Research Data Alliance: international sharing and exchange of research data [1]

  Submitted by bethplale on Mon, 2012-11-26 07:39   Tuesday, January 8, 2013 - 14:00 to 15:30
Event: Winter Meeting 2013 [2]
Session Type: Breakout [3]
Expertise Level: Beginner [4]
Abstract/Agenda:
The Research Data Alliance (RDA) is an emerging international organization whose goal is to accelerate data-driven innovation through the sharing and exchange of research data. With the rise of "Big Data" as a national and international priority, there is increased need for the infrastructure, policy, and practice that will power innovation. Efforts to develop and adopt common tools and infrastructure, harmonize data standards, and apply policy and best practice are needed to remove roadblocks to collaboration, and to accelerate data-driven research and discovery. The Research Data Alliance is being formed to facilitate such efforts. Membership in the RDA is open to everyone. The primary work of the Research Data Alliance will be conducted through short-term, outcome-oriented Working Groups, broader-focused Interest Groups, and effective international collaboration.

In this interactive session we will explore overlap and interest between RDA and ESIP working groups. The session will begin with a brief introduction of Research Data Alliance, with a focus on the current set of vanguard working groups. Through a moderated panel we will engage the audience in topics of interest including deeper understanding of vanguard working group activity, new topics for working group activity, and overlap. The panel will be made up of members from both RDA and ESIP.

Notes:
Stimulate dialog between ESIP and RDA.

Presentations

- Dr. Fran Berman, Research Data Alliance - Current Status
- Beth Plale - RDA @ ESIP Session
- Panel (Fran Berman, Beth Plala, Rebecca Koskela, Chris Lenhardt, Sayeed Choudhary, Ruth Duerr) - Exploring synergies between RDA and EISP

Actionables

- None mentioned... perhaps, any interested ESIP working groups / clusters / BoFs / committees look into submitting working group proposals to RDA

Discussions / Questions

- Introduction to RDA
- What are the opportunities for synergy / partnership between RDA and ESIP?
    - Revelent early interest groups for RDA - harmonization and use of persisten ID; metadata; legal interoperability; repository audit / certification
    - Points of overlap? Complementary activities?
- How can RDA help ESIP?
- How can a joint impact be made?
- How could RDA contribute to ESIP Winter Meeting Theme?
- What problems could RDA address that are out of scope for ESIP?
- Who are the first organizational partners? Should provide win-win scenario for both partners.
- RDA Early Activity Around Working Groups
    - Data Foundation - think, ISO network stack for data

- Communities and Engagement Group
- Metadata group - identifying and endorsing metadata solutions (balancing domain-specific and trans-disciplinary needs)
- Panel Statements
  - Sayeed - find optimal degree of overlap, lots ESIP could do for RDA (comment on docs / case statements, contribute experience)
    - What can RDA do for ESIP (i.e., the out of scope comment)? Try to avoid ESIP needing to extend itself or move into a new direction
    - Case: Systems geobiology in Yellowstone, need a mechanism to build bridges for multi-disciplinary science
  - Chris - represent ESIP in conversation
    - What is ESIP? (may help finding overlap between RDA and ESIP) It's all about data, core values include neutral, open, agile, etc. The members, people bringing experience and doing the value-added work. E.g., clusters form, graduate into working group, and eventually become committee. Air quality working group, became GEO air quality community of practice.   Success was due to putting the right people together at the right time. Data preservation standing committee, produced data citation guidelines, culminated in paper coming out of ESIP, now focusing on provenance. Energy and Climate working group...
    - Differences between RDA and ESIP, ESIP is organic and bottom-up, RDA has more structure applied. ESIP could be considered an innovation engine that feeds into RDA.
  - Rebecca - Why is DataONE involved in RDA?
    - DataONE is small, partnered with ESIP for education modules and joint workshops. Relationships with external organizations help overcome small size. DataONE is tasked with joining collaborations, attractive nature of RDA is the applied nature (12-18 month projects). DataONE chose topic planned on working on anyway (metadata), proposed metadata group wants to maintain a list of metadata standards. Two questions commonly asked (could be addressed by group), where do I put my data and what metadata standards should I use? Creating a minimum metadata standard is a longer term goal. Comment on proposed case statements!
  - Ruth - Experiences as a member of the ESIP community
    - Working with WGs including Preservation and Stewardship. Difference between RDA and ESIP, RDA is trying to be *very* cross-disciplinary, ESIP is concentrated in the Earth Sciences. RDA also more international. Overlap can occur, want to build something broadly applicable to data, go to RDA, if you want to specialize, go to ESIP. ESIP doesn't directly try to develop standards, but works on things that may someday become standards (which is where RDA could come in). ESIP is very bottom-up, very few formal processes. Good for getting things started, but sometimes a little formality is better (another place for cross-fertilization). ESIP working groups / clusters can be significantly longer than 12-18 months, ESIP is a good place for longer term (less *applied*) projects.
  - Audience Contributions
    - Brian Wee - One way for RDA to differentiate from ESIP, focus more on interdisciplinary efforts. ESIP has been domain agnostic in terms fo participants (but applications focused on Earth Sciences). RDA could work with domain scientists more. Put together groups of participants that span domains.
    - Rahul Ramachandran - data issues are specific domain, data discovery access and mining cluster in EarthCube, targeted domains within Earth Science with high volume datasets. Pushback from charette, geologists et al. have very different needs than high volume science.
    - Chris Lenhardt - All domains say that there problems are unique... but often when looking at the problem in more detail, they have similar problems
    - Mark Parsons - medical science doesn't care about polar projections, RDA can help bridge across domains, help get away from "not invented here"

