

## LDA 기반 뉴스 빅데이터 토픽모델링을 활용한 디지털 관련 이슈 분석과 정책적 함의

김 선 영\*

연구는 디지털 전환사회에서 디지털 접근성과 관련해 논의되고 있는 사회적 이슈를 살펴 우리 실정에 적용할 수 있는 디지털 포용에 관한 시사점 도출을 목적으로 하고 있다. 이를 위해 디지털 전환사회와 사회적 포용을 관련한 문헌을 고찰하고, 뉴스데이터 분석을 바탕으로 디지털 전환사회와 관련해 논의되고 있는 사안을 귀납적 접근을 통해 탐색하였다. 뉴스데이터 분석은 2020년 1월 1일에서 2023년 1월 31일까지를 시간적 범위로 빅카인즈에서 수집된 5,136개 뉴스 기사를 분석의 대상으로 하는 잠재디리클레할당 기반 토픽 분석을 수행하였다.

본 연구는 사회적 취약계층으로 표현되는 특정 집단뿐만 아니라 우리 사회의 모든 구성원을 두루 아우르는 디지털 포용정책의 필요성, 디지털 취약계층의 개념화, 사회적 포용으로부터의 디지털 포용에 대한 개념화의 필요성과 지능정보사회에서의 디지털이 지니는 의미 재고와 정책적 함의를 제시한다.

[주제어: 디지털 포용, 디지털 포용 사회, 디지털 격차, 사회적 배제, LDA기반 토픽분석]

### I. 들어가며

2020년 전 세계를 덮은 코로나-19 바이러스는 우리 인간의 삶이 급격히 디지털화되게 하였다. 제4차 산업혁명으로 인해 도래하고 있던 지능정보사회는 사회적 관계, 즉 직장이나 학교 생활 등을 유지하게는 했지만, 누구도 예측하지 못한 채 계속된 바이러스 확대에 의한 장기 간의 사회적 거리 두기는 디지털 사용을 급격하게 확대했고, 디지털화되는 사회 그 자체가 사회적 문제로 공론화되기 시작하였다. 디지털 접근의 어려움이 사회적 접근의 곤란으로 이어지는 현상이 빈번해지면서 야기된 사회적 소외 혹은 배제가 사회적 문제로 인식되면서, 사회적 맥락에서 디지털이 지니는 의미에 대한 재고의 필요성이 제기되는 계기로 작동하였다.

물론 팬데믹 이전에도 디지털 격차와 사회적 배제와의 관련성에 관한 연구는 존재했다(Helsper, 2008; Reber & Luginbuhl, 2016). “ATM기기 사용법을 모를 때는 은행 창구에서 돈을 찾아 썼다. 그런데 아들이 ATM기 사용법을 알려줘서 겨우 적응해서 살아가고 있는데 ATM기가 사라졌다”.<sup>1)</sup> 이는 한 기사에 나오는 인터뷰 내용으로 은행 창구와 ATM기가 모바일 은행으

\* 영남대학교 새마을국제개발학과 객원교수, 공주대학교 정책융합전문대학원 강사

로 대체되면서 은행거래가 어려워지고, 모바일에 익숙하지 않은 이들이 은행거래를 할 수 없게 될 수도 있다는 우려에 대한 보도로 디지털 접근과 사회적 배제의 연관성을 보여준 한 사례이다. 사회에서의 디지털의 영역이 확대될수록 디지털 접근의 문제는 불편함의 문제를 넘어 사회적 접근의 문제로 이어지게 되는데, 이러한 점에서 1990년대 이후 인터넷 망의 확대와 함께 정보 접근성 관점에서 이러한 논의가 확대되기 시작되었다. 그러다 코로나19 팬데믹 상황에서의 사회적 거리 두기를 통해 디지털기기를 사용 혹은 구매하지 못하는 디지털 접근 곤란이 단순한 ‘불편의 문제’를 넘어 직장 생활, 교육, 사회활동 등 ‘사회생활의 제한 문제’로 드러나기 시작하였다. 즉, 코로나19로 인한 사회적 거리두기가 사회를 급속하게 디지털화하면서, 디지털 접근성의 차이로 인한 디지털 격차와 사회적 불평등과의 직접적인 관련성을 드러내기 시작했고, 이로 인해 사회참여의 불평등 문제로 인식되게 된 것이(Bailey & Nyabola, 2021).

디지털 격차는 인터넷 사용을 통해 혜택을 받는 개인과 상대적으로 불리한 개인 사이에 존재하는 격차이다(Rogers, 2001). 즉, 인터넷 장치의 사용, 인터넷의 연결 속도, 이를 효과적으로 사용하는 데 필요한 기술력 등까지 포함하는 디지털을 통한 사회적 접근과 활용의 문제를 의미한다. 이에 대해 다양한 측면에서 디지털 격차 해소에 대한 노력이 이뤄지고 있다.

UN을 중심으로 하는 국제기구들은 국가 간의 디지털 격차 해소에 초점을 두고 다양한 사업을 진행하고 있다. 또한, EU 회원국을 비롯한 유럽의 여러 나라는 유럽이사회(EC)가 추진하는 디지털 포용 의제를 배경으로 각기 국가의 실정에 맞는 디지털 포용을 추진하고 있다. 이와 관련해 우리나라 역시 2021년의 디지털포용법(안) 등 다양한 디지털 포용정책이 추진되고 있다. 이는 계속해서 고도화되고 있는 디지털 기술과 디지털 환경에서 사회구성원의 활용역량과 접근 능력의 차이로 인한 경제·사회적 불평등 등과 같은 부정적 현상에 대한 법적·제도적 기본 틀을 바탕으로 하는 대응의 시도로 볼 수 있다. 그러나 매년 보고되고 있는 디지털 정보 격차 실태조사를 통해 일반 국민 전체의 디지털 격차 정도는 매년 조금씩 개선되고 있기는 하지만 여전히 ICT경쟁력에 비해 심각한 정도의 지역 간, 계층 간 디지털 격차를 알 수 있다.

코로나19는 디지털을 기반으로 하는 비대면 서비스를 급속히 확대하였고, 동시에 지능정보 사회화의 진전은 이를 더욱 가속하고 있다. 이러한 상황은 사회적 측면에서 디지털이 지니는 의미를 귀납적 접근을 통해 포괄적으로 살펴볼 필요를 의미한다. 즉, 최근 사회에서의 디지털 관련한 이슈를 객관적으로 살펴 이에 대한 대안을 파악할 필요하다. 이에 이 연구는 사회적 측면에서 디지털과 관련한 연구문헌, 언론자료, 웹사이트 등을 검토하고, 여기에 더해 뉴스데이터 분석을 통해 우리 사회에서의 디지털과 관련한 이슈를 탐색하고자 한다. 뉴스 기사는 빅카인즈에서 2020년 1월 1일에서 2023년 1월 31일까지를 시간적 범위로 디지털과 관련한 기사 391, 368개 중 제목에 디지털이 반드시 포함되는 기사 5,136건을 수집하였다. 이는 기사 내용의 정확성을 높이고자 했기 때문이다. 이를 바탕으로 잠재 디리클레 할당(Latent Dirichlet Allocation, LDA)을 기반으로 하는 토픽모델링 분석을 통해 우리 사회에서의 디지털과 관련한 이슈를 객관화해 살펴보고자 한다. 최근 3년간을 시간적 범위로 정한 이유는 코로나19를 지나는 과정에서의 디지털에 관한 다양한 이슈의 분석은 사회적 측면에서의 디지털의 측면을

1) 〈Financial Review〉 [소셜리뷰] 사라져가는 ATM기, 노인들은 어찌라고.... 2019.03.27. <http://www.financialreview.co.kr/news/articleView.html?idxno=16503>

이해하는 상당한 의미가 있다고 판단하기 때문이다. 이에, 본 연구는 우리 사회에서의 디지털과 관련한 이슈에 대한 이해를 바탕으로 지속할 수 있는 발전의 기틀을 갖춘 포용적 디지털 사회로의 전환을 위한 정책적 함의를 얻는 데 목적이 있다.

## II. 문헌연구

### 1. 사회적 포용 수단으로서의 디지털: 디지털 포용

디지털 사용 정도는 오랫동안 정보격차로 인식되었다. 정보격차의 문제는 대부분이 디지털 장치, 즉 컴퓨터의 사용에 초점이 있었다. 그러나 제4차 산업혁명과 함께 디지털 격차는 인터넷 장치의 사용, 인터넷의 연결 속도, 이를 효과적으로 사용하는 데 필요한 기술력을 포함한다는 점에서 정보격차와는 차이가 있다(Quaintance, 2018). Rogers(2001)는 디지털 격차로 표현하기는 하였지만, 실증연구를 통해 인터넷으로 혜택을 받는 개인과 그렇지 못한 개인 사이에 존재하는 데이터 접근의 차이가 만들어 내는 결과를 통해 정보격차를 보여주었다. 이는 네트워크의 속도, 모바일 연결 정도, 이를 사용할 수 있는 경제적 여건, 디지털 리터러시를 기준으로 개념화되고 있는(WEF, 2020) 디지털 격차와는 다소 개념적 차이가 있다. 디지털과 관련한 정책, 특히나 디지털 기술의 발전이나 비즈니스, 전자상거래 등은 경제적 측면의 경쟁력 강화와 밀접하게 관련되어 있다. 이런 이유로 디지털 기술력 확보나 경쟁력 강화가 디지털과 관련한 시민의 권리확보나 중요하게 고려되어 오늘날과 같은 사회적 배제의 원인으로 작동하는데, 정보통신의 확산은 사회의 다양한 계층 간의 격차와 밀접하게 관련되어 있기도 하기 때문이다.

디지털 격차 문제는 단순히 기기 소유의 여부나 인터넷 구독 여부와 같은 디지털 접근에 대한 논의에서 시작하는데, 이는 빈곤의 문제와도 밀접하게 관련되어 있다. 즉, 소득의 차이와 밀접하게 관련되어 있고, 이러한 차이는 컴퓨터 등의 디지털기기에 대한 물리적 접근의 격차로 처음 나타나기 시작하는데, 빈곤은 디지털 문맹으로, 결국에는 사회적 배제로 이어지게 된다. 결국, 빈곤은 디지털 격차를 동반하고, 그 결과가 사회적 배제인 셈이다. 이와 관련해 교육, 연령, 성별, 민족이나 인종 그리고 장애 유무와 같은 사회적 요소가 디지털 격차에 대한 다양한 예측 가능한 변수로 인정되고 있다(Parsons, & Hick, 2008: 2). 이는 EU의 여성의 디지털 접근 점수(Women in Digital Scoreboard)나 호주의 디지털 포용지수를 통해서도 알 수 있는데, 디지털 격차로 인한 디지털 접근의 문제는 신체나 정신적 장애로 사회참여 제한과 유사하게 고려되는 것이다.

EU는 디지털 포용을 모든 사람이 디지털 세계의 혜택을 누릴 수 있도록 하는 노력이라고 규정하고 있다. 이는 더 쉬운 인터넷 사용을 통해 가능한데, 누구나 접근 가능한 ICT, 장애인의 디지털 세계 접근을 위한 ICT 개발지원, ICT를 활용한 교육, 그리고 이를 통해 시민이 주 변화 및 사회적 배제에 대응 지원, 공공, 사회 및 경제활동에서 소외된 사람들의 참여율 제고

를 위한 사회적 포용 프로젝트 추진, 온라인에서의 언어적 장벽 제거를 위한 다국어 전환 등을 통해 디지털 포용을 정책적인 추진을 통해 개념을 실천하고 있다.<sup>2)</sup> 심지어 독일은 UN의 장애인 협약을 준용해 디지털 포용을 개념화해 관련한 정책을 추진하고 있다. 즉, 디지털로 인해 사회적 접근에 문제가 생기는 경우를 모두 장애로 간주하고, 디지털 포용정책의 범위를 확대하고 있는데, 이는 디지털화되는 사회에서 장애의 개념을 사회적 접근성으로 확대한 것으로 볼 수 있다.

