

디지털 큐레이션 플랫폼을 위한 신규 학술 콘텐츠 선정 지표 개발에 관한 연구*

A Study on Development of New Content Selection Indicators for Digital Curation Platform

이 석 형 (Seok-Hyoung Lee)** 이강산다정 (Kangsandajung Lee)***
김 재 훈 (Jayhoon Kim)**** 이 혜 진 (Hyejin Lee)*****

목 차

- 서론
- 선행 연구
- 디지털 큐레이션 플랫폼과 학술 콘텐츠
- 신규 콘텐츠 선정 지표 개발
- 결론 및 제언

초 록

연구자의 다양한 학술 콘텐츠 수요에 대응하기 위해 학술정보 제공기관에서는 새로운 유형의 콘텐츠를 발굴하고 디지털 큐레이션 플랫폼을 통해 구축·관리하고 있다. 그러나, 제한된 예산 및 인력과 콘텐츠 관리 프로세스 복잡화 등의 이유로 이용자의 신규 콘텐츠 요구를 효율적으로 대응할 수 있는 기준이 필요하다. 본 연구에서는 디지털 전환 기술을 적용하여 학술 콘텐츠의 수집, 구축, 관리를 자동화하여 연구자의 연구활동에 필요한 다양한 콘텐츠를 지원하는 디지털 큐레이션 플랫폼을 위한 신규 콘텐츠 선정 지표를 개발하고 국내 대표 과학기술분야 학술정보 유통기관인 A기관의 디지털 큐레이션 플랫폼에 시범 적용하였다. 신규 콘텐츠 선정 지표는 콘텐츠의 가치, 이용자의 요구도 및 만족도, 콘텐츠 수집, 구축, 관리의 용이성과 효율성을 포함하고 있으며, 디지털 자원 라이프 사이클 단계 중 입수 및 선정 단계에서 활용 가능하다.

ABSTRACT

Academic information providers are working on developing and establishing new types of content in response to the diverse academic content needs of researchers. To efficiently build and manage a wide range of content, these institutions are establishing digital curation platforms. These platforms apply digital transformation technologies to automate the collection, construction, and management of academic content, supporting researchers' activities by providing necessary content. To operate these platforms efficiently, it is essential to establish criteria that can effectively address the diverse content needs of users. We developed indicators for selecting new content and applied these selection priorities on a trial basis to the digital curation platform. The indicators for new content selection include content value, user demand and satisfaction, and the ease and efficiency of content collection, construction, and management. These indicators can be utilized during the acquisition and selection stages of the digital resource lifecycle.

키워드: 신규 콘텐츠 선정 지표, 디지털 자원 라이프사이클, 데이터 가치평가, 디지털 큐레이션 플랫폼
New Content Selection Indicators, Digital Resource Lifecycle, Data Value Assessment, Digital Curation Platform

- * 이 논문은 2024년도 한국과학기술정보연구원(KISTI)의 기본사업과제 “지능형 과학기술정보 큐레이션체제 구축”(K24L1M1C1-01)으로 수행한 내용의 일부를 재구성한 것임.
** 충남대학교 문헌정보학과 부교수(skyi@cnu.ac.kr / ISNI 0000 0004 6771 8906) (제1저자)
*** 한국과학기술정보연구원 디지털큐레이션센터 선임연구원
(lkdsj@kisti.re.kr / ISNI 0000 0004 7775 6308) (공동저자)
**** 한국과학기술정보연구원 디지털큐레이션센터 책임연구원
(jay.kim@kisti.re.kr / ISNI 0000 0004 6478 844X) (공동저자)
***** 한국과학기술정보연구원 디지털큐레이션센터 책임연구원/센터장
(hyejin@kisti.re.kr / ISNI 0000 0004 6490 0147) (교신저자)
논문접수일자: 2024년 8월 4일 최초심사일자: 2024년 8월 9일 게재확정일자: 2024년 8월 19일
한국문헌정보학회지, 58(3): 193-216, 2024. <http://dx.doi.org/10.4275/KSLIS.2024.58.3.193>

© Copyright © 2024 Korean Society for Library and Information Science
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. 서론

1.1 연구의 필요성 및 목적

디지털 전환 시대의 도래와 함께 연구자의 연구개발 행태는 크게 변화하고 있다. 빅데이터 분석과 인공지능 기술을 활용한 연구가 활발해지고 있으며 학문 분야간 경계가 허물어지면서 융합연구와 다학제적 연구가 증가하고 있다. 연구주제의 다양성과 연구개발 방법의 변화로 인해 연구자들은 다양한 유형의 학술정보 자원을 필요로 하는데 특히 논문, 특허, 보고서 등 전통적인 지식정보 콘텐츠 뿐만 아니라 데이터 기반 연구에서 많이 활용되는 연구데이터, 분석도구 및 학습모델과 같은 소프트웨어, 용어 사전, 그리고 실험 및 실습 인프라 등의 요구도 많아지고 있다(이석형, 이강산다정, 김재훈, 이혜진, 2024).

학술정보 제공기관들은 연구자들의 다양한 정보 및 데이터 수요를 충족시키기 위해 지속적인 콘텐츠 발굴이 필요하다. 펜데믹 시대에 오픈사이언스 운동의 확대로 수많은 정보와 데이터들이 개방되어 있지만, 제한된 인적·물적 자원으로 인해 적합한 콘텐츠를 선별하여 연구자에게 제공하는 것에 많은 제약이 따른다. 이러한 문제를 해결하기 위해 학술정보 제공기관들은 학술 콘텐츠 구축 관리 체계를 디지털 기술을 적용한 플랫폼 체계로 전환하여 구축 프로세스를 자동화하고 비용을 효율화하면서 콘텐츠 스펙트럼을 확대하고자 노력하고 있다.

특히, 국내의 대표적인 과학기술분야 학술정보 유통기관인 A기관은 국가 R&D 성과물과

학술논문 뿐만 아니라 인공지능 학습데이터와 학습모델, 용어사전 다양한 유형의 학술 콘텐츠를 디지털 기술을 적용하여 큐레이션을 자동화할 수 있는 디지털 큐레이션 플랫폼을 구축·운영 중에 있다. 디지털 큐레이션 플랫폼은 디지털 자원의 입수, 선정, 구축, 보존, 접근, 사용 및 재사용에 이르기까지의 라이프 사이클을 관리하는 활동을 지원하며, 이를 통해 다양한 형태로 연구를 수행하는 연구자의 학술 콘텐츠 수요에 대응하고자 한다.

이혜원, 윤소영, 박지영, 황혜경, 김재훈, 이혜림(2019)은 디지털로 표현된 학술 콘텐츠 대상의 큐레이션 활동을 큐레이션 라이프 사이클 모델(Curation Life-Cycle Model, 이하 CLCM)로 정의하여 제안하였다. CLCM의 순차적 활동을 큐레이션 업무의 개념화로부터 출발하여 콘텐츠의 생산 및 접수, 평가와 선별, 입수, 보존 활동, 저장, 접근, 이용 그리고 재사용, 이용자 경험 그리고 변환의 과정으로 설계하였다. CLCM 중 콘텐츠 요구가 발생(생산 혹은 접수)하면 큐레이션 대상 콘텐츠를 평가하고 선별하는 과정이 필요한데 A기관의 큐레이션 업무에서 보존 활동은 실행하고 있지 않으며 평가 및 선별 활동(특히 폐기와 관련하여)을 위한 구체적인 지침이나 기준은 마련되어 있지 않다고 하였다. 따라서, CLCM 기반의 순차적인 큐레이션 활동을 보장하기 위해서는 콘텐츠 수집 대상을 평가하고 선별할 수 있는 기준이 필요하다. 특히, CLCM을 구현한 디지털 큐레이션 플랫폼은 다양한 유형의 디지털 콘텐츠 큐레이션을 지향하고 있는데, 이들 자원의 효율적인 큐레이션을 위한 디지털 자원의 선정 지표가 고려되어야 할 것이다.

이에 본 연구에서는 디지털 큐레이션 플랫폼의 라이프 사이클을 지원하고 학술 콘텐츠 큐레이션의 효율적인 활동을 위해 콘텐츠 수집 대상을 평가하고 선별할 수 있는 신규 학술 콘텐츠 선정 지표를 개발하고 A기관의 디지털 큐레이션 플랫폼에 시범 적용하였다. 디지털 큐레이션 플랫폼이 학술활동에 필요한 모든 디지털 자원을 큐레이션 하는 것을 지향하는 시스템임을 감안하여 선정 지표 개발에 논문, 보고서, 특허 등 연구성과물부터 연구데이터 및 학습모델 등 데이터 기반 연구에 필요한 학술 콘텐츠까지 모두 고려하였다. 또한, 신규 학술 콘텐츠 선정 지표는 연구자 관점에서 학술 콘텐츠의 유용성 여부와 콘텐츠 수집, 정제, 구축 프로세스의 시간적, 비용적 효율성 여부를 모두 판단하여 수집 우선 순위를 산출할 수 있는 선정 지표를 제시하고자 한다.

1.2 연구 방법

본 연구는 디지털 큐레이션 플랫폼을 위한 신규 학술 콘텐츠 선정 지표를 개발하는데 그 목적이 있다. 본 연구에서는 연구개발 패러다임의 변화에 따른 연구자의 콘텐츠 요구 유형이 확대되고 있음을 고려하여 학술 콘텐츠의 범위를 논문, 특허, 보고서 등 연구성과물과 분야별 용어사전, 연구데이터, 그리고 학습데이터 및 학습모델 등도 포함하고자 한다. 선행 연구 분석 범위에 논문 등 연구성과물의 유용성 가치를 평가하는 연구 뿐만 아니라 데이터의 유용성과 가치를 평가하는 선행 연구를 분석하여 본 연구 목적에 부합하는 지표 구성과 지표 측정 방안을 도출하였다.

연구의 순서는 <표 1>과 같이 우선 신규 콘텐츠 선정 지표를 도출하기 위해 학술정보자원 개발 정책 및 콘텐츠 가치 평가 방법론 등 연구

<표 1> 연구절차 및 내용

단계	세부연구내용
1단계 이론연구	<ul style="list-style-type: none"> • 학술정보자원개발 정책, 콘텐츠 가치평가 방법론 선행 연구 분석
2단계 사례조사	<ul style="list-style-type: none"> • 디지털 큐레이션 플랫폼 추진 현황 및 콘텐츠 구축 현황 분석
3단계 지표도출 및 시범적용	<ul style="list-style-type: none"> • 학술 콘텐츠 유형 범주화 • 서비스 품질관리, 데이터 품질관리, 콘텐츠 가치평가 등을 참고한 신규 콘텐츠 선정 지표 도출 • 신규 콘텐츠 선정 지표 가중치 정의 • A 기관 디지털 큐레이션 플랫폼 시범 적용
4단계 시사점 도출	<ul style="list-style-type: none"> • 시사점 도출 및 향후 연구 주제 제안

목표와 관련한 선행 연구를 분석하고 디지털 큐레이션 플랫폼 추진 현황을 살펴보았다. 그리고, 디지털 큐레이션 플랫폼에서 수집, 구축 대상 학술 콘텐츠 유형에 대한 범위를 정의한 후, 학술 콘텐츠 이용의 유용성, 활용성 및 콘텐츠 관리의 효율성을 고려한 신규 학술 콘텐츠 선정 지표를 도출하였다. 도출된 선정 지표에 대한 지표 가중치를 설정하고 A기관의 디지털 큐레이션 플랫폼에 시범 적용한 후 시사점을 도출하였다.

2. 선행 연구

학술 콘텐츠 수집·구축의 효율화 측면에서 신규 학술 정보원이나 콘텐츠를 선정하기 위한 지표를 연구한 사례로는 크게 학술정보자원 개발 방법론과 과학기술정보 가치를 평가하는 방법론을 제시한 연구를 들 수 있다.

