LIBER Response to the Public Consultation on the review of the EU copyright rules

PLEASE IDENTIFY YOURSELF:

Name: Stitching LIBER (the Association of European Research Libraries)
EU transparency registry ID: 24973952940-04

TYPE OF RESPONDENT:

Other: We are answering on behalf of LIBER (www.libereurope.eu), the Association of European Research Libraries, and its network of over 400 member libraries from across Europe. These responses have been complied by the LIBER Copyright Working Group and approved by the LIBER Board.

LIBER's mission is to create an information infrastructure that enables research in LIBER institutions to be world class. For this infrastructure to thrive, we believe it must be part of an ecosystem that can accommodate and nurture the changing nature of research and innovation in the digital age.

As the infrastructure evolves to accommodate rapid advances in information technology, an explosion in the production of data and a culture shift towards collaboration and openness, so too must the surrounding policies and legislation. So far, however, the evolution of copyright and associated intellectual property legislation has not kept pace with the digital age. Without significant changes to European legislation, Europe's research potential will not be fully realised.

1. [In particular if you are an end user/consumer:] Have you faced problems when trying to access online services in an EU Member State other than the one in which you live?

Yes. In universities and other research institutions, researchers are usually granted access to copyrighted research material on the basis of individual or institution (IP) based identification. For access to this content, cross-border solutions have been found. Some restrictions can apply to content such as that digitised under public-private partnership, however, these are more often due to national content control than copyright issues.

More problematic is access to online multimedia content such as broadcasts, film, music and e-books, which are subject to regional licence control mechanisms. Restrictions are many and well documented (see Licences for Europe[1]), and are preventing researchers from accessing relevant linguistic and cultural content.

2. [In particular if you are a service provider:] Have you faced problems when seeking to provide online services across borders in the EU?

Yes. Besides the usual production and marketing costs, a content provider delivering content across border has to take three additional types of costs into account.

1. **Localisation Costs:** these costs cover tasks such as translation and adaptation of the content to local consumer preferences.

2. **Issuance Costs:** an amount which is unnecessarily high because of the diversity of legal systems in regard to Intellectual Property Rights (IPR), consumer protection, technical standardisation that content-providers must account for.

1 http://www.libereurope.eu/L4Ewithdrawal

3. **Maintenance-of-service Costs:** these mainly cover IPR aspects such as the risk of litigation and the cost to defend the service provided against copyright infringements.

Due to these high costs, a service provider may choose not to serve markets with smaller revenue generation potential. This puts researchers based in smaller European markets at a disadvantage because the content they require is not licenced for them to access from their country.

3. If you have identified problems in the answers to any of the questions above – what would be the best way to tackle them?

Industry should be encouraged to provide pan-European licences. By harmonising the legal and technical requirements for sharing content across member states, the cost of rolling out services across Europe could be reduced.

Inequality of access to content could also be addressed if libraries were able to provide an international interlibrary e-lending service on the same basis as that provided for analog formats (e.g. online document delivery for e-journal articles and e-books). See our answer to Question 36 for more information on how we feel e-lending could be better enabled.

4. Do you think that further measures (legislative or non-legislative, including market-led solutions) are needed at EU level to increase the cross-border availability of content services in the Single Market, while ensuring an adequate level of protection for right holders?

Yes. Licences for Europe has shown the willingness of service providers to move towards harmonisation of practices. The process also showed the limits inherent in self-commitment, considering that cross-border issues occur as much within the EU as between the EU and third countries. This makes, clear regulative action and agreements necessary on both a European and global scale.

5. Should the provision of a hyperlink leading to a work or other subject matter protected under copyright, either in general or under specific circumstances, be subject to the authorisation of the rightholder?

No. Education, learning and teaching – not to mention the functioning of modern-day society – requires that hyperlinks remain free of all intellectual property rights. They are one of the building blocks of knowledge, along with facts and data, catalogue and bibliographic records, and therefore must be freely usable and shareable by anyone for any purpose.

LIBER believes it would run counter to all the economic, social and educational interests of the Information Society to bring URLs into the scope of intellectual property laws. It would be entirely disproportionate and harmful to the European knowledge economy if URLs had to be licenced individually or collectively. Also, as creators of content and URLs, we do not believe it is in the interest of LIBER libraries to put a further barrier in place to exposing our collections via aggregators such as Europeana or OpenAire.

In general, we welcome the European Court of Justice (CJEU) ruling[2] in the Svensson case that providing a clickable link to content that has been made freely available on the Web does not require the authorisation of the rightsholder. We also agree that allowing member states to extend the concept of 'communication to the public' to included hyperlinking to publically available content would adversely impact the functioning of the internal European digital market.

6. Should the viewing of a web-page where this implies the temporary reproduction of a work or other subject matter protected under copyright on the screen and in the cache memory of the user's computer, either in general or under specific circumstances, be subject to the authorisation of the rightholder?

2 http://curia.europa.eu/juris/document/document.jsf?text=&docid=147847&pageIndex=0&doclang=EN&mode=req&dir=&occ=first&part=1&cid=135500

No. This is an extremely important question. It goes to the heart of whether a body of law that stretches back to 1709 to regulate the printing press is fit for purpose in a world where computers, servers and eBook readers must copy in order for a person to simply access creative works, whether those works are in copyright or in the public domain.
LIBER is, of course, very aware of Art 5.(1) of the Information Society Directive but does not believe that it serves any purpose at all to regulate the simple fact that machines copy in order for individuals to read, view, analyse or listen to something.

We strongly agree with the report "Copyright Reform for Growth and Jobs" that the right of reproduction, which is not defined in international treaties, needs to be more closely linked to the economic harm that is the result of communication to the public[3].

15. Would the creation of a registration system at EU level help in the identification and licensing of works and other subject matter?

Yes.

16. What would be the possible advantages of such a system?

Such a system would have the following advantages:

1. Organisations and individuals who want to search for rightsholders would be able to do so more easily and at a lower cost than is currently possible. This would in turn reduce the number of orphan works and increase the chance that creators could receive financial and / or reputational benefit from the reuse of their works through relicensing.

2. Authors and creators would be better able to take action against copyright infringement.

3. The creative economy would be better supported and access to knowledge would be improved because there would be a clear differentiation between content for which creators want to control reuse and works which people are not producing for commercial gain and could be freely reused by third parties for cultural and financial benefit. This is particularly pertinent for older works. As works age, it becomes increasingly difficult to find the rightsholders and works fall into limbo where no financial or reputational benefit accrues to the creator (see below).

Registration for such a system could be simple and need not be daunting for copyright holders. Given that online life in the 21st century requires registration for many online services, registration is no longer the barrier that it once was. In addition to the above points, we feel it is important to emphasise the following:

- The Term Directive 2011/77/EU was a significant lost opportunity for creators and licensors of sound recordings given that where works do revert there is no requirement for this information to be made public. If the Directive had included such a provision it would have facilitated the relicensing of works, that currently are not being licenced in any significant amount, thus depriving the creator of reputational and / or financial benefit.
- Value exists not only in the control of the exclusive rights by the creator, but also in their waiving of these rights and allowing third parties to reuse their works while still in copyright, as well as the process of copyright works entering into the public domain. For example, a 2005 Library of Congress study comparing rights related to audio recordings in the European Union and the United States shows that works no longer subject to control by the original rightsholder (in this case because they were in the public domain in Europe but not in the U.S.) meant that by a ratio of two to one there were more reissues available in the EU than the U.S.[4] This positive effect could be strengthened if a registration system were available that made it easy to locate works which were either in the public domain or for which their creators had voluntarily rescinded copyright.
- It is a well-known fact discussed many times in the context of the Orphan Works debate that as you go back further in time there are more orphan works and that more and more works fall out of commerce[5] [6]. Registration could help alleviate this issue as it would create smoother licensing for registered works, and a body of content unregistered that could be reused in a fashion that it currently is not possible.

[3] http://www.lisboncouncil.net//index.php?option=com_downloads&id=84
[4] Survey of Reissues of U.S. Recordings. T Brooks. Co-published by the Council on Library and Information Resources and the Library of Congress. 2005. http://www.loc.gov/rr/record/nrpb/pub133.pdf
[5] How Copyright Makes Books and Music Disappear (and How Secondary Liability Rules Help Resurrect Old Songs). Paul J Heald. 2013.
[6] Copyright Term Extension: Estimating the Economic Values. E Rappaport 1998.

17. What would be the possible disadvantages of such a system?

The main issue would be establishing a technical infrastructure to manage this and to maintain it over time, given the very long duration of copyright.

Another issue could be that some creators wish to assert copyright over their works but do not register because they are not aware of the database. However given that in the 21st century everyone who writes or creates online is a copyright holder, whether they want to be or not, we may have more people who want their work to be freely used by others than those who want to reserve their copyright. This is particularly evident in online communities and within the sciences.

