

**European Commission
Project:**

**Economic and Social Impact
of the Public Domain**

Introduction to Methodology

First COMMUNIA Conference 2008

Assessment of economic and social impact of digital public domain
throughout Europe

Louvain-la-Neuve, Belgium

*All members of the project team, which consists of the following people,
have contributed to the content of this paper:*

Rob Davies, MDR Partners (UK)

Martin Katuscak, Slovak National Library

Luiss Guido Carli University, Faculty of Law (Italy):

Emanuela Arezzo

Gustavo Ghidini

Rufus Pollock, Mead Fellow in Economics at Emmanuel College,
Cambridge (UK)

Rightscom Limited (UK):

Mark Isherwood

Hugh Look

Sue Sparks

Tom Rivers, Rivers Consultancy (UK)

Paul Stepan, Austrian Society for Cultural Economics and
Policy Studies (Austria)

Paul Torremans, University of Nottingham, Faculty of Law (UK)

Mikko Välimäki, Turre Legal (Finland)

1 Introduction

The project entitled Economic and Social Impact of the Public Domain in the Information Society commenced work on 1st January 2008 and is due to provide its final report to the European Commission by the end of April 2009. At this time, June 2008, the principle area of work for the project team is the gathering of data to enable the analysis to be undertaken to estimate the number, value and use of public domain works. We expect this data gathering process to continue through until September. Analysis is expected to commence in July and will very much be an iterative process with on-going data gathering activities.

In addition, the project team has been working on a stream of activity which will seek to assess the capacity of European cultural institutions to implement the principles for re-use as established in the Public Sector Information Directive. This is expected to be completed by the beginning of August. Finally, early stage data gathering has commenced on the stream of work which will focus on the identification and analysis of the currently available mechanisms for voluntary sharing of content.

The focus of the first few months of the project has been the development of methodologies and scoping of the requirements for data gathering. Given that there are no results to discuss, this paper focuses on the methodologies the project team will be taking with the various streams of work.

2 Estimating the size, usage and value of the public domain

The main quantitative part of the project is concerned with estimating the size, usage and, more importantly, value of the public domain within Member States of the European Union. Let us begin by making two general points. First, the intricacies and variations of copyright law across member states (even post-harmonisation) are such that the exact delineation of the public domain in each jurisdiction would be a major task. However, since our aim here is to obtain general quantitative *estimates*, we intend to take a 'broad' approach which ignores these subtleties and relies on a simple set of algorithms (for example, all works of authors who died before 1937, or simply all works published before, say 1870).¹ This will allow more time to be spent trying to estimate the value of the public domain, rather than having to spend time describing the intricate data requirements to potential data suppliers. Secondly, and as a related matter, these questions must obviously be treated separately for each of the primary categories of protected material; library works, musical compositions, sound recordings and film. Furthermore, within each of those categories it may be useful, and important, to distinguish further, for example between books and other printed media such as journals or newspapers, between sound recordings and compositions, or between multi-media and films.

2.1 The size and usage of the public domain

Broadly speaking then published literary, musical and artistic works by authors who died before 1937 will now be in the public domain. At the local level, moreover, cultural institutions have always compiled data records of material which has been acquired within a specified accounting period. With appropriate statistical sampling techniques, and taking account of particular events such as the two great wars of the twentieth century, it should be possible to produce reasonable estimates of the number of works in the public domain.

Sound recordings, fixations of performances and works of still photography and cinematography have to be considered separately. This is firstly because the technologies involved in capturing sound and still and moving images are relatively recent by comparison with the original Gutenberg technology, dating from the first quarter of the nineteenth century in the case of photography and from the last quarter in the case of sound recordings and cinematography. There are also formidable conservation

¹ Given the many other approximations that will no doubt be necessary the error induced by this particular assumption is likely to be, relatively, very small.

problems with film stock. So material of this kind in the public domain is a small category both in absolute terms and relative to subsequent production which remains rights-protected.

Thus, it is likely that data availability will vary significantly by the type of work. Returning to the example of photography, many photographs, both old and new, are anonymous and it is all but impossible to trace an author. The issue is further complicated by specific term provisions in the Directive relating to anonymous works. As such it will be impossible to proceed on the basis of an authorial death date. This point holds more generally: even for those works where authors are readily identifiable, depending on the quality of the available data it may be hard to match works with a specific authorial death date. If this is the case, it may be simpler, and almost as effective to simply do computations based on simple cut-offs: for example, that all work published before 1870 is in the public domain.²

With regard to seeking to measure the use of public domain material this, in all probability, will take place by virtue of case studies. Some information of this nature will become apparent as a result of the work being carried out with regard to the PSI Directive (see Section 3). Care will have to be taken in determining for this purpose what is “use” and what is “re-use”. Our preliminary view of the nature of the distinction is that “use” is where public domain material is simply exploited in its original form. This might include for example, the novels of Charles Dickens, re-printed in their original form with no editorial or new artistic contribution. It may therefore be possible to gain some information from publishers specialising in re-prints of this nature and a similar approach could be taken with public domain music.

