기관 2021-032

디지털 전환에 따른 미래사회 위험이슈 발굴 및 대응 전략 연구

A Study on Risk Issues and Response Strategies related to Digital Transformation

구본진

KISTEP

최종보고서

보안등급
일반[O], 보안[ ]

사업명		(AT)미래 이슈 발굴 및 성장전략 기획 연구					
기술분류	국가과학기술 표준분류	1순위 소분류 코드명	%	2순위 소분류 코드명	%	3순위 소분류 코드명	%
	부처기술분류 (해당 시 작성)	1순위 소분류 코드명	%	2순위 소분류 코드명	%	3순위 소분류 코드명	%

연구과제명		
	국문	디지털 전환에 따른 미래사회 위험이슈 발굴 및 대응 전략 연구
	영문	A Study on Risk Issues and Response Strategies related to Digital Transformation

연구책임자	성명	구본진	직급	부연구위원
	소속부서명	미래성장 정책센터	email	bonkoo@ kistep.re.kr
	최종전공	경영공학	국가연구자번호	11620232

연구기간		
	전체	2021. 02. 01 - 2021. 12. 31(0년 11 개월)
	단계 (해당 시 작성)	-
		-

당해연도 연구비	131,000(천원)
총계	131,000(천원)
단계 (해당 시 작성)	(천원)
	(천원)

위탁연구기관 등 (해당 시 작성)	기관명	책임자	직위	휴대전화	전자우편	기관유형
위탁연구기관						

실무담당자	성명	이미화	직급	책임전문관리원
	소속부서명	미래성장 정책센터	email	heyday@ kistep.re.kr
	최종전공	행정학	국가연구자번호	1154203

2021년 12월 31일

연구책임자: 구본진 (인)

한국과학기술기획평가원장: 정병선 (직인)

• **연구진**

- 연구책임자	구본진 (한국과학기술기획평가원 부연구위원)
- 참여연구원	유혜인 (한국과학기술기획평가원 연구원) 이미화 (한국과학기술기획평가원 책임전문관리원) 김진용 (한국과학기술기획평가원 연구위원)
• 외부연구진	심우민 (경인교육대학교 교수) 이종선 (명지대학교 교수) 손종진 (FAIR Labs. 대표)

기관 2021-032

디지털 전환에 따른 미래사회 위험이슈 발굴 및 대응 전략 연구

• 발행인 : 정병선
• 발행처 : 한국과학기술기획평가원
(27740) 충청북도 음성군 맹동면 원중로 1339
Tel) 043-750-2472
• http://www.kistep.re.kr
• 인 쇄 : 화신문화사

< 요 약 서 >

사업명		(AT)미래 이슈 발굴 및 성장전략 기획 연구							
기술분류	국가과학기술표준분류	1순위 소분류 코드명	%	2순위 소분류 코드명	%	3순위 소분류 코드명	%		
	부처기술분류(해당 시 작성)	1순위 소분류 코드명	%	2순위 소분류 코드명	%	3순위 소분류 코드명	%		
연구과제명		디지털 전환에 따른 미래사회 위험이슈 발굴 및 대응 전략 연구							
연구기간		*2021.2.1.~2021.12.31*							
연구비		총 131,000천원							

연구 목표 및 내용			
	최종 목표		디지털 전환에 대한 이해도 제고, 디지털 전환 관련 미래사회 위험이슈 구체화 및 대응 정책 방향 설정
	전체 내용		o 디지털 전환의 개념 정리 및 파급효과 분석 o 디지털 전환이 초래할 수 있는 미래사회 위험이슈 구체화 o 디지털 전환의 미래사회 위험에 대한 인식 조사 o 디지털 전환 관련 미래사회 위험이슈 대응 정책 동향 분석 o 디지털 전환 관련 미래사회 위험이슈 대응 정책 방향 제언
	1단계 (해당 시 작성)	목표	
		내용	
	n단계 (해당 시 작성)	목표	
		내용	

연구성과	디지털 전환의 개념 및 파급효과 정리, 디지털 전환 관련 미래사회 위험 세부 이슈 규명 및 인식 데이터 확보, 관련 정책 동향 정리 및 정책 수립 방향 제언
연구성과 활용계획 및 기대 효과	o 디지털 전환에 대한 이해도 제고를 통하여 정책 범주 명확화 o 디지털 전환의 역기능 방지 정책 수립을 위한 기초 자료 확보 o 디지털 전환 관련 미래사회 위험이슈에 대한 사회적·기술적 논의 종합 o 추상적 대원칙 선언 수준을 넘어 실행 가능성이 확보되는 정책 수립 방향 도출 o 향후 디지털 전환 역기능 방지 정책 및 세부 계획 수립의 기초 자료 제공 o 디지털 전환 관련 기술적·사회적 문제 해결에 대한 가이드 제공

연구개발성과의 등록·기탁 건수	논문	특허	보고서 원문	연구시설·장비	기술요약정보	소프트웨어	표준	생명자원		화합물	신품종	
								생명정보	생물자원		정보	실물

세부 정량적 연구개발성과 건수	과학적 성과			사회적 성과								기타
	논문게재	학술회의발표	보고서 원문	법령반영	정책활용	안건상정	제도개선	다른 연구에 활용	국제협력	(정책) 홍보	포상·수상	

국문핵심어 (5개 이내)	디지털 전환	디지털 전환의 역기능	인공지능	디지털 전환의 위험	
영문핵심어 (5개 이내)	digital transformation	negative effects of digital transformation	artificial intelligence	risk of digital transformation	

요 약 문

□ **(연구배경 및 필요성)** 정부는 경제/사회 전반에 빠르게 확산 중인 디지털 전환의 양면성을 이해하여 디지털 전환의 순기능 촉진과 동시에 역기능 대응 방안도 마련할 필요

- 디지털 전환은 경제/사회 발전 및 생활 편의성을 증진시키는 유용한 도구로 작용하지만 자칫 부정적 영향을 미칠 수도 있는 양면성을 내포
  - (긍정적 영향) 경제 시스템 진화, 단순노동 대체, 의사결정의 신속성/객관성 증대, 생산성 향상, 품질 향상, 운영 효율성 증대 등
  - (부정적 영향) 개인정보 오남용, 보안 위험, 시스템 신뢰도 저하, 지속가능성 저하, 경제적 격차 확대, 사회 갈등 증대, 윤리 문제 야기 등
- 현재 정부는 경쟁력 및 기술 패권 확보 등을 위하여 디지털 전환 촉진정책을 적극적으로 추진하고 있는 반면, 디지털 전환 관련 위험이슈에 대한 이해와 이를 방지할 수 있는 정책 추진은 상대적으로 더딘 경향
- 따라서 정부는 디지털 전환에 대한 이해도 제고를 통한 정책 범주 명확화, 디지털 전환 관련 위험의 세부 이슈 파악, 역기능 대응 정책의 수요영역 파악, 대응 정책 수단 검토 및 방향 설정 등을 통해 디지털 전환 미래사회 위험 대응 방안을 마련할 필요

□ (연구 성과 1) 본 연구는 문헌연구를 통해 디지털 전환의 개념 정리와 파급효과를 분석

- 디지털 전환의 학술적/실무적 정의 및 개념을 정리하여 정책 범주를 명확화
- 파편화되어 학술적/실무적으로 논의되고 있는 디지털 전환의 파급효과를 종합

□ (연구 성과 2) 본 연구는 디지털 전환이 초래할 수 있는 미래사회 위험이슈를 구체화

- 문헌연구를 통한 디지털 전환 역기능의 주요 유형 정리
- 관련 언론기사들을 대규모로 수집하고, Embedded Topic Modeling 분석을 적용하여 최근 야기/논의되고 있는 국내외 디지털 전환 역기능 세부 이슈들을 규명하고, 이를 빈도순으로 정리

※ (의의) 현재까지 전문가들을 중심으로 디지털 전환 미래사회 위험이슈(역기능)이 어떠한 형태로 발현될지에 대한 논의가 주를 이루었다면 본 연구는 한발 더 나아가 정량분석(언론기사 빅데이터를 Embedded Topic Modeling으로 분석)을 통해 실제로 우리 사회에서 발생/논의되고 있는 디지털 전환 미래사회 위험의 세부 이슈들을 규명

- 이를 통하여 정부가 역기능의 이슈별 대응 추진 시, 우선순위 설정에 필요한 근거자료를 확보

□ (연구 성과 3) 다음으로 디지털 전환에 대한 종합적인 인식수준 데이터를 확보

- 국내 기업에 재직 중인 성인들(1,824명)의 디지털 전환에 대한 일반적인 인식수준을 조사하여 연령별/산업별로 결과를 분석
- 나아가 연령별/산업별 디지털 전환 위험에 대한 인식수준 분석과 디지털 전환 역기능 세부 이슈별 정책적 대응 수요영역을 파악
- 해당 결과는 연령/산업 맞춤형 디지털 전환 촉진/역기능 대응 정책 설계의 기반 자료로 활용가능할 것으로 판단

□ (연구 성과 4) 아울러 본 연구는 주요국의 디지털 전환 역기능 대응 정책 동향을 정리

- 한국을 포함한 주요국들의 소극적/적극적 대응 정책 동향 분석
- 현 정책의 문제점 정리 및 정책적 개선 필요 사항을 제안

□ (연구 성과 5) 마지막으로 본 연구는 디지털 전환 역기능 정책에 대한 법리학적 검토와 함께 규제수준별 정책수단을 제시/분석

- 법리학적으로 디지털 전환의 핵심 기술인 인공지능을 정부가 규제하는 것이 타당한지 분석하여 향후 정책적 대응 근거를 확보
- 규제 강도에 따라 4개의 정책(규제) 수단을 제시하여 각 수단별 시나리오를 분석

※ (의의) 정부 입장에서 검토가능한 다양한 유형의 정책 수단을 제시하였고, 각 유형별 정책 수립/전개 시나리오와 유사 사례들을 제시함으로써 향후 종합적인 정책 수단 선택의 기반을 제공

구분	규제 강도
정책 수단 1: 전문가 윤리적 접근	낮음(低)
정책 수단 2: 인증체계 구축	
정책 수단 3: 개인적 권리 설정	
정책 수단 4: 직접적인 행정 규제 설정	높음

[그림] 규제 강도에 따른 디지털 전환 역기능 대응 정책 수단

목 차

Ⅰ. 디지털 전환에 따른 미래사회 위험이슈 발굴 및 대응 전략 연구 … 1

1. 서론 ····· 1
   (1) 연구배경 ····· 1
   (2) 연구목표 및 구성 ····· 2
2. 디지털 전환 관련 문헌연구 ····· 4
   (1) 디지털 전환의 정의 ····· 4
   (2) 디지털 전환의 경제적/사회적 파급효과 ····· 6
3. 디지털 전환 관련 미래사회 위험이슈 ····· 8
   (1) 디지털 전환의 미래사회 위험이슈 개요 ····· 8
   (2) 디지털 전환의 역기능 유형 (문헌연구) ····· 8
   (3) 디지털 전환의 역기능 세부이슈 분석 (국내: 이루다 케이스) ····· 10
   (4) 디지털 전환의 역기능 세부이슈 분석 (주요국 언론기사 정량분석) ···· 12
   (5) 소결 ····· 18
4. 디지털 전환 관련 미래사회 위험에 대한 인식조사 ····· 20
   (1) 인식조사 개요 ····· 20
   (2) 인식조사 구성 ····· 21
   (3) 인식조사 결과 ····· 22
   (4) 소결 ····· 29
5. 디지털 전환 관련 미래사회 위험이슈 대응 정책 동향 ····· 31
   (1) 소극적 정책 대응 ····· 31
   (2) 적극적 정책 대응 ····· 35
   (3) 소결 ····· 37
6. 디지털 전환 관련 미래사회 위험이슈 대응 정책 방향 설정 ····· 39
   (1) 개요 ····· 39
   (2) 인공지능(알고리즘)의 법규범적 속성 검토 ····· 39
   (3) 규제 강도에 따른 정책 시나리오 분석 ····· 43
   (4) 정책적 제언 ····· 48

참고문헌 ····· 49

표 차례

<표 1> 디지털 전환의 역기능 유형 ····· 8

<표 2> '이루다' 관련 디지털 전환의 역기능 위험 세부주제 도출 결과 ··· 11

<표 3> '이루다' 관련 디지털 전환의 역기능 위험 세부주제 도출 결과와 디지털 전환 역기능 유형 매칭 결과 ····· 12

<표 4> 임베디드 토픽 모델링 분석 결과(종합) ····· 13

<표 5> 임베디드 토픽 모델링 분석 결과(종합-세부이슈별 기술분류) ··· 13

<표 6> 임베디드 토픽 모델링 분석 결과(종합-세부이슈별 국가분류) ··· 14

<표 7> 미국의 디지털 전환 관련 세부 위험이슈 도출 결과 ····· 15

<표 8> 미국의 디지털 전환의 역기능 위험 세부주제 도출 결과와 디지털 전환 역기능 유형 매칭 결과 ····· 15

<표 9> 유럽의 디지털 전환 관련 세부 위험이슈 도출 결과 ····· 16

<표 10> 5G Networks 관련 세부 위험이슈 도출 결과 ····· 16

<표 11> AI 관련 세부 위험이슈 도출 결과 ····· 17

<표 12> AI 관련 디지털 전환의 역기능 위험 세부주제 도출 결과와 디지털 전환 역기능 유형 매칭 결과 ····· 17

<표 13> Blockchain 관련 세부 위험이슈 도출 결과 ····· 18

<표 14> 응답자 구성 ····· 20

<표 15> 조사항목 및 내용 ····· 21

<표 16> GDPR 前後 주요 변화 ····· 35

<표 17> GDPR 정보주체의 권리 ····· 36

<표 18> EU AI Act의 인공지능 위험 유형과 유형별 규제 사항 ····· 37

그림 차례

[그림 1] 델 테크놀로지스 디지털 전환 인덱스 ····· 1
[그림 2] 「Principled AI」주요 내용 ····· 9
[그림 3] 유형 매칭 결과 ····· 10
[그림 4] 디지털 전환에 대한 용어 이해도 (전체) ····· 22
[그림 5] 디지털 전환에 대한 이해도 (종사산업별) ····· 22
[그림 6] 디지털 전환에 대한 이해도 (연령대별) ····· 23
[그림 7] 국내 경제/산업 발전에 디지털 전환이 필요하다고 생각하는 정도 (전체) ····· 23
[그림 8] 국내 경제/산업 발전에 디지털 전환이 필요하다고 생각하는 정도 (종사산업별) ····· 24
[그림 9] 국내 경제/산업 발전에 디지털 전환이 필요하다고 생각하는 정도 (연령대별) ····· 24
[그림 10] 디지털 전환의 기업에 대한 긍정적/부정적 영향 인식 수준 (전체) ····· 25
[그림 11] 디지털 전환의 기업에 대한 긍정적/부정적 영향 인식 수준 (종사산업별) ····· 25
[그림 12] 디지털 전환의 기업에 대한 긍정적/부정적 영향 인식 수준 (연령대별) ····· 26
[그림 13] 5개 위험 유형별 위험성 인식 수준 ····· 26
[그림 14] 5개 위험 유형에 대한 위험 인식 우선순위 (1순위 및 2순위 선택 결과) ····· 27
[그림 15] 5개 위험 유형 外 심각한 디지털 전환 위험 요인 (자유응답 결과) ····· 27
[그림 16] 8개 대응 영역별 정책 수요 ····· 28
[그림 17] 8개 대응 영역에 대한 정책적 우선 순위 (1순위 및 2순위 선택 결과) ····· 28
[그림 18] 규제 강도에 따른 정책 시나리오 ····· 43

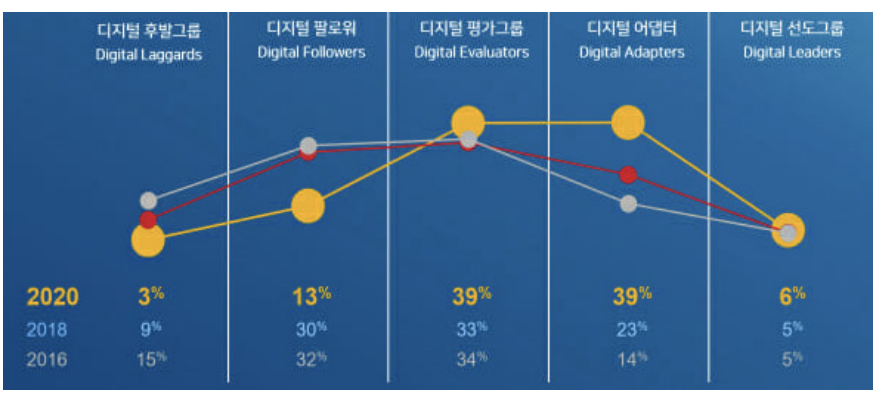
I

디지털 전환에 따른 미래사회 위험이슈 발굴 및 대응 전략 연구

1 서론

(1) 연구배경

□ 디지털 전환(digital transformation)[1]은 경제/사회 및 일상에 깊숙이 들어왔고, 이는 더욱 가속화될 전망[2]



