

빅데이터 분석을 이용한 세종시 건설 계획에 관한 여론 변화 A Changes of Opinion according to the Sejong City Construction Plan Using Media Big Data Analysis

조성수, 이상호
국립한밭대학교 도시공학과

Sung Su Jo(josungsu85@daum.net), Sang Ho Lee(Lshsw@hanbat.ac.kr)

요약

본 연구의 목적은 빅데이터 분석 기법을 이용하여 세종시 건설 계획에 대한 여론 변화를 분석하는 것이다. 연구 자료는 세종시 건설 계획과 관련된 한겨레와 동아일보, 한국일보 등 언론사의 신문기사이다. 세종시 건설 계획은 신행정수도, 행정중심복합도시, 세종시 수정안 등 3개 시기로 구분되었다. 분석 방법은 파이썬과 Gephi 0.9.2 프로그램을 활용한 빈도분석, 감성분석, 사회연결망분석이다.

세종시 건설에 따라 각 신문사의 쟁점에 대한 변화가 있었다. 첫째, 빈도분석 결과, 한겨레의 키워드는 세종시 건설 시기에 따라 찬성-찬성-반대의 특성을 나타내고 있었다. 동아일보는 반대-반대-찬성의 입장을 갖는 것으로 분석되었다. 한국일보는 반대-찬성-반대 경향을 갖는 것으로 분석되었다. 둘째, 감성분석 결과, 각 언론사는 논조의 특성이 변화되고 있었다. 한겨레는 긍정-긍정-부정 논조 특성을 보이고 있었고, 동아일보는 부정 - 부정 - 긍정으로 변화되었다. 한국일보는 부정-긍정-부정으로 변화되었다. 셋째, 사회연결망분석 결과는 다음과 같다. 각 언론사는 세종시 건설 시기에 진보와 보수, 중도 등 정치적, 이념적 특성에 따라 변화를 보이고 있었다. 한겨레는 지역균형발전에 초점을 맞추고 있었으며, 동아일보는 보수의 의견을 대변하고 있었다. 한국일보는 세종시 건설 시기에 여·야가 대립하는 이슈에 대해 부각 시키고 있었다.

■ 중심어 : | 빅데이터 | 감성분석 | 사회연결망분석 | 세종시 건설 | 여론 변화 |

Abstract

This study aims to analyze on the changes of opinion in terms of Sejong City construction using big data. The research data are newspaper articles related to the argument of construction in Sejong City. The newspaper article data was reported by Hankyoreh, Dong-A Ilbo and Hankook Ilbo. The arguments related to the construction of Sejong City was included the new administrative capita, multifunctional administrative city and amendments of Sejong City. The analysis method used in this study is frequency analysis, sentiment analysis and social network analysis using python and gephi 0.9.2 programs.

The results of the analysis are as follows. First, as a result of frequency analysis, the keywords of Hankyoreh showed the characteristics of consent - consent - dissent according to the construction period of Sejong City. The Dong-A Ilbo showed positions of dissent - dissent - consent. In addition, the Hankook Ilbo was analyzed to have the characteristics of dissent - consent - dissent tendency. Secondly, results of sentiment analysis, The Hankyoreh showed positive - positive - negative tone. The characteristic of Dong-A Ilbo is that the focus has changed from negative to negative to positive. The Hankook Ilbo showed that changed from negative to positive to negative. Finally, the results of social network analysis are as follows. At the time of the construction of Sejong City, each opinion of media was showed a changes in issues according to political and ideological characteristics such as conservative, progressive and moderation. In detail, Hankyoreh focused on balanced regional development. The Dong-A Ilbo represented the opinion of the Conservative Party. The Hankook Ilbo was highlighting the issues confronting the conservative party and progressive party during the construction of Sejong City.

■ keyword : | Big Data | Sentiment Analysis | Social Network Analysis | Sejong City Construction | Opinion Changes |

* 이 논문은 2019학년도 국립한밭대학교 교내학술연구비의 지원을 받았음

접수일자 : 2020년 06월 01일
수정일자 : 2020년 06월 30일

심사완료일 : 2020년 06월 30일
교신저자 : 이상호, e-mail : Lshsw@hanbat.ac.kr

I. 서론

역사적으로 지역에 기반을 둔 정치 세력간 갈등은 국내외를 막론하고 끊임없이 나타났다. 중국의 정치수도 베이징방(北京幫)과 경제수도 상하이방(上海幫) 그리고 한족의 고도 난징 사이에 정치적 갈등이 존재한다. 미국은 중부 대평원 중심의 공화당과 북동부 오대호에 기반을 둔 민주당의 정치적 갈등이 내재하고 있다[1].

역사적으로 보면, 공간과 정치는 시대를 막론하고 끊임없는 갈등을 자아낸다. 특히, 수도(首都) 이전에 대한 갈등은 첨예하다. 대표적으로 묘청의 서경천도운동과 노무현 전 대통령의 신행정수도 이전은 정치적 갈등이 존재하였다. 신재호 선생의 「조선사 연구초」는 묘청의 천도를 다음과 같이 평하였다. ‘서경천도는 낭·불 양가대 유가의 싸움이며, 국풍과 대 한학파의 싸움이고, 진취사상 대 보수사상의 싸움이니, 표청은 곧 전자의 대표요, 김부식은 후자의 대표였던 것이다.’

노무현 전 대통령의 신행정수도 이전은 중앙부처를 옮겨서 국토의 균형발전을 이루고, 수도권과 지방의 격차를 완화한다는 정책으로부터 시작되었다. 그러나, 당시 여당인 한나라당은 수도권 공동화론, 지역간 불균형 역(逆)침화 현상, 수도서울의 기능적 효율성 하락 등의 논리를 중심으로 신행정수도 이전에 반대하였다[2]. 신행정수도의 갈등은 정치적 갈등에서 지역적 갈등으로 확산되었다. 서울 및 수도권과 영남권, 강원권 등 지역과 호남권 및 충청권 등의 지역이 첨예한 대립을 보였다[3].

세종시 건설은 정치적, 지역적 갈등뿐만 아니라, 언론, 학자 등에서 찬성과 반대의 양면적 가치가 충돌하였다[3]. 특히, 유명돈, 마정미(2015)의 연구에서는 초기 신행정수도 추진이 일부 보수 언론의 극단적인 보도(‘헌법소원 움직임’, ‘천도론’, ‘국민투표 논란’ 등)로 인해 위기에 몰렸다고 주장하였다[4]. 이현숙(2013)은 세종시 이슈에 대해 보수와 진보 언론사이에, 중앙언론과 지역언론 사이에 서로 다른 보도 양상을 보이는 것에 주목하였다[5].

이처럼 세종시 건설과 관련되어 보수와 진보 신문사간 비교연구가 활발하게 진행되었다. 그러나 대부분의 연구는 직접 기사를 분석하고 직접 판단해내는 작업이

수반되었다. 이는 정확한 결과를 낼 수 있었겠으나, 데이터의 양적 분석 측면과 질적 분석 측면에서 한계가 있다. 연구자의 직접코딩 및 분석으로 연구자의 생각이 반영될 위험성 그리고 빅데이터 기사에 대한 소화 불가능이 바로 그것이다. 빅데이터 분석기법을 활용한다면, 기존 연구의 한계를 극복하여 신문기사 내용을 다각도로 조명할 수 있다[6]. 따라서 본 연구는 세종시 건설과 관련된 여론이 시기별로 어떻게 변화하였는지 빅데이터를 이용하여 분석하였다.

본 연구는 다음과 같은 연구 절차에 따라 수행되었다. 첫째, 기존 문헌을 고찰하고 연구의 범위를 신행정수도, 행정중심복합도시, 세종시 수정안 등 세종시 건설 기간 전체로 설정하였다. 둘째, 진보지, 보수지, 중도지 언론의 내용 및 논조 변화를 살펴보기 위해 한겨레, 동아일보, 한국일보 등 세 개 신문사를 선정하였다. 셋째, 빅데이터 분석에서 활용되는 파이썬 프로그램을 통해 뉴스기사 데이터 수집 및 자연어 처리를 진행하였다. 넷째, 언론사의 찬성 및 반대 초점을 분석하기 위해 빈도분석(Frequency Analysis)이 수행되었다. 다섯째, 언론의 논조 변화를 알아보기로자 감성분석(Sentiment Analysis)이 진행되었다. 다섯째, 언론 신문기사의 내용적 의미 구조를 분석하기 위해 사회연결망분석(Social Network Analysis)이 활용되었다.

