Access to digital collections of public domain works: Enclosure of the commons managed by libraries and museums

M. Dulong de Rosnay

Abstract

Digital copies of physical books and art objects curated by libraries and museums are being made available to the public online. Their access and reuse conditions are submitted to terms of use and policies defined by the institutions in charge of the development of databases and the digitisation of works, of which many are in the public domain. As no copyright applies to these works, their digital instantiation should in principle be freely accessible and reusable. However, in practice, some memory institutions databases contractual terms of use impose restrictions (for instance, reserving the commercial use of the version they digitalized or reserving the right to reuse photographies they funded), thus re-introducing physical barriers for resources which are supposed to be in the commons.

The article analyzes a sample of online databases policies of libraries and museums public domain collections in various countries, and provides examples of best practices of institutions and collective actions to avoid adding unnecessary restrictions to public domain works. A good governance of the digital commons including adequate partnerships will enhance access to knowledge.

Libraries, museums, public domain works, digitisation, enclosure, governance.

1. INTRODUCTION

Libraries and museums are digitizing their collections. Online databases of reproductions of manuscripts, books, paintings and other art works are being made available online though the websites of the libraries and museums or institutional portals. This research focuses on access and reuse conditions of public domain works archived in such databases. Public domain is to be understood in its strict definition as the state of works that are no longer covered by copyright, therefore authors can no longer exclude others from using their works. It corresponds to an open-access regime. Are out of the scope digitalized works which are still covered by copyright and which right holders have to be identified to negotiate authorization to perform the acts required by the digitization process and the making available of digital versions.

Creative works were in the public domain and free of use until the development of copyright legislation in the XVIIIth century. The barrier to access to these works was technological. Before the invention of the printing

press, it was long and difficult to reproduce works. Copyright has been developed to provide a limited monopoly and therefore an incentive to create and publish new works of authorship. With digital reproduction and online distribution, digital copies of works are considered as non-rival, non-exclusive goods. In economic terms, only the production of the first copy requires an incentive. Unlike to physical goods, their usage, being intensive or not, will not exhaust these public goods. The governance of digital commons does not have the same requirements than the governance of physical commons. But similarly, they need preservation and protection for the public interest. One opportunity offered by networked digital technologies is to provide access to knowledge to everyone who has access to a computer connected to the Internet. There is no risk of overusing digital commons, only a risk of underusing them, instead of facilitating access, comment, study, creative re-appropriation, transformation, research, education, online collaboration and unforeseen use cases.

Copyright opens a period of exclusive rights when authors and other creative actors may control the reproduction and the communication to the public of their works. Some acts remain outside of the realm of copyright. Their scope and nature vary among national legislations (fair use in the United States, exceptions for libraries, education, private copy, parody or other usages in civil law countries). So does copyright duration, which lasts up to 50 to 120 years after the death of the author or the publication of the work. When a work enters the public domain, it becomes part of some "common pool of resources" which can be freely reused by everybody for any purpose, without having to ask for permission or pay a remuneration, and without the rivalry and excludability characterizing physical common pool of resources. The public domain is beneficial to the public, including authors themselves who are finding inspiration and building upon previous works. But enclosures can restrict the public domain. Legal enclosures include copyright term extension, which will postpone the moment a work will reach the public domain, or the creation of a sui generis right for database producers to control the extraction and the reuse of substantial part of the database, in addition to the rights already pertaining to the elements of that database in case elements are copyrightable works.

This article describes some contractual and design enclosures which can be applied to digital collections of public domain works: the terms of use offered by libraries and museums, the technical characteristics of the platforms. Intuitively, digital representations of works which are in the public domain are also in the public domain. Indeed, the act of scanning is not an original creative process source of copyright. Besides, digitization is often funded by public money and it could be expected that publicly funded resources remain accessible to the public as taxpayer. However, public domain collections are most often not distributed to the public under a true public domain regime once they are part of a digital database.