[그림 1] 델 테크놀로지스 디지털 전환 인덱스[3]

1) (정의) 아날로그 데이터와 프로세스를 기계가 읽을 수 있는 형식으로 변환하는 디지화(digitisation)와 디지털 기술과 데이터, 이들의 상호연결을 활용하는 것을 의미하는 것으로 이는 새로운 행위 또는 기존 행위에 변화를 유도하는 디지털화(digitalisation)의 경제적/사회적 효과를 의미 (OECD, 2019), (구성 기술) artificial intelligence, blockchain, big data, IoT, cloud computing, 5G networks (OECD, 2019)

2) 디지털 전환 시장은 연평균 복합 성장률 16.5%로 `20년 4,698억 달러에서 `25년 1조 98억 달러까지 성장할 것으로 예측 (Research & Markets, 2020)

3) 델 테크놀로지스 디지털 트랜스포메이션 인덱스 2020 보고서

- 델 테크놀로지스가 `20년 전세계 18개국 대/중견기업 C레벨 및 관리직 임원 4,300여명을 대상으로 조사한 결과, 최근 3년간 디지털 선도그룹, 디지털 어댑터, 디지털 평가그룹은 지속적으로 증가
- 반면 디지털 성숙도가 낮은 디지털 팔로워 및 후발그룹은 감소
- 디지털 전환 관련 핵심 기술들(인공지능, 빅데이터, 5G network 등)의 빠른 발전과 Covid-19에 따른 디지털 수요 급증은 전분야의 디지털 전환을 가속

□ 정부는 디지털 전환의 양면성에 주목하고, 디지털 전환 촉진 정책과 함께 역기능에 대응할 수 있는 정책적 방안도 검토할 필요

- 디지털 전환은 경제/사회 발전 및 생활 편의성을 증진시키는 유용한 도구로 작용하지만 자칫 부정적인 영향을 미칠 수도 있는 양면성을 내포
  - (긍정적 영향) 경제 시스템 진화, 단순노동 대체, 의사결정의 신속성/객관성 증대, 생산성 향상, 품질 향상, 운영 효율성 증대 등
  - (부정적 영향) 개인정보 오남용, 보안 문제 증대, 시스템 신뢰도 저하, 지속가능성 저하, 경제적 격차 확대, 사회 갈등 증대, 기술(인공지능) 윤리 문제 야기 등
- 한편 정부는 경쟁력 및 기술 패권 확보 등을 위하여 디지털 전환 촉진정책을 적극적으로 추진하고 있는 반면, 디지털 전환 관련 위험이슈에 대한 이해와 이를 방지할 수 있는 정책 추진은 상대적으로 더딘 경향
- 따라서 정부는 디지털 전환 관련 미래사회 위험 대응을 위하여 디지털 전환에 대한 이해도 제고를 통한 정책 범주 명확화, 디지털 전환 관련 위험의 세부 이슈 파악, 역기능 대응 정책 수요영역 파악, 대응 정책 수단 검토 및 방향 설정 등을 추진할 필요

(2) 연구목표 및 구성

□ (목표 1) 디지털 전환의 개념 정리 및 파급효과 분석 (문헌 연구)

- 디지털 전환의 정의 및 범위 명확화
- 디지털 전환의 경제적/사회적 파급효과 파악

□ (목표 2) 디지털 전환이 초래할 수 있는 미래사회 위험이슈 구체화

- 디지털 전환 역기능의 주요 유형 정리 (문헌 연구)
- 국내외 디지털 전환 역기능의 세부 이슈 분석 (정량 분석)

□ (목표 3) 디지털 전환의 미래사회 위험에 대한 인식 파악 (설문 조사)

- 국내 기업에 재직 중인 성인들의 연령별/산업별 디지털 전환에 대한 인식수준 파악
- 연령별/산업별 디지털 전환 위험에 대한 인식수준 파악
- 디지털 전환 역기능 세부 이슈별 정책적 대응 수요 파악

□ (목표 4) 디지털 전환 관련 미래사회 위험이슈 대응 정책 동향 분석

- 주요국의 소극적/적극적 대응 정책 동향 분석
- 현 정책의 문제점 정리

□ (목표 5) 디지털 전환 관련 미래사회 위험이슈 대응 정책 방향 제언

- 법리학적 규제 가능성 검토
- 규제 강도에 따른 정책 시나리오 분석
- 정책 수립 방향 제언

2 디지털 전환 관련 문헌연구

(1) 디지털 전환의 정의

□ 디지털 전환은 합의된 정의가 부재하며 그 범주도 모호한 경향이 있기 때문에 본 절에서는 문헌연구를 통해 디지털 전환의 정의 및 범위를 종합

- (장훈, 2017) 디지털 전환은 디지털 기술이 바탕이 된 전환, 차별화된 변화를 의미하며 산업적 시각에서는 다음과 같이 정의할 수 있음; '다양한 디지털 기술을 바탕으로 기업의 전략이나 시스템 등을 근본적으로 변화시켜 새로운 가치를 창출하는 것'
- (한국마케팅연구원, 2019) 디지털 전환은 디지털 기술을 사회 전반에 적용하여 전통적인 사회구조를 혁신시키는 것으로 일반적으로는 기업에서 사물인터넷(IoT), 클라우드 컴퓨팅, 인공지능(AI), 빅데이터 솔루션 등 정보통신 기술(ICT)을 플랫폼으로 구축/ 활용하여 기존 전통적인 운영 방식과 서비스 등을 혁신하는 것을 의미
- (이동임, 2019) 정보통신기술을 활용하여 산업에서 기계와 프로세스가 지능적으로 연결되는 것
- (김민식 & 손가녕, 2017) 디지털 전환은 기업들이 최신의 디지털 기술을 활용하여 경쟁력을 확보하려는 노력이며 구체적으로는 모바일, 클라우드, 빅데이터, 인공지능 및 사물통신 등 디지털 신기술에 의해 촉발되는 경영 환경상의 변화에 적응하고 선제적으로 대응하여, 비즈니스 경쟁력을 근본적으로 제고하거나, 신규 비즈니스 모델을 만들어 새로운 성장동력을 확보하기 위한 활동을 의미
- (Verhoef 외, 2021) 디지털화는 현재의 비즈니스 프로세스에 IT나 디지털 기술을 도입하는 것인 반면, 디지털 트랜스포메이션(Digital Transformation)은 기업의 가치창조 프로세스에 근본적인 변화를 가져오는 것
- (Tabrizi 외, 2019) 디지털 전환은 (디지털 기술을 기반으로) 기존 비즈니스 모델을 혁신하는 것
- (한국정보화진흥원, 2019) 인공지능 (AI), 클라우드 (Cloud), 데이터 (Data) 등 디지털 기술기반 비즈니스 모델을 중심으로 산업 구조 재편하는 것
- (한국무역협회, 2019) 디지털 기반으로 고객경험, 운영.관리프로세스, 비즈니스 모델 등을 변화시키는 경영전략

- (임희종 외, 2021) 새로운 디지털 기술을 활용하여 고객데이터를 분석, 고객지향 가치를 창출하고, 이를 바탕으로 비즈니스 프로세스를 개선하거나, 새로운 비즈니스 모델을 만들거나, 새로운 비즈니스 기회를 통해 성장 동력을 만드는 것
- (신동수 외, 2021) 생산 유통 등 경제활동 전반이 정보통신기술과 데이터를 기반으로 이루어지는 체제로 탈바꿈하는 과정
- (김용진, 2018) 디지털 전환은 사물과 사물의 커뮤니케이션, 정보의 실시간 축적 및 분석, 제품의 서비스화 및 서비스의 제품화를 가져오는 기반
- (IBM, 2011) 디지털과 물리적인 요소들을 통합해 비즈니스 모델을 변화하고 산업에 새로운 방향을 정립하는 것
- (Agile Elephant, 2015) 자산(assets)의 디지털화와 조직의 생각하고 일하는 방식을 바꾸는 프로세스 전환, 리더십과 신규비즈니스 모델의 창출 그리고 이해관계자, 고객, 직원 등의 경험을 향상시키기 위한 기술의 활용까지 포괄하는 개념
- (IDC, 2015) 기업이 새로운 비즈니스 모델, 제품 및 서비스를 창출하기 위해 디지털 역량을 활용함으로써 고객 및 시장(외부 생태계)의 파괴적인 변화에 적응하거나 이를 추진하는 지속적인 프로세스
- (A. T. Kearney, 2016) ICBM (IoT, cloud, bigdata, mobile)+AI 등 디지털 신기술로 촉발되는 경영 환경상의 변화 동인에 선제적으로 대응함으로써 현행 비즈니스의 경쟁력을 획기적으로 높이거나 새로운 비즈니스를 통한 신규 성장을 추구하는 기업 활동
- (Altimeter Group, 2017) 끊임없이 변화하는 디지털 경제에서 효과적으로 경쟁하기 위해 고객과 직원을 위해 새로운 가치를 창출하기 위한 기술, 비즈니스 모델 및 프로세스의 재편성 또는 신규투자
- (김승래, 2021) 디지털 전환은 기업에서 사물인터넷(IoT), 클라우드 컴퓨팅, 인공지능(AI), 빅데이터 솔루션 등 정보통신기술(ICT)을 플랫폼으로 구축·활용하여 기존 전통적인 운영 방식과 서비스 등을 혁신하는 것을 의미 (디지털과 물리적 요소들의 통합, 전체 산업에 새로운 방향을 정립하는 것)
- (Vial, 2019) 정보, 컴퓨팅, 통신 및 연결 기술의 조합을 통해 속성에 대한 상당한 변경을 유발하여 개체를 개선하는 것을 목표로 하는 프로세스

- (디지털 전환의 정의) 디지털 기술(정보통신기술 기반)을 도입 또는 활용하여 조직/산업/사회의 경쟁력을 증가시키는 일련의 프로세스
- (디지털 전환 구성 기술/기술 범주) 인공지능, 빅데이터, 클라우드 컴퓨팅, 5G 네트워크, 사물인터넷, 블록체인, 컴퓨팅 파워

(2) 디지털 전환의 경제적/사회적 파급효과

□ 디지털 전환의 영향은 광범위하게 나타나고 있으며 본 절에서는 문헌연구를 통하여 디지털 전환의 영향을 경제적/사회적 파급효과 측면에서 정리

- (사업체계 재편) 디지털 전환은 기존 사업의 value chain에 ICT 기술을 접목하여 사업체계 전반을 재편시킴(예: 스타벅스의 디지털 벤처 부서 신설과 시스템의 디지털화, GE, 아디다스 등 전통 제조기업의 제조 공정 디지털화 등) (장훈, 2017)
- (운영 프로세스 자동화) 기업은 Robotic Process Automation(RPA) 도입으로 생산성 향상과 기업의 응답성(Responsiveness)이 향상되며, 민첩성(Agility) 향상을 도모할 것이며(Büyüközkan and Göçer 2018; Hartley and Sawaya 2019) 이를 통해 인건비 절감뿐만 아니라, 일관된 프로세스의 실행이 가능하게 될 것(임희종 외, 2021)
- (공급망 사슬 통합) 디지털 전환이 진행되면서, 공급망의 디지털화가 고도화되며, 정보의 공유 및 의사결정 조율의 수준이 높아질 것이고(Büyüközkan and Göçer 2018), 이러한 정보의 공유는 통합의 기초가 되며, 다음 단계로 조율(coordination)과 자원의 공유, 그리고 조직의 연결을 통해 리스크, 비용, 이익을 공유하게 될 것(Lee 2000)
- (프로세스 효율화) 기업은 디지털 전환을 통해 기존 아날로그 기반 또는 디지털화 기반의 프로세스를 혁신시켜 프로세스 전반의 효율성을 극대화(임희종 외, 2021)
- (BM 변화) 제품/서비스가 디지털화되고 전달체계가 디지털화되며, 생산 운영 체계가 디지털화되면 이들 비즈니스들은 플랫폼 기술을 요구하게 되고, 플랫폼 기술은 초연결성과 초지능성에 기반을 두기 때문에 사물인터넷(IoT), 클라우드, 인공지능 등을 통한 발전으로 O2O(Online to Offline) 등 새로운 '스마트 비즈니스 모델'의 등장을 촉진(김용진, 2018)
- (플랫폼 경제시대 도래) 플랫폼 경제는 제4차 산업혁명의 핵심, 플랫폼 서비스로서 디지털 기술과 네트워크를 기반으로 각 경제 주체 간에 다양한 생산과 소비가

이루어지는 것을 의미하며, 향후 디지털 전환을 통해 플랫폼 경제 확대로 사회, 정치, 경제 전반에 있어 기술적 확장과 상호 호환성 확보에 대한 필요성이 증가할 전망(김승래, 2021)

- (융합적 인재 중요성 증대) 디지털 전환은 인지능력과 비인지적/사회적 능력을 모두 갖춘 인재를 필요로 하므로 ICT 기술, 계량화 역량, 수리적 역량뿐만 아니라 자기조직화(self organization), 관리 능력, 소통 역량을 갖춘 인재의 가치가 증가할 것(이명화 & 최용인, 2017)
- (디지털 혁신생태계의 발달) 디지털 혁신생태계란 기존 산업생태계가 디지털 환경 기반으로 전환되면서 발전하는 것을 의미하고, 디지털 전환은 이러한 디지털 혁신생태계 발달을 촉진(플랫폼이나 시스템을 소유한 기업들이 생태계 리더 그룹으로 격상되고, 타 산업과의 연계/융합을 용이하게 하여 이종 산업생태계로의 진입/진출을 촉진시킴) (김승현, 2020)
- (조직문화 변화) 디지털 전환에 따라 조직원들이 갖추어야 할 필수 역량에 변화가 생기고 있음(관계적 민첩성(Relational Agility), 탄력성(Resilience) 등의 행동 양상과 인지/학습 유연성(Cognitive & Learning Flexibility) 등의 인지적 역량, 감성 조절성(Regulating Intense, Conflicting Emotion) 등 감성적 역량 등이 더욱 중요해지고 있음(Ashford et al. 2018; Petriglieri et al. 2018)). (임희종 외, 2021)
- (조직형태 및 구조 변화) 디지털 전환 가속화로 가상 협업, 온라인 커뮤니티, 오픈 소스 프로젝트, 긱 이코노미 등과 같은 새로운 조직 형태가 등장(Lomi et al. 2014).

3 디지털 전환 관련 미래사회 위험이슈

(1) 디지털 전환의 미래사회 위험이슈 개요

□ 디지털 전환은 인공지능 기술을 중심으로 잠재적 위험 이슈들이 예측되어 왔고, 일부는 이미 현실화되고 있음

- 디지털 전환은 순기능뿐만 아니라 역기능을 초래할 위험이 존재하고, 그중 일부는 이미 현실에서 문제로 나타나고 있음
- 아울러 이러한 역기능은 디지털 전환의 주요 기술들 中 핵심기술인 인공지능과 관련하여 발현되고 있음
- 이에 본 장에서는 문헌연구를 통해 디지털 전환 역기능의 주요 유형을 정리하고, 정량분석(Embedded Topic Modeling)을 통하여 국내와 주요국에서 어떠한 세부 문제들이 나타나고 있는지를 규명

(2) 디지털 전환의 역기능 유형 (문헌연구)

□ (KISTEP, 2020) Stanford Encyclopedia of Philosophy의 'Ethics of AI and Robotics'에서 언급한 10가지 이슈 中 디지털 전환과 연관된 5개 역기능을 유형화

<표 1> 디지털 전환의 역기능 유형

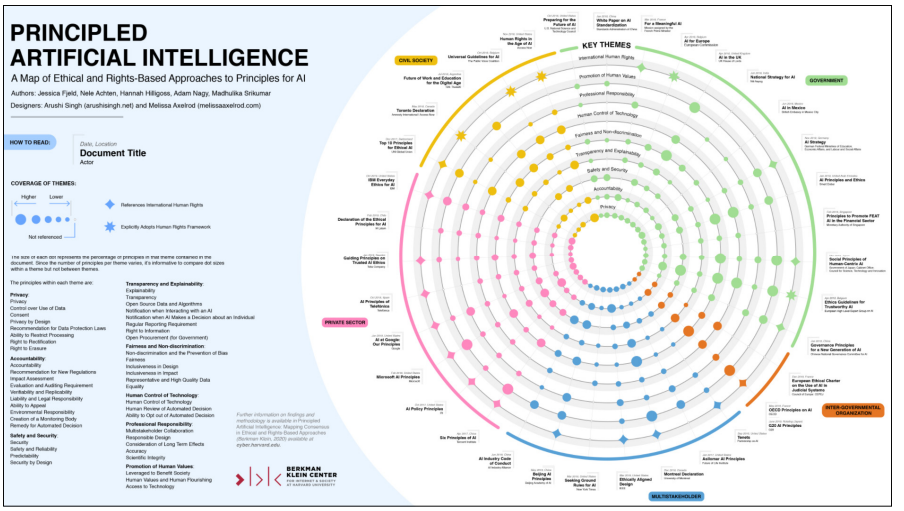
역기능 유형	개요 및 사례
프라이버시 및 감시	• 개인정보 데이터, 개인식별정보, 국가안보 데이터 등의 유출로 인한 문제 야기 • (사례) TikTok 어플리케이션은 Google에서 금지한 방법으로 사용자 개인정보 데이터를 무단으로 축적
의사결정 편향	• 인공지능 학습 데이터의 편향성으로 의사결정 결과 왜곡 또는 불공정성 야기 • (사례) 범죄자 예측 소프트웨어가 유색인종에 불리한 결과를 도출
조작	• 인공지능 기술 기반의 deepfake, psychological targeting을 통한 결과 조작으로 정치/국가 안보 위협 • (사례) 인공지능 기술을 통한 유명인 포르노 합성 제작/유포, SNS를 통한 개인의 정치성향 파악 및 편향된 정보 제공을 통한 선거 마케팅
자율 시스템의 불투명성	• 자율 시스템의 오작동, 인공지능 오용에 따른 잘못된 진단 결과 도출로 인한 문제 야기 • (사례) 테슬라 자율주행 시스템 오류로 인한 운전자 사망 등
자동화 및 고용	• 자동화된 거래에 따른 시장 변동성 증대, 인공지능 기반 작업으로 인한 사회/경제적 불평등 심화 문제 야기 • (사례) 알고리즘 트레이딩으로 인한 시장 변동성 가중

