Heritage Chains

Marco Cangemi

Heritage Chains

Marco Cangemi

S2111810

Heritage, media and museums

(1044M06Y)

Supervisor: Dr Llanes Ortiz

Heritage and Museum Studies(MA)

University of Leiden

Faculty of Archaeology

Leiden, 7 March 2018

Final Version

Index

Acknowledgements

The topic of this thesis were chosen in the early phases of my MA experience at Leiden University. However, during the time in which this work was developed, several things changed. My academic career, as well as the topic of the lectures I attended during this period, strongly influenced the discussions and approaches adopted in my work, and in this sense, this thesis has been subjected to phases of reorganisation during its development. Nevertheless, the main topic of the project has always sticked to the cornerstone on which all the discussions have been developed.

I would like to thank my Thesis Supervisor Llanes-Ortiz for his support and help and feedback. Without it, this thesis could not have achieved the same quality which could can show today.

Moreover, in the course of the development of this research, I could count on my friends Maia and Mattia which who in addition to their friendship have also supported me and offer substantial help for the development of my research.

Finally, I would like to thank all the staff of Leiden University, professors and my colleagues who have strongly influenced my academic experience.

Introduction

The justification for the thesis

This research investigates the role that cutting-edge developments in information communication technology (ICT) could play in the field of heritage studies and management, particularly in protecting and sharing digital heritage material. Having developed a strong interest in computational technology over the course of my academic career, I sought to exploit this within my MA thesis in order to propose a framework for developing a decentralised platform based on peer-to-peer (P2P) protocols, that would be able to offer a valid response to several issues afflicting the actual status of the field. With the title "Heritage Chains," the aim is to highlight the role which P2P technologies based and influenced by the blockchain could be used to discuss a framework for an environment in which heritage resources could be globally connected, preserved and shared by an indissoluble data system. Therefore, taking into account the primary role played by blockchain technologies and its derivatives, the Heritage Chains has resulted as the best title to describe the characteristics of the project. Specifically, I focuse my attention on digital libraries, which, since the establishment of the World Wide Web in 1991, have played a primary role in protecting and sharing heritage material in the digital space.

To offer a qualitative analysis through appropriate methodology, this research was developed on discussions emerging from the field of digital heritage from both theoretical and practical points of view. The main focus point of my research centres on open debates on the most pressing issues, highlighting the limits as well as the possibilities through an analysis of the current situation and existing of studies. This approach revealed specific problems repeating over the years that, more than others, have characterised the interest and the efforts of professionals, as well as everyone else who, for various reasons, found themselves becoming involved during their work with digital heritage.

Through the definition offered by UNESCO for the term 'digital heritage': '*The digital heritage consists of unique resources of human knowledge and expression. It embraces cultural, educational, scientific and administrative resources, as well as technical, legal, medical and other kinds of information created digitally, or*

converted into digital form from existing analogue resources. Where resources are 'born digital', there is no other format but the digital object' (UNESCO 2003, 75); it is clear that this area embraces in itself a series of highly different fields, which are the result of the application of computational sciences within the field of cultural heritage.

Unfortunately, throughout a decontextualized definition it is impossible to fully express the peculiarity and complexity of this field. To achieve a comprehensive vision of digital heritage concept, it could help the analysis of heritage meaning. Since the end of the last century, heritage concept and its study have started a process of detachment from the history field (Harrison 2010, 10) becoming an independent discipline with a specific methodology. Still today, heritage continues to maintain an unequivocal bond with history, but while in the past the two terms have been used as interchangeable synonyms, today in the term heritage the emphasis has been focused on the use of history for heritage purpose. Indeed, today history could be defined heritage only in case this has been used by communities to satisfy their social and memory needs.

In a broad sense, heritage is represented by the specific relations of a given community with the world around them, the past and the future. As defined by Harrison (2010), *"heritage is the ways in which we go about conserving things – the choices we make about what to conserve from the past and what to discard: which memories to keep, and which to forget; which memorials to maintain, and which to allow to be demolished; which buildings to save, and which ones to allow to be built over"* (Harrison 2010, 9).

Consequently, there is no single narrowed definition to categorize heritage, its meaning is context-dependent, shifting. It should be conceived as a set of relations built through a multitude of unique cultural resources. The value of these resources is connected to the role played within their relative consumer groups and consequently with their attribute as creators of identity and community.

Therefore, it is possible to discuss heritage as a consumer-defined resource (Ashworth 1994, 17-18). An interesting example of the development of social memory is the one reported by Anna Collard (1989) in relation to six communities of central Greece (Evritania) and their relationship with the Ottoman Empire period. These communities possess a strong link in their memory with the Ottoman period, while other periods chronologically and socially were more influential. For

instance, the German occupation during the World War II and the civil war (1946-1949) do not seem to have the same importance and attention in these communities' memories. In this sense, as indicated by Collard (1989), social memory preserves specific events and periods through a discourse in which the past and the present are combined to sustain the village identity. In this discourse, the past ('then') is connected with the memory of the Ottoman period. Still, this constructed past should not be considered as fixed or bounded, but is subjected to the mutability in the villages' life and social contingencies (Collard 1989, 98). The role of social memory has also played a fundamental role in the creation of modern nations and their national sentiments.

Such characteristics highlight the fact that heritage resources are not immutable, having to cope with the specific request of their consumers, these resources have continuously been adapted to the mutable exigence of their consumers. Therefore, it is possible to understand how the main question related to the understanding of heritage should not be 'what is heritage?' but rather 'whose heritage?' (Hall 2005). An interesting example which could help to better contextualise this discussion is the Karula National Park case studied by Kristel Rattus (year). In his work, Rattus developed an analysis of the multiple meanings that a heritage resource could have for different consumers. The Karula National Park is indeed one of the first parks in Estonia in relation to which a systematic debate on heritage issues took place. The park which is located in south-east Estonia has become the main tourist attraction, an element which along with the productive activities has played a main role in the economic flourishing of the local community. Even if it is possible to observe that the protection of the heritage site of Karula National Park has positively influenced the economic development of the area, it must be taken into account, to fully understand the actual situation, the confrontation between community members and institutions. In this regard, Rattus (2011) presents the case of Metsamoor (Wood Crane), one of the local people who live and work within the park. Metsamoor and her husband possess a series of tourist farms within the national park, and could be considered one of the main beneficiaries of the set of heritage protection rules applied within the park.

Nonetheless, as Rattus (2011) presents, Metsamoor argued with park administrators over the high level of restrictions imposed by the administration. Park administrators answered that such restrictions are needed to preserve the heritage

value of the park. As explained by Rattus *'Whereas the administration of the national park preferred the "realising"-principle, Metsamoor's wish was to explain the so-called old folk religion and old way of thought as well as how those old nice customs could be used today'* (Rattus 2011, 138). Even if it is generally possible to sustain the position of local against the imposition of external authority, in this case, both actors should be defined as local and in this sense, both opinions have to be considered on an equal position.

Such element highlights the fact that heritage should be intended as a discursive practice (Hall 2005, 5) over which communities and their multiple voices develop through the years as collective and social memory. In this regards, Heritage has been conventionally understood and defined as a set of resources related to history, traditions and culture. The notion is, however, a field in constant dispute and (re)construction, which is defined by a) communities; b institutions; and c) the state. Therefore "heritage" results in a dynamic notion as exemplified by the Evritania communities, Metsamoor, Karula National park case. Ultimately, however heritage is defined, communities, institutions and governments tend to associate it with objects, texts, information, and activities that once documented are integrated into a repository, archive or library. Especially taking into account the extensive process of digitalisation which has occurred in society since the end of the last century, it is necessary to focus the attention on the new digital solutions which could be used to increase the quality of integration of those resources as well the creation of new ones.

What is digital heritage?

Since the 1990s, human society has started a process of digitalisation which has led, in the course of the last two decades to digital infiltration within all the spheres of society. Indeed, several sectors have exploited the opportunities offered by such an instrument free from the limitations of the physical world. Moreover, the birth of the World Wide Web in 1991 has provided a new social space of interconnection to society; new social structures, methodologies of interaction, and communities were developed. Therefore, it is clear that in those years the field of heritage, as well as other academic branches, have started to focalise their attention in the digital world.

The digital world was initially seen as a valid and useful set tool and software at the service of various disciplines and projects, being able to increase the scope of actions like the protection, management and sharing of resources (Kalay 2007). In addition to this, the process of digitalisation has led to the creation of born-digital material which does not possess a physical counterpart. Taking into account that this material could come to acquire a meaningful cultural and heritage role, specific protection and maintenance processes and technique should be taken into account by heritage specialists. In this context, the concept of Digital Heritage and consequently, its field could be identified.

Digital heritage, as a field, comprehends all those digital techniques of preservation and elaboration of physical and intangible heritage resources as well as those resources digitised or developed in the digital context which possess heritage value and as such these have to be preserved for posterity (Cameron 2007, 172-173; Rahaman and Tan 2010, 93-94)

First of all, it is necessary to understand the role played by digitalisation in society and its impact on peoples' conception of and relation with the world around them. In this regard "classic" heritage institutions have adapted their techniques and practice to fulfil the new requirement of the groups they have to serve. Digitalisation and the expectation of individuals for digital resources have played a leading role, in the latest evolutions these forces have led institutions to modify their policy regarding heritage (Tang *et al.* 2018, 60-62). A striking example could be found in the approach of museums to the publication and sharing of their material through the digital medium. Until the second decade of this century, it was ordinary museums policy to forbid the public to take photos of their material. Museums have generally not established a clear strategy for sharing resources through the web. Today it is not possible to imagine a museum which is not willing to share its material over the internet. This practice has not only become generally applied but also has become fundamental for museums to continue being relevant in contemporary society (Kuan 2014, 48). Visibility within the web has acquired a central place in museum practice and policies. In this sense, people are nowadays incentived to take photos and share it over the internet along with comments, thoughts, and evaluations on their experience.

A further aspect that should be taken into account is the social expectation which people have developed thanks to the massive digitalisation of society and the

appearance of social media. Social expectation could be defined as the evolution of users' requirements resulting from the penetration of social networks within social dynamics. Consequently, it is possible to observe the appearance of a new set of requirements related to the enjoyment of digital experiences (Nisheva-Pavlova et al. 2015, 281; Economou and Pujol 2007, 242-244). Awareness of these expectations for the final satisfaction of the users offers an interesting perspective to investigate the relation of the public with digital heritage and to propose a further improvement in the experience provided to users. As discussed by H. Rahaman and B.K. Tan (year), three operations have to be offered by the digital institutions to be able to cope with users' requests and expectations: Exploration; Manipulation; Contribution.

Nevertheless, Rahaman and Tan in their article are more focalised over the practice of interpretation, which has been generally understood as systematic, applied in the same way by every user. Actually, thoughts and reactions to actions and situations are different from user to user and it is not possible to define a standardise practice of interpretation. Therefore, allowing users to explore, manipulate and contribute within these institutions will result in an increase of the quality of the services offered. Furthermore, institutions need to offer opportunity for interpretation in order to *"address the 'cultural uniqueness' of end-users and overcome the linearity and allow multiplicity in interpretation*" (Rahaman and Tan 2010, 94, their emphasis).

Consequently, interpretation comes to acquire a fundamental role within digital heritage. Indeed, to obtain the full understanding and elaborate on heritage requires an active interaction with resources and not only a mere observation. An interesting example could be the case of a virtual reconstruction of a heritage site. Even if reconstruction allows the free opportunity of navigation within the site's 3D reconstruction, still such experience could not be defined as adequate to understand the inherent significance of the heritage site. As traced by Rahaman and Tan (2010), a comprehensive interpretation relies upon three fundamental aspects: learning; provocation; satisfaction. All this considered, to establish valuable interpretation processes, heritage resources should be prepared to offer engaging interactions and cultural education, through the mediation of an effective presentation (Rahaman and Tan 2010, 99-100).

Having discussed the characteristics of digital heritage now, it is possible to examine the different aspects that constitute the digital heritage field. Even if it possible to define the field of digital heritage as a comprehensive reality of sub-branches with common practices, theories and challenges, each of those sub-branches possess peculiar characteristics and demands, and consequently should be also analysed as a unique case.

First of all, all applications of digital technology within the field of heritage should be considered, such as the virtual reconstruction and virtual environments of heritage sites and resources, Virtual reality (VR), an augmented reality (AR), digitalise heritage resources (Liritzis *et al.* 2015, 319-325). Digital-born heritage resources should be also taken into account. As defined by J. Ruan and J. McDonough (2009, 746), *'Born-digital heritage is born-digital materials of enduring value that should be kept for future generations.'* To be classified as born-digital, the original and only version of resources created have to be in electronic format.

A second category which should be taken into account includes those institutions which use the digital medium to cope with the dissemination of heritage resources. Institutions such as digital libraries of cultural heritage (DLCH), museums and online catalogues offer this type of services (Tang *et al.* 2018, 59). The main task of these institutions is the administration of digital heritage resources to be offered to users. These services could play an active role in the protection of heritage material in general, and digital heritage in particular, especially in those cases where the physical counterpart is under threat.

What are Digital Libraries of Cultural Heritage (DLCH)?

As it results clear through qualitative analysis, it is possible to retrieve some general characteristics and concepts for digital heritage. On the other hand, even if it is possible - collectively, with more time and resources- to offer a comprehensive discussion on digital heritage and its sub-branches, in this research the digital heritage debate will only be used to better contextualise the information discussed. Therefore, in this research, the attention will be focused on the digital heritage services and specifically over Digital Libraries of Cultural Heritage (DLCH), a sub-group of digital libraries which specifically focalise on the maintenance, sharing and protection of heritage resources in digital format

This choice is connected to the role that these instruments have acquired in Digital Heritage. Indeed, more than other services, digital libraries have played a fundamental role in sharing and protecting cultural heritage resources. Generally, digital libraries and digital collections are terms which erroneously have been used as synonyms, but actually, there is a meaningful difference between the two instruments that they represent. While each collection of digital raw resources could be defined as a digital collection, on the other hand, digital libraries should be conceived "as system that makes digital collections come alive, make them useful accessible, that make them useful for accomplishing work, and that connect them with communities" (Lynch 2002, 135). Through this definition, it is possible to observe that the core function of digital libraries is to be able to offer valid resources to cope with the needs of its reference community. Being community-oriented services, the digital libraries have deployed in these years several techniques and methodologies to best satisfy and serve their users. Taking into account the peculiarity of the demands of each community, it is easy to understand that the solution proposed by a particular digital library cannot always fit for other libraries and consequently with their users' requests.

As it is possible to understand DLCH should be intended only as a part of the field of digital libraries. Even if DLCH could be described through the general definition of digital libraries nevertheless, there are some peculiar characteristics which distinguish these from the other subcategories of DL.

In order to better contextualise the discussion on DLCH it is helpful to take into account the survey conducted by the Council on Library and Information Resources (CLIR) in 2002. Indeed, a valuable opportunity to retrieve the distinctive elements of a DLCH could be found in the list of its mission's core elements of Digital Cultural Heritage Initiative (DCHI) developed during this survey. It is indisputably that DLCH could be identified as a part of the DCHI and in this sense, the libraries should possess the same mission as other DCHI project. As pointed out by Zorich (2003, 12-13), it is possible to trace back four core elements to explain the purpose and the following practices implemented.

As the first point, Zorich (2003) places the commitment to serve the needs of a particular profession or discipline. In this sense, it is possible to observe that this element matches the definition of a digital library offered by Lynch.

The second core commitment is to develop and maintain a digital product. As in the first case, it is still challenging to propose a defined demarcation line which separates DLCH by the other types of DL. On the other hands, in her work Zorich indicates that the resources which should be maintained or developed have to possess a cultural heritage value.

The third mission point deals with the exploration of the digital arena and the promotion of beneficial digital cultural heritage policies. In this sense, the DCHI has the duty to actively analyse and take part in the broadest discussion regarding heritage and its relation with the digital medium. This active participation plays a fundamental role in increasing the quality of the services offered to users. At the same time, the involvement of this institution to the digital heritage discussion could provide the field with important data, reflections and proposals from those who deal with the maintenance and production of heritage resources in their daily businesses.

The final commitment of the DCHI mission is to contribute to the public good. Zorich's (2003) choice to conclude with this final commitment should not come as a surprise, considering that different DCHIs assert, more or less implicitly, a dedication to promoting the public good. The meaning of this commitment could be found in the opportunity for academics to improve their knowledge in and through the cultural heritage resources offered by these initiatives. In the broadest context, this could also be understood as a commitment to maintain and create the digital heritage resources used by specific communities and therefore offering a public service.

Having analysed the main elements which constitute the mission of DCHI projects, it is possible to observe that some distinctive aspects distinguish the DLCH from the other types of DLs. Nevertheless, the concept of DCHI could be too general to offer a qualitative definition of DCHL. In this regards, it could help to analyse the results of the Perseus project[1], discussed by Gregory Crane (2002). These results can offer an inside view and reflections over this peculiar and complex field of Digital Heritage (2002, 630-632).

1. The first point highlighted by Crane (2002) is the necessity to focus attention on the preservation of digital heritage sources. Even if this resolution requires overcoming several theoretical and practical issues,

[1] Perseus, established in 1985, is one of the oldest and most durable examples of online digital libraries. (http://www.perseus.tufts.edu/hopper/)

preservation is a core activity for serviceable digital libraries, and consequently, it is a priority to discuss and develop solutions to overcome those issues.

2. Crane indeed defines access to cultural heritage, and consequently to digital heritage, as a fundamental human right and criticised those policies which conceive access to cultural heritage as a privilege of specific groups. Heritage and its digital counterpart is defined human right and precondition for fostering dialogue and understanding across cultures and civilizations (Shaheed 2011). In this sense, DLCH has to take into account that a plurality of individuals and groups have to be able to relate to their material.
3. Therefore, the third point requires DLCH to be ready to serve the needs of diverse audiences offering tools and customisation options able to deal with the multiple uses which could be done of their resources.
4. The fourth element could be intended as conceptually strictly related to the previous one as it requires that the resources should be able to serve the needs and requests of different audiences[2]. Nevertheless, it is also necessary to conceive that such division it is unable to cope with the demands of who is using the same resources to retrieve information and data to investigate Thucydides speeches, philosophy, and ideas. (Crane and Wulfman 2003, 79)
5. The fifth elements pointed out by Crane is the fact that a DLCH has to be conceived as a laboratory in which reading is the main exercise. Consequently, these services should guarantee different tools for reading and contextualise resources both semantically and in the broadest context.
6. In this sense, Crane places as the sixth element the need to make digital heritage resources and their components reusable by users. Indeed, only allowing access to resources visualisation is not enough to ensure that information is properly acquired and understood. Access and use of resources' related data results in a necessary practice in DLCH to offer users

[2] A valuable example able to better contextualise this concept could be retrieved from the Peloponnesian War of Thucydides case discussed by Crane. This book presented in the classical division in books, chapters, and paragraphs could result useful to the users who aim to acquire historical information from the resources and to analyse the resource as text, source or literature.

the essential instruments to investigate and engage with Digital Heritage resources.

7. Focus the attention on the role played by standards and guidelines within DLCH. Indeed, today, when new standards or guidelines are published, DLCH have the function to adapt their infrastructure and resources to the prescription offered by these documents. In this sense, the changing of resources and their adaptation could cause the loss of valuable information which could play an essential role in the reconstruction of the historical evolution of these.

In this regard, Crane concludes that a DLCH file system which allows multiple versions of the resource to exist at the same time, known as a "versioning system," requires a high level of precision for references and semantic issues, elements which play a crucial role in the exegesis of digital heritage resources. Through these elements, it could be possible to acquire a better notion of the concept of DLCH.

After having discussed the prominent role of preservation and sharing of materials, which are the core elements for all digital libraries. These core concepts will be used as a paradigm for the analysis conducted in the course of this research, as well as a base to develop reflection and conclusion. Nevertheless, DLCH should be not understood as a completely separate category from the rest of DLs. Indeed, DLCHs possess several common elements with the rest of DLs. The DLCH branch of study should take into account the issues and opportunities highlighted in the broadest context of DLs to elaborate and adapt them to its purpose; to increase the quality of the services offered; to discover new opportunities which could help to reach state of the art services. Consequently, it is helpful to investigate how the evolution of the web occurred in the last two decades to fully understand the path of development of DLs; which on the internet have found a supportive instrument and sources to increase their scope and the quality of their services.

Research Questions

Having introduced in the previous paragraphs the main elements which will be discussed in the course of this research it is now necessary to present the questions

which this research project aim to answer. In this sense, my main question will be the following:

- What is the role that new information communication technology (ICT) could play within the field of Digital Heritage, particularly in protecting and sharing digital cultural heritage online?

My sub-questions are:

- What are the main challenges that digital libraries have in relation to security, accessibility, maintenance, protection and property rights?
- How have different professionals responded to these challenges in the specific field of data management online? Which proposes have been raised by academic in order to cope with this challenges?

Methodological Consideration

Before rushing to the main discussion of this research, I will point out some essential elements connected to this research project, specifically the methodologies used. In the first place in the course of this work, a framework based on P2P technologies able to solve some of the main issues connected to the field of DLCH and more generally to that of DL will be proposed. The proposal presented in this research, however, should be considered only as theoretical work. Indeed, even if the considerations and the practices proposed have been critically acquired from different scientific and academic articles, it should be taken into account that the draw up of this research was not anticipated by the creation of a proof of concept for the framework proposed. In this sense, it is impossible to offer direct data over the framework's functioning and public reception. This is essentially due to two major cause. First of all, it was not possible to build up a working proof of concept for the framework proposed in the one year's period in which the research have been developed.

Secondly, to offer a suitable framework for DLCH, it is necessary, even before the creation of a prototype, to set up a debate in which academic, engineers, experts of the field and possible users could discuss, propose, and develop solution to create a well-functioning environment able to respond to the request of different users and communities.

A further element which is necessary to be highlighted is the novelty of the field of study which I decide to investigate. Indeed, DLCH was the result of the massive digitalisation of society happened during the last two decades and compared to the other branches of heritage it is possible to observe a lack of academic discussions and scientific methodologies which could help increase the quality of the debate. Specifically, discussing the application of P2P technology in the context of DLCH it is necessary to conceive that just in the last year's P2P application in the field of DLs has started to acquire importance within the academic context. Therefore, the number of sources and material which have been used to support my proposal could be defined as scarce from a quantitative point of view.

A point which should be clarified is related to the need to talk about the history of the web and its connections to Digital Libraries of Cultural Heritage concepts and initiatives. Indeed, in Chapter 2, I discuss the evolution which has characterised the web from its birth in 1991 up to today. Such discussion, at first glance, could be defined as disconnected from the research which will be conducted. Nevertheless, it should be taken into account the role that internet has acquired in our society and consequently in digital libraries. Moreover, thinking about the scope of the innovations launched through the web it is necessary to offer an analysis of web evolution within this research. A final element which should also be taken into account is the fact that the P2P technologies over which is develop the proposed framework have been created as a consequence of the evolution of the web and to all the issues connected to the development and the actual state of the digital world.

