

2013년도 인권 관련 기사의 네트워크 분석*

- 포털사이트 다음(Daum) 뉴스페이지를 중심으로 -

김 찬 우**

I. 서론

정보화시대에 들어 기존의 종이신문보다 인터넷을 통해 뉴스를 접하는 비중이 늘어나고 있다.¹⁾ 인터넷으로 뉴스를 접하는 방법에서 언론사 홈페이지에 직접 찾아가 뉴스를 보는 것보다 포털사이트를 통해 뉴스를 접하는 비중 역시 점차 증가하고 있다.²⁾ 따라서 언론사 역시 포털사이트에 뉴스를 제공하기 위해 노력하고 있다. 따라서 포털사이트에 더욱더 많은 양의 뉴스가 존재한다.

그러나 많은 양의 뉴스가 포털사이트를 통해 제공된다고 해서 질까지 보장된다고 할 수는 없다. 최근 ‘세월호 사건’에 대한 언론의 보도에 대하여 비판의 목소리가 높은 것이 그 예라 할 수 있다. 각 인터넷 사이트나 포털 사이트에 기사의 클릭을 유도하기 위하여 과도한 취재나 오보까지 하여 많은 문제점이 드러났다. 특히 세월호 관계자들의 인권이나 프라이버스가 전혀 보장되지 않는 경우도 발생하였다.

예를 들어 고인의 생전의 모습이 담긴 사진이나 유품, 메시지 등을 공개하거나 사망

* 투고일자 : 2014. 5. 26 심사일자 : 2014. 6. 20 게재확정일자 : 2014. 6. 23

** 영남대학교 통일문제연구소 연구원

1) 한국언론진흥재단 ‘2013년 언론수용자 의식조사’에 따르면, 이용자들이 뉴스를 접하는 경로로 ‘스마트폰, 태블릿 PC 등 모바일 기기를 통해 인터넷으로 뉴스를 접한다’는 응답이 55.3%로 가장 높았고, ‘데스크톱 PC, 노트북 등 고정형 단말기를 통한 인터넷으로 뉴스를 접한다’는 응답의 경우 50.7%로 다음으로 높은 것으로 조사되었다. 또한 ‘종이신문으로 뉴스를 접한다’라는 응답은 2011년도에는 두 번째로 비중이 높았으나, 2012년도부터 세 번째로 밀려 2013년도에는 33.8%를 차지했다. 한국언론재단, 2013, “언론수용자 의식조사-제18회 <미디어 환경변화에 따른 이용자 행태조사> 조사분석,” 32쪽.

2) 인터넷으로 뉴스를 접하는 경우 데스크톱 등 고정형 단말기를 사용하는 사람(84.1%)과 스마트폰 등 모바일 기기 등을 이용하는 사람(76.2%)을 구분하여 각 언론사의 개별 홈페이지로 접하는 경우와 포털사이트를 통해 접하는 경우를 비교해보면, 특정 언론사 사이트가 아닌 포털 사이트를 경유하여 뉴스를 이용한 것으로 밝혀졌다. 또한 두 그룹 모두 60%이상이 뉴스의 작성/제공 언론사는 모른 채 이용하고 있다고 한다. 한국언론재단(2013), 33쪽.

이 확인된 학생의 일기장을 공개하는 등 언론이 ‘인권보도준칙’에 어긋나는 보도행태를 보였다.³⁾ 재난에 대한 언론의 보도준칙은 이미 2003년 대구지하철 참사를 통해 기자협회에서 재난·재해보도 가이드라인을 준비는 하였으나 제대로 진행되지 못했고, 2011년에 들어서 국가인권위원회와 한국기자협회가 공동으로 언론보도준칙을 제정하였다. 그러나 많은 시간이 흐른 지금에도 세월호 사건의 보도에서 보듯, 인권의 침해가 여전히 많아 다시금 언론 보도의 인권 침해에 대해 비판의 목소리가 높아지고 있는 것이다.

물론 언론은 비판과 감시의 기능을 가지고 있어 필연적으로 개인의 인격권 침해와 명예훼손의 가능성이 있다.⁴⁾ 또한 무의식 또는 고정관념이나 편견 등에서 비롯된 반인권적, 차별적 보도의 가능성이 있다. 그러나 언론에서 어떻게 사건을 묘사하느냐에 따라 사람들이 그 사건을 어떻게 볼 것인가에 대한 가치관이 바뀔 수 있다. 특히 아동, 여성, 성소수자 등 사회적 약자에 대해 뉴스에서 어떻게 보도하느냐에 따라 뉴스를 접하는 사람들이 해당 사건 혹은 사실에 대해 인식을 달리 할 수 있다. 이러한 뉴스의 영향력에 주목하여 뉴스가 인권과 관련하여 어떻게 보도하고 있는지 살펴보고자 한다. 특히 기자가 어떻게 인권을 인식하는가를 기사를 통해 표현할 수도 있다.

포털사이트 뉴스 데이터 중 인권 키워드가 들어간 뉴스의 빈도를 측정하고, 해당 기사의 중심도(centrality)를 측정하여 ‘인권’ 관련 기사에서 도출된 키워드들 사이의 관계망(network)을 분석한다. 특히 중심도의 경우에는 중요한 단어의 상·하위 순위의 키워드를 구분하여 분석한다. 각종 언론사가 운영하는 사이트가 아닌 포털사이트에 올라온 기사를 중심으로 살펴보는 것은 각종 언론사의 기사가 포털사이트에 모두 모여 있고, 기사를 접하는 사람들 역시 포털사이트를 많이 이용하기 때문이다.⁵⁾

따라서 이 연구는 포털사이트인 다음(Daum.net)에서 제공하는 뉴스기사 중에서 2013년 한 해 동안 인권이라는 키워드가 들어간 기사를 분석하는 것이다. 특히 해당 기사들을 수집, 분석하는 텍스트 마이닝(Text mining) 기법 중 시멘틱네트워크분석(semantic network analysis)을 활용하여 인권 기사의 패턴을 분석하고자 한다. 기존의 인권 관련 기사에 대한 분석은 일정 기간을 정하고 언론의 기사나 방송의 내용을 보면서 빈도를 분석하거나 해당 기사의 내용을 분석하는 것이 주를 이루었다. 그러나

3) 미디어오늘, 2013.04.23., “‘재난·재해보도 가이드라인’ 제정보다 ‘강제력’이 중요하다,”

<http://www.mediatoday.co.kr/news/articleView.html?idxno=116168>(검색일: 2014년 5월 31일)

4) 국가인권위원회, 2013, “주요 언론의 인권보도준칙 준수 실태조사,” 2013년 인권상황 실태조사 연구용역사업 보고서.

