

# JOpera and XtremWeb-CH in the Virtual EZ-Grid

**Cesare Pautasso**

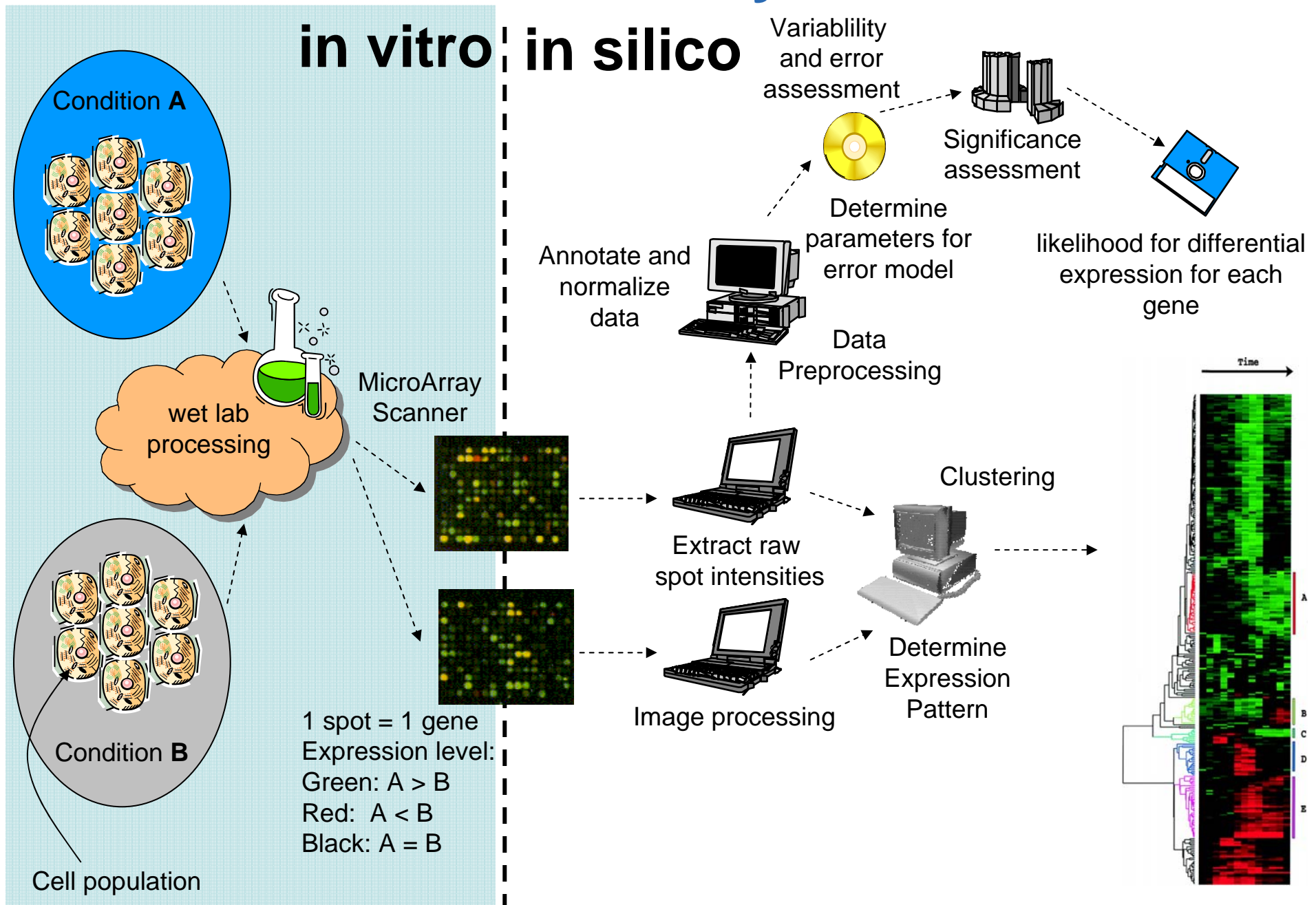
Faculty of Informatics

University of Lugano (USI), Switzerland

<http://www.pautasso.info>

# Virtual Laboratory Workflow

**in vitro** | **in silico**



# Building Virtual Laboratories

Copy & Paste  
between different Websites

NCBI  
PubMed Nucleotide Protein Genome Structure  
Search Genome for Go Clear  
Limits Index History

Caenorhabditis elegans  
Caenorhabditis briggsae

BLAST homepage  
BLAST overview  
BLAST FAQs  
BLAST news

**BLAST against Caenorhabditis elegans /  
Caenorhabditis briggsae sequences**

Compare your query sequence to DNA sequence or protein products.

