Forgotten as Data – Remembered through Information

Social Memory Institutions in the Digital Age: The Case of the Europeana Initiative

Attila Marton

Information Systems and Innovation Group
Department of Management
London School of Economics and Political Science
Houghton Street, London WC2A 2AE
UK

October 2011

Dissertation submitted in the fulfilment of the requirement
for the award of the degree of Doctor of Philosophy

DECLARATION

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

ABSTRACT

The study of social memory has emerged as a rich field of research closely linked to cultural artefacts, communication media and institutions as carriers of a past that transcends the horizon of the individual's lifetime. Within this domain of research, the dissertation focuses on *memory institutions* (libraries, archives, museums) and the shifts they are undergoing as the outcome of digitization and the diffusion of online media. Very little is currently known about the impact that digitality and computation may have on social memory institutions, specifically, and social memory, more generally – an area of study that would benefit from but, so far, has been mostly overlooked by information systems research.

The dissertation finds its point of departure in the conceptualization of information as an event that occurs through the interaction between an observer and the observed – an event that cannot be stored as information but merely as data. In this context, memory is conceived as an operation that filters, thus forgets, the singular details of an information event by making it comparable to other events according to abstract classification criteria. Against this backdrop, memory institutions are institutions of forgetting as they select, order and preserve a canon of cultural heritage artefacts.

Supported by evidence from a case study on the *Europeana* initiative (a digitization project of European libraries, archives and museums), the dissertation reveals a fundamental shift in the field of memory institutions. The case study demonstrates the disintegration of 1) the cultural heritage artefact, 2) its standard modes of description and 3) the catalogue as such into a steadily accruing assemblage of data and metadata. Dismembered into bits and bytes, cultural heritage needs to be re-membered through the emulation of recognizable cultural heritage artefacts and momentary renditions of order. In other words, memory institutions forget as binary-based data and remember through computational information.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my parents who have supported me throughout my entire studies and always found the right words of encouragement at the right time. For that I cannot thank them enough. “Köszönöm, hogy tanulhattam.” I would like to extend my gratitude to my two brothers and their families who have been a great source for inspiration throughout my entire PhD candidacy.

I would also like to express my deepest thanks to my supervisor, Prof. Jannis Kallinikos, for his invaluable advice and guidance through the world of academia. He took me on an intellectual journey that formed me into a better researcher than I ever hoped to become. I am looking forward to continuing this journey together. I would like to thank my second supervisor, Dr. Nathalie Mitev, for her support in matters of research and beyond.

I am indebted to the TIGAIRians who welcomed me into the team even before I started my PhD. I want to thank Prof. Giovan Francesco Lanzara for his great feedback on every word I have written during my candidacy. My thanks extend to my fellow TIGAIR PhD candidates, Aleksi Aaltonen and Jose-Carlos Mariategui, who shared the joys and struggles of the last five years.

My special thanks go to the team at the Europeana Office and to Jill Cousins for their time and, above all, for welcoming me into the domain of digital cultural heritage.

Last but not least, I would like to express my gratitude to everybody in the Information Systems and Innovation Group at the LSE for all the stimulating and enriching experiences I am so grateful for - especially Aaron Martin and Wifak Houij-Gueddana who were and will remain not only great colleagues but great friends.

LIST OF ABBREVIATIONS

ACE	Association Cinémathèques Européennes
ALEX	Historical Austrian Texts of Law Online
ANL	Austrian National Library
ANNO	Austrian Newspapers Online
APEnet	Archival Portal Europe
API	Application Programming Interface
ARROW	Accessible Registries of Rights Information and Orphan Works
BBC	British Broadcast Company
CASPAR	Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval
CENL	Conference of European National Librarians
CLIR	Council on Library and Information Resources
DC	Dublin Core
DCC	Dewey Decimal Classification
DMF	Direction des Musées de France
DPE	Digital Preservation Europe
DRO	Digital Representation Object
DSO	Digital Surrogate Object
EC	European Commission
EDL	European Digital Library
EDLnet	Thematic Network for the creation of the European Digital Library (later renamed into EuropeanaNet)
EDM	Europeana Data Model
eLib	Electronic Libraries Programme
EMA	European Museum Academy
ENPA	European Newspaper Publishers' Association
ESE	Europeana Semantic Elements
ESF	European Science Foundation
ESRC	Economic and Social Research Council
EUBAM	Arbeitsgruppe zu Europäischen Angelegenheiten für Bibliotheken, Archive, Museen und Denkmalpflege
EURB-ICA	European Regional Branch of the International Council on Archives
FOAF	Friends of a Friend
HLEG	High Level Expert Group (on Digital Libraries)
HTML	Hypertext Markup Language
ICOM-Europe	International Council of Museums Europe
IFLA	International Federation of Library Associations and Institutions

IFRRO	International Federation of Reproduction Rights Organisations
IFTA(FIAF)	International Federation of Television Archives (Fédération Internationale des Archives du Film)
ICT	Information and Communication Technology
IS	Information Systems
IT	Information Technology
IIPC	International Internet Preservation Consortium
JISC	Joint Information Systems Committee
JPG	Joint Photographic Experts Group Standard
KB	Koninklijke Bibliotheek (National Library of the Netherlands)
LIBER	Ligue des Bibliothèques Européennes de Recherche
LIS	Library and Information Science
LOD	Linked Open Data
LSE	London School of Economics and Political Science
MARC	Machine-Readable Cataloging
METS	Metadata Encoding and Transmission Standard
MICHAEL	Multilingual Inventory of Cultural Heritage in Europe
MLA	Museum, Libraries and Archives Council UK
NASA	National Aeronautics and Space Administration
OAI	Open Archives Initiative
OCLC	Online Computer Library Center
OCR	Optical Character Recognition
OPAC	Open Public Access Catalogue
PDF	Portable Document Format
PLANETS	Preservation and Long-Term Access Through Networked Services
PPP	Private-Public-Partnership
RDF	Resource Description Framework
REF	Research Ethics Framework
TEL	The European Library
TIGAIR	The Information Growth and Internet Research project
UGC	User Generated Content
URI	Unique Resource Identifier
URL	Unique Resource Locator
VIAF	Virtual International Authority File
W3C	World Wide Web Consortium
WWW	World Wide Web

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Based on a gigantic census commissioned by William the Conqueror in 1086, the so-called Domesday Book is a historical treasure as it compiles detailed information about the people of Medieval England and their property. For the 900th anniversary of its creation, the historical document received a special treatment by the BBC. In order to free the content of the book from the confines of the protective chest it is preserved in by the UK National Archives, the BBC invested £2.5 million to grant the Medieval document a second life in digital form (McKie and Thorpe 2002). The so-called Digital Domesday Book project digitized every page in high resolution and, enriched by multimedia functionalities, produced a set of discs in 1986.

However, during the years that followed, the computer system used for accessing the discs as well as the standards used for structuring the data stored on the discs had become obsolete. As a consequence, the digital version turned out to be inaccessible after only 15 years. As an expert was quoted in an article by The Guardian: "*It is ironic, but the 15-year-old version is unreadable, while the ancient one is still perfectly usable* […]. *We're lucky Shakespeare didn't write on an old PC*" (McKie and Thorpe 2002). It took the combined efforts and expertise in, what could be called, digital archaeology of the universities of Leeds and Michigan to make the data accessible again by means of developing specialized software emulating the hard- and software environment of the Digital Domesday Book (BBC 2002).

The example of the Digital Domesday Book is, unfortunately, not a unique case but rather symptomatic for the problems and challenges that arise when communication media and cultural heritage are cast into binary-based formats. Be it digitized or born-digital, an increasing portion of society's future past will be reconstructed based on digital objects (Kallinikos et al. 2010a). From this perspective, the example of the Digital Domesday Book highlights two issues. First of all, digital cultural artefacts rely on a layered architecture of hardware

devices, software applications and data standards based upon which data can be processed and emulated into a recognizable cultural artefact (Zittrain 2008; Yoo et al. 2010). Second, IT standards are prone to change and become obsolete in increasingly shorter periods of time risking the accessibility of culturally valuable digitized artefacts or born-digital data alike (Lyman 2002; Gladney 2006).

The testimonies of the past and the cultural heritage future generations will refer to will depend almost exclusively on computational operations and binary-based information systems. This is an unprecedented situation illustrated by the difference between the original Domesday Book, that can be used as it is, and its digital counterpart requiring an assemblage of appropriate hard- and software in order to be recognizable as the Digital Domesday Book. Thus conceived, the digitization of everyday as well as institutional life results in new developments in terms of how society will remember the past in the future – social memory. With the rise of ICT and, ultimately, the internet notions of cultural and social memory need to be reconsidered to address the changes practices of forgetting and remembering are going through. In detail, the reported research focuses on the changes memory institutions are going through as they step into the online world of web services, digital media and interconnectivity.

For centuries, memory institutions, such as libraries, archives and museums, have been playing a key role as guardians of knowledge and gatekeepers to information. Their institutionalized practices of selecting, classifying, preserving and mediating knowledge has had a profound impact on the structure of social memory and the ways society remembers and forgets. However, these practices are increasingly challenged by new ways of organizing knowledge. The revelatory single case study of this dissertation – Europeana (www.europeana.eu) – is a case in point. Being one of the flagship projects of the European Union Digital Agenda, Europeana brings together the digitization efforts of European memory institutions under a single framework thus granting direct access to millions of digitized cultural heritage artefacts. However, the findability of those items is partly provided by search engine and semantic web technologies – not by catalogues. In abstract terms, the structure of social memory is shifting from a

fixed set of ex-ante categories to ex-post computational calculations (Esposito 2002; Weinberger 2007).

The implications of this shift towards computational techniques of knowledge management and, as a consequence, the change of social memory institutions are manifold and wide-ranging. As an increasing part of cultural heritage is being migrated or born into this new paradigm of computational operations and algorithmic calculations, the question arises how memory institutions are changing in order to remain memory institutions?

1.1 Motivation and Scope of the Research

The research on memory institutional practices in a digital environment has its roots in a humble idea. As the author joined the community of IS researchers, the majority of research seemed to be dominated by a focus on the latest trend in ICT and the "newness" of the changes these trends bring about. In contrast to gazing into the future, the author looked back into the past in order to study one of the oldest information systems still in existence today – the library. At first, the inquiry into the history of libraries was driven by the appreciation of the resilience of librarian practices as they have evolved throughout and adapted to quite radical changes in the fabric of society. The invention of the printing press comes to mind leading to unprecedented forms of social interaction and organization. Admittedly, libraries changed as a consequence of the mass production of print media, but they still remained libraries.

One can easily draw parallels to the diffusion of the printing press with respect to the scope of changes and the force of impact the rise of ICT and binary-based digital media will most likely have on contemporary society. It is in this spirit that a first research interest developed based on a question of how libraries will change this time while still remaining libraries. In more general terms, the dissertation was motivated by a circular conceptualization of what is old and what is new or rather how "the old" finds itself in "the new" and "the new" in "the old". As the dissertation concludes, this circle will be interpreted as "the

old" being re-memembered into "the new". Be it old media or traditional ways of ordering artefacts, they are not abandoned and forgotten but rather re-membered into new media as well as new ways of ordering artefacts.

The relationship of the old and the new is reflected by the relationship between the past and the future structured by mnemonic operations. Conceived as a social phenomenon sui generis, social memory can be seen as the social construction of a past in order to fit an expected future. A construction performed in the present that is never the same, social memory provides a present-past and, therefore, a present-future neither of which are fixed or static. The past is constantly re-constructed and finds itself in the future or rather in the expectations of that which has not yet come to pass. It is in this sense that the past is in the future and the future is in the past caught in a circular loop of one feeding into the other. The research project set out to study how libraries are changing in order to remain libraries but, ultimately, resulted in an in-depth inqury into the ways society re-calls the past and calls the future. In other words, the dissertation developed an interest in *how* the so-called information society remembers and forgets (rather than *what*). As a consequence, the research broadened its scope from libraries to memory institutions more generally in order to address the contemporary form and structure of social memory in-formed and re-membered into the medium of bits and bytes.

The second motivation that has driven the research is the introduction of memory as a prolific concept for the domain of information systems (IS) research. At the same time as information technologies are capable to produce surprises and novel insights or, in more general terms, to construct information, they are also memory technologies. Thus conceived, the dissertation proposes the study of social memory and its technologies in order to shed light on the intricate relationship between ICT and digital media, on the one hand, and remembering and forgetting, on the other hand. Social memory is more than the mere storage and retrieval of data; it is an active operation of re-cognition – be it the re-cognition of the self or the other. Thus conceived, the dissertation can be seen as an attempt to create awareness about the difference between storage and remembering as well as deleting and forgetting. As the example of the Digital

Domesday Book illustrates, having the data stored does not result in the book being remembered. Indeed, the dissertation will argue that it is more appropriate to consider the storage of data as forgetting as data.

All these points taken together, the dissertation can be read as an argument against the viability of the storage metaphor of social memory. A metaphor that had a profound impact on mnemonic conceptualizations and practices in the past, social memory as a storage device has been revitalized with the rise of ICT and its immense capabilities to store binary-based data. As the increasing importance of search engine algorithms, folksonomies, datamining and other second-order technologies indicate, the focus shifts from the collection of data to its manipulation and processing in order to make the, by now, insurmountable heaps of data navigable and informative (Kallinikos 2006a; Gantz et al. 2008). However, this is not to be understood as contemporary society being on the verge of forgetting how to forget or even on the verge of destroying memory, as some propose (e.g. Assmann 2006; Bell and Gemmell 2009; Mayer-Schönberger 2009). Supported by empirical evidence and theoretical conceptualizations, the dissertation will argue that contemporary society is forgetting more by remembering more just as it is remembering more by forgetting more (Esposito 2002). It is forgetting as data and remembering through information.

1.2 Research Approach

The empirical study is based on the constructivist epistemology developed by Ernst von Glasersfeld (1997), Heinz von Foerster (2003), Humberto Maturana and Francisco Varela (1992), Niklas Luhmann (2002b) and others (e.g. Moser 2004a). Conceived as a self-referential system, scientific research is seen as the construction of viable explanations according to standards, best practices and criteria developed by science itself. By the same token, scientific reality presents itself as an accomplishment by science. In light of this argument, methods do not reveal certain aspects of a given reality but rather are the means to construct a scientific reality. Once methods become the topic of scientific discourse, science

observes itself and, thus, turns self-referential giving rise to a logos of methods on how to define science scientifically.

Against this backdrop, the research design as such can be seen as a way of constructing a phenomenon scientifically in order to provide for a viable description and explanation of the research field. In detail, the empirical study is a revelatory single case study (Yin 2003) on the Europeana initiative – an EU funded project on the digitization of European cultural heritage. Europeana serves as a case for the changes memory institutions are going through in order to remain memory institutions in the digital environment. Based on a qualitative methodology, the dissertation sets out to study the transformation of institutionalized social memory in depth in order to contribute new insights into the field of information systems research and social memories studies (Garde-Hansen et al. 2009b).

1.2.1 Empirical Study

Consistent with Yin's (2003:13) suggestion to use case studies when the phenomenon of interest cannot easily be delimited from its context, a case study research design was employed in order to investigate the digitization of cultural heritage, which is a highly interoperating and networked assemblage. Indeed, it is not only difficult to separate the phenomenon from its context but rather impossible as the assemblage itself is based on the interoperability of its elements, aspects, procedures and operations. By the same token, the technique of corpus construction was applied in order to develop a sampling structure gradually as the understanding of the field grew during the course of the research project (Bauer and Aarts 2000). In general, a corpus is basically a collection of data that may vary from simply collecting text to any kind of symbolic token. The methodological focus is on the variability in the population rather than representativeness derived from proportional demographic sampling.

In detail, the research design consisted of three stages;

1) The pilot study (Feb. – Sept. 2007) focused on the Austrian National Library (ANL) in order to explore the potential for studying the digitization of memory institutions as well as to develop a feasible set of expectations and focused research questions. The pilot was conducted during the author's MSc studies at the Information Systems and Innovation Group, LSE, and the results were published as a MSc dissertation available online (see Marton 2007). Hence, the pilot is not part of the PhD dissertation.

2) The main case study on the Europeana initiative (Sept. 2007 – June 2011) was designed as a revelatory case study since the pilot study confirmed Europeana's uniqueness with respect to its scale and organizational setup as an emerging online service provider. The dissertation focuses exclusively on the Europeana initiative and, therefore, on this phase of the overall research design. Because of limited resources, data collection focused on documents downloaded form Europeana's central document server.

3) The final stage (July 2011) served for communicative validation - a quality criterion recommended for qualitative research (Gaskell and Bauer 2000). The analysis and interpretation of the single case study was checked with employees of the recently founded Europeana Office. In detail, expert interviews were conducted in order to evaluate the viability of the results and explanations presented in the empirical study.