Chapter I: Digital Heritage and Digital Libraries: Definitions and Problems

1. Social Memory, Public Memory and Heritage

To develop a discussion within the field of heritage it is necessarily in the first place to analyse the main elements which characterise the heritage field discussion. Those elements, in addition to facilitating the development of a qualitative definition of heritage and its field, could also help contextualise the discussion which will be established in the next chapter.

Even if the main goal of this thesis is to analyse the role that ICT technology could play within the field of Digital Heritage and specifically for the DLCH, the analysis conducted in this chapter is of primary importance. Indeed, aiming to simplify the discussion, it is possible to affirm that Digital Heritage has to be understood as the application of ICT technology within the field of heritage. But what does heritage mean? And which are the main concepts which should be assimilated before being able to develop a valuable discussion on Heritage?

In order to offer a valuable answer to these questions, an analysis of the concept of memory and its role within society could be of help. Memory has always played a primary role in human societies, allowing individuals as well as communities to trace and strengthen their relation with the past and their own history. As pointed out by S. French (1995, 10), several historians have defined a fundamental distinction between their field of study and memory: mainly, history is framed as a discipline built over evidence, while memory is defined as a malleable guide to the past. While I agree over the discipline foundation of history, I argue against their division of history from memory. I argue that history is a memory which could be defined as authorised and official, and consequently which has acquired a prominent role in the discussion over the past. Nevertheless, as folklore, oral history, and biographic works etc., history should be approached as a branch of memory no less malleable than the other types of memory. Among the different branches of memory which could be investigated in this paragraph, the attention will be the focus over two specific types of memory: social memory and public memory.

The first element which should be discussed is necessarily social memory. This type of memory plays a fundamental role in human societies and in a peculiar way it is possible to see it as the foundation of social and community relationships. As discussed by Nietzsche, in the Animalia Kingdom the survival of the species is written on the genome of every single animal which instinctively puts into practice this genetic guideline. In the human context, the survival of the species also requires to investigate the meaning of human nature to maintain a legacy for future generations. As observed by Jan Assmann and John Czaplicka (1995, 126) to fulfil this requirement, humans have relied on social memory, seen as *"a collective concept for all knowledge that directs behaviour and experience in the interactive framework of society and one that obtains through generations in repeated societal practice and initiation"*. Therefore, social memory is a type of memory developed by non-specialist communities to serve the necessities of their societies and offer a set of knowledge and traditions as a legacy for future generations. Social memory does not necessarily need to be related to the public sphere of society, but it could also be developed within private groups as families, clan, etc., which possess private memory shared only within group's members (Biesecker *et al.* 2004, 22). Moreover, through social memory, those communities not only propose what should be remembered but also how it should be remembered (Fentress et al. 1992, 36-39). Differently from history, social memory refuses to pursue the 'objectivity myth,' offering space to multiple perceptions of the past (French 1995, 16).

The next type of memory which requires identification and analys is the public memory. Already through the terminology, it is possible to conceive that this type of memory is strictly linked to the public sphere. Public is an evolution of the Latin term 'Publicus' used to indicate something which is shared by each citizen and everything that could have an impact over the population of the administrative-political organization. An interesting analysis and definition of the Public sphere has been offered by Edward S. Casey who writes that *"public memory is radically bivalent in its temporality. Where other modes of remembering deal primarily with the past—with the notable exceptions of recognition and reminding -public memory is both attached to past and acts to ensure a future of further remembering of that same event"* (in Biesecker *et al.* 2004, 2). Working actively on the public sphere, the scope of the public memory is not limited to the past but it actually influences and interacts with the present and also shapes the way in which is conceive the

future. Public memory to be defined as such needs to be shared not only by a community but rather by an entire population. Through the analysis of several authors, Sara McDowell (2016, 40) was able to formulate a qualitative definition of public memory: *"public memory emerges from the intersection of official and vernacular cultural expressions…. A reflection of present political and social relationships... A fluid process that is not only negotiated by official or national groups but also by the media, academics, heritage institutions and local community organizations."* As it is possible to understand, public memory plays a fundamental role in society, indeed public memory tries to be accurate and to be able to conform to the exigence of society as a useful tool, a guiding force for society. In this context, Casey identifies five fundamental elements over which public memory is built (Biesecker *et al.* 2004, 32-36). The first element is the public place, addressed to all population public memory could not be developed within a private or a semi-private place, prerogative of a specific community. Secondly, public memory requires necessarily the presence of a public, a group of people brought together by a common purpose. The third element which plays a fundamental role is the public discussion which should be created. Indeed, after being reunited in a common place, people have to start a discussion and communicate over the specific object of public memory to interiorize it and be able to acquire the shared feeling and emotion from the community experience. Certainly, all these elements even if fundamental, are not enough to build public memory. In this sense, the fourth element is the presence of a common topic to be discussed. It does not mean that people must agree or to change their ideas, but rather to discuss a topic collectively while participants share their interest and different ideas for a particular topic. The last element which Casey identifies as fundamental for public memory is the commemoration in place. Indeed, it is required that people commemorate in the sense of remembering together, as part of a broader community. In this sense, even if one might not count on a direct relationship with the other members, he or she can feel to be part of the group thanks to the common purpose, to remember. As described by Tomislav Šola (2015, 44): *"The moment the need for passing on a certain experience to others was turned into an organized effort of the group or community- public memory was born. It was aimed at sharing and influence, trans-generational by character and contributing to survival and advancement of the community,"*

Throughout these two sections, it is possible to observe the prominent role played by memory within society. Memory does not only concern the relationship with the past, but it also has the power to dramatically influences the present and the shaping of the future. Being a crucial factor for the well-being of society and having acquired a prominent role in the academic debate, it was required to ensure the protection and access to memory to all members of communities. Furthermore, it is also needed to establish a relevant discussion to develop solutions and debates to increase the quality of memory and that of the related services. In this context, it is settled the birth of Heritage as the field of study as well of resources tied to memory processes.

In order to develop a valuable analysis of the concept of heritage it is helpful to take into account the definition offered by UNESCO in its document Draft Medium-Term Plan 1990-1995:

> *The cultural heritage may be defined as the entire corpus of material signs - either artistic or symbolic - handed on by the past to each culture and, therefore, to the whole of humankind. As a constituent part of the affirmation and enrichment of cultural identities, as a legacy belonging to all humankind, the cultural heritage gives each particular place its recognisable features and is the storehouse of human experience. The preservation and the presentation of the cultural heritage are therefore a corner-stone of any cultural policy* (Jokilehto 2005, 4-5).

Heritage resources are therefore all those resources which for a specific community are embodied or have acquired the memory value. Humankind heritage resources have to be protected and maintained to be usable and useful for those who throughout these resources could interact with memory and acquire knowledge. Indeed, the final task of heritage resources is to communicate the essence of human experience to its consumers (Šola 2015, 165). Therefore, the definition of heritage resources is strictly connected to the process related to them. Through consumption, interactions, modifications and negotiation, heritage resources can explicate their core function and acquire meaning and relevance as well as an actual meaning within society (McDowell 2016, 49).

Taking into account that heritage's primary objective is to communicate the essence of human experience to consumers, specialists have to play the role of mediators within this communication, by simplifying the fruition, promoting the dialogue and increasing the scope of the opportunity and transforming the knowledge in wisdom (Šola 2015, 215). Indeed, Tomislav Šola (2015) indicates that specialists have to study the nature of memory, identity and heritage and their relations the public need. The creation and development of a science of memory are indeed one of the main preoccupations of Šola (2015), who in his book *Mnemosophy* establishes a scientific theory for heritage. Such an approach which is defined by Šola as "Mnemosophy" is indeed what has generally been defined as heritology, museology or museography.

Nevertheless, Šola prefers to use the term Mnemosophy because it could offer a better understanding of the actual meaning and role of this theory. In this sense, Mnemosophy is defined as the theory related to the qualitative process of transforming the memory (Mneme) into wisdom (Sophia) by creating in this sense wisdom of memory and promoting a wise interaction and use of memory (Šola 2015, 215-216). In addition to offering a theoretical foundation for the science of heritage, Šola indicates three key processes which constitute the core elements related to the practice of heritage and generally, for all those institutions which deal with heritage. These practices defined as the 3C consist of the processes of collection, care, and communication of public memory. Such theory, as well as the methods proposed, can embrace the sectors of museums, archive and libraries. In this sense, these could be ascribed within a unique complementary area of expertise in which the practice of the different sectors could be maintained, discussed and improved. This new reality of heritage initiatives has been defined by Šola as the Total Museum (Šola 2015, 18).

2. Heritage and the Digital World

It is interesting that Paul F. Marty has echoed the same consideration proposed by Šola (2015) in the context of the total museum. Specifically, within the context of the debate developed during the Cultural Heritage Information Professionals workshop in 2008, Marty discusses a process defined as digital convergence. This convergence which was caused by the digitalisation process of society and the

consequently increasingly use and reliance on digital resources was able to blur the classical division between different methods of information organisation and therefore libraries, museums and archives (Marty 2009, 295). Such convergence has as primary objective to encourage a multidisciplinary discussion in which professionals from the three discipline could collaborate and combine their strategies, idea and frameworks to increase the quality of the services offered to users. Throughout the creation of a joint discussion for these institutions of heritage, it could be possible to establish a collective project in which each specialist and the relative institution could be engaged in a collaboration to "build a single vision of the future of globally networked data" (2009, 297).

In a later article, Marty (year) further developed and enriched this discussion by exploring the problems related to concretisation of the process of convergence. As he pinpoints, the convergence process was already presented by W. Rayward (year), who at the end of the nineties was affirming how the evolution of society will make the distinction between museums, libraries and archive insensate and irrelevant. Rayward's discussion is contextualised within the peak period of digitalisation of society, and the new-born web was revolutionising the concept of information and communication. Even if Marty generally agreed with Rayward's words, he does not show the same trust on the fact that a spontaneous process of convergence might happen.

Indeed, more than fifteen years after Rayward's proposal, Marty (year) highlights that the discussion on the convergence not only is still active but also that issues and doubts have been raised by academics and specialists. These issues regard the level of convergence that these institutions have to reach, as institutions have different methodologies and approaches to their resources. In this sense, de-contextualising resources from their ecosystems and methods could cause a diminution on the quality of the resources' presentation. Therefore, Marty (2014) supports the idea of a plurality of methodologies and institutions established to overcome specific issues and to respond to particular needs. On the other hand, he also remarks that this division is related to the backstage of information presentation and communication. On the front-end, the situation should be conceived as entirely different because most of the users have no interest on the processes related to the presentation of the information and the processes associated to it (Marty 2014, 618).

Therefore, Marty proposes to develop a convergence in which the traditional distinction between libraries, museums, and archives will be maintained for the process related to the back end. On the other hand, those processes should be routed to a shared service provider, a platform in which the access to different types of resources is ensured. To put in practice this project Marty recommends to establish a discussion in which convergence is examined, taking into account both internal and external perspectives and needs to find the most suitable solution (Marty 2014, 624).

A further powerful argumentation has been offered by David Bearman (1995) in his article regarding the necessity to standardise cultural heritage within the networked reality. In is article after having highlighted the keen interest which cultural heritage specialists have shown for the web since its birth, Bearman criticises the lack of standardisation for digital cultural heritage resources. Offering common standardised approaches, methods and systems represents for Berman the only solution to cope with the necessity of the users and the lack of interest both for the back-end operation and for the peculiars differences which characterise resources which belong to various fields of heritage. In this regards, Bearman affirms that to succeed and survive in the twenty-first-century heritage institutions should be able to provide *"easy, one-stop electronic access to their collections and programs"* (Bearman 1995, 281). Discussing standardisation Bearman pinpoints the fact that, in those years, the discussion over standardisation was already active and that an essential number of solutions were proposed.

Nevertheless, little progress was made on the practical context. For Bearman (1995), this situation results incredibly inefficient, relying on standards not only could increase the users' experience but also simplify the exchange of information and resources for both institutions and academics. Though Bearman's (1995) strategies might be seen as outdated and redundant in today's digital landscape, on the other hand, the concepts and ideas proposed in this article could offer food for thought regarding the value of standardisation within heritage field and the related convergence in the digital heritage field. It is interesting that Bearman highlights the influence of the web within the discussion and process of standardisation. Indeed, the web should be conceived not only as a valid instrument at the service of heritage field but as a source of inspiration on the discussion over the evolution of heritage institutions and for the development of new heritage

theories and methodologies. Such elements were well described by Šola (2015, 57), who while affirming that we live in a consolidated memory structure, remarks on the fact that *"Very recently, even in historical measures of time, this structure has been enveloped by the new, obvious but immaterial reality in its magnificent omnipresence, - the world wide web. It creates a global memory environment consisting of man-made electric impulses & social actions, forming a pulsating shell of a giant, primitive, hyperemesis brain."*

3. Digital Cultural Heritage Initiative

Even if the role played by the web is valuable to understand the evolution of digital heritage and its related institutions, it is also necessary to analyse the sectors in which digital heritage institutions could be divided and consequently the scholarship areas that constitute the field of digital heritage. First of all, it is important to exclude from this analysis those institutions and digital practices which even if connected to the branch of heritage should nevertheless, be inscribed within other fields of study. In this sense, projects as the digital applications in Archaeology, Archaeometry, Humanities are excluded as these are not directly related to the heritage area.

Discussing the digital convergence, the authors highlight the primary role played by libraries, archives and museums. In the last years, these institutions have primarily engaged in digitalisation to increase the scope of services offered to users as well to widen the scope of preservation and maintenance operations. In this sense, it was possible to assist to the establishment of a series of services as digital museum platforms as well as digital repositories like libraries, collections and archives. Even if often the establishment of these digital repositories is connected to physical institutions, it is also essential to take into account that even a large number of a digital-born institution were developed in these years.

Aiming to offer a presentation of the various branches which constitute the heritage field, it is useful to rely on the categorisation offered by Zorich (2003, 13-15) in the division of the digital heritage initiatives. In this sense in addition to the digital libraries and portal already, digital heritage also includes others types of institution and services like e-publishing initiatives which offer materials together with digital tools and functions which could not transfer to the physical copies. Proceeding it has to be included the group of educational and scholarly databases

which offer a set of resources already organised and optimised to be used for compilation, research and access. A further category of digital heritage initiatives is constituted by the references databases which offer to users powerful searching engine and tools specifically developed to deal with resources metadata. Finally, it should also be considered that those software tools specifically created to work with heritage resources as well as those additional resources like guidelines, standards, publications and all the other valuable materials develop to increase the quality of practice and the science behind digital heritage.

As it is possible to observe, the field of digital heritage includes in itself a broader range of services, platforms and tools. Each one of the subcategories shows peculiar characteristics, methodologies and theories which require a stand-alone analysis. Therefore, in this research the focus is specifically addressed to the digital libraries (DLs), the *"distinct types of digital information (for example, databases, Web sites, teaching resources) brought together in a product that, to the user, appears seamless and unified"* (Zorich 2003, 13). Again the concept of digital convergence acquires a central role in the discussion of the digital heritage institution, and correctly it should be understood as an adaptive necessity to conceive this category as a unified branch in which a series of common issues and practices could be retrieved.

To discuss this category, it is necessary to define the concept of digital library and consequently its relation with digital heritage. Simplistically, digital libraries could be defined as databases or information retrieval systems (Borgman 1999, 231). A more explanatory definition is offered by H. Lynch and H. Garcia-Molina (1995, 89) who state that *"Digital libraries were viewed as systems providing a community of users with coherent access to a large, organized repository of information and knowledge… The ability of the user to access, reorganize, and utilize this repository is enriched by the capabilities of digital technology."* Such a definition helps to better contextualise the concept of digital libraries and the peculiar characteristic of these databases.

The first element of differentiation is the role that each digital library plays in serving a community of users and their specific needs. Not only such characteristic distinguishes digital libraries from other databases but it also implies that particular strategies and methodologies have to be applied to serve each community. Thus, it is clear how it could even exist an enormous difference

between the appearance and functionalities offered by these information retrieval systems (Fox et al. 1995, 28; Lynch 2002, 138). Furthermore, it is also necessary to discuss the practice related to the objects offers in this database. Indeed, as highlighted by Barry Leinder (1998), from collection processes through the presentation of resources, digital libraries' core functions are focused on providing an increased number of resources and on the organisation and adaptation of those to fulfil their users' requirements. In this regard, the task to offer tools and direct and indirect support to their users should be understood as a further distinctive element of digital libraries.

Such support has not only to be able to cope with the direct request of users and consequently of human, but it also has to be able to serve automatise processes which have to assist users' requests (IFLA 2018, 1-2). Through these elements, it is possible to understand that there is a meaningful difference between a common database and a digital library. To be defined as a digital library it is necessary that within the database a series of processes are carried out (collections, maintenance, presentation and preservation) to offer a set of resources and services able to satisfy the needs of a specific community of users (Arms 2001, 209-2012; Lesk 2005, 29-30). Having retrieved a definition of digital library, we can discuss the relation between digital libraries and digital heritage. It is possible to state the existence of a digital libraries subcategory which have been established specifically to work with digital heritage resources, the digital libraries of cultural heritage (DLCH).

A further interesting example could be found in DELOS Network of Excellence established to guarantee access to the citizens of the European Union to heritage and cultural resources on digital format. Even if DELOS web page was frozen in 2009, nevertheless, DELOS has acquired a prominent role in the context of digital libraries thanks to its publication in February 2008 of the document A Reference Model for Digital Libraries Management Systems (Candela *et al.* 2008) which has influenced the evolution of the digital libraries debate. Remaining in the European context, it is interesting to take into account the Europeana project, a collective digital library in which have been shared cultural resource offered by various institutions of the member states of the European Union. Counting at its actual state more than 58 million resources (document, pictures, audio files, videos, and 3D models), Europeana is surely a grandiose project. In the North American context, it is worth discussing the American Memory project. Launched in 1990 as

digitalisation and diffusion project of the cultural material preserved in the Library of Congress, American Memory was initially based on the use of CD-ROMs to share contents among the public. In 1994, thanks to the prominent role which the web was acquiring in society, the project underwent a complete overhaul to be used as the leading project of the new-born National Digital Library program (American Memory 2018). American Memory was transformed into a web-based library which to this day contains more than one hundred thirty collections containing resources of USA heritage and memory. Finally, a further example which should be taken into account is the World Digital Library. Developed in a joint project between UNESCO and the Library of Congress, the World Digital Library is an open-access library built on the collaboration of 193 countries which have decided to offer free access to cultural and historical materials through a common platform. As explained by the project creator James Billington, the primary objective of the establishment of this international structure was to bring *'people closer together precisely by celebrating the depth and uniqueness of different cultures in a single global undertaking'* (Flood 2009). These examples can competently represent the category of DLCH constituted by digital library focused on protecting and sharing digital heritage resources and on offering tools, software, presentation to serve the needs of different communities (Crane 2002, 632-633).

Taking into account the existence of a subcategory of DLs specifically addressed on Cultural Heritage it is necessary to discuss the relationship between the broader group of digital libraries and the field of heritage. As digital libraries have also been built to organise, protect and share material not in possession of particular cultural, social or heritage values it is hastened to define a clear and close link between digital libraries and digital heritage. Nevertheless, the DL field acquires a fundamental role in the digital heritage context at the moment in which it wants to investigate the evolution process and to research possible solutions to increase the quality of service offer by these institution As a subcategory of digital libraries, DLCH is a field strongly influence by broadest evolution in the area of DLs. Furthermore, take into account the discussion developed by Diane M. Zorich (year) it is clear that heritage resources and processes are not only prerogatives of the DLCH and that a broader range of institutions are involved in digital heritage management. Therefore, limiting the scope of this research only to the DLCH branch could result in a counter-productive strategy which framed in classical

categorisation would be unable to fully embrace the complexity of DCHI. In this regard, the process of digital convergence which in the last twenty-five years has shaped the digital heritage sector, results favourable to establish a comprehensive discussion in which the evolution and problematic of Digital Libraries (DLs) as macro sector could be analysed and discussed.

Chapter II: Digital libraries history and the evolution of the web

1. Heritage resources in digital environments

The main objective of this chapter is to retrace the theoretical and technical evolution that the field of digital libraries has been undergoing from the mid-nineties to offer a coherent picture of the context in which the problems analysed in the following chapters were produced. To this end, the following research begins by analysing the historical background of digital libraries.

In the *Charter on The Preservation of Digital Heritage* composed by UNESCO in 2003 was stated that '*the digital heritage consists of unique resources of human knowledge and expression. It embraces cultural, educational, scientific and administrative resources, as well as technical, legal, medical and other kinds of information created digitally, or converted into digital form from existing analogue resources. Where resources are 'born digital', there is no other format but the digital object'.* (UNESCO 2003, 75). This section aims to affirm the intrinsic value that information stored in digital format comes to acquire as 'tangible' proof of human culture.

The impact of ICT within our society has been so massive and comprehensive that it can be considered a sort of revolution. Indeed, due to the digitalisation process, the access, consultation and even interaction with human knowledge have undergone complete overhauls in the last two decades.

In order to offer insight into this evolution, it is necessary to start the analysis from the nineties, a period in which Western society began an extensive process of digitisation. In order to retrace this evolution, it is analysed in this research via a series of relevant works and articles. To this end, this chapter traces the processes involved in digital heritage and those that have transformed it from a simple pioneering application to an academic branch.

In order to trace the evolution of digital libraries field from the nineties, it is necessary to consider the technological limits of the time. Less powerful systems imply a proportional increase in copy production time, and at the same time, an increased risk of file compromise. Second, it is further necessary to consider the

costs of both hardware and software, both of which suffered a fall in prices inversely proportional to the increase in technological progress. [3]

Furthermore, it is necessary to take into account the fact that, while the first web browser, the World Wide Web had already seen the light in 1991, and several browsers and portal were available in the nineties, at that time, web structure was characterised by static resources and cumbersome procedures for sharing and acquiring them.

On the other hand, the results and the potential suggested by these innovative tools resulted in these successful technologies being incorporated into the pre-existing networks of companies and institutions in the mid-nineties. It is since 1998 that we can start to conceive of the web, its fruition and consequently, the networks in more modern terms.[4]

Another fundamental factor involves the digitisation of cultural heritage, which, during the nineties, was progressively introduced within most of the institutes of industrialised countries. The corresponding reduction of costs allowed the digitisation to respond to needs and issues, such as the natural decay of paper and comprised an alternative solution to the continuous accumulation of paper material within the archives. The contained costs and the storage capacity made CD-ROMs the most common tool to store data in those years.

Twenty years later, it may be incomprehensible to conceive of the difficulties involved in the copying process. As such, I consider it necessary to dwell on the specific characteristics of the reference framework.

