

# 조현병과 정신분열병에 대한 뉴스 프레임 분석을 통해 본 사회적 인식의 변화\*

## Text Mining Driven Content Analysis of Social Perception on Schizophrenia Before and After the Revision of the Terminology

김 현 지 (Hyunji Kim)\*\* , 박 서 정 (Seojeong Park)\*\*\*  
송 채 민 (Chaemin Song)\*\*\*\* , 송 민 (Min Song)\*\*\*\*\*

### 목 차

- |           |          |
|-----------|----------|
| 1. 서론     | 4. 연구 결과 |
| 2. 이론적 배경 | 5. 결론    |
| 3. 연구 방법론 |          |

### 초 록

질환자에 대한 사회적 낙인을 제거하기 위해 2011년, 대한의사협회는 정신분열병을 '조현병'으로 개정하였다. 병명을 변경한 지 약 9년이 지났지만 실제로 사회적 인식이 어느 정도 변화하였는지 정량적으로 분석한 연구는 아직 없다. 이에 본 연구에서는 병명 개정에 따른 조현병에 대한 사회적 인식 변화를 확인하고자 네이버 뉴스 기사를 수집·분석하였다. 텍스트 분석에는 LDA 토픽 모델링, TF-IDF, 동시출현 단어, 감성분석 기법을 사용하였다. 분석 결과, 병명 개정 전보다 후에, 그리고 병명 개정 후 병용되는 정신분열병과 조현병 중 조현병에 대한 사회적 인식이 더 부정적인 것으로 나타났다. 즉, 병명 개정이 낙인을 해소하지 못했음을 알 수 있었다.

### ABSTRACT

In 2011, the Korean Medical Association revised the name of schizophrenia to remove the social stigma for the sick. Although it has been about nine years since the revision of the terminology, no studies have quantitatively analyzed how much social awareness has changed. Thus, this study investigates the changes in social awareness of schizophrenia caused by the revision of the disease name by analyzing Naver news articles related to the disease. For text analysis, LDA topic modeling, TF-IDF, word co-occurrence, and sentiment analysis techniques were used. The results showed that social awareness of the disease was more negative after the revision of the terminology. In addition, social awareness of the former term among two terms used after the revision was more negative. In other words, the revision of the disease did not resolve the stigma.

키워드: 정신분열병, 조현병, 텍스트 마이닝, 토픽 모델링, TF-IDF, 동시출현 단어, 감성분석  
Schizophrenia, Text Mining, Topic Modeling, TF-IDF, Word Co-Occurrence, Sentiment Analysis

- \* 본 연구는 2018년 대한민국 교육부와 한국연구재단의 지원을 받아 수행되었음 (NRF-2018S1A3A2075114).  
\*\* 연세대학교 문헌정보학과 석사과정(d2o2ji@naver.com / ISNI 0000 0004 7933 052X) (제1저자)  
\*\*\* 연세대학교 문헌정보학과 석사과정(pseojeong@yonsei.ac.kr / ISNI 0000 0004 7933 0749) (공동저자)  
\*\*\*\* 연세대학교 문헌정보학과 석사과정(chaemin11@yonsei.ac.kr / ISNI 0000 0004 7933 108X) (공동저자)  
\*\*\*\*\* 연세대학교 문헌정보학과 교수(min.song@yonsei.ac.kr / ISNI 0000 0000 8203 196X) (교신저자)  
논문접수일자: 2019년 10월 22일 최초심사일자: 2019년 11월 13일 게재확정일자: 2019년 11월 22일  
한국문헌정보학회지, 53(4): 285-307, 2019. <http://dx.doi.org/10.4275/KSLIS.2019.53.4.285>

## 1. 서론

‘정신분열병’ 또는 ‘조현병’이란 망상, 환청, 와해된 언어, 정서적 둔감 등의 증상과 더불어 사회적 기능에 장애를 일으키는 질환이다. 예후가 좋지 않고 만성적인 경과를 보이지만, 최근 약물 요법을 포함한 치료적 접근에 뚜렷한 진보가 있어 조기 진단과 치료에 적극적인 관심이 필요한 질환이다(서울대학교병원 2010). 2016년 정신질환실태 역학조사 보고서(삼성서울병원 2016)에 따르면 2016년 조현병 추정 환자는 약 11만 3천 명, 전체 국민의 약 0.22%로 추산된다. 조현병이 전 세계적으로 약 1%의 평생 유병률을 보이는 점을 고려했을 때, 현재 집계된 환자는 전체 환자의 약 5분의 1 수준이라고 유추할 수 있다.

정신분열병은 ‘분열병’이 가진 부정적인 어감 때문에 사회적으로 많은 오해를 샀지만 다음의 세 가지 이유에서 위험 질환이라고 단정 짓긴 어렵다. 첫째, 정신분열병은 초기 진단과 지속적 치료 시 자·타해 위험성이 매우 낮다. 실제로 정신질환자에 의한 범죄율은 일반인의 범죄율보다 훨씬 낮다. 대검찰청의 2018년 범죄분석(2018)에 따르면 2017년에 검거된 살인 범죄자의 47.3%가 정상이며 43.4%는 주취 상태, 정신질환이 있는 경우는 9.3%인 것으로 나타났다. 방화 범죄자 역시 13%만이 정신질환자였다. 물론 진주 아파트 방화·살인 사건, 강북삼성병원 교수 살인 사건 등 최근에 발생한 강력 범죄의 범인이 정신분열병 환자인 것으로 밝혀졌지만, 이들은 모두 범행 전 정신과 약물 복용을 중단하고 고립된 생활을 하면서 사회에 대한 반감을 키운 것으로 밝혀졌다. 즉, 정신분

열병 환자에 의한 범죄는 병 자체의 위험성보다는 지속적인 치료의 부재 때문으로 볼 수 있다. 둘째, 정신분열병은 환자마다 임상 양상이 매우 다양하게 나타난다. 임상 양상으로는 약물에 의해 잘 호전되는 불안, 불면, 짜증부터 환각, 환청, 망상 등이 있다. 마지막으로 정신분열병은 조기 진단 및 지속적 치료 시 별다른 장애 없이 사회로 복귀할 수 있다(서울대학교병원 2010).

이처럼 정신분열병은 스펙트럼이 넓고 회복 가능성이 크며 범죄로 이어질 확률이 낮음에도 불구하고, 환자에게 가해지는 낙인은 심각하였다. 분열환자들은 일종의 인격장애라고 불리는 반사회적 인격장애 환자나 다른 정신 질환 환자가 저지른 범죄까지 떠안아야 했다(대한조현병학회 2011). 여기서 낙인이란 “낙인찍히는 사람의 부정적인 속성만을 지칭하는 것이 아니라 이 속성으로부터 기인하는 인지적 고정관념(Stereotype)·감정적 편견(Prejudice)·행동적 차별(Discrimination)을 모두 포함하는 연속적인 과정”을 말한다(백혜진, 조혜진, 김정현 2017; Overton and Median 2008; Penn and Martin 1998). 사회적 낙인은 환자의 자아존중감 및 정신 건강 회복을 감소시키며, 환자가 지속적인 병원 방문과 약물복용을 지연하게 한다(Jorm 2000). 결과적으로 환자는 재발을 경험하고 만성화의 길로 들어서는 것이다.

결국, 치료 순응도와 재발에 영향을 주는 사회적 낙인을 제거하기 위해 2011년 3월, 대한의사협회는 정신분열병을 ‘조현병’으로 개정하였다. 병명을 변경한 지 약 9년이 지난 현시점에서 실제로 사회적인 인식이 어느 정도 변화하였는지 정량적으로 분석할 필요성이 있다.

이에 본 연구는 병명 개정에 따른 조현병에

대한 사회적 인식의 변화가 있는지를 확인하고자 한다. 또한, 병명 개정 후 병용되는 정신분열병과 조현병에 대한 사회적 인식의 차이를 확인함으로써 바람직한 보도 방향을 정립하고 올바른 치료환경을 구축하는 데 기여하고자 한다. 그러므로 본 연구에서는 정신분열병과 조현병에 관련된 네이버 뉴스 기사를 수집하고 다양한 텍스트 마이닝 기법을 수행하여 두 병명에 대한 대중의 인식을 거시적 측면과 미시적 측면에서 분석하고자 한다.

## 2. 이론적 배경

### 2.1 온라인 뉴스와 사회적 인식의 연관성

여론을 틀 짓는 언론의 기능을 프레임(Frame)이라 한다. 프레임은 뉴스 기사에서 특정 키워드와 은유, 개념, 시각적 이미지를 선택적으로 사용하여 이벤트를 전달함을 말한다. 언론은 프레임 설정을 통해 특정 입장을 지지하거나 반박하며, 이벤트에 대한 독자의 인식에 영향력을 행사한다(Entman 1991). 이러한 점에서 언론학자들은 프레임링을 미디어의 힘이라고 규정한다(최영해 2019).

뉴스 매체가 여론 형성에 결정적인 역할을 한다는 사실은 많은 연구에서 확인되었다. 이준웅(2005)는 공공 이슈를 둘러싼 언론매체의 뉴스 보도는 특정 프레임을 통해 수용자 시민의 의견 형성에 영향력을 행사함을 확인했으며, Dearing과 Everett(1996)은 뉴스 매체가 공공의제를 결정할 뿐 아니라 정책의제도 설정할 만큼 막대한 영향력을 지님을 확인하였다. 또

배영(2014)은 뉴스 보도가 문화콘텐츠에 대한 관심을 유지하고 소비를 촉발함을 밝혔으며, 김민철(2017)은 부정적 언론 보도로 인해 기업 이미지와 브랜드 태도의 손실이 나타남을 확인하였다.

