
연구데이터 보존 지침

2025. 1.

KIGAM 한국지질자원연구원
지질자원데이터센터

목 차

1. 개요	1
2. 연구데이터 보존 개념	2
3. 보존 데이터의 선택 및 평가	5
4. 데이터 권리 및 법적 준수	7
5. 데이터 리포지터리	8
부록. 디지털 자산 보존 프레임워크	12

표 목차

〈표 1〉 자료 유형 및 내용	4
〈표 2〉 장기보존 데이터 선택 기준	5
〈표 3〉 지구과학분야 리포지터리	9
〈표 4〉 수준별 디지털 자산 보존 프레임워크	12

1

개요

【 목 적 】

- ☐ 한국지질자원연구원 지질자원데이터센터에서 데이터 보존 시 적용 가능한 연구 데이터 보존 지침(안) 제시
- ☐ 내구성 있는 형식을 선택하고, 데이터를 리포지토리에 제출하여 장기간 보존하는 절차를 따르는 것

【 대 상 】

- ☐ 기탁된 연구데이터를 보존하려는 지질자원데이터센터 관리자

【 적용범위 】

- ☐ 원내에서 연구 활동을 통해 생산된 연구데이터와 외부 기관 및 개인으로부터 기탁된 연구데이터에 적용함

【 준 용 】

- ☐ 본 지침에서 정하지 아니한 사항은 국가과학기술연구회의 연구데이터 관리 가이드라인 및 한국지질자원연구원의 제·규정 준용 가능

【 연구데이터 보존의 의미 】

- ☐ 연구데이터의 장기적 생존(long-term viability)과 지속적인 접근성을 보장하기 위해 취하는 일련의 관리 활동
- ☐ 장기는 저장 매체의 손상, 기술 변경, 이전 및 새로운 매체 및 데이터 형식 지원, 이용자 커뮤니티 변경 등 리포지터리에 보관된 디지털 정보의 무결성 상실에 대해 걱정할 만큼 충분히 긴 시간을 일컬음

【 연구데이터 보존의 필요성 】¹⁾

- ☐ 디지털 데이터 보존은 모든 연구 프로젝트의 핵심 측면이어야 함. 일부 연구데이터는 고유하여 파괴 또는 손실될 경우 대체될 수 없음. 그러나 검증 가능한 데이터를 참조하는 것만으로도 연구가 건전하다고 판단할 수 있음
- ☐ 데이터의 효과적인 문서화
- ☐ 저장 매체의 성능이 저하되거나 자료가 손실될 수 있음
- ☐ 향후 소프트웨어 파일 포맷이 변경될 경우 데이터를 읽을 수 없음
- ☐ 데이터 파일을 위한 문서메이션이 남아있지 않으면 데이터를 이해하기 어려울 수 있음
- ☐ 새로운 소프트웨어로 데이터 파일을 열었을 때 연구를 지속할 수 없을 만큼 이해할 수 없거나 신뢰할 수 없도록 변경될 수 있음
- ☐ 한국지질자원연구원의 연구데이터관리규정 제5장 제16조에 의거 연구데이터의 보존 기간을 영구 보존으로 규정하고 있음

1) University of Leicester. (n.d.). Why preserve digital data?. Retrieved November 20, 2019. Available: <https://www2.le.ac.uk/services/research-data/keep-data/term-pres>

【 데이터 보존의 목표 】²⁾

- ☐ 데이터 관리(Data management): 불가피한 변경을 통해 디지털 레코드가 관리될 수 있도록 보장
- ☐ 접근성(Accessibility): 데이터를 쉽게 찾을 수 있고 접근할 수 있도록 보장
- ☐ 이용 가능성(Availability): 이용자가 필요한 방식으로 데이터 작업 수행 가능
- ☐ 데이터 문서화(Data documentation): 해당 데이터가 어떤 데이터이고 무엇에 관한 데이터인지 이용자의 이해를 도움
- ☐ 무결성(Integrity): 데이터 생애주기(Data Lifecycle) 동안 데이터의 신뢰성 보장

