



Astronomical Data Analysis Software and Systems 2007

Workflow in Astronomy, the VO France Workflow Working Group experience

André Schaaff, Observatoire de Strasbourg

Franck Le Petit, Observatoire de Paris

Philippe Prugniel, Observatoire de Lyon

Eric Slezak, Observatoire de la Côte d'Azur

Christian Surace, Laboratoire d'Astrophysique de Marseille



ADASS 2007, London
24 September 2007

Motivation

- **Many services are developed/deployed in the frame of the Virtual Observatory (registry, data services, Web Services, computing and Grid services, ...).**
- **Complex implementation and coordination of the services are possible through workflows**
 - **Evolution from an execution of one service to a combination of services (exchanging data, ...)**
 - **Question to be answered :**
 - Localisation and verification of the services (execution time, tests, results, ...)
 - Existing tools (how to take them into account in workflows) and constraints on the future developments
 - How to integrate the VO Standards ?
 -



Motivation (2)

- **Workflows are useful to capture scientific methodology and to provide provenance information for their results**
- **Workflows provide a formalization of the scientific analysis**
 - **routines to be executed, data flow, execution details, ...**
- **Workflow are structures useful to manage computation at a large-scale**
- **Collaboratively designed, assembled, validated, analyzed**
- **...**



Quick State of the Art

■ Workflow ...

(source : <http://www.gridworkflow.org>)

■ ... languages

- AGWL, BPEL4WS, BPML, DGL, DPML, GJobDL, GSFL, GFDL, GWorkflowDL, MoML, SWFL, WSCL, WSCI, WSFL, XLANG, YAWL, SCUFL/XScufl, WPD, PIF, PSL, OWL-S, xWFL, ...

■ ... language formalisms

- Petri net, UML activity diagram, BPMN, DAG, IPO, GPSG, Workflow Patterns, Pi Calculus, Finite-State Machine, Gamma-calculus, ...

■ ...mapping from abstract to concrete workflows

- CWG, ACWG, Grid Job Handler, GWES, ...

■ Workflow ...

■ ...engines

- BioPipe, BizTalk, BPWS4J, DAGMan, GridAnt, Grid Job Handler, GRMS, GWFE, GWES, IT Innovation Enactment Engine, JIGSA, JOpera, Kepler, Karajan, OSWorkflow, Pegasus (uses DAGMan), Platform Process Manager, ScyFLOW, SDSC Matrix, SHOP2, Taverna, Triana, wftk, YAWL Engine, WebAndFlo, WFEE, ...

■ ...composition/designing tools

- ilog's BPMN Modeller, CAT, GWUI, XBaye GUI for Workflow Composition, Triana, JOpera, Platform Process Manager, ...

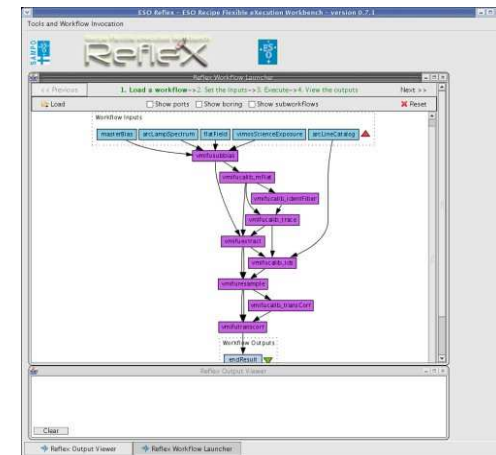


Examples of workflow builders in astronomy

The ESO Recipe Flexible Execution Workbench (based on Taverna)

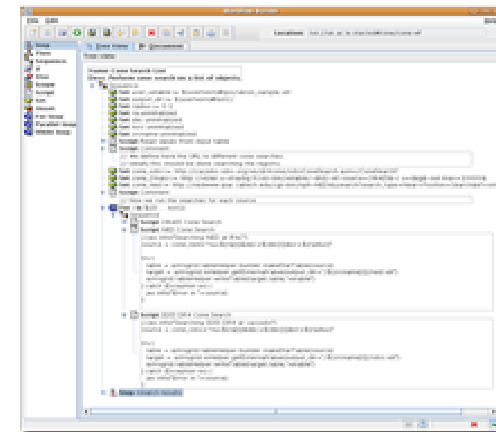
R. Hook et al.,
ESO

We had a useful discussion with R.
Hook and P. Järveläinen



www.eso.org/sampo/reflex

The AstroGrid workflow system



www2.astrogrid.org



VO France Workflow Working Group

- **In a first step it was necessary to give our own definition of a workflow :**
 - ...a sequence of tasks executed within a controlled context by an architecture taking into account VO standards

