Research Data Sharing without Barriers: The Role of the Research Data Alliance (RDA) and RDA-Austria

<table>
<thead>
<tr>
<th>Andreas Rauber</th>
</tr>
</thead>
<tbody>
<tr>
<td>Institute of Information Systems Engineering</td>
</tr>
<tr>
<td>Vienna University of Technology</td>
</tr>
<tr>
<td><a href="mailto:rauber@ifs.tuwien.ac.at">rauber@ifs.tuwien.ac.at</a></td>
</tr>
<tr>
<td><a href="http://www.ifs.tuwien.ac.at/~andi">http://www.ifs.tuwien.ac.at/~andi</a></td>
</tr>
</tbody>
</table>
Outline

- RDA – The Research Data Alliance
- RDA Europe (4.0)
- RDA-Austria (plus a breaking-news announcement!)
  - Why do we need RDA-AT?
  - What does RDA-AT do?
RDA in a Nutshell

Established **2013**, now more than **7300 members** from ~**140 countries**

**Vision**

Researchers and innovators openly share data across technologies, disciplines, and countries to address the grand challenges of society

**Mission**

RDA builds the **social and technical bridges** that enable open sharing of data

“Picking low-hanging fruit”

**Recommendations** and other outputs
Who is RDA

RDA Worldwide Growth

Members
7,372

rd-alliance.org/about-rda
RDA Geographical Distribution

RDA members come from 137 different countries

rd-alliance.org/about-rda

WWW.RD-@RESDATA.L berth
RDA Organisational & Affiliate

48 Organisational Members

8 Affiliate Members

https://rd-alliance.org/organisation/rda-organisation-affiliate-members.html
Outline

- RDA
- RDA Europe (4.0)
- RDA-Austria (plus a breaking-news announcement!)
  - Why do we need RDA-AT?
  - What does RDA-AT do?
RDA Europe

- European plug-in to the Research Data Alliance via a series of projects
- Launch of RDA Europe 4.0 in March 2018
- Mandated to become the centrepiece for an EU Open Science Strategy
- Establishment of a network of national nodes to foster adoption of RDA outputs in the region
- Calls to expand the network via a cascading grant funding mechanism (change to previous RDA-Europe grants)
RDA Europe

9 Project Consortium Nodes

• 4 nodes to be added 2018
• further 12 + 7 in round 2 and 3
Outline

- RDA
- RDA Europe (4.0)
- RDA-Austria (plus a breaking-news announcement!)
  - Why do we need RDA-AT?
  - What does RDA-AT do?
- What are these Working / Interest Groups?
National nodes discussion started in **March 2017** at RDA Plenary 11 in Berlin

**November 2017** RDA group for Austria established to represent the Austrian data management community

**December 2017** Constitutive General Assembly of RDA Austria

**March 2018** First Member Forum

Founding committee: Paolo Budroni, Raman Ganguly, Andreas Rauber, Barbara Sánchez, Tomasz Miksa (not in the picture: Seyavash Amini)
Workshop in Vienna, November 2017

- Data Stewardship Realized: From Planning to Action. Towards the Establishment of an Austrian Research Infrastructure
- Organized with the support of RDA Europe, the University of Vienna and the TU Wien
Endorsement of the EOSC

- On 11 April 2018, RDA Austria sent an endorsement to the October 2017 Declaration of the European Open Science Cloud
9th RDA WG/IG Chairs Meeting

- June 13-15 2018 at TU Wien
- Notes, slides, and summary available at the event website: https://www.rd-alliance.org/groups/9th-wgig-collaboration-meeting-vienna-13-15-june-2018
Application to become National Node

- Application submitted August 20 2018

Application for RDA Austria to become a National Node within the framework of RDA Europe

Related Call: RDA European Call for new nodes

Table of contents

Q1: Geographic area covered by the proposed node (e.g. country, region) 1
Q2: Key personnel involved in your application 1
Q3: Current RDA activities 3
Q4: Planned node activities to support the vision of RDA
  Q4.1 Mission 4
  Q4.2 Planned Working Groups and WG Branches 5
  Q4.3 Involvement in the European Open Science Cloud 5
  Q4.4 Outputs and KPIs 6
Q5: Community endorsements and support
  Q5.1 Pre-existing networks 6
  Q5.2 Composition of founding board 7
  Q5.3 Composition of Member Forum 8
Q6: Sustainability plans 9
Q7: Resources 10
Annex 11
Application to become National Node

