Final Report from West-Life Part B

1. Explanation of the work carried out by the beneficiaries and Overview of the progress

1.1 Objectives

The overarching objective of this project is to bring the world of complex data analysis in Structural Biology to a simple Web browser-based Virtual Research Environment (VRE), available to any laboratory involved in the experimental structural characterization of biomolecules and their complexes and assemblies.

Structural biology is the discipline that aims to elucidate the molecular basis of life. It contributes to the discovery of new therapeutics in many ways. For example, Latinović et al. determined the structure of a component of nose-horned viper venom, potentially leading to safer anticoagulants, using West-Life computational service for the work (doi: 10.1055/s-0038-1669785).

Structural biologists increasingly investigate large multi-state macromolecular machines rather than single gene products. They are typically experts in one or a few experimental techniques, for example Macromolecular X-ray crystallography (MX), Nuclear Magnetic Resonance (NMR), or cryo-Electron Microscopy (cryo-EM), but a full characterisation of large structures often requires a combination of several structural and computational methods. Instruct-ERIC, the European infrastructure for structural biology, encourages such work by facilitating visits to experimental facilities, and providing training to researchers learning new techniques. West-Life, a European Union Horizon 2020 e-Infrastructure Virtual Research Environment, facilitated the use of the associated computational methods.

The information technology infrastructure that supports these researchers must keep pace with these new demands. It must provide implementations of algorithms that can integrate data from different methods, and help track the processing steps and data in these increasingly complex projects. This e-infrastructure is provided as user-friendly web services. While for example an active crystallographer may find it worthwhile to invest time installing the CCP4 suite, a researcher using a particular technique for the first time may experience software installation as an overhead cost, and it is one that can be minimised.
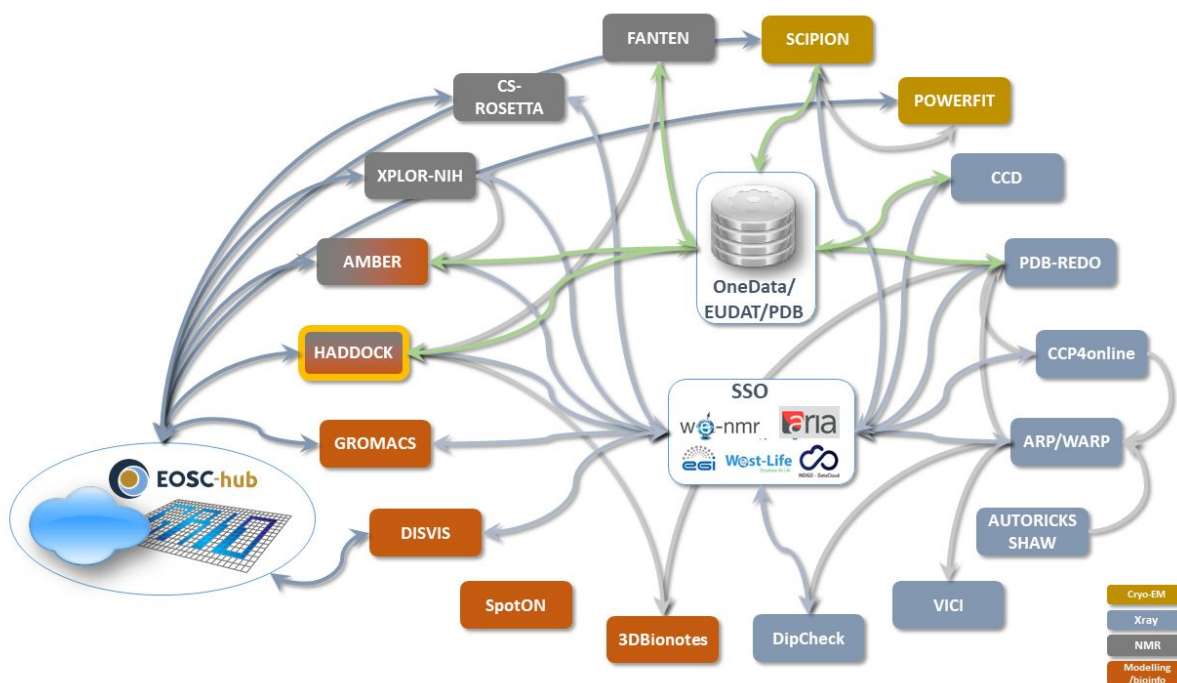
Figure 1.1.1

Objective 1: Provide analysis solutions for the different Structural Biology approaches

The West-Life project has enabled improvements to the following web services for structure determination: AMPS-NMR, ARP/wARP, DISVIS and PowerFit, FANTEN, HADDOCK, MetalPDB, PDB-REDO, ProteinCCD, and Scipion. These improvements cover extensions to handle different experimental data, changes to underlying algorithms, changes to improve overall speed, and modifications to user interfaces to aid usability. They include unifying changes such as the implementation of Single Sign On by individual services.

West-Life has also participated in the development of new services: DipCheck for validation of protein backbone geometry, PRODIGY and PRODIGY-LIG for the prediction of binding affinities for protein-protein and protein-ligand complexes, SpotOn for the identification of hot-spot residues in protein complexes, and 3DBIONOTES for annotating structures with biochemical and biomedical information. Major new versions of ARP/wARP (D7.4) and HADDOCK (D7.5) were released, to support building atomic models into cryoEM maps and protein-protein docking respectively.

The West-Life Portal https://west-life.eu provides links to these services. Comparing the statistics and their associated KPIs over the three years of the project, we can observe a sustained growth in both users and usage for the majority of services, which translates into a

sustained overall 25-40% increase rate per year (D5.8). The report on WP5 below gives further details.

The consolidated platform of West-Life leverages the resources provided by the EGI, which is currently supported with public European funds through EOSC-Hub. West-Life users get access to the EGI HTC e-infrastructure via the enmr.eu Virtual Organisation (VO). During 2018 EGI.eu and the enmr.eu VO have renewed the Service Level Agreement (SLA) signed in 2016, granting to enmr.eu VO until 31/12/2020 an amount of opportunistic computing time up to 53 Million of normalized CPU hours and opportunistic storage capacity up to 54 TB. More than 5.2 Million jobs (4.1 Million in the RP1) have run in RP2, translating into 30.7 Million (26.5 Million in RP1) normalized CPU hours consumed by the West-Life applications. In addition, four of the EGI Federated Cloud centres support the VO, and West-Life applications (mainly by Scipion and Gromacs) executed in RP2 more than 1,791 VMs (2,554 in RP1) and 175,290 CPU wall time hours (555,239 in RP1) (D4.5 and D4.6).

Objective 2: Provide automated pipelines to handle multi-technique datasets in an integrative manner

The project has facilitated the sequential use of multiple web services by building direct connections between web sites. These are shown in figure 1.1.1 above. For example, CCP4Online provides Balbes, MrBUMP and MoRDa services for solving crystal structures via molecular replacement. In each case, the output coordinate model should be correctly placed in the crystallographic unit cell, but as a first model may be insufficiently accurate for subsequent standard refinement to proceed. West-Life pipelines enable this partial solution to be sent to the ARP/wARP or PDB-REDO services for further building and refinement. PDB-REDO links in turn to the 3DBIONOTES visualization service. In 2018, this link was followed more than 1500 times. Construction of such pipelines varies with the specific services involved, but has typically required the development (or improvement) of programmatic access to allow the seamless transfer of data and metadata from one service to another. The adoption of Single Sign On mechanisms by services is also crucial, allowing the pipeline to operate without the researcher having to login multiple times. Since many services are computationally intensive and may require hours to complete, the ability to transfer data autonomously is very important. More details are provided in the WP4 report.

In addition to direct connections, pipelines may also be supported by passing data through centralised data stores. INFN has implemented the necessary support to West-Life SSO as an authentication method for Onedata. INFN and STFC worked on the integration of the West-Life Virtual Folder with Onedata, but the integration will only be "production ready" when a final release of Onedata is available.

Objective 3: Provide integrated data management for single and multi-technique projects, based on existing e-infrastructure

Typically, a structural biology project accumulates a large number of data files associated with different instruments and different pieces of software. Managing these data throughout the project can be difficult, and returning to a project (or inheriting someone else's project) at a later date even more so. West-Life has attempted to alleviate these potential problems through the development of two new facilities for data management.

The West-Life Virtual Folder addresses the issue of accessing data scattered among different data storage providers. It is a web-based application allowing a user to register data storage providers and connect accounts from multiple places in one common interface. The data are not physically aggregated, since many datasets can be large. Instead, the application allows a user to view all data and download selected data on request, or to get a publicly accessible link (conforming to the WEBDAV protocol) to either a particular file or a particular folder which can then be used by another tool to download, process and upload results. Currently Dropbox and WebDAV based data storage providers are supported (e.g. EUDAT B2DROP service, owncloud instances with WebDAV enabled or Pcloud commercial services) and connections to institutional repositories have also been demonstrated.

Secondly, we have developed a Repository for experimental data, which is installable locally or on cloud provision (D6.2). Large facilities already have data management services, for example ISPYB at several European synchrotrons. However at smaller facilities it is all too common that the user simply returns home with their data on a USB stick.

We have also facilitated the reuse of structural data by improvements to the search facilities of the PDBe (D5.5, D6.3) and text mining facilities (D7.8). The W3C standard PROV-O is designed to support the recording and sharing of provenance information. We have developed mechanisms for automatically saving this information, storing it, and viewing it, and these have been implemented in the Virtual Folder (D6.4).

Objective 4: Foster best practices, collaboration and training of end users

We organised three training events during Reporting Period 2 (D2.5), and participated in 41 workshops or conferences organised by others.

These interactions with other projects have been reflected in published journal articles which acknowledge multiple funding sources. West-Life has been acknowledged jointly with the following European projects:

- BioExcel grant no. 675728
- Bio-NMR no. 261863,
- CORBEL (INFRADEV-1-2014-1—Proposal: 654248),
- East-NMR no. 228461,
- EGI-Engage, grant no. 654142;
- ELIXIR-EXCELERATE (INFRADEV-1-2015-1—Proposal: 676559)
- INDIGO-DataCloud, grant no. 653549;
- iNEXT (project No. 653706)
- Instruct-ERIC, a landmark ESFRI project
- Instruct through grant reference RID:7 (Pilot EM Cloud Computing).
- Marie Skłodowska-Curie Individual Fellowships MSCA-IF-2015, MEMBRANEPROT 659826, and BAP-659025
- WeNMR (project# 261572)

Publications acknowledging West-Life also acknowledge 34 different national awards.

The web site https://west-life.eu is a centre for the dissemination of the work, and the @WestLifeSB twitter account has 231 followers. Webinars and lectures, collected by the WeNMR project, on the tools and services available to the users through West-Life are uploaded to the WeNMR YouTube Channel (https://www.youtube.com/c/WeNMRWestLifeVRC).

1.2 Explanation of the work for each work package.

1.2.1 WP1: Management

1.2.1.1 Summary and significant results

This work package is concerned with management of the West-Life project, communication and reporting. The project has completed successfully.

1.2.1.2 Objectives and tasks

The objectives of the management work package are closely aligned to the following tasks:

- Task 1.1: To project manage West-Life during Years 1-3
- Task 1.2: To manage project meetings including the Annual General Meetings for all partners, the Executive Steering Committee meetings and all other governance and operational meetings for the project.
- Task 1.3: To manage communication channels between partners, the Coordinator and the EC project officer, including the provision and operation of a project website. To manage collaborative relationships with industrial partners and third parties.
- Task 1.4: To manage all reporting to the Commission including financial and WP reports.
- Task 1.5: To develop a sustainability plan for the VRE that will see its broad adoption to service a user community spanning from the public to specialist.

1.2.1.3 Work carried out in P2

Task 1.1 Overall coordination of the project

Coordination of the work is achieved through the committees as described below. Also there are mailing lists provided by partner MU for technical discussion (http://internal-wiki.west-life.eu/index.php/Mailing_lists), a second mailing list for administration (https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=WEST-LIFE-ADMIN), and a wiki provided by partner LUNA (http://internal-wiki.west-life.eu/index.php).

Task 1.2 Formation and management of project boards and committees

The kick off meeting in January 2016 appointed an Executive Steering Committee chaired by Alexandre Bonvin (Utrecht University) and consisting of Martyn Winn (STFC), Ales Krenek (MU), Jose Maria Carazo (CSIC), and Lucia Banci (CERM). In RP2 this met in May 2017.

All work package leaders participate in a monthly Operations Board telephone conference to coordinate the work.

A small Project Management Board also meets monthly between meetings of the Operations Board.

The Scientific Advisory Board consists of Arwen Pearson (CFEL), Erwin Laure (KTH, BioExcel), and Piotr Sliz (Harvard, SBGrid). They attended the second All Partners' Meeting 19-20 Jan 2017, and Erwin Laure and Piotr Sliz attended the third All Partners Meeting 3-4 May 2018.

Task 1.3 Communication

Project members have participated in conferences and training workshops of the structural biology community, and of e-infrastructure providers, as reported in the WP2 report below. Outreach is also supported by the project web site http://west-life.eu.

Task 1.4 Reporting to the Commission

We reported via the web site SyGMA, and at the Periodic Review. All deliverables are also available via the project web site at https://about.west-life.eu/network/west-life/about/work-packages/deliverables.

During P1 PI Martyn Winn, Project Manager Chris Morris, and Antonio Rosato met with the Project Officer Victoria Tsoukala on 5th October 2016 to provide an informal update on the West-Life project.

Some amendments to the grant agreement were submitted as described in 1.2.1.5 below.

Task 1.5 Development of a sustainability plan

A sustainability plan was reported in D1.5. Subsequently, Instruct-ERIC has formed a Working Group on Computational Services, which held its first meeting on 20th November. This will take a strategic overview of computational services for structural biology, in the context of Instruct's integrative approach to combined and hybrid methods, thus continuing the mission of West-Life.

Other news is that 3DBIONOTES has become a "Recommended Interoperability Resource" of ELIXIR. The immediate impact in the West-Life context is that this web service, that is currently part of several W-L workflows, is susceptible to apply for additional funds in ELIXIR for future developments, assuring its sustainability. Additionally, much of the exploratory work done in W-L between partner CSIC and MU for the automatic deployment of Scipion (cryo EM) in the cloud has been included in a recently submitted proposal, "EOSC-Synergy", as work to be taken to the next stage of development, so as to allow generalized user access in EOSC (although with eligibility requirements to access a certain level of computational resources).

1.2.1.4 Progress towards Deliverables and Milestones

	Title	Lead	Type	planned	Progress
D1.1	Kick off meeting	STFC	Public report		RP1
D1.2	Project web pages	Instruct	Other public		RP1
D1.3	Periodic report	STFC	Public report		RP1
D1.4	Draft sustainability work plan	Instruct	Report		RP1
D1.5	Sustainability report	Instruct	Public report	Nov 2017	Delivered
D1.6	Final project conference	STFC	Public report	Oct 2017	Delivered, see 1.2.1.5 below
D1.7	Final Report	STFC	Public Report	Dec 2017	This document

Milestone	Date	Means of Verification
MS1 Kick off meeting	RP1	Report disseminated
MS2 Website launched	RP1	
MS3 First periodic report	RP1	Delivered
MS4 Sustainability report approved by Executive Steering Committee	12 May 2017	https://zenodo.org/record/1249372#.Wv6px1VKhhE

1.2.1.5 Deviations from Annex 1

The dates of a few deliverables and milestones required adjustment for reasons reported elsewhere in this document.

We decided that it would be more effective to hold the West-Life conference in May 2018, rather than in the final month of the project, so it can plan achievements rather than merely celebrate them. In particular, this enabled us to get guidance from our Scientific Advisory Board in time to act upon it.