“Re-use” we see as the creation of a new work where a public domain work has in some way been adapted or integrated with other material to create the new work. An obvious example of this technique is in music, particularly when “sampling” is involved. There is no doubt, however, that there will be types of usage which do not neatly fall into these definitions but the types of examples given will be useful ways of gaining some indications of the re-use of public domain material.

Outside of the cultural institutions and the types of commercial use alluded to above the other main area where there is use of public domain material is in education. The most obvious way in which public domain material might be used in this context is for the creation of course work material by teachers or pupils. Gaining widespread data relating to this type of use is likely to be somewhat difficult, so again, attempts will be made to gather data to create case studies. Apart from contact with individual schools it may be possible to gain some data from the examination boards in different countries which demonstrates the extent to which their curricula draw on public domain material. Similarly, in the UK the copyright legislation allows for the licensing of “off-air” recording of broadcast material by schools and colleges which is carried out by the Educational Recording Agency. Whilst the organisation’s focus is copyright material it may be that their sampling can throw some light on the extent and nature of the use of public domain material in educational establishments across the UK.

2.2 The value of the public domain

The more important, and the more difficult, aspect of this study is to estimate the *value* of the public domain.

2.2.1 THE GENERAL CONCEPT OF VALUE

When we talk of ‘value’ it is important to be clear about what we mean. Here the term value is interpreted to mean the social value (or welfare), which is the usual meaning attributed to the term value by economists. This value, be it of an apple or a novel, is the value derived by a user from its employment or enjoyment – often approximated in monetary terms by willingness-to-pay (WTP) – net of the costs of producing the good. For goods with an associated price, social value may in turn be divided into ‘user value’ (consumer surplus) which is defined as the value to the user of the good net

² One could obviously adjust the ‘conservativeness’ of this approach by varying the exact cut-off used. For example, one could compare the effect using a 1850 versus an 1870 versus a 1900 cut-off date.

the price paid for it, and ‘commercial value’ (producer surplus) which is defined as the price minus the cost of producing the good (the seller’s profit).

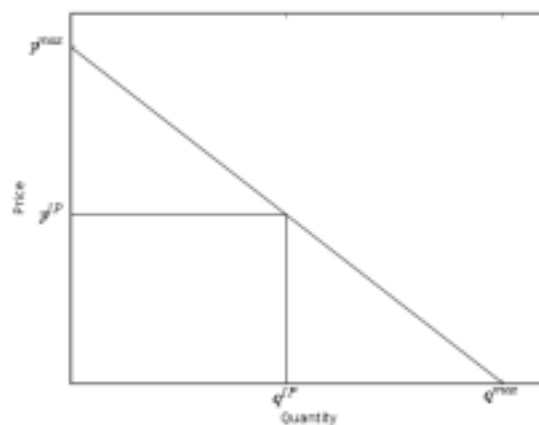
Thus the value of a good may be quite different from the price paid for it. This is an important distinction to keep in mind, for it is not unusual to see the value of an activity being equated to the income received rather than to the utility generated for society. To illustrate this difference consider the case of a novel that goes out of copyright and enters the public domain. Suppose before this occurred the novel was sold for £10 in bookshops but afterwards it is sold for £5 and is also available for free on the internet. Sometimes it is suggested that this results in a reduction in the value of that work for society since before the work was ‘worth’ £10 but now is ‘worth’ only £5 or even nothing.

From the above this can be seen to be completely false. The value of the work has not changed at all. All that has happened is that the price has dropped. A consumer who previously valued the book at, say, £15 and who paid £10 and was left with £5 of ‘consumer surplus’, now pays £5 (or £0) and is left with £10 (or £15) of ‘surplus’.

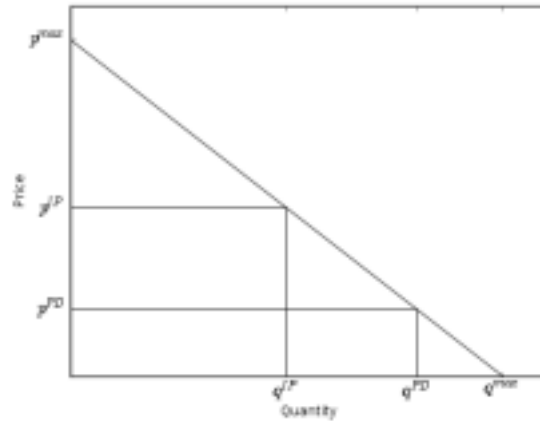
Furthermore, the reduction in price means that those consumers who valued the work at less than £10, and therefore did not buy at the original price, will now be able to purchase it. For each such consumer society gains the entire value they put upon the good (net of costs). Aggregating the valuations of all of these individuals who only get access at the lower price gives the total value to society of having this lower price. Conversely when a monopoly – or some other regulation – restricts access to a good there is a consequential loss to society (termed the deadweight loss).