마찬가지로, Reber & Luginbuhl(2016) 역시 연구를 통해 디지털화된 세상에서의 디지털 활용 곤란과 사회적 포용의 역관계를 설명하고 있다. 이에 대해 디지털 기술로 형성되고 운영되고 있는 사회에서, 가장 먼저 사회적 포용을 위해서는 사회적 맥락에서의 장애의 의미에 대한 재고를 제기하였다. 그들은 교정이나 특수교육에 있어 디지털은 장애의 개념에 대한 전환을 이끈다고 보았다. 즉, 육체적 장애에도 디지털을 이용할 수 있다면 사회활동 참여나 직장 혹은 교육 등 모든 공적 사회생활이 가능하지만, 디지털 사회에서 디지털 접근에 문제로 인해 온라인 교육이나 재택근무 등을 수행할 수 없다면, 디지털 세상에서 이는 사회적으로 배제될 수 있기 때문이다. 그러므로 디지털 포용은 산업이나 비즈니스 기회로 보기보다는 변화에 대한 사회적 참여라는 규범적 측면에서 접근되어야 할 필요가 강조되는 것이다.

UN은 사회에서 디지털의 의미를 더욱 강조하고 있다. 빅데이터 시대는 디지털 활용의 결과로 발전하는데, 휴대전화, 온라인 커뮤니티, 모바일 앱, 메시지, 웹사이트, 무인자동차, 사물인터넷 등 새로운 도구를 활용한 의사소통이나 정보 생산이나 교환 등이 모두 현대사회에서의 경제와 정치 행정체제의 근간으로 자리매김하고 있어 포용적 사회정책 차원에서 교육 등과의 연계는 강조되고 있다. 이는 UN이 사회적 포용을 실현하기 위한 수단으로 제시한 디지털 협력 로드맵이나 국제전기통신연합(International Telecommunication Union, ITU)의 ICTs와 관련한 사회적 인프라 개발과 환경개선을 통해 모두를 위한 디지털 포용(Digital inclusion of all)의 추진, 유엔개발계획(United Nations Development Programme, UNDP)이 제공하고 있는 종합적이고 체계적인 디지털 접근을 위한 지침 등<sup>3)</sup>과 더불어 인권을 위한 필수적 도구로 인터넷 활용을 규정하고 보편적 디지털 활용이 국가정책의 우선순위가 되어야 함을 강조하고 있는 유엔인권이사회(United Nations Human Rights Council, UNHRC)의 이니셔티브를 통해서도 디지털 포용의 의미를 알 수 있다.<sup>4)</sup> 이러한 여러 디지털 포용에 관한 논의에 비해서도 다소 진보적이긴 하지만 Witten(2021: 2)의 말을 빌려 디지털 포용정책의 필요성을 다음과 같이 정리할 수 있다.

디지털 세상에서는 장애를 좀 더 광의의 개념으로 이해할 필요가 있다. 누군가 참여의 기회가 제한당한다면, 육체적/정신적 장애로 인한 사회참여의 제한과 같다. 참여는 사회의 모든 영역에서 완전한 소속감을 느낄 권리이기 때문이다.

2) <EC>, Digital Inclusion, <https://digital-strategy.ec.europa.eu/en/policies/digital-inclusion>

3) <https://www.undp.org/?search=digital+inclusion> (2023. 02. 10. 검색).

4) <https://www.ohchr.org/en/search?query=digital%20inclusion> (2023. 06. 10. 검색).

디지털 포용은 사회적 약자의 사회적 배제의 원인과 유사한 원인으로 나타나는 디지털 격차 문제에 대응하기 위한 개념이다. 이러한 개념을 바탕으로 2021년부터 프랑스는 디지털 포용법(안) 발의를 통해 사회적 권리와 시민권과 같이 개인의 권리 유지와 의무이행을 위해 반드시 선행되어야 할 요소로 디지털 포용을 전제함을 공표하였다. 사회적 측면에서의 다양한 디지털 정책을 이 법을 통해 그 목적을 명료화하고자 개정된 법안은 2022년 7월에 발의되었고, 현재 실효성을 갖는다. 미국 역시 디지털 형평법이 발의 중이다. 한국도 마찬가지이다. 이러한 개념적 이해를 바탕으로 우리 실정에서 추진되고 있는 다양한 디지털 관련 정책 중 디지털 포용과 관련한 정책을 살펴보는 것은 의미가 있을 것이다.

## 2. 한국의 디지털 관련 정책

우리나라는 ITU가 정보통신기술에 대한 접근성, 이용도, 활용력 등에 관해 2017년 176개국을 대상으로 한 평가에서 상위 2위로 디지털 분야에서 경쟁적 우위를 점하고 있었다(UN, 2018: 89). 2010년대 초 시작된 스마트폰의 확산과 더불어 급속한 디지털 스마트화는 우리를 정보통신 강국으로 만들어줬지만, 성장의 이면에 기존과 다른 사회적 배제집단의 생성, 즉 디지털 배제집단을 인식하지 못했다. 이는 사회적 취약계층의 정보화 수준이 OECD 평균보다 상당히 낮다는 보고를 통해 알 수 있는데(관계부처합동회의, 2020a: 3-4), ITU의 평가에서 경쟁적 우위를 점하고 있는 우리의 실정과 견주어 생각해본다면, 우리 사회의 디지털 격차 정도의 심각성을 짐작할 수 있다.

우리의 급격한 디지털 정보화 및 스마트화 추진과정의 역기능 중 하나가 디지털 소외계층의 형성이다. 디지털 소외계층에는 대부분 고령층, 장애인, 여성 등과 사회적 취약계층의 사람들과 함께 다양한 계층의 사람들이 포함되어 있다. 이는 지역에 따라서도, 업무에 따라서도 다양하다. 즉, 대한민국 누구라도 사회적 배제와 관련될 수 있음을 알 수 있다. 이에 우리 정부도 유럽이나 호주나 뉴질랜드 등의 국가와 같이 성별, 지역, 계층, 나이와 관계없이 모든 국민이 사회적 차별이나 배제 없이 잘 사는 국가를 의미하는 혁신적 포용 국가 실현을 위한 방안으로 디지털 포용을 추진하였다.

우리의 디지털 정책은 코로나19 팬데믹 이전과 이후로 나뉘 살펴볼 수 있다. 우리나라는 국가정보화 기본법을 바탕으로 5년마다 ‘국가정보화기본계획’을 수립하고 있는데, 팬데믹 이전인 2018년 ‘제6차 국가정보화기본계획(2018-2022)’에서 최초로 정보 지능화 혜택을 모두가 누릴 수 있게 한다는 목적으로 디지털 환경 조성 구축이 제기되었다. 또한, 정보 지능화 영역의 인재 양성을 위한 교육 등이 포함됨으로 디지털 포용정책에 관한 국가 의지가 표현되었다. 그러나 장기화된 팬데믹으로 인한 사회적 거리두기는 전 지구를 대면에서 비대면사회로 급속한 전환을 이끌었고, 디지털 격차 문제는 사회적 배제의 문제로 거론되었다. 여기에 대응방안으로 중앙정부의 여러 부처는 부처의 성격에 맞는 다양한 디지털 관련 정책을 추진하기 시작했다. 특히 2020년과 2021년 사회적 거리두기 기간에 나타난 디지털 격차에 대응하기 위해 관계부처합동회의가 시행되고 이를 통해 디지털 포용과 관련한 여러 정책이 수립·추진되게 되었다.

첫 번째가 디지털 뉴딜에 포함된 디지털 포용이다. 한국판 뉴딜의 한 부문으로 2025년까지

계획된 디지털 뉴딜은 국민생활과 밀접한 분야의 데이터 구축, 개방 및 활용, 전 산업으로의 디지털 적용, 지능형 정부 그리고 K사이버 방역체계 구축을 통한 D.N.A 생태계 강화, 교육 인프라 디지털 전환, 비대면 산업의 육성, SOC디지털화를 통한 디지털 격차 해소를 중심으로 하는 이니셔티브의 추진을 중점으로 하고 있다. 특히 과학기술정보통신부 주관으로 하는 주민센터나 도서관 등을 활용한 수준별 맞춤형 디지털 교육을 지역의 특성과 수요에 따른 디지털 격차 해소를 목적으로 하는 정책으로 간주해 디지털 포용을 위한 정책으로 볼 수도 있기는 하다. 그러나 실질적인 디지털 뉴딜은 디지털 기술을 활용해 한국경제의 저성장 국면을 극복할 수 있는 혁신적인 성장동력 및 포용적 성장동력을 창출해 포용적 성장의 토대 마련을 목적으로 하는 정책 추진이기 때문에, 디지털을 통한 경쟁력 강화에 초점이 있다(노규성, 2020: 27-28). 이에 디지털 격차 해소 역시 사회적 배제 예방적 성격보다는 디지털을 통한 경쟁력 강화적 성격을 지닌다고 볼 수 있다.

또한, 2021년 한국판 뉴딜 2.0으로 발전해 추진되고 있는 디지털 뉴딜에는 디지털 포용정책의 하나로 고용사회안전망구축이 포함되어 있다. 이는 코로나19 충격이 저소득층 청년 등 취약계층에 집중되는 등 사회적 양극화의 심화, 특히 디지털 격차를 해소하는 방안으로 시행되는 디지털 교육은 실질적으로 취업률 제고를 기대하는 정책으로 디지털 취약계층의 사회적 배제에 대한 대응보다는 개인의 경쟁력 강화 수단으로서의 디지털 정책에 좀 더 초점이 있다고 볼 수 있다.

다음으로는 디지털 포용을 목적으로 하는 혁신적 포용국가 실현을 위한 디지털 포용 추진 계획(2020a)과 디지털 포용 추진계획: 다함께 누리는 디지털 포용 세상(2020b)과 같은 정책의 추진이 있다. 이들 정책은 모두 관계부처합동으로 추진되고 있어 사회에서 디지털이 차지하는 중요성과 함께 개별적 정책을 통해서 이를 해결하기 곤란함을 알 수 있게 한다. 이들은 디지털 격차에 대한 대응을 목적으로 하는 디지털 포용을 ‘국민 모두는 차별이나 배제 없이 디지털 세상에 참여하여 디지털 기술의 혜택을 고르게 누리기 위한 사회 전체의 노력’으로 정의하고 있다(관계부처합동, 2020b: 8). 디지털 역량 구축을 위한 교육, 취약계층의 디지털 접근성 강화를 위한 디지털기기의 지원이나 네트워크 인프라 확대, 취약계층의 사회참여와 일자리 지원과 관련된 정책의 추진을 통해 혁신적 포용 사회 실현에 디지털 포용을 매개로 한다는 것을 이해할 수 있다. 이는 문재인정부의 포용 국가 실현 의지의 발현으로 볼 수 있다. 그러나 이러한 정책은 중앙정부 중심의 지역 간 디지털 격차 해소를 목적으로 하는 농어촌 초고속 인터넷망 구축, 공공와이파이 구축과 같은 인프라 구축을 제외한 대부분 정책에서는 사회적 취약계층을 정책 대상으로 명시하고, 동시에 관련된 정책 역시 취업교육 등과 같은 유형으로 앞서 살펴본 국제기구와 디지털 정책을 적극적으로 추진하는 나라들과는 그 기본적인 개념이 다름을 알 수 있다. 즉, 우리의 디지털 포용정책은 디지털 역량 강화, 포용적 디지털 이용환경 조성, 디지털 기술의 포용적 활용 촉진이라는 이니셔티브와 이의 실천 방안으로 디지털 포용법 제정이 시도되고 있기는 하지만 부처별로 흩어진 디지털 정책은 사회적 취약계층을 정책 대상으로 하고 있어 정책의 보편적인 적용이 곤란하다는 한계를 지니고 있다. 준비되지 않았던 비대면사회로의 전환은 사회적 약자뿐만 아니라 디지털 사용에 익숙하지 않은 사회 속 모든 사람이 잠재적 사회적 배제의 대상이 된다. 이에, 디지털 격차로 인한

사회적 배제 혹은 사회적 소외에 대해 디지털 소외계층 혹은 디지털 취약개념과 같은 명시적 개념 설정의 필요성을 제기할 수 있다. 디지털 기반으로 사회경제적 패러다임이 전환되는 과정에서 누구라도 직면할 수 있는 디지털 접근의 곤란과 이로 인한 사회적 삶의 만족도 결여 혹은 사회적 배제에 대한 제도적 대응이 디지털포용법(안)이다.