5) 김위근 외 3명, 2013, “한국의 인터넷 뉴스: 언론사닷컴, 인터넷신문 그리고 포털 뉴스서비스,” 한국언론재단 연구보고서 2013-04.

이 글에서는 2013년 한 해 동안 노출된 인권 관련 기사를 전체 수집하여 분석함으로써 기간의 폭을 넓혔다. 그리고 인권 기사의 주요 단어들을 도출하여 키워드 간의 상관관계를 시각화하여 패턴을 찾을 수 있다는 점에서 의의가 있다.

따라서 연구는 2013년 한 해 동안 다음 뉴스사이트에 올라와있는 인권관련 기사를 수집하고, 유의미한 단어를 도출한 다음, 단어들 간의 상관관계를 UCINET 6를 이용하여 시각화, 분석한다. 먼저 2장은 언론 및 보도에서 인권의 문제를 검토한 기존 연구를 검토하고, 다음으로 이 연구의 방법론인 시멘틱네트워크분석에 대해서 살펴본다. 3장에서는 분석방법, 4장은 분석결과, 5장은 결론으로 구성한다. 4장의 분석결과에는 수집된 데이터들의 중심성(centrality)을 중심으로 분석하고, 상·하위 20개 단어들의 중심성을 비교 분석한다.

II. 기존연구 검토

1. 인권에 대한 언론보도 분석

국내 인권에 대한 연구는 분야별로 다양하게 이루어지고 있으나 인권에 대한 언론보도 분야에 있어 최근의 연구는 국가인권위원회의 보고서와 이승연의 성범죄보도관련 연구, 고선연의 북한 인권에 대한 보도 연구가 대표적이다.

먼저 국가인권위원회에서 발표한 보고서에는 인권보도 준칙 제정과정과 인권보도준칙 미준수 사례에 대한 현황 분석을 통해 인권보도의 중요성을 분석하였다.⁶⁾ 특히 2013년 6월과 9월 두 차례 신문과 방송의 인권보도 미준수 현황을 모니터링 하여 분석하였고, 신문, 방송기자 중 간부급 기자 등 10명에게 심층인터뷰를 통해 인권보도준칙에 대한 인식을 조사하였다. 먼저 모니터링 결과 6월에는 인권보도준칙 미준수가 1.7%, 9월에는 1.6%로 나왔으나 종합편성채널의 미준수가 많은 것으로 나타났다. 특히 TV 뉴스를 통해 매일 약 4건의 개인의 인격권을 침해하는 뉴스를 시청하게 된다고 분석하였다.

이승연의 경우에는 성범죄보도의 경우에 가해자와 피해자의 신상정보 공개문제를 지적하면서 기사작성 시 법률적 자문을 받는 등 엄밀한 과정을 거쳐 보도해야 함을 강조하였다.⁷⁾ 또한 고선연의 경우 북한 인권 문제를 진보 언론에서 어떻게 보는지 분석하

6) 국가인권위원회, 2013, “주요 언론의 인권보도준칙 준수 실태조사,” 2013년 인권상황 실태조사 연구용역사업 보고서(국가인권위원회).

였다.⁸⁾ 그리고 세월호와 같은 2011년 3월에 발생한 일본 대지진에 대해 한국, 일본, 미국의 주요 방송사들의 보도태도를 비교분석한 연구도 있다.⁹⁾

이러한 연구들은 인권의 개별 사안에 대한 분석이 대부분이고, 국가인권위원회의 분석은 전반적인 인권보도에 대한 분석은 이루어졌으나 일정기간 동안의 보도에 국한되어 있으며, 보도에 대한 평가는 있으나 빈도에 대한 분석만 수행했을 뿐 사용된 단어의 패턴을 도출하고 있지 않다. 따라서 이 글에서는 2013년 한 해 동안 인권 관련 기사들 전체를 그 대상으로 하고, 관련 기사들의 주요 단어들을 추출하여 상관관계를 분석하고자 한다.

2. 네트워크 분석이론

포털사이트에서 기사들을 수집하여 단어 간 의미를 분석하기 위해서는 텍스트 마이닝(Text mining)을 통해 데이터를 수집, 정제, 분석의 과정을 거쳐야 한다. 즉 포털사이트에 있는 기사(데이터)를 수집하여 정제하는 과정은 공학의 영역이라 여기서 설명하지 않는다. 다만 정제된 데이터를 분석하기 위해서는 시멘틱네트워크(semantic network)¹⁰⁾ 분석 방법이 필요하다. 시멘틱네트워크 분석이란 단어들의 빈도와 한 문장 안에서 동시에 사용되는 단어들의 관계를 통해 분석대상의 의미와 패턴을 분석하는 기법이다. 이때 단어나 구는 각각의 노드(node)를 형성하는 개념으로 놓고 개념 간의 연결 상태를 링크(link)로 나타내는데 여기서 링크로 드러나는 단어들의 공동출현(co-occurrence) 관계를 통해 의미를 해석한다. 단어들의 사용빈도와 관계는 텍스트에서 강조되는 상징성을 보여줄 수 있고 단어들의 결합을 통해 특정한 의미를 만들어내는 경향이 있기 때문이다.¹¹⁾

시멘틱네트워크 분석 방법을 통해 언론보도에 대하여 분석한 연구는 다음과 같다. 먼저 장하용은 시민단체 보고서와 매체비평지의 기사를 대상으로 시멘틱네트워크 분석을 실시하여 시민단체가 바라보는 언론비평의 특징을 분석하였고,¹²⁾ 박성희의 경우 제17

7) 이승선, 2012, “섬세하고 엄밀한 취재 없다면 차라리 보도하지 말아야 : 인권보호 등 성범죄보도 개선방안,” 『신문과 방송』 제503호(한국언론진흥재단), 28~32쪽.

8) 고선연, 2013, “‘진보’와 ‘진보’가 바라 본 북한 인권 문제: 한겨레신문과 경향신문 비교 분석,” ATE 제3권(이화여자대학교 스크랜튼학부), 35~55쪽.

9) 백선기·이옥기, 2013, “재난방송 보도에 대한 국가별 채널간 보도태도의 비교연구,” 『한국언론학보』 제57권 1호(한국언론학회), 272~304쪽.

