Report on Knowledge Exchange meeting with publishers on provisions in licences

Incorporating new developments in daily licensing practice: setting conditions for open access, patron driven access, and data- & text mining
Copenhagen, 9 March 2012

Summary
Fourteen representatives from national consortia (from Denmark, Germany, Netherlands and the UK) and publishers (Elsevier, SAGE and Springer) met in Copenhagen on 9 March 2012 to discuss provisions in licences to accommodate new developments. The main outcome was that it would be valuable to have a standard set of clauses which could used in negotiations, this would make concluding licences a lot easier and more efficient.
The comments on the model provisions the Licensing Expert group had drafted will be taken into account and the provisions will be reformulated. Data and text mining is a new development and demand for access to allow for this is growing. It would be easier if there was a simpler way to access materials so they could be more easily mined. However there are still outstanding questions on how authors of articles that have been mined can be properly attributed.

Introduction
In November 2011 the Knowledge Exchange (KE) Licensing Expert Group organised a one-day workshop on new developments in daily licensing practice1 in Munich. This workshop brought together experts from the four knowledge Exchange partner countries to explore ideas and exchange information about incorporating current developments in serials and other continuing resources in provisions for licences. Basis for the discussion were the licensing provisions the Licensing Expert Group has developed during its existence. The participants of the workshop encouraged the Licensing Expert Group to arrange for a meeting with several publishers to discuss the incorporation of the licence provisions with publishers in spring 2012. This workshop took place in March 2012 in Copenhagen and had as goal to discuss the inclusion of the Knowledge Exchange licence provisions in the licences of publishers to harmonise their licences in the partner countries of Knowledge Exchange.

Plenary: Introduction
Wilma Mossink from SURF welcomed participants to this workshop and explained the background to this meeting. Due to new antitrust rulings the discussions with the three publishers at the table were to take place in break out groups.
She first explained about the activities of the expert group. The group investigates the shift from toll access to open access. One of the activities has been a pilot investigating multinational licences. This pilot taught the partners that they use different models for the procurement of content and it is important to work on standard models for the parties involved.
Wilma Mossink presented background and ideas of the model provisions the Licensing Expert Group has developed and which were discussed at the workshop in November 2011. The model provisions had been shared prior to the workshop2. These included model provisions on continuous access and course packs. The model provision on open access was not yet drafted awaiting the discussions in this workshop.
Regarding continuous access there were some question marks. Is storing the materials in just one of the established archiving initiatives sufficient? Is the term ‘established’ clear enough or should e-journal archiving initiatives be mentioned by name? It was noted that the named trigger events are in line with the trigger events named by Portico. However, do these cover all the possible trigger events that could occur? And after a trigger event, under which licence agreement should the materials be made available?

---

1 A report on this workshop can be found at: http://www.knowledge-exchange.info/Default.aspx?ID=485
2 The model provisions discussed are an annex to this report
Regarding course packs the question was raised whether the definitions correctly reflect the meaning for all parties and whether the definitions should be written out as thoroughly as possible.

The presentation was followed by a plenary discussion on some general considerations regarding continuous access. Regarding the access to backfiles, there is a transfer working group already addressing this. There is also a code of practice already in place.

Offering continuous access is especially a challenge for the publisher when a learned society switches from the publisher. The challenge lies in the fact that you cannot licence what you do not own. The publisher has therefore to be careful to cover everything precisely in arrangements with a society. Publishers do work with checklists, but items still have to be negotiated with the society. This is especially complex for old journals with a history.

Licences also have to be transparent regarding archival rights. JISC has worked on an entitlement registry, to gain clarity on archival rights. So far, it has proved very difficult to collect a clear record. Publishers do not always have a record and do not keep track of archival rights. The Dutch National Library is running a project setting up a database around big deals. This also showed that it is important to have a proper record of the titles.

It is agreed that there is a need for a clear rights administration and it is not practical to maintain this by hand. There are also automated lists using ONIX-PL. However it is proving a challenge to connect with library information management systems.