18. What incentives for registration by rightholders could be envisaged?

As in the United States, rules could be put in place whereby certain elements of copyright protection are only available to rightsholders who have registered their works. For example, registration might be required for a rightsholder to start an enforcement action. Another incentive might be that rightsholders need to register their works in order to be eligible to collect royalties through collective rights management organisations.

19. What should be the role of the EU in promoting the adoption of identifiers in the content sector, and in promoting the development and interoperability of rights ownership and permissions databases?

Our main concern is that identifiers be based on open standards and be interoperable. The EU should encourage the adoption of open standards and foster interoperability.

20. Are the current terms of copyright protection still appropriate in the digital environment?

No. Whilst acknowledging copyright's social, economic and cultural functions, we believe that the minimum terms of protection in Europe (currently 70 years) should be reduced. Experts repeatedly point out that the current term of protection is too long[7].

The most significant argument against a long term of protection is the erosion of the public domain. As it is stated in the Europeana Public Domain Charter[8], *Having a healthy and thriving Public Domain is essential to the social and economic well-being of society*. The public domain is vital not only for librarians and researchers but also for creators and artists who, for centuries, have used and transferred works created by their predecessors.

Longer copyright terms exacerbate the issue of Orphan Works, which are created when copyright owners cannot be found or identified. Only a small percentage of copyright works created for commercial gain retain ongoing value sufficient to justify long copyright terms. The EU Directive on Orphan Works recognizes the significance of the orphan works problem and LIBER believes that a reduced term of protection for copyright works would greatly alleviate this issue. The need to carry out a diligent search for owners of orphan works is putting too much strain on library resources and is creating a cultural black hole as 'orphaned' content cannot be used by the public or researchers. This problem affects not only relatively old publications but also more recent materials. The National Library in Warsaw, for example, estimates that about 300,000 of the books, pictures and newspapers in its collection are 'orphaned', including a huge collection of clandestine literature from the Solidarity and Martial law period.

The terms of protection should be brought in line with the minimum terms of protection set out in the Berne Convention and the Agreement on Trade Related Intellectual Property Rights (TRIPS): life plus 50 years for copyrighted works and 50 years for neighbouring rights.

[7] I. Hargreaves, P.B. Hugenholtz, Copyright Reform for Growth and Jobs: Modernising the European Copyright Framework. The Lisbon Council Paper, 2013 http://www.lisboncouncil.net/publication/publication/95-copyright-reform-for-growth-and-jobs-modernising-the-european-copyright-framework.html

[8] http://pro.europeana.eu/c/document_library/get_file?uuid=d542819d-d169-4240-9247-f96749113eaa&groupId=10602

Further action should be taken at international level in order to investigate the optimal duration of the term of protection[9] [10].

21. Are there problems arising from the fact that most limitations and exceptions provided in the EU copyright directives are optional for the Member States?

Yes. There is to all intents and purposes no harmonisation in the EU for users of copyright works in the single market. This makes the cross-border use of copyright works problematic. From the perspective of a pan-European organisation such as LIBER, it is clear that heterogeneity of exceptions in national copyright legislation is blocking the development of an effective and efficient collaborative data infrastructure. Research in the digital age is becoming increasingly collaborative and international, yet exceptions relevant to researchers have not been adopted in all European countries. An uneven playing field exists for researchers in terms of their ability to access and exploit information in a collaborative infrastructure.

Libraries wishing to preserve and provide access to cultural heritage face similar issues. Exceptions benefitting publicly accessible libraries, museums and archives have not been implemented (uniformly) in all member states and this creates unnecessary uncertainties. One example of the impact of this uncertainty is the Europeana Newspapers Project, which aims to make historic newspapers available via Europeana. Due to varying national interpretations of copyright exceptions and limitations there are huge gaps between what and how the content from different institutions is made available.

22. Should some/all of the exceptions be made mandatory and, if so, is there a need for a higher level of harmonisation of such exceptions?

Yes. All those that relate to education, learning and access to knowledge should be mandatory. A harmonised approach to the adoption and implementation of limitations and exceptions across Europe is required to reduce legal uncertainty and to promote the circulation of knowledge in the single market.

23. Should any new limitations and exceptions be added to or removed from the existing catalogue? Please explain by referring to specific cases.

Yes. In light of the collaborative nature of research and the potential for digital technology to improve access to and the preservation of our cultural heritage, LIBER calls for the following:

- An exception for text and data mining for all research purposes (see answers to Q53-57 for elaboration).
- An exception for preservation networks between institutions. This is important because currently most preservation exceptions relate to preserving material in the library's own collection rather than having digital mirror sites and activities shared between institutions. Digital preservation is a costly exercise involving massive investment in technology, infrastructure and skills. Storage costs are not decreasing in line with the growth of digital information. The only way to ensure that our cultural heritage remains accessible in the future is to reduce duplication of effort and share infrastructure.
- Prevention from contracts undermining limitations and exceptions in copyright law. Without such an exception the harmonisation of exceptions is pointless as it will be undermined by licence agreements.
- Legal certainty for e-lending by libraries, similar to analogue lending.
- For publicly-funded research results to be made openly available regardless of contracts signed with a publisher. By increasing accessibility and availability, and by facilitating the use and reuse of content, open access will play an integral role in driving research excellence globally. Copyright law should not hinder its progress. An exception or legal instrument should be developed which allows researchers to make the results of publicly-funded research available openly regardless of contracts signed with a publisher. Open data is also an area of growth and potential driver of research and innovation. It is important to ensure that legal Intellectual Property instruments such as the Database Directive do not conflict with the realisation of this potential and negate the huge public investment in research infrastructures to support reuse of this data.

[9] R. Pollock, Forever minus a day? Calculating optimal copyright term, University of Cambridge, 2009, http://rufuspollock.org/papers/optimal_copyright_term.pdf

[10] N. Adilov, M. Waldman, Optimal Copyright Length And Ex Post Investment: A Mickey Mouse Approach, Economic Inquiry, 2013, http://ideas.repec.org/a/bla/ecinqu/v51y2013i2p1101-1122.html, vol. 51(2), p. 1101-1122

- Research exceptions that make no distinction between commercial and non-commercial purposes. Knowledge transfer is now at the heart of what research universities do. The lines between what is non-commercial and potentially-commercial research are blurred. Any attempt at definition will limit the impact of research activities.

No existing exceptions should be removed.

24. Independently from the questions above, is there a need to provide for a greater degree of flexibility in the EU regulatory framework for limitations and exceptions?

Yes. A general open-ended exception, compatible with the Three Step Test[11], is required. This will avoid obsolescence and ensure that any new copyright regime can accommodate future digital innovations, maintain European competitiveness and be interoperable with copyright regimes such as those in the United States. In an era of fast-moving technologies, we believe a closed list of exceptions is not flexible enough to allow European research institutions to compete with those in the United States and Asia (Israel, Singapore, South Korea and Taiwan), where fair use is already in place.

25. If yes, what would be the best approach to provide for flexibility? (e.g. interpretation by national courts and the ECJ, periodic revisions of the directives, interpretations by the Commission, built-in flexibility, e.g. in the form of a fair-use or fair dealing provision / open norm, etc.)? Please explain indicating what would be the relative advantages and disadvantages of such an approach as well as its possible effects on the functioning of the Internal Market.

The United States, Israel, Singapore and South Korea have an exception known as fair use. Rather than allowing a pre-defined activity, it is a test that can be applied to any activity that uses a copyright work to see if it undermines the legitimate interests of a rightsholder or not. A similar law is just one possible solution to the current situation in Europe.

Without an open ended exception, significant market distortion between Europe and the aforementioned countries could occur (e.g. Google books case in U.S.), potentially impacting on the research and development capacity of European countries.

26. Does the territoriality of limitations and exceptions, in your experience, constitute a problem?

Yes. Each non-global solution has the potential to generate market distortions which lead to the short-term reconfiguration of a market. Research, because it is often collaborative and can involve international teams, is highly sensible to such changes or geographic differences in regulatory frameworks or access to research content (and the modalities of its re-use).

One example of how the territoriality of limitations and exceptions can have an impact on research institutes is reflected in the findings of the U.S. District Court on the Google Books digitisation project. The court found that that the digitisation of university library collections by Google was fair use under U.S. copyright law. In his ruling, Judge Chin observed *"it advances the progress of the arts and sciences, while maintaining respectful consideration for the rights of authors and other creative individuals, and without adversely impacting the rights of copyright holders."* There is no exception in EU law that would allow European research libraries to engage in an activity at this scale. This means that researchers in the U.S. have easier access to more scholarly content than their European counterparts.

