

문헌정보학 교과목의 주요 키워드 관계 분석

Analysis of Keyword Relationships in Lectures of Library and Information Science

윤정민, 김나무, 김학래
중앙대학교 문헌정보학과

Jungmin Yun (cocoro357@naver.com), Namoo Kim (2913338@naver.com), Haklae Kim(haklaekim@cau.ac.kr)

요약

학문은 서로 독립적이지만, 동시에 서로 밀접하게 연결되어 있다. 문헌정보학은 전통적인 연구 주제와 정보기술 분야의 끊임없는 연계를 통해 융합 학문으로 진화하고 있다. 예를 들어, 데이터 사이언스와 빅데이터와 같은 기술은 컴퓨터 공학에서 활발하게 연구되고 있지만, 문헌정보학의 교과목에 빠르게 적용되고 있다. 교과과정은 내외부 환경 변화에 대응하는 대학의 전략이 반영된 결과이다. 일반적으로 교과과정의 개편은 새로운 교과목을 신설하거나 기존 교과목의 수업 내용을 업데이트하는 방식으로 진행된다. 본 연구는 문헌정보학을 구성하는 주요 학문 영역의 관계를 분석한다. 이를 위해 강의계획서에 기술된 정보를 추출하고 교과목의 특징과 교과목 사이의 연관성을 분석한다. 특히, 강의 설명에 포함된 한글 정보의 정확한 분석을 위해 개체명 인식 기법을 적용하고, 추출된 개체와 문헌정보학의 학문분류체계의 관계를 분석하기 위해 네트워크 분석 기법을 적용한다.

■ 중심어 : 데이터 분석 | 개체명 인식 | 네트워크 분석 | 문헌정보학 | 융합

Abstract

The sciences are independent of each other, but at the same time, they are closely connected to each other. Library and Information Science (LIS) is evolving into a multi-disciplinary subject by interlinking between traditional topics and information technology fields. For example, technologies such as data science and big data are being actively studied in computer science, but are rapidly being applied to subjects in LIS. A curriculum is the result of reflecting the university's strategy to respond to changes in the internal and external environment. In general, to reorganise a curriculum is divided into 1) a new course is newly established, or 2) the contents of an existing course are updated. This study analyses the relationships between major subjects of LIS. For this, the information described in syllabus is extracted and the relationship between the characteristics of the subjects. In particular, the entity name recognition technique is applied to accurately analyse a set of lecture descriptions in Korean, and the network analysis technique is applied to analyse the relationships between the extracted entities and the academic classification system of LIS.

■ keyword : Data Analysis | Named Entity Recognition | Network Analysis | Library & Information Science | Convergence

I. 서 론

융합의 사전적 의미는 다른 종류의 것이 녹아서 서로 구별 없이 하나로 합하여지는 것을 말한다. 서로 다른 학문의 결합은 새로운 지식을 창조하는 과정과 결과를 포함한다. 문헌정보학은 주로 문헌과 관련된 모든 사실이나 현상을 논리적·과학적으로 규명하고 사회적 적용 가능성을 탐구한다 [1]. 협의적으로 보면, 문헌정보학은 학문의 기초가 되었던 도서관학, 도서관의 고증과 해석을 중심으로 하는 서지학, 기록물의 관리와 보존에 대한 기록관리학, 정보의 조직과 분류를 위한 정보조직, 정보의 효과적 처리를 다루는 정보학으로 구성된다 [2]. 한국에서 문헌정보학은 도서관학으로 불렸으나, 1990년 이후 정보기술에 대한 패러다임의 변화에 따라 융합적인 학문으로 발전하고 있다. 한국에서 문헌정보학을 학과명으로 채택한 것은 1985년 전남대학교가 최초이다. 현재 문헌정보학과는 2년제 전문대학 5개, 4년제 학사과정 32개 설치되어 있다 [3].

인공지능, 데이터 사이언스, 빅데이터 등 디지털 패러다임의 변화는 융합을 지향하는 문헌정보학에 새로운 방향성을 탐색하게 한다. 문헌정보학은 대규모 데이터의 분석과 처리에 대한 다양한 이론의 융합을 통해 전통적인 경계를 넘어 다 학문적 체계로 진화하고 있다 [4]. 예를 들어, 데이터 마이닝, 빅데이터, 데이터 사이언스, 디지털 큐레이션 등 데이터 분석에 대한 다양한 이론을 정보학의 테두리에 포함하거나, 새로운 교과목을 개설하고 있다 [5][6].

