

텍스트마이닝을 활용한 중국 동북공정 관련 언론기사 빅데이터 분석

— SBS드라마 ‘조선구마사’ 방영 전·후 언론기사 헤드라인 빈도분석 —

이동훈*

目 次

- I. 서론
- II. 이론적 배경
 - 2.1 텍스트마이닝
 - 2.2 동북공정
- III. 연구문제와 연구방법
- VI. 연구결과
- V. 결론 및 제언

I. 서론

2021년 3월 26일 SBS 드라마 ‘조선 구마사’에서 역사적 인물을 왜곡 표현 논란을 일으켜 방영 2회 만에 폐지되었다. ‘조선 구마사’는 같은 해 3월 22일 첫 방영 직후 드라마의 역사 왜곡 논란이 국내 각종 커뮤니티와 SNS에 거론되며, 청와대 국민청원에 드라마 방영 중지 요구가 등장하고, 곧 이어 모든 광고주가 드라마 광고를 보이콧하며 드라마 폐지로 결론지어졌다. ‘조선 구마사’는 조선 시대를 배경으로 태종과 훗날 세종이 되는 충녕대군이 악령에 맞서 벌이는 혈투를 그린 퓨전사극 드라마이다. 중국 명나라를 통해 조선에 들어온 서역의 구마사를 맞이한 식탁에 월병, 피단, 중국식 만두 등이 올랐고, 극 중 의상과 군사들이 사용하는 검이 중국 것이라는 지적을 받았다. 또한 왕으로 오를 충녕대군이 연회장 한쪽에 서서 사신을 맞이하고 월병과 중국식 만두를 대접하는 장면이 모욕적이라는 비판을 받았다. 이와 같은 현상은 동북공정(東北工程)으로 커지는 반중 정서에 불을 지피며 시청자 불만이 폭발하자 삼성전자 등 주요 광고주들 사이에서 광고 철회가 이어졌다.¹⁾

한국 고대사를 자국 역사의 일부로 편입하려고 하는 중국의 동북공정(東北工程), 한국의 사드(고고도 미사일 방어체계) 배치 이후 시작된 중국의 한한령(限韓令) 등을 겪은 국내 시청자들은 “중국이 거대 시장과 자본을 앞세워 K-콘텐츠 전반에 ‘문화의 동북공정’을 벌이고 있다며 반감을 표현했다. 한국에서 제작된 드라마에서 나타난 역사 왜곡 현상에 시청자들의 반감은 더 크게 확산하였다.

2020년을 전후로 김치 원조 논란에서 시작하여 한복, 태권도, 상추쌈 문화 등으로 이어지는 한국 문화를 중국 문화의 일부로 흡수하려는 문화공정, 전파공정 현상은 2021년부터는 매우 다양한

* 이동훈: 동아방송예술대학교 콘텐츠제작학부 영상제작과 교수

1) 조선일보(2021), “중국풍 논란 SBS ‘조선 구마사’ 2회 만에 폐지”, 3월 26일, URL: <https://n.news.naver.com/article/023/0003604178>

분야에서 빠른 속도로 전개되어 논란을 일으키고 있다. 외국인 사업자들은 대규모 사이버 폭력²⁾과 경제적 협박에 굴복, 중국 측의 요구를 수용하며 한국 문화와 관련 내용을 지우거나, 이를 중국 문화라 인정하게 되는 사례가 늘어나고 있다. 이러한 왜곡과 문화 침공이 이루어지는 대표적인 장소는 유튜브나 트위터와 같이 세계인이 한자리에 모이는 웹 사이트, 그리고 게임 등 무의식적으로 상식을 학습하기 용이한 미디어이다. 방송·유튜브·게임에서 한국 문화가 중국산으로 둔갑할 때마다 한국의 대중은 반발했다. 중국 유명 유튜버가 “한복은 중국 전통 의복 한푸의 영향을 받았다”라고 주장하자 국내 네티즌들은 이메일, 댓글 등의 집단 항의를 펼쳤고, 2021년 2월 중국 모바일 게임에서 한복이 중국 전통 복식으로 소개되자, 방탄소년단은 미국 패션지 인터뷰에서 “대한민국 옷으로서 한복을 제일 좋아한다”라고 말하며 국내 팬들의 환호를 받았다.

