H2020 EINFRA-5-2015

www.bioexcel.eu

Project Number 675728

D5.4 – Governance and Business Plan

WP5: Governance and Sustainability

Document Information

Deliverable Number	D5.4
Deliverable Name	Governance and Business Plan
Due Date	2018-12-31 (PM38)
Deliverable Lead	FT
Authors	Holger Noack (FT), Ian Harrow (IHC), Alexandre Bonvin (UU), Adam Carter (EPCC), Rossen Apostolov (KTH), Erwin Laure (KTH)
Keywords	Sustainability, users, support, services, commercial
WP	WP5
Nature	Report
Dissemination Level	Public
Final Version Date	2018-12-21
Reviewed by	PMB
MGT Board Approval	2018-12-22

Executive Summary

This governance and business plan shows how we intend to build a sustainable and resilient organisation. We provide rough estimates of our self-cost throughout the document based on available data about the diverse kind of personnel at different institutions in different countries. This business plan is not a commitment to be executed exactly as is. Rather, every time we get to a point to actually deliver a service or a product, the efforts and the cost will need to be negotiated based on the information we have available then.

Section one gives a brief introduction into the field biomolecular computational research and explains, how we as a Center of Excellence positon ourselves in this environment.

In section two, we present the main stakeholders of the ecosystem and provide the estimate that our user community contains at least 15.000 active users. We discuss findings from surveys among academic and industrial users, showing that usability aspects, compatibility of software, and quality of prediction are major concerns. We further present data from HPC market studies, based on which we estimate that the European market for biomolecular HPC software has a size of ca €170 million in 2018.

In section three we discuss the added value that our center provides to the community. We argue, that through our work, users receive better software and have access to more support, which in turn improves the quality of research that is performed.

We break down the cost for the main activities of our center in section four, and argue that these cost need to be covered through public funding.

In section five, we motivate our intention to develop commercial offerings and describe three concrete business opportunities that we are currently exploring. This section is followed by a description of our marketing approach.

Our organisational structure and its governance is described in section seven, including the description of the process for the formation of a legal entity. In section eight we look at our resources and argue that our employees are our most valuable assets. We indicate how the social enterprise could move ahead to fund product development and initiate its operations, and we conclude this document in section nine with an outlook to BioExcel 2, the next phase of the Horizon 2020 project starting in 2019.

Contents

1 Introduction

Computational biomolecular research is indispensable to address current challenges in healthcare, agriculture and nutrition. Over recent years, the scale and complexity of molecular systems investigated by computational methods has increased significantly. This has been made possible by the development of new software tools and the availability of more powerful hardware environments.

As a consequence, it has become much more difficult for the individual biomolecular researcher to identify the most appropriate computational route to address a specific research topic. Today, accurate and efficient application of computational methodology requires the ability to design and deploy workflows, to choose the right software to accomplish a certain task, to understand intricacies of the underlying theoretical foundations, the impact of parametrization, and the ability to utilize the full scale of modern hardware environments. In other words, computational biomolecular research is an interdisciplinary type of research, which to a large degree depends on contributions from experts in molecular biology, theoretical modelling, and software development.

Within this context, BioExcel aims to assume a central role that allows us to facilitate cooperation between these various experts. Our goal is to increase the quality of biomolecular computational research that is performed in academic and industrial research environments. To achieve this, BioExcel has been actively working to (i) improve popular computational tools, and (ii) educate researchers in the best use of these computational tools. Our vision and mission statements are formulated to reflect these intentions:

Vision: *"A central hub for biomolecular modelling and simulations"*

Mission: *"Enabling better science by improving the most popular biomolecular software and spreading best practices and expertise among the communities."*

However, in order for us to be able to focus our efforts on our declared mission, BioExcel as an organisation needs to ensure that we can make independent decisions, even if they at some point deviate from the agendas of the host organisations of the current BioExcel partners. Consequently, this requires a financially independent entity, capable of surviving intermittent periods of funding. We think that the key to achieve this type of sustainability lies in the ability to remain relevant to our user communities. In other words, BioExcel's existence is directly linked to the value that we provide to the user communities.

When it comes to financial stability, we need to consider two major sources of revenue. Because the main beneficiary of our work and activities is the general public, it is appropriate that BioExcel continue to request public funding, either from European or national levels. However, in the case where additional functionality, usability or support is called for by an exclusive and/or industrial user group, we intend to charge commercial fees to finance our efforts. The work undertaken in direct contact with industry users to explore offerings with commercial potential has been presented in

deliverable D3.6. Based on these findings a forward-looking summary of possible interaction modalities with users was presented in deliverable D3.5.

The work presented in this sustainability plan was started by specifying our ambitions for certain aspects or areas that play into the overall sustainability goal:

i. To nurture a diverse user community, ideally consisting of individuals from academia and industry.

ii. To facilitate collaboration and expertise sharing among the different types of stakeholders.

iii. To provide training for research users as an important part of our activities. If requested, to provide user-specific support, either through collaborative research projects or in the form professional consultancy-type arrangements.

iv. To improve the performance and ease-of-use of important open-source software. This involves work on scalability aspects, as well as development of workflows that improve usability.

v. Finally, to establish a strong brand that increases our visibility in the communities and allows us to attract competence and funding.

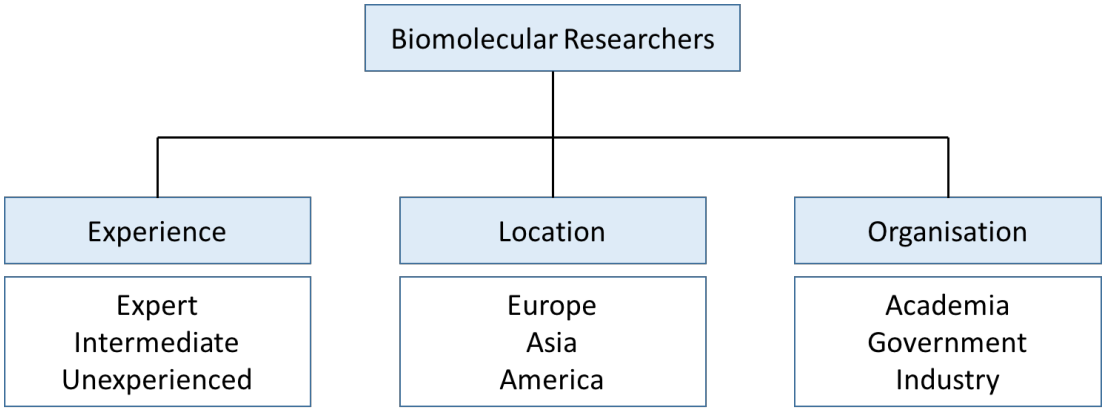
2 Ecosystem of computational research

2.1 Stakeholders and segmentation of user types

The three most prominent stakeholders in the ecosystem of computational research are:

1. Research users
2. Software providers, and
3. E-infrastructure providers

Research users: For BioExcel, the research users are in the center of attention. It is the conditions and quality of their work that we are trying to improve through provision of support and through selected development efforts of popular software. We can differentiate biomolecular researchers by their level of experience, location, and type of host organisation they are employed at.