출처: 2020년 과학기술혁신정책 핵심이슈 발굴 및 인텔리전스 기능 강화 연구 (KISTEP)

□ (메타분석) Harvard Berkman Klein Research Center는 `20년 「Principled AI」를 통해 36개의 유명한 AI 원칙 관련 문서를 대상으로 메타분석을 수행하여 8개 주요 이슈들*을 도출

* 프라이버시, 책임성, 안전 및 보안, 투명성 및 설명 가능성, 공정성 및 차별 금지, 기술에 대한 인간의 통제, 직업적 책임, 인간 가치 증진

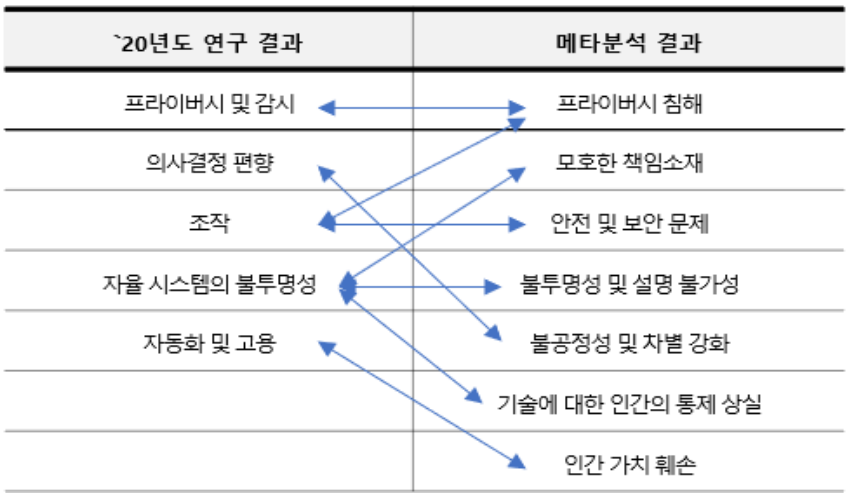
- 해당 8개 이슈들은 인공지능 기술로부터 파생될 수 있는 역기능들에 대응하기 위해 주요 국가/기관들이 설정하고 있는 인공지능 개발 가이드라인의 주요 항목을 의미
- 이를 통해 추정할 수 있는 디지털 전환 역기능 유형은 다음과 같음
  - 프라이버시 침해, 모호한 책임소재, 안전 및 보안 문제, 불투명성 및 설명 불가성, 불공정성 및 차별 강화, 기술에 대한 인간의 통제 상실, 인간 가치 훼손



출처 : https://wilkins.law.harvard.edu/misc/PrincipledAI_FinalGraphic.jpg

[그림 2] 「Principled AI」주요 내용

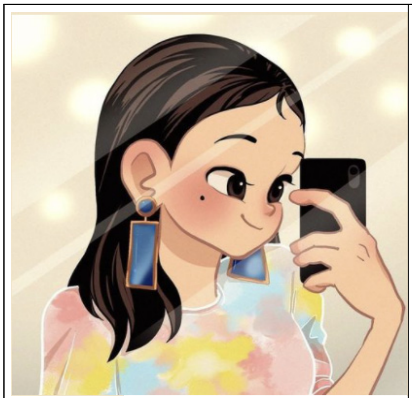
☞ (종합) 상기 2개 분류를 비교/종합해보면 [그림 3]과 같이 각 유형들이 매칭됨을 확인할 수 있음[4)]



[그림 3] 유형 매칭 결과

(3) 디지털 전환의 역기능 세부이슈 분석 (국내: 이루다 케이스)

□ 본 절에서는 국내에서 사회적으로 가장 크게 이슈가 됐던 디지털 전환의 역기능 사례('이루다' 케이스)를 정량분석하여 세부 위험 이슈들을 도출



인공지능 챗봇 서비스 '이루다'

- (개요) '이루다'는 스캐터랩(ScatterLab) 소속의 핑퐁팀이 개발한 딥러닝 알고리즘 기반의 열린 주제 대화형 인공지능 챗봇 서비스로 '20년 말에 공개 후 '21년 1월 초 사용자 수 약 40만명을 확보
- (논란) 개발사의 의도와 달리 개인정보 침해, 상식적 답변 오류, 특정 성별 혐오, 외설적 목적 사용 등의 문제 발생
- (경과) 각종 논란으로 '21년 1월 11일 서비스 중단, '21년 4월 개인정보보호위원회가 개인정보보호법 위반으로 '이루다' 개발사 스캐터랩에 총 1억330만원의 과징금 및 과태료 부과

4) 본 연구의 4장 인식조사에서는 5개 유형을 활용하여 설문조사 수행

- (방법: Embedded Topic Modeling) BERT 아키텍처에 사전학습된 언어모델(distiluse-base-multilingual-cased)을 활용한 LDA 토픽모델 활용
- (데이터) 2020.11.01 ~ 2021.04.30. 기간에 이루다 사건을 다룬 국내 언론 기사/사설 1,329개(전처리* 후 데이터: 1,100개)

  * ① 중복기사제거, ② 주제 연관성 낮은 기사 제거, ③ Mecab 토크나이저를 활용한 세그먼테이션, ④ BPE 기반 서브워드임베딩 적용

- (목적) 현재까지 국내에서 디지털 전환의 역기능 측면에서 가장 크게 이슈가 된 '이루다' 케이스를 정량분석하여 사회적으로 논란이 된 세부 위험 이슈들을 도출 및 분석
- (결과) 총 22개 세부주제들이 도출됐고, 연관주제별 군집화 및 연관성 높은 언론기사 내용 매칭을 통한 해석 작업을 거쳐 최종적으로 11개의 세부주제들과 이에 매칭된 키워드들을 정리

<표 2> '이루다' 관련 디지털 전환의 역기능 위험 세부주제 도출 결과

세부주제	관련 키워드	빈도
인공지능 윤리문제	인공지능, 윤리, 문제, 논란, 사회	293
혐오/성희롱 문제	혐오, 문제, 차별, 여성	67
혐오/차별 논란	문제, 차별, 사회, 혐오, 논란	59
(개인정보) 이용행위 과징금 부과	카카오톡, 이용, 과징금, 위반, 부과, 행위	35
(인공지능 기술의) 인권침해	침해, 인권위, 인공지능, 기술, 개발	33
인공지능 윤리 연구 필요성 제기	윤리, AI, 교수, 세미나, 토론, 연구	32
인공지능 윤리규범 마련	AI, 보호, 제도, 마련, 추진	27
개인정보 활용/보호	데이터, 개인, 정보, 활용	24
차별/증오 발언 근절	발언, 근절, 수립, 대응, 배척	16
(개인정보) 비식별 처리	이용자, 비식, 상시, 처리	15
(개인정보 무단 유출에 따른) 집단 소송/손해배상 청구	소송, 청구, 법무법인, 이용자, 제기, 무단	12

※ 관련 언론기사 임베디드 토픽 모델링 분석 결과

□ (결과분석) 이루다 케이스 정량분석 결과를 본 연구에서 상정한 8가지 디지털 전환 역기능 유형과 매칭한 결과 불공정성 및 차별강화, 프라이버시 침해 유형이 가장 많이 언급된 것을 확인

- '이루다'가 채팅 상황에서 특정 성별을 혐오했던 사례가 언론에서 가장 많이 지적되었고, 이는 불공정성 및 차별 강화 유형에 해당

● 아울러 이루다에서 학습에 활용됐던 비식별 처리 되지 못한 개인정보가 두 번째로 많이 지적되었음을 확인하였고, 이는 프라이버시 침해 유형에 해당

<표 3> '이루다' 관련 디지털 전환의 역기능 위험 세부주제 도출 결과와 디지털 전환 역기능 유형 매칭 결과

세부주제	디지털 전환 역기능 유형	빈도
인공지능 윤리문제	*(일반적인 사항이므로) 매칭 제외*	293
혐오/성희롱 문제	불공정성 및 차별 강화	67
혐오/차별 논란	불공정성 및 차별 강화	59
(개인정보) 이용행위 과징금 부과	프라이버시 침해	35
(인공지능 기술의) 인권침해	인간 가치 훼손	33
인공지능 윤리 연구 필요성 제기	*(일반적인 사항이므로) 매칭 제외*	32
인공지능 윤리규범 마련	*(일반적인 사항이므로) 매칭 제외*	27
개인정보 활용/보호	프라이버시 침해	24
차별/증오 발언 근절	불공정성 및 차별강화	16
(개인정보) 비식별 처리	프라이버시 침해	15
(개인정보 무단 유출에 따른) 집단 소송/손해배상 청구	프라이버시 침해	12

(4) 디지털 전환의 역기능 세부이슈 분석 (주요국 언론기사 정량분석)

□ 본 절에서는 주요국을 중심으로 언론기사에서 다루어진 디지털 전환의 주요 기술별/국가별 역기능 세부 위험 이슈들을 도출

● (방법: Embedded Topic Modeling) BERT 아키텍처에 사전학습된 언어모델(distiluse-base-multilingual-cased)을 활용한 LDA 토픽모델 활용

● (데이터) 2020.01.01 ~ 2021.08.23. 기간에 (미국, 중국, 유럽, 일본에서) 보도된 디지털 전환의 주요기술(인공지능, 빅데이터, 5G 네트워크, IoT, 블록체인)의 위험/역기능/문제 등을 다룬 언론기사/사설 2,175,893개(전처리 후 데이터: 349,076개)

● (목적) 주요국에서 논란이 되고 있는 디지털 전환의 세부 위험 이슈들을 정리하고, 나아가 디지털 전환의 기술별/국가별 세부 위험 이슈들을 정리

● (결과 1: 종합) 총 209개 세부주제들이 도출됐고, 연관주제별 군집화 및 연관성 높은 언론 기사 내용 매칭을 통해 9개 세부 이슈들을 정리

- (결과해석 1) 디지털 전환 역기능은 '무기를 위한 AI 사용', 'AI 프로파일링

편향', 'deepfake' 이슈, '보안' 이슈, '데이터 프라이버시 위험', '암호화폐 보안문제', '암호화폐 채굴로 인한 탄소배출 문제', 'AI 기반 대출 심사의 차별 문제'였음

- (결과해석 2) 상기 결과를 본 연구의 디지털 전환 역기능 유형에 매칭하여 해석 시 안전 및 보안 문제, 불공정성 및 차별 강화, 프라이버시 침해 유형이 가장 많이 회자됐음을 확인할 수 있음

<표 4> 임베디드 토픽 모델링 분석 결과(종합)

Label	Description	Total Count
military use of AI	AI for weapon	190
bias	biased AI profiling	127
deepfake	Facebook banning deepfakes	71
security	IoT vulnerabilities	63
ai ethics	Ethics in AI systems	62
privacy	Data privacy risk	41
hacker	crypto currency security (broken by hackers)	33
bitcoin sustainable	Bitcoin mining and carbon emission problem	24
AI discrimination	AI discrimination in credit lending	11

- (결과 2: 세부이슈별 기술분류) 도출된 9개 세부 이슈별 연관된 디지털 전환 기술들의 빈도분석을 수행한 결과

<표 5> 임베디드 토픽 모델링 분석 결과(종합-세부이슈별 기술분류)

Label	Description	5G networks	IoT	AI	big data	block chain
military use of AI	AI for weapon	1	10	175	2	2
bias	biased AI profiling		10	107	7	3
deepfake	Facebook banning deepfakes			66	0	5
security	IoT vulnerabilities		60	2	0	1
ai ethics	Ethics in AI systems			61	1	0
privacy	Data privacy risk			17	12	12
hacker	crypto currency security (broken by hackers)			3	0	30
bitcoin sustainable	Bitcoin mining and carbon emission problem			3	0	21
AI discrimination	AI discrimination in credit lending			11	0	0

- (결과해석) 디지털 전환 역기능의 세부 이슈들과 연관된 기술들을 분석한 결과 인공지능 기술이 가장 많이 연관되어 있었고, 다음으로는 IoT 기술이었음

※ 디지털 전환에 있어 인공지능 기술이 핵심기술로서 가장 많이 활용되기에 역기능 연관성도 가장 높았으며 IoT의 경우 다른 디지털 전환 구성기술 대비 실제 활용 제품/서비스와 가까운 기술이므로 역기능 연관성이 높게 나타난 것으로 추정됨

● (결과 3: 세부이슈별 국가분류) 9개 세부 이슈별 연관된 국가들의 빈도분석을 수행한 결과

<표 6> 임베디드 토픽 모델링 분석 결과(종합-세부이슈별 국가분류)

Label	Description	CN	EU	JP	US
military use of AI	AI for weapon	0	13	0	177
bias	biased AI profiling	3	22	2	100
deepfake	Facebook banning deepfakes	0	4	1	66
security	IoT vulnerabilities	0	14	0	49
ai ethics	Ethics in AI systems	0	5	0	57
privacy	Data privacy risk	0	4	0	37
hacker	crypto currency security (broken by hackers)	0	9	0	24
bitcoin sustainable	Bitcoin mining and carbon emission problem	0	7	0	17
AI discrimination	AI discrimination in credit lending	0	0	0	11

- (결과해석 1) 빈도분석 결과가 미국에 편중된 것으로 도출됐으나 이는 언론기사 구성이 영어기반의 미국 언론기사를 중심으로 구성되었기 때문
- (결과해석 2) EU와 미국의 각 이슈별 빈도분석 결과를 살펴보면 비슷한 비중을 확인할 수 있으나 유럽에서는 EU의 AI 규제와 IoT 기기의 보안 취약성 이슈가 미국에 비해 상대적으로 높은 비중으로 이슈화됐음을 확인

● (결과 4: 국가별 세부 위험이슈-미국) 미국 언론기사에 한정하여 디지털 전환 역기능의 세부 이슈들을 도출한 결과

- (결과해석 1) 디지털 전환의 역기능과 관련한 세부 이슈들 중에서는 IoT 기기의 사이버 보안, 사이버 공격 이슈가 가장 큰 비중을 차지
- (결과해석 2) 이외에는 무인 드론의 위험, 인종/성별 편향, 통신 보안, 개인정보/프라이버시 등이 큰 비중을 차지

- (결과해석 3) 디지털 전환의 세부 이슈들을 본 연구의 역기능 유형과 매칭한 결과, 안전 및 보안 문제가 가장 큰 비중을 차지하고 있음을 확인

<표 7> 미국의 디지털 전환 관련 세부 위험이슈 도출 결과

세부주제명	관련 키워드	빈도
(IoT 기기) 사이버 보안, 사이버 공격	cybersecurity, cyber, intrusion, attacks, ransomware, iot, vulnerabilities, malware	287
군사용 무인 드론	military, drones, unmanned, weapons, combat, enemy, warfare	207
영국의 화웨이 기기 금지	huawei, uk, telecoms, ban, suppliers	164
(Google의) 피부색에 따른 인종/성별 편향	skin, bias, Google, gender, biased, photos, images, color	72
(Facebook의) Deepfake video	deepfakes, videos, Facebook, disinformation, misinformation, factcheckers	55
미국의 Tiktok 사용 금지	tiktok, chinese, usgovernment, ban, restrictions	43
(감청, 스파이 행위 등에 의한) 독일의 화웨이 사용 금지	germany, Merkels, sabotage espionage, parliament, misused, denies, spying	35
경찰의 안면인식 카메라 활용에 따른 프라이버시 침해 논란	facial, recognition, police, enforcement, mask, surveillance, cameras, privacy	32