정신질환에 대한 여론도 예외는 아니다. 보통 공중은 언론을 통해 정신질환에 대한 다양한 사건 및 정책, 연구 결과를 접하고, 정신질환에 대한 인식이나 태도에 상당한 영향을 받는다(Goulden et al. 2011; Rhydderch et al. 2016; 백혜진, 조혜진, 김정현 2017). 예컨대 Corrigan 외 2명(2013)은 정신질환의 회복에 관한 기사가 사회적 낙인을 줄이고 긍정적인 태도를 증가시킨 반면, 공중 정신건강 시스템의 역기능에 관한 기사는 낙인을 심화하고 긍정적인 태도를 감소시킴을 증명하였다. 또한, 조수영, 김정민(2010)은 정신질환에 대한 부정적 혹은 잘못된 정보의 전달이 사회적으로 정신질환자에 대한 부정적 편견을 형성하며, 이러한 편견은 질환자의 정신과 치료를 방해하게 됨을 확인하였다. 이처럼 미디어는 정신건강에 대한 사회적 인식 조성에 중요한 역할을 수행하며, 건강정보를 폭넓게 전파하기 위한 효과적인 수단이라고 볼 수 있다(황애리, 나은영 2018).

한편, 기존 오프라인 신문과 달리 인터넷을 바탕으로 한 온라인 신문은 지면과 시간의 제약에서 보다 자유롭고, 상호작용성(interactivity)을 바탕으로 한 활발한 의견수렴이 가능하여 오프라인 신문보다 시민 저널리즘의 가능성이 크다(반현 2000). 그리고 뉴스 이용집중도 조사보고서 2016-2018(한국언론진흥재단 2018)에 따르면 포털 뉴스의 이용 점유율은 2015년 27.9%에서 2018년 35.8%로 대폭 상승하였으며, 모바일

일 인터넷을 통해 뉴스를 확인한다는 응답자는 19.5%에서 80.8%로 4배 이상 증가하였다. 온라인 뉴스의 영향력은 온라인 뉴스 기사의 수와 소셜미디어 메시지 수 간의 양의 상관관계를 통해서도 입증된 바 있다(Mollema et al. 2015).

이러한 선행연구들을 참고하여 본 연구에서는 뉴스 프레임을 사회적 인식으로 보고, 정신분열병 및 조현병 관련 온라인 뉴스 기사 분석을 통해 사회적 인식에 어떤 특징이 있는지 고찰하고자 하였다.

## 2.2 정신질환에 대한 언론 보도 분석 연구

지금까지 정신질환에 대한 언론 보도 내용 분석은 다양한 방식으로 행해져 왔다. 연구자가 직접 독해, 코딩하여 전반적인 뉴스 프레임과 기사 논조를 분석한 연구가 있는가 하면(백혜진, 조혜진, 김정현 2017), 네이버트렌드, 구글트렌드 등의 빅데이터 분석 도구(김혜선, 박도원, 홍영은 2018)나 텍스트 마이닝 기법을 사용해(감미아, 송민 2012) 보도 경향을 탐색한 연구도 있었다. 이 연구들은 모두 사전에 개발한 조작적 정의에 따라 수집한 데이터를 분류함으로써 분석의 객관성을 높이고자 하였다는 공통점을 가진다.

이에 본 연구에서도 주제 및 키워드의 객관적 분석을 위해 황애리와 나은영(2018)이 선행연구(노수진, 윤영민 2013; 정의철 2011; 조수영, 김정민 2010)를 근거로 제시한 뉴스 프레임의 조작적 정의를 수정, 활용하였다. <표 1>은 최종적으로 분석에 쓰인 6가지 뉴스 프레임의 정의이다. 선행연구의 예방 프레임과 위기 프레임은 의학적 프레임, 사회적 프레임과의 중복 가능성 때문에 제외하였다.

## 3. 연구 방법론

### 3.1 연구 질문

본 연구는 조현병 환자의 조기 진단과 지속적인 치료를 저해하는 요인이 될 수 있는 사회구성원의 인식이 병명 개정을 기점으로 어떻게 변화하였는지를 파악하고 이를 바탕으로 인식 개선 방안을 도출하고자 정신분열병 및 조현병 관련 뉴스 기사를 분석하였다. 이를 위한 연구 질문은 다음과 같다.

- RQ1. 병명 개정 전과 후의 조현병에 대한 사회적 인식은 어떠한 차이를 보이는가?

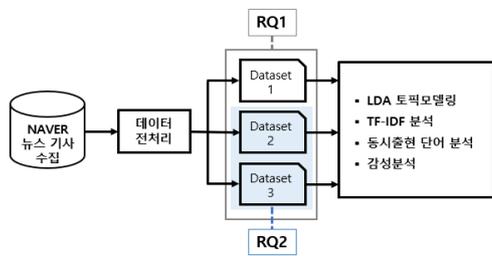
<표 1> 뉴스 프레임의 정의

프레임	내용
의학적	정신질환의 원인과 증상, 관련 약물, 검진 및 치료, 의학적 발견 및 새로운 치료법 소개
갈등	정신건강 관련 사건, 사고 혹은 법적/도덕적 분쟁 보도
사회적	정신건강 관련 실정, 사회문제 보도
정책	정신건강 관련 정부의 복지와 예산 등 정책 문제 보도
인간적 흥미	개인이나 유명인이 겪은 정신건강 관련 일화를 제시 및 이벤트 소개
상업적	특정 예술작품이나 기업, 서비스 관련 보도 등 정신건강 관련 정보를 상업적 관점에서 다룸

- RQ2. 병명 개정 후 정신분열병에 대한 사회적 인식과 조현병에 대한 사회적 인식은 어떠한 차이를 보이는가?

### 3.2 연구방법 및 절차

본 연구는 <그림 1>과 같이 데이터 수집, 전처리, 분석 순으로 진행하였다.



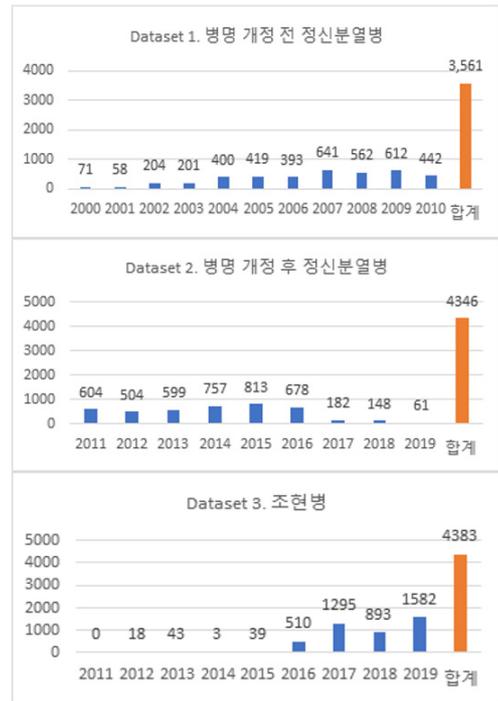
<그림 1> 연구 절차

데이터 전처리 후의 모든 분석은 병명 개정 전 정신분열병(증) 키워 데이터, 병명 개정 후 정신분열병(증) 키워 데이터, 조현병(증) 키워 데이터로 나누어 진행하였으며, 분석 기법으로는 LDA 토픽 모델링, TF-IDF, 동시출현단어, 감성분석을 사용하였다.

### 3.3 데이터 수집 및 전처리

본 연구에서는 언론 보도 분석을 위해 온라인 뉴스 기사를 활용하였다. 뉴스 기사는 네이버에서 검색결과로 제공하는 약 800여 개의 언론사가 2000년 1월 1일부터 2019년 5월 31일까지 보도한 뉴스 기사 본문을 대상으로 하였다. 네이버는 국내 이용률이 64%로 가장 높은 미디어 브랜드이며 이용자의 평균 정치 성향이

진보, 중도, 보수를 모두 아우른다(한국언론진흥재단 2018). 이처럼 네이버의 보편성은 연구 결과를 일반화하는 데 타당성과 신뢰성을 부여한다고 판단해 이를 수집 대상으로 결정하였다. 검색어는 '정신분열병'과 그 동의어인 '정신분열증', '조현병'과 그 동의어인 '조현증'으로 하였으며, Python을 활용하여 자동 수집하였다. 수집된 기사 총 20,332건 중 중복되는 기사와 본문에 두 검색어가 모두 등장한 기사는 분석에서 제외하였다. 최종적으로 분석에 사용된 뉴스 기사의 수는 <그림 2>와 같다.



<그림 2> 연도별 뉴스 기사 수

중복 기사를 제외한 병명 개정 전 정신분열병(증) 기사 3,561건, 병명 개정 후 정신분열병(증) 기사 4,346건, 조현병(증) 기사 4,383건, 총

12,290건의 뉴스 기사가 분석에 사용되었다. 연도별 최소 0건에서 최대 1,582건까지 다소 변동이 있으나 이는 병명 개정에 따른 자연스러운 현상이며, 비교 대상이 되는 각 구간의 합계는 일정한 분포를 보이므로 분석에 적합한 데이터셋(Dataset)으로 볼 수 있다.

분석에 앞서 데이터 전처리를 위해 KOMORAN 2.4 형태소 분석기로 내용에 해당하는 형태소를 추출하였다. 이때 내용어는 일반명사(NNG)와 고유명사(NNP)로 한정하였다. 그러나 형태소 자동 추출법은 분석기 코퍼스가 포함하고 있는 단어의 유한함 때문에 질병명, 약물명, 성분명, 인물명 등 뉴스 기사가 가지고 있는 다양한 고유명사 및 합성명사를 제대로 추출하지 못했다. 이를 해결하기 위해 n-gram 기법을 사용하여 연속된 n개의 일반명사(NNG), 고유명사(NNP), 의존명사(NNB)로 이루어진 단어를 뽑은 후 정의된 임계값(2로 설정) 미만인 단어들은 필터링하였다(이민철, 김혜진 2018). 필터링된 키워드 중 실제로 쓰이는 단어만 선별하여 코퍼스에 추가하였다. <표 2>는 코퍼스에 추가된 토큰의 예시이다.