Depending on the context a different granularity of this segmentation is required. For example, if want to design a course on calculation of free energies, we would need to define exactly what type of researcher this course is aimed at. In such case, we would need to take a closer look at what abilities researchers would need to have, and how those are linked to a certain degree of expertise in computational research. Accordingly, BioExcel has prototyped user personae that reflect different levels of user experience, and has performed a gap analysis, to determine the training needs of these different personae[1]. During 2019, these reference personae will be made available in the knowledge resource center on our website.

The type of host organisation or research environment of a researcher provides some information about specific needs. In general, academia has a broad support of qualified post-graduate students and postdocs, who usually are accustomed to open-source software. But academia also has an educational mandate, which in the case of computational research results in a recurring need to deploy virtual machines on remote computing infrastructures, to educate a number of largely inexperienced undergraduates. In comparison, researchers working at public non-academic organisations do not have educational obligations, and tend to work more towards applied research. These researchers are more likely to require stable tools that excel in performing repetitive tasks. Finally, biomolecular researchers from industry generally have higher requirements on data security to protect the intellectual property of molecules during analysis.

BioExcel's user community is composed of individuals that use the software adopted and further developed in BioExcel, with the two most prominent tools being GROMACS and HADDOCK. Surveys among these users show that the overwhelming majority of the BioExcel user community work in academic/non-profit environments. Interestingly, an average of circa 40% are located in Europe. The level of experience is more elusive, but is certainly correlated with the academic degree. For HADDOCK, we also know that 34% of the users had used the software for less than a year.

	GROMACS[2]	**HADDOCK[3]**
Experience	43% PhD students 23% Postdocs 15% Staff researcher	33% PhD students 20% Postdocs 36% Staff researcher
Location	50% Europe 19% North America 16% Asia	28% Europe 24% India 19% North America
Organisation	94% non-profit 6% for-profit	97% non-profit 3% for-profit

Table 1: Composition of BioExcel's user community based on surveys among users of HADDOCK and GROMACS.

[1] https://doi.org/10.5281/zenodo.264231

[2] GROMACS feature and usage survey, 2014

[3] HADDOCK online survey, access on April 2017: https://docs.google.com/forms/d/1WMzzvssuDMApHJW8ugMODXeo34tCI6CmXcdJsRctTO8/viewanalytics#start=publishanalytics

The size of this research community can be estimated by the number of registered software users, or from the number of publications that refer to the biomolecular software used. Based on an approach where we assume that each publication represents on average 3 active users, we estimate that our community consists of at least 15,000 active researchers.

Software providers: We can differentiate between open-source software that is developed in decentralized efforts mainly from academic research groups worldwide. In parallel, we have so called independent software vendors (ISVs) that develop and distribute specialized software for niched high-end computing applications. This specialization is a way to deal with competitive pressure. The majority of ISVs are small or medium-sized enterprises. Both open-source developers and ISVs must deal with a combination of operating systems, distributions, processors, and middleware environments. By implementing a policy of open programming interfaces and support libraries, BioExcel would ensure that software created within the open-source community allows for ISVs to provide added value, i.e. to develop complementary commercial variants and related software tools. Partnerships with BioExcel could be an interesting commercial opportunity for "co-design" of academic community and industrial applications software.

E-Infrastructure providers: E-infrastructure providers are often represented through system administrators. In the case of large HPC sites, they install and operate the hardware, monitor the performance, regulate access, and offer helpdesk or similar support services to researcher users. The quality of the software that runs on the hardware is part of the user satisfaction, and also impacts the efficiency on how the infrastructure is used. By matching-up the capabilities of BioExcel-supported software with available computing infrastructures, we are making sure that e-infrastructures are used efficiently. BioExcel is working closely with system administrators in deployment and testing of software performance updates.

2.2 Market size for biomolecular computational software

Simulations of biomolecular systems require effective computing techniques to achieve timely results. In their majority, computations are either performed as tightly coupled parallel jobs on high-performance computing infrastructures (HPC), or as loosely-coupled tasks on distributed networks of processors known as high-throughput computing (HTC).

The market size for biomolecular computational software solutions can be estimated by looking at HPC market studies that the International Data Corporation (IDC) has performed on behalf of the EC. According to the IDC market report from 2014-2015, HPC software generated global revenues of €4.4 billion in 2013, of which Europe's spending was estimated to €1.2 billion. It was further estimated that the European spending in this market will expand to €1.7 billion by 2018. Biosciences are the second largest HPC application segment, standing for circa 10% of all spending in this sector. Based on these numbers we can conclude that the European market size for biomolecular computational software is around €170 million. Some of the major

commercial players in this market are Schrödinger[4], BIOVIA 3DS[5], Chemical Computing Group[6], and OpenEye Scientific[7].

3 Value proposition

3.1 User needs

Most researchers start to regularly use computational software at some point during their postgraduate studies, e.g. when they become involved in a research project. A basic introduction to the software may or may not have been part of their graduate education. However, often it is the senior colleagues that are helping beginner users to get started with a particular choice of software. Due to the inherent complexity of computational research, every researcher has to make a serious effort to progress their ability of utilizing the available software to its full potential. There is thus a ubiquitous need for training, which currently is answered mainly through dedicated workshops and through provision of tutorials.

In the case of HADDOCK and GROMACS, users appreciate the performance of the software and the quality of the workshops. But when asked, they would like to see even more extensive tutorials and more occasions to attend workshops. The problem is that software development teams have limited resources to provide workshops. Increasing the number of workshops would require PIs to cut down on software development or research activities. It would also require to increase staff size dedicated for the provision of training, which is hampered by limited budgets. However, notably a majority of users said that they would accept to pay a small fee if a workshop is hosted in a nearby institute. This is interesting from sustainability perspective, i.e. if it is possible to attract sufficient revenues to cover a large part of person's salary, we would be able to increase the number of workshops.

Regarding the software supported by BioExcel, academic users made clear that apart from its performance, the free access of the software is of crucial importance to them. Besides further improvements in performance and efficiency, the three most common requests are:

1. Integration of analysis tools
2. Compatibility of software with more data types and other software
3. Provision of tools that support building molecular systems and setting up calculations

Researchers working in industry are a minority in our communities. To understand their needs, we conducted telephone interviews and went on site visits. The industrial researchers that we spoke with worked with virtual screening, antibody engineering, and lead optimization. They were mainly using industry-standard software such as Schrödinger (Maestro) or Chemical Computing Group (MOE) deployed on internal computing resources. Similar to their academic colleagues, scalability of software is

[4] https://www.schrodinger.com
[5] https://www.3ds.com/products-services/biovia
[6] https://www.chemcomp.com
[7] https://www.eyesopen.com

seen as important since it affects time to results. However, the quality of prediction is the major concern. In regards to usability, researchers see the need to facilitate:

1. Conversion between formats
2. Integration between software tools
3. Building and setting up systems, in particular:
    a) Embedding molecules in membranes
    b) Finding optimal parameters for force fields
    c) Adding ligands, which often affects a) and b) above

Our open-source computational software was known to most of these researchers, given that all of them have an academic background. Some even said that they had the impression that GROMACS was a very capable tool. However, in this particular case it was stated that the sheer complexity of this software deters them from using it, since they do not have time for extensive read-up of literature to be able to use it properly.

