# Defining a *lingua franca* for the ELIXIR/GOBLET e-learning ecosystem

Teresa K. Attwood, University of Manchester, United Kingdom
Brane L. Leskosek, University of Ljubljana, Faculty of Medicine, Slovenia
Jure Dimec, University of Ljubljana, Faculty of Medicine, Slovenia
Sarah Morgan, EMBL-European Bioinformatics Institute
Nicola Mulder, University of Cape Town, South Africa
Celia W.G. van Gelder, Dutch Techcentre for Life Sciences, The Netherlands
Patricia M. Palagi, SIB Swiss Institute of Bioinformatics, Switzerland

## Background

Today, the term **e-learning**[1] is widely used in a variety of contexts. Many of us are familiar with concepts like **Massive Open Online Courses** (MOOCs), such as those provided by Coursera, Udacity, edX or MIT OpenCourseWare; **Content**- or **Learning-Management Systems** (CMSs, LMSs), like Moodle or Blackboard; **Virtual Learning Environments** (VLEs), like EMBER; problem and tutorial portals, such as Rosalind and Train online; repositories for uploading and hosting **Educational Resources** (ERs), such as that provided by GOBLET; **ER aggregators** that harvest and disseminate training information, such as TeSS; and so on. Resources like these are seamlessly discussed under the umbrella term 'e-learning'; however, they are heterogeneous and do not mean the same thing from user, developer, trainer or content-provider perspectives. Confusion can therefore arise if those using the term generically have not grasped each other's specific meanings.

In the context of ELIXIR, this situation came to the fore in preparations for the EXCELERATE project, whose training programme (WP11) includes an explicit e-learning subtask (11.1.3) as part of the commitment to build an ELIXIR training infrastructure. The ultimate deliverable of the subtask is to develop an ELIXIR e-learning policy and deploy e-learning platform(s) at the ELIXIR level. But what does an '*e-learning platform for ELIXIR*' mean? Is it an instance of a **MOOC platform**, a customised LMS, a bespoke VLE, an ER repository, an aggregator, or something else? Clarification of this question was important, because each of the possible answers has very different implementation and resource requirements.

To try to address this question, and to facilitate communication both within ELIXIR, and between the ELIXIR and GOBLET trainer communities, an initial workshop – *Defining an e-learning lingua franca* – was held in Ljubljana (SI), 15-17 September 2015. At one level, the aim was to develop an overview of some of the e-learning approaches and applications developed or used by representatives from ELIXIR, GOBLET and other organisations – participants were therefore asked briefly to describe the systems they'd developed, to allow them, and their respective strengths and weaknesses, to be compared. At another level, the aim of the workshop was to try to converge on a common e-learning 'language', to help the ELIXIR and GOBLET trainer communities to communicate more coherently. To broaden the picture and gain a wider understanding of the challenges, a follow-up event was held during the GOBLET Annual General Meeting in Cape Town (ZA), 18-20 November 2015. Here, the outcomes of the first meeting were further discussed and refined.

Gaining a clearer view of the current e-learning landscape, and consensus on what we collectively mean by e-learning, were the necessary first steps towards formulating appropriate e-learning strategies for ELIXIR and GOBLET, and suggesting solutions that can feasibly be implemented. In the event, the issues were more deep-rooted than we

[1] Terms rendered in bold throughout this document are defined in the accompanying glossary.

expected, and many additional discussions were needed. This paper reflects the results of those discussions and of the deliberations of the workshop participants (whom are listed at the bottom of this document).

## Glimpsing a corner of the e-learning ecosystem

Realistically, two short workshops couldn't meaningfully review the entirety of the global bioinformatics e-learning ecosystem. Nevertheless, members of ELIXIR and GOBLET have been involved in developing a wide range of e-learning resources, and have hence gained valuable expertise and perspectives during the last 10-15 years. Presentations were therefore invited from this group, and balanced with talks from a number of practitioners from outside the ELIXIR/GOBLET communities. We were particularly interested to learn about the terminologies used by the presenters to describe their e-learning resources, the barriers to implementation or adoption they'd encountered, the standards they'd used, and any 'take home' lessons they could pass on from their various projects. A range of systems, with different levels of sophistication, aimed at different target audiences was presented, and the experiences of developing or using them are outlined in the sections below.

### *Lessons learned from MOOCs hosted on Coursera - Phillip Compeau*

Rosalind (http://rosalind.info) is a **problem-based learning** resource that promotes the development of bioinformatics programming skills via active problem solving (*e.g.*, counting DNA nucleotides, finding a motif in DNA, *etc.*). Rosalind problems (~280 at the time of writing) grow in biological and computational complexity. Scripts submitted with problem solutions are checked automatically, so the only resource required is an Internet connection. The problems are proposed by the project participants and by volunteers. In recent years, more than 38,000 users have attempted to solve at least one problem, and there have been almost 379,000 correct submissions. Rosalind has been used in classrooms by 90 different instructors more than 180 times. However, the platform wasn't created to be a standalone ER, but rather, as a self-paced teaching aid.

In 2013, Compeau and Pevzner created their first Coursera MOOC on Bioinformatics Algorithms. MOOCs provide traditional course materials, such as videos of lectures, reading materials, and problems to solve, with assessment of learning outcomes via quizzes and peer-to-peer evaluations. MOOCs also provide interactive fora to support interactions among students, professors and teaching assistants.

In 2015, this MOOC was extended to a full specialisation on Bioinformatics Algorithms, with seven modules: *Finding Hidden Messages in DNA*; *Genome Sequencing*; *Comparing Genes, Proteins and Genomes*; *Deciphering Molecular Evolution*; *Genomic Data Science and Clustering*; *Finding Mutations in DNA and Proteins*; and a capstone project, *Big Data in Biology* (sponsored by Illumina). Technically, MOOCs are quite demanding (a single module requires a minimum of 100 hours of work), the videos are not easily editable, and the interactivity between instructors and trainees is not optimal. Rosalind and the MOOCs began with funding and manpower from the University of California at San Diego, Saint Petersburg Academic University, the Howard Hughes Medical Institute, and a Russian Megagrant Award, with many professors and assistants providing content. An NIH grant is funding the 2015-2018 period. Both Rosalind and MOOCs reach thousands of users worldwide.

The Massive Adaptive Interactive Text (MAIT) concept arose to fill the gaps in the Rosalind platform and to overcome some of the issues encountered with MOOCs. MOOCs require massive development resources. **Active learning** and **adaptive learning** have been shown to increase student performance relative to **traditional instruction** (Freeman *et al.,* 2014). In active learning, assessments are integrated into the courses when they're needed, via stop-and-think questions, exercise breaks, code challenges, final data challenges, and so on. In adaptive learning, the courses are broken down into pieces, and remedial modules are inserted to overcome issues identified by students. The interactive CMS exploits research in intelligent tutoring systems and new educational projects that aim to improve student-student interaction. The textbook, *Bioinformatics Algorithms: an Active Learning Approach,* emerged both from the content of the Bioinformatics specialisation on Coursera and as a companion to this MOOC. Today, Coursera courses are powered by a fully interactive version of this textbook, with remedial modules, videos, slides and an e-book. It is a fully-fledged e-learning resource.

### *Metrics and quality assessment of e-learning methods and tools - Anthony Camilleri*

Various technological trends have had a direct influence on e-learning. These include:

1) ubiquitous computing, which brings access to processing power, anytime, anywhere;
2) open data movement, which brings any information anytime, anywhere;
3) learning analytics, which makes it possible to base teaching decisions on real data;
4) collaboration technologies, which make it possible to collaborate with anybody in real-time; and
5) personalisation technologies, which make possible educational solutions that are customisable to the individual.

Technological trends strongly influence social changes, and have placed increased demands on education programmes: to reach more students with increased efficiency, to be highly adaptable in a fast-changing society, ultimately, to provide more graduates - in short, to do more and better, but with less. In trying to cope with these pressures, the focus of educational professionals is often placed rather more on content management than on learning design. The ease of establishing a Web-based 'university' with just a handful of programmers has led to an exponential rise in the number of 'diploma mills', unconcerned by quality, coupled with an uncritical digital evangelism rather than pedagogical soundness. The popularisation of MOOCs is an example of such digital evangelism.

