

잠재디리클레할당 분석을 이용한 ‘노인일자리’ 관련 신문기사 토픽분석

이소정

남서울대학교 노인복지학과 조교수

Topic Modeling of Newspaper Articles on Government ‘Senior job program’ via Latent Dirichlet Allocation.

So-Chung Lee

Namseoul University, Department of Elderly Welfare, Assistant professor

요 약 본 연구는 노인일자리사업의 사회적 논의구조를 분석하기 위해 대표적인 대중매체인 신문기사에서 다루어지는 노인일자리 관련 주요 토픽들과 시계열적 특성을 분석하였다. 이를 위해 뉴스 통합 데이터베이스인 빅카인즈에 수록된 11개 중앙지와 8개 경제지의 노인일자리사업 관련 기사 1107개에 대해 잠재디리클레할당 방법을 이용한 토픽분석을 실시해 언론 기사에 내재된 노인일자리사업의 잠재토픽을 추출하였다. 분석결과 노인일자리사업에 대한 일반적 정보전달, 지자체 사업 홍보, 노후생활, 고용효과, 시장연계 등 5개의 잠재토픽이 추출되었는데 2015년까지 대부분의 언론 기사가 일반적 정보전달과 지자체 사업홍보에 국한되어 있어 노인일자리사업의 정체성에 대한 사회적 논의가 형성되지 못하였음을 알 수 있었던 반면 2015년 이후부터 노인일자리사업의 소득, 안전 등 노후생활 효과 관련 주제가 다루어지는 비중이 증가했으며 특히 문재인 정부 출범이후 고용효과와 관련된 기사가 압도적인 비중을 차지하게 되었음을 발견할 수 있었다. 본 연구는 이러한 결과에 근거해 향후 노인일자리사업의 질적측면 및 고용효과 측면을 증진시킬 수 있는 방안에 대한 고민의 필요성과 고용프레임 이외의 대안적 프레임 제시의 필요성을 제안하였다.

주제어 : 노인일자리, 사회적 논의구조, 신문기사, 토픽분석, 잠재디리클레할당

Abstract This study aims to find the structure of social discussion on government ‘Senior job program’ by analyzing 1107 newspaper articles on ‘senior job program’ from 11 major newspaper articles and 8 financial newspapers. Topic modeling via latent dirichlet allocation model was employed for analysis and as result, 5 latent topics were extracted as follows : general information, local government project propaganda, senior life related issues, employment creation effect and market relations. Until 2015, most of the articles focused on the first two topics, indicating not much discourse was formed concerning the characteristics of the program. However, after 2015, the third topic started to increase and after the launch of Moon Jae In government, there has been a drastic increase in the employment creation related topic indicating that current social discourse mirrored by the media is definitely focused on employment creation aspect of senior job program. Based on the result, this study suggests the necessity to increase the quality and also enhance employment aspects of Senior job program.

Key Words : Senior Job Program, Social discussion, newspaper articles, topic model, latent dirichlet allocation

*Funding for this paper was provided by Namseoul University year 2019

*Corresponding Author : So-Chung Lee(sjdswn@naver.com)

Received September 3, 2020

Revised September 25, 2020

Accepted October 20, 2020

Published October 28, 2020

1. 서론

우리나라는 2000년대 이후 급격한 저출산·고령화 현실에 대해 자각하며 이에 대응하고자 하는 대책을 마련하였다. 이러한 가운데 2004년 노인일자리사업이 처음 도입되었다. 노인일자리사업은 공적 소득보장제도가 미비한 가운데 빠른 속도로 증가하는 노인들을 대상으로 일을 통한 소득보전과 사회참여라는 두 마리 토끼를 잡기 위한 정책으로 도입되었다. 2004년 도입당시 3만5천여개의 노인일자리가 창출되었는데 이후 사업이 급격히 확대되어 2019년말 기준 약 64만개로 15년간 20배 이상 확대되었다.

이처럼 노인일자리사업이 지속적으로 확대될 수 있었던 배경은 두 가지로 분석해 볼 수 있다. 첫 번째는 노인일자리사업에 대한 국민적 합의가 존재하기 때문이다. IMF 외환위기 이후 우리 사회는 경제적·사회적으로 신자유주의적인 체질 전환의 과정을 겪게 되는데 동일한 맥락에서 복지정책 역시 공짜급여보다 일에 대한 대가로 제공되는 급여가 대중적인 공감을 얻게 된다. 이러한 변화와 급속도의 고령화에 대한 사회적 인식의 확산은 ‘노인일자리’에 대한 당위성의 토대를 형성하게 되며 고령사회의 주요한 대응수단으로 여겨지게 된다. 두 번째는 노인일자리사업이 갖는 정치적 속성이다. 노인에 대한 부정적 발언으로 정치적 생명을 달리던 정치인의 사례는 익히 알려졌는데, 이러한 사례가 보여주듯 고령인구의 증가와 함께 노인들은 정치적 득표에 유의미한 영향을 미치고 실력을 행사할 수 있는 집단으로 인식되기 시작하였다. 따라서 고령화의 주요 대응책으로 여겨지고 있는 노인일자리에 대한 축소는 정치권 및 정책입안자들에 의해 정치사회적 생명을 걸어야 하는 민감한 이슈이다. 반대로 노인일자리에 대한 확대는 중요한 유권자인 노인들의 지지를 얻고 국민적 공감을 얻는 이슈이기 때문에 정치권 및 정책입안자들은 노인일자리사업을 확대시키는 것에 공감하게 된다.

그러나 규모의 확대는 막대한 예산지출을 수반하며 2020년도에는 노인일자리사업에 약 1조2천억의 예산이 배정되었다. 노인일자리 양적 확충과 예산의 확대는 필연적으로 정책의 책임성과 효과성에 대한 관심을 불러일으킬 수 밖에 없으며 이러한 이유로 ‘과거기식 복지’ ‘통계착시’ ‘용돈벌이’ 등 노인일자리사업에 대한 비판적 여론이 확산되고 있다.

