

코로나19 사태와 온라인 정보의 다양성 연구 - 빅데이터를 활용한 글로벌 접근법

박한우^{1*}, 김지은², 주우봉²

¹영남대학교 언론정보학과, 디지털융합비즈니스학 및 동아시아문화학 협동과정 대학원,
²영남대학교 디지털융합비즈니스학 대학원, 사이버감성연구소

Online Information Sources of Coronavirus Using Webometric Big Data

Han Woo Park^{1*}, Ji-Eun Kim², Yu-Peng Zhu²

¹Interdisciplinary Graduate Programs of Digital Convergence Business and East Asian Cultural Studies,
Department of Media and Communication, YeungNam University

²Cyber Emotions Research Institute, YeungNam University

요약 이 연구는 웹보메트릭 빅데이터를 활용하여 코로나바이러스 확진 국가(지역)들의 온라인 정보원의 다양성을 조사했다. 구체적으로 2020년 2월에 코로나바이러스 확진자가 발생한 28개국을 대상으로 웹 데이터를 수집한 결과, 호주, 캐나다, 이탈리아 등의 온라인 가시성이 높게 나타나면서 관련 정보를 가장 많이 생산하고 있었다. 국가별 검색건수(hit counts)와 정보채널의 역할을 하는 도메인(domain) 빈도와는 통계적으로 유의한 상관성이 있었다. 한편 데이터 수집도구인 bing.com의 점유율이 평소에도 높은 국가들을 제외하고 다시 검토한 결과, 당시 확진자 수가 많았던 일본, 중국, 싱가포르 등이 코로나바이러스와 관련된 웹데이터를 주도적으로 올리고 있었다. 온라인 정보원은 n-헬릭스를 활용하여 분류되었다. n-헬릭스는 대학-기업-정부의 3주체간 상호작용과 혁신을 강조하는 트리플헬릭스모델(Triple Helix Model)에 기반한 확장된 분석틀이다. 그 결과, 정부기관이 18.1%를 차지하면서 코로나바이러스 정보의 최대 공급자로 나타났다. 2원성 네트워크 분석결과를 보면 언론사, 대학병원, 공중보건에 특화된 조직 등도 코로나바이러스 연구와 방역 정보의 온라인 유통에 적극적이었다. 웹페이지에 포함된 단어들을 중심으로 내용분석을 해 보니 건강, 학교, 가족, 공공, 방안 등의 단어가 중심성이 높게 나타나 코로나바이러스로 인한 개인별 예방수칙뿐만 아니라 생활 불편과 업무 장애로 인한 대처방안 등에 관심이 높다는 것을 알 수 있었다.

Abstract Using webometric big data, this study examines the diversity of online information sources about the novel coronavirus causing the COVID-19 pandemic. Specifically, it focuses on some 28 countries where confirmed coronavirus cases occurred in February 2020. In the results, the online visibility of Australia, Canada, and Italy was the highest, based on their producing the most relevant information. There was a statistically significant correlation between the hit counts per country and the frequency of visiting the domains that act as information channels. Interestingly, Japan, China, and Singapore, which had a large number of confirmed cases at that time, were providing web data related to the novel coronavirus. Online sources were classified using an N-tuple helix model. The results showed that government agencies were the largest supplier of coronavirus information in cyberspace. Furthermore, the two-mode network technique revealed that media companies, university hospitals, and public healthcare centers had taken a positive attitude towards online circulation of coronavirus research and epidemic prevention information. However, semantic network analysis showed that health, school, home, and public had high centrality values. This means that people were concerned not only about personal prevention rules caused by the coronavirus outbreak, but also about response plans caused by life inconveniences and operational obstacles.

Keywords : Coronavirus, Disaster, Big data, Webometrics, Media, Network Analysis, Triple Helix, n-Helix

본 논문은 2020년 방송문화진흥회 방송연구지원사업 연구과제로 수행되었음.

*Corresponding Author : Han Woo Park(Yeung-nam Univ.)

email: hanpark@ynu.ac.kr

Received July 13, 2020

Revised August 18, 2020

Accepted November 6, 2020

Published November 30, 2020

1. 서론

1.1 연구의 배경

오늘날 지식기반 사회에서 정보생산은 전통적 언론기관을 넘어서 여러 주체들이 참여하고 있다. 정보생산 과정이 복잡해짐에 따라 소위 '레거시'(legacy) 미디어의 이용시간이 지속적으로 감소하면서 전통적 언론사들은 여러 정보원들과 경쟁하는 상황에 직면해 있다. 특히 속보를 다루는 호우, 태풍, 폭설, 지진 등의 자연재해뿐만 아니라 안전사고와 전염병 등 사회적 재난이 발생하면 경쟁상황은 훨씬 더 심화되었다. 소셜미디어, 정부기관, 학·협회·단체, 지자체, 공공연구소 등은 매스미디어의 취재원이 되기도 하지만 그들 스스로가 정보를 생산하고 유통하는 주체가 되기에 상호 경쟁하는 체제가 형성된다.

그러나 지금까지는 대체재와 보완재, 주류미디어와 대안미디어, 매스미디어와 틈새미디어, 레거시와 뉴미디어 등의 이분법적 구도로 이러한 경쟁상황을 협소한 시각에서 접근해 온 경향이 강하다. 최대 다수의 이용자 만족을 목표로 하는 공리주의 시각이나, 서로 다른 의견의 수용이 민주적 사회를 견인한다는 입장에서 본다면 오늘날 우리가 목격하는 정보(원)의 다양화는 좀 더 열린 관점에서 분석하는 것이 더 적합할 수 있다. 더군다나 연구를 진행하는 현재 전 세계를 위협하는 코로나 바이러스에 대해서도 보편적이면서 다양한 콘텐츠를 제공하는 주체가 많을수록 사회적 면역체계는 더 강해질 수 있다.

따라서 이 연구는 코로나19 확진자가 발생한 국가들에서 온라인 출현도가 높은 정보원들을 빅데이터 기법을 통해서 수집한다. 대규모의 온라인 빅데이터에서 전통적 THM인 학계-비즈니스-정부(ABG, Academy-Business-Government)에 대한 일반화된 분석을 벗어나 보건기관, 언론사 등을 추가해서 다차원화적인 정보생산 환경을 규명하고자 한다. 이 연구는 긴급한 재난 상황에서 주요 국가의 사이버공간 정보전달원의 나선구조에 대한 체계적인 설명을 해 줄 수 있다는 측면에서 학술적, 사회적 가치가 높다.

2. 문헌검토

2.1 트리플헬릭스모델

이노베이션에 방점을 둔 트리플헬릭스모델이 인접 학문 분야에 알려지면서 여러 문제의식을 지닌 논문들이

발행되고 있다. 트리플헬릭스 모델은 대학, 기업, 정부의 네트워크를 중심으로 이루어진 삼중나선형의 움직임을 국가의 지식생산으로 파악한다[1]. 트리플헬릭스모델은 크게 두 분야로 나누어진다. WISDT (Webometrics, Informetrics, Scientometrics, Datametrics, Technometrics)에 중점을 두는 방법론주의 활동과 IISTS (Innovation In Science, Technology, and Society) 모델의 개발을 더 중시하는 이론주의 활동으로 나누어진다[2-3]. 본 논문에서는 사회 내외부의 '데이터', '정보', '지식', '지혜'의 '상호작용' 및 '관계교류'를 측정하는데 더 관심을 가지는 방법론과 궤를 같이 한다. 즉 코로나바이러스와 관련된 온라인 정보원들의 생태계와 거버넌스를 빅데이터를 기반으로 관계 형성, 유지, 개선 등을 찾고자 한다.