Contractual and technical restrictions are enclosing digital commons and making it impossible or at least difficult or expensive to perform some usages of works which are in theory made possible by their public domain status. These enclosures prevent to enjoy fully public domain works, which are the

cultural heritage of societies. After describing these enclosures, the article suggests alternatives for a better governance of digital public domain collections which can be led by the memory institutions and the law-makers, but also by the community of users in the form of collaboration with the memory institutions to better preserve and make available the public domain resources.

2. DIGITAL ENCLOSURES

2.1 Terms of use

Terms of use or licensing conditions are non-negotiable, unilateral contracts written by a service provider and that bind users. Database curators usually publish them on their website in order to inform the public of what they are allowed to do with digital items of the collection, and sometimes who they should contact to negotiate further usages. Institutions are defining their own terms without negotiation with the public nor control by the state. This practice has already been observed as publishers are dictating their conditions to libraries purchasing subscriptions to academic journals. For memory institutions, the rationale can be the desire to further control the destination of an object they own, or the quest for additional revenues for further investments. There can even be legal arguments in some countries, or even the perception that the law may support copyright claims on public domain works photographies and rights for database producers.

A review of terms of use of public domain works on some websites of libraries and museums and institutional portals reveals that digital versions of public domain works are most often not made available in the public domain. Memory institutions are adding legal restrictions to the reuse of public domain works they curate and digitized.

Reserving the commercial use of digital versions is the most widespread restriction, leading to the commodification of the public domain. A comprehensive policy with clear terms of use reserving commercial use can be found on the website of the Brooklyn Museum. Gallica, the digital library of French National Library hosting mostly public domain works, offers another example of terms of use, this time quoting legal provisions on which they state to ground their policy of granting free non-commercial use while submitting commercial use to a license based a detailed fee table.

Beyond reserving commercial exploitation is the decision to reserve the publication of versions which can be exploited commercially even by non-commercial entities. The British National Portrait Gallery started a legal procedure against a user who had downloaded thousands of high-resolution versions from public domain works and uploaded them on Wikipedia. The rationale for the museum is that they generate revenues through the sale of such copies. The legal question is whether photographies of public domain works are covered by copyright: not in the US, but well in English law. The claim also includes the circumvention of technical protection measures and breach of contract. But the case has not been brought to a court which could have discussed the copyrightability of copies of public domain works. The

same issue applies to the upload of photographies of public domain works by individuals and their upload under Creative Commons licenses, as the licenses are adding legal restrictions to copies of works which are free of use.

A good policy to signal a work is in the public domain is the notification though a Public Domain Mark. Such a tool is made available by Creative Commons to tag public domain works and is going to be used by Europeana portal, a digital library aggregating metadata of European collections to signal to the public that no copyright restrictions apply.

2.2 Technical access conditions

Not only price and legal conditions, but also technology can create barriers to access, redistribution and reuse of works and data. But technical choices can also help remove barriers and facilitate reuse. Technical Open Access can ensure that works can be actually and effectively downloaded, reused, processed and searched by both humans and search engines. Technical barriers can include the following: protection measures that prevent copying or downloading, compulsory registration, design features that add hidden costs to search and processing. For example, it can be more or less easy to interact with a document because of the publication format. HTML pages are more convenient to browse a large amount of articles compared to PDF files which require download. Poor indexing or lack of metadata also prevent some modes of use.

Domain-specific requirements for technical open access can be defined. For software, it is the release of the source code. For music, it could be delivering an audio file in a free and open format together with information to play and remix it: MIDI file and other data such as music notation or explanation for performers. In the case of large scientific databases, technical openness is ensured by the possibility of downloading the whole dataset and reusing and integrating data without cumbersome registration. A checklist for databases curator to ensure their database technical accessibility could ask the three following questions: Do you provide a link to download the whole database? Is the dataset available in at least one standard format? Do you provide comments and annotations fields allowing users to understand the data?

