

# 조현병 관련 주요 일간지 기사에 대한 텍스트 마이닝 분석

서울의료원 정신건강의학과,<sup>1</sup> 전남대학교 의과대학 정신건강의학교실<sup>2</sup>  
남희정<sup>1</sup> · 류승형<sup>2</sup>

## Text-Mining Analyses of News Articles on Schizophrenia

Hee Jung Nam, MD, PhD<sup>1</sup> and Seunghyong Ryu, MD<sup>2</sup>

<sup>1</sup>Department of Psychiatry, Seoul Medical Center, Seoul,

<sup>2</sup>Department of Psychiatry, Chonnam National University Medical School, Gwangju, Korea

**Objectives:** In this study, we conducted an exploratory analysis of the current media trends on schizophrenia using text-mining methods.

**Methods:** First, web-crawling techniques extracted text data from 575 news articles in 10 major newspapers between 2018 and 2019, which were selected by searching “schizophrenia” in the Naver News. We had developed document-term matrix (DTM) and/or term-document matrix (TDM) through pre-processing techniques. Through the use of DTM and TDM, frequency analysis, co-occurrence network analysis, and topic model analysis were conducted.

**Results:** Frequency analysis showed that keywords such as “police,” “mental illness,” “admission,” “patient,” “crime,” “apartment,” “lethal weapon,” “treatment,” “Jinju,” and “residents” were frequently mentioned in news articles on schizophrenia. Within the article text, many of these keywords were highly correlated with the term “schizophrenia” and were also interconnected with each other in the co-occurrence network. The latent Dirichlet allocation model presented 10 topics comprising a combination of keywords: “police-Jinju,” “hospital-admission,” “research-finding,” “care-center,” “schizophrenia-symptom,” “society-issue,” “family-mind,” “woman-school,” and “disabled-facilities.”

**Conclusion:** The results of the present study highlight that in recent years, the media has been reporting violence in patients with schizophrenia, thereby raising an important issue of hospitalization and community management of patients with schizophrenia. (Korean J Schizophr Res 2020;23:58-64)

**Key Words:** Media · News · Schizophrenia · Text-mining · Violence.

## 서 론

여러 언론을 통해 보도된 일련의 조현병 환자에 의한 사건 사고 뉴스에 대해 일반 대중은 높은 관심을 신속하게 나타내었다. 대형 포털 사이트에서는 ‘조현병’에 대한 검색량이 크게 증가하였으며, 상당 시간 동안 검색어 순위에 오르내리기도 하였다. 또한, 해당 기사에 대한 일부 네티즌들의 익명 댓글

들을 통해 조현병에 대한 부정적 인식을 거침없이 드러내는 모습이였다. 사건에 대한 대책으로 정부는 환자 동의 없이 추적 관리를 허용하는 등 중증정신질환자 지역사회 치료 지원 강화 방안을 내놓았고,<sup>1)</sup> 대한조현병학회는 조현병에 대한 사회적 낙인이 확산되는 것에 대한 우려를 성명서를 통해 발표하기도 하였다.<sup>2)</sup> 그러나 이미 신문과 방송 등 기존 언론은 물론 소셜 미디어와 포털 사이트 등 인터넷 언론을 통해 빠르고 넓게 전파된 사건 사고 뉴스들은 조현병 환자에 의한 폭력 문제를 일반 대중에게 인지시키는 역할을 하였다.

정신질환에 대한 사회적 편견은 조현병 환자들의 재활과 사회 복귀를 제한하는 주요 요인이다.<sup>3,4)</sup> 편견이란 부정적이거나 잘못된 정보에 의해 형성된 특정인에 대한 태도로, 과거에 편견은 주로 본인이 직접 경험하거나, 주변 사람들을 통해 학습되거나, 또는 사회집단에서 오래 동안 유지되어 온 것들로

Received: April 29, 2020 / Revised: July 23, 2020

Accepted: August 7, 2020

Address for correspondence: Seunghyong Ryu, Department of Psychiatry, Chonnam National University Medical School, 42 Jebong-ro, Dong-gu, Gwangju 61469, Korea

Tel: 062-220-5839, Fax: 062-225-2351

E-mail: ryush@chonnam.ac.kr

본 연구는 2018년도 대한조현병학회 연구기금 연구비의 지원을 받아 수행된 연구임.