In 1997, Margaret Hedstrom (Hedstrom 1997) denounced the deterioration process to which the hardware was subjected. Indeed, the most common storage drives were subject to a high degree of deterioration. Although storage drives such as optical disks were able to ensure longer life cycles, these were too costly to be widely used. Through recognising the limits involved in preserving the hardware, Hedstrom focused on analysing the potential of the software and the main characteristic of digital files, namely, the natural predisposition to copying processes. In fact, an infinite number of copies can potentially be produced from

[3] A striking example can be found in the $/GB ratio in the history of storage cost. In 1998, one GB of memory was worth almost $100, while in 2008, the price fell under $0.40 per GB. (Komorowski 2009)

[4] Mozila.org and the Google search engine came to light.

every file. In this sense, digital preservation operations have to depend upon a continuing copying process rather than on the survival of the physical storages:

'Librarians and archivists must prepare for reformatting as a regular step in information management' (Lesk, 2000).

In this regard, Hedstrom continued her analysis by focussing on refresh and migration, vital operations for preserving information stored in copied files. While refreshing ensures that the file does not contain any kind of data degradation in the copying process, migration responds to the need for compatibility between different formats (e.g. from .docx to .pdf). The difficulty involved in reducing the expenses related to hardware increased the influence of the standardisation process, the efforts of which aimed to simplify and reduce the costs of the copying process. The standardisation of digital systems substantialy improved communication as well as data sharing between different institutions.

The first years of 2000 marked a turning point in the ICT context. The impact of Napster and its P2P technology generated a considerable drop in the recording market. The scope of this event challenged concepts like privacy and ownership within the digital context.

The sudden changes taking place in the digital world at the turn of the millennium did not allow for easy data preservation, while in the digital libraries context, issues such as integrity and authenticity were raised (Barr *et al.* 2003). The intrinsic ability to easily copy any file suddenly became a problem, especially since this severely undermined the concept of intellectual property, as well as the value and authenticity of the information contained.

This context certainly influenced UNESCO's decision to publish the Charter on the Preservation of Digital Heritage in 2003. Among the most important tasks associated with the preservation of digital heritage, UNESCO underlined the importance of maintaining access to digital material. The nineties had, in fact, witnessed the birth and death of different operating systems, including the massive transition to graphic operating systems (e.g. MacOS, Windows, Ubuntu). Today, every operating system is characterised by the use of its own mark-up language. In this sense, the files produced by a machine could be unreadable by other models of machines that could not guarantee access. This issue deepens when the machine and the technologies used to create the file have long since died. As a result, the information stored inside files produced by an outdated machine could be lost

forever. To respond to this lack of compatibility between different rendering environments, several software was created to enable emulating the environments of origin. Through the creation of a virtual machine, the hardware and software functions used by the file system were recreated. In this way, emulation allows the restoration of access to information produced in obsolete systems.

To overcome the differences between the various systems, the creation of a Universal Virtual Computer (UVC) designed for archival purposes was proposed. This machine would emulate and migrate files in a simpler and more neutral format and therefore possess greater compatibility with the various operating systems.

Although several experiments were run in 2000 that allowed observing the implementation of a UVC and their effectiveness in faithfully recreating the processed files, at the same time, its inability to recreate the functions associated with these files became clear (Lorie and van Diessen 2005), create the original functions of the file, basically reproducing a static version aimed only at reading the contents. In this sense, the operations carried out by the UVC cannot be properly defined as emulation, because these machines are not able to recreate the complexity of the functions that characterised the environments of origin and the markup language used.

While the issue of long-time preservation came to hold a central role in digital libraries, new opportunities for interaction with the digital environment were also brought to light.

More specifically, the concept of metadata came to acquire a new role in light of the incessant process of digitisation that information was undergoing in those years. Because the creation of digital files involves the consequent development of a series of information (place, date and creator) within these files, this can play a decisive role in the analysis of the information contained within. The value that these can acquire is directly proportional to the time elapsed since the file was born. As W. L. Anderson clarified, metadata plays a central role in long-term preservation, as they respond to the need to provide the studied data with a valid, reliable and historical context (Anderson 2004). At the same time as clarifying how data access and preservation became an integral part of scientific practices, Anderson highlighted how the interdependence relations among traditional scientific disciplines were spreading among all academic branches, including archaeology and cultural heritage.

Despite having conceived the value that information digitalisation could possess within such an interdependent contest, Anderson could not imagine how an extensive process such as digitalisation would be applied and how ICTs would reach a predominant role within society due to a drastic change in the common concept of the web.

As a matter of fact, the same year that Anderson's paper was published was also characterised by certain events that, at a distance of almost fifteen years, can already be defined as of historical value. In February 2004, the first version of Facebook was launched at Harvard. In August, the world witnessed Google's IPO[5] symbol sudden raise the company in the web environment. Finally, by November, the first version of the open-source Firefox browser was launched.[6]

These events would lead to a progressive improvement in the tools available to the user within the web in the second half of 2000, and at the same time, to the birth of a new concept of interaction within digital platforms. However, this evolution did not bring any conclusive resolution to the elaborated issues concerning digital heritage and, in general, for concepts such as copyright protection, compatibility, accessibility and so on.

These issues and their application in the field of DLs formed the basis for the analysis conducted by Dr Yannis Ioannidis (Ioannidis 2005). He noted that it was significant how the evolution that was taking place within the digital world was also leading to a mutation of digital libraries. During this time, DLs began to be tightened in relation to the evolution these had undergone. Once composed mainly of text files and image files, they instead came to acquire new multimedia files, such as audio and video files, thanks to the new opportunity offered by Broadband (Savage and Waldman 2005, 216-217) which in those years was spreading around the world allowing in 2005 to witnessed the birth of YouTube (Burgess and Green 2018, 15-18). In this sense, these platforms acquired the role of digital counterpart even for museums and archives.

In the same years, referring to the context of computational sciences applied to museology, it is interesting to observe the considerations raised by Ross Parry in his article (Parry 2005). He noted that even if the museum computation was a

[5] Google went public on NASDAQ with a market cap of $ 23 billion.

[6] It is not considered necessary to discuss the value of these events because their repercussions are fully noticeable even today.

relatively dated field, it lacked interest, which was, in theory, a necessary element for both transmitting knowledges between generations and standardising the applications. Moreover, it is interesting how, basing his analysis over the total museum concept introduced by Professor Tomislav Šola, Parry re-proposed the idea of a platform intended as an information-space museum at a time when ICTs were drastically changing society. Although the so-called Wikis were already acquiring a central role in society at the time, these free-distributed open-source archives based on free sharing by users did not possess a high level of reliability. Parry also specified how, in his historical moment, the museum computation refers and acquire theoretical practices and approaches developed and established within other disciplinary contexts was able to develop its own theoretical approach that would be fundamental for further advancing the state of the art of museum environments. Indeed, the development of theorisation had not only led to the birth of the so-called digital cultural heritage intended as a specific area of research but also enabled new opportunities of collaboration between different research groups without the need to be in the same place. In conclusion, Parry stated how this evolution within the museum computation (now digital heritage) laid the foundations for an analytical approach aimed at examining the profound impositions that regulate the relationship between the museum and its public when, through the web, conceptual limits such as space and time were, for the first time, overcome.

Proceeding, it is interesting the work Santana-Quintero and Addison 2007. Aware of the enormous steps and goals that this branch had achieved in recent years, the authors focus on what could be done to protect and make accessible the enormous amount of information stored in digital format. Dr Addison noted how, in the current state of things, when seeking to perform a study, it was easier to create new data from scratch to employ information collected previously. Claiming that issues such as copyright, the lack of a standardised metadata system and the high costs of data recovery made it difficult to use the information already collected in the article was highlight that this situation, as opposed to most, was based on characteristics of digital heritage resources and the digital world in general (such as availability, reliability, access, sharing, etc.), and explained how much more needed to be done before these problems could be overcome. Among the advanced proposals, one, in particular, has attracted my attention, namely, the introduction of copyleft within academic production.

This license, based on and modifying the copyright (as well as its name), ensured free distribution and modification of the works protected by this license, as well as the future versions created.

As already noted by Parry, the field of digital heritage and specifically that of Digital Libraries needs to critically observe the digital environment's evolution in order to be able to acquire those practices that may turn out to be successful. For instance, copyleft played an important role, especially in the field of software (exemplary is the release of the GNU General Public License in 2007) and in their evolution in those years.

The 2007 results a decisive year in the evolution of the digital world, as this year witnessed the release of the first version of the iPhone, and the subsequent birth of the smartphone concept. It seems superfluous to analyse the historical significance of this event, and the influence that the smartphone as a tool would go on to play in our society over the next ten years.

The second decades of the twenty-first century would become characterised by the widespread use of smartphones, further increasing the digital influence on society.[7] Also in those years, several sector-leading companies began selling their Internet infrastructure through services called Infrastructure as a Service (IaaS). These enabled users to remotely exploit the services (server, DNS, computing, etc.) and to combine and scale them according to their requirements. The externalisation solution allowed companies and institutions to exploit these tools for creating a private network, helping them to save considerably on hardware costs and maintenance. In September 2011, this led the National Institute of Standards and Technology (NIST) to publish the 'Definition of Cloud Computing' (Mell and Grace 2011) which identified and demonstrated the main features of the cloud computing model. I find it interesting that this technology immediately exhibited an aptitude for becoming part of the DLs contest. Indeed, in the same year, Rupesh Sanchati and Gaurav Kulkarni demonstrated the potential of the cloud computing application within the field of DLs. Focussing their attention on university libraries as the cornerstone of academic study and scientific progress, the two academics noted how the functions of such platforms were no longer limited to simply preserving and sharing information. Having to meet the requirements of their users,

[7] ARMA International estimates that upwards of 90% of records created in 2009 were digital (Kirschenbaum *et al.* 2010, 2)

these platforms needed to provide 'appropriate, comprehensive and multi-level services for its users' (search engine, e-mail system, transfer protocol, etc.).
Their paper (Sanchati 2011, 38) primarily referred to applications in the university field. However, it is possible to observe that these formats can also apply to other types of digital archives. Authors have discussed how the quality of services offered by libraries can be significantly improved through the implementation of cloud computing, while the costs of maintaining these can be reduced at the same time. Sanchati and Kulkarni specifically noted how the application of this technology could be used to create a public cloud shared between universities. This solution could greatly simplify communications between the institutes while at the same time improving the amount of information that can be consulted by each user. In fact, although there are services such as the Open Public Access Catalogue (OPAC) and the Inter-Library Loan (ILL) aimed at providing access to DL services, the adoption of cloud computing in the university context would allow the creation of a unique access platform for digital heritage (Sanchati 2011, 39-40). Although the main aim of this application was to improve the quality of the services offered by university DLs, and consequently, improve academic research, it is also distinguished by its important ethical impact. For instance, the emergence of a shared cloud between universities could indeed provide an effective solution to cope with economic inequality between regions that can invariably influence the quality of services offered by academic institutions in less wealthy regions.

Equally interesting are the considerations regarding the application of digital forensics in the context of digital libraries and collections of cultural heritage raised by Matthew G. Kirschenbaum during the symposium *Digital Forensics and Born-Digital Content in Cultural Heritage Collections* held in 2010 by the University of Maryland. This interdisciplinary proposal (Kirschenbaum *et al.* 2010) arose from the moment in which both fields of study were distinguished by a strongly agnostic point of view and attention to themes such as preserving, processing, and interpreting collected data in order to make them available to the audience. Although, based on conceptually similar practices, these two fields have, for a long time, been seen as distant from each other. Obviously, digital forensics, as science focused on the recovery and analysis of digital material connected to crimes had long since acquired a central role in our society, and as a result, different technologies and tools had been developed within it. In this regard, the author noted

the peculiar characteristics of the technologies and tools developed in this sector, which exhibit high levels of adaptability with the context of DLs and its methodologies. Kirschenbaum's considerations were not only limited to the opportunities created by an interdisciplinary strategy but also aimed to analyse the role that a field like digital heritage acquires in the society at the moment in which most of the records are created in digital format. Nevertheless, this situation is not exempt from challenges such as protection, accessibility and sharing. Tools, as well as techniques offered by digital forensics, can provide a valid contribution to improving preservation from a technical point of view, but when analysing issues such as reliability, authenticity and general ethical issues, it became clear that these tools alone may not be sufficient to provide a decisive and compressive approach. However, it is interesting to note that Kirshenbaum also proposed a collaborative strategy between the institutes. In fact, due to the high levels that the digital forensics technologies have achieved, his article foresaw administrative cooperation aimed at creating regional services shared between institutes. Moreover, attention was also drawn to the need to impose common policies and strategies within the digital heritage in order to standardise digital platforms and simplify the communication between the institutes, thus improving the quality of their services.

This was mainly due to the predominant role that digitalisation had acquired in those years, which pervaded and drastically influenced the work, study, free time and sociality of individuals. In this context, it is understandable that issues such as privacy, security, intellectual property and protection not only assumed a central role within the academic context and public opinion but even underwent a profound change in their meaning. Such changes required solutions to be proposed and implemented that differed from those already explored, as well as real education and awareness of certain issues occurring in an interconnected world.

2. Digitalisation and sharing of heritage

In order to proceed, it might be necessary to contextualise the information analysed up to now regarding the evolution process that digital heritage has undergone over the years. This could provide a valid solution to establish a direct parallel between the contemporary evolution of the web and its intrinsic dependence on the digital heritage of this environment.

Wanting to discuss the processes of digitalisation of cultural heritage, as well as the protection and sharing of those resources, one cannot help but examine DLs evolution.

Having to deal with a shared reality and make its platforms accessible to their users even outside the confines of the local network, it is essential to rely on common languages and structures, such as those proposed by the World Wide Web Consortium (W3C),[8] the international organisation for web standards. These offer the best solution for standardisation within the broadest context, as well as the most effective way to be integrated within the web structure and exploit web resources and tools.

In order to provide a reliable explanation concerning the evolution that has characterised the web from its birth until the present day, the classic division has been applied in three versions (1.0, 2.0 and 3.0) to determine the various phases through which the web has evolved. At the same time, I consider it necessary to point out how this categorisation of the web is not connected to the linear evolution of the instrument in which one form of the web replaces the older version, but rather as an integrated social system in which the different stages of the web, and consequently their application, comes to coexist (Barrasi and Treré 2012, 1274).

2.1. Web 1.0

This first section discusses the web of documents, or Web 1.0, in order to identify the period in which the web was based on a hierarchical organisation of information. The sites were characterised by navigation based on menus, where the user had no chance of interacting with the web pages, but could instead only visualise and acquire content. Communication was unidirectional and based on mere visualisation of the required resource (Choudhury 2014, 8096-8097)

The first ADSL standard was published in 1998 (Ansi 1998), so it is necessary to take into account the speed limits that characterised the web for most of its first decade of life. Efficiently uploading and downloading resources and formats (audio, video, etc.) with large files was not supported, and so small, static files (documents, images, photos, etc.) were the most common resources of Web 1.0. As such, I consider it appropriate to refer to the description offered by Shivalingaiah

[8] https://www.w3.org

and Naik: ‘Web 1.0 is a system of interlinked, hypertext documents accessed via the Internet’. (Shivalingaiah and Naik 2008)

2.1.1. Digital Library 1.0

Since the birth of the web in 1991, digital heritage specialists have, rather pioneering, decided to provide new services based on this new means of communication. To this end, ArXiv.org[9] was created in that same year at Cornell University, intended to provide a portal through which for the first-time scientific resources could be shared in e-print versions. The innovations promoted by Arxiv formed the basis for the creation of various archives, both institutional and non-institutional, during the first half of the nineties. These portals were characterised by a menu structure that strongly limited the communication between the site and its users. More specifically, the services were provided in the form of an exportable catalogue without any possibility of user interaction.

This type of solution revealed an inability to meet the users’ needs. The first examples of digital archives that provided a platform for sharing and protecting several works did not provide any kind of cross-services or the ability to structure the research based on the user’s request. To overcome this lack of personalisation, the development of DLs focussed on creating services designed to make multiple search indexes available to the user. During the second half of the nineties, the Internet and the technologies associated with it were widely absorbed within companies and institutions (intranets) in order to increase the quality of the services they offered. Before that, private networks were characterised by peculiar languages and protocols. These changes made it possible to improve the services offered to their users and establish direct communication among these networks. As a result, companies’ increasing interest in the web led to a proliferation of institutional sites. In the field of DLs, the USA played a leading role within the international scene, launching the Digital Library Initiative (DLI) in 1994 (Fox 1999). Over the course of the second half of the nineties, this promoted the birth of DLs within several academic contexts. In October 1999, during the Santa Fe convention, the so-called ‘open archive initiative’ was launched. Themes such as the incessant growth of value for documents in digital format (specifically e-prints), Internet support for the academic world and the need to increase interoperability

[9] www.arxiv.org

among academic archives demonstrated how, on the eve of the 2000s, digital heritage was already having to deal with a set of challenges that continue to play a central role in the academic debate today (Sompel and Lagoze 2000)

The very early years of 2000 can be identified as a moment of transition for the web. Following the "gold rush" that had characterised the nineties, in which a myth of the web was developed, sites such as Napster and Wikipedia exhibited the first signs of a change, while in the field of DLs, the European Union promoted the DELOS Thematic Network project, aimed at developing innovations and improving the quality of services offered by the DLs of the communities (Candela et al. 2007)

2.2. Web 2.0

Due to the implication of the events characterising 2004 that were previously mentioned in this chapter, this year can be identified as a breaking point with the previous phase of the web. Among the main features of this phase of the web, a new concept of interaction stands out. Thanks to the new architecture of participation that distinguishes sites of this phase, for the first time, users began to play a central role in the web. Through new opportunity for collaboration, production and development users could interact within the web. This further added a new concept of openness encouraging a deeper level of interaction, allowing users to modify, use and build on existing elements. In this sense, the web was no longer an environment in which information was shared, but an instrument aimed at connecting people. The production of content was no longer the preserve of solely institutions or companies, but of the entire community, within which people began publishing content to be offered to the community. An example of this would be YouTube, founded in 2005.

Due to the difficulties connected to explaining the main changes that occurred within the web between 1.0 and 2.0, I believe it would be helpful to quote the definition of Web 2.0 that Tim O' Reilly provided for the first time in 2005 which he also coined the term Web 2.0 and shared it through his company site, the O' Reilly Media: *'Web 2.0 is the network as a platform, spanning all connected devices; Web 2.0 applications are those that make the most of the intrinsic advantages of that platform: delivering software as a continually updated service*

that gets better the more people use it, consuming and remixing data from multiple sources, including individual users, while providing their own data and services in a form that allows remixing by others, creating network effects through an 'architecture of participation', and going beyond the page metaphor of Web 1.0 to deliver rich user experiences' (O' Reilly 2005).

For the first time, the web structure allowed the user to be the protagonist of his navigation. By contrast, in the first building of the web, users only played the role of visitors, as they were only capable of exploring web resources without interacting with or through them. In this sense, Web 2.0 can be defined in a certain sense as humanising the web, as users gained the opportunity to actively interact within the digital world and become the absolute protagonists of the web experience.

2.2.1. Digital Library 2.0

In 2005, the European community presented the *i2010 Digital Libraries* project (Forster 2007) This project primarily aimed at facilitating access and improving the quality of services offered by DLs of the various members' states by establishing a direct collaboration between them and their teams. Among the important achievements connected to this project, Europeana resulted in one of the most iconic examples of collaboration and sharing among the states of the Union.

In 2007, with the publication of The DELOS Digital Library Reference Model (Candela *et al.* 2007, 17-18) it could be observed how the concept of DL was well established within e-learning, requiring greater standardisation in order to offer a product with certain qualities, abilities and features. Taking into account the leading role that the DLs have had and continue to have within the scope of protecting and sharing digital heritage, this publication should be understood as a fundamental step down the path of this field's evolution, with more and more people, including both specialists and non-academics, turning to DLs field to meet their needs. To satisfy the requirements of an increasingly growing public whose expectations increased simultaneously with the opportunities offered by the digitalisation of information, Users was no longer intended as a mere observer of the experience offered—as in Web 1.0— but as an absolute protagonist and, consequently, a fundamental element in the digital world.

These conclusions can be traced through the new aspect that began to distinguish the web in those years, as well as through its undoubted influence on the world of digital heritage, as expressed in Paul F. Marty's study on museum websites in 2007 (Marty 2008). By conducting a digital visitor study analysing the answer and reaction of a group of visitors to a standard questionnaire format, Marty was able to expose how the use of the digital museum directly connected to specific needs of its users, and how at the same time, the resources offered by such websites played a central role within the lives of the digital museum's visitors even after their experiences could be defined as concluded.

Maintaining the focus on the experience and the services offered to the users, I consider it necessary to analyse the considerations exposed by Duncan Hull and his team (Hull *et al.* 2008). By analysing the main features of computational "biology", the authors sought to trace a series of general applications of Web 2.0 technologies that were able to transform DLs from aseptic institutions into environments developed around the users' needs (exploration, manipulation, contribution). In this regard, software applications aimed at personalising and creating a social sphere within the research could be fundamental for providing a more complete and certainly more satisfying experience to users. This would allow them not only to read but also directly interact with the work in question, stimulating a broader debate that is developing autonomously in a public space.

As can be observed, Web 2.0 played a central role in the evolution of the problems related to DLs. In fact, after focussing for years on the collections contained within these environments, the academic debate at the time shifted its focus to the community's commitment within these environments. By the end of 2010, the role of social networks within society was considerable enough to influence the digital heritage and lead to the birth of new services within DLs focussed on responding to the social web expectations that the public began to show regarding DLs.

2.3. Web 3.0

Applying a systematic division between the various phases of the web could be unsuccessful and counterproductive (Barassi and Treré 2012, 1273). In its particular shape, the web has never put into disuse applications from its early stages. Rather, new technologies have been employed in order to adapt to the needs and possibilities that have come to light. A striking example can be found in e-mailing

applications. These have constituted one of the fundamental elements of the web since its early days, and even today, they still maintain a central role within the various applications offered by the web. Nevertheless, these applications should not be interpreted as static, because new technologies, ideas and applications for the web have strongly contributed to the contemporary aspect of these tools. An example in this sense would be the Semantic email addressing technology (Kassoff *et al.* 2009), which applies computational semantics to the e-mailing tools in order to improve their effectiveness and accuracy and at the same time automatise most of the processes. This example was chosen to reconnect the discussion with the phase analysed in this paragraph. Indeed, another name for Web 3.0 would be Semantic web (Shadbolt *et al.* 2006). In the field of computational science, the term semantic refers to processes put into practice by machines in order to comprehend the natural language and to act autonomously in reacting and best adapting to the external request.

Although this concept had already been developed and put into practice during the first decade of the 21st century[10], it was only following the turn of 2010 that the development of the computational semantic was able to considerably affect the wider public, being exploited within various applications of the web and IT in general.

When discussing the semantic web, it is necessary to consider the concept of computer ontology. By expressing the semantic relationship of the concepts, computer ontology regulates the understanding of data and inputs external to the machines in order to facilitate communication between machines and users. However, in order to properly understand ontology and its peculiar manner of working, it is helpful to first analyse the components that constitute each layer upon which ontology is built.

