

기본연구 | 25-16

AI 시대의 포용과 상생을 위한 사회적 의제 연구

문아람 외

2025. 12



정보통신정책연구원
KOREA INFORMATION SOCIETY DEVELOPMENT INSTITUTE

기본연구 | 25-16

AI 시대의 포용과 상생을 위한 사회적 의제 연구

문아람 외

2025. 12



정보통신정책연구원
KOREA INFORMATION SOCIETY DEVELOPMENT INSTITUTE

서 언

인공지능(AI)은 이제 거스를 수 없는 거대한 문명사적 흐름입니다. 생성형 AI의 등장과 기술의 비약적인 발전은 산업의 생산성을 획기적으로 높이고 우리의 일상을 편리하게 바꾸어 놓고 있습니다. 그러나 기술 혁신의 빛이 강할수록, 그 이면에 드리워진 그림자 또한 짙어질 수 있음을 우리는 경계해야 합니다. AI가 가져올 변화는 단순히 기술이나 경제의 영역에 머무르지 않습니다. 일자리의 구조가 재편되고, 정보의 생산과 유통 방식이 근본적으로 변화하며, 공공 서비스와 지역 사회의 모습까지 새롭게 정의되고 있습니다. 이러한 급격한 전환기 속에서 자칫 소외되는 계층이 생겨나거나 사회적 갈등이 심화된다면, AI 기술은 우리에게 축복이 아닌 위협이 될 수도 있습니다. 따라서 지금 우리에게 필요한 것은 기술 발전 속도에 발맞춘 사회적 숙의와 제도적 준비입니다.

이에 정보통신정책연구원은 『AI 시대의 포용과 상생을 위한 사회적 의제 연구』를 통해 AI 시대를 맞이한 우리 사회가 반드시 고민하고 해결해야 할 질문들을 던지는 ‘사회적 의제’를 발굴하였습니다. AI로 격차와 갈등이 심화될 노동, 이용자, 정보 생태계, 지역, 공공행정이라는 5대 핵심 분야를 선정하고 우리 사회의 지속가능성을 위해 시급히 논의되어야 할 쟁점들을 이 보고서에 담고자 하였습니다.

본 연구는 정보통신정책연구원의 문아람 연구위원이 총괄책임을 맡아 수행하였으며, 문정욱 디지털사회전략연구실장, 조성은 연구위원, 연소라 부연구위원, 김휘홍 부연구위원, 이으뜸 부연구위원, 전민경 연구원, 신진호 연구원이 참여하였습니다. 도출된 의제들이 향후 우리 정부가 ‘포용적이고 상생 가능한 AI 정책’을 수립하는데 견고한 토대가 되기를 기대합니다.

2025년 12월

정보통신정책연구원

원장 이 상 규

목 차

서 언	1
요약문	13
제 1 장 서 론	17
제 1 절 연구 배경과 필요성	17
1. 기회와 위기의 동시성	17
2. AI 시대의 사회적 의제 필요성	19
제 2 절 연구 목표와 추진 체계	20
제 2 장 사회적 의제 연구 현황	22
제 1 절 사회적 의제 연구/녹서의 개념적 정의	22
1. 사회적 의제의 개념	22
2. 녹서의 개념	23
3. 녹서(Green Paper)와 백서(White Paper)의 차이	26
제 2 절 사회적 의제 연구의 정책화 사례	31
1. 사회적 의제 연구와 녹서 발간 선행 사례	31
2. 전문가 주도의 사회적 의제 연구와 녹서 발간 선행 사례	37
제 3 절 사회적 의제 프레임워크	40
1. 사회적 의제 사례 분석	40
2. 사회적 의제 구성 요소	43

제3장 5대 분야별 사회적 의제 심층 탐색	45
제1 절 AI와 노동: 정형 편향적 기술 발전의 경로 전환	45
1. 논의의 배경	45
2. 주요 분석 내용	57
3. 식별된 사회적 의제안	77
제2 절 AI와 공공행정: AI 기반 공공·행정서비스 수용성 제고 방안	85
1. 논의의 배경	85
2. 주요 분석 내용	88
3. 식별된 사회적 의제안	137
제3 절 AI와 이용자: AI 일상화 시대의 리터러시와 격차 해소	148
1. 논의의 배경	148
2. 주요 분석 내용	150
3. 식별된 사회적 의제안	164
제4 절 AI와 정보: 정보 생태계 변화와 상호작용	173
1. 논의의 배경	173
2. 주요 분석 내용	183
3. 식별된 사회적 의제안	207
제5 절 AI와 지역: 혁신 확산의 지리적 병목	214
1. 논의의 배경	214
2. 주요 분석 내용	218
3. 식별된 사회적 의제안	226
제4장 5대 분야별 사회적 의제 검증	235
제1 절 조사 개요	235
1. 조사 목적 및 필요성	235

2. 조사 설계 및 방법	235
제2 절 5대 분야별 사회적 의제 전문가 검증 결과	237
1. AI와 노동: 의제 평가 및 주요 의견	237
2. AI와 행정: 의제 평가 및 주요 의견	246
3. AI와 이용자: 의제 평가 및 주요 의견	257
4. AI와 정보: 의제 평가 및 주요 의견	267
5. AI와 지역: 의제 평가 및 주요 의견	277
제5장 결 론	286
제1 절 AI 시대 상생과 포용을 위한 핵심의제 종합	286
1. 노동 분야 의제	286
2. 공공행정 분야 의제	289
3. 이용자 분야 의제	291
4. 정보 생태계 분야 의제	293
5. 지역 분야 의제	294
제2 절 정책적 시사점	296
참고문헌	300
[부록 1]	318
[부록 2]	320
Abstract	334

표 목 차

〈표 1-1〉 AI 발전 단계 구분	17
〈표 2-1〉 독서와 백서의 차이점	29
〈표 2-2〉 시민참여형 독서와 전문가 주도형 독서의 비교	39
〈표 3-1〉 5단계 동적 AI 노출 수준 프레임워크	60
〈표 3-2〉 실제 AI 활용 방식의 유형 구분	62
〈표 3-3〉 잠재적 AI 노출수준(AIOE) 기술통계: 5단계별	64
〈표 3-4〉 잠재적 AI 노출 수준 단계별 상위·하위 10개 직업	66
〈표 3-5〉 초기/최종 단계 AIOE 노출 수준 상위 10개 직종의 순위 변화	69
〈표 3-6〉 AIOE 수준 차이(S1 → S5) 상위 10개 직종	69
〈표 3-7〉 집단별 AI 노출 지표 및 수준 비교	71
〈표 3-8〉 잠재적 AI 노출 수준과 실제 AI 활용 수준 간 상관관계	72
〈표 3-9〉 직업별 실제 관찰된 자동화 및 보완 수준 기술통계	74
〈표 3-10〉 AI 공공·행정서비스 분야	90
〈표 3-11〉 분야별 AI 공공·행정서비스	91
〈표 3-12〉 국내외 AI 공공·행정서비스 분야 비교	92
〈표 3-13〉 내부 행정서비스-대민서비스 간 AI 공공·행정서비스 분야 비교	93
〈표 3-14〉 분야별 AI 공공·행정서비스 기능	94
〈표 3-15〉 ‘행정’ 분야 AI 공공·행정서비스 사례	97
〈표 3-16〉 ‘안전’ 분야 AI 공공·행정서비스 사례	99
〈표 3-17〉 ‘보건·복지’ 분야 AI 공공·행정서비스 사례	100

〈표 3-18〉 ‘교통’ 분야 AI 공공·행정서비스 사례	101
〈표 3-19〉 ‘교육/연구’ 분야 AI 공공·행정서비스 사례	103
〈표 3-20〉 ‘문화관광’ 분야 AI 공공·행정서비스 사례	104
〈표 3-21〉 ‘경제산업’ 분야 AI 공공·행정서비스 사례	104
〈표 3-22〉 ‘환경’ 분야 AI 공공·행정서비스 사례	105
〈표 3-23〉 측정문항 및 신뢰도 분석 결과	120
〈표 3-24〉 표본 특성	122
〈표 3-25〉 사회적 수용성의 인과기제 분석 결과	132
〈표 3-26〉 마지못한 수용성의 인과기제 분석 결과	136
〈표 3-27〉 전문가 인터뷰 대상자	157
〈표 3-28〉 「인공지능기본법」 제31조 제2항 규제 대상	184
〈표 3-29〉 조사 설계	192
〈표 3-30〉 실험 설계	195
〈표 3-31〉 1차 응답 이미지 평가 선형회귀분석 결과	197
〈표 3-32〉 AI가 전혀 활용되지 않은 이미지 이중차분 분석 결과	200
〈표 3-33〉 AI가 부분적으로 활용된 이미지 이중차분 분석 결과	201
〈표 3-34〉 AI가 전면적으로 활용된 이미지 이중차분 분석 결과	202
〈표 3-35〉 콘텐츠 유형별 공유 의도 이중차분 분석 결과	203
〈표 3-36〉 국내 주요 AI 관련 통계조사의 조사 내용 및 지역 단위	217
〈표 3-37〉 AI 산업 분류 기준	220
〈표 3-38〉 자기강화 메커니즘	223
〈표 3-39〉 AI 산업 확산의 지리적 병목 요인 분석	225
〈표 4-1〉 노동 분야 사회적 의제 전문가 검증 결과	237
〈표 4-2〉 노동분야 의제 전문가 정량 평가 결과	240
〈표 4-3〉 노동 분야 의제 전문가 우선순위 평가 결과	241
〈표 4-4〉 노동 분야 의제 전문가 우선순위 선정 사유	242

〈표 4-5〉 노동 분야 의제 전문가 정책과제 제안	243
〈표 4-6〉 노동분야 전문가 신규 제안 의제	246
〈표 4-7〉 공공행정 분야 사회적 의제 전문가 검증 결과	247
〈표 4-8〉 공공행정 분야 의제 전문가 정량 평가 결과	249
〈표 4-9〉 공공행정 분야 의제 전문가 우선순위 평가 결과	250
〈표 4-10〉 공공행정 분야 의제 전문가 우선순위 선정 사유	252
〈표 4-11〉 공공행정 분야 의제 전문가 정책과제 제안	255
〈표 4-12〉 공공행정 분야 전문가 신규 제안 의제	256
〈표 4-13〉 전문가 의견조사에 제시된 AI-이용자 영역 사회적 의제(안) ...	257
〈표 4-14〉 이용자 분야 의제 전문가 정량 평가 결과	260
〈표 4-15〉 이용자 분야 의제 전문가 우선순위 평가 결과	261
〈표 4-16〉 이용자 분야 의제 전문가 우선순위 선정 사유	263
〈표 4-17〉 이용자 분야 의제 전문가 정책과제 제안	264
〈표 4-18〉 이용자 분야 전문가 신규 제안 의제	266
〈표 4-19〉 전문가 의견조사에 제시된 AI-정보 영역 사회적 의제(안)	267
〈표 4-20〉 정보 분야 의제 전문가 정량 평가 결과	271
〈표 4-21〉 정보 분야 의제 전문가 우선순위 평가 결과	272
〈표 4-22〉 정보 분야 의제 전문가 우선순위 선정 사유	273
〈표 4-23〉 정보 분야 의제 전문가 정책과제 제안	275
〈표 4-24〉 전문가 의견조사에 제시된 AI-지역 영역 사회적 의제(안)	278
〈표 4-25〉 지역 분야 의제 전문가 정량 평가 결과	280
〈표 4-26〉 지역 분야 의제 전문가 우선순위 평가 결과	281
〈표 4-27〉 지역 분야 의제 전문가 우선순위 선정 사유	282
〈표 4-28〉 지역 분야 의제 전문가 정책과제 제안	283
〈표 4-29〉 지역 분야 전문가 신규 제안 의제	285
〈표 5-1〉 노동 분야 핵심의제	286

〈표 5-2〉 공공행정 분야 핵심의제	289
〈표 5-3〉 이용자 분야 핵심의제	291
〈표 5-4〉 정보 생태계 분야 핵심의제	293
〈표 5-5〉 지역 분야 핵심의제	294

그 림 목 차

[그림 1 - 1]	연구 추진 절차	21
[그림 3 - 1]	생산성 역설의 역사	46
[그림 3 - 2]	AI가 GDP에 미치는 영향: Goldman Sachs(2023) vs Acemoglu(2024)	48
[그림 3 - 3]	노동소득분배율 하락 요인 분해	49
[그림 3 - 4]	AI와 미래 일자리 영향 전망	51
[그림 3 - 5]	직업별 AI 노출 수준 분포 추이: 5단계	65
[그림 3 - 6]	잠재적 AI 노출 수준-실제 AI 활용 수준 산점도	73
[그림 3 - 7]	직업별 실제 관찰된 자동화·증강 수준 히스토그램	74
[그림 3 - 8]	직업별 실제 관찰된 증강 활용 비중 히스토그램	75
[그림 3 - 9]	AI 보완 잠재력 및 실제 보완 활용 수준 산점도	77
[그림 3 - 10]	국내외 AI 기반 공공·행정서비스 동향 분석 방법 및 절차	89
[그림 3 - 11]	국내외 AI 공공·행정서비스 분야 비교	92
[그림 3 - 12]	내부 행정서비스-대민서비스 간 서비스 분야 비교	93
[그림 3 - 13]	합리적 행위 이론(TRA: Theory of Reasoned Action)	107
[그림 3 - 14]	통합된 기술수용모형(UTAUT)	108
[그림 3 - 15]	통합된 기술수용모형2(UTAUT2)	109
[그림 3 - 16]	혁신저항모형	110
[그림 3 - 17]	혁신확산이론을 적용한 연구모형 사례	112
[그림 3 - 18]	가치기반 수용모형	113
[그림 3 - 19]	SMCRE 커뮤니케이션 모형	113
[그림 3 - 20]	연구모형	119

[그림 3-21] AI 기반 공공·행정서비스 이용 경험률 및 이용 의향	123
[그림 3-22] 분야별 적절한 AI 역할 비중	124
[그림 3-23] AI 기반 공공·행정서비스 평가	125
[그림 3-24] AI 기반 공공·행정서비스 분야 우려*정부 준비 수준 IPA	126
[그림 3-25] AI 기반 공공·행정서비스 유발 문제에 대한 평가	127
[그림 3-26] AI 기반 공공·행정서비스 발생 문제 우려*정부 준비 수준 IPA	128
[그림 3-27] AI 기반 공공·행정서비스의 윤리적 이슈에 대한 평가	129
[그림 3-28] AI 기반 공공·행정서비스 윤리 우선순위*정부 준비 수준 IPA	130
[그림 3-29] AI 기반 공공·행정서비스 개인 수용성 및 조건부 수용성	131
[그림 3-30] ‘혁신저항-사회적 수용성’ 간 관계에서 사회적 지지의 조건부 직접효과	135
[그림 3-31] ‘혁신저항과 마지못한 수용성’ 간 관계에서 휴리스틱의 조건부 직접효과 검증	137
[그림 3-32] AI 서비스 이용경험	155
[그림 3-33] 응답자 특성	193
[그림 3-34] 1차 응답 이미지 평가 도수분포표	196
[그림 3-35] 실제와 인식에 따른 온라인 공유 의도	199
[그림 3-36] AI 활용 수준 판별 정확도	199
[그림 3-37] 사회적으로 바람직한 AI 활용 여부 표시 라벨에 대한 인식 ..	205
[그림 3-38] 사회적으로 바람직한 AI 활용 여부 표시 방식에 대한 인식 ..	205
[그림 3-39] 미국 내 신기술 관련 직업의 지역적 확산	215
[그림 3-40] AI 산업 입지계수(2021년)	221
[그림 3-41] AI 산업 입지계수 변화	222

요 약 문

1. 제 목

AI 시대의 포용과 상생을 위한 사회적 의제 연구

2. 연구 목적 및 필요성

본 연구 과제는 AI 시대의 특정 정책을 제안하는 ‘백서(white paper)’가 아닌, 문제의 본질을 진단하고 사회적 논의의 기반을 마련하기 위한 의제를 모은 ‘녹서(green paper)’의 역할을 수행하기 위해 기획되었다. 따라서 본 보고서의 목적은 AI가 우리 사회의 5대 영역 - 노동, 공공행정, 사용자, 정보 생태계, 지역 - 에서 야기하고 있는 불균형과 불평등의 구체적인 양상과 그 근본적인 메커니즘을 심층적으로 분석하여 주목해야 하는 쟁점들을 제시하고자 한다. 본 보고서에서는 단일한 분야에 집중하지 않고 다수의 영역으로부터 제기되는 사회적 어젠다 간의 유사성과 차별성을 살펴볼 수 있다는 점에서 AI가 초래하는 사회적 문제를 파편화되지 않은 정보로 이해할 수 있다는 차별성과 독특성이 있다.

3. 연구의 구성 및 범위

본 과제가 집중하는 5대 영역은 그간 정보통신정책연구원에서 수행된 AI 사회 정책 연구와 직접 연계되는 주요 영역으로 선정되었다. 본 사회적 의제 연구 수행의 목적은 새롭게 부상하는 정책적 의제를 선제적으로 식별하여 향후 세부 정책 개발이나 방향, 대안을 도출해야 하는 후속 정책연구의 기반으로 기능하기 위함이다.

다. 따라서 AI 기술이 사회 구조와 개인의 삶에 가장 근본적인 영향을 미치는 주요한 통로이자, 불평등이 발현되고 심화되는 영역인 5대 분야의 사회적 의제는 현재 정부의 정책적 관심도가 반영된 기존 정책 과제의 차기 어젠다를 선제적으로 발굴한다는 점에서 의의가 있다고 할 수 있다.

본 연구는 다음과 같은 순서로 추진한다. 우선 사회적 의제 연구의 현황을 확인한다. 사회적 의제 연구의 정책적 활용 사례와 사회적 의제가 가지는 공통분모를 도출하기 위해 사례연구를 수행한다. 다음으로 5대 영역별로 심화 연구를 수행하여 현황을 진단하고 핵심 쟁점을 도출한다. 이와 동시에, 영역 간 상호작용과 복합적인 갈등 양상을 분석하기 위한 ‘갈등 분석 연구반’을 마련한다. 각 분야별 연구와 갈등 분석 연구반의 결과를 종합하여 1차적인 ‘분야별 의제 리스트’를 구성하고, 이후 각계 전문가의 평가와 심층 의견 수렴 과정을 거쳐 사회적 공론화의 토대가 될 최종 의제안을 도출할 예정이다. 이러한 과정을 통해 향후 심도 있는 사회적 합의를 형성하고 증거에 기반한 정책(evidence-based policy)을 설계하기 위한 객관적이고 신뢰할 수 있는 기초 자료를 제공하고자 한다.

4. 연구 내용 및 결과

AI로 인한 변화는 특정 계층의 문제나 기술 정책의 문제가 아니라, 사회 전체의 상생과 포용을 좌우하는 구조적 문제이며, 대응을 늦출 경우 격차·갈등·불신의 심화로 이어질 위험이 크다. 본 연구에서 도출된 5대 분야 의제와 전문가 검증 결과는 이러한 맥락에서 AI의 위험과 기회를 균형 있게 관리하고, 사회 구성원 모두가 이익을 공유할 수 있도록 제도적 토대를 재설계해야 함을 보여준다. 특히 노동권 보호, 공공행정서비스의 책임성과 투명성, 이용자 권리 확보, 신뢰 가능한 정보생태계, 지역의 지속가능성 강화는 AI 시대 상생과 포용의 조건이며, 각 분야별 선제적 대응 없이는 사회 전체의 지속가능성이 위협받을 수 있다. 본 연구로부터 도출한 5대 영역별 핵심의제는 다음과 같다.

〈노동〉

AI로 마모되는 근로자의 인적자본을 회복하기 위한 교육훈련 시스템, 어떻게 설계할 것인가?

지식·재즈업 직무 재편 가속
▶ 생애주기 소득 불안정 심화

#인적자본회복 #선제적재교육 #평생학습

AI가 진인 직무를 대체해 나가는 시대, '경력 단절 없는' 청년 고용은 보장할 수 있을까?

사무·데이터 등 초인 직무 자동화로 청년층의 초기 경력 형성 기회 박탈

#엔트리레벨소멸 #경력사다리붕괴 #초기경력보장

창조적 파괴 과정의 막대한 전환 비용은 누가, 어떤 원칙으로, 어떻게 분담할 것인가?

직무 파괴(분할)와 신규 일자리 창출(느림)의 시차로 인한 소득·속력 공백 발생

#창조적파괴 #비용분담 #사회적합의

알고리즘의 감시와 통제 강화 속에서, 근로자의 권리는 어떻게 보호할 수 있는가?

채용·평가·해고 전반에 AI 개입 확대
▶ 자율성 침해 및 '디지털 테일러리즘' 우려

#알고리즘관리 #디지털감시 #노동권보호

AI 독점으로 인한 초과 이익, 이익의 범위를 어떻게 정의하고 누구와 공유할 것인가?

AI 혜택의 기업 편중 vs 고용·임금의 낙수효과 약화
▶ 불평등 심화

#초과이익 #양극화해소 #새로운분배

〈공공행정〉

접근성 격차와 편향된 알고리즘으로부터 디지털 취약계층의 소외를 어떻게 막을 것인가?

고령층의 낮은 역량과 알고리즘의 편향성이 결합 → 취약계층이 공공 서비스 혜택 배제

#접근성격차 #알고리즘편향 #보편적관리

AI 행정 서비스 오류로 시민 피해 발생 시, 모호해진 '책임 공백'은 어떻게 해소할 것인가?

AI의 불투명성(Black Box)으로 인해 오류 발생 시 개발사, 공무원, AI 시스템 간 책임 소재가 불분명

#블랙박스 #책임소재규명 #피해구제

안전을 명분으로 한 AI 감시 기술 확산, '빅브라더 사회'로의 귀결을 막을 수 있는가?

설시간 대규모 감시 가능
▶ 시민의 사생활과 표현의 자유 등 민주적 기본권 위축 우려

#빅브라더 #기본권침해 #민주적통제

지능화된 해킹과 개인정보 유출 위험으로부터 국가 행정망의 기술적 안정성은 가능한가?

AI 시스템의 복잡한 구성 요소와 모델 자체의 취약점(데이터 추출)
▶ 행정망 대비 및 국가 안보 위협 증대

#해킹취약점 #개인정보보호 #국가안보

AI의 정책 결정 과정과 근거에 대해, 시민이 납득할 수 있는 '설명'을 제공할 수 있는가?

AI가 복지, 주택 등 중요 정책 결정에 개입하지만, 기술적 복잡성으로 인해 시민에게 결정 이유를 설명하기 어려움

#정책투명성 #설명요구권 #신뢰구축

〈이용자〉

모든 국민의 기본권으로서 'AI 리터러시 교육'을 보장하기 위해 정부는 무엇을 해야 하는가?

AI가 보편 기술이 되었으나, 개발자 양성에 치중된 정책으로 인해 일반 시민과 취약계층의 역량 공백 발생

#시리터러시 #교육기본권 #디지털포용

과도한 의존이나 두려움을 없애기 위한 '체험형 윤리 교육'을 어떻게 확산할 것인가?

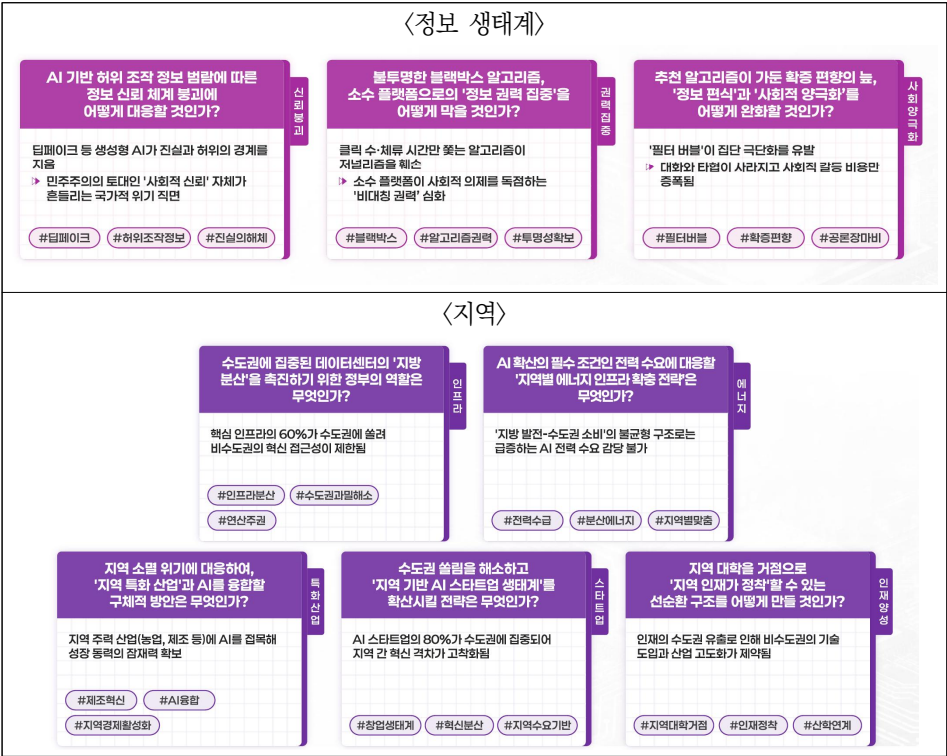
AI 경험 부족으로 인한 고령층 등 취약계층의 막대한 두려움과 적응에 대한 시간으로 기술 격차가 심화됨

#체험형교육 #비판적수용 #두려움해소

고령·장애인의 실질적 사용을 보장하는 법적·제도적 접근성 표준을 어떻게 마련할 것인가?

AI 접근성 불평등이 사회적 배제로 직결되며, 취약계층이 사용자가 아닌 '감시 대상'으로 전락할 위험 존재

#보편적설계 #법적무화 #자기결정권



5. 기대효과

본 연구는 노동, 공공행정, 이용자, 정보 생태계, 지역 등 5대 핵심 분야에서 AI 기술 확산이 초래할 구조적 변화와 사회적 갈등 요인을 심층적으로 진단하고, 이에 대응하기 위한 사회적 의제를 선제적으로 발굴하여 논의를 촉발한다는 점에서 정책적 의의를 갖는다. 특히 노동, 행정, 이용자, 정보, 지역 등 5대 분야에서 AI 확산에 따른 기회와 위기를 균형 있게 진단함으로써, 핵심 국정과제로 추진 중인 ‘AI 기본사회’ 실현을 통해 성장의 과실이 국민 모두에게 공유되는 ‘진짜 성장’의 토대를 마련하는 데 기여할 것이 기대된다. 또한 지역 간 AI 인프라 격차 해소를 위한 지역 균형 발전 전략과 디지털 전환 과정에서 소외되는 계층이 없도록 하는 ‘정의로운 일자리 전환’ 방안을 구체화하는 데도 중요한 참조자료로 활용될 수 있다.

제 1 장 서론

제 1 절 연구 배경과 필요성

1. 기회와 위기의 동시성

여타의 기술과 마찬가지로 AI가 가져오는 기회와 격차의 이중성은 단순한 장단점이 병렬적으로 나열되는 것이 아니라 기술의 본질적 특성에 기인해 발현된다. 즉 AI의 혜택을 창출하는 메커니즘은 동시에 그 위험을 유발하는 원인이 되며 이중적 영향은 시차를 두지 않고 동시에 발생하게 된다. 일례로, AI가 의료 영상 데이터를 분석하는 능력은 의료 서비스의 효율성을 획기적으로 개선하며 인류의 건강 증진에 기여하게 되는 명백한 혜택이지만, 동시에 영상의학과 전문의의 핵심 업무를 대체하여 그들의 고용 안정성을 위협하는 요인이 될 수도 있다. 따라서 AI의 영향에 대한 기회와 위험을 분리하여 위험에 대한 대응 방안을 다루기보다 기술의 특징으로 내재되는 근본적인 상충관계를 이해하여 발생 가능한 문제를 정의하는 것에서 출발해야 한다.

〈표 1-1〉 AI 발전 단계 구분

수준	DeepMind(Morris et al., 2024)	OpenAI(Bloomberg, 2024)
1	발아기(숙련되지 않은 사람과 동등)	챗봇(대화형 언어 AI)
2	유능가(숙련된 성인의 50%)	추론자(인간 수준의 문제해결)
3	전문가(숙련된 성인의 90%)	에이전트(행동하는 시스템)
4	장인(숙련된 성인의 99%)	혁신가(발명을 지원)
5	초인간(인간 100% 수준 능가)	조직(조직의 업무를 수행)

자료: Morris et al.(2024), Bloomberg(2024. 7. 12), “OpenAI Scale Ranks Progress Toward Human-Level Problem Solving”

파괴적 혁신을 가져오는 기술은 파괴와 혁신 사이에 시차가 있어 파괴로부터 영향을 받는 구성원과 기존의 사회 시스템이 적응할 시간을 가질 수 있던 것과 달리 AI의 영향은 기회와 위기를 동시에 유발하고 있다. 산업혁명 시기 기계화나 전기화 같은 범용기술이 수십 년에 걸쳐 보급되어 노동시장과 제도가 점진적으로 재편될 수 있었지만(Bresnahan & Trajtenberg, 1995), 생성형 AI를 비롯한 최신 AI 기술은 불과 지난 2~3년간 전 산업과 직종에 파급력을 미치고 있다. 이처럼 시장과 제도의 적응 기간이 거의 없는 AI의 확산은 생산성 향상, 새로운 시장 창출 등 긍정적 효과와 동시에, 고용 대체·소득 양극화·사회적 불평등 심화라는 부정적 충격을 병행하여 유발하게 된다.

구글 딥마인드, 오픈AI 등의 기술 기업이 주도한 연구는 AI를 동적인 개념으로 포착하기 위해 AI의 발전 단계를 구분하였으며, 현재의 기술 수준은 인간 전문가와 동등하거나 약간 능가하는 것으로 평가된다. 생산성 향상, 경제 성장, 그리고 사회 서비스 혁신의 잠재력을 유발하는 범용기술로의 발전은 의료, 교육, 금융 등 전문 서비스 분야의 생산성과 접근성을 획기적으로 개선하고 있으며(Brynjolfsson et al., 2023), 반복적이고 정형화된 업무를 자동화하여 인간이 보다 창의적이고 부가가치가 높은 활동에 집중할 수 있는 환경을 조성할 잠재력을 지닌다.

반면에 AI의 빠른 확산은 기존의 사회·경제적 불평등을 심화하고, 과거에는 볼 수 없었던 새로운 형태의 격차를 유발하는 강력한 동인으로도 작용한다. 고숙련자와 저숙련자 간의 소득 격차를 확대하기도 하고, 기술에 대한 물리적 접근성과 데이터의 편향성에 기인한 공공 서비스의 차별적 제공 가능성도 제기된다. 또한 새로운 리터러시를 요구함에 따라 사용자의 디지털 기술 활용의 접근성과 이로부터 편취할 수 있는 혜택의 격차도 벌어질 수 있다. 알고리즘에 의해 파편화된 정보 생태계, 그리고 AI 기술 패권을 둘러싼 국가 간, 지역 간 경제적 격차 등에 이르기까지 AI가 유발하는 불평등은 다층적으로 형성되고 있다(Cazzaniga et al., 2024; UNESCO, 2022).

2. AI 시대의 사회적 의제 필요성

문제는 이러한 동시다발적이고 복합적인 영향은 정책결정자와 사회 전반에 대응 체계를 마련해야 할 시간적 압박을 가한다는 점이다. 사회가 기술 충격을 소화하기도 전에, 정치·행정 측면에서는 산업과 시장의 변화에 뒤처지면 안 된다는 압박 속에서 정책 대안을 먼저 내놓는 경향을 보인다. 그러나 Korinek & Stiglitz (2017)가 지적하듯이, AI의 사회경제적 영향은 기술의 수용 수준, 제도적 환경, 산업 구조에 따라 경로가 크게 달라질 수 있으므로, 문제에 대한 정의 없는 처방은 오히려 정책 실패와 사회적 비용을 키울 위험이 있다. 기술 발전의 사회경제적 효과는 ‘혁신의 방향’과 ‘제도적 대응’이 상호작용하여 결정되므로, 제도적 대응 마련 과정에서 잘못된 문제의 정의가 설정한 방향 그 자체를 왜곡하지 않기 위한 노력도 병행되어야 한다(Brynjolfsson & McAfee, 2014).

따라서 AI 시대에 부상하는 사회경제적 쟁점은 성급한 ‘정책 솔루션 제시’보다 정밀한 이슈 분석과 쟁점 규명을 선행하는 것이 필요하다. 특히 노동시장 충격, 디지털 격차, 인재 쏠림 등과 같이 상충하는 이해관계가 얽힌 사안은 정책을 마련하기에 앞서 문제를 정의함으로써 이해관계자는 누구이며 상충 관계의 요인이 무엇인가를 분석하는 것이 사회적 합의와 정책의 수용도를 확장할 가능성을 높인다.

실제로 EU는 정책 주제에 대한 논의를 촉진하기 위해 유럽 위원회가 ‘녹서’를 발간하고, 관련 당사자가 제시한 제안을 바탕으로 협의 과정에 참여하고 토론하여 입법적 발전까지 이어질 수 있도록 채널을 마련해 두었다.¹⁾ 일례로 2005년과 2021년 각각 고령화와 소매금융 서비스에 관한 녹서를 발간한 바 있다(EC, 2021; 2015). 영국 정부 또한 2017년 AI를 포함한 미래 기술 전략에 대한 정책 마련 과정에서 산업전략 녹서를 통해 폭넓은 의견 수렴을 거친 바 있다(UK BEIS, 2017).

1) EU, “Green Paper,” [https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM:green_paper#:~:text=Green%20papers%20are%20documents%20published,European%20Union%20\(EU\)%20level](https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM:green_paper#:~:text=Green%20papers%20are%20documents%20published,European%20Union%20(EU)%20level).

제 2 절 연구 목표와 추진 체계

본 연구 과제는 AI 시대의 특정 정책을 제안하는 ‘백서(white paper)’가 아닌, 문제의 본질을 진단하고 사회적 논의의 기반을 마련하기 위한 의제를 모은 ‘녹서(green paper)’의 역할을 수행하기 위해 기획되었다. 따라서 본 보고서의 목적은 AI가 우리 사회의 5대 영역 - 노동, 행정, 이용자, 정보 생태계, 지역 - 에서 야기하고 있는 불균형과 불평등의 구체적인 양상과 그 근본적인 메커니즘을 심층적으로 분석하여 주목해야 하는 쟁점들을 제시하고자 한다. 본 보고서에서는 단일한 분야에 집중하지 않고 다수의 영역으로부터 제기되는 사회적 어젠다 간의 유사성과 차별성을 살펴볼 수 있다는 점에서 AI가 초래하는 사회적 문제를 파편화되지 않은 정보로 이해할 수 있다는 차별성과 독특성이 있다.

본 과제가 집중하는 5대 영역은 그간 정보통신정책연구원에서 수행하고 있는 AI 사회정책 연구와 직접 연계되는 주요 영역으로 선정되었다. 본 사회적 의제 연구 수행의 목적은 새롭게 부상하는 정책적 의제를 선제적으로 식별하여 향후 세부 정책 개발이나 방향, 대안을 도출해야 하는 후속 정책연구의 기반으로 기능하기 위함이다. 따라서 AI 기술이 사회 구조와 개인의 삶에 가장 근본적인 영향을 미치는 주요한 통로이자, 불평등이 발생되고 심화되는 영역인 5대 분야의 사회적 의제는 현재 정부의 정책적 관심도가 반영된 기존 정책 과제의 차기 어젠다를 선제적으로 발굴한다는 점에서 의의가 있다고 할 수 있다.

본 연구는 다음과 같은 순서로 추진한다. 우선 사회적 의제 연구 현황을 확인한다. 사회적 의제 연구의 정책적 활용 사례와 사회적 의제가 가지는 공통분모를 도출하기 위해 사례연구를 수행한다. 다음으로 5대 영역별로 심화 연구를 수행하여 현황을 진단하고 핵심 쟁점을 도출한다. 이와 동시에, 영역 간 상호작용과 복합적인 갈등 양상을 분석하기 위한 ‘갈등 분석 연구반’을 마련한다. 각 분야별 연구와 갈등 분석 연구반의 결과를 종합하여 1차적인 ‘분야별 의제 리스트’를 구성하고, 이후 각계 전문가의 평가와 심층 의견 수렴 과정을 거쳐 사회적 공론화의 토대가

될 최종 의제안을 도출할 예정이다. 이러한 과정을 통해 향후 심도 있는 사회적 합의를 형성하고 증거에 기반한 정책(evidence-based policy)을 설계하기 위한 객관적이고 신뢰할 수 있는 기초 자료를 제공하고자 한다.

[그림 1-1] 연구 추진 절차



자료: 연구진 구글 제미니AI 활용하여 생성

제 2 장 사회적 의제 연구 현황²⁾

제 1 절 사회적 의제 연구/녹서의 개념적 정의

1. 사회적 의제의 개념

사회적 의제란 사회 구성원 다수가 주목하고 논의하는 주제로서, 일정 시기에 사회 전반에서 문제로 인식되며 다양한 매체와 공적 담론 속에서 반복적으로 다루어져 대중이 그것을 중요하게 인식하게 된 사안을 말한다(Kingdon, 1984; McCombs & Shaw, 1972).³⁾ 사회적 의제에 대해 사회 구성원 간 합의를 형성하는 과정은 민주적 거버넌스의 핵심 요소로, 중요 사회 이슈를 해결하기 위한 국가 정책의 추진 동력을 확보하고 정치적 기반을 마련한다는 점에 그 중요성이 있다. 현대 민주주의 사회에서는 사회적 의제 형성 과정에서 의견 불일치와 다양한 해석이 발생할 수 있으며 때로는 가치의 다원성으로 인해 의제 형성의 합의가 어렵게 되지만, 이런 차이와 다양성이 오히려 사회적 의제 도출을 위한 심층적 토론과 균형 있는 합의 형성을 촉진하는 중요한 동력이 될 수 있다(하승태·조의현, 2008).

사회적 의제는 정부 정책 형성의 출발점으로, 정부가 특정 사회문제를 공론화하고 정책적으로 해결하고자 할 때, 그중 중요한 사안을 정책 의제(policy agenda)로 채택한다(정정길 외, 2023). 따라서 사회적 의제는 사회문제 해결 과정의 초기

2) 본 장은 한국행정연구원 조유선 부연구위원을 자문을 바탕으로 작성되었음.

3) Kingdon(1984)은 '사회적 의제' 대신 '정책 의제(policy agenda)'라는 용어를 사용하였는데, 정책 의제는 정책 입안자들이 검토하거나 실행하려고 고려하는 문제들의 집합이며, 이러한 이슈들이 공중(public)과 언론을 통해 강조되어야만 정책 의제로 부상할 수 있다고 설명함. McComb & Shaw(1972)도 대중이 중요한 문제로 인식하게 된 것이 공중 의제(public agenda)이며 이것들이 추후 정부가 다루어야 할 중요 문제인 정책 의제화된다고 설명함. 두 연구 모두 '사회적 의제'라는 용어를 직접적으로 사용하지는 않았으나 개념의 일치성을 고려하여 본 보고서에서 인용하였음.

단계에서 핵심적인 역할을 수행한다. 이에, 정부는 사회적 의제를 전략적으로 관리하고 선제적으로 발굴하기 위한 노력을 기울이고 있다. 특히 환경 변화에 대응하기 위해 각 분야의 정부출연연구기관을 중심으로 미래전략과 의제 발굴 연구가 활발히 진행되고 있으며, 대표적으로 경제·인문사회연구회가 2000년대 초반부터 국가 의제 연구를 지속적으로 추진해오고 있고, 국회미래연구원에서도 미래전략을 위한 정책 의제 연구를 수행해 오고 있다(NRC 국가의제연구단, 2023).

국제적으로도 사회적 의제라는 개념은 국가의 중요한 정책 방향을 제시하는 데 활용되고 있다. 유럽연합(EU)은 2005년 사회정책의제(Social Policy Agenda)를 발표하여 사회통합과 포용을 촉진하기 위한 전략 과제를 제시하였으며, 2017년에는 유럽 사회권리 기본원칙(European Pillar of Social Rights)을 채택하여 평등한 기회, 공정한 노동조건, 사회 보호와 포용 등 20개 핵심 원칙을 제시하였다.⁴⁾ 이 원칙들은 EU 회원국의 사회정책 의제 수립에 있어 공통의 준거틀(reference framework)로 작용하며, 사회적 의제 설정을 통해 회원국 간 정책 방향을 조율하고 사회적 가치를 확산하는 데 기여하였다. 이러한 사례는 사회적 의제가 국내외를 불문하고 사회문제에 대한 공론을 형성하고 정책 대응 방향을 구체화하는 핵심 수단임을 보여준다.

2. 녹서의 개념

사회적 의제를 공론화하고 정책으로 발전시키는 과정에서 세계 여러 나라의 정부는 녹서(Green Paper)라는 수단을 활용해 왔다. 녹서는 일반적으로 정부가 중요 정책을 결정하기에 앞서 해당 주제에 대한 공개적인 논의와 자문을 위해 발간하는 공문서를 의미한다.⁵⁾ 영국을 비롯한 유럽에서 폭넓게 활용되고 있으며, 아직

4) EU Social Policy Agenda(<https://www.eurofound.europa.eu/en/european-industrial-relations-dictionary/social-policy-agenda>, 검색일: 2025. 8. 14) 웹사이트 내용을 참고하여 작성.

5) EU-LEX 웹사이트 내 Green Paper(https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=LEGISSUM:green_paper, 검색일: 2025. 8. 14)에 대한 내용을

정책으로 확정되지 않은 사회적 의제와 대안을 제시하고, 이에 대해 이해관계자와 국민의 의견을 폭넓게 수렴하는 것을 목적으로 한다. 영국에서는 새로운 법률이나 정책을 도입하기 전에 우선 녹서를 발표하여 정부의 초기 구상이나 문제의식을 공개하고, 이에 대한 대중과 이해당사자의 의견을 청취한다.⁶⁾ 녹서는 입법을 향한 첫 단계로서 구체적인 행동에 대한 약속 없이 토론을 유도하는 문서이고 여러 대안적 정책 방향을 포함하고 있는 것이 특징이다.⁷⁾ 영국에서 발간된 녹서는 후속 단계인 백서로 이어지는데, 하나의 주제에 대해서 녹서로 의견 수렴을 거친 후 아이디어가 정리되면 백서로 정책 방향이 구체화된다.⁸⁾

EU에서도 중요한 정책 현안을 둘러싼 공식적 논의의 출발점으로 녹서를 활용한다. 녹서는 특정 사회 이슈 현황과 해결을 위해 고려해야 할 정책 방향을 제시하며, 이를 통해 해당 주제를 둘러싼 공개적 토론의 장을 마련한다.⁹⁾ 이런 문서는 관련 이해당사자뿐만 아니라 일반 국민에게도 의견 제출과 논의 참여를 촉구함으로써 다양한 시각과 이해를 수렴하고 정책 형성 과정에 반영할 수 있는 기반을 제공한다. 예를 들면, EU 집행위원회는 “2021년 인구 고령화에 관한 녹서(Green Paper on ageing: Fostering solidarity and responsibility between generations)”를 발표하고 이를 통해 고령화로 인한 사회의 장기적 영향(연금, 돌봄 등)에 대한 광범위한 의견 수렴을 거쳤다(European Commission, 2021). EU의 녹서는 향후 입법 또는 정책 발전으로 이어질 수 있는 논의의 토대를 제공하며, 이후 보다

참고하여 작성.

- 6) 영국 의회 홈페이지(<https://www.parliament.uk/site-information/glossary/green-papers/?id=32625>)를 참고하여 작성.
- 7) Shepherd+Wedderburn 웹페이지(<https://shepwedd.com/knowledge/what-difference-between-green-and-white-paper#:~:text=a%20final%20decision%20is%20made>, 검색일: 2025. 8. 12) 내용을 참고하여 작성.
- 8) UK Gov 웹페이지(<https://www.gov.uk/government/how-government-works#:~:text=Draft%20legislation>, 검색일: 2025. 8. 12.) 내용을 참고하여 작성.
- 9) EU-LEX 웹페이지 내 Green Paper(https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=LEGISSUM:green_paper, 검색일: 2025. 8. 14)에 대한 내용을 참고하여 작성.

구체적인 조치들은 백서로 정리된다는 점에서 영국과 동일한 맥락을 지닌다.

독일도 녹색과 백서 절차를 정책 수립 과정에 도입한 경험이 있는 국가로, 그중 가장 대표적인 사례가 ‘노동 4.0(Arbeiten 4.0)’이다. 독일 연방노동사회부(Bundesministerium für Arbeit und Soziales, BMAS)는 2015년 미래 노동 환경에 대응하기 위해 노동의 미래와 관련된 주요 변화 동향, 가치관의 전환, 정책과제 등을 개괄하고 사회 전반의 고용·근로 형태 변화에 대한 근본적인 질문을 제기하는 ‘노동 4.0’ 녹서를 발간하였다. 이후 독일 정부는 이를 토대로 학계, 노동계, 경영계, 시민사회 등 다양한 이해관계자와 전문가가 참여하는 폭넓은 사회적 대화를 약 2년에 걸쳐 추진한 뒤, 2017년 녹서 단계에서 제기된 핵심 질문에 대한 사회적 합의와 정부의 정책 대응 방향을 정리한 백서를 발간하였다.¹⁰⁾ 이 사례는 독일에서 녹서가 사회적 토론을 촉발하는 공론화 기제로 기능하고, 백서는 그 논의를 정책 대안과 실행계획으로 구체화하는 수단으로 기능했음을 보여준다(정운찬, 2018. 11. 19).

한국은 유럽식 녹서 제도를 공식적으로 운영하지는 않았지만, 그 개념과 필요성에 대한 논의는 꾸준히 이어져 왔다. 국내에서는 주로 정책 결정 이후 결과를 보고하거나 자료를 정리하는 백서 중심의 관행이 일반적이었으나, 미래 사회문제를 선제적으로 다루기 위해 녹서를 도입해야 한다는 요구가 지속적으로 제기되어 왔다(채준호 외, 2025; 홍인기, 2024. 8. 7). 특히 다가올 사회문제에 대해 질문을 도출하고, 정책 수립 이전 단계에서 관련 이해당사자 모두가 의견과 질문을 제시할 수 있는 기회를 보장하는 녹서 과정을 거침으로써 사회적 합의를 기반으로 한 대응 방안을 마련하고, 이를 정책 결정과 긴밀히 연계하는 것이 중요하다는 인식이 확산되고 있다.

10) 독일 연방노동사회부(BMAS) 홈페이지 내 웹페이지(<https://www.bmas.de/EN/Services/Publications/arbeiten-4-0-greenpaper-work-4-0.html#:~:text=With%20the%20dialogue%20process%20Work,partners%2C%20associations%20and%20operational%20practice>, 검색일: 2025. 8. 11)를 참고하여 작성.

3. 녹색(Green Paper)와 백서(White Paper)의 차이

녹서는 국가마다 활용 방식에 차이가 있지만, 공통적으로 정책 형성 초기 단계에서 정부가 국민과 이해관계자에게 다양한 질문을 제시하고 의견을 수렴하는 절차를 포함한다. 영국과 EU를 비롯한 유럽 국가에서는 이러한 과정을 중요한 국가 정책 결정 전에 도입하여 확대하고 있다. 한국에서도 녹서를 사회적 의제 설정과 정책 혁신을 위한 핵심 수단으로 인식하며, 그 도입 필요성이 지속적으로 강조되고 있다.

그동안 우리나라 정부 정책 결정 과정에서는 정책 형성 초기 단계에서 국민과 이해당사자의 의견을 수렴하는 ‘녹서’ 형태의 사회적 대화가 생략되는 경우가 많았다. 대신 정부의 주요 정책 방향을 정리하여 발표하거나 특정 사안에 대한 정부의 입장과 계획을 공식적으로 표명하는 백서 단계로 바로 이어지는 관행이 일반적이었다(채준호 외, 2025).

백서는 흔히 특정 사안이나 분야에 대해 정부가 조사·분석한 결과와 향후 구체적인 정책에 대한 정책 선언문으로, 입법 개정이나 새로운 법률 도입에 대한 제안을 담고 있는 경우가 많다.¹¹⁾ 영국에서는 백서를 정부의 정책 의도를 밝히는 공식 보고서(Command Paper)로 여기며, 계획 중인 법안의 초안이 포함되기도 한다. 이러한 백서에 대해 관련 이해관계자와의 추가적인 협의와 논의를 바탕으로 최종 수정을 거쳐 의회에 최종 법률을 제출하게 된다.

한국에서는 백서가 대개 연계 보고서 또는 사업 결과 보고서의 성격으로, 각 부처에서 정책 집행 경과나 성과를 정리해 국민에게 보고하는 문서로 활용되어 왔다(정운찬, 2018. 11. 19). “국방백서”, “산업통상자원백서”, “생명공학백서”, “정보통신백서” 등 해당 분야의 현황과 정부 조치들을 매년 또는 정기적으로 종합하여 발간하는 자료집으로 활용되어 오고 있다. 이는 미래 계획보다는 지난 기간의 정책 추진 내용과 평가를 담는 경우가 많아, 영국 및 EU 국가들이 활용하는 입법

11) 영국 일간지 《The Guardian》의 웹페이지(<https://www.theguardian.com/careand-support/reform/what-green-paper>, 검색일: 2025. 8. 11)를 참고하여 작성.

을 위한 정책 방향을 제시하는 백서와는 다소 차이가 있다. 그럼에도 기본적으로 백서는 정부가 특정 현안에 대한 입장과 정보를 투명하게 공개하고 국민과 국회에 설명하기 위한 공식 문서라는 공통점이 있다.

백서는 녹서에 비해 더 확정적이고 권위적인 성격을 띤다. 이런 측면에서 영국에서는 백서를 Command paper라 하며, 이해관계자로부터 의견 수렴을 위한 녹서는 ‘Consultation paper’라고도 한다. 녹서 단계에서 수렴된 다양한 의견과 대안이 정부 내부 검토와 사회적 논의를 통해서 수렴·조정되면, 이를 토대로 정부의 정책 구상이 백서로 정리된다. 특히 백서는 이후 법률 제정이나 정책 시행의 근거가 되는 경우가 많다.

이와 같이 살펴본 개념을 토대로 녹서와 백서의 차이점을 비교하면 다음과 같다. 먼저 작성 목적의 측면에서, 녹서는 특정 사회적 의제에 대해 ‘무엇이 문제이며 어떤 대안이 가능한가’라는 질문을 제기하고 논의를 촉발하는 것이 주된 목적이다. 이에 반해 백서는 그 논의의 결과를 바탕으로 ‘정부가 무엇을 하겠다’라는 구체적인 대응 방안과 정책 의지를 밝히는 데 목적이 있다(채준호 외, 2025). 즉 녹서가 문제를 제기하는 문서라면 백서는 해답을 제시하는 문서라고 할 수 있다.

작성 주체 측면에서는 두 문서 모두 공식적으로는 정부 부처 또는 공공기관이 작성하는 것이 일반적이지만, 녹서의 경우 정부기관 외에도 정부 연구기관 혹은 민간기관에서 학계, 전문가, 시민사회 등의 의견을 담아 발간하기도 한다.¹²⁾ 녹서 작성 과정에서는 학계, 시민사회 등 외부 전문가와 이해관계자의 참여가 활발하게 이루어지며, 백서는 전문가 집단과 이해관계자가 참여하기도 하지만 정부의 입장이 담기는 만큼 내부 검토와 조율을 거쳐 출간된다(박찬수 외, 2024).

12) 홍콩대 경영대학(The University of Hong Kong Business School)은 『The Hong Kong Economic Policy Green Paper 2025』를 발간하여 홍콩 경제의 활력 제고를 위한 정책 권고안을 제시하였다. 이 보고서는 부동산 가격 급등, 인재 확보, 기후변화, 공급망 재편 등 글로벌 경제 환경 변화에 대응하기 위한 방안을 담고 있으며, 홍콩의 미래 발전 방향을 모색하기 위해 학계와 전문가 집단이 함께 참여하여 편찬되었다(HKU Business School, 2025).

작성 시점 측면에서는, 녹서는 정책주기의 초기 단계에서 문제 정의와 공론화 단계에서 발간되며, 백서는 정책 결정 직전 또는 직후에 발간되어 정부 방침을 확정 짓는 단계에 활용된다(민주연구원 집단지성센터, 2025). 독일의 ‘노동 4.0 (Arbeiten 4.0)’ 사례에서 볼 수 있듯이 국가의 중요 정책에 대해 녹서 발표 후 약 1~2년간의 숙의를 거쳐 백서가 나오게 되는데, 이처럼 녹서는 초기 구상, 백서는 최종 구상이라는 시간 차가 존재한다(박찬수 외, 2024). 또한 시민 참여 측면에서도 녹서는 적극적인 참여를 내재화하고 있는데, 녹서 작성 과정에서의 의견 수렴뿐만 아니라(채준호 외, 2025), 녹서 발표 후에도 정부는 관련 주제에 대한 공개토론, 설문, 공청회 등을 통해 국민 의견을 수렴하며 경우에 따라 녹서 자체에 국민들이 다양한 의견을 제시할 수 있는 플랫폼을 만들기도 한다(Federal Ministry of Labour and Social Affairs, 2015). 반면 백서는 녹서 이후에 진행된 숙의 내용의 결과를 바탕으로 도출된 구체적인 정책안을 담고 있으므로 별도의 공식적인 참여 절차를 동반하기보다는 정부 정책안에 대한 발표에 더 중점을 둔다.

녹서와 백서 모두 정책자료이자 제안서의 성격을 지니기 때문에 법적 구속력은 없지만, 백서는 정부의 공식 입장이 발표된 이상 이후의 정책 추진 시 해당 방향을 따를 가능성이 높고 정부의 책임이 수반되기 때문에 상대적으로 효력이 더 높다고 할 수 있다. 영국의 경우 백서에 담긴 정책은 정부 입법안으로 구체화되어 의회에 제출되는 것이 통상적이며, EU의 경우에는 녹서에서 출발한 논의가 백서에 담기면 회원국의 정책에 어느 정도 반영되도록 하는 구속력을 지니게 된다. 반면 녹서는 이슈에 대한 열린 질문과 논의의 장이라는 측면에서, 녹서에서 논의된 아이디어라 하더라도 최종적으로 정책으로 채택되지 않을 수 있으며, 정부에서는 상대적으로 더 유연하고 편안하게 새로운 제안을 논의해 볼 수 있다.

마지막으로 활용 방식의 측면에서, 녹서는 사회 문제 해결을 위한 문제의 정의 및 해결을 위한 정책 아이디어의 발전과 의견 수렴을 위한 도구로 활용된다. 정부가 직접 의제를 주도하기보다는 쟁점을 제시하고 사회적 숙의를 이끌어내는 토론의 장으로 활용되며, 이를 통해 정책의 민주적 정당성을 높이고 정책 실패의 위험

을 낮추는 기능을 한다(채준호 외, 2025). 한편 백서는 이미 확정된 정책 방향을 국민과 의회에 설명하고 설득하는 한편, 향후 정책 수행의 근거자료로 활용되는 역할을 한다. 한국처럼 백서를 성과보고서로 활용하는 경우에는 정책 집행 과정과 성과를 투명하게 공개하고 기록을 남기는 기능을 하기도 한다.

〈표 2-1〉 녹색과 백서의 차이점

구분	녹서(Green Paper)	백서(White Paper)
작성 목적	이슈 제기 및 논의 촉발: 쟁점과 대안을 제시하고 다양한 의견을 수렴	정책 제시 및 입장 표명: 조사 결과와 대응방안을 정리하여 공식 보고
작성 주체	- 정부(부처·위원회 등) 주도로 작성하되 외부 전문가·이해관계자 의견 적극 반영 - 연구기관, 민간기관 등에서도 발간	정부(부처·내각 등)에서 최종 작성·승인 → 정부의 공식적인 입장을 담음
작성 시점	정책 결정 이전의 초기 단계(문제 정의 및 공론화 단계)	- 정책 결정 직전 혹은 이후(최종 방침 확정 및 입법 직전 단계) - 한국과 같이 백서를 성과보고서로 활용하는 경우 정책 집행 이후에 발간
이해관계자 및 시민 참여	높음	낮음(일부 전문가나 이해관계자가 제한적으로 참여하거나 필요시 후속 과정에서 반영)
법적 구속력	없음(아이디어 제안의 성격이며 정책 채택 여부 미확정)	없음(다만, 공식 입장으로서 이후 정책 추진의 준거점이 됨)
활용 방식	의제 형성 및 대안 탐색(사회 이슈에 대한 합의를 이끌어내는 숙의 민주주의의 도구)	정책 발표 및 기록(정부의 결정을 설명하고 향후 참고하는 공식자료)

자료: 연구진 작성

〈표 2-1〉에서 보듯이 녹색과 백서는 상호 보완적인 관계를 가지면서도 특성에 분명한 차이점이 존재한다. 녹서는 중요 사회 이슈에 대한 개방적인 질문을 던짐으로써 사회적 의제를 공론화하는 장으로서의 역할을 하며, 백서는 이런 논의과정에서 도출된 답변과 합의를 토대로 구체적인 정책 대안을 결정해 발표하게 된다.

특히 AI 시대와 같이 기술변화에 따른 사회적 영향이 큰 주제에 대해서는 녹서를 통한 선제적이고 포용적인 논의가 중요하며, 이를 바탕으로 백서를 통한 정책화와 제도화가 이어질 때 정책의 수용성과 효과성이 높아질 수 있다. 이와 같은 인식하에, 우리나라에서도 AI 시대의 노동시장과 고용환경 변화에 대응하기 위해 노사정이 협력하여 국가 AI 노동정책에 필요한 논의 주제를 담은 녹서를 2025년 10월에 발간할 예정이다.¹³⁾

13) 《아시아경제》(2025. 7. 17), ““노사정 협력해 AI 변화 대응해야”...노동정책 길잡이 ‘녹서’ 10월 발간”, <https://view.asiae.co.kr/news/view.htm?idxno=2025071718015036192>(검색일: 2025. 8. 14).

제 2 절 사회적 의제 연구의 정책화 사례

1. 사회적 의제 연구와 녹서 발간 선행 사례

정부 정책 과정에서 사회적 의제 연구나 녹서는 정책 결정을 앞두고 다양한 이해관계자와 국민의 의견을 수렴하기 위해 발간되는 제안서의 성격을 지닌다. 사회적 이슈나 문제 현상에 대한 여러 대안과 질문을 제시하고 이에 대한 사회 전체의 토론을 촉발하며 정부 정책 방향 제시에 도움이 되는 것을 목적으로 하고 있다. EU 등 선진국에서는 오래전부터 이러한 녹서 제도를 활용해 왔으며, 녹서를 통해 사회적 숙의와 공론화 과정을 거쳐 그 결과를 반영한 백서를 최종적으로 채택하는 구조가 정착되어 있다.¹⁴⁾ 이러한 과정은 사회적 의제를 투명하게 민주적으로 형성하고, 정책 결정에 사회적 합의와 정당성을 부여하는 데 기여한다.

가. 영국

영국 정부는 전통적으로 녹서와 백서를 활용한 정책 입안이 활발한 나라이다. 최근 10년간 영국에서는 주택, 복지, 산업전략, 사회통합 등 다양한 사회적 의제에 대해 녹서를 발간하고 정책 수립을 위한 의견 수렴의 과정을 거쳤다. 대표적인 사례로 영국 정부는 2017년 6월 그렌펠 타워 화재 참사¹⁵⁾를 계기로 촉발된 사회주택 개혁 논의에 대응하여 2018년 “A New Deal for Social Housing(사회주택을 위한 새로운 계획)”이라는 녹서를 발간하였다.¹⁶⁾ 영국 주택·지역사회·지방정

14) EUR-Lex 웹페이지 내 Green Paper(https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM:green_paper, 검색일: 2025. 8. 12)에 관한 내용을 참고하여 작성.

15) 런던 서부 래터머 로드에는 120가구, 약 600명의 주민이 거주하는 24층의 구청 소유 서민형 임대아파트에서 발생한 화재로 100여 명의 사상자가 발생한 사건으로, 도시 미관을 이유로 리모델링 시 사용된 플라스틱 외장재가 불을 키운 원인으로 지목되며, 정부의 건축 재정과 잘못된 주택 정책이 대형 참사로 이어졌다고 영국 국민이 크게 분노함(소방안전원 웹진 소방안전 플러스, <https://www.kfsi.or.kr/contents/webzine/201709/sub01-01.html>, 검색일: 2025. 8. 17).

16) 녹서 및 구체적인 정보는 영국 정부 내 웹페이지(<https://www.gov.uk/government>

부부(Ministry of Housing, Communities & Local Government)가 발간한 이 녹서는 사회주택(임대주택) 거주자의 안전과 권리 강화를 핵심 주제로 삼아, 정부의 사회주택 비전과 개혁 방안을 제시하였다. 구체적으로 거주자의 안전 확보, 불만 제기 절차 개선, 거주자 의견 청취와 책무성 확보, 사회주택 공급 확대, 자가 소유 지원 등 다섯 가지 원칙을 제안하였다. 정부는 녹서 발표와 함께 임차인을 대상으로 참여 행사 및 여론조사를 진행하여 2018년 11월까지 1만 건이 넘는 국민 의견을 수렴하였으며, 이 과정에서 전국 사회주택 거주자의 다양한 목소리가 접수되었다. 의견 수렴은 서면 제출(온라인·우편)과 지역 간담회 등 사회적 대화 방식을 통해 이루어졌다. 이런 절차를 거쳐 영국 정부는 2020년 11월 백서인 “The Charter for Social Housing Residents(사회주택 입주자를 위한 헌장)”를 발간하고 이를 통해 임차인의 권리 강화, 주택 품질 개선, 감독체계 개편 등의 구체적 정책을 제시하였다(Cromarty, 2023). 이 사례는 녹서에서 제기된 논의가 2년간의 공론화와 준비 과정을 거쳐 백서로 구체화된 대표적 사례로, 정부와 시민 간 소통을 통해 사회적 의제를 실질적 정책 성과로 연결한 모범적 사례로 평가된다.

영국 정부가 녹서를 활용한 또 다른 사례로는 2016년 보수당 메이 정부가 발표한 “Work, health and disability green paper: Improving lives”가 있다. 당시 고용연금부(Department for Work and Pensions, DWP)와 보건부(Department of Health, DH)는 장애인 고용 확대와 질병으로 인한 실직 방지를 주요 사회적 의제로 부각하며 공동으로 이 녹서를 발간하였다.¹⁷⁾ 장애인과 장기질환자가 더 오래 일할 수 있도록 지원하고, 이들의 일자리 활동을 통해 건강을 증진하는 방안을 모색하는 데 중점을 두었으며, 특히 비장애인 대비 약 30%에 달하는 장애인 고용률 격차 해소와 질병으로 인한 고용 단절의 악순환을 차단하는 데 정

/consultations/a-new-deal-for-social-housing, 검색일: 2025. 8. 17)를 참조하여 작성.

17) 녹서 및 구체적인 정보는 영국 정부 내 웹페이지(<https://www.gov.uk/government/consultations/work-health-and-disability-improving-lives/work-health-and-disability-green-paper-improving-lives>, 검색일: 2025. 8. 17)를 참조하여 작성.

책적 초점을 맞추었다. 녹서에는 현황 분석과 함께 고용지원, 직장 내 건강관리, 복지제도 개선 등에 관한 약 40개의 설문형 질문이 포함되었으며, 고용주·의료계·장애인 당사자·일반 국민을 대상으로 폭넓은 의견 수렴이 진행되었다. 예를 들면, ‘어떻게 하면 장애인이 더 오래 일할 수 있도록 지원할 수 있는가?’와 같은 질문을 던져 정책 아이디어를 구체적으로 도출하였다. 영국 정부는 2016년 10월부터 2017년 2월까지 4개월간 온라인 설문조사와 지역별 공청회를 병행하여 총 6,000여 건의 응답을 접수하였고, 기업, 복지단체, 장애인 당사자 등의 폭넓은 참여가 이루어졌다. 이러한 사회적 대화를 토대로 영국 정부는 2017년 11월 “Improving lives: the future of work, health, and disability”라는 백서를 발표하였다. 백서는 녹서에서 제기된 공개 질의에 대한 정부의 공식적 응답으로서, 장애인의 일자리를 10년 내 100만 개 추가 창출한다는 목표를 제시하였으며, 장애 수당 개편, 직장 건강관리 서비스 도입 등 40여 개의 구체적 실행계획을 포함하였다. 정책화 과정에는 장애인 단체 대표와 고용주 단체가 직접 참여하여, 전문가 중심 논의에 이해당사자의 목소리를 결합했다는 점에서 의미 있는 사회적 속의 모델로 평가된다.

이와 더불어 2017년 영국 디지털·문화·미디어·스포츠부(Department for Digital, Culture, Media & Sports)는 “Internet Safety Strategy Green Paper”를 발표하며 온라인상에서 발생하는 각종 위협(online harms)에 대응하기 위한 정책 논의를 본격적으로 시작하였다. 이 녹서는 디지털 경제가 제공하는 기회를 환영하면서도, 동시에 인터넷이 사용자에게 초래할 수 있는 위험을 최소화하기 위한 전략적 접근을 제시하였다. 녹서는 다양한 이해관계자, 전문가, 시민사회, 기업 등과의 협력을 통해 작성되었으며 발표 직후 일반국민을 대상으로 8주간의 공개 협의(consultation)가 진행되었다. 이런 공론화 과정을 토대로 영국 정부는 2019년 4월 “Online Harms White Paper”를 발간하여 구체적인 규제 정책 수단을 제시하였다.

나. 독일

독일은 사회적 대타협의 전통과 노사정 및 시민사회 간 대화가 활발히 발달한 국가로 평가된다. 그 대표적 사례가 바로 디지털 전환 시대 노동의 미래를 다룬 “노동 4.0(Arbeiten 4.0)” 녹서이다(정운찬, 2018. 11. 19). 이 녹서는 ‘디지털화라는 사회적 변동 속에서 좋은 노동의 기준을 어떻게 지켜낼 것인가’라는 근본적인 질문을 제기하며, 노동시간 유연화, 원격 근무, 평생학습, 사회보장 등 미래 노동시장을 둘러싼 10대 핵심 쟁점을 제시하였으며, 각 쟁점마다 정부의 문제 인식과 더불어 공론화를 위한 토론 질문을 함께 포함하였다(Federal Ministry of Labour and Social Affairs of Germany, 2015). 이 녹서의 가장 큰 특징은 국가 전체에 중대한 영향을 미치는 의제에 대해 전 국민적 차원의 숙의 과정을 거쳤다는 점이다. 노사정 대표, 업종별 협회와 기업, 학계 전문가, 일반 시민까지 폭넓은 참여가 이루어졌으며, 전국 단위 토론회와 공청회가 개최되었다. 또한 온라인 플랫폼을 통해 시민 의견을 수렴하여, 다양한 사회 구성원의 제안이 정책 논의에 반영될 수 있도록 했다(박찬수 외, 2024). 이와 같은 숙의 과정을 기반으로 독일 정부는 2017년 11월 “노동 4.0 백서”를 발표하였다. 백서에는 노동법 개정 방향, 노동 안전망 강화, 근로시간제 유연화 가이드라인 등 4차 산업혁명 시대 노동정책의 대응 전략이 구체적으로 담겼다. 이 백서는 녹서 발간 이후 2년 이상에 걸친 사회적 논의와 합의를 거쳐 정책을 구체화한 것으로, 심도 있는 문제 정의와 민주적 정당성을 확보함으로써 정책 추진 동력이 강력하게 뒷받침된 사례로 평가된다(아이뉴스, 2021. 1. 11).

다. EU

EU도 다양한 영역에서 녹서를 발간하여 범유럽 차원의 의견을 수렴하고, 이를 토대로 정책과 입법을 추진하고 있다. 대표적인 사례로 EU 집행위원회는 유럽 사회의 급속한 고령화 문제를 다루기 위해 2021년 “Green Paper on Ageing: Fostering solidarity and responsibility between generations”를 발간하였

다(European Commission, 2021). 이 독서는 고령화의 속도와 규모, 사회경제적 영향 등을 분석하고, 교육, 노동, 연금, 건강, 돌봄 등의 분야에서 고령화로 인한 영향 속에서 어떻게 세대 간 공평성을 유지하며 복지를 지속가능하게 만들 것인지에 대한 문제를 제기하였다. 독서 발간 후 12주간의 공개 의견 수렴을 시작하여 EU 회원국 정부, 각 분야의 전문가 단체, 시민사회, 일반 시민으로부터 폭넓은 의견 수렴을 진행하였으며, 온라인 설문과 토론을 통해 EU 각국의 의견서 및 추천 건의 시민 응답이 제출되었다. 다만, EU의 독서는 권고적 성격이 강하기 때문에 이후 구체적인 백서로 이어지지는 않았고, 회원국들이 국가 차원의 고령사회 전략을 수립하는 참고 자료로 활용되고 있다. 그럼에도 이 백서는 EU 공동의 문제를 회원국과 시민들이 함께 논의하는 장을 만들었다는 점에서 의의가 있고, 특히 세대 간 형평성 같은 민감한 주제에 대해 공론의 기반을 조성한 것으로 평가되고 있다(European Commission, 2021).

라. 미국

미국 연방정부 역시 독서를 활용하여 미래 기술 의제를 공론화한 사례가 있다. 미국에서는 공식적으로 ‘Green Paper’라는 용어를 자주 쓰지는 않지만, 정부 부처나 태스크포스(Task force)에서 발간한 정책 보고서가 독서의 역할을 하는 경우가 있는데, 2016년 8월 미국 상무부(The Department of Commerce) 산하 National Telecommunications and Information Administration(NTIA)에서 발표한 “Fostering the Advancement of the Internet of Things(사물인터넷 진흥을 위한 독서)”가 대표적 사례라 할 수 있다. 이 독서는 IoT(사물인터넷)가 국가 경제 성장, 공공서비스 혁신, 산업 경쟁력 강화에 미칠 수 있는 잠재적 기회와 위험 요인을 분석하고, 개인정보 보호 및 보안, 스펙트럼 및 인프라 자원 관리, 혁신 촉진을 위한 표준화, 공공 안전과 교통, 보건 분야에서의 활용 가능성 등을 핵심 의제로 제시하였다. 또한 각 의제별로 정책적 선택지와 규제 방향을 제안하기 보다 산업계·학계·시민사회·정부기관 간 토론을 촉진하기 위해 질문 형식으로 구

성된 것이 특징이다(NTIA, 2017). 녹서 작성 과정에는 기업(IT·제조업), 표준화 단체, 학계 연구자, 시민단체, 지방정부 관계자 등이 참여하였다. 이 과정에서 사이버보안 프레임워크, 자율주행차 안전 규제, 스마트시티 인프라 설계 등 구체적인 정책 쟁점들이 드러났다. 이후 미국 정부는 이 녹서에서 도출된 논의를 토대로 IoT 보안·프라이버시 리스크 관리 지침 등의 후속 정책·가이드라인을 도출하였다.¹⁸⁾ 비록 영국이나 독일과 같이 명시적인 백서로 이어지지 않는으나, 이 녹서는 미국 내 IoT 정책 개발의 기초 자료로 기능하였다.

마. 한국

한국의 정책문화에서는 유럽이나 영국처럼 제도화된 공식 녹서(Green Paper) 체계가 확립되어 있지 않지만, 유사하게 미래 의제 관련 연구보고서를 발간하거나 사회적 대화를 통해 녹서 발간을 시도한 사례들이 존재한다.

대표적인 사례로 광주광역시 광산구의 “지속가능한 일자리를 위한 녹서”(2025)를 들 수 있다. 이 보고서는 전국 기초지방자치단체 가운데 최초로 시민참여형 정책설계 과정을 통해 발간된 것으로, 제조업·서비스업·마을 일자리 등 다양한 분야의 주민들이 지역 일자리 문제에 대해 직접 고민하고 토론한 결과물을 담았다. 녹서의 핵심 질문에는 ‘일터에서의 세대 간 갈등을 어떻게 줄일 것인가’, ‘여성·장애인·이주노동자에게 좋은 일자리는 무엇인가’, ‘디지털 전환과 기후위기에 대응하기 위한 교육정책은 어떻게 마련해야 하는가’, ‘기후위기를 직접 체감하는 노동자들의 불안을 어떻게 해소할 수 있는가’, ‘광주 지역 청년들이 선호하는 양질의 일자리가 부족한 이유는 무엇인가’ 등 현장의 현실적 고민이 포함되었다(채준호 외, 2025).

녹서 제작을 위해 광산구는 2024년 5월 ‘지속가능 일자리 사회적 대화 추진단’을 출범시키고, 같은 해 6월부터는 시민 112명으로 구성된 ‘의제 발굴단’을 조직

18) 관련 내용은 Considerations for Managing Internet of Things(IoT) Cybersecurity and Privacy Risks(<https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8228.pdf>, 검색일: 2025. 8. 17)를 참조하여 작성.

하였다. 이들은 다섯 차례의 정례 토론과 함께 구 전역의 21개 동을 순회하며 총 12회의 ‘찾아가는 사회적 대화마당’을 개최하였다. 이러한 과정을 통해 총 1,436개의 질문이 도출되었고, 이후 이를 핵심 질문 20개로 압축하였다. 이 질문 목록을 바탕으로 정책담당 공무원과 전문가들이 함께 논의하여 정책 제언서를 완성하였다(무등일보, 2025. 5. 7).

이 사례의 의의는 기초지자체 단위에서 정책 결정을 위한 공식 문서로서의 구속력은 없더라도, 시민이 스스로 문제를 정의하고 질문을 제기하는 상향식 의제 설정을 제도적으로 실험했다는 데 있다. 행정 주도가 아닌 시민 참여형 접근을 통해 지역 민주주의의 활성화를 모색하고, 행정에 대한 주민 신뢰를 제고했다는 점에서 긍정적인 평가를 받고 있다. 또한 녹서라는 형식을 차용하여 사회적 대화를 공식 기록물로 남김으로써, 향후 정책 수립 과정에서 주민 의견을 반영할 수 있는 토대를 마련했다는 점에서 의미가 크다(채준호 외, 2025).

2. 전문가 주도의 사회적 의제 연구와 녹서 발간 선행 사례

사회적 대화 과정을 전면에 내세우는 유럽식 녹서와 달리, 높은 전문지식을 요구하거나 미래 불확실성이 커서 일반 시민이 직접 파악하거나 논의하기 어려운 사안일 때는 전문가 집단이 주도하는 사회적 의제 연구 혹은 전문가 집단 중심의 녹서가 주로 활용된다. 이런 유형의 녹서는 정책 형성 초기 단계에서 문제구조를 정밀하게 진단하고, 쟁점을 체계적으로 정리하는 데 초점을 두며, 공공의 참여보다는 전문지식 기반의 분석 및 대안 논의를 핵심으로 한다. 홍콩대 경영대학에서 발간한 “Hong Kong Economic Policy Green Paper 2025”는 전문가 주도 녹서의 대표적 사례라 할 수 있다. 홍콩은 국제 금융·무역의 허브로 대외 의존도가 높고, 미·중 관계변화, 디지털 전환, 인구구조 변화 등 복합적 변수로 인해 경제정책 전반에 중대한 불확실성이 발생한다. 이에 따라 해당 녹서는 시민참여형 토론회보다는 학계 전문가들이 중심이 되어 거시경제 안정, 혁신역량 강화, 노동시간 조

정 등의 핵심 의제를 설정하고 분석하였다.¹⁹⁾ 이러한 방식은 전문적 지식에 기반하여 문제 현황 및 정책 방향을 도출하는 데 효과적이라는 장점이 있다.

국내 경제·인문사회연구회의 ‘국가의제 종합연구’ 역시 전문가 집단이 주도한 사회적 의제 연구보고서로 볼 수 있다. 이 연구는 중장기적으로 한국 사회가 직면할 국가적 과제를 발굴하고 정책 방향을 제시하는 데 목적을 두고 있다. 이를 위해 국민참여형 설문이나 공청회를 활용하기보다는 각 분야 연구자들이 기존 자료와 정책환경을 심층 분석하여 사회·경제·환경·과학기술 전반의 핵심 의제를 정리하는 방식으로 진행되었다(NRC 국가의제연구단, 2023). 이는 체계적이고 학술적인 근거에 기반한 국가 의제 발굴이라는 점에서 의미가 크다고 할 수 있다. 특히 국민 다수가 체감하기 어렵지만 국가 경쟁력에 중요한 영향을 미칠 수 있는 인구구조 변화, 기술혁신 등 복잡한 주제를 다룰 때 전문가 중심 접근법의 필요성이 크다고 할 수 있다.

이와 같이 전문가 주도의 녹서는 시민이 체감하지 못하거나 참여하기 어려운 전문적이고 미래 지향적인 정책 의제를 다룰 때 강점을 발휘한다. 사회적 대화형 녹서가 민주적 정당성과 공론화를 통한 사회적 합의 확보에 기여한다면, 전문가 주도형 녹서는 정책 설계의 정밀성을 높이고 고도의 전문 지식을 바탕으로 복잡한 의제를 체계적으로 해석하며 중·장기적 사회 변화에 대응할 수 있는 전략적 로드맵을 마련하는 데 중요한 역할을 수행할 수 있다.

19) 관련 내용은 홍콩대 경영대학(The University of Hong Kong Business School)의 “The Hong Kong Economic Policy Green Paper 2025”를 참고하여 작성.

〈표 2-2〉 시민참여형 녹서와 전문가 주도형 녹서의 비교

구분	시민참여형 녹서	전문가 주도형 녹서
주요 목적	• 사회적 공론화와 합의를 통한 정책 정당성 확보	• 전문 지식을 기반으로 한 문제구조 진단 및 정책 방향 제시
주제 영역	• 시민이 직접 체감 가능한 생활·복지·노동 등 사회적 현안	• 고도의 전문성·불확실성이 큰 경제, 과학기술, 인구·환경 등 의제
참여 주체	• 일반 시민, 이해관계자, 시민단체, 지역사회 등	• 학계 연구자, 정부·공공기관 전문가, 산업계·전문 단체
운영 방식	• 공청회, 토론회, 설문조사 등 숙의 과정 중심	• 전문가 워크숍, 연구회의, 문헌분석 등의 연구·분석 기법 중심
장점	• 민주적 정당성 확보, 정책에 대한 사회적 수용성 증대, 다양한 이해관계자의 의견 수렴을 통한 정책 사각지대 해소	• 정책의 정확성·전문성 제고, 장기적 전략 수립 가능
한계	• 전문성이 요구되는 복잡한 의제에는 한계 • 시간과 비용의 부담	• 시민 및 이해관계자 참여 부족으로 정책 정당성 확보에 상대적 제약 존재
대표 사례	• 독일 Arbeiten4.0(2015) • 영국 Work, health, and disability(2016)	• Hong Kong Economic Policy Green Paper 2025 • 경제·인문사회연구회 국가의제 종합 연구

자료: 연구진 작성

제 3 절 사회적 의제 프레임워크

앞선 절에서 살펴본 바와 같이, 사회적 의제 연구와 독서는 공공정책의 논의와 형성 과정에서 핵심적인 역할을 수행한다. 본 절에서는 영국, 독일, 유럽연합(EU), 미국, 한국, 홍콩 등에서 발간된 사회적 의제 연구와 독서 사례를 검토함으로써, 이들 문서가 어떠한 체계와 구조로 구성되어 있는지를 분석하고자 한다. 나아가 이러한 비교 분석을 바탕으로, 사회적 의제 연구와 독서가 정책적 의사결정 과정에서 효과적으로 기능하기 위해 반드시 갖추어야 할 구성 요소와 조건을 구체적으로 제시하고자 한다.

1. 사회적 의제 사례 분석

영국의 “New Deal for Social Housing”(2018) 독서는 정부의 사회주택 비전과 함께 안전 강화, 거주자 권리 신장, 공급 확대 등 다섯 가지 핵심 원칙을 제시하고 각 원칙별 개혁 방안을 담는 구조로 작성되었다. 독서 본문은 화재 참사로 드러난 임대주택 문제들을 체계적으로 진단하면서 주택 안전과 품질 개선, 임차인 불만 처리 절차 개편, 주민 의견 청취와 책무성 확보, 사회주택 신규 공급 및 자가 소유 지원 등의 세부 주제를 장별로 다루었다. 특히 각 장에서 현황과 문제점을 서술함에 있어, 이에 대한 거주자들의 의견을 함께 제시한 것이 특징이라 할 수 있다. 현황-문제점-거주자 인터뷰 내용을 수록한 후 문제 현상의 개선을 위한 정부 정책 방향을 제시하는 방식으로 독서가 구성되었다(Ministry of Housing, Communities & Local Government, 2018).

또 다른 영국 사례인 “Work, Health, and Disability: Improving Lives”(2016) 독서는 장애인의 노동시장 참여 기회를 늘리고 건강 문제로 인한 경제활동 중단을 줄이기 위한 종합 대책을 논의함에 있어서, 현황 진단과 함께 고용지원, 직장 내 건강관리, 복지제도 개선 등 세부 주제별로 약 40개의 설문형 질문을 제시하는 구성을 취하였다. 각 장마다 정부의 문제 인식과 기존 정책을 간략히 소

개하고, 이어서 ‘어떻게 하면 장애인이 더 오래 일할 수 있도록 지원할 수 있는가?’와 같은 형태의 질문들을 열거함으로써 정책 아이디어를 모색하였다. 이러한 구성을 통해 해결책 제시보다는 각 질문에 대한 숙의를 촉진하는 데 중점을 두었다.

영국의 “Internet Safety Strategy Green Paper”(2017)에서는 서문에서 ‘디지털 시대 최안전 국가 수립’이라는 비전을 제시한 후, 본문에서 온라인 유해 콘텐츠, 아동보호 등의 광범위한 주제들을 다루었다. 각 주제별로 현황과 문제점을 기술하고, 정부가 고려 중인 대응 조치를 개략적으로 소개하면서도 최종 해법에 대해서는 다양한 이해관계자들의 추가 의견을 구하는 형식으로 보고서가 작성되었다.

독일의 ‘노동 4.0(Arbeiten4.0)’의 경우, 보고서 1장에서는 독일 노동환경 변화의 거시적 흐름(인구구조, 글로벌화 등)을 개괄한 후, 2장에서는 정부가 구상하는 ‘노동 4.0’의 개념과 방향을 제시하였다. 핵심이 되는 제3장에서는 본격적으로 노동시간 유연화, 숙련노동자 변화, 평생학습, 사회보장 등의 쟁점 분야를 선정하여 각 쟁점마다 정부의 문제의식과 함께, 토론을 위한 구체적인 질문들을 제시하였다. 제4장에서는 노사정 협력 등 기존 사회적 대화 제도의 현황을 다루었고, 제5장에서는 향후 사회적 대화를 통해 녹색에서 백서로 연계되는 단계를 제시하였다. 미래 노동정책 의제를 질문 형태로 제시함으로써, 백서로 연계되는 단계에서 논의될 대안의 폭을 열어두었다.

유럽연합의 “Green Paper on Ageing: Forstering solidarity and responsibility between generations”(2021)은 크게 두 부분으로 구성되어 있는데, 첫 부분에서는 유럽의 인구 고령화 속도와 규모, 사회경제적 영향을 데이터에 기반해 상세히 분석하고 이를 통해 기대수명 연장, 출산율 저하, 노령인구 부양비 급증 등의 현황과 추세를 제시하고, 이러한 변화가 연금·노동·건강·복지 시스템에 미치는 영향을 진단했다. 그 후 교육, 노동시장, 연금제도, 건강·요양 돌봄 등 여러 분야에서 세대 간 형평성을 유지하고 지속가능성을 확보하기 위한 정책 방향을 논의하였다. 예를 들면 ‘어떻게 세대 간 공정성을 보장할 것인가’ 등의 질문을 던지며, 다양한 대안을 모색하도록 구성되어 있으며, 보고서 결론 부분에는 EU 차원의 공

동 대응 필요성을 강조함과 동시에 회원국 정부와 시민들의 공개적인 토론을 요청하기도 하였다.

미국의 “Fostering the Advancement of the Internet of Things”(2016)도 앞선 사례들과 유사한 체계로, 사물인터넷이 국가 경제 성장, 공공서비스 혁신, 산업 경쟁력 강화에 가져올 기회와 위험을 분석하고, 이후 각 장에서 핵심 의제별로 현행 정부 정책과 민간 동향을 소개하고 제도적 공백이나 해결과제를 식별한 뒤, 마지막에 해당 분야에서 향후 논의가 필요한 쟁점을 질문 형태로 제시하였다.

마지막으로 국내 “광산구 지속가능한 일자리 녹서(2025)”는 주민 참여 과정과 문제의식을 서론에서 제시한 후, 본문에서 제조업, 서비스업, 마을 일자리 등 분야별로 주민들이 제안한 핵심 질문 20가지를 제시하였다. 각 질문마다 해당 문제에 대한 배경 설명과 현장 의견을 서술하였고, 이를 토대로 정책 제언 및 대안의 토의 결과를 제시하였다. ‘청년들이 선호하는 양질의 지역 일자리가 부족한 이유는 무엇인가’ 등의 질문을 제시하고, 질문별로 주민 토론에서 나온 아이디어와 전문가 검토 내용을 정리한 단락이 추가된 형태로 녹서가 정리되었다. 이런 구성은 주민들이 던진 질문을 중심에 두고 이에 대한 해답을 행정가와 전문가가 함께 모색하는 문제 중심형 보고서라고 할 수 있다(채준호 외, 2025).

시민참여형 중심의 녹서와 유사하게 전문가 중심의 사회적 의제 연구인 홍콩의 “Hong Kong Economic Policy Green Paper 2025”와 국내 경제·인문사회연구회의 ‘국가의제 종합연구’ 역시 이와 유사한 체계와 양식을 갖추고 있다. 경제·인문사회연구회의 ‘국가의제 종합연구’(2023)는 서론에서 연구 배경과 방법론을 설명한 후, 본문에서 경제, 사회, 기후·환경, 정부·거버넌스, 대외의 5개 분야로 나누어 각 분야별 핵심 정책 의제를 다루었다. 각 분야별로 현황 진단-의제 도출-정책 방향 제언의 일관된 서술 체계를 가지고 각 세부 과제를 다루는 방식으로 보고서가 작성되었다.

2. 사회적 의제 구성 요소

지금까지 기존 사회적 의제 연구와 녹서 사례를 살펴본 결과, 사회적 의제 연구와 녹서는 단순한 연구보고서가 아니라 정책의 논의와 사회적 대화를 촉진하는 문서라는 점에서 그 구성과 절차, 특징에 있어 일정 요소들이 필요하다. 가장 먼저 사회적 의제 연구 및 녹서에서는 서론부에 배경과 문제 인식을 명확하게 제시하고 있다. 특정 주제를 사회적 의제로 설정하게 된 정책환경 변화 혹은 문제의식을 명확히 밝히고, 이어서 연구의 목적과 정책적 목적을 제시하고 있다. 핵심은 사회적 의제 연구나 녹서가 단순히 정책 대안을 제시하는 것이 아니라 논의의 장을 마련하거나 정부 정책의 시급성을 제시하여 해당 이슈에 대한 공론화의 목적을 가진다는 점을 명확히 한다는 점이다.

또한 현황 진단 및 문제 정의 부분에서는 통계자료, 기존 연구 성과 등을 기반으로 현황을 분석하고 그 과정에서 해당 문제가 왜 중요한지, 어떠한 구조적 요인이 작동하고 있는지를 설명해야 한다. 이러한 현황 진단을 바탕으로 핵심 의제와 정책 질문이 도출되는데, 녹서의 경우 ‘어떻게 할 것인가’ 혹은 ‘무엇이 가장 효과적인 방안인가’의 형식으로 질문을 도출하는 반면, 사회적 의제 연구에서는 전문가의 의견을 바탕으로 보다 구체적으로 정책의 방향성에 대한 내용이 제시되기도 한다. 핵심 의제와 정책 질문들은 우선순위를 정하거나 범주화를 통해 제시되기도 한다. 마지막으로, 사회적 의제 연구 및 녹서 모두 정책 대안은 닫힌 형태의 결론이 아니라 다양한 가능성을 제시하는 형태로 논의된다.

녹서와 사회적 의제 연구는 무엇보다 문제 제시형 보고서라는 점이 가장 큰 특징이라고 할 수 있다. 답을 제시하기보다는 문제를 열어 놓고 사회적 논의를 촉진하는 기능을 수행해야 한다. 정부 주도가 아닌 전문가 집단 혹은 국민참여가 강조되어야 하며, 특정 해결책을 제시하기보다는 다양한 정책 방향을 함께 제시함으로써 논의의 장을 열어두는 체계가 필요하다. 무엇보다도 이러한 논의가 향후 법률 개정, 제도 개선 등으로 이어질 수 있도록 정책 연결성이 분명히 제시되어야 한다. 즉 문제 정의-질문 제시-토론 유도의 구조적 형식과 함께, 의제 발굴과 의견

수렴 과정에서 전문가 집단 및 이해관계자, 국민참여, 그리고 제도 개혁으로 이루어지는 정책 연결성에 대한 요건들이 충족될 때, 사회적 의제 연구와 녹서가 단순한 연구 결과물이 아닌 사회적 공론화의 장이자 정책 형성 과정의 촉매로서 작동할 수 있다.

제 3 장 5대 분야별 사회적 의제 심층 탐색

제 1 절 AI와 노동: 정형 편향적 기술 발전의 경로 전환

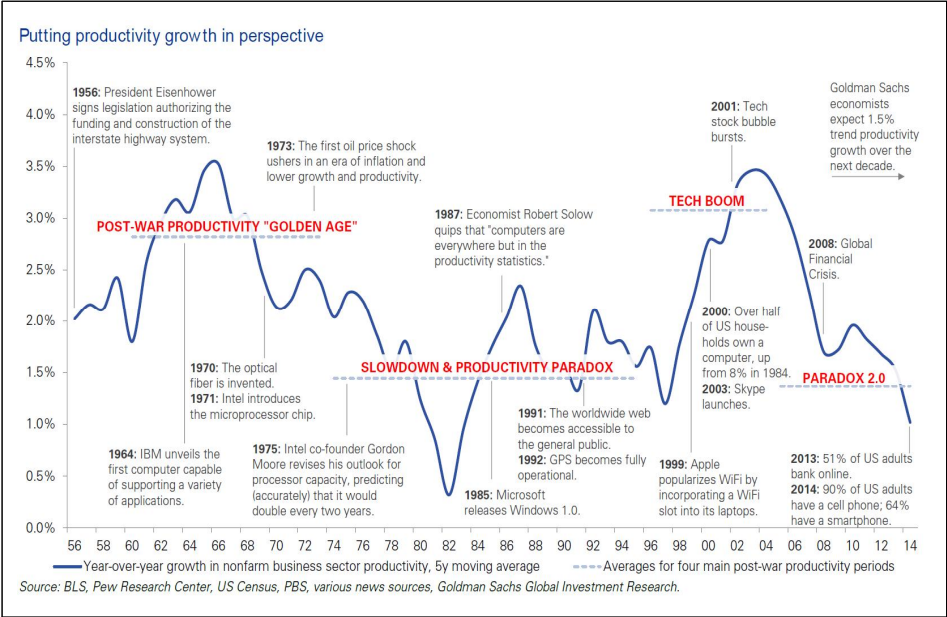
1. 논의의 배경

가. 생산성 효과: 평균의 함정

AI는 그간 지속적으로 관찰된 생산성 역설을 극복할 수 있는 기술로 기대되고 있다. 그간 기술의 발전은 효율성을 개선하며 생산함수 내 총요소생산성(total factor productivity, TFP)과 노동생산성을 함께 증가시켜 경제 성장에 기여할 것으로 기대되었으나 실제로 그 효과는 미미한 수준으로 포착되어 왔다. 이러한 현상에 대해 David Solow는 ‘생산성 역설(productivity paradox)’로 설명한 바 있다. 최근 AI의 등장 이후 다수의 연구를 통해 미시적 수준에서 생산성 개선 효과가 증명되고 있으며, 이러한 효과가 경제 전체에 누적된다면 통계상 주요 생산성과 경제 성장 지표가 향상될 것으로 예측된다(Brynjolfsson et al., 2023; Noy and Zhang, 2023; McKinsey, 2023).

장기적으로 기술은 진보하고 있음에도 경제의 생산성과 성장 지표에 반영되지 못하는 이 현상을 두고 기술경제학자인 Erik Brynjolfsson은 ‘생산성 J 곡선’ 개념을 통해 이를 극복할 수 있는 조건이 무엇인지 제시하였다(Brynjolfsson et al., 2021). 생산성 J 곡선은 X축의 기술 진보가 Y축의 성장 효과에 반영되는 데까지는 시차가 존재하여 우하향 곡선이나, 이후 제도적 보완과 사회적 수용이 개선되면 그 변곡점을 지나 우상향한다는 개념이다. 2022년 LLM의 상용화는 기술적 진전이 사회적 수용을 폭발시키며 생산성의 변곡점을 지나는 주요한 동인이 될 수 있다.

[그림 3-1] 생산성 역설의 역사



주: Goldman Sachs(2015), Picking apart the productivity paradox, Top of Mind, Issue 29, Goldman Sachs

생성형 AI의 등장으로 단순 반복되는 업무가 아닌 비정형적인 인지 업무에도 AI가 활용됨에 따라 업무의 생산성을 증가시킬 것이 기대되었다. 사회적 기대감을 반영하듯 2023년에는 AI가 지식노동 혹은 서비스 종사자의 일의 성과나 생산성을 증진하는가를 실증적으로 검증하는 연구가 다수 수행되었다. 《Science》지에 등재된 Noy and Zhang(2023)의 연구는 글 쓰는 업무가 직업인 프리랜서들의 작성 속도와 성과물의 질을 ChatGPT가 개선하는 것을 확인하였으며, Brynjolfsson et al.(2023) 또한 현장 실험을 통해 대화형 AI 어시스턴트가 고객센터 상담원의 시간당 해결 건수를 향상하는 것으로 분석하였다. 특히 이러한 효과는 노동자의 기존 숙련도에 따라 이질적으로 나타났다. 즉 경력이나 숙련도가 상대적으로 낮은 집단이 더 큰 성과 개선 효과를 경험하는 것으로 나타났다. 이와 유사한 일련의 연구들은 AI가 저숙련 그리고 신규 근로자에게 기존 숙련자의 암묵적인 지식을

전수하는 매개물의 역할을 수행할 수 있음을 보여준다.

반면 Humlum & Vestergaard(2025)가 덴마크의 11개 직종, 노동자 2만 5,000명을 대상으로 수행한 연구에서는 다른 가능성을 보여준다. AI 챗봇이 직장에 광범위하게 도입되었음에도 노동자들의 임금이나 기록된 근로 시간에 통계적으로 유의미한 변화를 가져오지 않았음이 확인되었다. 연구진은 이러한 결과의 원인을 두 가지로 분석했다. 첫째, 실험실 환경이나 특정 과업에 집중된 연구(RCT)에서 보고된 15~50%의 높은 생산성 향상과 달리, 실제 현장에서 노동자들이 체감하는 생산성 향상 효과는 평균 2.8%의 시간 절약 수준으로 훨씬 미미했다. 둘째, 그 미미한 생산성 향상마저도 임금 인상으로 이어지는 비율이 3~7%에 불과할 정도로 매우 약했다. 대신 생산성보다는 새로운 업무를 파생하며 신규 업무에 시간을 소요하는 것으로 나타났다. 또 다른 연구인 Kim and Moon(2024)은 생성형 AI는 환각이 발생하는 영역에서의 활용은 오히려 성과 개선을 저해한다는 점을 밝혀 생산성 효과가 미미할 수 있음을 제시한다.

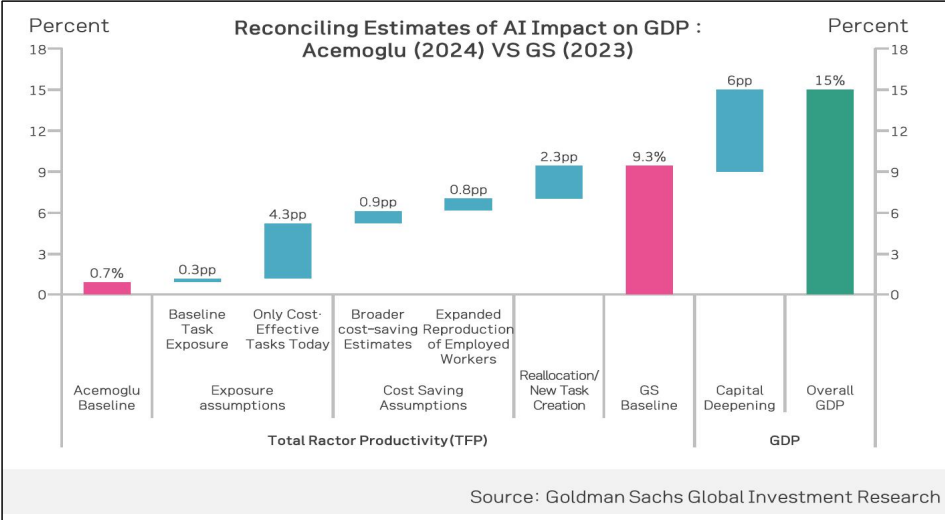
따라서 AI는 미시적 생산성을 개선할 수 있겠으나, ‘누구’의 생산성을 ‘어떻게’, 그리고 ‘얼마나’ 높이는가에 대해서는 불확실성이 존재한다. 결국 AI로 인한 생산성 향상은 모두에게 보장되는 것이 아니라는 점에서 기대했던 것과 달리 이번 AI 전환으로도 생산성 J 곡선의 변곡점을 통과하지 못할 가능성이 있다. Noy & Zhang(2023)이나 Brynjolfsson et al.(2023) 등의 기존 연구에서 나타난 미시적 수준의 업무 생산성 개선 효과는 AI 도입과 함께 기존 업무 프로세스의 재설계, 조직 내 시스템 최적화, 업무 관리 방식의 변화 등 보완적 투자가 동반되어야 나타날 수 있는 ‘최상의 시나리오’일 수 있다.

나. 성장 효과의 단절

AI의 경제 성장에 대한 기여도는 경로별로 상반된 효과를 유발하므로 최종적으로는 단절될 수 있다. AI가 GDP에 미치는 영향을 추정한 Goldman Sachs와 Acemoglu 모두 동일한 이론적 틀(Hulten's theorem)에서 출발하나 제도와 분배 구조에 대한 시각 차이로 생산성과 성장의 기여도에 차이가 나타난 바 있다.

Goldman sachs(2023)는 현재 자동화 가능한 업무 전체에 AI가 도입됨을 가정하고 AI에 대한 기업의 투자가 자본 심화를 통해 성장에 기여하여 향후 10년간 GDP가 약 15% 증가할 것으로 기대하였다. 반면에 Acemoglu(2024)는 이를 반박하며 실제 비용 절감이 가능한 업무에 AI가 채택되고, 그 과정에서 수반되는 구조적 실업의 가능성으로 노동소득분배율을 악화시켜 임금 효과로 이어지지 못함에 따라 (AI 도입으로 인한) GDP 증가율이 2~3% 수준에 그칠 것으로 추정하였다.