10) 시멘틱네트워크에 대하여 의미연결망이라고 번역하지만 아직 명확하게 정의된 것이 없기 때문에 이 글에서는 원문 그대로 시멘틱네트워크로 사용한다.

11) 박한우 & Leydesdorff, 2004, 한국어 내용분석을 위한 KrKwic 프로그램 이해와 적용, 『Jornal of The Korean Data Analysis Society』, 6권 5호, 1377~1388쪽.

대 대통령 선거 후보 토론 텍스트를 분석하여 각 이슈에 따라 후보별로 논제구성의 논리와 일관성, 그리고 차이를 찾아냈다.¹³⁾ 또한 박지영 등은 아이돌 스타들의 SNS 메시지에 드러나는 단어들을 정체성 지표로 간주하고 커뮤니케이션 방식을 규명하였으며,¹⁴⁾ 심홍진 등은 스마트폰과 소셜미디어 이용자들의 미디어 이용행태를 분석함에 있어 인터뷰 텍스트를 바탕으로 시멘틱네트워크 분석을 하였다.¹⁵⁾

신문기사를 분석 대상으로 한 시멘틱 분석은 남인용과 박한우의 연구와 김만재와 전방욱의 연구가 있다. 남인용과 박한우의 경우 대선 예비후보자에 대한 신문기사에서 사용된 단어와 연결망 구조를 통해 신문에서 만들어지는 후보 이미지를 추출하였고,¹⁶⁾ 김만재와 전방욱은 인간배아복제에 대한 신문기사를 분석대상으로 시멘틱연결망 분석을 통해 핵심적 위치에 있는 이슈와 주변부 위치에 있는 이슈를 구분하고 이를 통해 뉴스 보도의 특성을 밝혔다.¹⁷⁾ 또한 최윤정 등은 빅데이터 관련 신문기사를 수집하여 단어 간 연결망을 통해 사회의 인식변화를 연구하였다.¹⁸⁾

이러한 연구의 특징은 텍스트 즉 기사에서 주요 단어들을 추출하여 네트워크 분석을 한 것으로 이 글에서 보고자 하는 인권 보도의 네트워크 분석의 방법론을 제시하고 있다. 특히 시멘틱네트워크 분석을 위해 필요한 주요 단어 추출, 분석 값의 해석, 네트워크 시각화를 위해 필요한 요소 등을 제시하고 있다.

III. 분석방법

1. 분석대상 및 자료수집

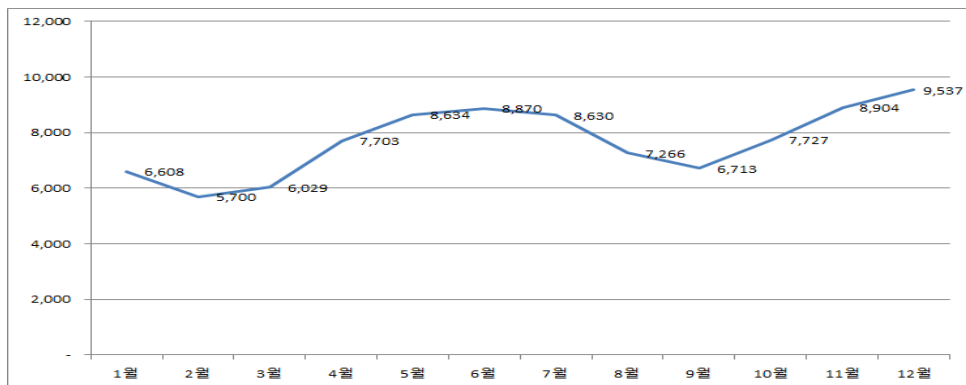
-
- 12) 장하용, 2001, “언론보도와 비평의 구조: 신문보도의 비평에 대한 네트워크 분석,” 『한국언론정보학』, 16권, 108~133쪽.
- 13) 박성희 2009, “제17대 대통령 후보 합동 토론 언어네트워크 분석,” 『한국언론정보학』, 45호, 220~254쪽.
- 14) 박지영·김태호·박한우, 2013, “의미연결망 분석을 통한 셀러브리티의 SNS 메시지 탐구,” 『방송통신연구』, 82호, 36~74쪽.
- 15) 심홍진·김용찬·손해영·임지영, 2011, “언어 네트워크 분석을 통한 스마트폰과 소셜미디어 이용자의 미디어 이용행태에 관한 탐색적 연구,” 『한국방송학보』, 25권 4호, 82~138쪽.
- 16) 남인용, 박한우, 2007, “대권 예비후보자 관련 신문기사의 네트워크 분석과 홍보전략,” 『한국정당학회보』, 6권 1호, 79~107쪽.
- 17) 김만재, 전방욱, 2012, “언어네트워크 분석 기법을 활용한 인간배아복제 신문보도 분석,” 『한국생명윤리학회』, 13권 2호, 19~34쪽.
- 18) 최윤정·권상희, 2014, “‘빅데이터’ 관련 신문기사의 의미연결망 분석,” 『사이버커뮤니케이션학보』, 제31권 1호, 241~286쪽.

이 연구는 다음 뉴스에서 ‘인권’ 단어가 들어가 있는 2013년 한 해 동안의 기사를 대상으로 하였다. 구체적으로 2013년 1월 1일부터 2013년 12월 31일까지를 분석기간으로 설정하였다. 분석대상은 다음 뉴스페이지에 있는 언론사의 인권 관련 기사의 제목과 부제이다. 포털사이트 다음(Daum.net) 뉴스페이지에는 특정 언론사의 기사만 있는 것이 아니라 대부분의 언론사에서 제공하는 기사가 있기 때문에 2013년 한 해 동안 ‘인권’ 기사를 모두 볼 수 있다. 또한 기사의 전문이 아니라 인권이 들어가 있는 제목과 기사 본문 내용에 인권 단어가 들어가 있는 기사들을 대상으로 하였다.

자료수집은 ‘인권’을 키워드로 검색하여 이루어졌으며, 다음 뉴스사이트에서 수집하였고, 수집도구는 TheScrm의 데이터 수집도구를 활용하였다. 2013년 매월 인권 관련 기사의 수는 아래의 표에 정리되어 있다.

