

RDA: What, Where, When and How

Françoise Genova



□ What to use RDA for?

- As an international, neutral forum to discuss topics related to data sharing with people from very different profiles
- To gather lessons learnt and best practices
- To discuss and work on topics of interest for you with other people interested, produce recommendations in collaboration and get these recommendations commented by the RDA community
- To examine the RDA outputs and adopt those of interest for you

□ Where does it happens?

- During the Plenary meetings
 - Working Group and Interest Group meetings
 - Bird-of-a-Feather sessions
 - Informal discussions
- In the RDA Working Groups and Interest Groups (teleconferences, mails, F2F meetings)
- On the RDA Web site – still being improved

□ When?

- Two Plenaries/year, organised in different regions of the world
- A new Group can be proposed any time
 - Request for Comments from the community
 - Review by TAB (and comments from OAB for WG)
 - Acceptance by Council
- WGs
 - Have 18 months to produce « implementable » deliverables
 - Can have follow-up Groups or implementation support activities
- IGs exist on the long term, can also produce outputs



RDA 9th Plenary Meeting

Data Infrastructures for Open Science

**5-7 April 2017, Barcelo Sants Hotel,
Barcelona, Spain**

Organised by Barcelona Supercomputing Center (BSC) with the support of RDA Europe



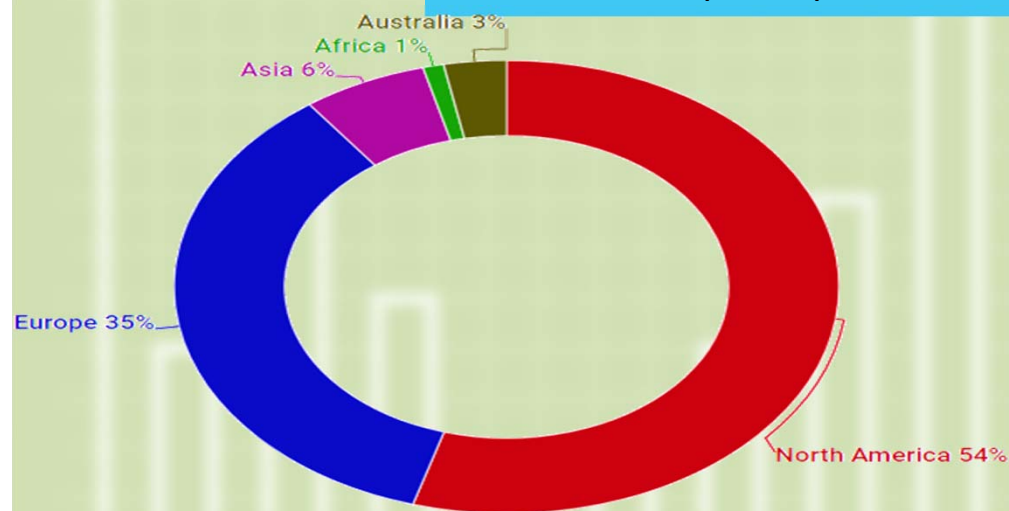
The 9th RDA Plenary Meeting will take place from 5th to 7th April 2017 at the Barcelo Sants Hotel, Barcelona, Spain. The plenary meeting is organised by the Barcelona Supercomputing Center-Centro Nacional de Supercomputación (BSC-CNS) with the support of RDA Europe.

<https://www.rd-alliance.org/plenaries/rda-ninth-plenary-meeting-barcelona>

Looking forward to seeing you all in Barcelona!



Over **430** participants from **30** countries



Increased female participation 44%
the highest sofar

73 Breakout meetings

of which 14 Working Groups

of which 37 Interest Groups

of which 6 Joint Working & Interest Groups

of which 16 Birds of a Feather

6 Outputs presented / 1 Final Release

60 Posters



rd-alliance.org/plenaries/rda-tenth-plenary-meeting-montreal-canada

WWW.RD-ALLIANCE.ORG
[@RESDATALL](https://twitter.com/RESDATALL)



CC BY-SA 4.0

Moving Towards Plenary 11: Berlin!