[그림 3-2] AI가 GDP에 미치는 영향: Goldman Sachs(2023) vs Acemoglu(2024)



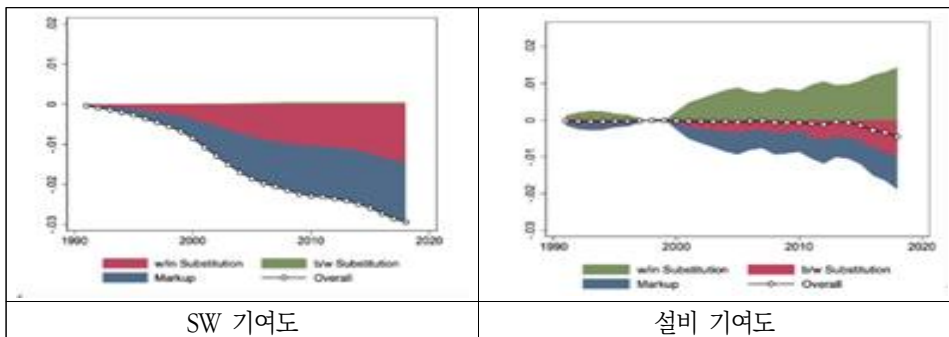
자료: Goldman Sachs(2024), “Top of Mind”

생산함수 내에서 자본과 노동 간의 보완성 조건을 충족하는 기술 진보나 자본 증가는 노동을 대체하지 않으면서도 생산성 향상을 유도할 수 있으면서 노동소득 분배율의 유지 또는 증가가 가능하다는 점에서 Acemoglu가 아닌 골드만 삭스의 추정치도 기대할 수 있을 것이다. 로봇, ICT 등의 생산성에 기여하는 기술은 자본 집약적 기술 진보라는 점에서 노동소득분배율 하락을 유도하는 하나의 원인으로 지목되어 왔다. 실제로 생산성(노동생산성, 총요소생산성)과 노동소득분배율은 음의 관계를 가지고 있으며, 이를 유발하는 여러 요인 중 기술적 동인은 생산성 향

상을 가져오는 기술이 새로운 자본재에 내재되어 상대적으로 자본소득에 집중된다는 점을 설명한다(Cho et al., 2025). AI로 인한 자본과 노동이 보완적이라면 동일 업무를 수행하는 단위 비용이 감소하는 ‘생산성 효과’가 발생하고 노동은 유지되므로 노동소득분배율은 장기적으로 개선되나, 노동 대비 기계 비용 하락으로 ‘대체 효과’만을 고려한다면 노동소득분배율은 하락하게 된다(Drozd and Tavers, 2024).

그렇다면 AI는 어떤 기술인가? Aum & Shin(2024)는 한국의 1990~2018년 기업 데이터를 통해 자본을 크게 설비와 소프트웨어로 구분하고 각 자본의 유형별로 노동과의 대체 탄력성과 노동소득분배율 증감의 기여 수준을 추정한 바 있다. 그 결과 소프트웨어와 노동 간의 대체탄력성은 1.6, 설비와 노동 간의 대체탄력성은 0.6으로 추정되어, 소프트웨어가 대체성이 큰 것으로 확인되었다. 또한 동기간 노동소득분배율 하락의 77%는 자본에 내재된 기술 진보에 기인한 것으로 분석되었으며, 그 가운데 67%는 소프트웨어 기반의 기술 진보가 원인이 되는 것으로 확인하였다. 로봇 등 물리적 기술과 결합하는 피지컬 AI의 도래가 기대되지만, AI의 본질이 다수의 애플리케이션을 창출하는 소프트웨어와 유사하다고 볼 수 있다면 AI 또한 노동소득분배율을 감소하고, 성장 효과를 단절하는 기술이 될 수 있다.

[그림 3-3] 노동소득분배율 하락 요인 분해



자료: Aum & Shin(2024)

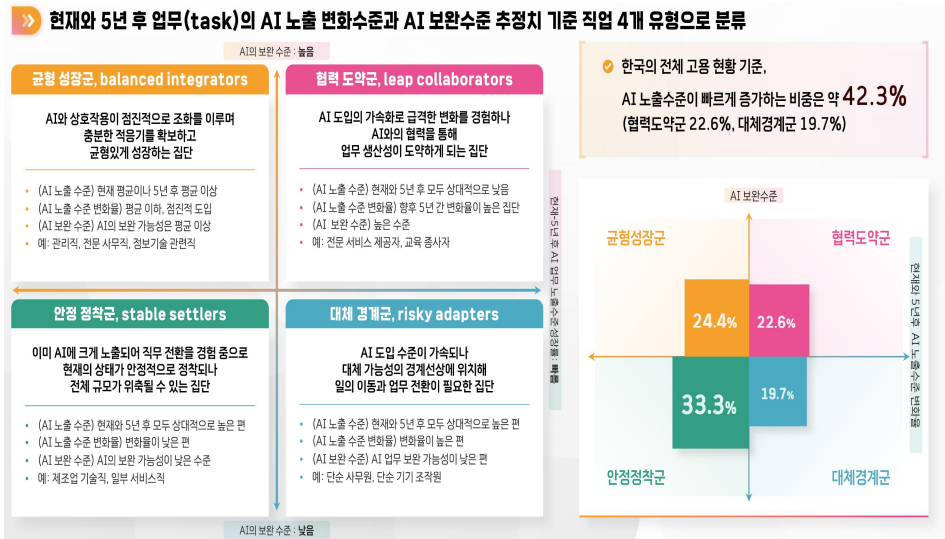
다. 노동 수요의 변화

AI가 노동 수요에 미치는 영향은 일자리를 구성하는 주요 업무(task)의 성격에 따라 ‘대체’ 혹은 ‘보강’이라는 양립하기 어려운 두 채널을 통해 발현되고 있다. Autor et al.(2024)은 미국 전체의 채용 공고 데이터를 분석하여 이러한 이중적 효과를 실증적으로 확인한다. 연구 결과, AI에 의한 자동화 가능성이 높은 상위 25% 직업(예: 정형화된 보고서 작성, 데이터 입력 등이 핵심 업무인 일자리)의 채용공고는 분기당 17% 감소했다. 반면, AI와 협업하여 인간의 능력을 증강시킬 수 있는 ‘보강’ 가능성이 높은 상위 25% 직업(예: 전략 기획, 창의적 설계 등)의 채용공고는 오히려 22% 증가했다.

이러한 결과는 생성형 AI가 노동 수요의 총량을 일방적으로 줄이기보다는, 노동의 성격과 구조 자체를 근본적으로 재편하고 있음을 시사한다. 이 과정에서 요구되는 스킬도 함께 변화한다. 온라인 채용공고 데이터 회사인 Lightcast는 AI 관련 스킬을 요구하는 직무는 그렇지 않은 직무에 비해 평균 28%의 임금 프리미엄을 제공하며, 2024년 기준 AI 관련 스킬 요구 직무의 51%가 IT 및 컴퓨터 과학 분야가 아닌 곳에서 발생하고 있음을 채용공고 데이터를 통해 확인하였다. 이는 노동 수요 측면에서는 이미 기존의 업무들이 AI로 ‘자동화가 가능한 직무’와 ‘AI가 보강하는 직무’로 구분하여 채용하고 있음을 시사한다.

기술 발전은 기존의 시장 질서를 재편하는 슈페터식의 ‘창조적 파괴’를 가속하나, 창조와 파괴 사이에는 항상 시차가 존재하므로 노동시장에서의 마찰로 기술적 실업이 발생한다. 구조적 실업이 발생하지만 동시에 새로운 산업과 직업이 등장하며 장기적으로는 균형을 찾아가게 되나, AI는 반복적이고 정형화된 업무에서 비정형적이고 인지적인 업무까지 영향을 준다는 점에서 기술적 실업이 나타나는 산업과 직군의 범위가 넓고 영향이 미치는 속도도 빠를 것이라고 분석되고 있다.

[그림 3-4] 시와 미래 일자리 영향 전망



자료: 문아람 외(2024), 『디지털 대전환 메가트렌드 미래전략 보고서』

국제 비교를 진행한 다수의 연구는 선진국의 경제 구조상 AI가 유발하는 노동 구조 변화에 예민하게 반응할 것으로 전망하고(Gymrek et al., 2023; 2024; IMF, 2024; WEF, 2025), 한국을 대상으로 한 다수의 연구에서도 현재 취업자 기준 AI에 노출된 비중이 최소 25%에서 많게는 50%를 상회하는 수준인 것으로 포착된다(한국은행, 2025; 장지연 외, 2024; 문아람 외, 2024).²⁰⁾

라. 가시화된 양극화와 진입 장벽

AI가 노동시장에 미치는 영향은 모든 노동자에게 균일하지 않다. 특정 기술을

20) 한국노동연구원(장지연 외, 2024, “AI와 노동”)는 AI 중 생성형 사전훈련 트랜스포머(GPT)에 대한 노출지표를 통해 한국 일자리 중 GPT에 노출된 고용 비중은 25.7%(증강 가능성 15.9%, 자동화 가능성 9.8%)로 추정함; 한국은행(2025, “AI와 한국 경제”)는 Felten et al.(2022; 2023)의 AI 노출지표를 활용하여 국내 일자리 중 51%가 노출되며, 이 중 보완수준도 높은 직군은 27%로 추정함; 정보통신정책연구원(문아람 외, 2024, 『ICT 기반 사회현안 해결방안 연구』)은 GPT를 기반으로 측정한 AI 노출지표를 활용하여 국내 고용시장의 56.7%가 영향을 받으며 이 중 보완수준도 높은 직군은 30.3%로 추정함.

보유한 소수, 특정 직업군, 그리고 특정 연령대에 속한 집단에 그 영향과 혜택이 집중되면서 새로운 형태의 양극화와 진입 장벽을 만들어내고 있다. 앞서 살펴본 바와 같이 분석 범위와 방법에 따라 AI가 고용 총규모에 미치는 영향에 차이가 있고 신직업의 출현을 반영하지 못하는 한계가 있으나, 한국은 AI로부터 대체될 가능성이 높은 직업보다는 보완 수준이 높은 직업을 좀 더 크게 포착한다는 점에서 희망적일 수도 있다. 다만, AI로 인한 기술적 실업의 가능성이 높은 집단은 여성일수록, 교육 수준이 낮을수록, 소득수준이 낮을수록 그 비중이 크다는 점에서 기존의 노동시장 내 취약 집단에게는 AI가 여전히 불리할 수 있다는 점도 공통적으로 발견된다.

현재 창조와 파괴 사이의 시차 가운데에서 그간 축적한 인적자본의 가치가 평가 절하되어 경력 사다리 붕괴가 가시화된 집단은 ‘청년’과 ‘신규 진입자’로 볼 수 있다. 이는 AI가 초급 수준의 경력이 요구되는 일자리를 자동화하는 효과와 나머지 일자리의 요구 역량을 고도화하는 효과가 동시에 발생하면서, 신규 인력이 노동시장에 진입하고 성장할 통로 자체를 막아버리는 위험한 병목 현상을 초래했기 때문이다. 과거 청년층이 노동시장에 첫발을 내딛고 현장 경험을 쌓는 중요한 통로였던 인턴십이나 신입사원 일자리는 대부분 정형화되고 반복적인 업무로 구성되어 있었다. 생성형 AI는 바로 이 영역에서 가장 강력한 대체 능력을 발휘하며, 기업들은 초급 인력을 채용하여 교육시키는 대신 AI 도구로 해당 업무를 자동화하거나, 즉시 AI를 활용할 수 있는 경력직 채용을 선호하는 경향을 보이는 것이다.

Anthropic의 CEO 다리오 아모데이는 AI의 발전으로 5년 이내 사회 초년생들의 사무직 일자리가 절반가량이 사라질 것이라고 언급한 바 있다.²¹⁾ 특히 기술, 금융, 법률, 컨설팅 등의 직군에서 요구되는 전문직 영역의 대졸자가 처음 진입하는 초급 직무들은 AI로 대체되므로 미국의 실업률이 10~20%가량 상승할 수 있다고 전망하였다. 또한 글로벌 전략 컨설팅 기업인 EY, 딜로이트, KPMG, PwC의 핵심 서비스인 조세, 재무, 회계 등의 AI 자동화가 가능해져 빅4 또한 오프쇼어링

21) Axios(2025. 5. 28), “Behind the Curtain: A white-collar bloodbath.”

방식으로 채용된 다수의 주니어 인력에 기반한 피라미드형 조직 구조를 유지할 필요성이 급격히 감소하고 있다고 인식하고 있다.²²⁾ 국내 온라인 채용공고 플랫폼인 잡코리아에서는 AI가 포함된 채용공고 수가 2022년 1/4분기 대비 28%가량 증가하였으며, 이 중 경력직은 53% 급증한 반면, 신입 채용은 단 5% 증가하는 데 그쳤다. 전환기의 노동시장은 즉시 투입 가능한 숙련 인력을 선호하므로 신규 인력을 채용하여 성장시킬 여력이 부족함을 시사한다.

이러한 초급 일자리의 소멸은 단순히 청년 실업 문제에 그치지 않는다. 이는 신규 인력이 현장에서 선배들의 멘토링을 받으며 암묵지를 습득하고, 실수를 통해 배우며 전문가로 성장해 나가는 전통적인 ‘경력 개발 경로’ 자체를 붕괴시킬 수 있다(Keenan, 2025). 경험을 쌓을 기회 자체가 사라지면, 미래의 숙련된 경력직 인력을 양성하는 파이프라인이 막히게 된다. 이는 중기적으로 산업 전체의 인력 수급 불균형을 심화하고, 장기적으로는 부모의 사회·경제적 지위가 자녀에게 전이되는 것을 막는 중요한 기제인 교육과 직업을 통한 ‘세대 간 이동성(inter-generational mobility)’을 심각하게 저해하며 구조적 양극화를 지속하는 위협으로 작용할 수 있다.

또 다른 가시화된 집단은 ‘엔지니어’로 이들은 개인의 능력과 환경에 따라 AI의 영향이 차등화되는 경험을 하고 있다. 그간 소프트웨어 개발자 직군은 AI에 의한 ‘증강’ 효과를 가장 극적으로 보여주는 사례로 분석되어 왔다. GitHub와 Accenture의 공동 연구에 따르면, Copilot을 사용한 개발자들은 코딩 속도가 최대 55% 빨라졌으며, 90% 이상이 직무 만족도가 향상되었다고 응답했다(GitHub, 2024). AI 도구는 반복적인 코드 작성, 버그 탐지, 테스트 케이스 생성과 같은 정형화된 과업을 자동화함으로써, 개발자들이 보다 창의적이고 복잡한 문제 해결, 시스템 아키텍처 설계, 비즈니스 로직 구현 등 고차원적인 과업에 집중할 수 있도록 도울 수 있기 때문이다(Microsoft, 2025).

소프트웨어 개발자 직군은 AI 도입의 증강 효과 이면에 있는 ‘역량의 고도화’라

22) Business Insider(2025. 5. 20), “AI is coming for big 4 too.”

는 도전 과제에 직면해 있다. AI가 단순 코딩을 대체함에 따라 단순히 코드를 작성하는 능력의 가치는 점차 하락하고, 대신 AI가 생성한 코드를 비판적으로 검토하고 수정할 수 있는 능력, 복잡한 시스템 전체를 조망하고 설계하는 아키텍처 역량, 그리고 AI 도구를 효과적으로 활용하여 최적의 결과를 이끌어내는 능력(예: 프롬프트 엔지니어링) 등 고차원적인 역량의 중요성이 그 어느 때보다 커지고 있다(신승윤 외, 2025). 이는 SW 개발자 직군 내에서도 AI를 효과적으로 활용하는 개발자와 그렇지 못한 개발자 간의 생산성 및 소득 격차가 벌어질 수 있음을 의미한다.

실제로 빅테크 기업은 스스로 자사가 생산하는 AI의 0번째 소비자가 될 수 있도록 조직을 변모하기 위해 엔지니어 직군의 대량 해고를 단행하고 있다. 마이크로소프트는 2025년 시장의 평가보다도 큰 순이익을 기록했음에도 여러 차례 대규모 정리해고를 시행하였으며, 2025년 5월 발표된 6,000여 명의 구조조정은 현재 소프트웨어 엔지니어링과 제품 관리 직군에 집중된 것으로 보인다.²³⁾

이 외에도 아마존, 메타, 세일즈포스 등 주요 테크기업의 기술 인력 감원 소식도 계속해서 확인할 수 있다. 이는 단순 자동화로 인해 불필요한 인력을 해고하여 전체 규모를 줄이는 차원이 아니라, 회사의 비즈니스 모델을 AI로 재편하는 과정에서 기존 인력이 수행하던 업무 자체가 사라졌기 때문으로 분석되기도 한다. 연초부터 1,000여 명의 직원을 감원할 것으로 밝힌 세일즈포스는 2025년 추가적으로 엔지니어를 충원할 계획이 없다고 하였으나, AI 상품 영업 인력은 추가로 채용하고 있다.

아울러 AI 기술 경쟁이 첨예해지며 소수의 최상위 AI 전문가와 개발자에 대한 수요가 폭발적으로 증가하였다. STEM 전문가인 엔지니어의 고용 불안정과 달리, AI 분야에서 창의적 사고를 하는 뛰어난 인재는 가장 희소하고 전략적인 자원으로 인식되는 ‘인재 인플레이션 현상’도 함께 관찰된다. 이로 인해 이들의 임금은

23) The Information(2025. 5. 17), “Behind Microsoft Layoffs, Automation Efforts Boom.”

천정부지로 치솟는 반면, 다른 분야의 인재 수요는 정체되거나 감소하여 같은 직종임에도 극심한 소득 격차가 존재한다. AI 분야를 주도하는 소수의 인재 확보를 위해 애플, 메타 등은 명문대 석·박사급 인재에 5억 원 이상의 초고연봉을 제시하고 있으며, 국내 기업인 KT도 AI 고급 인재를 확보하기 위해 개발자 급여 상한을 전격 폐지하며 20대 개발자도 억대 연봉으로 입사할 수 있는 길을 열었다. 이는 AI 개발 역량을 갖춘 소수에게 부가 집중되는 ‘승자독식’ 시장의 단면을 보여준다(KED Global, 2023).

이러한 인재 인플레이션과 인재 유출 현상은 국가 간 기술 격차로 심화된다. 스탠포드 대학의 ‘AI Index Report 2025’에 따르면, 2024년 AI 관련 민간 투자는 미국이 1,091억 달러로 압도적인 1위를 차지했으며, 이는 2위인 중국의 12배에 달하는 규모다. 최첨단 AI 모델 개발 역시 미국 기반 기관들이 주도하고 있어 기술 인재들을 유인하는 데 유리한 위치에 있다. 이러한 기술 패권의 물적·인적 자원 집중은 선진국과 개발도상국 간의 경제적 격차를 심화할 수 있다. IMF는 국가별 AI 노출도 수준을 정량화하여 비교한 결과, 디지털 인프라, 인적자본 등에서 앞선 선진국들은 AI 수용에 더 민감한 경제와 노동 구조를 가지고 있어 AI 시대를 먼저 경험하며 혜택을 누리게 되나, 그렇지 못한 국가들은 더욱 뒤처질 가능성을 제시하고 있다(IMF, 2024).

마. 업무 기반 접근법의 유효성

그간 경제학에서 기술 변화가 노동시장에 미치는 영향을 분석하는 이론적인 틀은 ‘숙련 편향적 기술 변화(Skill-Biased Technological Change, SBTC)’ 모델이었다. SBTC 모형은 1980년대 이후 컴퓨터와 같은 정보통신기술 확산이 고숙련 노동자의 생산성과 임금을 저숙련 노동자에 비해 상대적으로 더 크게 향상시켜, 대졸자와 비대졸자 간의 임금 격차(college wage premium)가 확대되는 현상의 채널을 성공적으로 설명하였다. 그러나 2000년대 이후 확인된 ‘일자리 공동화(job polarization)’, 즉 중숙련 직업군의 임금 정체 및 고용 감소가 나타나 현상을 설명하지 못하면서 분석 틀로서 한계를 드러냈다. SBTC 모델은 기술을 단

순히 숙련 노동을 보완하는 도구로 간주함으로써, 노동자가 수행하는 구체적인 ‘업무(task)’를 기술이 직접 대체하거나 보완하며 발생하는 메커니즘을 포착하지 못했다(Acemoglu and Autor, 2011; Acemoglu and Restrepo, 2019).

이러한 한계를 극복하기 위해 생산 활동을 제공하는 것이 노동자 유형이 아니라 일자리가 다양한 ‘업무’의 조합이라는, 업무의 분절성이라는 개념에 기반한 ‘업무 기반 모델(Task-Based Model)’이 제시되었다. 업무 기반 접근법에 따르면 기술이 노동 수요에 미치는 영향은 크게 4개의 채널을 통해 전이된다. 우선, 대체 효과(Displacement Effect)란 자동화 기술이 기존에 인간 노동자가 수행하던 업무를 자본으로 대체하는 직접적인 메커니즘이다. 이 채널은 그 자체로 노동 수요를 감소시킨다. 다음으로 ‘생산성 효과(Productivity Effect)’는 자동화로 인한 비용 절감과 효율성 증가가 경제 전반의 생산 규모를 확대시켜, 아직 자동화되지 않은 다른 업무에 대한 새로운 노동 수요를 간접적으로 증가시킨다. 세 번째 ‘복원 효과(Reinstatement Effect)’는 기술 발전보다 인간 노동에 비교우위가 있는 완전히 새로운 업무를 창출하여 노동 수요를 다시 회복시키는 상쇄의 메커니즘이다. 마지막으로 이미 자동화된 과업이 전체 일자리를 대체하여 일자리를 소거하는 ‘대체 효과(Substitution Effect)’이다(Acemoglu & Restrepo, 2018; 2019).

Autor et al.(2000)은 숙련편향 기술 발전 모델이 설명하지 못한 ‘중숙련 공동화 현상’을 정형 업무 기술발전(routine biased technological change, RBTC)으로 설명한다. Acemoglu & Restrepo(2019)도 지난 수십 년간 미국 노동시장의 불평등 심화로 ‘대체 효과’는 가속화된 반면, 새로운 일자리를 창출하는 ‘복원 효과’는 상대적으로 약화되어 새로운 업무를 창출하는 효과가 최근으로 올수록 상쇄되는 것을 실증적으로 확인한 바 있다.

업무 기반 접근법은 기술이 노동시장에 미치는 영향을 분석하는 데 가장 많이 활용되는 기본적인 프레임워크로 기능하고 있으며, 생성형 AI의 영향을 분석하는 데도 여전히 활용된다. 생성형 AI가 글쓰기, 코딩, 분석 등 과거에는 자동화가 불가능하다고 여겨졌던 비정형적이고 인지적인 업무를 수행할 수 있게 되면서, 업무 기반 접근법은 AI로부터 영향을 받는 업무를 정형·반복적 업무에서 비정형·인지

업무로 확대하여 분석하게 되었다(Hampole, 2025; Henseke et al., 2025; Pizzinelli et al., 2023; 문아람 외, 2024).

그러나 현재 AI를 업무에 활용하는 방식은 기존 업무 기반 모델의 개념적 확장을 요구한다. 첫째, AI는 업무를 ‘대체’하거나 새로운 업무를 ‘생산’하는 것을 넘어 업무 그 자체를 증강하고 있으나 이를 반영하지 못하는 한계가 존재한다. Autor et al.(2024)의 연구는 채용공고 분석을 통해 AI가 자동화 가능성이 높은 직무의 수요는 감소시키지만, 인간과 AI의 협업을 통해 시너지를 낼 수 있는 ‘보강’ 가능성이 높은 직무의 수요는 오히려 증가시킨다는 점을 실증적으로 보여주었다. 이는 노동시장에 미치는 영향을 분석하기 위해서는 대체와 복원이라는 이분법을 넘어, ‘보강’이라는 제3의 효과를 고려해야 할 필요성을 제기한다.

둘째, 업무 기반 접근법은 AI가 업무 수행에 요구되는 ‘전문성’의 가치를 변화시킬 수 있는 잠재력을 반영하기도 어렵다. AI는 노동을 대체하기도 하지만 소수 엘리트 전문가에게 집중되었던 고도의 판단 및 의사결정 업무를 보조함으로써, 더 넓은 범위의 노동자들의 역량을 증진하여 ‘전문가’의 역할을 수행하도록 돕는 ‘협업 도구’가 될 수 있다(Autor, 2024; Brynjolfsson et al., 2023; Noy and Zhang, 2023). 이는 AI가 업무의 수행 주체를 바꾸는 것을 넘어, 업무 자체의 난이도와 요구되는 역량을 변화시켜 노동시장의 진입 장벽과 경력 경로에 근본적인 영향을 미칠 수 있음을 시사한다. 따라서 향후 연구는 과업 기반 모델이, AI의 ‘보강’ 효과와 ‘전문성의 재편’이라는 새로운 차원을 반영할 수 있는지를 탐색할 필요가 있다.

2. 주요 분석 내용

가. 배경 및 목적

그간 국내외 다수 연구는 업무가 AI에 노출될 가능성을 전제로 직업이나 산업이 AI에 미치는 영향을 추정해 왔으나, 이러한 분석은 노출 수준이 실제로 AI의 사용과 그 사용의 성과로 연결되지 않는 한계를 내포한다. 기술이 채택되는 데는 개인이나 조직의 준비도, 데이터, 소프트웨어와 같은 무형자산의 축적, 적절한 업

무의 재설계, 조직의 거버넌스 적응 등이 갖춰줘야 하므로 AI 노출의 ‘가능성’과 실제 작업장 내에서의 ‘실현’ 사이엔 분명한 간극이 있다. 생산성 J 곡선 논의에서 보듯, 기술 도입 초기에는 생산성이 정체될 수 있으나, 일정한 학습과 보완적 투자 이후에야 그 효과가 실현된다. 이러한 현상을 고려한다면, 직업별 혹은 산업별 AI 노출 수준만으로 노동시장 변화를 예측하는 것은 과소 혹은 과대 추정될 수 있어 유의할 필요가 있다.

이에 본 연구는 실제 업무별 AI 활용 수준을 추정 가능한 새로운 정보를 도입하여, 한국의 일자리에서 업무별 실제 AI 사용 행태를 반영한 직업 변화를 확인할 수 있는 분석을 수행한다. 이를 통해 선행 연구에서 파악하기 어려운 직업 내 AI 활용 시 병목 발생 지점을 이해하여 노동 분야의 쟁점을 발굴해 보고자 한다. 주요한 연구 질문은 다음과 같다.

- 연구질문 1. AI가 고도화됨에 따라, 한국 직업들의 잠재적 노출 수준은 어떻게 변화하는가?
- 연구질문 2. AI에 대한 잠재적 노출 수준과 실제 활용 수준 간에 유의미한 격차가 존재하는가?
- 연구질문 3. 실제 AI 활용은 자동화와 증강 중 어떠한 패턴을 중심으로 나타나는가? 두 활용 방식은 상호 배타적인가?

나. 연구 방법론

1) 선행연구

- 동적 AI 노출 지표: Dominski & Lee(2025)

앞서 살펴본 바와 같이, AI가 노동시장에 미치는 영향은 최근까지도 업무 기반 노출법으로 접근해왔다. 기존 연구는 AI가 적용될 수 있는 업무의 기술적 가능성을 특허, 전문가, 생성형 AI 등을 통해서 평가하여 직업 단위의 노출도를 추정했다(Webb, 2020; Felten et al., 2023). 특히 최근 생성형 AI의 등장 이후, 평가의 주체에 AI가 적극적으로 활용되며 연구의 목적과 연구 질문에 따라 AI 업무 노출 지표도 새롭게 추정되어 분석에 활용되고 있다(Elendou et al., 2023; 문아

람 외, 2023). 이러한 업무 기반 접근법 진화의 연장선에서 Dominski & Lee (2025)는 AI 기술 진전 수준을 반영한 동적인 노출 수준을 측정하는 방법론을 제시하였다.

Dominski & Lee(2025)는 AI 발전 수준을 5개 단계(Stage 1~Stage 5)로 구분한 동적 AI 역량 프레임워크를 제시하고, 각 단계별로 업무 노출 가능성을 생성형 AI로 평가한 지표를 도출하였다. 각 단계는 실제 기술이 등장·확산된 시계열적 변화와 그에 따른 모델 기능의 질적 전환을 반영하며, 업무가 AI에 의해 수행될 수 있는 가능성을 평가하기 위한 기준으로 활용된다. 따라서 각 단계는 기술적 성능의 수준뿐 아니라, 실제로 해당 기술이 등장한 시점과 그 기술이 수행할 수 있는 기능적 범위를 기준으로 구분된다.

우선 1단계는 생성형 AI가 등장하기 이전의 전통적 기계학습(pre-LLM) 기반 기술 단계로 정의된다. 이 시기의 AI는 규칙 기반 또는 통계 기반 자연어 처리, 단일 목적에 특화된 딥러닝 모델 등의 기술로 대표되며, 텍스트·이미지·음성 등 각 영역을 별도의 모델이 처리하던 특징을 가진다. 이 시기는 최근의 AI 발전과 비교하기 위한 기준점으로 삼아, 생성형 AI 이전에도 이미 자동화가 가능했던 업무와 이후 새롭게 가능해진 업무를 구분할 수 있도록 설정한 단계이다.

2단계는 2022년 말 ChatGPT 출시를 기점으로 도입된 초기 대형 언어모델의 시기로 구분된다. 해당 단계의 모델은 자연어 이해와 생성 능력이 초기 형태로 실현되기 시작한 시기이며, 텍스트 기반 요약, 번역, 기본적인 콘텐츠 작성 등이 가능해졌다. 이 단계는 생성형 AI가 노동시장에 가하는 초기 충격을 설명하는 기준이 된다.

3단계는 AI가 텍스트를 넘어 이미지·비디오·음성 등 다양한 모달리티를 통합적으로 처리할 수 있는 멀티모달 LLM 단계로 진입한다. 2023년 말 DALL-E3 API 통합 등을 계기로 모델이 시각 정보를 이해하고 생성하는 기능이 확장되었으며, 시각적 판단·디자인·문서 레이아웃 생성과 같이 텍스트 기반 모델로는 어려웠던 업무들이 노출 범위에 포함되기 시작한다. 이 단계는 기술적 진전이 업무의 범위

를 어떻게 확장시키는지 파악하는 데 중요한 전환점으로 작용한다.

4단계는 고급 추론(reasoning) 모델이 등장하는 시기로, 2024년 말 출시된 OpenAI o1과 같은 모델들이 대표적이다. 해당 기술 단계에서는 복잡한 문제 해결, 논리적 추론, 계획 수립과 같은 고차원적 인지 기능이 구현되기 시작하며, 단순 텍스트 생성 이상의 전문적·분석적 업무를 수행할 잠재력이 크게 확대된다.

마지막으로 5단계는 자율적 워크플로우 실행 능력을 갖춘 에이전트 AI(agentic AI)의 단계로 설정되었다. 이 단계의 모델은 목표 설정, 계획 수립, 외부 도구 호출, 반복적 개선을 포함한 일련의 과정을 사용자의 추가 지시 없이도 수행할 수 있는 능력을 갖추게 된다. Dominski & Lee(2025)는 이 단계를 기술적 미래 정점에 대한 가정적 시나리오로 활용하여, 현재 시장에 완전히 구현되지 않은 기술 이더라도 향후 자율 에이전트가 인간 업무를 어디까지 대체하거나 수행할 수 있는지를 평가할 수 있도록 하였다.

〈표 3-1〉 5단계 동적 AI 노출 수준 프레임워크

단계	AI 역량 수준	시점	구분 근거	평가 목적
S1	Pre-LLM	2022. 11	규칙 기반·기계학습 자동화	초기 기준점 설정
S2	초기 LLM	2022. 12	자연어 생성	1차 생성형 AI 충격 평가
S3	멀티모달 LLM	2023. 10	이미지·비디오·텍스트 통합	비언어·시각 정보 기반 업무 평가
S4	추론 모델	2024. 12	문제 해결·추론·계획	전문가·사무직 고도화 충격 평가
S5	에이전트 AI	미정	자율적 작업 실행	장기적 자율화 시나리오 노출도 평가

자료: Dominski & Lee(2025) 기반 저자 재정리

• Anthropic Claude 대화 데이터: Handa et al.(2025)

잠재적 AI 노출 수준 측정이 발전되어온 만큼, 다른 한편에서는 실제 AI 활용 사례가 누적되며 업무 수행 과정에서 AI의 역할을 관찰할 수 있는 데이터도 공개되었다. 생성형 AI 서비스인 Claude를 제공하는 Anthropic은 2025년 Anthropic Economic Index(AEI)를 발표하며, 자사의 생성형 AI 모델이 실제 업무 수행 과정에서 어떤 역할을 하는지 실증적으로 관찰할 수 있는 최초의 데이터셋을 온라인에 공개하였다. AEI는 대규모 언어모델 Claude가 실제 사용자로부터 요청받은 업무 데이터를 기반으로 산출된 지수로, AI 활용의 실제적 강도와 사용 형태를 정량적으로 계측할 수 있다는 점에서 기존 연구 대비 중요한 진전을 제공한다(Handa et al., 2025).

AEI는 2024년 11월부터 2025년 8월까지 약 100만 건 이상 축적된 Claude의 실사용 로그를 바탕으로 구축하였다. 이 데이터는 단순한 질의·응답 기록을 넘어, 모델이 업무 수행 과정에서 어떤 종류의 도움을 제공했는지, 그 도움의 정도가 인간 업무의 대체 또는 보조 중 어디에 가까운지를 체계적으로 분석할 수 있도록 정제되어 있다. Anthropic은 사용자의 요청을 O*NET의 표준 업무 분류에 연계한 뒤, 모델이 수행한 기여의 유형을 자동화(automation)와 증강(augmentation)으로 구분하여 각각의 비중을 계량화하였다. 자동화 방식은 AI가 최소한의 인간 개입으로 과업을 직접 실행하는 패턴을 의미하며 여기에는 지시적 활용, 피드백 루프 활용이 해당한다. 증강 방식은 AI가 인간 역량을 강화하기 위해 협업하는 패턴으로 과업 반복, 학습, 검증이 해당한다.

〈표 3-2〉 실제 AI 활용 방식의 유형 구분

유형	세분화 유형	설명 및 예시
자동화	지시적 (Directive)	사용자가 AI에게 완전한 과업 위임을 지시(예: 이 기술 문서를 마크다운 형식으로 변환해 줘)
	피드백 루프 (Feedback Loop)	사용자가 주로(코드 실행 오류와 같은) 환경적 피드백을 전달하며 과업 완성을 유도(예: Python 스크립트가 IndexError를 반환했어. 수정해 줘... 이제 다른 오류가 발생하네)
증강	과업 반복 (Task Iteration)	사용자와 AI가 협업적 개선 과정 존재(예: 마케팅 전략 초안을 만들자... 좋은 시작이야. 여기에 구체적인 측정 항목을 추가해 줄래?)
	학습 (Learning)	사용자가 직접적인 과업 완수가 아닌 지식 습득 및 이해를 추구(예: 신경망이 어떻게 작동하는지 설명해 줘)
	검증 (Validation)	사용자가 자신의 작업을 확인하거나 개선하기 위해 AI를 활용(예: 중복 고객 레코드를 찾는 SQL 쿼리를 작성했는데, 내 로직이 맞는지 확인하고 개선안을 제안해 줘)

자료: Handa et al.(2025) 저자 재정리

각 업무에 대한 자동화·증강 비중은 모델의 응답 구조, 대화의 반복 양상, 산출물의 복잡성, 사용자 의존도 등을 기반으로 측정된다. 이 측정 과정은 정성적 판단 요소가 아닌, Anthropic이 사전에 정의한 시스템적 기준에 기반하여 일관적으로 적용되며, 그 결과 실제 AI 사용이 업무 내에서 차지하는 실질적 비중을 수치로 확인할 수 있다(Handa et al., 2025).

AEI의 도입은 기존 연구가 가진 중요한 한계를 보완하는 역할을 한다. 그간 대부분의 AI 노동시장 분석은 ‘특정 기술이 해당 업무를 수행할 수 있는가’라는 가능성(exposure)만을 고려해 왔으나, AEI는 ‘그 기술이 실제로 현장에서 사용되고 있는가’라는 실현(realized usage)의 차원을 명확히 구분할 수 있도록 한다. 이를 통해 AI의 잠재적 충격과 실제 노동시장 파급효과 사이의 괴리를 정량적으로 확인할 수 있게 되었으며, 기술 도입의 병목이 어디에서 발생하는지 파악할 수 있는 근거를 제공한다는 점에서 정책적·학술적으로 중요한 의미를 갖는다.

다만 AEI가 가진 특성상 몇 가지 한계도 존재한다. 첫째, Claude 모델을 기반

으로 구축되었기 때문에 특정 모델 생태계에서 나타난 사용 패턴이 전체 생성형 AI의 사용 패턴을 완전히 대표한다고 보기 어렵다. 둘째, 데이터가 글로벌 사용자 기반에서 수집된 만큼, 특정 국가 또는 산업의 맥락을 그대로 반영한다고 해석하기에는 제한이 있다. 그러나 이러한 제약에도 불구하고 AEI는 현재까지 공개된 자료 중 업무 수준에서 실제 AI 사용 강도를 직접 관찰할 수 있는 유일한 대규모 데이터로서, 노동시장 분석의 정밀도를 획기적으로 높일 수 있는 중요한 기초 자료로 평가된다.

• IMF의 구조적 보완수준: Pizzinelli et al.(2023)

IMF는 AI가 글로벌 노동시장에 미치는 영향을 분석하는 데 AI 노출 수준뿐만 아니라 보완성이라는 차원도 고려하였다. AI 업무 노출 수준(AI Exposure rate, AIOE)은 AI 응용 프로그램과 인간의 ‘능력’ 간의 중복 정도를 ‘노출’로 정의한 Felten et al.(2023)의 연구를 차용하였다. 다만, 노출 수준만으로는 AI가 노동을 대체할지 보완할지 알 수 없으므로, Pizzinelli et al.(2023)은 ‘잠재적 보완성 지수’를 구축하였다. 이 보완성 지수는 O*NET의 직업별 ‘작업 맥락(work contexts)’과 ‘숙련(skills)’ 데이터를 활용하는 것으로 특정 직무는 AI에 높게 노출되더라도, 물리적·사회적 맥락 때문에 인간 개입 없는 AI의 사용이 어렵다는 점을 반영하였다.

2) 연구 데이터

본 연구는 AI의 노동시장 영향을 다각도로 분석하기 위해, Dominski & Lee(2025)의 ‘잠재적 AI 노출 수준’, Anthropic의 ‘AI 실제 활용 수준’, 그리고 IMF의 ‘구조적 보완성’을 측정하는 세 가지 핵심 지표를 결합하여 한국 직업 분류에 맞게 재구성한 데이터를 활용한다. 세 가지 핵심 지표 모두 미국 노동부의 O*NET 데이터베이스에 등재된 개별 과업(Task)을 기준으로 측정되었으므로 이를 한국 노동시장에 적용하기 위해 미국과 한국의 직업분류체계 간 연계 작업을 선행하였다.

첫 단계로 O*NET 분석에 사용된 SOC19 코드를 미국 노동부가 공식 제공하는 SOC19-SOC18 대응표로 변환하고, ILO가 제공하는 공식 SOC-ISCO 대응표를

활용하여 직업을 국제표준(4-digit) 체계인 ISCO-08로 매핑하였다. 이후 통계청의 ISCO-한국표준직업분류 매핑표를 활용하여 한국표준직업분류(KSCO) 4자리 수준으로 변환하고, 마지막으로 한국고용정보원이 제공하는 한국고용직업분류와의 연계표로 변환하였다. 그 결과, 한국고용직업분류 소분류 수준에서 직업별 AI 노출 수준, AI 실제 활용 수준, AI 구조적 보완성 지표 데이터로 구축하여 분석에 활용하였다. 매칭된 최종 직업 수는 481개이다.

다. 분석 결과

본 연구는 Dominski & Lee(2025)의 단계별 AI 노출지표(AIOE), Anthropic의 실제 AI 사용지표(AEI), IMF의 구조적 보완성 지표(θ)를 한국 직업체계에 결합하여, 한국 노동시장의 직업별 AI 노출·활용·보완성 구조를 비교 분석하였다. 분석은 크게 (1) 잠재적 노출 수준 평가, (2) 잠재적 노출 수준과 실제 활용 수준 비교, (3) 직업별 자동화와 보완성 수준 비교 분석의 세 측면에서 이루어졌다.

1) AI 발전 단계별 잠재적 노출 수준 추이 변화

Dominski & Lee(2025)의 동적 지표를 활용한 한국의 직업별 노출 수준 추이를 살펴보면, AI가 진전될수록 한국 직업 대부분이 단계적으로 더 높은 수준의 노출을 경험하는 패턴을 보였다. LLM 직전인 Stage 1 평균은 30.7%, 에이전트 AI인 Stage 5의 평균은 84.2%로 약 54%p 증가하는 것으로 나타났다.

〈표 3-3〉 잠재적 AI 노출수준(AIOE) 기술통계: 5단계별

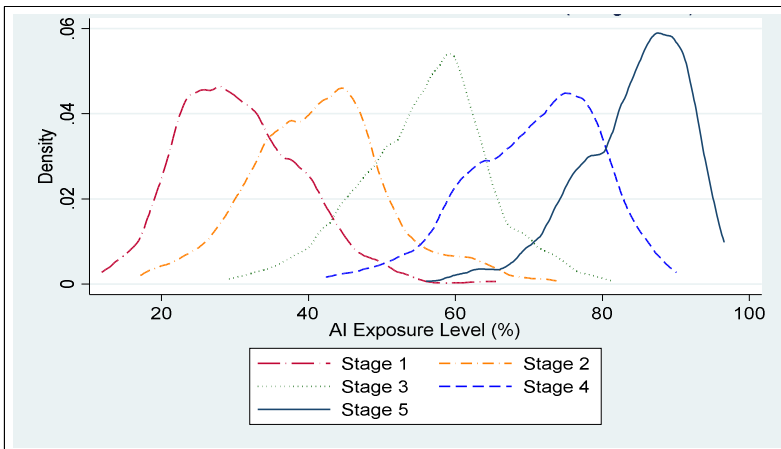
단계	평균(%)	표준편차	최소	최대	n
AIOE_s1	30.68	8.31	11.86	65.54	481
AIOE_s2	41.51	9.60	17.14	73.75	481
AIOE_s3	55.88	8.73	29.17	81.15	481
AIOE_s4	70.64	9.08	42.35	90.08	481
AIOE_s5	84.24	7.40	56.09	96.54	481

자료: 연구진 작성

한국고용직업분류 2025년 분류(KECO25) 수준에서 직업별 AIOE(%)가 5단계에 걸쳐 어떻게 변하는지 살펴보기 위해, 각 단계에 걸쳐 직업별 노출 수준의 분포가 어떻게 변화하는지 커널 밀도 추정으로 시각화하였다. [그림 3-5]에서 X축은 AI 노출 수준(0~100%)으로, 오른쪽으로 갈수록 해당 직업이 AI에 더 많이 노출됨을 의미하며, Y축은 밀도(density)를 나타내는 것으로 그래프의 높이가 높을수록 해당 X축 값(%) 근처에 더 많은 직업이 밀집해 있음을 의미한다. 단 Y축 값 자체가 직접적인 확률이나 직업 수를 의미하진 않음에 유의할 필요가 있다.

그 결과, 기술 발전과 함께 노출 수준의 평균이 상승하고 분포는 수렴하고 있으며, 초기에는 노출의 직업 간 편차가 컸지만, 장기적으로는 모든 직업이 고노출을 경험하게 되는 보편화 경향이 나타날 수 있음을 확인할 수 있었다. 우선 AI 기술이 발전함에 따라(S1 → S5), 한국 직업 시장 전체의 평균적인 AI 노출 수준이 일관되게 증가하고 있다. 5개 곡선(S1~S5)의 중심 위치를 살펴보면 생성형 AI가 상용화된 초기(1단계)에서 최근(5단계)으로 올수록 분포 전체는 명확하게 오른쪽으로 이동하고 있음을 볼 수 있다.

[그림 3-5] 직업별 AI 노출 수준 분포 추이: 5단계



자료: 연구진 작성

아울러 직업들이 가장 많이 몰려있는 노출 수준을 나타내는 분포 곡선별 봉우리에 주목한다면, 1단계에서는 봉우리가 20~30%대 근처에 폭넓게 형성되어 있고, 마지막 단계인 5단계에서는 80~90%대 근처에 매우 높고 뾰족하게 형성됨을 볼 수 있다. 이는 대부분의 직업이 초기(1단계)에는 20~30% 수준의 낮은 노출도를 가지다가, 이후 AI가 성숙하며 대부분의 직업이 높은 노출 수준에 밀집됨을 확인할 수 있는 결과이다. 아울러 곡선의 너비와 높이도 시간이 지날수록 좁고 높아지고 있어, 초기(1단계)에는 직업별 편차가 컸지만, 미래(5단계)에는 AI 노출도가 ‘일부 직업의 문제’가 아닌 ‘모든 직업의 보편적 현상’이 되는 것으로 해석할 수 있다.

다음으로 단계별 상위 10개 직종과 하위 10개 직종을 비교한 결과, 상위권에는 회계·경리·속기·증권·경영 기획 사무원이 지속적으로 포함되었으며, 하위권에는 미장공·석공·도배공 등 기능·육체노동 중심 직업군이 고착적으로 나타났다.

〈표 3-4〉 잠재적 AI 노출 수준 단계별 상위·하위 10개 직업

단계	상위 10개 직종			하위 10개 직종		
	직종 코드	직종명	AIOE	직종 코드	직종명	AIOE
S1	271	회계 사무원	65.54	4147	무용가 및 안무가	11.86
	272	경리 사무원	63.62	7022	방수공	11.99
	296	속기사	56.56	7021	미장공	12.50
	233	관세사	52.41	7015	건축 석공	12.73
	322	증권 사무원	52.24	4144	국악 및 전통 예능인	13.93
	284	자재관리 사무원	51.17	7014	콘크리트공	14.16
	262	영업 및 마케팅 사무원	51.16	5123	장례 지도사 및 장례 상담원	14.38
	261	경영 기획 사무원	51.16	7024	바닥재 시공원	15.16
	324	출납 창구 사무원	50.02	2212	변호사	15.79
	295	전산 자료 입력원	49.76	4145	지휘자·작곡가 및 연주가	16.00

단계	상위 10개 직종			하위 10개 직종		
	직종 코드	직종명	AIOE	직종 코드	직종명	AIOE
S2	271	회계 사무원	73.75	7022	방수공	17.14
	272	경리 사무원	72.68	7021	미장공	17.23
	296	속기사	69.60	7015	건축 석공	18.51
	261	경영 기획 사무원	68.88	7014	콘크리트공	18.69
	262	영업 및 마케팅 사무원	68.88	7024	바닥재 시공원	18.98
	322	증권 사무원	65.72	7017	조적공 및 석재 부설원	19.64
	6154	온라인 쇼핑 판매원	65.20	7011	강구조물 가공원 및 건립원	20.41
	6140	텔레마케터	65.20	7052	철로 설치 및 보수원	21.13
	264	교육 및 훈련 사무원	64.57	7025	도배공 및 유리 부착원	21.87
	292	고객 상담 및 모니터 요원	63.90	7051	광원·채석원 및 석재절단원	21.97
S3	271	회계 사무원	81.15	7022	방수공	29.17
	272	경리 사무원	79.99	7021	미장공	30.39
	296	속기사	77.53	7015	건축 석공	31.58
	261	경영 기획 사무원	76.71	7014	콘크리트공	31.81
	262	영업 및 마케팅 사무원	76.71	7017	조적공 및 석재 부설원	33.04
	6140	텔레마케터	75.97	5323	주방 보조원	34.23
	6154	온라인 쇼핑 판매원	75.97	7024	바닥재 시공원	34.28
	6155	상품 대여원	74.72	7011	강구조물 가공원 및 건립원	35.60
	322	증권 사무원	74.53	5123	장례 지도사 및 장례 상담원	35.65
	264	교육 및 훈련 사무원	73.65	7052	철로 설치 및 보수원	36.02
S4	271	회계 사무원	90.08	7022	방수공	42.35
	272	경리 사무원	89.62	7021	미장공	43.03
	261	경영 기획 사무원	88.60	5323	주방 보조원	43.83
	262	영업 및 마케팅 사무원	88.60	7014	콘크리트공	44.59
	6140	텔레마케터	87.26	7017	조적공 및 석재 부설원	44.99
	6154	온라인 쇼핑 판매원	87.26	7015	건축 석공	45.21
	264	교육 및 훈련 사무원	87.07	7024	바닥재 시공원	46.48
	6155	상품 대여원	86.46	7025	도배공 및 유리 부착원	47.95

단계	상위 10개 직종			하위 10개 직종		
	직종 코드	직종명	AIOE	직종 코드	직종명	AIOE
S4	322	증권 사무원	86.45	5632	세탁원 및 다림질원	48.09
	285	생산 및 품질관리 사무원	86.31	5123	장례 지도사 및 장례 상담원	48.89
S5	296	속기사	96.54	7022	방수공	56.09
	262	영업 및 마케팅 사무원	96.26	7021	미장공	56.56
	261	경영 기획 사무원	96.26	7017	조적공 및 석재 부설원	61.06
	272	경리 사무원	96.20	7024	바닥재 시공원	61.57
	271	회계 사무원	96.13	7015	건축 석공	61.80
	264	교육 및 훈련 사무원	95.59	5123	장례 지도사 및 장례 상담원	62.20
	6140	텔레마케터	95.13	5323	주방 보조원	62.32
	6154	온라인 쇼핑 판매원	95.13	7014	콘크리트공	62.54
	285	생산 및 품질관리 사무원	95.00	5632	세탁원 및 다림질원	63.52
	282	도로 및 철도운송 사무원	94.89	7025	도배공 및 유리 부착원	63.64

자료: 연구진 작성

초기 단계와 마지막 단계 상위 직종의 AI 노출 수준 증가폭을 비교한 결과, ‘회계 사무원’(271), ‘경리 사무원’(272) 등은 S1부터 이미 60% 이상의 매우 높은 노출도를 보였으며, S5에서도 여전히 최상위 노출 직업군에 포함되어 해당 직종은 AI 기술 발전 수준에 상관없이 고노출 직업군으로 고정된 유형으로 분류할 수 있다. 반면에, S5 노출 상위 10개 직종의 S1 순위를 살펴본 결과, 사무직·판매직에 해당하는 ‘텔레마케터’(6140), ‘온라인 쇼핑 판매원’(6154) 등은 초기 단계에서 51위권이었으나, S5에서는 7위권으로 순위가 44계단 급상승하였다.

〈표 3-5〉 초기/최종 단계 AIOE 노출 수준 상위 10개 직종의 순위 변화

단계	직종 코드	직종명	S1	S5	증가폭 (S5-S1)	S1 순위	S5 순위
초기 단계	271	회계 사무원	65.54	96.13	30.59	1	5
	272	경리 사무원	63.62	96.2	32.58	2	4
	296	속기사	56.56	96.54	39.98	3	1
	233	관세사	52.41	93.66	41.26	4	25
	322	증권 사무원	52.24	94.84	42.6	5	13
	284	자재관리 사무원	51.17	91.21	40.05	6	79
	262	및 마케팅 사무원	51.16	96.26	45.1	7	2.5
	261	경영 기획 사무원	51.16	96.26	45.1	7	2.5
	324	출납 창구 사무원	50.02	90.56	40.54	9	98
	295	전산 자료 입력원	49.76	90.94	41.18	10	87
최종 단계	296	속기사	56.56	96.54	39.98	3	1
	261	경영 기획 사무원	51.16	96.26	45.1	7.5	2.5
	262	영업 및 마케팅 사무원	51.16	96.26	45.1	7.5	2.5
	272	경리 사무원	63.62	96.2	32.58	2	4
	271	회계 사무원	65.54	96.13	30.59	1	5
	264	교육 및 훈련 사무원	46.08	95.59	49.51	24	6
	6140	텔레마케터	41.54	95.13	53.59	51.5	7.5
	6154	온라인 쇼핑 판매원	41.54	95.13	53.59	51.5	7.5
	285	생산 및 품질관리 사무원	49.27	95	45.73	13	9
	282	도로 및 철도운송 사무원	47.16	94.89	47.73	17	11

자료: 연구진 작성

〈표 3-6〉 AIOE 수준 차이(S1 → S5) 상위 10개 직종

직종 코드	직종명	S1	S5	차이 (S5-S1)	S1 순위	S5 순위	순위 변동
4163	아나운서 및 리포터	19.92	89.41	69.5	450	137	313
4111	출판물 전문가	21.73	90.96	69.23	423.5	85.5	338
4113	작가	21.73	90.96	69.23	423.5	85.5	338
2151	장학관·연구관 및 교육 관련 전문가	19.41	87.73	68.33	453	183	270

직종 코드	직종명	S1	S5	차이 (S5-S1)	S1 순위	S5 순위	순위 변동
4153	실내장식 디자이너	20.06	87.7	67.64	449	185	264
4146	가수 및 성악가	16	83.51	67.52	471.5	301.5	170
4145	지휘자·작곡가 및 연주가	16	83.51	67.52	471.5	301.5	170
2212	변호사	15.79	83.15	67.36	473	310	163
1333	웹 개발자	24.48	91.35	66.88	363	75	288
241	광고 및 홍보 전문가	24.59	91.37	66.78	361	74	287

자료: 연구진 작성

마지막으로 S1에서 S5로 갈수록 가장 큰 증가폭을 가진 직업들은 주로 콘텐츠 생산, 창작 등 지식노동 가운데서도 창의성이 요구되는 직종이었다. 이들은 초기 노출이 매우 낮았음에도, S5에서는 80~90%대까지 급상승하는 것으로 나타났다. 해당 직종들은 생성형 AI·멀티모달 AI·추론형 모델의 기능 확장에 직접적으로 영향을 받는 것으로 확인된다.

2) 잠재적 AI 노출 수준과 실제 활용 수준 간의 비교

다음으로는 앞서 분석한 AI에 대한 가설적 노출 지표인 단계별 AIOE가 실제 업무에서의 AI 활용(AEI) 행태와 일치하는가를 확인한다. 우선 분석 대상이 되는 481개 직종 중 실제로 업무에 AI를 활용하는 것으로 관찰된 직종은 367개에 해당하여, Anthropic의 Claude 대화 데이터에서 관찰되지 못한 나머지 114개의 직업을 위한 업무에는 활용되지 않았다. Anthropic의 Claude 대화 데이터에서 관찰된 직업(AEI 관찰)과 Anthropic의 Claude 대화 데이터에서 관찰되지 않은 직업(AEI 미관찰) 집단 간 잠재적 노출 수준과 잠재적 보완 수준의 차이를 확인한 결과, 두 집단 간 보완 수준과 노출 수준에 대부분 유의미한 차이가 있었다. 보완 수준의 경우, 관찰된 집단의 평균이 관찰되지 않은 집단보다 2.9점 높게 형성되었다. 단계별 노출 수준도 마찬가지로 기술이 발전할수록 실제로 관찰된 집단이 관찰되지 않은 집단보다 유의미하게 큰 값을 가졌다.

〈표 3-7〉 집단별 AI 노출 지표 및 θ 수준 비교

변수	전체 그룹		AEI 미관찰(A)		AEI 관찰(B)		(B-A)		
	Mean	S.D	Mean	S.D	Mean	S.D	mean	s.e	t stat
θ	59.9	6.8	57.7	6.7	60.6	6.7	2.9	0.72	4.02 (0.00)
AIOE_s1	30.7	8.4	29.9	7.7	30.9	8.6	1.1	0.90	1.17 (0.24)
AIOE_s2	41.5	9.6	36.5	8.5	43.1	9.4	6.5	0.99	6.64 (0.00)
AIOE_s3	55.9	8.7	51.2	8.3	57.3	8.4	6.2	0.89	6.90 (0.00)
AIOE_s4	70.6	9.1	65.4	9.0	72.3	8.5	6.9	0.92	7.47 (0.00)
AIOE_s5	84.2	7.4	81.1	7.7	85.2	7.0	4.1	0.77	5.25 (0.00)
Obs	481		114		367				

자료: 연구진 작성

다음으로 변수 간 상관관계를 살펴보면, 잠재적 노출 수준과 실제 AI 활용은 다르며 ‘AI에 노출될 것이다(가설)’와 ‘AI로 업무를 증강한다(현실)’ 혹은 ‘AI로 업무를 자동화한다(현실)’는 서로 다른 현상임을 확인할 수 있었다. 우선, 직업의 잠재적 AI 노출 수준(AIOE_s5)은 실제 AI 자동화 활용 수준(AEI_auto)과 실제 AI 증강 활용 수준(AEI_augm) 간 상관관계가 0.32, 0.36으로 상관관계가 약하게 드러났다. 아울러 IMF가 추정한 구조적 보완성 수준(θ)과 Anthropic의 실제 관찰된 증강 수준(AEI_augm)은 0에 가까운 상관관계를 나타내 두 변수는 통계적으로 거의 독립적임을 확인할 수 있었다. 즉 한 직업의 ‘보완’ 잠재력이 높다고 해서(높은 θ), 그 직업이 실제로 AI를 증강에 많이 활용(높은 AEI_augm)한다는 보장이 없음으로 해석할 수 있다.

반면에 Anthropic의 실제 데이터 내에서는 자동화(AEI_auto)와 증강(AEI_augm) 간의 상관관계가 0.745로 나타났다. 따라서 자동화와 증강이 서로 반대되거나 대

체되는 관계(zero-sum)가 아니라 AI를 자동화에 활발히 사용하는 직업은 증강에도 활발히 사용할 수 있는 것으로 해석된다.

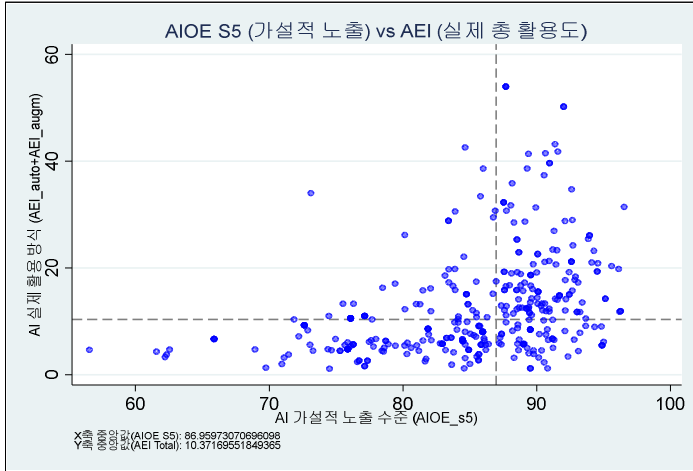
〈표 3-8〉 잠재적 AI 노출 수준과 실제 AI 활용 수준 간 상관관계

구분	θ	AIOE_s5	AEI_automation	AEI_augmentation
θ	1			
AIOE_s5	-0.0184	1		
AEI_automation	-0.0961	0.3223	1	
AEI_augmentation	0.0869	0.3594	0.7445	1

자료: 연구진 작성

마지막으로 잠재적 AI 노출 수준과 실제 AI 활용 수준을 산점도로 제시하고, 각 변수의 중앙값을 기준으로 4분면으로 구분하여 분석하였다. 1사분면(우상단)은 가설도 높고, 실제 총 활용도가 높은 집단으로 예측된 고효율군에 해당한다. 2사분면(좌상단)은 가설은 낮았으나, 실제 총 활용도가 높은 직종으로 의외의 고효율군으로 볼 수 있다. 3사분면(좌하단)은 예측된 저활용군으로 가설도 낮고, 실제 총 활용도도 낮은 집단이다. 4사분면(우하단)은 가설상 높게 노출되었으나, 아직 실제 총 활용도가 낮아, 향후 AI 도입이 폭발적으로 증가할 잠재적 그룹으로 볼 수 있다. 전체 367개 직업 중 1사분면은 129개, 2사분면은 55개, 3사분면은 128개, 4사분면은 55개 직종이 분포하고 있다. 즉 예측과 실재가 일치한 1사분면과 3사분면이 70%에 해당하였고, 나머지 30%는 불일치한 영역이다.

[그림 3-6] 잠재적 AI 노출 수준-실제 AI 활용 수준 산점도



자료: 연구진 작성

3) 직업별 자동화와 보완 수준 격차

앞서 관찰된 자동화와 보완 수준 간의 상관관계만을 살펴본다면 유사한 패턴으로 움직이고 있으나, 두 변수 간의 차이를 분석한다. 우선 기술통계를 보면, 실제 AI 활용은 ‘자동화’보다는 ‘증강’ 목적으로 1.6배 더 많이 사용($8.02 > 4.84$)되고 있었다. 또한 Anthropic 데이터에 포착된 직업 367개는 단 하나의 직업도 빠짐 없이 AI를 ‘증강’에 활용하고 있으나(최소값 0.879), 반면 AI를 ‘자동화’로 아예 활용하지 않는, 즉 최소값이 0인 직업은 20개 존재하였다.²⁴⁾

24) Handa et al.(2025)은 분석 대상인 수백만 건의 전체 AI 대화(conversations) 중 57%가 ‘증강(augmentation)’ 패턴을, 43%가 ‘자동화(automation)’ 패턴을 보였다고 보고했다. 이는 AI 총 활용량(total usage)의 과반이 증강 목적임을 의미한다. 반면, 본 보고서는 Handa et al.(2025)의 업무별 데이터를 O*NET 과업 가중치 등을 이용해 한국 ‘직업(occupation)’ 수준으로 가중 합산(aggregate)한 뒤, 각 ‘직업’ 내 AI 활용에서 자동화와 증강 수준을 살펴보았다.

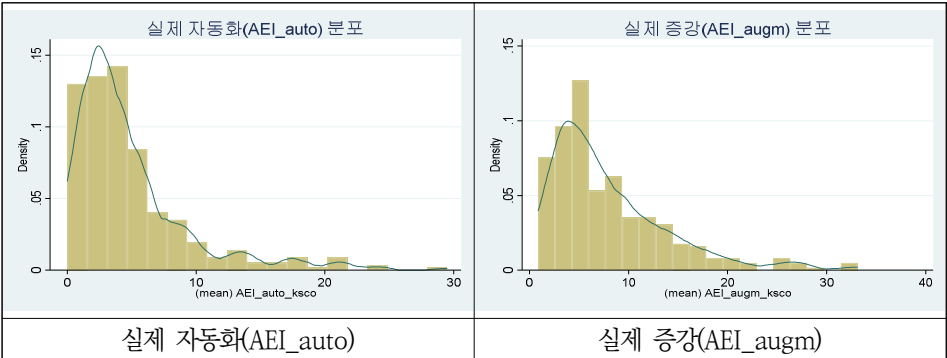
〈표 3-9〉 직업별 실제 관찰된 자동화 및 보완 수준 기술통계

변수	Obs	평균	표준편차	최소	최대
AEI_auto	367	4.838755	4.610684	0	29.48419
AEI_augm	367	8.021384	6.006217	0.878597	33.11255
AEI_total	367	12.86014	9.928148	1.153376	53.97077

자료: 연구진 작성

AI의 실제 활용 방식인 ‘자동화’(AEI_auto)와 ‘증강’(AEI_augm)의 분포가 본질적으로 다른 성격을 가지고 있음을 시각적으로 확인하기 위해 히스토그램을 제시하였다. 히스토그램의 X축은 각각 자동화와 증강, 증강 활용 비중이며, Y축은 해당 비중에 누적된 직업 수를 의미한다. 좌측 그래프인 실제 자동화 비중의 경우, 히스토그램이 상대적으로 0에 치우쳐 있다. 즉 10% 이상 의미 있는 수준의 자동화를 경험하는 직업은 상대적으로 소수이다. 우측 그래프는 실제 증강 비중으로, 자동화와 비교하면 3~10% 사이에 명확한 봉우리가 그려짐을 볼 수 있다. 이는 AI 활용이 관찰된 직업의 94.6%가 자동화를 0% 이상 경험하기는 하지만, 그 경험의 ‘강도’가 대부분 미미한 수준임을 의미한다.

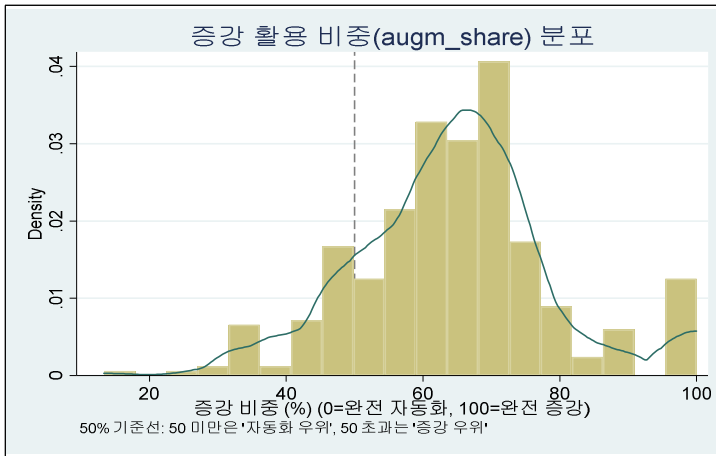
[그림 3-7] 직업별 실제 관찰된 자동화·증강 수준 히스토그램



자료: 연구진 작성

상대적인 증강 활용 비중(증강/(자동화+증강)) 히스토그램 또한 AI 활용은 ‘자동화’보다 ‘증강’이 중심임을 시사한다. 이 그래프의 X축(증강 비중)의 50% 기준선은 ‘자동화 우위’와 ‘증강 우위’를 나누는 기준으로, 히스토그램의 대부분이 50% 기준선의 오른쪽에 몰려있다. Anthropic 데이터에 포착된 대부분의 직업이 AI를 ‘자동화’보다는 상대적으로 ‘증강’의 목적을 중심으로 활용하고 있다. 또한 그래프의 가장 높은 봉우리는 X축의 60~75% 사이에 위치해 있다. AI를 활용하는 직업의 가장 보편적인 모습은 AI 총 활용 시간(혹은 양)의 약 2/3에서 3/4을 ‘증강’에 사용하고, 나머지 1/4에서 1/3은 ‘자동화’에 사용하는 것이라고 볼 수 있다.

[그림 3-8] 직업별 실제 관찰된 증강 활용 비중 히스토그램



자료: 연구진 작성

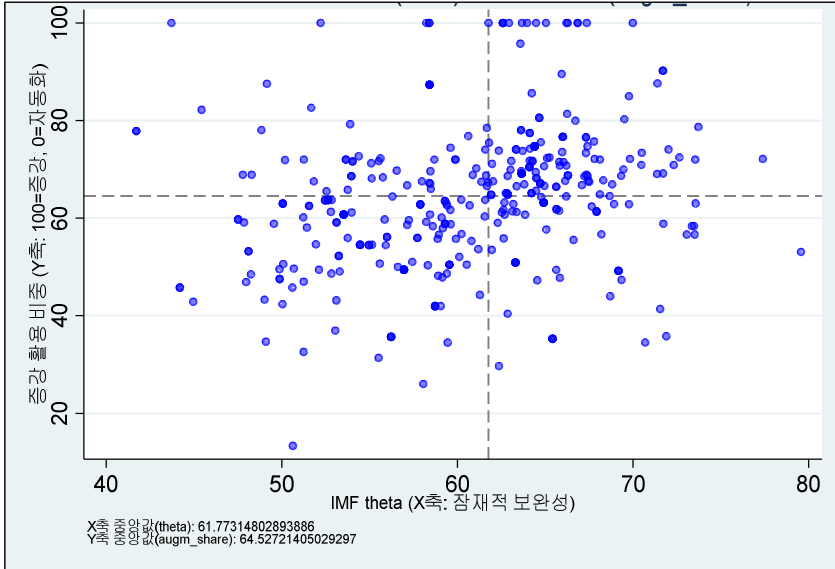
마지막으로 구조적 보완성(θ)과 실제 보완 수준(AEI_agum)의 산점도를 통해 가정적 보완성과 실제 증강 수준을 비교하였다. 그래프의 각 점은 Anthropic 데이터에서 활용이 관찰된 직업을 나타낸다. X축은 IMF(2024)가 제시한 ‘구조적 보완 잠재력’(θ) 지표를 의미하며, 이 값이 높을수록(우측) AI가 인간을 대체하기 어려운 직업이다. Y축은 Anthropic 데이터를 기반으로 본 연구에서 산출한 ‘실제 증강 활용 비중(augm_share)’이다. 이 값은 $[AEI\_augm / (AEI\_auto + AEI\_$

augm)]* 100으로 계산되며, 100%에 가까울수록(상단) 증강 우위, 0%에 가까울수록 자동화 우위의 활용 패턴을 보임을 의미한다.

두 변수의 중앙값(X축 $\theta=61.8$, Y축 augm_share=64.5)을 기준으로 사분면을 구분하였으며, 산점도에서 보듯이 두 변수 간 상관관계는 0에 가까워 통계적으로 독립적임을 알 수 있다. 이는 보완 잠재력이 높다고 해서 반드시 실제 증강 활용으로 이어지는 않음을 시사하며, 다음과 같이 4개 그룹으로 직업군을 유형화하는 정책적 근거를 제공할 수 있다.

- 1사분면(우상단): [선도적 증강] IMF가 예측한 ‘보완’ 직업군이, 실제로도 AI를 ‘증강’ 중심으로 가장 잘 활용하는 그룹
- 2사분면(좌상단): [의외의 증강] IMF는 ‘대체’ 위험이 높다고 봤으나, 실제로는 AI를 ‘증강’ 중심으로 활용하며 적응에 성공한 그룹
- 3사분면(좌하단): [자동화 우위 / 고위험] IMF의 ‘대체’ 위험 예측이 맞았고, 실제 활용도 ‘자동화’ 패턴에 치우친 그룹(정책적 관심 대상)
- 4사분면(우하단): [미래 잠재력 / 패턴 불일치] ‘보완’ 잠재력은 높으나(예: 의사, 판사), 현재의 AI 활용은 ‘자동화’ 중심인 그룹으로, 향후 ‘증강’ 중심으로 패턴이 이동할 잠재력이 가장 큰 집단

[그림 3-9] AI 보안 잠재력 및 실제 보안 활용 수준 산점도



자료: 연구진 작성

3. 식별된 사회적 의제안

가. (노동01) AI로 인한 직무 전환이 요구됨에 따라, 마모되는 근로자의 인적자본을 회복하기 위한 교육훈련 시스템은 어떻게 설계할 것인가?

1) 제안의 배경

AI 진전은 소프트웨어 도구를 넘어 에이전틱 AI, 물리 AI로 병합하므로 일부 지식 산업뿐만 아니라 제조 분야에 이르는 광범위한 산업의 일자리에 영향을 미치게 된다. 이는 생애주기 소득의 안정성을 영위하기 위해 평생직장이 아닌 평생학습을 요구한다는 점에서 기존 및 신규 근로자의 지속적인 직무 전환을 시급한 과제로 만들고 있다.

2) 현황 및 쟁점

생성형 AI 등장 이후 일자리 대체에 대한 우려로 세계경제포럼, 국제노동기구 등의 국제기구와 IMF, 한국은행 등 각국에서는 일의 미래를 양적으로 전망한 보

고서를 발표하며 중장기적으로 기존 일자리의 상당 부분이 AI로 대체되고, 일에 요구되는 직무 내용도 근본적으로 변할 것을 예측하고 있다. 현재 정부 주도의 직업 훈련은 대부분 실직자 중심이며, 재직자 훈련은 참여율이 저조한 경우가 많다. 특히 중장년층 혹은 저숙련 근로자의 경우 새로운 기술 습득에 대한 심리적·시간적 장벽이 높아 업스킬과 리스킬의 사각지대에 놓여있다.

3) 문제의 심각성

AI는 전 직종에 걸쳐 근로자가 축적한 인적자본의 마모를 가속한다. 마모된 인적자본을 회복하지 못하는 기존의 숙련 인력이 저숙련·저부가가치 노동으로 밀려 나가거나 노동시장에서 조기 이탈할 수 있다. 이는 기업과 국가 전반의 노동생산성 하락으로 연결될 수밖에 없다. 특히 숙련 인력의 하향 이동은 기술 변화에 적응하지 못한 실업이 장기화되면서 실업급여, 복지 재정 등에 대한 사회적 비용을 증대시킬 위험이 존재한다.

4) 기존 논의/대책의 한계

기존의 고용보험기금에 기반한 직업훈련 프로그램은 이미 실직한 경험을 한 이후에 작동될 수 밖에 없는 한계가 있다. AI 시대에는 실직하기 전, 즉 직무가 소멸하기 전에 선제적으로 전환 교육이 이루어져야 하나, 현재의 제도는 이러한 예방적, 선제적 재교육 시스템을 뒷받침하기에 경직되어 있다. 또한 훈련 내용이 산업계의 실제 AI 수요를 신속하게 반영하지 못하는 한계도 지적할 수 있겠다.

나. (노동02) AI가 청년층의 초기 경력 형성에 필수적인 일자리를 대체해 나가는 상황에서, ‘경력 단절’ 없는 청년 고용 시장을 어떻게 보장할 것인가?

1) 제안의 배경

2023년 직후 AI가 본격적으로 산업과 업무에 도입되고 에이전트로 진화하며 초임·보조 수준의 진입 직무(entry-level jobs)는 AI로 수행되는 자동화가 가속할 것이 전망된다. 청년들이 노동시장에 진입하며 경험을 쌓던 데이터 입력, 단순 분석, 초급 사무 보조 등의 직무가 급격히 대체되는 것은 단순히 실업의 양적 측면

을 넘어, 경력 사다리의 첫 번째 칸에 오를 기회 자체를 박탈하게 되므로 사회적 불안정을 야기할 수 있다. 이는 AI 기술이 유발한 마찰적 실업이 경기 사이클상의 단기적 문제가 아니라 세대 단위의 구조적 일자리 상실로 고착할 수 있다는 점에서 공론화와 제도적 대응의 필요성이 크다고 볼 수 있다.

2) 현황 및 쟁점

Brynjolfsson et al.(2025)²⁵⁾은 생성형 AI 도입 이후, AI 노출도가 높은 일자리를 가진 경력 진입 단계의 근로자(22~25세)는 약 13%의 고용 감소를 경험하였다고 분석한 바 있다. 한국도 마찬가지로 AI 확산에 따라 청년 고용이 크게 줄어 지난 3년간 약 20만 개의 일자리가 사라졌을 것으로 추정되었다(한진수·오삼일, 2025).²⁶⁾ 특히 2025년 2월 기준 청년층(15~29세)의 체감실업률은 16.4%로 4년 만에 가장 큰 폭으로 악화된 수치를 보였다는 점에서, 이미 한국의 청년층은 불안전 취업 상태를 경험하고 있는 것으로 보인다.

3) 문제의 심각성

초임 경력을 쌓지 못하여 고용 불가로 이어지는 악순환으로 임금 확보와 안정적 일자리가 손실되는 경로가 사회에 고착될 수 있다. 특히 청년들이 경력 초기에 적절한 훈련과 경험을 쌓지 못할 경우 국가의 인적자본이 축적되지 못하게 되며, 이는 장기적으로 노동생산성 하락으로 연결될 수 있다. 이는 장기 실업급여, 사회부조 등의 부담 증대로 귀결되어 사회적 비용 상승을 불러일으킬 수 있다.

4) 기존 논의/대책의 한계

기존의 청년을 위한 좋은 일자리를 공급하기 위한 고용 대책은 인턴십 확대, 단기 일자리 제공 등 양적 공급에 치중해 왔으나 AI로 인한 엔트리 레벨 직무의 구

25) Brynjolfsson, Chandar, & Chen(2025), *Canaries in the coal mine? six facts about the recent employment effects of artificial intelligence*, Stanford Digital Economy Lab.

26) 한진수·오삼일(2025), “AI 확산과 청년고용 위축: 연공편향 기술변화를 중심으로”, 《BOK 이슈노트》, 2025-30호.

조적 소멸이라는 근본적인 변화에 대응하기 어렵다. 일시적인 일자리 체험이 아닌, AI 시대에 필요한 핵심 역량을 체계적으로 개발하고 경력으로 연결할 장기적인 경로 설계가 부재한 상황이다.

다. (노동03) AI로 인한 노동시장의 창조적 파괴 과정에 적응하기 위해 발생하는 전환 비용은 누가, 어떤 비율로 분담하여야 하며, 배분의 우선순위는 어떤 사회적 기준으로 정할 것인가?

1) 제안의 배경

AI로 인한 노동시장의 창조적 파괴 과정에서, 직무의 파괴 혹은 소멸이 빠르게 집단적으로 발생하는 반면, 새로운 직무 창출은 비교적 느리게 분산적으로 일어난다. 일자리 감소와 증가 간 속도 및 시점의 불일치가 발생하므로 근로자들은 필연적으로 소득의 공백, 숙련의 공백, 그리고 지리적 공백에 직면하게 된다. 이러한 공백을 극복하는 과정에서 재훈련 비용, 이주 및 정착 비용 등 막대한 전환 비용과 책임을 배분하는 원칙은 부재하다.

2) 현황 및 쟁점

기술적 충격 이후 단기적 혹은 중기적 직무 전환의 실패로 노동시장에서는 창조보다 파괴의 규모가 더 크게 발생하는 것이 목격되어 왔다. 과거 미국의 제조업 분야는 중국과 수출입 경쟁에 더 많이 노출된 지역일수록 고용과 임금이 큰 폭으로 하락했으며, 신규 일자리가 창출되거나 새로운 지역으로의 이주 등을 통한 충격에서 회복하는 과정이 매우 느렸던 것으로 확인된 바 있다(Autor, 2013)²⁷⁾. 또한 산업용 로봇 확산 등으로 인한 1990년과 2007년의 자동화 과정에서 일자리 대체 효과가 신규 업무 창출이나 재배치 효과를 상회하는 것으로 나타난 바 있다(Acemoglu & Restrepo, 2020)²⁸⁾. AI도 마찬가지이다. AI에 의해 전문성이 대

27) Autor et al.(2013), "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*. 103(6), pp. 2121-2168.

체되는 근로자와 AI로 인해 전문성이 새롭게 가치 있게 되는 근로자가 서로 다른 집단 혹은 스킬을 요구하는 직무일 가능성이 높다는 것이 규명되고 있다(Autor, 2024)²⁹⁾. Autor(2024)는 AI가 전문성을 증강하는 방식으로 활용된다면 중산층이 재건할 수 있다는 점을 강조하고 있어, 전환 과정에 대한 상충되는 이해관계자 간의 집단적 선택에 따라 이는 기회가 될 수 있다.

3) 문제의 심각성

전환이 지연되면 가구소득과 지역 경제에는 구조적인 손실이 누적될 수밖에 없다. 과거 기술 혹은 무역 충격의 결과에서 확인하듯이 노동시장에서 자연스러운 배치는 일어나지 않으며, 장기적으로 임금과 고용 하락, 해당 지역의 경제활동참가율이 감소하게 된다. 또한 AI 도입으로 이익을 얻는 집단(기업, 고숙련 전문가)과 그 비용을 부담하는 집단(대체된 근로자, 쇠퇴 지역)이 불일치하기 때문에 사회적 갈등이 증폭될 가능성이 있다.

4) 기존 논의/대책의 한계

실업 상태를 지원해 온 기존의 실업급여나 고용보험은 단기적인 실업을 전제로 하므로 AI로 인한 구조적이고 장기적인 전환 과정에서 발생하는 복합적인 비용 전반을 감당하기에는 한계가 있다. 물론 적극적 노동시장 정책이 잘 갖춰진 스웨덴은 ‘전직보장협의회’와 같이 노사의 공동 재원으로 조기 전환을 지원하는 장치가 효과적으로 작동하고 있으나, 사회적 파트너십을 주도하는 제도가 없는 국가 혹은 산업 부문이 대부분이라는 점에서 결국 비용은 개인 혹은 중소기업에 편중될 수 있다.

28) Acemoglu, & Restrepo(2020), “Robots and jobs: Evidence from US labor markets,” *Journal of political economy*. 128(6), pp. 2188-2244.

29) Autor(2024). “Applying AI to rebuild middle class jobs (No. w32140),” National Bureau of Economic Research.

라. (노동04) AI가 구현하는 알고리즘 매니지먼트(채용, 업무 배분, 감시, 평가)가 노동 통제를 강화하고 노동의 질을 저해하는 상황에서, 우리는 근로자의 기본적인 권리를 어떻게 보호할 것인가?

1) 제안의 배경

AI는 근로자를 관리하고 통제하는 직무에도 개입하므로 인적 자원 관리 방식에도 영향을 준다. 업무 배분, 속도 조절, 성과 평가, 보상, 심지어 해고에 이르는 전 과정에 AI가 개입하는 알고리즘 매니지먼트는 노동의 질적 변화를 유발한다. 실시간 데이터 기반의 감시와 불투명한 의사결정은 근로자의 자율성을 침해하게 되며, 근로자의 권리와 인간적 존엄성에 대한 새로운 위협이 될 수 있다.

2) 현황 및 쟁점

알고리즘 매니지먼트는 배달·운송 등 플랫폼 노동 영역뿐만 아니라 콜센터, 물류창고, 사무직, 원격근무자에게까지 확산되고 있다. 특히 사무직·원격근무자를 대상으로 키보드 입력, 마우스 움직임, 웹캠을 실시간으로 감시하는 ‘보스웨어(Bossware)’는 근로자의 사생활을 심각하게 침해할 수 있다. 노동의 질적 저하 외에도 알고리즘 기반 채용의 경우, 편향성의 문제가 심각하게 발생한다는 점에서 공정성 이슈가 우려된다. 기존 플랫폼 노동자를 중심으로 업무를 관리하는 알고리즘의 투명성에 대한 관심이 높아져 제도 정비가 시도되고 있으나, 기업의 이해와 상충되는 지점이 있어 권리 보호는 미성숙한 상황이다.

3) 문제의 심각성

기업은 ‘생산성 향상’과 ‘객관적 성과 관리’를 명분으로 AI 기반의 감시와 통제를 강화할 유인이 있다. 반면, 근로자는 ‘디지털 파놉티콘(Digital Panopticon)’ 하에 극심한 심리적 스트레스와 자기 검열에 시달리며, 헌법상의 사생활 보호 및 인간적 존엄성을 침해받게 된다. 기존 인간 관리자와 달리 AI는 노동 과정을 초단위로 쪼개어 감독하는 것은 근로자의 자율성과 숙련 형성 기회를 오히려 박탈시켜 노동을 ‘비인간화(dehumanization)’하게 된다. 이는 기존 산업혁명 시대와

유사하게 ‘디지털 테일러리즘(digital taylorism)’을 부활시켜, 장기적으로 노동의 질과 창의성을 심각하게 저해할 수 있다.

4) 기존 논의/대책의 한계

현행 「근로기준법」은 인간 관리자의 명시적인 지시와 평가를 전제로 하고 있어, AI 알고리즘에 의한 간접적이고 시스템적인 통제와 감시, 불이익 처분에 대해서는 규율하기 어려운 측면이 있다. EU의 GDPR이나 「AI 법(AI Act)」 등에서 논의되는 자동화된 의사결정에 대한 설명요구권이나 알고리즘 영향평가 등의 제도적 장치가 도입되고 있으나, 국내에서는 아직 초기 단계로 근로자가 불이익의 이유를 알거나 구제받을 실질적인 경로가 없다. 또한 이러한 국가별 혹은 지역별 규제가 분절적으로 존재하고 있어, 다국적 혹은 다지역 사업장은 중복 규제와 준수의 부담이 커지고 있다.

마. (노동05) AI가 창출한 초과 이익이 소수 기업에 독점되고 임금과 고용상승으로 연결되는 밴드왜건 효과가 부재한 상황에서, 이 이익의 범위는 어떻게 정의하며 누구와 어떤 원칙으로 공유되어야 하는가?

1) 제안의 배경

AI로부터의 가치 창출은 거대 모델, 데이터, 컴퓨팅 파워와 같은 무형자산 중심의 자본에 근간하므로 초과 이익은 이를 보유한 일부의 기업에 귀속된다. 반면, AI 모델을 학습하는 데 필수적인 데이터를 그간 생성하고 제공한 대중이나 자동화로 인해 노동시장에서 소외되거나 직무 전환 비용을 부담하는 근로자는 이익 분배 과정에 포함되지 못하고 있다. 이전의 기술과 달리 AI는 고용 감소나 일자리의 질을 하락시킨다는 점에서, 기업이 획득한 자본소득 증가분이 고용과 임금 상승으로 연결되지 못하여 밴드왜건 효과를 누리기 어렵기 때문에 AI 시대 초과 이익 공유의 필요성에 대한 관심도 증가하고 있다.

2) 현황 및 쟁점

기술 전환에 따른 초과 이익 공유에 대한 관심은 디지털 경제 혹은 플랫폼 경제

에서도 존재해 왔다. 데이터가 주요한 자원인 디지털 경제에서는 개인 데이터에 대한 기여 보상에 대한 필요성이 제기되며, 캘리포니아 주지사는 데이터 배당의 논의를 구체화하였다. 빌 게이츠나 조셉 스티글리츠 등의 사업가나 학자는 자동화 기기나 AI 소프트웨어 사용에 대한 로봇세 혹은 AI세를 부과하여 전환 비용으로 사용할 것을 제안하기도 하였다. 최근 IMF는 생성형 AI 영향 대응을 위한 자본세, 적극적 노동시장정책 등 조세재정 수단에 따른 시뮬레이션 분석을 통해 세금과 기본소득의 효과를 전망한 연구보고서를 발표한 바 있다. 이러한 분배 방식은 각각의 목적과 운영 방식의 한계와 지속성에 차이가 있다.

3) 문제의 심각성

AI 혜택이 소수에 집중되는 경향이 축적되어, 노동이 제대로 된 경제적 가치로 인정받지 못하고 분배 장치의 부재가 지속될수록 경제적 불평등과 사회적 양극화가 발생하게 된다. 특히 근로자의 임금만은 사회 전체 수요의 기반이라는 점에서 총수요가 위축되어 AI로부터의 생산성 향상이 경제 성장으로 이어지지 못하고 오히려 장기 침체를 유발할 수 있다.

4) 기존 논의/대책의 한계

초과 이익 공유의 필요성에 대한 이해관계자 간의 의견 차이가 존재한다. 기업이나 자본을 보유한 민간은 초과 이익 공유의 취지에는 공감할 수 있겠으나 불리한 원칙은 혁신을 위한 투자를 위축시켜 국가의 경쟁력을 저해할 수 있음을 우려하게 된다. 이 외에도 기술적이나 행정적으로 초과 이익의 조작적 정의를 내리는 것이 수월하지 않다. AI나 로봇의 범위, 초과 이익을 측정 가능한 과세 단위 등 새로운 표준을 정하고 측정 방법을 마련하는 것은 기술적으로 쉽지 않다. 또한 새롭게 제시되는 기본소득이나 데이터 배당 등의 대안은 막대한 재원이 필요하여 실현 가능성에 대한 한계도 있다.

제 2 절 AI와 공공행정: AI 기반 공공·행정서비스 수용성 제고 방안

1. 논의의 배경

가. 연구 배경 및 필요성

인공지능(Artificial Intelligence, 이하 AI)은 21세기 사회 전반에 걸쳐 가장 급속히 발전하고 있는 기술 중 하나로, 산업 전반은 물론 행정·정치·교육·복지 등 공공부문에도 큰 영향을 미치고 있다. 특히 최근 몇 년 사이에 등장한 생성형 AI(Generative AI), 대규모 언어모델(Large Language Models), 자율적 의사결정 시스템 등은 기존 행정업무의 틀을 근본적으로 바꿀 수 있는 잠재력을 지니고 있다. 이에 따라 공공 영역에서도 AI의 활용은 더 이상 선택이 아닌 필수적 과제로 부상하고 있으며, 많은 국가들이 정부 차원의 전략을 수립하고 인프라를 구축하고 있다(문명재 외, 2025).

AI 기술은 공공행정의 효율성 향상, 예산 절감, 시민 서비스 개선, 정책 수립의 과학화 등 다양한 영역에서 긍정적 효과를 창출할 수 있다. 예컨대 민원 처리 자동화, 복지 수혜자 예측, 공공안전 사전 감지, 도시 문제 예측·관리 등 AI는 방대한 데이터를 분석하여 정교한 행정 결정을 가능하게 한다(Wirtschaftler, 2025). 이를 통해 공공행정의 대응 속도와 정책 적합성이 높아질 수 있으며, 시민 중심의 맞춤형 서비스 제공도 현실화될 수 있다.

그러나 이와 동시에 AI의 공공영역 도입은 수많은 사회적 쟁점을 야기하고 있다. 공공의 가치, 즉 공정성, 형평성, 포용성, 투명성, 책임성 등과 기술적 자동화 간의 긴장 관계가 형성되고 있으며, 기술이 오히려 사회적 불평등을 심화하거나 기존 행정의 민주성을 훼손할 수 있다는 우려도 제기된다(송경희, 2025). 이러한 문제의식은 단순한 기술 개발과 인프라 구축만으로는 해결되지 않으며, AI 도입이 미치는 사회적·제도적 함의를 면밀히 검토하고 대응 방안을 마련하는 것이 시급한 상황이다.

공공행정에서 AI 기술이 실질적으로 작동하고 정착하기 위해서는 무엇보다도 시민의 신뢰와 수용을 확보하는 것이 핵심이다(박승규 외, 2023). 특히 AI 기술은 비가시적이고 복잡한 작동 구조를 지니며, 그 결과 또한 시민 입장에서 직관적으로 이해하기 어려운 경우가 많다. 이는 기술에 대한 불신과 오해를 야기할 수 있으며, 공공정책의 정당성과 효과성을 저해할 수 있다. 따라서 기술적 완성도와는 별개로, 시민의 수용성(acceptability)을 제고하기 위한 노력이 병행되어야 한다(구윤철, 2025; 이임정, 2024).

한편 디지털 포용성(digital inclusiveness)은 기술 발전의 성과가 사회 전 계층에 공평하게 분배되고, 소외되는 집단 없이 모두가 기술 혜택을 누릴 수 있도록 하는 개념이다. 이는 고령자, 장애인, 저소득층, 저학력층, 농어촌 거주자 등 디지털 취약계층이 새로운 행정서비스로부터 배제되지 않도록 하는 정책적 접근을 필요로 한다. 실제로 AI 기술 기반의 행정서비스는 정보 접근성, 디지털 역량, 교육 수준 등에 따라 사회 계층별로 활용 격차를 보이고 있으며, 이는 기술 도입이 오히려 새로운 사회적 격차를 재생산하는 결과로 이어질 수 있다(김동욱 외, 2025).

따라서 AI 기술을 공공행정에 도입하는 과정에서 ‘기술의 민주화’라는 관점이 반드시 병행되어야 한다. 이는 단순히 서비스를 개방하는 것을 넘어서, 사회적 약자도 실질적으로 활용할 수 있는 제도적, 인프라적, 문화적 조건을 마련하는 것을 의미한다. 공공서비스는 모든 국민이 동등하게 접근할 수 있어야 하며, 기술 활용이 특정 집단에만 유리하게 작동해서는 안 된다. 이와 같은 디지털 포용성을 실현하지 못할 경우, 행정의 신뢰도 저하와 함께 정책 집행의 정당성 위기도 초래될 수 있다(WEF, 2025; BCG, 2025).

기존의 AI 관련 정책과 연구는 주로 기술적 관점, 산업 경쟁력, 국가 디지털 전략 등에 초점을 맞추는 경향이 강했다. 그러나 최근에는 기술의 사회적 수용성과 사회적 책임성에 대한 논의가 확대되면서, 기술의 실질적 작동 조건에 대한 이해가 필수적인 과제로 부각되고 있다. 즉 기술 ‘사용자’로서 시민, ‘조정자’로서 정부, ‘개발자’로서 민간 기업 간의 상호작용과 역할 분담을 포함한 사회적 맥락을

고려하지 않는 기술 도입은 지속가능하지 않다.

공공부문에서는 특히 그러한 경향이 두드러지며, 기술적 유효성만으로는 정책적 설득력을 확보할 수 없다. 행정서비스는 시민과의 신뢰 관계 위에 구축되어야 하며, 정책 집행의 정당성은 시민의 참여와 수용을 통해 보장된다. 따라서 AI 기반 행정서비스 도입과 관련해서는 기술 중심의 단일한 효율성 논리가 아닌, 시민 참여, 민주적 가치, 공공성 확보 등 다원적 가치 실현이 필요하다(Straub et al., 2023).

이러한 관점에서 본 연구는 단순한 기술 도입 효과 분석을 넘어, AI 기반 행정서비스가 실제로 어떻게 사회에 수용되는지를 탐색하고, 시민의 태도와 인식, 사회구조적 수용 조건, 제도적 장치 등을 종합적으로 고찰하고자 한다. 이를 통해 정부의 AI·디지털 전환이 진정한 의미의 ‘포용적 혁신’으로 이어질 수 있도록 정책 방향을 제시하고자 한다.

나. 연구 배경 및 필요성

본 연구의 핵심 목적은 AI 기반 공공행정서비스 도입 과정에서 나타나는 사회적 이슈를 구조화하고, 시민 수용성과 디지털 포용성이라는 두 개념을 중심으로 정책적 대응 방안을 제시하는 데 있다. 이를 위해 본 연구는 다음의 세부 목적을 설정한다. 첫째, 국내외 AI 기반 공공행정서비스 추진 동향 분석이다. 보다 구체적으로 AI 기술이 공공행정에 어떤 방식으로 도입되고 있으며, 각국의 정책적 접근은 어떠한지를 비교 분석하고자 한다. 이를 통해 우리나라의 정책적 위치를 파악하고, 제도적 정합성과 방향성을 모색한다. 특히 수용성 제고를 위한 정책 수단, 법제도 정비, 이해관계자 간 협력 체계 구축 방식 등을 검토함으로써 실질적 정책 설계를 위한 기초자료를 확보한다. 둘째, AI 도입에 대한 일반 시민의 인식과 태도 조사이다. AI 기반 행정서비스에 대한 일반 시민의 인식을 실증적으로 분석함으로써, 기술 수용에 영향을 미치는 요인을 규명하고자 한다. 이를 위해 인식 조사 및 통계 분석을 통해 사회집단별 기대 수준, 신뢰도, 기술 이해도, 공정성 인식, 프라이버시 우려 등을 다차원적으로 분석한다. 특히 성별, 연령, 교육 수준,

디지털 역량, 거주지역 등 인구통계학적 변수에 따라 수용 태도의 차이가 어떻게 나타나는지를 정량적으로 규명한다. 셋째, AI 공공행정의 수용성 전략을 도출하기 위한 전문가 및 공공기관 실무자 심층 인터뷰이다. AI 기술 도입의 실무적 관점과 정책 실행상의 과제를 파악하기 위해, 전문가와 실무자를 대상으로 한 심층 인터뷰를 수행한다. 이를 통해 기술 도입 시 촉진 요인과 장애 요인을 식별하고, 실질적인 정책 실행을 위한 개선 과제를 도출한다. 특히 서비스 설계·도입·운영 단계에서 행정적 제약, 조직 문화, 시민 참여 구조 등에 대한 실질적 분석을 통해 실천적 정책 제언을 마련한다. 마지막으로 AI 공공행정 활성화를 위한 정책 제언 및 제도 개선 방향 도출이다. 앞선 분석을 토대로, 정부 및 공공기관이 AI 기반 행정 서비스를 보다 포용적으로 설계하고, 시민 수용성을 제고할 수 있도록 하는 정책적·제도적 방향을 제시하고자 한다. 이는 AI·디지털 정부 추진 전략, 공공 AI 윤리 가이드라인, 시민 참여 플랫폼 설계, 행정 담당자 역량 강화 방안 등을 포괄하는 종합적 정책 프레임으로 이어질 것이다.

이와 같은 연구 목적하에 본 절은 총 5개 목차로 구성된다. 다음에서는 국내외 AI 기반 공공행정서비스 추진 사례를 분석하여 기술 도입의 정책적 함의를 도출하고, 일반 시민의 인식조사 결과를 통해 수용성의 조건을 검토한다. 그리고 전문가 및 실무자 심층 조사를 통해 정책적 과제를 도출하며, 마지막에는 연구 결과를 종합하여 향후 정책적 시사점을 제안하고자 한다.

2. 주요 분석 내용

가. 국내외 AI 기반 공공·행정서비스 동향

1) 조사 목적

공공부문에서 도입 및 운영되고 있는 AI 공공·행정서비스의 활용 실태와 동향을 파악하고자 하며, 이는 다음과 같은 두 가지 목적을 지닌다.

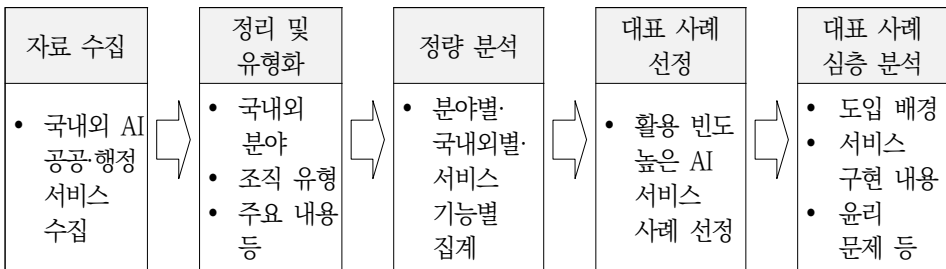
첫째, 2023년부터 2025년까지 최근 3개년간 공공부문에서 추진(계획) 중인 국내외 AI 공공·행정서비스 사례를 수집 및 분석하여 전반적인 현황을 파악하고, 서

비스 수혜 대상 및 목적, 분야, 기능 등 다양한 기준에 따라 유형화하여 AI 공공·행정서비스의 특성과 패턴을 도출하고자 한다. 둘째, 분석 결과를 바탕으로 설문 조사 문항 설계 기준이 될 서비스 유형과 활용 범위 등을 구체화하여, AI 공공·행정서비스에 대한 국민 인식 및 태도 조사를 위한 설문 문항 설계의 기초자료를 마련하고자 한다.

2) 분석 방법 및 절차

중앙정부 및 지방정부, 공공기관 등과 같은 공공부문 주도로 개발 또는 제공(계획) 중인 국내외 AI 공공·행정서비스 사례를 수집하기 위해, 2023년부터 2025년까지 최근 3년간 발간된 보도자료와 연구보고서 등을 살펴보았다. 이를 통해 수집한 국내외 AI 공공·행정서비스는 총 197개로, 수집한 사례는 서비스 분야, 서비스명, 추진기관, 주요 내용, 조직 유형(지방정부·중앙정부·공공기관) 등을 기준으로 정리하고 유형화하였으며, 분야별·국내외별·서비스 기능별로 집계하여 비교하였다. 이를 통해 분야별 활용 빈도가 높은 AI 서비스 대표 사례를 선정하고, 대표 사례의 도입 배경, 서비스 구현 내용, 윤리적 문제 등에 대한 심층적인 조사를 진행하였다.

