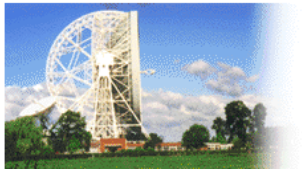
Workflows in Astrogrid with Taverna

Kevin Benson

kmb@mssl.ucl.ac.uk

(contact via e-mail if you wish to install plugins before official release)



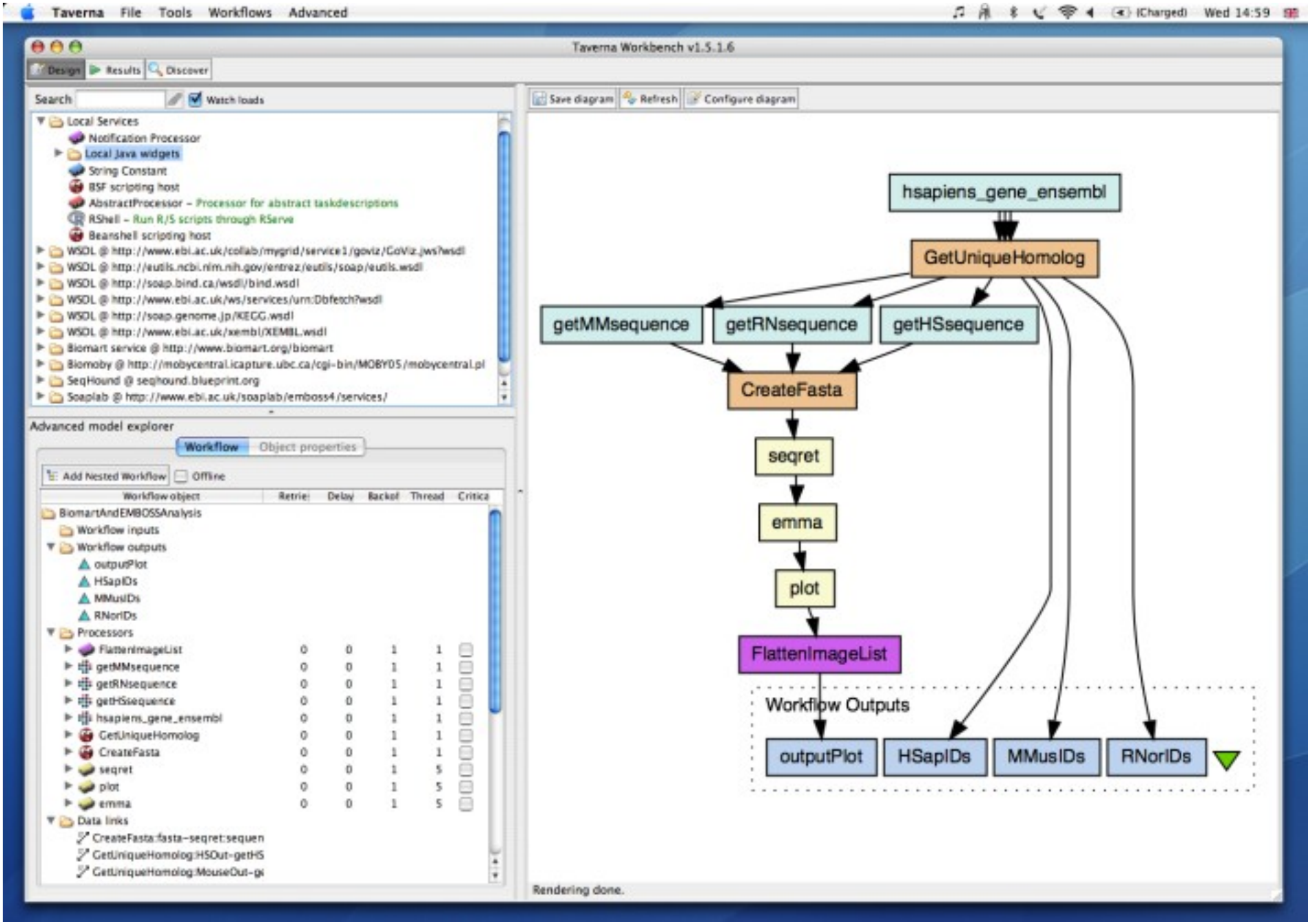
Jodrell Bank Observatory

Talk Outline

- Taverna
- AstroRuntime
- Demo

Taverna

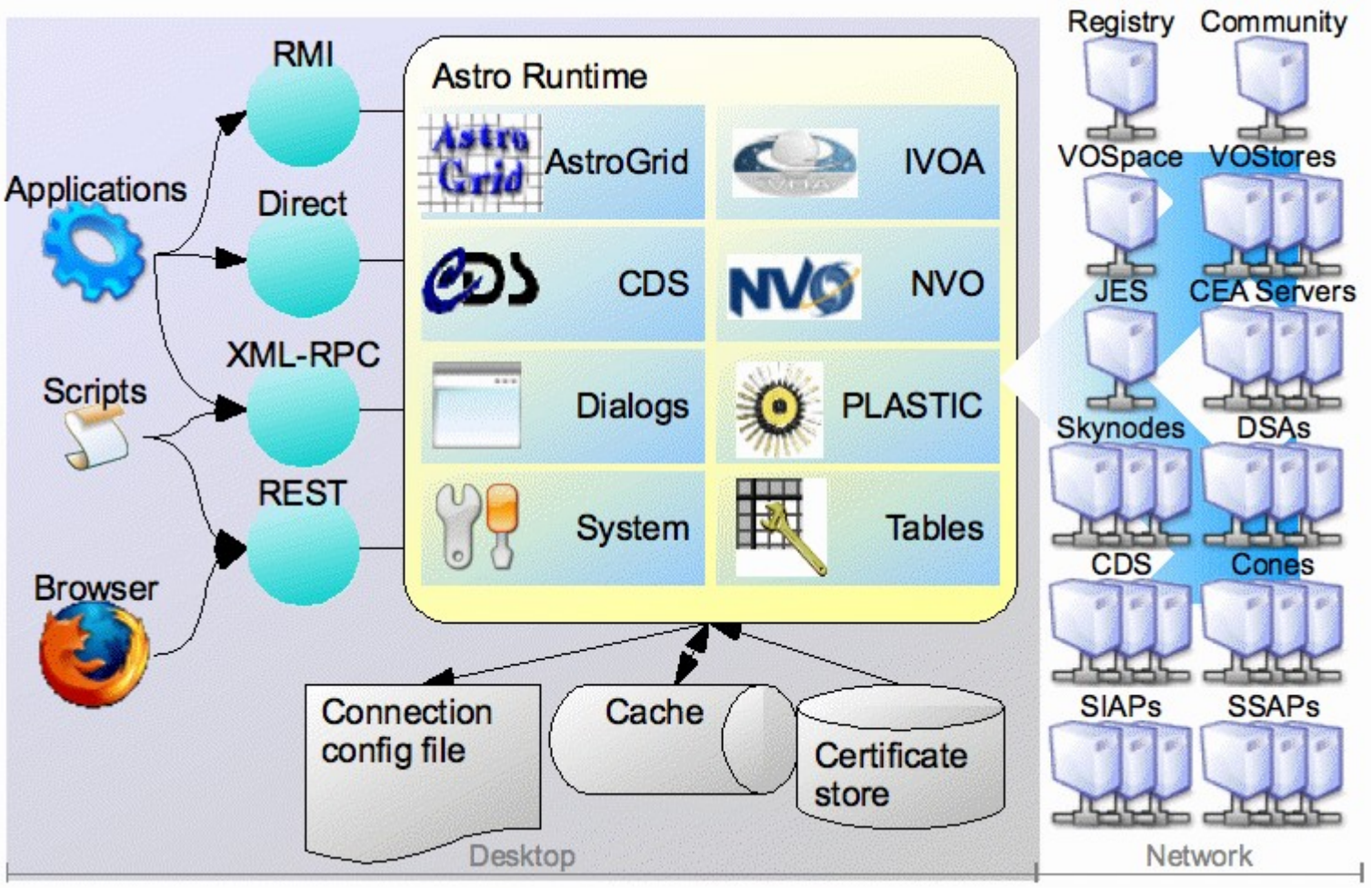
- The Taverna workbench is a free software tool for designing and executing workflows, created by the myGrid project.
- Allows a biologist or bioinformatician with limited computing background and limited technical resources and support to construct highly complex analyses over public and private data and computational resources, all from a standard PC, UNIX box or Apple computer.
- The myExperiment social web site supports finding and sharing of workflows.
- But with a nice plug-in architecture that Taverna has plugins are being built to connect to the VO community. ESO uses a tool called Reflex in implementing data reduction pipelines via Taverna. JVO will have a plugin connected to Taverna in the near future. AstroGrid have developed a plug-in adaptor for Taverna, thereby