

# **Submission to the Gowers Review of Intellectual Property on behalf of the Digital Curation Centre at the University of Edinburgh**

## **About the Digital Curation Centre**

Scientists, researchers and scholars across the UK generate increasingly vast amounts of digital data, with further investment in digitisation and purchase of digital content and information. The scientific record and the documentary heritage created in digital form are at risk from technology obsolescence, from the fragility of digital media, and from lack of the basics of good practice, such as adequate documentation for the data.

Working with other practitioners, the Digital Curation Centre supports UK institutions who store, manage and preserve these data to help ensure their enhancement and their continuing long-term use. The purpose of the centre is to provide a national focus for research and development into curation issues and to promote expertise and good practice, both national and international, for the management of all research outputs in digital format.

For more information about the Digital Curation Centre see [www.dcc.ac.uk](http://www.dcc.ac.uk)

## **GENERAL QUESTIONS**

### **2. How IP is used**

#### **(j) Have you encountered patents or other IP rights being used defensively?**

There are examples with the LOCKSS (Lots Of Copies Keep Stuff Safe) project which could be seen as defensive use. Publishers have denied the ability to use the system to preserve licensed content, despite LOCKSS being specifically designed to be publisher friendly. No one can use LOCKSS to get access to content they did not have a license to access in the first place. It does however mean that libraries can continue to get electronic access to the back issues they have paid for even if they cancel their subscription. If they were not LOCKSS users cancellation would often mean loss of access to back issues. Access to back issues may be the reason for the publishers' refusal to take part in the LOCKSS system as they wish to use non-access as a bargaining tool for selling other journals or bundles of journals. For more details about the LOCKSS project please see <http://www.lockss.org/>.

### **3. How IP is licensed and exchanged**

#### **(a) How easy is it to negotiate licenses to use other's IP for commercial or non-profit purposes?**

Please see answer to General Question 2(j) above and 3(e) below.

#### **(c) How easy is it to use others' IP for research purposes? Have you experienced difficulty around research exemptions?**

The exemption for research allows limited copying but research methods currently are much more complex. It is no longer a case of photocopying a journal in a library to read and develop ideas from. Researchers want to share large amounts of data and use digital formats, which have the potential for unlimited copying.

**(e) Are there barriers to licensing IP on grounds of cost? What drives these costs?**

Yes there are barriers. In particular, 'by the drink'/unit cost pricing is driven by pressures from wealthy markets and denies access to research users (who typically do not have discretionary budgets from which they can order such materials). An example (from 2001/2) was provided by the Electronic Article Supply Service (EASY) Project piloted by ingenta, Lancaster University and the Joint Information Systems Committee (<http://www.ingenta.com/easy/>). EASY aimed to set up a system for "electronic inter library loan" (ILL) whereby the publishers shared some revenue from the ILL (a source of revenue they would not otherwise have access to). Comparatively few people now use ILLs because it is paper based and slow, an electronic version would be more attractive. However, many publishers rejected the project preferring not to be involved in ILLs but to retain the likelihood of being able to charge people their full "copyright" price for electronic versions of articles on a pay per view basis at high rates. These costs make those articles inaccessible to many researchers but the publishers continue with that model because they can successfully command those prices in the wealthier markets. Lack of publisher cooperation essentially killed the project.

Nowadays costs are also driven up by the concerns of intellectual property owners, who in an increasingly digital world, and given the ease of copying, fear that the first copy of a digital work they sell may be the only one.

## **SPECIFIC ISSUES**

- **Current term of protection on sound recordings and performers' rights**

**(a) What are your views on this issue?**

The DCC suggests that an extension of the term would only be acceptable if a system of registration is put in place to qualify for that extension. In this method a right owner would register by paying a fee to get the extension which would avoid an automatic extension being the default position. This would allow owners of commercially valuable works to benefit from the extension, but at the same time, let less 'valuable' works fall into the public domain where, despite their lack of potential for exploitation, they may still be used as a resource to encourage further development.

- **Copyright Exceptions – fair use/fair dealing**

**(a) What are your views on the current exceptions on copyright law?**

Consultation with non-legal staff working within HE confirmed that many are not confident that they understand the exceptions. Simplification is required.

