없다.

AI 돌봄에서 나타나는 편익의 개인화와 부담의 사회화

예명대학원대학교 사회복지전공 교수

권진

들어가며: '돌봄 공백'을 메우는 AI?

이른 아침, 혼자 사는 노인의 방에서 기계가 먼저 말을 건넨다. 약 드실 시간이라고, 오늘 기분은 어떠냐고. 몇 해 전만 해도 낯설던 이 풍경은 지금 전국 곳곳에서 현실화 되어 있다. 사람이 미처 채우지 못한 돌봄의 빈자리를 기계가 먼저 메우기 시작한 것이다. 물론 그 배경에는 익숙한 사정이 있다. 2025년 한국은 초고령사회에 들어섰고, 그 속도는 세계에서 가장 빠르다. 돌봄 수요는 가파르게 늘어나는데, 요양보호사의 대규모 부족은 이미 예고되어 있다.

이러한 돌봄 공백을 메우는 유력한 해법으로 인공지능(AI)과 돌봄로봇이 떠오르고 있다. 정서지원 로봇 '효돌'·'다솜', AI 스피커·말벗 챗봇, 배설·이송(移乘) 보조 로봇, 응급안전 모니터링 시스템이 이미 지자체와 요양 현장에서 각광을 받고 있다. 그러나 기술이 들어오는 속도에 비해 "이것이 좋은 돌봄인가"라는 질문은 충분히 던져지지 않았다.

이 글에서는 AI 돌봄에서 편익은 개인화되는 반면, 그 부담—돌봄 노동의 재편, 접근의 격차, 그리고 데이터센터가 소비하는 전력·물·탄소—은 사회화·외부화된다는 점을 논한다. 편익과 부담이 이렇게 비대칭적으로 분배될 때, 이러한 부담은 언제나 아래로 흐른다. 노인에게, 요양보호사에게, 비수도권 주민에게, 그리고 미래세대에게로 흐른다.

그 흐름을 세 층위로 나누어 살펴보고자 한다. 부실한 토대(통합돌봄 전달체계) 위에 기술(돌봄로봇·생성형 AI)이 얹히고, 그 기술은 보이지 않는 하부구조(데이터센터)에 기대며, 세 층 모두에서 부담이 아래로 전가되는 구조라는 것이다. 돌봄 논의의 준거로는 기술철학자 마크 코켈버그의 돌봄 기술론을 삼는다. 그의 결론인 '로봇 윤리는 결국 로봇에 관한 것이 아니라 인간과 좋은 삶에 관한 성찰'이라는 점은, 좋은 돌봄의 문제를 '누가 영향을 받고, 누가 그 미래를 함께 쓰는가'라는 물음으로 되돌려 놓기 때문이다.

국내에서도 AI·로봇 돌봄의 윤리를 다룬 논의가 나타나고 있다. 돌봄윤리의 관점에서 자기결정권과 효율성, 안전 모니터링과 프라이버시, 인간관계와 기술 의존, 접근의 불평등이라는 윤리적 딜레마를 짚고 제도화를 제언한 연구, 돌봄로봇과 인간 돌봄이 공존할 조건을 묻은 연구가 대표적이다. 본고는 이러한 선행연구의 진단—특히 관계·프라이버시·불평등을 둘러싼 딜레마—을 상당 부분 공유한다. 다만 기존 논의가 주로 노인과 로봇이 만나는 '돌봄 상호작용'의 층위에 집중했다면, 본고는 그 아래의 부실한 전달체계(토대)와 눈에 보이지 않는 데이터센터(하부구조)를 함께 시야에 넣어, 문제를 '편익은 개인화되고 부담은 사회화된다는' 분배 정의의 물음으로 넓혀서 살펴보고자 한다.

논의의 준거: AI 시대의 '좋은 돌봄'을 묻는 다섯 개념

돌봄 기술을 다룬 논자는 적지 않지만, 이 글이 코켈버그를 준거로 삼는 데는 이유가 있다. 그는 돌봄로봇과 정보통신기술 기반의 노인 돌봄을 정면으로 파고든 몇 안 되는 기술철학자이면서, 논의의 자리를 '기계냐 사람이냐'를 따지는 데서 '무엇이 좋은 돌봄인가'를 묻는 것으로 전환시켰기 때문이다. 그는 돌봄로봇을 둘러싼 논쟁이 흔히 두 극단—기술이 모든 것을 해결한다는 유토피아

와, 기계가 인간을 대체해 돌봄이 완전히 비인간화된다는 디스토피아- 사이에서 오간다고 지적하며, 어느 쪽도 현실을 제대로 포착하지 못한다고 하였다. AI 시대에 ‘좋은 돌봄’을 묻는 일은 바로 이 소모적 이분법에서 벗어나는 데서 시작한다. 코켈버그의 논의에서 그 물음에 특히 유용한 다섯 개의 개념을 추리면 다음과 같다. 이 개념들은 이후 세 층위를 평가하는 척도로 등장한다.

