

코로나 백신으로 본 소셜미디어의 정치 양극화: 트위터 메시지의 단어 네트워크 분석과 토픽 모델링을 중심으로*

정 동 준†

인하대학교

2020년 이후 전 세계를 휩쓸고 있는 코로나는, 전염병이란 비정치적 이슈에도 불구하고 많은 나라에서 이념과 당파에 따라 정치화되고 있다. 이러한 정치화는 최근 민주주의 위기의 원인으로 거론되는 정치 양극화 현상을 부추기고 있다. 본 논문은 한국의 코로나 백신에 대한 논의가, 특히 오늘날 정치적 담론 확산의 주요 매개체인 소셜미디어에서 이념과 당파 집단에 따라 어떻게 양극화되고 있는지를 살펴본다. 코로나 백신에 대한 트위터 메시지를 단어 네트워크 중심성 분석과 토픽 모델링을 통해 살펴본 결과, 진보와 친여당 집단에서는 백신 관련 뉴스와 접종 예약에 대한 정보가 주로 트윗되고 있는 것으로 나타났다. 반면 보수와 반여당 집단에서는 백신 부작용과 혈전 문제가 주된 논의 주제로 밝혀졌고, 특히 반여당 집단에서는 정부의 백신 정책과 역할에 대한 논의도 활발한 것으로 나타났다. 트위터 상의 논의가 곧 여론이라 할 수는 없으나 이것이 곧 다른 온라인 미디어로 확산되고 오프라인의 여론과 정치적 담론 형성에도 기여한다는 점에서 이 같은 양극화 경향은 주의를 요한다.

주요어: 코로나 백신, 소셜미디어, 정치 양극화, 텍스트 네트워크 분석, 토픽 모델링

* 이 논문은 2020년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2020S1A5A8045323).

† 단독저자 : 정동준, 인하대학교 사회교육과 조교수, 인천광역시 미추홀구 인하로 100, E-mail : djjung@inha.ac.kr

■ 최초투고일 : 2021년 11월 26일 ■ 심사마감일 : 2022년 3월 30일 ■ 게재확정일 : 2022년 4월 15일

1. 서론

2020년 이후 우리의 삶을 가장 많이 바꾼 사건을 들라 하면 무엇보다도 코로나-19 바이러스(이하 코로나)의 확산을 들 수 있을 것이다. 2021년 들어 백신의 개발과 접종이 시작되면서 이 사태도 곧 마무리될 거란 희망이 생겨났지만, 델타와 람다 등 각종 변이 바이러스의 등장과 백신 효과에 대한 논란까지 가중되면서 2021년이 마무리되어가는 현재까지도 코로나 팬데믹(pandemic)은 좀처럼 끝날 기미를 보이지 않고 있다. 이러한 가운데 코로나에 대한 대응을 둘러싸고 집단마다 의견 차이를 보이고 있다. 코로나의 위험성 정도, 사회적 거리두기와 마스크 강제 착용과 같은 정부의 정책, 그리고 최근에는 백신 접종까지 코로나를 둘러싼 많은 이슈에 대해 각자 가진 이념과 당파적 성향에 따라 그 입장이 나뉘고 있다. 본질적으로 비정치적(apolitical) 이슈라 할 수 있는 코로나가 정치화(politicization)되고 있는 것이다(Druckman et al., 2020; Gilardi et al., 2021; Hart et al., 2020).

그렇다면 한국 사회에서 코로나 이슈는 어떻게 정치화되고 있는가? 정치 양극화는 오늘날 정치 학계에서 뜨겁게 논의되고 있는 주제 중 하나이다. 크게 ‘쟁점에 대한 입장’과 ‘상대 집단에 대한 감정’의 영역에서 관찰되는 정치 양극화는, 서로 다른 이념적 혹은 당파적 집단 간에 이러한 입장과 감정의 골이 깊어지는 것이다(장승진, 서정규, 2019; 정동준, 2018; Abramowitz & Webster, 2016; Bafumi & Shapiro, 2009; Iyengar et al., 2012; Mason, 2015). 특히 시민 차원에서 나타나는 정치 양극화의 원인으로는 많은 대내외적 요인이 있지만, 그 중에서 오늘날의 변화된 디지털 환경을 빼놓을 수 없다(김기동, 이재목, 2021; 장승진, 서정규, 2019; Iandoli et al., 2021). 특히

인터넷의 발달과 소셜미디어의 등장으로 자신들과 비슷한 입장을 가진 사람들과 활발하게 교류함에 따라 자신과 다른 의견은 차단하고 비슷한 의견만을 공유하는 ‘선택적 노출’과 ‘필터링’ 현상이 나타나고 있다(Asher et al., 2019; Bessi et al., 2016; Bozdog, 2020; Freden et al., 2020; Hameleers, 2020). 여기에 소셜미디어 특유의 ‘필터 버블’(filter bubble) 알고리즘까지 더해짐으로써 내 집단의 응집력은 강화되고 외집단과의 거리는 멀어지는 양극화가 일어나게 된 것이다(Pariser, 2011).

이렇듯 요즘 정치학계의 주요 화두는 크게 코로나, 정치 양극화, 그리고 디지털화(digitalization)로 압축해 볼 수 있다. 본 논문은 이러한 화두에 맞추어 오늘날 한국의 시민들이 코로나, 그 중에서도 가장 최근 이슈인 코로나 백신과 관련한 시민들의 의견이 온라인 상에서 어떻게 양극화되어 있는지를 살펴보고자 한다. 특히 본 논문은 대표적인 소셜미디어 플랫폼 중 하나인 트위터(Twitter) 공간에 주목한다. 트위터는 특히 극단적이고 포퓰리스트적인 견해가 활발하게 표출되고 공유되는 소셜미디어로(Tien et al., 2020), 트위터 상의 논의를 곧 시민 전체의 여론으로 볼 수는 없다. 하지만 트위터에서 인기를 얻은 게시물이 곧 다른 미디어에도 확산되고, 이러한 포퓰리스트적 담론이 주류 언론의 기사와 여론의 형성에도 영향을 미친다는 점에서(Hameleers, 2020; Mudde, 2016; Tufekci, 2014) 면밀한 관찰이 요구된다. 본 논문은 이념과 당파 집단 간에 트위터 메시지의 단어 네트워크(word network)와 세부 주제(topic)가 어떻게 다르게 나타나는지를 중심성 분석과 토픽 모델링(topic modeling)을 통해 살펴보고자 한다. 이를 통해 온라인 정치 공간에서 한국 시민들의 양극화에 대한 현주소를 진단해보고자 한다.

이를 통해 본 논문이 기여하고자 하는 바는 다음과 같다. 첫째, 코로나 백신에 대한 시민들의 논

의가 집단별로 어떻게 다르게 나타나는지를 살펴봄으로써 한국 사회에서 코로나 백신 이슈가 어떻게 정치화되어 나타나는지에 대한 실증 분석을 제공하고자 한다. 코로나 백신이 서로 다른 정치적 집단에서 어떻게 논의되고 있는지는 아직 많은 분석이 이루어지지 않았고, 특히 이에 대한 국내 연구는 거의 찾아보기 어렵다. 본 분석은 코로나의 정치화에 대한 연구에 이론적, 실증적 기여를 할 수 있을 것이다. 둘째, 트위터 메시지를 통해 소셜미디어의 정치적 구조와 특성을 탐색해보고자 한다. 트위터와 같은 소셜미디어는 최근 학계의 주목을 받으며 활발히 연구되고 있으나 소셜미디어 콘텐츠 자체를 이념적·당파적 집단에 따라 비교하는 국내 연구는 아직까지 많지 않다. 마지막으로, 이러한 분석을 통하여 한국 시민들의 정치 양극화가 현재 어떠한 양상으로 나타나고 있는지를 파악하고자 한다. 코로나 백신이 이슈에 대해 서로 다른 정치 집단이 얼마나 분극화된(polarized) 목소리를 내고 있는지를 살펴봄으로써 정치 양극화의 현주소를 진단하고자 한다.

2. 이론적 논의

1) 정치 양극화

정치 양극화는 최근 국내외 정치학계에서 가장 많이 논의되는 주제 중 하나이다. 정치 양극화는 1960년대 이후 미국을 비롯한 서구 사회에서 시민권 운동을 통해 등장한 여러 이슈들이 진보와 보수를 가르기 시작하면서 나타났다. 대공황과 2차대전을 겪으며 정부의 역할과 복지제도를 강화하는 쪽으로 진보와 보수 간에 한시적 합의가 이

루어졌던 ‘진보적 합의’(liberal consensus) 시기를 지나, 인종, 성(性), 환경 등 새롭게 나타난 쟁점들을 중심으로 이념적 양극화가 심화되기 시작한 것이다(Bafumi & Shapiro, 2009).

이러한 이념적 양극화는 특히 의회와 정당 정치인을 중심으로 한 엘리트 차원에서 두드러지게 나타나고 있다. 미국 의회에서 민주당원과 공화당원 사이의 이념적 거리가 해를 거듭할수록 멀어지고 있으며(Davis et al., 2014; Iyengar et al., 2012; Jones, 2010), 민주주의 역사가 오래된 서유럽에서도 주류정당의 극우화와 포퓰리스트 정당의 등장으로 정치 양극화 경향이 심해지고 있다(Cole, 2005; Grzymala-Busse, 2019; Mudde, 2016, 2019). 이 뿐 아니라 한 때 공고화된 민주주의를 이루었다고 평가받았던 동유럽 민주주의 국가들 역시 2008년 서유럽 금융위기 이후 시작된 집권당의 극우화로 ‘탈공고화’(deconsolidation) 현상이 나타나고 있다(정동준, 2018; Levitsky & Ziblatt, 2018). 우리나라에서도 많은 학자들이 국회를 위시한 엘리트 차원의 양극화가 점차 심해지고 있음을 지적하고 있다(가상준, 2016; 강원택, 2012; 정동준, 2017).

이렇듯 엘리트 차원의 정치 양극화는 어느 정도 학계의 합의가 이루어진 반면 시민 차원에서는 그렇지 못하다(Kevins & Soroka, 2018; Levendusky, 2009). 시민 차원의 이념적 양극화란 곧 중도층의 감소를 의미하는데 이에 대한 실증적 증거는 아직 부족한 편이다(장승진, 서정규, 2019; 장승진, 장한일, 2020). 따라서 일부 학자들은 시민 전체가 아닌 ‘당파적 지지자’만을 대상으로 하여 이들 사이의 이념적 거리의 증가 및 당파적 배열(partisan sorting)¹⁾로 오늘날의 정치 양극화를 설명하기도 한다(길정아, 하상웅, 2019; 정동준, 2018; 장승진,

1) 당파성과 이념, 쟁점 입장 간 상관관계가 높아지는 현상을 일컫는다.

서정규, 2019; Dinkelberg et al., 2021; Fiorina, 2017). 미국 유권자를 대상으로 한 최근 연구에서 던켈버그 외(Dinkelberg et al., 2021)는 공화당과 민주당 지지자들 사이의 이슈에 대한 이념적 입장이 특히 2016년과 2020년 사이에 멀어졌다고 주장한다. 즉, 일반 시민 차원의 양극화는 아직 관찰되지 않더라도 당파적 지지자 사이의 이념적 선호와 쟁점 입장이 과거에 비해 일관되게 나타남으로써 전반적으로 양극화되고 있다는 것이다.

하지만 이러한 주장도 오늘날 정당의 역할과 당파성의 규모가 과거에 비해 줄어든 것을 고려하면 다소 설득력이 떨어진다. 일반 시민들의 경제 수준과 교육수준이 높아지고 미디어 기술의 발달로 정당을 우회하여 정치에 참여할 수 있는 채널이 다변화됨에 따라 정당의 정치적 동원 기능이 줄어들었고, 이로 인해 당파적 지지자의 규모 역시 줄어든 것이다(Abramowitz & Webster, 2016; Dalton, 2007). 최근 학자들은 이러한 논리적 빈 곳을 메우기 위해 시민들 사이의 ‘정서적 양극화’(affective polarization)에 주목한다(김기동, 이재목, 2021; 장승진, 서정규, 2019; 장승진, 장한일, 2020; Hartevelde, 2021; Iyengar et al., 2019; Mason, 2015). 정서적 양극화란 자신이 “지지하는 정당과 지지하지 않는 정당에 대해서 느끼는 정서의 간극이 과거에 비해서 커”짐을 의미한다(장승진, 서정규, 2019, 8쪽). 시민들이 이념 성향이나 쟁점 입장이 분열되어 있지 않더라도, 감정적으로는 분열될 수 있는 것이다(Van Hiel et al., 2021). 특히 앞서 언급한 당파성의 감소와 함께 보자면 특정 정당에 대한 호감과 충성심은 과거에 비해 낮아졌더라도, 반대하는 정당에 대한 반감이 증가했을 수 있다(Abramowitz & Webster, 2016; Medeiros & Noël, 2014; Weber, 2021). 이러한 ‘부정적 당파성’(negative partisanship)의 증가 역시 오늘날 시민 차원의 정치 양극화를

낳는 하나의 현상이 될 수 있다(정동준, 2018).

2) 소셜미디어 분석

그렇다면 이렇게 시민들의 정서적 양극화와 부정적 당파성을 부추기는 요인은 무엇일까? 최근 여러 행위자 중심 연구들은 주류 정당의 양극화와 일련의 사회·문화·경제적 이슈를 중심으로 나타난 포퓰리스트 극단주의 정당의 출현과 같은 위로부터의 요인이나, 사회적 정체성을 공유하는 내집단의 응집력 강화와 기성정치에 대한 실망으로 나타난 포퓰리즘과 같은 아래로부터의 요인에 주목하고 있다(김기동, 이재목, 2021; Bracciale et al., 2021; Hartevelde et al., 2021; Roberts, 2021; Van Hiel et al., 2021). 이에 더하여 구조 중심적 요인으로는 경제적 위기와 사회경제적 양극화, 난민 문제와 EU 지역주의 등 여러 구조적 요인들이 거론되고 있는데, 그 중에서도 최근 학계의 주목을 받고 있는 것은 오늘날의 변화된 미디어 환경이다. 4차 산업혁명으로 일컬어지는 디지털 기술의 발달은 모든 분야에 거대한 변환을 가져오고 있고 정치 분야 또한 예외가 아니다. 디지털 기술이 정치 영역에 적용됨에 따라 디지털 시대의 정치에 대한 학계의 관심 또한 높아지고 있다(Karpp, 2020). 디지털 미디어를 통한 선거 운동, 디지털 플랫폼을 기반으로 한 정당의 조직과 운영, 그리고 전자정부와 전자가버넌스 등 디지털 기술은 위로부터의 정치를 바꾸고 있다. 아래로부터의 정치 또한 디지털 기술의 발달로 정치 참여에 들어가는 시간적, 경제적 비용이 획기적으로 줄어들어 시민들의 정치 참여가 확대되고 있다(Nelson et al., 2017). 즉, 온라인 공간에서 정치적 표현과 토론이 활성화되면서 시민들의 정치에 대한 관심과 지식, 그리고 정치효능감이 높아짐에 따라 정치 참여가 확대되고 있는 것

이다(Buchi & Vogler, 2017; Lee et al., 2018; Zhu et al., 2019).