In addition, many of the boundaries of the fair dealing exception are defined by case law. People cannot simply look to the statutes. The majority of users of IP and of the fair dealing exceptions are non-legal people who are not comfortable having to decipher case law.

**(b) Could more be done to clarify the various exceptions?**

Definitely. The exceptions need to be rethought for the digital age. See 'Other Issues', section 1 below.

In addition, 'non-commercial use' (as per section 29(1) of the CPDA) needs to be expanded upon, as it can often be difficult to determine whether something falls into this category.

**(c) Are there other areas where copyright exceptions should apply?**

For curation and preservation. See 'Other Issues', section 2 below.

**(d) Are the current copyright exceptions adequate or in need of updating to reflect technological change? Should UK law include a statutory exception for fair use?**

The current copyright exceptions are inadequate and they are definitely in need of updating to reflect technological change. See 'Other Issues', section 1 below.

In relation to private fair use we believe one option is that the work might be considered as independent of or separate from the carrier/format on which it is sold. Multiple copies of a work under the control of one person should be allowed, provided that only one is used at a time. However, they would not be allowed to transfer the work to another person without destroying their existing copies.

**(g) Are there issues concerning the archiving of material covered by copyright?**

Yes. In the first place, digital preservation (a more appropriate term than archiving in most cases) fundamentally requires the making of (preferably many) copies and potentially of derivative works ("migrated" across technology changes). This requires copyright permission which is rarely easy to obtain, and this means that continued access to many works is at risk.

Another example can be found in an Erpanet (Electronic Resource Preservation and Access Network) case study of the Theater Instituut Nederland (TIN) ([http://www.erpanet.org/studies/docs/erpastudy\\_TIN.pdf](http://www.erpanet.org/studies/docs/erpastudy_TIN.pdf) - Erpanet registration required, free). While this is not a UK body the situation would be similar here. In summary, TIN had copyright issues that prevented them from providing wider access to their resources by putting them online. They could provide a catalogue entry and description online, but in order to access some of the resources people had to actually visit the institute in person. This is a waste of valuable resources that, given today's technological capability, could be made easily available to many people

A second example is provided by PrestoSpace, a project developed to provide technical solutions and integrated systems for digital preservation of all types of audiovisual collections. The selection process for audiovisual content is often accompanied by serious rights issues, which tend to be complex. Audio visual content can contain a number of rights in the one work. This can prove difficult to manage and rights clearance remains a major unresolved problem in this area. The rights to use archive content are a significant factor when determining what to preserve.

We believe libraries, registered preservation repositories and other memory institutions should have the rights to copy and modify works (for example for migration purposes) that they either have licenses for or have legitimately under their care, provided that the copies or derivative works are used under essentially the same terms (unless rights have meanwhile expired). [Migration is the periodic transfer of digital materials from one configuration to another or from one generation of technology to another. Its purpose is to preserve the integrity of the object and retain the ability of users to retrieve, display and otherwise use that object in spite of constantly changing technology.]

Digital Rights Management (DRM) is also relevant here. More thought is needed on the requirements of DRM to ensure content is not 'locked-up'.

- **Copyright – digital rights management**

- (a) **Do you have a view on how the use of DRM technologies should be regulated?**

DRM must not be used in such a way as to disable use of legitimate exceptions or to affect the availability of a work after the expiry of copyright protection. In addition, the copying or modification requirements of preservation (whether through migration, emulation, re-creation or other technique) must be taken into consideration and regulation must ensure that DRM technologies do not unfairly restrict these.

- **Copyright – orphan works**

- (a) **Have you experienced any difficulties in identifying the owners of copyright content when seeking permission to use that content?**

Digitisation programmes often have problems with orphan works. This results in many works in which there is no commercial interest being unavailable for research or teaching use. Orphan works can present problems of ambiguity for a potential user or curator and this can create the fear of litigation. Although use of copyright works can be seen as a risk management business, many institutions are unable to or do not have the resources and expertise to quantify and therefore take those risks. This means that they are left unable to use not only the resources that are still protected (perhaps in spite of the fact they no longer have commercial value) but also those resources that may no longer be protected but that they cannot prove to be so.

