

디지털 전환의 미래사회 위험이슈 및 정책적 대응 방향: 인공지능을 중심으로

구본진**

<목 차>

1. 서론
2. 문헌연구
3. 디지털 전환의 미래사회 위험이슈 분석
4. 결과 종합 및 정책적 시사점
5. 결론

국문초록 : 디지털 전환(digital transformation)은 디깃화(digitisation)와 디지털화(digitalisation)의 경제적 및 사회적 효과를 의미한다. 디지털 전환은 경제/사회 발전 및 삶의 편의성을 향상시키는 유용한 도구로 작용하지만 부정적 영향(개인정보 오남용, 윤리문제 야기, 사회적 격차 심화 등)을 미칠 수도 있는 양면성을 보유하고 있다. 한편 정부는 산업 경쟁력 및 기술 패권 확보 등을 위하여 디지털 전환 촉진정책은 적극적으로 추진하고 있는 반면, 디지털 전환 관련 위험이슈에 대한 이해와 이를 방지할 수 있는 정책 추진에는 상대적으로 소극적인 상황이다. 이에 본 연구는 디지털 전환이 초래할 수 있는 미래사회 위험이슈를 Embedded Topic Modeling 방법론 기반의 언론기사 빅데이터 정량분석으로 체계화 및 구체화하고, 정책적 대응 방향을 제시하였다. 이를 위하여 먼저 주요국의 디지털 전환 역기능 세부 이슈들을 규명하였다. 다음으로 디지털 전환의 핵심 기술인 인공지능을 중심으로 주요국과 한국의 디지털 전환 역기능 세부이슈를 구체화하고, 비교분석하였다. 아울러 분석결과들을 종합하여 향후

* 이 논문은 한국과학기술기획평가원 연구비 지원에 의한 논문임 (AO22040, 2022년도 과학기술 혁신정책 핵심이슈 발굴 및 인텔리전스 기능강화 연구)

** 한국과학기술기획평가원 미래성장전략센터 부연구위원(bonkoo@kistep.re.kr)

정부의 디지털 전환 역기능 대응 정책 수립 방향을 제시하였다. 분석 결과에 근거한 정책적 함의는 다음과 같다. 첫째, 디지털 전환의 역기능은 기술 분야에만 한정되어 나타나지 않고 국가안보 및 사회 전반에 영향을 미치기 때문에 정부는 디지털 전환의 순기능 촉진뿐만 아니라 역기능 대응을 위한 정책도 마련해야 한다. 둘째, 디지털 전환의 미래사회 위험 세부이슈들은 국가에 따라 상이하게 나타나므로 정부는 국가적/사회적 맥락을 고려하여 디지털 전환 역기능 대응 정책을 수립해야 한다. 마지막으로 정부는 이해관계자들의 혼선을 최소화할 수 있도록 디지털 전환 역기능 대응 정책의 큰 방향을 설정하고, 실효성 있는 정책 수단을 마련해야 한다.

주제어: 디지털 전환의 역기능, 디지털 전환 위험이슈, 디지털 전환 역기능 대응 정책

A Study on Risk Issues and Policy for Future Society of Digital Transformation: Focusing on Artificial Intelligence

Bonjin Koo

Abstract : Digital transformation refers to the economic and social effects of digitisation and digitalisation. Although digital transformation acts as a useful tool for economic/social development and enhancing the convenience of life, it can have negative effects (misuse of personal information, ethical problems, deepening social gaps, etc.).

The government is actively establishing policies to promote digital transformation to secure competitiveness and technological hegemony, however, understanding of digital transformation-related risk issues and implementing policies to prevent them are relatively slow. Thus, this study systematically identifies risk issues of the future society that can be caused by digital transformation based on quantitative analysis of media articles big data through the Embedded Topic Modeling method. Specifically, first, detailed issues of negative effects of digital transformation in major countries were identified. Then detailed issues of negative effects of artificial intelligence in major countries and Korea were identified. Further, by synthesizing the results, future direction of the government's digital transformation policies for responding the negative effects was proposed. The policy implications are as follows. First, since the negative effects of digital transformation does not only affect technological fields but also affect the overall society, such as national security, social issues, and fairness issues. Therefore, the government should not only promote the positive functions of digital transformation, but also prepare policies to counter the negative functions of digital transformation. Second, the detailed issues of future social risks of digital transformation appear differently depending on contexts, so the government should establish a policy to respond to the negative effects of digital transformation in

consideration of the national and social context. Third, the government should set a major direction for responding negative effects of digital transformation to minimize confusion among stakeholders, and prepare effective policy measures.

Keywords: Negative effects of digital transformation, Risk issues of digital transformation, Response policies for problems of digital transformation

1. 서론

디지털 전환은 아날로그 데이터와 프로세스를 기계가 읽을 수 있는 형식으로 변환하는 디지화(digitisation)와 디지털 기술 및 데이터, 이들의 상호연결을 활용하는 것을 의미하며 이는 새로운 행위 또는 기존 행위에 변화를 유도하는 디지털화(digitalisation)의 경제적/사회적 효과로 정의된다(OECD, 2019). 또한 디지털 전환의 구성기술은 인공지능, 블록체인, 빅데이터, 사물인터넷, 클라우드 컴퓨팅, 5G 네트워크 기술이고(OECD, 2019), 이 중 인공지능 기술을 중심으로 디지털 전환은 경제/사회/일상에 빠르게 접목되고 있다¹⁾

