OPEN ACCESS

공개 학술 메타데이터 플랫폼을 활용한 기관 간 협업 네트워크 분석: 물 분야 사례*

Inter-Institutional Collaboration Network Analysis Using Open Scholarly Metadata Platform: A Case Study of the Water Sector

이승한 (Seunghan Lee)**

김성훈 (Sung Hoon Kim)***

초 록

본 연구는 서지정보를 기반으로 하여 물 분야에서 연구기관이나 주요 기업 간의 협력 연구 구조를 정량적으로 분석하는 것을 목적으로 한다. 이를 위해서 OpenAlex, Crossref, ROR(Research Organization Registry) 등에서 제공 가능한 메타데이터를 연계하여, 분석에 활용 가능한 논문 데이터를 수집하고 보강하는 프로토콜을 구축하였다. 해당 기관에서 제공하는 API를 통해 논문 메타데이터를 수집하였으며, 논문 저자의 소속기관 정보를 중심으로 정리하였다. 일부 메타데이터의 결측 값에 대해서는 Crossref와 ROR을 이용하여 보완하고자 하였다. 본 연구에서 대상으로 한 저널은 Water Research 이며, 25년간(2000~2024년)의 데이터 20,135건에 대해서, 5년 단위로 5개의 시기로 구분하여 시기별로 논문 발표의 강도가 높은 주요 기관과 공저 네트워크 변화를 검토하기 위해서 중심성 지표 3가지를 선정하여 비교 분석하였다. 그 결과, 본 사례연구에서는 고유벡터 지표를 기반으로 하는 선택의 경우 합리성이 있는 것으로 판단하였다. 서지 데이터의 저자의 소속기관 정보를 통해서 특정 분야의 연구에 대해 연구기관 간 협업 구조를 체계적으로 검토하는 프로토콜을 사례로 제시함으로써, 다양한 연구 분야에 대한 협업 허브를 기반으로 하는 주제 발굴 등을 위한 분석 기초자료로서의 활용 가능성을 확인하였다.

ABSTRACT

This study aims to quantitatively analyze collaborative research structures among research institutions and major corporations in the water sector based on bibliographic information. To achieve this, we established a protocol to collect and augment usable paper data by linking metadata available from OpenAlex, Crossref, and ROR (Research Organization Registry). Paper metadata was collected via APIs provided by these institutions and organized primarily around the affiliated institutions of paper authors. Missing DOI and RORID values were supplemented using Crossref and ROR. The journal targeted in this study is Water Research. For 20,135 data points spanning 25 years (2000-2024), divided into five periods in five-year increments, three centrality metrics were selected and compared to examine the major institutions with high publication intensity and changes in co-author networks per period. The results of this case study indicate that the selection based on the eigenvector metric is reasonable. By presenting a protocol that systematically examines the collaborative structure among research institutions for specific fields using authors' institutional affiliation information from bibliographic data, this study suggested the potential for utilizing foundational data to identify research topics based on collaboration hubs across diverse fields.

키워드: 계량서지학적 네트워크 분석, 공개 학술 메타데이터, 기관 간 협업, 물 분야, 중심성 분석
bibliometric network analysis, open scholarly metadata, inter-institutional collaboration, water sector, centrality analysis

---

* 본 논문은 한국수자원공사 R&D과제 '자연어처리 기반 물산업 분석 고도화를 위한 표준 프레임워크 개발(과제번호: G240148)'의 지원을 받아 작성된 것임.

** 한국수자원공사 K-water연구원 AI연구소 연구원(s0124kw@gmail.com) (제1저자)

*** 한국수자원공사 K-water연구원 AI연구소 소장(sunghoonkim@kwater.or.kr) (교신저자)

▪ 논문접수일자 : 2025년 11월 21일 ▪ 최초심사일자 : 2025년 12월 11일 ▪ 게재확정일자 : 2025년 12월 15일

▪ 정보관리학회지, 42(4), 201-220, 2025. http://dx.doi.org/10.3743/KOSIM.2025.42.4.201

※ Copyright © 2025 Korean Society for Information Management
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (https://creativecommons.org/licenses/by-nc-nd/4.0/) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

1. 서 론

증가하는 학술 논문의 수와 함께 국제적인 공동 연구나 기관 간의 연구 확산 등에 따라 국가 및 기관 차원의 연구 협업 구조를 정량적으로 분석하려는 시도가 이어져 왔다(Newman, 2001; Wagner & Leydesdorff, 2005). 이와 관련하여, 협업 네트워크 분석(Collaboration Network Analysis)은 학술 생태계에서 연구자 간 협력 구조를 해석하고 지식 확산의 흐름과 연구 허브의 형성, 그리고 네트워크의 구조적 특징 등을 이해하는 데 널리 활용되고 있으며(Donthu et al., 2021), 이러한 분석 결과를 정량적으로 표현하게 되면, 연구자나 기관의 영향력과 협업 역량을 설명할 수 있는 계량적 도구로 활용할 여지가 있다.

기존 연구들은 주로 저자 단위의 공저 네트워크를 구축하여 연구자들 간의 협력 구조와 네트워크의 전반적인 특성을 분석하는 데 초점을 두었다(Newman, 2001). 그러나, 기관 단위(Institution-level)의 협업 구조는 국가 간 연구정책, 공공과 민간 연구 협업, 연구자원의 분포 등 거시적 차원에서 더욱 중요한 의미를 갖는다(백영기, 2012; Wagner & Leydesdorff, 2005). 그럼에도 불구하고, 기관명 표기의 비표준화, 데이터베이스 간 불일치, 소속 정보 누락 등의 문제로 인해서 기관 단위 분석에서는 신뢰성 높은 데이터 구축이 어려워 제한적으로만 수행되어 왔다(이석형, 2014). 근자에는 이러한 메타데이터의 한계를 보완하기 위한 수단으로 연구자, 기관, 논문 등 학술 객체를 고유하게 식별하는 식별자인 영구식별자(PID: Persistent Identifier) 개념이 등장했다. 대표적으로 DOI (Digital Object Identifier)는 디지털 객체(논문, 데이터셋 등)를, ORCID(Open Researcher and Contributor ID)는 연구자를, 그리고 RORID (Research Organization Registry Identifier)는 기관을 각각 식별하는데 적용될 수 있다. 이 중에서도 2019년에 공개된 ROR은 전 세계 연구기관에 고유한 식별자를 부여하여 활용할 수 있도록 하고 있다. 최근에는 기관명을 RORID에 매핑하여 정규화 할 수 있도록 하는 AffilGood (Duran-Silva et al., 2024), 저자 소속 기관명과 RORID의 매핑 시스템 AffRO(Kallipoliti et al., 2025)가 개발되어, 비표준화되어 있던 기관명의 RORID 매핑 정확도가 향상되고 있다.

또한, 잘 알려진 Crossref는 DOI 등록기관으로서 논문 제목, 저자, 초록, 기관명, 출판연도 등 학술 메타데이터를 표준화된 형태로 제공하고 있다. OpenAlex(Priem et al., 2022)는 MAG (Microsoft Academic Graph)를 계승한 형태로서, 논문, 저자, 기관, 주제 정보를 통합한 개방형 네트워크를 구축하였다. 이와 관련하여, 최근에는 OpenAlex에서 발생하는 기관 식별자 (RORID)의 누락 문제를 분석하고, OpenAlex의 기관 객체와 내부데이터를 활용하여 누락된 기관 정보를 보완하는 절차가 제시되었다(Zhang et al., 2024). 이처럼 협업 네트워크 분석의 이론적 기초 및 기관명 표준화 문제의 해결, 오픈사이언스 분야의 인프라 개선 등을 통해 기관 간 연구 협업 구조의 분석 가능성이 개선되고 있으나, 이를 통합적으로 구현한 사례는 제한적이다. 따라서, 본 연구는 오픈 사이언스 환경에서 Crossref, OpenAlex, ROR을 기반으로 데이터를 수집하여 분석하는 방법론과 절차를 제시하는 것을 목적으로 하며, 이를 기반으로 특

정 분야의 사례에 대한 논문 메타데이터를 대상으로 적용함으로써 적용 타당성과 활용 가능성을 검토하고자 하였다.