<표 2> 2013년 ‘인권’ 기사 건수

월	빈도	월	빈도	월	빈도
1	6,608	5	8,634	9	6,713
2	5,700	6	8,870	10	7,727
3	6,029	7	8,630	11	8,904
4	7,703	8	7,266	12	9,537
합계			92,321		



<그림 8> 월별 인권 건수

2013년 1월 1일부터 12월 31일까지 각 언론사에서 ‘인권기사’ 수는 92,321건이다. 2013년 한해 중 가장 많은 인권 키워드가 언급된 달은 12월이며, 총 9,537건의 기사에서 인권이 언급되었다. 반면 가장 적게 언급된 해는 2월 달로 총 5,700건인데, 이 달에 특히 적은 것은 2월이 짧기 때문이다.

시기별로 살펴보면 인권 관련 기사가 2013년 1월에는 6,608건 있었고, 2월에는 기간이 짧아 기사 건수가 낮았으나 3월부터 증가하기 시작하여 7월 달 8,870건의 기사가 있었다. 이후 9월까지 다소 감소하다가 9월부터 증가하여 12월 달 9,537건의 기사가 있었다. 인권기사가 가장 많은 달은 12월 달로 9,537건의 기사가 한 달 동안 보도되었다. 일 평균 인권 관련 기사 수는 2월 203.6건, 3월 194.5건으로 3월 달에 가장 적었다.

2. 분석방법

기사의 수집은 TheScrm.com의 Textom¹⁹⁾을 통해 실시했고, 수집된 기사의 분석은 시멘틱네트워크 분석 기법을 활용하였다. 기간은 2013년 1월 1일부터 12월 30일까지 기사를 수집하였고, 더불어 개별 기사 제목(부제 포함)과 요약문에 등장하는 주요단어를 추출하였다. 다음으로 추출된 주요단어의 공동으로 출현하는 빈도수를 파악하고, 단어*단어의 행렬데이터로 구성하였다. 주요단어를 추출할 때 수집된 데이터에서 관사, 접속사, 문장의 어미 등은 제외했으며 유사한 맥락에서 사용된 단어들은 단일 단어로 일원화하는 과정을 거쳤다. 예를 들어, ‘인권위’, ‘국가인권위’ 등은 ‘인권위원회’로 통일시켜 빈도수를 산출하였다. 또한 ‘조례’, ‘제정’, ‘반대’, ‘개최’, ‘방안’, ‘증진’ 등은 낱말의 단어로 처리함으로 단순 주요 빈도수에 의해 누락되는 부분을 최소화하였다.

이러한 과정을 거쳐 만들어진 행렬데이터를 기반으로 단어들 사이의 연결구조와 특성을 파악하기 위해 네트워크 분석 소프트웨어인 UCINET6와 소프트웨어에 포함된 NetDraw의 기능을 활용하여 공동출현 단어들의 연결망을 시각화하였다.²⁰⁾ 이를 통해 연결망의 구조와 연결강도를 파악하였으며, 각 노드(단어)가 연결망에서 차지하는 구

19) Textom(www.thescrm.co.kr)은 대구 소재 TheIMC에서 개발한 웹상에서 수집 및 정제가 가능한 빅데이터 솔루션으로 웹과 SNS 상의 데이터를 수집하여 단어 간 매트릭스까지 일괄 처리되는 특징을 가지고 있다. 데이터 수집은 네이버, 다음, 구글, 트위터, 페이스북이며, 네이버의 경우 블로그, 카페, 뉴스, 웹문서, 지식IN, 이미지, 동영상, 전문정보(특학술자료, 특허/KS표준 등)의 채널에서 데이터를 수집할 수 있다. 또한 수집된 데이터의 정제 역시 웹상에서 이루어지며, 수집된 데이터에서 관사, 접속사, 문장의 어미 등이 자동으로 정제된다.

20) 손동원, 2002, 『사회 네트워크 분석』, 경문사.

적인 역할을 분석하기 위하여 프리만의 중심성 개념 정리와 측정 방법을 토대로 계량화된 지표인 연결중심성(degree centrality), 매개중심성(betweenness centrality) 그리고 근접중심성(closeness centrality)을 살펴보았다.²¹⁾

프리만의 중심성 개념은 연결망 안에서 각 노드의 영향력이 다양하게 나타나기 때문에 중심이 되는 노드가 생긴다고 보았다. 즉 연결망 중심 위치에서 여러 노드들과 직접적으로 가능한 한 멀리 연결되어 있다면 이 노드의 핵심적 영향력이 높다. 이러한 연결된 노드들의 합을 연결중심성이라고 정의했다. 또한 어떤 특정 노드를 통하지 않으면서 이와 연결된 노드들이 단절되어 의미가 끊어지는 경우도 있을 수 있다. 여기서 특정 노드는 노드들 간의 통제력이 높다는 것을 의미한다. 다시 말해 노드와 노드 사이에서 다리 역할을 하는 중개 영향력이라고 할 수 있으며 이를 매개중심성이라고 했다. 근접중심성은 노드와 노드의 직접적인 연결뿐 아니라 간접적인 연결 즉 하나의 노드와 연결된 모든 노드의 거리를 측정하는 것으로 근접중심성이 낮다는 것은 노드 간의 거리가 짧다는 것을 의미하며 네트워크 내에 영향력이 높다는 것을 의미한다.

따라서 연결중심성이 높다는 것은 다른 단어와 연결된 정도가 높다는 것으로 단어 간의 관계가 높다는 것을 의미한다. 반면 매개중심성은 한 단어가 다른 단어와의 네트워크를 구축하는 데 중개자 혹은 다리(bridge) 역할을 하는가로, 중개역할에 초점을 맞춘 중심성이다. 그리고 근접중심성은 단어 간의 거리(distance)가 핵심개념으로 단어 간 거리가 가까우면 쉽게 관계를 맺을 수 있다는 의미이다.

IV. 분석결과

1. 수집결과

2013년 한 해 ‘인권’이란 단어가 신문에 언급된 것은 기사에서 중요 단어를 정제하여 단어별 빈도를 분석한 값은 아래의 표와 같다.

21) Freeman, L. C., 1979, Centrality in social networks conceptual clarification, *Social networks*, No1. Vol 3, pp.215-239.