부터 형성되었다.<sup>5,6)</sup> 그러나 오늘날에는 매스 미디어에 의해 개인이 영위하는 집단과 경험의 범위가 빠르게 확대되고 있으며, 이를 통해 전달된 정보와 경험들은 편견을 형성하고 확산하는데 크게 기여할 수 있다.<sup>7)</sup> 앞서 언급한 바와 같이, 조현병 환자에 의한 사건 사고 소식들도 다양한 미디어를 통해 빠른 속도로 전파되고 있으며, 일반인들은 사건에 대한 간접 경험을 통해 조현병 환자에 대한 부정적 인식과 편견을 형성하게 된다. 기술의 혁명적 발전과 그에 따른 사회적 변화 가운데 사회 구성원으로서의 조현병 환자들의 입지는 앞으로 더욱 줄어들지도 모르는 현실 속에서, 관리 밖의 조현병 환자에 의한 사건 사고들은 조현병에 대한 부정적 인식과 사회적 편견을 빠르게 확산시키고, 결국 성실히 치료 받고 재활 중인 환자들의 사회 복귀를 더욱 어렵게 만들 수 있다.

요약하면, 조현병 환자에 의한 사건 사고는 언론의 주목을 끌 수 있는 소재로, 미디어를 통해 매우 빠른 속도로 전파되고 있으며, 이에 따른 조현병에 대한 부정적 인식과 편견의 확산이 우려되는 상황이다. 이에 본 연구에서는 조현병과 관련하여 어떠한 내용들이 언론 매체를 통해 전달되고 있는지를 확인해 보고자 하였다. 기존의 내용 분석(content analysis)에 기반한 언론 연구들은 많은 시간과 비용이 요구되고, 연구자의 주관이 반영될 수밖에 없다는 한계가 있었다.<sup>8)</sup> 대신, 본 연구에서는 텍스트 마이닝(text-mining)이라는 프로그래밍 기법을 활용하여, 보도 기사를 통해 수집된 비정형 텍스트 자료들로부터 특정 패턴과 의미 있는 정보를 추출하고자 하였다. 텍스트 마이닝은 인간의 언어로 이루어진 비정형 텍스트 데이터들을 자연어 처리(natural language processing) 방식을 이용하여 대규모 문서에서 정보를 추출하거나, 연계성을 파악하거나, 분류 혹은 군집화, 요약 등 데이터에 숨겨진 의미를 발견하는 기법을 말한다.<sup>9)</sup> 본 연구에서는 10개 주요 일간지에서 '조현병'을 검색어로 검색된 지난 2년간의 지면 기사를 수집하여, 빈도 분석(frequency analysis)과 동시 출현 단어 분석(co-occurrence network analysis), 그리고 토픽 모형 분석(topic model analysis)을 통해 조현병 관련 기사에서 등장하는 주요 키워드와 토픽을 파악하였다.

## 방 법

### 텍스트 데이터 추출

크롤링(crawling)은 웹 페이지를 그대로 가져와서 거기서 데이터를 추출해 내는 행위이다.<sup>10)</sup> 네이버 뉴스(news.naver.com)에서 제공하는 2018년 1월 1일부터 2019년 12월 31일까지의 10개 주요 일간지(경향, 국민, 동아, 문화, 서울, 세계, 조

선, 중앙, 한겨레, 한국) 지면 기사 중 '조현병'을 검색어로 검색된 579건 기사의 URL 주소 목록을 확보하고, R의 rvest 라이브러리(library)를 이용하여 해당 기사의 제목과 본문을 크롤링하였다.<sup>11)</sup>