- **Main goals**
 - Definition of use cases of general interest in different domains
 - Suggestion of solutions for designing and exploiting easily such workflows
 - Identification of the simplest workflow structure allowing portability
 - Definition of elementary bricks



Use cases

Illustration : some use cases presented and discussed in the frame of the working group



Image processing use case

■ operation

- detection and evaluation of related objects in 1 band image

■ subjacent model

- diffuse disjoined tasks in emission on a bottom slowly variable without defects

■ method

- cartography of the background
- thresholding by segmentation
- adjustment of an ellipse of form
- evaluation of the azimuth profile of brightness
- calculation of measurements of form and flow

E. Slezak
*Observatoire de
la Côte d'Azur*

...small part of a workflow use case which is under construction



Use case in Simulation

1 - PDR code

modélisation of diffuse clouds
gives theoretical absorption spectra

2 - Stellar spectra databases

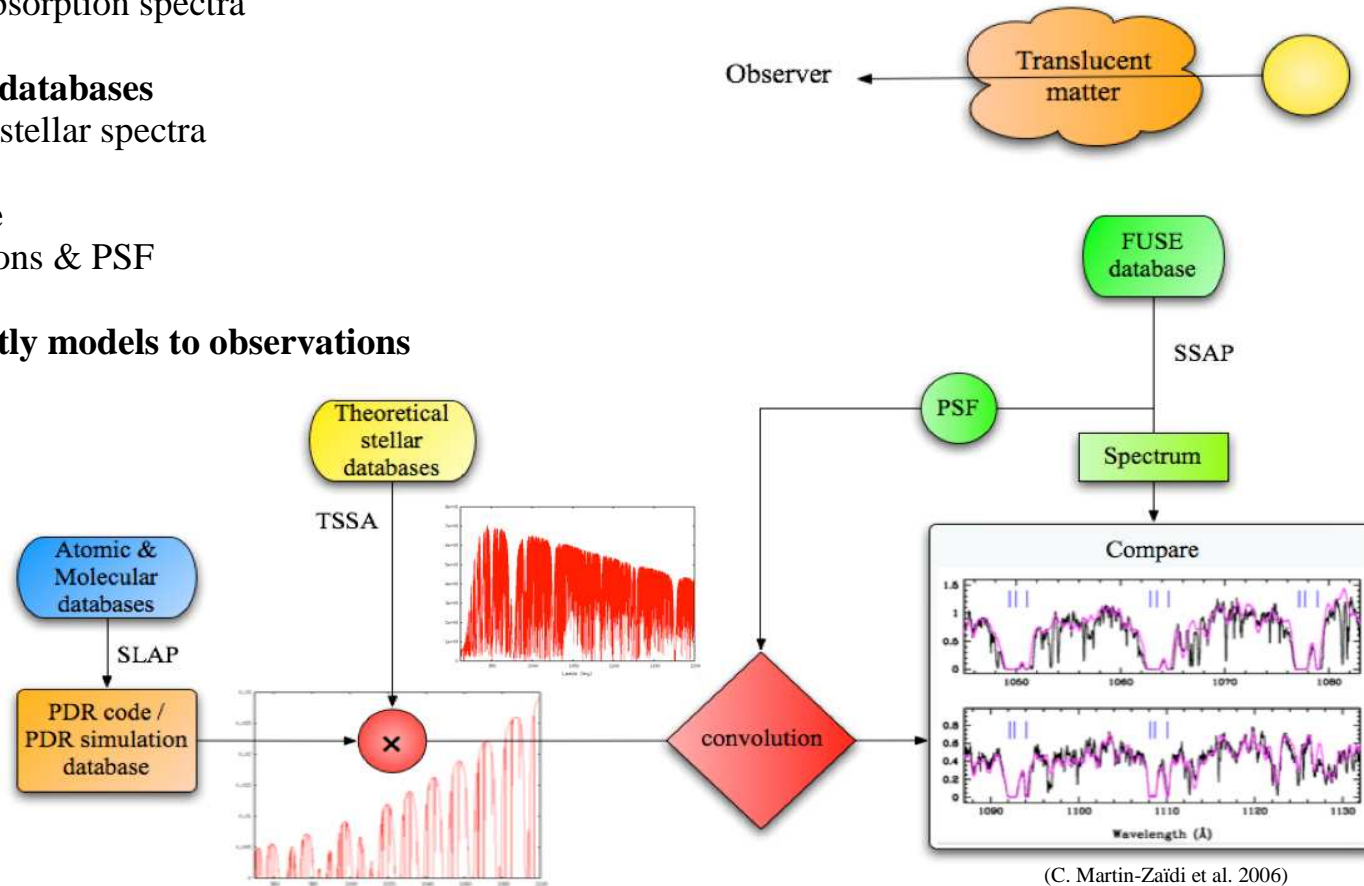
provides synthetic stellar spectra

3 - FUSE database

provides observations & PSF

4 - Compare directly models to observations

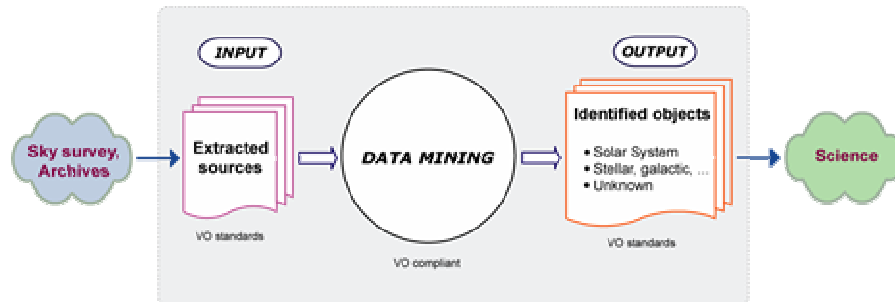
F. Le Petit et al.
*Observatoire de
Paris-Meudon*



(C. Martin-Zaïdi et al. 2006)