2.2.2 VALUING A WORK, AND HENCE THE PUBLIC DOMAIN: THEORY

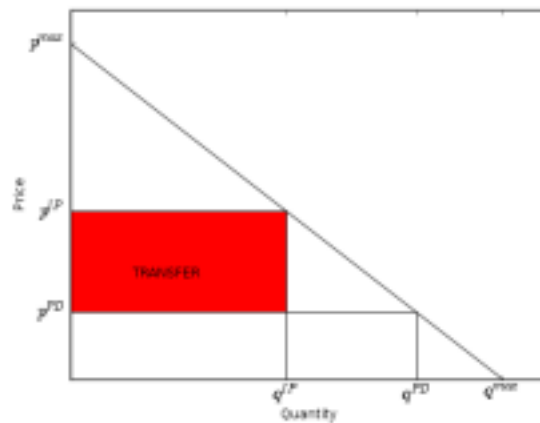
The value of a work (at a given price) then is the value of that work to each user summed across those users who gain access to that work (at that price).



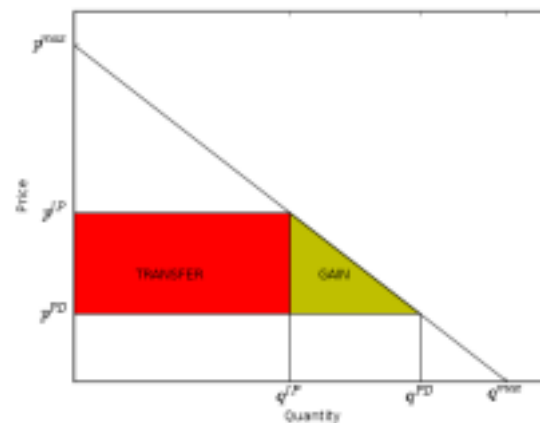
This is illustrated in terms of a classic linear demand curve. The first of these shows a single price and quantity (here superscripted IP but this can be ignored at the present). The value of this work would then be equal to the gross surplus generated by the work (equal to the area under the demand curve up to q^{IP}) minus the cost of producing it.



However here we are not simply interested in the value of the work but the value of public domain. Focusing on an individual work we take this to be the value generated by the work when it is in the public domain minus its value when under copyright. As shown by the figure when a work enters the public domain we expect its price to drop.



This has two effects. First, there is a transfer of value from the current owner of the exclusive right to users.



Second there is a gain in value as new users who previously did not purchase at p^{IP} gain access to the work (as this value is “lost” when access to the work is restricted by an IP right such as copyright).

Thus, the value of the public domain measured by reference to this work is the gain from greater access to it. In addition, strictly we should also take into account two additional related effects. First, one should take account of any reduction in the production of work arising from the income lost by the

owner of the IP right when the work enters the public domain. However, given the current length of copyright (and related rights) this effect is so slight that it can be ignored for the purposes of calculations in this study. Second, entry in the public domain of a work permits not only greater access, but greater re-use. The benefits of such re-use may not be adequately captured in the basic demand curve for the work³ and so the simple calculation described above would lead to an understatement of the public domain's value. Unfortunately, while this effect may well be important – after all there is substantial anecdotal evidence for the importance of public domain material in the production of new works – it is extremely difficult, at least with present data, to quantify.⁴ Hence, it is likely that this effect, just like the first, will generally be ignored in this part of the project (though it should be emphasised that other parts, especially those providing case-studies, will be dealing with this in some detail). Having established how to determine the value of the public domain for a given work the next step is to aggregate across works.

2.2.3 VALUING A WORK, AND HENCE THE PUBLIC DOMAIN: EMPIRICS

In a perfect world one would proceed by calculating a demand curve individually for each work in the public domain. At the same time one would determine the price and level of access both when in the public domain and when access is restricted by an IP right. Together this would permit us to calculate the net welfare gain and hence its public domain value. Finally we would then aggregate these values across all works. Unfortunately, given the state of available data such a task is clearly impossible. Rather, the aim must be to derive as good an *estimate* as possible given the material available.

First, it will likely be necessary to focus down on a few key types of works that are widely available and used, for example books and sound recordings. Next for each type of work it will probably be best to divide the analysis into two parts:

- (i) Estimate the deadweight loss of copyright (conversely the value of the public domain) generally. That is, estimate this as a function of a few key parameters defining the demand curve; and
- (ii) Combine this figure with estimates of how these underlying parameters vary across the population of works.

