

# The Tragedy of the Cultural Commons

## Research Report on the Project

Conducted at the Centre for Digitalization Research (CAIS, Bochum)  
within a Fellowship (October 2021 – March 2022)

May 2022

Dr. Jörg Lehmann, Eberhard Karls Universität Tübingen, [joerg.lehmann@uni-tuebingen.de](mailto:joerg.lehmann@uni-tuebingen.de)

DOI: [10.5281/zenodo.6513596](https://doi.org/10.5281/zenodo.6513596)

## Table of Contents

Executive Summary.....	2
Acknowledgments .....	2
Introduction .....	3
Methodology.....	5
The Cultural Commons Framework and Representative Research Questions.....	6
Data analysis .....	7
Findings and Results.....	8
Findings of the Systematic Comparison .....	8
The Background Environment .....	8
Resources .....	9
Governance .....	12
Patterns and Outcomes .....	17
Issue Analysis and Discussion.....	20
From Culture to Cultural Heritage to Digital Cultural Heritage.....	21
The Value of (Digital) Cultural Heritage .....	22
The Use of Cultural Heritage and Public Law.....	25
The Tragedy of the Digital Cultural Commons .....	27
The Future of Use: From Intellectual Property Rights to Remix Culture.....	31
Conclusions .....	34
Recommendations.....	36
Bibliography.....	38
Annex .....	41
Annex I: Code List for the Qualitative Analysis of Interviews Used in Atlas.ti .....	41

## Executive Summary

This research report provides core insights achieved during a research project conducted within a fellowship at the Center for Digitalization Research (Bochum, October 2021 – March 2022). The research project examined cultural heritage that has been digitized by European cultural heritage institutions as a commons, i.e. as a shared resource that is managed by a community. The first part of this research report systematically analyses the digital assets of a range of European cultural heritage institutions within a cultural commons framework. The results presented in this section are arranged with regard to the background environment of the commons, the characteristics of the resource and its governance, as well as typical patterns and outcomes, and they are mainly based on interviews with cultural heritage practitioners conducted during the fellowship. The second part of this report goes deeper into some of the issues arising with the availability of large amounts of digital cultural heritage, issues that have not been anticipated when digital transformation of cultural heritage began around the turn of the millennium, and which are related to the switch from the analogue to the digital. Firstly, one of the tasks of cultural heritage is to update cultural contents by creating culture anew. While this task is generally enabled and enhanced by the broad range of digital assets which are in the public domain and available in open access, intellectual property rights often form a barrier for the access to born-digital content and content which has been produced in the 20<sup>th</sup> century. This rights regime conflicts with the cultural practices established in communities in the creative sector working with digital material and with their expectations regarding the open accessibility of such digital assets. Secondly, the conversion of cultural heritage from analogue to digital has furthered its commodification. The availability of digital cultural heritage as Big Data Large pronounced its economic value and opened up previously unknown possibilities for exploitation. Large textual databases created out of digitized cultural heritage may serve as a foundation for machine learning applications and thus may feed machine translation services or support the establishment of large language models; or large image databases, enriched with crowd-sourced annotations, may support and enhance computer vision and a deepened understanding of such material by machines. From the viewpoint of a commons, the issue lies with the potential exploitation of these digital assets by private companies and thus in loss of communal benefits due to actions motivated by self-interest. This situation is termed the “Tragedy of the Cultural Commons”. Such a tragedy can be mitigated by cultural heritage institutions developing secondary products out of their digital assets on their own, like establishing machine learning models, providing parallel corpora for machine translation, establishing large language models for text generation etc. and by providing their users as well as companies with a licensed access to these models or products.

## Acknowledgments

The author would like to express his deep gratitude to the interviewees who dedicated their time and patience to participate in this research project, as well as to the participants in the workshop conducted during the fellowship, and finally to the fellows of his cohort at the Center for Digitalization Research (CAIS). Sincere thanks go to Assistant Professor Madelyn Rose Sanfilippo, who read a first draft of this research report and provided feedback.

## Introduction

As the 21<sup>st</sup> century advances, the profound impact of digitization on cultural heritage becomes more and more visible. Around the turn of the millennium, digitization was conceived of establishing a digital copy of an item belonging to cultural heritage as well as providing free access to it by everyone and from everywhere. With the unfolding of capabilities to process Big Data, with the use of artificial intelligence such as machine learning and pattern recognition, new possibilities for the exploitation of digitized or born-digital content have opened up. It becomes obvious that digitization has turned cultural heritage into a commodity which can be exploited by the private sector or for the sake of the society which has funded digitization. This observation draws the viewer's eye to the value of cultural heritage. The traditional conception regards cultural heritage items or collections as valuable with respect to the history and identity of a specific community or nation and judges their meaning alongside with such criteria as uniqueness, singularity, or scarcity of items. Cultural heritage institutions perform an important task by selecting specific items from the vast pool of cultural objects which may potentially become cultural heritage and incorporating them into the collections to be archived and preserved. Science supports this selection process by providing expertise on the context of such items, their meaning and materiality, or their provenance, thus keeping up the distinction between original and copy or original and imitation. Artificial intelligence approaches, by contrast, are not interested in the originality or the uniqueness of individual items. From this perspective, digital items are rather valuable as part of large corpora which can, for example be used for machine translation, machine learning applications, or the establishment of large language models.

While these diverging perspectives each in their own way underline the valorisation of cultural heritage, the commodification of digital cultural heritage highlights the characteristics of cultural heritage as a commons (German: "Allmende"), 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 as a whole. 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 (Kymlicka 2005: 213). However, since cultural heritage institutions play a crucial role in defining what cultural heritage is, in governing it and in being responsible for digitization or providing access to digital assets, their role and function within the digital economy of the 21<sup>st</sup> century is worth a closer look, an examination which will be performed in this study within a cultural commons framework. Such a perspective considers the legal and cultural context in which cultural heritage institutions operate as well as the characteristics of the resources which are created, obtained and to be maintained. Furthermore, the relationships amongst the members of the community and especially between cultural heritage institutions and their users have to be analysed, and governance mechanisms as well as patterns and outcomes need to be examined. With regard to risks associated with the commons, the role of the various participants is crucial: If users were to try to profit from the content without contributing at the same time to its preservation, the system could grind to a halt. The issue at stake is not a possible over-use of the resources, since digital assets are not at risk of becoming exhausted. Rather, the issue lies with the potential loss of communal benefits due to actions motivated by self-interest. Such a situation can be termed the "Tragedy of the Cultural

Commons". This notion forms the hypothetical horizon of the present study and aims at external users – especially private companies – which are not part of the cultural commons as such.

Even though cultural heritage institutions now serve a globalized community, the rules of access to digitized cultural heritage are still defined by national laws and by the individual cultural heritage institutions themselves. The traditional conception of cultural heritage still keeps the distinction between original and copy; within the digital sphere, however, this functional link is removed: The digital object itself can be copied endlessly without any loss, and rules and laws regulating access to the digital object may therefore be different from those cases where originality is bound to the materiality of the cultural object. Furthermore, cultural practices have changed with regard to digital assets. Access to cultural heritage was traditionally conceived in terms of a physical visit to a cultural heritage institution, where such a visit or access may stimulate the intellectual or artistic exchange with the physical object, and it may even inspire the creation of new cultural objects and thus enable use and re-use of cultural heritage. In the sphere of the digital, by contrast, access to the asset is often equated with the availability of this asset for download, use and re-use, such as it is practiced in music sampling, re-mixing and re-arrangements. Such cultural practices, which have emerged alongside with the availability of digital assets, often challenge existing legal frames, such as copyrights and intellectual property rights. An understanding of cultural heritage as a commons therefore consequently leads to an examination of where the conventional legal frame is seen to be still sufficient, and to an identification of the necessity to adapt this legal frame in order to make it fit for purpose in the 21<sup>st</sup> century.

Digital cultural heritage poses a challenge for the traditional conception of cultural heritage as something to be presented (the side of the sender), whereas the users are thought to be in the role of consumers (the side of reception). Digital transformation profoundly changes the relationship between cultural heritage institutions and their users, be it those who actually visit the institution and/or those who access the digital collections via the internet. We will soon have to think of the relationship between users and digital cultural heritage as a dense entanglement of actors and content or a socio-technical assemblage, where users contribute annotations to digital material which is then used for machine learning, or where users interact with models resulting from machine learning, and where users create new content out of such interactions which might then be again fed into the databases, alongside with content which has been made freely available by contemporary users. Such considerations illustrate the point that we have to reflect broadly on the impact of technology onto the cultural heritage sector and even on the cultural sector as such.

While the changes in the relationship between cultural heritage institutions (in their role as members of the community managing a resource) and their massively extended audience forms a first focal point of this examination of digital cultural heritage as a commons, a second one is formed by a bifurcation that has silently begun when cultural heritage institutions embarked on the digital transformation. The public sector has spent billions of Euros on the digitization of our cultural heritage. Examples of such large-scale digitization efforts are [Europeana](#) with more than 50 million digitized objects (books, music, fine arts, etc.), the digital library [Gallica](#) of the Bibliothèque nationale de France with more than 4,7 million digitized print products, or the [Zentrale Verzeichnis Deutschsprachiger Drucke](#) with more than 1,8 million digitized titles. The establishment of data is a costly endeavour, and

the provision of these data in open access can on the one hand be seen as a proof of the wealth of the societies which fund digitization: A wealth in terms of the culture being presented, of the long histories linked with it, and of the manifold inspirations which can be drawn from it. On the other hand – and this alternative has only recently been opened up – the availability of large volumes of data enables capabilities for exploitation which can be used by only a few already powerful institutions and companies. Big data approaches can turn societal and cultural wealth into profit and generate even more power, since only a minority of institutions dispose of the ingredients necessary to exploit such data. However, the cultural legitimacy of such approaches is challenged. Rather, cultural heritage is understood as a public good, which should be managed and governed by the communities which created this resource.

A commons approach was therefore seen as the best fit firstly to examine the current situation with regard to digital cultural heritage as a shared resource, and secondly analyse and discuss the issues and challenges arising from the digital transformation of cultural heritage within the current European legal framework.

## Methodology

Looking at cultural heritage institutions from the perspective of a commons framework provides the chance to systematically study individual cases by means of a structured approach, to inventory the structural similarities and differences of the various cultural heritage institutions and the services they provide, and thus to enable comparisons between these institutions as well as generalizable results. To enable such a systematic comparison of commons was at the heart of the Institutional Analysis and Development (IAD) approach developed by Elinor Ostrom (1990). While Ostrom and her collaborators have focused on commons arrangements in the natural environment, Frischmann et al (2014) have taken up the systematic IAD approach and transferred and adapted it to study knowledge commons. The knowledge commons framework was explicitly developed as a structured set of research questions to be applied to different cases within a cultural environment, where the central characteristic of those cases is the institutionalized establishment and sharing of resources amongst the community members. The knowledge commons framework seemed to be particularly suitable as an approach to analyse digital cultural heritage as well, since knowledge products can be understood as public goods that are marked by traits which can be seen as analogous to digital cultural heritage: “They are naturally nonrivalrous or nondepletable, meaning that consumption of the good does not deplete the amount available to other users, and nonexcludable, meaning that knowledge outputs are not naturally defined by boundaries that permit cheap exclusion of users” (Frischmann et al 2014: 6). Furthermore, responses to the narrative of the “tragedy of the commons” often narrow the set of possible solutions down to either government intervention and regulation, or to the exclusion of possible free-riders via a specific rights and licences regime. A framework approach, however, promises a more in-depth understanding of a commons, a thick description of the resource characteristics, and the provision of a solid basis for describing already existing solutions and making predictions on further developments.

The knowledge commons framework was explicitly designed to be adapted to other forms of commons as well, and this adaptation has been performed for digitized cultural heritage as

the cultural commons which is at the centre of this study. Consequently, the cultural commons framework asks for the legal and cultural background environment of the commons, it identifies and describes the basic attributes of the commons including its resources, the community members and the goals and objectives of the commons and its members, it takes governance structures into perspective and asks for patterns and outcomes (see “The Cultural Commons Framework and Representative Research Questions” below). The cultural commons framework serves two purposes: First, it guided the planning of the interviews with cultural heritage practitioners as well as the interviews with law scholars, all of which were conducted during the fellowship. Second, the framework was used for systematically analysing, comparing, organizing, and presenting the information gained from the interviews as well as from research literature on the topic.

## The Cultural Commons Framework and Representative Research Questions

### Background Environment

- What is the background context (legal, cultural, mission etc.) of this particular commons?
- What is the “default” status of the resources involved in the commons (public domain, copyrighted, intellectual property right, open, or other)? Is there some cultural heritage that is excluded from digitization by default?

### Attributes

#### *Resources*

- What resources are pooled and how are they created or obtained? (Are there nonobvious examples of resources, like research output or audio material?)
- What are the characteristics of the resources? Are they rival or nonrival, tangible or intangible? Is there a shared infrastructure?
- What technologies and skills are needed to create, obtain, maintain, and use the resources? Which search functionalities are offered to the users?