Industrial researchers perceive a higher pressure to shorten time-to-results. In addition, professional support and convenience involved are in this reluctance to use open-source software, given that there is a budget to purchase validated industry standard software that can be deployed and used more easily. In line with the priority to achieve tangible results in the shortest possible time, there is a general acceptance to purchase consultancy services when needed.

3.2 Added value of BioExcel vs. value of individual tools

BioExcel has brought together the user communities of leading open-source research software. Through BioExcel, software development of these previously loosely grouped programs is now more tightly aligned. BioExcel contributes by improving scalability, performance and usability of these tools. The work on scalability has made it possible to run this software on modern large-scale HPC and HTC environments and significantly improve the performance.

Usability and efficiency have been improved through new user interfaces, and a whole new concept of BioExcel building blocks that make it easier to design use-case specific workflows, and to move more easily between different tools. The ability to execute complex research tasks that rely on several tools in a semi-automatic fashion will benefit research users significantly. Consequently, we expect more users from both academia and industry to commit to the ecosystem of BioExcel-supported open-source computational software, under the assumption that it will be maintained and expanded in the future.

Another important part of the BioExcel added value is the increased number of training events and the provision of more formats, e.g. webinars, workshops, summer schools, community forums and hackathons.

All of our work has been paid for by public funding and is provided to the user communities for free. It is our ambition that free access to the basic tools and support offerings, especially for non-profit users, continues to be a core value of BioExcel. However, as part of our sustainability planning, we are investigating commercial

approaches that would allow us to maintain basic operations of the center during periods where we have no access to public funding.

4 Cost estimates for the center's main activities

The activities of the center can be grouped in three categories. The category software development comprises efforts to improve performance, usability and scalability of important open-source codes. Within user support, we provide online support and training events that aim to increase the competency of researchers to make best use of the computational tools available. Finally, community building refers to support-like activities that, however, are not targeted at specific user groups. Rather, these efforts provide access points to the community and help to increase our visibility.

Software development	User support	Community building
Improvement of codes	Webinars	Web Site
Workflow components	Workshops	White papers
Workflow design	Training events	Strategic partnerships
	Support forum	

Table 2: Overview over the main activities of the BioExcel CoE

The main component contributing to the cost for any of these activities is the salary of the employees. Gross salaries differ depending on a person's qualifications and work responsibilities, and they are also affected by the cost-of-living for the location, where the work is performed. Another factor contributing to the total cost of labour are payroll taxes for pensions savings, health-care insurance etc. The payroll-tax rules differ somewhat between the European countries, with the average tax rates ranging from 9% in Malta up to 49% in Sweden. However, even within the same country a tax rate is not static and may differ between institutions, e.g. offering generous pension payments may help to compete for competence.

Our total cost of labour is somewhat reflected in the person-month rates that were used to construct the Horizon 2020 budget for the BioExcel consortium. However, these rates are composed of contributions from different types of personnel working on different tasks in the project. To arrive at a cost breakdown that is more specific for certain activities, we started by looking at the cost of labour for different types of personnel. The rates specified in table 3 are taken from salary scales provided by a partner institution for each country, where BioExcel is operating[8] with the additional social and other institutional charges added. Gross salary scales do not specify fixed salaries for each employee category, but rather a range that provides some flexibility to factor in experience. For better comparison between the different countries, we have chosen numbers from the middle of a salary scale, even though some partners might have higher cost.

Total cost of labour in € / month*	UK	Netherlands	Germany	Spain	Sweden

[8] See appendix.

PhD student	3,650	4,253	5,375	2,565	4,570
Postdoc	4,648	5,271	5,825	3,915	5,637
Staff Researcher*	5,737	7,559	6,950	4,995	6,246
Full Professor	9,367	11,816	8,425	8,100	10,969

Table 3: Examples of total cost of labour for the five countries where BioExcel operates, composed of gross salary and with social fees added. Payroll taxes range from 28% to 56% between the countries in the table. The stated total cost of salary is neither intended nor suitable for calculating personal remuneration. *The category Staff Researcher is a collection of different professional designations with comparable responsibilities, such Lecturer (UK), Associated Professor (NL), and Research Associate (GER).

PhD students are not usually charged to projects. The bulk of the work is done by postdoc and staff researchers. PIs are to a higher degree involved in several projects, and hence, their contributions to BioExcel will be more selective. To reflect this different occupation for different types of activities, we will in the following use two rates. Rate 1 is based on the average cost of salary for postdocs and staff researchers from all partners, and rate 2 will in addition factor in the cost of salary for PIs.

Rate 1	5,678 €/month	35 €/h
Rate 2	7,031 €/month	44 €/h

Table 4: Average monthly cost of salary and hourly rates. Rate 1 uses example salaries from postdocs and staff researchers, rate 2 factors in the salary of professors. The hourly rate is based on 160 hours per month.

4.1 Software development

Improvement of scalability and performance of computational software is a continuous process, as is the development of building blocks, virtual machines and the design of entire workflows for specific use cases. Our cost related to these activities is thus a monthly running cost that scales with the number of people that are involved. It is a fair assumption that software development contains about equal contributions from postdocs, staff researchers and PIs. The table below gives the average total cost of labour for these three types of employees, using the example rates from table 3 for the locations where the majority of the work is done.

	HADDOCK	**GROMACS**	**PMX/MiMiC**	**Workflows / CWL**	
Location	Netherlands	Sweden	Germany	Spain	UK
PM Average	8,215	7,617	7,067	5,670	6,584

Table 5: Estimated average cost for software development (gross salary + payroll taxes) in €/person-month. Institutional overhead is not accounted for.

4.2 User support

User support comprises activities that occur with different frequency, and that require a varying extent of preparation. The individual support offerings often involve contributions from different BioExcel partners. A detailed description of how these user-support activities are delivered, including estimates of workload and related cost factors is given in D3.5 section 2 "*Core Services for the Wider Community"*.

For the webinars and the support forum we use the workload estimates presented in D3.5. The estimate for the 4-hour workshop is based on a structure consisting of 2 hours seminar followed by 2 hours hands-on exercise. The absolute minimum would be two persons involved in delivery, who each need to spent 6 hours on preparation, and additional 2 hours on post processing. A training day might be part of a larger event such as a summer school. There are usually more participants at these events, and the training material is more comprehensive, which requires a higher number of trainers. The workload estimate for the training day is based on a structure of 8 hours of seminars & hands-on exercise with 3 people involved in delivery. For each of these 24 person-hours of delivery we factor in two person-hours for preparation and post-processing.

	Workload	Rate	Other cost	Total	Comment
Webinar 1h	7h	35€/h	89€/month software	€293	per webinar
Workshop 4h	24h	44€/h	not accounting e.g. travel, catering, accommodation,	€1055	per workshop
Training 1 day	72h	44€/h	not accounting e.g. travel, catering, venue	€3164	per day
Support forum	8h	44€/h	-	€352	per month

Table 6: Cost estimates for selected, well defined user support activities. Institutional overhead is not accounted for. The stated workload estimate is the sum of contributions from several people involved in a support offering. In general, the estimates are on the low side. For some webinars, workshops or training events more personnel is required.