An amendment was submitted in August 2017, for a change of bank account. Another amendment was submitted in December 2017, and the following changes were agreed:

- Delays to three deliveries, because we could see opportunities to improve them if more time is allowed.
- Two partners had an increase in person months applied, with no change in budget, because they were able to assign more junior staff than originally envisaged.
- CIRMMP is contributing 3PM to D6.2. This was funded by a transfer from STFC, which suffered delays in hiring.
- INFN has transferred 30,000 Euros from travel expenses to labour, to increase their contribution to WP4.

An additional amendment was agreed during RP1.

1.2.2 WP2: Dissemination, Training and Outreach

1.2.2.1 Summary and significant results

 This work package aimed to disseminate the achievements of West-Life within and outside the existing structural biology community; with a special focus on outreach to training and engaging with new user communities and industry. Over the last reporting period of 18 months, the work package has organised 5 networking and 4 training events collaborating with other projects and networks as appropriate. A social media presence continued to grow and West-Life has been represented in a wide variety of platforms publicising the contribution of West-Life VRE to the scientific communities around the world. In particular, the work carried out in the project was presented at more than 70 international conferences around the world with 42 peer review publications referencing the project.

1.2.2.2 Objectives and Tasks

Objective 2.1: Training new and existing communities of users in the tools, services and methods offered by the West-Life VRE to achieve high impact research outcomes

Objective 2.2: Enabling developers to provide new solutions for currently intractable problems that limit integrative approaches in structural biology

Objective 2.3: Engaging all sectors from the public to research specialists and communicating the power of a virtual research environment to provide a seamless resource for the management, analysis, interpretation, archiving and protection of data acquired from structural biology methods.

Task 2.1 – Training programmes: user communities (Instruct, all).

This task addresses Objective 2.1 and will implemented a series of training events. Two training events in each of years 1, 2 and 3 will be tailored to the existing user structural biology community and provide information and practical experience of data handling, including the availability, access and use of existing tools and the data outcomes. The user workshops will interface with existing Instruct Training courses, which provide practical training on new and integrative structural biology technologies and will be extended to include best practices in data acquisition and handling at the research 'bench'. One training event per year (in years 1 - 3) will deal with archiving data, existing common data services (EUDAT http://www.eudat.eu), charters (Access charter). A programme for training components of West-Life will be developed by month 9, incorporating shared training events with stakeholders where relevant (M2.1).

Training events may include a virtual web-based component and face-to-face meetings. These will combine user groups and include a 'user clinic' at which practical management of problems, outstanding requirements and possible solutions can be discussed.

Task 2.2 – Training programme: developers (Instruct, all).

This task addresses Objective 2.2 by implementing a programme of training events for developers, which will build partly on the feedback from the user communities on problems and gaps in integrative resources and partly on the known issues in developing a full suite of integrated services. More specifically, these training events aimed at developers will focus on the solutions and on the presentation of the opportunities that West-Life has deployed and made available to the entire community of software developers and providers of computational services. By participating in the present training events, developers both within and outside the present partnership will have access to expertise, information, and guidelines to make their tools available via the West-Life platform and/or to exploit the computational resources gathered within the present project. This may lead to an expansion of services available to our users beyond the tools created by the partners.
The programme will implement at least three training workshops (one per year).

Task 2.3 – Dissemination and outreach (Instruct, all).

A key part of delivering services to the community at large is to promote understanding and importance of structural biology by providing information on the value of the research being supported, the resources underpinning the service and to share knowledge and experiences arising from the project with a broad stakeholder community. In specific terms, this task addresses Objective 2.3 and is devoted to developing and strengthening interactions with existing and new communities bridging the academic, industry, policy and decision makers and public sectors.
Implementation of this task will address the three target groups separately and also bring all parties together for the Kick-off meeting and the end of project meeting."

Task 2.1 – Training programmes for user communities

In the final reporting period of 18 months, West-Life organised 4 different training events targeting user communities in Europe and worldwide, when the workshop was organised in collaboration the partners are indicated. A summary of the training events is presented as a table below:

Title	Organiser	No. of Participants	Date
Hands-on in cryo Electron Microscopy image processing. São Paulo School of Advanced Science on Biophysical Methods to Study Biomolecular Interactions 2017	West-Life	80	15 – 19 Oct, 2017
Practical workshop on Cryo EM data analysis in Scipion	West-Life/MU/ CSIC	20	25-26 Jan 2018
INSTRUCT course on "Advanced methods for the integration of diverse structural data"	West-Life/Inst ruct - University of Florence	20	Feb 19 - 23, 2018
Webinar: "West-Life: Lessons learnt from developing a virtual research environment"	West-Life/Cor bel/Instruct	Online	23 Oct 2018

The workshop participants represented both members from existing structural biology communities and new user communities. The events were focused on providing practical knowledge of using scientific techniques and also, using tools and services made available through West-Life, including image processing and data analysis of structural data.

West-Life training contributions are not just restricted to events organised by the project. Scientists from the project partners conduct sessions at workshops or courses organised by collaborators. These sessions are aimed at training users in specific methodologies/tools/services enabled through West-Life.

There were 5 developer training and networking events organised during the last 18 months of the project. The objectives of the training workshops included sharing expertise within and outside the West-Life partners including building bridges with tools/services provided by eInfrastructures such as EUDAT (https://eudat.eu/services) and EGI (https://www.egi.eu). Courses also included approaches of cloud infrastructure setup and using software for distributed computing. A summary of the courses enabled through West-Life is below:

Title	Organiser	Type	No. of Participants	Date
West-Life iNEXT round table. Satellite meeting to Instruct Biennial Structural Biology Conference 2017	West-Life/ Instruct	Meeting	100	24 May 2017
West-Life all partners meeting	West-Life/ CSIC	Meeting	27	3-5 May 2018
West-Life Hackathon	West-Life/ Daresbury Laboratory	Workshop	4	13-14 Mar 2018
West-Life Hackathon	West-Life/ Instruct	Meeting	7	26 Sep 2018
Webinar: "West-Life: Lessons learnt from developing a virtual research environment"	WestLife/ Corbel/ Instruct	Webinar	Online	October 2018

West-Life also was represented at developer conferences such as DI4R along with programmers and services providers from various eInfrastructures and industrial stakeholders. These events presented opportunity to partnership developers to gain information and expertise in scientific developments and also to disseminate information on access and availability of tools and services that West-Life partners provide.

As agreed in the previous report one of the training workshops was a web-based training event. The webinar was aimed to both developers and users of West-Life.

Task 2.3 – Dissemination and outreach

During the duration of the project the website was kept current with up-to-date information and dynamic content as a way to engage with user communities. The West-Life website (https://west-life.eu/) was set up in collaboration with Instruct to share news and events which increased the reach to thousands of Instruct users (https://www.structuralbiology.eu/) in addition to disseminating important structural biology information to West-Life users. The website also has Twitter widget engaging with different user communities, partners and collaborators. The @WestLifeSB twitter account has 230 followers and has more than 1000 tweets engaging with policy makers, partners, collaborators and users, averaging excess of 100 profile views per month. Owing to the emergence of Twitter as a support platform, with specific instance of such support request in West-Life, the website also encourages users to contact West-Life using the twitter handle. Webinars and lecture, collected by WeNMR project, on the tools and services available to the users through West-Life are uploaded to the WeNMR YouTube Channel. More lectures from workshops will be added to this channel for further dissemination. We have also collaborated with Instruct and Corbel with a webinar to reach the extended biomedical science community.

While the tools and services provided through West-Life have their own established portals and support mechanisms, the forums and associated support pages in West-Life website will direct users to the correct help centre.

Work Package 2 produced the West-Life flyer after consultation with partners and it is made available for the everyone in the project to circulate among collaborators and at events. The flyer contents and the design were revisited through consultation with new user communities including public engagement expert, to make sure that the message and the text remain relevant to not just structural biology community but to a wider audience. The work package also produced template materials for presentations and posters for use by partners when representing the project at difference events. The flyer was also distributed to all Instruct training and outreach events reaching the extended structural biology community.

As an e-Infrastructure project, West-Life also benefited from continued presence at conferences and workshops for better dissemination of information about the project and the VRE. Partners of West-Life presented posters at different conferences and cited West-Life at more than 70 lectures at conferences worldwide (D2.5 https://zenodo.org/record/1475868). West-Life was also acknowledged in 43 published journal articles.

Project partners were continually encouraged to use West-Life branded templates for presentations and posters, use West-Life logos, namecheck other partners and distribute business cards when attending networking events.

West-Life also co-organised training and sustainability workshops along with Instruct to ensure that the results achieved are not lost to the community and will continue to provide service to the users. This is described further in the sustainability report but from the dissemination point of view is worth mention that the tools now available through the West-Life website will continue to be available through the Instruct-ERIC website.

Engaging with new user communities like Biomedical Sciences Research Infrastructures was seen as critical, particularly when structural biology data is becoming relevant and necessary for translational research. As part of this commitment, a foresight meeting from the milestone MS7 was co-organised with the round table meeting in coordination with WP3 as part of deliverable D3.2. The round table took place at Brno on May 24th 2017, as a special session of the 2nd iNEXT Annual User Meeting in Brno, just preceding the 3rd Biennial Scientific Conference of Instruct (Brno, May 24-26). More details about the round table are described in WP3 report. West-Life also continued to interact and engage with a varied variety of project and initiatives and WP3 reports on the achievements made in networking with these communities and stakeholder groups.

1.2.2.4 Deliverables and Milestones

	Title	Lead	Type	Planned	Progress
D2.1	Interim report on T2.1	Instruct	Public report		RP1
D2.2	Interim report on T2.2	Instruct	Public report		RP1
D2.3	Summary report of SB community engagement	Instruct	Public report		RP1
D2.4	Report on industrial user engagement	Instruct	Public report	Oct 2017	Delivered
D2.5	Engagement report	Instruct	Public report	Oct 2018	Delivered

Milestone	Date	Means of Verification
MS5 Training Programme Plan	RP1	Schedule for training programmes years 1-3 approved and published on the website
MS6 Engagement Plan	RP1	Outline plan and schedule for Engagement approved and published on the website
MS7 Foresight Meeting	RP1	Event information available from the website

1.2.2.5 Deviations from Annex 1

Instruct contributed extra effort to WP2, notably to enhancing the project web site.

1.2.3 WP3: Networking

The overall goal of this WP is to maximise communication and identify opportunities for collaboration with European and global infrastructures that may leverage Structural Biology data as well as with European and global initiatives addressing scientific data, such as the Research Data Alliance and other e-Infrastructures. A further objective of this WP is to interact with organizations involved in the development of science policy, including national and European scientific academies, funding bodies and associations thereof.

In practice, we aim to evaluate the needs expressed by regional, national, European or global infrastructures not directly involved in the present partnership, with regard to the availability and use of structural biology data. In this respect, Instruct has a central role by leveraging its close relationship with the other BMS RIs to foster the interaction between the large ESFRI community and West-Life. For interaction with data initiatives, a crucial role is that of the *Interest Group in Structural Biology*, chaired by three project partners (Lucia Banci, Chris Morris and Antonio Rosato), within the Research Data Alliance (RDA). The RDA is a global forum that comprehensively addresses all aspects of data science.

1.2.3.1 Summary and significant results

In the second reporting period, this WP extended its interaction with ESFRI infrastructures to define the applications as well as current and foreseeable use of structural biology data in the various domains of the biomedical sciences. In particular, one document on "The usage of structural data in different biomedical RI's" was co-authored with the CORBEL partnership (D3.5).

A follow-up event of the Round Table that was co-organized by iNEXT and West-Life in the frame of the 2nd iNEXT Annual Meeting (Brno, Czech Republic, May 2017) took place on the occasion of the second annual meeting of the CORBEL project. This second meeting was entitled "Visibility, usage and impact of research infrastructures - follow up of the Brno Round Table" and focused on the (re)evaluation of the main items outlined and the actions already initiated following the first RT. Eight different ESFRIs attended with their delegates.

The model of tight connection between the experimental Research Infrastructure (INSTRUCT) and its associated e-infrastructure (West-Life) developed in this project was crucial to define the design of the new application EOSC-Life which 13 biomedical infrastructures ((BBMRI, EATRIS, ECRIN, ELIXIR, EMBRC, EMPHASIS, ERINHA, EuBI, EU-OPENSCREEN, INFRAFRONTIER, INSTRUCT, ISBE, MIRRI) successfully submitted in response to the INFRAEOSC-04-2018 call.

The latter project will start on March 1ˢᵗ, 2019 and will develop common e‑infrastructure solutions for all biomedical RIs.

Networking activities carried out in P2 confirmed that enhancing the visibility and fostering usage of RI's, both physical and electronic, remains a high priority of all the stakeholders in the field. This is even more true in the context of the EOSC, because RIs are in the best position to guarantee the quality of data that will be available to the community. In other words, high-quality data provision to the scientific community via EOSC is a clear mandate of large RIs, particularly ESFRIs. This role of RIs must be disseminated to all biomedical researchers, in a cross disciplinary manner, so that researchers will aim to access the appropriate RIs or to use RI data especially when in activities that go beyond the borders of their own scientific subdomain. Therefore, RIs are teaming up both to promote the dissemination of infrastructure services and to encourage national and European funders to implement mechanisms fostering usage of RIs by funded projects.

1.2.3.2 Objectives and Tasks

The detailed objectives of this WP, as specified in the original application, are:

Objective 3.1: identify opportunities for collaboration with European and global infrastructures that may leverage Structural Biology data to enrich/extend its impact on the society at large

Objective 3.2: identify opportunities for collaboration with European and global initiatives addressing scientific data and other e-Infrastructures

Objective 3.3: interact with organizations involved in the development of science policy

The above objectives were be addressed through the following tasks:

Task 3.1 – Interaction with large infrastructures and European projects (CIRMMP, all). This task directly addressed Objective 3.1 by surveying and reporting on requirements and needs expressed by European or global infrastructures not involved in the present partnership, both physical and electronic, with regard to the availability and use of structural biology data.

Task 3.2 – Interaction with data initiatives (STFC, all). This task directly addressed Objective 3.2 by focusing on the involvement of project representatives in European and global fora related to the various aspects of the life cycle of scientific data. From the very beginning of the project, we will leverage the existence of an Interest Group in Structural Biology within the Research Data Alliance. The activities of this involved the participation of West-Life partners at relevant events with the aim of identifying synergies and opportunities for a more thorough exploitation of structural data.

Task 3.3 – Interaction with initiatives developing science policy (CIRMMP, all). This task directly addressed Objective 3.3. This task overviewed the development of science policies affecting the field of SB in general. We also evaluated opportunities to establish collaborations with or to bring specific issues/opportunities to the attention of science policy makers or funding bodies. Various project partners are already involved in similar endeavour with a series of

European Institutions and Consortia which identify needs of the various scientific communities and develop strategies to address new directions and challenges for research and innovation. This Task will constitute a venue where to compare the corresponding outcomes and develop synergistic strategies.

1.2.3.3 Work carried out in Period 2

Networking activities carried out in P2 allowed furthering and strengthening the engagement with other ESFRI RIs and other European projects. All Deliverables foreseen for this reporting period were met in a timely way. In particular, D3.5 consisted of a document on the usage of structural data in different biomedical RI's that was published jointly with the CORBEL partnership (https://doi.org/10.5281/zenodo.1034828). The Round Table held in Brno on May 24th was the starting point for the definition of a working group and for the identification of requirements regarding structural biology data shared by different infrastructures (MS10). Activities related both to the CORBEL and the recently approved EOSC-Life project (in response to the INFRAEOSC-04-2018 call) ensure that the engagement with biomedical ESFRI RIs will continue also beyond the end of the West-Life.