이렇게 제안된 법(안)은 디지털 포용을 사회의 모든 구성원이 소외와 차별 없이 지능정보 기술을 활용하고 지능정보서비스를 이용할 수 있도록 하여 지능정보사회 구성원의 삶의 질 향상과 지속 가능한 포용적 성장 증진을 추구하는 환경과 그 지향점으로 정의하고 있다(안 제2조). 디지털 포용법(안)은 디지털 접근과의 사회적 배제와의 관련성을 인지하고 있을 뿐만 아니라 디지털 활용을 통한 개인의 삶의 질에 초점을 두고 있어 기존의 디지털 포용정책보다는 정책 대상자의 범위를 확대하고 있다. 이는 디지털 전환의 가속화로 인한 디지털 격차와 이로 인해 나타날 사회적 문제에 대한 우려를 배경으로, 빅데이터, 네트워크, 인공지능 등 디지털 기술과 이로 인한 경제적인 경쟁력의 강화에 초점을 두기보다는 일상에서 디지털로 인한 국민 삶의 질 개선과 행복 증대를 목적으로 하고 있다는 점에서 큰 의미를 지닌다. 그러나 이러한 법(안)은 여전히 그 시행 여부를 알 수 없다.

그렇지만 윤석열 정부는 출범과 함께 과학기술정보통신부를 중심으로 하는 관계부처 합동 회의를 통해 4대 추진과제로 구성된 디지털 포용추진계획 의결 발표와 함께 다양한 포용적 디지털 전환사회 정책을 추진하고 있다. 여기에는 전국민 디지털 역량 강화, 포용적 디지털 이용환경 조성, 디지털 기술의 포용적 활용 촉진, 디지털 포용기반 조성과 함께 정책 수요자별 주요 추진 사항이 포함되어 있다. 그러나 2021년의 디지털 뉴딜과 상당히 유사성을 지니고 있다는 점에서 디지털 포용성의 한계를 보완할 필요가 있어 보인다.

### Ⅲ. 분석방법 및 분석과정

#### 1. 잠재 디리클레 할당(Latent Dirichlet Allocation, LDA)기반 토픽모델링

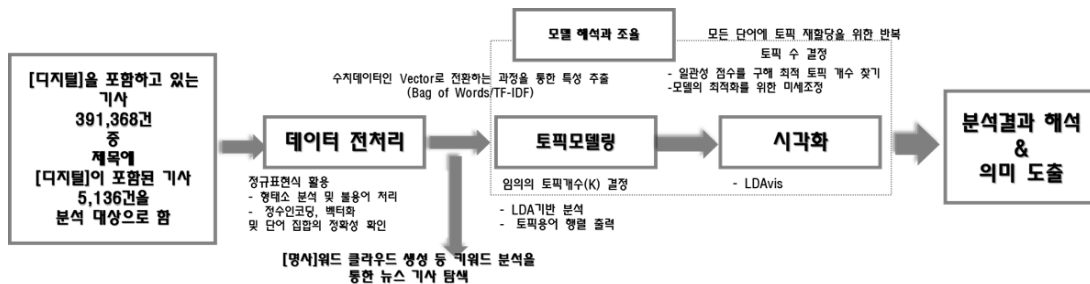
최근 우리나라에서 잠재 디리클레 할당(Latent Dirichlet Allocation, LDA)의 활용은 증가하고 있다(이대영·이현숙, 2021: 2). 이는 정책연구에서 빅데이터 활용의 증가를 통해서도 알 수 있다. 정책연구에 있어 빅데이터의 활용은 데이터의 정형화 여부와 관계없이 데이터 하위 모집단을 포괄적으로 살펴볼 수 있게 해, 우리가 흔히 사용하는 정제된 전통적인 데이터를 통해 발견하기 쉽지 않던 편향된 현상에서 사회 전반적인 현상을 두루 객관화해 살펴볼 수 있게 한다(김선영, 2020). 이러한 맥락에서 LDA토픽모델링은 사회의 현상을 반영하고 있는 뉴스 기사와 같은 텍스트 기반인 비정형 데이터를 통해 사회적 사안을 객관적으로 두루 살필 수 있도록 함으로 정책환경의 이해와 이를 바탕으로 현상에 대한 통찰을 가능하게 한다는 점에서 의미가 있기 때문이다. 이에 본 연구에서는 비정형 데이터 분석기법의 하나인 LDA토픽 분석을 통해 디지털 포용과 관련한 주요 이슈를 살펴보고자 한다.

토픽 모델링(Topic Modeling)은 텍스트로 구성된 문서 데이터에서 핵심 주제(Topic)를 찾는 텍스트마이닝 기법으로 문서 모델링을 통한 차원 축소 방법을 통해 주요 키워드를 스스로 선별하는 분석기법이다. 여기에서 가장 많이 사용되는 알고리즘은 LDA로 문서에 숨겨진 토픽(주제)을 알고자 할 때 사용되는 자연어처리 기법이다. 이는 문서와 문서, 혹은 단어와 단어 사이의 유사성에 대해 조건부 확률을 적용해 주제, 즉 토픽을 추론하는 방법(Steyvers & Griffiths, 2007)으로 문서 내 또 다른 문서 집합을 찾아낼 수 있다. Blei et al. (2003)이 개발한 LDA는 다른 토픽모델링에 비해 결과의 해석을 쉽게 할 수 있게 하고, 동시에 과적합 문제를 해결하기 때문에 방대한 비정형 데이터를 축소해 여러 토픽을 도출할 때 상당함 이점을 지닌다. 이런 이유로 자연어 분석과 관련해 다양한 영역에서 활용되고 있다. 이는 집합의 유사도를 측정해 이전에 알려지지 않은 의미 집합을 구성하는 생성 통계 모델(generative statistics model)을 기반으로 딥러닝을 통해 데이터를 비슷한 특징끼리 군집화하고 새로운 데이터에 관한 결과를 예측하는 방법인 비지도학습 모델로, 각각의 문서에서 드러나지 않은 숨겨진 의미를 찾기 위한 잠재변수 적용에 유용하게 사용되는 자연어처리 기법의 하나이기도 하다(Blei et al., 2003: 995). 또한, 각 부분의 통계적 패턴을 기반으로 유사도 분류가 가능한 요소에 적용될 수 있는 3단계 계층적 베이지안 모델을 적용한 것으로 단순한 주제의 분류보다는 주제에 포함된 주요 단어들을 제시하기 때문에 주제의 해석에 유용하다. 즉, 이는 문장에 직접 드러나지 않고 숨어있는 의미구조를 발견하기 위해 사용되는 텍스트 마이닝 기법으로, 일련의 텍스트에 숨겨진 정보를 찾아내 문서를 다양한 범주나 주제로 분류하는 기법인 토픽 분석이 발전한 형태이다. 이러한 토픽모델링을 통해 파악되는 토픽은 문서의 주제를 구성하는 용어들에 대한 연구자의 해석을 통해 이해될 수 있다.

## 2. 분석 대상 및 과정

이 연구는 LDA기반 토픽모델링 분석을 통해 우리 사회에서의 디지털 접근성과 관련한 사안을 객관화해 살펴보고, 이와 관련한 정책과 문헌연구를 바탕으로 우리의 디지털 접근성에 관한 함의를 찾는 것에 목적이 있다. <그림 1>은 LDA기반 토픽모델링 수행과정이다.

〈그림 1〉 LDA 기반 토픽모델링 분석과정





이를 위해, 뉴스 통합검색사이트인 빅카인즈를 활용해 MBC와 KBS를 포함한 54개 언론사의 2020년 1월 1일 이후 2023년 1월 31일까지 디지털과 관련한 기사 391,368개 기사를 수집하였다. 그리고 이를 바탕으로 제목에 디지털을 포함하는 기사 5,136건을 추출해 분석하였다. 제목을 기준으로 기사를 한 번 더 추출해 데이터의 정확성을 높였다. 수집된 기사는 정수화 인코딩 과정을 통한 토큰화 결과 89,265개 토큰으로 벡터화하는 데이터 정제과정을 수행하였다. 토큰은 의미를 지니는 최소 단위로, 이 연구에서는 단어 토큰화(Word-Tokenization)를 바탕으로 수행했는데, 이때 형태소는 명사로 변환하였다.

우리 사회에서의 디지털과 관련한 사안을 객관화한 결과를 나타내는 LDA기반 토픽모델링 분석은 데이터 항목들의 분포 확인과 데이터의 종합적 이해를 위해 탐색적 데이터 분석과 더불어 수행되었다.

## IV. LDA기반 토픽모델링 분석 및 분석에 대한 이해

### 1. 데이터 탐색

자연어처리(Natural Language Processing, NLP)에서 언급된 단어의 빈도수를 바탕으로 하는 시각화 기법인 워드 클라우드 분석을 시행하였다. 숫자로 표현되는 데이터가 아닌 경우를 모두 비정형 데이터라고 하는데, 이러한 데이터는 먼저, 정제과정을 통해 분석 가능한 형태의 데이터로 변환을 거친다. 이 과정에서 데이터의 일관성 및 신뢰성 확보를 담보하는 과정을 거쳐야 한다. 이를 위해 세종사전 실행을 통해 형태소를 분리하고, 명사를 추출한 후 분석을 시행하였다.<sup>5)</sup> 이를 바탕으로 하는 워드 클라우드는 〈그림 2〉를 통해 살펴볼 수 있다. 이는 순전히 단어의 빈도수 측정을 바탕으로 하기 때문에, 전체 데이터를 살펴보는 데 용이하다. 워드 클라우드가 네 가지로 표기된 이유는 형태로 분석을 통한 단어 정제과정에서 한글은 1음절 이상이면 의미를 지닐 수 있기 때문이다.<sup>6)</sup>

5) 형태소사전. <https://kbig.kr/portal/kbig/datacube/niadict.page> (2023. 09. 02. 검색).

6) 시계방향으로 1음절 이상, 2음절 이상, 3음절 이상, 4음절 이상으로 분류해 워드 클라우드 분석을 수행했으며, 기사에서 명사만을 추출해 분석의 대상으로 하였다.

〈그림 2〉 워드 클라우드



뉴스 기사에 디지털과 관련해 나타난 대부분은 과학기술을 중심으로 하는 산업과 이를 통한 국가경쟁력 확보를 위한 도구로 나타나고 있다. 관련해서 〈표 1〉은 디지털과 관련한 용어의 빈도를 워드 클라우드 보다는 조금 더 정량적으로 보여준다.

