H2020 EINFRA-5-2015

www.bioexcel.eu

Project Number 675728

D3.5 – Consultancy Modalities and Funding Options, Final Update

WP3: Consultancy & User Groups

Copyright© 2015-2018 The partners of the BioExcel Consortium

This work is licensed under a Creative Commons Attribution 4.0 International License.

The opinions of the authors expressed in this document do not necessarily reflect the official opinion of the BioExcel partners nor of the European Commission.

Document Information

Deliverable Number	D3.5
Deliverable Name	Consultancy Modalities and Funding Options, Final Update
Due Date	2018-12-31 (PM38)
Deliverable Lead	UEDIN
Authors	Adam Carter (UEDIN), Ian Harrow (IHC), Rossen Apostolov (KTH), Alexandre Bonvin (UU), Vera Matser (EMBL-EBI), Vytautas Gapsys (MPG), Emiliano Ippoliti (Juelich), Adrien Melquiond (UU), Holger Noack (FT), Stian Soiland-Reyes (UNIMAN), Mikael Trellet (UU)
Keywords	Modalities, Services, Community, Consultancy, Interaction
WP	WP3
Nature	Report
Dissemination Level	Public
Final Version Date	2018-12-18
Reviewed by	Rosa M Badia (BSC), Vera Matser (EMBL-EBI), Rossen Apostolov (KTH), PMB
MGT Board Approval	2018-12-22

Document History

Partner	Date	Comments	Version
UEDIN	2018-12-04	Complete first draft carried over from Google Doc	0.1
UEDIN	2018-12-12	Response to comments from other partners. Improved formatting.	0.2
UEDIN	2018-12-17	Updates in response to internal review	0.3
UEDIN	2018-12-19	Release candidate	0.4

Executive Summary

This document follows on from D3.3 "Consultancy Modalities and Funding Options, Half-Time Update". It is intended to be read in conjunction with D3.3, as it does not aim to repeat the content therein. The structure of the documents is similar, to allow cross-referencing. It aims to relate the more theoretical list of modalities presented in D3.3 with specific activities undertaken in the project.

BioExcel has used several modalities in activities aimed at the wider community including web pages, a discussion/feedback forum, workflow portals of different types (including software-as-a-service, and those providing access to workflow components), white papers, webinars, face-to-face meetings of various kinds (including Interest Group meetings, workshops, a community forum meeting and multiple face-to-face training courses), online meetings and calls.

Fewer of the modalities categorised as being important for professional products and services have been evaluated in BioExcel-1, as they are particularly difficult to offer successfully with a relatively small pool of experts. One modality however, *bespoke training*, has been piloted with an industry customer and deemed to be very successful.

The modalities described above were complemented by informal discussions at meetings and other academic interactions of the partners. Project partner UU also took part in the D3R challenge, raising their profile and that of the Centre of Excellence (CoE).

The modalities that have been used successfully in BioExcel-1 will be built on through the addition of several other modalities in BioExcel-2, which includes tasks focused on use cases, community participation, in-depth support, work on standards and best practice, and more.

BioExcel has explored a range of modalities during its work with the user community. Learning from BioExcel-1, BioExcel-2 plans to continue the most successful activities, and adopt several new modalities in order to continue its work to engage with the community. The CoE is clearly moving towards a second phase: In BioExcel-1 relationships have been established, and groundwork laid. In BioExcel-2 we will build on this to engage more deeply with the community.

A key lesson from BioExcel-1 has been the value of using multiple modalities to work with a given individual, organisation or group. There are certainly different modalities that work better with different groups, but probably more important is having the flexibility to combine these in ways which fit with the work that is being undertaken over a period of time. In BioExcel-2 we have sought to continue this process by placing a greater emphasis on a support team that can balance its time between engagement events and more specific support in response to user need.

Contents

1 INTRODUCTION 7

1.1 CONTEXT 7
1.2 DOCUMENT AIMS 7
1.3 DOCUMENT STRUCTURE 7
1.4 TYPES OF CONSULTANCY MODALITY 8

2 CORE SERVICES FOR THE WIDER COMMUNITY 9

2.1 WEB PAGES 9
2.2 DISCUSSION / FEEDBACK FORUM 9
2.3 HELPDESK 10
2.4 WORKFLOW PORTAL 10
2.4.1 HADDOCK WEBSERVER 10
2.4.2 PMX WEBSERVER 11
2.4.3 BUILDING BLOCK LIBRARY WEBSERVER 11
2.4.4 BIOEXCEL CLOUD PORTAL (AND EGI APP DB, BIOTOOLS) 12
2.5 FAQ RESOURCE 12
2.6 WHITE PAPERS & BEST PRACTICE GUIDES 12
2.7 WEBINARS 13
2.8 FACE-TO-FACE MEETINGS WITH USERS 15
2.8.1 INTEREST GROUP WORKSHOPS & MEETINGS 15
2.8.2 COMMUNITY FORUM MEETING 16
2.8.3 ACADEMIC CONFERENCE 17
2.8.4 TRAINING COURSES & SCHOOLS 19
2.9 ONLINE MEETINGS & CALLS 19
2.10 VIRTUAL TRAINING 20
2.11 IN-DEPTH REMOTE CONSULTANCY 21
2.12 STRATEGIC PARTNERSHIPS 21
2.13 COLLABORATIVE PARTNERSHIPS 21
2.14 CONSORTIUM CONTACTS 21
2.15 JOINING PROJECTS AS A PARTNER 21
2.16 DROP-INS / SURGERIES / CONSULTATIONS 22

3 PROFESSIONAL PRODUCTS AND SERVICES 22

3.1 TECHNICAL SUPPORT (FIELD SPECIALIST) 22
3.2 SCIENTIFIC SUPPORT (APPLICATION EXPERT) 22
3.3 BESPOKE TRAINING 22
3.4 MATCHMAKING ALLIANCES 23

4 OTHER RELEVANT MODALITIES 23

4.1 INFORMAL DISCUSSIONS & MEETINGS WITH INDIVIDUAL GROUPS 23
4.2 INDUSTRY PARTNERS 24
4.3 CHALLENGES 25
4.4 PRE-COMPETITIVE COLLABORATIVE ALLIANCE 25

5 FUTURE PLANS 25

5.1 USE CASES 25
5.2 COMMUNITY PARTICIPATION 26
5.3 IN-DEPTH SUPPORT AND WORK IN STANDARDS & BEST PRACTICE 26
5.4 USER & COMMUNITY NEEDS ANALYSIS 26

5.5 CHALLENGES **27**
5.6 RELATED ACTIVITIES IN TRAINING & DISSEMINATION **27**

6 CONCLUSION **27**

1 Introduction

1.1 Context

A key aspect of WP3 in the BioExcel project is the interaction between the Centre of Excellence (CoE) and the wider biomolecular research community. As we have previously described in D3.3, we are using the term "Consultancy Modalities" in a general sense to capture the many ways that we have been interacting with the community.