To elaborate a little; on the first point the key issue will be to estimate the demand curve, or some portion thereof, for the set of cultural works under consideration along with the likely price changes. Where data is insufficient to estimate the full demand system (as is likely to be the case) it may be possible to fall back to a reduced form approach in which one seeks direct estimates for key variables (price changes, demand elasticity etc.) and then combine these with a particular parameterisation for the demand curve. Even to estimate these parameters it is highly unlikely that one can proceed on an individual work by work basis. Instead, it will be necessary to use a whole set of works in order to deal with the very heterogeneous nature of cultural production.

For example, suppose one wishes to calculate the effect of copyright on price (an area on which little empirical work has yet been done). There are two obvious approaches to take. One approach is to track works as they move from the proprietary to the public domain. The second approach is to try and select a set of 'matching' works where some are in the public domain and some not and then compare the prices of the two sets of work. In both cases one needs a fairly large sample size to get any kind of robust estimate.⁵

2.2.4 SENSITIVITY ANALYSIS AND GROWTH OF THE PUBLIC DOMAIN UNTIL 2020

As discussed above all the figures obtained will necessarily be *estimates*. Hence it will be important to perform some degree of sensitivity and robustness analysis. This will be particularly true where there is substantial uncertainty about parameter values or the assumptions chosen. In all these cases

³ This is for several reasons, most obviously that downstream re-users are unlikely to be capturing the full surplus generated by their activities.

⁴ Remember that it is not sufficient simply to show that re-use of public domain material occurs but it is necessary to calculate the net benefit: i.e. the increased cost (and hence reduced creativity) of going to the next best alternative source.

⁵ There is a very large degree of price variation found for these types of products which is *not* a result of the work's copyright status but is due to other factors. To 'filter' this out then requires a large dataset (crudely).

sensitivity analyses will be conducted. This will provide additional ‘distributional’ information (e.g. confidence intervals) about outcome variables beyond the basic point estimates. Furthermore it will also provide full information about the robustness of the results.

For example, if there were uncertainty about the exact value of the elasticity of demand, instead of using a simple point estimate like -1.0 this would be replaced with some distribution function, for example that it was uniformly distributed between -0.5 and -1.5. This distribution would then be combined with others to yield an overall probability distribution for the variables of interest (social welfare etc.). Not only would this present the results in a clearer manner but it would also allow for robustness to be checked in a simple and straightforward manner.

As part of our sensitivity analysis we will also undertake projections of the growth (both in size and value) of the Public Domain over the period up until 2020. This work should be able to be performed using data already obtained for other parts of the project though it will obviously necessitate collection of data on the number of works for (approximately) another two decades after the public domain cut-off relevant for today.

3 *Re-use of content held by cultural institutions*

In December 2003 the EU adopted the PSI Directive (2003/98/EC). The Directive aimed to put in place a framework for the conditions governing re-use of public sector information in order to ensure fair, proportionate and non-discriminatory conditions for the re-use of such information. Under Article 13 of the Directive, the European Commission is required to carry out a review of the application of the Directive in which its scope and impact will be addressed. The results of this aspect of the study’s work are required to inform this review.

Under Article 2 of the Directive, certain types of content were exempted, including: ‘*documents held by cultural establishments, such as museums, libraries, archives, orchestras, operas, ballets and theatres*’. This exemption was incorporated during the proceedings in Parliament, as a result of strong opposition from some Member States and lobbying by some stakeholders, public content holders in particular. Other exemptions under Article 2 of the Directive relate to access, secrecy, educational and research organisations and intellectual property rights of third parties.

Given that the Directive is specifically aimed at valuable PSI with high exploitation potential this exemption is seen by some as running counter to the broad direction of the Directive’s overall policy: cultural institutions hold probably some of the most valuable PSI of all, which has already been identified as a catalyst for creativity, growth and jobs. In a recent independent study carried out for the European Commission it was estimated that, in 2004, more than 5 million people worked in the cultural sector, equivalent to 3.1% of the total employed population in the EU Member States. Various recent European policy documents have also emphasised the role of the cultural sector in contributing to innovation in other sectors of the economy by providing content for digital devices and networks.

Major technology providers such as Google and Microsoft have also demonstrated the huge value embedded in cultural institutions by investing in ‘mass digitisation’ projects at a number of large libraries. Public-private-partnerships (PPP) with public content holders, such as libraries, museums and archives may involve granting the private partners exclusive access or other privileges in return for infrastructural investments (but the exact nature of the “quid pro quo” is often not public knowledge). This form of cooperation is often considered to have general economic value within the EU Member States and PPP is in general terms advocated by the European Commission and its expert group as one means of meeting the cost of large-scale digitisation.

