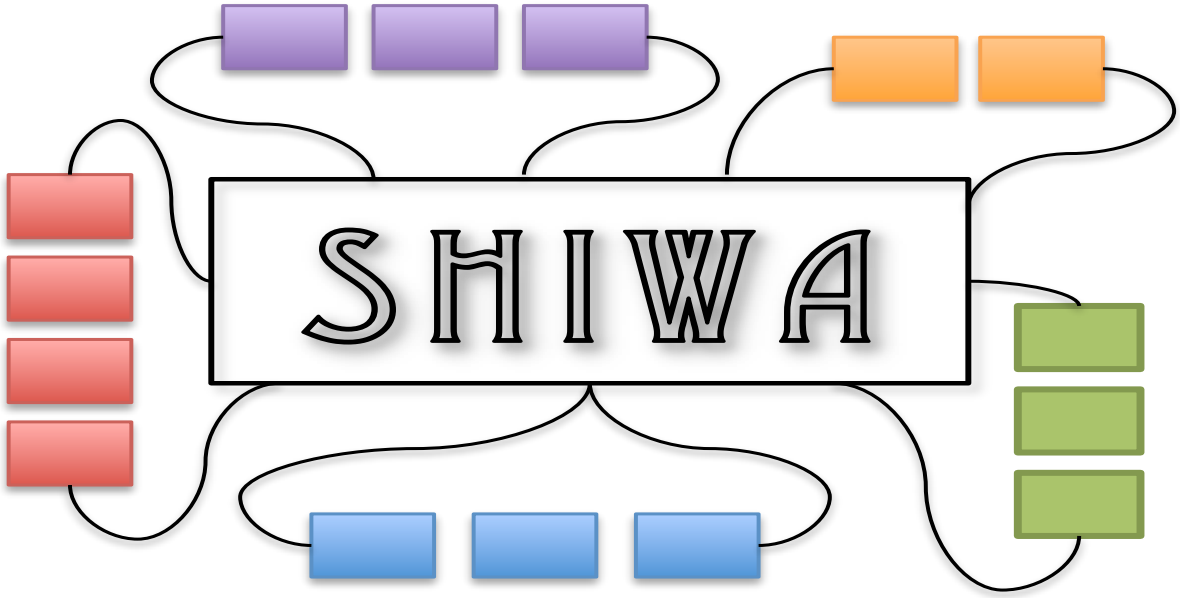


SHaring Interoperable Workflows for Large-Scale Scientific Simulations on Available DCIs

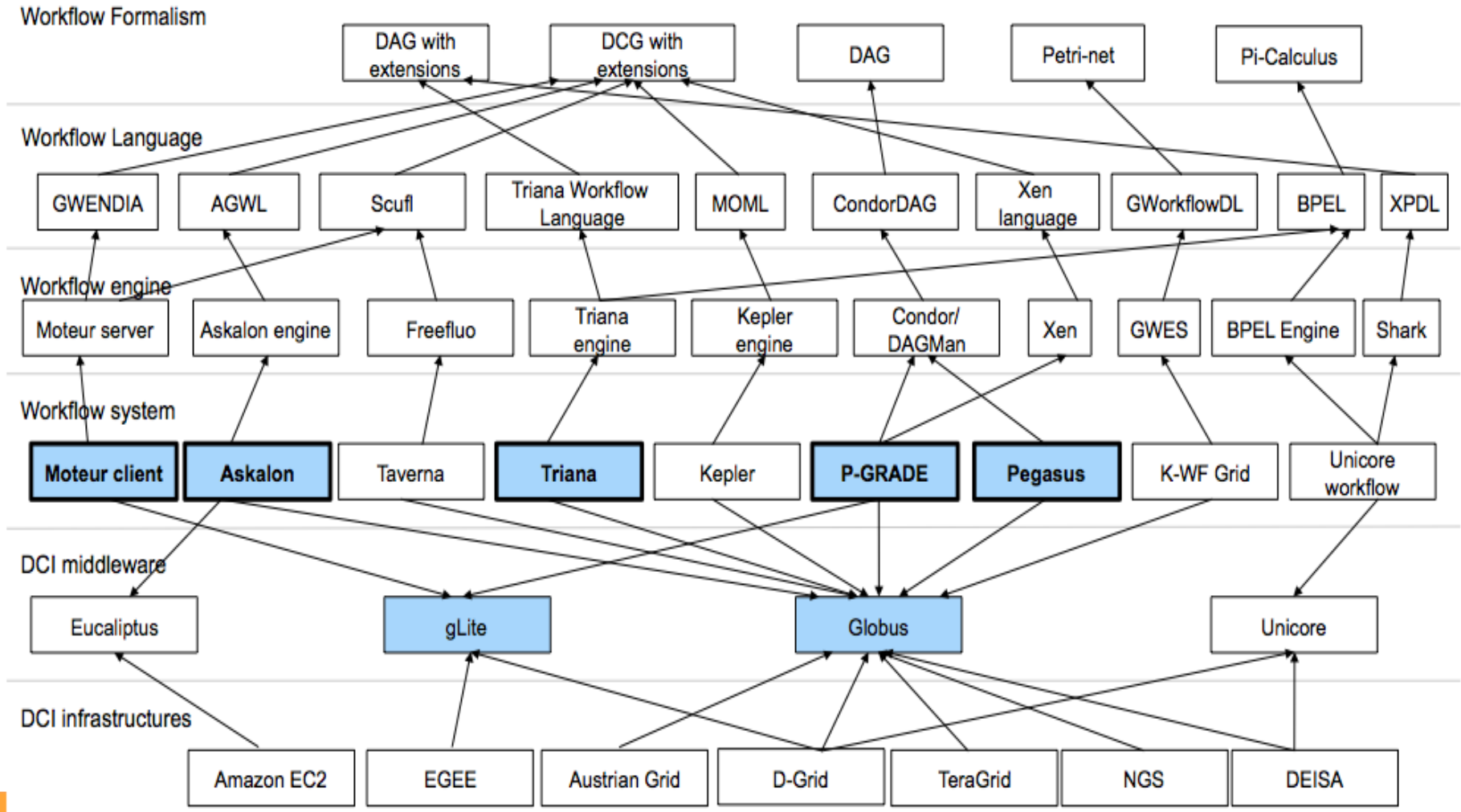
Workflow Interoperability Enables Interdisciplinary Research

Gabor Terstyanszky, University of Westminster

02nd February 2012

SHIWA is supported by the FP7 Capacities Programme under contract No. RI-261585

Workflow Interoperability Challenge



SHIWA Project

Partners:

Computer and Automation Research Institute, Hungarian Academy of Sciences	MTA SZTAKI	Hungary
University of Innsbruck	UIBK	Austria
Charité - Universitätsmedizin Berlin	C-UB	Germany
French National Centre for Scientific Research	CNRS	France
University of Westminster	UOW	United Kingdom
Cardiff University	CU	United Kingdom
Academic Medical Centre of the University of Amsterdam	AMC	Netherlands
University of Southern California	USC	USA

Duration:

July 2010 – June 2012 (extended to September 2012)

SHIWA Objectives and Services

Objectives:

- To create an environment which enables seamless execution of workflows of different workflow systems through workflow interoperability

