How to Publish to the VO

Marco Molinaro
Astronomical Data Center Forum
10-11 June 2013, Heidelberg
Putting together
Formats – Protocols – Models – Registry
to reach a goal

 interoperable astrophysical data resources

How?
Implementation/coding
Datasets/archives/DBs
Dynamical content, data policy
Resource Registry details
Getting started...

IVOA – Publishing in the VO

http://wiki.ivoa.net/twiki/bin/view/IVOA/PublishingInTheVONew

Publishing Data into the VO

- Publishing Data into the VO
  - 0. Introduction
  - 1. Questions & Answers about what type of data to be published into the VO
  - 2. Registry for VO Data Service Discovery
  - 3. Toolkits to publish data into the VO
  - 4. Developer's corner: other useful software tools and libraries for VO development
Getting started...

IVOA – Publishing in the VO

http://wiki.ivoa.net/twiki/bin/view/IVOA/PublishingInTheVONew

1. Questions & Answers about what type of data to be published into the VO

Q: I want to publish a combination of generic data (e.g., spectra and images), possibly with links to each other. Are there toolkits/services for doing so?

Q: I have a working database and I want to publish the data therein in the VO. What can I do?

Q: I am already using tools to leverage existing libraries. Where can I find them?

Q: I want to develop my own code to publish data. What tools can I use to publish it?

2. Registry for VO Data Service Discovery

Q: How does my service appear in Client Applications?

Q: I don’t want to write any code. Are there applications I can install and configure in order to serve my data?

Q: I don’t want to publish code. How do I publish my data? What tools can I use to publish it?

Q: I need to publish something, but I don’t want to write any code. What tools can I use to publish my data?
Getting started...

IVOA – Publishing in the VO

http://wiki.ivoa.net/twiki/bin/view/IVOA/PublishingInTheVONew

### 4. Developer’s corner: other useful software tools and libraries for VO development

<table>
<thead>
<tr>
<th>Tool / Library Name</th>
<th>Contact Point</th>
<th>Date of last release</th>
</tr>
</thead>
<tbody>
<tr>
<td>STIL</td>
<td>MarkTaylor</td>
<td>12/2012</td>
</tr>
<tr>
<td>STILTS</td>
<td>MarkTaylor</td>
<td>3/2013</td>
</tr>
<tr>
<td>SAVOT</td>
<td>AndreSchaaff</td>
<td>4/2010</td>
</tr>
<tr>
<td>CDS UWS Library</td>
<td>CDS cds-question (at) unistra.fr</td>
<td></td>
</tr>
<tr>
<td>CDS ADQL Library</td>
<td>CDS cds-question (at) unistra.fr</td>
<td>3/2012</td>
</tr>
<tr>
<td>UCD assignment tool</td>
<td>sebastien.derriere (at) astro.unistra.fr</td>
<td>10/2008</td>
</tr>
</tbody>
</table>

- **Tool / Library Name**: The name of the tool or library.
- **Contact Point**: The contact person for the tool or library.
- **Date of last release**: The date of the last release.

**DAL Services Supported**

- SIA/SSA/CS/TAP
- SIA/SSA/SCS/SLAP/TAP
- SCS, SIAP, SSAP: TAP

**Jobs**

- Implemented but not yet integrated
- SCS, SIAP, SSAP, TAP

**ADQL**

ADQL (Astronomical Data Query Language) is a SQL-like language. To sum up ADQL adds to the basic query a cone search around 10°5′ with a radius of 2′.
Your way to the goal

- Identify your path to goal, consider:
  - Technological expertise
    - Including coding longing
  - Dataset complexity
    - Single table, set of images, images and catalogues, multiple tablesets, tablestes with image and/or spectra access, ...
  - Structural starting point
    - From sparse files to fully structured archive
- but also...
  - Dataset content dynamics
    - One-shot DR / incremental dataset
  - Release data policy
    - Public, private, mixed
Lost in translation

- Identify your path helps in
  - Not loosing your mind while diving in the IVOA-Doc ocean
    - TECHNOLOGICAL EXPERTS must consider themselves good divers!

- First step in defining your strategy
  - Clear ideas of what you want to expose
  - How you want to structure the deployed resource(s)
Simple scenario 1: no expertise, only a bunch of files

- Toolkits
- Aided DC or toolkit publishing
  - National VO project contact
  - Existing VO involved DC
- Registration

- Issues
  - Data formats / metadata content
  - Dataset complexity
Simple scenario 2: Medium expertise, small archive

- Toolkits
- Libraries
- Specifications
- Publishing Registry?

- Issues
  - Starting dataset structure
  - Data policy
Simple scenario 3: Tech Expert with structured archive

- Standards & Protocols
- Code from scratch
  - Probably implementing existing library efforts
- Publishing (or full) Registry?

- Issues
  - Archive translation layer
  - Dataset dynamic upgrade
DC contributed examples

- Saada
- ESA
- CDS
- IA2
Saada: A tool for data publishing
Laurent MICHEL: Observatoire de Strasbourg
http://saada.u-strasbg.fr

Used by the XMM-Newton Survey Science Centre

- **Saada builds databases from data files**
  - No Code to write
  - Storage of heterogeneous dataset
  - Can host multiple data collections
  - Support persistent links between data
  - Meta-data tagging (ucd, units...) « by hand »
  - Access via Web interface or VO protocols

- **A Java layer on the top of an RDBMS**
  - PostgreSQL, MySQL or SQLite
SAADA Admin tool

Database Map
- Collections
  - Starting
    - TABLE
    - ENTRY
    - IMAGE
    - SPECTRUM
    - MISC
  - Ending
- Ending

Root Panel
- Starting.MISC

Data Management
- Create Collection
- Load Data
- Edit Filter
- Manage Data
- Manage Meta Data
- Manage Relationships

Data publication
- Database Installation
- Web Publishing
- VO Publishing
Centre de Données astronomiques de Strasbourg

Reference service for object names

Access to journal catalogue data

An interactive ‘portal’

VO interfaces alongside web and other interfaces
e.g. query VizieR via:

- dedicated web interface
- Topcat
- Aladin, via DS9
- scripts or web services
- TAP – Table Access Protocol

```sql
1  -- 2mass_around_M1
2  -- the objects from 2mass around M1 within 1 arcmin
4  FROM "II/246/out"
5  WHERE 1=CONTAINS(POINT('ICRS', raj2000, dej2000), CIRCLE('ICRS', 83.633083, 22.0145, 1/60))
```
VizieR TAP Interface  (also useable via other TAP clients)

The TAPVizieR service provides VizieR tables using ADQL (a SQL extension in Astronomy).

Search tables

Search by catalog, author's name, word(s) from title, position (resolved by Sesame), ...
  e.g: Veron, 2Mass, redshift, M31...
  Note: The vizieR capability takes advantage of METAdatat
VO-Dance

User Layer

Service Publisher restful endpoint

Service Publisher

SQL Composer

Universal Controller

Policy Manager

internal DB

VO-Dance Administration Interface

DB / Resource Layer

storage

storage

storage

storage

DB
Resource Registry

- Resource
  - Data collections
  - Services
    - Registry
- Curation and Content
- Capabilities and Interfaces
- Validation
- Where? (RofR, full harvestable, publishing)
Registry portals

http://registry.euro-vo.org
Thank you!