대학의 커리큘럼은 학문적 비전과 지향성을 반영한다. 개별 학과는 대내외 환경의 변화에 대응하기 위해 학과의 커리큘럼을 지속적으로 업데이트한다. 문헌정보학의 커리큘럼, 학교별 교과목 특성에 대한 연구는 다양하게 진행되고 있다 [7]-[9]. 대부분의 연구는 교과목 이름과 주제어 중심이거나 주기별로 변화되는 교과목을 분석하는 경향을 보인다. 그러나, 교과목의 변경은 일정 기간마다 이루어지는 특성을 갖고 있고, 강의의 실제 내용을 파악하는데 한계가 있다. 본 연구는 강의 계획서에 포함된 교과목명과 내용을 분석하고, 교과목 사이의 연관 관계를 탐구한다.

본 논문의 구성은 다음과 같다. 2장은 문헌정보학의 커리큘럼을 분석한 선행연구를 고찰하고, 본 논문에서

적용한 개체명 인식과 네트워크 분석 기법을 소개한다. 3장은 교과목 분석을 위한 데이터 수집, 분석 과정을 포함해 전반적인 연구 방법을 기술한다. 4장은 개체명 인식과 네트워크 분석을 통해 도출한 결과를 논의한다. 마지막으로 5장은 논문을 요약한다.

II. 관련 연구

1. 교과목 연구

교과목은 수업 설계 [10], 수요와 효과 분석[11], 신규 교과목의 개발 [12]-[13] 등 다양한 목적을 분석하는데 주요한 도구로 활용되고 있다. 특히, 학문의 융합성을 발견하기 위해 교과목을 분석하는 사례는 지속적으로 증가하고 있다 [14].

문헌정보학의 교과목 연구는 데이터의 수집 범위, 분석 기법이 다양하다[15]-[18]. 노영희는 2005년부터 2020년까지 약 15년간의 문헌정보학과 석박사 교과과정의 변화를 분석하고 있다. 이 연구는 교과과정을 7가지 학문주제분야로 구분하고 각 3~4년 단위로 각 분야의 변동율을 분석하고 있다 [15]. 모예림은 한국의 34개 대학에서 개설한 문헌정보학과 커리큘럼 현황을 분석한다. 이 연구는 교과목을 기술통계 기반으로 분석하고, 네트워크 분석 기법을 통해 문헌정보학 분류 체계와 개별 교과목의 관계성을 분석하고 있다. 그러나, 분석대상이 강의명에 한정되어 강의내용에 대한 구체적인 분석이 수행되지 않았다 [4].

2. 개체명 인식

개체명 인식 (Named Entity Recognition)은 문맥의 의미를 파악하고 비정형 텍스트에서 인물, 조직, 장소 등 미리 정의된 분류의 개체를 정의하는 자연어 처리 기술이다. 비정형 텍스트에 포함된 키워드는 형태소 분석을 통해 필요한 분석 단위를 구분할 수 있지만, 키워드의 정확한 의미를 정의하기 위해 개체명 인식은 필수적이다. 개체명 인식은 규칙 기반, 비지도 학습, 변수 기반 지도 학습 등 다양한 알고리즘이 연구되었고 [19], 최근 딥러닝 모델을 적용하는 연구가 활발히 진행되고 있다 [20][21]. 한국어 개체명 인식에 대한 연구도 LSTM [22], ALBERT [23] 등 딥러닝 기술의 적

용을 통해 인식을 향상시키는 연구가 있다. 한국어 개체명 인식을 위해 활용할 수 있는 데이터세트가 있고, 데이터세트에 따라 구분하는 개체는 차이가 있다. 예를 들어, 국립국어원의 데이터세트는 장소 (LC), 날짜 (DT), 기관 (OG), 시간 (TI), 인물 (PS)로 구분하고¹, 한국정보기술협회는 인물 (PS), 학문분야 (FD), 이론 (TR), 인공물 (AF), 기관 (OR), 지역 (LC), 문명 (CV), 날짜 (DT), 시간 (TI), 수량 (QT), 이벤트 (EV), 동물 (AM), 식물 (PT), 물질 (MT), 용어 (TM)로 구분된 데이터세트를 제공하고 있다².

3. 네트워크 분석

네트워크란 노드 (node)와 간선 (edge)으로 이루어진 자료구조이고, 네트워크 분석은 네트워크의 구조와 노드의 상호 관계를 파악하는 것을 말한다 [24]. 일반적으로 네트워크 분석은 그래프 이론을 바탕으로 모든 객체 사이의 관계를 정량적으로 분석하는데 적용된다 [25]. 사회 연결망과 같이 사회 구성원 사이의 관계를 분석하는 사례는 사회과학의 대표적인 연구 분야이다 [26]. 특정 학문을 구성하는 주제와 주제의 관계성을 분석하는 연구는 네트워크 분석의 대표적 응용 사례이다. 문헌정보학 [27], 간호학 [28], 의료 [29], 생태학 [30] 등 자연과학 분야에서 응용 연구가 활발하게 진행되고 있다.

