

통권 제325호

# 디지털 전환의 미래사회 위험이슈 및 대응 전략: 인공지능 역기능을 중심으로

KISTEP 미래성장전략센터 구본진





# 디지털 전환의 미래사회 위험이슈 및 대응 전략: 인공지능 역기능을 중심으로

A Study on Risk Issues and Response Strategies related to  
Digital Transformation: Focusing on Negative Effects of Artificial Intelligence

구본진

Bonjin Koo

I. 작성 배경

II. 디지털 전환의 정의 및 파급효과

III. 디지털 전환의 미래사회 위험이슈

IV. 디지털 전환 관련 미래사회 위험이슈  
대응 정책 동향

V. 디지털 전환의 미래사회 위험이슈 대응  
정책 제언

VI. 결론

[참고문헌]

I. Introduction

II. Definition of Digital Transformation  
and its Impacts

III. Risk Issues of Digital Transformation

IV. Policy Trends in Response to Risk  
Issues of Digital Transformation

V. Proposal of Policies to Respond to  
Risk Issues of Digital Transformation

VI. Conclusion

[References]





## 요약

### ■ 작성 배경

- 디지털 전환은 경제·사회 발전을 촉진하는 유용한 도구로 작용하지만, 자칫 부정적 영향을 미칠 수도 있는 양면성을 내포
- 정부는 경쟁력 및 기술 패권 확보 등을 위하여 디지털 전환 촉진정책을 적극적으로 추진하고 있는 반면, 디지털 전환의 위험에 대한 이해와 이를 방지할 수 있는 정책 추진은 상대적으로 더딘 경향
- 정부는 디지털 전환에 대한 이해도 제고를 통한 정책 범주 명확화, 디지털 전환의 세부 위험 파악, 대응 정책 수단 검토 및 방향 설정 등을 통해 디지털 전환 미래사회 위험 대응 방안을 마련할 필요

### ■ 문제 정의 및 분석

- 디지털 전환의 개념 정리 및 파급효과 분석 (문헌연구)
- 디지털 전환이 초래할 수 있는 미래사회 위험이슈 구체화 (국내외 뉴스 데이터 기반 Embedded Topic Modeling)
- 주요국의 디지털 전환 역기능 대응 정책 동향 정리 (문헌 연구)
- 디지털 전환 역기능 대응 정책에 대한 법적 검토 및 규제수준별 정책수단 제시/분석

### ■ 결론 및 정책제언

- 디지털 전환의 학술적·실무적 정의 및 개념을 종합하여 정책 범주 명확화
- 관련 언론기사들을 대규모로 수집하고, Embedded Topic Modeling 분석을 적용하여 최근 이슈화되고 있는 국내외 디지털 전환 역기능 세부 위험들을 규명 및 분류
- 한국을 포함한 주요국 정부의 대응 정책 동향 분석

- 디지털 전환 핵심기술인 인공지능에 대한 정부 규제 타당성 검토 및 향후 정책적 대응 근거 확보
- 향후 정부가 선택 가능한 규제 수준별 정책 수단 제시 및 분석

정책 수단	규제강도	내용	장/단점
전문가 윤리적 접근	가장 낮음	관련 기술 개발자/서비스 제공자가 전문가 집단의 참여/협력을 통해 자발적으로 통제를 추구하는 방식	<ul style="list-style-type: none"> <li>• (장점) 자율기반의 관련 기술/생태계 발전 촉진</li> <li>• (단점) 전문가 위원회 구성에 대한 객관성 결여 위험, 예측 못한 문제 발생 시 법적 책임 모호</li> </ul>
인증체계 구축	다소 낮음	정부가 이해관계자에게 자발적인 인증 획득을 유도하여 관련 분야 위험 대응 체계를 구축할 수 있도록 유인하는 방식	<ul style="list-style-type: none"> <li>• (장점) 물리적인 대응 체계 구축/운영 가능</li> <li>• (단점) 인증 주체 지정, 인증 수준 및 절차에 대한 이해관계자 합의에 오랜 시간 소요</li> </ul>
개인적 권리 설정	다소 높음	정부가 이용자들에게 알고리즘 구조 및 결과에 대한 설명을 기술개발자/서비스 제공자들에게 요구할 수 있는 권한을 부여하는 방식	<ul style="list-style-type: none"> <li>• (장점) 인공지능 기술/제품/서비스 개발자가 딥러닝의 인과관계 모호성을 최소화하여 제품/서비스를 개발할 수 있도록 유도</li> <li>• (단점) 기업 부담 가중, 설명 수준에 대한 규정 등을 설정하는데 상당한 시간 소요</li> </ul>
직접적인 행정 규제 설정	매우 높음	정부가 기술개발/서비스 제공 가능 영역을 설정하고, 이외 영역에 대해서는 법적 규제를 가하는 방식	<ul style="list-style-type: none"> <li>• (장점) 역기능의 피해 및 법적 분쟁 최소화</li> <li>• (단점) 진화하고 있는 인공지능 기술을 법으로 명확하게 규제하기 어려운 측면</li> </ul>

※ 본 이슈페이퍼는 2021년 KISTEP 기관고유 과제 ‘2021년 과학기술정책 핵심이슈 발굴 및 인텔리전스 기능 강화’로 수행한 내용으로 필자의 개인적인 견해이며, 한국과학기술기획평가원의 공식적인 의견이 아님을 알려드립니다.



