Why Publish to the VO?

Séverin Gaudet

Canadian Astronomy Data Centre
Canadian Advanced Network for Astronomical Research
IVOA Mission Statement

• “Allow astronomers to interrogate multiple data centers in a seamless and transparent way”
• “Provide new powerful analysis and visualization tools within that system”
• “Give data centers a standard framework for publishing and delivering services using their data.”
• “Standardization of data and metadata”
• “Standardization of data exchange methods”
• “Use of a service registry…”
• “A framework which enables data centers to provide competing and co-operating data services…”
IVOA Goal

To enable science!
- Data discovery
- Efficient data access
- Interoperable analysis tools
- Interoperable data
- Scalable visualization and computing
- Data Mining
The Data Centre Reality

Mission focus
- Serving primary users
- Requirements from user committees

Perception of VO services
- Increased scope
- An extra layer of software to build
- A secondary service delivery mechanism to be maintained
- Retrofitting is expensive
- Data engineering is hard
A Different Perspective

Why Publish in the VO?

Why Publish with the VO?
Why Publish with the VO

Use VO technologies in the data centre software stack

• Technical standards
• Format standards
• Data models
• Existing code bases
• Existing tools

Why?

• Leverage the intellectual investment from a wide community
  • Scientific
  • Technical
• Fundamental components for data centres
• No longer an “extra” layer
An example

The Canadian Astronomy Data Centre

If you have used CADC facilities for your research, please include the following acknowledgment:

The Canadian Astronomy Data Centre (CADC) is supported by the National Research Council of Canada with the support of the Canadian Space Agency.

Search by Target: All Archives | Search | Advanced Search

English | Français

NRC is currently renovating this website to improve its functionality in both official languages. Due to the complexity of the site, these improvements are taking place in stages. We regret the inconvenience to our users and will update this notice to report our progress.

In the meantime, if you are unable to access the information you require, please contact us at cado@nrc.ca for assistance in the official language of your choice.

Thank you for your patience.

Canadian Astronomy Data Centre - CoSADIE Data Centre Forum
Canadian Astronomy Data Centre

HST  FUSE  MOST

Heterogeneous collection:
- multiple missions and facilities
- multiple wavelengths

Pointed and survey observations

Many different data models

Continuous evolution

No Interoperability!
New approach

New perspective on VO standards

• VO has moved from periphery to core
• Re-using the intellectual investment of the IVOA

Applying to new CADC use cases:

• PI distribution: VOSpace, GMS, SSO
• Collaboration repositories: VOSpace, GMS, SSO
• Virtual machine repositories: VOSpace, GMS, SSO
• Sharing virtual machines: VOSpace, GMS, SSO
• Data publishing: VOSpace, GMS, SSO
• User databases: TAP, GMS, SSO
New approach
VOSpace – User Storage

- User storage
  - Programmatic access
  - Browser UI
- TAP query input and output
- Data publishing (DOI)
- Project shared repository
- Virtual machine image repository
- Cloud processing file persistence
- Telescope PI data distribution
The Challenges

New projects (designing)
- Look to VO technologies
- Participate to inform new standards

Existing projects (retrofitting)
- Look for opportunities for incremental additions
  - Adding new functionality
  - Upgrading capacity
  - New use cases
- Participate to inform new standards
The Challenges

Data engineering
- For discovery and access
- Internal data models vs. VO data models
- Data model translation: pre-generated or on the fly?
- The only path to interoperability
- ObsCore is a good starting point

Discovery and Operations (e.g. cutouts)
The Next Generation Virgo Cluster Survey