

## SUGGESTED THEMES

This document features suggested themes for submissions to the consultation, which is intended to serve as a guide for those interested in responding from the perspective of educational institutions, researchers or academics. Page and question numbers refer to the relevant section of the [consultation](#).

**For further detail please see submissions from educational / research and library bodies to the 2011 Hargreaves Review:** <http://www.ipo.gov.uk/ipreview/ipreview-c4e.htm>

### **The Hargreaves Review – Why Exceptions and a Modern Technologically Neutral Copyright Law is Important**

#### **Data and Text Mining** (p. 79-82 / Q.77)

*Please consider giving evidence where permissions to text mine have been refused, if granted unsatisfactory, took a long time or explaining why asking permission did not occur in the first place.*

Exabytes of data are produced each year. Humans cannot read and make sense of vast amounts of information on this scale and are increasingly using computers to analyse it for them – a technique known as data analytics, or text and data mining.

Under the current copyright system, researchers are often unable to use computers to analyse journals or other digital information (e.g web pages) to which they already have lawful access. Copyright law does not regulate a human reading a journal, but because a computer has to make a copy to read or mine the journal for you, copyright law becomes relevant (as well as any contract that governs access to the content. N.B. Very few standard contracts have terms and conditions for text and data mining.)

The Wellcome Trust for example estimates that over 80 per cent of the material housed in UK's main medical research database (UK PubMed Central) is unavailable for legal text and data mining as there is no copyright law exception, and the material is not open access.

Text-mining can already be legally undertaken in both the US and Japan. If we are to compete internationally, we cannot stand by and watch the rest of the world benefit from these technologies, while our own institutions are prevented from doing so by an outdated and complex copyright system.

I would also strongly agree with the Hargreaves submission made by AstraZeneca<sup>1</sup> that given that data analytics is to be performed on material that an organisation has legal access to (which in the case of scientific publications will be purchased material), revenue will not be harmed. On the contrary it encourages greater usage of journal and book content, and in a period of economic difficulty, helps justify the very large investment made in published information by UK universities, research institutes, and bodies like the National Health Service.

In summary, these text-mining tools enable us to do in a few minutes something that would take a person months or years to do. The proposed exceptions will enable researchers to use technology to its full potential to extract facts – accelerating discovery and its application to find new cures and treatments, and to deliver wider research and economic benefits.

In order to give the UK's R&D sector international competitive advantage, prior to an amendment at EU level to allow commercial data analytics, the CDPA should be amended as follows:

---

<sup>1</sup> <http://www.ipo.gov.uk/ipreview-c4e-sub-astrazeneca.pdf>

1. The extraction of facts and information from all works subject to copyright and related rights that the user has lawful access to is not an infringement. (S. 29 – Fair Dealing for Research and Private Study)
2. The making of copies specifically for the process of data analytics is not an infringement provided that the computer generated copy created for the purposes of data analytics is not communicated to the public. (S. 29 – Fair Dealing for Research and Private Study)
3. This, and any other exception cannot be overridden by contract law. (Chapter III – Acts Permitted in Relation to Copyright Works)

### **Digital Archiving** (p.70-72/ Q.72)

Currently the law does not allow sound and film to be legally preserved, as these works are currently excluded from the existing preservation exceptions. The law also does not allow multiple copies to be made while digitally preserving material – whereas format shifting is essential as is the use of multiple servers.

The proposed exception to allow digital archiving of all types of material is therefore a must.

### **Extend research copying (“fair dealing”) to sound and film, and allow librarians to make copies on behalf of researchers.** (p.74 – 77 /Q.75)

Currently the law does not allow sound and film to be legally copied for fair dealing purposes – namely private study and non commercial-research. This means that a scientist cannot make, or have made for them, a fair dealing copy of natural history programme in the course of his or her studies. Similarly a researcher of music can currently make a fair dealing copy of the lyrics to a song but not make a copy of the sound recording itself.

In the age of computing this is illogical and out of step with how people study and research in 2012.

### **Widen the existing copyright exception for quotation to allow information, analysis, argument or comment.** (p.105 / Q.94.)

A British Academy study a number of years ago showed the problems faced by academics when trying to quote from and reuse in-copyright works. (Copyright Research in the Humanities and Social Sciences. British Academy. 2006)

A clearer and stronger environment for criticism and review would support education and learning in the UK.

### **Facilitate mass digitisation of post 1870 in-copyright materials, including works whose copyright owner cannot be found (“orphan works”).** (p.14 – 39 / Q.1 – 43)

Much information and knowledge sits on the shelves of libraries. Enormous value can come from digitising this material and putting it on the web. We believe therefore that solutions for orphan works (works whose copyright owner cannot be identified or traced) and mass digitisation of content would very much support education and allow new paths of research to be opened up. (see p.46 of the Hargreaves Report and a case study relating to Malaria Research: <http://www.ipo.gov.uk/ipreview-finalreport.pdf>)

(N.B. Mass digitised content can also be data mined.)

**Update the existing copyright exceptions for educational establishments and teaching** (p.89-95/  
Q.85-89)

UK teaching exceptions are limited, and for example currently do not allow use of copyright material in the context of VLEs, do not allow examinations including third party copyright to be included in an institutional repository, the definition of an educational establishment is defined narrowly etc.

**Protecting Copyright Exceptions** (p.116-119 / Q.103)

The exceptions outlined above are important for education and research and in order to be effective they must be robust and enforceable. There is no point in providing a medical researcher the right to text and data mine in copyright law, or the right to use a sound recording for teaching if this is subsequently prevented by publishers' and third party contracts.

It is therefore essential that current legislation is amended to ensure that contracts no longer override copyright exceptions

We urge the government to implement the Hargreaves recommendations as swiftly as possible. Many of the above amendments are secondary legislation and should be done during the next session of parliament (2012 – 2013).