Digital Commons as a Model for Digital Sovereignty: The Case of Cultural Heritage
Lehmann, Jörg

DIGITAL COMMONS AS A MODEL FOR DIGITAL SOVEREIGNTY

THE CASE OF CULTURAL HERITAGE

Lehmann, Jörg
Berlin State Library
Berlin, Germany
joerg.lehmann@sbb.spk-berlin.de

ABSTRACT

This contribution looks at cultural heritage institutions and their digital assets from a commons perspective. Since the beginning of digitization in the late 1990s and with the change of the medium from the analogue to the digital, the role and mission of cultural heritage institutions has changed. Challenges for managing their assets in the sense of a commons arise, on the one hand, due to the current legislation on copyright and intellectual property rights, and, on the other, because of the availability of digital cultural heritage as Big Data, which opens up possibilities for economic exploitation of these assets by private companies. Should digital assets be available open access, or should access and use be regulated? This short paper discusses the possibilities for this model of sovereign data governance within the legal regimes of intellectual property rights and the public domain.

1 INTRODUCTION

A (digital) commons is a shared good or resource that is managed by a community for the benefit of its members, or, in a broader sense, is accessible for society or even for the global population. Digital cultural heritage can be understood as a commons, since digitization has mostly been funded by the public sector and because it is available open access via the internet. Such an understanding of digital cultural heritage as a commons is fostered by Article 27 of the Universal Declaration of Human Rights, which says that everyone has the right to freely participate in the cultural life of the community. The whole of cultural heritage institutions can therefore be conceptualized as the community governing and managing digital resources for the common welfare. Just like commons institutions, they neither pertain to the market nor to the state. The emphasis here is rather on sovereignty and self-organization, but the state—or in this case the European Union—provides the regulatory framework. The concept of the commons has been adapted in the 21st century to apply to the digital age, and the characteristics of a digital cultural commons have been carved out (Haux, 2021). A framework for the analysis and systematic comparison of commons institutions has been developed by Frischmann, Madison, & Strandburg (2014). Since the 1990s, millions of items of cultural heritage have been digitized. However, in many cases, cultural heritage institutions are not allowed to provide access to everything that is available in digital form; intellectual property rights prohibit this. Therefore, in a conventional understanding, digital cultural heritage refers to works which are in the public domain. Commercial use and the free re-use of digital assets which are not in the public domain are excluded by default. These restrictions therefore restrain the use of digital assets for the purpose of creating culture anew. Furthermore, the scale of digitization has turned cultural heritage into a commodity; such vast digital assets have a value as Big Data, since they can be used for machine learning applications, for machine translation, or the establishment of large language models. Now the question arises of who would benefit from this value: whether it serves the interests of private companies interested in optimizing their services and maximizing their profits, or whether it serves the common good. This issue has several twists: Because digital assets can be copied endlessly without the risk of the resource becoming exhausted, it is impossible to over-use the digital resource. However, there could be a potential loss of communal benefits due to actions motivated by self-interest (Yakowitz Bambauer, 2011): Private companies, for example, are not members of the commons, and the profits they might create out of the assets digitized mostly with taxpayers' money might not flow back into the commons. Such commercial use may preclude cultural heritage institutions from tapping into the potential of value creation and impair their digital sovereignty in managing the access to the digital assets as well as with regard to the maintenance of the commons.

The questions posed here have been discussed in a series of interviews with a range of cultural heritage practitioners and with law scholars, using the methodology developed by Frischmann et al. This short paper presents some of the key insights of a research project conducted at the Center for Advanced Internet Studies (CAIS, Bochum) during the winter term 2021/2022, the results of which were published, alongside the transcribed interviews, by Lehmann (2022).

2   DIGITAL SOVEREIGNTY WITHIN THE CURRENT LEGAL FRAMEWORK

Cultural heritage institutions fulfil an important task by selecting objects and collections from the vast pool of cultural products to preserve them and to provide access. These institutions therefore play a crucial role in defining what cultural heritage is and what of the totality of cultural products is going to be preserved. The value of cultural heritage is created by the expert knowledge and the procedures centered in the institutions that are responsible for identifying the cultural value of cultural goods—be it historical, artistic, scientific, architectural, archaeological, or otherwise—or that evaluate the meaning of a particular piece or collection for a specific community. Irrespective of whether physical objects or intangible cultural heritage are in the focus, the selection performed by cultural heritage institutions initiates a process of musealization and decontextualization (Lenski, 2013). The perspective shifts from the function of a good within its specific cultural context to the preservation of a tradition that is regarded as a form of cultural expression. Cultural heritage institutions perform the task of selection based on their expertise and by deploying the procedures their personnel—such as curators, archivists, librarians, conservators, or researchers—have learned in their specialist education. Galleries, libraries, archives, and museums (also called GLAM institutions) can often look back on a long pedigree and are endowed with high reputation and trust, which provides a certain quality assurance in the selection of the objects. By placing cultural products in these institutions, an ennoblement as cultural heritage is taking place. This function of valorization performed by cultural heritage institutions can be contrasted with the acquisitiveness of the big tech companies, which tend to collect each and every digital asset without any further differentiation and store them as Big Data in their data warehouses, and which do not have the means and procedures to address challenges central to the selection process and the forming of collections (Jo & Gebru, 2020). To a certain extent, interaction with the free market is evident, especially with the art market, which ranges from auction houses to the antiques trade; examples here include paintings and miniatures from art history, which are bought both by museums as well as by private collectors. The value of digital cultural heritage is indicated by its availability as Big Data. The vastness of the resources that are now available offers opportunities for economic exploitation for both private companies and

