

논문 2026-1-9 <http://dx.doi.org/10.29056/jsf.2026.03.09>

연구데이터 가치평가 지표 개발에 관한 연구: 국가연구개발사업 적용 사례

경태원*, 장효경**†

A Study on the Development of Research Data Valuation Indicators: Application Cases to National Research and Development Projects

Taewon Kyung*, Hyokyung Chang**†

요약

본 연구는 연구데이터의 가치를 “활용 가능성(utility)” 관점에서 평가하기 위한 지표 체계를 개발하고, 국가연구개발사업 사례를 통해 적용 가능성을 검증하였다. 객관성(재현 가능한 정량 측정), 확장성(다분야 재사용·유통 가능성), 실용성(R&D 성과관리 연동)의 3대 원칙을 기반으로 메타데이터 완성도, 품질 정합성, 최신성/시의성, 윤리·법적 적정성, 재사용 가능성, 표준화 준수도, 공개성·접근성, 인프라 연계성, 활용 기여도, 학술적 독창성의 10개 지표를 제안하였고, 각 지표를 점검 항목·증빙자료·점수 규칙으로 운영화하여 재현성을 확보하였다. NTIS-NRDP 연계 환경에서 “감염제어 컨버전스 연구센터” 사례에 적용한 결과, 표준 기반 문서화와 품질 관리, 갱신 이력, IRB 및 법 준수 등 객관성 근거가 명확했으며, CC 라이선스·다중 포맷·DOI·REST API 등은 확장성 증빙으로 확인되었다. 또한 논문·특허·기술이전과의 연계 가능성을 통해 연구데이터가 성과관리의 핵심 자산이 될 수 있음을 시사했으며, 최신성 해석, 표준 태그·코드북 보완, 접근 조건의 투명화 등 거버넌스 개선 과제도 도출되었다.

Abstract

This study develops a utility-oriented framework for valuing research data and validates its applicability through a national R&D case study. Based on three principles - objectivity (reproducible quantitative measures), scalability (cross-disciplinary reuse and dissemination), and practicality (linkage to R&D performance management) - we propose ten valuation indicators: metadata completeness, quality consistency, timeliness, ethical/legal compliance, reusability, standardization compliance, openness/accessibility, infrastructure interoperability, contribution to utilization, and scholarly originality. Each indicator is operationalized with check items, required evidence, and scoring rules to ensure reproducibility. The framework was applied in an NTIS-NRDP environment using the “Infection Control Convergence Research Center” case, scoring indicators based on evidence such as standard metadata, licensing, DOI/API availability, usage statistics, and links to outputs (papers, patents, technology transfer). Results showed strong evidence for standards-based documentation, quality control, update history, and compliance, while CC licensing, multiple formats, DOI, and REST APIs supported scalability. The study suggests research data can function as a core performance-management asset and highlights governance needs regarding recency interpretation, standard tags/codebooks, and transparent access conditions.

한글키워드 : 연구데이터, 연구데이터 가치평가, 가치평가 지표체계, FAIR 원칙, 데이터 거버넌스

keywords : Research Data, Research Data Valuation, Valuation Indicator Framework, FAIR Principles, Data Governance

* 한국특허전략개발원 국가전략기술특허지원실 접수일자: 2026.02.27. 심사완료: 2026.03.13.

** 한남대학교 컴퓨터공학과 게재확정: 2026.03.20.

† 교신저자: 장효경(email: chantellejang@hnu.kr)

1. 서론

오늘날 연구데이터는 과학기술 혁신과 경제·사회 발전의 핵심 자산으로 부상하고 있다. 데이터 중심 혁신(Data-driven innovation)은 생산성 향상과 신산업 창출의 원동력으로 평가되며, “데이터와 정보는 21세기의 가장 중요한 자산”이라는 인식이 확산되고 있다[1]. 특히 공공 재원으로 생산된 연구데이터의 개방과 공유는 투자 대비 성과(Return on Investment)를 극대화하고 과학적 발견을 가속하기 위한 필수 조건으로 여겨진다. OECD는 “데이터의 가치는 활용에 있다. 공공연구로 생성된 과학데이터에 대한 완전하고 자유로운 접근이 국제적 규범이 되어야 한다”고 권고하며 개방을 촉진하였다[2]. 실제로 개방된 연구데이터는 연구 투명성과 재현성, 학제간 협력 증진, 중복 연구 비용 절감 등 다양한 사회·과학적 편익을 창출하고 있다[3].