하지만 이러한 디지털 미디어의 발전이 순기능만 갖는 것은 아니다. 정당과 정치인들은 홈페이지와 소셜미디어를 통해 편향된 메시지를 전달함으로써 유권자들의 정치행태를 양극화시키고 있고(장우영, 2019; Bracciale et al., 2021; Hameleers et al., 2020; Jiang et al., 2020), 정파적 미디어(partisan media) 역시 자신의 이념과 당파적 입장에 따라 기사를 생산하고 프레임링(framing)함으로써 시민들의 정서적 양극화에 부추기고 있다(박지영, 2020; Lau et al., 2017; Prasetya & Murata, 2020). 시민들 또한 페이스북, 트위터와 같은 소셜미디어를 통해 자신과 이념적, 당파적 성향이 비슷한 사람들과 교류하고 선택적으로 정보에 노출됨(selective exposure)으로써 양극화에 편승하고 있다. 즉, 비슷한 성향의 집단끼리 교류하고 그렇지 않은 집단과는 분리됨으로써 집단의 동질성이 강화되고(Asher et al., 2019; Bessi et al., 2012; Bozdog, 2020; Freden et al., 2020), 상대에 대해 공격적이고 포퓰리스트적인 담론이 재생산됨으로써 집단 간의 이념적, 정서적 양극화가 심화되고 있는 것이다(Buder et al., 2021; Hameleers, 2020; Lee, 2018; Marchal, 2021; Paz et al., 2021). 이에 더하여 내가 소비하는 정보와 관련성이 높은 정보만이 계속해서 제공되는 소셜미디어 특유의 알고리즘으로 이러한 경향은 가속되고 있다(De Arruda et al., 2021; Pariser, 2011).²⁾

특히 트위터는 수많은 소셜미디어 중에서도 가장

활발히 연구되고 있는 플랫폼이다(Iandoli et al., 2021; Kubin & Von Sikorski, 2021; Yarchi et al., 2020). 트위터가 학계에서 널리 사용되는 이유는 다른 미디어처럼 수락에 기반한 친구맺기가 아니라 자유로운 팔로잉 관계를 통해 정보와 의견의 공유가 활발히 일어나기 때문이다. 이 같은 점 때문에 트위터는 특히 정치적 토론에 적합하다(Gruzd & Roy, 2014, p. 29). 또한 팔로잉과 답변(reply), 리트윗 등으로 유저 사이의 메시지 사이에 네트워크가 형성됨으로써 그 구조를 분석하기 용이하다는 장점도 있다(Gruzd & Roy, 2014; Iandoli et al., 2021; Yarchi et al., 2020). 연구자가 개방 API(application program interface)를 통해 손쉽게 대량의 데이터를 얻을 수 있다는 점 또한 유용하다. 물론 트위터가 극단주의적 성향을 가진 사람들이 활발하게 활동하는 소셜미디어란 점에서(Tien et al., 2020) 트위터 상의 논의를 곧 여론으로 볼 수는 없다. 하지만 트위터를 통해 공유되는 정보와 메시지가 공공의 담론 형성에 중요한 영향을 미친다는 점에서(Hameleers, 2020; Mudde, 2016; Tufekci, 2014) 오늘날의 정치 양극화를 이해하는 데 중요하다.

트위터에 대한 기존 연구는 크게 네트워크 구조에 대한 연구와 트위터 메시지의 내용에 대한 연구로 나누어 볼 수 있다. 먼저 네트워크 구조는 계정들이 서로 어떠한 팔로잉 관계로 연결되어 있는지(한규섭 외, 2013; Urman, 2019), 혹은 메시지가 서로 어떠한 답변과 리트윗 관계로 공유되고 있는지(이종임 외, 2019; Asher et al., 2019; Gruzd & Roy, 2014; Tien et al., 2020)로 분석

2) 물론 온라인 상의 교류가 다른 집단에 속한 사람과 만남의 기회를 높이고 다양한 정보에 우발적으로 노출시킴으로써(incident exposure) 오히려 양극화를 완화시킨다거나(Beam et al., 2018), 소셜미디어 사용이 반드시 양극화로 이어지는 것은 아니고(Feezella et al., 2021; Yang et al., 2017) 오히려 양극화가 소셜미디어 사용을 약기한다는(Nordbrandt, 2021) 연구결과도 있다. 하지만 소셜미디어와 정치 양극화에 대한 연구를 분석한 최근의 메타 연구에 의하면 소셜미디어가 정치 양극화를 약기한다는 데에 대부분의 연구가 합의를 이루고 있다(Iandoli et al., 2021; Kubin & Von Sikorski, 2021).

된다. 트위터 메시지에 대한 연구는 특정 이슈가 트위터 상에서 어떤 주제로 논의되고 있는지(노예영, 김기환, 2020; 설선미, 배정아, 2018; 정효정 외, 2016; Hameleers, 2020), 그러한 논의가 집단 간에 어떻게 다르게 나타나는지(장우영, 2013; Jiang et al., 2020; Paz et al., 2021), 그리고 메시지의 스타일이 포퓰리스트적인지 혹은 공격적인지(Bracciiale et al., 2021; Buder et al., 2021; Hameleers et al., 2020) 등을 중심으로 이루어지고 있다.

국내 학계에서도 트위터에 대한 연구가 비교적 활발히 진행되고 있기는 하지만, 정치학에서는 정치인의 트위터가 유권자의 투표행동과 선거결과에 미치는 영향에 대한 연구가 주를 이루어왔다. 장우영(2019)은 20대 총선에서 후보의 트위터 캠페인 참여가 당파적 지지자를 동원하고 득표율과 당선 확률을 높이는 데 기여하였음을 발견하였다³⁾ 남현주와 김하나(2014)는 18대 대통령선거 후보자들의 트위터 활용 방식, 목적, 그리고 내용을 비교 분석하였고, 이소영(2012) 또한 19대 총선 후보자들의 트위터를 통한 선거운동이 해당 후보의 득표율 및 유권자들의 투표율 상승에 기여하였음을 밝혀냈다.

이렇듯 국내 정치학계의 트위터 연구는 주로 선거운동에서 트위터의 영향력에 대한 연구가 주를 이뤘고, 정치적 이슈에 대한 트위터의 메시지나 네트워크 구조에 대한 분석은 상대적으로 많이 수행되지 않고 있다. 설선미와 배정아(2018)가 일련의 정책 이슈에 대한 시민들의 트위터를 분석하여 지방정치 이슈보다는 중앙정치 이슈가 네트워크의 중심을 이루고 있음을 발견하였고, 정효정 외(2016)는 ‘세월호 사건’을 키워드로 하여 트위터 이용자들의 네트워크와 메시지의 내용이 이념

적 집단 간에 어떻게 다르게 나타나는지를 분석하였다. 곽수정, 김현희(2019)와 이종임 외(2019)는 최근의 ‘미투 운동’에 대한 트위터 네트워크 구조와 메시지 토픽 분석을 통해 주요 정치 이슈가 소셜미디어에서 어떠한 구조와 내용으로 논의되고 있는지를 탐색하였다. 이처럼 최근 들어 몇몇 정치 이슈를 중심으로 트위터의 구조와 내용에 관한 연구가 진행되었지만 아직까지 그 수가 많다고 볼 수는 없다. 특히 시민 차원의 정치 양극화와 관련하여 이념과 당파에 따라 트위터 상의 논의가 어떻게 다르게 나타나는지에 대한 연구는 더욱 적다고 할 수 있다.

3) 코로나 이슈의 정치화

이러한 부족을 보완하고자 본 논문은 최근 가장 중요한 화두인 ‘코로나 백신’을 중심으로 트위터 상의 논의가 이념과 당파에 따라 어떻게 다르게 나타나는지를 분석하고자 한다. 코로나란 전염병 자체는 본질적으로 비정치적 이슈이지만 이에 대한 반응과 대처는 정치화되어 나타나고 있다. 이러한 코로나 이슈의 정치화에는 신문, 뉴스 등 정파적 미디어의 보도와 함께(박주현, 2020; Hart et al., 2020) 소셜미디어의 담론이 큰 영향을 미치는 것으로 분석되고 있다(Gilardi et al., 2021). 드럭맨 외(Druckman et al., 2020)는 실험 설문을 통해 시민들의 정서적 양극화가 코로나와 같은 비정치적 이슈에 대한 주관적 신념에 영향을 미침으로써 이를 정치화하고 있음을 밝혀낸 바 있다.

많은 기존 연구가 이념과 당파에 따라 코로나에 대한 반응이 다르게 나타남을 밝히고 있다. 전반적으로 진보 집단은 코로나를 보다 위협적으로 느껴 마스크 착용, 이동 제한, 그리고 백신 접종과

3) 다른 저자들과 함께 한 이전의 연구들에서도 비슷한 결과를 도출하였다(장우영, 이관률, 2017; 장우영김석주, 2014).

같은 정부의 보건 정책을 잘 따르는 반면, 보수 집단에서는 이러한 위협에 덜 민감하고 음모론과 연관 지음으로써 정부 정책에 부정적 반응을 보이는 것으로 나타났다(Calvillo et al., 2020; Clarke et al., 2021; Kerr et al., 2021; Mallinas et al., 2021; Pennycook et al., 2021; Weil & Wolfe, 2021).

코로나의 정치적 효과에 대한 국내 연구는 주로 유권자의 투표 선택과 정부 평가와 관련하여 진행되고 있다. 특히 코로나가 본격적인 정치 이슈로 자리잡은 21대 총선 결과 분석을 통해, 코로나에 대한 정부의 효과적 대응이 여당인 더불어민주당의 선거 승리에 주요한 영향을 미쳤다는 주장이 다수 제기되었다(길정아, 강원택, 2020; 송진미, 2020; 신정섭, 2020). 사례를 한국과 일본, 대만으로 확대한 박범섭과 신정섭(2021)의 최근 연구에서는 자국 정부의 코로나 대응을 다른 국가에 비해 긍정적으로 평가할수록 정부 신뢰, 민주주의 만족도, 국가자긍심이 높아짐을 발견하였다. 박선경과 신진욱(2021)은 정부의 코로나 정책을 긍정적으로 평가할수록 정부 역할의 확대를 지지하는 경향을 밝히기도 하였다. 김한나 외(2021)도 정부의 코로나 대응 정책에 대한 긍정적 평가가 백신 접종 의사를 높인다고 주장하였다.

이처럼 코로나와 관련된 국내 연구는 아직까지 투표 선택과 정부 평가에 집중되어 있고, 이념과 당파에 따른 정치화에 대한 연구는 드물다. 코로나 음모론을 믿는 사람일수록 백신 접종 의사가 약화된다는 김한나 외(2021)의 최근 연구도 있으나, 이 역시 정치 성향에 따른 코로나 백신의 인식 차이를 직접적으로 다루고 있지는 않다. 본 연구는 이 같은 공백을 매우기 위해 오늘날 시민 차원의 정치 양극화를 심화시키는 매개의 하나로서 소셜미디어 상의 정치 담론을 코로나 백신을 중심으로 분석해보겠다. 보다 구체적으로 코로나 백

신과 관련한 메시지를 트윗 혹은 답변(reply)한 계정들을 이념적, 당파적 입장에 따라 분류한 후, 각 집단 내에서 해당 메시지들이 어떠한 단어 구조와 토픽을 중심으로 논의되고 있는지를 비교·분석해보고자 한다.

3. 자료와 방법

코로나 백신을 둘러싼 트위터 메시지 분석을 위해 본 연구가 사용할 방법은 텍스트 네트워크 분석(text network analysis)이다. 텍스트 네트워크 분석은 사회연결망 분석(social network analysis)의 일종으로, 텍스트를 단어들 간의 네트워크로 보고 이러한 단어들의 출현 빈도, 위치, 거리, 연결 패턴 등의 형식적 요인을 통하여 텍스트의 구조와 내용을 분석하는 방법이다(Hunter, 2014). 텍스트 네트워크 분석은 크게 인접한 단어 간에 네트워크를 구성하여 그 구조를 분석하는 ‘구조중심적(structure-based) 분석’과 네트워크의 내용과 의미를 분석하는 ‘내용중심적(content-based) 분석’으로 나누어 볼 수 있다(Sebei et al., 2018). 본 연구에서는 두 가지 방식을 모두 사용하고자 한다. 먼저 트위터 메시지 상 인접한 단어들 간에 네트워크를 구성한 후, 해당 네트워크의 중심성(centrality) 분석을 통해 메시지의 핵심을 이루는 키워드가 무엇인지 파악한다. 다음으로 토픽 모델 분석을 통해 어떠한 코로나 백신 관련 토픽들이 트위터 논의의 주를 이루었는지 살펴보도록 하겠다. 두 분석 모두, 이념과 당파에 따라 트위터 계정을 분류한 후 각 집단별로 분석을 진행하겠다.

1) 자료 수집 및 전처리

분석 자료는 트위터 측에서 제공하는 개방 AP

I4)를 통해 수집하였다.⁵⁾ 자료 수집 및 전처리와 분석에는 NetMiner 프로그램(ver. 4.4.3.g)을 사용하였다. 수집 기간은 문재인 대통령이 방미를 통해 바이든(Joe Biden) 미(美)대통령과 한미 백신 파트너십 구축에 협의(한국시간 2021년 5월 22일 오전)한 이후 일주일 기간인 5월 21일부터 28일까지로, 백신을 둘러싼 논의가 활발히 일어나 시기 중 하나였다.⁶⁾ 특히 문대통령의 성과를 놓고 이념과 당과 간에 서로 다른 의견이 활발히 오갔다는 점에서 트위터의 양극화 현상을 관찰하려는 본 분석에 적절한 시기라 할 수 있다.⁷⁾ 분석 대상 트위터 메시지를 추리기 위해 ‘백신, 코로나 백신, 화이자, 아스트라제네카, 아스트라, 모더나, 안센, 접종’ 중 하나가 포함된 메시지를 수집하였고, 다른 형성에 보다 큰 영향력을 미칠 것으로 여겨지

는 팔로워 1,000명 이상의 계정에서 작성된 메시지만을 분석 대상으로 한정하였다. 이렇게 수집된 트위터 메시지는(최초로 생성된 트윗과 이 트윗에 대한 답변을 포함)⁸⁾ 총 11,198개이고, 여기에서 NetMiner의 자동 형태소분석 프로그램을 통해 명사만을 추출하였다.⁹⁾ 이렇게 하여 1차로 추출된 단어(명사)의 개수는 총 9,214개이다. 네트워크 분석에서 보다 직관적이고 의미있는 분석을 하기 위해서는 노드(node), 즉 단어와 링크(link)의 개수를 적절하게 제한할 필요가 있다. 이에 따라 수집된 단어 중 전체 문서(즉, 트위터 메시지)에서 최소 50번 이상 등장한 단어만을 대상으로 삼았다. 검색 키워드 자신, 조사와 같이 특별한 뜻이 없거나 분절되어 잘못 수집된 단어들 역시 분석에서 제외하였다.¹⁰⁾ 이러한 과정을 통해 최종

- 4) 트위터에서 제공하는 ‘Standard Search API’를 이용하여 수집하였다(<https://developer.twitter.com/en/docs/twitter-api/v1/tweets/search/api-reference/get-search-tweets>).
- 5) 개방 API의 사용에도 물론 단점은 존재한다. 데이터의 제공이 일방적으로 이루어져 트위터 측의 샘플링 방식에 따라 추출된 표본에 바이어스가 존재할 수 있기 때문이다(Tien et al., 2020). 하지만 API사용은 학계에서 널리 통용되는 방식이고, 그만큼 방대한 자료를 연구자 개인이 직접 구하기란 거의 불가능하다는 점에서 이는 어쩔 수 없는 한계로 여겨진다.
- 6) 코로나 백신 상황이 시시각각 변하여 논의의 대상과 주제가 시기마다 달라졌기에 수집기간을 일주일이라는 비교적 짧은 시기로 한정하였다. 실제로 같은 해 12월에 추가적으로 트위터 자료를 수집하여 같은 방식으로 분석한 결과, 샘플의 구성과 문서의 주제·키워드 등이 5월 결과와 너무 상이하여 의미 있는 시계열 분석을 수행할 수 없었다. 또한 트위터 계정의 이념적·당파적 입장을 연구자가 직접 코딩한 본 분석의 방법상, 분석 시기가 늘어질수록 계정의 수가 과도하게 많아지는 실질적인 문제도 있었다. 기존의 트위터 연구를 볼 때 분석 목적에 따라 수일의 짧은 분석 기간을 갖는 연구들도 상당수 존재한다는 점에서(곽수정, 김현희, 2019; 노예영, 김기환, 2020; Gruzd & Roy, 2014; Tien et al., 2020) 이같은 한계는 어느 정도 정당화될 수 있을 것이다.
- 7) 엄밀히 말해 양극화가 시계열적 변화를 내포하는 개념이란 점에서 이러한 횡단면 분석은 양극화를 논하기에 부족할 수 있다. 하지만 한 시점에서 집단 간의 행태가 서로 얼마나 분극되어(polarized) 있는지를 보는 것 역시 넓은 범위에서 양극화 논의와 연결되며, 이러한 횡단면 분석들이 쌓이면 시계열적 변화를 파악할 수 있다는 점에서도 의미가 있다. 실제로 본 연구가 인용한 많은 설문조사 분석(Harteveld, 2021; Malinas et al., 2021; Milosh et al., 2021 등)과 미디어 분석(Banks et al., 2021; Gruzd & Roy, 2014; Jiang et al., 2020; Paz et al., 2021; Tien et al., 2020; Urman, 2019 등) 역시 특정 시점에서의 횡단면 분석이다.
- 8) 동일한 메시지의 과도한 중복을 막기 위해 리트윗(RT) 메시지는 분석 대상에 포함하지 않았다.
- 9) 문서의 맥락을 온전히 이해하기 위해 명사 외의 품사도 고려해야 한다는 주장도 있으나(박종희 외, 2015) 본 분석과 같이 관찰값이 많은 경우 의미 있는 분석을 저해할 수 있다는 점에서 본 연구에서는 명사만으로 분석 대상을 한정하였다. 품사를 태깅하는 프로그램은 넷마이너에 내장된 ‘seunjeon’ 엔진(<http://eurjeon.blogspot.com/2013/02/blog-post.html>)을 사용하였다.
- 10) 분석의 효용을 높이기 위해 특별한 의미 없이 일상적, 반복적으로 사용되는 단어들 역시 대상에서 제외 수 있으나, 텍스트 분석에서 어떠한 단어가 실질적인 의미가 있는지를 판단하는 것은 매우 주관적인 문제이다.