- 1) 기만(deception)의 재해석: 스페로우 등 비판자들은 돌봄로봇이 노인을 기만한다고 본다. 로봇이 마치 감정을 가진 동반자인 것처럼 행세하여 노인에게 ‘가짜 세계’를 제공한다는 것이다. 코켈버그는 이 논리에 동의하지 않는다. 돌봄에서 일정한 ‘선의의 기만’은 반드시 부도덕한 것이 아니며, 인지 저하가 있는 노인을 대할 때 돌봄 제공자는 이미 일상적으로 자율성 존중과 당사자의 안녕(well-being) 사이에서 균형을 잡고 있다는 것이다. 핵심 질문은 ‘로봇이 기계임을 숨기는가’가 아니라 ‘로봇과의 상호작용이 사람의 삶의 질과 존엄을 지키는가’이다.
- 2) 실재 대 가상 이분법의 해체: 코켈버그는 로봇 돌봄을 ‘가짜 세계’로, 인간 돌봄을 ‘진짜 세계’로 나누는 구도 자체를 비판한다. 플로리다의 ‘온라이프(onlife)’ 개념을 토대로, 우리가 이미 온라인과 오프라인을 오가며 살고 있음을 지적한다. 기술로 매개된 돌봄은 덜 실재하는 돌봄이 아니라 다른 방식의 실재다. 따라서 문제는 매개 여부가 아니라 매개의 질이다.
- 3) 미래의 노인은 디지털 원주민: 오늘의 초고령 노인을 기준으로 로봇 돌봄의 수용성을 논하는 것은 오류라고 그는 말한다. 지금의 중장년, 나아가 우리 자신이 미래의 노인이며, 이들은 오히려 기술의 부재를 불이익으로 여길 수 있다. 이는 기술 도입 자체를 반대하는 논의가 왜 종종 초점을 빗나가는지를 설명해 준다.
- 4) 단절(disengagement)이라는 진짜 문제: 코켈버그가 기만보다 더 주목하는 위험은 단절이다. 보그만의 ‘장치 패러다임(device paradigm)’을 인용하여, 너무 손쉬운 기술은 우리의 관여·숙련·주의를 요구하지 않은 채 배경으로 물러나면서 사람과 사물로부터 우리를 떼어낸다고 본다. 중요한 것은, 이 단절이 돌봄 수혜자뿐 아니라 돌봄 제공자에게도 일어난다는 점이다. 효율 중심의 근대적 돌봄 문화 속에서 요양보호사는 이미 의미 있는 관계를 맺을 시간을 빼앗기고 있으며, 로봇은 이를 완화할 수도, 악화시킬 수도 있다.
- 5) 돌봄 장인정신(care craftsmanship)과 역량 접근: 코켈버그는 좋은 돌봄을 숙련된 관여를 요하는 장인적 실천으로 규정하고, 누스바움의 역량(capability) 접근을 끌어들인다. 좋은 돌봄이란 삶·건강·정서적 애착·사회적 유대·놀이 등 인간의 핵심 역량을 유지·증진하는 것이며, 로봇은 이 역량을 확장할 때에 비로소 가치롭다. 요컨대 로봇 윤리는 결국 로봇에 관한 것이 아니라 인간과 좋은 삶에 관한 성찰이며, 로봇은 인간을 비추는 거울이라는 것이 결론이다. 이 다섯 개념이 겨누는 질문은 하나로 모인다. 우리는 로봇이나 사람이나가 아니라 이 기술이 어떤 관계·역량·존엄을 만들어 내며, 그 편익과 부담은 누구에게 돌아가는가를 물어야 한다. 이 물음을 다음의 세가지 층위에 차례로 적용해보자.