1.2.2 Research Objectives

The increasing diffusion of the internet and digital media raises questions concerning the remembering of the past in the future. As traditional memory institutions (libraries, archives and museums) step into the online world, their role as guardians of knowledge and gatekeepers to information is being

renegotiated. At the most fundamental level, the image of memory institutions as societal storage compartments of cultural heritage needs to be re-evaluated. In contrast to, for instance, books, digital artefacts are computational processes that are only brought into a momentary form that resembles a fixed object (Kallinikos et al. 2010a). By the same token, the ordering of artefacts according to a set of fixed, ex-ante categorization rules and procedures, the central tenet of traditional memory institutions, are abandoned as they do not fit into the domain of algorithmic search engines and performative ordering through sophisticated data-base operations (Weinberger 2007). Taking the digitization of memory institutions as an example, the main objective of the dissertation is to convey a convincing argument to abandon the storage metaphor of social memory in favour of a metaphor of computation and processing. A transversal social memory, as it will be refered to in the dissertation.

In detail, the dissertation accomplishes the following objectives;

1. It introduces an alternative conceptualization of social memory based on the theory of social systems by Niklas Luhmann (1998). The dominating frameworks seem to be ill at ease with the relationship between individual memories and collective, cultural or social memory. The conceptual primacy of the storage metaphor leads to a theoretical dilemma as social and cultural memory is deemed to consist of individuals as remembering agents to be reminded by memory traces collected in storage compartments. Thus conceived, the field of social or cultural memory is not clear about the social aspects of remembering and forgetting.

2. It critically approaches the widespread notion of ICT as the ultimate remedy against forgetting. Based on conceptual and empirical work, the dissertation illustrates that social memory is a complex and dynamic social operation that does not fit into the technological framework of ICT the way it is conceptualized in the storage metaphor.

3. It aims at introducing the lively debate within the communities of librarians, archivists and curators with respect to the future of memory institutions. Indeed, ICT and the internet are by no means the end of libraries, archives and museums. From this perspective, the research also showcases the ways in which memory institutions are changing while remaining memory institutions.

All these three issues taken together result in the formulation of the following research question;

> *How does Europeana distinguish itself from the information habitat as a memory institution?*

In order to address the first objective, the research question refers to the distinction between Europeana and the information habitat as the environment into which Europeana is embedded. Thus, the research question is based on the fundamental tenet of the theory of social systems – (self)observation as an operation of making a distinction (e.g. Luhmann 2006). The second objective is addressed by reference to the concept of information habitat – a concept based on a computational, interoperating and process oriented approach towards ICT (e.g. Kallinikos 2006a; 2009b). Finally, the continuation of memory institutions in the digital domain (e.g. Kuny and Cleveland 1996) - the third objective – is addressed by observing Europeana as a memory institution. All these notions will be discussed in sufficient detail in the chapters that follow.

1.2.3 Research Limitations

From a conceptual perspective, the dissertation focuses on the study of libraries, archives and museums as memory institutions while leaving aside other institutionalized practices such as commemorations, rituals and so forth. By the same token, the dissertation does not reflect on the related domain of online privacy and data protection engaged in the analysis and critique of the long-term archiving of online personal data and the so-called digital footprint. Finally, the

research does not contribute to the broader field of organizational memory and recent developments with respect to organizational forgetting. While the author is aware of the intricate relationships between these fields of studies and the study presented in this dissertation, they are not relevant for answering the research question and, therefore, were not included into the research project in order to gain analytical clarity and focus. However, future research may address these issues.

From a methodological perspective, the single case study design inevitably leads to limitations with respect to the generalizability of the findings and therefore claims to provide for analytical generalization as proposed by Yin (2003). Finally, the focus on documentary evidence results in a lack of the kind of richness provided by personal accounts and opinions usually generated in an interview situation. Again, the author is aware of these methodological limitations, which will be discussed in detail and addressed as the dissertation unfolds the methodological approach and research design of choice.

1.3 Dissertation Outline

The dissertation is divided into eight chapters. The present chapter introduces the reader to the background developments, assumptions and reasons that have led to the inquiry into memory institutions in a digital environment. It also provides a short outline of the empirical study and the objectives it aims to accomplish.

Chapter 2 critically reviews the concepts of memory relevant for the social sciences – collective, cultural and social memory. While the first two concepts struggle with the separation of psychological and social memory, the dissertation introduces the concept of social memory as elaborated in the theory of social systems, which is based on a strict distinction between mental systems based on thought processes and social systems based on communication. The dissertation develops four paradigms in order to describe the emergence of a social memory in relation to innovations in communication media and record keeping; 1) memory of things, 2) places for musing, 3) institutionalized memory and 4)

information habitat. Followed by a discussion of memory institutions and related issues with respect to the digitization of cultural heritage, the chapter concludes with the formulation of a set of research interests addressing the transformation of memory institutions as they shift from the paradigm of institutionalized memory to the information habitat.

Chapter 3 focuses on the theory of social systems, which forms the basis for the conceptual framework applied in the dissertation. Beginning with an exploration of the notion of observation as distinction and indication, the chapter sets out to discuss the intricate relationship between observation, information, communication and memory. In particular, the chapter compares Shannon's engineering model of signal transmission with Luhmann's concept of communication as the autopoietic operation of social systems. Based on the theoretical framework, the research interests are further elaborated into a central research question and three sub-questions. Finally, the dissertation discusses the epistemological implications of the theory, which also serves as a bridge to the following chapter.

Chapter 4 delves into the methodological approach of the research. As a continuation of the previous chapter, the justifications for the choices in terms of research strategy, design, data collection and data analysis are presented. The chapter also documents how the various methods and techniques were put into practice. As a conclusion, the issue of quality criteria in qualitative research is discussed followed by the presentation of the quality criteria applied for the empirical study, which should allow an informed reader to evaluate the quality of the research.

Chapter 5 introduces the Europeana initiative as the case studied for the dissertation. After a historical outline of the prelude leading to the launch of the creation of a European digital library, the chapter discusses the political context and organizational structure of the Europeana initiative.

The case study analysis is presented in Chapter 6. The Europeana initiative is studied as an information space consisting of three layers – the digitized cultural

heritage artefact, the metadata it is structured and represented by and the context the metadata is embedded into. Two phases of the project are analysed in order to compare two different approaches towards the social memorization of cultural heritage artefacts – the portal based EDL prototype embedded into the paradigm of institutionalized memory and the API based Europeana service embedded into the paradigm of the information habitat.

Based on the evidence presented in the previous chapter, Chapter 7 discusses social memory as a transversal phenomenon in the digital domain. Thus conceived, the three layers of the Europeana information space are revisited in order to elaborate the concepts of transversal memory and forgetting-as-data. Finally, the dissertation argues against the viability of a storage metaphor of social memory in favour of a metaphor of computation and construction.

Chapter 8 concludes the dissertation. After a summary, the chapter discusses the contributions, limitations and, finally, the prospects for future research.

In particular, the case study demonstrates the unbundling and disintegration of three core aspects of memory institutional practices; these are the cultural heritage artefact, its description by means of so-called surrogates and, ultimately, the catalogue as such, which are dissolved into a steadily accruing assemblage of data and metadata. Dismembered into bits and bytes, cultural heritage needs to be re-membered through the emulation of recognizable cultural heritage artefacts and momentary renditions of order. As a consequence, memory institutions are caught up in a continuous cycle of forgetting as data and remembering through computational processes in order to remain guardians of cultural heritage in the information habitat. Thus conceived, the dissertation contributes an in-depth analysis of a major change in an important and venerable institution. Libraries, archives and museums and their practices of selection, order and preservation are not lost or abandoned but rather re-membered into a complex assemblage of data, metadata, computational operations and networked contextualization.

2. RESEARCH DOMAIN

2.1 Introduction

The domain of memory studies is in a process of becoming an interdisciplinary research endeavour combining a diverse range of disciplines such as neurology, psychology, sociology, philosophy, communication and media studies, anthropology, cultural studies, history, literary criticism, museum and library sciences and so forth (Radstone 2000a; Erll and Nünning 2008; Olick 2008; Roediger and Wertsch 2008; Suton 2009; Gudehus et al. 2010). The introduction of memory as a field for social science research can be traced back to the seminal works of Maurice Halbwachs in the early 20th century (e.g. Halbwachs 1992). An obscure endeavour in those days, since the grand theorists and founding fathers of modern social science as well as those who followed in their paths associated memory with tradition and a past irreconcilable with the rationalization of human affairs and the dawn of modernity that gave rise to the disciplines those grand theorists helped to establish – an attitude that was very much alive until the 1950s (Olick and Robbins 1998:107).

It was not until the 1980s that memory took a prominent role in academic, political and public life (Assmann 2002:400). Some authors even lament an inflation of discourses on memory leading to a "memory craze" in contemporary society (Baxter 1999; Olick 1999; Klein 2000; Berliner 2005). By now, there is a myriad of concepts and case studies that present quite a fragmented field of study (Olick 2008; Roediger and Wertsch 2008). As a result, memory studies, in general, and cultural or social memory studies, in particular, can only be described in a most generic sense.

> "*Used with various degrees of sophistication, the notion of memory, more practiced than theorized, has been used to denote very different things, which nonetheless share a topical common denominator: the ways in which people construct a sense of the past*" (Confino 1997:1386).

The reasons for the rising interest in memory are widely discussed among academics. One reason is seen to result from "traumatic" experiences caused by events that have been part of the public discourse for decades and will remain as

such even after their eye-witnesses passed away. The Holocaust in Germany and the Vietnam War in the USA are often invoked examples (Maier 1988; Wagner-Pacifici and Schwartz 1991; Perra 2010). Another related aspect is seen in the struggle of nation states with their respective pasts after the recent decline of their dictatorial regimes especially in Eastern Europe and Latin America (Borneman 1997; Olick 1999).

In more abstract terms, the increasing interest in social memory is connected with the deconstruction of the grand narratives of national identity but also of historiography's mission to uncover a "true" memory (Megill 1998; Olick and Robbins 1998; Klein 2000). By the same token, Nora (1989:7) famously stated that "[w]*e speak so much of memory because there is so little of it left.*" Due to an immense acceleration of everyday life through industrial and social modernization memory turns into history in an ever faster pace breaking down the walls between the expert historian and the lay memorizer (Kansteiner 2002; Poole 2008). Nora, following Halbwachs' notion that history is "dead" memory, concludes that what we call memory today is not real memory. It is not lived as *milieux* de memoires constantly evolving and caught in the dynamics of remembering and forgetting; it is rather a prosthesis memory organized and institutionalized in *lieux* de mémoires of monuments, archives, museums and so forth (Nora 1989; Cook 1997; Crane 1997). Traces of the past are obsessively archived due to an increasing anxiety of losing the past into oblivion - an anxiety brought about by the rapid acceleration of everyday life (Kallinikos 2010).

The observations outlined so far are closely aligned with observations on information to be developed in Chapter 3.5. Indeed, information and memory can be seen as two sides of the same coin. Thus conceived, this chapter is one half of a self-referential argument that oscillates between memory and information. In detail, the dissertation will walk the reader through the concept of memory first before it moves on to the notion of information in the next chapter on theory. Chapter 3.8 will revisit the points made in this introductory section, hence, bringing the conceptualization of the relationship between memory and information to a full circle. At this point, the dissertation will not delve into the question as to why memory is such a craze in contemporary society. Equipped

with additional theoretical concepts, this question will be revisited at the end of the theoretical chapter based on a better appreciation of the intricacies and subtleties of memory and information.

In what follows, the chapter will continue with the review of memory studies and social concepts of memory. An outline of the relationship between techniques and technologies of communication and record keeping will be used to illustrate the various ways social memory was structured into social memory paradigms. As Radstone (2000b:3) observers "*"memory" means different things at different times.*" The chapter then continues with an excursion into the research field – the world of libraries, archives and museums and their stepping into the digital domain of bits and bytes. After the identification of a gap in the literature and the formulation of the research interest, the chapter will conclude by aligning the social memory paradigms with the various concepts of memory, communication media, ways of ordering cultural artefacts and memory practices.

2.2 Social Memory Studies

2.2.1 Collective Memory and Social Frameworks

The study of social or cultural memory has its roots in the works of Maurice Halbwachs. A student of Emile Durkheim, he was one of the first to analyse and conceptualise memory as a social phenomenon in a systematic fashion. In his concept of collective memory, he discusses the formation of a group memory shared by its members. Remembering is a collective practice of reconstructing the past based on the present social framework (*cadres sociaux*) the group provides and its members find themselves in. In other words, the past is never quite the same since "*the past is not preserved but is reconstructed on the basis of the present* (Halbwachs 1992:40).

The social frameworks play a key role as they are the means through which a group accomplishes an image of the past in a way that is compatible with the dominant views of the very same group. For instance, language, the most

fundamental framework, does not only allow to talk about certain memories but enables to remember in the first place framed by what can and, more importantly, cannot be expressed with the words that are at an individual's disposal. The language one uses is not up for one to invent but is taken from the milieux one is born into (e.g. family), forced to join (e.g. military), chooses to become a member of (e.g. political party) and so forth (Halbwachs 1992). As a consequence, the connection between collective memory and identity becomes clear. Memory is repeated remembering perpetuating a sense of identity (Halbwachs 1992:47). The social framework is selective as a group tends to remember what fits or rather what the group makes to fit the framework; that is, to get rid of that which separates the members from one another. In today's terminology, the group decides what is remembered and what is forgotten (Wetzel 2011).

The individual's memory and the collective memory are interwoven and even though it is the individual that remembers, her memory is a social phenomenon. Hence, Halbwachs rejects the separation of the individual as the locus of memory from the collective (Connerton 1990:36). In the Durkheimian tradition, this is to be seen as a direct counterproposal to psychological and psychoanalytical concepts (Kansteiner 2002).

> "*There is no point in seeking where* [memories] *are preserved in my brain or in some nook of my mind to which I alone have access: for they are recalled to me externally, and the groups of which I am a part at any time give me the means to reconstruct them*" (Halbwachs 1992:38).

As it is the case with most collective concepts, it is difficult to establish whether a phenomenon is based on the aggregation of individuals or on the collective as such (Brockmeier 2002; Poole 2008; Kansteiner 2010). With respect to Halbwachs' concept, Olick (1999) observes a tension between collective memory being, on the one hand, an aggregation of socially framed individual memories or a collective phenomenon sui generis. He refers to the first as *collected* memory; it is individuals that remember while their social frameworks are merely reminders. The latter then is a *collective* memory with a degree of autonomy irreducible to psychological processes, which would also include the study of

stable symbolic systems preserved through technologies of memory such as writing (Olick 1999).

Halbwachs, however, does not incorporate material carriers of memory in a systematic way (Assmann 1995). For him, collective memory is lived in everyday life and experience and therefore remains relatively stable for a maximum of three generations, as it is the case when, for instance, grandparents, parents and children live in the same household. "Objectified" memory in the form of cultural heritage artefacts and historical records is "dead" memory or history – traces of the past we lost our "organic" connection with and buried in the graveyards of knowledge (Nora 1989; Olick and Robbins 1998). From this perspective, the tension between collected and collective memory becomes apparent in Halbwachs' work since he did not distinguish between memory and remembering performed by individuals.

2.2.2 Cultural Memory and Canon

Collective memory, as conceptualized by Halbwachs, is in fact a social-psychological phenomenon based on the lives and deaths of individuals. The discovery of memory as a field of study relevant to the humanities and, more importantly, culture, media and communication studies is usually ascribed to Jan and Aleida Assmann's proposal of a cultural memory – a memory that survives its human carriers based on external media (communication and memory technologies) as well as institutions such as libraries, archives and museums (Assmann 1999). Their work is based on the notion of social memory used by the art historian Aby Warburg. A contemporary of Halbwachs, he analyzed, although in an unsystematic fashion, artworks as symbolic expressions of ancient motifs, which have been passed on and, therefore, remembered over time (Confino 1997).

In order to minimize confusion with Halbwachs' concept, Jan and Aleida Assmann introduced a new terminology into the field. In their approach, they refer to Halbwachs' collective memory as *communicative* memory in order to

stress the non-institutional nature of everyday interaction and memory, which is not cultivated by specialists, not formalized and mostly found in spoken language and bodily gestures (Connerton 1990; Assmann 2008b). *Cultural* memory, on the other hand, is distanced from everyday life, maintained through cultural formations (e.g. texts, rites, monuments) and institutional communication (e.g. recitation, practice, observance) – a store of knowledge from which cohesive identities emerge (Assmann 1995). Indeed, it is hard to derive an identity for collectives such as the nation based on a concept of memory that does not go beyond the immediate interactions within groups (Basu 2011).