<표 3> 2013년 ‘인권’ 관련 신문기사에서 나타나는 주요 단어

단어	빈도	단어	빈도	단어	빈도	단어	빈도	단어	빈도
인권	11,659	일본	768	서울시	544	적극	404	인사	334
북한	3,686	위안부	743	학생인권조례	543	학교폭력	398	시행	332
침해	3,528	보장	738	평화	539	인권보장	398	설치	330
문제	2,825	민주당	737	추진	531	목사	395	사망	329
인권위원회	2,555	학교	737	직원	525	범죄	395	기업	329
여성	2,054	자유	733	우려	525	가치	392	러시아	327
단체	1,869	요구	725	예방	524	성소수	392	최초	326
미국	1,738	제정	709	인권교육	523	탄압	391	비난	326
유엔	1,585	피해자	698	전국	521	프로그램	391	선정	325
개최	1,467	민주	690	공개	520	발언	389	건설	324
장애인	1,460	참석	687	반대	519	사례	389	토론회	324
사회	1,386	장애	677	노동자	504	주제	387	시민단체	323
대통령	1,382	광주	667	실현	497	현재	384	설립	323
학생	1,319	인권유린	659	유린	496	가족	382	전교조	322
보호	1,243	회견	646	사실	493	복지	382	사태	322
중국	1,224	영화	642	계획	478	시장	382	국정원여직원	317
서울	1,111	지적	624	행위	475	이야기	382	이상	317
실시	1,079	행사	612	과정	472	외국인	381	일부	316
대표	1,071	새누리당	612	도시	471	사업	375	시설	315
한국	1,064	교수	610	운동	471	결정	374	행복	315
교육	1,051	시민	606	관계자	471	환경	373	송전탑	314
주장	1,041	제기	604	성폭력	470	언론	371	차원	313
활동	1,014	예정	597	존중	469	출범	363	방안	313
논란	991	내용	594	문화	463	책임	362	증진	312
개선	972	동성애	591	포함	460	대한민국	361	안전	310
운영	954	비판	585	생각	460	회장	361	처벌	310
지원	934	차별	582	교사	453	항상	359	구성	309
대상	909	필요	582	노동	447	삶	355	지속	309
탈북	883	이유	578	현장	442	시위	353	결혼식	301
진행	865	강조	572	역사	431	장관	352	실태	298
세계	863	발생	570	국정원	426	인정	351	사건	298
변호사	858	노력	569	권고	425	밀양	349	요청	296
사람	850	상황	568	조례	422	만델라	348	서울연합뉴스	292
인권운동	831	권리	566	감독	419	영국	348	확산	289
발표	814	관심	563	오늘	415	채택	347	거부	289
마련	807	방문	560	입장	410	참여	346	회원	288
광주시	801	지역	558	폭력	410	방향	346	제출	287
청소년	796	수사	557	확대	406	피해	345	고발	287
조사	795	시작	547	박근혜	405	기본	338	성명	287
주민	768	아동	545	법무부	405	위원장	335	평가	287

먼저 뉴스에서 ‘인권’ 단어로 검색했기 때문에 ‘인권’ 단어의 빈도가 11,659로 가장 높다. 다음으로 ‘북한’이 3,686개, ‘침해’가 3,528개, ‘문제’가 2,825개 등의 순으로 나왔다. 이러한 빈도의 결과를 보면 먼저 국내 인권문제를 담당하고 있는 국가인권위원회에 대한 기사의 내용이 높은 빈도(‘인권위원회’ 2,555개)를 차지하고 있는 것을 볼 수 있고, ‘여성’에 대한 언급도 2,054개로 상대적으로 높게 나타났다. 반면 낮은 빈도를 보

인 단어의 경우 ‘평가’ 287개, ‘고발’ 287개, ‘제출’ 287개 등으로 나타나 주로 인권문제를 해결하기 위한 행동에 대한 단어가 있었다.

또한 인권 문제에 있어서 ‘북한’ 인권에 대한 기사가 상대적으로 높게 나왔고(3,686), 다음으로 ‘여성’(2,054), ‘장애인’(1,460), ‘학생’(1,319), ‘탈북’(883), ‘동성애’(591), ‘아동’(545), ‘노동자’(504) 순으로 나타났다. 이 기간에 북한 인권에 대한 언급이 많았던 것은 유엔의 북한인권결의안이 채택되는 과정에 많은 보도가 있었기 때문이다.²²⁾ 따라서 ‘미국’ (1,738), ‘유엔’ (1,585)의 단어 빈도가 높은 것이다. 또한 김조광수 결혼식으로 인해 ‘동성애’ 591개, ‘성소수’ 392개, ‘결혼식’ 301개 등이 2013년 한 해 인권과 관련해 보도가 되었다.²³⁾

그 외 정당별로 인권기사와 연관을 보면 ‘민주당’(현 새천년민주연합) 737개, ‘새누리당’ 612개로 민주당이 인권과 관련해서 더 많은 발언이 있는 것으로 나타났고, ‘변호사’ 858개와 ‘영화’ 642개는 송광호 주연의 영화 <변호인> 때문에 빈도가 높게 나온 것이다. 그러나 노인 인권에 대한 단어가 없는 것은 낮은 빈도 때문에 주요 단어로 나오지 않은 것으로 노인 인권에 대한 기사가 많이 없다는 것을 의미한다.