토대: 부실한 통합돌봄 전달체계

AI 돌봄 플랫폼을 논하기 전에 확인해야 할 사실이 있다. 현재 한국 사회는 사람에 기반한 돌봄의 토대부터 부실하다는 것이다.

「의료·요양 등 지역 돌봄의 통합지원에 관한 법률」(이하 돌봄통합지원법)은 2026년 3월 27일 전면 시행되었다. 신청·발굴 → 조사·판정 → 개인별 지원계획 → 서비스 연계 → 모니터링으로 이어지는 5단계 전달체계를 지방자치단체가 직접 가동해야 하는 구조다. 방향은 옳다. 흩어진 의료·요양·복지를 개인 단위로 통합해 살던 곳에서의 노후를 지향한다는 취지다.

그러나 시행 초기의 신호는 우려스럽다. 첫째, 전담 인력의 공백이다. 국정감사에서는 통합돌봄을

많은 일선 공무원의 대다수가 기존 업무를 함께 떠안은 겸직 상태이고, 순수 전담 인력은 극소수에 그친다는 사실이 드러났다. 읍·면·동마다 최소 한 명을 두겠다던 공언과 달리, 시행 너 달이 지나도록 전담 인력을 한 명도 두지 못한 시·도가 여럿이다. 사람이 없는 전달체계는 문서 위의 체계일 뿐이다. 둘째, 재정의 미비다. 첫째 예산이 제도를 감당하기에 크게 부족하다는 비판이 시민사회와 장애계에서 이어졌다. 충분한 재정 없이는 통합돌봄이 현장에서 안정적으로 작동하기 어렵다는 것이다. 셋째, 인지도와 지역 격차다. 시행 100일이 지나도록 제도의 존재 자체를 모르는 국민이 상당수였고, 지역별 신청률의 편차도 뚜렷했다. 지방선거와 시기가 겹치면서 행정의 집중도마저 떨어졌다. 넷째, 구조적 난제다. 장기요양은 건강보험공단이, 의료는 민간이 주도해 상당 부분이 지자체의 통제 밖에 있고, 행정체계도 중앙집권적이다. 결정적으로 의료·요양·복지의 데이터가 제각기 다른 체계로 쌓여 있어, 정작 통합의 기반이 되어야 할 데이터 상호운용성이 확보되지 않았다. 시행 초기부터 제도의 정착을 걱정하는 목소리가 이어진 것도 무리가 아니다. 정부는 이러한 부실한 토대 위에 데이터 통합 기반의 AI 돌봄 플랫폼을 추진하고 있다. 그러나 전담 인력도, 예산도, 데이터 표준도 갖추지 못한 토대 위에 얹힌 기술은 통합의 실현과는 거리가 멀다. 코켈버그의 관점에서, 이는 관계적 돌봄의 단절을 기술로 메우려는 시도이며, 그 부담은 언제나 현장의 사람—특히 겸직에 내몰린 공무원과 요양보호사—에게 고스란히 무겁게 지워진다.

기술: 퍼지컬 AI와 생성형 AI

토대 위에 얹히는 기술은 성격이 다른 두가지 유형으로 나눌수 있다. 신체·공간에 작용하는 퍼지컬 AI(돌봄로봇)와, 언어·정서에 작용하는 생성형 AI(말벗 챗봇·대화형 시스템)다. 둘은 편익도 위험도 다르므로 분리해서 보아야 한다.

1) 퍼지컬 AI(돌봄로봇)

효돌·다솜 같은 정서지원 로봇이 지자체를 통해 독거노인에게 보급되어 약 복용·식사·운동을 알리고 말벗이 되며, 낙상 등 이상 징후를 관제센터로 연결한다. 지자체마다 보급이 늘었지만, 그 상당수는 구매가 아니라 몇 년 단위의 단기 대여여서 예산이 끝나면 회수된다. 몇 해 동안 인형에 이름을 붙이고 옷을 지어 입히던 노인이 사업 종료와 함께 그 인형을 도로 내주어야 하는 일이 반복된다는 지적도 나온다. ‘몇 대를 깔았는가’라는 성과 지표가 오히려 돌봄의 연속성을 해치는 셈이다. 응급안전안심서비스는 전국의 많은 독거노인 가구에 IoT 기기를 설치해 움직임과 응급 상황을 감지하고, 배설·이송 보조기기가 실증·상용화 단계에 있다.