From Data to Knowledge

To find out more visit: <https://www.rd-alliance.org/plenaries/rda-eleventh-plenary-meeting-berlin-germany>



□ How

- ... to become a member of the RDA?
- ... to create a RDA Group?
- ... to produce a RDA recommendation or output?
- ... to get information about adoption of RDA recommendations?

How to join?

Who Can Join RDA?

Any individual or organization, regardless of profession or discipline, with an interest in reducing the barriers to data sharing and re-use and who agrees to RDA's guiding principles of:

- *Openness*
- *Consensus*
- *Balance*
- *Harmonization*
- *Community-driven*
- *Non-profit and technology-neutral*



Individual Membership is free @
<https://www.rd-alliance.org/user/register>

□ How to create a RDA Group?

- First, have a look at existing Groups to see if they already do what you need – you can join and bring your expertise, requirements and ideas
- BoF: proposal for a Plenary session
- WG, IG:
 - Find interested colleagues from different regions and prepare a case statement by filling a template
 - A BoF session at a Plenary is a good place to identify interested colleagues and discuss possible activities for new Groups
- Remember: RDA works on the technical AND sociological bridges to enable data sharing

RDA Interest (IG) & Working Groups (WG) by Focus (1)

Total 88 groups:
30 Working Groups & 58 Interest Groups

Domain Science - focused

- ☐ Agrisemantics WG
- ☐ BioSharing Registry WG
- ☐ Fisheries Data Interoperability WG
- ☐ On-Farm Data Sharing (OFDS) WG
- ☐ Rice Data Interoperability WG
- ☐ Wheat Data Interoperability WG
- ☐ Agricultural Data IG (IGAD)
- ☐ Biodiversity Data Integration IG
- ☐ Chemistry Research Data IG
- ☐ Digital Practices in History and Ethnography IG

- ☐ Geospatial IG
- ☐ Global Water Information IG
- ☐ Health Data IG
- ☐ Linguistics Data Interest Group
- ☐ Mapping the Landscape IG
- ☐ Marine Data Harmonization IG
- ☐ Quality of Urban Life IG
- ☐ RDA/CODATA Materials Data, Infrastructure & Interoperability IG
- ☐ Research data needs of the Photon and Neutron Science community IG
- ☐ Small Unmanned Aircraft Systems' Data IG
- ☐ Structural Biology IG
- ☐ Weather, Climate and air quality IG

Community Needs - focused

- ☐ Certification and Accreditation for Data Science Training and Education WG
- ☐ RDA/CODATA Summer Schools in Data Science and Cloud Computing in the Developing World WG
- ☐ Teaching TDM on Education and Skill Development WG
- ☐ Archives & Records Professionals for Research Data IG

- ☐ Data for Development IG
- ☐ Development of Cloud Computing Capacity and Education in Developing World Research IG
- ☐ Early Career and Engagement IG
- ☐ Education and Training on handling of research data IG
- ☐ Ethics and Social Aspects of Data IG
- ☐ International Indigenous Data Sovereignty IG

RDA Interest (IG) & Working Groups (WG) by Focus (2)

Total 88 groups:
30 Working Groups & 58 Interest Groups

Reference and Sharing - focused

- ☐ Data Citation WG
- ☐ Data Description Registry Interoperability WG
- ☐ Data Security and Trust WG
- ☐ Empirical Humanities Metadata WG
- ☐ International Materials Resource Registries WG
- ☐ Provenance Patterns WG
- ☐ QoS-DataLC Definitions WG

- ☐ RDA / WDS Publishing Data Bibliometrics WG
- ☐ Repository Core Description WG
- ☐ Research Data Collections WG
- ☐ Research Data Repository Interoperability WG
- ☐ Data Discovery Paradigms IG
- ☐ National Data Services IG
- ☐ RDA/CODATA Legal Interoperability IG
- ☐ Reproducibility IG
- ☐ Sharing Rewards and Credit (SHARC) IG

Partnership Groups

- ☐ RDA / TDWG Metadata Standards for attribution of physical and digital collections stewardship WG
- ☐ RDA/WDS Scholarly Link Exchange Working Group
- ☐ ELIXIR Bridging Force IG
- ☐ RDA/NISO Privacy Implications of Research Data Sets IG
- ☐ RDA/WDS Publishing Data IG