[그림 3-10] 국내외 AI 기반 공공·행정서비스 동향 분석 방법 및 절차



자료: 연구진 작성

3) 분류기준

AI 공공·행정서비스는 수혜 대상을 기준으로 ‘대민서비스’와 ‘내부 행정서비스’

로 구분된다. ‘대민서비스’는 국민과 기업 등을 수혜 대상으로 하는 AI 공공·행정 서비스로, 국민이 일상생활에서 겪는 불편 사항을 개선하고 국민의 행정 만족도와 효율성을 제고하는 것을 목적으로 한다. ‘내부 행정서비스’는 정부와 공공기관 등을 수혜 대상으로 하는 AI 공공·행정서비스로, 중앙정부 및 지방정부, 공공기관 등의 내부 업무 효율성과 편의성 제고, 비용 절감을 목적으로 한다.

AI 공공·행정서비스 분야는 정부기능분류체계(Business Reference Model in Korea, BRM)³⁰⁾를 준용하여 17개 분야로 분류한 뒤, 설문 설계에 적합하도록 사용자 관점에서 유사한 분야들을 통합 및 재분류하였다. 이를 통해 도출된 분야는 1. 행정, 2. 안전, 3. 보건복지, 4. 교통, 5. 교육/연구, 6. 문화관광, 7. 경제산업, 8. 환경 8개이다.

〈표 3-10〉 AI 공공·행정서비스 분야

분야	정부기능분류체계(BRM)	
행정	1. 일반 공공행정	• 국가통계, 국민권익·인권, 국정운영, 국정홍보, 법제, 일반 행정, 재정·금융, 지방행정·재정지원, 개인정보
	2. 지역개발	• 산업단지, 수자원, 지역 및 도시
	3. 재정·세제·금융	• 금융, 기획재정, 세제
안전	4. 공공질서 및 안전	• 경찰, 법무 및 검찰, 안전관리, 해경, 고위공직자범죄수사
	5. 국방	• 병무행정
보건복지	6. 보건	• 건강보험, 보건의료, 식품의약품안전, 감염병 및 질병
	7. 사회복지	• 고용노동, 공적연금, 기초생활보장, 노인·청소년, 보건의료, 보육·가족 및 여성, 보훈, 사회복지 일반, 주택, 취약 계층지원

30) 정부기능분류체계는 중앙행정기관의 업무와 그와 관련된 정보를 체계적으로 관리·활용하기 위해 행정기관의 업무를 기능 중심으로 분류한 것으로, 정책분야·정책영역·대기능·중기능·소기능 등 그 기능의 수준에 따라 분류한 체계를 말한다(행정안전부, 공공데이터포털, <https://www.data.go.kr/data/15062615/fileData.do>).

분야	정부기능분류체계(BRM)	
교통	8. 교통 및 물류	• 도로, 물류 등 기타, 철도, 항공·공항, 해운·항만
교육/ 연구	9. 과학기술	• 과학기술진흥, 과학기술연구, 원자력기술
	10. 교육	• 고등교육, 교육일반, 유아 및 초·중등교육, 평생·직업교육
문화관광	11. 문화 체육관광	• 관광, 문화예술, 국가유산, 문화체육관광 일반, 체육
	12. 통신	• 방송통신, 우정
경제산업	13. 산업·통상, 중소기업	• 공정거래, 무역 및 투자유치, 산업·중소기업 일반, 산업금융, 산업진흥·고도화, 에너지 및 자원개발, 통상
	14. 농림	• 농업·농촌, 임업·산촌
환경	15. 환경	• 대기, 상하수도·수질, 자연, 폐기물, 해양환경, 환경 일반

자료: 국립세종도서관 정책정보포털, <https://policy.nl.go.kr/pages/information/guide.jsp>

4) 분석 결과

국내외 AI 공공·행정서비스 사례를 분야별로 살펴보면, 행정 분야가 87건(44.2%)으로 가장 많은 비중을 차지하였으며, 이어서 보건복지 33건(16.8%), 교통 24건(12.2%), 안전 19건(9.6%) 순으로 나타났다.

〈표 3-11〉 분야별 AI 공공·행정서비스

분야	건수	비율
행정	87	44.2%
보건복지	33	16.8%
교통	24	12.2%
안전	19	9.6%
환경	12	6.1%
교육/연구	12	6.1%
경제산업	6	3.0%
문화관광	4	2.0%
합계	197	100.00%

자료: 연구진 작성

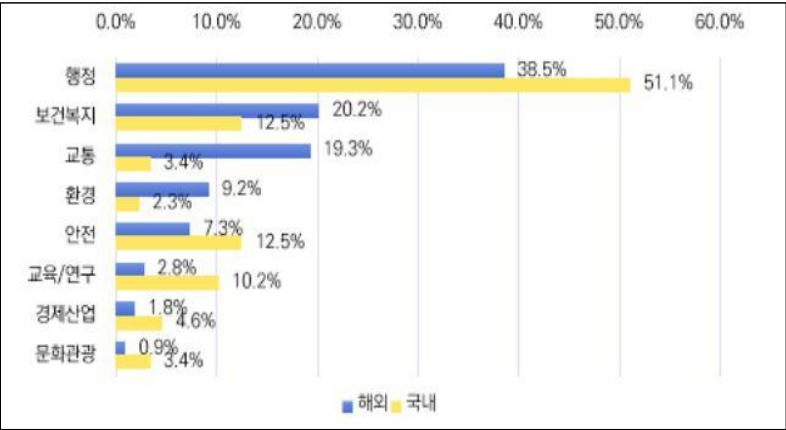
국내외 AI 공공·행정서비스 분야를 살펴보면, 해외의 경우 행정(42건, 38.5%), 보건복지(22건, 20.2%), 교통(21건, 19.3%), 환경(10건, 9.2%) 등의 순이며, 국내의 경우 행정(45건, 51.1%), 보건복지(11건, 12.5%), 안전(11건, 12.5%), 교육/연구(9건, 10.2%) 등의 순으로 나타났다.

〈표 3-12〉 국내외 AI 공공·행정서비스 분야 비교

분야	해외		국내	
	건수	비율	건수	비율
행정	42	38.5%	45	51.1%
보건복지	22	20.2%	11	12.5%
교통	21	19.3%	3	3.4%
환경	10	9.2%	2	2.3%
안전	8	7.3%	11	12.5%
교육/연구	3	2.8%	9	10.2%
경제산업	2	1.8%	4	4.6%
문화관광	1	0.9%	3	3.4%
합계	109	100.0%	88	100.0%

자료: 연구진 작성

[그림 3-11] 국내외 AI 공공·행정서비스 분야 비교



자료: 연구진 작성

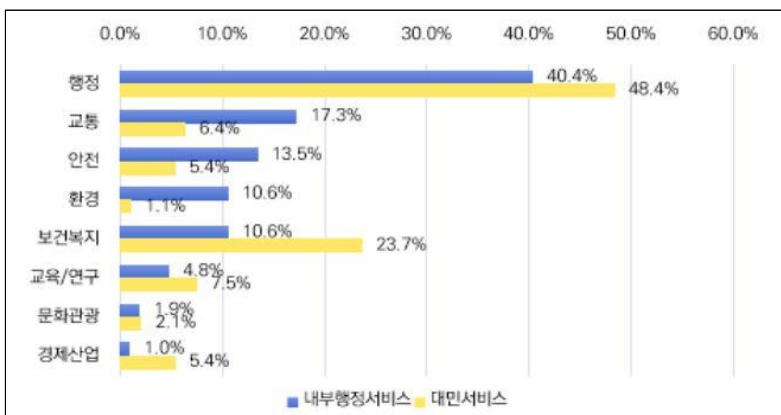
내부 행정서비스와 대민서비스를 비교하여 살펴보면, 내부 행정서비스의 경우 행정(42건, 40.4%), 교통(18건, 17.3%), 안전(14건, 13.5%) 등의 분야에서, 대민서비스의 경우 행정(45건, 48.4%), 보건복지(22건, 23.7%) 등의 분야에서 AI가 활발하게 도입 및 활용되는 것으로 나타났다.

〈표 3-13〉 내부 행정서비스-대민서비스 간 AI 공공·행정서비스 분야 비교

분야	내부 행정서비스		대민서비스	
	건수	비율	건수	비율
행정	42	40.4%	45	48.4%
교통	18	17.3%	6	6.4%
안전	14	13.5%	5	5.4%
환경	11	10.6%	1	1.1%
보건복지	11	10.6%	22	23.7%
교육/연구	5	4.8%	7	7.5%
문화관광	2	1.9%	2	2.1%
경제산업	1	1.0%	5	5.4%
합계	104	100.0%	93	100.0%

자료: 연구진 작성

[그림 3-12] 내부 행정서비스-대민서비스 간 서비스 분야 비교



자료: 연구진 작성

분야별 AI 공공·행정서비스 기능을 살펴보면, 행정 분야에서는 ‘민원 상담 및 안내 자동화’와 ‘행정업무 지원’에, 보건복지 분야에서는 ‘노인 돌봄’과 ‘의료진 업무 지원’에, 교통 분야에서는 ‘교통 운영 최적화’가, 안전 분야에서는 ‘범죄 예방 및 수사’에, 환경 분야에서는 ‘폐기물 관리’가, 교육/연구 분야에서는 ‘연구 지원’에, 경제·산업 분야에서는 ‘기업 지원’에, 문화관광 분야에서는 ‘맞춤형 관광정보 제공’에 AI가 주요하게 적용 및 활용되고 있는 것으로 나타났다.

〈표 3-14〉 분야별 AI 공공·행정서비스 기능

분야	건수	비율
행정	87	44.2%
민원 상담 및 안내 자동화(챗봇, 보이스봇, 로봇 등)	32	36.8%
행정업무 지원(문서 요약 및 작성, 정보검색 등)	28	32.2%
법률 정보 제공(관계 법령 및 판례 등)	10	11.5%
정책 수립 및 의사결정 지원(데이터 분석 등)	9	10.3%
공공 인프라 관리(국토변화 탐지, 건물 관리 등)	4	4.6%
서비스 접근성 강화(외국인 대상 통번역 등)	2	2.3%
시민 사회 참여 확대(시민참여 앱 개발, 문서 가독성 향상 등)	2	2.3%
보건복지	33	16.8%
노인 돌봄(응급상황 대응, 건강관리, 대화 등)	9	27.3%
의료진 업무 지원(질병 진단 지원, 의료정보 기록·관리 등)	8	24.2%
구인구직 지원(이력서 첨삭, 일자리 매칭 등)	5	15.2%
복지 혜택 대상자 선정 지원(심사 자동화 등)	4	12.1%
맞춤형 복지 혜택 정보 제공	4	12.1%
질병 예방·관리 지원(건강 예측 및 변화 감지, 건강정보 제공 등)	3	9.1%
교통	24	12.2%
교통 운영 최적화(실시간 최적 경로 설정, 신호 최적화 등)	9	37.5%
교통 인프라 관리 지원(도로 유지보수 등)	7	29.2%
교통 법규 위반 감시 및 단속(신호위반 등)	3	12.5%

분야		건수	비율
	교통사고 대응 지원 및 안전 강화(실시간 사고 감지 및 경보 등)	3	12.5%
	자율주행 기반 조성	2	8.3%
안전		19	9.6%
	범죄 예방·수사(범죄 실시간 모니터링 및 탐지, 수사 지원 등)	11	57.9%
	신속한 사고 대응 지원(사고 모니터링, 긴급신고체계 효율화 등)	7	36.8%
	군수물자 예측 지원	1	5.3%
환경		12	6.1%
	폐기물 관리(불법 투기 단속, 분류 등)	5	41.7%
	환경정보 모니터링 및 예측(기후, 대기질, 빙공해 등)	4	33.3%
	친환경 에너지 효율화(태양광, 스마트그리드 등)	3	25.0%
교육/연구		12	6.1%
	연구 지원(연구 정보 제공 등)	7	58.3%
	교직원 업무 지원(시험 평가 자동화, 학교생활기록부 작성 지원 등)	3	25.0%
	교육 및 학습 지원	2	16.7%
경제산업		6	3.0%
	기업 지원(시장·투자·인증·무역정보 제공, 컨설팅 등)	5	83.3%
	농작물 피해 조기 감지·예측	1	16.7%
문화관광		4	2.0%
	맞춤형 관광정보 제공(여행 일정 추천 등)	2	50.0%
	관광산업 활성화 지원(데이터 분석 등)	1	25.0%
	국가대표 선수 육성 지원(경기 영상 분석 등)	1	25.0%
총합계		197	100.0%

자료: 연구진 작성

5) 분야별 사례

앞선 분석에서는 공공부문의 각 분야에서 활용되는 AI의 기능(목적)에 대해 살펴보고, 이를 토대로 활용 빈도가 높은 AI 기능의 서비스 일부를 대표 사례로 선정하였다. 이에 따라 선정한 8개 분야(1. 행정, 2. 안전, 3. 보건복지, 4. 교통,

5. 교육/연구, 6. 문화관광, 7. 경제산업, 8. 환경)별 국내외 AI 공공·행정서비스 대표 사례는 다음과 같다.

가) 행정

행정 분야에서는 AI를 활용하여 단순 반복적인 민원 응대를 자동화하고, 24시간 이용 가능한 서비스 제공을 통해 국민의 서비스 접근성을 개선하였다. 또한 문서 요약 및 작성 지원이나 정보 검색 등과 같은 직원의 행정업무를 지원하고, 정책 수립 및 의사결정 지원에도 AI가 활용된다.

대표 사례를 살펴보면, 서울시의 AI 챗봇 ‘서울톡’이 있다. ‘서울톡’은 120 전화 상담에서 자주 문의하는 데이터를 학습하여, 서울시정에 대한 단순·반복적인 문의에 대해 24시간 자동 답변이 가능하다. 서울톡은 행정 분야 564종, 현장민원 접수 54종 등 주요 분야(교통, 행정, 복지, 환경, 문화, 경제, 안전, 주택 등)에 대한 정보를 제공하고, 공공시설 예약, 증명서 발급 절차, 도로명 주소 안내 등과 같은 서비스를 포함한다.³¹⁾ 2020년부터 운영된 서울톡은 2025년 8월부터 생성형 AI(Large Language model, LLM) 기술을 접목하여, 민원인과의 대화 맥락을 이해하여 자연스럽게 응답할 수 있는 고도화 사업인 ‘챗봇 2.0’을 추진한다. 생성형 AI 도입 과정에서 민감한 공공 데이터나 개인정보 유출에 대한 우려가 제기됨에 따라, 폐쇄망 운영으로 해당 위험성을 최소화할 계획이다.³²⁾

일본 이바라키현청은 청각 장애가 있는 공무원들의 근무 환경을 개선하고자 대화를 실시간으로 시각화하여 보여주는 서비스인 ‘Vuevo’를 도입하였다. 무선 마이크, 디바이스(스마트폰·태블릿·PC 등), 전용 앱 등으로 구성된 Vuevo는 무선 마이크의 위치 정보를 기반으로 발화자를 표시하고, 발언 내용과 청각 자료를 실

31) 한국지역정보개발원(2023), “현장 리포트 서울시 인공지능상담사 챗봇 ‘서울톡’”, 《KLID 웹진》, Vol. 139, http://klidwz.or.kr/webzine/vol139/sub_2_1_01.html.

32) Goover(2025), “서울시 ‘챗봇 2.0’으로 본 AI 행정 민원 혁신 전략”, 《Goover 리포트》, <https://seo.goover.ai/report/202508/go-public-report-ko-5db1943a-c3b5-40ab-84fe-c103a12c6aeb-0-0.html>.

시간으로 시각화한다. 또한 자동으로 회의 내용을 정리하여 요약문을 작성하기도 한다. 공무원들의 대화 및 회의 내용을 자동으로 기록하는 기능으로 인해 프라이버시 보호와 데이터 유출에 대한 우려가 있으며, 이에 대한 보완책으로 녹음 내용을 디바이스 외의 데이터베이스에 별도로 저장하지 않는 방식을 취한다.³³⁾

호주 뉴사우스웨일스주(New South Wales, NSW)는 공무원 인력이 부족하고 복잡한 승인 절차로 인해 주택 건설 승인이 지연되는 문제를 해결하고자 주택개발 승인 업무에 AI를 도입하였다. 해당 AI 시스템은 주택개발 신청자가 관련 서류를 제출하기 전과 후에 문서 형식, 규정 준수 여부 등을 점검하여 공무원의 주택개발 승인 업무를 지원한다. 심사 결과에 대한 책임 문제를 명확히 하기 위해 오타, 오류, 누락, 규정 위반 등과 같이 단순하고 명료한 사항에 대한 서류 검토는 AI가 맡고, 공무원은 AI의 검토 의견을 참고하여 승인에 대한 최종적인 의사결정을 내린다.³⁴⁾

〈표 3-15〉 ‘행정’ 분야 AI 공공·행정서비스 사례

구분	영역
한국 서울시, 민원 응대 AI ‘서울톡’	• 생성형 AI 모델과 연동하여 대화의 맥락을 이해하고 자연스럽게 민원에 응대할 수 있도록 기존의 AI 챗봇 고도화
일본 이바라키현, 실시간 대화 시각화 서비스	• 청각 장애가 있는 직원들의 근무 환경을 개선하고자 실시간 대화를 요약하고 시각화하는 서비스 ‘Vuevo’ 도입
호주 뉴사우스웨일스주, 주택 승인 업무에 AI 도입	• 주택개발 신청서가 제출되기 전이나 후에 서류를 검토하고 평가하여 승인 업무를 지원

자료: 한국지역정보개발원(2023); Goover(2025); 한국지역정보개발원(2024)

33) 한국지역정보개발원(2024), “생성형 AI 서비스 최근 동향: 멀티모달 서비스를 중심으로”, 《KLID 디지털 이슈 리포트》, 24-5호.

34) New South Wales Government(2024), “NSW Planning wants AI product solutions to make faster development applications,” Department of Planning, Housing and Infrastructure, <https://www.nsw.gov.au/media-releases/nsw-planning-wants-ai-product-solutions-to-make-faster-development-applications>.

나) 안전

안전 분야에서는 실시간 범죄 모니터링 및 탐지 등과 같은 범죄 예방·수사와 신속한 사고 대응 활동에 AI가 주로 활용되고 있다.

인도 우타르프라데시주(Uttar Pradesh, UP) 경찰청은 범죄 수사를 위해 생성형 AI를 활용한다. 경찰청의 범죄 수사 업무를 지원하는 ‘CrimeGPT’는 90만 개 이상의 범죄 데이터를 학습하여 용의자 추적 및 체포를 지원한다. 범죄자의 이미지, 음성, 텍스트 데이터를 입력하면, CrimeGPT는 해당 정보와 데이터베이스를 대조하여 용의자를 압축하거나 특정하는 표적 맞춤형 정보를 제공한다.³⁵⁾ 데이터 편향으로 인한 AI의 잘못된 결정을 방지하기 위해 최종 판단과 결정은 인간이 내리고, 필터링된 데이터만을 학습시켜 위험을 최소화한다.³⁶⁾

한국 광주광역시경찰청은 딥페이크 등과 같은 디지털 성범죄가 급속도로 확산함에 따라, 성범죄에 대해 신속하게 수사하고 피해 유형별로 맞춤형 해결책을 마련하기 위해 성범죄 수사에 AI를 도입하였다. 기존 성범죄 수사는 피해자가 기관을 직접 방문하여 피해 사항에 대해 전달하는 대면 방식으로 이루어지고 있어, 수사 과정이 길어지고 2차 피해가 발생할 수 있다는 지적이 있었다.³⁷⁾ 이러한 문제를 해결하기 위해 LLM 기술을 활용한 온라인 상담을 가능하게 하고, 수사절차 안내, 유사 판례 및 법률 정보 제공 등과 같은 개인 맞춤형 솔루션을 제공한다.³⁸⁾

35) 한국지정정보개발원(2024), “생성형 AI 서비스 최근 동향: 멀티모달 서비스를 중심으로”, 《KLID 디지털 이슈 리포트》, 24-5호.

36) 유진(2024. 5. 6), “인공지능 줌인: 크라임GPT ‘범죄와의 전쟁’에서 획기적 역할”, 《위키리스크한국》,
<http://www.wikileaks-kr.org/news/articleView.html?idxno=152275>.

37) 디지털플랫폼정부위원회 보도자료(2025. 2. 26), “공공부문 초거대 AI 도입·활용 가이드라인 2.0 발표”,
<https://www.dpg.go.kr/DPG/contents/DPG02020000.do?schM=view&id=20250226110028349393&schBcid=press>.

38) 디지털플랫폼정부위원회(2025), 『공공 AI 서비스 실증 사례집』.

〈표 3-16〉 ‘안전’ 분야 AI 공공·행정서비스 사례

구분	내용
인도 우타르프라데시주 경찰, ‘CrimeGPT’	• 범죄 용의자를 추적하고 체포할 수 있도록 지원하는 소프트웨어 ‘CrimeGPT’ 도입
한국 광주광역시경찰청, 성범죄 위기대응 AI	• 성범죄와 관련된 대량의 텍스트 데이터에 사전 훈련한 거대 언어 모델(LLM)을 활용하여 피해 진술 요약, 성범죄 유형 분류, 유사 판례 및 법률 정보 제공 등을 지원

자료: 위키리스크한국(2024); 디지털플랫폼정부위원회(2025)

다) 보건복지

보건복지 분야에서는 노인 돌봄을 목적으로 AI 기반 응급상황 감지 및 대응, 맞춤형 건강관리서비스 등을 제공하고 있으며, 데이터 기반 진단 보조, 의료정보 기록 및 관리 자동화 등을 통해 의료진의 업무를 효율적으로 지원하고 있다.

고령화율이 높은 일본 가나가와현 요코스카시가 노인의 치매 예방을 목적으로, 일상 대화에 특화된 생성형 AI 앱을 개발하기 위해 AI에 쇼와 시대(1926년 12월 26일~1989년 1월 7일) 뉴스들을 학습시켰다.³⁹⁾ 해당 앱은 노인과 주 2~3회 대화를 나누며, 뇌를 활성화하기 위해 과거의 기억을 떠올리게 돕는다. 시범사업을 통해 발견된 문제점을 개선하고, AI와의 대화 내용을 가족에게 공유하며, 건강상태를 분석해 알려주는 기능을 도입할 계획이다.⁴⁰⁾

싱가포르 정부는 의료 분야에서 AI를 활발하게 사용하는 나라로, 공공의료기관에서 환자의 질병을 조기 발견하고 맞춤형 치료를 지원하는 AI 영상 솔루션 ‘AimSG’를 도입하였다. 해당 AI 영상 솔루션은 AI가 의료 영상을 자동 분석하고 응급 치료가 필요한 환자를 분류하여 영상의학과 전문의가 더욱 정확하게 영상 보고서를 작성할 수 있도록 지원한다.⁴¹⁾ 의료 분야에 AI를 도입하는 것에 대해 프

39) 고대영·이나영(2024. 10. 11), “[초고령사회, 일본에 길을 묻다] ② 치매예방 말벗·스마트 양로원…‘AI 돌봄 도우미’에서 길 찾다”, 《이투데이》,
<https://www.etoday.co.kr/news/view/2408280>.

40) KBS(2025. 4. 9), “일본, ‘AI와 주고받는 대화’…치매 예방 실험”, 《KBS 뉴스》,
<https://news.kbs.co.kr/news/pc/view/view.do?ncd=8222719>.

라이버시 침해, 알고리즘 편향, 환자 동의 등과 같은 사회·윤리적인 문제가 우려된다. 이에 대해 싱가포르 정부는 의료 AI 개발자를 위한 원칙⁴²⁾을 마련하고, 윤리적·제도적·사회적 문제를 다루기 위한 공개적인 논의를 병행하고 있다.⁴³⁾

〈표 3-17〉 ‘보건·복지’ 분야 AI 공공·행정서비스 사례

구분	영역
일본 요코스카시, 치매 예방 AI 서비스	• 요코스카시는 고령화 문제 해결을 위해 Starley Inc.와 협력하여 대화형 생성 AI를 활용한 치매 예방 서비스를 개발
싱가포르, AI 기반 의료 진단	• 공공의료기관에서 질병을 조기에 발견하고 개인 맞춤형 치료를 지원하기 위해 AI 영상 솔루션인 ‘AimSG’를 도입

자료: 이투데이(2024); KBS(2025); Ang(2024); Gaw(2024)

라) 교통

교통 분야에서는 실시간 경로 안내 및 신호 제어를 통해 교통 흐름을 최적화하고, 교통 인프라 유지·관리를 지원하며, 교통 법규 위반에 대한 자동 감시 및 단속 등에 AI를 활용하고 있다.

체코 프라하시는 도시 전역에 AI 기반 교통 관리 시스템을 도입했다. 통신기술청(Technical Administration of Communications, TSK)에서 개발한 이 시스템은 고정된 패턴에 따라 신호를 제어하였던 기존의 시스템과는 달리, AI를 적용하여 실제 교통 상황에 맞춰 교통 신호를 실시간으로 최적화하도록 설계되었다.

41) Ang(2024. 1. 24), “싱가포르의 공공보건 분야에서 인공지능 도입의 배경”, 《Healthcare IT News》,
<https://www.healthcareitnews.com/news/asia/behind-singapores-widespread-ai-adoption-public-health>.

42) Ministry of Health, Artificial Intelligence in Healthcare Guidelines(2021),
[https://www.moh.gov.sg/docs/librariesprovider5/eguides/1-0-artificial-in-healthcare-guidelines-\(aihgle\)_publishedoct21.pdf](https://www.moh.gov.sg/docs/librariesprovider5/eguides/1-0-artificial-in-healthcare-guidelines-(aihgle)_publishedoct21.pdf).

43) Gaw(2024. 4. 29), “싱가포르의 인공지능과 생명과학: 융합과 규제 동향”, 국제변호사협회,
<https://www.ibanet.org/artificial-intelligence-life-sciences-singapore>.

자동차, 버스, 자전거, 보행자 등 도로 이용자에 따라 최대 이동 시간과 수요 등을 고려하여 신호를 조정하고, 가장 교통량이 많은 지역은 정지 횟수를 줄여 교통 체증 완화를 돕는다.⁴⁴⁾

영국 국도청(National Highways)은 AI를 통해 고속도로에서 운전자들의 부주의한 행동을 감시하고 불법행위를 적발한다. AI 기반의 카메라 시스템은 운전자의 안전벨트 미착용, 운전 중 휴대전화 사용이나 음식 섭취, 안전하지 않은 화물 적재 등과 같은 위험한 행동을 감지하고, 이를 관계 기관에 알린다. 2022년 12월 2주간 진행된 시범사업 결과, 헤이독의 M6 고속도로 23번 교차로 남쪽 도로 공사 현장에 설치된 카메라가 90초마다 위법 행위를 적발하였으며, 모든 현장은 평균 6분마다 1건을 적발한 것으로 나타났다.⁴⁵⁾

〈표 3-18〉 ‘교통’ 분야 AI 공공·행정서비스 사례

구분	영역
체코 프라하시, AI 교통관리 시스템	• 프라하시는 AI 교통 관리 시스템을 통해 전체 지역의 교통 상황을 평가하고, 원활한 교통 흐름을 위해 신호를 최적화
영국 국도관리청, 위험 운전자 단속	• AI 기반 카메라를 통해 운전자의 상태를 실시간 수집 및 분석하고, 운전 중 휴대전화 사용 여부 등을 단속

자료: Prague Morning(2024); BBC News(2024)

마) 교육·연구

교육·연구 분야에서는 맞춤형 학술정보 제공을 통해 연구 활동이나 학습을 지원하는 데 AI가 주요하게 활용된다.

한국 경제·인문사회연구회는 연구 및 정책 개발 역량을 높이기 위해 정책연구에

44) Prague Morning(2024. 11. 10), “New AI-Powered Traffic Lights to Improve Prague’s Streets,” <https://praguemorning.cz/new-ai-powered-traffic-lights-to-improve-pragues-streets/>.

45) BBC News(2024. 2. 15), “Vehicle-based littering solution pilot achieves 95% detection rate,” <https://www.bbc.com/news/uk-england-leicestershire-68575556>.

특화된 초거대 AI 서비스 구축을 추진한다. 해당 서비스는 연구의 계획-수행-활용 및 성과 확산 등과 같이 단계별로 연구를 지원한다. 구체적으로 연구계획 단계에서는 연구 주제나 방법론을 제안하고 연구계획 수립에 관한 질문에 대한 답변을 제공한다. 연구수행 단계에서는 기존의 연구계획서, 보고서, 선행연구 등을 요약 및 분석하여 제공한다. 연구보조 및 지원 단계에서는 연구 과정에서 진행한 회의의 회의록 생성, 문법·맞춤법·문맥 오류 탐지, 다국어 번역 등을 지원한다. 결과 활용 및 성과 확산 단계에서는 발표자료 생성 등과 같이 연구 결과를 공유하고 성과를 확산하기 위한 활용을 지원한다.⁴⁶⁾

대만 국립공공정보도서관(國家公共資訊圖書館, National Library of Public Information, NLPI)은 GPT-4o를 활용하여 이용자의 독서 취향에 맞는 도서를 추천하는 서비스를 제공한다. 해당 서비스는 이용자의 지난 2년간 대출 기록을 분석하여 대출과 관련된 이용자의 특징과 취향 등을 파악하고, 대출 기록이 없는 이용자는 몇 가지 질문을 통해 이용자의 취향과 관심 분야에 적합한 도서 목록을 제공한다. 해당 서비스는 개별 이용자가 15권의 책을 대출할 때마다 자동으로 재분석하여 추천 도서 목록을 업데이트한다.⁴⁷⁾

46) 경제·인문사회연구회(2024), “초거대 AI를 이용한 통합 연구자원 생성, 관리 서비스 개발”, 《미래정책 포커스》, 2024 가을호,

https://www.nrc.re.kr/board.es?act=view&bid=0044&issue_cd=42&list_no=178745&mid=a30200000000&nPage=1&tag=.

47) 문화체육관광부 국립중앙도서관 사서지원서비스(2025. 3. 7), “스마트도서관의 전담 독서 컨설턴트-대만국립공공정보도서관 ‘AI 도서 추천’ 서비스”,

https://librarian.nl.go.kr/LI/contents/L30201000000.do?schM=view&id=50552&schBcid=BBSMSTR_451.

〈표 3-20〉 ‘문화관광’ 분야 AI 공공·행정서비스 사례

구분	영역
한국 서울관광재단, AI 여행 플래너	• 생성형 AI 챗봇 도입을 통해 단순 질의응답을 넘어 관광에 특화된 개별 맞춤형 대화 서비스를 다국어로 제공

자료: 서울관광재단(2025)

사) 경제산업

경제산업 분야에서는 기업에 시장·투자·무역 정보 등을 제공하거나 컨설팅을 지원하는 데 AI가 주요하게 활용된다.

한국 대한무역투자진흥공사(Korea Trade-Investment Promotion Agency, KOTRA)는 기업의 수출을 지원하는 플랫폼을 구축한다. AI가 수출과 관련된 데이터를 수집·분석하고, 이를 바탕으로 수출 전 주기 프로세스에 걸쳐 기업의 수출을 지원한다. 수출기업은 AI가 추천하는 진성 바이어와 매칭되고, 131개의 해외 무역관에서 제공하는 AI 기반 시장정보 분석 서비스에 손쉽게 접근하여 획득할 수 있게 된다.⁴⁹⁾

〈표 3-21〉 ‘경제산업’ 분야 AI 공공·행정서비스 사례

구분	영역
한국 KOTRA, 지능형 무역투자 플랫폼	• 기업의 수출 전 과정에서 AI 분석 기반 맞춤형 수출지원 서비스를 제공하는 지능형 무역투자 플랫폼 구축

자료: COSMETIC INSIGHT(2025)

아) 환경

환경 분야에서는 환경정보 모니터링, 폐기물 관리, 불법 투기 단속, 스마트 에너지 그리드 등의 다양한 영역에서 AI를 도입하여 활용하고 있다.

49) 길태운(2025. 8. 21), “KOTRA, AI로 중소·중견기업 수출길 넓힌다”, 《COSMETIC INSIGHT》, <https://www.cosinkorea.com/news/article.html?no=55911>.

영국의 차른우드(Charnwood Borough Council)는 도로에 버려지는 쓰레기를 줄이기 위해 쓰레기 불법 투기 차량의 번호판을 자동으로 인식하는 AI 카메라를 워릭웨이(Warwick Way)에 설치하였다. 카메라가 차량에서 던져지는 쓰레기를 포착할 경우 차량 소유자에게는 벌금이 부과된다. 12주간 시범운영을 통해 해당 지역에 버려진 쓰레기 양이 37% 감소했으며, 28건의 벌금이 발행 및 납부되었거나 항소 또는 채권 추심이 발생했다.⁵⁰⁾

〈표 3-22〉 ‘환경’ 분야 AI 공공·행정서비스 사례

구분	영역
영국 차른우드, 쓰레기 무단투기 단속	• 워릭웨이에 배치되어 도로에 버려지는 쓰레기와 함께 번호판을 자동 인식하여 벌금 부과

자료: BBC News(2024)

6) 시사점

AI는 공공부문에서 업무처리 속도와 비용 절감, 서비스 접근성 향상 등을 가능하게 하여 다양한 분야에서 도입이 활발하다. 2023년부터 2025년까지 최근 3년간 국내외 AI 공공·행정서비스 사례를 살펴보면, 행정, 안전, 보건복지, 교통, 교육/연구, 문화관광, 경제산업, 환경 등 각 분야의 내부 행정서비스와 대민서비스에 AI가 도입되고 있다. AI가 공무원의 업무를 보조하거나 시민과 직접 소통하고 있으며, 심사나 평가 같은 의사결정을 내리는 사례도 살펴볼 수 있다.

공공부문에서 특히 AI 활용이 활발한 분야는 ‘행정’ 분야인 것으로 나타났다. 행정 분야에서는 공무원의 업무를 효율화하고 기존의 대민서비스를 개선하기 위해, AI를 기반으로 한 민원 상담 및 안내 자동화, 보고서 및 회의록 작성 등이 이루어지고 있다. 이는 공공부문에서의 근무 환경이 인간과 AI 간의 협업이라는 구조로 변화하고 있음을 의미하며, 공공서비스를 활용하고 공적인 혜택을 받는 시민

50) Isaac Ashe(2024. 5. 17), “AI camera system caught litter louts in Loughborough, AI camera system caught litter louts in Loughborough,” *BBC News*.

또한 AI를 활용해야 하는 환경에 놓여있음을 의미한다. 행정 분야 이외에도 일상생활과 밀접한 ‘보건복지’, ‘교통’ 등의 분야에서도 AI가 활발하게 도입되고 있어, 일상생활에서 AI 활용은 점점 불가피해질 것으로 보인다.

AI 공공·행정서비스 분야를 국내의 비교를 통해 살펴보면, 해외의 경우 행정, 보건복지, 교통 세 분야에서 AI가 활발하게 활용되는 반면, 국내의 경우 행정 분야에 집중되어 활용되고 있다. 국내와 해외를 비교하였을 때 가장 큰 차이를 보이는 분야는 교통 분야로, 국내에서는 교통 분야의 AI 활용도가 해외에 비해 낮은 수준인 것으로 나타났다. 이는 해외에서는 스마트시티 정책 또는 민간을 중심으로 AI가 각 분야에 적용되는 반면, 국내의 경우 공공 주도로 AI가 적용되고 있음을 시사한다.

각 분야별 AI 공공·행정서비스 사례를 살펴본 결과, AI의 프라이버시 침해, 데이터 유출, 감시 등과 같은 윤리적 문제 발생에 대한 우려가 제기되고 있다. 특히 시민의 일상생활과 밀접하게 관계있거나 민감정보를 다루는 서비스의 경우 해당 문제가 부각된다. 또한 보건, 안전 등과 같은 사람의 생명과 신체, 정신건강 등과 관련된 분야에서는 데이터 편향이나 오탐지, 설명 불가능성 등에 대한 우려가 더욱 강조되고 있다. 점차 AI의 적용 범위가 넓어지고 일상생활과 밀접해질 것으로 예상됨에 따라 단순히 효율성 등의 효과만을 고려한 AI 기술의 만연한 도입보다는, 신뢰성 확보를 위한 규범이나 가이드라인 정비, AI 접목에 따른 영향을 평가하고 이에 대한 대응방안을 마련하는 것이 중요해질 것으로 보인다.

나. AI 기반 공공·행정서비스에 대한 일반시민의 인식과 태도

1) 주요 이론에 관한 논의

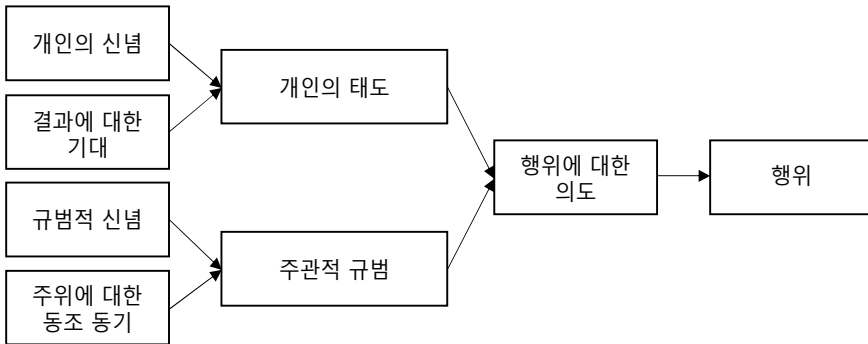
① 합리적 행위 이론(TRA)

Fishbein & Ajzen(1975)은 사람들의 행위를 설명하기 위한 모형으로서 ‘합리적 행위 이론’을 제안한 바 있다. 합리적 행위 이론은 인간의 행동이 즉흥적이거나 우연에 의한 것이 아니라, 자신의 신념과 결과에 대한 판단(개인의 태도), 그리고 주변을 의식(주관적 규범)하는 수준에 따라 행위 의도가 형성되고, 이러한

행위 의도는 종국에 행위로 나아감을 설명하고 있다.

구체적으로 살펴보면 1) 개인의 태도(attitude)는 ① 행위가 초래할 결과에 대한 개인의 신념(behavioral beliefs), 그리고 ② 행위가 초래할 결과에 대한 평가(evaluation of outcomes)에 의해 구성된다. 다음으로 2) 주관적 규범(subjective norms)은 ① 행위에 대한 주위의 태도를 의식하는 규범적 신념(normative beliefs), 그리고 ② 나와 가까운 주위에 동조 또는 순응하려는 동기(motivation to comply)로 구성된다. 즉 합리적 행위 이론은 개인의 태도와 주관적 규범이라는 두 가지 심리적 요인이 행위 의도를 설명하고, 행위 의도는 다시 실제 사람들의 행위를 설명할 수 있다는 논리적 구조를 제시하는 것으로 이해할 수 있다. 이를 도식화하면 [그림 3-13]과 같다.

[그림 3-13] 합리적 행위 이론(TRA: Theory of Reasoned Action)



자료: Fishbein & Ajzen(1975)

Fishbein & Ajzen(1975)의 합리적 행위 이론은 사람들의 행동을 인과적 논리에 따라 설명·예측할 수 있다는 점에서 행정학 및 정책학은 물론 사회과학 전반에 걸쳐 널리 사용되고 있는 이론이다. 그러나 해당 이론은 몇 가지 한계점을 내포하고 있다. 첫째, 사람들이 종종 보이는 즉흥적인 행동에 대해서는 설명하지 못하는 한계가 있다. 예를 들어 별다른 고민이 필요하지 않은 행동, 또는 다른 사람의 시선으로부터 자유로운 행동 등은 위 합리적 행위이론을 통해 설명하는 것이 바람

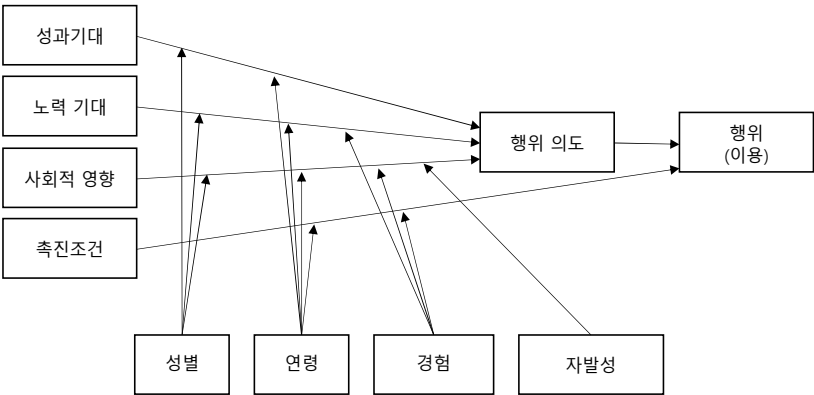
직하지 않다. 둘째, 합리적 행위 이론은 특정 행동을 위해 요구되는 개인의 내재적 역량, 또는 행위자를 둘러싼 외부 환경의 변화와 같은 상황에 대해서도 고려하지 못하는 한계를 지니고 있다(Ajzen & Fishbein, 1980).

② 기술수용모형(TAM)과 통합된 기술수용모형(UTAUT, UTAUT2)

기술수용모형은 기술에 대한 개인의 행동 의도와 실제 행동을 설명하기 위해 Davis(1989)가 제안한 모형으로 지각된 유용성과 지각된 사용 편의성이 사용 태도를 형성하고, 행동 의도가 최종적으로 실제 사용으로 이어진다고 본다. 기술수용모형은 새로운 기술이나 서비스를 개발하고 도입하는 과정에서 기술이 수용되기 위해 사용자 입장에서 무엇이 필요한지를 설명해 준다. 이후 다양한 모형으로 응용되어 활용되었으며, 통합기술수용모형(UTAUT) 역시 기술수용모형을 포함하여 다양한 모형을 통합하여 기술수용을 설명하고자 하였다.

통합기술수용모형(UTAUT)은 성과기대(performance expectancy), 노력기대(effort expectancy), 사회적 영향(social influence), 촉진 조건(facilitating conditions)이라는 네 가지 주요 개념을 제시하였고, 성별(gender), 연령(age), 경험(experience), 자발성(voluntariness of use)이 기술 수용에서 조절 역할을

[그림 3-14] 통합된 기술수용모형(UTAUT)

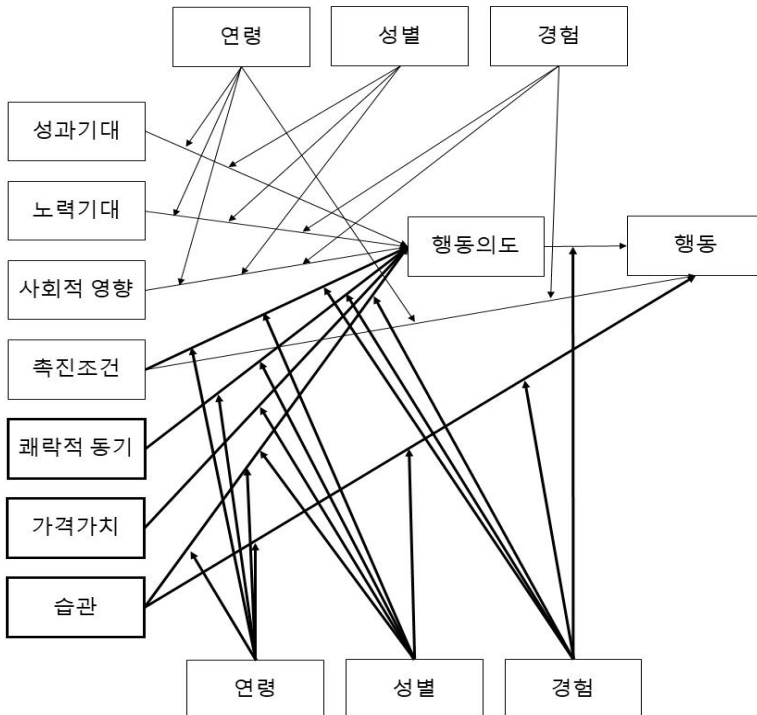


자료: Venkatesh et al.(2003)

하고 있음을 상정한다. 성과기대, 노력기대, 사회적 영향은 행동 의도에 대하여 영향을 주는 변인이며, 촉진 조건은 사용 행동에 직접적으로 영향을 주는 변인이다. 통합기술수용모형은 기존의 기술수용모형에 비해 높은 설명력을 지니는 것으로 확인되었다(Venkatesh et al., 2003).

UTAUT2는 일반 소비자 측면의 설명력을 높이기 위해 쾌락적 동기(hedonic motivation), 가격 가치(price value), 습관(habit) 변인을 새롭게 추가하여 구성한 모형이다. Venkatesh et al.(2012)은 UTAUT를 소비자 측면에 적용시켜 일반화의 가능성을 확장시키고자 하였다. AI 기반 공공·행정서비스를 사용하는 사용자 입장에서 수용성을 설명하기 위해서는 소비자 측면에서 제안된 개념들을 활용하는 것이 유용할 것으로 판단된다.

[그림 3-15] 통합된 기술수용모형2(UTAUT2)

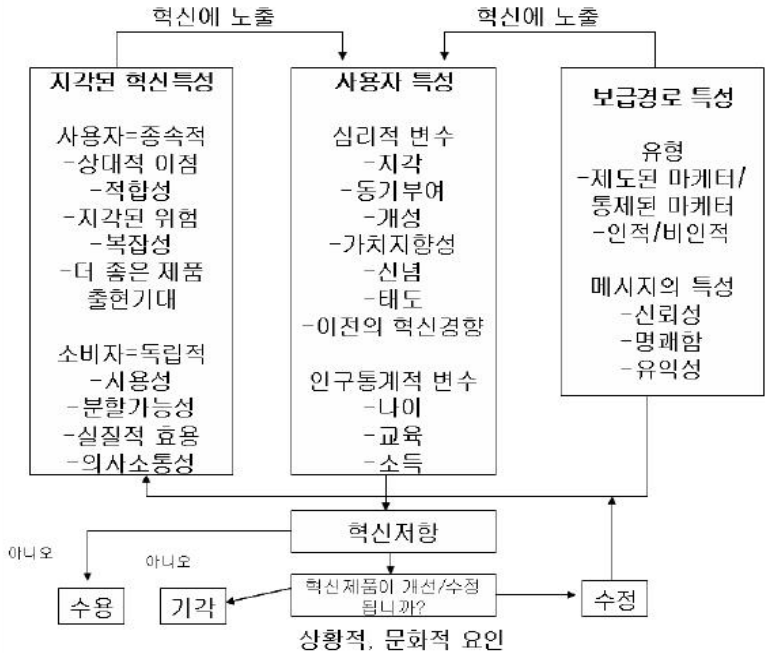


자료: Venkatesh, Thong, Xu(2012)

③ 혁신저항모형(Model of Innovation Resistance)

Ram(1987)은 혁신이 야기하는 변화에 대한 소비자들의 저항을 설명하기 위해 혁신저항모형을 제안하였다. 그는 기존의 연구들이 혁신이 ‘좋은 것’이라는 가정에 장점에만 초점을 맞추어 혁신 확산과 수용을 다루었다고 비판하며, 혁신 채택과 수용 과정을 잘 이해하기 위해 저항을 다루어야 한다고 주장하였다. Sheth(1981)의 혁신저항 개념에 기초해 혁신 특성, 사용자 특성, 보급경로 특성(촉진 특성)의 세 가지 요인으로 구성된 혁신저항모형을 제시하였다. 보급경로의 특성은 커뮤니케이션 모델에서도 주목하는 메시지(정보)와 매체(채널)의 특성과 연계된다. 혁신 저항 모형을 활용한 연구들은 주로 혁신 특성과 사용자 특성에 초점을 맞춰 연구가 진행되었다.

[그림 3-16] 혁신저항모형



자료: Ram(1987)

지각된 혁신 특성은 크게 소비자 종속적 변인과 소비자 독립적 변인으로 구분된다. 소비자 종속적 변인은 소비자가 각각의 속성에 대해 어떻게 인지하느냐에 따라 반응이 달라질 수 있는 변수를 의미하며, 소비자 독립적 변인은 모든 소비자들로부터 동일하게 나타나는 변수로 실증 모형에서는 제외된다. 소비자 종속적 변인에서 주로 활용되는 변인은 상대적 이점(relative advantage), 적합성(compatibility), 지각된 위험(perceived risk), 복잡성(complexity)이며, 소비자 특성 중 심리적 변수로는 혁신에 대한 지각(필요성, 호오), 동기, 개성(성격), 가치 지향성, 신념, 태도, 이전의 혁신 경험이 포함된다. 인구학적 특성으로는 연령, 교육, 소득이 포함되어 있다.

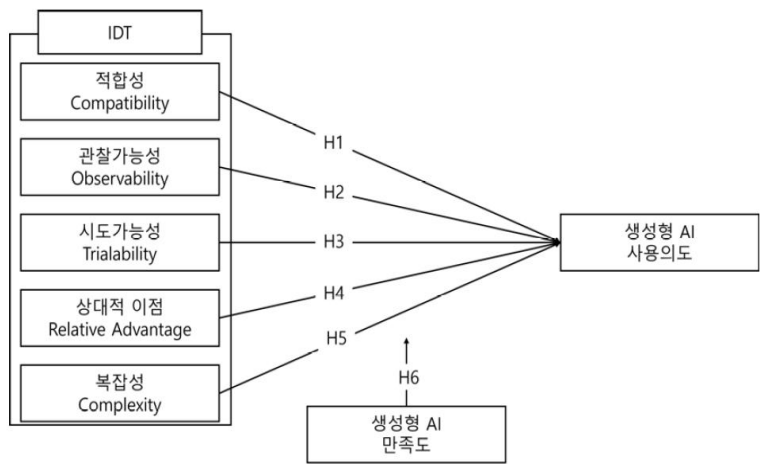
Ram(1987)이 제안한 혁신 저항 모형은 각각의 개념들이 어떻게 혁신의 수용에 영향을 미칠 수 있는지 기술하고, 이후 Sheth와 함께 마케팅 전략에 초점을 맞춘 후속 연구를 진행하였다(Ram & Sheth, 1989). Ram이 직접 실증 연구를 진행하지는 않았으나, 이후 신기술의 수용과 저항을 연구하는 학자들에 의해 혁신 저항 모형이 꾸준히 활용되고 있다.

④ 혁신확산이론(Innovation Diffusion Theory, IDT)

Rogers(2003)의 혁신확산이론⁵¹⁾은 새로운 기술이나 아이디어가 특정 커뮤니케이션 채널을 통해 시간의 흐름에 따라 사회체계 내에서 구성원들에게 어떻게 전파되는지를 설명하는 이론이다. 혁신채택에 미치는 영향요인으로 인구통계학적 특성, 내재적 혁신성, 혁신의 지각된 특성을 제시하였다. 특히 혁신의 지각된 특성인 상대적 이점(relative advantage), 적합성(compatibility), 복잡성(complexity), 이용 가능성(trialability), 관찰 가능성(observability)은 새로운 기술이나 아이디어에 대한 불확실성을 감소시킴으로써 혁신의 확산 속도에 영향을 미친다. 설재성(2025)은 혁신확산이론을 적용하여 시각디자인 분야에서 생성형 AI의 사용 의도에 영향을 미치는 요인들을 확인하였다.

51) 혁신확산이론은 1962년 Rogers에 의해 처음 제안되었고, 이후 1995년에 보다 정교하게 발전되었다(Miller, 2018).

[그림 3-17] 혁신확산이론을 적용한 연구모형 사례



자료: 설재성(2025)

⑤ 가치기반 수용모형(Value-based Adoption Model, VAM)

가치기반 수용모형은 Kim, Chan, and Gupta(2007)에 의해 모바일 인터넷의 수용을 설명하기 위해 제안된 모형이다. 새로운 기술을 수용할 때 사용자가 지각된 가치(perceived value)를 핵심 기준으로 삼게 되는데, 이는 기술을 통해 얻게 될 이익(benefit)과 감수해야 할 희생(sacrifice) 사이의 상충 관계를 종합적으로 고려하여 결정된다. 즉 기술이 주는 유용함이 사용자가 감수해야 할 복잡함이나 비용보다 더 큰 가치가 있는지를 고려해 수용 의도로 이어진다.

기술을 통해 얻게 될 혜택으로는 유용성(utility)과 즐거움(enjoyment)이 포함된다. 유용성은 사용자가 새로운 기술을 사용함으로써 인지하는 총체적 가치이며, 결과의 기대치이다. 즐거움은 예상되는 성과와 별개로 제품을 사용하는 활동 자체가 즐겁다고 인지되는 정도를 의미한다.

[그림 3-18] 가치기반 수용모형

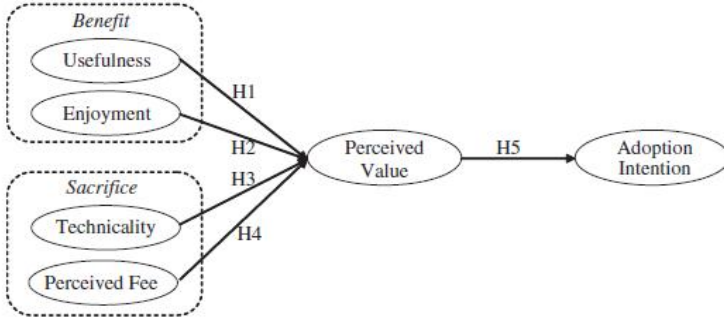


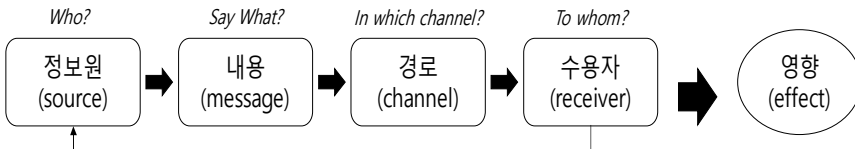
Fig. 1. Value based adoption model of technology.

자료: Kim, Chan, and Gupta(2007)

⑥ 커뮤니케이션 모형

Lasswell(1948)의 커뮤니케이션 모형은 의사소통 과정을 설명하기 위한 가장 쉬운 방법으로, 누가(who), 무엇을(say what), 어떤 매체를 통해(in which channel), 누구에게(to whom), 어떤 효과를 가져오는지(with what effect)를 설명하는 것이라고 주장하며 제안된 모형이다.

[그림 3-19] SMCRE 커뮤니케이션 모형



자료: 홍은영·이선이(2021)

커뮤니케이션 모형은 다양한 학문 분야에서 통용되고 있으며, 위험 커뮤니케이션 분야에서도 활발히 활용되고 있다. 다만 SMCRE 모형은 과정 모형으로 구성되었기에 주로 과거의 위험에 대한 커뮤니케이션 과정을 상세히 기술하기 위한 틀

로 사용되거나, 경로분석이 진행되기도 한다(예: 송해룡·김원제·정세일, 2010). 또한 위협의 종류나 적용 범위에 따라 차원별로 상이한 구성요인들이 설정되어 왔다. 예를 들어, 정보원(S) 차원에서는 소통 역량 및 노력, 전문성 등이 포함되었으며, 정보(M) 차원에서는 정보의 양적 측면(충분성 등)과 질적 측면(유용성, 용이성 등)이 다루어져 왔다. 채널(C)은 효율성과 양방향성, 신뢰성, 다양성 등으로 구성되었는데, 최근 뉴미디어의 발달로 인해 전통매체와 뉴미디어의 효과를 고려하는 경향이 있다(예: 류현숙·홍승희, 2015; 홍은영·박천희, 2023). AI 기반 공공·행정서비스 역시 편익과 위협이 공존하는 만큼 관련 정보가 어떻게 소통되는지에 따라 수용성에 중요한 영향을 미칠 수 있을 것으로 예상된다.

2) 변인의 설계

본 연구에서는 앞선 이론적 논의를 토대로 AI 기반 공공·행정서비스에 영향을 미칠 것으로 가정할 수 있는 주요 변인을 도출하였다. 연구모형의 간명성을 위해 기존 선행연구에서 강건하게 영향을 미치는 요인들을 중심으로 변인을 선정하였음을 밝힌다.

① 매개 및 종속변인: 저항 및 수용성

매개변인으로 설정한 혁신 저항은 통상적으로 개인적 수준의 것과 사회적 수준의 것으로 구분할 수 있다. 개인적 차원의 혁신저항은 이용자가 자신의 일상이나 업무에 혁신기술이 적용되는 것을 거부하는 내적 성향을 의미(예: 사적재화의 구매과정에서 발생하는 내적 갈등)한다. 반면 사회적 차원의 혁신저항은 우리 사회에 미칠 영향력이나 파급력을 우려하여 발생하는 집단 기반의 저항 성향을 뜻한다. 본 연구에서는 개인적인 차원보다는 사회적인 차원에 초점을 두었는데, 이는 AI 기반 공공·행정서비스가 지닌 광범위한 파급력을 고려하고자 했기 때문이다.

종속변인인 수용성은 사회적 수용성, 그리고 마지못한 수용성으로 분류하였다. 사회적 수용성은 우리 사회가 AI 기반 공공·행정 서비스를 사용해야 한다고 생각하는 정도를 의미한다. 마지못한 수용성은 개인적 차원에서 수용할 수밖에 없는

어쩔 수 없는 상황을 인지하고, 결국에 비자발적으로 AI 기반 공공·행정서비스를 수용할 것인지를 포함하는 개념이다. 마지못한 수용성은 사람들로 하여금 미래의 위험 인식에 대하여 인지 부조화를 야기할 수 있는 부정적인 개념이다(박천희·홍은영, 2023). 이러한 수용성은 사람들이 현재의 위험을 올바르게 인식하지 못하게 하며, 나아가 위험한데도 사람들이 낙관적 편향 태도를 갖도록 유발하는 원인이기도 하다. 본 연구에서는 이상의 두 가지 수용성을 중심으로 인과기제를 검토하기로 한다.

② 독립변인: 개인 특성, 커뮤니케이션 특성, 기술적 특성, 정부 신뢰

(a) 개인 특성은 행위에 대한 신념, 지각된 위험, 내재적·외재적 동기, 혁신성, 주관적 지식으로 구분하였다. 행위에 대한 신념은 합리적 행위 이론의 주요한 변인 중 하나이다. 이러한 신념은 특정 기술이나 서비스에 대한 태도를 형성하는 원인이기도 하다. 지각된 위험은 혁신저항모형에서 차용한 것으로서 특정 기술이나 서비스에 대한 위험 지각의 수준을 의미한다. 이러한 위험의 수준을 높게 지각할수록 대상에 대한 저항은 증가하고, 수용은 감소하는 결과로 나아가게 된다. 내재적·외재적 동기는 UTAUT, 혁신저항모형, 가치기반모형에서 차용한 것으로서 동기의 수준이 높을수록 저항 태도는 감소하는 방향으로 나아간다. 혁신성과 주관적 지식은 개인의 특성 차원에서 연구되곤 하는데, 혁신성과 지식의 수준이 높을수록 특정 기술이나 서비스에 대한 저항 태도는 낮아지는 것으로 연구된다(박천희·홍은영·문정욱, 2021).

(b) 커뮤니케이션 특성은 정보원에 대한 공적 신뢰, 정보가 지닌 질적인 특성, 뉴미디어(채널)에 대한 신뢰로 구분하였다. 정보원에 대한 신뢰는 전통적으로 정부 신뢰 영역에서 자주 다뤄왔던 개념이다(Poortinga & Pidgeon, 2003). 공식적인 정보를 전달하는 정보원에 대한 신뢰가 높을수록 사람들은 그로 인하여 발생하는 결과에 대해 수용하는 태도를 보이게 된다. 반면 정보원에 대한 신뢰가 부족할 경우 사람들은 전달 받은 메시지에 대해 불신하는 태도를 갖게 된다(김원제

외, 2009). 다음으로 정보의 질적 특성 역시 신기술에 대한 사람들의 인식에 영향을 미치는 주요한 변인으로 연구된다(정익재, 2014). 일반적으로 주어진 정보의 질적 우수성이 높은 경우에는 특정 기술이나 서비스에 대한 사람들의 부정적 태도는 감소하는 방향으로 나아가게 된다(홍은영 외, 2021). 이어서 채널에 대한 신뢰 역시 매우 유의미한 변인으로 연구되고 있다. 소위 ‘미디어가 메시지다’라는 McLuhan(1964)의 명제와 같이 채널 그 자체가 미치는 영향력은 상당하다. 특히 본 연구와 같이 새로운 기술로 인해 특정 위험이 불가피하게 발생할 수 있는 영역에서는 그 효과가 더욱 증가할 수 있다(Ader, 1995; Schaefer & Birkland, 2007).

(c) 기술적 특성은 Rogers(2003)의 혁신확산 이론에서 도출된 주요 개념 중 상대적 이점과 복잡성을 차용하였다. 상대적 이점은 ‘새로운 기술이 기존의 기술 또는 아이디어보다 더 나은 정도’를 의미(Rogers, 2003: 213)하므로 새로운 기술에 대한 혁신 저항에는 부(-)의 관계, 수용에는 정(+)의 관계를 갖는 것으로 검증되어 왔다(박천희·홍은영·문정욱, 2021). 디지털 전환에 대한 수용 태도를 다룬 Xioaqi & Zainal(2025)의 연구에서도 상대적 이점이 정적 효과를 갖는 것으로 확인되었다. 스마트워크 도입과 관련해서도 상대적 이점은 혁신저항의 수준을 낮추는 것으로 확인되었다(곽영·신민수, 2021). 이처럼 상대적 이점은 AI 기반 공공·행정서비스에 대한 혁신저항을 낮추고, 수용성에는 정적 영향을 미칠 것으로 예상된다.

복잡성은 혁신 저항에 정적인 영향을 미치는 것으로 확인되어 왔다(예: 윤수경·김명지·최준호, 2014, 신재권·이상우, 2016). 전자책 수용에 대한 윤수경 외(2014)의 연구에서는 복잡성이 혁신저항에 정적 영향을 미치고, 혁신저항은 수용의도에 부적 영향을 미치는 것으로 확인되었다. 신재권 외(2016)의 연구에서도 손목형 웨어러블 디바이스의 복잡성은 혁신 저항에 정적 영향을 미치고, 혁신 저항은 수용의도에 부적 영향을 미치는 것으로 확인되었다. 기존 연구의 결과들을 통해 복잡성은 AI 기반 공공·행정서비스에 대한 혁신저항에는 정적(+), 수용성에

는 부정(-) 영향을 미칠 것으로 예상된다.

(d) 마지막으로 정부 신뢰는 공정성과 가치 유사성으로 구성하였다. 신뢰는 상호 보완적인 하위 차원들로 구성된 다차원적 개념으로(Poortinga & Pidgeon, 2003), 학자들마다 상이한 차원의 개념과 측정지표를 다룬다(이대웅·손주희·권기현, 2018). Poortinga & Pidgeon(2003)은 일반적 신뢰 수준으로 역량, 신뢰성, 청렴성, 관심, 공정성, 개방성 등과 정부 정책에 대한 의심 수준, 그리고 가치 유사성으로 신뢰를 측정한다. Hurley(2011)는 특정 상황하에 신뢰 수준을 결정하는 관계적 요인으로 안전성, 공유가치의 유사성, 이해관계의 일치, 호의적인 배려 정도, 예측 가능성과 정직성, 역량, 커뮤니케이션 수준 등을 제시하였다.

공정성(Fairness)은 일반적 신뢰를 구성하는 필수적인 구성 요소로, 정부가 편향되지 않고 공평하게 정책과 규제를 집행할 것이라는 인식을 의미하며, 이는 수용에 정적 효과를 갖는다(Seidl & Drögemüller, 2024). 특정 기술이나 정책이 복잡하고 불확실할 때, 시민들 스스로 위험을 평가하기 어렵기 때문에 정부 및 규제 기관이 공정할 것이라는 신뢰가 수용성을 높일 수 있다(Poortinga & Pidgeon, 2003). 가치 유사성은 시민이 규제 기관이나 전문가 집단이 자신들과 중요한 목표나 이해관계, 도덕적 관점을 공유하고 있다고 인식하는 정도를 의미한다. 가치 유사성은 대중이 기관들을 심층적으로 평가하기 전에 직관적으로 판단하는 근거를 제공하며, 이는 기술 및 정책에 대한 수용성에 영향을 준다(Poortinga & Pidgeon, 2006).

③ 조절변인: 사회적 지지, 휴리스틱 판단

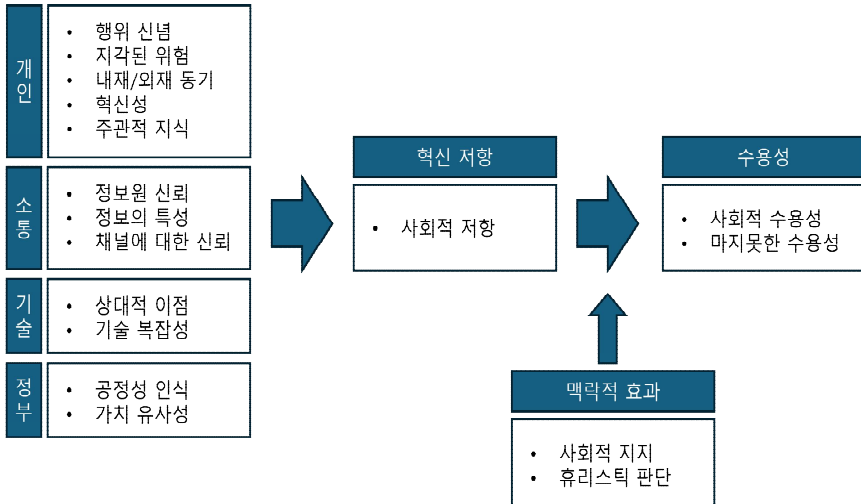
본 연구의 주요한 질문 중 하나는 동일한 저항 태도를 갖고 있음에도 누군가는 이를 극복하고 수용으로 나아가는 반면, 누군가는 왜 그러하지 못하는가를 밝히는 데 있다. 이를 검증하기 위해 혁신저항과 수용성 간 관계에 두 가지 조절변인을 가정하였다. 하나는 사회적 지지(social support)이며, 또 다른 하나는 휴리스틱(heuristic)이다.

사회적 지지란 주변 인물로부터 받는 정서적, 도구적 지지 수준을 의미한다(Chen & Chan, 2014). 사회적 지지는 지각된 유용성과 용이성에 긍정적인 영향을 미치고, 나아가 수용성에도 기여한다(Venkatesh & Davis, 2000). 이진명·박성지·전민선(2025)의 연구에서는 사회적 지지가 기술 유용성과 용이성을 매개로 수용성을 높이는 요인임을 확인하였다. 따라서 사회적 지지를 높게 지각하는 사람들은 혁신 저항 태도로 인한 사회적 수용성 감소가 완만하게 발생할 것으로 예상해 볼 수 있다.

휴리스틱-체계적 모델은 개인의 정보처리 과정이 두 가지 경로를 통해 이루어진다고 주장하는데(Eagly & Chaiken, 1993), 휴리스틱 처리는 복잡한 과제를 아주 간단한 판단 작업으로 단순화하여 의사결정을 하는 심리적 기제로, 심리적 지름길(mental shortcuts)을 통해 어림짐작으로 정보처리가 이루어지는 방식이다(Tversky & Kahneman, 1974). Trumbo & McComas(2003)에 의하면, 수신자가 동일한 정보원 신뢰를 지각했다 하더라도 정보를 체계적(systemic)으로 판단하는 사람과 감정적(heuristic)으로 처리하는 사람 간 다른 영향력을 행사할 수 있다고 보았듯이, 정보처리의 과정이 수용성에 다른 효과를 가져올 것으로 예상할 수 있다. 본 연구에서는 휴리스틱 판단을 하는 사람들의 경우, 혁신 저항 태도로 인한 마지못한 수용성 증가가 더욱 가파르게 나타날 것으로 가정하였다.

이와 같은 가정은 AI 기반 공공·행정서비스의 수용성 인과기제를 밝히는 데 도움이 될 뿐만 아니라, 실천적으로 의미 있는 함의를 도출할 수 있게 할 것이다. 이상의 논의를 토대로 AI 기반 공공·행정서비스 수용성을 설명하기 위한 연구모형은 다음과 같다.

[그림 3-20] 연구모형



자료: 연구진 작성

3) 설문 문항 개발

앞서 살펴본 이론적 배경의 주요 문헌을 토대로 본 연구에 활용한 개념의 조작적 정의를 진행하였다. 구체적으로 분석에 활용한 설문문항 및 신뢰도 분석 결과는 다음의 <표 3-23>과 같다.

〈표 3-23〉 측정문항 및 신뢰도 분석 결과

유형	측정 개념	진 술 문	신뢰도 계수
개인 특성 (x1)	행위 신념	AI 기반 공공행정 서비스를 활용한다면, 내 상황에 맞는 맞춤형 지원을 손쉽게 받을 수 있을 것이다.	.742
		AI 기반 공공행정 서비스를 활용한다면, 기존의 복잡했던 행정 절차를 쉽게 해결할 수 있을 것이다.	
		AI 기반 공공행정 서비스를 활용한다면, 사람에 의한 서비스보다 더욱 공정하게 업무가 처리될 것 같다.	
	지각된 위험	AI 기반 공공·행정서비스로 인해 개인정보 유출의 위험이 높아질 것이다.	.749
		AI 기반 공공·행정서비스로 인해 개인정보가 오남용될 가능성이 있다.	
	내재적 동기	AI 기반 공공행정 서비스를 이용하는 것은 나에게 재미있고 흥미로운 경험일 것 같다.	.781
		AI 기반 공공행정 서비스를 이용하는 것은 나의 일상생활에 즐거움을 주는 경험일 것 같다.	
	외재적 동기	AI 기반 공공·행정서비스를 이용하면 시간과 비용을 절약할 수 있을 것이라고 생각한다.	.745
		AI 기반 공공·행정서비스를 이용하면 정부가 제공하는 혜택이나 서비스를 더욱 쉽게 얻을 수 있을 것이라고 생각한다.	
	혁신성	나는 대체로 새로운 기술이 적용된 제품이나 서비스를 다른 사람보다 먼저 사용하는 편이다.	.845
		나는 첨단 기술이 적용된 제품이나 서비스를 먼저 사용하려는 시도를 하는 편이다.	
커뮤니 케이션 특성(x2)	정보원 신뢰	나는 인공지능 기술에 대해 잘 알고 있다.	.846
		나는 인공지능 기술을 다른 사람들에게 잘 설명할 수 있다.	
	정보의 질적 특성	공적 정보원(정부, 공공기관 등)은 AI 기반 공공·행정서비스와 관련된 정보를 정직하게 전달할 것이다.	.758
		공적 정보원(정부, 공공기관 등)은 AI 기반 공공·행정서비스와 관련하여 신뢰성 높은 정보를 전달할 것이다.	
	정보의 질적 특성	내가 접한 AI 기반 공공·행정서비스에 대한 정보는 이해하기 쉽다.	.673
	뉴미디어 신뢰	내가 접한 AI 기반 공공·행정서비스에 대한 정보는 나에게 도움이 된다.	
		뉴미디어(유튜브, 인스타그램, 페이스북, X 등)는 객관적인 정보를 제공한다.	.854

유형	측정 개념	진 술 문	신뢰도 계수
커뮤니 케이션 특성(x2)	뉴미디어 신뢰	뉴미디어(유튜브, 인스타그램, 페이스북, X 등)를 통해 전달 되는 정보는 믿을 수 있다.	.854
기술적 특성 (x3)	상대적 이점	AI 기반 공공·행정 서비스는 기존 행정서비스에 비해 더 좋 을 것 같다.	.737
		AI 기반 공공·행정 서비스는 기존 행정서비스에 비해 더 유 용할 것 같다.	
	복잡성	AI 기반 공공·행정서비스를 사용하는 과정은 복잡할 것 같다.	.829
		AI 기반 공공·행정서비스는 기능이 복잡할 것 같다.	
정부 신뢰 (x4)	공정성	정부는 공정한 행정서비스를 제공한다고 생각한다.	.747
		정부는 특정 집단의 이익이 아닌 다수의 이익을 추구한다고 생각한다.	
	가치 유사성	정부는 내가 생각하는 의견과 비슷한 수준의 정책/서비스를 제공할 것 같다.	.790
		정부는 내가 생각하는 아이디어와 유사한 문제 해결 방안을 제시할 것 같다.	
매개 (me)	혁신 저항	나는 우리 사회에 AI 기반 공공·행정서비스를 도입하는 것에 대해 반대한다.	.833
		나는 우리 사회에 AI 기반 공공·행정서비스가 도입되는 것이 우려스럽다.	
조절 (mo1)	사회적 지지	AI 기반 공공·행정서비스를 이용할 때 주변 사람들로부터 도 움을 받을 수 있을 것 같다.	.742
		AI 기반 공공·행정서비스 사용 중 어려운 문제가 발생하면 주변에 나를 도와줄 사람이 있다.	
조절 (mo2)	휴리스틱 판단	나는 AI 기반 공공·행정서비스와 관련된 이슈들을 느낌에 의 존해 판단하는 편이다.	.664
		나는 AI 기반 공공·행정서비스와 관련된 이슈들을 직관적으 로 판단하는 편이다.	
종속 (y1)	사회적 수용성	우리 사회는 AI 기반 공공·행정서비스를 적극적으로 도입할 필요가 있다.	.764
		우리 사회는 AI 기반 공공·행정서비스를 다양한 분야에 활용 해야 한다.	
종속 (y2)	마지못한 수용성	나는 AI 기반 공공·행정서비스를 어쩔 수 없이 이용할 수밖에 없을 것 같다.	.655
		나는 내 의사와 관계없이 AI 기반 공공·행정서비스를 마지못 해 이용하게 될 것 같다.	

자료: 연구진 작성

4) 자료수집 및 분석 방법

자료수집은 AI 기반 공공·행정서비스 이용 경험자 및 비경험자 전체를 모집단으로 하여 전문 조사 회사의 패널을 활용한 온라인 설문조사를 진행하였다. 표본추출 방법은 연령, 성별, 지역별 할당 방식으로 2,000명을 추출하고, 본조사가 진행되기 전 사전 조사를 100명 내외로 진행하여 문항의 신뢰도 및 타당도를 확보한 문항으로 본 조사를 진행하였다. 본 조사는 2025년 9월 19일부터 10월 2일까지(14일간) 진행되었으며, 수집된 표본의 특성은 다음의 <표 3 - 24>와 같다.

<표 3 - 24> 표본 특성

항목		n	%	항목		n %
성별	남자	(988)	49.4	직업	관리자	(133) 6.7
	여자	(1012)	50.6		전문가 및 관련 종사자	(146) 7.3
연령	만 19~29세	(284)	14.2		사무 종사자	(588) 29.4
	만 30~39세	(302)	15.1		서비스 종사자	(152) 7.6
	만 40~49세	(348)	17.4		판매 종사자	(79) 4.0
	만 50~59세	(396)	19.8		농림어업 숙련 종사자	(23) 1.2
	만 60~69세	(356)	17.8		가능원 및 관련 기능 종사자	(44) 2.2
	만 70세 이상	(314)	15.7		장차기계 조작 및 조립 종사자	(32) 1.6
거주 지역	서울	(372)	18.6		단순 노무 종사자	(69) 3.5
	인천/경기	(647)	32.4		군인(직업 군인)	(2) 0.1
	대전/충청/세종	(218)	10.9		학생	(76) 3.8
	광주/전라	(191)	9.6		전업주부	(298) 14.9
	대구/경북	(192)	9.6		취업준비 중	(59) 3.0
	부산/울산/경남	(295)	14.8		은퇴/취업의사 없음	(129) 6.5
사회 경제적 지위	강원/제주	(85)	4.3		무직	(118) 5.9
	낮음	(390)	19.5		기타	(52) 2.6
	중간	(1362)	68.1	월평균 소득	300만원 미만	(419) 21.0
최종 학력	높음	(248)	12.4		300만~400만 원 미만	(355) 17.8
	중학교 이하	(38)	1.9		400만~500만 원 미만	(262) 13.1
	고등학교	(463)	23.2		500만~700만 원 미만	(479) 24.0
	대학(4년제 미만)	(259)	13.0		700만~1,000만 원 미만	(313) 15.7
	대학교(4년제 이상)	(1057)	52.9		1,000만 원 이상	(172) 8.6
	대학원(석사과정)	(146)	7.3			
	대학원(박사과정)	(37)	1.9			

자료: 연구진 작성

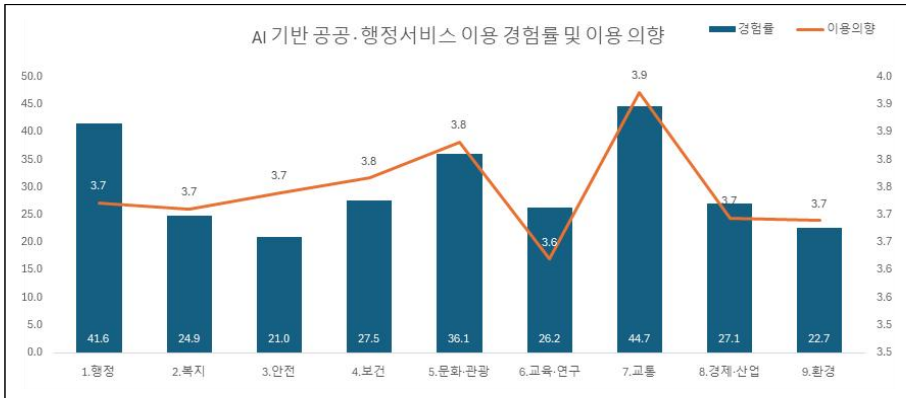
자료 분석 방법은 수집된 표본의 현황을 파악하기 위한 기초분석과 주요 변인들의 기술통계 분석, AI 기반 공공·행정서비스 관련 이슈들의 IPA(Importance Performance Analysis)⁵²⁾를 진행하였다. 또한 수용성 결정요인을 확인하기 위한 회귀분석(조절된 매개) 등을 진행하였다.

5) 분석 결과

① 기초분석

AI 기반 공공·행정서비스 이용 경험 유무를 분야별로 확인해 본 결과, ‘교통’ 분야의 경험률이 가장 높은 편(44.7%)이며, 차순으로 ‘행정’ 분야의 경험률도 41.6%로 나타났다. ‘안전’ 분야는 21.0%가 경험이 있다고 응답하여 가장 낮은 경험률을 보였으며, ‘환경’ 분야 역시 22.7%만이 경험이 있다고 응답하였다. 다음으로 이용 의향이 가장 높은 분야 역시 ‘교통’ 분야(3.9점/5점 만점)였고, 차순으로 ‘문화·관광’ 분야(3.8점/5점 만점)의 서비스를 이용할 의향이 높은 것으로 확

[그림 3-21] AI 기반 공공·행정서비스 이용 경험률 및 이용 의향



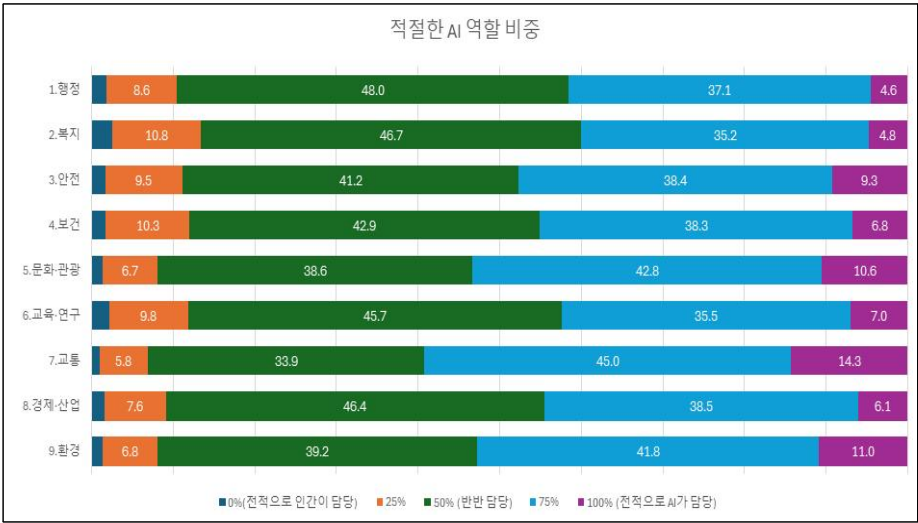
자료: 연구진 작성

52) 중요도와 만족도(성취도) 데이터를 바탕으로 이를 좌표로 표시하여 4분면에 나타난 위치에 따라 차별적 의미를 부여하는 방법으로, 본 연구에서는 우려 수준 및 우선순위(중요도)와 정부의 준비 수준을 기준으로 분석을 진행함.

인되었다. 모든 분야의 이용 의향이 보통(3점) 이상을 상회하고 있어, 응답자들은 전반적으로 AI 기반 공공·행정서비스를 이용하고자 하는 경향이 있는 것으로 판단된다.

다음으로 AI 기반 공공·행정서비스 분야별 적절한 AI의 역할 비중에 대한 응답을 살펴보면, 인간과 AI가 반반 담당해야 한다는 의견과 AI가 75%(인간 25%) 정도를 담당해야 한다는 의견이 지배적인 편이다. 특기할 점은 전적으로 인간이 담당해야 한다는 의견은 극소수에 불과하며, ‘교통’ 분야는 전적으로 AI가 담당해야 한다는 의견이 14.3%로 상당히 높은 비중을 차지하였다. 한편 ‘복지’ 분야는 상대적으로 전적으로 인간이 담당해야 한다는 의견(2.6%)과 AI가 25%(인간 75%) 정도 담당해야 한다는 의견(10.8%)이 높은 편이었다. AI의 역할 비중에 다소 보수적인 분야는 ‘복지’와 ‘교육·연구’로 나타났으며, 적극 활용해야 한다고 생각하는 분야는 ‘교통’과 ‘문화·관광’임을 확인할 수 있었다.

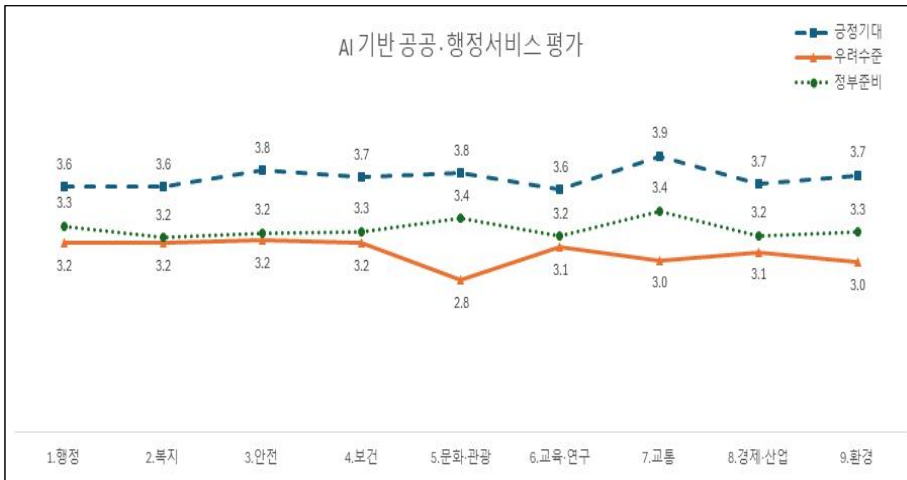
[그림 3-22] 분야별 적절한 AI 역할 비중



자료: 연구진 작성

AI 기반 공공·행정서비스 분야별 긍정 기대 수준과 우려 수준, 그리고 정부의 준비 수준에 대해 살펴보면, 전반적으로 긍정적 효과에 대한 기대 수준이 높은 편이었고, 상대적으로 우려 수준은 낮은 것으로 확인되었다. 우려 수준에 비해 정부의 준비 수준은 다소 높은 편으로 응답자들은 AI 기반 공공·행정서비스로 인해 발생할 것으로 예상되는 문제에 대해 정부가 상대적으로 준비를 잘하고 있다고 평가하였다. 가장 긍정적 효과의 기대 수준이 높은 분야는 앞선 결과들과 일맥상통하게 ‘교통’ 분야(3.9점/5점 만점)로 나타났으며, 차순으로 ‘문화·관광(3.8점)’, ‘안전(3.8점)’ 분야에 대한 기대가 높은 편이었다. 가장 우려 수준이 높은 분야는 ‘행정’, ‘복지’, ‘안전’, ‘보건’ 분야(3.2점/5점 만점)였으며, 정부의 준비가 가장 미흡하다고 평가받은 분야는 ‘복지’, ‘안전’, ‘교육·연구’, ‘경제·산업’ 분야(3.2점/5점 만점)로 확인되었다. ‘문화·관광’과 ‘교통’, ‘환경’ 분야는 상대적으로 우려 수준이 낮으며, 정부의 준비 수준은 높은 분야임을 알 수 있다.

[그림 3-23] AI 기반 공공·행정서비스 평가

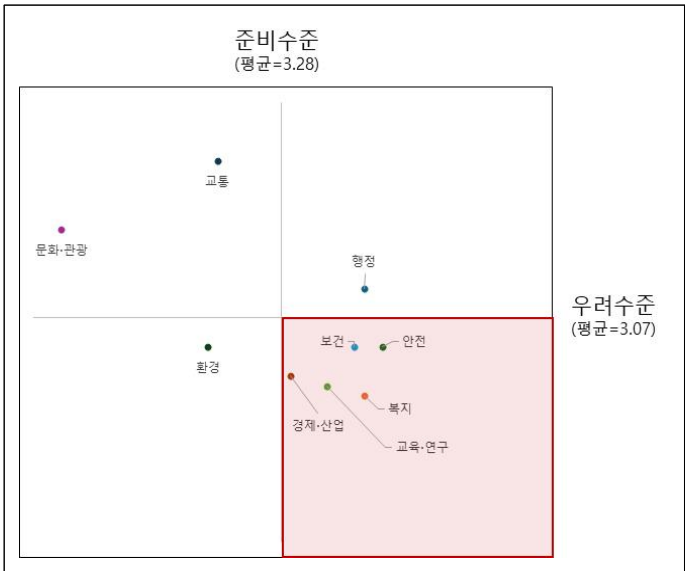


자료: 연구진 작성

AI 기반 공공·행정서비스에 대한 우려 수준과 정부의 준비 수준을 활용하여 IPA를 진행하였다. 모든 분야의 우려 수준 전체 평균은 3.07점이었고, 정부의 준

비 수준은 3.28점으로 확인되었다. 각 평균을 중심으로 사분면에 각 분야의 점수를 위치시킨 결과는 [그림 3-24]와 같다. 상대적으로 우려 수준이 높는데, 정부의 준비 수준은 평균보다 낮아 미흡하다고 판단되는 분야는 ‘보건’, ‘안전’, ‘경제·산업’, ‘교육·연구’, ‘복지’ 분야로 확인되었다. ‘행정’ 분야는 우려 수준은 다소 높지만, 정부가 준비를 잘하고 있다고 평가되었으며, ‘교통’과 ‘문화·관광’ 분야는 우려 수준이 낮고, 정부의 준비 수준이 높은 분야로 확인되었다.

[그림 3-24] AI 기반 공공·행정서비스 분야 우려*정부 준비 수준 IPA

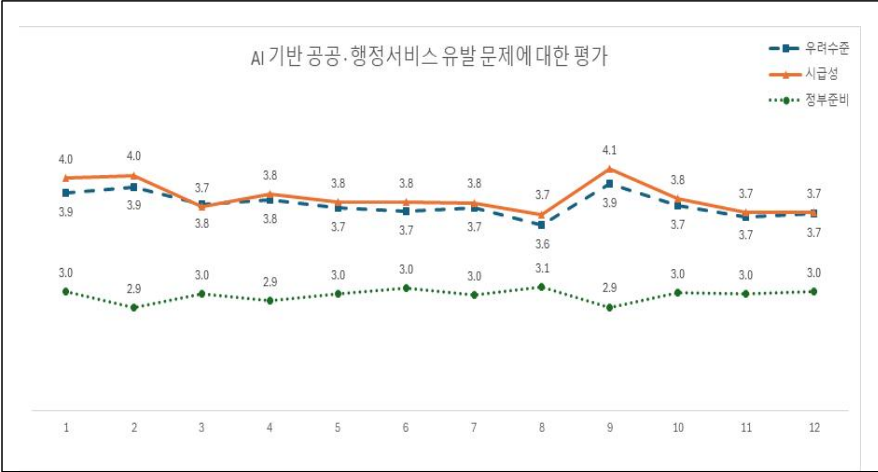


자료: 연구진 작성

다음으로 AI 기반 공공·행정서비스가 유발할 것으로 예상되는 문제들에 대한 우려 수준과 시급성, 그리고 정부의 준비 수준을 측정하였다. 전반적으로 우려 수준과 시급성은 매우 유사한 응답 경향을 보이고 있으며, 이를 통해 응답자들은 시급한 문제를 우려하는 것으로 해석할 수 있다. 응답자들이 가장 우려하면서도 시급하다고 생각하는 문제는 ‘지능형 신종 범죄 피해 발생’으로 확인되었으며, 그에 비해 정부의 준비 수준은 미흡(2.9점/5점 만점)한 것으로 평가하였다. 차순으로

‘개인정보의 오남용과 유출로 인한 문제’, ‘과도한 개인정보 수집으로 인한 사생활 침해 문제’ 등 응답자들은 개인정보와 관련된 문제를 가장 우려하는 것으로 나타났다.

[그림 3-25] AI 기반 공공·행정서비스 유발 문제에 대한 평가

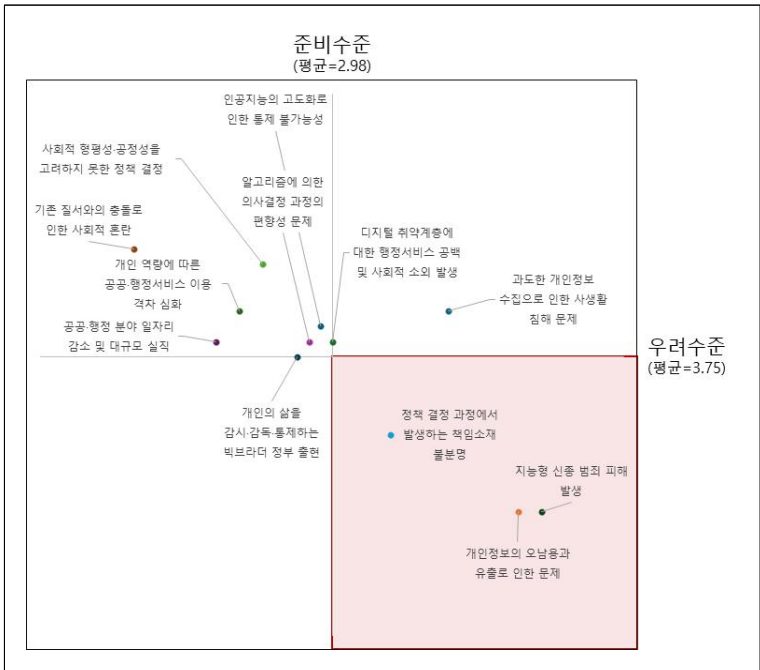


1	과도한 개인정보 수집으로 인한 사생활 침해 문제	7	개인의 삶을 감시·감독·통제하는 빅브라더 정부 출현
2	개인정보의 오남용과 유출로 인한 문제	8	기존 질서와의 충돌로 인한 사회적 혼란
3	디지털 취약계층에 대한 행정서비스 공백 및 사회적 소외 발생	9	지능형 신종 범죄 피해 발생
4	정책 결정 과정에서 발생하는 책임소재 불분명	10	인공지능의 고도화(초인공지능)로 인한 통제 불가능성
5	알고리즘에 의한 의사결정 과정의 편향성 문제	11	공공·행정 분야 일자리 감소 및 대규모 실직
6	사회적 형평성·공정성을 고려하지 못한 정책 결정	12	개인 역량에 따른 공공·행정서비스 이용 격차 심화

자료: 연구진 작성

AI 기반 공공·행정서비스로 인해 발생할 문제에 대한 우려 수준과 정부의 준비 수준으로 IPA를 진행한 결과, 상대적으로 시급히 정부의 대응이 필요한 문제는 ‘지능형 신종 범죄 피해 발생’, ‘개인정보의 오남용과 유출로 인한 문제’, 그리고 ‘정책 결정 과정에서 발생하는 책임소재 불분명’인 것으로 확인되었다. 전반적으로 모든 문제에 대해 정부의 준비 수준은 상대적으로 낮게 평가되고 있으나, 해당 문제들은 특히 우선순위에 두어야 할 것으로 판단된다.

[그림 3-26] AI 기반 공공·행정서비스 발생 문제 우려*정부 준비 수준 IPA



자료: 연구진 작성

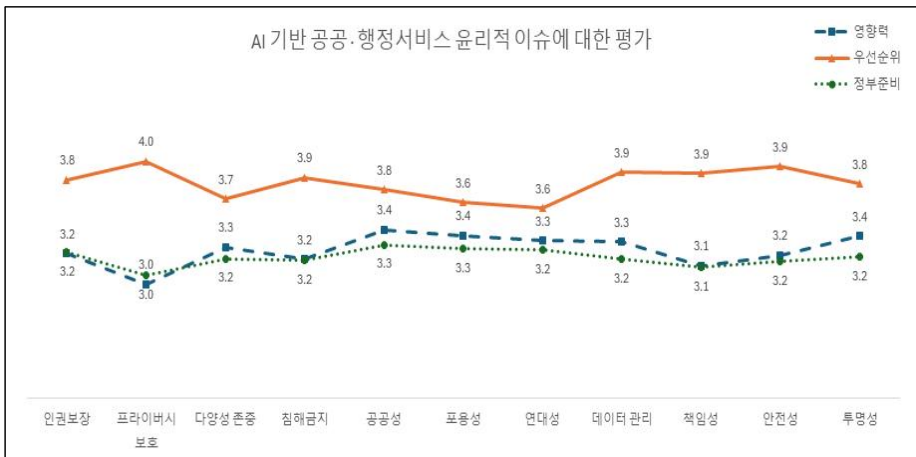
다음으로 AI 기반 공공·행정서비스가 촉발할 것으로 예상되는 윤리적 이슈들의 영향력(금/부정)⁵³⁾과 우선순위, 그리고 정부의 대응 수준에 대해 측정하였다. 전반

53) 영향력은 5점 척도로 측정되었으며, 점수가 높을수록 긍정적 영향력이 큰 것을 의미하며, 보통(3점)은 긍정적 영향력과 부정적 영향력이 비슷한 것을 의미함.

적으로 윤리적 차원에서의 영향력은 보통(3점) 수준을 약간 상회하여 부정적인 영향력보다는 긍정적인 영향력을 더 높게 인식하고 있는 것으로 판단된다. ‘프라이버시 보호’는 영향력 평균이 3점으로 긍정적 영향력과 부정적 영향력이 비슷하다는 의견으로 수렴하였다. 따라서 우선순위 역시 ‘프라이버시 보호’가 4.0점으로 가장 높은 것으로 확인되었다. 차순으로 ‘책임성’에 대한 영향력이 상대적으로 긍정적·부정적 영향력이 유사할 것으로 예상되었으며, 우선순위 역시 3.9점으로 높은 편이었다.

응답자들은 윤리적 차원에 대한 정부의 준비 수준에 대해 다소 만족하지 못하는 것으로 확인되었다. 가장 준비를 잘하고 있다고 평가받은 차원은 ‘공공성(3.4점)’과 ‘포용성(3.4점)’으로 확인되었으며, ‘프라이버시 보호(3.0점)’와 ‘책임성(3.1점)’은 상대적으로 정부의 준비 수준이 가장 낮은 차원이었다.

[그림 3-27] AI 기반 공공·행정서비스 윤리적 이슈에 대한 평가



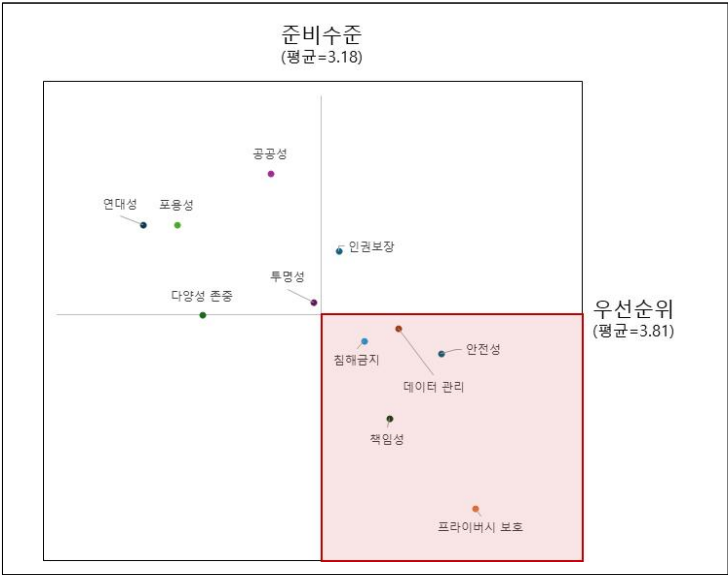
자료: 연구진 작성

윤리적 이슈의 우선순위와 정부의 준비 수준을 기준으로 IPA를 진행한 결과, ‘인권보장’ 차원은 우선순위가 높고, 정부의 준비 수준도 상대적으로 긍정적인 차원으로 확인된다. ‘연대성’, ‘포용성’, ‘다양성 존중’, ‘공공성’, ‘투명성’ 차원은 우

선순위에 비해 정부의 준비 수준이 상대적으로 높은 차원으로 나타났다.

우선순위가 높지만, 상대적으로 준비 수준이 미흡한 윤리적 차원은 ‘프라이버시 보호’, ‘안전성’, ‘데이터 관리’, ‘책임성’, ‘침해금지’로 도출되었다. 해당 차원의 윤리적 이슈들은 다른 차원에 비해 상대적으로 민감할 수 있어 주의가 필요할 것으로 예상된다.

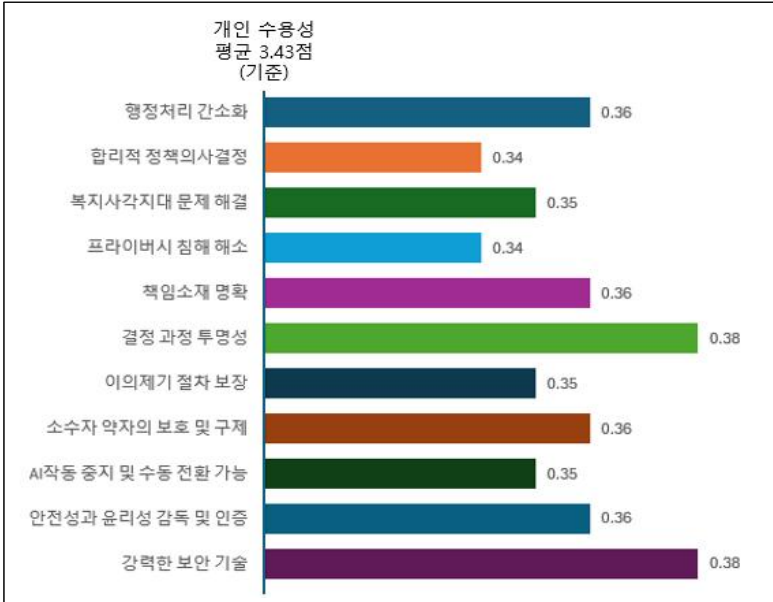
[그림 3-28] AI 기반 공공·행정서비스 윤리 우선순위*정부 준비 수준 IPA



자료: 연구진 작성

AI 기반 공공·행정서비스를 사용하는 것을 긍정적으로 생각한다는 개인적 수용 태도는 평균 3.43점으로 보통 수준을 상회하였다. 추가적으로 조건부 수용성을 측정하여 어떤 조건에서 개인들의 수용성이 증가하는지 확인하였다. 그 결과 AI 기반 공공·행정서비스의 결정 과정이 투명하게 공개된다는 조건과 강력한 보안 기술로 개인정보가 안전하게 보호된다는 확신이 있을 때, 수용성이 가장 큰 폭 (0.38점 증가)으로 증가하였다.

