H2020 EINFRA-5-2015

www.bioexcel.eu

Project Number 675728

D6.2 – Quality Assurance and Risk Management Plan
Revision 1

WP6: Management

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	D6.2 Revision 1
Deliverable Name	Quality Assurance and Risk Management Plan
Due Date	2016-01-31 (PM3)
Deliverable Lead	KTH
Authors	Rossen Apostolov (KTH)
Keywords	Project Management
WP	WP6
Nature	Report
Dissemination Level	Public
Final Version Date	2016-01-27
Reviewed by	Holger Noack (FT), Alexandre Bonvin (UU), Erwin Laure, Mark Abraham (KTH), Adam Carter (EPCC)
MGT Board Approval	2016-01-27
MGT Board Approval of Revision	2016-09-30

Document History

Partner	Date	Comments	Version
KTH	2016-01-07	First draft	0.1
KTH	2016-01-15	Comments from FT and UU addressed	0.2
KTH	2016-01-22	Additional comments from KTH, EPCC and UU	0.3
KTH	2016-09-22	Initial revision based on reviewers' comments	1.1
KTH	2016-09-26	Additional updates	1.2
KTH	2016-09-30	Comments by EBI, EPCC, KTH	1.3

Executive Summary

This document builds on D6.1 – Management plan (where we presented the operational structure of the project) and discusses the overall quality monitoring process for the project. Particular attention is given to the review process of deliverables and annual reports as well as milestone completion. Success criteria and KPIs are described and their status will be followed-up in the management reports. An additional section covers the overall contingency strategy where potential risks are identified for monitoring and regular assessment in the yearly management reports.

Contents

1   Introduction

The objective of the quality assurance and risk management plan is to define the processes, plans and metrics that shall apply throughout the BioExcel project in order to monitor the activities, to identify and eliminate potential risks, and to ensure the successful execution of the project. Deliverable review procedures and reporting timelines are also defined to guarantee the quality aspired for. The guidelines apply to all work packages.

2   Quality Assurance Action Plan

The quality assurance action plan addresses establishment of 1) an operational structure, 2) processes for monitoring of activities and their results, 3) mitigation strategies of potential risks.

In D6.1 we have described in details the project organization and management structure (Figure 1). Here we first give a brief overview followed by elaboration on the operational processes and assessment of results along main lines of center activities.

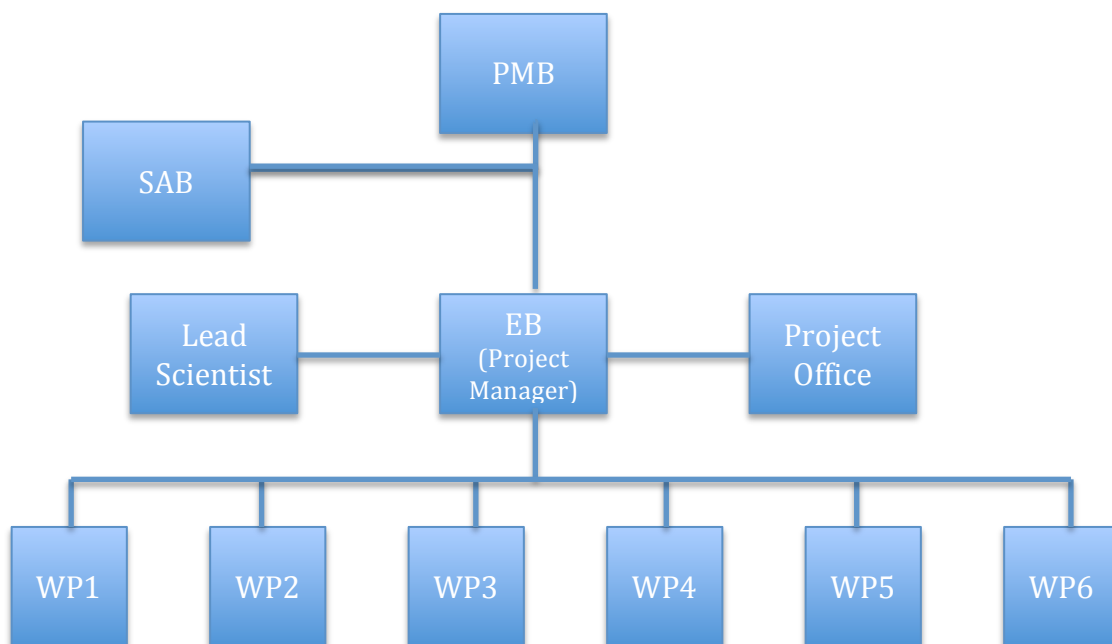**Figure 1. Center Management Structure**

2.1   Operational structure and general processes

The main driving body of the center activities is the Executive Board (EB) team. It is composed of all WP leaders, their deputies, the project manager and the lead scientist. In order to ensure appropriate involvement and communication channels, all partners are represented in the EB either as WP leaders or deputies.

