

토픽모델링과 키워드 네트워크 분석을 활용한 ‘문화콘텐츠’ 연구 경향 분석*

민요한, 김지영, 박옥남†

상명대학교 문헌정보학과

한국 원천콘텐츠에 관한 관심이 확대되고, 디지털 기술의 발달로 인해 문화콘텐츠의 제작·유통·활용 방식이 다양해짐에 따라, 문화콘텐츠는 우리가 살아가는 방식을 넘어 산업으로 자리 잡고 있으며, 국가의 부가가치 창출의 중요한 원천으로 강조되고 있다. 본 연구는 다양한 학문 분야에서 이루어지고 있는 문화콘텐츠에 대한 흐름을 살펴보는 것을 목적으로 한다. 문화콘텐츠의 연구 경향을 분석하기 위하여 2002년부터 2019년까지 KCI에 게재된 논문을 바탕으로 토픽모델링과 키워드 네트워크 분석을 수행하였다. 연구 결과, 한국, 개발, 스토리, 문학 등이 가장 중요한 키워드로 나타났다. 제1구간으로 구분된 2002년부터 2009년은 문화콘텐츠 연구의 도약기로, 문화콘텐츠와 정보기술의 결합으로 인해 문화콘텐츠가 국내에서 중요한 산업으로 진입하게 된 시기라고 볼 수 있다. 이를 통해 다양한 원천콘텐츠의 디지털화, 디지털 기술을 이용한 원천콘텐츠의 이용 가능성에 관한 연구가 활발하게 이루어진 시기라고 할 수 있다. 제2구간인 2010년부터 2014년은 성장기로 1구간의 디지털화 기술을 통해 입증된 문화콘텐츠의 활용 가능성에 대한 인정을 바탕으로 한류, 서사, 문학, 지역 등 다양한 영역으로 문화콘텐츠 연구가 본격적으로 성장한 시기로 해석된다. 2015년부터 2019년 사이의 3구간은 성숙기 단계로, 이 단계는 1~2구간을 통해 인정받는 문화콘텐츠 산업의 확산을 관광, 축제, 역사, 종교, 지역 및 공간, 교육 등 다양한 산업 분야로 성숙시키기 위한 다양한 연구가 수행된 시기로 해석된다. 본 연구는 약 20년간 수행되어 온 문화콘텐츠의 연구 동향을 토픽모델링과 네트워크 분석을 통해 다각적으로 조사하였다는 점에서 의의가 있다.

주요어: 문화콘텐츠, 연구동향, 네트워크 분석, 토픽모델링

* 본 연구는 2019년도 상명대학교 교내연구비를 지원받아 수행되었음.

† 교신저자(Corresponding Author) : 박옥남, 상명대학교 문헌정보학과 부교수, 서울시 종로구 홍지문길 20, E-mail : ponda@smu.ac.kr

민요한, 상명대학교 일반대학원 문헌정보학과 박사과정 수료(제1저자)

김지영, 상명대학교 일반대학원 문헌정보학과 박사과정 수료(공동저자)

■ 최초투고일 : 2021년 1월 29일 ■ 심사마감일 : 2021년 3월 29일 ■ 게재확정일 : 2021년 4월 4일

1. 서론

문화콘텐츠는 “문화산업이나 문화 주체의 문화 활동을 통해 생겨난 문화적 요소가 체화된 부호·문자·도형·색채·음성·음향·이미지 및 영상 등의 자료”를 지칭하는 용어(문화산업진흥기본법)로 창작물뿐만 아니라 창작물을 통해 파생된 모든 결과물을 포괄하는 개념이다. 문화콘텐츠 분야는 2001년 이후 괄목할 만한 성장을 거두고 있다. 한국의 영화, 음악, 드라마, 게임, 캐릭터, 출판 등 다양한 대중문화 콘텐츠의 해외 진출이 활성화됨에 따라, 한국 원천콘텐츠에 관한 관심이 확대되고, 디지털 기술의 발달로 인해 문화콘텐츠의 제작·유통·활용 방식이 다양해졌기 때문이다. 문화콘텐츠산업 통계조사에 따르면 2018년 콘텐츠산업 사업체는 10만 5,310개이며, 매출액은 119조에 이르고, 부가가치액은 약 47조로 약 39.7%의 부가가치율을 보이는 것으로 나타났다(한국콘텐츠진흥원, 2019). 이제 문화콘텐츠는 우리가 살아가는 방식을 넘어 산업으로 자리 잡고 있으며, 국가의 부가가치 창출의 중요한 원천으로 강조되고 있다. 문화콘텐츠라는 새로운 용어의 등장 이후 문화산업 범위가 확대됨에 따라 문화콘텐츠에 관한 관심이 커지고 용어의 정의 역시 확대되고 변형되고 있으며, 문화콘텐츠 분야는 특정 분야에서 연구가 한정적으로 이루어지는 것이 아닌 다양한 분야에서 학제적으로 연구되고 있는 것이 특징이다.

중앙정부, 지방정부 및 공공기관 등 다양한 정부 기관의 문화콘텐츠 산업에 대한 정책적 지원이 증가하고, 기업 등 산업계에서도 문화콘텐츠에 대한 참여가 점차 증가하고 있다. 대학 및 연구기관에서도 점차 문화콘텐츠와 관련된 학과 및 전공이 신설되고, 관련 연구가 본격적으로 시작되고 있으며, 특히 IT산업의 발전에 따

른 콘텐츠의 중요성 증가와 한류의 세계적 확산으로 인하여 문화콘텐츠의 중요성은 점차 커지고 있다.

이러한 문화콘텐츠의 중요성에도 불구하고, 문화콘텐츠에 관한 연구를 살펴보면 다양한 개념과 학제적 정의가 존재한다. 이는 문화콘텐츠라는 학문 분야가 어느 특정한 학문 분야에 속한다기 보다는 학제적인 성격으로 다양한 분야에서 연구가 진행되고 있기 때문이다. 이 때문에 다양한 분야에서 진행되는 문화콘텐츠 관련 연구는 통합된 개념으로서의 ‘문화콘텐츠’의 범주가 부족하고 각 학문 분야의 연구자가 자신들의 관점에서 문화콘텐츠를 정의하고 논의를 전개하는 것으로 살펴진다.

이러한 현실에도 불구하고 문화콘텐츠는 다양한 분야를 통합하는 개념으로써 위상을 가지고 있다. 다양한 학문 연구 분야에서 문화콘텐츠의 새로운 발전 방안을 모색하는 연구를 진행하면서 문화콘텐츠의 하위 분야가 확대되고 실질적인 위상이 상승하고 있으며, 이를 바탕으로 연구영역의 세분화, 확장 및 심화가 이루어지고 있다.

이에 본 연구에서는 다양한 학문 분야에서 이루어지고 있는 문화콘텐츠에 대한 흐름을 살펴보고자 한다. 문화콘텐츠 학문 분야에서 연구되고 있는 핵심 주제를 살펴보고, 주제간 연관 관계, 변화추이를 살펴봄으로써 학문의 발전 방향을 살펴볼 수 있다는 의의가 있다. 이를 위하여 토픽모델링 및 키워드 네트워크 분석을 통해 현재까지 진행된 문화콘텐츠의 연구 주제 및 핵심 키워드 등을 확인한다. 또한, 분석 결과를 바탕으로 문화콘텐츠의 과거와 현재를 살펴봄은 물론 추후 진행되어야 할 연구에 대해 살펴보고 이를 통해 문화콘텐츠의 학문적 발전을 촉진하고자 한다.

2. 이론적 배경

1) 토픽모델링 연구

토픽모델링이란 텍스트마이닝 기법의 하나로 비구조화된 텍스트 자료들의 문치로부터 토픽을 추출해주는 확률모델 알고리즘이다(Blei, Ng, & Jordan, 2003). 토픽모델링 기법은 방대한 텍스트 자료들로부터 맥락과 관련된 단서를 활용하여 토픽을 추출하고 각 토픽에 해당하는 문서를 파악하여 제공한다. 현재 토픽모델링은 디지털화된 문헌연구의 기법, 학술 동향을 분석하는 데에 많이 사용되고 있다.

나상태, 김자희, 정민호, 안주연(2016)은 토픽모델링을 이용한 시물레이션의 연구 동향 분석을 시행하였다. 이를 위하여 1992년부터 2015년까지 한국연구재단이 제공하는 KCI 데이터베이스에서 “시물레이션”이라는 한글 검색어로 검색하여 나온 문헌 11,895건의 제목, 초록 키워드, 게재연도를 크롤링하였다. 분석결과 시물레이션은 주로 공학 분야에서 사용되고 있으며, 사회과학 분야에서는 연구 결과가 빠르게 증가하고 있음을 파악하였다. 특히 공학 분야에서는 통신 및 전기분야에서 주로 사용되었고, 사회과학 분야에서는 시물레이션의 최신 활용 분야인 교육 및 오락, 교육, 국제교역 및 안전에 관한 연구에서 활발하게 사용되고 있음을 파악하였다.