The biomolecular research community is large and heterogeneous, and for many purposes it is better to think of several different, albeit overlapping, communities with which BioExcel interacts. BioExcel will therefore need to adopt a number of different modalities to make the most of our interactions with the different communities. A good understanding of the ways in which we are working with users is therefore very useful as the Centre moves into its fourth year, supported by a second phase of funding from the European Commission. In this document we use "BioExcel" to refer to the Centre of Excellence and to the EC project that supports it, for which this document is a deliverable. Where we want to explicitly differentiate between the first project and the follow-on project we use BioExcel-1 and BioExcel-2 respectively.

1.2 Document Aims

This document is intended to complement D3.3[1], and provide an update on consultancy modalities and funding options. Whereas D3.3 was more theoretical, this document aims to connect the list of possible consultancy modalities with specific activities that we have undertaken in the BioExcel project. It can be seen that in BioExcel-1 we have explored many but not all of the modalities in D3.3. In BioExcel-2 we have structured the project plans to build on the work undertaken in BioExcel-1 with additional effort to provide in-depth support and engagement.

It should be noted that we continue to differentiate "activities" from "modalities". To give an example, the single activity comprising the interactions between a company and the Industry Interest Group has involved online meetings, face-to-face meetings and site visits. Here the work with in the Interest Group has involved (at least) three different modalities. Similarly, "services" offered by the CoE could be provided through multiple modalities.

The document should be read in conjunction with D3.3 since it aims not to repeat the information included there.

1.3 Document Structure

We have used a similar structure to that of D3.3 to allow for comparison with the earlier document.

In Section 1 we will recap the modalities from D3.3 and their categorisation based on expected audience. In Sections 2-4 we will look at each modality in turn. For each modality we recap briefly what is involved in the offering, and then discuss in more detail what aspects of the modality have been explored in the project. In Section 5, we will

[1] BioExcel Deliverable D3.3, *Consultancy Modalities and Funding Options, Half-time Update*, https://doi.org/10.5281/zenodo.574613

briefly discuss plans for BioExcel-2, relating these to the modalities described in the rest of the document.

1.4 Types of Consultancy Modality

In Table 1 in D3.3 we categorised a number of consultancy modalities. An update to this list is included below in Table 1. The modalities are categorised according to the audience for which they are most relevant, but it should be noted several modalities would be relevant to multiple audiences.

Table 1: Types of Consultancy Modality

Core Consultancy Services

Target audience: Wider community*

Web site & online resources
- Web pages
- Discussion & feedback forum
- Helpdesk
- Workflow portal
- FAQ resource
- White papers & best practice guides
- Webinars
- Online Meetings & Calls
- Virtual training

In-Depth Remote Consultancy

Face to Face Meetings with users
- Interest Group workshops & meetings
- Community Forum Meeting
- Academic Conferences
- Training courses & schools

Partnerships & Collaborations
- Strategic partnerships
- Collaborative partnerships
- Consortium Contacts
- Joining Projects as a Partner

Consultations & Drop Ins

** as noted below, these modalities are often offered free, but do not need to be free*

Professional Consultancy Services

Target audience: Paying customers

Technical support
- Field specialist

Scientific support
- Application expert

Bespoke training

Matchmaking alliances

Other Relevant Modalities

Informal discussions with individual groups

Industry Partners

Challenges

Pre-Competitive Collaborative Alliance

2 Core Services for the Wider Community

In this section, we consider those modalities that would most commonly be used to deliver services for the wider community.

2.1 Web Pages

In keeping with our broad definition of consultancy modalities, we start with the most basic means of interacting with the community, namely the project's web site.

The BioExcel website *bioexcel.eu* was built at the very start of the project. It acts as the main platform for the center for dissemination of results and outreach to the communities. In the first phase of the project we undertook two major restructuring updates. Information has grown considerably and we are currently (Nov. 2018) working on a new major update of the website.

The main website is primarily a way to broadcast information to our users and potential users, but it also provides the mechanism for users to register for our newsletter and Interest Groups and, via GoToWebinar, for our webinars.

Being the centre's main outreach and dissemination platform, it is important that the web site continues to operate in the long-term. Maintenance costs are minimal (approx. 150 Euro/year) with occasional redesign updates. Substantial effort is required, however, for content curation, especially as activities increase and service offerings mature.

2.2 Discussion / Feedback Forum

BioExcel launched a feedback forum at the start of the project at *ask.bioexcel.eu*, using the Discourse platform. Each of the core software application was assigned a specific category, and certain other categories were initially set up for various Interest Groups. For HADDOCK this platform became the main support endpoint. For other codes (e.g. GROMACS, PMX) the forum complemented other support channels.

The forum has worked well for some communities, particularly those established around codes. In this case, the communities are large enough to establish a critical mass of users who contribute to and benefit from discussions.

As of today, HADDOCK category groups more than 240 topics including 81 with 500+ views and 17 with 1000+ views. Three of the core developers of HADDOCK actively contribute to the forum. An estimated 1-2 hours per week have been spent by the developers in discussions on the forum. It should be noted that considerable expertise is required to address many of the questions posted, so it is necessary for this effort to be provided by people familiar with the code and its application. These contributors have also ensured that the content remains well organised (reassigning, merging and deleting categories, for example). In addition, there is also some additional effort required to ensure that the platform is hosted on a stable platform with established backup and recovery processes.

2.3 Helpdesk

In BioExcel we considered providing a helpdesk, but rejected this as it requires considerable effort and availability to deliver a helpdesk service which can be relied upon by users. The discussion forum described above provides most of the value of a helpdesk, and responses have typically been quick.

A helpdesk *could* be funded as part of a paid-for support package for paying customers but a fairly large charge would be required to support this as it would need to be the primary role of those individuals delivering the service.

2.4 Workflow Portal

There are a number of related offerings that could be considered to be workflow portals, and BioExcel has worked on a number of these.

One powerful way to engage users is to offer access to a flagship software through a dedicated portal, providing the software as a service. Supported by BioExcel, both HADDOCK and PMX are offered through web portals which provide access to standard workflows, including some post-processing and analysis. These web portals aim to simplify the setup and run of a modelling task by providing intuitive web interfaces for the users to input their data. These portals offer an effective way of supporting the use of complex software. By providing the software on a specific platform it is easier to support; those who provide the support (for example through the forum described above) are familiar with the platform on which the software is being run, reducing the scope for complications due to external factors.

The fixed workflows provided by these portals offer notable usability benefits, particularly to entry-level users. More advanced users can benefit from building their own workflows, and this has been supported in BioExcel through other types of workflow portal, such as the BioExcel Building Block library and the BioExcel Cloud Portal. The portals also offer the potential to evaluate how the software is being used, although this data has not been systematically collected during the project. We now consider each of these portals in turn.

2.4.1 HADDOCK webserver

During the BioExcel project, the HADDOCK team has developed a new web portal built on recent technologies to renew and update its aging processing workflow. The new web portal introduces new features and improves the user experience. It relies on a new version of HADDOCK, namely HADDOCK2.4[2] and integrates the new capabilities of the software. In parallel of applying the necessary levels of security and performance, the design of the new portal followed some user feedback gathered via a former survey conducted by the Computational Structural Biology team at UU along the last few years. The technical aspects of the new developments are explained and detailed in the roadmap and development

[2] http://haddock.science.uu.nl/services/HADDOCK2.2

plan (D1.3)[3] and the document accompanying the final code releases in the project (D1.5)[4].

