

탐색적 논문

환자단체 보도자료와 미디어 보도기사 사이 주요 의제 차이 분석

텍스트 마이닝 기법을 중심으로

국문초록

국내의 환자 권리 운동은 2010년 한국환자단체연합회라는 환자단체들의 대규모 연합체가 결성되면서 더욱 가속화되었다. 이 과정에서 보도자료 배포를 통한 환자단체의 대언론 활동은 환자와 관련된 다양한 이슈를 미디어를 통해 사회에 알리는 데 중요한 역할을 하였다. 그러나 미디어는 보도자료 중 특정 이슈만을 선별해 의제화한다는 점에서 보도자료의 의도와 보도기사의 결과물은 서로 달라질 수 있다. 이에 본 연구에서는 한국환자단체연합회 창립 이후 12년 동안 배포한 보도자료 262건과 같은 기간 종합지와 공중파 TV 등 총 13개의 미디어에서 한국환자단체연합회와 관련해 보도한 210건의 보도 기사를 분석해 그 차이를 알아보았다.

본 연구에서는 빅데이터의 토픽 모델링(Topic Modeling) 기법을 연구 방법으로 채택하였다. 그 결과 보도자료와 보도기사들로 부터 각각 4개씩의 메인 토픽을 도출해 낼 수 있었으며, 의제의 목적, 태도, 접근 및 기술 방식 등에서 차이가 있음을 확인할 수 있었다. 이는 환자단체의 대언론 커뮤니케이션 활동의 방향성을 제시하고, 환자중심의 의료 환경을 만드는 데 도움을 줄 수 있을 것이다.

주제어: 환자, 환자단체, 의제설정, 빅 데이터, 텍스트 마이닝



김동석 | 서강대학교
신문방송학과 박사수료*



유현재 | 서강대학교
지식융합미디어학부
교수**



강현우 | 엔자임헬스
센터장***



황순욱 | 데이터클래스
대표****



손수지 | 부경대학교
기술경영대학원
석사과정*****

* 김동석은 현재 헬스 커뮤니케이션 회사 엔자임헬스의 대표로 재직 중이며, 서강대학교 신문방송학과에서 박사과정을 수료했다(E-mail: tomatocomm@naver.com, 제1저자).

** 유현재는 현재 서강대학교 지식융합미디어학부 교수로 재직 중이며, 서강헬스커뮤니케이션센터 센터장을 맡고 있다(E-mail: bus89@nate.com, 교신저자).

*** 강현우는 현재 엔자임헬스에서 헬스인사이트센터 센터장으로 재직 중이다(E-mail: hwkang@enzaim.co.kr, 공동저자).

**** 황순욱은 현재 빅데이터 전문회사 데이터클래스 대표로 재직 중이다(E-mail: swhwang81@gmail.com, 공동저자).

***** 손수지는 현재 부경대학교 기술경영대학원 석사과정에 재학 중이다(E-mail: thstnwl@gmail.com, 공동저자).

서론

개인 또는 조직이 해결해야 할 현안 이슈를 관계 공중에게 알리려고 할 때 미디어는 중요한 수단 이 될 수 있다. 특히 미디어에 뉴스 가치가 있는 정보를 언론에 제공하는 활동인 퍼블리시티(Publicity)는 저렴한 비용과 미디어의 신뢰를 활용할 수 있다는 점에서 널리 활용되고 있는 PR 방법 중 하나이다(권순만, 1999; 홍미현, 2004). 특히 미디어의 의제 설정(Agenda Setting) 기능은 퍼블리시티의 주요한 목적 중 하나라고 할 수 있다. 의제 설정은 미디어가 특정 이슈를 중요하게 다루면 대중들 역시 그것을 중요하게 인식하게 된다는 이론이다(McCombs, 2004). 여기서 더 나아가 미디어의 의제설정 기능은 단순히 무엇을 생각하게 하는 단계를 넘어 대중의 현실 인식에까지 영향을 줄 수 있다고 가정하기도 한다(반현 & McCombs, 2007).

국내 환자단체 역시 국내 의료 환경에서 환자들이 처한 열악한 현실을 알리기 위해 미디어를 통한 메시지 전달에 주력해 왔다. 과거 개인이나 작은 온라인 커뮤니티 정도에 머물러 있던 국내의 환자 운동은 2000년대 들어 백혈병환우회를 시작으로 한국환자단체연합회(이하, 환자단체연합회)라는 7개 환자단체가 연합한 조직이 결성되는 과정을 거치면서 환자를 둘러싼 다양한 갈등을 해결하고 공론화할 수 있는 전환기를 맞게 된다(김용, 2014; 김영희, 2019). 이 과정에서 환자단체연합회는 2010년 창립 이래 2022년 현재까지 국내 환자들이 직면한 다양한 이슈와 주장을 담은 보도자료를 미

디어에 배포함으로써 환자 개인이나 조직의 이슈를 미디어를 통해 공공 의제화하기 위해 노력해 왔다(김동석·유현재, 2022).

그러나 미디어는 수많은 보도자료들 중 뉴스 가치가 있는 것을 선별해 보도할 뿐만 아니라, 이를 가공, 변형하곤 한다. 즉 개인이나 단체의 의도가 담긴 보도자료는 미디어의 문지기(Gate Keeping) 기능을 통해 선별 과정을 거쳐 일부만 보도되거나, 미디어의 의제설정과 프레임 기능에 의해 원래 보도자료와는 다른 형태의 보도기사로 대중에게 전달될 수 있는 가능성을 내포하고 있다(김재희, 2009). 이는 환자단체연합회의 퍼블리시티 활동과 미디어의 보도에서도 비슷한 양상을 보였을 가능성이 있다. 환자단체연합회가 환자 개인이나 연관 단체들이 제기하는 다양한 이슈를 보도자료로 작성해 미디어에 보도하기 위한 퍼블리시티 활동을 전개했다고 하더라도, 그 결과물은 미디어에 의해 선택되고 가공되는 과정에서 다르게 해석될 수 있다는 것이다. 따라서 환자단체연합회의 보도자료가 주장하는 의제와 보도 기사를 통해 미디어가 그려낸 의제의 차이를 탐색하는 것은 보도자료를 배포한 주체, 이를 다시 선별해 보도한 미디어, 그리고 미디어를 통해 세상을 바라보는 대중 등 각각의 다른 관점을 이해하는 의미 있는 일이 될 수 있을 것이다.

본 연구에서는 이를 위해서 빅데이터에 기반한 텍스트 마이닝(Text Mining)을 연구 방법으로 채택하였다. 텍스트 마이닝은 대량의 축적된 자료를 활용해 표면적으로 드러나기 어려운 종합적 통찰을

제시할 수 있는 방법론으로 최근 사회과학 연구에서 많이 사용하고 있다. 텍스트 내에 가시적, 암시적 형태로 포함되어 있는 논의 및 표현 내용을 발견할 수 있다는 이점이 있을 뿐만 아니라, 뉴스와 관련된 축적된 빅데이터의 경우 정치, 경제, 사회, 문화 등 사회 전반의 이슈를 담고 있어 사회적 담론의 흐름을 파악하기에 적합하다고 알려져 있다 (DiMaggio et al., 2013; Chakraborty et al., 2014; 안예지·김동일, 2020).

특히 본 연구에서는 텍스트 마이닝의 여러 방법 중 토픽 모델링(Topic Modeling) 기법을 활용하여 환자단체연합회가 배포한 보도자료와 미디어의 보도기사 내 주요 토픽과 구조를 확인하고 그 공통점과 차이점, 그리고 함의를 알아보았다. 토픽 모델링 기법은 방대한 텍스트 자료에서 의미가 있는 토픽(주제)를 추출해 내는 방법으로, 전달하고자 하는 메시지를 텍스트 내에서 분석해 내는데 용이해 커뮤니케이션 분야에도 유용한 방법론으로 평가받고 있다(Blei et al., 2003; Jin & Spence, 2021; 권나현·최윤형, 2022). 국내에서도 관련 연구들이 활발히 진행되고 있는데 예를 들어 토픽 모델링을 활용해 국내 주요 일간지 기사를 분석해 대통령의 이슈 소유권을 분석해 내거나(박승정 외, 2017), 국내 언론 보도 기사에서 학생인권과 관련된 사회적 이슈를 추출하는 연구도 진행된 바 있다(유주영, 2022). 병원의 온라인 리뷰 데이터를 수집해 의료 서비스 만족도를 조사하는(이시환 외, 2017) 등 토픽 모델링을 활용한 연구는 대부분 수 년 또는 수십 년에 걸쳐 배포된 방대한 분량의 보도자료나

오랜 기간 동안 축적된 뉴스, 온라인 리뷰, SNS 등 대규모 데이터를 분석하는데 주로 활용되고 있다 (김은혜·박주호, 2022; 김선영, 2022).

본 연구는 환자 및 환자단체와 관련된 국내 주요 의제들을 보도자료와 보도기사 등 빅데이터를 활용해 객관적인 수치로 확인하고 비교하는 최초의 작업이 될 것이다. 특히 인력, 예산 등이 부족한 환자단체에게 미디어 보도를 통한 이슈의 사회적 쟁점화는 적은 비용으로 실행할 수 있는 중요한 PR 방법이 될 수 있다. 따라서 보도자료의 의도가 보도기사에서 어떻게 반영 되었는지 그 차이를 비교하는 것은 환자단체와 같은 시민단체가 효과적인 PR전략을 수립하는 데 도움이 될 것이다. 또한 이러한 작업은 환자와 환자단체 당사자들은 물론이고 환자와 관계 맺고 있는 정부, 의료계, 일반 대중 등 사회가 환자를 더 잘 이해하고 그들이 마음 놓고 치료 받을 수 있는 사회적 환경을 조성하는데 기여할 수 있다는 점에서 의의가 있을 것이다.

선행연구 및 이론적 배경

1. 퍼블리시티와 보도자료

PR(Public Relations)로 분류되는 다양한 커뮤니케이션 활동 중 미디어에 뉴스 가치가 있는 정보를 제공하고 보도기사를 이끌어 내는 퍼블리시티(Publicity)는 가장 일상적이고 중요한 PR 활동 중 하나로 간주되고 있다(권순만, 1999; 홍미현, 2004).

미디어는 뉴스가치가 높은 기사를 선택하고 이를 기사화해서 수용자에게 소개하는 역할을 하게 된다. 이 점에서 미디어에게 퍼블리시티는 “외부로부터 제공된 뉴스 가치가 있는 정보”라고 할 수 있다 (Cutlip & Center, 1978). 반대로 PR 담당자에게 퍼블리시티는 미디어에 기사가 실릴 자리를 얻는 활동이다. 이처럼 퍼블리시티는 일방적이기 보다는 PR 담당자와 미디어 사이 서로의 필요에 의한 상호적인 활동이라고 할 수 있다(김진환, 2003; 홍미현, 2004).

이때 ‘보도자료’는 퍼블리시티 활동을 가능하게 하는 중요한 매개 수단 중 하나가 될 수 있을 것이다. 보도자료는 개인이나 각종 단체가 미디어에 기사를 보도하기 위해, 기자에게 제공하는 공식적인 자료이며(임현수·이준웅, 2011; 이희성, 2020), 미디어를 통해 개인과 단체의 활동 및 주장을 사회에 알림으로써 자신들에게 유리한 환경을 조성하는데 도움을 줄 수 있는 중요한 정보 제공 수단이다 (Boumans, 2018). 또한, 보도자료를 이용한 PR 활동은 비용이 거의 들지 않고(Kiousis et al., 2009), 신뢰도 높은 정보를 빠르게 공중에게 전달할 수 있으며, 조직의 시각을 반영한 매우 신뢰도 높은 정보를 공중에게 신속히 전달할 수 있다는 장점이 있다. 이런 이유로 뉴스 소재를 개발하고, 보도자료를 작성해, 미디어에 배포한 후 그 결과물을 모니터링하는 일련의 활동은 PR 담당자의 중요한 업무로 여겨져 왔다(김진환, 2003).

그렇다고 보도자료가 공급하는 쪽에만 이익을 주는 것은 아니다. 보도자료를 공급받는 미디어 역시 가치 있는 정보를 안정적으로 공급받을 수 있는

혜택을 누린다. 특히 정부, 대기업, 공공기관 등 공공성과 신뢰도가 높은 조직에서 제공하는 보도 자료는 정보의 중요도와 정확성으로 인해 사회에 미치는 영향이 큰 중요한 자료가 될 수 있다(서병호·김춘식, 2001). 미디어가 보도자료를 활용한다는 것은 미디어 입장에서도 특별한 노력과 비용을 지불하지 않고도 중요한 정보를 획득하는 데 소요되는 시간과 비용을 줄일 수 있다는 것을 의미한다 (Kiousis et al., 2009). 즉 미디어에 보도자료와 같은 정보를 적극적으로 제공하는 조직의 활동으로 인해 미디어는 신속하고 저렴하게 양질의 정보를 제공받을 수 있으며, 정보를 제공한 조직 역시 그만큼 해당 조직의 정보가 뉴스로 선택될 가능성을 높일 수 있다(Turk, 1986; 이희성, 2020).

이처럼 보도자료는 상호에게 혜택을 주는 매개체이지만, 그렇다고 보도자료가 뉴스 보도에 그대로 반영되는 것은 아니다. 미디어의 보도 시간 및 보도 공간의 유한함이 갖는 현실적 한계뿐만 아니라, 제공된 모든 보도자료가 수용자에게 제공할 만한 가치가 있는 것은 아니기 때문이다. 이때 제공된 보도자료의 선택에 기자의 가치 판단이 작용하게 된다(Lee & Choi, 2017; 진명지 외, 2020). 또한 보도자료가 기자의 선택을 받았다고 하더라도, 보도자료의 제공 주체가 구성한 현실과 언론보도를 통해 재구성된 현실은 다를 수 있다(박승정 외, 2017). 즉 미디어는 보도자료에서 제시한 현실을 그대로 반영하거나, 현실을 새롭게 구성하기도 하고, 혹은 보도하지 않음으로 해서 현실을 구성하지 않을 수도 있다는 것이다(방성현·이건호, 2013).