- Application submitted August 20 2018
- Breaking news: On Thu, Nov. 8 2018 RDA Europe announced the 4 nodes selected for funding
  - Austria
  - Denmark
  - Portugal
  - Slovenia
- Elaborating detailed workplan KPIs and budget
- Signing formal agreement

Table of contents

| Q1: Geographic area covered by the proposed node (e.g. country, region) | 1 |
| Q2: Key personnel involved in your application | 1 |
| Q3: Current RDA activities | 3 |
| Q4: Planned node activities to support the vision of RDA |
| Q4.1 Mission | 4 |
| Q4.2 Planned Working Groups and WG Branches | 5 |
| Q4.3 Involvement in the European Open Science Cloud | 5 |
| Q4.4 Outputs and KPIs | 6 |
| Q5: Community endorsements and support |
| Q5.1 Pre-existing networks | 6 |
| Q5.2 Composition of founding board | 7 |
| Q5.3 Composition of Member Forum | 8 |
| Q6: Sustainability plans | 9 |
| Q7: Resources | 10 |
| Annex | 11 |
Mission of RDA Austria

- Serve cross border as a professional **connector** for global RDM and data sharing activities
- Act as node and **contractual partner** for RDA-Europe
- Disseminate outputs **outputs** from RDA global initiatives on a national level
- Organize **events** and trainings that promote best practices in RDM, inviting RDA WG/IG experts
- Promote local **adoption** of **RDA recommendations**
- Connect Austrian research data **stakeholders**
Planned Working Groups

1) Data citation, discovery and reuse
2) GDPR and research data
3) Research data management policies and services for academia, industry and SMEs
4) Machine-actionable data management plans
5) Specific challenges concerning research data in the digital humanities
6) <tba – it’s the members who decide what’s needed!>
Outline

- RDA
- RDA Europe (4.0)
- RDA-Austria (plus a breaking-news announcement!)
  - Why do we need RDA-AT?
  - What does RDA-AT do?
- What are these Working / Interest Groups?
Making Dynamic Data Citeable

March 2014 – September 2015

- Concentrating on the problems of large, dynamic (changing) datasets

Final version presented Sep 2015 at P6 in Paris, France

Endorsed September 2016 at P8 in Denver, CO

Since then: supporting adopters

Process for ICT Technical Specification

https://www.rd-alliance.org/groups/data-citation-wg.html
Identification of Dynamic Data

- Usually, datasets have to be static
  - Fixed set of data, no changes:
    no corrections to errors, no new data being added
- But: (research) data is **dynamic**
  - Adding new data, correcting errors, enhancing data quality, …
  - Changes sometimes highly dynamic, at irregular intervals
- Current approaches
  - Identifying entire data stream, without any versioning
  - Using “accessed at” date
  - “Artificial” versioning by identifying batches of data (e.g. annual), aggregating changes into releases (time-delayed!)
- Would like to identify precisely the **data** as it existed at a specific point in time
Granularity of Subsets

- What about the **granularity** of data to be identified?
  - Enormous amounts of CSV data
  - Researchers use specific subsets of data
  - Need to identify precisely the subset used

- Current approaches
  - Storing a copy of subset as used in study -> scalability
  - Citing entire dataset, providing textual description of subset -> imprecise (ambiguity)
  - Storing list of record identifiers in subset -> scalability, not for arbitrary subsets (e.g. when not entire record selected)

- Would like to be able to identify precisely the **subset of (dynamic) data used** in a process
Granularity of Subsets

- 14 Recommendations
  - versioned and time-stamped data
  - subsets via time-stamped queries
- 2-page flyer
- More detailed report:
- Adopter’s webinars and reports
DMP Common Standards WG

- Shortcomings of existing DMPs
  - Manual, free-form text, vague, not updated, …
- Machine-actionable DMPs
  - living documents
  - automate data management
    - collect information from systems
    - trigger actions in systems
- Goal: Define common data model for machine-actionable DMPs
  - to model information from standard DMPs
  - NOT a template
  - NOT a questionnaire
- 100+ members from around the world (DMP tool providers, funders, etc.)
- 18 months (Sep 2017 – Apr 2019)
- https://www.rd-alliance.org/groups/dmp-common-standards-wg
Machine-actionable Data Management Plans

- From PDF-DMPs...
- ...to maDMPs
- WG: define information model
- Focus on automation
Current DMPs – model questionnaires