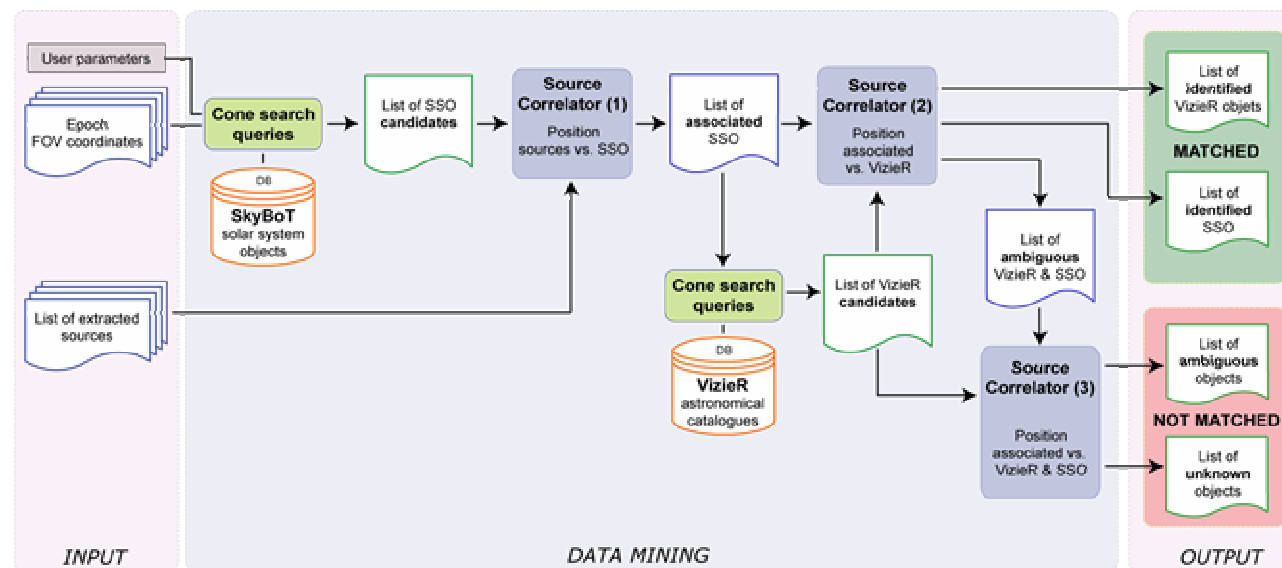


Use case in Data Mining

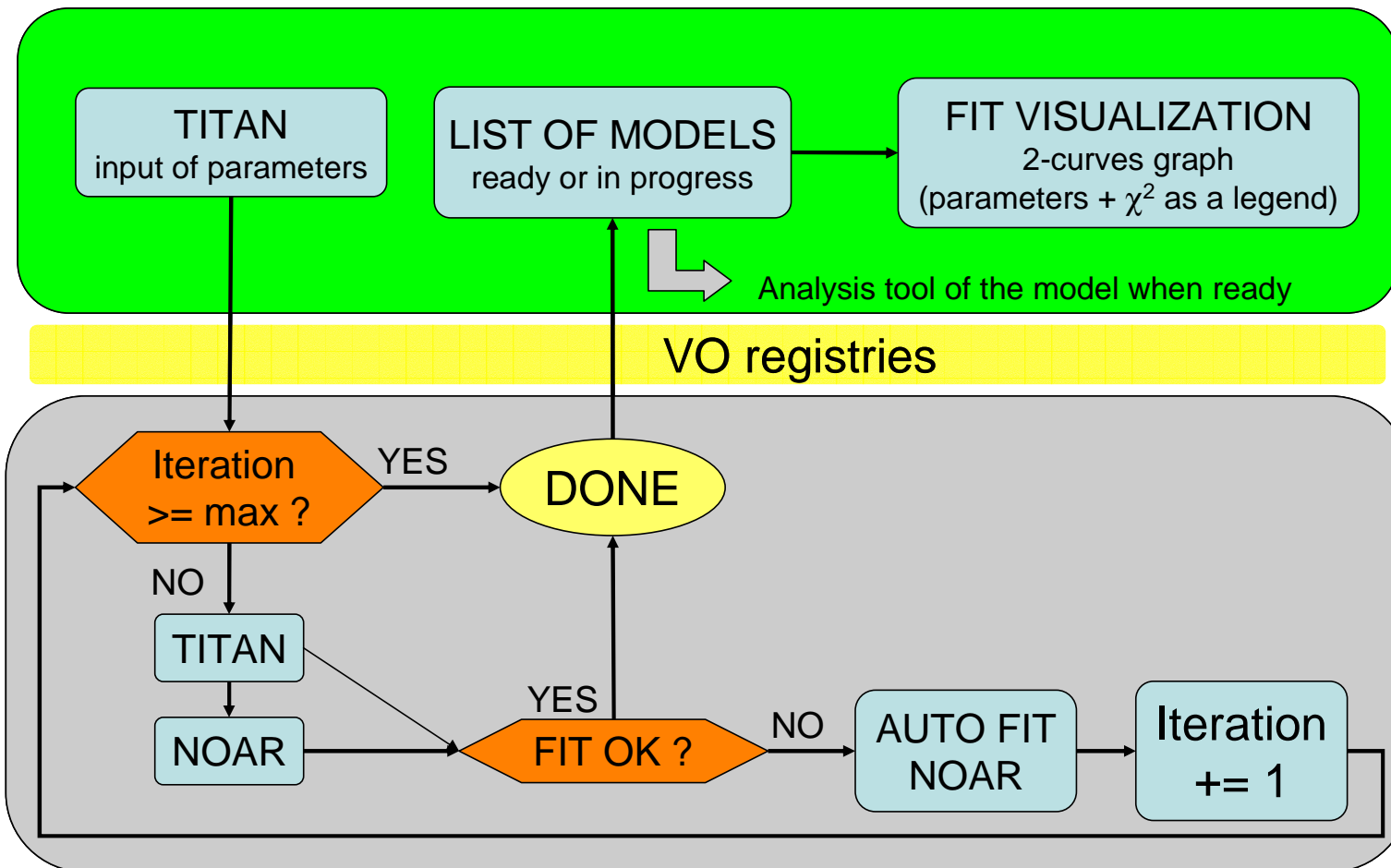


See poster P1.021

J. Berthier et al.,
IMCCE



Use case for TITAN and NOAR codes



L. Chevallier
Obs. Paris
Meudon



Use case implementation

- **The aim of the working group is not to decide which tool is better than another**
- **In a first step we decided to use a tool (AIDA – Astronomical Image processing Distribution Architecture) in which development some of us were involved**
 - **Easy to modify sources, to add data types, ...**
 - **VO standards enabling was discussed during the development**