중국의 동북공정(東北工程)에 대해서는 이미 많은 연구가 진행됐다. 최광식(2004)은 “동북공정의 배경으로 고구려사 왜곡”에 초점이 맞추어졌고, 김은국(2007)은 “고구려와 발해 등 한국 고대사와 관련된 주제”에 대한 연구가 진행되었으며, 이남주(2004)는 “동북공정의 논리”에 대한 연구도 이루어졌다. 역사와 영토를 중심으로 진행된 동북공정의 연구와 더불어 문화와 영상 콘텐츠를 중심으로 김성수(2020)는 “네이멍구자치구의 민족문화공정, 연혁과 전망”을 통해 중국 문화공정의 목적과 연혁을 연구하였으며, 이승호(2020)는 “TV 역사드라마를 중심으로 동북공정 이후 중국 대중 미디어의 고구려 역사 서사 방식”을 통해 영상 콘텐츠를 중심으로 중국이 타국의 역사를 자국 역사로 편입하려는 시도에 관한 연구가 이루어졌다. 정호섭(2013)은 “중국의 POST 동북공정(東北工程)과 고구려사 관련 동향 분석”을 통해 동북공정의 변화 양상을 연구하였으며, 또한 중국의 문화적 팽창주의와 문화공정에 관한 연구가 진행됐다. (이장원, 2007, 이천석, 2012). 이처럼 중국의 동북공정(東北工程) 추진 배경과 과정에 관한 연구와 중국의 역사드라마 서사 방식을 통한 역사 왜곡에 관한 연구는 진행됐다.

하지만 국내 언론에서 언급된 동북공정(東北工程)과 관련된 추이와 이슈에 대한 논의는 많지 않았다. 따라서 본 연구는 국내 언론에 보도된 중국 동북공정(東北工程) 뉴스의 헤드라인을 데이터화하여 ‘텍스트마이닝’ 기법을 통해 중국의 동북공정(東北工程)이 우리 사회에서 어떻게 해석되고 의미화되고 있는지를 파악하고자 한다. 많은 연구가 텍스트 마이닝 기법을 활용하여 언론 기사를 분석했다. 김민정, 김철주(2017)는 텍스트마이닝을 활용하여 승례문 관련 기사의 트렌드를 분석하였으며, 박철수(2019)는 북한 관련 뉴스의 기간별 고찰, 황서이, 박정배(2020)는 ‘한류’ 경향에 관한 국내 언론 기사의 빅데이터를 분석했으며, 조민정, 이신행(2021)은 코로나19 관련 언론 보도 프레임 분석을 연구했다.

본 연구는 ‘중국’, ‘동북공정(東北工程)’을 키워드로 하여 2003년부터 2021년까지 신문 기사에 대한 빈도분석을 통해 ‘중국’, ‘동북공정(東北工程)’과 관련해 어떤 키워드들이 자주 나타나고 있는지에 대해 파악하고자 한다. 이를 위해 뉴스 검색·분석 서비스 빅카인즈(BIG Kinds)를 통해 2003년부터 2021년 4월 23일까지 중앙지, 경제지, 지역종합지, 방송사 등의 언론사에 검색된 중국의 동북공정(東北工程) 관련 뉴스는 총 7,223건이었다. 이중 SBS 드라마 ‘조선 구마사’ 방영 전후, 신문 기사의 주요 키워드 추출을 통해 뉴스 기사의 주요 이슈를 살펴보고자 한다. 본 연구는 ‘중

2) 사이버불링 Cyber Bullying: ‘가상공간을 뜻하는 사이버(cyber)와 집단 따돌림을 뜻하는 불링(bullying)에서 생겨난 신조어로 사이버상에서 특정인을 집단적으로 따돌리거나 집요하게 괴롭히는 행위를 말한다.’ 시사·경제용어사전, 기획재정부, 2017

국', '동북공정(東北工程)' 관련 기사에 대해 텍스트마이닝 기법을 활용하여 빅데이터를 분석하는 것에 의미를 두고 있으며 동북공정(東北工程)' 관련 기사의 동향과 정보를 제공할 수 있을 것이라 기대한다.

