

Deliverable D3.4

Project Title:	World-wide E-infrastructure for structural biology	
Project Acronym:	West-Life	
Grant agreement no.:	675858	
Deliverable title:	Interim report on science policies	
WP No.	3	
Lead Beneficiary:	6 CIRMMP	
WP Title	Networking	
Contractual delivery date:	30 April 2017	
Actual delivery date:	28 April 2017	
WP leader:	Lucia Banci	CIRMMP
Contributing partners:	CIRMMP (P6), INSTRUCT (P7)	

Contents

Contents

1	Executive summary.....	3
2.1	The context.....	3
2.2	Science policies.....	5
2.2.1	The European Open Science Cloud.....	5
2.2.2	The FET Flagship program.....	8
2.2.3	Large research infrastructures.....	9
2.3	Conclusions.....	12

1 Executive summary

The West-Life project is providing virtual access to its services free of charge. Consequently, its operation beyond the lifetime of the current EC project depends on the capability of the consortium partners to attract, individually or in collaboration, further funding. Among the various factors affecting this capability, besides the necessary interaction with the user community and the software developers' community, there is the relevance of the services provided in the broad scientific scenario and in particular with respect to current and future science policies. The current report aims to provide an overview of the latter context, which constitutes a possible framework for the future development of West-Life.

The central role of large-scale research infrastructures (RIs), most notably those on the ESFRI roadmaps, and large-scale projects, such as FET Flagships, within European policies is particularly prominent. At the same time, the momentum of Open Science policies has increased significantly, as witnessed by the implementation of the European Open Science Cloud (EOSC) pilot project. The intersection of these two main strategic lines is the development of best practices and implementation of actual technical solutions for the management and stewardship of the high-quality data produced by RIs. West-Life is well positioned to tackle these challenges.

2 Detailed report on the deliverable

2.1 The context

Interactions with key stakeholders at both national and European level are fundamental in order to follow the development of science policy that could affect the SB research community and thus more specifically West-Life.

A relevant environment is that of the CORBEL project (<http://www.corbel-project.eu/>), an H2020 initiative of eleven ESFRI biological and medical research infrastructures (BMS RIs), aiming to create a platform for harmonised user access to their technological platforms, sample collections and other services. The eleven RIs in CORBEL are listed here: <http://www.corbel->



project.eu/participants.html. The INSTRUCT hub (P7), as well as the INSTRUCT centers CSIC (P5) and CIRMMP (P6), which are all members of the West-Life consortium, are partners of CORBEL and have been active in several of its activities. CORBEL is not explicitly involved in the development or the monitoring of science policies. Nevertheless, because of the high scientific and economic impact of ESFRI BMS RIs this consortium is a very important stakeholder and influencer with respect to European policy making.

The role of West-Life partner representatives in numerous national and international committees of funding agencies, large RPOs and EC-appointed science policy initiatives and working groups has provided the bulk of the information outlined in the present report. For example, Lucia Banci (P6) has been involved in the Building Blocks of Life assessment committee of the Netherlands Organisation for Scientific Research (NWO) and in the Italian committee of the National Research Council (CNR) for the evaluation of Excellent Projects. She is also the Italian representative in the Strategic Working Group “Health and Food” of ESFRI and contributed to the ESFRI Roadmap 2016 (see section 2.2.3). The Group is now working on the 2018 roadmap. Alexandre Bonvin (P8) was recently nominated as scientific advisory board member of the Molecular and Cellular Structure (MCS) cluster at the EMBL-EBI in Hinxton, which is providing a direct link and input channel with the Protein Data Bank in Europe (PDBe), the major repository of structural data. The first 2017 meeting of the MCS SAB took place on April 2017. Alexandre Bonvin is also member of the advisory board of the compute and data infrastructure programme of SURFSara, the main e-Infrastructure provider in the Netherlands. In the context of large structural biology projects, the role of Jose-Maria Carazo (P5) as scientific advisory board member of the European Synchrotron Radiation Facility (ESRF) helps assuring a close link between data producers and data analyzers (like WestLife). Furthermore, in relation to International initiatives, Jose-Maria Carazo is part of the International Scientific Advisory Committee of the Australian Center for Advanced Molecular Imaging and one of five members of the Top Users Program of the Center for Protein Science of the Chinese Academy of Sciences (at Shanghai).

Regarding science policy activities that are funded by or involve directly the European Commission, it is important to mention the project INRoad (<http://inroad.eu/>), which recently involved Lucia Banci (P6) as member of one external advisory panel (Reflection Group).