The first layer could be that which contains the Uniform Resource Identifier (URI),[11] constituted by a string of characters throughout which any resources are identified. Through the use of semantic triple, a term used to identify a series of three entities aimed at issuing a statement, the URIs are put into interaction through a fairly systematic codification (built through language as XML, N3 and Turtle The

[10] The term semantic web was coined in 2001 by the famous Tim Berners-Lee creator of the World Wide Web, who hypothesised a process of automation of the operations conducted by the machines in order to offer a better browsing experience on the web.

[11] The URL of a web page being one of the most common examples of a URI

second layer is based on the Resources Description Framework (RDF), a web standard through which a structure constituted by several triples is developed. Through the RDF Schema (RDFS), the instance, as well as the ability to be contextualised within classes (taxonomic groups), can also be connected to other URIs (relationships) imposing interrelation among a series of triples.

Although developed within RDFS structures, and always constituted by triples, the Web Ontology Language (OWL) constitutes a deeper level of ontology (Alotaibi 2010).

Specifically, in order to offer an exhaustive explanation of the intrinsic value of ontologies, it is appropriate to share the description of ontology offered by Ducharme '*Ontologies are formal definitions of vocabularies that allow you to define classes of resources, resource properties, and relationships between resource class members'* (Ducharme 2013, 39).

By implementing the OWL classes within the RDFs, it is possible to insert further metadata into the information in order to connect them and create a logic through the use of syllogistic logic. This makes the OWL one of—if not the main—pillars upon which the contemporary semantic web has developed. Furthermore, the intrinsic prerogative of such language, constituted by the *Open Word Assumption* formal system logic, which allows the machine not to perceive the absence of information as negative information, but rather as elements not expressed, in this sense creates an extended opportunity for interaction with the machine.[12]

2.3.1. Digital Library 3.0

Although the RDF was published in 1997, and the first version of the OWL only appeared in 2004, when the second version of OWL was released in 2009, the technological level, as well as the attention, were such to allow computational semantics to begin to be implemented outside the web structures and to influence the development of the various sectors. Therefore, all those fields connected to computational sciences, among which digital heritage and DLs, in particular, stand out, acquired the semantic technology in order to improve the quality of the services offered to users. In those years, several projects were brought to light for

[12] This would happen in the case in which a language-based closed-world assumption (CWA) logic is used, where non-information corresponds in a total negation, and in this sense, limits the establishment of an uncompleted system such as that of semantic language.

establishing semantic DLs, such as BRICKS,[13] Callimachus DL (García-Crespo *et al.* 2011,) and JeromeDL (Kruk 2005,), while entities such as Europeana and DELOS worked to implement a semantic enrichment to the services already offered.

As semantics were implemented in the context of DLs, a new opportunity for interacting with these systems was opened, while at the same time, the services offered to the user and the overall value of the DL were extended. The application of semantics can be conceived as an automation process through which most of the tasks that once required direct human intervention were instead managed directly by the machines through Application Program Interfaces (APIs). As can be observed, semantics was expressed in the creation of groups and relationships between objects. Applied to metadata, these procedures allowed the machine to understand the acquired information. This not only meant that all the data within the library could be autonomously categorised by the machine, but also that during the research phases, the machine was able to give more and precise results. A DL that had established a well-structured system of ontologies within the metadata of its files could, during the research phases of queries, offer results not only based on simple word matching, but also on their actual connection with the word meanings, with results based on the semantics of the term and the relationships created around it. Furthermore, based on standards, the application of semantics within DLs represented a new frontier in the interoperability between systems, allowing for new connections between the various DLs, which undoubtedly exerted a positive influence on the quality of services offered by each library. Finally, the semantic allowed users to increase the social value of DLs, as a system based on ontology could be used to share notes, comments and information between the various users of the community, thus improving user interactions, an element that positively influence the experience offered by the libraries by fostering further collaboration among users. In this sense, I argue that the DLs, as well as being a place where the culture was preserved, might also became a place of cultural production, acquiring new value from the academic point of view and, generally, in the field of heritage.

The social component specifically represents the further evolution of the semantic DLs to meet the needs of its users. Born with Web 2.0, in less than ten years, social networks achieved such an influence on society, they were able to

[13] www.brickscommunity.org

require the revising or even recreation of such concepts as sociality, sharing, information and even privacy.

In this context, the direction followed by DLs aimed at offering social solutions within their platforms is therefore clear. Specifically, ontologies were applied not only in the context of the repository but also in the features offered to create a connection between users

Even though in recent years, semantics have become widespread within most web applications, and the growing interest of academics, as well as business and institutions, has led to an increase in the potential of semantic-based systems, even today, such technology has not yet been fully understood and acquired by the public. Therefore, it is clear that there remains much to do before this technology can reach its full potential.

This chapter analysed the process of evolution that has characterised DLs over the last two decades. Obviously, this evolution cannot be summarised in only one chapter, but rather requires a thorough investigation in order to analyse the several aspects that have characterised it.

Nevertheless, the information provided by this chapter can be of help to anyone who wants to deal with this field, and at the same time, it provides a fundamental base on which the discussion of this research and its further analysis is developed.

Chapter III: DL's challenges today: storage, protection, access.

The previous chapter was analysed the field of digital libraries within the historical context in which it has evolved. Now, it is necessary to continue examining the salient aspects of this field, and particularly, the issues that have affected it and upon which the academic debate had focussed its attention. Specifically, DLs represent the main focus of this research, as these take on the responsibility of ensuring operations like the protection and sharing of resources in digital media.

Focussing attention on the digital heritage field, it is also necessary to define the concept of digital heritage in order to analyse adequate solutions for coping with its peculiar needs. In this sense, retrieving the lowest common denominator among the various aspects that constitute digital heritage could aid in discussing possible solutions for the development of related DLs.

In this chapter, the information retrieved over the course of this research is discussed to analyse the core elements representing the issues related to the process of sharing and protecting resources that afflict DLs. Therefore, after having deconstructed the concept of digital heritage to clarify its meaning and scope, it will be analysed the two main branches into which the core issues of DLs observed in the course of this research were divided: storage and cost issues, and protection and sharing issues. This categorisation could help to contextualise the nature of the analysed issues, which consist of establishing a balance between its constituent aspects. Moreover, taking into account that these issues also relate to the technologies used within DLs, it is necessary to highlight their role within this discussion.

1. What is digital heritage?

The discussion of the problems inherent in DLs linked to the heritage field should begin by understanding which resources these DLCH should maintain and share. Although it is clear that these collections must contain heritage resources in a digital version (definable as digital heritage), this statement may be irrelevant as long as the real meaning of the term digital heritage is not dispaced and fully analysed.

As such, I consider it necessary to take a step back and examine the definition provided by UNESCO for the term heritage: '*Heritage is our legacy from the past, what we live with today, and what we pass on to future generations. Our cultural*

and natural heritage are both irreplaceable sources of life and inspiration' (UNESCO 2017).

With the increase of attention, the academic debate has exhibited towards intangible heritage, and above all, towards safeguarding cultural diversity, it is possible to highlight the adaptation that the field of heritage was able to demonstrate at the turn of the millennium. The globalisation process constituted the main challenge of the new century, offering new issues as well as new opportunities, on which the current discussion is largely focussed. Several concepts, such as identity, communication and information, were subjected to mutation, which created a gap with the precedent for interaction with society (Hanna 2016, 3-8). The web further played a major role in the development of the globalized world shaping and common platform to connect the potentially everyone. Providing common tools and practices to interact with his platform the Web was able to establish a common philosophy of use and in this sense a peculiar human experience. Therefore, it is possible to discuss of the web not only as an instrument at services of different fields as that of heritage, but as an actual human experience connected to several aspects of society, and in this sense, part of our cultural legacy. This requires it to be protected in order to be preserved for future generations (Šola 2015, 57).

An important aspect was already highlighted in 2002 by Gregory Crane when he stated that access to the cultural heritage of humanity should not be understood as a privilege, but rather as the right of every citizen (Crane 2002, 630). In fact, as a prominent figure and editor-in-chief of the Perseus project,[14] he pointed out how the process of digitalisation, despite its intrinsic powers, was still too anchored in the logic that regulated the pre-digital period, and how culture still remained the prerogative of an elite who could afford to purchase and/or access such data. After almost ten years, the access to digital culture as a human right was re-proposed within the Special Rapporteur, published in 2011 by the United Nations Human Rights Council. Indeed, almost reaffirming Crane's words, they focussed entirely on the 'the right of access to cultural heritage' (Shaheed 2011), stating in an official and institutional manner that there was a need to reformulate the concept of access to culture, and how that problem remained still a long way from being overcome**.**

[14] The Perseus project is a DL created in 1995 with the aim of sharing cultural material with an open source license (http://www.perseus.tufts.edu/hopper/)

As highlighted in the report on the Implementation of Commission Recommendation 2011/711 / EU, drafted in 2016 by the European Commission, '*the fear of losing control, need to generate income and difficulties to assert public domain status*' (EU Commission 2016, 22) have been identified as some of the main obstacles determining the current situation and the limits regarding access to digital heritage. Among the various solutions proposed by member states of the European Union, the one proposed by Italy, and specifically by the Istituto Centrale per il Catalogo Unico (ICCU), is, in my opinion, one of the most valid solutions for overcoming the problems examined in the report: '*Use the most open possible licences when publishing the digital objects online, for spreading knowledge and fostering reuse*' (EU Commission 2016, 23). The interest that international and national institutions have revealed for this subject demonstrates just how much has been done and still needs to be done to overcome the problems and difficulties examined above, and to reach an institutional and social level that will allow an extensive application of universal access to cultural heritage, including in the digital context.

Concluding this discussion requires examining the relationship between heritage and the web. If the web is contextualised as a well-structured form of the cultural heritage of humanity in digital format, consequently, each user should be considered rights-holders over the intellectual property of the web. As foreseen by Tim Berners Lee, the opportunity connected to the wide availability of data will allow the user to be able to concretise the project in a common effort to develop an improved web structure able to cope with the specific needs of the society.[15]

2. Storage and costs

This section analyses the issues related to costs and storage. As mentioned in the first chapter, these issues are strongly interconnected, which plays a main role in the further analysis.

In order to discuss DLs within the context of digital heritage, or the academic world in general, it must first be taken into account that these institutions are largely based

[15] In this regard could result valuable the vision of the two TED talks published in 2009 and 2010, respectively, in which Tim Berners Lee presented his conception of the web and his related projects twenty years after the original presentation of the W3 original proposal (https://www.ted.com/talks/tim_berners_lee_on_the_next_web; https://www.ted.com/talks/tim_berners_lee_the_year_open_data_went_worldwide)

on public funds. Compared to their concrete counterparts, DLs represent an indisputably cheaper solution, thanks to the intrinsic characteristics of the digital medium, which reduce the costs of maintaining, copying and distributing the information contained in these archives to almost zero. The progressive decline has significantly reduced the costs for establishing and expanding the storage capacity of DLs. However, in the effort to reduce storage costs to zero, it is necessary to note that these account for only about 20% of the expenses, and therefore, approximately another 80% of expenses associated with DLs (Lesk 2005) should be considered.
These expenses, in addition to staff remuneration and the maintenance of the services offered by the DLs, refer to those associated with the concrete structures in which the DLs are based (offices, server rooms, etc.) and the costs related to their maintenance. Obviously, it could be said that through the use of cloud computing, the costs of the premises, such as those connected to scalability, can be somewhat reduced, as illustrated by Sanchati, R. and Kulkarni G. (Sanchati 2011). On the other hand, even if the authors have highlighted that this solution avoids the imposition of prohibitive expenses, it should not be forgotten that outsourcing also leads to an overall increase in costs related to DLs (Barthelemy 2001, 62-66).

Considering this, it is clear that, overall, the most economical solution for a DL is to maintain its functions and services within its own institution and avoid including third-party companies within its own processes or, moreover, entrusting the performance of certain activities to them.
Proceeding, it is necessary to analyse scalability, as this issue, as well as its economic implications, plays a central role in the discussion of DLs. The term scalability refers to the system's ability to be easily modified and to increase its features and the amount of information contained. In addition, this refers to the system's capacity to respond to the possibility of an increase in active users without suffering a drop in performance (Hill 1990; Chowdhury and Foo 2012, 21).
As discussed above, cloud computing allows for an excellent level of scalability because, relying on third-party networks in cases of a higher request for storage space, it does not need to buy new servers and new equipment, but only to subscribe to an upgraded version for the space available in the cloud. In addition, because it always relies on the same service and does not have to change the servers, operations such as migrating and copying data are no longer necessary. On the other hand, as already pointed out, cloud computing has not turned out to be the most

economically sustainable operation for a reality such as DLs. In this sense, this could provide a valid option for exploring new alternatives in order to find new answers that can offer concrete benefits for the field of DL, overcoming the difficulties associated with the hardware field, but at the same time without being engulfed in the logic of outsourcing.

Finally, to conclude the scope of the issues analysed in this sections, it is necessary to discuss the education of digital librarians and, in general, the staff who will be responsible for maintaining DLs, as well as ensuring the normal provision of services to users. In fact, staff education plays a major role in ensuring the proper functioning of these institutions and guaranteeing that new steps are taken in this area of research to achieve a state-of-the-art level. In this context, it is appropriate to report the considerations raised by the work of Terry Weech (Weech 2007). This study was based on the same initiative and studies conducted by the International Federation of Library Associations and Institutions (IFLA),[16] which focussed on drawing up documents and guidelines for educating digital librarians. In his work, Weech, according to the studies carried out until then, identified the skills and knowledge that digital librarians needed to possess in order to be able carry out their work in the best way. In this sense, in addition to the classical skills as a librarian, when working in the digital world, the staff must possess skills in computational sciences to cope with the specific problems that may appear when dealing with both analogue and digital resources (Abbas *et al.* 2006). Taking into consideration the research work conducted by Howard, the various representatives of Australian DLs were questioned regarding the skills most relevant to their work in order to determine how the ability to meet the needs of users resulted from the group of specific knowledge that was deemed most 'highly desirable' (Howard 2010, 269). The most quoted skills included flexibility, innovation (for the field of personal skills), initiation and critical thinking (for the field of general skills), as they were deemed the skills necessary to adapt to a constantly evolving world like that of DLs. On the other hand, it was, rather strange that the capacities belonging to the IT field were considered 'less desirable' for the majority of respondents (Howard 2010, 271-273). For the author, this may be due to how DLs offer librarians the potential support of teams of ITC specialists who are in charge of carrying out the more technical maintenance and tasks. Nevertheless, the lack of attention paid to IT is

[16] https://ifla.org/

counterproductive for DLs. In fact, although the idea of placing such tasks under the control of highly specialised personnel is valid, at the same time, digital librarians must be familiar with the technical counterparts of the instruments they use in order to be more involved in the discussion concerning the problems and opportunities inherent to this field.

3. Protection and Sharing

The analysis of the problems related to DLs cannot disregard examining the issues related to protecting and sharing digital material. Even if these operations may not seem to have a direct link at first glance, they actually represent two sides of the same coin, because when these concepts are considered in the DLs tasks, one cannot disregard the other.

On the one hand, wanting to concentrate attention solely on protection, it would be sufficient to keep the material in consideration in a microfilm, or better, in an archive which, even if distributed,[17] would be encrypted, and to enable access only for the smallest number of people who can be considered trusted (Skinner 2010). On the other hand, wanting to focus solely on sharing, it would be sufficient to share this material directly on the web, ensuring maximum usability of the material for anyone who wishes to use it. Obviously, this is a *reductio ad absurdum*, but it can help to understand how the work of DLs is perpetually hovering between the priorities of protecting and sharing information.

Focussing attention on the various techniques and tools that, in the course of human history, have been used to preserve and pass on information (such as clay tablets, papyri, books, digital media, etc.), it is well known that digital tools possess a clearly less extend life cycle compared to their predecessors (Hedstrom 1997, 197-198; Nelson *et al.* 2007). The issue is further complicated when the context of the web and the data it contains are examined. In this sense, the data reported by M. Costa and his team (Costa *et al*. 2017) can be illuminating, which revealed that after the first year of life, 80% of web pages are no longer in their original form, affecting the academic production that refers directly to these sites. This is further accentuated considering that the average lifespan of a web page is only forty-four days (de Lusenet 2002, 3). In this context, it is useful to examine the seventh article

[17] Through the distribution of the system, the protection of the material would be assured even in the event that one of the stored storages broke or the information within them was compromised

of the chapter 'Digital heritage as common heritage' from to UNESCO's guidelines for preserving digital heritage (National Library of Australia 2003). This article, entitled 'Defining what should be kept',[18] stated the need to preserve only the essential materials and what is strictly definable as heritage.

A further element that has played an increasingly prominent role in the relationship between users and the digital world in recent years, and specifically in Web 2.0, would be the concept of sociality and interaction between users. As already discussed social networks have strongly influenced the way people interact with the digital world, creating a real social expectation towards all the various applications of digital technology, including DLs. Specifically, as already discussed illustrate in the use of computational semantics in DLs has constituted a possible solution for meeting these expectations. In this sense, it is interesting to mention Nisheva-Pavlova *et al.* 2015, 273: *'Social semantic libraries reflect the changes in users' expectations resulting from the wide penetration of social networks in everyday life of a continuously widening variety of communities'*. Indeed, the semantics enable interacting with the documents by inserting annotations, corrections and comments that can be acquired by other users. In addition, by implementing an effective social media programme, DLs could strengthen the relationship of trust with its users and allow them to interact with each other. These elements would then lead to overall improvement in the quality of services offered, and at the same time, transform DLs into cultural production centres within which academics and non-academics can gather, share and collaborate (Kruk *et al.* 2008, 9-39).

Still focussing on the users, who ultimately appear to be the true protagonists of DLs, it is important to discuss the concept of usability and, in general, the relationship between the user and DLs. One of the main features of valuable DLs is its user-friendliness, which means being able to cope with the needs of its audience and making their work as simple as possible. Taking into account the fact that the public of DLs is highly varied in composition, incorporating users from different social, cultural and economic backgrounds, it is necessary to search for the lowest common denominators to make services accessible to everyone in the best way possible. To this end, the user interface (UI) plays a primary role within

[18] The chapter it is part of the sub-paragraph 'required measures' in National Library of Australia 2003,14.

the DLs experience (Sastry *et al.* 2011). The UI forms the point of connection between the user and libraries, and as such, must enable simple and intuitive use of the functionalities offered by the DLs field, which must be accessible to even less experienced users. Beyond this, a feature that often goes quietly unnoticed, but which plays an important role in democratising the task of DLs, lies in the weight of this service. As previously mentioned, the DLs' audience comes from different contexts, which means that not everyone can rely on a fast and high-power connection. As such, it is the duty of DLs to keep their services within everyone's reach, which means avoiding overloading the network traffic with unnecessary material, such as heavy graphics or other elements not useful for the actual functioning of the DLs. Summarising the importance of establishing a highly usable and user-friendly system, it should be pointed out that DLs resources lose any role they should possess when their users cannot exploit the data stored within them. User communities represent the true core of DLs, and as such, their needs play a primary role within DLs, which must evolve and be modelled according to these needs in order to achieve state-of-art on the field.

Discussing protection and sharing, it is necessary to examine the issues concerning intellectual property. Like any type of property right, several laws regulate and ensure the protection and use of data that can contain intellectual property material. Specifically, in 1968, the World Intellectual Property Organisation (WIPO)[19] was founded, a United Nations agency that focuses on regulation and legal protection for intellectual property. The national and international laws that regulate the intellectual property and the regulations promoted by WIPO play a role of primary importance in the work of DLs as institutions that have to deal with protecting and sharing material not directly belonging to them. In fact, rather than possessing the materials contained within them, DLs instead lease them, and operation based on license agreements established with the owners or those who hold the property rights. The DLs must perform their work, which consists of offering access to the materials stored within them to all their users, within the limits of such agreements, and in compliance with the current laws. Laws such as copyright play a primary role in the protection of intellectual property, as they ensure not only control and authenticity, but also limit the circulation of these materials affecting the development of scientific research.

[19] www.wipo.org

The copyright law places the right to decide concerning authorisations, reproduction and adaptation of the work directly under the control of the copyright owner, considerably limiting the initiatives of third parties, which, if not allowed by the author, are illegal. Meanwhile, it is necessary to specify that for academics or specialists, it is impossible to acquire a copyright license on a theory he has developed or a discovery he has made, as this would adversely affect the scientific progress of humanity. Copyright protection can only be used to protect the artistic component of a specific work, and therefore, the style and everything not necessarily connected to the clear understanding of the scientific component must be free to use for anyone who wants to learn and pursue new scientific progress (Kallinikou *et al.* 1993, 7).

Considering this, copyright laws create many difficulties in ensuring libraries' tasks. On the one hand, libraries have a duty to protect and ensure access to as many works as possible, while on the other, they are required to comply with copyright laws and to protect the intellectual property of these works. However, many copyright owners increasingly turn a blind eye to the activities carried out in and from DLs in a situation of semi-illegality. This results in a counterproductive and decidedly unwholesome scenario, requiring new solutions in order to overcome these problems, or at least contain them. One possible solution can involve open access (OA). It is important to understand how an OA license clearly diverges from public domain publication. Taking into consideration the creative commons license, OA allows the author or anyone who holds the rights of the work to impose different degrees of protection for their work (commercial use, distribution and publication related), and moreover, to publish the work under copyright through other channels. In clear contrast, a public domain publication implies that the author concedes every right regarding the use of the work to the public domain, and in this sense, no action can be performed against those who use such work without respecting the will of the author or the rights-holder (Bailey 2007).

Today, OA constitutes a very important resource for DLs and the academic world. Taking into account the massive amount of material that is published each day, DLs do not possess sufficient funds to purchase the new publications. However, with OA material, the DLs can offer their users new and updated publications, reducing the overall costs to near zero. Another important element is the protection of this material. Not having to comply with copyright laws, DLs can

ensure the material's preservation in the digital world without having to incur legal or ethical limitations. It is also important to emphasise that the concept of OA embodies the values of the democratisation of knowledge. This material can be useful for anyone without—generally—any limitation, positively influencing the dissemination of knowledge and providing anyone with the opportunity to enrich their cultural baggage without having to be affiliated with major institutions or to possess particular economic opportunities.

Therefore, it is possible that the scope of protection and sharing are not limited solely to the application of these operations within the DLs, but also to other questions related to fields such as jurisprudence and social sciences. Overcoming these problems, while a highly complex endeavour, is not just a question of improving the services offered, but a real priority for the area of the DLs that needs new implants to adapt to the fluid situation that characterises the present and future of this institution.

Chapter IV: What does P2P technology offer to DLs?