III. 데이터 수집과 연구방법

문헌정보학과 커리큘럼은 교육부 통계포털에서 '문헌정보학', '기록관리학' 주제에 대해 전국의 대학에서 개설한 2018년과 2019년 데이터를 수집한다. 수집한 데이터는 대학 34개, 교과목 1,519개이다. 데이터 분석을 위해 모든 교과목은 8가지의 세부분류 ('도서관경영', '정보조직', '문헌정보학일반', '정보학', '정보서비스', '서지학', '기록학', '기타')로 구분된다. 수집한 데이터에서 강의명과 강의설명은 중복되는 사례가 다수 발견된다. 예를 들어, 2018년과 2019년에 개설된 강의에서 306개가 중복되어 있고, 195개가 강의설명이 일

치한다. 한편, 2019년 데이터는 강의 설명이 공백이거나 강의 설명이 강의명과 일치하는 사례가 각각 35개, 18개이다. 즉, 2018년과 비교하면, 2019년의 강의 설명이 업데이트되지 않았고 (81.05%), 불완전한 상태이다. 이와 같은 이유로, 본 연구의 분석 대상은 2018년 데이터로 한정한다. 한편, 주제 분류의 기타 항목은 진로탐색, 교양교육과 관련된 교과목이 포함되어 문헌정보학의 고유 주제로 포함하는 않는다. 해당 항목을 제외시키고, 연구 대상은 최종적으로 34개 대학의 878개의 교과목으로 한정한다.

본 연구는 강의계획서에 기술한 정보를 분석해 교과목의 특징과 교과목 사이의 연관성을 분석한다. 이를 위해, 강의계획서에 있는 강의명과 강의설명 정보를 추출하고 분석한다. 특히, 강의 설명에 포함된 한글 정보의 정확한 분석을 위해 개체명 인식 기법을 적용하고, 추출된 개체와 문헌정보학의 학문분류체계의 관계를 분석하기 위해 네트워크 분석 기법을 적용한다. 한국어 개체명 인식은 한국전자통신연구원(Electronics and Telecommunications Research Institute: ETRI)의 언어 분석 API를 적용한다. ETRI가 제공하는 개체명 인식 API는 TTA 표준 개체명 태그셋³을 사용하여 특정 개체를 표현하는 단어의 의미 정보를 인식하는데 효과적이다. 수집된 데이터세트에서 한국어 개체를 추출하는 과정은 KoalaNLP⁴를 적용한다.

IV. 데이터 분석

1. 기술통계 분석

전국 34개 대학에 개설된 교과목은 878개이다. 정보학은 229개로 개설된 빈도가 가장 높다. 도서관 경영 (174), 정보서비스(144), 자료조직(140), 문헌정보학일반(104)도 비교적 높은 분포를 갖고 있지만, 서지학 (50), 기록학(37)은 상대적으로 낮다.

1 <https://tinyurl.com/tfke6ms2>

2 <https://tinyurl.com/4tvefqbw>

3 <https://tinyurl.com/47k9hjsa>

4 <http://koalanlp.github.io/koalanlp/>

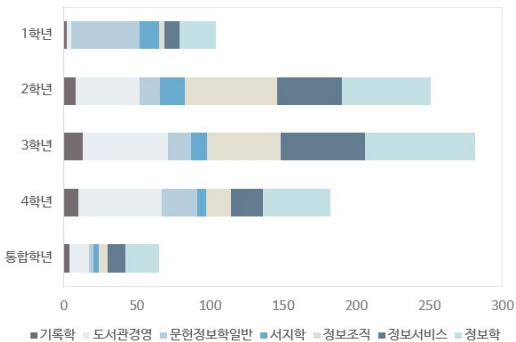


그림 1. 학년별 교과목 현황

각 대학은 전공 교과목을 2학년과 3학년에 주로 편성하는 경향이 있다. 그림 1에서 보듯이, 학년별로 개설된 교과목은 1학년(103개), 2학년(251개), 3학년(280개), 4학년(182개)이다. '빅데이터와 디지털 리터러시', '영화로 보는 도서관과 정보세상'과 같이 학년에 관계없이 개설되는 교과목도 다수 존재한다(62개). 전체 대학에 개설된 교과목은 학년별로 중점 과목이 존재한다. 예를 들어, 1학년은 문헌정보학일반(45.6%)의 비중이 높고, 2학년과 3학년은 각각 정보조직(25.0%)과 정보학(26.7%) 중심으로 편성되어 있고, 4학년은 도서관경영(31.3%)이 높은 비율을 갖고 있다.

