

Deliverable D5.3

Project Title:	World-wide E-infrastructure for structural biology	
Project Acronym:	West-Life	
Grant agreement no.:	675858	
Deliverable title:	Prototype of the new VRE portal functionality	
Lead Beneficiary:	LUNA	
Contractual delivery date:	Month 6	
Actual delivery date:	Month 6	
WP No.	5	
WP Title	Support	
WP leader:	Alexandre Bonvin	Partner: Utrecht University
Contributing partners:	Luna	

Deliverable written by Francois Ruty, Luna

1 Executive summary

The purpose of this deliverable is to set up a prototype of the VRE (Virtual Research Environment) within the project portal previously established in D5.1. The first version of the VRE is accessible at portal.west-life.eu (in the first instance, access is provided on demand).

The overarching goal of the VRE is to connect the many West-Life e-science portals to one another, as well as with academic storage repositories. Scientists must be able to use various WestLife e-science portals in an integrated manner, that is to say have a central location to manage all their data analysis and modelling. They must have the technical means to shuffle data between portals in an easy way, without for example having to transfer files via their own desktop over low-speed internet connections. Additionally, scientists shall have easy access to cloud resources, to scale and speed up cpu-intensive analyses or simulations. Thus, the aim is to provide a glue between existing academic services: West-Life e-science portals, academic storage repositories (e.g. EUDAT), and grid/cloud resources (e.g. EGI).

Three core features of the VRE are:

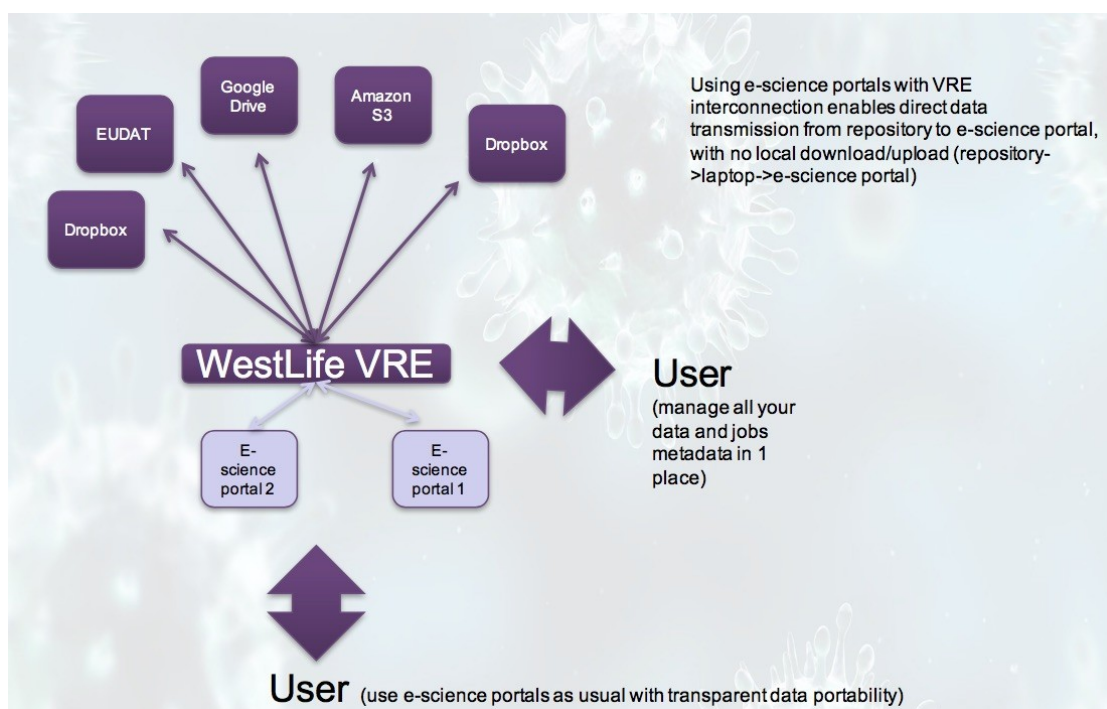
1. An aggregated view of all the scientist's data across personal and central repositories (EUDAT, Dropbox etc)
2. The ability to send job results directly from e-science portals to data repositories (no need to download/upload to a local desktop)
3. Storage of the metadata for all jobs submitted via the portal

These features are realised through specific data connections:

- The WestLife VRE connects to existing APIs of data repositories
- E-science portals connect to the WestLife VRE, using a simple procedure that minimises the disruption to existing services

The WestLife VRE shall provide:

- “Data location awareness” to e-science portals (where to fetch and dump data)
- Jobs metadata storage for e-science portals
- Aggregated jobs overview for end users
- Aggregated data overview for end users



This report describes a prototype for the West-Life VRE which addresses the above features. There is a particular focus on interconnection with academic storage repositories.