Task 3.1 – Interaction with large infrastructures and European projects

A round table involving several ESFRI RIs representatives took place at the very beginning of P2, took place in Brno, Czech Republic, on May 24th, 2017 (https://www.structuralbiology.eu/content/bringing-together-the-bio-medical-scientific-communities-the-role-of-research-infrastructures). The panelists included representatives of several BMS RIs and H2020 health related projects:
•   Lucia Banci (iNEXT, chair)
•   Serena Battaglia (ECRIN)
•   Steve Brewer (Edison)
•   Susan Daenke (Instruct)
•   Antje Keppler (EuroBioImaging)
•   Hugh Laverty (IMI)
•   Vitor Martins dos Santos (ISBE)
•   David Morrow (EATRIS)
•   Antonio Rosato (West-Life)
•   Bahne Stechmann (EuOpenScreen)
•   Merlijn Van Rijswijk (PhenoMeNal)
As the overall goal of the round table was to develop ideas to foster usage of Research Infrastructures and therefore to increase their impact and role for innovation, most of the discussion focused on the exploitation of Research Infrastructures in the area of Biomedical Sciences, both physical and electronic, by European researchers working in biological and biomedical sciences. The assessment of the services offered, the corresponding awareness by the target communities and the scientific community at large, existing/viable options for

integrated offers, how to increase appreciation of RI services, training of users were among the topics addressed. Presenters were asked to focus their presentations on specific selected points. This allowed to establish a common ground for the open discussion and make it easier to reach consensus on a set of specific topics. During the open discussion, several comments/questions came from the audience (about 100 attendees). This was important to obtain a better grasp of the point of view of infrastructure users, including their perceived bottlenecks and limitations of the current initiatives providing access to infrastructures (both electronic and physical).

A follow-up of the round table was organized on the occasion of the second annual meeting of the CORBEL project. This event took place in Amsterdam on October 26th, 2017. This second event "Visibility, usage and impact of research infrastructures - follow up of the Brno Round Table" focused on the (re)evaluation of some of the items outlined in the consensus document. The following participated in the second Round Table:

1. Serena Battaglia (ECRIN)
2. Susan Daenke (INSTRUCT)
3. Sven Fahrner (EMPHASIS)
4. Antje Keppler (EuroBioImaging)
5. Frauke Leitner (Euro-BioImaging)
6. Giovanni Migliaccio (EATRIS)
7. Francesca Morelli (West-Life)
8. Antonio Rosato (West-Life)
9. Edoardo Saccenti (ISBE)
10. Friederike Schmidt-Tremmel (ELIXIR)

As detailed in D3.6, from the discussion it emerged that it would be highly desirable to promote the dissemination of infrastructure services in a cross-disciplinary manner to the whole scientific community and especially to graduate students and postdocs, possibly via personal interactions (e.g., University career days). For this, it was suggested that initiatives such as CORBEL should foster the preparation of promotional material to allow publicity of all BMS RIs already at the level of a single RI node. West-Life could provide a perspective from users benefiting of data access and automated data analysis, to complement and better exploit the services provided by experimental facilities. After this second round table, a document summarizing the consensus of the discussion was circulated to the established working group (MS10). The working group agreed to collaborate to target several items related to the visibility and funding of research infrastructures. An example of shared output that has been produced in agreement with the above objectives is the CORBEL Catalogue of Services (http://www.corbel-project.eu/services.html). This is the first tool to list the main services of all these RIs at a glance and it aims to facilitate the utilisation of RI services by researchers from all over the world.

Extensive discussions related to the interaction of BMS RIs with the European Open Science Cloud (EOSC) took place also within the context of developing an application to the H2020 call INFRAEOSC-04-2018 "Connecting ESFRI infrastructures through Cluster projects" (EOSC-Life). This discussion leveraged the previous interactions that pinpointed the potential value of integrating structural biology data with data from other disciplines. The concept of the

application was initially designed in the context of the CORBEL partnership. Eventually the consortium of the new application involved 63 organizations associated with thirteen biomedical RIs (BBMRI, EATRIS, ECRIN, ELIXIR, EMBRC, EMPHASIS, ERINHA, EuBI, EU-OPENSCREEN, INFRAFRONTIER, INSTRUCT, ISBE, MIRRI). There has been a broad and general consensus that Structural Biology data are of relevance to basically all RI's in biomedical research and that the integration of structural data with the data produced by other communities in the life sciences must be pursued further. In this sense, the role of electronic infrastructures and Virtual Research Environments such as West-Life is crucial. West-Life constituted a model for the design of several parts of the application, which was selected for funding by the European Commission (see news at https://www.bsc.es/research-and-development/projects/eosc-life-providing-open-collaborative-space-digital-biology) and will start in March 2019. This provides a strong direct link that will ensure the transition of our results and vision to the European Open Science Cloud.

Task 3.2 – Interaction with data initiatives (STFC, all).

In this task we focus on the interaction with European and global fora related to the various aspects of the life cycle of scientific data. The first Deliverable – D3.1, a report on the "Assessment of the life cycle of structural data and comparison with other scientific data", was submitted according to the initial schedule, after extensive discussion with all the partners. The report is publicly available from the project web site. STFC collected comments on an initial version of the report also by the participants in the 5th RDA IG/WG Collaboration Meeting, which took place in Nottingham, UK on June 2016.

The organization of a session of the Interest Group in Structural Biology at the Ninth RDA plenary (M3.1) took place on April 5th 2017[1] and has fostered further discussion on the life cycle of structural data. Besides fulfilling M3.1, the additional input allowed us to publish the document as an open access article in the Data Science Journal (https://datascience.codata.org/articles/10.5334/dsj-2018-026/). Subsequently, a poster presentation with a more extensive of the achievements and impact of West-Life was presented at the Eleventh RDA plenary, in Berlin (March 2018).

As further activities in the domain of data science, West-Life participated in both editions of the Conference "Digital Infrastructures for Research, DI4R" that took place in 2017 and 2018. On both conferences, posters and presentations on West-Life were delivered by various partners (three talks in the last edition, see https://www.digitalinfrastructures.eu/content/di4r-2018-programme-now-online). These events also offered an opportunity to reinforce links between West-Life and the European Grid initiatives. In this context, it is particularly important to note that during the lifetime of West-Life, the core of its NMR services were selected for inclusion in the EOSC-Hub initiative as a part of the EOSC-Hub thematic services[23], under the name WeNMR suite for Structural Biology.

---

[1] https://www.rd-alliance.org/ig-structural-biology-rda-9th-plenary-meeting

[2]

https://marketplace.eosc-portal.eu/services?utf8=%E2%9C%93&location=&related_platforms%5B%5D=8&rating=&research_area=

Integration of the latter suite of services within the EOSC-Hub computational infrastructure began on January 1st, 2018.

Finally, it is worth mentioning the involvement of several West-Life partners in the preparation of an application to the ELIXIR infrastructure to seek the implementation of the 3DBioInfo community in structural bioinformatics. This application also involves several ELIXIR nodes, possibly up to 15[4].

Task 3.3 – Interaction with initiatives developing science policy (CIRMMP, all).

Interactions with key stakeholders at both national and European level continued in second period of the project, allowing us to monitor the development of science policy that could affect the SB research community. In addition to the participation of some of West-Life partners in the various networks and committee mentioned in the P1 report (e.g. Lucia Banci in ESFRI SWG H&F as the Italian representative or her involvement in the project INRoad as member of one external advisory panel; Susan Daenke as member of Medical Infrastructure/Users Forum), new interactions developed through the involvement of partners in new initiatives, such as the participation of Instruct-ERIC in the ERIC-Forum project.

The participation of Instruct in the ERIC-Forum initiative is particularly relevant in the present context. ERIC stands for European Research Infrastructure Consortium and is one of the 5 European legal instruments especially set up by the European Union to provide a legal framework for (multisite) European research infrastructures, such as ESFRIs. In fact, the large majority of ESFRIs operate under the ERIC legal form. Currently there are 19 ERICs established (not all are ESFRIs). The objectives of the Forum are

(1) to identify common challenges affecting the Forum Members as entities subject to the ERIC Regulation;

(2) to collectively response to these challenges;

(3) to contribute to the further development of the ERIC Regulation, ESFRI framework and European and international research context; and lastly

(4) to foster the visibility, impact and sustainability of ERICs.

---

[3] https://www.eosc-hub.eu/training-material/wenmr-suite-structural-biology

[4] https://www.elixir-europe.org/events/launch-meeting-proposed-elixir-community-structural-bioinformatics

1.2.3.4 Deliverables and Milestones

	Title	Lead	Type	planned delivery date	Progress
D3.1	Assessment of the life cycle of structural data and comparison with other scientific data	STFC	Report	RP1	
D3.2	Organization of a round table or joint meeting involving ESFRIs	Instruct	Report	RP1	
D3.3	Report on requirements by other RI	CIRMMP	Report	14 Sept 2017	Delivered
D3.4	Interim report on science policies	CIRMMP	Report	30 April 2017	Delivered
D3.5	Publication of a joint document on the usage of structural data in different biomedical RI's	CSIC	Report	31 October 2017	Delivered
D3.6	Update of requirements by other RI	CIRMMP	Report	1 june 2018	Delivered

Milestone	Date	Means of Verification
M3.1	RP1	https://www.rd-alliance.org/ig-structural-biology-rda-9th-plenary-meeting
M3.2	24 May 2017	https://www.structuralbiology.eu/content/bringing-together-the-bio-medical-scientific-communities-the-role-of-research-infrastructures
M3.3	26 Oct 2017	doi:10.5281/zenodo.1095444

1.2.3.5 Deviations from Annex 1

Deviations from Annex 1 are described in the Tables of the previous section. Partner MU found that salary costs were lower than planned and used the opportunity to allocate extra time.

1.2.4 WP4: Operation and maintenance of the computing and data infrastructure

1.2.4.1 Summary and significant results

The work of WP4 focused on the smooth operation of the e-infrastructure resources, underlying the services provided to the end users by WP5. Throughout the whole project, the state-of-the-art software solutions of both HTC (grid) and cloud computing paradigms were followed, gradually integrating formerly isolated solutions to a common framework. Thus, WP4 ensured smooth operation of existing services from day 1 of the project. The relevant KPIs indicate a sustained increase in the effective resource usage.

Principal achievements in RP2 are:

- Interfaces to the underlying infrastructure (job submission mechanism and data access) were consolidated and adapted to the evolving environment (phase-out of glite/EMI WMS job submission).
- Portal development carried by WP5 was continuously supported.
- Solutions leveraging cloud platform and emerging containerized applications were delivered.
- The unified security model was designed in a compatible way with respect to other infrastructures, and it was put in place.
- Unified data access was provided through West-life Virtual Folder.

1.2.4.2 Objectives and tasks

O4.1: Setup the project testbed, define interfaces used to provision hardware resources, and negotiate provisioning with the resource providers at the technical level

The project production resources use EGI infrastructure, both HTC and Cloud platforms, as well as additional resources contributed by partners (computing clusters at UU, CIRMMP and MU, and data storage at INFN). SLAs ensuring the resource availability were renewed periodically. Details are given below with Task 4.1.

O4.2: Define, implement, and deploy consolidated architecture for job submission and data access

Job submission was consolidated to a common mechanism using DIRAC4EGI in case of HTC platform, and Cloudify orchestration on top of OCCI in case of cloud platform. Details are given below with Task 4.2.

O4.3: Review existing security frameworks and define consolidated solution

The consolidated solution was defined in D4.2 and gradually implemented during the project. Details in Task 4.4. below.

O4.4: Ensure smooth migration of the legacy portals to the consolidated architecture

The portals using HTC platform were gradually migrated. Progress was reported in D4.3, D4.5, and D4.6.

An orchestration service (Cloudify CFM) was set up to demonstrate progressive use of the cloud technology. Gromacs portal was migrated to it as pilot, and the new Scipion portal was developed on top of it. Details are reported with Task 4.2 below and in WP5.

1.2.4.3 Work carried out in RP2

Task 4.1: Consolidation and operation of the infrastructure

The task was responsible for the infrastructure supporting the services provided by WP5 from the first day of the project. The project leverages resources provided by EGI in both HTC and cloud platforms. The resource usage is captured in the KPIs in the following tables. Numbers for RP1 and RP2 are given to be comparable.

Metric	RP1	RP2	Difference (%)
CPU-cores (best effort)	100,000	79,000	-21%
CPU-cores (SLA)	6,850	6,050	-12%
Storage (SLA)	265 TB	54 TB	-80%
No. of HTC jobs	4,111,560	5,210,764	+13%
HTC Norm. CPU hours	26,518,294	30,711,920	+16%
No. of Cloud VMs	2,554	1,791	-30%
Cloud elapsed hours	555,239	175,290	-68%

The most significant indicator is "HTC Norm. CPU hours" which relates directly to the payload computing capacity delivered to the users. The increase in this indicator is apparent. Similarly, we

observed an increase in "No. of HTC jobs", which indicates an increase in user demand for the services.

Availability of the resources to the community is ensured with Service Level Agreements -- contracts between the community and the resource providers. Such SLAs were in place before the project start, and they were renewed periodically, in 2018 most recently. Five resource centers committed to provide up to 53 million CPU hours until the end of 2020, which is, according to the resource usage so far, sufficient capacity to ensure smooth operation of the services.

Despite there is a slight decrease in the "CPU-cores (SLA)" indicator, it was fully compensated with a higher average performance of the cores and their better utilization, yielding the final increase in "HTC Norm. CPU hours". Hence the decrease does not affect the quality of services delivered to the users.

The absolute numbers on cloud resources usage are lower; those cover the new services (Scipion in particular) delivered by the project. EGI does not plan to abandon the HTC platform in short-term, therefore there was no need for urgent migration of the existing services from HTC to cloud.

Detailed statistics on resource usage and their discussion is given in Deliverables D4.5 and D4.6 (https://zenodo.org/record/1475835).

Task 4.2: Consolidation of job management mechanisms

HTC job submission. The common job management system, i.e. the software layer responsible for submitting computational jobs from the application portals to the underlying infrastructure was unified in RP1 as described in Deliverables D4.1 and D4.3. In RP2 it was maintained with no major upgrades per se.

On the other hand, due to the planned decommissioning of the gLite/EMI WMS job submission mechanism (end of 2017), the West-life application portals were gradually migrated to the DIRAC4EGI service. A more detailed report is given in Deliverables D4.5 and D4.6.

Application software deployment. The majority of the West-life application software is distributed via CVMFS system, which is a proven, stable solution.

Besides, two specific application virtual appliances are maintained:

- ScipionCloud, containing the Scipion desktop environment, and all the required image analysis software for CryoEM. In 2018, it was extended to support GPU and remote accelerated 3D rendering.
- VirtualFolder, providing pre-configured service to give unified access to multiple storage providers (Dropbox, B3DROP, webdav, ...)

