In part 1 of this series on distributed computing education, we introduced a list of components important for teaching environments. We outlined the first three components, which included development of materials for education, education for educators and teaching infrastructures, identifying current practice, challenges, and opportunities for provision. The final component, a supportive policy framework that encourages cooperation and sharing, includes the need to manage intellectual property rights (IPR).

E-science Education and the Importance of IPR

In many institutions where e-science is or could be taught, there’s a lack of appropriate teaching expertise. Computer scientists knowledgeable about e-science understand the technological issues but might be unable to show the relevance of e-science to the scientific disciplines. Similarly, educators within the scientific disciplines who understand their domain’s issues and challenges might not be aware of the potential for e-science to address these issues. Developing educational exemplars and resources for a new scientific discipline is a huge undertaking, and it’s simply not feasible for each institution to repeat this exercise. Instead, good educational content must be shared and developed collaboratively across institutions so a wide pool of experts can use it and contribute to it. This leads to the improved quality and relevance of educational resources.

At its core, e-science aims to be collaborative. The model of sharing and collaboration in creating and improving specialist educational content not only suits this well, but can actually provide students with a good case study of how distributed collaborative projects can work. Sharing educational content lowers the barriers of entry for e-science educators, or for educators in scientific or other disciplines hoping to teach e-science technologies in their courses. Teachers can share educational content rather than creating their own, so they increase the number of quality resources available to them. This in turn encourages uptake of e-science education and training, as well as creativity and sharing among educational professionals already active in e-science education. However, in sharing these resources, we must consider IPR issues such as copyright ownership and licensing.

Intellectual property rights correspond to the exclusive rights granted to creators of original works. IPR is an important topic because original works could stem from many sources and authors. Intellectual property includes copyright, trademarks, patents, designs, protection from passing off (one party representing their goods and services as those of another party who has an unregistered trademark), and protection of confidential information. Here, we introduce copyright, the branch of IPR relevant to all authors or content creators, and move on to discuss pertinent licensing issues in the context of digital libraries.

Copyright Law and the Digital Environment

Copyright comes into effect automatically when a work is created. In the UK, for example, content creators don’t need to register copyright; copyright law provides protection immediately after they
create a work. Copyright applies across media, protecting the expression of ideas that are "fixed" in a particular form. Literary, dramatic, musical, and artistic works as well as magazines and periodicals, sound recordings, films, databases, and computer programs fall under copyright protection. (For more information, see the UK Intellectual Property Office Web site at www.ipo.gov.uk/copy.htm.)

Copyright grants a copyright owner exclusive private rights to use the work through copying, distributing, broadcasting and/or modifying it. The work’s authors might hold the copyright. However, employers generally own copyright of work created during the course of employment. Some organizations choose to implement a lax IPR practice. For example, universities typically permit academics to take some form of IPR ownership; the academics are free to conduct IPR agreements for the sake of publications — for example, to give away the rights of research outputs to publishers. However, this practice shouldn’t be taken for granted, especially if the works concerned have financial and competitive implications.

The precise extent of exclusive private rights can vary across jurisdictions. (For further information on EU policy, please visit the European Commission’s IPR Helpdesk at www.ipr-helpdesk.org/index.html.) The 2001 EU Copyright Directive (Directive 2001/29/EC) is an attempt at standardizing, or harmonizing, copyright law among member states, keeping in mind certain modern requirements of the information society. As such it relates to educational materials that would be shared in the case of e-science (see the full text of the directive on the Foundation for Information Policy Research Web site at www.fipr.org/copyright/eucd.html). Considering a wider, international view, the well-established Berne Convention addresses the copyright issue, as does the World Trade Organization’s Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPs). (For more information, see the Berne Convention text at www.law.cornell.edu/treaties/berne/overview.html and the WTO’s TRIPs page at www.wto.org/english/tratop_e/trip_e/agrm7_e.htm.) The World Intellectual Property Organization (WIPO) also provides IPR frameworks that might be relevant. In particular, the 1996 WIPO Copyright Treaty sets out to clarify Article 20 of the Berne Convention in the face of new technologies, considering copyright in the context of ICT.

The challenges arising for e-science and the sharing involved in the use of e-infrastructures are relatively new and still in the process of being unraveled and addressed; work in this area is in its early stages, and understandings are only just emerging regarding how copyright law works in the digital environment. Consensus to support an international framework for the development and sharing of educational and training materials is essential, and lack of it will inhibit this development process. Focusing on the European context, JISC (Joint Information Systems Committee, UK) has voiced this conclusion about the need for harmonization on copyright in their draft comments on the EC Green Paper, ”Copyright in the Knowledge Economy:” “The full benefits of developments in ICT to support teaching, learning, research, and administration resulting from core Government and European investment is severely hindered by an archaic and out-of-step copyright system which lacks harmonization across Europe.” (For more information, see www.jisc.ac.uk/contentalliance.)