II. 이론적 배경

2.1 텍스트 마이닝

텍스트 마이닝은 자연어 처리 기술을 기반으로 직접적인 연관을 보여주지 않는 비정형 텍스트에서 숨겨진 관계 또는 패턴을 도출하여 의미 있고 활용 가치가 높은 정보 또는 지식을 창출하는 기법이다.³⁾

텍스트 분석에서 문서 중 단어의 중요도를 측정하는 방법은 특정한 단어가 한 문서 내에서 얼마나 자주 반복되는지, 그리고 문서 그룹 내에서 동일한 단어가 얼마나 많이 출현하는지를 측정할 수 있다. 텍스트 마이닝의 중요한 장점은 텍스트 내에서 실제적이고 잠재적으로 이루어지고 있는 메인의 논의 및 표현 내용이 무엇인지를 발견하는 것이다.⁴⁾

본 연구에서 활용한 텍스트마이닝 기법은 뉴스 빅데이터를 수집하여 의미 있는 정보를 추출하고, 용어의 빈도수와 양을 측정하여 유의미한 결과를 도출하는 것이라 할 수 있다. 언론 보도 기사에 대해 텍스트마이닝을 활용한 연구로는 감미아와 송민은(2012) 텍스트 마이닝을 활용한 신문사에 다른 내용 및 논조 차이점을 분석했으며, 윤호영(2020)은 남북 관계 뉴스 제목 44만 건을 텍스트 마이닝 분석을 통해 언론사별 논조 차이점 분석을 분석했다.

이외에도 텍스트마이닝 기법은 대량의 텍스트에 대한 분석 관련 연구로 이어져 특허정보, 학술지 논문정보, 기사, 소셜 미디어 등 다양한 분야에서 연구가 활발히 진행되고 있다.

2.2 동북공정

동북공정은 중국 동북의 역사와 현실 문제를 연구하기 위해 2002년부터 5년 동안 추진한 국가 프로젝트로 정식 명칭은 “동북변강역사와 현상계열연구공정(東北邊疆歷史與現狀系列研究工程)”이다.⁵⁾

3) 박철수, 「텍스트마이닝을 활용한 북한 관련 뉴스의 기간별 변화과정 고찰」, 2019: Judita, Stevenson, and Gaizauskas, 2015: Ronen and Daga, 1995)

4) 박철수, 「텍스트마이닝을 활용한 북한 관련 뉴스의 기간별 변화과정 고찰」, 2019: (Chakraborty et al., 2013)

5) 김지훈, 2020, 「2002년 전후 중국 동북공정의 추진과정과 추진 주체」 백산학보 117. 1, pp. 28-29

날짜	요약
2002년 2월	- 중국 정부에서 동북공정 프로젝트 시작 발표(기한 5원, 총 1,500만 위안 자금)
2003년 9월	- 국내 한겨레 신문에서 동북공정에 대해 보도 - 학자, 활동가 네티즌들 포함한 한국 각계에서 중국 정부의 동북공정에서 고구려 역사에 관한 주장을 엄하게 비판
2004년 2월	- 한중 양국 정부 모두 해당 문제를 학술적 측면에서만 토론을 진행한 것을 주장한 다는 입장 발표
2004년 3월	- 한국 정부에서 고구려연구원 개설
2004년 7월	- 북한에 있는 '고구려 무덤군'과 중국에 있는 '고대 고구려 왕조 수도 및 무덤군' 유네스코 세계문화유산에 동시 등재
2004년 8월	- 중국 홈페이지에서 고구려에 관한 언급 모두 삭제됨에 따라 한국 반한감정 - 중국 정부에서 한 유명 조선족 온라인커뮤니티 사이트를 정지시킴. 해당 사이트에서 동북공정에 대한 비판 여론 있음. - 동북공정 홈페이지 회원제 사이트로 바뀜 - 중국 외교부 차관 한국 방문 - 양국 5가지 '구두 합의' 달성
2006년 1월	- '해체된 중국이 만주를 병탄한다(中國解體吞併滿洲)'란 제목인 게시판 글이 중국 인터넷에서 유전. 이는 고구려 역사 및 영토 최고통제권이 중국 동북부에 있는 것 주장.
2006년 5월	- 주몽, 연개소문, 대조영, 대왕사신기에 포함한 일련 고대 한반도 사극 방송
2006년 9월	- 중국의 '고구려 역사 왜곡' 정지하는 것을 재촉하기 위해 한국 국회에서 결의안 채택. - 전 한국 대통령 노무현이 고구려사에 대한 중국의 견해에 전 중국 총리 온가보(溫家寶)에게 유감하다는 뜻을 표함.
2007년 1월	- 동북공정 공식 끝남, 지방정부에서 연구 지속 및 심화
2007년 2월	- 고구려의 설립자에 관한 드라마 '주몽' 홍콩 및 중국 남부지역에서 방영 - 중국 네티즌은 해당 드라마를 한국 정부 동북공정에 대한 전략적 대응으로 간주 - 고구려 문제에 관련 보도에 대한 가이드라인은 한 한국 신문지로 발표 - 중국 유명 게시판 사이트 '티엔야논단'의 '고구려바(高句麗貼吧)' 운영 정지

