

succeed★

Interoperability framework and Taverna



Succeed is supported by the European Union under FP7-ICT and coordinated by Universidad de Alicante.

Interoperability framework

- The IMPACT Interoperability Framework (IIF) provides the interoperability layer for the software components and content developed in the IMPACT project.
- The technical and research partners in IMPACT developed more than 20 different tools for various stages in the OCR process.
- IMPACT has therefore also developed an overall technical framework that allows for a loose coupling of these tools and the exchange of data between them.



Interoperability framework

- With this framework it is possible to form a pipeline of the tools, where the output of one tool is used as input for the next tool.
- There is no optimal combination of tools (called a workflow) for every purpose.
- Users have to be enabled to try and evaluate certain combinations to find their optimal workflow.



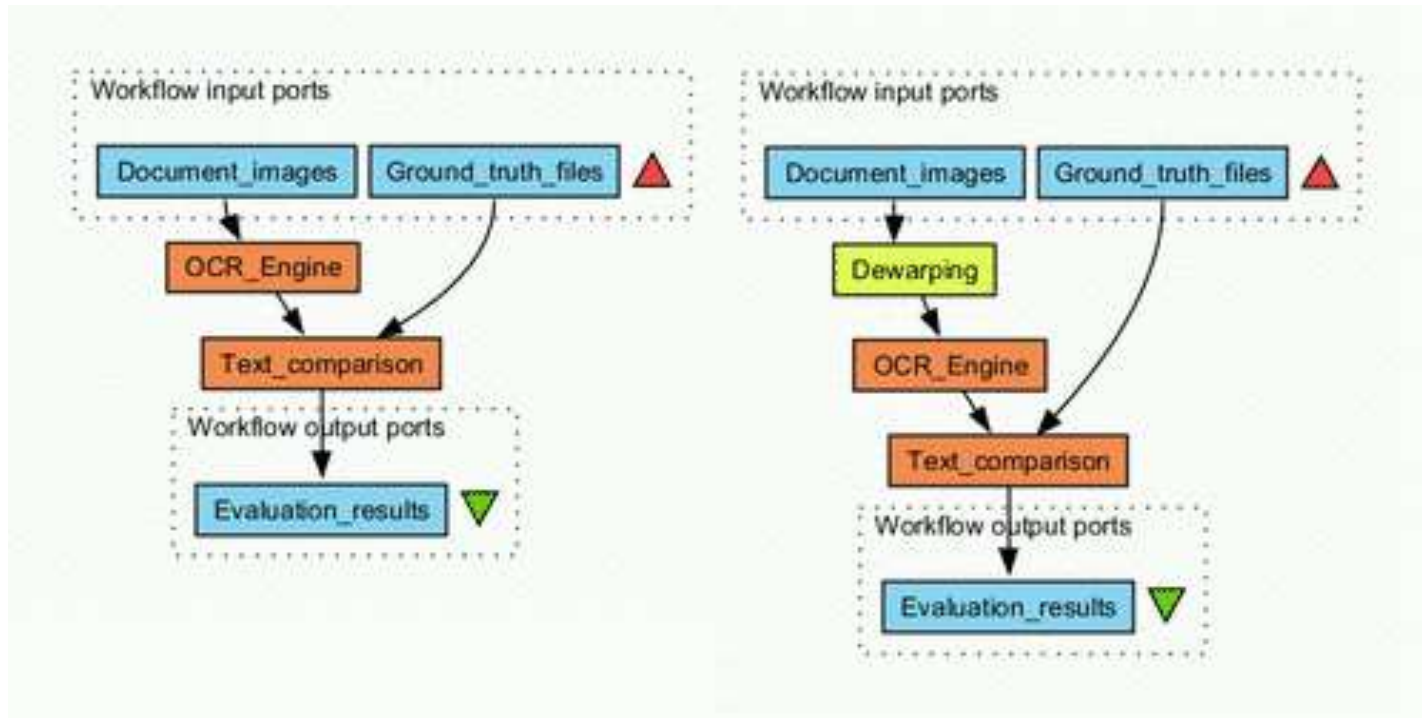
Interoperability framework

- To fulfil these requirements, each IMPACT tool was wrapped as a web service, using a generic Java-wrapper.
- All the web services together with common open source components from the [Apache Software Foundation](#) form the basic layer of the Interoperability Framework.