AmberTools, DisVis, and PowerFit were packaged as Docker containers (with the support of the INDIGO-DataCloud project), and integrated into a CI/CD pipeline to allow smooth distribution. Details are given in D4.5.and 4.6.
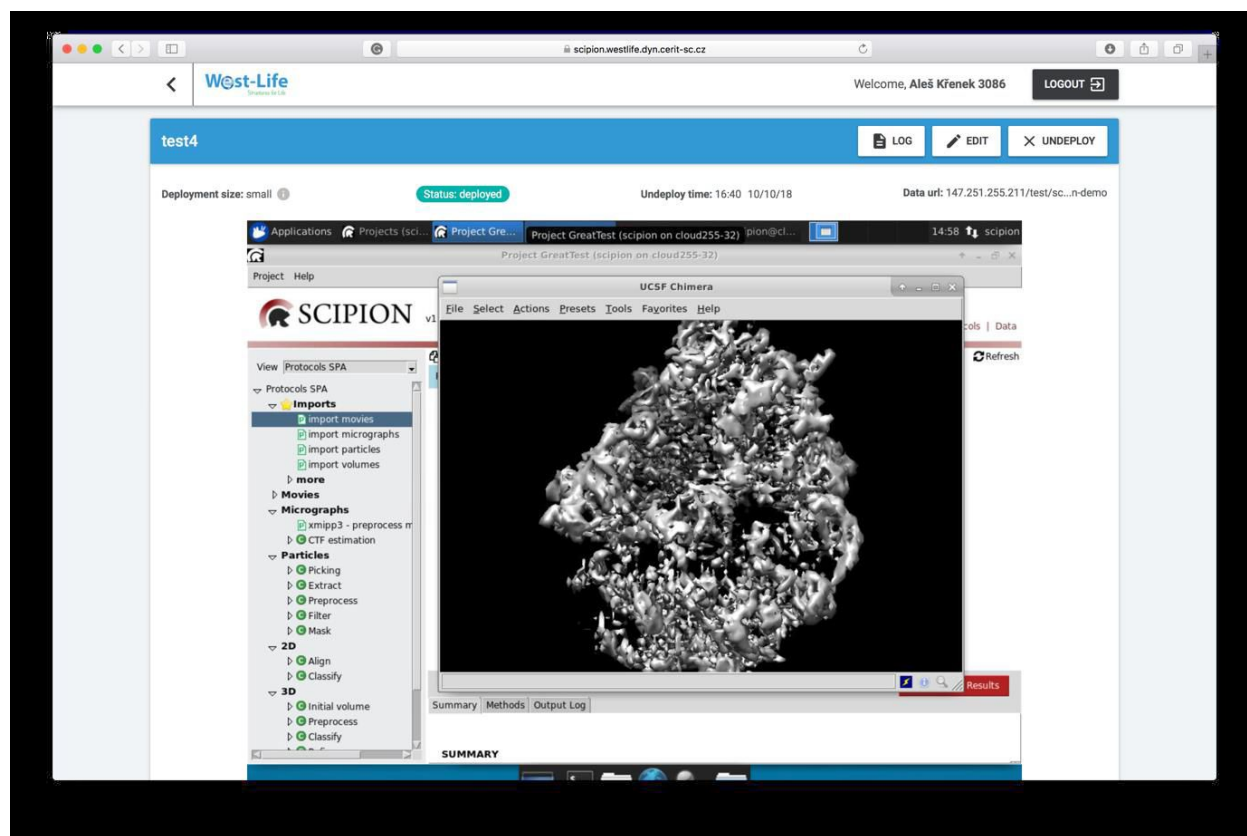
Cloud deployment and orchestration. Cloud orchestration services and their use pattern were described in D4.1 during RP1, and the application deployment was piloted on the Gromacs portal. In RP2 deployment recipes for Scipion were developed, and all the recipes we migrated to the emerging Cloudify version 4. Besides more streamlined use, the migration allows to fully leverage the Cloudify Manager service, which keeps track of the deployment lifecycle, including defined actions for failure recovery, automatic scaling etc.

More extensive development was done in wrapping the Scipion cloud deployment, which is still rather complicated to the end users, to a web portal.

The HADDOCK web portal was virtualized and wrapped in Docker containers, allowing the entire portal deployment in the cloud in few minutes.
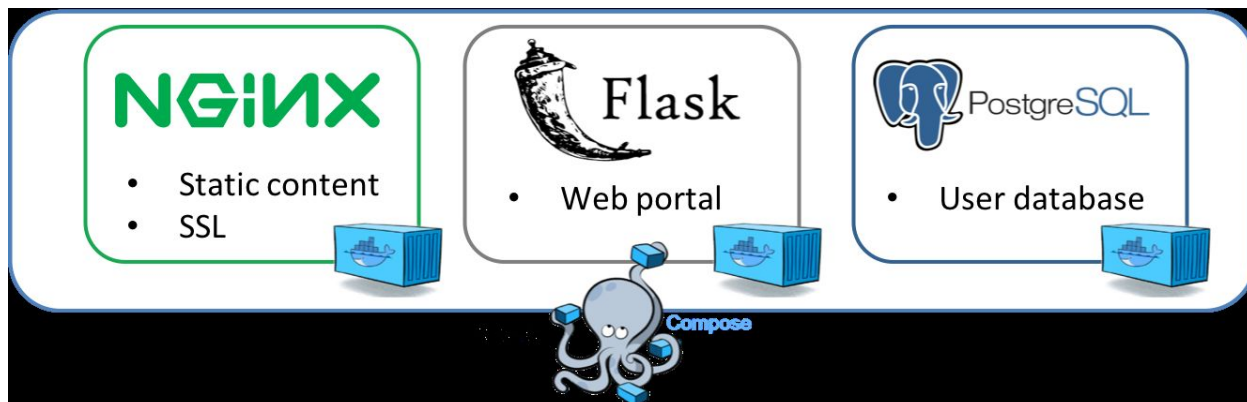
Details on all this work are given in D4.6.

Task 4.3: Programmatic Access to datasets

The majority of the work of the task was done in RP1, the outcomes are reported in D4.4.

The delivered core component, West-life Virtual Folder, integrates access to data from PDB and PDB-REDO databases seamlessly as a "special folder" allowing browsing the database entries (including visualization of metadata and 3D structure).

The Onedata service, operated by INFN, was upgraded to the most recent software version. Access to this data storage is supported with Virtual Folder as well. Moreover, the Scipion cloud deployment (described above) can leverage on storing the Scipion projects at OneData; those are staged in and out the cloud nodes seamlessly.

Task 4.4: Unified security and accounting model

The common security model to be used by Westlife services was defined in D4.2 in RP1. Further on in that period, a prototype of the authentication service (AAI) was set up. The security model follows the recommendations of "AARC blueprint". The design is interoperable, and it ensures compatibility with multiple identity providers.

During RP2, a production implementation was put in place and the Westlife services were migrated to it. Recently, the interface towards the services was extended from SAML to OpenID Connect, which is becoming popular due to more easy integration. At the time of writing D4.6 (where the details are given), the AAI registered 12 SAML and 8 OIDC services.

The development in RP2 was closely synchronized with similar developments in the field (Life-Science AAI pilot in AARC, and similar ELIXIR and BBMRI solutions) with the vision to provide compatible solutions which can be merged in future.

Direct integration with EGI CheckIn service was also demonstrated with several Westlife portals.

Following the adoption of EU regulations on data protection and privacy, West-life AAI components were reviewed to ensure compliance. The "West- Life AAI Acceptable Usage Policy" (AUP) and "West-Life AAI Privacy Policy" documents were updated (keeping as much compatibility with other life-science infrastructures as possible) and they were published. Existing users were notified on the collected personal information according to the legal requirements.

Accounting information on resource usage is currently implemented at the services individually. However, a proof-of-concept aggregation service was developed to show the direction of further integration (see D4.6).

1.2.4.4 Deliverables and Milestones

Number	Title	Lead	Type	Planned	Progress
D4.1	Consolidated architecture of job submission and interaction with infrastructure	Luna	Report		RP1
D4.2	Common security model design	MU	Report		RP1
D4.3	Report on experience with deployment of consolidated platform and its interaction with infrastructure	INFN	Report		RP1
D4.4	Overview of external datasets, strategy of access methods, and implications on the portal architecture	STFC	Report		RP1
D4.5	Report on the progress of the deployment of consolidated platform and its interactions with infrastructure	INFN	Report	31 Dec 2017	Submitted
D4.6	Final report on deployment of consolidated platform and the overall architecture	MU	Report	31 Oct 2018	Submitted

Milestone	Date	Means of Verification
MS11 Inventory of available resources and testbed setup	**RP1**	
MS12 First deployment of the consolidated platform	**RP1**	
MS13 Prototype access to selected datasets	**5 Oct 2017**	**https://zenodo.org/record/1002681**
MS14 Second deployment of the consolidated platform	**27 Nov 2017**	**https://zenodo.org/record/1066821**
MS15 Final deployment of the consolidated platform	**Sept 2018**	**https://zenodo.org/record/1438087**

1.2.4.5 Deviations from Annex 1

Partner INFN required more effort than planned to achieve the objectives of WP4, and funded this by a transfer from travel expenses, as agreed in an amendment. CIRMMP found that salary costs were lower than planned and used the opportunity to allocate extra time.

1.2.5 WP 5: Virtual Research Environment

1.2.5.1 Summary and Significant Results

This WP was centered on building and operating the Virtual Research Environment. In the first period we set up the web front end of West-Life using the ARIA content management system (CMS) provided and maintained by Instruct. The portal has been constantly improved during the second period. The portal offers documentation, tutorials and help desks to both end users and developers. During the second reporting period 6 new portals were added to the West-Life offering.

Usage statistics of the various portals are showing an increased number of new users over the three years of the project, with the largest relative increase in year 3 (+31% compared to the previous year). This goes hand in hand with an overall increase in the number of job submissions. The overall West-Life portal landscape offers over 18 portals, several of which with direct connections, establishing simple workflows to facilitate their use. The portals build on EOSC resources and various SSO mechanisms.

1.2.4.2 Objectives and Tasks

The objective of this WP is to integrate existing and operating services from the WeNMR Virtual Research Community and West-Life partners into a Virtual Research Environment (**Objective 5.1**). Within the scope of this research environment relevant information, training material, as well as user support are provided to engage and support the community (**Objectives 5.2 and 5.3**). New web interfaces for relevant software solutions will be added to the VRE to increase its versatility and to make it relevant for a broader audience (**Objective 5.4**). And finally, easy distribution and maintenance of services on local infrastructures or on EGI federated cloud resources is intended to be facilitated by customized end-user VMs and/or application containers (e.g. Docker), built for various scenarios (**Objective 5.5**).

To achieve those objectives, the following tasks were initially defined.

Task 5.1 – Deployment and operation of the West-Life portal (Luna, all)

This task directly addresses **O5.1**. It started by defining the baseline of existing services across all partners (such as X-ray crystallography from CCP4 and the corresponding ones for cryoEM from the CSIC) together with those of the WeNMR VRC. These have been integrated into a new VRE portal in the first period of the project, which provides end users with a friendly and dynamical entry point to all services, knowledge and support center. In this task, we also aimed at investigating and harmonizing user authentication and authorization mechanisms (AAI). Further the VRE portal should also implement tools and services related to data discovery and access (see WP6).

This task directly addresses **O5.2 and O5.3**. The existing knowledge and support center of WeNMR, covering NMR and SAXS services have been integrated into the new VRE portal developed in P1, and new components (tutorials, use cases, help center) to support X-ray crystallography, cryo-electron microscopy and the related integrative methods have been added. In the second period of the project this task was continued to constantly improve and add content to the VRE portal.

Task 5.3 – Development and integration of new service portals (UU, all)

This task addresses **O5.4**. All existing WeNMR, CCP4 and cryo-EM portals have been integrated in P1. During the second period, the task continued with the aim to integrate new portals as they became available. Care is taken to offer user-friendly interfaces, with a VRE-integrated AAI and the most suited submission mechanisms (e.g. DIRAC4EGI).

Task 5.4 – Customized end-users VMs (STFC, all)

This task addresses **O5.5**. Structural biology research has been targeting increasingly larger macromolecular machinery of the cell. Consequently, researchers need access to a wide range of techniques and expertise in order to truly exploit structural biology data. In most cases, however they are expert in only one or a few techniques and associated software. In this task the aim is to build custom VMs for different use cases, with all the necessary software, documentation and examples.

1.2.5.3 Work Carried Out In P2

Task 5.1 – Deployment and operation of the West-Life portal (Luna, all)

The West-Life portal was deployed during the first reporting period of the project. Since then we have been constantly monitoring the usage of the various portals in order to measure their impact. In P1 their baseline was defined based on statistics collected over 2015 (i.e. including a 10 month period prior to the start of the project) and reported in **Deliverable 5.2**. At the periodic review (and report) we compared the 2016 statistics with the defined baseline, demonstrating increased usage for most portals. We are here presenting an analysis of the usage statistics over 2017 and 2018 (extrapolating the last ones to the entire year). Those statistics and related KPIs were reported in two deliverables, **Deliverable 5.7** (month 24) reporting on the 2017 usage statistics and **Deliverable 5.8** (month 36) reporting on the extrapolated 2018 usage statistics.

The following table reports on the total statistics of all portals aggregated over the entire duration of the project.

Table 1.2.5.1: Overall usage statistics aggregated over of all portals over the entire duration of the project.

	Users total	New users	Job submissions	Grid/cloud jobs	%increase #Users total	%increase #new users	%increase #Job submissions	%increase #Grid/cloud jobs
Baseline 2015	16788	4437	43090	7692083	n.a.	n.a.	n.a.	n.a.
TOTAL 2016	21218	5534	52914	9246639	26%	25%	23%	20%
TOTAL 2017	24985	8081	72684	8935684	18%	46%	37%	-3%
TOTAL 2018	35695	10566	85800	7501906	43%	31%	18%	-16%

The percentages reported are always with respect to the previous period.

Overall we observe an increasing trend in both number of new users using the West-Life portals per year and number of job submitted indicating a sustained increased usage of our services over all years.

The only decrease observed is in the number of grid/cloud jobs submitted. This can be explained by some shifts in portal usage reflecting the fact that not all portals submit grid jobs and/or the same number of jobs. However, while the total number of jobs submitted to HTC resources has slightly decreased in 2018, it is still at a high level of 7.5 million jobs per year. Note also that actually the total CPU usage on the grid has increased in 2018 from around 18 to 21 million of normalized CPU hours, as reported in **Deliverable 4.6**. So while the total number of HTC jobs has decreased in 2018, their average duration has increased. For details statistics per portal please refer to **Deliverable 5.8**.

Content has been continuously added to the portal frontend (**Figure 1.2.5.1**), which now also provides a clear measure of project progress/impact with direct links to milestones and deliverables and impact in terms of number of presentations, publications and users trained.

Figure 1.2.5.1: View of the frontend of the west-life.eu portal

The services page presents all West-Life components in graphical view (**Figure 1.2.5.2**), with an indication of the various connections between West-Life services and with other eInfra services (e.g. EUDAT, ONEDATA, …).
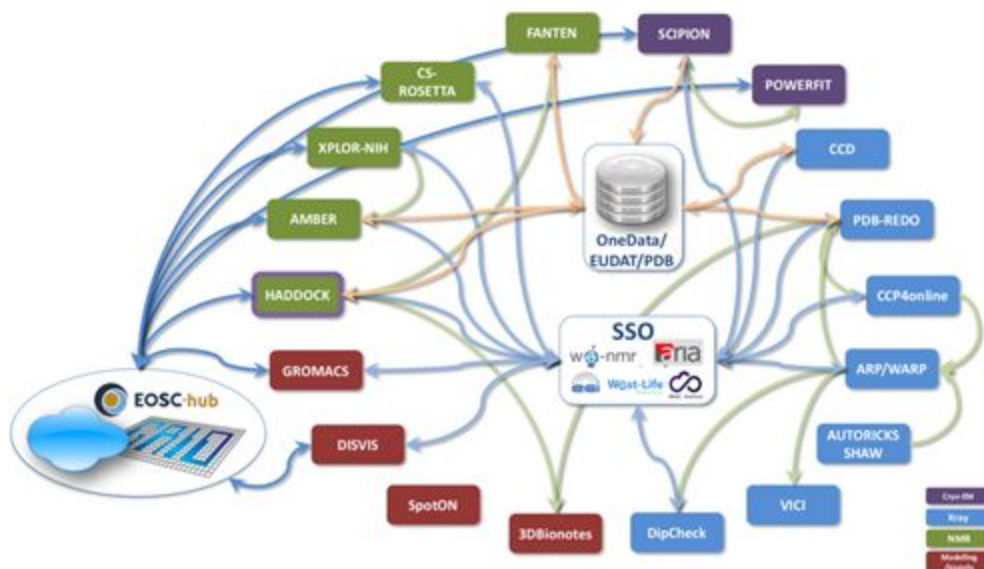
Figure 1.2.5.2: Schematic representation of the pipelines implemented between West-Life portals. Green lines indicate direct interactions between services, orange lines indicate interactions with centralised data stores, and purple lines indicate interactions with SSO services.

