# SHIWA and ER-flow: Services, use cases, recommendations

### Gergely Sipos, EGI.eu,

### Last update: 13/02/2013

This short document provides a summary of the technologies that were developed by the SHIWA project and are now used within the ER-flow project to build and expand workflow communities in Europe. The text explains the relationship among the different services that are included in the SHIWA technology package and explain the core set of use cases that these services can support. The use cases include different actors from principal investigators, to research fellow, trainers and students, scientific programmers and operators of services for scientific communities. The document also provides a few recommendations concerning the further development of some of these SHIWA services in order that the SHIWA technology package can better serve the ER-flow project in reaching new communities. The recommendations build on the assumption that structured scientific communities, who already operate web portals, would like to consume SHIWA services directly from these portals and the SHIWA technology should better support this. The recommendations therefore suggest the SHIWA community to extend the capabilities of the SHIWA Simulation Platform towards third party web portals that are operated by scientific communities, for example through the support these communities receive in the SCI-BUS project. The document has three goals:

1. Provide text for the ER-flow WP2 activity that can be reused on the ER-flow project website, the SHIWA User Forum website, and in various other promotion materials (leaflets, white papers, etc.). To help ER-flow reach new user communities these forums should have easy to understand descriptions on the use cases and the technologies that ER-flow uses to support these.
2. Help the ER-flow consortium reach a common understanding about the capabilities of the SHIWA services. This is especially helpful for WP5 members, who need to identify those use cases that their community needs to implement in the ER-flow project.
3. Initiate a dialog between the ER-flow project and the SHIWA community concerning the further development of some of the SHIWA services, so these could better support structured communities as well as simplify the monitoring of technology uptake for ER-flow.

## SHIWA technology in ER-flow

The SHIWA Simulation Platform (SSP in short) is one of the technical results of the SHIWA project. The SHIWA Simulation Platform is a service package, already used by various scientific communities, and used as a tool by the ER-flow project to expand existing workflow communities and to build new workflow communities in Europe. The SHIWA Simulation Platform includes the following services:

1. SHIWA Repository: A database where workflows and meta-data about workflows can be stored. The database is a central repository to discover and share workflows within and among communities[[1]](#footnote-1).
2. SHIWA Portal: A web portal that is integrated with the SHIWA Repository and includes a workflow executor engine that can orchestrate various types of workflows on various grid and cloud platforms. Workflow execution is orchestrated by the workflow systems[[2]](#footnote-2) that are interfaced with the portal.
3. SHIWA Desktop: A desktop environment that provides similar access capabilities than the SHIWA Portal, however it runs on the users’ desktops/laptops instead of a portal server.

The SHIWA Simulation Platform can be used in various scenarios that involve the development, execution, integration and sharing of workflows, or workflow systems. The core use cases that such scenarios are built from are summarised in the below table. Communities can apply the SHIWA Simulation Platform and support services from the ER-flow project to implement the use cases and to integrate these into larger, custom scenarios to the benefit their members.