RDA Interest (IG) & Working Groups (WG) by Focus (3)

Total 88 groups:
30 Working Groups & 58 Interest Groups

Data Stewardship and Services – focused

- ☐ Brokering Framework WG
- ☐ DMP Common Standards WG
- ☐ Exposing Data Management Plans WG
- ☐ RDA / WDS Publishing Data Workflows WG
- ☐ WDS/RDA Assessment of Data Fitness for Use WG
- ☐ Active Data Management Plans IG
- ☐ Data in Context IG
- ☐ Data Rescue IG
- ☐ Data Versioning IG
- ☐ Domain Repositories IG

- ☐ Libraries for Research Data IG
- ☐ Long tail of research data IG
- ☐ Physical Samples and Collections in the Research Data Ecosystem IG
- ☐ Preservation e-Infrastructure IG
- ☐ Preservation Tools, Techniques, and Policies IG
- ☐ RDA/WDS Certification of Digital Repositories IG
- ☐ RDA/WDS Publishing Data Cost Recovery for Data Centres IG
- ☐ Repository Platforms for Research Data IG
- ☐ Research Data Provenance IG
- ☐ Virtual Research Environments IG

Base Infrastructure – focused

- ☐ Array Database Assessment WG
- ☐ Data Type Registries WG
- ☐ Metadata Standards Catalog WG
- ☐ PID Kernel Information WG
- ☐ Data Fabric IG
- ☐ Data Foundations and Terminology IG
- ☐ Disciplinary Interoperability Framework IG
- ☐ Big Data IG

- ☐ Brokering IG
- ☐ Federated Identity Management IG
- ☐ Metadata IG
- ☐ PID IG
- ☐ Software Source Code IG
- ☐ Vocabulary Services IG

□ How to produce an Output?

- Work hard with your colleagues and produce an output of interest for others!
- WG proposed recommendations go to Request for Comments from the community
- Some Groups submit their outputs to external experts

RDA Recommendations & Outputs



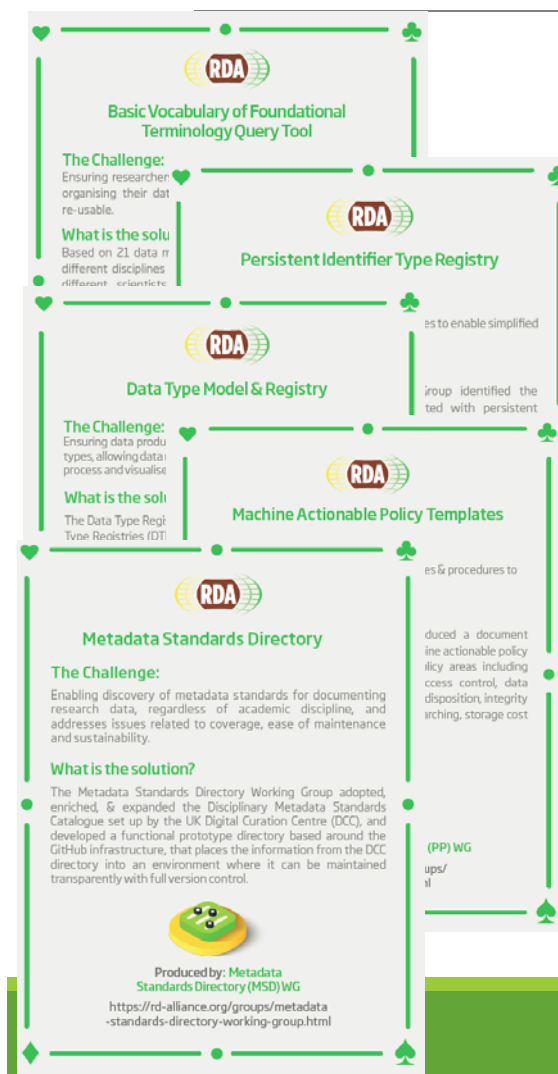
THE RDA OUTCOMES LEGEND

Recommendations: are the flagship outputs of RDA. They are RDA's equivalent of the "specifications" or "standards" that other organisations create and endorse. The process for creating and endorsing these is already defined.