2.4.2 PMX webserver

The PMX webserver[5] has been created to facilitate free energy calculation setup involving amino acid and nucleic acid mutations. The web-based infrastructure allows for an easy access to the most important command line features of the PMX package and enables rapid generation of hybrid structures/topologies without the need to install any additional software.

The service is free to use, and openly available. Since its launch the service has been accessed by ~300 unique users. This figure is an estimate based on the number of distinct IP addresses from which the service has been accessed. A decision was made not to require registration for the service in order to ensure a low barrier for entry to the system. There is potentially value in learning who is using the system, but the trade-off would have to be considered closely to determine whether it would pose too high a barrier for users of the system.

The main interest in using the webserver has been expressed by academic groups, while industrial collaborators prefer an in-house installation of PMX due to the strict security regulations on their side. The web server, however, allows industry users to evaluate the software before installing it locally.

In the future, the capabilities of the web server will be expanded by including ligand modification support and non-equilibrium free energy calculation setup. Having these features implemented would allow attracting new users primarily interested in the computer aided drug design.

An estimate for the cost of providing this service can be made by considering the amount of effort dedicated during BioExcel-1. Approximately 6PM of support and development effort was required to deliver this service in BioExcel-1.

2.4.3 Building Block Library webserver

A new web server[6], thought as an entry point for the BioExcel building blocks software library (biobb) is being developed in a mini-project launched during year 3 of BioExcel. The server will show the many different ways with which our biobb can be used (Python, Jupyter Notebooks, KNIME, Galaxy, CWL, etc.) and at the same time offering a chosen set of fixed workflows to be explored and run from the same portal.

Prototype building blocks as nodes on the KNIME analytics platform or as a fixed workflow with a Graphical User Interface (GUI) on our biobb web server were presented during a recent Focus Group meeting. This comprised of selected

[3] BioExcel Deliverable D1.3, *Roadmap of future hardware and long-term development plan for each pilot application*, https://doi.org/10.5281/zenodo.574605

[4] BioExcel Deliverable D1.5, *Final project releases of pilot applications*, in preparation

[5] http://pmx.mpibpc.mpg.de

[6] See https://bioexcel.eu/research/projects/biobb_web_server

representatives from the pharmaceutical companies, AstraZeneca, MedImmune and UCB at a venue in London.

2.4.4 BioExcel Cloud Portal (and EGI App DB, biotools)

BioExcel has also created the BioExcel Cloud Portal[7], working closely with ELIXIR to integrate the portal with the ELIXIR identity infrastructure. The cloud portal allows automatic deployment of pre-configured Virtual Machines on different cloud infrastructures. Access to the portal is done through ELIXIR AAI authentication, with a Single Sign On process with which just typical credentials from external tools (Google, LinkedIn, ORCID) are needed to login. Permissions for VMs deployment must be granted from the Cloud Portal managers. The BioExcel Cloud Portal has been tested in a workshop with 20 students. They were able to deploy VMs from the portal, work with them, and delete the instances, just using the web portal. Due to the great response to the technology behind the portal from the community and the workshop attendees, BioExcel-2 is planning to make use of the portal for different training activities, using pre-configured VMs and Docker containers.

2.5 FAQ Resource

A Frequently Asked Questions (FAQ) resource was not created in BioExcel-1. This is another modality that is under consideration for BioExcel-2. An evaluation of the effort required versus the benefit to the user will need to be conducted; an important factor is to consider what is the best source and material for FAQs (e.g. forum posts or tickets from a helpdesk) or if we need to build it from scratch. The content will need to be developed by the code experts and a decision will need to be made about whether such a resource should be developed slowly (e.g. as issues arise) or if an initial push should be made to provide as much coverage as feasible. The effort to maintain the FAQ should not be underestimated, especially with regular software releases. One way to lower the curation burden is to allow for community curation e.g. flag out of date posts). It is important to note that the content would be at the level of the software codes rather than the CoE. The FAQ could either be a traditional web-based FAQ or we could consider integrating this into the centre's Knowledge Resource Centre (KRC)[8] though this would require widening the scope of the KRC, which currently holds training resources (e.g. courses and tutorials) rather than FAQ-style material.

2.6 White Papers & Best Practice Guides

White papers and best practice guides were not widely used as a modality in BioExcel-1, but they have been planned as an activity in BioExcel-2. WP1 produced a white paper on Software Development Best Practices[9] which has already been downloaded 68 times from Zenodo since its publication in March. A good white paper requires significant effort to write, but once published it can have a wide audience. Two academic papers[10,11] published during the project have a best-practice element to them, but academic papers are slightly different as a modality.

---

[7] https://bioexcel.ebi.ac.uk
[8] http://krc.bioexcel.eu
[9] https://doi.org/10.5281/zenodo.1194634
[10] https://doi.org/10.1007/s41019-017-0050-4
[11] https://doi.org/10.5281/zenodo.1473157

A white paper on "Data Sharing for Biomolecular Simulations" is being planned as an outcome of a workshop with the same title held in Stockholm in Nov. 2019. This is a good example of how two modalities – workshops and white papers – work in tandem: The workshop itself emerged from a very active Twitter discussion around the question "What is the best way to share simulations data?" amongst a group of about a dozen people from different institutes in EU and US. The discussion evolved considerably and the group started working on a joint document, but they needed a forum to discuss the topic face-to-face (many of them have never met each other before the Twitter discussion). The researcher who started the original discussion reached out to BioExcel during one of our dissemination events (PRACE Days in Slovenia) for support. Thus, we funded a workshop with 30 participants who over 2.5 days engaged in active discussions. A follow-up event is also planned to take place in the USA towards the end of 2019. The main organisers of the initiative were very grateful for the support received from BioExcel and requested that they can carry on the BioExcel stamp on the initiative. Recordings from the event will be available on our YouTube channel.

2.7 Webinars

Throughout the project we have run a webinar series[12]. We are now typically running once or twice per month, depending on the time of year. The topics cover all areas related to the project activities - new software and best practices for its usage, use case success stories, workflow platforms and their applications. We have developed a robust and efficient process for organization, setup and post-processing of the webinars. By the end of BioExcel-1 we have hosted a total of 34 webinars. We were happy to see that other CoEs followed our example and started similar webinar series.

Webinars provide two different interfaces with the community: they are viewed by a wide group of people (over 2,300 YouTube views at the time of writing), and they also provide us a means to engage with other researchers and practitioners by allowing them to present their work as a BioExcel webinar. They also serve to support collaborative relationships (for example, we have run webinars in conjunction with the Pistoia Alliance and the IMAbs consortium).

