H2020 EINFRA-5-2015

www.bioexcel.eu

Project Number 675728

D4.2 – Competency framework, mapping to current training & initial training plan.

WP4: Dissemination and Training

Document Information

Deliverable Number	D4.2
Deliverable Name	Competency framework, mapping to current training & initial training plan.
Due Date	Original due date 2016-06-30 (PM8). Delayed until 2016-07-31 (PM9)
Deliverable Lead	EMBL-EBI
Authors	Vera Matser (EMBL-EBI)
Keywords	Training
WP	WP4
Nature	Report
Dissemination Level	Public
Final Version Date	2016-07-30
Reviewed by	Steven Newhouse (EMBL-EBI), Cath Brooksbank (EMBL-EBI), Rossen Apostolov (KTH), Erwin Laure (KTH), Bert de Groot (MPG), Alexandre Bonvin (UU)
MGT Board Approval	2016-07-30

Document History

Partner	Date	Comments	Version
EMBL-EBI	2016-06-15	First version	0.1
EMBL-EBI	2016-07-13	Second version incorporating EB comments and adding initial mapping and gap analysis report.	0.2
EMBL-EBI	2016-07-20	Third version incorporating additional feedback after adding further content	0.3
KTH	2016-07-26	Formatting and style changes	0.4
KTH	2016-07-30	Final approved version	1.0

Executive Summary

This document describes how the BioExcel Centre of Excellence has worked with its users to agree on the competencies that they require to make the most of HPC for their research. To enrich the competency profile we have, for each competence, defined what an individual will need to know and what skills they need to have to exhibit competence in a specific area, as well as listed what behaviours are suited and unsuited to an individual with that particular competency.

We have performed a training needs analysis to gather the competency requirements for BioExcel's users. These users include: entry-level users (e.g. bench-based molecular life scientists with a very limited computational background, working in both academic and industrial settings); expert users (e.g. structural bioinformaticians/computational chemists, medicinal chemists and others) and computational service providers, including systems administrators, applications experts and others involved in providing high-end computational services and software to biomolecular scientists.

An initial workshop was organised to outline the competencies required by these users to make the most of HPC, and the main outcome of the workshop was a matrix of the different target groups against the agreed competencies. A total of 31 competencies have been defined; 5 Generic Competencies, 13 Scientific Competencies; 8 Generic Computing Competencies and 5 Parallel Computing Competencies.

We are in the process of mapping these competencies to the learning outcomes of courses already provided by existing training programmes so that we can focus on developing new training that meets currently unmet needs. Though a large number of the competencies have existing courses mapped against them the coverage is often partial or fragmented. At the interim analysis stage (refers to D4.2 versus final analysis in D4.3 due in month 12) 10 out of the 31 competencies are insufficiently covered; based on having less than 5 training resources mapped against the competencies. Taking into account the partial mapping, the number of insufficiently covered competencies increased to 19.

Initial plans for the BioExcel Training Programme are presented with mapping against the relevant competencies. The work (task 4.1) underlying this deliverable formally runs until month 12 and the complete results with recommendations for the BioExcel Training Programme will be presented in D4.3 (PM12).

Contents

1 Introduction

The D4.2 deliverable builds on the ongoing work in Task 4.1 (which formally ends in month 12).

The aim is to define the competencies that users of BioExcel need to develop, to enable them to fully exploit the applications provided by the project. Figure 1 outlines the different steps in the development of the BioExcel Training Programme.

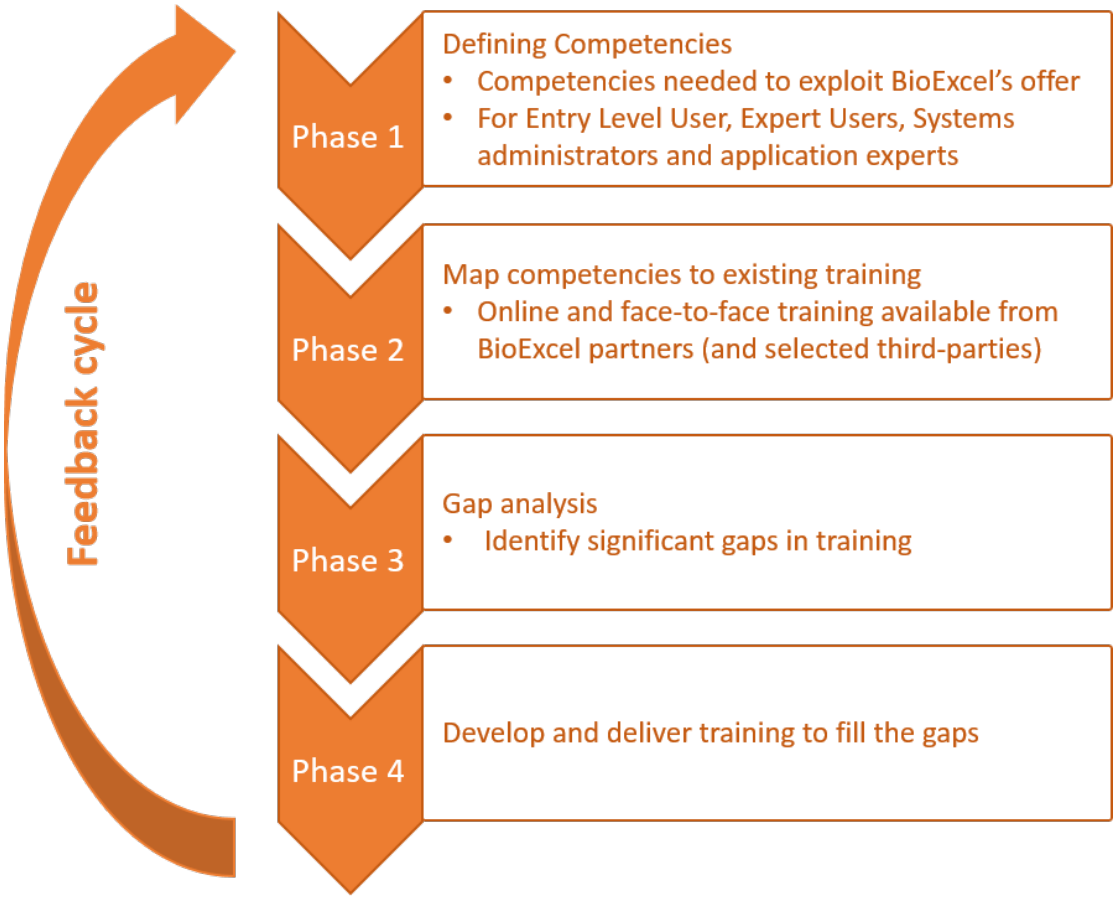


Figure 1: The steps taken to develop the BioExcel Training Programme

The aim of the work described here is, through cooperation with WP3, to consult with the potential users of BioExcel to identify the training needs and preferences of the academic and industrial users for training material and its delivery. It will thus guarantee that effort is spent in direction of addressing those exact needs.

Each competency will be mapped to training already provided by the BioExcel partners (and selected third-party resources) to create a matrix, which will allow us to identify training gaps.

The work outlined in this document as well as any additional work carried out on the competency framework, mapping and the training plan, together with

further consultation with our user groups and the wider computational biomolecular research community (survey and direct contact), will be used to construct a white paper describing our recommendations for the future development of a BioExcel training programme (D4.3) which will be delivered in month 12.

2 Competency Framework

2.1 Concept of Competency

A Competency is an observable ability of any professional, integrating multiple components such as knowledge, skills and behaviours.

The key aspects of the competency-based approach are:

• Competencies are observable, so acquisition can be validated objectively
• Evidence of competency can be collected in a competency portfolio
• Competencies are shared 'currency' applicable to learning of all types and at all career stages

2.2 Initial draft competency profile

Ahead of the workshop in Hinxton (see below) we prepared an initial draft profile based on existing competency profiles which had overlapping areas of competence (e.g. ISCB, CORBEL, RItrain) and discussions within the consortium. The aim of the initial draft profile was to speed up the progress that could be made during the workshop.

In preparation of the workshop the decision was made to use Entry Level Users, Expert Users and Systems Administrators / Application Experts as the initial target user groups. They will be part of the stereotypes and roles that will be used to define the service portfolio of the centre. The following profiles highlight what motivates these user groups.

Entry-level users:
An example might be a discovery biologist who wants to find potential interaction partners for a new target, or perhaps wants to see whether a library of small molecules could disrupt a potentially toxic protein-protein interaction. Potentially such a researcher might never have used modelling software before.

For this group, we need to deal with their 'unknown unknowns'; we wouldn't, for example, expect a discovery biologist to know how to tune a massively parallel computational problem to make it run efficiently. However, if we can teach them how to find out whether or not something is running efficiently, and who they

need to talk to in order to solve this problem, we would already be solving a major source of frustration.

Expert users:
They are modelling and simulation experts who want to do advanced and large scale modelling and simulations (for example, a computational chemist might want to run a huge chemical library past the entire PDB to find potential targets); perhaps they are using their cluster at the university and are eating up all the compute time; they might be looking for ways to fine tune their experiments, or to move them into a more computationally powerful environment (e.g. PRACE or EGI). The specifics of how these people optimise their simulations is specific to the application that they are using.

Systems Administrators and Application Experts (abbreviated to Systems Administrators):
Finally, we have the compute experts providing HPC services. They benefit from understanding enough biology to make sense of what their users want to do and why they want to do it; they need to understand how to optimise their compute architecture for the types of computational problems that biologists face (e.g. traditional high performance computing is not appropriate in situations where calculations are data-bound rather than cpu-bound. Data storage and management are crucial. Specific calculations depend on efficient input/output. Transmitting large data sets across a network may be problematic. Typically, the calculation is brought to the data, rather than the other way round.).

2.3 Competency Workshop

A 2-day workshop was organised at EMBL-EBI in Hinxton, 3-4th May 2016 titled "BioExcel: addressing training needs for advanced simulations in biomolecular research" to gain input into the competency framework.
The programme is available at

http://www.ebi.ac.uk/training/events/2016/bioexcel-addressing-training-needs-advanced-simulations-biomolecular-research.

Care was taken to make sure that the different BioExcel Interest Groups were well represented. A total of 19 people attended the workshop (Appendix 1).

A secondary purpose of the workshop was to provided participants with training in the three core BioExcel biomolecular modelling and simulation software packages: GROMACS, HADDOCK and CPMD. The user training also provided an incentive for participants to contribute to the training needs analysis.

In a number of sessions, we gathered information as input for the development of the BioExcel Training Programme:

- Expectations and bottleneck sessions: what do the participants expect from the workshop and what are the biggest bottlenecks they face in their research?
- Empathy mapping: what motivates different types of users and how can BioExcel help them (output in Appendix 2).
- Competency session part 1- defining the BioExcel competencies.
- Competency session part 2 - defining knowledge, skills and behaviour for each competency.

For each competency the participants rated whether they felt Entry Level and Expert Users needed Awareness, Working Knowledge, Specialist Knowledge or whether the competency was irrelevant. Definition of the levels can be found below:

- **Awareness:** the professional appreciates what is possible in this area and how the area impacts on their own work
- **Working knowledge:** the professional has a firm underpinning (Knowledge Resource Repository) in this area and applies it effectively in his or her day-to-day work
- **Specialist knowledge:** the professional actively contributes to advancement of the area, generating new understanding or new technology
- **Irrelevant:** the professional does not need skills or knowledge in this area to carry out his or her day-to-day work

There were no workshop attendees who felt strong affinity with the Systems Administrator group (polling during the workshop) so the breakout sessions focused on the Entry Level Users and the Expert Users. After the consolidation of the competency profile additional input was requested from representatives of the System Administrators group to make sure there are no gaps in the profile for that user group.

Major changes made to the competency profile, based on the workshop, were to split out (and expand upon) the computational competencies from the scientific competencies into generic computing competencies and parallel computing competencies.