박준형, 오효정(2017)은 토픽모델링 기법을 활용하여 국내 기록관리학의 연구동향을 분석하였다. 이를 위하여 1997년부터 2016년까지 국내 기록관리학 및 문헌정보학 관련 학술지 6종에서 발표된 기록관리학 관련 논문 1,027건을 각각 LDA 토픽모델링과 HDP 토픽모델링을 활용하여 분석하였다. 연구 결과 LDA 토픽모델링은 해당 분야를 대표하는 키워드의 빈도수를 바탕으로 대표되

는 토픽들의 키워드를 도출하는 데 효과적이었다. HDP 토픽모델링은 토픽별 특징을 파악할 수 있는 특수 키워드가 많이 도출되어 하위 토픽별 세부적인 핵심 키워드를 도출하는 데 효과적임을 확인하였다.

김창식, 최수정, 광기영(2017)은 토픽모델링과 시계열 회귀분석을 활용하여 정보시스템 분야의 연구동향을 분석하였다. 이를 위하여 2002년부터 2016년까지 Asia Pacific Journal of Information Systems, Information Systems Review, The Journal of Information System의 논문 1,245편의 초록을 선정하였다. 토픽모델링 분석 결과 “혁신역량”, “시스템구축”, “고객충성도” 등 주요 상위 20개의 토픽을 도출하였다. 시계열 회귀분석 결과 상승 추이를 나타내는 토픽으로 “고객충성도”, “소통혁신”, “정보보호” 등이 나타났고, 하락 추세를 나타내는 토픽으로는 “시스템구축”, “웹사이트” 등이 도출되었다.

2) 네트워크 분석 연구

네트워크 분석은 개인과 집단들 간의 관계를 노드와 링크로 모형화하여 그들의 구조나 확산과정을 계량적으로 분석하는 방법이다. 노드 간의 연관 관계에 초점을 맞춘 네트워크 분석은 다양한 분야에 걸쳐 응용되고 있다.

김의준, 권영현(2012)은 지역 문화산업을 대상으로 사회적 네트워크 분석을 시행하였다. 분석 자료는 2005년 우리나라 지역 산업연관표이며, 분석 대상 지역은 문화산업의 특성이 크게 나뉘는 수도권과 제주권을 대상으로 하였다. 분석 틀로는 단위 구조의 사회연결망 분석 방법을 활용하였다. 분석 결과 문화산업의 경제적 성장은 다른 연관 산업의 생산 활동을 유도할 수 있다고 볼 수 있다. 또한, 문화산업의 네트워크 밀도는 다른 산업에 비하여

높게 나타나 일자리 창출, 기업이나 산업간 협력 등에 유리한 특성이 있다고 볼 수 있음을 제시하였다. 광역권별로 살펴봤을 때 우리나라는 수도권 중심의 강한 문화산업 네트워크가 형성되어 있음을 확인할 수 있었고, 제주권은 지역 간 산업간 연계구조가 약하게 형성되어 있다고 볼 수 있음을 파악하였다.

이성신(2016)은 도서관마케팅 분야의 국내외 연구동향을 분석하기 위하여 저자 키워드를 활용하여 키워드 네트워크 분석을 시행하였다. 이를 위하여 국내 문헌정보학 분야 학회지 4종 및 국외 SCOPUS에 게재된 도서관마케팅 논문에 포함된 저자 키워드를 분석하였다. 연구 결과, 도서관마케팅 분야는 ‘공공도서관’ 키워드를 중심으로 연구가 수행되고 있음을 파악하였으며, ‘공공도서관’과 ‘대학도서관’ 키워드가 연결중심성과 매개중심성이 높게 나타났다.

박양우(2017)는 언어네트워크 분석을 통하여 국내 문화정책 연구 동향을 분석하였다. 이를 위하여 ‘문화정책논총’에 수록된 2008년 10월부터 2017년 1월까지 186편이 논문으로부터 도출된 키워드 832개를 대상으로 언어네트워크 분석을 시행하였다. 언어네트워크 분석은 키워드 빈도수, 밀도 분석과 중심성을 활용하여 분석하였으며, 이를 위하여 Netdraw 프로그램을 활용하여 시각화를 시도하였다. 분석 결과 가장 높은 빈도수를 보인 키워드는 ‘문화’였고, ‘문화정책/행정’, ‘문화산업/문화콘텐츠’, ‘정책’ 등이 가장 빈도수가 높게 나타난 키워드임을 확인하였다.

신유미, 박옥남(2019)은 장서개발관리 분야의 최근 연구 동향 분석을 위하여 2003년부터 2017년까지 15년간 문헌정보학 분야 4개 학회지에 게재된 논문 중 장서개발관리 분야 키워드를 포함한 논문을 검색하고, 저자 키워드를 분석하였다. 추출된 저자 키워드를 바탕으로 NetMiner 4 프로

그램을 이용하여 빈도분석, 연결중심성 분석, 매개중심성 분석을 실시하였다. 연구결과, ‘오픈액세스’, ‘기관 레포지터리’, ‘학술지’ 등이 장서개발 분야의 주요 키워드임을 확인하였으며, ‘대학도서관’의 경우 향후 지속적으로 연구가 될 분야의 키워드임을 파악하였다. 이를 통해 국내 장서개발관리 분야의 연구 경향을 분석하고 향후 연구 방향과 과제를 탐색하였다.

이외에 연구 동향 분석을 위하여 토픽모델링과 네트워크 분석을 동시에 수행한 연구도 있다. 김유빈, 조남옥(2017)은 사회연결망 영역의 연구 동향을 분석하기 위하여 토픽모델링과 네트워크 분석을 시행하였다. 이를 위하여 SCOPUS 학술데이터베이스에 2000년부터 2015년 사이의 게재된 약 56,000건의 문서를 이용하였다. 토픽모델링 분석 결과 20개의 토픽을 추출하고 토픽의 연도별 점유율을 도출하였으며, 이와 함께 연구자 네트워크를 연구 분야 네트워크로 치환하여 연구자 개인의 학문 분야를 파악하고 분석하였다. 김형석, 박민식, 강필성(2016)의 경우, 딥러닝 연구 동향을 분석하기 위하여 토픽모델링과 사회연결망을 활용하였다. SCOPUS와 Arxiv 데이터베이스로부터 4,800개의 논문을 수집하였다. 이를 통해 30개의 토픽 및 토픽 간 네트워크를 분석하였으며, 국가 간 학술 협업 관계를 파악하였다.

3) ‘문화콘텐츠’ 선행연구

문화콘텐츠 관련 연구는 문화콘텐츠 분야에서 스토리텔링 및 문화원형에 관한 연구 등 다양한 방법을 통해 이루어지고 있다. 본 연구에서는 이러한 문화콘텐츠 관련 연구들 가운데 문화콘텐츠 및 관련 연구 동향을 파악한 선행연구들을 고찰해보았다.

류준호, 윤승금, 이영주(2010)는 문화콘텐츠 관련 연구에 대한 메타분석을 실시하였다. 이를 위해 2001년부터 2009년 11월까지 국내 KCI 등재지 및 등재 후보지에 게재된 문화콘텐츠에 대한 연구논문을 분석대상으로 삼아 메타분석을 활용하여 분석하였다. 문화콘텐츠 관련 연구의 학문적 배경, 연구목적, 연구방법 등의 범주에 초점을 두었다. 분석 결과 문화콘텐츠 연구는 다양한 학문 분야에서 참여하고 있으며 특히 인문학 분야의 참여가 활발한 것으로 분석됐다. 또한, 문화콘텐츠 연구는 산업적 목적으로 수행되는 경우는 물론, 학문적 목적을 가진 연구도 점차 증가하고 있으므로 학문으로서의 입지를 다져가고 있는 것을 확인하였다. 마지막으로 문화콘텐츠 관련 연구에는 특정한 연구 방법론이 집중적으로 사용되며, 다양한 연구 방법론이 적용되지 못하는 것으로 나타났다. 또한, 기존 전통적 문화콘텐츠 연구분야와 신규 연구 분야 간의 관계 정립이 필요하고, 통합된 분야로서 문화콘텐츠학의 학문적 위치를 정립할 필요가 있음을 밝히고 있다.