In BioExcel, our Webinar series has also shown itself to be a useful seed activity from which other opportunities arise. One example of this is the webinar related to BioSimSpace, itself seeded from an informal discussion. Being invited to do a BioExcel webinar is not just a stand-alone dissemination activity, in this case the webinar worked as a door opener for further collaboration with the BioSimSpace group for the BioExcel/MolSSI Workflow workshop (see section *Informal Discussions* below). In another case, the webinar series helped to build a relationship with the developers of IBM Spectrum LSF. This collaboration started as an ad-hoc conversation with the IBM Spectrum LSF developers after the Common Workflow Language community learnt IBM was working on a CWL implementation. They were invited to the CWL community call, and then BioExcel was used as a vehicle to further mature the collaboration by inviting them to do a

[12] https://bioexcel.eu/webinars

webinar [13] . The webinar accelerated IBM's move to release their *cwlexec* implementation as open source. The BioExcel Webinar video is even embedded as documentation on the *cwlexec* GitHub repository[14]. This CWL implementation is now a strong contender to running large-scale CWL workflows on HPC, and has been considered by the ELIXIR CWL implementation study[15] for use on EMBL-EBI's LSF installation.

The systematic way that we run the webinars make further collaborations and developments possible. As our webinars are promoted and hosted publicly, then archived on YouTube[16] and SlideShare[17], they are available beyond the one-off session and can be reused by third-party projects as part of their tutorial and dissemination activities. We regularly tweet "reruns" of recorded webinars[18]. We have found that the Webinars can act as a forcing function for emerging projects to present themselves publicly, as we invite "early access" presentations of work in progress, enabling wider collaborations to initiate as a project is growing. Many times, the BioExcel webinar is the first webinar that project have done publicly, which they are keen to partake in. Projects and invited presenters also tend to promote[19] their BioExcel webinar to their own communities, which are then as a side-effect introduced to BioExcel. We therefore always start webinars with a general introduction to BioExcel and how to join our interest groups and newsletters.

During the project an opportunity arose through discussions with the company BiKi technologies, who were interested in presenting some of their work[20] to a BioExcel audience. To pilot the offering of website hosting as a service, we requested that BiKi made a payment to cover costs. On this occasion, we requested €270/webinar (based on an estimated 4 hours of work). In fact, despite increasing efficiency in terms of delivery as the project has progressed, this is probably on the low-side for the effort involved. A more realistic estimate would be that there is ~1 hour of planning (discussing content, finding a suitable date, etc.), ~1 hour of set-up an publicity (in GoToWebinar, on the project web site, and using Twitter), ~0.5 hour practice session with the host to ensure that the speaker is familiar with the platform, and to check connectivity and audio quality, ~2 hours for delivery (1 hour for the host, and 1 hour for a backup host at a second site), ~2 hours post-processing, posting online and publicising the post. In addition to the ~6.5 person hours, the cost of the GoToWebinar licence (€89/month) should be factored in to the cost of delivery.

---

[13] https://bioexcel.eu/webinar-cwlexec-a-new-open-source-tool-to-run-cwl-workflows-on-lsf
[14] https://github.com/IBMSpectrumComputing/cwlexec - cwlexec
[15] https://www.elixir-europe.org/about-us/implementation-studies/cwl-2018
[16] https://www.youtube.com/channel/UCd2hq8Q_ZyTU4YafEhweE5g/videos
[17] https://www.slideshare.net/BioExcel/presentations
[18] e.g., https://twitter.com/BioExcelCoE/status/1049290818329202689
[19] e.g., https://twitter.com/MolSSI_NSF/status/1044574871655731200
[20] https://bioexcel.eu/webinar-finding-a-trade-off-between-speed-and-accuracy-in-protein-ligand-binding-description-2018-04-26

Funding for webinars is most likely to come from core BioExcel funding, although there is the possibility to offer companies access to the series as described above. From the viewers' point of view, webinars are usually free to join. Conceivably, it would be possible to host webinars that could only be viewed by paying customers (either stand-alone, or as part of a package of services). For the live webinars, this could be managed using the GoToWebinar interface to control access to the live webinar and recordings, but there would be some additional workload involved in administering this, as there is no automatic link between the participants in GoToWebinar and any system which could control membership or payment. There might be scope to link access with a fee using the related project, *GoToTraining*, which is being considered for use in WP4 in BioExcel-2. Using GoToWebinar's interface for recordings would also potentially constrain what could be done in terms of post-processing of videos. An additional possible source of funding which has not been explored is sponsorship of webinars by commercial entities (that is, having a commercial sponsor that is *not* the presenter of the webinar, unlike the case with BiKi described above). Finally, it would be possible (subject to viewership numbers) to recoup a small amount of funding from the YouTube Partner Programme[21].

In the future, we also plan to run webinars on non-scientific topics, e.g. on equality, diversity and inclusivity; best practices for training; how to organize workshops etc.

2.8 Face-to-Face Meetings with Users

Meetings provide a useful means to engage with the wider community. There are several different ways to split up the possible modalities for meetings.

2.8.1 Interest Group Workshops & Meetings

During BioExcel we have organised and contributed to a number of workshops and meetings. Some of these have had direct connections with the project's interest groups, others have been organised in response to the needs of the wider community and opportunities identified by project partners. These meetings can be physically located at partners' sites, at a central location, or co-located with other meetings organised in the community. Using the frameworks of larger external workshops and conferences on related topics has two main advantages: 1) reducing costs for BioExcel, and 2) having the chance to promote the IG and BioExcel's activities to a larger audience. This was for example the strategy of the Hybrid Method Interest Group. This IG arranged its first face-to-face meeting in 2017 within the *First Workshop on Hybrid Methods in Molecular Simulation* [22] organised by the European HPC-LEAP joint doctorate program and the University of Cagliari, Italy and a second one in 2018 within the Human Brain Project (HBP) event Annual Meeting of the Co-Design Project 6 (CDP6) - Modeling for Drug Discovery[23] that was held in Jülich, Germany (Some members of the IG are actively collaborating with HBP researchers in the context of BioExcel's pilot use cases). The second meeting was attended by a combination of those who attended the first meeting and new members. The face-to-face meetings therefore served to

[21] See, e.g., https://support.google.com/youtube/answer/72857?hl=en-GB

[22] http://www.fz-juelich.de/ias/ias-5/EN/Conferences/Hybrid_Methods/Home.html

[23] http://www.fz-juelich.de/ias/ias-5/EN/Conferences/CDP6/Home.html

both engage with new members of the community and maintain relationships with existing members. This pair of meetings also illustrates how BioExcel funding was used to bring together a group of people who subsequently met again without funding from BioExcel (other than the time investment from BioExcel's participants). These IG meetings also led on to webinars.