디지털 전환은 경제/사회 발전 및 생활 편의성을 향상시키는 유용한 도구 작용하지만 그와 동시에 부정적 영향을 미칠 수도 있는 역기능을 보유하고 있다. 디지털 전환 역기능은 기술적 문제에 국한되지 않고, 민간이 통제하기 어려운 사회적 문제로 확장되어 나타난다. 예를 들어, 최근 특정 국가는 안면 인식 기술을 활용하여 국민, 기자, 유학생 등을 실시간으로 감시하는 시스템을 구축하여 논란이 됐고, 모 글로벌 기업은 이용자 동의 없이 생체정보를 수집 및 저장하여 거액의 합의금을 물게 되었다. 아울러 디지털 전환 역기능은 국가 안보까지 위협할 수 있다. 실제 특정 기업은 국가 통신망에 백도어를 설치하여 국가 기밀, 핵심 기술 등을 절취하여 논란이 됐다. 그러므로 정부는 디지털 전환의 양면성에 주목하고, 디지털 전환 촉진 정책과 함께 역기능에 대응할 수 있는 정책적 방안을 마련할 필요가 있다. 그러나 정부는 산업 경쟁력 및 기술 패권 확보 등을 위하여 디지털 전환 촉진정책은 적극적으로 추진하고 있는 반면, 디지털 전환 관련 위험이슈에 대한 이해와 이를 방지할 수 있는 정책 추진에는 상대적으로 소극적인 경향이 있다. 또한 관련 연구도 초기 단계로 정성적 논의 및 사례 분석을 중심으로만 진행되고 있는 상황이다.

따라서 본 연구는 이와 같은 연구 공백(research gap)을 메우고자 인공지능 기술을 중심으로 디지털 전환이 초래할 수 있는 미래사회 위험이슈를 구체화하였다. 본 연구는 기존의 관련 연구에서 사용한 정성적인 추정, 예측, 사례분석 등에서 한발 더 나아가 정량분석을 통해 미래사회 위험이슈를 체계적으로 분석하였다. 구체적으로는 Embedded Topic Modeling 방법론 기반의 언론기사 빅데이터 분석을 통하여 주요국의 디지털 전환 역기능 이슈들을 구체화 및 유형화하였다. 또한 분석결과 도출된 역기능 이슈들과 디지털 전

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

환 구성기술들과의 연관성을 분석하여 핵심기술(인공지능)을 규명한 뒤 인공지능 역기능을 구체화 및 유형화하였다. 다음으로 동일한 방법론을 적용하여 한국의 디지털 전환 역기능 이슈들을 구체화 및 유형화하였고, 이를 주요국 분석결과와 비교하여 정책적 시사점을 도출하였다. 나아가 분석결과들을 종합하여 향후 정부의 디지털 전환 역기능 대응 정책 수립 방향을 제시하였다.

2. 문헌연구

디지털 전환은 디지털 기술 중심의 전환, 차별화된 변화로 ‘다양한 디지털 기술을 바탕으로 기업의 전략이나 시스템 등을 근본적으로 변화시켜 새로운 가치를 창출하는 것’을 의미한다(장훈, 2017). 최근 많은 국가는 디지털 전환에 주목하여 관련 기술(인공지능 빅데이터, 클라우드, 컴퓨팅, IoT)개발 및 발전에 전력을 쏟고 있다(이명화 & 최용인, 2017).

또한 디지털 전환은 사회 및 산업 전반에 상당한 변화를 일으키고 있다. 임희중 외(2021)에 따르면 디지털 전환은 기존의 조직 프로세스가 디지털 기술로 효율화되고, 기존 고객들에게 가치가 전달되는 형태에 변화를 도모하는 모델혁신과 서비스를 재정의하여 새로운 산업 분야로 영역을 확장하는 전략혁신의 형태로 비즈니스 모델을 변화시킨다고 주장하였다. 동일한 맥락에서 김용진(2018)도 디지털 전환은 스마트 비즈니스 모델 등장을 촉진하고, 제조의 서비스화를 유도한다고 주장하였다. 아울러 디지털 전환은 조직의 운영 프로세스 자동화, 공급사슬 및 프로세스 통합, 공급사슬의 기술 혁신을 촉진한다(Büyüközkan et al., 2018; Wu et al., 2016).

나아가 관련 선행연구를 살펴보면 디지털 전환은 긍정적인 측면과 부정적인 측면이 공존함을 확인할 수 있다. 선행연구에서 언급된 디지털 전환의 긍정적 측면은 다음과 같다. 독일 공학한림원은 디지털 전환이 기업 운영에 있어 민첩성(주변 환경에 신속하게 대응, 근로자의 업무상 유연성 제공), 평생학습(기업의 생산성 향상, 직원의 업무 능력 및 고용 가능성 보장), 혁신중심의 공동 의사결정(근로자의 의사결정 참여, 기업의 요구 사항 조정 등)을 촉진할 수 있다고 발표하였다(장훈, 2017). 김승현(2020)은 디지털 전환이 기존 산업생태계가 디지털 환경기반으로 전환되어 발전하는 디지털 혁신생태계 발달을 유인할 수 있다고 주장하였다. 한재필(2018)은 디지털 전환은 서비스 산업의 영역 및 시장을 확대함으로써 침체된 서비스업의 도약을 촉진할 수 있다고 설명하였고, 신동수 외(2021)는 디지털 전환이 유통시간 단축과 재고 부담 완화를 통해 생산성을 향상할 수

있으며 새로운 고용기회 창출과 노동시장 효율성 제고, 물류시스템 최적화에 기여할 수 있다고 주장하였다.