정부는 내년도 예산을 확대해 노인일자리 80만개까지 늘리겠다는 목표를 제시하였다(파이낸셜뉴스 9월 1

일자). 지금까지 큰 반대여론 없이 노인일자리사업이 성장해 왔다면 앞으로는 노인일자리사업에 대한 찬반 논쟁이 더욱 본격화될 것으로 전망되며 사업의 효과성에 대한 잦은 지적이 예상될 것으로 보인다. 노인일자리사업이 덩치만 큰 사업이 아니라 보다 사회적 수요와 기대에 부응하는 정책으로 발전해 가려면 무엇보다 노인일자리사업에 대한 사회적 논의구조를 파악하고 이에 근거해 사업을 내실화시킬 수 있는 방안을 마련하는 것이 필요하다.

본 연구는 노인일자리사업에 대한 사회의 논의구조 및 그 변화의 흐름을 분석함으로써 향후 노인일자리사업의 발전방향에 대한 제언을 하고자 한다. 특정 주제에 대한 사회적 논의구조를 분석하는 작업은 양적·질적 차원을 넘나드는 복잡하고 심오한 주제이지만 본 연구에서는 대표적인 대중매체인 신문기사에서 다루어지는 노인일자리 관련 주요 토픽들과 시계열적 특성을 분석하는 것에서 시작하고자 한다. 기본적으로 미디어는 그 자체 복잡한 담론적 구성물로서 특정한 권력과 지식이 배설되는 장이기 때문에 특정 주제에 대한 광범한 사회적 논의구조를 때로는 선도하고 때로는 반영하기 때문이다[1]. 종합적으로 본 연구의 연구문제는 “노인일자리사업에 대한 신문기사에 담긴 잠재적 토픽은 무엇이며 시계열적으로 어떤 변화를 보이는가?”로 구체화될 수 있겠다.

2. 이론적 배경

2.1 노인일자리사업

노인일자리사업은 노인의 4품을 해소하고 다가오는 고령사회를 맞이하여 우리사회의 지속가능성을 높이기 위한 수단으로서 2004년에 도입되었다. 특히 국민연금 제도를 비롯한 노후소득보장 기제가 충분히 구축되지 않음으로 인해 우리사회는 OECD 국가 중 가장 심각한 노인빈곤 문제에 직면했다. 이러한 맥락에서 등장한 노인일자리사업에 대한 사회적 호응도는 높았으며 양적으로 급격히 성장하게 되었다. 2004년에 25,000개의 일자리 창출을 목표로 삼았으나 목표치보다 높은 35,127개의 일자리가 창출되었으며 정부재정지출 확대에 힘입어 매년 상향조정된 목표치, 그리고 이를 초과하는 달성률을 통해 2019년에 이르러 약 64만개의 일자리가 창출되었다. 2020년 현재 74만개의 일자리 창출을 목표로 하고 있으며 보건복지부는 2022년까지 100만개의 일자리를 창출하겠다는 목표를 제시하였다[2].

한편 법에 의해 정의된 노인일자리사업의 목적은 어

르신이 활기차고 건강한 노후생활을 영위할 수 있도록 다양한 일자리·사회활동을 지원하여 노인복지 향상에 기여하는 것이다[3]. 이를 고려하면 노인일자리사업은 단지 소득보충을 위한 정책에 제한되지 않으며 정책 대상 또한 노년기에 일자리 및 사회활동 참여 욕구를 가진 모든 노인을 대상으로 한다고 볼 수 있다. 이는 정책의 출발 이래 양적 팽창의 정책기조를 유지해온 노인일자리사업에 일정한 딜레마로 작용하는데, 한편으로는 100만개까지 확대해도 결국 욕구를 완전히 충족시킬 수 없으며 다른 한편으로는 양적 팽창에 주력하는 동안 상대적으로 질적인 부분은 간과되어 노인일자리 질적 발전에 대한 비판에 직면하게 된다는 것이다.

노인일자리사업은 2015년에 이르러 사업 명칭을 “노인일자리 및 사회활동 지원사업”으로 변경하고 일자리의 유형구분을 공익활동, 재능나눔활동, 취업활동, 창업활동, 경력유지활동으로 세분화시키게 된다. 여기서 노인일자리사업은 근로로서의 속성을 띄는 사업과 비근로, 즉 사회활동으로서의 속성을 띄는 사업을 명확하게 구분해 양자를 포괄해 진행하는 방향성을 표방하게 된다. 또한 2019년에는 ‘사회서비스형’이라는 새로운 유형을 도입, 시범사업을 실시하게 되는데 이는 참여 노인의 연령을 하향 조정하고, 사회서비스 영역에 종사할 수 있는 보다 전문성을 갖춘 양질의 일자리 개발의 취지로 볼 수 있다.

이처럼 노인일자리사업의 출발은 ‘복지적 욕구’에서 기인했으나 이후 노인의 근로 및 사회참여의 중요성이 부각됨에 따라 사회참여적 요소가 가미되어 노인일자리사업은 경제적 요소와 사회참여적 요소 모두를 갖는 사업이 되었다. 다시말해 노인들의 소득과 관련된 문제를 소득보장제도으로써 해결하지 못하고 일자리창출과 연계해 해결하고자 한 시도로서 노인일자리사업은 경제활동과 사회참여활동 사이의 애매한 경계를 유지한 채 양적으로 확대 발전되어 온 것으로 볼 수 있다. 이러한 상황 속에서 노인일자리사업의 명확한 정체성과 정책목표는 규명되지 않은채 양적 팽창을 거듭해 온 결과 막대한 정부 예산이 투입되는 거대한 노인복지정책으로 성장했지만 노인의 빈곤문제 해결, 안정된 양질의 일자리 창출, 사회활동을 통한 행복한 노후생활 가운데 어떤 것도 확실한 성과를 보여주지 못하는 애매한 정책의 면모를 벗어나지 못하고 있다.

2.2 선행연구 검토

지금까지 직접적으로 노인일자리사업에 대한 언론 기사분석을 수행한 연구는 없었지만 노인에 대한 기사분석

을 수행한 연구와 언론을 통해 나타난 기타 정책에 대한 사회적 논의구조 혹은 사회적 담론에 대한 연구는 진행된 바 있다.