[그림 3-29] AI 기반 공공·행정서비스 개인 수용성 및 조건부 수용성



자료: 연구진 작성

② AI 공공·행정서비스 수용성 회귀분석

본 연구에서는 AI 공공·행정서비스 수용성을 크게 두 가지로 구분하여 분석을 진행하였다. 하나는 ‘① 사회적 수용성’을, 그리고 다른 하나는 ‘② 마지못한 수용성’을 대상으로 분석하였다. 이와 같은 이유는 AI 공공·행정서비스가 새로운 기술 그 자체로서 우리 사회에서 수용되기 위한 조건을 탐색(사회적 수용성)하고자 함과 동시에, 전방위적으로 급속히 퍼져나가는 AI 기술로부터 회피할 수 없는 맥락적 특성을 반영한 수용성(마지못한 수용성)을 살펴보고자 했기 때문이다.

〈표 3-25〉는 사회적 수용성을 종속변수로 삼아 분석한 결과표이다. 우선 [모형 1]을 살펴보면, 개인 특성에서는 혁신성을 제외한 모든 변인들이 유의미한 결과로 나타나고 있다. 구체적으로 행위에 대한 신념과 내재적 동기, 외재적 동기는 혁신저항에 대하여 부적(-) 영향을 미치는 것으로 확인된다. 이는 AI 기반 공공·행정서비스에 대한 긍정적 신념이나 기대감이 혁신저항을 낮춘 것으로 해석할 수

있다. 반면 리터러시는 혁신저항과 정적(+) 영향으로 나타났는데, AI 기술을 잘 알고 있는 사람일수록 해당 기술의 사용에 대해 꺼려하는 태도가 증가하는 것으로 해석할 수 있다. 이러한 결과는 통상적인 위험(기술) 연구와 상반된다. 리터러시 또는 지식은 그 수준이 높아질수록 특정 기술이나 위험에 대하여 덜 민감하게 반응하는 것으로 연구되고 있다. 그러나 현시점의 AI 기술은 개인 간 관계에서는 물론 각종 미디어에서 그 부작용에 대하여 상당히 많이 다뤄지고 있다. 이와 같은 상황적 특성이 지식과 혁신저항 간 부적 관계를 형성한 것으로 예측된다.

〈표 3-25〉 사회적 수용성의 인과기제 분석 결과

개념	변인	모형1	모형2	조절된 매개 지수	조절된 매개지수의 Bias Corrected CI(95%)	
		X → M (a path)	M → Y (b path)		하한	상한
	cons	-.360	.706***	-	-	-
인구 사회학적 변인	성별(남→여)	.098**	-.023	.004	.000	.009
	연령	-.006***	.001	.000	.000	.000
	학력(저→고)	-.058	-.012	-.002	-.007	.001
	가구소득(1인당)	-.046	.027	-.002	-.005	.000
개인 특성	행위 신념	-.197***	.112***	-.008	-.016	-.001
	지각된 위험	.062**	.061***	.002	.000	.006
	내재적 동기	-.152***	.165***	-.006	-.013	-.001
	외재적 동기	-.142***	.104***	-.006	-.013	-.001
	혁신성	.055	.058**	.002	.000	.006
	주관적 지식	.127***	-.037	.005	.001	.011
커뮤니케이션 특성	정보원 신뢰	-.015	.054*	-.001	-.004	.002
	정보 질적 특성	.115**	-.003	.005	.000	.011
	뉴미디어 신뢰	.200***	-.049**	.008	.001	.016
기술적 특성	상대적 이점	-.213***	.278***	-.009	-.017	-.001
	복잡성	.470***	-.042*	.019	.002	.035

개념	변인	모형1	모형2	조절된 매개 지수	조절된 매개지수의 Bias Corrected CI(95%)		
		X → M (a path)	M → Y (b path)		하한	상한	
정부 신뢰	공정성 가치유사성	-.066*	.002	-.003	-.007	.000	
		.087**	.058**	.003	.000	.008	
매개(m)	혁신저항	-	-.065***	-	-	-	
조절(z)	사회적 지지	-	.028	-	-	-	
상호작용(mz)	매개*조절	-	.040*	조건효과(-1sd/mean/+1sd)			
				effect	t	s.e	
F value		71.523***	83.871***	저:	-.081***	-4.360	.019
R square		.388	.467	중:	-.061***	-3.768	.016
				고:	-.041*	-2.344	.018

주: 1) $p^* < .05$, $p^{**} < .01$, $p^{***} < .001$

2) bootstrap 10,000회

자료: 연구진 작성

이어서 커뮤니케이션 특성에서는 정보의 질적 특성(정보의 이해용이성 및 유용성)과 뉴미디어(유튜브, 인스타, 페이스북 등) 신뢰가 혁신저항에 대하여 정적(+) 영향을 미치는 것으로 확인되었다. 즉 개인에게 주어지는 정보의 질적 특성이 우수할수록, 그리고 뉴미디어에 의존할수록 혁신저항은 증가하는 것으로 나타난다. 이와 같은 결과 역시 앞서 언급한 바와 같이 전 세계적으로 언급되고 있는 AI의 잠재적 위험과 밀접한 관련이 있는 것으로 예측할 수 있다. AI로 인한 부정적 뉴스나 기사 등과 같은 정보, 그리고 이를 자주 접하게 되는 뉴미디어의 특성상 혁신저항과 밀접한 관련이 나타난 것으로 해석된다.

기술적 특성으로 반영한 상대적 이점과 복잡성 모두 유의미한 원인으로 확인되었다. AI 기술로 인한 혜택이 기존의 것에 비해 높다고 느낄수록 혁신저항은 감소하며, 반면 복잡하다고 지각할수록 혁신저항은 증가하는 것으로 확인된다. 이는 혁신확산이론에서 언급하는 내용과 일치하는 방향으로 나타나고 있다.

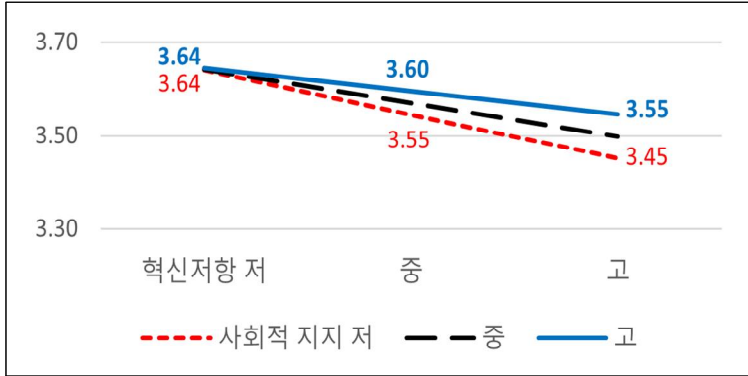
정부 신뢰를 측정하기 위하여 반영한 공정성, 그리고 가치 유사성은 모두 유의

미한 변인으로 확인되었다. 정부가 공정한 행정서비스를 제공한다고 지각할수록 혁신저항은 감소하는 것으로 확인된다. 반면 정부와 나의 의견이 비슷하다고 생각할수록 혁신저항은 증가하는 것으로 나타난다. 정부 및 공공기관에서 AI의 부작용에 관하여 주의해야 한다는 정보가 지속적으로 노출/홍보되고 있으며, 따라서 이에 동조하는 개인들은 혁신저항이 증가하는 것으로 이해할 수 있다.

이어서 매개변인인 혁신저항과 종속변인인 사회적 수용성 간 관계를 살펴보면, 강한 부적(-) 관계로 확인된다. 즉 혁신저항 태도가 증가함에 따라 사회적 수용성은 감소하게 된다. 그리고 이 과정에서 본 연구모형에서 제안한 사회적 지지의 조절효과가 발생하는 것으로 확인되었다([그림 3-30] 참조). 사회적 지지란 AI 기반 공공·행정서비스를 이용할 때 주변 사람으로부터 도움을 받을 수 있는가에 대한 지각 수준을 의미한다. 사회적 지지를 높게 지각하는 사람들일수록 혁신저항과 사회적 수용성 간의 부적 관계는 완만하게 감소하는 것으로 나타난다([그림 3-30]의 파란 실선). 즉 사회적 지지는 신기술의 도입으로 인해 필연적으로 발생하는 혁신저항의 부정적 효과(수용성 감소)를 완화하는 데 유용한 개념으로 이해할 수 있다.

한편 본 연구모형은 조절된 매개(moderated mediation model)이다. [모형1]과 [모형 2]가 모두 유의미한 경우 간접효과가 발생할 수 있으며, 이러한 간접효과가 조절변인(사회적 지지)에 의하여 그 크기가 변화되는지를 검증하는 모형이다. 이를 위해 조절된 매개 지수(index of moderated mediation)를 산출하였으며, 회귀분석표의 음영 처리된 부분은 조절된 매개 효과가 존재하는 경로이다. 예를 들어 행위 신념(행위에 대한 신념)의 조절된 매개 지수는 -.008로 확인되며, 부트스트랩 1만 번을 기준으로 하한값 -.016, 상한값 -.001로 이 사이값에 0이 존재하지 않는 것으로 확인된다. 따라서 사회적 지지라는 조절변인이 1단위 증가함에 따라 ‘행위 신념 → 혁신저항 → 사회적 수용성’의 간접효과 크기가 .008씩 유의미하게 감소하는 것으로 해석할 수 있다.

[그림 3-30] ‘혁신저항-사회적 수용성’ 간 관계에서 사회적 지지의 조건부 직접효과



자료: 연구진 작성

이어지는 다음의 회귀분석은 마지못한 수용성의 인과기제를 살펴보기 위한 분석이다. 독립변인 및 매개변인은 앞선 연구모형과 동일하기에 [모형3]의 분석 결과와 [모형1]의 결과는 동일하다.

본 회귀분석 결과표에서 초점을 두어 살피고자 하는 영역은 매개변인과 종속변인 간 관계에 있다. 매개변인인 혁신저항의 증가는 종속변인인 마지못한 수용성을 급격히 증가시키는 원인으로 확인되고 있다. 앞선 선행연구에서 언급했듯이 마지못한 수용성은 사람들로 하여금 현재 및 미래의 위험에 대하여 지나치게 낙관하게 만드는 요인이며, 그 결과 잠재적 위험에 대하여 과소평가하게 한다. 따라서 마지못한 수용성이라는 개념은 지극히 부정적인 것이며, 이를 완화할 수 있는 정책적 대응방안이 모색되어야 한다.

매개변인과 종속변인 간 정적(+) 관계에서 휴리스틱 변인은 유의미한 조건부 효과를 야기하는 것으로 확인되고 있다. 휴리스틱 판단에 의존하는 사람(고집단-파란색 실선)들은 혁신저항의 증가로 인해 발생하는 마지못한 수용성의 증가가 그렇지 않은 사람들에 비하여 상대적으로 가파르게 나타나고 있다. 휴리스틱이란 의사결정에 필요한 여러 이슈들을 느낌 또는 직관에 따라 결정하려는 태도를 의미한다. 즉 AI 기반 공공·행정서비스와 관련된 이슈에 대하여 적극적으로 정보를 찾아

보려는 노력을 하지 않고, 느낌이나 직관에 의존하는 사람들일수록 혁신저항의 증가에 따라 발생하는 마지못한 수용성의 증가폭이 더욱 크게 나타나는 것으로 이해할 수 있다. 이와 같은 발견은 정부의 리터러시 교육이 어떤 방향으로 설정되어야 하는가를 시사한다고 할 수 있다.

〈표 3-26〉 마지못한 수용성의 인과기제 분석 결과

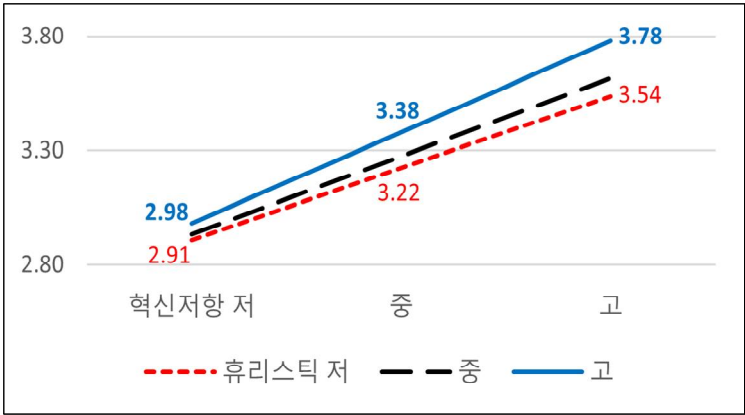
개념	변인	모형3	모형4	조절된 매개 지수	조절된 매개지수의 Bias Corrected CI(95%)	
		X → M (a path)	M → Y (b path)		하한	상한
	cons	-.360	2.125***	-	-	-
인구 사회학적 변인	성별(남 → 여)	.098**	.018	.006	.000	.014
	연령	-.006***	.001	.000	-.001	.000
	학력(저 → 고)	-.058	.029	-.003	-.010	.001
	가구소득(1인 당)	-.046	-.013	-.003	-.007	.000
개인 특성	행위 신념	-.197***	.004	-.013	-.028	-.001
	지각된 위험	.062**	.102***	.004	.000	.009
	내재적 동기	-.152***	.003	-.009	-.020	-.001
	외재적 동기	-.142***	.049	-.011	-.022	-.001
	혁신성	.055	-.054*	.002	-.001	.007
커뮤니케이션 특성	주관적 지식	.127***	.014	.007	.001	.015
	정보원 신뢰	-.015	.028	-.001	-.007	.003
	정보 질적 특성	.115**	-.082*	.006	.000	.014
기술적 특성	뉴미디어 신뢰	.200***	-.014	.011	.001	.023
	상대적 이점	-.213***	.071*	-.013	-.028	-.001
정부 신뢰	복잡성	.470***	.106***	.027	.004	.052
	공정성	-.066*	.035	-.004	-.010	.000
매개(m)	가치유사성	.087**	.038	.005	.000	.012
	혁신저항		.358***	-	-	-

개념	변인	모형3	모형4	조절된 매개 지수	조절된 매개지수의 Bias Corrected CI(95%)		
		X → M (a path)	M → Y (b path)		하한	상한	
조절(z)	휴리스틱 판단		.093**	-	-	-	
상호작용 (mz)	매개*조절	-	.057**	조건효과(-1sd/mean/+1sd)			
				effect	t	s.e	
F value		71.523***	42.007***	저:	.313***	11.450	.027
R square		.388	.305	중:	.342***	15.495	.022
				고:	.399***	17.242	.023

주: p* $\leq$.05, p** $\leq$.01, p*** $\leq$.001

자료: 연구진 작성

[그림 3-31] ‘혁신저항과 마지못한 수용성’ 간 관계에서 휴리스틱의 조건부 직접효과 검증



자료: 연구진 작성

3. 식별된 사회적 의제안

가. (공공행정①) 접근성 격차와 알고리즘 편향 관점에서 디지털 취약계층은 소외되지 않을 수 있는가?

1) 제안의 배경

공공·행정서비스는 국민의 기본권으로서 ‘보편적 접근성’을 핵심 가치로 한다.

보편적인 접근은 크게 두 가지 차원에서 논의될 수 있다. 첫째, 접근 배제에 대한 것으로 무인 키오스크나 모바일 앱 등의 활용이 어려운 취약계층이 소외되는 현상이다. 둘째는 AI 알고리즘 자체가 취약계층의 데이터를 편향되게 학습하거나 불공정한 기준을 내재화하여 복지 수혜나 공공 혜택의 대상에서 원천적으로 배제될 수 있는 문제가 있다.

2) 현황 및 쟁점

2024년 디지털 정보격차 실태조사에 따르면, 취약계층 중 고령층의 역량 점수는 55.9점(전체 65.6점)으로 다른 취약계층에 비해서도 한참 낮은 편이다. 단순 기기 및 인터넷 접근성은 점차 높아지고 있지만, 기기 활용 역량에서는 아직도 심각한 격차를 보인다. 예를 들어, 최근 상생페이백⁵⁴⁾은 본인 인증 후 몇 가지 정보만 입력하면 누구나 쉽게 신청할 수 있다고 하지만, 고령층은 스스로 본인 인증을 하기 쉽지 않다. 게다가 ‘디지털 온누리상품권’을 받더라도 어디에서 어떻게 사용해야 하는지도 알기가 어렵다. 즉 디지털 활용 능력이 없는 취약계층은 정책적 혜택에서조차 소외될 가능성이 높다. 국내에서는 2018년에서 2022년경 주요 공공기관에서 채용 과정의 공정성을 높인다는 명분으로 AI 역량검사(AI 면접)를 도입했다가 논란이 되었는데, 알고리즘 편향이 취업에서 특정 계층에 불리하게 작용한다는 것은 2014년 아마존 사례를 통해 이미 확인된 바 있다. 이처럼 공공·행정서비스 영역에서도 알고리즘의 편향으로 인해 특정 계층이 소외될 가능성이 있다.

3) 문제의 심각성

접근 격차는 눈에 보이는 불편과 소외를 야기하고, 알고리즘 편향은 눈에 보이지 않는 차별을 시스템화하고 영구화한다는 점에서 문제가 된다. 첫째, 디지털 취약계층은 물리적 접근성과 역량적 접근성 모두에서 격차를 겪는데, 필수적인 정보를 얻거나 공공서비스를 이용하는 데 심각한 제약을 받게 되면, 시민으로서의 기본 권리 및 기회 박탈로 이어지게 된다. 둘째, AI 모델은 과거의 행정 데이터를

54) 2024년 월평균 카드 소비액보다 2025년 9~11월 늘어난 월별 카드 소비액의 20%를 디지털 온누리상품권으로 돌려주어 소비진작 및 소상공인을 지원하는 사업.

학습하게 되는데, 이 데이터가 만약 특정 취약계층을 과소 대표하고 있다면, AI는 이들을 ‘정상 범주’에서 벗어난 극단치로 판단하거나 ‘신뢰할 수 없는’ 대상으로 학습할 수 있다. 그러면 복지 수급자 선정 과정에서 문제가 생기거나, 잠재적 ‘위험 가구’로 과도하게 분류되어 불필요한 행정적 감시의 대상이 될 수 있다.

4) 기존 논의/대책의 한계

디지털 격차를 해소하기 위한 기존 정책과 논의는 디지털 기기 보급 및 이용 교육 등 기술적 접근성 향상에 초점을 맞추고 있다. 따라서 디지털 리터러시 교육이나 공용 기기 보급 등이 이루어지고 있으나, 기술 발전 속도를 따라잡기는 어렵다. AI 서비스 설계 단계부터 소외계층의 접근성을 고려하는 ‘유니버설 디자인’ 원칙이 어떻게 적용될 수 있을지 지속적인 논의가 필요하다.

공공부문에서의 알고리즘 편향은 앞서 설명한 바와 같이 심각한 결과를 초래할 수 있다. 알고리즘 공정성, 투명성, 책임성 등을 확보하기 위한 논의와 규제가 진행되고 있으나, 해당 개념들을 정의하는 기준이 다양하고, 어떤 기준을 적용할지에 대한 사회적 합의가 제대로 이루어지지 않았다. 따라서 AI가 특정 집단(특히 취약계층)에 미칠 차별적 영향을 평가하는 ‘알고리즘 영향평가(AIA)’를 법적으로 의무화하는 제도가 필요하다⁵⁵⁾.

나. (공공행정②) AI 기반 공공·행정서비스에 오류가 발생하여 시민이 피해를 입었을 때, 책임 공백은 해소될 수 있는가?

1) 제안의 배경

공공·행정 서비스 영역에 AI가 도입되면 효율성과 신속성이 증대되는 만큼, 오작동, 데이터 편향 혹은 예측 불가능한 행동으로 시민이 금전적·비금전적 피해를 입을 가능성도 커진다. 딥러닝 기반의 AI는 스스로 학습하고 진화하며, 개발자조차 의사결정 과정을 완벽히 추적하기 어려운 ‘블랙박스(Black Box)’의 특성을 지닌다. 따라서 만약 AI 기반 공공·행정서비스 시스템(예: 복지 수급자 선정, 재난

55) <https://v.daum.net/v/20250924201502254>.

예측 등)이 치명적인 오류를 범해 시민에게 피해를 입히게 된다면, 이 책임이 누구에게 귀속되는지 법적, 윤리적으로 규명하기 매우 어렵다. AI가 개입하는 행정 결정은 시민의 권리와 의무에 직접적 영향을 미치기 때문에 오류가 발생하면 단순한 불편을 넘어 시민의 기본권 침해로 직결된다.

2) 현황 및 쟁점

본 조사에서 AI 기반 공공·행정서비스로 인해 발생할 것으로 예상되는 문제들 중 ‘정책 결정 과정에서 발생하는 책임소재 불분명’은 우려 수준 대비 정부의 준비 수준이 낮다고 평가받은 항목이다. 2024년에 온라인 민원 서비스인 ‘정부24’에서 다른 사람의 인적 사항이 뜨는 오류로 민감한 개인정보 1천여 건이 유출되었는데, 이 오류는 개발자의 프로그램 개발상 실수로 확인되었다⁵⁶⁾. 이러한 오류가 AI 기반 시스템에서 발생하게 된다면, 그때는 책임 소재를 명확히 하기 어렵고, 이해관계자가 모두 책임을 회피할 가능성도 배제할 수 없다. 또한 「행정기본법」 제20조(자동적 처분)에 따르면, 재량이 없는 처분에 한해 완전히 자동화된 시스템(AI를 적용한 시스템 포함)으로 처분을 할 수 있도록 근거를 마련하였는데, 오류 발생 시 피해 배상의 책임 주체와 절차에 대해서는 명확한 가이드라인이 제시되지 않아 한계를 갖는다.

3) 문제의 심각성

AI 기반 공공·행정서비스에 오류가 발생했을 때, 책임 주체는 알고리즘을 설계한 개발사, AI를 학습시킨 데이터를 제공한 기관, AI를 도입하기로 결정한 정부 부처, AI 판단을 검토 및 승인한 공무원, AI 시스템 자체 등 복잡하게 얽혀 있다. 만약 AI 기반 응급의료 시스템이 환자의 중증도를 잘못 분류해 구급차 배차를 지연시켜 환자가 사망했다면 누구에게 책임을 물어야 할지 판단하기 어렵다. 이 같은 치명적인 피해가 발생할 가능성을 내포한다는 것은 AI 기반 공공·행정서비스에 대한 불신으로 이어지며, AI 도입 자체를 꺼리거나(소극 행정), 반대로 모든 책

56) https://www.ytn.co.kr/_ln/0115_202405052157092595.

임을 AI 탓으로 돌리는 무책임한 행정으로 이어질 수 있다.

4) 기존 논의/대책의 한계

AI 기반 공공행정서비스에서 오류가 발생해 시민이 피해를 입을 경우를 대비해 정부는 책임소재를 명확히 하려는 다양한 대책을 제시해 왔다. 그러나 이러한 기존 대책은 몇 가지 중요한 한계를 갖는다. 첫째, 현재의 법·제도는 공무원, 시스템 개발자, 운영 주체 등 복수 행위자가 관여하는 AI 특성을 충분히 반영하지 못해 책임 주체가 모호해지는 문제가 지속된다. 둘째, 지난 2020년 12월 발표되었던 ‘국가 인공지능 윤리기준’ 등을 포함하여 정부 부처별로 분야별 특성에 맞는 알고리즘 투명성 확보나 설명 가능성 강화와 같은 원칙이 제시되었지만 실제 기술적 구현 의무가 다소 약해 피해 원인을 규명하기가 어려운 측면이 있다. 셋째, 사전 예방 중심의 관리체계가 강조되지만, 피해 발생 이후 신속한 구제 절차나 국가의 직접적 책임 인정 기준은 여전히 불명확하다. 마지막으로, 공공부문 특유의 외주·위탁 구조가 복잡해지면서 책임 분배를 위한 실질적 가이드라인이 미흡해 ‘책임 공백’이 구조적으로 반복될 가능성이 남아있다.

다. (공공행정③) ‘시민 안전’과 ‘범죄 예방’을 명분으로 개인의 일거수일투족을 감시하고 통제하는 ‘빅브라더’ 사회로의 귀결을 막을 수 있는가?

1) 제안의 배경

시민 안전, 범죄 예방, 효율적 행정 등의 공익적 명분 아래, 공공 CCTV 인프라와 AI 기술(안면 인식 및 행동 패턴 분석 기술 등)이 결합되면서 대규모·실시간·자동화된 감시가 가능해지고 있다. 2024년 기준 우리나라의 공공기관 CCTV는 195만 7,790대로 매년 전년 대비 약 10%씩이 증가하고 있다. 공식적으로 집계되지 않는 민간 CCTV, 자동차 블랙박스 등을 합치면 사실상 언제 어디서든 카메라에 찍히고 있다고 해도 과언이 아니다. 지능형 CCTV가 지속적으로 확산되고 대부분의 시민이 그 영향하에 있게 될 가능성이 농후한 만큼, 본 사회적 의제에 대한 논의가 시급하다.

2) 현황 및 쟁점

‘지능형 선별관제 시스템’은 AI가 실시간으로 ‘이상 상황’을 자동 감지하여 관제 요원에게 경고를 보내는 기술이다. 서울시는 2023년부터 AI 기능이 없는 일반 CCTV를 지능화하고 있으며, ‘AI 실종자 고속검색시스템’을 구축하는 등 전국 지자체에서 대규모 예산을 투입해 지능형 선별 관제 시스템을 도입·확대하고 있다. 시민 안전이라는 ‘공익’이 프라이버시라는 ‘기본권’보다 우선시되면서, 도입의 필요성 및 효과에 대한 사전 평가가 충분히 이루어지지 않은 채 행정 당국의 판단에 의해 일방적으로 결정되는 경우가 많다. 일례로 2025년 10월 국회 과학기술정보방송통신위원회 국정감사에서 윤 정부 당시 대통령경호처와 과학기술정보통신부가 공동 추진한 ‘군중 감시 인공지능(AI) 기술⁵⁷⁾’이 국민의 생체정보로 불특정 다수의 자유권을 침해할 수 있다는 우려가 제기되었다.

3) 문제의 심각성

경기도 부천시에서는 ‘군중 안전 솔루션’을 구축해 일정 면적 내 인구밀집도가 위험 수준에 도달하면 관계기관에 통보하고, 현장 안내방송 및 재난안전 문자를 발송하여 군집 사고를 선제적으로 예방하여 2024년 경기도 ‘CCTV 운영 우수사례 발표대회’에서 우수상을 수상했다. 현재는 시민의 안전을 명분으로 확산되고 있으나, 이는 CCTV에 포착되는 모든 이들의 정보를 실시간으로 수집하고 미래 행동을 예측할 수 있다는 점에서 논란이 될 수 있다. 원래의 목적을 넘어 특정 개인의 동선 추적이나 집회 참여자 식별에 사용되는 등 판옵티콘처럼 개인의 일거수일투족을 감시하고 통제하는 데 사용될 수 있다. 그렇게 되면 시민들은 자유로운 의사 표현, 집회 및 결사의 자유 등 민주주의의 근간이 되는 기본권 행사를 스스로 검열하고 포기할 수 있다. 더욱이 첨단 감시 기술의 복잡성은 일반 시민 및 시민사회 단체가 기술 도입 및 현황을 정확히 파악하고 비판 및 대응할 수 있는 역량을 약화할 수 있다.

57) 군중의 행동 패턴과 개인 생체신호를 분석해 ‘위험 인물’을 선별하는 기술.

4) 기존 논의/대책의 한계

‘빅브라더’ 사회로의 귀결 위험은 현행 법제가 AI 기반 감시 기술의 복잡성과 파급력을 효과적으로 통제하지 못하는 구조적 한계에서 비롯된다. 구체적으로 ‘감시’ 행위에 대한 명확한 법적 정의 및 규제 범위가 확립되지 않아, 정부나 지방자치단체의 도입 및 활용을 효과적으로 통제하기가 어렵다. 또한 「개인정보보호법」은 주로 유출 후 사후 처벌 및 배상에 초점을 맞추고 있어 ‘정보 수집 및 활용’ 그 자체로 인한 기본권 침해를 다루기에는 어려움이 있다. 전국 지자체 실무 현장에서는 개인정보 보호 대응에 어려움을 겪고 있으며, 이러한 어려움은 상대적으로 조직 및 예산이 작은 지자체일수록 가중된다. 특히 EU는 「AI 법」을 통해 ‘실시간 원격 생체 인식(안면 인식 등)’을 원칙적으로 금지(예외 조항 존재)하고 있지만, 우리나라는 국가인권위원회의 ‘원칙적 금지 권고’에도 불구하고 이를 규율할 명확한 상위법이 부재한 상황이다.

라. (공공행정④) 기술적 안정성 및 신뢰 차원에서 개인정보 보호(해킹 취약, 신종 범죄에 악용 등)는 가능한가?

1) 제안의 배경

AI 기반 공공·행정서비스를 제공하기 위해서는 방대한 양의 국민 개인정보(예: 주민등록, 의료, 소득 정보 등)를 처리하게 된다. 따라서 데이터 수집 및 학습 단계에서부터 심각한 프라이버시 침해 위험이 있다. 최근 민간 기업에서도 고객들의 개인정보가 대량으로 유출되면서 심각한 문제가 된 바 있다. 더욱이 공공 AI 시스템의 해킹은 개인정보의 유출을 넘어, 국가 행정망의 마비와 데이터의 신뢰성 붕괴를 의미한다. 예를 들어, 복지 수급 데이터가 조작되거나, AI 교통 시스템이 사고를 유발하도록 해킹된다면 그 피해는 돌이키기 어렵다.

2) 현황 및 쟁점

2024년 개인정보 유출 신고 원인의 과반이 해킹인 것으로 조사되었으며, 공공기관과 기업의 관리자 페이지를 겨냥한 해킹이 특히 많았다. 개인정보보호위원회

와 한국인터넷진흥원이 발표한 ‘2024년 개인정보 유출 신고 동향 및 예방 방법’에 따르면, 양 기관으로부터 접수된 유출 신고 307건 가운데 56%(171건)가 해킹이 원인이었다. 공공기관 유출 신고는 2023년 41건에서 2024년 104건으로 두 배 넘게 증가했다⁵⁸⁾. 서울시는 2024년 ‘AI 기반 사이버보안 종합계획’을 발표하고, 3년간 총 339억 원을 투입해 사이버 위협에도 시민 생활과 직결되는 행정서비스를 중단없이 제공할 것이라고 밝혔다. 그러나 AI를 활용하여 공격 대상 시스템의 취약점을 분석하고, 효율적인 공격 시나리오를 생성해 해킹에 소요되는 시간과 비용을 절감하는 등 AI 간의 싸움이 될 것으로 예측된다.

3) 문제의 심각성

AI 시스템은 데이터, 모델, 알고리즘, 인프라 등 다양한 구성 요소로 이루어져 있어, 기존 IT 시스템보다 잠재적인 해킹 경로(취약점)가 더 넓다. AI 모델 자체가 학습 데이터를 암묵적으로 기억하고 있어, 해커가 모델을 분석하거나 특정 입력을 반복함으로써 모델이 학습한 원본 개인정보를 추출해낼 위험이 존재한다. 만약 사법 시스템에서 피의자/피해자 자필 진술서, 범죄 기록, 수사 기록 등 고도의 민감 정보가 유출된다면, 사법 체계 전반의 기밀성과 신뢰성이 훼손된다. 피해자는 보복 위험에 노출되고, 증인이나 제보자가 신변 위협을 느껴 사법 정의 실현이 저해될 수 있다. 해킹 조직이 핵심 행정망을 교란하거나, 국가 기밀을 유출하는 등 국가 안보를 직접적으로 위협하는 도구로 악용될 수 있다. 또한 모든 데이터 활용 목적에 대한 완전한 이해와 동의를 구하는 것이 현실적으로 불가능하고, AI 개발 기업들이 모델 내부 작동 방식을 투명하게 공개하지 않는다면, 보안 취약점 점검 자체가 어렵다.

4) 기존 논의/대책의 한계

AI 기반 공공·행정서비스의 기술적 안정성과 개인정보 보호를 위한 노력이 계속되고 있으나, 급변하는 AI 위협 환경에 대응하기에는 한계가 많다. AI 시대 프

58) <https://www.yna.co.kr/view/AKR20250320066500530>.

라이버시 위협은 단순히 개인정보 유출과 같은 수동적 피해에 그치지 않고, 알고리즘에 의한 예측, 추론, 차별 등 개인이 통제할 수 없는 능동적 피해로 변화하고 있다. 따라서 데이터 처리의 투명성 및 동의에 초점을 맞추고 있는 현재 「개인정보보호법」으로는 한계가 있다. 즉 AI 시스템의 특수성을 반영한 접근이 필요하다는 것이다. 강력하게 개인정보를 보호하기 위한 방안으로 제시되는 것 중 하나가 ‘프라이버시 강화 기술(Privacy-Enhancing Technologies, PETs)’이다. 이는 데이터를 암호화하거나 난독화한 상태에서 분석, 처리, 활용이 가능하게 함으로써 데이터 활용과 프라이버시 보호를 동시에 가능하게 한다. 다양한 방법에 대해 논의가 활발히 이루어지고 있으나, 실질적인 도입은 매우 더디게 진행된다.

마. (공공행정⑤) AI의 정책 결정 과정(이유 및 근거)에 대해 시민이 납득할 수 있는 설명을 제공할 수 있는가?

1) 제안의 배경

시민의 권리와 의무에 직접적인 영향을 미치는 행정 처분은 그 결정의 이유와 근거가 명확히 제시되어야 한다. 이는 시민이 결정에 불복하고 이의를 제기할 수 있는 절차적 정당성의 핵심이다. 만약 AI가 모델 확률 40%에 의해 복지 수급 대상에서 탈락했다는 결과를 통보한다면, 시민은 무엇이 잘못되었는지, 무엇을 보완해서 재신청해야 하는지 알기가 어렵다. 그렇게 되면 이의제기권과 권리 구제의 기회를 원천적으로 박탈당하게 되고, 이는 행정의 정당성을 근본적으로 훼손할 뿐만 아니라, AI 시스템에 대한 불신과 저항으로 이어지게 된다. 즉 AI 정책 결정 과정이 왜, 어떤 근거로 내려졌는지에 대해 납득할 수 있는 AI 정책 결정 과정에서 왜 그러한 결정이 내려졌는지, 그리고 어떤 근거에 기반하고 있는지에 대해 시민이 납득할 수 있는 수준의 설명을 요구할 권리가 필요하다. 이는 행정의 민주적 정당성과 시민 권리 구제의 실효성을 보장하기 위한 핵심 전제 조건이다.

2) 현황 및 쟁점

2024년 7월 한국전자통신연구원(ETRI)은 오스트리아 국제응용시스템분석연구

소(IIASA)와 함께 공공정책 의사결정을 돕는 AI정책지능 연구 협력 체계를 구축하기로 했다고 밝혔다. 현실에서 다루기 힘든 정책 요소를 가상으로 설정하고 실험할 수 있어 과학적이고 신뢰할 수 있는 정책 의사결정에 도움이 될 것으로 보았다. 2025년 4월에는 영국 지방정부가 주택정책 결정에 AI를 도입하는 실험에 나섰다. 성공할 경우 공공행정 영역에서 AI 기반 의사결정 체계가 본격화될 것으로 전망된다. 이처럼 AI의 정책 의사결정은 앞으로 계속되고, 더 심화될 것으로 예상된다. 따라서 AI의 설명 가능성에 대한 논의가 활발하게 진행 중이지만, 시민이 납득할 수 있는 설명의 수준과 범위에 대한 사회적 합의는 부족하다.

3) 문제의 심각성

설명 가능한 인공지능(Explainable AI, XAI)은 AI 모델 내부를 해석하고 시각화하는 기술적 방법을 개발하는 것에 집중한다. 그러나 기술적으로 완벽한 설명을 비전문가인 시민이 이해할 수 있는 직관적이고 맥락적인 설명으로 전환하는 것은 상대적으로 미흡한 수준이다. 즉 시민의 인지적 수용 수준을 고려한 설명 설계가 부족하다는 것이다. 더불어 모델의 정확성과 설명 가능성 확보는 상충 관계(trade-off)에 있어 두 가지를 동시에 충족한다는 것은 어려운 일이다. AI 결정 과정이 불투명하면, 시민들은 자신의 불이익이 알 수 없는 알고리즘에 의한 것이라고 인식하여 행정 결정에 대한 불신을 갖게 될 수 있다. 이는 AI 공공·행정 서비스에 대한 수용을 낮추고, 궁극적으로 국가 정책 집행의 민주적 정당성을 훼손하게 될 가능성이 높다.

4) 기존 논의/대책의 한계

EU GDPR은 '자동화된 의사결정'에 대한 설명요구권을 반영하였으며, 우리나라에서도 「AI 기본법」에 고영향 AI 사용 시 설명 및 고지의무 등이 포함되는 등 법제화가 진행되고 있다. 그런데 법적으로 어디까지 설명해야 의미 있는 설명인지에 대한 사회적 합의가 어렵다. 개발자에게 유의미한 설명과 시민에게 유의미한 설명은 일치하기 어려울 수 있다. 예를 들어, 개발자가 변수 가중치 값을 중심으로 설명하게 되면 시민은 납득하기 어려울 수 있다. 또한 민간 기업이 개발한 AI

알고리즘을 공공부문에서 도입하게 되면, 기업이 영업 비밀을 이유로 핵심 로직 공개를 거부할 경우 정부조차 그 결정 과정을 설명하기 어려운 문제가 발생할 수 있다.

제 3 절 AI와 이용자: AI 일상화 시대의 리터러시와 격차 해소

1. 논의의 배경

인공지능 기반 서비스는 사회 전반의 일상과 업무, 교육, 공공서비스까지 확산하여 인간의 판단이나 선택 등에 직접적인 영향을 미치는 보편적 기술로 자리매김하고 있다. 특히 생성형 인공지능 기반 서비스들은 상호작용형 이용자 경험을 요구하고 있으며, 이에 따라 이용자의 역량 수준이 서비스 성능과 안전성, 나아가 사회적 신뢰 형성에까지 영향을 미치는 새로운 국면에 접어들고 있다. 이러한 변화 속에서 인공지능을 이해하고 활용하며 비판적으로 해석 및 수용하는 역량, 즉 인공지능 리터러시(AI literacy)는 단순한 기술 숙련도의 차원을 넘어서, 디지털 시민성(digital citizenship)의 핵심 요소로 부상하고 있다(OECD, 2025; UNESCO, 2024).

유네스코(UNESCO, 2021), OECD(2019, 2023) 등의 국제기구는 일찍이 인공지능 윤리 및 책임을 강조하는 정책권고에서 이용자 역량 강화(user empowerment)를 하나의 축으로 삼으며 리터러시 교육의 필요성을 명시적으로 강조해 왔다. 이에 따라 유네스코는 학생들을 위한 인공지능 역량 프레임워크 보고서(UNESCO, 2024)를 통해 인공지능 개발자로 성장하는 진로 교육뿐만 아니라 책임감 있는 인공지능 이용자로서의 역량 향상도 균형 있게 공교육 과정에 포함되어야 한다고 권고하고, 인공지능 윤리, 인간 중심 사고(human-centred mindset), 인공지능 기술 및 응용과 시스템 설계 등 네 가지 차원에서 12가지 역량을 개발하고 교육 과정 목표와 영역별 교육 방법론을 제시하였다. OECD도 유럽위원회(EC)와 공동으로 인공지능 리터러시 프레임워크를 개발하였으며, 이 역시 사람들이 인공지능에 대해 비판적 사고와 윤리적 책임을 갖고 사용하게 하기 위한 교육 프레임워크이다(OECD & EC, 2025).

그러나 다른 각도에서 보면, 사실상 인공지능 리터러시는 인재 양성 정책에 밀려 구체적인 실천 전략이 부재한 채 꽤 오랫동안 사회적 담론 수준에 머물러 있기

도 하였다. 각국의 국가 전략은 인공지능 기술개발자를 확보하기 위한 인재 양성 정책에 더 비중을 두고 있었고 일반 시민을 위한 인공지능 리터러시 확산 정책은 상대적으로 부족하였다(OECD, 2025). 최근에서야 공교육 체계에 인공지능 윤리 교육 등 인공지능 리터러시 함양을 위한 커리큘럼이 교육부 차원에서 개발되는 등 실천 단계로 들어서고 있다고 볼 수 있다. 초·중·고 과정에서 인공지능 교육 비중을 높이고 있는 우리나라의 경우 「2022 개정 교육과정」(교육부, 2022. 12) 총론 및 각론에서 ‘인공지능의 사회적 영향’과 ‘인공지능 윤리’를 주요 성취 기준에 포함하였고, 인공지능 기반 디지털 교과서 도입도 앞두고 있어서 앞으로 인공지능 리터러시 교육 비중이 증가할 것으로 기대된다.

인공지능 리터러시의 필요성은 학령기 아동에게만 국한되지 않는다. 이미 사회에 진출한 성인들 역시 급속히 변화하는 인공지능 기술 환경에 적응하고 이를 비판적으로 활용할 수 있는 역량이 필요하다. 성인 대상 인공지능 리터러시 교육은 공교육과는 다른 차원의 복잡성을 지닌다. 공교육에 인공지능 리터러시 과정을 도입하는 것은 교육과정 개편, 수업 시수 확대, 교사 연수 등 상대적으로 체계적이고 제도적인 접근이 가능하다. 반면, 성인 대상인 경우 여러 한계가 있다. 먼저 이들을 교육에 참여시킬 동력 확보가 어렵다. 직업적 필요성을 체감하지 않는 이상 이들의 자발적 참여를 기대할 수 없고, 필요성을 인식하더라도 시간적, 경제적 여건상 교육 참여가 현실적으로 제약된다. 교육 접근성의 문제도 있다. 지역별, 계층별, 연령별 차이로 인해 실질적으로 교육 기회에 접근할 수 있는 성인들은 한정적이며, 이는 결국 디지털 격차와 맞물려 여러 사회 이슈로 연결될 수 있다.

고령인이나 장애인 등 전통적 디지털 취약계층의 경우 상황은 더 심각할 수 있다. 이들은 기본적인 디지털 기기 사용도 어려워하는 경우가 많아 인공지능과 같은 한층 복잡한 기술 환경에 적응하기가 더 힘들 수 있다. 이들은 인공지능 기반의 디지털 서비스 전환에서 배제되거나 수동적 사용자로 머무를 수 있으며, 이는 정보 접근권과 표현의 자유, 자기결정권 등 헌법상 보장된 권리 실현에도 심각한 제약이 될 수 있다. 다른 한편으로, 인공지능은 이들이 신체적·인지적 제약을 극복하게 돕는 지능형 도구가 될 수 있고, 경제활동을 위한 새로운 기회가 제공될

수도 있다. 따라서 이들이야말로 인공지능 리터러시 교육이 특히 필요한 사회구성원이자 서비스 이용자이다. 기술의 편익을 누리는 과정에서 개인의 존엄성과 자율성을 잃을 위험에 노출될 수 있는 만큼, 스스로 자신을 보호할 수 있는 최소한의 역량이 요구된다.

이에 따라 본 연구에서는 인공지능 리터러시 정책이 이제 막 실천 단계로 본격화하는 현시점에서 디지털 격차를 최소화하고 사회 구성원 모두가 고르게 혁신 기술의 혜택을 받을 수 있기 위해 관련 정책을 어떻게 설계해야 할지에 대한 사회적 논의 주제를 끌어내 보고자 한다. 즉 현시점에서 우리 사회가 우선으로 논의하고 합의해야 할 핵심적인 의제들을 도출하여 제시하고자 한다.

2. 주요 분석 내용

가. 인공지능 리터러시와 사회적 필요

1) 리터러시 개념의 변화

리터러시(literacy)는 역사적으로 단순한 문자 해독 능력에서 출발하였으나, 사회·기술 환경의 변화에 따라 정보처리, 디지털 활용, 미디어 이해 등을 포함하는 복합적이고 확장된 능력으로 그 의미가 발전해 왔다. 가장 기초적인 정의는 읽고 쓰는 기능적 능력이다. 이후 교육학에서 비판적 문해력(critical literacy)으로 의미를 확장하였으며, 단순한 문자 해독 능력을 넘어 글을 비판적으로 분석하고 맥락적으로 이해하며 목적과 상황에 맞게 활용할 수 있는 능력을 의미하였다(Lee, 2011, 원출처 Freire, 1984).

이후 인터넷의 등장과 함께 리터러시에 새로운 역량이 추가되었다. OECD (2000)는 리터러시를 인쇄 또는 전자매체를 통해 정보를 이해하고 사용하는 능력으로 정의하고, 학교나 직장, 공동체 등 다양한 삶의 맥락에서 적용 가능한 문제 해결 능력의 핵심으로 간주하였다. 당시 일상의 보편 기술로 정보기술(information technology)이 주목받던 정보사회(information society)에서, 리터러시는 단지 문자를 해독하는 것이 아니라, 정보를 식별하고 평가하며 효과적으로

활용할 수 있는 능력으로 재정의된 것이다. 이와 관련하여 미국도서관협회(American library association, ALA)는 정보 리터러시(information literacy)라는 개념을 사용하였는데, 이는 언제 정보가 필요한지를 인식하는 것과 함께 효과적인 정보 접근과 평가, 그리고 정보 활용 능력을 의미한다(American library association, 1989). 이 정보 리터러시는 정보 과잉 사회에 대응하는 핵심 역량이었다. 그 밖에 유네스코(UNESCO, 2005)는 리터러시를 사람이 일상생활과 사회 참여를 위해 갖춰야 하는 필수 역량으로서 ‘읽고, 쓰고, 이해하고, 의사소통하는 능력’으로 정의하였다. 이 정의에서 리터러시의 주요 역량으로 의사소통 능력이 포함되는데, 이는 인터넷을 통해 시공간의 제약과 상관없이 소통할 수 있게 된 기술 환경 변화를 반영한 것으로 보인다.

이처럼 인터넷 시대 여과되지 않은 방대한 정보에 직접 노출되면서, 무엇보다 정보를 비판적으로 평가하고 목적에 맞게 활용할 수 있는 역량이 중요해졌다. 이에 더해 일상에서 컴퓨터 외에 다양한 디지털 기기의 활용과 함께 ‘디지털 리터러시’가 통용되었다. 디지털 리터러시는 Gilster(1997)가 개념화한 용어로, 단순히 온라인에서 필요한 정보를 찾는 행위를 넘어, 다양한 형식과 출처의 정보를 수집·이해하고 정보에 근거한 판단을 내리는 능력까지 포함한다. 즉 문제 해결 능력과 비판적 사고로까지 역량 범위가 넓어졌다.

한편 UNESCO(2011)는 기존의 정보, 디지털, 미디어 리터러시 개념을 통합하여 ‘미디어 및 정보 리터러시(media and information literacy, MIL)’라는 개념틀을 제시하기도 하였다. MIL은 개인이 정보와 미디어에 대해 비판적으로 접근하고 평가하며, 이를 기반으로 새로운 생각이나 아이디어를 형성하거나 의사소통하는 능력을 의미한다. 이 개념은 컴퓨터, 인터넷뿐만 아니라 사람들의 사회문화적 인식과 태도에 지대한 영향을 미치는 모든 미디어를 포괄한다. UNESCO(2011)는 이 능력을 민주주의에서의 시민 참여나 지속가능 발전을 위한 시민의 핵심 역량으로 간주하였다.

이처럼 리터러시는 정보에 접근하는 채널 확대 및 기술 환경의 변화에 따라 사회문화적으로 요구되는 역량을 포함하면서 개념을 확장해 왔다. 인공지능 리터러

시 역시 이러한 기존 리터러시 개념과 연계해 볼 수 있으며, 정보 이해와 활용을 넘어 인공지능 기술과 그 산출물을 비판적으로 수용하고 사회적 맥락에서 책임 있게 활용하는 능력으로 정의할 수 있다.

2) 인공지능 리터러시

인공지능 리터러시는 인공지능이라는 고유한 기술 특성과 사회적 영향력을 반영한 새로운 개념으로 이해되나 아직 신생 개념이라고 할 수 있다. 인공지능 기반의 새로운 디지털 서비스가 등장한 초기에는 인공지능 서비스의 활용 역량을 디지털 리터러시의 하위 개념으로 간주하여 디지털 리터러시에 인공지능에 대한 이해와 활용 능력을 포함하는 정도로 다루어왔다(Vuorikari, Kluzer, & Punie, 2022). 그러나 인공지능 기술의 발전과 활용 양상이 변화하면서 독립적인 개념으로서의 인공지능 리터러시가 주목받고 있다.

이는 인공지능 기술이 디지털 기술과 다른 고유한 특성을 지니고 있으며, 그로 인해 사회에 미치는 영향도 새로운 양상으로 나타날 수 있기 때문이다. 특히 인공지능시스템은 이용자의 개입 없이 혹은 이용자의 의도와는 별도로 결과물을 산출한다는 점에서, 인공지능이 어떤 과정을 거쳐 결과물을 산출하는지 이해하고 그 결과물을 어떻게 해석하고 활용할지, 그리고 인공지능의 잠재적 위험을 어떻게 확인하고 대응할지 아는 것이 중요해졌다.

인공지능 리터러시는 대략 ‘인공지능 기반 시스템을 이해하고 신뢰하며 윤리적으로 사용하는 데 필요한 인지적·기술적·사회적 역량의 집합’(OECD, 2021)으로 요약해 볼 수 있다. OECD(2021; 2024)는 단순한 기술 활용 능력을 넘어서 인공지능 시스템이 사회에 미치는 영향과 윤리적 함의를 인식하고 이에 대응할 수 있는 비판적 역량을 인공지능 리터러시 개념에 포함해야 한다고 강조하였다. 이에 따라 인공지능 리터러시의 핵심 역량을 인공지능 개념과 기능 이해, 인공지능 기반 서비스의 활용과 해석, 인공지능의 사회적·윤리적 영향 인식, 책임 있는 이용과 의사결정 참여 등 네 가지 영역으로 구분하였다. OECD(2024)는 이러한 인공지능 역량이 인공지능 시대에 경제 성장에 제 역할을 하는 필수 역량이라는 점을

강조하고 있다.

한편 UNESCO(2021)는 인공지능 리터러시를 모든 시민이 인공지능 기술을 안전하고 공정하게 사용할 수 있도록 하는 디지털 포용 정책의 핵심 기반으로 평가한다. 이에 따라 개발도상국과 사회적 취약계층을 위한 리터러시 교육 및 접근성 확보를 주요 정책 방향으로 제시하고 있다. 또한, 교육자와 학습자를 위한 역량 프레임워크를 구성하였으며, 여기에는 인공지능 개념과 원리, 인간-기계 상호작용, 윤리적 사고, 알고리즘 투명성, 편향 인식 등의 주제가 주요 학습 목표로 포함되어 있다.

학계에서는 인공지능 리터러시 개념을 교육 커리큘럼과 연결하여 접근하는 경향을 보인다. 예를 들어, Long and Magerko(2020)는 인공지능 리터러시를 인공지능 시스템의 작동 원리를 이해하고 인공지능이 생산한 정보를 분석하여 이에 기반한 비판적 의사결정을 수행할 수 있는 능력으로 규정하였다. 이들은 인공지능 리터러시를 ① 인공지능 개념 이해(conceptual knowledge), ② 도구로서의 인공지능 활용(practical skills), ③ 사회·윤리적 맥락에 대한 비판적 사고(critical and ethical reasoning)라는 세 가지 핵심 영역으로 구분하였는데, 이는 기존의 인공지능 교육 커리큘럼 구성과 유사한 특징을 보인다. 국내에서도 교육학적 접근에서 인공지능 리터러시를 다루는 연구가 주류를 이루고 있으며(이유미, 2022; 최숙영, 2022; 황현정·박지수·김승완, 2023), 기존의 디지털 미디어 리터러시에는 없었던 ‘비판’, ‘모니터링’, ‘(기술과의) 협업’, ‘윤리’ 등 새로운 키워드가 추가된 것에 주목한다(박선미, 2025). 이들 연구는 대체로 인공지능 리터러시 함양을 위해 성취해야 할 기준을 확인하는 과정에서 인공지능 리터러시의 개념과 구성요소를 탐구하였다.

지금까지 살펴본 문헌들을 종합하면, 인공지능 리터러시는 인공지능이 보편 기술이 된 사회의 구성원이라면 필수적으로 갖춰야 할 역량이자, 인공지능 사회를 살아가면서 민주주의와 공공선을 지키는 데 필요한 시민의 필수 소양으로 인식되고 있다(UNESCO, 2024; European Commission & OECD, 2025). 다만 이

인공지능 리터러시를 함양하기 위해 구체적으로 어떤 개별 역량이 필요한지, 어떤 교육 목표를 설정하고 커리큘럼을 개발할지는 여전히 탐구가 필요한 영역이다. 현재는 OECD나 UNESCO 등의 국제기구가 일차적인 기준을 제시한 초기 단계에 있기 때문이다.

나. 인공지능 리터러시 함양을 위한 정책

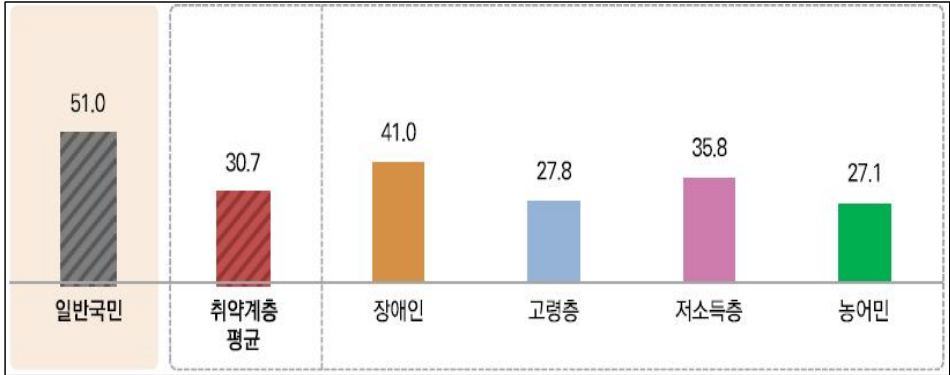
생성형 인공지능의 일상 침투 이후, 인공지능 기술개발자 양성을 위한 고급 역량보다는 인공지능에 대한 전반적인 이해와 활용 능력이 더 필요해졌지만, 우리나라를 비롯한 세계 각국의 인공지능 교육은 인공지능 기술 활용에 더 방점을 두고 있다(OECD, 2025). 이에 OECD(2025)는 정책적으로 인공지능 리터러시 증진을 위해 산업계와의 협력을 강화하고 더 포괄적인 학습 경로를 개발할 것을 권고한다. 산업계는 인공지능 기술개발 수준의 전문성보다는 업무에서 인공지능 도구를 잘 다루는 노동 인력이 필요하기 때문이다.

국내 상황을 보면, 정부 지원의 교육 정책에 일반 시민을 대상으로 한 인공지능 리터러시 함양을 위한 지원 정책은 미미하다. 기존의 디지털 교육 사업에 인공지능 관련 과목이 일부 개설되거나 교육과정 중 일부 시간을 할애하는 정도이며, 인공지능 리터러시 함양에 대한 명확한 교육 목표나 과정에 따른 구체적 커리큘럼이 정립되어 있지 않다. 이 와중에 고령인, 장애인 등과 같은 디지털 취약계층을 위한 인공지능 교육 프로그램은 더 찾기 힘들다. 이들은 기본적인 디지털 기기 사용 능력을 갖추는 데도 많은 노력을 해야 하기에 새로운 영역으로 교육 범위를 확대하는 것이 쉽게 이뤄지지 않는다.

다음의 [그림 3-32]에서 확인할 수 있듯이, 디지털 취약계층뿐만 아니라 일반 시민도 인공지능 서비스 이용 경험이 아직 충분하지 않다. 이는 인공지능 리터러시 교육의 정책적·사회적 준비 부족의 주요 이유일 수 있다. 일반인 및 디지털 취약계층 대상의 디지털 활용 교육 현장에서 학습자의 수요는 다양한 인공지능 서비스 툴 이용 방법을 배우는 데 쏠려 있고, 비판적이고 주체적인 활용 태도 정립과 관련 지식은 학습자의 수요에서 벗어나 있다.

[그림 3-32] AI 서비스 이용경험

(단위: %)



자료: 한국지능정보사회진흥원(2025. 3), 『2024 디지털정보격차 실태조사』, p. 295

체계화된 공교육에도 인공지능 관련 교육 시수 확대와 함께 리터러시 교육 방법을 모색하고 있는 상황에서, 일반 시민뿐만 아니라 고령인과 장애인을 고려한 성인 대상 인공지능 리터러시 교육 체계 개발은 요원할 수 있다. 그러나 학습 대상자별 순차적 정책 접근의 속도가 인공지능 기술의 발전 속도와 우리 사회로의 침투 속도보다 현저하게 느리다면, 심화된 디지털 격차로 더 큰 사회 문제와 비용이 발생할 수도 있다. 이에 다음에서는 인공지능 리터러시 정책을 개발, 추진하는 초기부터 디지털 취약계층을 고려할 당위를 모색해 보고자 한다.

다. 미래 인공지능 리터러시 정책 대상자 고찰

1) 고령인이 누릴 수 있는 인공지능 혜택과 리스크

고령인이 인공지능 시스템으로 누릴 수 있는 혜택을 꼽아보면, 신체적 및 정서적 지원뿐만 아니라 일상생활에서도 간단한 돌봄 지원을 받을 수 있다. 대화형 시스템을 활용한 상호작용에서 정서적 안정을 찾을 수 있고, 요양보호사를 대신하여 신체적 보조 역할을 지원받을 수 있다. 또 일상생활에서 약 복용 알림이라든가 혈압 체크, 안전 여부 확인 등의 기능을 수행할 수 있다(김정근, 2021; 국가인권위원회, 2022). 이들이 누릴 수 있는 사회 혜택은 곧 우리 사회의 돌봄 의무를 덜어주는 것이기도 하므로 사회 전반에 긍정적 효과를 가져올 수 있다.

다음으로 이들이 인공지능 시스템에 대한 이해나 비판적 태도를 갖추지 못한 채 이러한 시스템에 노출되었을 때 발생할 수 있는 리스크를 확인해야 한다. 먼저 인공지능 시스템과의 상호작용에 지나치게 몰입하였을 때 문제가 된다. 과몰입 행위에는 인공지능 시스템을 인간처럼 대한다든가 과도하게 신뢰하는 경향 등이 포함된다.⁵⁹⁾ 그밖에 생활 지원 시스템은 개인의 민감한 프라이버시를 외부에 유출할 리스크를 잠재하고 있으며, 신체 지원 시스템에는 작동 오류로 인한 물리적 피해 가능성이 내포되어 있다. 인공지능 모델의 데이터 편향으로 정서적 침해나 피해가 발생하였을 때 책임 규명 절차에 대한 무지 등으로 인한 불이익 등도 생각해 볼 수 있다.

2) 장애인이 누릴 수 있는 인공지능 혜택과 리스크

장애인도 고령인과 마찬가지로 인공지능 시스템으로 기존보다 더 지능적이고 편리한 신체적·정서적 및 생활 지원 서비스를 받을 수 있다. 이들의 나이가 청장년층이나 그 이하라면, 인공지능 시스템을 통한 장애 보완으로 훨씬 더 다양한 사회활동도 가능할 것이다. 이들을 위한 인공지능 보조 기술은 주로 시각 및 청각 장애인들을 위한 생활 지원, 그리고 학습장애 및 자폐 극복을 위한 진단 및 교육 지원에 집중되어 있다(권정민·이영선, 2020).

이들이 경험할 수 있는 리스크는 편향성, 프라이버시 침해 등 인공지능 시스템의 잠재적 위험뿐만 아니라 작동 오류나 제때 이루어지지 못한 시스템 업데이트 등으로 인한 피해에 노출될 수 있다. 인공지능 시스템의 보조를 받아 직장 근무를 포함한 여러 사회활동을 하고 있었다면, 이들이 경험하는 피해는 더 클 것이다. 예를 들어, 청각장애인이 운전하는 차량 공유 서비스 차량에서 통신 장애가 발생하였다면 손님과 소통이 원활하지 않아 서비스 제공에 어려움이 발생하고 중단해야 할 수 있다.

59) 공개된 실제 사건으로 미국의 60대 남성이 소금 섭취량을 줄이고자 챗GPT에 대체 방법을 문의하였고, 챗GPT가 제안한 화학약품을 검증 없이 섭취하다가 정신병원에 입원한 사례가 있었다(조선일보, 2025. 8. 14).

라. 전문가 자문을 통한 현황 및 이슈 탐구

여기에서는 지금까지 문헌 등을 참고하여 살펴본 사항들을 전문가 자문 인터뷰를 통해 다시 한번 고찰한다. 전문가 자문에는 단국대 특수교육대학원장 한경근 교수, 경기도 성은학교의 김원형 정보교육부장, 한국지능정보사회진흥원의 차하나 책임, 한신대학교 재활학과의 남세현 교수가 참여하였다.

〈표 3-27〉 전문가 인터뷰 대상자

전문가	소속	지위
한경근	단국대 특수교육대학원	대학원장/교수
김원형	경기도 성은학교	정보교육부장
남세현	한신대학교 재활학과	교수
차하나	한국지능정보사회진흥원 디지털역량개발팀	책임

자료: 연구진 작성

1) 디지털 취약계층의 인공지능 서비스 접근성 문제

먼저 일상에서의 인공지능 서비스 활용과 접근성 논의에서 장애 vs. 비장애의 이분법적 접근을 지양해야 한다는 논의가 있었다. 개인의 신체적, 정신적 장애의 문제가 아니라 생활 전반에 발생할 수 있는 사회적 장애 요인을 해소하는 관점에서 접근해야 한다는 주장이다.

“인공지능 활용과 접근성 논의는 단순히 ‘장애/비장애’라는 이분법을 넘어서야 합니다. 생활 전반에 발생하는 사회적 장애 요인을 줄이는 관점에서 접근해야 해요.”
(한경근 원장)

남세현 교수는 장애인의 편의를 위한 접근성이 물리적 접근성뿐만 아니라 온라인에서의 접근성도 법적으로 보장되어 있음을 강조하였다. 그러나 실제로 해당 접근성 규정이 충분히 반영되어 있는지는 또 다른 문제라고 지적한다.

“장애인차별금지법이나 편의증진법은 물리적 접근성뿐 아니라 정보통신 접근성도 규정하고 있습니다. 공공기관 홈페이지나 서비스는 시각 및 지체 장애인을 고려한 접

근성 인증이 필요하죠 하지만 정보 변환 기술(쉬운 글, 점자, 수화 등)은 기술적으로 가능하나 실제 현장 적용은 부족한 상황입니다.”(남세현 교수)

남세현 교수는 한경근 원장과 비슷한 시각으로 장애인의 눈높이에 맞춘 접근성 구현은 곧 일반적 상황의 대중 모두가 누리는 혜택이 될 것이라고 보았다. 시각장애인을 위해 복지 차원에서 제작되던 오디오북이 지금은 비시각장애인까지 확대되어 누구나 활용하는 서비스가 되었듯이, 장애인을 위한 편의는 비장애인을 위한 편의로 무리 없이 확대될 수 있다는 것이다.

이어, 남세현 교수는 디지털 취약 계층에게 긍정적 영향을 주고 나아가 일반인들에게도 잠재적 이점을 줄 수 있을 것이라고 다시 한번 언급하면서, 실제 출시된 여러 기술 도구를 소개하였다.

“인공지능은 이들에게 큰 도움이 될 수 있어요 예를 들어, 시각장애인을 위한 센서, 레이저 피드백 방식의 지팡이, 청각장애인을 위한 시각적 알람 시스템, 스마트가전의 개인 맞춤 서비스 등이 가능하죠” (남세현 교수)

“많은 혁신 기술들이 사회적 약자 지원과 직결되고 있습니다. 수어를 보여주는 아바타를 제작하거나 수어를 음성으로 변환하는 시스템, 지팡이형 시각장애 안내 디텍터, 휴머노이드 웨어러블 장치에 결합하는 방식으로 개발되기도 하고요. 비장애인 대상 기술에 프로토타입이나 모듈 같은 소규모 기술을 추가함으로써 접근성 기능을 탑재할 수 있습니다.” (남세현 교수)

남세현 교수가 정리한 분야별 실제 활용 사례는 다음과 같다. 첫째, 돌봄 분야에서 돌봄 제공자 부담을 경감하는 경우인데, 낙상·배변 등 안전이나 필요를 감지하는 센서와 인공지능 경보시스템의 결합인 바이탈 모니터링 기반 예방 및 정밀 진단 시스템이다. 둘째, 아동·신체 기능 보조 분야에서 웨어러블 보행 및 팔 동작 보조, 전자 의수·의족으로 자연스러운 동작 구현, 전동휠체어 자율주행화 등이 있다. 셋째, 감각 보조 분야에서는 시각장애인을 위한 카메라·AI 판독 기술이나 청각장애인을 위한 수어 인식 및 변환 시스템 등이 실사용 단계로 진입했다고 한다.

“시각장애의 경우 카메라와 인공지능 기술 연계로 사람, 차량, 장애물 등을 인식하고 표정, 연령, 성별 등 다양한 상황판단의 보조가 가능해요 청각장애는 한국어 음성 인식 정확도의 향상으로 회의나 원격수업에서 실시간 자막 보급이 가능하고, 경보음이나 위험신호를 진동, 빛, 자막 등으로 변환해서 안전 알람을 제공할 수 있습니다.”
(남세현 교수)

그밖에 언어장애나 발화 곤란의 경우 문장 예측과 TTS 기반 보완대체의사소통(AAC)으로 사용자가 선택, 편집하면 기기가 대신 발화할 수 있다고 한다. 인지장애나 발달장애, 치매 환자에게는 인공지능이 안전과 비안전을 판별하여 위험 상황을 예측하여 경고할 수 있고, 학습 수준과 진도를 실시간으로 조절하여 개인에 맞춘 학습도 가능하다고 한다.

다만, 인공지능 기반 서비스 확대에 따른 우려도 있다. 무엇보다 인공지능 기반 차별과 혐오 감지 시스템이 필요하다고 하였다. 즉 언론기사나 보도자료, SNS와 댓글 등에 있는 차별적 용어나 정보를 인공지능이 자동 검출하여 경고나 차단을 하는 모니터링 체계가 필요하다는 것이다. 이는 사용자를 보호하기 위한 것이다. 디지털 격차 우려에 대해서는, 디지털 접근성이 부족할 경우 배제와 차별이 심화될 위험이 있고, 현재처럼 키오스크 시스템 사용이 널리 퍼진 사회에서 키오스크 접근성이 미비한 경우 장애인의 일상 참여 자체가 제한된다. 서비스 제공자들은 현재 서비스에 디지털 포용에 걸맞는 시스템을 갖추기 위한 유예기간을 요청하지만, 그 사이에도 디지털 및 인공지능 기술 발전이 빠르게 진전되고 우리 일상에서 이 기술들을 기반으로 한 서비스가 확산되고 있는 상황에서 그 유예기간은 오히려 디지털/AI 격차만 심화하는 기간이 될 수 있다고 경고한다.

또한 이러한 디지털/인공지능 기기/서비스를 이용하는 장애인들을 보호하기 위해 개발사 및 서비스 제공자의 윤리나 인권 인식을 높이는 법·제도적 조치도 중요하다고 지적한다. 예를 들어 실종 방지나 배회 감지 기술을 악용하는 경우 프라이버시나 인권을 침해하는 기술이 될 수 있기 때문이다.

우리나라를 비롯한 세계 각국은 장애인의 일자리를 법적으로 보장하고 있다. 그러나 발달장애인이 주로 종사하는 단순 반복 작업 분야가 인공지능이나 여러 자

동화 기술로 빠르게 대체될 때, 장애인의 안정적 고용 기반이 붕괴될 가능성이 있다.

시각 및 청각 장애인과 비교하면 지적 장애인의 장애 스펙트럼이 매우 커서 학교라는 인위적 공간에서 지식 및 교육 커리큘럼을 일반화하기 어려운 상황이다. 이러한 현실에서 인공지능을 활용하였을 때 교육 효과와 다양성이 높아질 것이라는 기대감도 있었다.

“교사들이 인공지능을 수업에 활용할 때 표현이나 참여 기회가 확대될 수 있습니다. 그림이나 만화 생성 등을 통해서요 특히 지적장애 학생들의 경우, 학교라는 인위적 공간에서 지식의 일반화가 어려운데, 인공지능을 활용하면 현실적 경험을 시뮬레이션할 수 있어 교육 효과와 다양성이 높아집니다. 시장이나 편의점 같은 실제 상황을 가상으로 체험할 수 있거든요” (한정근 원장)

2) 특수교육 현장에서의 인공지능 리터러시 교육 현황

특수교육 현장에서도 인공지능 리터러시 교육의 필요성에 대해서는 충분히 공감하고 관심이 있다고 한다. 다만 시각장애, 청각장애와 같은 감각 장애가 있는 학생들뿐만 아니라 지적장애 및 자폐성 장애가 있는 학생들에게 인공지능 기술 및 리터러시 교육을 제공하는 데는 어려움이 많다고 한다.

“2022 개정 교육과정에서 관련 사항들이 포함되었고 교과서 편성도 완료되어 현장에서 적용 중입니다. 하지만 특수교사들이 아쉬워하는 부분이 많아요 장애 학생의 스펙트럼이 매우 넓거든요 학습이 전혀 불가능한 학생부터 전국 장애학생 정보경진대회에 참여할 만큼 정보화 역량이 뛰어난 학생들까지 있어서 교육적 요구 수준이 극단적으로 다양합니다. 그래서 일괄적 교육자료 적용에 어려움이 있어요” (김원형 교육부장)

교육부나 교육청에서 제공하는 인공지능 리터러시 관련 자료는 주로 초등 수준의 단순화된 형식으로 보급되는데, 주로 시청각 자료 형태여서 시각 및 청각 장애인은 사용하기 어려운 경우라는 지적이다. 점자나 음성 자료가 필요한데, 시각 및 청각 장애인 특히 공교육 과정에 있는 시청각 장애인 학생 수가 상대적으로 적어서 이들을 위한 교육 콘텐츠 개발이 수월하지 않은 현실이다.

인터뷰에 참여한 전문가 중 특히 교육 현장의 교사가 인공지능/디지털 리터러시 교육의 필요성을 가장 절실히 경험하고 있었다.

“장애학생들도 스마트폰을 사용하고 있어서 무분별한 자료 공유, 사진 전송, 보이스 피싱 등의 위협에 쉽게 노출됩니다. 무분별한 소셜결제 같은 실제 피해 사례도 있고 개인정보 유출 등의 사례가 증가하고 있어요. 장애 학생들은 더 쉽게 피해에 노출되므로 체계적이고 맞춤형 디지털/인공지능 리터러시 교육이 시급한 상황입니다.”
(김원형 교육부장)

그러나 현실은 리터러시를 가르쳐야 하는 교사들 사이에서도 인식 차이가 크다. 일부는 인공지능 리터러시 교육의 필요성을 공감하고 관련 자료를 적극 활용하려고 하지만, 다수는 관심이 부족하거나 상대적으로 필요성을 크게 인식하지 않는다고 하였다.

최근 교육청에서 교원의 인공지능 전문성 함양을 학교 평가의 의무 항목에 포함시키는 등 정책적 노력이 있지만, 실질적인 전문성 함양 효과에 대해서는 아직 검증되지 않은 상황이다. 학교 평가 요건을 맞추기 위해 교직원의 90% 이상이 연 15시간 이상 원격 연수를 수강하고 이수증을 제출해야 한다. 연간 15시간이라는 물리적으로 짧은 이수 의무 시간을 차치하더라도, 인공지능 교육의 필요성과 전문성 함양의 동력이 부족하다면 원격 연수는 오히려 해당 요건을 형식만 갖추는 요식 행위로 만들 수도 있다. 이 와중에 정보 교과가 필수 교과로 지정되지 않은 특수학교에서 기존 교과 내부 단원으로 디지털/인공지능 활용 요소를 포함하는 경우는 개별 교사의 의지에 크게 좌우될 수밖에 없다.

3) 고령층을 위한 인공지능 리터러시 교육 현황

고령층을 대상으로 디지털 리터러시 및 교육 콘텐츠를 개발하고 있는 한국지능정보사회진흥원의 차한나 책임은 현장의 인공지능 리터러시 교육 현황에 대해 체계적인 접근과 정책이 부재한 상황이라고 하였다. 기존 디지털 교육은 여러 디지털/인공지능 서비스 사용 방법을 학습하는 것에 집중하고 있으며, 리터러시 교육의 또 다른 중요 요소인 윤리 및 비판적 태도 교육에 관한 관심은 상대적으로 적

다는 것이다.

“인공지능 윤리교육을 단독 프로그램으로 진행하려면 수요가 낮아서 프로그램으로 개설되지 못해요. 더구나 인공지능 서비스에 대한 경험이 충분하지 않은 상태에서 고령층 등 디지털 취약계층에게 저작권이나 법적 문제 같은 내용을 소개하면 오히려 위협으로 받아들여서 폐쇄적 태도를 보일수도 있습니다. 따라서 활용 교육을 주로 하면서 리터러시 교육을 그 교육과정에 포함하여 활용과 연결해서 다루어야 합니다.” (차한나 책임)

현장에서의 윤리교육은 챗지피티(ChatGPT)나 제미니(Gemini) 등의 활용 방법을 학습할 때, 이를 통해 만든 자료를 외부에 사용할 때 출처를 표기해야 한다는 것, 프롬프트를 입력할 때 카드번호나 그 밖의 민감정보 등의 개인정보는 입력해서는 안 된다는 등의 주의 사항을 같이 전달하는 정도로서 그 비중이 높지 않다. 이처럼 윤리교육을 활용 교육에 결합하면 단순한 원론 전달이 아니라 실습 및 활용과 병행되기 때문에 효과가 더 크다는 것이다.

차한나 책임이 관여하는 디지털배움터의 강사 풀은 다양한 사회적 배경의 사람들로 구성되어 있다고 한다. 전업 활동 강사는 드물고, 최근에는 온라인 교육 방식을 선호한다고 한다. 이 강사진들은 인공지능 서비스 활용에는 익숙하지만, 그에 따른 리터러시를 어떻게 가르칠지에 대한 전문성은 부족하다고 하였다. 따라서 학습시키려는 서비스 활용 사례와 함께 고려해야 할 윤리적 이슈나 법제도 요소 등도 다룰 수 있도록 강사 교육 프로그램을 구성할 필요가 있다고 하였다.

차한나 책임은 최근 고령층을 위한 교육 콘텐츠 개발에서 체험형 및 상호작용형(interactive) 방식의 유용성을 전하였다. 고령층은 인공지능 서비스에 노출될 기회가 상대적으로 낮은 세대이기에 두려움과 같은 심리적 부담이 있다고 하였다. 복지관이나 주민센터를 통한 교육에서는 현장 참여가 부족할 수밖에 없다. 이에 부족한 현장 실습 기회를 디지털 기반의 가상 체험에서 찾고자 하였다.

“게임이나 시뮬레이션을 활용해서 메타버스, 키오스크, 웹 기반 등으로 참여자가 스스로 ‘나는 어떤 선택을 하는 인공지능인가’를 체험할 수 있도록 하는 것이죠 웹

기반 키오스크나 시뮬레이터를 활용하면 고령층도 텍스트 대신 체험으로 학습할 수 있어서 효과적입니다.” (차한나 책임)

차한나 책임은 최근 고령층의 키오스크 활용 교육 프로젝트에서 적용하였던 입문-실습-현장교육 3단계 설계를 이상적인 틀로 고려하고 있다. 교육 기관의 교육이 디지털 시뮬레이션 학습을 넘어 지역 거점의 오프라인 실습 비중도 확보해야 한다는 것이다. 이 과정에서 인공지능 윤리 체크리스트를 실습 흐름에 적용해 생성물 표기, 개인정보, 저작권 주의 문구 등을 자연스럽게 접하고 고민하게 하는 것이다. 이런 방식으로 2025년 인공지능 서비스에 대한 이해와 윤리기준, 그리고 활용을 통합한 체험형 교육 콘텐츠를 개발하고 있다고 한다.

4) 정책 개발을 위해 고려해야 할 점

한경근 원장은 인공지능 리터러시를 ‘이용자 개인의 역량 부족’ 이슈가 아니라 사회문제로 바라봐야 한다는 점을 특히 강조하였다. 이는 디지털 사회로 전환하면서 발생하는 사회문제라는 것이다. 따라서 법·제도적으로 보완해야 할 부분을 충실히 보완하는 데서부터 출발해야 한다고 주장한다. 예를 들어, 현재 법률에서 정보 접근 의무 조항이 인공지능 시대에도 적합한지를 면밀하게 검토하고 개선해야 한다고 제안한다. 또 장애 유형별 요구사항을 파악하고 이를 반영한 인공지능 활용 방안을 모색하여 프레임워크를 구축하고 순차적으로 접근해야 한다고 주장하였다.

남세현 교수는 장애인 관련 기술 개발 및 활용 시 인권 보장, 안정성 확보, 접근성 의무화와 같은 윤리·인권·안전 가이드라인 설계가 필요하다고 주장한다. 일자리 대체에 대해서는 극단적인 자본주의적 일자리 박탈 방지를 위해 인공지능 기업의 초과 이익을 환수하여 기본소득과 재취업 지원에 활용하는 방안을 검토해야 한다고 하였다. 인공지능이 일자리를 전면 대체하는 것이 아니라 장애인의 노동수행을 지원해서 고용의 기회를 창출하도록 유도하는 ‘협동 로봇’ 개념이 확산될 필요가 있다는 것이다.

또한 법·제도적 개선 방안과 관련하여 지원과 제재의 병행을 제안하였다. 예를 들어, 1997년 제정되어 2005년에 시행되었던 「편의증진법」은 당시 유예 및 예외

를 설정하고 신규 건물에 대해서만 의무화하였는데, 현재는 표준화되었다. 새로운 규제 도입은 저항이 있을 수밖에 없지만 저항이 있더라도 늦추지 말고 시행해야 결국 사회 총비용이 감소하므로, 정부는 비용 보전·장려와 같은 지원과 함께 기준 미준수 시 강한 제재를 하는 방향으로 신규 제도를 추진해야 한다는 것이다. 디지털 포용은 나중으로 미루면 갈등과 사회비용만 증가하게 될 것이라는 경고이다. 이와 함께 설계 초기부터 취약계층 포괄 기준을 원칙화하고 법과 가이드라인, 교육 등의 체계를 확립하여 실효성을 담보할 필요가 있다는 것이다.

3. 식별된 사회적 의제안

다음은 문헌조사 및 전문가 인터뷰, 인터뷰 이후 관련 현황 확인 등을 기반으로 연구진에서 제안하는 이용자 영역의 사회적 의제안이다.

가. (이용자01) [모든 국민의 권리로서의 AI 리터러시 교육 보장] 디지털 취약계층을 포함한 모든 시민이 AI 리터러시 교육에 평등하게 접근할 수 있도록 제도적 의무와 체계적 지원을 강화하는 데 정부는 정책적 지원과 자금을 얼마나 투입해야 하는가?

1) 제안의 배경

현재 AI 기술이 교육, 의료, 금융, 행정 서비스 등 시민 생활 전반에 걸쳐 확산되고 있지만, 이를 제대로 이해하고 활용할 수 있는 역량을 갖춘 시민은 매우 제한적이다. 특히 고령인, 저소득층, 장애인, 농촌 거주민 등 디지털 취약계층의 경우 AI 기술에 대한 접근성과 이해도가 현저히 낮아 새로운 형태의 디지털 격차가 심화되고 있다.

AI 리터러시는 단순한 기술 활용 능력을 넘어서 AI의 작동 원리 이해, 편향성과 윤리적 문제 인식, 개인정보 보호 등을 포함하는 종합적 역량이다. 이는 민주주의 사회에서 시민의 AI 기술의 영향을 받는 정책 결정에 참여하고, AI 서비스를 비판적으로 평가할 수 있는 기본권과 직결된다. 따라서 AI 리터러시 교육은 선택이 아

닌 모든 시민의 권리로 보장되어야 한다.

2) 현황 및 쟁점

OECD(2023) 조사 결과, 우리나라는 영국, 미국과 함께 AI 전문가 양성 교육에 정책 지원이 집중되는 국가 그룹에 속했으며, 일반 대중을 위한 AI 리터러시 교육은 매우 부족한 것으로 나타났다. 현재 과학기술정보통신부의 ‘모두의 AI’, 교육부의 AI 디지털교과서 도입, 고용노동부의 ‘디지털 새싹’ 등이 추진되고 있으나 대부분 학생이나 직장인을 대상으로 하며, 디지털 취약계층을 포괄하는 체계적인 AI 리터러시 교육 체계는 미흡한 상황이다.

3) 문제의 심각성

AI 기초 교육 부재는 AI 기술의 혜택이 특정 계층에만 집중되는 결과로 이어져 AI 역량 격차를 심화하고 새로운 불평등을 가중시킬 수 있다. 특히 AI 전환으로 인한 고용 악영향에 가장 취약한 저숙련 노동자 및 자동화 위험이 큰 직종의 종사자들이 AI 교육에 가장 접근하기 어려워져 계층 간 경제적·사회적 격차가 더욱 고착화될 위험이 있다.

또한 AI 기술에 대한 이해 부족으로 인해 시민들이 AI 서비스의 편향성이나 오류를 제대로 인식하지 못하고, 개인정보 침해나 조작된 정보에 무방비로 노출될 수 있고, 이는 개인의 피해를 넘어서 민주주의 제도의 근간을 흔들 수 있는 심각한 사회적 위험으로 발전할 가능성이 있다.

4) 기존 논의/대책의 한계

가장 중요한 문제는 전국 단위의 정책사업을 하기에는 부족한 예산 투입과 AI·디지털 기술의 발전과 함께 축적되는 지속되는 결여·격차이다. 현재 AI 리터러시 교육은 단발성 사업이나 시범사업 수준에 있으며, 모든 시민을 대상으로 하는 보편적 교육으로 확대하기 위한 장기적 재정 계획이나 법적 근거가 마련되어 있지 않다. 이에 따라 교육 기회의 불평등이 해소되지 못하고 있으며, AI 시대 시민 교육에 대한 국가적 책무가 제대로 이행되지 못하고 있는 상황이다.

나. (이용자02) [AI 기반 서비스의 보편적 접근성을 위한 법·제도적 보장] 고령인·장애인의 AI 서비스 활용 권리와 자기결정권을 실질적으로 보장하기 위해, AI 기반 기술/서비스 개발 시 보편적 디자인 가이드라인 등을 법적으로 표준화하고 이를 관리·감독할 거버넌스를 어떻게 구축할 것인가?

1) 제안의 배경

AI 기반 생활 편의 서비스가 확대될수록 디지털 접근성의 불평등이 새로운 사회적 배제 요인이 될 수 있다. 특히 고령인과 장애인은 기존 디지털 격차에 더해 AI 서비스 활용에서 추가적인 장벽에 직면하게 되며, 이는 기본적인 사회 서비스 이용 권리나 자기결정권 행사에 심각한 제약을 주게 된다.

이들은 또한 AI 기술의 ‘사용자’라기보다 관리 혹은 감시의 ‘대상자’로 고려되는 경우가 있다. 즉 안전·보호를 이유로 한 디지털 추적·관찰 기능의 대상이 되어 기본권 침해로 이어질 수 있다.

2) 현황 및 쟁점

가장 큰 쟁점은 기존 접근성 관련 법령이 AI 서비스의 특수성을 반영하지 못한다는 것이다. 예를 들어, 현재 공공기관 웹사이트에 대한 접근성 인증은 의무화되어 있으나, 일상에 점점 확산되고 있는 챗봇, AI 안면인식, 맞춤형 추천 등 AI 기반 서비스에 대한 접근성 인증 기준과 제도는 부재한 상황이다.

또한 민간 서비스의 경우 AI 접근성 확보는 법적 강제성이 없어 대부분 ‘권고’ 수준에 그치고 있다. 관련법을 준수하지 않더라도 시정 명령 외에 실질적인 제재 수단이 제한적이어서 기업의 자발적인 개선을 이끌어낼 동력이 부족할 수 있다. AI 서비스 초기 ‘설계’ 단계부터 접근성을 고려하도록 하는 법적 표준화가 요구된다.

3) 문제의 심각성

AI 서비스의 접근성 부족은 고령인과 장애인의 정보 접근권, 사회 참여권, 자기결정권을 직접적으로 침해할 수 있으며, 타인의 도움 없는 독립적인 생활을 더욱 어렵게 만들 수 있다. 또한 AI 기술이나 서비스의 설계가 비장애인·청년층 등 일

반 국민에게만 맞춰질 경우, 기존의 사회적 차별이 알고리즘 영역으로 그대로 전이되거나 오히려 증폭되는 악순환이 발생할 수 있다. 접근성이 보장되지 않는 AI 서비스의 확산은 결국 대안 서비스 제공 비용, 사회적 격차 해소 비용 등 사회 전체의 총체적 비용을 증가시킬 것이다.

4) 기존 논의/대책의 한계

현행 「지능정보화기본법」 등은 주로 정적인 웹사이트의 접근성을 다루고 있다. 이는 챗봇, AI 추천, 음성인식 비서 등 상호작용이 복잡한 AI 서비스의 고유한 장벽(예: 이해하기 어려운 AI의 응답, 잘못된 음성인식 등)을 포괄하지 못한다. 또 AI 기술의 빠른 발전으로 인해 대부분의 대응책은 유연하지만 법적 구속력이 없는 ‘가이드라인’과 같은 자율규제 형식을 취한다. 이는 기업의 자발적 참여에 의존하는 취약성이 있으며, 역동적인 변화 시기에도 고령인·장애인의 권리를 실질적으로 보장할 ‘최소한의 법적 표준’을 모색하고 개발하려는 정책적 노력이 필요하다.

마지막으로, 현시점에서는 AI 서비스가 출시되기 전후에 접근성 수준을 체계적으로 인증하고 관리·감독할 전담 기구나 거버넌스가 부재하다. 「AI 기본법」 실행 이후는 달라질 수 있으나 영향평가 등 기업 의무 내역에 포용적 접근성 수준을 주요 항목으로 하고 있지 않으면 표준 AI 서비스에서 고령인·장애인이 여전히 배제될 수 있을 것이다.

다. (이용자03) [AI 활용 경험과 연계된 체험형 윤리교육 설계·확산] 디지털 취약계층이 AI에 대한 과도한 의존이나 막연한 두려움 없이 AI를 비판적으로 평가하고 윤리적 문제를 인지할 수 있도록, 활용과 연계된 체험형 윤리교육을 어떻게 설계하고 보급할 것인가?

1) 제안의 배경

새로운 기술이 도입되는 초기 시점에는 신기술 채택을 두려워하거나 거부하는 사람, 혹은 과도하게 의존하는 사람 등이 존재한다. 현시점에서는 AI에 대한 막연한 불안과 거부감으로 기술 활용 기회를 스스로 피하는 경우도, 충분한 이해 없이

AI 서비스에 과도하게 의존하는 경우가 있으며, 이는 심각한 사회문제로 연결될 수 있다. 이러한 경향은 사회적으로 새로운 경험을 빨리 접하고 학습할 기회가 적은 고령인들에게 취약한 지점이다.

2) 현황 및 쟁점

학습자 중 고령인 비율이 높은 디지털배움터 현황 등을 보면 AI 기술 도구 활용 등에 대한 고령인들의 학습 욕구와 교육 수요가 있음을 알 수 있다. 그러나 고령인들의 AI·디지털 교육 수요를 체계적으로 파악하기 위한 전국 단위의 실증 조사는 부재하다.

또한 체험형 교육이라는 유용한 교육 유형을 발굴하였으나, 전국 단위로 고령인의 참여를 독려하기에는 실제 AI 서비스 실습 환경, 예시 데이터, 윤리적 상황 평가 도구 등의 체계적 확산 등 중장기 계획 등을 적극 추진할 사회적 동력이 부족하고, 이에 대한 정책 지원도 미비하다. 전문가 인터뷰에 따르면, 단기적으로는 디지털배움터에서 모범교재를 개발하고 있고 이를 온라인으로 배포할 예정이지만 이 성과가 전국 단위로 확산되기 위한 체계적 후속 조치가 부재하다. 또 온라인 형식의 배포는 디지털 접근 자체를 어려워하는 취약층에게는 여전히 접근 장애 요소일 수 있다.

3) 문제의 심각성

일상생활 서비스의 많은 부분이 비대면 디지털 서비스로 전환되고 있고, AI가 다양한 디지털 서비스와 결합되면서 신기술의 복잡성과 복합성을 이해하고 활용해야 하는 사회적 필요가 발생하고 있다. 하지만 경험과 학습 기회가 상대적으로 낮은 고령인들은 이러한 변화에 적응하는 데 어려움을 겪고 있다. 오프라인 서비스 비중이 높았거나 비슷한 비중으로 선택할 수 있던 과거와는 달리, 오프라인 서비스 비중이 빠르게 감소하는 사회 변화 속에서 고령인들이 경험하는 신기술 기반 서비스에 대한 두려움은 더 커질 수 있으며, 그러한 심리적 저항감으로 인한 기술 격차의 크기 역시 과거 어느 때보다 심각할 수 있다.

4) 기존 논의/대책의 한계

현시점에서는 AI 오남용을 예방하거나 기술을 비판적으로 해석하는 능력을 키우는 체험형·사례 기반 교육 모델이 부족하며, 교육기관 간 표준화된 커리큘럼도 부재하다. 고령인 대상으로 활용과 연계된 체험형 윤리교육 교재가 개발되고 있지만, 이와 별도로 취약계층의 특성을 반영한 맞춤형 윤리교육, 예를 들어 시각·청각 장애인을 위한 음성 기반 실습 환경, 고령층을 위한 단순화된 인터페이스 기반 학습, 일상 서비스 상황(예: 의료, 교통, 금융 등)별 시나리오 기반 훈련 등이 충분히 제공되지 않고 있다.

라. (이용자04) [노동 시장 재편에 대비한 취약계층 맞춤형 AI 활용 역량 재교육] AI 전환으로 고용 위협에 노출된 저숙련/장애인 노동자들을 대상으로, 생계형 일자리 유지 및 확산을 위한 맞춤형 AI 활용 교육 및 인간-AI 협업 기술 개발을 어떻게 지원할 것인가?

1) 제안의 배경

인공지능은 고용 시장에 변혁을 가져오고 있고, 기업들은 AI 기술 부족 해소를 위해 직원들의 업스킬링(upskilling)과 재교육(reskilling)을 요구하고 있다. AI 시대로의 사회 전환 과정에서 모든 국민이 AI 활용 교육 기회를 동등하게 누려야 하나, 특히 저숙련 노동자나 장애인들은 AI 대체율이 높은 일자리에 종사하는 와중에 변화하는 노동 환경에 적응하거나 독립적인 삶을 유지하기 위해 필요한 새로운 역량을 습득할 기회는 매우 부족하다.

2) 현황 및 쟁점

현재 정부는 K-디지털 트레이닝(국민내일배움카드) 등을 통해 AI 재교육을 지원하고 있으나, 저숙련·장애인 노동자가 실제 현업에서 AI를 활용하는 데 필요한 교육은 부족하다. 또한 장애인 노동자를 위한 교육은 AI 도구의 접근성(accessibility) 문제도 충분히 해결하지 못하고 있다. 예를 들어, 스크린 리더 호환성, 음성 명령 지원 등 장애 유형별 특성을 고려한 교육 과정 및 보조 기술 개발이 미흡한 실정이다.

3) 문제의 심각성

AI로 인한 노동 시장 재편 과정에서, 저숙련·장애인 노동자 계층의 실직이 현실화될 수 있다. 이는 단순히 개인의 경제적 문제를 넘어, 사회 양극화 심화, 복지 비용 급증, 계층 간 갈등 등 심각한 사회적 비용을 초래할 수 있다. 특히 장애인 노동자의 경우, AI가 새로운 장벽으로 작용하여 어렵게 확보한 경제적 자립 기반을 잃게 될 위험이 있다. 기술이 이들의 격차를 해소하는 것이 아니라 오히려 격차를 고착화하는 ‘기술적 배제’가 발생할 수 있는 것이다.

4) 기존 논의/대책의 한계

저숙련·장애인 노동자의 실제 직무(예: 서비스직, 단순 제조직)에 대한 커리큘럼이 부재하다. 이러한 상황에서 대부분의 재교육 정책이 새로운 AI 관련 일자리의 전환을 목표로 하는데, 저숙련·장애인 노동자에게는 현실적 대안이 될 수 없다. AI 서비스에 대한 이들의 접근성과 사용성이 우선순위에서 밀려나 있는 동안 격차는 더욱 심화될 것이다. 인간-AI 협업 기술을 고려할 때 장애인의 직무를 돕는 보조공학적 AI 기술에 대한 R&D 지원 및 정책적 유인이 필요한 시점이다.

마. (이용자05) [지역·계층 간 AI 리터러시 교육 격차 해소] 도시와 농촌, 소득 계층, 교육 수준에 따른 AI 리터러시 교육 기회의 격차를 해소하고, 디지털 취약계층이 거주지나 경제적 여건에 관계없이 양질의 AI 윤리 및 활용 교육에 접근할 수 있도록 하는 통합적 교육 인프라와 전달 체계를 어떻게 구축할 것인가?

1) 제안의 배경

인공지능 확산으로 AI 리터러시의 중요성이 커지고 있지만, 교육 접근성은 지역·계층별로 큰 격차를 보인다. 특히 고령층과 장애인은 물리적 이동의 어려움에 비해 대도시와는 달리 교육 장소인 복지관, 도서관 등 공공시설과 거주지 간 거리가 있고 대중교통도 부족하여 교육 기회가 제한된다. 교통비·교육비 등 경제적 부담도 저소득층의 참여를 가로막는다.

지역 간 교육 인프라 차이와 강사의 역량 편차로 인해 동일한 AI 리터러시 교육이라도 질과 내용이 크게 달라진다. 수도권과 광역시는 전문 강사 확보가 상대적으로 쉽지만, 지방 소도시나 농촌 지역에서는 디지털 교육의 전문성을 갖춘 강사를 찾기 어렵다. 온라인 교육이 대안으로 제시되지만, 디지털 기기 부족, 인터넷 환경 미흡, 기초 역량 부족 등으로 취약계층에게는 또 다른 접근 장벽이 될 수 있다.

이러한 격차는 단순한 교육 기회의 문제가 아니라 인공지능 시대 시민의 기본권 보장과 직결된다. 따라서 거주지나 경제적 여건과 무관하게 누구나 양질의 AI 리터러시 교육에 접근할 수 있도록 통합적·체계적인 교육 인프라 구축이 시급하다.

2) 현황 및 쟁점

취약계층 디지털·AI 교육의 대표 사업인 ‘디지털배움터’의 경우 도시 지역에 편중되어 있거나 온라인 과정 비중을 늘리고 있다. 코로나·팬데믹 기간을 거치면서 교육 신청 과정도 대부분 온라인으로 진행되고 있다. 읍·면 단위 지역에서는 물리적 교육 인프라뿐만 아니라 교육을 기획, 운영할 인적 자원도 상대적으로 부족하다. 예를 들어, 경남인재평생교육진흥원(2023. 11)⁶⁰⁾에 따르면, 경남 도내 읍·면 단위, 심지어 시·군 단위에서도 평생교육 전담 인력이 1명에 불과한 곳이 80%에 달했다고 한다. AI 리터러시와 같은 새로운 전문 교육을 기획하고 실행할 인적 자원이 구조적으로 부족할 수밖에 없는 상황인 것이다.

3) 문제의 심각성

현재의 교육 격차가 그대로 지속된다면 도시와 농촌, 수도권과 비수도권 간 AI 리터러시 격차는 점점 더 심각한 수준으로 벌어질 수 있다. 또 의료, 금융, 행정 서비스 등 필수 서비스가 디지털로 전환하고 인공지능을 탑재하는 추세에서 농촌 지역의 고령 인구는 AI 기술 혜택에서 배제되고 필수 서비스 이용에서도 어려움을 겪을 것이고, 이는 지역 소멸 가속화 및 국토의 균형발전 저해로 이어질 수 있

60) 경상남도인재평생교육진흥원(2023. 11), “찾아가는 시·군 평생학습 모니터링단 운영 결과보고서”.

다. 이들의 기술 배제는 디지털 기록정보 부족으로 인해 AI 정책 결정 과정에서도 배제되어 이들의 이익이 제대로 반영되지 않을 위험으로도 이어진다.

4) 기존 논의/대책의 한계

현재의 AI 교육 관련 대다수 프로그램은 단기·단발성 강의 중심, 기초 사용법 전달에 머무르는 수준으로 운영되고 있어서 이용자의 리터러시 및 실제 역량 향상에 충분하지 않다. 또 다양한 교육 실행기관(예: 중앙부처, 지자체, 평생교육기관, 민간 교육기관 등)이 유사한 수준의 교육을 반복 제공하는 경향이 있다. 누적 학습이 가능한 연속적인 교육체계가 거의 존재하지 않는 것은 조속히 해결해야 할 과제이다.

온라인 교육 확산 역시 취약계층의 접근성 문제를 고려하지 못하는 문제가 있다. 기초 디지털 역량이 부족한 고령층·장애인에게는 온라인 플랫폼 자체가 학습 장벽이 되며, 이에 대한 보조 설계 없이 기존 콘텐츠만 온라인으로 전환하는 방식으로서는 실질적 교육 효과가 제한적일 것이다.

제 4 절 AI와 정보: 정보 생태계 변화와 상호작용

1. 논의의 배경

가. 생성형 AI의 발전

AI 기술은 기계학습, 딥러닝, 자연어 처리, 컴퓨터 비전 등 핵심 분야를 중심으로 전례 없이 빠른 속도로 발전하고 있다. 특히 2023년 ChatGPT 출시 이후, 생성형 AI를 활용한 다양한 서비스가 등장하였으며, 맞춤형 서비스에 대한 이용자 선호 또한 지속적으로 확대되고 있다(이경선, 2023). ‘2024 지능정보사회 이용자 패널조사’에 따르면, 생성형 AI 서비스를 이용한 경험이 있다고 응답한 비율은 2023년 12.3%에서 2024년 24.0%로 약 두 배 증가하였다. 또한 생성형 AI 서비스 유료 구독 경험은 0.9%에서 7.0%로, 약 7배 증가한 것으로 조사되었다(방송통신위원회·정보통신정책연구원, 2025). 국내 기업의 생성형 AI 도입 역시 빠르게 확산되고 있다. 아마존웹서비스(AWS)가 발표한 ‘한국 생성형 AI 도입 지수’에 따르면, 국내 기업 94%가 생성형 AI 서비스 도입·활용을 시작하였으며, 63%는 AI를 활용해 기업 내 새로운 가치 창출과 혁신 역량 가속화를 주도할 최고AI책임자(CAIO) 직책을 신설한 것으로 나타났다(전자신문, 2025. 5. 13).