2. 의제설정과 프레이밍

미디어의 이러한 선택과 재구성의 기능은 게이트 키핑(Gate Keeping), 의제 설정(Agenda Setting), 프레이밍(Framing)과 같은 커뮤니케이션 이론들과 관련되어 있다. 특정 개인, 조직, 기관에서 미디어에 배포한 보도자료는 게이트 키핑, 의제설정, 프레이밍의 과정을 통해 대중에게 뉴스화된다고 할 수 있다. 세 가지 개념은 서로 밀접히 연관되어 있지만 차이 역시 존재한다(이정희, 2015; 김지현, 2020; 진명지 등, 2020).

게이트 키핑의 핵심은 미디어가 문지기처럼 뉴스를 '선택'한다는 점이다. 배포된 보도자료가 뉴스화 되기 위해서는 미디어, 특히 기자라는 게이트키퍼의 선택을 받아야 한다. 앞서 살펴본 바와 같이 미디어는 시간과 공간의 한계로 세상의 모든 것을 뉴스화할 수 없고, 제공된 모든 자료가 뉴스로서 가치가 있는 것도 아니기 때문에 미디어의 뉴스 선별 과정은 필수적이다. 미디어의 이러한 게이트 키핑 과정을 통과한 보도자료는 특정 사안에 대한 보도로 이어진다는 점에서 의제 설정과도 연결되어 있다고 할 수 있다(김방현, 2001; 박승정 외, 2017; 진명지 외, 2020).

의제 설정은 매스 미디어가 의식적으로 현재의 의제, 즉 논의할 이슈에 대한 대중의 생각과 의견을 새로 만들어 정하는 것을 말한다(McCombs & Shaw, 1972; McCombs, 2004). 매스 미디어에서 주목하지 않는 의제는 대중들의 생각 자체에서 소외될 수 있으며(이완수·손영준, 2011), 반대로 미디어

가 어떤 의제를 비중 있게 다루면 대중들은 그 논의를 중요한 것으로 인식하게 되어 적극적으로 중요한 의제로 주목을 받을 가능성이 높아진다. 다시 말해 매스 미디어가 사회적 파급 효과가 큰 이슈에 특별히 주목하고 미디어가 이를 자주 그리고 중요하게 다루게 되면 미디어 수용자들 역시 그 이슈를 중요한 의제로 인식하게 된다는 것이다(McCombs, 2004). 특히 의제 설정의 과정이 중요한 것은 미디어에서 중요하게 다룬 의제가 '공공의 의제'가 되어 사회적으로 큰 영향을 발휘할 수 있는 가능성이 높아지기 때문이다. 이처럼 미디어는 대중이 무엇에 관해 이야기하고 생각할 것인지를 결정하는데 영향을 줌으로써, 결국에는 사회적으로 중요한 공공의제에까지 영향을 미칠 수 있게 된다고 하였다(McCombs, 2014; 이정희, 2015).

전통적 의제 설정 이론은 미디어가 어떤 이슈를 비중 있게 취급하면 대중 역시 해당 이슈를 중요한 것으로 생각하게 된다는 현저성의 전이 정도의 수준에 머물러 있었다. 이에 비해 확장된 형태의 의제 설정 이론은 미디어가 대중들이 어떤 것을 중요한 것으로 인식하게 만드는 수준을 넘어 더 큰 영향을 행사한다고 보고 있다. 확장된 의제 설정 이론은 이슈 자체뿐만 아니라 이슈가 갖는 속성에까지 주목하고 있으며(반현 & McCombs, 2007; McCombs, 2014; 박승정 외, 2017), 그 속성은 기사나 미디어가 특정 이슈를 인식할 때 갖게 되는 일종의 틀, 즉 프레임(Frame)과 연관되어 있다고 본다(이정희, 2015).

프레임은 지각된 현실들 중에서 특정한 측면을

선택하여 이슈를 정의하고 기사 내용을 구성하는 방식이다(Gitlin, 1980). 이는 단순히 어떤 이슈를 논의의 전면에 위치시키는 것만을 의미하는 것은 아니다. 더 나아가 그 문제에 대해 생각하는 방식까지 제시하는 뉴스의 ‘틀 걸이’라고 할 수 있다(Entman, 1993). 의제 설정이 이슈를 선택하고 제외하는 것으로 의제를 운영하는 것과는 달리(Cobb & Elder, 1971; Lang, 1981) 뉴스의 프레임은 이슈가 어떻게 구성되고 전개되는가에 대한 논의에 초점을 맞춘다(Pan & Kosicki, 1993; 김재희, 2009). 이는 미디어의 보도가 현실에 대해 객관적으로 서술하는 것이 아니라, 미디어에 의해 특수한 방식으로 새롭게 구성된 주관적 현실이라는 것을 더 확연하게 보여준다고 할 수 있다(McCombs, 1997; 홍주현·김경희, 2010).

뉴스를 프레임하는 과정에서 미디어는 ‘선택’, ‘강조’ 그리고 ‘무시’의 방법을 활용하게 된다. 즉, 미디어는 특정 사건이나 이슈가 가지고 있는 다양한 특징 중에 자신들의 성향에 맞는 소수의 특성만을 선택하고 강조해 보도하며, 이와 반대되는 성향이나 의도에 적합하지 않는 특징은 무시하고 다루지 않는다는 것이다. 따라서 미디어의 프레임 기능은 특정한 사건이나 이슈에 미디어가 의도한 이미지를 생산해내는 역할을 한다고 이해할 수 있다(이정희, 2015). 이처럼 미디어에 배포된 보도자료는 게이트키퍼, 의제설정, 프레임 등 의 과정 속에서 최종적으로 하나의 보도기사가 되어 공중과 사회에 영향을 미친다고 볼 수 있다.

3. 의료 갈등과 이슈의 사회적 쟁점화

1990년 대 온라인에서 서로 정보를 주고받던 작은 커뮤니티 성격에서 시작된 국내 환자 모임은 2000년 만성골수성백혈병 치료제 ‘글리벡’의 국내 도입을 계기로 본격적인 환자운동으로 발전했다. 이후 2010년 2월 7개의 환자단체들이 모여 한국환자단체연합회를 결성하면서 국내에도 조직적인 환자 시민운동이 시작되었다(김용, 2014; 김영희, 2019; 한국환자단체연합회, 2022; 김동석·유현재, 2022). 그러나 환자와 관련된 시민운동의 시작이 늦은 만큼 오랫동안 잠재되어 있던 많은 이슈들이 환자단체연합회를 통해 분출하면서 이해집단 간 갈등도 많아졌다.

의료계는 다양한 관계 공중이 관여함으로 인해 어느 분야보다 갈등이 발생할 가능성이 높다고 할 수 있다. 갈등은 동일한 사안을 두고 관련 공중이 서로 배치되는 행위를 할 수 있는 상황으로 개인, 집단 사이에 발생하는 상호 대립의 과정이라고 하였다(Wagner & Hollenbeck, 2020). 갈등은 개인, 집단 간 상호 추구하는 가치, 목적, 목표 등이 다르고, 상대방의 입장이나 활동이 목표를 성취하는 데 어려움을 준다고 인식할 때 생겨날 수 있으며, 숨겨져 있거나 명시적으로 드러난 이해관계가 상호 대립적으로 나타나는 상태를 말한다(차희원, 2001).

의료 갈등은 하나의 사회현상으로 의료 소비자 측면에서는 환자를 포함한 의료 소비자들의 권리 의식 향상이나, 의사와 환자의 전통적인 관계의 변화 등과 같은 사회적 변화와 배경이 복합적으로

작용한다(황만성, 2005). 특히 의료사고를 둘러싼 책임 공방은 공급자인 의사와 소비자인 환자 사이에 가장 빈번하게 발생하는 갈등 중 하나이다. 1970년 대 까지 만 해도 의료는 국가에서 제공하는 시혜로 여겨졌다. 의료사고가 발생해도 개인의 운명으로 받아들여지곤 했다. 1980-1990년대 의료사고가 본격적으로 사회문제로 대두되어 의료 분쟁에 대한 행정적 차원의 대책이 마련되기 시작했지만, 의료 분쟁의 증가가 방어적 진료를 유도한다는 의료인의 입장을 대변하는 목소리가 더 힘을 얻거나, 의료사고가 의학에 대한 환자의 부족한 지식이나 경제적 동기로 분쟁화되고 있다고 간주되기도 했다(김정화, 1998). 이 같은 상황에는 의료 지식이 갖는 정보의 비대칭성이라는 근본적 문제나 환자의 권리를 대변해 줄 조직화된 단체가 부재한 것도 그 원인이 있었다. 하지만, 2010년 환자단체연합회의 출현은 환자 개인의 이슈를 사회적 이슈, 더 나아가 정책적 이슈로 발전시키고, 의료계와 환자 사이의 불균형을 해소하는 데 어느 정도 역할을 하게 된다(김영희, 2019; 김동석·유현재, 2022).

반대로 의료 공급자 측 요인으로는 의료의 상업화, 전문화, 대형화, 자동화로 인해 의료인이 환자를 대할 때 형식적이고 권위주의적인 태도를 취하게 되기 쉬우며 이는 의료 갈등으로 발전할 수 있다고도 볼 수 있다. 과도한 영리의 추구로 인하여 의사와 환자간의 인간적 신뢰관계의 악화 역시 의료 갈등을 촉진시키는 구조적 원인이 될 수 있다(황만성, 2005). 의료 갈등을 촉발 시키는 이러한 의료의 특수성은 의사 개인, 또는 병원 개별적인

문제라기보다는 사회 구조적인 문제로 의사와 병원 역시 환자와 다르지 않은 상황에 놓인 의료 갈등의 피해자라고 할 수 있다는 것이다.

이러한 갈등의 해법은 문제의 규정에서 시작되며, 문제는 공적 문제와 사적 문제로 구분할 수 있다고 하였다(김병진, 2000). 사적 문제란 개인, 즉 당사자에게 영향을 미치는 문제이며 그 해결은 개인 스스로가 담당해야 하는 문제를 의미한다. 반면, 공적 문제는 당사자 이외에 제 3자, 공공에게도 영향을 미치게 되는 문제이기 때문에 그 해결도 제3자, 즉 공공이 담당해야 할 문제라고 할 수 있다. 그러나 양자를 엄격히 구별한다는 것은 현실적으로 어렵다. 또한 많은 사람들에게 영향을 미치는 문제라고 모두 공공 문제가 되는 것은 아니며, 사회구성원의 불만족에 대한 용인의 정도를 넘어섰을 때 비로소 공공 문제가 된다고 하였다. 문제를 제기하는 쪽에서는 가능한 관련 이슈가 개인적인 범주가 아닌, 가능한 사회적, 공적 논쟁거리가 되도록 해야 할 것이다. 그래야만 문제를 해결 할 수 있는 결정권자나 대중의 관심을 끌고 해결을 필요로 하는 의제로 인정받을 수 있기 때문이다.

이런 이유로 환자단체와 같은 시민단체는 자신들이 직면한 이슈가 사회적 주목을 받기 위해 다양한 활동을 펼치곤 한다. 조직적으로 항의전화를 한다거나, 미디어를 통해 성명서를 발표하거나, 국회나 정부에 입법을 청원하는 활동 등 다양한 압력 수단을 통해 자신들의 이해관계를 관철시키려 한다(이원수 외, 2012). 특정 이슈와 관련된 이해 관계자들을 설득하거나 때로는 압력을 가하는 이러

한 커뮤니케이션 활동은 갈등 상황에서 해당 이슈를 순조롭게 풀어 가는데 중요하게 작용할 수 있다(Roloff, 1987; 차희원, 2001).

따라서 환자단체는 환자 개개인이나, 단체 차원에서 직면한 다양한 관련 문제들 중 특정한 이슈를 선별해 보도자료를 만들고 이를 미디어에 배포해 미디어가 관련 이슈를 의제로 채택하도록 노력하게 된다. 특정 이슈가 미디어에 의제로 채택되어 보도가 되면 일반 대중의 관심을 끌게 되고 미디어 의제에서 공공 의제로 발전할 수 있기 때문이다. 더 나아가 대중의 관심이 그 의제에 계속 집중되면 해당 의제는 공공 의제의 위상에서 정부 의제의 위상으로 진화해 해당 문제가 해결될 환경과 가능성을 높이는 데 기여하게 되는 것이다(박치성·명성준, 2009). 다시 말하면, 환자단체의 입장에서 미디어를 통한 의제 설정은 작은, 개인적, 조직 차원의 이슈를 큰 사회적, 공공적 이슈로 발전시켜 문제를 해결하는 결정적인 역할을 할 수 있다는 것을 의미한다.

4. 빅데이터와 텍스트 마이닝

빅데이터 분석이 도입되기 전까지 커뮤니케이션 연구에서 미디어 보도 기사의 내용을 분석하기 위한 시도는 주로 연구자의 수작업에 의존해 왔다. 커뮤니케이션 분야에서 대표적인 연구방법 중 하나인 ‘내용분석(Content Analysis)’은 연구자가 내용을 정독하여 그 함의를 정밀하고 세부적으로 분석해 낼 수 있다는 이점이 있지만, 연구자의 주관

성이 개입된다는 우려는 물론이고 자료의 규모에 대한 한계, 코더 간 신뢰도, 많은 시간과 노력을 요구한다는 단점 역시 있었다(박승정 외, 2017). 내용분석의 이러한 단점을 극복한 하나의 대안으로 특정 주제에 관한 일반적인 사회현상을 탐색할 수 있는 빅데이터 분석은 최근 사회과학연구 분야에서 중요하게 활용되고 있다(김선영, 2022).