정책적 맥락에서도 주요국들은 연구데이터 가치 극대화를 위한 개방 정책을 적극 추진하고 있다. 유럽연합(EU)은 FAIR(Findable, Accessible, Interoperable, Reusable) 데이터 원칙을 오픈사이언스의 핵심 기조로 삼아, 연구데이터를 찾기 쉽고, 접근 가능하며, 상호운용되고, 재사용 가능하도록 함으로써 데이터 활용 가치를 높이고자 한다[14]. 미국 역시 2023년부터 NSF와 NIH 등 주요 연구비 지원기관들이 데이터관리계획(DMP)과 데이터 공유를 의무화하며, 연구데이터를 공개·재사용하는 문화가 강화되고 있다[4]. 대한민국도 2018년 「혁신성장 촉진을 위한 연구데이터 공유·활용 전략」을 수립하여 범부처 차원의 연구데이터 관리·개방 체계 구축을 추진하였다[5]. 이처럼 각국 정부와 국제기구들은 연구데이터의 잠재가치를 인식하고, 데이터 가치 평가와 개방을 제도적으로 뒷받침하는 노력을 강화하고 있다.

한편, COVID-19 팬데믹은 연구데이터 공유의 가치를 단기간에 가시화한 대표 사례로 평가된다. 전례 없는 규모로 연구데이터가 신속하게 공개·공유되면서, 치료제·진단기술 개발 과정에서 연구 속도가 크게 단축되었기 때문이다[6]. 예를 들어, 전 세계 과학자들은 바이러스 유전체 서열 데이터를 공개 데이터베이스(GISAID 등)에 즉시 공유하였고, 이를 기반으로 국제사회는 단기간 내 표준 진단법을 마련할 수 있었다[6]. 이러한 경험은 “연구데이터의 가치는 활용될 때 발현된다”는 원칙을 실증적으로 보여주며, 위기 대응과 과학 발전에서 데이터 가치 평가와 공유의 중요성을 재확인시켰다[2].

본 연구는 연구데이터의 가치를 “단순한 양적 측정”이 아니라 “어떤 형태로 활용 가능한가”라는 관점에서 평가할 수 있는 지표 체계를 제안하고, 국가연구개발사업 사례에 적용하여 실증적으로 검증하는 것을 목적으로 한다. 이를 위해 다음의 연구문제를 설정하였다.

- (RQ1) 연구데이터의 특수성(공공성·재현성·재사용성)을 반영하면서도 재현 가능한 평가가 가능한 가치평가 지표 체계는 어떻게 구성될 수 있는가?

- (RQ2) 제안 지표 체계는 NTIS - NRDP 연계 환경에서 실제 국가연구개발과제 데이터에 적용될 때, 어느 수준의 실행 가능성과 설명력을 보이는가?

본 연구의 기여는 다음과 같다. 첫째, 객관성·확장성·실용성의 3대 설계 원칙을 기반으로 10개 지표와 점검 항목을 운영 가능한 형태로 구체화하였다. 둘째, 사례 적용을 통해 지표 체계가 단순 선언적 원칙이 아니라 정량·정성 통합형 평가 도구로 작동할 수 있음을 확인하였다. 셋째, 평가 과정에서 도출되는 취약 지점을 통해 연구데이터 거버넌스(갱신·표준·접근정책)의 개선 과제를 진단할 수 있는 실무적 함의를 제시하였다.

2. 연구데이터 가치 평가에 관한 선행 연구 및 프레임워크

연구데이터의 가치를 평가(Valuation)하는 방법에 대해서는 다양한 선행 연구와 프레임워크 제안이 존재한다. 전통적으로 데이터의 가치는 기업이나 기관에서 경제적 자산(Asset)으로 간주되어 왔으며, 경영학 분야에서는 무형자산 평가 기법을 응용한 가치평가 모델들이 개발되었다. 예를 들어 수익(이익) 접근법을 활용하여 데이터 자산이 창출하는 미래 현금흐름이나 비용절감 효과를 추정하는 연구들이 다수 이루어졌다. 선행 연구들은 데이터 자산이 공유 가능성이 높고 활용 기간이 길며, 데이터 품질(적시성, 정확성 등)에 따라 가치가 크게 좌우된다는 점을 강조한다. 따라서 데이터 가치평가에는 데이터 품질, 사용자의 활용 목적과 능력, 데이터 가치사슬 전반의 요소가 총체적으로 반영되어야 한다는 논의가 축적되어 왔다[7].

한편 학술 연구데이터의 가치를 평가하는 데에는 정량적 지표와 정성적 평가를 결합한 다양한 시도가 이루어지고 있다. 데이터 활용 빈도나 데이터 인용은 대표적인 정량 지표로 활용되며, Belter(2014) 등은 데이터 인용 분석을 통해 데이터 공유의 과학적 영향을 계량화하였다[8, 9]. 또한 Tu(2021)는 오픈리서치데이터 가치평가를 위해 접근성, 품질, 재사용성, 영향력 등을 통합한 모델을 제안한 바 있다[10]. 이러한 선행 연구들은 연구데이터 가치가 단일 차원의 지표로 포착되기 어렵고, 데이터 생애주기 전반(생산 - 관리 - 공개 - 재사용)에서 복합적으로 나타난다는 점을 공통적으로 시사한다.