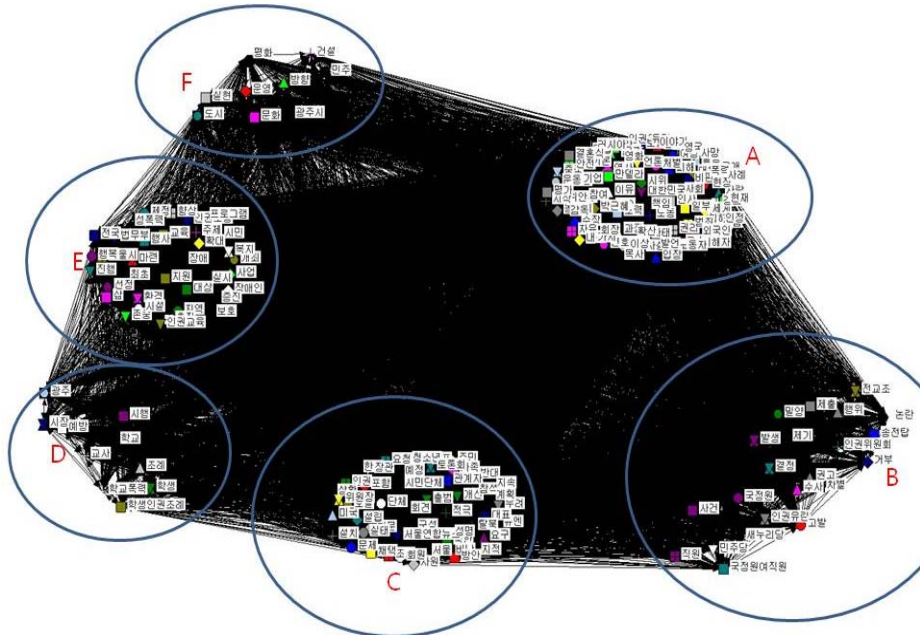
2. 주요단어 네트워크의 특성

1) 주요단어 네트워크

<그림 2>는 <표 2>의 주요 단어의 의미구조를 파악하기 위하여 총 200개 단어를 사용한 네트워크이다. 전체 네트워크에서 볼 수 있는 것은 단어가 독립적으로 존재하는 경우가 없다는 것과 네트워크의 외부에 있는 단어는 연결정도가 낮다는 것이다. 또한 <그림 2>의 경우 단어의 연결정도에 따라 크기를 나타낸 것으로 네트워크 중심에 가까울수록 기호의 크기가 크고 외부에 있을수록 작은 것을 볼 수 있다. 예를 들어 ‘인권’, ‘침해’, ‘인권위원회’, ‘문제’ 등은 주요단어 네트워크의 중심에 있으며 연결이 많은 것을 알 수 있다. 반면 전체 네트워크에서 ‘성소수’, ‘결혼식’, ‘만델라’, ‘송전탑’, ‘밀양’, ‘국정원여직원’ 등의 단어는 네트워크 외부에 있으며 크기도 작은 것을 볼 수 있다. 즉 2013

22) 연합뉴스, 2013.11.20, “‘인권상황 개선하라’...유엔 ‘북한인권결의안 채택,’”
<http://www.yonhapnews.co.kr/bulletin/2013/11/20/0200000000AKR20131120043600072.HTM>
 L?input=1179m(검색일: 2014년 5월 31일).

23) 뉴시스, 2013.08.08, “김조광수와 김승환의 ‘당연한 결혼식’ 기자회견,”
http://www.newsis.com/pict_detail/view.html?pict_id=NISI20130808_0008514439(검색일: 2014년 5월 31일).



〈그림 10〉 주요단어 네트워크의 CONCOR

CONCOR 분석을 통해 그룹을 나누면 크게 6개의 그룹으로 나누어지고, 단어가 많이 몰려있는 세 그룹과 상대적으로 적은 수의 단어가 있는 세 그룹으로 나눌 수 있다. 먼저 단어가 많은 그룹을 보면, A 영역에는 ‘인권’, ‘노동’, ‘기업’ 등 단어들이 하나의 그룹을 형성하고 있고, C 그룹은 ‘청소년’, ‘시민단체’, ‘출범’, ‘탈북’ 등의 단어들이 그룹을 만들고 있다. E 그룹의 경우 ‘서울시’, ‘성폭력’, ‘행복’, ‘장애인’ 등이 하나의 그룹이 되어 있다.

다음으로 적은 단어의 그룹을 보면, B 그룹은 ‘전교조’, ‘송전탑’, ‘밀양’, ‘국정원여직원’ 등의 단어가 그룹을 이루고 있고, D 그룹은 ‘광주’, ‘학생’, ‘교사’ 등 지역과 교육과 관련된 단어가 작은 그룹을 이루고 있다는 특징이 있다. 또한 F 그룹의 경우에는 ‘평화’, ‘민주’, ‘도시’ 등의 단어가 하나의 그룹을 형성하고 있다.

즉 CONCOR가 비슷한 행위자 간의 블록화를 분석하는 것으로 행위자를 단어로 바꾸면 2013년 한 해 동안 언론에서 인권 관련 보도는 위의 그룹별로 해당 단어들을 이용하여 보도하는 경우가 많았다는 점을 보여준다. 즉 ‘학교’와 관련해서는 ‘조례’, ‘예방’, ‘시행’과의 단어와 연관되어 보도가 많았으며, ‘국정원여직원’은 ‘새누리당’, ‘민주당’, ‘수사’, ‘권고’ 등과의 단어가 사용된 보도가 많았다는 것을 보여준다. 반면, ‘국가인권위

원회'가 '국정원여직원'과 '밀양 송진탑'과 연관되어 있음에도 불구하고 '논란', '거부' 등의 부정적 단어와 연관되어 있다는 것은 특징적이다.

2) 주요 단어의 중심성 분석

좀 더 자세하게 단어의 중심성을 분석해 본다. 중심성(centrality)은 네트워크 구조 속에서 개인, 혹은 하나의 단어가 가지는 권력과 영향력이라는 개념으로 개발되었으며, 네트워크 분석 지표 중에 가장 많이 사용되는 지표이다. 또한 중심을 측정하는 방법으로 연결정도중심성, 근접중심성, 매개중심성이라는 세 가지 유형이 있다. 이러한 값은 UCINET6를 통해 분석하고 분석한 값을 Netdraw를 사용하여 시각화한다. 또한 분석의 효율성을 감안하여 상위 20개, 하위 20개 단어의 중심성을 검토하고, 상호 비교한다²⁵⁾.

(1) 상위 20개 단어의 중심성 값 비교

<표 3>은 전체 단어들의 중심성을 상위 20개로 정리한 것이다. 먼저 연결중심성, 매개중심성, 근접중심성 모두 상위에 '인권'의 단어가 나타났으며 전체 네트워크 상에서 단어 자체적으로나 네트워크의 위치상 중심이 되는 단어인 것을 알 수 있다. 즉 '인권'은 빈도도 높은 동시에 다른 단어와 함께 전체 문장의 흐름을 주도하는 단어들이다. 이것은 기본적으로 자료를 수집하는 단계에서 키워드가 '인권'이기 때문에 높게 나온 것이다.

그러나 '인권'을 제외하면 각 중심성마다 다른 단어들이 나타난 것을 알 수 있다. '복한'은 높은 연결중심성을 가지고 있음에도 불구하고, 매개중심성과 근접중심성이 나타나지 않는다. 연결중심성이 3위인 '침해'의 경우 매개중심성은 18위, 근접중심성은 4위에 위치한다. 이다.

25) 상위 단어와 하위 단어를 구분하여 분석하는 것은 네트워크에서 먹합수 분포가 나타나기 때문이다. 먹합수 분포에서는 예외적으로 많은 링크(link)를 가지는 소수의 허브 또는 커넥터의 존재를 확인할 수 있다. 이수상, 2013, 『네트워크 분석 방법론』 (논형), 116~117쪽. 따라서 연결정도 가 높은 단어의 그룹과 낮은 단어의 그룹간의 비교를 통해 분석한다.

