

October  
2013

# Outcomes from the workshop for e-Infrastructure trainers

The workshop for e-Infrastructure trainers brought together 50 trainers from across the UK to work on plans for improving e-Infrastructure training, and to determine the benefits of creating a formal training community. The workshop took place on 14 August 2013.

This report summarises the discussions that took place at the workshop and lists the outcomes and plans for the future of the e-Infrastructure training community.



# About the Institute

The Software Sustainability Institute cultivates world-class research with software.

We help people build better software by working with researchers, developers, funders and infrastructure providers to identify key issues and best practice. The Institute is funded by the Engineering and Physical Sciences Research Council through grant EP/H043160/1.

For more information, visit [www.software.ac.uk](http://www.software.ac.uk).

# About this report

This report summarises the discussions that took place at the workshop for e-Infrastructure trainers and draws outcomes from these. It also details the benefits that would be gained from the creation of an e-Infrastructure training community. The report was written by Simon Hettrick.



This work by the University of Edinburgh on behalf of the Software Sustainability Institute is licensed under a Creative Commons Attribution 2.5 UK: Scotland Licence.

# Contents

<b>About the Institute</b>	<b>2</b>
<b>About this report</b>	<b>2</b>
<b>Overview</b>	<b>4</b>
Further information	4
Format	4
Acknowledgements	5
<b>Outcomes</b>	<b>6</b>
<b>Discussion group summaries</b>	<b>8</b>
Topics discussed at the workshop	8
How do we ensure a future for the e-Infrastructure training community?	8
Improving the understanding and status of training in research	9
How do researchers discover training?	10
Providing incentives to produce good training materials	10
Working with technical writers to produce high-quality training materials	11
Training the trainers - does it work and, if not, how can we make it work?	11
What statistics and stories do we need to make the case for training and how do we collect them?	12
What role should accreditation play in e-Infrastructure training?	12
What role should distance learning play in training?	13
The future of funding for training	14
<b>Conclusions</b>	<b>15</b>

## Overview

The future of research will rely on harnessing the power of computing, software and data - or *e-Infrastructure*. The effective use of e-Infrastructure is limited, because many researchers lack the required skills, and the fragmented nature of the e-Infrastructure training community makes it difficult to obtain them.

The field of e-Infrastructure is incredibly broad, so it is not possible for one group to train a researcher in every possible aspect of the technology. Instead, the responsibility for training is shared across many providers, which each specialise in different disciplines and are distributed across the UK. With each provider focussing on a specific aspect of training, it is easy to overlook the benefits that could be gained by collaborating with other providers.

Better communication between training groups would bring many benefits: sharing resources prevents duplication of effort, sharing best practice prevents duplication of mistakes, and collaborations bring a broader range of skills into a training programme. With these benefits in mind, the Workshop for e-Infrastructure trainers was organised on 14 August 2013 at the Hartree Centre in Daresbury. The workshop brought together e-Infrastructure trainers from across the UK to investigate the benefits of working together as a community, and to discuss how such a community would operate. These issues were investigated through a series of discussion groups, the results of which are summarised in this report.

The main issues that were raised at the workshop have been distilled from the discussions and are listed in the outcomes section on page 6.

## Further information

Everything about the workshop, from the agenda to the list of discussion topics, can be found online:

- <http://www.software.ac.uk/workshop-e-infrastructure-trainers>

Discussions between workshop attendees took place using a mailing list. The emails have been archived to create a permanent record of the discussions, and can be viewed online:

- <https://groups.google.com/forum/#!forum/e-infrastructure-trainers-2013>

## Format

The workshop was split into three sections: networking to make new contacts, lightning talks to showcase ideas, and discussion groups for in-depth conversations about subjects of importance to training.

To ensure that discussions were relevant to the current e-Infrastructure training community, the subjects for the discussion groups were nominated by the workshop attendees. During the workshop, attendees voted for the subjects in which they were interested, and only the most popular subjects were given time on the agenda.

## Acknowledgements

The workshop was organised by the Software Sustainability Institute in collaboration with the Digital Curation Centre (DCC), EPCC, Hartree Centre, Numerical Algorithms Group (NAG), and the STFC.

The author of the report would like to thank Ian Bush, Joy Davidson and David Henty for the benefit of their expertise and their invaluable help organising the workshop. Special thanks are extended to David Cable and Claire Devereux who, in addition to their expertise, provided the venue at which the workshop was held and funding for it.

