

# 데이터 리포지터리 인증 체계 분석 및 인증 전략에 관한 연구\*

- Coretrustseal을 중심으로 -

## A Study on Data Repository Certification Scheme Analysis and Certification Strategy: Focused on Coretrustseal

김 주섭 (Juseop Kim)\*\*

양성준 (Seong J. Yang)\*\*\*

김선태 (Suntae Kim)\*\*\*\*

### 목 차

- |                    |                 |
|--------------------|-----------------|
| 1. 서 론             | 4. CTS 인증 추진 전략 |
| 2. 이론적 배경          | 5. 결 론          |
| 3. 데이터 리포지터리 인증 체계 |                 |

### 초 록

이 연구의 목적은 데이터 리포지터리 인증 체계를 분석하고 Coretrustseal 인증을 추진 전략을 제안하는 것이다. 분석 대상 인증 체계에는 Coretrustseal, RatSWD, CLARIN certificate B, DIN 31644, ISO 16363, TDR 그리고 nestor 등이 포함된다. 그 중에서 Coretrustseal 인증은 타 인증의 기준 및 요구사항을 충족하고 있는 핵심 인증 체계이다. 따라서 본 연구에서는 Coretrustseal 인증 추진 전략을 단계별로 제안하였다. 제안된 인증 추진 전략에는 국내 데이터 리포지터리를 운영하는 기관에서 활용할 수 있도록 연구데이터 정책 및 가이드라인 관련 목록과 인증 신청서 예시가 포함되어 있다. 이러한 CTS 인증 추진 과정을 통해 각 기관에서는 연구데이터 정책 및 가이드라인을 확보할 수 있으며 CTS 인증을 획득하게 되면 해당 데이터 리포지터리에 대한 국제적인 신뢰성을 확보할 수 있을 것으로 판단된다.

### ABSTRACT

The purpose of this study is to analyze the data repository certification scheme and to propose a strategy to promote Coretrustseal certification. The certification system to be analyzed includes Coretrustseal, RatSWD, CLARIN certificate B, DIN 31644, ISO 16363, TDR and nestor. Among them, Coretrustseal certification is a core certification system that meets the standards and requirements of other certifications. Therefore, in this study, the Coretrustseal certification promotion strategy was proposed step by step. The proposed certification promotion strategy includes a list of research data policies and guidelines and examples of certification applications for use by institutions that operate domestic data repositories. Through this CTS certification process, each institution can secure research data policies and guidelines, and when CTS certification is obtained, it is judged that international reliability of the data repository can be secured.

키워드: 데이터 리포지터리, 인증, Coretrustseal, re3data, 연구데이터 정책

Data Repository, Certification, Coretrustseal, re3data, Research Data Policy

\* This research was supported by National University Promotion Program at Jeonbuk National University in 2021. 이 논문은 2021년 국립대학 육성사업의 지원을 받아 수행되었음.

\*\* 전북대학교 문헌정보학과 강사, 연구데이터융복합연구소 전임연구원  
(kimjuseop@jbnu.ac.kr / ISNI 0000 0004 7492 1806) (제1저자)

\*\*\* 전북대학교 통계학과 조교수(sjyang@jbnu.ac.kr / ISNI 0000 0004 9218 4176) (공동저자)

\*\*\*\* 전북대학교 문헌정보학과 부교수, 연구데이터융복합연구소장  
(kim.suntae@jbnu.ac.kr / ISNI 0000 0004 6492 6355) (교신저자)

논문접수일자: 2022년 4월 22일 최초심사일자: 2022년 5월 12일 게재확정일자: 2022년 5월 23일  
한국문헌정보학회지, 56(2): 209-229, 2022. <http://dx.doi.org/10.4275/KSLIS.2022.56.2.209>

\* Copyright © 2022 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. 서 론

해외 주요국가 및 우리나라에서는 4차 산업 혁명에 따른 연구데이터 확보를 위해 국가 차원의 데이터 전략을 수립하고 혁신 성장을 주도하고 있다. 이러한 동인으로 오픈사이언스의 등장과 데이터 집중형 연구라고 할 수 있는 연구 패러다임의 변화라고 할 수 있다. 특히 오픈 사이언스는 ‘오픈데이터’, ‘오픈액세스’ 그리고 ‘오픈협업’이라고 할 수 있는 키워드로 요약될 수 있다. 여기에서 ‘오픈데이터’는 데이터의 재사용과 검증을 위한 보존 필요성의 기반이 되었고 특히 연구데이터의 재사용과 투명성 확보는 데이터 리포지터리가 활성화되는 계기가 되고 있다. 여기에 2016년에는 네덜란드 및 유럽 학자들을 중심으로 오픈데이터를 위한 FAIR 원칙이 제안되었다. FAIR 원칙은 온라인상에서 연구데이터에 대한 수요와 공급이 증가하는 가운데 연구데이터를 공개함에 있어 찾을 수 있고(Findable) 접근할 수 있는(Accessible) 형태로 공개하되, 공개된 자료가 타 시스템과도 상호 호환가능하며(Interoperable) 재사용(Reusable) 또한 가능하도록 공개하자는 원칙이다(Wilkinson et al., 2016; 최희윤, 서태설, 2020).

데이터 리포지터리는 연구데이터를 수집, 저장 및 보존을 통해 재사용할 수 있는 수단이 되어가고 있다. 특히 연구의 재현성 및 검증 차원에서 재사용하고 공유할 수 있는 데이터를 어떻게 확보할 수 있는지가 오픈사이언스 그리고 데이터 집중형 연구의 성공으로 이끌 수 있는 원동력이라고 할 수 있다. 하지만 낮은 데이터의 품질은 연구 생산성을 떨어뜨릴 수 있으며 연구개발 비용과 연구 중복성을 증가시키는 등

결국에는 데이터 공유 및 재사용에 방해 원인이 될 수 있을 것이다. 따라서 개별 데이터가 아닌 데이터를 수집, 저장 그리고 보존하는 리포지터리 자체를 평가하는 것이 결국 데이터의 신뢰성을 보장할 수 있는 방법일 것이다. 또한 이러한 데이터 리포지터리 평가는 장기적으로 조직의 지속 가능성을 보장하고 업무의 관리 및 기술적 측면과 조직의 작업 방식에 대한 깊은 이해 나아가 향후에 데이터 품질 관리에 활용할 수 있다는 장점이 있다(Fernando et al., 2021).

본 연구의 목적은 이제까지 발표 또는 연구된 리포지터리 인증 체계를 분석하고 이 중에서 CTS(Coretrustseal, 이하 CTS)가 가지는 차이점을 분석하며 나아가 CTS 인증을 획득할 수 있는 전략을 제안하고자 한다. 연구에서 다루는 리포지터리 인증 체계는 CTS를 포함하여 RatSWD, CLARIN certificate B, DIN 31644, ISO 16363, TDR(Trusted Digital Repositories, 이하 TDR) 그리고 nestor 등이 있다. 이 중 CTS는 re3data 등록 리포지터리 중에서 가장 많이 획득한 인증 체계이며 타 인증 체계와의 호환성 그리고 갱신을 통한 관리가 가능하다는 점에서 본 연구의 핵심 연구대상으로 삼았다. 특히 CTS 인증을 획득하기 위하여 필요한 정책 및 가이드라인을 제시하였으며 인증 신청서를 예시로 작성하여 국내 데이터 리포지터리 운영 기관이 참고할 수 있도록 하였다.

데이터 리포지터리 인증 분석 및 인증 전략 관련한 연구가 아직 국내에서는 전무하여 본 논문이 가지는 시사점이 있지만 인증 추진 전략이 구체적이지 못한 점이 제한점이라고 할 수 있을 것이다.

## 2. 이론적 배경

이번 장에서는 데이터 리포지터리 인증 현황을 살펴보고 리포지터리 인증 관련 선행연구를 확인하여 본 연구의 시사점을 도출하고자 한다.

### 2.1 데이터 리포지터리 인증 현황

본 절에서는 데이터 리포지터리 인증 현황을 살펴보고 타 인증 체계 대비 CTS 인증 체계 이점을 살펴보고자 한다. 데이터 리포지터리 인증 현황을 살펴보기 위하여 연구데이터 리포지터리 레지스트리인 re3data 사이트를 분석하였다. re3data는 2012년에 출시된 연구 데이터 리포지터리를 색인화한 세계적인 레지스트리다. 이용자에게 2,450개에 달하는 데이터 리포지터리의 정보를 제공하며 연구자, 편집 기관, 도서관 및 출판사에게 리포지터리 환경과 유형에 대한 개요를 제공하고 있어 세계 연구데이터 인프라의 가장 대표적인 레지스트리다(Heinz,

P. et al., 2013).

다음의 <표 1>은 Re3data에 등록된 데이터 리포지터리 중 인증을 획득한 리포지터리 현황을 분석한 것이다.

re3data에서 확인한 결과 등록된 데이터 리포지터리 중 인증을 받은 리포지터리는 213개로 전체 2,813개 중 약 7.5%로 나타났다. 이 중에서 CTS 인증을 받은 리포지터리는 DSA(World Data System, 이하 DSA)와 WDS(World Data System, 이하 WDS)를 포함한 116개<sup>1)</sup>로 나타났다. 다음으로 독일 데이터 포럼의 RatSWD 인증을 획득한 리포지터리는 41개로 전체 리포지터리 중 19%를 차지했다. 세 번째 및 네 번째로 많은 리포지터리 인증 체계는 유럽 연구 인프라 CLARIN ERIC의 CLARIN certificate B가 그리고 네트워크 정보를 위한 독일 이니셔티브의 DINI Certificate가 각각 24개와 8개로 나타났다. 나머지 리포지터리 인증 체계는 DIN 31644, ISO 16363 그리고 Trusted Digital Repository로서 각각 1개씩 부여 받은 것으로 나타났으며

<표 1> 데이터 리포지터리 인증 현황

인증 체계	현황	비고
Coretrustseal (DSA+WDS)	116	54.46 % (DSA 15, WDS 42 포함)
RatSWD	41	19.25%
CLARIN certificate B	24	11.27%
DINI Certificate	8	3.76%
DIN 31644	1	DANS-EASY
ISO 16363	1	U.S. Government Information
Trusted Digital Repository	1	Polar Data Catalogue
other	21	TIB Digital Long-Term Archive, etc.
Total	213	

1) re3data에서 CTS 인증 체계를 살펴보면, 59개로 파악된다. CTS가 WDS(World Data System)와 DSA(Data Seal of Approval) 협력으로 출범한 데이터 리포지터리 인증 기관이므로 WDS 인증 및 DSA 인증은 CoreTrustSeal의 인증으로 대체된다.

