Incorporating new developments in daily licensing practice: Setting conditions for open access, patron driven access, and data & text mining

Report on Knowledge Exchange Workshop 22 November 2011, Munich

Executive Summary

New developments in higher education and research are having their repercussions in daily licensing practice. Examples are demands for perpetual access, usage of licensed content in course packs or virtual research environments, text mining and open access to publications. At the Knowledge Exchange workshop on Licensing Practice 20 experts discussed how these new developments could be incorporated in licensing.

The workshop consisted of four presentations on current developments in licensing followed by three parallel breakout sessions on the topics open access, new developments and data and text mining. This led to a lively exchange of ideas. Especially the aspect of data and text mining provided valuable insights in how this could be incorporated in licensing. The breakout groups fed back their findings and the workshop was closed Wilma Mossink.

The Knowledge Exchange Licensing expert group will work on how to implement all the provisions discussed. The suggestions will be incorporated in a meeting with publishers scheduled for March 2012. The various suggestions will be also shared with other international organisations working in this field.

Introduction

Experts from the four Knowledge Exchange (KE) partner countries came together in Munich. The meeting provided input to the participants on which new developments are taking place in higher education and ideas were explored how to incorporate these into licensing practice.

Several elements for a model licence containing broader access requirements were presented and discussed. The workshop was attended by participants from the KE Licensing Expert Group and experts from libraries and funding agencies involved in licence negotiations.

Provisions on continuous access, long term preservation and course packs

Wilma Mossink from the KE Licensing expert group started the workshop with a presentation on model provisions for continuous access, long term preservation which had been developed by the expert group. She identified the three elements addressed in the provisions: long time preservation, continuous access in case of trigger events and continuous access in case of termination of the agreement.

In order to ensure long time preservation of content, publishers should be obliged to archive in at least one of the established archiving initiatives for e-journals. Points of discussion are if the archiving initiative should be mentioned by name (e.g. CLOCKSS, Portico) and if just one is sufficient. Trigger events which would qualify for a fall back on these archives are: catastrophic and sustained

1 All presentations are available for download at: http://www.knowledge-exchange.info/Default.aspx?ID=488
failure of the publisher’s delivery platform, the dismissal of a title or if the publisher no longer offers the back issues or stops operation itself. If the agreement terminates, access to the earlier material should be possible without charge.

A further point of interest is the inclusion of content into course packs. Course packs are defined as being assembled by members of staff of the institution for use by students for the purpose of training, education and instruction either in printed or electronic form for a virtual learning environment that supports teaching and learning in an educational setting. There were discussions on inserting wordings about a removal after certain years and on the possibility for institutions to sell course packs to students. The used materials should carry appropriate acknowledgement of the source, title, author and publisher.

In the discussion the demands to publishers and technical development were specified. Long time preservation was described as taking the form of a dark archive in the case of trigger events, besides offering a hosting solution to deliver the content. Both aspects should be covered in the licences. Not all libraries want to be member of Portico due to its governance model. Portico is a US organisation and national network solutions would allow for a certain independence from this governance model.

**Strategies for Open Access provisions in licences**

Kristine Hillenkötter from the state and university library in Göttingen gave an overview on open access provisions in licences. Examples were initiatives from ARL, JISC, UKB and DFG. The German alliance licences are based upon a national open access policy of DFG and the alliance of German research organisations. When publishers accept an open access clause in their licence DFG funds 25% of the licence fee. This allows authors from authorised institutions, or institutions to promptly store their articles appearing in licensed journals generally in the form published by the publisher in a repository of their choice and to make them available in open access.

In the period of 2011-2013, 12 licences for journal collections were mentioned. During a session held at ICOLC Europe Conference in Istanbul, SPARC Europe and COAR discussed open access options and existing barriers for consortia. Among the challenges and opportunities for a national open access policy is the need for finding exact definitions and phrasings to reflect open access properly in a clause in the licence. Business models and funding concepts have to be developed.

The discussion showed that if authors give rights away, the open access clause in a licence will not work. The co-authors should be asked too depending on the legal framework. The contracts with authors should address these points. The embargo period is mostly 6 months, and the pdf usage is not a critical issue. Self-archiving of the published version was more attractive for scholars in Germany.