## Abstract

### ■ Introduction

- Although digital transformation acts as a useful tool to promote economic and social development, it may have negative effects.
- While the government is actively making policies to promote digital transformation in order to secure competitiveness and technological hegemony, understanding the risks of digital transformation and implementing policies to prevent it tend to be relatively slow.
- The government needs to prepare measures to respond to risks in the future society of digital transformation by clarifying policy categories with enhancing understanding of digital transformation, identifying detailed risks of digital transformation, reviewing response policy measures, and setting directions.

### ■ Problem Statement and Analysis

- Conceptual study of digital transformation and analysis of its impacts. (literature review)
- Reifying risk issues of digital transformation. (Embedded Topic Modeling based on domestic and international news big data)
- Summary of policy trends in response to negative effects in digital transformation in other countries. (literature review)
- Legal review of digital transformation dysfunctional response policies and suggestion/analyzing of policy measures.

## ■ Conclusion and Discussions

- Clarify policy categories by reviewing academic/practical definitions and concepts of digital transformation.
- By collecting related news data on a large scale and applying Embedded Topic Modeling analysis, the detailed risks of domestic and international digital transformation dysfunction that have recently become an issue are identified and analyzed.
- Analyzing policy trends of major countries including South Korea.
- Reviewing the feasibility of government regulations on artificial intelligence, a key technology of digital transformation, and securing grounds for future policy responses.
- Suggesting and analyzing policy measures for each regulatory level that the government can select in the future.



## I

## 작성배경

■ 디지털 전환(digital transformation)\*은 경제/사회 및 일상에 빠르게 접목되고 있고, 이는 더욱 가속화될 전망\*\*

\* (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)

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

- 델 테크놀로지스가 '20년 전세계 18개국 대/중견기업 C레벨 및 관리직 임원 4,300여명을 대상으로 조사한 결과, 최근 3년간 디지털 선도그룹(디지털 기술 활용 수준 1위 그룹), 디지털 어댑터(디지털 기술 활용 수준 2위 그룹), 디지털 평가그룹(디지털 기술 활용 수준 3위 그룹)이 지속적으로 증가 (델 테크놀로지스 디지털 트랜스포메이션 인덱스 2020)
- 디지털 전환 관련 핵심 기술들(인공지능, 빅데이터, 5G network 등)의 빠른 발전과 Covid-19에 따른 디지털 수요 급증은 전분야의 디지털 전환을 가속 중

■ 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 향상하는 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 전환의 위험에 대한 이해와 이를 방지할 수 있는 정책 추진은 상대적으로 더딘 경향
- 따라서 정부는 디지털 전환의 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

## II

## 디지털 전환의 정의 및 파급효과

- (디지털 전환의 정의) 디지털 전환은 합의된 정의가 부재하며 그 범주도 모호한 경향이 있기 때문에 본 절에서는 문헌연구를 통해 디지털 전환의 정의를 종합

〈표 1〉 디지털 전환의 정의

출처	정의
장훈 (2017)	디지털 전환은 디지털 기술이 바탕이 된 전환, 차별화된 변화를 의미하며 산업적 시각에서는 '다양한 디지털 기술을 바탕으로 기업의 전략이나 시스템 등을 근본적으로 변화시켜 새로운 가치를 창출하는 것'으로 정의 가능
한국정보화진흥원 (2019)	인공지능 (AI), 클라우드 (Cloud), 데이터 (Data) 등 디지털 기술기반 비즈니스 모델을 중심으로 산업 구조 재편하는 것
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

자료: 구본진 (2021) 중 일부 발췌

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

〈표 2〉 디지털 전환의 파급효과

출처	파급효과	내용
장훈 (2017)	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)
임희종 외 (2021)	운영 프로세스 자동화	기업은 Robotic Process Automation(RPA) 도입으로 생산성 향상과 기업의 응답성(Responsiveness)이 향상되고, 민첩성(Agility) 향상을 도모할 것이며 이를 통해 인건비 절감뿐만 아니라 일관된 프로세스의 실행이 가능해짐
	조직문화 변화	디지털 전환에 맞추어 직원들의 필수 역량이 변화(관계적 민첩성, 탄력성 등의 비중 증대)하고 있고, 이에 맞추어 조직문화도 변화하고 있음
	프로세스 효율화	기업은 디지털 전환을 통해 기존 아날로그 기반 또는 디지털화 기반의 프로세스를 혁신시켜 프로세스 전반의 효율성을 극대화
김용진 (2018)	BM 변화	제품/서비스가 디지털화되고 전달체계가 디지털화되며, 생산 운영체계가 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
김승래 (2021)	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
이명화 & 최용인 (2017)	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 증가
김승현 (2020)	디지털 혁신생태계 발달	디지털 혁신생태계란 기존 산업생태계가 디지털 환경기반으로 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 이중 산업생태계로의 진입/진출을 촉진)
Lomi et al. (2014)	조직형태 및 구조 변화	디지털 전환 가속화로 가상 협업, 온라인 커뮤니티, 오픈소스 프로젝트, 깃(Git) 이코노미 등과 같은 새로운 조직 형태가 등장
Wu et al. (2016)	공급망 사슬 통합	공급망의 디지털화가 고도화되고, 정보의 공유 및 의사결정 조율의 수준이 높아지게 되며 이러한 정보의 공유는 통합의 기초가 됨. 다음 단계로 조율(coordination)과 자원의 공유, 그리고 조직의 연결을 통해 리스크, 비용, 이익을 공유하게 됨