#### *Community Members*

- Who are the community members (of the infrastructure) and what are their roles? Who assumes the role of an aggregator?
- Are there distinctions in the degree and nature of openness with respect to each type of community member and the general public?

#### *Goals and Objectives*

- What are the goals and objectives of the commons and its members, including obstacles or dilemmas to be overcome?
- What are the history and narrative of the commons? (See also "mission"-question)

### Governance

- What are the relevant use cases and how do they relate to the goals and objective of the commons and the relationships among various types of participants and with the general public / the users?

- What are the governance mechanisms (e.g., membership rules, resource contribution or extraction standards and requirements, conflict resolution mechanisms, sanctions for rule violation)?
- Who are the decision makers and how are they selected?
- What are the institutions and technological infrastructures that structure and govern decision making?
- What informal norms govern the commons?
- How do non-members interact with the commons?
- What legal structures (e.g., intellectual property, contract, licensing) apply?

#### **Patterns and Outcomes**

- What benefits are delivered to members and to outsiders/users (e.g., dissemination to a broader audience, and social interactions that emerge from the commons)? What kind of innovations and creative output do you expect in the (near) future?
- What costs and risks are associated with the commons, including any negative externalities (e.g., legal proceedings regarding intellectual property rights / enforcing compliance with the rules)?

The main deviations and adaptations of this cultural commons framework from its predecessor as presented by Frischmann et al. (2014) can be seen in the 'background' section. with regard to the different legal context (public domain vs. intellectual property rights) to be applied here, as well as an explicit question about the mission of the individual cultural heritage institution and the addition of a question whether there is cultural heritage that is excluded from digitization by default. In the 'attributes' section, the qualification of the resources involved, questions about the existing infrastructure and the search functionalities provided as well as the exploration of the role of aggregators within the community are specific to the cultural commons framework. In the 'governance' section, the wording 'action arena' was replaced by the more applicable 'use cases' to investigate the interaction between institutions and users of digital cultural heritage. The section 'patterns and outcomes' was specified with regard to outside users, understood as non-members of the commons, and with questions exploring the (possible) use of digital resources by the institutions and their users, to focus on the state of the art and the creative re-use of digital cultural heritage; costs and risks associated with the commons was specified with questions regarding legal proceedings and the use of restrictive measures to regulate the use of the digital resources.

## **Data analysis**

The transcribed and anonymised interviews were coded and analysed with the help of the social science software Atlas.ti. A preliminary set of codes was created by referring to the list of interview questions. Each code within the code list was provided with a content description, its properties or brief examples for reference. These codes were applied deductively to the interview material, yet complemented by inductively emerging in-vivo codes. Ultimately, the code list comprised of altogether almost fifty codes (see Annex 1) spanning about 350 citations.



Codes were provided in such a way as to enable multiple interpretations, taking into account different layers and perspectives of analysis, keeping each other in check. Wherever possible, quite generally applicable codes were refined and specified. For example, the code “access” was used only where a general description of an access to the resources was given, with no further specification. The more detailed codes “access: complete”, “access: licences”, and “access: threshold” were applied where unrestricted access, the use of licences for access or the existence of thresholds to limit access was the case. The applied codes were established both following the key words of the cultural commons framework as well as topics and concepts that arose during the interviews. Ultimately, thickly coded interview transcripts were obtained. Memos written in Atlas.ti were very helpful for their analysis as they identified some of the overarching topics that might structure the presentation of the findings.

## Findings and Results

The following presentation of findings and results is mainly based on the interviews conducted during the fellowship. The structure of the following sections is organized according to the cultural commons framework, which allows a systematic comparison of the information gained from the interviews.

### Findings of the Systematic Comparison

#### The Background Environment

A commons, in the traditional meaning of the word, consists of natural resources to be shared and managed within a natural environment. In the case of a cultural commons, the description of the background environment is not that easy. The question about the legal background was answered quite easily by the interviewees by pointing to the currently dominant legal regimes of public domain and intellectual property rights, the latter of which may refer for example to images where personality rights apply. These two regimes can be identified as the two canonical background environments for cultural commons. In the former environment the digitized works are in the public domain or labelled with creative commons licenses mostly with the intention to specify a non-commercial use of these assets. Metadata most often have to be counted as part of this background environment, since it has become customary in Europe to label them as CC-0. Public domain, creative commons licences as well as CC-0 also mark the default status of the digital resources available on the websites and portals of the cultural heritage institutions. It is notable that the clear provision of information about what can be done with the digital assets (and what cannot be done) forms a distinctive feature of the services provided by cultural heritage institutions, especially in comparison to services like Google Image Search, which provides access to an abundance of images, but mostly lacks reliable information on the possible re-use of these images.

Correspondingly, in the latter environment, where intellectual property rights apply, the available material often forms a much smaller part of the digital assets, except for libraries which provide licenced material for their users. Cultural heritage institutions such as national libraries have found a couple of ways to provide access to digitized copyrighted material to



their users, for example by providing specific computer terminals on the premises of the institution where users can access the digital content without being able to download or copy it, or which enable to do the latter only in the case where the appropriate licences apply. Another possibility to provide access to such content consists in establishing agreements with individual copyright right holders or their organizations: “There are two areas where we have spent much time on both understanding and finding the solutions. On one hand literature books and on the other hand newspapers, because newspapers are so interesting to people. In both cases we have made quite a few agreements with copyright holders or publishers to be able to do things that we by default cannot do because of limitations by copyright or the old-fashioned legislations on libraries.” (I1) It seems advisable for cultural heritage institutions capable to perform the digitization process by themselves to propose collaborations with copyright holding institutions, since such collaborations may result in win-win situations: “And when it comes to for example newspapers, we often experience that they really want us to publish. We have relatively good agreements to do so. We have done some practical collaborations to do so like digitizing the newspaper and letting the newspaper collaborate also in the funding of all the work to do things faster in the given amount of time etc. Our willingness to do so has been reflected in good agreements, seen from the users' point of view.” (I1) It is obvious that dealing with material that may be protected by copyright is a cumbersome task for cultural heritage institutions, which requires a lot of time and effort. As one interviewee explained, a meticulous inquiry into rights is “the only way to really increase the amount of video material, for example, audio-visual material. The main blocker, at the moment, is clearly the conversation about rights. And then the problem is even more specific when you start looking at the time of the content, so we know that for everything related to the 20<sup>th</sup> century type of materials, we know that we have very few of those because this is usually blocked by copyrights applied to those materials”. (I10) It is not surprising that cultural heritage institutions are vexed by the complex legal situation around cultural heritage produced in the 20<sup>th</sup> century, and a broad range of literature providing guidance and advice for cultural heritage institutions in this matter bears witness of the uncertainties arising from difficult legal questions and the procedures established to apply the correct licences and publish the digital material (see here for example Klimpel 2019, Klimpel 2020a, Klimpel 2020b, Klimpel 2021; the best current overview of the intellectual property framework has been provided by the inDICES Project 2021).

The question whether or not some cultural heritage is excluded from digitization by default has mostly been answered with a clear “No”, which implied that material which is in the public domain will be digitized if the necessary funds are available. One of the rare examples where the legal regime does not reflect current social norms was given with objects from colonial contexts, of which one interviewee said that they were digitized but not shown for ethical reasons, as for example in the case of human remains.

## Resources

Generally speaking, resources are created by cultural heritage institutions themselves. If funds are limited and not everything which is in the public domain can be digitized, cultural heritage institutions usually perform a selection and take a decision of what should be digitized by prioritizing perishable material such as newspapers. The imperative of

conservation thus provides a guiding principle. Digitization may also include establishing electronic descriptions of the material available in the institution, such as in the case of Encoded Archival Descriptions (EAD) established in archives, which may later be enriched with links to digitized material. While individual cultural heritage institutions conduct the digitization process themselves or order companies to do it, it is often the case that the digital assets stay with the providing institution, where they are made available via a dedicated website or portal, while the metadata are passed on to a regional or national aggregator which may again pass it to another top-level aggregator.

In many contexts, multiple types of media are shared within a cultural commons. A characteristic of the commons is therefore that often a broad range of types of media are made available in electronic format, such as books, newspapers, photo collections, newsfeeds, documentaries, moving images, audio files, or files and records from archives, some of which may also house born-digital content like governmental publications or open data provided by the administration. With libraries, born-digital content like academic studies or linked open data are common. The technologies and skills necessary to establish the resources are often available in-house; beyond the know-how specific to the digitization process, these skills and capabilities comprise the establishment of websites and portals, data science skills, as well as the legal knowledge necessary to equip the material with the appropriate licences.

The question of who the community members are is decisive in understanding the specific character of cultural heritage as a commons. Here it is necessary to differentiate between those institutions which are publicly funded (like federal archives or the [Deutsche Digitale Bibliothek](#)) and those which are privately funded (often through foundations). These community members cooperate with each other by sharing digital assets, an exchange for which usually contracts, agreements, or Memorandums of Understanding between each two partners exist. Cultural heritage institutions might also share infrastructures or services, such as is the case with libraries jointly using discovery services. Further members of the communities may be formed by private companies sponsoring the digitization process and partnering with the cultural heritage institutions. Also, for these kinds of cooperation contracts are set up defining what is being digitized, which partner receives what kind of product and in which quality, and what might be done with processed data like OCR outputs. Beyond institutions and private companies, a part of the users can be considered as community members, namely those with a library or user card or those buying a reader's ticket. These users directly contribute to the maintenance of the commons, often by paying a fee, and they have to be distinguished from non-registered users or outsiders, whose role will be treated in more detail below.

The roles of the members within the community can be described according to their function within the sharing of the resources: There are cultural heritage institutions doing the actual digitization; they share their digital assets with aggregators (like the Deutsche Digitale Bibliothek) which usually cooperate with a broad range of cultural heritage institutions. Finally, there are aggregators (like Europeana) who nearly exclusively cooperate with other aggregators. Even though a direction of the data flows from cultural heritage institutions to top-level aggregators can be noted, the structure of the community can best be described as a non-hierarchical network of peer institutions. Within this network, shared infrastructures exist, within which the digital assets are open to each community member without any

restrictions, since the community members are seen to be on a par with each other. In the cases where there is no gradation between various types of users (i.e., registered users vs. outsiders), the general public is granted the same access to the resources as the community members, which speaks for a high level of openness of the resources which are undifferentiated by further degrees of access.

Asked about the goals and objectives of the commons, most interviewees provided as a standard answer that these consist in preserving the cultural heritage and providing access to it. As the main obstacles and challenges to overcome mostly technical issues were named, such as the broad variety of data formats which have to be managed, the volume of the data which has to be maintained and kept up to date, as well as improving the data quality, as the following interview citation shows in an exemplary way: "it is really about making sure that the data that are sent are fulfilling all those new requirements that are emerging due to different type of usage, different other emerging technologies, we need language tags associated with the data so that you can support the multilingual interface, we need more links in the data, because now we want to move away from having very textual data to more connecting data so to really have entities represented in the data, so structured knowledge as part of the data. So this is a big effort and it will continue to be because it is always an ongoing activity, but the idea is really to link the data quality efforts to all this new digital usage that is emerging." (I10)

Interestingly, the standard answer provided by the interviewees regarding the goals and objectives of the commons ("preservation and access") was complemented when they were asked about the innovations and creative output they expect for the near future. The answers provided here complement the traditional tasks of cultural heritage institutions with the enablement of re-use of the digitized material, mostly for creative and educational purposes. Several statements made by the interviewees support this observation, beginning from virtually presenting the digital assets: "we also have for example the virtual exhibition tool that we offer other partners from our network to create virtual exhibitions, and there of course our partner institutions try to create knowledge from the content we offer, and this is an approach we want to expand in the future." (I6) Other form of use and re-use include assembling digital material and integrating them into new products: "we have been launching several projects for re-use, like conversion to ePub format for digital collections, we also have a crowdsourcing project based on our digital collection, and we want to build a platform with different applications for them to create their own research environment, to enable the production of research publications on that platform." (I8) The emphasis on re-use completes the narrative of cultural heritage institutions and can be seen as a new goal, as can be read from the statement of another interviewee: "I think now the goals are really much more – it's less turned on the inside, you know, what we as an organisation can do together, but more like showing how the platform can be used, so now there was a switch then to really proving that what we were doing was relevant to any citizens, you know, so that the users, any citizens, they are using the digital content to do many different things, that we can prove that things are being used" (I10). The consequences of this new goal will be discussed further below.

## Governance

A digital commons can be defined as a (virtual) space as a place of social and cultural practices and the context of social interactions, and thus as a relation between the participants and the digital culture as well as the relationships existing between the participants (see Haux 2021, pp. 152–154). It is therefore imperative to describe the rules-in-use between the several members of the community and their use of digital cultural heritage. In our case, the relationships mainly in two areas have to be taken into perspective: Between cultural heritage institutions and the various kinds of users they serve; and the relationships between cultural heritage institutions themselves. While the imperative on relationships is a characteristic of all cultural commons, including digital commons, there are peculiarities of digital commons which will be investigated later in this research report, such as the re-usability of digital resources as well as its specific value as big data.