exposing the standard range of IVOA standard services, thus those for images (SIAP), spectra (SSAP), tables (TAP), and applications (CEA-Common Execution Architecture) to Taverna.
- http://taverna.sourceforge.net/
- http://www.myexperiment.org/


Taverna File Tools Workflows Advanced
Taverna Workbench v1.5.1.6
Design
Results
Discover
Search
Watch loads
Local Services
Notification Processor
Local Java widgets
String Constant
BSF scripting host
AbstractProcessor – Processor for abstract taskdescriptions
RShell – Run R/S scripts through RServe
Beanshell scripting host
Advanced model explorer
Workflow
Object properties
Add Nested Workflow
Offline
BiomartAndEMBOSSAnalysis
Workflow inputs
Workflow outputs
outputPlot
HSapIDs
MMusIDs
RNorIDs
Processors
FlattenImageList
getMMsequence
getRNsequence
getHSsequence
hsapiens_gene_ensembl
GetUniqueHomolog
CreateFasta
seqret
plot
emma
Data links
Save diagram
Refresh
Configure diagram
hsapiens_gene_ensembl
GetUniqueHomolog
getMMsequence
getRNsequence
getHSsequence
CreateFasta
seqret
emma
plot
FlattenImageList
Workflow Outputs
outputPlot
HSapIDs
MMusIDs
RNorIDs
Rendering done.