cultural heritage institutions, for example, by constructing large language models used for the improvement of their services, the attraction of more users and for increasing their revenues. The establishment of such large language models does not only raise the question of the ecological burden of computation and therefore of sustainability, but also of the consequences of the biases by which they are marked (Bender, Gebru, McMillan-Major & Shmitchell, 2021). In contrast to big tech companies, cultural heritage institutions have an excellent knowledge of the sources and domains from which the content they work with comes, they have metadata at hand which enable its careful curation, and they may therefore be able to provide high-quality products that consume less energy and serve their societies better than the models established by private companies would do (Lehmann, 2022).

The current legal framework in which cultural heritage institutions operate is mainly marked by the two rights regimes of public domain and intellectual property rights, both of which apply to the works under consideration here. The intellectual property rights regime has to be understood as a complement to the public domain part of cultural heritage, or, as James Boyle has put it, the public domain and the idea of the commons form the outside of intellectual property (Boyle, 2008). However, there are several transient zones between the two rights regimes, and consequently, there are several legal insecurities that arise out of the question how to deal with such material (like, for example, orphaned works, grey literature, leaflets and broadsheets etc.). But generally speaking, and in a conventional understanding, digital cultural heritage refers to works that are in the public domain.

In current public law, the purpose of cultural heritage is described as protecting and valorizing cultural traditions, but the aim of advancing cultural development is also cited (Lenski, 2013). This understanding of the function of cultural heritage was formed by the pre-digital age, where new cultural works were created through reception, be it through reading a book in a library or visiting a museum. The availability of cultural heritage in digital form, however, has changed the relationship between cultural heritage institutions and their users in multiple ways: Users have become accustomed to working with digital material that is in the public domain, but they are also interested in getting access to digital assets that have been produced in the past 70 years and may therefore be protected by intellectual property rights. This rights regime conflicts with the cultural practices established in communities in the creative sector who work with digital material and with their expectations regarding the open accessibility of such digital assets. Moreover, digital reproductions facilitate new modes of reuse that result out of qualities specific to the digital. Good examples here include the remixing of music, the animation of images or their conversion from 2D into 3D, and the creation of multi-modal books in electronic formats. Mission creep can be noted with regard to cultural heritage institutions: In comparison to the pre-digital age, their emphasis is no longer only

on preserving cultural heritage and providing access to it, but also on enabling the re-use of the cultural products which are available in digital format—with the aim of creating culture anew. A central challenge for cultural heritage institutions therefore consists in making available digital assets that are protected by intellectual property rights, with the purpose of stimulating the creation of culture. In so doing, cultural institutions would be managing their digital assets in a sovereign way.

Cultural heritage institutions have a range of possibilities and tools in this respect: They can negotiate with legators and the rights holders of the legacies to enable the re-use of material produced in the 20th century and of born-digital contents; this re-use may only be granted to registered users under certain conditions and in restricted spaces and not, as is current practice, to every possible user worldwide. Furthermore, cultural heritage institutions can engage with communities already working with digital assets; they can recognize their cultural practices as collective customs and traditions. An example of this has been given by the German UNESCO commission which has granted the status of intangible cultural heritage to the Demo scene, a subcultural computer art movement marked by comparably long traditions and customs (German Commission for UNESCO, 2021). Finally, cultural heritage institutions can invite such communities to create, curate, and pool their resources as digital cultural heritage and ask these communities to provide access to such digital resources, be it in the sense of open access for everyone or by enabling re-use of this material under certain conditions or with restricted access. In a certain way, the establishment of such relationships between cultural heritage institutions and the users of the digital assets they provide resembles a classical and historical conception of commons as closed spaces that contain resources to which only an elite has access. However, a trade-off has to be noted: While cultural heritage institutions can engage in sovereign management of their digital assets, they have to restrict access to them by excluding nonregistered users.

The second challenge for digital cultural heritage—the possible exploitation of digital cultural heritage as big data by private companies—can be addressed by using the possibilities given in the current legal framework. The European Directive on Copyright in the Digital Single Market (European Commission, 2019) has introduced two mandatory exceptions for cultural heritage institutions. The first exception allows institutions to make digital reproductions for the preservation of works that are permanently in their collections. Cultural heritage institutions can therefore digitize works that are still under copyright; however, they are not allowed to provide access to these digital assets. The second exception allows them to make use of their assets for the purpose of text and data mining; it thus enables the application of machine learning procedures and the development of artificial intelligence applications. Cultural heritage institutions can therefore develop such secondary products on their own, be they machine translation models or large language models created out of

167

the available massive textual databases. The European Directive on the Legal Protection of Databases (European Commission, 1996) provides the legal basis for cultural heritage institutions to license their contents and thus to regulate access for their users as well as for private companies. Moreover, and according to the Data Governance Act (European Commission, 2020, currently in its approval phase), it is possible for institutions providing data sets and models to demand fees and to realize profits; with respect to fees, it is even possible to differentiate between small and medium-sized businesses (SMBs) and larger companies, such as the big tech companies. While digitization is mostly state-funded, the significant maintenance costs associated with the management of the digital assets—such as technical equipment, electricity, and human resources costs—can therefore be covered, at least partly, through such fees.