<표 4> 상위 20개 단어들의 중심성 값 비교

순위	단어	연결중심성	단어	매개중심성	단어	근접중심성
1	인권	54,669	인권	16.292	인권	199
2	북한	17,571	문제	16.292	문제	199
3	침해	15,798	활동	16.292	활동	199
4	문제	13,870	예정	16.215	침해	200
5	인권위원회	9,704	발표	16.094	단체	200
6	여성	9,212	필요	16.061	대표	200
7	미국	8,015	포함	16.059	한국	200
8	유엔	8,007	주장	15.981	주장	200
9	단체	7,770	한국	15.872	논란	200
10	장애인	6,923	시작	15.866	진행	200
11	사회	6,829	요구	15.765	발표	200
12	학생	6,159	진행	15.672	예정	200
13	개최	6,117	대표	15.631	필요	200
14	보호	6,040	사람	15.425	포함	200
15	대통령	5,816	시민	15.385	개최	201
16	중국	5,520	대상	15.312	사회	201
17	교육	5,465	우려	15.214	지원	201
18	광주시	5,359	침해	15.136	요구	201
19	운영	5,351	단체	15.092	시작	201
20	실시	5,337	생각	15.04	인권위원회	202

즉 앞서 ‘북한’의 경우 다른 단어와 연결된 정도가 높지만(연결중심성), 단어 간의 중개역할과 다른 단어와의 거리가 멀다는 것을 의미한다. 또한 ‘침해’의 경우 다른 단어와의 연결정도가 3위에 위치하며, 중개역할은 18위이고, 다른 단어들과의 거리는 4위에 있다는 것이다.

(2) 하위 20개 단어의 중심성 값 비교

<표 4>는 전체 단어들의 중심성을 하위 20개로 정리한 것이다.

연결중심성이 가장 낮은 ‘사망’은 네트워크 외부에 있고, 매개중심성이 가장 낮은 결혼식은 단어 간의 중개역할을 하지 못한다는 것을 의미한다. 그리고 근접중심성이 가장 높은 ‘결혼식’은 다른 단어들과의 거리가 멀다는 것을 뜻한다. 특히 하위 20개 단어의 연결중심성과 매개중심성, 근접중심성 모두 포함된 단어는 ‘전교조’, ‘영국’, ‘법무부’, ‘사망’이 있다.

‘전교조’의 경우 노동조합 불법화의 문제와 학생인권조례에 대한 기사로 인하여 ‘노동조합’이나 ‘학생인권조례’ 등과 연결정도가 있으나 그 외 단어들과 연결정도가 낮아

낮은 연결중심성을 보였다. 그 외 ‘법무부’는 인사이드에 대한 기사가 많아 연결중심성이 낮았으며, ‘영국’은 자국 내 인권 문제, 영국 가디언지 등 신문사 이름이 기사에 언급됨에 따라 다른 단어들과의 연결정도가 낮게 나온 것이다. ‘사망’의 경우 시리아 내전에 발생한 사망에 대한 기사로 인해 인권-사망의 연결정도가 나온 것이고, ‘사망’ 단어로는 연결정도가 낮게 나왔다.

<표 5> 하위 20개 단어들의 중심성 값 비교

순위	단어	연결중심성	단어	매개중심성	단어	근접중심성
181	서울연합뉴스	1,417	법무부	5.321	고발	240
182	확산	1,416	건설	5.066	법무부	241
183	위원장	1,396	문화	5.044	조례	242
184	거부	1,392	영국	4.893	방향	242
185	회장	1,381	외국인	4.875	영국	243
186	처벌	1,369	도시	4.764	인권교육	247
187	전교조	1,367	전교조	4.717	도시	247
188	고발	1,338	조례	4.623	학교폭력	247
189	성명	1,338	학생인권조례	4.165	학생인권조례	251
190	영국	1,329	인권교육	4.102	국정원	255
191	구성	1,323	학교폭력	3.919	사망	255
192	제출	1,320	국정원	3.734	건설	257
193	법무부	1,318	성소수	3.454	전교조	257
194	선정	1,301	사망	3.293	성소수	267
195	만델라	1,297	밀양	2.515	러시아	268
196	이상	1,260	러시아	2.461	밀양	272
197	평가	1,241	송전탑	2.146	송전탑	280
198	인사	1,222	국정원여직원	1.454	만델라	300
199	회원	1,188	만델라	1.084	국정원여직원	302
200	사망	994	결혼식	0.526	결혼식	323

낮은 매개중심성을 보이는 ‘결혼식’, ‘만델라’, ‘국정원여직원’ 등은 다른 단어와의 네트워크에서 중개자 역할을 하지 못하는 것으로 독립적인 위치에 있다고 할 수 있다. 특히 ‘국정원 여직원’은 많은 논란이 있었던 단어임에도 불구하고 낮은 매개중심성을 가져 인권과 관련된 단어들과의 중개는 낮은 것으로 나타났다. 또한 근접중심성도 높아 다른 단어들과의 거리가 멀다는 것을 알 수 있다. 즉 여직원의 인권 문제가 논란이 되었으나 보도의 내용은 인권문제로 초점이 있지 않았다는 것을 볼 수 있다.

V. 결론

이 글은 2013년 한 해 동안 포털사이트 다음(Daum.net) 뉴스페이지에 등록된 인권 관련 기사들의 보도가 어떻게 이루어지고 있는가를 살펴보기 위해 시멘틱네트워크 분석 방법을 활용하여 관련 기사의 제목과 내용을 중심으로 분석하였다. 특히 최근 세월호 사건에서 ‘인권보도준칙’이 제대로 지켜지지 않았던 상황에서 2013년의 인권 관련 보도의 단어 간 의미들을 찾아보고자 했다.

이를 위해 데이터 분석 도구인 TheScrm.com을 이용하여 2013년 1월 1일부터 12월 31일까지 ‘인권’ 단어가 들어가 있는 기사를 수집하였고, 조사 등 정제 과정을 거쳐 의미 있는 단어를 도출한 후 단어*단어의 매트릭스를 구성하였다.

먼저 주요 단어 빈도의 특징은 높은 순위를 차지한 단어로 ‘북한’, ‘침해’, ‘문제’ 등의 순으로 나타났으며, 주요 인권 단어인 ‘여성’, ‘장애인’, ‘학생’, ‘탈북’, ‘동성애’, ‘노동자’ 등의 순으로 빈도가 나타났으나, ‘노인’ 인권에 대한 기사는 순위에 들어가지 않은 것이 특징이다.