먼저 한경혜외(2007)는 3개의 중앙일간지에 1997년에서 2006년 사이 ‘노인’ 또는 ‘고령화’를 포함하고 있는 1725개의 신문기사에 대한 분석을 수행하였으며 ‘성공적 노화’ 담론을 통해 언론이 노인에 대한 긍정적 프레임을 증가시켜온 과정을 보여주었다[4]. 김은준(2017) 역시 미디어에 재현된 노인담론을 분석하는데 신문기사 뿐만 아니라 텔레비전(리얼리티 프로그램 및 드라마), 정부정책에 나타난 담론 등 포괄적 의미의 ‘미디어’ 분석을 실시한다. 분석결과 미디어를 통해 사회경제적 주체로서 기능을 다하는 노인이 환영받는 시대의 노인상으로 재현되고 있으며 이는 국가와 사회의 돌봄 책무를 축소시키고 현실의 많은 노인들을 이상적 담론의 중심부로부터 밀어내 주변화된 타자로 고착시키고 있음을 주장하였다[5]. 이병호외(2019)는 소셜 빅데이터를 활용해 노인관련 담론을 분석하는데 고령사회 노인 삶과 관련된 주제어사전을 이론적 근거에 의해 먼저 구축하고 이에 대해 웹크롤링을 통해 시기별로 주제어 언급 빈도 추세를 분석하였다[6]. 이 연구는 “노인실태조사”라는 기존 노인 대상 설문조사로 확인할 수 없는 온라인이라는 공론장에서 일어나는 사회 전반적인 여론을 분석하고자 한 연구로써 수동적인 정책의 수혜자에서 능동적인 정책의 수요자로 바뀌고 있는 담론구조를 도출했다. 오현정 외(2019)는 노인에 대한 사회적 낙인 혹은 연령주의를 야기하는 사회적 기제로서 뉴스 미디어의 역할을 분석하고 있다. 2010년부터 2017년까지 8개 종합일간지 속 노인관련 뉴스 기사를 검색하여 1만6713건의 기사 중 무작위 표집을 통해 선정한 1567건의 기사를 내용분석하였으며 노인에 대한 보도경향 확인을 통해 노인에 대한 고정관념을 형성하는데 있어 언론이 특정한 영향력을 미치고 있다고 추론했다[7].

한편 여타 정부 정책에 대한 미디어의 프레임 및 담론 구조를 분석한 연구들도 존재하는데, 대표적으로 이재무 외(2018)의 문재인 정부 복지정책 및 예산에 대한 언론 프레임 분석 연구가 있다. 문재인 정부가 수립된 2017년 5월 11일 이후 180일간 ‘복지정책’ ‘복지예산’을 키워드로 하는 진보 및 보수로 구분되는 언론의 뉴스기사와 사설을 수집하였다. 분석 결과는 언론의 이데올로기적 성향에 따라 복지정책 혹은 예산에 대한 특정 프레임이 구축되고 있음을 보여주고 있다. 유수동 외(2017)는 사회적 이슈를 정부가 정책에 반영한다는 점에 주목해 아동복지

이슈에 대한 언론의 관심도와 정책결정에 대해 분석하고 있다. 김대중-이명박 정부 기간 동안 크게 관심을 받았던 아동복지 이슈에 대해 언론의 관심도를 측정한 결과 언론의 관심도가 증가할 때 이와 관련된 정책이 결정됨을 제시해 정책의제와 언론 관심도의 상관성을 밝혔다. 최인규 2012는 기초노령연금제도 결정과정에서 나타난 정책 이슈 확산과 언론보도형태에 대해 분석했는데 정부의제에 따라 언론 보도량 및 언론보도 프레임 구성에 변화가 있었음을 보여주고 있다. 김윤민(2016)은 언론이 형성한 국민기초생활보장 수급자 담론에 대해 분석하는데 보수 언론과 진보언론에서 수급자에 대해 생산하는 담론질서가 다르다는 사실을 제시하며 지배담론이 형성한 수급자 특성의 왜곡과 과장 문제를 공론화하고 있다[8-11].

이상의 연구들은 다양한 관점과 주제를 채택하고 있지만 종합해 보면 특정한 정책의제 혹은 정책 대상에 대해 언론이 형성하는 프레임, 담론, 논의구조가 존재하고 있음을 보여준다.

3. 연구방법

3.1 자료수집

본 연구의 분석에 활용된 신문기사는 통합뉴스검색사이트인 “빅카인즈”에서 수집되었다. 빅카인즈는 1990년부터 오늘날까지 54개의 뉴스미디어에 수록된 디지털 뉴스 콘텐츠 및 이에 대한 빅데이터 분석 결과를 제공하고 있다. 그러나 저작권 문제로 인해 기사 전문은 제공되지 않으며 기사에서 불용어가 제거되고 명사형태로 추출된 키워드와 형태소 분석 결과 가중치 상위 50개 키워드 등이 제공된다[12]. 본 연구에서는 전자의 키워드를 분석에 활용하였는데, 기사 본문에 수록된 대부분의 명사가 조사, 어미 등이 제거된 상태로 포함되어 있어 기사 원문과의 일치도가 높기 때문이며 오히려 텍스트마이닝 과정에서 필수적으로 수행해야 하는 불용어 제거, 명사추출 등의 작업을 덜어주기 때문이다.

2004년부터 2019년 사이에 기사 제목에 “노인일자리”를 포함하고 있는 11개 중앙지와 8개 경제지의 모든 기사들이 추출될 수 있도록 검색 알고리즘을 지정하였다. 검색 결과 1109개의 기사가 추출되었는데 이 가운데 사 진 기사를 제외하고 최종 1107개의 데이터가 분석에 활용되었다.

3.2 분석방법

본 연구는 텍스트마이닝 기법을 통해 분석을 수행하였으며 데이터 클리닝, 전처리, 분석의 전 과정을 위해 R 버전 3.6.2를 활용하였다. R 패키지 가운데 구체적으로 dplyr, ggplot2, KoNLP, lda, ldatuning, stringr, tidytext, tm, topicmodels(알파벳 순) 패키지를 활용하였으며 각 패키지의 활용 맥락은 다음과 같다.