The Workflows IG arranged a workflow training course[24] at BSC, introducing participants to workflow systems like *Galaxy, KNIME, Nextflow, PyCOMPSs* and *Common Workflow Language*. A side-effect of this meeting was that representatives of multiple workflow system communities were gathered in one place, and new collaborations and developments were spawned as a result, for instance Nextflow support for CWL[25] or BioExcel building blocks as KNIME nodes[26]. The joint symposium[27] with US-based *Molecular Sciences Software Institute* (MolSSI)[28] at the *Platform for Advanced Scientific Computing* (PASC18)[29] conference spawned further MolSSI collaborations leading to a two-part Workshop on Workflows in Biomolecular Simulations[30] that takes the opposite approach of the first workflow meeting in BSC, this time more of a deep-dive, putting real scientific use-cases to the test for workflow system providers to collaboratively find new solutions to scalability challenges identified at PASC.
The Training Interest Group also conducted a face-to-face meeting as part of the ELIXIR learning pathways study, as well as a meeting as part of the Community Forum (see below).

Fairly significant effort is required to organise, *and contribute to* face-to-face IG meetings. Additionally, ongoing effort investment is often required to maintain relationships between the meetings. In BioExcel-1 it was this — as opposed to the budget to pay for events — that proved to be the limiting factor. In BioExcel-2 we have sought to address this by increasing the effort available in WP3 and reducing the events budget. This is discussed further in Section 5.

2.8.2 Community Forum Meeting

In November 2017 we organised a Community Forum meeting[31] in Amsterdam. It was billed as "a networking event with a programme of small interactive working groups aligned with the BioExcel Interest Groups". The idea for the meeting came from trying to make best use of the effort available to organise interest group meetings. By co-locating these, it meant that a single venue could be organised, and the contributing project partners would need only travel to a single event.

The meeting was designed around a number of parallel sessions where each IG could hold meetings or mini-workshops. The general idea worked well, but experience from the event showed that there was, in fact, too much parallelism;

[24] https://bioexcel.eu/events/bioexcel-workflow-training-for-computational-biomolecular-research
[25] https://www.nextflow.io/blog/2017/nextflow-and-cwl.html
[26] https://bioexcel.eu/research/projects/biobb_knime
[27] https://bioexcel.eu/events/bioexcel-molssi-symposium-at-pasc-2018
[28] http://www.molssi.org/
[29] https://pasc18.pasc-conference.org/
[30] https://bioexcel.eu/events/bioexcel-molssi-workshop-on-workflows-in-biomolecular-simulations
[31] https://bioexcel.eu/events/bioexcel-community-forum-22-23-november-2017

up to four sessions were scheduled in parallel and there were several times where people wanted to be in more than one session at the same time or who preferred to attend a session that did not correspond to that which they were expected to attend. The meeting was also designed to provide the opportunity for the participants to move easily between the various sessions, and this aspect worked well. The meeting venue, the Lloyd Hotel, provided a number of interlinked, open, working spaces ("platforms") with a relaxed atmosphere which encouraged networking and discussion. This aspect worked well, and although it made on-the-day organisation hard work, we adapted the meeting on-the-fly to adapt to the interests of those present by combining some sessions and compressing others to allow people to attend the sessions that which they had an interest in.

One aspect of the meeting that worked particularly well resulted from the decision to bring everyone to the same event: the scientific experts and lead developers from the projects' codes could attend the industry round-table, for example. Time to talk with these experts was appreciated by the participants.

This kind of meeting needed considerable planning requiring significant effort and a long lead-time. Although the meeting was successful, the turn-out was lower than we hoped for. Registration numbers were low, and this was compounded by a number of no-shows. At a post event review, it was identified that there were several things that could be done differently in order to increase participation, including: (i) advertising the event earlier, (ii) advertising a *detailed* programme for the event earlier, (iii) charging a small commitment fee to register to ensure that registrants attended. A repeat of the community forum event was initially planned to run during 2018, but other project priorities dictated that this would not happen. Instead, it was decided to participate in a number of other individual events later in the year.

2.8.3 Academic Conference

Conferences were not considered specifically in D3.3, but they have been discussed in BioExcel-1 and proposed as an activity during BioExcel-2, so we discuss them here.

Whilst there was definitely value in the Community Forum meeting described above, its unusual nature made it harder to promote and made it harder to attract a large audience. When considering the best way that a CoE such as BioExcel could have an impact on a large number of people through a face-to-face event, it was concluded that a more traditional academic conference would provide a good platform to advance the goals of the CoE. Researchers have a good understanding of what a conference is. The opportunity to present their work in a talk or a poster, with the possibility that it could be published in conference proceedings provides a clear incentive for them to participate. It is also understood by the institutions (academic and commercial) that fund their work and that often make decisions regarding the value of attending an event.

In BioExcel-2 we have proposed a Community Conference. We use the term "community" to highlight that the aim would be to organise the conference with interested parties from outside BioExcel. The aim is *not* to have a "BioExcel"

conference, but to have a conference on "Computational Biomolecular Research" or a more specific field in this area.
Participation in external conferences is also considered a good way to engage with the wider community. In general, speakers at a conference will have to pay their own registration fee, travel and accommodation, unless they have been invited by the conference organiser. The cost is likely to be between €1750 and €3000 depending on the location of the conference. Such conferences provide good exposure for both the speaker and BioExcel.

Many external conferences also have a programme of official satellite meetings. These tend to come at no cost to the organiser of the satellite meeting and may include a limited number of free speaker registrations. All speakers are responsible for their own travel and accommodation costs, which would be paid by the project for invited external speakers. Participants pay to attend the session but that fee is paid to the conference organiser who is responsible for venue and catering costs. The costs for the CoE in running these meetings therefore typically consist of registration fees, support for invited speakers, and the costs of sending staff. This modality was used in BioExcel-1. Two successful satellite meetings at ISMB/ECCB[32],[33] were organised as part of WP4's dissemination activities, but these provided good opportunities for engagement with the participants and accessibility was supported by travel grants for some participants supported by WP3's events budget.

Without organising a specific event, it can also be beneficial to sponsor and support other conferences in areas that BioExcel aims to support. Conference sponsorship is typically defined in tiers, with more benefits being added in each tier; sponsorship with exhibition stand typically starts from €2000, rising to €6000 for additional benefits that could be worth considering. With the lower tier sponsorship package there will likely be a need to buy additional exhibitor or conference passes to project members attending the booth. Even when offering sponsorships or financial support for conferences there is effort involved in the administration of the payments. In BioExcel-1 we have supported a number of conferences and conference workshops including RO2018[34] and NETTAB 2018[35] (in which BioExcel was also represented during the talks[36]). Participation at these events has raised the profile of BioExcel but more concrete benefits have been harder to quantify. These events did, however, fall later in the project and follow-on work and opportunities may arrive at a later date.

[32] https://bioexcel.eu/events/sig-bioexcel-1
[33] https://bioexcel.eu/events/sig-bioexcel-2
[34] RO2018, Workshop at IEEE eScience 2018, Amsterdam, Netherlands (29 October 2018), http://www.researchobject.org/ro2018
[35] *NETTAB 2018: Building a FAIR Bioinformatics environment*, Genova, Italy (22-24 October 2018) http://www.igst.it/nettab/2018
[36] A. Hospital Gasch, *BioExcel & ELIXIR: towards interoperable and reproducible biomolecular research workflows*. Online at http://www.igst.it/nettab/2018/files/2018/10/NETTAB2018_Hospital.pdf

For larger conferences, there is also value in participating in a conference with a poster, talk or booth. In BioExcel-1 we attended a number of conferences; these are detailed in D4.4[37] and D4.6[38].