Supporting Outputs: are the outputs of RDA WGs and IGs that are fruit of RDA work, but are not necessarily adoptable bridges. "Upon request", these sort of outputs go through a community comment period and if no major objections or gaps are identified they get the RDA Brand.

Other Outputs: include workshop reports, published articles, survey results, etc. Anything a WG or IG wants to register and report. Upon request, these are published and discoverable on the RDA website but have no level of endorsement.

RDA Recommendations & Outputs



Data Foundation & Terminology: a model for data in the registered domain.

PID Information Types: a common protocol for providers and users of persistent ID services worldwide.

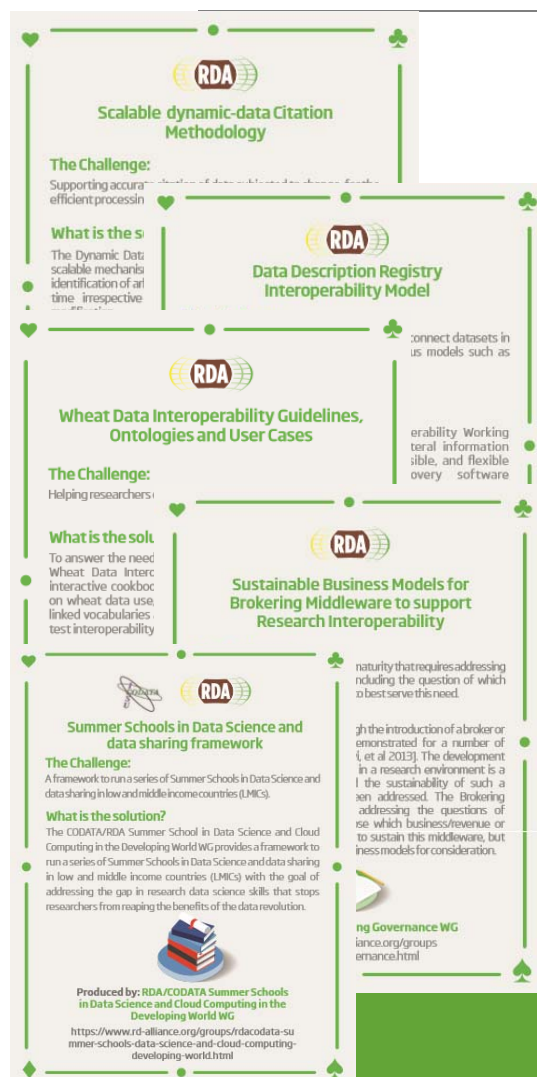
Data Type Registries: allowing humans and machines to act on unknown, but registered, data types.

Practical Policy: defining best practices of how to deal with data automatically and in a documented way with computer actionable policy.

Metadata standards directory: Community curated standards catalogue for metadata interoperability

rd-alliance.org/recommendations-and-outputs/all-recommendations-and-outputs

RDA Recommendations & Outputs



RDA
RESEARCH DATA ALLIANCE

Scalable dynamic-data Citation Methodology

The Challenge:
Supporting accurate and efficient processing of dynamic data

What is the solution?
The Dynamic Data scalable mechanism for identification of artifacts in respect to time

RDA
Data Description Registry Interoperability Model

connect datasets in various models such as

RDA
Wheat Data Interoperability Guidelines, Ontologies and User Cases

The Challenge:
Helping researchers

What is the solution?
To answer the need for Wheat Data Interoperability, an interactive cookbook on wheat data use, linked vocabularies, test interoperability

RDA
Sustainable Business Models for Brokering Middleware to support Research Interoperability

nature that requires addressing the question of which is best to serve this need.

RDA
Summer Schools in Data Science and data sharing framework

The Challenge:
A framework to run a series of Summer Schools in Data Science and data sharing in low and middle income countries (LMICs).

What is the solution?
The CODATA/RDA Summer School in Data Science and Cloud Computing in the Developing World WG provides a framework to run a series of Summer Schools in Data Science and data sharing in low and middle income countries (LMICs) with the goal of addressing the gap in research data science skills that steps researchers from reaping the benefits of the data revolution.