먼저, 본 연구의 분석 과정은 세 단계로 구분될 수 있는데 첫째, 수집된 텍스트에 대한 전처리 과정, 둘째, 토픽분석, 셋째, 추출된 토픽의 시계열적 변화 분석 등이다.

데이터의 전처리 과정이란 일반적인 양적연구에서 변수 리코딩 및 데이터 클리닝 과정에 비유될 수 있는데 한마디로 비정형화된 텍스트 자료를 텍스트 마이닝에 적합한 정형화된 자료로 변형하는 과정을 의미한다. 즉, 수집된 텍스트에는 분석에 활용될 명사뿐만 아니라 각종 부호, 숫자, 공란 등이 포함되어 있는데 이를 제거하여 분석에 투입될 텍스트를 정제하는 작업이라고 볼 수 있다 [13]. 그러나 양적분석과 비교해 텍스트 마이닝에서 전처리 과정은 매우 복잡하고 중요하며 전처리 방식에 따라 분석결과는 지대한 영향을 받는다. 따라서 텍스트마이닝 기법을 활용한 연구에서 수집된 텍스트에 대한 전처리 작업을 비교적 상세하게 소개하는 과정은 연구의 타당성 검증을 위해 필수적이다. 본 연구는 R의 tm 패키지 및 stringr 패키지를 활용하여 각종 부호 및 숫자, 공란등을 제거하는 작업을 수행하였다. 텍스트마이닝은 기본적으로 문서를 구성하는 단어와 단어를 구성하는 최소단위인 형태소를 양적으로 분석함으로써 문서(혹은 문서의 집합체인 말뭉치)의 특성을 규명하기 위한 방법론이다. 특정 문서 안에서 단어(또는 형태소)가 차지하는 중요도를 알 수 있는 가장 기초적인 방법은 단어의 등장빈도라고 생각할 수 있다. 그러나 분석대상인 말뭉치를 구성하는 문서들은 일반적으로 같은 주제에 관한 것이며(이 연구의 경우 ‘노인일자리’) 따라서 문서간 매우 높은 빈도로 공유되는 단어들(불용어가 아닌)이 존재하게 된다. 이러한 단어들은 말뭉치에서 토픽을 추출하는 과정을 교란하게 된다. 따라서 말뭉치 내에서 높은 빈도로 공유되고 있는 단어를 파악해 제거해 주는 작업이 필요한데 본 연구에서는 단어빈도-역문서빈도(Tf-idf)를 계산하는 방법으로 이 문제를 해결하였다. Tf-idf는 ‘특정 단어의 빈도와 중요도는 부적인 관계가 있다’는 지프의 법칙(Zipf’s law)에 근거하는데 말뭉치 내의 문서간 다빈도로 공유되는 단어의 경우 0 또는 0에 가까운 tf-idf 값이 도출된다

[14]. 본 연구에서는 tidytext 패키지를 활용해 말뭉치 내 단어들의 tf-idf 값을 계산했으며 계산 결과 0.01보다 작은 tf-idf 값을 갖는 8개의 단어를 제거하였다(제거된 단어는 일자리, 어르신(들), 노인들, 참여자, 고령자, 보건복지부, 사업단, 서비스 등이다). 또한 노인일자리에 관한 기사 중 다수가 지자체의 사업홍보 또는 성과에 관한 것이었기 때문에 텍스트에 등장하는 기초·광역 지자체의 명칭을 제거하였다.

두 번째 분석과정은 본 연구의 핵심인 토픽분석을 통한 잠재토픽을 추출하는 과정이다. DiMaggio et al.은 토픽분석을 통해 사회제도 또는 사회정책을 둘러싼 언어적 맥락을 유용하게 규명할 수 있음을 제시했다[15]. 특히 4차산업혁명 시대에 정보화기술이 비약적으로 발전함에 따라 디지털화되는 텍스트 양이 방대해지고 그것이 함축하는 사회적 의미의 중요성이 증가하는 상황 속에서 토픽분석은 광활한 언어의 바다에서 복잡성을 축약하고 해석을 부여할 수 있게 해준다는 측면에서 방법론적 유의성이 더욱 커지고 있다.

토픽분석을 위해 다양한 분석기법이 활용될 수 있는데 잠재디리클레할당 모형(Latent Dirichlet Allocation model, 이하 LDA)은 가장 널리 알려진 모형이다. LDA 모형의 기본 가정은 토픽분석에서 분석대상인 모든 문서는 말뭉치에 배속(embedded)되며 문서는 일련의 단어들의 집합이며 단어는 잠재된 특정 토픽을 나타내는 관측치라는 것이다. 이러한 가정에 기반해 LDA는 말뭉치를 구성하는 문서간 공기(co-occur)하는 단어 구조를 규명하는 확률모형이다[16-19]. 본 연구에서는 R의 topicmodels 패키지의 LDA() 함수를 이용해 분석을 수행하였다.

한편 토픽분석에서 가장 도전적인 과제는 말뭉치에 잠재된 토픽의 개수 k 를 정하는 것이다. k 값을 투입하는 방법은 연구자마다 상이할 수 있는데 가장 대표적인 방법은 이론에 바탕을 둔 연구자의 추정에 의해 투입되는 것이다. 그러나 이러한 가정은 “연구자의 자의성”에 대한 비판에서 자유로울 수 없다[20]. 이런 이유로 인해 연구자 추정에 근거하여 k 를 설정하지 않고 텍스트 데이터를 기반으로 최적의 k 를 도출하는 방법이 개발되어 사용되고 있다[13]. ldatuning 패키지의 FindTopicsNumber() 함수를 이용해 추정할 수 있는데 k 값의 적합도를 보여주는 Griffiths2004, Deveaud2014, CaoJuan2009, Arun2010 등 네 개 지수를 제시한다. 그러나 이처럼 데이터를 기반으로 얻는 k 값은 텍스트의 성격이나 해석을 근거로 얻은 것이 아니기 때문에 그 자체로 완벽하다고

할 수 없으며 무엇보다 알고리즘에 따라 추출되는 4가지 지수들이 제각각이기 때문에 연구자의 자의성을 완벽하게 제거할 수는 없다. 백영민(2019)은 이러한 이유로 데이터를 바탕으로 최적의 k 값들의 범위를 설정한 후 텍스트에 대한 연구자의 합리적 추론을 바탕으로 가장 적절한 k 값을 선택할 것을 제안하고 있다[13].