윤홍근(2011)은 학위논문의 메타분석을 통한 문화콘텐츠 연구 경향을 파악하였다. 이를 위해 초창기 문화콘텐츠학과를 개설한 국내 5개 대학의 학위논문을 대상으로 메타분석을 실시하였다. 연구 문제로는 분야, 연구목적, 연구방법, 학문적 배경, 교과과정의 연관성을 설정하였다. 분석 결과 첫째, 문화콘텐츠 분야의 연구는 다양한 분야에서 진행되고 있는 것으로 나타났다. 두 번째, 문화콘텐츠의 연구목적은 개발 소스 제공이나 연구 결과의 활용에 초점을 맞춘 실용적 목적이 대부분이다. 연구 방법론에서는 질적, 양적 방법을 혼합한 방법을 채택하고 있는 것으로 나타난다. 셋째, 문화콘텐츠 연구는 인문학을 중심으로 인접 학문 간 학제 간 융합이 일어나고 있음을 확

인할 수 있었다. 마지막으로 교과과정은 인문학 기반의 커리큘럼을 반영하고 있음을 확인하였다. 이 연구의 시사점은 문화콘텐츠 연구가 학제적 연구로 변화하고 있다는 것과 문화콘텐츠가 학문으로 자리 잡기 위해서는 다양한 학문의 연구 방법론과의 접목, 문화콘텐츠 학문만의 고유한 연구 방법론이 개발되어야 하며, 실무의 필요성을 반영한 교육과정 개편이 필요하다는 점이다. 마지막으로 연구자는 문화콘텐츠학의 역사가 길지 않으므로 연구 동향이나 지적구조를 전체적으로 분석하기에는 축적된 연구가 미비함을 이야기하며 연구 성과를 진단하는 후속 분석이 필요함을 지적하였다.

태지호(2016)는 문화콘텐츠는 학제연구의 특성이 있으며, 연구 분과 역시 매우 다양하다는 특징을 가지고 있다. 문화콘텐츠를 문화적 소재를 다양하게 가공하여 매체로 표현한 무형의 결과물로 정의하였다. 또한 '문화'와 '콘텐츠'를 다루는 방법론으로 활용 가능한 인접 학문의 연구 방법론을 제시하였다. '문화'와 '콘텐츠'에 대한 연구는 각각 '이해'를 위한 해석학적이고 비판적인 논의이며, '분석'을 위한 가치와 효과 중심의 논의라고 하였다. 또한, 양 단의 연구 방법론은 명확하게 구분되는 것이 아니므로 문화콘텐츠학의 독립적 정체성을 지닌 학문 영역으로 자리매김하기 위해 계속되는 조정과 연구가 필요함을 언급하였다.

오정심(2020)이 '문화콘텐츠'의 최근 20년간 수행된 관련 논문의 토픽모델링과 네트워크 분석을 수행하였다. 토픽모델링을 전 구간에 걸쳐 분석하였으며, 토픽모델링을 활용, 산업, 사회, 장르, 기술, 이론 및 체계 등으로 연구 분야로 구분하여 분류하였다. 또한, 문화콘텐츠 관련 학술논문으로는 인문학, 예술체육학, 복합학이 가장 높게 나타나는 것을 파악하였으며, 연구자 컴포넌트 분석

결과, 학제 간 연구가 다수를 차지하는 것을 파악하였다.

이를 종합할 때 토픽모델링과 네트워크 분석은 여러 분야의 연구 경향을 분석하기 위하여 활용됐음을 알 수 있다. 네트워크 분석은 키워드 간의 관계 및 주요 키워드의 중심성을 살펴봄으로써 연구 분야의 주요 키워드 및 관계 형성을 살펴볼 수 있는 이점이 있다. 또한, 토픽모델링은 텍스트에서 단어들을 토픽으로 추출하고, 토픽에서 추출되는 대표 논문을 살펴볼 수 있는 장점이 있다. 문화콘텐츠의 약 20년간의 연구 수행 기간을 고려할 때, 문화콘텐츠의 토픽모델링 및 네트워크 분석을 통해 그간의 메타분석 및 질적 분석 외의 계량적 분석을 통한 연구 동향을 살펴볼 필요가 있다. 오정심(2020)이 ‘문화콘텐츠’ 분야의 토픽모델링과 네트워크 분석을 수행한 바 있다. 그러나 오정심(2020)의 연구는 토픽모델링을 전 구간에 걸쳐 분석하였으며, 키워드 추출어에 기본 단어인 ‘문화콘텐츠’를 포함했으므로, 빈도수, 연결중심성, 위세중심성, 토픽모델링에서 문화, 문화콘텐츠, 콘텐츠가 반복해서 가장 높은 순위로 나타나고 있으며, 토픽모델링을 활용, 산업, 사회, 장르, 기술, 이론 및 체계 등으로 연구 분야로 구분하여 분류하였다. 이에 토픽모델링을 구간별로 구분하고, ‘문화콘텐츠’를 제외하여 관련 주제가 명확히 드러나게 하는 것이 필요하다.

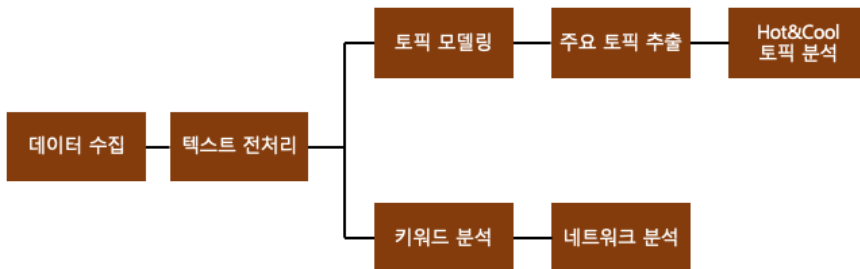
3. 연구 설계

1) 분석모델

본 연구에서 제안하는 문화콘텐츠의 연구 경향 분석을 위한 토픽모델링과 키워드 네트워크 분석은 다음과 같이 진행되었다(<그림 1> 참조). 첫째, 한국연구재단 KCI 데이터가 구축되기 시작한 2002년부터 2019년 11월까지 KCI에 등재된 저널의 논문을 대상으로 키워드 검색을 수행하여 논문을 수집하였다. 둘째, 수집된 논문을 대상으로 토픽모델링과 키워드 네트워크 분석을 수행하기에 적합한 포맷으로 변경하기 위하여 전처리 과정을 거쳤다. 마지막으로 전처리 과정을 거친 논문을 대상으로 각각 토픽모델링과 키워드 네트워크 분석을 시행하였다. 분석 도구는 분석 및 시각화를 동시에 시행할 수 있고 한글 데이터의 처리가 용이한 NetMiner를 분석 도구로 선정하였다. 분석기법은 빈도분석과 토픽모델링, 키워드 네트워크를 사용하였다.

2) 분석대상

본 연구는 문화콘텐츠의 연구 경향을 분석하기 위하여 토픽모델링과 키워드 네트워크 분석을 수행하였다. 문화콘텐츠라는 연구 분야가 특정 학문



<그림 1> 연구방법 및 과정

내에서 진행되는 연구가 아닌 학문의 구분 없이 진행되는 학제적인 연구 분야이기 때문에 분석 범위는 특정한 학문이 아닌 국내 학술지에 수록된 논문들을 대상으로 하였다. 이를 위하여 국내 학술지 및 게재논문에 대한 정보를 제공하는 KCI (Korea Citation Index)에서 키워드 검색을 수행하였다. 수집 기간은 문화콘텐츠 연구가 본격적으로 시작된 2002년부터 2019년 11월로 하였다. 검색결과는 총 1,693건으로 나타났으며, 기간별 연구 동향을 살펴보고자 3개 구간으로 나누어 분석하였다. 초창기에는 문화콘텐츠 연구가 시작하기 시작한 시기이므로 논문 수가 많지 않음을 고려하여, 2구간 및 3구간은 각 5년으로 구분하고, 나머지 1구간은 2002년부터 2009년까지 8년으로 선정하였다. 구간별 논문발행 건수를 살펴보면 2002년에서 2009년까지는 410건, 2010년에서 2014년까지는 588건, 2015년에서 2019년 11월까지의 695건으로 구간마다 계속 증가하는 것을 알 수 있다. KCI에서 ‘문화콘텐츠’를 키워드로 한 논문들을 검색한 결과는 <표 1>과 같다.

3) 전처리과정

국내외 문화콘텐츠 관련 학술논문의 키워드 분석 방법은 다음과 같다.

첫째, NetMiner를 통해 키워드 검색을 통해 자료를 수집한다.

둘째, 수집된 키워드는 다음과 같은 절차를 거쳐 표준화한다.

- (1) 검색어로 사용한 ‘문화콘텐츠’, ‘문화’, ‘콘텐츠’는 분석대상에서 제외한다.
- (2) 한 글자 단어는 제외한다. 예를 들면, 은, 는, 이, 가 등의 조사 및 응, 원, 해, 담 등의 의미를 유추하기 어려운 단어는 제외한다.
- (3) 같은 의미의 단어가 띄어쓰기 된 단어는 하나의 단어로 통일한다. 예를 들면, 스토리텔링, 스토리텔링은 스토리텔링으로 통일한다.
- (4) 동일한 의미로 쓰인 번역어는 하나의 단어로 통일한다. 예를 들면, 스토리와 이야기는 이야기로 통일한다.
- (5) 동일한 의미로 쓰인 유사어는 하나의 단어로 통일한다. 예를 들면, 전통, 전통문화는 전통으로 정보화, 정보기술은 정보 등으로 통일한다(<표 2> 참조).