주요 단어의 의미구조를 파악하기 위해 전체 네트워크를 시각화한 결과는 ‘인권’, ‘침해’, ‘인권위원회’, ‘문제’ 등이 네트워크의 중심에 있었으며, 네트워크의 외부에는 ‘성소수’, ‘결혼식’, ‘만델라’, ‘송전탑’, ‘밀양’, ‘국정원여직원’, 등의 단어로 밀양 송전탑 문제와 김조광수 결혼에 대한 기사가 외부에 위치해 있었다.

다음으로 주요 단어의 연결중심성, 근접중심성, 매개중심성을 상위 20개 단어와 하위 20개 단위로 구분하여 분석하였다. 먼저 상위 20개 단어의 중심성 분석 결과의 특징은 검색어인 ‘인권’ 단어를 제외하고는 단어간 연결·근접·매개 중심성이 다르게 나타났다. 예를 들어 빈도에서 2위를 차지한 ‘북한’은 연결중심성은 높지만, 매개 중심성이 낮아 다른 단어들과의 거리가 먼 것으로 나타났다.

하위 20개의 단어는 ‘결혼식’, ‘만델라’, ‘국정원여직원’ 등의 단어는 다른 단어들과 중개 역할을 하지 못하고 독립적인 위치에 있는 것으로 나타났다. 즉 성소수자의 결혼식이었던 김조광수 결혼식은 상대적으로 높은 ‘성소수자’ 단어의 빈도에 비해 그 연결이 낮았다. 또한 2012년 대선 당시 많은 논란이 있었던 국정원여직원의 인권 문제는 인권 기사와의 연결이 낮은 것을 알 수 있었다.

이러한 연구결과를 토대로 몇 가지 시사점을 도출해보면, 인권 기사에서 사용하는 단어는 주로 침해, 문제 등 문제에 대한 기사가 많다는 것을 알 수 있다. 즉 문제가 있는 인권에 대한 기사가 많음으로 인하여 인권의 부정적 프레임을 만들고 있다. 또한 인

권 보호의 여러 영역 중에서 노인 영역이 다른 부문에 비해 관심이 적다는 것을 알 수 있다. 또한 학생 인권의 문제에서는 ‘학교’, ‘학생’, ‘학생인권조례’ 등의 다양한 단어들과의 연결이 있었지만 구체적인 블록이 보이지 않는 것으로 보아 학생인권 보호를 위한 구체적 활동이 부족하다는 것을 알 수 있다.

이 글은 2013년 전 기간 동안의 인권 기사를 대상으로 기사의 단어 간 의미를 분석했다는 점에서 의의가 있다. 또한 보도에 있어서 인권 침해에 대한 관심이 증가하고 있는 상황에서 인권 관련 보도의 방향성을 검정했다는 점에서 함의를 찾을 수 있다. 그러나 연구 결과만으로 인권 보도의 문제점을 모두 분석했다고 하기는 힘들다. 또한 단어 간 의미를 분석함에 있어 연구자의 주관적 해석을 배제할 수 없었다는 점은 문제이다. 제목과 기사 본문 전체, 기자 등 다양한 요인들을 함께 분석하여 입체적으로 분석하는 것이 필요하다.

참고문헌

- 고선연, 2013, “‘진보’와 ‘진보가 바라 본 북한 인권 문제: 한겨레신문과 경향신문 비교 분석,” ATE 제3권(이화여자대학교 스크랜튼학부), 35~55쪽.
- 국가인권위원회, 2013, “주요 언론의 인권보도준칙 준수 실태조사,” 2013년 인권상황 실태조사 연구용역사업 보고서.
- 김만재, 전방옥, 2012, “언어네트워크 분석 기법을 활용한 인간배아복제 신문보도 분석,” 『한국생명윤리학회』, 13권 2호, 19~34쪽.
- 김위근 외 3명, 2013, “한국의 인터넷 뉴스: 언론사닷컴, 인터넷신문 그리고 포털 뉴스 서비스,” 한국언론재단 연구보고서 2013-04.
- 남인용, 박한우, 2007, “대권 예비후보자 관련 신문기사의 네트워크 분석과 홍보전략,” 『한국정당학회보』, 6권 1호, 79~107쪽.
- 박성희 2009, “제17대 대통령 후보 합동 토론 언어네트워크 분석,” 『한국언론정보학보』, 45호, 220~254쪽.
- 박지영·김태호·박한우, 2013, “의미연결망 분석을 통한 셀러브리티의 SNS 메시지 탐구,” 『방송통신연구』, 82호, 36~74쪽.
- 박한우 & Leydesdorff, 2004, 한국어 내용분석을 위한 KrKwic 프로그램 이해와 적용, 『Jornal of The Korean Data Analysis Society』, 6권 5호, 1377~1388쪽.
- 백선기·이옥기, 2013, “재난방송 보도에 대한 국가별 채널간 보도태도의 비교연구,” 『한국언론학보』 제57권 1호(한국언론학회), 272~304쪽.
- 손동원, 2002, 『사회 네트워크 분석』, 경문사.
- 심홍진·김용찬·손해영·임지영, 2011, “언어 네트워크 분석을 통한 스마트폰과 소셜 미디어 이용자의 미디어 이용행태에 관한 탐색적 연구,” 『한국방송학보』, 25권 4호, 82~138쪽.
- 이수상, 2013, 『네트워크 분석 방법론』, 논형.
- 이승선, 2012, “섬세하고 엄밀한 취재 없다면 차라리 보도하지 말아야 : 인권보호 등 성범죄보도 개선방안,” 『신문과 방송』 제503호(한국언론진흥재단), 28~32쪽.
- 장하용, 2001, “언론보도와 비평의 구조: 신문보도의 비평에 대한 네트워크 분석,” 『한국언론정보학보』, 16권, 108~133쪽.
- 최윤정·권상희, 2014, “‘빅데이터’ 관련 신문기사의 의미연결망 분석,” 『사이버커뮤니케이션학보』, 제31권 1호, 241~286쪽.

한국언론재단, 2013, “언론수용자 의식조사-제18회 <미디어 환경변화에 따른 이용자
행태조사> 조사분석,” 한국언론재단 조사분석서.

Freeman, L. C., 1979, Centrality in social networks conceptual clarification,
Social networks, No1. Vol 3, pp.215-239.

Textom(www.thescrm.co.kr)