Having discussed in the previous chapters the general concepts of heritage and DLs, the evolution of the field and the main challenges which yesterday and today characterise the area, it is now appropriate to focus the attention over the core theme of this research. In this sense, in this chapter, it will be discussed the P2P technologies and their application within the field of DLs and specifically within the DLCH branch. Therefore, it might be helpful to introduce the concept of Peer-to-Peer and related technologies. To establish a valuable analysis, it might result useful in the first place to check the qualitative definition of the term Peer-to-peer developed by R. Schollmeier: *"A distributed network architecture may be called a Peer-to-Peer (P-to-P, P2P) network, if the participants share a part of their own hardware resources (processing power, storage capacity, network link capacity, printers, etc.). These shared resources are necessary to provide the Service and content offered by the network (e.g. file sharing or shared workspaces for collaboration). They are accessible by other peers directly, without passing intermediary entities. The participants of such a network are thus resource (Service and content) providers as well as resource (Service and content) requestors (Servant-concept)"* (Schollmeier 2001, 101). Though such definition it might result easier to discuss P2P technology and their applications. Indeed, since the early years of these century P2P technology have been generally affiliated to practices of online piracy and unethical software and services which have correctly seen as menaces to the protection of IP of resources creator. Nevertheless, it is necessary to understand that is this is actually due to the opportunities provided by P2P which could be used to widen the scope of DL functions and in addition to offers solutions to cope with the same lack of IP protection that P2P technologies have pointed out.

In order to discuss P2P protocols, it is first necessary to provide a short recap of the protocols that, today, are used for web navigation and research: the HTTP (Hypertext Transfer Protocol). Since the birth of the World Wide Web in 1991, this family of protocols have been used to access web resources. The basic operation of HTTP protocol involves tracing the location of a specific resource within the web using its Uniform Resource Locator (URL). In so doing, the web browser receiving the URL in its search query is able to communicate with the server in which the

resource is located to request a copy of the resource. In this sense, in the current state of the web, the resources are bonded to the servers in which they are located and do not possess identification concerning what they actually are. The limits of this system are clear: when a resource is moved away from its original location, it is impossible to access that content again. One of the most well-known examples of this problem is represented by the famous web error message 404, which indicates that the browser is in communication with the server, but is unable to find the required resource. On the other hand, even if the owner has removed the resource from the web, it can still potentially be located in a cache folder held by some users that had access to that resource before it was taken down. If one of these users is online, hypothetically, it could be possible to acquire that resource from their cache folder.

In response to this issue, in 2001, computer programmer Bram Cohen created the innovative peer-to-peer (P2P) protocol for file sharing, published with the programme BitTorrent.[20] This protocol, and consequently, the programme based on it, allowed resources to be acquired directly from other users instead of using servers to create a direct connection among users. The P2P protocol did not base the research of a resource on the latter's location, but instead, on a specific identifier for the resource, represented by the famous *.torrent* file. With this file, which contains metadata and important information about the resource, a user, through the tracking protocol, could find other users connected to the system who possessed the resource, and were thus able to share it. As a result, while HTTP uses the location to find content, P2P employs a unique identifier to find the content. Moreover, without a single or limited server having to share the file with hundreds of users, the download speed is greatly increased. Thanks to the specific characteristic of the protocol, the resource file is divided into small packets, allowing it to be easily downloaded by users. Every time a user downloads a packet, he becomes himself a sender, allowing other people to acquire packets from his device.

Therefore, the P2P protocol is not only able to serve multiple people at the same time without any decline in the speed of bandwidth, but actually becomes more efficient each time a new user takes part in the process of sharing resources. The P2P protocols upon which several important applications have been built today

[20] www.bittorrent.com

such as Skype and Spotify, represents the paradigm of decentralised and distributed systems from Web 2.0 and on. Even if its first mainstream appearance with Napster created several legal and ethical issues that, even today, still play an important role in specialised debates, it is important to remember the role this technology has played within contemporary society.

These features should be considered in light of the fact that increasing the speed of the worldwide Internet connection has been revealed to be a slow process, characterised by several inequities between the different regions of the world.[21]This data thus contrasts with the technological path of evolution and improvement for services such as storage or calculation power. As a result, it is possible to observe that while the average weight of resources has received an important increase, the slight increase of bandwidth remains unable to cope with the increasing amount of power required by users.

1. IPFS: analysis and use

In 2014, Juan Benet, founder of Protocol Lab, published the third and final version of a paper (Benet 2014) presenting an innovative protocol for a distributed P2P data system for storing and sharing hypermedia. This file system advanced the technology developed within the context of both blockchain protocol and P2P file-sharing protocols to create an alternative to the HTTP-based system. The name of the new protocol was InterPlanetary File System (IPFS), which, in addition to being a catchy name, also explained the main difference from the other protocol used until that moment.[22]

The IPFS is content addressed, which means that the location of the resource loses any importance. Instead, the system uses the secure hash [23](identifier) of the content to search and localise resources within the network. Proceeding, in order to

[21] This slow increase is well demonstrated by the data on Akamai's state of the Internet report (Akamai 2017). Comparing the data of the q1 of 2015 against that of the q1 of 2017, it is possible to observe how in two years, the global average connection speed increased only from 5.0 Mbps to 7.2 Mbps.

[22] *'Designing protocols that work across vast distances like planets makes it possible for people around the Earth that are in distant areas, far away from the data centre, to connect as easily and nicely as people in the nice vast cities like this one (San Francisco). This is what we call the interplanetary principle. Design things for planets—planetary scale—and you will do very well on Earth*'. Juan Benet during his speech at the TEDx San Francisco in December 2016 (https://www.youtube.com/watch?time_continue=2&v=2RCwZDRwk48).

[23] Hash is a term used to define mathematical function created to protect (cryptography) and reduce data weight (compression) of resources

enable easy and efficient exchange of data (blocks) among users, the IPFS takes advantage from the BitTorrent peer-to-peer file distribution protocol, which with almost 20 years of history still proves to be efficient, reliable and extraordinarily scalable.

To organise and manage the exchange of resources, a protocol was developed called Bitswap (based on BitTorrent protocol), which was intended to efficiently control block exchange and distribution. Another fundamental role within an IPFS network is played by the Merkle Dag (direct acyclic graph) used within the GIT hub.[24]A Merkle Dag possesses a double function within a network: First, it manages the resources and objects within the network in a tree-shaped organisation. Furthermore, all objects possessing the same hash are stored as just one copy, avoiding an overload of the system. Second, the Merkle Dag allows for checking the validity and integrity of resources. As already observed, the IPFS, as a P2P system, is a distributed network in which members are anonymous and no central authority exists to ensure the reliability of the participants and the files shared by them. As a solution to this lack of security, a Merkle Dag could be used to check the file was erroneously modified or maliciously tampered.

A further protocol on which the organisation of the IPFS network security is based is the self-certified file system (SFS). The SFS simplifies the operation with the system, requiring each server to provide a private key to the users, which could be used to authenticate the node and create a point from which communication can be started with it. Due to the fact that the system lacks a central authority to preserve all the server keys, only the hash of these keys are shared in a process called the *self-certifying pathname.*

These comprise the main features of the IPFS protocol presented in 2014 by Juan Benet. Since then, the project has grown and several releases have been developed. On July 27, 2018, the first stable release was published. Together with the project, the attention of both the specialised and general public has also increased in these years.

Decentralisation and distribution comprise two of the main characteristics of the IPFS. Within a decentralised system, no authority wields control over the network and the operations carried out within it. Moreover, distribution implies that operations such as storing and processing resources happen in different nodes

[24] www.github.com/

contemporaneously. The division of resources into multiple chunks of data eliminates the previous relation between server-client in acquiring resources. Downloading the small package from several nodes allows the connection speed to be increased, and at the same time, prevents bandwidth overload. Other issues afflicting HTTP that the IPFS is able to resolve relate to security. A content-addressing network allows sites built on the IPFS to prevent an attack on the network, such as the famous DDoS (distributed denial-of-service attack) caused by flooding a server hosting a specific web page with requests.

From the above explanations, it is easy to understand how the IPFS has proposed a new revolutionary phase of the Internet, and generally in ICT evolution. To date, several applications have been built using the IPFS as a browser, storage platform, communication platform, social network and even operating system.[25] Regarding applications, the project has started to host sites developed specifically for the IPFS, demonstrated over the last year an increasing interest from the mainstream public concerning the IPFS protocol and, in general, the new way to approach the internet proposed and developed by the Protocol Lab.

2. Blockchain technology

The theory upon which blockchain technology was developed was published in 2008 by Satoshi Nakamoto, [26] and with it, the concept of cryptocurrency, which led in 2009 to the birth of Bitcoin. The hope for this system was to enable economic transitions in the digital world without the need for a trusted entity (e.g. the bank), which had the task of ensuring that there was no double expenditure of the amount involved in the transaction. The term blockchain refers to an *"online, digital payment systems based on peer-to-peer networking technology and public key cryptography. These systems allow users to an exchange value in a trustless setting"* as those created by open P2P systems (Hoy *et al.* 2017, 274).

To offer a clearer understanding of this technology, it is appropriate to focus on the salient features that compose it.[27] At the base of a blockchain system is

[25] For a complete list of applications developed in the IPFS, see the following: https://github.com/ipfs/awesome-ipfs/blob/master/README.md.

[26] Satoshi Nakamoto is a pseudonym of the individual or collective that created the concept of blockchain. To date, his or their identity still remains unknown.

[27] This analysis is based on the original publication that created the concept of blockchain (Nakamoto 2008).

located a timestamp server, which deals with recording in the hash, therefore every operation performed within the system and publishing a public report of these transitions. The simple fact that a transition has been included in this report indicates that it has occurred, and being public, all users of the system can be assured of its existence.

Based on a P2P system, it is necessary to provide an excellent security apparatus on which to place the reliability of the history of the transitions performed and, at the same time, to ensure that the system's democratisation does not fail by imposing a safeguard authority. In this sense, blockchain technology is based on protection called proof of work (PoW). Based on the idea of one vote per CPU,[28] the system ensures the preservation of democratic organisation as the majority will create a block, and making it outclass any parallel block, which could be the result of a minority or even misguided. Moreover, if malicious people wanted to modify the information of a chain, they would not only need to modify the block of interest, but would also have to modify the successive blocks, as they report the information of past blocks, thanks to the peculiar characteristics of hashing (Daintith and Wright 2008), causing complete modification of the chain. Entrusting the security of the system regarding the publication of transaction data, it is necessary to provide for the privacy of users. In this regard, the concept of private and public key comes into play. These keys, which are bound to each other, represent the digital signature system of the blockchain, through which is created a unique identifier aimed at identifying the author and his transition and offering an authentication system to the receiver. The senders use the private key, in the 'signing phase' to sign his own transition and encrypt his data. The receiver in the transition 'validation phase' uses the sender public key to verify and decrypt the data, thus ensuring that the data are correct and that third parties have not tampered the version received.

Finally, a vital element in the compression of a blockchain-based system is the network organisation. As blockchain is a distributed P2P system, the information is kept contemporaneously in all the nodes of the system. In this system, every node has the task of developing an increasingly complex PoW.

[28] In fact, the CPU is more valid than basing the system on the Internet protocol (IP) address, as these can be more easily referable to the same subject. Furthermore, while each CPU involves active participation by the user, IP addresses can be maintained without the owner actively participating in network operations.

Whenever a new PoW is found by a system, this and the relative block are shared with the rest of the system nodes.

These form the salient features of the blockchain technology, and it is based on these (at least conceptually) that all types of systems that use this technology are developed.

3. IPFS & Blockchain

Throughout the decade that followed Satoshi Nakamoto paper's publication, blockchain technology has acquired increasing relevance in our society. In the wake of Bitcoin, several protocols for cryptocurrency have been developed, each possessing peculiar features developed to cope with specific necessities. All these new protocols have also revealed that blockchain does not only mean cryptocurrencies but rather that this technology could be applied to several aspects of society, including the academic field. To summarise the main features of a blockchain-based network, it is possible to affirm that this type technology enables creating distribution-decentralised networks in which all the users' operations are rapidly saved in a shared, time-stamped archive, which exhibits a high level of security against accidental error and malicious attack.

These features have resulted in a fertile terrain on which a decentralised currency system has developed, but when examining the situation from a more general point of view, it is possible to observe how a new model for protecting and sharing data resulted in the main objective achieved by blockchain technology. In this sense, I think that could help to analyse some projects developed through blockchain technology within the academic, scientific and heritage field.

In 2016, Azra Aksamija, in collaboration with MIT (MIT Architecture), installed the Memory Matrix, an interactive monument that *"explores the possibilities for future heritage creation, employing new fabrication techniques and transcultural collaboration"* (Aksamija 2016). The monument was built on a metal frame, on which 20,000 fluorescent Plexiglas, called pixels, were applied. Each of those pixels, produced by various participants, had a hole cut with a laser in the shape of a vanished or undertreated heritage site. These pixels are able to together form the image of Septimius Severus' arc of triumph in Palmyra, which was destroyed in 2015 by the terrorist organisation ISIS.

All the pixels are connected to data stored within the Bitcoin blockchain network. The project Cryptographic Heritage, guided by Dr Dietmar Offenhuber and developed from the same consideration of Aksamija, focusses on the solutions implemented in order to store vanished heritage and immaterial cultural heritage in digital format. Offenhuber observed UNESCO's effort to store this data on a web page in order to ensure that the information would be protected and shared. However, he was very critical towards this solution due to the fact that websites remain *"just as vulnerable as the practices and buildings they are supposed to document"*(Offenhuber 2016). It is within this framework that the Cryptographic Heritage project was developed as a new perspective in the theme of heritage protection and sharing. Integrating heritage data within a transaction was demonstrated to be a useful solution for guaranteeing both a high level of protection and accessibility for these data. Blockchain creates a network database that is open to everyone but is also protected from tampering. Using the Bitcoin blockchain, the Cryptographic Heritage project was able to store the message from the creators of each pixel within the blockchain. Through the cryptographic encryption, only the pixel's owner, the one in possession of the private key, is able to fully access the data (transaction), and consequently, to modify the message. At the same time, this solution allows everyone to access and visualise the message through the use of the public key, which is printed on the lower part of the face of each pixel. Thanks to the peculiar distributed structure of the blockchain network, the data are stored in millions of nodes at the same time, and as a result, it is virtually impossible to fully delete the information from the networks. Thus, it can be said that Offenhuber's project found a solution to permanently protect and share intangible cultural heritage through a globally distributed encrypted database.

In an interesting article (Chen *et al*. 2017) the authors analysed the main processes inscribed in the establishment of a file system based on the IPFS and blockchain. The specific distributed conformation of an IPFS network could result in a speed increase within the processes of research and fruition of the data. In addition, the cryptography and distribution of data within the system responds to the issues related to protection against tampering, and also ensures that the data could remain permanently available. Furthermore, through the application of a blockchain layer over the data system, it is possible to implement an incorruptible distributed time-stamped archive in which each activity occurring within the system

is permanently stored. The base layer over which have been developed Chen and his team's model was the famous Bitcoin blockchain, which represents the first layer of the system model. The second layer, defined as the virtual chain, is comprised of all the activities and functions that could not run within the Bitcoin blockchain. The successive layer is defined as the routine layer due to the fact that (mutable) data used within routine operation are stored here. The final layers are those on which the (permanent) data are actually stored. In this sense, users could utilise the storage system without needing to trust the first two layers through the use of hash validated by the control plane. The final layer is storage, in which data are actually stored in the IPFS. Thus, it is possible to affirm that the article proposed the creation of a system that could cope with the exigencies of service providers, such as DLs. The article's proposals can also be understood as a valuable basis from which future analysis can be developed regarding the application of the IPFS and blockchain within the field of service providers, and consequently within the institutional and academic contexts.

The conclusion developed in Chen *et al.* 2017 was used as a basis for the article by Rajalakshmi and his team (Rajalakshmi *et al.* 2018,), which analysed the implementation of an IPFS and blockchain-based system to keep track of and secure research and academic records within the digital space. Specifically, their article analysed the establishment of a distributed data system should regulate the preservation and fruition of academic data without requiring the presence of a central authority. The aim of the article largely coincided with that of Chen *et al.* 2017. However, taking into account that the proposed environment would be focussed on the academic field, a peculiar solution has been proposed to cope with the specific exigencies and issues of this field. More specifically, focus on research records, the authors demonstrated how the availability and reliability of this data were fundamental for it to be used correctly. In this sense, they proposed developing a system built over blockchain and the IPFS in order to protect data from tampering and, at the same time, to ensure productivity for users without requiring a central authority to regulate and supervise the system's operation.

A fundamental element in Rajalakshmi and his team's project consisted of smart contracts. Much like a classic contract, a smart contract describes an agreement between two or more figures that runs until the requirements from all participants are fulfilled. Unlike a classic physical contract, however, when a

specific condition is reached, smart contracts are able to self-execute and do not require an authority to mediate disputes, thanks to the combined use of cryptography and distribution. The peculiar characteristics of smart contracts have made these programmes applicable within a blockchain system, which can utilise them to increase the scope of its opportunities, while at the same time better-regulating relation among users within its network. In 2015, this fertile terrain led Vitalik Buterin to develop Ethereum, a blockchain distributed environment. Ethereum was strongly inspired by the blockchain structure of Bitcoin, even if was developed specifically to integrate smart contracts within the digital ledger of the blockchain. In this sense, the Ethereum blockchain is not only able to track transactions within the networks in which it is applied, but can also programme them without the presence of a middleman to regulate and supervise the operation (Buterin 2013).

Returning to the work of Rajalakshmi and his team, their article proposed a framework for a DApp built over the Ethereum blockchain and utilising the IPFS as storage layers. The authors offered a framework designed to keep the research records and ensure their reliability and correct use. It is a perfect case to apply smart contract and distributed storage. This solution could enable automatise processes that will no longer be tied up in the trust of an institution but rather based on collective consensus concerning mathematical computations. To further understand this system, it would be helpful to analyse its specific characteristics in order to better comprehend the processes involved. The first element to take into account is a division of users into principal investigators (PIs) and junior researcher fellows (JRFs). This is used to determine the users' role within the system. The PIs are only able to upload a paper in the system, while JRFs can access the data, and if they need to propose a modification of resources, they can request permission to do so from the related PI. Resources collected in the system are uploaded in the IPFS. In order to restrict access only to the registered users, asymmetrical encryption is proposed to be applied to each resource using the Gnu Privacy Guard (Rajalakshmi *et al.* 2018, 1439-1440).

These encrypted data can be downloaded by everyone with access to the IPFS and the hash of the data, but only those with the decryption key (master key) can actually decrypt this data in order to use it. This master key is shared within the system, and both PIs and JRFs can decrypt the data. At the same time, data is

maintained over the IPFS, as every user is allowed to download, modify and even delete the work from its terminal without compromising the original version of the data.

The successive layer of the system is represented by the blockchain platform. The distributed application (DApp) built over the Ethereum blockchain collects all the metadata from the data stored in the IPFS, creating a provenance record file that includes such information as the identification of the PI, the hash of the data, the date of creation and so on. Each time a user accesses data, the system creates a transaction within the blockchain using smart contracts that retrieve and decrypt the document from the IPFS. This transaction also records the access (data and user) to the file and stores it permanently in the distributed blockchain ledger.

Having analysed its features, it is clear that the framework proposed by Rajalakshmi *et al.* 2018 could provide a valid opportunity to increase the quality of research retrieval and the protection of research records in the digital space.

The last project to be analysed is the Caltech Tomography Database, discussed in the article Ortega *et al.* 2018. This database was created to store the electron tomography (ETDB) dataset, recorded by Caltech's microbiology laboratory. This dataset, consisting of a series of 2D Transmission electron microscopy (TEM) projection images and the resulting 3D tomographic reconstruction, has created several issues regarding storing, protecting, and sharing these data. Each dataset could contain approximately 1 to 5 GB of material, and considering that the entire collection includes more than 11,000 datasets, the total amount of material exceeds 110 TB (Ortega *et al*. 2018, 5). The researcher was interested in the work of Alexandria,[29] which drew from blockchain (FLO[30]) and IPFS technology to establish a decentralised library, the Electron Tomography Database – Caltech (ETD - Caltech). The operation related organising the data within the system was regulated by a MySQL database hosted in an Ubuntu System as a server. The server was in control and ran a node within the FLO blockchain, and consequently worked in parallel with the Open Index protocol (OIP) operations. Data could be found through the use of queries and then downloaded directly from the blockchain and IPFS. To facilitate the operation, the ETD created a browser-

[29] https://www.alexandria.io

[30] https://flo.cash/

based interface[31] in order to facilitate access to everyone interested in their material. If the project's main goal was to allow access to as many people as possible, it was sufficient for the Caltech laboratory to share just their server data. However, the project's actual objective was to establish a data repository that could also be used by other laboratories around the world and was not controlled by a central authority.

This section demonstrated how the combination of blockchain technology and the IPFS could be able to overcome several issues affecting the actual state of the digital world. The analysed projects comprise only a part of the applications that these technologies could offer in contemporary society, and consequently, in the spread of culture. However, these analyses were able to highlight the strength of their application in creating distributed, permanent, tamper-proof archives.

4. Ripple Blockchain

In order to proceed with the next chapter, it is first necessary to introduce the Ripple protocol and the related XRP ledgers used as a fundamental core for the Heritage Chains framework model. As analysed in the last section of the second chapter, Bitcoin employs a solution called PoW, which is able to ensure a high level of security. At the same time, however, the operation of validation (mining) requires high computational power and a considerable amount of energy in order to be completed. This has enabled the creation of a flourishing market around the mining, but at the same time, PoW requires a long period for validation and a high cost for the operation. These problems were identified by the developers of Ripple as the main limits to overcome when creating a protocol of transaction verification. To this end, the Consensus protocol (Schwartz et al. 2014) was developed. The XRP ledger used a distributed agreement defined as the Ripple Protocol Consensus Algorithm (RPCA), conducted every few seconds by every validator node in order to ensure network protection through a democratic decision process.[32]

The consensus validation process represents Ripple's answer to the security operation performed by PoW. The validation process of candidate sets allows the system to create a shared time-stamped ledger of operations. Beyond the mere economical role, the RPCA also provides an efficient means to create a distributed

[31] https://etdb.caltech.edu/

[32] In order to better understand RPCA functioning, it could be helpful to watch this explicative video shared on Vimeo by the Ripple account (https://vimeo.com/64405422).

ledger archive in which all the information can be checked and retrieved by each node in the system. The main difference from the PoW resides in the cost distinguishing the two operations. While in the PoW, users need to use machines possessing advanced hardware power and to spend a considerable amount of energy to be able to mine (validate) each block (Hern 2018), the consensus process can be carried out by less powerful devices, and with a relatively reduced cost of power, meaning it is possible for a larger number of users to collaborate in the validation of ecosystem transactions. To run a validator node, the costs are comparable, in terms of power consummation, to those related to maintaining an email server (Ripple 2017). A second element that should also be taken into consideration is that in PoW-based systems, nodes compete to be the first to solve the PoW mathematical problem, while in the RPCA, all the nodes collaborate in the validation process. Thus, validation is no longer tied to the work of a single user, but rather to the collective consensus of the entire ecosystem.

Therefore, in the next chapter, it will be discussed the proposal of a framework model in which both Ripple and IPFS protocols are used in order to establish a scalable ecosystem in which could be offered valuable front-end services by multiple institutions at the same time.