<administrative_data>

<question>Who will be the Principle Investigator?</question>

<answer>The PI will be John Smith from our university.</answer>

</administrative_data>
Example (not the real model)

• Current DMPs – model questionnaires

<administrative_data>
  <question>Who will be the Principle Investigator?</question>
  <answer>The PI will be John Smith from our university.</answer>
</administrative_data>

• Machine-actionable DMPs – model information

"dc:creator": [ 
  {
    "foaf:name": "John Smith",
    "@id": "orcid.org/0000-1111-2222-3333",
    "foaf:mbox": "mailto:jsmith@tuwien.ac.at",
    "madmp:institution": "AT-Technische-Universität-Wien"
  },
]
Example (not the real model)

- Current DMPs – model questionnaires
  
  `<administrative_data>`
  
  `<question>`Who will be the Principle Investigator?`</question>`
  
  `<answer>`The PI will be John Smith from our university.`</answer>`
  
  `</administrative_data>`

- Machine-actionable DMPs
  
  "dc:creator":[
    
    "foaf:name":"John Smith",
    
    "@id":"orcid.org/0000-1111-2222-3333",
    
    "foaf:mbox":"mailto:jsmith@tuwien.ac.at",
    
    "madmp:institution":"AT-Technische-Universität-Wien"
  ]

  Reuse existing standards, e.g. Dublin Core, PREMIS, etc.
Example (not the real model)

• Current DMPs – model questionnaires
  
  <administrative_data>
    <question>Who will be the Principle Investigator?</question>
    <answer>The PI will be John Smith from our university.</answer>
  </administrative_data>

• Machine-actionable DMPs – model information
  
  "dc:creator": [ {
    "foaf:name": "John Smith",
    "@id": "orcid.org/0000-1111-2222-3333",
    "foaf:mbox": "mailto:jsmith@tuwien.ac.at",
    "madmp:institution": "AT-Technische-Universität-Wien"
  } ],

  Use PIDs whenever possible, e.g. ORCID
Example (not the real model)

• Current DMPs – model questionnaires

<administrative_data>
  <question>Who will be the Principle Investigator?</question>
  <answer>The PI will be John Smith from our university.</answer>
</administrative_data>

• Machine-actionable DMPs – model information

"dc:creator": [ 
  { 
    "foaf:name": "John Smith",
    "@id": "orcid.org/0000-1111-2222-3333",
    "foaf:mbox": "mailto:jsmith@tuwien.ac.at",
    "madmp:institution": "AT-Technische-Universität-Wien"
  }
],

Use controlled vocabularies
Example (not the real model)

• Current DMPs – model questionnaires

<administrative_data>

<question>Who will be the Principle Investigator?</question>

<answer>The PI will be John Smith from our university.</answer>

</administrative_data>

• Machine-actionable DMPs – model information

"dc:creator": [ {
  "foaf:name": "John Smith",
  "@id": "orcid.org/0000-1111-2222-3333",
  "foaf:mbox": "mailto:jsmith@tuwien.ac.at",
  "madmp:institution": "AT-Technische-Universität-Wien"
}
],

Develop own concepts and vocabularies only when needed

-----------------------------------------------
RDA/WDS WG Certification of Digital Repositories

- **Core**
  - CoreTrustSeal.org

- **Extended**
  - nestor
  - DIN 31644

- **Formal**
  - ISO
  - ISO 16363
Core Trust Seal (CTS)

- RDA Initiative merging
  - Data Seal of Approval (DAS), and
  - ICSU World Data System (WDS) Regular Members certification
- Self assessment based on a checklist
- Guidance
  - online tools
  - documentation and webinars
- Review of the self assessment by two reviewers
- Assessments publically available
- Renewal every three years
List of RDA WGs

- Agrisemantics
- Array Database Assessment
- Blockchain Applications in Health
- Brokering Framework
- Capacity Development for Agriculture Data
- Data Citation
- Data Description Registry Interoperability
- Data Type Registries
- Data Usage Metrics
- Data Versioning
- DMP Common Standards
- Empirical Humanities Metadata
- Exposing Data Management Plans
- FAIR Data Maturity Model
- FAIRSharing Registry: connecting data policies, standards & databases
- International Materials Resource Registries
- Metadata Standards Catalog
- On-Farm Data Sharing
- Persistent Identification of Instruments
- PID Kernel Information
- Provenance Patterns
- Public Health Graph
- RDA / TDWG Metadata Standards for attribution of physical and digital collections stewardship
- RDA/CODATA Summer Schools in Data Science and Cloud Computing in the Developing World
- RDA/WDS Publishing Data Workflows
- RDA/WDS Scholarly Link Exchange
- Research Data Collections
- Research Data Repository Interoperability
- Rice Data Interoperability
- Software Source Code Identification
- Storage Service Definitions
- WDS/RDA Assessment of Data Fitness for Use
- Wheat Data Interoperability WG
List of RDA IGs

