



EXCELERATE Deliverable D11.2

Project Title:	ELIXIR-EXCELERATE: Fast-track ELIXIR implementation and drive early user exploitation across the life sciences	
Project Acronym:	ELIXIR-EXCELERATE	
Grant agreement no.:	676559	
	H2020-INFRADEV-2014-2015/H2020-INFRADEV-1-2015-1	
Deliverable title:	Report on the training needs identified across the ELIXIR community	
	11	
Lead Beneficiary:	UEDIN	
WP Title	ELIXIR-EXCELERATE Training Programme	
Contractual delivery date:	31 August 2016	
Actual delivery date:	31 August 2016	
WP leader:	Chris Ponting and Patricia Palagi	43: UEDIN, 26: SIB
Partner(s) contributing to this deliverable:	Celia van Gelder [6 NBIC], Sarah Morgan [1 EMBL-EBI], Allegra Via [31 CNR], Rita Hendricusdottir [43 UEDIN], Eija Korpelainen [21 CSC], Chris Ponting [43 UEDIN], Terri Attwood [4 UNIMAN] and Patricia Palagi [26 SIB]	

Authors and contributors: Celia van Gelder, Sarah Morgan, Allegra Via, Rita Hendricusdottir, Eija Korpelainen, Chris Ponting, Terri Attwood and Patricia Palagi, on behalf of the Training Coordinators Group

Table of contents

1. Executive Summary	2
2. Project objectives	3
3. Delivery and schedule.....	3
4. Adjustments made.....	3
5. Background information	3
Annex 1: Report on the training needs identified across the ELIXIR community.....	7

1. Executive Summary

With this deliverable, we report on the efforts undertaken since 2014, and in particular since the start of the ELIXIR-EXCELERATE project, by the ELIXIR Training Platform regarding the identification of training needs in ELIXIR, the actions that were carried out to fulfil these needs, and outlines our plans for the years ahead.

The ELIXIR Training Coordinators Group (TrCG), which gathers representatives from all Nodes, and the ELIXIR training community play a pivotal role in all ELIXIR Training activities. EXCELERATE's and the Nodes' training activities are complementary and it is of utmost importance that these are strategically and coherently aligned, which is insured with the help of the TrCG. Furthermore, the ELIXIR Training Platform is well positioned among and actively engaging with global partners (e.g. SWC/DC, BD2K and GOBLET) to improve training and share best practices within and beyond ELIXIR's borders.

ELIXIR Training has identified three major target audiences: infrastructure users, infrastructure developers, and trainers. Several activities have already been provided for these groups in project year 1 and many more are planned for project year 2 to 4. Through the identification of priorities for ELIXIR's Training Platform, via Node surveys (2014) and through dedicated liaison persons, we were able to start addressing the critical training deficits important for these three groups - including members from academia as well as industry - across Europe. ELIXIR Training activities include workshops in the areas of Software or Data Carpentry, High-Performance Computing, Workflow Tools, Data Management and Stewardship, and Train-the-Trainer courses in which expert staff learn techniques and tools for teaching bioinformatics.

Moreover, the ELIXIR Training Infrastructure is gradually getting shaped. It is improving access to training via its online TeSS Training Portal and e-learning solutions, and a training framework with an initial set of guidelines, best-practices and metrics has started to be developed.

The core of this report consists of a preliminary analysis of a survey, which aimed to capture the most up-to-date inventory of the training landscape and training provision in ELIXIR Nodes, to quantify the Nodes' trainer capacity, and to identify specific areas of bioinformatics where training provision is low/needed, with special attention paid to the EXCELERATE use cases. More than 70 partner organisations from 16 Nodes with experience in providing training have responded to the survey. The preliminary results confirm a need for prioritising training for the EXCELERATE use cases within the marine metagenomics, crop and forest plant, rare disease and human data communities, and to improve the training capacity in the Nodes to deliver this particular training.

The survey has also enabled the definition of specific training needs of the individual use case communities, which will be refined further by more detailed surveys to be sent out soon. Further analysis of those results will allow the timely provision of specific training as

the use cases come up in full speed (from project year 2 and beyond). Also in this case, the TrCG and the training liaisons will play a major role in bridging to each of the Work Packages (both the use cases and other WPs).

Prioritising the training activities and areas is a challenge for the ELIXIR Training Platform and a critical factor in the overall effectiveness of training in the context of EXCELERATE. This is, however, very much dependent on the direction of travel and advances of the use cases and other EXCELERATE WPs, as well as the other ELIXIR Platforms. Training topics are emerging only now, after one year of development. Planning of training courses depends on the progress made in different ELIXIR and EXCELERATE activities, and courses will need to be timely and aligned well with the Nodes' training provision. Among other challenges the ELIXIR Training Platform will face in the subsequent project years are: i) devising and adopting an ELIXIR Industry training strategy; ii) securing future funding in order to sustain the growing ELIXIR Training community and beyond EXCELERATE; and iii) improving the visibility and discoverability of ELIXIR training courses, in particular by promoting the use of TeSS by trainers and trainees. The ELIXIR Training Platform, in close collaboration with the TrCG, will tackle these challenges from project year 2 and beyond.

2. Project objectives

With this deliverable, the project has reached or the deliverable has contributed to the following objectives:

No.	Objective	Yes	No
1	Build a sustainable training infrastructure for ELIXIR's community - encompassing both a technical infrastructure and training expertise, as well as mechanisms for guaranteeing quality of training		x
2	Develop and deliver training in topics selected as training gaps within the ELIXIR community in selected application areas	x	

3. Delivery and schedule

The delivery is delayed: Yes No

4. Adjustments made

No adjustments were made.

5. Background information

Background information on this WP as originally indicated in the description of action (DoA) is included here for reference.

Work package number	11	Start date or starting event:	month 1
---------------------	----	-------------------------------	---------

Work package title	ELIXIR-EXCELERATE Training Programme
Lead	Work Package Leads: Chris Ponting (UK) and Patricia Palagi (CH)
<p>Participant number and person months per participant</p> <p>1 – EMBL (14 PM), 2 – UOXF (12 PM), 4 – UNIMAN (30.8 PM), 5 – UTARTU (42 PM), 6 – NBIC (7.5 PM) [SARA 2 PM, VU 2 PM], 7 – CNIO (1 PM), 11 – UMA (2 PM), 13 – CSIC (0.8 PM), 16 – FCG (12 PM), 20 – CSC (10 PM), 24 – UiO (2 PM), 25 - SIB (14 PM), 26 – CNRS (4 PM), 30 – CNR (12 PM), 31 – LIU (4 PM), 32 - UL (42 PM), 35 – MU (7.2 PM), 40 – HUJ (4 PM), 42 – FORTH (6 PM), 43 - UEDIN (8 PM), 44 – UCAM (24 PM)</p>	
<p>Objectives</p> <p>Build a sustainable training infrastructure for ELIXIR’s community - encompassing both a technical infrastructure and training expertise, as well as mechanisms for guaranteeing quality of training (Task 11.1)</p> <p>Develop and deliver training in topics selected as training gaps within the ELIXIR community in selected application areas (namely WP6 to 9 Use Cases) (Task 11.2)</p>	
<p>Description of work and role of partners</p> <p>As medicine and the life sciences become increasingly centred on the generation, analysis and interpretation of big data, most science professionals will need to become more proficient in exploiting bioinformatics data and systems. The ELIXIR Community is estimated at 500,000 people, drawn from across the spectrum of individuals ranging from life scientists and bioinformaticians, to tool developers and infrastructure operators. This WP will up-skill European researchers focused on the WP6 to 9 Use Cases, who will be empowered to more effectively exploit the data, tools, standards and compute infrastructure provided by ELIXIR, and on ELIXIR-EXCELERATE developers and infrastructure operators.</p> <p>This Training Programme will be the foundation upon which international bridges will be built, in order to harmonise efforts, to share resources, to avoid duplication/redundancy and to maximise effectiveness. Specific organisations and initiatives with which to collaborate may be (but are not limited to): GOBLET (the Global Organisation for Bioinformatics Learning, Education and Training), ISCB (the International Society for Computational Biology), and Software Carpentry Foundation, and also the training programmes of other ESFRIs such as BBMRI, EATRIS, and ISBE. Collaboration with GOBLET essentially provides a ready-made global gateway to ELIXIR’s training resources, and will ensure that the foundations built by GOBLET are not duplicated – the leader of task 11.1.2 also leads GOBLET, assuring that their distinct roles will evolve in harmony. Close engagement with industry will be sought such as through the ELIXIR Industry Programme.</p> <p>ELIXIR is a new research infrastructure in its implementation phase. The ELIXIR community as represented in the EXCELERATE consortium is very diverse in a number of aspects, including for example maturity of individual Nodes and level of knowledge and expertise. For workshops held in the context of EXCELERATE, we will always try to identify the relevant expertise from within the consortium; however, as ELIXIR develops cutting-edge infrastructure for life sciences, external experts will need to be invited on a</p>	

relatively regular basis to contribute to the discussions. Their travel and subsistence will be supported by the grant.

A robust, high quality training infrastructure supporting ELIXIR resources and services will increase the impact and visibility of ELIXIR as a whole, ensuring that ELIXIR resources are introduced into industry and academia. It will catalyse and support the (self-) training of researchers, increase the impact of ELIXIR services, and facilitate scientific excellence.

All activities in this Training Programme will build on existing national training experience and expertise in the Nodes, and will complement, enhance, and disseminate these activities. Training activities will be closely aligned with the ongoing development of data infrastructure, standards, tools and services in each Node. This work paves the way for a later comprehensive, high quality and sustainable training programme to be rolled out over the entire ELIXIR Community.

The ELIXIR Training Programme targets the following distinct stakeholder groups:

- The ELIXIR community encompassing its partners, associated industry and SMEs (Task: 11.1)
- Developers and Infrastructure operators of ELIXIR's services (Task: 11.2.1)
- Users of ELIXIR's services (Task: 11.2.2)
- ELIXIR's trainers (Task: 11.2.3)

Task 11.1: Building an ELIXIR Training Infrastructure (143.3 PM)

Subtask 11.1.1: Assessing training quality, good practice and impact (45 PM)

ELIXIR Training needs to be timely, impactful, of high quality, and at scale. It will draw upon existing expertise in member countries that has yet to be combined under a unified structure. A Good Practice Coordinator will identify and deliver a framework of good practices throughout ELIXIR and ELIXIR-EXCELERATE. An initial workshop involving the ELIXIR Training Coordination Group (TrCG), industry/SMEs and other stakeholders will focus on defining specifications, metrics and key performance indicators to ELIXIR training. This subtask will build confidence for those seeking training (whether from academia, industry or other sectors) that ELIXIR training is being delivered to high standards, ensuring "best in class" training provision. It will also provide ELIXIR with a mechanism to capture and report on the impact of its training programme on the European and international level, and will harmonise across Nodes in standards, metrics and sharing of training materials. Development of well- structured training routes – through the use of workflows – embedded in online resources (WP11.1.2 and WP11.1.3) will specifically address urgent needs of WP6 to 9 Users and industry for cost-effective, time-effective, impactful training.