본 연구가 검토 대상으로 선정한 물 분야는 기후변화로 인한 물 부족 심화, 홍수, 가뭄 등 물 재해 위험의 증가, 그리고 데이터센터나 반도체 산업 등으로 대표되는 물 수요 구조의 변화로 인해 그 중요성이 지속적으로 확대되고 있다. 물 문제는 기술적 요소와 정책적 요소가 복합적으로 얽혀 있는 경우가 많아, 학제 간이나 기관 간 협력 연구가 중요하다. 또한 물 분야는 비교적 오랜 기간 동안 연구가 축적되어 온 영역으로, 다양한 학술 데이터가 보유되고 있다는 특징을 지닌다. 이러한 특성으로 인해 본 분야는 협력 연구 네트워크의 구조적 특성과 그 시간적 변화를 분석하고 검토하기에 적합하다고 판단되어 연구 대상으로 선정하였다.

2. 문헌 연구

2.1 협업 네트워크 분석을 위한 서지정보 활용 연구

협업 네트워크 연구는 사회연결망이론(Social Network Theory)과 계량서지학(Bibliometrics)의 결합을 통해 발전해 왔다. Newman(2001)은 과학자 간 공저 네트워크(Coauthorship network)가 짧은 경로 길이와 높은 군집 계수를 가지는 구조를 가지고 있어, 핵심적인 연구자가 전체 네트워크의 연결성을 주도한다고 하였으며, Wagner와 Leydesdorff(2005)는 국제적 공동연구(International collaboration)가 네트워크의 연결성과 중심성을 확장시킨다고 보았다. Das et al.(2018)은 다양한 중심성 지표(Centrality measures) 중에서 연결정도(Degree), 매개(Betweenness), 고유벡터(Eigenvector) 등을 중심으로 각각의 지표가 협업 네트워크 내 영향력 평가에서 어떤 의미를 갖는지 분석하였고, Donthu et al.(2021)은 연구자, 기관, 그리고 국가 수준에서의 네트워크 분석의 비교를 위한 서지정보 기반의 가이드를 제안하였다.

국내에서도 서지 정보의 활용 및 협업 네트워크의 구조적 특성을 파악하려는 시도가 지속되어 왔다. 이재윤 외(2011)는 국내외 학술 데이터베이스를 기반으로 계량서지 분석과 동시출현단어(Co-word) 분석을 수행하여 LED 제조 기술 분야의 주요 연구 주체(연구자, 기관, 국가 등)와 연구 주제 구조를 파악하고, 이를 해당 연구 분야에서의 네트워크 분석을 위한 기초자료로 구축하고자 하였다. 백영기(2012)는 KCI 논문 기반의 공저 데이터를 활용한 계량서지학 및 공저 네트워크 분석을 통해서 대학과 기업 간 협력 네트워크 구조와 혁신 창출 수준의 관계를 실증적으로 규명하는 연구를 수행하였다.

2.2 공개 학술 메타데이터 플랫폼의 활용과 기관 식별 표준화를 위한 연구

협업 네트워크 분석의 효용감과 성능을 높이기 위해서는 기관 간 협업 네트워크 분석의 보강이 필요하며, 이를 위해서는 기관 식별을 위한 정보를 통합하여 활용할 수 있는 방법을 적용하는 것이 필요하다. 이석형(2014)은 국내 학술논문 데이터베이스에서 기관명 정규화 과정

의 한계를 실증적으로 분석하여, 표준화된 식별체계의 부재가 기관 단위 연구의 확장성을 제약한다고 지적하였으며, 김진영 외(2016)는 기관명 자동 식별의 정확도가 전체 네트워크 구조의 정확도에 직접적인 영향을 미친다는 점을 보여주었다.

최근 오픈사이언스(Open Science)의 확산은 학술 메타데이터의 개방과 상호운용성 향상에 중요한 전환점을 마련하고 있다. 또한, 오픈사이언스는 연구의 투명성, 재현성, 그리고 공공 접근성을 보장하기 위한 정책 패러다임으로 자리잡고 있으며, 이는 데이터, 코드, 연구 성과의 공개를 통해 지식생태계의 공공성을 강화한다(United Nations Educational, Scientific and Cultural Organization, 2021). 대표적으로 Crossref는 전 세계 학술출판물의 DOI 등록과 메타데이터 공유를 담당하는 대표적 비영리 기관으로서, 학술논문의 제목(Title), 초록(Abstract), 저자(Author), 기관명(Affiliation), 출판연도(Year) 등 표준화된 메타데이터를 제공한다(Waltman et al., 2020). OpenAlex는 학술 저작물, 저자, 기관, 학술지, 주제 등 학술 객체를 이종 방향성 그래프(heterogeneous directed graph) 구조로 연결하여 통합적으로 제공하는데, 모든 객체에는 고유의 OpenAlex ID가 부여되며, DOI, ORCID, RORID 등의 PID를 활용함으로써 학술 메타데이터의 통합성과 연결성을 강화하고자 하였다(Priem et al., 2022).

특히 최근에는 계량서지정보의 표준화를 위한 도구들이 오픈소스 생태계에서 확장되어 연구되고 있다. Duran-Silva et al.(2024)는 기관명 명확화(disambiguation)를 위해 오픈소스 기반의 AffilGood 프레임워크를 개발하여 ROR과 호환이 가능하도록 하였고, Kallipoliti et al.(2025)은 AffRO(Affiliation strings to Research Organization) 시스템을 개발하여 원시 기관명 문자열(raw affiliation strings)을 RORID로 자동 매핑하는 알고리즘을 제시하였다. Orduña-Malea와 Bautista-Puig(2022)는 ROR이 기관명, 국가, 유형, 상위기관 관계를 포함하는 계층적 구조를 채택하고 있어, 기관 간 네트워크 분석의 연결성을 높였다고 주장한다.

국내의 경우, 이석형(2014)은 학술논문 메타데이터로부터 기관명 식별 데이터를 구축하는 과정에서의 문제점을 지적하였고, 김진영 외(2017)는 국내의 과학기술콘텐츠 데이터를 기반으로 기관명을 자동 식별하는 알고리즘인 형태소 기반 문자열 유사도 비교 방식을 제안한 바 있다. 이의 활용과 관련해서는, 이해성과 이혜진(2016)은 연구 유사도 기반의 네트워크를 구축하고 메타데이터를 통한 연구자 네트워크의 시각화 연구를 수행하였고, 임윤정 외(2020)는 학술논문 메타데이터를 분석하여 다양한 분야에서 연구자 간 협업 관계를 시각화하는 시스템을 구현한 바 있다.

2.3 영구식별자(PID)와 기관 간 네트워크 연결 방법론

연구자, 논문, 기관을 식별하는 PID 시스템은 학술 메타데이터의 연결성, 추적성, 재현성을 확보하고자 제안된 것이다. 예컨대 ORCID가 연구자 식별을 위한 방안으로 제시되면서, 연구자 정보의 중복 등록 및 동명이인 문제가 상당 부분 해결된 바 있다(Haak et al., 2012). 이와 같이, ORCID와 DOI를 중심으로 한 PID

연계 방식과 더불어 ROR이 공개되면서 영구 식별자를 활용한 네트워크 분석의 가능성이 확장되었다. 다음 <표 1>에서 PID의 유형별 구분과 식별 대상, 정의, 형식, 역할 및 기능을 정리하였다. Cousijn et al.(2018)은 DOI와 같은 영구 식별자(PID) 채택이 연구 데이터의 투명성, 추적 가능성, 재사용성을 높이는 핵심 요소임을 강조하였으며, Orduña-Malea와 Bautista-Puig (2022)은 RORID를 활용해 연구기관 웹사이트 간 하이퍼링크 연결성을 정량화하는 방법론을 제시하는 등 기관 식별자 기반 네트워크 분석의 유효성을 강조하였다.

또한, 기관명 표준화뿐만 아니라 메타데이터 보강 과정에서도 다양한 기술 발전이 이루어졌고, Duran-Silva et al.(2024)의 AffilGood과 Kallipoliti et al.(2025)의 AffRO는 비표준화된 저자 소속정보(affiliation string)를 정규화하고, 이를 ROR ID와 매핑하여 기관 식별의 정확도를 높인 사례이며, 특히 NLP(Natural Language Processing) 기반의 기관명 자동 식별 및 추론 기술 등이 기관명 인식(entity recognition)과 소속기관 추론(affiliation inference)에 활용되기 시작하였다(Lin et al., 2024). 한편, Zhang et al.(2024)은 OpenAlex 데이터에서 기관 정보(RORID)가 누락되는 현상을 분석하여 기관 간 협업 네트워크의 분석에 영향을 줄 수 있음을 보였다. Velez-Estevez et al.(2023)은 OpenAlex, Crossref 외에도 논문, 인용 현황 등을 제공하는 Dimensions 등 주요 계량서지 API를 비교 분석하였고, 각 데이터베이스의 대상 범위와 업데이트 주기 등이 상이하므로 상호보완적 활용이 필요함을 제안하였다. 이와 관련하여, Butt et al.(2021)은 Crossref와 DOI 간 인용 연결 정보를 제공하는 오픈 인용 인덱스인 OpenCitations (COCI)을 활용하여, 공개 학술 메타데이터 기반의 수집, 정제, 네트워크 분석 워크플로우를 제안하였다.