Organism: Caenorhabditis elegans  
Database: DNA sequence Program: blastn MegaBlast ☒  
Enter an accession, GI, or a sequence in FASTA format:

ETH  
Computational Biochemistry Research Group  
News | People  
Research | Education | Services  
Darwin | Bio-recipes | Challenges

ETH Zurich - D-InfK - ICOS - CBRG - Server Services - AIIAI: Related peptide sequence

Server Services  
AIIAI: Related peptide  
sequence  
Sequence Specification  
ASAS: Alerting of Structurally  
Ambivalent Sequences  
Run a DARWIN program  
Evolutionary Analysis  
FastaToDarwin  
Linear Classification  
MassSearch  
PepSearch  
PepSearch  
Phylogenetic Tree  
TreeGen  
Vertex Cover

## AIIAI: Related peptide sequence

Filling and submitting the following form will start an all against all of the given derived results. Sequences must be separated by commas and can be entered in the AIIAI: Related peptide sequence form.

NOTE: If your PepSearch results give an "ID not found error" in the AIIAI: Related peptide sequence form, use our [converting tool](#) to change your FASTA sequences to Darwin SGML.

See an [example](#) to get some sequences.

The result should be mailed to:  
(Email address):  
Title of your job (optional):

Sequences:



## Scalable Vector Graphics Plasmid Map

Plasmid Name:   
Line Thickness: 0.5  
Size (in base pairs):   
Shape: ☒ Circular ☐ Linear  
Clear Input

Markers:

Start:  Finish:  ☒ Arrow on

Marker type:  Fill-color: Red Line Thickness: 12

Add to list

Enzymes:

Add enzyme in the proper order for correct positioning (I'll fix this bug in later release)

Name:  Position:

Add to list

# Building Virtual Laboratories

## Copy & Paste

between different Websites

## Programming

Java, C++, C#, Fortran...

## (Shell) Scripts

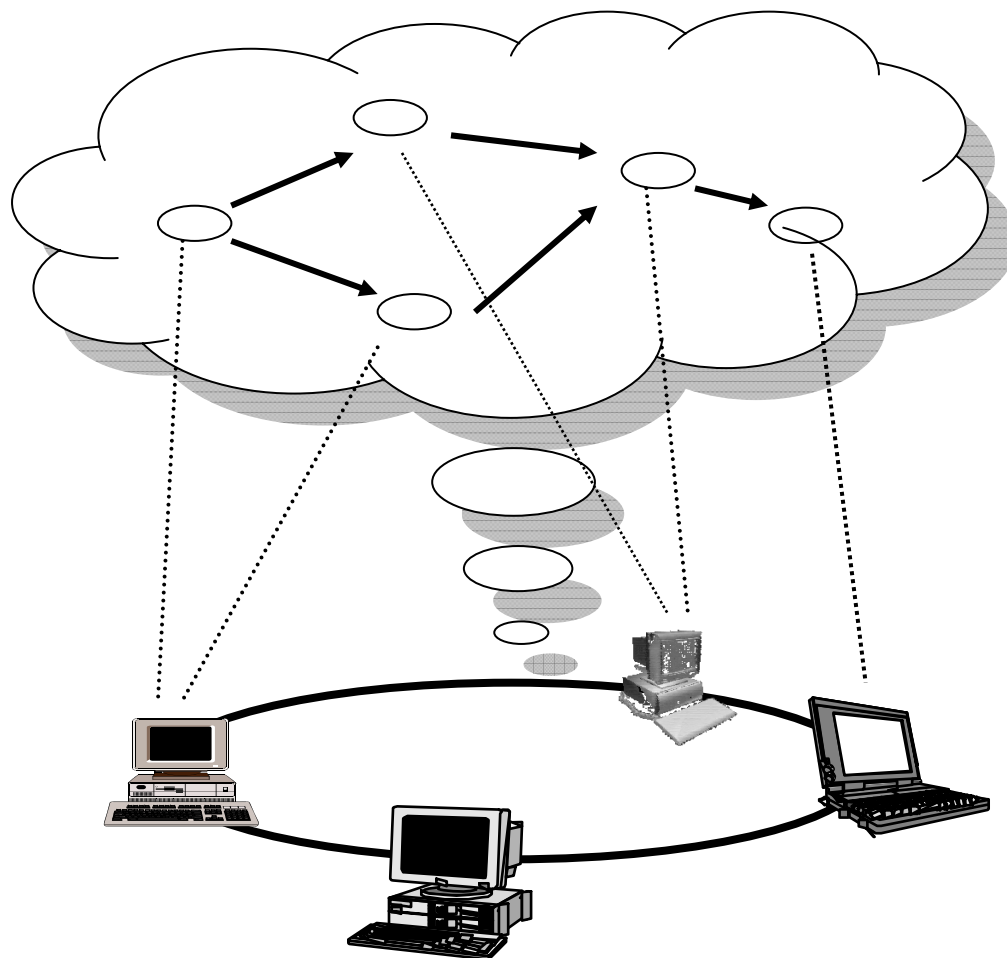
Tcsh, Bash, Makefiles,  
Python, Perl...