<표 8> 미국의 디지털 전환의 역기능 위험 세부주제 도출 결과와 디지털 전환 역기능 유형 매칭 결과

세부주제명	디지털 전환 역기능 유형	빈도
(IoT 기기) 사이버 보안, 사이버 공격	안전 및 보안 문제	287
군사용 무인 드론	기술에 대한 인간의 통제 상실, 인간 가치 훼손	207
영국의 화웨이 기기 금지	안전 및 보안 문제	164
(Google의) 피부색에 따른 인종/성별 편향	불공정성 및 차별 강화	72
(Facebook의) Deepfake video	프라이버시 침해, 안전 및 보안 문제	55
미국의 Tiktok 사용 금지	안전 및 보안 문제	43
(감청, 스파이 행위 등에 의한) 독일의 화웨이 사용 금지	안전 및 보안 문제	35
경찰의 안면인식 카메라 활용에 따른 프라이버시 침해 논란	프라이버시 침해	32

- (결과 5: 국가별 세부 위험이슈-EU) 유럽 언론기사에 한정하여 디지털 전환 역기능의 세부 이슈들을 도출한 결과
  - (결과해석) EU의 (한정된[5]) 언론기사들을 분석한 결과 안전 및 보안 이슈가 압도적으로 많은 비중을 차지

<표 9> 유럽의 디지털 전환 관련 세부 위험이슈 도출 결과

세부주제명	관련 키워드	빈도
IoT 기기의 보안 이슈	iot, devices, security, device, data, network, connectivity	156
영국의 화웨이 5G 네트워크 사용 관련 의사 결정	Huawei, 5g, networks, uk, government, telecoms, decision	148

- (결과 6: 기술별 세부 위험이슈-5G Networks) 5G Network과 관련한 주요국의 언론기사를 분석하여 디지털 전환의 역기능 세부 이슈를 분석한 결과
  - (결과해석) 5G Network 관련 기사의 디지털 전환 역기능 세부 유형은 주로 5G 네트워크 기기의 보안과 관련한 위험이었고, 이는 화웨이 기기의 국가 안보위협 이슈가 전국적으로 이슈화되어 국가 차원에서 이를 금지하게된 사건에 기인한다고 예상

<표 10> 5G Networks 관련 세부 위험이슈 도출 결과

세부주제명	관련 키워드	빈도
미국의 화웨이 5G 네트워크 사용 관련 보안 이슈	5G, US, networks, security, government, Chinese	295
스웨덴의 화웨이 5G 네트워크 사용 관련 이슈	Swedish, Huawei, 5G, networks, ZTE, Chinese	63
화웨이 네트워크 장비 보안 관련 위협/사용 금지	equipment, Huawei, networks, security, ban, fears, 5G, stop	29
EU의 화웨이 5G 장비 공급 관련 보안/위험 이슈	EU, 5G, suppliers, Huawei, risks, security	23

5) 유럽 국가의 영어기사 분량은 미국 대비 월등히 부족하였기에 상대적으로 결과를 일반화하기는 어렵다는 점에 주의

● (결과 7: 기술별 세부 위험이슈-AI) 인공지능과 관련한 주요국의 언론기사를 분석하여 디지털 전환의 역기능 세부 이슈를 분석한 결과

<표 11> AI 관련 세부 위험이슈 도출 결과

세부주제명	관련 키워드	빈도
EU의 생체 정보 사용 규제	EU, Vestager, regulation, commission, regulations, ban, biometric	232
사이버 안보, 사이버 공격/위협	security, cybersecurity, intrusion, threats, attacks, cyberattacks, risk	165
경찰의 감시 카메라 사용에 대한 인종차별, 편향성, 프라이버시 침해	police, recognition, crime, cameras, surveillance, bias, privacy	90
Facebook의 deepfake 비디오 사용을 통한 허위정보, 오보 문제	deepfakes, videos facebook, disinformation, misinformation	68
인종/성 편향 알고리즘 이슈	bias, algorithms, black, gender, race	63

- (결과해석 1) 인공지능 관련 기사의 디지털 전환 역기능은 예상대로 他 디지털 전환 구성 기술 대비 다양한 세부 유형들이 도출되었음
- (결과해석 2) 역기능 관련 세부 이슈 중에서는 생체 데이터 활용 문제가 가장 큰 비중을 차지하였음
- (결과해석 3) 다음으로 안보/보안 위험, 프라이버시 침해, 편향성 이슈가 디지털 전환 위험 세부이슈의 비중을 차지하였음
- (결과해석 4) 상기 세부 이슈들을 본 연구의 역기능 유형과 매칭한 결과, 프라이버시 침해 유형이 가장 큰 비중을 차지하고 있음을 확인

<표 12> AI 관련 디지털 전환의 역기능 위험 세부주제 도출 결과와 디지털 전환 역기능 유형 매칭 결과

세부주제명	디지털 전환 역기능 유형	빈도
EU의 생체 정보 사용 규제	프라이버시 침해	232
사이버 안보, 사이버 공격/위협	안전 및 보안 문제	165
경찰의 감시 카메라 사용에 대한 인종차별, 편향성, 프라이버시 침해	프라이버시 침해, 불공정성 및 차별 강화	90
Facebook의 deepfake 비디오 사용을 통한 허위정보, 오보 문제	프라이버시 침해, 안전 및 보안 문제	68
인종/성 편향 알고리즘 이슈	불공정성 및 차별 강화	63

- (결과 9: 기술별 세부 위험이슈-blockchain) 블록체인과 관련한 주요국의 언론 기사를 분석하여 디지털 전환의 역기능 세부 이슈를 분석한 결과
  - (결과해석) 블록체인 관련 기사의 디지털 전환 역기능 세부 유형은 주로 암호화폐 자체의 보안 관련 위험이었고, 이는 아직 블록체인 기술이 상품/서비스로 가시화된 사례가 암호화폐밖에는 없는 것에 기인하는 결과로 예상

<표 13> Blockchain 관련 세부 위험이슈 도출 결과

세부주제명	관련 키워드	빈도
해커/랜섬웨어에 의한 (암호화폐) 보안/도난 이슈	ransomware, hacker, stolen, security, cryptocurrency	46

※ 관련 언론기사 임베디드 토픽 모델링 분석 결과

(5) 소결

□ (의의) 본 장에서는 디지털 전환의 미래사회 위험이슈를 구체화하였음

- 현재까지 디지털 전환 미래사회 위험이슈(역기능)이 어떠한 형태로 발현될지에 대한 논의가 주를 이루었다면
- 본 연구는 한발 더 나아가 정량분석(언론기사 빅데이터를 Embedded Topic Modeling으로 분석)하여 실제로 우리/주요국 사회에서 일어나고 있는 또는 논의되고 있는 디지털 전환 미래사회 위험의 세부 이슈를 확인하였음
- 이를 통해 실제적인 미래사회 위험 세부이슈들을 구체화하였음

□ (시사점 1) 향후 디지털 전환의 미래사회 위험이슈 대응을 위해서는 인공지능 기술로부터 파생할 수 있는 역기능에 주목할 필요

- 디지털 전환의 미래사회 위험이슈는 인공지능 기술을 중심으로 논의되고 있음
- 이는 인공지능 기술이 다른 디지털 전환 구성기술*에 비해 상품/서비스 설계 및 구현에 기여하는 바가 높기 때문임

  * blockchain, big data, IoT, cloud computing, 5G networks
- 따라서 정부는 다른 디지털 전환 구성기술보다는 인공지능 기술에 집중하여 역기능 대응을 위한 정책을 고민할 필요가 있음

□ (시사점 2) 위험이슈 대응의 우선순위 설정 시 국가/사회적 맥락을 확인할 필요

- 정량분석 결과에서 알 수 있듯이 국가별로 가장 많이 이슈/논의된 디지털 전환의 위험 유형이 상이*하였음

* 미국은 안전 및 보안, 불공정성 및 차별 강화 문제가, EU는 프라이버시 이슈, 안전 및 보안 이슈가 가장 크게 사회적으로 논의가 되었음을 확인

- 따라서 정부도 디지털 전환 위험 이슈 대응 정책 설계 시 사회적 맥락을 고려하여 우선 대응 영역을 설정할 필요가 있음
- 이후 중장기적으로 위험 이슈 전반에 대응할 수 있는 방안을 모색하는 절차로 정책을 추진할 필요

4 디지털 전환 관련 미래사회 위험에 대한 인식조사

(1) 인식조사 개요

□ 본 인식조사는 디지털 전환 역기능 대응 정책수립에 필요한 인식 데이터 및 정책 수요 데이터 확보를 위한 목적으로 국내 재직 중인 성인을 대상으로 구조화된 설문조사를 설계 및 실시

- (목적) 디지털 전환 역기능 대응 정책 수립에 필요한 인식 데이터와 정책 수요 데이터를 확보
- (조사대상) 국내 기업에 재직 중인 성인 1,824명

  ※ (산업) 제조업, 건설업, 유통업, 서비스업, ICT, 금융업, 의료 및 보건, 기타

  ※ (연령) 20대~50대 이상
- (조사기간) '21년 11월 11일 ~ `21년 12월 06일 (약 4주)
- (조사방법) 구조화된 설문지를 활용한 온라인 패널조사

<표 14> 응답자 구성

구분	20대	30대	40대	50대 이상	계
제조업	59	87	93	93	332
건설업	16	34	74	67	191
유통업	10	54	76	71	211
서비스업	54	75	98	95	322
ICT	12	42	64	40	158
금융업	15	23	36	27	101
의료 및 보건	23	34	54	54	165
기타	62	78	95	109	344
계	251	427	590	556	**1,824**

(2) 인식조사 구성

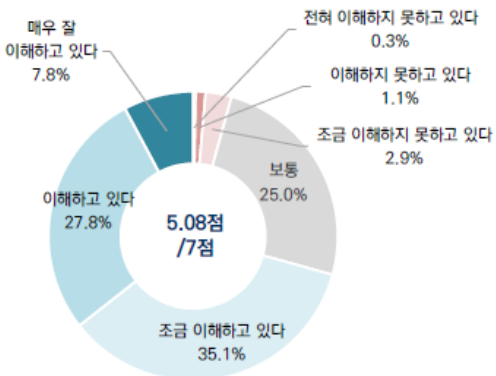
<표 15> 조사항목 및 내용

조사항목	내용
Demographic Variables (1)	응답자 이름, 연령, 성별
Demographic Variables (2)	종사 산업, 직급, 기업유형, 근속연수
기타 통제변수	성격, 성향
디지털 전환에 대한 일반적 인식	디지털 전환에 대한 용어 이해도
	국내 경제/산업 발전에 디지털 전환이 필요하다고 생각하는 정도
	기업에 대한 긍정적/부정적 영향도
디지털 전환 위험에 대한 인식	개인정보보호 및 사생활 침해 이슈에 대한 위험성
	의사결정 편향 이슈에 대한 위험성
	조작 이슈에 대한 위험성
	자율 시스템의 투명성 이슈에 대한 위험성
	자동화 및 고용 이슈에 대한 위험성
	디지털 전환에 의한 위험들 중 가장 위험하다고 생각하는 것
	그 외 심각한 디지털 전환 위험 요인(서술)
디지털 전환 위험 대응 정책 수요	개인정보보호 강화의 필요성
	책임성 강화에 대한 필요성
	안전과 보안 강화에 대한 필요성
	투명성 및 설명가능성 확보에 대한 필요성
	공정성과 평등 추구에 대한 필요성
	기술에 대한 인간의 통제 확보의 필요성
	전문적인 책임 강화의 필요성
	인간가치의 증진에 대한 필요성
	정책적 대응 중 가장 필요한 정보
	그 외 디지털 전환 관련 위험에 대응하기 위한 수단(서술)

(3) 인식조사 결과

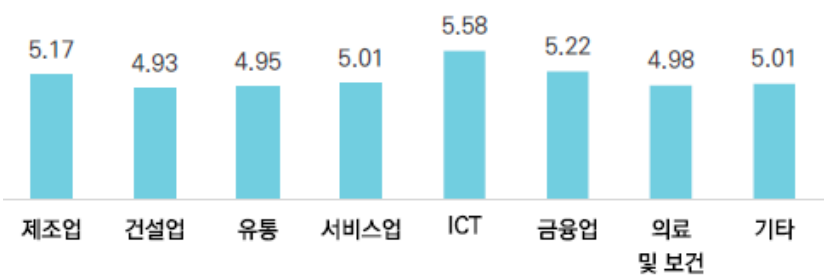
□ 디지털 전환에 대한 용어 이해도

- (전체) 디지털 전환 용어에 대한 이해도는 7점 평균 5.08점으로 높았고, 이해하고 있다(Top 3) 응답 비중이 70.7%였음



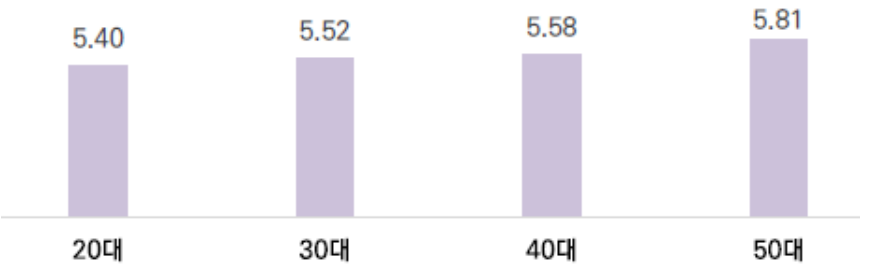
[그림 4] 디지털 전환에 대한 용어 이해도 (전체)

- (산업별) 디지털 전환에 대한 이해도는 산업별로 극명한 차이는 없었으나 ICT(5.58점) 및 제조업(5.17점) 종사자들이 타 산업 종사자보다 이해도가 높았음
  - 해당 결과는 ICT 및 제조업이 타 산업보다 디지털 전환이 상대적으로 빠르게 일어나고 있는 것에 기인했다고 예상



[그림 5] 디지털 전환에 대한 이해도 (종사산업별)

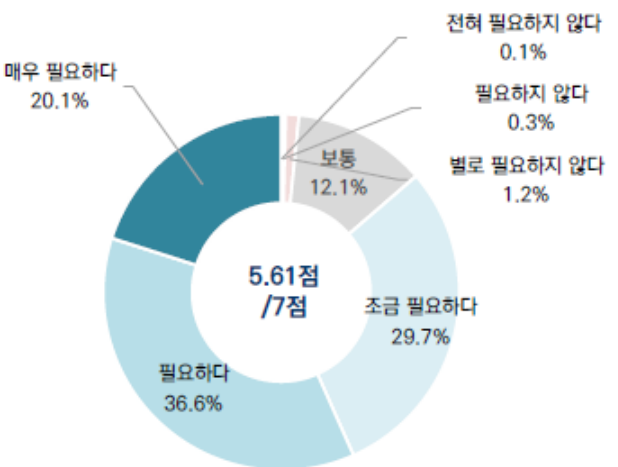
● (연령별) 디지털 전환에 대한 이해도는 연령별로 큰 차이가 없었으나 50대(5.81점)가 타 연령대 대비 약간 더 높은 이해도를 보유하고 있었음



[그림 6] 디지털 전환에 대한 이해도 (연령대별)

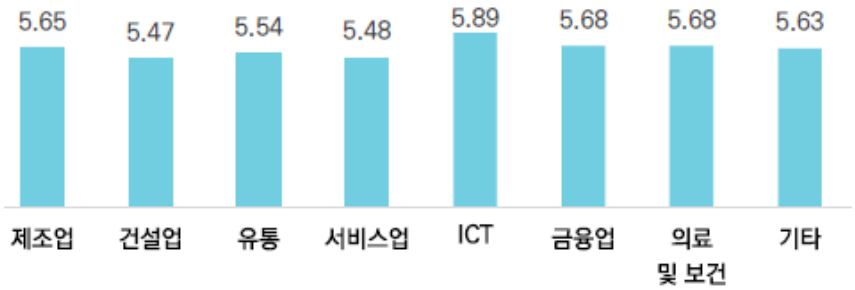
□ 국내 경제/산업 발전에 디지털 전환이 필요하다고 생각하는 정도

● (전체) 디지털 전환이 국내 경제/산업 발전에 필요하다고 인식하다는 정도는 7점 평균 5.61점으로 높게 나타났고, 필요하다(Top 3)는 응답비중은 86.4%였음



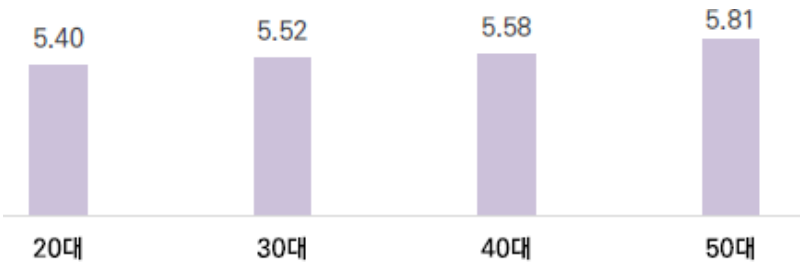
[그림 7] 국내 경제/산업 발전에 디지털 전환이 필요하다고 생각하는 정도 (전체)

- (산업별) 디지털 전환이 국내 경제 및 산업 발전에 필요하다고 생각하는 수준은 각 산업별로 큰 차이는 없었으나 ICT 산업(5.89점)이 타 산업보다 높았음