3. 연구방법

3.1 개요

본 연구에서는 공개 학술 메타데이터 플랫폼인 Crossref, OpenAlex, ROR을 연계 및 통합하여 메타데이터 중 기관 정보의 결측을 중심

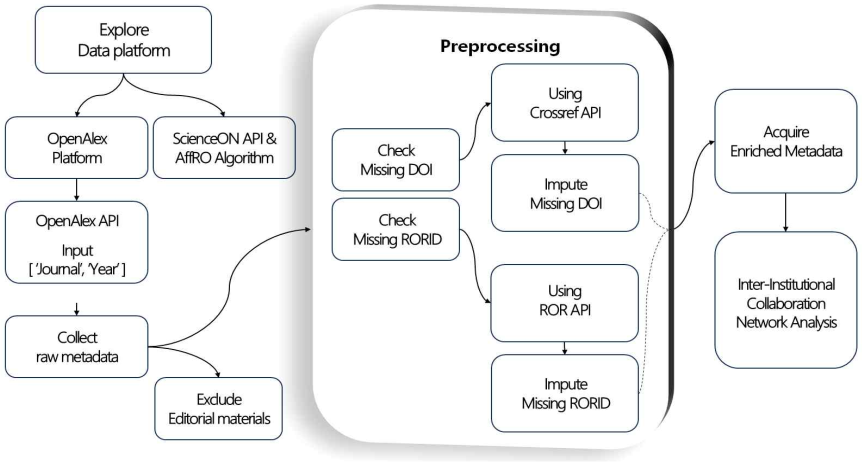
<표 1> 주요 영구 식별자(PID: Persistent Identifier) 개요

구분	PID(Persistent Identifier)		
	DOI	ORCID	RORID
식별대상	디지털 객체(논문, 데이터셋, 보고서, 책 등)	연구자(개인)	연구기관
정의	디지털 환경에서 영구적으로 식별되고 접근될 수 있는 객체(digital object)를 위한 식별자	학술 연구자를 고유하고 영구적으로 식별하는 16자리의 식별자(digit)	전 세계 연구기관에 대한 오픈 커뮤니티 기반의 글로벌 기관 식별자
형식	10.1000/182	0000-0002-1825-0097	https://ror.org/03rym5c26
역할 및 기능	학술 출판물의 영구 참조 등	저자 동명이인 문제 해소 등	기관명 표준화, 소속 정보 연계, 협업 네트워크 분석 지원

으로 보강하고, 기관 간 연구 협업 네트워크를 정량 분석하는 방법을 제안한다. 이를 통해 확보된 데이터를 기반으로 재현성이 높은 연구 협업 네트워크 분석 방법론을 제시하는 것을 목적으로 한다. 기관 정보 결측은 기관 식별자(RORID)가 단 한 건도 존재하지 않는 레코드를 기준으로 정의하였으며, 해당 레코드에 최소 1건 이상 기관 식별자가 보완된 경우를 보강으로 간주하였다. 사례연구로서, 최근 25년간(2000년~2024년) 물 분야에서 종합적이며 영향력이 높은 학술지인 Water Research(IF = 12.4, 2024년 기준)에 게재된 모든 논문을 대상으로 하였다. 구체적으로는, 기관 간 협업 네트워크를 구축하고 5년 단위로 시계열 비교를 수행하기 위해, 데이터 수집-정리-보강-네트워크 구축의 절차를 거쳤으며, 네트워크 분석의 효율성 확보를 위한 중심성 지표 기반의 방법론 검토를 포함하는 절차를 제시하고자 하였다. 그 순서는 다음 〈그림 1〉과 같이 나타낼 수 있는데, ① 연도 범위와 ISSN(복수 가능)을 입력 자료로 하여, ② OpenAlex를 통해 해당 저널의 논문 메타데이터를 1차적으로 확보한다. ③ 수집된 메타데이터의 구조를 정리하고 결측값을 보강한 후, ④ 표준화된 기관명을 바탕으로 ⑤ 동시 출현 네트워크를 구성하고, ⑥ 중심성 지표를 기반으로 평가하여 ⑦ 네트워크 시각화(툴 활용) 및 결과를 정리하도록 한다.

3.2 네트워크 분석과 중심성 지표

네트워크 분석은 노드(node)와 엣지(edge)로 이루어진 그래프 구조를 바탕으로, 분석하고자 하는 객체들 사이의 관계나 상호작용을 정량적으로 분석하는 방법이다. 학술 협업 네트워크에서는 연구자, 기관, 국가 등을 노드로, 공저(co-authorship)와 같은 공동 참여 관계를 엣지로 간주하여 협업 구조를 분석한다. 본 연구에서는 기관을 노드로 정의하고, 동일 논문에 공동으로 참여한 기관 쌍을 엣지로 구성하여 기관 간 협업 네트워크를 구축하였다. 한 기



〈그림 1〉 기관 간 협업 네트워크 분석을 위한 절차 개요도

관 쌍이 여러 편의 논문에서 반복적으로 공동 저자로 등장할 경우 해당 빈도를 엣지 가중치에 반영함으로써 두 기관 사이의 협업 강도를 표현하였다. 네트워크에서 특정 노드의 중요성을 정량적으로 평가하기 위해 다양한 중심성(centrality) 지표가 사용되는데, 이러한 중심성 지표는 네트워크 내에서 허브 역할을 하는 노드, 중개적 역할을 하는 노드, 영향력 있는 노드 등을 식별하는 데 활용된다(Das et al., 2018).

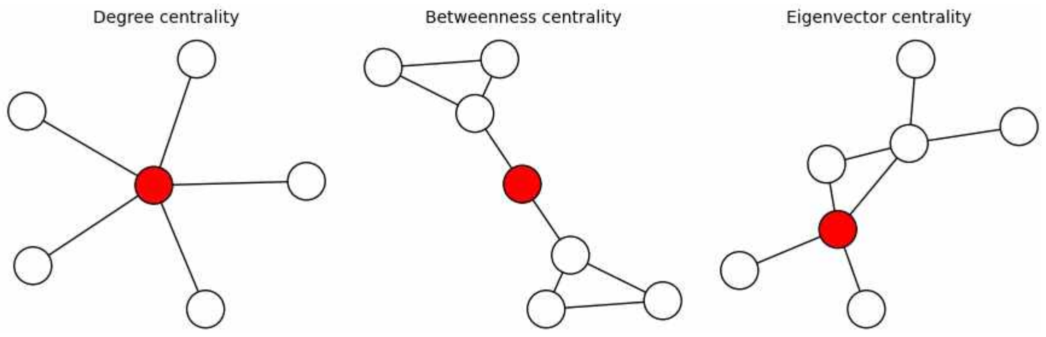
본 연구에서는 기관 간 협업 네트워크에서의 허브 기관과 협업 구조의 변화를 파악하기 위해 연결정도 중심성(Degree Centrality), 매개 중심성(Betweenness Centrality), 고유벡터 중심성(Eigenvector Centrality) 세 가지 지표를 선택하였고, 각 중심성 지표의 개념은 〈그림 2〉에 제시하였다. 먼저, 연결정도 중심성은 한 노드에 직접 연결된 이웃 노드의 수를 의미하며, 특정 기관이 얼마나 많은 다른 기관과 직접 협업 관계를 형성하고 있는지를 나타낸다. 이는 계산이 단순하고 직관적이라는 장점이 있으나 간접 연결이나 네트워크 전체 구조는 반영하지 못한다는 단점이 있다.

매개 중심성은 한 노드가 다른 노드 쌍 사이의 최단 경로 위에 얼마나 자주 위치하는지를 측정하는 지표로, 매개 중심성이 높은 기관은 여러 협업 경로의 중간다리 역할을 수행하며, 이는 네트워크 내에서 중개자 역할을 하는 기관을 식별하는 데 유용하다. 또한, 고유벡터 중심성은 단순한 연결 수뿐 아니라 영향력 있는 이웃 노드와의 연결을 함께 고려하는 지표로, 영향력 있는 노드와 얼마나 잘 연결되어 있는가를 반영한다. 이는 연결정도 중심성처럼 단순히 얼마나 많은 기관과 협업하는지를 넘어, 영향력 있는 기관들 사이의 핵심적인 위치를 파악하는 데 적합한 지표로 평가된다.