**New Developments and Licences**

As a kick off for the breakout sessions, Nol Verhagen from the University of Amsterdam spoke on the nature of a licence. He described it as a temporary right to use content or service under agreed conditions that restrict the use upon a period, for a well defined group of users, purposes and circumstances. Licensing contracts define minimum rights to the users that they don’t need ask permission for, and they overrule copyright as they are contracts. Open access and Creative Commons are new perspectives in this environment. Access should include remote access and offer no restrictions to access by mobile devices.

Usage should include downloading, printing, saving and possibly specific purposes like text mining and re-use. The storage should include post cancellation and perpetual access.
arrangements in the future. The prospect of sharing should consider usage in social networks and virtual research environments.

**Data and text mining**

John McNaught from the National Centre for Text Mining at the university of Manchester offered an accessible introduction into text mining and its implications for access. He provided a clear description of what text mining means. Processing huge amounts of texts enables discoveries human effort could not achieve. There is a need to use the full text, because only 8% of scientific claims found in the full article also appear in its abstract. However copyright restrictions do not cover text mining. There is an inconsistency that the human reader can search, but he cannot use a machine to do the same job for this. Trademark and copyright should get customised to the demands of text mining, and there should be a copyright exception for information analyses.

After these presentations three breakout groups investigated three topics. They provided plenary feedback on their discussion.

**Breakout group: Data and text mining**

A researcher is less interested in analysing the material of only one publisher, therefore data mining is useful only if a publisher cooperates with other publishers. As many trusted sources as possible should be consolidated. Access to institutional repositories is useful if the content is closed access otherwise. The data subset should be limited. In discovery service publishers cooperate with vendors to make the full text searchable, not only the abstracts. The PubMedCentral open access section could be copied and used for text mining. If licences do not have a clause for data and text mining, it should be renegotiated. It would be very useful to put an effort into setting up model licences, and to cooperate with publishers to get model licence for data and text mining. A working group should be established to discuss if a model licence should be an annex or an inclusion for existing licences.

**Breakout group: Incorporating new developments in licences**

This breakout group discussed both perpetual access and new practices in accessing content. Perpetual access to licensed content should be covered by licences. Long term preservation should be ensured by including an option that stores material in a dark archive. Portico was established as a default provider. However, digital preservation is not a question for the legal deposit libraries only. A new European infrastructure could offer storage of the content for the long term. There was a clear plea for good quality metadata.

The breakout group also discussed usage of licensed content through mobile devices or through social media. Students are not aware of copyright issues and should be informed. Conflicts arise from access through shared IP addresses and the question whether alumni or retired researchers are granted access by the licence. When re-using metadata of commercial vendors or suppliers the territory of open and linked data is entered. Compensations should be given if publishers withdraw titles from an e-book deal.

**Breakout group: Open access**

The last breakout group discussed open access clauses in licences. These should be implemented in subscription licensing agreements. This offers good prospects with large consortia but is harder to negotiate for small institutions. A consistent language for clauses for all countries would be very helpful. A working group should work on wording that is applicable for copyright laws in different countries. A guideline for authors is needed, and they should be contacted when the article is to be published, not afterwards. Persistent author IDs should be used, and a way to get the articles from publishers should be established. The terms of use should cover various open access aspects like copying, re-using and distributing.
Conclusions
Wilma Mossink summarised the most important conclusions of the day. The KE Licensing expert group will work on how to implement all the provisions discussed. The breakout group on text mining provided good questions for the planned meeting with publishers in March 2012. A model licence needs to be drafted and the aspect of discovery services is important. The outcomes of the discussion on new developments in licences should be discussed with LIBER. Conditions should be formulated why perpetual access is needed. The usage of legal deposit of content is another important topic. The European Commission has set up the Lisbon Agenda arguing for the growth of Europe. European infrastructure could solve the deposit challenge and provide a solution but this would require a good argument for funding. Finally the breakout group on open access clauses argued that an intertwining of the KE Open Access and Licensing groups could build a task force willing to work on a model licence.

This report was written by Isolde Förster, Bavarian State Library
Photograph by Keith Russell, Knowledge Exchange
The report is made available under a Creative Commons CC-BY 3.0 licence.