표 1. 동북공정 관련 일련 갈등 시간표⁶⁾

김정탁(2015)에 의하면 '동북공정이라는 용어의 사용은 2003년 9월 한겨레 신문에 보도했다'라고 거론되어 있으나, BIG KINDS 뉴스 검색을 통해 2003년 8월 중앙일보 '중국의 고구려사 빼앗기' 동북공정 연구목록 입수가 첫 보도의 기록으로 볼 수 있다. 이로 인해 한국 내에 한동안 격렬한 반중(反中)감정을 불러일으켰다. 2004년에 중국 외교부는 홈페이지에 있는 한국 소개란에 고구려 역사를 삭제한 것과 관련한 한국 외교부의 대응⁷⁾ 사실이 보도된 후의 한중(韓中)간의 외교 논란을 야기했다.

중국 조선족(朝鮮族; 차오셴주 (chaoxianzu))을 정신적으로 완전하게 편입시키기 위한 크게 보면 동북공정과 궤를 같이한다. 중국 공산당은 한족 외 55개 소수민족과 그 소수민족이 가진 특유의 역사마저도 중화사상이라는 개념에 동화시켜 사회주의 체제를 공고히 하려는 목적을 갖고 있다. 특히 위구르나 티베트의 분리·이탈 시도가 빈번해지면서 더 적극적이고 강력하게 중화사상 및 중국은 하나뿐인 조국이라는 사상을 교육·주입하고 있다. 그리고 중국의 55개 소수민족 중에서 특별

6) 김정탁, 2015, 「중국 내 반한감정 흐름에 관한 연구」, 성균관대학교 대학원, 25쪽

7) 한국 외교부 홈페이지, 2014, "중국의 고구려사 왜곡 문제", 8월 5일, URL: https://www.mofa.go.kr/www/brd/m_4080/view.do?seq=289957&srchFr=&srchTo=&srchWord=&srchTp=&multi_itm_seq=0&itm_seq_1=0&itm_seq_2=0&company_cd=&company_nm=

히 조선족은 자유민주주의의 시장 체제를 가진 경제 대국 겸 모국(한국)이 유독 가까운 곳에 존재하고 있다. 이런 상황에서 만약 “조선족의 문화는 중국식과는 구별되는 한국식 문화다.”라고 인정할 경우 이들이 정신적으로 체제에서 이탈하게 될 가능성이 있으므로 “조선족의 문화는 중국 소수민족의 문화 중에서 하나이니까, 곧 중국 문화이다.”라고 규정짓는 것이다. 이로써 ‘한국 문화는 중국 문화의 일부’라는 논리가 완성된다.

일대일로 등 중화 패권주의의 문화적 측면 강화 시도라고 할 수 있다. 중국은 세계 제2의 경제 대국으로 성장하며 막대한 영향력을 가지고 있으며, 남중국해를 통해 동남아시아로, 중동을 통해 유럽에 진출하려는 외교적 목표를 가지고 있다. 심지어 저 멀리 아프리카에도 발을 걸치면서 현지인의 원성을 들어가면서까지 투자를 많이 감행하고 있다. (중국-아프리카 관계) 이러한 중국의 패권주의에 위협을 느끼고 견제하려는 서방 국가(미국, 일본, 인도, 호주 등)들은 자유롭고 열린 인도-태평양 구상으로 대응하여 양측간의 갈등이 깊어지고 있으며, 한편으로는 중국의 영향력이 강한 지역(한국, 북한, 이탈리아, 동유럽 일대)에서도 중국인 관광객이나 부동산 투기 문제, 불법 체류자 문제, 코로나바이러스감염증-19 사태 등으로 반중감정이 확산하고 있다. 이에 중국은 물리적 영향력(경성 권력)뿐만이 아니라 문화적 영향력 부문에서 영향력을 강화하여 반중감정을 불식시키려는 정책이 진행 중이며, 이를 위해 폭넓은 문화 사업을 펼치고 있다.