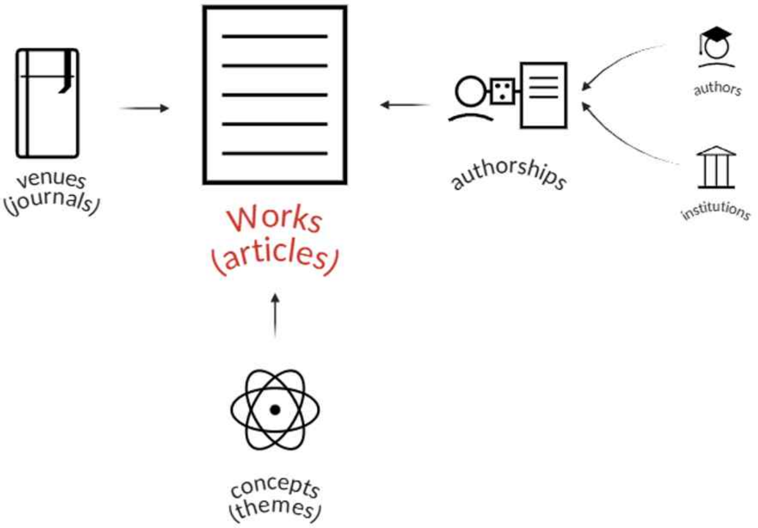
본 연구에서는 물 분야 상위 저널의 기관 간 협업 네트워크를 대상으로 세 가지 중심성 지표를 시기별로 적용하여, 각 시기의 핵심 기관과 협업 구조의 특징을 파악하고자 하였다.

3.3 메타데이터 수집 및 보강

본 연구에서는 OpenAlex가 제시하는 데이터 모델(Priem et al., 2022)을 활용하여 논문 메타데이터를 수집하였는데, 〈그림 3〉은 본 연구에서 사용한 OpenAlex 데이터 모델 구조를 데이



〈그림 2〉 중심성 지표(Degree, Betweenness, Eigenvector)에 따른 연결 개념도



〈그림 3〉 OpenAlex의 그래프 데이터 모델(Priem, Piwowar, & Orr, 2022)

터베이스 연결 맥락에 맞게 재구성한 것으로, 저자표기(authorships) 필드 내 소속(affiliations) 연결을 활용하여, 이후 기관 간 협업 네트워크 분석에 활용하였다.

3.3.1 논문 메타데이터 수집

이 과정에서는, ISSN과 연도 범위를 입력으로 받아 각 논문에 대해 OpenAlex가 제공하는 필드(openalex_id, doi, title, publication_date, journal_name, authorships)를 수집한다. 여기서 openalex_id는 각 논문 레코드에 대해 OpenAlex가 자체적으로 부여한 문헌 식별자이다. 본 연구의 네트워크 구축의 근간이 되는 Authorhips 필드에는 논문에 참여한 모든 저자에 대한 정보(저자의 이름, 순서, 국적, 종류(저자 순위), 소속기관 정보 등)가 리스트 형태로 제공되는데, 추후 결측 데이터에 대한 보강 수행에 활용할 수 있다. 또한, 실제 연구 논문이 아닌 문헌(Editorial material 등)은 수집 논문에서 제외하도록 하였다.

3.3.2 논문 메타데이터 보강

기관명은 국가, 언어, 출판사마다 상이하게 표기되기 때문에, 동일 기관이 서로 다른 문자열로 기록되는 경우가 빈번하게 발생한다(김진영 외, 2016; 이석형, 2014). 예를 들어 「Seoul National University」와 「Seoul Nat'l Univ.」는 사실상 동일 기관이지만, 표기상의 불일치로 인해 서로 다른 노드로 처리되는 경우가 있다. 따라서, 메타데이터 수집 직후에는 표기 차이로 인해 생기는 중복 데이터를 방지하여 분석 결과의 신뢰성을 제고하기 위해 논문의 제목과 저자명에서 불필요한 공백 및 구두점을 지우고, 문자는 소문자로 통일하며 날짜 형식을 표준화하여(YYYY-MM-DD)로 통일하는 작업을 수행하였다. DOI 보강은 OpenAlex를 통해 수집한 논문 메타데이터에서 DOI가 비어있는 레코드를 대상으로 이루어지는데, 논문의

제목과 제1, 제2저자 이름 및 순서가 일치하는 경우와, 최초 수집된 데이터셋에서 DOI 공통 접두사를 추출하여 이와 일치하는 경우에만 DOI 보강을 수행하도록 하였으며, DOI 결측 보강 시에는 Title, authorships 등, 원본 구조 및 내용은 그대로 두고 DOI 필드만 갱신하도록 구성하였다.

다음으로 기관 식별자인 RORID 보강은 ROR API를 활용해 수행하는데, 먼저 authorships 필드의 "institutions" 리스트 내에 OpenAlex 기관 ID(OpenAlex-ID)가 있는 경우에는 해당 값을 해당 플랫폼에서 찾아 매핑할 수 있도록 하고, 기관 ID가 없는 경우이거나 ROR 항목이 없는 경우에는 "affiliations" 리스트의 문자열을 활용하여 기관을 식별할 수 있도록 하였다. 해당 형식에 대한 사항은 <표 2>에 예시를 두어 참고하도록 하였다.

<표 2>와 같이 Authorships 필드 내 "affiliations" 리스트의 문자열이 단순한 경우에는 이미 OpenAlex에서 RORID를 보유하고 있어, 최초 수집 시 RORID를 보강할 필요가 없지만, <표 3>과 같이, 해당 문자열이 복잡하고 콤마(,)로

<표 2> 수집된 논문 메타데이터 셋 예시(1)

구분	Columns(컬럼)			
Field(필드)	DOI	ID	Display_name	Publication_date
Component (컴포넌트)	10.1016/j.watres.2016.05.042	W2404431829	Comparison of microbial ...	2016-05-28
Field(필드)	Authorships	Journal	ROR	Org_names
Component (컴포넌트)	[{... "institutions": [{"id": "https://openalex.org/I139264467", "display_name": "Seoul National University", "ror": "https://ror.org/04h9pn542" ...}], "affiliations": [{"raw__affiliation__string": "Seoul National University", ...}, ...], ...}]	Water Research	https://ror.org/04h9pn542	Seoul National University

<표 3> 수집된 논문 메타데이터 셋 예시(2)

구분	Columns(컬럼)			
Field(필드)	DOI	ID	Display_name	Publication_date
Component (컴포넌트)	10.1016/j.watres.2015.09.032	W2113606704	Lithium sorption ...	2015-09-21
Field(필드)	Authorships	Journal	ROR	Org_names
Component (컴포넌트)	[{... "institutions": [{"id": "https://openalex.org/I154130895", "display_name": "University of Auckland", "ror": "https://ror.org/03b94tp07", ...}], "affiliations": [{"raw_affiliation_string": "(Department of Civil and Environmental Engineering, University of Auckland, New Zealand)", ...}, ...], ...}]	Water Research	https://ror.org/03b94tp07	University of Auckland

길게 작성되어 있는 경우에는 RORID 보강이 정확하게 이루어지지 않기 때문에, 이를 분리하여 정확한 기관 매핑이 될 수 있도록 ROR API 호출 로직을 구성하는 작업이 필요하다. 이에 본 연구에서는 긴 문자열을 콤마 단위로 분리하여, ‘university’, ‘centre’, ‘institute’ 등과 같은 표준 연구 기관명을 포함하는 문자열 등을 ROR API 호출 쿼리로 사용하였으며, 국적을 나타내는 마지막 문자열을 조건 쿼리로 활용하여 매핑 시 정확도를 향상시키고자 하였다.

