A Report on the Research Data Alliance Plenary in March 2013
Gerald Lane, Susan Malaika
malaika@us.ibm.com

Abstract
The Research Data Alliance (RDA) is a newly formed organization to accelerate and facilitate research data sharing and exchange. The RDA's first meeting took place Gothenburg, Sweden on 18-20 March 2013. This document is a small summary of that meeting.

Introduction to the RDA
The Research Data Alliance (RDA) [1][2] is a newly formed organization to accelerate and facilitate research data sharing and exchange globally, including enabling research dataset access beyond university centers.

Three governments, the US, the European Union (EU) and Australia [1], have undertaken to provide funding to accelerate solutions to problems with research data sharing, and support the efforts of the RDA

Many kinds of research data are of interest to RDA, including geo-spatial, environmental, climatology, space, agricultural, food, and health. Some datasets are large – e.g., peta-bytes of astronomy data retrieved annually from one telescope.

Changing the way researchers approach dataset sharing is of interest to the RDA. For example, measurement systems are available to encourage researchers to "publish or perish". Similar systems for measuring the data that researchers publish and its quality, do not yet exist, but may be needed encourage dataset publication.

The RDA agenda is to establish Interest Groups and Working Groups [3] to draw individuals to the identified topics which require community definition, work, agreement and collaboration. The details are made available via postings on the RDA mailing lists. In general, RDA policy is to identify and adopt standards where they exist, but in some cases, e.g., for marine data, suitable standards are not available yet.

The RDA's first Meeting in Gothenburg, Sweden
Over 200 individuals from 31 countries were in attendance in Gothenburg on 18-20 March 2013 to represent their University, discipline or government. Countries represented included Australia, Germany, Denmark, South Africa, Sweden, UK, US etc.

Day one was the launch, day two consisted of working group activities [3] and day three focused on working-group feedback and then an open plenary session [4].

On day one, the launch speakers included:

- Farnam Jahanian, Director National Science Foundation in US, "*From Data to Knowledge*" [5]
- Neelie Krowes, EU Commissioner Digital Agenda, "*Opening Up Scientific Data*" [6]
- Duncan Lewis, Australian Ambassador to the EU and Nato*, " Towards an Open Access Research Data without Barriers or Borders – RDA Mission*" [7]

These officials initiated the launch by delivering inspirational talks about the need for solutions that share the data generated by government funded research in an organized and controlled manner.

The RDA is intended to operate as a community driven organization to accelerate global data sharing and exchange and discovery by facilitating research data sharing and exchange, use and re-use, standards harmonization, and discoverability. These goals will be achieved through the development and adoption of infrastructure, policy, practice, standards, and other deliverables. The current structure of the RDA is described in the Organization documents [8]. An outline of the RDA's principles and how it will operate was described in a presentation by Francine Berman, Ross Wilkinson, and Juan Bicarregui, entitled "*Research Data Alliance @ T=0: Goals, Principles, Component Groups*" [9] including the difference between Interest groups and Working groups and the Working group life-cycle as illustrated in the figure.
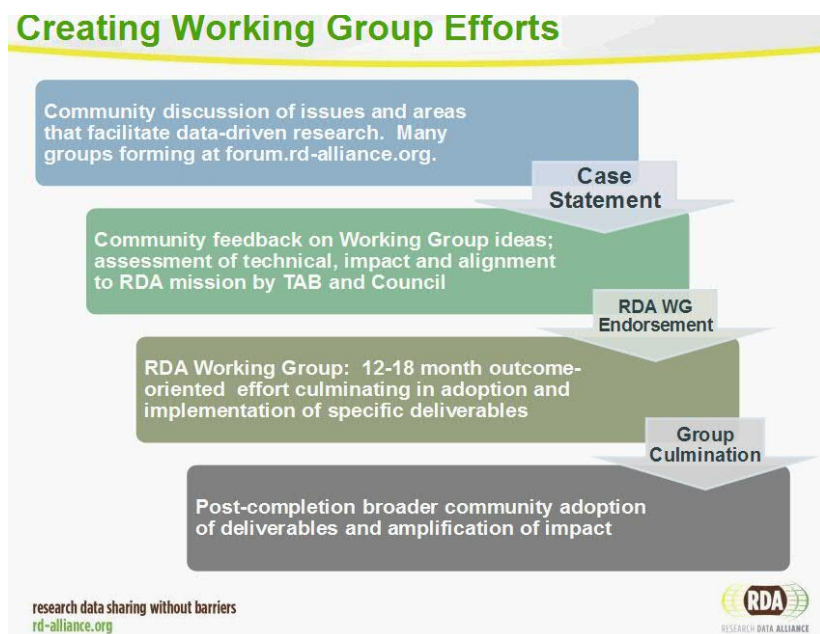
The RDA meeting in Gotherburg also included a presentation from Scott Brim on the IETF [10]. Chris Greer, Associate Director, IT Lab, National Institute of Standards and Technology and Francine Berman, Professor, Rensselaer Polytechnic Institute, have stated that Research Data Alliance (RDA) is to global data infrastructure what the Internet Engineering Task Force (IETF) is to the Internet. [11]