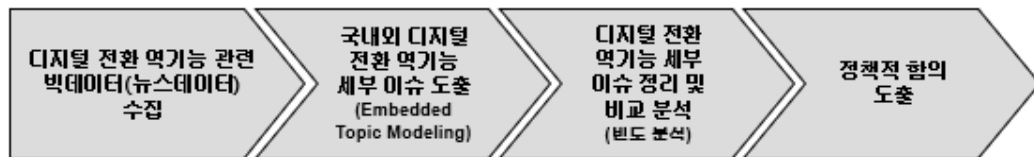
반면 선행연구에서 제시된 디지털 전환의 부정적 측면은 다음과 같다. 첫째, 디지털 전환은 사회 및 경제적 격차를 심화시킬 수 있다. 이명화 & 최용인 (2017)은 디지털 전환 관련 기술개발의 핵심인 민간 연구개발이 일부 선진국에 편중되어 있고, ICT 기술의 산업별 활용도가 균일하지 않으며 지역 및 개인 간 ICT 사용 격차가 크기 때문에 디지털 격차가 발생하게 되고 이는 결국 사회/경제적 격차를 심화할 위험이 존재한다고 주장하였다. 둘째, 디지털 전환은 일자리 수를 감소시킬 수도 있다. 디지털 전환은 자동화 영역을 확장하여 저숙련 및 중숙련 직종까지 기술이 노동을 대체하는 현상을 발생시켜 상당수의 근로자들이 일자리를 위협받을 수 있다(신동수 외, 2021; 여영준, 2020). 마지막으로 디지털 전환은 산업 간 경계를 모호하게 만드는 데 반해 관련 법/규제는 이러한 트렌드를 따라가지 못하여 이해충돌 및 갈등을 촉발할 수 있고, 법/제도 체계가 충돌할 위험도 존재한다(김승현, 2020).

디지털 전환 역기능 자체에 초점을 맞춘 학술 연구는 정성적 논의 및 사례분석 중심의 초기 단계에 머물러있다. 신영진 (2020)은 인공지능 기술을 중심으로 보안위협, 개인정보 침해사고, 윤리문제 등을 지적하고, 주요국의 역기능 방지 및 개인정보보호 사례를 분석하여 지능형 정부 구현 차원에서 필요한 법제도, 기술 등을 제안하였다. 윤상오 (2018)는 중앙정부, 지자체, 공공기관에서 활용하는 인공지능 기반 챗봇 서비스에 주목하여 이것의 역기능 및 쟁점을 연구하였고, 사례 중심의 분석을 수행하였다. 임규건 & 안재익 (2020)은 인터넷 보안에 포커스하여 신뢰 네트워크 기술인 Bright Internet을 통한 사이버 역기능 해소 가능성을 연구하여 관련 역기능을 세분화하였고, 이에 대한 인식수준을 조사하였다. 박승창 (2021)은 증강휴먼 인공지능 기술에 국한하여 이것의 역기능에 대한 사례연구를 수행하였다. 고태수 외 (2019)는 인공지능의 deep learning 기술의 한계를 중심으로 여기서 파생될 수 있는 법적 쟁점을 구체화하였고, 심우민 (2021)은 디지털 전환이 초래할 수 있는 사회갈등을 정성적으로 논의하였고, 이를 완화하기 위한 입법 실무절차 구성방안을 제시하였다.

3. 디지털 전환의 미래사회 위험이슈 분석

문헌연구에서 살펴본 바와 같이, 디지털 전환 역기능 연구는 비교적 최근에 시작되었고, 분석도 아직 정성적인 논의 및 사례를 중심으로 이루어지고 있음을 확인하였다. 이에 본 연구는 한 걸음 나아가 정량분석을 통해 디지털 전환 역기능을 구체화하고, 관련 정책 수립 방향을 제시함으로써 기존 선행연구와 차별화하고, 객관성을 확보하고자 하였다.

3.1 분석 개요



<그림 1> 연구 프레임워크

본 장에서는 정량분석을 통하여 주요국(미국, 중국, 유럽, 일본)과 국내에서 나타나고 있는 디지털 전환의 미래사회 위험이슈를 체계적으로 규명하였다. 정량분석 방법론은 Topic Modeling 방법론 중 Embedded Topic Modeling을 활용하였다. 본 연구가 Embedded Topic Modeling을 주요 방법론으로 선택한 이유는 해당 방법론이 객관성 확보를 위해 필요한 대규모 비정형 데이터의 정량분석에 적합한 방법론이기 때문이다. 객관성 확보를 위해서는 대규모 데이터를 특정 편향성 없이 이해 및 분석하는 작업이 선행되어야 한다. 그러나 데이터가 비정형이고, 규모가 일정 수준을 초과하게 되면 해당 작업을 사람이 수행할 수 없고, 컴퓨터가 수행해야 한다. 이때 적합한 방법론이 Topic Modeling이며 가장 개선된 Topic Modeling 방법 중 하나가 Embedded Topic Modeling이다.

구체적으로 Topic Modeling은 비정형 텍스트 데이터에 대한 구조적 분석을 통해 문서가 내포하고 있는 토픽들의 비중과 각 토픽을 구성하는 단어들의 분포를 제공함으로써 대용량 텍스트 데이터에 잠재된 토픽 정보를 인간이 해석 가능한 단어 조합의 형태로 제공한다. 토픽 모델 중 가장 활용도가 높은 Latent Dirichlet Allocation (LDA) 모델은 각 문서를 토픽의 확률적 혼합체(mixture of topic)로 간주하는 동시에 각 토픽을 단어의 분포로 간주하고 있다. LDA기반 토픽 모델링은 비지도 학습법으로 경험적 접근 없이도 관련 확률을 쉽게 계산하여 토픽 비중, 분포 등을 쉽게 계산할 수 있다는 장점이 있는 반

면, 단어 사전이 큰 경우(large vocabulary size) 토픽 모델 결과에 대한 해석이 어려워지는 문제가 발생하는 치명적인 단점이 있다. 이러한 단점 극복을 위해 고안된 방법론이 Embedded Topic Modeling이다. 이는 신경망 언어 모델(neural language model)을 활용하여 단어의 의미를 다차원 공간에 벡터화하는 단어 임베딩(word embedding) 방식을 활용하는 방식으로 기존 토픽모델링의 한계점을 크게 개선하였다 (Dieng et al., 2020). 기존 LDA 방법과는 달리 임베딩 기반의 토픽 모델(embedded topic model)의 경우 단어를 one-hot-vector와 같은 분절(discrete)된 저밀도 벡터(sparse vector)로 표현하는 것이 아니라 최신 딥러닝 방식이 채택하는 연속형(continuous) 고밀도 벡터(dense vector)로 표현하여 정보를 압축하는 임베딩 방법 (Mikolov et al., 2013)을 사용하고 있다. 이에 따라 Embedded Topic Model은 사전 학습된 워드 임베딩을 통해 단어와 토픽을 임베딩 벡터로 활용하여 대용량 데이터의 토픽모델링의 성능을 향상시켰다 (Dieng et al., 2020).

