SERVIR West Africa 2024 Training Program on Data Stewardship and the CoreTrustSeal Requirements: Session 5, R03 Requirement

Robert R. Downs, PhD

Center for International Earth Science Information Network (CIESIN)

Columbia Climate School, Columbia University

Coordinated by AFRIGIST

July 10, 2024, 13:00 – 14:30 GMT (9:00 - 10:30 a.m. EDT)

Brief Review of Previous Webinars

- Data Stewardship Concepts, Principles, and Certification Instruments
  - CoreTrustSeal, ISO 16363, Nestor DIN 31644
- Preservation, Designated Community, Sustainability
- CoreTrustSeal Certification requirements
- Organizational aspects of Data Stewardship
- CoreTrustSeal Requirements 2020-2022: R01-R16
- Setting the stage for Planning to Meet the CoreTrustSeal Requirements
- Motivation and R0 Requirement
- R0 Requirement and Introduction to R01 Requirement
- R01 Requirement and Introduction to R02 Requirement
- R02 Requirement and Introduction to R03 Requirement

SERVIR West Africa 2024 Training Program
on Data Stewardship and CoreTrustSeal Requirements

- Key topic:
  - R03 Requirement

- Instructional format
  - Lecture, questions, and discussion

- Session will be recorded
  - Recording will be posted on the SERVIR West Africa website
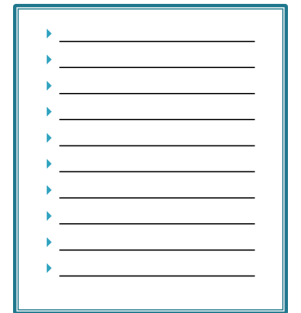
Outline for Current Session

- Progress Planning and Participant Introductions (5 minutes)

- Scheduling Training Sessions and CoreTrustSeal Application Preparation (5 minutes)

- Review of CoreTrustSeal R0, R01, R02, and homework (35 minutes)

- CoreTrustSeal Requirement 03 (30 minutes)

- Introduction to CoreTrustSeal Requirement 04 (10 Minutes)

- Preparing for Next Session (5)
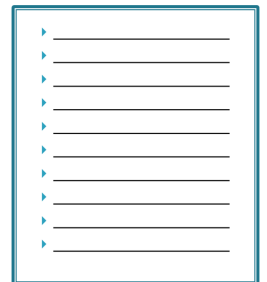
Progress Planning and Introductions

- CoreTrustSeal application template created for each repository
  - CERSGIS, ICRISAT FMNR, AFRIGIST, AGRHYMET
- Each repository uses template to progressively revise their self-assessment
- Please put additional repository names and your email address in Zoom chat
- Brief introductions by those who have not previously introduced themselves

Scheduling Training Sessions and CoreTrustSeal Application Preparation

- Plan to schedule sessions twice monthly enabling time for progress
  - Repositories complete template sections between sessions
  - Each session includes review and introduction to subsequent section
- Next session planned for Wednesday, July 31, 2024, 13:00-14:30 GMT
  - Topics: R04 Requirement and Introduction to R05 Requirement

Review of CoreTrustSeal Requirement R0 Context

Overview of CoreTrustSeal Requirements 2023-2025: Context

R0 Context

- General information about the repository

- Re3data identifier

- Repository Type. Select all relevant types from:
  - Generalist or Specialist; Specialist describes scope

- Overview – Brief Description of Repository

- Brief Description of the Designated Community

- Level of Curation Performed. Select all relevant types from:
  - Content distributed as deposited Basic curation – e.g., brief checking, addition of basic metadata or documentation. Enhanced curation – e.g., conversion to new formats during ingest, enhancement of documentation or metadata; Data -level curation – as in C above, but with additional editing of deposited data for accuracy

- Cooperation and Outsourcing to third parties, partners, and host organizations.

- Applicants renewing – Summary of Significant Changes Since Last Application

Discussion of Previous R0 Assignment

- Review requirements for R0 within the context of your repository

- Discuss among managers and staff within your repository

- Draft description for an re3data identifier for your repository
  - Requirements: https://www.re3data.org/suggest
  - Search for descriptions of other repositories: https://www.re3data.org/search

- Draft responses to R0 for your repository within your template

CoreTrustSeal Requirement 0, Background Information and Context

- Re3data identifier
  - Re3data is a **Re**gistry of **Re**search Data **Re**positories
  - Internationally recognized global registry of repositories
  - Assists potential depositors, users, and others in discovering data repositories
  - Users of Re3data can search for a repository by entering a topic as a query
- Repositories are described in the Re3data registry
  - Title of repository
  - Subject(s) covered by the content of the repository
  - Content type(s)
  - Country where the repository is located
  - Brief narrative description of the repository

CoreTrustSeal Requirement 0, Background Information and Context

- Repository Type. Select all relevant types from:
  - Select a repository type (Generalist vs. Specialist - specialist must specify domain and/or disciplines(s))
  - The repository needs to describe the disciplines represented by the content that the repository disseminates. Additional information related to the discipline(s) represented provides context for the reviewers to understand the repositories role in supporting the represented disciplines.

