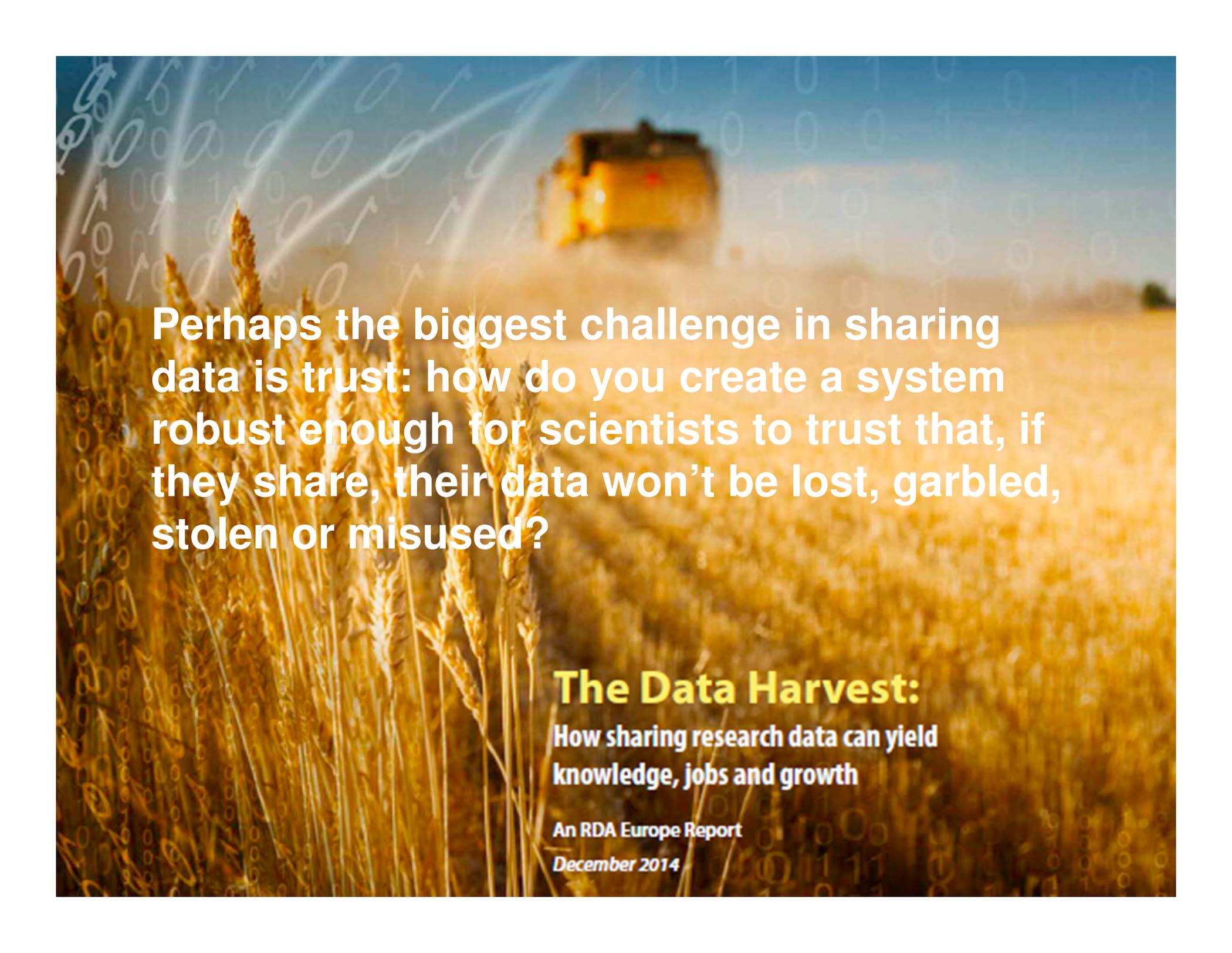




Certification

F. Genova (thanks to I. Dillo and Hervé L'Hours)

research data sharing without barriers
rd-alliance.org



Perhaps the biggest challenge in sharing data is trust: how do you create a system robust enough for scientists to trust that, if they share, their data won't be lost, garbled, stolen or misused?