The following actions were defined based on the workshop

- Consolidate the BioExcel Competency profile for the survey (completed; see Table 1).
- Ask for additional input for the Systems Administrators User Group (completed); 3 new competencies were added based on the input of the systems administrators (1.5, 3.7, 3.8).
- Create a survey to send out to a wider audience for feedback (completed; currently compiling contact list).
- Add missing Knowledge, Skills and Behaviours (ongoing).

Table 1 shows the version of the competency profile that was used as the basis of the survey.

Area	Competencies	Knowledge	Skills	Behaviours
1. Generic Competencies				
Aligned with ISCB competency profile	1.1 Function effectively in teams to accomplish a common goal.	**(to be added PM12)**	**(to be added PM12)**	**(to be added PM12)**
	1.2 Comprehend and comply with professional, ethical, legal, security, and social issues and responsibilities, and uphold these in the workplace as appropriate.	**(to be added PM12)**	**(to be added PM12)**	**(to be added PM12)**
	1.3 Communicate effectively with a range of audiences, technical and nonprofessional.	**(to be added PM12)**	**(to be added PM12)**	**(to be added PM12)**
	1.4 Engage in continuing professional development.	**(to be added PM12)**	**(to be added PM12)**	**(to be added PM12)**
	1.5 Identification and understanding of user needs.	**(to be added PM12)**	**(to be added PM12)**	**(to be added PM12)**
2. Scientific Competencies				
Best Practice	2.1 Apply computing expertise appropriate to the discipline and level of expertise (software as well as hardware).	**(to be added PM12)**	- Evaluates software - Understands applicability, optimisation and scope of different tools	**(to be added PM12)**
Best Practice	2.2 Apply expertise in natural sciences appropriate to the discipline.	- Has a deep comprehension of biological problems - Comprehends that a biological theory is not necessarily true	- Has an interdisciplinary view - Asks relevant, well-defined biological questions	- Takes a comprehensive approach to problems - Consults with colleagues - Looks for prior work - Displays scientific humbleness (own and other results) - Allows for unexpected results

Best Practice	2.3 Detailed comprehension of the scientific process and an ability to apply it (e.g. experimental design; inclusion of controls; processing, visualisation and interpretation of results).	- Knows where to find help - Comprehends the need for positive and negative controls - Comprehends the need for replicates (and what type, how many etc.) - Comprehends the principle garbage in, garbage out (GIGO) (for example generic or custom parameters for molecular dynamics) - Has a deep understanding of his/her own data	- Formulates a hypothesis - Tests a hypothesis - Accurately judges the validity of his/her results	- Asks for help and/or verification - Prepares before starting a new process
	2.4 Comprehension of, and compliance with, licensing policy.	- Is aware of the different types of licensing - Understands the significance and potential ambiguity of licensing depending on where you work (academia/industry, funding agency etc) - Understands the difference between open and closed licenses	- Chooses appropriate license for his/her own software	- Always checks the license information for software - Questions ambiguity in license information - Checks software is licensed
	2.5 Use a computer-based system, process, component, or program to meet desired needs in a biomolecular context.	**(to be added PM12)**	**(to be added PM12)**	**(to be added PM12)**
	2.6 Evaluate the ability of a computer-based system, process, component, or program to meet desired needs in a biomolecular context.	**(to be added PM12)**	**(to be added PM12)**	**(to be added PM12)**
	2.7 Comprehension of the local and global impact of high-performance	**(to be added PM12)**	**(to be added PM12)**	**(to be added PM12)**

	computing (HPC) and high-throughput computing (HTC) on individuals, organizations, and society.			
	2.8 Comprehension of how data-driven science, data analysis and computational modelling can be combined to generate and test hypotheses (e.g. machine learning, data mining, pattern recognition).	- Understands that science is an iterative process - Comprehends that models require experimental validation	- Accurately judges the validity of his/her results	- Comprehends that results might be difficult to interpret
Data	2.9 Identify and compile appropriate data sets to address specific research questions	- Has knowledge of different datasets - Is aware of the concept of unique identifiers - Comprehends the need to question / validate data	- Assesses data quality - Judges fit-for-purpose - Actively identifies new datasets - Tracks where data came from - Combines data from different sources - Data preparation/cleanup	- Advertises the existence of good quality datasets to others - Contributes to public datasets when appropriate
	2.10 Comprehension of, and compliance with, best practice in data management / organization / archiving and storage	- Knows and understands the data formats people are using - Is aware of the backup policy of the institution/compute resource - Ability to judge what data needs to be kept for storage	- Demonstrates knowledge of the file system structure - Tracks ever step in a process (traceability) - Version control - Backup - Documentation - Automates date analyses where appropriate	- Backs up his/her work in non-redundant ways - Documents his/her work - Has a version control system in place
	2.11 Comprehension of, and compliance with, best practice in distributed data management and data management planning	- Comprehends the need for standards (including ontologies) - Understands the importance of metadata - Comprehends the challenge surrounding legacy data (e.g.	- Annotates metadata (tools) - Uses ontologies - Structure data, where applicable submit to databases - Safely and efficiently moves	- Uses standards and ontologies - Looks for existing ontologies rather than create his/her own ontology

		inaccessibility, resourcing, IP) - When to move data, and when to run code elsewhere	data	- Captures sufficient metadata - Adds his/her data in supplimentary information of journal articles (e.g SMILES, InChI for compounds) - Actively promotes the use of standards, metadata annotation
	2.12 Search for, read and assess literature / manual	- Comprehend the need for a literature review (anything similar done? Data available?) - Is able to find relevant literature (different sources of) (e.g. PubMed, EuropePMC)	- Assessing quality (methods, results, etc) - Critically read scientific literature	- Conducts a literature review when appropriate - Will look at sources of literature outside his/her comfort zone / discipline
	2.13 Presenting your results to community (writing papers, conference presentation, YouTube)	- Knows which forums exist to present scientific results	- Write scientific journal paper - Write and present a conference talk - YouTube - Proficient at scientific poster design - Deep understanding of what to present and what not to present - Can target his/her presentation at the right audience / level	- Judge the appropriate medium to present through
3. Generic Computing Competencies				
	3.1 Comprehension of, and compliance with, good programming practice (as promoted, for example, by www.software-carpentry.org.	- Comprehends the need for best practice - Is aware of where (or with whom) to find examples of or written guidelines on best	- Inspires others to adhere to best practice	- Actively promotes best practice by leading by example - Will try and identify best practice in a relevant area

		practice in his/her field		
	3.2 Analyze a problem and identify and define the computing requirements appropriate to its solution (e.g., define algorithmic time and space complexities and hardware resources required to solve a problem).	**(to be added PM12)**	- Is able to benchmark a system to get an estimate of the required computing time	**(to be added PM12)**
	3.3 Apply knowledge of the operating system.	- Efficiently navigates his/her way around the Operation System (OS) - Is transitioning from Graphical User Interface (GUI) to command line (if not already proficient) - Demonstrates knowledge of the file system structure - Knows where to find the location of important configuration files & applications - Is aware of OS-related file format differences	- Creates and manages files and directories in a system using a Graphical User Interface (GUI) and using command line - Connects to remote systems with a GUI and the command line - Be able to use Unix/Linux features like pipes & redirection - Changes access permissions when required - Creates and manages files and directories in a system	- When you can't do something, search for it on the web - Back-up files - Uses access permissions appropriately
	3.4 Write/adapt computer programs (software development) for biomolecular simulations.	- Recognises common programming concepts like loops and function calls - Comprehends the pros and cons of different programming languages - Knowledge of existing code/libraries to re-use	- Adapt existing programs - Compile code - Run test - Recognises appropriate file types - Knows when quick and dirty is appropriate and when writing for reusability is appropriate - Program in an appropriate language	- Conform to best practices for writing re-usable code - Uses revision control - Uses a text editor with support for programming

			- Writes portable code - Writes efficient code - Writes appropriate tests - Debug code	
	3.5 Write his/her own scripts to perform tasks in context of biomolecular research.	- Knowledge of existing command/ibraries to re-use - Judges when a task should be automated	- Is able to automate the process of executing some processes remotely - Write & debug scripts	- Uses appropriate scripting language
	3.6 Install biomolecular simulation software on his/her computer.	- Knows where to download software & dependencies to install - Comprehend the difference between kinds of binaries & source - Awareness of different types of account (e.g. user, admin, etc.) and when they should be used - Deep understanding of his/her platform (hardware, OS version, ...)	- Select appropriately packaged code - Be able to use package managers - Be able to revert a system to a known state	- Reads the README - Checks licencing before installing/running software - Appreciates the impact of installing/running software on other users of the machine
	3.7 Deploy and test non-commercial software, including software that is built collaboratively and on a volunteer basis.	**(to be added PM12)**	**(to be added PM12)**	**(to be added PM12)**
	3.8 Apply knowledge of systems monitoring (e.g. queue monitoring, systems availability and optimisation, storage used; scheduling maintenance at appropriate times and communicating this to users).	**(to be added PM12)**	**(to be added PM12)**	**(to be added PM12)**
4. Parallel Computing Competencies				
Computational Workflow	4.1 Assess computational workflow	- Is aware that workflow	- Evaluates and selects an	- Seeks out and makes

Automation	systems and their potential benefits.	systems exist and what they do - Comprehends the pros and cons of different workflow systems	appropriate workflow system - Install workflow manager - Ablility to run a workflow - Write and modify a workflow	use of appropriate existing workflows and workflow components
	4.2 Apply knowledge of batch system	- Understands what resources are required per job - Demonstrates knowledge of the concept of queues and the runtime environment - Know about versions of software that are installed - Can estimate who is responsible for observed issues (a hardware issue? a software issue? a user environment issue?)	- Writes a good batch script making use of available batch system features	- Reads the documentation - Ask the right person for help
Software Development	4.3 Write computer programs that can run on a parallel computer	- Know available parallelisation libraries/languages and the architectures that they are best suited for - Knowledge of common parallelisation patterns	- Programs using an appropriate language/paradigm - Recognises (independent) units of work and potential parallelism (and any dependencies) - Uses appropriate tools to debug, profile, refactor parallel code	- Reads the documentation
Software Development/Deployment	4.4 Assess advantages and limitations for deploying, executing and optimising computations in a cloud/grid environment	- Evaluates the pros and cons of running computations in the cloud - Evaluates the pros and cons of using grid technologies - Demonstrates knowledge of container technologies (e.g. Docker)	- Routinely undertakes a cost-benefit analyis to understand whether cloud resources are the best value - Able to package software/data and deploy to the cloud - Selects an approriate	- Seeks out and makes use of appropriate existing Virtual Machines, etc.

		- Demonstrates knowledge of the concepts related to virtualisation	grid/cloud provider - Can build and provision a Virtual Machine (VM)	
Performance Optimisation	4.5 Apply knowledge of performance profiling to measure suitability of computing platforms	- Know that there are tools to help measure performance - Know how your machine's architecture and hardware can affect the performance of your code	- Be able to profile a code - Recognise bottlenecks in a code - Be able to refactor code in order to remove bottlenecks	- Run codes using appropritate resources (number of processors, etc.)

Table 1: The BioExcel Competency Profile (version used for the competency survey). The missing Knowledge, Skills and Behaviours will be added by month 12.

2.4 Competency Survey

To validate the BioExcel competency profile a survey will be sent out to the computational biomolecular community in collaboration with WP3 (due to be sent out in August).

We will ask respondents to choose whether they want to complete the survey for the Entry Level User, Expert User or Systems Administrator user group. If respondents feel that they can contribute for more than one type of user, e.g. an expertuser who employs entry level users, we ask them to fill in the survey multiple times.

The survey presents respondents with the BioExcel competencies one-by-one and they are asked to rate what level of expertise the user should have for this competency (Figure 2).

The same options were available as discussed during the workshop, however, we have added an **unknown** option in case respondents are not sure.

For each competency there is a text box to provide comments for that specific competency (e.g. the competency is unclear, should be split into two competencies or if an element needs to be added etc.).