- Overview – Brief Description of Repository
  - *Provide a short overview of key characteristics of the repository, reflecting the repository type selected. This should include information about the scope and size of data collections, data types and formats. Further contextual information may also be added.*

- # Brief Description of the Designated Community
  - *A clear definition of the Designated Community demonstrates that the applicant understands the scope, knowledge base, and methodologies-including preferred software/formats-of the group(s) of users at whom the curation and preservation measures are primarily targeted. The definition should be specific so that reviewers can assess whether that community is being served in the responses to other requirements.*

- Level of Curation Performed. Select all relevant types from:
  - A Content distributed as deposited
  - B Basic curation – e.g., brief checking, addition of basic metadata or documentation.
  - C Enhanced curation – e.g., conversion to new formats during ingest, enhancement of documentation or metadata;
  - D Data -level curation – as in C above, but with additional editing of deposited data for accuracy

- Level of Curation Explanation
  - Please add the description for your levels of curation.
  - The repository needs to describe how the level of curation or levels of curation are managed, as this will inform reviewers how the repository completes curation. Reviewers will look at the Curation Explanation frequently when reviewing responses to other requirements.

CoreTrustSeal Requirement 0, Background Information and Context
- Continued -

- (6) Cooperation and outsourcing to third parties, partners and host organizations.
    - List or describe insource partners within the host organization that contribute to the repository and its operations.
    - List or describe outsource partners that are external to the host organization that contribute to the repository and its operations.

Note: Insource partners may offer support, services, or guidance that the repository leverages to serve its Designated Community.
Outsource partners may be sponsors, contractors, or vendors that support the repository's services.

- **(7) Applicants renewing their CoreTrustSeal Certification**
  - *Summary of significant changes since last application. CoreTrustSeal certification has an expectation of continuous improvement over time. Repositories undergoing recertification should highlight briefly any significant changes including to technical systems, Designated Community or funding during the previous three years. This could include any steps taken to move from 'In Progress' to 'Implemented' Requirements since the last certification.*
  - *Only applicable to repositories that have receive CoreTrustSeal Certification*

This question should only be answered by repositories that are renewing CoreTrustSeal Certification.

Source: CoreTrustSeal Standards and Certification Board. (2022). CoreTrustSeal Trustworthy Digital Repositories Requirements 2023-2025 Extended Guidance (V01.00). Zenodo. https://doi.org/10.5281/zenodo.7051096

Review of Requirement R01 Mission and Scope

R01, Mission & Scope

- R01. The repository has an explicit mission to provide access to and preserve digital objects.

- Repositories take responsibility for the curation of digital objects, and for ensuring that materials are held in the appropriate environment for appropriate periods of time. For Trustworthy Repositories it must be clear to depositors and users that active preservation of and continued access to the digital objects is an explicit role of the repository.

- The response statement and evidence should include references to the following items:

    The mission to actively preserve and provide access to digital objects

    The level of approval that the mission has received.

- Evidence for this Requirement could include an approved public mission statement, roles mandated by funders, or a policy statement signed off by a governing board.

R01, Mission and Scope (Continued)

If preservation is not referred to in the mission of the repository or other relevant public documents provided as evidence, then the compliance level cannot be higher than "In Progress: the repository is in the implementation phase".

Example: SEDAC Mission Statement and Data Submission Guidance

SEDAC, the Socioeconomic Data and Applications Center, is one of the Distributed Active Archive Centers (DAACs) in the Earth Observing System Data and Information System (EOSDIS) of the U.S. National Aeronautics and Space Administration. Focusing on human interactions in the environment, SEDAC has as its mission to develop and operate applications that support the integration of socioeconomic and earth science data and to serve as an "Information Gateway" between earth sciences and social sciences.

There are many benefits to publishing your data in an open data repository. By publishing through a reputable domain-specific repository such as SEDAC, you will likely increase the citations of your published work and your data. SEDAC's process of quality assurance (QA) and documentation preparation adds value and may catch errors in your data or documentation. Users can more readily discover, understand and use your data both in the near future and in the long-term. Sponsors maximize their investment in data and knowledge generation and increase scientific transparency and replicability. In addition, scientists outside your project or discipline will be able to find, understand, and use your data to address additional questions, potentially in new or interdisciplinary research areas. As part of its commitment to support community efforts to understand human interactions in the environment, SEDAC preserves the data and related information resources that it disseminates to ensure their continuing access and use. Preservation of data and related information resources is accomplished by actively engaging in archiving, curation, and ongoing stewardship to support the current and future community of users in their quest to discover, explore, access, use, and cite the data and information resources that are managed by SEDAC.