대학별로 편성된 교과목은 그림 2와 같다. 전체의 18개 대학(52.94%)에서 정보학 관련 교과목이 가장 높은 빈도를 갖고 있기 때문에, 정보학에 대한 교육이 강조되고 있다고 해석할 수 있다. 개설 교과목을 기준으로 정보학 비중이 높은 대학은 성균관대학교(53.57%)와 연세대학교(50%)이다. 전북대학교는 개설된 교과목의 빈도는 19개로 가장 많지만, 개설된 교과목이 많기 때문에 정보학 분야의 비율은 상대적으로 낮다(29.23%). 실제 전북대학교는 정보학과 더불어 도서관경영, 정보서비스, 정보조직의 개설 과목도 다른 대학과 비교하여 높은 편이다. 반면, 기록학은 26개 대학(76.47%)에서 1개 이하의 강의를 개설하고 있다. 경북대학교는 기록학 분야의 과목을 5개 개설하고 있다. 다른 대학에 비교하면 비중이 높지만, 전체 교과목이 28개인 점을 고려하면 기록학 분야의 교과목은 상대적으로 낮다. 이와 유사하게, 서지학은 21개 대학(61.76%)에서 1개 이하의 교과목을 개설하고 있다. 서지학 분야는 전체 개설 교과목의 5.6%를 갖고 있지만, 대진대학

교에서 상대적으로 많은 교과목을 갖고 있다(22%).

	기록학	서지학	문헌정보학일반	정보조직	정보서비스	도서관경영	정보학	total
건국대학교	1	1	2	5	5	7	6	27
경북대학교	5	1	2	5	4	5	6	28
경성대학교	1	1	4	3	5	10	5	29
광주대학교	1	2	5	7	6	5	7	33
대구가톨릭대학교	3	1	5	2	5	4	3	23
덕성여자대학교	1	0	3	3	3	4	8	22
동덕여자대학교	1	3	4	4	2	2	10	26
부산대학교	1	4	5	3	4	5	8	30
상명대학교	1	2	2	5	4	7	5	26
숙명여자대학교	4	1	4	4	4	4	6	27
이화여자대학교	3	1	1	3	3	4	3	18
전북대학교	1	1	5	6	11	7	9	40
중앙대학교 서울캠퍼스	1	5	2	5	5	6	6	30
충남대학교	2	3	2	4	3	5	9	28
한남대학교	1	0	2	5	3	6	10	27
한성대학교	3	0	2	0	3	3	4	15
강남대학교	1	3	2	4	3	10	6	29
경기대학교	1	0	3	5	7	7	6	29
경일대학교	0	0	1	4	8	3	5	21
나사렛대학교	0	0	2	2	3	5	2	14
대구대학교	0	0	3	3	2	7	5	20
대진대학교	0	7	5	5	4	6	4	31
동원대학교	1	1	5	6	6	4	8	31
명지대학교 인문캠퍼스	0	3	2	3	3	6	8	25
서울여자대학교	0	0	2	3	4	3	6	18
성균관대학교	0	2	1	4	2	4	16	29
신라대학교	1	1	4	6	5	6	5	28
연세대학교	0	0	3	3	5	3	14	28
인천대학교	1	1	7	5	4	3	6	27
전남대학교	2	4	7	5	4	4	6	32
전주대학교	0	0	2	5	3	6	10	26
경주대학교	0	1	0	5	5	7	6	24
계명대학교	0	0	4	3	5	2	10	24
충부대학교	0	2	1	5	3	5	2	18

그림 2. 대학별 개설 교과목 현황

2. 개체명 인식

수집된 강의계획서에서 추출한 개체명은 1,733개이다. TTA 표준 개체명 태그셋은 다양한 분야에 등장하는 개체 유형을 중심으로 15개 대분류, 146개 세부 개체 유형을 정의하고 있다. 추출된 개체명은 TTA 표준 태그셋의 14개의 대분류와 54개의 세부 개체 유형을 갖는다. 한편, DATE(DT), ANIMAL(AM), QUANTITY(QT), MATERIAL(MT)와 같이 범용적으로 적용되는 개체명은 문헌정보학과 관련성이 낮아 분석 대상에서 제외시킨다. 그 결과, 10개 대분류와 42개의 세부 분야에 포함된 1,579개의 개체가 추출되었다. 한편, 추출된 개체명은 띄어쓰기, 한 글자 어휘, 복합어휘, 형용사 등 개체명으로 부정확할 수 있는 어휘를 정제하고, 최종적으로 1,506개의 개체명을 분석 대상으로 포함한다.

문헌정보학의 세부 분야와 개체명 태그 세트의 매핑 결과는 그림 3과 같다. 히트맵에서 짙은 푸른색은 해당 개체명의 비율이 높은 것을 의미한다. 상위 5개의 태그 세트는 CV_POSITION (20.4%), FD_SOCIAL_SCIENCE (16.7%), OGG_LIBRARY (15.1%), CV_OCCUPATION (10.5%), TMI_SERVICE (7.9%)이다. CV_POSITION은 직위나 직책에 대한 명

