

## STM contribution –JURI Copyright working group – 26 March European Parliament Brussels

### 1. Who is STM?

STM is the leading global trade association for academic and professional publishers. It has over 120 members in 21 countries who each year collectively publish nearly 66% of all journal articles and hundreds of thousands of monographs, text books, learning materials and reference works. STM members include learned societies, university presses, private companies, new starts and established players. STM's members publish some of the largest "hard" natural sciences, and also social sciences and humanities

The market –some key figures<sup>1</sup>:

Annual revenues generated from English-language STM journal publishing are estimated at about \$10 billion in 2013. The industry employs an estimated 110,000 people globally, of which about 44,000 are employed in the EU.

### 2. Why we are speaking today:

Global access to scholarly publications in the fields of science – all disciplines - has been facilitated in no small part by the efforts of STM publishers. Since the mid-1990s, STM publishers have made their published materials, scholarly and scientific journals, books and databases, available electronically and online. They are constantly innovating in means to facilitate access to information contained in the content that they publish. The industry has actively contributed in initiatives to enable online discoverability of scholarly publications and their authors. **The educational and research market is a primary market for STM publishers.**

### 3. What are our concerns? First some general remarks:

- The publishing industry depends on copyright protection to produce works and to innovate. This is even more so in the digital world.
- Copyright is about freedom: Freedom to publish, to speak and economic freedom. Scientific authors can choose the outlets where they believe it is best for them and their advancement of careers. Publishing in a high profile and prestige journals presents an indirect remuneration and is the cornerstone for scientific authors to advance their careers.

---

<sup>1</sup> [http://www.stm-assoc.org/2015\\_02\\_20\\_STM\\_Report\\_2015.pdf](http://www.stm-assoc.org/2015_02_20_STM_Report_2015.pdf)

- The scholarly publishing industry is highly innovative. Our members make constant investments into new technologies and solution. Therefore a steady flow of income is needed to sustain the system and future investments
- Exceptions – well and narrow crafted, along the 3 step test of Berne – are serving a special need. But broadening exceptions might be detrimental to the scholarly publishing market. **It is the high quality content we publish which is vital for the advancement of scientific progress.**
- The STM industry is at the forefront of the digital world. Over many years we developed a highly tailored licensing system to reflect the need of our customers. **Licensing of electronic content has become the very life blood of STM publishing.** But it is important that our members have contractual freedom, legal certainty and enforceability of agreements. STM's business relies on the ability to conclude enforceable contracts and licences.

### **Our concerns in particular : Educational exception, E-lending and Text and Data Mining (TDM):**

#### **a. Broadening the exception for education and research**

Existing exceptions and limitations have their place in copyright law, in particular where offline uses are concerned or where publisher and user cannot directly define contractual terms. However we are concerned that exceptions might lead to 2<sup>nd</sup> best solutions of re-purposing old content for new tools.

STM publishers are creating works protected under copyright solely targeted to the Higher Education and Research Market. One example is textbooks for the Higher Education sector.

With the advent and development of new innovative digital tools, scientific publications such as for example textbooks are morphing into smart but complex objects with data, web links, dynamic commenting, related blogs, wikis and videos. These technological developments allow an easy, seamless and cooperative sharing of information among students and researchers but require substantial investments. This content might disappear if publishers cannot afford the investment.

One has to be careful that the market is not pre-empted by further exceptions. This could harm the publishing sector which needs to continue to innovate and make further investments.

Licenses complement the current existing framework of exceptions. **They are flexible, adaptable and can be tailored to the need of the users and institutions.**

Most licenses used by STM members contain the list of permitted uses. Those clarify for the user what he or she can do under this agreement. This is regardless possible interpretations and exceptions.

One chief role of online licensing is to grant access to a defined user group. If you are part of this group – for example you are a student or researcher of a university - you can use the content anywhere and anytime. Content needs to

travel freely and across borders – we see the value and virtue of this but STM believes that this is best done through licences.

For this reason, you need to preserve freedom of contract. A Digital Single Market relies on a single licence being enforceable throughout the whole of the EU between willing buyer and seller. We need to take care that there **are no contract override provisions in exceptions that create disincentives to licensing in some parts of the EU, when in others the same licence would be valid.**

Since the inception of the internet, licensing of electronic content has become the very life blood of STM publishing. STM has consistently expressed the need for contractual freedom and legal certainty and enforceability of agreements. STM's business relies on the ability to conclude enforceable contracts and licences. Making it legally impossible for parties to a contract to agree and clarify the way in which content is going to be used, or rendering individual provisions of a contract unenforceable, will have a destabilising effect on STM publishing. It harms the widest possible availability of STM content on the most favourable terms and reduce the possibility to experiment with new access models.

## **b. E-lending**

STM has not yet issued a formal position on e-lending but please allow me to share some observations:

We note that courts in Germany have already pronounced that the principle of exhaustion relating to computer programmes, as decided by the Court of Justice of the European Union in 2012 in the *UsedSoft* case, **does not apply** to the licensing of digital products such as e-books. Questions on the licensing of such digital products are now pending before the Court of Justice. **In the digital environment every copy is a perfect copy.** There is no deterioration.