Partners: UK, EMBL-EBI

Additional resources required: Workshops €20,000

Subtask 11.1.2: TeSS Training Portal (51.3 PM)

A training portal, TeSS, will be developed in this task to be an active forum for aggregating, disseminating and coordinating information on ELIXIR-EXCELERATE's training activities/materials, including those relating to ELIXIR's Core Resources. Building on the TeSS prototype being piloted by ELIXIR-UK, this task will 'harden' the prototype, harness the outputs from WP6 to 9's Use Cases and 3rd-party content providers, and synergise globally with GOBLET. Importantly, training information from all Nodes will be pulled into the TeSS, ensuring ELIXIR- wide coverage. The TeSS will enable registration and discovery of training activities/materials through multi-centre information aggregation, it will allow users to collect ('package') sets of materials/tools/data required for training,

and offer workflows that allow related resources to be identified and harvested from source. The roles of GOBLET and TeSS are complementary: the former (a materials/course repository and trainer directory) acts as a feed to the latter (a resource aggregator and dissemination hub), obviating the need for ELIXIR-EXCELERATE to build its own repository. No mechanism currently exists via which users may either readily discover ELIXIR- EXCELERATE training events/ resources (many of which are dispersed on websites across ELIXIR member states and beyond) and/or determine their relevance (e.g., what is the audience of this course (beginner, advanced, etc.), what is its duration, where is it being held, which course should I take next, etc.?). The added value of this task is therefore in coordinating and making discoverable ELIXIR-EXCELERATE training activities/materials, and surfacing information in ways that support user decisions and choices. We will ensure wide uptake by all stakeholders (from trainers and trainees, to resource providers and developers, including those across industry/SMEs) via community-building events. This subtask will result in the release of the TeSS as an open resource, contributed to by the community, and shaped both by the community and by the outputs of ELIXIR- EXCELERATE's Use Cases (WP6 to 9).

Partners: UK, CH, PT, NL

Subtask 11.1.3: eLearning (47 PM)

Distance training has become essential to reach large audiences spread over many countries, as is the case for the ELIXIR community. This task will survey the current technological and pedagogical options, the existing e- learning expertise and technology in the Nodes, and possible international partners (Coursera, edX, Udacity, GOBLET etc.). In concert with the ELIXIR training community, we will decide for the best and suitable e- learning strategy for ELIXIR and then proceed into implementing this strategy, which will complement the TeSS training portal. During this process we will develop ELIXIR e- learning expertise, which will be exploited to create guidelines and recommendations for future trainers. A major part of this task will be devoted to derive scalable training materials related to the WP6 to 9 Use Cases and ELIXIR resources, which will all be made available to the ELIXIR community of trainers, developers and users. The trainers will benefit when delivering their courses; developers and users will benefit from the availability of materials. Partners: SI, NL, CH, PT, IL

Additional resources required: Workshops €20,000

Task 11.2: Delivering training to the ELIXIR-EXCELERATE community (116 PM)

Subtask 11.2.1: Train the Developer and Infrastructure Operator (40 PM)

This subtask targets all developers and infrastructure operators who are developing and maintaining ELIXIR services, distributed over all Nodes and ELIXIR WPs (most notably WP3, WP4, WP6 to 9, and WP10). The main activities are: a) mapping the competences across the Nodes and analysing collected training needs from the technical people through internet queries, discussion rounds at the annual ELIXIR All Hands Meeting and within the technical coordinators group; b) providing targeted training based on the mapped needs as face-to-face workshops or webinars where applicable; c) implement a dedicated virtual community space to ease the communication between the developers and infrastructure operators across the Nodes. The aim is to have a virtual coffee room where issues regarding resource development and ELIXIR services can be shared and re-occurring problems can be easily noticed and solved. The virtual community will allow buddying up individual developers across the Nodes who are working on similar tasks but otherwise would not meet nor have discussion possibilities. On top of all the information gathered, specific training can be quickly delivered in response to identified needs. Improving developers' skills and enabling easy knowledge exchange across the community will have immediate impacts on the quality of tools provided by the ELIXIR

community. With substantial competition among analysis tools and databases in the life sciences, better designed and executed services obtain greater visibility not only in Europe but worldwide thus further raising the awareness of ELIXIR services.

Partners: EE, EMBL-EBI, SE, IL, FI

Additional resources required: Training Procurement €60,000

Subtask 11.2.2: Train the Researcher (43PM)

It is essential that ELIXIR-EXCELERATE Users (e.g. in WP6 to 9) acquire skills that empower them to transform big data into meaningful knowledge. This Subtask will plug the training gaps using HPC facilities and expertise, specific cloud computing workflow training needs (WP4) and existing solutions, and the provision of training workshops. Life scientists, computational biologists, and bioinformaticians engaged in Marine Metagenomics, Plant Genotype- Phenotype interactions, Rare Disease and Biological Sample Collection research will be trained in domain-specific knowledge. To become independent and savvy users of the ELIXIR resources training will be given in: scientific programming, statistics, HPC, ontologies, workflows, data curation and annotation (WP3), as well as in the effective use of standards and FAIR data. Reproducibility in the life sciences is a rapidly growing problem. Targeted training is anticipated to improve experimental reproducibility, to empower scientists to address new avenues of research requiring big data analysis, and to cross-pollinate diverse disciplines with innovative analytical approaches. Training opportunities will be prioritised, in consultation with Industry/SMEs, and training delivery will be monitored (WP11.1).

Partners: ES, CH, UK, NL, SI, NO, PT, EL Additional resources required: Workshops €140,000

Subtask 11.2.3: Train the trainer (33PM)

For the ELIXIR training to achieve maximum impact, adequate capacity to train in all Nodes is required. This task will train the trainer pool that already exists – albeit in a fragmented manner – across ELIXIR Nodes and to identify and train new trainers. It will build a highly skilled and coherent community of training instructors by initially surveying available training capacity and developing an appropriate train the trainer (TtT) framework. Using targeted surveys and drawing upon knowledge and skills from Nodes with greater experience (and with WP11.1), expertise will be developed in Nodes, for ELIXIR Core Resources and Use Cases, where training opportunities have hitherto been scarce. The task will deliver a framework and associated material and guidelines for training new trainers across ELIXIR, from both industry and academia. To enhance delivery, training will take advantage of ready-to-run virtual machines (VMs, developed within ELIXIR and/or in collaboration for instance with Biolmg.org) and clouds, assisted by a workshop that gathers both trainers and infrastructure specialists from WP4.2. TtT sessions will accompany training courses in Marine Metagenomics (WP6), Plant Genotype-Phenotype (WP7) and Rare diseases (WP8) providing good practice, and guidance for prospective new trainers. TtT will build capacity whilst stimulating collaborations spanning ELIXIR Nodes and other stakeholders, including industry/SMEs which represent a currently unexplored resource.

Partners: IT, EMBL-EBI, SE, CH

Additional resources required: Workshops €154,000

Annex 1: Report on the training needs identified across the ELIXIR community

Report on the training needs identified across the ELIXIR community

Authors and contributors: Celia van Gelder, Sarah Morgan, Allegra Via, Rita Hendricusdottir, Eija Korpelainen, Chris Ponting, Terri Attwood and Patricia Palagi, on behalf of the Training Coordinators Group

Table of contents

1. Executive Summary	3
2. ELIXIR Training	4
2.1. ELIXIR's Training Coordinators' Group	4
2.2. ELIXIR Training builds on pre-existing activities in the national Nodes	5
3. Target audiences for ELIXIR Training	5
3.1. Infrastructure Users	6
3.2. Infrastructure Developers	6
3.3. Trainers	6
3.4. Range of training needs for different target audience and their domains	7
4. Identified training gaps, needs and solutions - 2014-2015	7
4.1. Process of identifying the needs and gaps 2014-2015	7
4.2. Identified gaps, needs and solutions for each of the ELIXIR Target audiences	8
5. ELIXIR Training efforts for ELIXIR Users	9
5.1. Software and Data Carpentry workshops	9
5.2. Workflow tools: Galaxy and Chipster	10
5.3. Training in HPC	11
5.4. Training for Use-Cases	11
5.5. Industry	12
5.6. ELIXIR Training and Data management & stewardship	12
6. ELIXIR Training efforts for ELIXIR Developers	12
7. ELIXIR Training efforts for ELIXIR Trainers	13
7.1. Train-the-Trainer	13
7.2. ELIXIR Training Infrastructure	13
<i>TeSS Training Portal: Findability of courses</i>	13
<i>E-learning</i>	14
7.3. Measuring ELIXIR training quality and impact	14
8. ELIXIR Training Survey 2016	15
8.1. Aims of the 2016 survey	15
8.2. The survey	15
8.3. Survey responses	17
8.4. ELIXIR members and EXCELERATE use-cases	17
8.5. Experience in training and audience	18

8.6. Training courses	19
8.7. Training resources and capacity	21
8.8. Next steps	22
<i>Use cases</i>	22
<i>Training in major topics</i>	22
<i>Numbers of trainers</i>	23
<i>Definition of ELIXIR membership and branding</i>	23
<i>Course advertising</i>	23
9. Collaborations with other Training initiatives and Programmes	24
10. Conclusions	26

Tables and figures

Table 1 Bioinformatics subject areas	16
Figure 1 ELIXIR training target audiences	6
Figure 2 Process of needs identification and timeline 2014-2016	8
Figure 3 Number of instructors trained per Node in Lausanne, January 2016	10
Figure 4 Numbers of responses across ELIXIR Nodes	17
Figure 5a,b Involvement of Node institutions/organisations in use-cases	17
Figure 6 Number of institutes across ELIXIR Nodes which run training courses; 8b: Period over which courses have been running	18
Figure 7 Overview of the target audience across three different communities (Academia, Industry, Healthcare/clinical)	19
Figure 8 Number of Nodes offering courses in each major topic	20
Figure 9 Number of courses offered by major topic	20
Figure 10 Most organisations have no major issues regarding training facilities (IT rooms, support, etc.) but lack trainers and monetary resources to run the courses they need. "Other" includes the following: more extended time planned for the courses; educational material; access to a better eLearning infrastructure; admin staff; requests	21

1. Executive Summary

With this deliverable, we report on the efforts undertaken since 2014, and in particular since the start of the ELIXIR-EXCELERATE project, by the ELIXIR Training Platform regarding the identification of training needs in ELIXIR, the actions that were carried out to fulfil these needs, and outlines our plans for the years ahead.

The ELIXIR Training Coordinators Group (TrCG), which gathers representatives from all Nodes, and the ELIXIR training community play a pivotal role in all ELIXIR Training activities. EXCELERATE's and the Nodes' training activities are complementary and it is of utmost importance that these are strategically and coherently aligned, which is insured with the help of the TrCG. Furthermore, the ELIXIR Training Platform is well positioned among and actively engaging with global partners (e.g. SWC/DC, BD2K and GOBLET) to improve training and share best practices within and beyond ELIXIR's borders.

ELIXIR Training has identified three major target audiences: infrastructure users, infrastructure developers, and trainers. Several activities have already been provided for these groups in project year 1 and many more are planned for project year 2 to 4. Through the identification of priorities for ELIXIR's Training Platform, via Node surveys (2014) and through dedicated liaison persons, we were able to start addressing the critical training deficits important for these three groups - including members from academia as well as industry - across Europe. ELIXIR Training activities include workshops in the areas of Software or Data Carpentry, High-Performance Computing, Workflow Tools, Data Management and Stewardship, and Train-the-Trainer courses in which expert staff learn techniques and tools for teaching bioinformatics.

Moreover, the ELIXIR Training Infrastructure is gradually getting shaped. It is improving access to training via its online TeSS Training Portal and e-learning solutions, and a training framework with an initial set of guidelines, best-practices and metrics has started to be developed.

The core of this report consists of a preliminary analysis of a survey, which aimed to capture the most up-to-date inventory of the training landscape and training provision in ELIXIR Nodes, to quantify the Nodes' trainer capacity, and to identify specific areas of bioinformatics where training provision is low/needed, with special attention paid to the EXCELERATE use cases. More than 70 partner organisations from 16 Nodes with experience in providing training have responded to the survey. The preliminary results confirm a need for prioritising training for the EXCELERATE use cases within the marine metagenomics, crop and forest plant, rare disease and human data communities, and to improve the training capacity in the Nodes to deliver this particular training.