이외 인증 받은 리포지터리는 21개로 확인되었다.

## 2.2 선행 연구

본 절에서는 데이터 리포지터리 체계를 중심으로 문현을 분석하여 본 연구와의 차이점을 기술하고자 한다. 다음의 <표 2>는 데이터 리포지터리 인증 관련 선행연구를 요약한 것이다.

Hervé, Mari, & Lisa(2019)는 CTS와 DSA, 그리고 WDS에 대해 일반적인 정보를 제공하고 있다. 연구자는 CTS가 인증의 편의성, 인증 효과 그리고 인증으로 인한 인식 측면에서 타 인증과의 차이점을 기술하고 있다. 편의성으로는 일반적인 요구사항을 제시하고 있어 융통성이 있다는 점과 내부 평가 절차로 이용될 수 있는 점 그리고 디지털 보존의 중요성에 대한 인식을 제고할 수 있다는 이점을 제시하고 있다.

Ingrid & Lisa(2018)는 연구데이터의 중요성 증대에 따른 리포지터리의 지속 가능한 모델

개발, CTS에 대한 일반적인 설명과 인증 시의 장점, 향후 방향 및 기대 효과를 제시하고 있다. 특히 CTS 인증은 데이터 생산자에게 데이터 및 관련 자료가 신뢰할 수 있는 방식으로 저장되고 재사용될 수 있고 편딩 기관에게는 데이터가 지속적으로 재사용 가능하다는 확신을 제공하고 있다. 또한 데이터 이용자는 데이터가 기탁된 리포지터리를 평가할 수 있으며 리포지터리는 효율적인 데이터 기탁 및 배포를 지원할 수 있다. 더불어 CTS 인증 획득 시 데이터에 대한 장기적이고 안정적인 액세스를 제공하며 정기적인 인증 주기에 따라 리포지터리의 영구적인 모니터링이 가능하다는 특징이 있다.

Mokrane & Recker(2019)는 CTS 리포지터리 인증과 FAIR 원칙이 구현하는 목표의 유사성, 중복 및 보완점을 조사하고, CoretrustSeal 요구사항이 FAIR 원칙 준수를 평가하기 위한 기준선 역할을 수행할 수 있는지에 대해 탐구하며, CTS 인증이 가지는 의의를 제시하였다. 연구자는 FAIR 원칙과 CTS 요구사항의 매핑을 통해 CTS가 FAIR 원칙을 준수하는지에 대한

<표 2> 데이터 리포지터리 인증 체계 선행연구

저자(연도)	제목	발행년도	비고
Hervé, Mari, & Lisa(2019)	CoreTrustSeal: 학술적 협업에서부터 지속 가능한 서비스에 이르기까지	2019	CTS (WDS, DSA 포함)
Ingrid & Lisa(2018)	CoreTrustSeal	2018	
Mokrane & Recker(2019)	CoreTrustSeal 인증 리포지터리: FAIR 데이터	2019	
Edmunds(2019)	학술 데이터에서의 신뢰성 구축: CoreTrustSeal 인증과 WDS 시스템	2019	
German Data Forum(2018)	RatSWD와 연구데이터 인프라: 현상 유지 및 품질 관리, RatSWD의 결과물	2018	RatSWD
Dobratz & Scholze(2006)	DINI 기관 리포지터리 인증과 그 외 사항들	2006	DINI
Donaldson(2020)	TDR 웹사이트 인증 정보: 내용 분석	2020	Trusted Digital Repository
nestor 인증 위킹 그룹(2013)	Nestor Seal 주석 보고서	2013	Nestor

평가 기반으로 활용될 수 있다고 제시하고 있다.

Edmunds(2019)는 DSA와 WDS가 CoreTrust Seal로 통합되는 과정을 설명하고, WDS가 수행하는 인증의 의의를 설명하였다. 현재 데이터 리포지터리에 대한 인증 표준은 Core, Extended, 그리고 Formal 세 가지 수준으로 구분되며, DSA와 WDS는 Core 수준의 인증 표준을 제공하였다. DSA 그리고 WDS는 일련의 신뢰 가능한 디지털 리포지터리 핵심 기준을 만들고 채택하여 현재의 CTS 인증 체계를 형성하였다.

German Data Forum(2018)은 「독일 데이터 포럼과 연구데이터 인프라」를 통해 인증 정보 기준을 제시하였다. 이러한 인증 정보 기준은 독일 데이터 포럼(RatSWD)에 이미 완료된 활동, 향후 계획된 활동 및 서비스 품질에 대한 개요를 제공하기 위해 고안되었다. 인증 정보 기준에는 '제공된 사회, 행동 및 경제 데이터의 범위 및 개발', '적절한 시기의 데이터 제공 방법', '도구 제공', '데이터셋의 품질 보증', '연구자의 이익을 충분히 고려한 데이터 보호', '서비스 개념', '기관 및 연구데이터 센터로 구성된 단일 엔티티', '기존 RDC(Research Data Centers)와 '중복 및 차별화된 특징', '연구 활동', '동일한 데이터의 다중 제공', '신청 처리 시간' 그리고 '직원' 등이 포함된다.

Dobratz and Scholze(2006)는 독일에서 오픈 액세스를 지원하는 수단으로 기관 리포지터리 인증에 대한 개요와 네트워크 정보를 위한 독일 이니셔티브인 DINI에서 개발한 DINI 인증서 2006에 대한 설명을 제공하였다. 문서 및 출판 리포지터리에 대한 DINI 인증(DINI Certificate for Document and Publication Repositories)은 디지털 문서의 잠재적 이용자 및 생산자에

게 리포지터리 운영 및 서비스에서 일정 수준의 품질이 보장됨을 보증하며 해당 인증서를 오픈액세스 개념을 지원하는 도구로 간주하였다. 따라서, DINI는 장기 보존 측면보다는 디지털 리포지터리를 운영하는 독일 고등교육 기관 간의 표준 사용, 상호 유통성 및 협력을 촉진하여 문서 및 출판 리포지터리 간 네트워크 구축을 목표로 한다.

Donaldson(2020)은 TDR(Trustworthy Digital Repository) 인증을 받은 91개의 리포지터리에 대한 인증 정보를 분석하였다. 연구자는 TDR 상태에 대한 설명, TDR 인증의 존재 및 위치, 하이퍼링크 여부, 인증 프로세스 그리고 검토 보고서의 공유 여부에 대해 분석하였다. 분석 결과 60% 이상이 검토 보고서를 게시하고 추가 인증 정보에 대한 링크를 표시하며 3분의 1 이상이 인증 과정을 설명하고 있었다.

nestor 인증 워킹 그룹(2013)은 DIN 31644 표준을 기반으로 개발되었으며 디지털 아카이브에 대한 확장된 자체 평가 프로세스를 제공한다. 자체 평가 프로세스는 디지털 아카이브가 '신뢰할 수 있는 디지털 아카이브에 대한 기준'에 정의된 대로 실용적인 방법을 제공하며 평가 결과가 긍정적인 경우 디지털 아카이브용 Nestor Seal을 이용하여 홍보할 수 있다.

이상으로 CTS, RatSWD, DINI, TDR 그리고 Nestor 등에 대한 인증 체계를 선행연구에서 살펴보았다. 각 연구에서는 인증 체계에 대한 설명과 인증 획득 기준에 대해서 기술하고 있다. 본 연구에서는 핵심 연구 대상인 CTS 인증 획득을 위하여 데이터 리포지터리를 운영하는 기관 측면에서의 인증 전략을 제안하고자 한다.

### 3. 데이터 리포지터리 인증 체계

이번 장에서는 CTS, RatSWD, CLARIN 인증 등 데이터 리포지터리 인증 체계를 살펴보고 한다. 각 인증 체계에서는 인증 정보에 대한 설명과 인증 획득을 위한 요구사항(기준)에 대해서 제시하였다.