At the European level, the case of Subito, illustrates how the territoriality of exceptions means that a service needed by researchers across Europe has been limited so that it serves only researchers based in a single country.

[11] http://en.wikipedia.org/wiki/Berne_three-step_test

27. In the event that limitations and exceptions established at national level were to have cross-border effect, how should the question of "fair compensation" be addressed, when such compensation is part of the exception? (e.g. who pays whom, where?)

There is a need for such a mechanism because limitations at national level do have cross-border effects (even if of a second order, economically speaking). Practically, however, it seems difficult to develop a working mechanism without redefining deeply the usual consumer understanding of content consumption (eg. the case of a student from country A enrolled in a MOOC from an institution in country B, using content whose IPR owners are in country C). Compensation schemes should only apply if there is evidence of specific harm to authors and if that harm cannot be remedied in other way.

28. (a) [In particular if you are an institutional user:] Have you experienced specific problems when trying to use an exception to preserve and archive specific works or other subject matter in your collection?

Yes. Although there is an exception for preservation within the Information Society Directive, it is difficult to apply to the purpose of digital preservation.

A key reason for this is the very nature of digital information. It is ephemeral and can be easily deleted, written over or corrupted. Because information technology such as hardware, software and digital object formats evolve so rapidly, it can be difficult to access and use digital materials created only a few years ago. At the same time as countless born-digital works are created every day, countless others are lost as they are removed, replaced, superseded or simply left untouched in obsolete formats and media.

Recommendations of the European Commission on the digitisation and online accessibility of cultural material and digital preservation (2011) acknowledges that digital material has to be managed and maintained to prevent its otherwise inevitable loss. For organisations, however, there are no less than 11 ways in which copyright-related issues prevent effect digital preservation from taking place.

1. **Design & Implementation of Preservation Exception:** Most national laws of EU Member States have implemented in one way or another the preservation exception provided by paragraph 2(c) of article 5 of the EU directive 2001/29/EC on the harmonisation of certain aspects of copyright and related rights in the information society. In most cases, however, the legislation has been created with physical preservation in mind and is ill fitted for the purposes of digital preservation. Preservation of physical works generally requires only occasional activities to repair or restore the copies of works, which is perfectly doable in the current legislative framework. Digital preservation, however, requires proactively performing a series of tasks (e.g. format shifting) to ensure that digital materials can be accessed for as long as necessary. For this reason the UNESCO Vancouver declaration on the memory of the world in the digital age (2012) calls for the elaboration of legal frameworks that support the preservation of and access to digitised cultural heritage.

2. **Reproduction Rights:** Digital preservation necessarily involves the exercise of one or more of the exclusive rights of the author or other right holder, the most prominent being the right of reproduction. Digital works cannot be preserved unless they can be copied or otherwise acquired by a digital archive or other preservation institution.

3. **Type of Institution:** One of the limitations of the current copyright regime is the range of institutions to which the preservation exception applies. This differs from one national jurisdiction to another. While the Directive 2001/29/EC provides that exceptions to reproduction can be applied to libraries, museums, archives and educational establishments, this often is not the case in national legislation. For example, UK copyright law provides an exception that applies only to libraries and archives, which means that other cultural organisations with valuable historical collections such as museums or galleries will infringe copyright if they copy for preservation without the express permission of copyright owners. According to the WIPO[12] study on Copyright Limitations and exceptions, single copy provisions also apply in Greece, Hungry, Latvia and Poland. Slovenia and Germany set a limit on the number of copies that can be made.

[12] http://www.wipo.int/copyright/en/limitations/

While the Directive 2011/29/EC does not limit the types of works which can be reproduced by aforementioned institutions, some national legislations have implemented limitations in this regard. For example UK copyright law only applies to literary, dramatic or musical works. This means that copying a sound recording, film, broadcast or artistic work for preservation is an infringement of copyright, unless expressly permitted by copyright owners.

4. **Single-copy Provision:** Long-term management of a digital work generally requires that multiple copies of the work be made over the course of its lifetime, both to retain one or more redundant copies in different locations and to perform migration of content from an old to a new technology to make sure the works can be accessed even when the technology has become obsolete. In some national legislations the number of copies permitted to be made by aforementioned institutions is limited. For example UK copyright law appears to only permit the making of a single copy, so multiple backup copies or the use of preservation technology that operates by making multiple copies are not covered by it. This limitation also impedes the operation of legal deposit institutions where the legal deposit law allows distribution of legal deposit materials from one legal deposit library to other deposit libraries that also have the right to these works. This limitation is recognised by the Recommendation of the European Commission on the digitisation and online accessibility of cultural material and digital preservation (2011) which asks Member States to make explicit and clear provision in their legislation for the multiple copying and migration of digital cultural material by public institutions for preservation purposes.

5. **Proactive Preservation:** Most of the national legislations also provide very specific conditions which have to be met in order for reproduction to be lawfully made. For example, UK copyright law provides that a reproduction can be made only if the item in question is damaged, deteriorating, lost, or stolen, or if the existing format in which the work is stored has become obsolete. Such provisions doesn't allow for proactive digitisation of works in order to create a digital copy of a work for preservation purposes, which is one the tasks the legal deposit institutions have been tasked with. Moreover, these provisions are meaningless in regard to the digital-born works. If they have been damaged (for example through bit rot) or lost, they cannot be copied anymore. Another limitation which fails to fulfil its intent is the provision that only copies of out-of-commerce works can be made. As argued above, the preservation of digitally born works must be proactive in the face of rapidly-changing technologies and the generally unstable state of internet resources. It can be argued that works published online can never become out-of-commerce, therefore preservation of these works becomes impossible in principle.

6. **Legal deposit and the Web:** The most troubling shortcoming of current implementations of the exception is that in most cases they provide that institutions tasked with preservation can reproduce only works in their holdings. While some EU Member States have adjusted their legal deposit legislation to also include works published online (including the obligation for legal deposit institutions to perform so-called web harvesting), the current legal-deposit legislation in different Member States is wildly inconsistent in this regard. For example, it is not clear if the regional nature of copyright allows for the harvesting of web pages belonging to national cultural heritage if they are hosted outside the respective country, created by residents of other country (for example, a web page programmed by a company in the U.S. while the content is created by residents of the respective country) or have a domain address outside of the national upper level domain.

7. **Streaming and Licenced works:** Even more uncertainty covers the works which have are made available by means of streaming, by commercial entities outside the country of residence of the author or works which can be only licenced. Examples include e-books made available on Amazon, musical works made available on iTunes, audiovisual works made available on YouTube, applications made available on Google Market and online computer games made available on Steam. There is no doubt that these works should be preserved for future generations. To this end the preservation exception should be extended to include also the works made available online and amendments to the legal deposit systems of Member States to be made. An even better solution would be to create an effective network of European legal-deposit libraries network for digital content, which would be fully in line with the Commission Recommendations of 2000, 2006 and 2011 for legal deposit and digital preservation.

   A severe challenge for research libraries is to ensure perpetual access to the electronic journals which form the most important source of scientific information. Since these journals are typically are made available on the grounds of subscription, the only solution to have a perpetual access to these works is to licence the reproduction rights of these journals. This is not always is possible. Even research works published in open access need to have a license attached that explicitly gives the user reproduction rights. Otherwise a licence has to be obtained in order to preserve the work and thus to avoid the risk that the publication might get removed at the original source while it still is necessary for researchers. This licensing work is a major cost element of digital preservation. The national library or other national

public infrastructure of the country could be made responsible for local hosting for preservation purpose, as well as for permanent access and use of the licenced materials by authorised users after the termination of a licence contract.

8. **Shared Infrastructure:** Since the preservation exception isn't very well harmonised across EU Member States, it seriously impedes an effective cross-border collaboration in digital preservation. Digital preservation is a very resource-intensive task, which requires advanced technical infrastructure and the training of experts in digital preservation. International centres for digital preservation would be more efficient, especially where developing a national centre of digital preservation would prove impractical and costly (e.g. in all three Baltic states). The lack of harmonised preservation exceptions also severely impedes a broader use of cloud preservation services. It could be argued, for example, that this is behind the low adoption rate of the Preservica digital preservation cloud service (created by the UK digital preservation experts Tesella) in the EU as compared to far higher levels of adoption in the United States.

   There is currently also some duplication of effort. This could be eliminated by means of national and international cooperation. For example, in Italy most of the Italian universities publish doctoral theses in their open access institutional repositories. Italian legal deposit libraries duplicate this effort by obtaining these works by means of web harvesting. This work might be done much more effectively by establishing a national digital preservation network and not by each institution digitally preserving its own collection. International collaboration on a European level would make a lot of sense in the field of web harvesting by allowing cross-border harvesting or by enabling the harvesting of web resources on behalf of preservation institutions in other EU countries.