II. 문헌고찰

1. 세종시 건설의 추진

세종시 건설은 2002년 당시 노무현 대통령 후보가 국가 균형발전을 위해 정부가 출선해야 한다는 점을 강조하며 제시된 공약이다. 당선 이후, 노무현 대통령은 2003년 12월 「신행정수도 건설 특별 조치법」(이하, 신행정수도법)을 국회에 가결 시켰으며, 2004년 5월 ‘신행정수도 건설추진위원회’를 출범시켰다. 일부 언론과 전문가들의 비판을 받기도 하였으나 신행정수도 건설은 대체로 순조롭게 진행되었다(권오혁, 2012). 2004년 08월에는 지역의 정책 전문가들이 참여하여 입지까지 선정을 완료하였다. 그러나 2004년 6월에 신행정수도법은 헌법재판소에 소원이 제기되었으며, 관습헌법에

위배된다는 판결(2004년 10월)을 받아 신행정수도 건설사업은 전면 중단되었다.

신행정수도 대안으로 2005년 1월 자족성을 갖춘 행정중심복합도시건설에 여야가 합의하였다. 이후, 2005년 3월 「신행정수도 후속대책을 위한 연기·공주지역 행정중심복합도시 건설을 위한 특별법」(이하, 행복도시법)이 국회에 가결되었다. 당시 노무현 대통령은 2005년 4월 ‘행정중심복합도시 건설추진위원회’를 공식 출범시켰다. 행복도시법은 수도 전체를 이전하는 대신 정부 행정기능의 일부인 12부 4처 2청 만을 이전하는 것이 핵심내용이다. 2005년 6월, 행복도시법은 헌법재판소에 헌법소원이 제기됐으나 당해 11월 각하되어 행복도시를 둘러싼 논란은 일단락되었다. 2006년 1월 ‘행정중심복합도시건설청’이 공식적으로 출범되었으며, 이후 도시기본계획(안)이 발표되었고, 도시의 명칭은 세종시로 정해졌다.

2008년 정권이 교체되면서 지난날 여야합의를 이끌어 추진되고 있던 행복도시시는 다시 한번 위기를 맞게 되었다. 당시 이명박 정부의 국무총리인 정운찬 총리가 2010년 1월 ‘세종시 수정안’을 발표하였다. 세종시 수정안은 기업과 과학문화가 복합된 교육과학 중심의 경제도시 건설이 핵심내용이다. 이로 인해, 여야가 대립하였으며, 여당 내에서도 친이명박계와 친박근혜계가 대립하였다[2]. 세종시 수정안 계획과 세종시 원안 계획 내용이 대립한 것이다. 이후 2010년 6월 세종시 수정안은 국회에서 부결되었다.

이처럼 세종시 건설은 크게 신행정수도, 행정중심복합도시, 세종시 수정안 등 세 가지 계획으로 구분된다. 따라서, 본 연구의 범위는 신행정수도, 행정중심복합도시, 세종시 수정안이다. 이는 세종시 건설이 진행되는 가운데 갈등이 증폭되었던 시기로서, 신행정수도법 통과 및 시행(2003년 12월 29일 ~ 2004년 4월 17일), 행복도시법 공포 및 시행(2004년 11월 15일 ~ 2005년 3월 18일), 세종시 수정안 발표 및 국회 부결 전(2010년 1월 11일 ~ 2010년 5월 29)으로 분류될 수 있다.

2. 세종시 건설 이슈 연구

세종시 건설 이슈 연구는 신행정수도로 인해 갈등이

표출되었던 2004년부터 진행되었다. 김춘식(2004)은 신행정수도와 관련된 이슈에 대해 조선일보, 중앙일보, 한겨레 등 9개 중앙일간지에서 2004년 7월 1일부터 7월 10일까지 보도된 기사를 중심으로 분석하였다. 기사 내용은 부정, 중립, 긍정으로 나누어 분석되었다. 분석 결과는 신행정수도와 관련된 기사는 부정적인 논조가 절반 이상(약 57%)을 차지하고 있었으며, 중립적 논조는 약 34%, 긍정적 논조는 약 9%로 분석되었다[7].

유승현, 황상재(2006)는 포털 다음 미디어와 조선일보, 한겨레를 중심으로 신행정수도 위헌에 대해 보도경향 프레임을 분석하였다. 기간은 신행정수도 위헌 기간 전·후로, 2004년 10월 1일부터 10월 31일까지 선정되었다. 분석 결과, 신행정수도 위헌에 대해 조선일보는 긍정적인 보도 태도를 보이고 있었으며, 한겨레와 포털 다음 미디어는 부정적인 보도 태도를 나타내고 있었다[8].

조용구(2012)는 신행정수도, 행정중심복합도시, 세종시 수정안 기간에 나타난 세종시 건설 이슈를 중심으로 찬반입장 변화를 분석하였다. 찬반입장 변화의 대상은 정부, 언론사 전문가·정당, 지역이다. 분석 결과, 세 가지 이슈에 모두 찬성을 보인 언론사는 한겨레, 대전일보 등으로 나타났으며, 반대입장을 보인 언론사는 조선일보, 중앙일보, 동아일보로 분석되었다. 전문가는 지역 소재의 연구원이 찬성을 수도권 소재의 연구원과 국가산하 기관 연구원이 반대입장을 갖는 것으로 나타났다. 지역은 충청권과 호남이 찬성을 표명하고 수도권과 영남, 강원권이 반대의 입장을 표명한 것으로 분석되었다[3].

박승민(2011)은 세종시 수정안으로 논란이 이슈로 부각되었던 2009년 9월 4일부터 2010년 6월 30일까지 KBS·MBC의 전국과 지방간 보도 TV뉴스 프레임을 비교 분석하였다. 분석결과, KBS·MBC 전국 보도는 세종시 수정안 찬성 보도가 많았으며, KBS·MBC의 대전·충남권은 반대의 보도가 더 많은 것으로 분석되었다[9]. 이현숙(2013)의 연구에서는 세종시 수정안 이슈 기간(2012년 9월 ~ 2013년 3월)에 동아일보, 경향신문, 대전일보, 중도일보 기사 내용을 분석하였다. 분석결과, 동아일보가 가장 부정적인 논조를 보이고 있는 것으로 나타났다. 경향신문은 부정과 긍정 논조를 동시에 가지고 있었다. 또한, 대전일보는 중립적인 논조를 보였으며, 중도일보는 긍정적인 논조를 갖는 것으로 분석되었

다[5].

유영돈, 마정미(2015a, 2015b)는 신행정수도, 행정 중심복합도시, 세종시 키워드를 중심으로 조선일보, 경향신문, 대전일보, 중도일보, 전북일보, 영남일보, 강원도민일보 등 전국 7개 일간지 기사를 수집하여 논조 및 뉴스 프레임 분석하였다. 분석 결과, 대부분 신문사가 중립적 논조를 지향하고 있다고 밝혔다. 그 중, 전반적으로 긍정적인 논조를 갖는 신문사는 대전일보, 중도일보, 전북일보로 나타났으며, 부정적인 논조를 갖는 신문사는 조선일보, 영남일보로 분석되었다. 특히, 중앙지와 지역지를 구분하지 않고 언론이 사회 통합자로서의 기능 보다 신문사의 내부 규범, 이데올로기, 정치적 성향, 지역적 편향성, 자본 논리의 기능 등이 더 강화되고 있음을 분석결과를 통해 보여주고 있었다[4][10].

세종시 건설과 관련된 선행연구는 시기의 분류, 신문 기사 자료를 통한 변화 분석 등 본 연구와 내용적인 측면에서 유사한 전개를 보이고 있었다. 그러나 몇 가지 한계점이 존재한다. 첫째, 기존의 연구는 분석되는 데이터 규모와 취득에 한계가 존재하였다. 둘째, 자료를 직접 코딩함으로써 신뢰도에 한계가 있었다. 셋째, 키워드 분류 및 연구내용의 해석에 있어 연구자의 주관성을 배제하기 어려운 현실적 제한이 존재하고 있었다[11]. 넷째, 대부분의 연구는 진보와 보수의 보도 태도, 프레임 등에 초점이 맞춰져 분석이 진행되었다.