2.8.4 Training courses & schools

BioExcel's training activities are discussed in more detail in D4.5[39] and D4.6. In BioExcel we have run six face-to-face training courses of varying length. Here we briefly discuss how these courses and schools work as a modality for interacting with the community and provide some indicative costs. In BioExcel we have aligned the training activities with other activities. The workflows training event described above which evolved from work in the workflows IG is a good example of this. Offering training is a very good way to directly engage with the community. Costs can be high, particularly for initial set-up, but courses and schools can be charged for to offset costs, or even generate income.

The effort required to run face-to-face training courses is high, both in administrative effort as well as content development effort. Here we provide some examples of the costs, the turnaround time is approximately 6 months and total costs heavily depend on the number of days in the training event and what is included in the day rate. As a guideline, the cost per training day for 30 participants is between 775 EUR and 3400 EUR. The lower figure assumes local trainers only, no dinner included and the use of a free-of-charge venue; the higher figure includes a dedicated venue, with 2 European trainers (needing 1 night accommodation and travel) and a course dinner for participants and trainers.

Face-to-face training is incredibly valuable due to the high level of interaction between the participants and the trainers as well as the participants with each other. In general, participants are willing to pay for face-to-face training, though full costing can be prohibitory depending on the country of residence and income level. We anticipate that face-to-face training events will always be part of the training activities of the CoE. Costs can be managed by a variety of measures, e.g. charging a registration fee, joint events with other initiatives. We would advocate charging a registration fee as a commitment fee, regrettably it is our experience that the no-show rate can be very high if the training is offered completely free of charge. Travel grants can be offered to remove the financial barriers for participants.

2.9 Online Meetings & Calls

As a distributed centre of excellence, there is real value in online meetings and calls in lieu of face-to-face meetings. In many cases, these are more feasible, particularly for some audiences such as industry representatives with limited time to offer to activities that are not at the core of their current work. In BioExcel,

---

[37] BioExcel Deliverable D4.4, *Dissemination Report and Updated Plan*, https://doi.org/10.5281/zenodo.574614

[38] BioExcel Deliverable D4.6, *Final Report on Dissemination and Training*. In preparation. To be made available at https://bioexcel.eu.

[39] BioExcel Deliverable D4.5, *Training Report and Updated Plan*, https://doi.org/10.5281/zenodo.574620

two interest groups in particular made use of online meetings and calls to engage with the relevant communities.

For the Industry IG, these calls were a good mechanism for bringing together multiple parties to discuss certain particular points. In several cases, these IG calls were the first direct interaction that these users had had with the CoE, after they had been invited to contribute. Unlike BioExcel's webinars, these meetings were by-invitation, as opposed to being openly advertised. They worked well as a way to meet a number of contacts and to get some general feedback on their needs, but it proved difficult to maintain engagement due to a great extent in those who were interested to be involved with the centre having specific questions that would be better addressed bilaterally, or asynchronously (in the case where input was required from developers and application experts in BioExcel who were not always available at the same times as the industry users). Later in the project, effort in the Industry IG was directed more towards establishing deeper links, for instance through site visits.

Online calls were also useful for maintaining existing relationships. Much of the collaborative work that took place during the preparation for the IMABs ITN grant proposal was conducted through such calls.

For the Training IG, online meetings and calls allowed BioExcel to work with collaborators, such as the ELIXIR Training Coordinators Group (TrCG)[40], the TrCG special interest group around HPC training and the IMABs consortium mentioned above. In addition, the Training IG is an external collaborator on an ELIXIR implementation study[41], where regular online meetings were used to maintain progress between face-to-face meetings.

2.10 Virtual Training

In the project to date, BioExcel's training activities have focused on face-to-face training, however plans to offer what we are referring to as "Virtual Training" are under active development and will be trialed in 2019. In D3.3 we referred to the Remote Training in §2.1.9. This has been moved from the section focused on online resources to highlight the fact that BioExcel's plans in this area will use the version of the modality which goes beyond putting material online, and also involves synchronous support activities.

In 2019 we anticipate running sessions with pre-recorded lectures, with pre-arranged live discussions sessions about the lectures and help sessions for the tutorials. We anticipate that the training session will be supported by either a VM or Docker container where the necessary software, data (potentially links to data) and recorded lectures are available.

This concept is still to be tested but it allows us to efficiently use the trainer time resources and to provide training support over a longer period. We believe that this will provide good value for the participants while at the same time lowering

[40] https://www.elixir-europe.org/platforms/training/how-organised
[41] https://www.elixir-europe.org/about-us/implementation-studies/learning-paths-2018

the participation costs (e.g. no travel costs, accommodation costs). We anticipate running this style of training alongside our face-to-face training events. This virtual training is currently being considered for development into a service that we can charge for. This is discussed further in D5.4.

2.11 In-Depth Remote Consultancy

This is a modality that we have explored in BioExcel-1 in the context of some of the project's pilot use cases, but in most cases, the remote interaction supported face-to-face consultancy. As described in §2.2 of D3.3, this modality is often coupled with helpdesk provision. Since there are no plans to offer a helpdesk in BioExcel-2, then there would need to be an alternative mechanism to gain access to this in-depth remote consultancy. This is something that could be offered through the support teams, as discussed further in Section 5.

2.12 Strategic partnerships

We have set up a clear process for the establishment of strategic partnerships with selected organizations. At the time of writing (Nov. 2018) we have signed eight. The partnership agreements outline the main directions of joint activities. Although community building in general is a task that requires considerable time and effort, the strong links that the Centre's partners have with other initiatives make the establishment of strategic partnerships relatively straightforward, with low cost and big leverage, particularly in terms of outreach and potential future collaborations.

The establishment of the partnerships gives a clear signal (both internally and externally) of those interests that the parties have in common. They expand the network and visibility of the CoE and provide opportunities for joint participation in future funding schemes.

2.13 Collaborative partnerships

Similar to the "Strategic partnerships", the collaborative ones establish close links with selected organizations. In this latter case though, there is a clear commitment for specific activities with a concrete timeline and milestones roadmap. Such partnerships help to structure efforts with the other party, and bring engagements to a deeper level. They add focus and stimulate long-lasting interactions. In general, we consider it beneficial to extend the number of such partnerships.

2.14 Consortium contacts

The informal professional networks of all partners are vital for extending the outreach, dissemination of results, popularizing service and product offerings, and overall positioning of the centre. Those networks of contacts have been and will continue to be heavily utilized in future.

2.15 Joining projects as a partner

Joint projects are understandably very beneficial for all involved parties. Such projects often arise as part of the strategic and collaborative partnerships. A lot of public funding schemes require multiple partners with complementary expertise, which is a strong incentive for both BioExcel and third parties to engage jointly. Such partnerships are very important for the long-term sustainability of the centre and we will continue exploring future opportunities.