Chapter V: A framework model for P2P in Digital Libraries

In this chapter, it will be presented prose a framework model for the application of The P2P technology proposed within the context of Digital Libraries. Furthermore, it will also establish a discussion over the opportunities offered by the framework able to cope with the issues highlighted in DL's challenges today: storage, protection, access (Chapter III).
In this sense, before proceeding with the actual presentation of the model, it is necessary to discuss its relation with heritage in light of the discussion developed in the previous chapters.

As discussed in the first chapter accordingly to Tomislav Šola it is possible to retrieve a general theory of heritage defines as Mnemosophy. Such concept in addition to offering an interpretative definition of the field of heritage it also expresses the core element of heritage specialist practice: the transformation of (public) memories in wisdom. To carry out this practice, specialists have to, first of all, take into account the requirements and perceptions of consumers to establish a quality connection with them. Specifically having focused the attention over the field of Digital Heritage and specifically over the operation of sharing and protection which constitute the core function of that broad category of institutions and service inscribed within the term DLs it is necessary taking into account the observation arise in the context of digital convergence. Indeed, as well stated by David Berman and by Paul Marty digital convergence should be intended as a convergence of institutions since now conceive as detached one from another. This process was mainly caused by the need to cope with a shared set of problems arise within the processes of digitalisation of society and to be able to satisfy the specific requests of the users. It is essential to consider that while maintaining on the back-end the traditional divisions among institution have proved to be the best solution to avoid that resources result decontextualised by their environments and related practices.

On the other hand, it is necessary to overcome those divisions at the moment in which these institutions want to offers their services to users through the digital medium. Indeed, in the front-end, the difference between practices and contexts provided by the different institutions does not play a fundamental role as in the backstage. It is undeniable that those different approach, processes and affinity

ecosystems maintain a crucial role to offer valuable services. Nevertheless, I argue that necessary to conceive that the majority the users which rely on these institutions are not interested in those aspects while the possibility to rely on a single service to access to multiple types of resources represents a very enticing perspective which will positively influence the relation and the approach of users to heritage resources. Therefore, it is possible to affirm that offering a unified platform to access a diversified set of resources has become a desirable and successful solution to satisfy users expectations. In the course of the previous chapter were presented some interesting P2P technologies and some examples of their application within the field of DLs. Those examples constitute the primary sources and base of the discussion developed in the next paragraphs. Indeed, the main aim of this chapter is to illustrate a proposal for a framework model to establish a multi-institutional ecosystem in which through the use of P2P technologies is offered to users the opportunity to access through one platform to several DLs' resources. Specifically, the proposed framework model it will be based in a conjunct use of Ripple blockchain and IPFS storage property to discuss the establishment of quality and functional ecosystem. Therefore, it is possible to define the presented model as a step in the direction of a new concept of museum: *a place where we can watch the practice develop itself, but also the moment when we want to be sure about what is the inherent logic of our mission in human society* (Šola, 1992, 394)

1. Framework[33]

Taking into account the technology analysed in this work and the project analysed in the last chapter, a framework is proposed that could be used within DLs to share and protect cultural heritage material in the digital form.

First, it is necessary to indicate the actors taken into account within this framework:

- **Nodes:** These form the main figures within the ecosystem. Each node will be run by every institution (library, university, museum, organisation, etc.) deciding to participate in the project. The nodes will form the grid of the system, fulfilling the role of maintaining the function of the ecosystem and offering their affiliates the opportunity to access the system.

[33] The use of the term framework has the role to clarify the fact that the model proposed in this research have not to be intended as complete and ready to be installed. Heritage Chains is a framework in the sense that it *'represents the cumulated experience of how the software architecture and its implementation for most applications in the domain should look like'* (Riehle 2000, 2)

- **Users:** Through their affiliation with an institution, people can access the ecosystem and the resources stored within it.
- **Contributors:** This term refers to users that actively contribute to the ecosystem through offering work or material they possess the right to for the digital chain archive. Each contributor maintains total control over his work, and in this regard, will be the only possessor of the private key of the resource.

In the framework, it was decided to relate each ecosystem node to each institution deciding to collaborate in the project. This decision is connected to the fact that these represent valid and largely trusted realities in their local context.
In addition to possessing adapt instruments and personnel to organise and manage these resources, the institutions employ long-established methods for organising realities and communities within the digital world. The account created by users within these institutions' digital counterparts could be used to ease the management of the ecosystem and to translate the identity verification, registration and access permission to the institutions, each of which will organise its own users. Thus, the same credential that each user possesses to access the institution (username and password) could be used to access the Heritage Chains, with the institution playing the role of intermediate and validator, equitable to that of the service provider.

2. Architecture: Core elements

To offer a description of the ecosystem architectures, it is first necessary to refer again to Chen *et al.* 2017. This work played a fundamental role in the development of this proposal, which takes advantage of the published work to establish a framework already tested and discussed by academics.

Focussing attention on a specific aspect of the digital world and aiming to solve specific issues analysed during the second chapter, for some elements, the proposed architecture diverges from their proposal. This proposed framework is divided into four layers (Fig. 1). In this sense, it could be defined as a modified version of the one proposed by Chen and his team during the 2017 IEEE International Conference on Big Data. Most of the core elements have not been subjected to any modification due to the fact that in this framework, institutions play the role of service providers enabling users to connect to the network. Nevertheless, while in Chen *et al.* 2017, the system allows single users to connect autonomously to the network, in order to offer a secure system for both resources and users, requiring each user's identity to

be related to an institution could provide the most effective solution to overcome any possible identity and security issues.

- **First Layer: Ledger**

This layer contains the ledger, which plays the role of a tamper-proof shared archive of the ecosystem's activities. As proposed by Chen *et al.* 2017, there are two possible ways to have and maintain a ledger. The first is to create a new one, and the second is to use one already established within a blockchain. Differently from their model, in this framework, the use of a pre-existing blockchain does not result in the best solution. Indeed, even if a well-established blockchain network could allow for considerable protection against malicious attack issues, such as that of the Byzantine fault, the ledger would also be full of unrelated data. Furthermore, considering that it was decided to use the Ripple protocol, the project member's decision would need to be completely autonomous.

Taking into account that the validation process is based on the democratic consensus of all participant nodes, if the ecosystem uses the Ripple public ledger as a base layer, it would be necessary to conceive that the nodes of the project would never be able to reach a number that allows them to ensure the authentication of the ecosystem operation. Even if approved by the supermajority of Heritage Chains participant nodes, it remains uncertain that the operation would be surely verified by other nodes of the network, thus creating serious inconveniences concerning the correct functioning of the ecosystem. Lucky, Ripple is an open-source project that allows everyone to use its protocol and technology for their own advantages.

Thus, it is possible to create a private network without any difficulties. When a node decides to create a new private ledger, all that is needed is to launch Ripple without the public server reference (--net), and a new ledger will be immediately created. Every time a server judges the server that created the new ledger as trustworthy, it will immediately follow the new ledger.

- **Second layer: Virtual chain**

The main role of a virtual chain could be compared to that of a virtual machine. Its role is to emulate the ledgers' activity and provide the opportunity to work on them without the risk of creating a problem in the actual ledger. In this sense, it offers considerable flexibility and security due to the fact that in case some operation

creates problems in the ecosystem, this would not influence the actual ledger, but only the virtual chain, which could be immediately reactivated or, in case of serious issues, a new one could be easily established while it is maintained and protected from any modification or collateral damages.

As in Chen *et al.* 2017, the virtual layer represents the core functional layer of the ecosystem. Its role is to check the validity of each operation executed within the system, such as resource uploads or access requests. Within the framework, using a system cryptographic key related to the resources could allow for verifying access to the resource for people who possess the public key and, at the same time, allow modification by the contributor possessing the resource's private key.

- **Third Layer: Routing**

Even if defined as an autonomous layer, routing comprises an integral part of the virtual chain layer. The routing operation's role is to hold the information of each node connected to the ecosystem, as well as all the resources connected to it. In this framework, each node relates to an institution, and in this sense, could be defined as trusted. However, the architectural conformation means that all the relative information could be traced back from the first and second layer, as indicated in Chen *et al.* 2017, 2655.

- **Fourth Layer: Storage layer**

The final layer is constituted by the IPFS, where all the resources shared with the system will be immediately uploaded. Even if shared within the IPFS, however, each resource uploaded in the ecosystem is asymmetrically encrypted, as demonstrated in Rajalakshmi *et al.* 2018, 1439. Because of this, it could not be accessed by the IPFS user because they do not possess the master key (public key) to open it. Only when a user makes a request within the ecosystem can they receive the master key to open these resources. Thus, even if resources are distributed within all the nodes forming the IPFS network, they remain protected from external and unauthorised use.

Throughout these data, the goal was to offer some core elements to discuss the development of the architecture of an ecosystem for DLs. As can be seen, the majority of the elements were acquired from the Chen *et al.* 2017 model, even if

some features have been modified and adapted in order to cope with the specific necessities and issues related to my project.

3. Sharing and access to the resources

The core of the proposed ecosystem activities involves sharing and accessing resources. This must be accomplishable by users with no difficulties. To ensure this, it was necessary to find the best solution to simplify the operation as much as possible. To this end, it was considered helpful to create a functional and intuitive user interface (UI), which could help people to organise their operation within the ecosystem and, at the same time, minimise the bandwidth overload, allowing everyone to access the system even with a low power connection

- **Upload file**

The first concern involves allowing everyone to easily upload a resource within the system. To this end, this research has taken inspiration from both the metadata model developed by the ETDB-Caltech project analysed in Ortega *et al.* 2018 and from the organisation and layout of different DLs.

The resultant framework (Fig. 2) can be interpreted as a draft of the possible UI for the resource-uploading phase. Establishing an adequate UI requires knowledge and practices not related to the theme of this research, however. This proposal aims solely to clarify the general features of this interface. Each contributor has to insert a structured series of data regarding the resource that it will upload in the ecosystem. This information will constitute the metadata of the resource and will be stored within the Ripple ledger. After a resource has been accepted to be uploaded in the ecosystem, it will immediately be encrypted, and using the combined data of the user and the uploaded resource, a pair of keys (public and private) will be created.

The final important element to take into account is the possibility for each contributor to decide the level of sharing for each uploaded resource. This could be public and available to each user of the system, or it could be private, which means that access could be limited to only the owner, to a single group, or to several groups, with only the users who are part of those groups being able to access the file.

- **Resource access**

As with the uploading phases, the access phase should also be simplified as much as possible in order to facilitate user interaction with the ecosystem. This also requires building a proper UI that is functional and intuitive for all users. As DLs and other institutions have already established valid UIs that could be used as an example for the Heritage Chains, it is possible to skip the discussion of this topic in the course of this work. Nevertheless, it remains necessary to analyse the operation conducted in order to allow users to access the resources stored in the ecosystem.
First, each user, after gaining access to the ecosystem, uses queries to research the resource in which he is interested. The queries send the user's request to an API, which starts to research the resource over the ecosystem. Using an API that is able to work with ontologies allows different types of research to be established (author, date, related library, related resources, etc.), and in this sense, a larger number of opportunities for retrieval can be offered to the users.

After finding the desired resource, the user sends a request of access to the ecosystem. This request is received by the virtual chain, which checks the legality of the request, and if the user has permission to access the resource, the machine immediately verifies the operation within the ledger. At the same time, the virtual chain retrieves the resource from the IPFS and sends an access link to the user together with the master key to decrypt the resource (fig.3).

4. Results

This section has proposed and described a framework for an ecosystem designed for DLs which store digital cultural heritage resources within an accessible platform in a secure, permanent and tamper-proof manner. Even if defined as a solely theoretical framework, Heritage Chains is able to highlight the new approach to heritage and Public Memory. Heritage Chains was principally conceiving to overcome the issues related to the fragmentation of heritage caused by the academic and institutional organisation. While in the back-end such fragmentation have a fundamental role thanks to the peculiar opportunities of investigation through the use of specific methodologies (museology, librarianship and archival science). In the front-end, the same fragmentation has severely constrained the opportunities of individuals to explore multiple types of heritage resources simultaneously, thus limiting the establishment of compressive approaches in which different heritage

resources might be used together in order to responded to the request of a better contextualisation in memory discussion (Šola 2003, 9-10)

Moreover, the process and results of this work are tied to already established projects and proposals, such as the ETDB-Caltech and the P2P File System Scheme based on the IPFS and blockchain by Chen et al. 2017, as well of that described in Rajalakshmi et al. 2018. At this point, it could be helpful to continue by discussing the opportunities of an ecosystem based on this framework.

To this end, the next section returns the discussion to the issues analysed during the second chapter in order to illustrate how this framework can cope with those issues and, in some cases, overcome problems that afflict the field of DLCH and DLs in general.

5. Analysis of opportunities

5.1. Storage and costs

As demonstrated in the previous sections, using the IPFS as a permanent Heritage Chains archive means using a well-established structure, which ensures several benefits. The first and most important is that the Heritage Chains resources are distributed all over the network, protecting them from being inaccessible. The IPFS is also content addressed, and as such, each resource receives an identifier hash that can be used to retrieve the file from all over the network. Institutions that already maintain the Heritage Chains resource in their storages do not have to modify their operation; each institution will continue to store the resources in the classical method, but at the same time, their encrypted version will be allowed to be shared over the IPFS. Thanks to the Heritage Chains' structure, this file can be retrieved and decrypted by the users. Moreover, each time users access and maintain the resources within their terminal, other users could acquire the resource from them, and in this sense, the bandwidth cost will be shared between the users, helping nodes to avoid overloading institution servers and decreasing the time required for operations such as requesting access and acquiring resources. Moreover, distributing resources over the network represents a key element for ensuring a high level of ecosystem reliability. In case the original node that has stored and shared the resource is unable to ensure the resource's availability (maintenance operations, loss of connection, censorship, etc.), all other possessors of the resource could continue to share it without causing any interference with the ecosystem activities.

A key element for discussing the feasibility of the Heritage Chains framework is the economical aspect. As stated previously, the institutions do not have to change their actual strategy regarding the sharing of material. It is possible for major institutions to help those institutions with fewer economic opportunities by offer their part of their storage space to ensure the reliability of their resources. Nevertheless, taking into account the consideration raised in the previous chapter, storage space is today a less problematic issue (at least for major institutions), and considering that the framework is proposed to connect as many institutions as possible, storage operations could be equally divided among all the members.

A second element that should be taken into account includes the costs related to the maintenance of the Ripple ledger. Taking into account that the proposed framework suggested establishing a Ripple private ledger, it is possible to simply refer to the minimum cost of the transaction in the Ripple ledger. Indeed, in Ripple, validators have the power to vote through consensus[34] to change transaction fees. In this sense, within the Heritage Chains, it is possible to maintain each transaction cost at its minimum level, which was settled at ten *drops* (0.00001XRP).[35]

Scalability comprises the last important element to take into account when discussing storage and costs. The previous section clarified how the opportunity of scalability has to been taken into account when establishing a new ecosystem. Thanks to the combined used of the IPFS and Ripple, the proposed framework is able to offer valuable opportunities for scalability. Indeed, implementing P2P technologies have been indicated to be the most adaptable solution to cope with the participant increase. Thanks to the features of two distributed systems, the workload can be divided equally among all the nodes that contributed together for the maintenance and routine operation. Actually, the framework proposed increasing participants instead of decreasing the general performance of the system, as this makes it possible to improve the user experience as well as the performance and reliability of the ecosystem. By increasing the number of institutions, and consequently, the nodes, the blockchain ledger can be contemporarily stored and run by a significant number of terminals, thereby ensuring an increased level of ecosystem security against errors or malicious attacks. Increasing the number of

[34] https://developers.ripple.com/fee-voting.html

[35] https://developers.ripple.com/transaction-cost.html. Taking into account that today, 1 XRP is exchanged for $0.44 USD (https://bitinfocharts.com/xrp/ accessed on 23 November 2018).

active users sharing and acquiring resources would allow the ecosystem to distribute the operation of sharing resources, reducing the risk of network congestion and nodes overload.

5.2. Protection and sharing

Protection and sharing represent the core function of every DL. As a brief description of the concept of a DL, this institution protects and shares a resource in the digital format. By operating within the digital world, DLs overcome the physical boundaries involved in providing access to everyone and everywhere, anytime.

Every institution that stores and shares resources in the digital world, such as museums, libraries, academic institutions, cultural centres and more can be defined as a DL. In some way, it is also possible to affirm that the objective of a total museum proposed by Tomislav Šola has been reached in contemporary society. However, even today, the enjoyment of knowledge and resources in the digital format are still not universality shared by humans all over the world. Boundaries such as political and economic restrictions have created discrepancy among the regions of the world. Resources are also not equally available worldwide, remaining tied to their location. Slow bandwidth power, distance from the server and economic limitations comprise only a part of the issues afflicting a large part of the globe, severely limiting the opportunity for a large percentage of humanity to access these resources. In addition, issues such as censorship and political control over the web access of entire nations further influence the current state of the digital world and impose a serious limit on the spread of culture and information called for by Šola.

Simultaneously, today's state of the web structure also presents issues related to protection. As observed by Dr Dietmar Offenhuber, web pages have the role of protecting and preserving heritage areas in danger as the material counterpart of the resource they aim to protect. Furthermore, as analysed by de Lusenet the estimated life of a web page is rather short (44 days), and even considering platforms that could offer significant conservation of time, they are still not able to ensure permanent protection and availability of the resource within the Internet, a fortiori, if HTTP's limitations are taken into account. Heritage Chains thus proposes

a new concept of protection and sharing[36] taking advantage of technologies that have demonstrated the ability to positively influence security and spread the resource within the digital world. Using a Ripple ledger structure to store information about resources (as a hash, metadata, etc.), and about the activities of users, makes it possible to ensure that the information of resources is not lost or tampered (by error or maliciously), increasing security for both the contributors and users. Moreover, the registration of each operation of the system in the ledger (in the form of transaction) could be used by each author to ensure that their right over the resource is respected, along with their intellectual property. In this regard, it could be useful to offer an example of a case of dispute regarding the use of a resource between a user and a contributor.

One such hypothetical scenario involves a contributor uploading a resource in the ecosystem, which is regulated by licence-free fruition for non-economical and academic purposes, but which also requires a licence to be acquired if this resource is used for profit. Every time a user accesses the resource, this operation is immediately recorded within the ledger, which is shared and the public. The information stored within the ledger, in addition to offering the contributor valid data regarding the use and fruits of his resource,[37] also offer to the contributor an unequivocal document allowing him to retrieve malicious users and assert his right over the resource. The transparent tamper-proof system, together with the public user identification, represents a valid disincentive for malicious activity within the Heritage Chains.

Using the IFPS as the storage layer for Heritage Chains makes it possible to offer several benefits to both users and institution participants. A resource identified through their hash (a unique resource identifier) could be traced back from all across the IPFS. Every time a user acquires a resource with his terminal, he immediately becomes an active contributor to the ecosystem's sharing process, allowing another

[36] Once again, protection and sharing are placed as a unique concept. This is due to the peculiarity of the area under discussion, which does not allow for these concepts to be examined as separate compartments, but rather requires this element to be analysed as unique in order to properly discuss and operate.

[37] It could be possible to provide each contributor with a specific historical archive for each individual resource, in which all accesses and requests of users and other such information are stored. On the other hand, this, like other features aimed at improving the user experience, can be omitted in the present work, as these can be developed within the proposed framework without modifying any fundamental aspect.

user to download chunks of the resource from his terminal. Conversely, implementing asymmetric encryption through GnuPG[38] also makes it possible to restrict the resource's fruition (decryption) to solely the Heritage Chains users that possess the public key (main key) of the resource.

Considering the contribution that the IPFS could offer for protection, in the IPFS, a resource can be stored persistently, as long as one or more users connected to the network possesses (within a cache folder) this resource. Simultaneously, it is also possible to affirm that resources are permanent (immutable), due to the fact that the IPFS is content addressing and each resource is identified with a URI (hash) related to the resource's intrinsic value. Modification of the resource causes the modification of its hash, and in the validation phase in which the hash of the resource is compared to the one stored in the Heritage Chains ledger, it is possible to immediately recognise any act of tampering. In addition to offering an irreproachable level of resilience, the features of the IPFS allow the Heritage Chains to establish censorship-resistant file storage.

Offering the example of the Wikipedia block operated by the Turkish government in April 2017 (Reuters 2017), it is possible to observe how the operation was accomplished without requiring particular effort from the government, which only had to, at all the country's connections, forbid access to the server in which the Wikipedia information was stored. The HTTP location-addressing operation has demonstrated to be particularly vulnerable to the operation of censorship, which can furthermore be executed easily by a central authority. Nevertheless, to respond to the government's action, a group of activists,[39] in collaboration with the protocol lab, immediately worked to establish a copy of the latest version of the Turkish Wikipedia over the IPFS.[40] This new version of Wikipedia provided a highly effective solution to bypass the Turkish censorship and forestall future censorship action.[41] This element could play an essential role in protecting and sharing resources put in practice within the Heritage Chains. The ecosystem is able to ensure that resources facing the threat of censorship (political, cultural and religious) could be shared by the contributor and remain available to

[38] https://gnupg.org/
[39] The Distributed Wikipedia Mirror is a global effort, independent from Wikipedia (https://github.com/ipfs/distributed-wikipedia-mirror)
[40] https://ipfs.io/ipns/tr.wikipedia-on-ipfs.org/wiki/Anasayfa.html
[41] https://ipfs.io/blog/24-uncensorable-wikipedia/

everyone. In this regard, the major institution could play a central role in the contrast, offering a permanent part of their storage to a resource under threat, ensuring that this could neither be lost nor destroyed.

This section strove to offer an overall picture of the features of this framework and to highlight opportunities related to its implementation. In addition to illustrating and clarifying the practical application of the framework, this analysis is able to highlight the role that the Heritage Chains could have in establishing an ecosystem capable of putting into practice the new concepts of museum and heritage offered by Šola. The Heritage Chains is not limited to a specific field; in this ecosystem, every type of data from every field of knowledge could be stored. In this sense, all kinds of institutions could participate[42] (universities, libraries, schools, museums, archives, cultural centres, community centres, scientific institutions, national and local institutions, etc.) in order to grant their users the opportunity to share and acquire data within the ecosystem.

6. Needs and Limits

While the last section analysed the opportunities related to the implementation of the Heritage Chains, it is necessary to point out and discuss the limits of the proposed framework in order to offer a complete picture on the aspects connected to this proposal.

A primary element that should be taken into account is the theoretical nature of this work. Even though it is based on projects already implemented or published, the proposed ecosystem has not been put into practice, and no proof of concept (PoC) has been developed. In addition, several elements would have to be considered and discussed before establishing a PoC for the proposed ecosystem. In this sense, the proposed framework should be understood as a theoretical basis, over which a discussion could be established regarding the new solutions for protecting and sharing digital resources offered by P2P branches. Moreover, the framework could be used as a foundation to continue to develop the concept established by Tomislav Šola and further analysed by Ross Parry, which was designed to

[42] Requests for participation can be assessed internally in the ecosystem. In fact, the possible candidatures can be evaluated by the institutions already within the ecosystem through a voting system.

revolutionise the concept of (digital) heritage and its relation with society and individuals.