Produced by: RDA/CODATA Summer Schools in Data Science and Cloud Computing in the Developing World WG
<https://www.rd-alliance.org/groups/dacodata-summer-schools-data-science-and-cloud-computing-developing-world.html>

Brokering Governance WG
www.rd-alliance.org/groups/brokering-governance.html

Data Citation: defining mechanisms to reliably cite dynamic data

Data Description Registry Interoperability solutions enabling cross platform discovery based on existing open protocols and standards

Wheat Data Interoperability impacting the discoverability, reusability and interoperability of wheat data by building a common framework for describing, representing linking and publishing wheat data

Brokering Governance WG: Sustainable Business Models for Brokering Middleware to support Research Interoperability

RDA/CODATA Summer Schools in Data Science and Cloud Computing in the Developing World WG: A framework to run a series of Summer Schools in Data Science and data sharing in low and middle income countries (LMICs)

rd-alliance.org/recommendations-and-outputs/all-recommendations-and-outputs

RDA Recommendations & Outputs



Repository Audit and Certification DSA-WDS: A convergent DSA-WDS certification standard to help eliminate duplication of effort, increase certification procedure coherence and compatibility thus benefitting researchers, data managers, librarians and scientific communities.

RDA/WDS Publishing Data Bibliometrics: improved research data metrics and corresponding services, with the final goal of increasing the overall availability and quality of citations and research data itself.

RDA/WDS Publishing Data Workflows: enhance the possibilities for greater discoverability and a more efficient and reliable reuse of research data benefitting other stakeholders like publishers, libraries and data centres.

RDA/WDS Publishing Data Services: A universal interlinking service between data and the scientific literature. **The Scholix initiative** a high level interoperability framework for exchanging information about the links between scholarly literature and data. It aims to build an open information ecosystem to understand systematically what data underpins literature and what literature references data.

rd-alliance.org/recommendations-and-outputs/all-recommendations-and-outputs

RDA Recommendations & Outputs



23 Things: Libraries For Research Data An overview of practical, free, online resources and tools that users can immediately take advantage of to incorporate research data management into the practice of librarianship.

Legal Interoperability of Research Data Principles and Implementation Guidelines: a set of principles and practical implementation guidelines offered as high-level guidance to all members of the research community —the funders, managers of data centers, librarians, archivists, publishers, policymakers, university administrators, individual researchers, and their legal counsel.

Matrix of use cases and functional requirements for research data repository platform Based on use cases, the matrix describes forty-four functional requirements identified for research data repository platforms and provides a score identifying relative importance.

BioSharing Recommendations Data repositories, standards and policies in the life, biomedical and environmental sciences

rd-alliance.org/recommendations-and-outputs/all-recommendations-and-outputs

Adoption & Implementation

*“Solving the problem must include **adopters** in the process, to ensure that real problems are addressed. Open problem solving is the key.”*

RDA Recommendations and Outputs take the form of technical specifications, code, policies or practices, harmonized standards or reference models. In the widest sense these aim for:

- ❑ Greater data sharing, exchange, interoperability, usability and re-usability;
- ❑ Greater discoverability of research data sets;
- ❑ Better management, stewardship, and preservation of research data;
- ❑ New data standards or harmonization of existing standards.

RECOMMENDATIONS & OUTPUTS

All Recommendations & Outputs

Adoption Use Cases

Become an RDA Adopter

Addressing data challenges

<https://www.rd-alliance.org/recommendations-and-outputs/all-recommendations-and-outputs>

75 Adoption Cases

<https://www.rd-alliance.org/recommendations-outputs/adoption-recommendations>

Find out how you can become an Adopter

<https://www.rd-alliance.org/recommendations-and-outcomes/become-rda-adopter>

rd-alliance.org/recommendations-and-outputs/all-recommendations-and-outputs

RDA MEETS RESEARCHERS IN CROATIA



RESEARCH INFRASTRUCTURES IMPLEMENTING RDA OUTPUTS FOR MAPPING METADATA STANDARDS

IMPLEMENTING RDA OUTPUTS FOR SCHOLARLY COMMUNICATION

RDA Adoption & Implementation Stories - Tell us yours!