# Outcomes

A number of issues were highlighted at the workshop, either because they were of significant importance or because they were repeatedly encountered by the trainers. Investigation of these issues constitutes the outcomes of the workshop.

The workshop provided the forum in which these issues were identified. It is now the responsibility of the e-Infrastructure training community to investigate and resolve these issues. In many ways, the outcomes of this report mark the start of the community which the attendees of the workshop voted to create.

## 1. Build an e-Infrastructure training community

- The e-Infrastructure training community needs a forum that provides a single point of contact for trainers, ensures better communication within the community, and can represent the community for the purposes of lobbying.
- The community will be created as a special interest group (SIG) for e-Infrastructure trainers. The possibility of an RCUK e-Infrastructure training working group should also be investigated.
- A SIG committee will be created which will assume the responsibility of operating and funding the SIG.
- Necessary communication infrastructure for the SIG will be created: website, mailing list, regular workshops.
- Where possible, the e-Infrastructure training community should work with existing communities with an interest in training, such as those related to HPC training, and IT services for research, and the groups that represent research technologists.

## 2. Increase recognition for the importance of e-Infrastructure training

- There is significant disconnect between the benefit to the research community of providing training, the time and effort it takes to produce training, and the reward a researcher receives for creating that training.
- Investigate the existing training provision and determine metrics that represent its success.
- Based on these success metrics, produce guidance on the relationship between the benefits of e-Infrastructure training and the strategic goals of the different stakeholders in the research community.
- Perform a cost-benefit analysis of the case for accrediting training, and investigate the success of existing schemes for accrediting e-Infrastructure training

## 3. Investigate the provision of training and make it more visible

- The current provision of e-Infrastructure training is distributed across many organisation which need to work together to promote available training, and create a continuous spectrum of training resources from beginner status to expert.
- Create a list of skills that underpin the effective exploitation of e-Infrastructure, and investigate existing training to identify duplicated resources and gaps in provision.

- Investigate the creation of a single-access point for training materials and courses, and determine support and funding for the concept.
- Investigate the success of champion schemes for broadening the uptake of training and promote them.

#### 4. Create and share best practice

- By working together, everyone in the e-Infrastructure training community can benefit from its shared expertise. A series of best practice guides should be created by bringing together working groups of experts within the community.
- Best practice for collecting metrics that demonstrate the success of training and the need for it.
- Best practice for the creation of training materials (such as, communication skills, proofing, design, provenance, intellectual property, etc.)
- Best practice for including a technical writer - or other specialists - on funding bids in order to improve the quality of training materials.

## Discussion group summaries

The majority of the workshop was dedicated to discussion groups. The workshop attendees voted for the topics in which they were interested, and then split into small groups (of around 3-10 people) to discuss them. The results of the discussions were captured in an email summary (which have been archived, see *Further Information* on page 4 for details) and it is these that form the basis of this report. In total, eleven subjects were investigated during the workshop, as described in the list below.

### Topics discussed at the workshop

- How do we ensure a future for the e-Infrastructure training community?
- Improving the understanding and status of training in research
- How do researchers discover training?
- Providing incentives to produce good training materials
- Working with technical writers to produce high-quality training materials
- Train the trainers - does it work and, if not, how can we make it work?
- What statistics and stories do we need to make the case for training and how do we collect them?
- What role should accreditation play in e-Infrastructure training?
- What role should distance learning play in training?
- The future of funding for training

### How do we ensure a future for the e-Infrastructure training community?

The workshop showcased the benefits that are gained from meeting together as a community: making new contacts, publicising work and sharing best practice. A vote held at the workshop showed - with near unanimity - that there was great interest in the creation of a formal e-Infrastructure training community.

There are various models that could be employed for the workshop, but the preferred option is a special interest group (SIG). A SIG provides a forum through which trainers can communicate, it organises regular meetings to strengthen the community, it provides a single point of contact against which issues about training can be raised and it has the weight of numbers needed to lobby for change. The SIG would be governed by a small committee and funded by contributions from the members or the member organisations. It was also suggested that the SIG could contribute to an RCUK working group on e-Infrastructure training.

The Software Sustainability Institute has agreed to provide effort to start the SIG, with the aim of handing over governance to a committee once some momentum has been built up.

Communities already exist within the field of e-Infrastructure training, but these do not represent the whole training community and they tend not to share information outside of their group - hence the need for a cross cutting e-infrastructure SIG. A number of these existing communities were identified (see below). These communities must be contacted and representatives should be asked to join the e-Infrastructure SIG to ensure representation of their views to the whole training community.