3.4 기관 간 협업 네트워크 구축 및 검토 방법

본 연구에서는 기관 간 협업 네트워크를 구축하기 위해 각 기관을 노드로, 동일 논문에 공저자로 참여한 기관 간의 연결을 엣지로 정의하였다. 엣지의 가중치는 기관 쌍의 동시 출현 빈도를 누적하여 산정하였으며, 단일 논문 내에 포함된 기관들의 가능한 모든 조합을 엣지로 형성하되, 동일 논문 내에서 기관명이 중복 표기된 경우는 1회로 간주하여 데이터의 중복을 배제할 수 있도록 하였다. 또한, 본 연구에서는 시기별 분석의 일관성과 정밀도를 확보하기 위해 기관 간 연결 빈도가 높은 상위 400개 기관을 기준으로 네트워크를 구축하여 분석하였다. 전체 네트워크에서 기관 협업의 영향력이 낮은 주변부 및 연결성이 미미한 노드를 제어함으로써, 실질적인 상위의 협업이 활발히 일어나는 핵심 군집을 중심으로 하는 네트워크의 위상학적 특성을 명확히 검토할 수 있다. 구축된 네트워크의 노드와 엣지 현황을 기초로 주요 허브 기관과 협업 강도가 높은 기관 쌍을 정량적으로 분석할 수 있도록 정리하였으며, 핵심 기관 및 협업 구조 도출을 위한 분석 도구로는 앞서 3.2절에서 논의한 세 가지 중심성 지표를 적용하였다.

4. 분석 결과 및 토의

본 장은 2000~2024년 Water Research 저널의 논문을 대상으로 공개 학술 메타데이터 플랫폼의 적용방식을 도입함에 따라 논의된 점들을 검토하고, 이후 개발된 프로토콜에 따라서 분석 절차를 적용하여 얻은 결과를 제시하고 논의한다. 전체 분석 기간은 5년 단위의 다섯 구간으로 나누어졌으며(1기: 2000년~2004년, 2기: 2005년~2009년, 3기: 2010년~2014년, 4기: 2015년~2019년, 5기: 2020년~2024년), 각 구간마다 논문 메타데이터 수집과 보강 및 기관 간 협업 네트워크 구축을 같은 규칙으로 수행한 결과를 기반으로 한다.

4.1 메타데이터 수집 및 보강 결과

구간별 메타데이터 수집 현황과 본 연구에서 제시한 보강 프로토콜의 적용 결과는 <표 4>와 같다. 사례연구 대상(Water Research)에 대해 API를 통해 수집된 문헌의 총 건수는 1기 2,810건에서 5기 6,686건으로 시기가 지날수록 지속적인 증가 추세를 보였다. 수집된 문헌 중 논문이 아닌 편집물(Editorial materials) 등을 제외하고 분석 대상 논문을 확정하였으며, 제외된 건수는 각 시기별로 약 60~110건 내외였다.

본 연구 대상 저널에 대해서 기관 정보 기준

〈표 4〉 사례연구 대상 논문(Water Research) 메타데이터 분석 결과

구분	연도				
	1기 (2000-2004)	2기 (2005-2009)	3기 (2010-2014)	4기 (2015-2019)	5기 (2020-2024)
문헌 총 수 (비논문 문헌 수)	2,810 (62)	2,733 (115)	3,306 (108)	4,600 (106)	6,686 (109)
기관 정보 기준 결측 수	126	73	62	335	73
보강 전 기관 정보 활용가능 비율	95.5%	97.3%	98.1%	92.7%	98.9%
RORID 기준 기관 정보 보완 수 (보완율)	85 (67.5%)	25 (34.2%)	10 (16.1%)	20 (6.0%)	8 (11.0%)
보강 후 기관 정보 활용가능 비율	98.5%	98.2%	98.4%	93.2%	99.0%

의 결측 비율은 최소 1.1%에서 최대 7.3%의 비율로 나타났으며, 전체적으로 보강 전의 기관 정보 활용가능 비율은 96.7%였다. 연구논문 수가 급증하기 시작한 4기(2015년~2019년)에는 결측 수가 335건으로 늘어나 활용 가능 비율이 92.7%로 일시적으로 하락하는 경향을 보였다. 이에 대해 Crossref와 ROR API를 활용한 보강 방법을 적용한 결과, 결측된 데이터 중 최소 6.0%에서 최대 67.5%가 복원되었다. 이를 통해 정보 활용 가능비율을 93.2%에서 99.0%까지 상승시킬 수 있었으며, 보강 후 전체 기간에 대한 평균 비율은 97.4%까지 상승되었다. 다만, RORID를 활용한 보강 과정에서 ROR에 미등록된 기관이거나 비영어권 기관명으로 표기된 경우에는 식별 및 보강에 다소 한계가 있음이 파악되었다.

4.2 기관의 연구 집중도 및 협업 네트워크 분석 결과

〈표 5〉와 〈표 6〉은 중심성 지표에 따른 상위 노드를 선정하기 위해 세 가지 중심성 지표에 따라 점수를 산정한 결과이다. 본 연구에서는 첫 번째 사례로서 기관 메타데이터의 보유 수준이 높아 활용 가능성이 가장 높은 5기(2020~2024년)를 기준으로 연구 집중도가 높은 상위 10개 기관(Top 10)을 도출하였고, 추가적으로 〈표 6〉에서는 다음으로 기관 정보 활용 가능 비율이 높은 1기(2000~2004년)의 결과를 제시하였다.

분석 결과, 최근인 5기에는 모든 중심성 지표에서 Chinese Academy of Sciences(중국과학원)를 필두로 한 중화권 기관들이 상위를 독점하는 경향이 뚜렷하게 나타났다. 특히 연결정도와 고유벡터 중심성뿐만 아니라, 매개 중심성에서도 중국 기관이 상위에 랭크되어 양적 성장과 더불어 네트워크 내 영향력 또한 확대된 것으로 파악된다. 다만, 매개 중심성을 기준으로 하는 경우에는 미국, 호주, 홍콩 등의 기관이 등장하여 최근 시점에서도 다양한 협업이 어느 정도 이루어지고 있는 것으로 판단된다.

반면에, 2000년대 초반(1기)의 경우에는 유

<표 5> 중심성 지표에 따른 집중도 상위 기관 지표 점수 분석 사례(Water Research, 2020~2024)

구분	중심성 지표 및 지표 점수					
	Degree(연결정도)		Eigenvector(고유벡터)		Betweenness(매개)	
Top 1	Chinese Academy of Sciences	0.5539	Chinese Academy of Sciences	0.6037	Chinese Academy of Sciences	0.1313
Top 2	University of Chinese Academy of Sciences	0.3308	University of Chinese Academy of Sciences	0.5183	University of Chinese Academy of Sciences	0.0922
Top 3	Institute of Environmental Chemistry	0.2180	Institute of Environmental Chemistry	0.4063	Geography Institute of China	0.0499
Top 4	State Key Laboratory of Pollution Control and Resource Reuse	0.2105	Geography Institute of China	0.1910	Hong Kong University	0.0492
Top 5	Hong Kong University	0.1980	State Key Laboratory of Pollution Control and Resource Reuse	0.1561	French National Centre for Scientific Research	0.0444
Top 6	Harbin Institute of Technology	0.1880	Tongji University	0.1259	UC Davis	0.0433
Top 7	Geography Institute of China	0.1754	Institute of Urban Environment	0.1096	The University of Queensland	0.0413
Top 8	Sun Yat-sen University	0.1704	Nanjing University	0.1011	State Key Laboratory of Pollution Control and Resource Reuse	0.0387
Top 9	Tongji University	0.1679	Shanghai Institute of Pollution Control and Ecological Security	0.0887	Institute of Environmental Chemistry	0.0380
Top 10	Tsinghua University	0.1679	Tsinghua University	0.0803	Nanjing University	0.0376