Higher education has reacted to these societal trends in various ways, particularly in terms of how they affect quality and Quality Assurance (QA):

- There has been a global rise in new e-learning quality labels, but a dearth of appropriate quality models. There are many models, but most are based on the properties of educational activities: quality of services (expressed as staff support and student support), quality of products (*e.g.*, curricula, course design, course delivery), and management. Quality models address:
    a) certification (certifying an e-learning course against a particular set of criteria or standards),
    b) benchmarking (comparing e-learning courses to other courses),
    c) accreditation (legal recognition of e-learning courses), and
    d) recommendation/advice (e-learning course design guides).
- There has been a rise in learning analytics in an attempt to improve quality. Learning analytics are methods that assess data collected by LMSs: *e.g.*, A/B testing, resource usage statistics, grading data, performance and student record monitoring. These provide information on how students interact with materials, peers and tutors, which materials and methods work best, and so on.

The assumption here is that by using real course- and teaching-related data, we can (in principle) assess the quality of teaching materials and teaching experts.

- There has been a significant and swift growth in the number of **Open Educational Resources** (OER). In April 2014, there were estimated to be more than 3,000 learning-material repositories, containing 12 million OERs, and MOOCs are growing at a rate of ~220% per year (in March 2015, there were 1,139 MOOCs in Europe).
- There has been increasing demand for recognition of course results and portability. But how is the quality of OERs and of their authors to be assessed (*e.g*., by comparison with classical publishing and citation metrics)? And how can the reuse of OERs by others be translated into a suitable quality metric?

Many recognition instruments exist. In **traditional education**, ECTS (European Credit Transfer System) points can be used for diplomas and degrees. Although widely used in higher education, however, the system is not well adapted to e-learning methods; hence recognition instruments have been developed specifically for these environments: *e.g*., certificates of completion, certificates of attendance, and badges (which are designed to reflect performance in specific learning activities, and are thus useful within courses). Students expect the same level of recognition in the e-learning world as with ECTS and classical diplomas. In terms of recognition, portability is key - students need be able to use their certificates in different situations (with employers, at different educational institutions, and so on). This is not yet a solved problem, and the quality of recognition instruments clearly warrants further discussion.

### *H3Africa eLearning Strategy - Victoria P. Nembaware*

Until recently, bioinformatics training in Africa was limited. New initiatives and training networks have changed this, but there is still not enough specialised training available to address the needs on the continent.

H3ABioNet (Mulder *et al*., 2015) – part of the Human Heredity and Health in Africa (H3Africa) consortium – is a pan-African bioinformatics network that aims to develop bioinformatics capacity in Africa. H3Africa is developing an 'eGenomics catalogue', the bioinformatics component of which is being developed by H3ABioNet (http://egenomics.h3abionet.org/). Here, links to existing Web-based courses are being collected, then annotated and sorted using EDAM ontology terms (Ison *et al*., 2013). A review system allows course participants to leave comments, using guidelines to try to achieve review consistency. The catalogue also provides links to other resources, such as books, webinars, journals and so on. Further, when H3ABioNet runs courses, the lectures are recorded and made available to the rest of the consortium, along with additional course materials. Where there is high demand, courses are live-streamed to multiple remote classrooms using the video system. Teaching assistants are placed in the distant rooms, and participants can address questions directly to the primary trainer. This has worked both for a 3-week course run in one location and live-streamed to two other countries, and for a trainer who wasn't able to travel, and who therefore taught remotely.

### *Lessons from GOBLET, TeSS, EMBER and A taste of bioinformatics - Teresa K. Attwood*

*1)* GOBLET's training portal (http://www.mygoblet.org) provides a repository to which trainers around the world may upload content, such as presentations, exercises, tutorials, *etc.* (Corpas *et al.*, 2015). The portal includes an announcements section (exploiting an iAnn plugin (Jimenez *et al*., 2013)) for publicising bioinformatics training events; a directory of trainers and their profiles; and course pages, via which trainers may advertise their courses, and with which they may associate ERs held in the repository, or to which they may provide links to ERs held elsewhere.

The GOBLET portal is a *global gateway* to and *repository* of training information, developed by, and primarily for, trainers. Although it provides access to ERs, it is not a structured **Distance Learning** (DL) tool, nor an LMS, VLE or MOOC platform (although, arguably, it has some CMS functionality). The portal is used by trainers and course organisers to store ERs, and to advertise up-coming events, and by trainees to discover them. Originally, content uploaded to the portal was not required to meet specific standards. However, GOBLET's Standards- and Learning, Education and Training (LET) Committees are working closely with ELIXIR and others to outline minimal criteria to which, in future, ERs should adhere in order to be considered 'GOBLET compliant'. This work is advancing as part of the Bioschemas initiative (http://bioschemas.org), which is exploiting schema.org semantic mark-up to describe events and ERs consistently in order to make them more discoverable.

By contrast with fully-fledged e-learning tools, the portal was not hugely expensive to create: it was originally developed in Drupal by volunteers and supported for a year by a full-time intern (funded by GOBLET subscriptions). As with all Web-based resources, however, it continues to have maintenance overheads. It is currently kept running on a volunteer basis; to fully evolve its functionality will require further dedicated funds.

The main lessons from building the GOBLET portal include the need to perform focused requirements-gathering and to procure *dedicated funds* in order to be able to deliver what users want: *i.e.*, to understand consumers' needs, and to deliver what they want to a professional standard.

*2)* ELIXIR's Training e-Support System, TeSS (https://tess.elixir-uk.org) is a platform for browsing and discovering ERs. Its primary aim is to give a snapshot of the ELIXIR training landscape, automatically aggregating data from ELIXIR Nodes, and surfacing information in ways that support user decisions and choices via training packages and workflows; it also harvests information from selected 3rd-party providers, like Software and Data Carpentry, Coursera, GOBLET, *etc.* A significant difference between TeSS and the GOBLET portal is that TeSS doesn't store the ERs it finds, but rather, provides *links* to them via their metadata.

By contrast with the GOBLET portal, the TeSS platform primarily offers a *European gateway* to and *registry* of training information. Although it provides access to ERs, it is not a structured DL tool, LMS, VLE or MOOC platform; its nascent workflow functionality does allow the creation of structured interactive 'tutorials', but the main focus of TeSS is its automatic aggregator functionality.

Originally, training information gathered by TeSS was not required to meet specific standards. However, as mentioned above, ELIXIR is working closely with GOBLET and others as part of the Bioschemas initiative, to outline minimal criteria to which ERs should now adhere in order to be considered 'ELIXIR' or 'GOBLET' compliant.

Implementation of TeSS (using the CKAN platform) was funded by the BBSRC as an ELIXIR-UK pilot, supporting two part-time developers for 18 months; its transition to 'production phase' (using Ruby on Rails) is funded by the EU EXCELERATE grant, also on a part-time basis. In common with all Web-based resources, TeSS has development and maintenance implications. To render the platform more sustainable, the platform gathers ERs largely automatically, limiting human intervention.

As with GOBLET, the main lessons from building TeSS include the need to perform focused requirements-gathering and to procure *dedicated funds* in order to be able to deliver services that users actually want. Ultimately, sustainable funding will be required to evolve the platform's functionality beyond the scope of EXCELERATE, to ensure that TeSS persists beyond the life-time of the grant and to encourage wider use.

*3)* The European Multimedia Bioinformatics Educational Resource, EMBER (https://ember.manchester.ac.uk), was an EU project, funded from 2001 to 2003; a 1-year 'no-cost' extension was needed to complete the project (Mabey and Attwood, 2001; Attwood *et al*., 2005). Building on a simple HTML-based tutorial developed in 1996 (http://www.biochem.ucl.ac.uk/bsm/dbbrowser/c32) and revised in 1999 (http://www.bioinf.man.ac.uk/dbbrowser/bioactivity/prefacefrm.html), the aim was to develop a more professionally packaged suite of ERs (including the interactive Web-based practical and an equivalent stand-alone version on CD-ROM) for under- graduate and early postgraduate studies, essentially to provide an introduction to bioinformatics (with particular emphasis on protein sequence and structure analysis).

EMBER was a 10-member consortium, including representatives from the UK, Switzerland, the Netherlands, Portugal and Belgium, and associates from South Africa and Canada. Two full-time developers were employed at the University of Manchester, and two others at the Expert Centre for Taxonomic Identification (ETI). The technical implementation of the Web-based product (including front-end python scripts and a back-end database) was professionally developed at the ETI. The system included a series of basic and advanced 'chapters', and various 'case studies'. Each chapter included a set of **learning objectives**; step-by-step, interactive practical tasks; background theory and literature to support the tasks; a self-assessment quiz, and so on. Initial and final quizzes were also used to evaluate user knowledge prior to completion of the tutorial and to gauge **learning outcomes** afterwards. While in development, user trials were conducted in Portugal, Belgium and the UK.