The RDA Working Groups and Major Themes
The working groups are in the early stages. Chairs have been assigned. The charters were reviewed in day two in Gothenburg. The goal is that working groups have short-term deliverables, so some working groups will start as interest groups in order to identify the short-term goals.

Major themes include universal dataset identification, metadata, trusted data collections and librarians.

- **Universal Dataset identification** themes include:
    - Higlighting existing global identification mechanisms – e.g., upc codes, ISO 14721
    - Relationship of identifier with URLs, and with work at the IETF

- **Metadata themes** include:
    - Thinking more broadly about what constitutes metadata, e.g., quality and provenance
    - Considering layers of data ownership: researcher; departmental; university or research institution; country or government
    - Characterizing data structures, formats, types, data model (extensible)
    - Making the metadata searchable

- **Data repository themes** include:
    - Trusted datasets certified by appropriate authorities, and associated with provenance quality profiles.
    - Community engagement and best practices for changing attitudes towards sharing data
    - Legal interoperability and respecting dataset rights across countries
    - Data citation mechanism to pinpoint datasets used to reach conclusions
    - Consolidating and compacting large datasets

- **Data librarian themes** include
    - Researchers and data librarians to keep the datasets fresh and available
    - Best practices and policies for data librarians
    - Dataset peer review including tools to support review of datasets and their quality

Conclusion and Next Steps

The attendees were enthusiastic about the direction of the RDA and the planned work. Working groups, met during day two and reports were shared on day three which concluded with an open plenary meeting to discuss feedback, and organizational growing pains, next steps and to announce a follow-on Plenary in the US in September 2013.

Related Materials

- [1] Research Data Alliance: http://rd-alliance.org/

- [2] Organization: http://rd-alliance.org/organisation/

- [3] Current Candidate Working and Interest Groups http://rd-alliance.org/working-groups/current-candidate-groups/

- [4] Agenda for Gothenburg Meeting http://rd-alliance.org/programme/

- [5] EU Commissioner, Neelie Krowes speech: http://admin.icordi.eu/Repository/document/Presentations/RDALaunch_Presentations/OpeningUpScientificData_NeelieKroes.pdf

- [6] NSF Director, Farnam Jahanian presentation: http://admin.icordi.eu/Repository/document/Presentations/RDALaunch_Presentations/FromDataToKnowledgeToDiscovery_FarnamJahanian.pdf

- [7] Duncan Lewis, Australian Ambassador to the EU and Nato speech: http://www.belgium.embassy.gov.au/files/bsls/HOM%20speech%20-%20towards-open-access-rd-without-barriers-borders-web.pdf

- [8] Research Data Alliance Organization Documents: http://rd-alliance.org/documents/

- [9] Francine Berman, Ross Wilkinson, Juan Bicarregui presentation: http://admin.icordi.eu/Repository/document/Presentations/RDALaunch_Presentations/ResearchDataAlliance_Goals%2CPrinciples%2CComponentGroups_FranBerman%2CRossWilkinson%2CJuanBicarregui.pdf

- [10] Scott Brim IETF Presentation http://admin.icordi.eu/Repository/document/Presentations/RDALaunch_Presentations/Internet%20EngineeringTaskForce_ScottBrim.pdf

- [11] Chris Greer, Associate Director, IT Lab, National Institute of Standards and Technology and Fran Berman, Professor, Rensselaer Polytechnic Institute http://www.cni.org/topics/ci/research-data-alliance/