#### 3.1 CTS(Coretrustseal)

CoretrustSeal은 RDA(Research Data Alliance, 이하 RDA) 협력기관인 ISC(International Science Council, 이하 WDS)의 WDS(World Data System, 이하 DSA)와 DSA(Data Seal of Approval, 이하 DSA)에서 출범한 인증 기관이다. 해당 인증 기관은 WDS-DSA 데이터 리포지터리 요구조건(DSA-WDS Core Trustworthy Data Repositories Requirements)의 카탈로그와 절차를 기반하여 이해관계가 있는 모든 데이터 리포지터리에 핵심 수준 인증(core level certification)을 제공한다. 해당 요구조건의 카탈로그는 신뢰할 수 있는 데이터 리포지터리의 핵심 특성을 반영하며, 데이터 리포지터리 인증을 통합한다. CoretrustSeal 데이터 리포지터리 요구사항의 모든 항목은 필수이며, 동일한 가치가 적용되는 독립 실행형 항목이다. 불가피한 일부 요소를 제외하고, 요구사항 중 증거의 중복은 가능한 한 최소한으로 유지되었다. 해당 요구사항의 확장 지침(CoreTrustSeal Extended Guidance)은 검토자의 작업을 용이하게 하며, 데이터 리포지터리 요구사항에 대한 인증을 받는 리포지터리에 대한 추가적인 지침을 제공한다. CoretrustSeal의 목적은 데이터 서비스 및

소프트웨어와 같은 여타 연구기관에 핵심 수준의 인증을 제공하는 것이다. 이러한 목적 하에 데이터 리포지터리에 전문적인 인증 도구 및 서비스를 제공하고, 검토자가 최적의 조건에서 검토를 수행할 수 있도록 지원하기 위한 비즈니스 모델을 개발하고 있다. CoretrustSeal 데이터 리포지터리 인증은 확장(extended) 수준 인증(nestor-Seal DIN 31644) 및 공식 수준 인증(ISO 16363)을 포함하는 리포지터리 인증을 위한 국제 프레임워크의 첫 번째 단계로 구상된다. 또한, CoretrustSeal 인증은 기존의 DSA 인증(DSA certification)과 WDS 일반 회원 인증(WDS Regular Members certification)을 대체하며, DSA와 WDS의 인증을 갱신하거나 재인증하려는 경우 CoretrustSeal 인증으로 전환하여야 한다(Coretrustseal, 2022). 다음의 〈표 3〉은 CTS 인증 획득을 위한 요구사항을 정리한 것이다.

CTS 인증 획득을 위한 요구사항은 크게 리포지터리 개요(1), 조직 인프라(6), 디지털 객체 관리(8) 그리고 기술(2) 등 R0부터 R16까지 17개 항목으로 구성되어 있다. CTS 인증을 획득하기 위해서는 요구사항에 대응되는 신청서를 작성하여 웹으로 신청하도록 되어 있다. 신청비용은 약 1,000 유로이며 비용이 납부되면 신청서에 대한 검토 프로세스가 진행되며 CTS 위원회에 따라 최대 5번의 검토과정을 통해 수정이 가능하도록 되어 있다. 인증 획득하기 위한 기간은 1년 6개월부터 2년 정도 소요된다.

#### 3.2 RatSWD

2004년에 설립된 독일 데이터 포럼(Rat für

〈표 3〉 CTS 인증 평가 항목(요구사항)

구분	항목	요구사항		구분	항목	요구사항	
배경	R0	<ul style="list-style-type: none"> <li>• 리포지터리 유형</li> <li>• 리포지터리 개요</li> <li>• 리포지터리 대상 커뮤니티</li> <li>• 큐레이션 수준</li> <li>• 인소싱/아웃소싱 파트너</li> <li>• 최근 인증 이후 변경사항</li> <li>• 기타 관련 사항</li> </ul>	디지털 개체 관리	R9	문서화된 보관 절차	<ul style="list-style-type: none"> <li>• 데이터 보존 정책</li> <li>• (아카이브 복사본 간에 일관성을 보장 위한) 검토 내용 (정책/시스템)</li> </ul>	
	R1			R10	보존 계획	<ul style="list-style-type: none"> <li>• 데이터 보존 정책</li> </ul>	
	R2			R11	데이터 품질	<ul style="list-style-type: none"> <li>• 데이터 품질 정책</li> <li>• 지정된 커뮤니티의 데이터 및 메타데이터에 대한 설명 및/또는 평가 기능</li> <li>• 관련 산출물에 대한 인용 또는 인용 지수에 대한 링크 제공</li> </ul>	
	R3			R12	워크 플로우	<ul style="list-style-type: none"> <li>• 워크플로우 정책</li> </ul>	
	R4			R13	데이터 검색 및 식별	<ul style="list-style-type: none"> <li>• 검색 기능 제공</li> <li>• 데이터 인용 권장</li> <li>• 사용 영구식별자 시스템</li> <li>• 검색 가능한 메타데이터 목록을 적절한(국제적으로 합의된) 표준을 준수하며 유지</li> <li>• 메타데이터의 자동 수집</li> <li>• 하나 이상의 분야별 또는 일반 자원 레지스트리에 리포지터리의 포함 여부</li> </ul>	
	R5			R14	데이터 재사용	<ul style="list-style-type: none"> <li>• 리포지터리에서 제공하는 메타데이터 템플릿</li> <li>• 데이터 재사용 정책</li> <li>• 마이그레이션 계획</li> </ul>	
	R6			R15	기술 인프라	<ul style="list-style-type: none"> <li>• 참조 표준</li> <li>• 리포지터리에서 인프라 개발 계획</li> <li>• 소프트웨어 인벤토리 유지 관리 및 시스템 문서 사용 가능 여부</li> <li>• 커뮤니티 지원 소프트웨어 사용 여부</li> <li>• 재해계획 및 비즈니스 연속성 계획 수립 여부</li> </ul>	
	R7			R16	보안	<ul style="list-style-type: none"> <li>• 운영 중단 시 필수 서비스의 신속한 복구 또는 백업을 제공하기 위한 절차 및 준비</li> <li>• 보안과 관련된 역할</li> <li>• 사용하는 모든 위험 분석 도구</li> <li>• 필요한 보안 수준 및 지원 방법</li> <li>• 시스템 접근 관리를 위해 사용되는 인증 및 권한 부여 절차</li> </ul>	
디지털 개체 관리	R8	<ul style="list-style-type: none"> <li>• 컬렉션 개발 정책</li> <li>• 메타데이터 제공 확인 절차</li> <li>• 리포지터리 선호 형식 목록</li> <li>• 수집 정책에 적합한 데이터만 허용되도록 마련된 절차</li> </ul>	기술	R0~R16, 17개의 요구사항			

Sozial und Wirtschaftsdaten, 이하 RatSWD)은 경험적 사회, 행동 및 경제 과학을 위한 연구 데이터 인프라 확장 및 개선과 관련하여 독일 연방 및 주 정부에 자문을 제공하는 독립 위원회이다. 연구데이터 센터(research data centers, RDC)와 데이터 서비스 센터(data service centers, DSC)는 독일 데이터 포럼(German Data Forum, RatSWD)의 인증 및 지원을 받아 사회, 경제 및 행동 과학을 위한 연구데이터 인프라 개선을 목표로 하며, 이러한 목표의 달성을 위하여 정부 통계에서 제공하는 인프라 범위를 넘어선 영역의 분야별 연구, 평가 연구, 연구 기반 설문 조사 및 공적 편성을 받는 연구와 같은 인프라 설치를 도모한다.

과학적 분석의 데이터 요구사항은 과학 및 정보 기술의 발전과 경제적, 사회적 변혁의 복잡성으로 인해 근본적으로 변화하였다. 통계 집계 뿐만 아니라, 심층적인 경험적 분석을 위해 마이크로데이터를 제공하는 방법론적 및 콘텐츠 관련 요구사항이 필요하며, 이와 밀접하게 관련된 것은 차별화된 경험적 분석을 수행할 수 있는 익명화된 마이크로데이터에 대한 액세스를 제공하는 것이다. 연구데이터 인프라를 지원하는 데 있어 RatSWD의 목표는 사용 가능한 정보를 최대한 활용할 수 있는 데이터 환경을 만드는 것이다(German Data Forum, 2018).

RatSWD의 핵심 과제 중 하나로, 시너지 효과를 창출하기 위해 확장된 연구 인프라를 조언 및 조정하는 역할이 있다. 독일 과학 및 인문학 위원회(German Council of Science and Humanitites)에서 강조한 바와 같이, 연구데이터 센터 생성의 역학 관계를 위하여 RatSWD는 지속적인 데이터 품질을 보장할 뿐만 아니라,

신속하고 사용자 친화적이며 오픈 데이터 액세스를 제공하여야 하며, 데이터 제공에 대한 과도한 요금 기반 모델을 지양하면서 이러한 요구 사항을 행정적으로 이해 및 이행할 수 있도록 RDC 또는 DSC로 인증하는 기준을 개발하였다. 다음의 〈표 4〉는 RatSWD의 인증을 받기 위한 연구데이터 센터(RDC)의 필수 정보 기준을 정리한 것이다(research data infrastructure of the German Data Forum, 2010).

RatSWD 인증 획득을 위한 RDC의 필수 정보 기준은 접근 경로, 도큐멘테이션, 장기 보존, 데이터셋 품질 보증, 익명화, 데이터셋 접근, 신청 처리 시간, 직원 및 인프라 등 16개의 항목으로 구성되어 있다. CTS 인증 요구사항과 다르게 RatSWD는 우선순위가 정해져 있는데 ‘최소 하나의 데이터 액세스 경로’, ‘충분한 데이터 도큐멘테이션 제공’ 그리고 ‘장기 데이터 이용가능성에 대한 개념’은 필수(M)이며 나머지 항목은 정보(I)로 구분하고 있다.