Orphan works may even be a problem for content intended to be freely available, such as web pages. For example, the UK Web Archiving Consortium requires a form to be signed asserting copyright status before web pages can be archived. However, the exact copyright status of any web page, given a long history of modification by authors perhaps from several sites, can be impossible to determine. So in theory one should not sign such a form, and the page cannot be archived.

- (b) **Do you have any suggestions on how this problem could be over come?**

The recent work by the US Copyright Office and their conclusion relating to Orphan works in the US provides an excellent starting point for considering possible avenues. We would suggest that the Gowers Review Team commission a full study to investigate how this issue might be dealt with in the domestic context taking into account all the interests involved.

## **OTHER ISSUES**

**1. Relevance of IP laws in Digital Age** – This is touched upon in the questions relating to copyright exceptions. Much of the law was developed in an analogue world but is now increasingly applied to digital formats. This can make application confusing, uncertain and unhelpful. The terms in the CPDA do not relate to how works are used today. For example, the concept of ‘a copy’. Copying is defined in the Act as “reproducing the work in any material form. This includes storing the work in any medium by electronic means.” (s17). However that does not relate accurately to the working of computers and clarification is required in this area. Indeed a significant re-think may be required. One contributor gave the opinion that “the notion of ‘making a copy’ is dead”. It is felt that computer users are often not clear if they are copying and if so how much they are copying. Many have no concern about how data is moved about and stored in various places (a sequence of web caches, not all on the user’s computer, for example) in order for them to see it on their screen, even though they realise it is likely that ‘technically’ copies are being made of that information.

Further, in relation to the Internet, every time a file is accessed online it is essentially downloaded and this could be interpreted as a copy being made. If the user saves it onto their computer, a further issue arises. They do not know if they infringe. We are also of the understanding that some programmes are one step ahead of the user. For instance, Google Earth copies what it thinks the user will want next ahead of being asked for it. In such a case there may be information that the user is not even interested in seeing, but by using Google Earth they have made copies of it.

Back up copies might also be considered. These are permitted under section 50A of the CPDA but limited to a computer program which it is necessary for a person to have for the purposes of their lawful use. Researchers frequently ‘synchronise’ a number of computers for safety and reasons of collaboration. Even if all the computers involved in the synchronisation belonged to them, under the present legislation this would be an infringement. Further, at the university a back up ‘copy’ is made of certain drives on staff computers every night. These backup copies are kept in a cycle (“grandfather, father, son”, for example), and some will be kept off-site for arbitrary periods. Once this copy has been made it cannot be expunged from that record. Is this an infringement?

Realistically researchers are infringing the Act on a daily basis. This is not a healthy environment in which to support innovation.

It would appear that a new definition of copying is needed. Copying should only be considered to have taken place should certain actions have been carried out such as, printing data, storing an easily accessible version on a computer or passing it to someone else although it is appreciated that a definition based on specific actions may become outdated as technology continues to progress. Alternatively, perhaps less emphasis should be placed on the concept of copying and effort made to identify a different central concept on which to base IP rights and which better serves the digital era.

**2. Curation and Preservation** – Copyright legislation impacts on the ability of organisations to undertake digital curation or preservation activities (from hereon we will use the word “curation”). This can affect most areas of digital curation, including copying and reformatting, preserving and providing long-term access to stored material in archives, repositories and libraries. Inherent in digital curation is the need to emulate and migrate. This is not always possible within the current confines of intellectual property law. The CEDARS project (CURL Exemplars in Digital Archives) has made some observations in this area. (For more please see <http://www.leeds.ac.uk/cedars/colman/CIW03.pdf>). Please see also the Copyright and Licensing for Digital Preservation project at Loughborough University ([http://www.lboro.ac.uk/departments/ls/disresearch/CLDP/Project\\_reports.htm](http://www.lboro.ac.uk/departments/ls/disresearch/CLDP/Project_reports.htm)).

The concern is that the legal complexities or barriers involved in curating digital objects will restrict centrally organised curation initiatives from curating resources and that owners of those resources will deliberately or inadvertently, through a simple failure to act, destroy the resources thereby removing the potential benefit of their future use.