지진, 태풍과 같은 자연재해뿐만 아니라 메르스 사태, 원자력 사고 등과 같은 일련의 재난 상황들을 겪으면서 학계에서는 재난과 관련된 언론 보도와 그 효과에 관한 연구들을 지속적으로 진행해 왔다. 또한 전통적인 미디어에서 온라인 기반 미디어의 이용이 증가함에 따라 온라인 데이터를 활용한 연구들도 점차 늘어나는 추세이다.

2.2 재난 보도 연구

권호천은 조선일보와 한겨레신문 기사를 중심으로 메르스 상황에 대한 신문보도 의미연결망을 분석하였다[4]. 매체별로 신문기사에 나타난 주요한 단어를 찾고 단어들 이 구성하는 네트워크의 구조적인 특징을 살펴보았다. 이를 통해 이념적 성향을 달리하는 언론사의 보도에 따라 국민들의 인식과 대처의 정도가 어떠한 차이를 보이는지 알아보하고자 했다. 이미나와 홍주현은 메르스의 확산에 따라 정부가 실시한 위기 대응 메시지를 대상으로 네트워크 분석을 실시했다[5]. 메르스의 확산 양상을 토대로 3 단계로 시기를 나누고 중앙정부와 지방자치정부의 공식 홈페이지에 게시된 메르스 관련 보도자료에 대해 분석했다. 진나영과 정정주는 한국의 메르스 상황에 대한 국내외 언론의 보도형식과 내용, 의미구조에 대한 의미망 분석을 실시했다[6]. Song 등 학자는 메르스 관련 정보를 전달하는 국내 온라인 채널-게시판, 트위터, 카페, 블로그, 뉴스에 따라 사람들이 느끼는 감성의 차이와 확산범위에 대해 연구했다[7].

선행된 연구들은 대부분 재난과 관련된 언론 보도에 대해 사회 네트워크 접근을 시도했다. 텍스트 의미구조에 대한 의미망 분석을 통해 재난 발생 시 국가 내에서 이루어

어지는 언론 보도의 방향과 프레임에 대한 분석 및 언론 보도에 따른 대중의 재난 인식에 미치는 효과를 파악하는 것에 치중되어 있다. 해외 정보원을 대상으로 분석을 진행한 연구에서도 국내와 해외의 언론 보도 차이를 비교하는 분석이 주로 이루어졌으며, 언론 이외 정보를 제공하는 주체의 다양성과 그 콘텐츠에 관한 연구는 미비하였다.

코로나바이러스와 같은 세계적인 재난 상황을 파악하고 대처하기 위해 본 논문에서는 국내뿐만 아니라 코로나바이러스 확진자가 발생한 여러 국가들을 대상으로 언론뿐만 아니라 얼마나 다양한 주체들이 재난과 관련된 정보를 제공하는가에 대해 알아보고자 한다. 뿐만 아니라 코로나 바이러스를 둘러싼 국가들 간에 글로벌 정보네트워크를 그려 관계적 특징을 살펴보고 이러한 현상에 n-Tuple 헬릭스 모델을 적용함으로써 국가의 사이버공간 정보전달원의 구조를 설명한다는 앞선 연구들과는 차이가 있다.

2.3 웹 정보원

인터넷이 급성장함에 따라 웹 데이터 분석은 양적 확대함과 내용적 다양성까지 중요해지고 있다. 특히 웹 정보원에 대한 분석은 여러 가지 유용한 사항과 다양한 측면을 발굴할 수 있다. 나아가 웹에서 얻는 정보는 일반적으로 주관적 요소의 개입이 적어서, 자기응답식(self-reporting) 전통적 연구방법에 비해 편견이나 과장될 위험성이 낮을 뿐만 아니라 해당 맥락을 제공할 가능성이 높다. 그리고 이전에는 알지 못했던 새로운 지식을 발견할 수 있는 디지털 세렌디피티(serendipity)도 존재한다. 예를 들어 Acharya와 Park은 네팔의 공공 데이터 관련된 웹 정보원 분석결과 국제기구 간 긴밀성이 다른 정부조직, 상업조직 등 비국제기구간의 긴밀성보다 더 큰 것으로 나타났다[8]. Ying 등은 Charleston의 여행에 관한 웹 정보원을 바탕으로 하이퍼링크 네트워크를 분석하여 여행업계 내의 네트워크가 성긴 것을 발견하였으며, 업종마다 이해관계자들의 온라인 행위 차이가 있었다는 것을 찾았다[9]. Park과 Park이 암호화폐의 정보원 네트워크 분석을 통해 발견한 것은 암호화폐와 관련된 정보생산과 소비가 유럽과 아시아에 집중되어 있으며, 최상위 도메인 수와 소셜네트워크 중심도 지표가 암호화 시가총액, 거래량에 관련되어 있다는 것이다[10]. 이처럼 웹 정보원에 대한 분석과 연구는 이미 여러 분야에서 유용한 결론을 도출하고 있다.

3. 연구방법

3.1 데이터 수집

이 논문 데이터 수집대상은 국제적으로 가장 많이 통용된 코로나바이러스를 키워드로 2020년 2월 당시 확진자가 발생한 28개 국가 및 지역이다(이하의 국가 및 지역은 국가로 표시했다). 해당 시기는 일본의 크루즈에서 확진자가 대거 발생한 시점이며, 우리나라는 대구의 신천지 신도로 인한 코로나바이러스 확산이 발생하기 이전이다. 데이터는 웹보메트릭 분석기(Webometric Analyst) 4.1을 이용하여 수집하였다. 웹보메트릭 분석기를 사용하여 데이터를 수집한 채널은 bing.com이다. 이 논문에서는 bing.com의 검색 에이피아이(API) 서비스를 통해 얻은 9149개의 웹페이지(URL 수)와 7003개의 도메인을 분석 대상으로 하였다.

3.2 분석방법과 소프트웨어

본 논문에서는 트리플헬릭스모델의 여러 방법론 가운데 웹보메트릭스를 활용한다. 웹보메트릭스란 웹(web)과 계량분석(metrics)의 합성어로 인터넷의 등장과 함께 선보인 새로운 방법론이다[11]. 웹보메트릭스를 통해 수집한 히트수와 도메인 수 분석을 통해 코로나바이러스에 대한 해당 국가의 관심도를 파악할 수 있다.

히트 수는 웹상에서 해당 키워드를 포함한 모든 페이지 수를 의미하므로 특정 국가의 코로나바이러스 관련 히트수가 높다는 것은 웹상에서 코로나바이러스에 대한 언급이 많이 일어났다는 것을 의미한다. 또한 도메인은 숫자로 된 IP 주소를 대신하는 문자로 된 인터넷상의 주소로써[12] 특정 국가의 코로나바이러스 관련 도메인 수가 많다는 것은 해당 국가에 코로나바이러스가 언급되고 있는 온라인 공간이 많이 존재한다는 것을 의미한다. 도메인을 좀 더 구체적으로 살펴보면 TLD(Top Level Domain, 최상위 도메인)로부터 하위의 도메인(subdomains, 2단계3단계 도메인)으로 구성되며 TLD는 .edu, .gov, .com, .org와 같은 gTLD(generic Top Level Domain: 일반 최상위 도메인)와 .au(Australia), .us(United States), .jp(Japan), .kr(South Korea)과 같은 ccTLD(country code Top Level Domain: 국가 코드 최상위 도메인)로 구분된다[9]. 이 논문에서는 TLD 중에서 ccTLD만을 이용해 코로나바이러스 확진자가 발생한 국가들의 히트수와 도메인 수를 수집 및 분석하였다.

미국, 프랑스, 캐나다 등 sTLD(Second Top-Level domain)가 존재하지 않는 국가를 제외한 나머지 국가에서는 sTLD가 gTLD와 같은 .edu, .gov, .com, .org 등의 형식으로 나타나는데 sTLD를 통해 정보를 제공하는 정보원을 파악할 수 있다. sTLD를 분류함으로써 국가별 정보원의 다양성을 파악하고 더불어 sTLD 빈도를 분석하면 코로나바이러스와 관련된 정보를 어떤 정보원들이 얼마나 생산해내는지 추정할 수 있다.