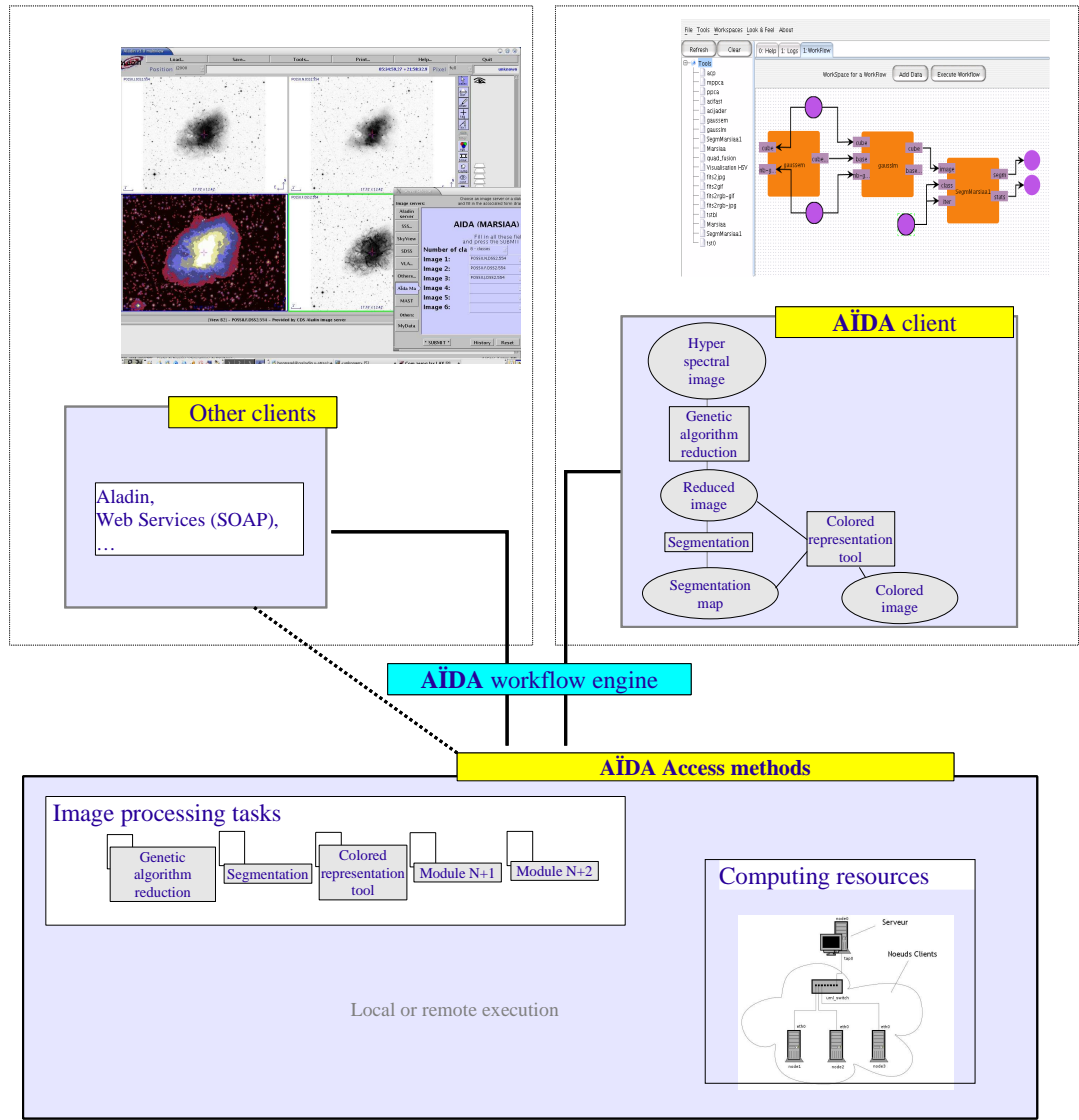


AIDA architecture

AIDA architecture
F. Bonnarel,
M. Louys,
A. Schaaff
CDS & LSIT

AIDA developments
J. Beugnot
J.-J. Claudon,
G. Mantelet,
