

Data Curation & Preservation Summary

*IVOA Interop Meeting
Sep 26-28
Hawaii*

Friday Sep 27 | 11:00-12:30 Naupaka 7

Research Data Alliance

- F. Genova serves on the Technical Advisory Board and has been participating in plenary meetings
- RDA is an initiative of the US (NSF), EU (iCordi) and Australia
- Goal: “The Research Data Alliance implements the technology, practice, and connections that make Data Work across barriers.”
- 8 Working groups, 18 Interest groups, participation is open to all
- Francoise recommends involvement by IVOA members to provide expertise, follow discussions, influence outcomes

Persistent Identifiers

Granularity, Versioning, Compound Objects

- A. Rots revisited the issue of PID in light of conversations at the RDA plenary meeting and personal experience at CXC
- Granularity: ideally Persistent Identifiers schemes should allow for different levels of Coarseness with a notion of hierarchy (or containment)
- Versioning: desirable, required for result reproducibility, but rarely used in ongoing archival operations
- Discoverability: strongly depends on the availability of descriptive metadata

AAS Update

- C. Biemesderfer discussed the Society's view on data publication and citation
- “Data sets are real scholarly products – but they require process to be valuable for a long time”
- Data publishing has a process: Peer review, Copy editing and markup, Identification, discovery, Presentation and access, Preservation
- “Citability in AAS articles depends on the heritage of the cited item”
- Complexity and curation needs may be significant for datasets in the long-tail of data products

Data Publishing Issues

- A. Accomazzi discussed issues related to the topic of “data publishing” as currently discussed in the wider community
- Many platforms have appeared in last few years, and astronomers are starting to use them
- Simple solutions (Figshare, Zenodo) are attractive but do not provide VO metadata and layered services
- Traditional solutions (VizieR, Astro archives) provide curation, discoverability but require more complex workflows, effort
- We should provide solutions, guidance, support to scientists ready to “publish” their data