### 3.3 CLARIN certificate B

2012년에 설립된 유럽 연구 인프라 CLARIN ERIC(European Research Infrastructure CLARIN ERIC)은 과학자, 강사 및 학생과 같은 시민에게 자연어 텍스트의 대규모 컬렉션 처리와 관련된 연구를 수행할 수 있는 전문 데이터 수집, 지식 및 기술적 능력을 제공한다. 인프라는 세계 각국의 다양한 위치에 분산되어 있으며, CLARIN ERIC 컨소시엄의 구성원들에게 자연어 분석을 위한 범용 도구와 언어 자원 및 소스 자료의 오픈 액세스 아카이브를 제공한다. CLARIN ERIC 인프라를 구성하는 센

〈표 4〉 RatSWD 인증 획득을 위한 RDC의 필수 정보 기준

기준	내용	기준	내용
최소 하나의 데이터 액세스 경로	<ul style="list-style-type: none"> <li>데이터를 사용할 수 있는 기본 액세스 경로(최소 경로)는 무엇인가?</li> <li>해당 액세스 경로를 선택한 이유는 무엇인가?</li> </ul>	기관 및 RDC로 구성된 단일 개체	<ul style="list-style-type: none"> <li>기관에 여러 RDC가 있는가?</li> <li>만약 그렇다면, 기존 RDC에 데이터를 통합할 수 없는 이유는 무엇인가?</li> </ul>
충분한 데이터 도큐멘테이션 제공	RDC는 데이터 문서를 제공하는가?	연구와 관련된 모든 데이터셋 제공	<ul style="list-style-type: none"> <li>사용자가 연구와 관련된 모든 데이터셋에 액세스할 수 있는가?</li> </ul>
장기 데이터 이용가능성에 대한 개념	RDC는 최소 10년(법적으로 규정된 보관 기간)동안 데이터 액세스를 보장하는가?	기존 RDC와 중복 및 차별화된 기능	<ul style="list-style-type: none"> <li>중복 가능성을 제공하는 RDC가 이미 있는가?</li> </ul>
사회 행동 또는 경제 과학 데이터 제공	<ul style="list-style-type: none"> <li>RDC에서 제공하는 자료가 다른 주제분야와 연구내용은 무엇인가?</li> <li>RDC는 어떤 기관에서 연구 데이터를 얻는가?</li> </ul>	연구 활동	<ul style="list-style-type: none"> <li>RDC는 RDC가 제공하는 데이터로 독립적인 연구를 수행하는 학문적으로 훈련된 직원을 고용하는가?</li> <li>제도화된(예: 계약으로 규제되는) 연구 협력이 있는가? 혹은 계획 중인가?</li> </ul>
도구 제공	RDC는 단순한 문서화 이상의 도구(예: 코드북, 변수 설명, 구문)를 제공하는가?	동일한 데이터의 다중 제공(다중 호스팅, 다중 사이트 호스팅 아님)	<ul style="list-style-type: none"> <li>연구데이터는 이미 다른 기관에서 사용할 수 있는가?</li> </ul>
데이터셋의 품질 보증	<ul style="list-style-type: none"> <li>데이터 검토(품질)가 RDC 책임의 일부인가?</li> </ul>	신청 처리 시간	<ul style="list-style-type: none"> <li>계약 작성에 필요한 모든 사용자 정보와 문서가 준비된 시점부터 데이터가 전송되는 시점까지의 평균 처리 시간은 얼마나 되는가?</li> <li>RDC는 데이터 액세스 제공에 대해 요금을 부과하는가?</li> </ul>
연구자의 이익을 고려한 데이터 보호	<ul style="list-style-type: none"> <li>데이터 보호와 연구자의 이익을 보호하는데 사용되는 절차 목록과 개인 데이터의 익명화에 사용되는 법적 근거 및 이에 상응하는 법적 근거</li> </ul>	직원	<ul style="list-style-type: none"> <li>RDC(예정)는 몇 명의 직원을 고용하는가?</li> </ul>
서비스 개념	<ul style="list-style-type: none"> <li>사용자가 사용할 수 있는 서비스(예: 텁당자, 자문 서비스, RDC 직원이 제공하는 교육, 워크샵 등)는 무엇인가?</li> </ul>	인프라 발전에 기여	<ul style="list-style-type: none"> <li>RDC의 고유한 점과 연구 데이터 인프라에서 사용할 수 없는 점은 무엇인가?</li> <li>RDC는 어디에서 조언이나 지식 공유가 필요하다고 생각하는가?</li> </ul>

터는 다음과 같이 세 가지 유형으로 구분된다 (CLARIN, n.d.).

- B 유형 센터: 유럽 CLARIN 인프라와 통합된 언어 처리 서비스, 컴퓨팅 장비 및 데이터 리포지터리를 유지 관리하는 기본 인프라
- C 유형 센터: 메타데이터 유형과 같은 특정 리소스 설명을 공유 및 제공하여 인프라에 협력하는 기관
- K 유형 센터(지식 공유 센터): 특정 연구 프로젝트를 지원하는 전문가 팀 네트워크

를 형성하기 위하여, 인프라 활동에 대한 지식을 홍보함. 또한 이용자가 인프라 운영에 관한 문제를 보고하고, 전문적인 도움을 받으며, 인프라 사용에 관한 문의를 처리함

이 중 B 유형의 센터는 네트워크의 기본 요소로, 허용된 표준을 엄격히 준수하여야 한다. B 센터로 등록하기 위하여, 센터 후보 기관은 필요한 서류를 첨부하여 센터 위원회의 위원장 또는 부의장에게 평가 요청 메일을 보내고, 센터 위원회에서 평가 위원회와 협의한 후 예상

평가 기간을 고지한다. 필요한 경우, 평가 위원회에서 센터 후보 기관에게 연락하여 불분명하거나 질문이 있는 사항을 처리할 수 있다. 평가가 완료되면, 센터 위원회는 결과를 센터 후보 기관에게 전달하며 CLARIN ERIC 웹사이트에 공개된다. 인증된 센터는 공식적으로 인정되었음을 나타내는 공식 로고를 웹사이트에 추가할 수 있다.

- 센터 레지스트리에 최신 항목이 있음
- B 센터 체크리스트를 채웠음
- CTS 절차를 완료 또는 진행 중임
- National Coordinator 또는 CLARIN 컨

소시엄 센터 위원회 대표가 작성한 센터 준수 선언문을 기재함

다음의 <표 5>는 CLARIN B 센터의 요구사항 및 체크리스트를 요약한 것이다.

CLARIN B 센터의 요구사항은 크게 ‘일반 요구사항’, ‘지적재산권 및 개인정보 보호’, ‘데이터 센터의 외부 평가’, ‘서버 인증’, ‘통합 ID 관리’, ‘메타데이터’, ‘영구 식별자’ 그리고 ‘통합 콘텐츠 검색’ 등 8개 대항목으로 구성되어 있다. 위 인증의 경우 CLARIN B 센터 인증 후 보로 등록되기 위해서는 CTS 인증이 필요하다는 특징이 있다.

<표 5> CLARIN B 센터 요구사항

요구사항	세부내용	요구사항	세부내용
1. 일반 요구사항	a. 센터 준수	• 센터는 CLARIN 커뮤니티에 유용한 서비스를 제공해야 함	3. 데이터 센터의 외부 평가 • 센터는 DSA 또는 MOIMS-RAC 접근 방식에서 제안한 바와 같이 적절하고 명확하게 지정된 리포지터리 시스템이 있어야 하고 품질 평가 절차에 참여하여야 함
	b. CLARIN 연결 가시성	• 각 센터는 웹사이트에서 CLARIN을 명시적으로 참조하여야 함	4. 서버 인증 • 센터는 보안 지침을 준수하여야 함. 즉. 서버에는 승인된 인증서가 있어야 함
	c. 편당지원	• 각 센터는 편당 지원 상태와 이와 관련된 관점에 대해 명시적으로 설명하여야 함	5. 통합 ID 관리 • 센터는 SAML2.0 및 신뢰 선언을 기반으로 하는 싱글 ID 및 SSO(Single Sign On)을 지원하기 위하여 가능한 경우 국가 ID 연합과 CLARIN 서비스 제공자 연합에 가입하여야 함
	d. 제공 자원 및 서비스	• 각 센터는 센터에서 사용할 수 있는 CLARIN 준수 자원 및 서비스에 대해 명시적으로 설명하여야 함	6. 메타데이터 • 센터는 CLARIN 계약에 따라 CCR5와 같은 승인된 레지스트리의 요소를 사용하는 구성요소 기반 메타데이터(CMDI)를 제공하여야 함 • 메타데이터는 OAI-PMH를 통해 수집 가능하여야 함
2. 지적재산권 및 개인정보 보호	a. 데이터 제공 및 지적재산권	• 각 센터는 데이터 및 서비스 제공 정책과 지적 재산권 문제 처리에 대해 명확하게 설명하여야 함	7. 영구 식별자 • 센터는 PID를 메타데이터 레코드와 연결(처리)하여야 함. 이러한 PID는 HTTP 수락 헤더를 고려하여 사람과 기계 해석 모두에 적합하여야 함 • 개별 파일(예: 텍스트, zip 또는 사운드 파일)은 식별자 또는 다른 PID와 결합된 메타데이터 레코드의 PID로 참조 가능함
	b. 개인정보보호 정책	• 센터는 연합 서비스 제공자 각각에 대해 GÉANT 데이터 보호 행동 강령(Data Protection Code of Conduct, DP-CoC)을 구현하여야 함	8. 통합 콘텐츠 검색 • 센터는 SRU/CQL 엔드포인트를 제공하여 결합과 함께 연합 콘텐츠 검색에 참여하도록 선택 가능함