장덕현(2007)은 과학기술분야 학술정보자원 개발 정책에 관한 연구에서 전자환경의 도서관 정보센터의 장서개발 정책이 기존의 인쇄매체 중심에서 전자자원 중심으로 변화해야 하며, 과학기술 분야와 같이 변화의 속도가 빠르고 학문 분야간 융합이 빈번한 경우 시의적절한 장서개발 정책이 필요하다고 하였다. 장서개발 정책에 있어 고려되어야 할 요소로 장서의 질, 수서 의지, 장서 목표, 자료의 유형, 주제 분야, 학술 유형, 발간 매체 형식, 사용 언어 및 발행 국가 등을 제시하였다. 또한, 입수 대상 자원 선정 프로세스에서 수집 자료의 가치성(SCI 수록 여부, 상용 데이터베이스 이름, Impact Factor 등)과 요구성(복사건수, 수요조사점수, 로그 데이터)을 포함

하여 학술지 평가에 반영하는 것을 제안하였다.

유승훈, 허재용, 안윤기(2009)는 지식정보의 확대·재생산과 효율적인 관리정책방향 수립에 기준을 지식정보 가치로 정의하고 지식정보의 가치를 측정할 수 있는 방안을 제시하고 지식정보 가치에 영향을 미치는 요인을 도출하였다. 지식정보의 가치를 이용적 가치, 보존적 가치, 내재적 가치로 구분하고 이용적 가치에 이용요금, 접근 용이성, 활용성, 이용관심도, 유용성, 부가가치창출, 상업성, 활용 범위를 세부 평가지표로 설정하였다. 보존적 가치에 한국적 고유성, 유일성, 학술성을 세부 평가지표로 설정하였고, 내재적 가치로는 권위, 자원 투입량, 공익성, 디지털화, 정확성, 완전성, 최신성, 객관성, 포괄성을 세부 평가지표로 설정하였다.

이선희, 김혜선(2014)은 국내 과학기술 연구자의 요구를 만족시키는 고품질 콘텐츠를 제공하기 위한 기초작업으로 이용빈도, 페이지뷰 등 평가지표 기반의 스코어링을 통해 논문 콘텐츠의 평가기준을 제시하였다. 논문 콘텐츠 가치를 이용성과 내재성으로 구분하고 이용성 가치 세부평가지표로 전문 접근성, 이용빈도, 언어가독성을 포함하였으며, 내재성 가치 세부평가지표로 권위, NDSL 이용지표, 완전성, 최신성, 정확성을 포함하였다. 이들 세부평가지표별 판단기준은 논문 메타데이터와 통계데이터를 정량화하는 것으로 설정하였다.

성태웅, 전승표, 변정은, 박현우(2016)는 과학기술정보 콘텐츠 가치평가모형 프레임워크 연구에서 특허, 논문, 보고서 등을 포함하는 전자적 데이터베이스와 같은 지식정보 콘텐츠의 시장가치를 객관적으로 평가할 수 있는 모형을 제시하였다. 일반적인 무형자산의 평가기법을 활

용하여 과학기술콘텐츠 유형별로 경제적 수명 주기, 현금흐름 추정 등을 적용한 가치평가 모형을 설계하고 평가 방법을 제시하였다. 콘텐츠의 경제적 수명은 기술순환주기, 콘텐츠 시장수요, 콘텐츠 등록 및 제작 이후 경과년수 등을 고려한 지표이며, 현금 흐름 추정 방법은 특허의 경우 기술사업화에 따른 매출액, 논문 및 보고서의 경우 투입인건비, 논문 제작비용, 논문 페이지 부, 다운로드 수 등을 고려한 콘텐츠 가치 정도를 산출하는 방법을 제안하였다.

김성훈, 도슬기, 한상은, 김재훈, 임석중, 박진호(2022)는 디지털 큐레이션 성숙도 모델 및 지표 개발에 관한 연구에서 디지털 큐레이션 플랫폼의 디지털 전환 성과를 측정할 수 있는 지표 개발을 시도하였다. 평가 척도를 기술, 데이터, 전략, 조직(인력), 영향력 등 5개 대분류를 중심으로 각 척도별로 측정 지표를 정의하였다. 특히, 본 연구와 관련하여 데이터 범주 내에 큐레이션 대상 데이터의 다양성, 적합성, 활용성, 적시성, 상호운용성, 보안, 유지보수성, 준비성, 완전성, 유용성 정도를 측정하여 디지털 큐레이션 플랫폼의 완성도를 측정하는 방법을 제시하였다.

이들 선행 연구는 논문, 특허, 보고서 등 학술 정보자원개발 정책과 콘텐츠 가치평가에 대한 중요한 기준들을 제시하였다. 그러나 최근의 데이터 기반 연구개발 패러다임으로 인해 다양한 유형의 학술 콘텐츠가 요구되고 있는데 이들 콘텐츠를 효율적으로 수집 구축하기 위해서는 학술정보자원개발 정책과 콘텐츠 가치평가 방법론이 종합적으로 고려되어야 할 것이다. 또한, 이용자 관점에서 콘텐츠의 유용성과 활용성 정도를 판단할 수 있는 지표를 도출하기 위해서는 LibQUAL과 같은 학술정보서비스의 품질평가

모델을 고려해야 하며 콘텐츠 관리의 효율성 정도를 판단할 수 있는 지표를 도출하기 위해서는 데이터 품질관리모델도 고려해야 한다. 이에 본 연구에서는 선행 연구에서 언급한 콘텐츠 평가 방법론들을 포함하여 학술정보서비스 품질평가 모델 중 송광석, 유한주, 김완중, 김종훈(2015)이 제시한 이용자 기반 NDSL정보서비스 품질 측정 모형과 이해진, 신재민, 배수영, 최희석, 이석형(2021)의 과학기술 지식인프라 통합서비스의 플랫폼의 품질평가모델, 데이터 품질관리 모델 중 한국지능정보사회진흥원(2022)의 공공데이터 품질관리 수준진단 평가 매뉴얼을 참고하여 신규 콘텐츠 선정 지표를 개발하였다.

3. 디지털 큐레이션 플랫폼과 학술 콘텐츠

본 연구는 이용자의 다양한 콘텐츠 요구에 효율적으로 대응하고 학술연구 활용 목적의 디지털 자원을 망라적으로 큐레이션 할 수 있는 디지털 큐레이션 플랫폼에 적합한 신규 콘텐츠 선정 지표를 개발하고 시범 적용하여 디지털 큐레이션 라이프 사이클을 지원하기 위한 기초 연구이다. 이에 본 장에서는 신규 콘텐츠 선정 지표를 시범 적용할 A기관의 디지털 큐레이션 플랫폼과 해당 플랫폼에서 다루고 있는 학술 콘텐츠 구축 현황을 분석하여 신규 콘텐츠 선정 지표 개발 범위를 정의하고자 하였다.

3.1 디지털 큐레이션 플랫폼

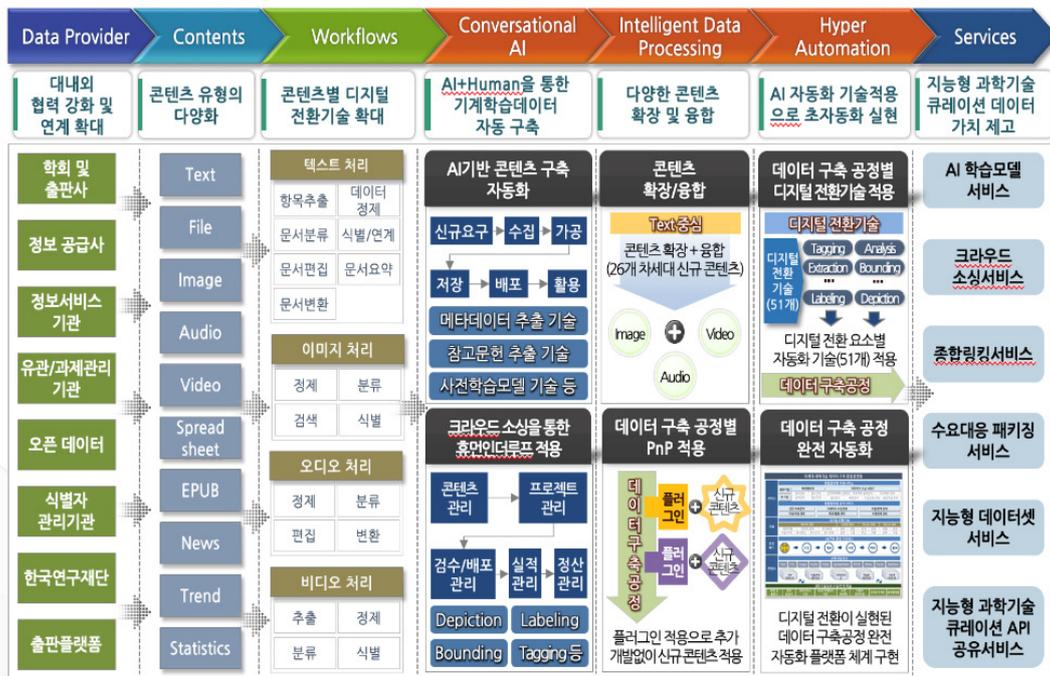
김주섭, 전예린, 김선태(2021)는 A기관의 학

술 콘텐츠의 구축 범위를 논문, 특허, 보고서, 연구데이터 국가 R&D 관련 정보 등으로 나눌 수 있으며 국내 학술논문과 해외 학술논문이 별도의 관리시스템을 통해서 구축된다고 분석하였다. 국가 R&D 성과물 보고서는 NRMS(National R&D Reports Management System)를 통해 등록되며 논문, 보고서 성과물들은 과학기술 지식인프라 통합서비스, 국가과학기술지식정보서비스 등을 통해 서비스되고 있다.

이후, 김성훈, 도솔기, 한상은, 김재훈, 임석중, 박진호(2022)는 디지털 전환시대 조직의 업무 환경 변화에 대응하기 위해 성숙도 모델의 개념을 차용하여 비즈니스 프로세스 재설계 과정을 거쳤고 그 결과 디지털 큐레이션 플랫폼 기반의 학술 콘텐츠 구축 관리 체계로 전환하고

있다고 하였다. 또한, 학술 콘텐츠 서비스 범위를 확대하고 AI 기술을 활용한 디지털 전환 플랫폼으로서의 데이터 기반 R&D를 위한 디지털 큐레이션 플랫폼을 단계적으로 구현하고 있다. <그림 1>은 A기관의 차세대 디지털 큐레이션 플랫폼의 워크플로우이다.

<그림 1>의 디지털 큐레이션 플랫폼의 가장 큰 특징은 논문, 보고서 중심의 학술 콘텐츠 구축 관리에서 콘텐츠 유형의 다양화와 AI 기술을 활용한 콘텐츠 구축의 자동화, 데이터 구축 공정별 플러그인 적용으로 신규 콘텐츠 적용의 유연성 확보 등을 들 수 있다. 디지털 큐레이션 플랫폼은 인간-인공지능 협업 방식의 대화형 인공지능, 데이터 확장성과 유연성을 지원하는 지능형 데이터 처리, AI로 완전 자동화가 실현



<그림 1> A기관 디지털 큐레이션 플랫폼 워크플로우

된 초자동화 기반의 큐레이션 플랫폼을 지향하고 있으며, 반대로 AI 기술에 활용될 수 있는 학습용 데이터의 구축 및 재사용성을 지원하는 체제를 지향하고 있다.

또한, 어떤 형태의 데이터에 대해서도 확장성이 용이하고 처리가 가능한 지능형 데이터 처리 플랫폼 구축을 위해 데이터베이스 구축 프로세스를 신규요구, 수집, 가공, 저장, 배포, 활용의 6단계로 표준화하여 정의하였으며 원문 중심의 학술 콘텐츠에서 텍스트, 이미지, 오디오, 비디오, 연구데이터, 통계데이터 등 다양한 유형의 콘텐츠를 제공하는 것을 고려하였다.