생성형 AI의 발전은 과거 고도의 프로그래밍 언어를 구사하는 전문 프로그래머만 수행할 수 있었던 작업의 진입 장벽을 획기적으로 낮추며 산업 구조 전반에 변화를 촉발하고 있다. 이제 일반 이용자는 별도의 프로그래밍 역량 없이도 생성형 AI 서비스를 활용하여 텍스트(ChatGPT, Gemini, Claude), 음성·음악(Suno AI, Synthesizer V, AIVA), 이미지(DALL·E, Midjourney, Stable Diffusion), 동영상(Runway Gen-3 Alpha, Pika Labs AI, Synthesia, D-ID), 코드(GitHub Copilot, Amazon Q Developer, Replit Ghostwriter) 등 다양한 형태의 데이터를 손쉽게 생성하거나 변환할 수 있게 되었다. 이러한 변화는 개인의 업무 활동을 자동화하는 기술 전반의 진전을 가속화한다. 맥킨지(McKinsey)는 2017년 기준으로 업무 활동 시간의 50%가 자동화되는 시점을 2053년으로 전망

하였으나, 2023년 동일한 시나리오에 생성형 인공지능을 반영한 결과, 그 시점을 2045년으로 추정하여 기존 전망보다 약 10년 앞당겨질 것으로 분석하였다(McKinsey, 2017; 2023). 이는 생성형 AI가 노동시장에서 인지 자동화를 촉진함으로써 자동화 기술 채택 시기를 앞당기고, 단순·반복적인 업무를 대체함으로써 인간이 보다 복잡하고 창의적인 과업에 집중할 수 있는 환경을 조성하고 있음을 시사한다(Demirci et al., 2025). 나아가 여러 실증 연구는 생성형 인공지능이 단순히 노동의 양이나 강도를 경감하는 것을 넘어 업무의 질과 생산성에도 긍정적인 영향을 미치는 것을 확인하였다(Brynjolfsson et al., 2023; Noy & Zhang, 2023; Peng et al., 2023, Kim & Moon, 2024).

나. 생성형 AI 결과물과 정보생태계의 위기

그러나 생성형 인공지능 기술의 발전은 동시에 우리 사회에 중대한 사회적·윤리적 과제를 제기하고 있다. 먼저, 학습 데이터 수집·분석 과정에서 가장 빈번히 지적되는 문제는 기존 데이터에 내재한 편향이 결과물에 그대로 반영되어 차별과 혐오를 재생산할 수 있다는 점이다. 대표적으로, 2016년 마이크로소프트가 출시한 챗봇 ‘테이(Tay)’는 사용자와의 상호 작용을 통해 자율 학습하도록 설계되었으나, 일부 사용자가 인종·성차별적 발언과 근거 없는 정치적 주장을 입력하자 이를 모방·확산하여 혐오 발언을 대량 생성하였고, 결국 출시 16시간 만에 서비스가 중단되었다(연합뉴스, 2016. 3. 25). 2021년에는 스캐터랩이 개발한 챗봇 ‘이루다’가 20대 여성이라는 페르소나를 기반으로 특정 연령과 성별에 대한 왜곡된 고정관념을 재현하고, 장애인·성소수자·인종 등 소수 집단에 대한 혐오와 편견을 드러냈다는 비판을 받으며 서비스가 중단되었다(연합뉴스, 2021. 1. 11). 최근에는 xAI가 개발한 챗봇 ‘그록(Grok)’이 유대인을 겨냥한 혐오 발언과 히틀러를 찬양하는 발언을 게시하여 국제적 비판을 초래하기도 하였다(연합뉴스, 2025. 7. 13).

아울러 최근 주목을 받은 ChatGPT-4o의 이미지 생성 기능은 생성형 AI의 학습 과정에서 발생할 수 있는 저작권 침해 가능성을 단적으로 보여준다. 해당 기능은 이용자가 직접 촬영한 사진을 업로드하고 특정 애니메이션 스타일로 변환을

요청하면 약 1분 내에 지브리, 디즈니, 픽사 등 특정 제작사의 고유한 화풍을 모사한 이미지를 생성·제공한다. 이 기능이 공개된 이후 전 세계적으로 ‘지브리풍’ 이미지 제작이 유행하였고, 소셜미디어에는 관련 이미지가 대량 공유되었다. 그러나 이러한 기술은 특정 작가나 제작사의 고유한 미술 스타일과 캐릭터 디자인을 원저작자의 사전 동의 없이 학습·재현함으로써, 해당 저작물의 변형·파생 작품이 무분별하게 생산·유통되는 문제를 초래할 수 있다. 지브리 스튜디오의 미야자키 하야오 감독은 인공지능이 제작한 애니메이션에 대해 “생명 자체에 대한 모독”이라고 강하게 비판한 바 있으며(이코노미스트, 2017. 1. 2), 멕시코 맥도날드가 지브리풍 이미지를 활용한 상업 프로모션을 진행하면서도 스튜디오 지브리에 어떠한 승인이나 대가를 제공하지 않았다는 점을 지적하며 생성형 AI 활용과 관련한 저작권 침해 및 윤리적 쟁점에 대한 사회적 관심을 촉발하였다(뉴스워커, 2025. 4. 2).

생성형 AI의 저작권 이슈와 관련하여, 2023년 1월 켈리 맥커넌(Kelly Mckernan), 칼라 오티즈(Karla Ortiz), 사라 앤더슨(Sarah Andersen) 등 저명한 아티스트들은 이미지 생성 인공지능 서비스를 운영하는 미드저니(Midjourney), 스테빌리티 AI(Stability AI), 디비언트아트(DeviantArt)를 상대로 저작권 침해 소송을 제기하기도 하였다. 원고 측은 해당 기업들이 원작 아티스트의 사전 동의 없이 이들의 화풍과 창작 스타일을 학습 데이터로 무단 활용하여, 이를 기반으로 유사한 새로운 이미지를 생성·배포함으로써 창작자의 권리를 침해하였다고 주장하였다. 특히 AI 이미지 생성기는 작가의 작품을 분해·재조합하여 새로운 창작물을 만드는 ‘콜라주(collage) 도구’에 해당하며, 이에 대한 어떠한 동의나 보상도 이루어지지 않았다는 점을 강조하였다(심나리, 2023). 뿐만 아니라, 2025년 6월에는 디즈니(Disney)와 유니버설(Universal)이 미드저니를 상대로 별도의 소송을 제기하였다. 두 기업은 피고가 ‘아이언맨’, ‘헐크’, ‘슈렉’, ‘미니언즈’ 등 자사의 대표 캐릭터를 학습 데이터로 무단 사용하였으며, 사용자가 단순한 프롬프트 입력 만으로도 해당 캐릭터와 동일하거나 현저히 유사한 이미지를 생성할 수 있도록 하여 직접적·간접적인 저작권 침해를 저질렀다고 주장하였다(이철남, 2025).

무엇보다 생성형 AI가 만들어낸 산출물, 특히 딥페이크(Deepfake) 기술을 활용해 실제와 구분하기 어려운 음향·이미지·영상 등을 사실인 것처럼 조작하는 행위는 개인의 명예와 권리를 심각하게 침해할 뿐만 아니라 사회적 혼란을 초래할 수 있다. 딥페이크를 이용한 불법 음란물 제작은 이미 중대한 사회 문제로 부상하였으며, 이외에도 금융·투자 사기, 보이스피싱 등 각종 범죄, 정치·사회적 여론 선동, 허위조작정보의 제작·유포 등 다양한 분야에서 AI 기술이 악용되는 사례가 증가하고 있다(문정욱 외, 2024).

이처럼 생성형 AI는 다양한 산업 분야에서 혁신과 효율성을 제고하고 창의적 활용을 가능하게 하는 등 긍정적인 효과를 제공하는 한편, 차별과 혐오의 재생산, 저작권 침해, 진실과 거짓의 경계 모호화와 같은 새로운 사회적 문제를 야기하고 있다. 특히 딥페이크를 포함한 AI 생성물의 확산은 정보의 생산·유통 구조를 왜곡하고 정보 생태계 전반의 신뢰를 잠식시키는 잠재적 위험을 내포하고 있다.

다. 생성형 AI 규제 동향

생성형 AI의 양면적 특성은 기술의 건전한 발전과 사회적 신뢰 기반 조성을 위한 규제의 필요성을 강력하게 뒷받침하는 근거이다. AI의 혁신 잠재력과 위험성에 대한 공감대는 각 국가가 마련한 AI 규제를 위한 법적 틀에서도 확인할 수 있다. EU는 「AI 법」에서 위험 수준에 따라 AI 시스템을 분류하고, 그에 상응하는 규제 수준을 정하고 있으며, 주로 기본권 보호 및 역내 통상 보호에 중점을 두고 있다.

미국은 2025년 초 트럼프 행정부의 전향적인 행정명령을 기점으로 바이든 행정부의 안전성 중심 접근과 대조되는 혁신 우선의 접근 방식을 취하고 있다. 이는 2025년 7월 백악관이 발표한 “미국 AI 액션플랜(America’s AI Action Plan)”에 잘 나타나고 있다. 액션플랜이 생성형 AI에 대해 제시한 유일한 기준은 최신 AI가 “자유 언론과 미국의 가치를 보호”하여야 한다는 점이다. 이는 연방 정부의 계약 상대를 이데올로기적 편향이 없는 시스템을 보장하는 대규모 언어모델(LLM) 개발 업체로 한정함으로써 규제 방향을 구체화하고 있다. 그 밖에 액션플랜은 개발과 배포를 촉진하기 위해 오픈소스 및 오픈웨이트 AI 모델을 장려하고 연방정부 내

생성형 AI 활용을 의무화하는 등 적극적 활용 정책을 추진하고 있다.

생성형 AI 규제와 관련하여 중국은 선도적 지위를 점하고 있으며, 개발 및 발전을 추구하고 있다. 글로벌 규제 동향에서 확인할 수 있듯이, AI 규제의 필요성에 대해서는 모든 국가가 합의하고 있다. 다만, 혁신 촉진과 기본권 보호 중 어떤 가치에 정책적 우선순위를 둘 것인지에 따라 구체적인 전략이 다르게 나타나고 있다.

이러한 글로벌 흐름 속에서 대한민국의 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법(이하, 인공지능기본법)」⁶¹⁾도 독자적인 균형점을 모색하고 있다. 「인공지능기본법」은 2024년 12월 26일 국회 본회의를 통과하여, 2025년 1월 21일 공포되었으며 2026년 1월 22일 시행을 앞두고 있다. 법의 목적은 인공지능의 건전한 발전과 신뢰 기반 조성을 위한 기본사항을 규정하여 국민의 권익과 존엄성을 보호하고 삶의 질 향상 및 국가경쟁력 제고에 기여하는 데 있다. 「인공지능기본법」은 진흥과 규제 양 측면을 모두 고려하여 AI 기술의 산업적 성장과 신뢰 기반 확보를 동시에 달성하기 위한 최초의 포괄법, 세계 최초의 ‘AI 기본법’이라는 상징성을 가진다. 그러나 법률은 여러 지점에서 비판의 대상이 되고 있으며, 법률이 시행되기 전부터 학계에서는 개정의 목소리가 높아지고 있다. 대표적으로 고영향 AI의 불분명한 정의를 예로 들 수 있다(정찬모, 2025). 법률은 고영향 AI를 정의하고 있지만(법 제2조 제4호), 지나치게 추상적으로 표현되어 있어 법적 불확실성 및 그로 인한 스타트업들의 AI 분야 기술 도입과 혁신 투자 위축이 문제점으로 지적되고 있다(박상철, 2024). 안전성·신뢰성 검·인증제도에 대해서도 규정의 구체성이 부족하다는 비판은 그대로 유효하다. 이에 덧붙여 검·인증 과정에서 개인정보 및 민감정보 유출, 사이버보안 위협 등의 부작용도 개선이 필요한 사항으로 언급되고 있다. 유사한 지적은 생성형 AI 관련 표시의무에서도 발견할 수 있다. 법률은 적용 범위와 기준에 대하여 명확한 방향을 제시하고 있지 못하며(정찬모, 2025), 생성물에 대한 일률적 적용에 따른 창작활동 위축이 우려된다(박상철, 2024).

법 개정을 통한 구조적 문제 해결이 가장 근본적인 접근방식임에는 틀림이 없

61) [시행 2026. 1. 22.] [법률 제21311호, 2026. 1. 20., 일부개정]

다. 그러나 아직 시행령을 통한 법률의 구체화 여지가 있고, 가이드라인을 통해 실무상의 불확실성을 해소할 가능성도 남아 있다. 다만, 그 과정에서 참고할 수 있는 풍부한 실증 모델 확보가 우선 이루어졌을 때 비로소 「인공지능기본법」의 본래 취지에 부합하며, 우리 AI 산업의 특수성을 반영하는 규제의 합리화도 기대할 수 있을 것이다.

라. 연구 목적

본 연구는 「인공지능기본법」 제31조(인공지능 투명성 확보 의무)를 헌법적 관점에서 분석하여, 해당 조항의 적용 과정에서 충돌이 예상되는 법익을 식별하고, 이익 조정의 최적화 방안을 도출하기 위한 실증적 근거를 탐색하는 것을 목적으로 한다.

1) 법제 분석의 필요성

「인공지능기본법」 제31조는 생성형 AI의 투명성 확보 조치를 다루고 있으며, 제2항은 AI 사업자에 대한 표시의무를, 제3항은 소위 ‘딥페이크’(엄주희, 2025)에 대응하는 AI 사업자의 고지 또는 표시의무를 규정하고 있다. 생성형 AI는 입력 데이터를 기반으로 글, 소리, 그림, 영상 등 다양한 형태의 결과물을 생성하는 AI 시스템을 가리키며, 제2항의 목적은 이용자가 생성형 AI 기반 서비스임을 명확히 인지하도록 하여 정보 비대칭성을 해소하고, 잠재적 위협에 대한 사전적 인식을 확보하는 데 있다. 제3항은 실제와 구분하기 어려운 AI 생성물의 영향을 감안하여 제2항에 비해 강화된 조치인 ‘명확한 고지 또는 표시’를 규정하고 있다. 이는 AI 생성 콘텐츠의 진위를 명확히 하여 허위 정보 유포나 오남용으로 인한 사회적 피해를 방지하는 데 목적을 둔다. 과학기술정보통신부 장관은 의무 위반 사실을 알게 되거나 신고를 받은 경우, 소속 공무원으로 하여금 사실조사를 개시하게 할 수 있으며(제40조 제1항), 법 위반 사실이 인정되면 시정조치를 명할 수 있다(동조 제3항). 그리고 시정명령을 이행하지 않은 경우에는 3천만 원 이하의 과태료가 부과된다(제43조 제1항 제3호).

특히 본 연구에서는 법 제31조 제2항 및 제3항의 투명성 확보 의무 내용을 분

석하는 데 초점을 맞추고자 한다. 「인공지능기본법」은 사실조사 및 과태료 부과라는 제재를 규정하면서 의무의 내용은 정교하게 설계하지 않고 있다. 물론 규제 설계자는 법률에(명확한) 표시 또는 고지의무를 명시하는 것으로 충분하다고 판단할 수도 있다. 그러나 이러한 판단이 AI 사업자의 입장과 반드시 부합한다고 볼 수는 없다. 오히려 의무 내용의 불명확성은 AI 기술의 현주소와 산업의 특성을 지나치게 단순하게 파악한 결과라는 비판을 불러일으킬 수 있다. AI의 역량과 그 파급효과를 충분히 이해했다면, 어느 수준으로 투명성 확보 조치를 이행해야 하는지 명확한 기준을 제시하여야 하기 때문이다(박상철, 2024). 그 결과 AI 사업자는 법적 불안정성을 안고 투명성 확보 조치 계획을 수립할 수밖에 없으며, 과도한 부담을 안게 된다. 문제는 여기에 그치지 않는다. 명확한 기준 없는 투명성 확보 조치 의무는 어떠한 형태로든 조치를 취하기만 하면 의무 이행에 충분하다는 신호를 시장에 전달하여, 애초의 취지인 기본권 보호도 전혀 달성하지 못할 수 있다.

규제의 본래 취지를 달성하기 위해서는 구체적인 사례 검토에 기반하여 실효적인 표시 수단 및 적용 방법을 도출하고, 이를 규범 정비 과정에 반영하여 규제 합리화를 이루어야 한다. 선행 연구들도 「인공지능기본법」 제31조에 의해 부과된 의무의 의미와 주체를 분석하는 것뿐만 아니라, 투명성 확보 조치의 구체화 필요성을 강조하고 있다. AI 생성물 표시의무는 결과물의 신뢰성 및 투명성 확보를 위한 핵심 조치이기는 하지만, 규정의 문언 및 체계 측면에서 주체별 역할 혼선, 불분명한 표현, 하위규정의 부재 등으로 인한 실효성 저하가 우려된다는 점을 지적하고 이용자 중심의 명확한 의무 부과 및 구체적 표시 수단 마련의 필요성을 역설하는 연구가 그 예이다(박상철, 2024). 또한 AI 생성물 표시의무와 관련된 입법과정 분석을 통해 적용 범위의 단계적 확대 경향을 확인한 후, 향후 법적 리스크 관리를 위해 국제적 규제 동향 및 「저작권법」과 「콘텐츠산업진흥법」 등 타법 개정 논의에 주목할 것을 강조하는 연구도 투명성 확보 조치의 구체화를 중심에 두고 있다(김형지·정선아, 2025).

물론 의무 내용의 명확성은 규제의 실효성과 직접 관련되어 있다. 누가, 무엇을 행해야 하는지 분명하지 않다면, 법률을 통해 의무를 부과한 취지인 투명성 제고

및 기본권 침해의 예방을 달성하기 어려울 뿐만 아니라, 실무에 혼란을 가져올 수 있기 때문이다. 특히 표시의무 제도의 실효적 적용에 필요한 ‘투명성 확보 조치를 어떻게 이행해야 하는지’는 거센 비판에도 불구하고 명확한 해결 방안 없이 남아 있다. 이때 실효성은 단순히 이용자 또는 대중이 AI 생성물임을 쉽게 확인하는 것을 의미하지 않는다. 표시방법의 과도한 강제는 AI 사업자의 기본권을 침해하는 위헌적 공권력 행사로 평가될 수도 있고, AI 기반 창작 산업을 위축시켜 AI 기술 발전을 저해하는 장애물로 작용할 수도 있다(손영선, 2024). 그러므로 핵심은 AI 사업자의 기본권을 최소로 침해하면서, 동시에 표시의무 제도를 도입한 의도를 달성하는 방법을 모색하는 것이다(엄주희, 2025).

2) AI 생성물 투명성에 대한 국민 인식 조사의 필요성

아울러 본 연구는 실제 국민들이 인식하는 AI 투명성의 의미와 필요성을 실증적으로 조사함으로써 해당 법의 제정 취지에 부합하는 방향으로 이를 구체화하고 발전시킬 수 있는 방안을 모색하고자 한다. 인공지능의 투명성 확보에 관한 법적·규범적 논의는 학계와 산업계에서 이미 활발하게 전개되고 있으나, 우리 국민이 인공지능 투명성의 의미와 필요성을 어떻게 인식하고 있는지에 대한 실증 조사 결과는 부족한 실정이다. 이러한 인식 조사는 다음과 같은 이유에서 필수적이다.

첫째, 정책 수용성 확보를 위함이다. 인공지능 투명성 관련 규제는 국민의 지지와 신뢰를 바탕으로 해야만 실효성을 담보할 수 있다. 정책 방향과 국민 인식이 괴리될 경우 법 집행 과정에서 저항과 불신이 발생하여 정책 효과가 저하될 수 있기 때문이다. 예컨대 2024년 7월 메타(Meta)는 AI 투명성 강화를 목적으로 ‘AI 생성물’ 라벨을 도입하였다. 그러나 실제로는 사진작가가 직접 촬영한 사진에까지 ‘AI로 제작됨(Made with AI)’ 표기가 부착되는 사례가 다수 보고되면서, 현업 사진가들의 반발과 대중의 혼선을 초래하였다. 이는 사진 자르기, 단순 잡티 제거 등과 같은 전통적인 사진 보정 작업마저 AI 생성물로 분류한 데에서 비롯된 것으로 밝혀졌다. 사진작가 노아 칼리나(Noah Kalina)는 “단순 보정 사진까지 ‘AI로 제작됨’으로 표기한다면 해당 용어는 사실상 무의미해진다”고 비판하였다(IT동아,

2024. 7. 8). 이러한 사례는 AI 생성물의 범위를 어디까지로 설정할 것인지, 그리고 국민이 어느 수준의 AI 활용부터 그 과정을 투명하게 공개받기를 원하는지라는 핵심 쟁점을 부각시키며, 향후 인공지능 투명성 제도의 설계와 집행 과정에서 반드시 반영되어야 할 정책적 과제를 시사한다.

둘째, 규제 강도와 범위 설계의 근거 제공이다. 국민이 기대하는 투명성 수준과 정보 공개 범위를 체계적으로 파악하는 것은 「인공지능기본법」 제31조의 구체적인 이행 방안을 설계하는 과정에서 필수적인 선행 단계이다. Wittenberg et al.(2024; 2025)은 생성형 AI 표시 정책을 수립하기에 앞서, 이해관계자들이 해당 정책이 달성하고자 하는 목표를 명확히 규정해야 하며, 설정된 목표에 따라 정책 설계 방향이 달라져야 한다고 강조한다. 이들은 AI 생성물 표시 정책의 목표를 크게 두 가지로 구분한다. 첫째, 절차 중심(process-based) 접근이다. 이는 콘텐츠가 제작·편집되는 과정에서 생성형 AI 도구가 사용되었는지를 명확히 밝히는 방식으로, 콘텐츠의 주제나 오도 가능성과 관계없이 모든 형태의 AI 생성물에 적용될 수 있다. 둘째, 위해성 중심(harm-based) 접근이다. 이는 콘텐츠의 사회적 위해 가능성, 즉 대중을 오도하거나 피해를 야기할 위험성이 높은 경우에 한정하여 표기를 부여하는 방식이다. AI 생성물 표시에 대한 사회적 기대 수준과 수용 가능성을 실증적으로 규명하는 것은, 우리 사회가 절차 중심 접근과 위해성 중심 접근 중 어느 방식을 보다 바람직하다고 인식하는지, 그리고 두 방식의 절충점을 모색할 경우 어떤 요소를 보다 중점적으로 고려해야 하는지에 관한 합리적 판단 근거를 제공하며, 궁극적으로 AI 생성물 표시 정책의 방향성과 세부 설계에 기여할 수 있다.

셋째, 사회적 신뢰 형성 및 민주적 거버넌스 구현을 위한 기반 마련이다. 기술 발전 속도가 빠른 인공지능 분야에서 투명성 논의가 전문가나 정책결정자 중심에만 머무를 경우, 일반 국민은 정책 형성 과정에서 소외될 우려가 있다. 국민 인식을 정책 설계에 반영함으로써 AI 거버넌스는 ‘위에서 부과하는 규제’가 아닌 ‘사회적 합의에 기반한 규범’으로 자리매김할 수 있으며, 이는 장기적으로 인공지능 활용 전반에 대한 신뢰를 강화하는 핵심 요소가 된다. 특히 AI 기술의 고도화에

따라 AI 생성물 탐지가 점점 어려워지는 상황에서, 사회적 합의 없이 도입된 AI 생성물 표시 의무는 실효성을 확보하기 어렵다. 표시 의무가 과도할 경우 불필요한 규제 부담과 사회적 반발을 초래할 수 있으며, 반대로 과소할 경우 사회적 위해 가능성이 높은 AI 생성 콘텐츠에 대한 경각심을 약화시켜 제도의 목적을 훼손할 위험이 있다.

무엇보다 콘텐츠 창작자가 AI 활용 여부를 자발적으로 표시할 유인이 부족한 현실을 고려할 때, 적절한 규제 수준을 도출하는 것은 제도의 수용성과 지속가능성을 담보하기 위해 매우 중요한 과제이다. 다수의 실증 연구 결과, AI 생성물 표시는 소비자로서 하여금 해당 콘텐츠의 신뢰도와 창의성을 낮게 평가하게 만드는 경향이 나타났다. 예를 들어, 고객 응대에서 챗봇임을 공개하지 않은 경우, 챗봇은 숙련된 직원과 유사한 수준의 구매 유도 효과를 보였으며, 경험이 부족한 직원보다 4배 높은 성과를 기록했다. 그러나 동일한 상황에서 챗봇임을 공개하면 구매율이 79.9% 감소하였다(Luo et al., 2019). 또한 AI를 활용하여 제작된 광고에 AI 생성물 표시가 부착될 경우, 소비자의 광고 신뢰도 인식이 낮아져 전반적인 광고 태도가 부정적으로 변화하는 경향이 나타났다(Baek et al., 2024). 글쓰기와 같은 창작 분야에서도, AI 활용 여부가 공개될 경우 논증적 에세이와 창작 스토리 모두에서 평균적인 품질 평가가 낮아지는 현상이 확인되었다(Li et al., 2024). 이러한 연구 결과는 콘텐츠 창작자와 소비자 모두가 수용하고 준수할 수 있는 표시 의무의 수준과 범위를 정교하게 설계하는 것이, 제도의 실효성을 확보하기 위한 선결조건임을 시사한다.

따라서 본 연구는「인공지능기본법」 제31조의 투명성 확보 의무를 둘러싼 헌법적 가치 충돌 상황에서의 규범적 분석과 함께 국민 인식에 대한 실증적 조사를 병행함으로써 정책적 정당성과 실효성을 제고할 수 있는 방안을 모색하고자 한다.

2. 주요 분석 내용

가. 법제 분석

「인공지능기본법」 제31조 제2항 및 제3항은 AI 생성 콘텐츠에 대한 투명성 규제의 핵심 조항으로 유럽연합의 「AI 법」과 국제적 기준을 반영하여 설계되었다(김정화 외, 2023; 김형지·정선아, 2025). 이는 생성형 AI와 딥페이크라는 구체적인 기술 양태 및 활용에 대한 직접적인 규제 조항으로서, 위반 시 행정 제재가 수반된다(「인공지능기본법」 제40조 제1항 및 제43조 제1항 제3호).

1) 「인공지능기본법」 제31조 제2항 및 제3항의 법적 구조

1-1) 제31조 제2항: 생성형 AI 결과물에 대한 포괄적 표시의무

「인공지능기본법」 제31조 제2항은 “생성형 인공지능 또는 이를 이용한 제품 또는 서비스를 제공하는 인공지능사업자는 그 결과물이 생성형 인공지능에 의하여 생성되었다는 사실을 표시하여야 한다”라고 규정하고 있다. 이 조항은 생성형 AI가 산출한 글, 이미지, 영상 등 결과물이 인간의 개입 없이 생성되었음을 명시적으로 드러내도록 요구한다. 이에 「인공지능기본법 시행령」⁶²⁾ 제23조 제2항은 AI 사업자가 결과물에 “사람이 인식할 수 있는 방법” 또는 “기계가 판독할 수 있는 방법(...)”으로 표시하도록 규정하여, 가시적 표시뿐만 아니라 워터마크 및 메타데이터와 같은 기계 판독 방식도 허용하고 있다. ‘인공지능 투명성 확보 가이드라인’(이하, ‘가이드라인’)⁶³⁾은 이를 구체화하여 영상 및 오디오 콘텐츠의 경우 전체 재생 시간 동안 표시할 필요는 없으며, 시작 부분 등에 안내하거나 기계 판독 기술을 적용하는 등의 유연한 방식을 제시한다.

그 입법 취지는 소비하거나 공유하는 콘텐츠가 인간이 사유를 통해 도출한 결과물이 아니라 기계적 알고리즘에 의한 확률적 조합의 결과물임(박상철, 2024)을 인지할 수 있는 기회를 보장하는 데 있다. 이는 AI 생성물에 대한 소비자의 오인

62) [시행 2026. 1. 22.] [대통령령 제36053호, 2026. 1. 21., 제정].

63) 과학기술정보통신부, 인공지능 투명성 확보 가이드라인, 2026. 1. 22.

을 방지하는 소극적인 보호를 넘어, 정보 생태계의 건전성을 유지하기 위한 적극적인 조치로 해석된다.

동 조항의 수범자는 ‘인공지능사업자’이다. 동 법 제2조 제7호에 따르면, 인공지능사업자는 인공지능산업과 관련된 사업을 하는 인공지능개발사업자 혹은 인공지능이용사업자 중 하나에 해당하는 법인, 단체, 개인 및 국가기관 등을 의미하며, 단순히 AI 서비스를 이용하여 결과물을 자신의 콘텐츠에 활용하는 일반 이용자(예: AI로 만든 이미지를 영화에 쓴 제작자)는 포함되지 않는다. 이는 매우 포괄적인 정의에 해당하며(비판적인 견해는 박상철, 2024), <표 3-28>과 같은 층위의 행위자들을 모두 규제 대상으로 삼는다.

<표 3-28> 「인공지능기본법」 제31조 제2항 규제 대상

구분	정의	해당 사례(예시)
인공지능 개발사업자	• AI 모델 자체를 연구·개발하여 제공 하는 자	• OpenAI(GPT), Google(Gemini), Naver(HyperCLOVA X)
인공지능 이용사업자	• 개발된 AI 모델을 API 등을 통해 이 용하여 제품·서비스를 제공하는 자	• AI 챗봇 상담 도입 기업 • 생성형 AI 기반 마케팅 툴 제공 스타 트업 등

자료: 연구진 작성

1-2) 제31조 제3항: 딥페이크에 대한 강화된 고지의무

1-2-1) 규제 대상의 구체성: ‘가상의 정보’

「인공지능기본법」 제31조 제3항은 규제 대상을 더 구체적으로 한정한다. “인공지능시스템을 이용하여 실제와 구분하기 어려운 가상의 음향, 이미지 또는 영상 등의 결과물(이른바 ‘딥페이크’)을 제공하는 경우”가 적용 요건이다. 핵심은 ‘실제와 구분하기 어려운’이라는 요건이다. 이는 단순한 포토샵 보정이나 명백한 애니메이션 캐릭터 생성 등은 규제 대상에서 제외될 수 있음을 시사하며, 인간의 인지 능력을 기만할 가능성이 높은 생성물에 규제의 초점을 맞추고 있음을 보여준다.

1-2-2) 의무의 내용: 명확한 인식 가능성

동조 제3항은 해당 결과물이 AI에 의해 생성되었다는 사실을 “이용자가 명확하게 인식할 수 있는 방식”으로 고지 또는 표시해야 한다고 명시한다. 「인공지능기본법 시행령」 제23조 제3항은 이 방식을 더욱 구체화하여, ① 이용자가 시각, 청각 등을 통하거나 소프트웨어 등을 이용하여 쉽게 내용을 확인할 수 있는 방법으로 고지 또는 표시하여야 하고, ② 주된 이용자의 나이, 신체적·사회적 조건 등을 고려하여 고지 또는 표시할 것이라고 규정한다. 단순히 서비스 이용 약관에 작은 글씨로 명시하는 것으로는 부족하며, 콘텐츠 자체에 워터마크를 삽입하거나 재생 전후에 명시적인 경고 문구를 띄우는 등 이용자가 직관적으로 인지할 수 있는 수단을 강구하여야 함을 의미한다.

1-2-3) 예외 조항: 예술적·창의적 표현의 자유 보장

제3항 후단은 중요한 예외를 두고 있다. “해당 결과물이 예술적·창의적 표현물에 해당하거나 그 일부를 구성하는 경우에는 전시 또는 향유 등을 저해하지 않는 방식으로 고지 또는 표시할 수 있다”는 단서 조항이다. 이는 헌법상 보장된 예술의 자유(제22조)와 표현의 자유(제21조)를 침해하지 않기 위한 비례성의 원칙이 반영된 결과로 볼 수 있다. 예컨대 영화 속에서 작고한 배우를 AI로 복원하거나, 풍자적 예술 작품에서 딥페이크를 활용할 경우, 화면 중앙에 거대한 경고 문구를 표시하는 것은 작품의 심미적 가치를 훼손할 수 있으므로, 엔딩 크레딧에 명시하거나 작품 설명란에 기재하는 등의 완화된 방식을 허용하겠다는 취지이다. 법률에 규정된 ‘예술적·창의적’이라는 개념은 향후 하위법령 또는 판례를 통해 구체화되어야 할 것이다.

2) 헌법적 관점에서의 쟁점 분석

생성형 AI 결과물과 딥페이크에 관한 위 규정은 기술적 규제를 넘어 헌법적 가치 간 충돌과 조정을 내포하고 있다.

2-1) 인공지능사업자의 제한되는 기본권

2-1-1) 표현의 자유에 대한 제한

「헌법」 제21조는 “모든 국민은 언론·출판의 자유를 가진다”고 규정하여 표현의 자유를 보장한다. 표현의 자유는 전통적으로 ‘사상 또는 의견의 자유로운 표명과 전파할 자유’를 의미하며, 표현의 형식과 방법, 시기 면에서도 제한받지 않을 것이 요구된다.⁶⁴⁾ 그러나 표현의 자유에는 자신이 말하고 싶지 않은 내용을 강제로 표현하도록 강요받지 않을 자유, 즉 ‘표현하지 않을 자유’도 포함된다. 「인공지능기본법」 제31조 제2항 및 제3항의 표시의무는 인공지능사업자에게 “이 콘텐츠는 AI가 생성한 것입니다”라는 특정한 메시지를 표현하도록 의무를 부과하는 규정이므로, 표현하지 않을 자유에 대한 제한으로 볼 수 있다.

2-1-2) 직업수행의 자유에 대한 제한

「헌법」 제15조는 “모든 국민은 직업선택의 자유를 가진다”고 규정하고 있으며, 이는 직업선택의 자유와 ‘직업수행의 자유(영업의 자유)’를 포함한다. 따라서 직업의 자유는 인간다운 삶을 영위하기 위하여 자신이 원하는 직업을 ‘선택’할 수 있고 스스로 선택한 직업을 자신의 뜻에 따라 ‘종사’할 수 있는 자유로 정의할 수 있다(한수웅, 2025). 「인공지능기본법」 제31조 제2항 및 제3항의 표시의무는 AI 생성 콘텐츠를 제공하는 사업자에게 해당 콘텐츠가 AI에 의해 생성되었음을 표시하도록 의무를 부과하고, 특히 딥페이크의 경우 이용자가 명확하게 인식할 수 있는 방식으로 표시하도록 규정하고 있다. 이는 사업자가 자유롭게 영업 방법을 선택할 수 있는 자유를 제약하는 규정으로 「헌법」 제15조 직업수행의 자유에 대한 제한에 해당한다(손영선, 2024).

2-2) 표시의무의 헌법적 정당성

그러나 표현의 자유도, 직업수행의 자유도 절대적인 것은 아니며, 「헌법」 제37조 제2항에 따라 “국가안전보장·질서유지 또는 공공복리를 위하여 필요한 경우”

64) 헌재 2002.4.25. 2001헌가27, 판례집14-1, 251면 이하 참조.

본질적인 내용을 침해하지 아니하는 한 법률로써 제한될 수 있다. 제한의 합헌성은 입법목적이 헌법상 정당한 공익을 추구하는지(목적의 정당성), 기본권 제한이 입법 목적 달성에 기여하는지(수단의 적합성), 덜 제한적인 대안은 없는지(침해의 최소성), 제한되는 사익과 보호되는 공익이 균형을 이루는지(법익의 균형성) 검토하여 판단한다. 여기에서는 목적의 정당성을 검토하고, 본 연구가 관심을 기울이고 있는 ‘수단의 적합성’과 ‘침해의 최소성’을 다루도록 한다.

2-2-1) 목적의 정당성

「인공지능기본법」 제31조 제2항 및 제3항의 입법목적은 정보 생태계의 투명성 확보, 국민의 알 권리 보장, 딥페이크로 인한 기본권 침해 방지 등이며, 그 정당성은 대개 쉽게 인정된다.

(알 권리) 현대 민주주의 사회에서 정보의 진실성은 올바른 여론 형성과 의사결정의 전제 조건이다. 헌법재판소는 “민주정치체제는 사상 또는 의견의 자유로운 형성 및 그 전달을 바탕으로 하는 건전한 여론 없이는 정상적인 기능이 발휘될 수 없으므로 언론·출판의 자유는 민주주의의 기초를 이루는 핵심적인 정신적 자유권”이라고 판시하였다.⁶⁵⁾ 이와 함께 재판소는 알 권리를 「헌법」 제21조의 표현의 자유에 포함되는 기본권으로 명확히 인정하면서, 이를 민주주의의 기본적인 전제로 파악하였다. AI 시대에 이러한 투명성의 요구는 더욱 강조된다. AI 기술이 텍스트, 이미지, 음성, 영상 등 다양한 형태의 정보를 대규모로 생성하기 시작함에 따라, 시민들은 자신이 접하는 정보의 출처를 명확히 알 수 있어야 한다는 필요성도 커지기 때문이다.

(정보 생태계의 투명성 확보) 알 권리의 실질적 보장은 곧 정보 생태계의 투명성 확보를 의미한다. 단순히 정보에 대한 접근 가능성을 보장하는 데 그치지 않고, 정보의 신뢰성을 평가할 수 있는 기준을 제공하는 것이 바로 실질적 보장의 핵심이다(김정화 외, 2023; 김형지·정선아, 2025). 이 기준에는 정보의 출처는 물

65) 헌재 2015.4.30. 2014헌마360, 판례집 27-1, 131; 헌재 1995.7.21. 92헌마177, 199, 판례집 7-2, 112 참조.

론 생산 방식, 전문성, 책임성 등이 속한다. 정보의 수신자는 이러한 기준을 바탕으로 정보의 신뢰성을 판단한다. 특정 언론사가 오랜 기간 정확한 보도를 했다면, 그 언론사의 기사가 신뢰를 획득하고, 특정 분야의 전문성을 가진 학자의 견해가 권위를 갖는 것과 마찬가지로이다. 정보의 질적 속성을 알 권리는 “민주정치체제는 사상 또는 의견의 자유로운 형성 및 그 전달을 바탕으로 하는 건전한 여론 없이는 정상적인 기능이 발휘될 수 없다”⁶⁶⁾고 판시한 헌법재판소의 판시 내용에서도 엿볼 수 있다. 건전한 정보적 기반이 마련되었을 때, 즉 정보의 출처와 생산 방식을 명확히 알 수 있을 때 비로소 국민은 비판적으로 정보를 평가하고, 합리적인 의사결정을 할 수 있기 때문이다.

(기본권 침해 방지) 특히 AI 기술을 활용한 딥페이크는 일반적인 AI 생성물과는 차원이 다른 위협을 야기한다. 딥페이크는 텍스트 중심이었던 기존 허위정보와 달리 시각적·청각적 요소를 결합하여 허위정보의 신뢰도를 더욱 높이는 방식으로 작동한다. 원본과 거의 구별할 수 없는 콘텐츠를 생성함으로써, 일반 시민들이 진실과 거짓을 판별하는 것을 불가능하게 만들어 정보 생태계를 근본적으로 교란한다. 이처럼 딥페이크로 인한 피해는 정보에 대한 전반적인 불신의 확산과 민주주의 체제 자체의 신뢰도 하락을 야기한다. 딥페이크의 폐해는 주관적 차원에서도 분명하게 드러난다. 특히 비동의 성적 합성물은 피해자의 인격권(초상권, 음성권, 명예권 등)에 중대한 침해가 될 수 있다(엄주희, 2025). 「헌법」 제10조 제1문의 ‘인간의 존엄과 가치’에서 도출되는 일반적 인격권은 자신의 얼굴과 음성이 자신의 의사에 반하여 조작되지 않을 권리를 포함한다. 「인공지능기본법」 제31조 제3항의 강화된 표시의무는 이러한 인격권 침해를 예방하고, 침해 발생 시 피해 구제를 용이하게 하는 최소한의 안전장치로서 규정된 것이다.

2-2-2) 수단의 적합성

AI 생성 콘텐츠나 딥페이크에 대한 표시의무 부과는 정보의 출처를 명확히 함

66) 헌재 2015.4.30. 2014헌마360, 판례집 27-1, 131 참조.

으로써 투명성을 확보하고 정보를 비판적으로 평가할 수 있도록 기반을 제공하는 데 적합한 것으로 보인다. 수단의 적합성은 입법목적의 달성에 그 수단이 효과적이고 적절한지를 판단하는 기준이며, 엄격한 경험적 검증이 아니라 합리적 관련성의 존재로 충분하므로, 법률의 위헌성 심사에 있어서 수단의 적합성은 보통 인정된다.

다만, 입법자 스스로 중대한 공익의 존재를 주장하고, 이에 대한 수단의 적합성을 일방적으로 확인한다면, 자유권에 대한 포괄적 제한이 가능하게 되므로 기본권 보장의 형해화 우려가 제기되기도 한다(한수웅, 2025). 그에 따르면, 입법형성권은 미래를 규율하고자 시도하는 행위이며, 따라서 현재 문제점과 규율 효과를 철저히 검토하여 미래를 예측하는 단계가 필요하다고 한다. 이러한 관점을 취하면, 입법권을 존중하여 합리성 심사에 그쳤던 도그마틱을 극복하고, 규율 대상의 특성과 규율의 효과에 대한 입법자의 예측 판단도 경험적으로 검증할 계기를 마련할 수 있다. 이때 떠올릴 수 있는 질문은 규율의 배경 및 효과, 「인공지능기본법」 제31조 제2항 및 제3항과 관련지어 보면, “정보 소비자들은 AI 생성물을 보고 정확히 판별할 능력을 갖추고 있는가?” 또는 “표시의무는 정보 생태계의 투명성 확보에 실제로 기여하는가?”이다. 물론 제한되는 기본권의 중요도와 제한 효과에 따라 경험적 검증의 강도도 달라져야 할 것이다(한수웅, 2025). 그러나 적어도 딥페이크가 타인의 인격을 도용하거나 현실을 왜곡하여 기만할 목적으로 사용될 경우, 이는 헌법상 보호받는 표현의 자유의 핵심 영역에서 벗어날 가능성이 크기 때문에(엄주희, 2025) 엄격한 경험적 검증까지는 필요하지 않다.

2-2-3) 침해의 최소성

침해의 최소성은 입법 목적을 달성하기 위하여 여러 수단 중 기본권을 가장 적게 제한하는 수단을 선택해야 한다는 원칙이다. 이와 관련하여 재판소는 “덜 제한적인 대안이 명백히 존재하는 경우” 해당 규제는 위헌이라고 판시하였다. 핵심은 대안이 ‘명백히’ 존재해야 하며, 그 대안이 입법 목적을 동일하게 효과적으로 달성할 수 있어야 한다는 점이다. 일단 「인공지능기본법」 제31조 제2항 및 제3항에 따른

표시의무는 AI 생성 콘텐츠나 딥페이크의 제작·배포 자체를 금지하는 것이 아니라, 단지 그것이 AI에 의해 생성되었음을 표시하도록 요구하므로 상대적으로 덜 침해적인 수단으로 평가할 수 있다. 이는 「공직선거법」이 딥페이크를 전면 금지하는 것과 비교하면 더 명확해진다.

다만, 표시 방법의 구체적 설계에 따라 인공지능사업자의 표현의 자유 및 영업의 자유 제한 정도가 현격히 달라질 수 있으며, 이는 침해의 최소성 판단에 결정적 영향을 미친다. 「인공지능기본법」 제31조 제2항 및 제3항은 구체적인 표시 방법을 규정하지 않고, 이를 시행령에 위임하고 있다(제4항). 그에 따라 「인공지능기본법 시행령」 제23조 제2항 및 제3항은 각각 AI 생성 콘텐츠 및 딥페이크의 경우 AI 생성물 표시 방법을 구체화하고 있다. 실제 수범자가 참고할 만한 자료는 ‘AI 투명성 확보 가이드라인’이다. 가이드라인에 따르면 AI 생성 이미지와 동영상의 경우 로고나 텍스트 등 가시적 또는 비가시적 워터마크 사용, 음성에 대해서는 가청적 방법 등 이용자가 명확히 알 수 있는 방법의 사용을 제시하고 있다. 영상과 음성의 경우, 콘텐츠 전체에 표시할 것 없이 이용자가 인식할 수 있는 수준이면 충분하다고 서술하고 있다. 다만, 가이드라인은 딥페이크 등에 악용될 수 있는 결과물에 대해서는 사람이 인식할 수 있는 방법만을 적용하도록 안내하고 있다.

표시 방법에 관하여 법률, 시행령, 가이드라인은 구체적인 방안을 제시하지 않고 있는 것으로 보이며, 이러한 상태에서는 침해의 최소성을 정확히 판단하기 어렵다. 선택한 수단이 입법 목적을 동일하게 달성하면서도 기본권을 덜 침해한다는 판단은 일차적으로 입법자의 판단에 맡겨져 있지만, 그러한 판단이 경험적으로 검증되었을 때 설득력을 더 획득할 수 있다(김형지·정선아, 2025). 본 연구는 이 점에 주목하고 다음과 같은 질문을 던지고자 한다: “어떠한 AI 생성물 표시 라벨이 ‘정보 생태계의 투명성 확보’라는 목표를 달성하면서도 덜 침해적인 수단인가?” 그리고 “어느 정도로 눈에 띄게 표시하여야 명확히 인지할 수 있는 방식에 해당하는가?” 물론 입법 목적의 달성과 기본권이 침해되는 정도는 ‘소비자’의 입장에서만 평가할 내용은 아니다. 전체 법질서를 아우르는 합리적 관련성을 기초로 여러 이해관계자의 관점을 포괄하는 객관적인 평가여야 할 것이다. 다만, 실제로 AI 생성

물 및 딥페이크를 마주하게 될, 그리고 정보 생태계의 대다수를 차지하는 소비자의 관점이 중요한 참고 자료가 될 수 있다는 점은 부인하기 어렵다.

정리하면, 「인공지능기본법」 제31조 제2항 및 제3항과 관련한 구체적인 표시 방법 모색은 “어떻게 표시하여야 하는가?”라는 단순한 기술적 문제가 아니라, 기본권 침해의 정도를 결정하는 핵심적인 헌법 문제이며, 경험적 방법론을 채용함으로써 침해의 최소성 원칙이 요구하는 ‘동등한 정도의 목적 달성’ 및 ‘덜 제한적인 대안’을 효과적으로 모색할 수 있다.

3) 소결

「인공지능기본법」 제31조 제2항 및 제3항의 표시의무는 인공지능사업자의 표현의 자유와 영업의 자유를 제한한다. 그러나 이는 헌법 제37조 제2항에 따라 정당한 공익 목적인 정보 생태계의 투명성, 알 권리, 기본권 보호를 위한 조치로서 과잉금지원칙 심사에서 정당화될 가능성이 높다. 다만, 표시의무가 실제로 이러한 입법목적을 달성하는 데 적합한지, 그리고 그 이행 방법을 어떻게 설계할 것인지에 따라 헌법적 평가는 결정적으로 달라질 수 있다. 특히 「인공지능기본법」이 추상적으로만 규정하고 있는 현시점에서는 구체적인 방안 모색이 필수적이다. 여기에는 효과적인 라벨링에 관한, 그리고 어느 정도를 명확한 표시 방법으로 여기고 있는지에 관한 일반시민의 인식이 유용한 참고 자료가 될 수 있을 것이다.

나. AI 생성물 표시 제도에 대한 일반시민의 인식 조사

1) 조사 목적

본 절에서는 AI 생성물에 대한 투명성 확보 필요성 및 적정 수준에 대한 일반 국민의 인식을 파악하는 것을 목적으로 수행한 설문조사를 소개한다. 최근 AI 기술을 활용한 이미지, 영상, 음원 등 다양한 콘텐츠가 급증함에 따라, AI 생성물에 대한 투명성은 창작자 및 소비자 보호, 콘텐츠의 신뢰성 및 책임성 확보, 그리고 딥페이크나 가짜 뉴스 등 사회적 혼란 방지를 위해 매우 중요한 이슈로 부상하고 있다. 특히 콘텐츠의 사용 맥락과 잠재적 영향력에 따라 투명성 수준을 차등 적용해야 한다는 논의가 활발한 가운데, 실제 사용자 측면에서의 인식 및 요구에 대한

체계적인 연구는 여전히 미진한 상황이다. 이에 본 조사는 AI 생성물에 대한 국민의 인식과 요구를 심층적으로 파악하여, 향후 관련 정책 및 가이드라인 수립에 필요한 객관적인 기초 자료를 제공하는 것을 목적으로 한다.

2) 표본 설계

본 조사는 AI 생성물의 투명성 및 적정 정보 제공 수준에 대한 국민 인식을 측정하기 위하여 온라인 설문조사 기반의 실험 연구 방법으로 설계되었다. 조사 표본은 주민등록인구통계를 기반으로 하여 전국 2개 권역, 2개 성별, 5개 연령대별로 나누어 할당 추출을 진행함으로써 대표성을 확보하고자 하였다. 조사는 2025년 11월 7일부터 11월 10일까지 총 4일간 진행되었으며, 이 기간 동안 총 752명

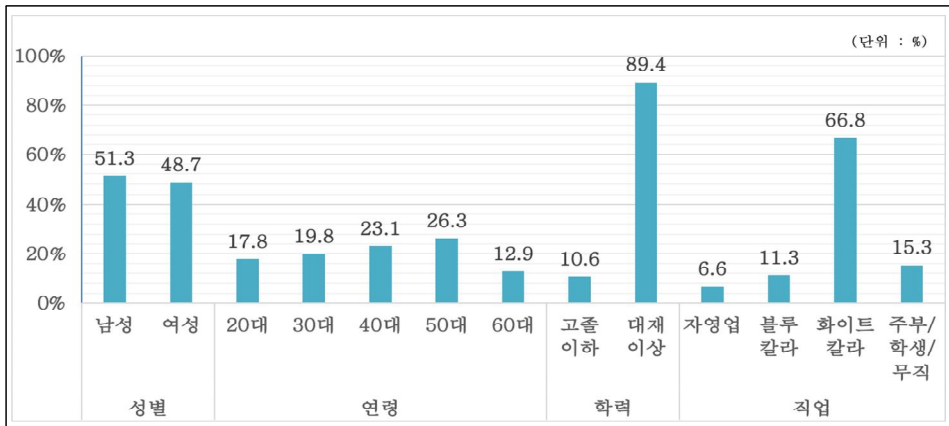
〈표 3-29〉 조사 설계

구 분	세 부 내 용
조사 대상	전국 만 20~64세 사이의 남녀
조사 지역	전국
조사 방법	온라인 설문 기반 실험 연구
표본추출방법	성/연령별 할당 추출
표본 수	752명
조사 기간	11월 7일~10일(4일간)
참고사항	1) 표본 추출 • 가장 최근 주민등록인구통계(행안부) 기반 할당 - 전국 2개 권역×2개 성×5개 연령대별 할당 추출 2) 응답자 답례 지급 • 설문 응답을 완료한 경우 5,000원의 기본 참여보수를 지급하며, 응답 내용에 따라 추가로 3,000원의 답례를 받을 수 있는 기회를 제공 - 응답 내용에 따른 추가보상 가능성을 부여하는 보상 구조를 통해 적극적으로 성실한 응답을 유도 - 응답 답례는 응답자의 선호에 따라, 현금 또는 모바일 상품권으로 사후 지급

자료: 연구진 작성

의 응답이 온라인으로 수집되었다. 응답자들에게는 적극적인 실험 참여를 독려하기 위해 응답 완료 후 선호에 따른 현금 또는 모바일 상품권 형태의 소정의 답례를 지급하였다.

[그림 3-33] 응답자 특성



자료: 연구진 작성

3) 실험 설계

본 조사는 AI 활용 수준 정보 제공 여부가 사용자의 콘텐츠 가치 평가 변화 및 투명성 인식에 미치는 영향을 비교 분석하는 것을 핵심 목표로 설계되었다. 이를 위해 실험에 활용할 이미지는 AI 활용 수준에 따라 세 가지 유형으로 구분하여 제작하였다. 첫 번째 유형은 AI가 전혀 활용되지 않은 이미지로, 연합뉴스 콘텐츠 플랫폼인 헬로아카이브⁶⁷⁾에서 실제로 연합뉴스 온라인 기사에 활용된 보도사진 원본 8장을 구매하여 실험자극물로 활용하였다. 두 번째 유형은 AI가 부분적으로 활용된 이미지로, 연합뉴스 보도사진 원본을 활용하되, 사진의 의미를 크게 훼손하지 않는 범위 내에서 일부 시각적 요소만을 AI로 변형했다. 구체적으로는, 생성형 AI에게 보도사진 원본을 제시하고 사진 배경의 시간대만 변경하도록 요청하여

67) 헬로아카이브, <https://www.helloarchive.co.kr/>.

이미지를 변형 제작하였다. 마지막 세 번째 유형은 AI가 전면적으로 활용된 이미지로, 어떠한 원본 사진도 제공하지 않고 오직 텍스트 프롬프트만을 이용하여 AI가 이미지 내 모든 시각적 요소를 완전히 생성하도록 하였다. 이에 따라 각 8개의 연합뉴스 보도사진 원본에 대해 AI 활용 수준에 따른 세 가지 유형의 콘텐츠(총 $8 \times 3 = 24$ 장)가 실험 자극물로 제작되었다.⁶⁸⁾ 각 응답자에게는 8개의 서로 다른 원본 이미지 각각에 대한 세 가지 AI 활용 유형 중 하나의 유형이 무작위로 선택되어 제시되었다. 즉 응답자들은 총 8개의 이미지를 무작위 순서로 보게 되며, 동일한 원본을 기반으로 한 여러 유형의 이미지를 중복해서 보지 않도록 설계되었다.

본 조사는 총 세 단계의 실험 절차로 구성되었다. 첫 번째 단계에서는 응답자들에게 제시된 이미지의 AI 활용 수준에 대한 정보가 전혀 주어지지 않은 상태에서 조사가 진행되었다. 응답자들은 제시된 이미지에 대하여 사실성(“이 이미지는 실제 일어난 일을 있는 그대로 보여준다”), 중요성(“중대한 사건/현상과 관련된 내용을 담고 있다”), 심미성(“이 이미지는 구도, 색감 등 시각적으로 아름답다”), 인상도(“이 이미지는 보는 사람에게 강한 인상이나 감정을 불러일으킨다”), 전반적 만족도(“귀하께서는 이 이미지에 대해 전반적으로 얼마나 만족하십니까?”)를 평가하였다. 이와 더불어, 해당 이미지가 AI 기술을 얼마나 활용하여 제작되었는지에 대한 응답자의 주관적인 예측을 측정하였으며, 해당 이미지를 온라인에서 보게 될 경우 다른 사람들과 공유할 것인지에 관한 의도도 함께 측정하였다. 이 과정은 응답자들의 AI 생성물에 대한 초기 가치 인식 수준을 파악함과 동시에, 일반 국민의 AI 생성물 판별 능력을 간접적으로 측정하는 목적으로 활용되었다.

두 번째 단계에서는 응답자들을 무작위로 대조군과 실험군에 배치하였다. 실험군에게는 제시된 이미지에 대한 실제 AI 활용 정보를 제공하였으나, 대조군에게는 해당 정보를 제공하지 않았다. 이는 대조군과 실험군의 비교를 통해 AI 활용 정보 제공 여부가 콘텐츠의 가치 평가 및 인식에 미치는 영향과 공유 의도 변화를 구체적으로 비교 분석하고, 정보의 투명성이 사용자의 판단과 행동 의도에 미치는

68) 본 실험 연구에서 활용된 실험자극물은 이현경 외(2025)의 연구에서 활용된 것과 동일하나 연구 목적 및 분석 초점이 상이한 별도의 연구로 수행되었음.

구체적인 효과를 측정하기 위함이다.

마지막 세 번째 단계에서는 두 집단 모두에게 AI 활용 수준 정보를 최종적으로 제공한 후 AI 생성물의 투명성 확보 방안에 대한 인식을 심층적으로 조사하였다. 구체적으로, AI 활용 정보 제공 방식으로 표시 문구의 적절성과 AI 활용 여부 표시 방식의 적정 수준에 대한 인식을 측정함으로써, 향후 정책 및 가이드라인 수립에 실질적으로 활용될 수 있는 국민적 요구사항을 파악하고자 하였다.

〈표 3-30〉 실험 설계

	조사 항목	
1단계	• (응답자 전체) AI 활용 수준 정보 제공 X - 콘텐츠 가치 평가(사실성, 중요성, 심미성, 인상도, 만족도) - AI 활용 수준 판별 능력 - 온라인 공유 의도	
2단계	• 대조군: AI 활용 수준 정보 제공 X - 콘텐츠 가치 평가(사실성, 중요성, 심미성, 인상도, 만족도) - AI 활용 수준 판별 능력 - 온라인 공유 의도	• 실험군: AI 활용 수준 정보 제공 O - 콘텐츠 가치 평가(사실성, 중요성, 심미성, 인상도, 만족도) - AI 활용 수준 판별 능력 - 온라인 공유 의도
3단계	• (응답자 전체) AI 활용 수준 정보 제공 O - AI 활용 수준 표시 문구의 적절성 - AI 활용 수준 표시 방식의 적정 수준	

자료: 연구진 작성

3) 조사 결과

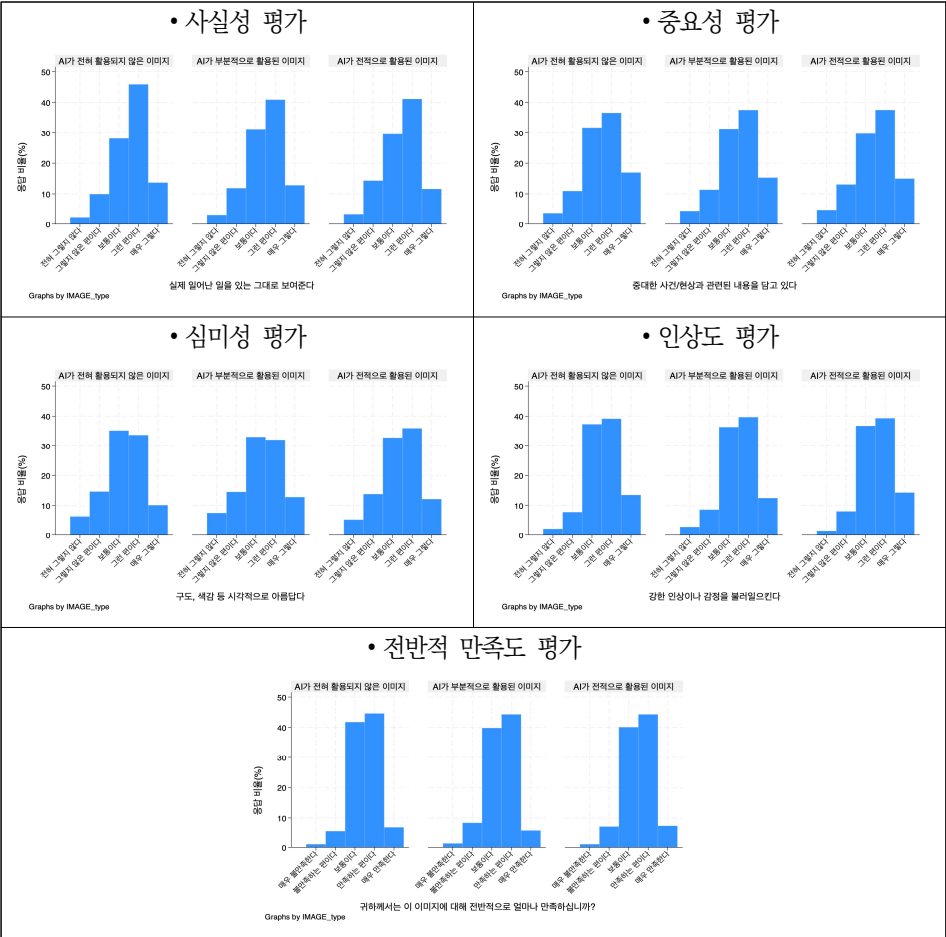
3-1) AI 활용 정보 공개 전: 1단계 응답 분석 결과

본 연구의 1단계 조사 응답을 분석한 결과 AI 활용 수준 표시가 없을 때 정보 생태계에서 발생할 수 있는 잠재적인 문제를 드러내는 중요한 통찰을 제공한다.

먼저, 콘텐츠의 실제 AI 활용 수준을 알지 못하는 상태에서 가치를 어떻게 평가하는지 살펴본다. 응답자들이 제시된 콘텐츠에 대해 평가한 가치 평가 항목별 응답 분포는 〈표 3-34〉에 제시되어 있다. 본 연구에서는 콘텐츠의 AI 활용 수준별

평가의 통계적 차이를 정량적으로 파악하기 위해 가장 부정적인 평가를 1, 가장 긍정적인 평가를 5로 수치화하여 선형회귀분석을 진행하였다. 분석 모델 구축 시, AI가 전혀 활용되지 않은 콘텐츠를 기준 집단(reference group)으로 설정하였다. 이를 기준으로, AI가 부분적으로 활용된 콘텐츠와 AI가 전면적으로 활용된 콘텐츠가 각 가치 평가 항목(사실성, 중요성, 심미성, 인상도, 전반적 만족도)에 미치는 영향을 회귀계수를 통해 추정하였다.

[그림 3-34] 1차 응답 이미지 평가 도수분포표



자료: 연구진 작성

〈표 3-31〉 1차 응답 이미지 평가 선형회귀분석 결과

	(1) 사실성	(2) 중요성	(3) 심미성	(4) 인상도	(5) 만족도
부분적 AI 활용	-0.102*** (0.0288)	-0.0414 (0.0303)	0.0143 (0.0289)	-0.0361 (0.0271)	-0.0633*** (0.0233)
전면적 AI 활용	-0.153*** (0.0301)	-0.0703** (0.0283)	0.0910*** (0.0292)	0.0284 (0.0260)	-0.00889 (0.0227)
상수	3.523*** (0.0392)	3.924*** (0.0381)	2.859*** (0.0414)	3.739*** (0.0354)	3.382*** (0.0321)
전체 관측수	6016	6016	6016	6016	6016
응답자 수	752	752	752	752	752

주: 괄호 안의 수치는 개인 수준으로 군집화한 표준오차이며, 실험자극물 고정효과(fixed effects)를 통제함

* $p<0.1$, $p<0.05$, $p<0.01$.

자료: 연구진 작성

분석 결과 AI 활용 수준 정보가 부재한 상황에서 응답자들이 콘텐츠를 평가하는 방식에서 몇 가지 중요한 경향이 확인되었다. 사실성 평가의 경우, AI가 전혀 활용되지 않은 콘텐츠에 비해 AI가 부분적으로 또는 전면적으로 활용된 콘텐츠의 평가가 통계적으로 유의미하게 낮게 나타났다. 이는 응답자들이 콘텐츠의 실제 AI 활용 수준을 정확히 알지 못하더라도, AI 활용이 의심되는 콘텐츠를 ‘진실된’ 정보로 받아들이는 데 주저함을 보였음을 시사한다. 한편 중요성 평가에서는, 특히 AI가 전면적으로 활용된 콘텐츠에 대한 평가가 다른 두 유형의 콘텐츠에 비해 유의미하게 낮게 나타났다. 이는 전달되는 내용이 동일할지라도, 해당 이미지가 AI 생성물로 의심되는 순간 그 콘텐츠의 내용적 중요성 자체를 낮게 평가하는 경향이 있음을 보여준다. 심미성에 대한 평가에서는 오히려 AI가 전면적으로 활용된 콘텐츠가 다른 유형보다 높게 평가되는 경향이 나타났다. 이는 AI가 전적으로 생성한 이미지가 기술적인 완성도나 시각적 구성 측면에서 인간의 작품보다 더 매력적으로 느껴질 수 있음을 의미한다. 다만, 인상도에 대한 평가에서는 콘텐츠

유형별로 통계적 차이가 나타나지 않았다. 이는 AI 생성물이 인간의 작품과 마찬가지로 감정적 혹은 시각적 인상을 불러일으키는 데는 일정 수준의 효과를 가질 수 있으며, 이는 AI 활용 수준과 무관하게 콘텐츠 자체의 특성에 더 의존할 수 있다는 해석을 가능하게 한다.

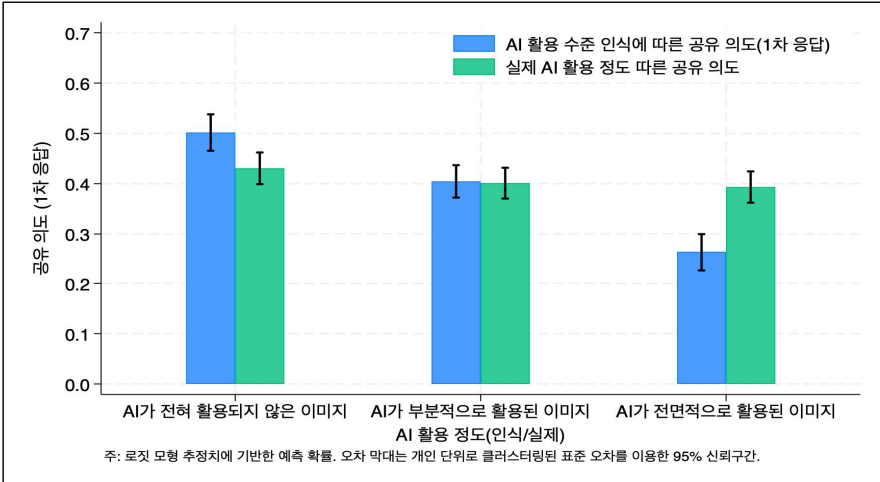
모든 가치 평가를 종합하는 “귀하께서는 이 이미지에 대해 전반적으로 얼마나 만족하십니까?”를 묻는 전반적 만족도에 대해서는 AI가 부분적으로 활용된 이미지에 대한 평가가 유의미하게 가장 낮게 나타나는 결과가 확인되었다. 이는 AI가 일부만 활용된 경우, 콘텐츠의 사실성이 손상되었다고 인식되면서도 심미성은 크게 향상되지 않아 전반적인 만족도가 가장 저해되는 결과로 이어졌다고 해석할 수 있다.

뿐만 아니라, 일반 소비자의 AI 생성물에 대한 인식과 실제 행태 사이에서 중요한 괴리가 발견되었다(그림 3-35) 참고). 응답자들은 해당 이미지가 ‘AI가 전혀 활용되지 않은 콘텐츠’라고 주관적으로 판단할수록 해당 이미지를 온라인상에서 다른 사람들과 더 많이 공유하고자 하는 경향을 보였다. 이는 소비자들 스스로가 AI 기술이 활용되지 않은, 이른바 ‘진짜(authentic)’ 콘텐츠만을 온라인에서 공유하고자 하는 높은 의도를 가지고 있음을 시사한다.

그러나 흥미롭게도 이러한 소비자들의 공유 의도와는 달리, AI 활용 콘텐츠이든 비활용 콘텐츠이든 실질적으로는 동일한 수준으로 공유하는 양상이 관찰되었다. 실제 실험에서 응답자들의 AI 생성물 판별 능력은 매우 부족하기 때문이다. AI 활용 정보를 제공받기 전의 초기 인식 조사 결과, 응답자들이 해당 이미지의 실제 AI 활용 수준을 정확하게 판별할 수 있었던 비율은 약 40%에 불과하였다(그림 3-36) 참고).

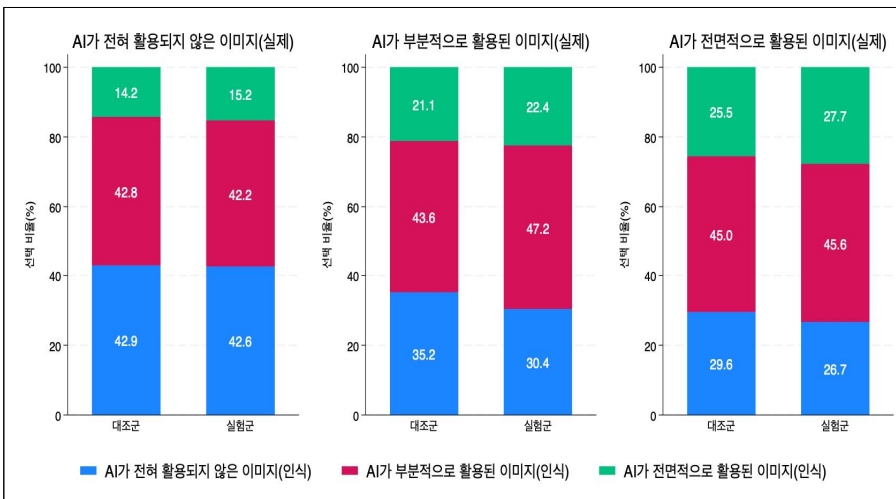
결론적으로, 본 결과는 일반 소비자들이 스스로 AI가 활용되지 않은 이미지만을 공유하려는 높은 의도를 가졌음에도 실제 행동 차원에서는 부정확한 판별 능력으로 인해 의도하지 않게 AI 생성물을 과도하게 공유하고 있을 가능성을 강력하게 보여준다. 이는 AI 생성물이 무분별하게 공유·확산될 수 있는 가능성과 함께, 향후 온라인상의 콘텐츠 신뢰도 저하 문제에 대한 정책적 고려의 필요성을 제기한다.

[그림 3-35] 실제와 인식에 따른 온라인 공유 의도



자료: 연구진 작성

[그림 3-36] AI 활용 수준 판별 정확도



자료: 연구진 작성

3-2) AI 활용 정보 공개 효과 분석: 1단계와 2단계 응답 변화 비교
 AI 활용 정보 공개의 인과적 효과를 보다 명확하게 추론하기 위해, 1단계 응답

과 2단계 응답의 차이를 비교하는 이중차분(Difference-in-Differences) 분석을 대조군과 실험군 간에 수행하였다. 이를 통해 AI 활용 정보를 투명하게 제공하는 것이 사용자의 콘텐츠 인식에 어떠한 영향을 미치는지 심층적으로 확인하였다.

분석 결과, 콘텐츠 유형별로 AI 활용 정보 공개의 효과가 확연하게 다르게 나타났다. 첫째, AI가 전혀 활용되지 않은 이미지의 경우, AI 활용 정보가 공개된 실험군에서 사실성, 중요성, 심미성, 인상도, 전반적 만족도의 모든 가치 평가 항목에 대한 1단계-2단계 간 변화분이 AI 활용 정보를 제공받지 않은 대조군의 변화분 대비 통계적으로 유의하게 증가하는 것으로 나타났다(〈표 3-32〉 참고). 이는 콘텐츠가 AI 활용이 없는 인간의 작품임이 명확히 확인되었을 때, 응답자들이 해당 콘텐츠의 가치와 전반적 만족도를 전반적으로 높게 재평가하는 경향이 존재함을 보여준다.

〈표 3-32〉 AI가 전혀 활용되지 않은 이미지 이중차분 분석 결과

	(1) 사실성	(2) 중요성	(3) 심미성	(4) 인상도	(5) 만족도
2차 응답	-0.0372** (0.0163)	-0.0110 (0.0162)	0.00876 (0.0186)	-0.0383** (0.0164)	-0.0186 (0.0120)
실험군	0.0305 (0.0539)	0.0268 (0.0565)	-0.0249 (0.0554)	0.0180 (0.0529)	0.0232 (0.0452)
2차 응답× 실험군	0.410*** (0.0379)	0.142*** (0.0291)	0.0951*** (0.0285)	0.128*** (0.0277)	0.152*** (0.0219)
전체 관측수	4022	4022	4022	4022	4022
응답자 수	736	736	736	736	736

주: 괄호 안의 수치는 개인 수준으로 군집화한 표준오차이며, 실험 자극물 고정효과(fixed effects)를 통제함

* $p<0.1$, $p<0.05$, $p<0.01$.

자료: 연구진 작성

반면, AI 활용이 확인된 콘텐츠에 대해서는 전반적인 가치 하락 현상이 나타났다. AI가 부분적으로 활용된 이미지의 경우, AI 활용 정보가 공개된 실험군에서 중요성 평가를 제외한 나머지 모든 항목의 가치 평가에 대한 1단계-2단계 간 변화분이 AI 활용 정보를 제공받지 않은 대조군의 변화분 대비 유의미하게 작게 나타났다(〈표 3-33〉 참고). 더 나아가, AI가 전면적으로 활용된 이미지의 경우에는 사실성, 중요성, 심미성, 인상도, 전반적 만족도 모든 항목에서 가치 평가의 변화분이 대조군 대비 유의미하게 작았다. 특히 AI가 전적으로 활용된 이미지에 대한 가치 평가 감소 폭은 AI가 부분적으로 활용된 이미지의 감소폭보다 더욱 크게 나타났다(〈표 3-34〉 참고). 이는 AI 활용 정보가 투명하게 공개되었을 때, 응답자들이 AI 생성물에 대해 심리적 반감 또는 낮은 신뢰도를 표출하며 콘텐츠의 가치를 전반적으로 낮추는 경향이 있음을 인과적으로 확인한 결과이다.

〈표 3-33〉 AI가 부분적으로 활용된 이미지 이중차분 분석 결과

	(1) 사실성	(2) 중요성	(3) 심미성	(4) 인상도	(5) 만족도
2차 응답	-0.00817 (0.0171)	-0.0445** (0.0176)	0.000907 (0.0188)	-0.00998 (0.0172)	0.0127 (0.0136)
실험군	0.0653 (0.0585)	-0.0474 (0.0560)	0.113* (0.0575)	0.0764 (0.0579)	0.107** (0.0479)
2차 응답×실험군	-0.175*** (0.0385)	-0.0218 (0.0288)	-0.0506* (0.0297)	-0.0585** (0.0297)	-0.0646*** (0.0239)
전체 관측수	4014	4014	4014	4014	4014
응답자 수	740	740	740	740	740

주: 괄호 안의 수치는 개인 수준으로 군집화한 표준오차이며, 실험 자극물 고정효과(fixed effects)를 통제함

* $p<0.1$, $p<0.05$, $p<0.01$.

자료: 연구진 작성

〈표 3-34〉 AI가 전면적으로 활용된 이미지 이중차분 분석 결과

	(1) 사실성	(2) 중요성	(3) 심미성	(4) 인상도	(5) 만족도
2차 응답	0.0250 (0.0164)	0.0130 (0.0192)	0.0190 (0.0161)	-0.00999 (0.0175)	-0.00599 (0.0129)
실험군	0.0784 (0.0602)	0.124** (0.0546)	0.0308 (0.0562)	0.0953* (0.0527)	0.0162 (0.0471)
2차응답× 실험군	-0.397*** (0.0439)	-0.153*** (0.0340)	-0.124*** (0.0298)	-0.108*** (0.0297)	-0.129*** (0.0274)
전체 관측수	3996	3996	3996	3996	3996
응답자 수	735	735	735	735	735

주: 괄호 안의 수치는 개인 수준으로 군집화한 표준오차이며, 실험 자극물 고정효과(fixed effects)를 통제함

* p<0.1, p<0.05, p<0.01.

자료: 연구진 작성

AI 활용 정보 공개가 콘텐츠 공유 의도에 미치는 영향을 살펴본 결과, 콘텐츠 유형에 따라 상반된 효과가 확인되었다. AI가 전혀 활용되지 않은 콘텐츠에 대한 공유 의도는 정보 공개 후 유의하게 증가한 반면, AI가 전적으로 활용된 콘텐츠에 대한 공유 의도는 유의하게 감소하였다. 한편, AI가 부분적으로 활용된 콘텐츠의 공유 의도는 유의미한 차이가 없는 것으로 나타났다(〈표 3-35〉 참고).

이러한 결과는 AI 활용 정보를 투명하게 공개하는 것이 응답자의 공유 행동 의사 결정에 중요한 역할을 함을 보여준다. AI 생성물이 가짜 뉴스나 딥페이크 등 사회적 위험을 야기하는 경우가 무분별한 온라인 공유 확산에서 비롯됨을 감안할 때, AI 활용 정보를 투명하게 공개하는 것은 위험성이 높은 콘텐츠의 확산을 억제하는 데 기여할 수 있음을 시사한다.

〈표 3-35〉 콘텐츠 유형별 공유 의도 이중차분 분석 결과

	(1) AI가 전혀 활용되지 않은 이미지	(2) AI가 부분적으로 활용된 이미지	(3) AI가 전면적으로 활용된 이미지
2차 응답	-0.0186** (0.00914)	-0.00544 (0.00715)	-0.00899 (0.00818)
실험군	-0.00696 (0.0326)	0.0213 (0.0315)	0.0516 (0.0324)
2차응답×실험군	0.0887*** (0.0152)	-0.00561 (0.0132)	-0.0662*** (0.0156)
전체 관측수	4022	4014	3996
응답자 수	736	740	735

주: 괄호 안의 수치는 개인 수준으로 군집화한 표준오차이며, 실험 자극물 고정효과(fixed effects)를 통제함

* $p<0.1$, $p<0.05$, $p<0.01$.

자료: 연구진 작성

3-3) 바람직한 AI 활용 표시 문구 및 방식에 대한 인식: 3단계 조사 결과 3단계 조사는 응답자들이 콘텐츠의 실제 AI 활용 정보를 모두 인지한 상태에서 향후 규범 및 정책 수립에 필요한 바람직한 투명성 수준과 방식에 대한 규범적 인식을 파악하는 데 중점을 두었다.

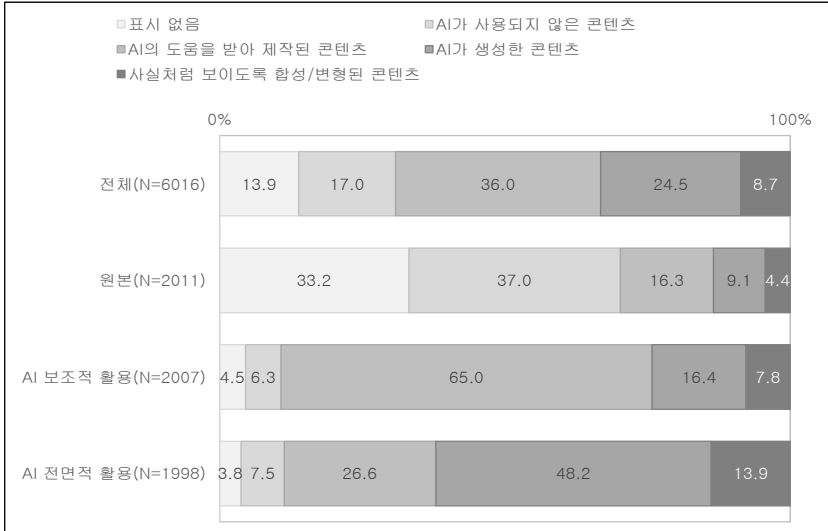
먼저 제시된 콘텐츠에 어떤 ‘AI 활용 여부 표시 라벨’을 부착하는 것이 사회적으로 가장 바람직한지 묻는 질문에 대하여, 응답자들은 콘텐츠 유형별로 상이한 선호도를 보였다. AI가 전혀 활용되지 않은 이미지에 대해서는, 전체의 약 33%가 아무런 표시도 하지 않는 것이 바람직하다고 응답하였으며, 이보다 4%p 높은 비율인 37%에 해당하는 응답자가 “AI가 사용되지 않은 콘텐츠”라는 명시적인 표시 라벨을 붙이는 것이 바람직하다고 응답하였다. 현실에서 ‘AI를 사용하지 않았다’는 표시를 굳이 하지 않는 것이 일반적인 관행임을 감안할 때, 오히려 명시적인 표시를 선택한 비율이 높게 나타났다는 점은 주목해 볼 만한 결과이다. 이는 응답

자들이 원본의 진위 여부를 명확히 구분할 수 있도록 투명하게 AI 활용 정보가 공개되는 것이 바람직하다고 판단하고 있음을 시사한다.

AI가 부분적으로 활용된 이미지에 대해서는 약 65%의 응답자가 “AI의 도움을 받아 제작된 콘텐츠”라는 표시 라벨을 바람직하다고 판단하였다. 반면 “AI가 생성한 콘텐츠”, “사실처럼 보이도록 합성·변형한 콘텐츠”와 같이 보다 강한 표현을 선택한 비율은 각각 16.4%, 7.8%에 그쳤다. 한편 AI가 전면적으로 활용된 이미지의 경우에는 약 절반에 해당하는 48.2%가 “AI가 생성한 콘텐츠”라는 표시 라벨이 바람직하다고 응답하였고, 26.6%는 “AI의 도움을 받아 제작된 콘텐츠”, 13.9%는 “사실처럼 보이도록 합성·변형한 콘텐츠”라는 표시 라벨을 선택했다. 이러한 결과는 AI 활용 수준 정도에 따라 적절하다고 인식되는 표시 문구가 달라지며, 국민들이 활용 정도를 구분해 표현할 수 있는 상이한 표시 라벨의 필요성을 인식하고 있음을 시사한다. 다시 말해, AI가 일부만 개입한 콘텐츠와 전적으로 생성·합성된 콘텐츠를 동일한 문구로 표시하는 것보다는, 활용 범위와 개입 정도에 따라 표시 라벨을 보다 세분화하고 표준화할 필요가 있음을 보여준다.

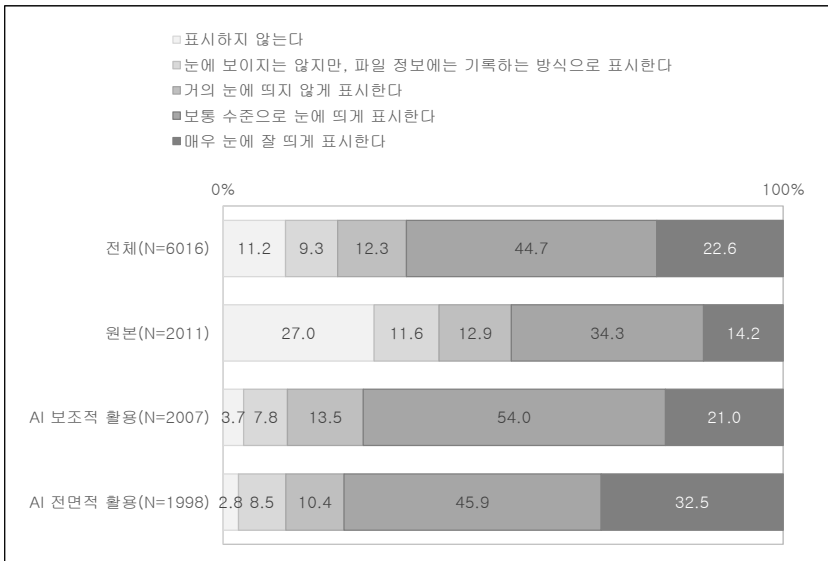
AI 활용 여부 표시 방식에 대해서는 콘텐츠 유형과 관계없이 “눈에 띄는 방식”이 바람직하다고 생각하는 경향이 지배적이었다. [그림 3-37]에 따르면, 전체 응답자의 44.7%가 “보통 수준으로 눈에 띄게 표시해야 한다”고 응답했으며, 22.5%는 “매우 눈에 잘 띄게 표시되어야 한다”고 응답하였다. 이를 종합하면, 응답자의 67.2%가 AI 활용 정보를 어느 정도 이상 명확히 인지할 수 있도록 표시되어야 한다고 요구하고 있음을 알 수 있다. 특히 AI가 전면적으로 활용된 이미지의 경우에는 “매우 눈에 잘 띄게 표시되어야 한다”는 응답이 32.5%로, 다른 유형의 콘텐츠 대비 가장 높게 나타났다. 이는 콘텐츠의 AI 활용 수준이 높아질수록, 소비자들이 해당 정보의 시각적 강조 수준 또한 높아야 한다고 요구하고 있음을 보여준다.

[그림 3-37] 사회적으로 바람직한 AI 활용 여부 표시 레벨에 대한 인식



자료: 연구진 작성

[그림 3-38] 사회적으로 바람직한 AI 활용 여부 표시 방식에 대한 인식



자료: 연구진 작성

4) 소결

본 실험 조사는 AI 활용 정보의 투명성이 일반 사용자의 콘텐츠 인식 및 공유 의도에 미치는 인과적 효과를 실증적으로 규명하였다는 점에서 중요한 의의를 갖는다. 주요 분석 결과, AI 활용 정보 공개는 AI 활용 수준에 따른 콘텐츠 유형별로 상이한 효과를 가져오는 것으로 확인되었다.

AI가 전혀 활용되지 않은 콘텐츠임이 명확해질 경우, 사실성, 중요성, 심미성 등 모든 가치 항목에 대한 평가와 공유 의도가 유의미하게 상승했다. 반면, AI가 부분적으로라도 활용되었음이 밝혀질 경우, 콘텐츠의 가치 평가와 전반적인 만족도는 유의미하게 하락하였다. 특히 AI가 전적으로 활용된 콘텐츠에 대한 가치 하락폭이 가장 크게 나타났으며, 공유 의도 역시 유의미하게 감소하였다. 이는 AI 활용 정보의 투명한 공개가 가짜 뉴스나 딥페이크와 같은 위험성이 높은 콘텐츠의 무분별한 온라인 확산을 억제하는 데 기여할 수 있다는 긍정적인 정책적 시사점을 제공한다.

또한 바람직한 표시 방식에 대한 규범적 인식 조사에서는, 응답자들이 인간의 창작물과 AI의 생성물을 명확히 구분할 수 있는 표시 라벨을 부착하는 것이 사회적으로 바람직하다고 인식하고 있음을 확인하였다. 더불어 AI 활용 수준이 높을수록 시각적으로 더욱 눈에 띄게 표시되어야 한다고 요구하는 ‘정보의 차등적 강조’ 필요성을 확인하였다. 이러한 결과는 향후 AI 생성물에 대한 표시 표준 및 가이드라인을 수립하는 데 중요한 기초자료로 활용될 수 있다.

다만, 본 실험에서는 온라인 뉴스의 보도사진을 실험자극물로 활용했음을 유의해야 한다. 즉 다양한 영역별 특성을 고려하지 않은 결과이므로 예술과 같이 창의성이 중요시되는 다른 영역에서는 이와 다른 결과가 나타날 가능성이 있으며, 따라서 이 결과를 모든 AI 생성물 영역에 일반화하기는 어렵다는 해석상의 한계가 존재한다. 향후 영역별 특성(예: 뉴스, 예술, 마케팅 콘텐츠 등)을 고려한 추가 연구가 필요할 것이다.

3. 식별된 사회적 의제안

가.(정보-1) AI 기반 허위 조작 정보 범람에 따른 정보 신뢰 체계 붕괴에 어떻게 대응할 것인가?

1) 제안의 배경

AI 기술, 특히 생성형 AI의 발전으로 누구나 텍스트, 이미지, 영상 콘텐츠를 손쉽게 생성할 수 있게 되었다. 이 기술의 발전 속도는 AI가 생성한 콘텐츠를 실제와 구분하기 어렵게 만들고 있으며, 이는 진실과 허위의 경계를 무너뜨려 사회 전반의 정보 신뢰 체계 붕괴를 초래할 수 있다. 특히 선거, 재난, 보건 등 중요한 사회적 의사결정 과정에서 조작된 정보가 유포될 경우, 그 피해는 개인을 넘어 사회 전체의 혼란과 기능 마비로 확산될 수 있다.

2) 현황 및 쟁점

딥페이크 영상과 AI 음성합성을 활용한 금융사기·보이스피싱, 개인의 얼굴·음성 등을 이용한 디지털 성범죄, 명예 훼손 등의 악의적 활용 사례가 다수 보고되고 있다. AI 기반 허위조작정보가 대규모 여론조작, 정치적 선전, 선거 개입 등으로 이어지며 민주적 의사결정의 공정성을 훼손할 우려가 현실화되고 있다. 핵심 쟁점은 악의적 생성자에 대한 책임 부과와 기술적 진위 검증 시스템 구축의 시급성이다.

3) 문제의 심각성

악의적 행위자는 손쉽게 대량의 허위조작정보를 생성·유포할 수 있으나, 이를 검증하고 대응하는 데는 막대한 시간과 자원이 필요하다. 한 번 유포된 허위조작정보는 삭제나 정정이 어렵고, 피해자의 회복이 사실상 불가능한 경우가 많다. 특히 디지털 리터러시가 낮은 계층(노인, 아동 등)이 허위조작정보에 더욱 쉽게 노출될 가능성이 크며, 이는 정치적 양극화와 사회 분열을 심화시킬 수 있다.

4) 기존 논의/대책의 한계

기존의 미디어 규제나 명예훼손 관련 법규는 인간이 수동으로 생성한 정보를 전제로 하고 있어, AI가 대규모로 빠르게 생성·유포하는 초국가적 허위정보에 대한 실효적 대응에 한계가 있다. 기술적으로도 워터마킹, 메타데이터 등 진위 판별 기술이 개발되고 있으나, AI 기술의 발전 속도를 따라가지 못하며 이를 우회하는 기술 역시 동시에 발전하고 있어 근본적인 해결책이 되지 못하고 있다. 플랫폼의 자율 규제에만 의존하기에는 그 책임 소재와 강제력이 미흡한 상황이다.

나. (정보-2) AI 기반 허위 과장 광고 및 소비자 기만 행위로부터 소비자의 합리적 선택권을 어떻게 보장할 것인가?

1) 제안의 배경

AI로 생성된 가짜 리뷰, 조작된 제품 이미지, 과장된 효과 영상 등이 온라인 쇼핑몰과 소셜미디어에 범람하고 있다. AI가 콘텐츠를 쉽게 생성하는 능력을 악용하여, 특히 건강기능식품, 화장품, 의료기기 등 소비자 건강 및 재산에 직결되는 분야에서 허위·과장 광고가 급증하고 있다. 이는 소비자의 합리적인 구매 결정을 심각하게 방해하고 시장 내 공정경쟁 질서를 훼손하고 있다.

2) 현황 및 쟁점

AI 가상 인물이 실제 사용 경험을 한 것처럼 연출하거나, 전문가·의료인 이미지를 무단 활용하여 상품의 신뢰도를 조작하는 사례가 증가했다. 더욱 심각한 것은 AI가 소비자 데이터 분석을 통해 건강·외모 콤플렉스 등 개인의 심리적 취약점을 겨냥한 맞춤형 감정 조작 광고를 확산시키고 있다는 점이다. 진위 검증이 어려운 광고 콘텐츠가 시장 내 공정경쟁 질서를 훼손하고, 정직한 사업자의 피해를 초래하고 있다.

3) 문제의 심각성

소비자의 합리적 의사결정 과정을 왜곡하여 불필요한 소비나 금전적 손실을 유발한다. 나아가 공정한 시장 경쟁을 저해하고, 정직하게 운영하는 사업자까지 피

해를 보게 만든다. 장기적으로는 광고의 진실성에 대한 소비자 신뢰 기반이 붕괴되어 디지털 상거래 전반의 지속가능성을 위협하고, 궁극적으로는 경제 시스템의 효율성을 떨어뜨린다.

4) 기존 논의/대책의 한계

현행 「표시광고법」이나 「소비자보호법」은 주로 명백한 사실 왜곡을 대상으로 하며, AI가 정교하게 생성한 ‘실제 경험처럼 보이는’ 가상 콘텐츠나 개인의 취약점을 겨냥한 ‘감정 조작형’ 광고에 대한 규율은 미흡하다. 또한 소셜미디어 등 다양한 플랫폼에 실시간으로 대량 유포되는 광고를 행정력만으로 일일이 모니터링하고 제재하는 데 물리적 한계가 존재한다.

다. (정보-3) AI 알고리즘의 불투명성이 야기하는 정보 편식과 사회적 양극화 현상을 어떻게 해소하고 공론장의 공정성을 확보할 것인가?

1) 제안의 배경

포털사이트의 뉴스 배열, 소셜미디어의 콘텐츠 추천, 검색엔진의 결과 순위 등은 모두 AI 알고리즘에 의해 결정되고 있다. 그 결과, 이용자는 알고리즘이 선별한 정보만 접하는 정보 편식 현상이 심화되고 있다. 이러한 알고리즘의 작동 방식은 블랙박스로 남아 있어, 편향되거나 조작되더라도 이를 확인하거나 시정할 수 있는 제도적 장치가 부재하며, 이는 사회적 대화의 기반을 위협한다.

2) 현황 및 쟁점

AI 알고리즘이 클릭률·체류시간 등 자극적 지표를 중심으로 콘텐츠 노출을 결정하면서, 선정적·감정적 콘텐츠가 과잉 노출되고 저널리즘의 공익성이 훼손된다. 자신의 기존 신념과 일치하는 정보만 반복 소비하는 필터 버블(Filter Bubble)과 유사한 성향의 이용자들이 폐쇄적 공간에서 상호 확증하는 ‘집단 극단화’ 현상이 심화되고 있다. 특히 청소년과 아동은 알고리즘이 큐레이션한 정보에 노출되어, 사회를 인식하는 방식 자체가 알고리즘에 의해 형성될 위험에 놓여 있다. 플랫폼이 사회적 의제 설정권을 사실상 독점했으나, 이에 상응하는 책임과 규제는 미비하다.

3) 문제의 심각성

정보 접근과 여론 형성이 소수 플랫폼의 불투명한 알고리즘에 의해 통제되는 구조는, 민주주의의 근간인 ‘자유로운 정보 유통’과 ‘공론장 형성’을 심각하게 위협한다. 이용자의 인식·의제 설정·가치 판단이 AI가 설계한 틀 안에서만 이루어지는 상황은 인간의 인지적 자율성과 비판적 사고 능력을 약화시킨다. 장기적으로는 사회 전체가 알고리즘이 조정하는 인지 구조 속에서 작동하게 되어, 정보생태계의 다양성과 인간 주체성이 근본적으로 침식될 우려가 있다.

4) 기존 논의/대책의 한계

알고리즘 투명성 확보를 위한 논의는 존재하지만, 영업 비밀 보호와의 상충 문제로 인해 플랫폼 기업들이 알고리즘의 핵심 작동 원리를 공개하는 것에 소극적이다. 기존의 「방송·신문법」은 전통적 매체를 중심으로 설계되어 있어, AI 알고리즘을 통해 콘텐츠를 유통하는 거대 플랫폼에 대한 규율 체계는 미흡하다. 또한 알고리즘을 투명하게 공개하더라도 일반 이용자가 그 복잡성을 이해하고 편향성을 파악하기 어렵다는 근본적 한계가 있다.

라. (정보-4) AI의 저작권 침해 및 무단 학습으로부터 창작 활동의 가치를 어떻게 보호하고 지속가능한 문화 생태계를 구축할 것인가?

1) 제안의 배경

생성형 AI가 기존 예술작품·디자인·사진 등을 무단으로 학습하여 유사하거나 동일한 창작물을 재생산함에 따라, 원저작자의 경제적 손실 및 명예 훼손이 발생하고 있다. 현행 「저작권법」 체계는 인간의 창작 행위를 전제로 설계되어 있어, AI 기반 창작물의 권리 보호와 학습 데이터 이용에 대한 책임 부과에 대응하기 어려운 구조적 한계에 직면했다.

2) 현황 및 쟁점

AI가 만든 이미지나 음악이 원작자의 작품과 구분되지 않거나 시장에서 더 높은 주목을 받는 상황에서, 예술가들은 직업적 정체성과 생계의 위협을 동시에 경

험하고 있다. AI 생성물은 상업적 트렌드를 모방하며 반복적으로 생산되기 때문에, 실험적이거나 비주류 예술 양식은 배제되어 문화의 다양성이 축소되고 있다. 핵심 쟁점은 AI 학습 데이터 이용의 공정 이용 원칙 적용 여부와 AI가 생성한 결과물에 대한 저작권 인정 기준이다.

3) 문제의 심각성

AI 시대에 기존 저작권 체계가 사실상 무력화되면서, “창작에 대한 보호와 보상”이라는 저작권의 본래 목적이 훼손되고 있다. 창작이 정당한 경제적 보상을 받지 못하는 활동으로 전락하면 예술가와 창작자의 생존 자체가 위협받으며, 이는 단순한 직업 위기 이상의 문화 생산 시스템의 붕괴로 이어질 수 있다. 창작 생태계가 붕괴되면 새로운 문화 콘텐츠의 원천이 고갈되고, AI는 과거 데이터를 반복 재생산하는 ‘자기복제적 기술’로 전락하여 문화의 정체와 퇴보를 초래할 수 있다.

4) 기존 논의/대책의 한계

대부분의 국가에서 AI 학습을 위한 저작물 이용에 대한 명확한 법적 기준이 부재하다. 저작권자의 허락 없이 학습하는 것을 금지할 경우 AI 산업의 발전이 저해될 수 있다는 우려와, 무단 학습을 허용할 경우 창작 생태계가 붕괴된다는 우려가 첨예하게 대립하고 있다. 또한, AI 생성물의 ‘인간의 기여도’를 객관적으로 측정하기 어렵기 때문에 저작권 침해 여부나 권리 귀속을 판단하는 것이 기술적·법률적으로 매우 복잡하다.

마. (정보-5) AI 활용 역량의 불균형이 야기하는 새로운 정보 격차와 사회적 불평등을 어떻게 해소할 것인가?

1) 제안의 배경

AI 기술이 빠르게 확산되면서 정보의 생산·유통·활용 방식이 근본적으로 변화하고 있으나, AI를 활용할 수 있는 역량의 차이는 개인 간, 세대 간, 지역 간 새로운 디지털 격차로 확대되고 있다. 특히 AI 생성물의 진위와 출처를 식별하고 비판적으로 사용할 수 있는 AI 리터러시의 중요성이 강조되고 있다. 이 활용 능력의

불균형은 단순한 기술 접근성의 차원을 넘어, 정보 이해력·비판적 사고·사회참여 기회의 불평등으로 이어지고 있다.

2) 현황 및 쟁점

AI 서비스 이용 경험은 주로 고학력·고소득층, 젊은 세대에 집중되어 있으며, 고령층·저소득층·농어촌 지역에서는 접근과 활용이 제한적이다. AI 도구를 효과적으로 활용하지 못하는 이용자는 정보 검색·의사결정·직업활동에서 불리한 위치에 놓이며, 일상적 기회 격차가 확대되고 있다. 핵심 쟁점은 AI 활용 능력이 낮을 수록 AI가 생성한 허위정보(딥페이크 등)나 불투명한 AI 생성물을 사실로 맹신하거나 오인할 위험이 높다는 점이다. 이는 AI 생성물의 투명성이 보장되지 않는 상황에서 더욱 심각해진다.

3) 문제의 심각성

AI 활용 격차는 단순한 ‘기술 숙련도’의 차이가 아니라, 사회적 기회·경제적 보상·정치적 영향력의 차이로 확산되는 복합적 불평등 구조로 발전하고 있다. 특히, AI 생성물의 불투명성과 리터러시 격차가 결합될 경우, 정보 접근과 해석 능력의 격차는 사회적 의제 형성 과정에서 특정 집단의 목소리를 배제시켜 디지털 민주주의의 기반을 약화시킨다. 이러한 불균형이 지속될 경우, AI는 사회통합을 촉진하는 도구가 아니라 오히려 계층 간 단절을 고착화하는 요인으로 작용하여 사회적 양극화를 심화할 위험이 존재한다.