For a description of newly integrated services in the second period refer to Task 5.3.

The West-Life Virtual Folder has been further developed during period 2. It addresses the issue of accessing data scattered among different data storage providers. It is a web-based application allowing a user to register data storage providers and connect accounts from multiple places in one common interface. The data are not physically aggregated, since many datasets can be large. Instead, the application allows a user to view all data and download selected data on request, or to obtain a publicly accessible link conforming to the WEBDAV protocol to either a particular file or a particular folder which can then be used by another tool to download, process and upload results. Currently Dropbox and WebDAV based data storage providers are supported (e.g. EUDAT B2DROP service, ownCloud instances with WebDAV enabled or Pcloud commercial service).

Aside from allowing access to sequence or structure data, the Virtual Folder also allows visualisation of structures in a web-based 3D viewer LiteMol. This is implemented for both user data in connected folders, and for structures in the PDB and PDB-REDO databases. Additional web-based visualisations for sequence data and value added annotation, alongside an interactive display of topology for a particular protein molecule, allows users a better understanding of protein structures. The web based components use the PDBe REST API

(pdbe.org/api) to access information from the PDBe database. Details of this integration of the PDBe web components and REST API have been provided in **Deliverable 5.5**.
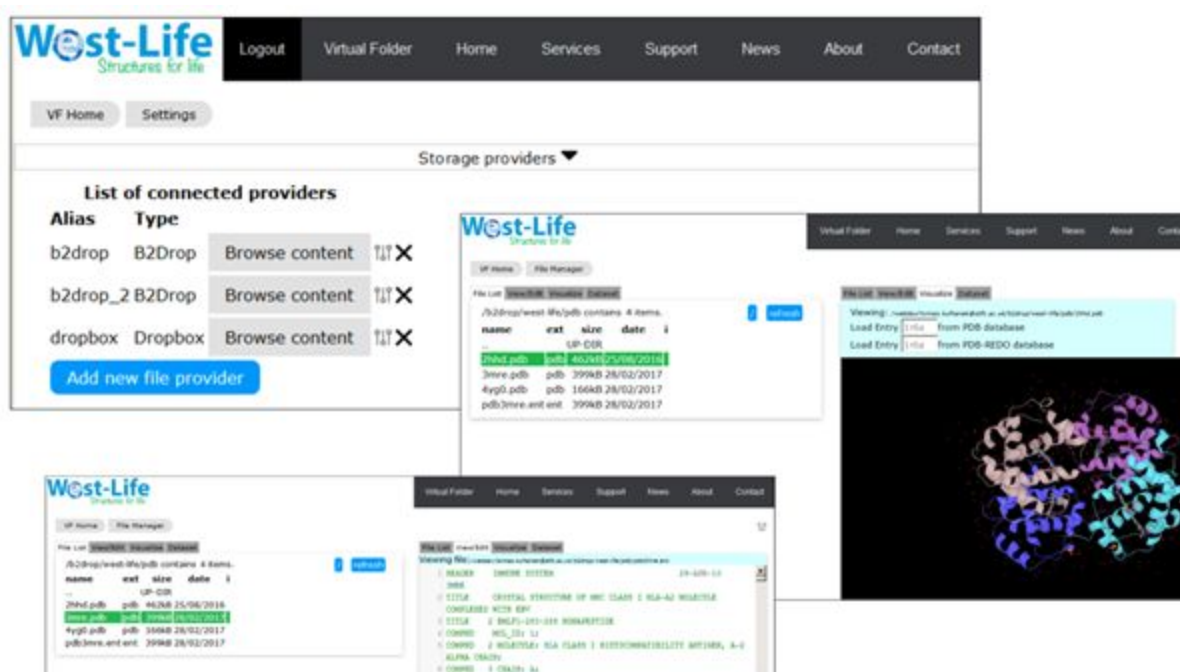
Figure 1.2.5.3: An instance of the Virtual Folder, accessible from the West-Life VRE portal after login. Three connected data providers are shown, together with a file listing from one. Standard file browsing utilities are available, together with bespoke tools such as a PDB viewer.

Finally, all portals requiring authentication are now supporting a Single Sign On mechanism, either through INSTRUCT ARIA, EGI CheckIn or the West-Life SSO as reported in **Deliverable 4.6**. They have also all been made GDPR compliant with clear terms of use and privacy statements.

Task 5.2 – Knowledge and support center (Instruct, all**)**

User support has long been an integral part of the services brought together in the West-Life VRE, and these services have well-established forums for user engagement. These support forums act as a repository of issues and solutions that is of help to future users. The West-Life web site provides navigation to these services and federated search of them. The support to users provided through the West-Life portal combines tutorials, documentation and support forums which provide answers by experts/developers covering areas of NMR, SAXS, X-Ray crystallography, Cryo-Electron Microscopy and related integrative methods in structural biology.

The West-Life portal effectively aggregates various support forums (e.g. the HADDOCK forum is operated in collaboration with the BioExcel CoE – http://ask.bioexcel.eu/c/haddock) and tutorials.

Two deliverables are providing a detailed overview of the knowledge and support center activities of the project: **Deliverable 5.4**, submitted at month 23, and **Deliverable 5.9** (Month 36) providing the final report on support activities.

Within West-Life, the support available is organized into a help desk/documentation page (https://about.west-life.eu/network/west-life/documentation). Users accessing the help desk can quickly and simply access the support they require for any West-Life tool/service from a single location. The help desk has been evolving throughout the West-Life project to integrate new features such as a search of support material (https://about.west-life.eu/network/west-life/Search-Support). The help desk has also expanded to add support for the new tools/services developed and added to the West-Life catalogue. For details refer to **Deliverable 5.9**.

Next to supporting end users, West-Life also aims to promote the development of new, cutting-edge structural biology software and resources. As a result, the pathway to creating a structural biology portal should be as easy as possible and this requires dedicated support for the portal developers themselves, not just for the end users of the portals. To achieve this there is a dedicated developers page on the West-Life website at https://about.west-life.eu/network/west-life/developer-help , accessible from a new main menu link. On this page developer can find:

- Links to the West-Life Wiki on how to implement the Single Sign On (SSO) with ARIA.
- Links to documentation on how to make your portal compatible with the West-Life virtual folder, a single point of access for cloud and local file storage for use with structural biology tools.
- Link to the DIRAC4EGI project documentation. DIRAC is a workload management service provided by EGI which manages access of users to compute services.
- Link to a page on best practices for web services.
- Contact form link to contact the West-Life team.

The most active support forum is the HADDOCK forum (http://ask.bioexcel.eu/c/haddock) with a sustained number of posts over the years (**Figure 1.2.5.4**)
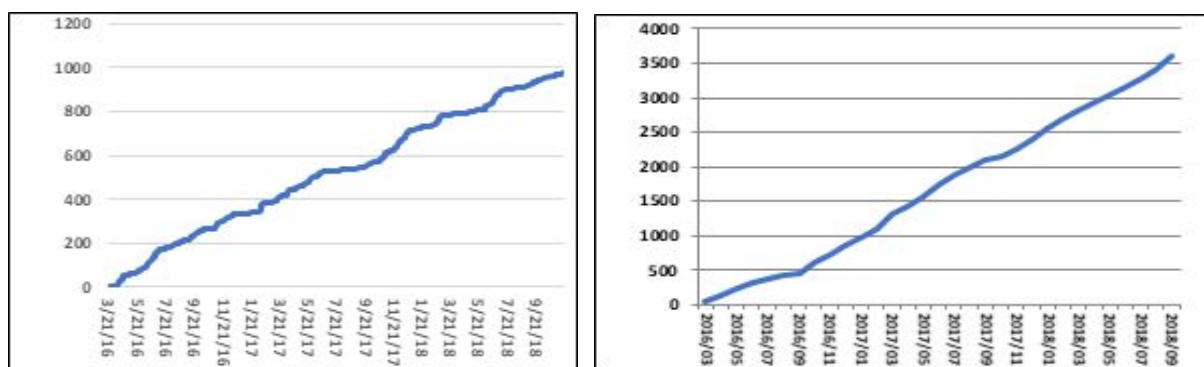
Figure 1.2.5.4: Cumulative number of posts in the HADDOCK forum at ask.bioexcel.eu (left) and to the CCP-EM mailing list (right)

Task 5.3 – Development and integration of new service portals (UU, all)

Since the midterm report, 5 new web portals have been developed and put in production. As detailed descriptions are provided in Deliverables 5.7 and 5.8, here we will only give a short overview of the new portals put into production in the second half of the project.

- **SpotOn** is a robust algorithm developed to identify and classify the interfacial residues as Hot-Spots (HS) and Null-Spots (NS) with a final accuracy of 0.95 and a sensitivity of 0.95 on an independent test set. The predictor was developed using an ensemble learning algorithm with up-sampling of the minor class and was trained on a large number of complexes and on a high number of different structural- and evolutionary sequence-based features.

- **3DBIONOTES** is a web application designed to automatically annotate biochemical and biomedical information onto structural models. Current sources of information include post-translational modifications, genomic variations associated to diseases, short linear motifs, immune epitopes sites, disordered regions and domain families. 3DBIONOTES has also been integrated in the output page of other West-Life services such as HADDOCK and close connections with PDB-REDO have been established so that newly proposed models can be annotated and visualized with 3DBIONOTES.

- **CCP4-SIMBAD** is a pipeline for sequence-independent molecular replacement, developed within CCP4 by the University of Liverpool and STFC, and available from the West-Life portal. It is good for identifying whether a crystal contains a contaminant protein, and can also help to find homologous structures in difficult-to-solve novel target cases.

- **DipCheck** is a validation tool for protein backbone geometry. DipCheck classifies the geometry into four categories: Favoured region, Allowed region, Generously allowed region and Disallowed region. DipCheck also classifies the overall geometry of a protein model, according to its DipScore distribution, in the same four categories. DipCheck adds a structure validation tool to the West-Life services portofolio
- **PRODIGY-Ligand** is an extension of the PRODIGY web-server for the prediction of binding affinity in protein-small ligand complexes. The predictive method has been successfully applied for the blind prediction of 102 protein-ligand complexes during the D3R Grand Challenge 2. PRODIGY-LIG has the advantage of being easy to use, generic and applicable to any kind of protein-ligand complex.

Next to portals developed within the West-Life consortium, the West-Life portal is also aggregating third party portals. These have been integrated upon request from the respective portal developers in order to provide exposure to their tools. Currently the portal provides 9 entries to such third party portals (**Figure 1.2.5.5**). This is a nice demonstration of the impact and visibility of West-Life.
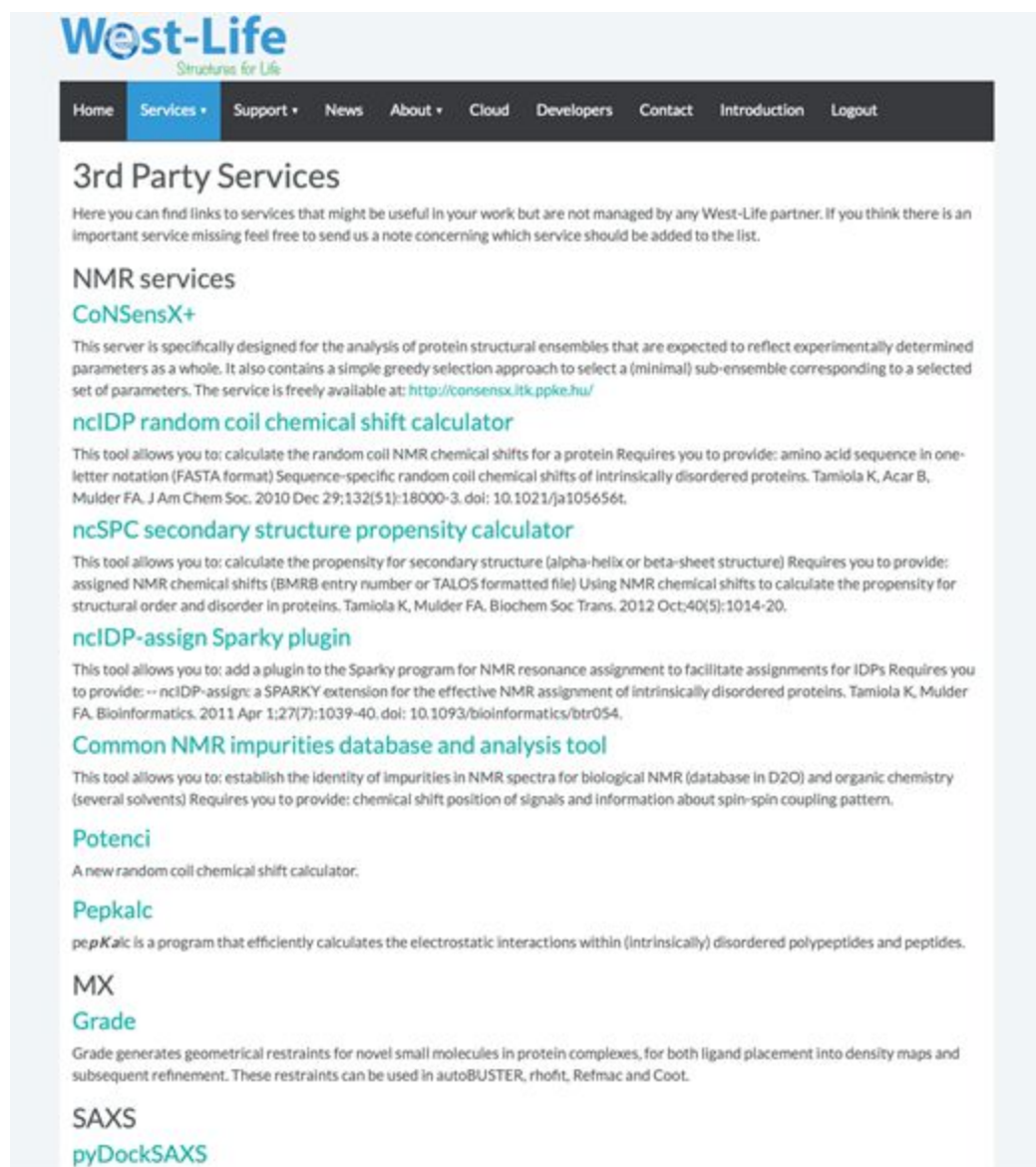
Figure 1.2.5.5: Snapshot of the third party portal page (https://about.west-life.eu/network/west-life/services/3rd-party-services)

Task 5.4 – Customized end-users VMs (STFC, all)

Three types of templates of virtual machines (VM) or containers were already reported in detail in the midterm report. Since then those have been updated and further developed. We will summarize here the main updates since the last report.