디지털 큐레이션 플랫폼은 김성훈, 도슬기, 한상은, 김재훈, 임석중, 박진호(2022)의 연구와 김성훈, 박진호(2023)의 연구를 통해 디지털 큐레이션 성숙도 모델을 기반으로 단계적으로 구축되고 있으며, 2023년 기준으로 국내논문 자동구축, 메타데이터/검색용 데이터 자동구축 프로세스 구현, 개체추출 데이터셋 구축 확대 등의 디지털 전환 사례 구현으로 디지털 큐레이션 성숙도 성장 수준을 달성하였다(한국과학기술정보연구원, 2023).

3.2 학술 콘텐츠 구축 및 관리 현황

A기관은 과학기술정보유통전담기관으로 국내외 학술논문, 국가R&D 연구성과물을 핵심 콘텐츠로서 구축하고 있으나, 최근 데이터 기반 R&D의 지원과 정보·데이터 융합 콘텐츠 제공 수요 등을 반영하여 인물/기관 식별데이터와 과학기술정보 분석용 기계학습데이터를 수집하거나 자체 제작하여 제공하고 있다.

2023년 기준으로 매년 약 국내논문메타 및 원문 4만건, 해외 논문메타 3백만건, 국가 R&D 보고서 메타 및 원문 2만건, 논문성과물 25만건 등이 수집 구축 대상이며, 인물/기관 식별데이터는 연간 15만건의 국내 학술논문을 대상으로 90만 인물/기관 개체를 구축하고 있다. 디지털 큐레이션 플랫폼을 통해 수집, 구축된 국내외 학술논문, 국가 R&D 연구성과물, 인물/기관 식별데이터, 기계학습데이터는 과학기술 지식인프라 통합서비스를 통해 이용자들에게 제공되고 있다. 기계학습데이터는 연구보고서내 표/그림 데이터 셋, 기관식별데이터, 국내논문 질의응답 데이터셋, 논문 전문 텍스트 데이터셋 등을 공개하고 있다.

한편, A기관은 교육부 산하 B기관으로부터 국내 학위논문을 수집하고 있으며, C기관으로부터 한국특허, 미국특허, 일본특허, 유럽특허 등 국내외 특허정보를 수집하여 서비스하고 있다. <표 2>는 2024년 7월 현재 A기관의 학술정보서비스 등을 통해 제공하고 있는 학술정보 구축 범위를 보인 것이다.

4. 신규 콘텐츠 선정 지표 개발

본 연구의 목적을 달성하기 위해 본 장에서는 디지털 큐레이션 플랫폼에서 처리할 수 있는 콘텐츠 유형 분류를 체계화하여 신규 콘텐츠 발굴 영역을 식별하고, 콘텐츠 유형에 적합한 콘텐츠 평가 방법론 선행 연구를 참고하여 본 연구에 맞는 신규 콘텐츠 선정 지표를 도출하였다.

〈표 2〉 학술정보유형별 구축 및 서비스 현황

학술정보유형	수집/구축 범위	수집/구축 방법	제공건수
국내학술논문	메타 및 원문	학협회 협약을 통한 수집/구축	3,625,833건
국내학술회의	메타 및 원문	학협회 협약을 통한 수집/구축	400,338건
해외학술논문	메타/원문라이선스	출판사 계약을 통한 수집/구축	112,242,899건
해외학술회의	메타/원문라이선스	출판사 계약을 통한 수집/구축	12,493,314건
국내학위논문	메타 및 원문URL	한국교육학술정보원 연계/수집	1,312,603건
연구보고서	메타 및 원문	성과물전담기관 연계/수집	352,888건
국내특허	메타 및 원문URL	특허청/특허정보원 연계/수집	6,688,077건
해외특허	메타	특허청/특허정보원 연계/수집	39,316,152건
동향정보	원문	자체구축	67,417건
인물	메타	자체구축	917,213건
기관	메타	자체구축	245,692건
기계학습데이터	데이터셋	자체구축	4,550,169건

4.1 학술 콘텐츠 범주화 및 신규 콘텐츠 발굴 영역 정의

〈그림 1〉의 디지털 큐레이션 플랫폼은 다양한 학술 콘텐츠 유형을 대상으로 디지털 기술을 적용한 큐레이션 활동을 지원하는 특징이 있으나 구체적으로 학술 콘텐츠의 유형을 명확하게 정의하고 있지는 않다. 큐레이션 대상 학술 콘텐츠 유형을 명확히 범주화하는 것은 정보원이 다르지만 동일 유형 콘텐츠를 구축·관리하는 프로세스의 표준 가이드라인을 정의하는데 필수적이라고 할 수 있다. 이에 본 연구에서는 연구자들이 필요로 하는 학술정보 콘텐츠 유형의 범주화에 대해 정의한 선행 연구를 근거로 콘텐츠 유형을 범주화하고 신규 콘텐츠 발굴 영역을 도출하였다.

케이티지(2002)의 과학기술정보 수요조사 및 정보가치지도 작성 연구에서는 과학기술정보유형을 학술저널, 전문도서 및 단행본, 연구개발보고서, 학술회의자료, 학위논문과 같은 과학기술 문헌정보와 기술/연구동향과 같은 기술동향정

보, 정책정보, 시장마케팅정보, 생산기술정보(통계/표준/규격/관측/측정자료/연구기자재/설비 등), 과학기술인력정보, 특허정보 등으로 정의하였다. 한국교육학술정보원(2004)의 학술정보 서비스 이용행태 분석을 통한 서비스 개선방안 연구에서는 학술문헌을 국내외 학술지논문, 국내외 단행본, 국내외 학위논문 등으로 정의하여 학술정보 서비스를 위한 이용자 인터페이스와 서비스 기능에 대한 개선 방향을 도출하였다.

권나현, 이정연, 정은경(2012)의 연구에서 생명·나노분야 연구자들이 주요 활용하는 정보유형을 논문, 단백질 데이터, 단행본, 동향정보, 이미지, 동영상 강의, 특허, 용어 정의 등으로 파악하였다. 반면, 과학기술정책연구원(2020)의 데이터 기반 연구개발 관리 혁신 방안에서는 논문, 특허, 보고서, 기술요약정보, 소프트웨어, 화합물, 생명자원, 신제품, 연구시설장비 정보 등 9대 연구성과물 정보로 범주화하고 성과확산을 위한 플랫폼 운영을 제안하였다. 이와 같이 연구자의 연구개발을 지원하고 연구생산성을 향상시키는데 필요한 학술 콘텐츠 유형을 연구한

결과를 <표 3>과 같이 상위/하위 분류 기준에 따라 매핑시켰다.

<표 3>의 콘텐츠 유형 중에서 학술논문, 학술회의자료, 연구보고서, 연구자정보, 연구기관 정보 등은 디지털 큐레이션 플랫폼에서 다양한 정보원으로부터 구축 중이며, 학위논문, 특허정보, 분석/정책/시장보고서, 동향정보, 산업기술 정보 등은 A기관의 과학기술 지식인프라 통합

서비스에서 직접 수집하여 연구자에게 제공하고 있다. 국가 R&D 정보와 연구시설장비, 표준/규격, 화합물, 생명자원, 신물질 등은 국가과학기술 지식정보서비스에서 수집하고 있으며, 연구데이터는 국가연구데이터플랫폼에서 수집·제공하고 있다. 반면, 용어, 동영상, 통계자료, 소프트웨어 등은 현재 미구축/미수집 중인 콘텐츠라고 할 수 있다. 미구축/미수집 중인 콘텐츠

<표 3> 디지털 큐레이션 플랫폼을 위한 학술 콘텐츠 범주화

상위분류	하위분류	참조영역	비고
논문	학술논문	케이티지(2002), 한국교육학술정보원(2004), 권나현 외(2012), 과학기술정책연구원(2020)	수집·구축중
	학술회의자료	케이티지(2002)	수집·구축중
	학위논문	케이티지(2002), 한국교육학술정보원(2004)	ScienceON 수집중
특허	특허정보	케이티지(2002), 권나현 외(2012), 과학기술정책연구원(2020)	ScienceON 수집중
보고서	연구보고서	케이티지(2002), 한국교육학술정보원(2004), 권나현 외(2012), 과학기술정책연구원(2020)	수집·구축중
	분석/정책/시장보고서	케이티지(2002)	ScienceON 수집중
동향	기술동향정보	케이티지(2002), 한국교육학술정보원(2004), 권나현 외(2012)	ScienceON 수집중
	정책동향정보	케이티지(2002)	ScienceON 수집중
산업기술	산업기술정보	케이티지(2002)	ScienceON 수집중
연구자	연구자정보	케이티지(2002)	수집·구축중
연구기관	연구기관	케이티지(2002)	수집·구축중
용어	용어	권나현 외(2012)	미구축/미수집
국가R&D	국가R&D사업	과학기술정책연구원(2020)	NTIS 수집중
	국가R&D과제	과학기술정책연구원(2020)	NTIS 수집중
	과제공고	과학기술정책연구원(2020)	NTIS 수집중
연구시설장비	연구시설장비	케이티지(2002)	NTIS 수집중
표준/규격	표준/규격	케이티지(2002)	NTIS 수집중
화합물	화합물	과학기술정책연구원(2020)	NTIS 수집중
생명자원	생명자원	과학기술정책연구원(2020)	NTIS 수집중
신물질	신물질	과학기술정책연구원(2020)	NTIS 수집중
소프트웨어	소프트웨어	과학기술정책연구원(2020)	미구축/미수집
동영상	동영상강의자료	권나현 외(2012)	미구축/미수집
통계자료	통계자료	케이티지(2002)	미구축/미수집
연구데이터	연구데이터	케이티지(2002), 한국교육학술정보원(2004), 권나현 외(2012), 과학기술정책연구원(2020)	DataON 수집

츠는 디지털 큐레이션 플랫폼 상에서는 신규 구축 콘텐츠 영역으로 볼 수 있으며 과학기술 지식인프라 통합서비스, 국가연구데이터플랫폼, 국가과학기술지식정보서비스에서 수집하고 있는 콘텐츠 유형도 디지털 큐레이션 플랫폼 내 콘텐츠와의 연계·융합이 가능하거나 신규 구축 콘텐츠 영역으로 확대할 수 있다.

4.2 신규 콘텐츠 선정 지표 도출

본 절에서는 디지털 큐레이션 플랫폼에서 논문, 특허, 보고서 등 연구성과물, 용어사전, 연구데이터, 소프트웨어 등과 같은 학술 콘텐츠의 수집·구축의 효율화를 위한 신규 콘텐츠 선정 지표를 도출한다. 이를 위해, 선행 연구 결과를 토대로 수집 대상 콘텐츠의 가치, 이용자의 요구도 및 만족도 지표 등 이용자 관점에서의 신규 콘텐츠 선정 지표와 콘텐츠 수집, 구축, 관리의 용이성과 효율성을 고려한 콘텐츠 제공자 관점에서의 콘텐츠 선정 지표를 고려하였다. 이용자 관점에서의 콘텐츠 선정 지표는 정보·데이터의 가치평가 모델, 정보서비스 품질평가 모델 등을 참고하여 구성하였고, 콘텐츠 제공자 관점에서의 콘텐츠 선정 지표는 디지털 큐

레이션 성숙도 모델, 데이터 품질관리 모델 등을 참고하여 구성하였다.

이용자 관점에서의 신규 콘텐츠 선정 지표 선정을 위해 NDSLQUAL(송광석, 유한주, 김완중, 김종훈, 2015)과 과학기술 지식인프라 통합서비스 플랫폼의 품질평가모델(이혜진, 신재민, 배수영, 최희석, 이석형, 2021), 과학기술콘텐츠 자산가치 평가모형(성태웅, 전승표, 변정은, 박현우, 2016)을 참고하였다.