<표 1> 구간별 논문수

구분	1구간(2002~2009)	2구간(2010~2014)	3구간(2015~2019)	총 건수
논문수	410	588	695	1,693

<표 2> 키워드 표준화 사례

구분	예시
검색어 제외	문화콘텐츠, 문화, 콘텐츠, 연구
한 글자 단어 제거	응, 원, 해, 담 등
띄어쓰기	스토리텔링 → 스토리텔링
번역어 통일	스토리 → 이야기
유사어 통일	전통, 전통문화 → 전통, 정보, 정보화, 정보기술 → 정보 등

4. 분석결과

1) 빈도분석

논문의 표제, 저자 키워드, 초록에서 나타난 키워드 출현 빈도수를 기반으로 키워드가 많이 연구된 분야를 조사하였다. KCI에서 ‘문화콘텐츠’를 키워드로 한 논문들을 검색한 결과 2002년부터 2009년 11월까지 발간된 논문은 총 1,693건이었으며, 그중 1개의 논문은 모든 키워드가 다른 논문에는 전혀 출현하지 않은 예외적인 경우였다. 이를 제외한 총 1,692건을 분석대상으로 하였다. 키워드 빈도분석은 2002년에서 2019년까지 전 구간과 연구논문의 건수 분포에 따라 3개의 구간으로 구분하여 2010년 이전을 1구간(2002~2009), 2000년 이후 5년 단위로 나누어 2구간(2010~2014), 3구간(2015~2019)으로 구분하여 구간별 키워드 분포 및 변화를 살펴보았다(<표 3> 참조).

3개 구간에서 공통으로 나타난 키워드는 ‘스토리’, ‘개발’, ‘교육’, ‘한국’, ‘지역’, ‘전통’이다. 2개

구간에서 나타난 키워드로는 ‘디지털’, ‘한류’, ‘문학’이 있다. 문화콘텐츠 연구로는 지역적 특색을 반영한 전통 및 교육의 개발에 관한 연구가 있었음을 알 수 있다. 1구간 및 2구간에서 디지털에 관한 키워드가 나타났으나 2015년 이후 3구간에서는 ‘디지털’에 대한 키워드가 상위 10개 안에 포함되지 않았으며, ‘지역’에 대한 키워드가 1구간, 2구간, 3구간에서 지속해서 상위 순위에 등장하고 있다. 구간별로 고유하게 등장하는 키워드로는 1구간에서 ‘원형’, 2구간에서는 ‘관광’과 ‘가치’가 등장하였으며, 3구간에서는 ‘역사’ 키워드가 등장하였다. 이를 종합해보면, 2002년부터 2009년까지는 문화산업이나 문화원형을 이용한 스토리텔링 또는 디지털 스토리텔링에 초점을 맞추고 있음을 알 수 있다. 2010년부터 2014년은 지역, 한류, 전통 등 문화콘텐츠의 재료가 될 수 있는 하위주제에 관한 연구가 활발하게 이루어졌음을 알 수 있다. 2015년부터 2019년 구간은 산업으로서의 문화콘텐츠와 역사, 관광 등 역사 스토리텔링, 관광 스토리텔링 등 문화콘텐츠의 영역이 학문적으로 점차 확장되고 있음을 알 수 있다.

<표 3> 구간별 키워드 비교(상위 10)

구간 순위	1구간(2002~2009)	2구간(2010~2014)	3구간(2015~2019)
1	스토리	스토리	지역
2	디지털	한국	스토리
3	개발	지역	한국
4	원형	교육	교육
5	교육	개발	개발
6	한국	디지털	문학
7	문학	한류	역사
8	지역	관광	관광
9	전통	전통	전통
10	인문학	가치	한류
키워드 수	1,408개	1,849개	2,192개

2) 토픽모델링 분석

토픽모델링은 비구조화된 문서 집합에서 잠재된 토픽들을 추출해주는 확률적 모델 알고리즘(Blei, Ng, & Jordan, 2003)으로 하나의 문서가 여러 토픽에 동시에 대응될 수 있다는 점에서 일반적인 군집화 기법과 차별되며, 현실 세계의 모델링에 더욱 적합한 기법으로 평가받고 있다. 토픽모델링의 대표적인 방법론으로 단어들의 확률을 이용하여 문서 집합 내의 잠재된 토픽을 찾아내는 기법인 LDA(Latent Dirichlet Allocation)가 있다. 초기 잠재의미분석(LSA)에서 시작하여 2003년 Blei가 고안한 LDA는 문서, 단어 등 관찰된 변수를 통해 문맥, 문서의 구조 등 보이지 않는 변수를 추론하는 방법으로 전체 문서집합의 주제, 문서별 주제 비율, 각 단어가 각 주제에 포함될 확률 등을 파악할 수 있다(박자현, 송민, 2013).

본 논문에서는 토픽모델링을 실시하기 위하여 KCI 학술지 논문 처리가 비교적 쉽고, 한글 처리가 용이한 NetMiner를 사용하여 토픽모델링을 수행하였다. 토픽모델링을 위한 알파값은 0.01, 베타값은 0.1로 설정하였고, Iterations는 10,000번으로 설정하였다. 또한, 빈도분석과 마찬가지로 2002~2009년, 2010~2014년, 2015~2019년까지의 기간을 나누어 분석하였다. 구간별 토픽모델링 결과, 빈도가 10회 이상 나타난 토픽의 분석 결과는 다음과 같다.

(1) 1구간(2002~2009) 토픽모델링

다음 <표 4>는 2002년에서 2009년까지의 토픽모델링의 결과를 나타낸 것이다. 이 기간에 발행된 논문에서는 스토리텔링, 디지털화, 한류, 지역 문화, 교육 순으로 토픽이 나타났다. 디지털 미디어를 활용한 스토리텔링이 가장 높은 순위에 있는

것을 고려할 때, 문화콘텐츠의 디지털화와 연구가 주로 이루어진 것으로 볼 수 있으며, 이 외에도 문화원형의 디지털 콘텐츠화, 한류 콘텐츠 교류, 지역관광 콘텐츠 개발 순을 볼 때 디지털화, 문화원형, 한류, 지역 문화와 관련된 연구들이 많이 이루어진 것으로 볼 수 있다. 토픽 17에 나타난 대표적인 논문으로 온라인 게임의 트랜스 미디어 스토리텔링에 관한 연구(류철균, 서성은, 2008), 토픽 8의 문화원형 디지털 콘텐츠화 사업에 대한 연구(안지형, 김운미, 2008), 토픽 18의 한-중, 한-일 문화콘텐츠 교류의 현황과 전망(김세훈, 2009) 등이 있다.

(2) 2구간(2010~2014) 토픽모델링

다음 <표 5>는 2010년에서 2014년까지의 토픽모델링의 결과를 나타낸 것이다. 이 기간에 발행된 논문에서는 전통문화, 고전문화, 스토리텔링, 지역 및 도시, 한류의 순서로 토픽이 나타났다. 전통문화와 고전문화가 가장 높은 순위에 위치하여 전통과 고전과 관련된 연구가 많이 수행되었다고 볼 수 있고, 1구간과 마찬가지로 스토리텔링과 한류가 여전히 상위 순서에 위치하여 연구가 많이 수행되는 것을 볼 수 있었다. 특이사항으로는 1구간에서는 등장하지 않았던 문화유산, 융합, 문화 이해와 관련된 토픽들이 상위로 등장하여 새롭게 연구가 수행된 것을 확인할 수 있었고, 디지털화는 순위가 많이 내려가 1구간보다는 연구가 덜 진행되는 것으로 볼 수 있다. 문화유산과 관련된 예로 여수지역의 해양 문화유산을 활용한 문화콘텐츠 개발에 관한 연구(노시훈, 2010)가 있으며, 문화융합과 관련된 대표적인 예로 문화예술과 콘텐츠산업의 융합에 관한 연구(김세훈, 2013; 홍병선, 2014)가 수행된 바 있다.