〈표 1〉 뉴스 기사 내 디지털과 관련한 용어

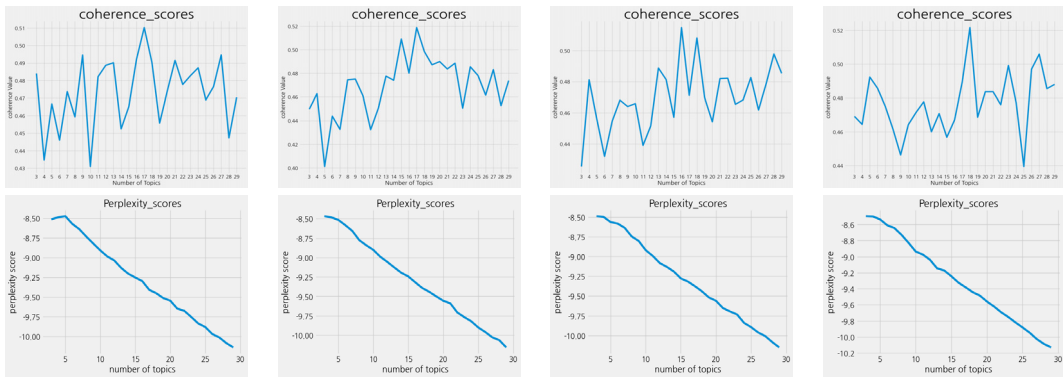
| 디지털 관련 용어 | 출현 건(출현 비)   | 디지털 관련 용어 | 출현 건(출현 비)   |
|-----------|--------------|-----------|--------------|
| 디지털       | 5120건(99.7%) | 디지털격차     | 426건(8.3%)   |
| 디지털지원     | 3037건(59.1%) | 디지털서비스    | 2046건(39.8%) |
| 여성        | 561건(10.9%)  | 노인        | 355건(6.9%)   |
| 사회진입제한    | 386건(7.5%)   | 소외        | 924건(18.0%)  |
| 국가경쟁력     | 1747건(34.0%) | 사회정의      | 433건(8.4%)   |
| 장애인       | 710건(13.8%)  | 취업        | 518건(10.1%)  |
| 디지털정책     | 2707건(52.7%) | 디지털행정     | 879건(17.1%)  |
| 사회취약계층    | 1617건(31.5%) | 저소득층      | 209건(4.1%)   |
| 학생        | 459건(8.9%)   | 고용        | 1109건(21.6%) |
| 배제        | 261건(5.1%)   | 차별        | 669건(3.0%)   |
| 녹색성장      | 445건(8.7%)   | 지속가능성     | 1943건(37.8%) |
| 갈등        | 426건(8.3%)   | 교육        | 1885건(36.7%) |

## 2. LDA 기반 토픽모델링 분석과정의 이해

데이터 정제 및 데이터 탐색 후 LDA기반 토픽분석의 수행은 연구자가 임의로 설정한 토픽 수 10개를 바탕으로 반복적으로 수행한 토픽 수에 대한 적절성 판단으로부터 시작한다. LDA

는 비지도학습 모형이기는 하지만, 토픽의 개수는 연구자가 지정해야 한다는 한계를 지닌다. 이에, 분석의 정확도를 높이기 위해 토픽 수를 10(N)개로 연구자가 임의로 지정해 분석을 시도한 후, 일관성 및 혼란도 계수를 확인하였다. 일관성 계수는 해당 모델이 실제로 의미 있는 결과를 나타내기 위해서 사용되는데, 높게 나타날수록 의미론적으로 일관성이 높음을 의미한다. 반면, 혼란도 계수는 로그지수로 지정된 범주의 숫자가 낮아진다면, 모델이 적정범위에 있음을 의미한다. <그림 3>은 반복적으로 수행한 일관성(Coherence)과 혼란도(perplexity) 측정 결과로, LDA기반 토픽모델링 기법의 수렴속도 확인 및 확률 모델의 개선 정도에 대한 평가를 보여주는 것이다.<sup>7)</sup> 즉, 이 연구에서 지정한 토픽 수의 타당성과 관련된다. 이러한 기계적 적절성과 함께 선행연구를 통해 얻은 정보를 바탕으로 토픽의 수를 17개로 할당하였다. 이는 연구자가 선행연구를 통한 사회에서의 디지털에 관한 이해와 반복적인 분석에서 보여주는 토픽과의 관련성에 대한 직관적 판단을 근거로 하고 있다.

〈그림 3〉 토픽 수 선정 근거<sup>8)</sup>



〈표 2〉과 〈표 3〉은 수집된 데이터를 Gensim<sup>9)</sup>을 활용해 비지도 학습을 수행시킨 결과이다. 먼저 〈표 2〉가 제시하는 상위 30개 중요 용어는 주어진 단어의 조건부 확률 계산 값으로 관찰된 단어  $w$ 가 잠재적 토픽  $T$ 에 의해 생성되었을 가능성을 의미한다. Chuang, et al.(2012: 75)의 연구를 통해 토픽 생성과정을 살펴볼 수 있다.

- 7) 일관성(Coherence) 측정은 토픽의 의미론적인 측면에서의 일관성을 의미하는데, 한 주제, 즉 하나의 토픽 안에 의미론적으로 유사한 단어가 많이 모여있다는 것을 의미한다. 반면, 토픽모델이 문서를 잘 반영하는 로그합수를 기반으로 하는 혼란도(perplexity)는 실제로 관측되는 값이 얼마나 정확하게 나타내는지를 보여주는 값으로, 동일 모델 내 파라미터에 따른 성능 평가할 때 주로 사용한다. 의미확률 모델이 결과를 얼마나 정확하게 예측하는지는 그 값이 낮을수록 정확한 예측을 의미한다. 그러나 의미상의 정확도를 반영하지 못한다는 한계를 지닌다.
- 8) 연구자는 토픽 수를 선정하기 위해 수십 회를 반복해서 학습시켰으며, 연구에서는 그 일부분만 토픽 수의 선정과 관련해 일부분만 보여주고 있다.
- 9) Gensim은 Python 라이브러리 중 하나로 문서를 의미 체계 벡터로 비지도 머신러닝 알고리즘을 사용하여 구조화되지 않은 raw data인 텍스트를 처리하도록 설계된 프로그램이다(Gensim, Topic Modeling for Humans). <https://radimrehurek.com/gensim/>

$$distinctiveness(w) = \sum_T P(T|w) \log \frac{P(T|w)}{P(T)}$$

즉, 주변확률  $P(T)$ 는 무작위로 선택된 단어  $w_0$ 은 토픽  $T$ 에 의해 생성될 가능성을 계산하는 것으로 생성 주제와 무작위로 선택된 용어  $w_0$ 의 비교를 통해 특정 항  $w$ 가 결정에 얼마나 유익한지를 측정하는 것이다. 여기에서 차별성(distictiveness)은 두 확률분포의 차이를 계산하는데에 사용하는 함수 Kullback-Leibler divergence(KLD)을 통해 정보 엔트로피 차이를 확률적으로 계산한 것이다. 이를 바탕으로 용어의 중요성(saliency)은 다음과 같은 식으로 표현될 수 있다.

$$saliency(w) = P(w) \times distinctiveness(w)$$

이를 바탕으로 하는 <표 2>는 5,136개 뉴스 기사에서 토픽을 구성하기 위한 바탕으로 사용된 주요 상위 30개 단어이다.

<표 2> LDA토픽을 구성한 주요 30개 단어

| 중요도 | 용어  | 중요도 | 용어  | 중요도 | 용어    |
|-----|-----|-----|-----|-----|-------|
| 1   | 디지털 | 11  | 지원  | 21  | 한국    |
| 2   | 경제  | 12  | 위기  | 22  | 국회    |
| 3   | 정부  | 13  | 기업  | 23  | 국가경쟁력 |
| 4   | 뉴딜  | 14  | 확대  | 24  | 정보    |
| 5   | 세계  | 15  | 교육  | 25  | 데이터   |
| 6   | 국민  | 16  | AI  | 26  | 예산    |
| 7   | 대통령 | 17  | 서비스 | 27  | 경영    |
| 8   | 금융  | 18  | 한국판 | 28  | 고객    |
| 9   | 지역  | 19  | 사업  | 29  | 검사    |
| 10  | 협력  | 20  | 구축  | 30  | 국장    |

<표 2>는 전체 기사의 파악한 상위 30개의 중요 용어를 기반으로 LDA에서 생성한 임의의 토픽과 토픽에 할당된 관찰된 용어와 수의 비중을 보여주고 있다. 전체용어 빈도에 따른 토픽용어 행렬의 출력값으로 주제 집합 내 관련 용어들로, 토픽용어 행렬 출력값은 분석 데이터 중 각 토픽에 속했다고 추정되는 기본 용어이다. 이들은 각 토픽에 속한 추정된 용어를 통해 각 집합의 주제를 좀 더 명확하게 유추할 수 있다.

그렇다면, <그림 2>, <표 1>과 <표 2>의 차이에 대해서도 살펴볼 필요가 있다. 탐색적 결과인 <그림 2>, <표 1>은 뉴스 기사의 토큰화를 바탕으로 한 단어의 빈도수를 측정한 결과이다. 반면 <표 2>는 어떠한 명령어나 조건도 없이 그저 주어진 데이터, 즉 문장의 숨겨진 의미를 기계학습한 결과이다. 즉, 관련 주제를 나타내는 용어들을 클러스터링하는 과정을 학습한 단어의 중요도에 따라 제시된 결과라는 점에서 빈도와 중요도라는 점에서 각각을 구분해서 살펴보는 것은 의미가 있다.