Cultural memory is directly related to this research since it introduces the notion of memory having an institutional existence being organized, among others, as libraries, archives and museums – a cultural heritage through which "*a society becomes visible to itself and to others*" (Assmann 1995:133). In general, memory is the basis for distinguishing myself, my family, my generation, my community, my nation, my culture, my religion from the others; "*Memory is knowledge with an identity-index*" (Assmann 2008b:115). Thus conceived, cultural memory is, quite literally, like visiting a museum. In this respect, the role of memory institutions is to maintain the potentiality of cultural memory to be actualized within a contemporary context (Assmann 1995; Kuhn 2010). They are the caretakers and preservers of cultural *heritage* artefacts; that is, they provide for the potentiality of their collections to be remembered. This is the reason why libraries, archives and museums can be referred to as memory institutions rather than educational institutions (Dempsey 1999). They may be connected to education and learning but they do not primarily teach.

Within this context, Aleida Assmann (2008a) further distinguishes between an active and passive cultural memory – the canon and the archive. While the canon preserves the past as present, the archive preserves the past as past. The term canon originates in the practices of the Catholic Church to form a selection of sacred texts, which are deemed authoritative and, hence, are not to be changed. As the term also found its place in literary criticism, arts and history, a canon became a collection of "classics" of a certain artistic tradition and a collective autobiography of, mostly, nations as it is told through textbooks, monuments and

so forth (Stuurman and Grever 2007; Assmann 2008a). By now, canon can be found as the subject of study in a wide variety of fields reaching from gender studies (e.g. Franke 2003) to even management studies (e.g. Gherardi and Nicolini 2002).

Aleida Assmann (2010) describes canonization as 1) a political *selection* of artefacts; 2) the ascription of *value* to the selected and, finally, 3) the preservation of the valuable selection in order to maintain *persistence* (see also Jones 1993; Cook 1997; Sandell 1998; Seixas 2007). Therefore, a canon can be defined as a selection of cultural artefacts, which are deemed valuable for a social formation – be it a group, collective or society; a faith, nation, artistic tradition or scientific discipline, to name a few. The key point is that a canon consists of artefacts, which are constantly kept relevant through education, storytelling, broadcasting, exhibiting and so forth. Thus conceived, a canon is the active aspect of cultural memory.

The archive, on the other hand, is the passive storage compartment of historical documents. In contrast to the canon, it consists of artefacts that have lost their cultural relevance and, therefore, do not contribute to the reconstruction of the past and present (Assmann 2008a). Archives store what has been forgotten, thus, forming a second-order memory. They store artefacts, which are not part of a canon but may be rediscovered at some moment in the future. In this sense, Assmann (2010) refers to a *historical* archive. By contrast, a *political* archive refers to bureaucratic control and administration. The term archive, after all, has its etymological roots in the Ancient Greek arkhō (αρχω) meaning to begin, to rule or to govern (Craig 1992). However, documents collected and processed by a governmental bureau may end up in a historical archive once they lose their utility. Indeed, anything can end up as a historical archive waiting to be discovered by a historian or archaeologist or even a geologist who treats geological sedimentations like an archive of earths history (Buckland 1997; Bowker 2005). In this sense, memory institutions are a combination of the two aspects of canon and archive. While a circulating library, for instance, offers access to books relevant within the context of a canon, a museum may exhibit only some artefacts keeping others secure in its archive for long-term

preservation. Of course, this setup is neither static nor fixed since archived artefacts may become part of a canon and canonized artefacts may be forgotten in an archive.

However, it is problematic to conceptualize a canon in such a fashion as it is based on an underlying assumption that cultural memory is representing a dominant view of the past. There is only one canon according to which a social identity is constructed. To accept the dominance of an official, so to speak, cultural memory as a given, does not provide for enough leeway to the notion that memory is contested (Olick and Robbins 1998). In fact, it does not consider that dominance is not an inevitable outcome of canonization (Stuurman and Grever 2007; Luyt 2011). Critical or minority viewpoints can develop an alternative canon of representative and authoritative texts, paintings and so forth as it is the case with the feminist countermovement in literature and science (Franke 2003). As a consequence, the concept also has its limitations with respect to the internet and online sociality, which denies the logic of a dominant canon (Assmann 2006; Weinberger 2007; Shirky 2008). This important argument will be revisited and explored in more detail after the internet is introduced as a new topic for the field of memory studies. For the moment, it is sufficient to simply conclude that cultural memory needs to be seen as based on a variety of canons, which may be complementary, contradictory or even ignorant of one another. Hence, cultural memory, in general, is not constrained to conformity but can be open to alternative expressions, parody or even subversion (Dijck 2007).

2.2.3 Social Memory and Forgetting

Given above made observations, the study of memory as a social phenomenon faces two problems, which can be summarized as follows. First, the popularity of memory in academic disciplines and public discourse has led to a field of study that, by now, addresses basically all walks of life. However, a primary focus on case studies and specific events has led to a research field without an organized memory of its own, as Olick and Robbins (1998) famously concluded. The field, as such, lacks gravitational centres in terms of conceptual work and theoretical

frameworks (Confino 1997:1387; Roediger and Wertsch 2008). The already mentioned "memory craze" and the wide range of terminologies used transform memory into an almost all encompassing term that is on the verge of losing any discriminatory power since anything can become an issue of memory (Klein 2000; Kansteiner 2002; Gudehus et al. 2010).

The second problem is the observation of the relationship between individuals and groups, collectives, societies as the "carriers" of social or collective memory – a tension already found in Halbwachs' work (Olick 1999; Brockmeier 2002; Kansteiner 2010). The fundamental contradiction revolves around the question whether conceptualizations of a collective memory are to be seen as a simile or as an analogy to individual or psychological memory processes. The first approach locates remembering within the individual who remembers in concert with other individuals. The collective remembers through its members *as if* it had a memory of its own. As a consequence, proponents of this approach are very sceptical towards the introduction of psychological and psychoanalytical methods into the realm of the social (Fentress and Wickham 1992; Irwin-Zarecka 1993; Assmann 1995). The latter, on the other hand, describes collective memory by applying psychological concepts such as trauma or amnesia in an analogous way to neurology, psychology and psychoanalysis (Olick 1999; Kansteiner 2002; Gook 2011). Hence, societies can have traumatic experiences, which can be suppressed into a collective sub-consciousness requiring therapeutic intervention (Klein 2000). In other words, the first approaches collective memory through the socialization of the individual, the latter through the psychologization of the collective.

Within this context, Olick and Robbins (1998) propose to use "social memory studies" as an umbrella term for research observing collective memory as a social phenomenon sui generis. It is an attempt to consolidate various research endeavours by distinguishing them as the study of social memory from the social study of memory – a distinction not always clearly drawn by collective memory studies (Olick 2010). The labelling is useful to some extent as it clarifies the object of study, however it does not really accomplish more. It basically follows the notion shared by many researchers that social, collective or cultural memory

is neither a metaphor nor an extension of individual memories and, hence, must not fall into the trap of what Jan Assmann (2002:401) calls "*collective mystifications*"; that is, to think of society as a collective body having a collective consciousness and, therefore, a collective memory. However, it does not address the problem of who does the remembering – the individual or the collective (Fentress and Wickham 1992; Assmann 2002). Used in this fashion, social memory simply replaces the term collective memory and extends the concept to incorporate material artefacts as reminders. Consequently, social memory studies and cultural memory studies can be used interchangeably (Olick and Robbins 1998).

In addressing this dilemma, sociology would be an obvious reference discipline as it has been struggling with questions of individualism and collectivism, agency and structure since its early days (e.g. Giddens 1984; for reference to memory and agency see Archer 1995:98). However, as Jan Assmann (2002) remarks, sociology has been remarkably reluctant in joining the discourse on social memory until today (Hagen 2011). An exception is the sociological branch of system theory developed by Niklas Luhmann (1998). Based on the second-order cybernetic research on memory by Heinz von Foerster (1967; 2003), Luhmann began to elaborate on the role of memory in his theory of social systems in the 1980s as well (Hagen 2011). Unable to finish his elaboration until his untimely death, it was mostly Elena Esposito (2002; 2008) who continued to analyse memory as an operation of social systems (Boyden 2003).

The advantage of a system theoretical approach is the strict separation between individuals and society or rather mental systems (e.g. consciousness), biological systems (e.g. body) and social systems (e.g. family) (Luhmann 1996b; Luhmann and Baecker 2002). As it will be discussed in more detail in Chapter 3.4.3, Luhmann proposes that social systems consist of communication rather than people. Social systems are therefore autonomous (not autarkic) while individuals or rather mental and biological systems are in their environment and vice versa. The communicative structures they embody are emergent phenomena, self-organized by and through communication with their own sets of attributes and

rules.[1] In short, social systems are defined as operationally closed systems based on communication (Luhmann 1996b; Luhmann and Baecker 2002).

As a consequence, a social system has a social memory that is communicative and, therefore, autonomous from mental operations of mental systems with their own memory capacities and structures. On a higher level of abstraction, memory is not a neurological, psychological or sociological concept but rather a systemic one; it is the faculty of every system (e.g. Ashby 1957; Bertalanffy 1973). Hence, mental systems and social systems have equivalent memory functions – equivalent according to abstract criteria defining the conceptual model "system" in general. In other words, neural, psychological and social capabilities of memory are instantiations of the same principles – they are structurally different but functionally equivalent (rather than simile or analogies).

To begin with, memory is defined as a complex system's function to maintain some kind of structural continuity in time, as its capability to preserve its internal form of self-organization – its in-formation - against external, environmental perturbations (Bussola 2011). In this sense, memory is a process or an operation rather than a collection of stored memories taking forms such as neural engrams, mental imagos, cultural artefacts or binary-based data (Riegler 2005). Conceptualizations of memory being more or less a passive recorder of information equate storage and retrieval with remembering, deletion or destruction with forgetting (Polastron 2007; Bell and Gemmell 2009).

In one of his essays, Heinz von Foerster (2003:101-31) refers to this notion as the "Mrs. X stores her coat during summer days" analogy. He criticizes that to see memory as a mere storage to be accessed by a retrieval mechanism is to conflate the vehicles for potential information with information itself – a point that will be

[1] At this point, the dissertation cannot go into more detail of Luhmann's theory of social systems and the conceptual rationale leading to this separation. The theory was constructed in a self-referential fashion, which means that, due to the linear structure of any narration, one is forced to accept certain concepts of the theory in order to be able to explain others. In this case, that would be the concept of social memory as a communicative operation of an operationally closed but structurally open autopoietic system. The reader is kindly asked for patience as this chapter will continue discussing memory as a faculty of systems in general and of social systems in particular. The discussion why social systems are made of communication and of communication only will be conducted in Chapter 3, which will feed back into this chapter.

revisited in Chapter 3.5. In other words, information is supposed to be stored and retrieved as it is, just like a winter coat can be stored during the summer and retrieved when the days are getting colder again. The more we store, the more we can retrieve and, consequently, the less we forget (Simon 1996:88). This analogy applies, however, to books, tapes or, more generally, to data - something that has the potential to inform (Assmann 1995).

Von Foerster, therefore, stresses the active and performative role of memory in terms of re-cognition rather than mere storage and retrieval. Coined the "Filet Mignon with French Fries and some salad, and an indefinable dessert" analogy, he states that recognizing the Filet Mignon as a Filet Mignon and the indefinable dessert as, at least, a dessert is already an accomplishment of memory. Memory constructs sameness into difference by filtering out the details of a singular event and, thus, making it comparable with other events. Recognizing a specific Filet Mignon as a Filet Mignon means to assign it to the abstract category of Filet Mignons. In other words, memory is the procedure of selecting what is remarkable and forgetting the rest. Forgetting, therefore, is the main function of memory, remembering the exception (Fentress and Wickham 1992; Esposito 2002; Assmann 2008a).

Filtering the singularity of events prevents the system from blocking itself (Foerster 1967; 2003).

> "[W]*ithout the ability to neglect most of the details and all the particulars that deviate from the remembered identity - that is, without the ability to forget - the faculty to remember would soon be overloaded*" (Esposito 2008:182).

Being a surprise or novelty, information, on the other hand, can only occur if an event is comparable with what is expected – a variety of what is remembered (Kallinikos 2006a:103). Hence, memory is not only oriented towards remembering but also, for the lack of a better term, anticipates what is yet to come; memory does not only re-call the past but also calls the future (Foerster 1967; Bussola 2011). "[M]*emory constantly modifies the past to connect it with a possible future in the present*" (Luhmann 1997:365). From this perspective, memory is a system's organization of observing information as it constructs sameness into difference and, as a result, recursivity and identity. In other words,

it balances redundancy and variety or repetition and novelty (Ashby 1958; Baecker; Esposito 2002:24).

Coming back to Luhmann (1998), these concepts can already be applied to language. Words or terms or rather semantics categorize the unique objects and singularities they refer to in a communicative, hence, social sense (Esposito 2002:21). For instance, the term "table" signifies all tables or rather the parameters that makes a table a table and not, for example, a chair irrespective of the individual features of each and every table there is. By the same token, classification systems strip that which is to be classified from its individual features according to the dimensions the classification system is built upon (Bowker and Star 1999). In this sense, Bowker's (2005:18) proposition that what is stored "*in archives is not facts, but disaggregated classifications that can at will be reassembled to take the form of facts about the world*" can be said about social memory in general – be it day-to-day communication or institutionalized cultural heritage (Assmann 1995). What is remembered are not actual tables but reconstructions of tables based on the parameters that define the category "table." Words correspond to concepts; they are generalizations filtering unfitting details as noise. This is, however, a communicative accomplishment, not a mental one (Esposito 2008). In the social domain, re-calling a table as a table is calling a table a table (Foerster 2003:104).

Esposito (2002) further elaborates the systemic interplay of social remembering and forgetting with respect to communication media and their ordering. She points at writing and, finally, printing as decisive innovations that resulted in the emergence of an autonomous social memory of a society too complex for a collective memory dispersed among individuals (Esposito 2008). With the developments of mass media and, ultimately, the internet, social memory has become ever more abstract in itself shifting to a higher level of order; a shift of focus from repetition to novelty and from remembering to forgetting. As it will be discussed in more detail further below, the paradigmatic example is the catalogue in the library. It does not remember books but rather classifications and instructions on how to find books in the library. From this perspective, memory turns out to be a self-referential phenomenon; the more elaborate the

classification system, the more details can be filtered out (Bowker 2005:17-21). However, the more elaborate the classification system, the more can be remembered as well. The more we forget, the more we remember. The more we remember, the more we forget.

To conclude, Esposito (2002; Boyden 2003) develops an evolutionary concept of social memory as a social operation of forgetting. The structure of social memory has changed over time with the diffusion of specific communication media and will continue to do so. Esposito's accomplishment is to show the changes of social memory as a process of forgetting. The following section will further explore Esposito's notion of the increasing abstraction of social memory and relate it to the history of memory institutions and the ordering of cultural artefacts.

2.3 Social Memory Paradigms

The emergence of social memory and organized practices of remembering and forgetting are coupled to developments in communication media spanning space and, more importantly in terms of memory, time (Cook 1997; Dijck 2007). In fact, it can be seen as a co-evolutionary relationship in terms of innovation and diffusion, meaning that the differentiation and diffusion of new communication media (e.g. books) is coupled to the differentiation and diffusion of social forgetting (e.g. catalogue). Memory, as a social phenomenon, is always mediated and, since modern times, mediatised (Esposito 2002; Dijck 2007; Hoskins 2009b; Perra 2010). Still, the following outline should not be read as the only way of describing the emergence of memory institutions nor as a unilateral cause-effect chain but rather as a feed-back loop between social memory and its communicative mediation by certain technologies (Orr 1977).

2.3.1 From the Memory of Things to Places for Musing

A communication medium, basically, separates the direct and mutual observation of communicative participants. Opposed to talking face to face, the utterance is de-contextualized and needs to be re-contextualized by the receiver in order to be (mis)understood. Hence, the written word gains a context-independent existence. However, that did not yet apply to the very first forms of pictographic writing like the hieroglyphs of ancient Egypt or early Sumerian clay tablets, since those do not mediate communication but rather representations of things (Fentress and Wickham 1992:17). Pictograms allow for a wide range of phonetic, semantic and grammatical interpretations that requires adequate knowledge in order to be read. Therefore, they are hardly able to mediate something the reader did not already know. Consequently, the very early pieces of writing were memory devices (aides-memoires) rather than communication technologies (Irwin-Zarecka 1993:162).

Aides-memoires present a paradigm of a memory of things; reminders rather than social memory as such. In the early days of pictographic record keeping, social memory was without any autonomy and as undifferentiated as the papyrus rolls and cuneiform tablets. In this sense, social memory can be seen as a social context for individual remembering the way Halbwachs (1992) conceptualized collective memory. Esposito (2002:44-97) refers to this paradigm as "*divinatorisches Gedächtnis*" translated as "*prophetic memory*" by Boyden (2003). The prime example is found in Greek mythology that defined memory as a divine phenomenon attributed to the goddess Mnemosyne (Lutz 1978; Polastron 2007:7). Embedded into the context of what is known and what is being talked about, prophetic memory does not accomplish a higher level of abstraction and, therefore, autonomy from the immediacy of face-to-face communication as it is the case with, for instance, storytelling.