자료: 구본진 (2021) 중 일부 발췌

# III

## 디지털 전환의 미래사회 위험이슈

■ (분석 개요) 본 절에서는 Embedded Topic Modeling\* 방법으로 국내외 디지털 전환 역기능 관련 뉴스 빅데이터를 분석하여 세부 이슈들을 도출 및 유형화\*\*

\* BERT 아키텍처에 사전학습된 언어모델(distiluse-base-multilingual-cased)을 활용한 LDA 토픽모델 활용

\*\* 「Principled AI」 (Harvard Berkman Klein Research Center, 2020) 분석 결과를 바탕으로 디지털 전환 역기능을 8개로 유형화(프라이버시, 책임성, 안전 및 보안, 투명성 및 설명 가능성, 공정성 및 차별 금지, 기술에 대한 인간의 통제, 직업적 책임, 인간 가치 증진)

※ (뉴스 빅데이터 선정 관련) 디지털 전환 미래사회 위험이슈는 비교적 최근에 부각된 이슈이므로 현재 관련 논문, 특허 등은 부족한 상황이며 이를 포착할 수 있는 정량화가 가능한 데이터는 뉴스 데이터가 유일. 물론 뉴스 데이터는 현재 발생하고 있는 문제들을 논하는 데이터이므로 미래를 예측하는 데이터로 활용하기에 한계가 존재하는 것이 사실이나 본 연구에서는 대규모 국내외 뉴스 데이터를 수집/활용함으로써 다양한 컨텍스트에서 발생하는 여러 가지 이슈들을 반영하였기에 향후 파생가능한 문제들을 추정해볼 수 있는 자료로 활용할 수 있다고 사료됨. 또한 본 연구는 대규모 뉴스 데이터를 활용함으로써 언론사별 편향성 등의 문제를 일정 수준에서 해결하였다고 판단됨

■ (국내 디지털 전환의 역기능 세부이슈 분석) 국내에서 사회적으로 가장 크게 이슈가 됐던 디지털 전환의 역기능 사례(‘이루다’) 관련 뉴스 데이터를 수집 및 정량분석\*하여 세부 위험 이슈들을 도출

\* (이루다 케이스 선정 관련) 연구 초반 국내 디지털 전환 역기능 분석도 주요국 디지털 전환 역기능 분석 방법과 동일한 방식으로 뉴스 데이터를 수집하였으나 다양한 연관 사건들과 여러 가지 관련 주제를 다룬 뉴스 데이터가 풍부한 주요국과는 달리 국내는 ‘이루다’ 사건 관련 뉴스가 지배적이었음. 따라서 해당 케이스가 대표성을 가지는 국내 디지털 전환 역기능 사례로 판단하여 이에 포커스하여 분석을 수행함

〈표 3〉 인공지능 챗봇 서비스 ‘이루다’

구분	내용
개요	‘이루다’는 스캐터랩(ScatterLab) 소속의 핑퐁 팀이 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

- (데이터) 2020.11.01 ~ 2021.04.30. 기간에 이루다 사건을 다룬 국내 언론 기사/사설 1,329개(전처리\* 후 데이터: 1,100개)

\* ① 뉴스 데이터 URL 기준 중복기사제거, ② 뉴스 제목 및 본문 첫 문장 기준 주제 연관성 낮은 기사 제거, ③ API 업체 제공 관련도 점수(relevancy score) 기준 하위 80% 기사 제거, ④ 본문 길이 5문장 이내의 뉴스기사 제거 등을 통해 연관성이 낮은 기사들을 최대한 제거한 뒤 데이터 분석 수행

- (분석) 수집 및 전처리한 뉴스 데이터를 Embedded Topic Modeling 방법으로 분석하여 디지털 전환 역기능 세부 이슈 도출
- (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 언론 기사 내용 매칭을 통한 해석 작업을 거쳐 최종적으로 11개의 세부주제들과 이에 매칭된 키워드들을 정리하고, 이를 유형화