<표 4> 1구간(2002~2009) 논문 토픽모델링 결과

No	토픽	핵심키워드	빈도
Topic-17	디지털 미디어 스토리텔링	스토리, 디지털, 서사, 미디어, 대중	89
Topic-8	문화원형 디지털 스토리	스토리, 디지털, 창작, 원형, 문학	45
Topic-18	한류 콘텐츠 교류	한국, 한류, 디지털, 향토, 교류	44
Topic-6	지역관광 콘텐츠 개발	지역, 축제, 관광, 개발, 장소	38
Topic-12	인문 콘텐츠 인력양성	교육, 인문학, 대학, 인력, 창의	32
Topic-19	민속문화 콘텐츠화	원형, 민요, 디지털, 불교, 진도	31
Topic-14	고전문학 콘텐츠화	문학, 공간, 교육, 고전, 지리	25
Topic-10	애니메이션 캐릭터 디자인	디자인, 캐릭터, 애니메이션, 제작, 공공	19
Topic-4	영웅서사시의 신화 스토리 개발	신화, 영웅, 스토리, 개발, 서사시	17
Topic-15	디지털 음악의 진흥	기술, 정책, 진흥, 디지털, 음악	14
Topic-3	정보서비스 모형	정보, 서비스, 모형, 영화, 시각	11

<표 5> 2구간(2010~2014) 논문 토픽모델링 결과

No	토픽	핵심키워드	빈도
Topic-19	전통문화	전통, 디자인, 지역, 디지털, 상품	51
Topic-20	고전문화	스토리, 고전, 문학, 소설, 민속	51
Topic-8	스토리텔링	스토리, 서사, 영화, 디지털, 영상	50
Topic-10	지역 및 도시	관광, 가치, 지역, 도시, 평가	48
Topic-13	한류	한류, 한국, 코드, 아랍, 영화	43
Topic-9	문화유산	관광, 문화유산, 해양, 지역, 세계	36
Topic-3	융합	교육, 예술, 융합, 창의, 인문	34
Topic-6	교육	교육, 전통, 한국, 디지털 예술	34
Topic-11	세계화	애니메이션, 정책, 한국, 중국, 디지털	30
Topic-18	디지털화	정보, 디지털, 기술, 문화유산, 지식	29
Topic-7	미디어	브랜드, 기업, 방송, 미디어, 국가	28
Topic-2	공간	문학, 공간, 박물관, 지역, 스토리	23
Topic-17	지역문화	역사, 방언, 서원, 보전, 경관	21
Topic-12	원형	공간, 소통, 삼국유사, 스토리, 미디어	19
Topic-5	문화이해	교육, 학습, 다문화가정, 지원, 교수	18
Topic-15	콘텐츠산업	캐릭터, 코드, 게임, 인도, 권역	16
Topic-16	제도	게임, 불교, 진흥, 재산, 유즈	16
Topic-1	지역문화	역사, 향토, 서비스, 정보, 남북	15
Topic-14	대중문화	대중, 소설, 시각, 감성, 욕망	13
Topic-4	브랜드	브랜드, 무예, 아이덴티티, 화성, 개발	11

(3) 3구간(2015~2019) 토픽모델링

다음 <표 6>은 2015년에서 2019년 11월까지의 토픽모델링의 결과를 나타낸 것이다. 이 기간에 발행된 논문에서는 교육, 지역 및 공간, 영상,

스토리텔링, 전통문화의 순서로 주제가 나타났다. 1구간과 2구간에서는 낮은 순서 위치에 있었던 교육이 3구간에서는 가장 높은 빈도를 나타냈고, 역사 및 공간과 영상 토픽이 그 뒤를 따랐다. 1구간과 2구간에서 상위권에 있었던 스토리텔링은

<표 6> 3구간(2015~2019) 논문 토픽모델링 결과

No	토픽	핵심키워드	빈도
Topic-18	교육	교육, 스토리, 학습, 고전, 한국	67
Topic-9	역사 및 공간	역사, 스토리, 지역, 공간, 문학	51
Topic-2	영상	디지털, 문화유산, 영상, 아카이브, 인문학	49
Topic-1	스토리텔링	스토리, 영화, 드라마, 신화, 제주	48
Topic-10	전통문화	전통, 디자인, 개발, 한국, 융합	48
Topic-19	축제	지역, 축제, 관광, 개발, 스토리	47
Topic-17	세계화	중국, 한류, 애니메이션, 지원, 비즈니스	43
Topic-13	한류	한류, 이미지, 한국, 의도, 국가	41
Topic-4	일본문화	일본, 요괴, 원형, 문학, 현대	38
Topic-16	정책	경제, 가치, 예술, 정책, 창조	35
Topic-12	관광	관광, 지역, 문화유산, 해양, 아리랑	31
Topic-11	문화해석	광고, 한국, 순례, 성지, 지역	30
Topic-14	문화유입	미디어, 트랜스, 매개, 대중, 한류	26
Topic-3	예술	예술, 공연, 공간, 문학, 마케팅	25
Topic-5	종교문화	종교, 불교, 의식, 박물관, 스토리	24
Topic-7	민속문화	스토리, 민속, 원형, 동화, 생태	22
Topic-6	콘텐츠	대중, 구조, 음악, 소비, 드라마	19
Topic-20	지역문화	동학, 지역, 혁명, 백제, 복식	19
Topic-15	유교문화	지역, 서원, 유교, 시조, 대전	17
Topic-8	방송	프로그램, 공연, 스포츠, 치유, 융합	15

여전히 높은 순위를 보여, 전 구간에서 높은 빈도로 등장하는 것을 확인할 수 있다. 또한, 1구간에서는 낮은 순위였고, 2구간에서는 존재하지 않았던 축제 토픽이 상위권에 있는 것을 볼 수 있어, 3구간의 연구 주제를 확인할 수 있었다. 이 외에 일본문화, 문화유입, 종교문화, 유교문화, 지역문화 등 문화콘텐츠의 다양한 토픽에 관한 연구가 수행되었음을 알 수 있다. 교육 토픽과 관련한 연구로는 스토리텔링 또는 콘텐츠화를 통한 한국문화 교육 방안에 관한 연구(박송희, 2016)가 수행되었다. 축제 토픽과 관련된 연구로는 문화콘텐츠 산업으로서의 지역축제 활성화 방안 및 사례분석(이용규, 2018) 등이 수행되었다.

2002년부터 2019년 11월까지 문화콘텐츠와 관련된 논문의 토픽모델링 결과 전 구간에서 높은 순위에 있는 토픽은 ‘스토리텔링’과 ‘전통문화’로 문화콘텐츠의 연구가 본격적으로 수행된 2002년부터 현재

까지 지속해서 연구가 수행되는 것을 확인할 수 있었다. 이 외에도 ‘한류’ 토픽은 1구간과 2구간에서는 연구가 많이 수행되었으나 3구간에서는 그만큼은 수행되지 못한 것으로 확인된다. 이 외에도 ‘교육’, ‘지역문화’ 등의 토픽 등이 지속적으로 상승하고 있는 것을 확인할 수 있었으며, 2구간에서는 문화융합 및 이해, 3구간에서는 축제, 종교, 유교 등 다양한 문화산업의 확대가 강조되고 있는 것을 알 수 있었다.

3) 키워드 네트워크 분석

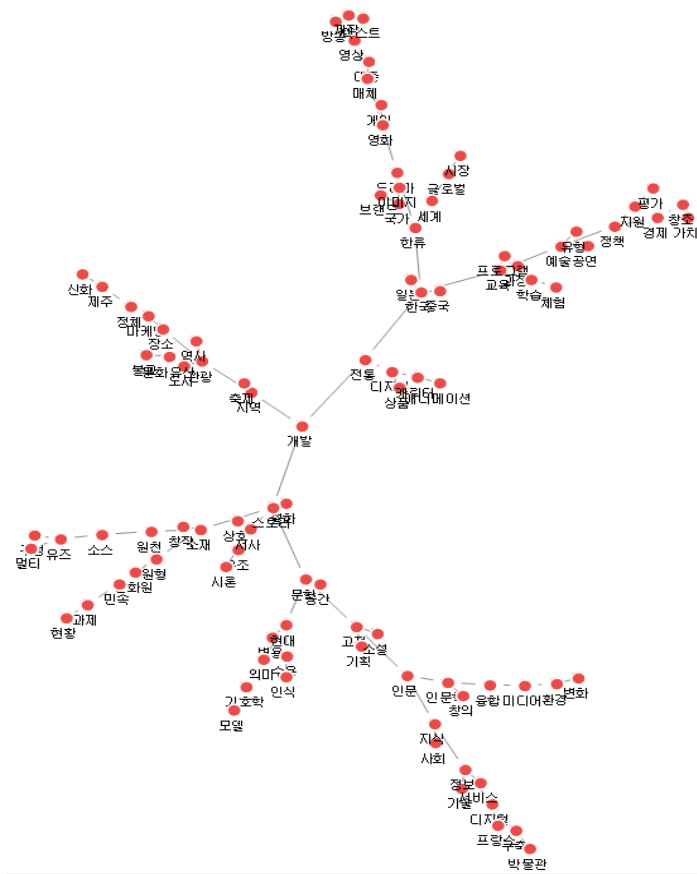
네트워크 분석에서 연결 구조의 특성을 파악하기 위한 대표적인 측정지표는 밀도, 중심성, 응집성 등이 있다. 그중에서 중심성(centrality)은 다수의 개체와 관계를 맺고 있는 노드의 영향력을 측정하는 것이며, 연결중심성은 한 노드의 중심성을 측정하는 것으로 한 노드가 다른 노드와 직접

연결되는 링크의 수로 파악할 수 있다. 즉 연결된 링크 수가 많을수록 연결중심성이 높으며, 이러한 연결중심성 분석으로 직접적인 영향력의 크기를 측정할 수 있다. 동시 출현단어 네트워크는 동시 출현 키워드 간의 관계를 분석함으로써 동시 출현 빈도가 높을수록 두 키워드 간의 연관성이 높은 것을 의미한다. 저자 키워드 간의 연관 관계를 분석하기 위하여 각 키워드를 동시 출현 관계로 분석하여 네트워크 분석을 시행하였다. 동시 출현 단어 네트워크는 유사도 분석법 중 내적치(Inner Product) 방식을 사용하여 동시 출현 관계를 구축하였으며, 관계 구성에서 같은 키워드끼리 연결

되는 셀프 루프(Self-Loop)는 제거하였다.