The survey has also enabled the definition of specific training needs of the individual use case communities, which will be refined further by more detailed surveys to be sent out soon. Further analysis of those results will allow the timely provision of specific training as the use cases come up in full speed (from project year 2 and beyond). Also in this case, the TrCG and the training liaisons will play a major role in bridging to each of the Work Packages (both the use cases and other WPs).

Prioritising the training activities and areas is a challenge for the ELIXIR Training Platform and a critical factor in the overall effectiveness of training in the context of EXCELERATE. This is, however, very much dependent on the direction of travel and advances of the use cases and other EXCELERATE WPs, as well as the other ELIXIR Platforms. Training topics are emerging only now, after one year of development. Planning of training courses depends on the progress made in different

ELIXIR and EXCELERATE activities, and courses will need to be timely and aligned well with the Nodes' training provision.

Among other challenges the ELIXIR Training Platform will face in the subsequent project years are: i) devising and adopting an ELIXIR Industry training strategy; ii) securing future funding in order to sustain the growing ELIXIR Training community and beyond EXCELERATE; and iii) improving the visibility and discoverability of ELIXIR training courses, in particular by promoting the use of TeSS by trainers and trainees. The ELIXIR Training Platform, in close collaboration with the TrCG, will tackle these challenges from project year 2 and beyond.

2. ELIXIR Training

"It is estimated that by 2018 there will be a talent shortage in the US alone of between 140,000 and 190,000 trained experts to analyse the increasing amount of data being generated." [McKinsey Global Institute](#)

Big data analysis is diversifying and transforming life sciences in industry and academia across Europe. Nevertheless, only by bridging the critical gap in highly skilled data professionals will the life sciences reach their full potential. In the UK, for example, big data professionals increased in number 10-fold between 2009 and 2014, and further steep increases are [expected](#). These widening recruitment gaps cannot be filled solely by new university graduates, because their numbers are too few and they have seldom acquired the scientific, leadership and interpersonal skills necessary to maximise big-data opportunities.

ELIXIR's Training Programme seeks to narrow Europe's critical gap in highly skilled life-science data-management professionals, specifically by addressing the challenges inherent in training a large and diversifying user base to fully exploit ELIXIR resources. The Programme further seeks to identify where its necessarily restricted training activities provide the greatest impact on health and wealth advances.

The ELIXIR Training Platform activities described in this report align with, and build on, the Training recommendations that were made during the [ELIXIR Preparatory Phase](#): these included coupling the development of resources to the provision of training, supporting the trainer community with infrastructure components, and tightly integrating with other training initiatives, both in Europe and beyond.

ELIXIR Training has identified important training gaps across the life sciences in Europe, relevant to both academia and industry. It is essential that these gaps are addressed as a matter of urgency, and that scientists are provided with high-quality training activities that provide greatest impact. Once this is achieved, scientists should be capable of transforming big data into knowledge and testable hypotheses, leading to an improved understanding of the biological processes underpinning many areas of life-science research.

In order to effectively deliver high-level and impactful training, ELIXIR counts on two strong communities of experts: the Training Coordinators' Group and the national Nodes.

2.1. ELIXIR's Training Coordinators' Group

ELIXIR's Training Coordinators' Group (TrCG) was established in September 2014 to deliver ELIXIR's training mission, and to harmonise training activities across member countries. The TrCG consists of training experts, each representing their own ELIXIR

Node, seeking to actively collaborate within ELIXIR, across Europe and beyond. TrCG meets monthly via video conference and twice-yearly face-to-face, which enables swift transfer of best practices from one European centre of training excellence to sites elsewhere.

The TrCG group shares information, identifies and implements best practices in training services, identifies training needs, and develops or delivers new training activities to fill these training gaps. In addition, the group liaises with each of the ELIXIR use-cases and platforms, and leads on engagement with other training initiatives, such as with GOBLET, BD2K TCC, other research-infrastructure projects and relevant training centres. This group of experts, and the extended network within their Nodes, was essential when ELIXIR training needs were being defined, making use of the surveys described below.

2.2. ELIXIR Training builds on pre-existing activities in the national Nodes

The ELIXIR Nodes and their ongoing activities sit at the heart of ELIXIR Training: existing national and local training infrastructures continue to be the primary producers and providers of new training activities in ELIXIR. The substantial added benefit of the ELIXIR Training infrastructure is in coordinating the Nodes' efforts, and in setting up a structured training infrastructure across Europe, including a framework of best-practices and detailed guidelines. ELIXIR Training thus builds and integrates the Nodes' foundational activities into a comprehensive Training Portfolio, focusing on training in subjects and for audiences identified by ELIXIR as priorities.

3. Target audiences for ELIXIR Training

Tailored training is required for targeting researchers with heterogeneous backgrounds in order that they gain full benefit from the services that ELIXIR provides. Within ELIXIR training, audiences range from infrastructure users (e.g., clinicians, wet-lab scientists and also bioinformaticians) to more established bioinformaticians/developers who deliver and maintain the services (infrastructure developers). In addition, a cohort of trainers needs to be trained who then can train others effectively, in this way working towards a scalable model for training large numbers of life scientists. Figure 1 illustrates the target audiences for ELIXIR training, and the overlapping groups of trainees.

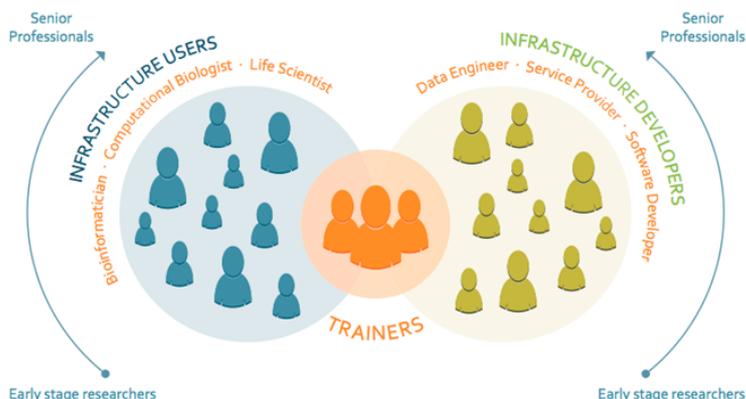


Figure 1 ELIXIR training target audiences

Below, we provide our working definitions of these three user groups:

3.1. Infrastructure Users

We define infrastructure users as those whom our resources are designed to serve. These researchers cover the entire spectrum, from experimental researchers in life sciences, who require basic skills in ‘small data’ analysis (including scripting and statistics), to ‘big data’ scientists/bioinformaticians collecting large data-sets, whose aim is to elicit new biological knowledge using integrative analysis.

3.2. Infrastructure Developers

We define infrastructure developers as the data engineers, service providers and software developers who develop and maintain ELIXIR services run from ELIXIR Nodes. ELIXIR services include: data resources, tools, compute provision and standards development. In addition, infrastructure developers support scalable and sustainable solutions for data management. Their aim is to provide support to the infrastructure users. The training needs are very diverse, and range from data management to software development, to maintaining ELIXIR services and resources.

3.3. Trainers

Trainers are competent users of a resource who can train other users. They need to have a deep understanding of users’ needs and learning styles in order to tailor their training accordingly, and structure their courses to be of high value to a diverse range of users. The trainers are users/researchers and/or developers themselves, who combine their training activities with their research/development activities. Building a network of new trainers around ELIXIR curricula and resources will increase the impact of ELIXIR as a whole, ensuring ELIXIR resources are introduced into both academia and industry, and potentially opening up new areas of engagement across Nodes.

3.4. Range of training needs for different target audience and their domains

Most research scientists are unable to take full advantage of the data bonanza that is arising from new technologies. Critical training needs across this community are, broadly, in the analysis and interpretation of genome-scale data of many types. At the other end of the spectrum, infrastructure developers need a wide range of technical training, as well as training in life-science subjects to be able to understand the biological rationale upon which large data-rich projects are designed.

In addition, users from different life-science domains may have very specific training requirements relating to ELIXIR services. For example, scientists analysing medical research data will require services that are able to handle sensitive data to protect the privacy of patients. ELIXIR Training accommodates these specific requirements by liaising closely with the different research communities of the four ELIXIR use-cases: rare diseases (WP8), biomedical and human genomics (WP9), marine sciences (WP6) and plant sciences (WP7).

Furthermore, when comparing the training needs between industry and academic users, an [ELIXIR-UK survey](#) has shown that industry users, a major user group of ELIXIR, have highly similar training requirements to academic users.

Given the large number of European scientists/developers working in data-rich science who require training, the heterogeneous target audience and training needs, ELIXIR Training has to define the training gaps throughout Europe and to prioritise the training efforts accordingly. This exercise has been achieved in 2 phases, in 2014-2015 and in 2016, as detailed below.

4. Identified training gaps, needs and solutions - 2014-2015

4.1. Process of identifying the needs and gaps 2014-2015

The process of identifying specific training gaps for ELIXIR, summarised in Figure 2, began in early 2014 and has been ongoing ever since. In January 2014, a survey was sent out to all Heads of Nodes (17 Nodes at the time) to obtain an insight into the Nodes' initial offers and needs in training. At that time, the majority of the Nodes indicated that they would offer training courses (both on their current and future Node resources), and the majority of Nodes were committed to share both their materials and their expertise. Other training-related components that were offered and needed were, for example, e-learning resources, training benchmarking and evaluation systems, outreach, training for trainers (train-the-trainer) for competencies development, and a training portal. Nodes also indicated that they were planning to train on resources, and on issues regarding data reliability, maintainability and consistency.

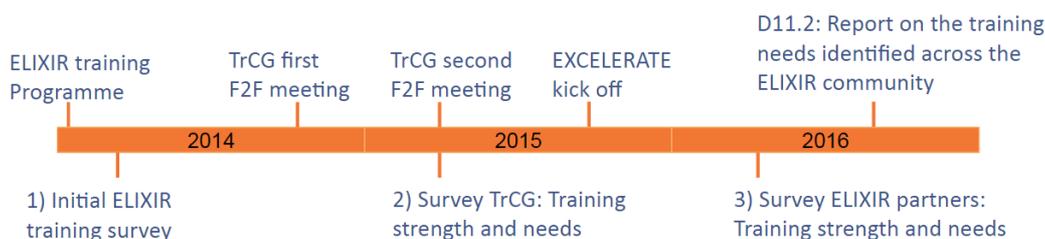


Figure 2 Process of needs identification and timeline 2014-2016

The outcome of the 2014 survey, together with the recommendations of the ELIXIR Preparatory Phase Document, served as a basis for defining the [9 Training Objectives of ELIXIR Training](#). At the end of 2014, when the TrCG was assembled, the Training Work Package plans for the EXCELERATE proposal were shaped by the outcomes of that survey and the pilot actions proposed by the TrCG.

As a result of this combined effort, six subtasks were identified for the EXCELERATE Training Work Package (WP11): of these, three address building the ELIXIR Training Infrastructure, the remaining three aiming to deliver training to the relevant target audiences (summarised in Box 1).

11.1 To build a sustainable training infrastructure for ELIXIR's community - encompassing both a **technical** infrastructure and **training expertise**, as well as **mechanisms** for guaranteeing quality of training.