- A community exists around advanced HPC training.
- A community exists around the people who provide IT services for research.
- A number of communities exist around services that are oriented to the support of research (the name *research technologist* for this role appears to be gaining some traction). These include research software engineers, data scientists and information professionals.

## Improving the understanding and status of training in research

There was a general perception in the group that undergraduates are not taught all of the skills they will require in research (postgraduate and beyond). This creates a skills gap that must be filled by trainers. However, it was argued that there is little incentive for researchers to take advantage of training, because there is a disconnect between the reward that results from learning new skills, and the time and effort that researchers are allowed to spend on acquiring them.

A specific problem is that the supervisors of researchers (i.e. senior academics and managers) prefer domain-specific training over e-Infrastructure training. This is understandable, because with its direct relevance to the researcher's work, it is easier to make the case for domain-specific training. The broad applicability of e-Infrastructure training makes it difficult for a supervisor to perceive a specific benefit. Countering this objection will require supervisors to appreciate the significant improvement to research that can be gained from harnessing e-Infrastructure, which requires the e-Infrastructure training community to better publicise these benefits.

One option for improving the provision of training across disciplines would be to provide more - and preferably free - summer schools, and training at domain-specific conferences (i.e. the conferences that the researchers attend as a matter of course). This would make it easier for researchers to gain training at a reasonable cost and without taking too much time out from their research work.

Many training schemes rely on volunteers from the research community to provide the training (i.e. to conduct the actual lectures). For example most summer schools use volunteer trainers and schemes such as Software Carpentry are completely reliant on volunteers. Using researchers as volunteer trainers is highly effective because they can provide a direct link between the skills they are teaching and the benefit they have observed to their research. However, there is little incentive for a researcher to give up time from their research to work as a volunteer trainer. This is because there is little recognition within academia for the benefit that volunteer trainers provide to the research community, and there is no reward either financially or in terms of career progression for researchers who train.

If institutions, the research community and funding organisations can be persuaded to recognise and reward volunteer trainers, there will be a significant growth in both the number of researchers who will be available to volunteer to train, and the range of skills available for train-

ing.

To improve the status of e-Infrastructure training, it would be useful to persuade professional bodies (IEEE, IOP, Royal Society, Wellcome Trust, Leverhulme, IChem, etc.) to recognise e-Infrastructure skills and include them as part of the accreditation of courses and degrees.

## How do researchers discover training?

There is a general campaign to get more researchers using e-Infrastructure, but the training that will make this possible is difficult to discover, because it is distributed over many different training providers. Researchers cannot gain skills if they cannot find the training that provides them.

There is a convincing case for effort and funds being invested into sharing training data. This would lead to a better understanding of the available training, which would prevent duplication of effort, make it easier to build on existing training and allow gaps in the provision of training to be identified. In turn, this would lead to a more efficient deployment of training effort and a greater reuse of materials, which would reduce the cost of training across the UK.

A single access point - a marketplace - for training courses and materials would make it significantly easier for researchers to discover training. There would be further benefits too. Trainers would be publicising the same marketplace, which would raise awareness of it across the research community - and greater awareness should broaden the uptake of training. The marketplace would also provide a forum for discussion about training, allowing community-based recommendations and, potentially, answers to researchers' questions.

A marketplace for training would require trainers to add information and materials to a third-party website, which will require a convincing case to be made for the benefits of doing so. Help will need to be provided with understanding and overcoming issues around licensing.

The use of champions was discussed as a way of increasing the reach of training into new research domains. Training champions are researchers who have received training and benefited from implementing it in their research. They are effective at dissemination, because they are trusted by the researchers: they use the right language and can show a direct benefit to their domain. A network of training champions could help to extend available training into new disciplines.

Although there is often prestige associated with being a champion, recruiting sufficient champions (and retaining them) will require further incentives, such as providing direct access to expertise from training providers.

## Providing incentives to produce good training materials

Good training materials make it easier to learn about a subject, and are more likely to be shared, but producing them takes time and skill. Currently, there is little recognition for the contribution that training makes to the research community, so there is little incentive to invest time and effort into producing good training materials. Hence, the first step in producing better training materials is to recognise the contribution that training makes to the research community and then reward the people who produce those materials.

A number of obstacles prevent the production of good materials. A lack of accreditation (as de-

scribed in a different discussion group on page 12), makes it difficult to select good materials over poor ones. A lack of understanding about provenance makes it difficult to trace the creator of materials, which means they lack recognition for creating them. Poor understanding of intellectual property prevents trainers from sharing their materials for fear of infringing copyright. Finally, a lack of access to skilled writers - people who can proof read and improve the clarity of writing - has a direct impact on the quality of the materials.