마지막으로 추출된 잠재토픽의 시계열적 변화 과정을 살펴보기 위해 노법래 외(2019)가 활용한 분석모형을 차용하여 각 문서를 특정 토픽에 할당한 후 토픽의 시계열적 변화가 분석되었다[21]. 구체적으로, 각 문서(기사)는 LDA 분석 결과 추정된 감마(gamma)값에 따라 0.8 이상의 감마값을 갖는 토픽에 배정되었으며 0.8 이상의 값을 갖는 토픽이 없을 경우 토픽 할당에서 제외시켰다.

4. 분석결과

4.1 잠재토픽 추출

가장 적합한 잠재토픽의 개수를 결정하기 위해 ldatuning 패키지의 FindTopicsNumber() 함수를 이용해 도출한 4개의 모형 적합도 지수(Griffiths2004, Deveaud2014, CaoJuan2009, Arun2010)는 Fig.1과 같다.

Griffiths2004와 Deveaud2014는 지수값이 클수록 LDA모형의 k 값이 더 적절한 잠재토픽개수가 되는 반면 CaoJuan2009와 Arun2010의 경우 지수값이 작을수록 적절하다. Fig.1의 상부는 k 값의 범위를 좁혀나가기 위해 실시한 1단계 분석 결과이다. k 값에 2부터 20 사이의 값을 투입한 결과 k 값이 14이후부터 Griffiths2004 지수값의 증가량이 거의 없어 적절한 잠재토픽의 개수는 2와 14사이에 있을 것으로 전제하고 2단계 분석을 실시하였으며 그 결과는 Fig.1 하단의 그림과 같다. 4개의 지수값이 일관된 방향으로 전개되지 않아 연구자의 판단에 의해 가장 적절한 k 값을 선택하게 되는데 4가지 지수를 가장 잘 수렴하는 적절한 잠재토픽 개수는 5와 7사이의 숫자임을 알 수 있다. 본 연구에서는 이 가운데 잠재토픽 개수를 5로 선택하였는데 그 이유는 “노인일자리사업”이 추상적인 개념이 아니라 개념적으로 명확한 국가 정책이며 정책의 역사도 제한되어 있어 관련된 잠재토픽의 개수가 많지 않을 것으로 판단되었기 때문이다.

잠재적 디리클레 할당(이하 LDA) 분석을 통해 분석된 5개 토픽의 핵심어들은 Table 1.에 제시된 바와 같다.

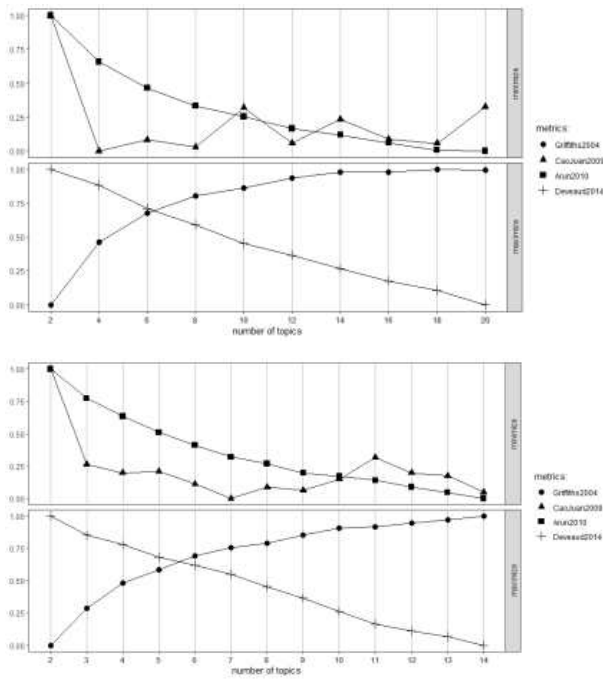


Fig. 1. Four Model Fitness Scores of k Topics(First and second estimation)

표에는 토픽별 베타값이 가장 높은 상위 10개 핵심어들이 제시되어 있다. Sutherland et al.은 LDA모형이 혼합모형(admixture model)이기 때문에 동일한 핵심어가 상이한 토픽에 반복적으로 등장하는 것이 가능함을 제시했다[22]. 따라서 핵심어만으로 각 토픽의 주제를 규명하기에는 제한적일 수 있으며 토픽을 대표하는 기사의 제목과 내용을 동시에 검토하는 것이 토픽의 주제를 밝

히는데 유용하다. Table 2는 토픽별 대표 기사문의 제목과 그 감마값을 보여준다. 감마값은 분석에 포함된 특정 텍스트가 각 토픽에 속하는 정도를 0~1의 값으로 표현한 것이며[21] 각 텍스트가 소속 토픽에 대해 대표성이 있는지 판단할 수 있는 준거가 된다.

첫 번째 토픽은 창출, 인턴십, 관계자, 제공 등의 핵심어를 포함하고 있는데 대표 기사들의 기사제목과 연결시켜 보면 첫 번째 토픽이 노인일자리사업에 대한 일반적 소개 혹은 사업량 변동 등의 ‘정보전달’ 성격을 가진 주제임을 알 수 있다.

두 번째 토픽의 핵심어는 구청장, 지자체, 발대식 등인데 두 번째 토픽으로 분류된 기사들은 대부분 특정 ‘지자체’에 국한된 기사들이었다. 노인일자리사업이 중앙정부의 사업이지만 사업시행의 핵심주체는 지자체이기 때문에 사실상 노인일자리사업에 관한 기사 중 다수는 지자체의 활동 및 업적에 대한 홍보성 기사들이다. Table 2.의 대표 기사 제목도 이와 같은 2토픽의 특성을 보여준다. 따라서 두 번째 토픽의 주제는 ‘지자체 사업홍보’로 정의할 수 있겠다.