〈표 3〉 토픽행렬 출력 및 잠재적 추정 용어

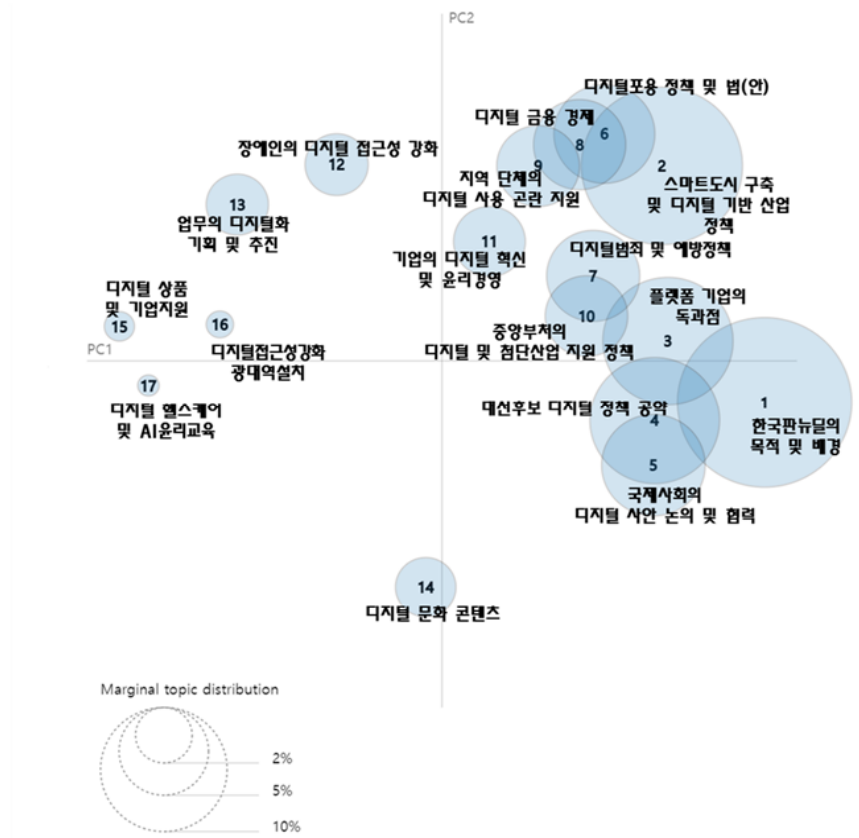
| 토픽<br>번호 | 토픽할당량<br>(%) | 전체용어 빈도에 따른<br>토픽용어 행렬 출력값   | 각 토픽에 속한 추정된 용어   |
|----------|--------------|--|---|
| 1        | 18.9         | 0.030*뉴딜 + 0.028*경제 + 0.020*국민 + 0.017*정<br>부 + 0.016*위기 + 0.016*한국판 + 0.015*세계 +<br>0.014*대통령 + 0.013*코로나 + 0.011*일자리       | 고용안전망, 국민보호대회, 사람투자, 직업훈련체계,<br>국가기발시설, 고용충격, 실직, 한국판뉴딜, 국가경<br>쟁력, 방역, 고용, 안전망, 선도국가, 사회안전망,<br>160조, 투자, 일자리, 기회                |
| 2        | 16.1         | 0.016*사업 + 0.012*지원 + 0.011*구축 + 0.011*추<br>진 + 0.011*계획 + 0.011*혁신 + 0.010*디지털 +<br>0.010*산업 + 0.009*도시 + 0.008*기업          | 물관리, 3D프린팅, 스마트도시, 공공시설, 테스트베<br>드, 연구용역, 포용도시, KOSPO, 스마트공장, 온라<br>인동영상서비스, 플래그십, 바우처, 사물인터넷,<br>IoT, 인프라, 융합, 생태, 에너지           |
| 3        | 10.6         | 0.013*사회 + 0.010*코로나19 + 0.008*기업 +<br>0.008*시대 + 0.007*디지털 + 0.007*변화 + 0.007*<br>경제 + 0.006*교수 + 0.006*한국 + 0.005*세계       | 조성욱교수, 독과점, Z세대, 사회보장제도, 불공정행<br>위, FATF, 경사노위, 혐오, 갑을, 편견, 자본주의,<br>아마존, 공정위, 컨퍼런스, 트렌드, 메타버스, 플랫폼                               |
| 4        | 10.5         | 0.014*정부 + 0.013*국민 + 0.011*대통령 + 0.010*<br>국회 + 0.009*정책 + 0.008*의원 + 0.008*정치 +<br>0.008*후보 + 0.008*대표 + 0.007*생각          | 더불어민주당, 국무의원, 국민의힘, 행복국가, 민심,<br>지지율, 후보, 지지율, 대변인, 문재인, 안철수, 이재<br>명, 이낙연, 법안, 규제, 통합  |
| 5        | 6.7          | 0.023*협력 + 0.019*한국 + 0.015*대통령 + 0.012*<br>중국 + 0.011*경제 + 0.009*국가 + 0.009*세계 +<br>0.009*미국 + 0.008*대응 + 0.008*강화          | APEC, CPTPP, WTO, SDG, 지속가능발전목표,<br>지구공동체, 경제동반자협정, 국가주식, 정상회의,<br>동북아시아, 사우디아라비아, 이니셔티브, 유엔총회,<br>시진핑, 협력체, 개도국, G20, 안보, 외교, 협력 |
| 6        | 6.3          | 0.092*디지털 + 0.020*정보 + 0.019*교육 + 0.019*<br>데이터 + 0.015*포용 + 0.015*AI + 0.013*활용 +<br>0.012*정부 + 0.011*사회 + 0.011*서비스        | 정보문화, 디지털배움터, KISA, 지능정보사회, 정보<br>통신전략위원회, 디지털플랫폼정부, 한국인터넷진흥<br>원, 정보격차, 디지털포용법, 클라우드, 스마트폰, 인<br>터넷                              |
| 7        | 5.4          | 0.032*교육 + 0.012*지원 + 0.012*여성 + 0.011*사<br>회 + 0.010*청소년 + 0.010*학교 + 0.008*아동 +<br>0.008*학생 + 0.007*정책 + 0.006*추진          | 학부모, 학력, 원격수업, 고교학점제, 미래교육, 성폭<br>력, 스토킹, 여가부, 교육부, 피해자, 아이, 여성, 가<br>족, 성범죄, 성평등, 행복, 안전, 피해자, 정책                                |
| 8        | 5.3          | 0.043*금융 + 0.012*디지털 + 0.012*서비스 +<br>0.010*대출 + 0.009*CBDC + 0.008*결제 + 0.007*<br>화폐 + 0.006*은행 + 0.006*상품 + 0.006*소비자      | CBDC, 디지털화폐, 금융안정, 파일럿, 토큰, 비트<br>코인, 토스뱅크, 금융포용, 암호화폐, 핀테크, 블록<br>체인, 카카오�뱅크, 가상자산, 현금, 대출, 마이데이<br>터, 금융회사, 리스크                 |
| 9        | 4.3          | 0.019*지역 + 0.019*금융 + 0.017*광주은행 +<br>0.014*은행 + 0.014*고객 + 0.011*코로나19 +<br>0.009*소상공인 + 0.009*강화 + 0.009*디지털 +<br>0.009*광주 | 광주은행, 사회공헌활동, 봉사활동, 충남스마트센<br>터, KMAC, 리딩뱅크, 플러스론, 포용금융센터, 지<br>역밀착경영, 지역민, ATM, NH, 소상공인, 지역민                                    |
| 10       | 4.2          | 0.025*예산 + 0.021*지원 + 0.014*정부 + 0.014*확<br>대 + 0.011*재정 + 0.011*내년 + 0.010*투자 +<br>0.009*편성 + 0.008*규모 + 0.007*추경           | 규제샌드박스, K-바이오, 누리호, 농식품부 해수부,<br>공공조달, 발사체, 예산편성, 예산안, 산업부, 경제회<br>복, 추경, 수출, 기재부, 부품, 해양   |
| 11       | 3.2          | 0.019*경영 + 0.016*회장 + 0.014*고객 + 0.013*기<br>업 + 0.013*대표 + 0.012*혁신 + 0.012*사장 +<br>0.011*조직 + 0.010*성장 + 0.010*ESG          | 기업은행, 한국수자원공사, BNK금융그룹, 웰니스,<br>애경케미칼, GM, 정보접근센터, 윤리경영, ESG, 혁<br>신경영, IBK기업은행, 신한금융, 화웨이, 혁신금융,<br>디지털                          |

| 토픽<br>번호 | 토큰할당량<br>(%) | 전체용어 빈도에 따른<br>토픽용어 행렬 출력값   | 각 토픽에 속한 추정된 용어  |
|----------|--------------|--|--|
| 12       | 2.5          | 0.022*탐장 + 0.018*미디어 + 0.015*장애인 + 0.015*방송 + 0.011*디지털 + 0.011*본부 + 0.009*서울시 + 0.009*정보 + 0.007*서울 + 0.007*서비스 | 자막, 지능데이터, 인프라본부, 방송통신, 인프라본부, 미디어, 방통위, 수어, 자상파, 키오스크, TV, 공공데이터, 단말기, 콘텐츠, OTT                 |
| 13       | 2.4          | 0.038*국장 + 0.034*검사 + 0.022*금융 + 0.018*과장 + 0.016*실장 + 0.014*부장 + 0.012*감독 + 0.012*전보 + 0.012*기획 + 0.010*본부장     | 경무기획정보화, 티앤씨재단, 서울중앙지검, 법무실, 관리국장, 센터장, 파견, 사무소장, 본부장  |
| 14       | 2.3          | 0.019*문화 + 0.013*세계 + 0.009*콘텐츠 + 0.008*예술 + 0.008*역사 + 0.007*전시 + 0.007*나라 + 0.006*작품 + 0.006*작가 + 0.006*평화     | 문화예술, 대한민국임시정부, 문화예술인, 클래식, M세대, 국가유공자, 뮤지엄, 미술관, 기념관, 민주공화국, 실리콘밸리, BTS, 역사, 영화, 음악, 국가경쟁력, 디지털 |
| 15       | 0.6          | 0.048*지점장 + 0.030*지점 + 0.011*부장 + 0.009*금융 + 0.009*여신 + 0.008*본부 + 0.008*센터장 + 0.007*센터 + 0.007*영업 + 0.007*팀장    | 지점장, 혁신부, 개발부, 전남대학교, 기획부장, 봉선이마트, 신탁, 하남공단, PB, 승진  |
| 16       | 0.5          | 0.022*와이파이 + 0.017*구축 + 0.017*공동 + 0.015*서비스 + 0.015*통신 + 0.014*KT + 0.014*농어촌 + 0.014*이용 + 0.013*5G + 0.009*공공  | 통신3사, 기지국, 주파수, 공동이용망, 기지국, 와이파이, LTE, GHz, Ap, 공동구축, 영상통화, 힙스터, 농어촌, 통신망, 광대역                   |
| 17       | 0.3          | 0.053*AI + 0.029*윤리 + 0.010*쉼터 + 0.010*카카오 + 0.007*기준 + 0.006*성동구 + 0.005*성동 + 0.004*원칙 + 0.003*혈압 + 0.003*제정    | 한국MS, 네트워크, 스마트쉼터, 이동노동자, 소프트웨어교육체험센터, 웹엑스, 리얼리티, AI, 윤리, 체험교육, AWS, 카카오피커, 원칙, 제어, 목소리, 혈압      |

이러한 분석의 결과는 <그림 4>를 통해 살펴볼 수 있다. 디지털, 정부, 지원, 한국판 뉴딜, 교육, 서비스, 코로나, 중국 등은 디지털 포용의 중요성이 제고된 배경과 디지털 포용정책에서 드러나는 용어로 주제를 형성하는데 필수적으로 작동한 용어인 <표 2>을 바탕으로 측정된 결과인 <표 3>은 <그림 4>로 시각화된다. 즉, <표 3>의 전체용어 빈도에 따른 토픽용어 행렬 출력값을 LDavis에 적용해 시각화한 결과이다.

〈그림 4〉 디지털 포용 관련 이슈에 대한 시각화

Intertopic Distance Map (via multidimensional scaling)



LDavis는 단어 개수만큼 존재하는 차원 수를 2차원으로 축소해 그 결과를 보여주는데, 주성분 분석(Principal Component Analysis)을 통한 차원의 반복적인 축소를 통해 주변 토픽의 분포(topic distribution)를 나타낸다. 〈표 3〉에 구성된 용어들은 디지털 포용과 관련한 정책과 선행연구를 바탕으로 해석해 〈그림 4〉에 주제를 표기하였다.

〈그림 4〉에 나타난 17개의 토픽은 각 토픽 사이의 거리에 따라 토픽과 토픽에 포함된 용어들의 관련성을 나타낸다. 즉, 1,4,5,9번 주제와 같이 집합이 겹치거나 거리가 가깝다는 것은 집합에 포함된 용어와 그 주제들의 높은 상관성을 나타낸다. 즉, 디지털 포용법(안)이 추진된 배경, 디지털 뉴딜, 사회적 배제, 국제사회에서의 디지털과 디지털로 인한 사회적 문제에 대한 논의는 상당히 높은 비율로 중복된 용어를 지니는 상관성 높은 주제임을 보여주고 있지만, 여전히 우리 사회에서의 디지털은 경제적 측면에서의 경쟁력 강화 수단과 관련한 이슈가 대부분임을 알 수 있다. 우리 사회에서의 디지털 포용과 관련한 이슈를 통해 디지털 포용정책 역시 가능할 수 있다. 즉, 디지털 포용정책은 이러한 사회적 이슈를 반영한다고 간주할 수 있다.

### 3. 분석결과 해석: 뉴스 기사를 통해 본 디지털 관련 이슈

코로나19 팬데믹과 함께 디지털 격차가 본격적으로 공론화되기 시작하면서, 이로 인한 사회적 배제의 심각성도 함께 심각성이 인지되기 시작하였다. 이에 대한 논의와 함께 디지털 포용과 관련한 여러 정책의 추진이 시도되었다. 그러나 뉴스데이터 분석은 우리 사회에서의 디지털과 관련한 이슈는 대부분이 다양한 측면의 경쟁력 확보 수단으로 강조되고 있음을 보여주고 있다.

〈표 3〉를 시각화한 결과인 〈그림 4〉에 대한 해석은 다음을 바탕으로 한다. 먼저, 토픽 간의 거리를 고려해야 한다. 토픽 간의 거리가 멀수록 판별 타당도가 높고 주제의 차이를 분명하게 보여주기 때문이다. 둘째, 토픽 간의 상관관계를 확인할 필요가 있다. 이는 토픽 마다의 유사성과 차이점을 파악하는 것은 토픽이 지닌 의미 해석에 필수적이기 때문이다.<sup>10)</sup> 즉, 이 연구의 경우 기본 30개의 단어가 상당 부분 거의 모든 토픽에 포함되어 있다. 그리고 그 단어들은 각각 다른 단어들과 연관되어 다른 토픽을 구성하고, 각 토픽은 고유한 의미를 지니지만, 〈그림 4〉를 통해 알 수 있듯이, 교집합, 즉 유사한 의미를 지니는 토픽 역시 상당하다. 즉, 클러스터 간의 가까운 거리로 각각의 토픽은 상당히 유사성을 지니고 있음을 뜻하기 때문이다. 이 토픽을 해석하면 다음과 같이 정리할 수 있다.

기사를 통해 살펴본 우리 사회에서의 디지털과 관련한 이슈는 한국판 뉴딜과 스마트도시 구축으로 크게 구분해 설명할 수 있는데 이 역시 경쟁력 확보와 밀접하게 관련되어 있음을 알 수 있다. 먼저, 토픽 1,3,4,5,7,10은 서로 밀접하게 이웃하고 있다. 즉, 2022년 치러진 대통령선거에서도 디지털은 중요한 이슈로, 사회경제문화를 모두 아우르고 있음을 알 수 있다. 이를 좀 더 구체적으로 살펴보면, 경제적 측면에서의 한국판 뉴딜과 플랫폼 기업, 그리고 국제사회에서 논의되고 있는 디지털 관련 주요 이슈의 국내 적용, 국가의 경쟁력 확보 방안으로서의 디지털첨단산업지원 정책과 함께 디지털과 관련한 사회적 이슈 및 디지털범죄가 서로 관련한 이슈로 나타나고 있음을 알 수 있다. 특히나, 디지털과 관련한 여러 정책은 우리 사회가 인식한 디지털범죄와도 밀접하게 관련되어 나타난다는 것은 주목할 만한 점이다. 그러나 디지털 격차와의 사회적 배제와의 관련성에 대해 논의되는 것, 여전히 우리 사회에서의 디지털은 국가와 개인의 경쟁력 확보를 담보하기 위한 수단으로 논의되고 있다는 것을 짐작할 수 있게 한다.