Within the community of European cultural heritage institutions which form the object of this study, a differentiation of users can be observed: There are on the one hand registered users (or users with a library card or a reader's ticket or the like), and on the other hand users which can be understood as the general public, who access the resources without any registration or log-in process and who may be described as outsiders to the commons. Finally, as a further group private companies can be identified. With respect to these different user groups, a broad range of rules-in-use are applied by cultural heritage institutions providing access to digital assets: One extreme is formed by the position that the public domain part of the content is openly available for everybody including private companies: "Everything that is in public domain is completely open, everyone can access it. We only have these restrictions for those documents that are not in public domain, and these are only accessible on our premises [...] We are always talking about material which is in the public domain. If that is the condition, everything is available for free, even for commercial use. So the only condition is to declare the source, that it comes from the national library, but that's it." (18) Open accessibility of the content even for big tech companies like Google is seen as an advantage: "Most of our visitors come today because they find something in Google or in our search engines. More than accessing the main website and looking in our search methods, they mainly, I don't have the exact figure, but I think the majority come directly to a work that they have found out on Google, but we are okay with that. We are not looking for profit. We are not a company, so we just want to increase our impact. So, we are not concerned by this type of usage, this type of access to our library." (13) Quite obviously this viewpoint does not see any danger of exploitation of these assets by a big tech company; rather, the cultural practice of using Google is seen as a welcome means to facilitate and enhance access to the resources provided.

The other extreme is formed by cultural heritage institutions which grant access to their digital assets only to registered users. One may think of a library here which provides library cards to their readers and asks them to provide log-in credentials whenever digital assets need to be accessed. In between these two extremes – open access for everyone including private companies vs. access only for registered users – a broad variety of differentiation can be found. A scientific library, for example, found a solution for the people not belonging to the university to which the library is attached: "The problem is that we are not focused only to the university members, which means students and the employees, but also to people who are interested in scientific literature as well and in a big city like ours this is an important part

of our work. More or less half of users in the library are not from the institution itself. So, it is easy for real, material books or things like this which you can store and which you can loan. But it is a big problem for all our electronic stuff. To have access to this electronic service is very important and in our licenses with the vendors we have special conditions for walk-in users, which means that users which come into the library have access to this electronic stuff, and we need technical facilities to provide this access and this is what we do, we have some kiosk-systems, and if you log in into the librarian system with your library account and from this onwards you have access to everything that we have bought electronically, and you can print it or take it as a copy to your dropbox or something like this." Another example is formed by the digital library Gallica of the Bibliothèque nationale de France, which provides free access to all digitized items and allows their download after having manually ticked a box. The content itself is secured using deep web technologies, thus protecting it from automatic access, for example by search machines like Google and private companies seeking to crawl the content. Other cultural heritage institutions differentiate between the content provided by them and the one they received from their partners: "there is a big difference between our content and external content, because external content was provided to us, or we participated in the creation of content under certain agreements, so as I said we have hundreds of agreements with different institutions and we are into the process of licensing them because we want to move in a more open... but it won't always be possible." (I3) In other cases, research communities may have received privileged access to digital assets, for example to large corpora, and for the purpose of creating products for education.

Governance mechanisms between the various cultural heritage institutions can be described via the exchanges taking place alongside the establishment of contracts, agreements, and Memorandums of Understandings. Digitizing cultural heritage institutions are consulted on the use of licences and public domain marks, they receive advice and encouragement: "So based on this that's why we provide a lot of recommendations and support to show them that – even though in their context they may have done specific choices, if they want to start sharing their digital assets, if they want to reach more users etc., maybe they should review their decisions, and it is about making them aware that basically the decision they are making on those aspects will influence their digital strategy and also how they are going to be looked at on the internet, how people are going to find their stuff, what they will be able to do with the content" (I10). Governance is exerted if cultural heritage institutions are seen not to provide the appropriate licenses: "so usually we have a lot of discussions happening at that stage where we receive the data from a partner, and then we might notice things in the data, and we try also to have these discussions to see if we can influence the partner in maybe sometimes changing things for more open licences, and also to make sure that their content can be better discovered and used" (I10). There are various explanations for the reluctance to provide open licences on the part of cultural heritage institutions which are delivering data. One is judged as an outworn understanding of ownership: "I could say the most challenging difficulty is that very traditional approach to cultural heritage content. So many institutions feel that they own their content, so that the content should not leave their institution. So they see it as a danger that people can download the content if they make it public at a different place. So they are still reluctant very often, and that includes not only public libraries which is indeed a problem, but also the publishers. Some publishers have old content, and they do not have plans to publish it again. However, they are reluctant to open the content or to allow us to digitize the content and to make a promotion or to promote or to

disseminate the content in internet. In some cases it was really amazing because we asked them 'Do you have plans to republish this content?' And they said 'No, but it is our content, and we don't want to leave it open.'" (I3) Another reason which was given is the lack of knowledge about the appropriate application of rights statements: "My impression is that they are struggling with – and these are especially smaller institutions – that they are not well informed. From my point of view there is really a big lack of information with respect to all these things. Maybe rights, maybe not so much about digitisation but about what we can make public, what we can't. How we treat for example objects from the NS era? And there is as far as I can see a very big lack of information, especially in smaller institutions, and then of course they are not so open for open data in a way and they tend not to give us their objects. This seems to be from my point of view the main problem at the moment." (I6) One of the means to put a soft pressure onto digitizing cultural heritage institutions are the provision of quality control mechanisms by aggregators: "When we were working with partners and publishing their data on our portal, we would usually give feedback on the quality, so we would use our publishing framework there to identify issues with the content, we would say okay, this content is entirely in this tier, or that tier [...] So that is the kind of feedback we give there. And I think one of the main motivating things for cultural heritage institutions to increase their quality according to our framework is when they're working on projects and the projects require a certain tier of the framework to be met. We work on European projects. The calls for these projects, say for example if you want to propose a digitization project then you have to make sure that the objects that you're digitizing with that funding are openly licensed. So it works a little bit to force them, basically, that in order to do digitization with this money, they have to adhere to the higher tiers of our publishing framework. So that definitely motivates them ..." (I10). Cultural heritage institutions providing advice to other institutions are well aware that the governance they exert has its limits, and that the arising difficulties are unavoidable corollaries of the digital transformation: "At the end of the day, the institutions are deciding what kind of rights they apply to their data, the role of [this specific cultural heritage institution] is really to provide some awareness on what the best practices are and should be. One of the missions of [this specific cultural heritage institution] is to really support institutions in really moving into the digital age, to really facilitate this transition." (I10)

As the structure of the community has been described above to be more of the sort of a decentralized network than a hierarchy with centralized decision-making bodies, institutions which could be regarded as decision makers are missing in this commons. Instead of this, advice, encouragement and soft pressure is exerted as described above. The taking of decisions is supported in a rather indirect fashion, e.g., by supporting the implementation of standards and protocols which form the technical complement to the contracts and agreements between the cultural heritage institutions. Even though being standards themselves, these tools may also present challenges where they are not sufficiently defined or open to interpretation: "Just quite the existence of conceptual standards for archival descriptions and also standards for the transmission of archival descriptions and the encoding of it in digital forms. There is a tremendous amount of variability in how people actually do this. There is also pretty much no use of common controlled vocabularies, to the extent that would make data integration easier. So, it's really a case when we actually receive material, the standards aren't really sufficient to enable it to be integrated without additional efforts. So, for example, in the case where institutions can actually produce Encoded Archival

Descriptions, there is a lot of different ways in which the standard can be interpreted & that makes it difficult to have two EADs from different places in the same database and expect it to work the same way. There is a lot of variation in how archives describe their material in terms of the number of different hierarchies they use.” (I15) Another example where decision making processes are supported is via the standardization of rights statements: “The work [our institution] did was really to try to standardize to a maximum the rights statements that are applied to the digital assets to try to create some cohesive landscape in Europe, so that’s why we had different initiatives which resulted in specifying a set of rights statements which we use at [our institution] which are the creative commons licences, so partners / institutions sending data to [our institution] can use those creative commons licences, or specific licences that were developed inside a consortium which is the [rightsstatements.org](http://rightsstatements.org) consortium.” (I10)

The technological infrastructures of this commons can be characterized with brief words. Usually, aggregators of digitized cultural heritage provide storage capacities, application programming interfaces (APIs) as well as standard protocols (such as the OAI-PMH Open Archives Protocol for Metadata Harvesting), tools for the exchange of metadata and digital objects, and tools for the validation of the data quality. These offers may be accompanied by tools for the establishment and management of controlled vocabularies and concordances, tools for data conversion as well as tools for data collections which may be used by smaller cultural heritage institutions which may not have the capacity to host a full productive workflow in-house. While this provision of an infrastructure is top-down, the cultural heritage institutions themselves provide a bottom-up delivery of data to the aggregator. Usually at least metadata are provided, often accompanied by thumbnail versions of the digitized content, and, depending on the licences which are applied, sometimes also the digital object itself is delivered. Often the core content stays in the cultural heritage institution itself; links in the metadata point to the website or portal of these institutions or directly lead to the digital object itself.

As has already become clear from the description of governance mechanisms between the diverse cultural heritage institutions, the ideal of openness and the enabling of open access to the resources can be seen as the informal norms which govern the commons. The general idea is that cultural heritage should be accessible for everybody worldwide without any restriction, and digitization is the vehicle to provide unrestricted access. However, while such a conception refers to cultural heritage which is in the public domain and which therefore forms the core part of the commons, at the same time it creates frictions with regard to cultural heritage to which intellectual property rights apply, since the latter rights regime has to be understood as the complement to the public domain part of cultural heritage as well as to the several transient zones between the two and the legal insecurities such spaces create (like, for example, orphaned works, grey literature, broadsheets etc.). The informal norm of a push towards openness resurfaced in the statements provided, endowed with the radiance of an understanding of the internet enabling free access for everyone, a conception that one of the interviewees, looking back at the early 2000’s, described with the following words: “It was really one thing that had happened at that time. We call it ‘the internet’. So the main consequence of the internet already at that time was that the user of the internet expected to find everything there. If it wasn’t on internet, it didn’t exist. And I believe that is even more true today. And the long sequel consequence of that observation was that there is only one thing to do: that is to digitize the entire collection and if possible, to make it available on the



internet.” (I1) While the internet is being seen as the medium enabling access, the ideal of openness is rooted even deeper: “if you look at it from a more social and political perspective, it is of course about the democratic right to access to information and to support everybody’s opportunity and ability to participate in community.” (I1) The informal norm of open access and resources published as public domain can also be read from the following statement, which includes soft criticism and disapproval of the cautiousness practiced by cultural heritage institutions: “But often as we can see for example in the case of newspapers, institutions are little bit shy to say they are public domain, because maybe there are still living authors that could have rights on these objects, so they often use a more restrictive license. Honestly speaking it makes no sense [laughs] because they are public domain, they are published, and they should be public domain. But often they have a license, and this is more because of the fear of cultural heritage institutions to get problems later.” (I6) The informal norm of digitized cultural heritage as a public good accessible to everyone as well as the tension this ideal creates with respect to the legal framework of intellectual property rights describes the current situation of the commons very well.

The differentiation of the various types of users discussed above has brought the fact to the forefront that a major part of those users (which may be described as the “general public”) have to be considered as outsiders or non-members to the commons. Other than registered users holding e.g., a library card and paying fees, these users do not directly contribute to the commons – the number of visits to a portal might form an argument for funding and can thus be seen as an indirect contribution – and they are also not involved in defining the rules-in-use. The main interaction of these non-members with the commons thus consists in the consumptive use of the available websites, portals and APIs. Further interactions may consist in asking for specific items, requesting corrections in the available metadata or demanding the de-publication of specific datasets: “At our institution, what is happening is that sometimes we receive what we call ‘take-down requests’ where we are being asked to remove, to de-publish some data, so sometimes de-publishing a specific item, but sometimes it is also about de-publishing specific datasets. This could be because the rights applied to the content is not the right one and there were some claims made, so this is why an institution is asking us to de-publish while they are addressing the issue.” (I10) From the perspective of a commons, the involvement of these non-members seems to be the exception rather than the norm. The potential that lies in the integration of such non-members was described by one of the interviewees who explained how their users were involved in enriching an image database with semantic information: “we have seen amazing examples of contributions like the first photo collection that I mentioned where people sit down and go through thousands of images and do description of every image of boats for example because that’s their interest and they provide a lot of information back which we are not able to produce ourselves. They have scale, they have knowledge that we cannot have ourselves.” (I1) From a commons standpoint, these outsiders or non-members therefore rather seem to be an untapped resource; especially with regard to the preparation of large datasets for machine learning, where e.g. annotations for texts or captions and descriptions for images are needed, crowdsourcing activities would form an obvious example for the engagement of these users.