## Workflows

Graphical, Drag & Drop  
and Connect Environments

# Grid Workflow Management

Make it easy to build Grid applications composed of **multiple jobs**



“Provide the scientist with a platform that takes care of all data handling and record keeping chores so that the user can concentrate on the *science* and not *computer science*”

# JOpera: Grid Workflow for Eclipse

JOpera Monitor – two\_step (0) – Eclipse SDK

File Edit View Navigate Search Project Run FieldAssist Window Help

Outline

- Version: 1.0
  - Parameters (12)
  - Tasks (13)
    - AcList2Trie
    - Decrement
    - Insilico spectro
    - Inspect
    - Mascot
    - MascotImport
    - Phenylx
    - Pidres2AC
    - Pidres2UnmatchedSp
    - Pidres2UnmatchedSp
    - Popitam
    - UpdateJoperaArgs
    - Xtandem
  - Views (7)
    - ControlFlow
    - DataFlow
    - DataFlow-CONFIGXML

two\_step (0)

Kernel Console

```

Unreachable workflow_v17_HeapSpaceChange.two_step.Mascot.0
Running workflow_v17_HeapSpaceChange.two_step.Xtandem.0
Finished workflow_v17_HeapSpaceChange.two_step.Xtandem.0
Running workflow_v17_HeapSpaceChange.two_step.Pidres2UnmatchedSpectra.0
Running workflow_v17_HeapSpaceChange.two_step.Pidres2AC.0
Finished workflow_v17_HeapSpaceChange.two_step.Pidres2AC.0
Unreachable workflow_v17_HeapSpaceChange.two_step.AcList2Trie.0
Unreachable workflow_v17_HeapSpaceChange.two_step.Inspect.0
Failed workflow_v17_HeapSpaceChange.two_step.Pidres2UnmatchedSpectra.0
Unreachable workflow_v17_HeapSpaceChange.two_step.Popitam.0
Unreachable workflow_v17_HeapSpaceChange.two_step.Pidres2UnmatchedSpectra1.0
Failed workflow_v17_HeapSpaceChange.two_step.0.0
Finished workflow_v17_HeapSpaceChange.two_step.Decrement.0
Finished workflow_v17_HeapSpaceChange.two_step.UpdateJoperaArgs.0
  
```

Instance ... Parameter ...