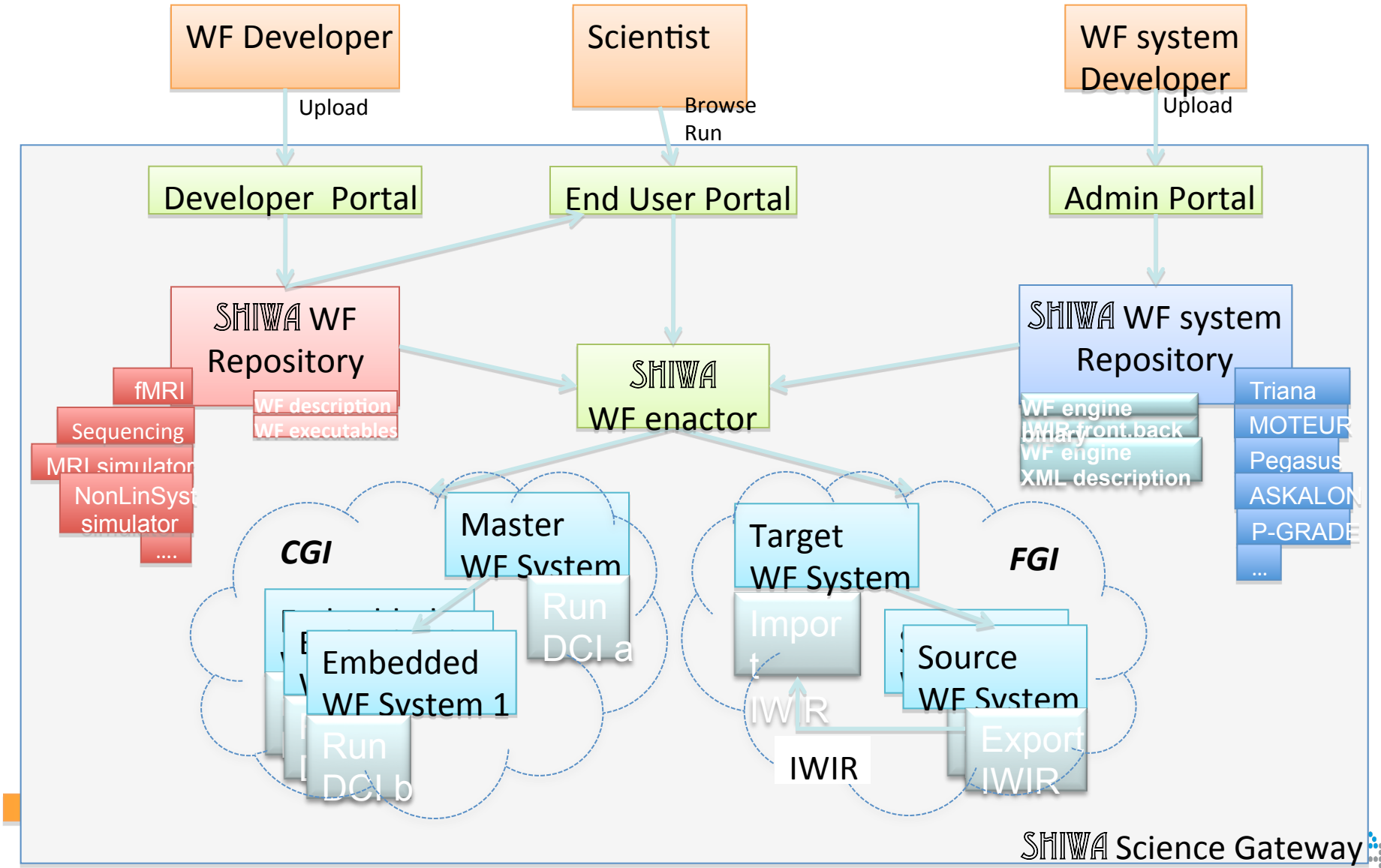
Services:

- supports the whole workflow lifecycle: editing, uploading, browsing downloading and executing workflows
- provides coarse- and fine-grained workflow interoperability solution
- offers Distributed Computing Infrastructure interoperability solution
- provides desktop computer and portal interfaces to manage workflows

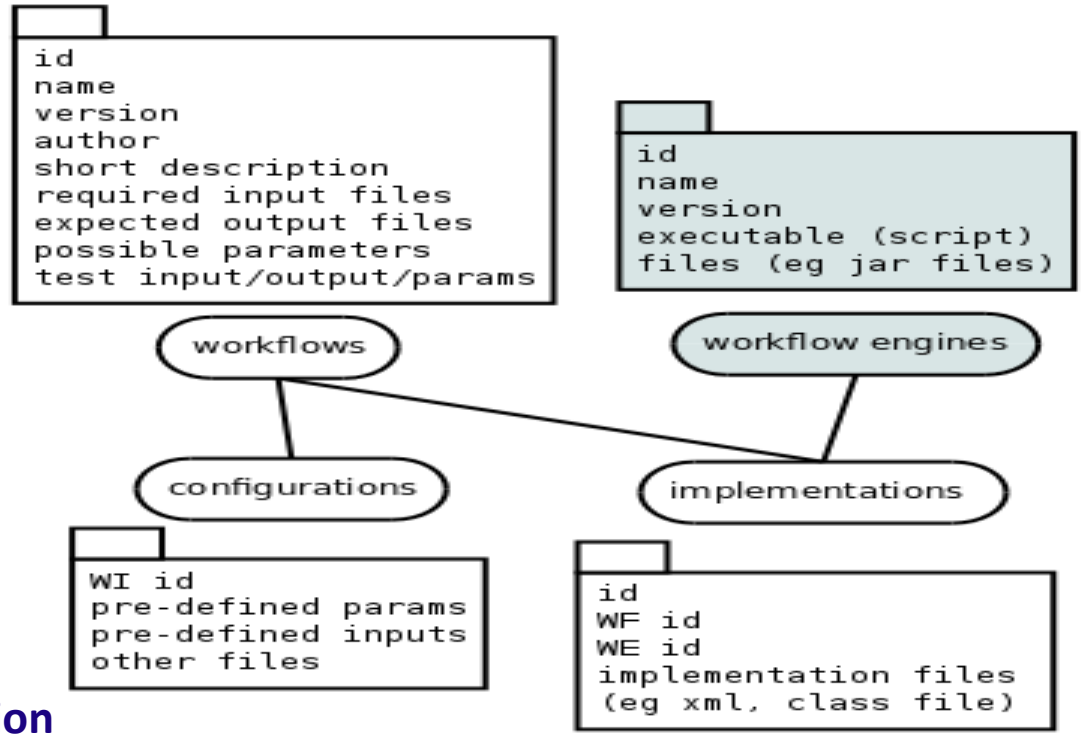
Key actors:

- researchers
- workflow engine developers
- workflow developers

SHIWA Solution



SHIWA Data Structure



Workflow description

- plain text which describes the inputs and outputs and explains what the workflow does

Workflow implementations

- It contains all implementation files or references to these files (via eg. URLs) and also holds other data/ metadata necessary to run the workflow on its associated workflow engine.

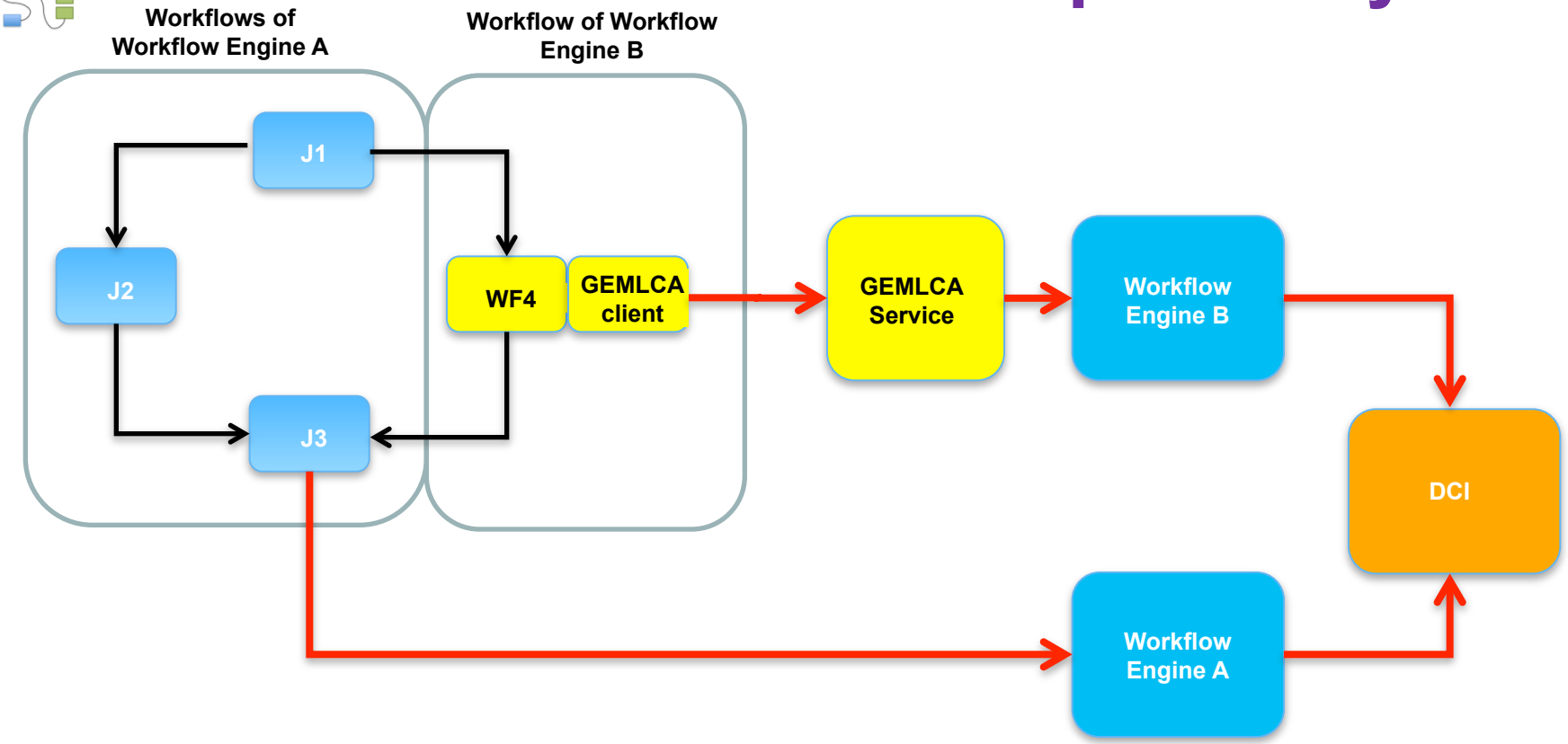
Workflow configurations

- It contains pre-defined input files and parameters or references to these files and parameters and other data/ metadata of a workflow.