빅데이터는 말 그대로 ‘대용량의 자료’라는 의미로 모바일 기기 등 커뮤니케이션 기술이 발달하면서 인간들이 남긴 흔적이 자료화되어 축적된 것을 말한다(서종근, 2021). 빅데이터 시대는 다양한 형태로 존재하는 자료의 축적을 가능하게 했고, 그 과정에서 텍스트 데이터(Text Data) 역시 그 비중이 높아지고 있다(Chakraborty, 2014). 텍스트 데이터 분석의 중요성이 강조되고 있음에도, 텍스트 데이터는 추상적 개념을 다루고 각 단어들 사이의 관계를 파악하기 어려운 단점이 있었다. 그러나 텍스트 데이터에서 키워드를 추출하고, 패턴을 분석해 의미 있는 정보를 추출해 내는 텍스트 마이닝(Text Mining) 기법이 개발되면서 복잡한 비정형 데이터인 텍스트 자료도 과학적으로 분석하는 것이 가능한 단계에 이르렀다(Hearst, 1999; 서종근, 2021).

텍스트 마이닝의 다양한 기법 중 키워드(주제어) 분석은 텍스트 마이닝에서 가장 기본적인 분석 방법으로 문서에 포함되어 있는 특정 단어의 출현 빈도와 가중치를 분석하는 방법이다(Clifton et al., 2004). 단어의 출현 빈도와 가중치는 TF(Term Frequency)와 TF-IDF(Term Frequency-Inverse

〈표 1〉 TF, TF-IDF의 의미

구분	의미	설명
TF (TermFrequency)	문서 내 특정 단어 단순 출현 빈도	문서 내 특정 단어가 얼마나 자주 출현하는 지를 나타내는 단순 빈도
TF-IDF (TermFrequency-InverseDocumentFrequency)	문서 내 특정 단어 중요도	문서 내 특정 단어가 얼마나 중요한 지를 보여주는 가중치

Document Frequency) 등으로 나타낼 수 있으며, 이때 TF는 특정 단어의 단순 출현 빈도를 의미한다. TF가 높으면 그 만큼 중요도가 높은 단어라고 할 수 있다. 그러나 반대로 특정 단어가 자주 언급된다는 것은 그 만큼 해당 단어가 일반적인 단어라는 의미이기도 하다. 이때 전체 문서 집합에서 한 단어가 얼마나 중요한 지를 수치적으로 나타낸 가중치인 TF-IDF를 사용하게 된다. TF-IDF 값은 한 문서에서 단어가 등장하는 빈도가 높을수록 커지고, 반대로 전체 문서에서 해당 단어를 포함하는 문서가 적을수록 TF-IDF값이 높아지게 된다(유주영, 2022).

본 연구에서는 TF-IDF 처리된 문서를 바탕으로 문서에서 토픽(주제)을 도출해 내는 토픽 모델링을 진행했다. 토픽 모델링 기법은 구조화되지 않은 방대한 문서의 집합에서 유사하게 범주화 될 수 있는 주제를 찾아내기 위한 알고리즘이다(Blei, 2012; 박승정, 2017). 명확한 토픽으로 할당되지 않는 일종의 노이즈 데이터들을 소수의 토픽에 부여함으로써 나머지 다수 토픽을 군집화 해 해석 가능성을 높인다(남춘호, 2016; DiMaggio et al., 2013; 김은혜·박주호, 2022). 즉, 주제와 맥락이 같은 것으로 보이는 키워드들을 구분해 한 덩어리로 묶어냄으로써 토픽

을 추출해 내는 분석 방법이라고 할 수 있다(백수미·문인호, 2021). 토픽 모델링 방법 중에서도 본 연구에서는 잠재 디리클레 할당(LDA, Latent Dirichlet Allocation, 이하 LDA)을 채택하였으며 그 이유는 LDA가 사회과학 연구에 적합한 토픽모델링 기법으로 알려져 있기 때문이다(DiMaggio et al., 2013). LDA는 주어진 문서 내 존재하는 단어들이 무리를 이뤄 각각의 토픽에 포함될 확률을 계산해 토픽을 구분해 낸다. LDA에서 단어의 의미는 해당 단어 단독이 아닌, 함께 사용하는 다른 단어들과의 관계에 따라 결정된다고 보고 있으며, LDA의 결과물 역시 한 문서나 문단에서 함께 사용된 단어의 동시 출현 여부를 중요하게 다룬다. 또한 LDA는 하나의 문서는 다양한 토픽을 포함하고 있다는 점에서, 하나의 텍스트가 다양한 관점을 반영할 수 있다는 사회과학의 관점과 맥을 같이 한다고 할 수 있다. 특히 토픽 모델링은 문서들을 관통하는 특정 토픽, 즉 주제를 분석함으로써 문서들이 담고 있는 전반적인 쟁점과 의견의 흐름을 분석할 수 있는 장점이 있다(안예지·김동일, 2020; 유주영, 2020)는 점에서 본 연구의 목적에 부합한다고 판단했다.

연구 문제

본 연구의 목적은 텍스트 마이닝 기법을 활용하여 2010년 창립 이래 2022년 현재까지 환자단체연합회가 보도자료를 통해 제기한 이슈들은 무엇이었는지, 해당 이슈들을 미디어는 어떻게 보도기사를 통해 의제화했는지 그 공통점과 차이점을 살펴보는 것이다. 그 결과를 통해 국내 환자와 환자단체가 직면한 핵심 의제들을 정량적이고 과학적으로 파악함으로써 정부에게는 환자 관련 정책을 입안하는데 중요한 객관적 자료를 제시하고, 의료인들의 입장에서는 의료 소비자인 환자들의 입장을 이해해 이를 진료 현장에 적용함으로써 환자 중심의 의료 실현에 도움을 줄 수 있을 것이다. 아울러 환자와 환자단체에게는 자신들이 제기한 이슈가 미디어를 통해 어떻게 의제화 되고 공공 의제화 되는지를 파악함으로써 환자단체가 환자들의 목소리를 효과적으로 커뮤니케이션할 수 있는 전략적 방향을 설정하는 데 중요한 근거 자료가 될 것으로 보인다. 이러한 연구 목적을 달성하기 위해 본 연구에서는 환자단체연합회가 창립 이후 2010년 10월 1일부터 2022년 5월 31일 약 12년 여 동안 배포한 보도자료와 한국언론재단에서 제공하는 기사 검색 서비스 빅카인즈(Bigkinds)에서 동일한 기간 동안 10대 일간지(경향, 국민, 동아, 문화, 조선, 중앙, 서울, 세계, 한국, 한겨레)와 3대 공중파(KBS, MBC, SBS) 방송에 보도된 관련기사들을 '환자단체연합회'를 키워드로 추출해 내어 분석 대상으로 삼았다.

한국환자단체연합회 홈페이지에 공개되어 있는 해당 보도자료는 총 262건이었다. 먼저 환자단체연합회가 미디어에 배포한 보도자료는 보도자료의 형식(홍보형, 정보제공형, 주장형)에 관계없이 전체 보도자료를 연구 대상으로 삼았다. 그 이유는 빅데이터를 활용한 텍스트 마이닝 연구라는 점에서 보도자료의 형식에 관계없이 배포된 모든 보도자료에 사용된 개별 단어와 문장이 모여 환자단체연합회의 의도가 반영되어 있을 것이라고 보기 때문이다.

빅카인즈를 통해 확보한 보도기사는 모두 210건이었다. 빅카인즈는 경제지와 지방지 등도 검색 서비스를 제공하고 있지만 매체의 특성에 따라 다루는 이슈가 편향될 수 있어 이번 연구 대상에서 제외했다. 예를 들어, 경제지의 경우 환자단체연합회와 관련된 이슈를 주로 경제적으로 바라보게 될 것이고, 지방지 역시 해당 지역에 초점을 맞춘 기사를 작성할 가능성이 많기 때문이다. 종합 일간지와 공중파 TV 역시 이념적 편향 등의 가능성이 없는 것은 아니지만 일반 대중에게 큰 영향력을 발휘하면서도 종합지, 종합방송으로서 뉴스를 종합적인 관점에서 제시한다는 점에서 10대 일간지와 3대 공중파 TV를 분석 대상으로 선정했다.

위 연구 자료를 기반으로 본 연구에서는 빅데이터의 텍스트 마이닝 기법을 활용해 아래와 같은 세 가지 연구 문제에 대한 해답을 얻고자 했다.

- [연구문제 1] 환자단체연합회가 배포한 보도자료의 주요 키워드와 의제는 무엇인가?

- [연구문제 2] 환자단체연합회와 관련하여 미디어가 보도한 보도기사의 주요 키워드와 의제는 무엇인가?
- [연구문제 3] 환자단체연합회 보도자료와 미디어 보도기사의 주요 키워드와 의제의 차이는 무엇인가?

연구 방법

본 연구에서는 파이썬 기반의 머신러닝 및 데이터 분석 도구 오렌지3(Orange3) 소프트웨어를 사용해 핵심키워드 추출과 토픽 모델링 연구를 진행했다. 오렌지3 소프트웨어는 머신 러닝, 텍스트 마이닝, 데이터 시각화, 통계 분석 등이 가능한 데이터 분석 도구로 코딩 없이 통계분석 및 시각화 작업이 가능하다는 장점(Blondel, et al., 2008; Lambiotte, et al., 2008)이 선택에 고려되었다.

본 연구는 다음과 같은 분석 절차로 진행했다.

- [1단계]: 연구 자료 수집 - 환자단체연합회 보도자료 및 13개 미디어 보도기사 수집
- [2단계]: 자료 데이터화 및 데이터 사전 처리 작업 - 수집 자료 데이터화, 유의어 및 제외어 등 사전 처리 작업
- [3단계]: 자료 분석 - 키워드 추출 분석, 토픽 모델링 분석
- [4단계]: 토픽명 부여 및 결과 논의 - 분석 결과 및 보도자료, 보도기사 원본 내용을 확인하며 토픽명 부여 및 차이점 확인

가장 먼저 환자단체연합회 홈페이지(<https://www.koreapatient.com>)에서 수집한 보도자료 원문을 엑셀(Excel) 스프레드시트에 입력했다. 보도기사는 빅카인즈에서 제공하는 엑셀파일 원문을 형식에 맞게 부분적으로 수정해 활용했다. 이러한 작업이 필요한 이유는 오렌지3가 엑셀(xls) 파일과 같은 문서를 인식하기 때문이다(Blondel, et al., 2008; Lambiotte, et al., 2008).

이어서 사전단계로 유의어와 제외어 등 불필요한 단어를 선별하는 데이터 클리닝 작업을 진행했다. 데이터 클리닝 작업은 데이터 마이닝 프로세스에서 중요한 단계로 수집된 자료를 더 이해하기 쉬운 형식으로 변환하고, 분석에 적합하게 가공하는 절차를 말한다. 이를 통해 비정형 자료에 대한 마이닝 알고리즘의 효율성을 향상시킨다. 유의어 선별 작업은 단어의 뜻은 같지만 사용 형태가 다른 경우 하나의 단어로 인식하게 하는 작업이다(백수미·문인오, 2021). 예를 들어, ‘한국환자단체연합회’를 ‘환자단체 연합회’처럼 띄어 쓰거나, ‘환자단체연합’, ‘환연’ 등으로 다르게 기술한 경우 모두 같은 의미이지만, 분석 프로그램 상에서는 각각 다른 뜻으로 인식된다. 따라서 의미가 같고 형태가 다른 단어들은 한 단어로 인식하도록 통일시키는 작업을 수행하였다. 또한, 제외어 선별 작업도 진행하였다. 제외어를 선별하는 사전 작업은 불필요하게 많이 사용된 단어나 큰 의미가 없는 단어 등을 인식하지 못하게 하는 작업이다. 예를 들어, ‘환자’ 등과 같이 모든 보도자료와 보도기사에서 빠지지 않고 등장하는 단어들은 분석에서 제외하였다.

본 연구에서는 오렌지3 프로그램을 이용해 TF(출현빈도), TF-IDF(중요도) 키워드를 추출한 후, 추출한 TF-IDF를 기반으로 테마를 찾아내기 위해 LDA 토픽 모델링 분석을 진행했다. 본 연구에서 채택한 LDA 토픽 모델링 분석을 실시하기 위해서는 먼저, 적절한 토픽 수를 결정하는 과정을 거쳐야 하며, 토픽 수는 연구자가 사전에 지정해야 한다. 토픽 수 지정을 위해 사용되는 방식은 연구자가 여러 개의 토픽 수를 지정하고 추출된 결과의 타당도를 검증하여 최적의 토픽 수를 결정하는 방법과 복잡도(Perplexity), 일관성(Coherence)과 같은 확률 분포의 지수 값을 이용하여 토픽 수를 결정하는 방법이 있다(DiMaggio et al., 2013; 박현정 외, 2017; 이수상, 2016; 김은혜·박주호, 2022). 본 연구에서는 위 두 가지 방법을 모두 활용해 토픽수를 최종 결정했다.

LDA 분석은 단어와 문서 사이에 존재하는 토픽이 단어와 문서에 어떠한 관계를 맺고 있는지에 대한 정보를 제공한다는 점에서 유용하지만(안예

지·김동일, 2020; 유주영, 2020), LDA분석 결과 자체로 해당 토픽이 어떤 의미를 가지는지에 대한 정보를 제공하지 않기 때문에 연구자가 판단하여 토픽에 명칭을 부여할 필요가 있다고 하였다(DiMaggio et al., 2013). 이에 따라 본 연구자는 특정 토픽에 포함되는 단어와 이를 포함하는 기사에 대한 분석을 통해 적절한 토픽명을 부여하였으며, 그 결과를 데이터 사이언스 전공 박사 1명, 석사 과정생 1명, 커뮤니케이션 전공 석사 1명의 검토를 받아 수정 보완하였다.