Sources: https://sedac.ciesin.columbia.edu/about and https://sedac.ciesin.columbia.edu/data-submission

Example: LAADS DAAC R01 Response

Organisational Infrastructure

R1 Mission & Scope (R01)

R01. The repository has an explicit mission to provide access to and preserve digital objects.

Compliance level:

Implemented: the requirement has been fully implemented by the repository - 1

Response:

As part of the larger EOSDIS ecosystem (https://earthdata.nasa.gov/about/daacs), the LAADS DAAC's mission is to serve as the steward for satellite-derived atmospheric science data on a global scale. It provides different levels of standard geophysical data products to address critical Earth system science investigations that have been vetted, quality-checked, and validated via established protocols and processes. As part of serving a varied, global user community, the LAADS DAAC provides mechanisms, tools, and services to order, acquire, interact with, understand, and use its atmospheric science product portfolio. Earth system science needs and requirements often cross disciplinary boundaries, and LAADS DAAC is well-positioned to serve users in such unique situations. For additional details, check the following URL: https://ladsweb.modaps.eosdis.nasa.gov/about/

As a NASA EOSDIS DAAC, LAADS is governed by the "Requirements for Archiving, Distribution and User Services in EOS Data and Information System (EOSDIS)" available via the following URL: https://cdn.earthdata.nasa.gov/conduit/upload/11349/423-10-69_ADURD_RevB.pdf

Example: AUSSDA DAAC R01 Response

R1 Mission & Scope (R01)

R01. The repository has an explicit mission to provide access to and preserve digital objects.

Compliance level:

Implemented: the requirement has been fully implemented by the repository - 1

Response:

AUSSDA's mission statement states its main objective: "We make social science data accessible and reusable." [1] and continues as follows: "AUSSDA - The Austrian Social Science Data Archive is a certified, national research infrastructure for the social science community. We offer sustainable and easy-to-use services in the field of digital archiving and preservation. The main beneficiaries are researchers, students, educational institutions and media professionals.

We implement international standards to make research data findable, accessible, interoperable and reusable according to the FAIR principles. AUSSDA supports the open science movement to maximize the potential for data reuse. We stand for integrity in archiving and advocate for compliance with data protection and ethical principles in research data management.

AUSSDA represents Austria as a national service provider in CESSDA ERIC, has locations at the Universities of Vienna, Graz, Linz and Innsbruck and works within a network of national and international partners." [1]

The mission statement was approved by AUSSDA's steering board ("Leitungsgremium"). The steering board consists of members of the AUSSDA consortium universities based in Vienna, Graz, Linz and Innsbruck, and a representative of the Austrian Federal Ministry of Education, Science and Research (see R8: Expertise & Guidance).

Links:

- [1] Mission Statement

Discussion of Previous R01 Assignment

- Review requirements for R01 within the context of your repository
- Discuss among managers and staff within your repository
- Draft responses to R01 for your repository within your template

Questions for Discussion

- Is the mission statement for the repository publicly accessible?

- What authority has approved the repository's mission statement?

- Does the mission statement or the repository's website describe preservation?

R02 Requirement Rights Management

Rights Management (R02)

- R02. The repository maintains all applicable rights and monitors compliance.

- The repository manages, and communicates to relevant stakeholders, all rights (permissions, prohibitions, obligations) covering data and metadata deposit, storage, preservation, access, and use.

- This requirement relates to the system, methods and artefacts (e.g. licenses, agreements, terms and conditions, and related policies and procedures) in place for rights management.

- The repository must obtain all necessary rights from the depositor, and demonstrate that there are sufficient controls in place to ensure they are applied and monitored.

- The response statement and evidence should include references to the following items:

  - The overall rights management approach to deposited files, data and metadata.
  - The rights to copy, transform, and store digital objects for preservation, as well as provide access to them
  - Conditions of use (e.g. intellectual property rights, distribution, intended use, protection of sensitive data, etc.).
  - Deposit and access agreements or licenses.
  - How rights metadata is managed for humans (e.g. license documents/files) or machines.
  - Monitoring of compliance at deposit, during curation/preservation, and during access and reuse. Describe any circumstances where compliance monitoring is not possible.
  - Measures in place if non-compliance is detected.

Rights Management (R02) – Continued

- Data and meta data, including 'open data', will usually have some rights attached even if there is no signed license artefact or formal agreement in place. This could include obligations such as citation and attribution of data and metadata used, or making secondary analysis openly available. If all data and metadata are made available without any conditions of access or use then this should be made clear in the response statement.

- Rights negotiations and transfer should be described under Deposit & Appraisal (R08). Any ethical codes of conduct, privacy measures, or legislation that influence rights management should be described under Legal & Ethical (R04).

