

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



VO data fetching generic use case

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)



Vodka



VO data fetching generic use case

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.



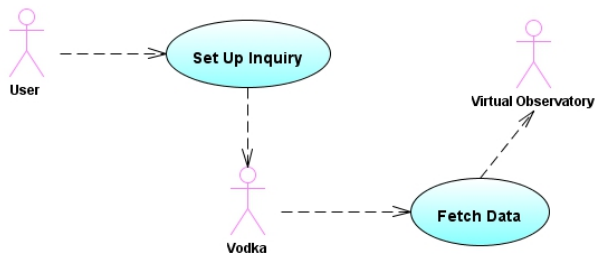
What if we could turn the arrows around?



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.



Vodka



Here comes the new gal in town

With Vodka users can be kept updated asynchronously and automatically.

Vodka



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



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

- 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.

Vodka



Inquiries

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

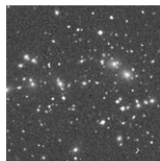
Welcome laurino

[New Inquiry](#)

[List Inquiries](#)

[Examples](#)











[Logout](#)



General Info									
Name	<input type="text" value="abell2255"/>	Description	<input type="text" value="Optical+UV catalogues"/>	Rate	<input type="text" value="Daily"/>				
Capabilities and Wavebands									
Select one or more capabilities	<input checked="" type="checkbox"/> Catalogues	<input type="checkbox"/> Images	<input type="checkbox"/> Spectra	<input type="checkbox"/> All					
Select wavebands	<input checked="" type="checkbox"/> Optical	<input type="checkbox"/> Infrared	<input checked="" type="checkbox"/> UV	<input type="checkbox"/> X-ray	<input type="checkbox"/> Radio	<input type="checkbox"/> Millimeter	<input type="checkbox"/> Gamma-Ray	<input type="checkbox"/> EUV	<input type="checkbox"/> All
Keywords									
Title	<input type="text"/>	Description	<input type="text"/>	Publisher	<input type="text"/>	Subject	<input type="text"/>		
Coordinates									
Source Name (e.g. M51)	<input type="text" value="abell2255"/>	Resolve							
RA	<input type="text" value="258.208"/>	DEC	<input type="text" value="64.053"/>	Radius	<input type="text"/>				
<input type="button" value="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.

Actions	Name	Description	First Snapshot	# Snap's	Next Snapshot
   	m51 x images		13/05/10 10.36	2	15/05/10 10.36
<input checked="" type="checkbox"/>   	abell2255		13/05/10 12.08	4	17/05/10 12.09
<input checked="" type="checkbox"/>   	m31 globclust		13/05/10 12.14	4	17/05/10 12.14

Vodka



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.

Actions	Name	Description	First Snapshot	# Snap's	Next Snapshot
	m51 x images		13/05/10 10.36	2	15/05/10 10.36
<input checked="" type="checkbox"/>	abell2255		13/05/10 12.08	4	17/05/10 12.09
<input checked="" type="checkbox"/>	m31 globclust		13/05/10 12.14	4	17/05/10 12.14

List of Snapshots

Actions	Creation Date	# resources	Status
	Thu May 13 12:20:12 CEST 2010	2	finished
	Fri May 14 12:21:39 CEST 2010	3	finished

Vodka



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.

Actions	Name	Description	First Snapshot	# Snap's	Next Snapshot
	m51 x images		13/05/10 10.36	2	15/05/10 10.36
<input checked="" type="checkbox"/>	abell2255		13/05/10 12.08	4	17/05/10 12.09
<input checked="" type="checkbox"/>	m31 globclust		13/05/10 12.14	4	17/05/10 12.14

List of Snapshots





Actions	Creation Date	# resources	Status
	Thu May 13 12:20:12 CEST 2010	2	finished
	Fri May 14 12:21:39 CEST 2010	3	finished

List of resources, i.e. VO services

Actions	Title	Publisher	Capability
	Astrophysics Data System	NASA Astrophysics Data System	ConeSearch
	XMM SUSX	XMM at MSSL	ConeSearch
	XMM-Newton Optical Monitor Serendipitous UV Source Survey Catalog	NASA/GSFC HEASARC	ConeSearch
	The NASA/IPAC Extragalactic Database	The NASA/IPAC Extragalactic Database	ConeSearch

Snapshot Differences

New Services

Actions	Title	Publisher	Capability
 	Cosmic Lens All-Sky Survey	The National Radio Astronomy Observatory (NRAO)	SimpleImageAccess
 	NRAO VLA Sky Survey at 1.4 GHz	NRAO	SimpleImageAccess

Old Services

<<< 1 2 3 4 >>>

Actions	Title	Publisher	Capability
 	MAST Image Scrapbook	MAST	SimpleImageAccess
	NRAO VLBA Calibrator Source Survey	NRAO	SimpleImageAccess
 	2nd Digitized Sky Survey (Blue)	NASA/GSFC HEASARC	SimpleImageAccess
 	First Digitized Sky Survey: Red Plates	NASA/GSFC HEASARC	SimpleImageAccess

Difference Files

Reference ID	New ID	#Old	#New	#Missing	OldFile	NewFile	MissingFile
572	584	1903	0	20	Download	Download	Download

Vodka



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);



- 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?



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.



Vodka