The legal structures applied to the commons have largely been described above, be it in the section on the background or with respect to the contracts and agreements between the cultural heritage institutions. However, it is worth mentioning that the many thorny issues

with which cultural heritage institutions have to struggle with respect to the appropriate applications of licences and rights statements also leave room for interpretation, and that the specific qualities of the digital sphere enables much more flexible approaches compared to cultural heritage in its physical format: “The legislation is placed on things and objects rather than the digital and when it comes to, let’s say – grey literature and also literature published by the government and there are quite a few examples, we are trying to ... rather to be close to the more dangerous edge... if we can call it that ... on the behalf of the citizen rather than protecting the rights for the potential right holder. So, if we are uncertain, we tend to look at things from the citizen’s point of view, from the democracy point of view rather than the pessimistic point and decide for the side of the potential copyright holder.” (I1) However, such an attitude demands a certain courage as well as an equanimity with respect to legal proceedings possibly arising out of the digital publication of a specific object.

### **Patterns and Outcomes**

The interviewees had no difficulties in describing the benefits delivered to the commons members and others. Broadly speaking, three groups benefitting from the commons were addressed: The users in general, the partnering cultural heritage institutions, and research and private companies. The benefits for the users of the websites and portals were described in general terms; one interviewee used the beautiful phrase “content is richness for society” (I3) to identify the public welfare effect of providing access to digitized cultural heritage. The interviewees generally were proud about the high quality of the services they provide to the public, such as delivering high-quality human-made indexing and cataloguing data facilitating precise information retrieval, providing a central access point for a broad range of highly diverse cultural heritage objects, or providing high quality digitisations of cultural heritage objects as opposed to low-resolution mass digitisation. Offering curated collections pulled together from several cultural heritage institutions or conducting community-specific activities – for example for educators – were seen as further benefits for the users in general.

The benefits for partnering institutions within the network were described as getting visibility via the presence of their digital objects on the portals of major aggregators or as part of their virtual exhibitions, and as getting the opportunity to offer the collections of a specific institution within a wider context. Providing an overview of unique items which are geographically dispersed but related to the same topic is definitely an advantage characteristic of digital access; as one interviewee put it: “it is useful to know what exists in what archives, and on top of that we do aim to interconnect material between archives as well” (I15). A further benefit for partnering institutions consists in the network effect of exchanging information on the appropriate application of rights statements and licences and the provision of openly available digitized items, which opens up possibilities for funding, thus strengthens cultural heritage institutions and diversifies their range of activities.

Enabling knowledge creation and research has always been a task of cultural heritage institutions, and the availability of large corpora of data has opened up the opportunity to foster the exchange with the scientific field as well as establishing new collaborations. One interviewee was “mainly thinking of the research community which often would benefit from the data if they have effective ways of exploiting them” (I3), whereas another interviewee “wanted to stress the intersection between the digital humanities and also AI research with

the library and information science field" (I2). Most of the interviewees were well aware of the opportunities large data corpora provide to form new alliances, as the following example shows: "Most of the research done in the National Supercomputing Centre is funded by a ministry which is interested in this kind of projects, and they are providing resources to institutions like the National Supercomputing Centre, and they can work with us just by signing an agreement between the institutions, with no contract or money from our side involved. They are going to have resources from the ministry, and we are providing the data" (I8). Moreover, doing research in-house was also seen as a field of activity, be it via the development of a knowledge graph which would enhance the services of the digital library by enabling a semantic approach to the various digital contents, or to develop large data corpora of high quality based on the huge amount of data available and the excellent knowledge of the available content in the institution: "I think that one of the major benefits of having for example a national library with a huge digital collection is that we have everything and we have also relatively [good] metadata on our collection, meaning that we can as the type of institution that we are, we can really aim the content of our collection at a certain type of machine learning, for example. And we can, probably we can compensate for some other dangers around – for example bias. Because we understand the content very well and we have a good control of our collection." (I1) Developing and providing machine learning models does not only aim at the research community, but also at private companies, which may for example use the models to establish systems assisting in human-machine communication. Even beyond these more practical application, the use of digitized cultural heritage content by private companies has been taken into perspective, as one of the interviewees stated: "We'd like to support also the cultural industry and in a sort of cycle process we'd like that users (or companies) help us develop different products based on our collections, so what we would like to offer is the access to our data and digital collections for the society to use them in the way they need or the way they want." (I8) From a commons perspective, the latter statements open up the question of how the commons would themselves benefit from the profits generated out of the large data corpora they provide.

With respect to the costs associated with the establishment of the commons, it has to be noted that digitization is mostly state-funded, while maintenance costs – technical equipment, storage capacities and human resources – have to be provided often by the cultural heritage institutions themselves. As has already been described above, significant human resources are needed to advise cultural heritage institutions on rights and licenses questions, an expertise ensuring the open availability of the digital objects. Moreover, it should not be forgotten that especially within the archival landscape many cultural heritage institutions lack the expertise to provide interoperable data, as one interviewee remarked: "I suppose if we had a huge amount of money we would try and introduce some specialist into a few key institutions. They could work in a concerted effort inside those institutions in order to do a lot of data clean up and integration work and then make that available to harvest" (I15).

It is not easy to realistically estimate the risks associated with digitized cultural heritage as a commons. Quite obviously cultural heritage institutions are cautious with regard to legal proceedings arising out wrongly labelled digital assets. This cautiousness might minimize the risk associated with digital cultural heritage; on the other hand, all of the interviewees had difficulties in remembering whether one of the institutions had to face legal proceedings. One of the interviewees answered in an exemplary way: "To be honest, no! As far as I know

[chuckles] we never had a problem with this. If we have problems, or if we have complaints, it is more about personal rights. [...] If people do not want that their year of birth is published on our website and they complain, we delete this of course. But with respect to intellectual property rights, as far as I know we never had any problems.” (I6) Therefore it can be concluded that this risk is seen to be low. The same can be said of the risk of a de-anonymization of data; an interviewee working with archival material which had been created until the mid-20<sup>th</sup> century judged the situation like this: “But, yes, I think there is a risk with the data we have, but the benefits outweigh the negatives to putting it up. As I said, in the beginning of the project there was a considerable concern that privacy-related issues based on de-anonymised data would be a problem, and for this reason we need to develop a suitable approach where people can access the material, but in practice it has not really been an issue mostly because most of the data we have is designed for dissemination anyway. To a large extent it is available elsewhere if you know where to look.” (I15) It can again be noted that the risk of de-anonymisation might be effectively mitigated by the cautiousness which the cultural heritage institutions apply before publishing the digital objects.

The digital sphere provides a broad range of evaluative criteria with which the patterns and outcomes can be measured. The obvious choice here are usage statistics of the portals, websites and APIs, the number of page views and of new users; also the tracking of activities on related social media sites. Moreover, surveys on specific features, collections, or the design can be placed on the portals to collect feedback. Beyond these data available to cultural heritage institutions, currently the Horizon2020 project [inDICEs](#) is being conducted, which is tasked with measuring the impact of the digital transformation onto both the cultural heritage sector and the creative sector. The project aims to develop suitable measures and will thus guide cultural heritage institutions in understanding the impact of their digitized content on the cultural and creative industries.

The cultural commons framework can be recapitulated in a flowchart. Initially, this flowchart visualized the Institutional Analysis and Development (IAD) approach developed by Elinor Ostrom (1990). It was then transferred and adapted to study knowledge commons by Frischmann et al (2014: 19).

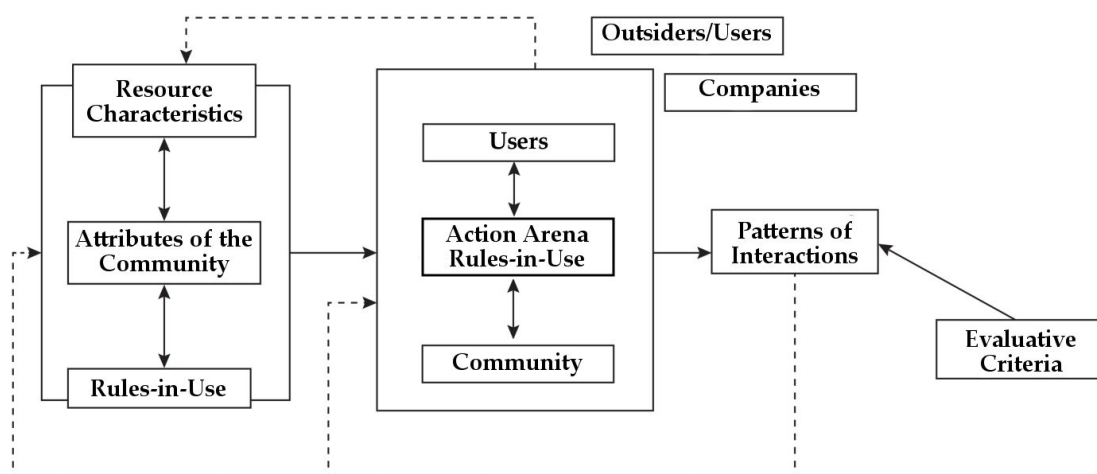


Figure 1: The Cultural Commons Flowchart

The main changes made to the original flowchart of the knowledge commons framework pertain to the 'Action Arena' in the centre; over here we have to acknowledge that a lot of interaction takes place between the community of cultural heritage institutions and their users, interactions which can best be described by the rules-in-use. Users who are placed in the central box can be understood as persons visiting a cultural heritage institution, users with a library card or registered users. Distinct from these users are those users and companies which were placed outside of the central box. This deviation from the original framework reflects the fact that digitization implies as a consequence an involvement of "outside users" and the possible use they might make of the digital assets provided by cultural heritage institutions. The existence of outside users which are not part of the commons again points to the possibility of a "Tragedy of the Cultural Commons", because these outsiders may create profits out of the assets digitized mostly with taxpayers' money, but these profits may not flow back into the commons.

## Issue Analysis and Discussion

After the interviews with cultural heritage institutions had been conducted and the findings had been systematized, four further interviews were held with German law scholars to achieve a better understanding of the legal framework in which cultural heritage institutions operate, and to chart the options which for example the several European directives related to the Digital Single Market offer for cultural heritage institutions. These interviews were conducted in German to keep the precision of the legal terminology; therefore, these interviews are not cited directly in this report, but only paraphrased; however, a reference to the interview transcript is provided in case that readers able to understand German want to delve deeper into this material. Moreover, an online workshop was conducted in which cultural heritage practitioners as well as the law scholars Professor Brett M. Frischmann (Charles Widger Endowed University Professor in Law, Business and Economics, Villanova University, Pennsylvania, USA), Michael J. Madison (Professor of Law and Faculty Director of the Future Law Project at the University of Pittsburgh School of Law, Pittsburgh, Pennsylvania, USA), and Assistant Professor Madelyn Rose Sanfilippo (School of Information Sciences, University of Illinois, Illinois, USA) participated. The author of this study is very grateful for the valuable insights and contributions provided by the participants to the workshop. In the following part of the study, some of the ideas discussed within the workshop may be taken up without providing a reference, but the author acknowledges the valuable contributions provided by several participants and expresses his deep gratitude.

The following part mainly discusses issues that arise from the relationship and interconnectedness between the state of the cultural commons as described above and the legal framework in which cultural heritage institutions operate, as well as the several uncertainties arising from them. Further topics arise from the challenges and unanticipated consequences brought about by digitization, such as related to the value of digital cultural heritage, its commodification, and the resulting change in the relationship between cultural heritage institutions and the users of the objects digitized by them.

## From Culture to Cultural Heritage to Digital Cultural Heritage

How does an artifact become a part of cultural heritage? It is hard to delineate the term “culture” and to distinguish it from other areas of life like, for example, “education” and “science”. Culture comprises a vast pool of material and intangible resources, and it is the process of musealization that takes cultural objects out of their original cultural context and declares them to be a part of cultural heritage. The process of “musealization” comprises of the selection, the acquisition, conservation, exploration, exposition and exhibition of these cultural objects. Cultural heritage institutions perform this process, and within it process the cultural objects receive a new social function. They lose their anchoring in everyday social life, in their primary social context, and also their original symbolic meaning (Lenski 2013: 4–6). Instead, these objects are re-coded as carriers of memory and meaning. We can say that cultural objects that were not appointed as cultural heritage from the outset become detached from their social context within this process of musealisation. Digitization can be understood as one further step of the process of decontextualization, but it also has further consequences: Artifacts that have been decontextualised can be transferred and received digitally and globally with their newly encoded content. It is the advancing of global connectedness that makes a global cultural exchange possible, and the internet is the paradigm for this exchange, comprising of the ideal of open accessibility for everyone worldwide.

It is not surprising that the processes of musealization and digitization resemble the “loss of aura” described by Walter Benjamin in his famous article “The work of art in the age of mechanical reproducibility” (Benjamin 1977: 136–169). Benjamin describes here how a work of art loses its aura by being taken out of its original context: “Even the most perfect reproduction of a work of art is lacking in one element: its presence in time and space, its unique existence at the place where it happens to be. [...] The whole sphere of authenticity is outside technical – and, of course, not only technical – reproducibility. [...] The authenticity of a thing is the essence of all that is transmissible from its beginning, ranging from its substantive duration to its testimony to the history which it has experienced. Since the historical testimony rests on the authenticity, the former, too, is jeopardized by reproduction when substantive duration ceases to matter. And what is really jeopardized when the historical testimony is affected is the authority of the object. / One might subsume the eliminated element in the term ‘aura’ and go on to say: that which withers in the age of mechanical reproduction is the aura of the work of art.”