〈표 4〉 ‘이루다’ 관련 디지털 전환의 역기능 위험 세부주제 도출 및 유형화 결과

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

- (결과 분석) 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 처리되지 못한 개인정보가 두 번째로 많이 지적되었음을 확인하였고, 이는 프라이버시 침해 유형에 해당
  - (소결) 한국은 윤리적 문제들이 디지털 전환 역기능 이슈의 주류를 차지
    - 한국은 (전술한 바와 같이) 본 연구의 데이터 수집 기간 중 '이루다' 사건이 거의 유일한 디지털 전환 역기능 관련 기사였음
    - 이로 인해 본 연구는 이루다 사건 뉴스에 집중하여 분석을 수행하였고, 분석 결과 인공지능 기술 기반 챗봇 서비스에 얽힌 역기능 이슈들이 세부주제로 도출되었음
    - 해당 세부주제들은 타 분야/기술과 결합되어 복합화/가시화되어 나타나는 역기능보다는 다소 원론적인 역기능(혐오, 차별, 인권침해 등)이 주를 이루고 있음을 확인
- (주요국 디지털 전환의 역기능 세부이슈 분석) 미국, 중국, 유럽, 일본 언론기사에서 다루어진 디지털 전환의 역기능 세부 위험 이슈들을 도출
- (데이터) 2020.01.01 ~ 2021.08.23. 기간에 주요국(미국, 일본, 중국, 유럽)에서 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)
 

\* ① 뉴스 데이터 URL 기준 중복기사제거, ② 뉴스 제목 및 본문 첫 문장 기준 주제 연관성 낮은 기사 제거, ③ API 업체 제공 관련도 점수(relevancy score) 기준 하위 80% 기사 제거, ④ 본문 길이 5문장 이내의 뉴스기사 제거 등을 통해 연관성이 낮은 기사들을 최대한 제거한 뒤 데이터 분석 수행
  - (분석 1) 전처리 과정을 거친 전체 데이터를 대상으로 Embedded Topic Modeling 분석을 수행(분석 결과에 대하여 연관주제별 군집화 및 연관성 높은 언론기사 매칭 작업을 통해 총 9개 세부 이슈들을 도출하여 이를 유형화하고, 빈도순으로 정리)
  - (분석 결과 1) 무기용 인공지능 문제와 인공지능의 편향성(bias) 문제가 높은 빈도를 차지하였고, 다음으로는 딥페이크 문제, IoT 기기의 보안 문제, 그리고 인공지능 윤리 문제가 사회적으로 많이 회자됐음을 확인

- (분석 2) 도출된 9개 역기능 이슈들과 가장 밀접하게 연관된 디지털 전환 구성 기술을 규명하기 위하여 유형별 연관 기술의 빈도분석을 수행
- (분석 결과 2) 기존의 예측 및 논의 결과와 동일하게 인공지능 기술이 디지털 전환의 미래사회 위험이슈와 가장 밀접한 관계가 있음을 확인(이는 인공지능 기술이 디지털 전환에 있어 핵심기술로 작용하기에 타 구성기술대비 역기능 연관성도 가장 높았던 것에 기인한 결과로 추정)

〈표 5〉 Embedded Topic Modeling 분석 결과(역기능 세부이슈별 기술 연관성 분석 포함)

세부주제	디지털 전환 역기능 유형	빈도수	5G networks	IoT	AI	big data	block chain
무기용 인공지능 (AI for weapon)	안전 및 보안	190	1	10	175	2	2
인공지능의 편향성 (biased AI profiling)	불공정성 및 차별 강화	127		10	107	7	3
딥페이크 (Facebook banning deepfakes)	안전 및 보안	71			66	0	5
보안 (IoT vulnerabilities)	안전 및 보안	63		60	2	0	1
인공지능 윤리 (Ethics in AI systems)	(매칭 제외)	62			61	1	0
프라이버시 (Data privacy risk)	프라이버시 침해	41			17	12	12
인공지능의 사회적 차별 심화 (AI discrimination in credit lending)	불공정성 및 차별강화	11			11	0	0

- (분석 3) 주요국 디지털 전환 역기능 관련 뉴스 데이터 중 인공지능 기술에 관련된 언론기사에 한정하여 Embedded Topic Modeling을 수행

〈표 6〉 주요국 인공지능 미래사회 위험 세부이슈 도출 결과

세부주제	디지털 전환 역기능 유형	관련 키워드	빈도수
생체 데이터 사용 문제에 대한 EU의 규제	프라이버시 침해	EU, Vestager, regulation, commission, regulations, ban, biometric	232



세부주제	디지털 전환 역기능 유형	관련 키워드	빈도수
사이버 안보, 사이버 공격/위협	안전 및 보안	security, cybersecurity, intrusion, threats, attacks, cyberattacks, risk	165
경찰의 감시 카메라 사용에 대한 인종차별, 편향성, 프라이버시 침해	프라이버시 침해, 인간 가치 훼손	police, recognition, crime, cameras, surveillance, bias, privacy	90
Facebook의 deepfake 비디오 사용을 통한 허위정보, 오보 문제	안전 및 보안	deepfakes, videos, facebook, disinformation, misinformation	68
인종/성 편향 알고리즘 이슈	불공정성 및 차별강화	bias, algorithms, black, gender, race	63