이에 본 연구는 Embedded Topic Modeling을 주요 방법론으로 선택하였다. 이를 활용하여 본 연구는 먼저 주요국(미국, 중국, 유럽, 일본)의 디지털 전환 미래사회 위험이슈들을 분석하였고, 다음으로 디지털 전환의 핵심기술인 인공지능에 초점을 두어 주요국과 국내 인공지능 기술의 미래사회 위험이슈를 도출한 뒤 결과를 비교분석하였다.

3.2 주요국 디지털 전환의 미래사회 위험이슈 분석

본 절에서는 Embedded Topic Modeling을 활용하여 주요국(미국, 중국, 유럽, 일본) 언론기사에서 다루어진 디지털 전환의 역기능 세부 위험 이슈들을 도출하였다. 데이터는 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 영문 언론기사/사설(뉴스 데이터) 2,175,893개를 수집 및 활용하였다. 구체적으로 뉴스데이터는 Newscatcher 데이터 API 서비스를 활용하여 디지털 전환의 주요기술 키워드 (artificial intelligence, big data, 5G networks, IoT, Blockchain)과 리스크 키워드인 (risk, problem, dark side)를 조합하여 검색 쿼리를 작성하여 초기 뉴스데이터를 수집하였다. 초기에 수집된 뉴스데이터는 검색 쿼리 간의 중복데이터 존재, 동일내용의 서로 다른 URL로 인한 중복데이터 존재, 검색 쿼리와 관련성 낮은 데이터 존재 등 데이터의 노이즈가 상당수 존재하였다. 이를 제거하기 위해 뉴스기사 URL 기준 중복제거, 뉴스 제목 및 본문 첫 문장을 기준으로 중복기사 제거, API 업체에서 제공하는 관련도 점수 (relevancy score)를 기준으로 상위 20%를 제외한 나머지 모든 데이터를 제거, 본문 길

이 5문장 이내의 뉴스기사 제거 등을 통해 토픽모델링 목적에 부적합한 데이터 노이즈를 최소화하였다. 이와 같은 전처리 작업을 거쳐 349,076개 분량의 언론기사 및 사설을 실제 Embedded Topic Modeling 분석에 사용하였다²⁾³⁾. 이를 통하여 주요국에서 논란이 되고 있는 디지털 전환의 세부 위험 이슈들을 정량적으로 규명하였다.

먼저 전체 데이터를 대상으로 Embedded Topic Modeling 분석을 수행하였다. 분석 결과에 대하여 연관주제별 군집화 및 연관성 높은 언론 기사 매칭 작업을 통해 총 9개 세부 이슈들을 도출하여 이를 빈도순으로 정리⁴⁾하였다. 정리 결과 무기용 인공지능 문제와 인공지능의 편향성(bias) 문제가 높은 빈도를 차지하였고, 다음으로는 딥페이크 문제, IoT 기기의 보안 문제, 그리고 인공지능 윤리 문제가 사회적으로 많이 회자됐음을 확인하였다.

나아가 도출된 9개 역기능 이슈들과 가장 밀접하게 연관된 디지털 전환 구성 기술을 규명하기 위하여 유형별 연관 기술의 빈도분석을 수행하였다. 분석 결과 기존의 예측 및 논의의 결과와 동일하게 인공지능 기술이 디지털 전환의 미래사회 위험이슈와 가장 밀접한 관계가 있음을 확인하였다. 이는 인공지능 기술이 디지털 전환에 있어 핵심기술로 작용하기에 타 구성기술대비 역기능 연관성도 가장 높았던 것에 기인한 결과로 추정된다.

-
- 2) 본 연구에 적용된 Embdded Topic Modeling 하이퍼파라미터는 관련 논문에서 소개된 기본값을 기초로 하되, 학습 반복횟수(epoch)를 기존 300에서 1000으로 상향조정하고, 학습률(learning rate)을 0.001에서 0.005 사이 구간에서 실험하여 주제 일관성(topic coherence)이 최대가 되는 학습률(0.005)을 채택하여 사용하였다. 마지막으로 토픽의 수를 설정하기 위해서 밀도기반 클러스터링(DBSCAN) 방식을 토픽 임베딩 벡터와 함께 활용하여 토픽 쌍의 임베딩 유사도가 0.9 이상인 토픽끼리 묶이도록 설정함으로써 최적의 클러스터 개수가 하이퍼파라미터 토픽 수로 설정되도록 하였다.
 - 3) Embedded Topic Modeling의 Topic 수는 분석 결과에서 도출된 키워드들을 검토하고, 분석 목적에 부합하는 키워드들의 조합들 및 관련 뉴스데이터를 확인하여 선택하는 방식으로 진행하였다. 구체적으로 연관주제별 군집화 및 연관성 높은 언론 기사 내용 매칭을 통한 해석 작업을 거쳐 Topic 수를 선택하였으며 이는 토픽 모델링에서 일반적으로 사용하는 방법이다.
 - 4) 디지털 전환 역기능과 관련하여 사회적으로 자주 언급된 이슈들 대다수는 명확한 정책적 해결 방안 또는 사회적 합의가 이루어지지 못한 이슈들이기 때문에 빈도수에 비례하여 사회적으로 문제가 되고 있고, 정책 대응이 시급한 이슈들로 볼 수 있다고 판단하였다.