본 논문에서는 코로나바이러스와 관련된 정보를 생산하는 정보원 간의 관계를 파악하기 위해 트리플헬릭스 모델의 주된 분석대상인 Academy(교육)-Business(기업)-Government(정부)에서 더 나아가 n-Tuple 헬릭스 모델을 적용하여 도메인을 Academy(교육), Business(기업), Government(정부), Media(언론), Health(보건), Other(기타) 6개의 그룹으로 분류했다. 분류한 6개의 그룹과 코로나바이러스 확진 국가와의 연결을 2-mode matrix(이원성 행렬)로 구성하여 국가별 정보원을 파악했다. 2-mode matrix는 매트릭스의 행과 열에 같은 개체가 배열되는 1-mode matrix와 달리 매트릭스의 행과 열에 서로 다른 개체가 배열된다[13]. 행과 열에 각각 6개의 그룹과 확진 국가를 두고 특정 그룹에 해당하는 해당 국가의 도메인들의 URL수를 입력했다.

코로나 바이러스에 대한 핵심 정보원뿐만 아니라 국가별로 제공하는 정보의 주제와 내용도 해당 국가의 사회문화적 배경과 보건커뮤니케이션 상황에 따라 다양하게 나타날 수 있다. 국가별 정보의 차이를 분석하기 위해 국가별로 사용하는 단어의 빈도분석을 실시했다. 단어가 출현하는 횟수가 많을수록 해당 단어들에 대한 관심도와 주목도가 높게 된다. 물론 빅데이터는 언급 빈도의 단순 집계와 통계적 분포뿐만 아니라 해당 단어들이 사용된 정서적 맥락과 제3의 단어와 가지는 관계성 등도 고려한다. 흔히 이러한 기법을 의미망(semantic network analysis)이라고 부른다. 텍스트에 포함된 구문의 요소에서 특정한 의미를 찾아 분석하는 방법이다. 이를 통해 전염병 위기상황에서 각국이 제공하는 정보원을 비교하고 나아가, 주제와 내용적 유사성과 차별성을 찾고자 한다.

4. 연구결과

코로나 확진 국가 중에서 히트수(hit counts)가 가장 많았던 국가는 호주(7980000)이다. 다음으로 캐나다(2590000), 이탈리아(1970000), 영국(1910000), 독일

(1680000), 프랑스(1460000), 스페인(1090000), 벨기에(831000), 일본(582000), 인도(478000), 중국(448000), 싱가포르(426000) 순서로 많았다.

2017년 Microsoft advertising의 자료[14]에 따르면 bing.com은 국가별로 검색엔진에서 미국 33%, 대만 26%, 영국 23%, 홍콩 20%, 프랑스 19%, 노르웨이 18%, 캐나다 16%, 독일 13%, 호주 12%, 스웨덴 12%, 스페인 10%, 이탈리아 9%의 점유율을 가지는 것으로 나타났다. 코로나 확진 국가 중에서 확진자의 수가 비교적 많지 않음에도 캐나다, 이탈리아, 영국과 같은 국가들이 상위권에 나타나는 것은 그 까닭이다.

미국의 경우 bing.com의 점유율은 높지만 미국으로 간주되는 gTLD인 .edu, .gov, .mil을 포함하지 않고 ccTLD로만 히트수를 수집하였기 때문에 등수가 낮게 나타났다. 호주가 1위를 차지한 배경은 bing.com의 점유율이 높은 집단에 속할 뿐만 아니라 중국이외의 최초로 코로나바이러스 배양에 성공했기 때문으로 추정된다. 이를 통해 코로나바이러스와 관련된 백신 개발의 희망을 가지게 되면서 관심이 증가한 것으로 보인다. bing.com의 점유율이 높은 집단을 제외하면 당시 확진자 수가 많았던 일본, 중국, 싱가포르가 그다음 순위로 나타나고 있다.

Fig. 1은 왼쪽부터 코로나바이러스 확진 국가들의 히트수와 도메인 수를 각각 지도에 표현한 것이다. 히트수의 특성상 특정 웹페이지에서 코로나바이러스에 대한 언급이 많아지면 히트수가 높아질 수 있다. 반면 도메인의 경우 코로나바이러스가 언급된 온라인 상의 실제 공간을 의미하므로 히트수와 도메인 수 사이의 차이가 존재할 수 있지만, 코로나바이러스 확진 국가들의 히트 수와 도메인 수 간에는 큰 차이가 없었다. 피어슨(Pearson) 상관분석에 따른 결과는 Table 1이다.

Table 1. Correlations of hitcounts and domains

	Hitcounts	Domains
Hitcounts	1	.564**
Domains		1

**p<0.01

코로나바이러스를 언급한 도메인의 수가 많을수록 코로나바이러스에 대한 해당 국가의 관심도는 높다고 간주할 수 있다. 왜냐하면 코로나바이러스가 언급되고 있는 온라인 공간이 그만큼 많이 존재한다는 것을 의미하기 때문이다. 예를 들면 일본은 644개의 도메인에서 코로나바이러스가 언급되어 2위를 차지했다. 데이터를 수집 할

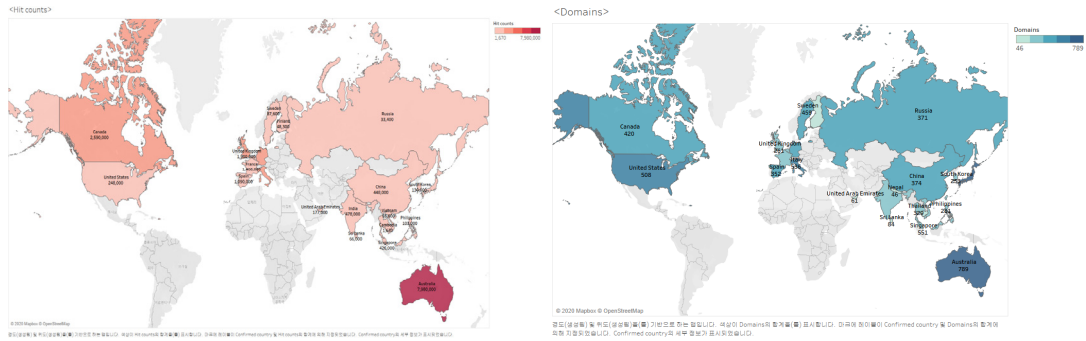


Fig. 1. Hit counts and number of domains of confirmed country

당시 일본의 크루즈에서 많은 수의 확진자가 발생한 것이 영향을 준 것으로 보인다. 뒤를 이은 싱가포르와 홍콩은 당시 확진자 수가 가장 많았던 중국과 일본 다음으로 코로나 확진자가 많이 발생했던 국가이다. 반면 한국은 신천지 교인으로 인한 코로나 확산이 발생하기 이전으로 253개의 도메인만이 링크되며 비교적 하위권에 머물렀다.