- **Virtual Folder VM** – A new VM template was introduced for OpenStack middleware and deposited in the EGI AppDB (https://appdb.egi.eu/store/vappliance/west.life.vm),

provided next to the existing OpenNebula instance (https://appdb.egi.eu/store/vappliance/d6.1.virtualfoldervm). It launches at boot time the most updated versions of OS based on Scientific Linux 7 and virtual folder from repository at `/cvmfs/west-life.egi.eu`. Binaries are released to the repository at `/cvmfs/west-life.egi.eu` on regular basis. The Virtual Folder on the public portal uses a customized Docker container which launches the most updated release from `/cvmfs/west-life.egi.eu`. Various virtual machines can be customized and built using scripts at https://github.com/h2020-westlife-eu/wp6-vm for testing, development and other special purposes with updated documentation at https://h2020-westlife-eu.gitbook.io/virtual-folder-docs/virtual-machines

- **ScipionCloud VM** – The ScipionCloud VM has been running on the EGI Federated Cloud, at CESNET-MetaCloud site, providing access to the Scipion Web Tools portal reachable at http://scipionwebtools.westlife.fedcloud.eu/m/services/ Details of its implementation have been reported in Deliverable 4.6.

1.2.5.4 Progress Towards Deliverables And Milestones

Number	Title	Lead	Type	planned	Progress
D5.1	Project Portal	Luna	Website		RP1
D5.2	Overview (baseline) of services and portals to be integrated into the new VRE	UU	Report		RP1
D5.3	Prototype of the new VRE portal functionality	Luna	Website		RP1
D5.4	Report on activities of the Helpdesk	Instruct	Report	April 2017	Delivered
D5.5	VRE-integrated PDBe search and query API's	EMBL	Repor		RP1
D5.6	Report on available VMs with associated documentation/use case for each of them	STFC	Report	April 2017	Delivered
D5.7	Report on access and usage statistics of the various services	UU	Report	Oct 2017	Delivered
D5.8	Report on access and usage statistics of the various services	UU	Repor	Oct 2017	Delivered
D5.9	Update Report on activities of the Helpdesk	Instruct	Repor	Oct 2017	Delivered

Milestone	Date	Means of Verification
M5.1 First version of new VRE portal online (Month 4)	**RP1**	The VRE portal implementation is available at http://portal.west-life.eu/
M5.2 First integrated, grid- or cloud- enabled X-ray service	**RP1**	http://internal-wiki.west-life.eu/index.php/D6.1: CCP4_integration
M5.3 First integrated, grid- or cloud- enabled cryo-EM service	Oct 2017	https://zenodo.org/record/1037961
M5.4 First implementation of data catalogue in VRE	May 2017	http://internal-wiki.west-life.eu/index.php?title= Data_catalogue

1.2.5.5 Deviations From Annex 1

MU were able to use more junior staff than originally planned, resulting in the contribution of more person months while staying within budget.

1.2.6 WP6: Data Management

This WP is concerned with the development of a data management layer, to handle the diversity and distribution of research data. The work of the first reporting period focussed on the creation of a Virtual Folder (VF) for unifying the different data sources available to the scientist. The VF has been integrated into the core architecture (WP4) and the VRE (WP5).

In RP2, we developed a repository for experimental data collected by visiting scientists (D6.2). This was delivered as virtual machine template in order to be installed in local or institutional repositories. In D6.4, this was extended to support more forms of metadata, especially supporting W3C PROV-O standard storing provenance record of a dataset.

D6.3 provided support for queries about macromolecular assemblies

The work on metadata for predicted models focuses on categorizing data from a well established macromolecular docking experiment. The EMBL-EBI first investigated user requirements for the search facets by incorporating feedback from the macromolecular docking community, planned the implementation of the facets and query interface (M6.4), and then developed it (D6.3).

1.2.6.2 Objectives and tasks

Task 6.1 Virtual Folder

The effort built on previous work by WeNMR and others to create a virtual folder view of scattered data, as reported below.

Task 6.2 Repository

For experimental facilities that are newly embarking on data management, we provided a reference implementation of a repository that supplies suitable metadata to the portal (D6.2).

Task 6.3 Provenance

This was then extended to a provenance view showing the contribution of each sample, experiment and dataset to the structural conclusions of the project, using the PROV-O standard (D6.4).

Task 6.4 Metadata for predicted models

Previously, predicted models have been less findable than experimental models, which obstructed collaboration between the computational modeling and structural biology communities. In RP1, we reported on existing standards for the relevant metadata (M6.3). This provided the basis for further progress in RP2, as reported next.

1.2.6.3 Work carried out in RP2

Task 6.1 Virtual Folder
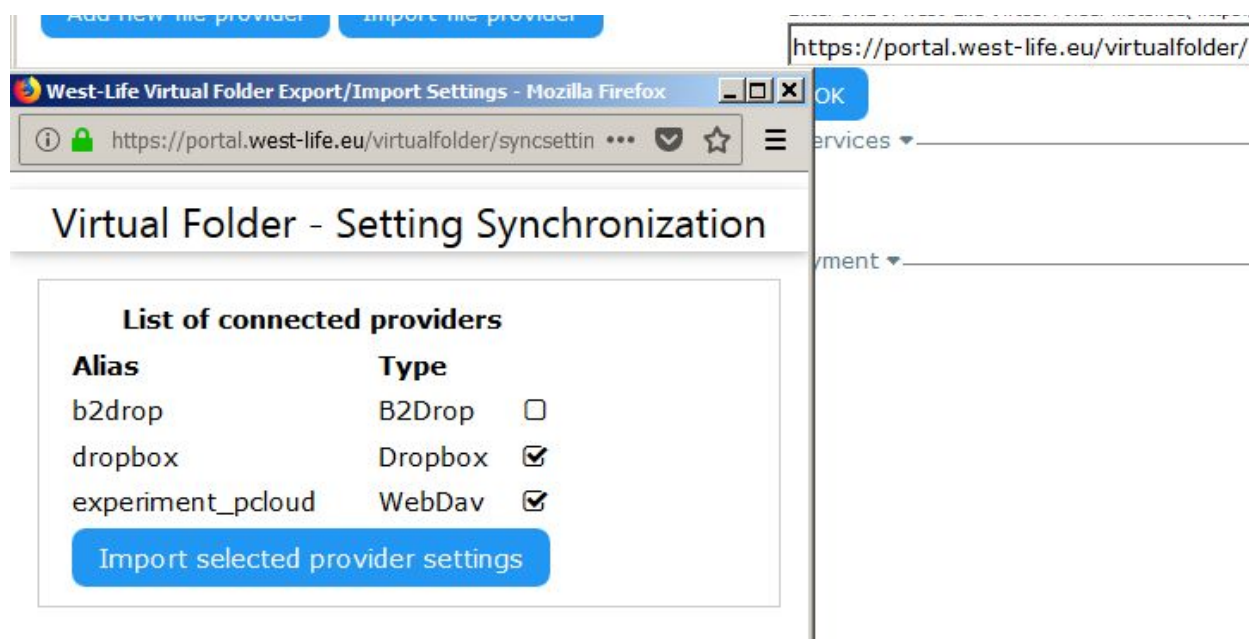
The Virtual Folder was established in RP1 and maintained as a web service in RP2. During RP2, it was integrated with West-Life SSO allowing user to log-in using ARIA, social network (Google) or WeNMR account. Additional support for metadata was implemented in the backend and REST API as well as web UI. The user can generate and modify metadata of a directory or
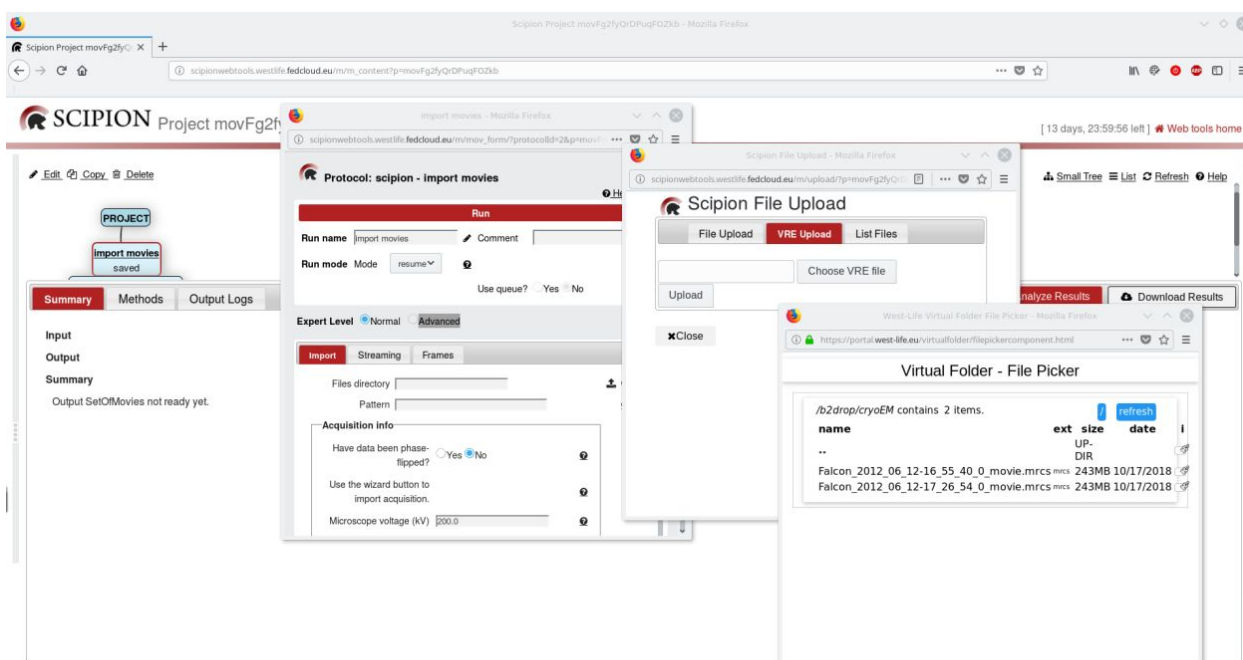
particular file including a first version of the provenance record, see Task 6.4 and Deliverable D6.4.

Local installation of the the Virtual Folder was supported with security fixes. VM templates based on CernVM 4 are updated by the vendor (CernVM team) continuously. Technical details reported in Deliverable D5.6. Additional feature was introduced in RP2 to export/import user's settings from one Virtual Folder (public portal) to another Virtual Folder (local).

Task 6.2 Repository

An exemplar Repository implementation was done using the Spring framework delivering a REST API and storing persistent data in a database. A flexible frontend UI obtains pieces of relevant data from the backend REST API or from third-party services - user info from West-Life SSO and project proposal info from ARIA API. The repository can be installed using bootstrap scripts on a clean derivative of RHEL 7 (Centos 7, Scientific Linux 7). Vagrant configurations are prepared to automate this installation in a local environment.

Binary distributions are released and distributed via CVMFS as a mounted directory /cvmfs/west-life.egi.eu. There is an option to install the repository from source code available from github. (Details in Deliverable D6.2).

A feature to harvest metadata from datasets available within the repository was introduced. These metadata are available to be exported or used by third party applications via the REST api.

Task 6.3 Provenance

We have developed mechanisms for automatically saving the provenance information, storing it, and viewing it. Once the necessary take up is achieved, so provenance data is recorded, this will allow a user looking at structural results to see the processing steps and, back along the chain of custody, the experimental data. Technical details are reported in Deliverable D6.4.

The user interface was enhanced in the metadata view of the dataset (folder) or file by a provenance section with the possibility to generate/edit and store the provenance record inside the virtual folder as another type of metadata or to export the provenance record to a third party provstore.

Task 6.4 Metadata for predicted models

The work on metadata for predicted models focuses on categorizing data from a well established macromolecular docking experiment. The EMBL-EBI first investigated user requirements for the search facets by incorporating feedback from the macromolecular docking community, planned the implementation of the facets and query interface (M6.4), and then developed it (D6.3). This service is now being tested at https://wwwdev.ebi.ac.uk/pdbe/widgets/capri-search/?advancedSearch:true=. A screen shot is shown below.
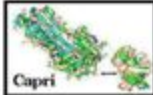
CAPRI: Critical Assessment of PRediction of Interactions

Community wide experiment on the comparative evaluation of protein-protein docking for structure prediction

Hosted By EMBL/EBI-MSD Group

Round 17

NOTE The Coordinates are not yet available for download

Target 38

T38 is a complex between centaurin-alpha 1 and the FHA domain (residues 448-545) of KIF13B. The complex and the coordinates of centaurin-alpha 1 (entry 3FEH, unreleased and confidential) are a gift of Dr. Hee-Won Park (Consortium for Structural Genomics and University of Toronto, Canada).

Predictor groups	40
High accuracy (***):	0
Medium accuracy (**):	0
Acceptable (*):	0
Incorrect:	370
Clashes:	29
	399
Clash threshold	59
average	19
std dev	20

Uploader groups	10
High accuracy (***)	0
Medium accuracy (**)	0
Acceptable (*)	0
Incorrect	662
Clashes (+)	337
	999

(+) includes 100 models of T37 - One uploader group submitted 100 models of T37 instead of T38

1.2.6.4 Progress towards Deliverables and Milestones

Number	Title	Lead	Type	planned	Progress
D6.1	Report on Virtual folder	STFC	Report		RP1
D6.2	Repositories	STFC	Report	April 2018	Delivered
D6.3	Report on Assembly queries	EMBL	Report	June 2018	Delivered
D6.4	Report on Provenance	STFC	Report	July 2018	Delivered

Milestone	Date	Means of Verification
M6.1 Virtual Folder	RP1	
M6.2 Repository	April 2018	D6.2
M6.3 Metadata standards for complexes	Nov 2016	http://internal-wiki.west-life.eu/images/4/44/Report_on_existing_standards.pdf

1.2.6.5 Deviations from Annex 1

D6.2 was delayed, by agreement. In addition, partner CIRMMP participated in this development and became the first site to install the repository.

1.2.7 WP7 Joint Research

1.2.71. Summary and significant results

The goal of this WP is to explore ways to extend, combine and enhance existing services, so that they can reach a wider audience. Others topics to be addressed include focused research on exploring Big Data approaches, together with a number of standardization issues.

The Work Package has been completed successfully, with all its Milestones and Deliverables achieved and even surpassed.

1.2.7.2 Objectives and tasks

Work Package 7, as a Joint Research Activity, is aimed at exploring new ways to use existing or close to existing services so that broader user communities can be reached. This overarching goal was to be accomplished through four Objectives:

1. Extending and benchmarking existing services, such as ProteinCCD, PDB_REDO, REFMAC and HADDOCK
2. Combining services into new workflows
3. Studying large sets of output data using Big Data approaches
4. Evaluating current and developing metadata standards for workflow definition

These have a one-to-one relationship with the four Tasks as follows.

Task 7.1 – Extending and benchmarking existing web services (EMBL-HA, CSIC, STFC, NKI, MU, CIRMMP, Instruct, UU)

Four "exemplary" and very popular services were chosen to concentrate our efforts on: ProteinCCD, PDB-REDO, REFMAC and HADDOCK.

ProteinCCD is aimed at the very initial stages of the Structural Biology pipeline, designing suitable genomic constructs to obtain good samples for further analysis.

PDB-REDO, in turn, focuses on the end of the data analysis pipeline, automatically optimizing atomic models from X-ray crystallography using the latest software.

REFMAC provides the core functionality for the refinement of atomic model refinement against experimental data, so that it is the basis of many other modelling approaches.

Finally, HADDOCK is one of the best known services for biomolecular docking (for High Ambiguity Driven biomolecular DOCKing).