<표 1> Embedded Topic Modeling 분석 결과(역기능 세부이슈별 기술 연관성 분석 포함)

역기능 유형	세부 내용	빈도수	5G networks	IoT	AI	big data	block chain
무기용 인공지능	AI for weapon	190	1	10	175	2	2
인공지능의 편향성	biased AI profiling	127		10	107	7	3
딥페이크	Facebook banning deepfakes	71			66	0	5
보안	IoT vulnerabilities	63		60	2	0	1
인공지능 윤리	Ethics in AI systems	62			61	1	0
프라이버시	Data privacy risk	41			17	12	12
암호화폐 해킹	crypto currency security (broken by hackers)	33			3	0	30
비트코인 채굴의 환경파괴	Bitcoin mining and carbon emission problem	24			3	0	21
인공지능의 사회적 차별심화	AI discrimination in credit lending	11			11	0	0

다음으로 인공지능과 관련한 디지털 전환의 역기능 세부 위험이슈들을 구체화하였다. 구체적으로는 인공지능 기술에 관련된 주요국의 언론기사에 한정하여 Embedded Topic Modeling을 수행하였다. 분석 결과 생체 데이터 활용 문제가 가장 큰 비중을 차지하였고, 다음으로 안보/보안 위험, 프라이버시 침해, 편향성 이슈가 디지털 전환 위험 세부이슈의 비중을 차지하였다.

<표 2> 주요국 인공지능 미래사회 위험 세부이슈 도출 결과

세부 위험이슈	관련 키워드	빈도수
생체 데이터 사용 문제에 대한 EU의 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

3.3 국내 디지털 전환의 미래사회 위험이슈 분석(인공지능 기술을 중심으로)

국내 디지털 전환의 역기능 세부이슈 분석은 핵심기술인 인공지능을 중심으로 2개의 분석을 수행하였다. 첫째는 인공지능 문제를 다룬 국내 언론기사들을 대상으로한 Embedded Topic Modeling 분석이고, 다른 하나는 국내에서 사회적으로 가장 크게 논란이 됐던 디지털 전환(인공지능) 역기능 사례인 ‘이루다’ 케이스에 대한 Embedded Topic Modeling 분석이다.

먼저 국내 인공지능 역기능 분석을 위하여 2020.01.01 ~ 2021.08.23. 기간에 국내 언론에서 보도된 인공지능 기술의 위험/역기능/문제 등을 다룬 언론기사 및 사설 16,562개를 수집 및 활용하였다. 이후 중복제거 등의 전처리 과정을 거쳐 14,419개의 언론기사 및 사설을 활용하여 Embedded Topic Modeling 분석을 통해 역기능 세부 이슈를 도출하였다.

세부이슈 도출 결과 한국은 인공지능 역기능 관련 세부이슈로 인공지능 윤리에 대한 이슈가 가장 높은 비중을 차지하였다. 다음으로는 국내에서 인공지능 역기능과 관련하여 가장 크게 이슈가 됐던 인공지능 기반 챗봇 서비스 ‘이루다’ 사건과 관련한 세부 이슈들(인공지능 윤리문제 소송, 개인정보유출 논란, 혐오 논란)이 높은 비중을 차지하였다. 또한 다른 국가들에서 발생했던 인공지능 역기능 이슈들(페이스북, 구글 등)과 딥페이크 범죄 이슈들이 낮은 비중으로 도출된 것을 확인하였다.

<표 3> 한국 인공지능 미래사회 위험 세부이슈 도출 결과

세부 위험이슈	관련 키워드	빈도수
인공지능 윤리 문제 일반	인공지능, 기술, 데이터, 윤리, 인식	182
(이루다) 윤리문제 소송	인공지능, 소송, 개인, 정보, 변호사	110
(이루다) 개인정보유출 논란	정보, 보호, 이용자, 점검, 서비스, 처리, 데이터, 카카오톡, db, 활용	91
(이루다) 인공지능 혐오 논란	발언, 대화, 혐오, 차별, 이용자, 증오, 논란, 표현	60
페이스북 유해 필터 알고리즘 논란	페이스북, 콘텐츠, 삭제, 유해, 발언, 흑인, 혐오, 폭력, 인종, 증오	31
인공지능 윤리 논쟁 - 기술 vs 규제	윤리, 인공지능, 사회, 차별, 지적, 제도, 논란	22
딥페이크 범죄 위험성 증가	영상, 얼굴, 합성, 범죄, 악용, 가짜, 이미지, 피해	22
구글 AI 윤리 연구자 해고 논란	구글, 논문, 윤리, 연구, 직원, 부사장, 사직	16
알고리즘 편향성 논란	뉴스, 검색어, 언론, 사용, 알고리즘, 중재법, 정보	15

다음으로 보다 심층적인 분석을 위하여 한국 사회에서 인공지능 기술의 역기능으로 가장 크게 논란이 됐던 ‘이루다’ 케이스를 분석하여 세부 위험 이슈들을 도출하였다. ‘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 40만 명을 확보할 정도로 매우 빠르게 성장했었다. 그러나 운영 과정에서 개발사의 의도와 달리 개인정보 침해, 상식적 답변 오류, 특정 성별 혐오, 외설적 목적 사용 등의 문제 발생하여 사회적으로 큰 논란이 됐고, 지탄의 대상이 됐다. 결국 각종 논란으로 ‘21년 1월 11일 서비스 중단하였고, ‘21년 4월 개인정보보호위원회가 개인정보보호법 위반으로 ‘이루다’ 개발사 스캐터랩에 총 1억330만원의 과징금 및 과태료 부과하였다.