The system has been maintained on an unfunded basis since 2003; keeping it up-to-date has therefore been a challenge. The only real 'barrier' to uptake has been the need for users to register with the system (but completion of even a simple registration form is enough to put some potential users off). Nevertheless, since its release, the EMBER VLE has been used as the practical component of the Bioinformatics Masters programme at Manchester, and in a range of *ad hoc* short courses and summer schools around the world; being freely **available**, it has also enjoyed a worldwide audience via the Internet.

The main lesson from building EMBER is that developing e-learning systems to a professional standard is time-consuming and costly. EMBER was awarded only about a third of the funds requested – with adequate time and funding, the final product could have been significantly more advanced (*e.g*., it could have been augmented with animations and simulations, options to select and combine specific chapters, *etc.*). Nevertheless, the resulting product was extremely useful, and has provided an almost unbroken service since 2003.

*4) A taste of bioinformatics* was an introduction to bioinformatics for school pupils (http://web.archive.org/web/20060224100344/http://www.royalsoc.ac.uk/play.asp?id=3823). The project involved members of the EMBER team and of the Centre for Science Education (CSE) at Sheffield Hallam University, one of the UK's leading academic groups in Science, Technology, Engineering and Maths (STEM) education. The CSE had been commissioned by the Royal Society to create a series of educational 'games' for school children to try to encourage interest in STEM in general, and in bioinformatics in particular. The game developed – *A taste of bioinformatics* – simulated real-time database searching, multiple sequence alignment and mutation identification; it was accompanied by a brief, animated tutorial, providing some of the background to the simulated tasks – as such, it could be considered a relatively 'light weight' VLE, without formal assessment methods.

The development team included story-boarders, who liaised with the scientists, plus graphic designers who developed the front-end visuals (implemented using Flash animations), and various technical gurus who implemented the back-ends. The games developed by the CSE proved very popular, until they were archived in order to make way for other projects on the Royal Society's website.

The main lesson here once again concerns the need for the right team, with the right level of expertise and the right level of funding in order to be able both to develop a professional resource, and to sustain it beyond the life-time of its current grant.

### *EMBL-EBI's Train online e-learning portal - Richard Grandison*

Train online (https://www.ebi.ac.uk/training/online) provides a portal to a range of stand-alone tutorials created to provide users with greater confidence in using EMBL-EBI resources. It was designed for a number of different audiences, with different learning styles and different needs. Launched in 2011, it currently contains more than 80 tutorials, ranging from functional genomics to chemical biology and literature resources. The key features of the tutorials are easy access (no need to register), self-paced learning (users may take an entire tutorial or just relevant sections, and may repeat the tutorial several times). The tutorials are available in different formats and depths (conceptual, quick tours, walkthroughs, videos, webinars), and are designed to be engaging and interactive, with guided examples, annotated screenshots, exercises and quizzes. The ERs are released under a Creative Commons Share-alike licence, and are thus open source. Since 2012, more than 700,000 unique users from 217 countries have accessed the portal. Currently, there are, on average, 20,000 unique users per month, with 35% returning users.

Train online is implemented in Drupal, the CMS used across EMBL-EBI websites. It was developed by consensus (Wright *et al*., 2010) with key requirements of various stake-holders (users, Web developers, domain experts) in mind: *i.e.*, that it should be self-paced, easy to use and update, interactive, open source, and developed in-house.

The development costs have been sustained with EMBL-EBI funds: the first 1.5 years required a Web developer (65% FTE), scientific training officer (150%) and a manager (20%); domain experts contributed content on a goodwill basis. Annual maintenance requires the same expertise and competencies, but less time: Web developer (15%), training officer (100%), manager (5%) and content providers (own time).

### *Scalable Personalised Education: data-science approach for evaluating open-ended assignments in MOOCs and providing personalised feedback - Ronen Tal-Botzer*

It is probably generally agreed that ideal education happens 1-on-1. This makes it possible to thoroughly evaluate a student's performance, and allows personalised feedback from the teacher. Normally, however, education takes place in a 1-on-many environment, which makes personalised feedback virtually impossible.

Realistic assessments of real processes tend to come from analyses of big data. In education, big data are assignments. In STEM, assignments are an important, integral part of learning. Mostly, they are open-ended and hence not very convenient for automatic analyses. Evaluation and personalisation are not scalable. Outside the e-learning world, assessing assignments takes >20% of educators' time; in large courses, assignments are not evaluated thoroughly, as detailed feedback cannot be given on individual basis. In e-learning environments, the situation is much worse.

The number of MOOCs is growing almost exponentially. On average, there are 7,200 students registered for a MOOC, 90% of whom will drop out. In such circumstances, "*Evaluation and feedback in MOOCs are virtually impossible. It is The Barrier to the MOOC revolution* (Prof. Gil Weinberg, Georgia Tech)". Automatic means of assignment analysis and assessment are required; the Sense system was designed to help do this.

Sense is based on machine-learning technology, which provides automatic assessment. The system analyses the assignments (*e.g.*, pieces of code), provides feedback about their weaknesses, and assigns them to predefined, graded groups. The system was trialled at Bar-Ilan University, the largest university in Israel. In the pilot study, three professors and five teaching assistants supervised two programming courses (150 students in total), a data mining and a medicine course (40 and 50 students respectively). This led the Rector's academic director to request that Sense be implemented at institutional level.

Ultimately, this begs the question, will machine learning ever replace the human educator? Probably not; but it's likely that educators who use machine learning will replace the ones who don't!

### *ELIXIR-SI's e-Learning Platform (EeLP) - Brane Leskošek*

EeLP (https://elixir.mf.uni-lj.si) was established by the ELIXIR-SI node to enable life-science students and scientists, laboratory technicians and clinical professionals to learn about bioinformatics tools and services, to perform practical work, and to analyse and store the resulting data, via a single interface. In addition to providing courses, the platform facilitates communication between students and teachers, enables course assessment (also for ECTS-accredited courses), plus connection to specific bioinformatics tools and services (*e.g.*, local clouds and clusters) via a Web-based interface to different single-sign-on systems. Development began at the beginning of 2014, based on more than 5 years of experience in developing e-learning tools in the University of Ljubljana's Faculty of Medicine. The platform is based on the Moodle LMS, with specific standard module extensions. Using Video-Conference (VC) facilities, the system can provide a flexible and scalable solution for disseminating existing Face-to-Face (F2F) courses, now transformed into blended synchronous courses, where teacher and students reside in distant locations.

### *e-Proxemis, eBiomics, Moodle and stand-alone e-learning modules - Grégoire Rossier*

e-Proxemis and eBiomics (http://ebiomics.sdcinfo.com/homepage, Lisacek *et al.*, 2012) were created to guide time-pressured life scientists in the use of bioinformatics tools and resources, focusing on proteomics- and omics-based examples. e-Proxemis began in 2006 as a private-academic partnership between Geneva Bioinformatics (GeneBio) SA and the SIB Swiss Institute of Bioinformatics. Implemented in the SPIP CMS, it supported self-paced learning by providing stand-alone tutorials, case studies, scenarios, exercises and quizzes, resulting in a multiple-entry-point, non-linear navigation system. Within 4 years, several hundred users had registered with the system. However, as bioinformatics is a fast moving field, continuous updates were required, and manpower was needed to deliver new releases every three months. Despite being popular with users, lack of funds eventually led to its discontinuation. Some years later, the concept was revitalised as eBiomics, and extended to encompass additional omics tools and resources. Funded via an EU Lifelong Learning Program grant, this project included partners in the Netherlands (Wageningen University and NBIC), Belgium (Université Libre de Bruxelles), France (University of Marseille and SDC), and Switzerland (SIB and HSeT). eBiomics comprises several interconnected sections, which can be accessed through different interactive activities, grouped as resources, conceptual charts, protocols, case studies and exercises. Although the system is still available, lack of funds and manpower to create and update its contents has left it on stand-by.

