RDA: State of the Art
Tobias Weigel, Peter Wittenburg, Rainer Stotzka

RDA-Deutschland-Treffen 2015
Potsdam, 25.-26.11.2015

research data sharing without barriers
rd-alliance.org
Research Data Alliance

Vision:
Researchers openly sharing data across technologies, disciplines, and countries

Mission:
Building the social and technical bridges

Guiding Principles
• Openness
• Consensus
• Balance
• Harmonization
• Community-driven
• Non-profit

> 700 active members
• Working Groups
• Interest Groups
• Plenaries

Glienicker Brücke: This work has been released into the public domain by its author, Uwca at the German Wikipedia project. This applies worldwide. Uwca grants anyone the right to use this work for any purpose, without any conditions, unless such conditions are required by law.
Experts that are engaged in creating deliverables that will directly enable data sharing, exchange, or interoperability.

- **Concrete problem**
- Short time frame (~18 months)
- Consensus-driven
- **Concrete output** improving data exchange, deliverables
  - Technical specifications and implementation practices,
  - Conceptual models or frameworks,
  - Policies, and
  - Other documents and practices
- Clear adoption by one or more specific communities

Working groups and their deliverables undergo a community review process.
41 RDA Interest Groups

Experts that are engaged in creating deliverables that will directly or indirectly enable data sharing, exchange, or interoperability.

- Group around a common interest
- Exist as long as they are active
- Coordination and communication
- Outputs:
  - Platform that leads to the formation of one or more Working Groups
  - Communication and coordination across Working Groups/Interest Groups
  - Communication and coordination with communities outside RDA
  - Important deliverables such as surveys, recommendations, reports
Group Clustering

Data providers

Data consumers

Social

Technical

Solutions

Beneficiary

Group Clustering
Group Clustering

- **Data providers**
  - Social/organizational solution aimed at data provider
  - Technical solution aimed at data provider

- **Data consumers**
  - Social/organizational solution aimed at data consumer
  - Technical solution aimed at data consumer

- **Solutions**
  - Governance, Certification, Cost Recovery, Legal, ...

- **Beneficiary**

---

Social/organizational and technical solutions are clustered in a 2D diagram with data providers and data consumers as axes. The diagram illustrates the relationship between solutions and data providers/consumers, emphasizing the alignment of social and technical solutions with beneficiary outcomes.
Outputs & Benefits

Data Citation
- Cite data that is subjected to change
- All changes are reflected in the citation information that includes a time-stamp & version history

DSA-WDS Repository Audit and Certification
- Convergent DSA-WDS certification standard

RDA/WDS Publishing Data Services
- Universal linking service between data and the scientific literature

Data Foundation & Terminology (DFT)
- Increased cross disciplinary data exchange and interoperability
- Common data model
- Foundational vocabulary and query tool
Group Clustering

Solutions

Social

Social/organizational solution aimed at data provider

Technical

Technical solution aimed at data provider

Data providers

Data consumers

Beneficiary

Repository, Fabric, Analytics, Identity, Management, ...
Outputs & Benefits

Data Type Registry
- “MIME-types” for data
- Information for tools to interpret, process, and display data

PID Information Types (PIT)
- Generic API
- Set of basic attributes

RDA/WDS Publishing Data Workflows
- Dataset: http://dx.doi.org/10.5281/zenodo.19107

Practical Policies
- Template of policy descriptions for automation
- Basic set of 11 policy areas
Group Clustering

Solutions

Interoperability, Harmonization, Integration, Metadata, ...

Data providers

Technical solution aimed at data provider

Data consumers

Technical solution aimed at data consumer

Social/organizational solution aimed at data consumer

Beneficiary
Outputs & Benefits

Data Description Registry Interoperability (DDRI)
- Find connections across research data registries and create global views of research data

Wheat Data Interoperability
- Framework to support the establishment of a global wheat information system
- Interoperability by:
  - Creating an interoperability framework
  - Providing guidelines on wheat data (cookbook)
  - Repository of linked vocabularies

Metadata Standards Directory
- Community curated standards catalogue for metadata interoperability
Group Clustering

Solutions

Social

Social/organizational solution aimed at data provider

Social/organizational solution aimed at data consumer

Education, Engagement, Bridging, ...

Q4

Data providers

Data consumers

Technical solution aimed at data consumer

Beneficiary
RDA/CODATA Summer Schools in Data Science and Cloud Computing in the Developing World

- First summer school course in the data science approaches and skills that are essential for 21st century research
RDA: Accelerate Data Sharing and Interoperability Across Cultures, Communities, Scales, Technologies

- **Technical parts of the data engine:**
  - Data type registries reference model
  - Wheat data interoperability framework

- **Rules of the road:**
  - Common agreement on data citation
  - Common practice for data repositories
  - Principles of legal interoperability

- **Better drivers**
  - Summer schools in data science and cloud computing in the developing world (with CODATA)
  - Active data management plan development and monitoring
Communiqué Meeting of the G7 Ministers of Science Berlin, 8-9 October 2015

Global Research Infrastructures

“... Further progress on sharing and managing scientific data and information should be achieved, especially by continuing engagement with community based activities such as the Research Data Alliance RDA. ...“