Search:

two\_step [1.0] (1/1)

Properties

Property	Value
pidresXtandem	/var/phenylx/results/3
SWISSPIT_CO	/home/swisspit/swiss
SWISSPIT_HO	/home/swisspit/swiss
System Input Par	
AM	NewAdapter
command	"\$JAVA6/bin/java" -C
System Output Par	
Output	length unmatchedPe PRINTLN: Comman [/bin/bash, -icfile, -/. Started execute() unmatchedPealList]
realtime	36370
retval	1
stderr	log4j:WARN No app log4j:WARN Please Pidres2UnmatchedP
stdout	length unmatchedPe PRINTLN: Comman [/bin/bash, -icfile, -/. Started execute() unmatchedPealList]

# Drag, Drop and Connect

JOpera Design - demo.oml - Eclipse SDK

125%

JOpera Mo...

**JOpera Navigator**

- demo
  - ws
    - demo.oml

**Outline**

- Processes (1)
  - Convert
    - Parameters (5)
      - country1 (xsd:string)
      - country (xsd:string)
      - symbol (xsd:string)
      - ConvertedPrice (Strin
      - OriginalPrice (xsd:flo
    - Tasks (3)
      - getQuote
      - getRate
      - Multiply
    - Views (2)
    - Programs (1)

**Overview**

Overview | Process: Convert | ControlFlow | DataFlow

**Problems** | **Properties**

0 errors, 2 warnings, 0 infos (Filter matched 2 of 23)

Description	Resource	In Fol
Outbox-Parameter "Convert	demo.or	de
Process-Output-Parameter '	demo.or	de