데이터 관리 및 공유로 인한 가치를 평가한 사례 연구들도 주목할 만하다. Beagrie와 Houghton(2014)은 영국의 주요 연구데이터센터에 대상으로 데이터 공유·보존 투자 대비 편익을 정량·정

성으로 분석하여, 데이터 인프라 투자 타당성을 실증적으로 제시하였다[12]. 이처럼 정량·정성 평가를 결합한 선행 연구들은 연구데이터 공유로 인한 과학적 생산성 증대와 경제적 효용을 보여 줌으로써 연구데이터 가치평가의 필요성을 뒷받침하고 있으며, 향후 지표체계 고도화를 위한 근거로 활용되고 있다.

3. 일반데이터 vs. 연구데이터 가치 평가 기준의 차이

일반데이터(상업/산업 데이터 등)와 연구데이터는 성격과 활용 목적이 다른 만큼 가치 평가의 기준과 관점에도 차이가 존재한다. 일반데이터는 시장 거래를 전제로 가치가 논의되는 경우가 많다. 기업이 보유한 고객 데이터나 위치 정보 데이터 등의 가치는 해당 데이터를 활용해 창출할 수익이나 절감할 비용으로 산정되며, 거래 시장에서 형성되는 가격(교환가치)에 의해 평가받기도 한다. 대한민국도 「데이터산업진흥 및 이용촉진에 관한 기본법」 기반으로 데이터 가치평가제도를 도입하면서, 데이터의 경제적 가치를 금액·등급·점수로 평가하는 가이드라인을 제시하였다[13]. 이 지침에서 데이터 가치는 “데이터의 생산·유통·활용에서 발생하는 경제적 효익의 측정치”로 정의되며, 사용 가치(value in use)에 중점을 둔다[13].

이에 비해 연구데이터는 상업적 거래보다는 학술적 활용 가치와 공익적 가치가 중시된다. 즉, 연구데이터의 가치는 당장의 금전적 수익보다 해당 데이터가 새로운 지식과 발견 창출에 얼마나 기여하는지(학술적 영향력, 사회적 편익)로 나타난다. 또한 연구데이터는 맥락 정보(Context)와 측정·분석 방법의 투명성, 재현 가능성 등 과학적 타당성이 핵심 품질 요소로 포함된다. 아울러 연구데이터는 ‘상품’이면서도 ‘성과물’이라는 이중적 성격을 지니므로, 일반 데이터에 적용되는 거래·

표 1. 일반데이터와 연구데이터 가치 평가 기준
Table 1. Valuation Standards for General and Research Data

항목	일반데이터	연구데이터	차이점/특징
생성 목적	상업적 유통, 통계	지식 창출, 실험/관찰 기반	연구 목적 중심, 유일성 높음
품질 지표	정합성률, 결측률 등	메타데이터, 측정방법 명확성, 재현성	기술적 품질 + 과학적 타당성 필요
활용도	타 데이터와 결합	연구 인용, 모델 훈련 등	논문·특허 기여도 등 과학적 활용도 중심
공개성/접근성	API, 포맷 중심	공유 조건, 익명화 여부, 재사용 조건	법·윤리적 제한 고려 필수
데이터 생애주기	지속적 업데이트	특정 시점의 결과값	일회성·정적인 경우 다수
경제성 평가	거래가능성, 수익환산	사회적 파급효과, 학술 기여도	수익보다는 공공적 기여 강조

가격 중심의 평가만으로는 충분히 설명되기 어렵다. 따라서 연구데이터의 특수성을 반영한 전용 가치평가 지표 체계 구축이 요구된다.

정리하면, 일반데이터 평가는 경제적 교환가치와 시장성을 중심으로 이루어지는 반면, 연구데이터 평가는 학술적·사회적 활용 가치, 과학적 품질, 개방성과 재사용성, 윤리·법적 적정성을 포괄적으로 고려하는 것이 특징이다[12, 13]. 궁극적으로 중요한 것은 “데이터가 어떤 목적으로 누구에게 얼마나 유용하게 쓰이는가”이며[13], 연구데이터의 경우 그 유용성이 지식과 공익의 형태로 발현된다는 점이 평가 기준의 핵심 차이라고 할 수 있다.

4. 연구데이터 가치 평가 지표 체계 및 프레임워크 사례

국제적으로 연구데이터의 관리와 가치를 높이기 위해 다양한 지표 체계와 프레임워크가 제시되어 왔다. 본 연구에서는 유럽(EU), 미국(USA), 일본(Japan), 한국(Korea)의 대표적 사례를 검토하였다.

4.1 EU의 FAIR 데이터 원칙과 오픈사이언스 프레임워크

EU는 오픈사이언스 정책의 일환으로 FAIR 데

이터 원칙을 대표 지침으로 채택하였다. 2016년 발표된 FAIR 원칙은 연구데이터가 “찾기 쉽고(Findable), 접근 가능하며(Accessible), 상호운용 가능하고(Interoperable), 재사용 가능(Reusable)”해야 함을 강조한다[14]. EU는 Horizon 2020 이후 데이터관리계획(DMP) 제출을 요구하고, 가능한 범위 내에서 연구데이터 공개를 촉진하면서 FAIR 원칙 준수를 핵심으로 삼았다. 또한 유럽 오픈사이언스 클라우드(EOSC) 구축을 통해 회원국 연구데이터의 연계·공유를 촉진하고 있으며, 이는 FAIR 구현을 위한 범유럽 인프라로 기능한다[1].