그 외 각종 언론사와 데이터 특성상 반복되는 '배포', '저작권', '취재', '허브'와 같은 단어 및 오타들은 불용어 처리하였다.

### 3.4 데이터 분석

본 연구에서는 정신분열병과 조현병에 대한 사회적 인식 분석을 거시적 분석과 미시적 분석으로 나누어 접근하였다. 거시적 분석으로는 LDA 토픽 모델링을 사용하여 뉴스 기사의 전반적 특성을 살펴보았으며, 미시적 분석으로는 TF-IDF, 동시출현 단어, 감성분석 기법을 사용하여 문헌 간의 관계와 맥락적 특성을 살펴보았다. 또한, 데이터셋은 병명 개정 전후, 병명 간의 사회적 인식 차이를 확인하기 위해 병명 개정 전 정신분열병 뉴스 기사(2000.01.01.~2010.12.31.), 병명 개정 후 정신분열병 뉴스 기사(2011.01.01.~2019.05.31.), 병명 개정 후 조현병 뉴스 기사(2011.01.01.~2019.05.31.)로 나누어 분석을 수행하였다.

#### 3.4.1 Latent Dirichlet Allocation(LDA) 토픽 모델링

LDA(Latent Dirichlet Allocation)는 토픽 모델링 기법 중 가장 많이 활용되는 문헌 생성 모델로, 문헌 단위에서 각 주제의 분포로 문헌을 표현하며, 이러한 각 주제는 다시 문헌집단 단위에서 추론 가능한 용어의 확률분포로 표현할 수 있다(송민 2017). 본 연구에서는 병명 개정 전

<표 2> 코퍼스에 추가된 토큰의 예시

토큰화 예리	코퍼스에 추가한 토큰
자(NNB) + 간(NNB) + 전(NNG) + 증(NNG)	자간전증
리스(NNG) + 페리(NNP) + 돈(NNG)	리스페리돈
리스(NNG) + 페달(NNG) + 콘(NNG) + 스타(NNG)	리스페달콘스타
마릴린(NNP) + 먼로(NNP)	마릴린먼로
현주(NNP) + 건조물(NNG) + 방화죄(NNG)	현주건조물방화죄

후에 정신분열병과 조현병에 관한 언론 보도의 주제가 어떠한 차이를 보이는지 확인하기 위해 LDA 토픽 모델링을 하였다. 모듈은 Python의 Gensim을 사용하였으며, 결과는 pyLDAvis로 시각화하였다.

### 3.4.2 TF-IDF와 동시출현 단어 분석

본 연구는 병명 개정 전후로 정신분열병과 조현병 뉴스 데이터의 핵심 키워드를 확인하기 위해 TF-IDF와 동시출현 단어 분석을 수행하였다.

TF-IDF(Term Frequency-Inverse Document Frequency: 역문서 빈도) 가중치 모델은 텍스트 마이닝을 위해서 문서 내부의 단어가 얼마나 중요한지를 평가하기 위한 표현방식이다. TF-IDF 가치 값이 큰 단어일수록 속해 있는 문서의 주제나 의미를 결정지을 가능성이 크며 이 측정치를 주요 키워드를 추출할 수 있는 척도로 활용할 수 있다(이성직, 김한 2009). TF-IDF 분석 단계에서는 먼저 데이터셋 별로 각 문헌당 상위 5개 단어의 TF-IDF 값을 산출하였다. 그다음 유의미한 분석을 위해 각 데이터셋에서 TF-IDF 값을 기준으로 내림차순 정렬을 한 뒤, 상위 30개, 하위 3개 단어를 비교하며 그 차이를 분석하였다.

한편, 각 단어를 단순 빈도 분석한 결과는 뉴스 본문에서 어떤 맥락으로 등장했는지 파악하기 어렵다는 한계가 있다. 이에 본 연구는 단어 간의 동시출현 네트워크를 구축하여 단어들이 사용된 맥락을 분석하고자 동시출현 단어 분석을 실시하였다. 동시출현 단어 분석은 공기어(Co-occurrence) 분석이라고도 하는데, 이때 공기어란 두 단어가 같은 문맥에서 함께 나타

나는 현상을 포괄하는 용어를 말하는 것이기에, 이들은 서로 밀접한 의미 관계를 갖는다고 여겨진다(송민 2017). 또, 네트워크 전체에서 가장 중요한 키워드가 무엇인지 확인하기 위해 다른 키워드와의 동시출현 빈도로 중심성을 나타내는 연결 중심성(Degree Centrality)을 측정하였다.

### 3.4.3 감성분석

감성분석은 문서 작성자가 가지고 있는 감정을 파악하고 이를 객관적으로 수치화하는 분석 방법으로, 분석 대상은 공식적인 문서와 비공식적인 문서로 나뉜다(주성빈 2019). 본 연구에서는 언론이라는 공식적 문서의 감성분석을 통해 정신분열병과 조현병에 대한 사회적 인식에 차이가 있는지를 파악하고자 하였다. 감성의 정도를 측정하는 기준은 목적에 따라 다양하지만, 감성 사전에 수록된 단어의 출현 빈도로 감성을 파악하는 것이 가장 보편적이다(주성빈 2019). 하지만 이 방법은 주제에 적합한 감성 사전을 구축하는 것이 매우 중요하다. 따라서 본 연구에서는 분석에 앞서, 기존에 구축되어 있던 한글 감성 사전인 yTextMiner(<http://informatics.yonsei.ac.kr:8080/yTextMiner/home.html>)의 koreanSentiWordNet(송민 2017; Esuli, Sebastiani 2007; Baccianella, Esuli, Sebastiani 2010)에 긍정/부정 키워드를 추가하였다. 예를 들어, 정신질환자의 '회복, 성취'와 관련된 용어는 긍정적 키워드 사전에, '범죄, 반사회성'과 관련된 용어는 부정적 키워드 사전에 추가하였다. 그리고 한 기사에서 긍정적 단어의 출현 빈도만큼 +1, 부정적 단어의 출현 빈도만큼 -1을 하여 계산한 최종값을 해

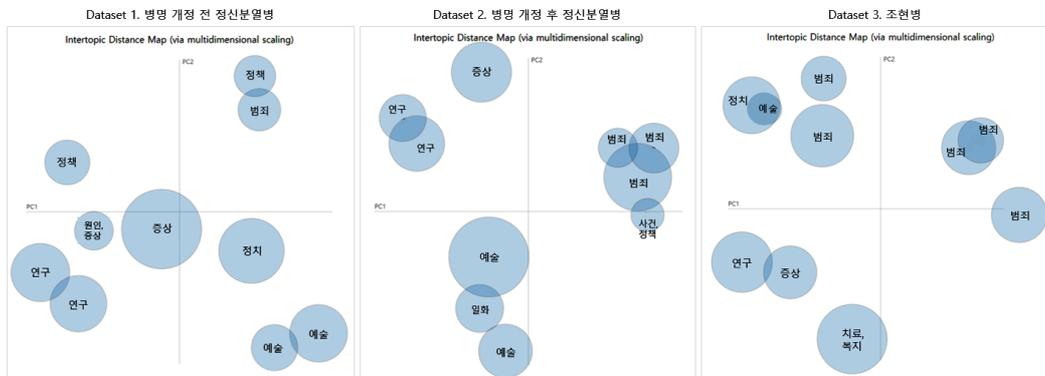
당 기사의 극성값(Polarity Score)으로 보았다. 극성값은 1에 가까울수록 강한 긍정을, -1에 가까울수록 강한 부정을 나타낸다.

## 4. 연구 결과

### 4.1 LDA 토픽 모델링

기간별(병명 개정 전/후), 검색어별(정신분열병·증/조현병·증)로 뉴스 기사를 나눈 데이터셋을 대상으로 LDA 토픽 모델링을 수행하였다. 토픽의 개수를 결정하는 방법으로 퍼플렉시티(Perplexity) 값 비교를 시도하였으나 값이 모든 구간에서 단조 감소하는 양상을 보였다. 그래서 본 연구에서는 토픽의 수를 5, 10, 20, 30으로 설정하여 토픽 모델링을 수행한 후, 토픽 간 유사도가 낮고, 동일한 토픽 내 주요 키워드들 간의 유사도가 높아 해석이 용이했던 토픽 10개를 최종 토픽 수로 하여 분석하였다. 토픽당 키워드는 20개를 추출하였으며, 구성 키워드의 연관성을 토대로 연구자 3명이 논의

하여 토픽에 주제명을 부여하였다. <그림 3>은 pyLDAvis 라이브러리를 사용하여 토픽을 시각화한 결과이며, <표 3>은 데이터셋별로 가장 많이 등장한 토픽 3개와, 각 토픽에 포함된 상위 5개 키워드 및 해당 키워드가 특정 토픽에 포함될 확률을 나타낸 표이다. <표 4>는 라벨링한 토픽을 뉴스 프레임의 관점에서 분석하기 위해 데이터셋에서 각 프레임이 차지하는 비중을 계산한 결과이다. 프레임의 비중은, LDA 알고리즘에 포함된 문서 내 주제의 비율 점수(“theta”)(Blei, Ng and Jordan 2003)를 합산한 다음 총 문서 수로 나눈 값을 a라고 할 때, 각 프레임에 속하는 토픽의 a 값을 모두 더하여 구하였다(Guo et al. 2016). 토픽은 다음과 같이 분류하였다: 의학적 프레임(증상, 연구, 원인, 치료), 갈등 프레임(범죄), 정책 프레임(정책, 복지), 인간적 흥미 프레임(일화), 상업적 프레임(예술). 또한, 데이터셋 1과 3의 ‘정치’ 토픽은 데이터 수집 시 기사 본문 하단의 ‘오늘의 인기 기사’, ‘추천 기사’ 등 정신분열병/조현병과 관련 없는 기사가 같이 수집되어 나온 것이기에 분석 시 고려하지 않았다.