EB telecom meetings are organized twice per month. During the meetings, each WP leader (or deputy in his absence) presents the status of ongoing activities and the plans for future ones. The project manager and EB members ensure that those activities are towards objectives described in the DoA, part A. Each of the WP leaders is responsible for monitoring and meeting the metrics (Table 2) , the milestones, preparing the deliverables and managing potential risks (Table 3)

Partners submit project internal quarterly reports to the project manager. The reports include:
1) description of accomplishments along the tasks assigned to the partner,
2) issues that are being observed
3) recommendation for improvement of the operations and
4) consumed effort during the reporting period.

Issues, which cannot be resolved at the operational EB level such as insufficient staff, consistent underperformance of personnel, global dangers for success of the project etc. are brought up to the corresponding PI, member of the Project Management Board (PMB), or the whole PMB, where applicable.

The project manager and the project office analyze the consumed effort and in the case of discrepancy the issue is discussed directly with the corresponding partner.

In the following sections we elaborate on some of the main procedures that have been established to ensure success of the project.

2.2   Deliverable review process
Every deliverable has a designated person responsible for its preparation, and that is typically the leader of the corresponding WP. The rest of the partners with effort in the given WP provide input to the leader for preparation of the document. The draft document is reviewed by at least two different partners who have little or no effort in the given WP while having related competency (Table 1).

Deliverable Numbers	Lead beneficiary	Main contributors	Reviewing partners
D1.1, D1.2, D1.3, D1.4	KTH	UU, Juelich, MPG	IRB, UMAN
D2.1, D2.2, D2.3, D2.4	IRB	BSC, EBI	KTH, UU
D3.1, D3.2, D3.3, D3.4, D3.5, D3.6	UEDIN	UMAN, KTH	BSC, EBI
D4.1, D4.4, D4.6	KTH	EBI	MPG, Juelich
D4.2, D4.3, D4.5	EBI	KTH	MPG, Juelich
D5.1, D5.2, D5.3	FT	KTH	IHC, UEDIN
D6.1, D6.2	KTH		FT

Table 1. Leaders and reviewers of deliverables.

The review process is as follows:

1. Deliverable leaders collect material from partners and provide the first draft for review **4 weeks** before submission deadline
2. Iterative **reviewing** process continues for the **following two weeks**
3. Second draft is submitted to the Project Management Board (PMB) **two weeks** before submission deadline
4. Comments are addressed and final version is ready **1 week** before submission deadline. The last week is used for final formatting changes and document is submitted to EC.

Important issues with deliverable preparation and submission are discussed within EB or raised to PMB, if needed.

2.3   Periodic reports and Milestones

In addition to the main deliverables, the project produces the following outcomes:

- Quarterly reports - internal
- Annual reports – to be submitted to EC
- Financial reports – internal, twice per year
- Financial annual reports – to be submitted to EC
- Milestones

Quarterly reports are submitted by each WP leader and serve as a base for deliverables as well as annual reports.

Annual reports are prepared according to the same process for the deliverables described above.

Financial reports are prepared by the Project Office. The bi-annual reports are used for internal monitoring of effort expenditure, while the annual ones are submitted to EC.

Milestones have clearly identified due dates in the DoA and responsible partners. When a milestone has been reached, the partner responsible needs to inform the EB and provide sufficient documentation which proves that the milestone has been reached. This documentation will subsequently be included in the next periodic report.

2.4   Software development

The consortium includes core developers of the three main applications as well as experts of the supported workflow systems. Quality assurance process for the developed code and adoption of best practices are explained in detail in deliverable D1.1.

2.5 User engagement

User engagement is an important but challenging task for the center. In order to ensure success and provide support to the Interest Groups (IG) leaders we have 1) defined a list of recommended actions to follow as means for interactions and stimulating activity with the communities, and 2) provided a rich set of tools for communication, engagement and support of the group members incl. forums, chat channels, blog space on our website etc. The action lists and tools selection will be updated and expanded if necessary.

The WP3 leader organizes regular, typically monthly, telecons with the IG leaders during which the latter present the current activities in the group, levels of engagement, future planned activities etc. The "health" of the group is assessed and, if necessary, alternative engagement actions are taken.

2.6 Training

Quality assurance in the training activities requires success in:
- Understanding the needs of the community
- Development of relevant training material to address those needs
- Establishment of wide dissemination channels with the wide community
- Delivery of a quality training program

To understand the needs, we have already connected with and surveyed the large user base of the three main codes and workflow systems. The currently available training material is assessed (https://goo.gl/AANSfm) and plans have been made for new development. The material will be described in D4.3 and included in the upcoming knowledge resource center.

We have compiled a database of contacts and dissemination channels to be used for advertisement of our training (as well as promotion) activities (internal). The database includes a range of organizations that represent the broad user communities (described in D3.1).

Assessment of the success of the training events is done via a template for standardized surveys (https://www.surveymonkey.co.uk/r/BioExcelTemplate) with slight modifications for each event.

WP4, similarly to other WPs, organizes regular telecons for organization and monitoring of the progress in the training and dissemination (next section) activities.