‘이루다’ 케이스도 Embedded Topic Modeling 방법론을 적용하여 분석을 수행하였고, 데이터는 이루다 사건을 다룬 국내 언론 기사 및 사설을 수집(기간: ‘20.11.01~‘21.4.30⁵⁾., 1,329개)하여 분석에 활용(전처리 후 1,100개 활용)하였다. 분석 결과 총 22개 세부주제들이

5) 본 데이터 수집 기간은 국내 언론에서 ‘이루다’를 다룬 기간(도입부터 폐지까지)임

도출됐고, 연관주제별 군집화 및 연관성 높은 언론 기사 내용 매칭을 통한 해석 작업을 거쳐 최종적으로 11개 세부 위험이슈들을 도출하였다. 도출된 세부 위험이슈들을 빈도순으로 정리한 결과 인공지능 윤리문제가 가장 높은 비중을 차지하였고, 다음으로 혐오/성희롱/차별 문제가 높은 빈도를 기록하였다. 또한 개인정보 이용 문제, 인권침해 문제 등도 주요 세부 위험이슈 유형으로 비중을 차지했다.

<표 4> ‘이루다’ 관련 디지털 전환의 역기능 위험 세부주제 도출 결과

세부 위험이슈	관련 키워드	빈도수
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

4. 결과 종합 및 정책적 시사점

앞장에서는 디지털 전환의 미래사회 위험이슈를 정량적으로 분석하여 세부 위험이슈들을 도출하였다. 가장 먼저 주요국의 디지털 전환 역기능 세부 이슈들을 분석한 결과 다양한 세부 위험이슈들(안전 및 보안 문제, 사회 불평등 심화 문제, 개인정보보호 침해 문제 등)이 도출됐고 이중 대다수는 인공지능의 역기능과 관련한 위험이슈들이었다. 세부이슈별 기술 연관성 분석 결과도 해당 결과를 지지하였다. 이를 통해 확인할 수 있었던 점은 첫째, 디지털 전환의 미래사회 위험이슈는 다른 디지털 전환의 구성기술들(blockchain, 5G network, cloud 등)보다 인공지능 기술이 중심이 되어 논의되고 있다는 점이다. 해당 결과는 인공지능 기술이 디지털 전환의 핵심 기반기술(Enabling Technology)

역할을 수행하고 있고(안준모, 2021), 사회를 규율하는 구조적 힘을 보유하면서 해당 기술의 혁신이 사회변동의 핵심 요인으로 작용하는(엄효진, 이명진, 2020) 특성이 분석 결과에 반영된 것이라 생각한다. 둘째, (인공지능 기술 중심의) 디지털 전환 미래사회 위험 이슈는 기술 분야에만 한정되어 나타나지 않았다. 해당 위험 이슈들은 국가 안보 및 보안, 사회 문제, 공정성 문제 등 사회 전반에 영향을 미치고 있었다. 또한 도출된 위험 이슈들은 명확한 솔루션이 존재하지 않고, 이해관계자들이 복잡하게 얽혀있는 이슈들이기 때문에 민간에서 쉽게 해결하기 어려운 특성을 갖고 있다. 그러므로 해당 결과는 정부가 디지털 전환 역기능 문제를 단순한 기술적 문제로 인식하여 대응하기보다는 사회적·국가적 문제로 이해하고, 이에 걸맞은 적극적인 정책적 대응이 필요한 영역임을 시사한다.

다음으로 디지털 전환의 핵심기술인 인공지능 기술에 초점을 맞추어 규명한 주요국과 한국의 인공지능 기술 중심의 디지털 전환 세부 위험이슈들의 비교 분석결과는 다음과 같다. 첫째, 주요국은 한국에 비하여 인공지능 기술 중심의 디지털 전환 세부 위험이슈들의 범주가 다양하게 도출되었다. 구체적으로 주요국은 생체 데이터 사용 문제, 사이버 안보 및 위협 이슈, 인종차별 및 편향성, 프라이버시, 허위정보 및 오보 문제 등 매우 다양한 문제들이 역기능으로 논의되고 있었던 반면, 한국은 인공지능 윤리 문제, 개인정보 유출, 인공지능 혐오 등 상대적으로 한정된 영역의 문제들이 인공지능 역기능으로 논의되고 있음을 확인하였다. 둘째, 도출된 인공지능 기술 중심의 디지털 전환 세부 위험이슈들의 특성이 국내 대비 주요국은 기술적 특성을 넘어 사회적, 국가적 특성을 갖는 이슈들이었다. 국내는 기술 자체의 문제점을 지적한 이슈들이 대다수였다면 주요국은 기술에 내재된 문제점이 사회 전반 나아가 국가 안보 및 보안에 어떠한 영향을 미칠지에 대한 이슈들이 다수를 차지함을 확인하였다.

위와 같은 결과가 도출된 것은 그동안 한국이 주요국 대비 디지털 전환(인공지능) 역기능에 대한 사회적 관심이 부족했고, 관련 논의가 사회적으로 다루어지기 시작한 시점도 상대적으로 늦었던 것에 기인했을 가능성이 있다. 본 연구의 분석 기간(2020.01.01 ~ 2021.08.23.)을 기준으로 한국은 아직 사회적으로 크게 이슈가 된 역기능 케이스는 단일 사례('이루다')였고, 해당 사례에서 가장 강조된 키워드들이 '윤리', '개인정보유출', '혐오'였기에 해당 키워드 외 다른 세부 이슈들은 강조되지 못했을 가능성이 있다. 반면 주요국은 인공지능 기술 중심의 다양한 디지털 전환 역기능 사례들과 이것이 초래할 수 있는 사회적 문제들에 대한 논의가 오랜 기간 축적됐고, 해당 역기능들이 실제 사회 문제로 발현된 케이스도 상당수 존재하였다. 이에 따라 보다 구체화된 세부 이슈들이 도출됐을 가능성이 존재한다고 생각한다.

