



RESEARCH DATA ALLIANCE



ICSU

WORLD DATA SYSTEM

DSA-WDS Partnership on Repository Certification Working Group

research data sharing without barriers
rd-alliance.org

Working Group Members

- Lesley Rickards (UK, PSMSL, WDS-SC) [Co-chair]
- Mary Vardigan (USA, ICPSR, DSA Board) [Co-chair]
- Kevin Ashley (UK, Digital Curation Centre)
- Michael Diepenbroek (Germany, Pangaea, WDS-SC)
- Ingrid Dillo (The Netherlands, DANS, DSA Board)
- Rorie Edmunds (WDS Programme Officer)
- Françoise Genova (France, CDS, WDS-SC)
- Hervé L'Hours (UK, UK Data Archive, DSA Board)
- Guoqing Li (China, CEODE, WDS-SC)
- Jean-Bernard Minster (USA, UCSD, Chair of WDS Scientific Committee)
- Mustapha Mokrane (WDS Executive Director)
- Paul Trilsbeek (The Netherlands, MPI for Psycholinguistics, DSA Board)
- Eleni Panagou, Ph.D. Candidate in Web Engineering, Democritus University of Thrace, Greece [RDA Early Career Researcher]


- Long term preservation of data in sustainable digital repositories is a key element of scientific data sharing
- The repositories should be TRUSTED
 - By scientists who deposit their data in a repository
 - By data users
 - By funding agencies who require that project results are shared
- An important topic to tackle within the RDA
- ICSU-WDS builds a community of quality-assured scientific data and data services, products, and information
- RDA/WDS partnership, hence the RDA/WDS Certification of Digital Repositories IG



The screenshot shows a web browser window displaying the website for the Repository Audit and Certification DSA-WDS Partnership Working Group. The browser's address bar shows the URL: <https://www.rd-alliance.org/groups/repository-audit-and-certification-dsa-wds-partnership-wg.htm>. The website has a navigation menu with links for Home, Organisation, Working and Interest Groups, Plenary Meetings, News & Events, Early Career Programmes, and About. A green banner below the menu says "View All Working Groups". The main content area features a breadcrumb trail: Home > Working and Interest Groups > Working Group > Repository Audit and Certification DSA-WDS Partnership WG. The title of the page is "Repository Audit and Certification DSA-WDS Partnership WG". To the left of the main text is a logo consisting of three stylized, overlapping curved lines in green, yellow, and red. To the right of the logo is a "Group details" section with the following information: Status: Recognised & Endorsed; Chair(s): Lesley Rickards, Mary Vardigan, Rorie Edmunds; Secretariat Liaison: Herman; TAB Liaison: Peter Wittenburg; Case Statement: Download. Below this is a paragraph of text explaining the importance of data quality and certification. To the right of the main text is a sidebar with several sections: "Group Wiki" with a link to "File Repository" and "Group Mailing list Archive"; a "Case Statement" section with a "Create new Case Statement" button; a "File Repository" section with a "latest files uploaded" section and a link to "Follow Up on Repository Audit & Certification DSA-WDS Partnership WG, 24 September 2014"; and a "Latest Webconference" section.

Home > Working and Interest Groups > Working Group > Repository Audit and Certification DSA-WDS Partnership WG

Repository Audit and Certification DSA-WDS Partnership WG



Group details

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

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 opportunities for long-term data sharing and corresponding services.

In recent years, a number of certification standards and accreditation procedures have been developed worldwide: Data Seal of Approval (DSA), Network of Expertise in long-term Storage and Accessibility of Digital Resources in Germany (NESTOR) seal / German Institute for Standardization (DIN) standard 31644, Trustworthy Repositories Audit and Certification (TRAC) criteria / International Organization for Standardization (ISO) standard 16363, and the International Council for Science World Data System (ICSU-WDS) certification of WDS Members.

The DSA and WDS certifications both offer a basic certification standard for trusted digital repositories. Their catalogues of requirements and their review procedures are based on the same principles of openness and transparency, and of striking the right balance between simplicity and robustness of the work and effort involved.

Group Wiki

- File Repository
- Group Mailing list Archive

Case Statement

Create new Case Statement

File Repository

latest files uploaded

Follow Up on Repository Audit & Certification DSA-WDS Partnership WG, 24 September 2014

Latest Webconference

The screenshot shows the homepage of the Data Seal of Approval website. At the top, there is a navigation menu with links for Home, Assessment, Community, News & Events, and a TOOL LOG IN button. A large red wax seal graphic on the left contains the text "Data Seal of Approval". Below this, a banner features the text "Towards sustainable and trusted data repositories" over a background image of documents. A text box on the left states: "There are 16 guidelines that together determine whether your data repository qualifies for the Data Seal of Approval." The main content area is divided into two sections: "News & Events" on the left, which lists "Finnish Social Science Data Archive Acquires the Data Seal of Approval" and "The Data Harvest", and "Seals Acquired Around the World" on the right, which includes a world map with red pins indicating the locations of data repositories that have received the seal.