Rights Management (R02) – Continued

- Stipulations on data access and use could be defined in a set of standard terms and conditions, or differentiated by depositor or type of digital object. Popular licence options include, but are not limited to, those offered by Creative Commons (https://creativecommons.org/) such as 'CC 0 Waiver' and 'public domain data' licences.

- For sensitive data, in particular, licences may specify limitations on use, usage environment (safe room, secure remote access), and types of users (approved researcher, minimum training requirements, etc.). More recently, the Local Contexts Traditional Knowledge and BioCultural Notices and Labels have emerged as a means for Indigenous peoples to allocate customizable statements of provenance, protocols and permissions to research objects.6

- While it may be challenging to identify instances of noncompliance , consideration should be given to the consequences if noncompliance is detected e.g. sanctions on current or future access/use of data and metadata. In the case of sensitive personal data disclosure, there may be severe legal penalties that impact both the user and repository. Ideally, repositories should have a public policy in place for noncompliance.

- For applicants that hold data or metadata with a disclosure risk, the target compliance level should be "Implemented: the requirement has been fully implemented by the repository".

Questions for Discussion

- How might the repository's approach to rights management be described?

- How might we describe the repository's approach to managing conditions for use of any data or information that might be considered sensitive or protected?

- How might we describe the repository's approach to licensing agreements or other formal agreements used to secure the rights of resources that are deposited, managed, and disseminated?

Rights Management (R02) Example from SEDAC: CIESIN Open Data Policy (Excerpt)

This CIESIN Open Data Policy has been established to promote the open and free exchange of data and information in support of research, decision making, education, and other applications. CIESIN intends this policy to support and comply with relevant open data policies, guidelines, and initiatives for scientific, government, and sustainable development data.i CIESIN's policy is to make data "open by default," with only narrow exemptions for genuine security, privacy, or legal concerns.

The data products covered by this policy include quantitative and textual data, digital maps, images, and other visualizations that CIESIN creates and publicly disseminates online, as well as CIESIN-authored documentation, metadata, and software code or scripts that are directly relevant to the data products (henceforth referred to as Data). In many instances, CIESIN disseminates Data created by, or in collaboration with, external individuals or organizations; for these Data, CIESIN will make best efforts to obtain permissions or releases in advance of the collaboration to ensure open access to these Data products. When Data are solely authored or

Rights Management (R02) Example from SEDAC: Description of Data Licenses for Data Producers

Data Licenses

SEDAC supports open scientific data, per the WDS Data Sharing Principles. We request that data providers accept the dissemination of their data under the Creative Commons Attribution 4.0 International License (CC-BY-4.0) unless there are extenuating circumstances such as data restrictions inherited from input data. Providers need to clearly indicate if the data are in the public domain, have a CC-BY license or equivalent fully open license, or have some other license applied to them.

SEDAC can work with data providers if, under certain circumstances, they are unable to provide the data under the CC-BY license, but under no circumstances can SEDAC publish the data in the absence of a signed statement granting SEDAC permission to disseminate the data. If your data are accepted for dissemination by SEDAC, you will need to download CIESIN's Open Data and Information Agreement, fill out the information required, and submit a signed copy to ciesin.info@ciesin.columbia.edu.

When completing the agreement, SEDAC recommends composing a concise data set title (125 characters or less) that is memorable, contains scientific context, geographical coverage (e.g. global) temporal extent (e.g. 1920-2020), and a version number if the version is greater than 1 (e.g. Version 2.01). Examples of data set titles can be viewed on the SEDAC website (https://sedac.ciesin.columbia.edu/) under Data.

Source: https://sedac.ciesin.columbia.edu/data-submission

Rights Management (R02)
Example from SEDAC:
CIESIN Open Data and
Information Agreement
(Excerpt)

COLUMBIA CLIMATE SCHOOL
CENTER FOR INTERNATIONAL EARTH SCIENCE INFORMATION NETWORK

Open Data and Information Agreement

This form is required to archive, preserve, disseminate, acknowledge authorship, and promote discovery of open data and information in accordance with the CIESIN Open Data Policy (see: http://www.ciesin.columbia.edu/documents/CIESINDataPolicy.pdf).

Title of Data or Information: _____

Version/Edition/Date: _____

Author(s) (Full name): _____

Brief Description: _____

1. My co-creators, if any, have authorized me to enter into this agreement on their behalf.

2. I warrant that my co-creators, if any, and I are the sole owner(s) of the above-referenced Data or Information. To the best of my understanding, I warrant that the Data or Information are original, do not infringe on the copyright or rights of any third parties, do not contain disparaging, sensitive, private, personal, confidential, or unlawful content, and are a truthful representation of the facts.

3. I agree that the listed Data or Information may be distributed under a Creative Commons Attribution 4.0 International License, available at http://creativecommons.org/licenses/by/4.0 and specified in legal code at http://creativecommons.org/licenses/by/4.0/legalcode.