3. 빅데이터 분석 방법론

빅데이터 분석은 비정형 빅데이터로부터 의미 있는 정보를 추출하는 방식으로 특정 단어의 빈도 및 연관성을 파악하여 키워드간 관계를 분석하는 방법이다. 빅데이터 기법을 활용하여 언론사별 논조 및 내용을 분석한 연구는 다양하게 수행되었다. 감미아, 송민(2012)의 연구는 한겨레, 동아일보, 경향신문 등 신문사별로 논조에 차이가 있음을 밝히기 위해 사회(4대강 사업), 국제(북한 핵), 정치(통합진보당) 등 6가지 분야 주제를 분석하였다. 단순빈도와 SNA를 활용한 클러스터링 분석기법이 활용되었다. 분석결과, 6가지 주제에서 경향신문과 한겨레는 유사한 키워드로 구성되어 있었다. 반면에, 동아일보는 이전 두 신문사와 다른 키워드로 구성되어 있었으며, 같은 현상을 다르게 바라보고 있는 것으로 분

석되었다[6].

양명철, 한혜경(2014)은 조선일보와 한겨레를 중심으로 이주민 보도에 대한 프레임 분석을 SNA 기법을 활용하여 시도하였다. 분석결과 조선일보는 이주민 외국인 노동자의 현상적 측면에 대한 사실을 전달하고 있었다. 한겨레 신문은 외국인 노동자의 현실과 정부의 정책, 제도적 측면 등을 포괄하여 보도하고 있는 것으로 분석되었다[12]. 김대옥, 최명일(2016)은 SNA를 활용하여 조선일보와 한겨레의 자살문제 이슈 보도에 대한 경향을 분석하였다. 분석결과, 각 신문사는 정치적인 성향에 따라 자살문제에 대해 다르게 접근하는 것이 확인되었다. 한겨레는 다양한 계층의 문제 또는 사회문제의 시각에 비중을 두었으며, 조선일보는 개인의 문제에 초점을 맞추어 보도하는 경향이 있었다[13].

박승정 외(2017)는 조선일보와 한겨레를 중심으로 박근혜 전 대통령과 관련한 사실을 토픽모델링 방법론을 이용하여 분석하였다. 분석결과, 두 언론사는 시간이 지남에 따라 강조하는 이슈에 차이를 보이고 있었다. 조선일보는 정치적, 이념적 성향을 강조하고 있는 반면에 한겨레는 사회, 경제, 정치 부문의 이슈에 대해 문제를 제기하는 보도 경향을 보였다[14]. 최진호 외(2019)는 4차산업혁명에 대한 뉴스 데이터를 보수지, 진보지, 경제지로 나누어 토픽분석 및 SNA를 실시하였다. 분석결과 보수지와 경제지는 4차산업혁명 핵심기술에 초점을 맞추고 있었다. 진보지는 대선국면, 정부정책 등에 중점을 두고 있었으며, 경제지는 산업적 이슈를 비중 있게 보도하고 있었다[15].

빅데이터를 활용한 기존 연구는 언론의 내용을 분석하기 위해 빈도분석, 토픽모델링, SNA 방법을 중심으로 각각 수행되고 있었다. 빈도분석은 언론에서 다루고 있는 이슈의 사회적 초점의 변화를 볼 수 있는 일차적인 분석이고[16], 토픽모델링은 텍스트에서 도출된 단어를 중심으로 잠재적인 토픽을 추출하는 것이다[17]. SNA는 핵심 단어간 특정한 형태로 결합하여 의미 구조를 분석하는데 적합한 방법이다[18]. 기존 연구에서 이 방법들은 언론의 변화를 분석하는데 적절히 혼합하여 사용되었다. 그러나 언론의 논조를 가장 탁월하게 분석할 수 있는 감성분석[19]은 잘 활용되고 있지 않았다. 또한, 다양한 연구 주제가 있었으나 빅데이터 분석 기

법을 이용해 세종시 건설 계획의 이슈와 관련된 연구는 미흡한 것으로 나타났다.

문헌고찰에 기반하여, 본 연구의 차별성은 다음 몇 가지로 정리될 수 있다. 첫째, 신행정수도 건설부터 행정중심복합도시, 세종시 수정안까지 연구의 범위를 확대하였다. 둘째, 연구의 범위에서 진보지, 보수지 언론의 변화뿐만 아니라 중도지 언론의 변화를 관찰하기 위해 진보지인 한겨레, 보수지인 동아일보, 중도지인 한국일보를 선택하였다. 셋째, 분석 데이터의 규모와 취득의 한계를 극복하고자 언론사 홈페이지에서 직접 크롤링하였다. 넷째, 데이터의 직접 코딩 및 분석에 따른 연구자의 주관적인 생각이 포함될 수 있는 위험성을 배제하고자 빅데이터 분석 기법을 선택하였다. 다섯째, 언론사 뉴스 기사를 다각적, 심층적으로 분석하고자 빈도분석, 감성분석, 사회연결망분석 방법을 동시에 활용하였다.

III. 분석 자료 및 방법

1. 분석자료 수집

본 연구는 진보지 한겨레, 보수지 동아일보, 중도지 한국일보 등 세 종합신문에 나타난 세종시 건설과 관련된 신문기사 제목을 수집하였다. 세 신문사를 선택한 이유는 하나의 주제에 대해 정치적 이념에 따라 신문사 간 서로 다른 입장 및 논조를 보이며, 그 변화를 자세히 관찰할 수 있을 거라 판단되었기 때문이다[6][12-15]. 신문기사 제목을 분석 자료로 수집한 이유는 최소한의 표현으로 많은 정보의 정확한 전달을 하는 표시성이 있기 때문이다[20]. 또한, 제목은 기사의 내용 자체를 이해하는데 핵심적인 정보가 담겨있으며[21], 사회적 분위기나 신문의 논조가 제목을 통해 나타날 수 있기 때문이다[21].

자료수집 기간은 문헌고찰을 기반으로 신행정수도법 통과 및 시행(2003년 12월 29일 ~ 2004년 4월 17일), 행복도시법 공포 및 시행(2004년 11월 15일 ~ 2005년 3월 18일), 세종시 수정안 발표 및 국회 부결(2010년 1월 11일 ~ 2010년 5월 29일)이다. 한겨레와 한국일보의 신문기사 자료는 한국언론진흥재단에서 제공되는 빅카인즈(www.bigkinds.or.kr)에서 수집되었다. 동아

일보 자료는 빅카인즈에서 제공되지 않아 동아닷컴 홈페이지(donga.com)에서 파이썬을 이용하여 직접 크롤링(Crawling)하였다.

신문기사 검색 키워드는 신행정수도, 행정중심복합도시, 세종시 수정안이다. 신문기사 자료는 신행정수도에, 한겨레가 247건, 동아일보가 466건, 한국일보가 314건이 수집되었다. 행정중심복합도시 자료는 언론사 별로 86건, 68건, 90건이 획득되었으며, 세종시 수정안은 634건, 298건, 436건으로 수집되었다[표 1]. 같은 기간에 신행정수도 이슈에서는 동아일보가 가장 많은 기사(466건)를 냈으며, 한겨레는 가장 적은 기사(247건)를 보도하였다. 행복도시 이슈에서는 한국일보가 가장 많은 기사(90건)를 보도하였다. 세종시 수정안은 한겨레에서 634건이 기사화되었으며, 동아일보에서 298건이 보도되었다.

2. 분석 방법

본 연구의 분석 방법은 크게 4가지의 단계로 구성된다. 첫째는 수집된 데이터를 정제하는 과정인 데이터 전처리하는 단계이다. 둘째는 데이터 전처리를 통해 도출된 단어를 활용한 빈도분석 단계이다. 세번째 단계는 감정사전을 통해 단어의 긍정 및 부정을 분류하여 분석하는 감성분석이다. 넷째는 DTM(Document-Term Matrix)을 구축하고 단어간 상관관계를 통해 내용적 연결 의미를 확인하는 사회연결망분석 단계이다.

데이터 전처리는 파이썬 패키지(KoNLpy, 꼬꼬마 세종 말뭉치)를 이용하여 형태소 분석이 진행되었다. 기사 제목에서 내용과 관련이 없는 특수문자(예: ■, ►, ※ 등), 조사(예: 은, 는, 이, 가 등), 의존명사, 공백 등 불필요한 데이터는 제거되었다. 형태소 분석에서 제거

표 1. 이슈에 따른 언론사별 수집 자료 (단위 : 건)

이슈	검색기간	한겨레	동아일보	한국일보
신행정수도	2003.12.29. ~ 2004.4.17.	247	466	314
행정중심복합도시	2004.11.15. ~ 2005.3.18.	86	68	90
세종시 수정안	2010.1.11. ~ 2010.5.29.	634	298	436

되지 않은 불필요한 단어들(동아, 겨레, 일보 등)은 불용어 사전에 추가하여 직접 제거하는 과정을 반복하였다. 데이터 전처리로 얻어진 정제된 자료는 세종시 건설 이슈 기간에 언론사별 기사가 어디에 초점이 맞춰지고 있는지 확인하고자 빈도분석 수행에 사용되었다.

