Data Curation & Preservation Summary

IVOA Interop Meeting
Nara, Japan
Dec. 7-11, 2010

Presentations

- Arnold Rots: DC&P in the VAO
  - Focus on curation of Astronomical Digital Objects (publications, datasets, objects)
  - Link ADO and make them available to applications (search, discovery)
  - Create prototype repository and workflow for capturing high-level data products associated with journal articles

- A. Accomazzi: Advanced Linking in the VAO
  - Formalize links and relationships between astronomy resources using ontologies
  - create knowledge base of literature, instruments, observational metadata
  - support semantic applications integrating metadata across archives, facilitating discovery and seamless search

Presentations (cont.)

- Jose Enrique Ruiz: Wf4Ever
  - EU-funded project enabling workflow preservation
  - Astronomy one of two case studies - VO Data Processing
  - Capture Research Objects and and preserve their integrity and authenticity
  - Uses Taverna and myExperiment

- Pierre Le Sidaner: Identifiers in the VO Paris Data Center
  - Attempts to define the use of IVOA identifiers for resources at ObsPM
  - IDs capture project, institute, organization, service
  - IDs can be used as OAIS internal keys
  - Challenge for IDs is long-term survival of institution, migration of data products

Presentations (cont.)

- Ray Plante & Alberto Accomazzi: IVOA identifiers
  - Registry built for resource resolution, not persistence and multiplicity
  - Dataset identifiers were designed to provide one level of indirection (like DOIs do), provide greater persistence
  - Norman Gray investigating long-term URIs (~100 years)
  - Upcoming work on this with NSF DataNet and at AAS Boston (May)

- Alberto Accomazzi: Publication of IVOA Standards
  - Proposal: submit IVOA standard document to arXiv
  - Automatically indexed in ADS, citeable, version control, great exposure
  - Once a REC, document is "published" as refereed paper in ADS
  - Exec, Document Coordinator and/or WG chairs to set policies, submit papers as appropriate

Conclusions

- DC&P is alive and well!

- New initiatives and policies in US, Europe geared towards preservation of scientific data will inform and influence our operations (e.g. NSF data policy)

- Publishing Data and linking to literature now a hot topic in many disciplines, expect new recommendations and/or mandates in next few years

- Time to review and revive our efforts with publishers, archives. Use mailing list and speak up to participate