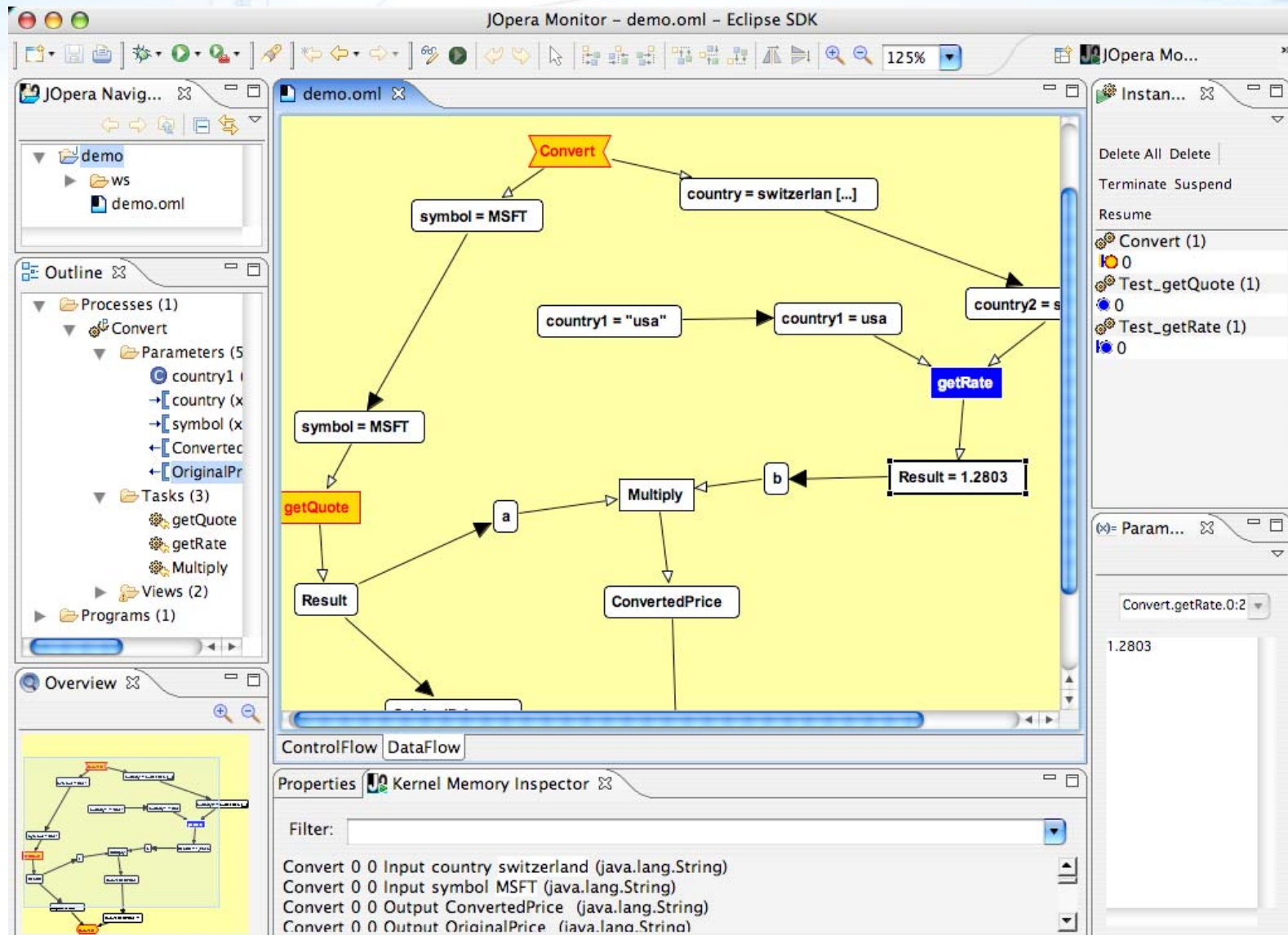
**JOpera Library**

Search:

- Convert
- NewProgram
- ws.services.xmethods.net.soap.urnxmethodsdel...
- Test\_getQuote
- getQuote

The main diagram shows a process flow for 'Convert'. It starts with a 'symbol' parameter, which is used in a 'getQuote' task. The result of 'getQuote' is a 'Result' object. This 'Result' is then used in a 'Multiply' task along with a 'country1' parameter (set to 'usa') and a 'country2' parameter. The 'Multiply' task produces a 'ConvertedPrice' output. The 'country1' parameter is also used in a 'getRate' task, which produces a 'Result' object. This 'Result' is then used in a 'b' task, which produces a 'Result' object. The 'b' task is connected to the 'Multiply' task. The 'country2' parameter is also used in a 'getRate' task, which produces a 'Result' object. The 'getRate' task is connected to the 'b' task. The 'country1' parameter is also used in a 'getRate' task, which produces a 'Result' object. The 'getRate' task is connected to the 'b' task. The 'country2' parameter is also used in a 'getRate' task, which produces a 'Result' object. The 'getRate' task is connected to the 'b' task.

# Run, Monitor, Steer and Debug



# A Growing User Community

Swiss Institute of Bioinformatics, Swiss National Supercomputing Center, Purdue University, National Center for Supercomputing Applications, Singapore Management University, McGill University, University of Kent, SINTEF (Norway), IWU Fraunhofer (Germany), University of Nebraska...