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!
- When the RDA/WDS IG established, exploring a partnership seemed natural

The screenshot shows a web browser window displaying the RDA/WDS Certification of Digital Repositories Interest Group page. The browser's address bar shows the URL: <https://www.rd-alliance.org/groups/rdawds-certification-digital-repositories-ig.html>. The page header includes the RDA logo and navigation links: Home, Organisation, Working and Interest Groups, Plenary Meetings, News & Events, Early Career Programmes, and About. A green banner below the navigation links contains the text "View All Interest Groups".

The main content area features a breadcrumb trail: Home » Working and Interest Groups » Interest Group » RDA/WDS Certification of Digital Repositories IG. Below this is the title "RDA/WDS Certification of Digital Repositories IG" and a placeholder image for the group logo. To the right of the logo is a "Group details" box containing the following information:

- Status:** Recognised & Endorsed
- Chair(s):** Micheal Diepenbroek, Ingrid Dillo, Mustapha Mokrane
- Case Statement:** Download

Below the group details is a section titled "Status: Recognised & Endorsed Joint RDA/WDS IG". The text in this section reads: "In order to guarantee data sharing, 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, making sure that the substantial investments in preparing and presenting the content and tools will not be lost. Researchers need to be sure that the resources the repositories offer remain meaningful and usable over time. Moreover, the repositories themselves need to have sustainable business models."

The next paragraph states: "Preservation and sustainability raise challenges in many areas. The main issues related to long term preservation and sustainability remain basically unresolved, as many organizational, technical, financial and legal aspects remain open. Certification is therefore fundamental in guaranteeing the trustworthiness of digital repositories and thus in sustaining the opportunities for long-term data sharing."

The final paragraph reads: "The Interest Group will build on previous work in the area of certification. It will deliver the global overview and the necessary recommendations and requirements that allow the effective implementation of certification of digital"

On the right side of the page, there is a sidebar with several sections:

- Group Wiki**
- File Repository**
- Group Mailing list Archive**
- Case Statement** (with a sub-link "Create new Case Statement")
- File Repository** (with a sub-link "Latest file attachment in this group repository empty")
- Latest Webconference** (with a sub-link "Webconference Repository Audit and Certification IG")

■ DSA

- 16 guidelines for Thrustworthy Digital Repositories (data producers/repositories/consumers)
- DSA granted for a period of 2 years
- 45 seals acquired, some 40 underway

■ WDS

- Assessment to allow membership
- 17 criteria
- Review every 3-5 years
- ~50 members

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

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

Harmonization of Requirements

- Mapped the DSA to the WDS, and the WDS to the DSA
- Found that lists have similarities and differences
 - DSA guidelines more concise; WDS has multi-part criteria
 - DSA focus on data management, not organizational stability
 - WDS certification includes membership in the WDS and certification of services, not in scope for the DSA

Common Requirements

- Context
- Appraisal
- Mission/scope
- Documented storage procedures
- Preservation plan
- Workflows
- Data discovery and identification
- Data reuse
- Data integrity and authenticity

Common Requirements, continued

- Technical infrastructure
- Security
- Licenses
- Continuity of access
- Data quality
- Confidentiality/Ethics
- Organizational infrastructure
- Expert guidance

Challenges Faced

- There were significant gaps between the DSA and the WDS in some areas:
 - Technical and organizational infrastructure
 - Confidentiality and ethics
- The group had to compromise to come to good solutions

- The group works on aligning procedures and is addressing:
 - Appeals process
 - Compliance ratings
 - Path to improvement
 - URLs required in evidence
 - Language
 - Renewal frequency
 - Transparency
 - Procedures for reviewers
 - Branding of new common requirements
 - How the DSA and WDS relate to each other

Next Steps

- Create mapping to Nestor and ISO standards
- Finalize the harmonized requirements with guidance and put them out to the community as Version 1
- Complete work on aligning procedures, determining the relationship of DSA and WDS to each other as organizations
- Create a testbed for certification
- Investigate shared pool of reviewers

- The common framework should be a first step in a larger continuum of certification schemes. How do we achieve integration?
- *Global geographic reach.* Basic certification must account for cultural and linguistic differences. The notion of repository certification is currently dominated by Europe and the USA. How do we reach out to others?
- **The framework can also be used for self-assessment!
An efficient pathway for repository improvement**

- C1 Context Org Infrastructure / Digital Object Management

Please provide context for your repository

- C2 Appraisal Digital Object Management

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

- C3 Mission/Scope Organizational Infrastructure

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

- C4 Documented storage procedures Digital Object Management
The repository applies documented processes and procedures in managing archival storage of the data

- C5 Preservation plan Digital Object Management
The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way

- C6 Workflows Digital Object Management
Archiving takes place according to defined workflows from ingest to dissemination

- C7 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

- C8 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

- C9 Data integrity and authenticity Digital Object Management

The repository guarantees the integrity and authenticity of the data

- C10 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(ies)

- C11 Security Organizational Infrastructure, Digital Object Management, Technology

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

- C12 Licenses Organizational Infrastructure

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

- C13 Continuity of access Org Infrastructure / Digital Object Management

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

- C14 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

- C15 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

- C16 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

- C17 Expert guidance Organizational Infrastructure

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