4.3 Community building

Community building is a task with many elements that are hard to quantify. To build a community it is necessary to understand the needs of the people who form the community, to learn how the community relates to other communities, to gain the trust of the people that make up the community and to maintain the relationship with the community through meaningful engagement. In many cases, this is an ongoing process, and activities related to community building are very interlinked with our user support and software development activities. This is illustrated, for example, in the wide variety of modalities, described in D3.5, that have been used in BioExcel to engage with the user community.

In the table below we look at activities that benefit the wider community, in contrast to the support activities mentioned above that benefit a subgroup of the community at a time. The workload estimate for the website content assumes an existing, fully functional website that receives regular weekly updates on community news and events. Publications of white papers does not account for all the effort that goes into accumulating the content for the white paper, the workload estimate is only looking at the final effort to compile the white paper. Our assumption is based on one person spending a total of 40 hours compiling the paper, supported by two more persons contributing circa 20 hours each. To build and maintain strategic partnerships we assume two persons each spending 2h per week on preparation of MoUs and networking activities.

	Workload	Rate	Other cost	**Total**	Comment
Website content	8h	35€/h	150€/year	€296	per month
White papers	80h	44€/h	-	€3,515	per white paper
Strategic Partnerships	16h	44€/h	-	€703	per month

Table 7: Cost estimates for selected community building efforts.

4.4 Other cost

Apart of the operational cost for the main activities of our center, we also have marketing cost to promote our services, as well as cost related to management and administration. We describe our marketing in section 6, with a brief breakdown of the associated cost. Management and administration includes activities such regular conference calls, internal meetings, bilateral discussions, sourcing and allocation of personnel, monitoring of KPIs, and accounting. We estimate that we spend circa 5% of our time on these activities, or 8h per month per consortium partner. With 11 partners, the total estimated cost for management and administration is estimated to circa €3900 per month using hourly rate 2.

5 Commercial applications

Industrial researchers are currently a minority of our community. The development of commercial offerings is foremost driven by the desire to broaden the impact of our work, i.e. to extend the reach of the BioExcel-supported tools to industrial research. Another driving factor is our intention to generate additional sources of revenue that have the potential to supplement public funding. In this context, industrial researchers are the most likely source of paying customers.

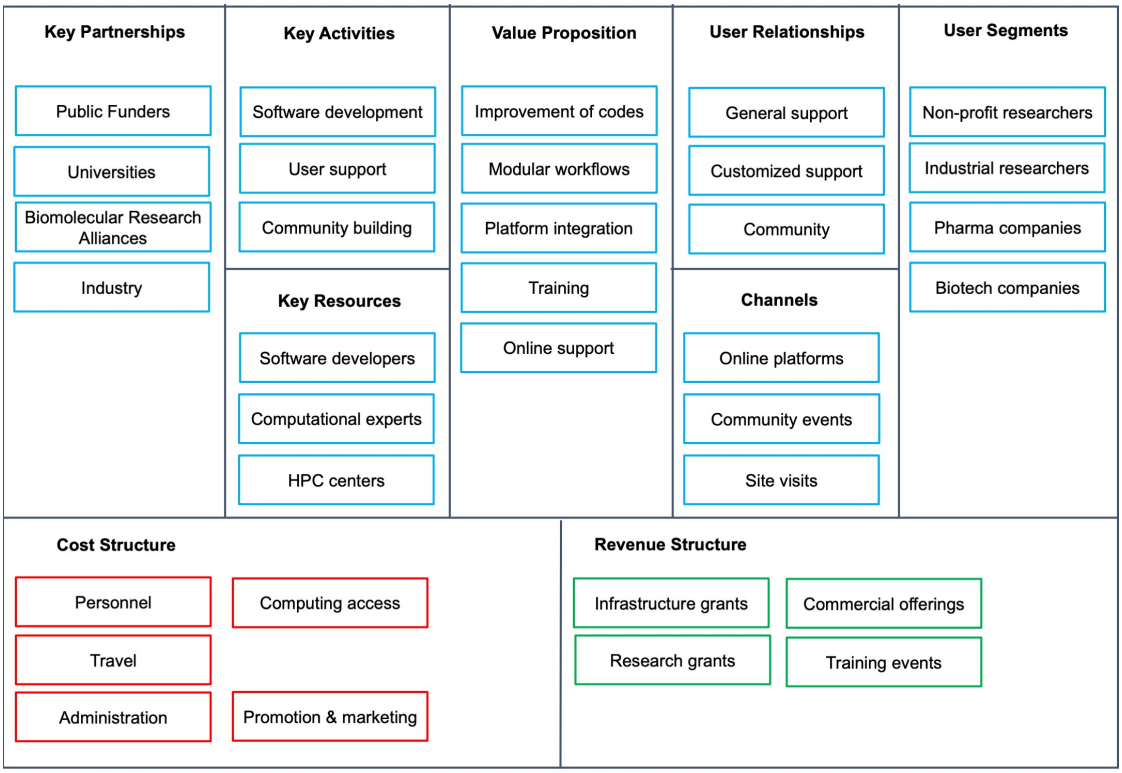
5.1 Business opportunities

From telephone interviews with industrial researchers that we conducted during 2016 and 2017, we know that their needs are quite similar to those their non-profit colleagues. The main difference, in our finding, is an emphasis on the importance of the quality of

prediction that trumped scalability or performance aspects. In 2018, we conducted site visits to two pharmaceutical companies, Janssen (part of Johnson & Johnson) and UCB in Belgium. The discussion we had deepened our understanding of the challenges and needs for molecular modelling and simulations in computer aided drug design. The reluctance to use free open-source codes is explained with the lack of enterprise level application support that providers of commercial solutions offer.

To successfully tap into the commercial space, BioExcel needs to leverage the edge it has over existing commercial actors. Our center is unique in that it unites the software development expertise of popular open-source codes that complement each other. GROMACS has an emphasis on MD calculations and is designed to exploit the computational power of the latest HPC infrastructures. HADDOCK on the other hand is used for docking applications and excels on distributed HTC computing platforms. We also have expertise in ab-initio QM methods and the calculation of free energies, which have always been hot topics in the communities. This technical expertise is combined with biomolecular research excellence through our key opinion leaders, and the possibility to access the latest high-performance computational infrastructures that are available in Europe. So, in summary, our edge is the ability to provide potential customers the full palette of technical and scientific support, ranging from biochemical reactivity studies of small molecules to studies of macromolecules and cellular mechanism. We have summarized our position in the marketplace / ecosystem in the business canvas model below.