Workflows

- Each tool(web service) is represented as a workflow module for Taverna.



Taverna

- .A workflow language based on a dataflow model.
- .A graphical editing environment for that language.
- .An invocation system to run instances of that language on data supplied by a user of the system.
- .When you download it you get all this rolled into a single piece of desktop software



Taverna

.To easily create complex workflows from single tools, the Taverna workflow system has been added as another layer to the Interoperability Framework.

.[Taverna](#) is an open source and domain-independent suite of tools to design and execute workflows.



Taverna

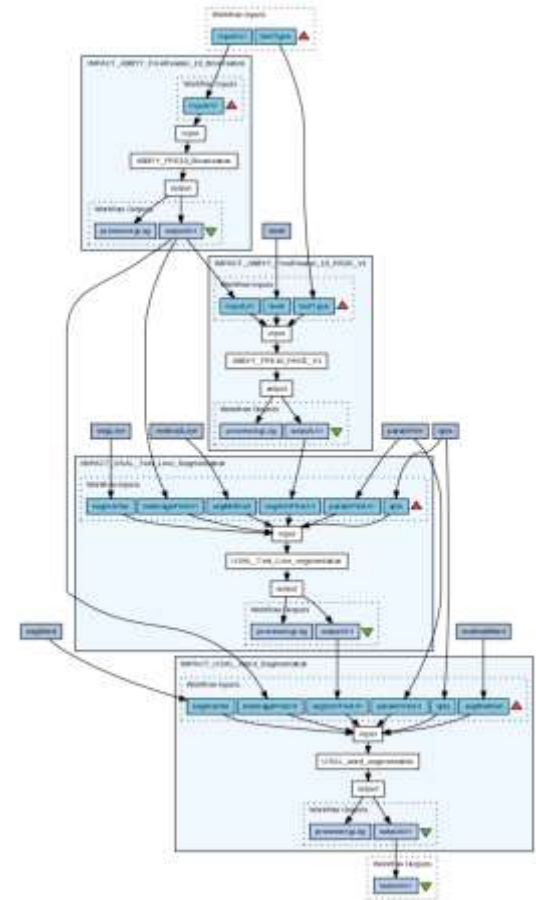
.The output of one tool can then be connected to the input of another by a simple drag-and-drop operation.

.Allows non-expert users to quickly create their own workflows and obtain results.



Workflows

- OCR workflow = data pipeline
- Building blocks = processing modules (nodes)
- Integration = interaction between nodes (mashups)



Community



- Ready-to-use and documented resources
- Community of experts
- Sharing of experiments and know how



The screenshot displays three workflow cards from the 'my experiment' community. Each card includes the following information:

- Workflow Title:** IMPACT IBM Concert UploadPagesUri BasicWorkflow (v1), Helper Transform XML using XSLT Basic Workflow (v1), and Abbyy OCR+Evaluation (v1).
- Metadata:** Created and last updated dates and times.
- Original Uploader:** Profile picture and name (Sivan, Clement, and Gerr).
- License:** Creative Commons Attribution-ShareAlike 3.0 Unported License.
- Workflow Diagram:** A small thumbnail image representing the workflow process.
- Instructions:** A yellow box containing specific instructions for using the workflow, such as 'Upload pages to the IBM Concert tool...' or 'The input XML document is converted using the input XSLT document...'
- Statistics:** Rating (e.g., 2.0 / 5), number of versions, reviews, comments, and citations.
- Usage:** Viewed and downloaded counts.
- Workflow:** A link to view the full workflow details.



Demo