Mechanical reproducibility implies for Benjamin a division into an original and its copy. According to Benjamin, the original still keeps qualities like authenticity and authority, whereas the production of a copy destroys these qualities, since the aura of a work of art rests on the uniqueness of artifacts in the context of its tradition. The very same can be said of the process of digitization, which at first glance resembles mechanical reproduction. But digitization has further consequences: It lowers the barriers for access, and even though the process of scanning or photographing does not, in itself, produce usable data, in the last consequence digitization comprises of a decisive implication: Digital reproducibility and availability is ultimately supplemented by the ability to change the digital asset at any time. While digitization often implies some further processing steps (like extracting the text from a scanned image with OCR), born-digital cultural objects already level the difference between original and copy. Whereas in the age of technical reproducibility reproduction was always a

copy, even this functional link to the original has been removed in the age of digital reproducibility. Alongside with accessibility and the possibility to download the digital assets, the aspiration and demand to use these materials in various way arises, a usage that goes beyond consumption because it aims at working with the digitized objects, at embedding them into new contexts and environments, and at creating new content on their basis.

Digitization thus has fundamentally changed the rules of the community of cultural heritage institutions: They are now serving a globalized community, but the rules of access are still defined by national laws and by the individual cultural heritage institutions themselves. And digitization also changes core functions of cultural heritage, some of which consist in the formation of identities or the presentation of historical relations and traditions, while another the function consists in updating cultural contents by creating culture anew.

### **The Value of (Digital) Cultural Heritage**

The first step in the process of musealization is the selection of the cultural heritage. Even though the uniqueness or singularity of a cultural object are amongst the selection criteria, this step can be understood as an artificial creation of scarcity with regard to the entirety of culture, a scarcity which also marks the value of cultural heritage. This value can be estimated from a range of different perspectives.

The process of selection is guided by the expert knowledge and the procedures housed by cultural heritage institutions, which are responsible for the identification of the historical, artistic, scientific, architectural, archaeological, or other cultural value of cultural goods (Lenski 2013: 35), or which evaluate the meaning of a particular piece or collection for a specific community. This selection is indifferent to the fact whether it is a question of cultural goods – this term is related to the materiality of an object – or whether intangible cultural heritage concerned. The latter term has been defined in the UNESCO Convention for the Safeguarding of the Intangible Cultural Heritage and designates forms of expression of culture not consisting in physical objects, meaning “practices, representations, expressions, knowledge, skills” which are “constantly recreated by communities and groups” (see UNESCO 2003a). With regard to the selection of intangible cultural heritage, the very same process of musealisation is initiated as with physical objects: If an intangible good loses its everydayness, the perspective shifts from the function of an everyday habit to the preservation of a tradition which is regarded as a form of cultural expression (see Lenski 2013: 48). In a similar vein, the Charter for the Preservation of the Digital Heritage (UNESCO 2003b) moves from a classic understanding that captures material aspects of cultural heritage onwards to a comprehensive consideration of immaterial goods that are identified to be part of cultural heritage. In this way, the Charter puts the immaterial digital content on the same level as the immaterial idea that was customarily put down in a physical object (Haux 2021: 164). However, cultural goods, intangible cultural heritage, and digital heritage do not form three distinct or clearly defined areas; rather, they are characterized by overlaps. An archivist receiving the legacy of a famous writer comprising of, amongst others, several hard disc drives with drafts of novels as well as other physical media with electronic content may classify such born-digital content just as any other physical cultural object, since it is mixed with other objects (papers, correspondence, photos) and is not part of traditions and customs practiced by communities and groups. The German Commission for UNESCO, on



the other hand, has put the demoscene – a subcultural computer art movement – onto their list of intangible cultural heritage, thus reflecting the long-standing tradition of a decentralized non-commercial community which has developed social customs and presents their cultural products on celebrations and festivals (German Commission for UNESCO 2021). As these examples show, there is definitely the need for a clear specification of the terms “intangible cultural heritage” and “digital cultural heritage” to better understand the process of selection and valorisation resulting out of appointing cultural goods as cultural heritage.

Cultural heritage institutions perform the task of selection on the basis of the expertise gathered in their house, often verifiable in the scientific education of personnel such as researchers, curators, archivists, librarians, or conservators. This expertise certainly supports the evaluation process; however, galleries, libraries, archives, and museums (or GLAM institutions) can often look back onto a long pedigree and have, as one of the interviewees remarked, their reputation as an asset, alongside with international appreciation as an institution. With their reputation, they stand in they for a certain quality assurance in the selection of the objects, and this notable combination of competences and qualities is something which, by contrast, big tech companies cannot provide (I12). As another interviewee remarked, the creation of value and the ennoblement of objects as cultural heritage is performed by institutions, most of which are not democratically legitimised (I13). These considerations illustrate that cultural heritage institutions perform their mission on the basis of the trust that people place in them, a trust which expresses people’s belief that these renowned and time-honoured institutions are taking the right decisions and making the right choices.

The value of cultural heritage is partly reflected by the free market, especially the art market, which ranges from auction houses to antiques trade. At this point, the inseparability of art and culture supports the commodification of culture in the most obvious way. Products of fine arts and of craftwork, items which have survived for centuries, or which have been collected by the houses of rulers, religious institutions or historically important persons provide typical examples of the overlap between the monetary valorisation objectified via the art market and the classification as cultural heritage provided by GLAM institutions. Even though the legacy of a single person may be hard to sell on the art market as a closed collection, there are many examples where single archival items – such as the autograph of a book which has become famous when it was printed later – have realised high prices on the art market. Furthermore, the interaction between cultural heritage institutions and the free market is observable in the provision of expertise from the side of cultural heritage institutions, for example with regard to the relation of individual items to specific archaeological sites, or with regard to the provenance of works of art: In this case, the provision of external evaluations from a disinterested third party supports pricing in the free market. However, the market only partly reflects the value of cultural heritage, since the mission and vision of cultural heritage institutions and therefore the interests in which they are acting may be incommensurable with the interests driving markets: Preservation, the provision of access and the tasks of public responsibility are by far removed from the logic of markets.

In a paradoxical way, the value of cultural heritage is also determined by the use that is being made of it. Use is conventionally understood as use in the form of access (Lenski 2013:

262/263). We may therefore think of individual items that are drawing the masses to an exhibition site and thus create a significant financial value which can be objectified in the form of entrance fees, such as is the case with the Mona Lisa in Paris or the Nofretete in Berlin. Though this value is certainly concrete, the paradox around it is that such objects are not marketable. Formulated the other way round: The better known an item of cultural heritage is, the less likely it is that it will be valued on a market, simply because usually museums don't sell their holdings. Rather, in such cases the value of cultural heritage is determined by an economy of attention. Attention is a scarce resource, and the digital sphere provides the means to measure and objectify it. The analysis of usage statistics, which is performed by all the cultural heritage institutions interviewed within this research project, thus serves a double purpose: It provides a legitimation for funding, as has been clearly expressed in the statement of one interviewee: "in the future we'll be more aware of what the impact is of what we do and how we can really say like 'that is really important' and we could justify our existence in that way" (I10). Secondly, usage statistics prove the appreciation allocated to each individual digital object within the economy of attention and thus conflates the valuation of cultural heritage with the creation of value (Lenski 2013: 458). While the value of cultural heritage is determined here by the use made of it by accessing it, the availability of digital cultural heritage provides a further, yet largely unexplored path to valorisation: the re-use of digital assets which arises from their capability to be modified, assembled, combined, or appropriated differently. The re-use of digital cultural heritage is largely confined to digital assets which are in the public domain or have received creative commons licences, and measures to determine the impact onto the cultural and creative industries or to estimate the re-use by groups or communities have still to be found. However, such re-use will certainly be of growing importance in the future and thus meaningfully influence the value of cultural heritage.

Finally, another perspective on the value of digital cultural heritage comes from its use as big data. The best example for such data are texts (digitized books), but also scans of images and digitally available music. Private companies as well as research institutions are definitely interested in bulk downloading and piling up huge amounts of data. Such use presents an opaque commodification of digital cultural heritage, since it has as a precondition the capability of such institutions to process the content and exploit it for their purposes. This is often not a task which can be easily performed by small companies for a couple of reasons: Some cultural heritage institutions do not provide the full text of books, but only the images compiled in pdf-files. This results in very large files, the download of which puts a heavy burden on server capacities. The Zentrales Verzeichnis Deutschsprachiger Drucke has therefore limited the download of their 1,8 million digitized titles to 16 GB per day, and the Biblioteca digital hispánica of the Biblioteca Nacional de España allows the download of their contents only item per item. If whole books are available only as images, they need to be OCR'd in order to extract the text. This is, for example, the case for manuscripts, a good part of books printed in German Fraktur, or for incunabula from the 16<sup>th</sup> and 17<sup>th</sup> century, which still stand in the way of the provision of high-quality OCR because of their non-standardized orthography and layout. And even if millions of texts are finally available in plain text format, interested institutions have to provide the computational power, enable the application of statistical models and technologies like machine-learning methods, and they have to employ capable developers of algorithms, all of which are necessary preconditions for dealing with big data. The fact that data being expensive and resource-intensive assets and that "only

already powerful institutions – corporations, governments, and elite research universities – have the means to work with them at scale” (D’Ignazio/Klein 2020: 41) was reflected in the interviews, when one of the interviewees pointed to “the National Supercomputing Centre, which is interested in exploiting this information to train artificial intelligence software to find different ways of natural language processing as well as different applications in the legal field or in medical research. They think that this big data corpus is a good training field for these artificial intelligence systems, and for the development of software” (I8). The value of digitized cultural heritage therefore does not only lie in the fact that the significant cost of digitization has been taken over by the taxpayers in the last instance, but also in the quality of the content. Several interviewees pointed out that they understand publishing houses as filters ensuring a certain quality of the content: “I think libraries and maybe that is the Google idea, libraries have high quality content in contrast to the unsupervised content they find out there. Internet is the source of terabyte of information but of different quality. So, my guess is that companies see books in digital libraries as a source of verified content. I mean as far as it can be verified. But at least there is an average quality which is above or at least above a lot of content that you can find out there” (I3, in the same vein also I13). The specific challenge of making use of digital cultural heritage as big data lies in the asymmetry it produces: While the general idea is that all users should be provided with access to the digital assets, it is only a small fraction of already powerful institutions which are capable to exploit such data for whatever use they are aiming at.

### **The Use of Cultural Heritage and Public Law**

Cultural heritage institutions act according to their missions and visions, as has become clear in the discussion of “goals and objectives” above. But how does public law conceive of the function of cultural heritage within the political system? What is the purpose of the protection of cultural goods which is most often part of European public law? And how does the digitization of cultural heritage relate to public law? German law scholar Charlotte-Sophie Lenski (2013: 299–316) has identified only two examples within European public law, where the meaning, purpose and function of cultural heritage and its protection have been positively defined in the law. The first example comes from Italy, which has a “Codice dei beni culturali e del paesaggio” (c.b.c.). Already in the first article two terms are introduced: “tutela”, which can be translated with protection, and “valorizzazione” of the cultural heritage, which can be translated with exploitation, improved access and public use, as well as with valorisation. Article 1 furthermore puts these two aims into an ideal context, since both aims serve the purpose of preserving the memory of the national community and its territory (“paesaggio”) as well as advancing cultural development (“fruizione”). While the protection of cultural goods facilitates their public use and valorisation aims more concretely at their use by increasing the possibility of appreciation (e.g., by dissemination and increase of knowledge) or the possibility of actual access to the cultural asset, the advancement of cultural development aims at a future-oriented perspective with the creation of new cultural works at its centre. The preservation, valorisation and creation of cultural assets are brought into a common context that results in two essential functions: On the one hand, a backward-looking function of cultural heritage, aimed at integrating the community, and on the other hand, a forward-looking function, aimed at the cultural development of the individual and the community.

The second example presented by Lenski is the Portuguese Lei do Património Cultural (LPC), which uses the terms “protecção e valorização” in a similar vein as the Italian public law has it. Though protection and valorisation may have different nuances in the meaning and in their assignment, both European public laws put their emphasis on the side of the reception, with the consequence that any public protective measure has to put its emphasis even-handedly onto the substance of the cultural good as well as onto its reception. In a quite general way, Lenski concludes from these examples that public cultural law as a whole describes a triad of measures, consisting of protection, value creation and promotion, referring at each point to the intellectual content of the cultural works. These three measures ultimately culminate in the creation of new cultural content, new cultural works which build on the preserved stock of cultural signs. Both protection and value creation can be advanced independently within the framework of targeted promotion, but also serve new creation. All three functions end in a development process that can relate either to the intellectual development of the individual, to the political development of the community, or to economic development (Lenski 2013: 316). Public cultural law is thus oriented not only to the creators of the intellectual content and the protection of their works, but also at the reception of cultural goods, with the double function of integrating the community and advancing the future cultural development of individuals and communities.