4.2 미국 NIST의 연구데이터 프레임워크 (RDaF)

미국은 연구데이터의 체계적 관리와 가치 제고를 위해 NIST를 중심으로 연구데이터 프레임워크(RDaF)를 제안하였다[1]. RDaF는 연구데이터의 생애주기 전반을 관리·평가하기 위한 요소를 구조화하여, 기관이 데이터 관리 성숙도와 가치 향상 전략을 점검할 수 있도록 지원한다. 또한 오픈사이언스 정책 기조 강화(예: 연방정부 지원 연구 성과물 데이터 공개 확대)와 맞물려, 데이터 공개·재사용을 촉진하는 운영 기반으로 활용될 수 있다[1].

4.3 일본 NII의 연구데이터 클라우드 및 정책 프레임워크

일본은 국립정보학연구소(NII)를 중심으로 연구데이터 인프라 구축과 정책 정비를 추진하고 있다. 대표 사례인 NII 연구데이터 클라우드(NII RDC)는 연구데이터 생애주기에 대응하여 연구데이터 관리(GakuNin RDM), 데이터 공개·출판(WEKO3), 검색·발견(CiNii Research) 플랫폼을 통합 제공한다[15]. 또한 데이터 거버넌스, 데이터 이력(Provenance) 추적, 보안 저장 및 연계 분석 등 부가 기능을 강화하여 데이터 신뢰성과 활용성을 제고하고 있다[15, 16].

4.4 한국의 NTIS 연계 연구데이터 관리·평가 체계

한국은 국가과학기술지식정보서비스(NTIS)를 중심으로 국가 R&D 성과 정보를 통합 관리해 왔으며[17], 최근에는 한국과학기술정보연구원(KISTI)의 국가연구데이터플랫폼(DataON/NRDP)과 연계하여 연구데이터 검색·공유 서비스를 강화하고 있다[18]. 이 과정에서 연구데이터의 등록·연계 기준과 메타데이터 체계 정비가 추진되었고, 과제 정보와 관련 연구데이터를 통합 탐색할 수 있는 기반이 확장되었다. 또한 정부 차원의 오픈사이언스 추진전략을 바탕으로, 연구데이터 공개·활용 실적을 성과관리와 연계하려는 정책적 방향성이 강화되고 있다. 이는 국가 차원에서 연구데이터를 공동 자산으로 인식하고 가치를 극대화하기 위한 제도·인프라를 정비해 가는 사례로 해석할 수 있다.

위 사례들은 공통적으로 (1) 표준 메타데이터와 식별자(예: DOI), (2) 접근·공개 정책과 권리(라이선스), (3) 상호운용성(API·포맷), (4) 재사

용 촉진(문서화·품질·거버넌스)을 핵심 축으로 삼고 있으며, 이는 본 연구의 지표 설계 원칙과도 정합적으로 연결된다[1, 14, 15].

5. 연구 데이터 가치 평가 모형 설계

5.1. 설계 원칙

연구데이터의 가치평가는 단순한 양적 축적이 아니라, 해당 데이터가 어떤 형태로 활용 가능한지를 기준으로 측정할 필요가 있다. 본 연구는 선행연구 검토를 토대로 연구데이터 가치평가 설계를 위한 원칙을 다음과 같이 수립하였다.

- 객관성: 재현 가능성 기반 계량 가능한 정량 데이터 측정 가능
- 확장성: 다양한 연구분야에 적용 가능한 과학적 활용성
- 실용성: 기존 R&D 성과관리 체계와 연동 가능한 학문적 기여도

5.2. 가치 평가 항목 및 지표 체계

연구데이터 가치 평가 항목 수립을 위해 앞서 제시한 세 가지 원칙을 바탕으로 표 2와 같이 10개 지표를 설계하였다[5, 7, 9, 10, 12, 13, 14].

객관성 관련 지표(1~4)는 수치 또는 체크리스트 기반으로 비교적 재현 가능한 평가가 가능하며, 확장성 관련 지표(5~8)는 과학적 활용 범위와 범용성을 반영한다. 실용성 관련 지표(9~10)는 성과관리 및 후속 기획과의 연계를 고려한다. 다만 일부 지표는 복수 원칙에 동시에 기여할 수 있으므로(예: 표준화 준수도는 객관성과 확장성 모두에 기여), 본 분류는 운영상 편의를 위한 기준으로 해석하는 것이 타당하다.