### 텍스트 데이터 사전 처리

일차적으로 연구자 2인의 검토를 통해 579건의 기사 중 인사, 부고, 운세 등 관련성이 낮은 기사 4건을 제외한 총 575건의 기사를 최종 선택하였다. 그리고 기사 본문 텍스트에서 신문사나 기자 이름 등 분석 시에 불필요한 단어는 텍스트 자료에서 삭제하였다. 텍스트 마이닝에 의한 텍스트 데이터 사전 처리 과정은 다음과 같다. 우선, R의 tm 라이브러리를 이용하여 기사 텍스트 자료에 대한 말뭉치(corpus)를 구성하였다.<sup>12)</sup> 그리고 말뭉치 텍스트 데이터에서 숫자, 알파벳, 문장부호, 그리고 특수문자들을 삭제하였다. 이후, NIADic 한글형태소 사전을 기반한 KoNLP 라이브러리의 함수를 이용하여 말뭉치 텍스트 데이터에서 2~8음절의 명사만을 추출하였다.<sup>13)</sup> 추출된 명사들의 빈도표를 확인하여 발현 빈도가 높은 단어 가운데, 분석에 불필요할 것으로 판단되는 단어는 추가로 삭제하였으며(예: '올해', '지난해', '과거' 등), 중복된 의미를 갖는 단어들은 하나의 단어로 통일시켰다(예: 신경정신과, 정신과 → 정신건강의학과, 정신건강복지센터, 정신건강센터 → 정신건강증진센터).

상기의 사전 처리를 거친 말뭉치에 대하여 tm 라이브러리를 이용하여 문서×단어 행렬(document-term matrix, DTM)과 단어×문서 행렬(term-document matrix, TDM)을 구축하였다. DTM과 TDM 모두 특정 문서에 등장하는 특정 단어의 등장 빈도를 행렬로 나타내는 것으로, 본 연구에서는 행렬에 대한 가중치(weighting)로 단어 빈도(term frequency, TF), 바이너리(binary) 가중치, 그리고 단어 빈도-역문서 빈도(term frequency-inverse document frequency, TF-IDF)를 분석 방법에 따라 적용하였다. TF는 특정한 단어가 어떤 범위 내의 문서에서 얼마나 자주 등장하는지를 나타내는 값으로, 특정 단어의 빈도수가 높으면 문서 내에서 해당 단어가 중요하다 생각될 수 있다. 바이너리 가중치는 해당 단어가 텍스트에 출현하면 1, 아니면 0을 반환하는 방법이다. 문서 빈도(document frequency, DF)는 특정 단어가 일정 문서들에서 얼마나 자주 사용되는지를 나타내는 값으로, 단어가 나타나는 문서의 수가 많을 수록, 그 단어는 보편적인 단어임을 의미한다. 역문서 빈도(inverse document frequency)는 DF의 역수이다. TF-IDF는 TF와 IDF의 곱으로, 여러 문서로 이루어진 문서군이 있을 때 어떤 단어가 특정 문서 내에서



리' 등의 단어가 상호 간 밀접한 관계를 나타냈으며, '경찰,' '사건,' '범죄,' '흉기' 등의 단어들은 '진주,' '아파트,' '주민,' '범행,' '협의,' '살해' 등의 단어 등과 가까운 관계를 갖고 있었다.

표 3은 10개의 잠재 토픽으로 분류된 상위 20개의 단어들과 기사 전체에서의 토픽 등장 확률의 총합을 나타낸다. 토픽의 이름은 각각의 토픽으로 분류된 단어들을 바탕으로 잠재 토픽의 의미를 살릴 수 있도록 정해졌다. 기사에서 가장 자주 등장했던 토픽은 '경찰-진주'이었으며, 이외에 '병원-입원,' '연구-결과,' '센터-관리,' '범죄-법원,' '조현병-증상,' '사회-문제,' '가족-심리,' '여성-학교,' '장애인-시설'의 토픽이 비슷한 수준의 등장 확률로 분류되었다.