상기 분석 결과를 통해 다음과 같은 정책적 시사점을 도출하였다. 첫째, 정부는 디지털 전환의 순기능뿐만 아니라 역기능에도 주의를 기울여야 한다. 현재 인공지능 기술을 중심으로 디지털 전환의 다양한 역기능들이 사회적으로 문제화되고 있고, 일부 유형(개인정보, 프라이버시 침해 등)의 역기능은 이미 국민의 일상에 가까워졌다. 따라서 정부는 인공지능 기술을 중심으로 우리 사회에 나타나고 있는 그리고 곧 문제가 될 역기능들을 파악하고, 선제적으로 대응할 수 있는 정책을 모색해야 한다.

둘째, 정부는 국가적/사회적 맥락을 고려하여 디지털 전환 역기능 대응 정책을 수립해야 한다. 본 연구 결과에서 알 수 있듯이 국가별로 문제로 나타나고 있는 역기능 세부 유형은 상이하다. 이는 디지털 전환 역기능이 국가적/사회적 맥락에 따라 다르게 나타날 수 있음을 시사한다. 예를 들어, 주요국에서 높은 비중을 보인 인공지능 역기능 세부 유형인 데이터 편향성에 따른 인종차별은 한국에서 동일하게 발현될 가능성이 작을 것이다. 이처럼 정부는 디지털 전환 역기능 대응 정책 수립 시, 국가적/사회적인 상황을 고려하여 대응 영역을 예측 및 선정해야 할 것이다.

셋째, 정부는 실효성 있는 디지털 전환 역기능 대응 정책수단을 고민해야 한다. 현재 한국 정부는 '19년도 말 「인공지능 국가전략」을 통해 역기능 방지 및 AI 윤리 마련 계획을 제시하였고, '20년 AI 활용 가이드라인 발표 등을 통해 관련 정책을 수립 및 추진하고 있다. 그러나 해당 정책들은 인공지능 개발자 또는 활용 업체들에게 권고사항 정도로만 작용하는 가이드라인 수준에 머물러있다. 물론 해당 정책들은 디지털 전환(인공지능) 역기능 대응 정책의 첫걸음으로써 의의가 있겠지만 시장에서 인지하는 실효성은 매우 낮은 상황이다. 따라서 정부는 인공지능 개발자/서비스 제공업체/이용자들의 권리를 확실하게 보호하면서 관련 산업발전은 저해하지 않는 실효성 높은 정책적 수단을 고민할 필요가 있다.

결과적으로 정부는 디지털 전환 역기능이 단순 기술적 문제가 아닌 사회 그리고 국가 차원의 심각한 문제를 초래할 수 있는 문제임을 인식해야 한다. 이러한 인식의 변화와 함께 국가 차원의 디지털 전환 역기능 대응 정책 방향을 설정해야 한다. 유럽연합과 같은 강력한 규제 중심의 대응, 미국과 같은 자율성 기반의 행위자 중심 규제를 통한 대응 또는 중간 수준의 규제 정책 등 거시적인 정책 방향을 설정해야 한다. 마지막으로 거시적인 정책 방향을 토대로 우리 사회에서 발현될 가능성이 큰 세부영역들을 예측하고 영역별 세부 대응 방안을 마련해야 한다.

5. 결론

본 연구는 그동안 학술적/실무적으로 강조되어온 디지털 전환의 순기능이 아닌 이것이 수반할 수 있는 그리고 수반하고 있는 역기능에 초점을 맞추었다. 현재까지 디지털 전환의 역기능에 관한 연구는 정성적인 차원에서 해당 문제가 어떠한 형태로 나타날지에 대한 논의 및 추정이 주를 이룬 데 반하여, 본 연구는 여기서 한발 더 나아가 데이터 기반의 정량분석을 통하여 해당 문제를 분석함으로써 실제 주요국 및 우리 사회에서 문제화되고 있는 그리고 논의되고 있는 디지털 전환 미래사회 위험의 세부이슈들을 규명하였다. 이를 통하여 상대적으로 연구가 부족했던 디지털 전환의 미래사회 위험이슈들을 구체화함으로써 향후 정부가 정책적 대응을 마련해야 할 세부영역들을 제시하였다.

아울러 본 연구는 정량분석 결과를 기반으로 정부의 디지털 전환 역기능에 관한 정책 수립 방향을 제시하였다. 현재 정부는 해당 영역에 있어 다소 소극적인 정책(AI 윤리 가이드라인 개발 및 배포 등)을 추진하고 있다. 이는 이해관계자들에게 강한 규제로 시장/산업발전을 저해하는 것은 최소화하면서도 역기능 방지에 있어 정부가 노력하고 있다는 시그널을 주기 위한 신중한 대응의 결과라 생각된다. 그러나 현재 정부는 디지털 전환 역기능 대응 정책의 큰 방향성에 관한 입장이 부재하고, 세부 대응 영역도 모호한 상황이다. 이로 인해 시장 내 이해관계자들은 향후 정책환경 변화 예측에 어려움을 느끼고 있다. 따라서 정부는 본 연구 결과와 같이 향후 디지털 전환의 미래사회 위험이슈 대응을 위해서는 인공지능 기술로부터 파생될 수 있는 역기능에 주목하고, 국가/사회적 맥락을 확인하여 위험이슈 대응의 우선순위를 설정해야 할 것이다.