2 Project objectives

With this deliverable, the project has reached or the deliverable has contributed to the following objectives:

No.	Objective	Yes	No
1	Provide analysis solutions for the different Structural Biology approaches	X	
2	Provide automated pipelines to handle multi-technique datasets in an integrative manner	X	
3	Provide integrated data management for single and multi-technique projects, based on existing e-infrastructure	X	
4	Foster best practices, collaboration and training of end users	X	

3 Detailed report on the deliverable

The VRE prototype is currently focused on interconnection with academic data storage repositories. Closer interconnection with the e-science portals provided by West-Life partners is the next step. We chose to address academic storage interconnection first since most of them come with an existing public API, while closer integration with the e-science portals will require small updates from portal operators.

Status

The current VRE prototype can connect to the EUDAT API, Dropbox, Google Drive and Amazon S3, and enables end users to have an aggregated view of all their data across all repositories (see screenshots in the Appendix).

It is also possible to add further non-commercial storage. The current VRE prototype can connect to any API, and so West-Life consortium members can set up an object store (with APIs) on the hardware they contribute to the project, and the VRE can include those storage services in the aggregated view it provides.

Users have a settings section on the VRE where they can allow the VRE access to their academic repositories.

Users can log on the web portal and view a unified folder structure that sums up what they have in all the repositories that they have connected to their VRE account. They can upload and download files, and create folders in any of their repositories, from the VRE. They can create shareable links to share files with other people.

Hardware requirements

The current VRE runs on Ubuntu 14.04 and does not yet store data locally – future versions will store metadata. The hardware requirements to run the web portal are low: a few 100s MB of disk space, less than 1GB of RAM and 1 CPU thread are necessary.

The lightweight nature of the West-Life VRE portal is a key requirement to make sure running costs are as low as possible, to ensure deployment is easy, and to ensure sustainability beyond the end of the project.

Software Technology

For the web portal, we use Python/Django for server code, and AngularJS for JavaScript code. It connects to the EUDAT B2DROP service by the standard WebDAV interface, and it uses the suppliers' custom APIs to connect to DropBox, Google Drive, and Amazon S3. An abstraction layer unifies the four into a single API for access to data repositories.

Other VRE Features

Some work was carried out to design the architecture of code snippets that West-Life e-science portals operators will be asked to include in their portals, in order to interface with the VRE. This has been circulated to other partners for comments. Those code snippets will allow e-science portals to query the VRE and fetch data locations. The VRE will thus abstract away the various storage repositories and will provide a single interface to e-science portals.

We expect the VRE to provide this unified interface to testbed WestLife portals within 3 months.

Note: all source code (API code snippet, web application etc) will be publicly available to ensure that the WestLife portal lasts beyond the end of the H2020 project.

Interaction with other WPs

The work carried out here on an aggregated view of multiple data repositories is tightly coupled with deliverable D6.1 on providing the user with a virtual folder of distributed data. Work is ongoing to make sure the two efforts complement each other well.

There are also on-going discussions with WP4 to make sure the small upgrades required from portal operators, for interconnection with the VRE, do not conflict with the backend job submission architecture evolution.

Appendix 1: Screen shots

Data Management

What is visible for partners: (after private login url)

Pype

Login

Username*

Password*

[Forgotten password?](#)

Pype **Cloud Drive** Publish Data fruty ▾

Cloud Folders

You haven't linked any storage providers.
[Click here to add one.](#)

Select a folder to view the files it contains.



On the screenshot above you can link your WestLife VRE account with your storage repositories (EUDAT, Google Drive etc....). The linking procedure is adapted to each provider. For instance, if you attempt to link a Google Drive account, it will open a new page where you have to login on your google account and allow the VRE access to your drive.