However, it is possible that the protection intended by the Directive in Articles 10 and 11 relating to non-discrimination and prohibition of exclusive arrangements may raise sensitive issues with respect to some of these arrangements. The precise force of arguments for exceptional circumstances to allow exclusive deals as necessary in the public interest and the frequency of their deployment are issues

closely relevant to the outcomes of this study which relate to the inclusion of cultural information under the scope of the Directive.

The question therefore remains, given that cultural information that is for all practical, social and economic purposes the same as other forms of PSI, what if anything differentiates the one from the other. The study will explore and assess the force of each of these arguments in practice.

The work to be undertaken by the study is therefore to gather and analyse the evidence for and against the current position of the Directive *vis a vis* content held by cultural institutions and to assess what policy actions, if any, need to be taken to change it. The study will gather data and views, from both the cultural sector and from actual and potential re-users, which will indicate whether or not the time is right for additional steps to be taken.

3.1 Document review

During March 2008, the provisions of major reports relevant to this area of the study were reviewed, including recent documents such as:

- (i) Report on the 2nd Law and Regulation Thematic Meeting held 10-11 September 2007 (Paphos, Cyprus): “e-Exploitation of Cultural Heritage Information – a need for a European hand?”:
http://www.epsiplus.net/reports/epsiplus_thematic_meeting_reports/law_and_regulation_meeting_2_report;
- (ii) European Digital Library Initiative High Level Expert Group – Sub-group on Public Private Partnerships:
http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg_minutes/ppp/interim_report_ppp.pdf; and
- (iii) The economy of culture in Europe. Study by KEA for the European Commission:
http://ec.europa.eu/culture/eac/sources_info/studies/economy_en.html.

3.2 Questionnaire survey

A questionnaire was compiled, targeted to national librarians, national museums, national archivists and a selection of local and regional cultural institution directors. The questionnaire was capable of adaptation to different levels of detail and mode of delivery, depending upon whether it was administered as a written document or whether it will be used to support a telephone or face-to-face interview.

This has been e-mailed to a wide selection of cultural institutions using the mailing list of the EDLNet (80 participants) and EDLocal projects (30 participants), which cover cross-domain cultural institutions (such as libraries, museums, archives and audio-visual archives) at the national and local/regional levels respectively in all 27 EU Member States and Norway.

This took place during April and was followed up by telephone for the purposes of clarification where necessary in individual instances. The findings were analysed during the first half of May 2008, and a series of up to 20 telephone interviewees were identified with the potential to provide a more detailed level of understanding. Targets for telephone interviews were identified mainly in terms of analysis of the number of resources that are available for re-use and/or the extent to which the email questionnaire responses provide ‘headline’ information about current practice and policies.

These telephone interviews are balanced approximately between key representatives of national libraries (5), national museums (4), national archives (4), national cross-domain cultural agencies (2), large audio-visual archives (2), public libraries (1), local/regional museums (1) and local/regional archives (1).

Finally the study has identify eight interviewees in four major cities in different member states where face-to-face interviews will be sought with national librarians, archivists or museum directors (at least

two of each). It is expected that the interviewees will include the Director of the Conference of European National Librarians (and the Deutsche Bibliothek). These interviews will take place during the first half of July 2008.

During the period of the study the views of other European level associations are being sought including the European Bureau of Library, Information and Documentation Associations (EBLIDA), the International Federation of Library Associations (IFLA), based in The Hague and eIFL.net, based in Rome, primarily by email and telephone. .

A separate questionnaire has been compiled, targeted to demand-side organisations and their associations. Its use will be limited to telephone and if possible, face-to-face interviews. The reviewees will include key European associations, including the Federation of European Publishers (FEP) and national associations representing re-users of cultural information. This will be followed up by interviews with individual re-use companies identified through desk research, word of mouth (e.g. during the survey of cultural heritage institutions and or identified as case examples by their representative associations). The interviews described in this paragraph will take place during July 2008.

The outcomes of this sequence of research will enable the project to assess and present the consensus of opinion within the cultural sector regarding its readiness for inclusion within the provisions of the PSI Directive, to summarise the perceived strength of the core arguments for and against this and to relate these arguments to quantitative and case-based evidence relevant to current practice and potential, weighed against the views of re-users.

The questionnaires to be used for the various different stakeholders are set out in the Appendix.

4 *Voluntary sharing schemes*

The role of voluntary schemes for sharing intellectual property has an important role to play in increasing the availability of content that can be used, re-used and re-purposed. Mechanisms for voluntary sharing are developing rapidly with new licences and tools and therefore this approach will become a normal part of the copyright licensing toolkit.

We will be taking an inclusive view of voluntary sharing schemes, rather than trying to restrict our analysis to models meeting a limited range of criteria. The focus will be on the effect of the licences in opening up content for use and re-use in the same way that the public domain does. There are different degrees of “openness” in such schemes⁶, and we do not believe that the study should impose its own criteria here, but should instead provide an objective description and analysis that will enhance overall understanding and allow readers to make their own decisions as to how the benefits of each might offer benefits similar to those found in public domain works.