세종시 건설 계획 시기에 각 언론사가 어떤 논조를 보이는지를 분석하기 위해 감성분석이 진행되었다. 감성분석은 언론사에서 작성된 기사 제목의 단어들이 어떠한 논조를 갖는지 판단할 수 있는 방법이다. 감정의 분류를 위해 KNU 한국어 감성사전이 활용되었다. 본 연구에서는 전처리를 통해 추출된 단어에 감성사전을 기준으로 긍정 및 부정에 대해 점수를 부여하였다. 이를 기반으로 세종시 건설 계획에 대해 각 언론사가 긍정적인 논조를 보이는지 또는 부정적인 논조를 보이는지 분석되었다.

사회연결망분석은 단어빈도-역문서빈도(TF-IDF, Term Frequency-Inverse Document Frequency) 가중치를 적용한 DTM(Document-Term Matrix)이 활용되었다. TF-IDF는 특정 단어가 문서 안에서 얼마나 자주 나타나는지의 단어빈도(TF, Term Frequency)와 그 단어가 문서 내에서 얼마만큼 공통적으로 출현하는지를 나타내는 문서빈도(DF, Document Frequency)에 역문서(IDF, Inverse Document Frequency)를 곱한 값이다. 이 값이 크면 클수록 특정 단어가 전체 문서에서 적게 등장한다는 것을 의미한다[23][24].

네트워크로 표현되는 사회연결망분석은 Gephi 0.9.2 버전을 활용하여 시각화되었다. 네트워크 내에서 각 단어(노드)가 얼마만큼 중요한 역할을 하고 있는지 확인하기 위해 중심성(Centrality) 분석이 수행되었다. 중심성 분석은 사이중심성(Betweenness Centrality)과 고유벡터중심성(Eigenvector Centrality)을 활용하였다. 사이중심성은 하나의 단어(노드)가 중심이 되어 주변 단어(노드)들을 얼마만큼 강하게 연결을 시켜주는지에 대한 능력을 값으로 표현하는 것이다. 즉, 주변에 있는 다른 단어(노드)들을 연결하는 빈도가 높으면 높을수록 그 단어(노드)는 중요한 역할을 한다고 할 수 있다. 고유벡터중심성은 중요한 단어(노드)에 연결된 단어(노드)일수록 그 중요도가 높은 단어(노드)라는 전체로 수렴된 중심성 지수이다. 즉, 덜 중요한 단어(노드)

보다 중요한 단어(노드)에 연결이 되면 고유벡터중심성 값은 커진다.

IV. 분석 결과

1. 빈도분석 결과

대규모의 빅데이터를 통해 가장 단순하면서도 유용한 정보를 획득하는 방법은 각 언론사에서 가장 많이 언급된 키워드를 분석하는 것이다. 다수 언급된 키워드는 언론사에서 사회적 관심이 어디에 초점이 맞춰지고 있는지 보여주는 가장 일차적인 지표로 사용된다[16]. 따라서 본 절에서는 세종시 건설 각 이슈에 대해 언론사별 초점에 변화가 있는지 확인하기 위해 빈도분석이 활용되었다. 빈도분석 결과는 각 이슈에 따라 언론사별로 상위 10개의 키워드로 나타났다[표 2].

상위 10개 출현 단어를 토대로 볼 때, 신행정수도 기간에 한겨레는 수도의 '이전', '대통령', '총선', '격돌', '박근혜', '국감' 등 정치와 관련된 키워드와 '투기', '부동산' 등의 부동산 관련 키워드가 주로 나타났다. 또한 '지역', '충남' 등 신행정수도 건설 위치와 관련된 키워드가 도출되었다. 동아일보는 '이전', '대통령' 등 정치 키워드와 '서울', '대전', '충남', '연기', '경기', '여론' 등 신행정수도를 둘러싼 지역의 여론과 관련된 키워드가 주로 나타났다. 특히, 신행정수도와 관련되어 '반대' 그리고 '위헌'과 같은 신행정수도 건설에 반하는 키워드가 나타난 것이 특징이다. 한국일보는 '이전'과 '대통령', '총선', '탄핵'이라는 정치적 키워드와 '투기', '분양', '토지', '아파트', '경제', '투자' 등 부동산 및 경제와 관련된 키워드가 주로 나타나고 있었다. 주목할만한 점은 '탄핵'이란 키워드의 등장이다.

세 개의 언론사 모두 상위에 출현한 키워드는 유사하지만, 그 순서와 출현한 키워드의 순서는 일부 차이를 보였다. 각 언론사는 신행정수도의 이전과 그 공약을 발표한 대통령에 초점을 맞추고 있었다. 하 순위 키워드로 갈수록 신문사간 특징을 확인할 수 있었다. 한겨레는 대선과 관련된 키워드(총선, 격돌 등)를 포함하고 있었으며, 한국일보는 부동산에 초점이 맞춰진 키워드(투기, 분양 등)가 나타났다. 동아일보는 각 지역(서울,

표 2. 이슈에 따른 언론사별 빈도분석 결과

주요 이슈	한겨레		동아일보		한국일보	
	키워드	출현 빈도	키워드	출현 빈도	키워드	출현 빈도
신행정수도	이전	29	이전	65	이전	51
	대통령	12	대통령	37	대통령	22
	투기	10	서울	36	투기	18
	충선	8	반대	29	분양	17
	부동산	8	대전	27	토지	14
	격돌	8	충남	26	충선	12
	박근혜	8	연기	16	아파트	12
	국감	8	경기	15	경제	11
	지역	7	여론	15	탄핵	11
행정중심복합도시	충남	7	위헌	15	투자	10
	이전	10	대전	7	이전	16
	토지	5	충남	7	확정	8
	연기	5	충북	5	지역	7
	공주	5	건설	4	충남	7
	이명박	4	토지	3	토지	7
	지방	4	계획	3	분양	7
	균형발전	4	현재	3	대책	5
	결정	4	반대	2	기업	3
세종시 수정안	오송	4	함의	2	각하	3
	투기	4	건설청	2	건설청	3
	원안	44	박근혜	27	박근혜	25
	대통령	42	대통령	15	지방	21
	박근혜	38	친박	15	갈등	19
	기업	29	민심	14	친박	18
	친박	27	기업	14	원안	17
	민심	24	혁신	12	대통령	17
	이전	24	지방	12	투표	16
	여당	26	여론	11	여론	13
	정치	22	원안	11	민심	12
	투표	22	반대	9	여야	11

대전, 충남 등)의 여론을 대변하고 있는 것으로 해석될 수 있다. 특히, 동아일보와 한국일보는 '반대', '위헌', '탄핵'이라는 키워드로 미루어 보았을 때, 신행정수도법 통과 및 시행 기간에 정치적 이슈를 한겨레보다 비중있게 다루고 있었다.

행복도시법의 시행과 헌법소원 각자가 진행되었던 행정중심복합도시 기간에 한겨레는 '이전', '이명박', '균형발전', '결정' 등 정치 및 정책적 키워드를 포함하고 있었다. 또한, '토지', '투기'라는 부동산 관련 키워드와 '연기', '공주', '지방', '오송' 등 지역적 키워드가 나타났다. 동아일보는 '대전', '충남', '충북' 등 지역과 관련된

키워드가 상위에 랭크 되어 있었으며, '건설', '토지', '계획', '합의', '건설청' 등 행복도시 건설과 관련된 키워드가 나타난 것으로 분석되었다. 또한, 헌법소원 각하와 관련된 키워드로 '현재'가 도출되었으며, 이와 함께 '반대'하는 키워드를 포함하고 있었다. 한국일보는 '이전', '확정', '대책', '기업', '각하' 등 정치적 키워드와 '지역', '충남' 등 지역적 키워드, '토지', '분양' 등 부동산과 관련된 키워드가 나타났다.

행복도시와 관련된 이슈에서 세 개 언론사 특징을 보면, 한겨레는 지역 균형발전과 지방 이전과 관련된 내용에 초점(이전, 지방, 균형발전 등)을 맞추고 있었다. 반면, 동아일보는 건설청의 행정중심복합도시 건설의 충청지역 입장과 여론을 대변(대전, 충남, 충북, 건설, 계획 등)하고 있었으며, 현재의 결정에 반대한다는 것에 초점을 맞추고 있는 것으로 분석되었다. 한국일보는 행복도시법이 헌법소원 각하로 인해 행복도시 건설이 확정되었다는 내용을 중점적(이전, 확정, 각하 등)으로 다루고 있었다.