표 2. 연구데이터 가치 평가 지표
Table 2. Evaluation Metrics for Research Data Value

구분	평가지표	지표설명	점검항목
객관성	1. 메타데이터 완성도	변수 정의·단위·측정방법 등 기술의 충실성	변수 설명, 측정 단위, 장비·환경 정보, 메타데이터 표준 준수
	2. 품질 정합성	오류·결측·중복 최소화 및 논리 일관성	결측률, 오류율, 중복률, 논리조건 위반
	3. 최신성/시의성	생성·갱신 시점의 유효성	생성일 경과, 업데이트일, 수집 주기, 최신 분석기술 반영
	4. 윤리·법적 적정성	개인정보·민감정보 보호 및 법·윤리 준수	익명화, IRB, 관련 법(개인정보보호법, PIPL 등) 준수
확장성	5. 재사용 가능성	제3자 재분석·활용 가능성	포맷 개방성, 익명화, 접근 조건, 활용 사례
	6. 표준화 준수도	표준 구조·분류체계 준수	DCAT, ISO 19115 등, 변수 코드화, 시맨틱 표기
	7. 공개성 및 접근성	접근 가능성과 이용 조건 명확성	라이선스(CCL), 접근 제약 수준, 포맷(CSV/JSON 등)
	8. 인프라 연계성	타 시스템·플랫폼 연동 가능성	API, DOI, 외부 서비스(Jupyter 등) 연계
실용성	9. 활용 기여도	논문·특허·기술이전 등 성과 기여	논문 인용, 특허, 기술이전, 공동연구
	10. 학술적 독창성	차별성·희소성·신규성	신규 변수, 신규 실험법, 최초 수집 희소 데이터

5.3. 지표 도출 및 운영화 절차

본 지표 체계가 선언적 수준에 그치지 않고 실제 평가에 적용되기 위해서는, 각 지표를 “점검 항목 - 증빙자료 - 점수 산정 규칙”으로 운영화하는 과정이 필수적이다. 본 연구는 (1) 선행연구·정책·프레임워크 검토를 통해 후보 요소를 도출하고, (2) 객관성·확장성·실용성의 3대 원칙에 따라 10개 지표로 구조화한 뒤, (3) 지표별 점검 항목과 증빙(메타데이터 필드/라이선스/DOI/API/로그/성과 식별자 등)을 매핑하였다. 또한 (4) 10점 만점의 부분 점수 구조를 적용하되, 존재 여부(O/X) 항목은 배점을 분할하여 합산하도록 설계하였다.

아울러 본 연구는 자동화(정량) - 전문가 정성평가의 결합을 운영 원칙으로 채택하였다. 즉, 메타데이터 충족률, DOI/API/라이선스 존재 여부, 업데이트 이력, 제공 포맷 등 플랫폼 기반으로 확인 가능한 항목은 자동 또는 체크리스트 기반(반자동)으로 1차 산정하고, 학술적 독창성(지표 10), 표

준화 준수도의 ‘일부 포함’ 판단(지표 6), 최신성/시의성의 맥락 판단(지표 3 일부) 등 해석이 필요한 항목은 증빙과 함께 ‘근거 문장(rationale)’을 작성하여 2차 평가한다. 정성 판단이 개입되는 항목은 최소 2인 독립평가 후 불일치 항목 합의(또는 전문가 검토)로 최종 점수를 확정하여 일관성을 확보하였다.

6. 연구데이터 평가 지표를 이용한 분석 사례

본 연구에서는 국가연구개발과제에 대해 수립한 10개 지표를 적용하여 평가를 수행함으로써 지표의 적절성과 적용 가능성을 검증하였다. 평가는 NTIS - NRDP 연계 환경에서 확인 가능한 자료 수집·가공을 바탕으로 수행하였다.

6.1. 평가 절차 및 점수 산정 방식

평가는 “증빙 기반 평가”를 원칙으로 하며, 각

지표는 10점 만점(총 100점)으로 산정한다. 구체적인 절차는 다음과 같다.

- ① 자료 수집: 과제 기본정보(NTIS)와 데이터 등록·메타데이터(NRDP/DataON) 및 성과 연계 정보(논문·특허·기술이전)를 수집한다.
- ② 증빙 정리: 지표별 점검 항목에 대응하는 증빙(메타데이터 필드, 라이선스 표기, API 제공 여부, DOI 부여, 다운로드 로그, 성과 식별자 등)을 기록한다.
- ③ 점수 부여: O/X형 항목은 사전 정의된 배점으로 분할 합산하고(예: 3점+3점+4점), 연속형·구간형 항목(다운로드 횟수 등)은 구간 기준에 따라 점수를 부여한다.
- ④ 검토 및 일관성 확보: 해석 편차가 발생할 수 있는 항목(예: 학술적 독창성)은 최소 2인 평가 또는 전문가 검토를 통해 판단 근거를 문장으로 남긴다.
- ⑤ 산출물화: 최종 점수와 함께 “증빙 위치/식별자”를 병기하여 감사 가능성과 재현성을 확보한다(표 3 개정안 참조).

6.2. 사례 분석

사례 과제로는 2017년 6월부터 2024년 2월까지(총 81개월) 수행된 국가연구개발과제 “감염제어 컨버전스 연구센터(과제고유번호: 1711182622)”를 선정하였다. 본 과제는 연구데이터의 생산·관리·공개·성과 연계가 비교적 명확히 확인되는 특성을 갖고 있어, 지표의 실증 적용과 평가 절차 검증에 적합한 사례로 판단하였다.