칭을 구분하고 있는데 학생, 교수, 사서와 같은 어휘가 이에 해당한다. FD_SOCIAL_SCIENCE는 사회과학 관련 학문 분야와 관련된 어휘를 분류하고 그 예로는 문헌정보학, 서지학, 정보학, 한국학 등이 있다. OGG_LIBRARY은 공공도서관, 대학도서관, 디지털도서관, 학교 도서관 등의 도서관이나 도서관 관련 기관 및 단체를 의미한다. CV_OCCUPATION은 교수나 정보전문가, 학생 등의 직업 명칭과 관련된 태그이다. TML_SERVICE는 IT 서비스 용어를 의미하며, 검색엔진, 웹사이트, 인터넷, 정보서비스 등을 포함한다. 식별된 개체명은 문헌정보학의 세부분류에 따라 차이가 있다. 문헌정보학일반에 해당되는 개체명들은 FD_SOCIAL_SCIENCE(37%)가 가장 높다. 도서관경영, 정보학, 서지학, 기록학, 정보서비스, 정보조직은 각각 OGG_LIBRARY(41%), TML_SERVICE(17%),

그러나 개체명 인식의 결과는 한계점을 갖고 있다. 첫째, 개체명 인식의 정확도가 완전하지는 못하다. 예를 들어, '초록'은 맥락상 'abstract'의 의미를 지니지만 동음이의어인 '초록(green)'으로 인식되어 색을 나타내는 TM_COLOR로 태깅되는 문제가 있다. '문맥(context)'이 세포나 조직 및 기관에 대한 명칭인 'TM_CELL_TISSUE'로 인식되는 오류가 있다 (2.31%). 한편 '독서토론'이 '서토론'으로 인식되는 등 어휘가 잘못 분절되어 개체명 인식이 올바르게 이루어지지 않는 경우도 존재한다 (1.67%). 둘째, 개체명이 사용된 맥락에 따라 하나의 개체가 여러 개체명 태그들로 나타난다 (0.81%). 예를 들어, '도서관학'이라는 개체는 'FD_SOCIAL_SCIENCE', 'OGG_EDUCATION', 'FD_ART' 등 다수의 태그로 인식되었다. 마지막으로, 개체명 인식의 태그는 개체명의 의미적인 범주만 나타낼 수 있지만 개체가 되는 단어 자체와 그에 대한 정확한 의미 파악이 어렵다. 따라서, 단어 사이의 관계를 파악하는 것이 어렵다. 이를 보완하기 위해 문헌정보학 세부분류별로 네트워크 분석을 시행하였다.

3. 네트워크 분석

문헌정보학 학문 분류와 강의계획서에서 추출한 개체명으로 구성된 네트워크는 334개의 노드와 460개의 간선으로 구성된다. 노드 사이의 평균 연결을 의미하는 링크 수의 평균 (average degree)은 2.754이다. 네트워크에서 노드 사이의 거리를 나타내는 네트워크 최대 거리(network diameter)는 4이다. 즉 각각의 노드 사이의 연결이 최대 4단계 이내에서 이루어진다. 또한 네트워크의 평균 경로 길이와 네트워크의 밀도는 각각 3.37, 0.008로 나타났다. 요약하면, 문헌정보학의 교과목 네트워크는 응집력이 낮고, 연결 거리가 상대적으로 높은 특징을 갖는다. 문헌정보학을 구성하는 학문이 독립적인 학문 영역이라는 것을 반영하고 있다고 해석할 수 있다. 반면, 네트워크 분석을 통해 문헌정보학의 융합적 특성을 발견할 수 있다. 그림 4에서 보듯이, 추출된 개체명은 특정한 학문 분류에 할당되고 있지만, 분류 사이를 연결해주는 개체명이 발견할 수 있다. 예를 들어, 도서관경영과 정보서비스는 학교도서관

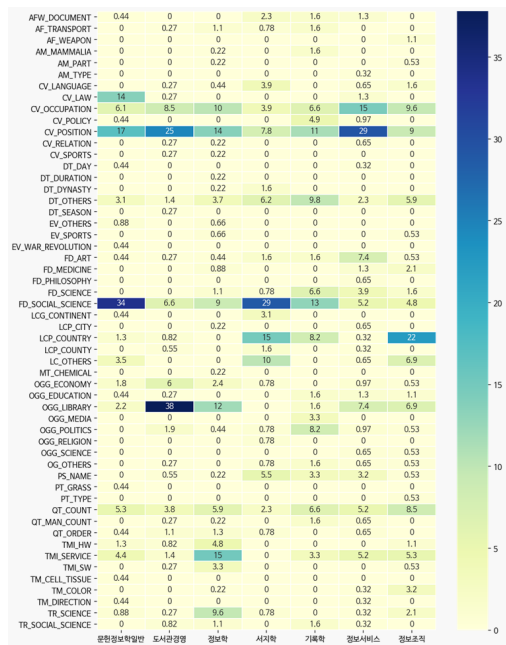


그림 3. 개체명 분석 결과와 히트맵을 이용한 시각화

FD_SOCIAL_SCIENCE(32%), FD_SOCIAL_SCIENCE(16%), CV_POSITION(32%), LCP_COUNTRY(27%, 정의: 국가명)가 가장 높은 비율이다. 상위 비율을 갖고 있는 개체명 인식 태그는 문헌정보학 세부분류의 분야별 주제범위의 분포를 보여 준다.