<그림 3> 토픽 분포

〈표 3〉 토픽 모델링 결과

토픽	병명 개정 전 정신분열병			병명 개정 후 정신분열병			조현병		
	증상	연구	예술	예술	범죄	증상	치료·복지	범죄	연구
1	치료 (0.031)	치료제 (0.024)	영화 (0.012)	사회 (0.011)	경찰 (0.036)	치료 (0.039)	환자 (0.035)	사건 (0.031)	연구 (0.023)
2	환자 (0.025)	개발 (0.015)	사랑 (0.008)	생각 (0.010)	여성 (0.024)	정신 (0.031)	치료 (0.034)	범죄 (0.026)	뇌 (0.014)
3	우울증 (0.014)	미국 (0.011)	삶 (0.007)	마음 (0.008)	사건 (0.020)	환자 (0.030)	입원 (0.027)	살인 (0.013)	교수 (0.009)
4	증상 (0.011)	신약 (0.010)	생각 (0.007)	부모 (0.005)	흥기 (0.016)	증상 (0.015)	건강 (0.021)	안인득 (0.012)	개발 (0.009)
5	자살 (0.011)	임상시험 (0.008)	그림 (0.006)	책 (0.004)	살인 (0.015)	우울증 (0.013)	센터 (0.016)	피해자 (0.011)	미국 (0.008)

〈표 4〉 뉴스 프레임별 비중

	의학적	인간적 흥미	갈등	상업적	정책	사회적	기타(정치)
정신분열병	58.20%	·	5.51%	14.04%	11.24%	·	11.01%
정신분열병	31.13%	8.80%	33.64%	22.2%	4.23%	·	·
조현병	27.25%	·	52.10%	2.93%	7.11%	·	10.61%

〈표 4〉에 따르면 병명 개정 전 정신분열병 기사에는 (1) 의학적 프레임이 가장 많았으며, 이어 (2) 상업적, (3) 정책, (4) 갈등 프레임 순으로 나타났다. 한편, 병명 개정 후 정신분열병 기사에는 (1) 갈등, (2) 의학적 프레임이 가장 많았으며, 이어 (3) 상업적, (4) 인간적 흥미, (5) 정책 프레임 순으로 나타났다. 마지막으로, 조현병 기사에는 (1) 갈등 프레임의 기사가 압도적으로 많았으며, 이어 (2) 의학적, (3) 정책, (4) 상업적 프레임이 나타났다.

세 데이터셋의 가장 큰 차이는 갈등 프레임에서 드러난다. 병명 개정 전에는 전체 기사에서 가장 적은 비중을 차지했던 범죄 기사가, 병명 개정 후 그 비중이 약 6배, 10배로 증가하는 것을 볼 수 있다. 이는 병명 개정 후 얼마 지나지 않아 조현병 환자의 강력 범죄 뉴스가 잇따

라 보도되면서 질환에 대한 경각심이 높아졌기 때문으로 보인다. 반면 의학적 프레임의 기사는 병명 개정 후 비중이 절반 정도로 감소하였다. 이는 병명 개정 전에는 정신분열병의 증상, 연구, 치료 등 의학적 지식과 정보가 지속적인 사회적 이슈였지만 병명 개정 후에는 이에 대한 관심이 줄어들었음을 보여준다. 한편, 병명 개정 후 정신분열병 기사의 프레임은 비교적 고른 분포를 보이는데, 이는 병명 개정 후에 정신분열병이 한 프레임에 국한되지 않고 다양한 관점에서 다뤄졌음을 보여준다.

#### 4.2 TF-IDF와 동시출현 단어 분석

〈표 5〉는 데이터셋별 TF-IDF 상위 30개의 단어와 그 값이다. 단어의 TF-IDF 가중치는 클

〈표 5〉 상위 30개 단어와 TF-IDF 값

병명 개정 전 정신분열병		병명 개정 후 정신분열병		조현병	
TF-IDF (1-15위)	TF-IDF (16-30위)	TF-IDF (1-15위)	TF-IDF (16-30위)	TF-IDF (1-15위)	TF-IDF (16-30위)
노홍철	마리화나	데드폴	피해망상	소녀상	새싹
0.4769	0.2777	0.513425	0.277293	0.633009	0.291425
우남	포스텍터	아만다 바인즈	이중섭	간지럼	흡연
0.4232	0.2722	0.442483	0.273954	0.387069	0.289572
스스무	피아노	피해망상증	성금	언행	전동차
0.3683	0.2707	0.433566	0.27358	0.377338	0.287731
유치원	기생충	간지럼	마리화나	협오범죄	포항
0.3230	0.2704	0.364372	0.273373	0.368862	0.281024
다빈치코드	금연	하품	손가락	아리피프라졸	피로
0.3146	0.2702	0.354045	0.271121	0.364197	0.268558
타우	로버트 고든	조퇴	제임스 조이스	젤다	비타민 D
0.3120	0.2653	0.339866	0.270966	0.352164	0.262602
박태원	선교사	아바타	협합	암페타민	니코 젠킨스
0.3111	0.2625	0.33653	0.270372	0.343396	0.260141
자간전증	베텔	저장강박증	차오	부산연계경찰서	당사자
0.3083	0.2611	0.306608	0.268391	0.331088	0.260094
존 듀퐁	쿠바	강박장애	지각	원숭이	혼잣말
0.3062	0.2578	0.304259	0.266535	0.328369	0.259368
임플란트	대마초	비타민 D	베개	리처드 로하스	관찰
0.2785	0.2578	0.304246	0.265319	0.324285	0.25786
동성애	대마초	법정 스님	한지	저스틴 설리번	노숙인
0.3012	0.2574	0.289401	0.26405	0.323981	0.25645
오메가-3	전자발찌	질그릇	강남역	경력	유전자
0.3008	0.2566	0.289401	0.256934	0.323221	0.25371
무용치료	링컨	왼손잡이	작업기억	부산 지하철	무함마드
0.2890	0.2564	0.283144	0.256636	0.303731	0.253458
모기지	혼인	대응계약	자책	그립자	소주병
0.2883	0.2560	0.280605	0.253122	0.302932	0.251559
천재	찰리	본드	편집증	차량	동화약품
0.2869	0.2527	0.278105	0.252463	0.298498	0.25075336

수록 단어가 포함된 문서 내에서 핵심적인 메시지를 담고 있다는 의미이며, 값이 낮을수록 많은 문서에서 다루고 있다는 의미이다. 보다 정확한 문맥 파악을 위해 〈표 6〉에서는 상위 30개 단어를 포함한 기사의 본문 내용을 확인 후 뉴

스 프레임을 분석하였다. 그 결과는 〈표 6〉에 제시하였다. 〈표 5〉와 〈표 6〉을 토대로 분석한 내용은 다음과 같다.

먼저 병명 개정 전 정신분열병 기사에서는 ‘노홍철, 우남, 스스무, 유치원, 다빈치코드’ 등

〈표 6〉 TF-IDF 값 상위 30개 단어가 출현한 뉴스 기사의 프레임

	의학적	인간적 흥미	갈등	상업적	정책	합계
정신분열병	12	10	4	2	2	30
정신분열병	19	5	2	3	1	30
조현병	11	6	13	·	·	30

다양한 주제의 단어가 높은 TF-IDF 값을 가졌다. 이는 당시 대중이 정신분열병의 다방면에 관심을 가졌음을 보여준다. 가중치가 상대적으로 낮게 나온 ‘피터팬(0.013)’, ‘리코펜(0.014)’, ‘대사증후군(0.014)’은 이 시기 대부분 기사가 의학적 키워드를 포함했음을 보여준다.

병명 개정 후의 정신분열병 기사에서는 상위 30개 단어의 절반 이상이 의학적 프레임과 관계된 키워드로 나타났다. 이를 통해 정신분열병의 증상 및 치료제가 자주 화제가 되었음을 알 수 있다. 한편, ‘산업(0.010)’, ‘시장(0.011)’, ‘인수(0.012)’의 가중치가 낮게 나온 점으로 보아 병명 개정 후 정신분열병 관련 뉴스에서는 제약회사의 인수합병과 관련된 상업적 내용이 기사에서 자주 언급되었음을 알 수 있다. 인간적 흥미 및 사회적 프레임은 이 데이터셋만이 가지는 특이점으로 나타났다. 병명 개정 전의 기사나 조현병 기사에서도 질환 관련 일화나 사회문제가 언급되긴 했지만, 이들은 부가적 정보였을 뿐 주제로 다루지진 않았기에 나타난 결과로 볼 수 있다.

정신분열병 기사와 달리, 조현병 기사의 상위 30개 단어에는 상업적 프레임, 정책 프레임과 관련된 키워드가 나타나지 않았다. 이는 정책이나 상업적 내용보다는 사건, 사고나 의학적 정보가 조현병과 관련된 주요 이슈였음을 암시한다. 반면 가중치가 가장 낮은 ‘경찰(0.013)’, ‘환

자(0.013)’, ‘치료(0.014)’와 같은 단어는 다수의 기사에서 다루어졌다고 볼 수 있는데, 이는 병명 개정 후 조현병 관련 사건이나 의학적 정보에 대한 관심을 보여준다.