2.7 Dissemination and outreach

Adequate dissemination of the results and outreach to the communities is of vital importance. The process for ensuring success includes:
- Compiling a database of dissemination channels to other organizations
    - Each channel has a designated contact point both from BioExcel's side and from the partner organization side.
- Social media strategy

- o Establishment of selected channels (Twitter, LinkedIn, G+). Putting in place and automated system for submission of newly added website content to those channels.
  - o Dedicated person is responsible for monitoring of the website content and the status of the social media channel engagement and development
- Web content
  - o Web content creation is divided in several topical categories, in general corresponding to the activities in the different WPs. WP leaders and deputies are responsible for planning of what content should be created in the given category, assigning people to create it and monitoring the execution.
  - o Web content addition is overseen by the web curator who is also responsible for monitoring of the social media channel activities.
- Promotion material
  - o Initial promotion material (flyers, poster, roll-up) is already created and distributed to partners. Once per year we are planning to revise the content of the material and update according to the progress in the project activities.

3   Success metrics and KPIs

Achievement of scientific excellence and impact are at the core goals of BioExcel and as such they underlie the performance indicators that will be used to monitor the success of project. Three pillars form the basis of our approach for achieving this excellence and impact (Figure 2).
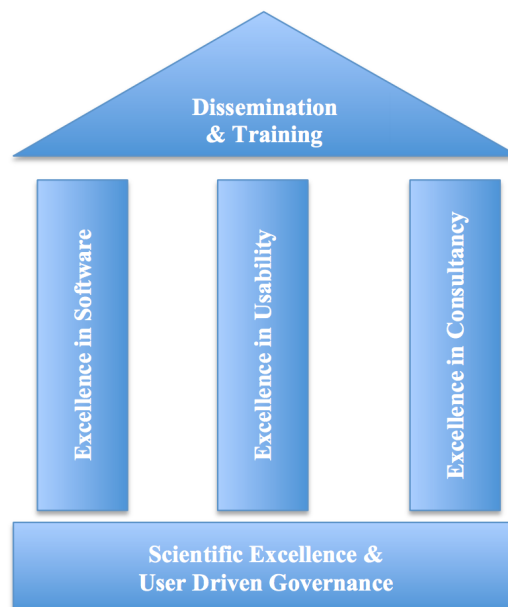

Figure 2. BioExcel Pillars of Excellence

In the following sections we will give a brief overview of the metrics for success in each of the pillars, while in section 3.4 we provide a more detailed description of the associated KPIs.

3.1 Software Excellence

Our software activities aim to improve the availability, efficiency, scalability and usability of the selected Life Science software packages of high importance and usage for the field.

Maintenance and possibly increase of scaling and efficiency on modern hardware as demonstrated by regular benchmarking, measured as absolute performance and relative scaling. This will depend heavily on the characteristics of next generation hardware. A baseline of 5% annual software improvement is expected from a subset of our applications, but actual improvement could be considerably higher. The baseline will be established with the first release of a benchmark suite.

Porting and improving performance/scaling of codes to new emerging architectures including accelerators. Depending on the availability and usefulness of new platforms we aim for two such architectures.

Increasing the quality of scientific software through professional Quality Assurance (specific targets for each application will be chosen after D1.1 at PM5). Quality will for instance be quantified by reporting on number and severity of bugs found and fixed through targeting testing on workflows found to be important to the members of the user groups, or reporting the annual increase in the fractions of the code bases covered by unit and regression tests.

Increasing use of HPC and HTC platforms and infrastructures (quantified through user-group reports of benefits derived from learning how to use the software better on current and new platforms and infrastructures).

Detailed description of the quality processes related to software development (e.g. repositories, code review, releases etc.) will be presented in D1.1 – Specification of software engineering, testing & QA.

3.2 Usability Excellence

A main objective of our activity along this pillar is to increase the usability of e-Infrastructures by enabling easy access to computing and data resources through workflow environments and associated data management tools and portals. Excellence in usability is ultimately linked to excellence in the software use and adoption. Those will be quantified through user-group reports of benefits derived from learning how to use the software better on current and new platforms and infrastructures.

3.3 Consultancy and training excellence

Our consultancy activity will enhance the uptake of computational technologies both in academic and industrial user groups and provide feedback and requirements to the other activities. The training and dissemination function will ensure that all stakeholders properly benefit from the CoE. This means targeting and interacting with academic and industrial end-users, software developers,

hardware and e-Infrastructure providers by developing workshops and training sessions for end users.

Increase effective usage of computational methods in the target sectors. This goal is difficult to quantify and set a numerical target mainly due to the following reasons. First, adoption of newly developed code is on a longer time-scale than the timeframe of the project. It would normally take several months even a year for users to be convinced in switching to a newer version; more time is needed to start using it exclusively; then apply for large-scale allocations on HPC resources on the basis of this version; and finally observe an increase in the usage. Second, it is difficult to obtain exact statistics from resource providers. Ideally, we'd like to see how many users are using which version of the software. However, some users don't use system wide modules rather own compiled binaries, and some centers don't have procedures for collecting statistics of what software has been used. Thus for our needs, we will monitor usage via surveys with user groups and outreach to HPC centers for collection of statistics.