6.3. 평가 결과 및 해석

분석 결과는 표 3과 같다.

위 사례 적용 결과, 본 지표 체계는 계량 가능한 정량지표 중심 구조를 통해 데이터 가치의 객

관적 평가를 가능하게 하였고, 동시에 정성 판단이 필요한 항목까지 “증빙과 근거 문장”을 병기함으로써 실행 가능성을 확보하였다. 특히 FAIR 원칙에 대응하는 요소(DCAT, DOI, API, 다중포맷, 라이선스)가 실제 평가에서 탐색 가능성·접근성·상호운용성·재사용성의 핵심 근거로 작동함을 확인하였다[14]. 또한 활용 기여도(논문·특허·기술이전)의 확인은 연구데이터가 보조 산출물이 아니라 R&D 성과관리의 핵심 구성요소로 편입될 수 있음을 시사한다.

다만 표준화 준수도(지표 6)는 “DCAT 및 ISO 19115 준수(O)”와 같이 준수 여부뿐 아니라, 코드북/시맨틱 태그/변수 코드화의 적용 범위(커버리지)가 점수에 직접 영향을 미친다. 따라서 본 연구는 지표 6을 평가할 때 (1) 표준 항목 충족 체크리스트(DCAT/ISO19115 매핑표), (2) 코드북(변수명·정의·단위·코드값) 파일, (3) 시맨틱 태그/변수 코드 목록을 증빙으로 수집하였다. 특히 “일부 포함”으로 표기되는 경우에는 전체 변수 수 대비 코드북/태그 적용 변수 수(가능 시 N/M)를 함께 제시하고, 누락 범위와 개선 필요성을 근거 문장으로 기록하여 사례 서술과 개선 과제를 연결하였다.

6.4. 거버넌스 개선 과제의 사례 기반 도출 및 증빙 연결

본 사례에서 제기된 거버넌스 개선 과제(최신성 해석, 표준 태그·코드북, 접근 조건 투명화)는 지표 평가 과정에서 확인된 분석 내용과 증빙의 취약지점에 근거한다. 즉, 가치평가는 점수 산출을 넘어 데이터 운영 체계의 개선 필요 영역을 진단하는 기능을 함께 수행한다.

첫째, 최신성/시의성(지표 3)은 생성·업데이트·수집주기는 확인 가능하나 “시의성” 해석이 데이터 성격과 사용 목적에 따라 달라질 수 있으므로, 데이터 최신성과 메타데이터 최신성을 분리

표 3. 연구데이터 가치 평가 예시
Table 3. Examples of Research Data Value Evaluation

평가지표	분석내용	점수예시	증빙
1. 메타데이터 완성도	변수 설명서(O), 단위(O), 측정 장비·방법(O), 시간·위치 기록(O), DCAT 기반 메타데이터 제공	각 2점 + 표준 준수 2점 → 10/10	NRDP 메타데이터 화면 캡처/필드 목록, DCAT 적용 여부 표시
2. 품질 정합성	결측률 2%, 오류율 1%, 중복 없음, 전수 검증 도구 통과	10/10	QC 리포트/검증 로그, 결측·오류 산출 근거
3. 최신성/시의성	데이터 생성 2020년, 2024년 3월 업데이트, 1개월 단위 수집 주기	업데이트 4점 + 주기 3점 + 시의성 3점 → 10/10	업데이트 이력, 수집 주기 정책/로그
4. 윤리·법적 적정성	개인정보 포함 없음(익명화), IRB 승인(O), 관련 법 준수	익명화 3점 + IRB 3점 + 법 준수 4점 → 10/10	IRB 번호, 비식별 처리 안내, 접근·제공 동의 근거
5. 재사용 가능성	총 150회 다운로드, 협력 연구자 재사용 사례, 접근 조건·포맷 개방성 확인	포맷 2점 + 익명화 2점 + 재사용(150회) 3점 + 사례 문서 3점 → 10/10	다운로드 통계, 활용 사례 문서/보고서
6. 표준화 준수도	DCAT 및 ISO 19115 준수(O), 코드북/시맨틱 태그 일부 포함	표준 준수 6/7 + 시맨틱 태그 2/3 → 8/10	표준 항목 충족 체크리스트, 코드북/태그 목록
7. 공개성 및 접근성	CC-BY 4.0, 로그인 기반 인증, CSV/JSON/XML 다중 포맷 제공	라이선스 3점 + 인증 3점 + 다포맷 4점 → 10/10	라이선스 표기 화면, 접근정책(인증 방식), 제공 포맷 목록
8. 인프라 연계성	REST API 제공(O), DOI 부여(O), R·Python 활용 가능	API 4점 + DOI 3점 + 외부 연동 3점 → 10/10	API 문서/엔드포인트, DOI 식별자, 예제 코드/연계 안내
9. 활용 기여도	논문 2편(Scopus/NTIS 등재), 특허 1건, 기술이전 1건	논문 4점 + 특허 3점 + 기술이전 3점 → 10/10	논문 DOI/식별자, 특허 번호, 기술이전 계약 근거(공개 가능 범위 내)
10. 학술적 독창성	신규 변수(미세항원 농도 등), 회소성 있는 현장·정밀 시계열 데이터	신규 변수 5점 + 회소성 5점 → 10/10	기존 데이터 대비 차별 설명(비교표), 신규성·회소성 근거 문장