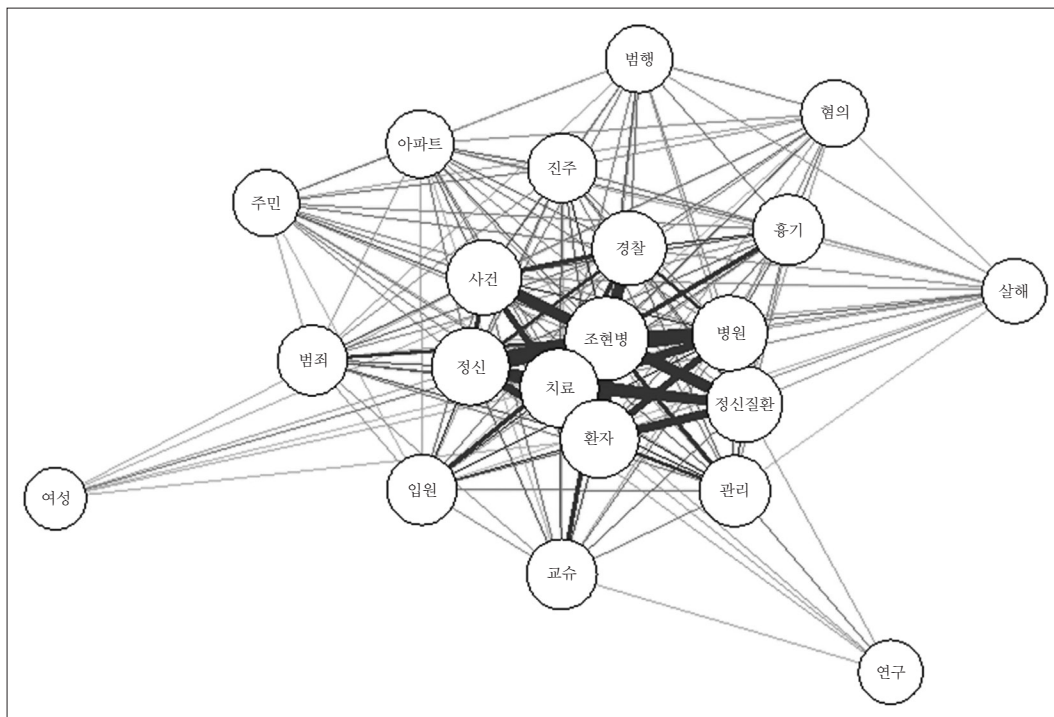
**Table 2.** Top 20 frequently occurring words based on term frequency-inverse document frequency (TF-IDF)

Rank	Word	TF-IDF*	Rank	Word	TF-IDF*
1	경찰	8.38	11	병원	3.62
2	정신질환	5.78	12	사건	3.54
3	입원	4.79	13	관리	3.48
4	환자	4.78	14	연구	3.47
5	범행	4.51	15	협의	3.46
6	아파트	4.45	16	정신	3.44
7	흉기	4.30	17	범죄	3.44
8	치료	4.27	18	살해	3.14
9	진주	3.86	19	교수	3.08
10	주민	3.86	20	여성	3.04

\*: Normalized value

## 고찰

본 연구에서 분석된 총 575건의 조현병 관련 기사 중 2019년에 보도된 기사 수는 427건으로 2018년 기사 수 148건 대비 약 3배에 달했다. 이는 2019년 한 해 동안 언론의 주목을 끈 조현병 관련 이슈들이 있었음을 시사한다. 주지하다시피 2019년 4월 경남 진주의 한 아파트에서 발생한 조현병 환자에 의한 방화 및 살인 사건은 큰 사회적 파장을 불러일으킨 바 있다. 본 연구의 빈도 분석에서 '경찰,' '아파트,' 그리고 '진주' 등의 단어가 높은 등장 빈도를 나타내고 있었으며, 동시 출현 네트워크 분석에서도 이들 단어가 '사건,' '범죄,' 그리고 '흉기' 등의 단어와 동반되고 있었다. 또한, 토픽 모형 분석도 '경찰-진주'가 기사 텍스트 중에서 가장 자주 등장한 토픽임을 확인하였다. 즉, 2019년 진주 사건이 최근 2년간 조현병 관련 언론 기사들 가운데 매우 큰 비중을 차지하는 이슈였음을 본 연구 결과는 보여주고 있다. 한편, 진주 사건 이외에도 최근 2년 동안 2018년 4월 방배 초등학교 인질 사건, 2018년 7월 경북 영천 경찰관 살인 사건, 그리고 2019년 6월 고속도로 역주행 사건 등 조현병 환자에 의한 몇몇의 폭력 사건이 언론을 통해 크게 다루어진 바 있다. 진주 사건과 함께 이같은 사건 사고와 이에 대한 사법 판결에 관한 기사들이 본 연구에서 조현병 관련 주요 단어 및 토픽으로 반영되고 있었다. 그리고 본 연구 결과는 사건 사고에 대한 단어나 토픽들과