요약하자면 병명 개정 전 정신분열병 기사에서는 의학적 키워드가 대부분의 기사에서 다루졌으며, 사회적 이슈가 비교적 다양한 주제에 걸쳐 나타났다. 병명 개정 후의 정신분열병 기사에서는 제약 시장 관련 키워드가 자주 언급은 되었지만, 의학 정보가 더 중요한 이슈였다. 마지막으로 조현병 기사에서는 전반적으로 등장하는 키워드도, 사회적 이슈도 모두 갈등 또는 의학적 키워드였다.

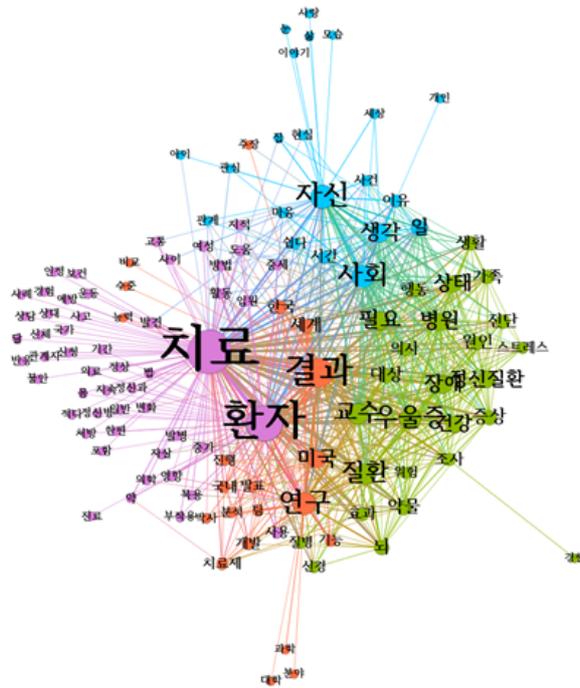
〈그림 4〉, 〈그림 5〉, 〈그림 6〉은 각 데이터셋에 동시출현한 단어를 노드(Node)로, 단어의 동시출현 쌍을 엣지(Edge)로 설정하여 그린 동시출현 네트워크이다. 시각화에는 Gephi(Bastian et al, 2009)를 사용하였다. 노드의 크기는 연결 중심성(Degree centrality)을 나타내며, 연결 중심성이 높을수록 크게 표현되도록 하였다. 연결 중심성이란, 네트워크를 구성하고 있는 하나의 노드가 네트워크 상의 다른 노드들과 연결되는 정도를 나타내는 값으로, 그 값이 높을수록 다른 노드들과 많이 연결되어 네트워크에서 높은 영향력을 가지는 것으로 해석할 수 있다(안주영, 안규빈, 송민 2016). 각 네트워크의 평균 연결 중심성은 〈표 7〉에, 연결 중심성이 가장 높

은 단어 20개는 <표 8>에 제시하였다. 노드의 색깔은 네트워크에서 커뮤니티 구조를 감지할 때 주로 쓰이는 기법인 모듈성(Modularity)의 계산 결과에 따라 관련성이 있다고 분석된 단어끼리 동일한 색으로 표시하였다. 효과적인 시각화를 위해 동시출현 단어 250회 이하, 모듈성 0.1% 미만인 커뮤니티는 네트워크에서 제외하였다.

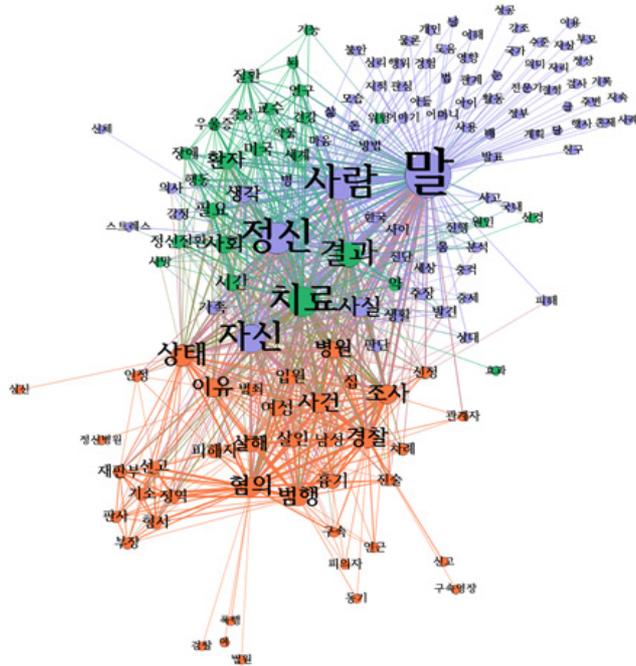
<그림 4>의 병명 개정 전 정신분열병 기사의 동시출현 네트워크는 총 130개의 노드와 753개의 엣지로 구성되었으며, 모듈리티 알고리즘 수행 결과 4개의 커뮤니티가 감지되었다. 먼저, '치료'와 '환자'를 중심으로 이루어진 좌측 커뮤니티의 경우, 정신분열병 환자의 치료와 처방에 관한 단어들로 구성되어 있음을 확인할 수

있다. '결과, 연구'를 중심으로 형성된 중앙의 커뮤니티는 치료제 개발 연구를 포함한 국내외 연구에 관한 단어들이 다수 출현한다. 또, 우측의 커뮤니티는 '우울증, 질환, 장애' 등의 증상 관련 단어가, 상단의 커뮤니티는 '자신, 사회, 이야기' 등 개인과 공동체의 실태 및 예술에 관한 단어가 주를 이룬다.

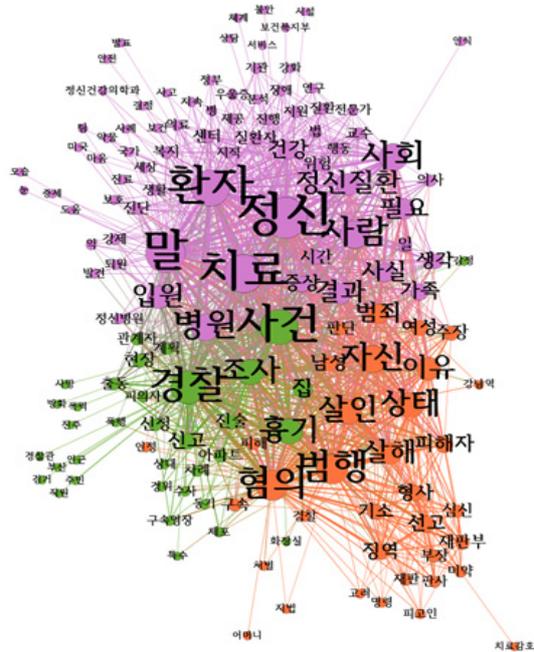
<그림 5>의 병명 개정 후 정신분열병 기사의 동시출현 네트워크는 총 169개의 노드와 950개의 엣지로 구성되었으며, 노드들은 크게 3개의 커뮤니티로 나뉘었다. 이 네트워크에서 주목할 점은 하단의 범죄 사건 관련 단어들로 구성된 커뮤니티이다. <그림 4>에선 찾아볼 수 없었던 '혐의, 범행'으로 대표되는 새로운 주제의 키워드가 해당 네트워크에서 다소 큰 비중을 차지



<그림 4> 병명 개정 전 정신분열병 관련 뉴스의 동시출현 네트워크



〈그림 5〉 병명 개정 후 정신분열병 관련 뉴스의 동시출현 네트워크



〈그림 6〉 조현병 관련 뉴스의 동시출현 네트워크

함으로 미루어 보아 개정 후 정신분열병 환자의 범행이 새로이 혹은 더욱 빈번히 보도되었음을 파악할 수 있다. 그 외 두 커뮤니티는 전자와 유사한 단어 분포를 보인다. 상단의 '말'과 '사람'을 중심으로 이루어진 커뮤니티에는 정신분열병의 증상을 나타내는 단어뿐만 아니라, 진단, 치료 및 예술 등 매우 다양한 주제와 연관된 단어들도 고르게 분포하고 있다. 또, 좌측 상단의 커뮤니티는 증상과 연구에 관련된 단어들로 구성되어 있다.

〈그림 6〉의 조현병 기사의 동시출현 네트워크는 총 162개의 노드와 1,494개의 엣지로 구성되었으며, 노드들은 3개의 커뮤니티로 나뉘었다. 이때 하단의 두 커뮤니티는 공통적으로 범죄 사건에 관한 키워드로 구성된다. 이에 둘을 범죄라는 하나의 주제의 커뮤니티로 본다면 그 비중이 전체 네트워크의 절반 정도를 차지함을 확인할 수 있다. 이는 정신분열병보다 조현병 관련 뉴스에서 사건, 사고가 상대적으로 많이 다뤄졌음을 보여준다. 반면, 나머지 반을 차지하는 상단의 클러스터는 환자의 치료 및 복지에 관련된 키워드들로 구성된다. '보건복지부, 센터' 등 복지와 직결된 단어는 조현병 뉴스 데이터만의 특징으로 볼 수 있다.

이에 정신질환 기사의 뉴스 프레임을 적용하자면 2011년 이전에 보도된 정신분열병 뉴스엔 의학적, 사회적, 상업적 키워드가, 2011년 이후 보도된 정신분열병 관련 뉴스엔 의학적, 갈등

외 여러 주제의 키워드가, 조현병 관련 뉴스에는 갈등, 의학적, 정책 키워드가 나타났음을 알 수 있다.

한편, 〈표 7〉에 따르면 세 네트워크 중 조현병 관련 네트워크의 평균 연결 중심성이 가장 높다. 이는 조현병 관련 기사에 등장한 단어들의 동시출현 단어가 해당 네트워크 상에서 영향력이 높음을 보여준다. 그 세부 내용은 〈표 8〉에서 확인할 수 있다.