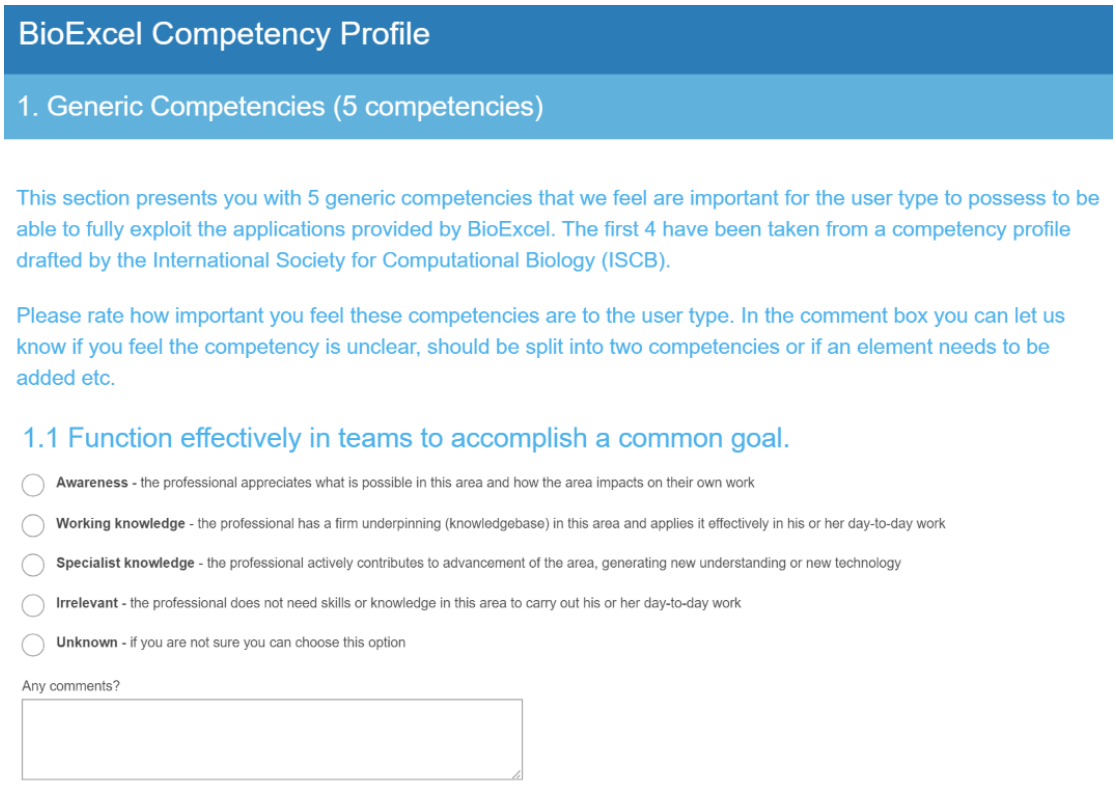


Figure 2: Sample (1st competency) of the draft BioExcel Competency Profile Survey. All other competencies are presented in a similar format.

In addition to the questions related to each of the competencies the following questions on training needs were added to the survey to inform the development of the BioExcel Training Programme:

- Please select what you would consider the highest priority training need for this user group (drop-down list of all competencies)
- Please summarise what this user group would be able to do (or be able to do better) as a consequence of the training (at least 50 words).
- What training formats do you feel would most benefit the user group?
- What resources do you currently go to when you need help or are trying to find a training course/material (e.g. google, specific training portals etc.)?

A pdf of the full survey can be accessed through the following link:

(ftp://ftp.ebi.ac.uk/pub/training/2016/BioExcel_2016/SurveyMonkey_Competency_BioExcel.pdf).

3 Mapping and Gap Analysis

The aim of defining the competency profile was to highlight what competencies users of BioExcel are required to have (or need to develop) to be able to fully exploit the applications provided by BioExcel. By mapping the competencies to training that is currently available, it will enable us to identify significant gaps in training. For this exercise we are focussing on training – both online and face-to-face courses/workshops – being made available through the BioExcel partners or selected third-parties such as training resources listed in major training portals (e.g. on-course®, www.on-course.eu and coursera, www.coursera.org).

The BioExcel CoE aims to highlight the training gaps and fill the training need if deemed appropriate. If BioExcel is deemed not to be the appropriate "body" to provide the training, we are able to highlight the need for training to collaboration partners.

3.1 Existing training

All BioExcel partners were asked to add their institutes relevant training courses to a google document; this will be an ongoing process when new courses become available. In addition to the BioExcel partner courses and online training material, relevant courses were extracted from the on-course® (www.on-course.eu) database. The initial selection was based on short courses (Continuing Professional Development - CPD) and a combination of the following five search terms: computational approaches, chemical biology, structural biology, systems biology, drug discovery. After cursory manual curation 43 courses were considered relevant to BioExcel users and these courses have been added to the existing training resources sheet. A total of 39 courses or specializations from the

coursera (www.coursera.org) online offer have been tagged as relevant. We anticipate that a lot of information collected in this google document can be used to populate the BioExcel Knowledge Resource Repository.

The list of existing training resources that was used for the initial mapping exercise is available in Appendix 3.

It is important to keep in mind that the aim of the mapping exercise is not to find all relevant training courses but to identify where the gaps are.

Similar to the competency profiling, we are using a methodology for the competency mapping that has been implemented by a number of projects (e.g. CORBEL, RItrain).

3.2 Initial Mapping and Gap analysis

Figure 3 provides an overview of the initial mapping of available training resources to the BioExcel Competencies. There are a number of very important considerations to be kept in mind when interpreting this figure:

- Figure 3 provides a falsely optimistic view of the training coverage owing to the fact that a large number of training resources provide only a partial map to the competencies, e.g. data visualisation under competency 2.3.
- The red line (5 training resources) is considered a minimum number of training resources that should be available for each competency. Some competencies will be considered insufficiently covered with higher numbers of training resources due to the partial coverage described above.
- For some competencies it may be hard to find specific training resources because they are likely to be incorporated into courses on different topics or competencies.
- The mapping exercise does not take into account the quality or frequency of the training resources.

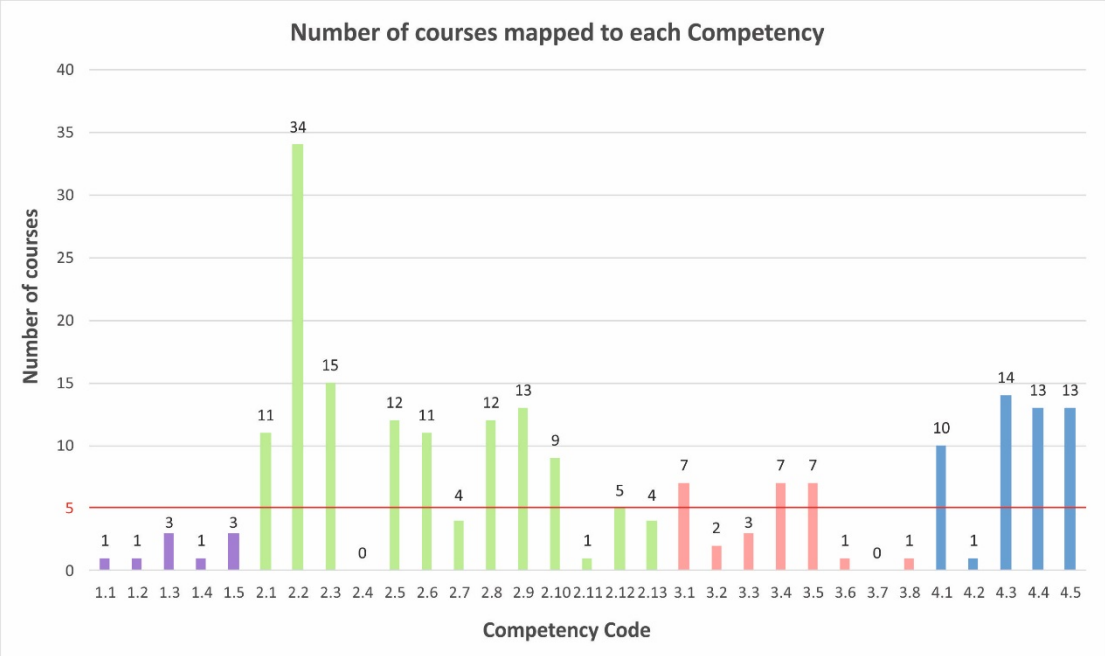


Figure 3: Initial mapping of available training resources to the BioExcel Competencies. The colours refer to the competence areas; purple = Generic Competencies, green = Scientific Competencies, pink = Generic Computing Competencies, blue = Parallel Computing Competencies

In Table 2, the following key has been used to summarise the coverage provided by the identified training resources. Where appropriate additional comments are provided to highlight important information pertaining to the gap analysis.

Coverage	Insufficient training resources identified but outside BioExcel scope	BioExcel or External training resources identified	Insufficient training resources identified

Competency (code and full description)	**Courses Mapped**	**Coverage**	**Gap Analysis Comments**
1.1 Function effectively in teams to accomplish a common goal.	1		Insufficient training resources identified but outside BioExcel scope*
1.2 Comprehend and comply with professional, ethical, legal, security, and social issues and responsibilities, and uphold these in the workplace as appropriate.	1		Insufficient training resources identified but outside BioExcel scope*
1.3 Communicate effectively with a range of audiences, technical and lay.	3		Insufficient training resources identified but outside BioExcel scope*
1.4 Engage in continuing professional development.	1		Insufficient training resources identified but outside BioExcel scope*
1.5 Identification and understanding of user needs.	3		User profiles, case studies etc *

2.1 Apply computing expertise appropriate to the discipline and level of expertise (software as well as hardware).	11		Mainly external resources
2.2 Apply expertise in natural sciences appropriate to the discipline.	34		Partial or fragmented coverage. Additional searches for specific courses suited to the BioExcel relevant domains would be beneficial
2.3 Detailed comprehension of the scientific process and an ability to apply it (e.g. experimental design; inclusion of controls; QA, QC, processing, visualisation and interpretation of results).	15		Partial or fragmented coverage. Not many courses take participants across the full scientific process. More resources needed on controls, QA, QC
2.4 Comprehension of, and compliance with, licensing policy.	0		More detailed search needed. Topic would lend itself to webinar, written guidance and online question sessions (e.g. ask.bioexcel.eu)
2.5 Use a computer-based system, process, component, or program to meet desired needs in a biomolecular context.	12		Partial or fragmented coverage. BioExcel resources on specific resources, mainly online tutorials
2.6 Evaluate the ability of a computer-based system, process, component, or program to meet desired needs in a biomolecular context.	11		Partial or fragmented coverage. BioExcel resources on specific resources, mainly online tutorials
2.7 Comprehension of the local and global impact of high-performance computing (HPC) and high-throughput computing (HTC) on individuals, organizations, and society.	4		Possible topic for BioExcel Webinar
2.8 Comprehension of how data-driven science, data analysis and computational modelling can be combined to generate and test hypotheses (e.g. machine learning, data mining, pattern recognition).	12		
2.9 Identify and compile appropriate data sets to address specific research questions	13		Partial or fragmented coverage. QA,QC aspect lacking
2.10 Comprehension of, and compliance with, best practice in data management / organization / archiving and storage	9		
2.11 Comprehension of, and compliance with, best practice in distributed data management and data management planning	1		More detailed search needed. Likely target for collaboration with other projects/institutes
2.12 Search for, read and assess literature / manual	5		Partial or fragmented coverage. Additional resources need for critical assessment of literature

2.13 Presenting your results to community (writing papers, conference presentation, YouTube)	4		Search for additional resources would be beneficial but training falls outside BioExcel scope
3.1 Comprehension of, and compliance with, good programming practice (as promoted, for example, by www.software-carpentry.org.	7		Partial or fragmented coverage. In scope for a computational skills course.
3.2 Analyze a problem and identify and define the computing requirements appropriate to its solution (e.g., define algorithmic time and space complexities and hardware resources required to solve a problem).	2		More detailed search needed. In scope for a computational skills course.
3.3 Apply knowledge of the operating system.	3		More detailed search needed. In scope for a computational skills course.
3.4 Write/adapt computer programs (software development) for biomolecular simulations.	7		Partial or fragmented coverage. In scope for a computational skills course.
3.5 Write his/her own scripts to perform tasks in context of biomolecular research.	7		Partial or fragmented coverage. In scope for a computational skills course.
3.6 Install biomolecular simulation software on his/her computer.	1		More detailed search needed. In scope for a computational skills course.
3.7 Deploy and test non-commercial software, including software that is built collaboratively and on a volunteer basis.	0		More detailed search needed. In scope for a computational skills course.
3.8 Apply knowledge of systems monitoring (e.g. queue monitoring, systems availability and optimisation, storage used; scheduling maintenance at appropriate times and communicating this to users).	1		More detailed search needed. In scope for a computational skills course.
4.1 Assess computational workflow systems and their potential benefits.	10		Available resources (PRACE/PATC) are at intermediate level and face-to-face
4.2 Apply knowledge of batch system	1		Available resources (PRACE/PATC) are at intermediate level and face-to-face
4.3 Write computer programs that can run on a parallel computer	14		Available resources (PRACE/PATC) are at intermediate level and face-to-face
4.4 Assess advantages and limitations for deploying, executing and optimising computations in a cloud/grid environment	13		Available resources (PRACE/PATC) are at intermediate level and face-to-face
4.5 Apply knowledge of performance profiling to measure suitability of	13		Available resources (PRACE/PATC) are at

computing platforms			intermediate level and face-to-face

Table 2: Overview of the initial mapping and gap analysis for the BioExcel competency profile.