둘째, 스마트 시티 구축 및 디지털 기반산업 정책에 관한 토픽 2는 디지털 금융경제, 지역 사회의 디지털 접근 곤란 지원 등 디지털 포용과 관련해 토픽 6,8,9와 관련되어 있다. 즉, 디지털 뉴딜, 사회적 포용, 디지털 격차 문제, 디지털 금융 역시 밀접하게 관련되어 있는데, 이는 디지털 격차를 디지털 뉴딜에서의 강조되는 교육과 같은 정책의 포함이 사회적 포용과 관련되어 있음을 나타내 주기는 하지만, 여전히 디지털 사회에서의 디지털 접근성으로 인해 야기될 문제에 대한 심각성은 나타나지 않고, 디지털 격차와 관련한 이슈는 디지털 금융경제에 대한 예방이 이슈화됨을 알 수 있다. 셋째, 토픽 11의 경우 기업의 디지털 혁신 및 윤리경영과 관련되어 있는데 이는 디지털 사회 전반과 관련되어 있어 한국판 디지털 뉴딜과 디지털

10) 학습 전 연구자의 의도에 따라, 연구자는 토픽에 포함된 단어의 수를 코딩을 통해 조정할 수 있다. 이 연구에서는 Blei et al(2003)와 같이 30개로 설정하였다.



시티 구축 및 디지털 기반산업과 모두 근접해 있다.

넷째, 사회적 취약계층의 디지털 접근성과 관련한 공론화이다. 이는 토픽12의 장애인 디지털 접근에 관한 사안으로 나타나고 있지만, 내용에는 사회적 취약계층의 디지털 접근으로, 디지털 격차와 관련한 사회적 이슈를 보여주고 있다. 다섯 번째로는 물리적인 디지털 환경의 구축과 관련된다. 이는 토픽 16을 통해 알 수 있다. 마지막으로 디지털상품 및 기업지원이나 업무의 디지털화, 디지털 헬스케어나 디지털 문화 콘텐츠 등으로 나타나고 있는데 각 토픽의 거리를 통해 그 관련성을 유추해 볼 수 있다.

이를 정리하면 우리나라의 디지털과 관련한 사회적 이슈는 디지털을 통한 편의성과 경쟁력 확보 수단으로 치우쳐져 있음을 알 수 있다. 즉, 개인의 사회참여 수단으로서의 디지털의 활용보다는 디지털을 통한 개인의 경쟁력확보, 그리고 사회가 요구하는 인력확보의 수단으로서의 성격이 강하게 나타나고 있어, 디지털 사회의 가속화로 인해 심각해질 사회적 배제와 이로 인해 야기될 문제는 여전히 사회적으로 인식되지 않고 있다.

국제사회에서 논의되고 있는 디지털은 단순한 경쟁력확보의 수단을 넘어 사회적 참여의 도구로 인정되고 있다. 그 예로 독일과 같은 국가는 UN의 장애인 인권협약을 준용해 디지털 접근 제약을 정신적/육체적 장애로 인해 야기되는 사회적 배제로 간주하고 있다. 지금과 같은 지능정보사회에서의 사회적 포용 수단으로서의 디지털과 관련한 사안이 공론화되고 있지 않다는 점을 해소하기 위해 사회적 포용 도구로 디지털에 대한 의미를 국가적 차원에서 재고할 필요가 있어 보인다. 이는 지능정보화사회에서의 디지털 접근성은 조직과 기관으로의 접근이자, 나아가서는 사회참여의 질에 관한 것으로 간주될 수 있기 때문이다.

## V. 정책적 함의 및 마무리

정보격차에 대한 오랜 논의는 디지털 격차로 이어지고, 이는 독일의 산업 4.0 발표 이후 발전하는 지능정보사회와 디지털에 관한 논의로 계속되고 있다. 디지털 기술개발을 통한 개인, 지역, 그리고 국가의 경쟁력 강화에 관한 논의에서 디지털 인프라 구축이나 개인의 사용역량까지 다양한 측면에서 디지털의 활용과 사회적 포용을 위한 수단으로 디지털에 관한 논의로 이어지게 된 것이다. 특히, 이 논의는 코로나19 팬데믹 이후 여러 가지 측면에서 차이를 지니는데, 특히나 사회적 측면에서 디지털은 그것을 활용한 경쟁력확보 수단이나 혹은 삶의 편의성 제고를 위한 활용적 측면에서의 논의에서 나아가 사회적 포용을 위한 수단으로 사회에서의 디지털이 지니는 의미는 전환되었다.

UN(2021)의 ‘사회적 포용을 위한 디지털 기술의 활용’과 같은 발표를 통해 알 수 있듯이, 코로나19 바이러스로 인한 사회적 거리두기로 인해 디지털은 사회참여의 기제로 공론화되었다. 디지털 격차로 인해 야기되는 사회적 배제와 이에 대한 대응으로 디지털 포용 추진이 사회적 측면의 디지털 정책의 화두로 자리 잡게 된 것이다. 이와 관련해 2021년 발의된 디지털 포용법(안)은 우리 사회에서의 사회적 포용 측면에서의 디지털은 경쟁력 제고 수단과 편리함을

위한 도구로서의 디지털 기술 진보, 경제적 발전, 지능정보사회에서의 디지털 접근 환경, 그리고 디지털을 사용할 수 있는 역량에 관한 논의의 결과이지만 여전히 그 시행 여부는 명확하지 않다. 이는 우리 사회가 여전히 디지털을 사회적 포용을 위한 도구적 측면보다는 경제적 발전 혹은 경쟁력 강화를 위한 도구에 비중을 두고 있기 때문으로 유추할 수 있다.

지능정보사회는 디지털 세상을 의미한다. 여기에서 디지털 접근의 곤란은 사회적 배제와 불평등으로 이어진다. 사회적 약자가 경험하는 사회적 배제에 대응하기 위해 사회적 포용이 마땅한 것과 같이 디지털 접근으로 인한 사회적 배제에 대한 대응도 마찬가지이다. 즉, 디지털 포용은 기술적 진보 과정에서 요구되는 사회적 포용정책의 하나이자 사회적 포용의 포괄적 적용으로 볼 수 있다. 그러나 2020년 1월 이후 디지털과 관련한 이슈는 디지털 뉴딜이나 디지털 경제 등 경제적 측면의 경쟁력 강화 수단과 관련한 사안이 대부분으로 디지털기기 사용 문제로 인한 사회적 접근 곤란이나 이러한 해결과 관련한 사안은 찾아볼 수 없다는 점이 아쉬운 부분으로 지적될 수 있다. 이 연구는 다음과 같은 함의를 제시하고자 한다.

첫째, 디지털 접근성은 물리적 환경도 중요하지만, 실질적으로 디지털을 사용할 수 있는 역량의 중요성이 강조된 정책이 요구된다. 지금의 중앙정부 중심의 디지털 포용은 디지털 인프라 구축을 통한 물리적 디지털 접근성 강화에 초점을 두고 있다. 이는 2020년 6월 과기부를 중심으로 발표된 디지털 포용 추진계획을 통해서도 알 수 있다.<sup>11)</sup> 인터넷 광대역의 설치 정도나 스마트폰이나 컴퓨터 보유나 혹은 전자정부의 완벽한 구축만으로 디지털 격차가 해소될 수 없어서, 지역공동체 차원의 디지털 격차 해소를 목적으로 하는 프랑스의 디지털 패스<sup>12)</sup>와 같은 정책 추진이 요구된다.

두 번째는 일상에서 디지털기기 보급 따라 기본적인 활용에 대한 적극적인 교육정책의 추진이다. 이는 디지털 취약계층에 대한 개념화를 전제로 해야 할 것이다. 물리적 환경 구축과 함께 디지털 문맹률 해소를 위한 지자체 중심의 정책이 요구된다. 현재 디지털 교육은 상당부분 적극적으로 이뤄지고 있지만, 디지털 포용에 포함된 상당수의 교육은 디지털을 통한 인재양성을 목적으로 하고 있다. 이는 한글을 제대로 이해하지 못하는 상태에 있음에도 불구하고 사회가 자격증 준비 교육을 적극적으로 시행하는 것과 같은 형식이라 디지털 격차 해소에는 한계가 있다. 이에, 영국과 같이 지역 도서관 활용이나 행정복지센터를 중심으로 스마트기기 도우미 설치를 고려할 수 있다. 물론 현재 우리도 지자체 차원의 디지털 교육을 시행하고 있기는 하지만, 정형화된 교육이 아니라, 누구라도 필요한 순간에 찾아가 도움을 요청할 수 있는 신뢰할만한 도우미의 설치가 필요하다.

11) <과학기술정보통신부>. '소외 없는 디지털 세상' 밈그림 확정. 디지털포용정책팀. 2020.06.03. <https://www.msit.go.kr/bbs/view.do?Code=user&mId=113&mPid=112&bbsSeqNo=94&nttSeqNo=2937336>

12) 디지털 사회(Société numérique)는 모든 프랑스 국민이 디지털과 관련한 새로운 기술을 익히고 이를 통한 취업 준비와 더불어 디지털 사회에서 시민의 권리와 의무의 공유를 목적으로 하고 있다. 개인이 디지털 사용에 중점을 두고, 누구나 디지털 기술을 익힐 수 있도록 디지털 도우미의 역할을 강조하고 있다.

- 지방정부와 중앙정부의 공동자금을 통한 디지털 패스 배포
- 지방정부의 디지털 전략 구현에 중앙정부의 지방정부 지원
- 디지털 지원 전문가를 위한 디지털기기 및 장치 지원
- 디지털 포용과 관련한 국가정책에 있어 국가적 조정 담당 등

마지막으로 디지털의 사회적 의미 제고를 통한 디지털전환사회에서의 포용적 디지털 활용 방안의 추진 제시이다. 대부분의 디지털 포용과 관련한 사안은 복지 정책적 차원에서 사회적 취약계층을 위한 지원정책으로 추진되고 있다. 즉, 디지털 전환사회에서 디지털 정책은 복지 정책 차원에서의 취약계층을 위한 지원 정책적 차원이 아니라 다양한 정책 차원에서의 포용적 디지털 정책이 요구된다. 즉, 2023년 3월에 종료된 싱가포르의 Digital for Life와 같은 정책에서 스마트국가 정책으로 확대된 디지털 정책으로의 확대의 필요하다. 이는 연령, 성별, 교육, 지역, 그리고 장애 유무와 같은 개인적 그리고 집단적 특징에 따른 다른 디지털 정책을 통한 개인의 역량과 지역사회 디지털 활용역량을 강화에서 시작될 수 있다. 이를 통해 디지털 활용의 가장 큰 장점인 유연성을 국가와 지자체 차원에서 체계적으로 활용할 수 있도록 하는 제도적 구축과 더불어 지역과 공동체 차원에서 디지털 인프라 활용을 가능하게 하는 맞춤형 정책의 추진이 더해진다면, 국가의 경쟁력을 담보할 수 있을 것으로 판단된다. 디지털 포용은 경제적 측면에서의 지속가능한 발전을 위한 시초이자 국가의 경쟁력 강화를 위한 불평등과 사회적 배제를 다루려는 시도로, 개인의 행동을 제한하는 구조를 변경하고 가장 적은 권한과 기회를 가진 행위자가 행동할 수 있도록 역량을 지닐 수 있는 환경을 만드는 일로 고려될 수 있다. 이는 싱가포르나 영국, 호주, 벨기에, 프랑스 등 세계 주요국의 여러 사례를 통해서도 알 수 있다.