   While some of the challenges described above can be remedied by improving national legal deposit legislation, some aspects (especially those related to the preservation of licenced content and the cross-border applicability of preservation exceptions) can be remedied only by changes in EU regulation. This suggestion has been confirmed by the study on the application of Directive 2001/29/EC on copyright and related rights in the information society, which was commissioned by the European Commission and carried out by De Wolf & Partners in 2013. One of the conclusions of the study was that the exceptions for libraries and archives at the EU level are in need of harmonisation to facilitate cross-border application.

9. **Technological Protection Measures:** Technological protection measures are closely connected with the reproduction of digitally-published works for preservation in libraries, archives and museums. Memory institutions have permission under most national legislations to bypass the technological protection measures of works in their holdings in order to take advantage of the preservation exception but in most cases this proves to be impractical. Moreover, these provisions typically don't cover the technological protection measures protecting the works which are licenced, not sold or acquired by means of legal deposit to libraries (eg. e-books protected by means of Adobe DRM technology).

10. **Adaptation Rights:** During the lifecycle of the digital preservation it sometimes becomes necessary to perform format shifting. For example, if any current technology becomes obsolete, the work has to be transformed to another format so as to be readable by current technological means. This can be interpreted as adaptation of work, which is not covered by the current preservation exception. Also, when format shifting complex data such as geographical data, it is inevitable that the functionality of the original data will be compromised and altered in some sense.

11. **Make-available Rights:** The ultimate goal of all preservation activities is to ensure that cultural heritage works can be accessed and used at a later date. Making digitally-preserved works available is therefore an integral part of the preservation process. Most of the current preservation exceptions in the legislation of EU Member States provide for limited opportunities to consult the digital works on premises of institutions who have performed the digital preservation. This is not universally true, however. Copyright law in the Netherlands does not allow the making available of preserved works. This use has to be licenced, which in turn means that digitised orphan works cannot be made available, at least not before the Directive 2012/28/EU on certain permitted uses of orphan works is implemented in the legislation.

The next two responses in this consultation give more detailed answers on the use of "make-available rights" in libraries. It should be noted at this point that the use of these rights in libraries, archives and museums is deeply interconnected. It is quite obvious that copying for preservation purposes does not in itself pose a risk to right holders, unlike providing access to these copies. However as long as preservation institutions have appropriate security and access control measures in place, the interests of right holders should not be threatened by more inclusive preservation exceptions or by more liberal regulations regarding access to the preserved copies of works.

29. If there are problems, how would they best be solved?

The solution to the problems described above is to expand the wording of the directive so that it also applies to the digital preservation activities and to make this exception mandatory across EU Member States. The latter has been suggested also by the De Wolf study on the application of Directive 2001/29/EC on copyright and related rights in the information society, on the grounds of the current inconsistent implementation of the exception in the EU Member States, the fundamental freedoms recognized by the EU Charter on Fundamental Rights and the European public interest.

The study argues that it would be a good idea to *"make mandatory some key exceptions that convey fundamental European public interests, such as the safeguarding of European cultural heritage or the development of a European research area and internal market for education. When those objectives are encumbered by the discrepancies resulting from diverging national implementations, harmonisation of exceptions, with proper conditions and restrictions, might be as important as harmonisation of exclusive rights that was achieved in 2001."*

30. If your view is that a legislative solution is needed, what would be its main elements? Which activities of the beneficiary institutions should be covered and under which conditions?

The preservation exception should apply equally to all non-profit libraries, archives, museums and other institutions as may be authorised by national law that are open to the public, provided they do not undertake these activities for any purpose of commercial advantage.

The preservation exception should apply equally to all categories of copyrighted materials, including literary, artistic, musical and dramatic works, as well as to motion pictures and sound recordings. It should also apply equally to all media and formats, whether hard copy or electronic, born digital or digitised for preservation. This includes the preservation of materials which are made available to the public in digital form or deposited by means of a legal deposit system, the legal ability to harvest publicly-available online content for preservation purposes and the preservation of licenced digital materials. To this end, licence clauses which override existing exceptions and limitations should be null and void.

The preservation exception should allow preservation institutions to proactively preserve copyrighted materials before they deteriorate, are damaged or are lost, and before any software or hardware required to access and use the material becomes obsolete. This includes also the right to perform format shifting and other adaptations for long-term preservation purposes. To this end, all licenced materials should be made available to preservation institutions either without technological protection measures or with the technological means necessary to remove the technological protection measures. The use of copies of the licenced materials should be permitted in case access to these materials is discontinued or suspended. Moreover, it should be established that the preservation exception cannot be over-ridden by contract terms of licences.

In order to effectively harmonise this exception, a detailed list of provisions would need to be put forward in the text of directive. Another option could be to transform the preservation exception into a limitation to copyright, defined by its objective and some general conditions and including a more open requirement that the use does not exceed what is necessary for its objective. Such an approach would make the copyright framework more fit for its purpose and understandable for both users and copyright owners. At the same time, if the requirements for this limitation are adequate and legitimate, it would not sacrifice the high level of protection of copyright and related rights that the EU law has adopted.

30. If your view is that a different solution is needed, what would it be?

It can be argued that all of the current problems can be solved by means of licensing. However, such approach would be extremely costly and burdensome both for memory institutions and rightsholders. A British Library/ARROW study on mass digitisation by libraries and archives found that it takes an average of nearly 6 hours to clear the rights to digitise a single book. It can be assumed that rights clearance of complex works consisting of a large number of separate works belonging to different owners (eg. film or broadcast) would take much longer than this. By using the data of the annual survey of hours and earnings of 2011, it was estimated that it costs £95 on average for libraries and archives to rights clear a single work. Information from the PwC report on educational copying suggests that it takes copyright owners 4,5 hours to licence a single work, at an average cost of £70. In face of such prohibitive costs many institutions would choose not to digitise these works, which in turn would create a high social and cultural cost.

The opinion that libraries and other memory institutions should not be overburdened by licensing is supported by the European Commission, which has pointed out in the document Copyright in the knowledge Economy: *"Libraries and universities underline the complexity and fragmentation of the current system of licensing agreements with publishers. A typical European university is required to sign a hundred or more licences governing the use of digital research material supplied by various publishers. Examining what each of these individual licences permit with respect to e.g. access, printing, storage and copying is a cumbersome process."*

LIBER strongly supports this position.

32. (a) [In particular if you are an institutional user:] Have you experienced specific problems when trying to negotiate agreements with rightsholders that enable you to provide remote access, including across borders, to your collections (or parts thereof) for purposes of research and private study?

Licences are frequently negotiated by national licensing consortia. It is common for licences to prevent cross-border access to digital content for research and study by imposing limitations on specific territories and categories of users on a take-it-or-leave-it basis. Licensors holding significant market power may have to re-negotiate licensing fees annually and there have been threats of refusing to grant licenses at all.

For digital documents it can be difficult to request remote access because the content may only be downloaded on a specific terminal or has a special licence provision and Technical Protection Measures in place that prevent more than one user from accessing the content.

Another barrier is that different platforms are used in different countries and by various publishers. This ties the library and end user to a multitude of platforms. IFLA has produced a useful background paper on the difficulties libraries experience in licencing and lending e-books[13].

Another pressing issue from LIBER's perspective is the ability of libraries to provide online access to collections they own. The exception that allows institutions to make works in their collections available 'for the purpose of research or private study, to individual members of the public by dedicated terminals on the premises' is not in line with our user's expectations.

33. If your view is that a legislative solution is needed, what would be its main elements? Which activities of the beneficiary institutions should be covered and under which conditions?

At the legislative level, to facilitate remote access the scope of the exception currently provided for in article 5(3)n of the InfoSoc Directive should be broadened so that libraries can make their collections available on-line via electronic networks such as an intranet for scientific and educational purposes without restriction to on-site terminals.

36. (a) [In particular if you are a library:] Have you experienced specific problems when trying to negotiate agreements to enable the electronic lending (e-lending), including across borders, of books or other materials held in your collection?

Yes. From a legal point of view e-lending has nothing to do with lending. Lending can be considered a specific use of the distribution right but e-lending can be considered a specific use of making available right in combination with the reproduction right. Contrary to lending, e-lending is an online service and not subject to exhaustion of rights after the first communication of the work to the public. In practice e-lending is implemented either by giving the user access to the work for a period of time, after which the access is denied, or by letting the user download a version of the work to his reading device. This work can then self-destroy after a pre-determined period of time.