4.1 **Top-down analysis**

The top-down analysis based on desk research will result in two detailed matrices, analysing the features of the leading schemes, and their benefits and drawbacks. An outline structure for each is shown below (Figure 2 and

Figure 1). The analysis of features will aim to compare the main features of the different voluntary sharing schemes. It will identify the principles of the licences, the ways in which they can be used and how their definition and usage is governed. It will also compare constraints such as territoriality and applicability to certain types of content. The analysis of benefits will compare the uses, for which the schemes are suited or otherwise, and the strengths and weaknesses.

4.1.1 *SCALE*

⁶ See, for example, the discussion of such schemes at <http://www.opendefinition.org/licenses>

Although statistics on voluntary sharing schemes are imprecise, we will aim to gain such empirical information as we can through search engines that identify shared content. We will also explore other routes such as contacting the governing body (if any) for the licence to find out if it collects usage information. This will provide an indication of which of the voluntary sharing schemes are most successful.

Main voluntary sharing schemes to be analysed - matrix of benefits (Draft 1)						
Benefits analysis						
Ref	Scheme	Variant	Legal foundation	Strengths	Limitations	Most relevant uses Unsuitable uses Additional comments
1	Creative Commons	Attribution (by)				
2		Attribution Share Alike (by-sa)				
3		Attribution No Derivatives (by-nd)				
4		Attribution Non-commercial (by-nc)				
5		Attribution Non-commercial Share Alike (by-nc-sa)				
6		Attribution Non-commercial No Derivatives (by-nc-nd)				
7	Science Commons					
8	Creative Archive					
9	AEShareNet					
10	Publishing and Alternative Licensing Model of					
11	Open Source & GNU General Public Licence					
12	GNU Free Documentation License (GFDL)					
13	Open Art Network					
14	Unesco Open Training Platform					
15	Licence Art Libre					
16	Open Database License					
17	Open Data Commons Public Domain Dedication					
18	Open Data Commons Community Norms					
19	O'Reilly Open Books project					
20	Click-use licences					
21	Design Science License					

Figure 1 – Benefits and drawbacks of voluntary sharing schemes

Main voluntary sharing schemes to be analysed - matrix of features (Draft 1)																		
Ref	Scheme	Variant	Features/characteristics															
			Type of work or content covered	Require attribution	Allow redistribution	Allow commercial use	Allow adaptation	Compatibility with authenticated access	Compatibility with DRM or other technical restrictions	Territorial limitations	Revocability	Time limitations	Significant limitations or discriminations (users)	Significant limitations or discriminations (usage)	Meet "open knowledge definition" criteria	Governance/ governing body	Reference URL	Example uses
1	Creative Commons	Attribution (by)																
2		Attribution Share Alike (by-sa)																
3		Attribution No Derivatives (by-nd)																
4		Attribution Non-commercial (by-nc)																
5		Attribution Non-commercial Share Alike (by-nc-sa)																
6		Attribution Non-commercial No Derivatives (by-nc-nd)																
7	Science Commons																	
8	Creative Archive																	
9	AEShareNet																	
10	Publishing and Alternative Licensing Model of Africa																	
11	Open Source & GNU General Public Licence software																	
12	GNU Free Documentation License (GFDL)																	
13	Open Art Network																	
14	Unesco Open Training Platform																	
15	Licence Art Libre																	
16	Open Database License																	
17	Open Data Commons Public Domain Dedication and Licence (ODC PDDL)																	
18	Open Data Commons Community Norms																	
19	O'Reilly Open Books project																	
20	Click-use licences																	
21	Design Science License																	

Figure 2 - Features of voluntary sharing schemes

4.2 Case studies

During the study we will focus on case studies of situations in which different forms of voluntary sharing schemes have been adopted or considered. These will include such schemes as Creative Commons, Science Commons and Open Access. The case studies will aim to establish both the legal theory being used, and also how it is applied in practice.

4.3 Outlook for voluntary sharing models

The information from the top-down review and the case studies will be analysed to provide a general overview of the context in which the licences are used, the extent to which they are gaining traction and which features seem to give the strongest advantage, including potential economic impact.

4.3.1 ECONOMIC IMPACT FRAMEWORK

Although it is not part of the project to attempt to quantify the economic impact of voluntary sharing agreements, we will develop a framework that will identify those areas in which an impact might be seen and how it might be quantified. This will be a top-down framework rather than data-driven, as this would serve to point out areas where data could usefully be collected in future.