세 번째 토픽에는 노인사업, 공익활동, 안전사고, 노후 생활, 활동비, 소득 등의 다소 넓은 범주의 핵심어들이 포함되어 있다. 대표 기사와 연결시켜 분석해 보면 세 번째 토픽에는 노인일자리 사업이 노인들의 소득, 생활에 미치는 영향 및 이와 관련된 이슈들로 구성되어 있음을 알 수 있다. 즉, 노인들의 노후소득 향상 수단으로서 노인일자리 사업의 효과성에 대한 기사, 일자리 사업 추진과정에서 나타나는 안전관련 이슈 등이 세 번째 토픽에 속한 기

Table 1. Topic Solution

Topic1		Topic2		Topic3		Topic4		Topic5	
keyword	Beta	keyword	Beta	keyword	Beta	keyword	Beta	Keyword	Beta
creation	0.012	head of Gu	0.024	elderly program	0.020	employed	0.042	exhibition	0.035
internship	0.009	local government	0.017	civic serice activity	0.015	statistics	0.017	silver delivery	0.017
related person	0.008	kick-off	0.016	safety accident	0.009	employment rate	0.012	program	0.012
provision	0.008	supporter	0.015	later life	0.009	unemployment rate	0.012	curriculum vitae	0.010
social	0.005	manpower dispatch type	0.013	activity expense	0.008	employment indicator	0.011	silver cate	0.009
continual	0.005	Education type	0.013	income	0.007	point	0.009	products	0.008
activity expense	0.005	civic serice type	0.011	basic education	0.007	unemployed	0.009	business agreement	0.007
stable	0.004	welfare type	0.011	activation	0.006	workers	0.008	consulting	0.007
labor cost	0.004	related person	0.011	social activity	0.006	workplace	0.007	mou	0.006
execution	0.004	carry out organizations	0.007	safety education	0.005	basic pension	0.006	companies	0.006

Table 2. Newspaper headlines of each topic

topic	gamma	headline
1	0.998	Ministry of welfare creates 60thousand senior jobs via community senior centers
	0.997	1.1million senior jobs created this year
	0.993	Senior job program increases older people's self-support
	0.916	welfare type jobs recruitment under process by senior welfare centers
	0.901	Government plans to increase senior jobs to 8 million by 2022
2	0.997	Seoul government creates 10thousand senior jobs, public work ttype and educational type doubles than last year
	0.994	Cangheung Dam becomes clean thanks to senior job program
	0.994	Kyunggi-do creates 27thousand senior jobs this year
	0.988	Eunpyeong-gu awards the best local government for senior job program this year
3	0.805	Gu-governments in Seoul propells senior jobs to enable senior citizens endow pocket money to grandchildren
	0.996	Senior job program turns out to be a condescending policy by giving as much as 2million won for 7 months
	0.996	Senior jobs should be increased as a source of old age incme
	0.991	People injured during Senior job program increases to 264 during the last 4 years
	0.988	Senior job program contributes to enhancing happiness of older people in an aging society
4	0.982	2million won for more than 10 years ; senior job payment is a 'passion pay'
	0.998	While 1.42million senior jobs increased, 3.2million youth employment decreased
	0.998	Despite senior job increase, the number of unemployed elders is largest ever
	0.997	Kim "Just an albeit work" vs Cho"Should not denigrate positive effects of senior jobs"
	0.997	The so-called employment rate improvement is an illusion caused by senior jobs
5	0.899	Employment rate in June is highest within the two decdes, thanks to senior jobs
	0.997	Senior job exhibition starts nation wide
	0.996	the FKI Appointed as senior work consultation organization
	0.993	Ministry of welfare created elderly jobs in cooperatino with companies
	0.880	Homeplus awarded by the Minister of Welfare for contributing to senior work program
	0.835	KORDI and SM group signs an MOU to create senior jobs in marine transportation industry

사들에 대표적으로 분포되어 있다. 따라서 세 번째 토픽은 노인일자리사업의 ‘노후생활’ 관련 토픽으로 정의할 수 있다.

네 번째 토픽의 핵심어는 취업자, 통계, 고용률, 실업률 등 노인일자리사업의 고용 관련 속성들을 표현하고 있다. 대표 기사제목도 이를 뒷받침하는데 노인일자리사업의 일자리 창출 효과 및 고용개선 효과에 대한 논의들이 주를 이루고 있다. 이처럼 네 번째 토픽은 일자리사업의 ‘고용효과’로 명명할 수 있겠다.

다섯 번째 토픽의 핵심어 역시 박람회, 실버택배, 프로그램, 이력서, 실버카페, 생상품, 업무협약, 컨설팅, mou, 기업체 등 다소 광범한 범주의 단어들 이 포괄되어 있다. 대표 기사제목을 살펴보면 이들 핵심어들이 노인일자리사업이 정부재정지원 일자리로서의 면모를 넘어 시장과 적극적으로 연계하고 협력하는 내용에 관한 것임을 알 수 있다. 따라서 다섯 번째 토픽은 노인일자리사업의 ‘시장연계’ 관련 이슈로 정의할 수 있다.

4.2 잠재토픽의 시계열적 변화

상술된 노인일자리사업의 다섯 가지 토픽의 시계열적 변화는 Table. 3와 Fig. 2에 제시된 바와 같다. 주목할 점은 노인일자리에 대한 기사의 양이 시계열에 따라 확연하게 증가하고 있음을 알 수 있는데 이는 두 가지에서 기인할 수 있다. 첫째는 노인일자리사업이 처음 시작된 2000년대에 비해 최근으로 올 수록 분석 자료인 디지털 말뚝치(디지털화된 신문기사) 자체가 증가했기 때문일 수 있으며 둘째, 노인일자리사업에 대한 사회적 관심도 증가가 기사량의 증가에 반영된 것으로 볼 수도 있다. 첫 번째 이유의 가능성을 염두에 두어야 하기 때문에 기사량의 양적 증가에 대한 지나친 의미부여는 경계해야겠지만, 특히 최근으로 올수록(문재인 정부 출범 이후) 확연한 양적 증가세는 일정부분 노인일자리사업에 대한 대중적 관심도 증가에서 기인한다고 볼 수 있다.