[13] http://www.ifla.org/publications/background-paper-on-e-lending-2012

As any use of the making available to the public right requires an authorisation from rightsholder for the purposes of enabling e-lending, libraries have to acquire a licence. As a result, the ability to acquire commercially published books for library collections is severely constrained. Unlike with physical books, publishers can refuse to grant a licence on e-books to libraries and often they choose to do so. This undermines the fulfilment of libraries' public mandate and threatens libraries' role in ensuring freedom of access to information for all, as enshrined in Article 19 of the Universal Declaration of Human Rights. That the acquisition policy of libraries may be decided by the rightsholder and not by the library is unacceptable.

Even when a licence for an e-book is granted, the licensing conditions are often very stringent. The use of physical works in libraries is strictly defined by the current framework of copyright limitations and exceptions, but this same framework does not have to be taken into account in licences that grant libraries the right to make e-books available. This allows rightsholders to enforce very strict access rules which would be unimaginable in the world of physical books. Quite often licences permit users to consult the e-books only on premises of library. Reproduction rights, including printing rights, may be excluded. A 2008 study by the British Library[14] of over 100 library contracts for electronic resources concluded that over 90% of contracts were more restrictive than exceptions in the copyright law. It can be argued that licensing terms that restrict access to works beyond the level of national copyright exceptions effectively allow rightsholders to circumvent the copyright system in order to diminish the rights of information users. This violates the spirit of copyright law.

The De Wolf study on the application of Directive 2001/29/EC on copyright and related rights in the information society recognises that "the traditional role of libraries in providing an alternative to get access to cultural content should be preserved in the digital environment and their mission should be extended to the provision of e-books and other digital content. To that end, their autonomy should be preserved. Relying only on the market to deliver e-books to library readers could potentially dictate unreasonable terms and conditions to libraries or transform public lending into another commercial service provided by the publishers".

The licensing requirements for research libraries and public libraries are markedly different. Research libraries use a wide array of subscription services in order to ensure economically feasible access to the most recent and most cutting edge research materials in the field of interest to the staff and students of the given university. While subscription to databases of research journals has been a common practice for research libraries for a couple of decades, subscription to e-book databases is gradually supplementing or even replacing the printed monographs. Taking into account the specific needs of their users, research libraries are mostly concerned about such aspects as the possibility for remote access to the works for authorised users and to copy or print parts of the work.

The most troubling aspect of e-book licensing is that the licences typically do not grant libraries the ownership of the copy of the work, or even the right to make a reproduction of the work for preservation purposes. If this existed, the library could ensure perpetual access to the licenced work. As the situation currently stands, libraries have experienced instances of publishers revoking licences to certain works, typically by removing certain titles from the subscription packages. If a distributor of e-books to libraries were to file for bankruptcy and cease trading, then libraries could theoretically be stripped of entire e-book collections overnight. This would obviously create a serious disruption to the work of research libraries. It is therefore of vital importance that libraries should be allowed by the preservation exception to make copies of licenced works (see the answers on questions regarding the preservation exception).

Additionally, we should consider that a critical library service is interlibrary loan, including international interlibrary loan. This ensures that individuals can access to any work published even if it is not carried by the local library. However, virtually no e-book licences granted to libraries allow these works to be loaned to other libraries, let alone libraries in other countries. Concerns have been voiced by publishers that this would allow for a work licenced to one library to be instantaneously supplied to library users anywhere in the world. While this certainly would be true if licences allowed international interlibrary lending, it would not create any threats to the economic interests of rightsholders if proper safeguard mechanisms were built into the agreements. For example, technology could be enabled to ensure that a work can be simultaneously lent out only to one reader, thus mimicking the physical reality.

[14] http://pressandpolicy.bl.uk/imagelibrary/downloadMedia.ashx?MediaDetailsID=2137

37. If there are problems, how would they best be solved?

It can be argued that the easiest and most effective way to remedy all the current problems with the lending of e-books by libraries would be to extend the framework of public lending to digitally published works. This could be done by ensuring that the principle of exhaustion of rights after the first sale of work also applies to digital works. This solution would, however, affect not only e-books but all digitally published works (eg. audiovisual materials and software) and not only the use of works by libraries but a wide range of uses, most notably resale of digital works. Therefore it is understood that such a drastic change in the copyright framework would require very careful consideration from all angles so as not to upset the balance of rightful interests of both rightsholders and users of works.

Another option is to consider less drastic changes in legislation which would still go a long way towards remedying the problems the problems described above. The main aim of these changes would be to make sure the existing copyright limitations and exceptions apply also to the contracts, thus granting libraries the right to copy a portion of the work, reformat the work for preservation purposes, provide an interlibrary loan copy, reformat the work to enable print disabled access, and bypass a technological protection measure for the purpose of exercising any non-infringing purposes.

Moreover, the exclusive rights that exist in regards to published works should not include the right of refusal to sell to libraries or to limit their services, both in physical and digital environment. In other words, the copyright directive should stipulate that contracts cannot override limitations and exceptions. Any terms of a contract violating existing limitations and exceptions should be considered null and void. This view is supported by the Hargreaves Review on Intellectual Property and Growth, commissioned by the UK Government, which explicitly recommends that the UK Government should legislate to ensure copyright exceptions and limitations are protected from override by contract. Moreover, this provision has to be made mandatory in order to harmonise the legislation of all the EU Member States and to make such services as international interlibrary loan of e-books lawful.

It has to be acknowledged that, just like the loan of works in tangible form, e-lending should be adequately remunerated to rightsholders. Any public-lending rights scheme should apply also to e-lending. The Digital Economy Act of the UK in 2010 has already extended the public lending scheme to audio books and loans of e-books on the premises of libraries, however it still has to be enacted. Moreover, further legislative changes should be made to take account of remote e-loans. At the same time, the methods used to establish the amount of money to be distributed towards the remuneration of rightsholders for lending should be reviewed when introducing a remuneration for e-lending. It would not be fair if the same amount of money were split among a broader base of rightsholders or would compensate for a bigger number of loans.

38. [In particular if you are an institutional user:] What differences do you see in the management of physical and online collections, including providing access to your subscribers? What problems have you encountered?

The management and availability of physical and online collections should be guided by the same principles, according to the public mandate of libraries. Libraries and other memory institutions will inevitably be required by society to provide broader and richer online services while at the same time maintaining the quality and availability of current services provided in person. Physical collections and services should therefore be recognised and treated as equivalent and complementary to digital collections and services. They should not be viewed as two contradictory ways of providing library services to the society.

40. [In particular if you are an institutional user, engaging or wanting to engage in mass digitisation projects, a right holder, a collective management organisation:] Would it be necessary in your country to enact legislation to ensure that the results of the 2011 MoU (i.e. the agreements concluded between libraries and collecting societies) have a cross-border effect so that out of commerce works can be accessed across the EU?

Yes. LIBER is a European organisation and can vouch that there is great interest amongst its members and their users across the EU in mass digitisation. Of the 50 European countries that we represent, we are only aware of 5 countries that have laws currently in place to allow mass digitisation (Norway, Sweden, Finland, Denmark and France). Germany passed a bill in 2013 to

introduce a law based on the MoU on Out of Commerce Works. The UK, meanwhile, is in the process of introducing Extended Collective Licensing (ECL) and may end up with a legal mechanism to allow mass digitisation but at the moment this depends on whether the UK collecting societies are in a position to, or choose to offer an ECL to facilitate mass digitisation.

The MoU has no legal status at all. The ability of a university or national library to lawfully digitise its collections and make them available online is entirely dependent on amendments to copyright law. As we have pointed out above, we are only aware of 5 EU countries that currently have such laws, plus the legislation in process in Germany.

Any legislation of this nature requires a cross border effect. There are a number of reasons for this:

- The digitisation may include rightsholders from more than just the country of origin;
- Mechanisms such as extended collective licensing and the French and German models are by definition representing orphan works and non-members and therefore a lawful cross-border effect is needed to ensure that their use outside the country of origin is lawful;
- As with the Orphan Works Directive 2012 and the Satellite and Cable Directive 1993, without legislation all Member States and their courts could determine that the communication to the public of digitised content from the above countries is a restricted act.

Not only do we believe that it is extremely important to have a vast range of material available online (not primarily American content as is currently the case) but there is a huge demand for this from citizens, scholars and publishers. Take Norway as an example. Its Bøkhylla book digitisation project has over 259,000 unique users (5.2% of the population), with 46 pages opened per visit. We strongly believe that the online availability of in-copyright works from across Europe will have a transformative and hugely positive effect on the educational, cultural, creative and financial life of this continent.

41. Would it be necessary to develop mechanisms, beyond those already agreed for other types of content (e.g. for audio- or audio-visual collections, broadcasters' archives)?

Yes. Researchers are interested in accessing not only written history but also our audio and audiovisual history, data and databases.

42. (a) [In particular if you are an end user/consumer or an institutional user:] Have you experienced specific problems when trying to use works or other subject-matter for illustration for teaching, including across borders?