Scheme 1: Business canvas model



The challenge for us to establish commercial offerings lies in the practicalities. Engagement with industrial actors is, apart from occasional site visits or workshops, not typically a focus in academia. Most institutions reject the idea that their employees divert from their academic agenda to pursue commercial activities. We conducted

bilateral talks with each partner to understand the capabilities, interests and limitations. We also had a joint workshop to discuss our options that resulted in a list of three potential offerings:

1. Provision of support, most likely as consultancy arrangement
2. Provision of training, including the use of online tools to reach a wider audience
3. Provision of workflow components or entire workflows

5.2 Consultancy support

The provision of consultancy and support is what industrial users were most interested in during our discussions. Consultancy would be provided by applications experts who would work on a client project at one of the BioExcel sites. Applications experts are people with strong background in software development of our codes, combined with biomolecular research experience. Consultancy arrangements could entail short-term projects to solve a specific biomolecular problem, design and execution of large-scale calculations making use of the powerful HPC infrastructures that BioExcel has access to, or simply dedicated support for clients that have decided to use the open-source codes. In other words, this intended provision of consultancy & support would be a driving factor for industrial users to convert to our codes. This transition would make it possible to offer commercial variants / modules to industrial actors as will be discussed in 5.4. below.

The challenge for BioExcel is that application experts need to be able to deal with customer questions on short notice, or take on projects that may come up any time. We did not have the capacity to provide this service during this first phase of BioExcel, since it would have required us to pull our developers from their assignments and neglect the work that we needed to do.

To be able to provide this type of service, we need to employ dedicated applications experts. It would be difficult to attract people with the right competency, if we were to offer them only short-term, insecure employment contracts in line with the expected low flow of assignments during the start-up phase. Moreover, most universities require the budget to cover the total cost of salary for a full year to be prepaid, before committing to employ a person. In other words, to be able to deliver this type of consultancy, we need to advance the funding for competitive salaries, comparable to that of software developers, even if we initially are unlikely to have full occupancy to justify the investment. To get started, for the second phase of BioExcel, the Horizon 2020 project starting in 2019, we have allocated part of the budget to employ application experts. The recruitment process has already been initiated, and we hope to get started with this service offering during Q2 2019.

To speed up the process towards a self-sustainable service, application experts could support the sustainability team to scout for suitable projects, preferably at small and medium size biotech companies. We might also consider to offer some selected pilot projects at a reduced price to build a pool of reference projects to attract future customers. The scouting should focus on clients that are not yet using computational methods, but that might benefit from the capabilities. Application experts are also expected to engage in other community and support activities such as participating in events and facilitating training activities. The table below summarizes the cost for this

service, using the same salary as was used for the estimate of our cost for software development from table 5. Again, we only look at the cost of the service, not accounting for sales, management etc.

	Application expert specialized on				
	HADDOCK	**GROMACS**	**PMX/MiMiC**	**Workflows / CWL**	
Location	Netherlands	Sweden	Germany	Spain	UK
PM Average	8,215	7,617	7,067	5,670	6,584
Self-cost in € / h	51	48	44	35	41
Overhead flat	25%				
Total € / h	64	60	55	44	51
Commercial € / h	130-190	120-180	110-170	90-130	100-150

Table 8: The hourly rates for application experts. We have factored in a flat rate for the overhead to arrive at the suggested commercial rates. This flat rate of 25% is based on the assumption that this service at some point would be delivered through a company. If instead this service would be delivered through a university partner, we would need to calculate with a flat rate of 60%.

The commercial rate is an estimate and will need to be negotiated with the client on a case by case basis. When working with industrial customers the margin should be at least 50% of the self-cost including overhead, and the resulting hourly rate should be within the range of the local market prices where the service is provided. As a guideline, in Western Europe it is not uncommon to charge 100-300 €/h for highly specialised consultancy services. These seemingly high prices account for the fact that the sales processes are very customer specific and time consuming as compared to more general types of consultancy.

In conclusion, our intention should not be to compete on price, but rather on quality/value of the service provided. Our differentiating factor with this service is the niche competency in high performance computing that we provide, and the cluster of expertise and key opinion leaders that we can draw from. To capitalize on this edge, we need to use the best experts available, and those are found in the same locations as our software development teams. We can still offer competitive pricing if we negotiate package deals for a larger number of hours to be consumed within a given timeframe. Similar, we could offer a monthly subscription to application support, which likely will become relevant mostly for larger Pharma companies with in-house computational research teams that have made the switch to BioExcel-supported software.

5.3 Training as service

With the training offering we consider foremost workshops that teach users how to use the software correctly. Workshops are composed of a seminar part delivered by the PI, followed by hands-on exercise that are supervised by staff researchers or postdocs. Different workshops can be designed to target different user types with different level of experience. In some sense, workshops could be viewed as a type of consultancy. The main difference to the consultancy assignments described previously, is that the content of the workshop sessions can be reused and will be valid for several users. Furthermore, we expect the customers for our training offering to predominantly originate from academia. Accordingly, we have to reach at a cost and pricing structure that is feasible for this target audience.

From surveys among our non-profit users we know that our workshops are highly regarded. The problem is that they benefit only a small group of participants at a time, most of them located in Europe, whereas at least 50% of our community are located outside of Europe. The bottleneck, however, is the availability of the key personnel, such as PI or lead developers.

The idea is to develop an online workshop-type experience that extends our reach and multiplies the number of users that benefit from the available time of our key personnel. Through our newly developed virtual machines, course participants could work on their hands-on exercises from their own laptop. The technical barrier to install these virtual machines is modest, some participants will be able to do it by themselves, others can be guided through the process with a tutorial. We consider to record the seminar sessions and send them together with the virtual machines to registered participants. Users can go through tutorials on their own, but will also have the possibility to join live sessions. A more detailed description of our initial plans is given in D4.6. Based on response from the community, we have narrowed down a list of three courses that are candidates for a first pilot:

1. Hands-on introduction to HPC for life-scientist
2. A HADDOCK course on docking
3. A Gromacs course on coarse-grained modelling

Our intention is to start with GROMACS, since this course would be most suited to test our virtual machines with users. To start with, we need to decide on the final structure of the course, the content of the different modules, the length of the individual sessions, and the type of exercises to include. Naturally, a lot of material is available, but it needs to be repackaged.

Depending on the reception, the training offerings can be further developed. Seminars could be streamed with live audiences that are allowed to ask questions, similar to webinars. We might also consider using an online portal, where users log in to view recorded sessions and download virtual machines, instead of sending them out. All of these options will be evaluated during the prototyping that is planned to commence in 2019. However, the overall goal is to reach at a structure that maximises the reach and the benefit to users, while being economical with time of our staff. The cost estimate below is based on a course design consisting of two recorded seminars and two live FAQ sessions, each with a length of 1 hour. The cost of salary is based on rate 2 from table 4.

Estimated cost to develop an online training offering		
Cost of Salary	€7,031	average per person month
PMs needed	2	Course design (content, modules, exercise)
	1	Course creation (recordings, preparation of VMs)
Total	**€21,092**	
Monthly running cost: 2 live sessions, 1h each		
Cost of Salary	€44	hourly rate
Preparation	4 h	1 person, 2h preparation for each live session
Broadcasting	4 h	1 person streaming, and 1 backup person
Post-processing	2 h	1 person, 1h post-processing for each live session
Email support	8h	1 person supporting 20 participants, 2h per week
Total	**€791**	

Table 9: Cost estimate for one online course.

Assuming that the course material is valid for a duration of 2 years, the development cost is approximately €10,546 per year. If we provide the same course once every month, the monthly cost for this course would be €1,670. Assuming furthermore that we have on average 20 participants per course, the charge per participant would need to be €83 to break even. This level of participation and the pricing might seem feasible; however, it is not clear how many of our users from e.g. India or China could afford to pay this rate, given that a substantial fraction of our user group is located in Asia.