syndrome for domains.
- Sayeed - Low-level engineering (e.g., identifiers) are shared amongst domains.
- Fran - Potential for RDA is to create things and try them, will not produce "esperanto" solutions, but avoid approaches where "everyone rolls there own" solution. RDA can help communities determine the interoperability plugs that other communities can use.
- ??? - What are the commonalities and differences between Earth Science and Computer Science work..
- Beth - Low-hanging fruit from engineering perspective (identifiers). Minimal metadata sets are more challenging.
- ??? - Is there a common level of metadata that can really work?
- Beth - everyone can implement Dublin Core, but everyone has different opinions.
- Mark Parsons - Point isn't to find least common denominator, but rather finding plugs / gateways for interconnection metadata systems.
- David ??? - Why strong limitation to things that can be accomplished by volunteers? Charter for working groups cannot require "extra resources". Effectiveness of IETF was partially due to NSF NET contributions / resources. Why deliberately limit RDA to things that can be achieved by volunteers?
- Fran - Peoples time for RDA does not have to be disjoint from there regular day-jobs. Why RDA doesn't pay b/c it is community driven, and does not have a sustainable business model (though it aims to get one). Many projects get to a point that you can't do something on your own. RDA is there to connect people who need to accomplish things to make progress. Many working groups will emerge from things people are doing in there companies, research labs, etc. Proximity of community organization with funding agencies will be a good thing for contributors (bring activities to the attention to the funding agencies). How to publicize results of working groups, which in term benefits RDA, working groups, and volunteers?
- Sayeed - As working groups complete work, funding agencies will deploy or spread results (result of working group could be a proposal).
- Sarah Graves - RDA's technology focus is its advantage.
- Fran - Get solutions that are good, not necessarily perfect.
- ??? - Institutions are faced with the task of building data management policies. If everyone shares, this task becomes less daunting. In the future it will be easier to implement data management policies of a specific type
- ??? - Everyone wants to define a set of standards. Focus on the adapters instead of the standards. Identify specific obstacles to answering cross-disciplinary questions. Look to see what the adapter is that needs to be written to connect to two disciplines.
- Fran (response) - Working group needs both those creating the adapter and the users who will eventually plug them in. Lots of room for new working groups, in a few years there will be a whole new set of problems people are looking into (as old problems are solved and new ones discovered).
- Sayeed - RDA is what we make of it. Some communities have places where they get together and discuss interoperability issues (e.g., ESIP for Earth Science). RDA could be a place where representatives from each interoperability community come to discuss broader issues.
- Chris - Need to continue to involve domain scientists, if its just hte technologists, that's a problem.
- Rahul R. - Anyone can submit RDA working groups. Might be a good idea for ESIP working groups / clusters to submit proposals to RDA for broader participation.

Actions:
Any interested ESIP working groups / clusters / BoFs / committees look into submitting working group

proposals to RDA.

Session Leads:

Name: [Beth Plale](#) [5]
Organization(s): [Indiana University](#) [6]

Notes takers:

Name: [Eric Rozell](#) [7]
Organization(s): [Rensselaer Polytechnic Institute](#) [8]

Creative Common License: Creative Commons Attribution 3.0 License
Teaser: RDA session explores overlap and interest between RDA and ESIP working groups
Keywords: [international](#) [9]
[Big Data](#) [10]
[Infrastructure](#) [11]
[practice](#) [12]
[policy](#) [13]
[data harmonization](#) [14]

Source URL: https://commons.esipfed.org/node/749

Links
[1] https://commons.esipfed.org/node/749
[2] https://commons.esipfed.org/taxonomy/term/464
[3] https://commons.esipfed.org/session-type/breakout
[4] https://commons.esipfed.org/taxonomy/term/260
[5] https://commons.esipfed.org/node/1221
[6] https://commons.esipfed.org/taxonomy/term/619
[7] https://commons.esipfed.org/node/292
[8] https://commons.esipfed.org/taxonomy/term/221
[9] https://commons.esipfed.org/taxonomy/term/504
[10] https://commons.esipfed.org/taxonomy/term/348
[11] https://commons.esipfed.org/taxonomy/term/347
[12] https://commons.esipfed.org/taxonomy/term/505
[13] https://commons.esipfed.org/taxonomy/term/506
[14] https://commons.esipfed.org/taxonomy/term/507