A further consideration that should be taken into account with the proposed framework concerns its intrinsic need to be established over a large group of institutions and users. Establishing the Heritage Chains without an adequate number of institutions— and consequently, nodes—could undermine the intrinsic properties discussed in the previous chapter. Indeed, the characteristic of P2P ecosystem allows it to express its full potential only when a large number of nodes participate to the project. Operations such as RPCA require a specific number of users in order to be run correctly and avoid possible malicious operation or error. In this sense, a minimum of twenty nodes should be established to put the system into practice, and at least sixteen of these should be validated in order to ensure correct and efficient functioning of the ecosystem.

Conclusion

In the course of this research I have discussed the role which P2P technologies could play within the field of Digital heritage and specifically within the area of digital cultural heritage initiatives. In this sense, the project attention was focused within the context of digital libraries of which digital libraries of cultural heritage could be defined as a sub-branch.

Therefore, in the chapter I, after having defined and discussed the central concepts related to heritage and memory, the analysis was focused on digital cultural heritage initiatives and on the definitions of digital libraries of cultural heritage to differentiate these categories from the rest of digital libraries and, at the same time, was highlighted the role played by digital libraries within the field of heritage. Proceeding, aiming to offer an exhaustive outlook over the field of DLs, in the chapter II, I charted the evolution of the evolution of DLs field by outlining and investigating academic debates and establishing a parallel discussion with the changes occurring within the web context. Having retrieved the main issues that afflict the area of digital libraries (storage, costs, protection, sharing and access) and consequently that of digital libraries of cultural heritage, chapter III has developed to analyse these problematics and to discuss the solution offered by different academics. In the course of chapter IV, after a brief introduction on the P2P protocols, the analysis was focused on the presentation of both IPFS and Blockchain protocols which play a fundamental role in the discussion developed in the fifth chapter. In this chapter, some examples of the application of both IPFS and blockchain within the field of museums and digital libraries have been offered to highlight the opportunity provided by a joint implementation of these technologies. As discussed the combined use of this technology have shown to be able to offer new protection and sharing solution as well as overcoming problems related to storage and costs. In addition, P2P technology have proved to be Abe to cope with inequalities in the heritage field, like the management of, access to, and presentation of collections so often shaped by Western people, institutions, and organizations (archaeologists museums, anthropologists, UNESCO). Finally, in chapter V and last chapter, a framework model (Heritage Chains) for an ecosystem based on IPFS and blockchain was proposed. This ecosystem was primarily designed to discuss valuable solutions to increase the quality of services connected with DLCH field. Indeed, through the ecosystem proposed it is possible to implement the front-end

digital convergence discussed by P F. Marty (2009; 2014) and by D. Berman (1995). and consequently make a step forward the total museum forecasted by Tomislav Šola: *"is not a place, it is arelationship between past and present, between what we have been and what we want to be. It cannot be isolated and limited to an institution and distributed to the masses. Is an idea, and yet it is a form – a form of relationship*" (2005,13).

Therefore, an element that must be taken into account is the fact that the implementation of the Heritage Chains framework proposed in this work should not be conceived as limited by DLCH but could be also be implemented by other types of DCHI. As discussed by Diane Zorich (2003), the field of DCHI includes several kinds of institution connected to the area of heritage. The research has suggested that DLCH plays a fundamental role within the context of sharing and protecting heritage resources in digital form, nevertheless it was necessary to conceive an ecosystem able to include also other types of projects which could be qualified as digital cultural heritage initiatives (as online educational and scholarly databases or academic and heritage e-publishing).

In this sense, the framework proposed has been developed as a front-end solution able to include, without limiting, the different back-end approaches which characterise the different digital cultural heritage libraries projects. Therefore, throughout this research I aimed to offer the main elements as a general background, an analysis of the technologies (even if very simplistic) and a framework over which could be established a discussion on the implementation of Peer-to-Peer technologies in the context of sharing and protecting digital heritage resources.

Therefore, it is important to mention that this research never had the objective to offer a concluded plan ready to be implemented. A discussion regarding the operations and implementation of these Peer-to-Peer technologies over the several organization's procedures have to be developed to avoid possible issues. Peer-to-Peer technologies, as shown in the fourth chapter, are indeed able to offer a large amount of opportunity thanks to peculiar features such as decentralisation, distribution and protection which are able to revolutionise the processes of protection and sharing. Nevertheless, it should not be forgotten that the same opportunities, if misused, could actually have catastrophic consequences such as an incorrect use caused by negligence, misjudgements or malicious actions. These misuses could actually jeopardise or even undermine the authenticity of resources,

their protection, and the intellectual properties of their owners. Undertaking a strict evaluation of the need as well as the added value of these P2P technologies is consequently the best solution to explore and proceed cautiously towards future implementations.

On the other hand, the effort required to investigate these technologies within the field of heritage will be repaid by the new opportunities for access, collaboration, protection and establishment of an ecosystem able to cope with the requests of contemporary society. Being able to sustain a dialogue between people worldwide has been intended as the primary objective connected to the implementation of these P2P technologies within the field of heritage. Thus, it is reductive to conceive Heritage Chain just as a digital ecosystem discussed to propose new techniques of sharing and protection. Heritage Chain is aiming also to create an increased interconnection among institutions of different countries to share capacity expenses and experiences and consequently to increase both the quality and the scope of their services. Moreover, this research was developed to propose a new way of conceiving the relation between users and digital heritage resources. Through the proposed framework, the discussion of a new and democratic connection with the heritage resources in digital spaces was foreseen. Overcoming issues such as censorship, digital international inequities and presenting a more inclusive environment in which everyone could be part and plays a role as a user, contributor or node (Chapter V section 1).

Such elements are able to highlight the role which the themes and discussions developed in this research are able to acquire within the heritage field. Heritage Chain is not just a mere proposal for including new technologies within the field of DLs, rather it should be understood as a step in the direction of the establishment of a total heritage reality and a Total Museum: *'a territory "covered" by awareness of its character by a sensitivity and knowledge that, because it is so complete, requires no assistance of institutions.'* (Šola 2015, 103).

Abstract

In the field of digital cultural heritage, digital libraries have played the role of preserving and sharing cultural material in digital format. These tools (portals, collections and archives) have been influenced by the evolution of the digital medium and by new technologies which, over the years, have been provided both by the areas defined as neighbours and by conceptually distant environments.

The birth of the web has played a fundamental role in today's concept of digital library, in fact this tool has allowed us to increase the scope of these collections and greatly improve their functions. In this sense it is possible to observe how digital libraries have evolved together with the web and how they have always exploited the technologies offered by cyberspace. This improves the possibilities offered to specialists in the sector and those who rely on these collections, their study, or simply to increase their cultural baggage.

These collections however are affected by numerous problems that undermine not only the possibilities offered to users and specialists but also their main functions of sharing and protection. Over the last few years, several specialists and academics have focused their attention on the analysis of these solutions. Nevertheless, many possible solutions have been offered in the past and today. This area suffers from various problems that undermine the operations connected to these collections.

The objective of this paper is to analyse the evolution of digital libraries in order to offer an analysis of the aforementioned problems and, as a consequence, to suggest a possible solution. Exploiting the new technologies offered by the digital medium is able to resolve, or at least curb, the weaknesses that grip this area. Through the proposal of an ecosystem for digital libraries of cultural heritage and digital cultural heritage initiatives based on the joint use of IPFS and the Ripple ledger, this work aims to offer a solution able to overcome the critical issues analysed during this study. Yet the proposed ecosystem should not be understood in any way as a conclusive point on the question but the main objective is to offer ideas. On these elements, we may build a debate on the potential of these technologies and their possible application within the field of digital libraries.

Appendix

- **Acronyms**

ADSL	Asymmetric Digital Subscriber Line
API	Application programming interface
ARMA	Association of Records Managers and Administrators
CLIR	Council on Library and Information Re-sources
CWA	Closed-World Assumption
DDoS	Distributed Denial of Service
DHT	Distributed Hash Table
DL	Digital library
DLI	Digital Library Initiative
DNS	Domain Name System
ETDB-Caltech	Electron Tomography Database- Caltech
FOAF	Friend Of A Friend (Ontology)
HCI	Human Computer Interaction
HTTP	Hypertext Transfer Protocol
IaaS	Infrastrutture as a Service
ICCU	Istituto Centrale per il Catalogo Unico
IEEE	Institute of Electrical and Electronics Engineers
IFLA	International Federation of Library Associations
ILL	Inter-Library Loan
IPFS	InterPlanetary File System
IPFS	Interplanetary File System
IPO	Initial public offering
ITC	Information and Communications Technology
LIS	Library and Information Studies
NIST	National Institute of Standards and Technology
OA	Open Access
OIP	Open Index Protocol
OPAC	On-line public access catalogue
OSDA	On-site Digital Archaeology
OWL	Web Ontology Language

P2P	Peer-to-peer
PoC	Proof of Concept
PoW	Proof of Work
RDF	Resource Description Framework
SFS	Self-Certifying File System
SGML	Standard Generalised Markup Language
SQL	Standard Query Language
TIFF	Tagged Image File Format
UNL	Unique Node List
URL	Uniform Resource Locator
UVC	Universal Virtual Computer
VCS	Version Control System
W3C	World Wide Web Consortium
WIPO	World Intellectual Property Organization
XML	eXtensible Markup Language

- **Figures**

For the purpose of this research I have created this figures (diagrams). The figures have the role to facilitate the compression of the framework proposed. In order to create this figures have been used some free icons acquired from Flaticon flaticon.com. All the rights of the icons are reserved to their respective authors.

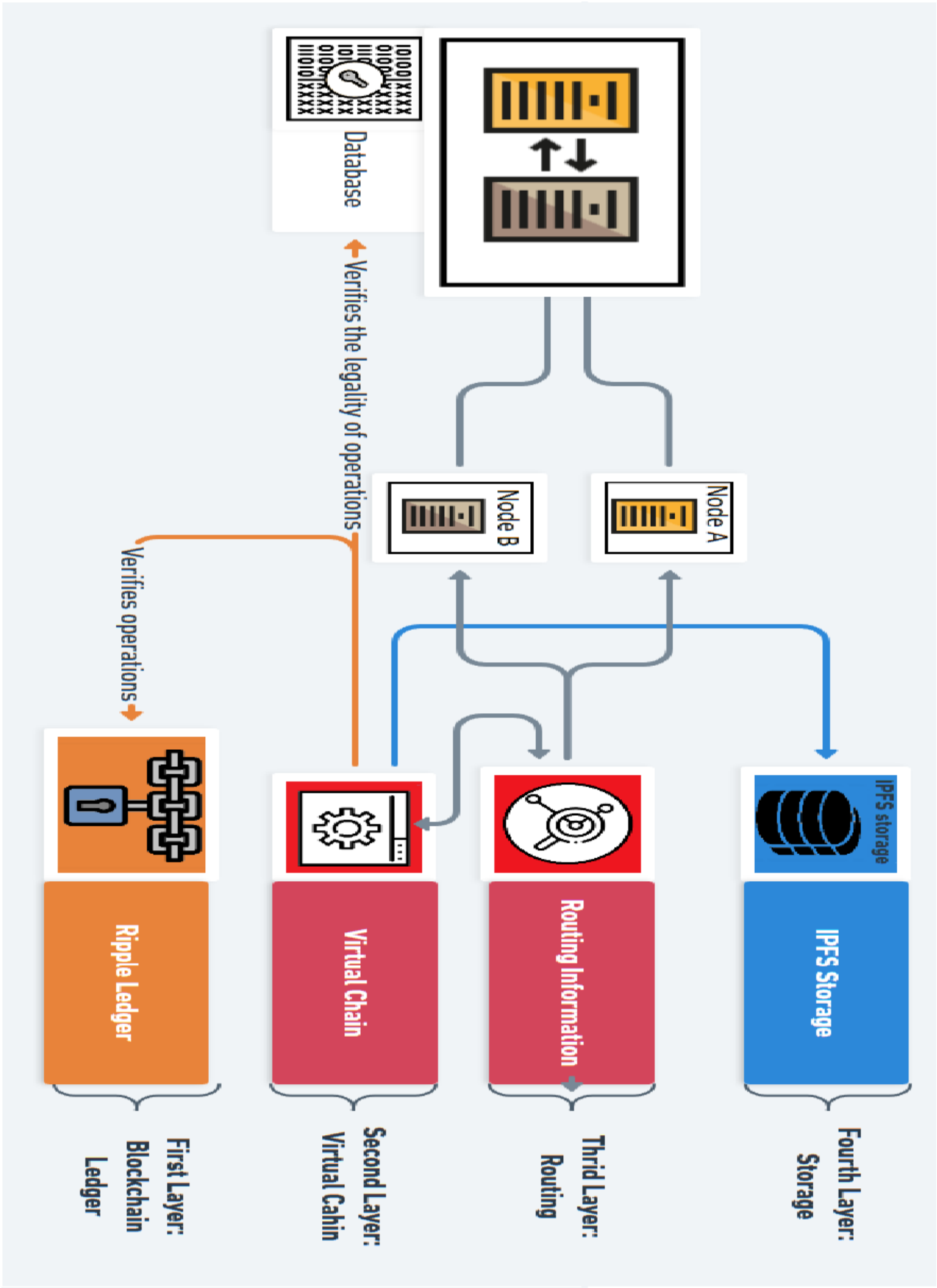

Figure 1: Ecosystem architecture

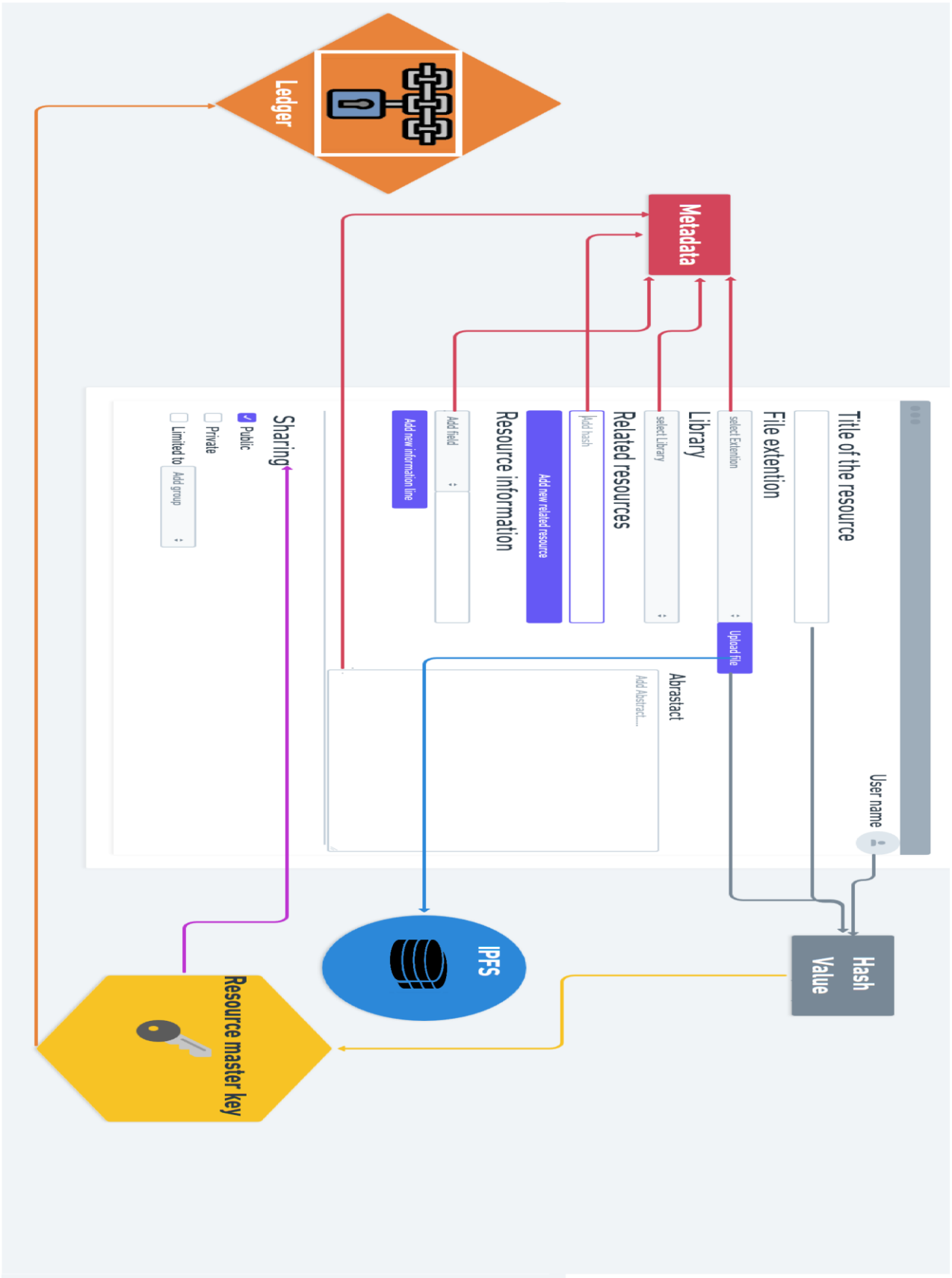

Figure 2: Upload resource operation

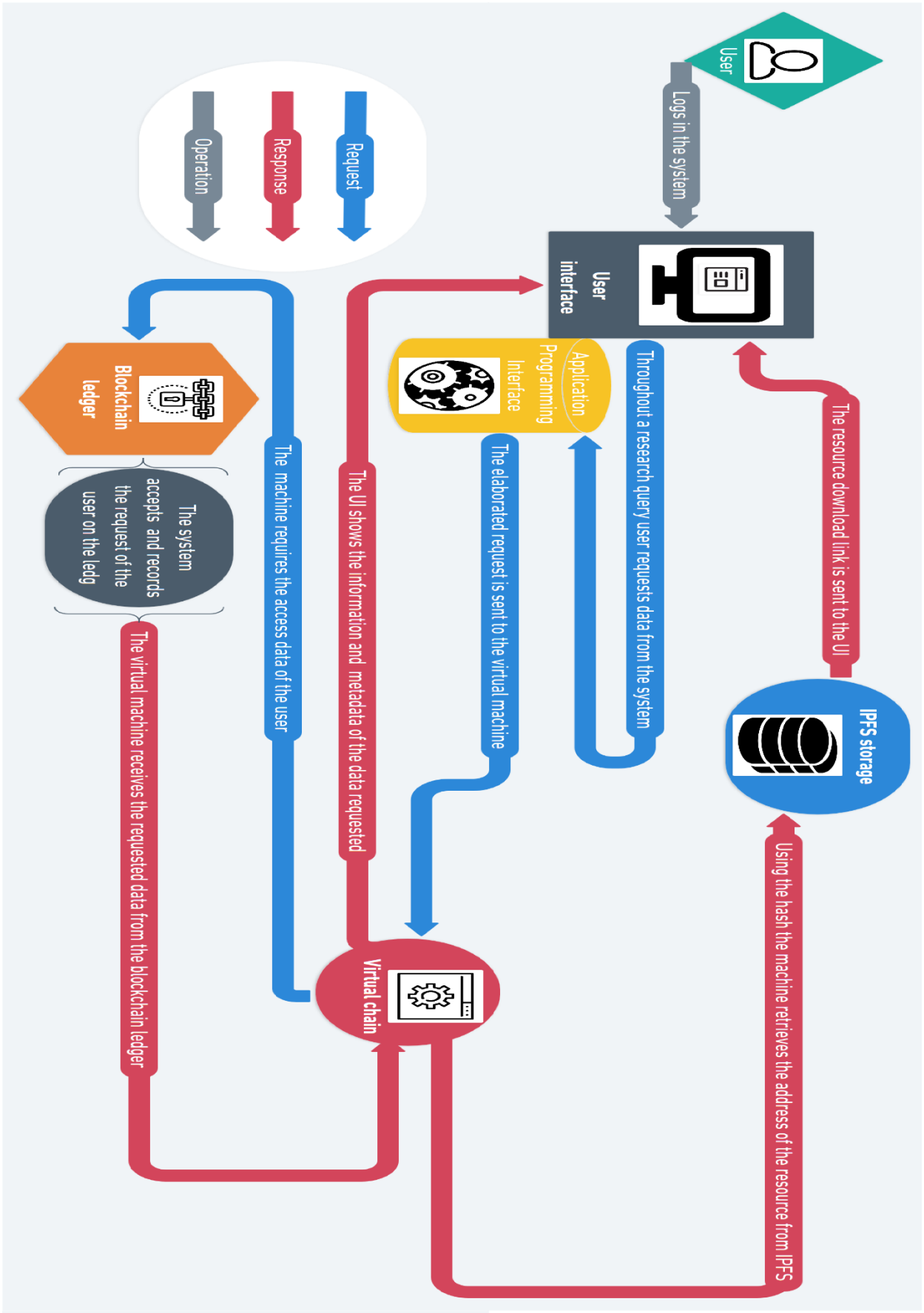

Figure 3 Resources access operation

Bibliography

Abbas, J., K. Kim, B. Wildemuth, Y. Choi, J. Mostafa, K. Brancolini, J. Pomerantz and A. Clobridge, 2006. Education for digital librarianship: Employers' needs and how they can be addressed, in Proceedings *of the American Society for Information Science and Technology* 43(1), 1-7. https://doi.org/10.1002/meet.14504301321

Alotaibi, S., 2010. Semantic Web Technologies for Digital Libraries: From Libraries to Social Semantic Digital Libraries (SSDL), Over Semantic Digital Libraries (SDL), in *The 4th Saudi International Conference. 30 - 31 Jul 2010.*

American National Standard for Telecommunications, 1998. *Network and Customer Installation Interfaces –Asymmetric Digital Subscriber Line (ADSL)Metallic Interface* (T1.) 413 (2). 413-1998.

Anderson, W., 2004. Some challenges and issues in managing, and preserving access to, long-lived collections of digital scientific and technical data. *Data Science Journal* 3, 191-201. https://doi.org/10.2481/dsj.3.191

Ashworth, G., 1994. From history to heritage–from heritage to identity, in Larkham, P.J. and G. J. Ashworth, (eds), *Building a New Heritage: Tourism, Culture and Identity in the New Europe*. London: Routledge, 13-30.

Assmann, J. and Czaplicka, J., 1995. Collective memory and cultural identity. *New german critique*, (65), 125-133. https://doi.org/125-133.10.2307/488538

Bailey, C.., 2007. *What is open access?* in Jacobs, N., (ed), *Open Access: Key Strategic, Technical and Economic Aspects*. Oxford: Chandos Publishing, 13-26

Barassi, V. and E. Treré, 2012. Does Web 3.0 come after Web 2.0? Deconstructing theoretical assumptions through practice. *New Media & Society* 14(8), 1269-1285. https://doi.org/10.1177/1461444812445878

Barr, J., B. Bradley, and B.T. Hannigan, 2003, Using digital watermarks with image signatures to mitigate the threat of the copy attack. In *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP'03). 2003 IEEE International Conference on* 3, 69-72. https://doi.org/10.1109/ICASSP.2003.1199109

Barthelemy, J., 2001. The hidden costs of IT outsourcing: lessons from 50 IT-outsourcing efforts show that unforeseen costs can undercut anticipated benefits. Understanding the issues can lead to better outsourcing decisions. *MIT Sloan management review*, *42*(3), 60-69.