**Fig. 2.** Co-occurrence network of 20 keywords based on term frequency-inverse document frequency (TF-IDF).

**Table 3.** Ten topics from latent Dirichlet allocation model within news articles on schizophrenia

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
경찰-진주	연구-결과	여성-학교	센터-관리	범죄-법원	병원-입원	장애인-시설	사회-문제	조현병-증상	가족-심리
경찰	연구	여성	관리	범죄	병원	정신질환	사회	치료	생각
아파트	교수	서울	지원	사건	환자	정신	문제	조현병	가족
진주	결과	시간	센터	범행	입원	장애인	국가	환자	마음
주민	이상	학교	정신건강	살해	정신	발생	한국	증상	아들
흥기	미국	남성	정신건강증진센터	상태	의사	정신장애	제공	중단	생활
사건	질환	상황	정부	판단	가족	시설	이유	편견	심리
경남	약물	사건	예산	인정	진료	기준	필요	증세	경험
신고	우울증	현장	지역	이유	강제	시행	국민	정신건강의학과	아버지
조사	발견	내용	위험	감정	정신건강의학과	전체	중요	망상	상당
차레	검사	사망	대책	법원	퇴원	부족	사용	악화	과정
혐의	가능성	폭력	마련	피의자	전문	장애	가능	사실	이해
출동	분석	목소리	서비스	피해자	의료기관	운영	심각	교수	모습
경찰관	국내	신청	지역사회	처벌	진단	환자	환경	행동	친구
범행	질병	주장	체계	능력	보호자	국회	안전	주변	어머니
이웃	설명	확인	대응	결정	거부	지적	인식	방치	고통
발생	장애	발생	중증	주장	의료	전문가	책임	정신분열병	세상
위험	건강	조현병	필요	국민	동의	불과	해결	본인	시작
현장	대상	상대	인력	범죄자	임세원	생활	부분	개선	정도
병력	진행	시작	강화	분석	위험	제도	다양	적절	부모
확인	영향	시민	퇴원	계획	필요	정도	의미	관리	가정
Occurrence ratio of the topic (%)									
12.20	10.06	9.27	10.05	9.98	10.20	9.15	9.60	9.92	9.57

연관되어 강제 입원이나 정신건강복지센터 관리에 대한 내용들이 기사에서 다루어지고 있음을 보여준다. 즉, 언론은 조현병 환자에 의한 사건 사고에 대한 논의 과정에서 환자들의 입원 치료와 지역사회 관리 이슈들을 사회 문제로서 거론하고 있었다. 이외에도 분석된 기사들 가운데에는 조현병 질환이나 연구에 대한 중립적인 내용도 포함되어 있었으며, 젠더나 장애인 문제가 조현병과 함께 다루어지기도 하였다.

본 연구에서는 감정 분석(sentiment analysis)을 수행하지 않아 주요 등장 단어나 토픽에 대한 주관적 가치 판단에 대한 내용은 확인할 수 없었다. 감정 분석은 텍스트에서 특정한 감정이 나타나는지 혹은 얼마나 나타나는지를 분석하는 것을 말한다.<sup>14)</sup> 그러나 본 연구 결과에서 조현병 환자에 의한 폭력 사건과 연관되는 단어나 토픽들이 많은 수 등장했다는 점은 최근 2년간 조현병 관련 언론 기사들이 대체로 부정적인 논조로 조현병에 대한 낙인 요인을 나타내었을 것으로 예상해 볼 수 있다. 실제로 기존 연구들을 통해 정신건강이나 정신질환 관련 국내 언론 보도 중에는 사건 사고 기사의 비중이 높으며, 기사 논조 또한 대체로 부정적인 경우가 많다는 사실이 확인된 바 있다. 1998년부터 2000년까지 2년간 ‘정신병’ 관련 신문 기사를 326건을 분석한 Kim 등<sup>23)</sup>의 연구에 따르