In P1 we provided extensive information on the work being done on these tools that, for completeness, will be briefly commented when presenting in more detail the progress during P2.

Task 7.2 – Combining existing services into new workflows (CSIC, EMBL-EBI, STFC, NKI, MU, CIRMMP, UU)

In this Task the focus is on quality measures, selecting two fields of activity. The first one was cryo-EM and the second one was aimed at the quality analysis of predicted complexes.

With respect to cryo-EM, the emphasis was placed on developing a simple but accurate workflow to provide local resolution estimation by a variety of techniques, techniques so diverse that they have never been thought to work in parallel. The final result was important enough to be accessible not only from West-Life Portal, but also from the Validation pages of the main cryo-EM database (EMDB, at the EBI, https://www.ebi.ac.uk/pdbe/emdb/)

As for the evaluation of predicted complexes, the goal was to assist the CAPRI (http://www.ebi.ac.uk/msd-srv/capri) community and to contribute to new methods development. It should be specially noted that CAPRI analysis protocols have been developed over a period of more than a decade, but that this is the first time that users have direct access to the assessment program via a web server.

Work on these topics has extended over P1 and P2, since the associated Deliverables were scheduled later in the project.

Task 7.3 – Handling and Mining Big Data (STFC, CSIC, MU)

This was a highly exploratory Task in which the experience of STFC in Big Data, combined with the collective experience in Structural Biology of the rest of the partners, was to provide the proper environment for the application of Big Data techniques to "selected" Structural Biology issues. It should be noted that we did not know at the start of the project which could be these "selected issues", and that a crucial preliminary work was to present and analyze suggestions from all partners. In this way two topics were highlighted for further exploration: (1) The application of Natural Language Processing (NLP) to increase the informational content of the main database in Structural Biology (PDB) and, (2), the use of deep learning approaches to analyze some features of cryo-EM maps. We also want to note that during the Mid Term evaluation we were provided with valuable insights on how we could relate to other NLP projects; naturally, these pieces of advice have been taken into account in the final work.

Considering that the Deliverable associated to this Task was towards the end of the project, the concrete work done in this respect was not presented in the P1 Report.

Task 7.4 –Assessing and Extending metadata formats (STFC, EMBL-EBI, NKI, MU, CSIC, CIRMMP, Instruct, UU)

The main activity here was the survey of available metadata standards relevant to the set of services offered by West-Life, identifying existing standards that could be used in the services

as well as "gaps" to be addressed. The project services were then to be updated using these standards.

The Deliverable associated to this task was programmed for the second half of the project, and therefore was not covered in P1 Report.

1.2.7.3 Work carried out in Period 2

Task 7.1 Extending and benchmarking existing web services

Work on extending existing services, especially the four exemplary ones previously indicated, was mostly carried out during the first part of the project, as initially planned. However, new and interesting work has also been accomplished in P2, even beyond initial project expectations. Going over ProteinCCD, PDB-REDO, REFMAC and HADDOCK, the final results of the project are as follows.

ProteinCCD. This service was completely re-implemented, from being a Java applet to a web application with a Flask/Biopython backend and a Bootstrap/Javascript front-end. On this new implementation, more functionality, security and increased speed performance was achieved as part of D7.2, as described in P1. Then, new scoring and ranking options have been added, by splitting the analysis and ranking in three tiers: a) use algorithms from the BioPython library to calculate the molecular weight, isoelectric point and absorption coefficients for all designed constructs; b) for providing solubility information we evaluate from external servers; c) for providing crystallizability information we also chose external servers. In this way, D7.3 has also been fully accomplished.

PDB-REDO. During P1 the planned parallelization of this service was fully accomplished (D7.1), and the work continued so as the web server engine was rewritten to incorporate ARIA based single sign-on and an API was created to allow other web servers in West-life to submit jobs in a multi-server workflow (also related to Task 7.2). Additionally, and making a connection with WP5 Task 5.3 (new services), PDB-REDO and 3DBIONOTES have been programmatically interconnected, so that PDB-REDO results can be visualized, annotated and analyzed using 3DBIONOTES in a totally transparent manner (Figure 1.2.7.2). This is an important new piece of work, accomplished during P2, and beyond project initial expectations.
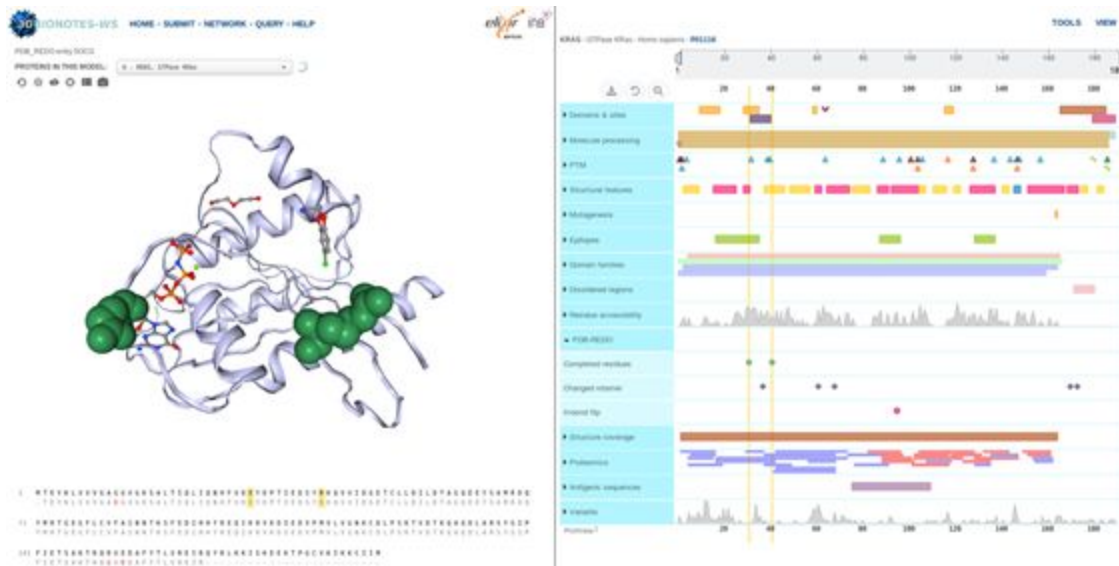
Figure 1.2.7.1 PDB-REDO refined structure of the KRas human protein. The figure displays 3DBIONOTES screenshot displaying the refined PDB-REDO entry 5ocg. The completed residues after the refinement process are displayed in green color. In 2018, over 1500 3DBIONOTES entries were requested from PDB-REDO web site.

REFMAC. The many changes that have taken place in REFMAC were largely covered in P1, specifically the remodeling of the service so as to provide remote access to the REFMAC engine in the context of the automated model-building suite ARP/wARP. Work has continued to fully deliver D7.4, incorporating cryo-EM map constraints. In this latter case the map and the initial model are converted to a minimum pseudo-crystallographic asymmetric unit-cell that in some cases may substantially reduce computation time. The map in a new coordinate frame is then used to calculate structure factors in the MTZ format, which is required by REFMAC. Finally, a REFMAC-based reciprocal-space refinement or a combined iterative ARP/wARP-REFMAC model building is performed depending on the users choice and the characteristics of the data. At the last step, a refined model is back-transformed to the original reference frame and provided to the user in PDB and mmCIF file formats, following the recommendations contained in D7.9 on metadata standards.

HADDOCK. This is also the case of a very popular service that has gone over a major re-engineering, resulting in HADDOCK2.4, fully supporting cryo-EM maps in addition of a whole set of new functionalities, extending even to the CNS core (simulation engine). Deliverable D7.5 has been fully accomplished. This new rewriting of the system has allowed the incorporation of new portals, such DisVis and PowerFit. We may add that PowerFit and the cryo-EM workflow for quality assurance have been interconnected, so that cryo-EM maps can be automatically exported to PowerFit. The small delay in this Deliverable was already reported

in P1 Report, and everything has proceeded according to this schedule during P2, including the addition of the named new portals and the new inter-service connections.

Besides extending the services mentioned above, that was the original work scheduled in West-Life, additional work towards the development of future services has also been performed in the project. We will highlight two examples:

1)   Work towards the development of new cryo-EM cloud services has been performed between MU and CSIC. In this context, a test development on a new approach to automatically deploy Scipion (the cryo-EM image processing developed in Madrid) on the cloud using Cloudify and puppet was first presented as a new Instruct user case during the recent IBERGRID conference (https://wibergrid.lip.pt/site/ibergrid2018/) and it is to be further developed in the context of a new EOSC-Synergy project, under review.

2)   We have developed a novel automated approach for efficient detection and 3D structural modelling of contamination in macromolecular crystallization. The potential impact of correct identification of the contaminant will help overcome currently troublesome issues in the interpretation of X-ray diffraction data. Soon after the West-Life project comes to an end, we expect to implement this development as a new web service linking the computational SIMBAD services provided by the CCP4 group at STFC and the ARP/wARP services at EMBL-HA.

Task 7.2 Combining existing services into new workflows

 Work in Task 7.2 is connected to other pieces of work reported in WP5 and to the development of services reported above. For the sake of clarity we will divide this section into two subsections: (1) Specific new workflows identified in the project as Deliverables (D.7.6 and D7.7), (2) Other combinations of services that have been also developed along the project

Regarding the first specific workflow, D7.6 was to focus on quality assurance in cryo-EM by providing quantitative estimation of local resolution. We initially considered ResMap (http://resmap.sourceforge.net/), but we have finally developed a service capable of integrating three totally different ways to estimate local resolution: Resmap, Blocres (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3837392/) and Monores (https://www.nature.com/articles/nmeth.4651). The need to consider these different estimation was due to the fact that not all programs performed equally well for all maps, especially since the differentiation of what was "signal" (the cryo EM map) and what was "noise" relied on the specification of an artificial boundary, often user-generated, normally referred to as a "mask". The degree of subjectivity introduced by this mask is not trivial, affecting the outer regions of a macromolecule, that are those involved in all types of interactions and are, therefore, especially important from a biomedical perspective.

Beyond these initially considered developments, we have also advanced in the area of relating a number of statistics generated during the process of angular refinement with the probability of reaching a correct (global optimum) or incorrect (local optimum) map (https://about.west-life.eu/network/west-life/documentation/scipion-web-tools/my-reliability-tool). This is especially useful at low resolution, when chain tracing is impossible.

The second workflow that was specified as a Deliverable, D7.7, related to quality assessment for predicted complexes. Its goal was to provide a web service capable of partially automatize assessment in the context of CAPRI (The Critical Assessment of Predicted Interactions (http://www.ebi.ac.uk/msd-srv/capri/), a community-wide initiative established in 2001 that has been at the forefront in this endeavor. Traditionally, the assessment of predicted complexes is carried out by an assessment team. The evaluation is based on a number of criteria, largely adopted by the CAPRI community. However, thanks to D7.7 users now have the "the "CAPRI Analyse" server, which can provide evaluation of the quality of the predicted models based on CAPRI protocols. The landing page of CAPRI Analyse is shown in the next Figure.

Then, additionally to these concrete workflows, West-Life has created several other new workflows combining existing tools. Among them we can mention:

- PDB REDO and 3DBIONOTES, that was already commented when presenting Task 7.1
- POWERFIT and Scipion, also commented in Task 7.1
- FANTEN and HADDOCK. FANTEN (http://abs.cerm.unifi.it:8080/) is a rigid body refinement protocol for certain types of NMR data, and work has focused on ways to use

FANTEN output models to HADDOCK input requirements in order to generate high quality models of protein-protein complexes.

After the scoping studies carried out in P1, we chose to focus on developing two prototypes during P2. The first one addressed the incorporation of new annotations to PDB thanks to the analysis of scientific literature, while the second prototype was aimed to cryo-EM map analysis using Deep Learning. The two of them are described in D7.8.

On the first prototype, the scientific literature analysis was to be conducted using Natural Language Processing (NLP), where first thousands, and then millions, of manuscript were considered. In general, an NLP challenge is defined by the corpus of texts to be processed and the question to be answered.

The corpus was initially defined as all papers which are linked in EuropePMC to PDB entries, and for which full text is available through EuropePMC. It consisted of 76120 papers at the time of writing D7.8

The question was defined as the identification of mentions of specific protein residues, and linking them to the relevant protein record in Uniprot. This question was specified by West-Life partner Sameer Velankar, Team Leader, Protein Data Bank in Europe (PDBe), as a useful complement to information already present in the PDBe.

The goal was defined as extracting information from all 76,120 papers in the corpus and pushing it to the EuropePMC annotation repository, and establishing a pipeline that will automatically annotate future papers.

D7.8 constitutes an excellent example of an exploratory work from which a sustainable service is in the process of being developed. Indeed, the initial set of several tens of thousands papers has now been expanded to several millions thanks to an agreement with EuropePMC by which the papers are accessed only for the "question" defined above, in a way that respects all IP rights. This huge corpus of data is now being used to generate new annotations into PDBe

It is to be noted that during the MidTerm review we were advised to contact the NLP project OpenMinTed (http://openminted.eu/) for possible synergies. Indeed we corresponded with them about this work. They offered tools for corpus extraction and processing. Unfortunately we needed a specialized criterion for corpus definition, as described above: papers linked to PDB-e entries. This selector is provided only by the Advanced Search facility of EuropePMC.

A second prototype was selected, using machine learning to recognize features in cryoEM maps. While machine learning has been applied several times for the particle picking problem, it has not been used for identifying features of single particle reconstructions. Applications include distinguishing protein vs nucleic acid maps, identifying missing components in noisy maps, or automatic recognition of side chains as input to model building.

For the prototype, we focused on the simple question as to whether a 2D slice from a cryoEM map contains protein or not. Although a very simple use case, it demonstrates several important features: preparation of the training dataset, effect of map blurring, choice of machine learning model architecture, and initialization and refining of weights. Initial results were presented in D7.8, using a 2.2Å single particle reconstruction of beta-galactosidase as an example. A training set of 2D slices was automatically annotated using the fitted model as an indicator of which class (protein or solvent) each slice belonged to. A 5-layer convolutional neural network was trained, and gave an accuracy of 98% for slices taken from a blurred map. The accuracy is slightly worse (94%) for slices taken from the original sharpened map, and much worse (25% and 59%) when making cross-dataset predictions.

Since D7.8 was submitted, work has continued to improve the machine learning model, to explore other cryoEM reconstructions, and to investigate other ways to characterise volume data. The CCP-EM group at STFC is teaming up with the new Scientific Machine Learning group at STFC to pursue machine learning approaches in cryoEM. The work done in this WP will be continued and developed further in that activity, ensuring sustainability of this theme.

Task 7.4 Assessing and Extending metadata formats

A review of metadata standards (D7.9) found that the greatest need was not new standards, but use of current ones. Indeed, this was not a surprising result, but D7.9 performed a prioritization/highlighting work, analyzing those areas that had to be specifically addressed in West-Life current offer of services as well as in all future work in the Structural Biology domain.

Probably the most important conclusion was the use of mmCIF format for import and export of structural models. Indeed, West-Life partners have made significant progress towards use of mmCIF:
- DipCheck now accepts mmCIF input
- ARP/WARP accepts mmCIF for ligands but not proteins (EMBL-Hamburg)
- The PDB-REDO databank now stores mmCIF files
- The PDB_REDO service reads and writes mmCIF
- Version 8 of PDB-REDO will use mmCIF internally (NKI).
- 3DBIONOTES reads mmCIF
- The new HADDOCK2.4 web portal nows reads and writes mmCIF files.

Another important consideration was the realization that the link between raw data and final structural models required additional work in the areas of Provenance and Workflow languages.