Table 3. Topic loading changes by year

	1	2	3	4	5
2004	1.004	1.993	1.006	0.013	2.984
2005	4.143	1.682	4.690	0.917	9.569
2006	1.044	3.140	5.610	1.825	14.380
2007	6.169	1.969	0.771	4.129	4.961
2008	3.903	6.903	1.221	4.429	14.544
2009	5.726	25.826	7.295	1.096	14.506
2010	5.950	23.469	6.920	8.330	15.332
2011	5.720	21.602	12.921	10.222	15.535
2012	17.325	36.629	10.169	3.372	15.505
2013	11.468	38.917	14.892	3.501	21.224
2014	10.493	42.023	15.543	3.471	19.470
2015	12.061	16.557	15.496	0.925	7.960
2016	23.117	30.848	27.819	5.757	21.460
2017	24.792	29.795	39.136	3.008	15.269
2018	12.981	28.070	25.533	15.668	13.748
2019	30.810	42.575	47.116	54.029	15.471

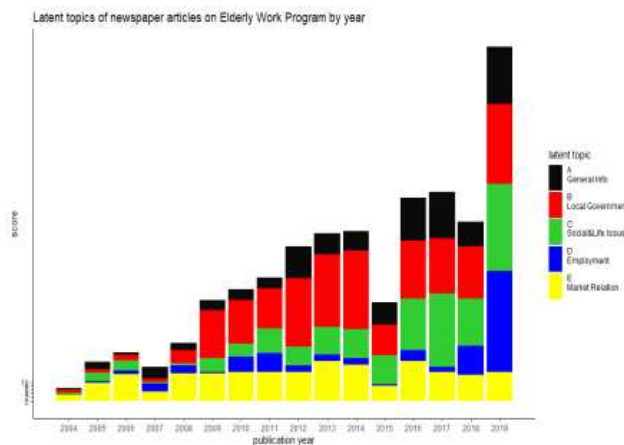


Fig. 2. Time-series dynamics of latent topics

이처럼 총량의 증가에 따라 각 토픽별 절대 수량도 시간에 따라 증가하는 경향을 보이지만 다섯 가지 토픽이 차지하는 상대적 비중에는 명확한 차이가 있다. 먼저 토픽2는 시기와 총량에 상관없이 전체 토픽 중 높은 비중을 차지하고 있는데 이것이 의미하는 바는 노인일자리에 관한 지자체의 활동 및 성과에 대한 홍보가 기사의 다수를 차지하고 있다는 것이다.

토픽1, 토픽2, 토픽5가 차지하는 비중은 시계열적으로 확연한 변화를 보이지 않고 비교적 일정한 반면 토픽3과 토픽4가 차지하는 비중에 있어서는 두드러진 변화가 나타난다. 토픽3의 경우 노인일자리에 도입되었던 초기인 2004년에서 2006년 사이(노무현 정부) 상대적으로 높은 비중을 차지하다가 2007년 이후 비중이 급감하며

상대적으로 토픽2의 비중이 증가한다. 그러나 박근혜 정부 중반인 2015년경부터 토픽1,2와 유사한 수준으로 비중이 늘어나며 2017년에는 가장 높은 비중을 차지하게 된다. 토픽4의 경우 2004년 이후 언론에서 거의 다루어 지지 않다가 2018년부터 급증하기 시작해 2019년에는 노인일자리에 관한 기사 중 가장 높은 비중을 차지하고 있다. 정리하자면 2004년 노인일자리에 시작된 이래 노인일자리에 사업량 증가, 참여자 모집 등 일반론적 정보전달에 대한 기사와 지자체 사업 홍보에 대한 기사가 주를 이루어 왔으나 박근혜 정부 중반 이후, 특히 문재인 정부 이후부터 최근까지 노인일자리에 노후 생활 및 소득, 고용창출과 관련된 이슈가 사회적 관심을 이루고 있음을 알 수 있다.

5. 결론

본 연구는 미디어를 통해 형성되고 있는 노인일자리에 대한 사회적 담론을 분석하기 위해 노인일자리 검색어로 추출된 신문기사에 대한 토픽분석을 실시하였으며 이에 대한 시계열적 변화를 분석하였다.

본 연구의 분석결과, 노인일자리에 대한 기사에 내재된 다섯 가지 토픽을 추출할 수 있었으며 이에 대한 시계열적 변화를 살펴봄으로써 노인일자리에 대한 사회적 담론의 이동을 분석할 수 있었다.

노인일자리에 시작된 2004년 이래 박근혜정부 중반까지 노인일자리에 대한 사회적 담론은 사업에 대한 일반적 정보전달과 지자체의 홍보에 머물러 있었으며 사회적 이슈가 제기되지 못했다. 그러나 박근혜정부 중반 이후부터 노인일자리에 노후소득과 관련된 이슈, 고용창출과 관련된 이슈 등 특정한 방향성을 내재하기 시작했으며 특히 최근에는 노인일자리에 고용 관련 이슈가 압도적인 사회적 관심 대상이 되고 있다. 왜 특정시점부터 노인일자리에 대한 정책에 대해 ‘소득’ ‘고용’과 같은 특정 프레임이 썩어지고 있는 것인지는 본 연구의 분석주제를 넘어서며 향후 연구로 미루어져야 할 것이다. 분명한 점은 여론은 노인들의 소득에 대한 명확한 대안으로 기능하거나 안정된 고용의 기능을 담당하는 정책대안으로서 노인일자리에 기대하고 있다는 것이며 현 정부에서는 ‘고용’의 프레임이 지배적으로 작동하고 있다는 점이다.