- Assessing training quality, good practice and impact
- TeSS Training Portal
- Elearning

11.2 To **develop and deliver training** in topics selected as training gaps within the ELIXIR community in selected application areas

- Train-the-developer and infrastructure operator (TtD)
- Train-the-researcher (TtR)
- Train-the-trainer (TtT)

Box 1 Six subtasks of the EXCELERATE Training Work Package (WP11)

4.2. Identified gaps, needs and solutions for each of the ELIXIR Target audiences

The major identified needs for our target audiences and the efforts that are being made to fill these gaps are briefly discussed below, and are summarised in Box 2. Furthermore, Annexes II and III list the ELIXIR Training activities focused on these gaps, which have already been organised, either since the beginning of Project Year 1 (PY1) or those planned for PY2 and beyond. These annexes show a condensed view of ELIXIR training events. The complete landscape of ELIXIR Training is much larger; there are many more training events in the Nodes that are not listed in these tables. To make them all findable for the diverse audiences, they are actively

announced in the [ELIXIR Training Portal](#), [TeSS](#), and on the ELIXIR website.

<p><u>Users</u></p> <ul style="list-style-type: none">• SWC/DC• Workflow tools (Galaxy, Chipster)• HPC training• Use-case needs• Industry needs• Data management & stewardship <p><u>Developers</u></p> <ul style="list-style-type: none">• Virtual coffee room - <i>Mechanism to build and connect the community</i>• Galaxy for developers• Data stewardship & data management <p><u>Trainers</u></p> <ul style="list-style-type: none">• TtT events and training• TtT framework to support trainers in the Nodes: materials and guidelines, incl metrics, standardisation of descriptors, etc.

Box 2 ELIXIR training target audiences and the efforts provided to fulfil their training needs

5. ELIXIR Training efforts for ELIXIR Users

5.1. Software and Data Carpentry workshops



Software and Data Carpentry (SWC/DC) workshops target researchers with little programming experience, to provide them with essential skills for dealing with and for analysing data using programming languages.



By focusing on core skills, and making sure that best practices are passed on in a useful and effective way, SWC/DC workshops act as a connector between researchers and service providers, allowing both sides to communicate better and work more effectively. The workshops are designed to improve overall data literacy: the main learning objectives are to be able to retrieve, view, manipulate, analyse and store one's own and other's data in an open and reproducible way.

To introduce the SWC/DC principles into the ELIXIR community the [ELIXIR Pilot: Software and Data Carpentry roll-out to ELIXIR Nodes](#) was run in 2015, aiming to: i) introduce the SWC/DC concept, curriculum and training model to as many ELIXIR Nodes as possible, ii) run a small number of Data Carpentry workshops, and iii) train a group of SWC/DC instructors within different Nodes (Figure 3).

In particular, this pilot:

- provided solid foundations for setting up a regular training programme across ELIXIR in computational skills for life sciences;
- helped to grow the certified instructor pool within ELIXIR;
- provided an opportunity to the ELIXIR training community to acquire and share best practices in training delivery;
- allowed the creation of ELIXIR-specific training materials ([16-17 March 2015](#) and [22-23 June 2015 Data Carpentry](#) material-preparation hackathons) that will be used in future SWC/DC ELIXIR workshops.

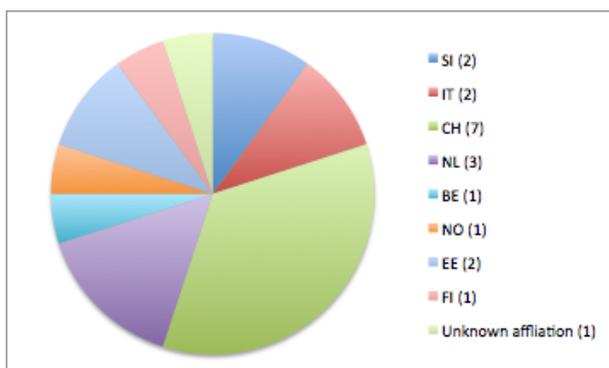


Figure 3 Number of instructors trained per Node in Lausanne, January 2016

As a result of the pilot project, ELIXIR has organised 11 further SWC/DC workshops in ELIXIR Nodes. The workshops trained around 240 life-science researchers from 13 ELIXIR Nodes. In addition, it created a pool of new SWC/DC instructors within ELIXIR, which will continue to grow in future.

In 2016, ELIXIR Training continued to build on the results of the ELIXIR SWC/DC pilot. Its ambition is to introduce the SWC/DC concepts and curriculum to all ELIXIR Nodes, and to grow an ELIXIR SWC/DC trainer community of certified instructors. To identify the awareness about SWC/DC in the Nodes, and to obtain insight into their SWC/DC needs, a [survey](#) was sent out to the TrCG in May 2016. Its [results](#) indicate that there is a strong interest within Nodes in running these workshops, in training new instructors, and in hosting such events in their countries. At the time of writing (August 2016), ELIXIR Training is working in closely with these Foundations to finalise a collaboration agreement, which will support and facilitate the spread of the SWC/DC model through all Nodes. The aim is to have the agreement in place from October 2016.

5.2. Workflow tools: Galaxy and Chipster

[Galaxy](#) and [Chipster](#) are excellent workflow tools for exposing a broad group of life scientists to an extensive range of bioinformatics tools, without the need for deep programming knowledge. They allow researchers that are not comfortable with programming languages to analyse data using bioinformatics tools via a graphical user interface.

When the ELIXIR Galaxy Working Group was formed, it became evident that there is a need to coordinate the Galaxy groups/initiatives spread throughout Europe, and to bring them together within one forum. Training was also identified as one of the [recommendations](#) from this group. Users and developers will need to be trained in using Galaxy to make the best use of its bioinformatics tools. Target audiences, for example, will be users of genomics, proteomics and metabolomics tools, but also developers, who have to set up and maintain Galaxy servers and pipelines.

In WP11, training in Galaxy is foreseen for both developers and administrators, as well as for life-science researchers. In this context, workshops will be organised across Europe where Galaxy specialists from several Nodes will be involved. During the ELIXIR All-hands meeting in Barcelona in 2016, a breakout session was organised to discuss Galaxy training needs. One of the outcomes of this session was the need for an ELIXIR Workshop for Galaxy Developers. This workshop will be held on 16-20 January 2017 in Strasbourg (FR), and it is organised by the [ELIXIR Galaxy Working Group](#), ELIXIR-FR and the WP11 TtD subtask. Furthermore, the ELIXIR

Galaxy WG, which is composed, amongst others, of technical and training coordinators of ELIXIR Nodes involved in Galaxy and Galaxy training, is actively liaising with the international Galaxy community.

[Chipster](#) is free, open source analysis software for high-throughput data, such as NGS. Many institutes have set up their own Chipster servers around the world, including 10 ELIXIR member countries. It is particularly well suited for training biologists, because it allows them to concentrate on understanding the analysis methods rather than struggling with practicalities, such as writing R code. Importantly, Chipster is easy to install because it is available as a ready-run virtual machine image, so the maintenance burden often associated with this kind of platform is minimal.

Chipster has a lot of training material available, including tutorial videos and a book on RNA-seq data analysis. Also a distance-learning course has been recently developed and tested in collaboration with other ELIXIR Nodes. Chipster also contains ready-made example sessions on the analysis of different types of NGS data, which can be used for training. The ELIXIR FI Node runs about 15 Chipster training courses every year in Finland and abroad. Some of these are EXCELERATE training workshops, and therefore focus on the EXCELERATE use-cases. In 2016, a workshop on rare diseases was organised in collaboration with WP8, and it gathered more than 40 people from different countries. The next EXCELERATE workshop is planned for 2017 in collaboration with the marine metagenomics use-case (WP6) and the PRACE (Partnership for Advanced Computing in Europe) Research Infrastructure.

5.3. Training in HPC

Training in High Performance Computing (HPC) extends the foundations laid out in SWC by introducing intermediate and advanced programming skills. These are used, for instance, in more complex algorithms, where calculations are carried out simultaneously using parallel computing in multiple cores. Training researchers in the use of parallel computing will upskill them in the more independent analysis of massive data-sets and, at the same time, will better exploit HPC facilities within ELIXIR. Training in HPC within EXCELERATE will commence from October 2016 onwards with a course organised by ELIXIR-ES. In addition, EXCELERATE collaborates in HPC training with the [PRACE Research Infrastructure](#). Collaboration is planned with several PRACE Advanced Training Centres (PATCs), and the first collaborative training course will take place at ELIXIR-FI in February 2017. Additionally there is opportunity to work with [BioExcel](#), an EU funded Centre of Excellence for support and training in HPC which will be developing training for biomolecular researchers.

5.4. Training for Use-Cases

Use-case training needs are gradually becoming clearer along the EXCELERATE plan, and the Training Platform is ensuring that these are captured in a timely fashion through, for instance, our liaisons and the TrCG. Some of the general training needs have been captured through the survey in 2016 (see following chapter), and other more specific needs have been recently captured by specific surveys sent to two use-cases, namely Marine Metagenomics (WP6) and Plants (WP7). In the future, these specific surveys will also be sent to researchers in the Rare Diseases (WP8) and Human Data (WP9) use-cases. In addition, several training events are being planned in collaborations between WP11, other EXCELERATE Work Packages and

the use-case communities. For instance, deliverable D8.3 describes in more detail the training plans for the users of WP8.

5.5. Industry

Industry users of ELIXIR services range from large multinationals to micro-SMEs, and cover fields including pharma, biotech, food and agriculture, and blue biotech. Given the number of research-intensive companies in Europe, and the increasing reliance within industrial R&D upon computational methods, stimulating innovation and supporting industry is a key priority for ELIXIR. In particular, supporting the bioinformatics training needs of industry is one of the key objectives defined in the [ELIXIR Industry Strategy](#) (WP13).

Two surveys have been undertaken to date, attempting to build a picture of the interests and training needs of wet-lab biologists and bioinformaticians in industry. These have helped guide the training strategy of ELIXIR UK and will be expanded to other Nodes in the future to capture a more comprehensive view. These surveys indicated a clear need for bioinformatics training, which is corroborated by similar results from GOBLET's survey (manuscript in preparation) and data reported by the Association of the British Pharmaceutical Industry (ABPI).

5.6. ELIXIR Training and Data management & stewardship

In the past, limited data-management expertise (“data management”) was needed in the life sciences; today, however, with data-intensive analyses, it has become essential. Also, external pressure is being applied to scientists in general to ensure that their experiments are reproducible, and the resulting data are accessible to, and reusable by, other scientists for a longer time (“data stewardship”). Data-management and stewardship plans are becoming prerequisites when applying for funds. ELIXIR has a key role in providing solutions, best practices and training in these areas, which have thus been prioritised for 2017 onwards. In particular, its training component will be built as a complement to the implementation study *Data Stewardship for Open Science - Training, Certification and Support Portal* that has been recently submitted for application to the ELIXIR Hub.

6. ELIXIR Training efforts for ELIXIR Developers

Training in bioinformatics frequently focuses on building researchers' scientific and technical expertise to exploit data and tools. However, there is a lack of organised training activities to support technical staff who develop and maintain the bioinformatics services used by those researchers. Improving developers' skills, and enabling easy knowledge exchange through this community, will certainly have a positive impact on the quality of tools provided by ELIXIR as a whole. ELIXIR developers and operators have thus been identified as a target audience for ELIXIR Training. For a start, the community has been gathered around a chat room where they can discuss their needs and exchange ideas. The first training workshops for developers will take place in September 2016 around “Behaviour-Driven Development”, an agile approach to software development, and in January 2017 around Galaxy for developers (already mentioned above).