연구 결과

1. 환자단체연합회 보도자료 분석 결과

1.1 TF(출현빈도), TF-IDF(중요도) 키워드 분석

환자단체연합회가 창립 이래 2010년 10월부터 2022년 5월까지 배포한 보도자료는 총 262건이었으며, 본격적으로 보도자료를 배포하기 시작했던

〈표 2〉 TF 기준 주요 키워드(보도자료)

순위	주제어	빈도	순위	주제어	빈도	순위	주제어	빈도
1	환자단체연합회	925	11	cctv	487	21	유족	355
2	의사	794	12	개정안	487	22	대한의사협회	330
3	건강보험	706	13	의료기관	467	23	설치	312
4	의료인	674	14	피해자	450	24	환자안전법	293
5	의료사고	640	15	치료	445	25	약	284
6	국회	636	16	의료법	436	26	안전	280
7	수술실	625	17	의원	413	27	식약처	275
8	병원	551	18	수술	394	28	환자안전사고	267
9	환자단체	491	19	복지부	382	29	국민	266
10	정부	488	20	생명	357	30	전예강	260

2012년 - 2021년 10년 동안 평균을 내면 매년 25.5건으로, 이는 매달 약 2건(2.125건)을 배포한 셈이었다. 본 연구에서는 전체 보도자료 262건을 대상으로 키워드 분석을 실시하였다.

키워드 분석은 자료 속 문장을 문법적, 어휘적 의미가 있는 최소 단위인 형태소로 분절해 이 중 등장 빈도가 높은 순으로 집계한 것이다. 먼저, 특정한 단어가 문서 내에서 얼마나 자주 등장하는지를 나타내는 TF(Term Frequency)를 기준으로 높은 출현 빈도를 보이는 키워드 30개를 제시하면 다음 <표 2>와 같다.

분석 결과, 보도자료에서 1회 이상 출현한 키워드는 6,354개였으며, 가장 빈번하게 출현한 키워드는 ‘환자단체연합회(925)’, ‘의사(794)’, ‘건강보험(706)’, ‘의료인(674)’, ‘의료사고(640)’, ‘국회(636)’, ‘수술실(625)’, ‘병원(551)’, ‘환자단체(491)’, ‘정부(488)’ 등의 순으로 높은 빈도를 보였다. 이는 환자단체연합회의 이슈가 주로 의사, 병원, 국회, 정부 등과 같은 관계 공중과 밀접하게 연계되었다는 것

을 예측할 수 있으며, 건강보험, 의료사고 등과 관련된 이슈에 대해 보도자료를 통해 자주 문제 제기했다는 것을 알 수 있다.

TF는 값이 높을수록 중요할 수 있지만, 그만큼 상투적으로 흔하게 등장한다고도 볼 수 있다. 따라서 문서 집합에서 한 단어가 얼마나 중요한 지를 나타내는 TF-IDF(Term Frequency-Inverse Document Frequency)를 알아보았다. TF-IDF 기준 상위 키워드를 정리하면 다음 <표 3>과 같다.

TF-IDF 분석은 TF 결과와 차이가 있었다. 표2에서 보는 바와 같이 ‘수술실’, ‘건강보험’, ‘CCTV’, ‘전예강’, ‘사망아’, ‘넥시아’, ‘의료사고’, ‘환자안전사고’, ‘환자 안전법’ 등의 키워드가 수위를 차지했다. TF와는 달리 새롭게 10위 권 내에 언급된 키워드는 ‘CCTV’, ‘전예강’, ‘사망아’, ‘넥시아’, ‘환자안전사고’, ‘환자안전법’ 등 6개에 달했다. TF에서 7위를 기록했던 ‘수술실’이 1위로 올라서고, 11위에 위치했던 ‘CCTV’가 3위로 올라서면서 환자단체연합회가 보도자료를 통해 수술실 CCTV 설치에 대한 이슈를

<표 3> TF-IDF 기준 주요 키워드(보도자료)

순위	주제어	빈도	순위	주제어	빈도	순위	주제어	빈도
1	수술실	3,516,221,235	11	식약처	2,054,550,811	21	수혈	1,812,247,535
2	건강보험	2,939,890,949	12	개정안	2,049,185,189	22	병원	1,806,132,369
3	cctv	2,901,574,474	13	루타테라	2,047,205,595	23	의료기관	1,789,565,946
4	의료인	2,865,779,076	14	설치	2,021,025,839	24	수술	1,765,323,54
5	전예강	2,838,459,995	15	유족	1,974,817,088	25	김재윤	1,730,436,031
6	사망아	2,635,399,689	16	의사	1,964,888,095	26	대한의사협회	1,680,528,778
7	넥시아	2,461,043,188	17	피해자	1,953,768,959	27	환자단체연합회	1,527,118,747
8	의료사고	2,452,510,076	18	의료법	1,933,086,305	28	대리수술	1,505,905,735
9	환자안전사고	2,435,906,908	19	면허	1,894,744,575	29	유령수술	1,498,222,543
10	환자안전법	2,130,157,929	20	의원	1,831,111,568	30	국회	1,486,874,411

특히 중요하게 생각했다는 것을 알 수 있다. 또한, TF에서 30위에 있던 의료사고로 사망한 어린이의 이름인 ‘전예강’, 그리고 ‘사망아’, ‘의료사고’, ‘환자 안전사고’, ‘환자안전법’ 등이 10위 권 내에 위치하면서 창립 이후 12년 동안 의료사고와 관련된 이슈는 환자단체연합회 보도자료 전체에서 매우 중요하게 다루어졌으며, 이는 전예강과 같은 어린이 환자의 사망이 계기가 되거나 연관이 있었을 것으로 예측할 수 있다.

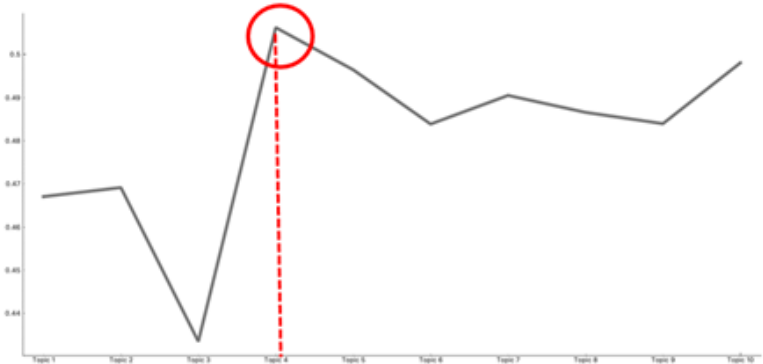
1.2 토픽모델링 분석

환자단체연합회의 보도자료에서 추출한 TF-IDF (문서 내 특정 단어의 중요도)를 기반으로 문서들을 관통하는 특정 의제를 분석해 내기 위해 토픽모델링을 진행했다(안예지·김동일, 2020; 유주영, 2020).

주제 일관성(Topic Coherence)과 혼잡도(Log Complexity) 값을 고려해 처음에는 토픽을 9개로 설정하였으나, 일반적으로 쓰인 단어의 빈도와 토픽 내에서 쓰인 단어의 빈도 사이에 차이가 컸고, 토픽이 분산되어 나타났다. 이에 추출되는 토픽과



<그림 1> 워드 클라우드(좌: TF, 우: TF-IDF)



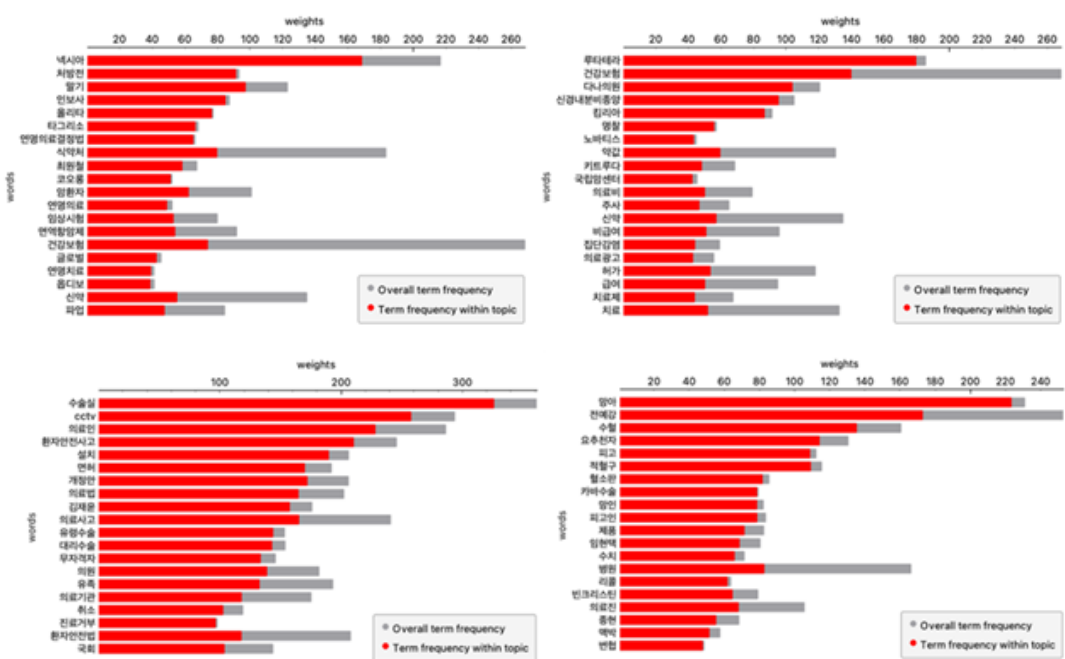
<그라프 1> 토픽 모델링 일관성 지수

키워드의 개수를 달리 설정하여 분석을 반복한 결과 4개의 토픽으로 설정하였을 때 키워드의 중복 정도가 낮으며, 분석된 주제 간의 배타성이 가장 잘 확보되었다. 또한 <그래프 1>에서 확인할 수 있는 것처럼 토픽 개수에 따른 일관성 지수 역시 4개의 토픽일 때 가장 높은 값을 보였다. 따라서 본 연구에서는 4개의 토픽과 각각의 토픽을 해석하는 데 필요한 키워드가 20개씩 추출되도록 설정하였다. 아래 <그래프 2>의 토픽모델링 결과에서 회색의 막대그래프는 일반적으로 키워드가 쓰인 정도이며, 빨간색은 해당 토픽 내에서 키워드가 쓰인 빈도를 나타낸다.

환자단체연합회의 보도자료에서 추출한 4개의 토픽은 <Topic 1: 환자 주권과 생명의 중요성>,

<Topic 2: 건강보험 급여화와 환자 피해 최소화>, <Topic 3: 수술실 CCTV 설치 법제화>, <Topic 4: 의료사고 예방>으로 테마를 각각 명명하였다.

<Topic 1: 환자 주권과 생명의 중요성>의 주요한 키워드들 중 가장 많이 언급된 '넥시아'와 넥시아의 개발자인 '최원철'은 효능 검증을 두고 논란이 많았던 한방항암제 넥시아와 관련된 키워드들이다. 환자단체연합회는 넥시아의 효능 검증에 미온적인 개발자, 병원, 정부당국을 대신해 환자단체연합회 스스로가 자체적인 검증을 벌이는 등 크게 이슈가 된 바 있다. 안전성 및 허가 이슈가 있었던 코오롱의 골관절염 유전자치료제 '인보사'와 관련된 키워드들은 투명한 정보 공개를 통한 환자의 알권리를 요구하는 내용이었으며, 키워드 '처방전'



<그래프 2> 토픽 모델링 결과(좌에서 우로 토픽 1, 2, 3, 4)

역시 환자의 알권리를 위해 약국 제출용 외에 환자 보관용 추가 처방전 발행의 필요성에 대한 문제 제기였다. ‘연명의료’, ‘연명치료’, ‘연명의료결정법’ 등에서도 가족이 아닌 환자 본인의 의사가 존중되는 연명의료 제도를 통해 환자들의 자기 선택권의 중요성을 강조한 바 있다. 즉 이들 키워드들은 공통적으로 환자의 알권리와 환자 스스로 판단하고 결정할 수 있는 환자 주권과 관련되어 있었다. 아울러 ‘넥시아’는 물론이고 ‘말기’, ‘암환자’, ‘면역항암제’, ‘올리타(비소세포페암 치료제)’, ‘타그리소(비소세포페암 치료제)’, ‘신약’ 등의 키워드들이나, ‘(의료계)파업’ 등의 키워드를 통해 말기 암환자 등 환자들의 생명이 제약사나 의료계의 이익과 타협의 대상이 될 수 없음을 강조한 것을 확인할 수 있었다.

〈Topic 2: 건강보험 급여화 및 환자 피해 최소화〉에서는 특히 건강보험 적용, 피해 보상, 환자 부담 기증 등 환자의 경제적 어려움과 관련된 키워드들이 많았다. ‘루타테라’, ‘키투루다’, ‘비급여’, ‘허가’, ‘급여’, ‘치료’, ‘약값’, ‘의료비’ 등의 키워드들은 건강보험 급여화와 직접적인 연관성을 갖는다. 특히 해당 토픽의 모든 키워드 중에서 가장 많이 출현한 신경내분비종양 치료제인 ‘루타테라’는 높은 약값으로 해외 원정치료를 받을 수밖에 없었던 환자들이 사회적 이슈가 되면서 건강보험 등재에 대해 환자단체연합회가 적극적으로 문제 제기한 부분이 반영된 것으로 보인다. ‘다나의원’, ‘집단감염’ 등의 키워드 역시 1회용 주사기를 반복 사용해 C형 간염 집단 감염사태에 처한 환자들의 피해보상과

C형 간염 치료제의 건강보험 적용으로 까지 이슈가 확산된 바 있다. 또한 ‘국립암센터’의 파업으로 인한 환자들의 혼란과 피해, ‘(연예인)의료광고’는 결국 최종 소비자인 환자들에게 부담으로 다가올 것이라는 주장 등 Topic 2는 환자들을 위한 올바른 치료 환경 조성을 위한 경제 제도적 지원과 관련된 것으로 분석됐다.