- (분석 결과 3) 생체 데이터 활용 문제가 가장 큰 비중을 차지하였고, 다음으로 안보/보안 위험, 프라이버시 침해, 편향성 이슈가 디지털 전환 위험 세부이슈의 비중을 차지
- (소결) 주요국은 다양하고 복합적인 문제들이 디지털 전환 역기능 이슈로 등장
  - 주요국은 데이터 수집 기간 동안 윤리적인 논의들이 주를 이루고 있던 국내와 달리 디지털 전환 역기능 세부 이슈들이 다양한 분야/기술과 결합되어 가시화되고, 복합화되어 나타나고 있었음
  - 이는 국내 대비 관련 기술수준이 높고, 연관 서비스들이 빠르게 론칭되어 일상에 정착된 영향도 있겠지만 이와 동시에 상대적으로 오래전부터 해당 역기능에 대한 사회적 관심이 있었음을 나타내는 결과

■ (종합 분석 및 시사점) 향후 디지털 전환의 미래사회 위험이슈 대응을 위해서 정부는 인공지능 기술로부터 파생할 수 있는 역기능에 주목할 필요가 있고, 컨텍스트에 맞춘 정책을 모색할 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 때문
  - \* blockchain, big data, IoT, cloud computing, 5G networks
- 따라서 정부는 다른 디지털 전환 구성기술보다는 인공지능 기술에 집중하여 역기능 대응을 위한 정책을 고민할 필요가 있음

- 아울러 국내도 관련 기술수준이 높아지고, 연관 서비스들이 확대되면 현재 주요국에서 논의되고 있는 역기능 이슈들이 국내 기술/환경/사회/정치적 맥락에 따라 시차를 두고 문제화될 가능성이 높음
- 이에 정부는 디지털 전환 위험 이슈 대응 정책 설계 시 사회적 맥락을 고려하여 우선 대응 영역을 설정하고, 이후 중장기적으로 위험 이슈 전반에 대응하는 방향으로 정책을 추진할 필요

## Ⅳ

## 디지털 전환 관련 미래사회 위험이슈 대응 정책 동향

■ (개요) 본 절에서는 주요국 및 국내 디지털 전환 관련 미래사회 위험이슈 대응 정책 동향을 2개 유형(소극적/적극적 정책 대응)으로 구분하여 정리

- 디지털 전환 관련 미래사회 위험이슈 대응 정책은 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 정책을 수립/추진
- 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 정책을 수립/추진

■ (소극적 정책 대응) 미국, 유럽 개별 국가, 중국, 일본, 한국 정부 등

- (미국) 미국은 국방부, 백악관, FTC를 중심으로 AI 사용 원칙, 가이드라인을 발표하며 낮은 규제 수준의 정책을 추진 중
  - '16년 미국연방거래위원회(Federal Trade Commission, FTC)는 빅데이터 또는 머신러닝을 사용하는 기업에게 공정성을 확보할 수 있도록 가이드라인을 발간
  - '18년 Defense Innovation Board(DIB)는 'AI Principles Project'를 통해 5개 AI 시스템 사용 원칙(책임감, 공정함, 추적 가능, 신뢰성, 관리 가능)을 제안
  - '19년 2월 도널드 트럼프 대통령은 “미국 AI 이니셔티브”를 시작하는 행정명령 발표
    - ※ (주요 내용) ①기술 혁신 추진, ②기술 표준 개발 추진, ③AI 기술 개발 및 적용 기술을 갖춘 작업자 교육, ④시민의 자유와 프라이버시를 포함한 미국의 가치를 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

- '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 제시(책임성(responsible), 공평성(equitable), 추적 가능성(traceable), 신뢰성(reliable), 통제 가능성(governable))
- (유럽-개별 국가) 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)
  - (독일) '18년 연방법무부 및 소비자 보호부와 연방 내무부는 개인 보호, 사회적 결속 유지, 정보화 시대의 변영 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.)
  - (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)
  - (스웨덴) AI 개발 및 사용에 필요한 정부의 평가를 요약한 "Artificial intelligence in Swedish business and society: Analysis of development and potential" 발표
- (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 신조를 포함하는 차세대 인공지능 거버넌스 원칙을 발표('19.6)
  - '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

※ (주요 내용) (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

- '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)

※ (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 지방 정부를 포함한 일본 사회 전반에 걸친 'AI-Ready Society' 원칙 제시

※ (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 보장 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)

- (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 AI 실현전략'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

## ■ (적극적 정책 대응) 유럽연합은 적극적인 디지털 전환 역기능 정책을 추진 중

- (EU: GDPR) EU는 EU 회원국에 일괄적으로 적용되는 개인정보 보호법(General Data Protection Regulation, GDPR)을 제정('16년) 및 시행('18년)
  - (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
  - (적용 대상) EU 내에 사업장을 운영하며, 개인정보를 처리하는 기업, EU 거주자에게 재화나 서비스를 제공하는 기업, EU 거주자의 EU 내 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(직장 내 업무수행, 경제 상황, 건강, 개인적 취향 등)을 분석, 예측하기 위한 자동화된 처리
  - (위반 시 과징금 부과 규정) 최대 과징금은 일반적 위반 사항인 경우 전 세계 매출액의 2% 혹은 1천만 유로(약 125억 원) 중 높은 금액이며, 중요한 위반 사항인 경우 전 세계 매출액의 4% 혹은 2천만 유로(약 250억 원) 중 높은 금액