A crucial change came with the innovation of the phonetic alphabet. The simplification of writing into a digital system of consonants in 2nd millennium B.C. Palestine and Syria, complemented with vowels roughly a millennium later in ancient Greece, allowed the text itself to become context free and autonomous

(Borgmann 1999:45-47). While reading ideograms still relies heavily on the contextual knowledge and memory of the reader, the reading of a text based on a phonetic alphabet relieves the reader from such burdens with the exception of how to pronounce the letters. In fact, a person can read such a text without understanding a single word. Although alphabetic writing was firstly used as a support for mnemonic techniques and oratory rhetoric,[2] it already meant the beginning separation of language from speaking (Fentress and Wickham 1992:18; Luhmann 1998:511) – a separation that can be traced back to the usage of signs as counting devices allowing to record quantities irrespective of what is to be counted (Borgmann 1999:40). One could convey content completely new to the reader allowing for the reader to be surprised – a very important aspect of information to be discussed in Chapter 3.3.

It is in those times that one finds the classical Great Library of Alexandria with an estimated amount of 743.000 scrolls collected through 3rd century B.C. (Jochum 1999; Polastron 2007). However, the collection was not an autonomous library but part of the temple of the muses – the Museion (Dahl 1958; Alexander and Alexander 2008:3). Although the scrolls themselves were already ordered by scholars according to subject criteria, the catalogue was merely an alphabetic inventory list of knowledge areas and authors (Strout 1956; Orr 1977:126). In other words, while the scrolls themselves already harboured the potential for reading to shift from remembering to learning, the catalogue still required a lot of contextual knowledge by the dedicated trustee who took care of just as many items as he could memorize (Assmann 2008b). One had to know what one was looking for.

While the rudimentary communication media changed from the usage of papyrus to parchment over the centuries, evolving into standardized documents and codices in the Roman Empire around 2nd century B.C. (Landheer 1957:18), the basic setup of the collections and their caretaking did not change until the final stages of the Medieval era (Strout 1956; Dahl 1958). Parchment was increasingly used for administrative records leading to the emergence of archives as

[2] Tropes were initially used by an orator to memorize his speech more easily. It is still used in writing for aesthetical reasons today.

caretakers of records and evidence (Cook 1997). The import of paper production techniques from China finalized the evolvement of the manuscript as a bound entity. However, throughout the Medieval Ages, libraries (in contrast to archives) were still part of religious institutions like monasteries or teaching mosques mostly dedicated to copying and preserving existing manuscripts. There was no need to change librarian routines of keeping an inventory, since a single library had to take care of a steady and rather small number of manuscripts. For instance, one of the largest monastery libraries – the Bobbio monastery library – had a collection of about 700 volumes in the 9th century (Dahl 1958:57). It was not until the rise of universities and the increasing number of quite large private collections throughout the Renaissance, that the dominance of monastery librarianship came slowly to an end.

Still, to call collections of texts throughout Ancient and Medieval times libraries in the modern sense of the word is a bit of a misnomer, since especially the distinguishing characteristic of libraries and museums as caretakers of mass media (e.g. books) and unique or rare objects (e.g. manuscripts) was not established yet (see Chapter 2.4.1). Textual artefacts were just as rare and, hence, valuable. In Ancient times, they were collected in places for musing – a museion in Greek or museum in Latin; in Medieval times it was the monastery or the mosque (Bedini 1965; Alexander and Alexander 2008:3). In fact, libraries remained to be part of museums until the 20th century, as it was the case with the British Library becoming independent from the British Museum as late as 1973.[3] Some authors propose that public museums are the continuation of spiritual and religious places for musing. They are the civic temples of today since they express "*in a modern idiom* […] *the essential values and world view of a community*" (Ames 1992:22) – a view of a civic world of documentation that co-evolved with other aspects of modernity such as industrialization, markets and so forth (Ekbia 2009b).

The prime example is Simonides' memory palace, an envisioned place of the mind a person can walk through to remember what has been memorized in the

[3] http://www.bl.uk/aboutus/quickinfo/facts/history/index.html, last access: 23 April 2011.

various locations of the palace (Yates 2001). Called "*rhetoric memory*", Esposito (2002:98-182) observes a first abstraction of social memory that allowed memory to be learned through mnemotechniques rather than to see it as a faculty of the divine. Memory was structured in places that required the knowledge of the topography from parts of the orator. In this sense, the museion was the worldly counterpart to the imagined memory palace. Constructed as walkable spaces, the key was the rationale according to which the items themselves were ordered. Weinberger (2007) refers to this setup as the first-order of order or the order of the things themselves, which still can be found in most public libraries and museum exhibitions today.

2.3.2 Institutionalized Memory

The invention of moveable type printing by Gutenberg in the middle of the 15th century can surely be called revolutionary. Although this technology was already well established in China, it was not very successful because of the extensive number of ideograms (4-5000 different types) required to print a book (Dahl 1958:84; Assmann 2002). In comparison, the highly granular system of relatively few Latin letters turned out to be an enabler for the mass production of media. In the first 50 years after Gutenberg's invention, 15-20 million books were produced (Weinberger 2007). Of course, this caused libraries to reorganize their items according to size rather than topic to save shelf space. Still, the printing press also brought qualitative changes. Besides the standardization of the book in terms of size and structure (e.g. usage of titles, paragraphs or page numbers), it was also necessary to standardize languages in order to minimize variety to reach as big an audience as possible – a process that led to the establishment of national languages from around the 16th century on and finally to the establishment of memory institutions as central instruments for the nurturing of national identities (Crane 1997). Most importantly, printing finalized the shift from what is known to what is new and interesting – from the oral tradition of repetition and devotion (still very much alive in the handling of manuscripts) to communication (Luhmann 1998:295-99).

The differentiation of books into a communication medium rather than a memory device co-evolved with the emergence of libraries as autonomous organizations dedicated to the collection of print media. By opening up the private collections of the aristocracy to the public from roughly the 17th century on, memory institutions were gaining an educational aspect in addition to the preservation of knowledge (Strout 1956; Landheer 1957:98; Alexander and Alexander 2008). Of course, at that time the public was limited to a relatively small elite (Dahl 1958:177). However, printing was not only dedicated to the production of books but also of pamphlets, leaflets and newspapers that addressed a growing literate portion of the population. The final innovation that popularized the reading of books was the novel in the 18th century. Accompanied by the invention of the paper-making machine in 1799 and the power press in 1810, reading became a comparatively cheap undertaking (Dahl 1958:220). In England for instance, the first circulating libraries rented out novels and further promoted literacy in the population to an estimated amount of 5 million in 1850 (Shapiro and Varian 1999:95). Hence, the beginning of an internal differentiation of libraries into public libraries as opposed to national and research libraries can be observed in those times. It is also the beginning of the shift from memory to information – from past tradition to future exploits (Connerton 1990).

The opening of memory institutions as educational facilities for the wider public marks the era of the modern library and museum accompanied by the professionalization of their caretakers from the mid 19th century on (Bedini 1965; Thompson 1982; Ames 1992:17; Crane 1997; Svenonius 2000:2; Alexander and Alexander 2008:9). Specifically books have proven themselves to be quite useful for the diffusion into a wider population as they are controlled through a publication process, mass produced and easy to use. Still, the notion that books preserve knowledge and hence libraries need to preserve books was and still is very much alive, which Ranganathan (1931:2) came to call a "*tendency to hoard books*". The user-friendliness of libraries, propagated by the British Museum Librarian Panizzi in the 1850s, towards open access to the items themselves collided with the need to store the items efficiently (Strout 1956; Dahl 1958; Weinberger 2007).

In general, public libraries allow direct access by way of organizing the items themselves. Books are being shelved according to the categories they are assigned to; hence, the way books are displayed is already informative. The dominant system, especially in Anglo-American libraries, is the infamous Dewey Decimal Classification System (DDC) developed by Melvil Dewey in the 1870s that classifies all the world's knowledge into 10 main categories and 100 sub divisions per category (OCLC 2003). Admittedly quite an ambitious task, it produces some biased or even embarrassing classifications because of its Anglo-American roots and the limited number of categories. For instance, the 100 sub divisions for the main category "Religion" are mainly dedicated to Christianity while there is only one division for Islam. By the same token, bias can be found in museums and the ways exhibitions were and still are curated (Jones 1993; Sandell 1998). As ambivalent as these open arrangements may be, they allow the user to browse through the collection and, therefore, to find something without knowing exactly what to look for. The library, rather than being an inventorized warehouse, becomes "walkable" for usage and discovery quite similar to walking through a museum exhibition *freely* - following a path laid out by the curator rather than being taken on a tour by an expert guide (Ames 1992:20; Alexander and Alexander 2008:10).

On the other hand, closed shelving applies to the immense growth in the number of books allowing indirect user access via various retrieval systems. The shift from book catalogues to card catalogues is the main result. Card catalogues, first formulated by the Smithsonian librarian Charles Coffin Jewett in 1852 based on the work of Panizzi (Strout 1956; Weinberger 2007:59), can organize an unlimited knowledge space since it is not bound by the limitations of book catalogues. However, browsing through a library's stock is not possible anymore, since the user needs to comply with the librarian way of describing items. The author catalogue is only usable when one knows exactly what one is looking for. From this perspective, a closed shelf library still remains a warehouse with an inventory that only answers questions like "Does the library have the book X by the author Y?" However, in order for a library to be used by patrons it needs to provide services enabling a user to find what she does not know (Ranganathan 1931; Thompson 1982:100). Subject catalogues are of limited help as they are

mostly based on a controlled vocabulary, meaning that items are indexed according to a standardized set of categories a user needs to familiarize with first. From this perspective, a library only provides a limited number of very specific paths for discovering items. In more general terms, librarianship refocuses from the problem of preserving knowledge to a problem of retrieval (Orr 1977:125; Hjorland 2000). As the concept of the catalogue diffused throughout the institutional landscape of modernity, it also found a place beyond libraries in museums, archives and, ultimately, in the domain of broadcasting providing new types of mass media to be archived by the respective broadcasting companies (Garde-Hansen et al. 2009a:3).

Be as it may, while industrial mass production finalized the standardization of formats such as books or newspapers, the card catalogue standardized how librarian items are to be described. Already to be found in Jewett's conceptualization, the standardization allowed for the merging of catalogues into union catalogues, first on a national and ultimately on an international level. As a consequence, loaning books between libraries became possible from the early 20th century on. Another consequence was the establishment of national libraries as reference libraries distributing printed catalogue cards in order to maintain the nationwide but also international quality of the catalogues (Dahl 1958:255). Especially among national libraries, cooperation became a standard practice particularly after the Second World War.

Esposito (2002:183-286) refers to this paradigm as "*culture as memory*" or "*modern memory.*" An explanation as to why culture organizes social memory would be too much of a distraction at this point and is not necessary for the furthering of the argument in this section. Instead, the dissertation will focus on, what Weinberger (2007) calls, the second-order of order. The prime example is the card catalogue constructed according to abstract criteria of categorization and order – according to a logic ("kata logos" in Ancient Greek) that is applicable to an infinite knowledge space mediated by communication media and cultural artefacts. In this sense, the card catalogue embodies a higher level of abstraction in social memory. The items themselves are forgotten. It is the rules of the catalogue that are being remembered. The order of the things themselves is

abandoned and replaced by the professional order of representations of things or, so-called, *surrogates*. Taking books as an example, their content is forgotten and filtered by the descriptive categories that fit onto a catalogue card.

2.3.3 The Information Habitat

Contemporary capabilities to store and to process immense amounts of binary-based data, combined with sophisticated tools for information retrieval and data mining, have led some observers to believe that society is forgetting how to forget. Enthusiasts proclaim a utopian future when information technology will relieve our modest mental memories from the strains of remembering and, ultimately, free human existence from the strains of its physical limitations – digital immortality (Bell and Gemmell 2009). Others warn that forgetting is an inherent aspect of the human condition. The internet and its services, on the other hand, may store and grant access to data for an indefinite period of time. As Mayer-Schönberger (2009) puts it, while humans forget, information technology remembers by default and may come back to haunt us with testimonies of our past mistakes.

By contrast, the communities of librarians, archivists and curators lament that today's society is on the verge of forgetting how to remember (Brindley 2009; Garde-Hansen et al. 2009a:6). Due to the ephemeral existence of, especially, online content, the rapid changes of technological standards and the absence of long-term preservation strategies, future generations may lack the necessary documentation to reconstruct the history of the late 20th and early 21st century (Russell et al. 1999; Bennett 2001). Indeed, the long-term perspective, specifically in terms of preservation, is a key criterion for distinguishing, say, a library from similar services such as Project Gutenberg (www.gutenberg.org) or Google Books (books.google.com) offering, to varying degrees, online access to digitized books. The WWW is not a library and Google is not its librarian (Green 2002). It is within this context that Aleida Assmann (2006:18) concludes that;

> "[i]*n the culture of the new media, memory is more likely to be destroyed than constructed. The internet, as we all know, is a medium that provides an unlimited plethora of information without actually storing it.*"

The situation proves to be more complicated as these two positions suggest (Radstone 2000b). Ironically, both positions are based on a storage metaphor of memory but reach opposing conclusions nonetheless (Locke 2000; Brockmeier 2002). While the first position relies on computer storage as a contemporary form of social long-term memory in contrast to the computers Random Access Memory (Riegler 2005; House 2009; Hagen 2011), the latter relies on the traditional archival storage as a declining form of social long-term memory in contrast to the canon as a working memory (Assmann 2008a; Olick 2010). Both agree that information and communication technology (ICT) has brought about a new technological paradigm (Dosi 1982; Kallinikos 2006a). They also agree that an increasing degree of cultural heritage is either being migrated or already born into digitality resulting in a new breed of cultural artefact that is dependent on computational operations (e.g. Manovich 2001; Dijck 2007; Kallinikos et al. 2010a).

However, the implications for social memory are still a terra incognita and draw more and more attention from the community of social memory studies summarized under the term *digital memory* (e.g. Garde-Hansen et al. 2009b; Hagen 2011). Reflecting the general state of memory studies, the existing conceptualizations seem to lack coherent conceptual frameworks to tackle digitality and online sociality.

> "*Thus, the existing paradigm* […] *and* [its] *associated traditions, theories and methods, is quickly becoming inadequate for understanding the profound impact of the supreme accessibility, transferability and circulation of digital content: on how individuals, groups and societies come to remember and forget*" (Garde-Hansen et al. 2009a:3).

In this spirit, it is of little help to re-cycle existing terms by simply combining them with new terms such as "digital" or "web" resulting in the study of online mementos, memorial web-pages, virtual candles, digital archives, online museums and ultimately digital memories without a deeper understanding of the technological paradigm shift from machinery to computation (Dosi 1982;

Manovich 2001; Kallinikos 2006a; Borgmann 2010; Kallinikos 2011). Hence, this section will focus on the most peculiar attributes of the new technological paradigm in order to contextualize the following conceptual chapters as well as the case study by means of relating it to new modes of ordering and, ultimately, remembering and forgetting.

Communication media, in particular, are increasingly in-formed into binary-based digits and formerly bounded, self-contained technological systems now share a common ground allowing for increased interconnectivity (Benkler 2006; Tilson et al. 2010; Yoo et al. 2010). Authors expect the impact of this leap to be as far-reaching as developments ascribed to the invention of the printing press and the mass production of print media (Luhmann 1998; Borgmann 1999; Shapiro and Varian 1999:95; Esposito 2002; Flusser 2002). Computers are powerful but, more importantly, purposeless machines affording immense potentialities for innovation (Esposito 2002; Zittrain 2008; Yoo et al. 2010). Hence, computational technology can be used for a vast variety of tasks (Eden 2007; Kallinikos et al. 2010a; Faulkner and Runde 2011). An analogous conclusion can be drawn with respect to the internet and its inherent agnosticism towards the services built on top. Conceptualized as a layered architecture of basically network hardware, services, applications and, finally, content, the internet allows for innovation on a variety of its layers without the necessity to know how the other layers work (Merali 2006; Zittrain 2008:68). From this perspective, the internet is equivalently without purpose as the PC and, therefore, facilitates surprises or functionalities the initial creators did not anticipate. Famously coined "generativity" by Jonathan Zittrain (2008), it affords innovations such as Wikipedia (Aaltonen and Lanzara 2010), open source software (Himanen 2001), mobile internet access (Sorensen 2010) or e-book readers (Yoo et al. 2010) as consequences of a complex and unpredictable information environment unintended by its designers (Benkler 2006; Marton 2009; Varian 2010).