## Working with technical writers to produce high-quality training materials

Trainers are well aware of the need to produce good training materials that not just cover the subject matter, but do so in an accessible and easy to follow manner. Why then are so many training materials perceived to be of a poor quality?

- Trainers generally come from a technical background and are unlikely to have been trained in communication and hence lack skills in writing, proofing and design. What's more, trainers who wish to improve their communication skills find it difficult to find appropriate courses.
- Good materials take a long time to prepare: a rule of thumb is eight hours preparation to one hour's worth of material. When combined with the lack of recognition and reward for producing training materials, there is little motivation to spend time on improving materials beyond the stage of *it does the job*.
- Trainers are too close to their subject matter, which makes it difficult to perceive training in the same way as a trainee (i.e. a beginner). This is partially due to a problem with communication skills and partially due to the lack of available resources making it difficult to secure the services of a proof reader.

One solution is to work with a technical writer. This separates the creation of training materials into two stages, each of which requires its own set of skills. The trainer provides expertise in the subject matter, and the technical writer provides communication skills and the perspective of a beginner. It was generally agreed that this separation of responsibilities would improve the quality of training materials, but this point is moot if there are no resources to employ a technical writer. It was noted during the discussion - but generally not well known - that technical writers are a legitimate cost on research council funding bids.

Resources for improving communication skills were discussed. Namely, there should be some effort to find an appropriate ISO standard for the creation of training resources. Tutorials are available from the Intel software college and from the Institute of Scientific and Technical Communicators [1].

[1]: <http://www.istc.org.uk/>

## Training the trainers - does it work and, if not, how can we make it work?

There is significant demand for training which can only be fulfilled if many more trainers are recruited. Finding people with the right skills is difficult: they must be domain experts, good

teachers and enthusiastic. Whereas basic training (on how to use a computer, say) needs only a generalist trainer, much e-Infrastructure training requires a trainer who is also an expert in the domain. This significantly limits the pool of people from which trainers can be recruited.

One way to overcome the shortage of trainers is to *train the trainers*. This is where expert trainers provide their expertise to apprentices and accelerate their acquisition of skills in the subject area and their knowledge about *how* to train. It was noted that a focus on teaching psychology (timings, gaps between exercises, attention spans, etc.) is often overlooked in e-Infrastructure training. The Goblet project [2] has set up a training trainers network, which has had some success and could be a model for future initiatives.

[2]: <http://www.mygoblet.org/>

## What statistics and stories do we need to make the case for training and how do we collect them?

Statistics are seen as a metric for funders (to gauge how well money has been spent), and not something that has value for the trainers themselves. Consequently, trainers find it difficult to dedicate the necessary time to collecting statistics about their work and, when they do, the collection is usually oriented towards justifying funding. It was felt that there is a lack of experience and clear guidance on what metrics should be collected and how the measurement should be performed. (For example, the people in the discussion group were unclear on the relative merits of quantitative and qualitative metrics.)

It was agreed that there should be a new focus on collecting metrics about training. In addition to qualitative metrics (“did you enjoy the training?”, “did you find it useful?”, etc.), quantitative metrics on the effectiveness of the training could be measured by assessing the aptitude of trainees before and after the training. This would make a convincing case for the effectiveness of the training. Where possible, follow up surveys should lead to a story that will raise awareness of future training.

The UK Professional Standards Framework (UKPSF) [3] was discussed as a useful resource that could help with recognising and benchmarking teaching and learning support. Resources for surveying were discussed - Survey Monkey and Google Forms - and LinkedIn was noted as being useful for mapping networks of communities.

[3]: <http://www.heacademy.ac.uk/UKPSF>

## What role should accreditation play in e-Infrastructure training?

It is difficult to choose training, because there is no reliable way of comparing the quality of different training courses or materials. This introduces an uncontrolled risk into choosing (and paying for) training, which can dissuade a researcher from attending training. By providing assurance of the quality of the training, accreditation of training might be the solution.

Accreditation would allow researchers to compare training packages and, in doing so, would reduce the perceived risk of investing in training. It would also have a beneficial impact on the amount of sharing, since training that achieves accreditation would be more likely to be trusted.

There are drawbacks. Confidence needs to be developed in the accreditation process, which takes time, although this could be accelerated by relying on existing - and trusted - accreditation bodies (see recommendations below). Accreditation adds a cost to the creation of training, which must be understood and balanced with the potential benefits in simplifying choice and increasing the sharing of materials.