2.16 Drop-Ins / Surgeries / Consultations

This modality has not been explored in BioExcel-1. The closest thing that we offered here were webinars related to the core codes with the Q&A session offering the opportunity to ask questions. Assuming a webinar series is running, the costs to offer such sessions would be minimal, other than the time investment from those with expertise to answer questions. With increased support effort available in BioExcel-2, it is recommended that these modalities be piloted then.

3 Professional Products and Services

In this section we consider a number of modalities that are more likely to be of interest to paying customers. Since this classification was made in D3.3, funding was secured for BioExcel-2, a project to help support the CoE for another three years. In this proposal, we have introduced a task to offer more in-depth support, and thus some of the modalities described below will be explored in the context of a publicly-funded project, however they appear to have considerable immediate value and could potentially be self-sustaining activities in the context of a wider CoE.

3.1 Technical Support (Field Specialist)

Detailed in-depth technical support was not offered in BioExcel-1. The centre, through participation in the HADDOCK forum on AskBioExcel and the GROMACS mailing lists provided a lot of support, but the type of detailed support that might traditionally be considered to be consultancy was not offered. These modalities are challenging to offer, but could be things that BioExcel could offer in the future. The critical point here is being able to find individuals with sufficient expertise and sufficient time to offer this support at short notice. It is either necessary to have a sufficient body of consultancy work to ensure that the consultancy or support team is kept permanently busy, or to have sufficient flexibility to drop or postpone other core development work so that people with the in-depth knowledge required to interact using this modality can be available when required. The situation in BioExcel-2 is improved from BioExcel-1 in that there is more effort available for this kind of support activity. There is also the intention to try to build a support team that overlaps with (or at least works alongside) development work in WP1. This remains a difficult thing to do, but it is hoped that it is possible to explore this modality further in the context of BioExcel-2, and beyond. As a long-term goal, the partners in the Centre want to be in a position to offer this modality.

3.2 Scientific Support (Application Expert)

Similarly to Technical Support, Scientific support was not a service that was offered in BioExcel-1. Much of what was described for technical support also holds true in this case. As with the technical support described above, there is scope to explore this modality further in the in-depth support task in BioExcel-2.

3.3 Bespoke Training

On the 26th of April 2018, we gave a full training day tailored to UCB pharma (five people from their computational group attended the training). We received a request from UCB about a specialised training to make use of HADDOCK on local

computers, and specifically, how to encode different type of experimental information for use in HADDOCK. A new specific tutorial was created explaining how to install and use a local version of latest 2.4 release of HADDOCK (available online[42], but not yet linked from our web pages), which took us a few days. In the future, if a different company would contact us with the same request and motivations, we would certainly be willing to provide the same service but we would first point them to the online tutorial. As an outcome of this training, there were discussions about possible future consulting on specific projects. Furthermore, since they purchased a HADDOCK license, they get some support as part of the license agreement for the first year.

A fee was charged for this training to help understand the steps required to negotiate and accept payment. The process ran smoothly and there appears to be scope to charge higher prices for such bespoke training in the future that could comfortably cover the costs of delivery. The costs, of course, vary depending on how related the bespoke training is to the stock tutorials, but the more courses that are run, the lower these development courses are likely to be. Training activities are therefore being carefully considered as services that could help support sustainability of the CoE in future. This is discussed further in D5.4.

3.4 Matchmaking alliances

Since Deliverable 3.3 we have reached out to the Pistoia Alliance (pre-competitive, not for profit organisation) to connect with the existing AbVance project on Antibody 3D structures. It was quickly realised that there was a mutual interest in collaborating on predictive molecular dynamics and docking for interactions between antigen and antibodies. To pursue this, it was agreed to form the integrative modern antibody (IMAbs) consortium which includes BioExcel partners (KTH, Utrecht and EMBL-EBI), academic (University of Oxford) and industry (AZ, MedImmune, and UCB) partners.

The main focus for the consortium was to submit a Marie Curie ITN proposal for funding in 2018. Although the funding was unsuccessful, the consortium has continued to grow, adding more partners and the improved and expanded proposal will be re-submitted to the 2019 round. The AbVance project manager, Dr Richard Norman will be a new partner for BioExcel 2 which includes antibody use cases and user support.

Bringing these parties together is an example of BioExcel acting as a matchmaker. In this case, partners from BioExcel are also participating in the consortium which ultimately would work as a pre-competitive alliance (see below).

4 Other Relevant Modalities

4.1 Informal Discussions & Meetings with individual groups

A modality that was not explicitly listed in D3.3, is informal discussions and meetings with individual groups. This is easily overlooked as a modality, and

[42] http://www.bonvinlab.org/education/HADDOCK-local-tutorial

arguably is a by-product of face-to-face meetings, events and visits, but it is worth noting that informal interactions, during meals, breaks and before events are often key to kick-starting and sustaining collaborations.

Informal discussions and networking are key benefits of BioExcel workshops and training events, particularly as core developers and researchers from BioExcel take part. Informal discussions arising from academic relations have many times been accelerated on a path to maturation thanks to BioExcel initiatives. For instance, UNIMAN was contacted by BioSimSpace developers [43] (CCPBioSim, HECBioSim) at the SSI Collaborations Workshop 2018 [44] because they knew UNIMAN was involved in BioExcel. An informal meeting was set up in Manchester to discuss overlaps between BioSimSpace approach and BioExcel workflows. Thanks to BioExcel as a driver, outcomes of this informal collaboration include a BioExcel webinar on BioSimSpace [45], code collaboration with BSC on BioExcel building blocks (D2.4), a BioSimSpace talk[46] at BioExcel's workshop Sharing Data from Molecular Simulations[47] in Stockholm and strong involvement of HECBioSim at the BioExcel/MolSSI workshop in Barcelona[48]. We are working on partnership MoUs with both HECBioSim and CCPBioSim, and plan to hold a join training session with the latter in early 2019.

4.2 Industry Partners

This modality establishes a longer-term relationship with partners in industry as a pre-competitive group, who may also contribute financially to cover costs.

Since Deliverable 3.3 we have organised an industry breakout session at the Community Forum which was attended by a few delegates from universities and also from the pharmaceutical company, Janssen (Part of Johnson and Johnson Group). This was followed up by a one day site visit to Janssen, near Antwerp. This brought together experts from BioExcel tool developers with molecular modellers at Janssen and two external collaborators. It resulted in understanding their capabilities and needs. Most importantly it led to collaboration on Free Energy Perturbations using the PMX tool.

A similar site visit to UCB near Brussels has also been undertaken which added to our understanding of their specific needs. This also resulted in collaboration on pmx and the docking tool, HADDOCK. We have also conducted a Focus Group in London which brought together representatives from AstraZeneca, MedImmune and UCB. This three-hour group setting enabled us to gain a deep understanding of challenges and needs through structured discussion. It also allowed the BioExcel developers to present on prototype ideas for workflow building blocks, fixed workflow web server and high-performance virtual screening. This was

[43] https://www.biosimspace.org

[44] https://www.software.ac.uk/cw18

[45] https://bioexcel.eu/webinar-biosimspace-filling-the-gaps-between-molecular-simulation-codes-2018-06-27

[46] https://drive.google.com/file/d/1J1BMOkcrjum_lz9cSQz0k2-pVmWdPx_7/view

[47] https://bioexcel.eu/events/workshop-on-sharing-data-from-molecular-simulations

[48] https://bioexcel.eu/events/bioexcel-molssi-workshop-on-workflows-in-biomolecular-simulations

followed by gathering detailed feedback from the participants through structured discussion.