The Vital-IT and the Swiss-Prot groups at SIB were commissioned by the Food and Agriculture Organization (FAO) of the United Nations and the International Atomic Energy Agency (IAEA) joint programme to produce e-learning stand-alone modules. The *Phylogenetics of animal viral pathogens and applications* module primarily aims to train veterinarians and molecular epidemiologists from diagnostic and research laboratories of developing FAO and IAEA member states. The *BLAST* and *Multiple sequence alignment* modules focus on sequence analysis by sequence similarity. Primarily targeting scientists in developing countries, these modules can be obtained on physical devices such as CD-ROMs or USB-keys upon request, and are also available via the dedicated Viral Zone website (http://viralzone.expasy.org/e_learning). Another stand-alone module developed in the SIB covers the Unix operating system and command-line basics. This module is a prerequisite for several SIB F2F courses today. These short, interactive (using animations and assessments), focused modules use the SCORM package for integration into LMSs, and can be accessed off-line.

To support its F2F courses, SIB also uses Moodle, mainly as an ER repository, which may be accessed by trainees both during and after courses. Moodle feedback forms are used to collect short- and long-term course evaluations. Overall, Moodle is used as an intranet for course organisers, participants and trainers.

### *Engaging Learning experiences - Pedro Fernandes*

Engagement in learning is a key component of any successful class. Engagement can certainly be facilitated by proximity, and is relatively easy to achieve in small learning environments; however, it is particularly difficult to achieve in asynchronous and geographically distributed settings, which is often the situation for e-learning. The ongoing movement to transform learning while maintaining engagement includes e-learning, but also encompasses blended learning, **flipped-classes**, **peer- instruction** and distance learning.

In Portugal, a distance-learning experiment with the University of Porto's Medical School gathered more than 250 students, more than 200 of whom attended remotely. The principal challenge was how to keep the same level of interaction for all students, regardless of their locations. Aside from allowing discussion of clinical cases **online**, an important asset for that community, this experience also allowed production of e-learning materials that could be used in other situations.

For successful delivery of distance learning courses, several important technical aspects were highlighted: the sound system used; the comprehensiveness of the slides provided; availability of an easy-to-use, reliable file-sharing system; the need for an appropriate chat room for questions and answers; a voting system where participants can express their opinions; and bi-directional video call facilities, where participants from different locations can see and interact with other participants.

During the workshop presentations outlined above, and in the discussions that followed, it was evident that there was a wide range of approaches and experiences, but little or no common usage of e-learning terminology within the represented communities. This, then, provided the motivation to try to reach some consensus on the definition of 'e-learning' and its attributes, in order both to facilitate future ELIXIR/GOBLET communication and collaboration, and pave the way for defining achievable e-learning activities for ELIXIR (and GOBLET) in future.

## Attributes of e-learning agreed during the workshop

Following the presentations, workshop participants, mainly trainers and organisers, discussed the following question: what should be considered *essential* attributes of 'e-learning' resources, and what attributes are just *nice to have*? The consensus is summarised in Table 1 (definitions of the attributes can be found in the Glossary at the end of this document, and online: https://elixir.mf.uni-lj.si/course/view.php?id=10).

**Table 1.** Essential and desirable attributes of e-learning resources.

| Essential attributes |
|---|
| Online |
| Interactive |
| Measurable learning outcomes |
| Prerequisites |
| Open access |
| Easy-to-use |
| Maintainable |
| Sustainable |

| Desirable attributes |
|---|
| Self-paced |
| Available on demand |
| Scalable |
| Interoperable |
| Quality control |
| Contextual |
| Support |

In its broadest sense, e-learning can be considered to be a mode of learning facilitated and supported by information and communications technologies, where ERs may be accessed via the Internet or via intranets. In this context, it was agreed that e-learning resources should be available **online**, that they should list **prerequisites** (as a guide to the level of knowledge or skills that would benefit the trainee most when engaging with the resource) and include **measurable learning outcomes** (to gauge how well specific learning targets have been met); importantly, they should also include **interactive** components (*i.e.*, should support **active learning**, with hands-on practical tasks, quizzes, and so on). In general, it was considered that e-learning resources should be **open access** (*i.e.*, should allow more-or-less unrestricted use and/or re-use, within specified copyright and licensing conditions) and be intuitive or **easy-to-use** (ideally, they should be **accessible**); they should also be **maintainable** (in terms of the feasibility and ease of making both technical and content updates), and **sustainable** (*i.e.*, have a long-term financial strategy for their maintenance).

The 'nice to haves' included being **self-paced** (allowing trainees to study in their own time at a rate comfortable for them) and available **on demand** (allowing access outside of timetabled sessions, at the point of need). It was also considered desirable for e-learning resources to be **scalable** (*i.e.*, feasibly expandable in scope and extendable to reach wider audiences), **interoperable** (*e.g.*, **SCORM**-compliant), to include some sort of **quality control** (*i.e.,* adhere to relevant standards and/or have some sort of metrics in place), and to be **contextual** (*i.e.*, integrate real-world examples to try to make the learning experience more meaningful to trainees). The final 'nice to have', where possible, is user **support** (whether at the level of dedicated help pages, context-sensitive help, help desks, discussion fora or 'clinics').

The desirability of linking e-learning systems to relevant tools and databases (via tasks, exercises, *etc.*) was also discussed, as hands-on interaction with bioinformatics resources helps to support active learning. Here, cross-connection with tools registries (*e.g.*, such as the Bio.tools registry (Ison *et al.*, 2016), the EMBRACE Web Service Registry (Pettifer *et al.*, 2010), BioCatalogue (Bhagat *et al.*, 2010), AppDB and other systems like BioMoby, Soaplab) was considered important: from a technical standpoint, this facilitates interoperation, as registries tend to adhere to standard communication protocols; from a user perspective, use of registries may also allow discovery of a wider range of tools and databases than are mentioned within the framework of a particular ER, and thereby perhaps provide a broader landscape for learning opportunities.

Another discussion point concerned **accreditation.** Although this is nice to have, it may require validation of standards at a national level; this is generally much harder to achieve in practice than **certification,** which may simply be used to confirm that certain levels of achievement have been met, without the necessity for external review. As accreditation and certification apply most readily to courses or programmes (especially in the context of higher education), and aren't generally applicable to all e-learning resources, they weren't included in the list of attributes in Table 1.

## What's missing?

Of the e-learning resources presented during the workshops, with the exception of e-Proxemis, all are still online in some form, but several are no longer funded, and maintenance is therefore sporadic (*e.g.*, EMBER) or frozen/non-existent (*e.g.*, eBiomics, *A taste of bioinformatics*). Moreover, all are very different in terms of their focus (and hence contents), the 'pedagogical' approaches and ERs they exploit or implement, and the technologies that underpin them. The TeSS and GOBLET platforms are centralised training portals, with primarily European and global outlooks, respectively, created to help trainers and trainees to store and discover ERs: neither was designed to provide a structured learning environment *per se* (although ERs may be structured within TeSS via packages and workflows). By contrast, EMBER was designed as a VLE, providing a self-contained, introductory practical protein sequence and structure analysis course that uses, amongst other things, various ELIXIR core resources, although it doesn't give practical introductions to the resources themselves. Conversely, Train online was designed specifically for this purpose, using standalone modules to give users greater confidence in using EMBL-EBI tools and databases. Rosalind is an online resource that helps users to improve their problem-solving skills, with a specific focus on bioinformatics programming techniques; MOOCs, on the other hand, are subject agnostic, but most are not available on demand.

Reflecting on the attributes that were considered essential and/or desirable for e-learning resources to have, it's interesting to consider the extent to which the presented resources fulfil those criteria. The results are summarised in Tables 2 and 3 – removed from this brief analysis were those resources that are no longer available or no longer actively maintained at some level.