NDSLQUAL(송광석, 유한주, 김완중, 김종훈, 2015)은 SERVEQUAL, e-SERVEQUAL, SiteOUAL, WebQUAL, LibQUAL 등 정보서비스 관련 문헌을 토대로 개발된 NDSL 품질평가모델이다. 구 NDSL(현 ScienceON)의 서비스 특성을 반영한 핵심 서비스 변수를 선정하고, 기대와 인지 수준에 대한 측정을 통해 서비스에 대한 갭 분석을 실시하여 서비스 품질지수를 제시하였는데 <표 4>의 주요 품질 차원 중에 신뢰성, 편리성, 시스템 사용성, 대응성, 보안, 서비스 문제 해결 등은 NDSL 서비스 인터페이스를 활용하는 서비스 이용 측면에서의 지표이며 정보품질은 NDSL에서 제공하는 정보의 최신성, 다양성, 완전성, 정확성 측면에서의 지표라고 볼 수 있다. 본 연구에서 신규 콘텐

<표 4> NDLSQUAL의 품질차원(송광석, 유한주, 김완중, 김종훈, 2015)

품질차원	주요 평가 요소
신뢰성	약속한 서비스를 믿을 수 있고 정확하게 실천하는 능력
편리성	NDSL이 제공하는 서비스에 대해 이용자들이 느끼는 이용의 편리성
시스템 사용성	NDSL의 유형성 특성인 사이트 디자인, 속도, 사이트의 구조적 특성
대응성	고객을 돕고 신속한 서비스를 제공하려는 의지
보안	NDSL의 개인정보 보안 수준에 대해 이용자들이 느끼는 정도
정보품질	NDSL에서 제공하는 정보의 최신성, 다양성, 완전성, 정확성
서비스 문제해결	NDSL이 제공하는 서비스 이용시 문제가 발생했을 경우 문제가 발생한 서비스를 해결하기 위한 일련의 조치와 상호작용

츠 우선순위 선정에 적합한 지표로는 정보 품질 지표를 참고할 수 있으며 정보의 최신성, 다양성, 완전성, 정확성 등이 연관되어 있다고 판단하였다.

과학기술 지식인프라 통합서비스 플랫폼의 품질평가모델(이혜진, 신재민, 배수영, 최희석, 이석형, 2021)은 연구자에게 신뢰성 있는 지식 인프라 서비스를 위해 지식인프라 서비스 품질

지표 및 측정항목을 개발하고 로그데이터를 활용한 평가방법론을 개발하였다. 통합서비스 플랫폼 품질평가모델의 품질차원은 크게 서비스/시스템, 콘텐츠, 이용자 측면으로 설계되었는데, 서비스/시스템 차원은 연구자의 연구활동을 지원하는 과학기술 지식인프라 서비스 인프라를 의미하여 본 연구의 신규 콘텐츠 수집 및 이용을 통한 가치 창출 측면에서 부합하지 않

<표 5> 과학기술 지식인프라 통합서비스 플랫폼 품질평가모델(이혜진, 신재민, 배수영, 최희석, 이석형, 2021)

측정 구분	품질 지표	설명
서비스/ 시스템	가용성	시스템 안정성, 시스템 가동율 등 서비스 품질 요소인 시스템 안정적 운영을 측정하는 지표
	보안성	ScienceON에서 제공하는 서비스/시스템의 개인정보 보호 및 웹사이트 보안 수준을 측정하는 지표
	사용편의성	명시된 조건에서 사용될 경우 사용자에게 의해 이해·학습·사용·선호될 수 있는 서비스/시스템 및 콘텐츠 제공여부를 측정하는 지표
	서비스활용성	이용자들에게 활용 가치가 있는 서비스/시스템 및 콘텐츠 제공여부에 대하여 정확성, 완전성, 적합성 등을 활용하여 측정하는 지표
	신뢰성	서비스/시스템 및 콘텐츠의 기술적 기능과 명시된 요구의 이행 정확성 및 적시성을 측정하는 지표
	연계용이성	다른 정보원 접근 편리성, 연계로그인 용이성 등 서비스 및 기능간 연계 정도를 측정하는 지표
	준비성	정기적인 업데이트로 최신 정보의 정도를 측정하는 지표
	협조체계	데이터/서비스 변경 및 장애 시 협조를 위한 관리체계 운영 정도를 측정하는 지표
이용자	사용편의성	명시된 조건에서 사용될 경우 사용자에게 의해 이해·학습·사용·선호될 수 있는 서비스/시스템 및 콘텐츠 제공여부를 측정하는 지표
	서비스활용성	이용자들에게 활용 가치가 있는 서비스/시스템 및 콘텐츠 제공여부에 대하여 정확성, 완전성, 적합성 등을 활용하여 측정하는 지표
콘텐츠	사용편의성	명시된 조건에서 사용될 경우 사용자에게 의해 이해·학습·사용·선호될 수 있는 서비스/시스템 및 콘텐츠 제공여부를 측정하는 지표
	서비스활용성	이용자들에게 활용 가치가 있는 서비스/시스템 및 콘텐츠 제공여부에 대하여 정확성, 완전성, 적합성 등을 활용하여 측정하는 지표
	신뢰성	서비스/시스템 및 콘텐츠의 기술적 기능과 명시된 요구의 이행 정확성 및 적시성을 측정하는 지표
	정기성	정기적인 업데이트로 최신 정보의 정도를 측정하는 지표
	최신성	최신 정보의 적시 제공/콘텐츠의 지속적 업데이트 등을 측정하는 지표

아 본 연구 범위에서는 고려하지 않았다. 콘텐츠 차원의 품질 측정 대상은 논문, 특허, 보고서, 동향, 연구자 정보 등에 대한 데이터베이스를 의미하며 이를 구성하고 있는 정보의 적시성, 정확성, 적절성 등을 평가함으로써 서비스/시스템에서 제공되고 있는 정보의 전반적 사항에 대하여 평가하는 지표를 의미한다. 이용자 차원의 지표들은 이용자가 서비스를 사용하며 느끼는 만족도와 같은 비계량적 지표뿐만 아니라 더불어 미시적/거시적 성과와 같이 계량적인 지표를 통해 측정될 수 있다.

과학기술콘텐츠 자산가치 평가모형(성태응, 전승표, 변정은, 박현우, 2016)은 데이터베이스, 지식정보, 특허정보 등 지식정보 자산의 경제적 가치를 평가할 수 있는 구조화된 평가 모델이다. 학술정보 측면에서 과학기술콘텐츠의 가치는 데이터베이스 관리 기준이나 인용지수에 따른 학술지, 학술정보의 가치를 측정하는 방식이었으나, 콘텐츠를 무형자산으로 간주하여 수명주기나 기여도, 가치평가 방법론에 따른 가치평가 모형을 제안하였다. 과학기술콘텐츠 가치평가의 핵심 변수로 콘텐츠 데이터베이스 자산의 수명주기, 데이터베이스 자산의 현금흐름 추정, 데이터베이스 자산의 할인율, 데이터베이스 자산의 기여도 등으로 정의하고 소득접근법 기반의 데이터베이스 자산 가치평가 모형을 제시하였다. 데이터베이스 자산 기여도 산출 방식에서 데이터베이스 자산 비중을 자산 기여도의 산출 인자로 포함하였는데 데이터베이스 자산비중 평가항목은 데이터베이스를 기반으로 지식정보의 가치를 정량적 지표와 정성적 지표로 구분하여 <표 6>과 같은 평가항목을 제시하였다.

<표 6> 과학기술콘텐츠 자산 비중 평가 항목(성태응, 전승표, 변정은, 박현우, 2016)

지표 구분	평가항목
정량적 지표	이용요금
	접근용이성
	활용(가능)성
	이용관심도
	유용성
	권위성
	정확성
	최신성
	포괄성
SW의존성	
정성적 지표	부가가치창출
	활용범위
	학술성
	자원투입량
	공익성
	객관성

콘텐츠 제공자 관점에서의 신규 콘텐츠 선정 지표 선정을 위해 본 연구에서는 공공데이터 품질관리 수준진단 및 평가 매뉴얼(한국지능정보사회진흥원, 2022)과 디지털 큐레이션 성숙도 모델(김성훈, 도슬기, 한상은, 김재훈, 임석중, 박진호, 2022)을 참고하였다.

행정안전부와 한국지능정보사회진흥원은 공공기관이 생성, 취득하여 관리하는 공공데이터의 품질을 확보하기 위해 체계적인 품질관리 활동을 수행하는지 여부를 진단하는 공공데이터 품질관리 수준진단 평가를 매년 실시하고 그 결과를 공표한다(한국지능정보사회진흥원, 2022). 진단 및 평가 대상기관은 중앙행정기관 45개, 광역자치단체 17개, 기초자치단체 226개, 공공기관 350개, 지방공기업 49개 등 687개 기관이며(2022년 기준) 등이며, A기관은 학술정보 제공 영역에서의 진단·평가 대상 기관이

다. 공공데이터 품질관리 수준진단 및 평가 항목은 데이터 관리체계와 데이터값 관리의 2개 영역 10개의 지표로 구성되어 있으며 그 내용은 <표 7>과 같다.

데이터 관리체계는 공공데이터 품질관리를 위해 기관의 품질관리 계획을 수립하고 이행하는 활동과 데이터베이스 구축 단계에서 준수해야 할 품질관리 활동으로 정의하고 있으며 품질관리 계획수립, 예방적 품질 관리진단, 데이터 표준 확산, 데이터 구조 안정화, 데이터 연계 관리 등으로 구성되어 있다. 데이터 값 관리는 각 기관이 구축 운영하고 있는 DB 및 개방데이터의 품질 제고를 위해 수행하는 품질 진단 및 개선 계획 수립, 오류 보완 조치 등 데이터 값의 품질 개선 활동으로 정의하고 있는데 데이터 품질 진단, 품질 진단 결과 조치, 데이터 값 오류율, 오류 신고 요구사항 분석 및 개선, 개방데이터셋 오류조치 등으로 구성되어 있다.

공공데이터 품질관리 수준진단·평가 지표는 공공기관의 데이터 관리 부분과 이용자

실제 데이터를 활용하는데 있어 고품질 데이터의 활용을 보장하기 위한 부분을 고려하여 구성된 것으로 파악된다. 디지털큐레이션 플랫폼 구축 방향성 수립에서 전반적인 학술정보·데이터의 관리체계에 대한 가이드라인이 마련되어 있으므로 품질관리계획이나 예방적품질관리진단 지표는 고려하지 않아도 될 것으로 판단되며 신규 학술 콘텐츠가 이용자에게 적시에 활용되어 가치 창출에 기여해야 한다는 목적성을 고려하면 신규 콘텐츠 수집시에 콘텐츠의 품질은 매우 중요한 요소라고 할 수 있다. 또한, 데이터 표준화, 구조화, 연계성 등이 신규콘텐츠 수집 우선순위에 있어 최우선적으로 고려해야 할 사항으로 판단된다.