5 Data sources

We recognise that key to the success in relation to estimating the number and value of the public domain is the provision of suitable data to enable reasonable economic analysis to be carried out. Section 6 identifies an exemplar selection of possible sources of the data that will be required. This is not intended to be a complete and further work has already been undertaken to considerably extend the list.

However, of equal, if not more, importance is having a clear set of requirements which define the nature and type of data that the project will require. “Open-ended” requests for data are likely to result in nothing at all or far too much data which then requires extensive work before it can be valuable to the economic analysis. Therefore one of the first tasks undertaken was to develop a set of data requirements. All of the project team was involved in this process to some degree or another.

Each was asked to document the nature of data likely to be required and identify possible sources of this data within their specialist or territorial area. This was then consolidated into one document. From this list of requirements an “instruction manual” for all the researchers in the project was created. The purpose of the instruction manual was to guide the researchers when explaining to potential sources the nature of data that is required. This will go some way to ensuring that the data search is focused and effective. The Commission has assisted in this by providing suitable letters of introduction.

In addition to the above we have identify certain commercial data aggregation services and are seeking to obtain data from them. A small contingency fund was created should it be necessary to make payments to these commercial services. Negotiation with such commercial services is being undertaken by the senior members of the project team. Throughout the period during which data acquisition is at its most intense, a regular check of the nature and quality of data that has been secured is being undertaken. This is to act as an early warning system should we experience difficulties in successfully securing data.

6 Initial list of potential data sources

The detailed work needed to identify possible sources of data has considerably extended this list which is given here for example purposes only.

6.1 “Public” data sources

- Libraries should have bibliographic catalogues;
- <http://musicbrainz.org> has a large database of mainly CD metadata;
- Death dates could be obtained from Philip Harper’s <http://kingkong.demon.co.uk> site;
- Austrian organisation, RTR KommAustria which is the regulator for Austrian collecting societies;
- Archives: Swedish National Archive;
- National Libraries: French National Library; Latvian National Library; Finnish National Library; Helsinki Municipal Library;
- National museums: British Museum, Albertina Museum, Kunsthistorisches Museum; Estonian National Museum; Slovak National Museum;
- Performing arts organisations: Konzerthaus, Musikverein;
- The Raita database is operated by the Finnish national library and includes at least 5,000 sound recordings that should be in public domain and are now fully searchable and downloadable as mp3 files on the internet;
- Teosto should be able to provide information on the use of works of composers that have recently passed or are about to pass into the public domain; and
- There is a Finnish digitisation project that covers all local newspapers and the aim is to make them freely available on the Internet (<http://digi.lib.helsinki.fi/sanomalehti/secure/query.html>); due to problems in defining public domain they have only digitised material up to 1890 but the breath of the material is quite impressive.

6.2 Individual contacts

- Wim van Drimmelen (KB) or his nominee
- David Dawson (MLA)
- Gunnar Urtegaard (NALMA, Norway)
- Rosella Caffo (Italian Ministry of Culture)
- Elisabeth Niggemann (Deutsche Bibilothek)
- Marinos Ioannides (Cyprus Institute)
- Dan Matei (Romanian cultural institute)
- Ioannis Trohopoulos (Veria Public Library, Greece)
- Spanish Ministry of Culture

6.3 Potential commercial data sources

- Nielsen have point of sale data for both books and recordings;
- Publishers have information on book production;
- Amazon Associated Web Service provides web service access to their database of product information, also possibly including sales and prices information;
- Muze has a catalogue database for musical products including details on over 732,000 popular music and 134,000 classical music items; and
- Google obviously has its own database used for Google Books

6.4 Potential case study sources

- *Cinematographic and audio-visual works*: the roles deemed to be “authors” of these works varies from territory to territory within the Union. Therefore, subject to the affects of the Term Directive, some works of this nature may be in copyright in one territory and in the public domain in others. If such examples can be identified evidence relating to the relative value (as defined in Section 2.2) of such works in the requisite territories could be found to make suitable comparisons;
- *Non-original photographs*: the Term Directive left Member States to legislate on the term for such works, as they wished. It may be possible to identify examples of the same work

in copyright in one territory and in the public domain in another and measure respective values in each;

- *Critical and scientific publications:* Similar provisions as applied to non-original photographs in the Term Directive apply to critical and scientific publications. Therefore it may be possible to identify similar examples and carry out similar analyses;
- *Finnish National Broadcasting Company:* Their main radio channel has operated for years now a nightly "Night Radio" program. The concept is very simple: they play classical music continuously all of which is in public domain. They have estimates on the number of listeners and they have also made calculations on how much they have saved in rights owners' royalties compared to broadcasting music under copyright or related rights; and
- *Finnish re-publishers:* three Finnish companies re-publish sound recordings after 50 years of protection from the publication of the original. The rights of performers and record labels in the sound recordings have expired. However, because the rights of composers (unless it's classical music) are managed by a collecting society the end result is that re-publishers can publish any sound recording that was published more than 50 years ago with a reasonable "flat" payment to the collecting society. Provided suitable data can be obtained a comparison could be undertaken between the value of the sound recordings when protected and their exploitation in this manner once that protection period has expired.