INRoad is a Horizon 2020 CSA project that aims at supporting RI policy development in terms of exchange of best practices for national roadmap drafting and evaluation procedures for RI, in order to promote comparability and synchronization of national procedures and harmonization of evaluation mechanisms.

Another Europe-wide channel to expose the West-Life partnership to current developments in science policy is via the participation in the RDA events, both at national and plenary levels. West-Life partners have taken part in activities such as the various meetings of the “RDA meets researchers” that have been organized throughout Europe (<https://www.rd-alliance.org/events.html>). Recently, a session of the Interest Group in Structural Biology took place within the Ninth RDA plenary (<https://www.rd-alliance.org/ig-structural-biology-rda-9th-plenary-meeting>). One of the aims of the session was to foster further discussion on the life cycle of structural data. Beyond this specific aspect, it allowed West-Life partners to engage in discussions with several international players in the broad field of data science. Among these, opportunities for Instruct (P7) and West-Life to collaborate with OpenAIRE (<https://www.openaire.eu/>) were discussed.

2.2 Science policies

2.2.1 The European Open Science Cloud

With the adoption of the Digital Single Markets strategy in 2015, the European Commission announced the launch of a cloud for research data. The 'European Open Science Cloud' aims to create a trusted environment for hosting and processing research data to support EU science in its global leading role. The EC appointed a High Level Expert Group on the European Open Science Cloud to advise on the scientific services to be provided on the cloud and on its governance structure. The EOSC is not a mere repository for scientific data, what was conceived from its very beginning as a stakeholder-driven infrastructure servicing science and innovation including all the technical elements necessary to provide a wide range of user-oriented services, ranging from data management to computation and data analytics. Collaboration with industry is an important aspect of the economic sustainability of the EOSC, implemented via private-public partnerships. A number of challenges for the successful development and implementation of the EOSC were identified in a workshop held in November 2015, whereas a workshop on the governance and funding model of the EOSC, involving Member State representatives, took place in June 2016 (<http://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud-workshop>).

Importantly, the EOSC is publicly funded and governed, thus guaranteeing its research-centric focus.

The EOSC initiative reinforces Open Science, Open Innovation and Open to the World policies, by enabling best practices of global data findability and accessibility (FAIR data), helping researchers get their data skills recognised and rewarded, allowing easier replicability of results. In other words, the EOSC constitutes the EU contribution to a future, global Internet of FAIR Data and Services underpinned by open protocols. A further important aspect of the EOSC is systematic and professional data management and long-term stewardship of scientific data assets and services. It is foreseen that half a million 'core data scientists' are needed to make the most of open research data in Europe.