효과는 일부 확인된다. 정서지원 로봇을 도입한 지역에서 노인의 우울감이 뚜렷이 줄었다는 실증 보고가 있고, 이용 노인 다수가 ‘혼자가 아니라는 느낌’을 되찾았다고 말한다. 그러나 정반대의 신호도 있다. 로봇을 손주처럼 대하며 정서적으로 과의존하는 경향, 그리고 ‘사람보다 낫다’는 진술이 역설적으로 드러내는 인간관계의 공백이다. 코켈버그의 틀로 보면, 보급이 단절을 완화하는지 심화하는지에 대한 평가 없이 보급 대수 중심으로만 진행된다는 점이 가장 큰 문제다.

더 근본적인 것은 접근의 불평등이다. 예비급여 시범사업에서 효돌·다솜이 첫 급여 대상이 되었지만, 적용 지역이 일부에 한정되었고 기초생활수급자는 신청 대상에서 아예 빠졌다. 요양시설의 로봇 구입비를 정부가 직접 보조하는 일본과 달리, 한국은 급여·보조의 윤곽이 여전히 흐릿하다. 로봇이 보편적 공공재가 아니라 소득·지역에 따라 차등되는 상품이 될 때, 기술은 격차를 완화하기는커녕 오히려 재생산한다.

2) 생성형 AI(말벗 챗봇과 대화형 시스템)

생성형 AI는 정서지원 로봇의 입을 훨씬 유창하게 만든다. 그러나 바로 그 유창함이 새로운 위험을 낳는다. 환각(hallucination)에 따른 부정확한 정보 제공은 복약·건강 안내에서 치명적일 수 있고, 자연스러운 대화는 코켈버그가 말한 기만과 과의존의 문턱을 낮춘다. 여기서 그의 기만 논의는 유효하되 그대로 적용되지는 않는다. 선의의 기만 자체보다, 부정확성과 데이터 축적이 결합될 때 문제가 발생할 수 있기 때문이다. 생성형 AI 말벗은 대화를 위해 음성·영상·건강·생활패턴 등 극도로 민감한 정보를 상시 수집하며, 이 데이터는 어딘가에 저장·학습된다. 실제 대가상의 질을 묻던 코켈버그의 물음은, 생성형 AI 시대에는 매개의 질에 더해 축적된 데이터의 통제권을 함께 물어야 한다.

두 층위가 공유하는 위험은 네 가지다. 첫째, 돌봄 노동의 재편과 탈속련이다. 돌봄 장인정신을 요하는 관계적 노동이 기기 관리 업무로 대체되며 소진을 부를 수 있고, 코켈버그가 강조하듯 이 부담은 여성과 이주노동자에게 불균등하게 전가되기 쉽다. 둘째, 책임의 공백이다. 응급 미감지·오작동으로 피해가 발생했을 때 개발사·공급자·기관·지자체·가족 사이에서 책임이 분산된다. 셋째, 프라이버시의 침식이다. 지속적 모니터링은 돌봄이라는 명분으로 정당화되기 쉬워 더 위험하며, 통제 없는 수집은 돌봄을 감시로 전락시킨다. 넷째, 인간 존엄의 훼손이다. 효율을 위해 따듯한 인간의 현존을 프로그램된 대체물로 바꾸는 것이며, 자기결정이 어려운 노인에게 동의 없이 기술을 적용할 때 특히 침해해진다.

하부구조: 데이터센터·전력·물·기후

세 번째 층위는 대개 논의에서 빠진다. 노인의 외로움을 달래는 말벗 챗봇 한 마디, 요양 데이터를 통합하는 플랫폼 한 번의 연산이 어딘가의 전력·물·탄소를 태운다는 사실이다. 이 하부구조야말로 편익의 개인화, 부담의 사회화가 여실하게 드러나는 지점이다.

2026년 'AI 데이터센터 특별법'을 둘러싸고 인권·노동·복지·환경 등 여러 시민사회 단체가 반대에 나섰다. 산업 진흥을 위해 국가가 공적 지원을 몰아주면서 기후·환경 부담은 지역과 공동체에 떠넘긴다는 것이 주요한 문제의식이다. 특히 화석연료로 데이터센터 전력을 충당할 길을 열어 준 점, 전력계통 영향평가를 면제한 점이 쟁점이 되었다. 시민사회는 이를 계통 부담과 기후 부채를 사회에 떠넘기는 특혜로 보았고, 설립 원칙·요건도 없이 속도전으로 처리된 입법 과정을 졸속이라 비판했다.