The question was raised whether the transfer of a title could be classed as a trigger event? It has an attraction, but a learned society might not allow this. The publisher would want to make the decision if it could be classed as a trigger event. However it should be clear if a transfer of a title is a breach of contract. In that case it should trigger a refund and this has to be transparent in the licence. If the publishers fails to provide an archival right there should be a financial penalty.

Publishers are working on providing access to licensed materials in other formats for people with a disability. National handicap organisations work with publishers and provide versions in other formats. The problem does seem to be fading away as new software allows reading aloud. This is not relevant to older formats. There is a growth in numbers of students asking for materials in other formats.

The publishers split up and went into separate rooms to discuss the terms and conditions in detail.

**Strategies for open access provisions in licences**

Over the last few years several initiatives have been undertaken to create a situation of open access through wording in consortia licences. One of the examples is the DFG alliance licences. In the discussions the underlying author agreement was also discussed. One of the issues discussed was how metadata can show the provenance of an article.

Lorraine Estelle from JISC presented the findings of the open access fees project\(^3\). In this project funders, institutes of higher education and publishers where interviewed. From the perspective of the institutes of higher education there where two challenges regarding transparency:

- **Double dipping**: it is unclear whether publishers are charging institutions twice
- **Metadata**: it is unclear who is the funder and who is the author. What is the licence attached to the article? It is also important to be able to track the funding. To document this there is a need for clear institutional, object, and author identifiers.

---

\(^3\) See: [http://www.jisc-collections.ac.uk/News/Open-Access-Fees-Project/](http://www.jisc-collections.ac.uk/News/Open-Access-Fees-Project/)
For publishers there is a challenge in collecting the micropayments for each article. Institutions welcome funding, yet at an institutional level it is complex to manage the transfer of article processing fees. A recommendation from the report was that a body should be identified to manage the documents and standards required. At present the processes and workflows at institutions often do not seem to be in place to support the transition. CROSSREF is working on a pilot project FundRef\(^4\) to crack the problem regarding funding identifiers. This involves publishers and funding agencies, to provide a standard way of reporting funding sources for published scholarly research. Once this is clear it would be possible to send the document back to the funder or in future maybe also the institution.

*Management of rights:* The Licensing Expert Group recommends using the licence to publish\(^5\) to arrange the agreement with the author. There is a fundamental difference in granting a licence to access and granting the rights to share materials with the world. Sending materials to a repository is a new service which is being plugged into an old business model. Combining the licence to publish with a subscription model is complex. There was a discussion on who is responsible for managing the rights of an article, especially the right to deposit in a repository. Is it the library or the publisher? There is a question whether the licence to publish needs to be altered specially for open access.

From the publishers perspective it was also added that metadata of the articles is also of interest. Offering basic access to the metadata of articles is one thing. However, if the institution is expecting extra services, the publisher may well charge for these services. It was added that if the institution wants a provision in a licence that is a perfectly reasonable request. However there may well be a cost to the publisher which would need to be covered.

The publishers split up and went into separate rooms to discuss the terms and conditions regarding open access.

**Do and if so how do new developments influence licences?**

PDAs, mobile devices and cloud computing are just a few of the developments consortia are currently confronted with. Nol Verhagen, university library of Amsterdam, presented developments in this field. His position was that licences should not restrict but rather extend the rights granted through copyright. They should also not restrict access on mobile devices. A further new development is trust federations for authorisation and authentication.

Researchers are making more and more use of text and data mining and would like to have easy access to machine read large numbers of articles. Also there is a demand to share articles through social media, social networks, Virtual Research Environments (also across national boundaries). Countries are also negotiating national licences so that for example small and medium enterprises can have access to the articles. A further new development is the storing and sharing of research data. The question was raised whether publishers are expecting to play a role in making research data available.

The participants discussed the developments Nol Verhagen had presented and how these could be addressed. One of the questions was whether all these new developments should be addressed in licences. Some of these are in fact already covered in licences. The problem is that you have to identify each single possible use which leads to an extremely long licence, it would probably be better to more general in the licence.