〈Topic 3: 수술실 CCTV 설치 법제화〉에서 수술실 CCTV 설치는 ‘유령수술’, ‘대리수술’, ‘무자격자’ 수술로 인해 의료사고가 빈번히 발생해 문제가 됐던 이슈이다. 또한 ‘환자안전사고’, ‘의료사고’, ‘(의사)면허’ 등의 키워드와 유족, 의료사고의 희생자인 ‘김재윤’ 어린이의 이름이 키워드에 포함되면서 의료사고 예방을 위한 수술실 CCTV 설치가 주요한 토픽이라는 것을 어렵지 않게 판단할 수 있었다. 여기에 더해 ‘환자안전법’, ‘의료법’, ‘개정안’, ‘국회’ 등의 키워드를 통해 환자단체연합회가 수술실 CCTV설치 법제화를 강력히 요구한 것을 알 수 있었다.

마지막으로 〈Topic 4: 의료사고 예방〉은 ‘사망아’, ‘사망인’, ‘전예강’, ‘종현’ 등 의료사고 희생자와 연관된 키워드들이 다수 출현했다. 특히 환자안전법을 촉발하는 계기가 되었던 예강이 종현이 등 어린이 피해자들이 집중 등장했다. 또한 ‘요추천자’, ‘빈크리스틴’ 등 의료사고의 원인이 키워드에 언급되어 있었으며, 신의료기술이지만 안전성 논란이 컸던 ‘카바수술’이나 의료기기와 의약품의 ‘리콜’ 등의 주제로 인해 Topic 4가 의료사고의 예방과 관련된 토픽이라는 것을 알 수 있었다.

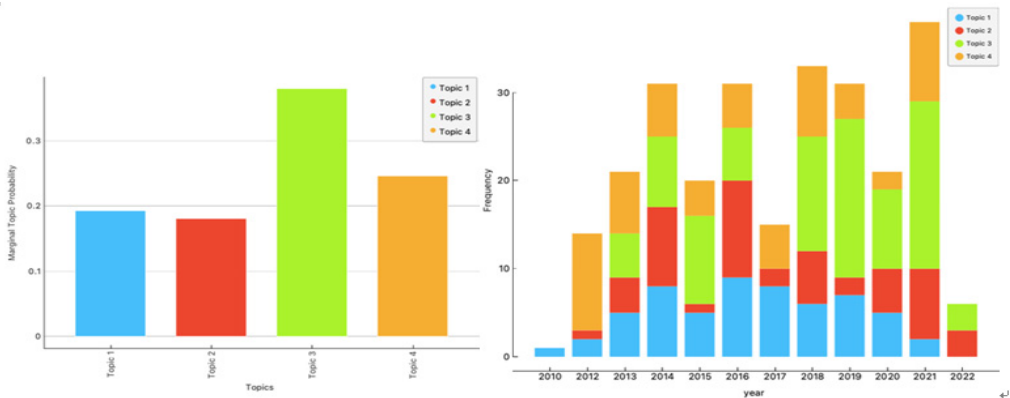
이들 4개 그룹의 토픽은 명확히 서로 다른 토픽으로 구분되었지만, 해당 토픽들 사이에서 공유하는 키워드를 함께 갖고 있기도 했다. 예를 들어 '건강보험'은 Topic 2의 주요한 키워드이기도 하지만, Topic 1의 키워드이기도 했다. 또한, 4개 토픽 모두 환자안전, 의료사고 등과 연관된 키워드들이 일정 부분 포함되어 있어 환자안전과 의료사고는 전체적으로 환자단체연합회의 중요한 토픽이었다는 것을 알 수 있다.

〈표 4〉와 〈그래프 3〉에 확인할 수 있는 것처럼 4개의 토픽 중 〈Topic 3: 수술실 CCTV 설치 법제화〉가 환자단체연합회 보도자료에서 가장 활발히 제기한 주제였으며, 그 다음은 〈Topic 4: 의료사고

예방〉이었다. 이어 〈Topic 1: 환자 주권과 생명의 중요성〉과 〈Topic 2: 건강보험 급여화 및 환자 피해 최소화〉는 거의 동일했다. 〈Topic 1: 환자 주권과 생명의 중요성〉은 2010년 처음 제기된 이후 지속적으로 상승하다가, 최근 2020년 대 들어서 줄어드는 경향성을 보였다. 〈Topic 2: 건강보험 급여화 및 환자 피해 최소화〉의 경우 2014년, 2018년, 2020년, 2021년에 특징적으로 관련 이슈가 두드러진 것으로 봐서 건강보험 급여화 이슈가 해당 기간에 집중적으로 논의되었다는 것을 알 수 있다. 〈Topic 3: 수술실 CCTV 설치 법제화〉는 2013년부터 꾸준히 제기되어온 이슈였으며, 특히 2018년 이후 최근 (2020년)에 관련 이슈에 대한 보도자료가 증가했

〈표 4〉 토픽별 연도별 보도자료 배포 분포

토픽/년도	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	총합계
Topic 1	1		2	5	8	5	9	8	6	7	5	2		58
Topic 2			1	4	9	1	11	2	6	2	5	8	3	52
Topic 3				5	8	10	6		13	18	9	19	3	91
Topic 4			11	7	6	4	5	5	8	4	2	9		61
총합계	1	0	14	21	31	20	31	15	33	31	21	38	6	262



〈그래프 3〉 토픽별 연도별 보도자료 배포 분포

던 것을 알 수 있다. 마지막으로 <Topic 4: 의료사고 예방>과 관련된 주제는 거의 매해 변화 없이 비중 있게 다뤄졌다. 다시 말해 의료사고 이슈는 환자단체연합회의 지속적이고, 중요한 이슈였다고 볼 수 있다.

2. 미디어 보도기사 분석 결과

2.1 TF(출현빈도), TF-IDF(중요도) 키워드 분석

2010년 10월부터 2022년 5월까지 한국환자단체연합회 빅카인즈 기사 검색 서비스에서 '환자단체연합회'를 키워드로 검색한 결과 총 210건의 보도기사가 검색되었다. 먼저, TF를 기준으로 높은 출현 빈도를 보이는 키워드 50개를 제시하면 <표 5>와 같다. 분석 결과, 보도기사에서 1회 이상 출현한 키워드는 6,938개였으며, 가장 빈번하게 출현한 키워드는 '병원(480)', '의사(468)', '건강보험(323)', '정부(303)', '수술(298)', '수술실(288)', '환자단체연합회(273)', 'CCTV(268)', '복지부(263)', '약(203)' 등이었다.

환자단체연합회의 보도자료와 비교했을 경우 약간의 순위 변동은 있지만 대부분의 키워드가 동일하게 1-10위 내에 위치했다. 눈에 띄는 변화는 보도자료에는 6위에 해당하던 '국회'가 보도기사에서는 18위로 내려갔고, '의료사고' 역시 5위에서 20위로 내려앉았다. 대신 '복지부'(19위 → 9위), 'CCTV'(11위 → 8위) 등이 10위에 올라섰고, '약'이나 '수술'과 같은 키워드가 새로 등장했다. 이는 환자단체연합회의 경우 국회를 통한 법제화의 필요성을 주장하거나, 의료사고와 관련된 이슈를 중점적으로 제기한 데 비해 미디어는 이 부분에 대해서 상대적으로 덜 다뤘던 것을 알 수 있다.

<표 6>에서 보는 바와 같이 TF-IDF 보도기사 분석에서는 'CCTV', '수술실', '건강보험', '의료진' 등의 키워드들의 경우 보도자료와 크게 변화가 없었다. 그러나 보도자료에서 8위에 올라있던 '의료사고'는 보도기사에서는 13위로 내려갔고, 다섯 번째에 위치해 있던 피해 어린이 '전예강'은 31위로 변동했다. 특히 보도자료에서 6번째로 중요한 키

<표 5> TF 기준 핵심 키워드(보도기사)

순위	주제어	빈도	순위	주제어	빈도	순위	주제어	빈도
1	병원	480	11	치료	190	21	의원	129
2	의사	468	12	의료진	174	22	의료	125
3	건강보험	323	13	안기종	169	23	의료계	123
4	정부	303	14	진료	160	24	설치	123
5	수술	298	15	문제	160	25	법안	116
6	수술실	288	16	의료기관	158	26	비급여	110
7	환자단체연합회	273	17	대한의사협회	146	27	교수	107
8	cctv	268	18	국회	140	28	서울	104
9	복지부	263	19	부담	139	29	감염	103
10	약	203	20	의료사고	132	30	사고	103

〈표 6〉 TF-IDF 기준 핵심 키워드(보도기사)

순위	주제어	빈도	순위	주제어	빈도	순위	주제어	빈도
1	cctv	2,215708655	11	설치	1,049459118	21	의료비	0,904819656
2	수술실	1,99582249	12	대한의사협회	1,026106438	22	치료	0,896075018
3	건강보험	1,851824837	13	의료사고	1,011646688	23	복지부	0,880068924
4	의사	1,566054208	14	법안	1,005745965	24	국회	0,869370842
5	수술	1,3518676	15	정부	0,972887833	25	부담	0,86316105
6	비급여	1,25854067	16	보장	0,942963419	26	존엄사	0,809237189
7	병원	1,222681544	17	진료	0,941892317	27	피해자	0,801688518
8	약	1,163840378	18	급여	0,9242006	28	의료계	0,795472326
9	감염	1,106546205	19	의료기관	0,905846205	29	항생제	0,775660058
10	의료진	1,051929501	20	킴리아	0,905328724	30	사고	0,766574016

워드클라우드를 분류했던 ‘사망아’의 경우 보도기사에서는 찾아 볼 수 없었다. 또한 7위에 있던 ‘넥시아’는 122위로, ‘환자안전법’(10위)은 237위로 큰 폭으로 하락해 ‘전예강’, ‘넥시아’, ‘환자안전법’ 등 특정 이슈를 명확히 지칭하던 키워드들의 하락세가 눈에 띄었다. 대신 ‘비급여’, ‘병원’, ‘약’, ‘감염’, ‘의료진’과 같은 일반적 주제어들이 10위권 안에 등장하는 특징을 보였다.

이는 보도자료의 경우 환자단체연합회가 해결해야 할 특정 이슈를 반복적으로 제기한 데 비해,

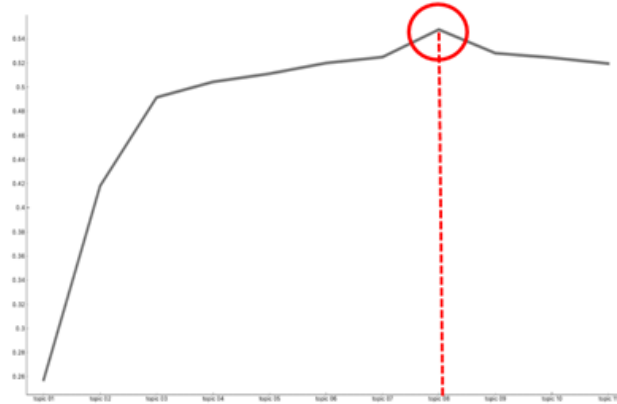
미디어는 특정 이슈에 집중하기 보다는 좀 더 포괄적으로 관련 이슈에 접근하거나, 보다 중립적인 입장에서 보도기사를 작성했던 것을 짐작할 수 있다.

2.2 토픽모델링

앞에서도 언급했듯이 토픽모델링에서 토픽의 개수를 선택할 때 혼잡도와 일관성 지수를 참고했다. 특히 토픽 개수에 따른 일관성 지수를 분석한 결과 8개의 토픽일 때 가장 높은 값을 보였으나, 4개-11개 토픽 사이에는 큰 차이가 없었다. 본 연구의 목적이 보도자료와 보도기사의 차이를 비교 분



〈그림 2〉 워드 클라우드(좌: TF, 우: TF-IDF)



〈그래프 4〉 토픽 모델링 일관성 지수

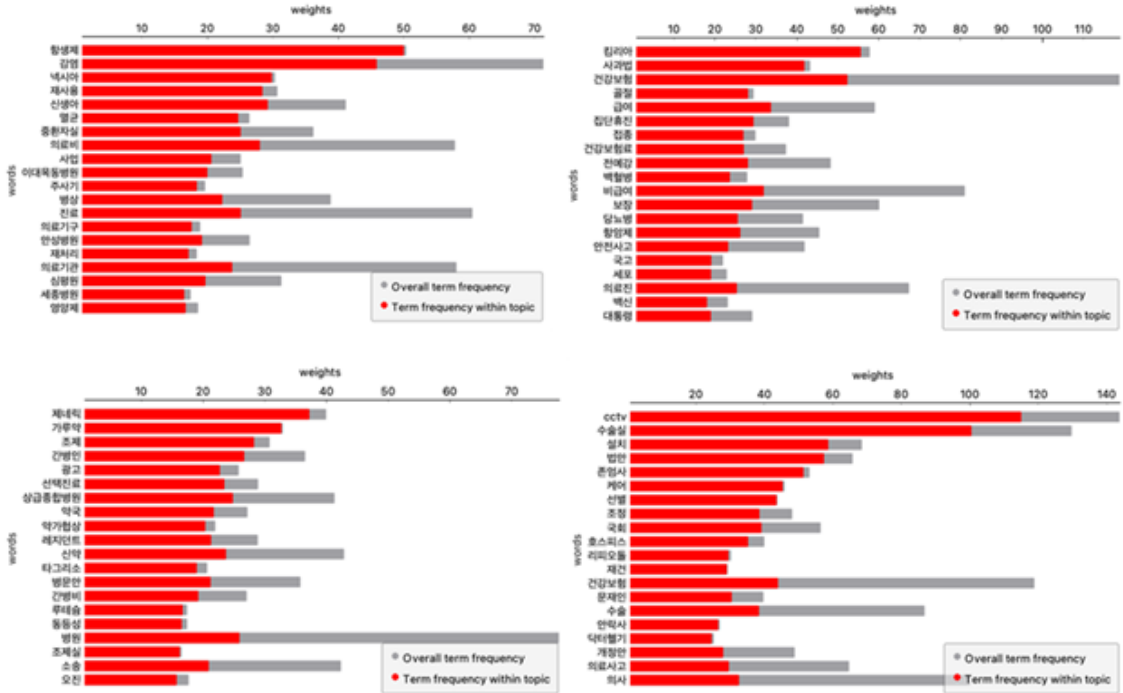
석하는 것이라는 점과 앞에서 분석한 보도자료가 4개의 토픽을 설정한 점, 그리고 보도기사의 경우 411개 토픽 사이 토픽의 일관성에서 큰 차이가 없었다는 점에서 보도기사 분석을 위한 토픽 수 역시 4개로 정하고, 각각의 토픽을 해석하는 데 필요한 키워드를 20개씩 추출되도록 설정했다.