C. Pestel,
CDS

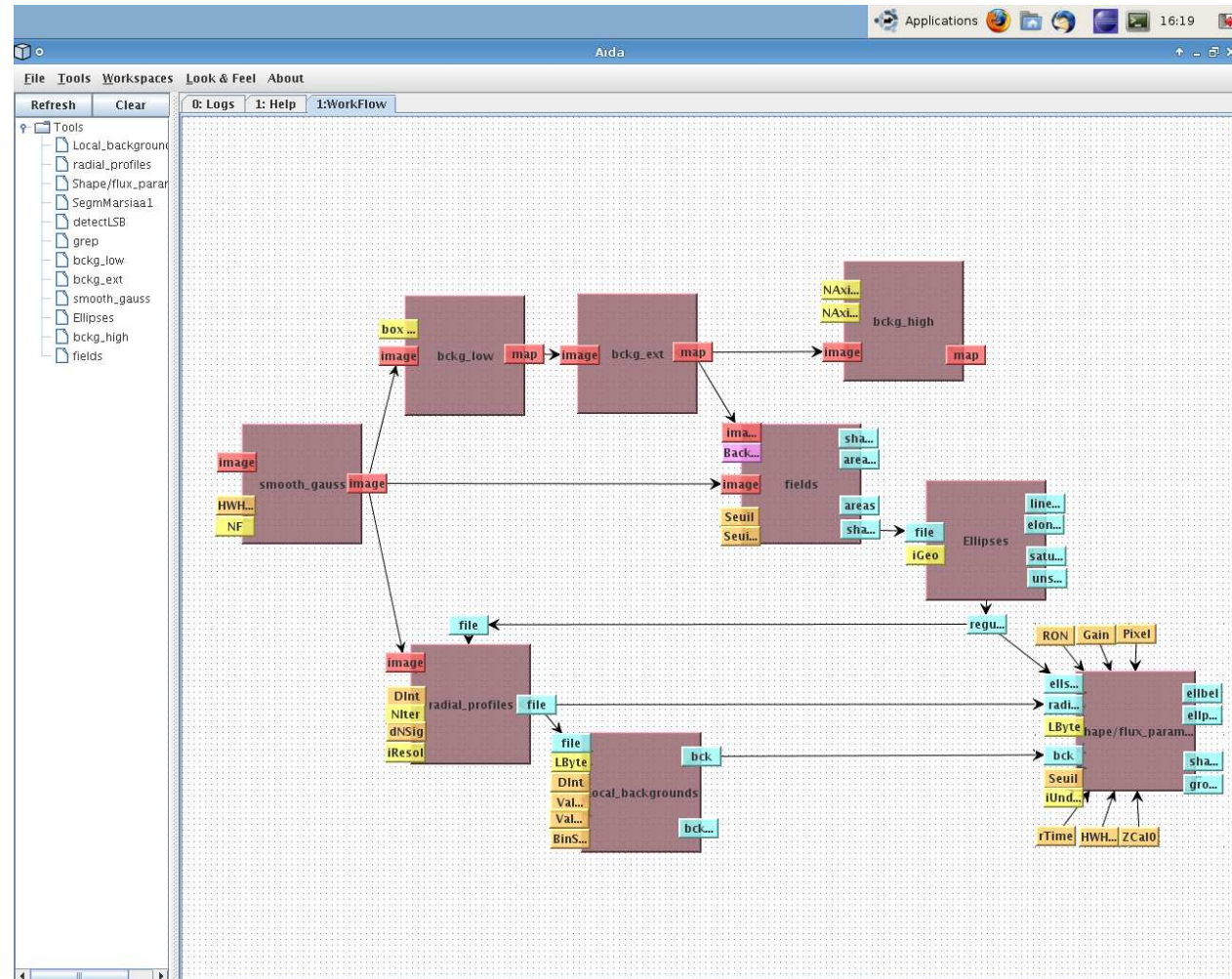
Work done in the frame of the French « Massive Data in Astronomy » project (2003-2006), OV France and Euro VOTech