It is obvious that the conception of cultural heritage in European public law does not take into account digitization or the access to digitized cultural heritage in the virtual space of the internet. With regard to the European legal framework, it should be noted here that the European Directive on Copyright in the Digital Single Market has introduced an exception for the preservation of cultural heritage which can be parallelized to the aim of protection in public law. Article 6 of the Directive (European Commission 2019) introduces a mandatory exception for cultural heritage institutions that allows for making reproductions for the preservation of works that are permanently in their collections. In other words: Cultural heritage institutions now obtain a high degree of legal certainty if they digitally reproduce, with the aim of preservation, the works they have acquired and which are thus a permanent part of their collections, irrespective of whether this work still falls under intellectual property rights restrictions or not. However, digitization for the sake of preservation (which can also be read as protection) does not enable accessibility of these digital assets.

Beyond this exception, it must be noted that public cultural law is marked by a traditional understanding of valorisation: Public use, access and exploitation of cultural heritage are conceived in a way that implicitly takes for granted that the use of traditional, material cultural heritage, its re-use and the creation of new cultural content originates from the reception granted by access. The thinking behind it is that you visit a cultural heritage institution, indulge in the contemplation of these physical objects, and are then inspired to create a new work. The accessibility of digital cultural heritage, by contrast, implies and demands the possibility to re-use the digital asset, and to re-mix, re-arrange, appropriate, complement, and adapt the content offered digitally. While this is possible for all digital assets which are in the public domain and, with restrictions such as non-commercial use, for assets which have received a creative commons licence, it is not possible for assets that in some form fall under the intellectual property rights regime. Therefore, a significant part of cultural heritage produced in the 20<sup>th</sup> and 21<sup>st</sup> century is excluded. The core deficiency that marks public cultural law is thus indicated by the difference between physical objects, which can be accessed in cultural heritage institutions irrespective of the rights regime which has

to be applied, and the part of digital cultural heritage falling under the intellectual property rights regime, which often cannot be accessed online, and which cannot be re-used.

This lamentable deficiency can be stated notwithstanding the observation that current copyright law can be regarded as dysfunctional. The best example for this comes from the United States of America, where the Project Gutenberg – a digital library which currently provides 56.000 eBooks, and which can be understood as a cultural commons – received a court order from the German publishing house S. Fischer, which complained that there were 18 works by three authors available electronically in Project Gutenberg. All these three authors died in the 1950s, and all of their works will be in the public domain by the end of this decade, but according to German law they are currently under copyright. As the Project Gutenberg Literary Archive Foundation has publicised in December 2021, both parties reached a settlement agreement including a compromise. However, it is difficult to see the gain of the publishing house; rather, the damage done to cultural heritage becomes visible, especially when seen from the perspective of public cultural law with its emphasis on reception.

While this court case shows in an exemplary way that intellectual property rights might stand in the way of the use of culture and thus also in the way of the creation of new cultural content, at the same time it shows that it is rather institutions (such as publishing houses) than individuals which insist on the protection provided by copyright. Copyright law has thus mutated into right of the marketing industry living on these works. It is thus moving further and further away from its role of protecting the creators and instead is protecting the economic interests of the industries that market their works (Lenski 2013: 14). While this is a general observation, Haux recounts a more specific case, where a piece of digital music, only five seconds long, has been re-used in a newly created piece of music, which ended in a lawsuit extending over about twenty years (Haux 2021: 3-29). The case of re-use of digital cultural heritage available in cultural heritage institutions and falling under the intellectual property rights regime is therefore blatant: If public cultural law ultimately aims at the creation of new cultural content and at facilitating new cultural works which build on the preserved stock of cultural heritage, how can public use and especially re-use of digital cultural heritage be harmonized with intellectual property rights? Can a public interest in digital cultural heritage as a common good be formulated, an interest which rests on the aim to create new cultural works and which supersedes digital assets marked by intellectual property rights? And finally: Which ways are there to pool resources – in the sense of a cultural commons – of digital assets which have been established within the past one hundred years but are ready for re-use?

### **The Tragedy of the Digital Cultural Commons**

If cultural heritage institutions decide to digitize their holdings, this decision implies a massive change with respect to their user groups. Consequently, they have to change from inward-looking institutions focusing on acquiring, preserving and administering their content to institutions serving global communities. Digitization implies globalisation and a significant increase of audiences. The systematic comparison presented above has indicated this augmentation by pointing to the “outsiders” to the commons, consisting of globally distributed users and companies. From a commons perspective, an increase in the number

of community members or neighbours always entails the danger of an over-use of the shared resources. Though not even Ostrom and her collaborators were able to provide a number indicating from where onwards the balance between the number of participants in the commons and the management of the commons is being overturned, an increase in the number of users also increases the likelihood that individual users might act according to their self-interest and cause the depletion of the resources. This cannot be said of digital resources, which are intangible, ubiquitous, naturally nondepletable and usable in non-rival way (Euler and Dreier 2004: 162). However, a “tragedy of the digital commons” can be imagined in two ways: Firstly, if private companies exploit the resource to improve their services and maximise their profits without contributing to the sustainability of the resource. With regard to open research data, the risk of a loss of communal benefits due to actions motivated by self-interest has been termed the “Tragedy of the Data Commons” (Yakowitz Bambauer 2011: 4). In the digital sphere, such a scenario is even more likely if digital assets are available in huge collections which can be viewed as big data, the establishment of which entails high economic costs which have been funded by taxpayers’ money, and the effective exploitation of which are limited to a few companies and research institutions. As for example the development of the language model BERT has shown, big tech companies like Alphabet Inc. make use of freely available resources and government-financed digitization efforts, since the language model was created out of data from an aligned MovieCorpus/BookCorpus dataset with 800 million words, provided by the University of Toronto, and English Wikipedia with 2,500 million words (Devlin et al. 2018). Though this may conform to current legal standards, the legitimacy of such an exploitation is questionable, since such a situation also stabilises the asymmetry between big tech companies and small and medium enterprises which currently characterises the market. Secondly, even though digital assets do not form a depletable resource in itself, it has to be considered that their ubiquity and dematerialisation depend on material conditions. Digital contents require physical media like server capacities, which can only be used in rivalry, or they necessitate the consumption of energy. Accordingly, a form of tragedy can be constructed out of these sustainability-related issues (Euler and Dreier 2004: 164).

With regard to the first aspect, cultural heritage institutions have several possibilities to restrict the acquisition of big data by private companies or to confine bulk download. Beyond technical restrictions such as the introduction of a download limit, cultural heritage institutions can always license the download of their contents. Licenses can be provided on the basis of the European Directive on the legal protection of databases (European Commission 1996), because the digitization of collections by cultural heritage institutions normally imply the creation of databases. Collections themselves may be understood as databases, or such databases would include information about each work included in the collections. Private companies would have to contact the individual cultural heritage institution to negotiate on the licences and the fee to be charged, a case which is not unlikely, as one interviewee revealed: “they have to contact us if they would like to bulk download, they really have to contact us to get that, but that would be possible” (I1). The legitimate interest of such users has to be proven (see interview I12) in order to acquire access to huge collections. Furthermore, cultural heritage institutions can differentiate between individual users and provide licences reflecting different usage scenarios and factual preconditions such as differing capacities of companies to process and analyse the data to which they provide access. The institution providing a licence and demanding fees can thus differentiate

e.g. between small and medium-sized businesses (SMBs) and big tech companies. This is allowed – but not proscribed – in the Data Governance Act (European Commission, 2020).

It might be argued here that the provision of licences and request of fees may be contrary to the idea of a commons openly providing access to all of the available resources, which has been described above as an informal norm. Demanding fees for accessing the content would imply that data are not ‘open’ anymore in the strict sense of the word; consequently, the term “shared data” would be more appropriate here (I14). However, licences do not provide the possibility to differentiate between individual users, and the request of fees is a non-compulsory option on the use of which cultural heritage institutions can decide by themselves after having balanced the arguments pro and contra (I14). From the perspective of a commons, it is not illegitimate to request fees for bulk download, since the basic principle applies that only those who contribute to the commons may participate in the use of the resources. In this way, the number of outside users is reduced since non-members of the community are integrated via a contractual agreement, and the maintenance of the resources is secured.

However, the general idea is not to exert a discriminatory treatment on the users. Even if we can assume that private companies like Google systematically collect openly available content from cultural heritage institutions, they cannot be considered as free riders. As has been noted above, Google contributes to the accessibility and findability of the content in cultural heritage institutions and therefore does not only profit from the data over there (I12). Charging fees for providing data is neither illegitimate nor unfair, and if an institution is of the opinion that economic giants like the big tech companies should not be strengthened further by supplying them with comparably cheaply acquired data which have been established with taxpayers’ money, this discussion should be shifted towards the political arena. As one interviewee remarked, it might be fairer and more reasonable to change the law of taxation or to remove tax advantages for big tech companies than to overcharge them with fees imposed on the acquisition of data (I14).

With respect to the second conceivable tragedy – rivalrous use of the physical infrastructure – the same solution applies: if companies request unusual large amounts of data, this would put an unusual high burden on the digital infrastructure, which justifies charging of fees to subsidise maintenance, provide computing capacity or enhance server capacities. Request of donations for this purpose seems to be a viable alternative. One interviewee underlined that the big tech companies have learned in the past twenty years that privatization and monopolisation are a dead end, and that they therefore have changed their policy: now they support open-source projects, open access and innovation sharing (I12). It seems convincing that large companies would remunerate cultural heritage institutions for providing large data corpora since they do not only profit from the data provided but may benefit from the creative output which is enabled by the data provided. Shared resources feed swarm intelligence which triggers innovations which even large companies cannot create. Finally, sharing of costs for the maintenance of the commons, e.g. via sponsoring agreements, can lead to a win-win-situation for both sides, since big tech companies can symbolically profit from the reputation of cultural heritage institutions, their authority in selecting cultural heritage to be preserved and their capacity to create value (I12).

Agreements on cooperation between cultural heritage institutions and private companies, on the other hand, were seen critically by the interviewees, since they necessarily imply



unresolvable dilemmas. Within partnerships aiming at the digitization of cultural heritage, for example, the interests of the parties involved are quite diverse. While private companies are ultimately interested in the monetarisation of the data, the establishment of which they subsidise, cultural heritage institutions put their focus on the preservation of cultural heritage and the provision of access, for the sake of the common welfare and quite abstract ideals such as the collection of items relevant for a community's identity or the provision of a database conserving the language of a minority. Private companies – and the Google Books project provides the best example for it – may finance digitization for a yet unknown purpose or a yet unanticipated application. If digitization is done within a cooperation, this approach to data creation and storage raises questions such as who the owner of the data will be after the end of the contract, who does the data analysis and who profits from its results, and who bears the expenses for the provision of data (I14). Cooperation agreements may therefore require careful considerations with regard to the consequences they have for cultural heritage institutions in the long term, as well as for the dependencies they create.

Cooperation between cultural heritage institutions, on the other hand, might provide a sustainable way of sharing the cost of the maintenance of the commons. If several institutions from the same domain in the cultural heritage sector agree on centralizing digitization and sharing the burden of physical infrastructures, they ensure data sovereignty and stay independent from private companies (I14). As one interviewee explained, the continuous preservation of digital content forms a challenge for cultural heritage institutions: “The content is static but technology to keep it available to is not static. It will be dependent, and it develops over time. You will need power, knowledge, hardware, software, and things like this to make it available over time.” (I4) This challenge can be met if several institutions cooperate, distribute the diverse tasks related to the diverse data formats, with the aim that each partner in the cooperation specializes on one or a few assets which are maintained: “We have the technical basis behind it, and another university offers for example editing projects. This may be a way, but to keep every material in one library, in a single technical environment will not fit, because you have not enough personnel and cannot scale up.” (I4) Again, this example shows that it might be easier for the community members of a specific common to cooperate with each other than with outsiders to the commons, since the interests of the community members may be similar.