**Table 2.** Presented resources measured against the 'essential' e-learning attributes. Green ticks denote attributes believed to be present; black ticks are attributes considered not fully satisfied; red crosses denote attributes that are not fulfilled; n/a is not strictly applicable.

| | Essential attributes | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Online | Interactive | Meas. LOs | Prereqs | Open Access | Easy-to-use | Maintainable | Sustainable |
| Train online | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| eGenomics Catalogue | ✔ | ✔ | n/a | n/a | ✔ | ✔ | ✔ | ✔ |
| EMBER | ✔ | ✔ | ✔ | ✗ | ✓ | ✔ | ✗ | ✗ |
| TeSS | ✔ | ✔ | n/a | n/a | ✔ | ✔ | ✔ | ✓ |
| GOBLET | ✔ | n/a | n/a | n/a | ✔ | ✔ | ✔ | ✓ |
| Rosalind | ✔ | ✔ | ✔ | ✗ | ✔ | ✔ | ✓ | ✓ |
| EeLP | ✔ | n/a | n/a | n/a | ✔ | ✔ | ✔ | ✓ |

**Table 3.** Presented resources measured against the 'desirable' e-learning attributes. Green ticks denote attributes believed to be present; black ticks are attributes believed to not be fully satisfied; red crosses denote attributes that are not fulfilled; n/a is not strictly applicable.

| | Desirable attributes | | | | | | |
|---|---|---|---|---|---|---|---|
| | Self-paced | On demand | Scalable | Interoperable | QC | Contextual | Support |
| Train online | ✔ | ✔ | ✔ | ✗ | ✔ | ✔ | ✔ |
| eGenomics Catalogue | ✔ | ✔ | ✔ | ✗ | ✔ | ✗ | ✔ |
| EMBER | ✔ | ✔ | ✗ | ✗ | ✗ | ✔ | ✔ |
| TeSS | n/a | ✔ | ✔ | ✔ | ✔ | n/a | ✔ |
| GOBLET | n/a | ✔ | ✔ | ✔ | ✔ | n/a | ✔ |
| Rosalind | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| EeLP | n/a | ✔ | ✔ | ✔ | ✔ | n/a | ✔ |

From the results in Tables 2 and 3, it's evident that no existing resource satisfies all of the essential and/or desirable attributes of a fully functional and supported e-learning resource. However, it's equally clear that the resources we've been discussing occupy very different niches in the e-learning ecosystem, and are hence not strictly comparable: *e.g.*, some are ER aggregators, while others are ER repositories; some are general purpose LMSs, while others are highly customised VLEs.

This diversity (some of which is captured in Table 4) helps to illustrate why the task of pinpointing what we collectively meant by 'e-learning' was so difficult: workshop participants had both different (technical) frames of reference and different mental models informing their particular views of what e-learning meant to them: for some, it was a MOOC, for others an instance of Moodle; for some it was repository, registry or catalogue of ERs, for others it was an immersive, subject-specific VLE. Reconciling such perspectives is hard; hence, the resulting classification captured in Table 4 is still debatable. Nevertheless, what is clear is that the granularity of e-learning resources is fundamentally different with respect to i) the modes of learning they either do or can potentially support; ii) the level of human input required to create and sustain them; iii) the level of technical infrastructure required to implement them; and iv) the level of funding required to create, maintain and sustain them. This cautions us to be more rigorous, and to try to define more clearly the level of granularity to which we're referring when we use the catch-all term 'e-learning'.

**Table 4.** Snapshot of the diversity of the e-learning ecosystem.

| Description | Technology | Characteristics | Instance | Content |
|---|---|---|---|---|
| Generic CMS/LMS | MOOC platform | General-purpose LMS, available 'off-the-shelf', and customisable for specific in-house use | Coursera bioinformatics specialisation | Customised course on bioinformatics algorithms |
| | Moodle | | EeLP | ERs & courses supporting systems biology, medicine, functional genomics, *etc.* |
| Registry | CKAN, Ruby on Rails | Automatic ER aggregator; stores ER meta-data; agnostic to ER type; primarily European focus | TeSS | Inventory of bioinformatics events & ERs (courses, materials, workflows, *etc.*) |
| Repository | Drupal | Manual ER store; upload open to all; | GOBLET | Inventory of bioinformatics trainers, |

| | | | | |
|---|---|---|---|---|
| | | agnostic to ER type; global focus | | events & ERs (courses, materials, *etc.*) |
| Catalogue/ portal | | Bespoke, manual collection of independent ERs relating to broad topic areas | eGenomics Catalogue | Collection of links to useful genomics resources, including training, journals, tools, etc. |
| | | | Rosalind | Pool of bioinformatics & programming problems |
| | | | eBiomics | Range of bioinformatics analysis workflows & various 'omics applications |
| | Drupal | | Train online | Standalone tutorials on EMBL-EBI resources, conceptual courses and recorded webinars. |
| VLE | PostgreSQL, Python | Bespoke, immersive resource, with interactive, focused, interdependent subject-specific practical tasks | EMBER | Self-contained hands-on introduction to protein sequence analysis, with live links to databases & software tools |
| | Flash | | *A taste of bioinformatics* | Self-contained hands-on introduction to protein sequence analysis, with simulated links to databases & tools |

## Recommendations for the ELIXIR/GOBLET strategy for e-learning, options and examples of best practices

ELIXIR is a pan-European endeavour to provide a sustainable data infrastructure for the life sciences. The 4-year EXCELERATE project, which commenced in September 2015, specifically aims to fast-track the implementation of the ELIXIR infrastructure and drive early user uptake. The focus of EXCELERATE WP11 is training: its principal tasks are to build a training infrastructure (11.1), and to deliver training to the ELIXIR community, via Train-the-Developer, Train-the-Researcher and Train-the-Trainer initiatives (11.2); among the subtasks are the TeSS platform (11.1.2) and e-learning (11.1.3). The focus of TeSS is on delivering a robust resource that seamlessly captures (with minimal human input) training information from ELIXIR Nodes (particularly in relation to EXCELERATE's use-cases) and 3rd-party providers, to help make ELIXIR's training events and ERs discoverable. By contrast, the focus of the e-learning sub-task is to review the existing expertise and technologies available across Nodes, and to derive scalable ERs relating to EXCELERATE's use-cases and ELIXIR's core resources, for different types of end-user. Working with TeSS and other WP11 efforts (including the analysis of training needs across Nodes), sub-task 11.1.3 is tasked with formulating an appropriate e-learning strategy for ELIXIR, providing concrete guidelines and recommendations on how best to implement ELIXIR's e-learning activities.

Toward this goal, one of the first steps was to organise the *Defining an e-learning lingua franca* workshop in Ljubljana, to help paint a picture of the expertise and technologies sequestered in the Nodes, and to better understand the challenges in building e-learning resources. From this workshop and the many follow-on discussions (including those in Cape Town), two fundamental observations stand out: first, that none of the e-learning resources described here (yet) provides a platform that delivers training on all of

ELIXIR's core resources *and* meets all of the essential (and desirable) attributes listed in Tables 2 and 3; second, that developing fully-fledged e-learning resources is time-consuming, expensive and difficult, requiring dedicated funds and highly proficient teams, with a spectrum of complementary pedagogical, design and technical skills, and expert subject knowledge.

Development of a fully featured e-learning resource wasn't costed into EXCELERATE (nor was it the aim of the GOBLET portal). This challenges us to determine i) what can feasibly be delivered via e-learning approaches/platforms, in the context of the very limited funds available in EXCELERATE and in light of other resources that already exist (like those presented in this paper); and ii) what assets can realistically be leveraged from other training activities orchestrated by the EXCELERATE WP11 Training Platform, such as those in CORBEL, RI-Train, *etc.* Ultimately, in light of the preceding discussions, what could '*an e-learning platform for ELIXIR*' deliver in the future and, in the shorter term, under the auspices of EXCELERATE?

The principal EXCELERATE-funded resources are EeLP and TeSS. TeSS wasn't developed as a pedagogical (e-learning) resource *per se*; rather, it was created to provide reference points for ERs available across ELIXIR and beyond: it aims to offer a comprehensive catalogue of training events and ERs, together with tools to create bespoke training packages and workflows. Two developers are currently funded on a part-time basis: their principal emphasis is i) to increase coverage of ELIXIR Nodes and 3rd-party providers, ii) to enhance links with the Bio.tools registry, and iii) to augment and build on TeSS' nascent workflow functionality. By contrast, EeLP is a Moodle-based LMS that uses VC facilities to provide, amongst other things, blended synchronous courses. Like TeSS, it is also supported by two part-time developers. In the context of EXCELERATE, it will deliver an e-learning infrastructure to store and provide easy access to ERs from ELIXIR's F2F and online courses, and to manage users.

In light of the outcomes of the workshops and discussions described here, and attempting to contribute to ELIXIR's over-arching e-learning strategy, a pragmatic approach would be for TeSS and EeLP to open channels of communication, to develop reciprocal links, and to harness the collective experiences gained from development of the Train online, eBiomics and EMBER systems to create a range of 'lightweight' e-learning tutorials that encourage hands-on engagement with ELIXIR's resources.