Table 2에서는 코로나바이러스 확진자 발생 국가들의 sTLD(Second Top-Level domain)를 분류하고 sTLD의 빈도분석을 통해 정보원별로 코로나바이러스와 관련된 정보의 생산 정도를 나타냈다. 코로나바이러스 키워드가 많이 나타난 sTLD의 순서는 다음과 같다. .gov (총 7003개의 도메인 가운데 1037개, 14.8%)가 가장 많았으며 .com(632개, 9.0%), .edu(535개 7.61%), .co(358개, 5.1%), .ac(276개, 3.9%), .org(266개, 3.8%), .go(194개, 2.8%), .or(87개, 1.2%) 순서였다.

sTLD 분석결과에서 눈에 띄는 몇 가지의 sTLD들이 있다. 먼저, .lg와 .ne이다. 이 sTLD들은 일본에서 사용하는 sTLD로 각각 Local Government Authorities와 Network Service Providers를 나타낸다. 예를 들면, 요코하마시의 홈페이지 www.city.yokohama.lg.jp와 후쿠오카시의 홈페이지 www.pref.fukuoka.lg.jp, 일본의 포털사이트 news.goo.ne.jp 등에서 사용되는 것을 볼 수 있다. .nhs는 영국에서 사용되며 National Health Service를 나타내기 때문에 www.nnuh.nhs.uk와 www.mkuh.nhs.uk처럼 대학병원 등의 도메인에 쓰이는 모습을 볼 수 있다. .gob는 Government를 나타내는 스페인의 sTLD로써 대사관 www.exteriores.gob.es이나 정부 도메인www.lamoncloa.gob.es을 예시로 들 수 있다.

정부를 나타내는 sTLD는 .gov, .go, .lg, .gob 등으로 도메인 총 1266개, 18.1% 비율로 가장 많았다. 그다음으

로는 기업을 나타내는 .com, .co가 990개, 14.1%, 교육을 나타내는 .edu, .ac가 811개, 11.5%의 비율을 차지했다. 이를 통해 전염병과 같은 국가적인 재난 상태에서는 정부가 정보생산자로서 중요한 역할을 한다는 것을 알 수 있다. 그러나 정부뿐만 아니라 기업과 교육 분야도 정부 못지않은 영향력과 역할을 가지고 있다는 것을 알 수 있다.

Table 2. sTLD rankings related to Coronavirus

STLD	Domains	%
gov	1037	14.8
com	632	9.0
edu	535	7.6
co	358	5.1
ac	276	3.9
org	266	3.8
go	194	2.8
or	87	1.2
lg(ip)	26	0.4
net	10	0.1
ne(jp)	9	0.1
nhs(uk)	9	0.1
gob(es)	9	0.1

total domains
7.003

7003개의 도메인은 총 9149의 URL을 가진다. URL은 온라인 주소인 도메인에 실제 정보의 위치 등의 경로가 추가로 담긴 실제 네트워크 주소이다. Table 3는 URL 값의 순서대로 도메인을 정리한 것이다. Top 3를 차지한 nsw.gov.au, moe.edu.sg, k12.ca.us는 모두 대학 이전의 공립 초등, 중등 교육과 관련된 도메인이다. 호주의 뉴사우스웨일스 지역과 관련된 nsw.gov.au는 573

개의 URL로 가장 높은 순위를 차지했다. 호주는 코로나 바이러스 발병 초기 시점에 확진자가 많은 국가 중 하나이다. 더불어 중국 대륙에 이어 최초로 살아있는 코로나 바이러스를 배양했는데 호주의 시드니대학교와 뉴사우스 웨일스대학교의 연구자들이 참여한 것이 영향을 미친 것으로 판단된다. 특히 뉴사우스웨일스 지역의 환자에게서 채취한 코로나바이러스로 진행된 실험이 성공을 이뤘기 때문에 코로나바이러스에 대한 지역의 관심이 더욱 높았던 것으로 간주된다.

k12.ca.us가 상위권에 오른 배경은 미국의 bing.com 점유율이 가장 높을 뿐만 아니라 당시 캘리포니아에서 6명의 확진자가 발생하며 미국 내 확진자 비율이 가장 높은 지역이었기 때문이다.

Table 3. Domains rankings related to Coronavirus

Domain	URLs	%	Domain	URLs	%
nsw.gov.au	573	6.3	fanpage.it	13	0.1
moe.edu.sg	130	1.4	mb.com.ph	13	0.1
k12.ca.us	125	1.4	china.org.cn	12	0.1
emb-japan.go.jp	74	0.8	knack.be	12	0.1
hku.hk	44	0.5	snu.ac.kr	11	0.1
ctvnews.ca	35	0.4	nhs.uk	11	0.1
moph.go.th	32	0.3	corriere.it	11	0.1
jal.co.jp	30	0.3	upm.edu.my	10	0.1
smu.edu.sg	29	0.3	k12.ga.us	10	0.1
chinadaily.com.cn	28	0.3	wrdsb.ca	10	0.1
k12.nj.us	26	0.3	lesoir.be	10	0.1
nus.edu.sg	23	0.3	belgium.be	10	0.1
repubblica.it	22	0.2	cmu.ac.th	9	0.1
dfa.gov.ph	20	0.2	kbs.co.kr	9	0.1
cuhk.edu.hk	19	0.2	yna.co.kr	9	0.1
k12.ny.us	19	0.2	visitkorea.or.kr	9	0.1
cri.cn	17	0.2	rti.org.tw	9	0.1
moh.gov.my	16	0.2	k12.ma.us	9	0.1
doh.gov.ph	16	0.2	lefigaro.fr	9	0.1
chula.ac.th	15	0.2	nic.in	9	0.1
people.com.cn	13	0.1	yandex.ru	9	0.1
hkbu.edu.hk	13	0.1	yle.fi	9	0.1
mahidol.ac.th	13	0.1	idg.se	9	0.1
k12.mi.us	13	0.1			
esteri.it	13	0.1			
total urls 9,149					

또한, emb-japan.go.jp와 jal.co.jp는 각각 일본의 대사관과 항공사 도메인이다. 일본의 크루즈에서 174명의 확진자가 발생했으며 크루즈 특성상 다양한 국가가 연관되어있는 상황이 반영되면서 상위권에 위치한 것으로 보인다. hku.hk는 홍콩대학교의 도메인이다. 홍콩 역시 데이터 수집 당시 50명의 확진자가 발생해 확진자의 수가 높은 국가 중 하나였다. 또한, South China Morning Post의 기사에 따르면 홍콩대학교의 연구팀은 당시 최초로 코로나 백신 개발에 성공한 상태였다. 홍콩대학교는 수년간 SARS 연구 등 바이러스 연구에 많은 관심을 쏟아왔으며 이번 코로나바이러스와 관련해서도 다양한 연구와 관심을 보이며 상위권에 오른 것으로 보인다.

ctvnews.ca는 캐나다의 언론사 도메인으로 캐나다뿐만 아니라 코로나바이러스와 관련된 다양한 국제 뉴스를 다루며 상위권에 위치했다.

Table 4는 코로나바이러스와 관련된 정보를 얼마나 다양한 주체가 제공하는지 분석하기 위해 코로나바이러스 확진 국가와 6개의 그룹으로 분류한 도메인 간의 연결을 근거로 2-mode matrix를 구성한 것이다. 2월 12일 기준 코로나바이러스 확진자 순으로 나라를 배열했으며 3개 이상의 URL을 가진 도메인을 대상으로 6개 그룹으로 분류하였다. 도메인은 government(정부), academy(교육), business(기업), media(언론), health care(보건기관), Other(기타)로 분류했다.

중국의 경우 6개의 도메인그룹 중 언론그룹이 107개로 가장 많은 URL을 가졌고 그다음은 교육그룹(15), 정부그룹(12) 순이었다. 일본은 정부그룹이 URL 81개로 가장 컸으며 기업그룹(43), 교육그룹(16) 그리고 언론그룹(15)이 뒤따라 나타났다. 홍콩은 교육그룹이 URL 87개를 가지며 가장 컸고 정부그룹(22), 언론그룹(11), 기업그룹(8)/보건그룹(8) 순으로 나타났다. 싱가포르의 교육그룹이 182개의 URL을 가지며 압도적으로 큰 비율을 보였다. 태국 역시 교육그룹이 90개의 URL을 가지며 가장 컸다.