The layers of this architecture present themselves as highly loosened and autonomous from one another (Blanchette 2011). In recent years, this fact has led to the fourth layer of content to develop features of generativity in terms of

content production rather than just mere content consumption (Shirky 2008; Yoo et al. 2010; Kallinikos et al. 2011; Kallinikos and Mariategui 2011). Referred to as Web 2.0 or user generated content (UGC), online users are able to embark upon activities so far reserved to professional experts as the often evoked example of Wikipedia, in contrast to the Encyclopaedia Britannica, illustrates (Benkler 2006; Shirky 2008; Pentzold 2009; Aaltonen and Lanzara 2010). A more relevant example with respect to this dissertation, is the social tagging of online content resulting in amateur folksonomies commonly produced and openly shared, which provide for a stark contrast to the professionally designed taxonomies of librarians, archivists and curators (Weinberger 2007; Gray et al. 2011). By the same token, social network sites pose a new hybrid of personal and public spheres bringing about hybrid memory traces, which are semi-private or quasi-public – a *social network memory* (Dijck 2007; Hoskins 2009b:30; Kreps 2010).

Brought forth by the pervasive and universal medium of bits, digital content can be separated from the devices, software applications and material media used for access, processing, manipulation and storage (Yoo 2010; Faulkner and Runde 2011). E-books are a telling and very relevant example as the content of the book (data) is separated from the way it is processed and displayed (metadata) on various devices with their own specifications – be it a PC, an iPad, smartphone or dedicated e-book readers such as Kindle. This is, of course, enabled by the interoperability and interconnectivity *between and within* all layers of the architecture mediated by bits and bytes. A generativity emerging as an inherent attribute of a layered and modular infrastructure (Hanseth and Lyytinen 2010; Tilson et al. 2010), it allows for recombination, reshuffling and repurposing across organizational and institutional boundaries (Sassen 2004; Kallinikos 2006a; Kallinikos et al. 2010a; Yoo et al. 2010). It is in this sense that memory traces can be shared across platforms and, with the rise of mobile technology, independent of the whereabouts of the sharers (Reading 2009). Information, thus, is ubiquitous (Yoo 2010) rendering its memorization ubiquitous in a way that is very different from the media ecology dominated by broadcasting, in general, and the television, in particular (Hoskins 2009a; Kallinikos and Mariategui 2011).

> "*In our post-broadcast age, the relationship between media and audiences are transposed and transformed, affording visibility on a new ubiquitous mediatised past, literally a "new memory"*" (Hoskins 2009b:29).

In addition, second-order technologies are running automated operations in the background handling the, by now, insurmountable heaps of data (Kallinikos et al. 2011). Most prominently, the increasing role of search engines in the online world seems to convey to them a position as gatekeepers to information equivalent to memory institutions (Dreyfus and Spinosa 1997; Introna and Nissenbaum 2000; Shaker 2006; Kallinikos et al. 2010a). However, the algorithmic search functionalities through which online content is made findable and, therefore, accessible is very different from the traditional practices of categorization and cataloguing (Locke 2000; Sluis 2010).

First of all, the search engine brings a momentary order. The search engine results page is a momentary rendition based on algorithmic calculations and data-base functionalities lost the moment it is closed by the user (Kallinikos et al. 2010a; Truran et al. 2011). By the same token, search engines do not categorize according to a pre-existing classification system but rather rank according to sophisticated algorithms. In other words, ex-ante order is exchanged by ex-post ordering. Esposito (2002:287-368) sees the search engine as a paradigmatic example for a new relationship between remembering and forgetting she terms "*telematic memory.*" The navigation through the information habitat is guided by the search results page, which is created, based on algorithmic calculations, for a specific user every time a search query is processed. In this sense, Dijck (2007:50) proposes that memory or rather remembering is becoming less of a process of recall but rather a topological skill of navigation. Similar to walking through the memory palace, with the exception that the layout of the palace is constructed based on a specific search query and disassembled after the user left the building. Hence, with search engines, social memory does forget fixed categorizations and the selective, persistent avenues paved for information discovery. Instead, it is enough if the algorithms are remembered.

The abstraction of the repository based on representational properties of a fixed order is exchanged by a higher degree of abstraction based on performative ordering only to be forgotten the moment the ordering is abandoned – the moment the results page is closed. This is not to say that the ordering is lost but rather it is stored as data. In a sense, each search query results in a momentary catalogue for a user only to be recalculated even when the same search query is repeated (Kallinikos et al. 2010a). This algorithmic memory (Sluis 2010) is a performative order*ing* rather than a static order. In other words, contemporary society is remembering even more by forgetting even more. Information is forgotten as data and (re)constructed (i.e. remembered) by means of computational operations (Kallinikos 2009b; Kallinikos et al. 2010b). Thus conceived, data mining tools or online search engine services are, in fact, technologies of remembering. They reconstruct, in a highly granular fashion, events stripped of their singularity by being stored as data based on the very dimensions according to which they were collected and stored in the first place.

To conclude, the new social memory paradigm is emerging out of an interconnected information environment of networked data and user involvement – an information landscape Kallinikos (2006a) refers to as the *information habitat*. A key observation for the following chapters is the shift from an ex-ante and static order to an ex-post and performative ordering. Be it social tagging or search engine indexing, the information habitat is based on a third-order of order with the filter on the way out rather than on the way in, as Weinberger (2007) proposes. It is the ordering of everything rather than an order of something. Within this context, the community of digital memory studies is engaged in a variety of digital memories and related practices trying to make sense of the "new memory" of the information habitat, as Hoskins (2009b:29) observes. A key conclusion to be drawn at this point, however, is that social memory structured according to the rationale of third-order of order is indeed forgetting more and, therefore, remembering more (Esposito 2002).

2.4 Memory Institutions

The concept of memory basically addresses the relationship between remembering and forgetting. It was Aleida and Jan Assmann's (2008b) contribution to observe memory institutions as crucial aspects of the structuring of this relationship on a societal level, hence, their practices as instrumental in the development of social identities. Elena Esposito (2002) reminds us, as those practices were gradually developed and fine-tuned, they also shifted from a focus on remembering to forgetting. Bringing these two approaches together, memory institutions can be conceptualized as institutions of forgetting.

As Dempsey (1999) observes, there is no definite umbrella term for libraries, archives and museums. Following his suggestion to use "memory institutions", the notion will be further explored in combination with social memory as primarily an operation of forgetting rather than remembering (Esposito 2002). To begin with, memory institutions are engaged into the formation of a canon consisting of cultural heritage artefacts. In general, canonization can be described as the selection, order and preservation of cultural artefacts. For this research, therefore, memory institutions are first and foremost libraries, archives and museums. Other social mnemonic notions such as public memorial days or monuments, which could be seen as being part of a broader memory institutional field, are left aside in order to gain a sharper analytical focus.

2.4.1 Libraries, Archives, Museums

As discussed in Chapter 2.2.2, Aleida Assmann (2010) defines canonization as 1) a political *selection* of artefacts; 2) the ascription of *value* to the selected and finally 3) the preservation of the valuable selection in order to maintain *persistence* (see also Jones 1993; Cook 1997; Sandell 1998; Seixas 2007). In this respect, the WWW is not a library nor are search engines librarians (Kuny and Cleveland 1996; Graham 2005). The domain of cultural heritage remains, so far, the sole domain of memory institutions. For instance, it is a library that elevates books from being mass produced consumer goods, embedded into a network of

authors, readers, publishers, printing industry, educational institutions and so forth, to being cultural *heritage* artefacts as part of a cultural memory (Assmann 2008). This is the difference between being a library and having a library. In light of this argument, it becomes clear that the online world lacks the institutional practices that would ascribe a canonical status to websites, blogs, wikis, emails, flash-videos, software applications and so forth. Even though search engines may provide for findability and order in a different but equivalent way to the practices of memory institutions, they are not constructing a canon. The same applies to the description of online content by means of commons based social tagging and folksonomies (Benkler 2006; Weinberger 2007).

An important starting point, Assmann's definition underestimates the importance of forgetting. Complemented by Weinberger's (2007) concept of three orders of ordering, the process can be further specified as selecting, ordering and preserving. Thus conceived, the aspect of value, as proposed by Assmann, is replaced by order. Since the observation that cultural heritage artefacts are deemed valuable is quite redundant, it is more appropriate to give order a more prominent role as a social process of forgetting. After all, the rationale according to which artefacts are being ordered and therefore made accessible reflects what has been one of the core themes of the librarian, archival and curator communities for over a century and, therefore, should be taken into consideration accordingly (Strout 1956; Craig 1992; Svenonius 2000; Alexander and Alexander 2008). In what follows, the three steps of selecting, ordering and preserving will be explored in greater detail in order to sharpen the conceptual gaze and understanding of what makes memory institutions into memory institutions and how libraries, archives and museums can be differentiated.

Selection

Selecting what becomes part of a canon and what is left to the ravages of time is an obvious way of forgetting. In general, one can say that libraries are selecting copies such as of books, newspapers or CD-ROMs – i.e. mass produced and packaged communication media. Codified in the Functional Requirements for Bibliographic Records (FRBR), the librarian domain introduces a hierarchical typology of the entities it collects (IFLA 1998:18).

Type	Definition	Example
Work	A distinct intellectual or artistic creation	Shakespeare's Macbeth
Expression	The intellectual or artistic realization of a work	Macbeth expressed in writing, as a stage performance, opera or a film
Manifestation	The physical embodiment of an expression of a work	Macbeth as found in different editions of books, recordings of different stage performances or films created by different directors
Item	A single exemplar of a manifestation	The actual book or DVD found on the shelves of libraries

Table 1: The typology of librarian entities (IFLA 1998).

Crucially, the work is defined as an abstract concept or category filled with fitting items. A library selects and, as a consequence, orders both – works and items (Svenonius 2000). As it is the case with every classification system, the FRBR typology constructs sameness into difference as well as difference into sameness and, therefore, performs as a mnemonic social operation. Assigned to the same category of the work, the items are further differentiated into sub-categories as they are catalogued accordingly.

> "*The placing of a given edition in its organizational context within the bibliographic universe is not unlike making a definition: first one states its genus (the work to which it belongs) and then, in a systematic way, its differentia*" (Svenonius 2000:11).

The distinction of work and item is an outcome of the mass production of media and is, therefore, a consequence of the original/copy distinction mass media is based upon. In this sense, a library takes care of a copy out of many. By contrast, the distinction between work and item is unnecessary in museums and archives since the selected items are rare, if not unique originals. Museums and archives select objects to become part of their collections according to considerations of uniqueness and value. As a curator outlined in a presentation collected for analysis (XE-004:1), those considerations can be outlined as summarized in Table 2.

Uniqueness	Value	Example
Unique in form	Valuable out of context	Valued arts objects, aesthetic minerals, curiosities
Unique by context	Not valuable out of context	Illustrations, symbols, John Lennon's t-shirt
Not unique	Not particularly valuable	Examples in natural history museums
Unique by rarity	Valuable as evidence	Most objects in paleontological museums and historical archives

Table 2: Categorization of museum and archival artefacts according to uniqueness and value (XE-004:1).

Hence, a contemporary print edition of Shakespeare's Macbeth ends up in a library. Since the first folio print of Shakespeare's Macbeth is rare and valuable as an artefact itself (not its content), it is taken care of in a museum or archive.

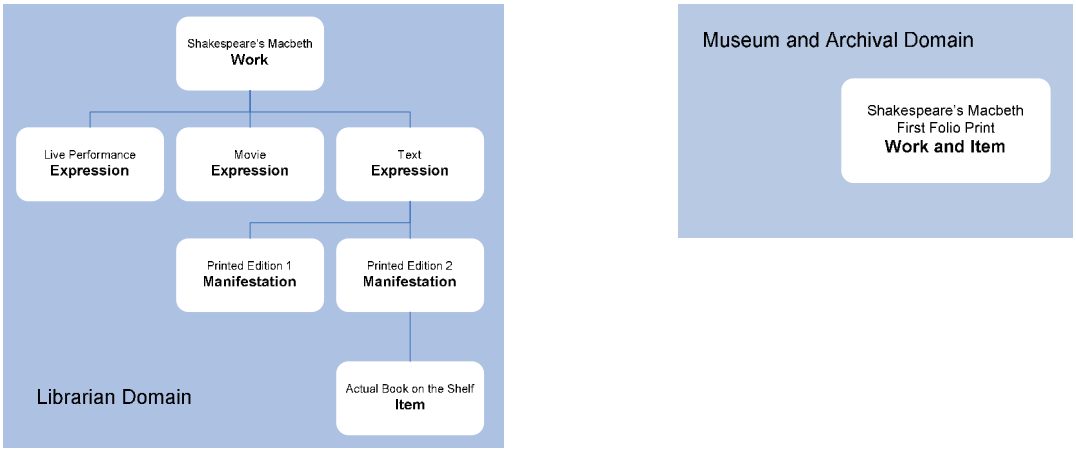


Figure 1: Relationship between work and item in the library, archive and museum.

An exception to the distinction depicted in Figure 1, the selection of examples for classes is close to the work/item distinction such as conserving a specific animal as an illustration for a type of animal exhibited in a natural history museum. Still, even items in a natural history museum are considered to be cultural artefacts as the following example illustrates;

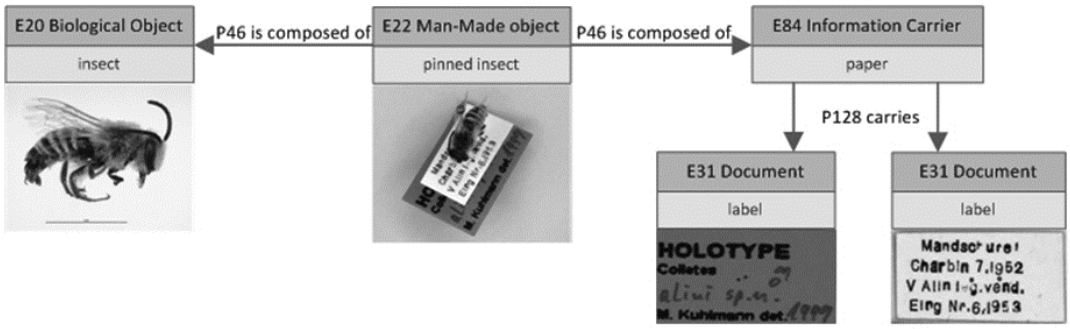


Figure 2: An insect as a cultural heritage artefact (XE-001:2).

Figure 2 is taken from a presentation given at a Europeana conference in 2008 and is part of the data corpus of the research (XE-001). It depicts the transformation of an insect into a cultural artefact (referred to as a "man-made object") composed of the insect itself and its documentation and labelling. Certainly, this example touches upon a grey area, since it may seem too much of a stretch to treat the documentation of nature not only as resulting in cultural artefacts but in cultural heritage artefacts (Buckland 1997). However, as an insect is selected, ordered according to a pre-defined typology and preserved for future generations, it may be more appropriate to see it as a document not of the history of nature but rather of the historiography of nature. In other words, the documentation of nature reflects the way a society observes nature rather than nature itself.

Order

Traditionally, the ordering of cultural heritage artefacts relies on the arrangement of the material artefacts complemented or replaced by the ordering of the descriptions of the artefacts (Strout 1956). In this sense, the term includes the open-shelf arrangement found in public libraries as well as exhibitions in museums. However, it is due to the innovation and diffusion of the second-order of order that ordering becomes an operation of social forgetting (Esposito 2002; Weinberger 2007). As already stated above, the card catalogue allows for the forgetting of an, in principle, unlimited amount of artefacts according to a set of professionally developed rules and procedures of ex-ante categorization and cataloguing. By the same token, remembering becomes a process of information retrieval following those very same rules and procedures.

In abstract terms, ordering refers to the construction of associations between artefacts that allows for avenues of discovery. In contrast to museums, the librarian domain as a whole has been quite successful in codifying the process of ordering due to the standardization of its mass produced collectibles. Representing the librarian domain, the International Federation of Library Associations and Institutions (IFLA 1998) defines the so-called bibliographic objectives as enabling a user to 1) find, 2) identify, 3) select and 4) acquire an entity.

Bibliographic objective	Explanation
find entities that correspond to the user's stated search criteria	To locate either a single entity or a set of entities in a file or database as the result of a search using an attribute or relationship of the entity
identify an entity	To confirm that the entity described in a record corresponds to the entity sought, or to distinguish between two or more entities with similar characteristics
select an entity that is appropriate to the user's needs	To choose an entity that meets the user's requirements with respect to content, physical format etc., or to reject an entity as being inappropriate to the user's needs
acquire or obtain access to the entity described	To acquire an entity through purchase, loan etc., or to access an entity electronically through an online connection to a remote computer

Table 3: Bibliographic objectives according to Svenonius (2000:17).

Traditionally, the bibliographic objectives were achieved through associations by creator and subject classification "hard-wired" into the card catalogue. As outlined in Table 3, the introduction of data-based catalogues did not fundamentally change this approach. In their revised version, the bibliographic objectives refer to entities and to the means of their discovery in a most general sense that can be extended to data-based operations.