UN을 비롯한 국제기구와 상당수의 국가는 각자의 국가 형편에 맞는 디지털 포용정책을 추진하고 있다. 우리도 마찬가지로 여러 가지가 추진되고 있다. 2021년에 발의는 되었지만, 여전히 그 시행 여부가 확실하지 않은 디지털 포용법(안)이나 취업을 목적으로 하는 디지털 포용 추진, 혹은 한국지능정보사회진흥원에 속한 디지털 포용본부의 지능정보화기본법에 근거한 취약계층 디지털역량강화사업 운영지침이나 디지털 포용대상을 사회적 취약계층을 대상으로 한정하고 있는 등 여러 가지 한계가 있다. 이는 디지털 격차로 인해 야기될 수 있는 잠재적인 사회적 문제에 대한 대응 미비로 볼 수 있다. 매년 보고되고 있는 디지털 정보격차 실태조사에서 구분된 디지털 격차를 통해 사회적 취약계층과 관계없이 누구라도 디지털 취약계층이 될 수 있음을 알 수 있다. 이에, 디지털에서 소외된 취약계층 외의 사람들이 경험하는 사회적 배제에 대한 대응의 강화가 요구된다. 이를 위해 사회적 포용관점에서 디지털의 의미를 재고하고, 디지털 취약계층에 대한 논의를 통한 지역 간, 계층 간 인구학적인 특성을 반영해 구체적인 정책적 사안이 요구된다. 앞서 논의된 여러 국가의 정책에서 보듯이, 디지털 포용은 취약계층의 디지털 접근 강화보다는 모두의 디지털 접근 곤란성에 초점을 두어야 잠재적 디지털 포용 사각지대로 인한 사회적 배제 예방과 디지털을 기반으로 발전하고 있는 사회의 지속 가능한 진화를 담보할 수 있을 것이다.

이 연구는 2023년 1월까지의 정책과 이슈를 LDA기반 토픽분석을 통해 우리 사회에서의 디지털과 관련된 사회적 이슈 전반을 객관화 살펴봤다는 점에서 의미가 있다. 그러나 분석내용이 다분히 이슈 중심이고 새로운 정부의 새로운 정책 방향을 설정하는 데 요구되는 구체적인 정책적 함의를 제시하지 못한다는 점에서 한계가 있다. 이에 비교 정책적 관점에서의 전환적 디지털 포용사회와 관련한 연구를 통해 이를 보완할 필요가 있다.

## 참 고 문 헌

- 과학기술정보통신부. (2020). 2021, 디지털 뉴딜 국민 체감 성과 실현의 원년으로!. 한국판 뉴딜, 디지털뉴딜 보도자료.
- \_\_\_\_\_. (2022). 「디지털포용법」 제정 공청회 개최: 디지털 대전환 시대, 디지털포용 사회 구현을 위한 기본법제 마련 시동. 보도자료(1월 12일). 디지털포용정책팀.
- 과학기술정보통신부 · 한국지능정보사회진흥원. (2022). 2022 디지털 정보격차 실태조사. 국가통계 승인번호 제120017호.
- 관계부처합동. (2018). 2018년 국가정보화기본계획 추진실적 점검보고서(안).
- \_\_\_\_\_. (2020a). 혁신적 포용국가 실현을 위한 디지털 포용 추진계획. 과학기술정보통신부 정보통신정책실 정보통신정책관 디지털포용정책팀.
- \_\_\_\_\_. (2020b). 디지털 포용 추진계획: 다함께 누리는 디지털 포용 세상.
- \_\_\_\_\_. (2020c). 지역과 함께하는 「지역균형 뉴딜」 추진방안.
- \_\_\_\_\_. (2020d). 「한국판 뉴딜」 종합계획. 한국판뉴딜 국민보고대회(제7차 비상임경제회의).
- \_\_\_\_\_. (2021a). 내 삶을 바꾸는 2021년 한국판 뉴딜 추진계획.
- \_\_\_\_\_. (2021b). 한국판 뉴딜 2.0: 미래를 만드는 나라 대한민국.
- 기획재정부. (2021). 「10월 한국판 뉴딜 주요사업 추진계획」 배포.
- 김선영. (2020). 증거기반 정책에서의 빅데이터에 관한 연구. 「한국정책학회보」, 29(1): 69-91.
- 노규성. (2020). 포용적 혁신성장과 일자리 창출을 위한 디지털 뉴딜 전략에 관한 연구. 「디지털융복합연구」, 18(1): 23-33.
- 디지털포용법안(의안번호 7422). (2021). 강병원의원 대표발의. 발의자:강병원 · 김민기 · 조승래 · 황희 · 김병욱 · 강훈식 · 홍성국 · 김영호 · 김정만 · 서영석 · 이탄희 · 이광재 · 유동수 · 윤영찬 · 김민석의원(15인).
- 방승기반국. (2022). 방통위, 디지털 미디어 복지 실현을 위한 「22년 소외계층 미디어 포용 세부추진계획」 수립. 한국판 뉴딜, 디지털뉴딜 보도자료.
- 이대영 · 이현숙. (2021). LDA 토픽 모델링의 적정 토픽 수 결정 방법 탐색: 혼잡도와 조화평균법 활용을 중심으로. 「교육평가연구」, 34(1): 1-30.
- 이부하. (2021). 지능정보화 기본법의 문제점과 입법 제안. 「법학연구」, 24(1): 33-59.
- 조기열. (2021). 디지털포용법안 검토보고, 디지털포용을 위한 법적·제도적 기반 마련, 제384회국회(임시회) 제2차 과학기술정보방송통신위원회.
- 이해영 · 김윤희 · 김선영 · 이우철. (2022). 「포용적 디지털 전환을 위한 주요국 정책·법제 비교·분석 연구」. 한국행정학회. 과학기술정보통신부.
- ADIA. (2020). A National Digital Inclusion Roadmap. digitalinclusion.org.au.
- Bailey, L. E., & Nyabola, N. (2021). Digital Equity as an Enabling Platform for Equality and Inclusion. For Peaceful, Just and Inclusive and Societies. Policy Brief.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning*

- Research*, 3(Jan): 993-1022.
- Chuang, J., Manning, C. D., & Heer, J. (2012). Termite: Visualization Techniques for Assessing Textual Topic Models. In Proceedings of the International Working Conference on Advanced Visual Interfaces (pp. 74-77).
- Dickel, S., & Franzen, M. (2015). Digitale Inklusion: Zur sozialen Öffnung des Wissenschaftssystems/Digital Inclusion: The Social Implications of Open Science. *Zeitschrift für Soziologie*, 44(5): 330-347.
- Edler, C. (2020). e-Inclusion-Inklusive-Partizipative Forschung und Entwicklung, User-Centred Design und Empowerment, Orientierungen für einen Ansatz der Forschung und Entwicklung (F&E) gemeinsam mit Menschen mit kognitiven Behinderungen. Doctoral dissertation, Pädagogische Hochschule Ludwigsburg.
- Ehlers, A. et al. (2021). Pakt für Inklusion 2021. *Behindertenpädagogik*, 60(2): 210-214.
- El, C. H., SOON, C., & TAN, D. (2021). State of Digital Inclusion in Singapore. Institute of Policy Studies Lee Kuan Yew School of Public Policy National University of Singapore & Principal Investigator NUS Centre for Trusted Internet and Community.
- Ekbia, Hamid R. (2016). Digital Inclusion and Social Exclusion: The Political Economy of Value in a Networked World. *The Information Society*, 32(3): 165-175.
- European Commission. (2021). Digital Economy and Society Index (DESI) 2021 Thematic chapters.
- European Commission. (2021). Europe's Digital Decade: Commission Sets the Course towards a Digitally Empowered Europe by 2030.
- Fang, M. L., Canham, S. L., Battersby, L., Sixsmith, J., Wada, M., & Sixsmith, A. (2019). Exploring Privilege in the Digital Divide: Implications for Theory, Policy, and Practice. *The Gerontologist*, 59(1): e1-e15.
- Gabryelczyk, Renata. (2020). Has COVID-19 Accelerated Digital Transformations? Initial Lessons Learned for Public Administrations. *Information Systems Management*, 37(4): 303-309.
- Gidley, J., Hampson, G., Wheeler, L., & Bered-Samuel, E. (2010). Social Inclusion: Context, Theory and Practice. *The Australasian Journal of University-Community Engagement*, 5(1): 6-36.
- Goggin, G. (2018). Afterword: Why Digital Inclusion Now?. Author Version of Chapter in Digital Inclusion: Be on the Right Side of the Digital Divide, edited by Massimo Ragnedda and Bruce Mutsaers (Lanham, MD: Lexington, 2018).
- Granata, P. (2020). SDG Digital Inclusion Framework: Advancing UN Global Goals through Equitable Digital Access. the Media Ethics Lab, St. Michael's College, University of Toronto.
- Griffiths, T. L., & Steyvers, M. (2004). Finding Scientific Topics. *Proceedings of the National Academy of Sciences*, 101(suppl\_1): 5228-5235.
- Guterres, A. (2020). Report of the Secretary-General Roadmap for Digital Cooperation.
- Hartnett, M. & Alison, F. (2019). Digital Inclusion in New Zealand. *Journal of Open, Flexible and Distance Learning*, 23(2): 1-4.
- Heisler, D., & Meier, J. (2020). Digitalisierung am Übergang Schule Beruf: Ansätze und Perspektiven in

- Arbeitsdomänen und beruflicher Förderung.
- Helbig, N., Gil-García, J. R., & Ferro, E. (2009). Understanding the Complexity of Electronic Government: Implications from the Digital Divide Literature. *Government Information Quarterly*, 26(1): 89-97.
- Helsper, E. & Alexander van D., (2017). Do the Rich Get Digitally Richer? Quantity and Quality of Support for Digital Engagement. *Information, Communication & Society*, 20(5): 700-714.
- Helsper, E. (2008). Digital Inclusion: An Analysis of Social Disadvantage and the Information Society. Department for Communities and Local Government.
- Hudes, K. (2021). Dimensions of Digital Inclusion: A Review of Research and Practice, Part 2. Prepared for the City of Toronto Research Project. Digital Access: Who is Underserved and Why.
- ITU. (2021). Regional Good Practices: Practices Accelerating Innovation, Entrepreneurship and Digital Transformation in Europe.
- Kronauer, M. (2013). Soziologische Anmerkungen zu zwei Debatten über Inklusion und Exklusion (pp. 17-25). W. Bertelsmann Verlag.
- Leshner, M., Gierten, D., & Attrey, A. (2020). Going Digital Integrated Policy Framework. OECD Digital Economy Papers. No.292.
- Matinmikko-Blue, M., Aalto, S., Asghar, M. I., Berndt, H., Chen, Y., Dixit, S., ... & Ziegler, V. (2020). White paper on 6G drivers and the UN SDGs. arXiv preprint arXiv:2004.14695.
- Michael, L. (2021). Government Urged to Address the Digital Divide. Australia PRObono<https://probonoaustralia.com.au/news/2021/03/government-urged-to-address-the-digital-divide/>
- New Zealand Government. (2019). The Digital Inclusion Blueprint. the Department of Internal Affairs Te Tari Taiwhenua. [digital.govt.nz](https://digital.govt.nz)
- O'Reilly, D. (2005). Social Inclusion: A Philosophical Anthropology. *Politics*, 25(2): 80-88.
- Osimo, D., De Luca, A., & Codagnone, C. (2010). E-Inclusion Initiatives from Private and Non-profit European Entities.
- Parsons, C., & Hick, S. F. (2008). Moving from the Digital Divide to Digital Inclusion. *Currents: Scholarship in the Human Services*, 7(2): 1-16.
- Quaintance, Z. (2018). The Quest for Digital Equity: A Look at the Evolution of the Challenge to Ensure Advances in Technology Bring Benefits to Everyone. <https://www.govtech.com/civic/the-quest-for-digital-equity.html>
- Rawal, N. (2008). Social Inclusion and Exclusion: A Review. *Dhaulagiri Journal of Sociology and Anthropology*, 2: 161-180.
- Reber, C., & Luginbühl, M. (2016). Inklusion ohne digitale Medien ist nicht mehr denkbar. Schweizerische Zeitschrift für Heilpädagogik, 22(4): 13-18.
- Rogers, E. M. (2001). The digital divide. *Convergence*, 7(4): 96-111.
- Schulz, L. (2021). Kultur der Diklusivität: Auf dem Weg zu einer digital-inkluisiven Schulgemeinschaft. *Schule verantworten / führungskultur\_innovation\_autonomie*, (2): 64-71.
- Sievert, C., & Shirley, K. (2014). LDAvis: A Method for Visualizing and Interpreting Topics. In Proceedings

