

## 20161123 - Notes from the RD-A & Australian DP IG meeting

### Participants:

Australian Astronomical Observatory (AAO): Simon O'Toole  
Australia's Academic and Research Network (AARnet): Ingrid Mason  
Australia National Data Services (ANDS): Rowan Brownlee, Mingfang Wu  
Curtin University: Muhammad Azeem Sadiq  
National Computational Infrastructure: Jinbo Wang  
NSW Department of Primary Industries: Timothy Smith  
RDA DP IG Chair/The iSchool at Illinois: David Dubin  
Terrestrial Ecosystem Research Network: Siddeswara Guru  
University of Queensland: Helen Morgan

Chair: David Dubin

Notes: Mingfang Wu

### Notes:

#### RD-A Data Provenance Interest Group (DP IG) update (Dave):

- [RD-A DP IG](#) is working on a proposal for spinning off DP working group (WG). The WG to be started will focus on provenance solution patterns - the idea is to adapt the concept of design pattern from software development to match a particular provenance use case to various solution options and provenance vocabularies. Here are a couple of points about the design patterns from [Dave and Nick's presentation](#) at the last (8th) RD-A plenary.
  - Some ways of doing things in provenance are better than others. This activity is to generate provenance design patterns (for any provenance task such as representation, transmission, use etc.) perhaps in response to a series of provenance use cases that we would generate.
  - The patterns should relate to core RDA interests, perhaps data/data and data/people relationships.
- The RD-A IG has a longer time frame and a broader scope, while WG has a limited time frame (about 1.5 years) with a specific goal. Both IG & WG have outputs. The RD-A DP IG has worked on a range of demonstrations, chaired sessions at RD-A plenaries, collaborated with, provided consultation and background information/readings to other RD-A IGs & WGs.
- The proposal for the (DP) WG includes charter and various plans: value proposition, engagement, work plan and initial membership. Resources about starting the DP WG are publically available at Github: <https://github.com/RDAProvIG/WGplanning>. There are two steps to engage and contribute:
  - Anyone who would like to contribute the WG development, please email Dave to express your interest. Dave will then send you an invitation that gives you contributor access right to the Github repository.
  - Dave & Nick will put together a team to work on the proposal. The high level view of the proposal is available by following the above github link. This group

may contribute to specific activities required by the proposal, especially for activities in the Value Proposition section.

- It is planned to have a draft work plan for creating the DP WG by end of Feb. 2017.
- It would be good to get any example of how provenance vocabulary and documentations have been applied by this group for the purpose of addressing a particular need.
- Dave will try to put together a list of design patterns and use cases. In future we may have an agenda item per meeting to discuss a user case and its design patterns.

### **Updates for Australian DP IG Participants:**

Tim:

- Tim and his team are looking at getting into this (provenance) space. A provenance problem had in past is having multiple versions of a project and don't know what have been done to each. It would be useful to have an identity for a project, it will be really useful to tie up all identifies of different versions, so changes of a project can be tracked.

Ming:

- Andrew (Treloar) and Ming wrote an article "Provenance in support of ANDS' four transformations". The article is published in International Journal of Digital Curation. [doi:10.2218/ijdc.v11i1.416](https://doi.org/10.2218/ijdc.v11i1.416)

Ingrid:

- Representing AARNet working on [data life cycle framework](#), which is a collaboration between Nectar, RDS, ANDS, AARNet and AAF. The project is discussing system identifiers. It would be help to get a list of use cases and provenance design patterns.

Simon:

- AAO is currently building a large data archive. The International Virtual Observatory Alliance is currently developing provenance data model, which is based on W3C PROV-DM. AAO is looking ways of implementing the data model into its systems.

Jingbo:

- Working with Nick closely on ocean color provenance capture report. Near future plan is to generate in-situ provenance report: whenever users run processing pipeline, activities in a pipeline will be captured in a log file, a provenance report will be generated from the log file. The provenance report will be both human and machine readable.