Use cases supported by ER-flow using the SHIWA technology:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| USER PROFILE | USER GOAL | TASKS | TASK GUIDANCE | AVAILABLE TOOLS | BENEFITS |
| Principal investigator, demonstrator, research student, research fellow | Run a pre-defined scientific calculation, for example, to repeat an experiment or to demonstrate a scientific method during a training. | Execute a workflow | Select a workflow that encapsulates the simulation. Import the workflow into the execution environment, then run it with the default, or with an updated set of input parameters. | SHIWA Workflow Repository and SHIWA workflow execution environment.  Both are accessible through the SHIWA Portal or the SHIWA Desktop. There are a growing number of community specific portals that also provide access to these SHIWA services. | Services of the SHIWA Simulation Platform make scientific calculations easily accessible and repeatable by non-technical users. These users can browse, select and execute scientific calculations from either desktop or portal environments. |
| Research fellow, scientific programmer | Capture a scientific calculation into a process that can be repeated at a later stage in an unmodified, or in some customised form. | Develop a workflow | Use a workflow editor system to define the steps of the calculation, input data for the calculation, and data dependencies among the computational steps. | The SHIWA Simulation Platform includes various workflow systems. Workflows can be built in these systems, then made compatible with the SHIWA workflow executor system using the P-GRADE workflow interface. | The SHIWA Simulation Platform offers technological freedom in choosing a system for workflow development. The P-GRADE system inside the SHIWA Simulation Platform guarantees that workflows built in any of the supported workflow languages are executable through the SHIWA Portal and through third party portals that are integrated with the Platform. |
| Research fellow, scientific programmer | Integrate multiple scientific calculations into a single computation, for example to compare different approaches or to achieve results that can be delivered only with an integrated system. | Integrate workflows | Use a workflow editor that can encapsulate workflows that are written in different workflow languages with different tools. | One of the unique feature of the SHIWA Portal is that its P-GRADE workflow editor can integrate workflows built with different workflow tools into a single workflow, and support the execution of such complex workflows using the native workflow systems. | The SHIWA technology enables the reuse of workflows outside of their native environments, as well as the integration of workflows with other workflows that are built in a completely different language and system. These features improve the reuse of workflows inside, and across communities. |
| Research fellow, scientific programmer | Share scientific calculation with research collaborators or with the general public, so they can repeat and reuse these on demand. | Share a workflow | Use the Repository to store and share a workflow with colleagues or with the general public. Provide a description and other meta-data for the workflow so others can find and use it effectively. | SHIWA Workflow Repository. It is accessible from the SHIWA Desktop, the SHIWA Portal, or from community specific portals. | Using the SHIWA Workflow Repository your workflow becomes accessible for scientific research communities around the world. The workflows are directly executable through the SHIWA Portal, through third party portals, or downloadable for offline use and archiving. Workflow meta-data ensure that the shared items are accessible and useable by those, who are targeted with your work. |
| Project/community representative, scientific programmer | Support scientific communities to define, execute and share novel types of data processing workflows. | Integrate a workflow system | Integrate your workflow system with the SHIWA platform to make your system compatible with other workflow environments, and with various cloud and computing grid infrastructures. | Graphical and scripting interfaces exposed by the SHIWA Simulation Platform for workflow system integration. | The SHIWA technology provides a platform for workflow systems to reach new scientific communities.  Through the platform these workflow systems can access computing and storage resources from various grids and cloud infrastructures in Europe and beyond. Through the SHIWA technology these workflow systems can support simulations that require the integration of multiple workflows may built in different workflow systems. |
| Project/community representative, scientific programmer | Enrich a community specific portal with workflow development, sharing and execution capabilities. | Integrate a community portal | Interface your portal environment with SHIWA services. The Repository provides a storage for workflows and their meta-data. The | The SCI-BUS Science Gateway Framework includes portlets that can be use in Liferay based Web portals to interact with services of the SHIWA Simulation Platform. | By integrating the services of the SHIWA Simulation Platform into community portals, these SHIWA services become directly accessible to community members. The SHIWA services can be useful for these portal users in various scenarios that involve the development, execution, integration and sharing of workflows built with any of the workflow languages and systems that are supported in SHIWA. |

## Workflow systems in the SHIWA Simulation Platform

One of the most important questions when developing a workflow is “which workflow system to use?” The choice depends on various factors, most importantly on the structure and type of the calculations that one wants to capture and connect together in the form of a computational workflow. The SHIWA Simulation Platform is already integrated with various workflow systems, and acts as an open platform, so other workflow systems can be also connected to it. The native workflow system that is used by the SHIWA Simulation Platform is P-GRADE. P-GRADE can be used not only to develop new computational workflows, but also to develop workflows that integrate other workflows that have been prepared in any of the other systems. Note that only P-GRADE has a graphical interface inside the SHIWA Simulation Platform. The other systems can be used by their own native interfaces outside SHIWA.

## Recommendations for SHIWA

After the SHIWA project is over the SHIWA Simulation Platform is supported by the SHIWA community[[3]](#footnote-3). Services of the Platform can be accessed and consumed from the SHIWA Portal and SHIWA Desktop environments. While this access method is sufficient for individual users, it is less attractive to communities who already operate portals for their members, and who would rather consume SHIWA services directly in these portals than in environments provided by a third party. ER-flow aims to support various scientific communities in Europe, and such a use case is likely to be raised very soon. To meet this demand the services of the SHIWA simulation Platform need to become customisable and embeddable into third party portals. This section provides a set of recommendations to the SHIWA community on how this could be achieved.