The Data Harvest:

How sharing research data can yield knowledge, jobs and growth

An RDA Europe Report

December 2014

What is a trustworthy repository?

- mission to provide reliable, long-term access to managed digital resources to its designated community, now and into the future
- constant monitoring, planning, and maintenance
- understand threats to and risks within its systems
- **regular cycle of audit and/or certification**



The certification landscape

4

- 4 certification standards available



DIN 31644



ICSU
WORLD DATA SYSTEM



ISO 16363

European certification framework

- **Basic Certification** is granted to repositories which obtain DSA certification – ICSU-WDS at the same level
- **Extended Certification** is granted to Basic Certification repositories which *in addition* perform a structured, externally reviewed and publicly available self-audit based on DIN 31644/nestorSeal
- **Formal Certification** is granted to repositories which *in addition to* Basic Certification obtain full external audit and certification based on ISO 16363

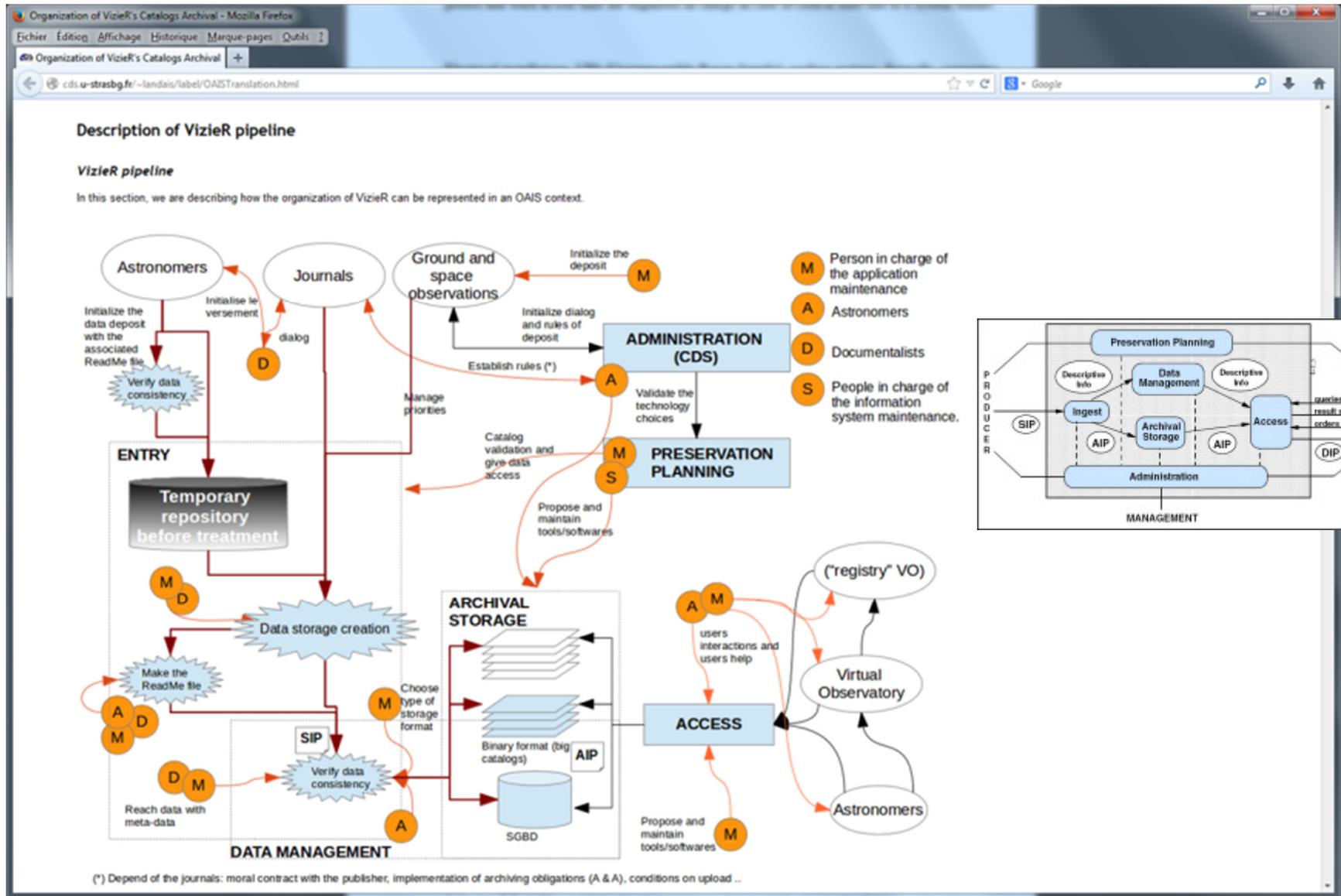


Why a formal audit/certification?

6

- Yes, why, when our users trust us?
- The example of CDS
 - Trusted by users, data providers and agencies
 - CDS underwent 'basic' certification in the WDS and DSA contexts
 - An in-depth work on our procedures when checking the criteria
 - A very interesting team work
 - Finally no change in our process but an end-to-end description, clarification of licensing aspects, etc
 - Very positive reaction from our authorities, the journal we work with, etc