Such a juxtaposition of cultural heritage institutions with big tech companies highlights the changing functions of cultural institutions in the 21st century—they have gone from preserving physical assets to establishing and administering outputs of machine learning along with providing quality assurance for these products, for example, by preparing "Model Cards for Model Reporting" (Mitchell et al., 2019). Moreover, the aim of cultural heritage institutions to develop secondary products on their own opens up the possibility of strengthening their bonds with registered users, for example, by including them in the establishment of machine learning procedures. Users may become engaged in crowdsourcing activities, such as annotating images or collectively putting captions on them, labelling data, or enriching metadata. Such approaches foster the traditional idea of the commons, where members of the community are obliged to fulfil specific duties, they strengthen social sustainability and the maintenance of resources, and thus contribute to the sovereign and self-organized management of the commons.

3 CONCLUSION

Both challenges for the sovereign management of digital assets identified here point in the same direction, which may be described as a movement from the open to the closed. It is generally possible to maintain the commons that enables digital sovereignty and self-organization of cultural heritage institutions within current legal regimes. But whereas digitization was begun before the turn of the millennium following the ideal of providing full open access to digital assets, the protection granted by copyright and intellectual property rights can only be lifted by the provision of closed spaces in which users can access the digital material under certain conditions. The above-described developments have consequences in the form of reviving a classical and historical conception of commons as closed spaces with resources to which only an elite has access. In these elitist communities, self-regulation, trust in the normative framework, and the importance of

obligations serve to maintain the resources, while non-members of the community must ask for access to licensed content and have to pay fees. The downside clearly lies in a compartmentalization of the internet and in a disbanding of the idea of open access for everyone.

4 ACKNOWLEDGMENTS

5 REFERENCES

1. Bender, E., Gebru, T., McMillan-Major, A. & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots. Can Language Models be Too Big? Conference on Fairness, Accountability, and Transparency (FAccT '21), March 3-10, Canada. ACM, New York, NY, USA, pp. 610 –623. https://doi.org/10.1145/3442188.3445922.

2. Boyle, J. (2008). The Public Domain. Enclosing the Commons of the Mind. New Haven / London: Yale University Press.

3. European Commission (1996). Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases. Brussels: European Commission. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31996L0009:EN:HTML.

4. European Commission (2019). Directive on Copyright in the Digital Single Market – Directive (EU) 2019/790. Brussels: European Commission. http://www.europarl.europa.eu/doceo/document/A-8-2018-0245-AM-271-271_EN.pdf.

5. European Commission (2020). Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act) – COM/2020/767. Brussels: European Commission. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0767&from=EN.

6. Frischmann, B. M., Madison, M. J., Strandburg, K. J. (2014). Governing Knowledge Commons. In B.M. Frischmann, M. J. Madison & K. J. Strandburg (Eds.), Governing Knowledge Commons. Oxford: Oxford University Press, pp. 1–43.

7. German Commission for UNESCO (2021). Bundesweites Verzeichnis Immaterielles Kulturerbe: Demoszene – Kultur der digitalen Echtzeit-Animationen. https://www.unesco.de/kultur-und-natur/immaterielles-kulturerbe/immaterielles-kulturerbe-deutschland/demoszene.

8. Haux, D. H. (2021). Die digitale Allmende. Zur Frage des nachhaltigen Umgangs mit Kultur im digitalen Lebensraum. Zürich/St. Gallen: Dike Verlag. https://doi.org/10.3256/978-3-03929-012-3.

9. Jo, E. S. & Gebru, T. (2020). Lessons from Archives: Strategies for Collecting Sociocultural Data in Machine Learning. FAT*'20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, pp. 306-316. https://doi.org/10.1145/3351095.3372829.

10. Lehmann, J. (2022). The Tragedy of the Cultural Commons. Research Report and Data Publication. https://doi.org/10.5281/zenodo.6513596.

11. Lenski, S.-C. (2013). Öffentliches Kulturrecht. Materielle und immaterielle Kulturwerke zwischen Schutz, Förderung und Wertschöpfung (= Ius Publicum, 220). Tübingen: Mohr Siebeck.

12. Mitchell, M. et al. (2019). Model Cards for Model Reporting. FAT*'19: Proceedings of the Conference on Fairness, Accountability, and Transparency, pp. 220–229. https://doi.org/10.1145/3287560.3287596.

13. Yakowitz Bambauer, J. R. (2011). Tragedy of the Data Commons. Harvard Journal of Law and Technology 25:1, pp. 1–67. https://doi.org/10.2139/ssrn.1789749.