We feel that all stakeholders should be given the freedom to experiment and develop new business and licensing models. We are in a still evolving e-book market. Therefore e-lending should be approached with great care.

STM has the following contribution to make to the topic of e-lending:

- 1 For science and humanities publishers, the relevance of library sales in STM/Academic textbooks is completely different from the sales of books in general trade. Sales of books and 2<sup>nd</sup>-hand print books are totally different from trading in e-books licensing, where every e-book is an original copy, not a 2<sup>nd</sup> hand copy.
- 2 There is a need for a full diversity of models for e-lending. Exceptions would box these many models in and constrain the market unnecessarily.
- 3 STM has an open invitation to discuss and make the system work better for all through exchange of views.

The final topic is TDM – Text –and Data Mining:

## **c. TDM**

Although there is no uniform definition we – STM define Text and Data-Mining as follows:” TDM is a computational process whereby text or datasets are crawled and analyzed by software that recognizes new entities, relationships and actions.”

**TDM is NOT a simple search but a series of different actions/steps.**

**TDM leads to new relationships if successfully performed,** and it needs trusted and reliable text and data.

To keep it simple the following steps are needed:

1. Definition of the questions that one – the researcher - wants to solve
2. Compilation and assembling of the content which should be mined. **In practical terms for scholarly publishers – define the articles/works which should be considered for the TDM process** and exclude those that should not.
3. Process the preselected content so that the TDM Software which recognizes the entities and relationships can properly function

[And if necessary re-fine the process e.g. refine the question and/or content]

**It is particular step 2** where STM publishers have a stake and are concerned: we are the producers of high quality content. We invest to create this content. We are the guarantors of the trustworthiness and reliability. **TDM is about high quality content.**

In order to facilitate TDM to subscribed and non-subscribed content, certain technical preparations are needed to “make the content fit for TDM”. Therefore STM publishers are supporting various industry solutions. Those are CrossRef Text and Data mining service<sup>2</sup> and the service of CCC<sup>3</sup> (Copyright Clearance Center). Crossref leverages existing infrastructure STM publishers have built over the years to allow a cost-effective, yet reliable form of TDM in the public interest. The CCC bespoke TDM platform can be customised for different high-intensity research sectors. Those are the pharmaceutical sector and other industry sectors with different requirements.

For non-commercial researchers wanting to mine content that they/their institutions are subscribing to, STM publishers have committed to provide licenses under standard terms at no additional costs.

**Ladies and Gentlemen, stringent ambitious commitments were made under the L4E framework, including that non-commercial subscribing institutions would enjoy access for TDM at no additional costs<sup>4</sup>.** In part these commitments were made so that the true demand for non-commercial TDM can still surface.

Circumstances are very different when commercial players want to mine copyright protected works and also in this case an exception is not needed as this is already a rapidly developing market. In a market study “Text Analytics

---

<sup>2</sup> <http://www.crossref.org/tdm/>

<sup>3</sup> <http://www.copyright.com/>

<sup>4</sup> [http://www.stm-assoc.org/2013\\_11\\_11\\_Text\\_and\\_Data\\_Mining\\_Declaration.pdf](http://www.stm-assoc.org/2013_11_11_Text_and_Data_Mining_Declaration.pdf)

2014, by Seth Grimes of Alt Plana<sup>5</sup> the author conclude that "the technology is pervasive, even if not ubiquitous.... Text analytics has found its greatest success in four industry groupings: consumer facing businesses, public administration and government, life sciences and clinical medicine, and scientific and technical research".

Interestingly the study notes that "many text analytics users don't understand that they are doing text analytics, nor do they need to." With respect to the market size the study says that "2014 market size surely exceeds the \$2 billion figure suggested by applying a compounded 25% annual growth rate to the 2011 figure.

Concerning the competitive advantage of TDM in US versus Europe, it can be said this is remaining intact.

In a report commissioned by the European Commission in 2014 the following statement is being made: *"When it comes to the deployment of TDM, there are worrying signs that European researchers may be falling behind, especially with regard to researchers in the United States...."*

It has been alleged that the fair use principle in the US allows miners to mine any content they wish for commercial and non-commercial purposes. This is far from the truth and anyone mining (or simply indexing) websites of STM publishers, does it through a licensing scheme.

The fact is that the number of TDM references in the EU and the USA were similar until 2007. Then the EU28 started growing at a rate of 12.1%, higher than the USA (6.8%) (2007-2013 CAGR). A query performed in Scopus delivered the following results: EU28 publishing 1011 papers in 2013, vs 709 for the US. Even so, STM does not believe that counting research papers involving or about TDM is a suitable proxy to assess progress in TDM.

To conclude on TDM:

There is a growing and functioning broad market for TDM covering various market places. STM has the opinion that successful TDM is about **"tools and not rules"**.

#### **Conclusion :**

**STM considers that the public interest of research and education is best served by encouraging the creation of new publications and information services for its markets**

---

<sup>5</sup> <http://www.digitalreasoning.com/resources/Text-Analytics-2014-Digital-Reasoning.pdf>  
International Association of Scientific, Technical and Medical Publishers  
Prima House, 267 Banbury Road, Oxford OX2 7HT, UK  
[www.stm-assoc.org](http://www.stm-assoc.org)  
Registered at UK Companies House FC0274745  
Page: 5