〈표 7〉 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

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

〈표 8〉 EU AI Act의 인공지능 위험 유형과 유형별 규제 사항

인공지능 위험 유형	위험 유형별 정의	위험 유형별 규제
수용불가 위험	<ul style="list-style-type: none"> <li>• AI 시스템이 사람 행동을 조작하는 경우</li> <li>• 사회적 약자를 이용/공격하는 경우</li> </ul>	<ul style="list-style-type: none"> <li>• 시장 출시 불가</li> <li>• (위반시) 최대 3000만 유로, 세계 연간 매출액의 6% 중 높은 금액의 벌금을 부과</li> </ul>
고위험 AI	<ul style="list-style-type: none"> <li>• AI 시스템이 운송/교통, 기계, 무선장비 및 의료기기 등 사람의 안전과 관계된 경우</li> <li>• 가스, 전기 등 중요 인프라에 이용되는 경우</li> <li>• 잠재적 가해자 또는 피해자가 될 위험 평가, 거짓말 탐지 등 감정 상태 확인, 증거 신뢰성 평가, 범죄 분석에 사용되는 경우 등</li> </ul>	<ul style="list-style-type: none"> <li>• 허용하되 위험관리시스템을 기획/설계/구축할 의무 부과</li> <li>• 데이터 거버넌스 구축, 기술문서 마련, 기록 투명성 확보, 인간에 의한 감독, 정확도, 보안의 확보 등이 요구</li> <li>• 감독기관의 모니터링 대상이 되고, 상용화 이후에도 문제 상황의 발생에 따른 보고 의무 등의 의무 부과</li> </ul>
저위험, 최소위험	<ul style="list-style-type: none"> <li>• 게임 등 권리침해, 안전위험이 최소화된 경우</li> </ul>	<ul style="list-style-type: none"> <li>• 규제하지 않음</li> </ul>

자료: EU 인공지능법(안) 시사점과 대응전략, 전자신문('21.6.29)

- (종합 분석 및 시사점) 주요국은 디지털 전환 기술 중 인공지능 기술의 양면성에 주목하여 이것의 역기능 대응을 위한 정책 마련을 신속하게 추진하였고, 향후 효과적인 디지털 전환 위험 이슈 대응 정책 마련을 위해서는 규제 수준에 따라 다양한 정책적 수단을 검토할 필요
- 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

- 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 중
- 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



## V

## 디지털 전환의 미래사회 위험이슈 대응 정책 제언

■ (개요) 본 절에서는 디지털 전환의 위험 대응 정책 수립을 위하여 인공지능의 법규범적 속성을 검토하고, 규제 강도에 따른 정책 수단을 제시

- 디지털 전환 관련 미래사회 위험이슈 대응을 위해서는 일방적 규제보다는 다양한 정책 수단과 국내 사회적 맥락을 검토할 필요
  - 자율규제 형식의 정책은 산업 및 기업의 성장을 촉진할 수 있다는 장점이 있으나 예상치 못한 큰 경제/사회적 문제에 선제적으로 대응이 어려워 문제 발생 시 국가와 사회에 상당히 큰 피해를 입힐 수 있다는 단점이 존재
  - 반대로 강한 법적 규제 형식의 정책은 파생될 수 있는 미래사회 위험을 최소화할 수 있는 장점이 있는 반면, 산업/기업의 성장을 저해할 수 있고, 주요국 대비 기술 경쟁력을 저하시킬 수 있다는 단점이 존재
- 이에 본 절에서는 디지털 전환의 핵심 기술인 인공지능의 법규범적 속성을 검토하고, 이에 대하여 정부가 고려가능한 다양한 정책 수단을 도출

■ (인공지능의 법규범적 속성 검토) 인공지능 기술은 알고리즘의 법규범적 불확정성으로 인하여 규제에 한계가 존재하므로 향후 법적 대응은 알고리즘의 투명성 확보 여부를 중심으로 이루어질 것

- 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
- 인공지능 알고리즘이 불확정적인 속성을 가진다고 했을 때, 가장 유효한 대응방식은 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
- 이와 관련하여 쟁점이 되는 2가지 사항은 알고리즘 투명성 확보에 있어 인간의 개입이 불가피함을 시사
  - (차별과 편향) 특정 계층으로부터만 추출된 불완전한 데이터가 빅데이터 알고리즘에 입력되면, 여기에 데이터가 반영되지 못한 이들은 빅데이터 기술의 혜택으로부터 소외될 가능성이 높고, 이런 데이터 피드백이 반복되어 누적되면 사회적 차별이 고착화될 가능성이 존재하고, 알고리즘 설계 과정이 투명하지 않거나 자동화된 알고리즘이 ‘블랙박스’로 남게 되면, 그 안에 내재된 차별적 요소를 발견해내거나 그에 저항하는 일이 어려워짐(Executive Office of the President (2016), FTC (2016) 참고)
  - (프라이버시 침해 가능성) 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 되었고, 이와 같은 개인정보의 처리에 인간의 개입이 최소화되고 있음(즉, 인공지능 알고리즘이 특정 개인에 관하여 내린 판단이 어떠한 기준과 근거를 가지고 이루어진 것인지를 알지 못하는 경우가 빈번하게 발생할 가능성이 높음)