As for Provenance, the W3C recommendation PROV-O (the PROV Ontology) has been considered adequate for Structural Biology, and that two critical West-Life infrastructures such as the Virtual Folder and the Repository should implement support for PROV-O . We made a

successful application to become an EUDAT Data Pilot. In addition, Instruct has requested recognition as an EUDAT community for B2SHARE, with the core PROV-O concepts added as community metadata terms.

In terms of workflow specification, the issue is very complex, since many West-Life services have been developed without using external workflow engines, so that they have created specific, ad hoc, implementations. Indeed, multiple workflow engines exist, and the choice of one of them is not obvious. Therefore, the identification of this need is clear, although its solution is not; further work in this area has to be done in the context of the new EOSC (European Open Science Cloud), specifically within the coming ESOC-Life project. The new "Common Workflow Language" is likely to help.

1.2.7.4 Progress towards deliverables and milestones

	Title	Lead	Type	planned	Delivered
D7.1	Multi-core implementation of PDB-REDO server	NKI	Public Report	RP1	
D7.2	ProteinCCD with new analysis options	NKI	Public Report	RP1	
D7.3	ProteinCCD with construct scoring and ranking	NKI	Public Report	31 Oct 2017	27 Nov 2017
D7.4	A REFMAC server for EM and NMR	EMBL	Public Report	30 April 2018	18 May 2018
D7.5	A HADDOCK server for EM	UU	Public Report	30 April 2018	25 May 2018
D7.6	EM quality assurance workflow	CSIC	Public Report	31 Oct 2017	30 Oct 2017
D7.7	Quality analysis workflow for predicted complexes	EMBL	Public Report	30 April 2018	11 May 2018
D7.8	Report on prototypes constructed using Big Data approaches	STFC	Public Report	30 April 2018	11 May 2018
D7.9	Report on existing metadata standards, and proposals for new vocabularies	STFC	Public Report	30 April 2018	18 May 2018

	Date	Means of Verification
MS24 Ranking in ProteinCCD	**RP1**	
MS25 Integration of PDB_REDO server and database	**RP1**	
MS26 The workflow and i/o of the Rfmac web service defined	**RP1**	
MS27 The workflow and i/o of the HADDOCK web service defined	**RP1**	
MS28 The EM Quality Assurance workflow defined	**RP1**	
MS29 The quality analysis workflow for predicted complexes defined	**23 Oct 2017**	**https://zenodo.org/record/1035144**
MS30 Big Data software introduced	**3 Nov 2017**	**https://zenodo.org/record/1040458**
MS31 Existing metadata services identified	**RP1**	
MS32 Existing services updated to existing data types	**RP1**	

1.2.7.5 Deviations from Annex 1

Extensions for D7.4 and D7.5 were agreed in amendment AMD-675858-13, because there were opportunities to improve them with the extra time.

MU were able to use more junior staff than originally planned, resulting in the contribution of more person months while staying within budget.

CSIC required some extra person months to achieve their objectives in WP7. Due to the loss of a highly experienced skilled member of staff, and maternity leave for another, they face the expense of training a new recruit in the rapidly changing field of computation for electron microscopy.

Some person months were transferred to NKI to enable success in tasks 7.1, 7.2, and 7.4, in particular D7.3.

1.3 Progress beyond the state of the art, expected results until the end of the project and potential impacts (including the socio-economic impact and the wider societal implications of the project so far)

Structural biology is a discipline within the life sciences, one that investigates the molecular basis of life by discovering and interpreting the shapes of macromolecules. Structural biologists are choosing harder targets each year as they seek to understand the macromolecular machinery of life. Expertise in a single experimental method, and the associated software, is not enough to solve these systems. Structural biologists first prepare a sample by a complicated series of laboratory operations. Then they visit synchrotrons for X-ray diffraction experiments, and/or use Nuclear Magnetic Resonance instruments, and also make increasing use of Electron Microscopes, as well as a range of other instruments. Next they must process the obtained data with computers.

West-Life has helped by making software available as web apps, and by integrating existing apps into combined pipelines, as the diagram shows. West-Life has thus matched services from the emerging European Open Science Cloud to the emerging research methods in structural biology. The following recently published research all used West-Life web services.

Andres Ramos at the University College of London researched into a recently discovered mechanism for cell–cell communication that is important to the development and function of a broad range of tissues, exosomal miRNA transfer. They discovered that the protein Syncrip recognise a class of miRNA using a novel RNA binding domain, the NURR domain, selecting those RNAs for exosomal loading. miRNA transfer via exosomes has been implicated in a number of diseases, including cardiomyopathies, neurological diseases and cancers, and a deeper understanding of the selectivity of the mechanism may lead lead to new treatments.

Birthe Kragelund, University of Copenhagen, studied the antibiotic Gentamicin, which is used to treat some severe infections. Unfortunately it has side effects in the kidneys and ears. She and her collaborators showed how it enters the patient's own cells, by interaction with the cell receptor megalin. This insight may support the structure-based design of a complementary drug to block the side effects.

Nelly Morellet, University of Paris, studied an enzyme called piggyBac transposase, which is promising as a tool for genetic therapies. They discovered how it binds to and reacts with DNA, which will be helpful for the design of more powerful variants of the enzyme.

Zorica Latinović at the Jožef Stefan Institute in Ljubljana studied a glycoprotein in the venom of the horn-nosed viper. They identified one of its targets in the blood coagulation system, showing

that it belongs to a class of anticoagulants which are intensively sought, as they would be safer for medical application than existing drug. The glycoprotein is too large to be suitable for oral use, so the investigated the structural details of its anticoagulant activity, discovering details that could guide the design of a small molecule that could be administered in a tablet rather than by injection.

References

Fruzsina Hobor, Andre Dallmann, Neil J. Ball, Carla Cicchini, Cecilia Battistelli, Roksana W. Ogrodowicz, Evangelos Christodoulou, Stephen R. Martin, Alfredo Castello, Marco Tripodi, Ian A. Taylor & Andres Ramos, A cryptic RNA-binding domain mediates Syncrip recognition and exosomal partitioning of miRNA targets, 10.1038/s41467-018-03182

Robert Dagil, Charlotte O'Shea, Anders Nykjær, Alexandre M. J. J. Bonvin, Birthe B. Kragelund, Gentamicin Binds to the Megalin Receptor as a Competitive Inhibitor Using the Common Ligand Binding Motif of Complement Type Repeats 10.1074/jbc.M112.434159

Nelly Morellet, Xianghong Li, Silke A Wieninger, Jennifer L Taylor, Julien Bischerour, Séverine Moriau, Ewen Lescop, Benjamin Bardiaux, Nathalie Mathy, Nadine Assrir, Mireille Bétermier, Michael Nilges, Alison B Hickman, Fred Dyda, Nancy L Craig, Eric Guittet, Sequence-specific DNA binding activity of the cross-brace zinc finger motif of the piggyBac transposase 10.1093/nar/gky044

Zorica Latinović, Adrijana Leonardi, Lidija Kovačič, Cho Koh, Jernej Šribar, Alenka Bakija, Divi Venkateswarlu, R. Kini, Igor Križaj, The First Intrinsic Tenase Complex Inhibitor with Serine Protease Structure Offers a New Perspective in Anticoagulant Therapy 10.1055/s-0038-1669785

1.4. Access provisions to Research Infrastructures

This is not applicable to west-Life, which is not an experimental project.

2. Update of the plan for exploitation and dissemination of result

There is no update. The plan was implemented and is reported in the WP2 report above.

3. Update of the data management plan

There is no update. This is not an experimental project, so it does not generate data. The new software codes developed by West-Life are open source - see https://github.com/h2020-westlife-eu. Some West-Life deliveries have been achieved by extensions to existing codes. All of these are free for academic use, but not all of them are open source.

4. Follow-up of recommendations and comments from previous review(s)

The reviewers gave "Recommendations concerning the period covered by the report" as follows:

- Recommendations are related to project management in relation to enhancing the project visibility and the general measuring of Key performance Indicators and the risk identification and mitigations report.
- No actions of resubmission of deliverables or milestones are requested, although comprehensive collation and presentation through the project website of those deliverables that have been completed is requested.

These are addressed point by point below, along with their "Recommendations concerning future work".

Management

Visibility

We have improved the formal reporting in the project website, making it easier to find the project deliverables and milestone results, here https://about.west-life.eu/network/west-life/about/workpackages/deliverables and https://about.west-life.eu/network/west-life/about/workpackages/milestones.

Ongoing improvements to the web site have also addressed the reviewers' comments about usability.

KPIs

The reviewers commented *"… the Key Performance Indicators measuring must be enhanced."* After reviewing the guidance offered to European projects, we made a new analysis of possible KPIs. These are reported in sections 1.2.4.3 and 1.2.5.3.

Risks

The reviewers noted that *"Risk identification and mitigations … must be enhanced."* A new risk analyses was been carried out on the occasion of an All Partners Meeting, resulting in an updated risk assessment document. The document is made available on https://docs.google.com/spreadsheets/d/119B80kzo2_5enJ42qaztAmY5jXQk7-9PPMbIKOfyU5w/edit?usp=sharing_eil&ts=5954e9ba to which all partners were invited to contribute.

Presentation through the project website of deliverables

Deliverable and Milestone reports are now available in Zenodo, with links at https://about.west-life.eu/network/west-life/about/work-packages/deliverables and https://about.west-life.eu/network/west-life/about/work-packages/milestones.

Recommendations concerning future work

IPR Management

The reviewers also said *"West-Life needs to generate an IPR management plan, which we see as an essential requirement for attracting the commercial engagement that may be needed to provide for financial sustainability beyond the project life-time."*

The IP newly created by West-Life is be available on an Open Access basis, in line with the general IP policy of H2020. Where West-Life is facilitating extensions to existing software, an existing license applies. The Description of Action gave some details: "The structural biology field has a large number of data analysis programs, with a range of deployment routes and licensing conditions … For example, the CCP4 software suite for X-ray crystallography is licenced free to non-profit users and at a modest cost to for-profit users. Notwithstanding this, the CCP4 suite is a large collection of programs, some of which are available separately under open source licences. Another example of open source application software is XMIPP for cryoEM from CSIC, while the HADDOCK software from UU is free of charge for non commercial users."

Potential commercial support for West-Life is being considered as part of the sustainability plan.We recognise that having clear IPR management of the West-Life components is an important precursor to such future activities. We should also note here that West-Life as eInfrastructure project is offering access to the services and various software, but does not own any software.

Commercial use of web services depends among other things on a clear understanding that the service provider claims no ownership of the results. However, giving such a declaration creates a contract between user and provider, so potentially creating liability for the provider. We therefore developed an Acceptable Use Policy, to which researchers assent when registering for a West-Life userid. This is available at https://auth.west-life.eu/aai/AUP.pdf. We also made West-Life services compliant with the GDPR.

Outreach

The reviewers pointed out *"Dissemination, training and outreach need to be improved, summarized, displayed and reported… For these tasks, some aspects of the web site should be improved."* The improvements to the web site include enhanced reporting of dissemination here https://about.west-life.eu/network/west-life/outreach. There are more details elsewhere in this report.

Sustainability

The reviewers observed *"… the long-term operation and, importantly, the economical sustainability planning needs to be further considered and strengthened. … the different models should be enumerated, the model most appropriate for West-Life should be determined, and the resulting strategy defined. … Discuss internally and go to other members of the e-infrastructure community that are already advanced in this area and get help. These decisions should not be delayed to the end of the project…."*

Partner Instruct has organised a meeting on 13th October 2017 which addressed options for sustainability of both West-Life and iNEXT services beyond the project end dates and formed the key structure for the sustainability report. The report included plans for the longevity of structural biology infra- and e-services in terms of technical maintenance and updating, and plans for the continued financial and operational support for them.

Some of the West-Life services that have been in operation for many years making use of the EGI e-Infrastructures have been selected as thematic services in the new EOSC-Hub H2020 project, which started in January 2018. This is providing additional support for their operation and continuous improvement beyond West-Life.

Cloud

The reviewers also said that *"It does not appear that all possible options of cloud resources to support the project have been considered, and those that have been chosen are not sufficiently well integrated into coherent systems. The choice of cloud computing infrastructure should be taken as an opportunity for positioning of the West-Life platform in the EU scientific research einfrastructure ecosystem."*

West-Life has been making use of the EGI FedCloud resources, which are covered by a SLA with EGI. Those are freely accessible and there has not been any need to look for other providers.

Partner CSIC has developed ScipionCloud, a gateway for running Scipion software on private and public clouds. ScipionCloud is available as an Amazon Machine Image (AMI), that can be used on AWS platform, and as a EGI Virtual Appliance. ScipionCloud has been prepared and tested using resources at the CESNET-MetaCloud, IISAS GPUCloud and INFN-PADOVA-STACK sites due to

their high availability of instances and support. ScipionCloud is directly accesible from West-Life pages under the VM section at https://about.west-life.eu/network/westlife/services/vm.

Partner MU held a workshop in April 2017 on automated deployment of services (application portals in particular) using cloud orchestration techniques. The last part of the workshop was done in collaboration with Indigo Datacloud project. All the workshop materials are available at http://internal-wiki.west-life.eu/w/index.php?title=Cloud_Orchestration_Training. Earlier, Andrei Tsaregorodtsev, EGI, trained West-Life developers in the use of DIRAC.

Big Data

The reviewers say: *"The scope and substance of the work package that relates to "big data" needs to be refined and collaborative work with other more experienced projects is recommended. … it will require additional discussion among partners to see what they can do to avoid duplications."*

Chris Morris attended session led by OpenMinted during the Open Science Fair 2017 in Athens, and Rob Firth attended their concluding conference in May 2018 in Brussels. We produced a named entity recognizer for residues (amino acids within a protein chain). EuropePMC has included it in their text mining platform.

We also performed some pilot work to use Convolutional Neural Networks for cryo-Electron Microscopy. STFC has now advertised a post to continue this work with internal funding.

5. Deviations from Annex 1 and Annex 2

5.1 Tasks

All tasks were implemented.

5.2 Use of resources

Partner EMBL made an adjustment to the costs for RP1, as a result of booking travel costs for trips that happened in that period, and recalculation of salary costs.

There were several variations from plan in RP2. Partner CSIC required some extra person months to achieve their objectives in WP7. This need arose due to the loss of a highly experienced skilled member of staff, and the change to a child-care part time state of a key person, highly experienced both in cryo-EM and cloud computing. Although new personnel was promptly hired to compensate for these labour changes, CSIC faced the expense of training a new recruit in the rapidly changing field of computation for electron microscopy. Given that the project as a whole is within budget, we propose that these extra costs are admitted.

Several partners found that salary costs were lower than planned and used the opportunity to allocate extra time: UU in WP5, MU in WP3 and WP7, and CIRMMP in WP2, WP4, and WP5. Salary costs depend on the number of years of research expertise of the hired researchers. A total of 116 extra person months was contributed.

Nearly ten person months were transferred to NKI to enable success in tasks 7.1, 7.2, and 7.4, in particular D7.3 and the integration of the Virtual Folder (WP6) with the developments in PDB-REDO and Protein-CCD (WP7).

There were some other, smaller divergences from the original plan. The larger of them are explained as follows. Instruct contributed extra effort to WP2, notably to enhancing the project web site. Six person months in WP6 were transferred to CIRMMP, so that they could become the first users of the repository (D6.2). Luna's contribution to WP4 was reallocated to WP5. INFN required more effort than planned to achieve the objectives of WP4, and funded this by a transfer from travel expenses, as agreed in an amendment.

5.2.1 Unforeseen subcontracting

No work was subcontracted.

5.2.2 Unforeseen use of in kind contribution from third party against payment or free of charges

There were no unforeseen in kind contributions.