### 3.4 DINI Certificate

DINI(Deutsche Initiative für Netzwerkinformation, Deutsche Initiative für Netzwerk Information)는 독일 고등 교육 인프라 및 서비스 기관 및 학회가 결성한 이니셔티브이다. DINI 자체는 편당 지원 기관이 아니지만, 독일 연구 재단(German Research Foundation, DFG) 또는 독일 교육 연구부(German Ministry for Education and Research, BMBF)와 같은 독일 편당 지원 기관과 협력한다. DINI의 주요 목표는 표준화되고 상호운용할 수 있는 정보 서비스 및 대학 간 통신 네트워크에 대한 권장사항을 수립하는 것으로, 주요 과제로써 네트워크로 연결된 디지털 출판 서비스의 구조화 및 구축을 도모한다. DINI는 우수 또는 모범사례를 평가 및 보급하고 지역, 국가 및 국제 협력을 시작하고 강화한다. 또한 독일 내에서 OAI-PMH(Open Archives Initiative Protocol for Metadata Harvesting)와 같은 기술이 사용될 수 있도록 한다. 따라서 DINI의 초점은 정보 인프라의 개발 및 지원이며, 여러 워킹그룹의 운영을 통해 이를 달성하고자 한다. 문서 및 출판 리포지터리에 대한 DINI 인증(DINI Certificate for Document and Publication Repositories)은 리포지터리를 일반 기관의 웹 서버와 구별하고, 디지털 문서의 잡재적 이용자 및 작성자에게 리포지터리 운영 및 서비스에서 일정 수준의 품질이 보장됨을 보증한다. 또한, DINI는 인증서를 오픈액세스 개념을 지원하는 도구로 간주한다. 따라서, DINI는 장기 보존 측면보다는 디지털 리포지터리를 운영하는 독일 고등교육 기관 간의 표준 사용, 상호 운용성 및 협력을 촉진하여 문

서 및 출판 리포지터리 간 네트워크 구축을 목표로 한다. 다음 <표 6>은 DINI 인증의 기준 및 세부사항 중 최소 요구사항만을 요약한 것이다 (Dobratz & Scholze, 2006; dini, n.d.).

DINI 인증은 ‘서비스의 가시성’, ‘지침(정책)’, ‘저자 및 출판사 지원’, ‘법적 측면’, ‘정보 보안’, ‘인덱싱 및 인터페이스’, ‘개방형 지표(사용 통계)’ 그리고 ‘장기 아카이빙’ 등 8개 기준으로 구성되어 있다. 특히 최소 및 권장요구사항을 제시하고 있어 인증을 준비하는데 구체적인 가이드라인으로 활용될 수 있다는 특징이 있다. DINI 인증은 연구자 간의 의사소통과 출판물의 신뢰성, 지속가능성 및 겸증 가능성을 보장한다는 측면에서 타 인증과 유사성을 가지고 있다.

### 3.5 기타

DIN 31644는 독일의 DIN 표준 위원회(DIN-Normenausschuss Information und Dokumentation)에서 디지털 보존을 위한 독일 역량 네트워크(Nestor) 워킹그룹의 권장 사항을 기반으로 신뢰할 수 있는 디지털 아카이브에 대한 기준을 채택하며 등장하였다. DIN 31644는 독일어로 되어 있으나, 영어권에서도 사용이 가능하며, Nestor가 진행하는 확장 인증 절차는 약 3개월이 소요된다. 해당 절차는 ISO 16363을 따르는 공식 인증인 전체 외부 감사(full external audit)가 아니며, DIN 31644는 평가 옵션으로 두 가지 항목이 존재한다. 해당 표준은 조직 프레임워크, 객체 관리, 그리고 인프라 및 보안이라는 세 가지 부분으로 구성되며, 각 항목에 따라 신뢰할 수 있는 리포지터리에 대한

〈표 6〉 CLARIN B 센터 요구사항

기준	최소 요구사항	기준	최소 요구사항
기준 1. 서비스의 가시성	<ul style="list-style-type: none"> <li>온라인에서 이용 가능</li> <li>기관 홈페이지에서 참조 가능</li> <li>영구 URL 사용</li> <li>웹사이트에 내 오픈액세스 출판물에 대한 명확한 표시</li> </ul>	기준 5. 정보 보안	<ul style="list-style-type: none"> <li>시스템에 보안 개념</li> <li>기술적 유지보수 계획</li> <li>시스템 운영에 필요한 문서</li> <li>백업 절차에 의한 정기 저장</li> <li>정기적 모니터링</li> <li>문서의 무결성</li> <li>영구식별자 할당</li> <li>영구식별자로 제공</li> <li>영구 보존</li> <li>TSL 기술을 이용한 데이터 교환</li> </ul>
기준 2. 지침(정책)	<ul style="list-style-type: none"> <li>서비스 정책 공개</li> <li>운영자의 권리와 의무 정의</li> <li>저자 및 출판사의 권리와 의무 정의</li> <li>문서 유형에 대한 설명 내용 및 기술 품질에 대한 요구사항</li> <li>콘텐츠를 사용할 수 있는 최소 기간과 해당 보증에 대한 사양</li> <li>장기 아카이브에 대한 설명</li> <li>서비스의 기술적 운영에 대한 설명</li> <li>오픈 액세스에 대한 진술</li> <li>문서 버전 및 삭제 방법에 대한 정보</li> </ul>	기준 6. 인덱싱 및 인터페이스	<ul style="list-style-type: none"> <li>색인화 규정의 문서화</li> <li>더블린코어를 이용한 색인</li> <li>DDC 분류</li> <li>문서 또는 출판 유형 설명 포함</li> <li>OAI-PMH 2.0 및 DINI OAI 지침 준수</li> <li>직접한 데이터 형식으로 제공</li> </ul>
기준 3. 저자 및 출판 사 지원	<ul style="list-style-type: none"> <li>웹사이트를 통한 연락처 및 지원서비스 접근</li> <li>리포지터리 업로드 방법 사용 가능</li> <li>전자출판 관련 기술적 정보</li> </ul>	기준 7. 개방형 지표 (이용 통계)	<ul style="list-style-type: none"> <li>일관된 액세스 로그 유지</li> <li>웹 서버 로그는 익명으로 처리</li> <li>자동 액세스는 기록되지 않음</li> <li>통계 작성 기준 및 표준 설명에 필요한 문서</li> </ul>
기준 4. 법적 측면	<ul style="list-style-type: none"> <li>저자와 출판사 및 운영기관 간 법적 관계는 정식 계약을 준수함</li> <li>운영자는 기탁 라이선스를 게시</li> <li>출판물의 전자 저장과 대중이 이용할 권리</li> <li>제3자에게 통지 및 전송할 권리</li> <li>복사 및 추가, 전자적 또는 물리적 형식으로 변환할 권리</li> <li>작성자에 의한 라이선스 선택 가능</li> <li>2차 출판물에 대한 배포 의사의 문서화 및 검증 가능한 방식으로 표현</li> <li>제3자의 권리가 침해되지 않을 것을 운영자에게 보증</li> <li>법적 요구사항을 충족하는 웹사이트에 법적 고지</li> <li>운영자는 게시된 문서의 법적 속성을 각자의 문서에 기록</li> </ul>	기준 8. 장기 아카이빙	<ul style="list-style-type: none"> <li>최소 5년 이상의 보존 기간</li> <li>원본 파일 및 사본에 대한 기술적인 보호가 없음</li> <li>문서 삭제에 대한 규정 존재</li> </ul>

요구사항을 제시하고 있다(Digital Preservation Coalition, 2021).

ISO 16363(CCSDS 652.0-M-1)은 PTAB (Primary Trustworthy Digital Repository Authorization Body, 이하 PTAB)이 만들고

검토하는 국제 표준으로, '신뢰할 수 있는 디지털 리포지터리의 검토 및 인증 - OAIS를 기반으로 아카이브가 수행해야 하는 작업에 대한 포괄적인 측정 기준 설정'이다. 이와 유사한 표준으로, ISO 14721(OAIS - 아카이브가 디지

털 정보의 장기 보존을 제공하는 데 필요한 항목에 대한 참조 모델)와 ISO 16919(신뢰할 수 있는 후보 디지털 리포지터리의 감사 및 인증을 제공하는 기관에 대한 요구 사항 - 감사 기관에 대한 역량 및 요구 사항 지정)이 있으며, 이 세 가지 표준은 국제적으로 인정 및 인증된 '신뢰할 수 있는 디지털 리포지터리 세트'를 구축하는 데 밀접하게 관련된 주요 집합(family)을 형성한다. 디지털 보존은 단순한 작업이 아니며, 이러한 표준을 통해 보존 작업을 담당하는 리포지터리의 신뢰도를 판단할 필요가 있다(PTAB, n.d.).