Use by mobile devices is covered by the regular licence and does not require a separate clause. However allowed usage does differ from publisher to publisher and from licence to licence. As libraries cannot accommodate for each specific case libraries are only passing on limited access. It would be important to make it clear to the user what use they can make of the material. SERU\(^6\) uses with this concept. Have separate wordings allows for a shorter contract and it is easier to alter than changing legal clauses. It would be helpful if there was a public place on a website where standard clauses could be found. This would allow for easier re-use of clauses. The publishers were interested in discussing a roadmap for working on such a collection of clauses.

---


5 [http://copyrighttoolbox.surf.nl/copyrighttoolbox/authors/licence/](http://copyrighttoolbox.surf.nl/copyrighttoolbox/authors/licence/)

In connection with this you could also consider drafting different versions of texts for users with different levels of knowledge. For example one for the researcher who wants to know what he/she can do with the material and one for the lawyer who needs to know the legal agreement.

Regarding research data, publishers do not wish to monopolize the data or hold rights over the data. The question is whether publishers are able to add value to the data. The publishers see a role for libraries in storing and managing the research data.

**Text and data mining**

Wilma Mossink introduced the discussion on data and text mining. She raised the question whether a way of working could be established which would be acceptable to researchers, librarians and publishers. In Japan this has been solved by including an exception for data and text mining for research in the copyright legislation. However this is not the case in European countries.

Data and text mining is a field which is developing rapidly and challenges are emerging. One of these is that researchers want to mine across the boundaries of publishers and do not want to have to approach each publisher separately. This is resulting in publishers receiving individual requests from researchers to mine articles. For publishers it is a challenge to handle all these separate requests. There is also a technical issue involved as there are different ways of accessing the articles. For publishers it would be very useful if there were standards for how data and text mining took place and these standards were being shared.

A study by Eefke Smit and Maurits van der Graaf for the Publishing Research Consortium showed that most publishers are willing to allow data and text mining. On the other hand, when negotiating the 2012 Nesli licence a clause was added. 50% of the publishers took this clause out and only 25% (especially the smaller publishers) left this clause in the licence. There is an STM group working on a model clause for text and data mining to be included in the STM licences. This model clause still needs a broader discussion.

Regarding the fragmented collection of articles, it would be great if the publishers would provide a standard doorway, or collect the material in one place. The latter was also a recommendation from the Smit & de Graaf study but seems unlikely to be supported at present. Nonetheless there are places where a lot of content is available, for example Portico, or the Dutch National Library. Providing access through these organisations would mean these would become a licensing organisation on their own and take on a different role.

Publishers are working on discovery services based on a database of full text and metadata. The researcher may well want to mine everything, not just the licensed content. Perhaps a solution would lie in decoupling data mining and subscription. An approach could be to identify a trusted body that could coordinate and meanwhile work on a licence to offer access the full text. Trust management would have to be in place to administer rights.

Once content has been mined and new content created, there are also questions regarding the rights on the derivative works. It should be made clear in the end user licence what is allowed. This also relates to the attribution problem. How can the authors of thousands of articles be properly attributed? This could be registered by showing the object identifiers of the articles (for example the DOI) that were extracted. There might be moral rights of integrity involved as it might be important to attribute the authors in name. A use case would be helpful to be able to examine this further.

Derivative works is a concern for publishers. They are concerned that articles will be mined and the text will be shared with the world. In the long term the demands for data and text mining are an argument for gold open access. Through making articles available in open access the process would be a lot less complicated to arrange.

**Conclusions and discussing a roadmap for inclusion of the model provisions in the publishers’ licences**

Wilma Mossink thanked all the participants for their thoughts and comments. The first step will be that the comments from the plenary and breakout discussions will be incorporated in the model provisions which will be redrafted. These will be shared again for discussion and a check by lawyers will be required. If all parties agree on these they can be broadcast. The publishers are then free to endorse the provisions.