Also, not included in this cost estimate are cost related to the whole virtual learning environment, such as server hosting, technical support, live chat assistance or payment gateways. Initially, until we have a more detailed structure in place, it might be safe to factor in an overhead flat rate of 40%, assuming that the online training is provided through a company with otherwise average overhead of 25%.

Nevertheless, the commercial feasibility will increase over time if we succeed to build up a set of online courses on different topics. In this case, we could use pricing models, where users get access to all course material for a monthly rate. Even if we should not

break even entirely, and still need some co-funding from public sources to maintain this service, the result would still be that we drastically extend our reach to the larger community independent of location, thereby maximising the impact of the investments from public sources that have been made to improve the BioExcel-supported software.

5.4 Workflow-related products

The motivation to develop a software product is that once finalised, it could provide a steady revenue without requiring substantial workload on our side. The decision to choose workflows as vehicle is that they allow us to capitalize on the functionalities of GROMACS and HADDOCK, yet without interfering with the freely available nature of those codes. In principal, any actor in the marketplace could develop commercial software that is linked to these codes. However, our advantage is that the developers of these codes are part of BioExcel, and hence we can expect a smooth prototyping resulting in a better performance of linked software. Workflows are furthermore interesting since they simplify the user experience, and hence enable non-expert users to make best use of the software.

We are currently investigating two types of workflow-related products, a KNIME integration and a GUI web server. KNIME is a popular platform for development of generic computational workflows but very extensively applied to drug design. Integration of our workflow building blocks into KNIME allows us to tap into an existing user community of more than 20.000 active users, of which approximately 40% work with pharmaceutical research. KNIME is a java-based environment, but it is possible to integrate nodes that are wrapped in Python for prototyping. Based on the response that we receive from potential customers, we will select certain workflow modules for which we are going to write the Java layers. Or alternatively, we could choose these modules based on a requested complete workflow. Creating Java layers is workload intensive, but it will allow users to connect nodes and modify workflows, simply by dragging and dropping.

Our second workflow-related product idea is a web server hosting a fixed workflow that will allow users to run MD calculations to refine energies of molecular mutations. The server should offer an intuitive GUI that allows users to modify molecules and execute the installed workflow in a guided process. Furthermore, through this web server, non-expert users will be able generate the needed input files to make use of large-scale HPC resources. We also explore the possibility of directly connecting this GUI with supercomputers. If successful, this tool will increase the leverage of European investments into computing infrastructure by increasing the number of users that can make useful application thereof.

To test these product ideas with prospective users, we recently conducted a focus group event in London with representatives from AstraZeneca, MedImmune and UCB. We have received overall positive feedback, i.e. participants liked the idea of flexibility that the integration with KNIME nodes provides, and also the GUI web server that will be useful for non-expert users. In conclusion, the focus group meeting confirmed that there could be real interest. However, more such meetings will be needed during later stages of prototyping to ensure that the final products meet the needs of their users. One take-home message was again the emphasis on the importance of enterprise-level support for any type of product.

The development costs are difficult to estimate, since at this point the prototyping structure is not entirely clear. The effort needed for the development of Java nodes will depend on the type of each node. We base our estimate on the effort that would go into building an entire workflow containing a number of Java nodes. For our second product idea, the GUI web server, we assume a two-stage approach consisting of development of a first working version with basic functionality, plus a second phase where we increase functionality and features. As most of the development work related to workflows was done in Spain and the UK, and we base our cost estimate on the average cost of salary for developers from both locations, specified in table 5, i.e. €5,670 and €6,584 respectively.

	Workflow with KNIME Java nodes	GUI web server hosting fixed workflow(s)
Average cost of salary	€6,127	
PMs needed	6-9	9-15
Total	ca. €37,000 to €55,000	ca. €55,000 to €92,000

Table 10: A first estimate of the development cost for KNIME Java nodes and the GUI web server.

In addition to the development activities, we will need conduct several more industry focus groups in order to gather requirements and feedback as we go along. Based on our experience with the first focus group meeting in London this year, we should count with at least €2,000 per meeting just for flights, location and catering, and another €1,000 for reimbursement of costs of the local participants. Assuming that we would need 3 to 4 more focus groups just to be able to build one decent product, we count with a minimum of €10,000 in addition to the estimated development cost in the table.

Given that these costs are rather high, we might need to explore various financing models to move forward with the product development. We could apply for public funding on the national level that is targeted at innovation and development activities. Another option would be to attract private capital, in which case, however, dividends from future earnings would need to be paid to the investors. Finally, we could try to have prospective industrial customers co-finance the development of our products, in which case we likely would need to offer some kind of exclusivity in return.

6 Marketing approach

Details of our ongoing marketing efforts are given in deliverable D4.6 "*Final report on dissemination and training*". There we present and discuss website usage, mailing lists, newsletters, the number of followers on LinkedIn and Twitter, the number of conferences that we attended and sponsored etc. Here, in this section, we provide a brief overview over the different marketing channels that we use.

6.1 Branding

A strong brand forms the basis of all marketing efforts, and the BioExcel brand is now recognised in biomolecular research circles. To increase recognition the BioExcel logo has been used in all media. A tag line “Supported by BioExcel” was used for sponsored events, and we consider using "Powered by BioExcel" for our commercial offerings. We try to use a unified presentation for all material that is provided online, regardless of which BioExcel-supported code it relates to. This will make it easier for biomolecular researchers to recognize tools that are supported through BioExcel.

6.2 Online media

Online marketing has the advantage that people will be able to click on an add and immediately view information about a product. It is the simplest way to reach an audience that is already quite technology savvy, and provides a level of targeting that is difficult to achieve by offline media. We have already developed a following on social media like Twitter and LinkedIn during this first phase of BioExcel, where our posts regularly reach hundreds of people. BioExcel will therefore continue to maintain these channels by ensuring that only relevant content is posted. Other ways of reaching our community is through our website and through our regular free webinars. Both can be used to occasionally promote our commercial offerings described in section 5.

6.3 Partnerships

Another channel that we have been using is word-of-mouth marketing through our network and personal contacts. The organisations that signed up to our community interest groups and our strategic partners will help us to multiply our marketing efforts. At the same time, we will continue looking for other industry contacts and organisations in order to re-enforce our network and also reach potential new markets.

Participation in biomolecular events is an important part of our marketing campaign. These include workshops, symposia, conferences and similar activities where industry and academia meet in order to discuss latest research trends and the future of business. In general, whenever possible without travelling, members of the centre should go to their closest business events to represent BioExcel and our products.

6.4 Direct sales

A pipeline of leads will be developed from the sustainability team in collaboration with all the partners and through the use of a customer relationship management (CRM) tool. We have already tested one CRM tool during this first phase of BioExcel. It will be most relevant to keep a registry for our industrial contact, and to manage consultancy assignments. Also, when piloting our workflow-related offerings, we will use the leads to check our assumptions with regards to pricing. The aim is not to have hundreds of leads, but to have a few good ones that we can progress with along the path to commercialisation.