【 데이터 관리 계획과 보존 】

- ☐ 데이터 관리 계획에는 다음과 같은 보존 관련 사항이 명시되어야 함³⁾
 - 데이터 보존 책임 담당자 지정
 - 생산할 데이터 포맷 기술
 - 생산할 데이터셋의 크기
 - 데이터의 저장 위치
 - 연구 분야 또는 연구 기관의 데이터 리포지터리가 존재하는지 기재하고 활용 여부 설명

【 데이터 파일 정리 및 기술 】

- ☐ 데이터 보존은 연구데이터의 장기적 생존(long-term viability)과 지속적인 접근성을 보장하기 위해 취하는 일련의 관리 활동이므로 데이터 파일 정리와 데이터 기술을 포함하는 활동임
- ☐ 연구데이터의 지속적인 접근과 잠재적인 재사용을 고려하여 데이터 파일의 형식(format)은 가능한 범위에서 독점적이지 않고 공개된 표준을 따라야 함
- ☐ 메타데이터와 문서화를 이용하여 보존할 연구데이터를 기술하여야 함

2) BGS. (n.d.). BGS Digital Preservation Policy Retrieved August 7, 2019. Available: <https://www.bgs.ac.uk/downloads/start.cfm?id=3173>

3) USGS. (n.d.). USGS Data Management Plan Checklist. Retrieved August 7, 2019. Available: <https://prd-wret.s3-us-west-2.amazonaws.com/assets/palladium/production/s3fs-public/atoms/files/data-management-checklist.pdf>

□ 다음 <표 1>은 자료 유형에 따른 지침을 나타낸 것임

<표 1> 자료 유형 및 내용

자료 유형	지침 내용 ⁴⁾
데이터 파일	<ul style="list-style-type: none"> • 기계 판독이 가능한 형태, 즉 소프트웨어로 데이터의 개별 내용 또는 내부구조를 확인하거나 수정, 변환, 추출 등 가공할 수 있는 상태의 연구데이터⁵⁾ • 기탁하고자 하는 데이터셋 또는 데이터 파일은 향후 재사용을 고려하여 널리 이용되는 포맷으로 제공하거나 해당 도메인 분야 커뮤니티에서 보편적으로 이용되는 특정 포맷으로 제공할 것을 권장함
도큐멘테이션	<ul style="list-style-type: none"> • 데이터 파일과 함께 데이터 파일을 해석하는 데 필요하거나 설명할 수 있는 도큐멘테이션 파일을 제공해야 함 • 도큐멘테이션 파일의 예로는 코드북, 데이터 수집 도구, 요약 통계, 프로젝트 요약 및 데이터 관련 출판물 목록이 포함될 수 있음 • 그 외에도 다음과 같은 내용을 포함할 수 있음⁶⁾ • 프로젝트 배경 및 목적 • 방법론에 관한 정보 • 이용한 소스 • 관련된 연구 • 샘플링 절차 • 데이터셋의 내용과 구조 • 데이터의 기술(description)과 파일명 목록 • 데이터를 작업하거나 읽을 때 필요한 도구 또는 소프트웨어 • 데이터의 알고 있는 오류 또는 약점에 관한 설명 • 프로젝트와 관련되거나 결과로 출판된 출판물에 대한 레퍼런스 • 레코드, 데이터 변환 또는 형식 변경(format change)에 관한 문서
메타데이터	<ul style="list-style-type: none"> • 데이터 파일과 함께 데이터 파일의 내용을 기술하고 있는 메타데이터를 제공해야 함 • 메타데이터에는 프로젝트 타이틀, 책임 연구자명, 요약, 배포자, 주제어, 지리적 범위, 시간적 범위, 기탁자를 기술함

4) ICPSR(2019). Retrieved November 12, 2019. Available: <https://www.icpsr.umich.edu/icpsrweb/deposit/>

5) 연구데이터관리규정(안)_v2.1 제4장 연구데이터의 등록 및 관리 제12조(연구데이터의 등록)