Example of implementation for an image processing workflow

14

Use case
E. Slezak
Observatoire de la Côte d'Azur



Simulation use case (F. Le Petit) will be implemented next month

ADASS 2007
24 September 2007

Workflow builder plugin in Aladin

Based on AIDA
graphical libraries
and engine

C. Pestel
T. Boch
CDS



Aladin workflow builder v. 0.3 - Using JLOW library

```

WED.m51.14.0'=get NED m51 14.0'
Simbad.m51-1=get Simbad m51
XMatch-1=xmatch 2 1 4
XMatch-2=xmatch 2 3 4
sync: export "XMatch results2" toto
  
```

Aladin v4.0 *** BETA VERSION (based on v4.013) ***

Position ICRS Pixel full

m51

4 planes, 1 view, 7Mb

ADASS 2007
24 September 2007

Remarks about the work done

- **Useful to describe workflow use cases and to be able to implement them with a light and flexible workflow tool**
 - Real help for astronomers
 - Easy to reuse in “new” workflows “old” applications developed in the past by trainees, Ph D. students, engineers and astronomers in different languages
 - ...
- **Good training before using more “heavy” tools**
- **Experience sharing with people in different domains (image, spectroscopy, data mining, simulation, ...)**
- ...



Ongoing work

- **Study how to involve the Characterization IVOA standard in the workflow process**
- **Work on bricks development**
 - extraction of subparts in n dimensions hypercube
 - separation of structures in n dimensions hypercube
 - resolution adaptation for imagery and spectrum
 - multi components fit
 - image addition, spectrum addition
 - ...
- **Interoperability of languages, checking with IT community what are the finalities of the exchanges between workflows, the best approach for standardization, etc.**



VO France Workflow WG

Thanks to all the contributors and participants to the working group :

Christophe Barache – Obs. Paris-Meudon, Jérôme Berthier - IMCCE, Thomas Boch - CDS, Frédéric Boone - LERMA, François Bonnarel - CDS, Loïc Chevallier – Obs. Paris-Meudon, Jean-Julien Claudon - CDS, André Csillaghy – HES Switzerland, Bernard Debray - LAOB, Jean-Michel Desert - IAP, Pierre Didelon - CEA, Marie-Lise Dubernet – Obs. Paris-Meudon, Anabela C. Goncalves - LUTH, Jesus Iglesias - IMCCE, Gaëlle Labourot - UVSQ, Martin France - CRAL, Pierre Le Sidaner – Obs. Paris-Meudon, Mireille Louys - CDS-LSIIT, Guillaume Mella - LAOG, Areg Mickaelian – BAO Armenia, Johan Montagnat - ESSI Sophia Antipolis, Nicolas Moreau - LERMA, Steven Morin - UVSQ, Jonathan Normand – Obs. Paris-Meudon Cyril Pestel-CDS, Fabrice Roy - LUTH, Lena Sargsyan - Yerevan State University , Alain Sarkissian - Service d'Aéronomie, Renaud Savalle - Obs. Paris-Meudon, Françoise Tran Minh - LERMA Obs. Paris-Meudon, Frédéric Vachier – IMCCE, Bruno Voisin - NUI Galway, ...

<http://www.france-ov.org/twiki/bin/view/GROUPEStravail/Workflow>

Contact : schaaff@astro.u-strasbg.fr