디지털 큐레이션 성숙도 모델(김성훈, 도슬기, 한상은, 김재훈, 임석중, 박진호, 2022)은 A기관 디지털 큐레이션 플랫폼의 디지털 전환 성과를 측정할 수 있는 지표로 구성된 모델이다. 평가 척도를 기술, 데이터, 전략, 조직(인력), 영향력 등 5개 대분류를 중심으로 각 척도

<표 7> 공공데이터 품질관리 수준진단·평가 지표(한국지능정보사회진흥원, 2022)

영역	진단·평가지표	
데이터 관리 체계	품질관리 계획수립	기관 데이터 품질 관리 계획 수립 및 이행
	예방적 품질 관리 진단	예방적 품질관리 진단 및 진단결과 반영
	데이터 표준 확산	데이터 표준 정의 및 적용률
	데이터 구조 안정화	데이터 구조 정의 및 산출물 관리
	데이터 연계 관리	연계데이터 정의 및 송·수신 내역 관리
데이터 값 관리	데이터 품질 진단	품질진단 기준 정의 및 진단
	품질진단 결과 조치	품질진단에 따른 개선계획 수립, 개선조치
	데이터 값 오류율	데이터 값 최종 오류율
	오류 신고 요구사항 분석 및 개선	데이터 품질 오류 분석 및 문제점 개선
	개방데이터셋 오류 조치	개방데이터셋 및 오픈API 품질오류 조치결과
가점 지표	예방적품질관리 컨설팅 비정형 데이터 메타데이터 품질 진단 대상DB 초과 수행	

별로 측정 지표를 정의하였는데 <표 8>은 성숙도 모델의 지표 분류를 보인 것이다. 디지털 큐레이션 성숙도 모델은 데이터 큐레이션을 위한 업무 프로세스의 디지털 전환 수준을 측정하는 목적을 가지고 있으며 본 연구의 신규 콘텐츠 선정 지표와는 데이터 범주에서 데이터 품질과 데이터 관리와 관련된 지표와 연계성을 가질 수 있다. 따라서 데이터 품질 차원의 데이터 다양성, 적합성, 활용성, 적시성, 상호운용성, 보안, 유지보수성과 데이터 관리 차원의 준비성, 완전성, 유용성이 연관 지표로 고려될 수 있다.

4.3 신규 학술 콘텐츠 선정 지표

신규 학술 콘텐츠 선정 지표를 개발하기 위해 앞 절에서 도출하였던 지표 항목들을 대상으로 지표 풀(Indicator Pool)을 구축하고 불필요하거나 관계성이 낮은 항목은 삭제하였다. 또한, 지표간 공통성, 유사성에 따라 중복되는 측정항목 또는 표현방식은 다르나 유사한 설명을 갖는 측정항목들은 삭제 또는 통합하는 정제 과정을 거쳤다. 선행 연구 분석에서 다른 학술정보서비스 품질평가모델, 콘텐츠 자산가치평가 모형, 디지털 큐레이션 성숙도 모델 등은 그 목적과

<표 8> 디지털 큐레이션 성숙도 모델 최종 지표(김성훈, 도슬기, 한상은, 김재훈, 임석중, 박진호, 2022)

대분류	중분류	소분류
기술	연구개발	R&D투자 (연구개발과정에서 사용된 투입물 평가)
		기술개발역량 (신기술 기반으로 시스템을 개발하는지 여부)
	IT인프라	혁신역량 (신기술을 파악, 도입하기 위한 조직체계 및 역량)
데이터	IT인프라	보안 (조직 내외부 연계에 대한 보안 수준)
		네트워크 (기관의 네트워크 범위, 외부자원의 종류와 정도)
		규제 환경 (조직 업무 관련 법 제도의 이해 및 업무 적용)
	데이터 품질	다양성 (실제 세상의 데이터와의 유사특성의 데이터 확보)
		적합성 (데이터 유형과 활용 목적에 맞는 규격, 속성)
		활용성 (재활용성, 검색가능성, 유용성 등)
		적시성 (시기적으로 적절하고 유효한지의 여부)
		상호운용성 (메타데이터, 표준에 따른 상호운용 지원 수준)
		보안 (데이터 보안 수준의 유지)
		유지보수성 (데이터 변경사항에 대한 이력 관리)
데이터 관리	준비성 (데이터 관리 정책 및 지침)	
	완전성 (데이터 구축 공정 과정에서 데이터의 완전성)	
	유용성 (기관의 요구, 업무 정의에 적합성 여부)	
사회적 영향력	디지털 격차 해소 기여	모든 이용자를 위한 서비스를 수행하고 있는지 평가
	경제적 효과	국가과학기술발전, 특허 생성, 기술 수출에 이바지한 정도를 평가
	교육적 효과	교육자료 제공 및 연구자들의 평생교육에 이바지한 정도를 평가
	데이터 개방 정도	데이터 개방 정도에 대한 이바지 정도를 평가
	전반적인 만족도	제공하는 시스템의 UX/UI, 속도, 이용자 불편사항 제고 등의 정성적 만족도를 평가

적용 대상에 따라 지표가 상이한 부분이 있어 본 연구의 목적과 부합하여 콘텐츠 수집 및 구축 측면과 이용자 활용 목적 측면을 고려하여 관련성 있는 지표 중심으로 선별하였다.

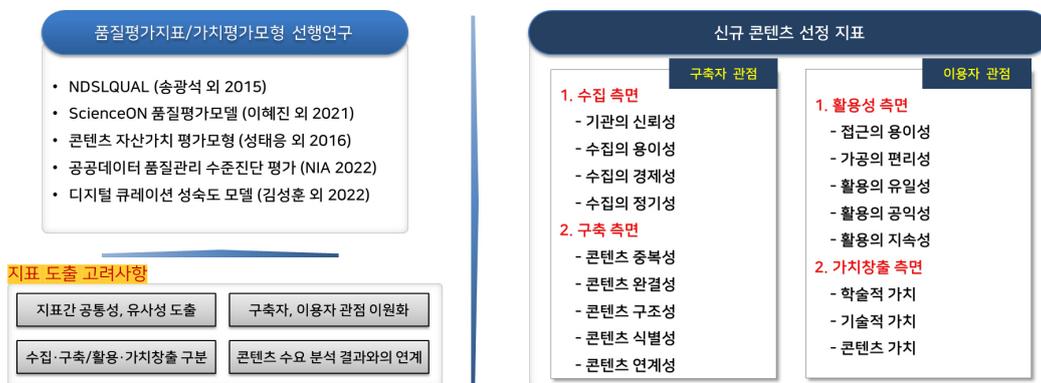
〈그림 2〉와 같이 콘텐츠 구축 관점에서의 지표는 디지털 큐레이션 성숙도 모델과 디지털큐레이션 플랫폼의 구성 요소를 최우선적으로 고려하였으며 이용자 관점에서의 지표는 NDSLQUAL, 과학기술 지식인프라 통합서비스 품질평가모델, 콘텐츠 자산가치 평가모델 등을 고려하여 〈표 9〉와 같이 수집, 구축, 활용성, 가치창출 등 4개 영역에 17개 세부지표로 구성된 신규 콘텐츠 선정 지표를 정의하였다.

수집 영역은 생산 기관의 신뢰성, 수집의 용이성, 수집의 경제성, 수집의 정기성 지표로 구성되어 있으며 구축 영역은 콘텐츠 중복성, 콘텐츠 완결성, 콘텐츠 구조성, 콘텐츠 식별성, 콘텐츠 호환성 지표로 구성되어 있다.

생산 기관의 신뢰성은 콘텐츠 생산 기관의 공신력과 신뢰도를 의미하며 정보·데이터 제공 기관이 국내외 공공기관, 민간기관, 학/협회, 생산기관 미상인지 여부를 판단하여 정보원의 신

뢰도를 지표화한 것이다. 수집의 용이성은 생산 기관의 협의에 따라 DB to DB, 파일 전송 프로토콜(FPT), API, 웹 크롤링 방식의 방식을 활용한 수집의 자동화 정도를 측정하는 지표이다. 수집의 경제성은 새로운 콘텐츠 수집시 발생하는 비용, 구매 계약 절차를 참고한 무료 여부를 판단하여 비용 효율성을 측정하는 지표이다. 수집의 정기성은 생산 기관의 정보·데이터 생산·갱신 및 제공 주기가 정기적인지 혹은 비정기적인지 여부에 따라 신규 콘텐츠의 적시 제공 여부를 판단할 수 있는 지표이다.

구축 영역의 콘텐츠 중복성은 A기관에서 기존에 수집하고 있는 콘텐츠인지 여부를 판단하여 중복 구축을 배제하기 위한 목적의 지표이며 완전중복, 부분중복, 신규 콘텐츠 등으로 구분하여 측정할 수 있다. 콘텐츠 완결성은 수집 콘텐츠 메타데이터가 모두 기술되어 있고 원문/본문이 포함되어 있는지에 따라 메타+본문/원문, 메타, 본문/원문, URI 수준 등의 완결성을 측정하는 지표이다. 콘텐츠 구조성은 수집 콘텐츠 메타데이터가 존재하고 콘텐츠 항목이 정의되어 있는지 등을 파악하여 콘텐츠의 확장



〈그림 2〉 신규 학술 콘텐츠 선정 지표 도출

〈표 9〉 디지털 큐레이션 플랫폼을 위한 신규 콘텐츠 선정 지표

영역	지표 항목	내용
수집	기관의 신뢰성	콘텐츠 제공 기관의 역할과 미션이 명확하고 지속가능한 기관인지 여부
	수집의 용이성	수집 프로토콜이 정의되어 있거나 협의를 통해 프로토콜 구현/적용이 가능한지 여부
	수집의 경제성	수집의 유/무료 여부, 구매나 계약 절차의 복잡성 등
	수집의 정기성	콘텐츠 생산 기관이 주기적으로 콘텐츠를 생산하는지 여부 정기적으로 수집이 가능한지 여부
구축	콘텐츠 중복성	기존에 수집·구축된 콘텐츠와 동일 혹은 유사한 유형의 콘텐츠인지 여부
	콘텐츠 완결성	콘텐츠 필수항목이 모두 기술되어 있고, 콘텐츠의 정확도, 신뢰성이 보장되는지 여부
	콘텐츠 구조성	콘텐츠 유형별로 적용 가능한 메타데이터가 존재하는지, 디지털큐레이션 플랫폼의 콘텐츠 구조와 호환/적용이 가능한지 여부
	콘텐츠 식별성	국제표준식별자 부여가 가능한지, 타 기관의 국제표준식별자를 적용할 수 있는지 여부
	콘텐츠 연계성	식별자 기반으로 타 콘텐츠 유형과의 연계가 가능한지 여부
활용성	이용자 요구도	이용자 요구도가 높은 콘텐츠 유형인지 여부
	접근의 용이성	획득한 콘텐츠를 변환하거나, 응용프로그램에서 호출하거나, 분석 단계에서 수월하게 활용 가능한지 여부
	활용의 유일성	활용하고자 하는 콘텐츠가 A기관 원내 서비스 사이트에서만 접근가능한 것인지, 다른 사이트에서 서비스 가능한 것인지의 여부
	활용의 공익성	지불 비용 없이 활용이 가능한지 여부, 공적인 목적으로 활용이 가능한지 여부
	활용의 지속성	콘텐츠를 연계, 공유하거나 타인이 재사용 가능한지 여부
가치창출	학술적 가치	수집/활용되는 콘텐츠가 논문, 보고서, 새로운 이론 창출 등에 기여할 수 있는 콘텐츠인지 여부
	기술적 가치	수집/활용되는 콘텐츠가 특허, 소프트웨어, 신기술 개발 등에 기여할 수 있는 콘텐츠인지 여부
	콘텐츠 가치	수집/활용 콘텐츠가 다른 콘텐츠와 연계 융합될 경우 더 큰 가치를 유발할 수 있는지 여부

성과 연계성을 판단하는 지표로서 표준메타데이터 포함, 콘텐츠 항목 정의, 비정형 형식 등의 구조화 여부 등으로 측정한다. 콘텐츠 식별성은 수집 콘텐츠가 DOI, ORCID, ISNI, ROR, GRID, ISSN, ISBN 등 표준 식별자가 존재하는지를 판단하여 콘텐츠의 호환성 보장 여부를 측정하는 지표이다. 콘텐츠 연계성은 수집 콘텐츠가 디지털 큐레이션 플랫폼에서 관리되고 있는 다른 콘텐츠와 의미적으로 연계가 가능한지를 측정하는 지표이다.