ADOPTING RDA OUTPUTS FOR ... CLIMATE DATA

DKRZ adopts 6 RDA outputs for climate data modelling

DKRZ is integrating parameter identifiers for use cases supporting precise data tracking, automated identification and reanalysis, custom and early data access into the Earth System Grid Federation data infrastructure and supports the ISO 15926 data modelling. This requires elemental RDA information to be interoperable across multiple services and tools and formulating community-specific RDA profiles. Furthermore, future automated processing workflows

could leverage such information as well if bound to specific data types and broken through a dedicated interface. To give services as possibly huge number of objects and their identifiers, the services and tools involved can also benefit from a possible RDA recommendation on research data collection.

The Challenge

"Current data management practices still rely largely on managing files and directories in file systems. Factors such as the relative increase of data volumes compared to available network bandwidth and the easy availability of remote and on-demand computing resources are drivers behind

bringing processing and data closer together. National and international policy changes in Earth Science funding may also cause a shift in the skills and expectations of environmental data service users."

Says Tobias Weigel, a Computer scientist at the adopting organisation, Deutsches Klimarechenzentrum (DKRZ)

Together, these factors lead to scenarios where it will be increasingly difficult to manage data on a specific or proprietary basis and deal with data transfer, replication and the data management as a comparatively low level of automation. Future tools may increasingly hide the location and structure of data objects from the user, requiring more intelligence from backend services. Services that provide access to processing and processing and more data management transparency may be particularly valuable for interdisciplinary users and/or with established community practices.

Weigel continues: "With such solutions that increase automation, cost of maintaining services will increase, which would have a devastating effect on service quality and reduce resources available for developing new services required for future user demands. Data experience has shown that many tasks such as data transfer or replication suffer from manual intervention required as long as no comprehensible data handling solution is in place. Such tasks may take up even more resources due to the data volumes and number of operations to manage increases exponentially."

RDA RECOMMENDATIONS ADOPTED

Basic Identification and Terminology strictly implemented and community specific basic identifiers such as digital object and persistent identifiers.

RDA Information Types: describe RDA record properties/benefits from types possibly defined across communities.

Basic Metadata: gain a better understanding of software modularity and make data management, processing operations and workflow development.

Basic Types Register: register RDA information types and eventually biotypes with processing services.

Keywords: Better Classification, data citation processing and access information systems for potentially variable data.

ANSWERING COMMUNITY NEEDS

The community that benefits from reduced implementation costs data management across multiple organisational boundaries. In particular attention within the European context is the European Helmholtz for Earth System Modelling (HMES). The breadth of scientific and data home use requires beyond the core climate modelling community can involve other disciplines such as climate impact research, adaptation and mitigation, public services, agriculture and so on. In the climate change is global phenomenon and challenge and conduct on a huge number of areas.

WHY RDA

RDA, not just found as a venue where experts from different disciplines and shaped by different community practices can combine their collective knowledge to build reusable, valuable solutions within manageable time. The RDA outputs to move to the adoption were selected by providing future application scenarios and long-term challenge and finding good matches with manageable goals with the existing solutions in DKRZ.

Find out more

Visit RDA @ rd-alliance.org
Email: enquiries@rd-alliance.org

20

<https://www.rd-alliance.org/newsroom/rd-alliance-adopts>

Adopting RDA Outputs for Climate

RDA in a Nutshell

WWW.RD-ALLIANCE.ORG/
[@RESDATALL](https://twitter.com/RESDATALL)



24 October 2017

RDA Global

Email - enquiries@rd-alliance.org

Web - www.rd-alliance.org

Twitter - [@resdatall](https://twitter.com/resdatall)

LinkedIn - www.linkedin.com/in/ResearchDataAlliance

Slideshare -

<http://www.slideshare.net/ResearchDataAlliance>

RDA Europe

Email - info@europe.rd-alliance.org

Web - europe.rd-alliance.org

Twitter - [@RDA_Europe](https://twitter.com/RDA_Europe)

RDA US

Twitter - [@RDA_US](https://twitter.com/RDA_US)