Data-bases provide for more options, such as title searches or Boolean-Operators, but basically remain within the librarian rationale since the data-base contains the same descriptions or, as it is called today, descriptive metadata as the cards in the card catalogue. Indeed, some advanced data-base functionalities were first formulated by librarians before the diffusion of computers (Weinberger 2007).

By the same token, librarian principles of subject classification were applied by early search engines as they manually categorized web sites to offer online services similar to catalogues (Introna and Nissenbaum 2000; Locke 2000).

Figure 3: Second-order of order discovery system.

In general, the second-order of order is based on the creation of a surrogate that is separated from and, thus, stands for the actual artefact it represents (Howarth 2006). A catalogue is simply the arrangement of those surrogates according to a specific rationale, as the etymological origin "kata logos" testifies, meaning "according to a certain rationale or logic" in Ancient Greek (Strout 1956). Hence, a brochure of products arranged in a certain way in order to sell to customers is a catalogue just as much as a collection of catalogue cards. Still, the difference is that memory institutions apply the catalogue as a means for canonization rather than for marketing purposes.

Preservation

The preservation of cultural heritage artefacts is a key aspect of memory institutions (Assmann 2008a). A broad range of societal practices rely on enduring and persistent artefacts authenticated, canonized and collected by dedicated authorities for reasons of documentation, reference and identity (Allison et al. 2005; Assmann 2008a). Archives have played a crucial role in this context and continue to do so with respect to online content, in general, and web-pages, in particular (Lyman 2002; Cox 2007). Committed to providing persistent access to reliable testimonies of social facts, archives have been entrusted with the key tasks of collecting cultural items, documenting their provenance and preserving their authenticity. The ways these tasks have been carried out reflect a longstanding process of institutionalization revolving around two key principles – authenticity and provenance (Hirtle 2000; Factor et al. 2009).

To be precise, one has to differentiate between conservation and preservation. While the first is about taking care of the artefact itself, the second is about taking care of the content of the artefact (Arms 2001:254). Obviously, a museum conserves the items themselves in contrast to a library that could preserve the content of a book by simply photocopying the pages or even by acquiring a new copy of the book. With respect to archives, it may be either way depending on the rationale for archiving as well as on the types of artefacts being archived. In this sense, an archive may be best described as documenting themes, topics, procedures and so forth entrusted with the preservation of the authenticity and provenance of the document (Buckland 1997). In the archival domain, authenticity does not refer to accuracy but rather to the guarantee that the document was not tempered with after it was archived. A document may be a forgery but still authentic as it provides evidence of actions of an agent (Craig 1992; Hirtle 2000). By the same token, the catalogue needs to be preserved as well in order to provide for persistent findability and accessibility. Hence, preservation, in general, combines the preservation of content, documentation and the discovery system.

In this sense, conservation or preservation does not necessarily mean the prevention of forgetting but the prevention of disintegration and loss (Connerton 2008). The card catalogue, for instance, forgets the items, which are, nevertheless, conserved or preserved. Thus conceived, canonization can be conceptualized as a process of forgetting since forgetting is not only to be equated with the destruction (Polastron 2007) or deletion (Mayer-Schönberger 2009) of physical artefacts or computer files. By the same token, remembering is not to be equated with the mere storage of artefacts but involves discovery and accessibility; as it will be discussed in Chapter 2.4.5, a notion that turns out to be quite a challenge with respect to digital artefacts.

To conclude, libraries, archives and museums approach their role as memory institutions in quite different ways. Certainly, libraries, archives and museums pose a complex field of study that can be approached and distinguished from a variety of perspectives. For this dissertation, however, and with respect to

selection, order and preservation discussed in this section, libraries shall be nominally defined as ordered and preserved selections of packaged, mass produced communication media (e.g. Landheer 1957; Orr 1977) while museums are ordered and conserved selections of unique or rare artefacts (Alexander and Alexander 2008). An archive shall be understood as a selection of conserved or preserved documentations, which serve as socially accepted evidence (Buckland 1997; Cook 1997; Hirtle 2000). In contrast to having a private collection of books, documents or curiosities, memory institutions are involved into the canonization of a cultural heritage. It is in this sense that memory institutions are institutions of social forgetting.

2.4.2 Digital Libraries

The exponential growth in the mass production of media, what would be referred to as information growth today (Kallinikos 2006a), canonized by memory institutions is, in fact, an immense problem. Various authors came to question whether libraries will be capable not only to store books but also to maintain a certain level of usability (Ranganathan 1931; Thompson 1982). Technological innovations, like microfilming, may have had temporarily solved the storage problem, but the biggest hope is put into computerization and digitization not only in terms of storage and preservation but also in terms of access and retrieval (Schwartz 2000).

First steps were the image scanning of catalogue cards later followed by OCR scanning made accessible via Online Public Access Catalogues (OPAC). Data-based catalogues are now able to use a wide variety of IT enabled search and retrieval functionalities such as Boolean operators or ranking and relevance feedback (Hahn 1998). However, the ultimate step is the digitization of the items themselves, which enables users not only to retrieve descriptions but rather to access items via the internet. Projects such as Project Gutenberg (www.gutenberg.com), the Internet Archive (www.archive.org), the World Digital Library (www.wdl.org), Europeana (www.europeana.eu) but also private-public partnerships such as Google's book project (books.google.com), the Open

Content Alliance (www.opencontentalliance.org) and many more bear witness to the immense cultural but also business interests revolving around the access to cultural heritage artefacts (Bearman and Trant 2005; Graham 2005; Landon 2009; Tanner et al. 2009). Especially the involvement of Google is cause for a lot of controversy as to how access will be granted to digitized and born-digital items, mostly books and journals, in the future (Bearman 2006; Bjorner 2006; Waller 2009).

The term digital library is used in a very broad sense as an organized collection of digital items – be it digitized or born digital (Schwartz 2000; Arms 2001). Hence, the term also covers digital repositories of archives and museums (e.g. Knell 2003). Libraries, archives and museums, therefore, *have* digital libraries alongside their traditional collections. Following Rusbridge's (1998) suggestion, these compositions of traditional memory institutions, data-base driven cataloguing and digital items can be referred to as *hybrid* libraries. Having-a-digital-library is still embedded into the traditional notion of libraries but also archives and museums being "*a specific place with a finite collection of tangible information and it is geographically constrained*" (Oppenheim and Smithson 1999:99). In this sense, digitization simply complements the existing collection and practices.

By contrast, being-a-digital-library addresses fundamental changes in the ways cultural artefacts are being organized and made findable. This distinction will be shown in the case study as the abandonment of an ex-ante logic of categorization and the disintegration of the catalogue. In fact, formations of this kind can hardly be called digital libraries in the first place since the distinguishing characteristics that have separated libraries from archives and museums are not valid anymore (Howarth 2006:47). It may be more appropriate to speak of digital memory institutions rather than of digital libraries, digital archives and digital museums respectively (Marty 2008; Waibel and Erway 2009).

This observation is the basis for the research question to be developed in Chapter 3.7. Without going into too much detail, the basis can be formulated as the question of how memory institutions are changing in order to remain memory

institutions. Similar notions can be found in the library and information science (LIS) community as it is the case with Kuny and Cleveland (1996:1) stating that "[t]*echnological progress has changed how libraries do their work, not why.*" The LIS field obviously observes and discusses a process of change of librarianship ascribed to the rise of ICT and the internet (Davis and Lagoze 2000); this includes changing user behaviour (Peterson Bishop et al. 2000), the increasing importance of documentation and preservation of online communication, especially in science and research (Ercegovac 1997), and comparisons with new information service providers of which Google is the most prominent one (Schwartz 2000; Bearman 2006; Bjorner 2006; Li 2006). An indicator for the ongoing discussion is the lack for a standard definition of digital libraries (Meyyappan et al. 2000). Schwartz (2000), for instance, discovered 64 different definitions of digital libraries.

Compared with a traditional definition of a library as presented by Oppenheim and Smithson (1999) above, the conclusion to be drawn could be that either the library does not change significantly or it changes tremendously. Either a library simply includes digital works into their collections and offers corresponding services or it leaves "the place" behind (Oppenheim and Smithson 1999:99) offering any document any time to anyone in any place (Covi and Kling 1996:672) – a library without walls (Lee 2000b; Schwartz 2000). Which one to pick depends very much on the focus in terms of *what* makes a library a library or rather a memory institution a memory institution (Baker 2006).

The question also addresses the very important and still evolving relationship between institutionalized memory and ICT. Digital libraries are often positioned somewhere between these two domains (Oppenheim and Smithson 1999; Hjorland 2000; Schwartz 2000; Barreau 2001; Graham 2005). The discussion unfolds along three major dimensions; digitization and preservation of cultural heritage, accessibility for users and interoperability of cataloguing. The first point refers to the difficulties of not only digitizing analogue media but also to keep the digital content retrievable at any time (Petschar 2002). Memory institutions have to think in long terms. Decisions made during digitization (e.g. file structure) have a huge influence on the preservation policy and vice versa.

That is the case, because digital documents are unable to "*care for themselves"* (Russell et al. 1999:277), for digital media require the appropriate hard- and software in order to be accessible (Kallinikos et al. 2010a). All solutions, as it will be discussed further below in Chapter 2.4.5, do have considerable advantages and disadvantages (Russell et al. 1999; Bennett 2001).

Accessibility mostly refers to the interaction between the memory institution and the user (Eschenfelder and Agnew 2010). ICT enables not only multi-dimensional search functionalities through the repository but also new meeting points for face-to-screen interaction and collaboration. This is another novel issue for the field of librarianship that requires, like all the three points discussed here, quite some technical know-how. The discussion covers a wide range from usability in terms of screen- and interface design (Thong et al. 2002) and the enhancement of the readability of texts on screen (Greene et al. 2000), the design of systems to improve the relevance of results of user enquiries (Marcum 2001; Tuominen et al. 2003) to new tools for especially scientific users in terms of collaborative research and user behaviour (Peterson Bishop 1999; Wilensky 2000; Bollen et al. 2005; Holley 2010) – sometimes referred to as library 2.0 (Law 2008; Gerolimos and Konsta 2011). The final point mainly discusses concerns regarding the interoperability between formerly quite independent standards of classification and metadata (Petschar 1997; Suleman and Fox 2001; Woldering 2004). Metadata standards play a key role in the future development (Baker 2006) since they form the basis for search, retrieval, delivery, rights management and preservation (Russell et al. 1999; Smiraglia 2006; Massart et al. 2010). Hence, all three issues are highly interrelated.

As a summary, institutionalized memory is in a stage of renegotiating its role as guardian of knowledge and provider of information (Young 1996; Primary Research Group 2008). So far most discussions concern existing libraries and, to some degree, archives adapting to technological challenges but also as a reaction to Google's own project of mass-digitizing books (Coyle 2006; Schnapp 2008). It is only in recent years that museums are also discovering the internet for displaying their digitized collections (see for instance Google's collaboration

with leading art museums at www.googleartproject.com) (Elings and Waibel 2007; Waibel et al. 2010).

2.4.3 Metadata

As "data about data", the term metadata originated in the field of computer science and referred to electronic records only, thus distinguished from cataloguing as practiced in memory institutions (Heery 1996). Today, however, the communities of librarians, archivists and curators use these terms interchangeably as surrogates and, to an increasing degree, cultural artefacts take a digital format. Indeed, there is a plethora of metadata standards, schema and types most of them developed for specific projects and collections serving a variety of functionalities (Dekkers 2001; Greenberg 2006; Massart et al. 2010).

The two most important types for memory institutions and, therefore, for this dissertation are schema used for resource discovery and use; i.e. descriptive and structural metadata respectively (Arms 2001; Campbell 2006). Descriptive metadata is the successor of the catalogue card. Hence, it enables an entity to be found either by being stored in a separate data-base, as it is the case with an OPAC, or by being a part of the entity it describes, as it is the case with HTML web-pages (see Figure 4). Structural metadata, on the other hand, is a novelty for memory institutions as it describes how a digital object is to be processed and, ultimately, displayed. It ranges from a simple definition of a file being a .pdf document or .jpg image to complex digital objects assembled from a variety of data sources and services (Arms 2001:225-41). As the structural metadata of complex digital cultural heritage artefacts are part of the analysis in Chapter 6, this section will discuss descriptive metadata with a particular focus on standards relevant for the case study.

The implementation of descriptive metadata schema into memory institutions goes as far back as the 1960s. Under the leadership of the Library of Congress, the Machine-Readable Cataloging (MARC) standard was developed to allow the description and sharing of the bibliographic records of monographs via magnetic

tapes (Ristow and Carrington 1971; Avram 1975). Extended to all bibliographic items over time, it is the worldwide standard format within the domain of libraries by now. In the USA, MARC rejuvenated the sharing of catalogue records and the quite labour intensive and expensive cataloguing that finally culminated in the foundation of the Online Computer Library Center (www.oclc.org) in 1967.[4] Being in charge of the online union catalogue WorldCat (www.worldcat.org), it has been developed into a global pool of bibliographic records with libraries from all over the world adding their records to a collection of, by now, more than 100 million entries representing more than 1 billion bibliographic items.[5] The MARC standard also led to the development of remote access services enabling a library user to query a catalogue directly from a workstation – a feature used by reference management software tools such as Endnote – as well as the creation of online portals allowing a user to search through a catalogue via an OPAC web interface (Arms 2001:40-50).

The advantages of MARC are quite obvious as it lowers the boundaries for sharing bibliographic records and, as a consequence, lowers the costs for cataloguing. However, with the rise of the WWW and new formats of digital documents, it is reaching its limits as well. MARC was initially conceived for experts and their high standards for documentation and descriptive meta-data, which are, of course, not met by the lay population engaged in publishing web-pages and other born digital material. In order to address the latest developments in online media, the OCLC initiated a new metadata scheme called Dublin Core (DC) (www.dublincore.org) in 1995 – named after the hometown of the OCLC, Dublin, Ohio (Baker 2000; Arms 2001:192-94; Coleman 2006). The goal was to develop a scheme simple enough for lay people to be used for describing their own web resources. The result was a scheme of 15 elements (e.g. subject, creator, publisher etc.) as the lowest common denominator for describing any web resource.

The simplicity of Dublin Core, though designed to be its strength, was also its weakness, since it did not take into account the specific requirements of memory

[4] http://www.oclc.org/uk/en/about/default.htm, last access: 26 April 2011.
[5] http://www.oclc.org/uk/en/worldcat/about/default.htm, last access: 26 April 2011.

institutions. As a consequence, the original version – Simple Dublin Core – was opened to further extensions allowing the documentation of issues such as provenance or rights holders referred to as Qualified Dublin Core (Sugimoto et al. 2002). Figure 4 shows an example of Dublin Core used to describe an HTML document embedded into the document header. The Dublin Core elements are declared as “DC”.

```
<meta name=”DC.subject” content=”dublin core metadata element set”>
[…]
<meta name=”DC.publisher” content=”OCLC Online Computer Library Center, Inc.”>
<meta name=”DC.creator” content=”Weibel, Stuart L., weibel@oclc.org.”>
<meta name=”DC.creator” content=”Miller, Eric J., emiller@oclc.org.”>
<meta name=”DC.title” content=”Dublin Core Elements Set Reference Page”>
[…]
<meta name=”DC.form” scheme=”IMT” content=”text/html”>
<meta name=”DC.language” scheme=”ISO639” content=”en”>
<meta name=”DC.identifier” scheme=”URL”
content=”http://purl.oclc.org/metadata/dublin_core”>
```

Figure 4: Example of Simple Dublin Core embedded into an HTML document (Arms 2001:197).

Initiatives such as Dublin Core clearly show the increasing importance of online and born digital artefacts and the changes to be taken into account by the memory institutional field (Elings and Waibel 2007). For instance, the archiving of online scholarly publications has lead to the development of a wide-ranging interoperability framework - the Open Archives Initiative (OAI, www.openarchives.org). The OAI is a project promoting the development of metadata standards, complex object models and tools for metadata harvesting and exchange endorsed by libraries, archives and museums alike including Europeana, the case study of this dissertation (Sompel and Lagoze 2000).

Another example is being developed beyond the memory institutional field by the World Wide Web Consortium (W3C). The Resource Description Framework (RDF, www.w3.org/RDF/) is a standard for describing relationships between web resources based on a very simple syntax (Baker 2000; Byrne and Goddard 2010; Powell et al. 2010). To an increasing degree also used by memory

institutions, RDF allows for the qualified linking of data in contrast to the unqualified linking of documents as it has been the case with the WWW and hypertext (Shadbolt et al. 2006). Defined as web resources and identified by Unique Resource Identifiers (URI), any data can be addressed and associated with other data hosted by separate organizations. By now, RDF has been used by Wikipedia, BBC, the Library of Congress and many other organizations to make their data interoperable between one another through the so called Linked Open Data (LOD, www.linkeddata.org) community (Berners-Lee 2006); an important development to be revisited in the case study in Chapter 6.5.