* We will be able to point users at RItrain and CORBEL; owing to their remit, these projects are more suited to developing training for this area. Through collaboration with these projects we are already aware of training that is being developed. Many host institutions will have their own training courses for employees. What BioExcel will be able to do is point to a number of good examples if institutes are trying to develop training to cover these competences.

The following competencies are insufficiently covered (clustered by area):

1. Generic Competencies

1.5 Identification and understanding of user needs.

2. Scientific Competencies

2.4 Comprehension of, and compliance with, licensing policy.
2.7 Comprehension of the local and global impact of high-performance computing (HPC) and high-throughput computing (HTC) on individuals, organizations, and society.
2.11 Comprehension of, and compliance with, best practice in distributed data management and data management planning.
2.13 Presenting your results to the community (writing papers, conference presentation, YouTube).

3. Generic Computational Competencies

3.2 Analyse a problem and identify and define the computing requirements appropriate to its solution (e.g., define algorithmic time and space complexities and hardware resources required to solve a problem).
3.3 Apply knowledge of the operating system.
3.6 Install biomolecular simulation software on his/her computer.
3.7 Deploy and test non-commercial software, including software that is built collaboratively and on a volunteer basis.
3.8 Apply knowledge of systems monitoring (e.g. queue monitoring, systems availability and optimisation, storage used; scheduling maintenance at appropriate times and communicating this to users).

4. Parallel Computing Competencies

4.2 Apply knowledge of batch system

In Section 5.2 of this document - BioExcel Training Programme - we discuss how BioExcel plans to address the gaps in the competency mapping through the BioExcel Training Programme (initial recommendations).

The mapping and gap analysis phase of Task 4.1 is still ongoing; over the coming maps we anticipate extending the work along the following points.

- The mapping and gap analysis needs to reflect the partial mapping of resources to competencies

- Extend the mapping and gap analysis to reflect whether training resources are online or face-to-face
- Suitability of the training resources for the BioExcel User Groups (entry-level users, expert users, and systems administrators).
- After the initial mapping we can complete a targeted search for training resources for the competencies that are currently considered to be insufficiently covered.
- Incorporate feedback from the Competency Survey
- There are a number of training resources that map to a high number of BioExcel competencies (≥4). The aim would be to have a detailed look at these resources and specifically recommend them to BioExcel users if appropriate and/or explore the possibility of collaborating with the provider (if an external provider). A collaboration could involve making a BioExcel tailored version of the course, or providing a similar course in a different format/setting.

One practical problem we foresee is that there may be face-to-face courses that are very relevant, based on the BioExcel competencies, but are one-off external events. The implications of this issue will be discussed in D4.3 but some of the ways BioExcel can ameliorate this is by taking the following actions:

1. Advertise one-off external courses that map to BioExcel competencies through aggregators such as on-course and through fast-paced channels such as Twitter
2. Keep a record of relevant external events on the BioExcel website to encourage relevant events to be re-run in the future (potentially by a different third-party).

3.3 Collaborations

There are currently a number of relevant projects that have strong interest in training:

RItrain (http://ritrain.eu/) - mission is to improve and professionalize the training of managerial and leadership staff in research infrastructures (RIs).
CORBEL (http://www.corbel-project.eu/) - is an initiative of eleven new biological and medical research infrastructures (BMS RIs), includes competency-based training Work Package aimed at technical operators of BMS RIs.
ELIXIR-UK (http://www.elixir-uk.org/) - the United Kingdom node of ELIXIR is dedicated to training and hosts the main ELIXIR Training Portal (TeSS) (https://tess.elixir-uk.org/).
ELIXIR (www.elixir-europe.org) - unites Europe's leading life science organisations in managing and safeguarding the increasing volume of data being generated by publicly funded research.
PRACE Advanced Training Centres (http://www.prace-ri.eu/) - The PRACE Advanced Training Centres (PATCs) provide top-class education and training opportunities for computational scientists in Europe and are the primary source for PRACE training portal materials. The BioExcel partners BSC and EPCC are PATCs and can act as liaisons to PRACE.

RItrain and CORBEL are using the same competency-based approach to develop their training programmes. Joint strategy meetings are taking place to ensure that where possible the same process is being followed. Each project will complete a gap-analysis to assess training needs; we will coordinate between the projects to ensure that training gaps are filled by the most appropriate project and to avoid duplicated effort.

This list is not complete and we will keep extending this list as we make contact with training coordinators in other projects.

3.4 Training Interest Group

To aid collaboration with other training professionals we are looking into the feasibility of starting a Training Interest Group within BioExcel; the Interest Group (IG) could include trainers of specific resources (e.g. HADDOCK, GROMACS, CPMD etc.), training coordinators but could also be a way to reach out to related external projects (e.g. other CoEs).

The BioExcel Training IG will have the following aims:

- Improve visibility of training initiatives and individual courses/resources. Members of the BioExcel Training IG will be able to act as liaisons to other training initiatives.
- Facilitate collaboration between projects to promote best practice and efficient use of resources.
- Improve communication between "computational trainers" and "life science trainers". The overlap between the domains is currently not very large.
- Promote the importance of high quality training

Next steps in establishing the Training IG:

- Write a factsheet for the Training IG
- Compile a list of people who we would like to invite to join the Training IG (we anticipate only a subset will actively join)
- Organise a webinar to launch the BioExcel Training Interest Group and present the BioExcel competency profile. The Training IG/Webinar would target individual people, though everyone with an interest in training is welcome to join.
- Determine the timeline for the launch of the Training IG, this could coincide with presenting the results of the competency survey as well.

4 Initial plans for the BioExcel Training Programme

This section will provide an overview of how the conclusion from the gap analysis have been implemented in the BioExcel Training Programme. This overview is preliminary and final recommendations will be presented in D4.3 (due in month 12). In addition to the outline of the training programme we also present what training has been delivered by the project partners to date.

4.1 Previous training and dissemination events

The User Training task (Task 4.4 PM6-PM36) commenced in PM6 but the BioExcel partners have been actively engaged in training since (and before) the start of the project.

4.1.1 Face-to-face training courses and dissemination events

BioExcel or the BioExcel partners have been involved in 26 events, and an additional 4 events are planned by the end of the calendar year. Events are categorised based on their primary purpose (Training (T)/Dissemination (D)/ Interest Group meeting (IG) or dual purpose). Including the planned events, we have 15 dissemination events, 10 training events, 1 Interest Group event and 3 dissemination/training events as well as an Interest Group/Training meeting (full list is provided in Appendix 4).

4.1.2 BioExcel Webinar Series

BioExcel CoE has launched a series of educational webinars for computational biomolecular research (collaboration between WP3 & WP4). The series covers broad topics related to: the latest developments in BioExcel's major software packages; their application to modelling and simulation; best practices for performance tuning and efficient usage on HPC and novel architectures; introductory tutorials for novel users and much more.

The webinars include an audience Q&A session during which attendees can ask questions and make suggestions. They are a great opportunity to interact with the main code developers. Videos of the webinars are made available through the BioExcel website and the BioExcel YouTube channel.
http://bioexcel.eu/category/webinar/
https://www.youtube.com/channel/UCd2hq8Q_ZyTU4YafEhweE5g

Since the beginning of the webinar series, five webinars have taken place (Table 3).

Date	Title	Speaker	# Attendees
29th April 2016	Webinar #1: Integrative modelling of biomolecular complexes with HADDOCK	Alexandre Bonvin	18
11th May	Webinar #2: Performance Tuning	Mark Abraham	83

2016	and Optimization of GROMACS		
25th May 2016	Webinar #3: Atomistic Molecular Dynamics Setup with MDWeb	Adam Hospital	46
10th June 2016	Webinar #4: Mutation free energy calculation with pmx	Bert de Groot	34
30th June 2016	Webinar #5: QM/MM approaches in CPMD	Emiliano Ippoliti	16

Table 3: BioExcel Webinar Series

4.1.3 Online training material

A number of online training resources are already available (see list below); these are mainly centred around BioExcel's three major software packages (HADDOCK, GROMACS & CPMD). The exception is the EMBL-EBI train online portal (http://www.ebi.ac.uk/training/online/) which has a much wider remit.

A full list of training resources is available in Appendix 3 (used for the mapping exercise)

- Homology modelling / MD with Gromacs / docking with HADDOCK

http://www.bonvinlab.org/education/molmod/
https://github.com/haddocking/molmod

- HADDOCK CASP-CAPRI T70 ab-initio docking tutorial

http://www.bonvinlab.org/education/HADDOCK-CASP-CAPRI-T70/
https://github.com/haddocking/CASP-CAPRI-T70-tutorial

- WeNMR lecture on information-driven docking with HADDOCK

https://www.youtube.com/watch?v=InwxoiVy3QQ

- Demonstration of the WeNMR HADDOCK webportal (easy interface)

https://www.youtube.com/watch?v=5eS8BAOJ6es

- GROMACS.org Tutorials

http://www.gromacs.org/Documentation/Tutorials

- CPMD tutorial

http://cpmd.org/cpmd-tutorial

4.2 BioExcel Training Programme

Based on the competency profile and the early stages of the mapping exercise we plan to initially target the BioExcel training at the three user groups identified prior to the workshop.

1. Entry Level Users: We anticipate that the training need for this group will be primarily focussed on bridging the gap between life-science and HPC/HTC computing (users transitioning from a bench-based molecular life science background).
2. Expert Users: the training needs for this group will likely be quite specific for each use case (e.g. specific software packages, optimisation). As an additional source of information we anticipate requesting input from the Interest Groups (IG) as to what the highest priority training needs are (through IG leads & WP3).
3. Systems Administrators: this group has a double aim, to enable the previous groups to be able to communicate better with the systems administrators as well as to provide training for this group directly. We anticipate that trying to engage this group in training could be challenging, webinars would be likely be the best option. Training of interest could be software installation webinars.

The BioExcel Training Programme will combine the following elements:

- Face-to-face training
- eLearning
- Knowledge resources repository for training material and external courses

Note that all suggested training listed below are initial thoughts and subject to resources and full discussions within the consortium. Prioritisation will depend on the training need as well as availability of experts. D4.3 will provide detailed

4.2.1 Initial plans for face-to-face training

Year	Date	Title
		Competencies
Year 1	3-4th May 2016 (EMBL-EBI)	BioExcel: addressing training needs for advanced simulations in biomolecular research
Year 1	20-21st October 2016 (BSC)	BioExcel: workflow training for computational biomolecular research
		4.1
Year 2	10-13 April 2017 (KTH)	BioExcel – PRACE Spring School; Focus is on GROMACS and related codes, analysis and visualization tools
		2.3, 2.6 (TBC)
Year 2	TBC	Advanced face-to-face training on the three BioExcel codes HADDOCK, GROMACS, CPMD
		2.5; 2.6 (TBC)
Year 2	TBC	Computational skills for life scientists (partially

		project based)
		Generic Computing Competencies (3.1 – 3.8)
Year 2 or 3	TBC	Introduction to HPC computing for life scientists
		4.2 – 4.5 (building on from "Computational skills for life scientists")
Year 3	TBC	Advanced face-to-face training on the three BioExcel codes HADDOCK, GROMACS, CPMD
		2.5; 2.6 (TBC)

Table 4: Initial plans for face-to-face training mapped to the relevant BioExcel competencies. Mapping is approximate since it will be dependent on exact content.