4) 기존 논의/대책의 한계

기존의 정보화 교육이나 디지털 격차 해소 정책은 주로 인터넷 접근성과 기본적인 컴퓨터 사용 능력을 높이는 데 초점을 맞추었다. 그러나 AI 시대에는 생성형 AI에 의해 제작된 정보의 진위를 판별하는 능력, AI의 편향성 이해, 비판적 사고와 같은 고차원적인 AI 리터러시 교육이 필수적으로 요구되기에, 기존의 교육 프로그램으로는 이러한 심화된 역량 수요를 충족시키기 어렵다. 더욱이, AI 생성물의 출처, 제작 과정, 학습 데이터 등에 대한 투명성 정보가 충분히 제공되지 않는

구조적 문제가 존재한다. 이러한 정보적 제약 속에서는 아무리 리터러시 교육을 강화하더라도, 이용자가 AI 생성물의 진위를 정확히 판단하고 비판적으로 접근하는 데 본질적인 한계가 따른다.

제 5 절 AI와 지역: 혁신 확산의 지리적 병목

1. 논의의 배경

인공지능(AI) 기술은 최근 비약적으로 발전하며 국가 경제와 산업 전반의 구조를 재편하고 있다. 대규모 언어모델, 생성형 AI, 자율주행, 스마트 제조, 정밀 의료 등 다양한 분야에서 AI는 생산성 향상과 신산업 창출의 핵심 동력으로 자리 잡을 것으로 예상된다. 이에 따라 미국은 최근 AI 행동계획(America's AI Action Plan)을 발표하며 국가 차원의 전방위적 AI 육성 전략을 추진하고 있고, 중국은 차세대 인공지능 발전 계획을 바탕으로 2030년까지 세계 AI 선도국으로 도약하기 위한 노력을 지속하고 있다. 우리나라도 인공지능 대전환을 통해 AI Big 3 진입을 목표로 하고 있다.

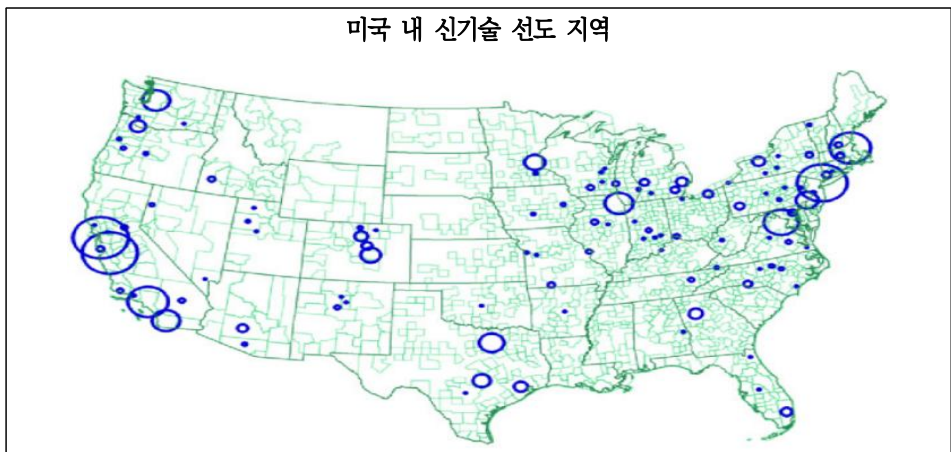
그러나 AI의 발전은 기술혁신과 산업 고도화를 촉진하는 긍정적 효과와 함께, 특정 지역에 집중될 경우 경제력·고용·혁신 기회의 격차를 심화시킬 수 있다는 우려를 낳는다. OECD(2024)는 AI 발전이 지역 간 격차를 심화시킬 수 있다고 경고하며, AI 발전 정책과 함께 지역 균형 발전 정책의 병행을 강조했다. Muro and Methkupally(2025)는 미국 내 지역별 AI 준비도 지수(AI 기술을 개발·활용할 수 있는 지역의 역량)를 구축하여 분석한 결과, AI 준비도가 일부 대도시에 집중되어 있어 향후 지역 불평등이 심화될 가능성을 지적하였다. 이러한 국제 사례는 AI 산업 발전 과정에서 대도시 편중과 지역 불균형이 현실적 위협이라는 점을 보여주며, 이는 한국에도 동일하게 적용될 수 있는 경고다.

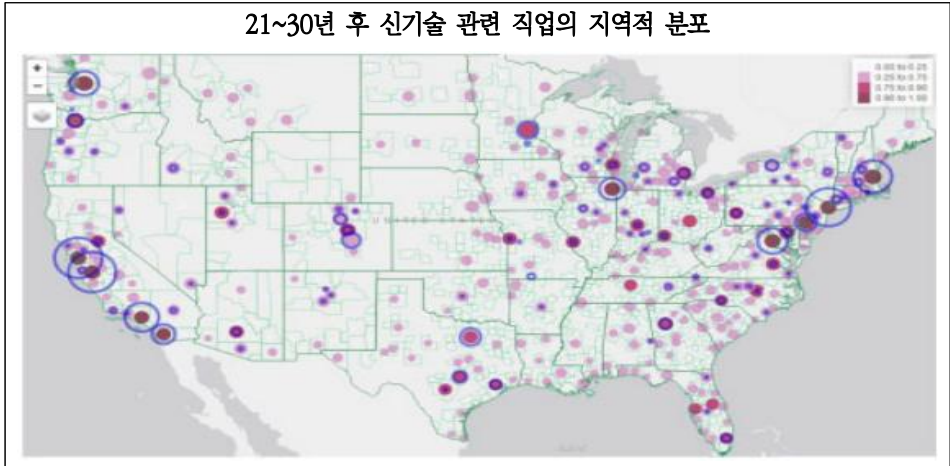
AI 기술의 지역적 편중 가능성은 경제학에서 오래전부터 논의되어 온 집적효과(agglomeration economies)에서 기인한다. 집적효과는 한 지역에 모인 경제주체들이 규모의 경제가 작동하는 생산요소를 공동 활용하는 공유 효과(sharing), 인력과 기업 간 연결의 질을 높이는 매칭 효과(matching), 그리고 대면 교류를 통해 아이디어가 확산되는 지식 스프illo버(knowledge spillovers)로 구성된다(Marshall, 1890). 이러한 세 가지 메커니즘은 경제주체들이 특정 지역에 집적될

수록 생산성을 향상시키기 때문에 지리적 집중을 심화시키는 요인으로 작용한다. Moretti(2021)는 이러한 집적효과가 생산성을 높이는 효율성의 원천이지만, 지역 간 소득·고용 격차를 심화시키는 형평성 문제를 야기함으로써 효율성과 형평성 간의 근본적 상충관계를 형성한다고 지적한다. 특히 AI 산업은 대규모 연산 인프라에 대한 높은 수요, 전문 인력·기업 간의 매칭, 그리고 지식 스�필오버가 두드러지는 지식 집약적 산업으로서, 집적효과와 이에 따른 지역 불평등이 발생할 가능성이 크다.

문제는 이러한 효율성과 형평성의 상충이 단기적인 현상에 그치지 않고 장기적으로 고착될 수 있다는 점이다. 이는 신기술이 지역적으로 확산되는 방식과 그 속도에 밀접하게 관련된다. 첫째, 선행연구에 따르면 신기술의 지역적 확산은 매우 국소적이다. 예를 들어, Jaffe et al.(1993)은 인용 특허와 피인용 특허의 위치를 비교해, 특허 인용이 동일한 주(state)와 도시권(Standard Metropolitan Statistical Area) 내에서 집중적으로 발생한다는 점을 밝혔다. 또한 Rosenthal and Strange(2008)는 집적효과가 임금에 미치는 영향을 분석하며, 지식 스펠오버의 효과가 발생지로부터 몇 마일만 떨어져도 급격히 약화됨을 보였다.

[그림 3-39] 미국 내 신기술 관련 직업의 지역적 확산





자료: Kalyani et al.(2025), p. 1324, 1326

둘째, 신기술의 지역적 확산은 매우 더딘 속도로 이루어지며, 이는 초기 우위 지역의 기술적 우세가 장기간 지속되는 경향으로 이어진다. Kalyani et al.(2025)은 1976~2014년 사이에 개발된 신기술 관련 직업이 지역적으로 어떻게 확산되었는지를 분석하였다. [그림 3-39]의 첫 번째 그림은 신기술 선도 지역(등장 초기 특히 활동이 가장 활발했던 지역)을, 두 번째 그림은 21~30년이 지난 시점의 신기술 관련 직업의 지역적 분포를 나타낸다. 해당 연구는 신기술 관련 직업이 선도 지역에 장기간 편중되며, 전국적으로 확산되는 데 약 50년이 걸릴 것으로 추정하였다. 이러한 선행연구에 비추어볼 때, AI 산업 또한 강한 집적효과로 인해 지역적 확산이 매우 국소적이고 느릴 가능성이 높으며, 이에 따른 지역 불평등이 장기간 지속될 수 있다.

〈표 3-36〉 국내 주요 AI 관련 통계조사의 조사 내용 및 지역 단위

통계명	작성기관	조사 단위	AI 관련 조사 내용	지역단위
인공지능산업 실태조사	소프트웨어 정책연구소	기업	AI 공급기업의 매출액, 인력 현황 등	전국 (지역 정보 없음)
기업활동조사	통계청	기업	AI 개발 여부·방법 및 활용 분야	시도
기업정보화 통계조사	한국지능정보사회 진흥원	기업	AI 활용 여부·목적 등	시도
중소기업 정보화수준조사	중소기업 기술정보진흥원	기업	중소기업의 AI 활용 수준	전국 (지역 정보 없음)
디지털산업 실태조사	정보통신정책연구원	사업체	디지털산업 내 AI 개발·활용 여부	전국 (지역 정보 없음)
한국미디어 패널조사	정보통신정책연구원	가구	생성형 AI 이용 현황	시도
지능정보사회 이용자패널조사	정보통신정책연구원	가구	생성형 AI 이용 현황 및 인공지능에 대한 인식	시도

자료: 각 조사 작성기관의 공식 통계 및 발간물을 참고하여 연구진이 정리

이러한 국제적 논의와 중요성에도 불구하고, 국내에서는 AI 산업의 지역적 분포·확산 및 지역 불평등에 대한 연구가 여전히 부족하다. 주요 이유 중 하나는 지역 수준에서 AI 관련 데이터가 제한적이기 때문이다. 예를 들어, AI 공급기업(AI 소프트웨어·서비스·하드웨어 생산)을 대상으로 한 인공지능산업실태조사에는 지역 정보가 포함되지 않는다. AI의 개발 및 활용에 대한 정보를 포함한 기업활동조사는 지역 정보가 시도 수준에 그친다. 그러나 주요 선행연구(Jaffe et al., 1993; Kalyani et al., 2025; Rosenthal and Strange, 2008)에 따르면, 신기술의 지역적 분포·확산 양상은 매우 국소적으로 나타난다. 따라서 AI 산업의 지역적 분포·확산을 연구하려면 시도보다 더 세분화된 지역 단위(예: 시군구)의 분석이 필요하다. 그럼에도 국내의 디지털 산업·지역 관련 연구는 이러한 데이터의 한계로 인해 주로 시도 수준에서 수행되었다(이준영·김영민, 2024; 장재영·박소연, 2022). 이에 더해, 디지털 격차를 다루는 한국미디어패널조사와 지능정보사회 이

용자 패널조사 역시 시도 단위로만 정보를 제공한다. 이로 인해 지역 간 디지털 격차 연구 또한 시군구 수준에서는 이루어지지 못한 실정이다(고흥석, 2023; 박준혁·심홍진, 2023).

본 연구는 이러한 문제의식에서 출발하여 다음과 같은 과업을 수행한다. 첫째, 시군구 단위에서 AI 산업의 지역 분포를 파악한다. 둘째, AI 산업의 지역 간 확산을 제한하는 공간적 병목을 실증적으로 분석한다. 셋째, 분석 결과를 바탕으로 AI 시대 지역 간 포용과 상생을 위한 사회적 의제를 도출한다. 본 연구는 AI 산업의 성장과 지역 균형 간의 조화를 모색하기 위한 정책 논의의 기초자료를 제공할 것이다.

2. 주요 분석 내용

가. AI 산업의 지역 분포

본 연구는 2021~2023년 전국사업체조사 자료를 이용해 시군구 단위에서 AI 산업의 지역 분포를 분석한다. 분석 기간(2021~2023년)은 생성형 AI 확산과 함께 AI 산업이 빠르게 성장한 시기로, 최근 AI 산업의 지리적 분포와 변화 양상을 분석하기 위해 선정하였다. 본 연구에서는 AI 산업을 AI 기술을 공급하는 기업(AI 소프트웨어·서비스·하드웨어 생산)뿐 아니라 AI 기술을 활용하는 기업까지 모두 포함하는 개념으로 정의한다. AI 기술을 공급하는 산업은 인공지능산업실태조사의 분류 기준을, AI 활용도가 높은 산업은 Calvino et al.(2024)의 AI 집약도(AI intensity) 기준을 따른다. 각 분류 기준에 해당하는 산업 목록을 <표 3-37>에 제시한다.

AI 산업은 대규모 연산 인프라에 대한 높은 수요, 전문 인력·기업 간 매칭, 지식 스펠오버가 두드러지는 지식 집약적 산업으로서 집적효과가 작동할 가능성이 높다. 이러한 집적효과는 국소적인 공간에서 발생한다(Jaffe et al., 1993; Kalyani et al., 2025; Rosenthal and Strange, 2008). 이에 본 연구는 AI 산업의 지역 분포를 분석하기 위해 광역 단위가 아닌 시군구 수준의 세분화된 공간 단위를 활

용한다.

AI 산업의 지역적 분포를 분석하기 위해 Kalyani et al.(2025)의 방법을 따라 입지계수(location quotient)를 계산한다. 연도 t 시점에 지역 r 에서 산업 i 의 입지계수 LQ_{rit} 는 다음과 같이 정의한다.

$$LQ_{rit} = \frac{N_{rit} / \sum_r N_{rit}}{\sum_i N_{rit} / \sum_i \sum_r N_{rit}}$$

여기서 N_{rit} 는 연도 t 에 지역 r 에서 산업 i 에 종사하는 노동자 수를 나타낸다. 위 식에서 분자는 산업 i 고용에서 지역 r 이 차지하는 비중이며, 분모는 전체 산업 고용에서 해당 지역이 차지하는 비중이다. LQ_{rit} 값이 1보다 크면 연도 t 시점에 지역 r 이 산업 전체 대비 산업 i 에서 더 집중되어 있음을, 1보다 작으면 산업 전체 대비 집중되어 있지 않음을 의미한다.

AI 산업의 입지계수를 산출하여 AI 산업의 지리적 집중을 분석한다. [그림 3-40]은 분석 기간의 시작점인 2021년 기준 AI 산업 입지계수를 시각화한 것으로, 첫 번째 지도는 전국을, 두 번째 지도는 서울·경기 지역을 확대해 보여준다. 입지계수에 따르면 AI 산업은 일부 지역에 집중되어 있으며, 서울·경기 지역과 대전 유성구에서 높게 나타난다. 한편 서울·경기 지역을 확대해 보면, 해당 지역 내에서도 입지계수가 지역별로 상이하다. 서울·경기 지역의 중심부(서울 강남구·서초구, 경기 성남시 등)에서 입지계수가 가장 높고, 남부 지역이 대체로 높은 반면 북부 지역은 전반적으로 낮은 수준을 보인다. 서울 강남구·서초구, 경기 성남시 등 일부 지역에서 나타나는 국소적 집중은 집적효과의 국소성과 부합한다.

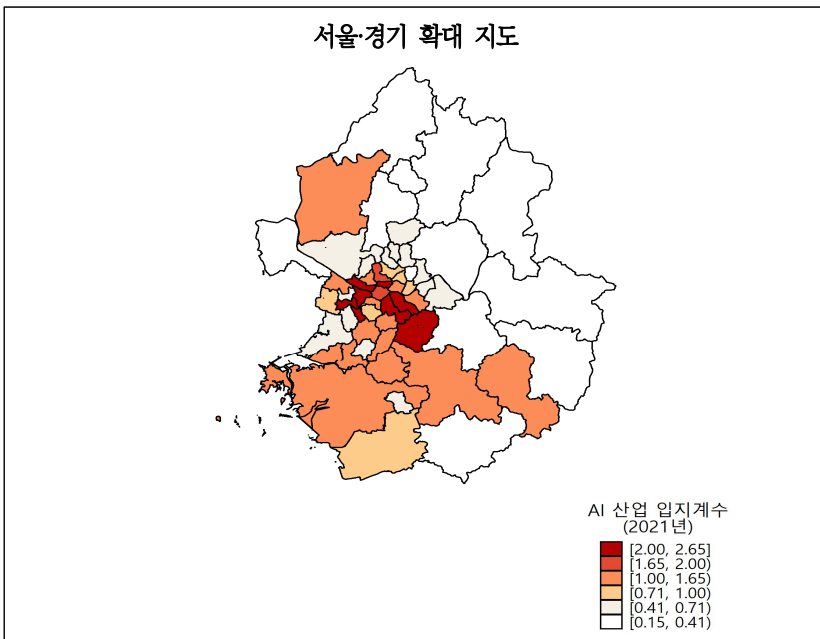
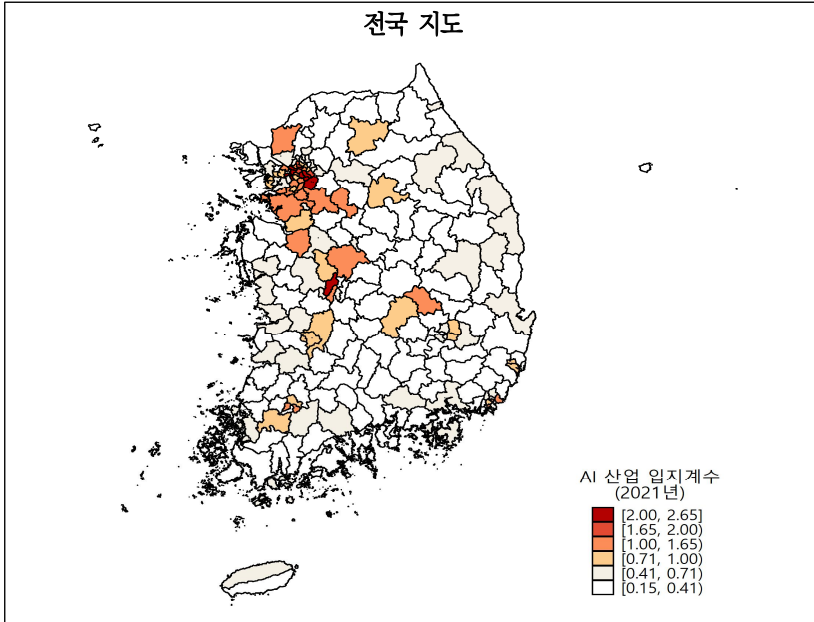
〈표 3-37〉 AI 산업 분류 기준

분류기준	한국표준산업분류(10차)	산업특성
인공지능산업 실태조사	• (58221) 시스템 소프트웨어 개발 및 공급업 • (58211) 유선 온라인 게임 소프트웨어 개발 및 공급업 • (58212) 모바일 게임 소프트웨어 개발 및 공급업 • (58219) 기타 게임 소프트웨어 개발 및 공급업 • (58222) 응용소프트웨어 개발 및 공급업 • (62010) 컴퓨터 프로그래밍 서비스업 • (62021) 컴퓨터 시스템 통합 자문 및 구축 서비스업 • (62022) 컴퓨터시설 관리업 • (62090) 기타 정보기술 및 컴퓨터 운영 관련 서비스업 • (63111) 자료처리업 • (63112) 호스팅 및 관련 서비스업 • (63120) 포털 및 기타 인터넷 정보매개 서비스업 • (63910) 뉴스 제공업 • (63991) 데이터베이스 및 온라인 정보 제공업 • (63999) 그 외 기타 정보 서비스업 • (26111) 메모리용 전자집적회로 제조업 • (26112) 비메모리용 및 기타 전자집적회로 제조업 • (26121) 발광 다이오드 제조업 • (26129) 기타 반도체소자 제조업	AI 기술 공급 산업
Calvino et al.(2024)	• (26) 전자부품, 컴퓨터, 영상, 음향 및 통신장비 제조업 • (58) 출판업 • (59) 영상·오디오 기록물 제작 및 배급업 • (60) 방송업 • (61) 우편 및 통신업 • (62) 컴퓨터 프로그래밍, 시스템 통합 및 관리업 • (63) 정보서비스업 • (64) 금융업 • (65) 보험 및 연금업 • (66) 금융 및 보험 관련 서비스업 • (70) 연구개발업 • (71) 전문 서비스업	AI 집약도 상위(high) 산업

주: Calvino et al.(2024)은 산업별 AI 집약도(AI intensity)를 AI 관련 인적자본, 혁신활동, 산업 활용 장벽을 반영한 AI 노출도, AI 활용을 기준으로 측정함. 네 가지 요소 중 최소 두 가지가 상위 25%에 해당하고, 나머지 두 가지는 하위 25%에는 속하지 않는 경우가 AI 집약도 상위(high) 산업으로 정의됨

자료: 인공지능산업실태조사 및 Calvino et al.(2024) 기반 연구진 재정리

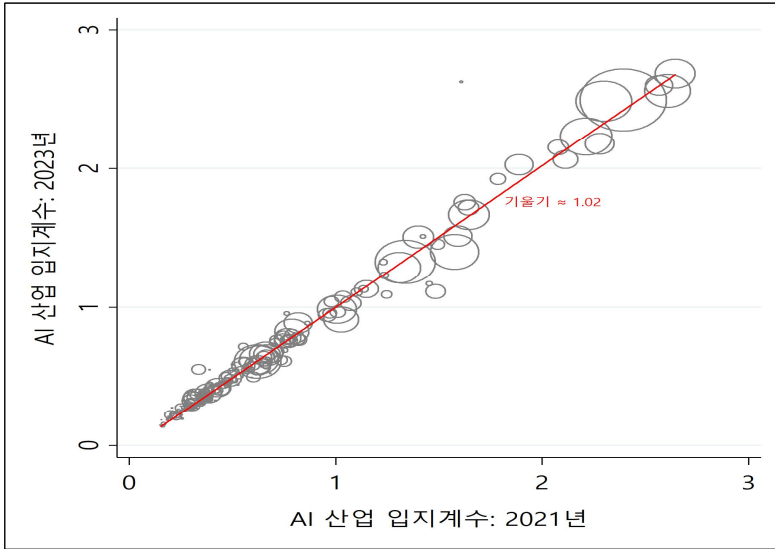
[그림 3-40] AI 산업 입지계수(2021년)



자료: 전국사업체조사를 연구진이 정리

다음으로, AI 산업의 지역 집중도 변화를 살펴보기 위해 2021~2023년 사이의 입지계수 변화를 분석한다. [그림 3-41]에서 x축은 2021년, y축은 2023년의 AI 산업 입지계수를 나타내며, 각 원의 크기는 지역별 고용 규모를 의미한다. 추정된 기울기는 1.02로, 2021년에 입지계수가 높았던 지역이 2023년에도 높은 수준을 유지했음을 보여준다. 이러한 패턴은 신기술 산업의 지리적 집중이 지속된다는 Kalyani et al.(2025)의 분석과도 일관된다.

[그림 3-41] AI 산업 입지계수 변화



자료: 전국사업체조사를 연구진이 정리

나. AI 산업 확산의 지리적 병목 분석

위에서는 AI 산업이 특정 지역에 집중되어 있으며 이러한 지역적 집중 현상이 지속되고 있음을 보였다. 이는 Krugman(1991)으로 대표되는 신경제지리학(New Economic Geography)에서 강조하는 자기강화(self-reinforcing) 메커니즘과 관련된다. 초기 단계에서 한 지역이 특정 산업에서 우위를 갖게 되면 기업과 노동자가 유입되고, 이러한 유입은 매칭·공유·지식 스펀오버와 같은 집적효과

를 통해 생산성을 높인다. 높아진 생산성은 다시 기업과 노동자의 추가 유입을 유인하여 초기의 우위를 더욱 공고히 하는 순환적 인과관계(circular causation)를 형성한다.

〈표 3-38〉 자기강화 메커니즘

구분	주요 내용	작동 방식
집적효과	집적 → 생산성 향상	매칭, 공유, 지식 스퍼illo버
경로 의존성	초기 우위가 강화되어 지속	초기 우위 → 기업·노동자 유입 → 집적효과에 따른 생산성 향상 → 기업·노동자 추가 유입의 순환

자료: 연구진 작성

위에서 논의한 자기강화 메커니즘에 기반하여, AI 산업의 지역 간 확산을 제약하는 병목을 실증적으로 분석한다. 회귀식은 다음과 같다.

$$LQ_{r, AI, 2023} = \alpha + X_{r, 2021} \cdot \beta + \epsilon_r$$

위에서 LQ 는 입지계수, r 은 시군구를 의미한다. 종속변수는 2023년 지역 r 에서 AI 산업의 입지계수를 뜻한다. $X_{r, 2021}$ 은 2021년 지역 특성으로 구성된 설명변수 벡터이다. 2021년 변수를 사용하는 이유는 초기 우위가 자기강화 과정을 통해 2023년에도 지속되는지를 확인하기 위함이다. ϵ_r 은 오차항이다. 오차항은 광역 시도 내에서는 상관되나 광역시도 간에는 독립적이라고 가정하며, 이에 따라 표준오차는 광역시도 단위로 군집화하였다. 회귀분석은 지역 고용 규모를 반영하기 위해 지역별 고용자 수를 가중치로 부여하여 수행하였다.

본 연구는 초기 우위를 나타내는 설명변수로 네 가지 요인에 초점을 맞춘다.

1) AI 관련 인적자본

AI 인재가 집적된 지역일수록 기업-노동자 간 매칭이 원활하게 이루어져 생산

성이 높아진다. AI 관련 인적자본은 컴퓨터공학·수학·통계학 학위 보유자를 기준으로 식별하며, 지역별 고용조사 자료를 활용한다.

2) 데이터센터

데이터센터는 전력·냉각·서버·부지 등 막대한 초기 고정비가 투입되는 시설로, 규모의 경제가 작동하는 AI 산업의 핵심 생산요소이다. 해당 변수는 한국데이터센터연합회 자료로 측정한다.

3) 벤처캐피털 투자액

급성장하는 AI 산업에서 벤처캐피털은 혁신 자본 인프라로 작용하여 기업 성장을 촉진한다. 본 연구에서는 AI 스타트업에 대상으로 한 벤처캐피털 투자액을 이용하여, The VC 자료를 활용한다.

4) 산업 구조

지역마다 산업 구조가 상이하며, AI 기술이 생산성을 높일 수 있는 산업이 더 많이 분포한 지역일수록 AI 산업 성장의 잠재력이 크다. 본 연구는 이러한 산업 구조를 AI 집약도 상위 산업의 고용 비중으로 측정하며, AI 집약도 상위 산업의 정의는 Calvino et al.(2024)을 따른다.

위 네 가지 변수는 모두 로그 변환하여 사용하였다. 이 외에 지역의 인구통계학적 특성(성별·연령 구성과 인구 규모)은 통제변수로 포함한다.

회귀분석 결과는 네 가지 초기 조건이 자기강화 메커니즘을 통해 AI 산업의 지역 집중을 지속시키는 병목 요인으로 작동함을 시사한다. 2021년에 AI 인적자본, 데이터센터, 벤처캐피털 투자액, AI 집약도 상위 산업 고용 비중이 높은 지역은 2023년에도 AI 산업 입지계수가 높게 나타난다. 이는 초기 우위 → 기업·노동자 유입 → 집적효과를 통한 생산성 향상 → 기업·노동자 추가 유입으로 이어지는 자기강화 메커니즘과 부합한다. 즉, AI 산업 확산의 지리적 병목은 초기 조건의 지역별 격차와 이를 강화하는 집적효과에서 비롯되는 것으로 해석된다.

〈표 3-39〉 AI 산업 확산의 지리적 병목 요인 분석

종속변수: AI 산업 입지계수(2023년)	
AI 인적자본 (로그, 2021년)	0.1059** (0.0446) [0.2111]
데이터센터 (로그, 2021년)	0.0936*** (0.0116) [0.1824]
벤처캐피털 투자액 (로그, 2021년)	0.0087** (0.0035) [0.1421]
산업 구조 (로그, 2021년)	0.5952*** (0.0233) [0.7575]
통제변수	✓
R^2	0.8780
관측치 수	229

주: 종속변수는 2023년 AI 산업 입지계수이며, 설명변수는 모두 2021년 기준임. 성별·연령 구성 및 인구 규모는 통제변수로 포함함. 각 행에서 첫 번째 수치는 회귀계수, 소괄호 안 두 번째 수치는 표준오차, 대괄호 안 세 번째 수치는 표준화 회귀계수임. 표준오차는 광역시도 단위로 군집화되었으며, 회귀분석은 지역별 고용자 수를 가중치로 부여하여 수행됨. **, ***는 각각 5%, 1% 수준에서 통계적 유의성을 의미함

자료: 연구진 작성

위의 회귀분석을 통해 초기 조건의 지역별 격차가 AI 산업 확산의 병목으로 작용함을 확인하였다. AI 산업의 지역적 집중이 지속될 경우, AI 산업의 혜택이 일부 지역에 집중되고, 비수도권의 산업 전환 및 혁신 역량이 제약될 가능성이 있다. Kalyani et al.(2025)는 1976~2014년 사이에 개발된 신기술 관련 직업이 선도 지역에 장기간 편중되며, 전국적으로 확산되는 데 약 50년이 걸릴 것으로 추정하였다. 이는 AI 산업 육성 정책뿐 아니라, 지역 간 혁신 확산을 촉진할 수 있는 정책이 병행되어야 함을 시사한다. 이에 따라 후속 부분에서는 인적자본, 데

이더센터·에너지, 벤처캐피털, 산업 구조 측면에서 지역 간 혁신 확산을 위한 사회적 의제를 제시한다.

3. 식별된 사회적 의제안

가. (지역01) 데이터센터의 지역 분산을 촉진하기 위한 정부의 역할은 무엇인가?

1) 제안의 배경

데이터센터는 AI 모델 학습, 클라우드 서비스, 고성능 연산 등을 위해 필수적인 핵심 기반 인프라이며, 2024년 한국데이터센터연합회 자료에 따르면 국내 데이터센터의 약 60%가 수도권에 집중되어 있다. 이러한 집중 구조는 연산 인프라 접근성을 지역별로 불균등하게 만들고 있으며, AI 기술 확산 및 산업 생태계 형성 과정에서 수도권 중심 구조를 더욱 강화할 가능성이 제기된다. 기업·연구기관·전문 인력이 수도권에 집중된 상황에서 연산 인프라까지 동일하게 집적될 경우, 비수도권이 혁신 활동을 수행할 유인을 잃고 기술·산업 발전의 지리적 격차가 확대될 위험이 존재한다.

2) 현황 및 쟁점

하이퍼스케일 데이터센터는 특히 대규모 전력 공급, 송전망 안정성, 넓은 부지 확보가 요구되므로 물리적 요건만 고려할 때 지방 입지가 상대적으로 유리하다. 그러나 지방은 수도권과 비교했을 때 초저지연 네트워크 접근성이 낮고, 설계·운영·보안 등 전문 인력을 안정적으로 확보하기 어려운 구조적 제약을 갖고 있다. 이러한 요인으로 인해 데이터센터 운영의 필수 요소인 네트워크 품질, 서비스 연계성, 전력 공급 문제 등이 지방 입지의 핵심 병목으로 작용한다.

또한 데이터센터의 운영 과정에서 발생할 수 있는 소음, 열 배출, 전력 수요 증가에 대한 지역 주민의 우려가 지속되고 있으며, 일자리 창출 효과가 제한적이라는 인식이 결합되어 지역 주민의 수용성 문제가 나타난다. 그 결과, 일부 지방 지자체에서는 유치 계획이 지연되거나 백지화되는 사례가 발생하고 있으며, 물리적 입지

가능성이 실제 투자로 이어지지 않고 있다.

3) 문제의 심각성

데이터센터 구축은 AI 산업 성장의 기반이지만, 갈등·규제·입지 문제로 인해 연산 인프라 공급 속도가 수요 증가 속도를 따라가지 못하는 상황이 지속되고 있다. 데이터센터 입지가 수도권에 집중된 상태가 지속될 경우, AI 모델 개발, 학습, 응용 서비스 배포 등 핵심 가치사슬이 수도권에 더욱 결집할 가능성이 높아지고, 이로 인해 비수도권의 산업 성장 여건이 구조적으로 제한될 수 있다. 장기적으로 이러한 집중 현상은 지역 간 생산성, 혁신 역량, 고급 일자리 창출 측면에서 격차를 확대할 위험을 내포한다.

4) 기존 논의/대책의 한계

기존 정책은 전력 공급, 부지 확보, 환경 인허가 등 물리적 인프라 확충에 초점을 두어 왔지만, 지방 입지의 핵심 병목으로 지적되는 초저지연 네트워크 구축, 전문 인력 확보, 운영 생태계 형성 등 구조적 문제에 대한 대응은 충분히 이루어지지 않았다. 또한 세제 혜택, 전기요금 감면 등 경제적 인센티브가 제공되고 있으나 대형 클라우드 사업자가 요구하는 네트워크 품질과 확장성 측면에서는 실효성이 제한적이다. 더불어 데이터센터가 지역경제에 미칠 수 있는 산업 연계 효과나 장기적 경제 파급효과에 대한 정량적 근거가 부족하여, 지역 주민 및 지자체를 설득할 수 있는 정책적 커뮤니케이션 기반이 충분히 구축되지 못하고 있다.

나. (지역02) AI 확산을 위한 지역별 에너지 인프라 확충 전략은 무엇인가?

1) 제안의 배경

AI 산업의 고도화로 인해 연산 수요가 급증하면서 전력 소비량이 빠르게 증가하고 있으며, 이를 뒷받침하기 위한 안정적 전력 공급 체계 마련이 핵심 과제로 부상하고 있다. 기존 전력 생산은 비수도권 지역의 대형 발전설비에 상당 부분이 의존하고 있으나, 수도권 및 대도시를 중심으로 AI 산업이 집적되는 현 상황에서는 이러한 공급 구조만으로 수요를 충족하기 어려운 한계가 나타난다. 지속적인 AI

산업 성장을 위해서는 전력 공급 확대뿐 아니라 전원 구성의 다변화, 지역별 특성에 기반한 에너지 인프라 확충 전략이 필요하다는 요구가 제기된다.

2) 현황 및 쟁점

전력 생산 잠재력이 높은 비수도권 지역은 발전설비를 추가 구축하여 공급 능력을 강화할 수 있으며, 대규모 부지를 활용한 재생에너지 발전원 확대도 가능하다. 반면, 전력 수요가 집중된 수도권 및 산업 집적 지역은 분산형 전원, 소규모 모듈형 발전, 에너지 저장장치(energy storage system) 등으로 국지적 수요에 대응할 수 있다. 그러나 신규 발전설비 구축 과정에서는 환경 규제, 경제성 확보 문제, 전력시장 구조 등 제도적 제약이 존재하며, 특히 입지 인근 지역 주민의 수용성 문제가 지속적인 난관으로 지적된다. 또한 송전망 확충을 위한 대규모 투자와 장기적 계획 수립이 요구되지만, 비용 부담과 사회적 갈등이 병행되어 정책 추진이 지연되는 경우가 발생한다.

3) 문제의 심각성

수도권 중심의 전력 소비 구조가 유지될 경우, 부지 부족과 송전망 과부하 문제가 누적되어 전력계통의 안정성이 저하될 가능성이 높다. 특히, 초대규모 AI 모델 학습, 데이터센터 가동, AI 서비스 운영 등에서 전력 수요가 예측보다 빠르게 증가할 경우, 기존 전력 공급 구조가 이를 따라가지 못해 병목이 발생할 위험이 존재한다. 전력 공급 확보가 미비할 경우 연산 자원 부족, 서비스 성능 저하, 산업 투자 위축 등 AI 혁신의 확산 과정에서 구조적 제약이 발생할 수 있으며, 에너지 기반시설의 불균형이 지역 간 산업 성장 격차를 더욱 심화시키는 결과로 이어질 우려가 제기된다.

4) 기존 논의/대책의 한계

지금까지의 논의는 재생에너지 확대, 송전망 보강, 발전설비 증설 등 물리적 인프라 확충에 주로 초점을 두어 왔으나, 지역별 전력 수급 구조와 산업 수요 특성을 반영한 전략적 설계는 부족하였다. 특히 비수도권의 발전 잠재력을 수도권 및

산업 밀집 지역의 수요와 연계하는 전력계통 운영 전략, 분산형 전원 및 자가 발전 도입을 유도하기 위한 경제적 유인 설계, 전력망 안정성을 확보하기 위한 저장·전력조정 기술 적용 등 공급-수요-계통을 통합적으로 고려한 접근이 미흡하다. 또한 발전설비 구축이 지역사회에 가져올 편익과 부담을 체계적으로 분석하고 이를 기반으로 주민 수용성을 확보하려는 정책적 소통이 충분하지 않았다. 나아가 AI 산업 특성을 반영한 전력 수요 전망과 국가 단위 장기 에너지 계획 간 정합성이 낮아, 산업 성장 속도와 전력 공급 능력 간 괴리가 발생할 수 있다는 한계점이 존재한다.

다. (지역03) 지역 특화 산업과 AI 융합을 통한 지역 경제 활성화 방안은 무엇인가?

1) 제안의 배경

지역은 각기 다른 산업구조, 자원 기반, 역사적 발전 경로를 보유하며, 이러한 특성은 지역별 비교우위와 산업 발전 방향을 결정하는 핵심 요인으로 작용한다. AI 기술은 생산 공정 자동화, 데이터 기반 의사결정, 제품·서비스 혁신 등을 촉진함으로써 기존 산업의 생산성을 향상시키고 새로운 부가가치 창출을 가능하게 한다. 이에 따라 지역별 특화 산업과 AI 기술을 연계하는 전략은 지역 경제를 활성화하고 산업 생태계의 지속가능성을 확보하는 데 중요한 정책 방향으로 제기된다. 특히 인구 감소와 산업 기반 약화가 심화되는 비수도권 지역에서는 고유 산업을 고도화하는 방식으로 새로운 성장 동력을 마련할 필요성이 부각된다.

2) 현황 및 쟁점

다수의 비수도권 지역에서는 기존 주력 산업이 글로벌 경쟁 심화, 수요 정체, 인력 부족 등의 이유로 성장세가 둔화되고 있으며, 이를 대체할 새로운 산업 기반 발굴이 지연되고 있다. 농업·제조·해양·바이오·문화관광 등 다양한 산업 분야에서 AI를 활용한 스마트팜, 스마트팩토리, 데이터 기반 관광 서비스 등 개별 프로젝트가 추진되고 있으나, 이러한 사례는 산업 생태계 전반으로 확산되기에는 제도

적·기술적 기반이 충분하지 않다. 특히 지역 기업의 디지털 역량, AI 기술 도입 비용, 전문 인력 확보 문제 등이 병목으로 작용하고 있으며, 산업 전반을 지원하는 데이터 인프라, 기술 중개 플랫폼, 연구·교육 체계가 미비하여 확산 속도가 제한되고 있다.

3) 문제의 심각성

지역 특화 산업의 AI 전환이 지연될 경우 생산성 향상과 고부가가치 창출이 이루어지지 못해 산업 경쟁력이 지속적으로 약화될 가능성이 크다. 이는 기업 투자 감소, 양질의 일자리 축소, 지역 내 경제 기반 붕괴로 이어질 수 있으며, 장기적으로는 지역 인구 유출과 산업 생태계 축소로 연결될 위험을 내포한다. 반면 수도권 및 일부 성장 지역에서는 AI 기반 산업 고도화가 빠르게 진전되고 있어, 지역 간 혁신 역량과 경제 성과의 격차가 더욱 확대될 수 있다. 이러한 흐름은 지역 균형발전 뿐만 아니라 국가 전체의 혁신 역량과 생산성 향상에도 제약요인으로 작용할 가능성이 제기된다.

4) 기존 논의/대책의 한계

지금까지의 정책은 지역 산업을 지원하기 위한 R&D 사업, 클러스터 조성 등 개별 프로젝트 중심으로 추진되어 왔으나, 지역 특화 산업과 AI 기술을 체계적으로 연계하는 전략적 접근은 부족하였다. 기술 공급자 중심의 지원 프로그램이 다수였으며, 지역 기업이 실제로 AI 도입을 실행할 수 있도록 비용·인력·데이터·플랫폼을 통합적으로 지원하는 구조가 미비하였다. 또한 산업별 특성에 맞춘 데이터 활용 체계, 표준화된 기술 적용 모델, 지역 대학·연구기관과의 연계 체계 등이 부족하여 기술 확산이 제한되었다. 더불어 지역 단위에서 AI 전환이 가져올 산업 구조 변화, 고용 효과, 경제 파급효과에 대한 정량적 근거가 충분히 구축되지 않아 정책 설계와 예산 배분에 활용하기 어려운 한계가 존재한다.

라. (지역04) AI 스타트업 생태계의 지역 간 확산을 촉진하기 위한 전략은 무엇인가?

1) 제안의 배경

AI 산업의 혁신을 주도하는 스타트업은 기술 실험, 신사업 모델 개발, 연구자·개발자 인재 풀 형성 등 산업 생태계 전반에서 핵심적 역할을 수행한다. 그러나 2024년 기준, The VC 자료에 따르면 국내 AI 스타트업의 약 80%가 수도권에 집중되어 있으며, 이는 자본·인재·기술·조직 경험의 수도권에 집중되는 결과로 이어진다. 이러한 집중은 집적효과를 통해 생산성과 혁신 속도를 높이는 긍정적 측면이 있지만, 동시에 지역 간 혁신 역량의 격차를 심화시키고 비수도권의 신산업 성장 기회를 제약할 가능성이 제기된다. 이에 따라 AI 스타트업 생태계가 특정 지역에 한정되지 않고 전국적으로 확산될 수 있도록 하는 전략적 접근이 요구된다.

2) 현황 및 쟁점

2024년 기준, The VC 자료에 따르면 AI 벤처캐피털 투자액의 약 70%가 수도권 소재 스타트업에 집중되고 있으며, 대학·연구소·대기업·창업보육기관 등 혁신 주체가 수도권에 밀집함에 따라 초기 창업 자금 조달, 산학 연계, 기술 개발 협력, 시장 검증 등 생태계 기반이 수도권 중심으로 형성되어 있다. 반면 비수도권 지역은 AI 전문 투자기관, 기술 검증 환경, 초기 스타트업을 지원할 중간 지원조직 등이 부족하여 창업 자체가 제한되거나 성장 단계로의 전환이 지연되는 사례가 발생한다. 또한 지역별 산업구조, 데이터 접근성, 창업 인프라 수준에 따라 AI 비즈니스 모델을 구현할 기회가 불균등하게 분포하고 있으며, 지역 대학과 연구기관의 기술 사업화 역량도 지역 간 편차가 크다. 그 결과 비수도권은 유망 창업 아이템과 인재가 수도권으로 이전되는 누적적 불균형 구조를 경험하고 있다.

3) 문제의 심각성

AI 스타트업 생태계가 수도권에 집중될 경우 AI 혁신의 성과가 특정 지역에 편중되어 국가 전체의 균형적 기술 발전을 저해할 위험이 존재한다. 이는 지역 산업

의 AI 전환 속도와 새로운 일자리 창출 능력을 약화하고, 장기적으로는 지역 산업 기반 자체가 축소되는 결과로 이어질 수 있다. 또한 지역 기반 AI 창업이 활성화 되지 못할 경우 지역별 데이터·산업·수요 특성을 반영한 AI 서비스 개발이 어려워지고, 이는 AI 도입 효과의 지역 간 편차를 확대하는 요인으로 작용할 수 있다. 나아가 혁신 역량의 지리적 병목이 지속될 경우 AI 산업이 국가 차원의 균형 발전을 견인하는 신성장 동력으로 기능하기 어렵다는 우려가 제기된다.

4) 기존 논의/대책의 한계

지금까지의 정책은 창업 지원, 엑셀러레이팅, 클러스터 조성 등 개별 프로그램 중심으로 추진되어 왔으나, AI 특성을 반영한 기술·데이터·인재·시장 접근성 기반을 지역 단위에서 통합적으로 구축하는 전략은 미흡하였다. 스타트업 지원 정책이 수도권 중심으로 운영되거나 기술 공급자 중심으로 설계된 경우가 많아 지역 기업의 실제 창업·성장·투자 유치로 이어지는 선순환 구조가 약하게 형성되어 있다. 또한 지역 산업구조와 연계된 AI 사업모델 발굴, 지역 대학과 연구기관의 기술 사업화 지원, 지역 특화 데이터·테스트베드 구축 등 지역별 비교우위를 기반으로 한 차별화된 전략이 부족하다. 더불어 지역 창업 생태계 확대에 따른 고용·투자·산업 효과에 대한 정량적 근거가 충분히 축적되지 않아 정책 우선순위 설정과 예산 배분에서 지역 확산 전략의 정당성이 충분히 확보되지 못한 상황이다.

마. (지역05) 지역 대학을 중심으로 한 효과적인 AI 인재 양성 방안은 무엇인가?

1) 제안의 배경

AI 분야의 혁신을 주도하는 핵심 요소는 고급 인적자본이며, 인재가 어느 지역에 축적되는가에 따라 해당 지역의 기술 도입 능력, 산업 고도화 속도, 연구 생산성 등이 크게 달라진다. 현재 AI 관련 전문 인력은 수도권에 집중되어 있어 비수도권 지역이 AI 기반 혁신 활동을 추진하거나 산업 전환 전략을 실행하는 데 제약을 받고 있다. 이에 따라 지역 대학을 중심으로 지역 인재를 양성하고, 이를 지역 산업과 연계해 지역 내에 머물도록 하는 전략의 중요성이 강조된다.

2) 현황 및 쟁점

정부는 지방대학 육성 정책을 추진하며 지원을 확대하고 있으나, AI 관련 학과 신설, 융합 교육 프로그램 운영, 산학 공동 연구 기반 조성 등의 측면에서 지역별 편차가 크다. 일부 대학은 커리큘럼 개편과 연구센터 설치를 통해 대응하고 있으나, 상당수 지역 대학은 교수진 확보, 연구 인프라 구축, 산학 협력 네트워크 형성에서 수도권 대비 열위한 상태에 놓여있다. 또한 지역 산업 구조와 연계된 AI 교육 모델이 충분히 확립되지 않아, 교육이 실제 지역 산업 수요와 연결되지 못하는 사례가 발생하고 있다. 그 결과 교육·연구·산업 간 선순환 구조가 형성되지 못해 지역 차원의 인적자본 축적이 제한된다.

3) 문제의 심각성

지역 대학이 AI 교육 및 연구 기반을 갖추더라도 졸업생의 상당수가 취업 기회를 찾아 수도권으로 이동할 경우, 지역 내 인적자본이 축적되지 못하고 대학의 교육·연구 성과가 지역 산업 발전으로 환류되지 않는 문제가 발생할 수 있다. 이는 지역 대학이 단순한 교육기관에 머물고 지역 혁신 생태계의 실질적 기반으로 기능하지 못하게 만들며, 결과적으로 지역의 기술 채택 능력과 산업 고도화 속도를 저해할 수 있다. 이러한 현상이 누적될 경우 AI 산업 확산 과정에서 지역 간 역량 격차가 더욱 확대되고, 지역 균형발전 정책의 효과가 제한될 가능성이 존재한다.

4) 기존 논의/대책의 한계

기존 정책은 지방대 혁신 지원, 데이터 기반 교육 인프라 구축, 청년 인재 육성 사업 등 개별 과제 중심으로 추진되어 왔으나, 지역별 산업 수요와 연계된 AI 교육·연구 체계를 체계적으로 설계하는 전략은 부족하였다. 대학 교육 강화와 인재 배출에 초점을 둔 정책이 많았지만, 졸업생이 지역 산업으로 유입되고 혁신 주체로 성장할 수 있도록 지원하는 취업·창업·연구 연계 프로그램은 미비하다. 또한 지역 대학 간 협력, 지역 연구기관·공공기관·기업 간 연계, 지역 수요 기반 프로젝트 발굴 등을 통해 지역 생태계 전체를 강화하려는 접근이 부족하다. 더불어 지역 인재 유출의 원인, 인재 정착 유인을 높일 수 있는 요인, 대학의 산업 연결 효

과 등에 대한 정량적 근거가 충분히 축적되지 않아 정책 설계의 기반이 제한적인 상황이다.

제 4 장 5대 분야별 사회적 의제 검증

제 1 절 조사 개요

1. 조사 목적 및 필요성

본 장에서 수행한 전문가 심층조사는 연구진이 도출한 의제(안)들이 AI 시대 우리 사회가 공통으로 논의해야 할 핵심 쟁점과 과제를 적절히 포괄하고 있는지 검증하는 데 목적이 있다. 본 조사는 3장에서 식별한 AI 확산이 초래하는 사회 변화에 대응하기 위해 노동, 행정, 이용자, 정보 생태계, 지역의 5대 분야별로 5개의 핵심 사회적 의제(안)를 대상으로, 해당 의제의 적절성·타당성·시의성을 전문가 관점에서 평가하도록 하였다.

앞선 연구 단계에서는 각 분야별로 AI 확산이 초래하는 구조적 변화, 사회적 갈등 요인, 제도적 공백을 탐색하고, 이를 바탕으로 각 분야별로 5개, 총 25개의 사회적 의제안을 도출하였다. 그러나 의제안이 전문가 집단의 경험적 판단 및 실제 정책 현장에서의 적용 가능성을 충분히 반영하기 위해 사회적 공공성, 의제의 타당성, 논의의 필요성, 사회적 파급성을 전문가 시각에서 평가하고, 구체적인 정책 방안에 대한 정성적 의견을 수렴하고자 하였다. 따라서 본 조사는 각 분야별 전문가로 구성된 검증 패널을 통해 의제의 실효성을 평가하고, 정량·정성적 의견을 반영하여 최종 사회적 의제 체계를 보완하는 것을 목적으로 한다.

2. 조사 설계 및 방법

조사는 5대 분야(노동, 공공·행정, 이용자, 정보 생태계, 지역)의 현안에 정통한 전문가 집단을 대상으로 심층조사 방식으로 진행되었다. 각 분야의 특성을 고려할 수 있도록 학계, 공공 연구기관, 법조계 등에서 전문성(전공)과 활동 분야가 일치

하는 전문가 5인을 선정하여 조사를 의뢰하였다. 조사지는 모든 분야에서 동일한 기본 구조와 평가 기준을 사용하되, 의제 문구만 각 분야 특성에 맞춰 차별화되도록 설계하였다. 조사지의 주요 내용은 아래와 같다. 조사는 비대면 서면조사로 진행하였다.

전문가 조사지 문항 체계

1. 의제안 개요 제시

- ✓ 각 분야별로 연구진이 도출한 5개의 사회적 의제안을 제시
- ✓ 의제별로 제안 배경-현황-문제의 심각성을 요약 제시하여 평가의 공통 이해 기반 제공

2. 의제별 정량 평가(5점 척도)

평가 항목은 다음 4개 기준으로 구성

- ✓ 문제의 공공성: 사회 전체 또는 특정 집단이 공유해야 하는 공공적 중요성
- ✓ 의제의 타당성: 현황·사례 등 실증적 근거 기반 여부
- ✓ 논의의 필요성: 공론화 및 정책 논의의 시급성
- ✓ 사회적 파급성: 논의 확산 가능성과 정책 적용 파급력

3. 의제별 정성 의견 수집

- ✓ 가장 중요하다고 판단되는 의제 3개 선정 및 선정 이유
- ✓ 해당 의제를 해결하기 위한 정책과제·방안 제시
- ✓ 제시 의제의 한계·보완점 의견
- ✓ 연구진이 미처 제시하지 못한 새로운 의제 제안

제 2 절 5대 분야별 사회적 의제 전문가 검증 결과

1. AI와 노동: 의제 평가 및 주요 의견

가. 평가자 구성

본 연구는 3장에서 도출한 5대 분야별 사회적 의제(안)의 타당성과 적절성을 검증하기 위해 분야별 전문가 5인을 대상으로 심층조사를 진행하였다. ‘AI와 노동’ 분야는 노동경제학, 사회학, 법학, 경영학 등 노동 관련 법률, 경제, 사회, 고용 분야의 산학연 전문가 5인이 조사에 참여하였다. 본 조사에서 전문가에게 제시한 노동 분야 5개의 의제는 다음과 같다.

〈표 4-1〉 노동 분야 사회적 의제 전문가 검증 결과

(노동-1) AI로 인한 직무 전환이 요구됨에 따라, 마모되는 근로자의 인적자본을 회복하기 위한 교육훈련 시스템은 어떻게 설계할 것인가?	
제안의 배경	• 에이전틱 AI, 물리 AI로의 진화로 지식 산업과 제조업 등 산업 전반에서 직무가 빠르게 재편 • ‘평생 직장’으로서의 일자리를 위한 ‘평생 학습’이 필수 과제가 됨 • 생애주기 소득의 안정성을 위해서 현직자를 위한 상시 학습 인프라 구현이 시급
현황 분석	• 중장기적 직무 대체와 재편을 전망됨에도, 정부 주도 직업 훈련은 실업자에 편중되고, 재직자의 참여율은 저조한 편 • 중장년층과 저소득층의 심리적·시간적 장벽으로 학습 격차는 확대되고 재교육의 사각지대는 확대
문제의 심각성	• 전 직종에 걸쳐 근로자가 교육과 경력을 통해 축적한 인적자본은 AI 기술이 빠르게 진화함에 따라 급격하게 마모가 시작 • 숙련 인력의 저숙련직 하향 이동이나 노동시장의 조기 이탈로 국가와 기업의 장기적 생산성이 저하되고 사회적 비용 급증 위험
(노동-2) AI가 청년층의 초기 경력 형성에 필수적인 일자리를 대체해 나가는 상황에서, ‘경력 단절’ 없는 청년 고용 시장을 어떻게 보장할 것인가?	
제안의 배경	• AI가 사무, 데이터 입력 등 엔트리 레벨 업무를 빠르게 자동화 • 첫 직무 경험의 기회가 축소되며 청년이 경력 사다리의 ‘첫 번째 칸’에 오를 기회 자체가 박탈 • 이는 단기적·마찰적 실업이 아닌, 세대 단위의 구조적 일자리 상실로 고착화될 위험 증가

현황 분석	• 국내외 연구(Brynyolfsson, 2025; 한진수·오삼일, 2025) 결과, AI 노출도가 높은 초급 일자리에서 청년 고용 급감 확인 • 2025년 2월 기준, 한국 청년층 체감실업률 16.4%(4년래 최고치)로 이미 심각한 고용 불안 상태
문제의 심각성	• 초기의 경력 부재가 장기적으로 고용 불가로 이어지는 악순환 우려 • 청년층의 경력 축적 실패는 국가의 인적자본 손실로 장기적인 노동생산성 하락으로 연결 • 또한 장기적으로 실업 급여, 사회 부조 등의 사회적 비용 급증 초래
(노동-3) AI로 인한 노동시장의 창조적 파괴 과정에 적응하기 위해 발생하는 전환 비용은 누가, 어떤 비율로 분담되어야 하며, 배분의 우선순위는 어떤 사회적 기준으로 정할 것인가?	
제안의 배경	• AI로 인한 직무의 '파괴'(빠름/집단적)와 '창출'(느림/분산적) 간에는 속도·시점의 불일치가 발생 • 막대한 전환 비용(재훈련, 새 일자리로의 이주 등)의 발생으로 근로자의 소득과 숙련의 공백, 지리적 격차가 필연적으로 발생
현황 분석	• 차이나 쇼크(Autor, 2013), 로봇(Acemoglu & Restrope, 2020) 충격은 대체 효과가 창출 효과를 상회하거나 회복이 느림이 확인 • AI 시대에도 대체되는 근로자와 수혜를 입는 근로자가 서로 다른 집단일 가능성이 높음 • AI를 노동 증강형으로 활용해 중산층을 재건할지는 이해관계자 간의 집단적 선택의 문제(Autor, 2024)
문제의 심각성	• 기술 충격 이후의 노동시장의 회복은 느리거나 실패할 수 있어 장기적 임금과 고용 하락 및 경제활동참가율의 감소가 고착화될 가능성 • AI 도입의 이익 집단과 비용 부담 집단의 불일치 지속 • 전환 비용의 부담이 편중될 경우, 불평등과 사회적 갈등 증폭 우려
(노동-4) AI가 구현하는 알고리즘 매니지먼트(채용, 업무 배분, 감시, 평가)가 노동 통제를 강화하고 노동의 질을 저해하는 상황에서, 우리는 근로자의 기본적인 권리를 어떻게 보호할 것인가?	
제안의 배경	• 채용-배치-감시-평가-해고까지 알고리즘 매니지먼트가 스며들며 노동의 질과 절차적 권리(설명과 이의제기)까지 직접 영향을 받게 됨 • 실시간으로 수집되는 데이터 기반 과도한 모니터링과 불투명한 의사결정은 근로자의 자율성과 존엄성 침해 위험
현황 분석	• 알고리즘 매니지먼트는 플랫폼 노동을 넘어 콜센터, 물류창고, 사무직 등 전 영역으로 확산 • 보스웨어 등을 통한 감시와 사생활 침해와 AI 채용의 편향성과 공정성 관련한 시비 발생 • 근로자의 알권리에 알고리즘 투명성을 포함하는 제도적 정비의 시도가 있으나 기업 이해와 상충되므로 실제 규범은 여전히 미성숙

문제의 심각성	- 기업의 경우, 생산성을 위해 감시와 통제를 강화할 유인이 있으며, 근로자는 디지털 파놉티콘 하에 스트레스와 존엄성이 침해 - AI 통제는 노동의 비인간화하며, 디지털 테일러리즘을 부활시켜 근로자의 자율성과 창의성에 기반한 숙련 형성 기회를 박탈
(노동-5) AI가 창출한 초과 이익이 소수 기업에 독점되고 임금과 고용상승으로 연결되는 밴드 왜건 효과가 부재한 상황에서, 이 이익의 범위는 어떻게 정의하며 누구와 어떤 원칙으로 공유되어야 하는가?	
제안의 배경	- AI로부터의 이익은 거대 모델, 데이터, 컴퓨팅 파워 등 무형 자산을 보유한 소수 기업에 편중 - 데이터를 생성하는 주체인 대중과 직무 전환의 비용 부담이 요구되는 근로자는 이익 분배 과정에서 배제 - AI는 노동시장을 위축하여 AI로부터의 이익이 임금과 고용 상승으로 연결되지 못해 새로운 분배 원칙에 대한 고민이 필요
현황 분석	- 전환 비용 재편 마련을 위한 다양한 방안이 제기되나 각각의 방안은 운영 방식의 한계와 재원의 지속성에 대한 장단점이 혼재 - 일례로, 캘리포니아주에서는 디지털 경제의 대중의 기여에 대한 보상을 위한 데이터 배당제 도입을 시도 - 전환 비용 재원을 마련하기 위한 로봇세, AI세 부과와 기본 소득 등의 대안적 분배 방식도 논의
문제의 심각성	- AI 혜택이 소수에 독점되어 경제적 불평등과 사회적 양극화 심화 - 근로자의 임금 감소는 사회 전체의 총수요 위축으로 연결

자료: 연구진 작성

나. 정량 평가 결과

‘AI와 노동’ 분야 전문가 5인을 대상으로 5개 의제(안)에 대해 4개 항목(문제의 공공성, 의제의 타당성, 논의의 필요성, 사회적 파급성)을 5점 척도로 평가하도록 하였다. 전체적으로 식별한 노동 분야 5개 의제 모두 4개 항목에서 전반적으로 ‘높음(4.0점)’ 이상의 높은 평균 점수를 획득하여, 본 연구진이 도출한 의제(안)들이 AI 시대 노동 분야의 핵심 쟁점을 타당하게 포괄하고 있음을 확인하였다.

의제별 종합평균을 기준으로 보면 ‘(노동-2) 청년 고용’ 의제는 5인 평균 4.60점으로 가장 높게 평가되었으며, ‘논의의 필요성’(4.80점)과 ‘사회적 파급성’(4.60점)에서 최고점을 기록하였다. 이는 AI로 인한 청년층의 초기경력 형성 기회 박탈 문제가 가장 시급하고 중요한 사회적 과제로 부상하고 있음에 공감대가 형성된

결과로 볼 수 있다. ‘(노동-1) 교육훈련 시스템’과 ‘(노동-3) 전환 비용’ 의제도 4.40점의 높은 평균 점수를 받았다. ‘(노동-4) 알고리즘 통제’(4.35점)와 ‘(노동-5) 이익 공유’(4.25점) 의제 역시 모든 항목에서 4.0점 이상의 높은 점수를 받아 의제로서의 타당성을 인정받았다.

항목별로는 ‘(노동-3) 전환 비용’과 ‘(노동-4) 알고리즘 통제’가 ‘문제의 공공성’ 측면에서 4.80점으로 가장 높은 공감대를 얻었으며, ‘(노동-4) 알고리즘 통제’ 의제는 ‘사회적 파급성’ 항목에서 3.60점으로 상대적으로 낮게 평가되었다.

〈표 4-2〉 노동분야 의제 전문가 정량 평가 결과

구분	문제의 공공성	의제의 타당성	논의의 필요성	사회적 파급성	의제별 평균
(노동-1) AI 직무 전환 → 인적자본 마모 대응	4.60	4.40	4.60	4.00	4.40
(노동-2) 청년 초기경력 단절 대응	4.40	4.60	4.80	4.60	4.60
(노동-3) 전환 비용의 분담과 사회적 기준	4.80	4.20	4.40	4.20	4.40
(노동-4) 알고리즘 매니지먼트와 노동권 보호	4.80	4.60	4.40	3.60	4.35
(노동-5) AI 초과이익의 정의 및 공유 원칙	4.60	4.40	4.00	4.00	4.25
평가항목별 평균	4.64	4.44	4.44	4.08	

자료: 연구진 작성

다. 정성 의견 분석

5개 의제 중 가장 중요하고 우선순위가 높아야 할 의제 3개를 선정한 결과 ‘(노동-2) 청년 고용’ 의제는 5인의 전문가 전원이 Top 3 의제로 선정하여 가장 높은 시급성과 공감대를 형성하였다. ‘(노동-1) 교육훈련 시스템’ 의제는 2명의 전문가가 1순위로 지목하는 등 총 4회 선택되어 그 뒤를 이었다. 반면 ‘(노동-3) 전환 비용’ 의제는 1명의 전문가만 2순위로 선택하여 타 의제 대비 우선순위에서 밀리

는 것으로 나타났다.

〈표 4-3〉 노동 분야 의제 전문가 우선순위 평가 결과

구분	1순위 선정횟수	2순위 선정횟수	3순위 선정횟수	총 선정횟수
(노동-1) AI 직무 전환 → 인적자본 마모 대응	2	0	2	4
(노동-2) 청년 초기경력 단절 대응	1	3	1	5
(노동-3) 전환 비용의 분담과 사회적 기준	0	1	0	1
(노동-4) 알고리즘 매니지먼트와 노동권 보호	1	1	1	3
(노동-5) AI 초과이익의 정의 및 공유 원칙	1	0	1	2

자료: 연구진 작성

각 의제를 선정한 전문가들의 선정 사유는 아래와 같이 정리하였다. 5개 의제 중 (노동-2) ‘청년 고용’을 5인 전원이 선택하며 가장 시급하고 중요한 과제로 선정하였으며, AI가 엔트리 레벨 직무를 대체하며 청년층의 경력 사다리 진입 자체를 박탈하고, 이는 장기적인 국가 경쟁력 하락과 세대 간 불평등으로 이어질 수 있으나 사회적 경각심은 낮다는 점이 핵심 선정 사유로 제시되었다. 다음으로 (노동-1) ‘교육훈련 시스템’이 1순위로 2회 지목되었는데, 기존의 취약한 대응 시스템을 넘어 현재 고용보험 기반 직업교육 시장의 ‘독과점적 비정상 운영’을 혁신해야 하는 ‘매우 높은 정책적 난이도’를 가진 근본적인 문제라는 점이 강조되었다. 또한 (노동-4) ‘알고리즘 통제’는 현행 노동법이 AI의 시스템적 통제를 규율하지 못하는 한계와 노동자의 기본권 침해 때문에, (노동-5) ‘이익 공유’는 기업의 ‘비용 사회화’와 ‘AI 양극화’에 대응하는 장기적 과제라는 측면에서 중요하게 다루어져야 함이 강조되었다.

〈표 4-4〉 노동 분야 의제 전문가 우선순위 선정 사유

의제번호	전문가 근거 요약
노동01	• 기존 직업훈련 구조는 ‘실업자 중심’, ‘단기 과정 중심’, ‘평생교육과의 단절’ 문제 존재 • 현재 고용보험 기반의 직업교육 시장이 ‘독과점 체계로 운영’되고 있음을 지적, 교육의 내용적 혁신을 넘어 기존 교육시장 구조를 혁신해야 하는 높은 정책적 난이도가 요구 • AI·로봇 시대에는 하드스킬+소프트스킬 결합형 역량체계가 필요 • 교육 콘텐츠·평가·인증체계 전환을 포함한 ‘시스템 레벨의 overhaul’ 필요
노동02	• AI가 화이트칼라 직종의 엔트리 레벨 업무를 대체하며 신규 인력의 숙련 개발 기회를 제한하고 있으며, 이는 장기적으로 산업 및 국가 경쟁력 하락에 치명적 요인 • AI로 인한 신입 채용 감소가 청년층의 노동시장 진입 자체를 가로막아, “세대 간 불평등 문제”로 비화될 수 있음을 우려 • 기업들은 단기 효율성 때문에 청년 신규 고용을 축소하고 있으며, 이는 “세대 단위의 구조적 일자리 상실”로 이어질 수 있음 • 청년층의 경력 단절이 빈곤 증가와 비정규직의 원인
노동03	• AI 전환으로 인해 재교육 및 고용서비스 대상자가 급격히 증가할 것이며, 이에 따른 “재원(고용보험 기금)도 한계에 이를 것”으로 예상 • 증가할 비용을 고려하여 “사회적 분담을 어떤 방식으로 가져가야 할지” 논의 필요
노동04	• AI가 채용, 평가, 해고 등 전통적 노동법 규율 영역과 근태관리 등 기업 재량 영역 모두에 개입하며 관리의 실질을 강화 • 현행 노동법은 ‘AI 기반 통제’의 새로운 위험을 포착하기 어려움 • 알고리즘 매니지먼트가 노동을 개별화시켜 기존의 동질성을 바탕으로 한 연대나 조직화를 어렵게 만들 • EU나 ILO와 달리, 한국에서는 정보주체로서 노동자의 권리에 대한 보장 논의가 거의 없음을 지적하며 논의의 필요성을 강조
노동05	• (노동-3) ‘전환 비용’이 단기적 문제라면, (노동-5) ‘이익 공유’는 “장기적 문제”라고 규정 • AI 기반 기업이 비용의 사회화, 공정거래질서의 훼손, 사회 규범의 약화를 유발하고 있어, 이에 대한 근본적인 접근이 필요 • AI의 기술적 이익은 대기업에 집중되고 전환의 사회적 비용은 취약계층에 집중되는 “AI 양극화” 문제에 대응해야 하며, 현재 이 논의가 정책 아이디어 차원에 머물러 있음

자료: 연구진 작성

전문가들은 우선순위로 선정한 의제들에 대해 다음과 같은 구체적인 정책과제를 제안하였다. (노동-1)에 대해서는 기존의 실업자·학원 중심 훈련을 재직자·대학 중심의 ‘평생학습 시스템’으로 전환하고, 하드스킬이 아닌 소프트스킬 교육으로 대전환하며, 나아가 기존 직업교육 시장의 독과점 구조를 파악하고 기술 저항감 해소를 위한 ‘심리상담’ 등을 포함한 포괄적 지원을 강조한 의견도 제시되었다. (노동-2) ‘청년 고용’ 문제는 기업의 사회적 책임을 강조하고 인센티브를 제공하며, (노동-4) ‘알고리즘 통제’는 기술 도입 시 민주적 거버넌스와 노동자의 데이터 권리를 보장하는 방향으로 「근로기준법」 등 현행법 개정이 필수임이 강조된다. 마지막으로 (노동-3, 5)와 관련해서는 정부·지자체·산업계의 전환 비용 분담 원칙을 세우고, ‘이익 환수제’나 ‘AI세’ 등 새로운 재원 마련 방안에 대한 실험적 연구가 필요함이 제기되었다.

〈표 4-5〉 노동 분야 의제 전문가 정책과제 제안

의제번호	전문가 정책과제 제안
노동01	• 시스템의 전면 개편: 기존의 실업자·학원 중심 훈련에서 벗어나, 재직자 중심 및 ‘대학 중심’의 ‘평생학습 시스템’으로 전면 개편. 기업 현장 중심의 실무 재교육을 통해 ‘일과 학습의 통합’을 추구. 기존 직업교육 시장 개혁 • 교육 내용의 대전환: 단순 코딩 등 하드스킬(Hard skill) 교육을 넘어, 의사소통, 협업, 문제해결 등 AI 시대에 더욱 중요해지는 ‘소프트스킬(Soft skill)’ 체득화 교육으로의 대전환 • 대상자 맞춤 지원: 기술 수용성이 낮은 중장년·저숙련 노동자에게는 성과지표와 상관없이 기술 자체를 충분히 사용하고 익숙해질 시간과 기기 보급을 보장 • 거버넌스 구축: 노·사·정·학이 유기적으로 협력하여 새로운 교육체계를 개발할 수 있는 거버넌스를 구축
노동02	• 기업의 역할 유인: 민간 기업이 AI로 저숙련 업무를 대체하더라도, 미래 숙련자 양성이라는 ‘사회적 책임’ 관점에서 청년 채용을 유지하도록 인센티브(고용보험료 감면, 세제 혜택 등)를 제공 고려 • 노동 가치 및 경력 재설계: ‘기계-인간 분업’ 양상이 변화하는 것을 전제로 기술로 대체할 수 없는 인간 노동(예: 현장 교육 직무)의 ‘사회적 가치를 재조정’하고, 새로운 커리어 패스와 숙련 체계를 설계 • 정부의 연계 프로그램: 프로젝트와 일 경험을 인증하는 ‘커리어인증제’와 ‘고용서비스-직업훈련-인턴십’을 연계한 패키지형 프로그램 추진

의제번호	전문가 정책과제 제안
노동02	• 부처 간 협력: 기존 고교/대학의 인턴십 프로그램을 재검토하고 ‘교육부와 고용노동부’ 간의 긴밀한 협력 체계를 구축 • 다층적 사회안전망: 청년층을 대상으로 ‘소득보장’(현금급여)과 ‘직업훈련·상담’(서비스)을 다층적으로 도입하는 사회안전망 재설계
노동03	• 비용 분담 주체: 재교육, 고용안정 지원금 등 다양한 항목의 비용을 정부, 지자체, 산업계 모두가 분담하기 위한 원칙 수립 • 재원 마련 방안: AI 전환으로 부가가치와 이익이 급상승하는 부문(플랫폼, IT 등)에 대해 ‘재건축초과이익환수제’와 유사한 형태의 이익 환수에 대한 연구와 논의가 필요하다고 제안. 다만, 이러한 제도가 자본주의 경제체제에 부합하는지에 대해서는 사회적 논의와 합의 필요
노동04	• 노동자 데이터 권리: EU의 GDPR(정보 이동권)과 같이, 노동자가 자신의 데이터를 통제하고, 노동조합 등에 이관하여 집합적 권리 주장을 위해 사용할 수 있도록 하는 ‘정보주체로서의 권리’와 디지털 시민권 프레임 정립 • 법제도 개선: 알고리즘 통제 문제 해결을 위해서는 「근로기준법」, 「노동조합법」 등 현행 법률 개정이 필수적이며, 이를 위해 ‘사회적 대화’와 ‘정치권의 협조’를 우선 확보 • 민주적 거버넌스: 기업이 기술을 도입하는 경영상 결정 단계부터 근로자 측의 참여를 촉진하는 ‘민주적 거버넌스 체계’를 마련
노동05	• 선행 연구: AI 양극화의 구체적 실태(수혜/피해 집단) 파악을 위한 정량적 연구가 다수 선행될 필요 • 정책 아이디어 실험: 로봇세, AI세, 데이터 배당제 등 다양한 정책 아이디어에 대한 ‘실험적 적용’과 기존 분배 정책과의 연계 방안을 검토 • 기업의 사회적 책임: 기업이 AI 기반 생태계 구축 과정에서 사회에 전가된 비용에 대한 사회적 책임 측면에서의 대응 필요성 강조

자료: 연구진 작성

라. 의제별 핵심 보완 의견

전문가들은 연구진이 제시한 5개 의제의 방향성에는 대체로 동의하면서도, 의제들이 간과하고 있는 핵심 이해관계자와 관점에 대한 보완이 필요하다고 지적했다. 우선 ‘기업’과 ‘노동조합’의 역할에 대한 강조이다. 한 전문가는 5개 의제 모두에 ‘기업’과 ‘전통적인 노동계(노동조합)’를 고려할 필요성을 제기하였다. 의제 해결을 위해서는 기업의 자발적인 협력이 필수적이므로, 기업에 대한 ‘동기부여 방안’이 함께 논의되어야 하며, 동시에 막대한 영향력을 가졌으나 신규 노동자 조

직화에 실패하고 있는 기존 노동조합을 어떻게 설득하고 변화시킬 것인지에 대한 논의가 필요하다고 보았다.

다른 한편에서는 의제에서 추상적 예측보다 ‘구체적 사실’이 필요함을 제언하였다. ‘(노동-3) 전환 비용’과 ‘(노동-5) 이익 공유’ 의제는 ‘추상적 예측’에 기반한 측면이 있으므로, AI가 노동시장에서 실제로 만들어내고 있는 ‘구체적인 사실(facts)’로도 의제를 살펴볼 필요가 있다는 의견이 제시되었다. 이와 마찬가지로 (노동-5) 의제는 ‘밴드왜건 효과 부재’나 ‘초과 이익 독점’이라는 전제가 논란의 소지가 있다는 지적이 제기되었다. ‘초과 이익’의 정의와 ‘공유의 당위성’에 대한 사회적 합의가 전제되어야 하므로, 의제화를 위해서는 근거 보완이 필요하다는 의견이다.

(노동-1) ‘교육훈련’ 의제가 세분화될 필요성도 제시되었다. ‘(노동-1)’ 의제는 AI 리터러시를 근로자의 지위와 역할(최고결정자, 관리자, 실무자)에 따라 매트릭스를 만들어 세분화하고, 기술 수용성이 낮은 중장년층에게는 기술 자체에 ‘익숙해지는 시간’을 보장하는 등 콘텐츠의 질적 변화를 강조해야 한다는 의견이 있었다. 마지막으로 5개 의제를 아우르는 ‘통합적인 패러다임’ 설계가 필요하며, ‘청년층’ 중심의 접근을 넘어 향후에는 산업별, 기업 규모별, 고용 형태별 등 다양한 ‘계층별’ 접근으로 확대할 필요가 있다는 의견도 있었다.

전문가들은 AI 시대 상생과 포용을 위한 노동 분야에 새로운 의제안으로 다음과 같은 안을 제시하였다.

〈표 4-6〉 노동분야 전문가 신규 제안 의제

의제 제목	제안 내용 및 배경
AI 기술과 젠더 불평등	• 제안 내용: “AI 기술이 성별분업을 어떻게 심화시키는가?” • 제안 배경: AI 개발 직종의 성별화, AI가 적용되는 직종(사무직 vs 현장직), 플랫폼 노동(배달 vs 가사) 등에서 이미 성별 분업이 뚜렷하다. AI 기술의 개발과 적용 과정을 젠더 관점에서 평가하고 이것이 노동시장 불평등을 심화시킬 요소를 점검해야 한다.
AI 도입 과정에서의 민주적 거버넌스	• 제안 내용: “기업과 정부가 AI 기술을 도입하고 정책을 추진할 때, 노동자 등 이해관계자의 참여를 촉진하기 위하여 필요한 민주적 거버넌스 체계를 어떻게 구축할 것인가?” • 제안 배경: AI의 사회적 효과는 결국 기술이 아닌 ‘조직’과 ‘의사결정’의 문제이며, 현장 노동자가 배제된 AI 도입은 저항과 불평등을 심화시킬 수 있다. 따라서 알고리즘 관리 도입 시 근로자의 참여와 정보 접근권을 보장하고, 국가 차원에서도 정책 결정 과정에 이해관계자 참여를 제도화하는 것이 시급하다
전 국민의 AI 인재화	• 제안 내용: “전 국민의 AI 인재화를 위한 사회 운동은 어떻게 전개할 것인가?” • 제안 배경: AI 일상화에 따라 AI 역량은 지식정보화 시대의 인터넷망과 같은 사회적 인프라가 되었다. 이는 저출생 시대의 국가 경쟁력 회복과 직결되므로, 기술 교육을 넘어선 사회 운동(캠페인) 차원의 접근이 필요하다.
노동조합의 변화와 사회적 대화 주제	• 제안 내용: “AI로 인한 노동시장 재편 국면에서, 노동조합은 어떻게 변화되어야 하는가? AI와 노동 의제를 논의할 수 있는 사회적 대화의 주체는 어떤 기준으로 확정할 수 있는가?” • 제안 배경: 현재의 전통적 노동조합은 새로운 노동자들을 대표할 역량이나 인식이 현저히 부족하다. 따라서 기존 노동조합의 혁신을 촉구하고 새로운 조직의 등장을 촉진하여, 실질적인 사회적 대화 주체를 정립하는 논의가 필요하다.

자료: 연구진 작성

2. AI와 행정: 의제 평가 및 주요 의견

가. 평가자 구성

3장에서 다루었던 ‘AI와 행정’ 분야의 사회적 의제(안)의 타당성과 적절성 검증에 전문가 5인이 참여하였다. 참여 전문가는 AI 정책 및 정보보호정책 교수, 행정학 교수, 법학 및 입법연구자, 스마트 행정 연구자, 인공지능 안전 정책 및 글로벌 협력 연구자이다. 본 조사에서 전문가에게 제시한 공공행정 분야 5개 의제(안)은 다음과 같다.

〈표 4-7〉 공공행정 분야 사회적 의제 전문가 검증 결과

(행정-1) 접근성 격차와 알고리즘 편향 관점에서 디지털 취약계층은 소외되지 않을 수 있는가?	
제안의 배경	• AI로 인한 취약계층의 접근 배제와 알고리즘의 불공정한 기준 내재화 위험은 소외 문제를 심화시킬 수 있음
현황 분석	• 고령층의 디지털 역량 점수가 다른 취약계층 대비 현저히 낮은 상황으로, 정책적 혜택조차 활용 능력 부족으로 소외될 위험이 높음. 또한 편향성을 내재한 알고리즘으로 차별적 의사결정이 발생함
문제의 심각성	• 접근 격차는 불편과 소외를 야기하며, 시민으로서의 기본 권리 및 기회 박탈로 이어짐. 알고리즘 편향은 눈에 보이지 않는 차별을 시스템화하고 영구화하는 결과를 초래할 수 있음
(행정-2) AI 기반 공공·행정서비스에 오류가 발생하여 시민이 피해를 입었을 때, 책임 공백은 해소될 수 있는가?	
제안의 배경	• AI의 블랙박스(Black Box) 특성으로 인해 시민에게 피해 발생 시 책임 소재 규명이 어려움. 따라서 AI 시스템 오류로 시민에게 치명적 피해가 발생할 경우, 책임 공백이 발생할 가능성이 높음
현황 분석	• 일반시민 조사 결과, 시민들은 ‘정책 결정 과정에서 발생하는 책임소재 불분명’에 대해 우려 수준 대비 정부의 준비 수준이 낮다고 평가함. AI 기반 시스템의 활용에 대한 근거는 마련되는 한편, 오류 발생 시 피해 배상의 책임 주체와 절차에 대한 명확한 가이드라인은 부재한 상황임
문제의 심각성	• AI 기반 서비스 오류 시 책임 주체는 알고리즘 개발사, 데이터 제공 기관, AI 도입 정부 부처, 공무원, 그리고 AI 시스템 자체 등 복잡하게 얽혀 있어 판단이 매우 어렵고, 책임을 회피할 가능성이 있음. 치명적인 피해 발생 가능성은 AI 공공 서비스에 대한 불신으로 이어지며, AI 도입을 꺼리는 소극 행정이나, 모든 책임은 AI로 돌리는 무책임한 행정을 초래할 가능성이 있음
(행정-3) ‘시민 안전’과 ‘범죄 예방’을 명분으로 개인의 일거수일투족을 감시하고 통제하는 ‘빅브라더’ 사회로의 귀결을 막을 수 있는가?	
제안의 배경	• 공공 CCTV와 AI 기술 결합으로 대규모 실시간·자동화 감시가 가능해지면서, 사실상 모든 시민이 감시 영향 하에 있을 가능성이 높아지는 만큼 기본권 보호 논의가 시급함
현황 분석	• 대부분의 지방자치단체는 대규모 예산을 투입하여 CCTV와 AI 기술을 결합하고 있으나, 기술 도입의 필요성과 효과에 대한 사전 평가가 충분히 이루어지지 못한 채 행정 당국의 판단에 의해 일방적으로 결정됨
문제의 심각성	• 군집 사고 예방에 기여하는 등 긍정적인 측면이 있지만, CCTV에 포착되는 모든 이들의 정보를 실시간으로 수집하고 미래 행동을 예측할 수 있다는 점에서 논란의 여지가 있음. 본래 목적을 넘어 특정 개인의 동선 추적, 집회 참여자 식별 등에 활용될 수 있다고 판단되면, 시민들의

문제의 심각성	자유로운 의사 표현, 집회 및 결사의 자유 등 민주주의 근간이 되는 기본권 행사를 스스로 검열하고 포기하게 될 수 있음
(행정-4) 기술적 안정성 및 신뢰 차원에서 개인정보 보호(해킹 취약, 신종 범죄에 악용 등)는 가능한가?	
제안의 배경	• AI 공공·행정서비스는 국민의 개인정보를 활용하기 때문에 해킹 발생 시 국가 행정망 마비로 이어질 수 있음
현황 분석	• 2024년 개인정보 유출 신고 원인의 과반(56%)이 해킹으로 인한 것이며, 특히 공공기관 관리자 페이지를 겨냥한 해킹이 증가함. 서울시는 'AI 기반 사이버보안 종합계획'을 수립하는 등 대응 노력을 하고 있으나, AI를 활용하여 해킹 공격을 수행하는 등 'AI 간의 싸움'으로 예측되어 보안 환경은 더 복잡해질 것으로 전망됨
문제의 심각성	• AI 시스템은 다양한 구성 요소로 이루어져 있어, 기존 IT 시스템보다 잠재적인 해킹 경로(취약점)가 더 넓음. AI 모델 자체가 학습 데이터를 암묵적으로 기억하는 특성이 있어, 해커가 모델 분석을 통해 원본 개인정보를 추출해낼 위험이 존재함. 나아가 핵심 행정망 교란 등 국가 안보를 직접적으로 위협하는 도구로 악용될 수 있음
(행정-5) AI 기반의 정책 결정 과정(이유 및 근거)에 대해 시민이 납득할 수 있는 설명을 제공할 수 있는가?	
제안의 배경	• AI 기반 정책 결정의 이유와 근거가 명확히 제시되어야 시민의 이의제 기권과 권리 구제가 보장될 수 있는데, 모델의 정확성과 설명 가능성은 상충 관계임
현황 분석	• 한국전자통신연구원이 정책 의사결정을 돕는 AI 정책지능 연구 협력체계를 구축하고, 영국 지방정부가 주택정책 결정에 AI를 도입하는 등 AI 정책 의사결정은 점차 심화될 것임. 이에 따라 AI 설명 가능성에 대한 논의가 활발히 진행 중이나, 시민이 납득할 수 있는 설명의 수준과 범위에 대한 사회적 합의가 부족한 실정임
문제의 심각성	• 설명 가능한 인공지능(XAI)은 AI 모델 해석에 집중하지만, 기술적으로 완벽한 설명이 비전문가인 시민이 이해할 수 있는 직관적이고 맥락적인 설명으로 전환되는 수준은 상대적으로 미흡함. 모델의 정확성과 설명 가능성 확보는 상충 관계에 있어 두 가지를 충족하기 어려운데, AI 결정 과정에 대한 설명이 어려울 경우 시민들은 행정 결정에 대한 근본적인 불신을 갖게 되며, 이는 AI 공공 서비스 수용을 낮추고 국가 정책 집행의 민주적 정당성을 훼손할 가능성이 높음

자료: 연구진 작성

나. 정량 평가 결과

‘AI와 행정’ 분야 전문가 5인이 각 의제(안)의 공공성, 타당성, 논의의 필요성, 사회적 파급성 등 4개 항목을 5점 척도로 평가하였다. 각 의제(안)의 종합 평균을 비교해 보면, 먼저 ‘(행정-4) 기술적 안정성 및 개인정보 보호의 취약성 심화’ 의제가 가장 높게 평가되었다(4.60). 그다음으로는 ‘(행정-1) 접근성 격차 및 알고리즘 편향에 따른 디지털 취약계층 소외’ 의제(4.55)이다. 그다음으로 ‘(행정-2) 오류 발생 시 책임 공백 해소 방안 부재’ 의제(4.25)가 뒤를 이었다.

항목별로는 ‘(행정-1) 접근성 격차 및 알고리즘 편향에 따른 디지털 취약계층 소외’가 ‘논의의 필요성’ 측면에서 5점으로 가장 높은 공감대를 얻었으며, ‘(행정-5) AI 정책 결정 과정에 대한 시민의 납득 가능한 설명 제공 가능성’ 의제는 ‘사회적 파급성’ 항목에서 3.8점으로 상대적으로 낮게 평가되었다.

〈표 4-8〉 공공행정 분야 의제 전문가 정량 평가 결과

구분	문제의 공공성	의제의 타당성	논의의 필요성	사회적 파급성	의제별 평균
(행정-1) 접근성 격차 및 알고리즘 편향에 따른 디지털 취약계층 소외 문제	4.8	4.4	5	4	4.55
(행정-2) AI 기반 공공·행정서비스 오류 발생 시 책임 공백 해소 방안 부재 문제	4.4	4.2	4.2	4.2	4.25
(행정-3) 첨단 감시 기술 확산에 따른 ‘빅브라더’ 사회 귀결 위험 문제	4	3.8	4	4.4	4.05
(행정-4) AI 시대 기술적 안정성 및 개인정보 보호의 취약성 심화 문제	4.4	4.8	4.4	4.8	4.60
(행정-5) AI 정책 결정 과정에 대한 시민의 납득 가능한 설명 제공 가능성 문제	4.4	4	4	3.8	4.05
평가항목별 평균	4.40	4.24	4.32	4.24	

자료: 연구진 작성

다. 정성 의견 분석

5개 의제 중 가장 중요하고 우선순위가 높아야 할 의제 3개를 선정한 결과 ‘(행

정-1) 디지털 취약계층 소외’ 의제는 2명의 전문가가 1순위로 지목하는 등 총 5회 선택되어 가장 높은 시급성과 공감대를 형성하였다. 반면 ‘(행정-5) AI 정책 결정에 대한 시민의 납득 가능한 설명 제공’ 의제는 1명의 전문가만 2순위로 선택하여 타 의제 대비 우선순위에서 밀리는 것으로 나타났다.

〈표 4-9〉 공공행정 분야 의제 전문가 우선순위 평가 결과

구분	1순위 선정횟수	2순위 선정횟수	3순위 선정횟수	총 선정횟수
(행정-1) 디지털 취약계층 소외	2	0	3	5
(행정-2) 오류 발생 시 책임 공백 해소 방안 부재	1	2	0	3
(행정-3) ‘빅브라더’ 사회 귀결 위험	0	1	2	3
(행정-4) 기술적 안정성 및 개인정보 보호의 취약성	2	1	0	3
(행정-5) AI 정책 결정에 대한 시민의 납득 가능한 설명 제공	0	1	0	1

자료: 연구진 작성

각 의제를 선정한 전문가들의 선정 사유는 아래와 같이 정리하였다. 5개 의제 중 (행정-1) ‘디지털 취약계층 소외’ 의제를 가장 시급하고 중요한 과제로 선정하였다. 전문가들은 디지털 역량 격차를 단순한 기술 활용 문제를 넘어, 기본권, 참여권 등으로 연결될 수 있는 주요한 이슈라고 보았다. 고령층, 저소득층 등 취약계층의 공공서비스 접근에 대한 제약은 지역간·세대간 격차를 심화할 수 있으며, AI의 알고리즘 편향이 제도적으로 보호해야 할 대상을 배제하거나 차별의 정도를 악화시킬 수 있다는 점에서 논의의 필요성이 높다고 보았다. 특히 최근에는 AI 챗봇을 기반으로 한 민원 서비스가 확대되면서 복잡한 상담이 필요한 경우에는 오히려 서비스의 질이 저하되거나, 정책 수혜 자격이 있어도 디지털 신청 절차를 완수하지 못해 권리를 포기하는 고령자의 사례가 발생하고 있음을 지적하였다. 그다음으로 (행정-4) ‘기술적 안정성 및 개인정보 보호의 취약성’이 1순위로 2회 지목되었다. 그 이유로 전문가들은 내부 행정서비스와 대민서비스에 AI 활용도가 높

아지고 AI 학습에 개인정보가 활용될 유인과 가능성이 증가함에 따라, 개인정보 보호와 안전성이 정책·제도적으로 담보되지 않으면 책임 없는 활용으로 흘러갈 우려를 제기하였다. 특히 개인정보 침해로 인한 잠재적인 피해 대상이 전체 시민이라는 점과, 개인정보 침해로 인한 사이버 범죄와 경제적 피해의 심각성이 사회적 실제로 확인된다는 점을 지적하였다. 그다음으로는 (행정-2) ‘오류 발생 시 책임 공백 해소 방안 부재’ 의제가 뒤를 이었다. 전문가들은 AI가 공공의사결정의 핵심 도구로 활용될 경우, 책임 귀속 체계의 부재를 가장 논의가 시급한 구조적인 문제로 평가하였다. 책임 체계가 불명확하면 공무원들의 소극 행정이나 AI가 결정했다는 식의 무책임한 행정, AI 기술 도입에 대한 부정적인 인식 확산 등으로 이어질 수 있음을 우려하였다. 특히 복지 급여 결정이나 인허가 심사 등과 같이 국민의 권익에 직접적인 영향을 미치는 영역에서 오류가 발생하고도 적절한 피해 구제 절차가 마련되어 있지 않다면, AI 행정에 대한 신뢰가 근본적으로 무너질 수 있음을 강조하였다. 다음으로는 (행정-3) ‘빅브라더 사회 귀결 위험’ 의제이다. 전문가들은 아직 빅브라더 관련 문제가 국내에서 사회적 실재로서 일상적이고 보편적으로 확인되고 있지 않지만, 국외에서 발생한 빅브라더 관련 문제가 국내에서도 발생할 수 있다는 우려가 상당하다는 점을 지적한다. 또한 국내에서 AI 기반의 안면인식 등에 대한 공공사업 및 정책이 활발하게 추진되고 있으나 이에 대한 평가나 합의 등의 과정은 충분하지 않으며, 특정기관이나 상위 행정주체와 같은 소수의 특정 집단이 다량의 데이터를 독점할 경우 정보의 비대칭성과 권력 불균형이 심화될 가능성이 있다고 보았다. 나아가 감시 기술이 통제되지 않을 경우, 사회구성원들의 자율성 제한으로도 이어질 수 있다고 보았다. 마지막으로 (행정-5) ‘AI 정책 결정에 대한 시민의 납득 가능한 설명 제공’ 의제가 선정되었다. XAI 기술(Explainable AI, 설명 가능한 AI)이 제공하는 기술적인 설명과 시민이 이해할 수 있는 실질적인 설명 간의 간극이 큰 것을 강조하였다. 현재 GDPR의 설명요구권, AI 기본법의 고지의무 등이 도입되고 있으나, 어느 수준까지가 충분한 설명인가에 대한 사회적 합의가 없는 상태이며, 민간 기업의 영업비밀 보호와 충돌하는

문제도 발생할 수 있음을 지적하였다.

〈표 4-10〉 공공행정 분야 의제 전문가 우선순위 선정 사유

의제번호	전문가 근거 요약
행정01	• 디지털 역량 격차가 공공서비스 접근권, 참여권 등 기본권 침해와 시민권 격차로 이어질 우려 존재 • 고령층·저소득층의 이용 장애가 심화되어 권리 포기 사례까지 발생 • 디지털 자산이 이용 가능한 세대나 계층과 그렇지 않은 계층의 격차가 점차 확대됨에 따라, 투입·재생산되는 디지털 데이터에도 양극화 발생 가능 • 알고리즘 편향은 제도적으로 보호해야 할 대상을 배제하게 되고 차별의 정도를 악화시킬 가능성이 있음
행정02	• 공공의사결정에 AI가 관여할 때 책임 귀속 체계가 부재하면, 공무원의 소극 행정·무책임 행정 발생 가능 • 특히 복지·인허가 등 권익 직결 영역에서 오류 구제 절차가 미비하면 AI 행정 신뢰가 붕괴될 우려가 있음
행정03	• 해외 사례처럼 국내에서도 감시·데이터 독점 위험이 발생할 가능성이 높음 • 안전인식 등 공공 AI 사업의 평가합의 절차가 미흡하고, 통제 부재 시 자율성 침해 우려 존재
행정04	• AI 활용 확대로 공공부문에서의 개인정보 활용 유인이 증가하여 개인정보에 대한 보호 및 안전성 제도 강화가 필수적 • 개인정보 침해 시 피해 규모가 전체 시민으로 확대되고 실제 범죄·경제적 피해 사례가 존재
행정05	• 설명 가능성은 AI 기반 정책 결정의 민주적 정당성 확보를 위한 핵심 조건이나, XAI 설명은 기술적 설명에 머물러 시민이 실제로 이해할 설명과 큰 간극이 존재함 • 아직 충분한 설명 기준에 대한 사회적 합의가 부족하며, 기업의 영업비밀과도 충돌 가능성이 있음

자료: 연구진 작성

전문가들은 우선순위로 선정한 의제들에 대해 다음과 같은 구체적인 정책과제를 제안하였다.

(행정 1)에 대해서는 첫째, 디지털 취약계층 개개인의 필요와 특성에 맞춤형된 교육·훈련 프로그램을 구성·운영하고, 해당 교육·훈련 프로그램의 제공 경로를 다양화한다. 둘째, 디지털 취약계층도 AI 서비스에 쉽게 접근하고 활용할 수 있도록

공공 AI 서비스 개발 시 접근성 가이드라인 준수, 음성 인터페이스, 쉬운 언어, 다국어 지원 기본 탑재 등을 의무화한다. 셋째, 취약계층 테스트 그룹을 운영하고 피드백 반영 절차를 마련한다. 구체적으로는 디지털 권익옹호(Digital Rights Advocate) 제도, 주민센터, 복지관 등에 디지털 취약계층 지원 전문 인력 배치, AI 서비스 이용 대행 및 권리구제 지원 역할 강화 등이 있다. 넷째, 다양한 정책 등에 파편적으로 규정되어 있는 디지털 취약계층 보호 관련 규정을 일원화하여, 디지털 취약계층에 대한 종합적이고 체계적인 보호 방안을 포함하는 법률 또는 지침을 제정한다.

(행정 2)에 대해서는 첫째, AI 시스템 생애주기(기획-개발-도입-운영-폐기)별 책임 주체를 명시하고, 개발사-알고리즘 결함, 데이터 제공자-학습데이터 품질, 도입기관-적용 적절성, 운영자-모니터링 의무 등의 내용을 담은 책임 분산 프레임워크를 법제화하고 가이드라인을 마련한다. 둘째, 고위험 AI 시스템(복지급여, 인허가 등) 도입 전 영향평가, 독립적 검증기관 지정 및 알고리즘 감사 체계 구축 등 사전검증 제도를 마련한다. 셋째, 입증책임 전환 규정(피해자가 AI 사용 사실만 입증하면, 무과실 입증은 행정기관 책임), 집단소송 허용 및 징벌적 손해배상 도입, AI 피해구제 전문 ombudsman 제도 운영 등 AI 결정으로 인한 피해 발생시 이를 구제할 수 있는 법제를 마련한다.

(행정 3)에 대해서는 제시한다. 첫째, 신규 첨단 감시 기술 도입 또는 기존 첨단 감시 기술 고도화 시 사전에 개인 권리에 미칠 영향을 평가하고 개인 권리 침해 위험의 최소화 방안을 마련한다. 둘째, 데이터 및 개인정보 보호를 위한 법률 및 지침에 데이터 주권(개인·기업·국가와 같은 정보 주체가 자신의 데이터·정보를 소유하고 어떤 목적·방식·장소·시기에 따라 사용할 것인지를 결정할 수 있는 권리)의 실질적인 구현을 위한 원칙과 규정을 강화한다. 셋째, 대규모 데이터 수집·처리 주체의 과도한 데이터 수집·처리를 감시할 수 있는 독립적인 기구를 설치한다. 특정 주체의 데이터 독점을 방지하기 위해 다양한 주체로 구성된 공공·지역 커뮤니티를 중심으로 데이터를 분산 관리하거나, 데이터의 분산을 위해 해당 기술 보유 주체 이외의 다양한 주체들이 중심이 되는 감시·평가 체계를 구축한다. 넷째,

개인 식별 데이터의 제한 또는 익명화를 위한 기술 및 데이터 보호를 위한 기술을 개발 또는 고도화한다.

(행정 4)에 대해서는 다음과 같다. 첫째, AI 과정에서 발생할 수 있는 위험을 진단하고 실제로 발생한 위험을 해결하기 위해 개인정보 처리의 적합성을 평가하여 적정 수준에서의 개인정보 수집 및 처리가 이루어질 수 있도록 하는 위험관리체계를 운영한다. 둘째, AI의 기술적 특성을 반영하는 방식으로 AI에 특화된 개인정보 관련 법률·지침을 제정하거나 개인정보 보호 관련 법률·지침을 수정·보완한다. 셋째, AI를 통해 처리되는 개인정보의 생태 구축 및 개인정보 처리과정 검증 등을 위한 별도의 전문기관 등을 설치한다. 넷째, 개인정보 비식별조치 관리·감독 강화를 통해 개인정보 안전성을 제고하고, AI 서비스에 필요한 데이터의 품질관리를 위한 기준을 마련한다.

(행정 5)에 대해서는 다음과 같다. 첫째, 시민이 납득할 수 있는 설명을 제공하기 위한 일관된 기준이나 체계를 마련한다. 예를 들어, 일반 시민(Level 1)에게는 결정 결과와 주요 고려 요소(예: ‘소득, 가구원 수가 주요 기준’)를 제공하고, 이의제기자(Level 2)에게는 개인 사례 적용 방식과 대안 시나리오를 제시하며, 소송단계(Level 3)에서는 알고리즘 로직 및 가중치를 공개하는 방식이다. 둘째, 자동화 결정 여부, 사용 AI 모델명, 주요 판단 기준, 이의제기 방법 명시, 비전문가도 이해 가능한 시각화 자료 첨부 의무화 등 AI 결정서 표준 양식을 개발한다. 셋째, 공공조달 계약 시 설명가능성 조항 필수 포함, 알고리즘 에스크로(Escrow) 제도 운영 등과 같이 민간 알고리즘 공공 도입 시 투명성을 확보할 수 있는 장치를 마련한다.