면 부정적 기사의 비율이 69.9%로 높았으며, 부정적 기사 중 ‘정신병 환자는 위험하거나 난폭하며 범죄를 잘 저지른다.’는 내용을 포함한 기사가 거의 절반을 차지하고 있었다. 정신건강, 정신질환, 그리고 자살에 대한 2009년 인터넷 기사 1,495건을 분석한 Lee 등<sup>24)</sup>의 연구에서는 부정적 관점으로 분류된 기사가 전체 기사 중 23.9%를 차지했으며, 부정적 관점의 기사 중 93%는 사건 사고와 관련된 기사로 나타났다. 2006년부터 2016년까지 10년간 5대 주요 일간지의 1,028개 정신질환 관련 기사를 분석한 Paek 등<sup>25)</sup>의 연구에 따르면 전체 기사 중 41.1%가 정신질환 관련 낙인 요인을 보도하고 있었다. 2016년 1월부터 2017년 9월까지 13개 일간지에 보도된 정신건강 관련 기사 1,011건을 분석한 Hwang과 Na<sup>26)</sup>의 연구에 따르면, 중립적 논조를 제외하고 봤을 때 부정적 논조의 기사 수가 긍정적 논조의 기사에 비해 약 2배 많았다(14.6% vs. 7.5%). 그리고 해당 연구에서 ‘조현병’으로 검색된 82건의 기사만을 놓고 보았을 때에는, 부정적 논조의 기사 수는 25건(30.5%)로 긍정적 기사 5건(6.1%)의 5배에 해당하였으며, 약 절반의(42건) 조현병 관련 기사들에서 정신건강 관련 사건, 사고 혹은 관련 법적 도덕적 분쟁을 보도하는 갈등 프레임을 다루고 있었다.<sup>26)</sup>

앞서 언급한 언론 분석 연구들은 주로 코드북(codebook)과 코더(coder)를 이용한 내용 분석을 통해 수행되었다. 그러나 오늘날 대규모 데이터를 분석하는데 있어서 이와 같은 전통적 분석 방법은 코드북의 타당성과 코더의 신뢰도 확보 문제, 그리고 방대한 텍스트를 분석하는데 필요한 시간과 비용 등 한계를 갖는다.<sup>8)</sup> 반면, 텍스트 마이닝 기법은 대규모 텍스트 데이터에서 의미를 구성하는 최소 단위의 텍스트 구성 요소를 파악하고, 이 요소들에 수학적 알고리즘을 적용하여 어떤 텍스트 구성 요소가 텍스트를 주도적으로 설명하는지 또는 텍스트의 의미를 예측하는데 효과를 나타내는지를 쉽게 정량화 해낼 수 있다.<sup>14)</sup> 이에 최근에는 인문, 사회 및 경제학 분야에서 텍스트 마이닝을 이용한 많은 연구들이 수행되고 있다.<sup>27,28)</sup> 그러나 정신건강 또는 정신질환에 대한 언론 보도 형태를 텍스트 마이닝 기법으로 분석한 시도는 국제적으로도 그 수가 많지 않으며, 본 연구는 국내 기사에 대한 최초의 텍스트 마이닝 연구에 해당한다. Koike 등<sup>29)</sup>은 1985년부터 2013년까지 일본 4개 신문사와 1개 TV 뉴스에서 보도된 조현병 관련 기사 총 6,361건을 분석하여 2002년 일본에서 시행된 '통합실조증(統合失調症)'으로의 병명 개정이 언론 보도에 미친 영향을 텍스트 마이닝 기법을 이용하여 분석한 바 있다. 상기 연구의 빈도 분석에서 등장 빈도 상위 20개의 단어 중에는 '지방병원', '판결', '피고', '죽임(killing)', '투옥', '타살', 그리고 '자상(stabbing)' 등 폭력 사건 사고와 관련된 단어들이다수 포함되어 있었으며,<sup>29)</sup> 이는 본 연구에서 확인된 국내 조현병 관련 기사의 다빈도 등장 단어들과도 같은 맥락에 있다고 볼 수 있다.