콘텐츠 활용성 영역은 이용자 요구도, 접근의 용이성, 활용의 유일성, 활용의 공익성, 활용

의 지속성으로 구성되어 있다. 이용자 요구도는 학술정보 서비스 이용자의 수요나 활용이 많은 콘텐츠 유형을 기준으로 이용자 요구도를 측정하거나 로그를 분석하여 산정할 수 있다. 접근의 용이성은 원내 서비스에서 검색이나 API 형태로 활용 가능한지, 직접 다운로드가 가능한지 여부에 따라 메타데이터와 원문의 완전 접근 가능, 메타 수준에서 접근 가능, 생산기관을 통한 접근 등의 기준으로 접근성 정도를 측정하는 지표이다. 또한, 이용자가 획득한 콘텐츠를 응용 프로그램에서 별도의 변환 없이 확인이 가능한지, 연구 목적에 따라 변환이 용

이한지 여부에 따라 완전 가공/호환 가능, 부분 가공/호환 가능, 별도 프로그램 설치 등의 기준으로 가공의 편리성 정도를 측정할 수 있다. 활용의 유일성은 이용자가 검색한 콘텐츠가 원내 서비스에서만 획득 가능한 것인지, 다른 기관에서 중복적으로 서비스하는 것인지 등을 측정하는 지표이다. 활용의 공익성은 콘텐츠 획득 비용의 유료/무료, 기관내 활용 가능 여부 등을 측정하여 활용에 제한이 없는지를 판단하는 지표이다. 활용의 지속성은 재사용 가능/불가능 여부를 측정하여 제약없이 콘텐츠 재사용이 가능한지를 판단하는 지표이다.

가치창출 영역 지표는 학술적 가치, 기술적 가치, 콘텐츠 가치 등으로 구성된다. 학술적 가치는 새로운 이론 창출, 학문 분야 발전 등에 기여할 수 있는 콘텐츠인지 여부, 기술적 가치는 특허, 소프트웨어, 신기술 개발 등에 기여할 수 있는 콘텐츠인지 여부, 콘텐츠 가치는 다른 콘텐츠와 융합하여 더 큰 가치를 유발할 수 있는지 여부를 판단하는 지표이다.

4.4 지표 가중치 설정 및 측정 방식 정의

〈표 9〉에서 정의한 신규 콘텐츠 선정 지표를 적용하는 과정에서 디지털 큐레이션 플랫폼의 성과목표에 따라 지표별 가중치가 달리 설정될 수 있다. 지표별 가중치는 시기별로 업무성과와 이용자 만족도 중 어느 영역에 비중을 두는지에 따라 동적으로 변할 수 있다고 할 수 있다. 이에 본 연구에서는 디지털 큐레이션 플랫폼 구축 로드맵의 방향성에 부합하여 영역별 지표 가중치를 부여한 사례를 제시하였다.

수집 영역에서 큐레이션 업무의 효율성을 고

려한다면 수집의 용이성, 수집의 경제성 지표가 우선적으로 고려될 수 있으며 이용자 만족도를 고려한다면 기관의 신뢰성, 수집의 정기성 지표가 가중치가 상대적으로 높다고 볼 수 있다. 디지털 큐레이션 플랫폼 구축 로드맵은 큐레이션 콘텐츠의 활용 확대를 성과지표로 관리하고 있는 점을 고려하여 기관의 신뢰성, 수집의 정기성 지표에 상대적으로 높은 가중치를 부여하였다.

구축 영역에서는 디지털 큐레이션 플랫폼의 성과목표가 신규 콘텐츠 개발 방향이 기존의 콘텐츠와 차별화된 새로운 유형의 콘텐츠 개발을 통한 학술 콘텐츠 스펙트럼 확장이 핵심 목표로 콘텐츠의 중복성, 콘텐츠 식별성, 콘텐츠 연계성이 콘텐츠 완결성이나 콘텐츠 구조성 보다 가중치가 높은 것으로 판단하였다. 반면, 활용성 영역에서는 정보서비스의 효율화, 콘텐츠 이용의 활성화 측면에서 이용자 요구도, 활용의 유일성, 활용의 공익성 지표가 접근의 용이성, 활용의 지속성 보다 가중치가 높은 것으로 판단하였다.

가치창출 영역은 학술적 가치, 기술적 가치, 콘텐츠 가치는 콘텐츠의 유형, 특성, 활용 목적 등에 따라 가치의 중요도가 상이할 수 있으며 이는 이용자 관점에서의 가치의 정도를 판단하는 지표이므로 본 연구에서는 별도로 가중치를 두지 않는 것으로 정의하였다.

〈표 10〉은 디지털 큐레이션 플랫폼 구축 로드맵을 고려한 신규 콘텐츠 선정 지표 가중치 부여 결과 및 지표 측정 방식과 점수 부여 기준을 정의한 것이다.

4.5 신규 콘텐츠 선정 지표 시범 적용

〈그림 3〉은 디지털 큐레이션 플랫폼에서 관

〈표 10〉 신규 콘텐츠 선정 지표 측정 방식 및 가중치 부여

영역	지표항목	가중치	측정 방식 및 점수 부여 기준
수집	기관의 신뢰성	30	- 정보원의 유형별 점수 부여 예) 정부/공공기관(10), 정보센터/도서관/출판사 등(10), 민간기관(8), 비영리기관/단체(6), 기관 미상(2)
	수집의 용이성	20	- 수집 프로토콜 유형 별로 점수 부여 예) API 연계(10), DB-to-DB(8), 파일 전송(6), 웹 크롤링(8), 직접 입력(4)
	수집의 경제성	20	- 수집 비용 유무에 따라 점수 부여 예) 무료 수집(10), 조건부 무료 수집(8), 유료 수집(6), 조건부 유료 수집(4)
	수집의 정기성	30	- 주기적 수집 가능 여부에 따라 점수 부여 예) 주기적 제공(10), 비정기적 제공(8), 일회성 제공(6)
구축	콘텐츠 중복성	25	- 기존 수집 대상 콘텐츠인지 여부에 따라 점수 부여 예) 기존콘텐츠 중복 없음(10), 중복비율 50% 미만(6), 중복비율 50%이상(3)
	콘텐츠 완결성	15	- 콘텐츠 메타+원문/본문 수집 모두 가능 여부, 항목의 완결성 여부에 따라 점수 부여 (평균 측정) 예) 메타+원문/본문 수집가능(10), 메타만 수집/원문본문 URL 제공(8), 메타만 수집(6) and 항목정의(10), 항목 미정의(6)
	콘텐츠 구조성	15	- 표준메타데이터 적용 가능 여부에 따라 점수 부여 예) 표준메타데이터 적용가능(10), 표준메타데이터 적용 불가능(4)
	콘텐츠 식별성	25	- 국제표준식별자 부여 및 호환 가능, 자체식별자 포함 여부 등에 따라 점수 부여 예) 국제표준식별자 부여 가능(10), 자체식별자 보유(8), 식별자 미보유(4)
	콘텐츠 연계성	20	- 식별자 기반으로 타 콘텐츠 유형과의 연계가 가능한지 여부에 따라 점수 부여 예) 학술정보/데이터/소프트웨어 등과의 연계 가능(10), 연계 불가능(2)
활용성	이용자 요구도	40	- 이용자의 콘텐츠 수요 조사 등을 통한 이용자 요구도 정도에 따라 점수 부여 예) 학술논문, 회의자료, 보고서, 동향, 데이터셋, 소프트웨어, 편집정보, 용어사전 등(10), 국내의 특허, 학위논문, 정책입안자료, 장비 등(4), 기타(7)
	접근의 용이성	10	- A기관 원내서비스를 통해 획득한 콘텐츠를 호환 가능한 응용 프로그램을 통해 열람이 용이하거나 데이터 변환이 가능한 정도에 따라 점수 부여 예) 별도의 응용프로그램 설치 없이 콘텐츠 열람/가공 가능(10), 범용 응용 프로그램을 활용하여 콘텐츠 열람/가공 가능(8), 특정 응용 프로그램 설치/활용하여 콘텐츠 열람/가공 가능(6), 상용 프로그램을 설치/활용하여 콘텐츠 열람/가공 가능(4)
	활용의 유일성	20	- 활용 콘텐츠가 A기관 원내 서비스 사이트에서만 접근 가능한 것인지, 다른 사이트에서 서비스 가능한 것인지의 여부에 따라 점수 부여 예) A기관 원내서비스에서만 접근 가능(10), A기관 원내서비스+콘텐츠 제공기관에서만 접근 가능(8), 복수의 서비스에서 접근 가능-포털제외(6), 포털 서비스에서 접근 가능(4)
	활용의 공익성	20	- 활용 콘텐츠가 무료/유료로 활용 가능한 것인지, 특정 IP에서만 활용 가능한 것인지에 따라 점수 부여 예) 장소제한 없이 무료 활용 가능(10), 특정 장소에서 무료 활용 가능(8), 장소제한 없이 유료 활용 가능(6), 특정 장소에서 유료 활용 가능(4)
	활용의 지속성	10	- 콘텐츠를 연계, 공유하거나 타인이 재사용 가능한지 여부에 따라 점수 부여 예) 제한없이 공유가능(10), 공유제한있음(4)
가치창출	학술적 가치	100	- 활용 콘텐츠가 논문, 보고서, 새로운 이론 창출 등에 기여할 수 있는 콘텐츠인지를 수집부서에서 정성적 판단에 의해 점수 부여
	기술적 가치		- 활용 콘텐츠가 특허, 소프트웨어, 신기술 개발 등에 기여할 수 있는 콘텐츠인지를 수집부서에서 정성적 판단에 의해 점수 부여
	콘텐츠 가치		- 활용 콘텐츠가 다른 콘텐츠와 연계 융합될 경우 더 큰 가치를 유발하는지 여부를 수집부서에서 정성적 판단에 의해 점수 부여