The main idea behind the proposal is to operated the services of the SHIWA Simulation Platform centrally for every SHIWA consumer (including the ER-flow project), and develop APIs for the services so these consumers could access SHIWA functions from any client they want. The APIs could come with client libraries and/or Web gadgets[[4]](#footnote-4) to simplify the integration of SHIWA services into third party websites. The APIs should expose the key functions that services of the SHIWA Simulation Platform includes. A possible set of these functions and relates APIs could be:

* Repository API: an interface to interact with the SHIWA Repository in order to upload, download, search for and and manage workflows and meta-data. The Repository API should expose the services of the SHIWA Workflow Repository.
* Simulator API: A service to run a specific workflow from the SHIWA Repository. The simulator API should expose those services of the SHIWA Simulation Platform that are currently responsible for workflow execution. There is no dedicated service at the moment in SHIWA for this function. Workflow execution is tightly coupled with the SHIWA Portal.

Both the Repository and the Simulator APIs should support third-party accounts, meaning that when the libraries are used within a community portal, then existing users of this portal can access the Repository and Simulator functions without requesting an additional account, and without being asked for re-authentication when accessing these SHIWA services.

Additional benefits of implementing the recommendations:

* Monitoring of SHIWA Simulation Platform usage: If all the SHIWA front ends (including the SHIWA Portal and SHIWA Desktop) would start consuming SHIWA services through the API, then the SHIWA servers that implement the API could monitor and collect statistical data about the use. This information can be useful not only to the SHIWA Community, but also to some of the SHIWA consumers, such as the ER-flow project that needs to collect various metrics concerning the impact of its workflow community building activities.
* Lower cost of maintenance: The proposal recommends the SHIWA Community to centralise functions into services that are centrally provided for all the SHIWA user communities. Consequently, SHIWA technology providers need to (hopefully) distribute a simpler and smaller software package to SHIWA users, who will therefore run into less installation and configuration problems.
* Be prepared for commercial use: Offering online services that are accessible and flexibly usable through APIs puts the SHIWA Community into a position that business models for these services can be relatively easily developed and implemented. Because all the user needs to go to these central services, the servers can provide service based on agreed SLAs (e.g. to certain IPs, during a certain time period, to an agreed number of workflows, workflow executions, etc.)

## Requirements for the SHIWA Repository

The Repository should be extended with an API and possibly with a client library and/or gadget to interact with the service. The API (Repository API) should expose the key repository functions:

* + Listing, filtered listing of workflows that are stored in the repository
  + Obtaining meta-data about a workflow stored in the repository
  + Uploading a new workflow with meta-data into the repository
  + Updating a workflow and/or its meta-data in the repository
  + Deleting a workflow from the repository

The API would enable third party clients (for example portals setup in SCI-BUS) to interact with the SHIWA Repository. The SHIWA community should remain responsible for the maintenance of the API and any libraries and gadgets coming with it. The SHIWA community must ensure that the library and gadgets are compatible with the Repository server and that whenever the API exposed by the server changes a new version of the library and the gadget are released and made available to user communities (including the ER-flow project).

In the long term, if there is demand, this concept can be expanded even further. The Repository service could allow user communities to request ‘Repository instances’ from the SHIWA Repository Server, which service would host multiple, community-specific virtual repositories. Repository instances should be logically separated from each other. Repository instances can be requested by community representatives and, after the setup of an instance, the owner (the community representative) can decide who can have user account on the instance. Those, who have accounts can use the instance to:

* Upload a workflow but keep it private (useful during workflow development and for archiving workflows)
* Upload a workflow and make it visible to other users of the same repository instance (useful to collaborate with members of the same community)
* Upload a workflow that is publicly visible (useful to collaborate with members of the European workflow community)
* Browse workflows from the repository (a workflow that is private to me, that is shared by a user of the same repository instance, a workflow shared by users of other repository instances)

1. Besides the SHIWA Portal and the SHIWA Desktop, the SHIWA Repository can be also accessed through its native web interface. However the SHIWA Portal includes every feature that the Repository native Web interface has so for the sake of simplicity ER-flow should promote only the SHIWA Portal, but not the Repository native interface. [↑](#footnote-ref-1)
2. The following workflow systems are already integrated with the portal: ASKALON, Galaxy, GWES, Kepler, MOTEUR, P-GRADE, PEGASUS, ProActive, Taverna, Triana. [↑](#footnote-ref-2)
3. SHIWA community = those who develop, operate, use or in any other way benefit from SHIWA services. The community is coordinated by MTA SZTAKI. [↑](#footnote-ref-3)
4. A web gadget is a small application that can be embedded into any third party web sites written in html. More about the topic and gadgets from the EGI Community are available at <http://go.egi.eu/gadgets>. [↑](#footnote-ref-4)