하고 유형별 판단 기준을 명문화할 필요가 있다. 둘째, 표준화 준수도(지표 6)는 준수(O/X)뿐 아니라 코드북·시맨틱 태그·변수 코드화의 커버리지도 핵심이므로, 체크리스트·코드북·태그/코드 목록을 증빙으로 제시하고 전 변수 확대 및 부여 규칙 표준화를 추진해야 한다. 셋째, 공개성·접근성(지표 7)과 윤리·법적 적정성(지표 4)은 제한이 있더라도 접근 수준, 승인 기준·절차·소요기간을 구조화해 투명성을 높여야 한다. 종합하면, 가치평가 지표는 점수 산출을 넘어 운영 품질 진단과 거버넌스 개선 과제 도출에 활용될 수 있다.

7. 결론 및 시사점

본 연구는 연구데이터의 가치를 “양적 축적”이 아닌 “활용 가능성” 관점에서 평가하기 위해 객관성(재현 가능한 정량 측정)·확장성(다분야 적용 가능한 활용성)·실용성(기존 R&D 성과관리와의 연동)의 3대 원칙을 수립하고, 이를 기반으로 10개 가치평가 지표 체계를 제안하였다. 제안 지표는 메타데이터 완성도, 품질 정합성, 최신성/시의성, 윤리·법적 적정성, 재사용 가능성, 표준화 준수도, 공개성 및 접근성, 인프라 연계성, 활용 기여도, 학술적 독창성으로 구성되며, 점검 항목을 체크리스트와 정량 지표로 구체화하여 평가의 재현성과 실행 가능성을 확보하고자 하였다.

또한 NTIS - NRDP 기반 자료를 활용해 “감염

제어 컨버전스 연구센터” 과제를 사례로 평가한 결과, DCAT 기반 메타데이터 제공, 결측·오류·중복 관리, 정기적 업데이트, IRB 및 법 준수 등 객관성 지표에서 높은 수준이 확인되었다. 더불어 CC-BY 4.0 및 다중 포맷 제공, REST API·DOI 연계, 논문·특허·기술이전 등 성과 연동도 증빙 기반으로 확인되었다. 이는 본 지표 체계가 국가 R&D 데이터 환경에서 정량·정성 통합형 평가 도구로 작동할 수 있음을 시사한다.

정책적·실무적 시사점은 다음과 같다. 첫째, 본 지표 체계는 FAIR 원칙과 정합적으로 설계되어 DOI 부여, DCAT 메타데이터 준수, API 연계, 다중 포맷 제공 등의 요소가 실제 평가에서 탐색가능성·접근성·상호운용성·재사용성의 핵심 증빙으로 기능함을 확인하였다. 이는 향후 국가 차원의 연구데이터 공유 생태계를 강화할 때, “표준·연계·개방”을 권고 수준이 아니라 정량 평가요소로 내재화하는 것이 효과적임을 의미한다. 둘째, 활용 기여도(논문·특허·기술이전 등)가 데이터와 연계되어 확인된다는 점은 연구데이터가 보조 산출물이 아니라 R&D 성과관리의 핵심 구성요소로 편입될 수 있음을 시사한다. 가치평가 결과를 성과평가 및 인센티브 체계와 연동할 경우 데이터 기반 연구문화 확산에 기여할 수 있다. 셋째, 운영 관점에서는 최신성 관리, 표준 태그·코드북의 누락, 접근 제한 조건의 투명성 등 보완 필요성이 확인되었으며, 이는 지표 적용이 곧바로 데이터 거버넌스 개선 과제를 도출하는 진단 도구로 활용될 수 있음을 보여준다.

마지막으로 본 연구의 한계와 향후 과제는 다음과 같다. 단일 사례 기반 평가는 분야별 데이터 특성(민감정보 포함 여부, 생성 주기, 표준 적용 난이도 등)에 따른 편차를 충분히 반영하기 어렵다. 후속 연구에서는 다양한 분야·유형의 과제를 대상으로 표본을 확대하고, 지표 간 가중치(예: 윤리·법 준수의 중요도) 설정, 자동화 가능

한 평가항목(메타데이터 충족률, DOI/API/라이선스 존재 여부 등)과 전문가 정성평가가 필요한 항목(학술적 독창성 등)의 결합 방식을 고도화할 필요가 있다. 그럼에도 본 연구가 제시한 10개 지표와 점검 항목은 NTIS - NRDP 연계 환경에서 연구데이터 가치평가를 제도화하고 데이터 기반 성과관리 체계로 확장하기 위한 실질적 기반을 제공한다.