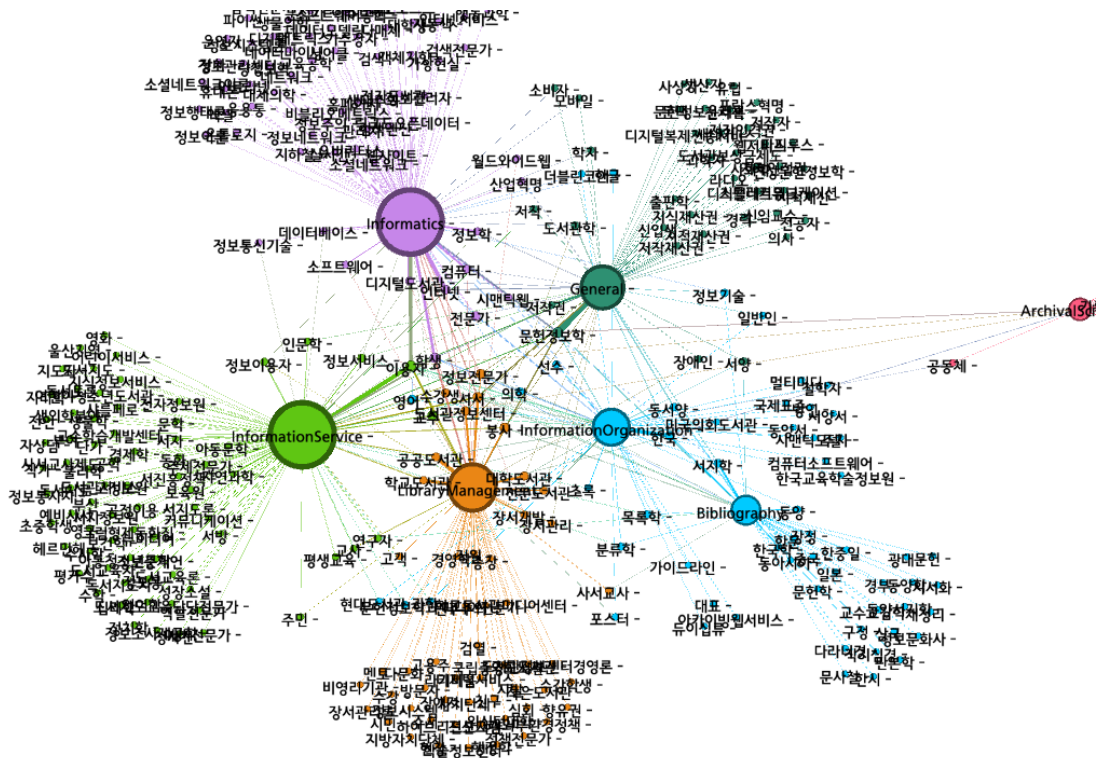


그림 4. 문헌정보학 세부 분류의 키워드 네트워크

관, 공공도서관, 사서, 봉사와 같은 주제를 공유하고 있고, 정보조직과 서지학은 한국, 동서양, 목록학과 관련된 주제로 연결되어 있다. 정보조직의 의학, 선수, 정보전문가와 같은 개체명은 도메인에 특화된 주제를 포함하는 사례로 볼 수 있다. 특히, 정보기술은 대부분의 교과목에 공통적으로 나타나고 있다. 디지털도서관, 컴퓨터, 인터넷, 월드와이드웹은 정보학 관점의 기본이론과 배경을 나타내는 개체명이다. 반면, 시맨틱 웹과 링크드 데이터는 문헌정보를 의미적으로 표현하기 위한 기술로, 이와 같은 개체명이 도서관경영, 정보서비스, 문헌정보학일반 영역에서 공통적으로 출현하고 있다는 점은 흥미로운 사실이다. 사이 중심성이 높은 개체명은 네트워크에서 특정 노드의 짝(pairs)이나 그룹(groups)들 간의 연결에 있어서 중요한 역할을 한다는 것을 의미한다. 사이 중심성의 상위 10개에 포함된 개

체명은 이용자, 교수, 학생, 사서, 봉사, 전문가와 같이 문헌정보 서비스의 실질적인 주체가 포함되어 있고, 대부분의 교과목에서 공통적으로 활용하고 있다.