Workflow engines

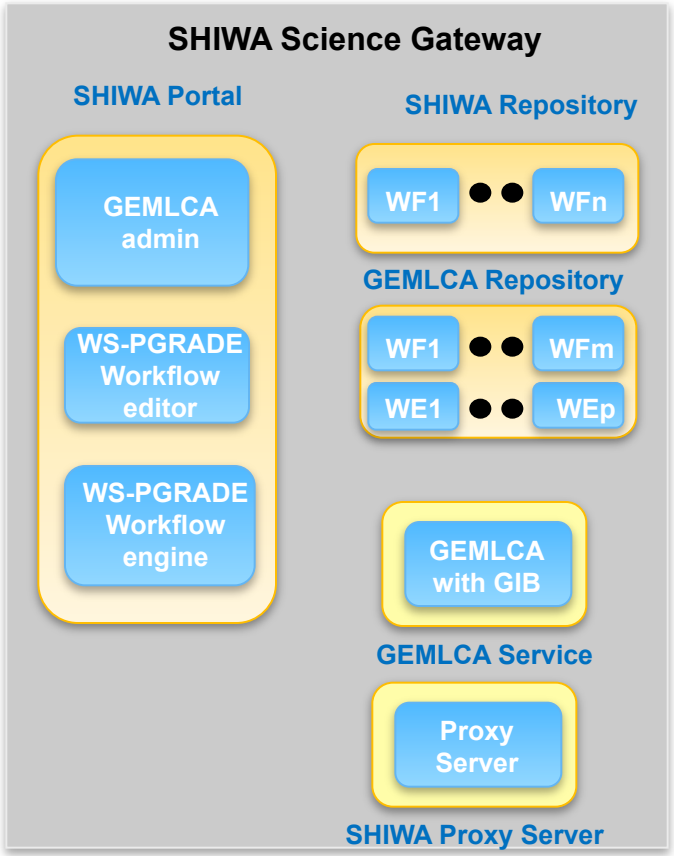
- It contains files and other data/metadata necessary to execute a workflow engine on a grid site or references to them

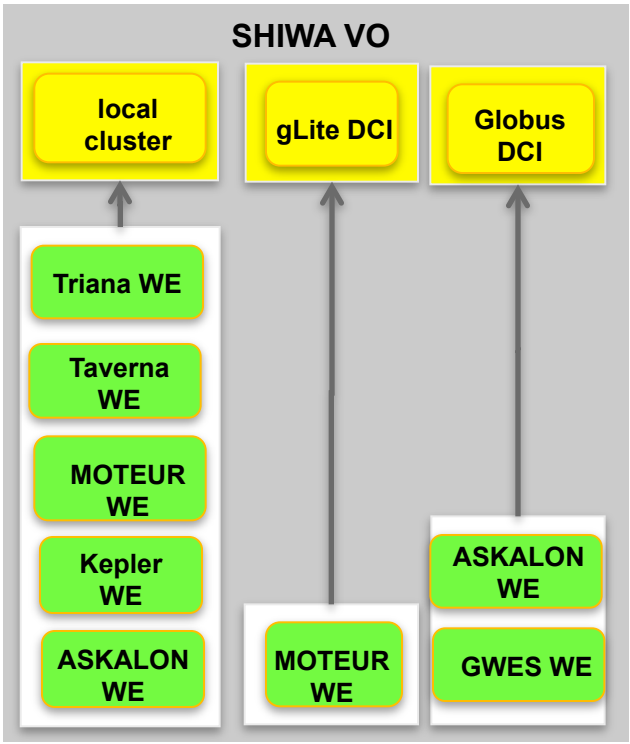
Coarse-Grained Interoperability



- native workflows: J1, J2, J3
- non-native workflows: WF4
  - black boxes which are managed as legacy code applications

CGI Infrastructure





SHIWA Science Gateway

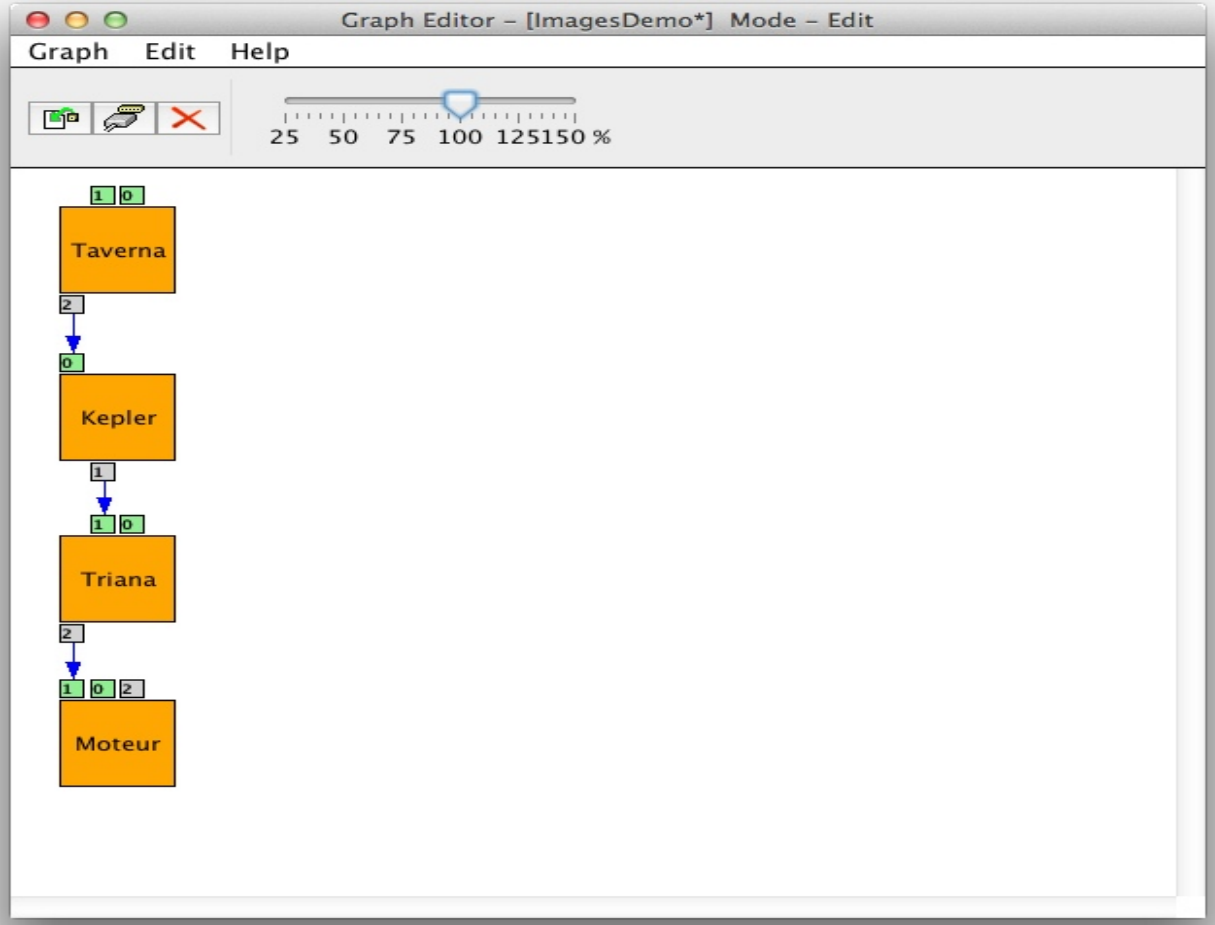
native WE	PGRADE
portal	PGRADE v2.4.1
repository	GEMLCA + SHIWA repo
submitter	GEMLCA with GIB
proxy management	SHIWA Proxy Server

Resources

local resources: invocation of locally deployed WEs
WE submission to local cluster