One of the outcomes of the discussion was that a bank of standard clauses would save all parties time and effort. This would be like the SERU wording but it would be in proper licensing terms.

**Continuous access provision**

Publishers would like to limit the options listed in the provision, an institution should not be left any possible option. It should be clear that an institution only has right to continuous access if they have purchased this. Institutions sometimes prefer to receive a copy but do not appear to take any action with this digital copy.

**Course Packs**

Regarding course packs, the wording is not quite clear, a more precise definition is required.

---


Open Access provision
It is important to make the concepts clear, and to provide clarity on what authors and institutions are allowed to do.

Data and text mining provision
It would be valuable to work on use cases, to provide insight into what is being done and what is still required. A licence for data and text mining is required, for this a model text can be drafted.
Annex: Model provisions for discussion
Version 3 February 2012 – The version discussed at the workshop

Model provision regarding continuous access

1. Long term preservation
The Publisher will archive the Licensed Material in at least one of the established e-journal archiving initiatives at the option of the Institution.

2. Continuous access in case of ‘trigger events’
The Publisher guarantees continuous access to and use of the Licensed Material which was published and paid for during the term of this and any preceding Agreements between the Publisher and the Institution through one of the established e-journal archiving initiatives if one or more of the following events (‘trigger events’) occur:

- a catastrophic and sustained failure of the Publisher’s delivery platform or
- the Publisher stops operations, or
- the Publisher ceases to publish a title, or
- the Publisher no longer offers back issues.

3 Continuous access in case of termination of the Agreement
In the case of termination of the Agreement, except when such termination is due to a breach of the Agreement by the Institution, the Publisher will provide (at the option of the Institution) the Institution and its Users with continuous access to and use of the Licensed Material which was published and paid for during the term of this and any preceding Agreements between the Publisher and the Institution, without charge, either by one or more of the following options:

- continued online access to archival copies of the same Licensed Material on the Publishers’ server;
- granting access to one of the aforementioned accepted e-journal archiving solutions;
- supplying archival copies of the same Licensed Material to a archiving facility shared by more Institutions;
- supplying archival copies of the same Licensed Material to an Institution to be uploaded to the protected network of the Institution.

4 Continuous access in case of transfer of a journal title
In the case when a journal title moves between publishers and access to the journal title is no longer supported by the Publisher, the Publisher guarantees continuous access to and use of the Licensed Material which was published and paid for during the term of this and any preceding Agreements between the Publisher and the Institution by legally binding the receiving Publisher of the journal title not to remove content that was previously deposited in accordance with the archiving arrangements even if the receiving publisher will not be continuing to deposit content in the archive. The Publisher will legally bind the receiving Publisher to continue the existing archiving arrangements for the journal title after the effective date of transfer of the journal title.

When one or more trigger events occur or in the case of termination of the Agreement or in the case of a journal title transfer continuous access implies that the Institution is permitted to use the Licensed Material under the same terms and conditions as provided for in the Agreement.

Model provision regarding Course Packs

Definitions:
Course Pack: a multi-source collection or compilation of information (e.g. book chapters, journal articles, abstracts, multi-media materials) assembled by members of staff of the Institution for use by students for the purpose of training, education and instruction, either in printed, electronic or non-print perceptible (audio or Braille) form;
Virtual Learning Environment: a system designed to support teaching and learning in an educational setting;

Permitted use:
Members of staff are permitted to download and print out parts of the Licensed Materials free of charge in Course Packs in connection with courses for academic credit and distribute these to the students of the Institution;
Members of staff and students may incorporate links to the Licensed Materials as well as parts of the Licensed Materials in Course Packs in connection with courses for academic credit, free of charge to be distributed or made available to the students of the Institution via Virtual Learning Environments or within an e-mail communication.
The Institution may offer Course Packs in audio or Braille to students who are in the reasonable opinion of
the Institution visually impaired.

Acknowledgement
The parts of the Licensed Materials used in Course Packs shall carry appropriate acknowledgement of the
source, title, author and publisher.

Model provision regarding Open Access to Publications

To be discussed.