6) ESSA(n.d.). ESSA guidelines for depositors. Retrieved

【 장기보존 데이터 선택의 필요성 】⁷⁾

- ☐ 데이터 스토리지의 비용이 많이 소요하지 않더라도 모든 데이터를 저장하지 않고 장기보존 데이터를 선택해야 하는 이유는 다음과 같음
- 디지털 데이터의 급격한 증가로 인해 모든 것을 저장하는 비용은 적지 않음
 - 적절한 미러링과 백업 시스템 없이는 디지털 보존 방식을 유지할 수 없으며, 결국에는 백업과 미러링은 보존비용을 증가시킴. 즉, 이는 스토리지 비용을 적어도 2배 이상 증가시킴을 의미함
 - 모든 데이터를 저장한다면 어떤 데이터가 검색의 대상인지 확인하기 위해 추가적인 노력이 필요할 수 있으며 데이터를 선별하여 저장하면 이러한 과정을 줄일 수 있음
 - 많은 데이터 관리 및 보존비용이 소요됨으로 보존 메타데이터를 생성하고 관리하는 비용과 보존해야 할 데이터에 대한 보존비용을 고려해야 함

【 장기보존 데이터 선택 기준 】⁸⁾

- ☐ 자원의 한계로 인해 모든 데이터를 장기 보존하기 어려우므로 다음의 기준을 이용하여 장기보존 가치를 가진 데이터를 선정할 수 있음
- ☐ 다음 <표 2>는 장기보존 데이터 선택 기준을 나타낸 것임

<표 2> 장기보존 데이터 선택 기준

항목	내용
법적 고려사항	<ul style="list-style-type: none"> • 데이터를 보유해야 하는 법적 사유가 있는가? • 소송, 공개 조사, 경찰 수사 또는 법률적으로 이의를 제기할 수 있는 보고서나 논문에 데이터를 사용하였거나 사용할 수 있었는가? • 데이터를 보유해야 하는 재정적 의무 또는 계약적 의무가 있는가? • 논문 실적을 등록한 논문 작성에 사용한 논문 데이터인가?
과학적 가치 또는 역사적 가치	<ul style="list-style-type: none"> • 데이터가 다른 사람들에게 유용한 지리적 범위 또는 시간적 범위를 가지고 있는가? • 데이터가 역사적 가치를 가지고 있는가?(예: 과학적 발견의 랜드마크로 제시될 수 있는가?) • 데이터가 처리 방법, 새로운 표준 또는 전례의 변화를 포함하고 있는가?

7) DCC (2014). 'Five steps to decide what data to keep: a checklist for appraising research data v.1'. Edinburgh: Digital Curation Centre. Available: <http://www.dcc.ac.uk/resources/how-guides>

8) BGS. (n.d.). NGDC Data value checklist. Retrieved August 23, 2019, Available: <https://www.bgs.ac.uk/services/ngdc/documents/DVCNGDC.pdf>

항목	내용
	<ul style="list-style-type: none"> • 데이터가 과학의 트렌드 또는 최신 프로젝트를 지원하고 있는가? • 관련된 과학 분야에서 더 많은 연구가 있을 가능성이 있는가? • 과학계의 미래 요구/방향에 부합할 가능성이 있는가? • 데이터가 광범위한 수집에 기여하고 있는가? • 데이터의 재사용 가능성이 있는가? • 데이터가 출판물에 인용되어 있는가?
원본	<ul style="list-style-type: none"> • 데이터가 고유(unique)한가? • 데이터가 변경되지 않고 기존의 무결성을 유지하고 있는가? • 데이터를 재생산하거나 재수집하는 데 큰 비용을 소모하는가? • 이것이 이 데이터의 원본(primary copy)이라고 생각되는가? • 이 데이터의 사본이 위협에 처해 있는가?
조건	<ul style="list-style-type: none"> • 데이터에 관련 메타데이터가 첨부되어 있는가? • 비과학적 가치 데이터보다 과학적 가치 데이터가 비교적 많은가? • 데이터가 추가적인 처리 과정(감별, 형식 변환 등) 없이 수집될 수 있는가? • 데이터가 컬렉션(collection)에 추가하기에 적합한 상태인가?(즉, 읽기 쉽고, 손상되지 않으며, 다룰 수 있을 만큼 견고한 상태)
저장 및 보존	<ul style="list-style-type: none"> • 특별한 요구사항(디지털 또는 하드카피) 없이 데이터를 저장할 수 있는가? • 특별한 요구사항(디지털 또는 하드카피) 없이 데이터를 보존할 수 있는가?
접근/이용	<ul style="list-style-type: none"> • 지적재산권 또는 저작권에 대한 제한 없이 데이터를 기탁할 수 있는가? • 외부 자원에 의해 부과된 조건이나 기존의 이용 약관 없이 데이터를 기탁할 수 있는가? • 데이터의 이용에 있어 시간적인 제한 없이 데이터를 기탁할 수 있는가?
형식/기술적 제한	<ul style="list-style-type: none"> • 기탁이 허용 가능한 데이터 포맷인가? (예시9) • 데이터가 전문(일반적으로 이용할 수 없는) 소프트웨어 없이 접근 가능한가? • 지질자원데이터센터에서 전문 소프트웨어를 즉시 이용할 수 있는가? • 허용 가능 데이터 형식이 아닌 경우, 데이터를 적절한 스토리지/보존으로 또는 공통 이용 형식으로 전송할 수 있는가?