Signature: _____ Date: _____

Rights Management (R02) Example: Excerpt from AADC R02 Response

The AADC Data Policy (https://data.aad.gov.au/about/data-policy) lays out the rights, obligations and responsibilities of researchers participating in the Australian Antarctic Science Program (AASP).

The Conditions of Use page (https://data.aad.gov.au/about/conditions-of-use) provides information to users of our data.

All of our metadata records/datasets are also labelled with a CC-BY 4.0 licence - see the "Use Constraints" section of our metadata. For example:

- https://data.aad.gov.au/metadata/records/AAS_4590_Rock_Texture
- https://data.aad.gov.au/metadata/AAS_4590_Rock_Texture (new website - still undergoing development).

Essentially:

1) All data collected as part of the AASP remain the property of the Commonwealth of Australia.

2) All data collected as part of the AASP are expected to be catalogued and archived in the AADC (or an alternative approved repository).

3) All data collected as part of the AASP are expected to be made publicly available within an appropriate timeframe.

4) Anyone using data catalogued and archived in the AADC are expected to correctly cite the data, and provide appropriate credit to the listed researchers.

For more information, see the Data Policy, and Conditions of Use.

Rights Management (R02) Example: Excerpt from ICSD R02 Response

Access to the ICSD data is subject to a license agreement between the customer and FIZ Karlsruhe. Licenses distinguish between single users and multiple-users or whole entities. The regular licenses offer access and usage of the whole data for one subscription year, but include also a limit on the download of data, whereas the usage and download of the whole dataset is covered by the API license. The user interface has mechanisms in place to restrict the download of more than 1000 datasets. The number of downloads per institution is monitored in ICSD Web. Furthermore there are multi-annualand perpetual licenses available, as well as consortial licenses. The following ICSD versions are available: ICSD Web, ICSD Desktop, ICSD Intranet, ICSD API

ICSD WEB

ICSD Web is a service enabling web access to ICSD. Its servers run at FIZ Karlsruhe and can be accessed with a common browser and an appropriate license. Authentication for single users and multiple users is handled by logins/passwords. Campus/Site users are authenticated based on their IP address.

ICSD Desktop

ICSD Desktop is a Windows-based solution for individual users as well as small research groups that would like to install ICSD locally on their computer/laptop. ICSD Desktop installs stripped-down servers to run the required services on the local machine. Access from other machines is prohibited. Access via a network is not possible.

ICSD Intranet

ICSD Intranet is an in-house solution based on the ICSD Web technology. It allows users to offer their own web-based access to ICSD within their university or company. Authentication is performed via a separate LDAP module, but can also be integrated into an existing LDAP solution.

Discussion of Previous R02 Assignment

- Review requirements for R02 within the context of your repository
- Discuss among managers and staff within your repository
- Draft responses to R02 for your repository within your template

CoreTrustSeal Requirement R03: Continuity of Service

R03. The Repository has a plan to ensure ongoing access to and preservation of its data and meta data.

- The repository must have measures in place to address the risks inherent in changing circumstances, including in mission and/or scope. This Requirement covers the stable management of repository services over time (business continuity) and the response when services have problems (disaster recovery). It also includes preparations for handover of digital objects and services to another repository (succession planning). The deposit, storage, preservation, and access services offered by the repository to depositors and users are all in scope.

- The response statement and evidence should include references to the following items:
  - The functions and services offered by the repository to depositors and users.
  - The approach to rapid changes of circumstance and long-term planning.
  - The options for relocation or transition of the activity to another repository. For example, the case of cessation of funding due to an unexpected withdrawal of funding, or a shift of host institution interests.
  - The repository approach to managing policies, procedures and other business information over time.

CoreTrustSeal Requirement R03: Continuity of Service, Continued

- Even though succession agreements may be hard to achieve it is important to acknowledge the possibility that a repository will cease to function or exist. If there is no formal, written agreement between the repository and a successor then the compliance level cannot be higher than "In Progress: the repository is in the implementation phase".

- Any technical aspects of business continuity, and disaster and succession planning should be covered in R15 (Technical infrastructure).

CoreTrustSeal Requirement R03: Continuity of Service, Continued

- Repositories must ensure continuity of their collections and assume responsibility in the case of a temporary or permanent break in service. Responses and evidence should demonstrate the level of responsibility taken for digital objects, the level of risk for the repository, and the level of succession planning for the future of the data collection.

- Relevant information could include whether the applicant is the primary or only custodian, whether the depositor shares some responsibility for the future of the digital objects and any service level guarantees or minimum guaranteed time periods (e.g. for retention or preservation) in place.