TDR(trusted digital repositories, 이하 TDR)은 연구도서관 그룹(Research Library Group, 이하 RLG)과 국립 기록보관소(National Archives and Records Administration)와 협력하여 만들어진 신뢰할 수 있는 디지털 리포지터리 인증을 위한 평가 체크리스트(An Audit Checklist for the Certification of Trusted Digital Repositories)에서 시작되었다. '현재와 미래에 지정된 커뮤니티에 관리되는 디지털 리소스에 대한 안정적이고 장기적인 액세스를 제공하는 것'을 목표로 하며, 1996년에 그 필요성이 제기된 이후 TDR에 대한 여러 국제 표준이 개발되었다. TDR 프레임워크는 다음과 같이 신뢰성을 높이는 일련의 세 가지 수준으로 구성된다.

- 기본 인증(Basic Certification) : DSA 인증을 획득한 리포지터리에게 부여함
- 확장 인증(Extended Certification) : 기본 인증 상태인 리포지터리 중, ISO 16363 또는 DIN 31644에 따라 구조화되고 외부에서 검토되며, 공개적으로 사용 가능한 자체

평가를 수행하는 리포지터리에게 부여함

- 공식 인증(Formal Certification) : 기본 인증 외에 ISO 16363 또는 DIN 31644를 기반으로 전체 외부 평가 및 인증을 획득한 리포지터리에게 부여함

Nestor(network of expertise in long-term storage of digital resources in Germany, 이하 nestor)는 2003년부터 2009년까지 시행된 BMBF (German Federal Ministry for Education and Research) 후원 프로젝트로 설정되었으며, 2009년 7월부터 독립적으로 운영되었다. Nestor의 인증 워킹그룹(Certification WG)은 표준 DIN 31644 "신뢰할 수 있는 디지털 리포지터리에 대한 기준(Criteria for trusted digital repositories)"을 기반한 자체 평가로써 인증 프로세스의 초안을 작성하였다. Nestor 인증은 아카이브가 준수하여야 하는 34개 기준의 복잡하고 포괄적인 세트로 구성되며, 해당 인증은 디지털 리포지터리 평가 및 인증을 위한 유럽 프레임워크에 따라 확장된 수준의 신뢰성을 보장함을 의미한다. DIN 31644 "신뢰할 수 있는 디지털 리포지터리에 대한 기준"을 기반으로 개발 및 제공하는 디지털 아카이브에 대한 확장된 자체 평가 프로세스는 디지털 아카이브가 신뢰할 수 있는지 확인할 수 있는 조화롭고 실용적인 방법을 제공한다. 검토된 평가에서 긍정적인 결과가 도출될 시, 신뢰할 수 있는 디지털 아카이브용 Nestor 인증으로 이를 홍보할 수 있다. 인증 수수료는 500유로이며, 이러한 인증은 독립 실행형 솔루션으로 취득이 가능하지만, 유럽 평가 및 인증 프레임워크(European Framework for Audit and Certification)에도 적합하며, DSA 기본 인

증으로부터의 확장 인증을 부여하는 역할도 수행한다(nestor, 2020; EuroCRIS, 2016; nestor n.d.).

### 3.6 종합

이번 장에서는 1절부터 5절까지 데이터 리포지터리 인증 체계를 살펴보았다. 다음의 <표 7>은 3장에서 분석한 데이터 리포지터리 인증 체계의 시사점을 정리한 것이다.

본 연구의 핵심 대상인 CTS 인증과 타 인증 체계와 비교하면 먼저, CTS 인증은 공식 인증은 아니지만 데이터 리포지터리 레지스트리인 re3data에서 확인한 인증 체계 중 가장 많이 획득한 인증으로서 디팩토라고 할 수 있다. 특히 CTS 인증은 타 인증 대비 다음의 <표 8>과 같이 인증 편의성과 효과 그리고 인식 측면에서 이점을 가지고 있다(Dillo, I. und de Leeuw, L. 2018; L'Hours, H., Kleemola, M., & de Leeuw, L., 2019).

편의성 측면의 경우 CTS 인증이 자체 평가로 이용될 때 평균적으로 2~4일이 소요되며 기밀정보를 보호할 수 있는 접근성을 제시할 수 있다는 점과 효과 측면에서는 CTS 인증이 내부 평가 절차로 활용될 수 있으며 마지막으로 인식 측면으로 디지털 보존의 중요성과 OAIS 표준 준수의 필요성에 대해 인식을 강조하고 있다는 점이 특징이라고 볼 수 있다.

더불어 CTS 인증은 확장 수준 인증(nestor-Seal DIN 31644) 및 공식 수준 인증(ISO 16363)을 포함하는 리포지터리 인증을 위한 글로벌 프레임워크의 첫 번째 단계로 계획되었으므로 해당 인증을 획득하는 것은 확장과 공식 수준 인증을 내포하는 것으로 판단할 수 있다(Coretrustseal, 2022).

## 4. CTS 인증 추진 전략

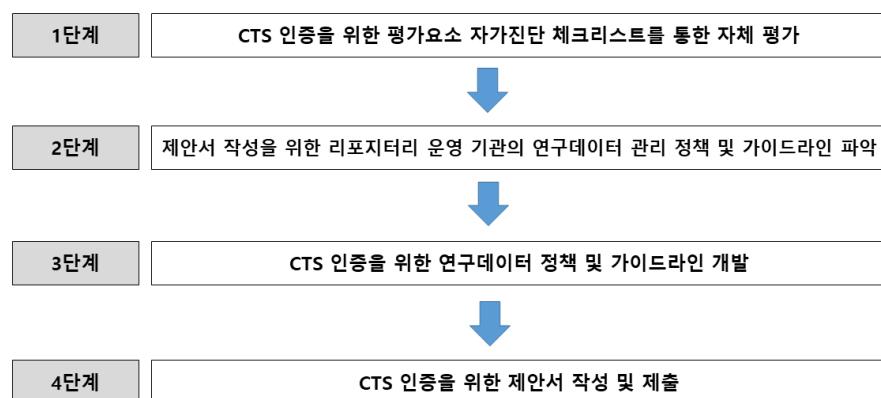
이번 장에서는 데이터 리포지터리 인증을 위

<표 7> 데이터 리포지터리 인증 체계 및 시사점

인증 체계	시사점
CTS	<ul style="list-style-type: none"> <li>• 공식/확장 리포지터리 인증을 위한 국제 프레임워크의 첫 번째 단계</li> <li>• 3년마다 재인증 요구</li> </ul>
RatSWD	<ul style="list-style-type: none"> <li>• 지속 가능하고 신뢰할 수 있는 데이터 액세스에 대한 FAIR 원칙 및 인증을 수립하여 연구데이터 센터 지원</li> </ul>
CLARIN	<ul style="list-style-type: none"> <li>• CLARIN의 B 센터 인증 후보로 기재되기 위해서는 CTS 인증이 필요</li> </ul>
DINI Certificate	<ul style="list-style-type: none"> <li>• 오픈 액세스 리포지터리 및 출판 서비스를 위한 인증 체계</li> </ul>
DIN 31644	<ul style="list-style-type: none"> <li>• 확장 인증 절차(약 3개월 소요)</li> </ul>
ISO 16363	<ul style="list-style-type: none"> <li>• 공식 수준 인증</li> <li>• 향후 2년 동안 매년 사후평가 필요, 3년차에는 완전한 재인증 심사 필요</li> </ul>
TDR	<ul style="list-style-type: none"> <li>• 기본, 확장, 공식 수준 인증 존재</li> <li>• 다양한 상황, 아키텍처 및 기관의 책임을 수용할 수 있을 만큼 충분히 광범위</li> </ul>
nestor	<ul style="list-style-type: none"> <li>• DIN 31644 기반의 확장 수준 인증</li> <li>• 아카이브가 준수하여야 하는 34개 기준의 복잡하고 포괄적인 세트로 구성</li> </ul>

〈표 8〉 CTS 인증 획득시 이점

구분	내용
인증 편의성 측면	<ul style="list-style-type: none"> <li>• 자체 평가 수행에 소요되는 시간이 짧음</li> <li>• 일반적인 요구사항 제시</li> <li>• 정보 공개의 예외성</li> <li>• 인증 개선 시, 데이터 리포지터리에 진행 상황이 표시됨</li> </ul>
인증 효과 측면	<ul style="list-style-type: none"> <li>• 문서 및 프로세스의 개선 도모</li> <li>• 데이터 이용자의 리포지터리 평가 지원</li> <li>• 데이터의 효율적인 기탁 및 배포 시, 데이터 리포지터리 지원</li> <li>• 내부 평가 절차로 이용 가능</li> <li>• 연속 및 장기 계획의 필요성 · 유용성 재확인 및 관리 의제 상에서의 명시성 확보</li> <li>• DIN 31644 인증을 신청하기 위한 기반으로 이용 가능</li> </ul>
인증 인식 측면	<ul style="list-style-type: none"> <li>• 디지털 보존의 중요성에 대한 인식 제고</li> <li>• 신뢰성 있는 이미지 형성</li> <li>• 인식 및 투명성의 고취</li> <li>• 동일한 표준 하에 작업을 수행하여 커뮤니티 구축에 도움이 됨</li> <li>• OAIS 표준 준수의 필요성 강조</li> <li>• 응용 프로그램의 입증 자료 작업 시, 검토자와의 상호작용의 중요성을 인식시킴</li> </ul>



〈그림 1〉 CTS 인증 신청 준비 절차

한 국제 프레임워크의 첫 번째 단계인 CTS 인증을 획득하기 위한 전략을 제시하고자 한다. CTS 인증 절차는 다음의 〈그림 1〉과 같다.