Develop the various nascent bio-science communities who are only beginning to use HPC enabled modelling, simulation and data analytics. This goal is also difficult to quantify numerically. We will monitor via surveys and we expect to observe an increase in the number of participants to training workshops and number of communities they represent.

3.4 Key Performance Indicators (KPIs)

To assess the impact of the project, Key Performance Indicators have been defined and will be monitored and reported in the management reports. Table 2 provides an extensive list of the metrics, their likelihood, impact etc.

A note about WP1 KPIs. Our indicators stem from the Software Sustainability Institute's criteria-based assessment (https://www.software.ac.uk/sites/default/files/SSI-SoftwareEvaluationCriteria.pdf), a list of criteria used to evaluate quantitatively the quality of a software in a number of areas. We have adopted to use only metrics that are applicable to all software endorsed by BioExcel. Because it is hard to predict "numbers" that we expect to achieve in the course of BioExcel, hard to measure in a way that connects to BioExcel activity, and expected slow uptake before the project's final review, we have defined both reachable and stretch targets as a comparison with the current performances. For instance, a reachable target could be "sustainable growth rate" or "moderate performance improvement" while a stretch target could be "substantive growth rate" or "substantial performance improvement".

Table 2. Key Performance Indicators

ID	KPI	Justification	Additional details	Reachable target	Stretch target
	WP1				
WP1_KPI01	User documentation	Users need to know how to use both old and new code	Comprehensive documentation (in multiple formats) for all major features, e.g. with equations, background theory, links to scientific literature, useful parameter ranges, example applications, links to tutorials ...	90%	100%
WP1_KPI02	Developer documentation	Developers need to understand how the code works, and how the development process works	inline documentation (e.g. Doxygen)	All new methods have inline documentation	New methods and a number of appropriate old methods have inline documentation
WP1_KPI03	Developer policy	Developers need to know the development procedure and the requirements for the code	Developer procedure is fully documented, up to date, and available online	100%	100%
WP1_KPI04	Architectural overview	Clear map of the code workflows helps users and developers understand the impact of their choices	Software-interaction diagrams are available for high-level code	at least 70% of the modules implemented in or used by the software are described in the map	100%
WP1_KPI05	Buildability/Installability	Users and system administrators need to be able to install BioExcel software on supported platforms/OS	Distributions have instructions for compiling (if needed) and installing the software. All mandatory third-party dependencies are listed with suitable versions and licences	Clear simple-case installation instructions are available for all supported platforms/OS	Background information for handling more complex installation scenarios is available
WP1_KPI06	Portability	Users need the software to work on multiple platforms	List of platforms supported and tested - Use open standards and function with third party software	sustain the number of still-viable platforms supported/tested for each formal release	increase the number of still-viable platforms supported/tested for new release
WP1_KPI07	Testability	Developers need confidence that their changes don't break old functionality. Users need to know that new versions will continue to be at least as reliable as old ones.	Project has unit tests, integration tests, workflow tests	All new code has appropriate test coverage, e.g. unit, integration, regression or domain-specific.	Continuous integration is supported – tests are automatically run whenever the source code changes

WP1_KPI08	Performance	When not introducing new features, code changes should improve the performance of the software	Simulations in the benchmark set improve relative performance with former release. Improvement should be measured on identical hardware between old and new code versions running in their respective optimal configurations.	sustainable/moderate improvement in performance (could be elapsed time and/or accuracy) over the benchmark set	significant improvement in performance over the benchmark set (>10%)
WPI1_KPI09	Usability	User interfaces should make simple things simple, and powerful things possible, without being too complex to test, document and understand	Setup workflows (and/or web servers) with flexible set of available parameters and tested workflow usage	Pre-defined scenarios embedded in easy-to-use frameworks with possibilities for fine-tuning for advanced usage	Number of pre-defined frameworks addressing each of the main fields of application
WPI1_KPI10	Availability	Regular releases of new versions with bug fixes and changed functionality is useful to users and developers (possibly in addition to BioExcel formal deliverables of released code)	Releases of new major version / bug-fix-only versions	1 every 2 years / 2 per year	1 every year / 5 per year
WPI1_KPI11	Maintainability	Some changes might occur at the core of the software and make its usage significantly different from the previous versions. We need to handle this transition in the smoothest way to minimize the impact on the user experience	Make previous stable versions of the software available on demand with proper documentation associated. Might need self-contained binaries for outdated dependencies	Stable version running over the last 2 years	All stable versions released previously
	WP2				
WP2_KPI01	Number tools registered in Bio.tools having the BioExcel tag	The number of tools identified and registered by BioExcel defines and publicizes the power of the centre of excellence software infrastructure. Elixir's Bio.tools registry assures to reach a wider audience within the bioinformatics community.	registration	15 total for the project	25 total for the project
WP2_KPI02	Number of VMs and Docker containers stored in EGI having the BioExcel tag.	Building blocks registered and stored as either VMs or Docker containers in EGI AppDB guarantee an easily	availability	10 total for the project	15 total for the project