9) NGDC. (2023.). NGDC Depositing Data - Preferred Digital Formats. Retrieved May 19, 2023 Available: <https://www.bgs.ac.uk/download/ngdc-depositing-data-preferred-digital-formats/>

【 디지털 보존을 위한 권리 보장 】

- 지오빅데이터 오픈플랫폼은 장기 보존을 위해 디지털 객체를 복사, 변환 및 저장하는 데 필요한 권리를 보장하기 위해 구조화된 프레임워크에서 운영되고 있음
 - 데이터 기탁자와의 계약: 데이터 제출 프로세스 동안 기탁자는 데이터 공유, 라이선스 부여 및 변환 및 보존을 허용할 법적 권리를 보유하고 있음을 확인해야 함. 이 확인에는 보존 활동에 부여된 권리를 자세히 설명하는 플랫폼의 이용 약관에 동의하는 것이 포함되어 있음
 - 라이선스 프레임워크: 기탁자는 Creative Commons 라이선스를 포함한 다양한 라이선스 옵션 중에서 선택하거나 플랫폼의 요구 사항에 맞는 사용자 지정 저작권 조건을 지정할 수 있음. 선택한 라이선스는 플랫폼이 형식 마이그레이션 및 메타데이터 향상을 포함한 필요한 보존 활동을 수행할 수 있도록 허용해야 함
 - 정책 준수: 모든 작업은 데이터 처리를 위한 법적 및 윤리적 프레임워크를 설명하는 KIGAM 데이터 관리 규정을 준수함. 특정 규정은 장기 보존 및 관련 권리를 다루고 있음

【 디지털 보존을 위한 권리 보장 】

- 데이터 기탁은 기탁자가 플랫폼에서 데이터를 관리, 변환 및 저장할 수 있는 충분한 권한을 부여했는지 확인하기 위해 엄격한 검토 프로세스를 거치고 있음. 규정을 준수하지 않는 제출은 거부되거나 수정을 위해 기탁자에 반환될 수 있음

【 예치자 및 사용자 교육 】

- 투명성을 보장하기 위해 기탁자에게 제출된 데이터에 대한 적절한 라이선스 및 권한을 보장하는 것을 포함하여 책임에 대한 자세한 지침이 제공되고 있음. 사용자가 액세스하는 데이터의 라이선스 약관을 이해하는 데 도움이 되는 교육 리소스도 제공할 수 있음

【 데이터 리포지터리의 정의 】

- ☐ 리포지터리는 온라인 데이터베이스 서비스로, 디지털 자원의 장기 저장과 보존을 관리하고 탐색과 접근을 위한 카탈로그를 제공하는 아카이브임¹⁰⁾

【 데이터 리포지터리의 선택 시 고려사항 】¹¹⁾

- ☐ 제출한 데이터셋에 영구적인 식별자 제공
- ☐ 데이터셋 탐색 후 해당 데이터셋의 내용을 확인하고 이용할 수 있도록 지원하는 메타데이터를 데이터셋의 랜딩페이지로 제공
- ☐ 데이터 이용에 관한 추적 지원
- ☐ 커뮤니티 요구에 대응하거나 ‘신뢰할 수 있는 데이터 리포지터리’로 인증됨
- ☐ 데이터 보호와 같은 법적 요건을 충족하고 불필요한 라이선스 조건 없이 데이터 재사용 가능