- Active Data Management
- Agricultural Data Interest Group
- Archives and Records Professionals for Research Data
- Big Data
- Biodiversity Data Integration
- Brokering
- Chemistry Research Data
- CODATA/RDA Research Data Science Schools for Low and Middle Income Countries
- Data Discovery Paradigms
- Data Fabric
- Data for Development
- Data Foundations and Terminology
- Data in Context
- Data policy standardisation and implementation
- Data Rescue
- Development of cloud computing capacity and education in developing world research
- Digital Practices in History and Ethnography Disciplinary Collaboration
- Domain Repositories
- Early Career and Engagement
- Education and Training on handling of research data
- ELIXIR Bridging Force
- ESIP/RDA Earth, Space, and Environmental Sciences
- Ethics and Social Aspects of Data
- Federated Identity Management
- From Observational Data to Information
- Geospatial
- Global Water Information
- Health Data Interest Group
- International Indigenous Data Sovereignty
- Libraries for Research Data
- Linguistics Data
- Long tail of research data
- Mapping the Landscape
- Marine Data Harmonization
- Metadata
- National Data Services
- Open Questionnaire for Research Data Sharing Survey
- Physical Samples and Collections in the Research Data Ecosystem
- PID
- Preservation e-Infrastructure
- Preservation Tools, Techniques, and Policies
- Quality of Urban Life
- RDA/CODATA Legal Interoperability RDA/CODATA Materials
- Data, Infrastructure & Interoperability
- RDA/NISO Privacy Implications of Research Data Sets
- RDA/WDS Certification of Digital Repositories
- RDA/WDS Publishing Data Cost Recovery for Data Centres
- RDA/WDS Publishing Data
- Repository Platforms for Research
- Reproducibility
- Research Data Architectures in Research Institutions
- Research data needs of the Photon and Neutron Science community
- Research Data Provenance
- Sharing Rewards and
- Small Unmanned Aircraft Systems’ Data
- Software Source Code
- Structural Biology
- Virtual Research Environment
- Vocabulary Services
- Weather, climate and air quality

FACULTY OF INFORMATICS
Contact RDA Austria - Board

<table>
<thead>
<tr>
<th>Name</th>
<th>Position</th>
<th>Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andreas Rauber</td>
<td>TU Wien, Chairman</td>
<td><a href="mailto:rauber@ifs.tuwien.ac.at">rauber@ifs.tuwien.ac.at</a></td>
</tr>
<tr>
<td>Barbara Sánchez Solís</td>
<td>TU Wien, Chairman Deputy</td>
<td><a href="mailto:barbara.sanchez@tuwien.ac.at">barbara.sanchez@tuwien.ac.at</a></td>
</tr>
<tr>
<td>Raman Ganguly</td>
<td>University of Vienna, Secretary</td>
<td><a href="mailto:raman.ganguly@univie.ac.at">raman.ganguly@univie.ac.at</a></td>
</tr>
<tr>
<td>Tomasz Miksa</td>
<td>TU Wien, Treasurer</td>
<td><a href="mailto:miksa@ifs.tuwien.ac.at">miksa@ifs.tuwien.ac.at</a></td>
</tr>
<tr>
<td>Paolo Budroni</td>
<td>University of Vienna, Secretary Deputy</td>
<td><a href="mailto:paolo.budroni@univie.ac.at">paolo.budroni@univie.ac.at</a></td>
</tr>
<tr>
<td>Seyavash Amini</td>
<td>Ivocat, Legal Advisor</td>
<td><a href="mailto:amini@ivocat.de">amini@ivocat.de</a></td>
</tr>
</tbody>
</table>

Slides „RDA in a Nutshell„ from RDA Global, Slideshare - [http://www.slideshare.net/ResearchDataAlliance](http://www.slideshare.net/ResearchDataAlliance)

Contact RDA Global:
Email: enquiries@rd-alliance.org
Web: [www.rd-alliance.org](http://www.rd-alliance.org)
Twitter: @resdatall
Thanks