4.2.2 Initial plans for online Training

1. Webinars: below is a list of training webinar topics that are under discussion and the competencies they map to.

- An entry level webinar that introduces users to concepts underlying biomolecular simulations (e.g. docking, molecular dynamics, free energy calculations) – 2.1; 2.2
- Where to get compute resources and the benefits and challenges of different types – 3.2; 4.4
- Mathematical basis of different modelling approaches – 2.2
- Short seminars with questions and documentation on how to install specific software packages/codes (own environment & cloud deployment) – 3.6; 3.7; 4.4

We anticipate that the first training webinar could take place early in year 2 (before the end of December 2016).

2. Online training courses: we anticipate creating online courses using the EMBL-EBI train online platform (and potentially other platforms).

- An Entry level course, potentially based on modules PIs currently teach – 2.2

3. Blended-learning courses: this type of course combines face-to-face learning with online learning. Participants would come to an initial workshop (maybe 2-3 days), and stay in contact with their trainers for the next 6 months through an online learning environment (with discussion boards, online courses and tasks), after that they come back for a final workshop.

We think this type of learning could potentially work really well for BioExcel users because it lends itself well to topics that are very complicated, requiring assimilation of new concepts over time. This approach also works well for workflow-based approaches, so could dovetail very well with BioExcel's use cases. This format would allow us to give users the initial knowledge face-to-face and assist them online while they try to turn this knowledge into a skill with their own data and in their own environment. The "downside" to this type of learning is that it is time intensive to develop and requires more resources to

maintain as well as a dedicated training platform. Discussions are currently underway to decide whether or not sufficient resources are available for this approach and which potential courses/users would most benefit from this approach.

4.2.3 Impact assessment

An important aspect of the BioExcel training Programme will be to analyse the impact of the training we deliver. The following KPIs related to training were listed in D6.2:

- Number of organised training events
- Number of people trained (academic and industrial)

The KPIs listed in D6.2 are currently under discussion following the project review [PM8]. An updated list of KPIs related to training will be published in D4.3. As far as possible, we intend to align these with ELIXIR's training KPIs, currently under development as part of the ELIXIR-Excelerate project.

It is routine practice for course providers to gather feedback through a survey after each training course (Appendix 5); we intend to do this for BioExcel as it will allow us assess whether the course should run again and, if so, make improvements to future iterations of each course. Measuring the impact of training on a person's ability to do his or her job is significantly harder. We anticipate sending out post-6-month surveys (subject to participant consent, see section "About you" in Appendix 5) to try and capture the long-term impact of our training programme. This approach has been tested and found useful by EMBL-EBI and will be implemented within ELIXIR as part of ELIXIR-Excelerate's training work package.

5 Next steps

Throughout this document we have tried to list what the next steps are for each section. Below is the full list annotated with expected completion times.

5.1 Competency profile

- [PM10] Compile contact list for the competency survey (reused for the invitation to Training IG).
- [PM10] Send out the Competency Survey.
- [PM12] Add missing Knowledge, Skills and Behaviours.
- [TBC] In collaboration with WP3 extend the competency profiling to the "stereotypes" that are to be defined.

5.2 Mapping and Gap Analysis

- [PM12] The mapping and gap analysis needs to reflect the partial mapping of resources to competencies
- [PM12] Extend the mapping and gap analysis to reflect whether training resources are online or face-to-face
- [PM12] Suitability of the training resources for the BioExcel User Groups (entry-level users, expert users, and systems administrators).
- [PM12] After the initial mapping we can complete a targeted search for training resources for the competencies that are currently considered to be insufficiently covered.
- [PM12] Incorporate feedback from the Competency Survey

5.3 Training IG

- [PM10] Write a factsheet for the Training IG with justification as to the aim and target audience.
- [PM10] Compile a list of people who we would like to invite to join the Training IG (we anticipate only a subset will actively join).
- [PM10] Determine the timeline for the launch of the Training IG, this could coincide with presenting the results of the competency survey as well.
- [PM13] Organise a webinar to launch the BioExcel Training Programme and present the BioExcel competency profile. The Training IG/Webinar would target individual people, though everyone with an interest in training is welcome to join.

5.4 BioExcel Training Programme

- [PM12] Draft recommendations for the BioExcel training programme.
- [PM12 – all-hands meeting] Discussion with partners to determine the timing of the proposed elements of the Training Programme.
- [PM15] There are a number of training resources that map to a high number of BioExcel competencies (≥4). The aim would be to have a detailed look at these resources and specifically recommend them to BioExcel users if appropriate and/or explore the possibility of collaborating with the provider (if an external provider). A collaboration

could involve making a BioExcel tailored version of the course, or providing a similar course in a different format/setting.

5.5 Upcoming Milestones and Deliverables

- [PM12] D4.3: Recommendations for BioExcel training programme. This document will report on the outcome from the evaluation of the training needs and the plans for further training activities.

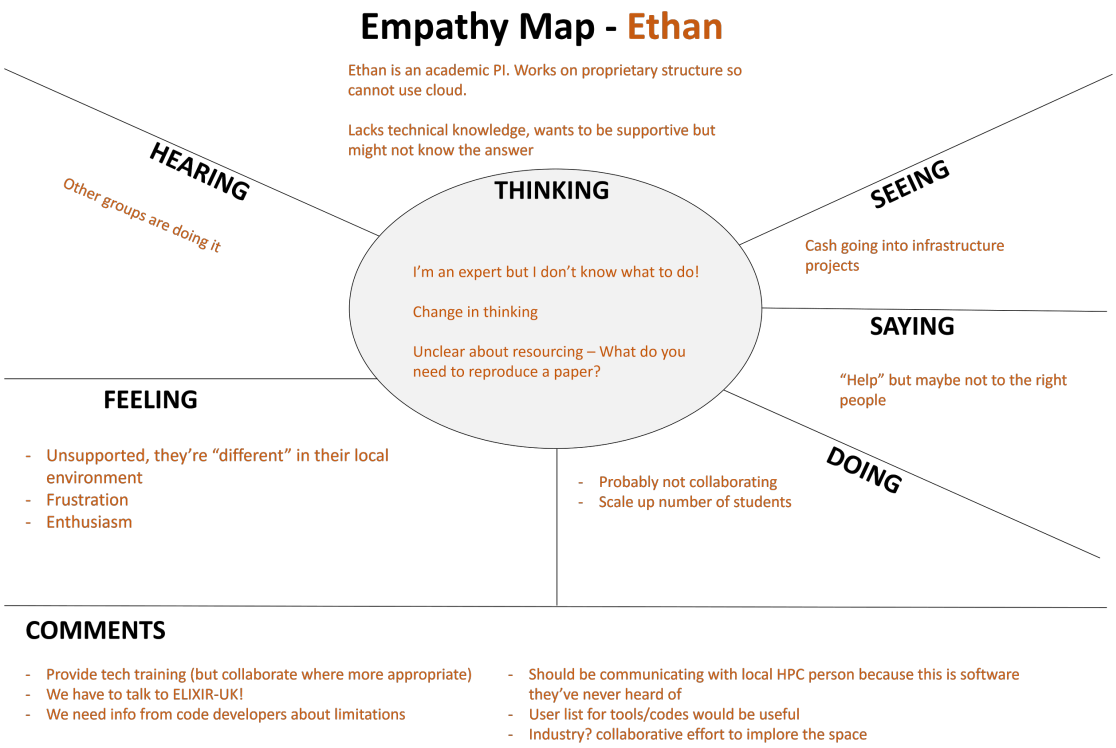
6 Appendix

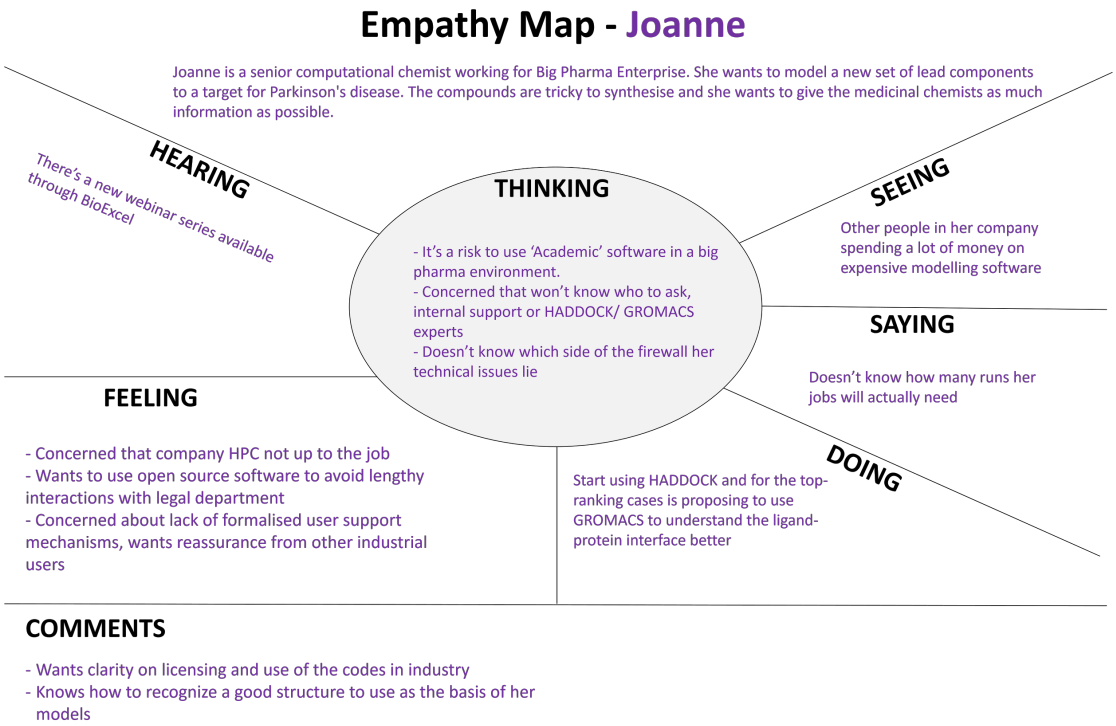
Appendix 1: List of Workshop participants for BioExcel: addressing training needs for advanced simulations in biomolecular research, 3-4th May 2016, EMBL-EBI, Hinxton, UK