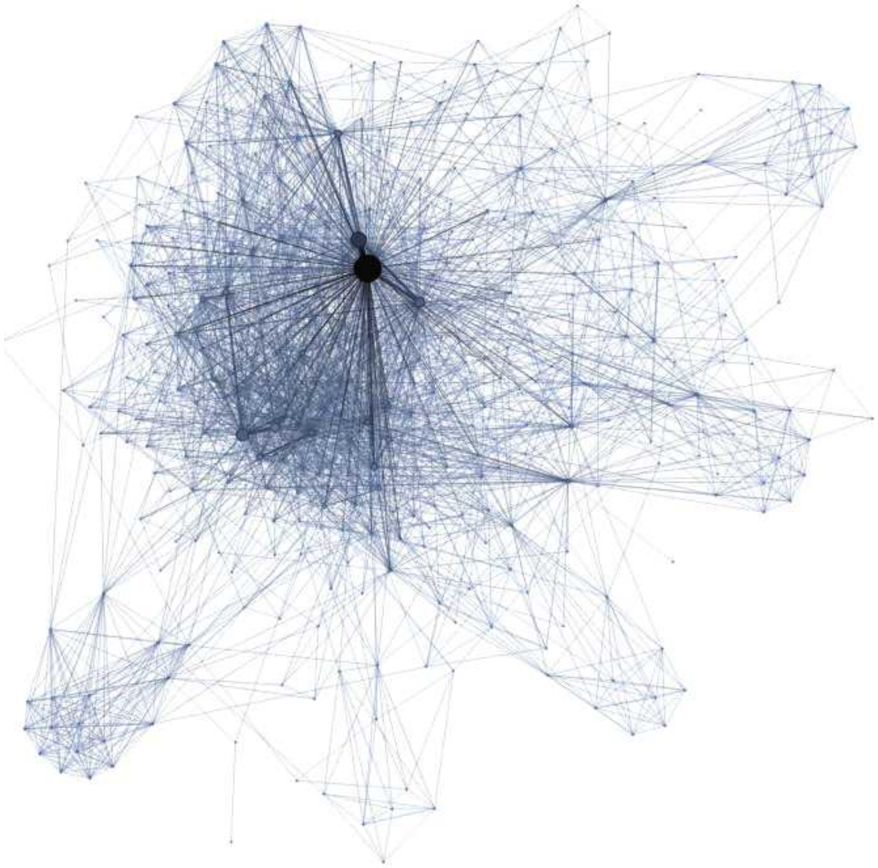
<표 6> 중심성 지표에 따른 집중도 상위 기관 지표 점수 분석 사례(Water Research, 2000~2004)

구분	중심성 지표 및 지표 점수					
	Degree(연결정도)		Eigenvector(고유벡터)		Betweenness(매개)	
Top 1	French National Centre for Scientific Research	0.1128	Hong Kong University	0.6968	University of Illinois	0.2452
Top 2	Laboratory of Physical Chemistry and Microbiology for Materials and the Environment	0.0501	Hong Kong University of Science and Technology	0.6932	McGill University	0.1689
Top 3	McGill University	0.0476	NTU	0.1072	French National Centre for Scientific Research	0.1610
Top 4	Chinese Academy of Sciences	0.0451	King Edward VII College of Medicine	0.0543	Environmental Protection Agency	0.1090
Top 5	Environment Agency	0.0426	Hokkaido University	0.0538	Swiss Federal Institute of Aquatic Science and Technology	0.1061
Top 6	Finnish Environment Institute	0.0426	The University of Queensland	0.0526	CU-Boulder	0.0912
Top 7	University of Lorraine	0.0426	Indian Institute of Technology Roorkee	0.0523	University of Lorraine	0.0812
Top 8	University of Barcelona	0.0351	Iowa State	0.0520	Iowa State	0.0737
Top 9	CU-Boulder	0.0351	Sydney University	0.0518	Tsinghua University	0.0694
Top 10	Environmental Protection Agency	0.0351	Dalian University of Technology	0.0517	Ghent University	0.0651

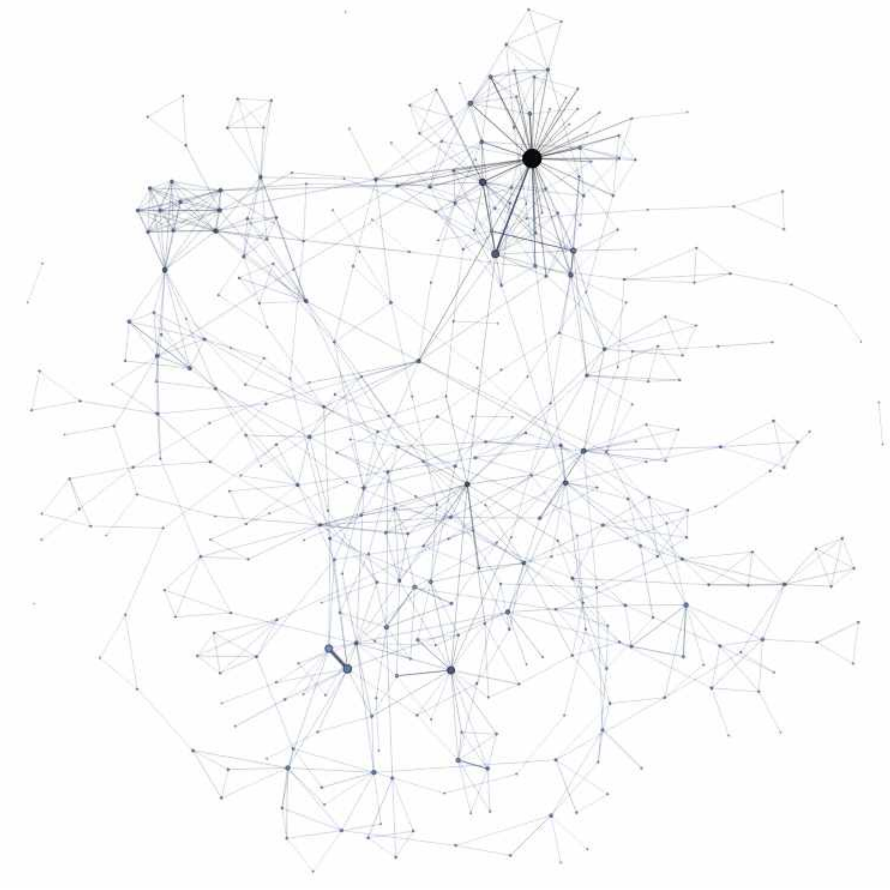
럽, 호주, 중국, 홍콩, 미국 등 다양한 지역의 연구기관 및 대학들이 지표별로 상이하게 상위를 차지하여, 다각화된 연구 활동이 이루어졌음을 알 수 있다. 특이한 점으로는, 상위 3개 기관만을 한정하여 볼 때, 연결정도 중심성과 매개 중심성이 높은 공통의 기관(French National Centre for Scientific Research와 McGill University)이 있었지만, 고유벡터 중심성을 기준으로 할 때는 홍콩과 싱가포르의 기관들이 상위에 랭크되는 차이점이 있었다. 이 시기에는 상대적으로 해당 국가들의 연구 활동이 활발해지면서 기관 상호 간의 협업이 강화된 것으로 분석할 수 있다.

〈그림 4〉와 〈그림 5〉는 각각 5기와 1기에 대한 네트워크 분석 시각화(NetworkX 기반) 결과를 나타낸 것이다. 네트워크 거점의 다양성은 최근 시기(5기)와 2000년대 초반(1기) 모두에서 확인되지만, 앞서 분석한 바와 같이 최근 시기의 경우 특정 거점으로의 집중 경향이 상대적으로 더욱 두드러지는 것으로 나타났다.

〈표 7〉은 시기별로 중심성 지표에 따른 최상위 집중도(Top 1)에 대한 분석 결과로서, 각 지



〈그림 4〉 네트워크 시각화 분석 사례(Water Research, 2020~2024)



〈그림 5〉 네트워크 시각화 분석 사례(Water Research, 2000~2004)

〈표 7〉 중심성 지표에 따른 최상위 집중도 분석 (Top 1) 결과

중심성 (Centrality)	시기(연도)				
	1기 (2000-2004)	2기 (2005-2009)	3기 (2010-2014)	4기 (2015-2019)	5기 (2020-2024)
Degree (연결정도)	French National Centre for Scientific Research (0.1128)	French National Centre for Scientific Research (0.1128)	French National Centre for Scientific Research (0.1404)	Chinese Academy of Sciences (0.2030)	Chinese Academy of Sciences (0.5539)
Eigenvector (고유벡터)	Hong Kong University (0.6968)	French National Centre for Scientific Research (0.4591)	Swiss Federal Institute of Aquatic Science and Technology (0.5519)	Chinese Academy of Sciences (0.4754)	Chinese Academy of Sciences (0.6037)
Betweenness (매개)	University of Illinois (0.2452)	Swiss Federal Institute of Aquatic Science and Technology (0.1663)	French National Centre for Scientific Research (0.2325)	Chinese Academy of Sciences (0.2567)	Chinese Academy of Sciences (0.1313)

표별로 산정된 점수를 함께 제시하였다. 앞선 분석 결과와 연계하여 볼 때, 1기(2000~2004년)에는 유럽, 중화권(홍콩 포함), 미국 등의 기관이 혼재되어 다양한 집중도를 보였으나, 2기~3기에는 중심성 지표와 관계없이 유럽 기관이 최상위를 모두 차지하는 경향을 보였다. 이후 4기(2015~2019년)부터는 중국의 단일 기관(Chinese Academy of Sciences)이 모든 지표에서 최상위 집중도를 유지하는 위치를 점하고 있는 것으로 나타났다.