세종시 수정안은 정권이 바뀔에 따라 행복도시 계획의 변경을 다루었던 기간이다. 이 기간에는 한겨레, 동아일보, 한국일보 등 세 개의 언론사 모두 상위에 나타난 키워드가 정치와 관련되어 있는 것을 확인할 수 있었다. 주요 특징으로, 한겨레는 세종시 수정안이 아닌, 원안으로 진행되어야 한다는 내용에 초점(박근혜, 친박, 민심, 이전, 투표 등)이 맞춰지고 있었다. 동아일보는 세종시 수정안 이슈에서 '원안', '반대', '기업', '혁신'이라는 키워드를 중심으로 보도하고 있었다. 또한, 한국일보는 '지방', '갈등', '원안', '투표', '여론', '민심' 등의 키워드를 통해 세종시 수정안 이슈로 정치적, 사회적 갈등이 표출되고 있다는 것을 간접적으로 보여주고 있다.

2. 감성분석 결과

빅데이터를 이용한 감성분석은 텍스트 문서에서 감정, 태도, 의견, 성향이 어떻게 표현되고 있는지 빈도분석보다 비교적 정확하게 확인할 수 있다. 특정한 주제에 대해 긍정 또는 부정적 표현들이 포함되어 있는지 확인하는 분석 방법이다[19]. 세종시 건설 계획에 대해 언론사별 논조에 변화가 있었는지 확인하고자 감성분석이 수행되었다[표 3].

표 3. 언론사별 감성분석 결과

구분	한겨레			동아일보			한국일보		
	신행정수도	행정중심 복합도시	세종시 수정안	신행정수도	행정중심 복합도시	세종시 수정안	신행정수도	행정중심 복합도시	세종시 수정안
감성 분석 결과									
구분	신행정수도	행정중심 복합도시	세종시 수정안	신행정수도	행정중심 복합도시	세종시 수정안	신행정수도	행정중심 복합도시	세종시 수정안
긍정	0.0964	0.0930	0.0583	0.0652	0.0133	0.0523	0.0573	0.0874	0.0573
부정	0.0219	0.0581	0.0772	0.0766	0.0266	0.0377	0.0796	0.0526	0.1100
종합	0.0745	0.0348	-0.0189	-0.0114	-0.0133	0.0145	-0.0222	0.0333	-0.0527

신행정수도와 행정중심복합도시, 세종시 수정안에 따라 나타난 한겨레 감성분석 결과는 다음과 같다. 신행정수도(긍정 0.0964, 부정 0.0219), 행정중심복합도시(긍정 0.0930, 부정 0.0581), 세종시 수정안(긍정 0.0583, 부정 0.0772) 등으로 각 시기의 감성 점수가 도출되었다. 신행정수도(종합 0.0745)와 행정중심복합도시(종합 0.0348)에는 긍정적인 논조를 보였다. 반면에 세종시 수정안(-0.0189) 시기에는 부정적인 논조를 갖는 것으로 분석되었다. 행정중심복합도시 보다 신행정수도 시기에 긍정적인 논조가 더 강한 것으로 분석된 것이다. 즉, 한겨레는 세종시 건설 시기에 따라 긍정-긍정-부정으로 보도의 논조가 변화되었다.

동아일보는 신행정수도(긍정 0.0652, 부정 0.0766), 행정중심복합도시(긍정 0.0133, 부정 0.0266), 세종시 수정안(긍정 0.0523, 부정 0.0377)으로 감성분석 결과가 도출되었다. 신행정수도 이슈(종합 -0.0114)와 행정중심복합도시 이슈(종합 -0.0133)에서 부정적인 논조를 보이고 있었다. 세종시 수정안 이슈는 긍정적 논조(종합 0.0145)로 확인되었다. 이 결과는 세종시 건설 계획 시기에 따라 동아일보의 논조가 반대-반대-찬성으로 변화되었다는 것을 나타낸다. 세 가지 시기에 따른 한국일보의 감성분석 결과는 다음과 같다. 한국일보는 신행정수도(긍정 0.0573, 부정 0.0796)와 행정중심복합도시(긍정 0.0874, 부정 0.0526), 세종시 수정안(긍정 0.0573, 부정 0.1100)으로 분석값이 도출되었다. 따라서, 한국일보는 부정(종합 -0.0222)-긍정(종합

0.0333)-부정(종합 -0.0527)으로 논조가 변화되었다.

세종시 건설 기간을 중심으로 보면, 신행정수도에서 한겨레는 동아일보와 한국일보 보다 더 긍정적인 논조를 보이는 것으로 나타났다. 한국일보는 동아일보보다 신행정수도 이슈를 더 부정적으로 조명하고 있었다. 행정중심복합도시는 한겨레와 한국일보가 긍정적인 논조를 갖고 있었으며, 동아일보는 부정적인 논조를 갖고 있었다. 세종시 수정안 기간에는 한국일보가 가장 부정적인 보도 태도를 보이고 있었다. 동아일보는 긍정적인 논조를 나타내고 있는 것으로 분석되었다. 따라서, 감성분석 결과를 종합하면, 신행정수도, 행정중심복합도시, 세종시 수정안 등 세종시 건설 계획에 따라 세 언론사의 논조는 변화되고 있었다. 특히, 한겨레와 동아일보는 정치적, 이념적 신념에 따라 정책의 변화에 찬성과 반대의 입장을 명확히 보여주고 있었다.

3. 사회연결망분석 결과

본 절에서는 세종시 건설 이슈를 중심으로 언론사별 내용에 차이가 있는지 알아보기 위해 사회연결망분석이 진행되었다. 사회연결망분석은 신문기사의 내용적 흐름을 파악하고자 모듈성(Modularity)과 사이중심성, 고유벡터중심성을 통해 수행되었다. 모듈성은 커뮤니티 구조를 파악하기 위한 가장 적절한 분석 방법으로, 연결망의 주요 흐름을 파악할 수 있다. 또한, 연결중심성과 사이중심성을 동시에 고려하는 분석 방법이다[25]. 사이중심성은 특정 노드가 다른 노드들 간의 중개자 역

할을 수행하는 정도를 의미한다. 즉, 전체 연결망에서 전반적인 주장을 갖는 키워드가 사이중심성 값이 높게 나타난다고 볼 수 있다[26]. 고유벡터중심성은 연결된 노드의 개수보다 얼마나 중요한 노드와 연결되어 있는지를 파악하는데 사용되는 지표이다. 즉, 하나의 노드가 다른 중요한 노드에 연결되면 그 노드의 의미는 중요하다는 것을 나타낸다.

본 절의 사회연결망분석 결과에서 노드의 개수는 최소 28개에서 최대 50개까지 표현되었다. 이는 10 이상의 연결정도(Degree)를 갖는 노드만 추출하여 분석한 것이다. 엣지(Edge)는 연결강도(Degree Weight)가 3 이상 표현되도록 설정하였다. 노드의 색상과 엣지의 굵기는 모듈성, 노드 크기는 고유벡터중심성이 반영되었다.

3.1 신행정수도 기간의 SNA 결과

먼저 신행정수도 이슈의 한겨레 SNA는 263*263 단어행렬을 통해 분석되었다[표 4]. 그중 28개 노드와

133개 엣지가 표현되었다. 사이중심성과 고유벡터중심성이 가장 높은 키워드 상위 5개는 수도(65.21, 1.00), 행정(50.47, 0.90), 이전(21.64, 0.54), 한나라(20.55, 0.36), 대통령(11.32, 0.32) 등이다. 커뮤니티 군집은 신행정수도 이전과 이를 둘러싼 정치적, 정책적 이슈 등 3개이다. SNA의 분석결과, 신행정수도 이전이 노무현 대통령으로부터 추진 되었으며, 서울과 경기 지역을 중심으로 반대 의견이 모이고 있음을 분석결과를 통해 짐작할 수 있었다. 또한, 한나라당과 우리당이 신행정수도 이전과 관련되어 격돌하고 있었으며, 대통령 탄핵과 부동산의 이슈 내용을 포함하고 있었다. 이러한 분석 결과는 신행정수도 이전에 초점을 맞추고 있는 것으로 해석할 수 있다.