미디어의 보도기사에서 추출한 4개의 토픽은 〈Topic 1: 병원의 잘못된 의료 관행〉, 〈Topic 2: 환자 중심성〉, 〈Topic 3: 환자의 치료 접근권〉, 〈Topic 4: 환자의 인권과 생명의 중요성〉으로 토픽명을 부여했다.

〈Topic 1: 병원의 잘못된 의료관행〉에 해당하는 키워드들 대부분은 ‘이대목동병원’, ‘안성병원’, ‘밀양세종병원’ 등 병원의 이름 및 병원과 연관되어 있었다. ‘이대목동병원’은 신생아중환자실 환자 사망 사건과 관련해 병원명 이외에 ‘중환자실’, ‘감염’, ‘항생제’, ‘주사기’, ‘영양제(허위청구)’ 등의 키워드가 함께 출현한 것을 관련 보도기사 내용을 통해서

확인할 수 있었다. ‘밀양세종병원’은 환자를 신체 보호대로 결박해 화재 시 참변을 당한 사건과 관련해 진상 규명 및 재발 방지에 대한 이슈와 연관이 있었다.

또한 병원명이 직접 언급되지는 않았지만, ‘주사기’, ‘재사용’의 키워드를 통해 다나의원의 C형 간염 집단감염 사태가 Topic 1에 자주 등장했음을 알 수 있었다. ‘넥시아’ 역시, 해당 키워드가 환자단체연합회의 보도자료에서 한방암치료제 효능에 대한 자체적인 검증에 초점이 맞춰진 것과는 다르게, Topic 1 보도기사에서는 넥시아의 개발자가 속해 있는 단국대병원이 함께 언급된 것을 확인할 수 있었으므로, 이 역시 병원과의 연관성에 포함되는 키워드라는 것을 알 수 있었다. 그렇다고 환자가 병원과 대립하는 이슈만 있었던 것은 아니다. ‘안성병원’의 경우 환자들이 지속적으로 주장한 수술실 CCTV설치를 실행한 병원이었기 때문이다. 하지만 이 역시 유령수술, 대리수술, 무자격수술



〈그래프 5〉 토픽 모델링 결과(좌에서 우로 토픽 1, 2, 3, 4)

등 수술실 CCTV 설치가 의료계의 부정적인 면과 연관되어 있다는 점에서 병원의 잘못된 의료관행과 연관된 것으로 분류할 수 있었다.

〈Topic 2: 환자 중심성〉에서는 환자 중심 안전 문화, 환자중심 연구, 환자중심 치료환경, 환자중심 정책과 제도 등 다양한 주제에서 환자 중심성을 다루고 있었다. 예를 들어 ‘당뇨병’이라는 키워드는 보도자료 원문을 확인한 결과 환자중심 의료 연구 사업에서 당뇨병 환자의 개인정보를 활용한 환자단체의 연구가 배제되면서 발생한 ‘환자가 빠진 환자중심 의료연구’ 논란과 연관된 주제로 파악되었다. ‘건강보험’, ‘건강보험료’, ‘비급여’, ‘급여’, ‘국고’, ‘보장’, ‘백신’, ‘김리아(백혈병치료제)’, ‘접

중’, ‘집단휴진’ 등은 환자중심의 치료 환경을, ‘전예강’, ‘사과법’ 등은 의료사고 시 의료진이 환자에게 사과하는 것이 법적으로 문제가 되지 않게 하는 법의 제정과 관련된 것으로 환자중심의 안전 문화와 관련된 키워드들이었다. ‘대통령’은 대선과 관련해 대통령 후보들에게 환자중심의 정책에 대해 호소한 보도기사 내용이 포함된 것을 확인할 수 있었다.

〈Topic 3: 환자의 치료 접근권〉은 환자들의 해외 원정 치료를 야기했던 노바티스의 방사선 치료제 ‘로테슈’, 폐암표적항암제 ‘타그리소’의 국내 철수 등의 이슈와 관련된 키워드들로 환자들의 의약품 접근권과 치료 받을 권리와 관련이 있었다. 토

픽 내에서 가장 많이 출현한 것으로 나타난 ‘제네릭’, 그리고 ‘신약’, ‘동등성’ 등의 키워드는 경제적 상황에 관계없이 환자들의 의약품에 대한 치료 접근권이 향상되어야 한다는 점과 연관되어 있었으며, ‘간병인’, ‘간병비’ 등은 병원 입원 치료 시 간병의 문제점 개선을, ‘가루약’은 일부 약국에서가루 약 조제를 기피함으로써 환자들이 의약품 접근권을 침해당하고 있다는 내용의 관련 기사를 확인할 수 있었다. 또한 키워드 ‘레지던트’는 2020년 전공의(인턴, 레지던트)파업과 관련해 환자들의 치료받을 권리가 논의된 것을 근거로 Topic 3을 <환자의 치료 접근권>으로 분류하였다.

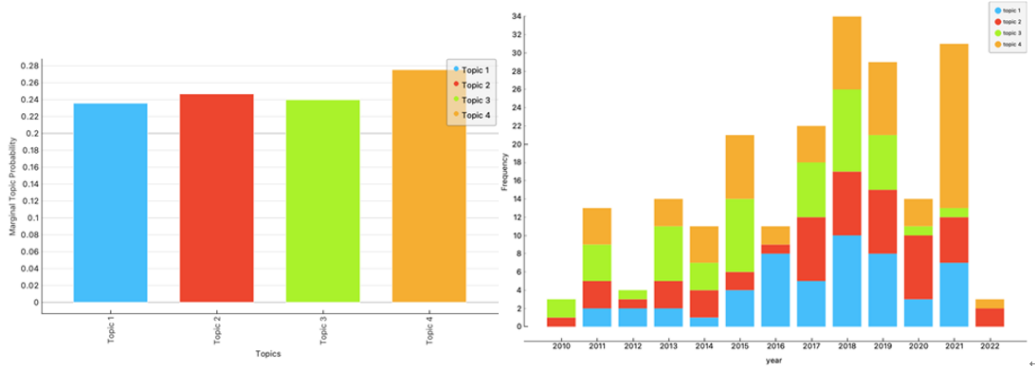
<Topic 4: 환자의 인권과 생명의 중요성>에서 가장 대표적인 키워드인 ‘수술실’, ‘CCTV’, ‘설치’는 수술실 내 환자의 인권과 의사들의 인권과 관련해서 첨예하게 대립했던 문제 중 하나였다. 환자단체 연합회의 보도자료에서 수술실 CCTV 설치가 무자격, 유명수술 및 법제화와 관련된 키워드들과 함께 출현함으로써 해서 ‘수술실 CCTV 설치 법제화(Topic 3)’로 분류된 것과는 달리, 미디어 보도기사에서는 ‘존엄사’, ‘안락사’, ‘호스피스’ 등 환자의 인권과 관련된 키워드들과 함께 출현한 것을 알 수 있다. 따라서 보도기사에서의 수술실 CCTV 설치는 환자

의 안전과 인권 침해와 관련된 목적이 더 강조된 것으로 판단할 수 있었다. ‘존엄사’, ‘호스피스’, ‘안락사’ 등의 키워드는 환자의 인권뿐만 아니라, 생명과도 밀접한 관계가 있는 키워드들이라고 할 수 있다. ‘리피오돌’(간암치료제)의 경우에는 제약사가 약가 인상을 요구하며 공급 중단을 선언하자 환자의 생명을 담보로 한 약가협상이라는 비난이 있었으며, ‘닥터헬기’의 도입으로 공공의료의 강화의 필요성을 강조한 것 역시 환자들의 생명과 관련된 키워드라는 점에서 Topic 4를 <환자의 인권과 생명의 중요성>으로 분류했다.

미디어 보도기사는 4개의 토픽을 비교적 큰 편차 없이 골고루 다뤘지만, 그 중 <Topic 4: 환자의 인권과 생명의 중요성>을 다른 토픽보다 좀 더 많이 다뤘다. Topic 1, 2, 3은 그 편차가 거의 없었다. 시기별로는 <Topic 1: 병원의 부적절한 의료 관행>의 경우 2011년 등장한 이래 지속적으로 최근까지 논의가 된 것으로 봐서 초기부터 병원과의 대립이 불가피했음을 알 수 있다. <Topic 2: 환자 중심성>의 경우 미디어에서 꾸준히 다루었지만 2013년 2014년에 증가했다가, 특히 2017년 이후 주요 이슈로 급부상해 지속됐던 것을 알 수 있다. 이는 어린 환자들의 죽음과 관련해 환자안전법이 이슈가 됐

〈표 7〉 토픽별 연도별 보도기사 배포 분포

토픽/년도	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	총합계
Topic 1		2	2	2	1	4	8	5	10	8	3	7		52
Topic 2	1	3	1	3	3	2	1	7	7	7	7	5	2	49
Topic 3	2	4	1	6	3	8		6	9	6	1	1		47
Topic 4		4		3	4	7	2	4	8	8	3	18	1	62
총합계	3	13	4	14	11	21	11	22	34	29	14	31	3	210



〈그래픽 6〉 토픽별 연도별 보도기사 게재 분포

던 시기와 연관된 것으로 보인다. 〈Topic 3: 환자의 치료 접근권〉의 경우 환자단체연합회 창립 이래 주요한 주제였지만 2020년 이후에는 해당 주제가 크게 줄어든 것을 알 수 있다. 〈Topic 4: 환자의 인권과 생명의 중요성〉은 반대로 2021년 이후 크게 증가한 것을 알 수 있다. 그 이유는 2021년 사회적으로 연명의료법이 본격 논의되면서 언급이 늘어났던 것으로 보인다.

논의 및 시사점

환자단체연합회의 보도자료와 미디어의 보도 기사 사이에 가장 큰 차이점은 다음 세 가지 정도로 요약할 수 있겠다. 첫째, 환자단체연합회가 ‘법제화’를 목표로 관련 이슈들을 의제화하려 노력한 것과는 달리 미디어는 법제화 보도에 대해서 환자단체만큼 적극적이지는 않았다는 점이다. 둘째, 환자단체연합회가 ‘의료사고’에 대해 집중적으로 의

제화 하려 했지만, 미디어는 중립적으로 해당 이슈들을 바라보려 한 것으로 보인다. 셋째, 환자단체연합회는 직면한 이슈 해결에 보다 집중한 것에 비해, 미디어는 개별 이슈를 통해 보다 거시적 방향성을 제시하려 했다는 점 등이다.

1. 환자단체, 제도적 정착을 위한 법제화 목표

환자단체연합회가 배포한 보도자료의 키워드 출현 빈도(TF) 분석에서 ‘환자단체’, ‘환자단체연합회’와 같은 환자단체를 상징하는 일반적인 키워드를 제외한 상태에서 10위 권 내에 가장 빈번하게 출현한 키워드는 ‘의사(2위)/의료인(4위)’, ‘건강보험(3위)’, ‘의료사고(5위)’, ‘국회(6위)’, ‘수술실(7위)’, ‘병원(8위)’, ‘정부(10위)’ 등이었다. 이를 통해 환자단체연합회는 건강보험, 의료사고, 수술실(CCTV 설치) 등의 의제를 집중 거론했으며, 이들 논의에 의사, 병원, 국회, 정부 등의 관계 공중들(Stakeholders)이 연관되어 있었다는 것을 알 수 있

다. 특히 환자단체연합회 보도자료에서 ‘국회(6위)’, ‘개정안(12위)’, ‘의료법(16위)’, ‘환자안전법(24위)’ 등의 키워드의 빈도가 높았다는 것은 환자단체연합회가 자신들의 호소가 단순히 일회성 주장에 그치지 않고 이를 ‘법제화로 연결 시켜 사회적 제도로 정착시키기 위해 노력했다는 것을 암시해 준다. TF-IDF 분석을 통해 추출된 키워드에서도 ‘환자안전법(10위)’, ‘개정안(12위)’, ‘의료법(18위)’, ‘국회(30위)’ 등 ‘환자안전법’과 ‘국회’의 순서만 바뀌었을 뿐 이 같은 경향은 그대로 유지되는 것을 확인할 수 있었다.