2020년 이래로 중국이 중국 내 소셜 미디어 플랫폼이 아닌, 글로벌 소셜 미디어 플랫폼, 유튜브·트위터 등에 김치·한복·판소리·아리랑·삼계탕 등이 중국문화라고 주장하는 콘텐츠들을 업로드하기 시작했다. 또한 한 채널의 TV 퀴즈쇼에서는 ‘태극기도 중국인이 만들었다’라고 왜곡하고 하고 있으며 무술 관련 경연 프로그램에서는 ‘중국무술이 한국 태권도의 근본이며 그 뿌리는 중국무술이다’라고 언급을 하기도 했다. 공산당 기관지인 ‘인민일보’의 자매지인 환구시보는 2020년 11월 김치는 중국에서 유래한 것이라고 주장했다. 이후 유엔 주재 중국대사도 중국 정부의 계정으로 된 트위터에 김치를 중국 것이라고 했다. 환구시보나 인민일보 영문 자매지인 글로벌타임스에서 왜곡 보도하고 (카톡 같은) 위챗이나 (단문을 전달하는 SNS인) 웨이보를 통해 전파되고 있다.

III. 연구문제와 방법

3.1 연구 문제

연구 문제 1. 최초 중국 정부의 동북공정(東北工程) 시작이 발표된 2002년 이후, 국내 언론에 보도는 어떤 키워드들이 자주 나타나고 있는가?

연구 문제 2. 2021년 3월 SBS 드라마 ‘조선 구마사’ 폐지를 전후로 어떤 문제들이 이슈화되어 보도되었는가?

3.2 연구 방법

본 연구에서 자료 수집은 신문 기사 데이터베이스 포털인 ‘BIG KINDS’⁸⁾에서 2003년부터

8) <http://www.bigkinds.or.kr>

2021년 4월까지 보도된 중앙지, 경제지, 지역종합지, 방송사 기사를 대상으로 ‘중국’과 ‘동북공정’이라는 키워드로 검색된 뉴스 제목의 데이터를 수집하였다. 이를 통해 총 7,546건의 기사를 대상으로 분석하였으며, 이중 중복되는 기사는 제거했다. 중국 정부의 동북공정(東北工程) 시작이 발표된 2002년부터 2021년까지 18년간 보도된 기사들을 포함했어야 했으나, BIG KINDS 뉴스 검색 데이터가 2003년 이후부터의 데이터가 검색되어 2003년부터 2021년 4월까지를 검색 기간으로 설정했다.

또한, 2021년 3월 SBS 드라마 ‘조선 구마사’ 폐지를 전후로 어떤 이슈들이 보도되고 있는가를 분석했다. ‘조선 구마사’라는 키워드로 검색된 뉴스 제목의 데이터를 통해 용어의 빈도수와 양을 측정하여 유의미한 결과를 도출하는 것이라 할 수 있다.

전처리 과정은 신문 기사 데이터베이스 포털 ‘BIG KINDS’에서 수집된 자료를 R-STUDIO를 통해 뉴스 기사 제목(HEADLINE)의 빈도 분석을 실시했다. 신문 기사 텍스트의 빈도를 분석하고 그 결과 값을 시각화하기 위해서는 네 단계의 분석과정을 거쳐야 한다.

1. 텍스트 전처리 과정
2. 토큰화 과정
3. 빈도수 분석과정
4. 시각화 과정

첫 번째, readxl 패키지를 통해 ‘빅 카인즈’에서 수집된 엑셀 파일을 로드한 후 함수 stringr 패키지를 통해 공백, 구두점, 데이터의 노이즈를 제거하는 텍스트 전처리 과정을 실행한다. 두 번째, 한국어 형태소 분석을 위해 multilinguer 패키지를 통해 한국어 텍스트 전처리를 위한 형태소 분석기 RcppMeCab를 설치 후 텍스트를 단락, 문장, 단어, 형태소 단위로 나누는 토큰화 과정을 실행한다. 세 번째, 신문 기사의 제목인 텍스트의 빈도수를 분석하기 위해 dplyr 패키지를 통해 단어 빈도 분석 과정을 실행한다. 네 번째, ggplot2 패키지를 통해 그래프 및 워드 클라우드 형태의 데이터 시각화 과정을 실행한다.