〈표 4-11〉 공공행정 분야 의제 전문가 정책과제 제안

의제번호	전문가 정책과제 제안
행정01	• 디지털 취약계층 교육·훈련: 취약계층 개개인의 필요와 특성에 맞춘 교육·훈련 프로그램을 구성·운영하고 제공 경로를 다양화 • 공공 AI 접근성: 디지털 취약계층도 쉽게 AI 서비스를 이용할 수 있도록 접근성 가이드라인 준수, 음성 인터페이스, 다국어 지원 등을 의무화 • 취약계층 피드백 반영: 디지털 취약계층 테스트 그룹을 운영하고 피드백을 반영하며 권리구제 체계 강화 • 규정 일원화: 디지털 취약계층 보호 관련 규정을 일원화하여 종합적·체계적인 보호 방안을 법률 또는 지침으로 마련
행정02	• AI 책임 분산: AI 시스템 생애주기별 책임 주체를 명시하고 개발사, 데이터 제공자, 운영자 등 역할을 법제화하여 책임 분산 프레임워크를 마련 • 고위험 AI 사전검증: 복지급여, 인허가 등 고위험 AI 시스템 도입 전 영향평가, 독립적 검증기관 지정, 알고리즘 감사 등 사전검증 제도를 마련 • AI 피해구제 법제: 피해자가 AI 사용 사실만 입증하면 행정기관이 무과실 입증하도록 규정하고, 집단소송·징벌적 손해배상·전문 ombudsman 제도를 통해 피해를 구제
행정03	• 감시 기술 권리 영향 평가: 신규 첨단 감시 기술 도입 시 개인 권리에 미치는 영향을 평가하고 권리 침해 위험 최소화 방안을 마련 • 데이터 주권 강화: 데이터 및 개인정보 보호를 위한 법률과 지침에 개인·기업·국가 등 정보 주체의 데이터 소유와 활용 권리를 실질적으로 구현 • 데이터 독점 방지: 대규모 데이터 수집·처리 주체의 과도한 데이터 활용을 감시하고 공공·지역 커뮤니티 중심 분산 관리 및 평가 체계를 구축
행정04	• 개인정보 보호 기술 개발: 개인 식별 데이터 제한·익명화 및 데이터 보호를 위한 기술을 개발·고도화 • 개인정보 위험관리 체계 구축: AI 과정에서 발생할 수 있는 위험을 진단하고 개인정보 처리 적합성을 평가하여 안전한 수집·처리 체계를 운영 • AI 특화 개인정보 법제화: AI 기술 특성을 반영한 개인정보 관련 법률·지침을 제정하거나 기존 법률·지침을 수정·보완 • 개인정보 전문기관 설립: AI 처리 개인정보 생태 구축과 처리과정 검증을 위해 별도의 전문기관을 설치 • 개인정보 안전성 관리 강화: 개인정보 비식별조치 관리·감독을 강화하고 AI 데이터 품질관리 기준을 마련
행정05	• 설명 기준 체계 구축: 시민이 납득할 수 있는 설명 제공을 위해 의사결정 결과, 주요 고려 요소, 알고리즘 로직 공개 등 단계별 체계 마련 • AI 결정서 표준화: 자동화 여부, 모델명, 판단 기준, 이의제기 방법, 이해 가능한 시각화 자료 첨부 등 표준 양식을 개발

의제번호	전문가 정책과제 제안
행정05	• 투명성 확보 장치 마련: 공공조달 계약 시 설명가능성 조항 필수 포함, 알고리즘 에스크로 제도 운영 등을 통해 투명성을 확보

자료: 연구진 작성

라. 의제별 핵심 보완 의견

전문가들은 연구진이 제시한 5개 의제의 방향성에는 대체로 동의하면서도, 다음과 같은 보완이 필요하다고 지적했다.

〈표 4-12〉 공공행정 분야 전문가 신규 제안 의제

의제 제목	제안 내용 및 배경
AI 기반시설 장애로 인한 사회기능 마비	• 제안 내용: “AI 기반시설 장애로 인한 사회기능 마비 문제를 어떻게 해결할 것인가?” • 제안 배경: AI 기반시설(예: 데이터센터) 장애는 막대한 경제적·사회적 피해를 초래할 뿐만 아니라, AI 기반 시스템·서비스의 안정성·지속가능성을 손상시킬 수 있기 때문에 이에 대한 대응책을 선제적으로 마련해야 한다.
AI 기반 행정의 민주적 통제	• 제안 내용: “AI 기반 행정의 민주적 통제를 위해 무엇을 고려해야 하는가?” • 제안 배경: AI가 공공 의사결정 구조 자체를 어떻게 변화시키는가에 대한 민주주의적 관점이 필요하다. 핵심 문제는 “누가 AI 도입을 결정하고, 어떤 방식으로 시민 의사가 반영되는가”이기 때문이다. 고위험 공공 AI 도입 전 시민 속의 절차 법제화, 시민·전문가공무원 혼합 구성의 알고리즘 위원회 제도, 의회 소속 독립 AI 감사기구 설립, 그리고 AI 의사결정의 사회적 의미를 이해하는 시민 참여 역량 강화 등이 필요하다.
공공부문 고용 불안 및 공무원 역량	• 제안 내용: “AI 도입에 따른 공공부문 고용 불안 및 공무원 역할 전환의 사회적 갈등 문제?” • 제안 배경: 내부 행정의 효율성, 효과성 향상은 AI 기술의 행정 도입에 따라 예상되는 가장 핵심적인 긍정적 효과 중 하나이다. 하지만 해당 현상에 따른 공무원의 직무 불안, 역할 축소, 업무 개편 등과 관련한 해당 공공부문 구성원들의 부정적 인식과 수용 의지와 관련한 내용은 심도 있게 다루어지지 않고 있다.
공공행정영역의 고영향 AI 대응	• 제안 내용: “공공기관은 AI기본법의 고영향 AI를 안전하고 신뢰할 수 있게 제공하고 있는가?” • 제안 배경: 2026년 1월 발효를 앞둔 「AI 기본법」에서 제시한 고영향 AI의 대부분이 공공행정 분야에 해당되어 공공행정기관이 이행해야 할 해당 책무를 명확하게 정의하고 실천방법을 제시할 필요가 있다.

자료: 연구진 작성

앞서 제시한 공공행정 영역의 다섯 가지 AI 사회적 의제 외에도, 전문가들은 AI 기반시설 장애로 인한 사회 기능 마비, AI 기반 행정의 민주적 통제 강화, 공공부문 고용 불안 및 공무원 역량 변화, 공공행정 영역의 고영향 AI 대응 체계 구축 등의 추가 의제를 제안하였다. 이들은 공공행정영역 AI의 사회적 이슈로 매우 중요하고 시급하게 대처해야할 주제들이며 이는 후속 연구에서 심층 분석과 면밀한 검토를 통해 각 의제의 대응 방안을 구체적으로 도출할 필요가 있다.

3. AI와 이용자: 의제 평가 및 주요 의견

가. 평가자 구성

3장에서 다루었던 ‘AI와 이용자’ 분야의 사회적 의제(안)의 타당성과 적절성 검증에 전문가 5인이 참여하였다. 참여 전문가는 특수교육과 교수, 특수학교 정보부장(교사), 재활학과 교수, 디지털 취약계층 교육 관련 정책사업 운영자, 관련 정책 연구자 등이다. 본 조사에서 전문가에게 제시한 이용자 분야 5개 의제(안)는 다음과 같다.

〈표 4-13〉 전문가 의견조사에 제시된 AI-이용자 영역 사회적 의제(안)

(이용자-1). [모든 국민의 권리로서의 AI 리터러시 교육 보장] 디지털 취약계층을 포함한 모든 시민이 AI 리터러시 교육에 평등하게 접근할 수 있도록 제도적 의무와 체계적 지원을 강화하는 데 정부는 정책적 지원과 자금을 얼마나 투입해야 하는가?	
제안의 배경	• 생성형 AI의 일상 침투로 AI가 보편 기술로 자리 잡아가고 있지만, 그에 걸맞는 시민의 AI 리터러시는 부족함. 그동안 정부의 AI 관련 교육 정책은 개발자 양성에 치중되었고, 일반시민, 특히 디지털 취약계층을 위한 체계적 AI 리터러시 교육 기회는 현저히 부족한 상태임
현황 분석	• OECD(2023) 조사 결과, 우리나라는 영국, 미국과 함께 AI 전문가 양성 교육에 정책 지원이 집중되는 국가 그룹에 속했으며, 일반 대중을 위한 AI 리터러시 교육은 매우 부족한 것으로 나타남
문제의 심각성	• AI 기초 교육 부재는 AI 기술의 혜택이 특정 계층에만 집중되는 결과로 이어져 AI 역량 격차를 심화하고 새로운 불평등을 가중시킬 수 있음 • 특히 AI 전환으로 인한 고용 악영향에 가장 취약한 저숙련 노동자 및 자동화 위험이 큰 직종의 종사자들이 AI 교육에 가장 접근하기 어려워 계층 간 경제적·사회적 격차가 더욱 고착화될 위험이 있음

(이용자-2) [AI 기반 서비스의 보편적 접근성을 위한 법·제도적 보장] 고령인·장애인의 AI 서비스 활용 권리와 자기결정권을 실질적으로 보장하기 위해, AI 기반 기술/서비스 개발 시 보편적 디자인 가이드라인 등을 법적으로 표준화하고 이를 관리·감독할 거버넌스를 어떻게 구축할 것인가?	
제안의 배경	• AI 기반 생활 편의 서비스가 확대될수록 디지털 접근성의 불평등이 새로운 사회적 배제 요인이 될 수 있음 • 또한 이들을 AI 기술의 ‘사용자’라기보다 관리 혹은 감시의 ‘대상자’로 보고 서비스되는 기술 도구가 있으며, 결과적으로, 안전·보호를 이유로 한 디지털 추적·관찰 기능의 대상이 되어 기본권 침해로 이어질 수 있음
현황 분석	• 기존 접근성 관련 법령은 AI 서비스의 특수성을 반영하지 못하고 있음. 공공기관 홈페이지 접근성 인증은 의무화되어 있으나 AI 기반 서비스에 대한 접근성 인증은 부재하며, 민간 서비스의 경우는 미약하거나 권고 수준에 그침. 또한 위반 시 제재 수단이 제한적이어서 실질적 개선 동력도 부족한 편임
문제의 심각성	• AI 서비스 접근성 부족은 고령인과 장애인의 정보 접근권, 사회 참여권, 자기 결정권을 침해할 수 있으며, 타인의 도움 없는 독립적 생활이 더 어려워질 수 있음 • AI 기술/서비스의 디자인이나 정책이 일반 국민에게만 맞춰지거나 기존의 사회적 차별을 알고리즘 영역으로 넘겨버리는 악용 등이 발생할 수 있음 • 접근성이 보장되지 않는 AI 서비스는 결국 대안 서비스 제공 비용, 격차 해소 비용 등 사회 전체의 비용을 증가시키게 됨
(이용자-3) [AI 활용 경험과 연계된 체험형 윤리교육 설계·확산] 디지털 취약계층이 AI에 대한 과도한 의존이나 막연한 두려움 없이 AI를 비판적으로 평가하고 윤리적 문제를 인지할 수 있도록, 활용과 연계된 체험형 윤리교육을 어떻게 설계하고 보급할 것인가?	
제안의 배경	• AI 기술이 키오스크, 모바일 앱, 공공 웹사이트 등 일상 서비스 전반에 확산되고 있지만, 경험 및 학습 기회가 상대적으로 부족한 고령인들은 막연한 두려움 때문에 더 경험 및 학습 기회를 잃는 악순환에 있음 • 이에 활용과 연계된 체험형 윤리교육 교재 개발 등 시범적 노력이 시도되고 있음
현황 분석	• 고령인들의 AI·디지털 교육 수요가 체계적으로 파악되고 공유된 경우가 없어서 사회 전반의 인식이 부족함 • 담당 부처 및 기관에서 시범적 노력이 있지만, 그 성과가 사회 전반으로 확산되어 활발하게 활용되기 위한 후속 조치가 부재함
문제의 심각성	• 과거 선택 가능했던 오프라인 서비스 비중이 빠르게 감소하면서 고령인들이 경험하는 신기술 기반 서비스에 대한 두려움이 더 커지고, 그러한 심리적 저항감으로 인한 기술 격차 역시 과거 어느 때보다 심각해질 수 있음
(이용자-4) [노동시장 재편에 대비한 취약계층 맞춤형 AI 활용 역량 재교육] AI 전환으로 고용 위협에 노출된 저숙련/장애인 노동자들을 대상으로, 생계형 일자리 유지 및 확산을 위한 맞춤형 AI 활용 교육 및 인간-AI 협업 기술 개발을 어떻게 지원할 것인가?	

제안의 배경	AI는 고용 시장에 변혁을 가져오고 있고, 기업들은 AI 기술 부족 해소를 위해 직원들의 업스킬링(upskilling)과 재교육(reskilling)을 요구하고 있음. AI 시대로의 사회 전환 과정에서 모든 국민이 AI 활용 교육 기회를 동등하게 누려야 하나, 특히 저숙련 노동자나 장애인들은 AI 대체율이 높은 일자리에 종사하는 와중에 변화하는 노동 환경에 적응하거나 독립적인 삶을 유지하기 위해 필요한 새로운 역량을 습득할 기회는 매우 부족함
현황 분석	- 저숙련 노동자나 자동화 위협이 큰 직종 종사자들은 AI 교육에 접근할 가능성이 낮음 - 발달 장애인들이 주로 종사하는 단순 노무 분야는 기계화 및 AI 기술 도입으로 인해 일자리가 박탈될 위험에 직면함
문제의 심각성	AI에 의한 일자리 대체는 취약계층의 생존권을 직접적으로 위협하며, 특히 발달장애인 등은 일자리를 잃을 경우 다른 고급 기술을 습득하기 어려워 결국 저숙련 시장으로 재진입하여 기존 노동자와 충돌하거나 사회적 고립을 심화시킬 수 있음
(이용자-5) [지역·계층 간 AI 리터러시 교육 격차 해소] 도시와 농촌, 소득 계층, 교육 수준에 따른 AI 리터러시 교육 기회의 격차를 해소하고, 디지털 취약계층이 거주지나 경제적 여건에 관계없이 양질의 AI 윤리 및 활용 교육에 접근할 수 있도록 하는 통합적 교육 인프라와 전달 체계를 어떻게 구축할 것인가?	
제안의 배경	- 기존 디지털·AI 교육은 대도시의 교육기관, 복지관, 도서관 등을 중심으로 제공되었고 농촌 및 소도시 거주자들의 접근은 상대적으로 어려움 - 최근 온라인 교육이 대안으로 제시되지만, 안정적인 인터넷 환경, 적절한 디지털 기기, 온라인 학습에 필요한 기초 역량 등이 전제되어야 한다고 할 때 취약계층의 실질적인 해결책이 되지 못할 수 있음
현황 분석	- 취약계층 디지털·AI교육의 대표 사업인 ‘디지털배움터’의 경우 도시 지역에 편중되어 있거나 온라인 과정 비중을 늘리고 있음. 교육 신청 과정도 대부분 온라인으로 진행됨 - 읍·면 단위 지역에서는 물리적 교육 인프라뿐만 아니라 교육을 기획, 운영할 인적 자원도 상대적으로 부족함
문제의 심각성	- 현재의 교육 격차가 그대로 지속된다면 도시와 농촌, 수도권과 비수도권 간 AI 리터러시 격차도 심각한 수준으로 벌어질 수 있음 - 농촌 지역의 고령 인구는 AI 기술 혜택에서 배제되어 의료, 금융, 행정 서비스 등 필수 서비스 이용에서 어려움을 겪을 것이고, 이는 지역 소멸 가속화 및 국토의 균형발전 저해로 이어질 수 있음 - 이들의 기술 배제는 디지털 기록정보 부족으로 인해 AI 정책 결정 과정에서도 배제되어 이들의 이익이 제대로 반영되지 않을 위험도 있음

자료: 연구진 작성

나. 정량 평가 결과

‘AI와 이용자’ 분야 전문가 5인이 각 의제(안)의 공공성, 타당성, 논의의 필요성,

사회적 파급성 등 4개 항목을 5점 척도로 평가하였다. 흥미롭게도 해당 의제(안)이 모든 국민에게 영향을 미치는 보편적 이슈인 경우 모든 항목에서 높게 평가된 반면, 특정 집단(예: 발달장애인, 지역)으로 한정되는 이슈인 경우는 상대적으로 낮은 평가를 받았다.

각 의제(안)의 종합 평균을 비교해 보면, 먼저 ‘(이용자-1) 모든 국민의 AI 리터러시 교육’ 의제가 가장 높게 평가되었다(4.85). 그다음으로 ‘(이용자-2) 보편적 접근성을 위한 법·제도’ 의제(4.40)와 ‘(이용자-3) 윤리교육 설계·확산’ 의제(4.30)가 뒤를 이었다. 「장애인차별금지법」, 「장애인고용촉진 및 직업재활법」 등에 따른 장애인 고용 현황이 AI 도입으로 급격히 무너질 것을 염두에 두고 도출된 의제 ‘(이용자-4) 취약계층 맞춤형 AI 활용 역량 재교육’(4.00)과 수도권에 쏠린 교육 인프라 등에 비판적인 의제 ‘(이용자-5) 지역·계층 간 격차 해소’(3.85)는 상대적으로 낮은 평가를 받았다. 이 두 의제는 특히 ‘사회적 파급성’ 항목에서 낮았다(각 3.60). 그다음으로 낮은 평가 항목은 공공성이었는데, 각각 4.00과 4.20의 평가를 받아서 해당 항목의 평균인 4.44보다 낮았다.

〈표 4-14〉 이용자 분야 의제 전문가 정량 평가 결과

구분	문제의 공공성	의제의 타당성	논의의 필요성	사회적 파급성	의제별 평균
(이용자-1) 모든 국민의 권리로서 AI 리터러시 교육 보장	5.00	4.80	4.80	4.80	4.85
(이용자-2) AI 기반 서비스의 보편적 접근성을 위한 법·제도적 보장	4.60	4.40	4.40	4.20	4.40
(이용자-3) AI 활용 경험과 연계된 체험형 윤리교육 설계·확산	4.40	4.20	4.40	4.40	4.30
(이용자-4) 이용자 시장 재편에 대비한 취약계층 맞춤형 AI 활용 역량 재교육	4.00	4.20	4.20	3.60	4.00
(이용자-5) 지역·계층 간 AI 리터러시 교육 격차 해소	4.20	3.80	3.80	3.60	3.85
평가항목별 평균	4.44	4.28	4.32	4.08	

자료: 연구진 작성

다. 정성 의견 분석

다음으로 제시된 5개 의제(안) 중 가장 중요하고 우선순위가 높아야 할 의제(안) 3개를 선택하고 그에 대한 의견 제시를 요청하였다. 먼저 의제(안)들의 우선순위 결과는 다음과 같다.

〈표 4-15〉 이용자 분야 의제 전문가 우선순위 평가 결과

구분	1순위 선정횟수	2순위 선정횟수	3순위 선정횟수	총 선정횟수
(이용자-1) 모든 국민의 권리로서 AI 리터러시 교육 보장	3	2	0	5
(이용자-2) AI 기반 서비스의 보편적 접근성을 위한 법·제도적 보장	2	2	0	4
(이용자-3) AI 활용 경험과 연계된 체험형 윤리 교육 설계·확산	0	1	2	3
(이용자-4) 이용자 시장 재편에 대비한 취약계층 맞춤형 AI 활용 역량 재교육	0	0	2	2
(이용자-5) 지역·계층 간 AI 리터러시 교육 격차 해소	0	0	1	1

자료: 연구진 작성

앞서 4개의 평가 항목 모두에서 가장 높은 평가를 받았던 ‘(이용자-1) 모든 국민의 AI 리터러시 교육’ 의제(안)는 1순위 횟수와 총 우선순위 횟수 모두에서 가장 많이 선정되었다. 그다음으로 2안과 3안이 뒤따랐는데, 이 결과는 앞서 정량 평가 결과와 같은 유형이다.

의제들의 우선순위를 선정한 이유를 살펴보면, 선정의 타당성을 알 수 있다. 먼저 1안과 관련하여, ‘모든 국민의 권리로서의 AI 리터러시 교육 보장’에는 취약계층도 포함되므로 우선순위로 선정하였음을 확인하였다. AI의 급속한 확산 과정에서 대부분의 성인이 AI와 AI를 현명하게 활용하는 역량인 AI 리터러시에 대한 교육을 받지 못했다는 상황도 주된 이유였다. AI 시대에 개인의 권리를 보호하고 민주적 참여 과정 등을 포함한 중요한 사회 활동을 위해 모든 국민이 체계적 교육을

받아야 한다는 것이다. 그리고 의제 3안, 4안, 5안의 문제는 하위 영역으로 포괄하여 함께 해소될 수 있을 것으로 기대하였다.

2안의 주요 선정 이유는 AI의 접근 보장으로 인해 얻는 정책 효과에 대한 기대였다. 신기술로 인해 발생할 수 있는 여러 사회 격차를 해소하고 그만큼 사회비용을 줄일 수 있다는 것이다. 신기술이 사회에 처음 등장하고 확산하는 초기 단계에서는 신기술 채택에 대한 사회적 저항감이 가장 크게 작동한다. 이 저항감이 강하고 오래 지속될수록 신기술로 인한 여러 사회 격차가 벌어지고 차별과 배제가 확산되기에, 신기술에 대한 사회적 수용성을 높이는 방안으로 보편적 접근 보장이 중요하다는 것이다. 또한 정책 대상이 일반 대중보다 범위가 좁은 고령인·장애인으로 한정될 경우, 정책 효과에 대한 기대가 낮게 인식될 수 있다. 그러나 기준을 취약계층에 맞추어 설계할수록 오히려 일반 대중 역시 추가적인 학습이나 비용 부담 없이 해당 혜택에 자연스럽게 접근할 수 있으며, 그 효과는 사회 전반으로 확산된다는 점도 간과할 수 없다.

3안 선정의 주된 이유는 이론 교육보다 현장과 밀접한 실천 교육의 효과였으며, 또 그러한 교육이 ‘책임 있는 활용’으로 이어질 것이라는 기대감이었다. AI 생성 콘텐츠를 무비판적으로 받아들였을 때 사회에 쌓이는 불신, 정보 왜곡 등의 부작용과 AI의 사회적, 윤리적 영향을 이해하는 역량을 각각 고려했을 때, 체험형 실천 교육으로 모두가 어느 정도의 이해와 역량 수준을 갖춰야 한다는 것이다.

그 밖에 4안은 사회 노동인구인 장애인이 노동시장에서 배제되지 않도록 통제해야 한다는 필요에 대해, 5안은 보편적 접근의 중요성만큼 개인의 특수성을 고려한 접근도 동시에 고려해야 한다는 것이 언급되어 있다.

〈표 4-16〉 이용자 분야 의제 전문가 우선순위 선정 사유

의제번호	전문가 근거 요약
이용자01	• AI 리터러시 교육은 단편적 기술교육의 차원을 넘어 개인의 권리를 보호하고 민주적 참여 과정에도 적용되므로 모든 국민에게 체계적 교육을 받을 권리가 보장되어야 함 • 교육현장에서 AI를 활용하는 학습자들마다 AI 접근성 차이가 존재하므로 형평성 보장을 위한 보편적 AI 교육 체계 마련 시급 • 보편적 AI 서비스 접근성이 일정 수준으로 보장되면 전 계층 대상 AI 리터러시 교육권 인식과 실제적 교육의 실현 가능
이용자02	• 디지털 취약계층을 포함할 수 있는 실질적이고 보편적인 AI 리터러시 교육 체계의 마련은(이용자-3), (이용자-4), (이용자-5)의 문제를 하위 영역으로 포괄하여 함께 해소될 것으로 기대
이용자03	• AI 대전환이 급속도로 전개되는 상황에서 포용성을 보장하는 표준이 초기에 도입되지 못하면 계층간 격차 발생, 차별과 배제의 확산이 가속될 수 있음 • AI의 보편적 사용을 시민의 권리로 인정하는 제도적 마련은 다른 의제를 실현할 수 있는 전제 조건 • AI 기술이 다양한 영역에 영향을 미치는 만큼 특정 대상에 대한 불이익을 재생산할 가능성이 있으므로 단순 권고 수준을 넘어서는 법적·제도적 의무 마련을 통해 보편적 접근성을 보장할 필요가 있음 • AI 기반 기술 서비스는 개인의 선택이나 환경적 요인으로 AI 사각지대를 형성할 가능성이 있으므로 법과 제도 정비를 통해 AI 사각지대를 방지하고 해소할 정책을 마련할 필요가 있음
이용자04	• AI의 사회적·윤리적 영향을 이해하고 책임 있게 활용하는 실천적 역량이 필수적 • AI가 생성한 정보를 무비판적으로 수용하게 되면 기술 불신, 알고리즘 편향, 정보 왜곡 등의 사회적 부작용 초래 • AI 시대에 초래될 수 있는 윤리적 위험을 인식하고 이를 대처하기 위해서는 이론적 교육보다 현장과 밀접한 실천적 교육이 필요함
이용자05	• AI 취약계층의 노동소외는 소득 불평등과 사회참여 기회의 감소로 이어질 가능성이 존재 • 취약계층의 노동시장 재진입을 지원하는 사회안전망 구축이 필수적, 특히 저숙련 노동자, 장애인 등 취약계층은 상이한 학습장벽을 가지고 있으므로 노동시장 재편에 대비한 대상별 맞춤형 AI 활용 교육 필요
이용자06	• (이용자-1)이 보편적 접근이라면 (이용자-5)는 특수한 접근으로서 보편-특수가 함께 이루어져야 함 • 다층지원체계(MTSS)를 형성하여 (이용자-1)을 바탕으로 (이용자-5)를 이름으로써 서비스의 양적 측면과 질적 측면 모두의 향상을 기대

자료: 연구진 작성

이어 전문가들이 각자 우선순위로 선정한 의제(안)에 대해 제안한 정책 과제는 다음과 같다. 먼저 1안과 관련하여, 취약계층이 활용할 수 있는 보조공학기기, 무료 교육 플랫폼 등을 개발, 제공하는 정책 사업, 기존 교육 인프라 확장, 현장 교원·강사의 관련 역량 강화 지원 등의 제안이 있었다. 흥미롭게도 유치부, 초등부, 중고등 교육 등 아동·청소년의 AI 리터러시 교육에 대한 제안도 높은 비중을 차지하였다. 이는 제시된 의제(안)이 성인, 그중에 전통적 디지털 취약계층인 고령인과 장애인에 집중되었던 것을 염려한 결과라고 짐작된다.

2안에 대해서는 사회 기초·기본 서비스에 대한 보편적 접근성을 의무로 하는 현행 법령에 AI 서비스 관련 내용도 구체적으로 의무화하는 법 개정, 장애단체 등 이해당사자가 참여하는 법제도 개발 절차, 표준화 및 가이드라인 개발 및 주기적 갱신, 디지털 포용법 실효성 제고, 사회적 합의 형성을 위한 조치 및 노력 등이 제안되었다.

3안과 관련한 정책과제에는 다양한 장애 유형과 정도, 소득 수준 등을 모두 고려한 교육 콘텐츠 및 보조공학기기 개발·보급 사업이 제안되었다. 또 공공·민관 협력 기반의 교육 확산 모델, 교육기관의 지속가능성을 담보하는 평가·인증 체계, AI 윤리교육 표준과정 개발, 취약계층을 위한 생활밀착시설 소규모 실습교육 등이 제안되었다. 4안에서는 취약계층의 더욱 불리한 재취업 기회 및 상황을 반영한 기존 재교육 정책 및 고용센터/평생학습관 운영 등에서의 확장안이, 5안에서는 관련 전담 조직 구성 및 교육자료·방법 연구의 필요성이 제안되었다.

〈표 4-17〉 이용자 분야 의제 전문가 정책과제 제안

의제번호	전문가 정책과제 제안
이용자01	• 공교육 정규 과정에 AI 리터러시 교육 관련 사항 편입: 초·중·고등학교에 AI 리터러시 교육 의무적 실시하여 관련 역량 강화 • 정보 취약계층 우선적 접근성 마련: 취약계층이 AI 리터러시 교육에 소외되지 않도록 보조공학기기 및 무료 교육 플랫폼 제공 • 현장 교원 AI 관련 역량 강화: 교육부, 교육청, 교육지원청 차원의 체계적이고 지속적인 교원 역량 강화 프로그램 개발 및 운영, 현장의 요구에 맞는 다양한 연수 기회 제공 및 우수사례 공유

의제번호	전문가 정책과제 제안
이용자01	• 디지털 포용법 개정: AI 리터러시 교육을 국민의 기본권으로 명시하고 국가 및 지방자치단체의 교육 제공 의무 법제화 • 교육 인프라 확장: 'AI 디지털 배움터'를 AI 리터러시 교육 거점(허브)으로 전환·고도화 • 유치부·초등부 디지털 시민성 교육: AI 시대에 자기주도권을 가진 AI 활용법 조기교육 필요 • AI 리터러시 교육 전담 조직 지정 또는 신설: 학령기 의무교육과정은 물론 평생 교육 차원에서 열린 교육기회를 마련하고 기관 수준 프로그램 개발·인프라 마련
이용자02	• 법률상 접근성 의무 확대: 기존 장애인 접근성, 정보통신 접근성 의무를 AI 서비스로 확장 • 장애인체 등 이해당사자가 규제·검증: 현실에 맞는 법과 제도 마련을 위해 관련 당사자들의 요구사항을 반영할 수 있는 노력이 필수적 • 표준화 및 가이드라인 개발·주기적 갱신: 관련 법과 제도를 현장의 사용자들에게 안내할 수 있는 노력과 주기적인 갱신 필요 • 재정·기술 지원 관련 전문인력 양성 필요: 법 또는 제도적으로 지원방안 마련 및 지속적 전문 인력 양성 필요 • 디지털 포용법 실효성 제고: 제도적 장치 및 기술표준 개발·보급, 기술/서비스 공급자·개발자 인식 개선, 포용적 기술 개발 역량 강화 • 법·제도 기반 구축 및 장기적 사회적 합의 형성: 디지털 격차 확대 및 디지털 장애 집단 형성을 막기 위한 AI 리터러시 교육 및 학술적 논의 필요
이용자03	• 포용적 맞춤형 교육 설계 시스템 구축: 다양한 장애 유형·정도 및 소득수준에 따른 학습 접근성을 보장하는 콘텐츠 및 보조공학기기 개발·보급 • 공공-민관 협력 기반 지역 확산 모델 구축: AI 윤리 체험 교실 등 공공기관 중심 정기적 프로그램 편성 및 운영, 민간 차원 후원 등의 참여 유도 • 지속가능한 제도화 및 평가·인증 체계 마련: 국가 차원 정책 사항 반영, 시설 및 교육기관 연계 AI 윤리교육 강사 양성 과정 운영 • AI 윤리교육 표준과정 개발: 과기정통부·교육부 연계 AI 윤리교육 국가표준안 개발·배포, 현장 연계 체험형 교육 커리큘럼 개발 • 취약계층 대상 AI 윤리교육 제공: 생활밀착시설 소규모 실습교육 실시, 민간기업 사회공헌 연계·협력 체계 구축
이용자04	• 맞춤형 AI 재교육 체계 구축: 대상별 특성을 고려한 맞춤형 AI 활용 교육과정 설계, 고용노동부 주관 지역 고용센터/평생학습관 등 인프라 연계·추진 거버넌스 구축 • 고비용 AI 활용 자본 이윤 추구 극대화 예방책 마련: 대규모 실업 또는 노동 대체 발생 시 보상적 대안 모색, AI-인간 협업 기술 우선지원 등 적정수준의 사회적 합의 도출 필요
이용자05	• 특수계층 접근을 위한 다층 지원 체계 작동: 기술 접근성이 낮은 집단을 위한 별도 대책을 동시 마련 필요, 취약계층 전담 조직 구성 및 교육자료·방법 연구 필요

자료: 연구진 작성

라. 의제별 핵심 보완 의견

마지막으로 전문가들이 새롭게 제안한 의제들을 제시한다. 첫째, 현시점에서 이미 뉴스에 보도되는 등 사회문제가 된 위험으로부터 AI 취약계층을 실질적으로 보호하기 위한 사회제도적 환경과 역량을 어떻게 갖출 것인지에 대해 논의하자고 제안하였다. 딥페이크를 활용한 피싱 범죄, 개인정보 탈취, 로맨스스캠 등 실질적 피해의 대상자는 대개 AI 취약계층인데, 이들의 자체 방어 역량 강화와 함께 사회 시스템을 구축하는 방안을 모색해야 한다는 것이다.

그 밖에 제안된 의제안들은 <표 4-18>에서 살펴볼 수 있다.

<표 4-18> 이용자 분야 전문가 신규 제안 의제

의제 제목	제안 내용 및 배경
AI 취약계층 보호	• 제안 내용: “AI 기술 확대에서 기인하는 다양한 위험성으로부터 AI 취약 계층을 보호할 수 있는 환경과 역량을 어떻게 강화할 것인가?” • 제안 배경: AI 리터러시가 부족한 취약계층이 경험하는 위험요인(예, 딥 페이크를 활용한 피싱 범죄, 개인정보 탈취, 로맨스 스캠 등), 비의도적 불법 행위 노출(지적재산권 침해, 명의도용 등) 등으로부터 보호하기 위한 사회적 시스템을 어떻게 구현하고, 개인의 방어 역량을 강화하기 위해 어떤 정책 지원이 필요한지 논의할 필요가 있다.
AI 알고리즘 차별과 권리보장	• 제안 내용: “AI 알고리즘 차별로부터 취약계층을 보호하기 위한 권리보장 체계를 어떻게 구축할 것인가?” • 제안 배경: 기존 5개 의제는 주로 교육을 통한 개인의 AI 리터러시 역량 강화에 초점을 두고 있어서 AI 리터러시 격차를 완전히 해소할 수 없기에, AI 리터러시가 낮은 취약계층이 AI 시스템으로부터 차별받지 않도록 보호하는 제도적 안전장치 마련이 필요하다.
AI 활용 의사결정 공정성과 투명성	• 제안 내용: “AI 활용 과정에서 의사결정의 공정성과 투명성을 어떻게 확보할 것인가?” • 제안 배경: ‘알고리즘 책무성(Algorithmic Accountability)’과 관련하여 AI가 사회복지, 채용, 교육, 신용평가 등에서 점점 더 결정 분야와 권한을 다양하게 가지게 되면서, 그 결정 과정의 불투명성(black box)이 심각한 문제가 된다. 사람들은 자신의 데이터가 어떤 판단에 쓰였는지 알 권리(right to explanation)를 보장받을 수 있어야 한다.

의제 제목	제안 내용 및 배경
AI 개발 시민참여 및 사회적 거버넌스	- 제안 내용: “AI 개발·운영 과정에서의 적극적인 시민 참여 및 사회적 합의 거버넌스가 가능할 것인가?” - 제안 배경: 그동안 AI 정책과 기술 표준은 국가 및 기업의 전문가 중심으로 형성되어 왔다. 소수의 전문가가 다양한 특성과 요구를 지닌 모든 사용자를 대변하기에는 한계가 있을 것이다. 앞으로는 여러 사용자가 직접 AI 거버넌스에 참여하고, 사회적 가치(공정·안전·책임성)를 논의할 수 있는 구조가 필요하다.

자료: 연구진 작성

4. AI와 정보: 의제 평가 및 주요 의견

가. 평가자 구성

‘AI와 정보’ 분야는 경제학, 법학 등 관련 분야 전문가 5인이 조사에 참여하였다. 본 조사에서 전문가에게 제시한 정보 분야 5개 의제는 다음과 같다.

〈표 4-19〉 전문가 의견조사에 제시된 AI-정보 영역 사회적 의제(안)

(정보-1) 정보 신뢰 붕괴	
제안의 배경	- 생성형 AI로 인해 누구나 텍스트·이미지·영상 콘텐츠를 손쉽게 생성할 수 있게 됨 - 기술의 발전으로 AI가 생성한 콘텐츠를 실재와 구분하기가 점점 더 어려워지고 있으며, 이는 진실과 허위의 경계가 무너뜨리고, 정보 신뢰 체계를 붕괴시키는 결과를 초래할 수 있음 - 특히 선거, 재난, 보건 등 중요한 사회적 의사결정 과정에서 조작된 정보가 유포될 경우, 그 피해는 개인을 넘어 사회 전체로 확산될 수 있음
현황 분석	- 딥페이크 영상과 AI 음성합성을 활용한 금융사기·보이스피싱, 개인의 얼굴·음성 등을 이용한 디지털 성범죄, 명예 훼손 등의 기술의 악의적 활용 사례가 다수 보고 - AI 기반 허위조작정보가 대규모 여론조작, 정치적 선전, 선거 개입 등으로 이어지며 민주적 의사결정의 공정성을 훼손할 우려 - 허위조작정보 확산으로 사회적 신뢰 붕괴 및 공론장 왜곡 현상 심화
문제의 심각성	- 악의적 행위자는 손쉽게 대량의 허위조작정보를 생성·유포할 수 있으나, 이를 검증하고 대응하는 데는 막대한 시간과 자원 필요 - 한편 유포된 허위조작정보는 삭제나 정정이 어렵고, 피해자의 회복이 사실상 불가능한 경우가 많음

문제의 심각성	• 디지털 리터러시가 낮은 계층(노인, 아동 등)이 허위조작정보에 더욱 쉽게 노출될 가능성 • 정치적 양극화·사회 분열 심화를 조장할 수 있음
(정보-2) 허위 과장 광고 및 소비자 기만	
제안의 배경	• AI로 생성된 가짜 리뷰, 조작된 제품 이미지, 과장된 효과 영상 등이 온라인 쇼핑몰과 소셜미디어에 범람 • 특히 건강기능식품, 화장품, 의료기기 등에서 AI를 활용한 허위·과장 광고 급증
현황 분석	• AI 가상 인물이 실제 사용 경험을 한 것처럼 연출하거나, 전문가·의료인 이미지를 무단 활용하여 신뢰를 조작하는 사례 증가 • 소비자 데이터 분석을 통해 건강·외모 콤플렉스 등 개인의 심리적 취약점을 겨냥한 맞춤형 감정 조작 광고 확산 • 진위 검증이 어려운 광고 콘텐츠가 시장 내 공정경쟁 질서를 훼손하고, 정직한 사업자의 피해를 초래
문제의 심각성	• 소비자의 합리적 의사결정 과정을 왜곡하고, 공정한 시장 경쟁을 저해 • 장기적으로 소비자 신뢰 기반이 붕괴되어 디지털 상거래 전반의 지속가능성을 위협
(정보-3) AI 알고리즘 불투명성 및 정보 편식	
제안의 배경	• 포털사이트의 뉴스 배열, 소셜미디어의 콘텐츠 추천, 검색엔진의 결과 순위 등은 모두 AI 알고리즘에 의해 결정되고 있음 • 그 결과, 특정 정보는 과도하게 노출되고 중요한 정보는 묻히며, 이용자는 알고리즘이 선별한 정보만 접하게 되는 정보 편식 현상이 심화 • 이러한 알고리즘의 작동 방식은 블랙박스로 남아 있어, 편향되거나 조작되더라도 이를 확인하거나 시정할 수 있는 제도적 장치 부재
현황 분석	• AI 알고리즘이 클릭률·체류시간 등 자극적 지표를 중심으로 콘텐츠 노출을 결정하면서, 선정적·감정적 콘텐츠가 과잉 노출되고 저널리즘의 공익성 훼손 • 자신의 기존 신념과 일치하는 정보만 반복적으로 소비하게 되어 유사한 성향의 이용자들이 폐쇄적 공간에서 상호 확증하는 ‘집단 극단화’ 현상이 나타날 우려 증대 • 특히 청소년과 아동은 형성기부터 알고리즘이 큐레이션한 정보에 노출되어, 사회를 인식하고 사고하는 방식이 알고리즘에 의해 형성될 위험 • 주류·인기 콘텐츠만 추천되어 소수 의견과 비주류 정보는 배제되고, 사회적 대화와 타협의 공간이 축소됨 • 결국 플랫폼은 사회적 의제 설정권을 사실상 독점하게 되었으나, 이에 상응하는 책임과 규제 부재
문제의 심각성	• 정보 접근과 여론 형성이 소수 플랫폼의 불투명한 알고리즘에 의해 통제되는 구조는, 민주주의의 근간인 ‘자유로운 정보 유통’과 ‘공론장 형성’을 심각하게 위협

문제의 심각성	• 이용자의 인식·의제 설정·가치 판단이 AI가 설계한 틀 안에서만 이루어지는 상황은 인간의 인지적 자율성과 비판적 사고 능력 약화 • 장기적으로는 사회 전체가 알고리즘이 조정한 인지 구조 속에서 작동하게 되어, 정보생태계의 다양성과 인간 주체성이 근본적으로 침식될 우려
(정보-4) AI 저작권 침해 및 창작 활동 위축	
제안의 배경	• 생성형 AI가 기존 예술작품·디자인·사진 등을 무단으로 학습해 유사하거나 동일한 창작물을 재생산함에 따라, 원저작자의 경제적 손실 및 명예 훼손 발생 • 현행 저작권법 체계는 인간의 창작 행위를 전제로 설계되어 있어, AI 기반 창작물의 권리 보호와 책임 부과에 대응하기 어려운 구조적 한계
현황 분석	• 인간 창작의 노력과 독창성의 가치가 평가절하되고, 창작 활동이 ‘무임승차’ 가능한 영역으로 인식되는 경향 확산 • 특히 예술가들은 AI가 만든 이미지나 음악이 본인 작품과 구분되지 않거나 시장에서 더 높은 주목을 받는 상황에서 직업적 정체성과 생계의 위협을 동시에 경험하고 있음 • AI 생성물은 상업적 트렌드를 모방하며 반복적으로 생산되기 때문에, 실험적이거나 비주류 예술 양식은 배제되고 문화의 다양성 축소
문제의 심각성	• AI 시대에 기존 저작권 체계가 사실상 작동하지 않으면서, “창작에 대한 보호와 보상”이라는 저작권의 본래 목적이 무력화되고 있음 • 창작이 정당한 경제적 보상을 받지 못하는 활동으로 전락하면 예술가와 창작자의 생존 자체가 위협받으며, 이는 단순한 직업 위기 이상의 문화 생산 시스템의 붕괴로 이어질 수 있음 • 창작 생태계가 붕괴되면 새로운 문화 콘텐츠의 원천이 고갈되고, AI는 과거 데이터를 반복 재생산하는 ‘자기복제적 기술’로 전락함으로써 문화의 정체와 퇴보를 초래할 수 있음
(정보-5) AI 활용 역량 및 정보 격차	
제안의 배경	• AI 기술이 빠르게 확산되면서 정보의 생산·유통·활용 방식이 근본적으로 변화하고 있는 반면, AI를 활용할 수 있는 역량의 차이는 개인 간, 세대 간, 지역 간 새로운 디지털 격차로 확대되고 있음 • AI의 원리·한계·위험을 이해하고 비판적으로 사용할 수 있는 AI 리터러시의 중요성이 강조되고 있으며, AI 활용 능력의 불균형은 단순한 기술 접근성의 차원을 넘어, 정보 이해력·비판적 사고·사회참여 기회의 불평등으로 이어지고 있음
현황 분석	• AI 서비스 이용 경험은 주로 고학력·고소득층, 젊은 세대에 집중되어 있으며, 고령층·저소득층·농어촌 지역에서는 접근과 활용이 제한적 • AI 도구를 효과적으로 활용하지 못하는 이용자는 정보 검색·의사결정·직업 활동에서 불리한 위치에 놓이며, 일상적 기회 격차가 확대되고 있음 • AI 활용 능력이 낮을수록 AI가 제공하는 정보를 맹신하거나, AI가 생성한 허위정보를 사실로 오인할 위험이 높음

문제의 심각성	• AI 활용 격차는 단순한 ‘기술 숙련도’의 차이가 아니라, 사회적 기회·경제적 보상·정치적 영향력의 차이로 확산되는 복합적 불평등 구조로 발전하고 있음 • 정보 접근과 해석 능력의 격차는 사회적 의제 형성 과정에서 특정 집단의 목소리를 배제시켜 디지털 민주주의의 기반을 약화 • 이러한 불균형이 지속될 경우, AI는 사회통합을 촉진하는 도구가 아니라 오히려 계층 간 단절을 고착화하는 요인으로 작용할 위험 존재

자료: 연구진 작성

나. 정량 평가 결과

‘AI와 정보’ 분야 전문가 5인을 대상으로 5개 의제(안)에 대해 4개 항목(문제의 공공성, 의제의 타당성, 논의의 필요성, 사회적 파급성)을 5점 척도로 평가하도록 하였다. 정보 분야 의제 전문가 정량 평가 결과, 도출된 5개 의제(안)는 4개 평가 항목(문제의 공공성, 의제의 타당성, 논의의 필요성, 사회적 파급성)에서 모두 ‘높음(4.0점)’ 이상의 높은 평균 점수를 획득하여, 본 연구진이 도출한 의제(안)들이 AI 시대 정보 분야의 핵심 쟁점을 적절히 반영하고 있음을 확인하였다.

의제별 종합평균을 기준으로 보면 ‘(정보-1) 정보 신뢰 붕괴’ 의제가 5인 평균 4.70점으로 가장 높게 평가되었다. 이 의제는 특히 ‘문제의 공공성’과 ‘사회적 파급성’ 항목에서 각각 4.80점으로 최고점을 기록하였으며, 이는 정보 신뢰 문제가 가장 시급하고 중요한 사회적 과제로 부상하고 있음에 전문가들의 높은 공감대가 형성된 결과로 볼 수 있다. ‘(정보-3) 알고리즘 불투명성 및 정보 편식’ 의제는 4.45점, ‘(정보-5) AI 활용 역량 및 정보 격차’ 의제는 4.25점, ‘(정보-2) 허위 과장 광고 및 소비자 기만’ 의제는 4.15점의 높은 평균 점수를 받았다. 한편, ‘(정보-4) 저작권 침해 및 창작 활동 위축’ 의제는 평균 3.75점으로 5개 의제 중 가장 낮게 평가되었으며, 특히 ‘문제의 공공성’과 ‘의제의 타당성’ 항목에서 3.60점으로 최저점을 기록하였다.

항목별로는 ‘(정보-1) 정보 신뢰 붕괴’와 ‘(정보-5) AI 활용 역량 및 정보 격차’ 의제가 ‘문제의 공공성’ 측면에서 4.80점으로 가장 높은 공감대를 얻었다. 평가 항목별 평균 점수는 ‘문제의 공공성’이 4.40점으로 가장 높았으며, ‘의제의 타당성’이 4.04점으로 가장 낮았다. 특히 논의의 필요성 항목의 전체 평균은 4.36점

으로, 4개 평가 항목 중 두 번째로 높은 평균 점수를 기록하였다. 이는 전문가들이 정보 분야의 5개 의제 모두에 대해 시급한 사회적 논의가 필요하다는 데 폭넓게 공감하고 있음을 시사한다.

〈표 4-20〉 정보 분야 의제 전문가 정량 평가 결과

구분	문제의 공공성	의제의 타당성	논의의 필요성	사회적 파급성	의제별 평균
(정보-1) 정보 신뢰 붕괴	4.80	4.60	4.60	4.80	4.70
(정보-2) 허위 과장 광고 및 소비자 기만	4.20	4.00	4.20	4.20	4.15
(정보-3) 알고리즘 불투명성 및 정보 편식	4.60	4.20	4.60	4.40	4.45
(정보-4) 저작권 침해 및 창작 활동 위축	3.60	3.60	4.00	3.80	3.75
(정보-5) AI 활용 역량 및 정보 격차	4.80	3.80	4.40	4.00	4.25
평가항목별 평균	4.40	4.04	4.36	4.24	

자료: 연구진 작성

다. 정성 의견 분석

5개 의제 중 중요도 및 시급성을 기준으로 우선순위가 높아야 할 의제 3개를 선정한 결과, ‘(정보-1) 정보 신뢰 붕괴’ 의제가 5인의 전문가 전원에게 선택되어 총 5회 선정되었다. 특히 이 의제는 1순위로 2회, 2순위로 3회 지목되어 가장 높은 시급성과 공감대를 형성하였다. 그 뒤를 이어 ‘(정보-3) 알고리즘 불투명성 및 정보 편식’ 의제가 1순위 1회, 2순위 1회, 3순위 2회 등 총 4회 선택되어 높은 우선순위를 보였다. ‘(정보-5) AI 활용 역량 및 정보 격차’ 의제는 총 3회 선택되었으며, ‘(정보-4) 저작권 침해 및 창작 활동 위축’ 의제는 총 2회, ‘(정보-2) 허위 과장 광고 및 소비자 기만’ 의제는 총 1회 선택되는 데 그쳐 타 의제 대비 우선순위에서 밀리는 것으로 나타났다.

〈표 4-21〉 정보 분야 의제 전문가 우선순위 평가 결과

구분	1순위 선정횟수	2순위 선정횟수	3순위 선정횟수	총 선정횟수
(정보-1) 정보 신뢰 붕괴	2	3	0	5
(정보-2) 허위 과장 광고 및 소비자 기만	1	0	0	1
(정보-3) 알고리즘 불투명성 및 정보 편식	1	1	2	4
(정보-4) 저작권 침해 및 창작 활동 위축	0	0	2	2
(정보-5) AI 활용 역량 및 정보 격차	1	1	1	3

자료: 연구진 작성

각 의제를 선정한 전문가들의 근거 요약은 다음과 같이 정리하였다.

‘(정보-1) 정보 신뢰 붕괴’ 의제는 민주적 절차의 기본 구조이자 다른 의제들의 기반 구조를 훼손하는 복합적 위기로 규정되었다. 정보의 진위 판별 과정에서 누락되는 시민의 인지적 피로는 비판적 사고 포기와 냉소적 무관심으로 이어져 사회·제도 신뢰 약화 및 공론장 붕괴를 초래할 수 있다는 점이 핵심 선정 사유로 강조되었다. 이는 경제·사회 전반의 신뢰를 흔들고 사회 전체의 비용으로 이어지므로 국가적 위기 관리 차원의 핵심 의제로 다루어져야 한다고 결론 내렸다.

‘(정보-3) 알고리즘 불투명성 및 정보 편식’ 의제는 불투명한 알고리즘이 개인의 선택을 가장한 필터버블·에코챔버 현상을 강화하여 정보 접근 편식과 집단 극단화를 발생시키고, 이는 공론장 기능 약화 및 민주주의 제도에 대한 신뢰 붕괴로 이어질 구조적 위험성을 내포하고 있다고 진단하였다. 특히 공론장의 게이트키퍼 역할을 하는 AI 알고리즘이 정보 흐름을 보이지 않는 손으로 지배하고 있으며, 기술 발전에 따라 투명성 확보가 어려워짐에 따라 AI 공공성·책임성 확보를 위한 감독체계 구축 등 대응 과제 도출이 시급함을 강조하였다.

‘(정보-5) AI 활용 역량 및 정보 격차’ 의제는 다른 네 가지 의제의 선행 조건으로서 중요하며, AI 포용 정책의 핵심이라고 평가되었다. AI 리터러시 및 디지털 접근성 강화는 탈진실 대응, 소비자 기만 방지, 알고리즘 편향 완화 등에 기여하는 기반 역량으로 필요하며, AI 활용 능력의 격차는 노동시장 양극화 및 정보 권

력의 비대칭과 같은 사회 불평등으로 이어질 우려가 크다는 점이 선정 사유로 제시되었다.

‘(정보-4) 저작권 침해 및 창작 활동 위축’ 의제는 AI 모델 개발 과정에서의 저작물 이용과 창작자 권리 사이의 충돌로 인해 제도적 해결이 어려우며, AI 효율성에 대한 인식 확대로 예술적 창작 활동의 동기 자체가 위축될 수 있다고 보았다. 글로벌 규범과의 정합성을 고려한 정당한 보상 체계 구축 및 사회적 대타협의 필요성이 강조되었다.

‘(정보-2) 허위 과장 광고 및 소비자 기만’ 의제는 소비자 기만이 존재하나 관련 제도와 규제에 대한 논의가 부족하며, 다른 의제들과 연결되는 동시에 사회에 미치는 구체적 영향을 평가할 논의가 필요하다는 점이 지적되었다.

〈표 4-22〉 정보 분야 의제 전문가 우선순위 선정 사유

의제번호	전문가 근거 요약
정보01	• 다른 의제들의 원인이자 기반구조로, 정보 신뢰는 사회적 합의, 정책 실행, 민주적 절차의 기본 구조 • 정보 신뢰 붕괴는 민주주의의 절차적 정당성의 훼손 가능성으로 이어짐 • 정보의 진위 판별 과정에서 누적되는 인지적 피로는 시민의 비판적 사고 포기 및 냉소적 무관심으로 이어지며, 이는 사회·제도 신뢰 약화와 공론장 붕괴로 이어짐 • 정보 신뢰 붕괴는 공공성, 사회적 파급, 경제적 비용 등이 결합된 복합적 위기이며, 경제/사회 전반의 신뢰를 흔들고 사회 전체의 비용으로 이어지므로 국가적 위기 관리 차원의 핵심 의제
정보02	• 소비자 기만이 존재한다고 판단되나, 관련 제도와 규제에 대한 논의 부족 • 다른 의제들과 연결되는 동시에 사회에 미치는 구체적 영향을 평가할 구체적인 논의 필요
정보03	• 불투명한 알고리즘이 개인의 선택을 가장한 필터버블·에코챔버 현상을 강화하여 정보 접근 편식과 집단 극단화 현상 발생 • 공론장 기능 약화와 민주주의 제도에 대한 신뢰 붕괴로 이어질 구조적 위험성 존재 • (정보-01), (정보-02) 관련 생산자에 의도에 의해 선별되는 정보에 따라 아동·청소년기 사회인지 및 인식 왜곡 가능성 있음 • 다만 구체성 측면에서 모호성이 있어 공론화 대책 필요

의제번호	전문가 근거 요약
정보03	• 공론장의 게이트키퍼 역할을 하는 AI 알고리즘은 상세 내용이 공개되지 않아 정보 흐름을 보이지 않는 손으로 지배 • AI 공공성·책임성 확보를 위한 감독체계 필요 • 기술 발전에 따라 투명성 확보가 어려워짐에 따라 대응 과제 도출이 시급
정보04	• AI 모델 개발 과정에서의 저작물 이용과 창작자 권리 사이의 충돌로 저작권 침해 및 창작 활동 위축 문제에 대한 제도적 해결이 어려움 • AI 효율성에 대한 인식 확대로 예술적 창작 활동의 동기 자체가 위축 • 글로벌 규범과의 정합성을 고려한 정당한 보상 체계 구축 및 사회적 대타협 필요
정보05	• 다른 네 가지 의제의 선행 조건으로서 중요하며, AI 포용 정책의 핵심 • AI 리터러시, 디지털 접근성 강화는 탈진실 대응, 소비자 기만 방지, 알고리즘 편향 완화, 저작권 보호에 기여하는 기반 역량으로 필요 • 인공지능 활용 능력의 격차는 사회불평등으로 이어질 우려 • 노동시장 양극화, 정보권력의 비대칭 등의 문제 가능성 • AI 기술 활용 정도가 사회 계층 재편에 영향

자료: 연구진 작성

전문가들은 우선순위로 선정한 의제들에 대해 다음과 같은 구체적인 정책과제를 제안하였다. ‘(정보-1) 정보 신뢰 붕괴’ 의제에 대해서는 AI 투명성 검증 체계 및 시민의 비판적 정보 역량 강화를 병행하는 원칙 중심의 프레임워크 구축이 강조되었다. 또한 콘텐츠 출처 표시 의무화 및 위험 유형 기반 규제를 통해 플랫폼의 책임을 명확히 하고, 딥페이크 탐지/검증 기술 개발을 위한 R&D 지원과 플랫폼·언론·정부의 공동 대응 체계 마련을 촉구하였다.

‘(정보-3) 알고리즘 불투명성 및 정보 편식’ 의제에 대해서는 AI 공공성·책임성 확보를 위한 감독 체계 수립이 시급하며, 알고리즘 영향 평가 의무화 및 이용자의 ‘알고리즘 접근권’ 보장이 중요하다고 제언했다. 나아가 정부의 독립적 감사와 투명성 강화를 추진하고, 「공정거래법」상 기만·불공정 요인을 알고리즘에 적용하는 법규 연계를 모색해야 한다고 강조하였다.

‘(정보-5) AI 활용 역량 및 정보 격차’ 의제와 관련해서는 AI 리터러시 교육을 기본권 차원에서 논의하고 정규 교육 과정에 포함하는 것이 핵심이며, 고령층, 저소득층 등 취약계층의 접근 인프라를 강화하는 ‘AI 포용 정책’의 강력한 추진이

강조되었다.

‘(정보-4) 저작권 침해 및 창작 활동 위축’ 의제에 대해서는 AI 이용에 기반한 저작자에게 정당한 보상을 지급할 수 있도록 일정 부담금 납부와 수익으로 기금 설치 방안 등에 대한 논의를 제기하였다. 또한 AI 생성 콘텐츠 출처 및 생성 방식 명시 의무화를 통해 기존 저작물 학습에 대한 보상권을 보장하는 법제도 정비가 시급함을 강조하였다.

‘(정보-2) 허위 과장 광고 및 소비자 기만’ 의제와 관련해서는 AI를 악용한 소비자 기만 행위에 대응하기 위해 소비자의 합리적 의사 결정을 왜곡하는 AI 알고리즘의 영향을 점검해야 하며, 관련 제도와 규제에 대한 사회적 공론화와 논의가 필요함을 지적하였다.

〈표 4-23〉 정보 분야 의제 전문가 정책과제 제안

의제번호	전문가 정책과제 제안
정보01	• 개별 사례 대응을 넘어선 원칙 중심의 프레임 정립 필요 • AI 투명성 정보 검증 체계 및 시민의 비판적 정보 역량 강화 병행 • 콘텐츠 출처, 생성 방식 표시 의무화 시스템 구축 • 고위험 분야(선거·보건·재난) 우선 위험 기반 규율 • 플랫폼·언론·정부의 공동규제책 마련 • 알고리즘·광고 추천의 투명성 책임 강화 • 연구 기관·대학의 증거 기반 정책연구 생태계 구축 • 허위 조작 정보 유통 방지를 위한 플랫폼 책임 강화 제도 마련 • 딥페이크 탐지/검증 기술 개발을 위한 R&D 투자 지원
정보02	• 플랫폼 기업의 거대화와 AI 알고리즘 발전이 소비자의 합리적 의사 결정을 왜곡하고 공정한 시장 경쟁을 저해하고 있음을 재평가 • 플랫폼 및 AI 발전에 따른 사회적 비용 절감과 우려를 종합적으로 고려 • 유관 국가 연구 기관과 대학의 공동 연구 마련 • 공론화 과정을 소셜미디어, 카드뉴스, 방송 등으로 일반 대중에게 전파
정보03	• 디지털 플랫폼의 콘텐츠 노출 기준 공개 • 알고리즘의 투명성 확보와 불가능성 제어를 위한 국가적 차원 지원 • 정부의 독립적 감사와 투명성 규제 강화 • 공정거래법상 기만·부당유인 기준을 알고리즘에 적용 • 선거·보건 등 고위험 맥락에서의 신속한 심사·차단 절차 마련

의제번호	전문가 정책과제 제안
정보03	• 관련 기관 및 대학의 증거 기반 공동 연구를 통해 AI 및 플랫폼 계량 평가 정례화 및 결과의 공문화
정보04	• AI 개발에 이용된 저작물의 저작권자에게 정당한 보상책 마련 • AI 개발 및 이용 사업자에 대한 일정 부담금 납부 및 수익으로 기금을 설치하여 창작자 지원 • 개발과 규제 사이의 사회적 대타협과 상호 양보·이해 • 인간 창작자의 권리 보호·투명한 창작 환경 조성 • AI 생성 콘텐츠 출처 및 생성 방식 명시 의무화 • 기존 작품 학습에 대한 보상권 보장
정보05	• 디지털포용법 관련 정책에 대한 사회적 공문화 필요 • 디지털 역량 함양을 위한 균등한 교육 기회 제공 • AI가 초래하는 불평등 해소를 위한 적극적인 정책과 제도적 기반 마련 • 전국민 대상 AI 리터러시 기본 역량 교육 제공 체계 구축 • 취약계층·지역 대상 AI 활용 및 접근 인프라 확충 • AI 리터러시 교육권을 기본권 차원에서 논의

자료: 연구진 작성

라. 의제별 핵심 보완 의견

전문가들은 제시된 5개 의제가 시급성과 타당성을 갖추었음을 인정하면서도, 논의의 효율성을 높이기 위해 의제 간 체계 정립, 범위 조정 등의 보완 가능성에 대해 제언하였다.

먼저, 전문가들은 의제 간의 상호 연관성과 개념적 중복성이 높다고 지적하며, 논의의 효율성을 위한 통합 및 범위 조정을 제안하였다. ‘(정보-1) 정보 신뢰 붕괴’ 의제가 ‘(정보-2) 허위 과장 광고 및 소비자 기만’ 의제를 충분히 포괄할 수 있는 상위 개념으로 판단되었으며, 일부 전문가들은 두 의제를 통합하거나(정보-2)를 (정보-1)의 하위 개념으로 설정하는 것이 적절하다는 의견을 제시하였다. 또한 ‘(정보-3) 알고리즘 불투명성 및 정보 편식’ 의제의 경우, 논의의 초점이 허위조작정보 자체의 문제인지, 알고리즘 증폭 메커니즘의 문제인지, 아니면 소비자 기만 문제인지를 더욱 명확히 구분하여 각 의제의 핵심 초점을 설정할 필요성이 강조되었다.

또한 의제들의 논의 층위가 다름을 지적하며, 정책 수립의 효율성을 위해 의제들을 총론적 의제와 각론적 의제로 구분하는 체계적 관점 도입이 필요하다고 제

언되었다. 정보 생태계 패러다임 변화, 거시적 공공성 및 신뢰 확보 문제 등 구조적 문제에 해당하는 (정보-1), (정보-3), (정보-5)를 총론적 의제로, 개별법 영역 문제, 규제 체계 정립 문제 등 구체적인 법제·정책적 문제에 해당하는 (정보-2), (정보-4)는 각론적 의제로 구분하여 접근하면 효율적인 논의가 가능할 것으로 보았다. 특히 (정보-5)는 정보 생태계의 문제에 국한하지 않고, AI 발전에 따른 불평등 전반을 포괄적으로 검토하는 의제로 확대할 필요가 있으며, 장래의 노동 이슈와 사회보장 제도 변화까지 아우를 수 있는 방향으로 논의를 심화해야 한다는 점이 강조되었다.

마지막으로, 다음과 같은 추가 의제 검토가 필요함을 제기하였다. 첫째, AI 공론장 거버넌스 확립 문제가 있다. 이는 AI가 공론장 기능과 사회 통합에 미치는 광범위한 영향을 고려하여, 기술적·사회적·제도적 측면을 아우르는 민주적 거버넌스 체계를 새롭게 확립하는 문제가 필요함을 시사하였다. 둘째, AI 주도 정보 생산 시대에서의 전통적 전문가 역할 재정의 문제이다. AI가 정보의 생산과 분석을 주도함에 따라, 언론인, 법률가 등 전통적 전문가 집단의 역할, 책임, 권위를 새롭게 정립해야 하는 근본적인 논의가 필요하다고 보았다. 나아가 의료·법률 등 고위험 분야에서의 검색 알고리즘 내용 설명 및 제시 문제가 있다. 오진이나 오판 등 직접적인 피해를 야기할 수 있는 고위험 분야에 AI 알고리즘을 적용할 때, 그 판단의 근거나 과정을 어떻게 투명하게 제시하고 법적으로 규율할지에 대한 구체적인 논의가 필요하였다. 셋째, AI를 활용한 사기, 피싱 범죄의 국가적 대응 체계 강화 문제이다. 뎡페이크 등 AI 기술을 악용한 고도화된 사기 및 피싱 범죄가 증가함에 따라, 기존의 수사 체계를 넘어서는 국가적 차원의 예방 및 대응 시스템을 강화하는 문제가 시급하게 다루어져야 한다고 강조되었다.

5. AI와 지역: 의제 평가 및 주요 의견

가. 평가자 구성

‘AI와 지역’ 분야는 인공지능, 지역경제학, 노동경제학, 공공경제학 분야의 전문

가 5인이 조사에 참여하였다. 본 조사에서 전문가에게 제시한 지역분야 5개의 의제는 다음과 같다.

〈표 4-24〉 전문가 의견조사에 제시된 AI-지역 영역 사회적 의제(안)

(지역-1) 데이터센터의 지역 분산을 촉진하기 위한 정부의 역할은 무엇인가?	
제안의 배경	• AI 산업의 핵심 인프라인 데이터센터의 약 60%가 수도권에 입지(한국데이터센터연합회, 2024) • 데이터센터의 지역적 집중이 AI 연산 인프라 접근성의 지역 격차를 심화시켜 수도권 중심의 산업 구조를 고착화할 우려가 제기됨
현황 분석	• 하이퍼스케일 데이터센터는 대규모 전력공급·부지를 갖춘 지방 입지가 유리하나, 지방은 저지연 네트워크 접근성이 수도권에 비해 낮고, 전문 인력 확보가 어려움 • 지역 주민의 수용성(납비현상, 일자리 창출 효과에 대한 회의감) 문제로 지방 유치가 지연되는 사례 존재
문제의 심각성	• AI 산업 발전에 필수적인 인프라 구축이 사회적 갈등으로 지연되고 있음 • 데이터센터의 수도권 집중이 지속될 경우 비수도권의 AI 산업 성장 여건이 악화되고 지역 간 혁신 격차가 심화될 가능성이 존재
(지역-2) AI 확산을 위한 지역별 에너지 인프라 확충 전략은 무엇인가?	
제안의 배경	• AI 산업의 고도화로 전력 수요가 급증하는 가운데, 기존 전력 공급량이 많은 비수도권 지역에 의존하는 것만으로는 한계가 있음 • 지속적인 AI 산업 성장을 위해서는 전력 생산의 확대와 다변화가 필요하며, 지역별 여건에 맞는 에너지 인프라 확충 전략이 요구됨
현황 분석	• 전력 생산 잠재력이 높은 지역은 발전설비를 추가 구축하고, 전력 수요 집중 지역은 분산형 전원 기술을 통해 전력 공급을 확대할 수 있음 • 그러나 규제, 경제성, 주민 수용성의 문제로 추가 전력 공급원 확보에 어려움이 있음
문제의 심각성	• 수도권 중심의 전력 소비 구조가 지속될 경우, 부지 부족과 송전망 부담 증가로 인해 전력계통의 안정성이 저하될 가능성이 높음 • AI 산업의 확산 속도에 비해 추가적인 전력 공급 확충이 뒤처질 경우, 전력망 병목으로 이어질 우려가 제기됨
(지역-3) 지역 특화 산업과 AI 융합을 통한 지역 경제 활성화 방안은 무엇인가?	
제안의 배경	• 지역별로 고유한 자원과 산업 기반이 존재하며, AI 기술은 이러한 지역 특화 산업의 생산성 향상을 이끌 핵심 동력으로 작용 가능 • 지역 특화 산업을 반영한 맞춤형 AI 융합 전략을 통해 지역 경제를 활성화하고, 지방 소멸 문제를 완화할 필요가 제기됨

현황 분석	• 다수의 비수도권 지역에서 주력산업의 침체로 새로운 성장 동력 발굴이 지연되고 있는 것으로 나타남 • AI를 접목한 스마트팜·스마트팩토리·관광 콘텐츠 등 사례가 있으나, 산업 생태계 전반으로의 확산은 제한적임
문제의 심각성	• 지역 특화 산업의 AI 전환 지연은 생산성 향상 정체와 경쟁력 저하로 이어져 지역 경제 기반이 약화될 우려가 제기됨 • 이에 따라, 비수도권 지역의 혁신 역량과 산업 생태계가 위축되어 지역 간 격차가 확대될 가능성이 있음
(지역-4) AI 스타트업 생태계의 지역 간 확산을 촉진하기 위한 전략은 무엇인가?	
제안의 배경	• 2024년 기준, AI 산업의 혁신 동력인 스타트업의 약 80%가 수도권에 집중되어 있음(The VC, 2024) • 이러한 지리적 집중은 집적 효과를 통해 생산성을 높이지만, 지역 간 혁신 역량의 격차를 심화시킬 우려가 제기됨
현황 분석	• 2024년 기준, AI 벤처캐피털 투자액의 약 70%가 수도권 소재 스타트업에 집중되어 있는 상황(The VC, 2024) • 이처럼 투자를 비롯한 초기 창업 자금·산학협력 네트워크 등 스타트업 생태계가 수도권을 중심으로 형성·발전되어 있음
문제의 심각성	• 수도권 중심의 스타트업 생태계는 지역 간 혁신 격차를 심화시키며, 지역 산업의 AI 전환 혜택이 일부 지역에만 집중될 수 있음 • 이에 따라 혁신의 지리적 병목이 지속되면, AI 기술이 지역 균형 발전의 새로운 동력으로 작용하기 어려워질 가능성이 있음
(지역-5) 지역 대학을 중심으로 한 효과적인 AI 인재 양성 방안은 무엇인가?	
제안의 배경	• 혁신의 핵심 주체 중 하나는 인적자본이며, 인적자본의 지역적 분포는 지역 간 혁신 격차에 영향을 미침 • AI 관련 인재가 수도권에 집중되어 있어, 비수도권의 AI 혁신 활동과 산업 고도화가 제약될 우려가 있음
현황 분석	• 지방대학 육성 정책이 추진되고 있으나, AI 관련 교육·연구 프로그램이 충분히 구축되지 못한 지역이 많음 • 교수진 확보, 산업 연계, 연구 인프라 등에서 수도권과의 격차가 존재하며, 인적자본 개발을 위한 생태계가 미흡한 상태
문제의 심각성	• 지방대학이 AI 교육 기반을 갖추더라도, 졸업생의 상당수가 수도권으로 이동할 경우 지역 내 인재 축적이 어려울 수 있음 • 이로 인해 지방 대학이 지역 혁신 역량 강화로 이어지지 못하고, AI 산업 확산의 지리적 병목이 심화될 가능성이 있음

자료: 연구진 작성

나. 정량 평가 결과

‘AI와 지역’ 분야 전문가 5인을 대상으로 5개 의제(안)에 대해 4개 항목(문제의 공공성, 의제의 타당성, 논의의 필요성, 사회적 파급성)을 5점 척도로 평가하도록 하였다.

5개 의제 중에서 ‘(지역-2) 지역별 에너지 인프라 확충 전략’ 의제는 5인 평균 4.40점으로 가장 높게 평가되었으며, ‘사회적 파급성’(4.60점), ‘문제의 공공성’ 및 ‘논의의 필요성’(각각 4.40점)에서 높은 점수를 기록하였다.

다음으로, ‘(지역-3) 지역 특화 산업과 AI 융합’과 ‘(지역-1) 데이터센터의 지역 분산’이 각각 5인 평균 3.70점 및 3.60점을 기록하였다. 특히, ‘(지역-3) 지역 특화 산업과 AI 융합’은 ‘의제의 타당성’에서 4.00점을, ‘(지역-1) 데이터센터의 지역 분산’은 ‘문제의 공공성’에서 4.00점을 받았다.

한편 ‘(지역-4) AI 스타트업 생태계의 지역 간 확산’과 ‘(지역-5) 지역 대학 중심 AI 인재 양성’은 각각 평균 3.00점을 기록하였으며, 점수 편차가 커 전문가 간 관점 차이가 존재하는 것으로 해석된다.

〈표 4-25〉 지역 분야 의제 전문가 정량 평가 결과

구분	문제의 공공성	의제의 타당성	논의의 필요성	사회적 파급성	의제별 평균
(지역-1) 데이터센터의 지역 분산	4.00	3.20	3.60	3.60	3.60
(지역-2) 지역별 에너지 인프라 확충 전략	4.40	4.20	4.40	4.60	4.40
(지역-3) 지역 특화 산업과 AI 융합	3.60	4.00	3.60	3.60	3.70
(지역-4) AI 스타트업 생태계의 지역 간 확산	3.20	3.20	2.80	2.80	3.00
(지역-5) 지역 대학 중심 AI 인재 양성	2.60	3.60	2.80	3.00	3.00
평가항목별 평균	3.56	3.64	3.44	3.52	

자료: 연구진 작성

다. 정성 의견 분석

5개 의제 중 가장 중요하고 우선순위가 높아야 할 의제 3개를 선정한 결과, 5인의 전문가 중 4명이 ‘(지역-1) 데이터센터의 지역 분산’과 ‘(지역-2) 지역별 에너지 인프라 확충 전략’을 Top 3 의제로 선정하여 가장 높은 중요도를 보였다. ‘(지역-3) 지역 특화 산업과 AI 융합’의 경우 5인의 전문가 중 3명이 Top 3로 선정하였으나, 모두 3순위로 선정하였다. 마지막으로, ‘(지역-4) AI 스타트업 생태계의 지역 간 확산’과 ‘(지역-5) 지역 대학 중심 AI 인재 양성’은 5인의 전문가 중 2명이 Top 3로 선정하였다.

〈표 4-26〉 지역 분야 의제 전문가 우선순위 평가 결과

구분	1순위 선정횟수	2순위 선정횟수	3순위 선정횟수	총 선정횟수
(지역-1) 데이터센터의 지역 분산	2	1	1	4
(지역-2) 지역별 에너지 인프라 확충 전략	3	1	0	4
(지역-3) 지역 특화 산업과 AI 융합	0	0	3	3
(지역-4) AI 스타트업 생태계의 지역 간 확산	0	2	0	2
(지역-5) 지역 대학 중심 AI 인재 양성	0	1	1	2

자료: 연구진 작성

각 의제를 선정한 전문가들의 선정 사유는 아래와 같이 정리하였다. 전문가들은 AI 산업 확산과 관련하여, 안정적인 전력 공급이 선결 조건이라는 점에서 ‘(지역-2) 지역별 에너지 인프라 확충 전략’을 가장 중요한 사회적 의제로 보았으며, 데이터센터 입지가 산업·인재·서비스의 공간적 분포를 구조적으로 결정한다는 점에서 ‘(지역-1) 데이터센터의 지역 분산’ 역시 핵심 의제로 평가하였다. 또한 지역 고유 산업과 AI를 결합하는 전략은 비수도권 산업 침체와 지방소멸 문제에 대응하며 인프라 구축 이후 실제 지역 경제 활성화로 이어지는 경로라는 점에서 ‘(지역-3) 지역 특화 산업과 AI 융합’도 의미가 있다고 보았다. 더불어 지역별 수요 기반으로 AI 서비스 개발을 촉진할 수 있다는 점에서 ‘(지역-4) AI 스타트업 생태

계의 지역 간 확산'이 중요하다고 인식되었으며, 장기적으로 지역 대학을 거점으로 한 인재 양성과 정착 생태계 구축이 지역 균형 발전에 기여할 수 있다는 점에서 '(지역-5) 지역 대학 중심 AI 인재 양성'도 필요하다고 평가되었다.

〈표 4-27〉 지역 분야 의제 전문가 우선순위 선정 사유

의제번호	전문가 근거 요약
지역01	• 데이터센터는 AI 산업의 핵심 물리 인프라로, 입지가 지역 산업·인재 분포를 구조적으로 결정할 수 있음 • 지방이 부지·전력 측면에서 유리하나 네트워크·인력·수용성 제약으로 실현되지 못하고 있어 정부 조정 필요 • 단순한 기술 문제를 넘어 초기 분산이 장기적 지역 균형 발전에 중요하며, 전력 인프라 논의와 연계해 접근 필요
지역02	• AI 확산으로 전력 수요 급증, 에너지 공급이 AI 산업 성장의 선결 조건 • 수도권 중심 전력 구조가 송전망 부담 및 산업 경쟁력 저하로 이어질 우려 • 지역별 발전 잠재력·수요·송전망 조건이 상이하므로 지역 맞춤형 전력 공급 전략 필요 • AI 서비스 수요 확대 시 에너지 부족이 AI 도입의 병목으로 작용할 수 있음
지역03	• 지역 고유 산업과 AI 결합을 통해 비수도권 성장동력 확보 및 지방소멸 대응 가능 • 전통 산업 침체와 지역 간 생산성 격차 확대를 완화하는 핵심 전략 • 지역 특화 산업과 AI 융합은 데이터센터와 전력 인프라와 같은 공급 측 기반을 실제 산업 성장으로 연결하는 후속 의제로서, 지역 경제 활성화와 생산성 제고에 직접적으로 기여할 수 있음
지역04	• 지역 산업 성장과 데이터센터 분산을 촉진할 수 있는 해결 방안 중 하나 • 지역 수요 기반 AI 서비스 개발을 위해 지역 생태계 조성이 필요 • 다만 중앙정부 개입보다는 지역 산업·대학·지자체 역할이 더 중요하다는 관점 존재
지역05	• 지역 대학을 통한 인재 양성 및 정착 생태계 구축은 장기적 균형 발전에 필수 • 대학 부지 내 데이터센터 입지 등 산학 협력 강화 시 시너지 가능 • 다만 인재 수요가 형성될 경우 자연스럽게 해결될 수 있다는 시각도 존재

자료: 연구진 작성

전문가들은 우선순위로 선정한 의제들에 대해 다음과 같은 구체적인 정책과제를 제안하였다. ‘(지역-2) 지역별 에너지 인프라 확충 전략’ 관련하여, 전문가들은 AI 확산에 필요한 전력 수요 증가에 대응하기 위해 지역별 발전 잠재력과 송전망 여건을 반영한 맞춤형 에너지 공급 전략을 마련하고, 데이터센터 입지와 연계해 전력 인프라를 선제적으로 확충할 필요가 있다고 보았다. 한편 ‘(지역-1) 데이터 센터의 지역 분산’에 대해서 전문가들은 데이터센터는 부지·전력 등 물리적 요인이 지방에 유리함에도 네트워크 접근성·수요 기반·주민 수용성 문제로 입지가 제한되고 있으므로, 경제적 인센티브 제공, 특화단지 조성, 이익 공유 모델 마련 등을 통해 지방 분산을 촉진해야 한다는 의견이 제시되었다. ‘(지역-3) 지역 특화 산업과 AI 융합’을 위해서는 지역별 주력산업에 맞춘 기술 적용 로드맵을 수립하고 실증 단지를 조성하며, 중소기업 AI 전환을 위한 바우처 및 맞춤형 플랫폼을 구축해 산업 기반의 생산성 향상을 유도해야 한다는 제안이 있었다. 아울러 ‘(지역-4) AI 스타트업 생태계의 지역 간 확산’에 대해서는 지역 산업 수요와 연계하여 육성하고, 클라우드 크레딧·금융지원·산학협력 강화를 통해 지역 기반 혁신 기업을 발굴·정착시키는 정책이 필요하다는 견해가 확인되었다. 마지막으로 ‘(지역-5) 지역 대학 중심 AI 인재 양성’ 관련하여 지역 대학을 중심으로 AI 특화 교육 프로그램, 계약학과, 산학 프로젝트, 데이터센터·기업과 연계한 취업·연구 구조를 구축해 지역 내 인재 정착과 산업 연계 역량을 강화해야 한다는 제안이 제시되었다.

〈표 4-28〉 지역 분야 의제 전문가 정책과제 제안

의제번호	전문가 정책과제 제안
지역01	• 지방 입지 유인을 위한 부지·세제·전력 요금 감면 등 경제적 인센티브 패키지 제공 • 저지연 네트워크·AI 연산·클라우드 연계 등 인프라를 갖춘 데이터센터 특화단지 조성 • 주민 수용성 제고를 위한 이익 공유(지역기금·지역 청년 채용·지역 환원 구조 등) 마련 • 유형별 데이터센터 특성(산업형 vs 기업 자체형)에 따른 맞춤형 분산 전략 추진

의제번호	전문가 정책과제 제안
지역02	• 지역별 발전 잠재력 기반 '맞춤형 에너지 로드맵' 수립 및 전력 공급원 다변화 • 데이터센터 밀집 지역에 분산형 전원·SMR·연료전지·PPA 등 온사이트 전력 확보 추진 • 송전망·발전설비 관련 인허가 간소화, 규제 개선 및 신속한 구축 체계 마련 • 지역 환원·전기요금 혜택 등 주민 수용성 강화를 위한 이익공유제 도입
지역03	• 지역 주력산업별 AI 적용 로드맵 및 실증 단지 조성(농업·제조·관광 등) • 중소기업 AI 도입을 위한 컨설팅·솔루션 구매 바우처 확대 • 산업 공통 AI 플랫폼 구축(예: 농업·제조·유통 데이터 통합 플랫폼) • 초기 성공 기업을 집중 지원해 확산 모델을 만드는 '선도기업 육성 프로그램' 추진
지역04	• 지역 특화 산업과 연계한 AI 창업 육성 및 수요 기반 서비스 개발 지원 • AI 스타트업 지원을 위한 클라우드 크레딧·금융·보조금 등 인센티브 제공 • 산학협력 및 지역 기업·지자체 연계로 수요·시장 기반 생태계 조성
지역05	• AI 특화 계약학과 신설 및 산업체 연계 실습 프로그램 확대 • 대학 부지 내 데이터센터 유치 등 캠퍼스형 AI 클러스터 구축 • 지역 기업 디지털 전환 컨설팅·현장 참여 프로그램으로 지·산·학 연계 강화 • 지역 산업 수요 기반 인재 정착 모델 구축(졸업 후 지역 취업 연계 등)

자료: 연구진 작성

라. 의제별 핵심 보완 의견

전문가들은 제시된 의제의 방향성에는 대체로 동의하면서도, 정책 설계 시 추가적으로 고려해야 할 관점이 있다고 보았다. 먼저, ‘(지역-1) 데이터센터의 지역 분산’과 ‘(지역-2) 지역별 에너지 인프라 확충 전략’이 지역 균형 발전을 위해 중요하지만, 인프라 입지를 결정할 때는 지역 균형 발전이라는 목표보다 국가 경쟁력, 비용 효율성, 전력 공급 여건 등 경제적 최적화를 우선 고려해야 한다는 점을 강조하였다. 데이터센터 위치가 AI 서비스 확산과 직접적으로 연결된다는 가정에는 일부 회의적인 견해도 있었으며, 단순한 부지 유치보다 산업·인력·스타트업 생태계와 연계해 실질적 경제효과로 이어지도록 설계할 필요가 있다고 보았다. 에너지 인프라 의제는 전력 수요 전망과 지역별 전력 구조를 뒷받침할 수 있는 데이터가 마련되면 정책 논의의 실효성이 높아질 수 있으며, 데이터센터·산업수요·에너지공급 간 선후관계와 정책 로드맵을 명확히 해야 한다는 의견이 있었다.

또한 ‘(지역-3) 지역 특화 산업과 AI 융합’은 자동화, 데이터 기반 의사결정, 신

산업 창출 등 다양한 의미로 해석될 수 있으므로, 정책 방향을 보다 명확히 정의할 필요가 제기되었다. 산업별 구체적 적용 모델, 기술 요건이 논의되어야 하며, 중소기업의 도입 장벽과 전환 비용을 고려한 실증 중심 접근이 필요하다는 의견이 있었다. ‘(지역-4) AI 스타트업 생태계의 지역 간 확산’과 ‘(지역-5) 지역 대학 중심 AI 인재 양성’은 국내 AI 경쟁력·핵심 인재 확보가 선행되지 않은 상황에서 시기상조일 수 있으며, 산업 수요 기반 없이 추진될 경우 지역 정착 효과가 낮을 수 있다는 문제점이 지적되었다. 나아가 정책 주체 간 역할 분담, 주민 수용성, 국제 비교, 중장기 성과지표 마련 등 정책 실행을 위한 구체화가 필요하다는 의견이 제시되었다.

전문가들은 AI 시대 상생과 포용을 위한 지역 분야에 새로운 의제안으로 다음과 같은 안을 제시하였다.

〈표 4-29〉 지역 분야 전문가 신규 제안 의제

의제 제목	제안 내용 및 배경
AI 거점 간 연계 전략	- 제안 내용: “지역별 데이터센터·AI 기관 간 연계 및 분업 체계 구축 필요” - 제안 배경: 현재 지역 거점들이 고립된 생태계로 존재하여 국가 단위 협업 부재
공공투자 기반 연산자원 배분	- 제안 내용: “공공투자로 구축된 연산자원(GPU·데이터센터) 통합 관리 및 분배 기준 마련” - 제안 배경: 정부 지원 기반 인프라 확충 시 산업·기업·지역별 접근권 배분 문제 발생
지역 주민 대상 포용적 AI 교육·접근성 강화	- 제안 내용: “비수도권·중장년층 등 디지털 소외 계층 대상 AI 교육접근성 정책 마련” - 제안 배경: AI 인프라 구축만으로는 혜택 확산 어려우며 지역 내 불평등 심화 우려
AI 기반 지역 공공서비스 격차 해소	- 제안 내용: “의료·교육·행정 등 공공서비스 영역의 접근성 강화 및 수요 기반 설계” - 제안 배경: 공공서비스가 대도시 중심으로 도입되어 지역 주민 삶의 질·기본권 격차 심화

자료: 연구진 작성

제 5 장 결 론

제 1 절 AI 시대 상생과 포용을 위한 핵심의제 종합

1. 노동 분야 의제

〈표 5-1〉 노동 분야 핵심의제

(노동01) 문제 AI로 인한 직무 전환이 요구됨에 따라, 마모되는 근로자의 인적자본을 회복하기 위한 교육훈련 시스템은 어떻게 설계할 것인가?
• (제안의 배경) AI가 에이전트 AI, 물리 AI로 진화하며 지식 산업과 제조 분야 등 광범위한 산업의 직무 재편을 촉발. 이는 생애주기 소득 안정성을 위해 ‘평생 직장’이 아닌 ‘평생 학습’ 기반의 상시적 직무 전환 체계를 요구하며, 근로자의 지속적인 직무 전환이 시급 • (현황 분석) 국내외 주요 연구로부터 중장기적으로 기존 일자리의 상당 부분이 AI로 대체되고 직무 내용의 근본적 변화가 예측되지만, 현재 정부 주도 직업 훈련은 실업자 중심이며 재직자 참여는 저조. 특히 중장년층과 저숙련 근로자는 심리적, 시간적 장벽으로 인해 재교육의 사각지대 존재 • (문제의 심각성) AI는 근로자가 축적한 인적자본의 마모를 가속. 숙련 인력이 저숙련 노동으로 하향 이동하거나 노동시장에서 조기 이탈할 경우, 개인은 장기 실업과 소득 공백이 확대되고 국가 전반의 노동생산성 하락 우려. 또한 실업급여 등 복지 재정에 대한 사회적 비용을 증대 부담 • (기존 논의의 한계) 기존 고용보험기금 기반 훈련은 ‘실직한 이후’에 작동하는 한계가 있음. AI 시대에는 직무가 소멸하기 전 ‘선제적’ 전환 교육이 필요하나, 현재 제도는 경직되어 있으며 훈련 내용이 산업계의 실제 AI 수요를 신속히 반영하기 어려움. 하드스킬, 소프트스킬 등을 결합한 전환기 맞춤형 역량 체계가 부재
(노동02) AI가 청년층의 초기 경력 형성에 필수적인 일자리를 대체해 나가는 상황에서 ‘경력 단절 없는’ 청년 고용 시장은 어떻게 보장할 것인가
• (제안의 배경) AI가 데이터 입력, 단순 사무 보조 등 초임·보조 수준의 진입 직무(entry-level jobs)를 빠르게 자동화. 데이터 입력, 단순 분석, 보조 수준의 사무직 등 청년의 첫 일자리로 기능해 온 직무가 빠르게 축소되며 청년층이 경력 사다리의 첫 번째 칸에 오를 기회 자체를 박탈하며, 단기적 마찰적 실업이 아닌 ‘세대 단위의 구조적 일자리 상실’로 고착화될 수 있어 공론화가 시급

- (현황 분석) 해외 연구에서는 AI 도입 이후 노출도가 높은 경력 진입 단계 근로자의 고용이 13% 감소했음을 보고, 국내 연구(한진수·오삼일, 2025) 역시 AI 확산으로 청년 일자리가 크게 위축되었음을 추정.
- (문제의 심각성) 초기 경력 부재는 장기적 고용 가능성을 낮추고, 임금 경로·직장 안정성·생애소득에 구조적 손실과 '고용 불가'로 이어지는 악순환이 고착. 청년층의 경력 축적 실패는 국가 인적 자본의 손실로 이어져 장기적 노동생산성 하락을 초래하며, 사회부조 등 사회적 비용 상승으로 귀결
- (기존 논의의 한계) 기존의 인턴십 확대, 단기 일자리 제공 등 '양적 공급' 대책은 AI로 인한 '엔트리 레벨 직무의 구조적 소멸'이라는 근본적 변화에 대응하기 부족. 또한 '일경험·훈련·취업'이 체계적으로 연결되지 않아 지속가능한 경력 경로 설계가 부재하며, 기업의 청년 신규채용 유인이 갈수록 감소

(노동03) AI로 인한 노동시장의 창조적 파괴 과정에 적응하기 위해 발생하는 전환 비용은 누가, 어떤 비율로 분담되어야 하며, 배분의 우선순위는 어떤 사회적 기준으로 정할 것인가?

- (제안의 배경) AI로 인한 직무의 '파괴'(빠름/집단적)와 '창출'(느림/분산적) 간에는 속도와 시점의 불일치가 발생. 이 과정에서 근로자는 '소득의 공백', '숙련의 공백'에 직면하며 막대한 '전환 비용'(재훈련, 이주 비용 등)이 발생하나, 이 비용과 책임을 배분하는 원칙은 부재
- (현황 분석) 과거 기술/무역 충격(차이나 쇼크, 로봇 도입) 연구에서도 일자리 대체 효과가 창출 효과를 상회하거나 회복이 매우 느렸음이 확인. AI 시대에도 대체되는 근로자와 수혜를 입는 근로자가 서로 다른 집단일 가능성이 높으며, AI를 노동 증강형으로 활용해 중산층을 재건할지는 이해관계자 간의 집단적 선택의 문제
- (문제의 심각성) 전환이 지연되면 가구 소득과 지역 경제에 구조적 손실이 누적된다. 과거 사례처럼 자연스러운 노동 재배치는 일어나지 않으며, 장기적 임금/고용 하락이 고착화
- (기존 논의의 한계) 기존 실업급여 등은 단기 실업을 전제로 하므로 AI로 인한 구조적, 장기적 전환 비용을 감당하기 어려움. 스웨덴의 '전직보장협의회'처럼 노사 공동재원으로 선제적 전환을 지원하는 제도가 있으나, 강력한 사회적 파트너십이 부재한 대부분의 국가에서는 비용이 개인이나 중소기업에 편중

(노동04) AI가 구현하는 알고리즘 매니지먼트(채용, 업무 배분, 감시, 평가)가 노동 통제를 강화하고 노동의 질을 저해하는 상황에서, 우리는 근로자의 기본적인 권리를 어떻게 보호할 것인가?

- (제안의 배경) AI가 업무 배분, 성과 평가, 해고에 이르는 인적자원 관리 전반에 개입하며 노동의 질적 변화를 유발. 실시간 데이터 기반의 감시와 불투명한 의사결정은 근로자의 자율성을 침해
- (현황 분석) 알고리즘 매니지먼트는 플랫폼 노동을 넘어 콜센터, 물류창고, 사무직, 원격근무자에게까지 확산. 키보드 입력이나 웹캠을 감시하는 '보스웨어(Bossware)'는 심각한 사생활 침해를 유발하며, AI 기반 채용에서는 편향성으로 인한 공정성 이슈가 우려
- (문제의 심각성) 기업은 생산성 향상을 명분으로 감시와 통제를 강화할 유인. AI로 초 단위 감독은 노동의 '비인간화'를 초래하며, '디지털 테일러리즘'을 부활시켜 장기적으로 노동의 질과 창의성 저해

- (기존 논의의 한계) 현행 근로기준법은 ‘인간 관리자’의 ‘명시적’ 지시를 전제로 하므로, AI에 의한 ‘간접적’, ‘시스템적’ 통제를 규율하기 어렵다. EU 등에서 논의되는 ‘설명요구권’이나 ‘알고리즘 영향 평가’ 등이 국내에는 초기 단계에 머물러, 근로자가 불이익의 이유를 알거나 구제받을 실질적 경로가 없음

(노동05) AI가 창출한 초과 이익이 소수 기업에 독점되고 임금과 고용상승으로 연결되는 밴드왜건 효과가 부재한 상황에서, 이 이익의 범위는 어떻게 정의하며 누구와 어떤 원칙으로 공유되어야 하는가?

- (제안의 배경) AI의 가치 창출은 거대 모델, 데이터 등 무형 자산을 보유한 소수 기업에 귀속. 반면, 데이터 생성을 제공한 대중이나 직무 전환 비용을 부담하는 근로자는 이익 분배에서 배제. AI가 고용 감소나 질 하락을 유발하여 밴드왜건 효과를 누리기 어렵다는 전망이 제기됨에 따라, 초과 이익 공유의 필요성이 대두
- (현황 분석) 디지털 경제에서 데이터 기여에 대한 보상(데이터 배당) 논의가 있어왔으며, 빌게이츠 등은 자동화 기기에 대한 로봇세(AI세)를 부과하여 전환 비용으로 사용할 것을 제안했다. IMF 등 국제기구에서도 자본세, 기본소득 등 조세재정 수단을 통한 대응 효과를 분석하고 있으나, 각 분배 방식은 한계와 지속성에 차이
- (문제의 심각성) AI 혜택의 독점은 사회적 불평등과 산업 생태계 왜곡을 심화시키고, 노동자의 임금 감소는 총수요 위축으로 이어질 수 있음. 이익의 편중이 고착될 경우, AI 기반 산업 구조 자체가 사회적 지속가능성을 상실할 우려
- (기존 논의의 한계) 초과이익 공유 필요성에 대해 이해관계자 간 의견 차이가 존재. AI나 로봇의 범위, 초과 이익을 측정할 과세 단위 등 새로운 표준을 정하고 측정 방법을 마련하는 것이 기술적·행정적으로 용이하지 않음

자료: 연구진 작성

2. 공공행정 분야 의제

〈표 5-2〉 공공행정 분야 핵심의제

(행정01) 접근성 격차와 알고리즘 편향 관점에서 디지털 취약계층은 소외되지 않을 수 있는가?

- (제안의 배경) AI 공공·행정서비스는 국민의 민감한 개인정보를 처리하므로 해킹 발생 시 국가 행정망 마비로 이어질 수 있음
- (현황 분석) 2024년 개인정보 유출 신고 원인의 과반(56%)이 해킹으로 인한 것이며, 특히 공공기관 관리자 페이지를 겨냥한 해킹이 증가함. 서울시는 ‘AI 기반 사이버보안 종합계획’을 수립하는 등 대응 노력을 하고 있으나, AI를 활용하여 해킹 공격을 수행하는 등 ‘AI 간의 싸움’으로 예측되어 보안 환경은 더 복잡해질 것으로 전망됨
- (문제의 심각성) AI 시스템은 다양한 구성 요소로 이루어져 있어, 기존 IT 시스템보다 잠재적인 해킹 경로(취약점)가 더 많음. AI 모델 자체가 학습 데이터를 암묵적으로 기억하는 특성이 있어, 해커가 모델 분석을 통해 원본 개인정보를 추출해낼 위험이 존재함. 나아가 핵심 행정망 교란 등 국가 안보를 직접적으로 위협하는 도구로 악용될 수 있음
- (기존 논의의 한계) AI 기반 서비스의 기술적 안정성과 개인정보 보호 노력은 지속되고 있으나, 급변하는 AI 위협 환경에 대응하기에 한계가 있음. AI 시대의 프라이버시 위협은 단순 유출을 넘어, 알고리즘에 의한 예측, 추론, 차별 등 개인이 통제할 수 없는 능동적 피해로 변화하고 있으므로, 현행 「개인정보보호법」만으로는 한계가 명확함

(행정02) AI 기반 공공·행정서비스에 오류가 발생하여 시민이 피해를 입었을 때, 책임 공백은 해소될 수 있는가?

- (제안의 배경) AI로 인한 취약계층의 접근 배제와 알고리즘의 불공정한 기준 내재화 위험은 소외 문제를 심화시킬 수 있음
- (현황 분석) 고령층의 디지털 역량 점수가 다른 취약계층 대비 현저히 낮은 상황으로, 정책적 혜택조차 활용 능력 부족으로 소외될 위험이 높음. 또한 편향성을 내재한 알고리즘으로 차별적 의사결정이 발생함
- (문제의 심각성) 접근 격차는 불편과 소외를 야기하며, 시민으로서의 기본 권리 및 기회 박탈로 이어짐. 알고리즘 편향은 눈에 보이지 않는 차별을 시스템화하고 영구화하는 결과를 초래할 수 있음
- (기존 논의의 한계) 기존 정책은 디지털 기기 보급 및 이용 교육 등 기술적 접근성 향상에 초점을 맞추고 있어, 기술 발전 속도를 따라잡는 데 한계가 있음. AI 서비스 설계 단계부터 소외계층을 고려하는 ‘유니버설 디자인’ 원칙 적용과 ‘알고리즘 영향평가’의 법적 의무화 등에 대한 논의가 필요함

(행정03) ‘시민 안전’과 ‘범죄 예방’을 명분으로 개인의 일거수일투족을 감시하고 통제하는 ‘빅브라더’ 사회로의 귀결을 막을 수 있는가?

- (제안의 배경) AI의 블랙박스(Black Box) 특성으로 인해 시민에게 피해 발생 시 책임 소재 규명이 어려움. 따라서 AI 시스템 오류로 시민에게 치명적 피해가 발생할 경우, 책임 공백이 발생할 가능성이 높음

- (현황 분석) 조사결과, 시민들은 '정책 결정 과정에서 발생하는 책임소재 불분명'에 대해 우려 수준 대비 정부의 준비 수준이 낮다고 평가함. AI 기반 시스템의 활용에 대한 근거는 마련되는 한편, 오류 발생 시 피해 배상의 책임 주체와 절차에 대한 명확한 가이드라인은 부재한 상황임
- (문제의 심각성) AI 기반 서비스 오류 시 책임 주체는 알고리즘 개발사, 데이터 제공 기관, AI 도입 정부 부처, 공무원, 그리고 AI 시스템 자체 등 복잡하게 얽혀 있어 판단이 매우 어렵고, 책임을 회피할 가능성이 있음. 치명적인 피해 발생 가능성은 AI 공공 서비스에 대한 불신으로 이어지며, AI 도입을 꺼리는 소극 행정이나, 모든 책임은 AI로 돌리는 무책임한 행정을 초래할 가능성이 있음
- (기존 논의의 한계) 정부는 '책임 있는 AI'를 위한 가이드라인과 원칙을 제시하고 있으나, AI 도입의 각 단계별 책임 분산 및 귀속에 대한 구체적인 논의는 부재함

(행정04) 기술적 안정성 및 신뢰 차원에서 개인정보 보호(해킹 취약, 신종 범죄에 악용 등)는 가능한가?