- **It was worth it**

OAIS-like description of our services



Basic certification

- Real interest for the data providers
- But... What to do in practice when you are a data provider?
- An important topic to tackle within the RDA
- RDA established partnership with WDS on that topic (and others): a common Interest Group
- WDS + DSA : RDA/CDS Certification of digital repositories WG

Certification WG background

- Data Seal of Approval and World Data System both lightweight mechanisms for repository assessment
 - Self-assessment, no on-site visit
 - Peer-reviewed assessment supervised by the DSA Board and the WDS Scientific Committee
- DSA began in social science and humanities, WDS in natural and physical sciences but both expanding in scope
- Over past years, both groups began to see synergies
- Common members!
- DSA/WDS Certification WG

WG Goals, which were achieved

10

- Develop common catalog of criteria for basic repository assessment
- Develop common procedures for assessment
- Implement a shared testbed for assessment
- i.e. *alignment*

- *For a later stage* : Ultimately, create a shared framework for certification that includes other standards as well, including Nestor and ISO 16363/TRAC

Common requirements

11

- 16 common criteria
- Each criterion comes with guidance and self assessment level
- **Context: an essential element. The trustworthiness evaluation depends on the data repository mission!**
- Three topics addressed
 - Organisational infrastructure
 - Digital object management
 - Technology
- List of criteria in annex at the end of the talk