Barriers to digital curation are also caused by widespread uncertainty in the legal environment which is in flux as it moves to accommodate rapidly changing technologies. It is difficult to make long-term plans and take decisive curation action against such a background. In such a way a single major overhaul of the legislation that is designed to withstand future technological change would be of benefit to digital curators as opposed to piecemeal changes.

**3. Data and Databases** – Much of science depends on the collection of data and other information and the subsequent re-use of that material. This data is generated on an exponential basis and held within vast databases. The progress of science depends on the reuse of that data for a variety of purposes. It can be difficult for the non-scientist to appreciate the size and importance of these databases to the scientific community. The database right has the potential to cause problems here, as has the practice of ‘locking’ data within private databases to which the key may only be available at a price.

Many databases within research environments and HE are produced by extracting content from other databases. The value lies in bringing the information altogether in one place and restructuring it into a more useful format. There is a potential infringement of the database right here (depending on whether the database which data was taken from qualifies for the right and how much is being extracted or reutilised). Also, these databases often involve the work of a large number of contributors, so rights ownership of parts of the database may be unclear. Issues of identification of the various rights holders arise in this circumstance. Clarification would be welcomed.

It is questionable whether the economic imperatives that drive the entertainment and commercial database industries are best suited to the progress of science. Even reverting to the most basic justifications, scientists, particularly where publicly funded, often have little interest in the incentive copyright and the database right are designed to engender (see below for more on this). Proprietary claims made on the contents of databases can only inhibit the free flow of information and as such seem the antithesis of the sharing, collaborative ethos on which science thrives.

**4. Motivation/Incentivisation of Academics** – Part of the justification of copyright is to give the creator a property right thereby creating a financial incentive to produce. However this argument doesn't necessarily follow in academia (leaving aside the text book market for the moment). Academics traditionally (and increasingly are required to) assign the copyright in their scholarly work to publishers who then make money from the work by including it in journals, which they sell (or in the electronic environment, lease access to). But the publishers do not pay for the articles so the academics do not benefit financially. The key motivator for academics is attribution – they want people to reference their work. To this end, even after assigning copyright to a publisher, some academics will post the same article on their personal web page, despite the fact they have assigned copyright. The financial incentive of copyright does not operate here to the same degree.

**5. Open Access** – is a matter of importance to those working in research. The report of the Science and Technology Committee – 'Scientific Publications: Free for all?' (<http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/399/39902.htm>) and the RCUK Position Statement on Access to Research Outputs (<http://www.rcuk.ac.uk/access/index.asp>) are significant in this area. We do not believe that these have gone far enough in advocating open access. In particular, although the RCUK Position Statement says that a copy of any journal article or conference proceedings should be deposited in an appropriate e-print repository it makes this "*subject to copyright and licensing arrangements*". Further the Government's response to Science and Technology Committee report was largely negative. The RCUK Position Statement highlighted the importance of long-term preservation, something we the DCC are specifically interested in. As discussed in Section 2 above, the obstacle that copyright presents in this area needs to be addressed.

**6. Open Access for Data and Databases** – Less well documented and discussed is open access as it relates to data and databases, particularly in terms of economic sustainability. This is of huge relevance to the research community. The open access model may be viable for journal articles but databases are dynamic and require continual maintenance, which necessitates financial input even if this is minimal. If databases are to be made available free of charge, who is going to subsidise them? What would be the best economic model? We recognise that this is a question of policy and economics but in order to answer it we must consider the appropriate legal infrastructure/environment to support that model.

**7. Public Domain** – This is a hugely valuable resource for scholarship, teaching and indeed for the information economy. Any process of creation depends not only on the existence of a wide variety of sources on which a creator can draw, but also upon the accessibility and re-usability of those sources. In this process, works are often cumulative. The law is there to mediate between property rights and the commons. Denying works entry to the public domain or allowing confusion around dates of entry to the public domain denies the public access to a valuable resource. We would suggest considering not only ensuring that intellectual property rights do not lean too far in favour of the property right owner (in terms of duration, scope etc) but also creating a simple way in which authors may choose to dedicate their work to the public domain and that this be recorded centrally and made clear to potential users of the work. Addressing the issues created by orphan works will also be valuable.