반면, 미디어의 보도기사에서는 법제화와 연관되어 있는 단순 언급 키워드(TF)들 중 ‘국회(18위)’가 가장 많이 출현하기는 했지만, 이는 환자단체연합회의 보도자료에서 ‘국회’가 6위에 올라 있었던 것에 비해 크게 하락한 것을 알 수 있다. 뒤를 이어 ‘법안(25위)’, ‘개정안(32위)’ 등의 키워드 정도가 법제화와 관련해 50위 내에 언급되었을 뿐이었다. 미디어의 보도기사에 대한 TF-IDF 분석에서도 ‘법안(14위)’과 ‘국회(24위)’, ‘개정안(32위)’ 등의 법제화와 관련된 키워드들이 중요하게 출현했지만, 환자단체연합회 보도자료에서 ‘환자안전법(10위)’, ‘개정안(12위)’, ‘의료법(18위)’ 등 법제화 관련 키워드가 대부분 20위 권 내에 위치했던 것과는 격차가 있었다. 토픽모델링을 통해 얻은 결과물에서도 이 같은 경향성은 크게 다르지 않은 것을 발견할 수 있었다. 환자단체연합회의 보도자료에서 추출한 4개의 토픽에 <Topic 3: 수술실 CCTV 설치 법제화>와 같이 법제화와 관련된 토픽이 포함된 것과는

달리, 미디어의 보도기사에는 법제화와 관련된 토픽은 존재하지 않았다. 이는 환자단체연합회의 희망과는 달리 미디어가 관련 제도의 법제화와 관련된 보도에 환자단체연합회만큼 적극적이지 않았다는 것을 알 수 있다. 이러한 결과는 환자단체연합회가 법제화를 주요한 해법으로 제시했지만, 미디어는 세부적 해법의 제시보다는 포괄적이고 거시적 관점에서 관련 이슈를 사회적 이슈로 부각시키는 것에 보다 중점을 두었던 것으로 보인다. 이는 보도자료의 의도가 보도기사에 정확히 반영되지는 않을 수 있다는 것을 다시 한 번 확인해 준 결과라고 할 수 있다.

2. 미디어, 의료사고에 대한 중립적 태도

법제화를 둘러싼 차이 외에도 ‘의료사고’에 있어서도 환자단체연합회와 미디어는 서로 다른 태도를 확인할 수 있었다. 환자단체연합회의 보도자료에 언급된 키워드 출현 빈도 분석(TF)에서 ‘의료사고’는 5번째로 많이 언급된 키워드였다. ‘의료사고’ 키워드 외에 10위 권 밖에서도 ‘피해자’(14위), ‘유족’(21위), ‘안전’(26위), ‘환자안전사고’(28위), ‘전예강’(30위) 등 의료사고와 연관되어 있는 다양한 키워드들 역시 높은 빈도를 보였다. TF-IDF 분석에서도 ‘의료사고’ 키워드는 8위를 차지할 정도로 중요한 키워드 중 하나였다. 특히 ‘의료사고’라는 직접적 키워드를 제외하고도, 10위 권 내에 위치한 키워드들(수술실, CCTV, 전예강, 사망아, 넥시아, 환자안전사고, 환자안전법) 거의 대부분이 의료사

고와 직접, 혹은 간접적으로 연관된 키워드들이라는 점은 환자단체연합회가 의료사고와 관련된 이슈를 집중적으로 의제화 하려 했다는 것을 알 수 있게 해 준다.

그러나 미디어 보도기사의 키워드 언급량(TF) 분석에서는 환자단체연합회 보도자료에서 5위로 나타났던 ‘의료사고’ 키워드가 20위로 큰 하락을 보였다. ‘감염’(29위), ‘사고’(30위), ‘피해’(31위), ‘피해자’(33위) 등 연관 키워드들이 있기는 했지만 출현 빈도가 환자단체연합회의 보도자료보다는 높지 않았다. 또한 환자단체연합회 보도자료에서 30위를 기록한 의료사고의 피해자였던 ‘전예강’ 어린이의 이름과 ‘사망아’ 키워드 역시 미디어 보도기사에서는 50위권 내에서는 찾아 볼 수 없었다. 대신 미디어의 보도자료에 출현한 키워드(TF)들은 ‘병원’(8위→1위), ‘정부’(10위→4위), ‘수술’(18위→5위), ‘복지부’(19위→9위), ‘약’(25위→10위) 등 일반적이고 중립적인 단어가 10위권 내로 새로 진입한 것을 알 수 있다. 특히 미디어의 보도기사에 ‘병원’이라는 키워드와 ‘복지부’라는 키워드가 10위권 내로 진입한 것으로 봤을 때 미디어는 환자단체연합회와 병원 사이의 관계에 더 관심을 기울였을 가능성이 있으며, 환자단체연합회가 법제화를 위해 ‘국회’라는 기관에 주목했던 것과는 달리 미디어의 경우 국회 대신 보건의료정책의 수립과 집행 역할을 하는 ‘복지부’와 연결 시켜 환자관련 의제를 다루었을 것으로 추정할 수 있다.

토픽모델링 분석 결과에서도 환자단체연합회와 미디어 사이의 이 같은 차이는 유사하게 나타났

다. 환자단체연합회의 보도자료에 <Topic 4: 의료 사고 예방>이라는 토픽이 추출된 것과는 달리, 미디어의 토픽에서는 ‘의료사고’로 분류된 토픽이 존재하지 않는다. 이 같은 현상이 나타난 배경에는 의료사고와 같은 병원과 환자 간 양측의 입장이 서로 첨예하게 대립할 수 있는 이슈에 대해서는 미디어가 자주 기사화하거나 언급하는 것을 자제하는 중립적인 태도를 가지고 있었던 것으로 볼 수 있다. 그럼에도 불구하고 미디어가 의료사고에 대한 토픽을 외면했다고 할 수는 없다. 보도자료에서 보다는 그 빈도가 하락했지만 ‘의료사고’(20위), ‘감염’(29위), ‘사고’(30위), ‘피해’(31위), ‘피해자’(33위) 등 의료사고 관련 키워드 순위가 여전히 높은 빈도를 차지하고 있으며, ‘병원’(1위), ‘수술’(5위) 등의 키워드 역시 직간접적으로 의료사고와 연관이 있기 때문이다. 따라서 환자단체연합회는 의료 사고에 대한 미디어의 중립적 태도, 병원과 환자와의 관계에 관심을 갖는 미디어의 성향을 고려해 보도자료를 구성할 필요가 있다고 하겠다.

3. 의제 접근 및 기술 방식의 차이

위에서 논의했던 것처럼 환자단체연합회의 보도자료에서 추출한 키워드들 중 ‘법제화’와 연관된 키워드들은 환자단체연합회의 커뮤니케이션 활동의 목적이 무엇인지 가늠할 수 있게 해 준다. ‘의료 사고’와 관련된 다양한 키워드들 역시 환자단체연합회가 ‘의료사고’와 관련된 이슈들을 중점적으로 다뤘다는 것을 알 수 있다. 이 같은 의료 현장에서

환자단체연합회가 겪고 있는 현실적 이슈 중심의 커뮤니케이션은 토픽모델링 결과에도 반영되어 있다. <Topic 1: 환자 주권과 생명의 중요성>, <Topic 2: 건강보험 급여화와 환자피해 최소화>, <Topic 3: 수술실 CCTV 설치 법제화>, <Topic 4: 의료사고 예방> 등 4개의 토픽은 <Topic 1: 환자주 권과 생명의 중요성>을 제외하고는 대부분 구체적 이고 현실적인 문제들과 관련이 있어 보인다.

반면 미디어의 보도기사는 앞서 살펴본 것처럼 빈도 분석에서 ‘병원’, ‘정부’, ‘수술’, ‘약’ 등 구체적 이슈를 지칭하기 보다는 일반적인 키워드가 자주 출현하는 것을 확인할 수 있었다. 토픽모델링 기법을 활용해 미디어의 보도기사에서 추출한 4개의 토픽 역시 <Topic 1: 병원의 잘못된 의료 관행>, <Topic 2: 환자 중심성>, <Topic 3: 환자 치료 접근 권>, <Topic 4: 환자의 인권과 생명의 중요성> 등으로, <Topic 1: 병원의 잘못된 의료 관행>을 제외 하고 나머지 3개 토픽의 경우 세부적이고 실천 지향적 테마보다는 인권, 권리, 생명, 환자 중심성 등 포괄적, 개념적인 테마라는 것을 확인할 수 있다. 다시 말해, 환자단체연합회의 경우 해결이 필요한 직면한 이슈 중심의 커뮤니케이션을 진행했다면, 미디어는 개별 이슈의 해결에 집중하기 보다는 개선되어야 할 거시적 방향성에 집중했다고 볼 수 있다. 그렇다고 환자단체연합회와 미디어가 서로 다른 토픽을 다루었다고 보기는 힘들다. 환자단체연합회의 보도자료에도 미디어의 보도기사가 다뤘던 거시적 의제인 ‘환자 중심성’, ‘환자치료 접근권’, ‘환자의 인권과 생명’이라는 테마 역시 담겨

있었을 것이다. 다만, 환자단체연합회의 보도자료에는 이런 거시적 의제보다는 실천적이고 현실적 의제가 더 부각되도록 내용이 기술되었을 것이라고 보는 것이 타당하겠다.

환자단체연합회는 미디어의 이러한 보도 성향을 이해하고, 보도자료 소재를 선정하고 작성하기 전에 거시적인 의제까지 염두에 둔 기획을 진행한다면, 미디어에 중요한 의제로 인식되어 보도기사로 선택될 가능성을 높일 수 있을 것이다. 미디어 역시 환자단체의 커뮤니케이션 목적이 환자들의 직면한 현실적 어려움의 시급한 해결에 있기도 하지만, 그 해법이 일회성에 그치기보다는 법제화를 통해 사회에 제도적으로 정착 시키는 것이라는 점을 인식한다면 환자들의 목소리를 보도기사에 반영하는데 도움이 될 수 있을 것이다. 의료 시장에서 환자와 의사는 서로 공생과 협력할 수밖에 없는 가장 중요한 파트너라고 할 수 있다. 또한 소비자 주권 시대를 맞아 의료 소비자인 환자의 목소리를 경청하고 이들을 이해하려 노력하는 것은 의료 서비스 발전의 시작이자 반드시 필요한 과정이기도 하다. 그런 의미에서 의료계는 환자단체를 서로 대립하는 적대적 이해관계자가 아닌, 의료계를 함께 발전시킬 수 있는 파트너로 인정하는 열린 마음을 가질 필요가 있겠다. 물론 이러한 태도는 환자단체에도 해당하는 자세일 것이다. 마지막으로 정부 및 정부 관련 부처에게 있어 환자와 미디어의 의제화 방향을 파악할 수 있다는 것은 정책 수요를 올바르게 파악해 정책 개발에 적극 반영할 수 있는 기회가 될 것이며, 국회는 환자단체가 의료 현장에서

겪고 있는 어려움을 입법화함으로써 환자 중심의 체계를 마련하는 데 도움을 줄 수 있을 것이다.

4. 연구의 한계와 후속 연구

본 연구는 분석 대상 미디어를 13개 일간지 및 공중과 방송으로 한정하였다. 주지하다시피, 환자단체연합회가 미디어에 배포한 보도자료 262건에 대해 13개 미디어에서 보도한 보도량이 210건이라는 것은 종합지와 공중과 TV의 영향력과 접근의 어려움을 고려한다고 하더라도 많은 보도량은 아니라고 할 수 있다. 이는 종합지와 공중과 TV의 한정된 공간과 시간 때문일 수도 있지만, 환자 관련 이슈들이 대부분 전문적인 지식을 필요로 하거나, 대중적이기 보다는 특수한 지식의 영역에 속하기 때문일 수도 있다. 따라서 향후 추가 연구에는 환자 관련 이슈에 대한 높은 이해도와 관심을 갖고 있는 의학 전문 미디어를 포함한 대규모 연구가 진행될 필요가 있겠다.

또한, 본 연구가 한국언론재단의 빅인즈 검색 프로그램에서 '환자단체연합회'를 키워드로 검색해 관련 기사를 추출해 낸 것이어서, 환자와 관련된 미디어의 모든 보도를 포함했다고 하기는 어렵다. 미디어는 환자단체연합회가 제기한 이슈 외에도 자체적으로 다양한 환자 관련 이슈를 의제화했을 것이다. 환자단체연합회 역시 환자와 관련된 모든 이슈를 보도자료화 할 수 없었을 것이라는 점에서 연구에 활용한 데이터가 환자와 관련된 모든 내용을 포괄하고 있다고 보기 어려울 수 있다.

하지만, 환자단체연합회가 국내 유일의 환자단체 연합체로 오랜 기간 환자와 관련된 대부분의 중요한 이슈들에 대해 적극적으로 미디어에 입장과 주장을 개진해왔다는 점에서 일부 누락된 부분이 있을 수 있지만 연구의 결과에 영향을 미칠 수 있는 정도는 아닐 것이라 판단된다.

마지막으로 본 연구에서 핵심적으로 사용한 텍스트 마이닝 기법 중 하나인 LDA은 뚜렷한 분석적 장점에도 불구하고 기술적인 한계도 존재한다고 할 수 있다. LDA를 활용한 텍스트 분석은 문장의 주술관계나 형용사, 부사와 같은 난해한 키워드 분석보다는 명사 의존적이라는 점에서 보도자료와 보도기사의 감성까지 잡아 낼 수 없는 기술적 한계를 인정할 수밖에 없다(박승정 등, 2017). 그럼에도 불구하고 기존 내용분석의 한계를 어느 정도 극복하고 객관성과 과학성을 담보해 내고 있다는 점에서 여전히 의미 있다고 하겠다.