Yes. In German universities, professors often complain about the uncertain rights situation concerning the use of images in slides/presentations, as well as citing other material or providing students with additional material. The German High Federal Court just recently ruled that no more than 10-20% of a book may be provided to students as additional teaching material. Moreover, publishers claim that there is no adequate remuneration for making use of their material (German law provides for a limitation with remuneration for right holders concerning the use of material for teaching).

Directive 2001/29/EC enabled the implementation of limitations or exceptions for the purpose of teaching, however the implementation of those limitations was not obligatory. This lack of obligation led to a different approach in every Member State. A study conducted by Professor Xalabarder[15] shows the heterogeneity of teaching exceptions all over the EU.
In countries like Spain, the teaching limitation has been narrowed. Teachers are only allowed to use small fragments of a work in a physical classroom if those works are not textbooks and if they are from a formal education institution. This implementation excludes many kinds of education many works. Moreover, the misunderstanding on the use of works for educational purposes has led to some Spanish universities being sued by collecting societies.

The directive allowed the implementation of an exception for reproduction in libraries when the copies were made for teaching and educational purposes but this exception does not exist in Spain and libraries are only allowed to make copies for research or conservation.

[15] http://www.wipo.int/meetings/en/doc_details.jsp?doc_id=130393

We would also point out that Europe's students are increasingly studying across borders and universities are increasingly establishing campuses abroad and teaching through the internet via Virtual Learning Environments. According to the UK Higher Education Statistics Agency there are 598,925 students studying for degrees outside the UK. Of these some 78,000 are in the European Union.

Studies from the Observatory on Borderless Higher Education show that American universities are increasingly targeting the EU as a good market for the growth of foreign campuses. While U.S. universities will no doubt assert the cross-border applicability of U.S. limitations and exceptions, including fair use where technology has developed ahead of specific limitations and exceptions, we believe that European internet-based learning will be at a clear competitive disadvantage given the fragmented nature of member state copyright laws and the lack of clarity around their cross-border applicability. There are of course a number of reasons why MOOCs in the U.S. are far ahead those of Europe (e.g. language) but one of the factors is certainly the flexibility of American limitations and exceptions.

43. If there are problems, how would they best be solved?

Legal uncertainly in the education sector must be addressed. A broader mandatory exception for teaching is necessary. It should be one that does not limit the kind of works or the kind of educational premises. The exception in Art 5.3. should be implemented uniformly across Member States. Introducing a European wide limitation for copying, teaching and making material available in the context of a virtual learning environment, intranet or extranet (combined with adequate remuneration) would be the best solution for dealing with cross-border issues.

Moreover, educational institutions in Europe are having to compete with educational institutions internationally because of the growth of distance learning online courses and MOOCs. This is difficult in Europe because eLearning across borders is difficult to support as licences tend to be regional.

44. What mechanisms exist in the market place to facilitate the use of content for illustration for teaching purposes? How successful are they?

Current solutions include collective licensing or agreements with publishers at national level. Such mechanisms are not comprehensive in terms of content or uses covered. They can be very expensive and time consuming to implement. Indeed, such is the lack of clarity around the coverage of these mechanisms that institutions can end up requesting permissions for uses which do not need to be licensed.

The different mechanisms available at national level (e.g. a flat fee per student system in Spain versus a collective licencing scheme in Finland) have resulted in inconsistencies in the use of material for teaching purposes across Europe.

45. If your view is that a legislative solution is needed, what would be its main elements? Which activities of the beneficiary institutions should be covered and under what conditions?

Any legislative solution should include a broad mandatory exception for teaching which would facilitate the fair and proportionate use of in-copyright works. As pointed out above VLEs are already widely used in education in Europe.

This mandatory educational exception should cover all uses of all types of works for teaching purposes, regardless of the type of educational institution. It is important to stress that uses of computer programs, databases and multimedia works (such as video games) should be expressly included.

Activities related to eLearning should also be covered (enlargement of limitation to citation) and should apply to all Member States.

Another complementary solution is that any author should be allowed to reuse her own material, overriding copyright agreements with a publisher, in educational setting such as classrooms or in a VLE.

46. If your view is that a different solution is needed, what would it be?

Besides the aforementioned legislative solution, any teaching material or educational resource funded by public money should be disseminated openly.

47. (a) [In particular if you are an end user/consumer or an institutional user:] Have you experienced specific problems when trying to use works or other subject matter in the context of research projects/activities, including across borders?

Yes.

SUBITO CASE. International document supply by libraries.
This case illustrates the legal uncertainty in providing international researchers and libraries with copies of articles for non-commercial research .While most, if not all, EU countries allow a library to supply articles to users for non-commercial / research purposes, the cross-border situation is unclear given the lack of uniformity of Member State copyright laws. This legal uncertainty forces libraries to use licensing as the only way to guarantee a cross-border supply of documents for research purposes. Publishers are able to exclude titles and countries in the license and charge fees that are simply not appropriate for students and teachers but rather set at commercial rates for businesses.

SUBITO is a library document delivery service of the German speaking countries. SUBITO delivered copies of documents to final users and libraries (both foreign and national) for (among other reasons) non-commercial uses such as research. Legal action was taken against SUBITO in 2006, questioning the compliance of its electronic document (pdf) delivery services with German copyright law and foreign copyright laws when the delivery was cross border. In relation to international deliveries, the court said the legal situation was determined according to both German and foreign copyright law, making it in practice extremely complicated for Subito to comply with all national laws.

Currently, international delivery to libraries in Germany, Switzerland, Austria and Liechtenstein is allowed for non-commercial purposes. Subito has negotiated licences with publishing houses to enable Subito to offer copies (by post, fax or electronically with Technological Protection Measures). The agreement provides for royalties to be paid for each copy that is sent and it allows the publishing houses the possibility to exclude certain publications from the licence. According to Subito's site: "the alternative solution of seeking to achieve a sufficient degree of legal certainty by having the legal situation verified in all of those countries to which subito e.V. delivers would have entailed devoting a disproportionate amount of time and effort".

While as an economy we should be promoting the knowledge economy we find it at odds with the principle of a single market that an individual in Germany can be supplied with an article from a library in Berlin for their own non-commercial research purposes, but a student in Dublin, Athens, Madrid or Warsaw would not be able to receive the same article.

OpenAire
OpenAIRE is an EU-funded open-access infrastructure for research that supports the European Commission's Open Access policies (FP7 pilot on OA and H2020 OA to publications policy and Open Data pilot). The infrastructure aims to support open scholarly communications and provide worldwide access to the research output of European projects and open-access content (publications and research data and datasets) from European repositories, data centres and publishers. The project gathers open-access content from almost all European Union countries.

OpenAIRE has published a study[16] that analyses the potential impact that the current copyright and database directives might have on this research e-infrastructure and the re-use of data.

- **Copyright directive and open access research e-infrastructures:** In the event that copyright protected data is used within an open access scientific information e-infrastructure, exclusive rights such as reproduction, communication and adaptation might be infringed. The study concludes that even though the InfoSoc Directive gives Member States the opportunity for exceptions or limitations to the reproduction right and the public communication right for scientific purposes, these exceptions are adopted very differently. Under the current copyright framework, it is not possible to create an e-infrastructure that respects copyright laws and falls under exceptions and limitations in every European country.

[16] http://www.openaire.eu/en/about-openaire/publications-presentations/public-project-documents/doc_download/621-openairelicensing-studysummary

- **Sui Generis database protection right and open access research e-infrastructures:** If the data used in the infrastructure is part of a database (protected by sui-generis right), rights of the of extraction and re-utilisation are restricted. As with the copyright directive, Member States can include an exception to these rights in their national laws for research purposes, but these exceptions are not harmonised within the EU, making it impossible for the infrastructure to comply with every national law.
- **Specific uses of the open access research-infrastructure by end users:** Some cases of exploitation of the information offered contained in the infrastructure have been described as potentially infringing copyright and the database directive, and national copyright and databases laws such as access, linking (e.g. papers to research data), mining and re-use.

48. If there are problems, how would they best be solved?

According to the above-mentioned study, the barriers to fully exploiting an open-access e-infrastructure for research described above could be solved by:

1. Introducing a new broader mandatory research exception at the European level and notably by widening the limitations to the database directive to infrastructure operators such as non-commercial scientific databases.

2. Making limitations of the database directive mandatory so that data (and databases) can be made available to the public on a non-commercial basis.

Also, allowing publicly-funded research results to be made openly available regardless of contracts signed with a publisher would contribute to knowledge being widely used for research purposes. By increasing accessibility and availability, and by facilitating the use and reuse of content, open access will play an integral role in driving research excellence globally. Copyright law should not hinder its progress. An exception or legal instrument should be developed which ensures that researchers can retain copyright in order to make the results of publicly-funded research available openly, regardless of contracts signed with a publisher.