CTS에서는 리포지터리 운영 기관을 위한 평가요소 자가진단 체크리스트를 제공하고 있다. 자가진단 체크리스트는 평가항목과 동일한 내용

으로 요구조건별 점수를 0~4점<sup>2)</sup>까지 부여할 수 있도록 제시되어 있다. 자가진단 체크리스트를 통해 현재 리포지터리의 현황을 파악한 다음으로 CTS 신청서 작성을 위한 기관의 정책 및 가이드라인을 파악해야 한다. 다음의 〈표 9〉은 R1~R16의 CTS 요구사항을 통해 인증 획득을

2) 0점은 '해당 없음', 1점은 '고려하지 않음', 2점은 '이론적 개념', 3점은 '구현 단계' 마지막으로 4점은 '완전한 구현'을 나타낸다.

〈표 9〉 CTS 인증을 위한 연구데이터 관리 정책

구분	정책	비고	구분	정책	비고
R1	데이터 접근 및 보존에 관한 사명 선언문	기타	R9	데이터 보관 프로세스	데이터 보존
R2	라이선스 정책	데이터 공유/활용	R10	데이터 제출자와 리파지토리 간 계약서	데이터 수집
R3	데이터 제출자와 리파지토리 간 계약서	데이터 수집	R11	데이터 품질	데이터 관리
R4	공개 위험이 있는 데이터 관리 정책	데이터 관리	R12	워크플로우	데이터 관리
	공개 데이터 사용 지침	데이터 공유/활용			
R5	인력 개발	기타	R13	메타데이터	데이터 수집
R6	커뮤니케이션 정책/방법	기타	R14	메타데이터	데이터 수집
				마이그레이션 계획	데이터 보존
				데이터 재사용	데이터 공유/활용
R7	데이터 신뢰성	데이터 수집	R15	커뮤니케이션 정책/방법	기타
	데이터 무결성	데이터 관리		참조 표준	기타
	데이터 변경 전략	데이터 관리		인프라 개발 계획	기타
	버전 관리 전략	데이터 관리			
R8	리포지터리 선호 형식 목록	데이터 수집	R16	복구 및 백업 절차	데이터 보존
	컬렉션 개발 정책	데이터 수집		데이터 보안	데이터 보존

위해 필요한 정책을 연구데이터 관리 측면에서 정리한 것이다.

〈표 9〉에서 분석한 바와 같이 CTS 인증을 위해 필요한 연구데이터 관리 정책을 데이터 수집, 관리, 보존 그리고 공유/활용으로 정리한 결과 다음의 〈표 10〉과 같다.

CTS 인증 획득을 위한 다음 전략으로는 데이터 리포지터리를 운영하는 기관의 정책과 CTS 인증에 필요한 정책을 비교하는 것이다. 다음의 〈표 11〉은 모 기관이 운영하는 데이터 리포지터리가 CTS 인증을 추진한다는 가정하에 CTS 요구조건 중 조직 인프라에 해당하는 정책을 비교한 것이다.

예를 들어 ‘데이터 기탁자를 위한 지침’, ‘데이터 제공 동의서’ 그리고 ‘데이터 값 체크리스트’ 등과 같은 CTS 요구 정책이 기관 정책에

포함되어 있지 않다면 CTS 인증 추진을 위해서 해당 정책 또는 가이드라인을 개발할 필요성이 있을 것이다. 마지막 인증 추진 전략으로 CTS 인증을 위한 신청서 작성 및 제출이다. 다음의 〈표 12〉는 CTS 인증 요구 조건 중 R1~R3까지의 신청서를 예시로 작성한 것이다.

〈표 12〉는 데이터 리포지터리를 운영하는 기관이 해당 리포지터리에 대한 CTS 인증 추진 시 참고할 수 있는 신청서 일부 내용을 작성한 것이다. 예시로 작성하기 전 작성방법에 가이드를 확보하기 위하여 CTS 인증을 획득한 리포지터리의 신청서를 분석한 결과 R0~R16까지 문항에 대한 답변과 해당 정책의 출처를 기술한 것을 확인할 수 있었다. 특히 인증을 획득하기 위해서는 본문에도 기술한 바와 같이 약 1년 6개월에서 2년 정도 기간이 소요되며 약 3~5회

〈표 10〉 CTS 인증을 위해 필요한 연구데이터 정책

연구데이터 정책	세부 내용
데이터 수집	<ul style="list-style-type: none"> <li>• 데이터 제출자와 리포지터리 간 계약서</li> <li>• 메타데이터</li> <li>• 데이터 신뢰성</li> <li>• 컬렉션 개발 정책</li> <li>• 리포지터리 선호 형식 목록</li> </ul>
데이터 관리	<ul style="list-style-type: none"> <li>• 큐레이션 수준</li> <li>• 데이터 품질</li> <li>• 워크플로우</li> <li>• 공개 위협이 있는 데이터 관리 정책</li> <li>• 버전 관리</li> <li>• 데이터 변경</li> <li>• 데이터 무결성</li> </ul>
데이터 보존	<ul style="list-style-type: none"> <li>• 마이그레이션 계획</li> <li>• 데이터 보관 프로세스</li> <li>• 접근/보존 연속성 계획</li> </ul>
데이터 공유 및 활용	<ul style="list-style-type: none"> <li>• 데이터 재사용 정책</li> <li>• 라이선스 정책</li> <li>• 공개 데이터 사용 지침</li> </ul>
기타	<ul style="list-style-type: none"> <li>• 접근/보존 사명 선언문</li> <li>• 인프라 개발 계획</li> <li>• 참조 표준</li> <li>• 인력 개발 정책</li> <li>• 전문가 커뮤니케이션</li> <li>• 커뮤니티 정책</li> </ul>

〈표 11〉 CTS 요구 정책 및 기관 정책 비교(예시)

Coretrustseal 요구 정책		기관 정책 및 가이드라인	
조직 인프라	사명 선언문	-	연구기관의 역할 연구데이터 관리 규정
	데이터 수집 정책	데이터 수집 프로세스 데이터 기탁자를 위한 지침 데이터 제공 동의서	연구데이터 수집 및 저장 (승인 및 검토) - -
	데이터 보존 정책	데이터 값 체크리스트	-
	라이선스 정책	지적재산권	연구데이터 이용 협약 [연구데이터]
			연구데이터 관리 가이드라인

정도의 검토 과정을 거쳐 최종적으로 인증이 부여된다. 하지만 R0~R16까지 17개의 문항이 모두 동일한 가중치를 가지고 있기 때문에 어느 하나 답변이 부족하거나 심사위원의 검토에 미

치지 못한다면 인증 부여가 쉽지가 않다. 따라서 모든 문항에 대한 답변을 충실히 작성해야 하며 해당 답변에 대한 근거자료로서 기관의 정책 및 가이드라인 역시 준비되어야 하며 해당

〈표 12〉 CTS 인증을 위한 신청서 작성 예시

구분	신청서 내용	
R1	리포지터리의 임무	<ul style="list-style-type: none"> <li>• 체계적인 연구데이터 유통관리</li> <li>• 효과적인 검색활용 환경 제공</li> <li>• 클라우드 기반의 데이터 분석 지원</li> </ul>
	리포지터리의 필요성	<ul style="list-style-type: none"> <li>• 연구데이터 중요성 증대 국가 경쟁력의 원천으로 대두</li> <li>• 연구데이터 공유 융합을 통한 가치 창출 필요성 증대</li> <li>• 연구데이터 공유 활용을 위한 국가적인 환경 구축 필요</li> </ul>
	리포지터리의 비전	<ul style="list-style-type: none"> <li>• 통합정보시스템의 기능 개선 및 관리</li> <li>• 통합정보시스템에 수집된 자료 또는 정보의 관리 및 제공</li> <li>• 국가연구개발사업의 추진 및 정책수립에 필요한 정보 및 통계자료의 생산 및 분석</li> <li>• 그 밖에 통합정보시스템의 운영 및 관리를 위하여 필요하다고 과학기술정보통신부장관이 인정하는 업무</li> </ul>
R2	사용 중인 라이선스는 무엇인가?	<ul style="list-style-type: none"> <li>• OOO에서는 Creative Commons License를 사용하고 있음</li> </ul>
	해당 라이선스 사용 조건은 무엇인가?	<ul style="list-style-type: none"> <li>• Creative Commons License 중 주로 CC-BY, CC BY-SA, CC BY-ND, CC BY-NC, CC BY-NC-SA, CC BY-NC-ND를 사용하고 있으며 해당 라이선스의 사용 조건은 아래와 같음. CC BY-NC-ND의 경우 (<a href="https://creativecommons.org/licenses/by-nc-nd/4.0/">https://creativecommons.org/licenses/by-nc-nd/4.0/</a>)           <ul style="list-style-type: none"> <li>• 카페와 포맷 변경이 가능</li> <li>• 영리적 목적으로 이용 금지</li> <li>• 상업적 목적으로 이용 금지</li> <li>• 변경 금지</li> </ul> </li> </ul>
R3	데이터 보장기간은 얼마인가?	<ul style="list-style-type: none"> <li>• 데이터는 영구 보장됨</li> </ul>
	데이터 관리 중기(3~5년) 및 장기(5년 초과) 계획이 있는가?	<ul style="list-style-type: none"> <li>• 데이터 관리 계획은 현재 개발 중임</li> </ul>
	데이터 관리 담당자가 있는가? 몇 명인가?	<ul style="list-style-type: none"> <li>• 데이터 관리 담당자는 1명이며, 리포지터리의 데이터 관리 및 보존 업무를 맡는다.</li> </ul>

정책이나 가이드라인이 부족하거나 존재하지 않을 경우 CTS 요구조건에 맞춘 정책 및 가이드라인을 개발해야 할 것이다.