Possible accreditation bodies:

- The Open University: <http://www.open.ac.uk/>
- Intellect: <http://www.intellectuk.org/>
- The British Computer Society: <http://certifications.bcs.org>
- The CPD certification service: <http://www.cpduk.co.uk>
- The Higher Education Academy: <http://www.heacademy.ac.uk/professional-recognition>

There is some experience in this area within the e-Infrastructure community: HPC Wales provides a PGCert in HPC, and Software Carpentry are investigating Mozilla *Open Badges*. Massive open online courses (MOOCs) use online tests, like blackboard or moodle, for accreditation. The success of these schemes should be investigated.

Finally, it might be preferable or easier to accredit trainers rather than training.

## What role should distance learning play in training?

Distance learning can provide a simple way of scaling training, and it is generally less costly than face-to-face training. However, it was felt that distance learning was not suited to specialist training, because this often requires the skills of an on-hand trainer. This may limit the applicability of distance learning to e-Infrastructure training.

The up-front investment in online training is lower (less expense, no travel), which may encourage more researchers to try out the training and hence provide an opportunity to showcase the benefits of e-Infrastructure to a new market. Although it can be difficult to attract people to online training - simply because there is so much training available - this problem could be countered if the e-Infrastructure training community worked together to promote specific online resources.

Various types of training can be classified as distance learning:

- Massively Open Online Courses (MOOCs)
- Distance learning of masters and undergraduate degrees
- Subscription-based training where trainees pay for a short online course, which might include Q&As or screencast lectures.
- Open, community training such as the training available on websites like YouTube.

## The future of funding for training

Trainers have experienced a significant demand for their expertise: courses, summer schools and workshops are often over subscribed - sometimes significantly so. Despite this demand, securing funding to run training is difficult. The e-Infrastructure community must make the case for why it is important to fund training by demonstrating demand, proving value and lobbying funders with this information.

The first and most obvious way of funding training is to charge the researchers who attend the courses. However, the research community generally expects training to be provided for free - especially if that training is provided by a group that is seen as being part of academia. One consequence of this expectation is that research bids rarely contain adequate (or indeed any) funding for staff training, and this makes it difficult - if not impossible - for researchers to attend paid-for training. A second option for funding training is to get funders to pay for the training, which means that researchers can attend for free.

The training community must either campaign to change researchers' expectations about paying for training, or secure funding for training from funding organisations. Of the two options, the latter seems to fit more easily with the accepted practices of the research community.

Groups who have been successful in gaining funding for training should work together to produce guidance on how to demonstrate the benefits of training and hence make the case for funding it. These arguments must be made in the light of furthering the strategic goals of the research councils and other funders.

The case for funding could be strengthened by reviewing the training provision to identify what training is needed and where effort is being duplicated. To this end, a gap analysis should be conducted across e-Infrastructure training. Once the results of the gap analysis are known, and the benefits of funding training have been demonstrated, this information should be used by the community to lobby for the funding of training.

## Conclusions

The criticism levelled at the e-Infrastructure training community is that it is fragmented and isolated. It is perhaps surprising then that so much common ground was discovered during the workshop for e-Infrastructure trainers.

The most significant outcome from the workshop is that the community has a strong desire to formalise. A vote at the workshop showed almost unanimous support for the creation of a special interest group (SIG). The creation of this SIG will provide a single point of contact for the trainers in the community, which is vital for dissemination and lobbying, and should make it significantly more likely that trainers will share resources and best practice. The next step is to raise the issue of the SIG with the community and canvas for people to join the committee.

Recognition of training as a vital contribution to the research community should be one of the first issues that is challenged by the SIG. This will require convincing arguments and the facts needed to support them, so a first step must be the collection of success metrics.

Contributing to the lack of recognition of training is the relative difficulty of discovering courses and resources. The creation of a single access point to training would be a significant improvement over the current spread of resources across many different organisations. It would also allow researchers to build up knowledge and confidence in a single resource for training. Expansion of champion schemes should also be investigated since they appear to provide significant benefits - especially at overcoming e-Infrastructure scepticism within a domain - for relatively moderate cost.

If the goal is to combat fragmentation, then the sharing of resources and best practice may prove to be the main benefits gained from the creation of an e-Infrastructure community. By working together, the training community will increase the quality and visibility of resources, making it easier for researchers to acquire the skills needed to harness e-Infrastructure for their research.