〈표 8〉의 연결 중심성 상위 단어의 차이는 각 데이터셋에서 중요하게 다루고 있는 토픽의 차이를 보여준다. 가장 큰 차이는 병명 개정을 기점으로 '범행, 조사, 혐의' 등 범죄 사건에 관한 단어의 중요성이 급증한다는 점이다. 치료, 연구, 증상에 관한 의학적 용어가 핵심인 병명 개정 전 데이터에서는 범죄 키워드를 전혀 찾아볼 수 없었는데 개정 후 이들이 급부상한 점으로 보아, 병명 개정 전보다 후에 정신분열병 및 조현병 환자의 범위가 훨씬 중요하게 다뤄졌으며 사회적으로 관심을 받았다고 할 수 있다. 또한, 병명 개정 후 정신분열병 기사보다 조현병 기사에서 범죄 관련 키워드의 연결 중심성이 3배 정도 높게 나타났는데, 이는 병명 개정 후, 정신분열병보다 조현병 기사에서 환자의 범행이 더욱 부각 되었음을 보여준다. 하지만 범죄 키워드가 유사하며 두 데이터셋에서 차지하는 비중이 비슷함을 통해, 보도의 주 내용은 범죄 사건과 경찰 조사에서 크게 벗어나

〈표 7〉 평균 연결 중심성

병명 개정 전 정신분열병	병명 개정 후 정신분열병	조현병
11,585	11,243	18,444

〈표 8〉 연결 중심성 상위 20개 단어

병명 개정 전 정신분열병		병명 개정 후 정신분열병		조현병	
상위단어	연결 중심성	상위단어	연결 중심성	상위단어	연결 중심성
치료	253.04	말	203.25	정신	395.98
환자	177.60	정신	190.89	치료	256.02
연구	151.20	사람	176.03	환자	249.83
결과	149.49	치료	162.47	경찰	225.39
자신	147.45	자신	160.12	사건	219.90
질환	126.57	결과	136.37	말	208.13
사회	87.63	상태	97.32	협의	202.08
교수	78.79	사건	83.93	범행	190.75
우울증	77.82	범행	78.59	사람	165.89
미국	73.96	환자	77.99	병원	162.02
장애	68.46	이유	76.92	조사	138.93
병원	59.12	사회	74.83	살인	135.99
생각	56.52	사실	66.58	흥기	132.37
정신질환	50.34	조사	60.13	자신	132.08
건강	49.65	협의	55.24	사회	131.54
상태	49.27	병원	49.08	상태	129.28
필요	45.57	경찰	47.27	살해	113.01
일	41.23	살해	45.69	정신질환	102.17
진단	40.88	필요	44.14	이유	99.28
증상	30.74	여성	36.80	필요	96.49

지 않는다고 해석할 수 있다. 반면 병명 개정 전후 모두에서 ‘정신’과 ‘치료’가 상위에 위치함으로써 보아, 병의 치료는 병명과 관계없이 중요한 문제로 취급되었다고 해석할 수 있다.

### 4.3 감성분석

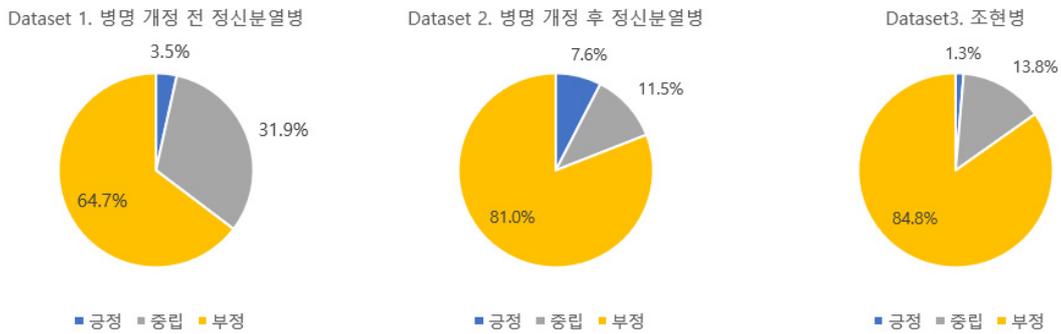
주제 맞춤형 감성 사전을 구축하여 감성분석을 한 결과는 〈표 9〉, 〈그림 7〉, 〈표 10〉과 같다. 먼저, 〈표 9〉는 병명 개정 전, 후의 정신분열병 기사와 조현병 기사의 평균 극성값이다. 극성값은 중립값 0을 기준으로 1에 가까울수록 강한 긍정을, -1에 가까울수록 강한 부정을 나타

낸다. 세 데이터셋은 모두 극성값이 음수이므로, 언론에서는 정신분열병/조현병에 대해 긍정적인 시각보다는 부정적인 시각이 더 많다고 볼 수 있다. 한편, 평균 극성값이 가장 낮은 조현병 기사가 정신분열병 기사보다 병의 위험성이나 관련 범죄 등 부정적인 측면을 더 부각했다고 볼 수 있는데, 이는 〈그림 7〉에서 확인할 수 있다.

〈그림 7〉은 병명 개정 후 정신분열병과 조현병 기사에서의 긍정, 중립, 부정적 기사 건수와 비율이다. 이때 (극성값) >= 0.01이면 ‘긍정’, -0.01 < (극성값) < 0.01이면 ‘중립’, (극성값) < -0.01면 ‘부정’으로 하여 기사를 분류하였

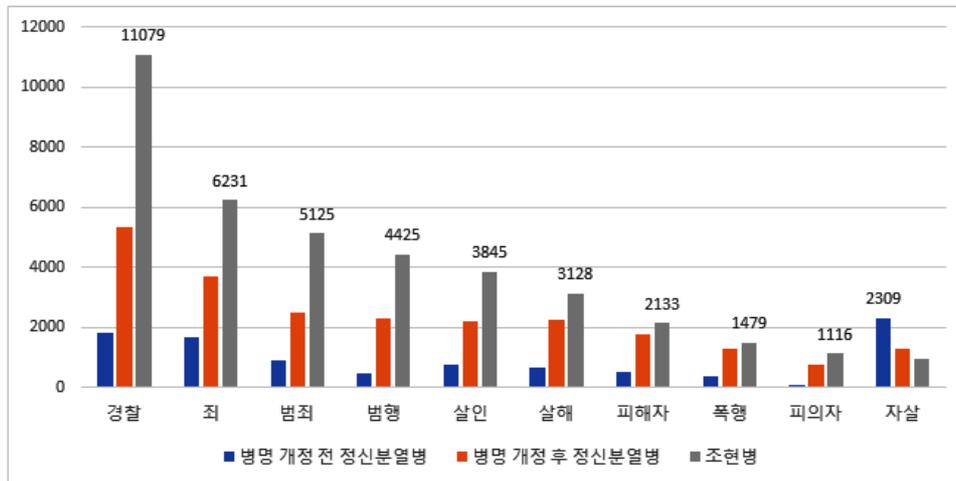
〈표 9〉 병명 개정 전·후의 평균 극성값

	병명 개정 전 정신분열병	병명 개정 후 정신분열병	조현병
평균 극성값	-0.02197	-0.03864	-0.04511



〈그림 7〉 병명 개정 전·후 긍정/중립/부정적 기사의 비율

〈표 10〉 병명 개정 전·후의 부정적 단어 출현 횟수



다. 주관성을 최대한 배제해야 하는 뉴스 기사의 텍스트의 특성상 극성값이 낮을 수밖에 없다는 점을 고려하여, 텍스트의 긍정/부정성이 비교적 뚜렷이 나타나기 시작하는 기사의 극성값인 0.01을 경계값으로 하였다. 〈그림 7〉에

따르면 세 데이터셋 중 조현병 데이터셋에서 긍정적 기사의 비중이 가장 낮고, 부정적 기사의 비중은 가장 크다. 즉 언론은 조현병과 관련하여 질환자의 회복 가능성, 성공 스토리와 같은 긍정적인 내용은 거의 보도하지 않고,

그 위험성만 더욱 강조하여 보도함으로써 질환자에 대한 사회적 거부감을 심화했다고 볼 수 있다.

반면, 병명 개정 전 정신분열병 데이터셋에서는 중립적 기사가 비교적 높은 비중을 차지한다. 이를 통해 개정 후보다 전에 사회적 편견을 조장하지 않는 범위 내에서 사실을 객관적으로 전달하고자 하는 경향이 더 컸음을 확인할 수 있다. 또, 병명 개정 후 정신분열병 데이터셋의 긍정 기사 비율이 가장 높은 것으로 보아 개정 후 정신분열병의 긍정적 측면을 기사에 반영하려는 시도가 이루어지고 있음을 알 수 있다. 긍정적 기사의 대부분은 예술작품에 관한 보도인데, 이는 병명 개정에도 불구하고 예술 분야에서는 여전히 해당 질환이 조현병보다 정신분열병으로 더 자주 일컬어짐을 보여준다.

〈표 10〉은 병명 개정 후 정신분열병과 조현병 기사 전체에서 출현 횟수 차이가 가장 큰 부정적 단어 상위 10개이다. 범죄와 관련된 키워드인 '경찰, 죄, 범죄' 등은 병명 개정 전 정신분열병 기사에서 최저, 조현병 기사에서 최고 빈도로 나타난 반면, '자살'은 상반된 양상을 보였다. 이는 병명 개정 후, 특히 2016년 5월에 발생한 강남역 살인 사건 이후로 정신분열병/조현병 질환자의 강력 범죄가 연이어 보도되면서 세간의 관심이 병의 증상에서 질환자의 범행으로 옮겨갔기 때문으로 볼 수 있다. 이처럼 개인의 죽음보다 대중에게 더 위협적으로 느껴질 수 있는 범죄와 관련 키워드가 더욱 빈번하게 등장함은, 병명 개정 후 정신분열병/조현병 질환자는 잠재적 범죄자일 것이라는 사회적 낙인 형성에 크게 영향을 미쳤을 것으로 보인다.