		deployment and usage. EGI AppDB has been chosen as a well-known and stable infrastructure for software distribution.			
WP2_KPI03	Software library coverage. Number and percentage of defined building blocks used.	The software library collected (D2.1) covers a large portion of what is needed for the development of the CoE portable environments for computing part and in particular for the progress of the pilot use cases. Coverage of this library is linked to the completeness of BioExcel software infrastructure.	usage (by BioExcel/community?)	70% (from the list of modules of D2.1)	90% and above (updated list)
WP2_KPI04	HPC and cloud interaction. Number of demonstration workflows installed in HPC supercomputers (e.g. Marenostrum) and/or virtualized environments (e.g. Cloud computing).	Deployment and running of the different pilot use cases in HPC and/or virtualized environments is a measure of what the centre (and the partners involved) can do in computational biomolecular transversal studies (HPC, clouds, workflow managers, databases, etc.) , and also proves the expertise that can be offered to the users of the CoE.		5 (corresponding to the project pilot use cases)	10
WP2_KPI05	Workflow manager coverage. Number of workflow managers used in BioExcel demonstration workflows.	A CoE needs to have a broad expertise in the state of the art tools. Workflow managers are becoming crucial for the current computational biomolecular research projects. Examples of real workflows using different workflow managers will be very helpful for our users.		4 (corresponding to the ones stated in the D2.1 list of modules)	6
WP2_KPI06	Number of workflows implemented in BioExcel Project	The number of workflows up and running completed in the BioExcel project is a measure of success for the centre in terms of		5 (corresponding to the project pilot use cases)	10

		both share of expertise and tools findability, accessibility and interoperability.			
	WP3				
WP3_KPI01	Number of healthy interest groups running by end of project. Healthy implies that it meets all of the KPIs stated in the list of IG KPIs below (unless a given IG states otherwise in its plans).	We need enough IGs to cover the space in which the CoE hopes to operate (but not so many that we cannot put sufficient effort into engagement). This KPI is intended to reflect the extent to which we cover the wider biomolecular community. Interest groups are generally considered "healthy" if they meet the KPIs for Interest Groups. These are listed separately at the end of this table.		6	10
WP3_KPI02	Number of people using the core public services, including consultancy	Over the course of the project we expect to pilot consultancy offerings. Resources are limited here but running this number would give us experience of this kind of offering, and help raise awareness of the centre.		15	30
WP3_KPI03	Number of people applying to calls for services for each of the three core applications.	We propose calls for various services at two levels: (i) core public services and (ii) potential enterprise services for paying "clients". These calls will be part of our customer research to understand user need, set priorities and to determine the demand for charged services. The number of people responding to such calls measures the interest in such services for each of the three core applications.		15	30

WP3_KPI04	Number of academic institutions represented in IGs (Total across IGs for project duration)	This KPI demonstrates the breadth of the impact of the IGs and measures one aspect of impact on the whole biomolecular communitiy.		20	30
WP3_KPI05	Number of commercial organisations represented in IGs (Total across IGs for project duration)	This KPI demonstrates the breadth of the impact of the IGs and measures one aspect of impact on the whole biomolecular community.		20	30
WP3_KPI06	Number of face-to-face IG meetings held (Total across IGs for project duration)	Face-to-face meetings are important for strengthening the feeling of membership of a group. These targets correspond to the 20 meetings proposed in the DoA. 18 meetings correspond to 1 meeting every 2 months on average.		15	25
	WP4				
WP4_KPI01	Scope of available training material	Training material needs to cover sufficiently the available functionality of the software		80% of newly developed functionality within the project or older but popular one, based on user feedback and IG input	100% + additional as requested by the community
WP4_KPI02	Number of CoE training events	Training events need to cover the main areas of expertise in the centre		2 per year, 6 in total	8 in total
WP4_KPI03	Number of people who have attended BioExcel courses	Courses need to cover topics of interest for the users		80% of maximum number of places	100% of maximum places
WP4_KPI04	Course feedback	Rating of the training course based on the participant course feedback survey		> 60% responses "Good" or "Excellent" (based on 30 - 50 % response rate for survey filled in at the event)	> 80% responses "Good" or "Excellent" (based on 30 - 50 % response rate for survey filled in at the event)
WP4_KPI05	Long-term impact, measured through post-6 month survey	Has the training received made the user more effective in their day-to-day activities e.g. have they gained a higher level of competence		> 60% responses "Useful" or "Very Useful" for day-to-day activities (based on 15% - 25% response rate)	> 80% responses "Useful" or "Very Useful" (based on 15% - 25% response rate)