[그림 8] 국내 경제/산업 발전에 디지털 전환이 필요하다고 생각하는 정도 (종사산업별)

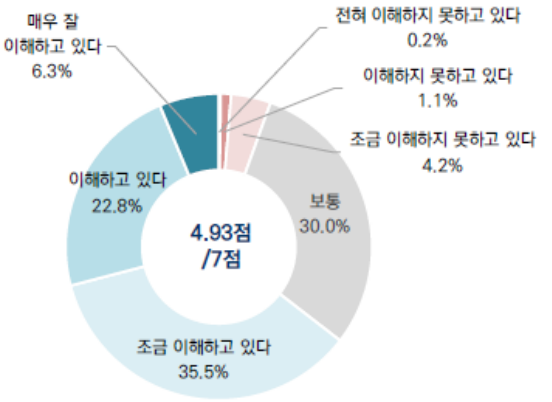
- (연령별) 디지털 전환이 국내 경제 및 산업 발전에 필요하다고 생각하는 수준은 50대(5.81점)가 타 연령대 대비 가장 높게 응답하였음



[그림 9] 국내 경제/산업 발전에 디지털 전환이 필요하다고 생각하는 정도 (연령대별)

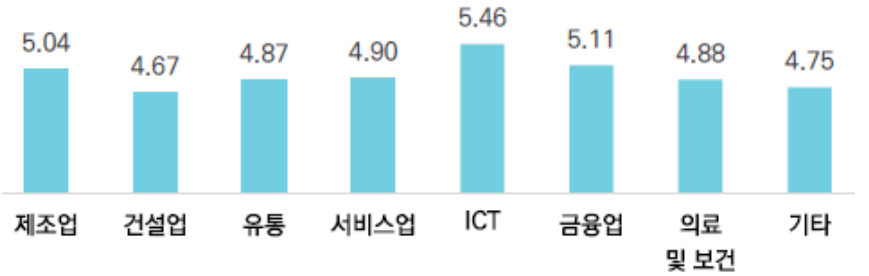
□ 디지털 전환의 기업에 대한 긍정적/부정적 영향 인식 수준

- (전체) 디지털 전환의 기업에 대한 긍정적/부정적 영향 인식 수준은 7점(부정적~긍정적) 기준 평균 4.93점으로 다소 긍정적으로 나타났고, 긍정적이다(Top 3) 응답 비중은 64.5%였음



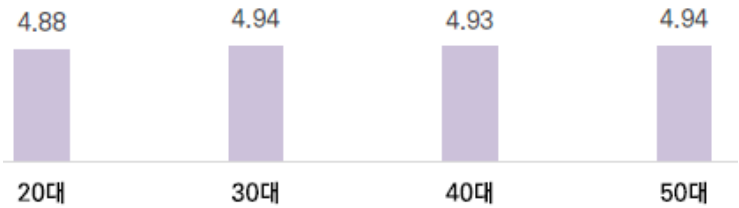
[그림 10] 디지털 전환의 기업에 대한 긍정적/부정적 영향 인식 수준 (전체)

- (산업별) 디지털 전환에 대한 인식은 ICT 산업 종사자들이 타 산업 종사자들에 비해 상대적으로 더 긍정적으로 인식하고 있었고(5.46점), 건설업 종사자들이 타 산업 대비 가장 부정적으로 인식하고 있었음(4.67점)



[그림 11] 디지털 전환의 기업에 대한 긍정적/부정적 영향 인식 수준 (종사산업별)

● (연령별) 디지털 전환에 대한 인식은 대부분이 보통 이상의 긍정 수준으로 인식하고 있었으며 30대와 50대에서 각 4.94점으로 가장 높았고, 40대(4.93점), 20대(4.88점) 순으로 높게 나타났음

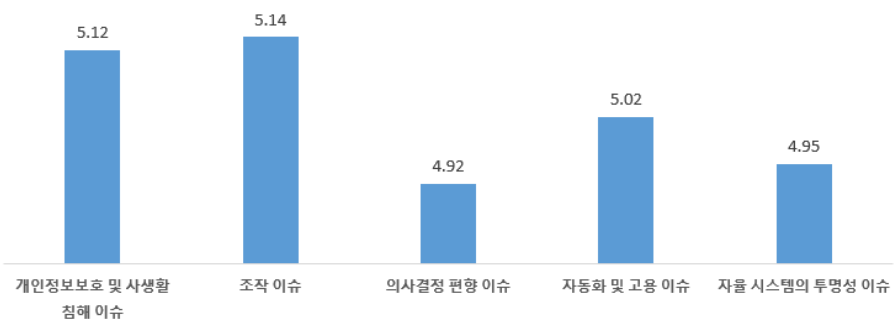


[그림 12] 디지털 전환의 기업에 대한 긍정적/부정적 영향 인식 수준 (연령대별)

□ 5개 위험 유형*별 위험성 인식 수준

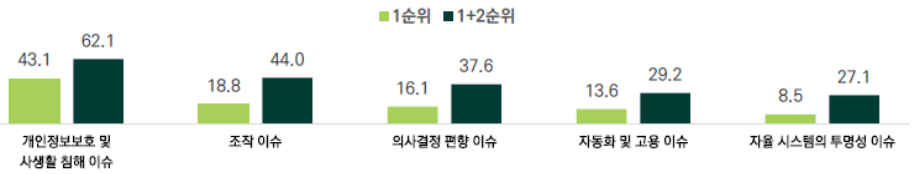
* 개인정보보호 및 사생활 침해 이슈에 대한 위험성, 의사결정 편향 이슈에 대한 위험성, 조작 이슈에 대한 위험성, 자율 시스템의 투명성 이슈에 대한 위험성, 자동화 및 고용 이슈에 대한 위험성

● 전반적으로 모든 위험 유형에 대하여 중요성을 높게 인식하고 있었고(4.9점 이상), 5개 유형 중에서도 개인정보보호 및 사생활 침해이슈(5.12점)와 조작이슈(5.14점)에 대한 위험성을 보다 높게 인식하고 있었음



[그림 13] 5개 위험 유형별 위험성 인식 수준

- 5개 위험 유형에 대한 우선순위(1순위 및 2순위) 응답 결과 개인정보보호 및 사생활 침해 이슈에 대한 위험을 가장 크게 인식하고, 다음으로 조작 이슈, 의사결정 편향 이슈 순으로 가장 큰 위험이라고 인식하고 있었음



[그림 14] 5개 위험 유형에 대한 위험 인식 우선순위 (1순위 및 2순위 선택 결과)

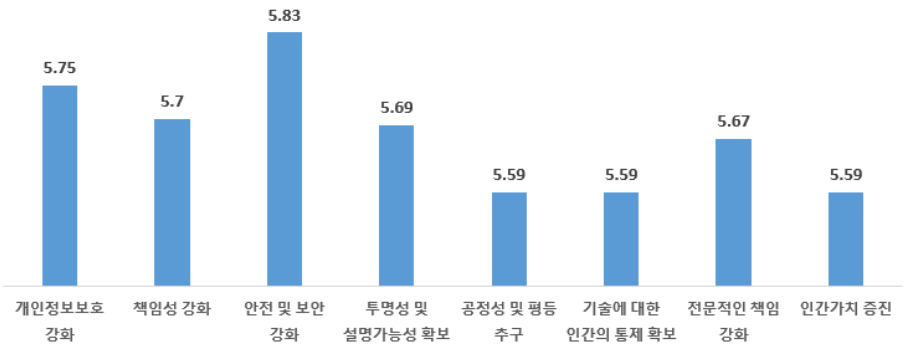
- 아울러 그 외 심각한 디지털 전환 위험 요인에 대한 자유응답에서 역시 '해킹, 개인정보 유출로 인한 피해'에 대한 경험이 가장 많았고, 이 외에는 '일자리 감소 문제', '오작동 및 악의적 사용으로 인한 피해규모 증가' 등의 위험 경험이 높았음

[그림 15] 5개 위험 유형 外 심각한 디지털 전환 위험 요인 (자유응답 결과)

□ 8개 대응 영역*별 정책 수요

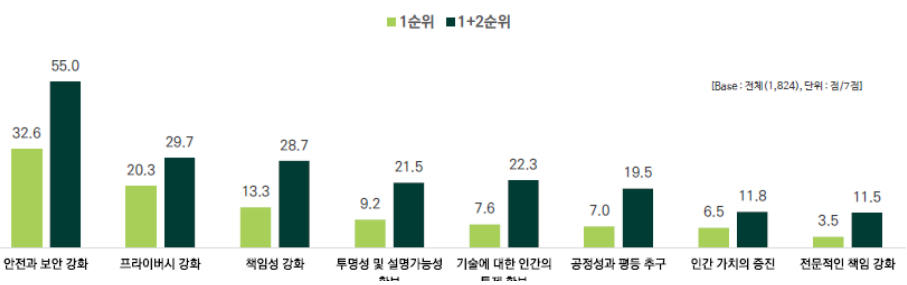
* 개인정보보호 강화, 책임성 강화, 안전 및 보안 강화, 투명성 및 설명가능성 확보, 공정성 및 평등 추구, 기술에 대한 인간의 통제 확보, 전문적인 책임 강화, 인간가치 증진

- 대응 영역 전반에 대한 정책 수요가 모두 높았으며(5.59점 이상) 안전 및 보안 강화(5.83점)와 개인정보보호 강화(5.75점)에 대한 수요가 다른 영역보다 약간 더 높은 경향이 있었음



[그림 16] 8개 대응 영역별 정책 수요

- 8개 대응 영역에 대한 정책적 우선순위(1순위 및 2순위) 응답 결과 안전과 보안 강화와 프라이버시 강화가 가장 높은 비중을 차지하였고, 이는 8개 대응 영역별 정책 수요 결과와도 일치



[그림 17] 8개 대응 영역에 대한 정책적 우선 순위 (1순위 및 2순위 선택 결과)

(4) 소결

□ (의의) 본 인식조사를 통해 디지털 전환에 대한 종합적인 인식수준을 확인할 수 있었음

- 현재까지 조사들은 대부분 디지털 전환에 대한 일반적 인식조사 수준에 머물렀으나 본 연구에서는 이를 연령별/산업별로 세분화하였고,
- 나아가 디지털 전환 위험에 대한 인식 그리고 이에 대한 영역별 정책적 대응 수요까지 확인할 수 있었음
- 이를 통해 종합적이고, 세부적인 디지털 전환에 대한 인식수준을 파악할 수 있었음

□ (시사점 1) 일반적 인식조사 결과 디지털 전환 기술과 밀접한 산업일수록 중요성과 긍정적인 영향을 더 높게 인식하였음

- ICT 산업은 타 산업대비 디지털 전환 기술이 산업의 핵심 기술로 활용되고 있기에 해당 산업 종사자들은 디지털 전환을 익숙한 개념으로 인식하고 있을 가능성이 높음
- 이로 인해 해당 산업이 디지털 전환에 대한 용어 이해도, 국내 경제/산업 발전에 필요한 정도, 기업에 대한 긍정적 영향에 있어 타 산업대비 인식수준이 높았음
- 반면 디지털 전환 기술과 접점이 상대적으로 적은 건설업은 해당 인식조사 항목에서 가장 낮은 수준을 보였음

□ (시사점 2) 위험에 대한 인식조사 결과 모든 유형의 디지털 전환 위험에 대한 인식수준이 높은 경향이 있었음

- 유형별 충분한 설명과 예시로 설문 문항을 구성하였기에 인식수준이 높게 도출됐을 가능성을 배제할 수는 없으나,
- 이점을 감안하더라도 다소 생소할 수 있는 디지털 전환 위험에 대한 인식 수준이 전반적으로 높게 도출된 점은 국민들도 해당 위험들을 중요하게 생각하고 있음을 시사
- 한편 디지털 전환 위험은 기업/연구자/사용자들이 개인차원에서 충분한 대응 체계를 갖추기 어려운 특성이 있음을 고려했을 때 해당 결과는 디지털 전환 위험 대응 정책 마련이 필요함을 암시

□ (시사점 3) 영역별 대응 정책 수요조사 결과 전 영역의 디지털 전환 대응 정책의 필요성을 높게 인식하고 있었음

- 응답자들이 직/간접적으로 경험해보았을 가능성이 높은 안전 및 보안강화에 대한 수요가 가장 높았고, 다른 영역들에 있어서도 전반적으로 높은 수요를 보였음
- 해당 결과는 디지털 전환의 미래사회 위험이슈 대응에 있어 선택적인 정책적 대응보다는 종합적인 정책적 대응이 필요함을 암시

5 디지털 전환 관련 미래사회 위험이슈 대응 정책 동향

□ 본 장에서는 주요국 및 국내 디지털 전환 관련 미래사회 위험이슈 대응 정책 동향을 2개 유형*으로 구분하여 분석

* (1) 소극적 정책 대응, (2) 적극적 정책 대응

- 디지털 전환 관련 미래사회 위험이슈 대응 정책은 주로 인공지능 기술을 중심으로 수립/추진되고 있는 경향
- 소극적 정책 대응은 낮은 규제 수준의 정책적 대응을 의미하며 주로 국가/정부 차원에서 법적 구속력이 약한 개발 가이드라인, 윤리 원칙 배포 등의 형식으로 정책이 수립/추진 中
- 적극적 정책 대응은 강한 규제 수준의 정책적 대응을 의미하며 관련 법안 상정을 통해 기술/서비스/제품 개발의 기획 단계에서부터 특정 기준을 준수할 것을 요구하는 형식으로 정책이 수립/추진 中

(1) 소극적 정책 대응

□ (미국) 미국은 국방부, 백악관, FTC를 중심으로 AI 사용 원칙, 가이드라인을 발표하며 낮은 규제 수준의 정책을 추진 중[6)]

- '16년 미국연방거래위원회(Federal Trade Commission, FTC)는 빅데이터 또는 머신러닝을 사용하는 기업에게 공정성을 확보할 수 있도록 가이드라인을 발간[7)]
- '18년 Defense Innovation Board(DIB)는 'AI Principles Project'를 통해 5개 AI 시스템 사용 원칙*을 제안

  * 책임감, 공평함, 추적 가능, 신뢰성, 관리 가능

- '19년 2월 도널드 트럼프 대통령은 "미국 AI 이니셔티브"를 시작하는 행정명령 발표[8)]
  - (주요 내용) ①기술 혁신 추진, ②기술 표준 개발 추진, ③AI 기술 개발 및 적용 기술을 갖춘 작업자 교육, ④시민의 자유와 프라이버시를 포함한 미국의 가치를

6) `19년 '알고리즘책임법'이 발의된 경우가 있으나 이는 EU 인공지능 법안과 달리 알고리즘 설계/제공 관계자의 강력한 처벌을 통한 통제보다는 문제 발생 시 시정/설명을 요구하도록 구성
7) "Big Data: A Tool for Inclusion or Exclusion?" Fedral Trade Commission (2016)
8) https://futureoflife.org/ai-policy-united-states/

보호하고 AI 기술에 대한 대중의 신뢰 증진, ⑤ AI에서 미국의 기술 우위를 보호하는 동시에 혁신을 지원하는 국제 환경 촉진

- ’20년 미국연방거래위원회(Federal Trade Commission, FTC)는 법 집행 조치, 연구 및 지침에 AI 도구의 사용이 투명성, 설명 가능성, 공정성, 책임감을 강조[9)]
- ’20년 4월 미국 국방부(Department of Defense)는 AI의 군사적 활용에 대한 위험을 방지하기 위해 국방혁신위원회(Defense Innovation Board)의 제안을 통해 5가지 윤리 원칙*을 제시[10)]

  * 책임성(responsible), 공평성(equitable), 추적 가능성(traceable), 신뢰성(reliable), 통제 가능성(governable)

□ (유럽-개별국가) 디지털 전환의 미래사회 위험이슈 대응을 위해 관련 조직을 설립하고, 윤리 원칙 등을 발표하는 형식으로 정책을 추진 中

- (영국) 의회는 인공지능 발전의 경제적, 윤리적, 사회적 영향을 더 고려하고 견고히 하기 위해 AI 선택위원회(Select Committee on AI)를 설립(’17.6)[11)]
  - ’18년 4월 위원회는 “AI in the UK: Ready, willing and able?” 보고서를 발표하며 AI 윤리에 대한 경쟁력 제고
  - 보고서의 5가지 원칙 키워드는 ①인류의 공동 이익, ②명료성과 공정성, ③데이터 권리와 프라이버시, ④교육받을 권리, ⑤악(惡)한 목적을 인공 지능에 부여 금지
- (독일) ’18년 연방법무부 및 소비자 보호부와 연방 내무부는 개인 보호, 사회적 결속 유지, 정보화 시대의 번영 보호 및 증진을 목표로 하는 행동에 대한 구체적인 권장 사항뿐만 아니라 윤리적 기준과 지침을 개발하는 ‘데이터 윤리위원회’를 발족
- (프랑스) ’18년 대통령이 ‘인류를 위한 AI’라는 제목으로 AI 국가 전략 발표(2018.3.29.)[12)]하며 투명성과 공정성을 강조
- (이탈리아) ’17년 AGID(Agency for Digital Italy)에서 AI Task Force를 출범(’17.9)시켰고, ’18년 “AI at the service of citizens” 백서를 발표(’18.3)[13)]
  - (백서 윤리부문 주요 키워드) 데이터 품질과 중립성, 책임성, 투명성 및 개방성, 개인정보보호[14)]