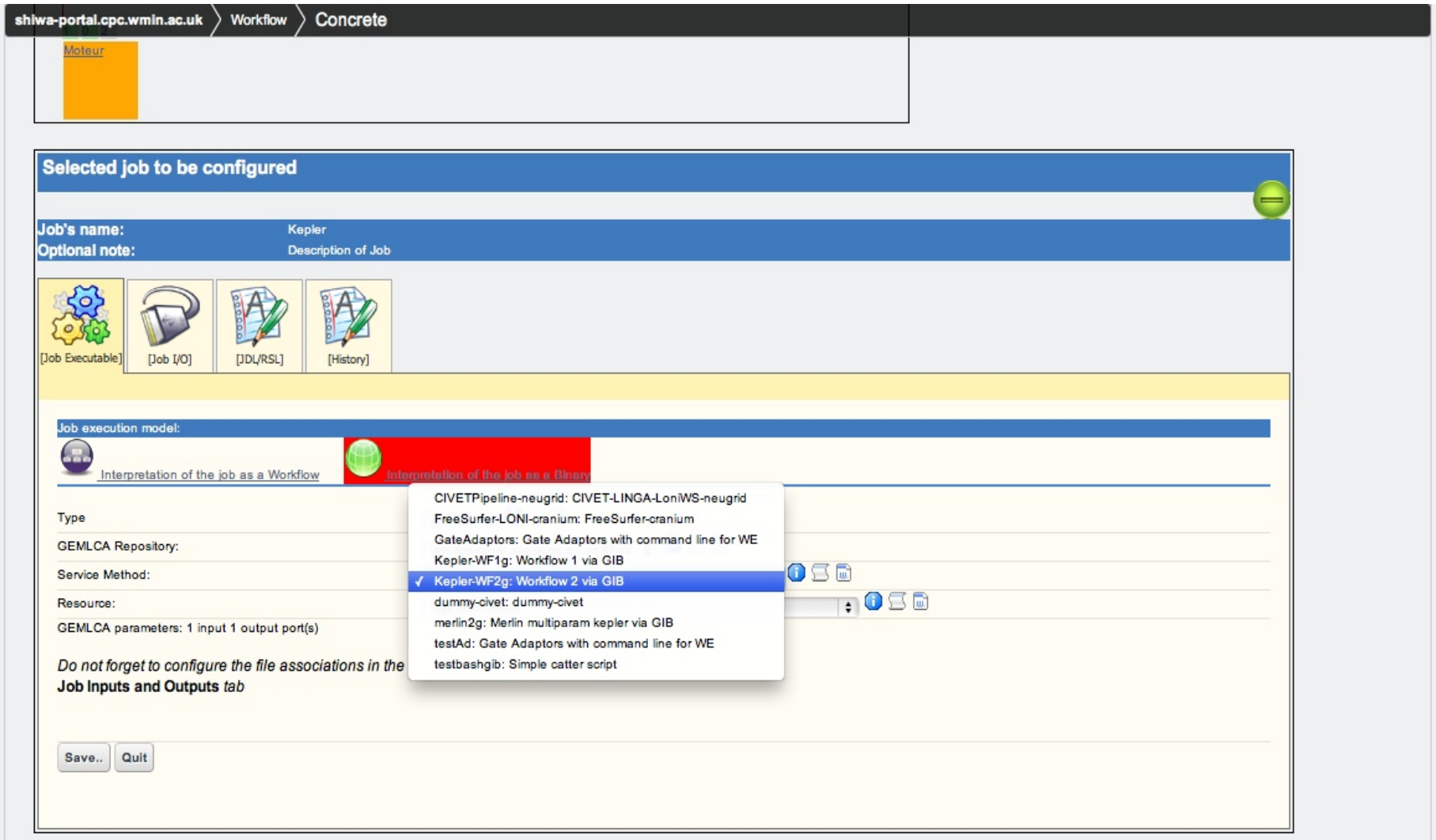
remote resources: through remotely pre-deployed WEs to gLite and Globus DCIs