IV. 연구 결과

언론 기사를 통해 대중적인 관점에서 이슈에 대한 대중의 전반적인 인식과 주요 사회 의제를 파악할 수 있다. 앞서 제시한 두 가지 연구문제 ‘중국 정부의 동북공정(東北工程) 시작이 발표된 2002년 이후, 국내 언론에 보도는 어떤 키워드들이 자주 나타나고 있는가?’와 ‘2021년 3월 SBS 드라마 〈조선 구마사〉 폐지를 전후로 어떤 문제들이 이슈화되어 보도되었는가?’를 분석하기 위해 신문 기사 데이터베이스 포털 ‘BIG KINDS’에서 수집된 자료에 대한 키워드 빈도분석 및 워드 클라우드(Wordcloud)를 도출했다.

2002년 중국 측에서 동북공정을 발표한 시점으로 첫 번째, ‘역사 왜곡’ 규탄과 대책 마련을 촉구하는 기사들이 주류를 이루었고, 두 번째, 그에 따른 대책으로 ‘한국의 고대사 바로 알기’, ‘고구려사 바로 세우기’ 등 중국의 동북공정에 대응하기 위한 구체적인 방안과 해결책이 제시되었고, 세 번째, 중국의 동북공정 전략이 영토의 역사 왜곡에서 문화와 문화콘텐츠의 영역으로 확장되는 문화공정의 변화 양상으로 파악할 수 있다.

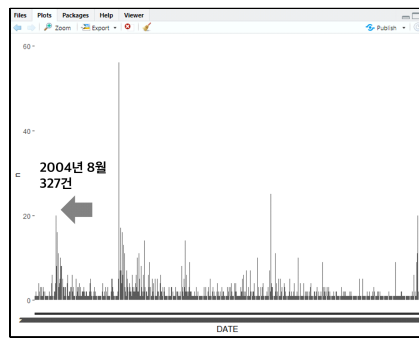


그림2. ‘중국, 동북공정’
키워드 빈도분석 결과 - 2004년 8월

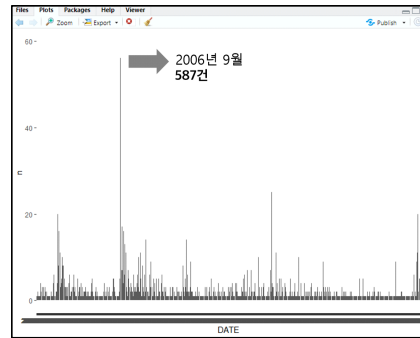


그림3. ‘중국, 동북공정’
키워드 빈도분석 결과 - 2006년 9월

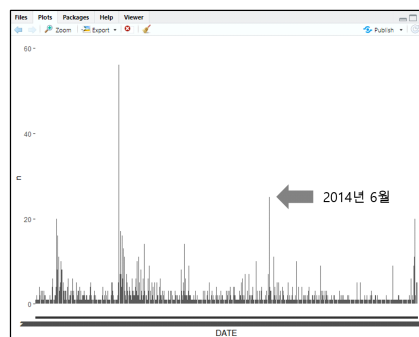


그림4. ‘중국, 동북공정’
키워드 빈도분석 결과 - 2014년 6월

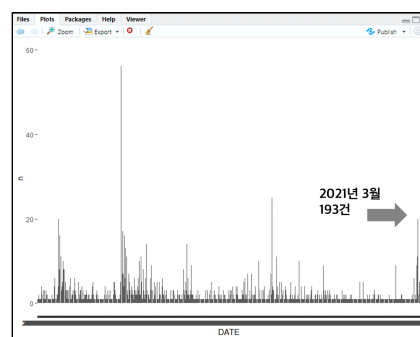


그림5. ‘중국, 동북공정’
키워드 빈도분석 결과 - 2021년 3월

2002년 중국 동북공정 발표 이후 2021년까지 ‘중국’, ‘동북공정’ 관련 뉴스 보도량이 크게 증가한 시점은 네 부분으로 정리된다. 첫 번째(그림2), 2004년 7월 북한과 중국이 고구려 유적을 ‘세계문화유산’으로 각각 등재 결정이 이루어지면서 8월부터 국내 언론에서 집중 보도되기 시작했다. 이 시기는 중국 동북공정의 역사 왜곡 현상이 고구려 유적 ‘세계문화유산’ 등재라는 사안으로 구체화되어가고 있으며, 중국 동북공정이 중국의 국가 주도로 치밀하게 계획되어 진행되고 있다는 것을 파악하는 시점이라고 할 수 있다.