7. ELIXIR Training efforts for ELIXIR Trainers

7.1. Train-the-Trainer

The expansion of training under ELIXIR relies on the development of new trainers. Across Europe, the need for courses outweighs the number of places available, notably in subjects such as NGS data analysis. The call for such courses has not yet abated, and as the life sciences expand into newer territories, yet more trainers will be needed.

Whilst the need for bioinformatics competencies (and hence training) is well recognised, the ability to provide such training is not yet well developed in all ELIXIR Nodes. New trainers are therefore required across both academia and industry, but a programme was needed that not only provides them with the tools and tips for providing an enriching learning experience to their trainees, but also provides guidance on course development and access to a wider support network, with a focus on ELIXIR-derived resources and infrastructure.

Also in the light of the enormous number of potential trainees, trainer capacity-building is crucial; as such, Train-the-Trainer (TtT) courses were indicated as an important need. Only through a proper TtT mechanism we can make a significant contribution towards meeting the high demand for training.

The EXCELERATE TtT programme started in 2016 by holding three pilot courses for trainers, and by building a seed set of training [materials](#) that will continue to be improved in the next six months. Other TtT courses are planned for forthcoming years. This project is carried out in collaboration with GOBLET, and was identified as a key area of collaboration between these two organisations.

Clouds, virtual machines and Docker are gaining momentum in bioinformatics. A knowledge-sharing workshop organised by ELIXIR Finland gathered more than 30 bioinformatics trainers and technical specialists from 13 different countries to learn and discuss the needs and possibilities of these technologies in bioinformatics training. A guideline document for trainers will be prepared, and the materials and [presentation videos](#) are publicly available.

7.2. ELIXIR Training Infrastructure

In order to effectively deliver training in a timely way, fulfil the needs of ELIXIR users, and coherently evaluate its impact within ELIXIR, the pillars of the ELIXIR Training infrastructure needed to be reinforced. A number of actions and tools have been put in place within EXCELERATE to support and strengthen this infrastructure.

TeSS Training Portal: Findability of courses

All of ELIXIR's training efforts will be wasted if they are not appropriately disseminated within ELIXIR and beyond. [TeSS](#) has been developed to ensure that training events, courses and materials, both from ELIXIR and from 3rd-party providers, are exposed both within the boundaries of national ELIXIR Nodes and internationally. TeSS is a centralised portal that aggregates ELIXIR-relevant training information across Europe, built to complement other platforms, such as On-Course and GOBLET's training portal.

TeSS started as an ELIXIR-UK pilot project in 2015, and has recently been released as [version 1.0](#). This release includes fixes and improvements to the user interface and support for small devices, full support for training workflows, support for linking

resources in TeSS with other external resources, as well as an increased number of content providers and training resources, harvested via custom-built scrapers. In the future, TeSS entries will be crosslinked to the Tools Registry, increasing its compliance with the [FAIR principles](#).

At the time of writing (August 2016), TeSS contains more than 3,000 training resources (events and materials) from 22 content providers. It has received strong buy-in from ELIXIR Nodes, who are enthusiastically using and sharing content via TeSS.

E-learning

Distance training and e-learning have become essential to reach large audiences spread over many countries, as is the case for the ELIXIR community. Their added values and benefits over “classical” training are evident and easy to reach with the development of ICT tools and services. The benefits of e-learning resources include that they are more suited to users’ needs and availability; more cost effective; provide a multiplication effect on the number of concurrent trainees; allow progress tracking; offer lower carbon footprints, *etc.*

Several Nodes have indicated, in the preliminary surveys, an interest in profiting from the new educational technologies, and are moving in the direction of exploiting distance and e-learning in their courses. This interest has been confirmed by their participation in the first e-learning related workshop, organised in Slovenia in September 2015. The results of this workshop are currently being gathered in a white paper, which will include an ELIXIR training strategy for e-learning.

In summary, e-learning will count on the extensive use of TeSS registry and GOBLET portal to provide reference points for e-learning resources available across the ELIXIR and GOBLET communities. It will be complemented with an ELIXIR dedicated e-learning platform (under development by ELIXIR-SI) that can be exploited by all Nodes in need of e-learning infrastructure. This platform will allow trainers to upload ELIXIR course materials and other course-related resources, and make them available for trainees. This platform has been tested by ELIXIR-FI (RNA-seq), ELIXIR-SE (Unix/Linux tutorial) and ELIXIR-IT (Python).

7.3. Measuring ELIXIR training quality and impact

The provision of training activities is a key outcome of ELIXIR as a whole, and it needs to be timely, current and impactful. To ensure that this objective is met, a robust method of assessment, monitoring and reviewing of ELIXIR courses needs to be adopted by all ELIXIR Nodes. It is thus very important to define a common approach across ELIXIR to allow the same core data-set to be collected in all cases, allowing easy comparison across ELIXIR courses and providing not only a benchmarking process, but also the ability (through the collection of appropriate metrics and key performance indicator (KPIs)) to determine the impact of training being delivered by ELIXIR.

This process started with a two-day workshop in January 2015, where ELIXIR trainers and TrC coordinators evaluated the methods, KPIs and metrics currently being used across the Nodes and its partners. As a result, an initial set of descriptors has been identified belonging to three main general categories: i) descriptors for training events, ii) quantitative metrics for individual training events; and iii) questions for short-term feedback. The full descriptors are listed in the workshop [report](#).

In the next few months, all ELIXIR training providers should adopt these indicators across ELIXIR Nodes, allowing us to capture a minimal core and consistent set of

information about ELIXIR training events. These indicators will also continue to be refined and expanded along the EXCELERATE project, until it achieves its optimum format reflecting the best mechanism to effectively capture and report on the quality and impact of ELIXIR training.

Preliminary information gathered in recent months regarding training audiences and needs has allowed creation of the first building blocks of the ELIXIR Training Platform. Bioinformatics is a moving target, and ELIXIR continues to grow, with new Nodes joining every year, together with new users, trainees and trainers. In order to have a more up-to-date, extensive and comprehensive view of the current training needs within ELIXIR, a broader and systematic survey has recently been sent out, which corresponds to the heart of Deliverable 11.2. The analysis of the outcomes is reported in the following section.

8. ELIXIR Training Survey 2016

8.1. Aims of the 2016 survey

Since the initial surveys in 2014-2015 (as described in “identified training gaps, need and solutions...” above), several new ELIXIR Nodes have been created, whose training activities needed to be captured. In addition, it was anticipated that training activities offered by pre-existing Nodes would have evolved over the intervening period.

The 2016 survey was originally designed to provide information on the planning and delivery of the train-the-trainer (TtT) subtask in the EXCELERATE Training Work Package WP11. More specifically, it was intended to capture information on three areas of expertise across ELIXIR Nodes:

1. Identify Nodes currently offering TtT provision: this information would be helpful in initiating an ELIXIR-wide programme.
2. Identify Nodes in greatest need of new trainers: this information would help to prioritise the initial sets of TtT courses.
3. Identify specific areas of bioinformatics in which the number of individuals able to provide training is currently low: this would assist us in focusing TtT courses on specific topics and to increase the number of trainers.

To maximise the potential of the survey, its remit was then extended to also capture the landscape of training provision within the Nodes:

4. Describe the general training offering across ELIXIR: providing a comprehensive and up to date view of the training landscape.
5. Review the training strengths of each Node: If a Node is offering training in a certain area, then one could assume that it is an area of strength for them. This information is additionally important for EXCELERATE Capacity Building (WP10).
6. Review the current training being offered and needed in relation to the four EXCELERATE use-cases (WP6-9): to identify those Nodes with pre-existing training expertise in these scientific areas so that the use-case projects can build upon, rather than reinvent, training activities.

8.2. The survey

This was designed and built using SurveyMonkey, which allowed both easy access by all Nodes and easier analysis, because all data were collected electronically. The survey consisted of 26 questions, split into six areas:

1. Organisation information
2. EXCELERATE use-case involvement
3. Experience, by target audience, of the organisation in providing training
4. Training courses 2015-2017 (ran, being run, would run), divided by bioinformatics subject areas
5. Training resources and capacity
6. Final comments

With regard to the target audience, we identified seven categories for each of three major types of institution (Academia, Industry, Healthcare/Clinical): Undergraduate students, Postgraduate students, PhD students, Postdoctoral researchers, Technical staff, and Core facility staff.

The themed list of bioinformatics subject areas in which training could be sought (and should therefore be on offer) was produced through discussion with the TrCG (Table 1 below). Each subject area was further split into more specific course topics to gain a more detailed picture of course offerings and gaps across all institutions and Nodes.

Table 1 Bioinformatics subject areas

Subject area
Chemical biology and Metabolomics
Computational science skills
Data integration
Health information
High-throughput sequencing data processing, analysis and interpretation
Proteins and proteomics
Resource management and sharing
Sequence bioinformatics
Statistics
Structural bioinformatics
Systems biology
Training and teaching
EXCELERATE use-cases

The survey was initially distributed via the Training co-ordinators group who were then asked to send it on to those relevant organisations within their Node who provided short training courses (**not** traditional, semestered educational courses), and to maximise the number of institutions/organisations responding to the survey.

8.3. Survey responses

The first step of the analysis of the survey responses is given below:

Organisations

Responses were received from 74 organisations across 16 ELIXIR Nodes. For six Nodes, only one institution provided responses; for others, numbers varied between two and 16 institutions (Figure 4).

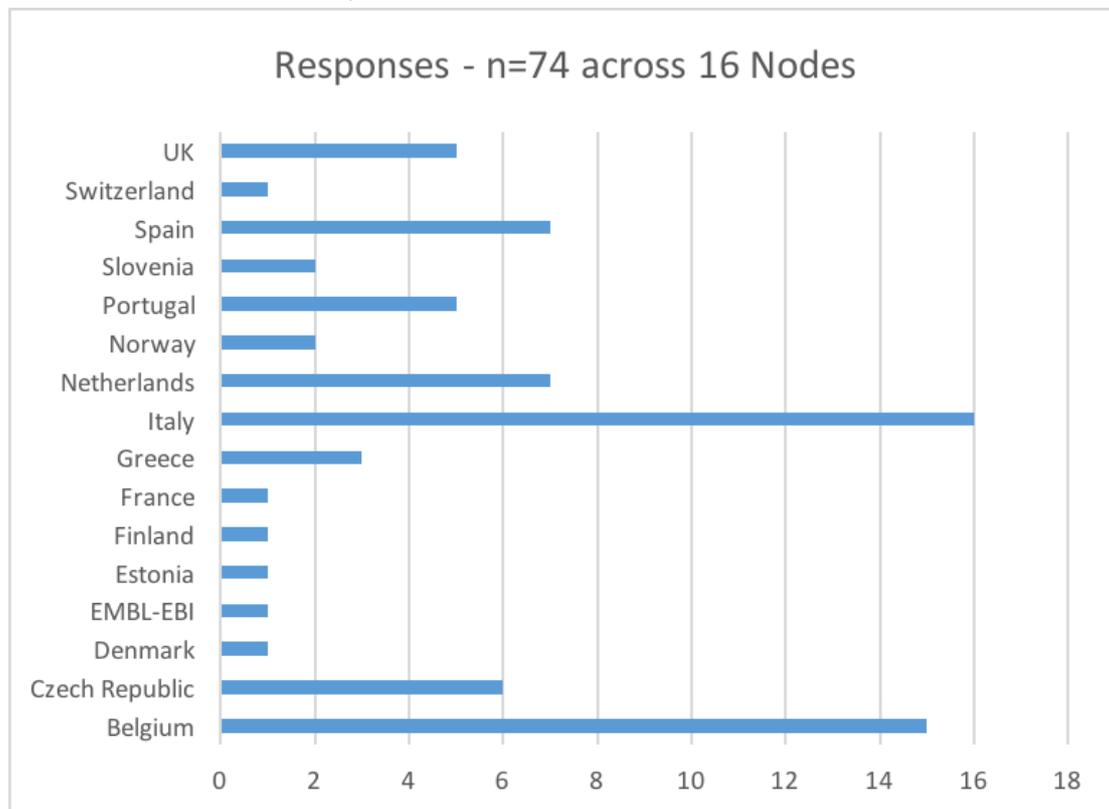


Figure 4 Numbers of responses across ELIXIR Nodes

8.4. ELIXIR members and EXCELERATE use-cases

Of those who responded, 12/74 did not consider themselves to be official members of ELIXIR.