6.5 Cost of marketing

The costs for marketing are very flexible and can vary in line with available community events, progress in product development, or strategic decisions of the board. During some periods, we might increase online marketing that requires us to build graphic material to be used on Twitter and LinkedIn. During other periods, we might need to increase our networking activities, or design brochures and roll-ups to attend community events. On average, we should count with an average of 10 hours per week

to promote our center to the user community. Based on rate 1 from table 4 our cost for marketing is estimated to circa €1419 per month. If we factor in the need to occasional purchase graphical design services, we can count with a total of circa €2,000 per month.

7 Organisational structure & Governance

The BioExcel consortium is composed largely from academic partners that work in the interest of the general public, and whose operations are funded with public money. This situation makes it difficult to engage in activities that have a commercial aspect to them. Some institutions do not allow their employees to conduct any additional activities outside of their employment. Others may have set a limit of e.g. 5% of an employee's time that this person can use to pursue non-academic activities.

Furthermore, if we are successful with the development of our commercial offerings, and if we take on consultancy assignments from industrial users, we need to be able to invoice our clients. Again, most academic partners do not have a structure in place to conduct such activities. And those that do, take out a fee of circa 25%, which we would need to factor into our pricing.

The solution is thus to start up a separate, independent entity that is capable of conducting commercial activities without the restrictions that academic institutions impose. Since our intention is to reinvest any eventual earnings to support BioExcel's mission and to co-finance the center's activities, we have investigated foundations, and economic interest groups, that both offer some advantages regarding taxation. However, the least complicated and least costly choice is to form a company. It is our intention that this company operates like a social enterprise, meaning that eventual bottom line earnings are not distributed among its shareholders, but instead reinvested into the company to support BioExcel.

7.1 Independent legal entity - process

All personnel involved in the initial phase of BioExcel, as well as those joining BioExcel during the second phase of the Horizon 2020 project are invited to become a founding member or shareholder, if they are willing contribute to the company. Those that are initially reserved or hesitant will have the option to join the company at a later stage. We have send out an invitation to all partners during December this year, and we hope to have a first online meeting with all potential co-founders during January or early February.

The initial founding members will write the company' articles of association. They will also vote on the composition of the first management board that will steer the strategy of this social enterprise.

In general, the company should be founded in a country, where we have BioExcel people in place. One option is to start this company in Sweden, since the share capital for a limited company is comparably low there. In addition, Sweden offers several opportunities to apply for innovation funding from national agencies such as e.g. VINNOVA or ALMI. The start-up process and all business dealing can be handled entirely online in Sweden, with the exception of the banking business that requires at least one personal meeting with representatives. The process is the following:

1. An online form has to be filled at the national agency *Bolagsverket* (through verksamt.se), where the business is briefly described and the founding members and the company's signatory people are specified. We would likely choose a setting where two signatory persons are needed to sign official documents in the name of the company.
2. The signatory person(s) register a bank account for the company where the start-up share capital of SEK 50,000 (circa €5,000) is deposited. We have spoken to several of the main banks, and Nordea Bank has shown to have the most competent support for entrepreneurs while offering normal market prices for their services (circa €110 per year).
3. All founding members pay their part of the share capital into the account, and a share register is created in the form of a simple excel file. Within 5 days after deposit of the share capital the company will officially be registered by Bolagsverket and is operational.

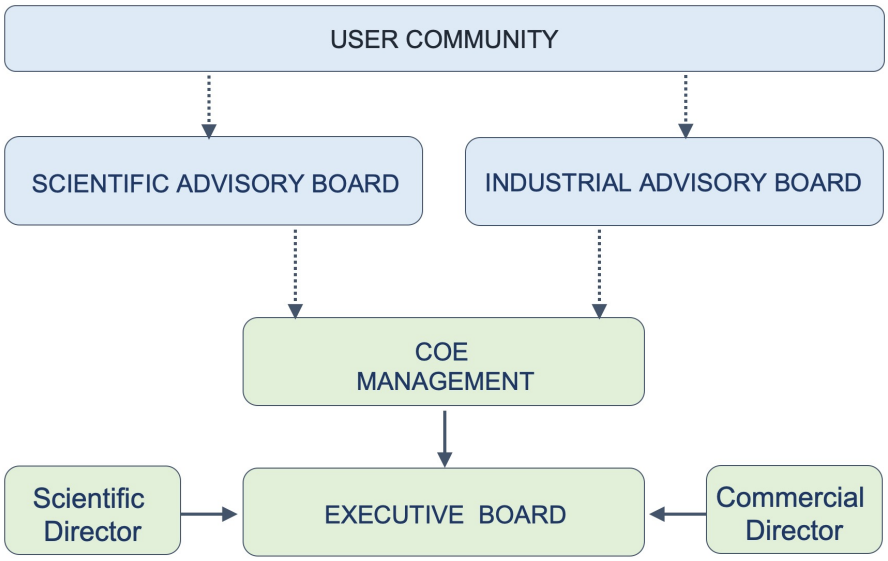
As a side note, the share capital is not intended to be consumed for the company's operations, instead, it serves as a security. Should it happen that more than 50% of the share capital are consumed, the company needs to be liquidated. In other words, all cost that the company will face during the start-up phase need to be covered through additional investments, or ideally, through paying customers. However, until the company employs personnel or generates recurring revenues, the cost can be kept close to zero.

Some shareholders might prefer to sign an insurance that provides additional security from personal liability. An insurance to cover damage of up to €5 million will cost around €200 per year.

7.2 Governance

With the foundation of the company, the brand BioExcel will consist of two strands. The consortium will deal with all non-profit activities that are paid for by public funding as outlined in section 4. Any service and product prototypes that have commercial potential, such as the ones outlined in section 5, can be handed over to the company once the product concept is mature, and assuming that the consortium's project management board (PMB) has no objections. Moreover, the company can decide to initiate its own product development projects if it can secure the necessary funding.

The BioExcel consortium is operational as long as public funding is provided, and its proposed governance structure is presented in the scheme below.



Scheme 2: Governance of the BioExcel consortium

The CoE management board is the ultimate authority of the consortium. Each partner of the consortium provides one member to this board. The work of this management board is overseen by a scientific and an industrial advisory board that are each composed of representative members from the user community. This structure will ensure that interests and the needs of the community are taken into consideration in strategic decisions of the center. The executive management of the center's activities will be coordinated by the executive board that is composed of work package leaders, and supported by a scientific director and commercial director, the later will monitor the results of the project and assess their innovation and exploitation potential.

The strategy of the BioExcel company will be aligned to that of the consortium, given that the shareholders of the company are composed from the same people that are, or have been, part of the BioExcel consortium. The company will be governed by a board of directors that is elected by its shareholders. The daily operations of the company will be overseen by its management, which initially can be composed of 1 to 2 people. During times when no public funding is provided and the consortium ceases to exist, the company will continue to uphold the BioExcel brand and its network.

8 Resource management

8.1 Assets

Our current most valuable assets are the people, the software developers and the scientific experts, who are part of the various BioExcel activities described above in section 4. To ensure a continued high quality of our work, we need to be able to keep these people as part of our center, and to attract more personnel with similar competencies. This can be done by paying competitive salaries, i.e. making use of the upper range of the salary scales, but also by providing security in the form of employments that stretch over 3 years, or possibly longer, combined with generous vacation and pension payments.