- If sustainability partially depends on a host or parent organisation, or another organisation has guaranteed that it will take over the responsibility in the case of a service discontinuity, this should be clearly indicated. Identifying and entering a formal agreement with a successor organisation that can undertake to deliver the same levels of care and service is acknowledged as a challenge for many repositories. For this reason a continued status of 'in progress' may be accepted as sufficient during renewal of certification if clearly explained.

- If there is no formal, written agreement between the repository and a successor organisation this requirement cannot be assessed as' implemented'.

R03 Continuity of Service Example: Excerpt from AADC R03 Response

Implemented: the requirement has been fully implemented by the repository - 1

The primary function of the AADC is to catalogue and archive data from the AASP and to make these data available via search tools to the general public. In the unlikely event that something were to happen though, our metadata and data holdings, and the associated capabilities of search and access, would be secure through the following means:

1) As specified by the AADC Preservation Strategy 2024, a formal arrangement exists with Environment Information Australia (EIA), which is part of DCCEEW, who will take all our metadata and data holdings should the AADC cease to exist. The AADC Preservation Strategy 2024 states that EIA are an official approver of the document (page 4). https://data.aad.gov.au/about/preservation-strategy

2) The metadata catalogue would be preserved at the Global Change Master Directory (GCMD), which maintains a copy of our metadata records, as well as metadata from other National Antarctic Programs. Copies of all our metadata records would also be available from the Australian Research Data Commons (ARDC), and subsets would exist at the Australian Ocean Data Network (AODN) and various other organisations.

3) Data would be transferred to https://data.gov.au - which is an initiative of the Australian Government to provide data management services for its various departments - catering primarily to those departments that do not have a dedicated data repository.

4) Data would also be transferred to the Institute for Marine and Antarctic Studies (IMAS). IMAS is part of the University of Tasmania, and has its own Antarctic data centre for managing data arising from University funded Antarctic programs.

5) Ocean related data would be transferred to the AODN Data Portal, and to the Marine National Facility.

6) Some of our data holdings are also already available through other institutions like the Global Biodiversity Information Facility and the Ocean Biodiversity Information System.

R03 Continuity of Service Example: Excerpt from CSIRO R03 Response

Compliance level: In Progress: the repository is in the implementation phase - 0

The CSIRO Data Access Portal undertakes to make the files located in the archive available for the long term (1). The Scientific Domain the Pulsar Observation Search, a component of the CSIRO Data Access Portal and a function of the Australia Telescope National Facility states that data stored in the archives are for an indefinite period (2).

CSIRO has a long history as an organisation within Australia. CSIRO began as the Advisory Council of Science and Industry in 1916. CSIRO is constituted by and operates under the provisions of the Australian Government Science and Industry Research Act 1949 (3). CSIRO's main funding is directly from the Australian Government and this funding is more than AUD900 million per year (4) and is based on a triennial funding program.

Under the Archives Act 1983 (the Act), CSIRO must retain and destroy records in accordance with disposal authorities approved by the National Archives of Australia (5). In February 2023, the National Archives of Australia announced the Large-Big Data Project. The project aims are:

- Transfer, preservation and access to large/big datasets – requirements under the Archives Act 1983
- Options for a distributed custody model agreement.

We will work with the National Archives of Australia over the next few years to develop a succession plan ... ✓

R03 Continuity of Service Example: Excerpt from LAADS DAAC R03 Response

The long-term stability of the LAADS DAAC depends on funding from NASA, our sponsoring agency. NASA has clearly stated that it believes that long-term stewardship of NASA-collected remote sensing and field campaign data are essential. … As indicated in the ESDIS Project Plan, the functional and performance requirements for the ESDIS Project include the following:

- Provide active archive and distribution services for data until a peer review for science merit removes the need to maintain active status.
- Secure and preserve the unique (irreplaceable) and essential (designated for archive in perpetuity) data and information until such time as they are provided to permanent archival agencies.

Each DAAC is governed by the "Requirements for Archiving, Distribution and User Services in EOS Data and Information System (EOSDIS)" available via the following URL: https://cdn.earthdata.nasa.gov/conduit/upload/11349/423-10-69_ADURD_RevB.pdf

This document states: "Requirements to archive and distribute the data for a given mission extend as long as required by the ESDIS Project. This duration is dependent on the active use of the data by NASA-funded investigators and the provisions of long-term archiving as determined by ESDIS. [DAACs] will need to ensure that the data are transitioned to the appropriate Long-Term Archive when NASA notifies the [DAAC] that it is appropriate to do so." In the unlikely event that the NASA-EOSDIS LAADS DAAC is relocated to a different host institution, the LAADS DAAC will follow NASA procedures for activity transition. …

Continuity of Service R03: Questions

- How might a repository ensure that its holdings will continue to be preserved, accessible, and usable if its current funding ends?
- What agreements can the repository establish with others to ensure that its holdings will continue to be preserved, accessible, and usable if its current funding ends?