# Why Users Like JOpera

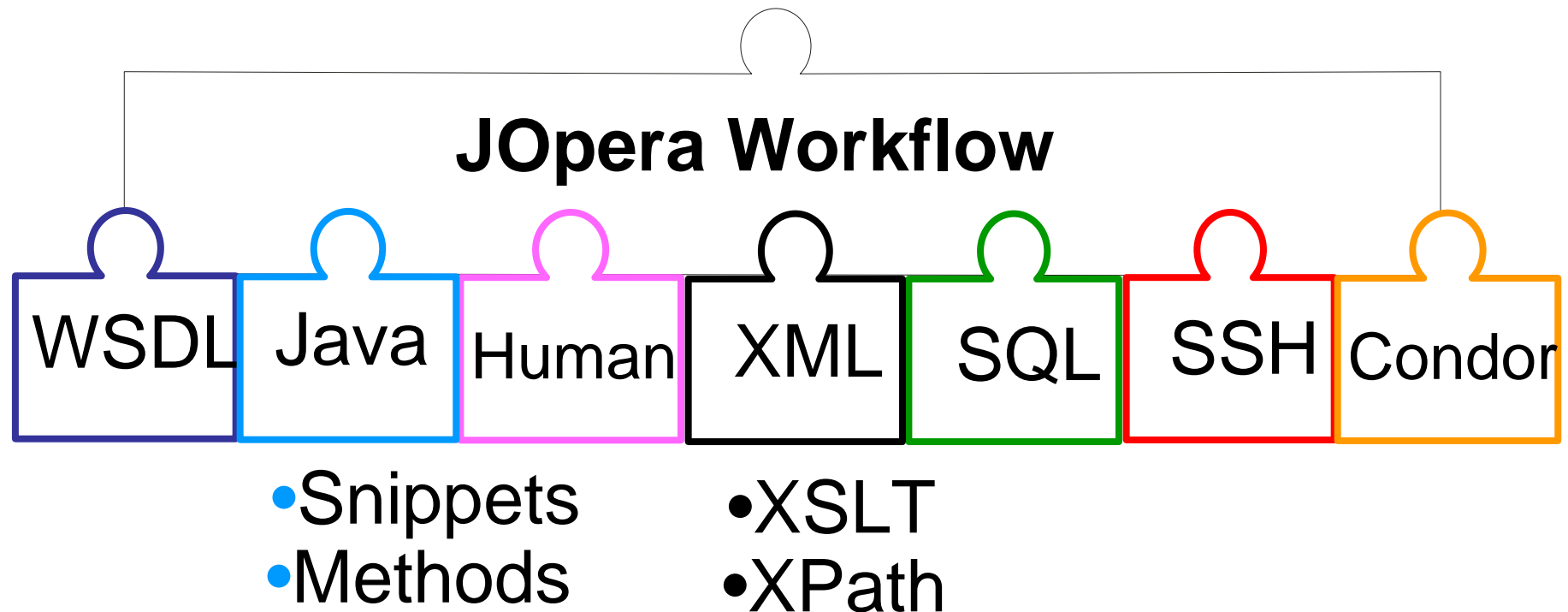
- **High Level Workflow Language**
  - Data and Control Aspects (Visual Representation)
  - Recursion, Iteration, Parallelism and Pipelining
- **Open and Extensible Component Model**
  - Run existing code without changes
  - Synchronous, Asynchronous, and Streaming interaction
  - Web services support (Axis, WSIF)
  - Secure access to remote file systems and hosts (SSH)
  - Easy to integrate with existing schedulers (e.g. Condor)

# Why Users Like JOpera

- **High Level Workflow Language**
  - Data and Control Aspects (Visual Representation)
  - Recursion, Iteration, Parallelism and Pipelining
- **Open and Extensible Component Model**
  - Run existing code without changes
  - Synchronous, Asynchronous, and Streaming interaction
  - Web services support (Axis, WSIF)
  - Secure access to remote file systems and hosts (SSH)
  - Easy to integrate with existing schedulers (e.g. Condor)
- **Strong Eclipse Foundation**
  - Platform Independent (Eclipse/Java)
  - Flexible, Extensible, Modular and Embeddable

## Extensible Component Model

- Combine in the same workflow jobs implemented using an open and extensible set of technologies



## Sharing Workflows as a Service

- JOpera processes are automatically published to clients using a variety of access protocols