2.4.4 Copyright

The above described implementations and developments can be seen as data-based successors of the cataloguing practices of describing cultural heritage artefacts. The final and current step is the mass digitization of the collected items themselves that began in the 1990s. For instance, the French National Library launched its digital library Gallica (http://gallica.bnf.fr) in 1997 focusing on the digitization of rare or unique works of outstanding French artists and authors. By the same token, the Electronic Libraries Programme (eLib) was initiated by the Joint Information Systems Committee (JISC) in the UK in 1994. Although it did not aim at the development of a single digital library, it has funded a variety of studies, pilot projects and implementations of operational services related to the digitization of cultural heritage to this very day (Raitt 2000). As an open source project, Project Gutenberg (www.gutenberg.org) goes even further back in time. Launched in 1971, it began digitizing out-of-copyright books by volunteers typing the content of a book into a text file. Provided in various formats, it is a freely accessible repository of about 36,000 e-books.[6]

It was, however, Google's launch of its mass digitization of books project that brought the notion of copyright to the attention of artists, publishers and librarians alike. "Google Print", announced in December 2004 and later renamed

[6] http://www.gutenberg.org/wiki/Gutenberg:The_History_and_Philosophy_of_Project_Gutenberg_by_Michael_Hart#The_Project_Gutenberg_Philosophy, last access: 13 June 2011.

into “Google Book Search” in November 2005 (http://books.google.com), is about the digitization and indexing of major libraries.[7] Google provides full text searchability of the content as well as differentiated access depending on the copyright status of the content. In a sense, Google returned to the origins of its ranking algorithm initially developed by the founders of Google, Sergey Brin and Larry Page, to rank the books in the University of Stanford digital library according to a citation analysis (Leetaru 2008). The basic idea is to evaluate the relevance of a book or paper based on the number of times it is cited by others. As it was applied to web-pages and to the number of incoming links from other web-pages, Google’s success story began.

Google Book Search caused a lot of turmoil in terms of copyright and the future role of libraries (Bearman 2006; Bjorner 2006; Lessig 2006; Duguid 2007). Google’s attitude of “digitize first and ask later” with respect to the copyright status of the books digitized resulted in associations of artists and publishers to sue Google for copyright infringement. The lawsuits that followed are still at court and brought to the attention of the public through continuous reporting (e.g. Page 2011). During the authoring of this dissertation, the courts have not come to a definite judgement upon this issue.

As a rule of thumb, works can be seen as being out-of-copyright 70 years after the death of the longest living creator or its publication (European Commission 2010a:10). However, the legal situation varies to such a considerable degree from state to state that, in practice, only works from before 1870 are safe to be deemed out-of-copyright. The situation is getting even more complex once the grey area of orphan and out-of-print works are taken into consideration (Hirtle 2008). Works are referred to as orphaned when the copyright holder is either not identifiable or locatable. By contrast, out-of-print works are under copyright but not produced anymore (Lessig 2006; Lavoie and Dempsey 2009).

The European Commission estimates that, within Europe, about 3 million books (13% of the total number of in-copyright books) and 225,000 films are orphaned

[7] See an up-to-date list of partner libraries at http://books.google.com/googlebooks/partners.html, last access: 13 June 2011.

(European Commission 2010a:5) – a considerable amount of cultural heritage. In a final report commissioned by the EC (HLEG on Digital Libraries 2009), a high level expert group recognized the importance of addressing orphan and out-of-print works. One of the main conclusions of the report was;

> "*that digitisation and online accessibility needs to be achieved in full respect of the current copyright rules, while for cultural institutions there is the need for copyright reform and further harmonisation at European level to create the appropriate conditions for large scale digitisation*" (HLEG on Digital Libraries 2009:6).

In line with this argumentation, the EU launched its own digitization project called Europeana, which is the case study of this dissertation and will be discussed in detail further below.

As a reaction to Google, major technology companies and not-for-profit organizations came together to form the Open Content Alliance (OCA, www.opencontentalliance.org) in October 2005 (Leetaru 2008). Having well known companies and Google competitors such as Adobe and Microsoft as its supporters, the OCA only digitizes books with the explicit permission of the rights holders (Coyle 2006). The digitization itself is processed by the Internet Archive – a not-for-profit organization mostly known for archiving the WWW (Howell 2006; Kallinikos et al. 2010a).

In addition to Google's lack of asking for the permission to digitize in-copyright books, a second line of criticism addresses the de-facto privatization of public domain books. While books scanned by the OCA and hosted by the Internet Archive remain in the public domain, Google acquires exclusive rights to the digitized books. The library holding the print version used for scanning is allowed to give access to the digitized version via its own online service. Anything that goes beyond that, however, requires Google's permission (Leetaru 2008). As a result, Google would own a monopoly with respect to out-of-print as well as public domain books if its digitized version was the only one available (Samuelson 2010). The contracts between the various libraries and Google are not always made public and, therefore, it is not clear for how long Google receives exclusive rights on the digitized content. According to a report

commissioned by the EC, the standard agreement contains a 15-year period (Niggemann et al. 2011).

As a consequence, Google may not only gain an unfair advantage over its competitors but also may restrict access, even censor or shut-down the service entirely if profit-interests demand it. For instance, Google's collaboration with the Chinese government in terms of blocking certain web sites, although discontinued by now, is still well known (Branigan 2010). Less well known is Google's plan to close its video archive. Prevented by a public outcry, the announcement caused some concerns that a similar fate may await Google's book archive (Barron 2011). While it is not the intention of this section to paint Google as an evil-doer, it does want to raise attention to the fact that Google is not a custodian of cultural heritage by any means. Technological developments in the mass-digitization of cultural artefacts are an important aspect of the future of our past. High-tech companies such as Google play a crucial role as innovators in terms of access and findability (Coyle 2006). However, memory institutions as guardians of cultural artefacts and, ultimately, of heritage provide for more than just mere access to things – digitized or not. Apart from the issues raised in this section, there is one considerable point that distinguishes one from the other. That point being preservation and persistence, it will be the focus of the following section.

2.4.5 Digital Preservation

The preservation of digitized as well as born-digital artefacts and their persistent accessibility are probably the most challenging tasks for memory institutions (Russell et al. 1999; Bennett 2001; Becker and Rauber 2011). The reasons lie in the makeup of digital artefacts consisting of bits, standards and their correct processing by corresponding soft- and hardware (Kallinikos et al. 2010a; Blanchette 2011; Faulkner and Runde 2011). Some digital artefacts, such as hypertext, may not have boundaries or may be constantly updated (Buckland 1997; Hjorland 2000). Material carriers such as CD-ROMs or hard disks have relatively short life-spans making them unsuitable for long-term storage. Still,

copying bits and bytes from one carrier to another does not solve the problem either, since ICT standards are prone to become obsolete in ever shorter periods of time preventing the data from being accessed (Kuny and Cleveland 1996; Lyman and Kahle 1998). There are already examples of historically valuable data having become unreadable over the period of a few decades; examples include the BBC, NASA and the White House (Wiggins 1996; BBC 2002; McKie and Thorpe 2002).

In light of these observations, Russel et al. (1999) discuss three possibilities for the preservation of digital cultural heritage artefacts. The first possibility is to conserve the actual soft- and hardware along with the data, thus, allowing access to the cultural artefacts in their original format. This is, however, the most impractical solution as it would turn every memory institution into a technology museum. Second, the authors propose to emulate past soft- and hardware standards on contemporary platforms requiring the detailed documentation of the digital cultural artefact and the standards used for their processing. The last possibility is the migration of the data from an obsolete to a newer standard. The third and, to a lesser degree, the second possibility, however, may end up distorting the data. Migration, on the one hand, inevitably changes or may even lead to loss of data while emulation, on the other hand, may present data not quite the way it would have been presented with the original soft- and hardware. As a result, memory institutions would not be able to guarantee the authenticity of the cultural artefacts meant to be preserved (Hirtle 2000; Lyman 2002; Allison et al. 2005; Klump 2011).

Within this context, online content poses a significant problem (Lyman 2002). An increasingly important part of social interaction and cultural expression, online content is recognized by memory institutions as part of our heritage, which needs to be canonized and archived (Lyman and Kahle 1998). Suffice to say, the preservation of online content is quite a daunting task. It is not only the sheer amount of websites and other web resources that poses a challenging problem, but also the fact that online content is ephemeral and constantly updated (Allison et al. 2005; Kallinikos 2006a). So far, the task has been addressed mostly on a national level under the term web-archiving (CLIR and

Library of Congress 2002; Gladney 2007). For instance, the archive of UK websites, led by the British Library (www.webarchive.org.uk), selects websites nominated by partnering institutions and the public in general. The Austrian National Library is in the process of building its own so-called Web@rchive Austria[8] - a service that automatically harvests every .at domain, complemented by theme and event specific websites. Recently, other internet services have also been recognized by memory institution as illustrated by twitters donation of its whole public archive of tweets to the Library of Congress in 2010.[9]

Digital artefacts are, in fact, computational operations based on data and their correct processing (Saidis and Delis 2007; Ekbia 2009a; Kallinikos et al. 2010a; Faulkner and Runde 2011). By contrast, the traditional concepts of conservation, preservation and authenticity were developed with respect to material objects raising the question whether they still apply in the digital domain (Hirtle 2000; Marton 2010). A telling example is the Internet Archive (www.archive.org). In its efforts to archive the WWW, the Internet Archive makes snap-shots of web-pages as they are rendered in the web browser. Each time a modification of an archived web-page is registered, another snap-shot is taken and stored with the previous versions (Lyman and Kahle 1998; Howell 2006). A snap-shot, however, is not the web-page. In fact, only the content and lay-out of the web-page is being preserved. This works well for static web content but does not capture the underlying functionalities of dynamic web-pages assembled through data-base operations. For instance, one may see the Google homepage from as far back as 1998 but it is not possible to make a search query that would deliver the search results from that period of time. From this perspective, the Internet Archive does not preserve the WWW but rather the documentation of the WWW.

Given above stated arguments, memory institutions may need to, first of all, abandon the notion of conservation entirely. Preservation, on the other hand, needs to be reconceptualised (Gladney 2006). In its original meaning, preservation refers to the content of, in the end, written or printed documents.

[8] http://www.onb.ac.at/about/webarchivierung.htm, last access: 01 June 2011.
[9] See the press release at http://www.loc.gov/today/pr/2010/10-081.html, last access: 01 June 2011.

With digital cultural artefacts, it is the preservation of data and meta-data required for the emulation of a recognizable cultural artefact. This is sufficient for digitized and static born-digital cultural artefacts, since their only requirement is to be displayed correctly. A dynamic born-digital cultural artefact, however, needs to be transformed or, one could say, frozen into a persistent form, which, as the example of the Internet Archive illustrates, results in a different digital artefact altogether. Hence, future generations may need to rely, to a considerable degree, on the documentation of digital cultural artefacts for the canonization of their heritage rather than on digital cultural artefacts themselves (Lyman 2002).

Preservation is more than just the storage of bits and bytes. At least for memory institutions, it also involves the preservation of recognizable digital cultural heritage artefacts. By the same token, memory institutions also need to take care of the descriptions of the digital artefacts in order to provide for persistent findability as well (Hitchcock et al. 2007). Considering the fact that data as well as metadata cut across institutional boundaries, communication genres and artefact types, so must also their preservation. In this sense, the preservation of digital cultural heritage artefacts is not only up to memory institutions alone but requires a concerted effort of technology companies, data providers, governments, the legal system, research and development and so forth. Multidisciplinary research projects such as the International Internet Preservation Consortium (IIPC, www.netpreserve.org), Digital Preservation Europe (DPE, www.digitalpreservationeurope.eu), Preservation and Long-term Access through NETworked Services (PLANETS, www.planets-project.eu), Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval (CASPAR, www.casparpreserves.eu) and many more are a testimony for the crucial role of preservation and, ultimately, the future of our past.

2.5 Discussion

Conceptually, the wider field of social memory studies oscillates between collected and collective memory (Olick 1999). As discussed above, this presents itself as a problem since the question whether society or rather social operations

can remember or forget are indeed forgotten in a blind spot of the research field at large. The blind spot is the locus of remembering in the minds and/or mnemonic practices of persons being individuals and members of a collective at the same time while the concepts of cultural (Assmann 1999) or social (Olick and Robbins 1998) memory, as they are mostly applied within the field, turn out to be collections of reminders only. In other words, cultural or social memory is observed as externalized personal memories sustained by practices of canonization or stored in archives (Assmann 2008a). Canonization, in this sense, is the continuous reminding of members of a certain collective of what it means to be a member of that very same collective. The archive, on the other hand, is the storage of forgotten externalized memories, which are, indeed, graveyards of knowledge bereft of their capabilities to remind.

In this sense, social or cultural memory is, in fact, a social psychological memory that, as a concept, falls back on the individual as an agent for remembering what is kept in society's storage compartments of memory traces (Archer 1995:98). By contrast, social systems theory makes a clear distinction between minds and the social and therefore provides for an alternative view on social memory as an autonomous, systemic social operation that remembers and, more importantly, forgets (Esposito 2008). Social memory, therefore, is not a mere storage compartment. In this spirit, the theoretical chapter will elaborate on the notion of social systems as communicative systems rather than a collection of individuals (Luhmann 1996b). Thus conceived, the dissertation offers an alternative approach to the phenomenon of interest in order to contribute a different perspective and, as a result, to offer new insights. By no means is this to be understood as a claim for having a better view on this conceptual gap than those provided by other frameworks.

The second gap identified in the literature of memory studies is the conceptualization of digital memory. The study of digital memory is still a terra incognita (Garde-Hansen et al. 2009a). For one, the immense possibilities for data storage and retrieval at the disposal of contemporary society has led to a revitalization of the storage metaphor in the studies of social memory observing data-bases as the new storage compartments of memory traces (Riegler 2005;

Assmann 2006; Bell and Gemmell 2009; House 2009; Mayer-Schönberger 2009). Equating storing and retrieving with remembering and deleting with forgetting, however, underestimates the performative aspects of social memory, in general, as well as the continuous updating of data-bases required for their operational maintenance (Foerster 2003:101-31; Kallinikos 2006b). On the other hand, the study of the information habitat as the new social memory paradigm revolves around the primacy of search engines and social tagging (Esposito 2002; Weinberger 2007). As a consequence, the role of memory institutions in the so-called information society is marginalized if not declared to be an old-fashioned way of doing things.

However, as the discussion of the literature on memory institutions demonstrated, the field of memory institutions is stepping into the online world driven by a lively and active debate within and between the respective communities of librarians, archivists and curators. To interpret the shift from a social memory paradigm of institutionalized memory to the information habitat as the end of memory institutions as such is vocally debunked as a myth by these communities (Kuny and Cleveland 1996). Quite the opposite, the role of memory institutions, though changed, may be even more important with respect to the ephemeral and mutable existence of born-digital, online content (Lyman and Kahle 1998; Lyman 2002). Still, memory institutions and their development into digital libraries are mostly seen as a technological enterprise. The paradoxical relationship between remembering and forgetting is transformed into a technical problem that can be solved by state-of-the-art IT infrastructures (Bishop et al. 2003; Graham 2005). Similar to the equating of social memory with data-bases, computational storage and retrieval is deemed to solve the problem of forgetting – for better or worse (Bennett 2001; Mayer-Schönberger 2009). Given these observations, there is quite a gap between the conceptualization of digital memories as new ways of remembering and the emergence of digital libraries only bridged by the common storage metaphor (Assmann 2006; House 2009). From this perspective, the dissertation aims to position the study of digital libraries within the wider context of the paradoxical relationship between remembering and forgetting.

Based on the analysis of the research domain of social memory studies and library and information science, the research interest can be defined as follows. Given the paradigm shift in social memory from memory institutions to the information habitat, from ex-ante classification to ex-post ordering, the question arises as to what happens to libraries, archives and museums as institutions of forgetting.

2.6 Conclusion

For Halbwachs (1992), the individual remembers in unison with other members of the same collective. As Esposito (2008) points out, collective memory has become more and more limited with the increasing complexity of society. She suggests that a social memory emerged based only on social operations without the involvement of mental memory processes for its own sustainment. Due to developments in communication technologies from writing and printing to relatively recent innovations in ICT, social interaction has been enabled to be mediated in increasingly decontextualized ways leading to a high degree of variability in terms of who is communicating with whom about what topic over what period of time. For instance, writing frees the communicating persons from the necessity of being at the same place at the same time. Mass media constructs its message in a way that is understood by an anonymous audience. Finally, the internet or rather many of the services built on top of it enable the kind of many-to-many communication where creators and audience are one and the same (Weinberger 2007).