분석 대상이 된 단어는 총 209개이다.

다음으로 중요한 과정은 트위터 계정을 이념과 당파에 따라 분류하는 일이다. 이념 성향과 당파성은 높은 상관관계를 가지지만 동일한 개념은 아니다(정동준, 2018; Mason, 2015). 우리나라의 경우에도 진보 정당이 정권을 잡고 있는 시기에 코로나 사태가 발생하였고, 백신을 포함한 여러 정책이 진보 여당에 의해 주도되었다는 점에서 양자 간의 개념 구분이 필요하다. 즉, 기존 연구에 따르면 일반적으로 진보 집단은 보수에 비해 코로나의 위험에 더 민감하고 따라서 백신 접종에 보다 긍정적 태도를 가질 것으로 예상된다. 하지만 코로나 백신은 이념에 따라 정치화된 이슈이기도 하지만 동시에 매우 당파적인 이슈이다. 같은 진보라 하더라도 현 정부와 여당을 지지하지 않거나 반대하는 사람이라면, 정부가 주도하는 백신 정책에 반대하고 백신 접종에 보다 부정적 입장을 가질 수 있다. 반대로 보수라 하더라도 현 정부를 긍정적으로 평가하고 있다면 백신에 대해서도 긍정적 입장을 가질 수 있다. 밀로쉬 외(Milosh et al., 2021)는 최근 연구에서 당파성이 코로나 공공정책에 대한 미국 시민들의 태도를 결정하는 가장 중요한 요인임을 밝히기도 하였다. 본 연구는 이러한 점을 고려하여 이념 성향과 별도로 당파에 따라서도 집단을 분류하였다. 특정 정당에 대한 입장이 명확히 알기 힘든 트위터 메시지의 특성

상 정당별로 당파성을 측정하기에는 무리가 있다. 이에 따라 정부와 여당인 더불어민주당에 대해 긍정적 입장을 가지는 집단(친(親)여당), 중도적 입장을 가지거나 정부와 여당에 특별한 입장을 보이지 않는 집단(중도), 그리고 정부와 여당에 반감을 표시하는 집단(반(反)여당)으로 분류하였다.

계정의 이념적, 당파적 성향을 판단하는 방법으로는 팔로잉 관계를 보거나, 계정 생성 시 등록된 사용자의 사회배경 정보를 보거나, 계정의 메시지를 직접 보고 평가하는 방법 등이 있다. 첫 번째의 경우, 팔로우하는 정당이나 정치인의 이념적, 당파적 성향이나(한규섭 외, 2013; Urman, 2019) 혹은 팔로우하는 언론의 정파적, 이념적 성향(Tien et al., 2020)를 통해 계정의 성향을 판단한다. 두 번째의 경우, 계정 생성 시 입력하는 지역(location)과 같은 사회배경 정보를 근거로 판단하는 방법이다(Jiang et al., 2020). 두 방법은 분류가 간편하고 객관적 정보를 분류의 기준으로 삼는다는 장점이 있다. 하지만 특정 계정의 팔로우나 특정 사회배경 출신이 아닌, 일반적인 트위터러안을 대상으로 하는 본 연구에서 두 방법은 표본 편향의 문제를 야기할 수 있다. 또한 팔로잉 관계와 사회배경 정보가 직접적으로 계정의 이념적, 당파적 입장을 나타낸다 보기도 어렵다. 따라서 본 연구에서는 분석 대상 트위터 계정 1,575개의 타임라인에 있는 메시지의 내용을 질적으로 판단, 코딩하여¹⁾ 이념과

특히 네트워크 분석의 특성상, 단독으로는 특별한 의미가 없을지 몰라도 다른 단어와의 연결을 통해 의미를 갖게 될 수 있다는 점에서 주관적인 기준은 최대한 배제하고 본문에 제시된 기준에 따라 분석대상을 선정하였다. 예를 들어 ‘카카오톡’, ‘네이버’, ‘다음’과 같은 고유명사의 경우 내포한 의미를 판단하기 쉽지 않을 수 있으나, ‘잔여’, ‘예약’, ‘출처’, ‘병원’ 등과 같은 연결된 단어들을 통해 그 의미의 방향성을 충분히 파악할 수 있다. 확인을 위해 이러한 단어들을 제외하고 추가분석을 실시한 결과, 본 분석의 결과와 크게 다르지 않았다.

- 11) 집단 분류의 타당성을 확보하기 위해 연구조교 1인과 함께 코딩을 진행하였고, 다수의 온/오프라인 회의를 통해 두 코더 간의 집단 분류기준을 최대한 일치시키고자 하였다. 이를 통해 최종적으로 얻은 코더간 신뢰도(inter-coder reliability)는 이념 집단의 경우, 퍼센트 일치도 85.1%, 카파 지수(Cohen's Kappa) 0.713, 크리펜도르프 알파(Krippendorff's Alpha) 0.712, 당파 집단의 경우 각각 83.6%, 0.732, 0.732로, 코딩 기준에 관한 두 코더 간 합의가 충분히 이루어진 것으로 나타났다(Cohen, 1960; Hayes & Krippendorff, 2007). 코딩 작업에 참여한 천슬기 조교에게 감사를 표한다.

당파 집단을 분류하였다(정효정 외, 2016).¹²⁾

비공개이거나 폐쇄된 계정, 혹은 지나치게 상업적이거나 부적절한 계정 79개를 제외한 나머지 1,496개 계정의 이념과 당파 집단 분류는 아래 <표 1>과 같다.¹³⁾ 표에서 나타난 바와 같이 진보와 친여당 성향의 사용자가 그렇지 않은 사용자에 비해 훨씬 더 많은 것으로 조사되었다. 우리나라에서 트위터가 주로 진보의 정치 공간으로 사용되고 있다는 기존의 연구결과(장우영, 2013; 황유선, 2013)가 본 분석 표본을 통해서도 나타난 것이다.¹⁴⁾

2) 분석 방법

본 논문에서 실시한 분석은 크게 두 가지이다. 첫째, 집단별로 트위터 메시지에 나타난 단어 간에 네트워크를 구성한 후, 어떠한 핵심 키워드를 중심으로 네트워크가 형성되어 있는지를 보는 것이다. 이를 위해 각 집단 내에서 단어 네트워크

를 구성¹⁵⁾한 후, 해당 네트워크에 대한 중심성 분석을 수행하였다. 중심성의 종류에는 여러 가지가 있는데, 핵심 키워드를 찾으려는 본 분석의 목적에 따라 아이겐벡터(Eigenvector) 중심성 분석을 수행하였다. 아이겐벡터 중심성은 단순히 연결된 이웃 노드가 많을 뿐 아니라 그 이웃의 중심성이 높으면 재귀적으로 자신의 중심성도 높아지는 성질을 가지고 있어, 네트워크 전반에 미치는 영향력이 큰 핵심 노드를 찾기에 유용하다(김용학, 김영진, 2019, 126-127쪽).

이러한 구조중심적 분석은 핵심 키워드를 중심으로 단어들 간 관계를 파악하는데 용이하고, 특히 시각화를 통해 분석 결과를 직관적으로 이해할 수 있다는 장점을 가진다. 하지만 키워드를 넘어 텍스트의 토픽이 무엇인지까지는 알기 어렵고, 토픽들 간 중요도 등 보다 자세한 정보는 알 수 없다는 한계를 지닌다. 본 연구는 토픽 모델링을 통해 이러한 점을 보완하고자 한다. 토픽 모델링은 ‘문서를 구성하는 단어들은 무작위로 선택된

<표 1> 집단별 분석 대상 트위터 계정 개수(개)

		이념적 입장			계
		진보	중도	보수	
당파적 입장	친(親)여당	803	0	0	803
	당파적 중도	256	303	0	559
	반(反)여당	13	11	110	134
계		1,072	314	110	1,496

- 12) 내생성의 오류를 범하지 않기 위해 코로나와 관련된 트윗 내용을 근거로 집단을 분류하지 않는 것을 기본 원칙으로 하였다. 기타 자세한 코딩 기준은 별도 첨부한 [부록 1] 참고.
- 13) 극단적 견해가 표출되는 경향이 강한 트위터 특성상 보수 성향이 현재의 진보 여당에 대해 호감을 표하는 트윗 메시지를 발견하기는 어려웠다. 이에 따라 보수 입장의 계정은 모두 반(反)여당으로 분류하였다.
- 14) 본 분석의 목적은 집단별 트위터 메시지의 비중이 아닌, 각 집단 내 메시지의 내용과 구조가 서로 어떻게 다른지를 보는 것이다. 전체 표본이 아닌 각 집단 내 트위터 메시지를 분석하고 그 결과를 집단 간에 비교한다는 점에서 집단별 표본 크기의 차이는 문제가 되지 않을 것이다.
- 15) 이는 기존에 형성된 ‘단어-문서(트위터 메시지)’의 2모드 네트워크를 ‘단어-단어’의 1모드 네트워크로 변환하는 것으로, 같은 문서 내 바로 인접한 단어들(윈도우 크기: 2) 사이에 링크를 형성하였다(링크에 방향성 부여하지 않음).

것이 아니다'란 전제 아래, 관찰 가능한 단어와 문서 사이를 매개하고 있는 토픽이라는 잠재 변수를 찾아내는 분석 방법이다(노예영, 김기환, 2020; Blei, 2012). 다시 말해 문서 내 단어들은 '토픽의 단어 분포'와 '문서의 토픽 분포'의 결합으로 생성된다 가정하고, 토픽 내 단어의 등장 확률과 문서 내 토픽의 등장 확률을 통해 두 분포를 추정하는 기법이다. 추정에는 토픽 모델링에서 널리 사용되는 잠재적 디리클레 할당(Latent Dirichlet Allocation, LDA) 모형을 사용하였다(Blei et al., 2003).

4. 분석 결과

1) 단어 네트워크 중심성 분석

먼저 집단별 단어 네트워크에 대해 아이겐벡터 중심성 분석을 수행하였다. 단어 네트워크에서 노드 간의 과도한 연결로 인한 시각적 혼동을 줄이고자 각 집단에 맞게 링크의 개수를 조절하였다.¹⁶⁾ 분석 결과는 직관적 해석을 위해 스프링 맵(Spring map) 방식으로 시각화하였다.¹⁷⁾ 시각화 맵에서 개별 노드의 크기는 해당 단어의 중심성 수치에 비례하는 것으로, 크기가 클수록 네트워크 상에서 보다 중요하고 핵심적인 단어라 할 수 있다. 지면의 제약으로 중심성 수치가 큰 단어들을 중심으로 끊어지지 않고 서로 직간접적으로 연결된 컴포넌트(components)만을 추려 집단별로 <그림 1>과 <그림 2>에 제시하였다.¹⁸⁾

먼저 <그림 1-1>부터 <그림 1-3>은 이념 집

단별 결과이다. 진보 집단의 분석 결과인 <그림 1-1>의 경우, '뉴스', '다음', '네이버', '종합', '카카오', '출처' 순으로 중심성 수치가 높게 나타났다. 이들은 트윗 메시지에 같이 등장하며 나란히 연결된 구조를 보였다. 이 밖에 '앱', '잔여', '동네', '병원' 등의 단어도 비교적 높은 중심성 수치를 보이며 서로 연결되어 있었다. 진보 집단에서는 백신 관련 뉴스, 잔여 백신에 대한 집중, 그리고 이러한 집중을 할 수 있는 앱(application)과 동네 병원에 관한 정보가 가장 많이 논의되고 있는 것이다. '걱정'과 같은 부정적 단어도 함께 등장하긴 하였으나 그 중심성 수치가 상대적으로 낮고, 단어 구조 상 '잔여'와 연결되어 있는 것으로 볼 때 백신 부작용에 대한 걱정보다는 백신의 공급 부족으로 쉽게 예약하지 못하는 상황에 대한 걱정으로 보인다. 즉, 진보 집단에서는 백신 관련 정보에 대한 메시지들이 주로 트윗되고 있고, 백신에 대한 부정적 우려나 정부에 대한 불신과 같은 비판적 메시지는 상대적으로 적은 것으로 나타났다. 이러한 양상은 중도 집단 결과인 <그림 1-2>에서도 비슷하게 관찰되었다. '카카오', '네이버', '앱', '잔여' 순으로 높은 중심성 수치를 보여 백신 접종 정보와 예약 방법에 대한 논의가 가장 많은 것으로 나타났다. 하지만 중도에서는 '먹통', '물량', '실패'와 같은 부정적 단어들도 상대적으로 높은 중심성 수치를 보여, 당시의 백신 공급 상황에 대한 불편과 염려가 어느 정도 거론되고 있는 것으로 나타났다.

이에 반해 보수 집단 결과인 <그림 1-3>에서 가장 핵심적으로 논의된 키워드는 백신 접종의

16) 같은 트윗 메시지에 동시등장(co-occurrence)하는 횟수를 기준으로, 기준 이하의 링크는 삭제하였다. 상대적으로 관측치가 많은 진보 및 친여 집단은 5회 이상으로, 관측치가 적은 중도 및 보수, 반여 집단은 2회 이상의 링크만을 남겨 시각화하였다.

17) 자세한 집단별 각 단어의 중심성 수치는 [부록 2] 참고.

18) 전체 단어 네트워크에 대한 그림은 [부록 3] 참고.

나타났다. 백신 부작용 문제와 원인에 대한 트윗이 수위를 차지했던 보수와 달리, 반여당인 <그룹 2-3>에서는 ‘뉴스’, ‘네이버’, ‘다음’, ‘카카오톡’이 가장 높은 중심성 수치를 보여 다른 두 집단과 마찬가지로 백신 관련 뉴스와 예약 정보에 대한 논의가 가장 활발하게 나타난 것이다. 특이한 점은 ‘경제’가 바로 다음으로 높은 중심성을 보였다는 점이다. 코로나와 백신으로 인한 경제적 상황이 이들의 주된 관심사로 나타난 것이다. ‘부작용’, ‘혈전’, ‘원인’에 대한 트윗이 그 뒤를 이으며 하나의 컴포넌트를 이루었고, ‘국제’, ‘사회’, ‘협력’과 ‘정부’, ‘대한민국’, ‘대통령’ 등도 활발히 논의된 주제였다.