WP4_KPI06	Presence in media	Raise awareness about our services to the wider community		* 1 publication per year in general / broad public journals * 8 tweets per month	* 2 publications per year in general / broad public journals * 10 tweets per month
WP4_KPI07	Number of website visits	This KPI helps to determine whether we are increasing our reach as well as when peak interaction periods are		600 per month	1200 per month
WP4_KPI08	Promotion of scientific publications	Are we effectively highlighting the achievements of the project		blog + tweet per publication	video recording per publication
WP4_KPI09	Number of Twitter followers	Similar to WP4_KPI08, determine whether our reach is increasing		100/130/170 for PY1/2/3 (i.e. 30% increase)	100/150/230 for PY1/2/3 (i.e. 50% increase)
WP4_KPI10	Number of subscribers to newsletter	Similar to WP4_KPI09, WP4_KPI10 determine whether our reach is increasing		300/400/500 for PY1/2/3 resp. (i.e. 30% increase)	300/450/650 for PY1/2/3 resp. (i.e. 50% increase)
WP4_KPI11	Number of publications acknowledging the BioExcel project	Measurement of the impact the project has through publications		4 per year	8 per year
	WP5				
WP5_KPI01	Number of HPC centers contacted	To get estimates of the number of the biomolecular researchers using HPC, current challenges in deployment of open-source software, technological trends, use of HPDA, challenges or approaches of HPC cloud delivery		3 per year	6 per year
WP5_KPI02	Number of academic and government users approached	To understand interest in proposed support and service offerings, challenges in HPC use, type of HPC use		4 per year	8 per year
WP5_KPI03	Number of Independent Software Vendors (ISVs) approached	To understand interest and requirements for software co-development, service requirements		2 per year	4 per year
WP5_KPI04	Number of industrial users approached	To understand interest in proposed support and service offerings,		2 per year	4 per year

		challenges in HPC use, type of HPC use			
WP5_KPI05	Number of organisations similar to BioExcel studied	To compare different options of organisational setup, identify overlapping services/communities		2	4
WP5_KPI06	Number of other CoEs engaged with	To discuss sustainability planning, sharing of knowledge		2	4
	IGs specific				
IG_KPI01	Number of posts per month on forums in ask.bioexcel related to each IG	The forum needs to be actively used in order for it to be a useful resource, and for people to consider posting and reading. Average activity is more important than activity of any given user.		10	15
IG_KPI02	Number of members for each IG	IGs need a minimum number of participants to justify support and resources. These numbers would make it worthwhile to hold a face-to-face event.		30	40
IG_KPI03	Number of contributing members for each IG	Not all IG members need to contribute, and those that do need to contribute that often. This KPI is intended to exclude members who express an interest, but do not contribute. A contribution is considered to be a post on the forum (including responses) or attendance at an IG meeting.		15	20
IG_KPI04	Number of attendances at each type of IG-related webinar and online meetings over a 12 months period	Attendance at a webinar is a good indicator for how interesting the IG is. This is total number of attendances (=number of attendees * number of webinars for an IG)		100	200
IG_KPI05	Number of attendances at each IG-related face-to-face events per year	A single IG-related event with 20 participants would make the IG worthwhile. Additional events with ongoing		20	40

		participation demonstrates ongoing engagement.			

4 Risk Management

Establishing a center of excellence is a complex task that carries risks not only associated with the execution of the project but also ones that are related to achieving the ultimate goal of a sustainable organization. The consortium has the experiences and required know-how in biomolecular Life Science, parallel applications, HPC (exascale) and HTC technologies, and providing support for academic and industrial user groups in the field, for a successful execution of the project work-plan and the setup of a sustainable CoE.

Yet, it is crucial to implement risk management processes in order to mitigate undesired outcomes. In Table 3 we present a list of identified potential risks, both for the execution of the project and the longer-term success of the CoE. We have put in place the following methodology for monitoring, analysis and action plan regarding those risks:
- WP leaders are responsible for continuously monitoring the progress along the objectives in their workpackages. The monitoring relies on sufficient communication with the corresponding WP members. Identified issues are brought up at the next executive board (EB) meeting. If there's been a consistent problem with intra-workpackage communication, the case is escalated to the responsible partner PIs, which are members of the project management board (PMB).
- During EB meetings the issues are brought up and their scope, severity of impact, priority for resolving and suggested action plan are discussed. If the risk concerns a single WP, its leader will be in charge to execute actions, as agreed by the board. If the scope covers more than a single workpackage, leaders of all affected ones are required to implement coordinated efforts and the necessary actions for resolving them.
- In case of conflicts, the issues are raised to the PMB and dealt with according to the conflict resolution strategy based on the Consortium Agreement, as outlined in D6.1 "Management plan".
- Risks with high rating (= Likelihood x Impact, see table XXX), are noted as "high risk factors" and are given special attention for monitoring and implementation of preventive or counteractive measures.
- Risks will be evaluated at all regular reporting periods, i.e. quarterly internal reports and yearly periodic reports.

The risk management plan thus ensures that risks are identified, brought to the attention of concerned personnel and counter measures are put in place. Close interactions between WP leaders within the EB, as well as at the higher management level of PMB, will guarantee that mitigation strategies are executed in due time. Monitoring and planning are supported by the roadmap (based on

the description of work) of milestones and deliverables, as well as the internal periodic reporting. We will also setup a database for tracking issues that have arised and the actions taken and their outcome.

Table 3. Risks and mitigation actions. WP is the affected workpackage. Its leader is responsible for monitoring the corresponding risk.