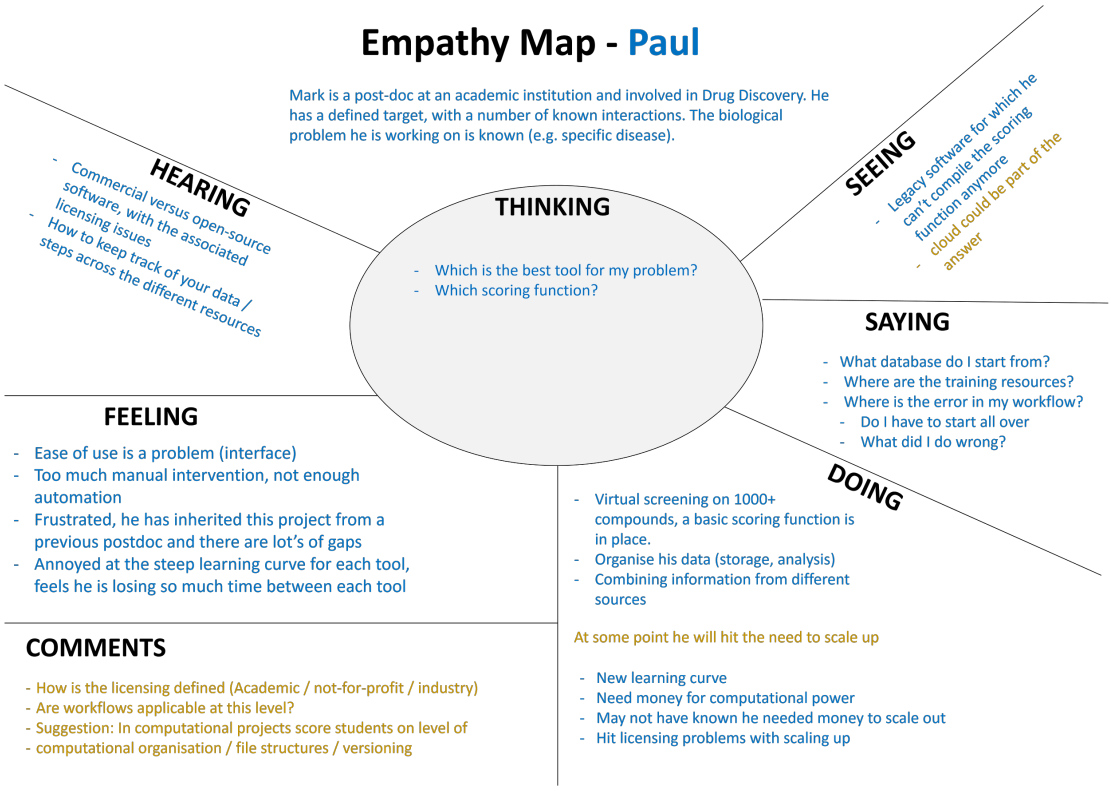
First Name	Last Name	Position	Affiliation
Mark	Abraham	Staff Scientist	KTH
Alexandre	Bonvin	Principal Investigator (Group Leader)	Utrecht University
Cath	Brooksbank	Principal Investigator (Group Leader)	EMBL-EBI
Adam	Carter	Other	EPCC, The University of Edinburgh
Mark	Forster	Industry	Syngenta
Emiliano	Ippoliti	Experienced Researcher (Post Doc)	IAS-5 / INM-9 Computational Biomedicine
Ramyarooban	Kanapathywasam	Industry	Unilever
Predrag	Kukic	Experienced Researcher (Post Doc)	University of Cambridge
Lee	Larcombe	Staff Scientist	ELIXIR UK University of Edinburgh
Caroline	Low	Staff Scientist	Imperial College London
Lok Hin	Lui	Early Stage Researcher (PhD)	University College London
Vera	Matser	Staff Scientist	EMBL-EBI
Steven	Newhouse	Principal Investigator (Group Leader)	EMBL-EBI
Ilias	Soumpasis	Industry	Unilever
Andrea	Spitaleri	Experienced Researcher (Post Doc)	Istituto Italiano di Tecnologia
Marta	Strumillo	Early Stage Researcher (PhD)	EMBL-EBI
Mikael	Trellet	Experienced Researcher (Post Doc)	Utrecht University
Sameer	Velankar	Principal Investigator	EMBL-EBI

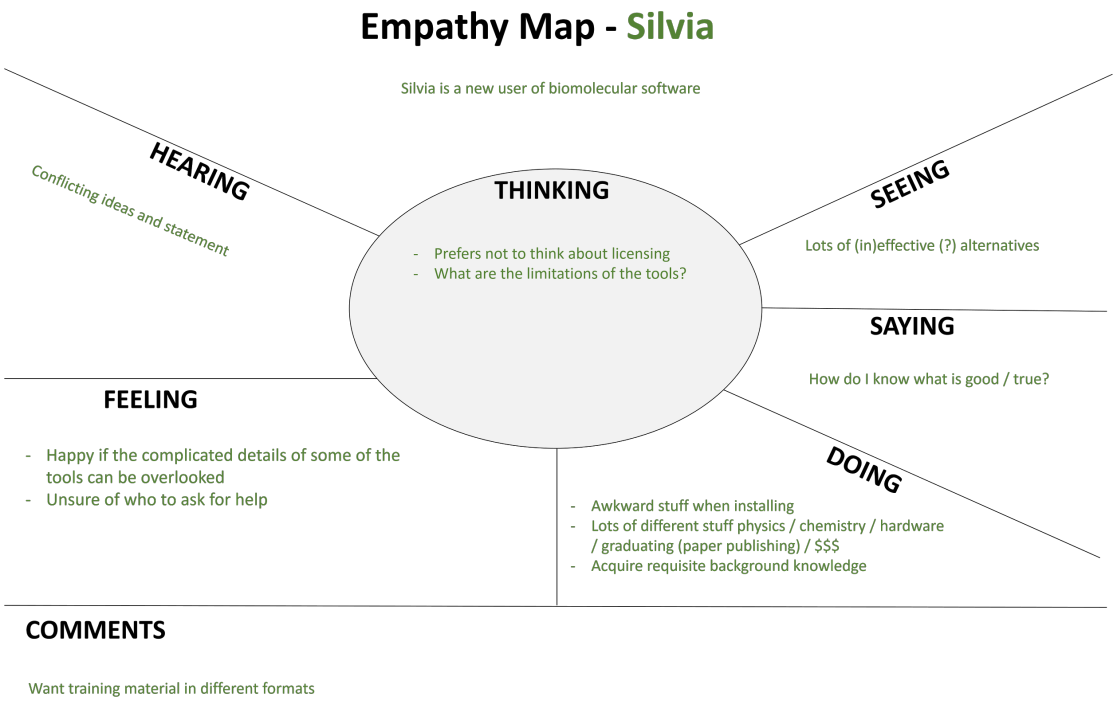
		(Group Leader)	
Gitanjali	Yadav	Principal Investigator (Group Leader)	National Inst of Plant Genome Research

Appendix 2: Empathy Maps created during the BioExcel: addressing training needs for advanced simulations in biomolecular research course, 3-4th May 2016, EMBL-EBI, Hinxton, UK.






Empathy Map - Paul
Mark is a post-doc at an academic institution and involved in Drug Discovery. He has a defined target, with a number of known interactions. The biological problem he is working on is known (e.g. specific disease).
HEARING
Commercial versus open-source software, with the associated licensing issues
How to keep track of your data / steps across the different resources
THINKING
Which is the best tool for my problem?
Which scoring function?
SEEING
Legacy software for which he can't compile the scoring function anymore
cloud could be part of the answer
SAYING
What database do I start from?
Where are the training resources?
Where is the error in my workflow?
Do I have to start all over
What did I do wrong?
FEELING
Ease of use is a problem (interface)
Too much manual intervention, not enough automation
Frustrated, he has inherited this project from a previous postdoc and there are lot's of gaps
Annoyed at the steep learning curve for each tool, feels he is losing so much time between each tool
DOING
Virtual screening on 1000+ compounds, a basic scoring function is in place.
Organise his data (storage, analysis)
Combining information from different sources
At some point he will hit the need to scale up
New learning curve
Need money for computational power
May not have known he needed money to scale out
Hit licensing problems with scaling up
COMMENTS
How is the licensing defined (Academic / not-for-profit / industry)
Are workflows applicable at this level?
Suggestion: In computational projects score students on level of
computational organisation / file structures / versioning


Empathy Map - Silvia
Silvia is a new user of biomolecular software
HEARING
Conflicting ideas and statement
THINKING
Prefers not to think about licensing
What are the limitations of the tools?
SEEING
Lots of (in)effective (?) alternatives
SAYING
How do I know what is good / true?
FEELING
Happy if the complicated details of some of the tools can be overlooked
Unsure of who to ask for help
DOING
Awkward stuff when installing
Lots of different stuff physics / chemistry / hardware / graduating (paper publishing) / $$$
Acquire requisite background knowledge
COMMENTS
Want training material in different formats

Appendix 3: Existing training resources (date: 12/07/2016) used for the mapping of training resources to BioExcel Competencies.

N°	Competency mapping	Training Resource Name	Type	URL	BioExcel Partner
1	2.5; 2.6	Homology modelling / MD with Gromacs / docking with HADDOCK	Tutorial	http://www.bonvinlab.org/education/molmod/ https://github.com/haddocking/molmod	UU
2	2.5; 2.6	HADDOCK CASP-CAPRI T70 ab-initio docking tutorial	Tutorial	http://www.bonvinlab.org/education/HADDOCK-CASP-CAPRI-T70/ https://github.com/haddocking/CASP-CAPRI-T70-tutorial	UU
3	2.5; 2.6	WeNMR lecture on information-driven docking with HADDOCK	YouTube	https://www.youtube.com/watch?v=InwxoiVy3QQ	UU
4	2.5; 2.6	Demonstration of the WeNMR HADDOCK webportal (easy interface)	YouTube	https://www.youtube.com/watch?v=5eS8BAOJ6es	UU
5	2.5; 2.6	CPMD tutorial	Tutorial	http://cpmd.org/cpmd-tutorial	Juelich
6	4.2; 4.3; 4.4; 4.5	ARCHER intro / driving test		https://www.archer.ac.uk/training/course-material/online/index.php	UEDIN
7	2.7; 4.3; 4.4; 4.5	Practical Introduction to High Performance Computing (PItHPC)	Online	http://www.drps.ed.ac.uk/15-16/dpt/cxpgph11093.htm	UEDIN
8	2.5; 2.6	GROMACS.org	Tutorial	http://www.gromacs.org/Documentation/Tutorials	KTH
9	n/a	BioExcel Webinar Series	Webinar	http://bioexcel.eu/category/webinar/	KTH
10	3.2; 4.5	Best bang for your buck: GPU nodes for GROMACS biomolecular simulations	Journal Paper	https://doi.org/10.1002/jcc.24030	KTH
11	n/a	PRACE Training Portal (Partnership for Advanced Computing in Europe)	YouTube	https://www.youtube.com/user/PRACECourses/featured	External

12	n/a	Train online - repository of free online training courses on EMBL-EBI and collborating data resources	Training portal	http://www.ebi.ac.uk/training/online/	EMBL-EBI
13	n/a	OnCourse	Training portal	https://www.on-course.eu/	EMBL-EBI
14	n/a	TeSS - ELIXIR's Training Portal	Training portal	http://elixir-uk.org/training-platform	External (EMBL-EBI)
15	n/a	Goblet - Training portal	Training portal	http://mygoblet.org/training-portal	External (EMBL-EBI)
16	n/a	Find CPD	Training portal	http://www.findcpd.com/	External
17	n/a	Coursera	Training portal	https://www.coursera.org/	External
18	2.10	Data Carpentry - introductory computational skills needed for data management and analysis in all domains of research	Data Carpentry	http://www.datacarpentry.org/	External, UNIMAN
19	2.5; 2.6	MDWeb - Molecular Dynamics on Web Tutorials (Protein Atomistic/Coarse-Grained Molecular Dynamics Setup, Analysis, Run)	Online Tutorials / PDF	http://mmb.irbbarcelona.org/MDWeb/htmlib/help/MDWeb_Tutorial.pdf http://mmb.irbbarcelona.org/MDWeb/help.php?id=tutorial http://mmb.irbbarcelona.org/MDWeb/help.php?id=tutorialAnalysis http://mmb.irbbarcelona.org/MDWeb/help.php?id=tutorialRun	IRB
20	2.5; 2.6	NAFlex - Nucleic Acids Flexibility Server Tutorials (Nucleic Acids Atomistic/Coarse-grained Molecular Dynamics Setup, Analysis)	Online Tutorials / PDF	http://mmb.irbbarcelona.org/NAFlex/htmlib/help/NAFlex_Tutorial.pdf http://mmb.irbbarcelona.org/NAFlex/help.php?id=tutorials	IRB
21	2.5; 2.6	FlexServ - Protein Flexibility Server	Online Tutorials / PDF	http://mmb.irbbarcelona.org/FlexServ/tutorial.php	IRB
22	4.3; 4.4; 4.5	COMPSs tutorials and manuals	Training material	http://www.bsc.es/computer-sciences/grid-computing/comp-superscalar/downloads-and-documentation	BSC