보수 집단과 반여당 집단 사이의 이 같은 차이는 기존 연구로 어느 정도 설명할 수 있다. 보수 집단에서 백신의 부작용 문제가 가장 핵심적으로 거론되고 있다는 것은, 보수일수록 코로나의 위험성이 과장된 것으로 보고 코로나 관련 여러 사회보장 정책에 반대하는 경향이 강하다는 기존의 연구 결과(Calvillo et al., 2020; Clarke et al., 2021; Kerr et al., 2021; Mallinas et al., 2021; Pennycook et al., 2021)와 일맥상통한다. 코로나가 그다지 위험한 질병이 아니며 오히려 백신으로 인한 부작용과 위험이 문제라는 쪽으로 논의가 흐르고 있는 것이다. 반면, 당파성은 정부에 대한 신뢰 및 평가와 연관된 개념이다(길정아, 강원택, 2020; 박선경, 신진욱, 2021; Jiang et al., 2020; Klein & Robison, 2020). 즉, 여당에 부정적인 사람은 백신 자체의 문제보다도 그로 인한 경제 상황과 정부의 대응에 초점을 두고 이에 대

한 부정적 평가를 내리려는 경향이 강하기에, ‘경제’, ‘정부’, ‘대한민국’, ‘대통령’과 같은 단어가 중요한 키워드로 나타난 것이다.

2) 토픽 모델 분석

지금까지 단어 네트워크 중심성 분석을 통해 어떤 단어가 핵심 단어인지를 알아보았다면, 이제 이러한 단어들이 어떠한 토픽들을 중심으로 문서, 즉 트위터 메시지를 이루고 있는지를 살펴보겠다. 토픽 모델링에서 중요한 것은 해당 문서들이 몇 개의 토픽으로 구성되어 있는지를 정하는 일이다. 적정 토픽 개수에 대한 합의된 기준은 현재까지 없고, 토픽들의 내용이 서로 중복되지 않을 정도로 적고, 같은 토픽 내에 지나치게 다양한 주제의 단어들이 포함되지 않을 정도로 많은 수준에서 연구자의 재량에 의해 결정된다(이수상, 2016). 혼잡도(perplexity)와 응집도(coherence) 같이 잠재된 토픽의 개수를 확률적으로 추정하는 방식도 있으나 이 또한 각각의 한계를 가지고 있고(이대영, 이현숙, 2021), 본 분석에서 사용한 넷마이너 프로그램이 이를 추출하는 패키지를 현재로서는 제공하고 있지 않다.¹⁹⁾ 따라서 본 분석에서는 각 집단별로 토픽의 개수를 5개로 할 때 범주화가 가장 잘 이루어진다고 보았다. 분석에 필요한 기타 값들은 선행 연구를 참고하였다²⁰⁾ <표 2>~<표 7>은 집단별 분석 결과에 따라 연구자가 주관적으로 부여한 ‘토픽 라벨(label)’과 각 토픽을 구성하고 있는 주요 ‘키워드’²¹⁾를, 토픽에서 각 키워드가 차지하는 비중이라 할 수 있는 키워드별 ‘확률’

19) 대안으로 토픽 내 단어 간의 클러스터링 수준을 알려주는 ‘실루엣(Silhouette) 계수’(Mehta et al., 2014)를 구할 수 있었는데, 3~10개 사이의 토픽 개수에서 각 집단별 실루엣 계수를 구한 결과 그 값이 크게 다르지 않았다.

20) 잠재 변수의 추정방식으로는 MCMC(Markov Chain Monte Carlo)를 사용하였고(Chang, 2011, p. 262), ‘문서의 토픽 분포 하이퍼파라미터(α) 값’과 ‘토픽의 단어 분포 하이퍼파라미터(β) 값’은 선행 연구를 따라 가장 최적의 값이라 여겨지는 0.5와 0.01로 각각 설정하였다(Naili et al., 2017, p. 56).

과 함께 제시하였다.

먼저 이념 집단 별 결과를 보면, <표 2>의 진보 집단에서는 ‘백신 관련 뉴스’를 토픽으로 다른 문서가 가장 많은 것으로 확인되었다. 전체 3,022개의 트윗 중 26.7%에 해당하는 808개의 문서가 ‘뉴스’와 ‘다음’ 등의 키워드를 중심으로 한 백신 관련 정보와 관련된 트윗이었다. 물론 ‘이상’, ‘하루’, ‘국민’ 등의 키워드도 있으나 상대적 분포 확률이 높지 않고 의미의 연관성이 낮았다. 다음으로는 ‘대한민국 국가와 정부’, ‘백신 부작용과 이상’, 그리고 ‘잔여 백신이 있는 동네 병원’과 ‘예약 방법(앱)’의 주제가 그 뒤를 이었다. 중도 집단은 다른 집단에 비해 토픽의 범주화가 분명하지 않게 나타났다. <표 3>에 따르면 ‘사람’, ‘마스크’, ‘생각’과 같은 다소 추상적이고 일관성 없는 키워드로 구성된 토픽이 가장 많은 비중(전체 트윗 중 40.2%)을 이루었고, 뒤를 이어 ‘예약 가능 잔여 백신’, ‘부작용 문제와 뉴스’, ‘잔여 백신 예약 방법’과 ‘백신 이상과 코로나 예방’ 등의 주제가 높은 비중을 차지하였다.

보수 집단 분석 결과인 <표 4>의 가장 두드러진 특징은 ‘혈전 등 백신 부작용’에 관한 트윗이 가장 많은 비중을 차지하고 있다는 점이다. 전체

218개의 문서 중 23.9%에 해당하는 52개의 문서가 해당 주제를 담고 있었다. 물론 다른 집단에서도 ‘부작용’과 같은 키워드가 발견되었지만 보수 집단과 같이 ‘혈전’ 등 구체적인 부작용 증상에 대해 활발한 논의가 이루어지지 않는 않았다. 기타 ‘백신 관련 뉴스’, ‘대한민국과 각국의 백신 상황’, ‘코로나 바이러스 집단 면역’, ‘잔여 백신 예약 방법’에 관한 주제도 높은 비중을 보였다. 특히 ‘미국’, ‘중국’과 같은 다른 나라들의 코로나 상황에 대한 논의는 다른 집단에서는 발견되지 않는 특징이었다.

정리하면 진보 집단에서는 백신 관련 뉴스와 관련 정보의 검색이 가장 활발히 논의된 토픽이었던 반면, 보수 집단에서는 혈전 등 백신 부작용과 관련된 토픽이 가장 많이 비중을 차지한 것으로 나타났다. 중도 집단에서는 상대적으로 토픽이 명확히 구별되지 않았다. 이 같은 결과는 위의 단어 네트워크 중심성 분석 결과와도 일맥상통한다. 진보에서는 백신 관련 뉴스와 예약 방법이, 보수에서는 혈전 등 백신 부작용 문제와 원인이 가장 핵심적인 단어로 연결되었던 앞선 분석과 유사한 결과가 토픽 분석에서도 발견된 것이다.

당파 집단에 따른 토픽 모델 결과 역시 앞선 단

<표 2> 이념적 진보 집단 트윗의 주요 토픽

토픽 라벨	문서 수(개)	문서 수(%)	키워드 1	확률	키워드 2	확률	키워드 3	확률	키워드 4	확률	키워드 5	확률
1 백신 관련 뉴스	808	26.7	뉴스	0.117	다음	0.094	이상	0.039	하루	0.036	국민	0.035
2 대한민국 국가와 정부	591	19.6	대한민국	0.077	정부	0.062	국가	0.058	대통령	0.045	언론	0.042
3 백신 부작용과 이상	565	18.7	사람	0.097	마스크	0.05	부작용	0.049	생각	0.046	이상	0.041
4 잔여 백신 동네 병원	562	18.6	잔여	0.195	병원	0.089	예약	0.078	사람	0.057	동네	0.046
5 잔여 백신 예약 방법	496	16.4	잔여	0.2	예약	0.15	네이버	0.051	가능	0.048	카카오톡	0.045

21) 지면의 제약으로 토픽별 상위 5개의 키워드만을 정리하였다. 보다 자세한 결과는 [부록 4] 참고.

<표 3> 이념적 중도 집단 트윗의 주요 토픽

토픽 라벨	문서 수(개)	문서 수(%)	키워드 1	확률	키워드 2	확률	키워드 3	확률	키워드 4	확률	키워드 5	확률
1 사람과 마스크	229	40.2	사람	0.119	마스크	0.066	생각	0.043	주사	0.042	이야기	0.035
2 예약 가능 잔여 백신	103	18.1	잔여	0.2	예약	0.111	가능	0.09	오늘	0.06	신청	0.06
3 부작용 문제와 뉴스	87	15.3	뉴스	0.108	확진	0.052	부작용	0.051	문제	0.043	다음	0.04
4 잔여 백신 예약 방법	76	13.4	잔여	0.152	예약	0.136	카카오톡	0.088	네이버	0.088	병원	0.077
5 백신 이상과 코로나 예방	74	13.0	시작	0.106	이상	0.087	예방	0.064	질병	0.039	연령	0.035

<표 4> 이념적 보수 집단 트윗의 주요 토픽

토픽 라벨	문서 수(개)	문서 수(%)	키워드 1	확률	키워드 2	확률	키워드 3	확률	키워드 4	확률	키워드 5	확률
1 혈전 등 백신 부작용	52	23.9	부작용	0.089	확보	0.078	정부	0.077	혈전	0.074	사회	0.074
2 마스크와 백신 관련 뉴스	51	23.4	마스크	0.156	뉴스	0.057	출처	0.054	국민	0.039	사람	0.038
3 대한민국과 각국의 백신 상황	44	20.2	미국	0.18	대한민국	0.077	국가	0.077	중국	0.072	이상	0.065
4 코로나 바이러스 집단 면역	37	17.0	국민	0.112	면역	0.099	집단	0.076	바이러스	0.055	문제인	0.041
5 잔여 백신 예약 방법	34	15.6	예약	0.116	잔여	0.11	네이버	0.047	뉴스	0.042	출처	0.035

어 네트워크 분석과 비슷하게 나타났다. <표 5>에 따르면 친여당 집단에서는 진보와 마찬가지로 ‘잔여 백신 병원 예약’, ‘잔여 백신 예약 방법’, ‘백신 관련 뉴스와 불안’과 같은 정보성 토픽이 가장 활발하게 논의된 가운데, ‘백신 이상과 코로나 예방’, 그리고 ‘대한민국 국가와 정부’의 역할에 대한 논의가 뒤를 이었다. 당파적 중도 집단인 <표 6>에서도 마찬가지로 ‘백신 관련 정보’와 ‘예약 방법’에 대한 주제가 주를 이루었으나, 이념적 중도 집단과 마찬가지로 같은 토픽 아래 다양한 주제의 키워드들이 섞여 있어 토픽의 범주화가 명료하게 이루어지지 못했다.

반여당 집단의 경우에도 보수 집단과 비슷하게 ‘혈전 등 백신 부작용’ 주제가 가장 많은 비중을 차

지하였다. <표 7>에 의하면 해당 토픽이 나타난 문서의 비중은 24.1%로 보수 집단에 비해서도 소폭이나마 높았고, ‘부작용’, ‘혈전’의 분포 확률은 더욱 높게 나타났다. 이 외 반여당에서 두드러지게 나타난 토픽은 두 번째로 많은 비중을 차지한 ‘정부의 백신 확보’와 다섯번째인 ‘백신 이상 반응과 국가 정책’이었다. 이는 보수 집단의 트윗 문서에서는 등장하지 않은 토픽들로, 이러한 차이는 앞서 언급한 바와 같이 두 집단이 가진 특징에서 기인한다 볼 수 있다. 즉, 보수 집단은 백신의 부작용 문제를 민감하게 인식하는 반면, 반여당 집단은 이에 대한 정부의 역할과 대응에 보다 큰 관심을 갖는 것이 이러한 차이로 이어진 것이다.

<표 5> 친여당 집단 트윗의 주요 토픽

토픽 라벨	문서 수(개)	문서 수(%)	키워드 1	확률	키워드 2	확률	키워드 3	확률	키워드 4	확률	키워드 5	확률
1 잔여 백신 병원 예약	586	24.4	잔여	0.223	예약	0.113	사람	0.08	병원	0.07	노쇼	0.056
2 잔여 백신 예약 방법	495	20.6	잔여	0.151	예약	0.107	가능	0.066	네이버	0.054	카카오톡	0.043
3 백신 관련 뉴스와 불안	474	19.7	뉴스	0.128	다음	0.095	불안	0.041	사람	0.04	국민	0.035
4 백신 이상과 코로나 예방	450	18.7	이상	0.083	예방	0.041	마스크	0.041	노인	0.04	시작	0.038
5 대한민국 국가와 정부	397	16.5	대한민국	0.078	국가	0.068	정부	0.065	대통령	0.054	미국	0.044

<표 6> 당파적 중도 집단 트윗의 주요 토픽

토픽 라벨	문서 수(개)	문서 수(%)	키워드 1	확률	키워드 2	확률	키워드 3	확률	키워드 4	확률	키워드 5	확률
1 잔여 백신 예약 방법	391	34.8	잔여	0.204	예약	0.192	카카오톡	0.076	노쇼	0.071	네이버	0.067
2 동네 신청 가능 잔여 백신	197	17.5	잔여	0.23	신청	0.084	가능	0.059	동네	0.047	일본	0.044
3 사람과 마스크	189	16.8	사람	0.101	마스크	0.072	뉴스	0.068	다음	0.058	부작용	0.051
4 백신 접종 병원	188	16.7	병원	0.11	생각	0.059	전화	0.047	주사	0.044	연락	0.037
5 백신 이상과 예약	160	14.2	시작	0.085	이상	0.063	예약	0.053	사람	0.044	노인	0.038

<표 7> 반여당 집단 트윗의 주요 토픽

토픽 라벨	문서 수(개)	문서 수(%)	키워드 1	확률	키워드 2	확률	키워드 3	확률	키워드 4	확률	키워드 5	확률
1 혈전 등 백신 부작용	68	24.1	부작용	0.127	혈전	0.092	문제	0.066	원인	0.061	대상	0.061
2 정부의 백신 확보	56	19.9	국민	0.12	확보	0.073	사회	0.073	면역	0.068	정부	0.064
3 잔여 백신 예약 방법	56	19.9	잔여	0.107	예약	0.1	뉴스	0.083	출처	0.066	네이버	0.062
4 대한민국과 각국의 백신 상황	51	18.1	미국	0.139	마스크	0.099	대한민국	0.098	중국	0.095	문재인	0.073
5 백신 이상 반응과 국가 정책	51	18.1	이상	0.107	국가	0.078	인센티브	0.059	반응	0.032	생각	0.03

5. 결론 및 함의

본 논문에서는 코로나 백신을 중심으로 한 소셜 미디어의 논의가 이념적, 당파적으로 어떻게 다르

게 나타나는지를 트위터 메시지를 중심으로 분석하였다. 분석 결과, 코로나 백신을 둘러싼 트위터 메시지는 이념과 당파에 따라 서로 다른 단어 구조와 주제를 갖는 것으로 나타났다. 먼저 단어 네

트위크 중심성 분석 결과, 진보적 성향을 가진 집단의 트윗은 백신 관련 뉴스, 잔여 백신 확인과 예약 방법 등 백신에 관한 정보를 공유하는 단어들이 핵심을 이뤄 연결되어 있었다. 반면 보수 집단의 트윗에서는 백신의 부작용과 원인에 대한 단어들이 높은 중심성 수치를 보이며 연결되어 있었다. 당파를 중심으로 집단을 나눈 경우, 친여당 집단은 진보 집단과 마찬가지로 백신 관련 뉴스와 예약 방법에 대한 단어들이 핵심 단어를 구성하고 있었다. 반여당의 경우에는 보수 집단과 조금 다른 결과가 나왔는데, 백신 부작용과 원인에 대한 단어가 중요한 단어 구조를 이루고 있는 것은 동일하나 그 중심성 수치가 상대적으로 낮게 나왔고, 대신 경제 문제나 국제사회의 협력, 정부와 국가의 역할에 대한 단어들이 핵심적인 키워드를 구성하고 있는 것으로 나타났다.