9) Federal Trade Commission (2020), “Using Artificial intelligence and Algorithms”
10) U.S. Department of Defense, “DOD Adopts 5 Principles of Artificial Intelligence Ethics”
11) https://futureoflife.org/ai-policy-united-kingdom/
12) https://futureoflife.org/ai-policy-france/, https://www.aiforhumanity.fr/en/
13) https://futureoflife.org/ai-policy-italy/
14) Italy, AI Task Force(2018), LIBRO BIANCO SULL’INTELLIGENZA ARTIFICIALE AL SERVIZIO DEL CITTADINO

- (스웨덴) AI 개발 및 사용에 필요한 정부의 평가를 요약한 “Artificial intelligence in Swedish business and society: Analysis of development and potential)” 발표[15)]
  - (주요 내용) 윤리적이고 지속 가능한 안전한 AI를 위한 국내외 규정과 규범을 인정하면서 표준과 원칙을 개발하려는 정부의 목표 강조

□ (중국) 다양한 정부기관과 위원회에서 윤리 규범/가이드라인/원칙 제시

- ’17년 국무원은 「차세대 AI 발전계획」을 발표하며 AI 연구개발, 산업화, 인재 개발, 교육 및 기술 습득, 표준 및 규정 설정, 윤리 규범 및 보안을 위한 이니셔티브와 3단계 목표를 제시
- ’19년 차세대 AI 거버넌스 전문가위원회(New Generation AI Governance Expert Committee)*는 8개 신조**를 포함하는 차세대 인공지능 거버넌스 원칙을 발표(`19.6)[16)]

  * ’19년 3월 과기부(MOST)에서 설립

  ** 조화와 친근함, 공정성과 정의, 포용성과 공유, 개인정보보호 및 보안, 제어 가능성, 책임 공유, 열린 협력, 민첩한 거버넌스

- ’19년 국가 차세대 AI 추진실은 AI 거버넌스 전문가위원회의 AI 개발 촉진 관련 법률/규정 및 윤리 규범 공식화를 선언
- 아울러 글로벌 AI 주도권을 확보하고자 “베이징 AI 원칙(Beijing AI Principles) (’19)” Beijing Academy of Artificial Intelligence (BAAI) (2019)을 발표[17)]
  - (AI R&D 원칙) ①선(善)한 목적에 활용, ②인류를 위해 활용, ③책임감, ④AI 시스템의 성숙도·견고성·신뢰성 및 제어 가능성, ⑤윤리적, ⑥다양성·포용성 반영, ⑦AI 개방형 플랫폼 구축
  - (AI 활용 원칙) ①적절한 목적에 활용하여 AI 오용·남용 방지, ②정보에 입각한 동의, ③AI 심리적·정서적·기술적 교육과 훈련
  - (AI 거버넌스 원칙) ①인간-AI를 고려한 고용 최적화, ②AI 거버넌스를 통한 공생 최적화, ③AI 발전을 고려한 원칙·정책·규정 조정, ④AI 수명주기를 고려한 세분화 및 구현, ⑤AI 잠재적 위험을 고려한 전략적 설계

□ (일본) 타 국가대비 가장 빠르게 관련 회의기구를 창설하여 위험이슈를 검토하고, 관련 가이드라인을 작성/배포 중

15) https://futureoflife.org/ai-policy-sweden/
16) https://futureoflife.org/ai-policy-china/
17) https://www.baai.ac.cn/news/beijing-ai-principles-en.html

- ’16년 총무성은 AI가 초래할 수 있는 문제에 대해 다중 이해관계자(산, 학, 관, 소비자)가 함께 논의하는 회의기구인 ‘AI 네트워크 사회를 향한 컨퍼런스’를 창설
- ’17년 AI Network Society는 윤리 및 투명성 등을 강조한 ‘「국제 토론을 위한 AI R&D 가이드라인 초안」’을 발표(’17.7)[18)]
  - (기본철학) 인간 중심 사회, 소프트 법률 가이드라인, 이점과 위험성의 균형, 지속적 검토 및 갱신
  - (9가지 원칙) 협업, 투명성, 제어성, 안전성, 보안성, 프라이버시, 윤리(인간의 존엄성과 개인의 자율성), 사용자 지원, 책임성
- ’19년 내각부는 『인간 중심의 AI 사회 원칙안(’19)』 발표를 통해 국가 및 지방정부를 포함한 일본 사회 전반에 걸친 ‘AI-Ready Society’ 원칙 제시[19)]
  - (AI 사회 7대 원칙(안)) 인간중심 원칙(The Human-Centric Principle), 교육/문해 원칙(The Principle of Education/Literacy), 개인정보보호 원칙(The Principle of Privacy Protection), 보안 보장의 원칙(The Principle of Ensuring Security), 공정 경쟁 원칙(The Principle of Fair Competition), 공정성·책임성·투명성의 원칙(The Principle of Fairness, Accountability, and Transparency), 혁신의 원칙(The Principle of Innovation)[20)]

□ (한국) 한국은 「인공지능 국가전략 (2019)」, 「디지털 뉴딜 (2020)」 정책을 중심으로 디지털 전환 및 관련 기술/산업 발전을 지원하고 있으며 디지털 전환 관련 역기능 방지를 위한 정책을 마련 중

- ’18년 과기부 및 한국정보화진흥원은 「지능정보 사회 윤리 가이드라인 (2018)」을 제정하여 사람 중심의 지능 정보 사회를 실현, 지능형 정보 기술 및 서비스의 개발자 및 공급자의 윤리적 책임 강화, 사용자의 오용 방지를 위한 AI 기술의 역기능 방지책을 마련
- `20년 말 과기부는 「인공지능 국가전략(2019)」의 추진과제로 제시된 역기능 방지 및 AI 윤리와 관련한 「인공지능 윤리기준」 마련
- 또한 과기부는 `21년 5월 '신뢰할 수 있는 AI실현전략'을 발표
- ’21년 5월 개인정보보호위원회는 AI 개인정보보호 자율점검표를 확정 및 공개
  - (주요 사항) 적법성, 안전성, 투명성, 참여성, 책임성, 공정성 등 업무처리 전 과정에서 지켜야 할 6가지 원칙과 점검해야 할 16개 항목, 54개 확인사항을 제시

18) https://futureoflife.org/ai-policy-japan/
19) Japan (2019), “Social Principles of Human-Centric AI”
20) Japan (2019), “Social Principles of Human-Centric AI”, p.7-11

(2) 적극적 정책 대응

□ (EU: GDPR) EU는 EU 회원국에 일괄적으로 적용되는 개인정보 보호법(General Data Protection Regulation, GDPR)을 제정('16년) 및 시행('18년)[21]

- (제정 목적) ①디지털 단일시장에 적합한 통일되고 단순화된 프레임워크 마련*, ②권리와 의무 강화**, ③현대화된 개인정보보호 거버넌스 체계 구축***

* 단일 개인정보보호법 적용, 원스톱샵 메커니즘

** 정보주체 권리 확대(동의 요건 강화, 이동권/잊혀질 권리 등 도입), 기업의 책임성 강화(DPO 지정, 개인정보 유출 통지 신고제 등 도입)

*** 개인정보 감독기구 간 협력 강화, 법 적용의 일관성 보장을 위한 European Data Protection Board 설립(`18), 신뢰할 수 있고 비례적인 제재 부과

<표 16> GDPR 前後 주요 변화

	Before (Directive 95/46/EC)	After (GDPR)
기업의 책임	개인정보 최소 처리 처리목적 통지	DPO 지정, 영향평가 등 추가
정보주체 권리	열람 청구권 등	정보이동권 등 새로운 권리 추가
과징금 부과	회원국별 자체 법규에 따라 부과	모든 회원국이 통일된 기준으로 부과

- (적용 대상) EU 내에 사업장을 운영하며, 개인정보를 처리하는 기업, EU 거주자에게 재화나 서비스를 제공하는 기업, EU 거주자의 EU 內 행동을 모니터링하는 기업

※ 명백하게 EU 시장을 염두하고 있을 때 적용되며, 단순 접근 가능성은 GDPR 적용 근거가 되지 않음

- (정보주체의 권리: 신설/강화된 조항) 삭제권, 처리 제한권, 개인정보 이동권, 프로파일링*을 포함한 자동화된 의사결정

* 개인의 사적인 측면(직장내 업무수행, 경제 상황, 건강, 개인적 취향 등)을 분석, 예측하기 위한 자동화된 처리

21) https://www.privacy.go.kr/pic/nation_eu.do

<표 17> GDPR 정보주체의 권리

권리	내용
정보를 제공받을 권리	• 컨트롤러는 공정하고 투명한 처리 원칙을 보장하기 위해 정보주체에게 본인의 개인정보를 어떻게 처리하고 있는지에 관한 정보를 명확하고 쉬운 언어로, 무상으로 알려주어야 함
정보주체의 열람권	• 컨트롤러는 정보주체가 개인정보 처리 내용과 그 적법성을 확인할 수 있도록 정보주체의 요구가 있을 경우 자신의 개인정보 및 다음의 모든 정보에 대해 열람할 수 있도록 조치하여야 함
정정권	• 정보주체는 개인정보가 부정확하거나 불완전하다면 이에 대한 정정을 요구할 권리가 있고, 정보주체의 정정권리를 보장할 수 있도록 필요한 조치를 하여야 함
삭제권(잊힐권리)	• 정보주체는 본인에 관한 개인정보의 삭제를 컨트롤러에게 요구할 권리를 가지며 컨트롤러는 개인정보 처리목적의 달성, 정보주체의 동의 철회 등의 경우에 개인정보를 삭제하여야 함
처리 제한권	• 정보주체는 자신에 관한 개인정보의 처리를 차단하거나 제한할 권리를 가지며, 개인정보 처리가 제한되면 컨트롤러는 그 정보를 보관하는 것만 가능하며 처리 제한을 해제하는 경우 그 사실을 정보주체에게 고지
개인정보 이동권	• 정보주체나 다른 컨트롤러에게 자신의 데이터를 제공할 것을 요구할 수 있는 권리
반대권	• 정보주체는 프로파일링 등 본인과 관련한 개인정보의 처리에 대해 언제든지 반대할 권리를 지님 • 컨트롤러는 정보주체에게 최초 고지하는 시점에 반대권에 대한 내용을 알려주어야 하며, 이러한 사항은 다른 정보와 분리하여 분명하게 제시해야 함
프로파일링을 포함한 자동화된 의사결정	• 법적 효력을 초래하거나 이와 유사한 중대한 효과를 미치는 사항에 대하여 프로파일링을 포함한 자동화된 처리에만 근거한 인적개입 없는 결정의 적용을 받지 않을 권리

- (위반 시 과징금 부과 규정) 최대 과징금은 일반적 위반 사항인 경우 전 세계 매출액의 2% 혹은 1천만 유로(약 125억원) 중 높은 금액이며, 중요한 위반 사항인 경우 전 세계 매출액의 4% 혹은 2천만 유로(약 250억원) 중 높은 금액

□ (EU: AI Act) 인공지능 법안은 `21년 4월 EU 집행위원회가 발표한 법안으로 주요 골자는 인공지능이 사람에 미치는 위험 정도에 따라 위험을 3가지 유형으로 나누어 규제하는 것임

● (규제 목표) 인공지능 시스템의 안전과 기본권, 기존 법률 준수, 법적 불확실성 제거, 기본권 및 안전요건의 거버넌스 향상, 합법적이고 안전하며 신뢰할 수 있는 인공지능 어플리케이션 단일시장 발전 촉진 등

● (위험 기반 규제 접근) 인간의 기본권과 안전을 중심으로 인공지능 시스템의 영향을 고려하여 위험을 산정하고, 위험관리를 수행하도록 제안

- (위험 유형) 수용불가 위험, 고위험, 저위험/최소위험

<표 18> EU AI Act의 인공지능 위험 유형과 유형별 규제 사항[22)]

인공지능 위험 유형	위험 유형별 정의	위험 유형별 규제
수용불가 위험	• AI 시스템이 사람 행동을 조작하는 경우 • 사회적 약자를 이용/공격하는 경우	• 시장 출시 불가 • (위반시) 최대 3000만 유로, 세계 연간매출액의 6% 中 높은 금액의 벌금을 부과
고위험 AI	• AI 시스템이 운송/교통, 기계, 무선장비 및 의료기기 등 사람의 안전과 관계된 경우 • 가스, 전기 등 중요 인프라에 이용되는 경우 • 잠재적 가해자 또는 피해자가 될 위험 평가, 거짓말 탐지 등 감정 상태 확인, 증거 신뢰성 평가, 범죄 분석에 사용되는 경우 등	• 허용하되 위험관리시스템을 기획/설계/구축할 의무 부과 • 데이터 거버넌스 구축, 기술문서 마련, 기록 투명성 확보, 인간에 의한 감독, 정확도, 보안의 확보 등이 요구 • 감독기관의 모니터링 대상이 되고, 상용화 이후에도 문제 상황의 발생에 따른 보고 의무 등의 의무 부과
저위험, 최소위험	• 게임 등 권리침해, 안전위험이 최소화된 경우	• 규제하지 않음

(3) 소결

□ (의의) 주요국의 디지털 전환 관련 미래사회 위험이슈 대응 정책을 규제 수준에 따라 구분하여 조사

● EU는 가장 선제적으로 강력한 수준의 정책(규제)를 통한 디지털 전환 미래사회 위험이슈 대응을 추진 中

● 반면 한국을 포함한 다른 국가들은 가이드라인/윤리 원칙 제정 수준에서 낮은 수준의 대응 정책을 추진하고 있음

● 현시점에서는 어떠한 정책 수단이 옳은지 판단하기는 어려우나 이를 구분하여 살펴봄으로써 각 규제 수준별 장/단점* 유추 가능

* (낮은 수준의 규제) 디지털 전환 기술의 자유로운 발전 및 관련 산업 성장 촉진/추후 예측하지 못한 문제 발생 시 국가차원의 신속한 대응이 어려움, (높은 수준의 규제) 정부 차원에서 관련 위험을 미리 검토하여 대응책을 마련해 놓음으로써 안전한 산업 생태계 조성 가능/규제로 인한 기술 및 산업 발전 저해

22) EU 인공지능법(안) 시사점과 대응전략, 전자신문('21.6.29)

□ (시사점 1) 주요국은 디지털 전환 기술 中 인공지능 기술의 양면성에 주목하여 이것의 역기능 대응을 위한 정책 마련을 신속하게 추진하였음

- 美/日/中/EU 주요국가들은 `16~`17년도부터 인공지능 기술에 초점을 맞추어 디지털 전환 역기능 대응을 위한 정책적 기틀을 마련한 반면,
- 한국은 상대적으로 약간 늦은 시점인 `19년 말부터는 주요국과 같이 인공지능 기술에 포커스하여 디지털 전환 역기능 대응 정책을 수립/추진 중

□ (시사점 2) 효과적인 디지털 전환 위험 이슈 대응 정책 마련을 위해서는 규제 수준에 따라 다양한 정책적 수단을 검토할 필요

- 현재 미국/중국은 낮은 수준의 자율규제 형식을 통한 역기능 대응으로, EU는 강한 수준의 법/규제를 통한 역기능 대응으로 방향을 설정하여 관련 기술/산업 생태계를 조성 中
- 그러나 한국은 아직 다양한 정책 수단에 대한 깊이 있는 고민이 부족하고, 이에 대한 사회적 합의 및 정책의 거시적 방향 설정이 부족한 상황
- 이에 다양한 정책 수단을 검토하여 국가/사회적 맥락에 맞추어 정책 방향을 설정할 필요

6 디지털 전환 관련 미래사회 위험이슈 대응 정책 방향 설정

(1) 개요

□ 디지털 전환 관련 미래사회 위험이슈 대응을 위해서는 일방적 규제보다는 다양한 정책 수단과 국내 사회적 맥락을 검토할 필요

※ 5장에서 살펴본 바와 같이 디지털 전환 관련 미래사회 위험이슈 대응은 그 핵심기술인 인공지능 기술을 중심으로 추진되고 있기에 본 장에서도 인공지능 기술을 중심으로 대응 정책 방향을 논의

- 자율규제 형식의 정책은 산업 및 기업의 성장을 촉진할 수 있다는 장점이 있으나 예상치 못한 큰 경제/사회적 문제에 선제적으로 대응이 어려워 문제발생 시 국가와 사회에 상당히 큰 피해를 입힐 수 있다는 단점이 존재
- 반대로 강한 법적 규제 형식의 정책은 파생될 수 있는 미래사회 위험을 최소화 할 수 있는 장점이 있는 반면, 산업/기업의 성장을 저해할 수 있고, 주요국 대비 기술 경쟁력을 저하시킬 수 있다는 단점이 존재
- 이에 본 장에서는 미래사회 위험이슈 대응 정책 수립을 위하여 ①인공지능의 법규범적 속성을 검토하고, ②규제 강도에 따른 정책 시나리오를 분석하였음

