Research Data Management for data on renewable materials and products

Dr. Ana Slavec

Consulting statistician at the InnoRenew CoE

RDA Europe Ambassador for Engineering/Renewable materials

Visit of Faculty of Forestry and Wood Technology at Mendel University

Brno, 4. 11. 2019

Breaking the ice

- What type of data do you collect in your research?
- In what file formats do you save your data?
- Where do you store your research data?

Outline

- Why should this interest us?
- What is Research Data Management?
- What is the Research Data Alliance?
- RDA Recomendations and Outputs
- How to get involved?

Why should this interest us?

1. **Ethics**
   - RDM is part of the responsible conduct of research, i.e. the practice of scientific inviestigation with integrity
   - Reproducibility crisis in science (well-managed and accessible data allows others to validate and replicate findings)
2. **Requirements imposed by funders and publishers**
   - DMP as a deliverable in Horizon2020
   - Open Science will be one of the pillars of the Horizon Europe, the next EU framework programme for research and innovation
3. **Saves times and resources (in the long run)**

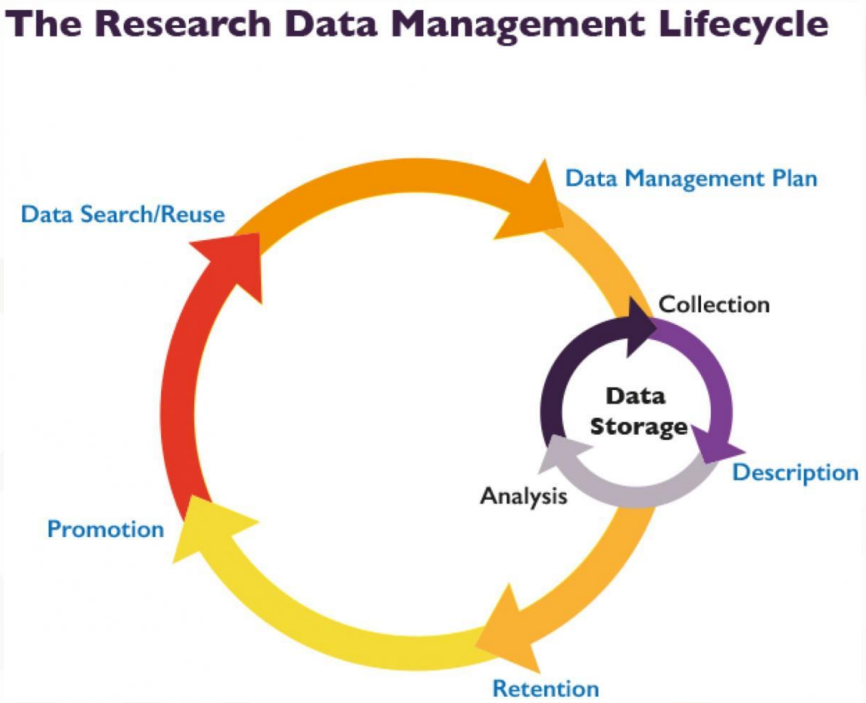
Key data challenges in engineering

- A lot of research is done in collaboration with industry which is reluctant to share data
- Fear of losing competitive advantage
- Fear of data being misused and misinterpreted
- Lack of metadata standards and ontologies (for certain engineering fields)
- Lack of domain-specific repositories (for certain engineering fields)

Research data management (RDM)

RDM lifecycle

- RDM describes the organisation, storage, preservation, and sharing of data collected and used in a research project.
- It involves decision about how data will be preserved and shared after the project is completed.



Data management plan (DMP)

- A formal document that outlines how data are to be handled during a research project, and after the project is completed
- What should a DMP include?
  - Description of data to be collected/created (content, type, format, volume)
  - Standards/methodologies for data collection and management
  - Ethics and intellectual property (highlight any restrictions on data sharing)
  - Plans for data sharing and access
  - Strategy for long-term preservation
- Useful resources: http://www.dcc.ac.uk/resources/data-management-plans

The FAIR data principles

- **Findable:** metadata and data should be easy to find for both humans and computers
- **Accessible**: once the users finds the required data, they need to know how can they be accessed, possibly including authentication and authorisation
- **Interoperable**: data usually need to be integrated with other data be interoperable with applications or workflows for analysis, storage, and processing
- **Reusable**: metadata and data should be well-described so that they can be used in different settings

Metadata

- „Data about data“
- Metadata helps to cite and disambiguate data
- Example: persistent identifier (e.g. DOI), pubication date, title, authors, description, keywords, licence, funding, related idnetifiers, etc.
- General (Dublin Core) or discipline specific (e.g. DDI)
- Metadata directory: http://rd-alliance.github.io/metadata-directory/