우리나라도 교육그룹이 51개의 URL을 가지며 가장 컸고 언론그룹(29), 정부그룹(20) 순으로 나타났다. 말레이시아는 교육그룹과 보건그룹의 URL이 많았으며, 대만은 교육그룹과 언론그룹의 도메인이 많았다. 호주는 정부그룹이 591개의 URL을 가지며 전체 국가 중에서 가장 많았다. 캐나다와 이탈리아, 인도, 러시아는 비교적 언론그룹의 도메인이 교육그룹과 정부그룹보다 많았다.

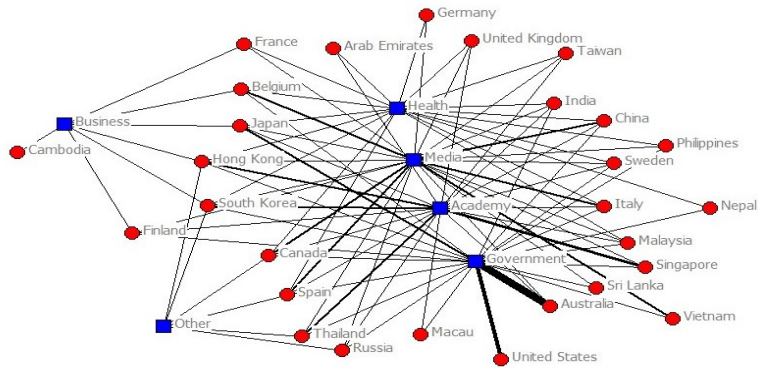


Fig. 2. Network of institutional actors and country

가장 많은 URL을 가진 그룹은 정부(총 1143 도메인)였으며, 언론(677), 교육그룹(618)의 비중이 다음으로 높았고 보건그룹(117)이 뒤를 이었다. 보건그룹의 경우는 보건복지부와 같은 공적인 기관이 대부분을 차지했다. 예외적으로 독일에서는 대학병원, 캐나다에서는 사설병원 그리고 스페인에서 간호조합과 같은 민간협회도 나타났다. 또한 비교적 확진자가 많은 국가 및 지역(홍콩, 싱가포르, 태국, 대한민국 등)에서는 교육그룹이 큰 비율을 차지하는 반면, 확진자가 적은 나라(캐나다, 이탈리아, 스페인, 벨기에 등)일수록 언론그룹의 비율이 크게 나타났다.

Fig. 2는 UCInet을 이용하여 2-mode matrix를 시각화한 것이다. UCInet의 Netdraw에서 레이아웃 옵션으로 Graph theoretic layout-spring embedding을 이용하였다. 이 레이아웃은 중심이 되는 노드가 중앙에 위치하도록 시각화되는 특징을 가진다. 따라서 URL이 많은 그룹인 정부와 언론, 교육그룹 그리고 보건 그룹이 정중앙에 위치한 것을 볼 수 있다. 국가와 그룹 간에는 총 95개의 관계선이 존재하며, 선의 굵기는 URL의 수가 클수록 두껍게 표시되었다. 이를 통해서 코로나바이러스 확진자가 발생한 국가 내에서 어떤 정보원이 관련된 정보를 생산하는지 그리고 얼마나 많은 정보를 생산하는지 파악할 수 있다. 코로나바이러스 확진자가 발생한 대부분의 국가들이 정부와 언론, 교육그룹과 관계가 있었으며 그 중에서도 언론그룹이 가장 많은 국가와 관계를 가졌다. 다시 말해, 코로나바이러스 확진자가 발생한 대부분의 국가들에서 정부, 언론, 교육기관이 코로나바이러스와 관련된 정보를 생산하는 정보원으로서의 역할을 하고 있다는 것이다. 일본, 태국, 중국, 홍콩, 싱가포르의 순으로 betweenness 값(매개중심성)이 가장 높게 나타났으며 후주의 degree 값(연결중심성)이 가장 높다. 그룹 측면

에서는 health care(보건 기관) 그룹의 betweenness 값이 가장 높으며 government(정부) 그룹의 degree 값이 제일 높은 것으로 나타났다.

좀 더 자세히 살펴보면 중국은 공산국가임에도 불구하고 언론으로 분류되는 도메인이 많았는데 chinadaily.com.cn가 28개의 URLs를 가지며 가장 많았다. 그 뒤로는 cri.cn와 people.com.cn, china.org.cn이 나타났는데 chinadaily.com.cn는 영어로 발간되는 중국공산당 홍보부 산하의 뉴스매체였고 cri.cn는 국영 국제라디오방송이었다. people.com.cn 역시 중국공산당 중앙위원회의 공식신문이었으며 china.org.cn은 국영 웹 포털로 상위에 위치한 4개의 도메인 모두 중국 정부와 관계된 언론기관이었다. 따라서 중국의 경우 언론기관임에도 불구하고 정부의 성격을 가지는 언론매체가 주로 정보를 생산했으며, 해외와 관련된 정보생산에 더 민감한 것으로 보인다. 포털의 경우 텔레비전이나 신문과 같은 전통적인 언론매체와 성격이 상이할 수 있으나 코로나와 같은 재난 상황에서 재난의 추이를 확인할 수 있으며 공신력 있는 정보를 제공한다는 점에서 언론과 같은 역할을 한다고 판단하여 언론으로 분류하였다.

일본의 경우 정부그룹 도메인 중에서 대사관 도메인인 emb-japan.go.jp가 74개의 URL을 차지했다. 이는 데이터 수집 당시 일본의 크루즈에서 확진자가 대거 발생했으며 크루즈 특성상 다양한 국가와의 소통이 필요한 상황이 영향을 끼친 것으로 보인다. 그다음 비중을 차지한 기업그룹에서도 역시 항공회사인 jal.co.jp가 그 중 30개의 URL을 가지며 오프라인에서의 이슈가 온라인의 정보생산에 큰 영향을 미친다는 사실을 보여주고 있다. 그 뒤로는 언론그룹에 속하는 일본 포털사이트인 구의 도메인goo.ne.jp와 야후 도메인 yahoo.co.jp가 각각 6

개, 5개의 URL을 가졌다.

홍콩은 교육그룹이 가장 큰 비율을 차지했는데 그중에서도 홍콩대학교의 도메인인 hku.hk가 44개의 URL을 가지며 가장 많았다. 그다음으로는 중문대학교cuhk.edu.hk와 침례대학hkbu.edu.hk의 도메인들이 많은 URL을 가졌다. 홍콩대학교와 중문대학교는 모두 홍콩에서 상위권에 속하는 대학이면서 홍콩대학교는 바이러스 연구에서, 중문대학교에는 호흡기 의학에서 전문성을 보이는 곳이었다. 전염병과 같은 국가적 재난 상황에서는 정부 등 관련 부처의 대처뿐만 아니라 질병에 관한 연구도 중요하게 여겨지는 것으로 보인다. 특히 이번 코로나바이러스는 폐렴과 연관성이 깊기 때문에 해당 분야에 전문성을 지닌 대학들이 큰 관심을 보이는 것으로 추정된다. 그 다음으로는 보건그룹에 속하는 공립병원관리기관 도메인 ha.org.hk가 8개의 URL을 가졌다. 더불어 홍콩에서는 기타그룹에서 천주교 도메인인 catholic.org.hk가 4개의 URL을 보였다. 이는 홍콩의 가톨릭 비율이 28.6%(출처: KOTRA)로 가장 큰 비중을 차지하는 종교이기 때문인 것으로 추정된다.