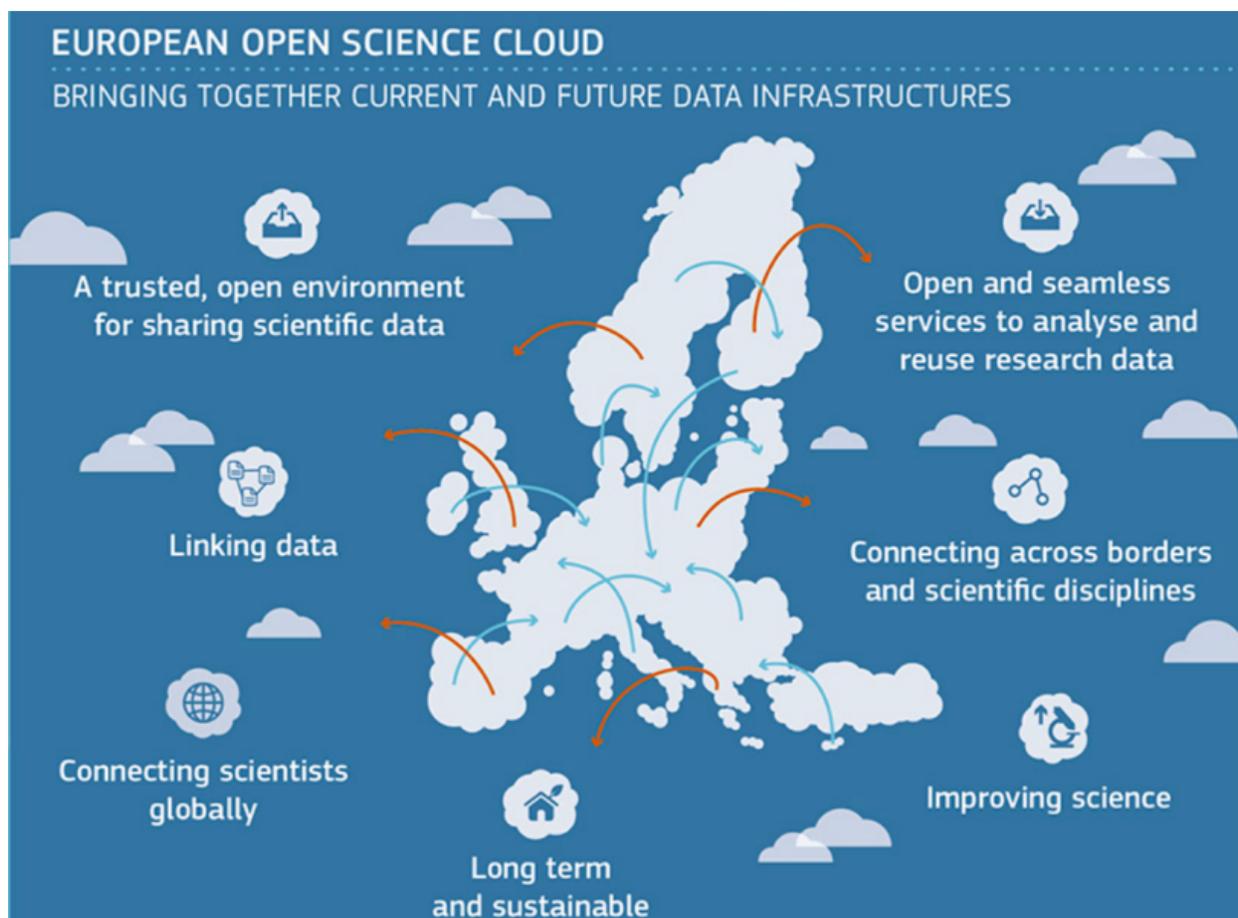


Figure 1. A schematic representation of the European Open Science Cloud initiative

Open and seamless services, such as those provided by the West-Life partnership, that allow users from all over the world to manage, analyse and reuse research data are a key component of the EOSC. The mature services of West-Life (the WeNMR ones in production since a long time) have been included as thematic services in the eInfra_12 EOSC-Hub proposal, submitted by EGI/EUDAT/INDIGO-Datacloud.

The Communication to the European Parliament officially defining the strategy towards the European Cloud Initiative and the EOSC was released about one year ago (COM/2016/0178, <http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:52016DC0178>). Notably, it identified the following sources of EU funding for the European Cloud Initiative:

- Horizon 2020 Framework Programme for Research and Innovation (Horizon 2020)
- Connecting Europe Facility (CEF)
- European Structural and Investment Funds (ESIF)

- European Fund for Strategic Investments (EFSI)

Existing funding under Horizon 2020 has been allocated to support the European Open Science Cloud and to kick-start the European Data Infrastructure. The initial estimate of the required additional public and private investment was €4.7 billion over 5 years. This includes €3.5 billion for data infrastructure, €1 billion for a large-scale EU-wide Quantum Technologies flagship (to be launched in 2018) and €0.2 billion for actions on widening access and building trust. On this basis, in October 2016 the High Level Expert Group recommended to close discussions about the ‘perceived need’ of a science cloud and to take immediate action on the EOSC in close concert with Member States, building on existing capacity and expertise. This was done in a report entitled “Realising the European Open Science Cloud” (http://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf). Consequently, the EOSC Pilot (<https://eoscpiilot.eu/>) was launched on January 1st 2017 until the end of 2018. During these two years, it will deliver a number of building blocks, including technology demonstrators and a first draft of a multi-stakeholder structure which can accommodate a mix of different users.

West-Life has an optimal position to collect the requirements and organize the interaction of research groups that are active in structural biology with the EOSC pilot and, later on, with the fully implemented EOSC. In particular, West-Life has the opportunity to provide tools and formalize approaches to allow structural biologists to properly manage and curate their data. Such approaches could be integrated to the computational services that are central to the offer of West-Life in order to streamline the experience of users and facilitate their data sharing in the EOSC.