53. (a) [In particular if you are an end user/consumer or an institutional user:] Have you experienced obstacles, linked to copyright, when trying to use text or data mining methods, including across borders?

Yes. There are three separate mechanisms which severely limit the act of text and data mining in Europe.

1. **Copyright Law:** Text and data mining is concerned with the extraction and subsequent analysis of facts and data. In order to do this, it is necessary to make a copy of the content in order for a machine to extract the relevant facts and data. This means that text and data mining is effectively precluded under copyright law in Europe.

2. **Database Directive:** This is another legal barrier to text and data mining because it also protects against the copying of large portions of databases.

3. **Technical protection measures (TPMs):** These prevent the downloading of large amounts of content and are also preventing the application of text and data mining techniques.

The combination of the above factors means that the potential results and benefits of text and data mining research are currently being suppressed because of a lack of legal clarity.

Even if content is openly available on the Web, it is unclear if the copying of this content for the purpose of text and data mining is legal (unless permission is expressly given). Content licenced under open access licences such as cc-by-sa or cc-by-nc may also inadvertently prevent users from mining the content. This can be because of the lack of clarity in the law about the use of facts and data and their requirement to acknowledge, or because conditions such as Share Alike are almost impossible to implement when mining across a corpus.

As highlighted by both the Japanese and UK governments (the only two countries in the world who respectively have or are introducing a specific limitation and exception for text and data mining), this is an area where a law designed for the printing press does not seem able to cope proportionately or adequately with new technologies.

The following points show that in the case of data analytics it is not enabling of innovation to "stretch" a law designed to regulate paper printing and the artistic expression of an author.

- Some forms of text and data mining produce probabilities / hypotheses from what is not written about in the text. We would argue that a law that regulates artistic expression is not applicable to an activity that uses what has not been expressed.
- The activity is lawful if performed using a pen and paper as only facts and data are extracted – it is only because the EU Copyright Directive regulates the fact that a computer must copy, irrespective of the financial harm the simple act of copying with no communication to public over and above what is lawful, that copyright law and database law is relevant.
- Text and Data mining does not use the artistic expression that copyright law aims to protect but analyses the underlying facts (that one has lawful access to already).
- Copyright law allows rightsholders to control how their works are distributed and to receive financial rewards in return for their investment in the production of the work. These financial rewards then lay the foundation for further investment and innovation to take place. In the case of text and data mining, where access to the work has already been provided, it is difficult to argue that innovation will be encouraged by further restricting access to the work.
- Applying the logic of the Svensson judgement, that no infringement took place because there was no reuse of the work by "a public that was not taken into account by the copyright holders when they authorised the initial communication to the public, it cannot be argued that the reading (using robots or otherwise) of online content by those who have lawful access to it was not foreseen by the licensor / copyright holder.

We believe that text and data mining really has nothing to do with the principles behind copyright and database rights. It needs to be recognised that the single reason it does is because mere copying by machines and an arrangement of content (irrespective of whether any harm is done or not during the course of their use) has become subject to copyright law.

54. If there are problems, how would they best be solved?

For the sake of transparency and reproducibility, researchers must be able to share the results of text and data mining, as long as these results are not substitutable for the original copyright work. This ability to share should be irrespective of copyright law, database law or contractual terms to the contrary. A specific exception to this effect allowing the copying of content for the purpose of text and data mining is necessary. Such an exception should not distinguish between commercial and non-commercial purposes as the definition of 'non-commercial' is impossible to clarify and, for research institutions, would certainly prevent knowledge transfer.

55. If your view is that a legislative solution is needed, what would be its main elements? Which activities should be covered and under what conditions?

A legislative solution should take the form of an exception which allows the copying of content and the circumvention of TPMs for text and data mining purposes. This exception must not be over-rideable by contracts and should not distinguish between commercial and non-commercial purposes.

56. If your view is that a different solution is needed, what would it be?

Licencing will not work as an alternative solution. The negotiation of licences is resource intensive and unscalable. TDM licence offerings, such as that provided by Elsevier, are unreasonable in that they (1) require researchers to provide data to a third party about their research activities, (2) limit the manner in which their research output may be made available, and (3) place arbitrary limits on the number of articles an institution may mine in a given timeframe.

Licences are also unable to address the current ambiguity around the mining of the open Web.

Some publishers have argued that licences are necessary because they fear that increased traffic and downloading resulting from crawling for the purpose of text and data mining will reduce the performance of their infrastructures. LIBER has consulted with several open access publishers on this issue. Although their business model is different from subscription publishers, their infrastructure is similar. In fact, because their content is freely available for researchers to mine it could be argued that they may have more experience in dealing with crawling for the purpose of mining. These publishers claim that content mining is, in fact, a trivial and easily managed source of traffic when compared to human traffic.

In support of our response to the questions on text and data mining, PLOS, eLife and Figshare have cosigned a document which outlines several potential solutions to manage any significant rise in traffic that may occur as a result of mining. These are best-practice solutions which LIBER strongly supports and can help to implement. The document is included in the Annex to this response. The solutions outlined in the document are briefly described in the following points:

1. **Data Dumps:** Because of our central role in digital preservation and the provision of infrastructure for open access depositing, many libraries already have the infrastructure in place for the depositing of data and could, if necessary, take dumps of content from publishers in order for researchers to crawl that content locally.

2. **Best Practice for Crawling:** It is in nobody's interest to deploy robots for text and data mining irresponsibly. Many websites provide information on the required crawl delays which indicate how many seconds should be left between the downloading of new pages. If publishers were to provide this information to researchers, the researchers could modify their crawls accordingly. As technology and infrastructure improves, the crawl delays should shorten. Libraries are willing to facilitate and communicate the development of best practice in the use of robots for text and data mining.

3. **Technical measures:** There are measures currently in place to deal with rogue activity. If best practice for the use of robots has been agreed and crawl delays are not respected, then it is perfectly reasonable for the content provider to deploy technical measures to throttle this activity.

It is in the interest of research libraries and the researchers they support to ensure that their information infrastructure works efficiently and is not compromised by irresponsible research practices.

57. Are there other issues, unrelated to copyright, that constitute barriers to the use of text or data mining methods?

There is still a gap in the availability of technology and skills for text and data mining. As legal barriers are reduced, market-led solutions to fill these gaps will emerge. An increase in the availability of open-access content (articles and data) and infrastructure will also make text and data mining easier.

76. In particular, is the current legal framework clear enough to allow for sufficient involvement of intermediaries (such as Internet service providers, advertising brokers, payment service providers, domain name registrars, etc.) in inhibiting online copyright infringements with a commercial purpose? If not, what measures would be useful to foster the cooperation of intermediaries?

No. Any legislation that relates to intermediaries needs to carefully consider the roles, obligations and safe harbours and how they pertain to the many different types of online intermediary. We believe that the European Commission and Parliament could learn much from the mistakes made by the UK government who, in the drafting of the Three Strikes Legislation, did not differentiate between intermediaries such as ISPs and mobile telephone companies, and organisations such as libraries, educational establishments, local governments, cafes and hotels. All of the above types of organisations can act as internet intermediaries.

We believe that a concept such as "public intermediary" may well be a useful concept when and if further consideration of the roles, responsibilities and liabilities is undertaken by the European legislature.

78. Should the EU pursue the establishment of a single EU Copyright Title, as a means of establishing a consistent framework for rights and exceptions to copyright across the EU, as well as a single framework for enforcement?

No opinion. LIBER is tentative in its support of the establishment of a single EU Copyright Title in order to overcome territoriality issues in the internal market. Harmonisation and mandatory exceptions should be prioritised over the establishment of a copyright title. If such a title were to be developed its scope should be limited to addressing the internal market issues (such as mandatory exceptions) and leave flexibility for cultural, social and educational issues to be addressed at Member State level.

79. Should this be the next step in the development of copyright in the EU? Does the current level of difference among the Member State legislation mean that this is a longer term project?

A pragmatic approach should be taken. A general move towards harmonisation (within the EU and globally) should be supported, whereas priorities in relation to European regulation should be set according to the ability of the market to self-regulate. Reasons to regulate should be restricted to inefficient markets and not be guided by calls from actors in such markets who fear losing their position through market reconfiguration. The development of a copyright title should not be rushed and should involve further consultation with stakeholders.

LIBER believes that the first priority should be the revision of the Information Society Directive to include additional exceptions and to clarify and make existing exceptions mandatory. The Database Directive should also be revised.

Exceptions must be harmonised across Member States. Interpretive guidelines should be produced in order to encourage EU countries to promote more flexibility and a user-friendly approach. Harmonisation must not result in the reduction of flexibility in current exceptions. In our view a less flexible copyright regime would seriously hinder research and innovation in Europe, more than is already the case.