- (제안의 배경) 공공 CCTV와 AI 기술 결합으로 대규모 실시간 자동화 감시가 가능해지면서, 사실상 모든 시민이 감시 영향 하에 있을 가능성이 높아지는 만큼 기본권 침해 논의가 시급함
- (현황 분석) 대부분의 자치단체는 대규모 예산을 투입하여 CCTV와 AI 기술을 결합하고 있으나, 기술 도입의 필요성과 효과에 대한 사전 평가가 충분히 이루어지지 못한 채 행정 당국의 판단에 의해 일방적으로 결정됨
- (문제의 심각성) 군집 사고 예방에 기여하는 등 긍정적인 측면이 있지만, CCTV에 포착되는 모든 이들의 정보를 실시간으로 수집하고 미래 행동을 예측할 수 있다는 점에서 논란의 여지가 있음. 본래 목적을 넘어 특정 개인의 동선 추적, 집회 참여자 식별 등에 활용될 수 있다고 판단되면, 시민들의 자유로운 의사 표현, 집회 및 결사의 자유 등 민주주의의 근간이 되는 기본권 행사를 스스로 검열하고 포기하게 될 수 있음
- (기존 논의의 한계) '감시' 행위에 대한 명확한 법적 정의 및 규제 범위가 확립되지 않아 AI 감시 기술의 파급력을 효과적으로 통제하기 어려움. 또한 '개인정보보호법'은 주로 유출 후 사후 처벌 및 배상에 초점을 맞춘다는 한계가 있으며, EU가 「AI Act」를 통해 실시간 원격 생체 인식(안면 인식 등)을 원칙적으로 금지하는 것과 달리 우리나라에는 이를 규율할 상위법이 부재함

(행정05) AI의 정책 결정 과정(이유 및 근거)에 대해 시민이 납득할 수 있는 설명을 제공할 수 있는가?

- (제안의 배경) AI 정책 결정의 이유와 근거가 명확히 제시되어야 시민의 이의제기권과 권리 구제가 보장될 수 있는데, 모델의 정확성과 설명 가능성은 상충 관계임
- (현황 분석) 한국전자통신연구원 등이 정책 의사결정을 돕는 AI 정책지능 연구 협력체계를 구축하고, 영국 지방정부가 주택정책 결정에 AI를 도입하는 등 AI 정책 의사결정은 점차 심화될 것임. 이에 따라 AI 설명 가능성에 대한 논의가 활발히 진행 중이나, 시민이 납득할 수 있는 설명의 수준과 범위에 대한 사회적 합의가 부족한 실정임
- (문제의 심각성) 설명 가능한 인공지능(XAI)은 AI 모델 해석에 집중하지만, 기술적으로 완벽한 설명이 비전문가인 시민이 이해할 수 있는 직관적이고 맥락적인 설명으로 전환되는 수준은 상대적으로 미흡함. 모델의 정확성과 설명 가능성 확보는 상충 관계에 있어 두 가지를 충족하기 어려운데, AI 결정 과정에 대한 설명이 어려울 경우 시민들은 행정 결정에 대한 근

본적인 불신을 갖게 되며, 이는 AI 공공 서비스 수용을 낮추고 국가 정책 집행의 민주적 정당성을 훼손할 가능성이 높음

- (기존 논의의 한계) EU GDPR은 ‘자동화된 의사결정’에 대한 설명요구권을 반영하였으며, 우리나라에서도 ‘AI 기본법’ 등에 설명 및 고지 의무가 포함되는 등 법제화가 추진되고 있음. 그러나 법적으로 어디까지, 어떻게 설명해야 유의미한 설명인지에 대한 사회적 합의가 어려운 상황임. 또한 민간 기업이 개발한 AI 알고리즘을 공공부문에 도입할 경우, 기업이 로직 공개를 거부하면 정부조차 그 결정 과정을 설명하기 어렵다는 한계가 있음

자료: 연구진 작성

3. 이용자 분야 의제

〈표 5-3〉 이용자 분야 핵심의제

(이용자01) [모든 국민의 권리로서의 AI 리터러시 교육 보장] 디지털 취약계층을 포함한 모든 시민이 AI 리터러시 교육에 평등하게 접근할 수 있도록 제도적 의무와 체계적 지원을 강화하는 데 정부는 정책적 지원과 자금을 얼마나 투입해야 하는가?

- (제안의 배경) 생성형 AI의 일상 침투로 AI가 보편 기술로 자리 잡아가고 있지만, 그에 걸맞는 시민의 AI 리터러시는 부족함. 그동안 정부의 AI 관련 교육 정책은 개발자 양성에 치중되었고, 일반시민, 특히 디지털 취약계층을 위한 체계적 AI 리터러시 교육 기회는 현저히 부족한 상태임
- (현황 분석) OECD(2023) 조사 결과, 우리나라는 영국, 미국과 함께 AI 전문가 양성 교육에 정책 지원이 집중되는 국가 그룹에 속했으며, 일반 대중을 위한 AI 리터러시 교육은 매우 부족한 것으로 나타남
- (문제의 심각성) AI 기초 교육 부재는 AI 기술의 혜택이 특정 계층에만 집중되는 결과로 이어져 AI 역량 격차를 심화하고 새로운 불평등을 가중시킬 수 있음. 특히 AI 전환으로 인한 고용 악영향에 가장 취약한 저숙련 노동자 및 자동화 위험이 큰 직종의 종사자들이 AI 교육에 접근하기 어려워져 계층 간 경제적·사회적 격차가 더욱 고착화될 위험이 있음
- (기존 논의의 한계) 전국 단위의 정책사업을 하기에는 부족한 예산 투입과 AI·디지털 기술의 발전과 함께 축적되는 지속되는 결여·격차이며, 이를 위한 장기적 재정 계획이나 법적 근거를 논의할 필요가 있음. ‘모든 국민의 보편적 권리’라고 하였을 때 기준점을 일반 평균이 아니라 취약계층에 둬서 정책의 포용성을 담보해야 하는데, 취약계층의 지원은 별도로 생각하는 사회인식이 정책개발 논의에도 적용되어, 이러한 보편 권리 논의에 고령인·장애인을 배제하는 우를 범하기도 함

(이용자02) [AI 기반 서비스의 보편적 접근성을 위한 법·제도적 보장] 고령인·장애인의 AI 서비스 활용 권리와 자기결정권을 실질적으로 보장하기 위해, AI 기반 기술/서비스 개발 시 보편적 디자인 가이드라인 등을 법적으로 표준화하고 이를 관리·감독할 거버넌스를 어떻게 구축할 것인가?

- (제안의 배경) AI 기반 생활 편의 서비스가 확대될수록 디지털 접근성의 불평등이 새로운 사회적 배제 요인이 될 수 있음. 또한 이들을 AI 기술의 '사용자'라기보다 관리 혹은 감시의 '대상자'로 보고 서비스되는 기술 도구가 있으며, 결과적으로, 안전·보호를 이유로 한 디지털 추적·관찰 기능의 대상이 되어 기본권 침해로 이어질 수 있음
- (현황 분석) 기존 접근성 관련 법령은 AI 서비스의 특수성을 반영하지 못하고 있음. 공공기관 홈페이지 접근성 인증은 의무화되어 있으나 AI 기반 서비스에 대한 접근성 인증은 부재하며, 민간 서비스의 경우는 미약하거나 권고 수준에 그침. 또한 위반 시 제재 수단이 제한적이어서 실질적 개선 동력도 부족한 편임. 의도하지 않은 사회적 약자에 대한 대상화 이슈가 개별적으로 제기되나 사회문제로 논의되는 단계까지 나아가지 못함
- (문제의 심각성) AI 서비스 접근성 부족은 고령인과 장애인의 정보 접근권, 사회 참여권, 자기결정권을 침해할 수 있으며, 타인의 도움 없는 독립적 생활이 더 어려워질 수 있음. AI 기술/서비스의 디자인이나 정책이 일반 국민에게만 맞춰지거나 기존의 사회적 차별을 알고리즘 영역으로 넘겨버리는 악용 등이 발생할 수 있음. 접근성이 보장되지 않는 AI 서비스는 결국 대안 서비스 제공 비용, 격차 해소 비용 등 사회 전체의 비용을 증가시키게 됨
- (기존 논의의 한계) 현행 「지능정보화기본법」 등은 주로 정적인 웹사이트의 접근성을 다루고 있어서 복잡한 AI 서비스의 고유한 장벽(예. 이해하기 어려운 AI의 응답, 잘못된 음성인식 등)을 포괄하지 못함에도 개정 논의가 더딤. 또한 AI 기술의 빠른 발전에 따른 유연성에 강조점을 두는 동안, 고객 집단 비중이 낮은 장애인 등을 위한 기업의 자발적 참여 유도에는 취약성이 있음. 역동적인 변화 시기에도 고령인·장애인의 권리를 실질적으로 보장할 '최소한의 법적 표준'으로 사회적 논의가 진행되어야 함. 이와 관련한 인증 및 관리·감독 전담 기구나 거버넌스 구축 논의도 병행되어야 함

(이용자03) [AI 활용 경험과 연계된 체험형 윤리교육 설계·확산] 디지털 취약계층이 AI에 대한 과도한 의존이나 막연한 두려움 없이 AI를 비판적으로 평가하고 윤리적 문제를 인지할 수 있도록, 활용과 연계된 체험형 윤리교육을 어떻게 설계하고 보급할 것인가?

- (제안의 배경) AI 기술이 키오스크, 모바일 앱, 공공 웹사이트 등 일상 서비스 전반에 확산되고 있지만, 경험 및 학습 기회가 상대적으로 부족한 고령인들은 막연한 두려움 때문에 더 경험 및 학습 기회를 잃는 악순환에 있음. 이에 활용과 연계된 체험형 윤리교육 교재 개발 등 시범적 노력이 시도되고 있음
- (현황 분석) 고령인들의 AI·디지털 교육 수요가 체계적으로 파악되고 공유된 경우가 없어서 사회 전반의 인식이 부족함. 담당 부처 및 기관에서 시범적 노력이 있지만, 그 성과가 사회 전반으로 확산되어 활발하게 활용되기 위한 후속 조치가 부재함
- (문제의 심각성) 과거 선택 가능했던 오프라인 서비스 비중이 빠르게 감소하면서 고령인들이 경험하는 신기술 기반 서비스에 대한 두려움이 더 커지고, 그러한 심리적 저항감으로 인한 기술 격차 역시 과거 어느 때보다 심각해질 수 있음
- (기존 논의의 한계) 지금까지 디지털배움터 등의 정책사업으로 거주지 중심 접근성을 확보했다면, 앞으로는 유용한 학습 방법 개발에도 집중할 필요가 있음. 즉, 체험형·사례기반 교육 모델 및 교육기관 간 표준화된 커리큘럼 개발 논의로 확대될 필요

4. 정보 생태계 분야 의제

〈표 5-4〉 정보 생태계 분야 핵심의제

(정보01) AI 기반 허위 조작 정보 범람에 따른 정보 신뢰 체계 붕괴에 어떻게 대응할 것인가?

- (제안의 배경) 딥페이크, 생성형 콘텐츠 등 AI 기술의 발전으로 정보의 진위를 가려내기 점점 더 어려워지면서, 정보 생태계의 신뢰 기반이 근본적으로 흔들리고 있음
- (현황 분석) 정치·금융·상업 등 다양한 영역에서 AI를 활용한 허위조작정보가 대량 생산·유포되고 있음. 유포된 허위조작정보는 삭제나 정정이 어렵고, 피해 회복이 사실상 불가능한 경우도 많아 사회 전반에 상당한 피해를 초래
- (문제의 심각성) 정보 신뢰 붕괴는 민주주의의 절차적 정당성과 사회적 신뢰와 합의의 기반을 훼손하는 국가적 위기임. 시민의 인지적 피로를 가중시켜 공론장에 대한 냉소적 무관심을 초래하고, 이는 장기적으로 사회적 자본(신뢰)의 고갈 및 거대 경제적 비용으로 이어질 수 있음
- (기존 논의의 한계) 현행 규제는 전통적인 언론의 오보나 단순한 허위 사실 유포에 초점을 맞추어 AI 기반의 지능적·기술적 조작을 규율하기 어려움

(정보02) AI 알고리즘의 불투명성에 따른 정보 권력의 집중을 어떻게 막을 것인가?

- (제안의 배경) AI 알고리즘의 작동 방식이 블랙박스로 남아 있으며, 이는 AI의 공공성과 책임성 확보를 저해하는 주요 방해 요소
- (현황 분석) 많은 플랫폼에서 AI 알고리즘이 클릭 수·체류 시간 등 자극적 지표를 중심으로 콘텐츠 노출을 결정하면서 선정적·감정적 콘텐츠가 과도하게 확산되고, 저널리즘의 공정성이 훼손되고 있음
- (문제의 심각성) 알고리즘의 불투명성은 정보 권력의 비대칭성을 극단적으로 심화하고, 소수 플랫폼이 사회적 의제 설정 권력을 사실상 독점하게 만들 우려가 큼
- (기존 논의의 한계) 기술적으로 복잡한 AI 알고리즘의 작동 방식을 어떤 기준으로 규제·감독할지에 대한 명확한 규범이 부재. 기업들은 영업비밀을 이유로 알고리즘 정보 공개를 거부하고 있어, ‘알고리즘 접근권’ 등 정보주체의 권리 보장은 여전히 매우 미흡한 실정

(정보03) AI 추천 알고리즘이 초래하는 정보 편식과 사회적 양극화를 어떻게 완화할 것인가?

- (제안의 배경) 알고리즘 불투명성이 초래하는 사회적 결과에 논의의 초점을 두고, 공론장 기능 약화 및 사회 통합 저해라는 거시적 문제를 다루기 위해 독립 의제로 설정
- (현황 분석) 추천 알고리즘이 개인의 선호에 맞춘 정보만을 제공하는 ‘필터 버블’ 및 ‘에코 챔버’ 현상을 강화하여, 이용자가 다양한 관점에 노출될 기회를 박탈. 이로 인해 유사한 성향의 이용자들이 폐쇄적 공간에서 상호 확증하는 ‘집단 극단화’ 현상이 나타남
- (문제의 심각성) 정보 편식은 집단 극단화를 심화시켜 사회 구성원 간 이해와 관용을 저해하며, 민주주의의 핵심 기능인 건전한 공론장을 마비시키고 사회적 갈등 비용을 증대시키는 구조적 위험을 내포함

- (기존 논의의 한계) 개인의 자유로운 선택으로 포장된 알고리즘의 추천 행위를 제도적으로 규율하는 것에 대한 사회적 합의가 미흡. 또한, 정보 편식이 사회 전체에 미치는 부정적 파급 효과에 대한 실증적이고 계량적인 연구가 부족하여 정책 우선순위 설정에 한계 존재

자료: 연구진 작성

5. 지역 분야 의제

〈표 5-5〉 지역 분야 핵심의제

(지역01) 데이터센터의 지역 분산을 촉진하기 위한 정부의 역할은 무엇인가?

- (제안의 배경) 데이터센터는 AI 서비스·모델 학습의 핵심 인프라임. 국내 시설의 약 60%가 수도권에 집중되어 있어 연산 자원 접근성과 혁신 역량이 지역별로 불균등하게 축적될 가능성이 존재함
- (현황 분석) 비수도권은 부지·전력 공급 등 입지 조건이 유리하나 초저지연 네트워크 접근성, 전문 인력 확보, 운영 생태계 부족 등이 병목으로 작용함. 또한 소음·열·전력 수요 증가에 대한 주민 수용성 문제로 유치가 지연되는 사례가 존재함
- (문제의 심각성) 연산 인프라 집중이 지속될 경우 AI 개발·학습·배포까지 가치사슬 전체가 수도권에 고착될 가능성이 높으며, 비수도권은 산업·고용·생태계 축적이 제한될 우려가 있음
- (기존 논의의 한계) 부지·전력 등 물리적 인프라 중심 접근이 주류였으며 네트워크 품질·전문 인력·경제 파급효과 근거 확보 등 구조적 제약 해소가 미흡함

(지역02) AI 확산을 위한 지역별 에너지 인프라 확충 전략은 무엇인가?

- (제안의 배경) AI 확산으로 전력 수요가 급증하는 가운데 비수도권 발전·수도권 소비 구조만으로 산업 성장을 뒷받침하기 어려운 한계가 존재함. 이에 지역별 전원 구성 재설계 필요성이 제기됨
- (현황 분석) 비수도권은 재생에너지·대규모 발전 확장 가능성이 높고 수도권은 분산형 전원·ESS로 대응 가능함. 다만 환경 규제·경제성·송전망 투자·주민 수용성 등이 주요 제약요인으로 작용함
- (문제의 심각성) 수요 집중 지역의 병목이 누적될 경우 전력계통 안정성 저하, 연산 자원 부족, 데이터센터 투자 지연 등이 발생할 위험이 존재하며 이는 지역 간 산업·생산성 격차를 확대할 우려가 있음
- (기존 논의의 한계) 공급 확대 중심 접근이 우세했으며 지역별 산업 수요·전력망 운영·분산형 전원 유인 등 통합 설계가 미흡했음. 또한 주민 수용성 기반 의사소통 전략도 부족함

(지역03) 지역 특화 산업과 AI 융합을 통한 지역 경제 활성화 방안은 무엇인가?

- (제안의 배경) 지역은 상이한 산업 기반을 보유하고 있으며 AI 기술은 기존 산업 고도화를 통해 비수도권 성장 동력 확보에 기여할 수 있음. 이에 지역별 비교우위를 활용한 AI 융합 전략 필요성이 제기됨
- (현황 분석) 스마트팜·스마트팩토리 등 개별 사업은 존재하나 지역 기업의 디지털 역량 부족, 전문 인력 부족, 데이터·플랫폼 생태계 미비로 산업 전반 확산은 제한됨

(지역03) 지역 특화 산업과 AI 융합을 통한 지역 경제 활성화 방안은 무엇인가?

- (문제의 심각성) AI 전환 지연 시 산업 경쟁력 약화 → 투자 감소 → 고용 축소 → 인구 유출로 이어지는 악순환이 발생할 위험이 있으며, 수도권 대비 혁신 격차가 구조적으로 확대될 가능성이 있음
- (기존 논의의 한계) 개별 사업·공급자 중심 R&D 지원에 치중됨. 산업 연계형 데이터 인프라, 표준 기술 모델, 산학 네트워크 기반 확산 전략 등이 부족하며, 경제효과 측정 근거가 미흡함

(지역04) AI 스타트업 생태계의 지역 간 확산을 촉진하기 위한 전략은 무엇인가?

- (제안의 배경) AI 스타트업의 약 80%가 수도권에 집중되어 있어 혁신 역량이 특정 지역에 몰릴 가능성이 있으며, 지역 기반 창업 생태계 구축 필요성이 제기됨
- (현황 분석) 수도권은 VC·대학·기업 네트워크가 밀집해 있으나 비수도권은 투자·기술 검증·지원조직 부족으로 창업 및 성장 단계 전환이 제한됨. 이에 따라 인재·기술이 수도권으로 이동하는 구조가 누적되고 있음
- (문제의 심각성) 지역 기반 창업이 부재할 경우 지역 수요 기반 AI 서비스 개발이 지연되고 AI 확산 효과가 지역별로 불균등하게 나타날 우려가 있음
- (기존 논의의 한계) 창업 지원 프로그램은 존재하나 AI 특성 기반으로 데이터·테스트베드·산학 사업화·지역 수요 모델을 통합적으로 구축하는 전략이 부족함

(지역05) 지역 대학을 중심으로 한 효과적인 AI 인재 양성 방안은 무엇인가?

- (제안의 배경) AI 핵심 인재가 수도권에 집중되어 비수도권의 기술 도입과 산업 고도화가 제약되고 있으며, 지역 대학 기반 인재 양성과 지역 내 정착 유도 필요성이 커지고 있음
- (현황 분석) 일부 대학은 대응 중이나 다수 지역 대학은 교수진·연구 인프라·산학 네트워크가 부족해 교육이 실제 산업 수요와 연계되지 못하는 사례가 존재함
- (문제의 심각성) 졸업생이 수도권으로 이동할 경우 지역 내 인적자본이 축적되지 않고 대학 성과가 산업 혁신으로 환류되지 않아 지역 혁신 역량 격차가 확대될 우려가 있음
- (기존 논의의 한계) 교육 강화 위주 접근이 많았으며 취업·창업·연구 연계 정책이 부족했고 지역 간 협력·정착 유인 설계·효과 측정 근거가 미흡함

자료: 연구진 작성

제 2 절 정책적 시사점

AI 확산은 산업 구조, 행정 운영, 정보 이용, 지역사회 등 사회 전 영역의 규칙과 균형을 근본적으로 재편하고 있다. 이러한 변화는 기존 제도 설계가 전제해 왔던 사회적 합의, 권리 구조, 분배 방식이 더 이상 온전히 작동하기 어려운 환경을 마련한다. 지금 이 시점에서 사회적 의제를 식별하고 핵심 질문을 제기하는 이유는, 단순히 위험을 진단하기 위함이 아니라, AI 전환이 불러올 구조적 충격 속에서 ‘어떻게 해야 모두가 함께 변화에 적응하고, 포용적이고 지속가능한 사회로 전환할 수 있는가’라는 사회적 당면과제를 보다 명확히 정의하기 위해서이다.

AI로 인한 변화는 특정 계층의 문제나 기술 정책의 문제가 아니라, 사회 전체의 상생과 포용을 좌우하는 구조적 문제이며, 대응을 늦출 경우 격차·갈등·불신의 심화로 이어질 위험이 크다. 본 연구에서 도출된 5대 분야 의제와 전문가 검증 결과는 이러한 맥락에서 AI의 위험과 기회를 균형 있게 관리하고, 사회 구성원 모두가 이익을 공유할 수 있도록 제도적 토대를 재설계해야 함을 보여준다. 특히 노동권 보호, 행정서비스의 책임성과 투명성, 이용자 권리 확보, 신뢰 가능한 정보 생태계, 지역의 지속가능성 강화는 AI 시대 상생과 포용의 조건이며, 각 분야별 선제적 대응 없이는 사회 전체의 지속가능성이 위협받을 수 있다. 이에 본 절에서는 각 분야의 핵심 의제가 제시하는 정책적 시사점을 정리하고, 향후 제도 설계를 위한 방향을 제시하고자 한다.

노동 분야에서 제시된 핵심 의제는 AI 전환으로 인한 직무 재편, 청년층의 경력 붕괴, 전환 비용의 불평등한 부담, 알고리즘 매니지먼트 확산, AI 초과이익의 집중 등 노동시장의 구조적 불안정과 분배 문제에 초점을 둔다. 위와 같은 핵심 의제를 다루기 위한 정책은 단기 대책을 넘어 직업훈련·전환지원·노동권 보호·분배 체계 등 노동시장 전반의 시스템 재설계가 요구된다. 구체적으로는 직무 소멸 이전에 선제적으로 개입하는 상시적·예방적 재교육 체계 구축, 청년층을 위한 새로운 ‘초급 경력 경로’ 창출, 전환 비용의 사회적 부담 원칙 확립, 알고리즘 기반 통

제에 대한 투명성·설명권·데이터 이동권 등 노동자의 디지털 권리 보장, 그리고 AI 생산성 이익의 사회적 공유 메커니즘 마련이 핵심이다. 이러한 정책이 실현될 경우, 기술 변화의 속도가 격차와 배제를 심화하는 경로가 아니라 노동자의 역량 회복·경력 지속성·공정한 전환·노동권 존중·포용적 분배 구조로 이어지는 긍정적 전환 효과를 기대할 수 있다. 결국 노동 분야의 의제는 AI 시대에 모두가 일할 권리, 학습할 기회, 전환을 견딜 안전망을 갖춘 상생·포용적 노동시장을 구축하기 위한 핵심 토대라 할 수 있다.

공공행정 분야에서 제시된 핵심의제는 AI 시대 기술적 안정성 및 개인정보 보호의 취약성 심화 문제, 접근성 격차 및 알고리즘 편향에 따른 디지털 취약계층 소외 문제, AI 기반 공공·행정서비스 오류 발생 시 책임 공백 해소 방안 부재 문제, 첨단 감시 기술 확산에 따른 ‘빅브라더’ 사회 귀결 위험 문제, AI 정책 결정 과정에 대한 시민의 납득 가능한 설명 제공 가능성 문제이다. 각각의 핵심의제를 다루기 위한 대응 방향과 정책 방안은 다음과 같다. 먼저 AI 시대의 기술적 안정성과 개인정보 보호 취약성에 대해서는, AI 처리 과정의 위험을 체계적으로 진단·관리할 수 있는 위험관리체계 구축이 필요하다. 아울러 AI 특성을 반영한 개인정보 보호 법·지침을 정비하고, 개인정보 처리 과정과 데이터 생태를 검증할 전문기관 설치, 비식별·품질관리 기준 강화 등을 통해 개인정보 보호 기반을 한층 강화해야 한다. 디지털 취약계층의 접근성 격차와 알고리즘 편향 문제에 대해서는 개별 특성에 맞춘 교육·훈련 프로그램의 제공 경로를 확대하고, 공공 AI 서비스 개발 시 접근성 기준·쉬운 언어·음성 인터페이스 등 기본 요소를 의무화할 필요가 있다. 또한 취약계층 테스트 그룹 운영 및 지원 인력 배치 강화, 파편적 규정을 통합한 종합 보호 법제 마련이 요구된다. AI 기반 공공서비스 오류 시 책임 공백 문제는 생애주기 단계별 책임 주체를 명확히 하고, 고위험 분야에 대한 사전검증 제도를 강화해야 한다. 더불어 입증책임 전환, 집단소송, 징벌적 배상, 전문 ombudsman 등 실효적 피해구제 장치를 마련할 필요가 있다. 첨단 감시 기술 확산으로 인한 ‘빅브라더’ 위험을 완화하기 위해서는 기술 도입·고도화 시 사전 영향평가를

의무화하고, 데이터 주권을 실질적으로 구현할 수 있는 법적 원칙을 강화해야 한다. 또한 독립적 감시 기구 설치, 데이터 분산관리 구조 도입, 익명화·보호기술 고도화 등이 요구된다. 마지막으로 AI 정책 결정 과정의 설명가능성 문제는 시민 수준에 따른 설명 기준을 정립하고, 자동화 여부·모델명·판단 기준·이의제기 방법 등을 포함한 표준 결정서를 마련해야 한다. 특히 공공 조달 시 설명 가능성 조항 반영, 알고리즘 에스스로 도입 등 투명성 확보 장치 강화가 필요하다.

이용자 분야에서는 인구의 다수를 차지하는 보편 시민이 아니라, 상대적으로 취약한 계층에 초점을 두었다. 이는 현재 정부 정책이 주로 ‘평균적 시민’을 대상으로 설계되는 현실에 대한 문제의식에서 출발하였다. 이 경우 고령인, 장애인 같은 전통적 취약계층은 정책 대상에 배제되고 이들의 디지털 격차는 더 심화될 수 있다고 본 것이다. 같은 맥락으로 이번 연구의 전문가 인터뷰에서 시민 대상 AI 리터러시 정책의 기본 기준을 낮추어 추진할 경우, 디지털 포용을 동시에 실현할 수 있다는 의견이 있었다. 이용자 분야 핵심 의제는 AI 리터러시 교육을 국민의 권리로 제도화, AI 기반 서비스 접근성의 보편적 보장을 위한 법·제도 개선, 고령층 맞춤형 AI·디지털 교육 방안을 개발·확산할 필요성이다. 일정 수준의 학력과 소득을 갖춘 은퇴자라 하더라도 새로운 기술을 활용할 기회가 부족해 디지털 격차를 경험하는 사례가 적지 않다. 특히 직장 생활에서 디지털 기술을 활용해 왔던 이들도 은퇴 이후에는 적절한 학습 기회와 동력을 얻기 어려운 현실이다. 이에 실생활 경험 기반의 학습 환경을 구축하자는 제안이 제기되었고, 이는 이미 일부 교육 현장에서 가상 현실 기반 프로그램 등으로 시도되고 있다. 앞으로 정책 대상은 ‘디지털을 전혀 모르는 고령층’이 아니라, ‘적합한 학습 방식에 접근하지 못한 이들’로 재정의할 필요가 있으며, 은퇴 후 삶의 리듬에 맞춘 새로운 교육 시나리오와 사업 설계에 대한 사회적 논의가 요구된다.

정보생태계 분야의 핵심 의제를 다루기 위해서는, 플랫폼과 AI 사업자에 대한 책임성과 투명성 의무를 축으로 한 규범·제도 정비에 우선적으로 추진될 필요가 있다. 구체적으로는 「AI 기본법」 등에 따라 AI가 생성해내는 결과물에 대한 표시

제도가 실효성 있게 집행될 수 있도록 표시 체계를 구체적으로 설계하고 표준화하여 AI 투명성 확보에 힘써야 한다. 또한 AI 허위조작정보에 대한 추적 및 피해 구제 체계를 구축해야 한다. 아울러 주요 플랫폼·서비스에 대해 알고리즘 영향평가 및 외부 감사, 핵심 지표·기준 공개 등 알고리즘 투명성·설명 가능성 의무를 부과해야 한다. 나아가 알고리즘이 공적 정보나 다양한 관점을 일정 비율 이상 노출하도록 하는 공공성 기준을 마련하고, 공영·공익 플랫폼 및 다양한 정보원에 대한 접근을 지원함으로써, 필터 버블·에코 챔버 현상을 완화하고 균형 잡힌 정보환경을 설계하는 방향으로 정책을 패키지화할 필요가 있을 것이다.

지역 분야에서 제시된 핵심 의제는 데이터센터·전력망 등 AI 인프라의 공간적 배치, 지역 산업의 AI 전환, 지역 기반 창업 생태계 형성, 지역 대학을 통한 인재 정착 등 AI 확산 과정에서 발생하는 지리적 불균형과 혁신 격차에 초점을 둔다. 이에 따라 정책은 물리적 인프라 확충만이 아니라, 네트워크 품질·전문 인력·산학 연계 등 구조적 병목을 해소하는 방향으로 설계될 필요가 있으며, 지역별 에너지 공급 체계 재편, 분산형 데이터센터 특화 단지 조성, 지역 주력산업 맞춤형 AI 실증·플랫폼 구축, 지역 수요 기반 AI 창업 지원, 지역 대학과 산업을 연결하는 정착형 인재 양성 체계 등이 요구된다. 이러한 정책이 실현될 경우 AI 인프라와 혁신 역량이 수도권에만 집중되지 않고 전국적으로 확산되며, 지역 산업 경쟁력 향상·신규 일자리 창출·비수도권 성장동력 확보로 이어지는 지역 균형 발전 효과를 기대할 수 있다. 궁극적으로 지역 분야 의제는 AI 시대에 모든 지역이 인프라 접근권, 산업 성장 기회, 창업·고용 생태계를 공유하는 포용적 혁신 기반을 구축하기 위한 핵심 토대라 할 수 있다.

참 고 문 헌

1장

[해외 문헌]

Acemoglu, Daron, and Pascual Restrepo(2019). “Automation and New Tasks: How Technology Displaces and Reinstates Labor.” *Journal of Economic Perspectives*. 33(2), pp. 3-30.

Bresnahan, T. F. & Trajtenberg, M.(1995). “General purpose technologies ‘Engines of growth’?” *Journal of Econometrics*. 65, pp. 83-108.

Brynjolfsson, E. & McAfee, A.(2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W. W. Norton & Company.

Erik Brynjolfsson, Danielle Li, and Lindsey Raymond(2025). “Generative AI at Work.” *The Quarterly Journal of Economics*. Volume 140, Issue 2, May, pp. 889-942.

European Commission(E.C)(2015). “GREEN PAPER on retail financial services Better products, more choice, and greater opportunities for consumers and businesses, COM/2015/630 final.”

_____(2021). “GREEN PAPER ON AGEING Fostering solidarity and responsibility between generations, COM/2021/50 final.”

Korinek, A. & Stiglitz, J. E.(2017). “Artificial intelligence and its implications for income distribution and unemployment.” NBER Working Paper No. 24174.

Mauro Cazzaniga, Florence Jaumotte, Longji Li, Giovanni Melina,

Augustus J. Panton, Carlo Pizzinelli, Emma J. Rockall, and Marina Mendes Tavares(2024). “Gen-AI: Artificial Intelligence and the Future of Work.” Staff Discussion Notes 2024, 001.

OECD(2024). “Governing with Artificial Intelligence: Are governments ready?” OECD Artificial Intelligence Papers. No. 20, OECD Publishing, Paris,

UK BEIS(2017). “Building our Industrial Strategy: Green Paper.” Department for Business, Energy & Industrial Strategy, UK Government.

2장

[국문 문헌]

《무등일보》(2025. 5. 7), “1천436개 질문이 정책으로…‘녹서’에 담긴 일자리 해법”, <https://m.mdilbo.com/detail/NSWCCT/742380> (검색일: 2025. 8. 13).

민주연구원 집단지성센터(2025), 『녹서 2025(Green Paper 2025)』, 민주연구원.
박찬수 외(2024), 『스마트생산 열린혁신랩 운영 및 연구사업(5차년도)』, 과학기술정책연구원.

《아이뉴스24》(2021. 1. 11), “[박태웅 칼럼]눈을 떠보니 선진국이 돼 있었다”, <https://www.inews24.com/view/1333621> (검색일: 2025. 8. 16).
정운찬(2018. 11. 19.), “독일의 ‘노동 4.0’ 정책이 주는 가르침”, <https://www.khan.co.kr/article/201811192113005> (검색일: 2025. 8. 11).

정정길 외(2023), 『정책학원론』, 대명출판사.

채준호 외(2025), 『광산구 지속가능한 일자리를 위한 녹서』, 광산구.

하승태·조의현(2008), “중요한 사회적 의제(MIP)에 대한 공적 합의”, 《한국언론정

보학보》, 통권 41호, pp. 41-74.
홍인기(2024. 8. 7), “상상과 질문”,
https://www.khan.co.kr/article/202408072042005/?utm_source=urlCopy&utm_medium=social&utm_campaign=sharing (검색일: 2025. 8. 11).
NRC 국가의제연구단(2023), 『2023 국가의제 종합연구』, 경제·인문사회연구회.

[해외문헌]

Cromarty, H.(2023). *Social housing reform in England: What next?* House of Commons Library.
European Commission(2021). “Green Paper on Ageing: Fostering solidarity and responsibility between generations. European Commission.” <https://op.europa.eu/en/publication-detail/-/publication/d918b520-63a9-11eb-aeb5-01aa75ed71a1/language-en>.
Federal Ministry of Labour and Social Affairs of Germany(2015). *Re-Imagining Work: Green Paper Work 4.0. Germany*.
HKU Business School(2025). *Hong Kong Economic Policy Green Paper 2025*. HKU Business School.
Kingdon, J. W.(1984). *Agendas, Alternatives, and Public Policies*. Boston: Little, Brown & Co.
McCombs, M. E. & Shaw, D. L.(1972). “The Agenda-Setting Function of Mass Media.” *The Public Opinion Quarterly*. 36(2), pp. 176-187.
Ministry of Housing, Communities & Local Government(2018). *A New Deal for Social Housing*. Ministry of Housing, Communities & Local Government.
NTIA(2017). *Fostering the Advancement of the Internet of Things*. U.S. Department of Commerce, National Telecommunications and

Information Administration.

3장 1절

[국문 문헌]

- 신승윤·윤보성·진희승(2025), “SW 개발자 채용시장의 변화와 생성형 AI의 영향”,
《소프트웨어정책연구소 이슈리포트 IS-193》, 소프트웨어정책연구소.
- 문아람 외(2024a), 『ICT 기반 사회현안 해결방안 연구』, 정보통신정책연구원.
- _____ (2024b), 『디지털 대전환 메가트렌드 미래전략 보고서』, 정보통신정책연구원.
- 한진수·오삼일(2025), “AI 확산과 청년고용 위축: 연공편향(seniority-biased) 기술변화를 중심으로”, 《BOK 이슈노트》, 제2025-30호, 한국은행.

[해외 문헌]

- Aum, S. & Shin, Y.(2025). “The Labor Market Impact of Digital Technologies.” NBER Working Paper No. w33469. National Bureau of Economic Research.
- Brynjolfsson, E., Rock, D., & Syverson, C.(2017). “Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics(No. w24001).” National Bureau of Economic Research.
- Brynjolfsson, E., Rock, D., & Syverson, C.(2021). “The productivity J-curve: How intangibles complement general purpose technologies.” *American Economic Journal: Macroeconomics*. 13(1), pp. 333-372.
- Cho, W., F. Manaresi, and M. Reinhard(2025). “Superstars, shooting stars, and the labour share: Cross-country evidence.” OECD Science, Technology and Industry Working Papers, No. 2025/01. OECD Publishing, Paris.

- Droz, L. & Tavares, M. M.(2024). “Generative AI: A Turning Point for Labor’s Share?” *Economic Insights*. 9(1), pp. 2-11. Federal Reserve Bank of Philadelphia.
- Dominski, J. & Lee, Y. S.(2025). “Advancing ai capabilities and evolving labor outcomes.” arXiv preprint arXiv:2507.08244.
- Felten, E. W., Raj, M., & Seamans, R.(2023). “How will Language Modelers like ChatGPT Affect Occupations and Industries?” arXiv preprint arXiv:2303.01157.
- Hampole, M.(2025). “Artificial Intelligence and the Labor Market.” NBER Working Paper No. w33509.
- Handa et al.(2025). “Which Economic Tasks are Performed with AI? Evidence from Millions of Claude Conversations.” arXiv:2503.04761.
- Henseke, G., Davies, R., Felstead, A., Gallie, D., Green, F., & Zhou, Y.(2025). “How Exposed Are UK Jobs to Generative AI? Developing and Applying a Novel Task-Based Index.” arXiv.
- Humlum, Anders and Vestergaard, Emilie(2025. 4. 15). “Large Language Models, Small Labor Market Effects.” University of Chicago, Becker Friedman Institute for Economics Working Paper No. 2025-56.
- Kim, D. G. & Moon, A.(2024). “From helping hand to stumbling block: the ChatGPT paradox in competency experiment.” *Applied Economics Letters*. 1-5.
- Shakked Noy, Whitney Zhang(2023). “Experimental evidence on the productivity effects of generative artificial intelligence.” *Science*. 381, pp. 187-192.
- Pizzinelli, C., Panton, A., Tavares, M. M., Cazzaniga, M., & Li, L.(2023).

“Labor Market Exposure to AI: Cross-country Differences and Distributional Implications.” IMF Working Paper 23/216.

Webb, M.(2019). “The Impact of Artificial Intelligence on the Labor Market.” Stanford University Working Paper.

3장 2절

[국내 문헌]

곽영·신민수(2021), “스마트워크 도입 요인과 사용자 혁신저항 및 이직의도 간의 관계에 대한 연구: 조직통제 조절효과를 중심으로”, 《경영정보학연구》, 23(4), pp. 23-43.

경제·인문사회연구회(2024), “초거대 AI를 이용한 통합 연구자원 생성, 관리 서비스 개발”, 《미래정책 포커스》, 2024 가을호,

https://www.nrc.re.kr/board.es?act=view&bid=0044&issue_cd=42&list_no=178745&mid=a30200000000&nPage=1&tag=.

김원제·이창주·하연희·조항민(2009), “위험커뮤니케이션의 구성요인 및 요인 간 영향관계에 관한 탐색 연구: SMCRE 모델 적용을 통한 원자력 위험커뮤니케이션의 구조분석을 중심으로”, 《스피치와 커뮤니케이션》, 11, pp. 80-123.

디지털플랫폼정부위원회(2025), 『공공 AI 서비스 실증 사례집』.

디지털플랫폼정부위원회 보도자료(2025. 2. 26), “공공부문 초거대 AI 도입·활용 가이드라인 2.0 발표”,

<https://www.dpg.go.kr/DPG/contents/DPG02020000.do?schM=view&id=20250226110028349393&schBcid=press>.

문화체육관광부 국립중앙도서관 사서지원서비스(2025. 3. 7), “스마트도서관의 전담 독서 컨설턴트-대만국립공공정보도서관 ‘AI 도서 추천’ 서비스”, <https://librarian.nl.go.kr/LI/contents/L30201000000.do?schM=view>

&id=50552&schBcid=BBSMSTR_451.

박천희·홍은영(2023), “디지털 전환에 대한 위험 인식 연구: 일자리 부문의 마지못한 수용성 개념을 중심으로”, 《정보사회와 미디어》, 24(2), pp. 79-124.

류현숙·홍승희(2015), “소셜미디어를 통한 위험인식과 위험소통에 관한 탐색적 연구”, 《Crisisonomy》, 11(8), pp. 93-116.

서울관광재단(2025. 2. 24), “(스마트관광팀) 비지트서울(Visit Seoul) 생성형 AI 활용 서울관광 특화 ‘여행 플래너’ 서비스 시범 오픈”,

https://www.sto.or.kr/press/%EC%8A%A4%EB%A7%88%ED%8A%B8%EA%B4%80%EA%B4%91%ED%8C%80-%EB%B9%84%EC%A7%93%EC%84%9C%EC%9A%B8visit-seoul-생성형-ai-활용-서울관광-특화-‘여행-플래너’-서비스-시범-오픈_/15163.

윤수경·김명지·최준호(2014), “혁신특성과 사용자특성이 전자책 수용에 미치는 영향: 혁신저항모형을 중심으로: 혁신저항모형을 중심으로”, 《한국콘텐츠학회논문지》, 14(8), pp. 61-73.

유진(2024. 5. 6), “인공지능 줌인: 크라임GPT '범죄와의 전쟁'에서 획기적 역할”, 《위키리스크한국》, <http://www.wikileaks-kr.org/news/articleView.html?idxno=152275>.

이진명·박성지·전민선(2025), “고령자 기술수용모델(STAM)을 적용한 돌봄로봇 수용의도 결정요인 분석”, 《소비자학연구》, 36(4), pp. 145-167.

고대영·이나영(2024. 10. 11), “[초고령사회, 일본에 길을 묻다] ② 치매예방 말벗·스마트 양로원…‘AI 돌봄 도우미’에서 길 찾다”, 《이투데이》, <https://www.etoday.co.kr/news/view/2408280>.

한국지역정보개발원(2023), “현장 리포트 서울시 인공지능상담사 챗봇 ‘서울톡’”, 《KLID 웹진》, Vol. 139, http://klidwz.or.kr/webzine/vol139/sub_2_1_01.html.

_____ (2024), “생성형 AI 서비스 최근 동향: 멀티모달 서비스를 중심으로”, 《KLID 디지털 이슈 리포트》, 24-5호.

- 홍은영·문정욱·박천희(2023), “디지털 전환 시대의 정보 격차 개념에 관한 비판적 논의: 지능정보기술 격차 모형의 제안 및 실증 연구”, 《한국행정학보》, 57(3), pp. 291-326.
- 홍은영·박천희(2023), “디지털 전환 시대, 새로운 위험에 대한 대응행동 의도 결정 요인 분석: 위험 지각 요인(두려움, 주관적 지식)의 이중매개효과를 중심으로”, 《한국지역정보학회지》, 26(2), pp. 27-56.
- 박천희·홍은영·문정욱(2021), “지능정보기술의 이용의도 결정 메커니즘에 관한 연구: 혁신저항과 리터러시의 기능을 중심으로”, 《한국정책학회보》, 30(2), pp. 317-348.
- Ang, A.(2024. 1. 24), “싱가포르의 공공보건 분야에서 인공지능 도입의 배경”, 《Healthcare IT News》,
<https://www.healthcareitnews.com/news/asia/behind-singapores-widespread-ai-adoption-public-health>.
- 길태윤(2025. 8. 21), “KOTRA, AI로 중소·중견기업 수출길 넓힌다”, 《COSMETIC INSIGHT》,
<https://www.cosinkorea.com/news/article.html?no=55911>.
- Gaw, B.(2024. 4. 29), “싱가포르의 인공지능과 생명과학: 융합과 규제 동향”, 국제변호사협회,
<https://www.ibanet.org/artificial-intelligence-life-sciences-singapore>.
- Goover(2025), “서울시 ‘챗봇 2.0’으로 본 AI 행정 민원 혁신 전략”, 《Goover 리포트》,
<https://seo.goover.ai/report/202508/go-public-report-ko-5db1943a-c3b5-40ab-84fe-c103a12c6aeb-0-0.html>.
- KBS(2025. 4. 9), “일본, ‘AI와 주고받는 대화’…치매 예방 실험”, KBS 뉴스,
<https://news.kbs.co.kr/news/pc/view/view.do?ncd=8222719>.

[해외문헌]

- Ader, C. R.(1995). “A longitudinal study of agenda setting for the issue of environmental pollution.” *Journalism & Mass Communication Quarterly*. 72(2), pp. 300-311.
- Ajzen, I. & Fishbein, M.(1980). *Understanding attitudes and predicting social behavior*. Englewood cliffs, NJ: Prentice-Hall.
- Ajzen, I.(1991). “The theory of planned behavior.” *Organizational behavior and human decision processes*. 50(2), pp. 179-211.
- BBC News(2024. 2. 15). “Vehicle-based littering solution pilot achieves 95% detection rate.”
<https://www.bbc.com/news/uk-england-leicestershire-68575556>.
- Ashe, Isaac(2024. 5. 17). “AI camera system caught litter louts in Loughborough, AI camera system caught litter louts in Loughborough.” *BBC News*.
- Chen, K. & Chan, A. H. S.(2014). “A review of technology acceptance by older adults.” *Gerontechnology*. 13(1), pp. 1-10.
- Eagly, A. H. & Chaiken, S.(1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt.
- Poortinga, W. & Pidgeon, N. F.(2003). “Exploring the dimensionality of trust in risk regulation.” *Risk Analysis: An International Journal*. 23(5), pp. 961-972.
- _____(2006). “Prior attitudes, salient value similarity, and dimensionality: Toward an integrative model of trust in risk regulation.” *Journal of Applied Social Psychology*. 36(7), pp. 1674-1700.
- Prague Morning(2024. 11. 10). “New AI-Powered Traffic Lights to

- Improve Prague's Streets.”
<https://praguemorning.cz/new-ai-powered-traffic-lights-toimprove-pragues-streets/>.
- McLuhan, M.(1964). *Understanding media: The extensions of man*. NY: McGraw Hill.
- Ministry of Health(2021). “Artificial Intelligence in Healthcare Guidelines.”
[https://www.moh.gov.sg/docs/librariesprovider5/eguides/1-0-artificial-in-healthcare-guidelines-\(aihgle\)_publishedoct21.pdf](https://www.moh.gov.sg/docs/librariesprovider5/eguides/1-0-artificial-in-healthcare-guidelines-(aihgle)_publishedoct21.pdf).
- New South Wales Government(2024). “NSW Planning wants AI product solutions to make faster development applications.” Department of Planning, Housing and Infrastructure.
<https://www.nsw.gov.au/media-releases/nsw-planningwants-ai-product-solutions-to-make-faster-development-applications>.
- Ram, S.(1987). “A Model of Innovation Resistance.” *Advances in Consumer Research*. 14(1), pp. 208-212.
- Rogers, E. M.(2003). *Diffusion of innovations(5th Ed.)*. New York: Free Press.
- Schaefer, T. M. & Birkland, T. A.(2007). *Encyclopedia of media and politics*. Cq Pr.
- Seidl, R. & Drögemüller, C.(2024). “Procedural fairness and safety in the acceptance of nuclear waste disposal in Germany: an empirical study.” *Journal of Risk Research*. 27(8), pp. 969-985.
- Sheth, J. N.(1981). “Psychology of Innovation Resistance: The Less Developed Concept(LDC) in Diffusion Research.” *Research in Marketing*. 4, pp. 273-282.
- Trumbo, C. W. & McComas, K. A.(2003). “The Function of Credibility in

- Information Processing for Risk Perception.” *Risk Analysis*. 23, pp. 343-353.
- Tversky, A. & Kahneman, D. A.(1974). “Judgment under uncertainty.” *Science*. 185, pp. 1124-1131.
- Venkatesh, V. & Davis, F. D.(2000). “A theoretical extension of the technology acceptance model: Four longitudinal field studies.” *Management Science*. 46(2), pp. 186-204.
- Xiaoqi, Xu & Zainal, Siti(2025). “Relative advantage, technology innovativeness: Digital transformation acceptance attitude and employee creativity.” *Environment and Social Psychology*. 10(3).

[기타]

국립세종도서관 정책정보포털.

<https://policy.nl.go.kr/pages/information/guide.jsp>.

3장 3절

[국내 문헌]

- 교육부(2022. 12), 『2022 개정 교육과정』.
- 국가인권위원회(2022), 『디지털 격차로 인하나 노인의 인권상황 실태조사』, 국가인권위원회.
- 권정민·이영선(2020), “장애인을 위한 인공지능 활용 동향”, 《한국초등교육》, 제31권 특별호, 서울교육대학교, pp. 187-202.
- 김정근(2021), “코로나19 팬데믹 시대 미국의 AI/로봇을 활용한 노인 돌봄 사례와 이슈”, 《국제사회보장리뷰》, Vol. 16, pp. 16-26.
- 이유미(2022), “디지털 시대 새로운 패러다임과 리터러시: 디지털 리터러시와 AI

- 리터러시를 중심으로,” 《교양학연구》, 20집, pp. 30-60.
- 《조선일보》(2025. 8. 14), ““소금 대신 먹어” 챗GPT 말 들었다가… 정신병원 입원한 60대, 왜?”,
https://www.chosun.com/international/international_general/2025/08/15/OCKRHSNLVBADNJBZKASFU7BPVE/.
- 조성은 외(2019), 『ICT 해결방안 연구: 디지털 포용 정책을 중심으로』, 정보통신정책연구원.
- 최숙영(2022), “AI 리터러시 프레임워크에 대한 연구”, 《한국컴퓨터교육학회지》, pp. 73-84.
- 황현정·박지수·김승완(2023), “국내외 AI 리터러시의 연구 동향 분석: 체계적 문헌 고찰 방법을 중심으로”, 《교육문화연구》, 29권 3호, pp. 135-164.
- 《한겨레신문》(2023. 6. 14), “청각장애인 기사가 운전하는 차량 공유 서비스”,
<https://www.hani.co.kr/arti/science/future/1095905.html>.

[해외문헌]

- American library association(1989). *Presidential committee on information literacy: Final report*.
- Gilster, P.(1997). *Chapt. 1. Literacy for the Internet age in Digital Literacy*. Wiley Computer Publishing.
<https://archive.org/details/digitalliteracy00gils/page/n15/mode/2up>.
- Lee, C. J.(2011). “Myths about critical literacy: What teachers need to unlearn.” *Journal of Language and Literacy Education*. [Online] 7(1), pp. 95-102. 원출처: Freire, P.(1984). *Pedagogy of the oppressed*. New York: Continuum.
- Long, D. & Magerko, B.(2020). “What Is AI Literacy? Competencies and Design Considerations.” In *Proceedings of the 2020 CHI*

- Conference on Human Factors in Computing Systems. Association for Computing Machinery, pp. 1-16.
- OECD(2000). *Measuring student knowledge and skills: The PISA 2000 assessment of reading, mathematical and scientific literacy*. OECD publishing.
- _____(2019). *OECD Principles on Artificial Intelligence*.
- _____(2021). *OECD digital education outlook: Pushing the frontiers with artificial intelligence, blockchain and robots*. OECD publishing.
- _____(2023). *The state of implementation of the OECD AI principles four years on*. OECD publishing.
- _____(2025). *Bridging the AI skills gap: Is training keeping up?* OECD publishing.
- OECD & EC(2025). *New AI literacy framework to equip youth in an age of AI*. OECD publishing.
- UNESCO(2005). *Education for all: literacy for life; EFA global monitoring report, summary*. UNESCO.
- _____(2011). *Media and information literacy curriculum for teachers*. United Nations Educational Scientific and Cultural Organization.
- _____(2021). *AI and education: Guidance for policy-makers*. United Nations Educational Scientific and Cultural Organization.
- _____(2024). *AI Competency Framework for Students*. United Nations Educational Scientific and Cultural Organization.

3장 4절

[국내 문헌]

김형지·정선아(2025), “인공지능기본법 통과에 따른 표시의무 도입과 국내 입법 동향”, 《리걸타임즈》, 통권 제193호.

문정욱 외(2024), 『AI 윤리·신뢰성 확보를 위한 실천 방안 및 정책연구』, 과학기술정보통신부·정보통신정책연구원.

박상철(2024), “인공지능 기본법의 시행 전 개정 필요성 - 규제 조항의 체계·축조상 문제점을 중심으로-”, 《정보법학》, 제28권 제3호, 한국정보법학회.

방송통신위원회·정보통신정책연구원(2025), “2024 지능정보사회 이용자 패널조사”.

손영선(2024), “생성형 인공지능과 개인정보 보호 - 헌법 및 개인정보 보호법적 검토 -”, 《법학연구》, 제34권 제2호, 연세대학교 법학연구원.

심나리(2023), “[미국] 아티스트 3인, 이미지 생성 AI의 저작권 침해를 주장하며 소송 제기”, 《저작권 동향》, 제2호, 한국저작권위원회.

엄주희(2025), “인공지능 딥페이크와 인권 그리고 헌법적 구제”, 《법학연구》, 제35권 제1호, 연세대학교 법학연구원.

이경선(2023), “생성형 AI가 가져올 미래와 도전과제”, 《KISDI Perspectives》.

이철남(2025), “[미국] 디즈니와 유니버설, 이미지 생성 AI ‘미드저니’를 상대로 저작권 침해 소송 제기”, 《저작권 동향》, 제7호, 한국저작권위원회.

이현경 외(2025), “ICT기반 사회현안 해결방안 연구”, 과학기술정보통신부·정보통신정책연구원.

정찬모(2025), “인공지능기본법 주요 내용의 검토와 향후 과제”, 《법학연구》, 제28집 제1호, 인하대학교 법학연구소.

《전자신문》(2025. 5. 13), “AWS, “韓 기업 94%가 생성형 AI 도입…63%는 CAIO 선임””, <https://www.etnews.com/20250513000185>.

《연합뉴스》(2016. 3. 25), “인공지능 세뇌의 위험…MS 채팅봇 테이 차별발언으로 운영 중단”, <https://www.yna.co.kr/view/AKR20160325010151091>.

- 《연합뉴스》(2021. 1. 11), “성희롱 혐오 논란에 3주만에 멈춘 이루다 AI 윤리 숙제 남기다”, <https://www.yna.co.kr/view/AKR20210111155153017>.
- _____(2025.7.13.). “머스크의 xAI, 그록의 히틀러 소환에 진심사과 ... 시스템 수정”, <https://www.yna.co.kr/view/AKR20250713004100091>.
- 《이코노미스트》(2017. 1. 2), “인공지능 기술은 생명에 대한 모독”, <https://economist.co.kr/article/view/ecn201701020028>.
- 《뉴스워커》(2025. 4. 2), “지브리 프사 열풍에 화난 미야자키 하야오... ‘맥도날드 너 마저’”, <https://www.newsworker.co.kr/news/articleView.html?idxno=373862>.
- 《IT동아》(2024. 7. 8), “메타, ‘AI 사진’ 꼬리표에 사진작가들 불만 ... AI 도구 확산에 혼란 가중”, <https://www.donga.com/news/It/article/all/20240708/125831564/1>.
- 심나리(2023), “[미국] 아티스트 3인, 이미지 생성 AI의 저작권 침해를 주장하며 소송 제기”, 《저작권 동향》, 제2호, 한국저작권위원회.
- 이철남(2025), “[미국] 디즈니와 유니버설, 이미지 생성 AI ‘미드저니’를 상대로 저작권 침해 소송 제기”, 《저작권 동향》, 제7호, 한국저작권위원회.
- 문정욱 외(2024), 『AI 윤리·신뢰성 확보를 위한 실천 방안 및 정책연구』, 과학기술정보통신부·정보통신정책연구원.

[해외문헌]

- The White House(2025). *America's AI Action Plan*.
- Allenby, G. M. & Rossi, P. E.(1998). “Marketing models of consumer heterogeneity.” *Journal of econometrics*. 89(1-2), pp. 57-78.
- Batura, O., Van Gorp, N., & Larouche, P.(2015). “Online Platforms and the EU Digital Single Market: A Response to the Call for Evidence by the House of Lords’ Internal Market Sub-Committee.” e-

Economics.

McKinsey Global Institute(2017). *Jobs Lost, Jobs Gained: Workforce Transitions in a Time of Automation*. McKinsey & Company.

_____(2023). *The Economic Potential of Generative AI. The next Productivity Frontier*. McKinsey & Company.

Demirci, Ozge, Jonas Hannane, & Xinrong Zhu(2025). “Who Is AI Replacing? The Impact of Generative AI on Online Freelancing Platforms.” *Management Science*.

Brynjolfsson, E., Li, D., & Raymond, L. R.(2023). “Generative AI at Work(No. w31161).” National Bureau of Economic Research.

Noy, Shakked & Whitney Zhang(2023). “Experimental evidence on the productivity effects of generative artificial intelligence.” *Science*. 381, pp. 187-192.

Peng, Sida, Eirini Kalliamvakou, Peter Cihon, & Mert Demirer(2023). “The Impact of AI on Developer Productivity: Evidence from GitHub Copilot.” arXiv preprint.

Kim, Duk Gyoo & Ahram Moon(2024). “From helping hand to stumbling block: the ChatGPT paradox in competency experiment.” *Applied Economics Letters*, 32(18), 2669-2673.

Wittenberg, Chloe, Ziv Epstein, Adam J. Berinsky, & David G. Rand(2024). “Labeling AI-Generated Content: Promises, Perils, and Future Directions.” An MIT Exploration of Generative AI, March.

Wittenberg, Chloe, Ziv Epstein, Gabrielle Peloquin-Skulski, Adam J. Berinsky, & David G.(2025. 6) “Rand, Labeling AI-generated media online.” *PNAS Nexus*. Vol. 4, Issue 6.

Luo, Xueming, Siliang Tong, Zheng Fang, & Zhe Qu(2019). “Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot

- Disclosure on Customer Purchases.” *Marketing Science*. 38(6), pp. 937-947.
- Baek, Tae Hyun, Jungkeun Kim, & Jeong Hyun Kim.(2024). “Effect of disclosing AI-generated content on prosocial advertising evaluation.” *International Journal of Advertising*, pp. 1-22.
- Li, Zhuoyan, Chen Liang, Jing Peng, & Ming Yin(2024). “How Does the Disclosure of AI Assistance Affect the Perceptions of Writing?” *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pp. 4849-4868.

3장 5절

[국내 문헌]

- 고홍석(2023), “미디어 리터러시 역량의 지역 격차”, 《KISDI STAT Report》, Vol. 23-22, 정보통신정책연구원.
- 박준혁·심홍진(2023), “지능정보사회 이용자 패널조사를 통해 살펴본 국내 정보 격차 실태 및 특성”, 《KISDI Perspectives》, No. 5, 정보통신정책연구원.
- 이준영·김영민(2024), 『디지털 전환 기술 도입의 지역·산업별 기업 효과』, 연구자료 2024-05, 산업연구원.
- 장재영·박소연(2022), 『ICT 투자가 지역 균형 발전에 미치는 영향』, 기본연구 22-05, 정보통신정책연구원.

[해외문헌]

- Calvino, F., Dernis, H., Samek, L. and Ughi, A.(2024). *A Sectoral Taxonomy of AI Intensity*. OECD Artificial Intelligence Papers No. 30. Paris: OECD Publishing.

- Jaffe, A. B., Trajtenberg, M., and Henderson R.(1993). “Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations.” *The Quarterly Journal of Economics*. 108(3), pp. 577-598.
- Kalyani, A., Bloom, N., Carvalho, M., Hassan, T., and Lerner, J.(2025). “The Diffusion of New Technologies.” *The Quarterly Journal of Economics*. 140(2), pp. 1299-1365.
- Krugman, P.(1991). “Increasing Returns and Economic Geography.” *Journal of Political Economy*. 99(3), pp. 483-499.
- Marshall, A.(1890). *Principles of Economics*. 8th ed. London: Macmillan and Co.
- Moretti, E.(2021). “The Effect of High-Tech Clusters on the Productivity of Top Inventors.” *American Economic Review*. 111(10), pp. 3328-3375.
- Muro, M. and Methkupal, S.(2025). “Mapping the AI Economy: Which Regions Are Ready for the Next Technology Leap.” *Brookings Institution Report*. Washington, DC: Brookings Institution.
- OECD(2024). *Job Creation and Local Economic Development 2024: The Geography of Generative AI*. Paris: OECD Publishing.
- Rosenthal, S. S. and Strange, W. C.(2008). “The Attenuation of Human Capital Spillovers.” *Journal of Urban Economics*. 64(2), pp. 373-389.

[부록 1]

실험자극물(1) 서울 송례문

	AI가 전혀 활용되지 않은 이미지 (연합뉴스 온라인 기사에 사용된 보도사진 원본)	AI가 부분적으로 활용된 이미지 (연합뉴스 온라인 기사에 사용된 보도사진 원본의 일부 시각적 요소를 AI로 변형한 이미지)	AI가 전면적으로 활용된 이미지 (이미지의 모든 시각적 요소를 AI로 생성한 이미지)
현장성 이미지			
연출성 이미지			

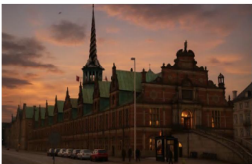
실험자극물(2) 강원도 낙산사

	AI가 전혀 활용되지 않은 이미지 (연합뉴스 온라인 기사에 사용된 보도사진 원본)	AI가 부분적으로 활용된 이미지 (연합뉴스 온라인 기사에 사용된 보도사진 원본의 일부 시각적 요소를 AI로 변형한 이미지)	AI가 전면적으로 활용된 이미지 (이미지의 모든 시각적 요소를 AI로 생성한 이미지)
현장성 이미지			
연출성 이미지			

실험자극물(3) 덴마크 코펜하겐 구 증권거래소

	시가 전혀 활용되지 않은 이미지 (연합뉴스 온라인 기사에 사용된 보도사진 원본)	시가 부분적으로 활용된 이미지 (연합뉴스 온라인 기사에 사용된 보도사진 원본의 일부 시각적 요소를 시로 변형한 이미지)	시가 전면적으로 활용된 이미지 (이미지의 모든 시각적 요소를 시로 생성한 이미지)
현장성 이미지			
연출성 이미지			

실험자극물(4) 브라질 리우데자네이루 국립박물관

	시가 전혀 활용되지 않은 이미지 (연합뉴스 온라인 기사에 사용된 보도사진 원본)	시가 부분적으로 활용된 이미지 (연합뉴스 온라인 기사에 사용된 보도사진 원본의 일부 시각적 요소를 시로 변형한 이미지)	시가 전면적으로 활용된 이미지 (이미지의 모든 시각적 요소를 시로 생성한 이미지)
현장성 이미지			
연출성 이미지			

[부록 2]

디지털 이미지 인식 파악을 위한 조사

안녕하십니까? 권슈머인사이트에서는 정보통신정책연구원의 의뢰로 디지털 이미지 인식 파악을 위한 조사를 실시하고 있습니다. 본 조사는 디지털 이미지에 대한 개인의 인식과 디지털 이미지 표시에 대한 태도를 알아보기 위한 연구로, 바쁘시겠지만 잠시만 시간을 내주시면 큰 도움이 되겠습니다. 감사합니다.

SQ1. 귀하의 성별은 어떻게 됩니까?

- 1. 남성
- 2. 여성

SQ2. 귀하의 출생년도는 어떻게 됩니까?

년 → 만20~64세까지만 진행

SQ3. 귀하께서 현재 살고 계시는 지역은 어디입니까?

1.서울	2.부산	3.대구
4.인천	5.광주	6.대전
7.울산	8.세종	9.경기도
10.강원도	11.충청북도	12.충청남도
13.전라북도	14.전라남도	15.경상북도
16.경상남도	17.제주도	

=====

18.외국에 거주한다 → 조사종단

SQ4. 본 조사에는 재난 현장 이미지가 포함될 수 있으며, 이는 정서적 불편감을 유발할 수 있습니다.
본 조사에 참여 여부는 자발적으로 이루어집니다. 귀하께서는 본 조사 참여에 동의하십니까?

- 1.예
- 2.아니오 → 조사종단

SQ5. 귀하께서는 아래 개인정보 수집에 동의하십니까?

- ※ 수집된 모든 개인 정보는 익명으로 처리되어 답례 제공 목적으로만 사용됩니다.
 ※ 귀하께서는 동의를 거부할 권리가 있으며, 동의하지 않는 경우 조사는 중단됩니다.

[개인정보 수집, 이용 동의] 응답자 확인 및 답례 지급에 필요한 개인정보 수집 안내문

- ① 개인정보의 수집·이용 목적 : 응답자 확인, 답례 제공 등에 필요한 개인정보 수집·이용
 ② 수집 대상 개인정보 항목 : 응답자 이름, 휴대폰 번호, (현금 지급시) 주민등록 번호 및 계좌번호
 ③ 개인정보 보유 및 이용기간 : 조사 종료 후 3개월

1. 동의한다
 2. 동의하지 않는다 → **조사종단**

귀하께서는 본 조사의 응답자로 선정되었습니다. 본 설문에 대한 귀하의 응답은 이미지 표시 제도 연구에 중요한 자료로 활용될 것입니다. 이에 귀하의 성실하고 솔직한 응답 부탁드립니다. 감사합니다.

A. 인공지능(AI) 활용 이미지 표시 방법 평가

먼저, 다음 안내문을 읽고, 질문에 응답해 주십시오.

1. [질문 내용 안내]

지금부터 귀하께 8개의 이미지가 순차적으로 제시됩니다.

각 이미지는 AI 활용 정도에 따라 다음 세 가지 유형 중 하나에 해당합니다.

- AI가 전혀 활용되지 않은 이미지는 실제 연합뉴스 온라인 기사에 사용된 보도사진 원본입니다.
- AI가 부분적으로 활용된 이미지는 실제 연합뉴스 온라인 기사에 사용된 보도사진 원본의 일부 시각적 요소를 AI로 변형한 이미지입니다.
- AI가 전면적으로 활용된 이미지는 이미지의 모든 시각적 요소를 AI로 생성한 이미지입니다.

- ① (1차 응답) 귀하께서는 각 이미지에 대하여, AI가 얼마나 활용되었는지를 추측하여, 응답합니다.

■ 제시 이미지 및 SI 활용 수준 질문 예시

이미지 설명	귀하께서는 이 이미지 제작에 SI가 얼마나 활용되었다고 생각하십니까?
서울 송례문의 아름다운 모습이다. 	1. SI가 전혀 활용되지 않은 이미지이다 2. SI가 부분적으로 활용된 이미지이다 3. SI가 전면적으로 활용된 이미지이다 ※ 참고사항 1. SI가 전혀 활용되지 않은 이미지는 실제 연합뉴스 온라인 기사에 사용된 보도사진 원본입니다. 2. SI가 부분적으로 활용된 이미지는 실제 연합뉴스 온라인 기사에 사용된 보도사진 원본의 일부 시각적 요소를 SI로 변형한 이미지입니다. 3. SI가 전면적으로 활용된 이미지는 이미지의 모든 요소를 SI로 생성한 이미지입니다.

② (2차 응답) 귀하께는 동일한 이미지가 다시 제시되며, SI가 얼마나 활용되었는지 다시 한 번 추측하여 응답합니다.

2. [기본 답례 안내]

본 조사에 끝까지 응답하신 경우, 모든 응답자에게 기본 답례로 5,000원을 드립니다.

3. [추가 답례 안내]

- ① 조사가 종료된 후, 귀하에게 제시된 8장의 이미지 중 1장을 무작위로 선택한 후,
- ② 선택한 이미지의 SI 활용 수준에 대한 귀하의 1차 또는 2차 응답 중 하나를 무작위로 선택합니다.
- ③ 선택한 귀하의 응답이 실제 SI 활용 수준과 일치할 경우, 귀하께 추가 답례 3,000원이 지급됩니다.
- ④ 즉, 제시된 8장의 이미지의 SI 활용 수준을 정확히 추측할수록, 귀하께서 추가 답례를 받을 가능성이 높아집니다. 이에 귀하의 성실한 참여와 신중한 응답을 부탁드립니다.

※ 귀하께서는 최대 8,000원(기본 답례 5,000원 + 추가 답례 3,000원)의 답례를 받을 수 있습니다.

※ 기본 답례 및 추가 답례는 현금 또는 상품권으로 지급됩니다. 답례 지급 방식(현금 또는 상품권 여부 및 종류)은 본 조사 마지막에 질문드릴 예정입니다.

다음은 8개의 이미지 중, **N번째** 이미지에 대한 질문입니다.

이미지 설명



평가 이미지 동시 제시

문1. (전체 응답자)

귀하께서는 이 이미지 제작에 **SI**가 얼마나 활용되었다고 생각하십니까?

※참고사항

1. **SI**가 전혀 활용되지 않은 이미지는 실제 연합뉴스 온라인 기사에 사용된 보도사진 원본입니다.
2. **SI**가 부분적으로 활용된 이미지는 실제 연합뉴스 온라인 기사에 사용된 보도사진 원본의 일부 시각적 요소를 **SI**로 변형한 이미지입니다.
3. **SI**가 전면적으로 활용된 이미지는 이미지의 모든 요소를 **SI**로 생성한 이미지입니다.

1. **SI**가 전혀 활용되지 않은 이미지이다
2. **SI**가 부분적으로 활용된 이미지이다
3. **SI**가 전면적으로 활용된 이미지이다

평가 이미지 제시

문2. (전체 응답자)

귀하께서 위 이미지를 온라인에서 보게 된다면, **다른 사람과 공유**하시겠습니까?

1. 공유하겠다
2. 공유하지 않겠다

문2-1. (다섯번째 이미지 평가에만 제시)

귀하께서는 이 이미지에 **SI** 활용 여부를 표시하는 것이 얼마나 필요하다고 생각하십니까?

이 문항은 응답자의 주의력을 확인하기 위한 문항입니다. 아래 보기 중 '그렇지 않은 편이다'를 선택해 주십시오.

전혀 그렇지 않다	그렇지 않은 편이다	가끔 봤다	그런 편이다	매우 그렇다
1	2	3	4	5

평가 이미지 제시

문3. (전체 응답자)

이 이미지에 대해 다음 4개의 항목별로, 귀하의 생각과 가장 가까운 수준에 응답해 주십시오.

전혀 그렇지 않다	그렇지 않은 편이다	보통 이다	그런 편이다	매우 그렇다
1	2	3	4	5

1. 이 이미지는 왜곡이나 과장 없이 실제 일어난 일을 있는 그대로 보여준다	1	2	3	4	5
2. 이 이미지는 사회적으로 중대한 사건·현상과 관련된 내용을 담고 있다.	1	2	3	4	5
3. 이 이미지는 구도, 색감 등 시각적으로 아름답다	1	2	3	4	5
4. 이 이미지는 보는 사람에게 강한 인상이나 감정을 불러일으킨다	1	2	3	4	5

평가 이미지 제시

문4. (전체 응답자)

귀하께서는 이 이미지에 대해 전반적으로 얼마나 만족하십니까?

매우 불만족한다	불만족하는 편이다	보통이다	만족하는 편이다	매우 만족한다
1	2	3	4	5

AI 활용 정도 공개



위의 이미지는 (아래 3개 보기 중, 실제 활용 수준 제시 : 별도 파일 참조)

1. AI가 전혀 활용되지 않은 이미지입니다.
2. AI가 부분적으로 활용된 이미지입니다.
3. AI가 전면적으로 활용된 이미지입니다.

※ 참고사항

1. 시가 전혀 활용되지 않은 이미지는 실제 연합뉴스 온라인 기사에 사용된 보도사진 원본입니다.
2. 시가 부분적으로 활용된 이미지는 실제 연합뉴스 온라인 기사에 사용된 보도사진 원본의 일부 시각적 요소를 시로 변형한 이미지입니다.
3. 시가 전면적으로 활용된 이미지는 이미지의 모든 요소를 시로 생성한 이미지입니다.

앞서 귀하께서 응답하신 이미지에 대해 다시 질문드리겠습니다

📖 평가 이미지 제시

문5. (전체 응답자)

다음은 앞서 귀하께서 응답한 이미지에 대해 다시 묻는 질문입니다.

귀하께서 위 이미지를 온라인에서 보게 된다면, 다른 사람과 공유하시겠습니까?

1. 공유하겠다
2. 공유하지 않겠다

📖 평가 이미지 제시

문6. (전체 응답자)

다음은 앞서 귀하께서 응답한 이미지에 대해 다시 묻는 질문입니다.

이 이미지에 대해 다음 4개의 항목별로, 귀하의 생각과 가장 가까운 수준에 응답해 주십시오.

전혀 그렇지 않다	그렇지 않은 편이다	보통 이다	그런 편이다	매우 그렇다
1	2	3	4	5

1. 이 이미지는 왜곡이나 과장 없이 실제 일어난 일을 있는 그대로 보여준다 1 2 3 4 5
2. 이 이미지는 사회적으로 중대한 사건·현상과 관련된 내용을 담고 있다. 1 2 3 4 5
3. 이 이미지는 구도, 색감 등 시각적으로 아름답다 1 2 3 4 5
4. 이 이미지는 보는 사람에게 강한 인상이나 감정을 불러일으킨다 1 2 3 4 5

📖 평가 이미지 제시

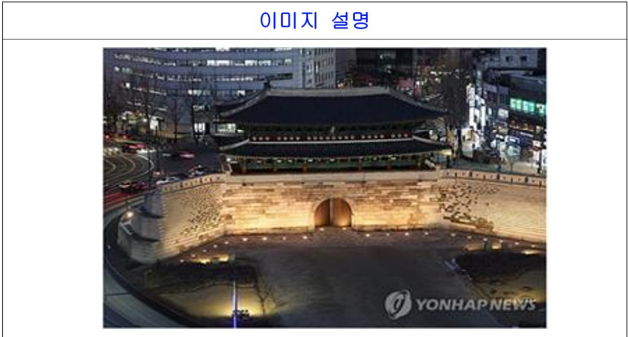
문7. (전체 응답자)

다음은 앞서 귀하께서 응답한 이미지에 대해 다시 묻는 질문입니다.

귀하께서는 이 이미지에 대해 전반적으로 얼마나 만족하십니까?

매우 불만족한다	불만족하는 편이다	보통이다	만족하는 편이다	매우 만족한다
1	2	3	4	5

▣ AI 활용 정도 공개



위의 이미지는 (아래 3개 보기 중, 실제 활용 수준 제시 : 별도 파일 참조)

1. AI가 전혀 활용되지 않은 이미지입니다.
2. AI가 부분적으로 활용된 이미지입니다.
3. AI가 전면적으로 활용된 이미지입니다.

※ 참고사항

1. AI가 전혀 활용되지 않은 이미지는 실제 연합뉴스 온라인 기사에 사용된 보도사진 원본입니다.
2. AI가 부분적으로 활용된 이미지는 실제 연합뉴스 온라인 기사에 사용된 보도사진 원본의 일부 시각적 요소를 AI로 변형한 이미지입니다.
3. AI가 전면적으로 활용된 이미지는 이미지의 모든 요소를 AI로 생성한 이미지입니다.

문8. (전체 응답자)

귀하께서는 이 이미지에 어떤 ‘AI 활용 여부 표시 라벨’을 부착하는 것이 사회적으로 가장 바람직하다고 보십니까?

1. 표시 없음 *“(표시 없음)”을 선택한 경우, 아무 라벨도 표시되지 않습니다
2. AI가 사용되지 않은 콘텐츠 (No AI used)
3. AI의 도움을 받아 제작된 콘텐츠(AI-assisted)
4. AI가 생성한 콘텐츠(AI-generated)
5. 사실처럼 보이도록 합성·변형된 콘텐츠(Altered or synthetic content that is realistic)

문9. (전체 응답자)

귀하께서는 이 이미지에 ‘AI 활용 여부 표시 라벨’을 어떤 방식으로 부착하는 것이 사회적으로 가장 바람직하다고 보십니까?

1. 표시하지 않는다
2. 눈에 보이지는 않지만, 파일 정보에는 기록하는 방식으로 표시한다
(예: 파일 속성을 확인해야만 알 수 있도록 '메타데이터' 또는 '비가시적 워터마크' 형태로 기록)
3. 거의 눈에 띄지 않게 표시한다
(예: 추가 정보창을 클릭시에만 확인 가능하게 하거나 이미지를 자세히 봐야만 보이는 정도로 모서리에 아주 작은 글씨나 아이콘으로 표기)
4. 보통 수준으로 눈에 띄게 표시한다
(예: 이미지의 중앙 또는 하단에 반투명한 작은 글씨나 아이콘으로 표시)
5. 매우 눈에 잘 띄게 표시한다
(예: 이미지 중앙 또는 하단에 선명한 색상의 큰 글씨나 아이콘으로 표시)

B. AI 생성물 표시 제도에 대한 인식

다음은 AI 생성물 표시 제도에 대한 귀하의 의견을 묻는 질문입니다.

문1. (전체 응답자)

귀하께서는 AI 생성물(예: 딥페이크 이미지·영상, AI로 작성된 기사나 글 등)을 온라인에서 실제로 접해본 경험이 있습니까?

- ☐ 1. 예(있다)
☐ 2. 아니오(없다) → **문3.으로 이동**
☐ 3. 잘 모르겠다



문2-1. (문1.에서 1. 예(있다)에 응답한 응답자)

최근 1년 동안, 귀하께서는 AI 생성물(예: 딥페이크 이미지·영상, AI로 작성된 기사나 글 등)을 얼마나 자주 보았습니까?

전혀 보지 못했다	거의 보지 못했다	가끔 봤다	자주 봤다	매우 자주 봤다
1	2	3	4	5

문2-2. (문1.에서 1. 예(있다)에 응답한 응답자)

귀하께서는 AI 생성물(예: 딥페이크 이미지·영상, AI로 작성된 기사나 글 등)을 어디 어디에서 봤습니까? 모두 응답해 주십시오.

1. 온라인 뉴스 / 포털 기사
2. 유튜브, 틱톡 등 동영상 플랫폼
3. 인스타그램, 페이스북 등 SNS
4. 커뮤니티 / 게시판 (예: 디시인사이드, 에펠키키아 등)
5. 메신저 / 채팅방 (예: 카카오톡, 인스타그램 DM 등)
6. 광고나 캠페인 콘텐츠
7. 기타 (구체적으로 응답해 주십시오.)

문2-3. (문1.에서 1. 예(있다)에 응답한 응답자)

귀하께서는 AI 생성물(예: 딥페이크 이미지·영상, AI로 작성된 기사나 글 등)을 봤을 때, 주로 어떤 느낌을 받으셨습니까? 다음 보기 중, 가장 가까운 하나에 응답해 주십시오.

1. 흥미롭고 신기했다
2. 불안하거나 불쾌했다
3. 실재와 구분하기 어려워 혼란스러웠다
4. 별다른 감정이 없었다
5. 기타 (구체적으로 응답해 주십시오.)

문3. (전체 응답자)

귀하께서는 AI 생성물(예: 딥페이크 이미지·영상, AI로 작성된 기사나 글 등)을 실제 사람이 만든 콘텐츠와 구별할 자신이 얼마나 있으십니까?

전혀 없다	없는 편이다	보통이다	많은 편이다	매우 많다
1	2	3	4	5

문4. (전체 응답자)

귀하께서는 'AI 생성물 투명성'이란 개념을 어떻게 이해하고 있습니까?

1. 모든 콘텐츠에 대해 AI가 어느 단계에서, 어떻게 활용되었는지 등 모든 정보를 공개하는 것
2. 모든 콘텐츠에 대해 AI의 주요 활용 여부를 명시하는 것
3. 일부 위험이나 오해의 소지가 있는 콘텐츠에 한해 AI가 어느 단계에서, 어떻게 활용되었는지 등의 모든 정보를 공개하는 것
4. 일부 위험이나 오해의 소지가 있는 콘텐츠에 한해 대해 AI의 주요 활용 여부를 명시하는 것
5. 기타 (구체적으로 응답해 주십시오.)

2026년 1월, 국내에서 AI 기본법이 본격적으로 시행될 예정인 가운데, 방송·영상·미디어·콘텐츠 분야에서 AI 기술이 활용된 경우 해당 콘텐츠에 ‘AI 생성물’임을 표시하는 의무화 방안이 추진되고 있습니다.

「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」

[시행 2026. 1. 22.] [법률 제20676호, 2025. 1. 21., 제정]

제31조(인공지능 투명성 확보 의무) ㉠ 인공지능사업자는 고영향 인공지능이나 생성형 인공지능을 이용한 제품 또는 서비스를 제공하려는 경우 제품 또는 서비스가 해당 인공지능에 기반하여 운용된다는 사실을 이용자에게 사전에 고지하여야 한다.

㉡ 인공지능사업자는 생성형 인공지능 또는 이를 이용한 제품 또는 서비스를 제공하는 경우 그결과물이 생성형 인공지능에 의하여 생성되었다는 사실을 표시하여야 한다.

㉢ 인공지능사업자는 인공지능시스템을 이용하여 실제와 구분하기 어려운 기상의 음향, 이미지또는 영상 등의 결과물을 제공하는 경우 해당 결과물이 인공지능시스템에 의하여 생성되었다는 사실을 이용자가 명확하게 인식할 수 있는 방식으로 고지 또는 표시하여야 한다. 이 경우 해당결과물이 예술적·창의적 표현물에 해당하거나 그 일부를 구성하는 경우에는 전시 또는 향유 등 율저해하지 아니하는 방식으로 고지 또는 표시할 수 있다.

㉣ 그 밖에 제1항에 따른 사전고지, 제2항에 따른 표시, 제3항에 따른 고지 또는 표시의 방법 및 그 예외 등에 관하여 필요한 사항은 대통령령으로 정한다.

문5. (전체 응답자)

AI 생성을 표시 의무화를 생각했을 때, 귀하에게 가장 먼저 떠오르는 생각이나 느낌은 무엇입니까? 자유롭게 작성해 주시기 바랍니다.

(예 : ○○해서 필요하다 / ○○해서 우려된다 / 잘 모르겠다 등 어떤 답변이든 괜찮습니다. 생각이 떠오르는 대로 적어주세요.)

응답란 : _____

문6. (전체 응답자)

귀하께서는 AI 생성물 표시 의무화가 시행될 경우, 어떤 긍정적 효과나 기대되는 점이 있을 것으로 생각하십니까? 자유롭게 작성해 주시기 바랍니다.

응답란 : _____

문7. (전체 응답자)

귀하께서는 AI 생성물 표시 의무화가 시행될 경우, 어떤 부정적 영향이나 우려되는 점이 있을 것으로 생각하십니까? 자유롭게 작성해주시기 바랍니다.

응답란 : _____

문8. (전체 응답자)

귀하께서는 AI 생성물 표시 의무화가 시행될 경우, 다음과 같은 긍정적 영향이 나타날 가능성이 얼마나 있다고 생각하십니까?

전혀 그렇지 않다	그렇지 않은 편이다	보통 이다	그런 편이다	매우 그렇다
1	2	3	4	5

1. 콘텐츠 제작과 유통의 투명성이 높아질 수 있다
- 1 2 3 4 5
2. 온라인 정보에 대한 신뢰성과 책임성이 강화될 수 있다
- 1 2 3 4 5
3. 건전한 AI 활용 문화가 조성될 수 있다
- 1 2 3 4 5
4. AI 기술의 사회적 수용성과 신뢰도가 향상될 수 있다
- 1 2 3 4 5
5. 이용자가 AI 생성물임을 명확히 인식하여, 딥페이크나
가짜 뉴스 등으로 인한 피해를 예방할 수 있다
- 1 2 3 4 5

문9. (전체 응답자)

귀하께서는 AI 생성물 표시 의무화가 시행될 경우, 다음과 같은 부정적 영향이 나타날
가능성이 얼마나 있다고 생각하십니까?

※ 앞 질문은 긍정적인 영향을 묻는 질문이며, 이 질문은 부정적인 영향을 묻는 질문입니다.

전혀 그렇지 않다	그렇지 않은 편이다	보통 이다	그런 편이다	매우 그렇다
1	2	3	4	5

1. 표현의 자유가 침해되거나, 창작 활동이 위축될 수 있다
- 1 2 3 4 5
2. 표시 기준이 모호하여 불필요한 혼란이나 분쟁이
생길 수 있다
- 1 2 3 4 5
3. 창작자가 AI 활용 사실을 숨기려는 행위가 늘어날
수 있다
- 1 2 3 4 5
4. 이용자가 AI 생성물에 대한 과도한 불신이나
부정적인 인식을 가지게 될 수 있다
- 1 2 3 4 5
5. AI 활용 자체가 위축되어 기술 혁신이 지연될 수 있다
- 1 2 3 4 5

문10. (전체 응답자)

AI 생성을 표시 제도를 설계할 때, 가장 중요하게 고려되어야 할 요인은 무엇이라고 생
각하십니까?

응답란 : _____

문11. (전체 응답자)

귀하께서 생각하는 바람직한 AI 생성물 표시 제도는 어떤 모습입니까?

응답란 : _____

C. 디지털 서비스 이용 현황

다음은 귀하의 디지털 서비스 이용 현황에 대한 질문입니다.

문1. (전체 응답자)

귀하께서 최근 3개월 동안 아래의 플랫폼 서비스를 이용해 본 경험이 있는지 여부를 응답해 주십시오.

	이용 경험 여부	
	응답란	
1. SNS(소셜미디어 : 인스타그램, 페이스북, X, 카카오톡 등)	1. 있다	2. 없다
2. 동영상 공유/스트리밍(유튜브, 틱톡, SOOP 등)	1. 있다	2. 없다
3. 온라인 커뮤니티(디시인사이드, 에펠키키아 등)	1. 있다	2. 없다

문2. (전체 응답자)

귀하께서는 평소 온라인(SNS, 온라인 커뮤니티 등)에서 얼마나 적극적으로 참여하십니까? 다음 보기 중, 해당되는 항목에 모두 응답해 주십시오. (복수응답)

1. 게시된 내용을 주로 열람하거나, 정보를 검색한다
2. 게시된 내용에 댓글을 작성한다
3. 게시된 내용에 '좋아요'나 '찬성/반대' 등의 공감 표시를 한다.
4. 게시된 내용을 가족, 친구, 지인과 공유한다
5. 직접 글이나 영상을 게시한다
6. 기타 (구체적으로 응답해 주십시오.)

문5. (문3.에서 12. 대화형 AI에 응답한 응답자)

귀하께서 최근 3개월 동안 제시한 생성형 AI 서비스를 이용해 본 경험이 있는지 여부를 응답해 주십시오.

	이용 경험 여부	
	응답란	
1. 텍스트 생성 (챗GPT, Google Gemini, Anthropic Claude, Meta Llama 3 등)	1. 있다	2. 없다
2. 음성·음악 생성 (Suno AI, Synthesizer V, VALL-E, Voicemod AI 등)	1. 있다	2. 없다
3. 이미지 생성 (DALL-E 3, Midjourney V6, Stable Diffusion 3 Medium, Adobe Firefly 등)	1. 있다	2. 없다
4. 동영상 생성 (Runway Gen-3 Alpha, Pika Labs AI, Synthesia, D-ID 등)	1. 있다	2. 없다
5. 코드 생성 (GitHub Copilot, Amazon Q Developer, Replit Ghostwriter 등)	1. 있다	2. 없다
6. 기타 (구체적으로 응답해 주십시오.)	1. 있다	2. 없다

D. 응답자 특성

다음은 자료분류와 답례 제공을 위한 질문입니다.

D1. 귀하의 최종 학력은 어떻게 됩니까?

1. 고등학교 졸업 이하
2. 전문대 재학/졸업
3. 대학교 재학/졸업
4. 대학원 재학/졸업 이상

D2. 귀하의 직업은 무엇입니까?

1. 자영업 (종업원 9명 이하 소규모업소 주인/가족종사자)
2. 판매/서비스직 (상점점원, 세일즈맨 등)
3. 기능/숙련공 (운전사, 선반/목공, 숙련공 등)
4. 일반작업직 (토목 현장작업/청소/수위/육체노동 등)
5. 사무/기술직 (일반회사 사무직/기술직, 교사, 항해사 등)
6. 경영/관리직 (5급 이상 공무원/교장/기업체 부장 이상 등)
7. 전문/자유직 (대학교수/의사/변호사/예술가/종교가 등)
8. 전업주부
9. 학생
10. 기타 (구체적으로 응답해 주십시오.)

D3. 현재 귀하의 월평균 총수입은 얼마나 됩니까? 보너스, 이자수입 등 모든 수입을 합해서 응답해 주십시오.

1. 월 99만원 이하
2. 월 100 - 199만원
3. 월 200 - 299만원
4. 월 300 - 399만원
5. 월 400 - 499만원
6. 월 500 - 599만원
7. 월 600 - 699만원
8. 월 700 - 799만원
9. 월 800 - 899만원
10. 월 900 - 999만원
11. 월 1,000만원 이상

D4. 응답자 답례 제공을 위해, 귀하의 이름과 휴대폰 번호를 응답해 주십시오.

개인정보 항목	응답란	응답 확인란
이름		
휴대폰 번호		

D5. 본 조사 참여 답례(기본 답례 + 추가 답례)는, 귀하의 희망에 따라, 현금 또는 모바일 상품권으로 지급해드릴 예정입니다. 귀하께서는 본 연구 참여에 대한 답례를 현금으로 받으시겠습니까? 모바일 상품권으로 받으시겠습니까?

※ 현금으로 받으려면, 귀하의 계좌번호(은행명 포함)를 입력해야 하며, 모바일 상품권으로 받으려면, 희망하는 모바일 상품권 종류를 응답해야 합니다.

└── 1. 현금으로 받겠다(귀하의 계좌로 입금)

| 2. 모바일 상품권으로 받겠다(귀하의 휴대폰으로 발송) → **D6-2.로 이동**



D6-1. (D5.에서 1. 현금으로 받겠다(귀하의 계좌로 입금)에 응답한 응답자) 귀하께서는 본 연구 참여에 대한 답례로 현금을 응답하였습니다. 참여 답례 입금을 위한 귀하 명의의 계좌번호와 지출증빙을 위한 귀하의 주민등록번호를 응답해 주십시오.

입금정보 항목	응답란
계좌주	
은행명	
계좌명	

D6-2. (D5.에서 2. 모바일 상품권으로 받겠다(귀하의 휴대폰으로 발송)에 응답한 응답자) 귀하께서는 본 연구 참여에 대한 답례로 모바일 상품권을 응답하였습니다. 귀하께서 받고 싶은 모바일 상품권을 응답해 주십시오.

※ 모바일 상품권은 앞서 귀하께서 응답한 귀하의 휴대폰 번호로 발송할 예정입니다.

1. GS편의점
2. CU편의점
3. 해피콘(SPC)
4. GS칼텍스 주유권

끝까지 조사에 협조해 주셔서 감사합니다.

Abstract

Moon, Ahram; Moon, Jung-Wook; Cho, Sung-Eun; Youn, Sora;
Kim, Hwi-Hong; Lee, Eutteum; Jeon, Minkyung; Shin, Jin-Ho

■ Title

A Study on Social Agendas for Inclusion and Co-prosperity in the AI Era

■ Purpose of Research

This research project is designed to function as a 'Green Paper' that gathers agendas to diagnose the essence of problems and lay the foundation for social discussion, rather than a 'White Paper' that proposes specific policies for the AI era.

■ Main Outcomes of Research and Policy Implications

Accordingly, the purpose of this report is to deeply analyze the specific patterns of imbalance and inequality caused by AI in five major areas of our society—Labor, Public Administration, Users, Information Ecosystem, and Region—and to present key issues that require attention by examining their fundamental mechanisms. This

report is distinct and unique in that it allows for an understanding of the social problems caused by AI as non-fragmented information by examining the similarities and differences among social agendas raised from multiple domains, rather than concentrating on a single field.

정보통신정책연구원 기본연구 안내

■ 2023 기본연구

- 기본연구 23-01 데이터 경제 활성화를 위한 민관 역할분담과 정책 개입영역 연구
(윤성욱, 박소연)
- 기본연구 23-02 해외 주요국 신규 사업자 이동통신시장 진입 효과 분석(김민희)
- 기본연구 23-03 유사국 사례를 통해 본 북한 통신시장 발전 방안 연구(임동민, 서소영,
이중화, 조대근, 서흥수)
- 기본연구 23-04 메타버스 시대 기본권 보호에 관한 연구: 인격권 논의를 중심으로
(권은정, 한혜지, 계인국, 김법연, 이승민)
- 기본연구 23-05 디지털화폐 생태계 변화에 대응한 중앙은행 디지털화폐(CBDC) 도입
정책 연구(박동욱)
- 기본연구 23-06 ICT 확산에 따른 노동시장 임금격차(최지은, 이은영, 최세림, 이현옥)
- 기본연구 23-07 데이터 생산·관리 역량 평가를 위한 성숙도 모형 개발 연구(노희용,
장신재, 박지원)
- 기본연구 23-08 인터넷 생태계 환경 변화에 따른 네트워크 인프라 비용 분담에 관한
연구(염수현, 강인규, 이상규, 김태오)
- 기본연구 23-09 미디어 이용 조사의 모드 효과 비교(신지형)
- 기본연구 23-10 방송미디어분야 자율규제 제도화 방안 연구(이종원)
- 기본연구 23-11 이동통신 네트워크 인프라 산업 생태계의 전환 방향 연구(여재현,
박지현, 윤도원, 장희선, 김선우, 정인준, 박의환)

■ 2024 기본연구

- 기본연구 24-01 AI 반도체 생태계 조사(윤성욱, 김민식, 민대홍)
- 기본연구 24-02 인공지능 감시에 의한 권력의 확대와 그 규범적 대응방안 연구
(문광진, 허진주, 전민경)
- 기본연구 24-03 시청각미디어 콘텐츠의 공정한 거래를 위한 정책 방안 연구: 수익
배분 관련 이슈를 중심으로(강준석)
- 기본연구 24-04 온라인 광고 시장의 경매 경쟁에 대한 연구(백소성)
- 기본연구 24-05 기술진보에 따른 산업별 업무 변화 연구(서영선)
- 기본연구 24-06 5G 요금제 특성이 이용자 편익에 미치는 영향: 단말기와 요금제의
결속을 중심으로(박진환)

- 기본연구 24-07 방송 프로그램 장르 구분의 타당성 검토: 방송법과 방송사업자 및 수용자의 장르 분류 기준 차이를 중심으로(황현정)
- 기본연구 24-08 생성형 AI가 미디어 분야에 미칠 영향에 대한 탐색적 연구(곽동균, 김남두, 주성희, 황현정, 강하연, 김예빈)
- 기본연구 24-09 디지털 경제 활성화를 위한 실물자산토큰(Real World Asset) 시장의 탈중앙화 금융 알고리즘 적용 가능성 연구(김 찬)
- 기본연구 24-10 AI 반도체 정책의 효과성 제고 방안 연구(정현준, 김민식, 오정숙)

■ 2025 기본연구

- 기본연구 25-01 통신시장의 경쟁활성화를 위한 공동지배력 평가 방법론 연구 (라성현)
- 기본연구 25-02 디지털 심화에 따른 몰입형 서비스 생태계 연구 (이재영, 심홍진, 성윤희, 이승환, 구자근)
- 기본연구 25-03 중국 유통플랫폼의 글로벌 확장과 대응방안 (김성옥, 장신재, 손가녕, 전민경)
- 기본연구 25-04 디지털 융합 시대 정보통신법제의 통합과 재편에 관한 연구 (권은정, 김법연, 김지훈, 선지원, 박상미)
- 기본연구 25-05 디지털 환경 변화를 반영한 데이터 트래픽 예측 연구(김민희, 정광재)
- 기본연구 25-06 글로벌 미디어 플랫폼 확산 대응을 위한 국내 미디어 서비스 경쟁력 분석 및 강화 전략 연구(강준석, 권용재)
- 기본연구 25-07 월드뱅크 등 국제기구와의 협력 연구 (황준호, 양종민, 김지원, 문용일, 이효원)
- 기본연구 25-08 생성형 AI의 생산성 분석(민대홍)
- 기본연구 25-09 디지털 전환에 따른 소매패턴 변화와 정책 방향(장신재, 이선희)
- 기본연구 25-10 인공지능·자율시스템 기반 도시의 윤리적 설계 방안 연구 (이현경, 문광진, 전민경)
- 기본연구 25-11 디지털 기술 활용이 기업 성과에 미치는 영향 분석과 정책 시사점 연구(손녕선)
- 기본연구 25-12 생성형 인공지능 서비스 채택 선행요인에 관한 탐색적 연구 (주성희, 임연수, 김예빈)
- 기본연구 25-13 주파수 경매에서 네트워크 구축조건 도입 방식에 대한 연구(백소성)
- 기본연구 25-14 AI 기술주권과 국가경쟁력 제고 방안 연구

(이경선, 김성옥, 이경은, 오장민, 윤혜선)

기본연구 25-15 플랫폼 확산의 인플레이션 영향에 대한 연구(김경은, 조수진, 심명규)

기본연구 25-16 AI 시대의 포용과 상생을 위한 사회적 의제 연구

(문아람, 문정옥, 조성은, 연소라, 김휘홍, 이으뜸, 전민경, 신진호)

기본연구 25-17 ICT 혁신이 사회·경제에 미치는 영향(최지은, 서영선, 노희용)

기본연구 25-18 차세대 통신기술이 통신시장 경쟁구조 및 네트워크 투자에 미치는 영향
분석(김민희, 김경모)



● 저 자 소 개 ●

문 아 랍

- University of Southern California
경제학 박사
- 현 정보통신정책연구원 연구위원

문 정 옥

- 고려대 행정학 박사
- 현 정보통신정책연구원 실장

조 성 은

- Rutgers, the State University of New Jersey
Communication, Information&Library
Studies 박사
- 현 정보통신정책연구원 연구위원

연 소 라

- Texas A&M University 경제학 박사
- 현 정보통신정책연구원 부연구위원

김 휘 홍

- 뮌헨대학교(LMU) 공법학 박사
- 현 정보통신정책연구원 부연구위원

이 으 뜸

- University of Virginia 경제학 박사
- 현 정보통신정책연구원 부연구위원

전 민 경

- 서울과학기술대학교 IT정책전문대학원
정책학 석사
- 현 정보통신정책연구원 연구원

신 진 호

- 한국학중앙연구원 한국학대학원 사회학
석사
- 현 정보통신정책연구원 연구원

기본연구 25-16

AI 시대의 포용과 상생을 위한 사회적 의제 연구

2025년 12월 일 인쇄

2025년 12월 일 발행

발행인 이 상 규

발행처 정 보 통 신 정 책 연 구 원

충청북도 진천군 덕산읍 정통로 18

TEL: 043-531-4114 FAX: 043-535-4695~6

인 쇄 인 성 문 화

ISBN 979-11-7000-439-4 93320

〈비매품〉