The example of large language models indicates another way to mitigate the tragedy of the commons. The establishment of a large language model is a costly endeavour involving a high environmental burden as well as the risk of reproducing unwanted biases with yet unknown consequences for society at large, which is, amongst others, a result of the unqualified compilation of textual resources (Bender et al. 2021). Large language models are finally good examples of how private companies might make use of cultural heritage data to optimise their services and maximise their profits. Cultural heritage institutions could use their digital assets alongside with the available metadata to them to establish high-quality language models by themselves. This opens up the chance to deal with biases in the data, or at least to reach a good judgment of such biases: “And we can, probably we can compensate for some other dangers around – for example bias. Because we understand the content very well and we have a good control of our collection. While what you see in many places when they train transformer models – they need very much data and they do an uncritical harvesting of the internet which of course mirrors the society to some extent, but on the other hand it also mirrors the fact that some of the active people on internet are not too

friendly. They are not very polite while the majority might be somewhat polite. But when you look at the internet it is kind of different, but if you look at our collection, it is, I think it mirrors our society pretty well.” (I1) For such purposes, cultural heritage institutions can even make use of digitally available copyrighted content, since the Directive on Copyright in the Digital Single Market (European Commission, 2019) provides an exception for text and data mining. If cultural heritage institutions develop secondary products out of their digital assets on their own, if they establish machine learning models, provide parallel corpora for machine translation, establish large language models for text generation etc., they could provide their users as well as companies with a licensed access to these models or products. Such an approach, which is costly, may include user engagement as crowdsourcing activity (Ridge et al. 2021) and which may only be feasible for larger institutions such as national libraries, would preclude the loss of communal benefits due to actions motivated by self-interest as it would exploit the comprehensive databases at hand for the common welfare, which at same time fits the mission of cultural heritage institutions and minimizes the risk of a “tragedy of the cultural commons”.

### **The Future of Use: From Intellectual Property Rights to Remix Culture**

An opposition between the two main legal regimes – public domain and intellectual property rights – and the claims to use even copyrighted material is not particular to digital cultural heritage. James Boyle has described public domain and the idea of the commons as the outside of intellectual property and characterized the battles on the use of copyrighted material conducted in the internet as the range wars of the information age (Boyle 2008). Jonathan Zittrain has clairvoyantly the function of the big platforms as gatekeepers controlling content and code and thus prohibiting the free unfolding of generative ideas and the re-use of code, in this way pointing to the confinement of innovation (Zittrain 2009). And Flath et al. (2017) have investigated the emergence of creative remixing patterns enabled by the re-use of existing knowledge shared in open online communities. All these contributions identify discussions conducted in parallel to the ones in the cultural heritage field, thus underlining that the challenge identified here is not unique to digital cultural heritage.

In the near future, the digitization of cultural heritage will change with regard to several important dimensions. The most significant change may not necessarily pertain to the quantity of the available digital objects. With museums each holding millions of objects and archives containing kilometres of document shelving, the ambition to digitise will always outstretch the available resources to do so. But there will be significant leaps in quality. Technical advances will enable high-quality digitisations e.g., of three-dimensional objects, and the use of state-of-the-art techniques like motion capture, photogrammetry, linear and laser scanning, panoramic video, stereographic panoramas will become more common. Generally, the quality of the digital assets provided will be improved in a way that alleviates the use of such material, which may inspire a broad range of creative uses. In addition, there will be a steep increase in the number of born-digital material handed in to cultural heritage institutions, which also broadens the possibilities for the re-use and analysis of these assets. Material posted on social media, for example, necessitate a consideration of the context in which it was placed, as well as the reactions it triggered, such as likes and comments, retweets, adaptations, satires etc.

While these changes in quality bring along with them a heightened attention, a higher acceptance, and a more intense engagement with digital cultural heritage, this will also affect the re-use of the digital material according to the expectations and visions of cultural heritage institutions. The availability of digital assets comprising enhanced capabilities for re-use and providing the full breadth of possibilities for exploitation characteristic of digital media does not only match the cultural practices of communities producing digital cultural content, it also will certainly trigger the desire of such users, and especially those working in the creative industries and as cultural entrepreneurs, to engage not only with material which is in the public domain and has been created a long time ago, but also with assets from the recent past or, more generally, from material produced in the 20<sup>th</sup> century. In other words: they will be looking for digital objects to which they have some personal relationship, which they perceive to be part of their own cultural environment, its history and tradition, and the culture of which they want to roll forward and recreate. As Mayo Fuster Morell has carved out, the emphasis of such communities is on use and re-use, which may be generally seen as a characteristic of digital commons, understood as “information and knowledge resources that are collectively created and owned or shared between or among a community and that tend to be non-exclusive, that is, be (generally freely) available to third parties. Thus, they are oriented to favor use and reuse, rather than to exchange as a commodity.” (Fuster Morell 2010: 5). Digital media have been around for more than a quarter of a century now, and the intercourse with it has not only informed cultural practices at large but has also had a decisive influence on the self-conception of cultural heritage institutions when they started their digitization endeavours at the turn of the millennium and decided to provide their holdings in an openly accessible manner. If cultural heritage institutions therefore aim to support such user requests, their commitment would foster and reinforce exactly such an understanding of culture out of which they have been acting themselves in the past decades.

But access to digital cultural heritage which has been created in the recent past is mostly prohibited by intellectual property rights, which may result in an under-use of such assets and therefore lead to an underproduction in the cultural sector and a cultural stagnation. However, especially cultural heritage institutions have a range of means at hand by which they can strengthen the claims and aspirations of their audiences. First of all, they can negotiate with the legators, with rights holders of the legacies being handed in, with living individuals as well as with institutions providing data on the conditions according to which born-digital contents as well as assets which may be digitized can be re-used. Such an approach might open up restricted spaces where exceptions from intellectual property rights can be made and the re-use of the digital assets for the benefit of culture created anew is enabled. By privileging cultural heritage which is donated including rights to re-use it, cultural heritage institutions could profit from their power of ennobling cultural objects, and legators would emerge as benefactors of a new kind. An interesting example for this is provided by contemporary art. While the digitization of such cultural products is already enabled by the European Directive on Copyright in the Digital Single Market (European Commission 2019) for the sake of preservation, contemporary artists, copyright holders or legators who are familiar with this specific segment of the art market may be well suited to understand the benefit the contemporary art scene (and for culture writ large) of making such works available digitally.

A second approach for cultural heritage institutions can be established if they engage with communities already working with digital data in manifold forms. Occupying themselves with

the cultural practices exerted by such communities would on the one hand open up the possibility for cultural heritage institutions to better understand the approaches and the needs of potential re-users of digital cultural heritage, such as how such communities relate to the material they work with, how they conduct their discussions about production processes, perform cooperation and exchanges, or how they evaluate contributions and recreations. On the other hand, this occupation with a specifically digital form of cultural production by cultural heritage institutions would broaden the recognition of such communities and sharpen the understanding that even cultural practices that may be seen as recent developments have not only established new forms of interaction and spaces of cultural production, but also in the meantime coined habits and established traditions themselves. It may at first glance seem to be awkward to regard Snapchat as a part of a youth culture which has established and formed traditions (I12), or to understand the tradition of re-using memes as a cultural practice (I14). But clearly such a conception builds bridges between contemporary digital culture, cultural heritage institutions and the current and future user groups they intend to work with. At the example of the Demoscene provided above, cultural heritage institutions could support the recognition of the traditions, customs, and the collective character of such digital communities by national UNESCO commissions. This means on the one hand that contemporary cultural practices of creating culture are themselves attributed the status of intangible cultural heritage, and on the other hand it broadens acceptance of the fact that the creation of new culture and the advancement of culture often implies the recourse onto prior works, a feature which is characteristic of what might be termed “remix culture” and which runs counter to an understanding of the creator as a genius independent from her social environment, which lies at the heart of European intellectual property rights. In an indirect way, the recognition of communities practising remix culture would therefore fuel the discussions around the overdue reformation of copyrights and their focus on “works”, and it would move the focus of these discussions to a more contemporary conception of cultural production marked by collective endeavours and performed within an environment characterized by sociotechnical assemblages.

Thirdly, cultural heritage institutions can extend their engagement with creative communities not only by supporting the recognition of their collective modes of creation as intangible cultural heritage, but by inviting them to collectively create and curate digital resources which can openly be shared without interfering with copyright or other restrictive licenses. Such digital resources can be incorporated into the pool of digital assets and would thus be regarded as digital cultural heritage in itself. In this way, a specific resource could be created which no longer has to be attributed to one of the two standard legal regimes of public domain or intellectual property rights in the common sense of these terms, but which rather resembles the collective wisdom and collected knowledge of indigenous groups, to which no individual intellectual property rights apply. If such creative communities do not wish to provide open access to the resources they contribute, but prefer to restrict access and re-use to the members of their community, this would provide an example of a commons (or “Allmende”) in the traditional sense, i.e. a digital cultural commons resource maintained by an elitist community, characterized by the importance of obligations, self-regulation, and trust into the normative framework established alongside the resource (see also Haux, Chap. 7, 149–202).

Finally, cultural heritage institutions can collect and consolidate good practices around the re-use of digital cultural heritage, as well as lessons learned about the re-use of digital assets

by the cultural industry. In Germany, the Coding da Vinci events serve as both hackathons working with digital assets coming from digitization endeavours in the cultural domain, as well as a publication platform enabling cultural heritage institutions to publish open cultural data for public re-use. If such examples were pulled together, this collection of good practices may help identifying patterns in the re-use of data and thus characterizing typical approaches to the creation of new culture. Such collections of good practices would in turn support policymakers in formulating exceptions from copyright specifically for digital cultural heritage. Comparable to the exceptions for “fair use” in US-American copyright (Madison 2004), such an exception could amend the Directive on Copyright in the Digital Single Market (European Commission 2019). It would be based on the patterns of re-use documented as cultural practices by underlining the importance of societal transmission of culture and the collective dimension inherent to it (I 13), rather than the individualistic dimension of “property” (Lenski 2016: 141). In this way, variability could be introduced into the strict regime of intellectual property rights.

## Conclusions

In this study, digital cultural heritage has been examined within a cultural commons framework as a resource that is being established and maintained by cultural heritage institutions. But this particular commons extends beyond the resource and community members in the strict sense, and includes various types of users as well, irrespective of the fact whether they should be regarded as part of the commons or as “outsiders” to it. In this way, the commons framework conceptualizes the digital cultural commons as a virtual space characterized by the social and cultural practices exercised within it. It thus encompasses various interactions between the diverse user groups and the digital resources provided by cultural heritage institutions. This approach captured important shifts that have taken place in the time span since the beginning of the first digitization activities by cultural heritage institutions. The first shift is marked by a functional change of the digital assets: for cultural heritage institutions, the task is now not longer to just push digitized objects ‘out there’, but to provide material which is being used and re-used for the creation of (digital) culture. This is on the one side facilitated and alleviated by improvements in the quality of the digital objects; on the other side, cultural heritage institutions have to face the dilemma that the provision of high-quality digital content enabling re-use is limited, especially with regard to digital-born content or content produced within the past 30 years, both of which are most of the time protected by intellectual property rights. There are certainly a couple of ways enabling cultural heritage institutions to bypass such a rights regime, but it is also obvious that the more feasible of them point to the creation of closed virtual spaces in which binding rules and obligations exist which enable the re-use of the digital assets for creative purposes. This result can be seen as a peculiar outcome of this study: whereas digitization was begun after 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 unhinged by the provision of closed spaces in which users can access the digital material under certain conditions. This latter model strikingly resembles the classical and historical conception of a commons: a shared resource managed by a community for the benefit of its members, where the latter have to be understood as an elite to which access to the resource

is granted in exchange to various duties and obligations, rules and practices which are necessary to maintain and recreate the resource in order to keep it sustainably.

The second shift recorded by this study points in the same direction, which can be described as a move from the open to the closed. While digitization in the beginning was motivated by the imperatives “preservation and access”, the accumulation of huge amounts of digitized objects has fuelled the transformation of digital cultural heritage into a commodity which is regarded as invaluable in the time of big data. The establishment of large amounts of high-quality data is a costly endeavour, even more if these data are processed into machine learning models, large language models or generative pre-trained transformers which are capable to create new visual or textual content. Both big data as well as high-quality models generated out of them are items which can easily be commercialized. This potential of digitized cultural heritage has only been unrolled in recent years, and the transformation of cultural heritage into a commodity is certainly an unintended consequence of digitization. And again, the potential commercialization of digital cultural heritage points into the same direction of an elitist conception of a closed commons: access to the resource is only granted in exchange for compensation, for example by way of licencing accompanied by fees, thus avoiding a “Tragedy of the Cultural Commons”.

A move from the open to the closed collides at first glance with the ideal of open access for everybody worldwide which several of the interviewees underlined as an informal norm, and it may point into the direction of the compartmentalisation of the internet which can currently be observed in many places. However, it should be noted that the examination of digital cultural heritage as a cultural commons in any case revealed a differentiation of users (and companies respectively), with the consequence of a basic distinction between users / companies regarded as members of the community and those to be regarded as outsiders. Both shifts described above speak for a further differentiation of the users and companies which can be regarded as members of the community. From a commons perspective, a strengthening of the bonds between the community members – in this case cultural heritage institutions, users and companies – also includes the chance to strengthen the commons as a whole and to foster its self-organization and management. If registration is needed for users to be granted access to specific parts of the resource (to provide special provisions and obligations), this also opens up the chance to integrate such users for the benefit of the resource, for example by contributing to crowdsourcing activities which are needed to establish the training data enabling machine learning, or to contribute to the provision of descriptions of images, an activity which is manageable for a significant amount of images only if you a broad pool of contributors is available. Regarding the integration of private companies into the commons, long-term sponsorship agreements may form an alternative to demanding fees for the use of the resource and, if applicable, of secondary products created out of it. Such agreements enable a sustainable maintenance of the resource and to secure for example the funding of the technical infrastructure, comprising servers, energy costs and technical personnel, etc. From the perspective of a commons, such a differentiation of the various roles of users and companies and the specification of the relationships between the various members of the community including obligations and duties provide obvious solutions to the issues identified in this study.