【 데이터 리포지터리의 예시 】

- ☐ 범용 리포지터리
 - FigShare(<http://figshare.com>)
 - Dryad(<https://datadryad.org>)
 - Zenodo(<http://zenodo.org/>)
 - DataHub(<http://datahub.io>)
 - DANS(<http://www.dans.knaw.nl/>)

10) The University of Sheffield. (n.d.). Research data repositories. Retrieved November 15, 2019.
Available: <https://www.sheffield.ac.uk/library/rdm/repositories>

11) DCC (2014). 'Five steps to decide what data to keep: a checklist for appraising research data v.1'.
Edinburgh: Digital Curation Centre. Available: <http://www.dcc.ac.uk/resources/how-guides>

□ 다음 <표 3>은 지구과학분야 리포지토리를 나타낸 것임

<표 3> 지구과학분야 리포지터리

리포지터리	설명
National geoscience data centre (NGDC)	<ul style="list-style-type: none"> National Geoscience Data Centre(NGDC)는 영국의 BGS(British Geological Survey)의 데이터 셋을 관리하는 리포지터리로서 지구과학적 데이터와 정보를 수집하고 보존하여 커뮤니티에서 장기적으로 사용할 수 있도록 함 http://www.bgs.ac.uk/services/ngdc/
Centre for Environmental Data Analysis (CEDA)	<ul style="list-style-type: none"> CEDA는 영국 대기 과학 및 지구 관측 커뮤니티를 위해 NERC((Natural Environmental Research Council)를 대신하여 대기 및 지구 관측 데이터 센터 기능을 운영함 https://www.ceda.ac.uk/
UK Polar Data Centre(UK PDC)	<ul style="list-style-type: none"> UK Polar Data Center(UK PDC)는 영국의 북극 및 남극 환경 데이터 관리의 중심점으로 NERC(Natural Environmental Research Council)의 환경 데이터 센터 네트워크의 일부임 https://www.bas.ac.uk/data/uk-pdc/
PANGAEA	<ul style="list-style-type: none"> PANGAEA는 지구, 환경 및 생물다양성 과학의 지리 참조 데이터를 보관, 게시 및 배포하기 위한 개방형 액세스 라이브러리로서 약 30년의 역사를 가지고 있음 https://www.pangaea.de/
TOAR Surface Observation Database	<ul style="list-style-type: none"> TOAR(Tropospheric Ozone Assessment Report) 데이터베이스는 세계에서 가장 광범위한 표면 오존 측정 데이터베이스임 https://toar-data.fz-juelich.de/
Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC)	<ul style="list-style-type: none"> ORNL DAAC(Oak Ridge National Laboratory Distributed Active Archive Center)는 NASA(National Aeronautics and Space Administration) 지구과학 데이터 및 정보 시스템에서 관리하는 EOSDIS(Earth Observing System Data and Information System) 데이터 센터 중 하나임 https://daac.ornl.gov/
Norwegian marine data centre (NMD)	<ul style="list-style-type: none"> NMD(Norwegian Marine Data Centre)는 해양 환경 및 수산 데이터의 전문 처리 및 장기 저장과 데이터 제품 생산을 전담하는 국가 데이터 센터임 https://www.hi.no/en/hi/forskning/research-groups-1/the-norwegian-marine-data-centre-nmd
International Council for the Exploration of the Sea (ICES)	<ul style="list-style-type: none"> ICES는 정부 간 해양 과학 기구로 해양 환경과 그 생물 자원에 대한 과학적 지식을 높이고 공유하며 이를 활용하기 위해 보존, 관리 및 지속 가능성에 대한 조언을 제공함 https://ecosystemdata.ices.dk/

리포지터리	설명
ICTS SOCIB Data Repository	<ul style="list-style-type: none"> ICTS SOCIB(Balearic Islands Coastal Observing and Forecasting System)은 해양 데이터 제품 스트림과 모델링 서비스를 제공하는 다중 플랫폼 분산 및 통합 시스템임 https://www.socib.es/data/