동아일보의 신행정수도 SNA 결과는 다음과 같다. 829*829 단어행렬로 분석되었으며, 33개 노드와 164개 엣지로 표현되었다. 사이중심성과 고유벡터중심성이 가장 높은 키워드는 수도(219036.4, 1.00), 이전(137198.8, 0.85), 수도권(117638.5, 0.34), 대통령

표 4. 신행정수도에 대한 언론사별 사회연결망분석 결과

한겨레			동아일보			한국일보		
키워드	사이 중심성 값	고유벡터 중심성 값	키워드	사이 중심성 값	고유벡터 중심성 값	키워드	사이 중심성 값	고유벡터 중심성 값
수도	65.21	1.00	수도	219036.4	1.00	수도	106002.1	1.00
행정	50.47	0.90	이전	137198.8	0.85	대통령	48149.1	0.39
이전	21.64	0.54	수도권	117638.5	0.34	이전	33298.7	0.64
한나라	20.55	0.36	대통령	59556.7	0.30	행정	30199.5	0.65
대통령	11.32	0.32	서울	50191.1	0.32	부동산	22784.5	0.37
우리당	11.28	0.26	행정	41530.8	0.45	정부	21621.1	0.30
서울	10.04	0.25	칼럼	29586.0	0.11	총선	19276.1	0.21
경기	8.43	0.24	주민	28676.8	0.30	정치	18312.3	0.15
국감	7.50	0.25	정부	26775.9	0.25	투기	16530.1	0.29
전망	5.95	0.17	사실	22607.1	0.17	한나라	16351.7	0.25

(59556.7, 0.30), 서울(50191.1, 0.32)로 나타났다. 커뮤니티 군집은 크게 4개로 구분되었으며, 주제는 신행정수도 이전의 반대, 수도권 지역에서 신행정수도에 대한 논란, 충청권 지역의 투기시장 등이다. 동아일보에서 중심성 키워드와 주제를 통해 본 신행정수도 SNA 의미는 다음과 같이 해석될 수 있다. 신행정수도 이전 추진에 따라 당시 여당인 한나라당을 중심으로 반대 의견이 나타나고 있었다. 대전·충남·충북 등 충청권에서의 부동산 투기에 대한 우려 섞인 목소리가 담겨있었다.

한국일보의 신행정수도 이슈의 SNA는 167*167 단어행렬과 39개의 노드, 226개의 엣지로 표현되었다. 사이중심성과 고유벡터중심성 상위 키워드는 수도(106002.1, 1.00), 대통령(48149.1, 0.39), 이전(33298.7, 0.64), 행정(30199.5, 0.65), 부동산(22784.5, 0.37), 정부(21621.1, 0.30)로 분석되었다. 주제의 군집은 크게 4개의 군집으로 나타났으며, 신행정수도 이전에 대한 현재의 판결, 부동산 투자와 투기, 대통령 탄핵 등 이다.

SNA 분석결과, 한국일보는 신행정수도를 부동산과 관련된 내용에 부처 이슈를 부각시키고 있었다. 특징은 대통령 탄핵을 직접적으로 언급하고 있다는 것이다.

신행정수도에서 각 언론사의 내용 차이를 정리하면 다음과 같다. 부동산 투기의 내용이 공통으로 다뤄지고 있었다. 한겨레는 서울 및 경기 중심으로 신행정수도를 반대하고 있는 내용을 기사화하였다. 동아일보는 당시 여당인 한나라당의 당론을 대변하는 것으로 볼 수 있다. 한국일보는 신행정수도는 정치적 목적으로 이루어진 결과이며, 직접적인 대통령 탄핵 내용을 전면에 표현하고 있었다.

3.2 행정중심복합도시 기간의 SNA 결과

행정중심복합도시 이슈의 한겨레 SNA는 125*125 단어행렬을 통해 분석되었으며, 33개 노드와 102개 엣지로 표현되었다[표 5]. 사이중심성과 고유벡터중심성이 가장 높은 키워드는 행정도시(78.98, 1.00), 행정

표 5. 행정중심복합도시 이슈에 대한 언론사별 사회연결망분석 결과

한겨레			동아일보			한국일보		
키워드	사이 중심성 값	고유벡터 중심성 값	키워드	사이 중심성 값	고유벡터 중심성 값	키워드	사이 중심성 값	고유벡터 중심성 값
행정도시	78.98	1.00	행정도시	9699.7	1.00	행정도시	8078.3	0.69
행정	52.60	0.84	사실	4609.5	0.16	이전	6114.4	0.72
수도권	60.13	0.77	충남	3682.4	0.48	토지	5401.3	0.39
이전	58.11	0.66	부동산	2097.5	0.13	충남	5074.2	0.22
균형발전	28.77	0.60	충북	2059.0	0.25	행정	4513.3	1.00
도시	13.90	0.56	수도	1735.5	0.22	대상	2596.0	0.10
중심	21.88	0.52	분양	1370.8	0.13	확정	2056.4	0.38
토지	87.75	0.48	계획	1321.5	0.08	대책	2039.9	0.54
공주	6.10	0.47	인사	1230.0	0.16	부동산	2009.1	0.37
연기	6.10	0.47	부처	1210.6	0.17	공주	1912.0	0.03

(52.60, 0.84), 수도권(60.13, 0.77), 이전(58.11, 0.66), 균형발전(28.77, 0.60) 이다. 커뮤니티 군집은 6개로 확인되었으며, 주제는 행정중심복합도시 충청권 건설, 행정수도 이전, 연기 및 공주 등 충청권의 투기 제한, 지역균형발전 등으로 나타났다. 한겨레는 행정도시건설을 통해 행정기능이 이전되어야 함과 지역의 균형발전 필요성을 이슈화시키고 있었다. 또한, 주택 등의 부동산 투기에 대해서는 제한되어야 함을 기사 제목에 담아 표현하고 있음을 알 수 있었다.

동아일보의 행정중심복합도시 이슈 SNA는 88*88 단어행렬과 48개 노드, 163개 엣지가 결과로 나타났다. 행정도시(9699.7, 1.00), 사설(4609.5, 0.16), 충남(3682.4, 0.48), 부동산(2097.5, 0.13), 충북(2059.0, 0.25) 등이 사이중심성과 고유벡터중심성이 높은 키워드로 분석되었다. 주요 군집은 5개로 나타났다. 주제는 행정도시의 현재결정여부, 행복도시건설청, 행복도시건설에 따른 자연 훼손, 행정수도 서울, 부동산 투기와 관련된 이슈 등으로 나타났다. 동아일보의 SNA 결과는

행정수도서울 그리고 계룡산 훼손 등의 키워드를 통해 행정중심복합도시 건설을 비판적 시각으로 보는 것으로 해석될 수 있다. 또한, 부동산과 관련된 이슈를 크게 부각시키고 있는 것으로 나타났다.

54*54 단어행렬과 37개 노드, 143개 엣지로 구성된 한국일보의 행정중심복합도시 이슈 SNA 결과는 다음과 같다. 사이와 고유벡터중심성 값이 큰 주요 키워드는 행정도시(8078.3, 0.69), 이전(6114.4, 0.72), 토지(5401.3, 0.39), 충남(5074.2, 0.22), 행정(4513.3, 1.00) 등으로 분석되었다. 주요 커뮤니티는 4개로 나타났으며, 공공기관 이전 확정, 여야에서 합의가 된 행정중심복합도시, 행복도시법의 각하, 토지와 부동산 대책 등의 내용을 담고 있다. SNA의 의미를 살펴보면, 행복도시에서 한국일보는 행복도시 이전과 관련되어 행정, 중심, 이전, 도시, 행정도시 등 다양한 키워드로 표현 및 기사화하고 있었다. 행정도시법이 각하되면서 토지, 부동산 시장의 대책을 고려가 필요하다는 주장을 제기하고 있는 것으로 나타났다. 결과를 전반적으로 살펴보면

표 6. 세종시 수정안 이슈에 대한 언론사별 사회연결망분석 결과

한겨레			동아일보			한국일보		
키워드	사이 중심성	고유벡터 중심성	키워드	사이 중심성	고유벡터 중심성	키워드	사이 중심성	고유벡터 중심성
세종시	257750.4	1.00	세종시	148189.5	1.00	세종시	270971.9	1.00
수정안	72938.0	0.73	수정안	30118.0	0.58	수정안	55071.9	0.65
정부	33905.8	0.39	박근혜	22438.8	0.32	총리	23277.6	0.26
선택	24943.4	0.22	총리	13998.3	0.25	박근혜	22993.2	0.17
박근혜	19227.0	0.21	정치	13493.1	0.21	갈등	17190.1	0.21
대통령	17433.1	0.20	대통령	10694.5	0.18	지방선거	14994.7	0.17
후보	12097.4	0.18	한나라	10362.9	0.28	국회	14241.0	0.28
한나라	11323.0	0.17	사설	10339.6	0.18	사설	13128.5	0.21
폭풍	10676.9	0.38	민주	8231.1	0.17	후보	12965.0	0.09
청와대	10607.6	0.17	경기	8009.5	0.08	대표	11870.1	0.19

을 때, 행복도시와 관련되어 긍정적인 내용이 담겨있는 것으로 해석될 수 있다.