SHIWA Portal: Workflow Editor



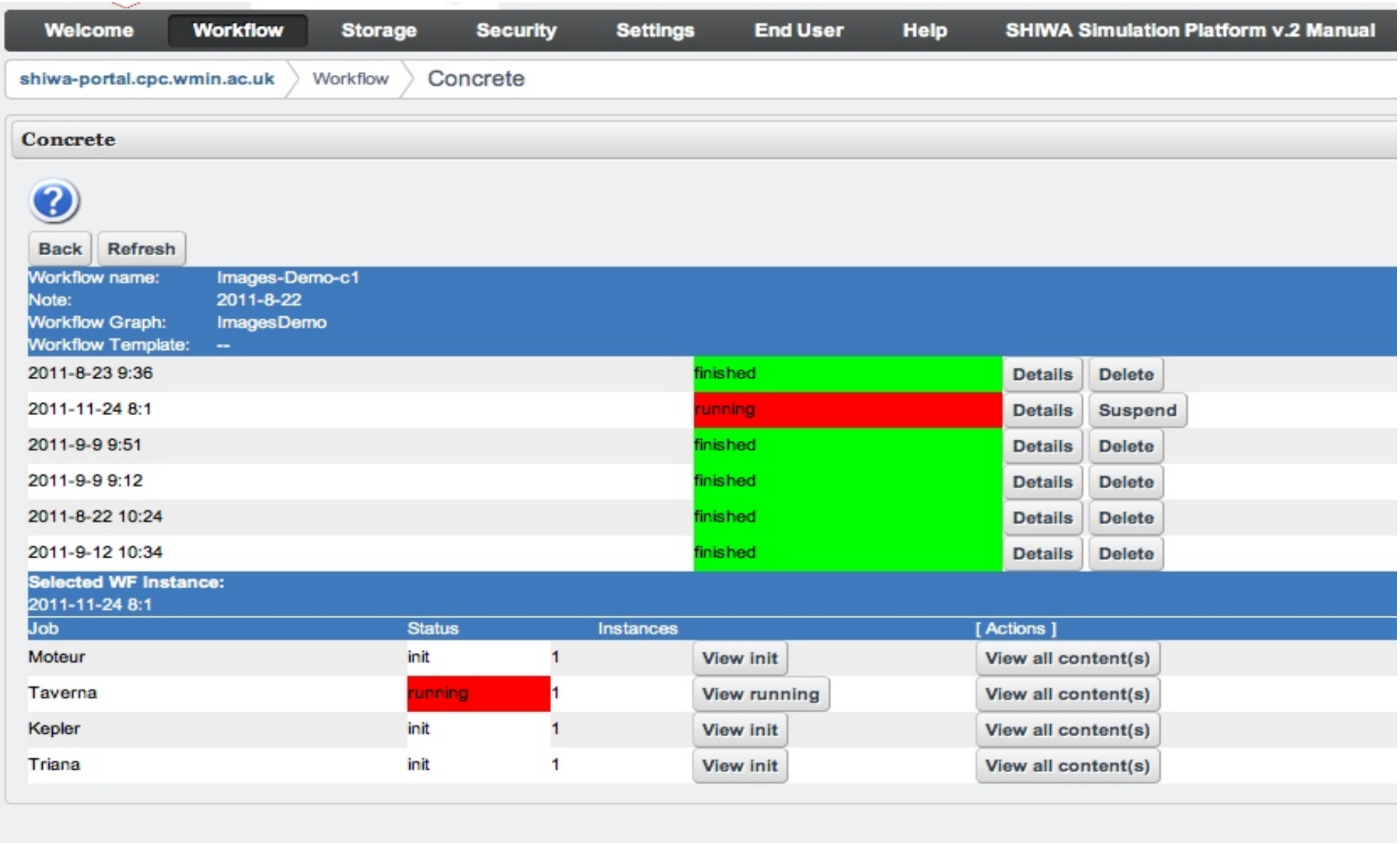
e-infrastructure

SHIWA Portal: Configuring Workflow



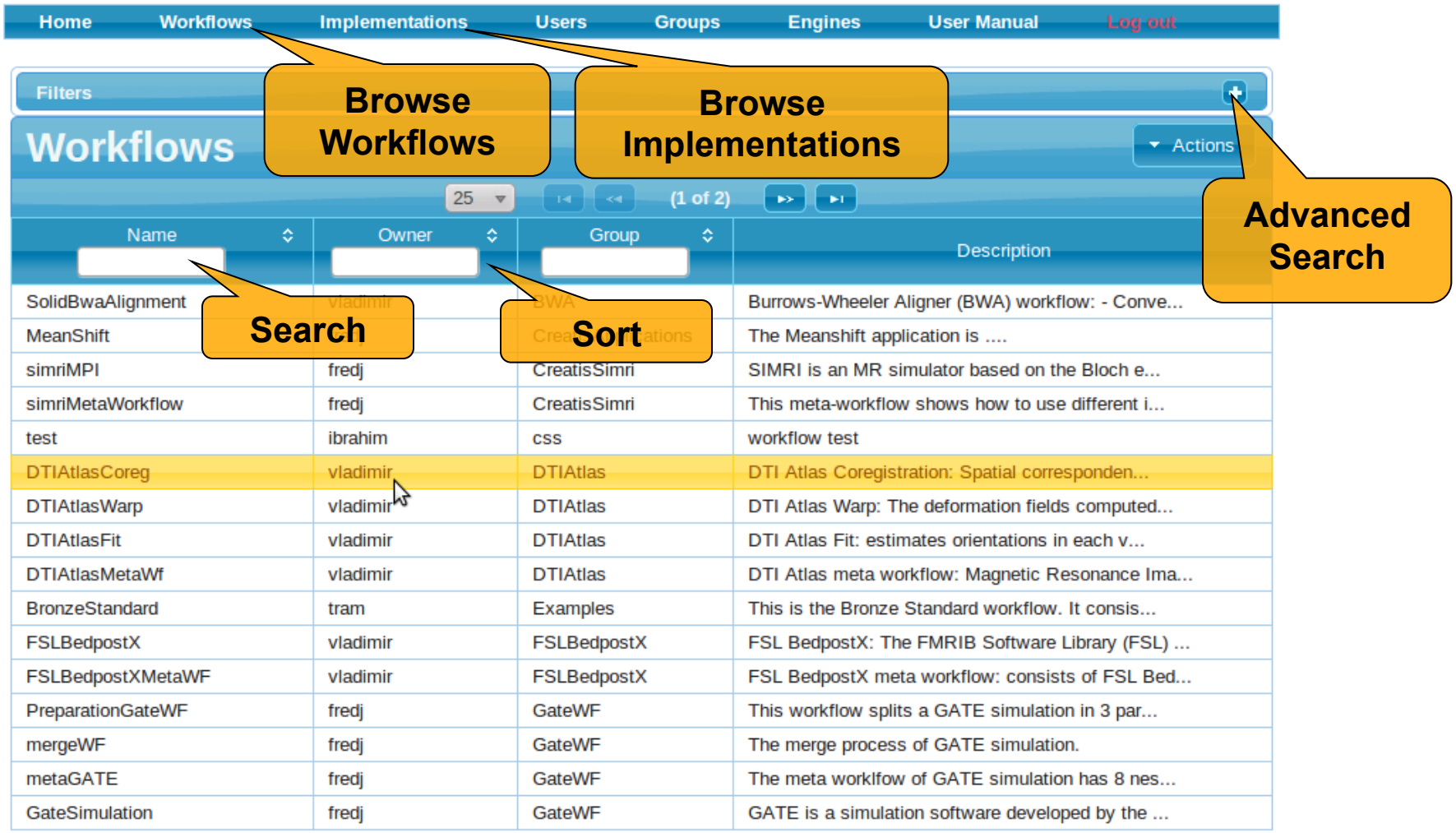
e-infrastructure

SHIWA Portal: Executing Workflow

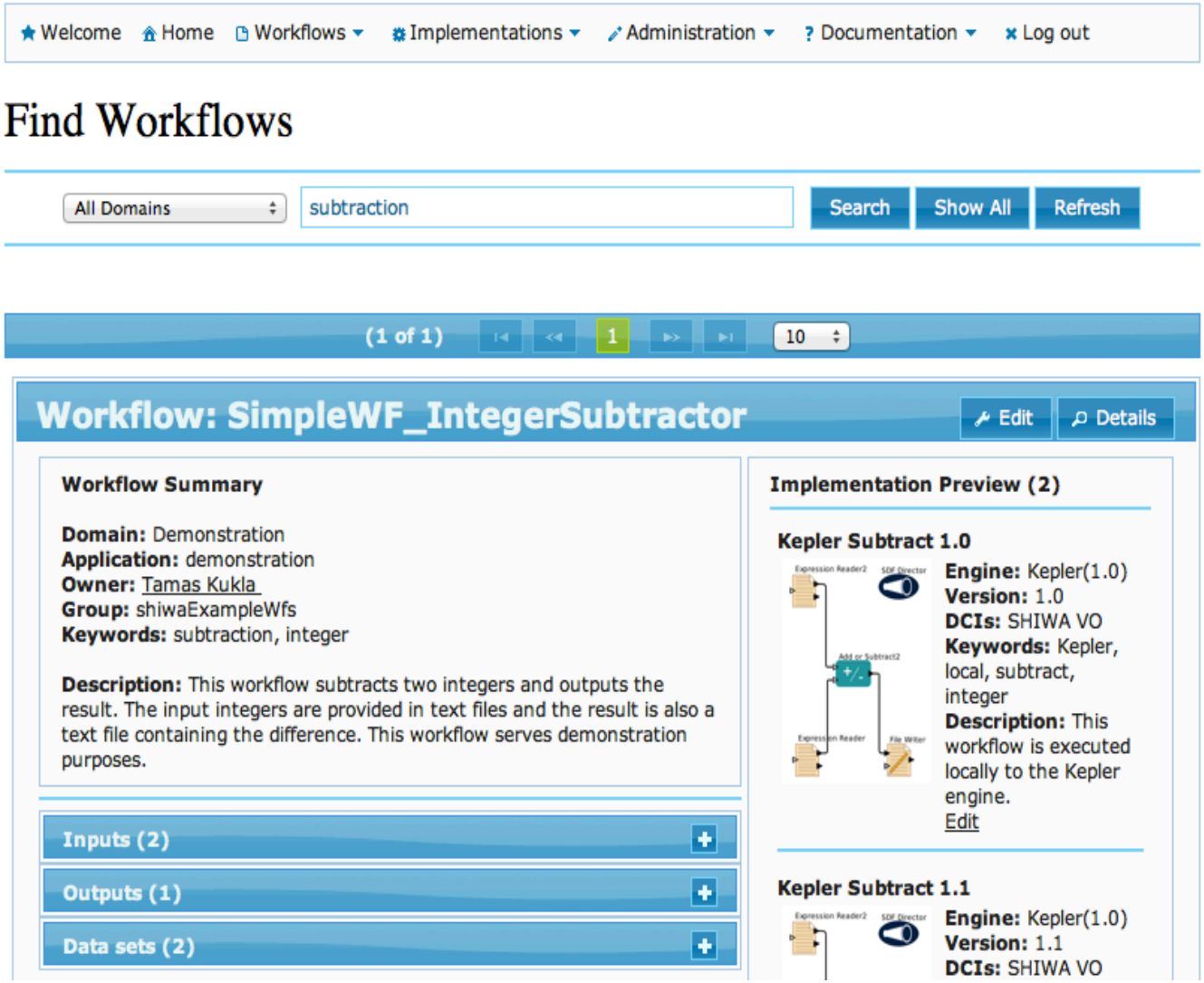


e-infrastructure

SHIWA Workflow Repository

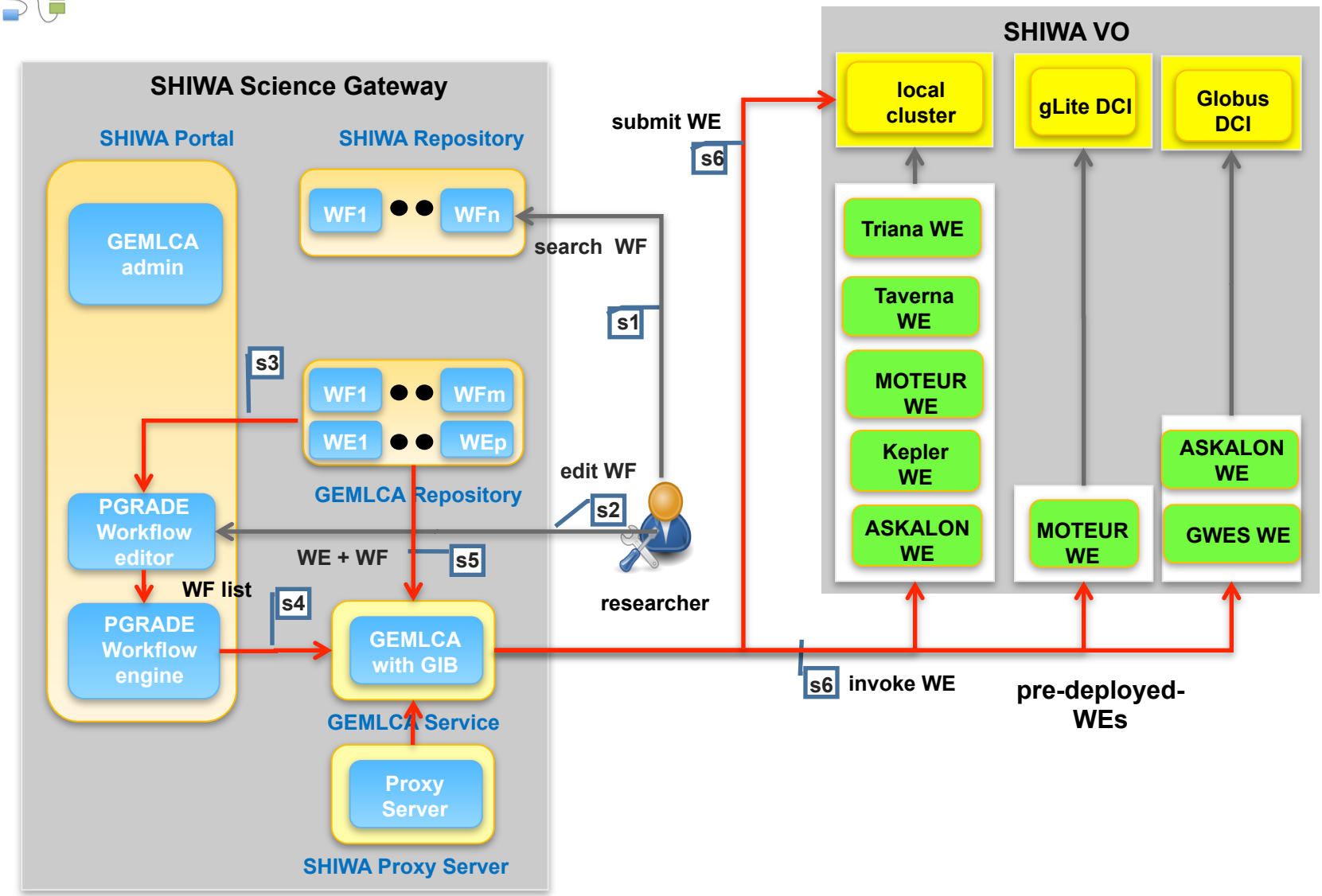