- of the Workshop on Interactive Language Learning, Visualization, and Interfaces (pp. 63-70).
- Sorkin, K. (2022). Smart Thinking on Digital Inclusion from Singapore overnment. Global Government Forum.
- Steyvers, M., & Griffiths, T. (2007). Probabilistic topic models. In Handbook of Latent Semantic Analysis (pp. 439-460). Psychology Press.
- Tilley, S. (2020). In the Name of 'Digital Inclusion': The True Cost of the Automation and Privatization of Australia's Social Security System. *Social lternatives*, 39(1), 28-38.
- UNDP. (2022). Digital Strategy 2022-2025. [digitalstrategy.undp.org](https://digitalstrategy.undp.org)
- UNESCO. (2017). Case study by UNESCO-Pearson Initiative for Literacy: Improved Livelihoods in a Digital World, Education Sector, Education 2030.
- UNESCO. (2018). Designing Inclusive Digital Solutions and Developing Digital Skills. Education Sector, Education 2030.
- United Nations. (2021). Leveraging Digital Technologies for Social Inclusion. United Nations Department of Economic and Social Affairs, Policy Brief 92.
- Vall, R. (2020). RAPPORT D'INFORMATION: au nom de la mission d'information sur la lutte contre l'illectronisme et pour l'inclusion numérique (1) sur la lutte contre l'illectronisme et pour l'inclusion numérique. SÉNAT , SESSION EXTRAORDINAIRE DE 2019-2020, N° 711.
- Van Dijk, J., & Hacker, K. (2003). The Digital Divide as a Complex and Dynamic Phenomenon. *The Information Society*, 19(4): 315-326.
- Velinov, Emil et al. (2020). Digital Skills and Social Inclusion in Adult Education: Evidence from EU and Non-EU Countries. International Multidisciplinary Scientific Conference on Social Sciences and Arts SGEM, 539-546.
- Wagg, S. & Boyka, S. (2021). A Policy-level Perspective to Tackle Rural Digital Inclusion. *Information Technology and People*. (early access). DOI:10.1108/ITP-01-2020-0047.
- Wilson, C. (2020). Digital Inclusion in Australia. *Interaction*, 48(3): 11-15.
- Witten, U. (2021). Barrierefrei im Digitalen?. *Zeitschrift Für Disability Studies*. DOI 10.15203/ZDS\_2021\_1.03.
- World Economic Forum. (2020). Accelerating Digital Inclusion in the New Normal. In collaboration with Boston Consulting Group.
- 〈4차산업혁명위원회〉. <https://www.4th-ir.go.kr/>. (검색일: 2022.01.10.)
- 〈Australian Digital Inclusion Index〉. <https://www.digitalinclusionindex.org.au/>. (Retrieved on Jan, 20, 2022.)
- 〈BigKinds〉. <https://www.bigkinds.or.kr/>. (검색일: 2023.02.10.)
- 〈Bundesministerium für Bildung und Forschung〉. Inklusion durch Digitalisierung. <https://www.qualifizierungdigital.de/qualifizierungdigital/de/projekte/inklusion-durch-digitalisierung>. (Retrieved on Jan, 10, 2022.)
- 〈Digital for Life〉. <https://www.imda.gov.sg/digitalforlife/data-for-all>. (Retrieved on Jan, 10, 2022.)

- 〈EC〉. Shaping Europe's digital future. [https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/shaping-europe-digital-future\\_en](https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/shaping-europe-digital-future_en). (Retrieved on Jan. 20, 2022.)
- 〈Financial Review〉 [소셜리뷰] 사라져가는 ATM기, 노인들은 어찌라고....2019.03.27. <http://www.financialreview.co.kr/news/articleView.html?idxno=16503>. (검색일: 2022.01.20.)
- 〈Good Thongs Foundation〉. Digital Inclusion in Libraries in England. <https://www.goodthingsfoundation.org/insights/digital-inclusion-in-libraries-in-england/>. (검색일: 2023.06.24.)
- 〈Gov.UK〉. Government Digital Inclusion Strategy <https://www.gov.uk/government/publications/government-digital-inclusion-strategy/government-digital-inclusion-strategy>. (Retrieved on Jan. 20, 2022.)
- 〈Infocomm Media Development Authority〉. A Digital Future for All. <https://www.imda.gov.sg/About-IMDA/Who-We-Are/Digital-For-Life> (Retrieved on Nov. 10, 2022.)
- 〈International Telecommunications Union〉. Digital inclusion of all. <https://www.itu.int/en/mediacentre/backgrounders/Pages/digital-inclusion-of-all.aspx>. (Retrieved on Jan. 20, 2022.)
- 〈OHCHR〉. <https://www.ohchr.org/en/search?query=digital%20inclusion> (2023. 06. 10. 검색).
- 〈SÉNAT〉. Lutte contre l'illectronisme et pour l'inclusion numérique. 2021년 4월 14일. [http://www.senat.fr/espace\\_presse/actualites/202104/lutte\\_contre\\_lillelectronisme\\_et\\_pour\\_linclusion\\_numerique.html#:~:text=Mercredi%2014%20avril%202021%2C%20le%20S%C3%A9nat%20a%20examin%C3%A9%20la%20proposition,Rassemblement%20D%C3%A9mocratique%20et%20Social%20Europ%C3%A9en](http://www.senat.fr/espace_presse/actualites/202104/lutte_contre_lillelectronisme_et_pour_linclusion_numerique.html#:~:text=Mercredi%2014%20avril%202021%2C%20le%20S%C3%A9nat%20a%20examin%C3%A9%20la%20proposition,Rassemblement%20D%C3%A9mocratique%20et%20Social%20Europ%C3%A9en). (Retrieved on Feb. 10, 2022.)
- 〈Skills for Digital Tomorrow〉. <https://all-digital.org/#/members-directory>. (Retrieved on July. 10, 2023.)
- 〈Smart Nation Singapore〉. <https://www.smartnation.gov.sg/> (Retrieved on July. 10, 2023.)
- 〈Société numérique〉. <https://societenumerique.gouv.fr/en/>. (검색일: 2022.02.22.)
- 〈UN. Department of Economic and Social Affairs Social Inclusion〉. Leveraging digital technologies for social inclusion. <https://www.un.org/development/desa/dspd/2021/02/digital-technologies-for-social-inclusion/>. (Retrieved on May 11, 2022.)
- 〈UNDP〉. <https://www.undp.org/?search=digital+inclusion> (2023. 02. 10. 검색).
- 〈World Economic Forum〉. How digital communities can build universal digital inclusion. 2021년 12월 14일. <https://www.weforum.org/agenda/2021/12/digital-communities-digital-divide/>. (Retrieved on Nov. 13, 2022.)
- 〈World Economic Forum〉. The first alliance to accelerate digital inclusion. 2022년 1월 17일. <https://www.weforum.org/impact/tackling-digital-deserts-the-first-cross-sector-alliance-to-close-the-digital-gap-launches-at-the-davos-agenda>. (Retrieved on Nov. 13, 2022.)
- 〈World Economic Forum〉. These are the invisible barriers to tackling the digital divide. 2021년 11월 10일. <https://www.weforum.org/agenda/2021/11/costs-literacy-design-invisible-barriers-tackling-digital-divide>. (Retrieved on Nov. 13, 2022.)
- 〈과학기술정보통신부〉. '소외없는 디지털 세상' 밑그림 확정'소외없는 디지털 세상'밑그림 확정- 정보통신전략위 '디지털 포용 추진계획'의결 -. 2022년6월23일. <https://www.msit.go.kr/bbs/view.do?sCode=user&mId=113&mPid=112&bbsSeqNo=94&nttSeqNo=2937336> (검색일: 2022.06.26.)



- 〈대한민국 정책브리핑〉. '22년 디지털 뉴딜 실행계획 사전 브리핑. <https://www.korea.kr/news/policyBriefingView.do?newsId=156493220>. (검색일: 2022.01.20.)
- 〈디지털 뉴딜〉. 디지털 뉴딜 2020-2025: 국민이 체감하는 디지털 뉴딜 대표사업 성과 로드맵. <https://digital.go.kr/front/main/main.do>. (검색일: 2022.01.20.)
- 〈디지털플랫폼정부위원회〉. 기본원칙. <https://dpg.go.kr/document/rule>
- 〈연합뉴스〉. WHO, 코로나19에 팬데믹 선언...통제될 수 있다(종합). 2020년 3월 12일. <https://www.yna.co.kr/view/AKR20200312002152088>. (검색일: 2021.12.28.)
- 〈한국지능정보사회진흥원〉. 디지털 포용본부. [https://www.nia.or.kr/site/nia\\_kor/ex/bbs/ListBusiness.do?businessMnCd=23000400](https://www.nia.or.kr/site/nia_kor/ex/bbs/ListBusiness.do?businessMnCd=23000400). (검색일: 2022.01.27.)
- 〈혁신적포용국가〉. 포용국가 사회정책 추진계획. <http://www.inclusivekorea.go.kr/info3000.jsp>. (검색일: 2022.02.22.)

○ 김선영(金善映): Kiel대학교 사회학과에서 사회정책학 박사학위를 취득하고, 현재 공주대학교 정책융합전문대학원과 영남대학교 새마을국제개발학과에서 강의하고 있다. 주요 관심분야는 포용적 디지털 전환사회와 지능정보사회에서의 과학 및 노동사회정책, 디지털 ODA이다. 최근 논문으로는 LDA 토픽모델링을 활용한 ChatGPT의 사회경제적 이슈 분석: 지역사회 갈등조정과 민원배심제-대구 수성구민원배심제 운영의 경험 등이 있다(sylikelily@hanmail.net).

투 고 일 자: 2023년 8월 20일

심 사 일 자: 2023년 8월 23일(초심), 2023년 9월 21일(재심)

게재확정일자: 2023년 9월 21일

of planning, application, screening and selection, payment, and case management. Nevertheless they need also prepare for a plan to derive an integrated policy direction that respects autonomy and decentralization in terms of policy establishment and evaluation participation. Fourth, both provincial and local governments establish and execute budgets at a matching ratio of 7:3, and local tax revenues are the main source of the project. The project requires a flexible application according to the financial independence of each city and county as well as an overall review of tax imposition standards to secure stable financial resources. In conclusion, it is difficult to give a positive evaluation on the policy of the youth basic income project in Gyeonggi Province, but it can be meaningful that the welfare paradigm has changed with the introduction of a new universal welfare policy called basic income for youth.

**Key Words:** Youth Basic Income, Universal Welfare, Gilbert & Terrell's Analysis Framework, Gyeonggi Province

## Discovering Policy Implications from Analysis of News Big-data Related to Digital Issues Based on LDA Topic-modeling

..... Kim, Sun Young

This study aimed to discover the implications of digital inclusion applied to our social & policy environment by reviewing the literature on the issues being discussed in relation to digital transformation society and social inclusion to explore issues, focused on policies about digital inclusion in our society. And then, issues related to digital inclusion in our society were explored with the news big data analysis. For the big data analysis, LDA topic modeling was conducted on 5,136 news data related to digital between Jan. 01, 2022, and Jan. 31, 2023. The collected data was from BigKinds. As a result of the analysis, it was described that digital inclusion in our society is related to digital issues from an economic perspective rather than any other perspective. The social issues seemed to affect the digital inclusion policies. This study implies the need for digital inclusion following; 1) applying to all members of our society from a specific group expressed as the socially disadvantaged class, 2) the need to conceptualize digital inclusion as distinct from social inclusion, and 3) to prescribe the social meaning of digital inclusion in the digital inclusion policy.

**Key Words:** Digital Divide, Inclusive Digital Transformation Society, Digital Inclusion, LDA Topic Modeling