Bringing these perspectives together, within the constraints of the funding available, we recommend:

- building on the nascent workflow functionality of TeSS;
- ensuring that the TeSS and EeLP developers exchange best practices and experiences;
- ensuring that TeSS and EeLP are seamlessly interoperable;
- making e-learning resources findable via TeSS (including those systems discussed here that are still 'alive');
- creating e-learning modules using a consistent implementation language, such that they can be accessible in EeLP and shared through TeSS;
- ensuring that any resources developed satisfy the essential and most of the desirable attributes described here;
- adherence to standards (*i.e.*, SCORM-compliance), making it possible to reuse such resources in other e-learning settings;
- increased coordination across Nodes that have existing e-learning platforms (SI, CH, NL, FR, UK…) – *i.e.*, building an ELIXIR community with expertise in and/or a need for e-learning; and

- addressing use-case needs, endeavouring to re-use, wherever possible, rather than building from scratch.

Within the limits of EXCELERATE, we believe that following these recommendations would be the most efficient, cost-effective approach for supporting the training- and capacity-building platforms, for the benefit of the ELIXIR and GOBLET communities.

Outside the confines of EXCELERATE, there is scope to be more creative. One solution for the future could be to develop a joint grant proposal with a range of interested parties (potentially involving various ELIXIR Nodes, GOBLET, BD2K, H3ABioNet, CODATA-RDA, *etc.*), to create a powerful e-learning system, encompassing not just the range of essential and desirable attributes listed in Table 1, but including a more comprehensive set of features and supporting a wider set of learning modalities.

As proofs of concept, i) the EMBER skeleton structure has already been implemented as a TeSS workflow (in the coming months, the workflow structure will be populated with existing ERs and links to Bio.tools, whenever possible); and ii) an ELIXIR-FI F2F course on RNA-seq with Chipster (by Eija Korpelainen) has been transformed into an e-learning resource within EeLP, and executed in the ELIXIR-SI and ELIXIR-CZ nodes, and two further courses are in preparation (the Unix/Linux command line, prepared by ELIXIR-SE and ELIXIR-SI, and Python for Life Scientists, prepared by ELIXIR-IT, ELIXIR-SI, ELIXIR-PT, ELIXIR-UK and GOBLET).

To build on this process of exchange and tight collaboration, a workshop is to be organised early in 2017, to bring together members of ELIXIR-UK, ELIXIR-NL, ELIXIR-CH, ELIXIR-SI, *etc*. to plan the next integration steps and to prioritise courses.

## Conclusions

Fully-fledged e-learning resources are time-consuming to create and hard to do well. Experiences with systems like Train online, e-Proxemis, eBiomics, EMBER, TeSS, EeLP and so on, demonstrate that highly skilled teams are required both to design and implement e-learning resources, and to maintain them once they've been deployed 'in the wild'. The costs of developing and maintaining such systems to a professional standard are high.

Nevertheless, e-learning resources can reach large, geographically dispersed audiences, and bring efficiencies of cost and scale. They can furnish training opportunities at the point of need, at a pace that suits learners, and can allow effective progress tracking. As such, they are particularly suited to delivering training to widespread communities like those represented within ELIXIR. Logically, therefore, an e-learning resource of some kind should contribute to ELIXIR's training infrastructure.

As we've seen, if the ultimate vision were to develop an accessible, easy-to-use, open-access e-learning system that provided self-paced, on-demand, interactive training on ELIXIR's core tools and resources, with full user support, this would require substantial resourcing: it would require a diverse team, funded for a sufficient period of time to be able to deliver a robust product. The team would need to include members with i) subject-specific expertise; ii) relevant pedagogical skills; iii) expertise in graphic design; iv) experience with CMS/LMS and database development; and so on. It would also require a close collaboration with the tool developers and users of those resources. This would likely require a multi-million Euro investment.

Currently, neither ELIXIR-EXCELERATE nor GOBLET have sufficient resources to be able to achieve this. In the short term, we therefore recommend a pragmatic approach involving TeSS and EeLP, exploiting lightweight solutions that transform i) existing ELIXIR F2F courses into e-learning resources within EeLP, and ii) existing e-learning modules into TeSS workflows. The topics to be prioritised should be those identified in the training needs surveys distributed by the ELIXIR Training Platform (in relation to EXCELERATE's rare disease, marine metagenomics and plant use-cases). Longer term, TeSS and EeLP should become ELIXIR's reference platforms, by means of which users can discover ELIXIR training events and ERs. Of course, we welcome the addition of new resources to ELIXIR's e-learning portfolio, but encourage would-be developers to join the e-learning team, and help to disseminate good practices and adherence to standards. Finally, we recommend that dialogues between ELIXIR Nodes, GOBLET, BD2K, H3ABioNet, CODATA-RDA, *etc.*, remain open in order both to explore and exploit future opportunities to establish more formal collaborations, and to build on the foundations already established by these organisations.

## References

Attwood, T.K., Selimas, I., Buis, R., Altenburg, R., Herzog, R., Ledent, V., Ghita, V., Fernandes, P., Marques, I. and Brugman, M. (2005) EMBER - A European Multimedia Bioinformatics Educational Resource. *Bee-J,* 6-4.

Bhagat J, Tanoh F, Nzuobontane E, Laurent T, Orlowski J, Roos M, Wolstencroft K, Aleksejevs S, Stevens R, Pettifer S, Lopez R, Goble CA. (2010) BioCatalogue: a universal catalogue of web services for the life sciences. *Nucleic Acids Res*., **38**(Web Server issue), W689-94. doi: 10.1093/nar/gkq394.

Corpas, M., Jimenez, R.C., Bongcam-Rudloff, E., Budd, A., Brazas, M.D., Fernandes, P.L., Gaeta, B., van Gelder, C., Korpelainen, E., Lewitter, F., McGrath, A., MacLean, D., Palagi, P.M., Rother, K., Taylor, J., Via, A., Watson, M., Schneider, M.V. and Attwood, T.K. (2015) The GOBLET Training Portal: A Global Repository of Bioinformatics Training Materials, Courses and Trainers. *Bioinformatics*, **31**(1), 140-142. DOI:10.1093/bioinformatics/btu601

Ison, J., Kalas, M., Jonassen, I., Bolser, D., Uludag, M., McWilliam, H., Malone, J., Lopez, R., Pettifer, S. and Rice, P. (2013) EDAM: an ontology of bioinformatics operations, types of data and identifiers, topics and formats. *Bioinformatics*, **29**(10), 1325-32. doi: 10.1093/bioinformatics/btt113.

Ison, J., Rapacki, K., Ménager, H., Kalaš, M., Rydza, E., Chmura, P., Anthon, C., Beard, N., Berka, K., Bolser, D., Booth, T., Bretaudeau, A., Brezovsky, J., Casadio, R., Cesareni, G., Coppens, F., Cornell, M., Cuccuru, G., Davidsen, K., Vedova, G.D., Dogan, T., Doppelt-Azeroual, O., Emery, L., Gasteiger, E., Gatter, T., Goldberg, T., Grosjean, M., Grüning, B., Helmer-Citterich, M., Ienasescu, H., Ioannidis, V., Jespersen, M.C., Jimenez, R., Juty, N., Juvan, P., Koch, M., Laibe, C., Li, J.W., Licata, L., Mareuil, F., Mičetić, I., Friborg, R.M., Moretti, S., Morris, C., Möller, S., Nenadic, A., Peterson, H., Profiti, G., Rice, P., Romano, P., Roncaglia, P., Saidi, R., Schafferhans, A., Schwämmle, V., Smith, C., Sperotto, M.M., Stockinger, H., Vařeková, R.S., Tosatto, S.C., de la Torre, V., Uva, P., Via, A., Yachdav, G., Zambelli, F., Vriend, G., Rost, B., Parkinson, H., Løngreen, P. and Brunak, S. (2016) Tools and data services registry: a community effort to document bioinformatics resources. *Nucleic Acids Res*., **44**(D1), D38-47. doi: 10.1093/nar/gkv1116.