SHIWA Workflow Repository

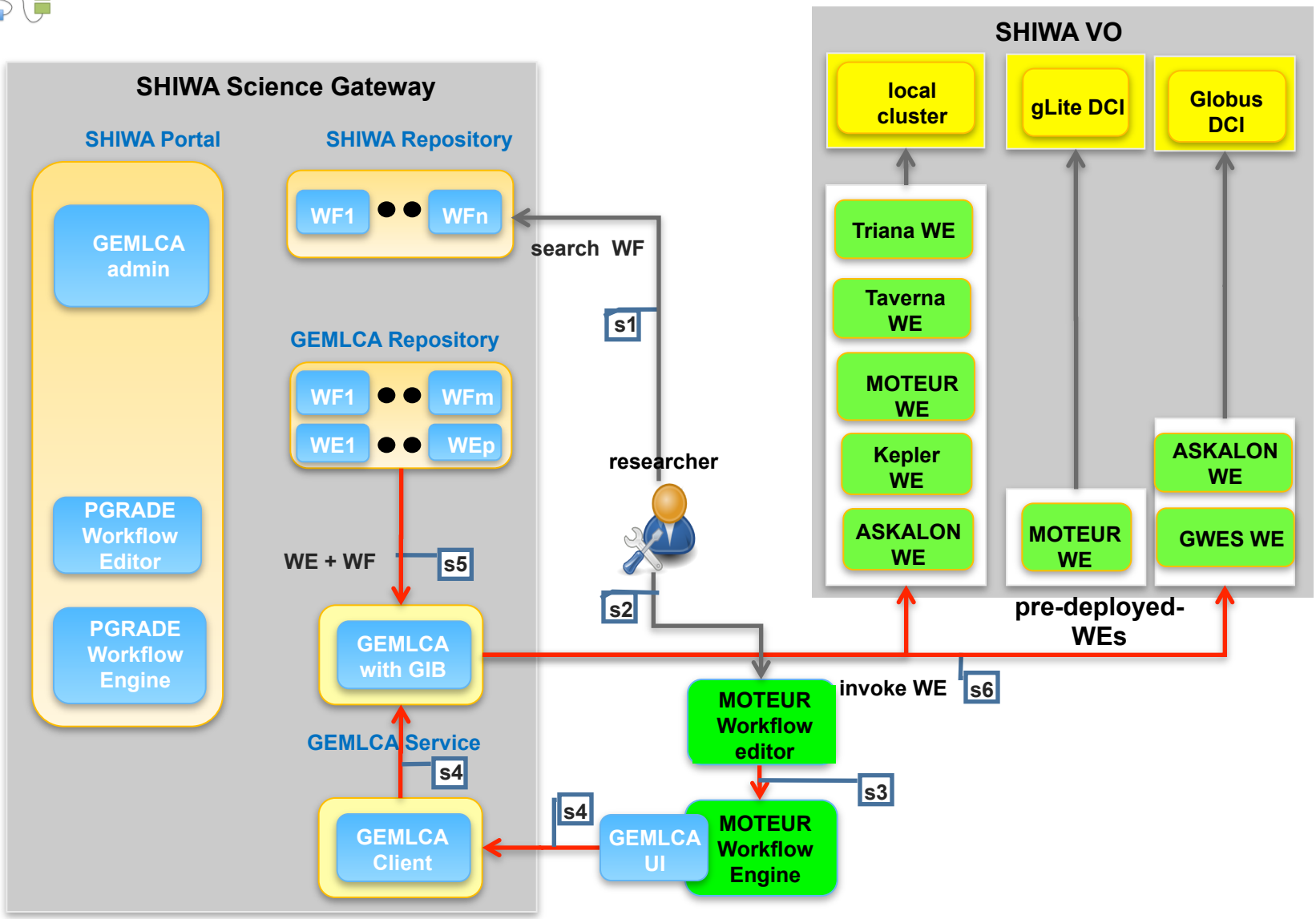


e-infrastructure

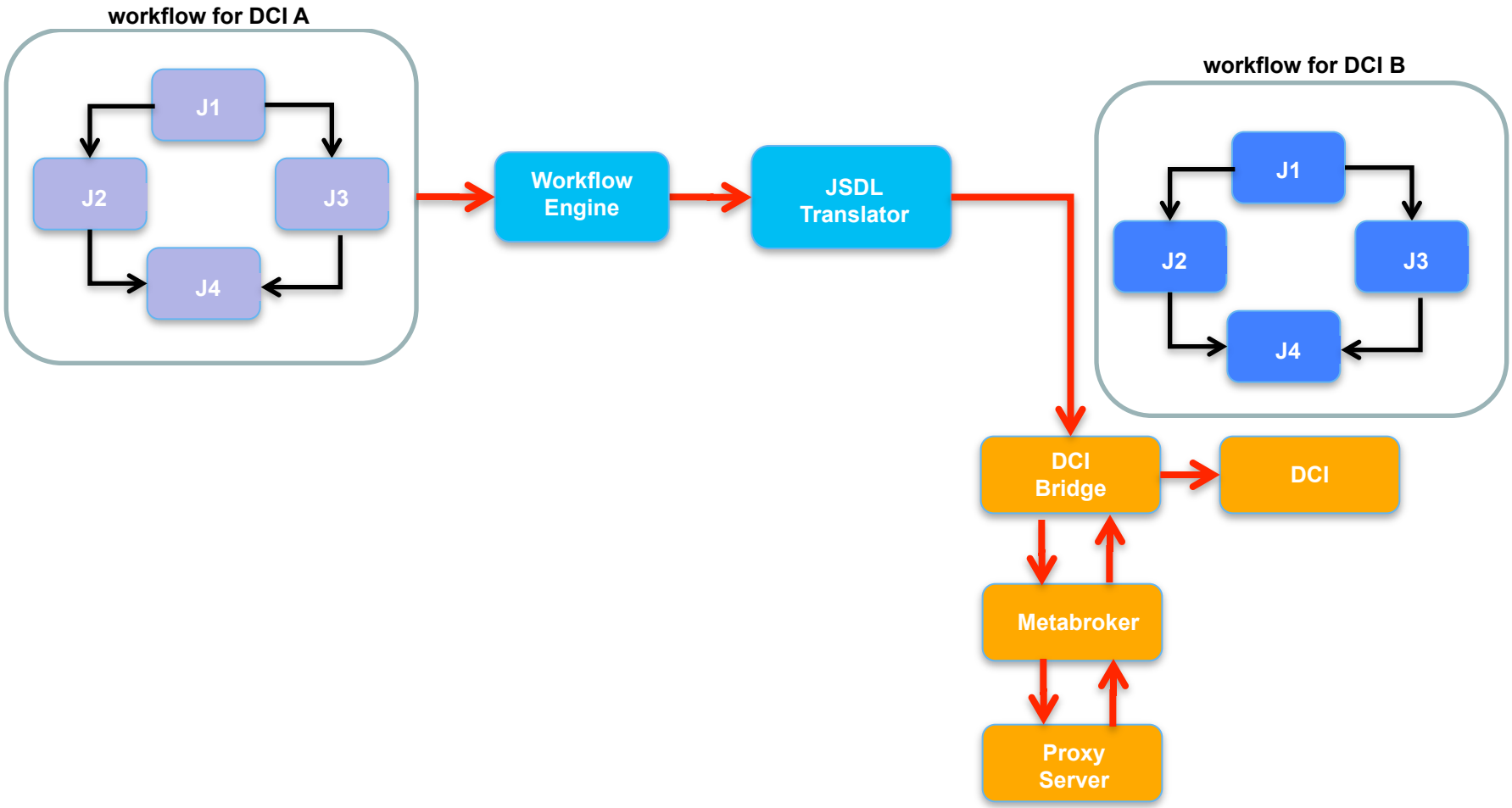
CGI User Scenario: Native WE



CGI User Scenario: Non-native WE

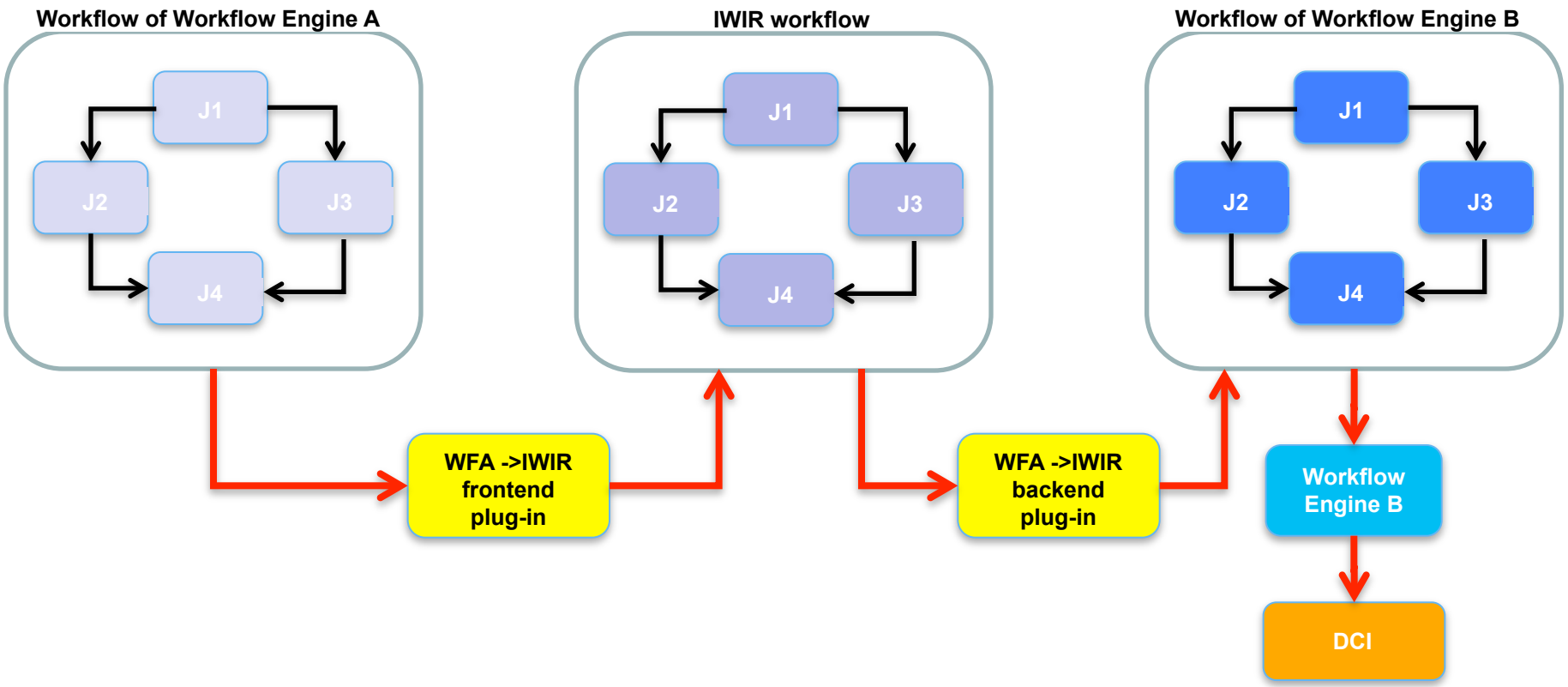


DCI Interoperability



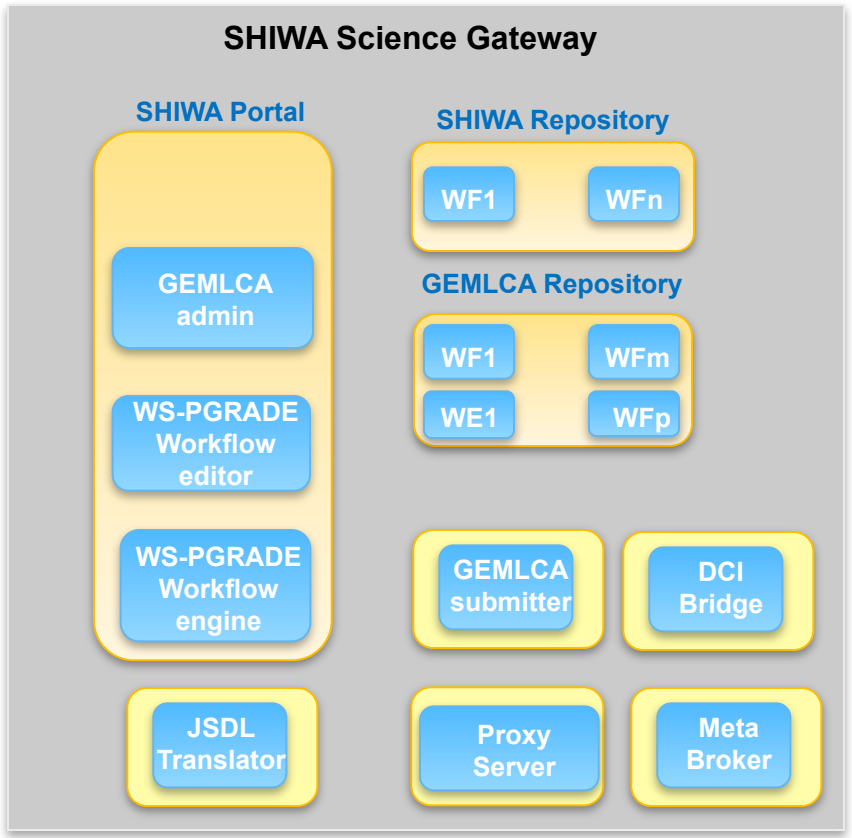
e-infrastructure

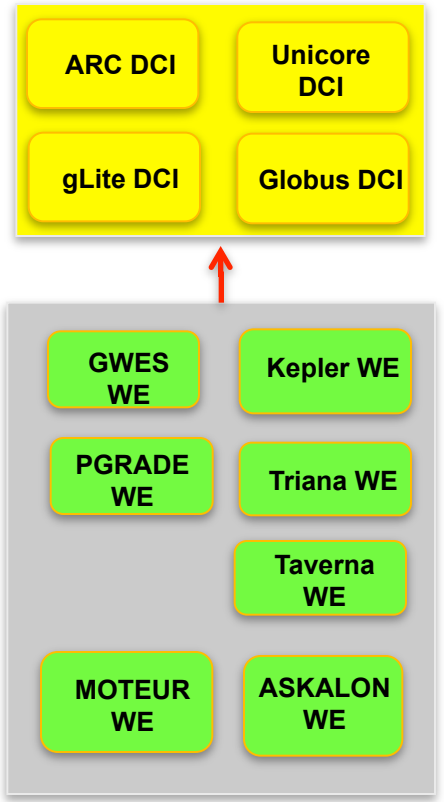
Fine-Grained Interoperability



- front-end plug-in: converts workflow of workflow engine A into IWIR representation
- back-end plug-in: converts from IWIR representation into workflow of workflow engine B

SSP DCI & FGI Infrastructure





SHIWA Science Gateway

native WE	ASKALON, MOTEUR, WS-PGRADE
portal	WS-PGRADE v3.2.2
repository	GEMLCA + SHIWA
GEMLCA	
proxy server	