두 번째(그림 3)는 중국의 2006년 9월을 전후로 뉴스가 보도량이 증가하였다. 이 시기는 2006년 9월 10일 아시아유럽정상회의(ASEM)에 참석한 노무현 대통령이 중국의 고대사 왜곡 움직임인 동북공정과 관련하여 원자바오 중국 총리에게 직접 유감을 표명했던 시기이다. 원 총리는 “정부 차원에서 필요한 조치를 취하겠다”라는 의사를 전달하고 향후 동북아 고대사와 관련된 논의를 이어가자는 원론적인 답변이 있었다. 2002년 2월 시작된 중국의 동북공정은 5년 기한으로 시기적으로 이미 종료 시점이 이르고 있었으며, 중국 동북공정에 대처하지 못한 정부의 무능을 비판하는 기사

와 학술 차원의 연구 및 실제 파악을 촉구하는 기사들이 주류를 이루고 있다.

세 번째<그림 4>은 2014년 6월 중국 자본의 힘으로 한류 콘텐츠와 한류 스타를 역이용한 첫 사례라고 할 수 있다. 한류 스타인 김수현과 전지현이 중국의 장백산 생수 광고 모델 발탁 여부로 다량의 기사가 작성되었으며, 이 시기는 중국의 동북공정 전략이 영토의 역사 왜곡에서 문화와 문화 콘텐츠의 영역으로 확장되는 문화공정이 시발점이 되는 시기였음을 알 수 있다.

네 번째<그림 5>는 2021년 3월을 전후로 증가한 뉴스 보도량 증가 그래프이다. 2021년 3월 26일 SBS 드라마 '조선 구마사'의 한 장면에서 노골적으로 중국풍 소품과 의상을 사용하고, 역사적 인물을 왜곡 표현하면서 논란을 일으켜 방영 2회 만에 폐지되었다. 드라마를 통해 "이미 중국 네티즌들은 웨이보(중국 눈)를 통해 '당시 한국의 전형적인 모습'이라며 드라마 장면을 옹호하기 시작했다"라는 비판적 관점의 기사와 이처럼 "드라마에 과잉 반응하는 그것이야말로 동북공정의 성공을 의미하고 있다"라는 옹호적 기사로 양분되어 나타나고 있다. 결국 '조선 구마사'의 사례처럼 중국 자본의 문화침투라는 새로운 형태의 동북공정 의도에 맞서 우리의 역사는 우리 스스로 지켜나가야 한다는 자성론적 목소리도 대두되고 있다.

결론적으로 2002년 중국 동북공정 발표 이후 2021년까지 '중국', '동북공정' 관련 언론 뉴스를 키워드로 분석해 본 결과, 첫째, 역사적 기록을 의도적으로 왜곡 함으로써 한국의 고대사를 중국 자국의 역사로 편입시키려는 의도를 분석, 이에 대한 대응 방안을 촉구하는 기사의 흐름이 있었다. 둘째, 역사 왜곡 패러다임을 지나 문화콘텐츠 부분에 침투된 중국 자본에 대한 위험성을 보도하는 흐름으로 정리할 수 있다.



그림 6. 2021년 3월, SBS 드라마 '조선 구마사' 보도 관련 키워드 빈도분석 워드 클라우드 시각화

V. 결론 및 제언

본 연구는 2021년 3월 26일 SBS 드라마 '조선 구마사'의 내용 중 조선의 역사적 인물을 왜곡 표현으로 인한 역사 왜곡 논란을 통해 '중국', '동북공정(東北工程)'을 키워드로 하여 2003년부터 2021년까지 18년간 언론 기사에 보도된 내용은 무엇인지, 그리고 우리가 미처 인식하지 못하는 이슈의 흐름을 살펴보고자 했다.