For images, the availability of a high-resolution version will allow reuse in printed format or otherwise. For text, the availability of an OCR (optical character recognition) version will allow full-text search, which is not possible with the image of a text. Important technical features facilitating the search and the access to digital representations of literary and artistic works are the presence of metadata describing the work. Metadata should include rights metadata, licensing conditions should be made visible for each item, and not only on a terms of use webpage. Therefore, it will be easier to keep them attached to the work once it will be republished elsewhere and continue to display the work with relevant information on it. The Brooklyn museum publishes notices with metadata including a Creative Commons Attribution Non Commercial license as rights information.

Metadata by themselves, information about the work (author, year, museum,

size, etc), are a valuable resource, which can also be made available in the public domain. Data are usually not copyrightable, but databases can be protected. Bibliographic records datasets are informational commons, both as individual data and aggregated under the format of a catalogue or a database of metadatas. Metadata when released under public domain conditions will allow further processing and the development of other projects by the community. For instance, the British Library dedicates its metadata in RDF/XML format to the public domain with a Creative Commons tool, and in MARC standard for non-commercial purposes.

3. GOVERNANCE OF THE DIGITAL PUBLIC DOMAIN

3.1 Policy recommendations

Contractual and technical restrictions are implemented on purpose or by accident and lack of knowledge. In both cases, leading studies and providing educational material on the economic, social and cultural impact of such practices should facilitate informed decision on governance and the design of an adequate environment for databases of public domain works.

A public domain manifesto has been drafted within Communia, the European thematic network on the digital public domain, based on the input of researchers, activists and actors of memory institutions. Communia policy recommendations are stating that works which are in the public domain should remain in the public domain. No contractual or technical restriction should be introduced to limit access to and reuse of public domain works. Besides refraining to introduce digital enclosures, a more proactive role is proposed to cultural heritage institutions in the marking of public domain works, as part of the mission of preservation and making available of "shared knowledge and culture" to the society.

Finally, it could be recommended to lawmakers to draft provisions to prevent the enclosure of the public domain. Such a protection could be located in law on public sector information reuse, and to database legislation, to include an exception to databases of public domain works and databases of data, the metadata catalogues.

3.2 Partnerships

A reason to restrict access to public domain works is the belief that the monetization of reserved usages will bring additional funding to an expensive digitization process. This assumption is uncertain, as transaction costs may deter initiatives and be higher than royalties. This question deserves further research. Besides, the public domain generates value through its reuse for further creation, research, education, private consultation or commercial use. Memory institutions are receiving public funding for the preservation and distribution of their collections and need to sustain the expensive process of developing digital collections. Public-private partnerships can help institutions to save on these costs and improve the process.

Google Books project of digitization of libraries collections has been widely

commented, especially for the in-print and in-copyright part of the project but some restrictions are remaining even for the public domain component. Metadata are not available in an open format. It is possible to download a PDF of public domain books, but when looking for a public domain book title in the search interface, it is not easy to access to the public domain edition among the most recent editions which are not in the public domain.

Successful examples of collaboration with external actors are found in the Wikipedia GLAM (Galleries, Libraries, Archives and Museum) project. They reflect a renewed vision of public-private partnership and constitute a structured collective action in the participation of a user community to the conservation of the public domain.

Wikipedia relies on a large network of volunteers, or wikipedians, who write, edit, link, discuss and categorize articles. They are also contributing photos to Wikimedia Commons to illustrate Wikipedia articles or to be reused in external contexts. Photos uploaded on Wikimedia Commons must be either in the public domain or under a Creative Commons Attribution Share Alike license. Examples of collaboration between public institutions and Wikipedia chapters include the donation of images and books to Wikimedia Commons and Wikisource by archives or museums, and the making of photographies in museums for Wikimedia Commons or other archives. Further cross-fertilization practices could enrich the two communities based on respective expertise and skills are: the drafting of notices describing works, metadata enrichment by volunteers once they are made available under public domain conditions by GLAM institutions, or the restoration of damaged digitized version by skilled volunteers.