(1) 전 구간(2002~2019) 키워드 네트워크 분석

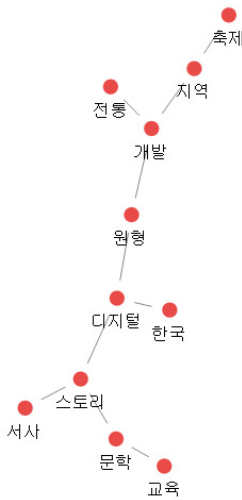
2002년부터 2009년 11월까지 발간된 논문은 총 1,693건 중 키워드 검색을 통해 2회 이상 출현 단어를 포함한 논문은 1건을 제외하고 총 1,692건의 논문에서 각 키워드가 25개 이상의 논문에 등장한 단어들만을 추출하였다. 전 구간 키워드 네트워크 분석을 통해 한국, 개발, 스토리, 문학 등의 주요 노드를 중심으로 네트워크가 연계되어 있음을 알 수 있다(<그림 2> 참조).



<그림 2> 전구간 키워드 네트워크

(2) 1구간(2002~2009) 키워드 네트워크 분석

2002년부터 2009년까지 8년간의 2회 이상 25개 이상 논문에 동시출현한 키워드를 분석한 결과는 <그림 3>과 같다. 상위 10개의 중심키워드는 ‘스토리’, ‘디지털’, ‘개발’, ‘원형’, ‘교육’, ‘한국’, ‘문학’, ‘지역’, ‘전통’, ‘인문학’ 순이었으며 동시출현 키워드 상위 총 10쌍은 분석 결과는 <표 7>과 같다.



<그림 3> 1구간 키워드 네트워크

이를 살펴보면, 디지털, 스토리, 개발 노드가 가장 중심성이 높게 나타났으며, 이들을 중심으로 다른 키워드와 연결되는 것을 볼 수 있는데, 디지털은 한국, 원형, 개발과 연계되고 있으며, 개발은 지역, 전통, 원형과 연결되어, 지역 콘텐츠 개발, 전통 콘텐츠 등의 연구가 진행되고 있다. 마지막으로, 스토리는 디지털, 서사, 문학 등과 연계되어 다양한 영역의 콘텐츠 스토리텔링에 관한 연구가 진행되었다.

가장 중심성이 높은 디지털, 스토리, 개발 등의 노드는 동시출현 빈도 역시 높게 나타나고 있음을 알 수 있다. 이를 통해 ‘문화콘텐츠의 디지털 스토리 개발’이 제1구간에서 가장 중요한 연구 키워드로 나타나고 있음을 알 수 있다. 대표적인 연구로는 문화원형의 디지털 콘텐츠화 소재개발 및 콘텐츠화 사업 모형에 대한 연구(예: 정재인, 2008; 황동열, 2003), 문화원형 디지털콘텐츠 사례 및 디지털콘텐츠 활성화에 관한 연구(예: 배수을, 2008; 최혜실, 2008), 콘텐츠의 디지털 아카이브즈(예: 서혜란, 2006; 함한희, 박순철, 2006)에 대한 주제가 연구되었다.

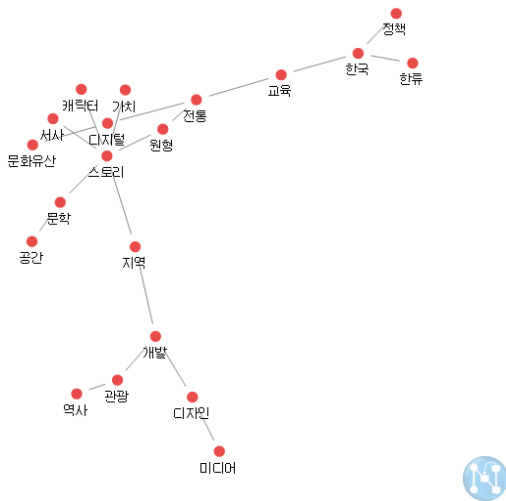
(3) 2구간(2010~2014) 키워드 네트워크 분석

2010년부터 2014년까지 5년간의 2회 이상 25개

<표 7> 1구간 키워드 연결중심성 및 동시출현 키워드(상위 10순위)

순위	키워드	Degree Centrality	동시출현 키워드	Criterion Value
1	스토리	0.066809	(디지털, 스토리)	17
2	디지털	0.065387	(디지털, 원형)	14
3	개발	0.057569	(개발, 스토리)	13
4	원형	0.054726	(개발, 원형)	13
5	교육	0.053305	(디지털, 한국)	12
6	한국	0.051883	(문학, 스토리)	12
7	문학	0.051883	(지역, 축제)	12
8	지역	0.039090	(개발, 지역)	11
9	전통	0.028429	(개발, 교육)	10
10	인문학	0.024876	(개발, 한국)	10

이상 논문에 동시 출현한 키워드를 분석한 결과는 <그림 4>와 같다. 상위 10개의 중심키워드는 ‘스토리’, ‘한국’, ‘지역’, ‘교육’, ‘개발’, ‘디지털’, ‘한류’, ‘관광’, ‘전통’, ‘가치’ 순이었으며 동시출현 키워드 상위 총 10쌍은 분석 결과는 <표 8>과 같다.



<그림 4> 2구간 키워드 네트워크

이를 살펴보면, 스토리, 한국, 지역이 가장 중심성이 높게 나타났으며, 이들을 중심으로 다른 키워드와 연결되는 것을 볼 수 있는데, 스토리는 서

사, 디지털, 문학, 원형과 연계되고 있으며, 한국은 한류, 정책, 교육과 연결되어, 한류 콘텐츠 개발, 문화콘텐츠 활성화 정책 등의 연구가 진행되고 있다. 마지막으로, 지역은 스토리, 개발 등과 연계되어 다양한 지역 콘텐츠의 활성화에 관한 연구가 진행되었음을 알 수 있다.

1구간 키워드 분석에서는 디지털, 스토리, 개발 등 가장 중심성이 높은 키워드 사이의 동시출현 빈도 역시 높게 나타났으나, 2구간 키워드 분석에서는 이들 노드가 서로 다른 노드와 연결되고 있는 것을 볼 때, 문화콘텐츠 연구영역이 지역, 한국 및 한류, 문학 및 지역 스토리텔링 등 다양한 영역으로 확대되고 있음을 알 수 있다. 한국 및 한류에 관한 연구로는 한류 문화콘텐츠의 영향, 활용 및 확산(예: 송정은, 2014; 안지현, 정철, 2014), 지역 키워드와 관련해서는 지역 문화콘텐츠의 영향, 활용 및 활성화 방안(예: 강연호, 2011; 류용재, 2012), 지역 문화콘텐츠를 활용한 지역 마케팅 방안(예: 이정준, 2011; 전영록, 박상훈, 2013)에 대한 연구가 있다.