Introduction to CoreTrustSeal Requirement R04: Legal and Ethical

Introduction to CoreTrustSeal Requirement R04: Legal and Ethical

R04. The repository ensures to the extent possible that data and metadata are created, curated, preserved, accessed and used in compliance with legal and ethical norms.

- This requirement relates to repository awareness and processes around legal and ethical issues, including privacy and confidentiality, that impact the creation, curation, and use of digital objects.

- To maintain the trust of those who agree to have their digital objects held by the repository, evidence should demonstrate practices that reflect the legal status and sensitivity of digital objects, including guidance for depositors and users.

- The response statement and evidence should include references to the following items:
    - How the repository identifies and manages relevant legal and ethical standards that impact operations.
    - Compliance with specific legal and/or ethical discipline or domain standards.
    - Information requested from depositors to confirm that data collection or creation was carried out in accordance with legal and ethical criteria in the relevant geographical location or discipline (e.g. Ethical Review Committee/Institutional Review Board or Data Protection legislation).
    - Any data or metadata with disclosure risk e.g. depositor/user information, personal, cultural, or environmental information.

Preparing for Next Session

Preparing for Next Session: Activities in Your Repository

- Template assigned to each repository for self-assessment
- Review requirements for R03 within the context of your repository
- Discuss among managers and staff within your repository
- Draft a mission statement for your repository
- Draft responses to R03 for your repository within your template
- Review draft responses to other requirements for consistency

Questions?

Review of the CoreTrustSeal Requirements

CoreTrustSeal Requirements 2023-2025: Context

R0 Context

- General information about the repository

- Re3data identifier

- Repository Type. Select all relevant types from:
  - Generalist or Specialist; Specialist describes scope

- Overview – Brief Description of Repository

- Brief Description of the Designated Community

- Level of Curation Performed. Select all relevant types from:
  - Content distributed as deposited Basic curation – e.g., brief checking, addition of basic metadata or documentation. Enhanced curation – e.g., conversion to new formats during ingest, enhancement of documentation or metadata; Data -level curation – as in C above, but with additional editing of deposited data for accuracy

- Cooperation and Outsourcing to third parties, partners, and host organizations.

- Applicants renewing – Summary of Significant Changes Since Last Application

Source: CoreTrustSeal Standards and Certification Board. (2022). CoreTrustSeal Requirements 2023-2025 (V01.00). Zenodo. https://doi.org/10.5281/zenodo.7051012

CoreTrustSeal Requirements 2023-2025: Organizational Infrastructure

- R01 Mission & Scope
  - The repository has an explicit mission to provide access to and preserve digital objects.
- R02 Rights Management
  - The repository maintains all applicable rights and monitors compliance.
- R03 Continuity of Service
  - The Repository has a plan to ensure ongoing access to and preservation of its data and metadata.
- R04 Legal and Ethical
  - The repository ensures to the extent possible that data and metadata are created, curated, preserved, accessed and used in compliance with legal and ethical norms.
- R05 Governance & Resources
  - The repository has adequate funding and sufficient numbers of staff managed through a clear system of governance to effectively carry out the mission.
- R06 Expertise & Guidance
  - The repository adopts mechanisms to secure ongoing expertise, guidance and feedback-either in-house, or external.

CoreTrustSeal Requirements 2023-2025: Digital Object Management

- R07 Provenance & Authenticity
  - The repository guarantees the authenticity of the digital objects and provides provenance information.
- R08 Deposit & Appraisal
  - The repository accepts data and metadata based on defined criteria to ensure relevance and understandability for users.
- R09 Preservation Plan
  - The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way.
- R10 Quality of Service
  - The repository addresses technical quality and standards compliance, and ensures that sufficient information is available for end users to make quality-related evaluations.
- R11 Workflows
  - Digital object management takes place according to defined workflows from deposit to access.
- R12 Discovery & Identification
  - The repository enables users to discover the digital objects and refer to them in a persistent way through proper citation.
- R13 Reuse
  - The repository enables reuse of the digital objects over time, ensuring that appropriate information is available to support understanding and use.

CoreTrustSeal Requirements 2023-2025: Information Technology & Security

- ## R14 Storage & Integrity
  - The repository applies documented processes to ensure data and metadata storage and integrity.

- ## R15 Technical Infrastructure
  - The repository is managed on well-supported operating systems and other core infrastructural software and hardware appropriate to the services it provides to its Designated Community.

- ## R16 Security
  - The repository protects the facility and its data, metadata, products, services, and users.