(2) 인공지능(알고리즘)의 법규범적 속성 검토

□ 인공지능 기술은 알고리즘의 법규범적 불확정성으로 인하여 법규범을 통한 예방적 견지에서 규제하기에는 한계가 존재

- 인공 신경망의 판단 구조는 궁극적으로 인공지능의 판단 및 행위의 규범적 차원의 예견 가능성이 현저하게 떨어질 가능성이 높음
- 다소 고정적인 기술 환경이 전제된 상황에서 해당 기술이 발생시킬 위험 및 법익 침해의 문제는 그러한 기술에 관한 확정적인 기술적·관리적 보호조치 및 각종 인허가 요건 등을 법령 등에 명시함으로써 대부분의 사안에 법규범적 대응을 수행할 수 있음
- 그러나 현재의 인공지능 기술은 태생적 기술 구조(불확정성)이 전제되어 있기에 법규범을 통한 규제는 한계가 있음을 감안할 필요

□ 이와 같은 인공지능 기술의 법규범적 불확정성을 감안했을 때 법적 대응은 알고리즘의 투명성 확보 여부를 중심으로 이루어질 것

- 인공지능 알고리즘이 불확정적인 속성을 가진다고 했을 때, 가장 유효한 대응 방식은 시시각각 유동적으로 변화하는 알고리즘 자체가 투명성을 확보하여 제기 가능한 위험성을 최소화할 수 있는 대응 방안들을 지속적으로 강구해 나가는 것임
- 현실 법규범적 논의에 있어 알고리즘 투명성의 요청은 알고리즘의 구성 과정과 그것을 통한 기계적 판단 결과가 가지는 영향을 사전에 예측할 수 있도록 하는 방안에 관한 논의임
- 물론 자율적인 기계 학습을 통해 구성되는 알고리즘을 사전에 모두 정확하게 예측하는 데에는 한계가 있을 수밖에 없으므로 이러한 투명성 확보가 용이한 것은 아님
- 결과적으로 이러한 알고리즘 투명성 요청은 다분히 규범적 이상향 또는 지향점을 의미한다고 보는 것이 타당함

□ 알고리즘 투명성 요청 관련 쟁점이 되고 있는 첫 번째 사항은 차별과 편향임

- 미국은 '16년 1월 미국 연방거래위원회(Federal Trade Commission: FTC)가 발간한 "빅데이터: 포용의 도구인가 배척의 도구인가?"[23]와 같은 해 5월 백악관이 내놓은 "빅데이터: 알고리즘 시스템, 기회, 그리고 시민권"[24]인공지능 기술의 근간이 되는 빅데이터 기술이 개인정보보호에 대한 우려를 넘어 현재의 경제적 차별을 영속화하거나 새로운 차별을 만들어 낼 수 있다는 새로운 문제를 제기
  - 첫째, 특정 계층으로부터만 추출된 불완전한 데이터가 빅데이터 알고리즘에 입력되면 여기에 데이터가 반영되지 못한 이들은 빅데이터 기술의 혜택으로부터 소외될 가능성이 크고, 이런 데이터 피드백이 반복되어 누적되면 사회적 차별이 고착화될 가능성이 존재

(예) 국내 대형 포털 사이트들의 알고리즘에 기반한 검색 순위 및 기사 배열의 문제
- 알고리즘 방식을 채용한다는 것은 인간의 개입을 가급적 제약하여, 객관적인 데이터에 입각하여 검색 순위 및 기사 배열의 순서를 정하겠다는 의도를 가지고 있으나 인터넷 상에 존재하는 데이터와 이용자들의 검색 빈도는 그 자체가 가치 편향을 가지는 정보들일 가능성이 높음
- 검색 순위와 기사 배열은 그 자체로 하나의 정보이기도 하지만, 특정 쟁점에 관한 여론 및 사회적 가치판단을 좌우할 수 있는 중요한 역할을 한다고 볼 수 있음
- 따라서 데이터에만 의존하는 인공지능 학습을 통해 알고리즘을 구축하게 되는 경우,

23) Federal Trade Commission, Big Data: A Tool for Inclusion or Exclusion?, January, 2016
24) Executive Office of the President, Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights, May, 2016

결국에는 특정 편향을 고착화시킬 수 있는 상황이 매우 빈번하게 발생할 수밖에 없음

- 둘째, 알고리즘 설계 과정이 투명하지 않거나 자동화된 알고리즘이 '블랙박스'로 남게 되면 그 안에 내재된 차별적 요소를 발견해내거나 그에 저항하는 일이 어려워짐

(예) 인공지능 기술기반 개인 대출 심사
- 금융권의 대출 가능 여부는 개인의 경제활동 이력에 대한 데이터를 기반으로 한 신용평가 점수(credit score)에 의해 결정
- 미국인의 약 11%가 경제적 어려움으로 신용 이력이 충분치 않은 신용 사각 (credit invisible) 지대에 있으며, 이들은 인종이나 거주지역 등을 근거로 금융거래를 거부당하는 '디지털 레드라이닝(digital redlining)'이라는 신종 디지털 차별의 피해자들이라고 할 수 있음
- 결과적으로 이들에 대한 편견은 현재의 빅데이터 기술을 통해 더욱 고착화될 가능성이 농후

(예) 인공지능 기술기반 법원의 판결[25)]
- 실제 가중처벌이 이루어지는 누범(累犯) 등을 판단하는 데 있어 특정 통계적 알고리즘을 활용하는 사례
- 이 경우 소송 당사자들은 그러한 판단 알고리즘이 내린 결정의 근거를 알고 싶어 할 가능성이 높음
- 그러나 법원 또는 판사의 입장에서는 많은 경우 그러한 알고리즘의 판단이 개연성이 높다는 이유를 명확하게 이해하거나 설명하기 힘든 경우가 많음

□ 알고리즘 투명성 요청 관련 쟁점이 되고 있는 두 번째 사항은 프라이버시 침해 가능성임

- 인공지능 알고리즘의 활용은 궁극적으로 특정 개인에 최적화된 개별화 서비스를 지향하는 것이며 이를 위해서는 특정인을 대상으로 한 다양한 제반 데이터들을 수집·학습함은 물론이고, 이에 기반하여 서비스 알고리즘을 신속하게 구축해야 함
- 환언하면 인공지능 알고리즘의 유용성은 얼마나 많은 소비자 또는 이용자 데이터를 확보하고 있는지 여부와 직결되기 때문에 바로 이 지점에서 문제시되는 것이 개인정보 또는 프라이버시의 문제임

25) 최근 미국 법원에서 콤파스(Compas)라는 소프트웨어가 문제시 된 바 있었고, 본 판결의 문제점에 관해 명확히하고 있는 사이트로는 하기 링크 참조
https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

- 원론적인 차원에서 개인정보의 활용은 정보주체의 동의를 받을 수 있다면 크게 문제되지 않을 가능성이 높으나, 인공지능의 학습 및 분석과정을 통하여 자연스럽게 당초에는 식별 가능성이나 프라이버시 침해 위험이 없던 정보가 침해 가능성을 가지는 정보로 언제든지 변화할 수 있는 상황이 되었음
- 이와 같은 개인정보 및 프라이버시 정보의 활용에 있어 인공지능 기술이 더욱 문제시되는 것은 과거와는 달리 개인정보의 처리에 인간의 개입이 최소화된다는 점임(즉, 인공지능 알고리즘이 특정 개인에 관하여 내린 판단이 어떠한 기준과 근거를 가지고 이루어진 것인지를 알지 못하는 경우가 빈번하게 발생할 가능성이 높음)
- 이러한 측면에서 EU와 일본은 개인정보 보호법제 개선에 관한 논의들을 상당히 진척시켜 나가고 있음[26)]

EU의 GDPR
- 이 규칙은 기본적으로 빈번해진 데이터 활용 급증이 불러올 수 있는 프라이버시 침해 위협을 직시하고 있으며 그 가운데에서도 데이터의 안전한 활용 방안을 유도하기 위한 규정들을 담고 있음
- 대표적으로 원칙적 사전 동의 원칙, 가명처리(pseudonymisation) 규정, 프라이버시 중심 설계(privacy by design) 규정, 프로파일링 규제 규정 등을 명확히 했을 뿐만 아니라 EU 역내 국가들의 프라이버시 규범을 강화하기 위한 성격을 가짐

일본의 개인정보보호법 개정
- 일본의 경우에는 민간영역에 적용되는 「개인정보보호법」을 새롭게 개정하여 '익명가공정보' 개념을 도입했다는 특징이 있음
- 이는 개인정보의 활용과 보호를 조화시키기 위한 방안인 것으로 평가할 수 있음
- 실제로 일본의 익명가공정보 규정은 상당부분 개인정보 규제를 완화하고 있는 것으로도 보이는 것이 사실이지만, 익명가공정보의 활용과 관련한 다양한 규제들을 설정하고 있음
- 이와 더불어, 일본은 개인정보 보호의 문제를 더욱 엄격하게 접근하기 위하여 개인정보보호위원회의 역할과 위상을 재정립하기도 하였음

□ 상기 논의를 종합해보면 인공지능 알고리즘의 투명성을 확보하기 위하여 주안점을 두어야 하는 쟁점 요소는 두 가지(학습 데이터의 건전성과 알고리즘의 안전성/공정성)이며 이를 달성하기 위해서는 인간의 개입이 불가피

26) 물론 EU와 일본은 관련 규정을 전면적으로 개편함으로서 실무적인 제도개선 조치가 완성된 것처럼 오해를 불러일으킬 수 있으나 EU GDPR은 향후 실제 작용과정에서 세부적인 적용기준들이 구체화될 것으로 보이며, 일본의 경우에도 법개정과 관련한 논란이 지속되고 있어 아직 법제 정비를 완수했다고 보기 힘든 측면이 있음

- 모종의 인공지능 서비스를 통해 영업을 영위하고자 하는 사업자들은 인공지능이 비교적 정확하고 공정한 데이터를 학습할 수 있도록 개입할 필요가 있고,
- 학습 데이터 선별 또는 초기 알고리즘 구축 등 무엇을 원인으로 하든, 차별적이거나 위험성 있는 판단 결과가 도출되는 경우에도 그것을 통해 서비스를 제공하고자 하는 이들은 그 결과를 통제하기 위해 개입할 필요가 있음
- 결국 인공지능 알고리즘 투명성을 확보하기 위한 시도는 인공지능 서비스 개발자, 관리자 및 사업자 등을 기본적인 규율대상으로 하는 것이라고 할 수 있음

> (유사사례) EU의회가 EU집행위원회에 대한 권고로서 의결한 Civil Law Rules on Robotics에서 다음과 같이 기술[27]
> - "T. 아시모프의 법칙은 자율성과 자체학습 기능이 내장된 로봇을 포함한 로봇의 설계자와 생산자, 운영자에게 적용되는 것으로 간주되어야 한다. 이러한 법칙은 기계 코드로 변환될 수 없기 때문이다."
> - 이는 결국 알고리즘 투명성을 확보하기 위한 책임 또는 책무 논의는 개발자, 관리자 및 사업자 등에게 귀결될 수밖에 없음을 의미

(3) 규제 강도에 따른 정책 시나리오 분석

구분	규제 강도
시나리오 1: 전문가 윤리적 접근	낮음(低)
시나리오 2: 인증체계의 구축	⇕
시나리오 3: 개인적 권리의 설정	
시나리오 4: 직접적인 행정 규제 설정	높음

[그림 18] 규제 강도에 따른 정책 시나리오

□ 본 연구에서는 규제 강도에 따른 정책 시나리오를 4개로 유형화하였음

- 첫째, 다소 강제적인 법률적 규제보다는 인공지능 개발 및 서비스 사업자들의 윤리적인 자율 통제를 기대하는 방식

27) European Parliament, European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), 2017.2.16.

- 이 방식은 기본적으로는 법적 대응이라고 보기 힘든 측면이 있기 때문에 상정하고 있는 규율 강도가 가장 낮은 단계
- 다만 법률적 차원에서 이러한 윤리적 접근을 유도하는 등의 규정을 둘 수 있다는 측면에서 온전히 비법률적 접근이라고 보기는 힘든 측면

● 둘째, 서비스 및 상용화된 인공지능 서비스들에 관해 활용할 수 있는 공적인 인증체계를 법규범적으로 구축하는 방안

- 국내에서 이러한 인증체계를 활용하는 방식은 개인정보 또는 정보보호 분야에서 빈번히 활용되어 왔는데 이들은 대부분 국가(공공기관) 주도의 인증체계였음
- 이 방법은 직접적인 행정적 의무나 책임을 부과하는 것이 아니라는 측면에서 다소 완화된 법률적 접근방법일 수 있음

● 셋째, 법규범상 직접적으로 특정 권리를 설정하거나 의무를 규정하는 방식

- 권리나 의무 규정을 통한 접근 방식은 전통적인 법적 접근방식이라고 할 수 있으며 권리나 의무의 설정은 다소 강화된 기본권 제한 방식
- 특히 권리 부여의 경우에도 그에 대응하는 자에게는 의무가 부과되는 것도 동일하다는 점에서 타방의 기본권 제한이 발생

● 넷째, 직접적으로 인공지능 알고리즘 활용상의 법적 요건들을 설정하는 방식

- 예를 들면, 현행 정보보호 법제들이 그에 관한 기술적·관리적 보호조치 요건들을 법령상 규정하고 또한 운영하는 상황과 유사한 행정 규제
- 이러한 행정규제를 정립하여 활용하려면 인공지능 알고리즘 기술이 명확화되어 법적 판단이 일정 부분 확정적으로 이루어질 수 있어야 함

□ (시나리오 1: 전문가 윤리적 접근) 전문가 집단의 자발적인 참여와 협력을 전제로 스스로 전문가 통제를 달성할 수 있도록 하는 방안

● 일반적으로 매우 빠른 기술 변화가 존재하거나 사안의 파악을 위해 매우 전문적인 식견이 요구되는 경우에는 외부적 규제는 실효성을 가지기 힘든 측면이 존재

※ 전문가가 아닌 제3자적 지위에서 관련 사안의 정확한 실체적 관계를 파악하기도 힘들거니와 이해하기도 어렵기 때문

● 인공지능 연구의 경우 상당히 빠르게 변화를 거듭하는 영역일 뿐만 아니라 매우 전문적인 기술 영역이라고 볼 수 있어 이러한 윤리적 차원의 접근이 타당한 측면

(예) 「생명윤리 및 안전에 관한 법률」 제7조 등의 국가생명윤리심의위원회 및 제10조 기관생명윤리위원회 등의 규정
- 생명윤리 분야는 인공지능 분야와 마찬가지로 매우 첨단 기술의 영역이고, 관련 위험이 매우 전면적으로 확산될 가능성을 가짐
- 따라서 「생명윤리 및 안전에 관한 법률」은 국가적 차원의 전문가 윤리 통제 업무를 수행하는 국가생명윤리심의위원회를 규정하고 있으며 개별 관련 연구 기관별로 기관생명윤리위원회를 두도록 규정
- 국가 위원회를 두고 기관 위원회를 별도로 두도록 하는 것은 연구기관의 자율성을 보장하여 연구의 수월성을 보장한다는 취지도 가지고 있는 것[28]

□ (시나리오 2: 인증체계 구축) 인허가 또는 행위 규제 요건 등을 법령상 규정하여 직접적으로 규제하는 방식이 아니라 자발적인 인증 획득을 유도하여 관련 분야 위험 관리 등의 대응 체계를 구축할 수 있도록 유도하는 방식

- 인증체계는 국가 및 공공기관 등이 주축이 되어 운영되는 경우도 있고, 법령상 근거 없이 민간 인증사업자들이 운영하는 경우도 있는 방식
- 우리나라의 경우에는 이제까지 민간의 자발적인 참여를 유도한다는 취지를 가지고 있음에도 불구하고, 국가 및 공공기관 주축으로 인증체계 및 제도가 운영되어 왔기에 국가 중심적 행정규제로서의 성격이 강한 측면이 존재[29]
- 일반적으로 주요 국가들의 인증 관련 제도화 동향은 민간 인증사업자들의 인증을 전제로 운영되는 경우가 대다수

(예) 일본[30]
- 일본은 인공지능에 관한 민간 인증사업자의 인증을 제도화 또는 국제 표준화하기 위한 작업을 수행 中
- 이러한 일본의 관점은 인공지능 투명성 확보 작업이 비단 개별 일개 국가 차원에서 이루어질 수 없음을 인식하고 있는 것이라고 볼 수 있음
- 또한 규제에 관한 글로벌 스탠다드를 주도하기 위한 장기적인 포석이라고 볼 수 있음

※ 이러한 인증에 관한 문제에 대해서는 모종의 규제체계에 관해 다소 소극적인 입장을 가진 것으로 알려진 미국의 경우에도 그 필요성을 일부 언급하고 있다는 점에서도 이러한 법규범적 접근이 현실적인 대안 중 하나로 주목받고 있다는 점을 확인할 수 있음[31]