본 연구는 다음과 같은 제한점을 갖는다. 첫째, 본 연구는 2018년부터 2019년까지 비교적 짧은 기간 동안 보도된 조현병 기사를 분석 대상으로 하였으므로, 연구 결과는 해당 기간 동안 주목을 끌었던 몇몇 조현병 관련 이슈들을 반영하고 있다. 이에 본 연구 결과가 조현병에 대한 국내 언론의 전반적인 보도 양상을 설명한다고는 볼 수 없다. 둘째, 본 연구에서는 형태소로서 명사만 분석되었으며, 동사나 형용사는 분석에 포함되지 않았다. 또한 한글형태소 사전의 한계로 고유 명사나 전문 용어 일부도 분석에 포함되지 않았을 가능성이 있다. 셋째, 동시 출현 단어를 분석하기 위하여 단어 간 네트워크를 구성하고 이를 시각화 하였으나, 중심성(centrality) 분석 등 세부적인 네트워크 분석은 본 연구에 포함시키지 않았다. 이는 본 연구의 수행 범위를 넘을 것으로 판단되며, 향후 추가 연구를 통해 다루도록 하겠다. 넷째, 가용한 한국어 감성어 사전(sentiment lexicon)이 없는 관계로 본 연구에서는 위에서 언급한 바와 같이 감정 분석을 수행되지 않았다.

이와 같은 제한점에도 불구하고, 본 연구는 조현병 관련 언론 기사를 텍스트 마이닝 기법으로 분석한 국내 최초의 연구로, 최근 언론에서 다루어진 조현병 관련 이슈를 정량적으로 확인하였다. 향후 연구를 통해 텍스트 데이터에 대한 수학적 분석 방법을 이용하여 조현병 관련 사회 이슈와 그에 대한 언론 동향, 그리고 소셜 미디어에서의 여론을 통합적으로 분석해 볼 필요가 있다.

## 결론

본 연구는 최근 수 년 동안 언론에서 조현병 환자에 의한 사건 사고, 특히 폭력 문제가 조현병 관련 주요 이슈로서 보도되었음을 텍스트 마이닝 분석을 이용하여 정량적으로 확인하였으며, 이에 대한 논의에서 조현병 환자의 입원이나 지역 사회 관리에 관한 토픽이 함께 제기되고 있음을 보여주었다. 최근 일련의 조현병 관련 사건 사고 위주의 보도 기사들은 기존의 정신질환에 대한 언론의 부정적 논조와 낙인 요인을 더욱 가중시킬 수 있다는 점에서, 이에 대한 언론 매체와 정신의학계 간 공동의 논의와 숙고가 요구된다.

**중심 단어:** 기사·언론·조현병·텍스트 마이닝·폭력.

## ORCID iDs

Hee Jung Nam <https://orcid.org/0000-0001-8375-7516>  
Seunghyong Ryu <https://orcid.org/0000-0001-6127-760X>