통합돌봄의 졸속과 데이터센터 특별법의 졸속은 닮은꼴이다. 둘 다 사회적 합의와 토대 점검을 건너뛰어 채 속도와 산업논리를 앞세웠고, 그 결과 부담이 아래로—현장 인력과 지역 주민에게—지워진다. 그 물리적 규모도 만만치 않다. 대형 AI 데이터센터 한 곳이 웬만한 발전소 한 기에 맞먹는 전력을 요구하기도 하고, 전력·물 사용을 둘러싼 갈등과 소송도 늘고 있다. 대형 데이터센터가 소비하는 물의 양은 막대하고 입지도 수도권에 크게 쏠려 있어, 실효가 불투명한 정크 AI에 대한 경계마저 나온다. 주민이 모르는 사이에 입지가 결정되어 분쟁으로 번지는 사례도 보고된다. 그리고 여기서 세 층위가 하나로 이어진다. 돌봄을 위한 AI가 소비하는 전력·물·탄소의 부담은 비수도권과 미래세대에게 전가된다. 그런데 미래세대는 다름 아닌 미래의 돌봄 대상이다. 돌봄을 위한 AI가 기후위기를 통해 정작 돌봄받아야 할 이들을 위협하는 역설. 이것이 하부구조 층위가 드러내는 문제다. 코켈버그가 로봇을 ‘인간을 비추는 거울’이라 했다면, 데이터센터는 우리가 ‘좋은 삶’의 비용을 누구에게 청구하고 있는지를 비추는 또 하나의 거울이다.

대응 방안: 보급에서 좋은 돌봄으로, 개인에서 정의로

앞선 진단을 종합하면, 대응은 여덟 갈래로 정리된다. 이 여덟 방향을 관통하는 하나의 원칙은 ‘영향받는 자의 참여’, 즉 노인·요양보호사·지역주민·미래세대가 기술의 설계와 평가에 함께 참여해야 한다는 것이다.

- 1) 통합 거버넌스와 질 중심 평가체계: 복지부(서비스)·과기정통부(플랫폼)·산업부(로봇)·개인정보위(프라이버시)로 흩어진 사업을 조정할 국가 단위 거버넌스가 필요하다. 보급 대수가 아니라 돌봄의 질·윤리·책임구조를 함께 점검하고, 로봇이 이용자의 역량(자율·유대·건강)을 실제로 증진했는지 평가하는 성과 기반 지표로 전환해야 한다.
- 2) 돌봄로봇 특화 윤리 가이드라인과 영향평가 의무화: AI 기본법(2026.1 시행)의 고영향 AI 영향평가를 돌봄 영역에 구체적으로 적용하고, 침해금지·안전성·투명성·설명가능성·지속가능성을 담은 돌봄로봇 윤리 가이드라인을 정부 규범으로 제도화하여야 한다. 취약계층 대상 기술에는 EU식 설계 단계 안전·권리 보호(safety/privacy by design) 원칙을 적용한다.
- 3) 보완재 원칙의 명문화: 기술은 인간 돌봄을 대체하는 것이 아니라 보조·증강하는 수단으로 자리매김해야 한다. 로봇이 요양보호사의 신체 부담(이승·배설 보조)을 덜어 관계적 돌봄에 쓸 시간을 확보해 주는 방향-코켈버그의 돌봄 장인정신을 회복시키는 방향-으로 설계·도입되어야 한다.
- 4) 돌봄 노동자 재교육과 참여적 설계: 일본의 실패 사례를 감안했을 때, 현장 수용성 없는 도입은 방치로 끝난다. 돌봄 노동자·이용자·개발자가 함께 기술을 시험·설계하는 참여적 절차, 그리고 오작동·근접위험 대응을 포함한 활용 교육·안전 매뉴얼을 제도화한다.
- 5) 공공재로서의 접근권 보장: 돌봄로봇을 공적 급여·보조금과 연계해 소득·지역 격차를 해소한다. 예비급여를 넘어 본급여화로 나아가되, 기초생활수급자 등 최취약계층이 배제되지 않도록 설계하는 것이 중요하다.
- 6) 이용자 동의·설명권과 돌봄 리터러시: 데이터 수집 범위·목적에 대한 알기 쉬운 설명과 동의 절차, 그리고 노인·가족·종사자가 기술을 비판적으로 이해하는 ‘돌봄 리터러시(care literacy)’ 교육을 병행한다. 기술의 한계와 오류 가능성을 아는 것이야말로 과의존을 막는 첫 단계다.
- 7) 토대 우선 원칙: AI·플랫폼을 투입하기 전에 통합돌봄 전달체계·전담인력·데이터 상호운용성부터 확충해야 한다. 기술이 부실한 토대의 알리바이가 되지 않도록, 사람 기반 돌봄의 최소 요건을 기술 도입의 전제 조건으로 명시한다.
- 8) 기후정의와 하부구조 책임: 돌봄 AI 도입의 사회적 비용 산정에 데이터센터의 전력·물·탄소를 포함시킨다. 재생에너지 전환, 지역 분산, 그리고 입지 지역 주민이 참여하는 거버넌스 원칙을 연계해, 돌봄의 편익을 누리는 사회가 그 환경 부담도 함께 지도록 설계한다.