【 연구데이터의 저장 】

- ☐ 저장할 연구데이터 사본의 개수와 동기화 방법 제시
- ☐ 연구데이터 저장을 위한 스토리지 제공
- ☐ 클라우드 기반 서비스에 저장된 데이터 전송 시스템의 백업을 위한 사본 사이트 보유
- ☐ 정전 시 백업 사이트에서 데이터 다운로드 서비스 제공
- ☐ 스토리지(저장공간) 솔루션 비교 기준 제시
- ☐ 데이터 백업 시 무결성 및 접근성 보장

【 백업 및 복구】

- ☐ 데이터의 손실 및 손상을 방지하고 보호하기 위하여, 연구자는 데이터가 여러 위치에 정기적이고 자동으로 백업하여야 할 책임이 있음
- ☐ 지오빅데이터 오픈플랫폼의 백업시스템은 InnoStor Appliance(ISA-2000)와 Quantum Scalar i500로 이중 백업으로 구성되며, 서비스 스토리지로부터 주기적으로 백업된 데이터를 백업시스템에 저장하여 백업을 수행함
 - 백업 대상: 지오빅데이터 오픈플랫폼의 데이터, 데이터베이스 및 사용자 데이터 파일에 대한 백업 수행
 - 백업 주기
 - 사용자 데이터파일, 데이터베이스 및 시스템 데이터 백업: 매일
 - 사용자 데이터, 데이터베이스 및 연구데이터(파일) 전체 백업: 토요일
- ☐ 복구 정책 및 지침
 - 시스템 소프트웨어 및 응용 소프트웨어는 로컬 GIT 리포지터리로 부터 복구 수행
 - 연구데이터, 데이터베이스, 사용자 데이터 파일에 대한 복구는 백업장치에 저장된 데이터로 복구 수행
 - 백업된 데이터가 문제가 발생하였을 경우, 소산지의 테이프 백업을 통한 복구 수행

【 연구데이터 아카이빙 및 보존 】

- ☐ 연구데이터의 보존을 목적으로 주기적으로 연구데이터를 아카이빙(자기테이프)하여 보존함
- ☐ 아카이빙 및 보존(Archiving and Preservation) 방법: 테이프 백업을 통하여 백업 테이프를 원격지에 소산 및 아카이빙하여 보존(yearly)
- ☐ 아카이빙(Archiving) 테이프의 보존 기한은 최소 5년 이상으로 보존함

【 기술 및 절차 안정성을 위한 보존 전략 】¹²⁾

- ☐ 마이그레이션(Migration): 비교적 일반적이지 않은 파일 혹은 더 이상 이용되지 않는 파일 포맷에서 현행 파일 포맷으로의 파일 포맷 변환
- ☐ 에뮬레이션(Emulation): 구형 또는 더 이상 이용되지 않는 컴퓨터의 기능을 모방하는 것으로 에뮬레이션을 이용하면 컴퓨터에서 이전 파일 형식을 읽은 다음 현재 파일 형식(에뮬레이션과 마이그레이션을 결합)으로 저장하거나 향후 오래되고 더 이상 이용되지 않는 파일을 읽고 이용하기 위한 기술
- ☐ 정규화(Normalization): 보존을 위해 데이터 포맷을 일반적인 포맷 형식(예: 텍스트 파일을 개방형 문서 포맷 또는 Word 포맷으로 제한)으로 제한하거나 소프트웨어 종속 파일 포맷을 변환한 소프트웨어 종속 파일 포맷(예: SPSS 시스템 파일) 또는 소프트웨어 독립파일 포맷(예: ASCII 또는 XML 기반 포맷)으로 변환

12) Inter-university Consortium for Political and Social Research (ICPSR). 2009. "Principles and Good Practice for Preserving Data", International Household Survey Network, IHSN Working Paper No 003, December 2009.

【 디지털 자산 보존 프레임워크 개요 】¹³⁾

- 디지털 자산 보존 프레임워크는 2013년 National Digital Stewardship Alliance가 제시한 것으로 다음의 <표 4>를 사용하여 디지털 보존 수준을 평가할 수 있음
- 해당 부록은 디지털 자산 보존 수준을 평가하기 위한 안내로 리포지터리의 보존 상태를 평가하여 향후 수준을 어디까지 높여야 하는지 연도별로 제시할 수 있음