SBS 드라마 ‘조선 구마사’ 방영 전후, 신문 기사의 주요 키워드 추출을 통해 뉴스 기사의 주요 이슈를 분석했다. 본 연구는 ‘중국’, ‘동북공정(東北工程)’ 관련 기사에 대해 텍스트마이닝 기법을 활용하여 빅데이터를 분석하는 것에 의미를 두었다. 분석 결과 국내 언론에 가장 많이 언급된 키워드로는 역사 왜곡(779건), ‘고구려, 고구려사(616건)’, ‘한국사, 한국(292건)’, 논란(177건), 김수현(156건)으로 순으로 나타났다. 이를 통해 각 키워드의 세부적인 기사 흐름을 분석해 본 결과 ‘동북공정’은 중국 정부에서 2002년 ‘동북공정’ 계획을 공식적으로 발표한 이후 2004년 유네스코 세계문화유산 등록을 통해 고구려 역사를 자국 역사의 일부로 편입시키는 ‘역사 왜곡’의 1기와 2014년 이후부터는 한류 콘텐츠의 배우들을 자국의 콘텐츠에 출연시켜 한국 배우의 이미지를 차용하여 한국 문화를 중국 문화의 일부로 흡수하려는 ‘문화 왜곡’의 2기로 정의할 수 있으며, 최근 2021년 3월 SBS ‘조선 구마사’ 논란 전후로 문화콘텐츠 제작의 자본 참여를 통한 ‘역사문화공정’의 3기로 분류할 수 있다.

본 연구는 텍스트마이닝 기법을 활용하여 지난 18년간의 기사 제목 빈도수 분석을 통해 ‘중국’, ‘동북공정’을 보도 내용의 흐름과 이를 통해 도출된 특정 키워드들을 통해 ‘중국 동북공정’의 과정을 파악할 수 있었다. 하지만 키워드 빈도수 분석만으로 내용분석의 흐름을 파악하는 데 한계가 있었으며, 객관적으로 유의미한 통계 결과를 도출하지 못하였다. 향후 보도된 기사의 제목 빈도수 분석 외 기사 본문의 감정분석, 의미연결망 등의 빅데이터 분석을 통해 보다 유의미한 언론 기사 분석이 필요할 것이다.

【참고문헌】

- Chakraborty, G., M. Pagolu and S. Garla, Text Mining and Analysis: Practical Methods, Examples, and Case Studies Using SAS, SAS Institute, 2013.
- 최광식(2004), 「東北工程의 배경과 내용 및 대응방안 -고구려사 연구동향과 문제점을 중심으로-」 - 한국고대사 연구 33, 5-21쪽.
- 이남주(2004), 「중국 동북공정의 논리와 대응 방향」 황해문화 45, 219~230쪽.
- 김은국(2007), 「중국의 동북공정 상에 보이는 발해사 서술」 아시아문화연구 12, 103-132쪽.
- 이장원(2007), 「중국의 문화적 팽창주의 동북아 질서에 대한 중국의 의도 분석」, 한국학술진흥재단, 국제정치논총, 48(2), 33~52쪽.
- 강윤정, 윤재열, 임지연, 김웅모(2011), 「뉴스 기사의 연관 단어 텍스트 마이닝을 이용한 스타의 분야별 기여도 순위 비교기법」, 18(02), 1,191-1,194쪽.
- 김미아,송민 (2012) 「텍스트 마이닝을 활용한 신문사에 따른 내용 및 논조 차이점을 분석」, 지능정보연구, 18(03), 53-77쪽.
- 이천석(2012), 「중국 동북 강변 강정 책과 문화공정, 국제정치연구」 15(2) : 47~69쪽.
- 정호섭(2013), 「중국의 POST 東北工程과 고구려사 관련 동향 분석」 韓國史 學報 51, 37-70쪽.
- 노형임(2014), 「빅데이터 텍스트 마이닝-정치연설을 중심으로」, 화법 연구, 0(26), 289-325쪽.
- 박철수(2019), 「텍스트마이닝을 활용한 북한 관련 뉴스의 기간별 변화과정 고찰」.
- 김성수(2020), 「네이밍구자치구(內蒙古 自治區) 민족문화공정, 내력과 전망」, 동북아 역사 논증, 0(68), 85-150쪽.
- 윤호영(2020), 「남북 관계 뉴스 제목 텍스트 마이닝 분석: 227개월 44만 건 보도 관심주제」, Journal

- of the Korean Data Analysis Society(December 2020), 22(6), 2,517-2,534쪽.
- 이승호(2020), 「동북공정 이후 중국 대중미디어의 고구려 서사 방식: TV 역사드라마를 중심으로」, 동서 비교문학 저널, 0(53), 281-304쪽.
- 황서이, 박정배(2020), 「한류' 경향에 관한 국내 언론 기사 빅데이터 분석연구」, 한국 엔터테인먼트 산업 학회논문지 14(5), 1-14쪽.
- 조민정, 이신행(2021), 「코로나19 관련 언론 보도 프레임 분석: 자료기반 자동화 프레임 추출 방법을 중심으로」, 한국 소통학보, 20:1, 65-107쪽.