Astro Runtime

- A client-side library of functions to access VO
  - scope: integrate all VO standards, popular ad-hoc services, suitable helper functions
- Library is exposed as a desktop service
  - accessible from almost all programming languages
    - XMLRPC, HTTP, RMI, C-binding
- API designed to use consistent abstractions and types
  - shallow learning curve
  - Insulates from change.

Astro Runtime - Schematic



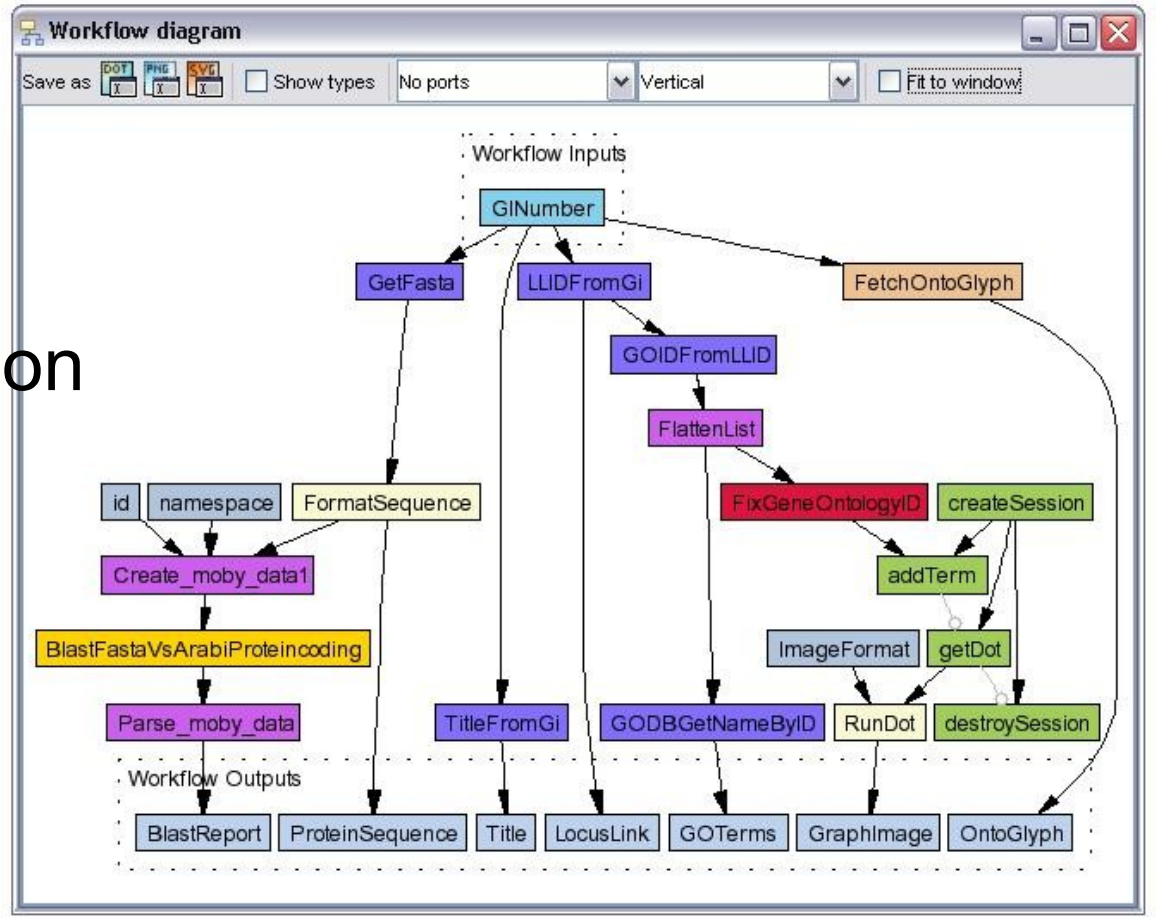
http://www2.astrogrid.org/desktop/astro-runtime
http://www.astrogrid.org

Client-side Science Scripting

- Command-line access to VO
  - through Python scripts (or perl, or other language) calling AstroRuntime routines.
  - astronomer can integrate VO services with their existing toolset.
- Power of the for loop.
  - remote app. execution
    - sextractor script, montage mosaic tool
  - image search at multiple positions
  - cone search over multiple archives

Client-side Workflow

- Taverna workflow engine
  - Graphical interface
  - Step-through execution
- Connects to Astro Runtime to call VO services.



Demo and Links

http://www.astrogrid.org
http://www2.astrogrid.org/desktop/astro-runtime
http://taverna.sourceforge.net
http://www.myexperiment.org/