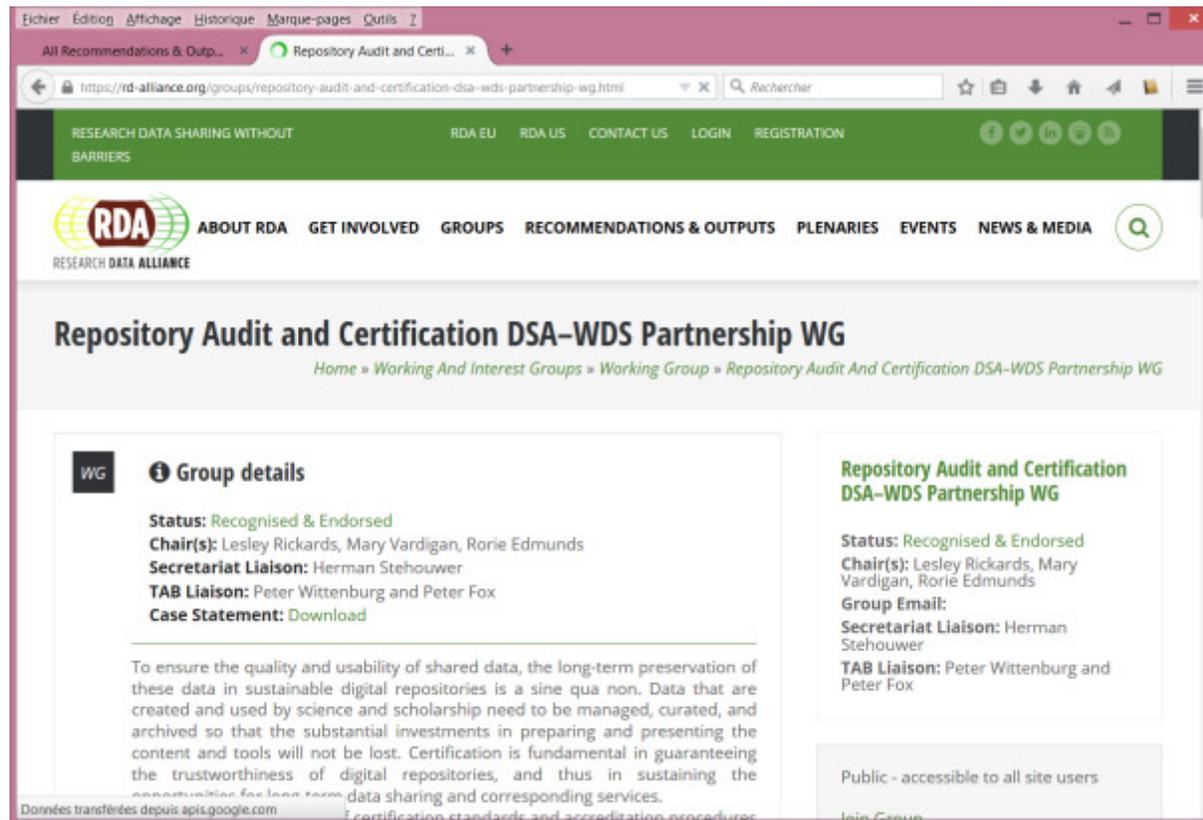
Strong message

12

- Very useful for self-assessment even if not submitted for external review!
- All criteria come with guidance and self-assessment of compliance level (not to be used by the external reviewers but here to help the applicants)

The Certification Working Group

13



The screenshot shows a web browser window displaying the RDA website. The page title is "Repository Audit and Certification DSA-WDS Partnership WG". The navigation menu includes "RESEARCH DATA SHARING WITHOUT BARRIERS", "RDA EU", "RDA US", "CONTACT US", "LOGIN", "REGISTRATION", "ABOUT RDA", "GET INVOLVED", "GROUPS", "RECOMMENDATIONS & OUTPUTS", "PLENARIES", "EVENTS", and "NEWS & MEDIA". The main content area features a "Group details" section with the following information:

- Status:** Recognised & Endorsed
- Chair(s):** Lesley Rickards, Mary Vardigan, Rorie Edmunds
- Secretariat Liaison:** Herman Stehouwer
- TAB Liaison:** Peter Wittenburg and Peter Fox
- Case Statement:** Download

Below the details, there is a paragraph of text: "To ensure the quality and usability of shared data, the long-term preservation of these data in sustainable digital repositories is a sine qua non. Data that are created and used by science and scholarship need to be managed, curated, and archived so that the substantial investments in preparing and presenting the content and tools will not be lost. Certification is fundamental in guaranteeing the trustworthiness of digital repositories, and thus in sustaining the data sharing and corresponding services."

On the right side, there is a summary box with the same details as the "Group details" section. At the bottom right, there is a "Public - accessible to all site users" label and a "Join Group" button.