Jimenez, R.C., Albar, J.P., Bhak, J., Blatter, M.C., Blicher, T., Brazas, M.D., Brooksbank, C., Budd, A., De Las Rivas,J., Dreyer, J., van Driel, M.A., Dunn, M.J., Fernandes, P.L., van Gelder, C.W., Hermjakob, H., Ioannidis, V., Judge, D.P., Kahlem, P., Korpelainen, E., Kraus, H.J., Loveland, J., Mayer, C., McDowall, J., Moran, F., Mulder, N., Nyronen, T., Rother, K., Salazar, G.R., Schneider, R., Via, A., Villaveces, J.M., Yu, P., Schneider, M.V., Attwood, T.K. & Corpas, M. (2013) iAnn: An Event Sharing Platform for the Life Sciences. *Bioinformatics*, **29**(15), 1919-1921. DOI:10.1093/bioinformatics/btt306

Lisacek, F., Koks, P., Berthault, P., Rossier, G., Bottu, G., van Helden, J., Sylvestre, J., Kraehenbuhl, J.P. & Leunissen, J. (2012) eBIOMICS - A Website Promoting Good Practices and Specific Methods for using Bioinformatics Resources. In *Proceedings of the International Conference on Bioinformatics Models, Methods and Algorithms,* 386-391, 2012, Vilamoura, Algarve, Portugal

Mabey, J.E. and Attwood, T.K. (2001) EMBER: a European Multimedia Bioinformatics Educational Resource. *CAL-laborate. A collaborative publication on the use of Computer Aided Learning for tertiary level physical sciences*, **6**, 13-16.

Mulder, N.J., Adebiyi, E., Alami, R., Benkahla, A., Brandful, J., Doumbia, S., Everett, D., Fadlelmola, F.M., Gaboun, F., Gaseitsiwe, S., Ghazal, H., Hazelhurst, S., Hide, W., Ibrahimi, A., Jaufeerally Fakim, Y., Jongeneel, C.V., Joubert, F., Kassim, S., Kayondo, J., Kumuthini, J., Lyantagaye, S., Makani, J., Mansour Alzohairy, A., Masiga, D., Moussa, A., Nash, O., Ouwe Missi Oukem-Boyer, O., Owusu-Dabo, E., Panji, S., Patterton, H., Radouani, F., Sadki, K., Seghrouchni, F., Tastan Bishop, Ö., Tiffin, N., Ulenga, N.; H3ABioNet Consortium (2016) H3ABioNet, a sustainable pan-African bioinformatics network for human heredity and health in Africa. *Genome Res.*, **26**(2), 271-7. doi: 10.1101/gr.196295.

Pettifer S, Ison J, Kalas M, Thorne D, McDermott P, Jonassen I, Liaquat A, Fernández JM, Rodriguez JM; INB-Partners, Pisano DG, Blanchet C, Uludag M, Rice P, Bartaseviciute E, Rapacki K, Hekkelman M, Sand O, Stockinger H, Clegg AB, Bongcam-Rudloff E, Salzemann J, Breton V, Attwood TK, Cameron G, Vriend G. (2010) The EMBRACE web service collection. *Nucleic Acids Res.*, **38**(Web Server issue), W683-8. doi: 10.1093/nar/gkq297.

Wright VA, Vaughan BW, Laurent T, Lopez R, Brooksbank C, Schneider MV. (2010) Bioinformatics training: selecting an appropriate learning content management system - an example from the European Bioinformatics Institute. *Brief Bioinform.*, **11**(6), 552-562.

## Glossary

See also online https://elixir.mf.uni-lj.si/course/view.php?id=10

**Accessibility**: a design feature of a product, technology or environment that enables direct (unassisted) access and indirect access (*i.e.*, using assistive technology, such as a screen reader) for people with disabilities (including visual, hearing, motor or cognitive impairments).

**Accreditation**: the process by which courses are nationally recognised and meet established industry, enterprise, educational or legislative standards. It is commonly used in the context of higher education, where the quality of educational programmes is assessed by an external body; where standards are met, accredited status is granted by the agency.

**Active learning:** a mode of learning in which trainees actively, rather than passively, participate in the learning process. Learning is thus *experiential*: *e.g.*, it may involve reading papers, writing reports, completing tasks, problem solving, *etc*., and hence encourages learners both to do things and to think critically about what they're doing (see Wikipedia).

**Adaptive learning**: a mode of learning that exploits computers as interactive teaching devices, and in which the provision of ERs (and trainer input) is adjusted to match trainees' individual learning styles or needs. The adaptive process is orchestrated via trainee responses to questions, tasks and experiences. The technology draws on a variety of research fields, from computer science to neuroscience, education and psychology (see Wikipedia).

**Asynchronous learning**: a mode of learning whose defining characteristic is that students and teachers don't interact in person in real-time. An advantage of asynchronous learning is that it is adapted to learners' available time and learning styles. Typically, students utilise materials that have been pre-prepared and made accessible in a CMS/LMS.

**Availability**: describes whether an ER is accessible by trainees at all times (on demand), or whether access is time-limited (as, for example, with semestered courses in many MOOCs).

**Blended learning**: a mode of learning that combines characteristics of both synchronous and asynchronous learning, and is probably the most frequent mode in which e-learning is deployed. The typical blended-learning scenario is a course taught in a classroom, augmented with online materials and technologies; alternatively, it may be an online module interspersed with F2F components mediated by a tutor.

**Certification:** the process by which characteristics of an object, person or organisation are affirmed. Such affirmation may be provided by external review, assessment or audit, but external validation is not mandatory. In the case of professional certification, for example, a person's competence to complete a job/task may be confirmed by passing an exam, completing a course or educational programme) (see Wikipedia).

**Content Management System (CMS)**: a computer application specialised for organising digital content, often supporting project management or meeting users' information needs. Drupal, Joomla, Wordpress and Concrete5 are CMSs.

**Contextual teaching and learning**: involves making learning meaningful to students by connecting to the real world. It draws upon students' diverse skills, interests, experiences and cultures, and integrates these into what and how students learn, and how they are assessed (*e.g.*, an ER that has clear descriptions, gives examples, and is adapted to the context and situation of learners).

**Distance Learning (DL)**: a mode of learning, usually facilitated by information and communications technologies, in which students and teachers are not in regular F2F contact. The term is often used synonymously with distance education, e-learning, **online learning**, *etc.*; its defining characteristic, however, is geographic distance/displacement of teachers and learners. In the UK, DL has its roots in correspondence courses, such as those pioneered by the Open University.

**Educational Resource (ER)** or **Learning Object**: a digital resource designed for learning and teaching: ERs may be static, intended for reading or watching (Powerpoint slides, book chapters, video podcasts, webinars, assignments, Web pages, *etc.*), or dynamic, intended for active user participation (*e.g.*, lessons that adapt to the user's knowledge, tests, blogs, forums, *etc.*). Most ERs are learning objects that cover focused topics; these may be combined into wider structures (*e.g.*, courses) – hence, ERs may not necessarily include learning objectives or assessment functionality.

**Ease-of-use** (**usability** or **user friendliness**): describes the degree to which a resource is straightforward to use, navigate or understand (*e.g.*, the intuitiveness of a Web interface).

**E-learning**: a mode of learning facilitated and supported by information and communications technologies, in which ERs may be accessed via the Internet or via intranets.

**ER Aggregator**: a resource or software application whose primary function is to collect information (metadata) about ERs from multiple sources and to disseminate it via a single interface. Aggregators can actively search the Internet for ERs, or use established data-exchange links with ER producers, *pulling* metadata from them using standard information-exchange protocols (*e.g.,* OAI-PMH), or receiving metadata through *push* mechanisms from 3rd parties. TeSS is an example of an ER aggregator.

**ER Repository**: a resource or database system whose primary function is to store ERs. GOBLET is an example of an ER repository.

**Flipped-Class**: an educational strategy that uses a blended-learning approach, reversing the traditional paradigm by both delivering educational content (lectures, exercises, *etc.*) and encouraging peer-to-peer collaboration online, or providing opportunities to perform research activities in students' own time, and exploiting the classroom setting to engage in follow-up discussions under the guidance of a tutor (see Wikipedia).

**Gamification**: the introduction of computer-game elements into ERs to improve learners' (players') motivation and engagement. Gamification strategies often reward players who accomplish the learning tasks, and/or exploit the human tendency to compete.

**Interactive**: containing elements that allow bi-directional flow of information between a resource and its user (*e.g.*, a practical task, such as a real-time database search, or a quiz that provides real-time feedback on trainee responses): the resource interacts with the human user to obtain data or commands, and gives immediate results or updates information.

**Interoperability of ERs**: a property that renders ERs re-usable across different LMSs, generally by compliance with FAIR (Findable, Accessible, Interoperable, Reusable) principles and adherence to interoperability standards, such as SCORM and Tin Can.

**Learning Management System (LMS)**: a computer application specialised for dissemination and administration of ERs, including support for assessment of learning outcomes. An LMS may be considered a CMS specifically developed for management of ERs, with pronounced interaction capabilities and sophisticated user management, allowing participants to be organised into cohorts, groups and roles. Moodle and Blackboard are examples of LMSs; EeLP is a customised Moodle instance.