80. Are there any other important matters related to the EU legal framework for copyright? Please explain and indicate how such matters should be addressed.

1. A major concern for LIBER and its members is that technological protection measures may prevent copyright exceptions from being enforced. Provisions included in licence agreements can have the same effect (please see analysis from the British Library[17]).

   Much has been made of the necessity for a digital single market at a European level. Effort invested in creating a legislative framework to facilitate this market will be for nothing if that framework can be undermined by heterogeneous licence agreements.

   Libraries and the researchers they support need to have confidence that they can operate and innovate within the legal framework. Exceptions reinforce this confidence and should not be undermined by supplicating legislation to licences.

   On a practical note, it is unrealistic to expect individual researchers to become experts on licensing terms and their implications (libraries employ dedicated full-time staff for this), yet this is exactly what click-through licences for text and data mining, for example, expect of researchers. This type of approach can mean that research results are repressed because of legal uncertainty.

   In summary, LIBER strongly recommends that the European legal framework explicitly prevents licences from overriding exceptions and that the copyright framework should also recognise the occasional need for libraries to use or create software that can help circumvent these measures for the purpose of making and facilitating lawful use of a work.

2. As reflected in our answers, the copyright and IP framework can have a major impact on research and cultural heritage activities. In particular we believe that any revision of the copyright framework should be considered within the context of the major investments and policies being developed in relation to research infrastructures at a European level and should be developed in collaboration with other Commission Directorates such as DG Research & Innovation and DG Connect.

[17] http://pressandpolicy.bl.uk/imagelibrary/downloadMedia.ashx?MediaDetailsID=2137

ANNEX 1: Best Practice In Enabling Content Mining

A pure Open Access publisher perspective, as signed by PLOS, eLife and Figshare.

A. Introduction

Enabling the discovery and creative re-use of content is a core aim of Open Access and of Open Access publishers. For those offering Open Access publication services enabling downstream users to discover and use published research is a crucial part of the value offering for customers. Content mining is an essential emerging means of supporting discovery of research content and of creating new derivative works that enhance the value of that content.

Content mining generally involves the computational reading of content, either by obtaining specific articles from a publisher website or by working on a downloaded corpus. Computational access to a publisher website has the potential in theory to create load issues that may degrade performance for human or other machine users.

This issue needs to be actively managed to enable the full potential of content mining. Traffic management of this sort is part and parcel of running a modern public-facing website and there are well established means of doing this.

Downloads that result from crawling and content mining contribute a trivial amount to the overall traffic at one of the largest Open Access publisher sites and are irrelevant compared to other sources of traffic. This is true both of average traffic levels and of unexpected spikes.

Managing high traffic users is a standard part of running a modern web service and there are a range of technical and social approaches to take in managing that use. For large scale analysis a data dump is almost always going to be the preferred means of accessing data and removes traffic issues. Mechanisms exist to request automated traffic be kept at certain levels and these requests are widely followed - where they are not technical measures are available to manage these problematic users.

B. Scale and scope of the problem

PLOS receives around 5 million page views per month to a corpus of 100,000 articles. Scaling this up to the whole literature suggests that there might be a total of 500 million to 5 billion page views per month across the industry, or up to seven million an hour, largely from human users. As noted below the largest traffic websites in the world provide guidance that automated agents should limit retrieving pages to a specified rate. Wikipedia suggests one page per second or 360 an hour, PLOS requests a rate of twelve an hour.

PLOS infrastructure routinely deals with spikes of activity that are ten times the average traffic and is designed to manage loads of over 100 times average traffic without suffering performance problems. Thus it would require hundreds of thousands of simultaneously operating agents to even begin to degrade performance.

Content mining is a trivial and easily managed source of traffic compared to other sources, primarily coverage on popular social media sites. Coverage of an article on a site like Reddit often leads to tens of thousands of requests for a single page within an hour. By contrast automated crawling usually leads to a smaller number of overall downloads and is spread out over longer time periods making it much easier to manage. As an example there are attempts made to artificially inflate article download counts, which involve tens of thousands of requests for the same article. We do not even attempt to catch these at the level of traffic spikes because they would be undetectable, they are detected through later analysis of the download data.

Sources of traffic that do cause problems are generally rogue agents and distributed denial of service attacks where hundreds of thousands or millions of requests occur per second. These sources of traffic are the main source of service degradation and need to be managed based on the scale of traffic and the likelihood of being a target for such attacks. The scale of content mining traffic for any given publisher will be dependent on the scale of interest in the content that publisher is providing.

C. Management approaches

There are broadly three complementary approaches to supporting content mining in a way that does not have any impact on user experience. While all of these approaches are implemented by effective scholarly publishers, it is worth examining these approaches in the context of a truly high-traffic site. Wikipedia is an excellent example of an extremely high traffic site that is also subject to large scale mining, scraping, and analysis.

1. **Providing a Data Dump:** The first and simplest approach is to provide a means of accessing a dump of all the content where it can be obtained for off line analysis. Generally speaking the aim of analysis is to mine a whole corpus and

enabling the user to obtain a single dump and process this offline improves the experience for the miner while removing any risk of impact to website performance. Wikipedia provides a regular full dump of all content for all language versions and recommends that this be the first source of content for analysis. Open Access publishers adopt a similar strategy utilising deposition at Pubmed Central or on their own websites as a means of providing access to a full dump of content. PLOS recommends that those wishing to parse the full corpus use PMC or EuropePMC as the source of that content.

This approach is especially useful for smaller publishers running their own infrastructure as it means they can use a larger third party to handle dumps. Of course for smaller publishers with a relatively small corpus the scale of such a data dump may also be such that readily available file sharing technologies suffice. For a small publisher with a very large backfile the imperative to ensure persistence and archiving for the future would be further supported by working with appropriate deposit sites to provide both access for content miners and preservation. Data dumps of raw content files are also unlikely to provide a viable alternative to access for human readers so need not concern subscription publishers.

2. **Agreed Rates of Crawling:** It is standard best practice for any high traffic website to provide a "robots.txt" file that include information on which parts of the sites may be accessed by machine agents, or robots, and at what rate. These files should always include a 'crawl-delay' which indicates the time in seconds that an agent should wait before downloading a new page. Wikipedia's robot.txt file says for instance "Friendly, low-speed bots are welcome viewing article pages, but not dynamically-generated pages please" and suggests a delay of at least one second between retrieving pages. This is not enforced technically but is a widely recognised mechanism that is respected by all major players - not following this is generally regarded as grounds for taking technical measures as described below.

   PLOS request a crawl delay of 30 seconds currently, Biomed Central asks for one second, eLife for ten. For a large publisher crawl delays of this magnitude means that it is more sensible for large scale work to download a data dump. Where a smaller number of papers are of interest, perhaps a few hundred or a few thousand then the level of traffic that results from even large numbers of content mining agents that respect the crawl delay is trivial compared to human traffic from other sources.

3. **Technical Measures:** Some actors will not respect crawl-delays and other restrictions in robots.txt. In our experience this is rarely the case with content miners and much more frequently the result of malicious online activity, rogue robots, or in many cases the result of security testing at research institution which sometimes involves attempts to overload local network systems.

   Whether the source is a spike in human traffic, malicious agents, or other sources of heavy traffic maintaining a good service requires that these issues be managed. The robots.txt restrictions become useful here as when it is clear that an agent is exceeding those recommendations it can be shut down. The basic approach here is to "throttle" access from the specific IP address that is causing problems. This can be automated, although care is required because in some cases a single IP may represent a large number of users. For PLOS such throttling is therefore only activated manually at present. This has been done in a handful of cases, none of which related to text mining.

   At larger scale automated systems will be needed but again this is part of running any highly used website. Load-balancing, monitoring incoming requests and managing the activity of automated is a standard part of running a good website. Identifying and throttling rogue activity is just one part of the suite of measures required.

Summary

Enabling content mining is a core part of the value offering for Open Access publication services. Downloads that result from crawling and content mining contribute a trivial amount to the overall traffic at one of the largest Open Access publisher sites and are irrelevant compared to other sources of traffic. This is true both of average traffic levels and of unexpected spikes.

Managing high traffic users is a standard part of running a modern web service and there are a range of technical and social approaches to take in managing that use. For large scale analysis a data dump is almost always going to be the preferred means of accessing data and removes traffic issues. Mechanisms exist to request automated traffic be kept at certain levels and these requests are widely followed - and where they are not technical measures are available to manage these problematic users.

There are sources of traffic to publisher websites that can cause problems and performance degradation. These issues are part of the competent management of any modern website. Content mining is not, however, a significant source of issues at the current time.