언론사별 행정중심복합도시 이슈에 대한 내용적 차이는 다음과 같다. 신행정수도 이슈와 마찬가지로 부동산 투기 우려와 대책이 필요하다는 것에 공통적인 내용을 보도하고 있었다. 한겨레는 지역균형발전을 위한 행정중심복합도시의 추진 필요성에 대해 기사화하고 있었다. 동아일보는 행복도시로 인해 자연훼손이 우려된다는 내용과 행정수도는 서울이라는 내용을 보도하고 있었으며, 한국일보는 행정중심복합도시 건설에 대해 여야가 합의를 이루었다는 내용에 초점이 맞춰져 있었다.

3.3 세종시 수정안 기간의 SNA 결과

세종시 수정안 이슈 한겨레 SNA는 591*591 단어행렬을 통해 분석되었으며, 32개의 노드와 110개 엣지로 구성되었다. 세종시(257750.4, 1.00), 수정안(72938.0, 0.73), 정부(33905.8, 0.39), 선택(24943.4, 0.22), 박근혜(19227.0, 0.21) 등은 사이 및 고유벡터 중심성 값이 높게 나타난 키워드이다. 커뮤니티 군집은 2개로 나타났으며, 주요 주제는 세종시 수정안 백지 확정, 부처이전확정, 원안 추진, 한나라당 내의 친박과 대통령 갈등 등으로 확인되었다. 세종시 수정안에서 한겨레는 당시 정부에서 세종시 수정안을 선택하였지만, 민심, 친박, 박근혜, 대구 등 주변 정치적 이슈 등으로 세종시 수정안이 백지화 및 원안 추진 확정되었음을 키워드를 통해 내용을 전달하고 있었다.

세종시 수정안 이슈 동아일보 SNA는 474*474 단어행렬로 구성되었으며, 36개 노드와 199개 엣지로 표현되었다. 사이중심성과 고유벡터중심성 상위 키워드는 세종시(148189.5, 1.00), 수정안(30118.0, 0.58), 박근혜(22438.8, 0.32), 총리(13998.3, 0.25), 정치(13493.1, 0.21) 순으로 나타났다. 커뮤니티 집단은 크게 3개로 분류되었으며, 주제는 세종시 수정안 국회 부결, 박근혜·정몽준·김무성 중심의 세종시 원안, 대통령·총리·단신 등으로 볼 수 있다. 동아일보는 세종시 수정안 부결에 반대하는 것으로 해석될 수 있다. 이는 세종시 수정안에 찬성한다는 뜻으로 볼 수 있다. 또한, 박근혜, 정몽준, 김무성과 이명박 대통령의 정치적 갈등을

기사화하고 있었다.

한국일보의 세종시 수정안 이슈는 559*559 단어행렬, 50개 노드와 134개 엣지로 SNA가 분석되었다. 사이중심성과 고유벡터중심성 값이 큰 주요 키워드는 세종시(270971.9, 1.00), 수정안(55071.9, 0.65), 총리(23277.6, 0.26), 박근혜(22993.2, 0.17), 갈등(17190.1, 0.21) 등으로 나타났다. 커뮤니티는 3가지로 분류되었으며, 세종시 수정안에 따른 내외부적 정치적 갈등이 주된 내용을 다루고 있었다. SNA 결과의 세부 내용은 다음과 같다. 세종시 수정안이 정치적, 사회적, 지역적으로 논란 및 갈등을 불러일으키고 있다는 내용으로 기사화되고 있었다. 한국일보는 정치적으로 한나라당 친박 의원과 당시 이명박 대통령이 세종시 수정안과 원안을 놓고 갈등이 있다는 것을 주된 내용으로 다루고 있었다.

세종시 수정안 이슈는 공통으로 당시 박근혜 대표와 이명박 대통령간 갈등이 주요 내용으로 기사화되었다. 언론사별 내용적으로 미묘한 차이가 나타났다. 한겨레는 세종시 수정안 백지화 그리고 원안 추진이라는 내용에 초점이 맞춰져 있었으며, 동아일보는 세종시 수정안 부결이라는 키워드를 사용하여 기사화하고 있었다. 한국일보는 세종시 수정안을 중심으로 나타나는 지역적, 정치적, 사회적 갈등에 초점을 두고 있었다.

V. 결론

본 연구의 목적은 빅데이터 분석 기법을 이용하여 언론사별 세종시 건설 이슈의 내용 및 논조 변화를 분석하는 것이다. 자료 중 보수지 동아일보 신문기사는 동아닷컴에서 직접 수집하였다. 이외 진보지 한겨레와 중도지 한국일보 자료는 한국언론진흥재단 빅카인즈에서 수집되었다. 검색 키워드는 신행정수도, 행정중심복합도시, 세종시 수정안 이다. 빅데이터 분석기법을 통해 데이터는 전처리 되었으며, 자료는 신행정수도법 통과 및 시행(2003년 12월 29일 ~ 2004년 4월 17일), 행복도시건설법 공포 및 시행(2004년 11월 15일 ~ 2005년 3월 18일), 세종시 수정안 발표 및 국회 부결 전(2010년 1월 11일 ~ 2010년 5월 29) 기간의 신문기

사이다.

자료 분석은 파이썬과 Gephi 0.9.2 프로그램을 활용하였고, 빈도분석, 감성분석, 사회연결망분석 방법이 사용되었다. 이러한 과정을 통해 세종시 건설 계획이 변화됨에 따라 각 신문사의 보도에 대한 논조 및 내용 변화를 분석하였다. 빈도분석 결과를 종합하면 다음과 같다. 신행정수도-행복도시-세종시 수정안의 세종시 건설 이슈에 따라, 한겨레는 찬성-찬성-반대 경향의 특성을 나타내고 있었다. 동아일보는 반대-반대-찬성 입장을 갖는 것으로 분석되었다. 한국일보는 반대-찬성-반대의 특성을 갖는 것으로 나타났다. 각 언론사는 세종시 건설 기간에 서로 다른 사회적 이슈에 초점을 맞추고 있었으며, 찬성 또는 반대 등 다양한 입장 특성을 갖는 것으로 분석되었다.

감성분석 결과, 언론사는 정치적, 이념적 신념 또는 성향에 따라 세종시 건설에 대한 찬성과 반대의 입장을 명확히 보여주고 있었다. 한겨레는 긍정-긍정-부정의 논조를 보였으며, 동아일보는 부정-부정-긍정의 보도 태도를 갖고 있었다. 중도일보는 부정-긍정-부정의 보도 특성을 갖고 있었다. 한국일보는 기존 연구와 다른 결과로 도출된 것이 특징이다. 조용구(2013) 연구에서 한국일보는 행복도시 기간에 반대 입장을 갖는 것으로 나타났다[3]. 본 연구에서는 한국일보가 행복도시 기간에 찬성 논조를 나타내는 것으로 분석되었다. 이는 신행정수에서 언급된 것과 같이 대량의 데이터를 직접 분석할 때에 나타나는 한계점으로 판단된다.

언론사별 사회연결망분석 결과는 다음과 같다. 전반적으로 세종시 건설 기간에 각 언론사는 유사한 내용을 공통으로 포함하고 있었다. 그것은 세종시 부동산과 관련된 내용이다. 신행정수도 기간의 경우, 한겨레의 내용적 특성은 서울 및 경기권에서 반대하고 있는 것으로 도출되었다. 동아일보는 당시 여당인 한나라당을 중심으로 신행정수도에 대해 반대하고 있었다. 한국일보의 경우 대통령 탄핵이라는 내용을 중심으로 보도 특성이 도출되었다.

행정중심복합도시 이슈에서는 한겨레가 지역발전 내용에 초점을 맞추고 있었다. 동아일보는 도시건설에 따른 자연훼손이 쟁점으로 부각 되었다. 한국일보는 행복도시의 여·야 합의 내용을 중심으로 보도하고 있었다.

세종시 수정안 기간에서 한겨레는 세종시 수정안 백지화 및 원안 추진이라는 내용적 특성에 초점을 맞추고 있었다. 동아일보는 세종시 수정안 부결이라는 쟁점을 통해 한겨레보다 세종시 수정안 관련 쟁점을 축소하는 경향을 보였다. 한국일보는 세종시 수정안 기간에 지역, 정치, 사회 등 다양한 갈등을 불러일으키고 있다는 것에 초점을 둔 것이 특징이다.