2.2.2 The FET Flagship program

The FET Flagship programme represents a huge investment by the EC in strategic and collective research and innovation. It aims at undertaking research and innovation activities that will transform science into technology with significant industrial and societal impacts, exceeding the sum of individual efforts found in national initiatives. Through supporting such work, Flagships will foster the growth and enhance the economic impact of European companies,

including SMEs and start-ups. A Panel of high-level experts was convened to undertake an interim evaluation of the instrument, and its implementation through the already ongoing Graphene and Human Brain Flagship Projects. The outcome of this evaluation highlighted that the Flagships have demonstrated their effectiveness in delivering excellent science, whereas their future effectiveness in supporting innovation still needs to be demonstrated. It is thus apparent that it not yet easy to achieve, in one instrument, both excellent science and excellent innovation. This aspect links to the efficiency of the strategic and operational management structure of so large initiatives, also in terms of linking to national initiatives. Another recommendation was that funding models and funding time-scales should be adapted to the special nature of the Flagships, in particular by implementing a longer funding cycle to improve the 'in year' flexibility needed to respond to changing circumstances and new opportunities in a way that does not require the funding to be spread evenly across financial years. Longer time-scales would also help reduce the amount of time spent by Flagship participants in looking for further funding.

The scientific interests of West-Life are closer to the interests of the Human Brain Project (HBP). The HBP provided strong examples of interdisciplinary work integrating modelling with data sets, an activity that is reminiscent of the approaches in integrative modelling of macromolecular structures of higher order and complexity. Of the six ICT research infrastructure platforms launched by the HBP in 2016, three are attracting their first users. Again, the HBP is working on the development of a European federated infrastructure for data hosting. As noted previously, this is an area of application that is central to the activities of West-Life, with regards to data management for integrative structural biology. Structural biology provides the molecular basis for the understanding of all cellular processes, and thus is important for the multi-scale nature of the HBP. Interactions in this sense were pioneered by the MoBrain competence center of the Horizon 2020 project EGI-Engage (<https://www.egi.eu/>), which is a federated e-Infrastructure set up to provide advanced computing services for research and innovation. In this context, it is also relevant that the evaluation of the Flagships recommended a closer interaction with H2020 projects, such as West-Life, and beyond. The Flagships should also engage research communities more broadly.

2.2.3 Large research infrastructures

“Research infrastructures (RIs) are facilities, resources and services used by the science community to conduct research and foster innovation. RIs can be single-sited (a single resource at a single location), distributed (a network of distributed resources), or virtual (the service is provided electronically).” (https://ec.europa.eu/research/infrastructures/index_en.cfm?pg=about)

Large RIs of European relevance have been addressed extensively by the European Strategy Forum on Research Infrastructures (ESFRI). ESFRI aims to develop the integration of RIs across Europe and strengthen them to enhance the quality of the activities of European scientists and attract to Europe the best researchers from around the world (<http://www.esfri.eu/about-esfri>). A cornerstone of ESFRI activities has been the publication of the first Roadmap for pan-European research infrastructures in 2006. The ESFRI Roadmap has been updated a few times, the last of which in 2016 (<http://www.esfri.eu/roadmap-2016>). The ESFRI Roadmap has had a strong impact on the development of National roadmaps for RIs by the Member States. In particular, the provision of a specific EU Regulation allowing for a bespoke legal entity for European Research Infrastructures (the ERIC) emphasizes the relevance of RIs to Europe. Importantly, a key performance measure of ERICs is their impact on science policy, through their research work on challenges which have high socio-economic impact. In addition to ESFRI, Horizon2020 has deployed numerous initiatives regarding RIs. Specifically, Horizon2020 has contributed to structuring the Research Infrastructures landscape by an extensive array of activities, including support to the Design, Preparatory phase, Integration, and Implementation of RIs. Coordination and improvement of international outreach are also taken into account, for example by the CORBEL and B3Africa initiatives, respectively. As pointed out by the Science Europe (<http://www.scienceurope.org/>) association of European Research Funding Organisations (RFO) and Research Performing Organisations (RPO) large-scale RIs cannot typically be built by one Member State alone. In this sense the action of Horizon 2020 has been crucial to facilitate the development of such facilities, by catalysing common investment from Member States by supporting the preparatory phase of RIs. Among other recommendations, Science Europe proposed that further support should target emerging communities, which are at the very start of their RI lifecycle. Another important point is that synergies and complementarities should be investigated both at scientific and technological level, and at the funding level, e.g. by combining different European funding schemes or

European and National funding schemes. Science Europe also advocated for the EC to implement cross-directorate collaborative mechanisms as a means to maximise the opportunities for RI management and operation. For example, a collaborative mechanism between the RI and Marie Skłodowska-Curie Actions programmes could enable the mobility of RI managers and operators across Europe and beyond, with the principal aim of knowledge transfer and sharing best practice. Specific attention (and funding) could be on training and education, data management, and specific incentives (for example schemes to hire business development personnel or to support staff mobility).