Appendix

Questionnaire for Cultural Institutions

- 1 Were you aware of the European Directive on PSI re-use (2003/98/EC)?
- 2 Did you already know that cultural heritage institutions such as libraries, museums and archives were excluded from the scope of the Directive?
- 3 What would you see as the main costs of cultural heritage institutions being brought within the scope of the Directive (legal, processing etc)
- 4 What would you see as the main benefits of cultural heritage institutions being brought within the scope of the Directive?
- 5 What do you think are the main barriers to cultural heritage institutions being brought within the scope of the Directive?
- 6 Does your institution charge for:
 - a) access to any of the content which it holds
 - b) licences to re-use any of the content which it holds?
- 7 If the answer to either 3a) or 3b) is yes, how are prices for re-use of the content determined?
 - a) according to a published tariff available on the Web
 - b) according to a published tariff available on request
 - c) by negotiation with individual re-users
 - d) in some other way (please describe)
- 8 Why do you charge for re-use of content?
 - a) to support revenue/meet general costs
 - b) to fund new activities
 - c) to enable digitisation
 - d) another reason (please describe)?
- 9 How many re-use transactions did your institution agree during each of the last three financial years?
- 10 How many applications for re-use did your institution refuse in each of the last three years?
- 11 What was the gross amount of income your institution received from licensing or otherwise providing content for re-use in each year?
- 12 How many staff (full time equivalent) do you employ to support the preparation and licensing of content to re-users? Please estimate the total costs per annum of this activity?
- 13 What type(s) of licence do you issue to re-users (e.g. Creative Commons)? Are licences provided electronically?
- 14 Do you have any arrangements (e.g. public-private partnerships) whereby private sector suppliers digitise the content of your institution in return for the right to re-use it? Please describe the main provisions of any such arrangements as far as possible.
- 15 Describe briefly the types of content (format, subject etc.) which you usually license for re-use. Please attach any precise or detailed data which is available on this.
- 16 What type of technical and descriptive infrastructure do you maintain for making content available for re-use (e.g. repositories, databases, metadata) including your approach to long-term preservation?
- 17 What type of interface do you use to process re-user access/licensing of content (e.g. a page on main portal/website, special database, email address etc)?
- 18 How would you define the term 'public domain' for the purposes of your institution? What do you see as your institutional responsibilities in terms of the public domain?
- 19 What proportion of the content which you provide for re-users is in the public domain? Do you make any charges for this? If so what is your pricing basis/ tariff?
- 20 Would you consent to a more in-depth interview on this subject, if requested?

Questionnaire for Cultural Institutions' re-users and their representative bodies

QUESTIONNAIRE FOR REPRESENTATIVE BODIES

- 1 Were you aware of the European Directive on PSI re-use (2003/98/EC)?
- 2 Did you know that cultural heritage institutions such as libraries, museums and archives were excluded from the scope of the Directive, (alongside documents held by educational and research establishments and public service broadcasters)?
- 3 Do you hold or know of any data on the quantity or value of business or transactions involving re-use of information held by cultural heritage institutions conducted by your membership?
- 4 What are your general impressions of the way in which business involving re-use of content or information held by cultural institutions functions?
- 5 Are there any specific problems in the area of licensing or otherwise accessing content for re-use which you believe need to be addressed. If so, what are they?
- 6 Are there any specific cases exemplifying this which you can identify?
- 7 Can you direct the study towards member companies who are involved in this area so that we can contact them?
- 8 Finally, the provisions of the Directive on PSI re-use are to be reviewed in 2008. Do you feel that cultural heritage institutions should be brought within its scope? Please explain your reasons for or against.

QUESTIONS FOR INDIVIDUAL RE-USE COMPANIES

- 1 Were you aware of the European Directive on PSI re-use (2003/98/EC)?
- 2 Did you know that cultural heritage institutions such as libraries, museums and archives were excluded from the scope of the Directive, (alongside documents held by educational and research establishments and public service broadcasters)?
- 3 Can you provide any data on the quantity or value of business or transactions involving re-use of information held by cultural heritage institutions conducted by your company?
- 4 What is your general experience of the way in which business involving re-use of content or information held by cultural institutions functions?
- 5 Are there any specific problems in the area of licensing or otherwise accessing content for re-use which you believe need to be addressed. If so, what are they?
- 6 Are there any specific cases exemplifying this which you can identify?
- 7 Finally, the provisions of the Directive on PSI re-use are to be reviewed in 2008. Do you feel that cultural heritage institutions should be brought within its scope? Please explain your reasons for or against.