<https://rd-alliance.org/groups/repository-audit-and-certification-dsa%E2%80%93wds-partnership-wg.html>

The WG Recommendations are on line

14

The screenshot shows a web browser window displaying the RDA website. The page title is "Repository Audit and Certification DSA-WDS Partnership WG Recommendations". The main content area includes a "Chairs:" section listing Lesley Rickards, Mary Vardigan, and Rorie Edmunds. A summary paragraph states: "The Repository Audit and Certification DSA-WDS Partnership WG has produced a two-part recommendation. The first part consists in a Catalogue of Common Procedures for certification, while the second part is focused on a Catalogue of Common Requirements. The Catalogue of Common Procedures was developed by the DSA-WDS Partnership Working Group on Repository Audit and Certification, a Working Group (WG) of the Research Data Alliance (RDA). The goal of the effort was to create a set of harmonized Common Procedures for certification of repositories at the basic level, drawing from the procedures already put in place by the Data Seal of Approval (DSA) and the ICSU World Data System (ICSU-WDS). These". A sidebar on the right provides additional details: "Repository Audit and Certification DSA-WDS Partnership WG", "Status: Recognised & Endorsed", "Chair(s): Lesley Rickards, Mary Vardigan, Rorie Edmunds", "Group Email:", "Secretariat Liaison: Herman Stehouwer", and "TAB Liaison: Peter Wittenburg and Peter Fox". A "Join Group" button is visible at the bottom of the sidebar.

<https://rd-alliance.org/group/repository-audit-and-certification-dsa%E2%80%93partnership-wg/outcomes/dsa-wds-partnership>

Organisational infrastructure

15

- Mission/scope
- Licenses
- Continuity of access
- Confidentiality/Ethics
- Organisational infrastructure
- Expert guidance

- Data Integrity and authenticity
- Appraisal
- Documented storage procedure
- Preservation plan
- Data quality
- Workflows
- Data discovery and identification
- Data reuse

Technology

17

- Technical infrastructure
- Security

- R0 Context

Please provide context for your repository

- R1 Mission/Scope

Organizational Infrastructure

The repository has an explicit mission to provide access to and preserve data in its domain

- R2 Licenses

Organizational Infrastructure

The repository maintains all applicable licenses covering data access and use and monitors compliance

- R3 Continuity of access Organizational infrastructure

The repository has a continuity plan to ensure ongoing access to and preservation of its holdings

- R4 Confidentiality/ethics Organizational Infrastructure

The repository ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with disciplinary and ethical norms

- R5 Organizational infrastructure Organizational Infrastructure

The repository has adequate funding and sufficient numbers of qualified staff managed through a clear system of governance to effectively carry out the mission

- R6 Expert guidance Organizational Infrastructure

The repository adopts mechanism(s) to secure ongoing expert guidance and feedback (either in-house, or external, including scientific guidance, if relevant)

- R7 Data integrity and authenticity Digital Object Management

The repository guarantees the integrity and authenticity of the data

- R8 Appraisal Digital Object Management

The repository accepts data and metadata based on defined criteria to ensure relevance and understandability for data users

- R9 Documented storage procedures Digital Object Management

The repository applies documented processes and procedures in managing archival storage of the data

- R10 Preservation plan Digital Object Management

The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way

- R11 Data quality Digital Object Management

The repository has appropriate expertise to address technical data and metadata quality and ensures that sufficient information is available for end users to make quality-related evaluations

- R12 Workflows Digital Object Management

Archiving takes place according to defined workflows from ingest to dissemination

- R13 Data discovery and identification Digital Object Management

The repository enables users to discover the data and refer to them in a persistent way through proper citation

- R14 Data reuse Digital Object Management

The repository enables reuse of the data over time, ensuring that appropriate metadata are available to support the understanding and use of the data

- R15 Technical infrastructure Technology

The repository functions on well-supported operating systems and other core infrastructural software and is using hardware and software technologies appropriate to the services it provides to its Designated Community

- R16 Security Technology

The technical infrastructure of the repository provides for protection of the facility and its data, products, services, and users