The EC-funded InRoad initiative (<http://inroad.eu/>) pointed out that the nature and complexity of the societal challenges require a global approach for the design and operation of RI. International cooperation is also highly strategic when pooling of resources is necessary for construction and operation of RI and in order to achieve scientific excellence. This is in complete agreement with the views expressed by Science Europe, as outlined above. The European Commission supports the internationalization of the Research Infrastructures and addresses international cooperation also as a tool to support or complement the EU external policy and contribute to Science Diplomacy. The InRoad project tries to contribute to a better harmonisation and synchronisation of priority-setting, funding evaluation and life-cycle management of EU Research Infrastructures (RI) at European and national level. Another specific focus is on the comparison and synchronization of the procedures for the development and evaluation of national roadmaps for RIs.

Globally, it appears that the ESFRI has been the main driver of strategic thinking about RI development and, later on, RI prioritization, mainly thanks to the publication of its roadmaps and updates thereof. The ESFRI roadmap fostered the development of corresponding roadmaps at the national level in most Member States (<http://www.esfri.eu/national-roadmaps>). The MERIL portal has been developed with funding by the European Commission to provide a mapping of the existing RIs across Europe (<https://portal.meril.eu/meril/> and more recently, the EU-funded RISCAPE (<http://www.riscape.eu/> was tasked with mapping major research facilities worldwide to identify synergies and opportunities for engagement between RIs beyond Europe. The funding scenario for RIs development, implementation and operation is relatively complex, as it necessarily relies on the successful synergistic combination of multiple funding schemes that are often managed at different levels (regional, national, European). The multifaceted scientific and technological landscape of RIs in Europe warrants the implementation of specific actions to

improve training of RI operators, development and sharing of best practices for operation and management aspects and, in general, to provide career options in this context. This could in principle be coordinated with similar developments at the national level and linked to the development of harmonized procedures for the preparation and evaluation of national RI programs. Such harmonization would be beneficial also to enable a more effective use of different funding schemes for RI support, especially by reducing administrative overheads. In parallel, there is consensus also on the desirability of a better exploitation of technical and scientific synergies among RIs (taking into account their different levels of maturity, where relevant). Production of high-quality data and data management/stewardship have been explicitly pinpointed by the recent “Science Europe Policy Brief” and the previous “Science Europe Position Statement”. West-Life is providing data services for the Structural Biology community, in collaboration with the ESFRI Instruct. Such services typically feature two components, one related to data management and the other related to the procurement of computing resources. Solutions developed in regards of both components are not inherently tied to the Structural Biology community but could be packaged in a form that allows reuse by other RIs thereby fostering the adoption of common protocols and simultaneously providing a potential basis for data use and analysis across different RIs. Notably, there have been already steps in this direction at the level of RI management, exemplified by the adoption of the ARIA system for the management of scientific applications by the CORBEL project for its first open call.

Many large scale structural biology facilities (e.g. synchrotrons) provide international access including new modes of automated control and remote access. There is a growing need for multi-site, virtual participation in experiments that can benefit national and international collaboration. Harmonised data management protocols are key to achieving this and will complement initiatives to establish jointly owned equipment that generated large amounts of primary and processed data. This has been recognized by the expansive e-Infrastructure programme of Horizon2020 which focuses a significant funding commitment in support of this objective and in turn influences national policies on e-infrastructure investment and compatibility. Indeed, research data infrastructure should support the development of collaboration environments using trusted data, from data capture through to simulation and modelling and to dissemination, including data storage and management, cloud services and associated expertise and skills development. These e-infrastructures should be comparable to

and be able to interface and leverage further development of international opportunities (e.g. through EOSC) and other global initiatives with emphasis on interoperability.

2.3 Conclusions

Current science policies clearly highlight the central role of RIs. In particular, the international dimension of large-scale RIs, including those on the ESFRI roadmaps, is recognized. There are extensive needs for a corresponding synchronization of several aspects in the implementation, operation, and evaluation of existing and future RIs. Such needs include training RI staff and operators on aspects such as data stewardship and data management.

At the same time, the European policies towards Open Science have gained significant momentum, culminating with the funding of the EOSC pilot project. There is an important link between such policies and the development of practical solutions for data stewardship by RI operators, in order to enable the sharing of the high-quality data produced by the RIs.

West-Life is well positioned to address this link by providing procedures and technical solutions, initially validated by the ESFRI Instruct and the structural biology community and subsequently disseminate these achievements to other RIs, particularly within the biomedical sciences.