Pyte Cloud Drive Publish Data test

Cloud Folders

- ▼ B2Drop 1
 - ▼ afolder
 - ▶ afolder2
 - ▼ anotherfolder
 - ▶ documents
 - ▶ music
 - ▶ photos
- ▼ gdrive
 - ▼ Luna
 - ▶ Missions
 - ▶ old
 - ▶ Projets
 - ▶ Software Q&A
 - ▶ Luna - business
 - ▶ METEO SPACE
 - ▼ Takeout
- ▼ My S3 account
 - ▶ Some folder

Takeout Refresh Upload File + New Folder









Filename	Uploaded	Size	State
Searches-20150513T213145Z.zip	Fri, Jan 8, 2016 4:41 PM	38.2 KB	Ok  

Pyte Cloud Drive Publish Data fruty

Cloud Folders

- ▼ my grive
 - ▼ Luna
 - ▶ Missions
 - ▶ old
 - ▶ Projets
 - ▶ Software Q&A
 - ▶ Luna - business
 - ▶ METEO SPACE
 - ▶ Takeout

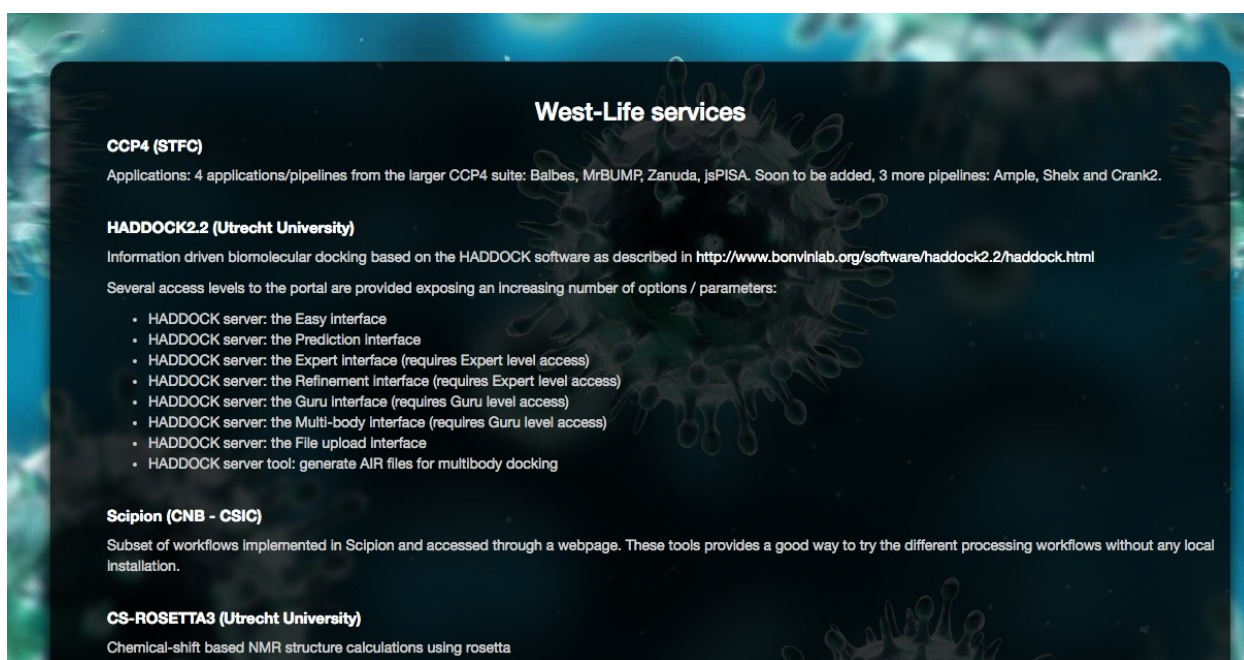
Luna Refresh Upload File + New Folder

Filename	Uploaded	Size	State
dst_last_96_hrs.txt	Tue, Jan 26, 2016 2:18 PM	16.3 KB	Ok  
Monthly Strategy	Tue, Jan 26, 2016 2:18 PM	N/A	Ok  
Shotgun_root.zip	Tue, Jan 26, 2016 2:18 PM	230 MB	Ok  
Weekly Strategy	Tue, Jan 26, 2016 2:18 PM	N/A	Ok  

On the 2 screenshots above, you can see an example of an aggregated view with Google Drive, EUDAT B2DROP and Amazon S3 accounts. Academic partners can set up further object storage. Most of those systems expose an API very similar to Amazon S3 so integration of custom academic storage services is easy.