<표 8>은 시기별로 상위 연결성을 가지는 협업 쌍(엣지)을 분석한 것이다. 협업 빈도수는 협업 쌍 아래에 기재하였으며, 최상위 연결성 협업 쌍은 동일한 빈도가 나타나는 경우는 없이 모두 하나씩 도출되었다. <표 7>과 <표 8>을 종합하여 보면, 4기 이후 최상위 집중도를 보여주는 기관(Chinese Academy of Sciences)이 5기 이후에는 협업에서도 최상위 랭크를 차지하고 있어 전체적인 연구 집중도에서 중국의 영향력이 커짐을 알 수 있었다. 또한, 노드 분석과 엣지 분석을 시기별로 비교함으로써 노드의 최상위 기관과 협업 최상위 기관이 겹치는 경우를 찾고자 하였다. 분석 결과, 고유벡터를 중심성 지표로 선정한 경우에 1기에는 Hong Kong University가, 3기에는 Swiss Federal Institute of Aquatic Science and Technology가, 5기에는 Chinese Academy of Sciences가 노드와 엣지 모두에서 공통 기관으로 나타나 연관성이 높은 것으로 검토되었다.

이러한 결과는 선호적 연결(Preferential Attachment) 원리 측면에서 해석될 수 있다. 선호적 연결 원리는 네트워크가 성장하는 과정에서 새로운 노드가 이미 많은 연결을 가진 노드와 연결될 확률이 높다는 이론으로, 그 결과 연결도가 높은 노드가 시간이 지남에 따라 더 많은 연결을 축적하며 네트워크 내에서 허브 노드로 발전하게 됨을 설명한다(Newman, 2001). 본 연구의 분석 결과는 물 분야의 최상위 학술지를 대상으로 한 기관 협업 네트워크 분석 결과에서도 허브 노드 현상이 관찰됨을 보여주며, 해당 분야의 기관 협업 구조를 이해하는 데 있어 선호적 연결 원리가 적용 가능하다는 근거를 제공할 수 있다.

<표 8> 시기별 상위 연결성 분석 결과

구분	시기(연도)				
	1기 (2000-2004)	2기 (2005-2009)	3기 (2010-2014)	4기 (2015-2019)	5기 (2020-2024)
Top 1	Hong Kong University – Hong Kong University of Science and Technology (26)	ETH Zurich – Swiss Federal Institute of Aquatic Science and Technology (24)	ETH Zurich – Swiss Federal Institute of Aquatic Science and Technology (51)	State Key Laboratory of Pollution Control and Resource Reuse – Tongji University (122)	Chinese Academy of Sciences – University of Chinese Academy of Sciences (339)
Top 2	French National Centre for Scientific Research – University of Lorraine (13)	Hong Kong University – Hong Kong University of Science and Technology (23)	State Key Laboratory of Pollution Control and Resource Reuse – Tongji University (40)	Chinese Academy of Sciences – Institute of Environmental Chemistry (113)	Chinese Academy of Sciences – Institute of Environmental Chemistry (246)

구분	시기(연도)				
	1기 (2000-2004)	2기 (2005-2009)	3기 (2010-2014)	4기 (2015-2019)	5기 (2020-2024)
Top 3	Nanjing University – State Key Laboratory of Pollution Control and Resource Reuse (8)	Chinese Academy of Sciences – Institute of Environmental Chemistry (18)	Hong Kong University – Hong Kong University of Science and Technology (37)	Nanjing University – State Key Laboratory of Pollution Control and Resource Reuse (106)	State Key Laboratory of Pollution Control and Resource Reuse – Tongji University (244)
Top 4	French National Centre for Scientific Research – University of Poitiers (7)	Delft University of Technology – IHE Delft Institute for Water Education (14)	Chinese Academy of Sciences – Institute of Environmental Chemistry (37)	Chinese Academy of Sciences – University of Chinese Academy of Sciences (94)	Shanghai Institute of Pollution Control and Ecological Security – Tongji University (170)
Top 5	Chinese Academy of Sciences – Institute of Environmental Chemistry (7)	French National Centre for Scientific Research – University of Lorraine (14)	Catalan Institute for Water Research – University of Girona (27)	Hong Kong University – Hong Kong University of Science and Technology (82)	Nanjing University – State Key Laboratory of Pollution Control and Resource Reuse (166)
Top 6	NSERC – UWaterloo (7)	French National Centre for Scientific Research – University of Poitiers (12)	Nanjing University – State Key Laboratory of Pollution Control and Resource Reuse (24)	ETH Zurich – Swiss Federal Institute of Aquatic Science and Technology (73)	Shanghai Institute of Pollution Control and Ecological Security – State Key Laboratory of Pollution Control and Resource Reuse (156)
Top 7	Environmental Protection Agency – University of Cincinnati (6)	State Key Laboratory of Pollution Control and Resource Reuse – Tongji University (11)	Delft University of Technology – KWR Water (24)	State Key Joint Laboratory of Environment Simulation and Pollution Control – Tsinghua University (69)	Institute of Environmental Chemistry – University of Chinese Academy of Sciences (148)
Top 8	NSERC – Polytechnique Montréal (6)	Australian Water Quality Centre – SA Water (11)	Swiss Federal Institute of Aquatic Science and Technology – Swiss Federal Institute of Technology in Lausanne (24)	Institute of Environmental Chemistry – University of Chinese Academy of Sciences (67)	Chinese Academy of Sciences – Geography Institute of China (121)
Top 9	Chinese Academy of Sciences – Institute of Hydrobiology (6)	French National Centre for Scientific Research – Pierre and Marie Curie University (9)	French National Centre for Scientific Research – University of Poitiers (24)	Shanghai Institute of Pollution Control and Ecological Security – Tongji University (58)	State Key Joint Laboratory of Environment Simulation and Pollution Control – Tsinghua University (121)
Top 10	French National Centre for Scientific Research – Laboratory of Physical Chemistry and Microbiology for Materials and the Environment (6)	State Key Joint Laboratory of Environment Simulation and Pollution Control – Tsinghua University (8)	Federal University of Toulouse Midi-Pyrénées – French National Centre for Scientific Research (22)	Shanghai Institute of Pollution Control and Ecological Security – State Key Laboratory of Pollution Control and Resource Reuse (54)	Hong Kong University – Hong Kong University of Science and Technology (110)

5. 결론 및 향후 연구 과제

본 연구는 오픈 사이언스 환경에서의 다양한 공개 자료를 기반으로 데이터를 수집하여 분석하는 표준화된 방법론을 제시하기 위해, 특정 분야의 주요 저널(Water Research)을 대상으로 논문의 메타데이터 중 기관 정보를 절차에 따라 추출하도록 하였다. 그 결과 95% 이상의 기관 정보가 확보되었으며, 이를 활용하여 해당 저널에 대해 시기별로 구분하여 네트워크 분석을 수행하고, 기관 간 연구 집중도와 협업도의 변화상을 파악할 수 있도록 하였다. 본 연구에서 제시된 방법론은 각 공개 학술 메타데이터 플랫폼(OpenAlex, Crossref, ROR)을 기반으로 하여 규칙을 설정함으로써 타 연구자도 활용할 수 있도록 재현성을 확보하고 필요시에는 보완하여 사용할 수 있도록 하였다. 다만, 본 연구는 상기 제시된 공개 플랫폼만의 메타데이터를 주로 활용하고 있기 때문에, 다양한 분야의 저널과 시기에 대해서 해당 플랫폼이 내재적으로 가지고 있는 한계의 영향력을 배제할 수 없는 특성이 있다. 따라서 향후 다수의 사례 연구를 추가함으로써 제시된 기법들의 타당성과 안정성 및 저널 메타데이터 정보 보강의 효율성을 보완할 수 있을 것이다. 또한, 저널 분야에서의 오픈 사이언스 형태의 확장에 맞추어, 공개 가능한 범위에서의 연구 초록 등을 수집하고 인공지능 언어모델(LM)이나 자연어처리 기법(NLP)을 병행 활용함으로써 연구 네트워크의 분석을 넘어, 목표로 하는 분석 분야의 핵심 기술 분류나 기술적 동향을 보다 정량화하여 객관적이고 효과적으로 파악하는 데 활용할 수 있을 것으로 기대된다. 또한, 본 연구에서는 물 분야 사례연구 적용을 위해서 물 분야 상위의 단일 저널을 대상으로 시기에 따른 연결정도, 고유벡터, 매개 중심성을 기반으로 하는 네트워크 분석을 수행하고 연구 집중도의 변화상을 추정하여 설명하는 것이 가능함을 보여주었다. 이에 향후 연구에서는 해당 분야의 적용 저널 수를 확대하거나, 상수도 처리, 하수 처리, 홍수・가뭄 등의 재난관리, 수생태 관리, 댐・하천 등 인프라 관리, 해수 담수화, 지하수 관리 등 광범위한 물 분야에서의 필요 영역에 대한 연구 집중도 및 협업 관계를 분석하고, 분야별 특성에 따른 네트워크 분석 방법을 마련할 수 있도록 하는 등, 보다 실무 활용성을 높이는 도구로서의 활용 방안을 마련할 필요성이 있다.