노인일자리에 가운데 압도적인 비중을 차지하는 ‘공익활동’은 그동안 ‘알바성 일자리’, ‘허드렛 일’이라는 비

판에 직면해 왔는데 ‘고용’의 프레임으로 바라봤을 때 향후 비판은 더욱 가속화될 것으로 보인다. 따라서 향후 노인일자리사업이 사회적 기대에 부응하는 정책효과를 거두기 위해서는 두 가지 전략 방향 가운데 선택하는 것이 불가피할 것으로 보인다. 첫째, 공익활동의 양적 확대 뿐만 아니라 질적 발전의 측면에서 적극적인 변화를 시도하고 지속적인 고용창출과 연계될 수 있는 방안을 고민하는 것이다. 둘째, 노인일자리에 대한 관점을 ‘고용’의 프레임에서 이탈시키고 새로운 관점으로 조망하게 만들기 위한 노력이다. 새로운 프레임이 무엇인지에 대한 논의는 본 연구의 범위를 벗어나지만 사업의 명칭인 “노인일자리 및 사회활동 지원사업”에서 추론해 볼 수 있는바, 노인들에게 또다른 의미를 부여하는 ‘사회활동’으로서의 가치를 부각시키고 그 필요성에 대해 사회적으로 설득하고 논의해 나가는 것은 하나의 의미있는 대안으로 제시될 수 있겠다.

REFERENCES

- [1] S. Kim. (2015). The Politics of Representing the Pain of Others : Regarding the Sewol Ferry Disaster, *Media and Society*, 23(4), 67–119.
- [2] Korea Labor Force Development Institute for the Aged, www.kordi.or.kr
- [3] Ministry of Government Legislation, www.moleg.go.kr
- [4] G. Han & S. Yoon. (2007). Critical Review of Discourse on Aging in Korean Newspapers, *Journal of the Korean Gerontological Society*, 27(2), 299–322.
- [5] E. J. Kim. (2017). How Media Makes the Elderly into Welcoming Citizens in the Aged Society, *Korean Journal of Journalism and Communications*, 61(3), 157–188. <http://doi.org/10.20879/kjcs.2017.61.3.005>
- [6] B. H. Lee & G.Y. Kim. (2019). The Themes and Trends of Discourse on the Elderly in Korea Identified by Analyzing Social Media Big Data, *Social Welfare Policy*, 46(3), 17–201.
- [7] H.J. Oh & K.A. Shin. (2019). How does Korean News Media Cover News Stories of Older Adults? A Content Analysis of Korean News Articles Published After 2010, *Journal of Public Relations*, 23(4), 40–68. DOI: 10.15814/jpr.2019.23.4.40
- [8] J. M. Lee. & Y. S. Park. (2018). A frame analysis of the conservative and progressive media on the Moon Jae In Government’s welfare policy and budget, *Journal of Budget and Policy*, 7(2), 51–80.
- [9] S.D. Yoo & J.S. Lee. (2017). Media Interest and Policy Decision of Child Welfare Issue : An Empirical Study on the Downs’ Issue–Attention Cycle, *Journal of Government Administration*, 13, 29–60.
- [10] I. K. Choi. (2012). Analysis on the types of the Reporting Behavior based on the Diffusion of Policy Issues : Focused on the Decision–making Process of the Basic Old Age Pension System, *Journal of Regional Studies and Development*, 21(1), 155–191.
- [11] Y. M. Kim. (2016). Discourse about Recipients of the National Basic Livelihood Security by the Press, *Journal of Critical Social Policy*, 53, 282–325.
- [12] Big Kinds, www.kinds.or.kr
- [13] Y. M. Baek. (2019). *Text–Mining Using R* : Hanul Academy.
- [14] Silge, J. & Robinson, D. (2017). *Text Mining with R : A Tidy Approach* : O’Reilly.
- [15] DiMaggio, P., M. Nag & D. Blei. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture : Application to newspaper coverage of U.S. government arts funding, *Poetics*, 41, 570–606. <http://dx.doi.org/10.1016/j.poetic.2013.08.004>.
- [16] Blei, D.M., A. Ng, & M. Jordan. (2003). Latent Dirichlet Allocation, *Journal of Machine Learning Research*, 3, 993–1022.
- [17] Blei, D.M. & J. D. Lafferty. (2006). Dynamic topic models. In : Pohoreckhy, A., Bottou, L., & Littman, M.L.(eds.), *Proceedings of the International Conference on Machine Learning*. 113–120.
- [18] Liu, L., L. Tang, W. Dong, S. Yao & W. Zhou. (2016). An overview of topic modeling and its current applications in bioinformatics, *SpringerPlus*, 5(1608), <http://doi.org/10.1186/s40064-016-3252-8>
- [19] McFarland, D.A., D. Ramage, J. Chuang, J. Heer, C.D. Manning & D. Jurafsky. (2013). Differentiating Language Usage Through Topic Models, *Poetics*, 41, 607–625.
- [20] Hall, D., D. Jurafsky. & C. Manning. (2008). Studying the history of ideas using topic models. In: Laputa, M., Ng, H.T.(Program Co–Chairs), *Proceedings of the 2008 Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, Stroudsburg, PA, 363–371.
- [21] B. R. Roh. & K. E. Yang. (2019). Text Mining Analysis of South Korea’s Birth–rate Decline Issue in Newspaper Articles : Transition Patterns over 18 Years, *Korean Journal of Social Welfare*, 71(4), 154–176.
- [22] Sutherland, I., Y. Sim., S. K. Lee, J. Byun & K. Kiatkawsin. (2020). Topic Modeling of Online Accommodation Reviews via Latent Dirichlet Allocation, *Sustainability*, 12(182). DOI: 10.3390/su12051821.

이 소 정(Lee So-Chung)

[정회원]



- 1998년 2월 : 서울대학교 사회복지학과 (문학학사)
- 2007년 2월 : 서울대학교 사회복지학과(사회복지학 박사)
- 2012년 3월 ~ 현재 : 남서울대학교 노인복지학과 조교수
- 관심분야 : 노인학, 노인일자리

· E-Mail : sjdsw@naver.com