Online Guidance and Related Resources for Improving Data Stewardship and Attaining CoreTrustSeal Certification 2023-2025

- CoreTrustSeal Standards and Certification Board. (2022). CoreTrustSeal Requirements 2023-2025 (V01.00). Zenodo. https://doi.org/10.5281/zenodo.7051012

- CoreTrustSeal Standards and Certification Board. (2022). CoreTrustSeal Trustworthy Digital Repositories Requirements 2023-2025 Extended Guidance (V01.00). Zenodo. https://doi.org/10.5281/zenodo.7051096

- CoreTrustSeal Standards and Certification Board. (2022). CoreTrustSeal Trustworthy Data Repositories Requirements: Glossary 2023-2025 (V01.00). Zenodo. https://doi.org/10.5281/zenodo.7051125

- CoreTrustSeal Standards and Certification Board. (2022). CoreTrustSeal Revision Working Group Change Log and Associated Materials (v01.00). Zenodo. https://doi.org/10.5281/zenodo.7051237

- Apply for CoreTrustSeal Certification: https://www.coretrustseal.org/apply/

- CoreTrustSeal Requirements (& webinars): https://www.coretrustseal.org/why-certification/requirements/

- CoreTrustSeal Administrative Fee (& waiver): https://www.coretrustseal.org/apply/administrative-fee/

- CoreTrustSeal: https://www.coretrustseal.org/

Online Resources for Improving Data Stewardship and Attaining CoreTrustSeal Certification (Continued)

- Consultative Committee for Space Data Systems. 2012. Reference Model for an Open Archival Information System (OAIS). CCSDS 650.0-M-2. Also published as ISO 14721:2012. https://public.ccsds.org/Pubs/650x0m2.pdf

- Consultative Committee for Space Data Systems (CCSDS). 2011. Audit and Certification of Trustworthy Digital Repositories. CCSDS 652.0-M-1. Also published as ISO 16363:2012. https://public.ccsds.org/pubs/652x0m1.pdf

- Group on Earth Observations (GEO). 2016. Global Earth Observations System of Systems (GEOSS) Data Management Principles. http://earthobservations.org/open_eo_data.php#

- Group on Earth Observations (GEO). 2016. Global Earth Observations System of Systems (GEOSS) Data Sharing Principles. http://earthobservations.org/open_eo_data.php#

Online Resources for Improving Data Stewardship and Attaining CoreTrustSeal Certification (Continued)

- Carroll, S.R., et al. 2020. The CARE Principles for Indigenous Data Governance. Data Science Journal, 19(1). https://doi.org/10.5334/dsj-2020-043

- Lin, D., et al. 2020. The TRUST Principles for digital repositories. Scientific Data 7, 144. https://doi.org/10.1038/s41597-020-0486-7

- Wilkinson, M, et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3, 160018. https://doi.org/10.1038/sdata.2016.18

Online Resources for Improving Data Stewardship and Attaining CoreTrustSeal Certification (Continued)

- International Science Council (ISC) CODATA international discussion list
  - http://lists.codata.org/mailman/listinfo/codata-international_lists.codata.org

- Research Data Alliance (RDA) https://www.rd-alliance.org/
  - https://www.rd-alliance.org/groups/rdawds-certification-digital-repositories-ig.html

- Group on Earth Observations (GEO)
  - https://earthobservations.org/

- World Data System (WDS)
  - https://www.worlddatasystem.org/

Online Resources for Improving Data Stewardship and Attaining CoreTrustSeal Certification (Continued)

- Downs, 2021. Improving Opportunities for New Value of Open Data: Assessing and Certifying Research Data Repositories. Data Science Journal. https://doi.org/10.5334/dsj-2021-001

- Downs RR. 2019. International Standards for Trustworthy Data Repositories. National Institutes of Health (NIH) Trustworthy Data Repositories Workshop. Rockville, MD. https://doi.org/10.7916/d8-h7xf-ha70

- Downs RR. 2021. Adopting the TRUST Principles for Digital Repositories with the GEOSS Data Management Principles and the GEOSS Data Sharing Principles. ESIP 2021 Winter Meeting (Virtual). https://doi.org/10.6084/m9.figshare.13607747.v1

- Downs RR, Chen RS. 2016. A Portfolio Approach to a Sustainable Business Model for Scientific Data Stewardship. SciDataCon 2016, Denver, CO. https://doi.org/10.7916/d8-fae5-cz67

- Downs RR, Chen RS. 2015. Bridging Disciplines: Assessing the Interdisciplinary Impact of Open Data. 41st IASSIST Annual Conference, Minneapolis, MN. https://doi.org/10.7916/D8J38SDZ

- Downs RR, Lehnert KA. 2020, Opportunities for Sharing Resources Among Research Data Repositories. Research Data Alliance 16th Plenary (RDA P16). https://doi.org/10.7916/d8-kv81-sv37

- Downs RR, Lehnert KA, Zhang L, Donaldson DR, Silva JR. 2020. Selecting Research Data Repository Platforms for Open Science. RDA P16. https://doi.org/10.7916/d8-051c-y379

Thank you!

rdowns@ciesin.columbia.edu