VO Data Keeping-up Agent
A data fetching agent for the VO

Omar Laurino  Riccardo Smareglia

INAF - OATS, Italy

IVOA InterOp, Victoria - May 17th 2010
Users have to trigger their interaction with the VO in order to find and fetch data (notice the natural direction of the arrows in the diagram)
Users have to trigger their interaction with the VO in order to find and fetch data (notice the natural direction of the arrows in the diagram).

So, the interaction with the VO (in order to fetch data) is basically synchronous. But the Virtual Observatory is getting more and more alive! New services and new data continuously pop up, especially when the time domain comes into play (e.g. ELTs: LSST, OWL, ALMA). Also, complex queries may take some time to run.
How to improve this use case

What if we could turn the arrows around?

Omar Laurino (INAF - OATS)
How to improve this use case

What if we could turn the arrows around?

We can’t... but we can introduce a new actor: an agent.

Omar Laurino (INAF - OATS)
Here comes the new gal in town

With Vodka users can be kept updated asynchronously and automatically.
Here comes the new gal in town

With Vodka users can be kept updated asynchronously and automatically.

Look! We have turned the arrows around! ;-)
VOdka goals

However, playing with arrows is not our only goal, we also want to:

- expose the power of the VO but not its complexity;
- make users perceive that the Virtual Observatory is alive, and easily understand whether the VO is useful to them or not;
- try and pick the best features of the best existing VO data fetching tools;
- give the user a quick glimpse of what he can find inside the VO (see live examples!);
- save user’s inquiries;
- develop data mining specific tools for building datasets (see you in the KDD session, btw!).

Vodka is on her way to fully achieve all these goals, but I will show you we are working hard on it.
**Vodka Targets**

**Vodka targets**

- **VO newbies**: no apps to download in order to start, automatic updates, live examples (no SQL, ADQL or other buzzwords whatsoever);
- **VO frequent flyers**: many datasets (maybe inquiries) to manage, keeping up with new data;
- **VO developers**: SOAP webservice interface and client API;
- **Data miners**: multi-λ cross matching, multi-λ BoK extraction.

Omar Laurino (INAF - OATS)
Inquiries and Snapshots

Inquiries

An inquiry is defined by its searching criteria. It may carry only resources (Registry Inquiry) or also data (Data Inquiry)

Welcome Laurino

New Inquiry
List Inquiries
Examples
Logout

General Info
Name abell255
Description Optical+UV catalogues
Rate Daily

Capabilities and Wavebands
Select one or more capabilities
- Catalogues
- Images
- Spectra
- All
Select wavebands
- Optical
- Infrared
- UV
- X-ray
- Radio
- Millimeter
- Gamma-Ray
- EUV
- All

Keywords
Title
Description
Publisher
Subject

Coordinates
Source Name (e.g. M51) abell255
RA 258.208
DEC 64.053
Radius

Submit
An inquiry consists of several snapshots;  
A snapshot consists of several resources;  
If it is a data inquiry each resource will have its own file.

<table>
<thead>
<tr>
<th>Actions</th>
<th>Name</th>
<th>Description</th>
<th>First Snapshot</th>
<th># Snap’s</th>
<th>Next Snapshot</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>m51 x images</td>
<td></td>
<td>13/05/10 10.36</td>
<td>2</td>
<td>15/05/10 10.36</td>
</tr>
<tr>
<td>✓</td>
<td>abell2255</td>
<td></td>
<td>13/05/10 12.08</td>
<td>4</td>
<td>17/05/10 12.09</td>
</tr>
<tr>
<td>✓</td>
<td>m31 globclus</td>
<td></td>
<td>13/05/10 12.14</td>
<td>4</td>
<td>17/05/10 12.14</td>
</tr>
</tbody>
</table>
An inquiry consists of several snapshots; 
A snapshot consists of several resources; 
If it is a data inquiry each resource will have its own file.

<table>
<thead>
<tr>
<th>Actions</th>
<th>Name</th>
<th>Description</th>
<th>First Snapshot</th>
<th># Snap’s</th>
<th>Next Snapshot</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>m51 x images</td>
<td></td>
<td>13/05/10 10.36</td>
<td>2</td>
<td>15/05/10 10.36</td>
</tr>
<tr>
<td>✔️</td>
<td>abell2255</td>
<td></td>
<td>13/05/10 12.08</td>
<td>4</td>
<td>17/05/10 12.09</td>
</tr>
<tr>
<td>✔️</td>
<td>m31 globclus</td>
<td></td>
<td>13/05/10 12.14</td>
<td>4</td>
<td>17/05/10 12.14</td>
</tr>
</tbody>
</table>

List of Snapshots

<table>
<thead>
<tr>
<th>Actions</th>
<th>Creation Date</th>
<th># resources</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Thu May 13 12:20:12 CEST 2010</td>
<td>2</td>
<td>finished</td>
</tr>
<tr>
<td>✔️</td>
<td>Fri May 14 12:21:39 CEST 2010</td>
<td>3</td>
<td>finished</td>
</tr>
</tbody>
</table>
An inquiry consists of several snapshots;
A snapshot consists of several resources;
If it is a data inquiry each resource will have its own file.