싱가포르는 다른 국가들보다 특히 더 교육그룹이 컸는데 초등, 중등의 교육을 담당하는 도메인 moe.edu.sg가 130개의 URL로 많이 나타났다. 더불어 싱가포르 경영대학교의 도메인인 smu.edu.sg가 29개의 URL을 가졌고 싱가포르 국립대학 nus.edu.sg이 23개의 URL을 가진 것으로 나타났다. 이는 코로나바이러스 확산을 막기 위해 휴교를 고려하거나 실시한 대부분의 국가들과 달리 싱가포르에서는 휴교 없이 정상적으로 학교를 운영할 것이라는 교육부의 지침이 있었기 때문에 교내 안전을 위해 더 많은 정보를 생산했을 것으로 추정된다.

태국 역시 chula.ac.th, mahidol.ac.th, cmu.ac.th 등 대학교 도메인들이 많이 나타나며 교육그룹이 큰 비중을 차지했다. 그러나 정부그룹에 속한 식품의약부의 도메인 moph.go.th이 가장 많은 URL을 가지는 것으로 나타났다. 우리나라에서는 국내에서 큰 대학병원을 운영하고 있는 대학교인 서울대학교 snu.ac.kr, 연세대학교 yonsei.ac.kr, 아주대학교ajou.ac.kr가 많은 URL을 가지며 교육그룹의 비중을 높였다. 다음으로 큰 언론그룹에서는 KBS kbs.co.kr, 연합뉴스 yna.co.kr가 많은 URL을 가졌고, 정부그룹에서는 한국관광공사 visitkorea.or.kr와 서울시 seoul.go.kr가 많이 나타났다.

Table 4. Distribution of URLs per institutional actors

	Government	Academy	Business	Media	Health	Other
China	12	15	0	107	8	0
Japan	81	16	43	15	4	0
Hong Kong	22	87	8	11	8	4
Singapore	3	182	0	4	0	0
Thailand	37	90	0	5	0	3
South Korea	20	51	4	29	6	3
Malaysia	5	18	0	8	16	0
Taiwan	0	35	0	17	4	0
Germany	0	0	0	12	3	0
Australia	591	15	0	4	0	0
Vietnam	5	0	0	48	0	0
United States	234	0	0	0	0	0
France	0	0	7	46	4	0
Macau	13	14	0	0	0	0
Arab Emirates	0	3	0	0	4	0
United Kingdom	0	6	0	4	11	0
Canada	8	39	0	55	7	3
Italy	13	3	0	91	6	0
Philippines	30	0	0	24	16	0
India	9	3	0	24	3	0
Russia	7	4	0	21	0	6
Spain	6	4	0	52	7	3
Nepal	3	0	0	0	4	0
Cambodia	0	0	5	0	0	0
Belgium	10	0	3	60	3	0
Finland	6	11	3	9	0	0
Sweden	25	22	0	26	3	0
Sri Lanka	3	0	0	5	0	0

국가 및 그룹의 2-mode matrix를 분석하는 과정에서 UCInet를 사용하여 중심성 분석을 실시했다. 그 결과는 Table 5와 Table 6이다. Degree 값은 노드가 이웃 노드들과 직접 연결된 정도를 파악할 수 있고 2-Local 값은 중심성을 측정할 수 있다. Degree와 2-Local의 값이 가장 높은 3개의 나라는 Australia(호주), United States(미국), Singapore(싱가포르)이고, 가장 높은 3개의 그룹은 government(정부), media(언론)과 academy(교육)이다. 그러므로 국가 측면에서는 호주, 미국, 싱가포르 등이 인터넷에서 코로나 정보의 양이 크고, 그룹 측면에서는 정부, 언론과 교육에 관련한 기관의 코로나 정보가 더 많다.

Table 5. 2-mode matrix centrality for country (among institutional actors)

Degree			2-Local		
1	Australia	101.667	1	Australia	4092.209
2	United States	39	2	United States	1592.036
3	Singapore	31.5	3	Singapore	706.03
4	Japan	26.5	4	Japan	691.863
5	China	23.667	5	Thailand	603.345
6	Hong Kong	23.333	6	China	573.577
7	Thailand	22.5	7	Hong Kong	523.613
8	South Korea	18.833	8	Italy	470.369
9	Italy	18.833	9	South Korea	446.851
10	Canada	18.667	10	Canada	424.798
11	Belgium	12.667	11	Sweden	357.881
12	Sweden	12.667	12	Belgium	313.214
13	Spain	12	13	Philippines	311.964
14	Philippines	11.667	14	Spain	270.351
15	France	9.5	15	Vietnam	227.446
16	Taiwan	9.333	16	Taiwan	200.042
17	Vietnam	8.833	17	France	191.196
18	Malaysia	7.833	18	India	171.071
19	India	6.5	19	Russia	147.75
20	Russia	6.333	20	Malaysia	143.613
21	Finland	4.833	21	Macau	139.946
22	Macau	4.5	22	Finland	118.857
23	United Kingdom	3.5	23	Germany	50.446
24	Germany	2.5	24	United Kingdom	45.851
25	Sri Lanka	1.333	25	Sri Lanka	40.56
26	Arab Emirates	1.167	26	Nepal	23.196
27	Nepal	1.167	27	Arab Emirates	13.821
28	Cambodia	0.833	28	Cambodia	2.173

Table 6. 2-mode matrix centrality for actors

Degree			2-Local		
1	Government	40.821	1	Government	2678.542
2	Media	24.179	2	Academy	529.173
3	Academy	22.071	3	Media	377.756
4	Health	4.179	4	Business	54.452
5	Business	2.607	5	Health	52.173
6	Other	0.786	6	Other	12.405

국가별로 코로나바이러스와 관련된 어떤 정보가 주로 전달되고 있는지 파악하기 위해 웹메트릭스에서 수집한 웹 내용 데이터를 바탕으로 FullText 툴을 사용해 단어빈도분석을 실시했다[15]. Fig. 3은 국가별로 100회 이상 출현하여 빈도율이 1%가 넘는 단어를 제시하였다. “coronavirus”가 검색어이기 때문에 해당 단어는 그림에서 제외하였다. 그림에서 국가와 지역은 사각형 노드, 단어는 원형노드로 나타났다. 호주의 degree, betweenness, eigenvector(위세중심성) 값이 가장 높게 나타나며 미국의 degree, eigenvector 값이 2위이고 betweenness 값이 3위를 차지했다. 캐나다의 degree, eigenvector 값이 3위이고 betweenness 값이 2위로 나타났다. 단어 측면에는 degree와 eigenvector 값이 높은 것은 novel, virus, health이고 betweenness 값이 높은 것은 novel, China, virus이다. 전체적으로 novel(신종), China(중국 대륙), health(건강), virus(바이러스), public(공공), school(학교) 등의 단어 빈도

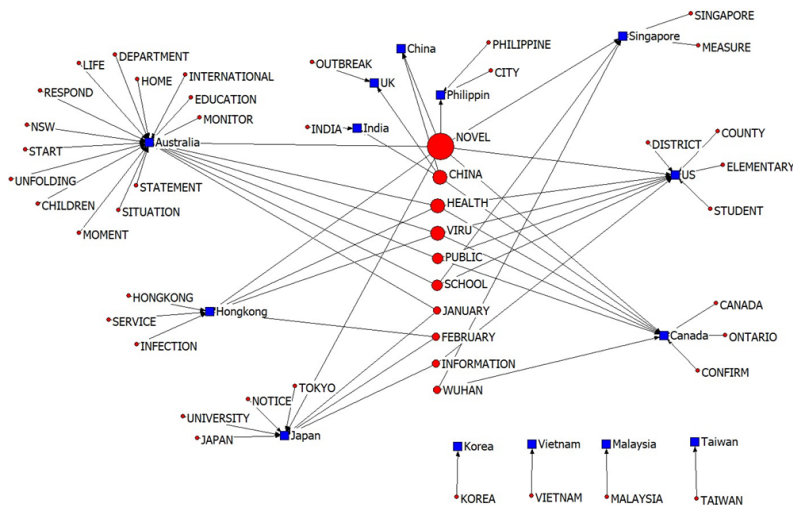


Fig. 3. A semantic network among high frequency words of web contents (n>100)

가 높게 나타났다. 이 밖에도 호주는 Home(가정), Education(교육), Life(생활), Monitor(모니터) 등이, 홍콩은 Service(서비스)와 Infection(감염), 일본은 University(대학교)와 Tokyo(도쿄), 캐나다는 Ontario와 Confirm(확진자), 미국은 District(지방)와 Elementary student(초등학생), 싱가포르의 Measure(방안), 필리핀은 City(도시) 등의 단어의 빈도가 높아 해당 분야에 관심을 쏟고 있는 것을 알 수 있다. 한국, 인도, 베트남, 말레이시아 등의 국가에서는 바이러스를 제외하고 가장 높은 빈도를 가진 단어가 자신의 나라와 지역이었다. 자국의 상황에 보다 많은 관심을 보이는 것으로 보인다.