■ (규제 강도에 따른 정책 시나리오 분석) 규제 강도에 따른 정책 수단을 제시 및 분석

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

[그림 1] 규제 강도에 따른 정책 수단

- (정책 수단 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
- (정책 수단 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 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
  - 일반적으로 주요 국가들의 인증 관련 제도화 동향은 민간 인증사업자들의 인증을 전제로 운영되는 경우가 대다수
- (정책 수단 3: 개인적 권리의 설정) 알고리즘을 개발 및 서비스하는 사업자로 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
- (정책 수단 4: 직접적인 행정 규제 설정) 정부가 인공지능 기술 개발 및 서비스 제공 가능 영역을 설정하고, 이외 영역에 대해서는 기술 개발과 서비스 제공을 금지하는 방식
- 직접적인 행정규제를 설정하는 방식은 현 단계에서 심도 있는 논의가 이루어지고 있지는 못한 상황
- 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
  - ※ (로봇공학 기술자를 위한 윤리 행동 강령: 가역성) 제어성의 필수 조건인 가역성은 로봇을 안전하고 안정적으로 동작하도록 프로그래밍 할 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

## VII 결 론

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

- 디지털 전환의 미래사회 위험이슈 대응은 인공지능 기술의 잠재적 위험 대응을 중심으로 설계될 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

- 현재 디지털 전환 미래사회 위험이슈에 대한 정책적 대응 방향은 아직 정답이 없는 사안
- 이에 본 연구는 규제 수준에 따른 4개 유형의 정책 수단을 제시하였고, 각 유형별 법적 근거/사례 및 장/단점을 심도 있게 분석하였음

〈표 9〉 정책 수단별 주요 사항

정책 수단	규제강도	내용	장/단점
전문가 윤리적 접근	가장 낮음	관련 기술 개발자/서비스 제공자가 전문가 집단의 참여/협력을 통해 자발적으로 통제를 추구하는 방식	<ul style="list-style-type: none"> <li>• (장점) 자율기반의 관련 기술/생태계 발전 촉진</li> <li>• (단점) 전문가 위원회 구성에 대한 객관성 결여 위험, 예측 못한 문제 발생 시 법적 책임 모호</li> </ul>
인증체계 구축	다소 낮음	정부가 이해관계자에게 자발적인 인증 획득을 유도하여 관련 분야 위험 대응 체계를 구축할 수 있도록 유인하는 방식	<ul style="list-style-type: none"> <li>• (장점) 물리적인 대응 체계 구축/운영 가능</li> <li>• (단점) 인증 주체 지정, 인증 수준 및 절차에 대한 이해관계자 합의에 오랜 시간 소요</li> </ul>
개인적 권리 설정	다소 높음	정부가 이용자들에게 알고리즘 구조 및 결과에 대한 설명을	<ul style="list-style-type: none"> <li>• (장점) 인공지능 기술/제품/서비스 개발자가 딥러닝의 인과관계 모호성을 최소화하여</li> </ul>

정책 수단	규제강도	내용	장/단점
		기술개발자/서비스 제공자들에게 요구할 수 있는 권한을 부여하는 방식	제품/서비스를 개발할 수 있도록 유도 • (단점) 기업 부담 가중, 설명 수준에 대한 규정 등을 설정하는데 상당한 시간 소요
직접적인 행정 규제 설정	매우 높음	정부가 기술개발/서비스 제공 가능 영역을 설정하고, 이외 영역에 대해서는 법적 규제를 가하는 방식	• (장점) 역기능의 피해 및 법적 분쟁 최소화 • (단점) 진화하고 있는 인공지능 기술을 법으로 명확하게 규제하기 어려운 측면

- 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

## 참 고 문 헌


- 구본진 (2021), 디지털 전환에 따른 미래사회 위험이슈 발굴 및 대응 전략 연구, KISTEP 2021-032호
- 김승래 (2021), 디지털 전환시대 플랫폼 노동자의 법적 보호방안. 법학연구, 21(2), pp. 1-30.
- 김승현 (2020), [혁신성장 전망\_창업·중소·중견기업 정책 전망] 디지털 혁신생태계로의 전환을 준비하자. 과학기술정책연구원.
- 김용진 (2018), 디지털전환 시대의 변화와 기업의 대응 방안. ie 매거진 대한산업공학회, 25(1), pp. 20-24.
- 델 테크놀로지스 (2020), 디지털 트랜스포메이션 인덱스 2020 보고서
- 이명화, 최용인 (2017), [OECD] OECD 과학기술산업 지표를 통해 본 디지털 전환 동향과 도전 과제. 과학기술정책, 27(12), pp. 16-21.
- 이상직, EU 인공지능법(안) 시사점과 대응전략, 전자신문, 2021. 6. 29.
- 임희중, 최보름, 송지희 (2021), 기업의 디지털 전환 (DT) 경쟁력 분석 모형개발 및 적용: 공기업 10개의 사례를 중심으로. Korea Business Review, 25(3), pp. 61-100.
- 장훈 (2017), [유럽] 디지털 전환과 노동의 미래. 과학기술정책, 27(11), pp. 10-13.
- 한국정보화진흥원 (2019), 디지털트랜스포메이션 성공전략, IT & Future Strategy, 5호.
- Agile Elephant (2015), <http://www.theagileelephant.com/what-is-digital-transformation/>
- A.T. Kearney (2016), Digital Transformation
- 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?”
- Harvard Berkman Klein Research Center (2020), <https://cyber.harvard.edu/topics/ethics-and-governance-ai>
- IBM (2011), Digital Transformation, IBM Global Business Services
- IDC (2015), <https://www.idc.com/itexecutive/research/dx>

