Upgrades, Clean-ups, and the Registry of Registries
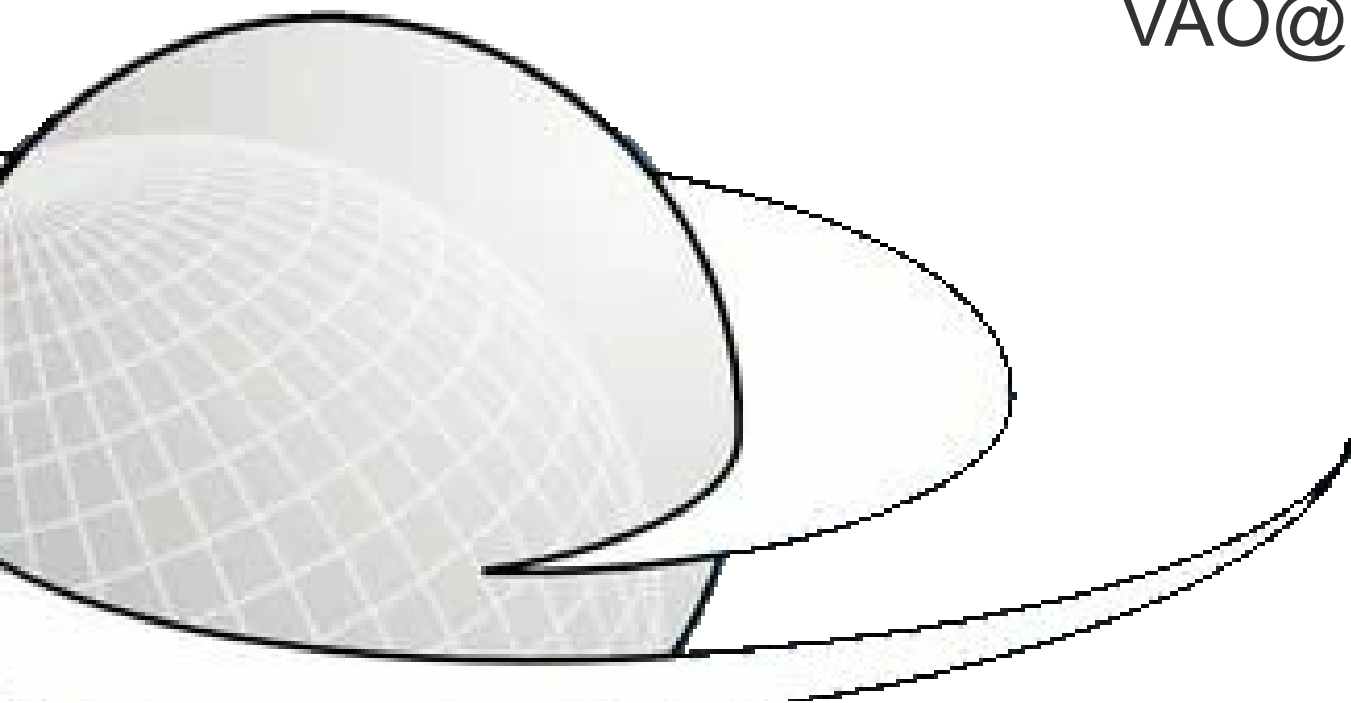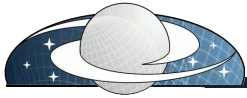
Ray Plante
VAO@NCSA

Upgrading to latest schemas

- Evolution
  - VODataService v1.1 => Rec in 2010
  - TAPRegExt, SimpleDALRegExt both require VODS v1.1

- Some registries (incl. VAO) still uses VODataService v1.0
  - It's okay—nothing breaks, but...
  - VAO is upgrading as part of move to RegTAP

- Recommend Registries upgrade records
  - Stylesheet on TWiki: RegUpgradeToVODS11
  - Uses XSL2; saxon engine provided

VAO Upgrade Phase 2

- Phase 2 will focus on record content
  - Coordinated with release of our new publishing tool
  - Campaign to invite publishers to review their records, improve content
- We will make recommendations to publishers
  - Create draft records that with improvements which they can review, improve and/or accept.
  - **We plan to institute a common pattern for registering resources from a publisher**
    - → Recall past presentations by M. Demleitner

Motivation for Re-org

- The registry framework allows multiple ways of registering a set of resources
  - This flexibility causes real confusion for users
- Issues:
  - Collections vs. Data Services
    - → Some have registered collections by themselves
    - → Some have registered only the services that access a collection
    - → Some have done both
  - Mirrored and derived collections
- What users see
  - Discovery tools (DDT, VAO Directory) usually display search results as a flat list
  - Users may see multiple all incarnations of a collection
    - → The difference between them is not obvious
    - → Some they can search, others they can't, ...
    - → Lack of consistency means tools can't easily help user understand results.

A Common Registration Pattern

- Encourage/enforce a single pattern for registering resources
  - Want tools to be able to recognize the use of this pattern to provide a more meaningful display of search results.

- Basic Features
  - Every publisher registered as an Org; authorities registered
  - All data is registered as a Collection
    → Science keywords associated with collection description
  - Services that access a collection registered separately
    → Make use of relationships to connect collection and services
  - Identify resources that a mirror or derived (with relationship link back to parent collection)
  - Make use of VOSI to expand available metadata (table columns)
  - Stategic insertion of keywords & "micro formats"

Whose with us?!

- How far do we need to go?
  - Hope: fixing a sufficient number of smaller publishers (non-Vizier/HEASARC) will be enough for registry tools to display smarter results
  - Programmatic users will have simpler/consistent set of standard queries for getting what they want
- Can we help other registries?
  - GAVO is already doing this
  - ESAC is being rebuilt

Registry of Registries (RofR)

- Mechanism for discovering other registries
- IVOA Note: Registry of Registries
  - Explains recipe for discovering and harvesting from registries
  - Contains some errors! Needs a revision.
  - Does this need to become a standard?
  - Reference Software: IVOAHarvester
    - → http://trac.us-vo.org/project/nvo/wiki/IVOAHarvester
    - → Implements RofR Note, validates records
- Maintenance
  - Ray Plante runs it on behalf of IVOA
    - → Software at http://trac.us-vo.org/project/nvo/wiki/RofR
  - Consider future beyond 2014
  - Developments:
    - → Allow Registry Admins update their Registry Records (currently done by hand).
  - What is the will of the WG?