<table>
<thead>
<tr>
<th>Actions</th>
<th>Name</th>
<th>Description</th>
<th>First Snapshot</th>
<th># Snap’s</th>
<th>Next Snapshot</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>m51 x images</td>
<td></td>
<td>13/05/10 10.36</td>
<td>2</td>
<td>15/05/10 10.36</td>
</tr>
<tr>
<td></td>
<td>abell2255</td>
<td></td>
<td>13/05/10 12.08</td>
<td>4</td>
<td>17/05/10 12.09</td>
</tr>
<tr>
<td></td>
<td>m31 globclust</td>
<td></td>
<td>13/05/10 12.14</td>
<td>4</td>
<td>17/05/10 12.14</td>
</tr>
</tbody>
</table>

List of Snapshots

<table>
<thead>
<tr>
<th>Actions</th>
<th>Creation Date</th>
<th># resources</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Thu May 13 12:20:12 CEST 2010</td>
<td>2</td>
<td>finished</td>
</tr>
<tr>
<td></td>
<td>Fri May 14 12:21:39 CEST 2010</td>
<td>3</td>
<td>finished</td>
</tr>
</tbody>
</table>

List of resources, i.e. VO services

<table>
<thead>
<tr>
<th>Actions</th>
<th>Title</th>
<th>Publisher</th>
<th>Capability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Astrophysics Data System</td>
<td>NASA Astrophysics Data System</td>
<td>ConeSearch</td>
</tr>
<tr>
<td></td>
<td>XMM SUSS</td>
<td>XMM at MSSL</td>
<td>ConeSearch</td>
</tr>
<tr>
<td></td>
<td>XMM-Newton Optical Monitor Serendipitous UV Source Survey Catalog</td>
<td>NASA/GSFC HEASARC</td>
<td>ConeSearch</td>
</tr>
<tr>
<td></td>
<td>The NASA/IPAC Extragalactic Database</td>
<td>The NASA/IPAC Extragalactic Database</td>
<td>ConeSearch</td>
</tr>
</tbody>
</table>
## Snapshot Differences

### New Services

<table>
<thead>
<tr>
<th>Actions</th>
<th>Title</th>
<th>Publisher</th>
<th>Capability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cosmic Lens All-Sky Survey</td>
<td>The National Radio Astronomy Observatory (NRAO)</td>
<td>SimpleImageAccess</td>
</tr>
<tr>
<td></td>
<td>NRAO VLA Sky Survey at 1.4 GHz</td>
<td>NRAO</td>
<td>SimpleImageAccess</td>
</tr>
</tbody>
</table>

### Old Services

<< < 1 2 3 4 >>

<table>
<thead>
<tr>
<th>Actions</th>
<th>Title</th>
<th>Publisher</th>
<th>Capability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MAST Image Scrapbook</td>
<td>MAST</td>
<td>SimpleImageAccess</td>
</tr>
<tr>
<td></td>
<td>NRAO VLBA Calibrator Source Survey</td>
<td>NRAO</td>
<td>SimpleImageAccess</td>
</tr>
<tr>
<td></td>
<td>2nd Digitized Sky Survey (Blue)</td>
<td>NASA/GSFC HEASARC</td>
<td>SimpleImageAccess</td>
</tr>
<tr>
<td></td>
<td>First Digitized Sky Survey: Red Plates</td>
<td>NASA/GSFC HEASARC</td>
<td>SimpleImageAccess</td>
</tr>
</tbody>
</table>

### Difference Files

<table>
<thead>
<tr>
<th>Reference ID</th>
<th>New ID</th>
<th>#Old</th>
<th>#New</th>
<th>#Missing</th>
<th>OldFile</th>
<th>NewFile</th>
<th>MissingFile</th>
</tr>
</thead>
<tbody>
<tr>
<td>572</td>
<td>584</td>
<td>1903</td>
<td>0</td>
<td>20</td>
<td>Download</td>
<td>Download</td>
<td>Download</td>
</tr>
</tbody>
</table>
What users can do

- Set up inquiries and decide the updating rate;
- receive updates directly to their mailbox;
- view inquiry details, i.e. the criteria and the list of snapshots;
- view snapshot details, i.e. the list of resources of a specific snapshot;
- view the history of incremental time differences between snapshots, both in terms of resources and data;
- download a single votable for the entire snapshot;
- download a single votable for each resource in a snapshot, as it appeared when the snapshot was taken.
- download incremental files (new data, old data, missing data);
Future improvements

- Smart form validation and new searching criteria (but before doing that we need to design the setup form so to be powerful and yet user friendly);
- data mining specific tools for, e.g., BoK extraction;
- specific clients for most active services (e.g. simbad, ned, ads);
- add more capabilities (e.g. TAP);
- finalize SOAP web service and client API packages (Java, Python);
- SAMPify (maybe through the WebSampConnector);
- integration with VOSpace;
- plugins?
Try it!

Where you can find Vodka

At the moment you can reach Vodka at this temporary location: http://spock.oats.inaf.it:8080 By signing up you will also be able to receive news about Vodka.

Thank you!