## 5. 결 론

본 연구의 목적은 데이터 리포지터리 인증 체계를 조사 및 분석하고 인증 체계 중 가장 많은 리포지터리를 포함하고 있는 CTS 인증 추

진 전략을 제시하는 것이다. 분석한 인증 체계는 'CTS', 'RatSWD', 'CLARIN certificate B', 'DINI Certificate', 'DIN 31644', 'ISO 16363', 'TDR' 그리고 'Nestor Seal' 등 8개의 데이터 리포지터리 인증에 대하여 조사하였다. 특히, 'CTS', 'RatSWD', 'CLARIN certificate B' 그리고 'DINI Certificate' 등에 대해서는 인증 요구조건을 분석하였다. 이 중 'CTS'는 17개의 요구사항에 존재하며 각각의 요구사항은 가중치가 동일하여 어느 하나 충족하지 못하면 인

증을 받을 수 없다. 또한 CTS는 타 인증 대비 데이터 리포지터리를 정기적으로 관리할 수 있으며 신뢰할 수 있도록 보장하는 단일 지침 세트로 구성되어 있다. 특히, CTS는 확장 및 공식 인증을 위한 프레임워크의 첫 번째 단계로서 3년마다 정기적 갱신을 통해 관리될 수 있으며 타 인증의 요구사항 중 핵심 요구사항을 모두 포함하고 있어 표준 호환성을 가지고 있다 는 장점이 있다.

이에 따라 본 연구에서는 CTS를 중심으로 데이터 리포지터리가 인증 획득하기 위한 전략을 제안하였다. 특히 국내에서는 CTS뿐만 아니라 타 인증까지도 획득한 사례가 전무하기 때문에 인증 획득 관련하여 본 연구에 대한 의의가 있다고 할 수 있다. CTS 인증 추진 전략은 다음과 같다.

- 1단계: CTS 인증을 위한 평가요소 자가 진단 체크리스트를 통한 자체 평가
- 2단계: 신청서 작성을 위한 리포지터리 운영 기관의 연구데이터 관리 정책 및 가이드라인 파악
- 3단계: CTS 인증을 위한 연구데이터 정책 및 가이드라인 개발
- 4단계: CTS 인증을 위한 신청서 작성 및 제출

1단계와 2단계를 통해 기관에 필요한 정책

및 가이드라인을 파악한 다음 데이터 리포지터리를 운영하기 위한 정책 및 가이드라인을 개발하여 마지막으로 신청서 작성 및 제출로서 CTS 인증 추진 전략 프로세스를 제안하였다. CTS 인증을 위해 필요한 정책 등과 예시로 작성한 신청서는 본문에 제시하였으며 도출된 정책과 신청서 예시는 데이터 리포지터리를 운영하고 있는 기관에서 충분히 활용할 수 있다고 판단된다.

국내에서는 최초로 시도한 데이터 리포지터리 인증 체계 분석과 인증 추진 전략이라는 점에서 본 연구가 의의를 가질 수 있다고 판단되지만 구체적이고 실용적인 내용이 부족하다는 점이 본 연구의 제한점이라고 볼 수 있다. 하지만 데이터 리포지터리 인증 추진 과정을 통해 기관의 연구데이터 정책 및 가이드라인을 확보할 수 있을 뿐만 아니라 효과적인 데이터 관리 방안을 수립할 수 있으며 해당 리포지터리 인증으로 연구데이터 서비스의 질적 경쟁력을 확보할 수 있을 것으로 기대할 수 있다.

국내에는 CTS 인증을 포함하여 데이터 리포지터리 인증을 받은 사례가 없다. 비록 인증 전략으로 보기에는 본 연구가 구체적이지 못한 내용을 담고 있지만 앞으로 CTS와 같은 데이터 리포지터리 인증 신청 과정을 통해 부족한 부분을 파악하고 또한 해외 인증 모범사례를 체계적으로 분석하는 등 인증 전략 관련 실용적인 연구가 필요할 것으로 판단된다.

## 참 고 문 헌

- [1] 최희윤, 서태설 (2020). 글로벌 연대와 상생의 길 오픈사이언스. 서울: 한국과학기술정보연구원.
- [2] CLARIN (n.d.). About Us. CLARIN. Available:  
<https://clarin-pl.eu/index.php/en/about-us/>
- [3] CoretrustSeal (n.d.). About. CoretrustSeal. Available: <https://www.coretrustseal.org/about/>
- [4] CoretrustSeal (n.d.). Data Repositories Requirements. CoretrustSeal. Available:  
<https://www.coretrustseal.org/why-certification/requirements/>
- [5] Digital Preservation Coalition (2021). Audit and certification. Digital Preservation Coalition.  
Available:  
<https://www.dpconline.org/handbook/institutional-strategies/audit-and-certification>
- [6] Dillo, I. und de Leeuw, L. (2018). CoreTrustSeal. Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare, 71(1), S. 162-170. doi: 10.31263/voebm.v71i1.1981.. IASSIST Quarterly 43 (1), pp. 1-17. DOI: <https://doi.org/10.29173/iq936>
- [7] DINI Certificate (2016). Open access repositories and publishing services. Deutsche Initiative für Netzwerkinformation. Available:  
<https://dini.de/dienste-projekte/dini-zertifikat/english/about-the-certificate/>
- [8] Dobratz, Susanne & Scholze, Frank (2006). DINI institutional repository certification and beyond. Library Hi Tech. 24. 583-594. DOI: 10.1108/07378830610715446.
- [9] Donaldson DR. (2020). Certification information on trustworthy digital repository websites: A content analysis. PLoS ONE 15(12): e0242525.  
<https://doi.org/10.1371/journal.pone.0242525>
- [10] EuroCRIS (2016). DANS archives receives the NESTOR Seal for Trustworthy Digital Archives. EuroCRIS. Available:  
<https://eurocris.org/news/dans-archives-receives-nestor-seal-trustworthy-digital-archives>
- [11] German Data Forum (2018). The German Data Forum (RatSWD) and Research Data Infrastructure: Status Quo and Quality Management, RatSWD Output, No. 1 (6), Rat für Sozial-und Wirtschaftsdaten (RatSWD), Berlin, <http://dx.doi.org/10.17620/02671.30>
- [12] L'Hours, H., Kleemola, M., & de Leeuw, L. (2019). CoreTrustSeal: From academic collaboration to sustainable services. IASSIST Quarterly, 43(1), 1-17.  
<https://doi.org/10.29173/iq936>
- [13] Mustapha Mokrane & Jonas Recker (2019). CoretrustSeal certified repositories - Enabling

Findable, Accessible, Interoperable, and Reusable (FAIR) Data, 16th International Conference on Digital Preservation. Amsterdam, The Netherlands.

- [14] nestor (2020). nestor. Available:  
[https://www.langzeitarchivierung.de/Webs/nestor/EN/nestor/nestor\\_node.html](https://www.langzeitarchivierung.de/Webs/nestor/EN/nestor/nestor_node.html)
- [15] nestor (n.d.), nestor Seal for Trustworthy Digital Archives. nestor. Available:  
[https://www.langzeitarchivierung.de/Webs/nestor/EN/Zertifizierung/nestor\\_Siegel/siegel.html?nn=182566](https://www.langzeitarchivierung.de/Webs/nestor/EN/Zertifizierung/nestor_Siegel/siegel.html?nn=182566)
- [16] nestor Certification Working Group (2013). Explanatory notes on the nestor Seal for Trustworthy Digital Archives. [https://files.dnb.de/nestor/zertifizierung/notes\\_nestor\\_Seal.pdf](https://files.dnb.de/nestor/zertifizierung/notes_nestor_Seal.pdf)
- [17] Pampel, H., Vierkant, P., Scholze, F., Bertelmann, R., Kindling, M., Klump, J., Goebelbecker, H. J., Gundlach, J., Schirmbacher, P. & Dierolf, U. (2013). Making Research Data Repositories Visible: The re3data.org Registry. *PloS one*. 8. e78080. 10.1371/journal.pone.0078080.
- [18] PTAB (n.d.). Home. PTAB - Primary Trustworthy Digital Repository Authorisation Body Ltd. Available: <http://www.iso16363.org/>
- [19] research data infrastructure of the German Data Forum (2010). Criteria of the German Data Forum (RatSWD) for the Establishment of Research Data Infrastructure. RatSWD. Available:  
[http://www.share-project.org/fileadmin/pdf\\_documentation/RatSWD\\_FDZCriteria\\_1\\_.pdf](http://www.share-project.org/fileadmin/pdf_documentation/RatSWD_FDZCriteria_1_.pdf)
- [20] Wilkinson, M., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., Bonino da S. S., Luiz O., Bourne, P., Bouwman, J., Brookes, A., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C., Finkers, R. & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*. 3. 10.1038/sdata.2016.18.

• 국문 참고자료의 영어 표기

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

- [1] Choi, Hee-yoon & Seo, Tae-Su (2020). Open Science as a Path of Global Solidarity and Coexistence. Seoul: Korea Institute of Science and Technology Information.