이 같은 결과는 이어진 토픽 모델 분석에서도 유사하게 나타났다. 진보 집단의 트윗에서 가장 활발히 논의된 토픽은 백신 관련 뉴스였던 반면, 보수에서 가장 많이 논의된 주제는 혈전 등 백신 부작용에 관한 것이었다. 당파에 따른 분석에서도, 친여당은 잔여 백신의 예약과 백신 관련 뉴스가 토픽의 수위를 차지한 반면, 반여당의 경우에는 혈전 등 부작용이 가장 자주 논의되는 주제로 나타났다. 토픽 분석에서도 반여당과 보수 집단 사이에 약간의 차이가 관찰되었는데, 반여당에서는 보수에서 관찰되지 않던 각국의 백신 상황과 정부의 역할, 그리고 정부의 백신 정책에 대한 논의들이 주요 토픽을 이루고 있었다.

이러한 결과는 기존의 코로나 정치화와 관련된 연구와 일맥상통한다. 이념적으로 보수일수록 코로나의 위험성이 과장되었다 여기고 이에 대한 정부의 정책과 백신 접종에 부정적 입장을 가진다는 기존의 연구결과가 본 분석에서도 일부나마 재현된 것이다. 진보의 경우 반대로 백신의 위험

성을 비교적 사실에 근거하여 인식함으로 이에 대한 조치와 백신 접종에 보다 협조적이란 결과 역시 본 분석을 통해 확인할 수 있었다. 이념적 입장과 정부 여당에 대한 입장이 다를 수 있다는 점에서 실시한 당파에 따른 분석도 전반적으로 비슷한 결과를 보였으나, 보수와 반여당 집단 사이에 약간의 차이가 발견되었다. 백신의 부작용 문제에 중점을 둔 보수와 달리, 반여당 집단에서는 정부의 역할과 정책에 대한 논의가 보다 활발히 이루어진 것이다. 이러한 결과 역시, 정부와 여당에 대한 평가가 당파적 입장의 형성에 영향을 미친다는 기존의 연구 결과로 설명될 수 있다.

정치 양극화가 오늘날 민주주의의 위기를 가져오고 있는 주된 요인으로 지목되는 상황에서 이러한 결과는 의미하는 바가 크다. 코로나와 같은 비정치적 이슈마저도 이에 대한 인식과 대처에 있어 이념과 당파에 따라 입장이 다르게 나타나는 양극화 경향이 우리나라에서도 관찰되고 있는 것이다. 비록 우리나라에서는 아직까지 미국이나 프랑스 등 서구 민주주의와 같이 코로나를 둘러싼 폭력 시위까지는 발발하고 있지는 않지만, 정치적 담론 형성에 중요한 매개체 역할을 하는 소셜미디어 상에서 이렇게 이념적, 당파적으로 분열된 모습이 나타나고 있는 것은 언젠가는 그러한 사태가 일어날 수 있는 불씨가 있음을 보여준다. 실제로 온라인에서의 갈등이 오프라인 공간에서의 갈등으로 번질 수 있음이 앞선 연구를 통해 밝혀지기도 했다(Gallacher et al., 2021). 특히 2022년 대선을 앞두고 주요 후보들이 백신 이슈를 더욱 정치적으로 이용할 수 있다는 점에서 이러한 양극화에 대한 지속적인 관찰과 주의가 요구된다.

본 연구는 이러한 코로나 백신의 정치화에 대한 한국 사례의 분석을 통해 학계의 논의에 이론적, 실증적으로 기여하였다는 데 의의가 있다. 특히 이러한 정치화가 디지털 시대에 주목받고 있는 온라

인 정치 공간에서 어떠한 양상으로 나타나는지를 살펴봄으로써, 현재 우리나라 정치 양극화의 현주소를 부분적으로나마 파악할 수 있었다. 이러한 의의에도 불구하고 본 연구는 다음과 같은 한계를 지닌다. 먼저 첫째로, 본 분석이 이념과 당파에 따른 코로나 백신에 대한 입장 차이를 밝혀내긴 하였으나 이것을 곧 양자 간의 인과관계로 보기는 어렵다. 이는 무엇보다도 계정, 즉 사용자(user)의 성별, 나이, 소득, 교육수준 등 사회배경 변수나 코로나 백신에 대한 태도에 영향을 미칠 만한 제3의 요인들이 통제되어 있지 않기 때문이다. 공개 API를 통해 얻은 자료만으로는 이러한 요인들에 대한 정보를 알 수 없었다. 이 같은 문제는 서버이 등을 통해 사용자에게 다양한 정보를 얻고, 이를 바탕으로 집단을 구분하여 집단 간 메시지를 비교 분석함으로써 해결될 수 있을 것이다. 즉, 실제 사용자에게 대한 설문과 그들이 작성한 메시지에 대한 내용 분석, 이 둘을 혼합해야 하는데 그 방법의 어려움으로 본 연구를 포함한 대부분의 연구에서 둘 중 한 방법만을 사용하고 있다. 따라서 본 논문은 소셜미디어 상의 정치 양극화에 대

한 탐색적(exploratory) 연구의 하나로 볼 수 있고, 보다 엄정한 인과관계는 향후 연구 방법의 보완을 통해 밝혀낼 수 있을 것이다.

또한 본 연구는 여러 소셜미디어 플랫폼 중 트위터만을 분석함으로써 편향된 표본을 사용하였다는 한계가 있다. 주지하듯이 트위터는 극단주의적 견해가 많이 표현되는 공간이고, 이는 다른 소셜미디어와 비교하더라도 그렇다는 것이 기존 연구의 공통적 견해이다. 이에 더하여 본 분석에서는 전체 표본 중 진보 집단이 차지하는 비중이 보수에 비해 압도적으로 높았다. 물론 본 연구의 목적이 전체에서 각 집단이 차지하는 비중을 비교하는 것이 아닌, 각 집단 내 수행된 분석 결과를 서로 비교한다는 점에서 그 자체로 문제가 될 수는 없다. 또한 아무리 치우쳤다 하더라도 이러한 논의가 결국 다른 소셜미디어와 온라인 공간으로 확산되고 이것이 전반적인 정치 담론에 영향을 미친다는 점에서 충분히 분석할만한 가치를 지닌다. 그럼에도 트위터 자료만이 갖는 편향된 특징을 보완하기 위해서는 향후 다른 소셜미디어 플랫폼 및 온라인 공간에 대한 분석이 동반되어야 할 것이다.

참 고 문 헌

- 가상준 (2016). 정책영역별로 본 국회 양극화. <OUGHTOPIA>, 31(1), 327-354.
- 강원택 (2012). 제19대 국회의원의 이념 성향과 정책 태도. <의정연구>, 18(2), 5-38.
- 곽수정 · 김현희 (2019). 텍스트 마이닝과 토픽 모델링을 기반으로 한 트위터에 나타난 사회적 이슈의 키워드 및 주제 분석. <KTSDE>, 8(1), 13-18.
- 길정아 · 강원택 (2020). 제21대 국회의원선거에서의 회고적 투표: 대통령의 코로나 대응 평가와 당파적 편향.<한국 정당학회보>, 19(4), 101-140.
- 길정아 · 허상용 (2019). 당파적 편향에 따른 책임 귀속: 여야간 갈등인식과 정당 호감도를 중심으로. <의정연구>, 25(1), 46-78.
- 김기동 · 이재목 (2021). 한국 유권자의 당파적 정체성과 정서적 양극화. <한국정치학회보>, 55(2), 5787.
- 김용학 · 김영진 (2019). <사회 연결망 분석>. 제4판 중판. 박영사.

- 남현주 · 김하나 (2014). 연구논문: 정치적 소통도구로서의 SNS 활용: 제18대 대선후보자를 중심으로. <공공정책과 국정관리>, 8(1), 31-68.
- 노예영 · 김기환 (2020). 트위터에 반영된 2018 평창 동계올림픽의 사회적 시정 이슈 분석: 구조적 주제 모형 (Structural Topic Modeling)의 적용. <한국스포츠산업 · 경영학회지>, 25(6), 1-17.
- 박선경 · 신진옥 (2021). 코로나19와 국가역할 확대에 대한 국민 인식: 이념성향, 재난피해, 정부대응, 재난재원의 영향. <한국과 국제정치>, 37(1), 153-189.
- 박중희 · 박은정 · 조동준 (2015). 북한 신년사(1946-2015)에 대한 자동화된 텍스트 분석. <한국정치학회보>, 49(2), 27-62.
- 박주현 (2020). 언론의 이념성향에 따른 '코로나19' 보도 프레임 비교 연구. <한국언론학보>, 64(4), 40-85.
- 박지영 (2020). 빅데이터 분석을 통해 살펴본 미디어의 정치적 편향성 및 선택적 미디어 노출로 인한 정치적 양극화: '검찰개혁'을 바라보는 상반된 인식을 중심으로. <한국정치연구>, 29(3), 213-243.
- 설선미 · 배정아 (2018). 민선 7기 지방선거과정에 나타난 정책이슈의 역동성 분석: 정당별 보도 자료 및 트위터 (Twitter)의 연결망 분석을 중심으로. <지방행정연구>, 32(3), 35-63.
- 송진미 (2020). 재난 상황에서의 이슈 선점과 정당 지지: 코로나19 발생 전 · 후 패널데이터 분석을 중심으로 <한국 조사연구학회 조사연구>, 21(4), 63-91.
- 신정섭 (2020). 코로나19가 제21대 국회의원 선거 투표선택에 미친 영향: 정부대응 평가와 개인피해를 중심으로. <한국정치연구>, 29(3), 155-182.
- 이소영 (2012). 4.11 총선과 소셜 네트워크 정치캠페인: 총선 후보자의 트위터 선거캠페인을 중심으로. <21세기정치학회보>, 22(3), 287-312.
- 이수상 (2016). 독후감 텍스트의 토픽모델링 적용에 관한 탐색적 연구. <한국도서관 · 정보학회지>, 47(4), 1-18.
- 이종임 · 홍주현 · 설진아 (2019). 트위터에 나타난 미투 (#Me Too)운동과 젠더 갈등이슈 분석: 네트워크 분석하 의미분석을 중심으로. <미디어, 젠더&문화>, 34(2), 99-146.
- 장승진 · 서정규 (2019). 당파적 양극화의 이원적 구조: 정치적 정체성, 정책선호, 그리고 정치적 세련도. <한국정당학회보>, 18(3), 5-29.
- 장승진 · 장한일 (2020). 당파적 양극화의 비정치적 효과. <한국정치학회보>, 54(5), 153-175.
- 장우영 (2013). 트위터스피어의 선거 이슈와 동원: 19대 총선을 사례로. <21세기정치학회보>, 23(3), 247-268.
- 장우영 (2019). 소셜미디어 선거캠페인 연구 동향과 쟁점. <정보화정책>, 26(1), 3-24.
- 장우영, 김석주 (2014). 트위터 선거캠페인과 정치동원: 2011년 서울시장 재보궐선거 사례를 중심으로. <한지역정보학회지>, 17(1), 93-123.
- 장우영, 이관률 (2017). SNS 활용과 선거캠페인 효과. <연구방법논총>, 2(1), 53-78.
- 정동준 (2017). 한국 정치공간의 시민과 대표 간 이념적 일치: 개념화와 측정. <의정연구>, 23(2), 67-108.
- 정동준 (2018). 2018년 지방선거 이후 유권자들의 정치 양극화: 당파적 배열과 부정적 당파성을 중심으로. <오토피아>, 33(3), 143-180.
- 정효정 · 배정환 · 홍수린 · 박찬웅 · 송민 (2016). 정치적 이념에 따른 트위터 공간에서의 집단 간 의견차이 분석: 세월호사건을 중심으로. <한국언론학보>, 60(2), 269-302.

- 한규섭 · 박주용 · 이덕재 · 이혜림 (2013). 트위터 팔로잉 관계에 대한 대표성과 양극화에 대한 논의 검증: 한국과 미국의 의회구성원들의 트위터 팔로워들 네트워크 비교 연구. <사이버커뮤니케이션학보>, 30(1), 295-336.
- 황유선 (2013). 진보적 정치 성향과 정치 지식은 정치 대화를 촉진하는가?: 트위터 매개 정치 대화와 트위터연계 정치 대화에 관한 영향력 탐색. <한국언론학보>, 57(3), 221-248.
- Abramowitz, A. I., & Webster, S. (2016). The rise of negative partisanship and the nationalization of U.S. elections in the 21st century. *Electoral Studies*, 41, 12-22.
- Asher, M., Leston Bandeira, C., & Spaiser, V. (2019). Do Parliamentary Debates of e Petitions Enhance Public Engagement With Parliament? An Analysis of Twitter Conversations. *Policy & Internet*, 11(2), 149-171.
- Bafumi, J., & Shapiro, R. Y. (2009). A New Partisan Voter. *Journal of Politics*, 71(1), 1-24.
- Beam, M. A., Hutchens, M. J., & Hmielowski, J. D. (2018). Facebook news and (de)polarization: reinforcing spirals in the 2016 US election. *Information, Communication and Society*, 21(7), 940-958.
- Bessi, A., Petroni, F., Vicario, M. D., Zollo, F., Anagnostopoulos, A., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). Homophily and polarization in the age of misinformation. *The European Physical Journal*, 225, 2047-2059.
- Blei, D. M. (2012). Probabilistic Topic Models. *Communications of the ACM*, 55(4), 77-84.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Bozdog, C. (2020). Managing Diverse Online Networks in the Context of Polarization: Understanding How We Grow Apart on and through Social Media. *Studier i Modern Språkvetenskap*, 6(4).
- Bracciale, R., Andretta, M., & Martella, A. (2021). Does populism go viral? How Italian leaders engage citizens through social media. *Innovation: The European Journal of Social Science Research*, doi:https://doi.org/10.1080/1369118X.2021.1874472.
- Buchi, M., & Vogler, F. (2017). Testing a Digital Inequality Model for Online Political Participation. *Socius: Sociological Research for a Dynamic World*, 3, 1-13.
- Buder, J., Rabl, L., Feiks, M., Badermann, M., & Zurstiege, G. (2020). Does Negatively Toned Language Use on Social Media Lead to Attitude Polarization?. *PsyArXiv*, October 22, doi:10.31234/osf.io/dx9ws.
- Calvillo, D. P., Ross, B. J., Garcia, R. J. B., Smelter, T. J., & Rutchick, A. M. (2020). Political Ideology Predicts Perceptions of the Threat of COVID-19 (and Susceptibility to Fake News About It). *Social Psychological and Personality Science*, 11(8), 1119-1128.
- Chang, E. Y. (2011). *Foundations of Large-Scale Multimedia Information Management and Retrieval: Mathematics of Perception*, Springer Science & Business Media.
- Clarke, E. J., Anna, R. K., & Dyos, E. (2021). The role of ideological attitudes in responses to COVID-19 threat and government restrictions in Australia. *Personality and Individual Differences*, 175.

- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37-46.
- Cole, A. (2005). Old Right or New Right? The Ideological Positioning of Parties of the Far Right. *European Journal of Political Research*, 44, 203-30.
- Dalton, R. J. (2007). Partisan Mobilization, Cognitive Mobilization and the Changing American Electorate. *Electoral Studies*, 26, 274-86.
- Davis, T., Frost, M., & Cohen, R. (2014). *The Partisan Divide: Congress in Crisis*. Campbell, CA: Premiere.
- De Arruda, Henrique Ferraz, Felipe Maciel Cardoso, Guilherme Ferraz de Arruda, Alexis R. Hernández, Luciano da Fontoura Costa, & Yamir Moreno (2021) Modelling how social network algorithms can influence opinion polarization. *Information Sciences*, 588, 265-278.
- Dinkelberg, Alejandro, Caoimhe O'Reilly, Pádraig MacCarron, Paul J. Maher, & Michael Quayle. (2021). Multidimensional polarization dynamics in US election data in the long term (2012-2020) and in the 2020 election cycle. *Analyses of Social Issues and Public Policy*, 21(1), 284-311.
- Druckman, J. N., Klar, S., Krupnikov, Y., Levendusky, M., & Ryan, J. B. (2020). How Affective Polarization Shapes Americans' Political Beliefs: A Study of Response to the COVID-19 Pandemic. *Journal of Experimental Political Science*.
- Feezell, J. T., Wagne, J. K., & Conroy, M. (2021). Exploring the effects of algorithm-driven news sources on political behavior and polarization. *Computers in Human Behavior*, 116.
- Fiorina, M. (2017). *Unstable Majorities: Polarization, Party Sorting, and Political Stalemate*, Stanford: Hoover Press.
- Fredén, A., Rheault, L., & Indridason, I. H. (2020). Betting on the underdog: the influence of social networks on vote choice. *Political Science Research and Methods*, 1-8.
- Gallacher, J. D., Heerdink, M. W., & Hewstone, M. (2021). Online Engagement Between Opposing Political Protest Groups via Social Media is Linked to Physical Violence of Offline Encounters. *Social Media + Society*: <https://doi.org/10.1177/2056305120984445>
- Gilardi, F., Gessler, T., Kubli, M., & Müller, S. (2021). Social Media and Policy Responses to the COVID-19 Pandemic in Switzerland. *Swiss Political Science Review*, 27(2), 243-256.
- Gruzd, A., & Roy, J. (2014). Investigating Political Polarization on Twitter: A Canadian Perspective. *Policy & Internet*, 6(1), 28-45.
- Grzymala-Busse, A. (2019). The Failure of Europe's Mainstream Parties. *Journal of Democracy*, 30(4), 35-47.
- Hameleers, M. (2020). Augmenting polarization via social media? A comparative analysis of Trump's and Wilders' online populist communication and the electorate's interpretations surrounding the elections. *Acta Politica*, 55, 331-350.

- Hameleers, M., Schmuck, D., Bos, L., & Ecklebe, S. (2020). Interacting with the ordinary people: How populist messages and styles communicated by politicians trigger users' behaviour on social media in a comparative context. *European Journal of Communication*, 36(3), 238-253.
- Hart, P. S., Chinn, S., & Soroka, S. (2020). Politicization and Polarization in COVID-19 News Coverage. *Science Communication*, 42(5), 679-697.
- Harteveld, Eelco (2021). Fragmented foes: Affective polarization in the multiparty context of the Netherlands. *Electoral Studies*, 71: <https://doi.org/10.1016/j.electstud.2021.102332>.
- Harteveld, Eelco., Philipp Mendoza, & Matthijs Rooduijn (2021). Affective Polarization and the Populist Radical Right: Creating the Hating?. *Government and Opposition*: <https://doi.org/10.1017/gov.2021.31>
- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1), 77-89.
- Hunter, S. (2014). A Novel Method of Network Text Analysis. *Open Journal of Modern Linguistics*, 4, 350-366.
- Iandoli, L., Primario, S., & Zollo, G. (2021). The impact of group polarization on the quality of online debate in social media: A systematic literature review. *Technological Forecasting & Social Change*, 170.
- Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, J. S. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, 22, 129-146.
- Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, Not Ideology: A Social Identity Perspective on Polarization. *Public Opinion Quarterly*, 76(3), 405-31.
- Jiang, J., Chen, E., Yan, S., Lerman, K., & Ferrara, E. (2020). Political polarization drives online conversations about COVID-19 in the United States. *Human Behavior and Emerging Technologies*, 2(3), 200-211.
- Jones, D. R. (2010). Partisan Polarization and Congressional Accountability in House Elections. *American Journal of Political Science*, 54(2), 323-37.
- Karpf, D. (2020). Two provocations for the study of digital politics in time. *Journal of Information Technology & Politics*, 17(2), 87-96.
- Kerr, J., Panagopoulos, C., & Linden, S. V. D. (2021). Political polarization on COVID-19 pandemic response in the United States. *Personality and Individual Differences*, 179.
- Kevins, A., & Soroka, S. N. (2018). Growing Apart? Partisan Sorting in Canada, 1992-2015. *Canadian Journal of Political Science*, 51(1), 103-133.
- Klein, E., & Robison, J. (2020). Like, Post, and Distrust? How Social Media Use Affects Trust in Government. *Political Communication*, 37(1), 46-64.

- Kubin, Emily, & Christian von Sikorski (2021). The role of (social) media in political polarization: a systematic review. *Annals of the International Communication Association*, 45(3), 188-206.
- Lau, R. R., Andersen, D. J., Ditonto, T. M., Kleinberg, M. S., & Redlawsk, D. P. (2017). Effect of Media Environment Diversity and Advertising Tone on Information Search, Selective Exposure, and Affective Polarization. *Political Behavior*, 39(1), 231-255.
- Lee, C., Shin, J., & Hong, A. (2018). Does social media use really make people politically polarized? Direct and indirect effects of social media use on political polarization in South Korea. *Telematics and Informatics*, 35(1), 245-254.
- Lee, F. (2018). Impact of social media on opinion polarization in varying times. *Communication and the Public*, 1(1), 56-71.
- Levendusky, M. (2009). *The Partisan Sort: How Liberals Became Democrats and Conservatives Became Republicans*, Chicago, IL: University of Chicago Press.
- Levitsky, S., & Ziblatt, D. (2018). *How Democracies Die*, New York: Crown.
- Mallinas, S. R., Maner, J. K., & Plant, E. A. (2021). What factors underlie attitudes regarding protective mask use during the COVID-19 pandemic?. *Personality and Individual Differences*, 181.
- Marchal, Nahema (2021) "Be Nice or Leave Me Alone": An Intergroup Perspective on Affective Polarization in Online Political Discussions. *Communication Research*: <https://doi.org/10.1177/00936502211042516>
- Mason, L. (2015). "I Disrespectfully Agree": The Differential Effects of Partisan Sorting On Social and Issue Polarization. *American Journal of Political Science*, 59(1), 128-45.
- Medeiros, M., & Noël, A. (2014). the Forgotten Side of Partisanship: Negative Party Identification in Four Anglo-American Democracies. *Comparative Political Studies*, 47(7), 1022-46.
- Mehta, V., Caceres, R. S., & CarterEvaluating, K. M. (2014). Topic Quality using Model Clustering. In *2014 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, IEEE, 178-185.
- Milosh, Maria., Marcus Painter., Konstantin Sonin., David Van Dijke., & Austin L. Wright (2021). Unmasking partisanship: Polarization undermines public response to collective risk. *Journal of Public Economics*, 204: <https://doi.org/10.1016/j.jpubeco.2021.104538>.
- Mudde, C. (2016). *On Extremism and Democracy in Europe*. New York, NY: Routledge.
- Mudde, C. (2019). The 2019 EU Elections: Moving the Center. *Journal of Democracy*, 30(4), 20-34.
- Naili, M., Chaibi, A., & Ghezala, H. (2017). Arabic topic identification based on empirical studies of topic models. *ARIMA Journal*, 27, 45-59.
- Nelson, J. L., Lewis, D. A., & Lei, R. (2017). Digital Democracy in America: A Look at Civic Engagement in an Internet Age. *Journalism & Mass Communication Quarterly*, 94(1), 318-334.
- Nordbrandt, Maria (2021). Affective polarization in the digital age: Testing the direction of the relationship between social media and users' feelings for out-group parties. *new media & society*, 1-20:

- <https://doi.org/10.1177/14614448211044393>.
- Pariser, E. (2011). *The Filter Bubble: What the Internet is Hiding from You*. Penguin UK.
- Paz, María Antonia., Ana Mayagoitia-Soria., & Juan-Manuel González-Aguilar (2021) From Polarization to Hate: Portrait of the Spanish Political Meme. *Social Media + Society*: <https://doi.org/10.1177/20563051211062920>.
- Pennycook, G., McPhetres, J., Bago, B., & Rand, D. G. (2021). Beliefs About COVID-19 in Canada, the United Kingdom, and the United States: A Novel Test of Political Polarization and Motivated Reasoning. *Personality and Social Psychology Bulletin*, 1-16.
- Prasetya, H. A., & Murata, T. (2020). A model of opinion and propagation structure polarization in social media. *Computational Social Networks*: <https://doi.org/10.1186/s40649-019-0076-z>
- Roberts, Kenneth M. (2021). Populism and Polarization in Comparative Perspective: Constitutive, Spatial and Institutional Dimensions. *Government and Opposition*: <https://doi.org/10.1017/gov.2021.14>.
- Sebei, H., Taieb, M. A. H., Aouicha, M. B. (2018). Review of social media analytics process and Big Data pipeline. *Social Network Analysis and Mining*, 8(30). <https://doi.org/10.1007/s13278-018-0507-0>
- Tien, J. H., Eisenberg, M. C., Cherng, S. T., & Porter, M. A. (2020). Online reactions to the 2017 'Unite the right' rally in Charlottesville: measuring polarization in Twitter networks using media followership. *Applied Network Science*, 5(10).
- Tufekci, Z. (2014). Big questions for social media big data: Representativeness, validity and other methodological pitfalls. In Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media (ICWSM 2014), 505-514.
- Urman, A. (2019). Context matters: political polarization on Twitter from a comparative perspective. *Media, Culture & Society*, 42(6), 857-879.
- Van Hiel, Alain., Jasper Van Assche., Tessa Haesevoets., David De Cremer., & Gordon Hodson (2021). A Radical Vision of Radicalism: Political Cynicism, not Incrementally Stronger Partisan Positions, Explains Political Radicalization. *Political Psychology*: <https://doi.org/10.1111/pops.12785>.
- Weber, Till (2021). Negative voting and party polarization: A classic tragedy. *Electoral Studies*, 71. <https://doi.org/10.1016/j.electstud.2021.102335>.
- Weil, Audrey M., & Christopher R. Wolfe (2021). Individual differences in risk perception and misperception of COVID-19 in the context of political ideology. *Applied cognitive psychology*, 36(1), 19-31.
- Yang, J. H., Barnidge, M., & Rojas, H. (2017). The politics of "Unfriending": User filtration in response to political disagreement on social media. *Computers in Human Behavior*, 70, 22-29.
- Yarchi, M, Baden, C., & Kligler-Vilenchik, N. (2020). Political Polarization on the Digital Sphere: A Cross-platform, Over-time Analysis of Interactional, Positional, and Affective Polarization on Social Media. *Political Communication*, 38(1-2), 98-139.

Zhu, A. Y. F., Chan, A. L. S., & Chou, K. L. (2019). Creative social media use and political participation in young people: The moderation and mediation role of online political expression. *Journal of Adolescence*, 77, 108-117.

[부록 1] 트위터 계정의 이념적·당파적 입장 분류 기준

□ 이념적 입장: 아래의 기준에 따라 진보, 중도, 보수의 세 가지로 분류

- 본 분석의 목적이 코로나 이슈에 대한 집단별 차이를 보고자 한다는 점에서, 내생성의 오류를 범하지 않기 위해 코로나 관련 트윗 내용을 근거로 집단을 구분하지 않음
- 트윗 메시지의 내용을 통해 쟁점입장에 따라 진보, 중도, 보수로 분류
 - 진보: 환경보호, 동물보호, 페미니즘, 동성애 및 차별금지법 찬성, 검찰개혁, 언론개혁, 세월호 추모, 노동권 보호 등을 주장하는 트윗(리트윗 포함)이 많은 경우, 혹은 일부 있더라도 진보 계정들을 많이 팔로잉한 경우
 - 보수: 기업규제 반대, 종교적 보수주의, 국가안보, 사회질서 및 치안, 동성애 및 차별금지법 반대 등을 주장하는 트윗이 많은 경우, 혹은 일부 있더라도 보수 계정들을 많이 팔로잉한 경우
 - 중도: 특별히 정치적 쟁점에 대한 트윗이 없거나, 이념적 입장을 알기 어려운 경우
- 사적 계정이 아닌 공적(기관 및 공인) 계정의 경우,
 - 언론사는 기존에 알려진 언론사의 이념 성향에 따라 분류(명확하지 않은 경우, 중도로 분류)
 - 주요 정부 부처는 현재 정부의 성향에 따라 진보로 분류하고, 지방자치단체의 경우 현재 지자체장의 당적에 따라 진보 또는 보수로 분류
 - 기타 관공서 및 공공기관의 경우는 중도로 분류
 - 정당과 정치인은 당적에 따라 진보 또는 보수로 분류

□ 당파적 입장: 아래의 기준에 따라 친여당, 중도, 반여당의 세 가지로 분류

- 마찬가지로 코로나 관련 트윗 내용을 근거로 집단 구분을 하지 않음
- 트윗 메시지의 내용을 통해 쟁점입장에 따라 친여당, 중도, 반여당으로 분류
 - 친여당: 문재인 대통령, 현 정부, 여당인 더불어민주당에 우호적인 트윗 메시지(예: 문재인 대통령에 대한 지지와 호의 표현, 현 정부의 정책에 대한 지지 및 호의 표현, 더불어민주당 경선에 대한 관심 및 대권 후보에 대한 지지 및 호의 표현 등)가 많은 경우, 혹은 일부 있더라도 친여 계정들을 많이 팔로잉한 경우
 - 반여당: 문재인 대통령, 현 정부, 여당인 더불어민주당에 반대하거나 반감을 나타내는 트윗 메시지가 많은 경우, 혹은 일부 있더라도 반여 계정들을 많이 팔로잉한 경우
 - 중도: 문재인 대통령, 현 정부, 여당인 더불어민주당에 특별한 입장을 보이지 않거나, 알기 어려운 경우
- 사적 계정이 아닌 공적(기관 및 공인) 계정의 경우,
 - 언론사는 언론사의 당파적 성향에 따라 분류(명확하지 않은 경우, 중도로 분류)
 - 주요 정부 부처는 현재 정부의 성향에 따라 친여로 분류하고, 지방자치단체의 경우 현재 지자체장의 당적에 따라 친여 또는 반여로 분류
 - 기타 관공서 및 공공기관의 경우는 중도로 분류

- 정당과 정치인은 당적에 따라 분류

□ 분류 제외(결측치 처리)

- 다음의 계정은 분석 대상에서 제외함
 - 지나치게 상업적이거나 유해한(음란, 이단종교 등) 계정
 - 한국어를 사용하지 않은 외국 계정
 - 검색일 기준으로 폐쇄되거나 비공개로 전환된 계정

□ 분류 작업 과정

- 2021년 5~6월: 분류 원칙 및 세부기준 마련
- 2021년 7~8월: 연구자(1차 코더)의 분류 작업 수행
- 2021년 9~10월: 연구조교(2차 코더)의 분류 작업 수행, 수차례의 회의(9월 1일, 9월 8일, 9월 27일)를 통해 코더 간 분류기준에 대한 합의 도출
- 2021년 10월: 1차 분류 작업 완료 및 코더간 신뢰도(inter-coder reliability) 수행, 코더 간 상이한 분류에 대해 연구자의 2차 분류 작업을 통해 최종 분류 완료