참고문헌

- 권나현, 최윤형 (2022). 코로나19 확산 시기별 대통령 연설문에 나타난 위기 커뮤니케이션 분석: 구조적 토픽 모델링과 감성분석을 중심으로. *광고PR실학연구*, 15(2), 7~41.
- 권순만 (1999). 일간신문의 병원과 관련된 기사 분석. *한국병원경영학회지*, 4(2), 60~84.
- 김동석, 유현재 (2022). 국내 의료 환경에 대한 환자단체의 '문제인식'과 '제약인식'에 대한 연구: 한국환자단체연합회 창립 이후 배포된 보도자료 분석을 중심으로. *PR연구*, 20(4), 25~49. doi:10.15814/jpr.2022.26.4.25
- 김방현 (2001). *보도자료 분석을 통한 행정홍보의 문제점에 관한 연구*. 석사학위논문, 한남대학교 지역개발대학원.
- 김병진 (2000). *정책학 개론*. 서울: 박영사.
- 김선영 (2022). 새마을운동 관련 사회적 이슈 탐색 및 의미에 관한 연구: 뉴스 빅데이터의 LDA기반 토픽분석을 중심으로. *사회적경제와 정책연구*, 12(2), 151~178. doi:10.22340/seps.2022.05.12.2.151
- 김영희 (2019). *환자운동을 통한 환자안전법(중환이법) 제정 과정 연구*. 석사학위논문, 대구가톨릭대학교 대학원.
- 김용 (2014). *한국 의료커뮤니케이션의 변화과정 연구: 의료서비스의 환경과 활동을 중심으로*. 박사학위논문, 고려대학교 대학원.
- 김은혜, 박주호 (2022). 교육부 대입 정책 보도자료(2003-2021) 기반 토픽 모델링을 통한 정권별 특징 분석. *열린교육연구*, 30(2), 63~89. doi:10.18230/tjye.2022.30.2.63
- 김재희 (2009). *의학관련 보도의 프레임 연구*. 석사학위논문, 이화여자대학교 대학원.
- 김지현 (2022). 한국 언론이 코로나19 보도자료를 활용하는 방식: 조선일보와 한겨레의 '사회적 거리두기' 기사 분석. *정보사회와 미디어*, 23(1), 1~39. doi:10.52558/ISM.2022.04.23.1.1
- 김진환 (2003). *보도자료의 기사화에 관한 연구*. 석사학위논문, 중앙대학교 신문방송대학원.
- 남춘호 (2016). 일기자료 연구에서 토픽모델링 기법의 활용가능성 검토. *비교문화연구*, 22(1), 89~135.
- 박승정, 전진오, 김선우, 김성태 (2017). 국내 주요일간지의 대통령 이슈소유권에 대한 빅데이터 분석. *정치정보연구*, 20(3), 25~55. doi:10.15617/psc.2017.10.31.3.25
- 박치성, 명성준 (2009). 정책의제 설정과정에 있어 인터넷의 역할에 관한 탐색적 연구. *한국정책학회보*, 18(3), 41~69.
- 박현정, 김한나, 홍유정 (2017). 토픽모델링을 활용한 학생인권조례의 사회적 이슈 분석. *아시아교육연구*, 18(4), 683~711. doi:10.15753/aje.2017.12.18.4.683
- 반현, Maxwell, E. McCombs (2007). 의제설정 이론의 재고찰. *커뮤니케이션 이론*, 3(2), 7~53.
- 방성현, 이건호 (2013). 일간지 기사와 보도자료 프레임 비교를 통한 비구성적 현실 탐색: 4대강 사업 사례분석. *한국언론학보*, 57(1), 163~186.
- 백수미, 문인오 (2021). 텍스트마이닝 및 CONCOR 분석을 활용한 환자안전문화 융복합 연구주제 분석. *디지털융*

- 복합연구, 19(12), 359~367. doi:10.14400/JDC.2021.19.12.359
- 서병호, 김춘식 (2001). 정부의 대인론 홍보에 관한 연구: 재정경제부의 보도자료 분석과 평가를 중심으로. *한국언론학보*, 45(2), 216~249.
- 서종근 (2021). *빅데이터를 활용한 노인장기요양보험 키워드 네트워크 및 LDA 분석*. 석사학위논문, 연세대학교 보건대학원. doi:10.21032/jhis.2021.46.3.257
- 신기영, 이종혁 (2013). 소셜 네트워크 분석을 이용한 인사채용 검증기법. *정보과학회논문지: 데이터베이스*, 40(6), 370~376.
- 안예지, 김동일 (2020). 발달장애 학생 언론 보도기사의 주제어 및 토픽분석. *특수교육학연구*, 54(4), 27~50. doi:10.15861/kjse.2020.54.4.27
- 유주영 (2022). 학생인권에 관한 사회적 이슈 분석: 토픽모델링을 활용한 국내 언론 보도기사 분석을 중심으로. *교육학연구*, 60(1), 205~234. doi:10.30916/ker.60.1.205
- 이수상 (2016). 독후감 텍스트의 토픽모델링 적용에 관한 탐색적 연구. *한국도서관정보학회지*, 47(4), 1~18. doi:10.16981/kliiss.47.4.201612.1
- 이시환, 조아람, 이훈영 (2017). 온라인 병원 리뷰자료의 Latent Dirichlet Allocation 분석을 활용한 의료서비스 만족 요인에 관한 연구. *서비스경영학회지*, 18(5), 23~44. doi:10.15706/jksms.2017.18.5.002
- 이완수, 김찬석, 이민규 (2012). 국내 입법 커뮤니케이션의 구조적 특징. *한국언론정보학보*, (60), 52~74.
- 이완수, 손영준 (2011). 북한 핵실험 이슈에 대한 언론의 의제 구성. *한국언론정보학보*, (56), 175~193.
- 이정희 (2015). *인터넷 신문의 정책커뮤니케이션 기능과 사회 아젠다 진화과정 분석*. 석사학위논문, 한국과학기술대학교 대학원.
- 이희성 (2020). Ctrl+C, Ctrl+V, 국내 온·오프라인 신문의 처널리즘 분석: 카피킬러를 이용한 보도자료와의 문장 일치도 분석을 중심으로. *사회과학연구*, 27(4), 61~79. doi:10.46415/jss.2020.09.27.4.61
- 임현수, 이준웅 (2011). 보도자료 기사화 과정에서의 영향요인에 관한 연구: 정부 보도자료에 대한 조선일보, 한겨레 기사 분석을 중심으로. *한국언론학보*, 55(2), 6~31.
- 진명지, 이서현, 최낙진 (2020). '코로나19' 관련 보건복지부 보도자료의 언론보도 활용에 관한 연구. *지역과 세계*, 44(4), 125~154.
- 차희원 (2001). 시민단체(NGO)와의 바람직한 관계형성을 위한 쌍방균형적 홍보전략에 관한 연구. *홍보학연구*, 5(1), 133~174.
- 차희원 (2001). *시민단체와 대응 조직체간 갈등해소를 위한 승승적 PR 전략에 관한 연구: 시민단체의 커뮤니케이션 선택성과 갈등결과, 그리고 영향요인을 중심으로*. 박사학위논문, 이화여자대학교 대학원.
- 한국환자단체연합회 (2022). *단체소개*. Retrieved 8/1/2022 from <http://www.koreapatient.com/ab-1716>
- 홍미현 (2004). *보도 자료가 기사화에 미치는 영향 및 의존도 연구-중앙일간지 온라인제임 담당기자를 대상으로*

. 석사학위논문, 동국대학교 대학원.

- 홍주현, 김경희 (2017). 언론의 정치적 성향이 뉴미디어 정책 관련 사설의 의제 및 보도 태도에 미치는 영향. *한국콘텐츠학회논문지*, 17(3), 162~177. doi:10.5392/JKCA.2017.17.03.162
- 황만성 (2005). 의료인과 환자간의 갈등해소를 위한 법적 제도적 방안. *형사정책연구원 연구총서*, 13~200.
- Blei, D. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77~84. doi:10.1145/2133806.2133826
- Blei, D., Ng, A., & Jordan, M. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993~1022.
- Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), P10008. doi:10.1088/1742-5468/2008/10/P10008
- Boumans, J. (2018). Subsidizing the news? Organizational press releases' influence on news media's agenda and content. *Journalism Studies*, 19(15), 2264~2282. doi:10.1080/1461670X.2017.1338154
- Chakraborty, G., & Pagolu, M. K. (2014). Analysis of unstructured data : Applications of text analytics and sentiment mining. *SAS Global Forum*, 288~2014.
- Clifton, C., Cooley, R., & Rennie, J. (2004). Topcat: Data mining for topic identification in a text corpus. *IEEE Transactions on Knowledge and Data Engineering*, 16(8), 949~964. doi:10.1109/TKDE.2004.32
- Cobb, R. W., & Elder, C. D. (1971). The politics of agenda-building: An alternative perspective for modern democratic theory. *The Journal of Politics*, 33, 892-915. doi:10.2307/2128415
- Cutlip, S. M., & Allen, H. C. (1978). *Effective public relations*. Englewood Cliffs, NJ: Prentice-Hall, p. 31. doi:10.1080/00913367.1979.10717979
- DiMaggio, N., & Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding. *Poetics*, 41(6), 570~606. doi:10.1016/j.poetic.2013.08.004
- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51~58. doi:10.1111/j.1460-2466.1993.tb01304.x
- Gitlin, T. (1980). *The whole world is watching*. Berkely: University of California Press. doi:10.1017/S0084255900042169
- Hearst, M. A. (1999, June). Untangling text data mining. *In Proceedings of the 37th Annual meeting of the Association for Computational Linguistics*, College Park, Maryland. doi:10.3115/1034678.1034679
- Jin, X. (2021). Understanding social-mediated disaster and risk communication with topic model. In R. Djalante,

- & R. Shaw (Eds.), *Integrated research on disaster risks* (Chapter 19). Nature Switzerland AG: Springer. doi:10.1007/s11948-013-9502-z
- Kiousis, S., Kim, S. Y., McDevitt, M., & Ostrowski, A. (2009). Competing for attention: Information subsidy influence in agenda building during election campaigns. *J & MC Quarterly*, 8(3), 545~562. doi:10.1177/107769900908600
- Lambiotte, R., Delvenne, J. C., & Barahona, M. (2008). *Laplacian dynamics and multiscale modular structure in networks*. arXiv preprint arXiv:0812.1770. doi:10.1109/TNSE.2015.2391998
- Lang, G. E., & Lang, K. (1981). Watergate: An exploration of the agenda-building process. In G. C. Wilhoit, & H. De Bock (Eds.), *Mass communication review yearbook* (pp. 447~468). Beverly Hills, CA: Sage.
- Lee, S. H., & Choi, J. B. (2017). How was the local government's press release reflected in the national and local newspaper?: Focusing on comparison of frames used on the heavy snowfall disaster in Jeju. *Journal of Communication Science*, 17(1), 28~55. doi:10.14696/jcs.2017.03.17.1.28
- McCombs, M. E. (1997). New frontiers in agenda setting: Agendas of attributes and frames. *Mass Communication Review*, 24(1/2), 32~52.
- McCombs, M. E. (2004). *Setting the agenda: The mass media and public agenda*. Malden, MA: Blackwell. doi:10.1093/poq/nfj003
- McCombs, M. E. (2014). *Setting the agenda: Mass media and public opinion*. Hoboken, New Jersey: John Wiley & Sons.
- McCombs, M. E., & Donald, L. S. (1972). The agenda-setting function of mass media. *POQ*, 3(2). doi:10.1086/267990
- Pan, Z., & Kosicki, G. M. (1993). Framing analysis: An approach to news discourse. *Political Communication*, 10(1), 55~75. doi:10.1080/10584609.1993.9962963
- Roloff, M. E. (1987). Communication and conflict. In C. R. Berger, & S. H. Chaffe (Eds.), *Handbook 01 communication science* (pp. 484~534). CA: Sage.
- Turk, J. V. (1986). Information subsidies and media content: A study of public relations influence on the news. *Journalism Monograph*, 100.
- Wagner, J. A., & Hollenbeck, J. R. (2020). *Organizational behavior: Securing competitive advantage* (Eds.). New York, NY: Routledge. doi:10.4324/9781003009580
- Weaver, D. E., & Ellitt, S. N. (1985). Who sets the agenda for the media: A study of local agenda-building. *Journalism Quarterly*, 62(1), 87~94. doi:10.1177/107769908506200

Abstract

Analysis of Major Agenda Differences between Press Releases by Patient Groups and Media Articles: Focusing on Text Mining Analysis

Kim, Dong Seok

PhD Candidate, School of Communication, Sogang University

Yoo, Hyun Jae

Professor, College of Media, Arts and Science, Sogang University

Kang, Hyun Woo

Head, Health Insight Center, Enzaim Health

Hwang, Sun Wook

President, Data Classic

Son Su Ji

MA Candidate, Graduate School of Management of Technology, Pukyong National University

The domestic patient rights movement, which began in the 2000s, is accelerating further with the formation of a large coalition in 2010 of patient groups called the Korea Alliance of Patient Organization. As a result, the organization's media activities, such as distribution of press releases, and the coverage of those activities by the media now play an important role in informing our society of the various issues affecting patients. However, the intent of those press releases and contents of media articles covering them are bound to differ in that the media selects only specific issues they want to cover from the content provided in the press releases. Therefore, in this study, 262 press releases distributed over the 12 years since the foundation of the Korea Alliance of Patient Organization and 210 press articles published by 13 media were analyzed to determine differences. To this end, this study employed the Topic Modeling technique, a big data research method and differences were identified in their purpose, attitude and approach. This study suggest patient organizations' media communication strategies, and help the public to better understand patients, patient groups, and the media.

keywords: Patient, Patients group, Agenda setting, Big data, Text mining