## Recommendations

### For Cultural Heritage Institutions

**1) Enable access to copyrighted material in non-downloadable form:** In order to pursue the aims of enabling access as openly as possible, thus facilitating use and reception of the cultural works, and of supporting their re-use and the creation of new cultural content even from copyrighted material, cultural heritage institutions could provide such material in a non-downloadable form. The Norwegian “Bokhylla” or bookshelf agreement provides a precedent for this: Users can virtually ‘open a book’ and read a certain number of pages within it, but they cannot download any content, and they need to have a Norwegian IP-address. Paradigms for such an approach are books and newspapers, for which licenses from copyright holders or publishers were asked or specific agreements with them were made. Compensations may be needed by printing houses, like it is common practice in libraries. Collaborations with copyright holders may be another way to enable access, for example if the material is not yet digitized, the cultural heritage institution could do the digitization and ask for a contribution to the funding of the digitization endeavour. The collection of model contracts is helpful, since smaller institutions holding copyrights will be more willing to cooperate if there are model contracts available that have been negotiated with larger institutions.

**2) Pool digital resources which are available for re-use:** The provision of pools of contemporary high-quality or born-digital content exempt from intellectual property rights enhances the re-use of the digital assets and stimulates creating culture anew. Cultural heritage institutions can negotiate with legators, devisors and the rights holders handing in the legacy and ask for fair access and re-use of the digital material in exchange for the increase in value created by the ennoblement of such material as cultural heritage. Cultural heritage institutions can invite communities accustomed to work with digital material to create and openly share these assets as a common resource. In this way, current cultural practices which have been established around digital material receive recognition and acceptance, thus shifting the focus from the work of an individual genius, which is at the centre of the current intellectual property rights regime, to contemporary collective practices of re-mixing, re-arranging, appropriating, complementing and adapting content which is available digitally. Collections of best practices around the re-use of digital cultural heritage further stimulate the cultural field as well as the creative industries to work with the established pools.

**3) Develop secondary products out of the available digital assets:** Cooperation between cultural heritage institutions enables not only sustainable ways of sharing the cost of the maintenance, but also the pooling of resources and the delegation of tasks within the cooperation. Suitable institutions can assume the task to develop secondary products out of these shared resources, such as machine learning models or large language models for text generation in chatbots etc., or they can cooperate with research institutions to do so. Within the European Union, the establishment of language banks and parallel corpora for machine translation is an important endeavour which supports the integration of the Union and which therefore will receive funding. Furthermore, the development of secondary products provides the chance to integrate users via crowdsourcing activities, especially where their

contributions cannot be matched by the human resources available in the cultural heritage institutions. This is, for example, the case with regard to the preparation of large datasets for machine learning, where annotations for texts or captions and descriptions for images are needed.

## **For Policy Makers**

**Work towards a legal framework recognizing contemporary forms of collective cultural production:** Intellectual property rights form a main blocker for re-using digital cultural heritage with the aim of creating culture anew. However, the European Directive on Copyright in the Digital Single Market contains an exception from copyrights which allows digitization for the purpose of preserving cultural heritage. This exception could be reformulated and amended with the aim of creating new cultural content by re-using the stock of digital cultural heritage which is still under copyright. Such an exception may be made under certain conditions; access may be restricted to digital assets held by cultural heritage institutions, it may be limited to communities with a proven record of digital and collective cultural practices, and it may restrict re-use to non-commercial purposes. The legal recognition of the remix culture in the musical scene, for example, would foster a contemporary conception of cultural production marked by collective endeavours and performed within an environment characterized by sociotechnical assemblages. Such an exception would identify and formulate a public interest in digital cultural heritage as a common good, an interest that aims at the creation of new cultural works and the unleashing of creativity for the common welfare.



## Bibliography

- Beer, David. 2013. Algorithms: Shaping tastes and manipulating the circulations of popular culture. In: Beer, David (ed.), *Popular Culture and New Media: The Politics of Circulation*. London: Palgrave Macmillan, pp. 63–100.
- Bender, Emily, Gebru, Timnit, McMillan-Major, Angelina, and Shmitchell, Shmargaret. 2021. On the Dangers of Stochastic Parrots. Can Language Models be Too Big? *Conference on Fairness, Accountability, and Transparency (FAccT '21)*, March 3–10, 2021, Virtual Event, Canada (pp. 610–623). ACM, New York, NY, USA. DOI: <https://doi.org/10.1145/3442188.3445922>
- Benjamin, Walter. 1977. Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit. In W. Benjamin, *Illuminationen. Ausgewählte Schriften I*. Frankfurt am Main: Suhrkamp Verlag, pp. 136–169. Translation into English: <https://www.marxists.org/reference/subject/philosophy/works/ge/benjamin.htm>
- Boyle, James. 2008. *The Public Domain. Enclosing the Commons of the Mind*. New Haven / London: Yale University Press.
- D'Ignazio, Catherine and Klein, Lauren F. 2020. *Data Feminism*. Cambridge, MA/ London: MIT Press.
- Devlin, Jacob, Chang, Ming-Wei, Lee, Kenton, and Toutanova, Kristina. 2018. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. arXiv:[1810.04805v2](https://arxiv.org/abs/1810.04805v2).
- Euler, Ellen and Dreier, Thomas. 2004. Creative Commons – iCommons und die Allmendeproblematiken. In: Bourcier, Danièle and Dulong Rosnay, Mélanie de (Hg.): *International Commons at the Digital Age. La création en partage*. Paris: Romillat, pp. 155–169.
- 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>
- 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](http://www.europarl.europa.eu/doceo/document/A-8-2018-0245-AM-271-271_EN.pdf)
- 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>
- Flath, Christoph M., Friesike, Sascha, Wirth, Marco, and Thiesse, Frédéric. 2017. Copy, Transform, Combine: Exploring the Remix as a Form of Innovation. *Journal of Information Technology*. 32:4, pp. 306–325. DOI: <https://doi.org/10.1057/s41265-017-0043-9>
- Frischmann, Brett M., Madison, Michael J., and Strandburg, Katherine J. 2014. Governing Knowledge Commons. In: Frischmann, Brett M., Madison, Michael J., and Strandburg, Katherine J. (eds.), *Governing Knowledge Commons*. Oxford: Oxford University Press, pp. 1–43.

- Fuster Morell, Mayo. 2010. *Governance of online creation communities: Provision of infrastructure for the building of digital commons*. PhD Thesis. [https://www.onlinecreation.info/outline\\_design/](https://www.onlinecreation.info/outline_design/)
- 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>
- Haux, Dario Henri. 2021. *Die digitale Allmende. Zur Frage des nachhaltigen Umgangs mit Kultur im digitalen Lebensraum*. Zürich / St. Gallen: Dike Verlag. DOI: <https://doi.org/10.3256/978-3-03929-012-3>
- InDICEs Project. 2021. *Deliverable 2.1. Mapping of the relevant European IP legal framework*. DOI: <https://doi.org/10.5281/zenodo.5141439>
- Klimpel, Paul. 2021. *Urheberrechtsreform 2021. Neue Chance für das kulturelle Erbe*. Berlin: Zuse Institute Berlin. URN: <https://nbn-resolving.org/urn:nbn:de:0297-zib-84315>
- Klimpel, Paul. 2020a. *Kulturelles Erbe digital – Eine kleine Rechtsfibel*. Hrsg. vom digiS Forschung- und Kompetenzzentrum Digitalisierung. Berlin: digiS. DOI: <https://doi.org/10.12752/2.0.004.0>
- Klimpel, Paul. 2020b. Neue Chancen für das kulturelle Erbe, in: *Bibliotheksdienst* 2020, 54 (6–7): pp. 559–564.
- Klimpel, Paul. 2019. Overview Article: European Cultural Heritage and Copyright, European IP Helpdesk, October 2019. [https://intellectual-property-helpdesk.ec.europa.eu/regional-helpdesks/european-ip-helpdesk/europe-ip-specials/europe-ip-specials-copyright-and-cultural-heritage\\_en](https://intellectual-property-helpdesk.ec.europa.eu/regional-helpdesks/european-ip-helpdesk/europe-ip-specials/europe-ip-specials-copyright-and-cultural-heritage_en)
- Kymlicka, Will. 2005. The moral foundations and geopolitical functions of international norms of minority rights. A European case study. *Anales de la Cátedra Francisco Suárez* 39, pp. 209–241.
- Lenski, Sophie-Charlotte. 2016. Öffentliches Kulturrecht und Urheberrecht als wechselseitige Auffangordnungen, in: Goldhammer, Michael / Grünberger, Michael / Klippel, Diethelm (Hrsg.), *Geistiges Eigentum im Verfassungsstaat. Geschichte und Theorie*, Tübingen: Mohr Siebeck, pp. 141–157.
- Lenski, Sophie-Charlotte. 2013. *Öffentliches Kulturrecht. Materielle und immaterielle Kulturwerke zwischen Schutz, Förderung und Wertschöpfung (= Ius Publicum, 220)*, Tübingen: Mohr Siebeck.
- Madison, Michael. 2004. A Pattern-Oriented Approach to Fair Use. *William & Mary Law Review* 45:4, pp. 1525–1690.
- Ostrom, Elinor. 1990. *Governing the Commons*. Cambridge, UK: Cambridge University Press.
- Ridge, Mia, et al. 2021. *The Collective Wisdom Handbook: Perspectives on Crowdsourcing in Cultural Heritage*. DOI: <https://doi.org/10.21428/a5d7554f.1b80974b>
- UNESCO. 2003a. *Convention for the Safeguarding of the Intangible Cultural Heritage*. <https://ich.unesco.org/en/convention>
- UNESCO. 2003b. *Charter on the Preservation of the Digital Heritage*. <https://unesdoc.unesco.org/ark:/48223/pf0000229034.locale=en>

Yakowitz Bambauer, Jane R. 2011. Tragedy of the Data Commons. *Harvard Journal of Law and Technology* 25:1, pp. 1–67. DOI: <https://dx.doi.org/10.2139/ssrn.1789749>

Zittrain, Jonathan. 2009. The End of the Generative Internet. *Communications of the ACM* 52:1, pp. 18–20. DOI: <https://doi.org/10.1145/1435417.1435426>

## Annex

### Annex I: Code List for the Qualitative Analysis of Interviews Used in Atlas.ti

<b>Code Name</b>	<b>Code Definition</b>
access	Access to cultural heritage
access: complete	Examples in which there are no restrictions for access
access: licenses	Use of licences for access
access: threshold	Existence of a threshold to limit access
AI (Artificial Intelligence)	Use artificial intelligence applications
benefits	Benefits of digitization for various participants
British CDPA 1988	British Copyrights, Designs and Patents Act 1988
copyright	Copyright restrictions
copyright: exception	Agreements made on copyrighted materials to grant access to it
costs + human resources	Costs arising from the commons
cultural practices	Text passages referring to cultural practices
differentiation of commons	Difference between various forms of commons
dilemmas	Dilemmas cultural heritage institutions have to face
Directive on Digital Single Market	Mention of the DSM
goals + objectives	Goals pursued with establishing the commons
governance	Governance of the commons
historical commons	Reference to earlier forms of commons
incentives	Incentives to participate in the commons
information policy	Policy used by governments
intangible cultural heritage	As opposed to material cultural heritage
intellectual property right	Rights restricting use
knowledge creation	Knowledge creation performed in the commons
legitimacy	Text passages describing legitimacy of cultural heritage institutions
maintenance	Measures securing maintenance of the commons
monopoly / antitrust	Text passages referring to antitrust measures
obstacles + challenges	Obstacles and challenges faced with regard to the commons
open data policy	Text passages referring to policy
participants	Community members
preservation	Goal in establishing the commons
profit: common welfare	Goal of cultural heritage institutions
public domain	Label used for digital cultural heritage
public good vs. common property	Differentiation between two legal concepts
recsys	Recommender systems

regulations	Regulations pertaining digital cultural heritage
resources/assets (characteristics of)	Text passages describing the resource
risks	Risks associated with the commons
rules-in-use	Relationships between participants and users
sociotechnical assemblage	Current environment of users
use	Use of digital cultural heritage
use (2): creation	Creation of new cultural heritage
use (2): creation - ancillary copyright	Ancillary rights arising out of newly created cultural content
use (2): differentiation 'private'-commercial	Text passages referring to a differentiation between possible uses
use (2): education	Educational use made of cultural heritage
use (2): market	Market-related use of cultural heritage
use (2): capabilities for exploitation	Exploitation of digital cultural heritage
value of cultural heritage	Text passages specifying the value of cultural heritage
word embeddings	Text passages referring to this natural language processing technique