Web Clients

WS Clients

Eclipse RCP  
Clients

REST

WSDL

Java

**JOpera Workflow**

WSDL

Java

Human

XML

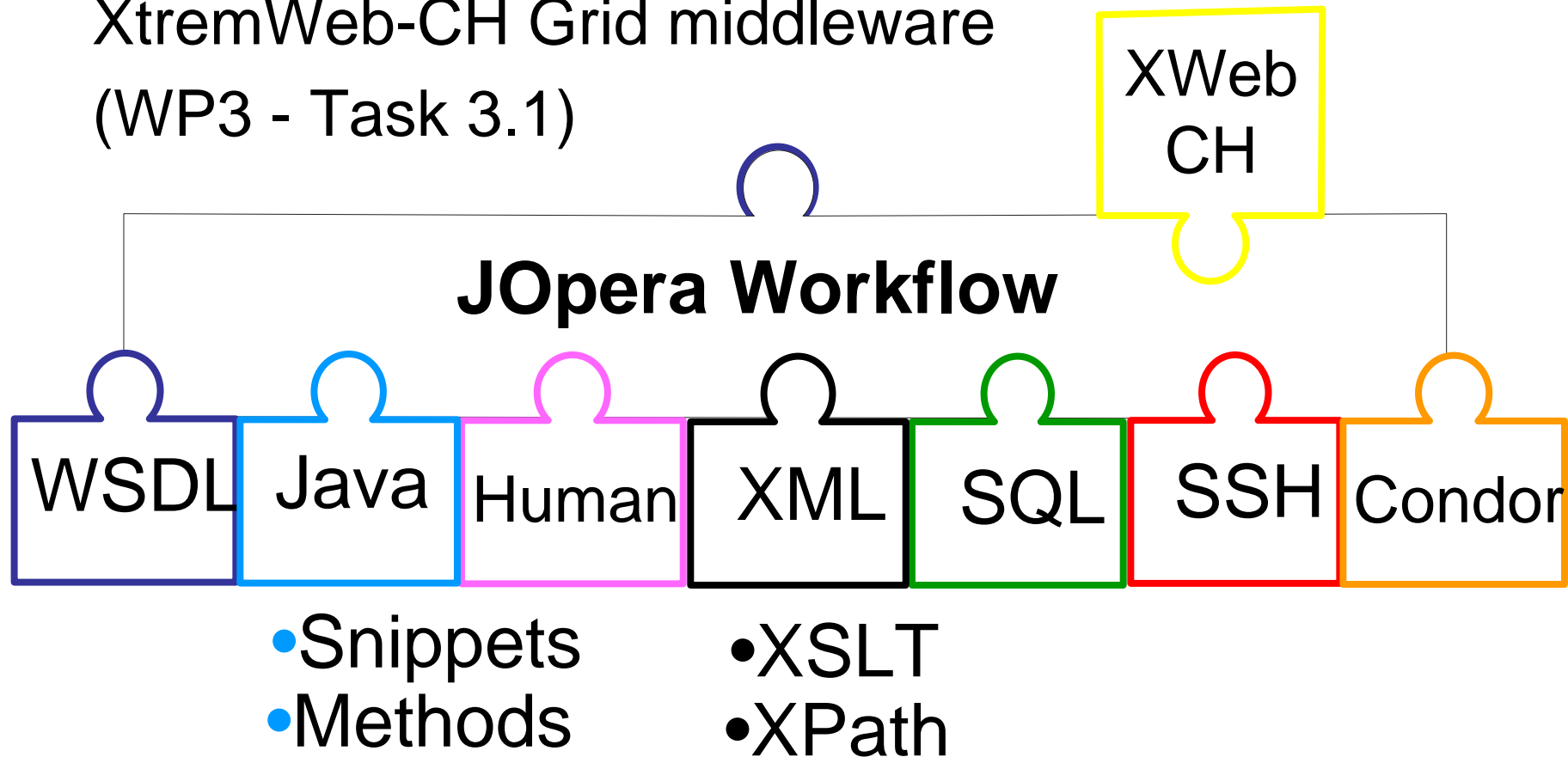
SQL

SSH

Condor

## JOpera in the Virtual EZ-Grid Project

- Develop a new adapter for the submission of jobs to XtremWeb-CH Grid middleware (WP3 - Task 3.1)



## JOpera and XtremWeb-CH

1. Definition of job submission parameters based on the API of XtremWeb-CH
2. Development of a JOpera adapter plugin to connect the two systems:
  - Start the job by submitting it to the Grid middleware
  - Monitor the job execution
  - Retrieve the job results (and its success/failure outcome)
3. Evaluate and test the integrated system on the Virtual EZ-Grid infrastructure

# Roadmap 2009

- **Swiss National Grid Deployment**
  - ARC/Nordugrid Job Submission
  - Virtual EZGrid: Integration with XWebCH
  - Grid Portal for Workflow Monitoring
- **New Scientific Workflow Applications**
  - Computational Chemistry Workflows
- **Web 2.0 Mashup Development**
  - RESTful Web Service Composition
  - Pipelined processing over RSS Feeds

# Free Download

<http://www.jopera.org/latest>

## More Information

<http://www.jopera.org/docs/publications>