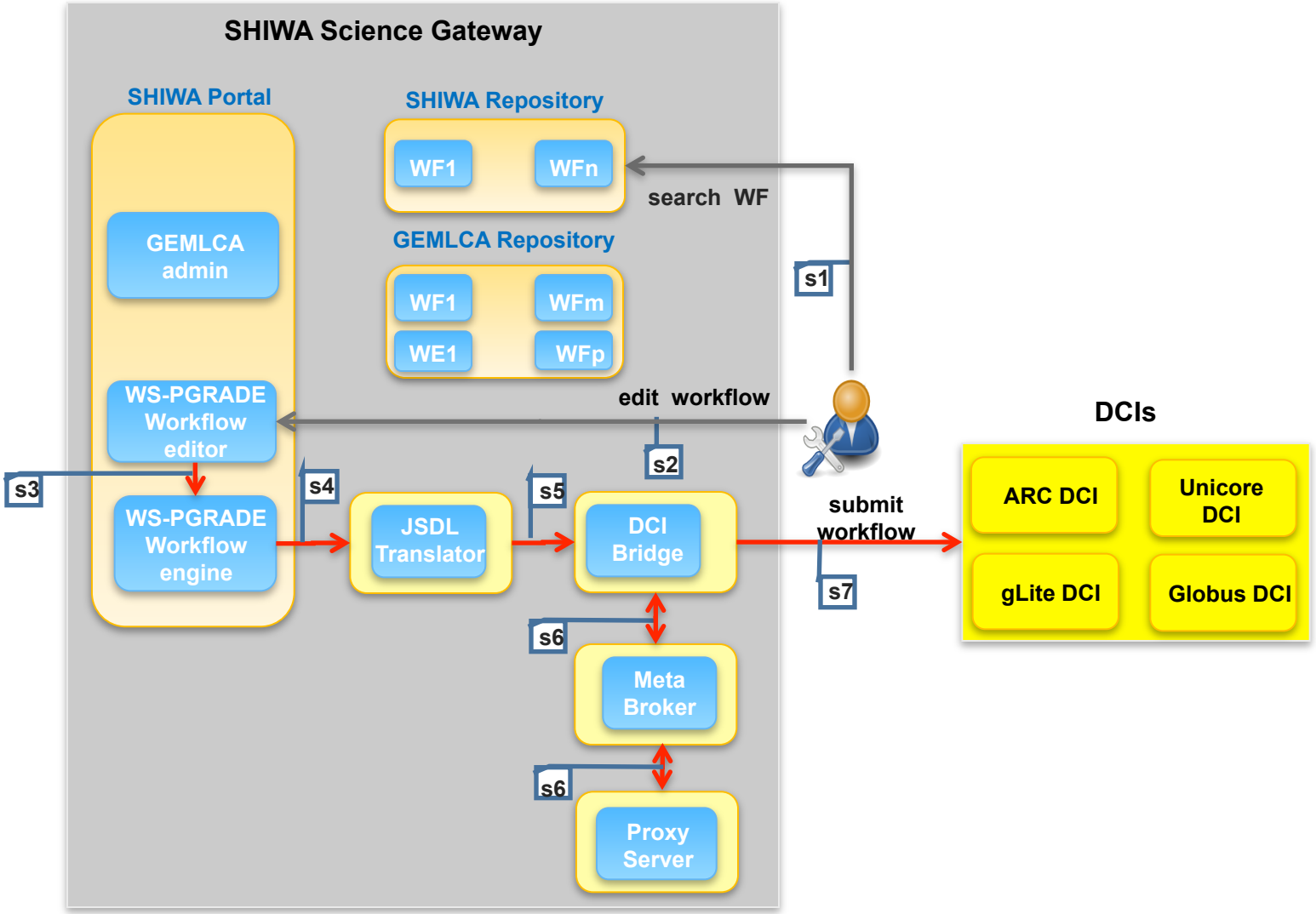
Resources

local WEs: Kepler, MOTEUR , Taverna & Triana WEs
submit to the local cluster

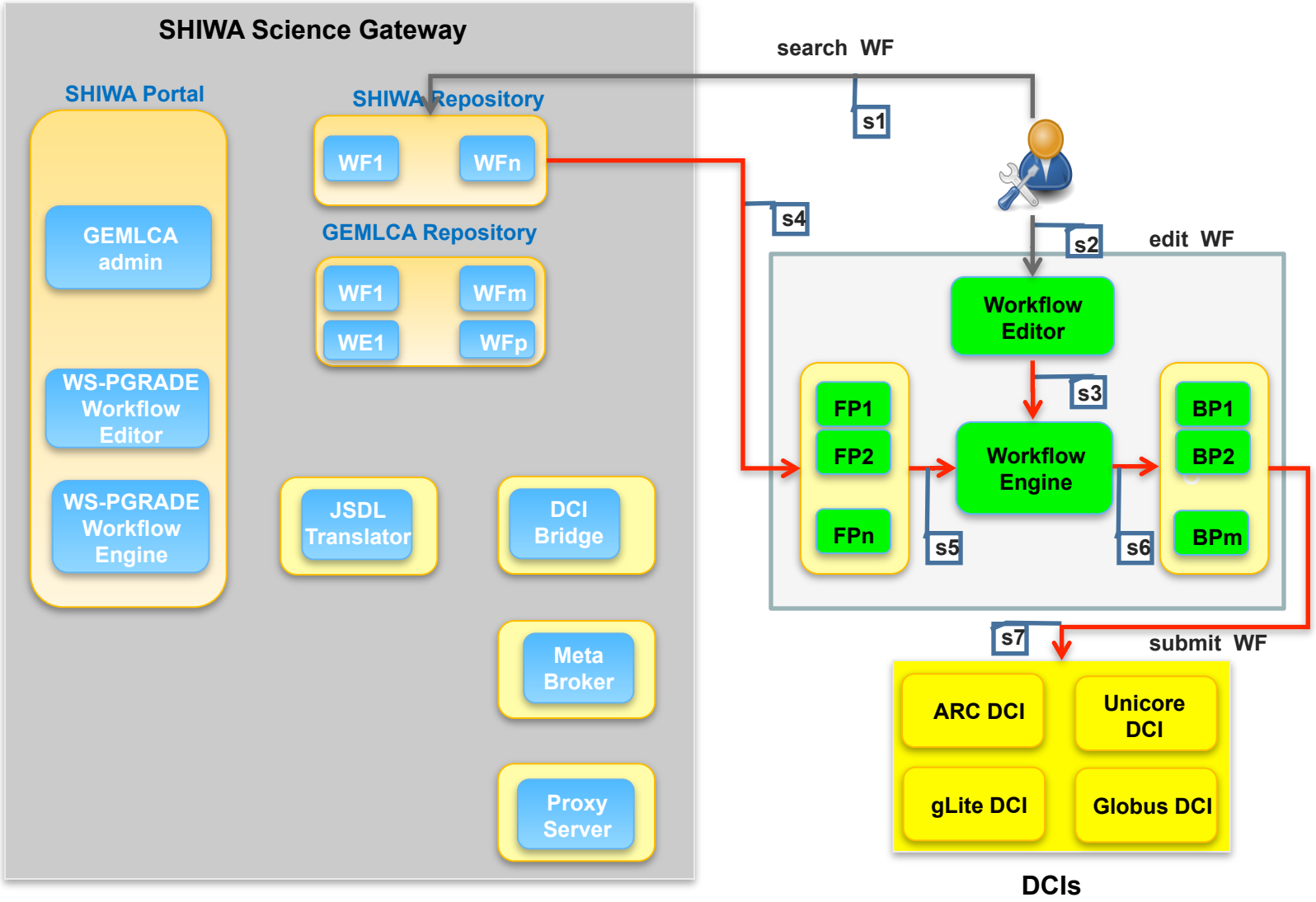
remote WEs: ASKALON, GWES & MOTEUR WEs
ARC, gLite, Globus and Unicore DCIs

e-infrastructure

DCI Interoperability Scenario



FGI Interoperability Scenario



e-infrastructure

SHIWA Simulation Platform: Access

SHIWA Science Gateway (SHIWA Portal + SHIWA Repository)

- portal and repository account given by the science gateway administrator

DCI resources

- certificate given by the relevant DCI administrator, for example the VO sysadmin

Access to the SHIWA Portal

shiwa-portal.cpc.wmin.ac.uk/liferay-portal-6.05

Access to the SHIWA Repository

shiwa-repo.cpc.wmin.ac.uk

SHIWA Simulation Platform: Selling Points

- the simulation platform supports the whole lifecycle of workflows, i.e. creating, testing, uploading, browsing, downloading and running workflows
- users can use workflows of their own and of other workflow systems through the same user interface, i.e. they can re-cycle or share workflows created by other research teams or even other research communities
- workflow developers (or e-scientists) can create, run, upload workflows through the simulation platform
- researchers (astro physicists, earth scientists, bio scientists, etc.) can browse the workflow repository, select and run workflows through the simulation platform
- the simulation platform provides access to major European e-infrastructure, the European Grid Initiative (EGI) infrastructure which offers large pool of compute and storage resources