세종시 건설 기간에 따른 언론사 입장에 대한 변화를 내용적 측면에서 보았을 때, 본 연구의 결과는 기존에 수행되었던 연구 결과와 궤를 같이하고 있는 것으로 분석되었다[3][5][8]. 즉, 각 언론사는 진보와 보수, 중도로 분류되어 세종시 건설 이슈에 대해 정치적, 이념적 성향을 신문 기사에 포함하고 있다는 것이다. 방법론적 측면에서 보면, 언론사별로 세종시 건설 이슈에 따라 핵심 키워드를 제외한 키워드는 다른 키워드로 구성되어 어있는 것이 확인되었다. 즉 언론사의 정치적인 성향에 따라 핵심 키워드를 중심으로, 주변에 서로 다른 키워드가 포진되어 세종시 건설에 대한 쟁점을 언론사별로 다른 시각으로 접근하고 있다는 것이다[6][13].

본 연구는 빅데이터 분석을 활용하여, 세종시 건설 기간에 따라 각 언론사 시각이 다르게 변화되고 있음을 분석하고 논의한 것에 의의가 있다. 본 연구는 다음과 같은 한계가 있다. 첫째, 기사 제목을 중심으로 분석한 것이다. 기사 제목은 표시성 즉 대표성을 갖고, 핵심적인 정보가 담겨있다는 측면에서 중요한 역할을 한다[20][21]. 그러나 기사 내용의 문장과 문장, 문장과 단어 간 관계와 의미를 해석하기에는 한계가 있다[13].

둘째, 세종시 건설 이슈는 2002년부터 지금까지 현재 진행형이다. 방대한 시기를 3개의 시기로 나누고, 단기 데이터를 활용하여 분석한 것은 본 연구의 한계이다. 향후, 세종시 건설의 이슈를 세분화하고 언론을 포함한 연구기관, 학자 등의 입장이 어떻게 변화되었는지 연구의 범위를 확대하여 분석할 필요가 있다. 본 연구에서는 빅데이터 분석 방법이 언론을 분석하는데 효과적인 방법임을 확인할 수 있었다. 빅데이터 분석은 여론을 분석하는데 있어 직접설문보다 용이하며, 오프라인 조사를 보완할 수 있는 합리적인 방법이다. 이처럼 빅데이터 분석을 도시계획 또는 도시건설 시 시민 및 언론 의견 분석에 활용한다면, 분석결과에 대한 객관성

을 담보할 수 있을 것으로 기대된다.

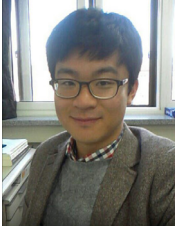
참 고 문 헌

- [1] 이상호, *공간을 말하다*, 북바이북, 2020.
- [2] 권오혁, “세종시 건설에 대한 검토와 대안의 모색,” 한국경제지리학회지, 제15권, 제4호, pp.464-480, 2012.
- [3] 조용구, *세종시건설의 찬반 쟁점과 특성 분석*, 한밭대학교, 석사학위논문, 2013.
- [4] 유영돈, 마정미, “사회적 갈등이슈에 대한 전국 7개 일간지 보도 비교: 세종시 갈등 보도 내용분석을 중심으로,” 언론과학연구, 제15권, 제2호, pp.285-316, 2015a.
- [5] 이현숙, “세종시 뉴스 보도에 나타나는 프레임 분석,” 정치정보연구, 제16권, 제1호, pp.229-264, 2013.
- [6] 감미아, 송민, “텍스트 마이닝을 활용한 신문에 따른 내용 및 논조 차이점 분석,” 지능정보연구, 제18권, 제3호, pp.53-77, 2012.
- [7] 김춘식, “대안 없는 부정적 보도, 취재원은 수도권 편중 ‘행정수도 이전’ 보도,” 신문과 방송, 제8권, pp.78-81, 2004.
- [8] 유승현, 황상재, “포털미디어의 뉴스 프레임에 대한 탐색적 연구,” 사이버커뮤니케이션학보, 제20호, pp.197-232, 2006.
- [9] 박승민, *전국과 지방 TV 뉴스간 갈등 이슈 보도 비교: 세종시 수정안에 대한 프레임 분석*, 경희대학교 대학원, 석사학위논문, 2011.
- [10] 유영돈, 마정미, “세종시 갈등에 대한 뉴스 프레임 연구 : 7개 전국 일간지 기사 분석을 중심으로,” 한국언론학보, 제59권, 제3호, pp.29-59, 2015b.
- [11] 박한우, Loet Leydesdorff, “한국어의 내용분석을 위한 KrKwic 프로그램의 이해와 적용 : Daum.net에서 제공된 지역혁신에 관한 뉴스를 대상으로,” Journal of The Korean Data Analysis Society, Vol.6, pp.1377-1387, 2004.
- [12] 양명철, 한혜경, “한국미디어의 이주민보도 프레임 분석,” 동북아 문화연구, 제41권, pp.71-89, 2014.
- [13] 김대욱, 최명일, “의미연결망분석을 이용한 2005~2014년 자살보도 분석,” 한국언론학보, 제60권, 제2호, pp.178-208, 2016.
- [14] 박승정, 전진오, 김선우, 김성태, “국내 주요일간지의 대통령 이슈소유권에 대한 빅데이터 분석,” 정치정보연구, 제20권, 제3호, pp.25-55, 2017.
- [15] 최진호, 이해수, 진은형, “4차산업혁명에 관한 뉴스 빅데이터 활용 토픽 분석: 언론사 유형 및 주요시기에 따른 비교를 중심으로,” 사이버커뮤니케이션학보, 제36권, 제2호, pp.173-219, 2019.
- [16] 김일환, “인문학을 위한 신문 빅 데이터와 텍스트 마이닝,” 어문론집, 제78권, pp.41-62, 2019.
- [17] D. M. Blei, A. Y. Ng, M. L. Jordan, and J. Lafferty, “Latent Dirichlet Allocation,” Journal of Machine Learning Research, Vol.3, No.4/5, pp.993-1022, 2003.
- [18] 박한우, 이연옥, “복합적 텍스트 분석을 이용한 포털 댓글에 관한 연구 -17대 대통령 선거 기간 미디어 '다음'에 게시된 광운대 BBK 동영상-,” 한국자료분석학회, Vol.11, No.2, pp.731-744, 2009.
- [19] Nasukawa and Yi, IEEE INTERNATIONAL CONFERENCE ON DATA MINING, Vol.3, pp.427-434, 2003.
- [20] 송경화, 강범모, “신문 기사의 언어 사용 양상,” 인지과학, 제17권, 제4호, pp.255-269, 2006.
- [21] 조수선, “온라인 신문 기사의 제목과 개요 효과,” 한국언론학보, 제49권, 제2호, pp.5-32, 2005.
- [22] B. S. Brooks, J. Z. Sissors, and F. K. Baskette, *The art of editing*, Allyn & Bacon, 2001.
- [23] SCHÜTZE, Hinrich, MANNING Christopher D. and RAGHAVAN Prabhakar, *Introduction to information retrieval*, Cambridge University Press, 2008.
- [24] G. Salton and C. Buckley, “Term-weighting approaches in automatic text retrieval,” Information processing & management, Vol.24, No.5, pp.513-523, 1988.
- [25] M. E. Newman, “Modularity and community structure in networks,” Proceedings of the national academy of sciences, Vol.103, No.23, pp.8577-8582, 2006.
- [26] 박치성, 정지원, “텍스트 네트워크 분석: 사회적 인식 네트워크(socio-cognitive network) 분석을 통한 정책이해관계자 간 공유된 의미 파악 사례,” 정부학연구, 제19권, 제2호, pp.73-108, 2013.
- [27] www.bigkinds.or.kr
- [28] www.donga.com

저 자 소 개

조 성 수(Sung Su Jo)

정회원



- 2012년 2월 : 국립한밭대학교 도시공학과(공학사)
- 2014년 2월 : 국립한밭대학교 대학원 도시공학과(공학석사)
- 2018년 3월 ~ 현재 : 국립한밭대학교 대학원 도시공학과 박사과정

〈관심분야〉 : 빅데이터, 스마트시티, 계량도시계획

이 상 호(Sang Ho Lee)

정회원



- 1993년 8월 : 연세대학교 도시공학과(공학박사)
- 1993년 12월 ~ 1995년 2월 : 삼성그룹비서실 삼성경제연구소 선임연구원
- 1995년 2월 ~ 현재 : 국립한밭대학교 도시공학과 교수

〈관심분야〉 : 빅데이터, 스마트시티, 계량도시계획