Of the 62 ELIXIR members, 23 are involved in a use case, seven being involved in two or more (Figure 5a,b). There was representation from all four use cases for this survey.

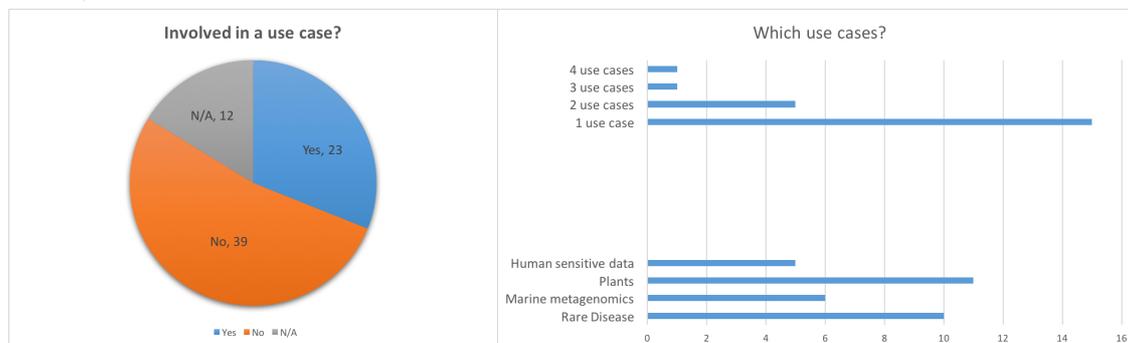


Figure 5 a,b Involvement of Node institutions/organisations in use cases

8.5. Experience in training and audience

Most institutions have pre-existing experience of training, the majority having offered training for more than five years (Figure 6 a,b). Therefore, we have representation within the Nodes from organisations that do have significant experience in offering relevant training.

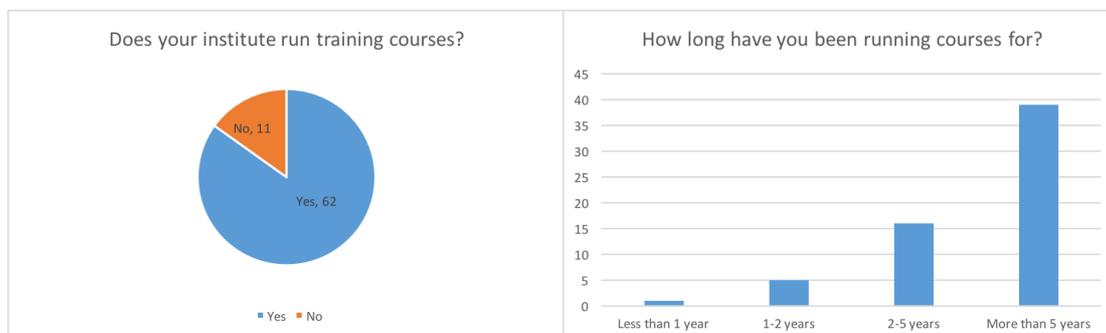


Figure 6 a,b Number of institutes across ELIXIR Nodes which run training courses; 8b: Period over which courses have been running

The major audience for these courses is the **academic community**, focusing on postgraduate students, PhD students and postdocs (Figure 7). This is likely to reflect the career stage with greatest demand for training. Nevertheless, we note an increasing interest in this area from the health and clinical communities, which represents an area for potential future growth: the number of institutions offering training aimed at industrial and clinical/health audiences is approximately half of those providing academic training. In all areas, the lowest number of training offerings is for principal investigators, technical staff and core facility staff. Some technical staff training will be provided under the ELIXIR-EXCELERATE project (“train-the-developer subtask”), although this is an area where ELIXIR’s links to projects such as [CORBEL](#) should be leveraged to ensure that these groups are able to access appropriate training.

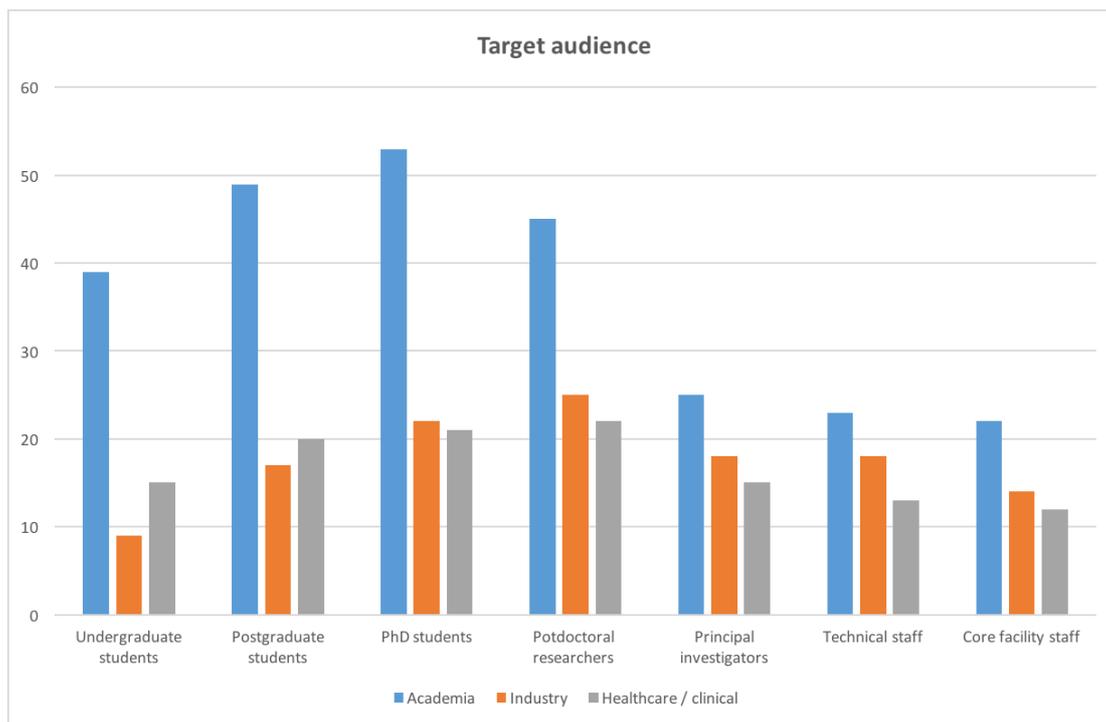


Figure 7 Overview of the target audience across three different communities (Academia, Industry, Healthcare/clinical)

8.6. Training courses

This was a major section of the survey and analysis to date (and presented below) has focused on Node level training offerings at the level of the major topics as shown in Table 1 (above).

Within this section, institutions were asked to state what training courses they:

7. Had offered in 2015
8. Were offering in 2016
9. Planned to offer in 2017
10. Would offer in 2016 / 2017 if they had appropriate resource
11. in a set of course subjects which were presented under the major bioinformatics topics from Table 1.

Unsurprisingly, the largest number of Nodes offer courses relating to high-throughput sequence analysis, with the least frequent offerings being metabolomics and chemical biology, and teaching & training (Figure 8). There is a core group of Nodes that provides training across the majority of these themes. It is difficult to gauge the degree of expertise available, however, as the number of organisations represented by a national Node offering courses within a topic can be quite varied.

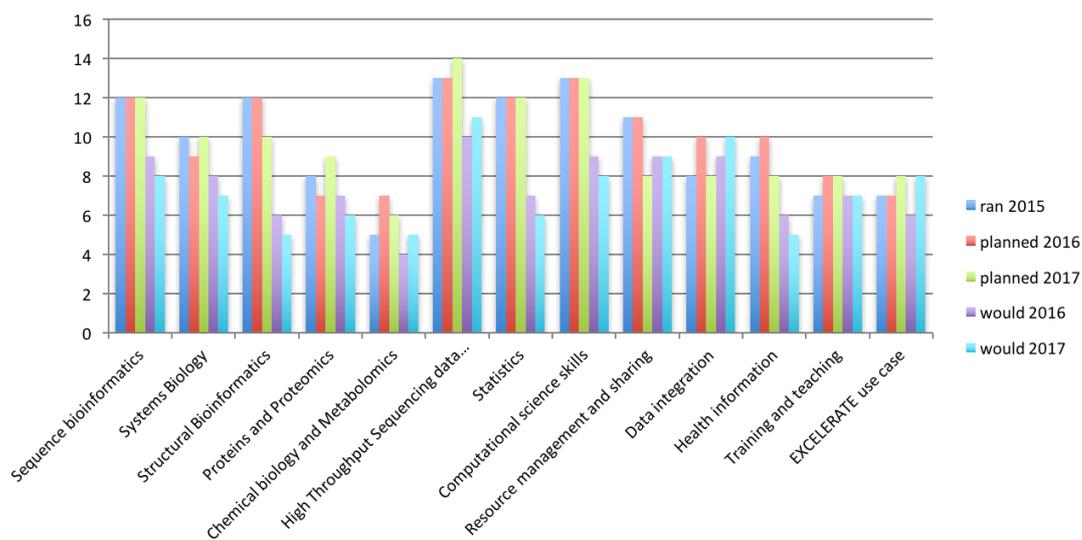


Figure 8 Number of Nodes offering courses in each major topic

A view on the numbers of courses available across these topics (Figure 9) can be taken, again highlighting the trends described above, but demonstrating in much more detail the low numbers of courses offered in some of these areas. There are some caveats to this data: in some topics there may be more course subjects on offer; and organisations completing the survey were not asked to provide actual numbers of courses on offer, rather to simply indicate a positive or negative response on whether a course subject was offered by them, so multiples of courses held within one year will not be accounted for.

It is interesting to note that a number of keys areas for EXCELERATE (e.g. use cases, training and teaching) and ELIXIR more widely (e.g. resource management and data sharing; data integration) are lacking in training offerings. Further, more detailed analysis of this data is required, and will be undertaken over the next few months in conjunction with the capacity building task and leaders of the use-cases.