**IPR and Digital Libraries – Content and Depositor Licenses**

Digital libraries can provide students and teachers of e-science, and more specifically distributed computing, access to valuable education and training materials. A digital library is a collection of managed resources available over the Internet and, as in a "traditional" library, these resources are organized and catalogued for easy reference by users who share library content. There are various user roles and therefore different IPR considerations within the use scenarios of digital library provision. In general, all end users must observe and abide by the use conditions associated with the content under specific licenses, such as making the appropriate attribution (linking to the original works) when deriving works. Compared to other groups of users, depositors are confronted with the most significant IPR implications when they upload materials to a digital library environment. Depositors should only submit materials for which they hold the copyright or to which they are otherwise entitled to submit. Typically, depositors can be either the copyright holders — that is, the authors (content creators) — or designated proxies such as librarians, whose role lets them submit materials within the bounds of the law.
The options for handling ownership of materials in digital repositories correspond to the two generic mechanisms for sharing copyrighted content:

- licensing, in which depositors retain copyright ownership but grant either exclusive or non-exclusive licenses, and
- assignment, in which depositors transfer copyright ownership to the digital library or a designated organization.

Under the licensing mechanism, an exclusive license would give the licensee a right that no other person has (for example, exclusive broadcast). A non-exclusive license lets the licensor retain the right to enter into additional non-exclusive licenses with third parties. You also have several options when drawing up licenses; for example, you can employ a lawyer and legal firm or you can adapt standard licenses.

If the copyright is assigned to the digital library or a designated organization, the original copyright holder can’t change the license terms under which the material is made available at a later date. Experiences on the Jorum project (www.jorum.ac.uk), which provides a free online repository service for educators, have pointed to a marked reluctance among UK institutions to agree to transfer copyright ownership. The digital library would assume the publisher role in the case of assignment. Alternatively, the digital library can recommend an appropriate organization to which it would transfer copyright, but the depositor would be free to refuse this. Within the contexts of institutional employment, it’s often unclear who is permitted to assign copyright. It would be necessary to determine who within the universities or other institutions is responsible for giving permission to transfer copyright. The risk is reduced if the depositor licenses the work rather than assigning ownership, as organizations and institutions are more likely to permit licensing rather than ownership transfer. For a digital library facilitating content sharing among multiple organizations and international contexts, the solution might be too costly. This model might, however, be feasible for projects such as EGEE (Enabling Grids for E-SciencE, http://egee1.eu-egee.org), which has conducted prior IPR negotiation and obtained permission regarding copyright assignments from its participating institutions.

For digital libraries in general, a more practical approach is to employ a licensing mechanism via one of the following use cases during content submission:

- Depositors submit their own licenses or select from a list of off-the-shelf standard licenses such as the Creative Commons (CC, http://creativecommons.org).
- The repository specifies a default license to which depositors must consent.

If depositors supply their own licenses, then the repository should display the license along with the download and have consumers agree to it before they can download the materials. Using a single license makes it easier for consumers to understand and know their rights. However, most content producers already have licenses and might not want to distribute their work under a different license. Requiring a particular license could dissuade them from submitting material.

Licenses aren’t exclusive, so copyright owners can license their material under multiple licenses. For instance, Wikipedia articles are licensed under the GNU Free Documentation License, but individual contributions can be licensed under whatever copyright that author prefers, such as CC (the GFDL license also covers the individual contribution).

In addition to licensing, the digital library must address other considerations. The digital library must take reasonable precautions to ensure that depositors are entitled to deposit materials and that no third-party claims exist. However, as there is no widely used copyright registry such as exists for patents, checking this isn’t often possible. Instead, the most common approach is to have depositors agree to a declaration at deposition time, in the form of a “deposit license.”
While content licenses grant rights to consumers, deposit licenses are an additional requirement for a digital library facilitating a repository service. It's an agreement between the digital library and the depositors who are now responsible for obtaining the relevant permissions. By agreeing to this license during file submission, the depositors are declaring that they either are the owner of the copyright (the author) or have permission as a proxy to submit the content. The license sets out the terms under which the digital library will host materials on the copyright holder’s behalf. This includes getting permission to distribute multiple copies of the material over the Internet. It also includes other statements such as a disclaimer about IPR liability.

The ICEAGE Digital Library and Creative Commons Licenses

The ICEAGE (International Collaboration to Extend and Advance Grid Education) digital library (http://library.iceage-eu.org) developed out of a collaboration between the ICEAGE project and EGEE. The ICEAGE project’s primary goal was to support the growth of grid and e-science education within the EU context. The creation of an open repository of teaching materials served as one means to achieve this goal. The library houses distributed computing educational materials, including lectures, presentations and tutorials from the summer school series (International Summer School on Grid Computing, or ISSGC) and the first winter school (IWSGC 08), which we discussed in earlier installments of this department.

The current IPR solution for the ICEAGE digital library is to have a recommended license, but to let depositors use their own licenses if they prefer. Certain CC licenses are available, and depositors are encouraged to use these, but any license will be accepted. CC licenses are standard and useful for sharing educational content; non-exclusive CC licenses are used throughout the world, with the licenses ported to 34 international jurisdictions including the US and 9 major EU countries. Generic/unported licenses are also available.