- Lomi, A., Conaldi, G., Tonellato, M., & Pallotti, F. (2014). Participation motifs and the emergence of organization in open productions. *Structural Change and Economic Dynamics*, 29, pp. 40-57.
- OECD (2019), *Going Digital: Shaping Policies, Improving Lives*
- OECD (2019), Artificial intelligence, blockchain, big data, IoT, cloud computing, 5G networks
- Research & Markets (2020), <https://www.globenewswire.com/news-release/2020/08/14/2078517/0/en/The-World-s-Digital-Transformation-Industry-2020-2025-Trends-Opportunities-and-Competitive-Landscape.html>
- Wu, L., Yue, X., Jin, A. & Yen, D.C.(2016). Smart supply chain management: A review and implications for future research, *The International Journal of Logistics Management*, 27(2), pp. 395-417.



## KISTEP 이슈페이퍼 발간목록

발간호	제목	저자
2022-04 (통권 제324호)	대전환 시대의 과학기술혁신 정책 이슈	변순천, 구본진, 김성진, 김진하, 김현오, 박노언, 배용국, 오서연, 이원홍, 신동평, 정선민, 최창택 (KISTEP)
2022-03 (통권 제323호)	2030 국가온실가스감축목표에 기여할 10대 미래유망기술	이동기 (KISTEP)
2022-02 (통권 제322호)	국내외 환경변화에 따른 과학기술혁신 총괄기능 강화 방향	이정재 (KISTEP)
2022-01 (통권 제321호)	KISTEP Think 2022, 15대 과학기술혁신정책 아젠다	손병호·손석호 (KISTEP)
2021-20 (통권 제320호)	국가 R&D 기술사업화 핵심 영향요인 분석 및 시사점	황인영(KISTEP)
2021-19 (통권 제319호)	탄소중립 달성을 위한 정부 연구개발 정책 및 투자방향	한응용, 전은진, 손영주(KISTEP)
2021-18 (통권 제318호)	유럽연합의 임무주도형 혁신정책의 특징과 시사점	강진원(KISTEP)
2021-17 (통권 제317호)	대학 기술지주회사제도 개선방안	정동덕(KISTEP)
2021-16 (통권 제316호)	규제자유특구 운영 현황 분석 및 제도 개선 제언	이재훈, 박일주(KISTEP)
2021-15 (통권 제315호)	성장동력 현황 분석 및 정책 제언 -D.N.A와 BIG3-	김진용 외(KISTEP)
2021-14 (통권 제314호)	R&D인프라의 실증 데이터 활용을 위한 주요 이슈와 정책제언	유형정, 김선재, 권정은, 이승필(KISTEP)
2021-13 (통권 제313호)	신입과학기술인의 역량 인식 차이 분석	김지홍, 주혜정(KISTEP)
2021-12 (통권 제312호)	공공 R&D 투자의 사회·경제적 파급효과 분석	엄익천(KISTEP), 황원식(전북대학교)
2021-11 (통권 제311호)	지역대학 위기와 새 정부 고등교육정책 거버넌스 방향	오세홍, 안지혜, 유지은 (KISTEP)
2021-10 (통권 제310호)	기술개발지원 지역 R&D의 효율성 개선 방향 제언	박석종(KISTEP), 염성찬
2021-09 (통권 제309호)	바이오헬스 산업 성장가속화를 위한 정부R&D의 역할 및 예산배분 전략	홍미영, 김주원(KISTEP)
2021-08 (통권 제308호)	2045년을 향한 미래사회 전망과 핵심이슈 심층분석	정의진 외(KISTEP)



## 필자 소개

### ▶ 구본진

- 한국과학기술기획평가원 미래성장전략센터 부연구위원
- T. 043-750-2472 / bonkoo@kistep.re.kr

---

## KISTEP ISSUE PAPER 2022-05 (통권 제325호)

---

|| 발행일 || 2022년 5월 18일

|| 발행처 || 한국과학기술기획평가원 전략기획센터  
충청북도 음성군 맹동면 원종로 1339  
T. 043-750-2300 / F. 043-750-2680  
<http://www.kistep.re.kr>

|| 인쇄처 || 주식회사 동진문화사(T. 02-2269-4783)

---