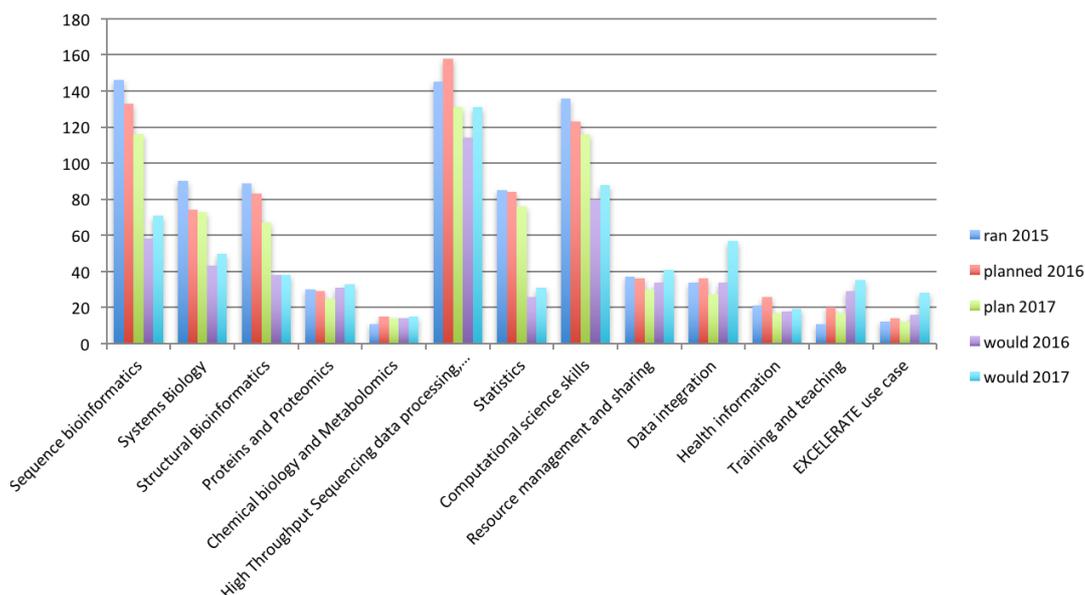


Figure 9 Number of courses offered by major topic

8.7. Training resources and capacity

When asked what resources organisations lack that would otherwise allow them to deliver courses they would like to run, people (trainers/organisers) and money were the two major resources identified (Figure 10).

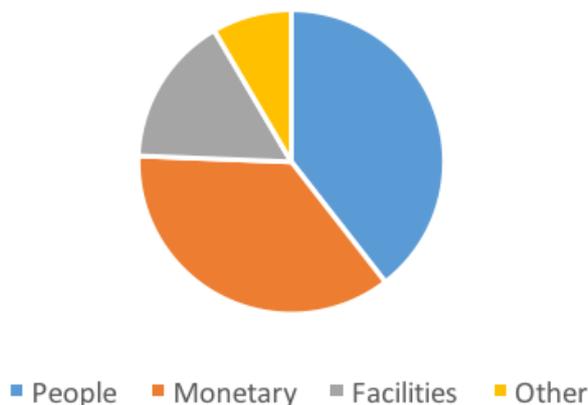


Figure 10 Resources required for running training courses. Most organisations have no major issues regarding training facilities (IT rooms, support, etc.) but lack trainers and monetary resources to run the courses they need. "Other" includes the following: more extended time planned for the courses; educational material; access to a better eLearning infrastructure; admin staff; requests

The deficit in human resources (trainers) is an area we are working to improve through the TtT activity within ELIXIR, to enable enthusiastic researchers who have experience in specific areas of analysis to deliver effective training to others. Increasing the number of instructors in Nodes, where they lack the training expertise but have the necessary subject expertise, is a significant move towards the delivery of more courses. Lack of money to support the development of new training courses is a potentially larger stumbling block, but is not a new issue. Although there are funding bodies who provide specific funds for training, such as EMBO and FEBS, they are few in number, and, over recent years, requests for funding in general have increased, yet the amount they are able to provide has not necessarily increased proportionately.

As for trainers, the numbers available in specific subject areas largely reflect the course offerings: where the number of courses offered is low, so too are the number of available trainers (Figure 11). In order to identify which subject areas are in greater need for new courses and trainers, we have to compare these results with those from the analysis on what courses Nodes would run if they had resources available to them. For prioritisation of effort, these need to be balanced with the use-case related training needs and offerings, to ensure appropriate provision of training in those areas identified as focal points for the ELIXIR community at this stage. One example highlighted through previous discussion in relation to the plant use-case is the need for genome annotation training (both within this use-case and more widely); there was a perceived lack of such training, a finding supported by this survey, but we are now making a step towards a solution in that we have now identified institutions that provide such training, who can be approached to share their training expertise with other Nodes.

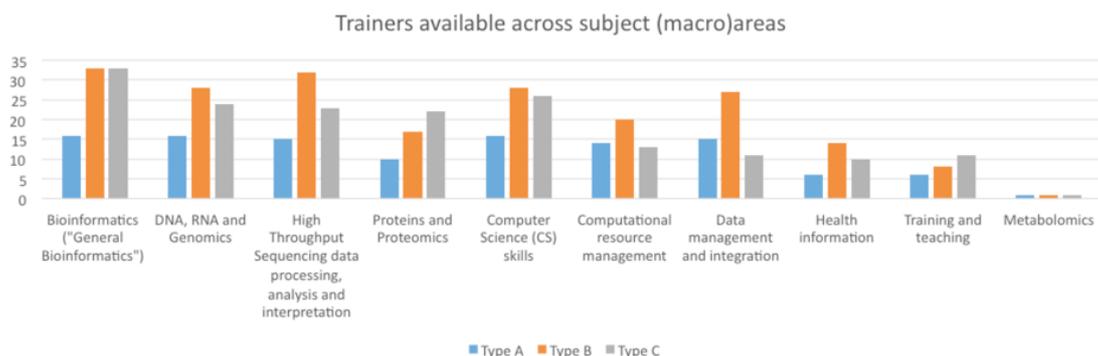


Figure 11 Trainer Type A: scientists with no previous training/teaching experience who would like to be involved in training; Trainer Type B: scientists with some (even little) training/teaching experience; Trainer Type C: "certified" trainers (e.g. People who have received formal TtT training or certified Software Carpentry Instructors) or trainers with extensive training/teaching experience

8.8. Next steps

This survey enabled us to gather a large amount of data both at general and very detailed levels. Here, we have reported a general initial overview of the results, but more detailed analyses will be required in order to maximise the outcomes of this survey as guidelines for future decisions (such as on where ELIXIR needs to focus TtT courses) and actions (such as the linkage of the training needs of specific Nodes to training capacity available in others).

In particular, here is a description of our plans following this report.

Use cases

- We observed low numbers of courses currently being offered in relation to the use cases (through direct question and when considering the relation of specific course topics to these). To ensure that effective training for the use cases is delivered we need to cross reference (i) who is offering courses with those who are part of the EXCELERATE use cases, and (ii) where there are other providers, bring them into discussions as training needs become clearer.
- We need to clearly identify Nodes that have indicated there is need for courses in given subject areas but due to the lack of people resource cannot provide these courses. Subsequently we would like to link them with Nodes that already have the trainers and expertise and prioritise these Nodes and subject areas for TtT courses.

Training in major topics

- There is great variety in the "amount" of training offered across the subject areas, but given the nature of trends in research this is to be expected. This survey does however provide information when planning priorities for training development (whether for developer, researcher or trainer) at two levels:
 - It provides an ELIXIR level view, where the programme can focus on areas viewed as priority at the level of the whole project.

- It provides a Node level view which Nodes can use to focus on their specific areas of need, which may not match those of ELIXIR as a whole.

Numbers of trainers

- We observed that there are low numbers of trainers in areas relating to infrastructure development e.g. computational resources, data management, training and teaching, and extremely low in metabolomics.
- We will identify Nodes more in need of trainers in these subject areas and will discuss with them how the EXCELERATE TtT programme can help to increase their numbers of qualified trainers.
- As for other areas, we may explore the possibility to work with e.g. CORBEL to help increase numbers of trainers.

Definition of ELIXIR membership and branding

It is clear from the survey that not all institutes completing the survey are (or consider themselves to be) an official member of ELIXIR. Defining ELIXIR membership and ELIXIR branding are issues that extend beyond the ELIXIR Training platform but may have implications for the ease of accessing the training offerings we have identified. They will have to be resolved on Hub level. As stated previously, within WP11 we currently have a working definition of ELIXIR Training and ELIXIR node training, which will continue to evolve as ELIXIR's growth continues.

Course advertising

Though not commented on above, on behalf of the hub we asked where organisations advertised courses. Highest on the list was institutional websites, followed by emails / mailing lists.

- It is clear from the responses that there needs to be a focused encouragement of advertising via TeSS as the number of organisations using this mode of advertising was low (6/74). The number of organisations using TeSS as a place to advertise events is steadily increasing and its most recent release (Summer 2016) has seen an increase in the number of content providers involved.
- One mode of advertising that was popular (18/74 responses) was [On-Course](#), a very rich (manually filled) course catalogue. On-Course is due to lose its funding when the parent project (EMTRAIN) ends in September. Access to training information to trainees might diminish if On-Course cannot be updated and maintained any longer.

Key outcomes of the ELIXIR Training Survey 2016

Aim:

- Inventory of training landscape and training provision in ELIXIR nodes
- Quantify trainer capacity in ELIXIR nodes
- Identify specific areas of bioinformatics where training provision is low/needed, with special attention for the use cases

First results:

- 74 organisations (in 16 ELIXIR nodes) responded, the majority (>40) of which have > 5 years of training experience
- Major audience for the courses: academic community
- Detailed analysis of offerings, plans and trainer capacity per node for all the defined bioinformatics subject areas (Figures 10 and 11)
- Trainer capacity and lack of monetary resource are highlighted as major obstacles in organising courses

Next steps, detailed analysis (fall 2016):

- use-case courses: more clearly define training needs for each and cross-reference with nodes that provide such training, also provide support for specific nodes that have indicated a use case need
- Node level analysis of training needs: to help the development at a node level of appropriate training
- tailor ELIXIR TtT programme to dedicated trainer needs in the nodes.

9. Collaborations with other Training initiatives and Programmes

In order to fulfil ELIXIR's training challenges, and to ensure embedding in the global training landscape, ELIXIR Training is actively reaching out to several organisations. Two of them, the SWC/DC Foundations and the Galaxy Foundation have already been mentioned above. In addition, several other collaborations are ongoing and, for each of those, a summary of the current status of collaborative efforts is given below.



The mission of [GOBLET](#) (Global Organisation for Bioinformatics Learning, Education and Training) is to provide a global, sustainable support and networking infrastructure for bioinformatics trainers and trainees ([Attwood et al., 2015](#)). This includes a training portal

([Corpas et al., 2015](#)), allowing trainers to share materials, tools and techniques, guidelines and best-practice documents, and resources to help train trainers and teachers. In addition, GOBLET is: i) fostering the international community of bioinformatics/computational biology trainers through networking events (particularly in collaboration with ISCB, through joint workshops and tutorials at ISMB/ECCB, and joint development of the Community of Special Interest (COSI) in Computational Biology Education), ii) facilitating bioinformatics capacity development across the globe, particularly through its train-the-trainer and train-the-teacher initiatives, and iii) developing best-practice standards and guidelines for bioinformatics training ([Via et al., 2013](#)).

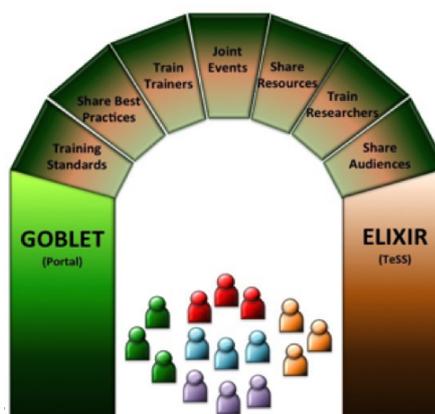
GOBLET currently contains around 40 members, including representatives from organisations in Europe, Africa, Asia, Australasia, and North and South America, together with individual members. Many of the ELIXIR Training Coordinators and EXCELERATE subtask leaders are the GOBLET representative for their institutes (CH, FI, PT, IT, NL, UK).

As GOBLET is positioned to provide an interface between the ELIXIR and world-wide bioinformatics training communities, this close collaboration ensures many, if not all, training efforts can be shared, thus avoiding duplication of cost and effort.