Bearman, D., 1995. Standards for Networked Cultural Heritage. *Archives and Museum Informatics* 9 (3), 279-307. https://doi.org/10.1007/BF02770469

Benet, J., 2014. IPFS - Content Addressed, Versioned, P2P File System (DRAFT 3).

Biesecker, B., S.H. Browne, E.S. Casey, R.A. Eberly, H. Heinrich, A. Kiewe, C. E. Morris, B. Schwartz, C. E. Scott, and K.R. Phillips 2004. *Framing Public Memory*. Tuscaloosa: University of Alabama.

Borgman, C.L., 1999. What are digital libraries? Competing visions. Information Processing and Management 35, 227-243. https://doi.org/10.1016/S0306-4573(98)00059-4

Burgess, J. and J. Green, 2018. *YouTube: Online video and participatory culture*. John Wiley & Sons.

Cameron, F. 2007. The politics of heritage authorship. The case of digital heritage collections. in Kalay, Y., Kvan, T. and Affleck, J. eds. *New heritage: New media and cultural heritage*. London: Routledge, 170-184.

Candela, L., D. Castelli, N. Ferro, Y. Ioannidis, G. Koutrika, C. Meghini, P. Pagano, S. Ross, D. Soergel, M. Agosti, M. Dobreva, V. Katifori, and H. Schuldt, 2007. *The DELOS Digital Library Reference Model. Foundations for Digital Libraries*. ISTI-CNR.

Chen,Y., H. Li, ,K. Li, and J. Zhang, 2017. An improved P2P file system scheme based on IPFS and Blockchain, in *Big Data 2017 IEEE International Conference on*, 2652-2657. https://doi.org/10.1109/BigData.2017.8258226

Choi, Y. and Rasmussen, E., 2009. What qualifications and skills are important for digital librarian positions in academic libraries? A job advertisement analysis. *The journal of academic librarianship*, *35*(5), 457-467. https://doi.org/10.1016/j.acalib.2009.06.003

Choudhury, N., 2014. World wide web and its journey from web 1.0 to web 4.0. *International Journal of Computer Science and Information Technologies*, *5*(6), 8096-8100.

Chowdhury, G.G. and S. Foo, 2012, Digital Libraries and Open Access. in in Chowdhury, G. and S. Foo, 2012, *Digital libraries and information access: Research perspectives*. United Kingdom: Facet Publishing. 143-154

Collard, A., 1989. Investigating 'social memory' in a Greek context in Tonkin, E., M. McDonald, and M. Chapman. *History and ethnicity*. London: Routledge, 89-103.

Costa, M., D. Gomes and M. Silva, 2017. The evolution of web archiving. *International Journal on Digital Libraries* 18(3), 191-205. https://doi.org/10.1007/s00799-016-0171-9

Crane, G. and C. Wulfman, 2003. Towards a cultural heritage digital library, in IEEE Computer Society, in *Proceedings of the 3rd ACM/IEEE-CS joint conference on Digital libraries*, 75-86 https://doi.org/10.1109/JCDL.2003.1204845

Crane, G., 2002. Cultural Heritage Digital Libraries: Needs and Components, in Agosti M., Thanos C. (eds) *Research and Advanced Technology for Digital*

Libraries. ECDL 2002. Lecture Notes in Computer Science 2458. Heidelberg: Springer, Berlin, 626-637.

de Lusenet, Y., 2002. Preservation of digital heritage. *Draft Discussion Paper Prepared for UNESCO*. https://doi.org/10.1353/lib.2007.0053

DuCharme, B., 2013. *Learning SPARQL: querying and updating with SPARQL 1.1*. Sebastopol: O'Reilly Media, Inc.

Economou, M. and Pujol, L., 2007. Educational tool or expensive toy? Evaluating VR evaluation and its relevance for virtual heritage. in Kalay, Y., Kvan, T. and Affleck, J. (eds). *New heritage: New media and cultural heritage*. London: Routledge, 242-260.

Fentress, J., C. Wickham and American Council of Learned Societies, 1992. *Social memory* (New perspectives on the past). Oxford; Cambridge (MA): Blackwell.

Forster, H., 2007. The i2010 digital libraries initiative: Europe's cultural and scientific information at the click of a mouse. *Information Services & Use* 27(4), 155-159. https://doi.org/ 10.3233/ISU-2007-2740

Foucault, M., 2012. *Discipline and punish: The birth of the prison*. Vintage.

Fox, E.A., 1999. Digital Libraries Initiative (DLI) Projects 1994-1999. *Bulletin of the American Society for Information Science and Technology* 26(1), 7-11. https://doi.org/10.1002/bult.135

Freedman, M.J., E. Freudenthal, and D. Mazieres, 2004. *Democratizing Content Publication with Coral*. *NSDI* 4, 1-18.

French, S., 1995. What is social memory? *Southern Cultures*, 2(1), 9-18. https://doi.org/10.1353/scu.1995.0049

García-Crespo, Á., J.M. Gómez-Berbís, R. Colomo-Palacios, and F. García-Sánchez, 2011. Digital libraries and web 3.0. The CallimachusDL approach. *Computers in Human Behavior* 27(4), 1424-1430. https://doi.org/10.1016/j.chb.2010.07.046

Hall, S., 2005. Whose heritage? Un-settling 'the heritage', re-imagining the post-nation. Naidoo. *The Politics of Heritage: The Legacies of 'Race'*. Taylor & Francis. 23-35.

Hanna, N.K. (ed), 2016. Mastering digital transformation: Towards a smarter society, economy, city and nation, in Hanna, N.K. (ed), *Mastering Digital Transformation: Towards a Smarter Society, Economy, City and Nation*. Emerald Group Publishing Limited.

Harrison, R., 2010. What is heritage, in Harrison, R. (ed), *Understanding the politics of heritage*. Manchester: Manchester University Press, 5-42.

Hedstrom, M., 1997. Digital preservation: a time bomb for digital libraries. *Computers and the Humanities* 31(3), 189-202. https://doi.org/10.1023/A:1000676723815

Hill, M. 1990. What is scalability? *ACM SIGARCH Computer Architecture News* 18(4), 18-21. https://doi.org/10.1145/121973.121975

Howard, K., 2010. Programming not required: skills & knowledge for the digital library environment. *Australian Academic & Research Libraries* 41(4), 260-275. https://doi.org/10.1080/00048623.2010.10721480

Hoy, M.B., 2017. An introduction to the Blockchain and its implications for libraries and medicine. *Medical reference services quarterly* 36(3), 273-279. https://doi.org/10.1080/02763869.2017.1332261

Hull, D., S.R. Pettifer, and D.B. Kell, 2008. Defrosting the digital library: bibliographic tools for the next generation web. *PLoS computational biology* 4(10). https://doi.org/10.1371/journal.pcbi.1000204

Ioannidis, Y., 2005. Digital libraries at a crossroads. *International Journal on Digital Libraries* 5(4), 255-265. https://doi.org/10.1007/s00799-004-0098-4

Jokilehto, J., 2005. Definition of cultural heritage: References to documents in history. *ICCROM Working Group 'Heritage and Society*, 4-8.

Kalay, Y., 2007. Introduction: Preserving cultural heritage through digital media in Y. E. Kalay, T. Kvan, and J. Affleck, (eds), *New heritage: New media and cultural heritage*. London: Routledge, 17-26.

Kallinikou, D., J.M. Papadopoulos, A. Kaponi, and V. Strakantouna, 1993. Intellectual Property issues for digital libraries in the Internet networked public sphere, in *8th International Conference Computer Ethics: Philosophical Enquiry* 334. http://dx.doi.org/10.2139/ssrn.2619734

Kassoff, M., C. Petrie, L.M. Zen, and M. Genesereth, 2009. Semantic Email Addressing: The Semantic Web Killer App?. *IEEE Internet Computing* 13(1). 48-53 http://dx.doi.org/10.1109/MIC.2009.20

Kirschenbaum, M., R. Ovenden, G. Redwine, and R. Donahue, 2010. *Digital forensics and born-digital content in cultural heritage collections*. Washington, D.C: Council on Library and Information Resources.

Koontz, C. and B. Gubbin, (eds), 2010. *IFLA public library service guidelines* 147.Berlin: De Gruyter.

Kruk, S.R., Decker, S. and Zieborak, L., 2005. JeromeDL-Reconnecting Digital Libraries and the Semantic Web. in Andersen K.V., Debenham J., Wagner R. (eds) Database and Expert Systems Applications. DEXA 2005. Lecture Notes in Computer Science, vol 3588. Berlin, Heidelberg: Springer. https://doi.org/10.1007/11546924_70

Kruk, S. R., B. McDaniel, (eds). 2008. *Semantic digital libraries*. Heidelberg: Springer Verlag.

Kruk, S.R., S. Decker, and L. Zieborak, 2005. JeromeDL–adding semantic web technologies to digital libraries, in *International Conference on Database and Expert Systems Applications.* Springer, Berlin, Heidelberg, 716-725. https://doi.org/10.1007/11546924_70

Kuan, C., 2014. 7 Lessons Learned for Digital Culture in Din, H.,and S. Wu (eds), *Digital Heritage and Culture*. London: World Scientific, 41-52.

Leiner, B., 1998. The scope of the digital library. in *DLib Working Group on Digital Library Metrics*.

Liritzis, I., F.M. Al-Otaibi, P. Volonakis, and A. Drivaliari, 2015. Digital technologies and trends in cultural heritage. *Mediterranean Archaeology & Archaeometry* 15(3), 313-332. https://doi.org/10.5281/zenodo.33832

Lorie, R. and R. van Diessen, 2005. UVC: A universal virtual computer for long-term preservation of digital information. *IBM Res. rep. RJ*, 10338. https://doi.org/10.1007/978-3-642-27549-4_20

Lynch, C., 2002. Digital Collections, Digital Libraries & the Digitization of Cultural Heritage Information. *Microform & Imaging Review*, 31(4), 131-145. https://doi.org/10.1515/MFIR.2002.131

Lynch, C. and H. Garcia-Molina,1996. Interoperability, scaling, and the digital libraries research agenda. *Microcomputers for Information Management*, *13*(2), pp.85-132.

Marty, P.F., 2008. Museum websites and museum visitors: digital museum resources and their use. *Museum Management and Curatorship* 23(1), 81-99. https://doi.org/10.1080/09647770701865410

Marty, P.F., 2009. An introduction to digital convergence: libraries, archives, and museums in the information age. *Museum Management and Curatorship*, *24*(4), 295-298. http://dx.doi.org/10.1080/09647770903314688

Marty, P.F., 2014. Digital convergence and the information profession in cultural heritage organizations: Reconciling internal and external demands. *library trends*, *62*(3), 613-627. https://doi.org/10.1353/lib.2014.0007

Maymounkov, P. and D. Mazieres, 2002, March. Kademlia: A peer-to-peer information system based on the xor metric, in *International Workshop on Peer-to-Peer Systems*, 53-65. https://doi.org/10.1007/3-540-45748-8_5.

McDowell, S., 2016. Heritage, memory and identity, in Graham, B.J. and P. Howard (eds), *The Ashgate research companion to heritage and identity*. Ashgate Publishing, Ltd, 37-53

Mell, P. and T. Grance, 2011. The NIST definition of cloud computing. *Association for Computing Machinery. Communications of the ACM* 53(6), 50-53. https://doi.org/10.12691/jcsa-2-3-1

Muir, A. 2006. *Preservation, Access and Intellectual Property Rights Challenges for Libraries in the Digital Environment*. London: Institute for Public Policy Research.

Nelson, M.L., F. McCown, J.A. Smith, and M. Klein, 2007. Using the web infrastructure to preserve web pages. *International Journal on Digital Libraries* 6 (4), 327-349. https://doi.org/10.1007/s00799-007-0012-y

Nisheva-Pavlova, M., D. Shukerov, and P. Pavlov, 2015. Design and implementation of a social semantic digital library. *Information Services & Use* 35(4), 273-284. https://doi.org/10.3233/ISU-150784

Ortega, D.R., C.M. Oikonomou, H.J Ding,. P. Rees-Lee, and G.J Jensen,. 2018. *ETDB-Caltech: a blockchain-based distributed public database for electron tomography*. https://doi.org/10.1101/453662

Parry, R., 2005. Digital heritage and the rise of theory in museum computing. *Museum management and Curatorship* 20(4), 333-348. https://doi.org/10.1080/09647770500802004

Rahaman, H. and B.K. Tan, 2010. Interpreting digital heritage: Considering end-user's perspective, in *15th International Conference on Computer Aided Architectural Design Research in Asia, CAADRIA*, 93-102.

Rahaman, H., 2018. Digital heritage interpretation: a conceptual framework. *Digital Creativity*, *29* (2-3), 208-234. https://doi.org/10.1080/14626268.2018.1511602

Rajalakshmi, A., K. V. Lakshmy, M. Sindhu and P. P. Amritha, 2018. A Blockchain and IPFS based framework for secure Research record keeping. *International Journal of Pure and Applied Mathematics* 119 (15), 1437-1442.

Rattus, K., 2011. Which Heritage? Which Landscape? Defining the Authenticity of Cultural Heritage in Karula National Park. *Journal of Ethnology and Folkloristics* 1(1-2), 131-143.

Riehle, D., 2000. *Framework design: A role modeling approach* (Doctoral dissertation, ETH Zurich). https://doi.org/10.3929/ethz-a-003867001

Ruan, J. and McDonough, J.P., 2009, August. Preserving born-digital cultural heritage in virtual world. In *2009 IEEE International Symposium on IT in Medicine & Education* 1, 745-748. https://doi.org/10.1109/ITIME.2009.5236324

Sameti Pour, A. 2017. *Ethnographic Study of Facebook Usage among the Persian and Arabic- Speaking Refugees and Asylum seekers in Asylum Center*, Utrecht (unpublished Master thesis Utrecht University).

Sanchati, R., 2011. Cloud computing in digital and university libraries. *Global Journal of Computer Science and Technology* 11(12), 36-42.

Santana-Quintero, M. and A.C. Addison, 2007, September. Digital tools for heritage information management and protection: The need of training, in *International Conference on Virtual Systems and Multimedia*, 35-46. https://doi.org/10.1007/978-3-540-78566-8_4

Sastry, H.G., G. Manjunath, and L.C. Reddy, 2011. User interface design challenges for digital libraries. *International Journal of Computer Applications* 15(6), 7-13. https://doi.org/1010.5120/1955-2616

Savage, S.J. and D. Waldman, 2005. Broadband Internet access, awareness, and use: Analysis of United States household data. *Telecommunications Policy*, *29*(8), 615-633. https://doi.org/10.1016/j.telpol.2005.06.001

Schollmeier, R., 2001. A definition of peer-to-peer networking for the classification of peer-to-peer architectures and applications, in *Peer-to-Peer Computing, 2001. Proceedings. First International Conference on*, 101-102. https://doi.org/10.1109/P2P.2001.990434

Schwartz, D., N. Youngs, and A. Britto, 2014. The Ripple protocol consensus algorithm. *Ripple Labs Inc White Paper* (5).

Shadbolt, N., T. Berners-Lee, and W. Hall, 2006. The semantic web revisited. *IEEE intelligent systems* 21(3), 96-101. http://doi.ieeecomputersociety...IS.2006.62

Shivalingaiah, D. and U. Naik, 2008. Comparative Study of Web 1.0, Web 2.0 and Web 3.0, in *International CALIBER-2008.* https://doi.org/10.13140/2.1.2287.2961

Skinner, J., 2010. Learning from the past: Digitization and information loss. *B Sides*. 1-7.

Šola, T., 1992. The future of museums and the role of museology. *Museum Management and Curatorship* 11(4), 393-400. https://doi.org/10.1080/09647779209515332

Šola, T., 2005. What theory? What heritage? Some excerpts from the current book project on heritology. *Nordisk museologi*, (2), 3-16. http://dx.doi.org/10.5617/nm.3311

Šola, T., 2015. Mnemosophy: An Essay on the Science of Public Memory. Zagreb: European Heritage Association.

Sompel, H.V.D. and C. Lagoze, 2000. The Santa Fe convention of the open archives initiative. *D-Lib Magazine* 6(2). https://doi.org/10.1045/february2000-vandesompel-oai

Stribling, J., Councill, I.G., Li, J., Kaashoek, M.F., Karger, D.R., Morris, R. and Shenker, S., 2005. OverCite: A cooperative digital research library, in *International Workshop on Peer-to-Peer Systems*. Berlin, Heidelberg: Springer, 69-79.

Szabo, N., 1997. Formalizing and securing relationships on public networks. *First Monday* 2(9). https://doi.org/10.5210/fm.v2i9.548

Tang, Y., L. Zhou, J. Cao, J. Li, and X. Nie, 2018. Integration of Digital Cultural Heritage Resources in China: Understanding Public Expectations. *Libri: International Journal of Libraries and Information Services*, *68*(1), 59-70. https://doi.org/10.1515/libri-2017-0027

Weech, T., 2007, Competencies Needed for Digital Librarianship. *Digital Library for the Maghreb Workshop. Retrieved April* 23, 2-9.

Zorich, D.M., 2003. *A Survey of Digital Cultural Heritage Initiatives and Their Sustainability Concerns. Managing Economic Challenges*. Washington, DC: Council on Library and Information Resources.

List of internet page

Akamai 2017. *State of internet*, Akamai.com 2017, http://www.akamai.com/uk/en/about/our-thinking/state-of-the-internet-report/, accessed on 6 December 2018.

Aksamija, A., 2016. *Memory Matrix*, mit.edu 2016, http://www.architecture.mit.edu/art-culture-and-technology/project/memory-matrix, accessed on 6 December 2018.

American Memory 2018. *Mission and History*, memory.loc.gov, http://memory.loc.gov/ammem/about/index.html, accessed on 27 January 2018.

Benet, J., 2014. IPFS-content addressed, versioned, P2P file system. *arXiv preprint arXiv:1407.3561*, Arxiv.org, 2014, https://arxiv.org/pdf/1407.3561.pdf, accessed on 6 December 2018.

Buterin, V., 2013. *Ethereum White Paper a next generation smart contract & decentralized application platform, cryptorating.eu 2013*, https://cryptorating.eu/whitepapers/Ethereum/Ethereum_white_paper.pdf, accessed on 6 December 2018.

Daintith, J. and E. Wright, 2008. *A dictionary of computing*. Oxford University Press, Inc., oxfordreference.com 2008, http://www.oxfordreference.com.ezproxy.leidenuniv.nl:2048/view/10.1093/acref/9780199234004.001.0001/acref-9780199234004-e-2326?rskey=7QeYHy&result=2522, accessed on 6 December 2018.

EU commission, 2016. *Cultural heritage Digitisation, online accessibility and digital preservation. Report on the Implementation of Commission Recommendation 2011/711 / EU*, ec.europa.eu 2016, http://ec.europa.eu/information_society/newsroom/image/document/2016-43/2013-2015_progress_report_18528.pdf, accessed on 6 December 2018.

Flood, A. 2009, *Free-access World Digital Library set to launch,* theguardian.com, https://it.search.yahoo.com/search?fr=mcasa&type=E110IT91082G91207&p=Free-access+World+Digital+Library+set+to+launch, accessed on 27 January 2019.

Hern, A. 2018. *Energy cost of 'mining' bitcoin more than twice that of copper or gold*, theguardian.com 2018, http://www.theguardian.com/technology/2018/nov/05/energy-cost-of-mining-bitcoin-more-than-twice-that-of-copper-or-gold, accessed on 6 December 2018.

IFLA 2018. *IFLA/UNESCO Manifesto for Digital Libraries*, ifla.org 2018, https://www.ifla.org/publications/iflaunesco-manifesto-for-digital-libraries, accessed on 6 December 2018

Komorowski, M., 2009. *a history of storage cost,* mkomo.com 2009, http://www.mkomo.com/cost-per-gigabyte, accessed on 6 December 2018.

Lesk. M., 2000. *Preservation of New Technology*, Lesk.com 1992 www. http://lesk.com/mlesk/cpa/cpa2.html, accessed on 6 December 2018.

Lesk. M., 2005. *How to Pay for Digital Libraries*, Lesk.com 2002, http://www.lesk.com/mlesk/derek02/derek2.html, accessed on 6 December 2018.

Nakamoto, S. 2008. *Bitcoin: A Peer-to-Peer Electronic Cash System*, Bitcoin.org 20089, http://www.bitcoin.org/bitcoin.pdf, accessed on 6 December 2018.

National Library of Australia, 2003. Guidelines for the preservation of digital heritage, unesdoc.org 2003, https://www.unesdoc.unesco.org/ark:/48223/pf0000130071, accessed on 6 December 2018.

O' Reilly, T. 2005. Web 2.0: Compact Definition, Radar 2005, http://www.radar.oreilly.com/2005/10/web-20-compact-definition.html, accessed on 6 December 2018.

Offenhuber, D., 2016. *Cryptographic Heritage*, *memorymatrix.org* http://www.memorymatrix.org/cryptographic-heritage%e2%80%8a-%e2%80%8ausing-bitcoin-to-store-, accessed on 6 December 2018.

Reuters 2017. *Turkey blocks Wikipedia under law designed to protect national security*, theguardian.com 2017, http://www.theguarian.com/world/2017/apr/29/turkey-blocks-wikipedia-under-law-designed-to-protect-national-security, accessed on 6 December 2018.

Ripple 2017, *Validators and Unique Node Lists,* XRP ledger developers portal 2017, http://www. developers.ripple.com/technical-faq.html, accessed on 6 December 2018.

Shaheed, F., 2011. *Access to cultural heritage as a human right*, ohchr.org 2011, http://www.ohchr.org/EN/Issues/CulturalRights/Pages/Consultation10Feb2011.aspx, accessed on 6 December 2018.

Šola, T. 1998. The innovation of ecomuseums, Academia.edu 2016, https://www.academia.edu/20040838/The_Innovation_of_Ecomuseums, accessed on 6 December 2018.

UNESCO 2003. The Charter on The Preservation of Digital Heritage in *Records of the General Conference* 32nd Session*: Resolution*, Volume 1, unesco.org 2003, http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/charter_preservation_digital_heritage_en.pdf, accessed on 6 December 2018.

UNESCO 2017. World Heritage, unric.org 2017, https://www.unric.org/en/latest-un-buzz/30237-our-legacy-from-the-past accessed on 6 December 2018.

WIPO 2011, What is Intellectual property? Wipo.int 2011, http://www.wipo.int/edocs/pubdocs/en/intproperty/450/wipo_pub_450.pdf, accessed on 6 December 2018.

List of figures