23	4.3; 4.4; 4.5	COMP Superscalar	Video	https://www.youtube.com/channel/UCxx5hojMjCNqKRJclMaJbHQ	BSC
24	4.1	Apache Taverna documentation	Documentation, tutorials, videos	https://taverna.incubator.apache.org/documentation/	External, UNIMAN
25	4.1	Taverna developer guide	Tutorial	http://dev.mygrid.org.uk/wiki/display/developer/Developers+Guide	UNIMAN
26	4.1	Taverna tutorials	Training material	https://github.com/taverna/taverna-tutorials	UNIMAN
27	4.1	Common Workflow Language user guide	Tutorial	http://www.commonwl.org/draft-3/UserGuide.html	External, UNIMAN
28	2.9	Using the Open PHACTS API	Webinar	https://www.youtube.com/watch?v=8XV0kC7PuSU	UNIMAN
29	2.9	Open PHACTS YouTube channel	Webinars	https://www.youtube.com/playlist?list=PLQ_ieWFru1xHMUoWtHypLvbVjZLoQMev3	External, UNIMAN
30	2.9	Open PHACTS support portal	Support portal	http://support.openphacts.org/support/home	External, UNIMAN
31	2.9	Open PHACTS API docs	Documentation	https://dev.openphacts.org/docs	External, UNIMAN
32	2.9	Open PHACTS developer portal	Documentation	https://dev.openphacts.org/resources	External, UNIMAN
33	2.9	Open PHACTS ops.js API docs	Documentation	http://openphacts.github.io/ops.js/Openphacts.html	UNIMAN
34	4.1	Open PHACTS example workflows	Examples	http://www.myexperiment.org/groups/1125	External, UNIMAN
35	4.1	Taverna example workflows	Examples	http://www.myexperiment.org/packs/609	UNIMAN
36	2.5; 2.6	Presentations from GROMACS developers/users workshop	Presentation recordings	http://www.mpibpc.mpg.de/grubmueller/gromacs2016	KTH, MPG, External
37	4.5	PATC PERFORMANCE ANALYSIS WORKSHOP		http://www.vi-hps.org/training	External
38	2.7; 4.3; 4.4;	Hands-on Introduction to HPC @ EPCC		https://events.prace-ri.eu/event/479/	External

	4.5				
39	4.3	MPI - PRACE Training Portal		http://www.training.prace-ri.eu/training_material/?tx_pracetmo_pi1[category]=15 http://www.training.prace-ri.eu/training_material/?tx_pracetmo_pi1[category]=17	External
40	4.3	Debugging, Profiling and Optimization Tools (10) - PRACE Training Portal		http://www.training.prace-ri.eu/training_material/?tx_pracetmo_pi1[category]=33	External
41	4.5	Performance Analysis - PRACE Training Portal		http://www.training.prace-ri.eu/training_material/?tx_pracetmo_pi1[category]=1	External
42	4.5	Performance Analysis Workshop @ EPCC at University of Cambridge		https://events.prace-ri.eu/event/490/	External
43	4.3	PATC@BSC - Parallel Programming Workshop (Oct 2016)	Face-to-face	http://www.bsc.es/marenostrum-support-services/hpc-trainings/prace-trainings	BSC
44	4.4;4.5	PATC@BSC - Administration of Petaflop Machine (Jan/Feb 2017)	Face-to-face	http://www.bsc.es/marenostrum-support-services/hpc-trainings/prace-trainings	BSC
45	4.3	PATC@BSC - Programming of Petaflop Machine (April 2017)	Face-to-face	http://www.bsc.es/marenostrum-support-services/hpc-trainings/prace-trainings	BSC
46	4.3;4.4; 4.5	PATC@BSC - HPC core training (May 2017)	Face-to-face	http://www.bsc.es/marenostrum-support-services/hpc-trainings/prace-trainings	BSC
47	4.4; 4.3;	PATC@BSC - Programming Distributed Computing Platforms with COMPSs (Feb 2017)	Face-to-face	http://www.bsc.es/marenostrum-support-services/hpc-trainings/prace-trainings http://www.bsc.es/marenostrum-support-services/hpc-education-and-training/patc-training/2015-23-27-nov-patc-parallel	BSC
48	4.4;	PATC@BSC - HPC simulations for Science and Engineering	Face-to-face	http://www.bsc.es/marenostrum-support-services/hpc-trainings/prace-trainings	BSC
49	4.1; 4.4	PATC@BSC - Introduction to simulation	Face-to-face	http://www.bsc.es/marenostrum-support-	BSC

		environments for Life sciences (March 2017)		services/hpc-trainings/prace-trainings http://www.bsc.es/lifescience-patc	
50	4.3; 4.4	PATC@BSC - Introduction to programming in CUDA (June/July 2016, 2017)	Face-to-face	http://www.bsc.es/marenostrum-support-services/hpc-trainings/prace-trainings	BSC
51	4.5; 4.4; 4.3	PATC@BSC - PUMPS (July 2017)	Face-to-face	http://www.bsc.es/marenostrum-support-services/hpc-trainings/prace-trainings http://www.bsc.es/marenostrum-support-services/hpc-education-and-training/pumps-summer-school	BSC
52	2.3; 2.8; 2.9; 2.10; 2.11	PATC@BSC - Big Data Analytics (Feb 2017)	Face-to-face	http://www.bsc.es/marenostrum-support-services/hpc-trainings/prace-trainings http://www.bsc.es/big-data-analytics	BSC
53	1.3; 2.13	Media and Communications Training: Masterclass for Senior Executives	online	https://www.on-course.eu/courses/media-and-communications-training-masterclass-for-/	External
54	2.12	Effective Reading and Speed Reading	online	http://www.government-knowledge.com/seminar/1876/	External
55	2.13	Writing in the Sciences	online	https://www.coursera.org/course/sciwrite	External
56	2.13	How to Write and Publish a Scientific Paper (Project-Centered Course)	online	https://www.coursera.org/learn/how-to-write-a-scientific-paper	External
57	2.1; 2.7; 2.8; 3.4; 4.4; 4.4; 4.5	Clouds, Distributed Systems, Networking	online	https://www.coursera.org/specializations/cloud-computing	External
58	2.8	Machine Learning Specialization (6 courses)	online	https://www.coursera.org/specializations/machine-learning	External
59	2.3; 2.8	Data Mining Specialization (6 courses)	online	https://www.coursera.org/specializations/data-mining	External
60	2.3; 2.8; 3.5	Data Science Specialization	online	https://www.coursera.org/specializations/jhu-data-science	External
61	2.3; 2.7; 2.8; 2.10; 2.13	Data Science at Scale Specialization	online	https://www.coursera.org/specializations/data-science	External

62	2.3; 2.8; 2.10	Big Data Specialization	online	https://www.coursera.org/specializations/bigdata	External
63	2.1; 3.2; 3.4	Data Structures and Algorithms Specialization	online	https://www.coursera.org/specializations/data-structures-algorithms	External
64	1.5	User Experience: Research & Prototyping	online	https://www.coursera.org/learn/user-research	External
65	2.1; 2.2	Biology Meets Programming: Bioinformatics for Beginners	online	https://www.coursera.org/learn/bioinformatics	External
66	2.1	Fundamentals of Computing Specialization	online	https://www.coursera.org/specializations/computer-fundamentals	External
67	2.1; 2.2	Bioinformatics Specialization	online	https://www.coursera.org/specializations/computational-biology	External
68	3.1	Agile Development Specialization	online	https://www.coursera.org/specializations/agile-development	External
69	2.1; 2.2; 3.1; 4.1	Genomic Data Science Specialization	online	https://www.coursera.org/specializations/genomic-data-science	External
70	2.1; 3.3	The Hardware/Software Interface	online	https://www.coursera.org/course/hwswinterface	External
71	2.1; 2.5; 2.6	Build a Modern Computer from First Principles: From Nand to Tetris (Project-Centered Course)	online	https://www.coursera.org/learn/build-a-computer	External
72	3.1; 3.4; 4.3	Parallel programming	online	https://www.coursera.org/learn/parprog1	External
73	3.4; 3.5	Python for Genomic Data Science		https://www.coursera.org/learn/python-genomics	External
74	2.2; 3.3	Command Line Tools for Genomic Data Science		https://www.coursera.org/learn/genomic-tools	External
75	2.2	Molecular Machines: Integrative Structural and Molecular Biology	Conference	http://www.embl.de/training/events/2016/HYB16-01/	EMBL
76	2.2	ChEMBL: Exploring bioactive drug-like molecules	online course	https://www.on-course.eu/courses/chembl-exploring-bioactive-drug-like-molecules/	EMBL-EBI

77	2.2	PDBeChem: Searching for small molecules and small molecule fragments	online course	https://www.on-course.eu/courses/pdbechem-searching-for-small-molecules-and-small-m/	EMBL-EBI
78	2.2	Structural Bioinformatics	Face-to-Face	https://www.on-course.eu/courses/embo-practical-course-on-computational-structural-/	External
79	2.2	In Silico Systems Biology	Face-to-Face	https://www.on-course.eu/courses/biopredyn-the-systems-biology-modelling-cycle-buil/	External
80	2.2	Cheminformatics	online course	https://www.on-course.eu/courses/cheminformatics/	External
81	2.2	Structural Bioinformatics	online course	https://www.on-course.eu/courses/structural-bioinformatics/	External
82	2.2	Drug Design	online course	https://www.on-course.eu/courses/drug-design/	External
83	2.2	Practical Drug Discovery: Case Studies	online course	https://www.on-course.eu/courses/practical-drug-discovery-case-studies/	External
84	2.2	Medicinal Chemistry	online course	https://www.on-course.eu/courses/medicinal-chemistry-n2/	External
85	2.1; 2.3	Python for Everybody Specialization (4 courses)	online course	https://www.coursera.org/specializations/python	External
86		FitSM Advanced Training in Service Planning and Delivery		https://www.on-course.eu/courses/fitsm-advanced-training-in-service-planning-and-de/	External

87		FitSM Foundation Training		https://www.on-course.eu/courses/fitsm-foundation-training/	External
88		FitSM Advanced Training in Service Operation and Control		https://www.on-course.eu/courses/fitsm-advanced-training-in-service-operation-and-c/	External
89	2.3; 3.5	R Programming	online course	https://www.coursera.org/learn/r-programming	External
90	3.5	Python for Genomic Data Science		https://www.on-course.eu/courses/python-for-genomic-data-science/	External
91	2.5; 3.1; 3.4; 3.5	Computing Skills for Drug Discovery		https://www.on-course.eu/courses/computing-skills-for-drug-discovery/	External
92	2.2; 2.9; 2.12	Introduction to Modelling Biological Systems		https://www.on-course.eu/courses/introduction-to-modelling-biological-systems/	External
93	2.2	Functional Genomics and Systems Biology		https://www.on-course.eu/courses/functional-genomics-and-systems-biology/	External
94	2.2	Structure Determination of Drug Targets		https://www.on-course.eu/courses/structure-determination-of-drug-targets/	External
95	1.3 (Instructor Training); 1.5 (Instructor Training); 2.10; 3.1;	Software Carpentry Workshop	face-to-face	http://software-carpentry.org/	External

	3.3; 3.4; 3.5; 3.6;				
96	2.1; 2.10	Databases & Data Modelling		https://www.on-course.eu/courses/databases-data-modelling/	External
97	2.2	Biointeractomics: from Bimolecular Interactions to Networks		https://www.on-course.eu/courses/biointeractomics-from-bimolecular-interactions-to-/	External
98	2.2; 2.3; 2.10	Principles of Digital Biology		https://www.on-course.eu/courses/principles-of-digital-biology/	External
99	2.2; 2.9	PDBe: Searching the Protein Data Bank		https://www.on-course.eu/courses/pdbe-searching-the-protein-data-bank/	External
100	2.2; 2.9	ChEBI: the online chemical dictionary for small molecules		https://www.on-course.eu/courses/chebi-the-online-chemical-dictionary-for-small-mol/	External
101	1.1; 1.4; 2.12	Professional Skills in Drug Discovery		https://www.on-course.eu/courses/professional-skills-in-drug-discovery/	External
102	3.1	System Design & Analysis Using UML		https://www.on-course.eu/courses/system-design-analysis-using-uml/	External
103	2.2	Measuring Drug Binding		https://www.on-course.eu/courses/measuring-drug-binding/	External
104	2.3; 2.10	Introduction to Integrative Omics		https://www.on-course.eu/courses/introduction-to-integrative-omics/	External