Abbreviations: L (Likelihood); I (Impact); R (Rating=Likelihood x Impact); Ops (Operational); Comm (Communications); Ext (External). Color coding of rating: Green=low, Yellow=medium, Red=High

Risk #	Description	WP	L	I	R	Proposed Mitigation actions
Risks for the project execution						
R1	Changes to future hardware roadmaps	WP1	1	5	5	If processor architectures change significantly, BioExcel will focus on the most important products in the market, or those for which the broadest end-user impact is expected. In this case we will alter priorities between accelerators, heterogeneous parallelization and scaling to focus on the largest impact for end users.
R2	Computational engine of HADDOCK (CNS - third party software) no longer developed / supported on new architectures	WP1	2	4	8	A complete rewrite / porting of the code to another computational engine will be required, coming probably at the cost of functionality. This would mean allocating efforts to this end, in close collaboration with the GROMACS developers (a possible alternative to CNS)
R3	CPMD does not support accelerators. If the presence of the accelerators will be widespread everywhere in supercomputer architectures in the next future, their usage could become crucial for using efficiently those machines, and CPMD is currently unable to use them.	WP1	3	3	9	CPMD, which deals with the quantum part of the calculations of its QM/MM interfaces, does not support accelerators. This could be mitigated by allocating efforts to implement it in the main code, in close collaboration with the IBM developers, the main maintainers of CPMD
R4	Discovery of missing requirements or finding of unsolvable steps during the development of pilot projects.	WP2	4	3	12	The development of the pilot projects will guide the generation of project requirements. The modular nature of the computational framework used will ease adaptation to new requirements.
R5	Workflow systems prove difficult to exploit HPC platforms of the project (supercomputers, grid, cloud).	WP2	3	3	9	Partners have extensive experience in a range of workflow and distributed system platforms, and a mixed portfolio of systems will mitigate the risk.
R6	Workflow systems prove difficult to run in the platform infrastructures of the project (virtual machines, docker containers, web interfaces).	WP2	3	3	9	Partners have extensive experience in the most used infrastructures, and a mixed portfolio of systems will mitigate the risk.
R7	Inability to automatically check for incorrect intermediate input/output data in workflows	WP2	4	4	16	Modular tools as well as entire workflows will have software testing procedures integrated, following the rules established by Elixir project.
R8	Workflow systems prove difficult to run with intermediate user interaction	WP2	4	3	12	User interaction possibility will be taken into account during the design and development of the different workflows in the project. Particular workflow managers having this capability will need to be chosen in these certain cases.
R9	Interest Group membership proves difficult to attract	WP3	3	4	12	Target audiences reviewed and greater effort apportioned to developing the communities. Refocus offer to membership.

R10	Initial services offered by CoE are not attractive to users	WP3	3	4	12	Discuss regularly with WP5 and WP6 how user requirements can be met by the service offering. Ensure that service roll-out includes monitoring to allow us to learn from initial roll-outs. Change service offering if necessary. Ensure service offering is clearly communicated to potential users.
R11	Lower than expected engagement with Ask Bioexcel Forums	WP3	4	3	12	Regularly monitor forum engagement. Ensure all questions on forums are addressed. Seed forums with questions and discussion points. Link to forums from all available sources, including specific forum posts related to each webinar. Finally, close un-used forums and re-focus on others where there is more interest.
R12	Potential users locked-in to existing communities and support structures	WP3	3	3	9	Ensure promotion of our offering specifically targets useful features of our offering that others don't have. Look for means to integrate with existing support infrastructure. Finally, if the external offering remains more compelling, consider withdrawing from active support in areas where it is difficult to compete, and promote existing support mechanisms.
R13	Aim to address broad user community allows less time to focus on important user groups	WP3	3	3	9	Monitor engagement with IGs. Focus on a smaller number if necessary to ensure that popular groups are fully supported. Consider different types of IG and/or community engagement mechanisms that require less effort.
R14	Consultancy proposals may not align with proposed business models	WP3	2	4	8	Frequent communication with WP5 will help to mitigate this. Community engagement can help to manage user expectations. Business models can be adapted in reponse to user-driven consultancy proposals. Ensure user input includes priorisation and could be used to evidence demand for services; look for consultancy models & content that people would be prepared to pay for.
R15	Failing to deliver project results and training to relevant audience	WP4	2	4	8	Training activities will be closely aligned with those of related initiatives such as CECAM, PRACE PATCs, EGI, EMBL-EBI, Elixir/Instruct etc. Further, the quarterly internal reports will allow tight monitoring of the progress and allow to take early action when required.
R16	Low numbers of delegates at training events or workshops	WP4	3	5	15	Effective promotion of events beginning several months before the event; developing training that is specifically aimed towards filling perceived (and real) gaps in training
R17	Lack of involvement of other partners/WPs in Training (WP4)	WP4	3	4	12	WP4 is liaising with the other partners and WPs to make sure that training events are planned when sufficient resources are available. Events will be planned over a 6-months timeframe. External trainers will be brought in when needed
R18	Failure to engage key audiences	WP6	3	4	12	We have a database of contact points with the different communities described in D3.1, and each contact point is connected with a center partner. In case of lack of engagement, the situation will be evaluated by the EB and corrective engagement strategy will be executed via the contact point personnel.
R19	Non-availability of the BioExcel website or the associated software/codes	WP6	3	5	15	The website is hosted in Sweden with automated backups taken daily. Additional server is available in Spain and can be used as a backup in case of emergency. All web-infrastructure (website, forums, social channels etc.) has multiple people with administrative privileges for support.
R20	Partner is not competent to carry out allocated tasks.	WP6	1	5	5	Partners have been carefully selected based on different required expertise (high performance and high throughput computing, biomolecular life science, software development, scientific & industrial applications), track record in their field (number of scientific publications and citations for research partners; level of innovation; academic excellence) and balance of the consortium. The consortium agreement includes measures to be taken if a partner still would not deliver for some reason, such as replacement of this partner by another one and a corresponding budget reallocation.