국가 및 웹내용의 2-mode matrix를 분석한 결과는 Table 7과 Table 8에서 제시하였다. Degree와 2-Local의 값이 가장 높은 나라는 호주, 미국, 캐나다이다. 고 빈도 단어의 Degree 값이 가장 큰 것이 novel이고 virus, health, China, school, public 순서로 나타났다. 2-Local 값은 novel이 가장 크고 virus, health, China, school, public 순서로 나타났다. 이를 통해 건강, 학교, 공공에 관련한 주제가 더 주목 받는 이슈라는 것을 알 수 있다.

Table 7. 2-mode matrix centrality for country (among high frequency words of web contents)

Degree			2-Local		
1	Australia	0.417	1	Australia	0.174
2	US	0.208	2	US	0.043
3	Canada	0.188	3	Canada	0.035
4	Japan	0.167	4	Japan	0.028
5	Hongkong	0.146	5	Hongkong	0.021
6	Singapore	0.104	6	Singapore	0.011
7	Philippin	0.063	7	Philippin	0.004
8	India	0.042	8	India	0.002
9	UK	0.042	9	UK	0.002
10	China	0.042	10	China	0.002
11	Korea	0.021	11	Korea	0.000
12	Malaysia	0.021	12	Malaysia	0.000
13	Taiwan	0.021	13	Taiwan	0.000
14	Vietnam	0.021	14	Vietnam	0.000

Table 8. 2-mode matrix centrality for words (top 10)

Degree			2-Local		
1	Novel	0.571	1	Novel	0.095
2	Virus	0.286	2	Virus	0.068
3	Health	0.286	3	Health	0.068
4	China	0.286	4	Public	0.058
5	School	0.214	5	School	0.052
6	Public	0.214	6	January	0.042
7	Wuhan	0.143	7	Unfolding	0.030
8	January	0.143	8	Statement	0.030
9	Information	0.143	9	Start	0.030
10	February	0.143	10	Situation	0.030

4. 결론

2020년 코로나가 전 세계적으로 대유행하는 배경에서, 본 연구는 웹보메트릭 방법을 이용해 위기 기간 동안 코로나바이러스를 키워드로 웹 정보를 수집했다. 수집된 도메인을 분류하고 분석한 결과, 코로나바이러스를 언급한 도메인 수와 히트수가 많을수록 코로나바이러스에 대한 해당 국가의 관심도가 높은 것을 알 수 있었다.

데이터를 전체적으로 보면, 정부 도메인 총 1266개, 18.1% 비율로 가장 많았다. 그 다음에 기업 도메인은 811개, 11.5%의 비율을 차지했다. 전염병과 같은 국가적인 재난 상태에서는 정부가 정보생산자로서 중요한 역할을 한다는 것을 알 수 있다. 전염병과 같은 국가적 재난이 발생했을 때 정부의 집행력과 반응 정도는 재난의 통제에 매우 중요하다[16,17]. 또한 기업과 교육 분야도 정부 못지않은 영향력과 역할을 가지고 있다. 전염병이 발생했을 때, 통상적으로 기업과 교육은 가장 큰 영향을 받는 업종에 속하는데 전염병으로 인해 기업의 업무 중단을 초래하고 학교는 휴교를 하게 되기 때문이다. 도메인에 실제 정보의 위치 등의 경로가 추가로 담긴 실제 네트워크 주소 URL를 분석한 결과, Top 3를 차지한 도메인은 모두 교육과 관련된 도메인이었다. 일본 대사관과 항공사, 홍콩대학교 및 뉴스 도메인의 URL도 당시에 높게 나타났다. 이유를 살펴보면 코로나바이러스 확진자가 대거 발생한 사건과 관련된 것이거나 코로나바이러스 연구와 관련된 핫이슈에 관한 것이 대부분이다.

코로나바이러스 확진 국가와 6개의 그룹으로 분류한 도메인 간의 연결을 근거로 2-mode matrix를 통해 코로나바이러스와 관련된 정보의 다양한 주체를 살펴봤다.

그 결과 코로나바이러스 확진자가 발생한 대부분의 국가들에서 정부, 언론, 교육기관이 코로나바이러스와 관련된 정보를 생산하는 정보원으로서의 역할을 하고 있다는 것이다. 더 자세히 살펴보면 중국의 경우 언론기관임에도 불구하고 정부의 성격을 가지는 언론매체가 주로 정보를 생산했으며, 해외와 관련된 정보생산에 더 민감한 것으로 보인다. 일본의 경우 정부그룹 도메인 중에서 대사관 도메인이 가장 높다. 이는 데이터 수집 당시 일본의 크루즈에서 확진자가 대거 발생했으며 크루즈 특성상 다양한 국가와의 소통이 필요한 상황이 영향을 키친 것으로 보인다. 홍콩은 교육그룹이 가장 큰 비율을 차지했다. 교육그룹을 살펴보니 바이러스 연구와 호흡기 의학에서 전문성을 보이는 곳들이었다. 이를 바탕으로 해당 분야에 전문성을 지닌 대학들이 큰 관심을 보이는 것으로 추정된다. 싱가포르는 다른 국가들보다 특히 더 교육그룹이 컸는데 코로나바이러스 확산을 막기 위해 휴교를 고려하거나 실시한 대부분의 국가들과 달리 싱가포르에서는 휴교 없이 정상적으로 학교를 운영할 것이라는 교육부의 지침이 있었기 때문에 교내 안전을 위해 더 많은 정보를 생산했을 것으로 추정된다.

마지막으로 우리가 수집한 웹페이지의 내용을 의미망 분석을 했는데, 국가별 웹사이트 내용의 단어 빈도를 보면 건강, 학교, 가족, 공공, 방한 등이 가장 주요한 몇 가지 주제가 되었고 대부분의 나라들이 자신의 지역과 관련된 주제를 더 많이 제공했다는 것을 알 수 있다.

본 연구의 한계점은 내용분석을 실시하면서 영어로 된 내용만 고려해서 연구했다는 점이다. 또한 2월의 데이터를 분석했기 때문에 현재 더 많은 국가에 확진자가 발생했으므로 향후 연구에서 이 방법을 바탕으로 더 많은 데이터를 분석해서 더 큰 가치를 얻기를 기대한다.

현재 닥친 코로나바이러스 위기에서 벗어나기 위해서는 데이터의 다양성이 더 많이 확보되고 더 복잡한 모델이 필요할 수 있다. 예컨대 소셜 미디어, 뉴스 미디어, 건강 학술 논문, 그리고 다른 이질적인 데이터들의 집계는 이 논문이 극복해야 할 벽찬 과제다. 그럼에도 불구하고 이 논문은 보건 캠페인의 전략적 메시지 작성과 채널 선택 시스템 구성을 위한 기초적 자료가 될 수 있다. 웹메트릭 분석이 질병 예측과 확산을 위한 만능 도구는 아니지만 대중들의 이야기와 공유된 뉴스로부터 얻게 되는 함의와 그 잠재력을 과소평가해서는 안 된다.