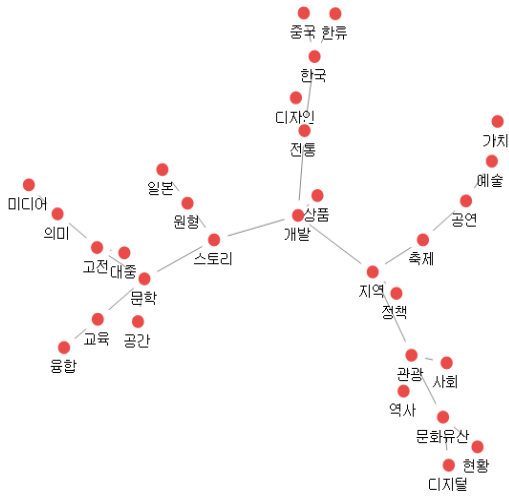
(4) 3구간(2015~2019) 키워드 네트워크 분석

2015년부터 2019년까지 5년간의 2회 이상 25

<표 8> 2구간 키워드 연결중심성 및 동시출현 키워드(상위 10순위)

순위	키워드	Degree Centrality	동시출현 키워드	Criterion Value
1	스토리	0.261905	(한국, 한류)	20
2	한국	0.234848	(서사, 스토리)	18
3	지역	0.185606	(교육, 한국)	13
4	교육	0.174242	(스토리, 지역)	13
5	개발	0.172078	(디지털, 스토리)	12
6	디지털	0.169372	(개발, 지역)	11
7	한류	0.134199	(개발, 관광)	10
8	관광	0.128788	(개발, 스토리)	10
9	전통	0.127706	(디지털, 전통)	10
10	가치	0.126623	(멀티, 유즈)	10

개 이상 논문에 동시출현한 키워드를 분석한 결과는 <그림 5>와 같다. 상위 10개의 중심키워드는 ‘지역’, ‘스토리’, ‘한국’, ‘교육’, ‘개발’, ‘문학’, ‘역사’, ‘디지털’, ‘관광’, ‘전통’, ‘한류’ 순이었으며 동시출현 키워드 상위 총 10쌍은 분석 결과는 <표 9>와 같다.



<그림 5> 3구간 키워드 네트워크

이를 살펴보면, 지역, 스토리, 한국 노드가 가장 중심성이 높게 나타났으며, 이들을 중심으로 다른

노드와 연결되는 것을 볼 수 있다. 지역, 스토리, 한국이 중심성이 높게 나타난 것은 2구간과 동일하며, 다만 한국보다 지역이 가장 높은 중심성을 보이는 등 순위의 차이가 있음을 알 수 있다. 지역과 관련된 키워드는 개발, 관광, 축제였으며, 스토리와 동시출현 빈도가 높게 나타난 키워드는 개발, 한국, 문학으로 나타났다. 마지막으로 한국과 동시출현 빈도가 높게 나타난 키워드는 한류, 스토리, 전통이었다.

지역과 관련한 키워드는 2구간의 경우 스토리, 개발이었으나, 3구간은 개발, 관광, 축제로 나타난 것은 2구간은 지역 문화콘텐츠의 사례 및 활용 가능성에 관한 연구가 집중된것에 비해, 3구간은 지역 문화콘텐츠 기반의 관광 및 축제 활성화 방안 또는 관광 및 축제의 지역 문화콘텐츠 활용 가능성에 대한 연구(예: 송은아, 임준목, 2018; 최락인, 2019) 등으로 지역 콘텐츠에 관한 연구가 세분되고 있기 때문으로 해석된다. 스토리와 동시 출현빈도가 높은 키워드는 2구간의 경우 서사, 디지털, 문학, 원형으로 나타난 것에 비해, 3구간은 개발, 한국, 문학으로 나타나서, 3구간은 한국 및 문학 스토리의 확장 가능성에 집중된 것(예: 조현구, 2017)이 특징이다.

<표 9> 3구간 키워드 연결중심성 및 동시출현 키워드(상위 10순위)

순위	키워드	Degree Centrality	동시출현 키워드	Criterion Value
1	지역	0.064354	(개발, 지역)	26
2	스토리	0.058877	(개발, 스토리)	20
3	한국	0.045185	(스토리, 지역)	19
4	교육	0.041534	(개발, 전통)	17
5	개발	0.041077	(관광, 지역)	16
6	문학	0.030580	(한국, 한류)	16
7	역사	0.027841	(스토리, 한국)	15
8	관광	0.027841	(전통, 한국)	15
9	전통	0.026472	(문학, 스토리)	13
10	한류	0.025559	(지역, 축제)	13

5. 결 론

본 연구는 2002년부터 2019년까지 문화콘텐츠 연구 경향을 분석하는 것을 목적으로 한다. 이를 위하여 KCI 등재된 논문을 검색하여, 키워드 네트워크 분석 및 토픽모델링을 실시하였다. 총 1,693건의 논문을 3구간으로 구분하여 분석한 결과는 다음과 같다.

첫째, 키워드 빈도분석 결과, ‘스토리’, ‘개발’, ‘교육’, ‘한국’, ‘지역’, ‘전통’의 키워드가 전체적으로 꾸준히 많이 연구되고 있는 핵심 키워드로 나타났다. 1구간은 스토리텔링 또는 디지털 스토리텔링이 중요하게 강조되었으며, 2구간은 한국, 한류, 지역, 교육이 강조되었다. 3구간은 2구간과 마찬가지로 지역, 한국, 교육, 역사 등이 강조되고 있음을 알 수 있다.

둘째, 토픽모델링 분석 결과, 1구간은 스토리텔링, 디지털화, 한류, 지역 문화에 대한 토픽이 중요하게 나타났으며, 이 중 디지털 스토리가 가장 순위가 높은 토픽으로 나타났다. 2구간은 전통문화, 고전문화, 지역 및 도시, 한류에 대한 토픽이 높은 순위로 나타났으며, 디지털화는 1구간에 비해 순위가 내려간 것을 알 수 있다. 3구간은 교육, 지역 및 공간, 영상 등의 토픽이 높은 순위를 차지하고 있으며, 축제 토픽이나 종교문화, 유교문화 등 다양한 종류의 문화콘텐츠에 대한 토픽이 중요하게 부각되었음을 파악하였다.

셋째, 키워드 네트워크 분석을 통해 한국, 개발, 스토리, 문학 등이 전 구간에 주요 노드로 조사되었다. 1구간의 경우, 문화콘텐츠의 디지털 스토리 개발이 가장 중요한 노드로 나타났다. 2구간은 한국, 한류, 지역, 교육 등의 노드가 부각되었으며, 연구영역이 지역, 한국 및 한류, 문학 및 지역 스토리텔링 등 다양한 영역으로 확대되고 있음을 파악하였다. 3구간의 경우, 지역, 스토리, 한국 노드

가 계속 강조되어 나타났으며, 축제, 관광 등의 노드가 부상한 것이 특징이다. 이는 축제 및 관광 노드가 지역 노드와 결합하여 지역 문화콘텐츠 기반의 관광 및 축제 활성화에 관한 연구로 세분되고 있기 때문으로 해석된다. 또한, 스토리의 경우 개발, 한국, 문학 등의 노드와 동시출현 빈도가 높게 나타났는데, 이는 3구간에서 문화콘텐츠 연구영역이 한국 및 문학 스토리의 확장 가능성에 집중되고 있는 것을 의미한다.

연구결과, 문화콘텐츠 연구 동향은 다음과 같다. 빈도분석, 키워드 네트워크 분석, 토픽모델링 분석을 통해 한국, 개발, 스토리, 문학 등이 모든 영역에서 중요하게 나타났다. 제1구간으로 구분된 2002년부터 2009년은 문화콘텐츠 연구의 도약기로, 문화콘텐츠와 정보기술의 결합으로 인해 문화콘텐츠가 국내에서 중요한 산업으로 진입하게 된 시기라고 볼 수 있다. 이를 통해 다양한 원천콘텐츠의 디지털화, 디지털 기술을 이용한 원천콘텐츠의 이용 가능성에 관한 연구가 활발하게 이루어진 시기라고 할 수 있다. 제2구간인 2010년부터 2014년은 성장기로 1구간의 디지털화 기술을 통해 입증된 문화콘텐츠의 활용 가능성에 대한 인정을 바탕으로 한류, 서사, 문학, 지역 등 다양한 영역으로 문화콘텐츠 연구가 본격적으로 성장한 시기로 해석된다. 2015년부터 2019년 사이의 3구간은 성숙기 단계로, 이 단계는 1-2구간을 통해 인정받는 문화콘텐츠 산업의 확산을 관광, 축제, 역사, 종교, 지역 및 공간, 교육 등 다양한 산업 분야로 성숙시키기 위한 다양한 연구가 수행된 시기로 해석된다.

4차산업혁명 시대의 ICT와 결합한 문화콘텐츠 산업의 발전 가능성을 고려할 때, 문화콘텐츠와 관련 연구는 앞으로 더욱 확대될 것이며 다양하게 연구될 것이다. 본 연구는 약 20년간 수행되어 온 문화콘텐츠의 연구 동향을 토픽모델링과 네트

워크 분석을 통해 다각적으로 조사하였다는 점에서 의미가 있다. 다만 본 연구는 회귀분석을 시행하지 않았으므로, 문화콘텐츠의 상승토픽과 하락토픽을 도출하지 않았다는 한계가 있다. 향후 지속적인 연구동향에 대한 분석을 통하여 문화콘텐츠의 발자취와 발전 방향을 고찰할 필요가 있다.