ID	Risk	WP				Mitigation
R21	Key staff leave or are unexpectedly absent (WP leaders, EB members etc)	WP6	3	4	12	Regular communication, deputies if applicable and in case of resignation a hand-over period will be used to facilitate a smooth transition. Staff might be more likely to leave towards the end of funding (e.g. for new projects) - to mitigate more than 1 staff at each partner should be kept in the loop (e.g. lurk on mailing lists, local update meetings).
R22	Individual consortium members do not effectively manage competing demands, timetable delays as a result of inadequate project management	WP6	3	2	6	As Work Packages and the effort at each partner is not too tightly coupled, the effect of this risk factor is lowered. However there are higher risk closer to project deadlines (e.g. multiple deliverables). We counter the risk by continual board updates and multiple draft/review cycles on deliverables.
R23	Shared documents are not efficiently maintained, e.g. access problems, inability to locate document or correct version)	WP6	1	5	5	We have already established a good working procedure using Trello, Google Drive folder and archiving in document management system.
Risks for the long-term success of the CoE						
		WP1				
R24	Alternative codes make existing codes less relevant	WP1, WP2	3	3	9	While it's likely that the popularity of codes will change over time; WP2's code-agnostic architecture can adapt new tools as they come along. As for WP1, the optimization and benchmark work will still be benefitial to existing code users; but WP1 priorities might have to be revised after testing the performance and scope of the new code.
R25	Some of the software components are not available anymore, or fail to fulfil the project requirements	WP2	4	3	12	The software components have been selected carefully but in case of unavailability of any of them, alternative components will be used. Capturing software tools using Docker and VM images ensures continued availability.
R26	Inability to integrate newly created tools and workflow managers in the CoE infrastructure; Inability to stay technologically up-to-date in the computational biomolecular research field.	WP2	2	4	8	Use the different field expertise of the project partners (workflow managers, simulation tools, genomics, web interfaces, etc.) to be informed and always up-to-date in the different fields, and work together in integrating solutions.
R27	Lack of acceptation of the workflow solutions offered by the CoE.	WP2	4	4	16	Partners are extremely well positioned in their respective fields of expertise. Their present tools have already a significant user community that will easily accept the newly generated pipelines as a natural evolution of the already existing tools. On top of that, the workflow development will be directly connected to Elixir recommendations to reinforce the visibility, integration, interoperability and availability.
R28	Consultancy project funding low	WP3	3	4	12	Business model reviewed and remedial actions taken. The HPC centre partners all have experience of operating in a multi-funded environment. Key management task will be to carefully monitor the development over time of consultancy and other funding lines.
R29	Users prefer less centralised and/or more focused support mechanisms	WP3	3	3	9	Re-focus on areas where we can make most impact. Ensure that we make the most of opportunities afforded by integrated service offerings. Consider collaboration and integration with other popular support infrastructures.
R30	Insufficient level of engagement with related national/international initiatives	WP4	3	3	9	Outreach and promotion with potential collaborators will be done via direct communication rather than using mass-market approaches such as mailing lists. Such personalised engagement is expected to nurture long-term interactions and create opportunities for more projects and high ROI. Training IG will encourage engagement.
					0	
R31	Not enough users are taking advantage of BioExcel to reach critical mass for sustainability.	WP5	3	5	15	The major focus of the initial implementation and roll-out plan of the CoE is the establishment of an efficient marketing strategy to reach potential users across all target segments. If these efforts are not sufficient to maintain a required user base, the CoE will adapt the

						business model and services offered to match user demand.
R32	BioExcel is unable to recruit sufficient number of FTEs through partners to satisfy short-term project demand.	WP5	4	2	8	Link with external service providers and consultants to access external pay per service expertise. Adjust business model to match project capacity to internal and external resources.
R33	Public funding limitations	WP5	3	5	15	The budget requires that membership fees and consultancy fees are complemented by public funding in order to sustain the centre of excellence. If this funding cannot be attracted a smaller version of the centre will have to be executed instead.
R34	Economic climate disfavours investment in HPC use in general and software expertise in particular.	WP5	3	5	15	Adjust business model to meet current project demand.