The open-source software that we support is another type of asset, namely intellectual property (IP) which is protected from unauthorized commercial exploitation through proper licenses. In general, the current licenses for GROMACS and HADDOCK permit non-profit usage, whereas commercial distribution is prohibited. It is, however, possible to create commercial software that is linked to these open-source codes, which we intend to do with the commercial workflow applications described in section 5. The successful development of the web server GUI and the Java nodes will constitute new IP that we need to protect through proper licenses and copyright. And finally, we have our goodwill asset, the BioExcel brand, that has gained reputation among the community.

8.2 Funding

During the cost breakdown presented in section 4 and section 5, we have argued that public funding will remain the major source to fund the operations of the BioExcel consortium. Even the company can, during its start-up phase, tap into available public funding aimed at innovation activities. However, on the national level this type of funding is usually paid as co-funding, meaning that shareholders are often required to contribute an equal part to the funding requested. For the company, provision of consultancy services is therefore the most immediate way to generate income that can be reinvested into product development. The problem is, however, that in order to provide consultancy services, the company would need to hire application experts, and, as we have discussed above, this requires some budget to provide competitive salaries on secure terms of employment. The provision of online training through the company could be a way to generate initial revenues to employ personnel to the company, and might be the right way to get started.

Another option forward would be the attraction of private capital from business investors to finance the development of our commercial offerings. However, the process is not simple and investors are usually looking for a relatively quick and substantial return. A product prototype needs to be in place to demonstrate the concept, and a pitch-book describing commercial risks, benefits and exit options needs to be compiled and defended against critical investors that try to spot flaws. The effort going into networking and meetings with potential investors will take a toll on our time, and one could argue that this time is better spent in trying to secure a sufficiently large pipeline of consultancy assignments instead. Moreover, given that our limited company will be set up as social enterprise, we would likely need to create an additional company specifically for the product that the business angels would finance. In this case, this

additional company would be owned by the BioExcel social enterprise and by the private investors, and in case of commercial success, the investors and the BioExcel company would both receive dividends.

9 Conclusion

With this governance and business plan we have tried to give a comprehensive picture of our plans, and the direction in that we are heading. It is as much a document for external readers to understand our center, as it is for the BioExcel partners to define and internally communicate the common denominator for our sustainability planning.

The cost estimates that have been presented throughout the document are in their majority pure operational cost for specific activities, specified as monthly running cost without factoring in overhead. This was done to allow easy adjustment of these estimates depending on type of organisation where the work will be performed. For instance, if the software development mentioned in section 4 is performed at a university, we can safely add 50% overhead without overshooting the estimate. If a developer receiving exactly the same benefits and social securities would be working under the roof of a company instead, we would factor in overhead of 25%.

The description of the planned commercial offerings is work in progress. We will continue working on it, and the timeline from hereon will align to the milestones in BioExcel 2, the second phase of this project starting in 2019. We are furthermore happy to announce that we have found a highly competent partner with a long experience in business development of larger pan-European organisations, who will lead the sustainability efforts during this next phase.

There are still many details that we need to solve, regarding e.g. the process of transferring product prototypes from the consortium to the company, or how to deal with part-time contributions from some consortium partners to the company until we can employ dedicated people inside the company. However, we have an optimistic outlook and will do our best to find reasonable solutions that benefit everyone involved as well as our user community.

10 Appendix

10.1 Cost of salaries

United Kingdom University of Edinburgh	salary group	gross salary[9] £ / year	total cost of salary** £ / year	gross salary € / month[10]	total cost of salary € / month
PhD*	UE06	30,395	38,714	2,866	3,650
PostDoc	UE07	38,460	49,295	3,626	4,648
Lecturer	UE08	47,263	60,845	4,457	5,737
Senior Lecturer	UE09	56,403	72,836	5,318	6,868
Professor	UE10	76,609	99,345	7,224	9,367
Average PM rate used in BioExcel, 2019-2022					6,800

*In the UK, PhD students typically receive studentships. If they are charged to projects their salary would correspond to UE06. ** The payroll tax is based on actual numbers[11]. For reference, these numbers correspond to a rate of 29%

Netherlands University of Utrecht	salary group[12]	gross salary € / month	total cost of salary* € / month
PhD	P	2,835	4,253
PostDoc	10.6	3,514	5,271
Staff (assoc. prof.)	12.5	5,039	7,559
Staff (lecturer)	14.5	6,056	9,084
Full Professor	H1.8	7,877	11,816
Average PM rate in BioExcel, 2019-2022			6,000

* The payroll tax at Dutch universities varies between 30-50%. Our BioExcel partner UU is applying a rate 50%,

[9] UEDIN gross salary scale: http://www.docs.csg.ed.ac.uk/HumanResources/Pay/UE01_to_UE10_Nov18.htm

[10] Using the average rate during 2018 of 1.1315 GBP:EUR https://www.ecb.europa.eu/stats/policy_and_exchange_rates/euro_reference_exchange_rates/html/eurofxref-graph-gbp.en.html

[11] UEDIN payroll taxes: http://www.docs.csg.ed.ac.uk/HumanResources/Pay/OncostsAugust2018.pdf

[12] Dutch Universities gross salary scale from 2019: http://www.vsnu.nl/files/documenten/CAO/Salarisschalen%20per%201%20feb%202019.pdf

Germany Max-Planck Society	salary group[13]	total cost of salary* € / year	total cost of salary € / month
PhD	E13 level 2 E14 level 1	64,500	5,375
PostDoc	E13 level 3 E14 level 2	69,900	5,825
Staff (assoc. prof.)	E14 level 4 E15 level 4	83,400	6,950
Full Professor	W	101,100	8,425
Average PM rate in BioExcel, 2019-2022			6,500

*The payroll tax is included in these numbers. For reference, the average payroll tax in Germany is 28%.

Spain Institute for Research in Biomedicine	gross salary € / month	total cost of salary* € / month
PhD	1,900	2,565
PostDoc	2,900	3,915
Research Associate	3,700	4,995
Full Professor	6,000	8,100
Average PM rate used in BioExcel, 2019-2022		4,200

* The average payroll tax for Spain of 35% was used[14].

[13] German DFG gross salary scale: http://www.dfg.de/formulare/60_12_-2018-/60_12_en.pdf

[14] Taken from: Statistisches Bundesamt (Destatis), 2018 https://www.destatis.de/Europa/DE/Thema/BevoelkerungSoziales/Arbeitsmarkt/HoeheLohnnebenkosten.html

Sweden Kungliga Tekniska Högskolan	gross salary SEK / month	gross salary € / month[15]	total cost of salary € / month*
PhD	30,000	2,930	4,570
PostDoc	37,000	3,613	5,637
Staff Researcher	41,000	4,004	6,246
Professor	72,000	7,031	10,969
Average PM rate used in BioExcel, 2019-2022			7,300

* KTH applies a payroll tax of 56%.

[15] Using the average rate during 2018 of 10.2400 SEK:EUR
https://www.ecb.europa.eu/stats/policy_and_exchange_rates/euro_reference_exchange_rates/html/eurofxref-graph-sek.en.html