All CC licenses follow a set of baseline rights and can be customized according to three conditions: whether the work can only be used noncommercially, whether it can be modified, and whether it must be licensed under the same license if derivative works are redistributed. The combinations of the optional conditions have resulted in six standard CC licenses, all of which require attribution to or acknowledgment of the copyright owner. Here are some characteristics of CC licenses:

- They enable and encourage reuse.
- The copyright owner must be acknowledged.
- Use restrictions are feasible: noncommercial use only, no derivative works.
- Derivative works can be distributed under the same license (ShareAlike) or under no specific condition (license doesn't include conditions for distributing derivative works).
- They're flexible and can be used as a template for other forms of license, such as the BBC Creative Archive License (restricts content use to UK only).

The ICEAGE digital library has used the Attribution-Noncommercial-ShareAlike CC license for all content the project generates. This license permits noncommercial reuse of material so long as the copyright owner is acknowledged and shared derivative works use the same license (BY-NC-SA) (http://creativecommons.org/about/licenses). ICEAGE also acknowledges that in some cases, depositors might wish to opt for other licenses. The digital library currently lets depositors pick other CC licenses and other project licenses such as one from EGEE.

The most important baseline requirement of all CC licenses is attribution, or proper citation of the authors of content submitted to the repository when it is reused. The CC license deed typically contains a hyperlink to the original work, which all derivative work must cite. Apart from this, the license includes recommendations for other citation mechanisms, such as the use of logos or branding (a generic statement acknowledging the work) within the derivative works. Currently, the CC license deed (http://creativecommons.org/licenses/by/3.0/) contains the names of the content creators (the Web site refers to authors or licensors — copyright holders) as a hyperlink to the original work if the deed is referenced from within the ICEAGE digital library (the original work details).
digital library has a mechanism to enable the CC deed to automatically include all creator names. The description of the original work (details page) can also include appropriate branding.

Providing e-science training and education is challenging. Not only does it depend on synthesizing the outcomes from a rapidly advancing domain, it must be accomplished in environments where knowledge largely originates in cross-institutional, multidisciplinary, and international contexts. Facilitating knowledge transfer in such environments is problematic and labor intensive. Making educational content accessible online with mechanisms for enabling content sharing, derivation, and reuse thus becomes vital to enable the achievement of economies of scale in large-scale collaborative production of content among e-science communities. This requires an understanding of IPR issues because the aforementioned actions involve the exclusive rights of copyright holders.

These IPR recommendations for digital repositories derive from our experiences with the ICEAGE digital library and our knowledge of copyright:

- Licenses produced by the Creative Commons should be recommended for digital repositories containing e-science education and training materials.
- Where appropriate, depositors should be permitted to upload their own usage licenses, but when this is done, the repository must make all reasonable efforts to explain this usage license to users downloading the material.
- A separate deposition license or agreement is also required and must be clearly displayed whenever a depositor attempts to upload new materials; depositors should have to explicitly agree to this license before they can upload materials.
- Any digital repository of e-science education and training materials should make it easy for users downloading materials to cite them. This can be done by including appropriate metadata containing, for example, the details of the author and copyright holders. This is particularly important if the materials themselves don’t identify the author, which is often the case.
- All metadata should be based on international standards such as Dublin Core to facilitate querying, harvesting, federation, and so on.
- Repositories should let consumers search materials by license type or with a license type filter in place to ensure that materials available under a particular license are easy to find. This should reduce the danger of, for example, a consumer accidentally modifying and distributing materials that were originally distributed under a license prohibiting this.
- Repositories should provide depositors with an FAQ or short explanation of common problems that might arise when depositing materials. For example, it should point out that the depositor’s employer might hold the copyright for the materials, and it should remind depositors to check the licenses of any material that they might have reused in creating their material.
- The repository must include a simple mechanism to remove or block materials and derivative works in case of disputes.

We’ve illustrated the importance of attending to IPR issues to allow for the smooth development and use of digital repositories to support education and training in distributed computing and e-science. But, fundamentally, we must also address the issue of curriculum development before we conclude this series with a look at broad strategy and policy recommendations. What’s the current state of curricula, and how can we improve on existing content and its delivery? The next article in this series will begin to answer these questions, presenting work carried out at international workshops and curriculum development sessions supported by ICEAGE, the e-IRG (e-Infrastructure Reflection Group), and OGF (Open Grid Forum).
Boon Low is systems developer at the National e-Science Centre, Edinburgh. Contact him at boon@nesc.ac.uk.

Kathryn Cassidy is a member of the Computer Architecture and Grid Research Group at University of Dublin (Trinity College). Contact her at Kathryn.Cassidy@cs.tcd.ie.

David Fergusson is deputy director of training, outreach, and education at the National e-Science Centre, Edinburgh. Contact him at dfmac@nesc.ac.uk.

Malcolm Atkinson is director of the e-Science Institute and e-science envoy at the National e-Science Centre, Edinburgh. Contact him at mpa@nesc.ac.uk.

Elizabeth Vander Meer is the education and training policy officer at the National e-Science Centre, Edinburgh. Contact her at evmeer@nesc.ac.uk.

Mags McGeever is a legal services associate at the Digital Curation Centre, Edinburgh. Contact her at mags.mcgeever@ed.ac.uk.