디지털 큐레이션 플랫폼을 위한 신규 학술 콘텐츠 선정 지표 개발에 관한 연구 211

기관	정보유형	수집						구축				활용				가치창출			합계	순위									
		기관의 신뢰성		수집의 용이성		수집의 경제성		콘텐츠 중복성		콘텐츠 완결성		콘텐츠 구조성		콘텐츠 식별성		콘텐츠 연계성		이용자 요구도			접근의 용이성	활용의 유일성	활용의 공공성	활용의 지속성	학술적 가치	기술적 가치	콘텐츠 가치		
		0.3	0.2	0.2	0.3	0.25	0.15	0.25	0.2	0.4	0.1	0.2	0.4	0.1	0.2	0.2	0.1	0.2			0.2	0.1	0.2	0.1	0.2	0.1	0.2	0.1	
한국특허전략개발원	특허정보	10	8	6	8	3	8	10	10	10	4	4	2	8	10	6	8	8	6	6	6	6	6	6	6	7.2875	39		
	분석보고서	10	8	8	6	10	8	8	6	8	10	6	8	8	10	6	8	8	10	6	6	6	6	6	6	7.7	19		
	통계자료	10	8	6	6	10	6	6	10	6	7	4	8	8	10	6	8	10	6	6	8	10	6	6	8	7.75	14		
한국기초과학지원연구원	연구시설장비	10	8	6	8	3	6	8	8	6	4	6	8	8	8	6	6	4	6	6	6	6	6	6	4	6.6125	62		
	학술논문	10	8	6	8	3	4	8	8	6	7	8	6	8	8	6	6	4	6	6	6	6	6	6	4	6.7875	59		
한국산업기술진흥원	분석보고서	10	8	6	6	6	8	8	8	6	10	4	6	8	4	6	6	6	6	6	6	6	6	6	6	6.4125	44		
	기술동향정보	10	8	8	8	6	8	8	8	6	6	6	8	8	8	8	6	6	6	6	6	6	6	6	6	7.175	41		
	산업기술정보	10	8	6	6	6	6	8	8	8	6	8	8	6	8	8	6	4	6	6	6	6	6	6	6	7.025	47		
한국화학연구원	신물질	10	6	6	6	3	8	8	8	6	7	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6.4125	63		
	화합물	10	6	6	6	3	8	6	8	6	7	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6.4125	63		
	연구보고서	10	8	8	10	10	8	8	6	8	10	6	8	6	8	6	8	6	6	6	6	6	6	6	6	7.85	10		
	동영상강의자료	10	8	4	4	10	6	6	6	4	8	7	6	6	10	4	4	4	4	4	4	4	4	4	4	6	6.625	61	
한국생명공학연구원	생장자원	10	8	8	8	6	6	6	6	8	6	7	4	6	6	6	6	6	6	6	6	6	6	6	6	6.825	56		
	학술논문	10	8	8	8	6	8	8	8	6	7	6	6	8	6	8	6	8	8	8	8	8	8	8	8	7.625	21		
정보통신산업진흥원	소프트웨어	10	6	6	6	8	8	8	6	8	8	10	4	4	6	8	4	8	4	8	4	8	4	8	4	7.525	23		
한국표준과학연구원	표준/규격	10	6	6	6	8	8	8	8	8	4	6	4	8	4	6	8	4	6	8	4	6	8	4	6	7.05	45		
AI-Hub	연구데이터	10	6	6	6	8	8	6	6	6	10	8	6	10	8	6	8	6	8	6	8	6	6	6	6	7.7	17		
통합데이터지도	연구데이터	10	6	6	6	8	8	6	6	6	10	8	6	10	8	6	8	6	8	6	8	6	8	6	6	7.7	17		
	소프트웨어	10	6	6	6	8	6	6	6	6	10	6	6	6	6	6	6	6	6	6	6	6	6	6	6	7.325	35		
국회도서관	연구데이터	10	8	8	8	4	6	8	8	8	8	8	6	4	6	8	4	8	4	8	4	8	4	8	4	7.475	30		
	연구자정보	10	8	8	8	6	6	8	8	8	8	8	8	8	8	8	8	6	6	6	6	6	6	6	6	7.4	31		
	연구기관	10	8	8	8	6	6	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	7.525	26		
국립중앙도서관	인명전기	10	8	8	8	4	8	8	8	8	8	8	8	8	8	8	8	6	4	6	6	6	6	6	6	7.35	33		
	단체명전기	10	8	8	8	8	6	8	8	8	8	8	8	8	8	8	8	6	6	6	6	6	6	6	6	7.525	26		
	주제명전기	10	8	8	8	8	6	8	8	8	8	8	8	8	8	8	8	6	6	6	6	6	6	6	6	7.525	26		
한국교육학술정보원	학술논문	10	6	6	6	8	8	6	8	8	3	4	4	6	8	4	8	4	8	4	8	4	8	4	8	4	6.825	58	
	통계데이터	10	8	6	6	10	6	6	10	6	7	4	8	8	10	6	8	10	6	6	6	6	6	6	6	7.75	14		
	연구데이터	10	6	6	6	8	8	6	8	8	10	4	4	6	8	4	8	4	8	4	8	4	8	4	8	4	7.525	23	
기계로봇연구정보센터	기술동향정보	10	8	8	8	10	8	8	8	6	10	8	6	8	8	6	8	8	6	6	6	6	6	6	6	4	7.775	13	
	연구데이터	10	8	8	8	10	8	8	8	8	10	6	6	8	8	8	6	6	8	8	6	6	6	6	6	6	7.825	11	
	용어사전	10	8	8	8	8	8	8	8	8	10	6	8	8	8	8	6	6	6	6	6	6	6	6	6	6	7.8	12	
	학술논문	10	8	8	8	8	8	8	8	8	7	8	8	8	8	6	4	6	6	6	6	6	6	6	6	6	7.5	29	
	연구자정보	10	8	8	8	8	8	8	8	8	10	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	7.55	22	
생물학연구정보센터	기술동향정보	10	8	8	8	8	8	8	8	8	10	8	8	8	6	6	6	6	6	6	6	6	6	6	6	6	7.7	16	
	연구자정보	10	8	8	10	8	8	8	8	8	10	6	8	8	8	6	6	6	6	6	6	6	6	6	6	6	7.95	7	
	학술논문	10	8	8	10	8	8	8	8	8	10	6	6	6	8	8	8	8	8	8	8	8	8	8	8	8	8.25	2	
의과학연구정보센터	생명자원	10	8	8	8	6	6	6	8	6	7	4	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6.825	56	
	논문	10	8	8	8	8	8	8	8	8	8	6	6	8	8	6	8	6	8	8	8	8	8	8	8	8	7.95	6	
전자정보연구정보센터	데이터	10	6	6	6	8	8	6	8	8	10	4	4	6	8	4	8	4	8	4	8	4	8	4	8	4	7.525	23	
	논문	10	6	6	6	8	8	8	8	8	7	6	4	8	4	6	8	4	6	8	4	6	8	4	6	8	7.35	33	
한의학융합연구정보센터	데이터	10	8	4	4	6	6	6	6	6	10	6	6	8	8	6	8	8	6	8	6	8	6	8	6	6	7.2	40	
	논문	10	8	4	4	6	6	6	6	6	7	6	6	8	8	6	8	8	6	8	6	8	6	8	6	6	6.9	53	
	뉴스	10	8	4	4	8	6	6	6	6	7	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6.875	54	
	동향	10	8	8	8	4	6	8	8	8	10	6	4	6	8	6	6	6	6	6	6	6	6	6	6	6	7.175	41	
화학연구정보센터	연구자	10	8	8	8	4	6	4	8	6	10	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6.975	52	
	동향	10	8	8	8	8	8	8	8	8	10	8	8	8	8	8	6	6	6	6	6	6	6	6	6	6	7.85	9	
화학공학소재연구정보센터	동향	10	8	8	8	10	6	8	8	8	10	8	8	8	8	8	6	6	6	6	6	6	6	6	6	6	7.9	8	
국가바이오데이터스테이션	데이터	10	8	8	8	8	10	8	8	8	8	8	8	8	8	8	6	4	8	6	4	8	6	4	8	6	8.175	4	
국립의과학지식센터	논문	10	8	8	10	10	6	8	8	8	7	8	8	8	8	6	6	6	6	6	6	6	6	6	6	6	7.7	19	
	의학검색엔진	8	8	8	8	10	6	8	8	8	6	8	8	8	8	6	8	8	8	6	6	6	6	6	6	6	7.3	36	
	IP용어사전	8	6	6	6	8	8	6	8	8	6	4	4	6	8	4	8	4	8	4	8	4	8	4	8	4	6.975	48	
	정보통신용어사전	10	10	6	6	8	6	6	10	8	10	8	8	8	8	10	6	6	8	6	6	6	6	6	6	6	8.225	3	
	보안용어사전	8	6	6	6	8	8	6	8	8	6	4	4	6	8	4	8	4	8	4	8	4	8	4	8	4	6.975	48	
	IT용어사전	8	8	8	8	6	6	8	8	8	6	8	8	8	6	8	8	6	6	6	6	6	6	6	6	6	7.05	45	
	온라인건축용어사전	8	8	8	8	8	6	8	8	8	6	8	8	8	8	8	6	6	6	6	6	6	6	6	6	6	7.175	41	
	산림임업용어사전	8	8	6	6	6	8	8	8	8	6	8	8	8	8	8	6	4	6	6	6	6	6	6	6	6	6.875	54	
	식품용어사전	8	8	8	8	10	6	8	8	8	6	8	8	8	8	8	6	6	6	6	6	6	6	6	6	6	7.3	36	
	항공기상용어사전	8	8	8	8	10	6	8	8	8	8	6	8	8	8	8	6	6	6	6	6	6	6	6	6	6	7.3	36	
	물리학용어집	8	6	6	6	8	8	6	8	8	6	4	4	6	8	4	8	4	8	4	8	4	8	4	8	4	6.975	48	
	의학용어	10	8	8	8	8	6	6	10	6	10	8	8	8	8	10	6	6	8	10	6	6	6	6	6	6	8.275	1	
	과학백과사전	8	6	6	6	8	8	6	8	8	6	4	4	6	8	4	6	8	4	8	4	8	4	8	4	8	4	6.975	48
	대기환경용어사전	8	8	8	8	8	6	6	8	6	6	4	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6.7	60	
네이버지식백과	10	4	8	8	8	8	8	8	8	10	6	8	8	8	8	6	8	8	8	8	8								

리하고 있지 않은 신규 콘텐츠 영역을 발굴하고 해당 콘텐츠에 대해 4.3절과 4.4절에서 설정한 신규 콘텐츠 선정 지표를 시범 적용한 도구를 보인 것이다. 신규 콘텐츠 선정 지표 도구는 디지털 큐레이션 플랫폼 운영 및 관리자가 활용하며 발굴한 신규 콘텐츠 정보원과 정보 유형, 학술콘텐츠 범주에 따라 지표별로 점수를 측정하고 지표별 가중치를 곱하여 수집, 구축, 활용성, 가치 창출 영역별 점수 합이 평균값으로 지표값을 산출하고 우선순위를 도출한다.

신규 콘텐츠 선정 지표를 시범 적용하기 위해 우선 이석형, 이강산다정, 김재훈, 이혜진(2024)의 연구의 학술정보 및 데이터 요구 분석 결과를 토대로 연구자가 연구 개발 과정에서 많이 활용하는 학술논문, 연구보고서, 회의자료와 부족하다고 느끼는 데이터셋, 강연자료 및 발표자료, 용어사전 등을 대상으로 정보원을 발굴하였다.

이용자 요구도에 따라 현재 A기관 디지털 큐레이션 플랫폼에 탑재되어 있지 않은 콘텐츠 정보원은 한국기초과학지원연구원, 한국화학연구원, 한국생명공학연구원 등 정부출연연구기관, 생물학연구정보센터, 기계로봇연구정보센터 등 주요 분야 전문연구정보플랫폼, 국립중앙도서관, 국회도서관, KERIS, 한국특허전략개발원, 네이버 등이며 각 기관별로 학술논문, 용어사전, 데이터셋, 학습모델 등을 발굴하였다(〈그림 3〉의 1,2,3행). 이들 신규 콘텐츠별로 수집, 구축, 활용, 가치창출 측면에서 〈표 9〉와 〈표 10〉의 측정 방법 및 가중치에 따라 점수화한 결과 의학용어사전, 생물학연구정보센터의 학술논문, 정보통신용어사전, 국가바이오데이터스테이션의 연구데이터, 네이버 지식백과, 의과학연구

정보센터의 논문, 생물학연구정보센터의 연구자 정보, 화학공학소재연구정보센터와 화학연구정보센터의 동향정보, 한국화학연구원의 연구보고서 순으로 우선순위가 높았다.

이러한 결과는 이석형, 이강산다정, 김재훈, 이혜진(2024)의 이용자 설문 결과 학술논문과 연구보고서의 활용도가 높다는 점과 협업 연구에 필요한 연구자 정보, 실험에 필요한 데이터에 대한 요구도가 높다는 연구 내용과 부합한다고 볼 수 있다. 단, 〈그림 3〉의 우선 순위 중 의학용어사전, 정보통신용어사전, 네이버 지식백과 등 용어사전의 우선순위가 상대적으로 높은 이유는 이용자의 요구도도 높지만 디지털 큐레이션 플랫폼에서 용어사전 구축이 이루어지지 않아 중복성 이슈가 없다는 점, 용어사전을 기반으로 주제 식별 및 콘텐츠 연계의 용이성 측면에서의 점수가 높기 때문으로 판단된다.