Recommended file formats

Type of data	Recommended formats	Acceptable formats
Tablular data	csv, .tab, .por, .xml	.txt, xls, .dbf, .ods, .sav, .dta, .mdb
Geospatial data	.shp, .shx, .dbf, .prj, .sbx, .sbn, .tif, .tfw, .dwg, .gml	.mdb, .mif, .kml, .ai, dxf, .svg
Textual data	.rtf, .txt, .xml	.html, .doc
Image data	.tif	.jpg, .gif, .tif, .tiff, .raw, .psd, .bmp, .png, .pdf
Audio data	.flac	.mp3, .aif, .wav
Video data	.mp4, .ogv, .ogg, .mj2	.avchd
Documentation and scripts	.rtf, .pdf, .xthml, .htm, .odt	.txt, .doc, .xls, .xml

Source: https://www.ukdataservice.ac.uk/manage-data/format/recommended-formats

Research data repositories

- General purpose repositories (e.g. Zenodo, Figshare)
- Institutional data repositories
- Domain specific repositories

- Registry of research data repositories: https://www.re3data.org/

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

Research Data Alliance (RDA)

What is RDA?

RDA is an international **member based organization** focused on the development of infrastructure and community activities that reduce barriers to data sharing and exchange, and the acceleration of data driven innovation worldwide.

With more than 8,800 members globally representing 137 countries, RDA includes **researchers, scientists and data science professionals** working in multiple disciplines, domains and thematic fields and from different types of organisations across the globe.

RDA is building the social and technical bridges that enable open sharing of data to achieve its vision of researchers and innovators openly sharing data across technologies, disciplines, and countries to address the grand challenges of society.

What does RDA do?

Members come together through self-formed, volunteer, focussed Working Groups, exploratory Interest Groups to exchange knowledge, share discoveries, discuss barriers and potential solutions, explore and define policies and test as well as harmonise standards to enhance and facilitate global data sharing & re-use.

RDA members collaborate together across the globe to tackle numerous infrastructure & data sharing challenges related to:

- Reproducibility
- Data preservation
- Best practices for domain repositories
- Legal interoperability
- Data citation
- Data type registries
- Metadata
- and so many more!

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 at https://www.rd-alliance.org/user/register

Why Join RDA as an Individual Member?

Individual Member Benefits

- ***Contribute*** to acceleration of data infrastructure development
- Work and ***share experiences*** with collaborators throughout the world
- ***Access*** to extraordinary network of colleagues with various levels of experience, perspectives and practices
- Gain greater ***expertise*** in data science regardless of whether one is a student, early or seasoned career professional
- ***Enhance*** the quality and effectiveness of personal work and activities
- ***Improve*** one's competitive advantage professionally and positioning oneself for leadership within the broader research community

Individual RDA Members 8,810

How to get involved?

RDA IGs and WGs

	Interest groups	Working groups
Number (Aug 19)	61	33
Members	Experts from the community	Experts from the community
Purpose	Platform for exhange on various topics	Focus on specific goal (concrete output)
Outputs	Supporting outputs	Recommendations and supporting outputs
Duration	Active over longer periods	Fixed period 12-18 months
Groups relevant for engineers	• Research Data Management in Engineering IG • RDA/CODATA Materials Data, Infrastructure & Interoperability IG • Physical Samples and Collections in the Research Data Ecosystem IG	• International Materials Resource Registries WG • Persistent identification of instruments WG • International Materials Data Registries WG

Plenary Meetings

- Organised around the world every 6 months
- exciting & productive events bringing together a unique community of **data science professionals, from multiple disciplines and domains**;
- help move the community forward in **creating tangible deliverables** that improve data sharing across disciplines, technologies, and countries;
- heart of the plenaries are working meetings of **RDA Working & Interest groups** and new potential groups through **Birds of a Feather** meetings
- presentation of new **Outputs and Adoption** cases

15th Plenary Meeting

RDA in Czech Republic

- Website: https://www.rd-alliance.org/groups/rda-czech-republic

- Contact persons: **Tomas Mildorf** (University of West Bohemia) and **Karel Charvat** (Czech Centre for Science and Society)

Acknowledgments and resources

- Jeuse, A. 2019. Adopting RDA Recommendations and Outputs Across the research data lifecycle: https://www.rd-alliance.org/rda-outputs-overview-presentation
- RDA in a Nutshell (August 2019). https://www.rd-alliance.org/sites/default/files/attachment/RDA-in-a-nutshell-August-2019.pptx

This work has been produced with the support of the RDA Europe Ambassador programme. It has received funding from the European Union's Horizon 2020 (H2020) research and innovation programme via the RDA EU 4.0 project (ref. GA no. 777388).

Let‘s stay in contact

Ana Slavec

InnoRenew CoE

ana.slavec@innorenew.eu

@aslavec

RDA Europe Ambassador for Engineering/Renewable materials: https://www.rd-alliance.org/rda-disciplines/rda-europe-ambassadors

This work has been produced with the support of the RDA Europe Ambassador programme. It has received funding from the European Union’s Horizon 2020 (H2020) research and innovation programme via the RDA EU 4.0 project (ref. GA no. 777388).