Job Submit

WestLife portals are operational from day 1



West-Life services

CCP4 (STFC)
Applications: 4 applications/pipelines from the larger CCP4 suite: Balbes, MrBUMP, Zanuda, jsPISA. Soon to be added, 3 more pipelines: Ample, Shelx and Crank2.

HADDOCK2.2 (Utrecht University)
Information driven biomolecular docking based on the HADDOCK software as described in <http://www.bonvinlab.org/software/haddock2.2/haddock.html>
Several access levels to the portal are provided exposing an increasing number of options / parameters:

- HADDOCK server: the Easy Interface
- HADDOCK server: the Prediction interface
- HADDOCK server: the Expert interface (requires Expert level access)
- HADDOCK server: the Refinement interface (requires Expert level access)
- HADDOCK server: the Guru interface (requires Guru level access)
- HADDOCK server: the Multi-body interface (requires Guru level access)
- HADDOCK server: the File upload interface
- HADDOCK server tool: generate AIR files for multibody docking

Scipion (CNB - CSIC)
Subset of workflows implemented in Scipion and accessed through a webpage. These tools provides a good way to try the different processing workflows without any local installation.

CS-ROSETTA3 (Utrecht University)
Chemical-shift based NMR structure calculations using rosetta

HADDOCK2.2
WeNMR/West-Life GRID-enabled web portal

WeNMR home NMR services SAXS services HADDOCK tutorials WeNMR Support Center RSS

WELCOME TO THE WENMR WEB PORTAL >>

HADDOCK (High Ambiguity Driven protein-protein DOCKing) is an information-driven flexible docking approach for the modeling of biomolecular complexes. HADDOCK distinguishes itself from ab-initio docking methods in the fact that it encodes information from identified or predicted protein interfaces in ambiguous interaction restraints (AIRs) to drive the docking process. HADDOCK can deal with a large class of modeling problems including protein-protein, protein-nucleic acids and protein-ligand complexes.

More information about HADDOCK2.2 can be found on the HADDOCK2.2 website

Read also what an independent review by Moreira *et al.* has to say about our software...

HADDOCK is one of the flagship software in the EU H2020 BioExcel Center of Excellence for Biomolecular Research.

HADDOCK WEBSERVER

REGISTRATION: The use of the HADDOCK WeNMR GRID-enabled docking server is free for academic users. Access to the server is managed through Single Sign On (SSO) authentication using your WeNMR account. Old style HADDOCK web server accounts are still supported. How to proceed:

PROFILE >>

we-nmr

West-Life

bioexcel
Center of Excellence for Computational Biomolecular Research

e-infrastructure

SERVICES

The WeNMR web portal is an easy

WestLife portals are already operational and available from the West-Life portal, but they are not yet integrated into the VRE. The VRE currently only provide an aggregated view of all data. In the near future, users will go to their usual e-science portals, and instead of having to upload data on those web interfaces, they will be able to select data from the VRE aggregated view, directly in the e-science portals interfaces.

Scipion Web Tools - West-Life

Single particle analysis tools

Create your initial volume

My first map

Align your movies

My movie alignment

Analyze your maps

My resolution map

© 2015 · Scipion is licensed under the GNU General Public License

home >> AMPS-NMR



AMPS-NMR

(Including paramagnetic restraints plugin)

WeNMR GRID-enabled web portal

WeNMR home NMR services SAXS services WeNMR Support Center

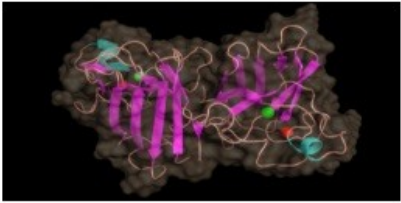
WELCOME TO AMPS-NMR WEB PORTAL PROFILE >>

Supported browser:   RSS  ATOM

AMBER-BASED PORTAL SERVER FOR NMR STRUCTURES (AMPS-NMR)

Amber (acronym to Assisted Model Building with Energy Refinement) is a suite of programs that allow users to perform molecular dynamics (MD) simulations on biological systems.

This web portal makes available the entire functionality of AMBER, in particular (but not only) using NMR-derived information as restraints for MD.



To use AMPS-NMR you have to register to WeNMR grid infrastructure

You can access a trial version of the service using username *quest* and password *quest*

we-nmr
West-Life
e-infrastructure

: Access to AMPS-NMR :

Username:

Password:

Forgot your password?
New to AMPS-NMR? Sign up.
If you don't have credentials but