## 5. 결론

조현병 환자의 낙인은 환자의 정신건강 회복을 저해하고 사회 복귀를 방해하여 고립된 삶을 심화하는 등 여러 가지 부정적 결과를 초래한다. 이에 본 연구는 조현병 환자의 낙인에 영향을 미치는 온라인 뉴스 기사를 분석하여 낙인 현상을 극복하기 위해 언론이 나아가야 할 바람직한 방향을 제시하는 기초자료를 제공하려고 하였다. 이를 위해 2011년 병명 개정을 기준으로 전, 후 약 10년간 네이버 뉴스에서 제공한 정신분열병/조현병 관련 기사 12,290건을 병명 개정 전, 후로 나누어 내용 분석을 하였다. 연구의 주요 결과는 다음과 같다.

먼저, 병명 개정 전 정신분열병 기사의 주요 프레임은 의학적 프레임으로 나타났다. 6가지 프레임 중 의학적 프레임의 기사가 58.2%로 가장 많이 보도되었으며, 키워드의 출현 빈도 및 중요도 역시 치료, 연구, 증상과 관련된 의학용어가 가장 높았다. 하지만 이러한 주요 흐름 속에서 다양한 프레임의 키워드가 이슈화되곤 하였다. 또한, 부정적인 뉘앙스가 전체의 64%로 우세하였지만, 중립 감성으로 판명되는 기사 역시 31.9% 정도로 적잖이 나타났다.

병명 개정 후 정신분열병 기사의 주요 프레임은 전체의 약 30%씩을 차지하는 갈등 프레임과 의학적 프레임으로 나타났다. 정신분열병의 증상, 관련 물질, 치료법, 제약회사 등 다양한 의학 정보가 많은 기사에서 중점적으로 다뤄졌다. 강남역 살인사건과 같은 범죄 콘텐츠도 다뤄지긴 했으나 그 중요성은 여전히 의학 정보에 미치지 못했다. 또한, 병명 개정 전보다 부정적 뉘앙스는 심화 되었지만 동시에 환자의 성공 스토리,

회복 사례 등 정신분열병의 긍정적 측면을 보도하는 기사가 증가하여 병의 양면이 언론에 비춰졌다.

마지막으로 병명 개정 후 조현병 기사의 주요 프레임은 갈등 프레임으로 나타났다. 갈등 프레임은 전체의 52%로 가장 많이 보도되었을 뿐 아니라, 그 연관어의 중요도 또한 다른 데이터셋보다 최소 3배 이상 높게 나타났다. 갈등 프레임 기사의 주 내용은 질환자의 범행과 경찰 조사였다. 의학적 키워드 역시 높은 빈도로 출현하긴 했지만, 기사의 주제보단 부가적 정보로 언급된 횟수가 더 많았다. 한편, 조현병 기사에서 부정적 뉘앙스가 가장 많이 드러났으며 긍정적 뉘앙스는 1% 정도로 거의 나타나지 않았다.

이런 점에서 연구문제에 대한 결론은 다음과 같다.

- RQ1. 전체적으로 병명 개정 이후 정신분열병과 조현병에 갈등 프레임과 범죄 관련 키워드의 비중은 급증한 반면, 의학적 프레임의 기사 비중은 급감하였다. 또, 기사의 부정적인 주관이 기사에 반영되는 경향이 증가했다. 이는 대중이 정확한 의학 지식은 알려고 하지 않은 채 정신분열병/조현병 환자는 위험하다는 편견만 더 가지게 되었음을 보여준다.
- RQ2. 부정적 기사의 비율, 기사 내 부정적 키워드의 중요도 및 다양성, 기사에 드러난 감성을 종합해 보았을 때 병명 개정 후 정신분열병보다 조현병에 대한 사회적 인식이 더 부정적임을 알 수 있었다. 이러한 인식에는 갈등, 의학적 프레임의 증감뿐 아니라 병명 개정과 관계없

이 계속해서 해당 질병을 정신분열병으로 일컫는 문화예술 보도의 관습도 기여한 것으로 나타났다.

결론적으로, 정신분열병에서 조현병으로의 병명 개정이 사회적 인식을 개선하였다고 보기는 어렵다. 사회적 인식과 동일시되는 언론 보도에서 정확한 정보를 전달하고 편견 해소에 기여하는 의학적 프레임은 감소하고, 불안과 편견을 조장하는 갈등 프레임만 증가하였기 때문이다. 이처럼 낙인 불식이라는 긍정적 파급효과를 기대하고 병명 개정을 추진하였지만 연이은 강력 범죄 보도로 인해 조현병에 대한 사회적 인식은 오히려 나빠졌다. 하지만 단순히 언론 보도 때문에 부정적 인식이 증가했다고 보기는 어렵다. 사실 정신질환자에 대한 편견과 차별은 오래전부터 동서양을 불문하고 존재하였으며(보건복지부 2007; Angermeyer and Matschinger 2005; Link et al. 1999) 정신질환에 대한 개인의 인식은 인구통계학적 특성을 포함한 다양한 요인에 의해 영향을 받기 때문이다. 그럼에도 불구하고 본 연구는 인식 개선의 방안으로 '바람직한 보도 방향의 정립'을 제안하고자 한다. 서구에서는 편견 해소를 위해 1990년대부터 대중매체를 활용하였으며 효과성이 입증된 여러 사례가 존재함으로 보아 그 파급력이 작지 않음을 알 수 있기 때문이다.

정신분열병과 달리 병명 자체의 부정적 어감은 없는 조현병이지만, 해당 질병이 계속해서 범죄 등의 부정적인 정보와 연관되어 보도된다면 사회적 인식은 병명과 관계없이 절대로 호전되지 않을 것이다. 따라서 인식개선을 위해서는 조직적, 개인적 차원의 노력이 필요하다. 먼저

조직적 차원에서 국내 언론사는 '정신질환자에 관한 보도 가이드라인'을 마련하여 무분별한 보도와 그에 따른 편견 조장을 방지해야 할 것이다. 개인적 차원으로는 정신분열병/조현병 관련 기사를 보도하는 언론인의 보도 지침 준수와 사회적 편견을 조장하지 않는 객관적인 기사 작성이 필요할 것이다. 이와 같은 다차원의 노력이 병행된다면 질환자에 대한 사회적 낙인이 해소될 수 있을 것으로 기대된다.

본 연구는 국내 최초로 텍스트 마이닝 기법을 활용하여 지상파 군과 종편 군을 모두 포함한 장기간의 정신분열병/조현병 기사 내용을 분석했다는 점에서 의의가 있다. 또한, 다양한 텍스트 마이닝 기법에 사전에 정의된 뉴스 프레임 유형을 접목하여 대량의 문헌 데이터를 구체

적 기준에 따라 통계적, 시간적으로 분석하고, 전통적인 정성적 내용 분석에서 제기되었던 객관성 확보의 한계를 보완했다는 점에서 의의가 있다. 본 연구의 한계점으로는 형태소 분석기가 고유명사와 합성명사를 모두 반영하지 못해 일부 의학 전문용어나 신조어가 분석 결과에서 빠졌을 가능성이 있으며, LDA 토픽 모델링의 결과로 도출된 주제 라벨링 시 연구자의 주관이 불가피하게 반영된다는 점을 들 수 있다. 향후 연구에서 객체명 인식기(NER)를 사용하여 명사를 추출하고, 주제 라벨링에 분야 전문가의 도움을 받는다면 보다 정확한 정신분열병/조현병에 대한 사회적 인식의 파악이 가능할 것으로 판단된다.

## 참 고 문 헌

- [1] 감미아, 송민. 2012. 텍스트 마이닝을 활용한 신문에 따른 내용 및 논조 차이점 분석. 『지능정보연구』, 18(3): 53-77.
- [2] 김민철. 2017. 스포츠유명인의 부정적 언론 보도가 스폰서의 기업이미지, 브랜드태도에 미치는 영향력 및 광고 적합성의 완충효과. 『한국스포츠산업경영학회지』, 22(4): 39-58.
- [3] 김혜선, 박도원, 홍영은. 2018. 정신장애 범죄에 대한 언론보도 경향과 범죄위험성 인식. 『장애의 재해석』, 208-250.
- [4] 노수진, 윤영민. 2013. 우울증에 관한 언론 보도 분석: 온라인 뉴스 미디어를 중심으로. 『한국언론정보학보』, 61: 5-27.
- [5] 대한조현병학회. 2011. 『(정신분열병) 병명 개정 백서: 조현병-정신분열병의 새로운 이름』. 서울: 대한조현병학회.
- [6] 반현. 2005. 제3주제: 독자들이 바라보는 신문위기: 질적연구. 『한국언론학회 심포지움 및 세미나』, 51-86.
- [7] 배영. 2014. 문화 콘텐츠의 유형별 소비패턴과 소셜미디어의 역할. 『정보와 사회』, 15(2): 33-65.