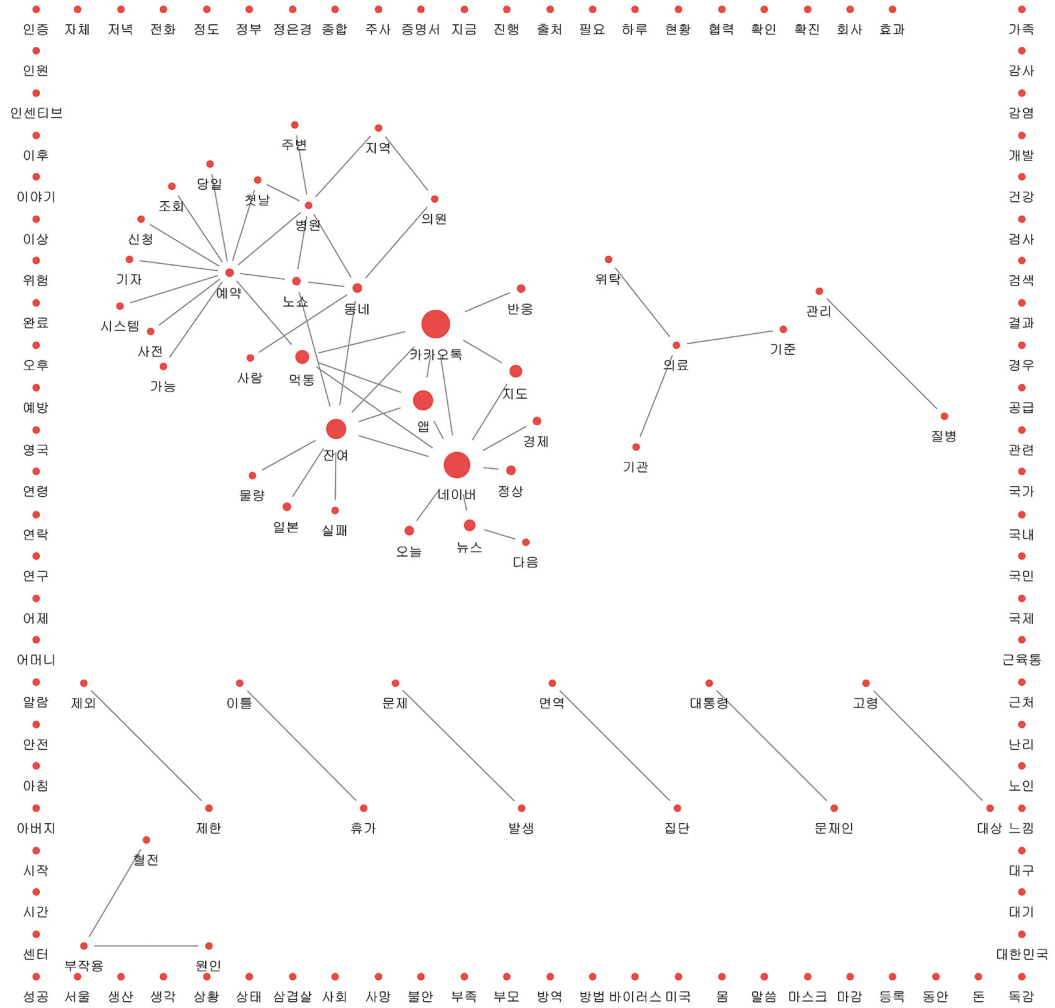
□ 분류의 한계

- 이념과 당파성을 구분하여 집단 분류를 하기는 하였으나, 개념적으로도 양자 간의 명확히 구분되는 것을 보지 못하여, 트윗 내용만으로 양자를 구분하기에는 어려움이 있었음
 - 따라서 당파적 입장의 경우, 명확하게 친여당 혹은 반여당의 입장을 확인할 수 있는 메시지 혹은 팔로잉 관계가 관찰되지 않은 경우, 이념적 입장에 따라 진보는 친여당으로, 보수는 반여당으로 분류하였음
- 중도 집단에는 이념적·당파적 입장이 중도적인 경우, 혼재되어 있는 경우, 그리고 알기 어려운 경우를 모두 포괄하였음
 - 엄밀히 말해 이념적·당파적 입장을 명확히 알기 어려운 경우는 ‘모름’으로 처리하는 것이 맞으나, 그 경우 너무 많은 결측치가 생겨남
 - 이념적·당파적 입장을 명확히 알기 어려운 정도로 그 입장이 강하지 않다고 해석할 수 있다는 점에서 중도 집단으로 분류함
- 이념적·당파적 입장의 ‘강도’는 트윗 메시지에 대한 주관적 판단만으로 정확히 측정할 수 없어 본 분석에서는 고려하지 않음
- 데이터 수집 기간(5월 21일~28일)과 계정의 이념적·당파적 입장에 대한 코딩 기간(7월~9월)이 일치하지 않아 그 사이에 입장의 변화가 있었을 가능성을 배제할 수 없음
 - 하지만 이념적·당파적 입장이 그렇게 단시간에 변하는 성향은 아니란 점에서 정당화될 수 있을 것

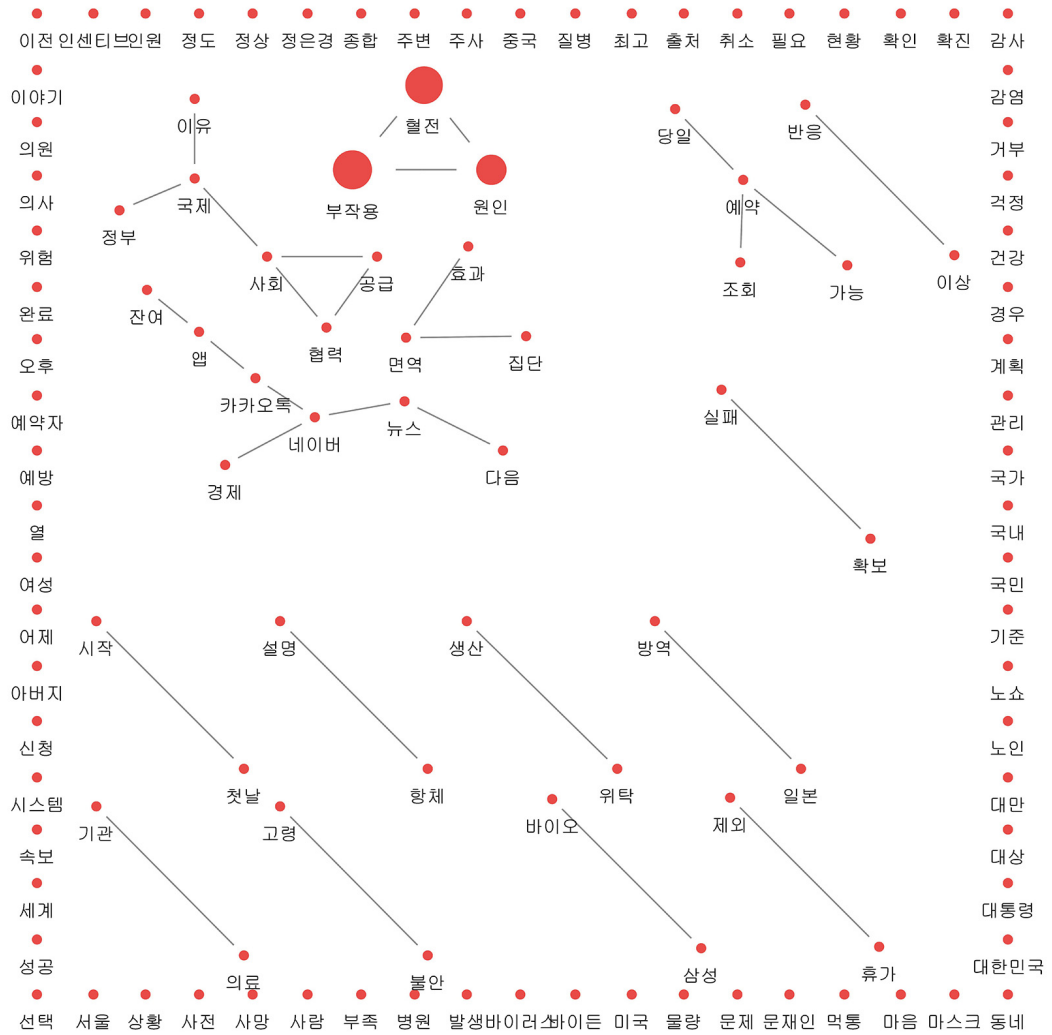
[부록 2] 집단별 아이겐벡터 중심성 수치 상위 20개 단어

순위	이념적 성향						당파적 입장					
	진보		중도		보수		친여		중도		반여	
	단어	중심성 수치	단어	중심성 수치	단어	중심성 수치	단어	중심성 수치	단어	중심성 수치	단어	중심성 수치
1	뉴스	0.703166	카카오톡	0.582415	부작용	0.637859	뉴스	0.703563	카카오톡	0.568807	뉴스	0.631682
2	다음	0.696025	네이버	0.539211	혈전	0.616275	다음	0.695764	네이버	0.501175	네이버	0.599935
3	네이버	0.12217	앱	0.371866	원인	0.459737	네이버	0.123271	잔여	0.416967	다음	0.3638
4	종합	0.049758	잔여	0.349864	면역	0.031486	종합	0.049086	앱	0.357805	카카오톡	0.259233
5	카카오톡	0.037844	먹통	0.186739	집단	0.031058	카카오톡	0.032702	동네	0.152369	경제	0.15353
6	출처	0.027444	지도	0.173861	효과	0.005176	출처	0.029808	지도	0.142161	앱	0.128637
7	협력	0.027239	뉴스	0.12717	뉴스	0.000000	협력	0.024632	예약	0.13244	먹통	0.033177
8	속도	0.022617	오늘	0.083634	네이버	0.000000	속도	0.024543	먹통	0.116334	잔여	0.016462
9	앱	0.011479	동네	0.074136	다음	0.000000	앱	0.008451	병원	0.100793	부작용	0.002372
10	잔여	0.010111	정상	0.062725	카카오톡	0.000000	잔여	0.007905	주변	0.084958	혈전	0.002229
11	지도	0.004769	반응	0.045113	경제	0.000000	동네	0.001392	노쇼	0.06515	원인	0.001692
12	동네	0.002083	노쇼	0.044131	앱	0.000000	사회	0.001059	뉴스	0.061508	국제	0.000051
13	사회	0.00126	예약	0.042507	잔여	0.000000	공급	0.001044	사람	0.05971	사회	0.00005
14	병원	0.001251	경제	0.041817	사회	0.000000	병원	0.000868	일본	0.056295	정부	0.000024
15	공급	0.001063	일본	0.040693	국제	0.000000	걱정	0.000502	오늘	0.055952	면역	0.000021
16	걱정	0.000592	물량	0.027129	공급	0.000000	지역	0.000335	정상	0.040484	집단	0.000021
17	주변	0.000592	실패	0.027129	협력	0.000000	기관	0.000289	시작	0.036688	공급	0.000021
18	지역	0.00046	병원	0.016994	정부	0.000000	국제	0.000127	전화	0.03597	협력	0.000018
19	기관	0.00034	다음	0.01477	이유	0.000000	의원	0.000069	근처	0.033895	이유	0.000008
20	정부	0.000338	첫날	0.011918	예약	0.000000	예약	0.000055	물량	0.033777	대한민국	0.000008