## REFERENCES

- 1) Kim A. Government announces measures to deal with mental illness in light of rise in crimes. The Korea Herald 2019 May 15. Available from URL: <http://www.koreaherald.com/view.php?ud=20190515000759>.
- 2) Korean Society for Schizophrenia Research. Statement from the Korean Society for Schizophrenia Research on the police officer killed. The Korean neuropsychiatric association newsletter 2018 July 16. Available from URL: <http://www.knpanews.or.kr/news/articleView.html?idxno=69>.
- 3) Corrigan P. How stigma interferes with mental health care. Am Psychol 2004;59:614-625.
- 4) Lauber C, Rossler W. Stigma towards people with mental illness in developing countries in Asia. Int Rev Psychiatry 2007;19:157-178.
- 5) Corrigan PW, Penn DL. Lessons from social psychology on discrediting psychiatric stigma. Am Psychol 1999;54:765-776.
- 6) Park JI, Jeon M. The stigma of mental illness in Korea. J Korean Neuropsychiatr Assoc 2016;55:299-309.
- 7) Goepfert NC, von Heydendorff SC, Dreßing H, Bailer J. Effects of stigmatizing media coverage on stigma measures, self-esteem, and affectivity in persons with depression-an experimental controlled trial. BMC Psychiatry 2019;19:138.
- 8) Krippendorff K. Content analysis: an introduction to its methodology. Thousand Oaks: Sage;2013.
- 9) Rajman M, Besançon R. Text Mining: natural language techniques

- and text mining applications. In: Spaccapictra S, Maryanski F, editors. *Data mining and reverse engineering*. Boston: Springer;1998. p.50-64.
- 10) Cho J, Garcia-Molina H, Page L. Efficient crawling through URL ordering. *Comput Netw Isdn Sys* 1998;30:161-172.
  - 11) Wickham H. rvest: Easily Harvest (Scrape) Web Pages. R package version 0.3.4. 2019.
  - 12) Feinerer I, Hornik K. tm: Text Mining Package. R package version 0.7-6 2018.
  - 13) Jeon H. KoNLP: Korean NLP Package. R package version 0.80.1. 2016.
  - 14) Baek YM. Ext-mining using R. Paju: Hanulmplus;2017.
  - 15) Lang D, Chien GT. wordcloud2: create word cloud by ‘htmlwidge’. R package version 0.2.1. 2018.
  - 16) Lee SS. A Content analysis of journal articles using the Language Network Analysis Methods. *J Korean Soc Inform Manage* 2014; 31:49-68.
  - 17) Epskamp S, Cramer AOJ, Waldorp LJ, Schmittmann VD, Borsboom D. Qgraph: network visualizations of relationships in psychometric data. *J Stat Softw* 2012;48:18.
  - 18) Blei DM, Lafferty JD. A correlated topic model of science. *Ann Appl Stat* 2007;1:17-35.
  - 19) Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. *J Mach Learn Res* 2003;3:993-1022.
  - 20) Roberts M, Stewart B, Tingley D. Stm: R package for structural topic models. *J Stat Softw* 2019;91:1-40.
  - 21) Chang J. Lda: collapsed Gibbs sampling methods for topic models. R package version 1.4.2. 2015.
  - 22) Griffiths TL, Steyvers M. Finding scientific topics. *Proc Natl Acad Sci U S A* 2004;101(Suppl 1):5228-5235.
  - 23) Kim SW, Yoon JS, Lee MS, Lee HY. The analysis of newspaper-articles on psychosis. *J Korean Neuropsychiatr Assoc* 2000;39:838-848.
  - 24) Lee MS, Ko JA, Oh JH, Shin YS, Park JS. Trend of internet media report for mental health and suicide. *J Korean Assoc Soc Psychiatry* 2009;14:97-104.
  - 25) Paek HJ, Cho HJ, Kim JH. Content analysis of news coverage on stigma and attribution regarding mental illness. *Korean J Journalism Commun Stud* 2017;61:7-43.
  - 26) Hwang A, Na EY. Content analysis of news coverage with regard to mental health: focusing on news frame and tone. *Health Soc Welf Rev* 2018;38:290-322.
  - 27) Choi JW, Han HS, Lee M, An JM. The prediction of corporate bankruptcy using text-mining methodology. *Product Rev* 2015;29: 201-228.
  - 28) Kang B, Song M, Jho W. A study on opinion mining of newspaper texts based on topic modeling. *J Korean Soc Lib Inf Sci* 2013; 47:315-334.
  - 29) Koike S, Yamaguchi S, Ojio Y, Ohta K, Ando S. Effect of name change of schizophrenia on mass media between 1985 and 2013 in Japan: a text data mining analysis. *Schizophr Bull* 2016;42:552-559.