Besides sharing the mission to make knowledge accessible to the public and the exchange and acquisition of competences, both sides can find many advantages in such partnerships. An incentive for the institution is free labor, partially answering to the need of additional revenues leading to the idea to commercialize the public domain. A benefit is also extended exposure. For instance, the Rosetta Stone is viewed five times more on Wikipedia than on the British Museum online collection.
Incentives for wikipedians which have been provided in some of these projects are prices for best photography, participating to a contest or a fun event, meetings with the curators and learning from the expertise of the institutions staff members during residency.

4. CONCLUSION

Terms of use of digital collections of public domain works hosted by museums and libraries most often authorize their non-commercial usage for free, but license their commercial use for a fee. Such policies are creating new limitations on the public domain. These enclosures are based on arguable legal considerations and the desire to secure additional revenues. However, some institutions are beginning to mark public domain works with standardized metadata, encouraging the public to reuse the cultural heritage. Collaboration between memory institutions and user communities provides an

interesting option towards sustainability of digitization projects and governance of the public domain.

5. REFERENCES

Boyle, James, The Second Enclosure Movement and the Construction of the Public Domain. Law and Contemporary Problems, Vol. 66, pp. 33-74, Winter-Spring 2003.

Cohen, Noam,Venerable British Museum Enlists in the Wikipedia Revolution, *The New York Times*, June 4, 2010.
http://www.nytimes.com/2010/06/05/arts/design/05wiki.html?_r=1

Dulong de Rosnay, Melanie, "Opening Access in a Networked Science", in *Publius Project, Essays and conversations about constitutional moments on the Net collected by the Berkman Center*, June 2008.
http://publius.cc/2008/06/13/melanie-dulong-de-rosnay-opening-access-in-a-networked-science/

Dulong de Rosnay, Melanie, "Check Your Data Freedom: Defining a Taxonomy for Access and Reuse of Life Science Data", *Nature Precedings*, July 2008.

Guadamuz, Andres, "National Portrait Gallery copyright row", *Technollama*, July 19, 2009. http://www.technollama.co.uk/national-portrait-gallery-copyright-row

Hess Charlotte, Ostrom Elinor, "Ideas, Artifacts, and Facilities: Information as a Common-Pool Resource", *Law and Contemporary Problems* 66, p. 122.

British National Portrait Gallery copyright dispute Wikipedia notice
http://en.wikipedia.org/wiki/National_Portrait_Gallery_copyright_dispute

British Library bibliographic data policies
http://www.bl.uk/bibliographic/datafree.html

Brooklyn Museum Image Service
http://www.brooklynmuseum.org/image_services.php
Brooklyn Museum notice for the Head of a Ptolemaic Queen
http://www.brooklynmuseum.org/opencollection/objects/3799/Head_of_a_Ptolemaic_Queen

Communia public domain manifesto http://publicdomainmanifesto.org/
Communia project http://www.communia-project.eu/

Creative Commons Attribution Non Commercial license
http://creativecommons.org/licenses/by-nc/3.0/
Creative Commons Public Domain Mark 1.0
http://creativecommons.org/publicdomain/mark/1.0/
Creative Commons Public Domain Dedication
http://creativecommons.org/publicdomain/zero/1.0/
Creative Commons Attribution Share Alike license
http://creativecommons.org/licenses/by-sa/3.0/

Europeana and Creative Commons Launch New Public Domain Mark

http://version1.europeana.eu/web/guest/news/-/blogs/europeana-and-creative-commons-launch-new-public-domain-mark

Gallica terms of use http://gallica.bnf.fr/conditions?ArianeWireIndex=true

GLAM Wikipedia project notices http://en.wikipedia.org/wiki/Wikipedia:GLAM
http://meta.wikimedia.org/wiki/GLAM-WIKI_Recommendations