참 고 문 헌

- [1] Robert J. Hanisch, Debra L. Kaiser, Bonnie C. Carroll, Research Data Framework (RDaF): Motivation, Development, and A Preliminary Framework Core, National Institute of Standards and Technology, Feb., 2021. <https://doi.org/10.6028/NIST.SP.1500-18>
- [2] Dirk Pilat and Yukiko Fukasaku, OECD Principles and Guidelines for Access to Research Data from Public Funding, Data Science Journal, Volume 6, Open Data Issue, 17 June 2007. <https://www.oecd.org/en/publications>
- [3] Open Science - Enabling Discovery in the Digital Age, OECD Going Digital Toolkit, 2021. <https://www.oecd.org/en/publications>
- [4] Sarah M. Nusser, The Role of Statistics in Promoting Data Reusability and Research Transparency, Annual Review of Statistics and Its Application, Vol. 10:145-164, March, 2023. <https://doi.org/10.1146/annurev-statistics-033121-105114>
- [5] Promotion of research data sharing and utilization strategies, KOCEC Newsletter, Vol. 29, July, 2018. https://www.koced.or.kr/newsletter/v29_180712/html/spc4.html
- [6] Eric R G R Aguiar, Jesús Navas, Luis G C Pacheco, The COVID-19 Diagnostic Technology Landscape: Efficient Data Sharing Drives Diagnostic Development, Front Public Health. 2020 Jun 18;8:309.

- <https://doi.org/10.3389/fpubh.2020.00309>
- [7] Okki Kim, Cheon Woong Park, Wan-Sup Cho, and Jung Park, "Data Asset Valuation Model Review", The Korea Journal of BigData, vol. 6, no. 1, pp. 153 - 160, Aug. 2021.
DOI : 10.36498/kbigdt.2021.6.1.153
- [8] Grace C. Donovan and Madison L. Langseth, Are Researchers Citing Their Data? A Case Study from The U.S. Geological Survey, Data Science Journal, Vol. 23, Sep. 2024.
DOI: 10.5334/dsj-2024-024
- [9] Belter, C. W., Measuring the value of research data: A citation analysis of oceanographic data sets. PLoS ONE, 9(3): e92590. 2014. DOI:
<https://doi.org/10.1371/journal.pone.0092590>
- [10] Zhifang Tu, A Synthesized Model for Evaluating the Value of Open Research Data, iConference 2021 Poster Descriptions, March 2021.
<http://hdl.handle.net/2142/10968>
- [11] Bertin Martens, Data Access, Consumer Interests and Social Welfare: An Economic Perspective, SSRN Electronic Journal, Jan. 2020. DOI:10.2139/ssrn.3605383
- [12] Beagrie, N. and Houghton J.W., The Value and Impact of Data Sharing and Curation: A synthesis of three recent studies of UK research data centres, March 2014.
https://repository.jisc.ac.uk/5568/1/iDF308_-_Digital_Infrastructure_Directions_Report%2C_Jan14_v1-04.pdf
- [13] Yoon & Yang, Ministry of Science and ICT Publishes Data Valuation Guide, Jul. 2024. https://www.hwawoo.com/newsletter/2024_07_10_240710_k_t.pdf
- [14] Mark D Wilkinson et al., The FAIR Guiding Principles for scientific data management and stewardship, Sci Data, Vol. 6, No. 6, Mar 2016. doi: 10.1038/sdata.2016.18
- [15] Research Center for Open Science and Data Platform (RCOS), National Institute of Informatics, Overview of the NII Research Data Cloud,
<https://rcos.nii.ac.jp/en/service/#~:text=The%20NII%20Research%20Data%20Cloud,started%20its%20operation%20in%202021>
- [16] NII RCOS, NII Research Data Policy and Institutional Data Governance, 2023. https://www.nii.ac.jp/en/upload/NII_Overview2023_en.pdf
- [17] National Science & Technology Information Service(NTIS), <https://ntis.go.kr>
- [18] Gwangnam Choe, Construction of linkage system between NTIS and Research Data, KISTI, 2020.
12.<https://scienceon.kisti.re.kr/srch/selectPORSrchReport.do?cn=TRKO202100005913>

저자 소개



경태원(Taewon Kyung)

1998.2 호원대학교 전자계산학과 졸업
2002.2 경희대학교 전자계산공학과 석사
2008.8 경희대학교 산업공학과 박사
2008.8-2010.7 한국생산기술연구원
2010.7-현재: 한국특허전략개발원 전문위원
<주관심분야> 특허 빅데이터 분석, IP-R&D, 트리즈, 프로젝트관리



장효경(Hyokyung Chang)

1991.2 한남대학교 전자계산공학과 졸업
2012.2 한남대학교 컴퓨터공학과 박사
2021.3-현재: 한남대학교 조교수
<주관심분야> 데이터베이스, 데이터 사이언스, 유비쿼터스 컴퓨팅, 노코드