5. 결론 및 제언

본 연구에서는 연구자의 연구개발 활동에 필요한 다양한 학술 콘텐츠를 지원하기 위한 신규 콘텐츠 선정 지표를 개발하고, A기관의 디지털 큐레이션 플랫폼을 대상으로 신규 콘텐츠 선정 우선 순위를 시범 도출하였다. 연구의 목적을 달성하기 위해 선행 연구 분석을 통해 정보·데이터의 가치평가 모델, 정보서비스 품질평가모델 등을 참고하여 콘텐츠의 가치, 이용자의 요구도 및 만족도 지표 등 이용자 관점에서의 신규 콘텐츠 선정 지표를 정의하였으며, 디지털 큐레이션 성숙도 모델, 데이터 품질관

리 모델 등을 참고하여 콘텐츠 수집, 구축, 관리의 용이성과 효율성을 고려한 콘텐츠 제공자 관점에서의 콘텐츠 선정 지표를 정의하였다.

신규 학술 콘텐츠 선정 지표는 수집, 구축, 활용성, 가치창출 등 4개 영역에 17개 세부지표로 구성되었다. 수집 영역은 생산 기관의 신뢰성, 수집의 용이성, 수집의 경제성, 수집의 정기성 지표로 구성하였으며, 구축 영역은 콘텐츠 중복성, 콘텐츠 완결성, 콘텐츠 구조성, 콘텐츠 식별성, 콘텐츠 호환성 지표로 구성하였다. 콘텐츠 활용성 영역은 이용자 요구도, 접근의 용이성, 활용의 유일성, 활용의 공익성, 활용의 지속성으로 구성하였으며, 가치 창출 영역 지표는 학술적 가치, 기술적 가치, 콘텐츠 가치 등으로 구성하였다.

주제별, 분야별로 다양한 유형의 학술 콘텐츠 수집을 위해서는 제한된 물적, 인적 자원을 고려해야 하는데 신규 콘텐츠 선정 지표를 적

용하여 이러한 자원을 효과적으로 사용할 수 있다. 또한 가치 있고 신뢰성 있는 학술 콘텐츠를 우선적으로 수집, 구축할 수 있는 기회를 제공할 수 있어 장기적으로 이용자의 만족도를 높이고 콘텐츠 활용성을 높일 수 있다. 또한, 플랫폼에 탑재된 콘텐츠와 중복되지 않은 새로운 콘텐츠를 식별할 수 있어 디지털 큐레이션 플랫폼 운영의 효율성을 유지하는데 도움이 될 것으로 판단된다.

본 연구에서는 디지털 큐레이션 플랫폼의 성과목표와 로드맵을 고려하여 신규 콘텐츠 선정 지표의 가중치를 부여한 사례를 제시하고 점수 측정 기준을 정의하였다. 향후 연구로는 큐레이션 활동에 따른 업무 성과목표에 따라 가중치를 변경 관리하는 방안과 그 결과의 타당성을 살펴보고, 신규 콘텐츠 선정 지표 적용에 따른 디지털 큐레이션 활동의 효과성을 검증할 예정이다.

참 고 문 헌

- 과학기술정책연구원 (2020). 데이터 기반 연구개발 관리 혁신 방안. 서울: 과학기술정책연구원.
- 권나현, 이정연, 정은경 (2012). 과학기술분야 R&D 전주기 연구: 국내 생명 및 나노과학기술연구자를 중심으로. 한국문헌정보학회지, 46(3), 103-131. <https://doi.org/10.4275/KSLIS.2012.46.3.103>
- 김성훈, 도슬기, 한상은, 김재훈, 임석중, 박진호 (2022). 디지털 큐레이션 성숙도 모델 및 지표 개발에 관한 연구: KISTI 디지털큐레이션센터를 중심으로. 정보관리학회지, 39(4), 269-296. <https://doi.org/10.3743/KOSIM.2022.39.4.269>
- 김성훈, 박진호 (2023). AHP 기법 기반 디지털 큐레이션 성숙도 모델 · 지표 가중치 연구: KISTI 디지털 큐레이션 성숙도 모델을 중심으로. 정보관리학회지, 40(2), 243-262. <https://doi.org/10.3743/KOSIM.2023.40.2.243>
- 김주섭, 전예린, 김선태 (2021). 학술 콘텐츠 종합링킹체제 구축에 관한 연구: 한국과학기술정보연구원

- 을 중심으로. 한국문헌정보학회지, 55(1), 493-519.
<https://doi.org/10.4275/KSLIS.2021.55.1.493>
- 성태웅, 전승표, 변정은, 박현우 (2016). 과학기술정보 콘텐츠의 가치평가모형 프레임워크 연구. 한국콘텐츠학회논문지, 16(11), 421-433. <https://doi.org/10.5392/JKCA.2016.16.11.421>
- 송광석, 유한주, 김원중, 김종훈 (2015). 이용자 기반 NDSL정보서비스품질측정 모형개발 연구. 서비스경영학회지, 16(1), 99-123. <https://doi.org/10.15706/jksms.2015.16.1.005>
- 유승훈, 허재용, 안윤기 (2009). 지식정보 가치평가 모형의 실증적 개발. Journal of Information Science Theory and Practice, 40(1), 113-132.
- 이석형, 이강산다정, 김재훈, 이혜진 (2024). 데이터 기반 R&D 지원을 위한 연구자의 학술정보 및 데이터 요구 분석 연구. 한국문헌정보학회지, 58(1), 255-283.
<https://doi.org/10.4275/KSLIS.2024.58.1.255>
- 이선희, 김혜선 (2014). 국가과학기술정보센터(NDSL)의 논문 콘텐츠 가치평가에 관한 연구. 한국콘텐츠학회 2014 추계종합학술대회, 389-390.
- 이혜원, 윤소영, 박지영, 황혜경, 김재훈, 이혜림 (2019). 연구콘텐츠라이프 사이클 모델 개발: 한국과학기술정보연구원 콘텐츠 큐레이션 센터를 중심으로. 정보관리학회지, 36(3), 203-228.
<https://doi.org/10.3743/KOSIM.2019.36.3.203>
- 이혜진, 신재민, 배수영, 최희석, 이석형 (2021). 과학기술 지식인프라 통합서비스의 품질평가모델 연구. 디지털콘텐츠학회논문지, 22(5), 867-876. <https://doi.org/10.9728/dcs.2021.22.5.867>
- 장덕현 (2007). 과학기술분야 정보자원개발정책에 관한 연구. 한국비블리아학회지, 18(2), 107-124.
- (주)케이티지 (2002). 과학기술정보 수요조사 및 정보가치지도 작성연구. 대전: 한국과학기술정보연구원.
- 한국과학기술정보연구원 (2022). 지능형 과학기술정보 큐레이션 체제 구축. 대전: 한국과학기술정보연구원.
- 한국과학기술정보연구원 (2023). 지능형 과학기술정보 큐레이션 체제 구축. 대전: 한국과학기술정보연구원.
- 한국교육학술정보원 (2004). 학술정보서비스 이용 행태 분석을 통한 서비스 개선 방안 연구. 서울: 한국교육학술정보원.
- 한국지능정보사회진흥원 (2022). 2022년 공공데이터 품질관리 수준진단 평가 매뉴얼. 대구: 한국지능정보사회진흥원.

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

Chang, Durkhyun (2007). A study on the collection development policy for science and technology.

- Journal of the Korean Biblia Society for Library and Information Science, 18(2), 107-124.
- Kim, Ju-Seop, Jeon, Ye-rin, & Kim, Seuntae (2021). A study on the establishment of the comprehensive academic contents linking system: focusing on KISTI. Journal of the Korean Society for Library and Information Science, 55(1), 493-519.
<https://doi.org/10.4275/KSLIS.2021.55.1.493>
- Kim, Seonghun & Park, Jinho (2023). A study on weighting for digital maturity model & indicators based on AHP technique: focusing on digital curation maturity model of Korea institute of science and technology information. Journal of the Korean Society for Information Management, 40(2), 243-262. <https://doi.org/10.3743/KOSIM.2023.40.2.243>
- Kim, Seonghun, Do, Seulki, Han, Sangeun, Kim, Jayhoon, Lim, Seokjong, & Park, Jinho (2022). A study on development of digital curation maturity models and indicators: focusing on KISTI. Journal of the Korean Society for Information Management, 39(4), 269-296.
<https://doi.org/10.3743/KOSIM.2022.39.4.269>
- Korea Education and Research Information Service (2004). A Study on Service Improvement Plans through Analysis of Academic Information Service Usage Behavior. Seoul: Korea Education and Research Information Service.
- Korea Institute of Science and Technology Information (2022). Construction of an Intelligent Science and Technology Information Curation System. Daejeon: Korea Institute of Science and Technology Information.
- Korea Institute of Science and Technology Information (2023). Construction of an Intelligent Science and Technology Information Curation System. Daejeon: Korea Institute of Science and Technology Information.
- KTG Co., Ltd. (2002). Research on Science and Technology Information Demand Survey and Information Value Map Preparation. Daejeon: Korea Institute of Science and Technology Information.
- Kwon, Na-Hyun, Lee, Jung-Yeon, & Jung, Eun-Kyung (2012). Full cycle research in science and technology R&D - focusing on domestic life and nanoscience technology researchers. Journal of the Korean Society for Library and Information Science, 46(3), 103-131.
<https://doi.org/10.4275/KSLIS.2012.46.3.103>
- Lee, Hyejin, Shin, Jaemin, Bae, Suyeong, Choi, Heeseok, & Lee, Seokhyoung (2021). A study on the quality evaluation model for S&T knowledge infrastructure integrated service. Journal of the Digital Content Society, 22(5), 867-876.
<https://doi.org/10.9728/dcs.2021.22.5.867>

- Lee, Hyewon, Yoon, Soyoung, Park, Ziyong, Hwang, Hyekeyong, Kim, Jayhoon, & Rhee, Hea Lim (2019). Development of a research content lifecycle model: focused on the content curation center of the Korea institute of science and technology information. *Journal of the Korean Society for Information Management*, 36(3), 203-228. <https://doi.org/10.3743/KOSIM.2019.36.3.203>
- Lee, Seok-Hyoung, Lee, KangSan-DaJung, Kim, Jay-Hoon, & Lee, Hye-Jin (2024). A study on the scholarly information and data requirements of researchers for data-driven research and development. *Journal of the Korean Society for Library and Information Science*, 58(1), 255-283. <https://doi.org/10.4275/KSLIS.2024.58.1.255>
- Lee, Sun-Hee & Kim, Hye-Sun (2014). A study on the value evaluation of the paper contents of the national science and technology information center (NDSL). *Proceedings of the 2014 Fall Conference of the Korea Contents Association*, 389-390.
- National Information Society Agency (2022). *2022 Public Data Quality Management Level Assessment Manual*. Daegu: National Information Society Agency.
- Science and Technology Policy Institute (2020). *Data-Driven Innovation Strategies for R&D Management*. Seoul: Science and Technology Policy Institute.
- Seong, Taeung, Jun, Seungpyo, Byun, Jeongeun, & Park, Hyunwoo (2016). The study on frameworks of valuation models for the contents of science and technology information. *Journal of Korea Content Association*, 16(11), 421-433. <https://doi.org/10.5392/JKCA.2016.16.11.421>
- Song, Kwang-Seok, Yoo, Han-Joo, Kim, Wan-Jong, & Kim, Jong-Hoon (2015). Development of a user-based model for measuring the quality of NDSL information services. *Journal of Service Management*, 16(1), 99-123. <https://doi.org/10.15706/jksms.2015.16.1.005>
- Yoo, Seung-Hoon, Heo, Jae-Yong, & Ahn, Yoon-Ki (2009). Empirical development of a knowledge information value evaluation model. *Journal of Information Science Theory and Practice*, 40(1), 113-132.