105	2.2; 2.3	Quantitative Methods for Scientific Research I		https://www.on-course.eu/courses/quantitative-methods-for-scientific-research-i/	External
106	2.2; 2.3	Quantitative Methods for Scientific Research II		https://www.on-course.eu/courses/quantitative-methods-for-scientific-research-ii/	External
107	2.2; 2.9	UniProt: Exploring protein sequence and functional information		https://www.on-course.eu/courses/uniprot-exploring-protein-sequence-and-functional-/	External
108	2.3; 2.12	Analytical Methods		https://www.on-course.eu/courses/healthcare-science-2-analytical-methods/	External
109	2.2; 2.3	Functional Genomics and Research Skills Analytical Techniques		https://www.on-course.eu/courses/functional-genomics-and-research-skills-analytical/	External
110	2.2; 2.9	PDBe: Exploring a Protein Data Bank (PDB) entry		https://www.on-course.eu/courses/pdbe-exploring-a-protein-data-bank-pdb-entry/	External
111	2.2; 2.9	Bioinformatics for Systems Biology		https://www.on-course.eu/courses/bioinformatics-for-systems-biology/	External
112	2.2	Principles of Protein Structure		https://www.on-course.eu/courses/principles-of-protein-structure/	External
113	2.2	PDBePISA: Identifying and interpreting the likely biological assemblies of a protein structure		https://www.on-course.eu/courses/pdbepisa-identifying-and-interpreting-the-likely-b/	External
114	2.2	PDBeFold: Searching for structural homologues of a protein		https://www.on-course.eu/courses/pdbefold-searching-for-	External

				structural-homologues-of-a-/	
115	2.2	Druggable Systems		https://www.on-course.eu/courses/druggable-systems/	External
116	2.2	Techniques in Structural Molecular Biology		https://www.on-course.eu/courses/techniques-in-structural-molecular-biology/	External
117	2.3; 2.10; 3.1; 3.5	Data Carpentry Workshop	face-to-face	http://www.datacarpentry.org/	External
118	2.3; 2.8; 4.1	KNIME User Training		http://www.knime.org/training	External
119	2.8	KNIME Text Mining Training		http://www.knime.org/training	External
120	2.8; 3.4	KNIME Developer Training		http://www.knime.org/training	External
121	3.8; 4.1	KNIME Server Training		http://www.knime.org/training	External
122	2.1; 2.2	Mathematics for Metabolic Modelling		https://www.on-course.eu/courses/mathematics-for-metabolic-modelling/	External
123	2.8	Data Mining Techniques		https://www.on-course.eu/courses/data-mining-techniques/	External
124	1.3	Effective Virtual Working		https://www.on-course.eu/courses/effective-virtual-working/	External
125	1.2	Biomedical data: Ethical, legal and social implications	online course	http://www.ebi.ac.uk/training/online/course/biomedical-data-ethical-legal-and-social-implicati	EMBL-EBI
126	1.5	User experience design	online course	http://www.ebi.ac.uk/training/online/course/user-experience-design	EMBL-EBI
127	2.12	Europe PMC: Quick tour	online course	http://www.ebi.ac.uk/training/online/course/europe-pmc-quick-tour-0	EMBL-EBI

128	2.8	Text mining: Key concepts and applications	online course	http://www.ebi.ac.uk/training/online/course/text-mining-key-concepts-and-applications	EMBL-EBI

Appendix 4: List of training and dissemination events by BioExcel and BioExcel partners (greyed out are future events).

Date	Event name	Location	Website	Primary activity dissemination (D) / training (T) / Interest Group meeting (IG)
09 Nov 2015	11th Concertation Meeting for H2020 e-Infra Projects	Brussels, Belgium	https://ec.europa.eu/digital-single-market/news/11th-e-concertation-meeting-european-e-infrastructure-projects	D
04 Feb 2016	PATC Course: Programming Distributed Computing Platforms with COMPSs	Barcelona, Spain	https://www.bsc.es/marenostrum-support-services/hpc-education-and-training/patc-training/2015-23-27-nov-patc-parallel	T
18-19 Feb 2016	Open PHACTS 2-day conference on Linking Life Science Data: Design to Implementation, and Beyond	Vienna, Austria	http://www.openphactsfoundation.org/event-announcement-linking-life-science-data-design-to-implementation-and-beyond/#more-3135	D
22 Feb 2016	Symposium on computation applied to Life Sciences (BSC/IRB)	Barcelona, Spain		D
29 Feb 2016	Workshop: In silico tools in drug design and target discovery	Barcelona, Spain	http://www.bsc.es/resdrugdes2016/	T
7-8 Mar 2016	EMBL-EBI SME Forum	Hinxton, UK	http://www.ebi.ac.uk/industry/sme-forum	D
14-15 Mar 2016	PRACE Course. Simulation Environments in Life Sciences	Barcelona, Spain	http://www.bsc.es/marenostrum-support-services/hpc-education-and-training/patc-training/2016-14-	T

			15-mar-simulation	
17-18 Mar 2016	VIB Conference: Applied Bioinformatics in Life Sciences	Leuven, Belgium	http://www.vibconferences.be/event/applied-bioinformatics-in-life-sciences	D
5-22 Apr 2016	HPC-LEAP School on numerical analysis and algorithms at the exascale: Classical N-body methods for complex systems on massively parallel architectures: CoS-1	RWTH Aachen, Germany & FZJ Juelich, Germany	https://hpc-leap.eu/index.php/events/upcoming-events.html?id=25	T
11-15 Apr 2016	INSTRUCT practical course on "Advanced methods for the integration of diverse structural data with NMR data"	Utrecht, Netherlands	https://www.structuralbiology.eu/support/whats-on/calendar-events/instruct-practical-course-advanced-methods-integration-diverse	D/T
26-29 Apr 2016	EASC2016	Stockholm, Sweden	https://www.pdc.kth.se/easc2016	D
03 May 2016	EMBL-EBI workshop on addressing training needs	Hinxton, UK	http://www.ebi.ac.uk/training/events/2016/bioexcel-addressing-training-needs-advanced-simulations-biomolecular-research	T
9-13 May 2016	HPC Summit, organized by EXDCI	Prague, Czech Republic	https://exdci.eu/events/hpc-summit-week-exdci-workshop	D
10-12 May 2016 => HPC Summit at the same time	PRACEdays16	Prague, Czech Republic	http://www.prace-ri.eu/pracedays_16	D
12-May 2016	ETP4HPC workshop	Prague, Czech Republic	http://www.etp4hpc.eu/en/events/european-hpc-summit-week-may-12016-prague_307.html	D
16-17 May 2016	ExTASY Tutorial: Tools for Advanced Sampling of	Edinburgh, UK	https://www.epcc.ed.ac.uk/extasy-tutorial-1617-may-	T

	Macromolecular Systems, 16/17 May 2016		2016	
16-20 May 2016	PRACE Spring School and E-CAM tutorial on Molecular & Atomic Modelling	University College Dublin, Ireland	https://events.prace-ri.eu/event/497/	T
18-20 May 2016	Tools registry hackathon	Institut Pasteur, 25 rue du Docteur Roux, 75015 Paris, France	http://tinyurl.com/registryhackathon8	D
19-20 May 2016	Workshop: GROMACS on highly parallel and heterogenous systems	Gottingen, Germany	http://bioexcel.eu/events/workshop-gromacs-on-highly-parallel-and-heterogeneous-platforms/	T
27-28 May 2016	Workshop: Computer Simulation and Theory of Macromolecules	Hunfeld, Germany	http://www.mpibpc.mpg.de/grubmueller/huenfeld	T
6-10 Jun 2016	School on Molecular Modelling for Life Sciences	Pula, Sardinia, Italy	http://molmodel2016.dsf.unica.it/wp/	T
15-17 Jun 2016	BITS 2016	13th Annual Meeting of the Bioinformatics Italian Society	http://bits2016.bioinformatics.it/index.html	T
19-23 Jun 2016	ISC-HPC		http://www.isc-hpc.com/	D
4-9 Jul 2016	EMBO course on integrative modelling	Barcelona, Spain	http://events.embo.org/16-biomol-interact/	D/T
5-7 Jul 2016	EMBL conference: Lifelong Learning in Biomedical Sciences	Heidelberg, Germany	http://www.embl.de/training/events/2016/LLL16-01/index.html	D
24-29 Jul 2016	Gordon Research Conference: Theory and Simulation Across	Girona, Spain	http://www.grc.org/programs.aspx?id=11141	D

	Scales in Molecular Science			
3-7 Sep 2016 (4 Sep = SIG-BioExcel)	ECCB	The Hague, Netherlands	http://www.eccb2016.org/	D/T
17-19 Oct 2016	GROMACS Hackathon (17-19 October at BSC)	Barcelona, Spain		IG
20-21 Oct 2016	BioExcel: workflow training for computational biomolecular research	Barcelona, Spain		IG/T
13-18 Nov 2016	SC 16	Salt Lake City, Utah, USA	http://sc16.supercomputing.org/	D

Appendix 5: Template of the EMBL-EBI on-site course feedback survey. A number of template surveys are used within EMBL-EBI; on-site, off-site, industry. We will select the most appropriate template and customise for each course or workshop.

EMBL-EBI

A blank template - On-site course (Sept 2013)

Course feedback survey

EMBL-EBI collects feedback from every course and workshop we run. The survey is a way for you to inform us about the course you have participated in, what you enjoyed, what you found useful and how we can make improvements. This information is also used to inform the development of new courses and workshops.

Where possible, please take the time to provide written answers and explain as much to us as you can.

EMBL-EBI

A blank template - On-site course (Sept 2013)

About you

This section asks for a few personal details.

We may want to contact you about future EMBL-EBI courses and to gather long-term feedback.

You do not need to provide these details if you do not want to. EMBL-EBI will not use your personal details for any purpose other than that stated above. We will not pass your details to any third party.

1. Name

2. Email

3. May we contact you by email in future to take part in usability testing/ interviews/ surveys or other user research activities ?

- Yes
- No

EMBL-EBI

A blank template - On-site course (Sept 2013)

Course content

This section asks for your thoughts on the content and delivery of the course.

* 4. Please tell us your overall rating for the entire course.

Poor Satisfactory Average Good Excellent

* 5. Please rate each section of the course.

	Did not attend	Poor	Satisfactory	Average	Good	Excellent
Session 1						
Session 2						
Add rows as needed						

Comments

* 6. What was the best part of the course?

* 7. What was the worst part of the course?

* 8. The balance of theoretical and practical content across the course was

Too practical About right Too theoretical

* 9. Have you used the resources covered in the course before?

Unaware of them Used other service Occasionally Frequently

Other (please specify)

* 10. Will you use the tools/resources covered in the course in your future work?

Yes No Maybe

Comments

* 11. Would you recommend this course?

Yes No Maybe

Comments

EMBL-EBI

A blank template - On-site course (Sept 2013)

Course logistics

This section asks for your thoughts on the organisation of the course.

* 12. The overall course organisation (including: registration, accommodation, EMBL-EBI assistance) was:

Poor Satisfactory Average Good Excellent

Comments

* 13. The catering provided during the course was:

Poor Satisfactory Average Good Excellent

Comments

* 14. The accommodation you stayed in during the course was:

Poor Satisfactory Average Good Excellent

Comments

EMBL-EBI

A blank template - On-site course (Sept 2013)

15. Any other comments?

EMBL-EBI

A blank template - On-site course (Sept 2013)

Thank you

Thank you on behalf of EMBL-EBI's training team for taking the time to complete this survey. It is your feedback that helps us to improve and maintain high training standards.