4.3 Challenges

Challenges were not discussed in D3.3, but were identified in the project as being a useful way to engage with the community. A successful submission to the challenge can gain a lot of exposure for the team (and hence the CoE) and results in a concrete contribution to advances in the field which are subsequently published. As a stand-alone modality, challenges mostly correspond to competitively working on a common problem. The outcomes from this work are subsequently shared in publications and face-to-face meetings. These challenges do have some significant effort investment, but the exposure is good, and the likelihood of this was sufficient to justify the participation of the partner UU in D3R Grand Challenge 3, supported in part through BioExcel. The publication describing this contribution[49] has been downloaded over 600 times at the time of writing. BioExcel will be a platinum sponsor of the CAPRI challenge in 2019 and will support the organisation of the event: One partner (Prof. Bonvin from Utrecht University) is a member of the CAPRI management board and the meeting will make use of the facilities at EMBL-EBI.

4.4 Pre-competitive collaborative alliance

This is usually a not-for-profit alliance which enables collaboration on pre-competitive projects. It is particularly important for commercial sectors (e.g. pharmaceuticals and agricultural/food) where there is significant risk of collusion between competitors to manipulate favourable market prices.

The IMAbs consortium described above could be considered to be a form of pre-competitive alliance, but at present the work of this alliance has focused on the writing of a proposal together, and so the involvement of BioExcel to date has been more related to the matchmaking modality.

5 Future Plans

BioExcel-2's directly user-facing activities will continue to be focused in WP3, which is entitled "Community, Support and Use Cases" to better reflect the areas in which BioExcel plans to work. In writing the proposal for BioExcel-2 a balance was sought between identifying key activities that we have committed to undertake and to provide a function to allow the project to remain user-driven and responsive to users' needs.

5.1 Use Cases

BioExcel-2 will explore four use-cases. Compared with BioExcel-1, which relied in several cases on the work of others to implement the use cases, BioExcel-2's use-cases will be explored primarily with project effort. They therefore do not correspond so closely with consultancy modalities. That said, they have been developed with likely end-users in mind; for example, Janssen, UCB, MedImmune and Abvance have all expressed interest in the use case "Biomolecular interactions

[49] https://doi.org/10.1007/s10822-018-0148-4

engineering - application to antibody design" and it is likely that interactions with some of these companies will take place in the course of exploring the use case.

5.2 Community Participation

BioExcel-2 has a specific task related to Community Participation. BioExcel will attend and contribute to community events through talks to present the CoE's work, and sessions to channel input from the community. In addition, BioExcel will complement these by organising several events throughout the course of the project, in response to the needs of the community. This task will thus involve several of the face-to-face meeting modalities such as workshops, conferences, and discussion meetings similar to the Interest Group meetings described above (note that in BioExcel-2 it has been decided to phase out formal Interest Groups). A meeting similar to the Community Forum is anticipated to be one of the meetings organised by BioExcel, along with a community conference which will be organised in conjunction with others in the biomolecular research community. As part of this community participation work, a continuation of BioExcel's webinar series is planned. The primary focus here will be on attending and organising events (face-to-face and, where appropriate online) as opposed to engaging with activities in the wider community.

5.3 In-Depth Support and work in Standards & Best Practice

Working with and for others is the scope of two tasks "In-Depth Support" and "Standards and Best Practice". The in-depth support is envisioned as being proactive where possible, as opposed to reacting to simply reacting to helpdesk tickets. This task will allow us to explore the "In Depth Remote Consultancy" modality described above. The expectation here is that a number of pilot activities, defined in conjunction with WP5 to support sustainability, will be undertaken with third-parties to address their problems. These are expected to be smaller in scope and timescale than the Pilot Use Cases from BioExcel-1 or the use cases activities in BioExcel-2 and could each use a number of modalities, including a site visit, followed up with remote in-depth support. This in-depth support activity also has potential to explore other modalities such as matchmaking alliances.

As described above, best practice guides were identified as something that a CoE would be expected to provide. With a specific task in BioExcel-2 there is effort available to do the underlying preparatory work required to write authoritative best practice guides informed, for example, by measuring the performance of the project's core codes on different systems in different circumstances on different problems. This task is thus likely to involve engagement with collaborating partners (such as collaborations in HPC training with ELIXIR) using both face-to-face and online modalities as well as producing end-products such as best practice guides that can be widely accessed by the wider community.

5.4 User & Community Needs Analysis

This task in BioExcel-2 plans to use questionnaires, interviews and interactions in forums at events to collect user needs. This activity will use online questionnaires linked from the project's website and will also rely on face-to-face meetings and the forum to interact the community. This task will also concertedly bring together user data from multiple sources to help keep track of how users are interacting with us through the multiple modalities we are using.

5.5 Challenges

Building on successful participation in the D3R[50] in BioExcel-1, there are plans in BioExcel-2 to participate in future challenges and to sponsor or organise one of the associated meetings. As a first step. BioExcel will support the CAPRI challenge in 2019, as described above.

5.6 Related activities in Training & Dissemination

In addition to WP3's plans, several of the modalities described above will be used during WP4 activities including face-to-face and virtual training activities, and attendance at relevant conferences. There is also a specific task in BioExcel-2, entitled "Collaboration", which will actively engage with related national and international activities to raise awareness of BioExcel services, share best practice and create and assess opportunities for collaborative projects.

6 Conclusion

As can be seen from this document, BioExcel has explored a range of modalities during its work with the user community. Learning from BioExcel-1, BioExcel-2 plans to continue the most successful activities, and adopt several new modalities in order to continue its work to engage with the community.

The CoE is clearly moving towards a second phase: In BioExcel-1 relationships have been established, and groundwork laid. In BioExcel-2 we will build on this to engage more deeply with the community, particularly those with whom we have started to work with in BioExcel-1, including industry contacts. The growing maturity of the centre is reflected in the use of different modalities over time. Those that can reach a broad number of people such as support forums for core codes and webinars have help us demonstrate the value of the CoE, and the pilot use cases have served the dual purpose of providing work that we can showcase to attract the interest of the community and to build relationships with those with whom we have worked.

A key lesson from BioExcel-1 has been the value of using multiple modalities to work with a given individual, organisation or group. There are certainly different modalities that work better with different groups, but probably more important is having the flexibility to combine these in ways which fit with the work that is being undertaken over a period of time. In BioExcel-2 we have sought to continue this process by placing a greater emphasis on a support team that can balance its time between engagement events and more specific support in response to user need.

[50] https://drugdesigndata.org/about/grand-challenge