참 고 문 헌

김진영, 이석형, 서동준, 김광영 (2016). 국내 과학기술콘텐츠 전거데이터 구축을 위한 소속기관명 식별 방법과 시스템에 관한 연구. 디지털콘텐츠학회논문지, 17(6), 555-563. https://doi.org/10.9728/dcs.2016.17.6.555

김진영, 이석형, 서동준, 김광영, 윤정선 (2017). 국내 과학기술콘텐츠 저자의 소속기관명 식별을 위한 자동 식별 알고리즘에 관한 연구. 디지털콘텐츠학회논문지, 18(2), 373-382.

http://doi.org/10.9728/dcs.2017.18.2.373
백영기 (2012). 대학과 기업 간 상호협력에 따른 혁신창출－계량서지학적 분석. 한국경제지리학회지, 15(4), 493-514. https://doi.org/10.23841/egsk.2012.15.4.493
이석형 (2014). 학술논문 저자 소속기관 식별데이터 구축에 관한 연구. 사회과학연구, 25(4), 391-410. https://doi.org/10.16881/jss.2014.10.25.4.391
이재윤, 김판준, 강대신, 김희정, 유소영, 이우형 (2011). 계량서지적 기법을 활용한 LED 핵심 주제영역의 연구 동향 분석. 정보관리연구, 42(3), 1-26.
이해성, 이혜진 (2016). 연구 유사도 가중치 소셜 네트워크 생성과 연구자 관계 정보 도출을 위한 학술논문 메타데이터 분석 시스템 개발. 한국정보기술학회논문지, 14(1), 143-153. https://doi.org/10.14801/jkiit.2016.14.1.143
임윤정, 송규원, 박혜원, 조민상, 박병현, 정현준 (2020). 학술논문의 메타데이터 분석을 통한 연구자 관계성 분석. 한국정보과학회 학술발표논문집, 88-89.
Butt, B. H., Rafi, M., & Sabih, M. (2021). A systematic metadata harvesting workflow for analysing scientific networks. PeerJ Computer Science, 7, e421. https://doi.org/10.7717/peerj-cs.421
Cousijn, H., Kenall, A., Ganley, E., Harrison, M., Kernohan, D., Lemberger, T., Murphy, F., Polischuk, P., Taylor, S., Martone, M., & Clark, T. (2018). A data citation roadmap for scientific publishers. Scientific Data, 5, 180259. https://doi.org/10.1038/sdata.2018.259
Das, K., Samanta, S., & Pal, M. (2018). Study on centrality measures in social networks: A survey. Social Network Analysis and Mining, 8, 13. https://doi.org/10.1007/s13278-018-0493-2
Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research, 133, 285-296. https://doi.org/10.1016/j.jbusres.2021.04.070
Duran-Silva, N., Accuosto, P., Przybyła, P., & Saggion, H. (2024). AffilGood: Building reliable institution name disambiguation tools to improve scientific literature analysis. Association for Computational Linguistics. Proceedings of the Fourth Workshop on Scholarly Document Processing (SDP 2024), 135-144. https://doi.org/10.18653/v1/2024.sdp-1.13
Haak, L. L., Fenner, M., Paglione, L., Pentz, E., & Ratner, H. (2012). ORCID: A system to uniquely identify researchers. Learned Publishing, 25(4), 259-264. https://doi.org/10.1087/20120404
Kallipoliti, M., Chatzopoulos, S., Baglioni, M., Adamidi, E., Koloveas, P., & Vergoulis, T. (2025). From raw affiliations to organization identifiers. arXiv:2505.07577 [cs.DL]. https://doi.org/10.48550/arXiv.2505.07577

Lin, L., Xie, W., Chen, Z., & Hao, T. (2024). An end-to-end entity recognition and disambiguation framework for identifying author affiliation from literature publications. Association for Computational Linguistics. Proceedings of the Fourth Workshop on Scholarly Document Processing (SDP 2024), 120-129. https://doi.org/10.18653/v1/2024.sdp-1.11

Newman, M. E. J. (2001). The structure of scientific collaboration networks. Proceedings of the National Academy of Sciences of the United States of America, 98(2), 404-409. https://doi.org/10.1073/pnas.98.2.404

Orduña-Malea, E. & Bautista-Puig, N. (2022). Measuring web connectivity between research organizations through ROR identifiers. STI 2022 "From Global Indicators to Local Applications", 1-13. https://doi.org/10.5281/zenodo.6948453

Priem, J., Piwowar, H., & Orr, R. (2022). OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. arXiv:2205.01833 [cs.DL]. https://doi.org/10.48550/arXiv.2205.01833

United Nations Educational, Scientific and Cultural Organization (2021). UNESCO Recommendation on Open Science. https://doi.org/10.54677/MNMH8546

Velez-Estevez, A., Perez, I. J., García-Sánchez, P., Moral-Munoz, J. A., & Cobo, M. J. (2023). New trends in bibliometric APIs: A comparative analysis. Information Processing & Management, 60(4), 103385. https://doi.org/10.1016/j.ipm.2023.103385

Wagner, C. S. & Leydesdorff, L. (2005). Network structure, self-organization, and the growth of international collaboration in science. Research Policy, 34(10), 1608-1618. https://doi.org/10.1016/j.respol.2005.08.002

Waltman, L., Kramer B., Hendricks, G., & Vickery, B. (2020, September 25). Open Abstracts: Where are we? Crossref Blog. https://doi.org/10.64000/94km7-h9a72

Zhang, L., Cao, Z., Shang, Y., Sivertsen, G., & Huang, Y. (2024). Missing institutions in OpenAlex: Possible reasons, implications, and solutions. Scientometrics, 129(10), 5869-5891. https://doi.org/10.1007/s11192-023-04923-y

• 국문 참고문헌에 대한 영문 표기
(English translation of references written in Korean)

Beck, Young Ki (2012). The contribution of university-business interaction to innovation: Bibliometric analysis. The Economic Geographical Society of Korea, 15(4), 493-514. https://doi.org/10.23841/egsk.2012.15.4.493

Im, Yun-Jeong, Song, Gyuwon, Park, Hye-Won, Cho, Min-Sang, Park, Byung-Hyun, & Jung, Hyun-Jun (2020). Analyzing the relationship of researchers through metadata analysis of academic paper. Korean Institute of Information Scientists and Engineers, 88-89.

Kim, Jinyoung, Lee, Seok-hyong, Suh, Dongjun, & Kim, Kwang-Young (2016). A study on identifying institutional names for building authority data of Korean scientific and technological contents. Journal of Digital Contents Society, 17(6), 555-563. https://doi.org/10.9728/dcs.2016.17.6.555

Kim, Jinyoung, Lee, Seok-hyong, Suh, Dongjun, Kim, Kwang-Young, & Yoon, Jungsun (2017). A study on the identification algorithm for organization's name of author of Korean science & technology contents. Journal of Digital Contents Society, 18(2), 373-382. https://doi.org/10.9728/dcs.2017.18.2.373

Lee, Hae-Sung & Lee, Hye-Jin (2016). Development of the journal paper metadata analysis system for the creation of the research similarity-weighted social network and the extraction of researcher relationship information. The Journal of Korean Institute of Information Technology, 14(1), 143-153. https://doi.org/10.14801/jkiit.2016.14.1.143

Lee, Jae-Yun, Kim, Pan-Jun, Kang, Dae-Shin, Kim, Hee-Jung, Yu, So-Young, & Lee, Woo-Hyoung (2011). A bibliometric analysis on LED research. Journal of Information Science Theory and Practice, 42(3), 1-26.

Lee, Seok-hyong (2014). A study on the construction of identified data of author's affiliation in academic papers. Journal of Social Sciences, 25(4), 391-410. https://doi.org/10.16881/jss.2014.10.25.4.391