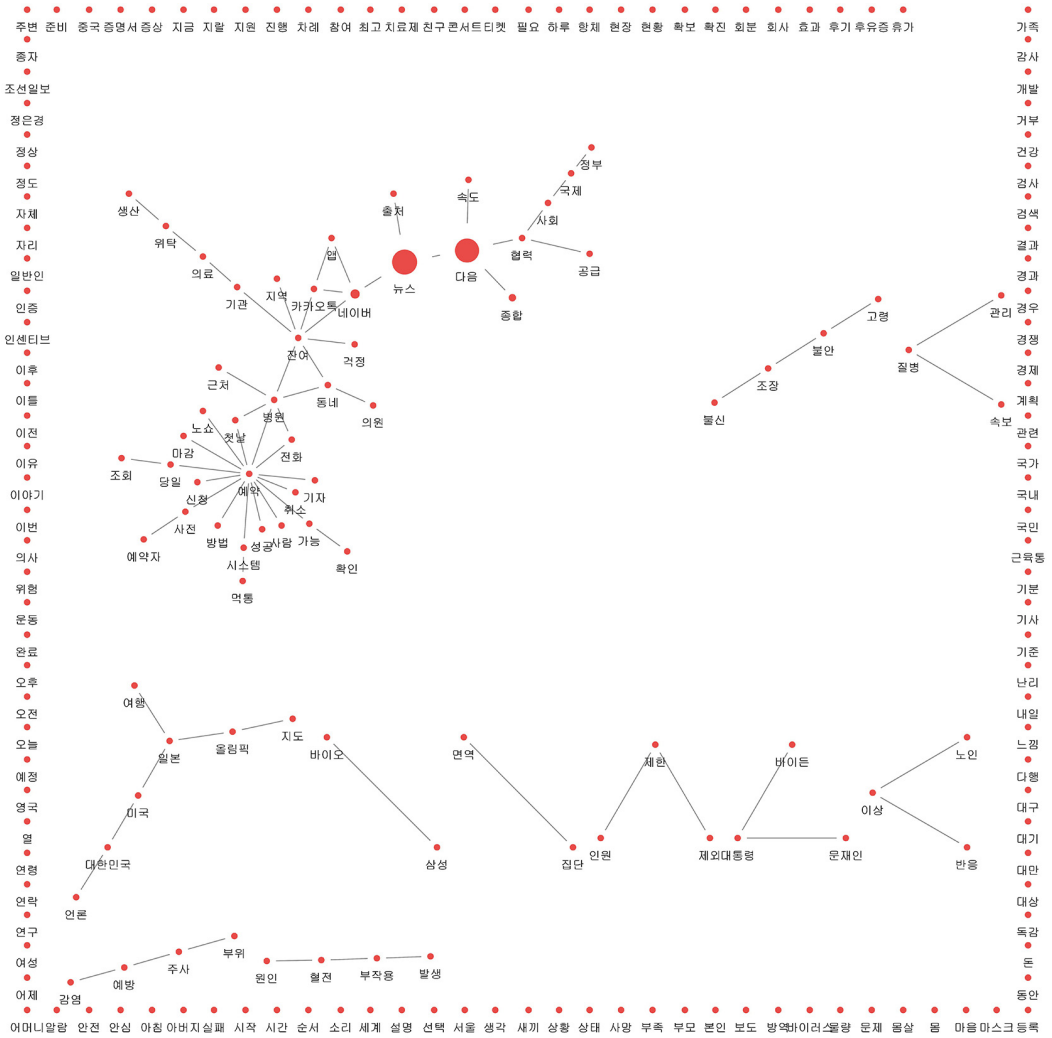
3-2. 이념적 중도 집단



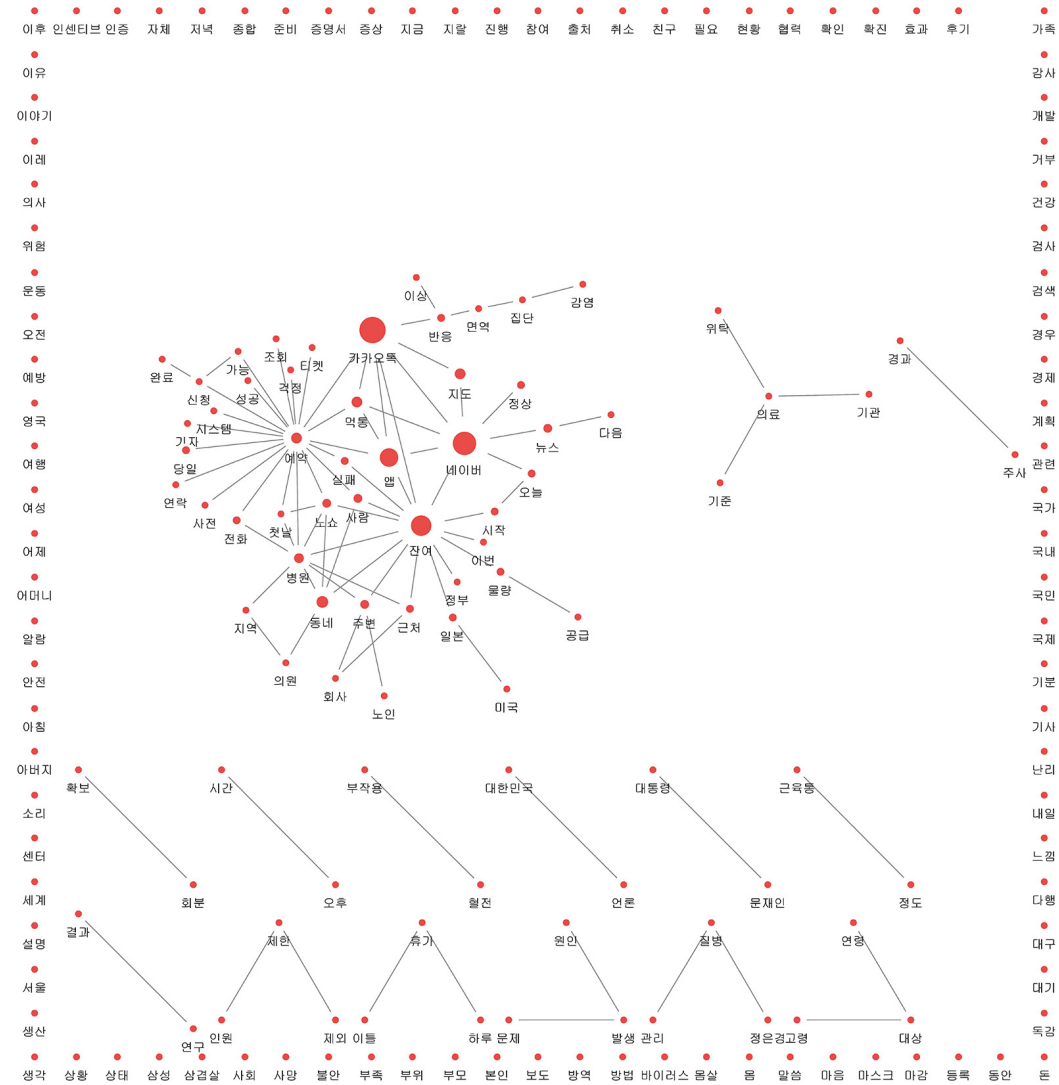
3-3. 이념적 보수 집단



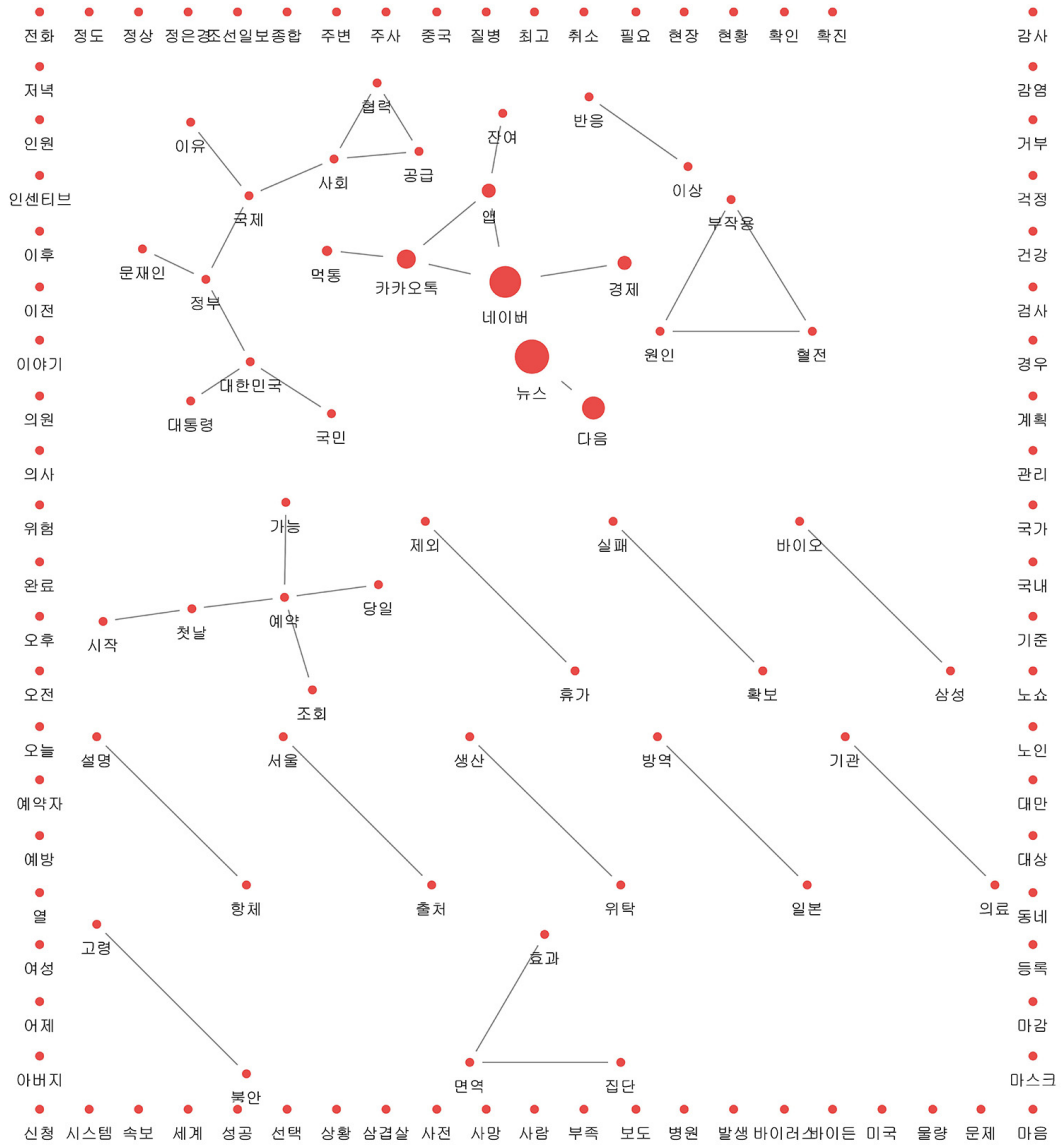
3-4. 당파적 친여 집단



3-5. 당파적 중도 집단



3-6. 당파적 반여 집단



[부록 4] 집단별 토픽 라벨과 토픽별 상위 10개의 키워드

4-1. 이념적 진보 집단

토픽 라벨	문서 수(개)	문서 수(%)	키워드 1	확률	키워드 2	확률	키워드 3	확률	키워드 4	확률	키워드 5	확률	키워드 6	확률	키워드 7	확률	키워드 8	확률	키워드 9	확률	키워드 10	확률
백신 관련 뉴스	808	26.7	뉴스	0.117	다음	0.094	이상	0.039	하루	0.036	국민	0.035	일본	0.033	시작	0.028	불안	0.024	예방	0.021	질병	0.021
대한민국 국가와 정부	591	19.6	대한민국	0.077	정부	0.062	국가	0.058	대통령	0.045	언론	0.042	국민	0.034	미국	0.033	생산	0.031	방역	0.03	세계	0.023
부작용과 이상	565	18.7	사람	0.097	마스크	0.05	부작용	0.049	생각	0.046	이상	0.041	노인	0.038	정도	0.03	어머니	0.028	완료	0.028	반응	0.026
잔여 백신 동네 병원	562	18.6	잔여	0.195	병원	0.089	예약	0.078	사람	0.057	동네	0.046	전화	0.041	신청	0.04	노쇼	0.026	대기	0.026	연락	0.024
잔여 백신 예약 방법	496	16.4	잔여	0.2	예약	0.15	네이버	0.051	가능	0.048	카카오톡	0.045	노쇼	0.042	예방	0.029	확인	0.025	앱	0.024	오늘	0.024

4-2. 이념적 중도 집단

토픽 라벨	문서 수(개)	문서 수(%)	키워드 1	확률	키워드 2	확률	키워드 3	확률	키워드 4	확률	키워드 5	확률	키워드 6	확률	키워드 7	확률	키워드 8	확률	키워드 9	확률	키워드 10	확률
사람과 마스크	229	40.2	사람	0.119	마스크	0.066	생각	0.043	주사	0.042	이야기	0.035	가족	0.026	어머니	0.024	시간	0.022	의사	0.021	국민	0.021
예약 가능 잔여 백신	103	18.1	잔여	0.2	예약	0.111	가능	0.09	오늘	0.06	신청	0.06	동네	0.039	휴가	0.031	하루	0.023	확인	0.02	일본	0.02
부작용 문제와 뉴스	87	15.3	뉴스	0.108	확진	0.052	부작용	0.051	문제	0.043	다음	0.04	생각	0.032	감염	0.03	출처	0.028	대한민국	0.028	혈전	0.025
잔여 백신 예약 방법	76	13.4	잔여	0.152	예약	0.136	카카오톡	0.088	네이버	0.088	병원	0.077	노쇼	0.067	앱	0.06	첫날	0.032	먹통	0.029	조회	0.021
코로나 예방과 이상	74	13.0	시작	0.106	이상	0.087	예방	0.064	질병	0.039	연령	0.035	대상	0.03	의료	0.025	증명서	0.025	관리	0.024	기준	0.023

4-3. 이념적 보수 집단

토픽 라벨	문서 수(개)	문서 수(%)	키워드 1	확률	키워드 2	확률	키워드 3	확률	키워드 4	확률	키워드 5	확률	키워드 6	확률	키워드 7	확률	키워드 8	확률	키워드 9	확률	키워드 10	확률
혈전 등 백신 부작용	52	23.9	부작용	0.089	확보	0.078	정부	0.077	혈전	0.074	사회	0.074	협력	0.055	원인	0.051	국제	0.045	이유	0.041	공급	0.041
마스크와 백신 관련 뉴스	51	23.4	마스크	0.156	뉴스	0.057	출처	0.054	국민	0.039	사람	0.038	위험	0.031	이야기	0.03	다음	0.029	감염	0.029	필요	0.025
대한민국과 각국의 백신 상황	44	20.2	미국	0.18	대한민국	0.077	국가	0.077	중국	0.072	이상	0.065	인센티브	0.04	상황	0.036	지원	0.034	문재인	0.034	정상	0.018
코로나바이러스 집단 면역	37	17.0	국민	0.112	면역	0.099	집단	0.076	바이러스	0.055	문재인	0.041	의사	0.035	효과	0.032	인센티브	0.031	이상	0.03	정도	0.027
잔여 백신 예약 방법	34	15.6	예약	0.116	잔여	0.11	네이버	0.047	뉴스	0.042	출처	0.035	카카오톡	0.034	가능	0.032	시작	0.03	당일	0.025	첫날	0.024

4-4. 당파적 친여 집단

토픽 라벨	문서 수(개)	문서 수(%)	키워드 1	확률	키워드 2	확률	키워드 3	확률	키워드 4	확률	키워드 5	확률	키워드 6	확률	키워드 7	확률	키워드 8	확률	키워드 9	확률	키워드 10	확률
잔여 백신 병원 예약	586	24.4	잔여	0.223	예약	0.113	사람	0.08	병원	0.07	노쇼	0.056	신청	0.046	동네	0.044	전화	0.031	대기	0.025	지역	0.023
잔여 백신 예약 방법	495	20.6	잔여	0.151	예약	0.107	가능	0.066	네이버	0.054	카카오톡	0.043	일본	0.03	확인	0.027	앱	0.026	올림픽	0.026	당일	0.026
백신 관련 뉴스와 불안	474	19.7	뉴스	0.128	다음	0.095	불안	0.041	사람	0.04	국민	0.035	기사	0.033	하루	0.033	언론	0.032	질병	0.031	생각	0.029
백신 이상과 코로나 예방	450	18.7	이상	0.083	예방	0.041	마스크	0.041	노인	0.04	시작	0.038	반응	0.027	완료	0.026	어머니	0.023	부작용	0.023	확진	0.021
대한민국 국가와 정부	397	16.5	대한민국	0.078	국가	0.068	정부	0.065	대통령	0.054	미국	0.044	국민	0.037	생산	0.032	방역	0.029	세계	0.026	확보	0.024

4-5. 당파적 중도 집단

토픽 라벨	문서 수(개)	문서 수(%)	키워드 1	확률	키워드 2	확률	키워드 3	확률	키워드 4	확률	키워드 5	확률	키워드 6	확률	키워드 7	확률	키워드 8	확률	키워드 9	확률	키워드 10	확률
잔여 백신 예약 방법	391	34.8	잔여	0.204	예약	0.192	카카오톡	0.076	노쇼	0.071	네이버	0.067	앱	0.043	오늘	0.034	첫날	0.027	시스템	0.022	병원	0.021
동네 신청 가능 잔여 백신	197	17.5	잔여	0.23	신청	0.084	가능	0.059	동네	0.047	일본	0.044	하루	0.039	주변	0.037	휴가	0.037	국민	0.028	미국	0.027
사람과 마스크	189	16.8	사람	0.101	마스크	0.072	뉴스	0.068	다음	0.058	부작용	0.051	확진	0.026	대한민국	0.025	방역	0.025	정부	0.024	상황	0.022
백신 접종 병원	188	16.7	병원	0.11	생각	0.059	전화	0.047	주사	0.044	연락	0.037	어머니	0.036	시간	0.036	사람	0.03	일반인	0.029	정도	0.029
백신 이상과 예약	160	14.2	시작	0.085	이상	0.063	예약	0.053	사람	0.044	노인	0.038	예방	0.035	대상	0.032	질병	0.029	대한민국	0.027	연령	0.026

4-6. 당파적 반여 집단

토픽 라벨	문서 수(개)	문서 수(%)	키워드 1	확률	키워드 2	확률	키워드 3	확률	키워드 4	확률	키워드 5	확률	키워드 6	확률	키워드 7	확률	키워드 8	확률	키워드 9	확률	키워드 10	확률
혈전 등 백신 부작용	68	24.1	부작용	0.127	혈전	0.092	문제	0.066	원인	0.061	대상	0.061	사람	0.059	연구	0.051	마스크	0.036	이틀	0.03	사망	0.027
정부의 백신 확보	56	19.9	국민	0.12	확보	0.073	사회	0.073	면역	0.068	정부	0.064	집단	0.056	협력	0.053	방역	0.045	국제	0.045	이유	0.044
잔여 백신 예약 방법	56	19.9	잔여	0.107	예약	0.1	뉴스	0.083	출처	0.066	네이버	0.062	다음	0.037	카카오톡	0.035	시작	0.034	정은경	0.025	첫날	0.024
대한민국과 각국의 백신 상황	51	18.1	미국	0.139	마스크	0.099	대한민국	0.098	중국	0.095	문제인	0.073	정부	0.045	대만	0.033	상황	0.024	일본	0.018	정상	0.018
백신 이상 반응과 국가 정책	51	18.1	이상	0.107	국가	0.078	인센티브	0.059	반응	0.032	생각	0.03	사람	0.027	완료	0.025	지원	0.024	기준	0.023	이야기	0.02

Political Polarization on Social Media Conversations about COVID-19 Vaccination: Evidence from the Word Network Analysis and Topic Modeling of Twitter Messages in South Korea

Dong-Joon Jung

Inha University

The COVID-19 pandemic, an apolitical issue, has been politicized across many countries since 2020. This politicization of COVID-19 aggravates the current political polarization of many democracies, which is regarded as one of the main culprits of today's crisis of democracy. This paper explores how discussions on COVID-19 vaccination in South Korea differ between ideological and partisan groups by focusing on social media, today's main channel of the distribution of political discourse. From the results of the word network centrality analysis and topic modeling of Twitter messages, this study finds that news on vaccination and information on vaccine reservation are the most often discussed topics among the liberal and pro-government groups, while vaccine side-effects and blood clots are reported as the prominent issues among the conservative and anti-government groups. For the anti-government group specifically, concerns about the government's vaccine policies are also widely discussed.

Keywords: COVID-19 Vaccination, Social Media, Political Polarization, Text Network Analysis, Topic Modeling