<표 4> 수준별 디지털 자산 보존 프레임워크

요소	레벨 1 (데이터 보호)	레벨 2 (데이터 인식)	레벨 3 (데이터 모니터링)	레벨 4 (데이터 복구)
스토리지 & 지리적 위치	<ul style="list-style-type: none"> 물리적으로 서로 분리되어 저장된 두 개의 전체 복사본 이기종 미디어(광학 디스크, 하드 드라이브 등) 위의 데이터의 경우 해당 매체에서 스토리지 시스템으로 콘텐츠 전송 	<ul style="list-style-type: none"> 최소 3개의 완전한 복사본 최소 1개의 사본이 다른 지리적 위치에 존재 스토리지 시스템, 스토리지 미디어 및 스토리지 사용 시 필요한 것 문서화 	<ul style="list-style-type: none"> 다른 재해 위험이 있는 지리적 위치에 하나 이상의 사본(예: 허리케인 지역 대 지진 지역) 스토리지 시스템 및 미디어에 대한 노후화 모니터링 프로세스 유지 	<ul style="list-style-type: none"> 서로 다른 재해 위험이 있는 지리적 위치에 최소 3개 이상의 사본 현재 접근이 가능한 시스템 및 미디어에 파일과 메타데이터를 보관하는 포괄적인 계획 보유
파일 고정성 및 데이터 무결성	<ul style="list-style-type: none"> 수집 시 파일 무결성 확인(제공된 경우) 제공되지 않은 경우, 체크섬(checksum) 생성 모든 콘텐츠에 대한 바이러스 검사 	<ul style="list-style-type: none"> 모든 데이터 수집 시 무결성 검사 원본 미디어 작업 시 읽기 전용 고위험 콘텐츠 바이러스 검사 	<ul style="list-style-type: none"> 일정한 간격으로 무결성 검사 무결성 로그 유지; 필요에 따라 감사 정보를 제공 손상된 데이터를 감지하는 절차 유지 모든 콘텐츠에 대한 바이러스 검사 	<ul style="list-style-type: none"> 특정 이벤트 또는 활동에 대응하여 모든 콘텐츠의 무결성 검사 손상된 데이터를 교체하거나 수리하는 절차 유지 한 사람이 모든 사본에 대한 쓰기 접근 권한을 가지고 있지 않은지 확인
정보 보안	<ul style="list-style-type: none"> 개별 파일의 읽기, 	<ul style="list-style-type: none"> 콘텐츠에 대한 문 	<ul style="list-style-type: none"> 삭제 및 보존 작 	<ul style="list-style-type: none"> 로그 감사 수행

13) Phillips, M., Bailey, J., Goethals, A., & Owens, T. (2013, January). The NDSA levels of digital preservation: Explanation and uses. In Archiving Conference (Vol. 2013, No. 1, pp. 216-222). Society for Imaging Science and Technology.

요소	레벨 1 (데이터 보호)	레벨 2 (데이터 인식)	레벨 3 (데이터 모니터링)	레벨 4 (데이터 복구)
	쓰기, 이동 및 삭제 권한이 있는 이용자 식별 • 개별 파일에 대한 권한 제한	서 접근 제한	업을 포함하여 파일에 대한 작업을 수행한 이용자의 로그 유지 관리	
메타 데이터	• 콘텐츠의 인벤토리 및 스토리지 위치 • 인벤토리 정보의 백업 및 물리적 분리 보장	• 관리 메타데이터 저장 • 변형(transformative) 메타데이터 및 로그 이벤트 저장	• 표준 기술 (technical) 및 설명(descriptive) 메타 데이터 보존	• 표준 보존 메타데이터 저장
파일 형식	• 디지털 파일 작성에 입력할 수 있는 경우 알려진 공개 형식 및 코 덱의 제한된 사용권장	• 사용 중인 파일 형식 인벤토리	• 더 이상 지원되지 않는 파일 형식 모니터링	• 포맷 마이그레이션, 에뮬레이션 및 유사한 작업 수행

버전	일자	내용
0.1	2023. 03. 20.	문서 Outline 작성
0.6	2023. 04. 28.	분야별 초안 작성
0.8	2023. 05. 08.	지침 기관 검토
1.0	2023. 05. 19.	검토 의견 반영 및 보완
1.1	2025. 01. 15.	검토 의견 반영 및 보완(CTS)