네트워크 분석은 문헌정보학의 학문적 연계와 융합을 해석하는데 유용할 수 있다. 문헌정보학의 전통적 연구 분야인 도서관경영, 서지학, 정보조직은 분류학, 목록학, 서지학을 포함한 기초 이론을 중심으로 연결되어 있다. 반면, 정보학과 정보서비스는 데이터베이스, 소프트웨어 등 정보기술을 중심으로 문헌정보학의 주제들과 융합되는 경향을 보여준다.

V. 결론

본 연구는 문헌정보학을 구성하는 주요 학문 영역의 관계를 분석하고 있다. 학문 영역은 서로 독립적인 주

제이지만, 동시에 서로 밀접하게 연결되어 있다. 문헌정보학은 도서관경영, 자료조직, 기록학과 같은 전통적인 학문 분야와 정보학, 정보서비스와 같은 융합적 성격의 학문으로 구성되어 있다. 본 연구는 강의계획서에 나타난 교과 내용을 분석하여 학문의 융합적 특징을 검토하고 있다. 연구 결과에서 보듯이, 자료조직, 도서관경영, 서지학은 자료의 분류와 조직과 관련된 주제를 공유하고 있고, 정보학과 정보서비스는 도서관경영과 자료조직과 다수의 주제를 공유하고 있다. 특히, 도서관과 관련된 서비스 제공자와 사용자를 대표하는 주제가 교과 내용의 전반에서 나타나고, ‘소프트웨어’, ‘월드와이드웹’, ‘시맨틱웹’과 같은 기술 용어가 다수의 분류에서 공통적으로 나타나는 것은 의미 있게 해석할 수 있다.

본 연구는 융합의 과정이나 대응을 분석하지 못하고 있다. 본 연구의 데이터세트는 강의계획서를 기본으로 하고 있다. 그러나, 문헌정보학에 개설된 강의는 2년동안 크게 변화하지 않았고, 강의계획서가 학기에 따라 업데이트되지 않는다는 한계가 있다. 교과목의 개편을 포함하려면 데이터의 범위를 넓게 확대하거나, 새로운 연구 방법을 검토할 필요가 있다. 한편, 수집된 데이터를 지속적으로 업데이트하고, 공유하기 위해 데이터를 구조적으로 표현하고 표준화하기 위한 방법을 검토해야 한다. 향후 연구는 개별 학과의 웹사이트에 있는 강의계획서를 수집하기 위한 방안, 문헌정보학 분야의 주요 개체명을 사전 구축, 문헌정보학과 인접된 학문의 관계성을 분석하기 위한 방안을 검토할 필요가 있다.

참 고 문 헌

- [1] 서혜란, “국내 문헌정보학 교육의 미래를 위한 제언,” 한국문헌정보학회 학술발표논문집, pp.9-35, 2016.
- [2] 최형욱, 최예진, 남소연, “문헌정보학 분야의 지적구조 및 연구 동향 변화에 대한 시계열 분석: 2003년부터 2017년까지,” 정보관리학회지, 제35권, 제2호, pp.89-114, 2018.
- [3] Yelim Mo, Euntaek Seon, Goun Park and Haklae Kim, “Course Analysis of Library and Information Science in Korea,” Information, Vol.11, No.1, 2020.
- [4] 강지혜, “문헌정보학과와 데이터 사이언스 커리큘럼 개발 실태와 방향성 고찰,” 한국도서관·정보학회지, 제47권, 제3호, pp.343-363, 2016.
- [5] 유사라, “메타리터러시 관점에서의 문헌정보학 전공 커리큘럼 진단연구,” 한국문헌정보학회지, 제52권, 제2호, pp.191-220, 2018.
- [6] 이재운, “데이터 사이언스와 데이터 리터러시,” 한국정보관리학회 학술대회 논문집, pp.11-15, 2015.
- [7] 노영희, 안인자, 최상기, “한국 문헌정보학 교과과정의 신규교과목 개설추이 분석 연구,” 한국문헌정보학회지, 제46권, 제1호, pp.29-53, 2012.
- [8] 최상기, 안인자, 노영희, 김주섭, “국내외 문헌정보학 주요교과목 강의계획서분석을 통한 수업내용 및 방법 비교 연구,” 한국문헌정보학회지, 제47권, 제2호, pp.223-245, 2012.
- [9] 노영희, 안인자, 최상기, “한국문헌정보학 교과과정 운영모형 및 표준교과목 개발에 관한 연구,” 한국문헌정보학회지, 제46권, 제2호, pp.55-82, 2012.
- [10] 남진숙, “PBL(Project based Learning)을 활용한 융복합 교양교과목의 설계 및 운용,” 한국사교와표현학회 학술대회 논문집, pp.23-33, 2020.
- [11] 주석진, “플립러닝을 활용한 대학생 전공 진로지도 교과목의 효과분석,” 진로교육연구, 제33권, 제1호, pp.1-20, 2020.
- [12] 황인아, 김인경, “‘교양적 소양(Cultivated Literacy)’을 기르기 위한 수학적 추론’ 교과목 개발에 대한 연구,” 교양교육연구, 제14권, 제6호, pp.161-173, 2020.
- [13] 이지용, “창의융합인재 양성을 위한 교양 교과목 개발 사례 연구,” 교양학연구, pp.227-256, 2020.
- [14] 최미정, 진상우, “간호대학생을 위한 창의성기반 창의융합교과목 개발,” 한국엔터테인먼트산업학회논문지, 제14권, 제3호, pp.83-91, 2020.
- [15] 노영희, 곽우정, “국내 문헌정보학과 대학원 교과과정에 대한 연구,” 한국문헌정보학회지, 제54권, 제4호, pp.123-146, 2020.
- [16] 이혜원, 한승희, “데이터과학 교육과정에 대한 분석적 연구,” 한국문헌정보학회지, 제54권, 제1호, pp.365-385, 2020.
- [17] 오상희, 김수정, 이종욱, “미국 문헌정보학 분야 대학원의 메이커스페이스 교과목 내용 분석,” 사회과학연구, 제31권, 제1호, pp.263-278, 2020.
- [18] 최상희, “교과 키워드 분석을 통한 문헌정보학과 교