참 고 문 헌

- 강연호 (2011). 디지털 시대, 지역문화콘텐츠 개발의 의미와 방향. <열린정신 인문학연구>, 12(2), 26-284.
- 김세훈 (2009). 한국 영상콘텐츠 개발의 활성화를 위한 한-중, 한-일 문화콘텐츠 교류의 현황과 전망. <애니메이션연구>, 5(3), 44-62.
- 김세훈 (2013). 문화예술과 콘텐츠 산업 융합 활성화 방안: 정책적 차원을 중심으로. <문화산업연구>, 13(8), 45-53.
- 김유빈 · 조남욱 (2017). 토픽모델링과 사회연결망분석을 이용한 사회연결망분야 연구 동향 분석. <한국콘텐츠학회 학술대회논문집>, 5680-5691.
- 김의준 · 권영현 (2012). 지역문화산업의 네트워크 분석: 가치사슬의 제한적 적용. <한국지역개발학회지>, 24(8), 25-42.
- 김창식 · 최수정 · 광기영 (2017). 토픽모델링과 시계열회귀분석을 활용한 정보시스템분야 연구동향 분석. <디지털콘텐츠학회논문지>, 18(6), 1143-1150.
- 김형석 · 박민식 · 강필성 (2016). 토픽 모델링과 사회연결망을 통한 딥러닝 연구동향 분석. <대한산업공학회 춘계공동학술대회 논문집>, 1877-1899.
- 나상태 · 김자희 · 정민호 · 안주언 (2016). 토픽 모델링을 이용한 시물레이션 연구 동향 분석. <한국시물레이션학회 논문지>, 25(3), 107-116.
- 노시훈 (2010). 여수지역 해양문화유산의 문화콘텐츠 개발에 관한 연구. <호남문화연구>, 47, 25-50.
- 류용재 (2012). 지역 문화콘텐츠산업의 문제와 전망. <우리춤과 과학기술>, 8(1), 127-160.
- 류준호 · 윤승금 · 이영주 (2010). 문화콘텐츠 관련 연구에 대한 메타분석: 연구 분야, 목적, 방법을 중심으로. <언론과학연구>, 10(1), 124-165.
- 류철균 · 서성은 (2008). 온라인게임의 트랜스미디어 스토리텔링 연구: <메이플스토리>를 중심으로. <어문학>102, 439-470.
- 문화산업진흥기본법. 2조 3·4호. 국가법령정보센터.
- 박송희 (2016). 스토리텔링을 통한 한국 고전문학 교육방안: <박씨전>을 중심으로. <우리문학연구>, 52, 159-188.
- 박양우 (2017). 언어네트워크분석을 통한 국내 문화정책 연구동향 분석(2008-2017). <한국콘텐츠학회논문지>, 17(11), 371-382.
- 박자현 · 송민 (2013). 토픽모델링을 활용한 국내 문헌정보학 연구동향 분석. <정보관리학회지>, 30(1), 7-32.
- 박준형 · 오효정 (2017). 국내 기록관리학 연구동향 분석을 위한 토픽모델링 기법 비교: LDA와 HDP를 중심으로. <한국도서관정보학회지>, 48(4), 235-258.
- 배수을 (2008). 무용문화원형의 디지털콘텐츠 개발사례와 활성화 방안. <무용예술학연구>, 23(23), 5986.

- 서혜란 (2006). 문화원형콘텐츠의 장기보존에 관한 연구: 디지털 소리자료를 중심으로. <한국비블리아학회지> 17(2), 65-82.
- 송은아·임준묵 (2018). 문화산업으로서 축제콘텐츠의 활성화방안. <한국엔터테인먼트산업학회논문지>, 12(8)55-63.
- 송정은 (2014). 경제한류와 문화한류가 한국 브랜드 이미지 형성에 미치는 역할. <한류비즈니스연구>, 1(1)109-143.
- 신유미·박옥남 (2019). 장서개발관리 분야 최근 연구동향 분석에 대한 연구. <정보관리학회지>, 36(2)105-129.
- 안지현·정철 (2014). 한류의 지속가능성과 한류문화콘텐츠의 성공과정에 대한 탐색적 고찰. <관광학연구>, 38(7), 215-238.
- 안지형·김운미 (2008). 문화원형 디지털콘텐츠화 사업을 중심으로 본 무용콘텐츠 활성화 방안 연구. <우리춤과 과학기술>, 6(1), 99-129.
- 오정심 (2020). 빅데이터 토픽모델링 및 네트워크 분석을 통한 문화콘텐츠학 지식구조 연구. <문화정책논총>, 34(2), 35-69.
- 윤홍근 (2011). 학위논문의 메타분석을 통한 문화콘텐츠 연구동향: 학국외대, 한양대, 인하대, 고려대, 동국대등 5개 학교의 학위논문을 중심으로. <글로벌문화콘텐츠>, 7, 149-180.
- 이성신 (2015). 키워드 네트워크 분석을 통한 도서관마케팅 연구 경향 분석. <한국문헌정보학회지>, 50(3)383-402.
- 이용규 (2018). 지역사회의 지역문화콘텐츠 개발 및 활용사례 분석. <글로벌문화콘텐츠>, 37, 89-108.
- 이정준 (2011). 지역문화콘텐츠를 활용한 도시마케팅 보드게임 제작 및 활용. <디지털콘텐츠학회지>, 12(4) 531-539.
- 전영록·박상훈 (2013). 지역문화콘텐츠를 통한 지역이미지 개선 및 지역발전성과 기대 비교: 실경산수유지컬 왕의 나라 관람객을 중심으로. <관광연구>, 27(6), 367-384.
- 정재진 (2008). 처용설화의 문화원형 디지털콘텐츠화에 관한 연구. <디지털콘텐츠학회지>, 9(4), 591-600.
- 조헌구 (2017). 융합적 문화콘텐츠 구축을 위한 고전문학연구의 확장가능성. <어문논총>, 72(72), 191-223.
- 최락인 (2019). 지역문화관광 콘텐츠를 통한 관광자원개발 활성화정책 방안. <예술인문사회융합멀티미디어논문지>, 9, 365-374.
- 최혜실 (2008). 한국문화의 세계화를 위한 한류의 지속 및 확장 방안 연구: 문화원형 디지털 콘텐츠 사업 사례 중심으로. <Comparative Korean Studies>, 16(2), 549-568.
- 태지호 (2016). 문화콘텐츠 연구 방법론의 토대에 대한 모색. <인문콘텐츠>, (41), 75-96.
- 한국콘텐츠진흥원 (2019). 2019콘텐츠산업 통계조사보고서. 서울: 한국콘텐츠진흥원
- 함한희·박순철 (2006). 디지털 아카이브즈의 문제점과 방향: 문화원형 콘텐츠를 중심으로. <한국비블리아학회지> 17(2), 23-42.
- 홍병선 (2014). 문화예술에서 문화콘텐츠산업 융합연구의 방향과 그 대안 모색. <교양교육연구>, 8(1), 43-450.
- 황동열 (2003). 문화원형의 디지털 콘텐츠 개발 모형에 관한 연구. <한국비블리아학회지>, 14(1), 46.
- Blei, D. M., Andrew Y. Ng, & Michael I. J. (2003), Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3(1), 993-1022.

The Trend Analysis of ‘Cultural Contents’ Research Using Topic Modeling and Keyword Network Analysis

Yohan Min, Giyoung Kim, Ok Nam Park

Sangmyung University

As interest in original Korean content expands, and due to the development of digital technology, the production, distribution, and utilization of cultural content varies. The area of cultural content has become an important industry beyond the way we live, and it has been emphasized as a vital source of value-added creation in the country. The purpose of this study is to examine flow in the area of cultural content for various academic disciplines. To analyze the research trends of cultural content, topic modeling and keyword network analysis were conducted based on articles published in KCI from 2002 to 2019. The study has shown that Korea, development, stories, and literature are among the most important keywords. The first period (2002 to 2009) was a leap forward in cultural content research, and the combination of cultural content and information technology led to cultural content becoming an important industry in the Republic of Korea. It was a time when research on the digitization of a variety of original cultural content, and the availability of original content using digital technology, was actively conducted. The second period (2010 to 2014) is interpreted as a growth stage for cultural content research in various areas—Chinese culture, Hallyu (Korean wave), narrative, literature, and regions—based on recognizing the possibility of using proven cultural content identified through digitization technology in the first period. The third period (2015 to 2019) was the maturity stage, in which various studies were conducted into the expansion of the cultural content industry as recognized through the first and second periods in a variety of industries, including tourism, festivals, history, religion, region and space, and education. Topic modeling and network analysis shows this study is meaningful in that the research trend has continued for about 20 years in various ways.

Keywords: Cultural Content, Research Trends, Network Analysis, Topic Modeling