**Learning objectives**: statements that define the intentions of a teacher in terms of the purpose and goal of a curriculum, course, lesson or activity. They are also referred to as instructional objectives or learning goals.

**Learning outcomes (LOs)**: statements that define the demonstrable/measurable skills, competencies or knowledge that learners will have acquired having completed a course, lesson or activity – they are the tangible evidence that the learning objectives have been achieved.

**Maintainable**: the ability to support the technical infrastructure of a resource (in line with routine software and hardware updates) *and* to support updates to its content (in line with new knowledge or the availability of new functionality).

**Massive Open Online Course (MOOC)**: an online course aimed at an unlimited number of participants, with open access via the Internet. MOOCs provide traditional course materials, such as filmed lectures, readings and problems. Assessment of learning outcomes is achieved by quizzes and by peer-to-peer evaluations. MOOCs also provide interactive fora to support interactions between students and teachers. MOOCs have emerged as the fastest growing trend in the e-learning ecosystem.

**Measurable learning outcomes:** an outcome or level of attainment that can be measured or assessed in some way (*e.g.*, via a quiz; an oral, written or practical exam; *etc.*).

**MOOC platform:** a particular form of LMS that aggregates and hosts MOOC courses. Examples include edX, Coursera and Udacity.

**On-demand**: the availability of an ER by trainees at all times (*i.e.*, unrestricted by time-limits) at the point of need, allowing them to 'dip in and out' at any time, and to re-engage with it at any point or at any stage.

**Open access (OA):** unrestricted ability to access, use and re-use an online resource (*i.e.*, without subscription fees and/or certain copyright or licence conditions).

**Online**: accessible via the Internet or intranet for large or restricted communities of users.

**Online learning**: a mode of learning in which ERs are accessed via the Internet or intranet.

**Open Educational Resource (OER)**: an ER intended for free use. The levels of openness of OERs vary from complete or near-complete, in the case of Creative Commons or GNU General Public licences, to free only in certain circumstances (*e.g.*, for non-commercial use).

**Peer instruction**: a student-centred mode of learning in which students are given questions or concepts upon which to reflect, and are then required to formulate their own ideas or work out solutions within small groups. In the process of explaining their ideas to each other, students are forced to critique their own arguments, allowing them (and their tutors) to assess their understanding before leaving the classroom (see Wikipedia).

**Prerequisites**: the skills or level of knowledge required as a prior condition to engaging with a particular learning object, course, system, *etc*. (*e.g.*, the level of knowledge required to be able to understand a text-book; the practical skills required to be able to complete a course; the programming skills required to be able to set up an e-learning platform).

**Problem-based** (or **Scenario-based**) **learning**: a mode of learning that promotes independent, active task solving rather than passive learning. Given an assignment (which may include links to ERs), learners attempt different tasks, and make progress on the basis of positive outcomes from each step.

**Quality control**: the process by which the quality of a resource is assessed according to defined standards, specifications or metrics (*e.g.*, measures of learning effectiveness, trainee satisfaction, trainer satisfaction).

**Scalability**: the technical and/or financial ability to extend or expand the scope, functionality or the size of the user-base/audience of a resource, relative to the ease, efficiency and cost-effectiveness of the change required.

**Self-paced**: ability to complete a course, tutorial, exercise, *etc*. at a trainee's own speed.

**Sharable Content Object Reference Model (SCORM)**: the *de facto* industry standard for e-learning interoperability. SCORM specifies how ERs and LMSs communicate. If an ER is SCORM-compliant, it can be used in any SCORM-compliant LMS, and *vice versa*.

**Support**: synchronous or asynchronous help available within a resource or course (*e.g.*, contextual help embedded in Web pages, a Q&A forum, a help-desk, a trainer-trainee 'clinic').

**Sustainable**: the ability (both technically and financially) to maintain the content and functionality of a resource, and its technical infrastructure, in a long-term, cost-effective way.

**Synchronous learning**: a mode of learning whose defining characteristic is that students and teachers are engaged at the same time, whether face-to-face in a classroom or online. Online lessons are typically delivered via video-conferencing equipment or live-streamed podcasts.

**Traditional instruction** (or **traditional education**): a method of instruction that is teacher-centred, rather than learner-centred, in which the focus is on rote learning and memorisation; it was the predominant practice until educational reforms in the 1980s encouraged the use of more holistic methods (see Wikipedia).

**Virtual Learning Environment (VLE)**: a (usually bespoke) immersive, interactive learning environment, with focused, practical tasks that aim to achieve specific outcomes (ability to perform database searches, to interpret output results, *etc.*), providing all the necessary information and (access to) resources for performing those tasks. The environment is thus *experiential* and puts theory into behavioural *practice*. Examples include EMBER and the Royal Society game, *A taste of bioinformatics*.

Speakers and organisers of the ELIXIR and GOBLET workshop held on 15-17 September, 2015 in Ljubljana, Slovenia

Phillip Compeau, Computational Biology Department, Carnegie Mellon University, USA
Anthony F. Camilleri, The Knowledge Innovation Centre, Swieqi, Malta
Victoria Nembaware, H3Africa Consortium, South Africa
Teresa K. Attwood, University of Manchester, United Kingdom
Richard Grandison, EMBL-EBI, Hinxton, United Kingdom
Ronen Tal-Botzer, Bar Ilan University, Israel
Brane L. Leskosek, University of Ljubljana, Faculty of Medicine, Slovenia
Grégoire Rossier, SIB Swiss Institute of Bioinformatics, Switzerland
Pedro Fernandes, Instituto Gulbenkian de Ciência, Portugal
Patricia M. Palagi, SIB Swiss Institute of Bioinformatics, Switzerland
Celia van Gelder, DTL and Netherlands Bioinformatics and Systems Biology research school (BioSB), The Netherlands
Jure Dimec, University of Ljubljana, Faculty of Medicine, Slovenia
Eija Korpelainen, CSI, Finland
Sarah Morgan, EMBL-European Bioinformatics Institute
Rita Hendricusdottir, University of Oxford, United Kingdom
Lee Larcombe, Oxford University, United Kingdom
Chris Ponting, University of Oxford, United Kingdom
Alex Upton, UMA, Spain
Hedi Peterson, ELIXIR-EE, Estonia
Allegra Via, Sapienza Università, Rome, Italy
Vojtech Spiwok, UCT Prague, Czech Republic
Dimitris Kafetzopoulos, Greece
Alexander Botzki, VIB, Belgium
Nygård Ståle, University of Oslo, Norway
Sara Light, Stockholm University, Sweden
Henriette Bak-Jensen, University of Copenhagen, Danemark

Niall Beard, University of Manchester, United Kingdom
Manuel Corpas, The Genome Analysis Centre, United Kingdom
Miroslav Ruda, CESNET, Czech Republic

Speakers and organisers of the GOBLET and ELIXIR workshop held on 19 November, 2015 in Cape Town, South Africa

Annette McGrath, CSIRO, Australia
Bruno Gaeta, UNSW, Sydney, Australia
Nicola Mulder, University of Cape Town, South Africa
Michelle Brazas, OICR and bioinformatics.ca, Canada
Celia van Gelder, Dutch Techcentre for Life Sciences and Netherlands Bioinformatics and Systems Biology research school (BioSB), The Netherlands
Eija Korpelainen, CSI, Finland
Cath Brooksbank, EMBL-EBI
Judit Kumuthini, CPGR, South Africa
Teresa K. Attwood, University of Manchester, United Kingdom
Fourie Joubert, Galaxy, USA
Pedro Fernandes, Instituto Gulbenkian de Ciência, Portugal
Pascale Gaudet, SIB Swiss Institute of Bioinformatics, Switzerland
Fran Lewitter, ISCB
Manuel Corpas, TGAC, United Kingdom
Peter van Heusden, SANBI, South Africa
Erik Bongcam-Rudloff, SGBC-SLU, Sweden
Patricia M. Palagi, SIB Swiss Institute of Bioinformatics, Switzerland
Javier De Las Rivas, Cancer Research Center (CiC-IBMCC, CSIC/USAL/IBSAL), Spain
Vicky Schneider, TGAC, United Kingdom
Vicky Nembaware, H3Africa, South Africa
Gabriella Rustici, University of Cambridge, United Kingdom
Pamela Black, Welcome Trust, United Kingdom