본 연구는 상기와 같은 학술적/정책적 기여가 있는 반면, 다음과 같은 한계점도 갖고 있다. 첫째, 디지털 전환의 세부 위험이슈들 중 현존하지 않는 문제는 상대적으로 덜 반영이 되었을 가능성이 있다. 본 연구는 디지털 전환 역기능과 관련한 대규모의 언론기사들을 수집 및 정량분석하여 디지털 전환의 세부 위험이슈들을 구체화하였기에 현재 사회적으로 논의되고 있는 그리고 문제로 나타나고 있는 세부 위험이슈들을 포착할 수 있었다. 그러나 분석 데이터가 언론기사에 한정되어 있기에 현시점의 세부사항들은 확인할 수 있었으나 예측 기반의 현존하지 않는 세부 위험이슈들은 상대적으로 덜 반영된 측면이 존재한다.

둘째, 산업/상품/서비스에 특화되어 나타나는 디지털 전환 역기능 세부 이슈들은 포착하지 못했을 가능성이 존재한다. 본 연구는 데이터 수집 시, 디지털 전환의 역기능을 다룬

기사/사설들을 수집하기 위하여 해당 이슈를 전반적으로 커버할 수 있는 키워드들을 사용하였다. 이로 인해 본 연구에서 활용한 키워드로 나타나지 않은 아주 세부적인 산업/상품/서비스 특화 역기능 이슈들이 데이터로 반영되지 못했을 가능성이 있다.

따라서 향후에는 데이터의 범주를 언론에서 논문, 연구보고서 등의 다양한 매체로 확장하고, 데이터 수집 키워드를 보다 확장하여 디지털 전환의 미래사회 위험이슈를 구체화할 필요가 있다. 이러한 연구 확장을 통하여 예측 기반의 현존하지 않는 세부 위험이슈들에 대한 규명과 산업/상품/서비스에 특화되어 나타나는 미래사회 위험 이슈들에 대한 확인이 가능할 것으로 사료된다.

참고문헌

(1) 국내문헌

- 고학수·정해빈·박도현. (2019), “인공지능과 차별”, 『저스티스』, 제171호, pp. 199-277.
- 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, 제25권 제1호, pp. 20-24.
- 박승창 (2021), “증강휴먼 AI 기술의 역기능과 예방대책에 관한 연구”, 대한전자공학회 학술대회, pp. 1595-1598.
- 신동수·이규환·이재진·주연희 (2021), “디지털 전환이 생산성 및 고용에 미치는 영향”, 해외경제포커스 제 2021-22호, 한국은행 국제경제리뷰
- 신영진 (2020), “안전한 지능형 정부 구현을 위한 개선방안 연구-인공지능기술 적용에 따른 역기능 방지 및 개인정보보호를 중심으로”, 『한국범죄정보연구』, 제6권, 제2호, pp. 135-162.
- 심우민. (2021), “디지털 전환과 사회갈등-입법학적 분석과 대안”, 『유럽헌법연구』, 제37호, pp. 481-522.
- 안준모 (2021), “인공지능을 통한 행정의 고도화: 기회와 도전”, 『한국행정연구』, 제30권, 제2호, pp. 1-33.
- 엄효진·이명진 (2020), “인공지능(AI) 기반 지능정보사회 시대의 노동시장 변화: 경제사회학적 접근을 중심으로”, 『정보사회와 미디어』, 제21권, 제2호, pp. 1-20.
- 여영준 (2020), “포스트 코로나 시대 기술변화와 혁신정책 방향성 재정립: 창조적 학습사회 전환을 중심으로”, 『기술혁신연구』, 제28권, 제4호, pp. 153-165.
- 윤상오. (2018). “인공지능 기반 공공서비스의 주요 쟁점에 관한 연구: 챗봇 (ChatBot) 서비스를 중심으로”, 『한국공공관리학보』, 제32권, 제2호, pp. 83-104.
- 이명화·최용인 (2017), “[OECD] OECD 과학기술산업 지표를 통해 본 디지털 전환 동향과 도전 과제”, 『과학기술정책』, 제27권, 제12호, pp. 16-21.
- 임규건·안재익 (2020), “Bright Internet 신뢰네트워크 도입에 따른 지능정보사회의 사이버 역기능 해소에 대한 인식 분석”, 『Information Systems Review』, 제22권, 제3호, pp. 99-118.
- 임희중·최보름·송지희 (2021), “기업의 디지털 전환 (DT) 경쟁력 분석 모형개발 및 적용: 공기업 10개의 사례를 중심으로”. 『Korea Business Review』, 제25권, 제3호, pp. 61-100.
- 장훈 (2017), “[유럽] 디지털 전환과 노동의 미래”, 『과학기술정책』, 제27권, 제11호, pp. 10-13.
- 한재필 (2018), “디지털 전환 시대의 서비스 혁신”, 국토연구원.

(2) 국외문헌

- Blei, D. M. (2012), “Probabilistic topic models”, *Communications of the ACM*, Vol. 55, No. 4, pp. 77-84.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003), “Latent dirichlet allocation”. *Journal of Machine Learning Research*, 3(Jan), pp. 993-1022.
- Büyüközkan, Gülçin, and Fethullah Göçer (2018), “Digital Supply Chain: Literature review and a proposed framework for future research”, *Computers in Industry*, Vol. 97, pp. 157-177.
- Dieng, A. B., Ruiz, F. J., & Blei, D. M. (2020), “Topic modeling in embedding spaces”, *Transactions of the Association for Computational Linguistics*, Vol. 8, pp. 439-453.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013), “Distributed representations of words and phrases and their compositionality”. *Advances in Neural Information Processing Systems*, Vol. 26, pp. 3111-3119.
- OECD (2019), “Going Digital: Shaping Policies, Improving Lives”
- Wu, L., Yue, X., Jin, A. and Yen, D.C. (2016), “Smart supply chain management: A review and implications for future research”, *The International Journal of Logistics Management*, Vol. 27, No. 2, pp. 395-417.

□ 투고일: 2021.12.27. / 수정일: 2022.02.17. / 게재확정일: 2022.02.27.