육 주제 연구 - 한국·미국 정보기술관련 교과 중심으로 -, "한국도서관·정보학회지, 제50권, 제2호, pp.43-60, 2019

- [19] 송은영, 최희련, 이흥철, "Word Embedding에 PCA를 적용한 개체명 인식 모델을 위한 효율적인 학습방법 연구," 대한산업공학회지, 제45권, 제1호, pp.30-39, 2019.
- [20] 김홍진, 박성식, 김학수, "공유계층을 이용한 형태소 분석과 개체명 인식 통합 모델," 정보과학회논문지, 제48권, 제2호, pp.167-173, 2021.
- [21] 민진우, 나승훈, 신종훈, 김영길, "한국어 개체명 인식 & 링킹 & 의존파싱 통합모델," 한국정보과학회 학술발표논문집, pp.392-394, 2019.
- [22] 정예원, 이종혁, "Bidirectional LSTM-CRF 기반 한국어 개체명 인식을 위한 유사 접사 자질을 이용한 단어 표상 확장," 정보과학회 컴퓨팅의 실제 논문지, 제26권, 제9호, pp.408-413, 2020.
- [23] 이영훈, 나승훈, 최윤수, 이해우, 장두성, "ALBERT를 이용한 한국어 자연어처리: 감성분석, 개체명 인식, 기계독해," 한국정보과학회 학술발표논문집, pp.332-334, 2020.
- [24] 광기현, 윤정섭, "지능정보기술 분야에서의 글로벌 기술 지식 경쟁력 분석 : 한국을 중심으로," 한국콘텐츠학회논문지, 제21권, 제1호, pp.24-38, 2021.
- [25] 정세원, "빅데이터와 유렵연구: 사회 및 의미연결망 분석 사례를 통한 유렵연구의 가능성 고찰," 통합유렵연구, 제11권, 제3호, pp.355-380, 2020.
- [26] 김문옥, "빅데이터를 활용한 방송뉴스 제목의 다학제적 의미연결망 연구 - KTV 대한뉴스를 중심으로 -, " 한국과학예술융합학회, 제39권, 제1호, pp.17-31, 2021.
- [27] 정유경, "국내 문헌정보학 분야 학술지의 인용 네트워크분석," 한국문헌정보학회지, 제54권, 제4호, pp.221-238, 2020.
- [28] 이하나, 안순태, "우울감 표현에 대한 지각된 혜택과 장애 요인: 의미 연결망 분석을 활용한 탐색적 연구," 성인간호학회지, 제32권, 제6호, pp.571-583, 2020.
- [29] 이장석, 신명호, "'여성'과 '의료'간 의미연결망 분석: 학술연구를 중심으로," 문화산업연구, 제20권, 제4호, pp.27-33, 2020.
- [30] 강완모, 박찬열, "그래프 이론을 적용한 서울시 녹지 연결망의 정량적 분석," 한국환경생태학회, 제25

권, 제3호, pp.412-420, 2011.

저 자 소 개

윤 정 민 (Jungmin Yun)

정회원



■ 2018년 2월~현재 : 중앙대학교 문헌정보학과 재학
<관심분야> : 텍스트 마이닝, 자연어 처리, 데이터 사이언스

김 나 무 (Namoo Kim)

정회원



■ 2018년 2월~현재 : 중앙대학교 문헌정보학과 재학
<관심분야> : 도서관, 데이터 시각화, 데이터 사이언스

김 학 래 (Haklae Kim)

정회원



■ 2010년 6월 : 아일랜드 국립대학교 (공학박사)
■ 2019년 3월~현재 : 중앙대학교 문헌정보학과 교수
<관심분야> : 지식공학, 인공지능, 데이터 사이언스