Furthermore, GOBLET can help to connect ELIXIR with a range of international organisations, allowing outreach to communities beyond ELIXIR's member countries.

In Spring 2015, GOBLET and ELIXIR formed a [Joint Training Strategy](#), which sets out a number of areas for collaboration:

- ELIXIR's training portal, [TeSS](#), and [GOBLET's training portal](#);
- train-the-trainer and train-the-researcher activities (amongst the first collaborative events that took place in 2015 and 2016 were workshops dedicated to e-learning in Slovenia and [South Africa](#) and a [metagenomics materials hackathon](#) in the UK);
- Joint exploration of training 'accreditation' mechanisms; and
- Sharing of best practice and expertise on professionalising bioinformatics training (including, for example, work on course descriptors and on metrics and evaluation mechanisms).



The NIH's Big Data to Knowledge (BD2K) Training Consortium seeks to empower current and future generations of researchers with a comprehensive understanding of the data-science ecosystem: the ability to explore, prepare, analyse, visualise and interpret Big Data. The TCC helps to coordinate the diverse activities occurring within the BD2K Training Consortium into a synergistic training effort. The TCC and ELIXIR share many of the same goals and, therefore, forming a partnership to share educational materials and standards is naturally mutually beneficial for our respective efforts.



TCC
BD2K Training
Coordinating Center

Therefore, in spring 2016, BD2K TCC and ELIXIR training formed a [joint training collaboration](#) covering the following general aspects:

- Training portals: jointly defining metadata standards for easy and automated communication between BigDataU and TeSS, the respective training portals; engage content providers; using machine-learning/statistical techniques; collaborate on training workflows and sharing code;
- Collaborate on the ISCB competences;
- Summer School in collaboration with RDA-CODATA;
- Collaboration with GOBLET;
- International Interactions and Frameworks for "Big Data" Training Standards.



EDISON
building the data
science profession

[EDISON](#) is a 2-year project (started September 2015) with the purpose of accelerating the creation of the Data Science (DS) profession. Major activities include

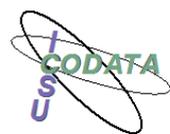
the building of a Competence Framework (CF-DS) and development of Model Curricula (MC).

ELIXIR is EDISON's partner for the life science domain. ELIXIR has a formal representative (Allegra Via, TrC of ELIXIR-IT) in the Expert Liaison Group for Data, and, in summer 2016, an EDISON-ELIXIR working group has been formed to establish a closer relationship between both organisations. The purpose of this collaboration is to capture life science and bioinformatics needs in terms of basic and advanced training on digital data management (also termed Data Science): amongst others, in identifying and validating domain-specific needs and professional profiles, and to validate the tools and knowledge EDISON can provide for the benefit of the life-science community.

The [RDA-CODATA Working Group on Research Data Science](#) is addressing the worldwide need for individuals with the combination of skills necessary to optimise use of high-throughput data sets; to this end, the Working Group has built a Introductory Research Data Science Curriculum, containing subjects such as Open Science, Open Research Data, SWC/DC, Visualisation,



Analysis and Computational Infrastructures. This curriculum is being rolled out in their Data Science Summer Schools, the first of which took place in Trieste (1-12 August, 2016).



ELIXIR Training has been liaising with this Working Group since Spring 2015. Taking the lead, RDA-CODATA (with input from H3Africa, GOBLET and ELIXIR) successfully applied for funding the 2017 edition of the Summer school, which will include bioinformatics flavoured modules. We are actively exploring further joint funding opportunities.



In line with ELIXIR efforts towards collaborating with other biological and medical research infrastructures, the ELIXIR Training Platform is actively reaching out to, for example, CORBEL. CORBEL will set up a training programme for Research Infrastructure (RI) operators and key users to drive rapid implementation of shared services into operation.

The RITrain project is aimed at the managers and service operators of life-science RIs, with a particular emphasis on distributed research infrastructures. The project will develop a training programme enabling RIs across all domains to gain expertise on governance, organisation, financial and staff management, funding, IP, service provision and outreach in an international context. The ELIXIR Training Platform is represented in both projects by EMBL-EBI.



10. Conclusions

The preliminary needs analysis made in the previous years have clearly indicated a need for training provisions in Software or Data Carpentry, High-Performance Computing, Workflow Tools, Data management and stewardship, and Train-the-Trainers. Thanks to a very pro-active, enthusiastic, and productive ELIXIR training community, led by the TrCG, many courses have already been done in the past 12 months to fulfil these areas and many more are planned for the years ahead (annexes II and III). Several initiatives and collaborations have already started and they will contribute to improve quality and impact of ELIXIR training.

The latest survey (described in this deliverable) has gathered a large amount of data both at general and very detailed levels. The results of its primary analysis confirm a need for prioritising training for the use cases, identified previously, and to improve the training capacity in the Nodes to deliver this particular training. For the latter, we project to build a “Bridging programme” where Nodes, which have the human resources and expertise, coach Nodes with specific identified need. The more precise training needs of the use case communities have started to be defined with this survey and will be clarified with more specific future surveys, allowing to timely provide the specific training when these projects will be in full speed (from PY2 and beyond). Besides, the TrCG and the training liaisons will play a major role in bridging between all involved Work Packages (both the use cases and other WPs).

Prioritising the training topics, from an extensive list of ELIXIR resources and data driven topics, is an issue for ELIXIR Training and is very much dependent on the advances of the other EXCELERATE WPs and ELIXIR Platforms as well. EXCELERATE has closed its first development year, and only now training topics are emerging. Planning training courses depends on the advancements of ELIXIR projects and needs to be aligned and timely provided. Besides, ELIXIR Training’s Portfolio and Nodes Training’s Portfolio walk hand in hand. They complement each other and their strategies need to be well aligned.

Other challenges that ELIXIR Training is facing are to devise the appropriate and impactful ELIXIR training strategy for the Industry and to find a sustainable funding model for training beyond EXCELERATE. We will also need to ensure an improved visibility and discoverability of ELIXIR training courses, in particular by promoting the use of TeSS by trainers and trainees. These challenges will be tackled from PY2 and beyond.

Planned Training workshops funded or co-funded by ELIXIR-EXCELERATE

From Sep 2016

WP11 Sub-task #	WP11 Sub-task names
11.1.1	Impact & Quality
11.1.2	TeSS
11.1.3	elearning
11.2.1	Train the Developer
11.2.2	Train the Researcher
11.2.3	Train the Trainer

Event title	Funding resources	Dates	Location
Metagenomics training	ELIXIE-BE	15 Sep 2016	Brussels, BE
Workshop for developers: Software faster: From months to minutes	EBI	20 Sep 2016	TBC
elearning course (synchronous): linux/unix	WP11.1.3	Oct/Nov 2016	TBC
TeSS workshop with trainers, content providers and developers	ELIXIR-UK & WP11.1.2	Autumn 2016	TBC
Train the trainer event	WP11.2.3	Oct 2016	Rome, IT
Train the trainer event	WP11.2.3	Nov 2016	Slovenia
Data Carpentry Workshop	ELIXIE-BE	7-8 Nov 2016	Brussels, BE
Train the trainer event	WP11.2.3	Nov/Dec 2016	TBC
Workshop related to elearning subtask	WP11.1.3	End 2016/Early 2017	TBC
HPC Workshop	WP11.2.2	Dec 2016/Jan 2017	Malaga, SP
Galaxy Workshop for developers	WP11.2.2	16-20 Jan 2017	Strasbourg, FR
Train the trainer event	WP11.2.3	Jan 2017	TBC
Train the trainer event	WP11.2.3	Feb 2017	TBC
elearning course (synchronous) Python	WP11.1.3	Early 2017	TBC
Metabolomics Workshop	Elixir funding & Elixir-FR	Jun 2017	Paris, FR
Clinical metagenomics	Elixir funding & Elixir-FR	Sep 2017	Paris, FR
Data Curation Training for tranSMART	ELIXIR-LU & IMI-eTRIKS	September 2016	Luxembourg
SWC workshop	ELIXIR-UK	Q4 2016/Q1 2017	Oxford, UK
DC workshop	ELIXIR-UK	Q4 2016/Q1 2017	Bradford, UK
SWC workshop	ELIXIR-UK	Q4 2016/Q1 2017	Nottingham, U

Galaxy Workshop	WP11.2.2 & co-funding	Q2/Q3 2017	The Netherland
Train the trainer event	WP11.2.3	Q4/2017	Switzerland
HPC workshop	WP11.2.2	Q4/2017	Switzerland
Train the Developer Workshop related to Data Stewardship	WP11.2.1	2017	TBC
SC/DC workshop	ELIXIR-SI, WP11.2.2 & SC/DC	2017	Slovenia

**Training workshops funded or co-funded by ELIXIR-EXCELERATE
Sep 2015 - Aug 2016**

WP11 Sub-task #	WP11 Sub-task names
11.1.1	Impact & Quality
11.1.2	TeSS
11.1.3	elearning
11.2.1	Train the Developer
11.2.2	Train the Researcher
11.2.3	Train the Trainer

Event title	Funding resources	Dates	Location
ELIXIR/GOBLET workshop: defining an e-learning lingua franca	WP11.1.3 & ELIXIR Hub	15-17 Sep 2015	Ljubljana, SI
RNA-seq data analysis with Chipster (Elearning course)	ELIXIR-SI	18-Sep-15	Ljubljana, SI
TeSS workshop with training standards group BioSchema	ELIXIR-UK	1-Oct-15	Birmingham, UK
Data Carpentry Workshop	ELIXIR-BE	2-3 Nov 2015	Brussels, BE
Software and Data Carpentry's "Instructor Training" Workshop	ELIXIR - UK	23-24 Nov 2015	Manchester, UK
ELIXIR Data and Software Carpentry Instructor Training	ELIXIR SC/DC Pilot	13-14 Jan 2016	Lausanne, CH
Excelerate Train the trainer and Training Impact workshop	WP11; 11.1.1 & 11.2.3	19-21 Jan 2016	Hinxton, UK
RNA-seq data analysis with Chipster (Elearning course)	WP11.1.3	Feb-16	Prague, CZ
"Training" and "metadata searching" satellite meetings @ELIXIR All Hand	ELIXIR-UK	Mar 2016 (All hands)	Barcelona, SP
ELIXIR and DARIAH "AAI Workshop for service and resource providers"	ELIXIR-UK	15-16 Mar 2016	Manchester, UK
ELIXIR/GOBLET hackathon for metagenomic training material re-use	ELIXIR-UK	7-8 Apr 2016	Hinxton, UK
Hackathon Bioinformatique	WP 1.3	24-23 Mar 2016	Paris, FR
Train the trainer event	WP11.2.3	10-11 May 2016	Cambridge, UK
Using clouds and virtual machines in bioinformatics training	WP11.2.3	23-25 May 2016	Espoo, FI
BIIT web-tools for high-throughput data analysis from ELIXIR-Estonia	ELIXIR-EE	12 May	Cambridge, UK
ELIXIR-UK: Data Carpentry	ELIXIR-UK	16-17 May 2016	Cambridge, UK
Elixir Technical Hackathon : Tools, Workflow and Workbenches	WP 1.3	18-20 May 2016	Paris, FR
ELIXIR-UK: Statistics Course	ELIXIR-UK	26-27 May 2016	Cambridge, UK
RNA-seq and CHIP-seq data analysis with Chipster	ELIXIR-CZ	1-2 Jun 2016	Prague, CZ
Variant analysis workshop	WP11.2.2	13-15 Jun 2016	Espoo, FI
Train the trainer event	WP11.2.3	10-15 Jul 2016	Oeiras, PT