또한 최근 의료 분야에서 증강 지능 시스템이 질병의 검출과 임상진단을 돕는 중요한 접근방식으로 널리 인용되고 있으나[18], 이 논문의 목적은 이러한 AI 시스템에

데이터를 제공하여 전염병 발생을 예측하는 데 있지 않다. 그보다는 웹메트릭 빅데이터 분석을 통해 전략적 커뮤니케이션의 필요성과 기회를 강조하고 있다. 이러한 분석이 공중보건 공무원의 고용을 대체할 수 없지만, 빠르게 전파하는 정보를 수집하여 그 구조를 검토하면, 복잡한 의사결정에 있어 공중보건 전문가들을 도울 수 있다.

References

- [1] H. W. Park, L. Leydesdorff, H. D. Hong, S. J. Hong (2004). "Triple-Helix Indicators for the knowledge-based innovation system: A comparison between South Korea and The Netherlands", *Journal of the Korean Data Analysis Society*, Vol.6, No.5, pp.1389-1402, Oct. 2004.
- [2] G. F. Khan, H. W. Park, "Editorial: Triple-Helix and innovation in Asia using scientometrics, webometrics, informetrics". *Scientometrics*, Vol.90, No.1, pp.1-7, Jan. 2012.
DOI: <https://doi.org/10.1007/s11192-011-0506-7>
- [3] C. J. Chung, H. W. Park, "Mapping triple helix innovation in developing and transitional economies: Webometrics, scientometrics, and informetrics", *Scientometrics*, Vol.99, pp.1-4. Apr. 2014.
DOI: <https://doi.org/10.1007/s11192-013-1105-6>
- [4] H. C. Kwon, "A study of semantic network analysis of newspaper articles on MERS situation: Comparing conservative and progressive news media", *Health Communication, the Official Journal of Korean Academy on Communication in Healthcare*, Vol.11, No.1, pp.63-80. 2016.
DOI: <http://doi.org/10.15715/kjhcom.2016.11.1.63>
- [5] M. Lee, J. H. Hong, "Semantic network analysis of government's crisis communication messages during the MERS outbreak", *Journal of The Korea Contents Association*, Vol.16, No.5, pp.124-136. May 2016.
DOI: <http://doi.org/10.5392/JKCA.2016.16.05.124>
- [6] N. Y. Jin, C. J. Chung, "Semantic network analysis of domestic and overseas media coverage regarding Korea MERS". *Journal of Communication Science*, Vol.18, No.2, pp.222-262, Jun. 2018.
DOI: <http://doi.org/10.14696/jcs.2018.06.18.2.222>
- [7] J. Y. Song, T. M. Song, D. C. Seo, D. L. Jin, J. S. Kim, "Social big data analysis of information spread and perceived infection risk during the 2015 Middle East Respiratory Syndrome outbreak in South Korea", *Cyberpsychology, Behavior, and Social Networking*, Vol.20, No.1, pp.22-29, Jan. 2017.
DOI: <https://doi.org/10.1089/cyber.2016.0126>
- [8] S. Acharya, H. W. Park, "Open data in Nepal: a webometric network analysis", *Quality & Quantity*, Vol.51, pp.1027-1043, May 2017.

DOI: <https://doi.org/10.1007/s11135-016-0379-1>

- [9] T. Ying, W. C. Norman, Y. Zhou, "Online Networking in the Tourism Industry: A Webometrics and Hyperlink Network Analysis", *Journal of Travel Research*, Vol.55, No.1, pp.16-33, May. 2014.
DOI: <https://doi.org/10.1177/0047287514532371>
- [10] S. Park, H. W. Park, "Diffusion of cryptocurrencies: web traffic and social network attributes as indicators of cryptocurrency performance", *Quality & Quantity*, Vol.54, pp.297-314, Feb. 2020.
DOI: <https://doi.org/10.1007/s11135-019-00840-6>
- [11] H. W. Park, "Mapping the e-science landscape in South Korea using the webometrics method", *Journal of Computer-Mediated Communication*, Vol.15, No.2, pp.211-229, Jan. 2010.
DOI: <https://doi.org/10.1111/j.1083-6101.2010.01517.x>
- [12] H. J. Park, "A Study on Legal Nature of Domain Name", *Law Review*, Vol.26, pp.237-262, May 2007.
- [13] G. Y. Kim, "Analysis on destination network characteristics of using 2-mode network analysis : Case of Japanese tourists visiting Korea", *Journal of Tourism Studies*, Vol.31, No.4, pp.23-46, Nov. 2019.
DOI: <https://doi.org/10.21581/jts.2019.11.31.4.23>
- [14] Microsoft Advertising, Did we mention Bing is bigger than you think?, Twitter, c2017 [cited 2017 August 8], Available From: <https://twitter.com/MSFTAdvertising/status/898208047578849280> accessed May 12, 2020
- [15] L. Leydesdorff, I. Hellsten, "Metaphors and diaphors in science communication: Mapping the case of stem cell research", *Science Communication*, Vol.27, No.1, pp.64-99, Sep. 2005.
DOI: <https://doi.org/10.1177/1075547005278346>
- [16] H. W. Park, S. Park, M. Chong, "Conversations and medical news frames on Twitter: Infodemiological study on COVID-19 in South Korea", *Journal of medical Internet research*, Vol.22, No.5, e18897, May 2020.
DOI: <https://doi.org/10.2196/18897>
- [17] H. C. Park, H. W. Park, "Disaster Communication Networks on Twitter : Gyeongju Earthquake", *Journal of the Korean Data Analysis Society*, Vol.19, No.1, pp.291-302, Feb. 2017.
DOI: <http://doi.org/10.37727/ikdas.2017.19.1.291>
- [18] J. B. Long, J. M. Ehrenfeld, "The role of Augmented Intelligence (AI) in detecting and preventing the Spread of Novel Coronavirus", *Journal of Medical Systems*, Vol.44, No.3, e59, Feb. 2020.
DOI: <https://doi.org/10.1007/s10916-020-1536-6>

박 한 우(Han Woo Park)

[종신회원]



- 1995년 2월 : 한국외국어대 미디어&커뮤니케이션학 (정치학사)
- 1997년 2월 : 서울대 언론정보학 (문학석사)
- 2002년 6월 : 뉴욕주립대학교 (버팔로) 정보학/커뮤니케이션학(Ph.D)
- 2003년 3월 ~ 현재 : 영남대학교 교수

<관심분야>

빅데이터, 웹보메트릭스, 블록체인, 네트워크분석, 과학기술정보학, 소셜미디어, 트리플헬릭스

김 지 은(Ji-Eun Kim)

[준회원]



- 2017년 2월 : 영남대학교 언론정보학과, 경영학과(인문학, 경영학 학사)
- 2019년 12월 ~ 현재 : 영남대학교 사이버감성연구소 연구원
- 2020년 3월 ~ 현재 : 영남대학교 디지털융합비즈니스학과 석·박사 통합과정

<관심분야>

빅데이터, 네트워크분석, 관광학, 웹보메트릭스

주 우 봉(Yu-Peng Zhu)

[정회원]



- 2014년 8월 : 계명대학교 기계자동차공학(기계자동차공학학사)
- 2016년 8월 : 계명대학교 기계공학(공학석사)
- 2020년 8월 : 영남대학교 디지털융합비즈니스학(Ph.D)

• 2018년 9월 ~ 현재 : 영남대학교 사이버감성연구소 블록체인정책연구센터 연구원

<관심분야>

빅데이터, 블록체인, 네트워크분석, 과학기술정보학, 웹보메트릭스