- [8] 백혜진, 조혜진, 김정현. 2017. 정신질환의 낙인과 귀인에 대한 언론 보도 분석. 『한국언론학보』, 61(4): 7-43.
- [9] 서울대학교 의과대학. 2007. 『2007년도 정신질환실태 역학조사』. 세종: 보건복지부, 서울대학교 의과대학.
- [10] 삼성서울병원. 2016. 『2016년도 정신질환실태 역학조사』. 세종: 보건복지부, 삼성서울병원.
- [11] 서울대학교병원. 2010. N의학정보 조현병. [online] [cited 2019. 8. 22.]  
<<http://www.snuh.org/health/nMedInfo/nView.do?category=DIS&medid=AA000356>>.
- [12] 송민. 2017. 『텍스트 마이닝』. 서울: 청람.
- [13] 안주영, 안규빈, 송민. 2016. 텍스트 마이닝을 이용한 매체별 에볼라 주제 분석: 바이오 분야 연구 논문과 뉴스 텍스트 데이터를 이용하여. 『한국문헌정보학회지』, 50(2): 289-307.
- [14] 이민철, 김혜진. 2018. 텍스트 마이닝 기법을 적용한 뉴스 데이터에서의 사건 네트워크 구축. 『지능정보연구』, 24(1): 183-203.
- [15] 이성직, 김한준. 2009. TF-IDF의 변형을 이용한 자뉴스에서의 키워드 추출기법. 『한국자거래학회보』, 14(4): 59-73.
- [16] 이준웅. 2005. 갈등적 사안에 대한 여론 변화를 설명하기 위한 프레임 모형 검증 연구: 정부의 통일 정책에 대한 뉴스 프레임의 형성과 해석적 프레임의 구성을 중심으로. 『한국언론학보』, 49(1): 133-162.
- [17] 정의철. 2011. 유방암 관련 언론보도의 프레임 특성에 대한 연구. 『언론과학연구』, 11(2): 402-440.
- [18] 조수영, 김정민. 2010. 정신건강 및 정신질환에 대한 지상과 TV 뉴스 분석. 『한국언론학보』, 54(5): 181-204.
- [19] 주성빈. (2019). 언론의 범죄 보도 행태에 관한 감성분석. 『한국범죄심리연구』, 15(2): 127-140.
- [20] 최영해. 2019. 『대통령의 중점정책 변동 과정에서 언론과 여론의 영향 관계 분석』. 박사학위논문, 성균관대학교 국정전문대학원 행정학과.
- [21] 한국. 대검찰청. 2018. 『2018 범죄분석』. 서울: 대검찰청.
- [22] 한국언론진흥재단. 2018. 『뉴스 이용집중도 조사보고서 2016-2018』. 여론집중도조사위원회.
- [23] 황애리, 나은영. 2018. 한국 언론의 정신건강 보도에 관한 내용 분석 연구: 뉴스 프레임과 기사 논조를 중심으로. 『보건사회연구』, 38(4): 290-322.
- [24] Angermeyer, M. and Matschinger, H. 2005. "The stigma of mental illness in Germany: a trend analysis." *Int J Soc Psychiatry*, 51: 276-284.
- [25] Baccianella, S., Esuli, A. and Sebastiani, F. 2010. "Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining." *Lrec*, 10(2010): 2200-2204.
- [26] Bastian, M., Heymann, S. and Jacomy, M. 2009. "Gephi: an open source software for

- exploring and manipulating networks.” In Proceedings of the Third international AAAI conference on weblogs and social media.
- [27] Blei, D. M., Ng, A. Y. and Jordan, M. I. 2003. “Latent dirichlet allocation.” *Journal of machine Learning research*, 3(Jan): 993-1022.
- [28] Corrigan, P. W., Powell, K. J. and Michaels, P. J. 2013. “The effects of news stories on the stigma of mental illness.” *The Journal of nervous and mental disease*, 201(3): 179-182.
- [29] Dearing, J. and Everett, R. 1996. Agenda-Setting. Communication Concepts. Vol 6: Thousand Oaks, CA: Sage Publications.
- [30] Entman, R. M. 1991. “Framing US coverage of international news: Contrasts in narratives of the KAL and Iran Air incidents.” *Journal of communication*, 41(4): 6-27.
- [31] Esuli, A. and Sebastiani, F. 2007. “SentiWordNet: a high-coverage lexical resource for opinion mining.” *Evaluation*, 17(1): 26.
- [32] Goffman, E. 1974. Frame analysis: An essay on the organization of experience. Harvard University Press.
- [33] Goulden, R., et al. 2011. “Newspaper coverage of mental illness in the UK, 1992-2008.” *BMC public health*, 11(1): 796.
- [34] Guo, L., et al. 2016. “Big social data analytics in journalism and mass communication: Comparing dictionary-based text analysis and unsupervised topic modeling.” *Journalism & Mass Communication Quarterly*, 93(2): 332-359.
- [35] Link BG, et al. 1999. “Public conceptions of mental illness: labels, causes, dangerousness, and social distance.” *Am J Public Health*, 89: 1328-1333.
- [36] Mollema, L., et al. 2015. “Disease detection or public opinion reflection? Content analysis of tweets, other social media, and online newspapers during the measles outbreak in The Netherlands in 2013.” *Journal of medical Internet research*, 17(5): e128.
- [37] Jorm, A. F. 2000. “Mental health literacy: Public knowledge and beliefs about mental disorders.” *The British Journal of Psychiatry*, 177(5): 396-401.
- [38] Overton, S. L., and Medina, S. L. 2008. “The stigma of mental illness.” *Journal of Counseling and Development*, 86(2): 143.
- [39] Penn, D. L., and Martin, J. 1998. “The stigma of severe mental illness: Some potential solutions for a recalcitrant problem.” *Psychiatric Quarterly*, 69(3): 235-247.
- [40] Rhydderch, D., et al. 2016. “Changes in newspaper coverage of mental illness from 2008 to 2014 in England.” *Acta Psychiatrica Scandinavica*, 134(S446): 45-52.

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

- [1] Kam, M. and Song, M. 2012. "A Study on Differences of Contents and Tones of Arguments among Newspapers Using Text Mining Analysis." *Journal of Intelligence and Information Systems*, 18(3): 53-77.
- [2] Kim, M. 2017. "The Effects of Sports Celebrities' Negative Media Coverage on Sponsor's Corporate Image and Brand Attitude and the Buffering Effect of Ad Relevance." *Korean Journal of Sport Management*, 22(4): 39-58.
- [3] Kim, H., Park, D. and Hong Y. 2018. "Newspaper report trends for mental disorder crime and perception of crime risk: based on newspaper analysis and survey method." *Reinterpretation of Disability*, 208-250.
- [4] Roh, S. and Yoon, Y. 2013. "Analyzing Online News Media Coverage of Depression." *Korean Journal of Communication & Information*, 61: 5-27.
- [5] Korean Society for Schizophrenia Research, 2011. (Schizophrenia) Disease Name Revision Report: New Name for Schizophrenia-Schizophrenia.
- [6] Bahn, H. (2005). Topic 3: Newspaper Crisis Perceived by Readers. In Proceedings of the Symposiums and Seminars of the Korean Society For Journalism And Communication Studies, 51-86.
- [7] Bae, Y. 2014. "Consuming patterns by type of cultural content and the role of social media." *The Korean Journal of the Information Society*, 15(2): 33-65.
- [8] Paek, H., Cho, H. and Kim, J. 2017. "Content Analysis of News Coverage on Stigma and Attribution Regarding Mental Illness." *Korean Journal of Journalism & Communication Studies*, 61(4): 7-43.
- [9] Seoul National University College of Medicine. 2007. *The National Mental Illness Report in 2007*. Seoul: Seoul National University College of Medicine.
- [10] Samsung Medical Center. 2016. *The National Mental Illness Report in 2016*. Seoul: Samsung Medical Center.
- [11] Seoul National University Hospital, 2010. NMedical Information Schizophrenia. [online] [cited 2019. 8. 22.]  
<<http://www.snuh.org/health/nMedInfo/nView.do?category=DIS&medid=AA000356>>.
- [12] Song, M. 2017. *Text Mining*. Seoul: Chungram.
- [13] An, J., Ahn, K. and Song, M. 2016. "Text Mining Driven Content Analysis of Ebola on

- News Media and Scientific Publications.” *Journal of the Korean Society for Library and Information Science*, 50(2): 289-307.
- [14] Lee, M. and Kim, H. 2018. “Construction of Event Networks from Large News Data Using Text Mining Techniques.” *Journal of Intelligence and Information Systems*, 24(1): 183-203.
- [15] Lee, S. and Kim, H. 2009. “Keyword Extraction from News Corpus using Modified TF-IDF.” *The Journal of Society for e-Business Studies*, 14(4): 59-73.
- [16] Rhee, J. 2005. “Testing the Framing Model of Public Opinion Process: News Frames, Interpretive Frames, and Framing Effects on Opinion.” *Korean Journal of Journalism & Communication Studies*, 49(1): 133-162.
- [17] Jung, E. 2011. “Media’s breast cancer news framing and its implications.” *Journal of Communication Science*, 11(2): 402-440.
- [18] Cho, S. and Kim, J. 2010. “Network TV News Coverage of Mental Illness.” *Korean Journal of Journalism & Communication Studies*, 54(5): 181-204.
- [19] Joo, S. 2019. “Sentimental Analysis of Crime News Data - Focused on the Comparison before and after Regulation of Media Report.” *Korean Criminal Psychology Review*, 15(2): 127-140.
- [20] Choi, Y. 2019. *Analysis of the relationship between the press and public opinion in the process of the presidential projects change: Focused on the administrative capital and 4 major rivers projects*. Ph.D. diss. Graduate School of Governance, Sungkyunkwan University, Republic of Korea.
- [21] Republic of Korea. Prosecution Service, 2018. Analytical Statistics on Crime.
- [22] Republic of Korea. Ministry of Culture, Sports and Tourism, 2018. News Concentration Survey Report.
- [23] Hwang, A. and Na, E. 2018. “Content Analysis of News Coverage with Regard to Mental Health: Focusing on News Frame and Tone.” *Health and Social Welfare Review*, 38(4): 290-322.