※ 현재 각 국가들은 아직까지 명확한 인공지능 알고리즘 인증체계 구축에 관한 대안을 제시하고 있는 상황은 아니지만 향후 인공지능의 위험 및 역기능이 더욱 현실화되는 경우에는 이와 같은 인증체계를 구축해 나갈 가능성이 크다고 볼 수 있음

28) 이와 유사한 맥락에서 인공지능에 관한 규범적 탐색을 의도한 EU의 Civil Law Rules on Robotics 권고의 경우에도 '연구윤리위원회' 등과 같은 전문가 윤리적 접근 방식에 관하여 언급하고 있다는 점에 주목할 필요

29) 물론 「정보통신망 이용촉진 및 정보보호 등에 관한 법률」상 정보보호 관리체계 인증(ISMS)의 경우에는 특정 규모 이상 사업자들에게는 법률상 인증 의무를 부과하고 있으나 이는 인증체계의 기본 취지에 부합하지 않는다는 평가가 지배적

□ (시나리오 3: 개인적 권리의 설정) 알고리즘을 개발 및 서비스하는 사업자들로 하여금 이용자들에게 알고리즘 구조와 영향을 알려주도록 하고, 이용자들의 입장에서는 그에 대한 설명을 요청할 수 있도록 하는 규제방식

- 이는 전통적으로 개인정보자기결정권을 보장하고 있는 법제들이 취하고 있는 보편적인 방식으로 개인정보를 정보주체로부터 수집, 이용 및 제공하기 위해서는 수집 목적 등 관련 사항에 관한 정보를 정보주체에게 제공해 준 연후에 정보주체의 동의를 받아야 함(well informed consent)

- 종래 좀 더 포괄적인 이용자의 설명 요청권의 유형은 우리나라의 「약관의 규제에 관한 법률」 제3조상의 약관에 관한 설명의무 등을 예시로 확인할 수 있음

- 그런데 이 규정의 경우, 사업자측의 약관에 관한 설명의무가 존재하는 반면 설명을 요청할 수 있는 권리가 이용자측에 부여되어 있는 것인지는 모호한 측면이 있음[32]

(예) 설명 요청권
- 최근 법제화 후 그 구체적인 실현방식에 관해 논의 중인 인공지능 알고리즘 등에 관한 설명요청권(right to explanation)이라는 용어가 있음
- 이는 아직까지 권리개념으로 통용되는 용어라고는 볼 수 없음
- 다만 설명 요청권이 현재 GDPR 상에 규정되어 있다는 점을 지적하는 논의가 제기된 바 있으며[33], 이에 대해서는 상당한 논란이 발생하고 있음[34]
- 현재 제기되고 있는 설명요청권은 GDPR 프로파일링 관련 조항에 있어 특정 알고리즘이 가지는 로직과 그 영향을 정보주체에게 알려주어야 한다는 규정 내용으로 비롯됨(물론 관련 규정 본문에 이에 대해 명확하게 권리성을 부여하고 있는 것은 아니지만, 해석 및 추론을 통해 그 권리성을 확인할 수 있다는 입장으로 이해해 볼 수 있음)
- 이와 관련한 조문은 GDPR상 수차례 반복되며[35] 그 내용을 보면 명확하게 정보관리자에게 설명을 요청할 수 있는 권리라는 법문을 사용하고 있는 것은 아니지만 그러한 권리가 규정되어 있는 것으로 충분히 해석될 여지가 있는 것이 사실[36]

30) 일본은 직접적으로 AI 개발을 규제하기 위한 기준을 제시하기 위한 시도보다는 제3자가 해당 AI의 개발원칙에 대한 적합성을 평가하여 인증하는 제도를 상정하고 있다. ＡＩネットワーク社会推進会議事務局(総務省情報通信政策研究所調査研究部), 「ＡＩ開発ガイドライン」(仮称) の策定に向けた国際的議論の用に供する素案の作成に関する論点, 2016.12.28, 51면.

31) Executive Office of the President, Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights, 2016.5, pp. 22-24.

32) 다만 법문의 반대해석상 그러한 권리가 이용자에게 존재한다는 사실을 추론해 볼 수 있는 상황

33) Bryce Goodman, Seth Flaxman, “European Union regulations on algorithmic decision-making and a "right to explanation"”, 2016

34) 이러한 논의를 보다 상세히 소개하고 있는 논문으로 다음의 논문을 참고 Sandra Wachter, Brent Mittelstadt, Luciano Floridi, “Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation”, International Data Privacy Law 7(2), 2017.

35) 이에 대해서는 특히 GDPR 제13조, 제14조, 제15조에 규정되어 있는데, “프로파일링을 포함한 자동화된 의사결정의 존재, 그리고 적어도 그러한 경우, 이에 사용되는 로직에 관한 의미 있는 정보와 해당 처리가 정보 주체에 대해 갖는 중요성과 예상 결과”를 정보주체에게 알려야 한다는 법문이 동일하게

□ (시나리오 4: 직접적인 행정 규제 설정) 보다 직접적인 행정규제를 설정하는 방식은 현 단계에서 심도 있는 논의가 이루어지고 있지는 못한 상황[37]

- 그 이유는 아직까지 인공지능 기술의 기술 방식 중 보편적 지위를 가지고 있는 것이 존재하지 않을 뿐만 아니라 향후 기술발전의 맥락이 어떠한 방향으로 흐를지도 명확하게 예측하기 힘들기 때문
- 다만 인공지능 알고리즘 기술의 불확정성 또는 블랙박스적 성격에 주목하여 이에 대한 규제방향을 포괄적으로나마 언급하고 있는 문헌들은 EU의 Civil Law Rules in Robotics에서 확인해 볼 수 있음

> 표준화, 안전 및 보안
> - 22. 표준을 정하고 상호운용성을 부여하는 문제가 인공지능과 로봇공학 기술 분야에서 향후 경쟁에 핵심이라는 점을 강조한다. 혁신을 촉진하고, 역내시장의 분열을 방지하며 해당하는 경우 작업 환경의 최소 안전기준을 포함한 고도의 제품 안전과 소비자 보호를 보장하기 위해 특히 유럽표준화기구 및 국제표준화기구와 함께 기술 표준의 국제적인 조화를 위해 계속 노력할 것을 집행위원회에 촉구한다. 혁신의 가치를 극대화하고 로봇의 상호 소통을 보장하기 위한 합법적 역엔지니어링과 개방형 표준(reverse-engineering and open standards)의 중요성을 강조한다. 이 점에서 로봇공학 관련 표준 개발을 전담하는 'ISO/TC 299 로봇공학' 같은 특별 기술위원회를 설치하는 것을 환영한다.

> 로봇공학 기술자를 위한 윤리 행동 강령
> 가역성
> - 제어성의 필수 조건인 가역성은 로봇을 안전하고 안정적으로 동작하도록 프로그래밍할 때의 기본 개념이다. 가역성 모델은 로봇에게 어떤 동작을 어떻게 되돌릴 수 있는지 알려준다. 마지막 작업이나 일련의 작업을 실행 취소할 수 있는 능력을 통해 사용자는 원하지 않는 작업을 취소하고 '정상' 작업 상태 단계로 되돌아갈 수 있다.

- 상기 EU의 논의 상황을 분석해 보면, 아직까지 직접적인 법적 규제를 설정하기 위한 작업을 진행하고 있는 것은 아니지만 충분히 알고리즘 투명성 확보를 위한 규제가 가능하다는 점을 가역성 또는 역엔지니어링이라는 용어를 통해 보여주고 있음[38]

반복

36) EU GDPR, recital (71) 정보 주체는, 온라인 신용 거래 신청 자동 거절이나 인간의 개입이 없는 전자 채용 절차 등과 같이, 자동화된 처리만을 바탕으로 정보 주체의 개인적 요소를 평가하는 조치를 포함할 수 있으며 자신과 관련한 법적 영향 또는 이와 유사하게 중대한 영향을 미치는 결정의 대상이 되지 않을 권리를 가져야 한다. (중간생략) 어떤 경우든 그러한 처리에는 적합한 보호장치가 적용되어야 하며, 여기에는 정보 주체에 대한 구체적 정보 제공, 인간의 개입을 받을 권리, 자신의 견해를 표현할 권리, 그러한 평가 후 도달된 결론에 관해 설명을 들을 권리, 결정에 이의를 제기할 권리가 포함되어야 한다. 그러한 조치는 어린이와는 관련이 없어야 한다.

37) EU의 AI Act는 발표 이후 의결/법제화가 된 사항은 아니기에 본 장에서는 논외로 다루었음

38) 그러나 이러한 역엔지니어링을 법적 규제를 직접적으로 설정하는 데에는 현실적인 어려움이 있을

(4) 정책적 제언

□ 정부는 인공지능 기술을 중심으로 디지털 전환의 미래사회 위험이슈 대응 방향을 모색할 필요

- 디지털 전환의 미래사회 위험이슈 대응은 인공지능 기술의 잠재적 위험 대응을 중심으로 설계 필요
- 인공지능 기술의 잠재적 위험 대응은 그 규범적 이상향/지향점인 알고리즘의 투명성 확보 여부를 중심으로 이루어질 공산이 높음
- 이에 국내 기술적/사회적 컨텍스트에 적합하고, 이해관계자들이 동의할 수 있는 수준의 인공지능 기술 투명성 기준을 모색하고, 인공지능 기술 개발자/서비스 제공자들이 해당 투명성을 확보하도록 어떻게 유도할 것인지를 고민해야 함

□ 정부는 디지털 전환의 미래사회 위험이슈 대응을 위한 다양한 방식의 규제 또는 규율을 검토하고, 큰 방향성을 설정할 시점

- 디지털 전환 미래사회 위험이슈에 대한 정책적 대응 방향은 정답이 없는 사안
- 이에 본 연구에서는 규제 수준에 따른 4개 유형의 정책 수단을 제시하였고, 각 유형별 법적 근거/사례 및 장/단점을 심도있게 분석하였음
- 정부는 해당 결과를 바탕으로 (현재의 인공지능 가이드라인 제작/배포 등의 정책에서 한 걸음 더 나아가) 국내의 사회적/규범적/법적 여건에 부합하는 다양한 검토 수단들을 검토해야 함
- 검토 이후에는 역기능 대응 정책의 큰 방향성(규제 수준 및 수단)을 설정하고, 이에 맞추어 세부 수단들을 정비하고, 구축해 나아가야 할 것으로 판단됨

※ 만약 정책적 대응 영역의 우선순위를 설정해야 한다면 본 연구에서 정량 분석한 국가별/기술별 디지털 전환 미래사회 위험의 세부 이슈 결과와 인식조사 결과 참고 가능

것으로 보임. 아직까지 알고리즘 판단 근거와 과정을 명확하게 보여줄 수 있는 소위 설명 가능한 인공지능(Explainable AI) 기술 구현이 일반적으로 가능한 것으로 언급하기 힘든 상황. 즉, 가역성 확보나 역엔지니어링이 가능하기 위해서는 이를 기술적으로 구현할 수 있어야 하지만 아직까지 그러한 기술은 연구 초기단계라고 할 수 있음.

참 고 문 헌

- 개인정보보호위원회, 국가정보-EU, https://www.privacy.go.kr/pic/nation_eu.do
- 김민식, 손가녕. (2017) 제4차 산업혁명과 디지털 트랜스포메이션 (Digital Transformation) 의 이해. 정보통신방송정책 29권, 3호: 26-32.
- 김승래, (2021), 디지털 전환시대 플랫폼 노동자의 법적 보호방안. 법학연구, 21(2), 1-30.
- 김승현. (2020), [혁신성장 전망_창업·중소·중견기업 정책 전망] 디지털 혁신 생태계로의 전환을 준비하자. 과학기술정책연구원.
- 김용진, (2018), 디지털전환 시대의 변화와 기업의 대응 방안. ie 매거진 대한산업공학회, 25(1), 20-24.
- 델 테크놀로지스, (2020), 디지털 트랜스포메이션 인덱스 2020 보고서
- 변순천, (2020), 2020년 과학기술혁신정책 핵심이슈 발굴 및 인텔리전스 기능 강화 연구, 한국과학기술기획평가원.
- 신동수, 이규환, 이재진, 주연희. (2021), 디지털 전환이 생산성 및 고용에 미치는 영향, 해외경제포커스 제 2021-22호, 한국은행 국제경제리뷰
- 이동임, (2019), 독일의 디지털 전환과 직업교육훈련. 한국기술혁신학회 학술대회, 1895-1922.
- 이명화, 최용인. (2017), [OECD] OECD 과학기술산업 지표를 통해 본 디지털 전환 동향과 도전과제. 과학기술정책, 27(12), 16-21.
- 이상직, EU 인공지능법(안) 시사점과 대응전략, 전자신문(`21.6.29), https://m.etnews.com/20210628000161
- 임희종, 최보름, 송지희. (2021), 기업의 디지털 전환 (DT) 경쟁력 분석 모형개발 및 적용: 공기업 10개의 사례를 중심으로. Korea Business Review, 25(3), 61-100.
- 장훈. (2017), [유럽] 디지털 전환과 노동의 미래. 과학기술정책, 27(11), 10-13.
- 한국마케팅연구원. (2019), Digital Transformation(디지털 전환) 전략과 문제점. 53(6) , 42-52.
- 한국정보화진흥원, (2019), 디지털트랜스포메이션 성공전략, IT & Future Strategy, 5호.

- AI for humanity, (2018), https://www.aiforhumanity.fr/en/
- Asia Society, (2020), AI PRINCIPLES IN CONTEXT, https://www.baai.ac.cn/news/beijing-ai-principles-en.html
- Bryce Goodman, Seth Flaxman, (2016), “European Union regulations on algorithmic decision-making and a "right to explanation"”
- European Parliament, (2017), European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL))
- Executive Office of the President, (2016), Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights
- Federal Trade Commission, (2016), “Big Data: A Tool for Inclusion or Exclusion?”
- Federal Trade Commission (2020), “Using Artificial intelligence and Algorithms”
- Finances Online, (2021), https://financesonline.com/digital- transformation - statistics/
- Future of Life Institute, (2020), AI POLICY - FRANCE, https://futureoflife.org/ai-policy-france/
- Future of Life Institute, (2020), AI POLICY - UNITED KINGDOM, https://futureoflife.org/ai-policy-united-kingdom/
- Future of Life Institute, (2019), AI POLICY - UNITED STATES, https://futureoflife.org/ai-policy-united-states/
- Future of Life Institute, (2020), AI POLICY - ITALY, https://futureoflife.org/ai-policy-italy/
- Future of Life Institute, (2020), AI POLICY - SWEDEN, https://futureoflife.org/ai-policy-sweden/
- Future of Life Institute, (2020), AI POLICY - CHINA, https://futureoflife.org/ai-policy-china/
- Future of Life Institute, (2020), AI POLICY - JAPAN, https://futureoflife.org/ai-policy-japan/

- Italy, AI Task Force(2018), LIBRO BIANCO SULL'INTELLIGENZA ARTIFICIALE AL SERVIZIO DEL CITTADINO
- Japan (2019), "Social Principles of Human-Centric AI"
- Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica, Machine Bias, ProPublica(`16.5.23), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing
- OECD, (2019), Going Digital: Shaping Policies, Improving Lives
- OECD, (2019), Artificial intelligence, blockchain, big data, IoT, cloud computing, 5G networks
- Sandra Wachter, Brent Mittelstadt, Luciano Floridi, (2017), "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation", International Data Privacy Law 7(2)
- Tabrizi, B. N., Lam, E., Girard, K., & Irvin, V,. (2019), Digital transformation is not about technology. HarvardBusiness Review, 13.
- U.S. Department of Defense, (2020), "DOD Adopts 5 Principles of Artificial Intelligence Ethics"
- Verhoef, P. C., Broekhuizen, T., Bart, Y., Bhattacharya, A., Dong, J. Q., Fabian, N., & Haenlein, M,. (2021), Digital transformation: A multidisciplinary reflection and research agenda. Journal of Business Research, 122, 889-901
- Vial, G. (2019), Understanding digital transformation: A review and a research agenda, The Journal of Strategic Information Systems, Volume 28, Issue 2, 118-144.
- ＡＩネットワーク社会推進会議事務局(総務省情報通信政策研究所調査研究部), (2016),「ＡＩ開発ガイドライン」(仮称)の策定に向けた国際的議論の用に供する素案の作成に関する論点, p. 51.

1. 이 보고서는 한국과학기술기획평가원에서 수행하는 연구보고서 입니다.

2. 이 보고서 내용을 발표할 때에는 반드시 한국과학기술기획평가원에서 수행한 연구결과임을 밝혀야 합니다.

3. 국가과학기술 기밀유지에 필요한 내용은 대외적으로 발표 또는 공개하여서는 아니됩니다.

KISTEP 한국과학기술기획평가원
Korea Institute of S&T Evaluation and Planning

(27740) 충청북도 음성군 맹동면 원중로 1339
TEL : (043) 750-2300 FAX : (043) 750-2680
http://www.kistep.re.kr