나가며

코켈버그는 돌봄의 미래를 쓰는 일이 철학자만의 과제가 아니라 모든 이해당사자가 함께 편집해 가는 이야기라고 말한다. 이 통찰은 우리 현실에 그대로 적용된다. 초고령사회의 돌봄 공백은 실재하며, AI와 로봇은 그 공백을 완화할 잠재력이 분명히 있다. 그러나 기술은 스스로 좋은 돌봄이 되지 않는다. 우리에게 필요한 것은 유토피아나 디스토피아냐라는 소모적 이분법이 아니라, 코켈버그가 제안한 대로 구체적 기술을, 구체적 상황에서, 영향받는 당사자와 함께 평가하는 태도다. 그리고 부담이 아래로만 흐르는 이 구조를 뒤집어, 편익과 책임이 함께 분배되도록 설계를 수정해야 하는 일이다. 돌봄의 미래는 로봇이 쓰는 것이 아니다. 그것은 마땅히, 영향받는 모든 사람이 함께 써야 한다.

참고문헌

마크 코켈버그, 『AI 윤리에 대한 모든 것』, 아카넷, 2024.

김희강(2024). ‘돌보는 로봇’?: 돌봄로봇과 로봇돌봄의 윤리. 정부학연구, 31-59.

이소윤. (2025). AI/로봇 기반 노인 돌봄의 윤리적 딜레마: 돌봄윤리 관점에서의 문헌연구. 상담심리교육복지, 12(5), 235-250.

이은영. (2024). 인간 돌봄과 인공지능로봇 (AI) 돌봄 사이에서*초고령시대에 돌봄 로봇과 인간 돌봄의 공존성 연구를 중심으로. 한국간호윤리학회지, 4(1), 1-11.

한겨레신문, “지자체 통합돌봄 인력 82%가 겸직...”

https://www.hani.co.kr/arti/society/rights/1274650.html?utm_source=copy&utm_medium=copy&utm_campaign=btn_share&utm_content=20260831

경향신문, “읍·면·동마다 전담 인력 1명’ 공언하더니...통합돌봄 시행 넉 달 지나도록 전국 151명뿐”

https://www.khan.co.kr/article/202608210600061/?utm_source=urlCopy&utm_medium=social&utm_campaign=sharing

투데이타임즈, “AI 돌봄 로봇, 독거노인 우울감 25% 낮춘다...초고령사회의 새 돌봄 모델로 부상”

<https://www.todaytimes.co.kr/news/496877>

헬스조선, “‘사람보다 낫다’는 AI 돌봄 로봇... 독거 노인에 필요하지만, 과 의존 주의도”

https://health.chosun.com/site/data/html_dir/2025/08/22/2025082201939.html

세계일보 심층기획, “‘말벗 AI’에 의존하는 노인들... ‘챗봇 넘어 진짜 돌봄 로봇 필요’ [심층기획-두 얼굴의 AI 의료·돌봄]”. <https://www.segye.com/newsView/20260602514178>

Nguyen, N. C., & Saito, M. (2025). Issues in applications of nursing care robots, and in the training of care workers in their use in Japan. *Frontiers in medicine*, 12, 1459015.

한국NGO신문, “전국 41개 시민사회단체 'AI 데이터센터 특별법 반대...기후위기, 환경 부담 확대”.

<https://www.ngonews.kr/news/articleView.html?idxno=228274>

오마이뉴스, “AI 데이터센터 유치 열풍, 지역엔 무엇을 남기나”. <https://omn.kr/2it9x>