The process of increasing abstraction of communication from face-to-face interaction (language from speaking) goes hand in hand with an increasingly abstract structure of social memory (Esposito 2002; Boyden 2003). For instance, the beginning of writing lies in lists as reminders for individuals; hence, social memory was rudimentary or rather undifferentiated from individual memories as it is the case in Halbwachs' concept of collective memory. It was the individual that had to remember the content. Alphabetic writing, at first, was used as a memory device supporting the mnemotechniques and rhetorical techniques of the

orator. Memory was a "walkable" place. In a metaphorical sense, it was the "memory palace" – a memorized place the orator imagined himself to walk through in order to retrieve the memory stored in different corners. In a literal sense, it was the places for musing – the museions of ancient times as well as the monasteries and mosques of medieval times. Be it in a literal or metaphorical sense, it was not the content that was memorized but its location in the palace. Hence, it was more important to know where to find content rather than the content itself.

The increasing autonomy of social memory, beginning with alphabetic writing, came to its conclusion with printing and the popularization of reading. By the same token, libraries, archives and museums emerged engaged in the institutionalization of social memory through canonization. Due to the rising number of cultural artefacts, specifically mass print, the paradigm of location was abandoned in favour of an abstract second-order of surrogates. In addition to content not being remembered, the location of the content was forgotten as well. Instead, it was the ex-ante rules of the card catalogue and the underlying taxonomy – incorporated into the material makeup of the card catalogue itself – that allowed the retrieval of the items represented. Finally, networked and binary-based, digital media allows for an immense variety of ways order can be brought into a heap of data after its collection. The crucial shift is the computationally sustained and algorithmically processed shift from ex-ante rules and rationales to ex-post search queries and data mining. The catalogue is forgotten as well, what remains are algorithms and the continuous up-dating of data-bases in order for them to remain informative.

Thus conceived, the relationship between communication media, the ordering of artefacts and mnemonic devices can be summarized into social memory paradigms each understanding memory in a specific way distinct from the other (Radstone 2000b). As summarized in Table 4, the social memory paradigms of social context and memory palace focus on oral remembering by means of mnemonic devices or mnemotechniques. In contrast, the paradigms of memory institutions and information habitat focus on mediated memory taking the forms of material artefacts and binary-based digital media.

Social memory paradigm	**Related concepts**	**Paradigmatic medium**	**Order of order** Weinberger (2007)	**Paradigmatic examples**
Social context as reminder for individuals No autonomous social memory	- Collected memory (Olick) - Collective memory (Halbwachs) - Communicative memory (Assmann and Assmann) - Prophetic memory (Esposito)	- Ideogramms - Pictograms	None	- Storytelling - Rituals - Myths
Places for musing Memory palace	- Collective memory (Olick) - Rhetorical memory (Esposito)	- Alphabetic writing	- First-order Order of the things themselves	- Lists - Mnemotechniques - Rhetoric - Inventory (Book catalogue) - Museion, monasteries, mosques
Institutionalized memory	- Cultural memory (Assmann and Assmann) - Prosthetic memory (Nora) - Modern memory (Esposito)	- Mass media (print, radio, television)	- Second-order Order of representations of things	- Canonization - Taxonomies - Cataloguing (Card catalogue, OPAC) - Libraries, archives, museums,
Information habitat	- Algorithmic memory (Sluis) - Telematic memory (Esposito) - Digital memory (e.g. Garde-Hanse, Hoskins and Reading) - Social network memory (Hoskins)	- Binary-based, digital medium	- Third-order Ex-post, performative ordering	- Search engines - Folksonomies - Facetted browsing - Internet, WWW

Table 4: Social memory paradigms, media and order.

The distinction into social memory paradigms is an analytical tool. The categories are not explicitly historical or mutually exclusive. Storytelling, for instance, is still to be found in contemporary society but also finds a new life in, say, fairy tales read to children as a bed time story or in the institutionalized form of oral history. Still, with the rise of ICT and the internet a new paradigm of social memory is emerging addressed by the field of digital memory studies. Within this context, this chapter discussed the peculiarities of the information habitat as a paradigm of binary-based computational operations as well as networked artefacts and relationships. As a second area, the discourse on digital libraries was introduced covering the developments within memory institutions as they step into the information habitat.

Finally, the discussion of the literature revealed following gaps;

1. The conceptual primacy of the storage metaphor leads to a theoretical dilemma as social and cultural memory is deemed to consist of individuals as remembering agents to be reminded by memory traces collected in storage compartments. Thus conceived, the field of social or cultural memory is not clear about the social aspects of remembering and forgetting. An alternative theoretical framework, provided by social systems theory, allows for a different approach based on the conceptualization of social systems as autonomous (but not autarkic) communicative operations with a distinct structure for social remembering and forgetting.
2. The reliance on IT and its immense capabilities for data storage, processing and retrieval redefines the paradoxical and self-referential relationship between remembering and forgetting into a technical problem. Thus deemed as the solution for forgetting, the dynamics of memory are underestimated. As discussed in this chapter, the intricacies of social memory are more complex and do not easily fit into the technological framework of IT.
3. The primacy of search engines and so-called web 2.0 services within the study of digital memories ignores the transformation process memory institutions are going through as they are stepping into the information habitat. Memory institutions are indeed changing in order to remain

memory institutions. Therefore, the study of this change will contribute to the understanding of the new social memory paradigm at large.

Based on these conceptual and methodological gaps, the research interest was defined broadly as the changes libraries, archives and museums as institutions of forgetting go through when they step into the information habitat and, as a consequence, shift from ex-ante classification to ex-post ordering. In the following theoretical chapter, the conceptual gap will be addressed leading to the formulation of specific research questions. The research questions will, then, build the foundation for the chapter on methodology and the empirical study.

3. THEORY

The theory of social systems was already briefly discussed with respect to the broader field of cultural and social memory studies. Based on the work of Niklas Luhmann (1996b; 1998), social memory was proposed as an autonomous operation of social remembering and forgetting distinct from mental or individual memories (Esposito 2002). Thus conceived, the theory offers an alternative concept in contrast to the traditional dichotomy of collected and collective memory as well as the storage metaphor (Gregor 2006). Although a key theoretical building block in Luhmann's later works, memory is definitely not one of the leading notions of the theory of social systems (Dimbath 2011). This chapter, therefore, is dedicated to the exploration of Luhmann's highly abstract theorization of society and to the positioning of memory within the wider conceptual landscape of social systems theory.

3.1 Introduction to Niklas Luhmann

Luhmann's development of his theory of social systems can be roughly divided into three stages of a journey towards explaining and understanding society as a self-referential phenomenon. The first stage is directly linked to Talcott Parsons' system theory and his analysis of the structure of social action (e.g. Parsons 1968). A student of Parsons, Luhmann shared a passion for grand theories claiming to be universal in their explanatory power. However, in the early 1980s Luhmann made a decisive step away from Parsons's framework by shifting from action to communication as the elements of social systems culminating in the publication of "Soziale Systeme" (Social Systems) in 1984 - probably his most well known book worldwide (Luhmann 1996b; Stichweh 2000). The shift was a result of Luhmann's growing frustration with sociological theorization or rather the lack thereof. The discipline failed to theorize contemporary phenomena, hence his criticism. By constantly reinterpreting the grand theorists and founding fathers of sociology, the sociological frameworks applied remain inadequate with respect to unprecedented issues such as the ecological movement and high-risk

technologies (Luhmann 1991; Luhmann and Baecker 2002). Reaching beyond the boundaries of sociology, he incorporated developments from the field of general systems theory, specifically second-order cybernetics that had brought a paradigm shift into general systems theory during the decades that followed World War II (Sciulli 1994; Froese 2010).

General systems theory is a transdisciplinary effort to provide for a unifying terminology and methodology across all scientific disciplines – from physics to biology, social science and humanities (Bertalanffy 1973). Integrating theories such as information theory (Shannon and Weaver 1998) and cybernetics (Wiener 1948; Ashby 1957), general systems theory mostly focused on so-called "open systems" - systems that do not seem to be governed by the laws of thermodynamics. "Closed systems", such as bottled gas, inevitably tend towards entropy – maximum disorder or noise. By contrast, open systems are open towards their environment and are subject to a continuous inflow of matter, energy or information processed by the organized complexity that is the system - a faculty that has been coined negentropy (Beniger 1986; Shannon and Weaver 1998). The paradigmatic example is, of course, a living organism that maintains its inner structure until its death.

In their simplest form, open systems were described by basic input-output-models such as Shannon's model of information transmission discussed below. This changed with the self-referential turn introduced by second-order cybernetics. Systems theory shifted from observ*ed* systems to observ*ing* systems that involved the observing cybernetician as an observing system observing observing systems (Foerster 2003:289; Brier 2008; Froese 2010). A second important change was the shift from a whole/parts distinction to a distinction of system/environment. The system is defined by its difference to its environment or rather the system *is* the difference between system and environment (Luhmann and Baecker 2002:66-90). Thus conceived, the environment is a conditional aspect of the existence of the system without determining it. The system, in turn, is structurally autonomous but not autarkic from the environment.

Luhmann incorporated these notions by defining social systems as autonomous structures defined by their difference to their respective environment based on communication (Luhmann 1996b). In detail, he translated the self-referential concept of self-organizing systems, which maintain their borders based on their own internal operations. In contrast to trivial machines that, given the same input, always deliver the same output, self-organizing systems are non-trivial and historical, hence, the same input may deliver different outputs just as different inputs may deliver the same output (Foerster 2003). While the first is a-temporal, the latter refers to time and, as it will be discussed in more detail further below, memory in order to distinguish between before and after (Luhmann 1997). Either way, it is the system that determines what counts as environment since it is only open to certain input while being agnostic to whatever else happens "outside" of the system. Both, trivial and non-trivial, are determined by their internal structure but non-trivial systems have a level of complexity that makes them unpredictable.

In his third stage, Luhmann finalized his theorization on self-referentiality and social systems by incorporating the theory of autopoiesis into his thinking (Stichweh 2000). The theory was originally developed by Humberto Maturana and Francesco Varela as a phenomenology of the living defining living beings (in contrast to "non-life") as a process of continuous self-reproduction.

> "*Our proposition is that living beings are characterized in that, literally, they are continually self-reproducing. We indicate this process when we call the organization that defines them an autopoietic organization*" (Maturana and Varela 1992:43).

Although intended to be a biological concept, the paradigm of second-order cybernetics was enriched with the concept of autopoiesis making it, according to Luhmann, a general characteristic of all complex systems – biological systems (life), mental systems (psyche) and social systems (Luhmann 1986). Thus, Luhmann finalized his conceptualizations of self-referentiality by describing social systems not only as self-organizing – systems that maintain their borders and, hence, themselves – but also as self-reproducing – systems whose input and output are themselves.

Luhmann's autopoietic turn found its most coherent and elaborate expression in his opus magnum "Die Gesellschaft der Gesellschaft" (The Society of Society) published in 1997 – one year before his untimely death due to illness (Luhmann 1998; Lee 2000a). Unfortunately, the book was never translated into English, which may have clarified some misunderstandings in the wider community of academics concerning the conceptualization of social systems as autopoietic systems (for an example see Mingers 2002; for a critical reply to Mingers see King and Thornhill 2003). Luhmann's elaborations on autopoiesis were not a one-to-one adaptation of a biological concept based on a conclusion by analogy (Varela 1981); that is, social systems are not like living organisms in an analogous way but rather functional equivalences according to abstract criteria (Sciulli 1994). To abstract characteristics from a specific research context into more general categories is a methodological approach typical for general systems theory. In other words, for autopoiesis to come from biology does not necessarily make those criteria biological (King 1993; Urrestarazu 2011).

The abstraction of autopoiesis out of its biological context results basically in two criteria. For one, autopoietic systems are operationally closed. That is, they do not only create their own structure but also they are operationally autonomous from their environment (Luhmann and Baecker 2002:110). In other words, autopoietic systems do not only self-organize the elements they are made of but self-reproduce the elements themselves. From a temporal perspective, the elements are operations forming "*a network of dynamic processes whose effects do not leave that network*" (Maturana and Varela 1992:89; see also Luhmann 1998:65). There is no living outside of life, no thinking outside of the mind and, as Luhmann defines communication as the basic operation of society, no communication outside of society.

The second criterion derives from the term "poiesis". After a discussion with a friend and Aristotelian philosopher, Maturana learned about the term poiesis as an activity that produces something – such as a piece of art – in comparison to "praxis" as an activity that is an end in itself – such as being virtuous (Luhmann and Baecker 2002:111). The neologism "autopoiesis", therefore, was created by Maturana to stress that the living was the continuous *result* of molecular

processes rather than the molecular processes themselves, which would have been more appropriate to be referred to as "autopraxis" (Maturana et al. 2011). The term indicates "*that the system is the product of* […] *its own activity, and not simply self-sufficient activity as such*" (Sciulli 1994:41). As a consequence, autopoiesis is a process of *production* of a system out of itself and is to be distinguished from self-*creation* – an important distinction to be revisited in terms of the epistemological foundation of this dissertation. Translated into social systems theory, the term autopoiesis leads to a, probably counterintuitive, conclusion; it is communication that communicates or rather it is communication that produces communication. As it will be discussed below, social systems self-reproduce themselves as communicative systems (rather than as human agent collectives) through communication (Luhmann 1986; Luhmann 1996b).

The theoretical framework of the dissertation is solely based on the final phase of Luhmann's work and therefore reflects his thinking as it is expressed in "Die Gesellschaft der Gesellschaft" (Luhmann 1998) and other directly related publications. The following concepts, therefore, will introduce the theory of social systems in its latest incarnation based on Luhmann's completed autopoietic turn.[10]

3.2 Observation and Distinction

There are many roads to Niklas Luhmann's theory of social systems (Arnoldi 2001). One possible way is to begin with the basic concept of observation. Based on the work of George Spencer Brown (1972), observation is defined as an operation of making a distinction. An observation, so to speak, can only indicate or focus on one side of the distinction referred to as "marked" while the other side of the distinction is "unmarked". Purposefully defined in such abstract

[10] Hence, the criticism of Jürgen Habermas that led to a famous debate between the two social theorists and also culminated in the publication of a book (Habermas and Luhmann 1971) will be left aside. In the aftermath of the debate, both addressed, clarified and even implemented critical points raised by the other into their academic work and theories, which renders their initial criticism practically invalid (Bausch 1997; Kjaer 2006). Today, their debate may be of interest for historians of sociology but, outdated as it is, an in-depth discussion of the critical points raised by Habermas would be unnecessarily confusing and, thus, pose a diversion.

terms, observation does not only apply to the cognitive capabilities and thought processes of a consciousness but also to social systems (Luhmann 1998). As it will be discussed in detail below, communication is a social operation of making consecutive distinctions.

Observation, thus defined, is a constructive act of forming a form – a difference between what is observed and what is not observed. As such, this definition is not very informative, as it does not refer to anything of essence. A difference is where one "entity" ends and another begins. It is when future turns into past – the now. It is, so to speak, the border that makes a form observable. In this sense, one can understand an observation as differance – the formation of a form (Derrida 1976:63; Luhmann 1996b). Hence, this research is based upon a philosophy of difference (Heidegger 1969; Derrida 1982; Luhmann 2006), which will be the guide throughout the entire dissertation – be it conceptualization and theory, research question and methodology. Consequently, whenever things, entities, identities and so forth are mentioned, they shall be understood as forms that are differences.

For instance, from a social perspective it is the difference between me (Ego) and the other (Alter-Ego). I am me *because* I am not you (Mead and Morris 1934; Maturana 1988). From a factual perspective it is this and not that; from a temporal perspective it is before and after - past/future (Cooper 2005:1691). However, a difference has no place of its own; it is not out there to be found but it is brought into the world. Consequently, a holistic approach is simply impossible (Luhmann 2002a:101), since there is always a difference of what-is-observed and what-is-not-observed. For instance, by writing about difference itself, another difference was introduced that is the distinction between difference/essence. Thus conceived, the unity of the distinction or rather the distinction itself cannot be observed since it is only a difference between one and the other. In other words, making a distinction and marking one side is one single operation – one cannot distinguish without marking and one cannot mark without distinguishing (Luhmann 1998:69).

The distinction itself is the co-called blind spot – signified by the "/" sign (Luhmann and Baecker 2002). The metaphor of the blind spot was introduced by von Foerster referring to the spot of the eye's retina where the nerve fibres lead into the brain. Hence, that area is not covered by photo receptors. The eye is literally blind at that spot but still;

> "*this localized blindness is not perceived as a dark blotch in our visual field (seeing a dark blotch would imply "seeing"), but this blindness is not perceived at all, that is, neither as something present, nor as something absent: Whatever is perceived is perceived "blotchless""* (Foerster 2003:212).

The eye is blind to its own blindness. In more abstract terms, observation is possible *because* not everything can be observed at once. In the end, there is no final, absolute synthesis incorporating every and any as well as potential and actual difference(s). Every observation (including this dissertation) is based on distinctions (Maturana 1988).

The notion of making a distinction, however, is also to be applied to the observer who is making the distinction. As already stated above, entities are not essentials but differences. Therefore, observing an observer is observing a difference between what is observed (as an observer) and what is not observed (as an observer). In other words, *the observer is just another observation* – a distinction made and marked by the same observer observing him/her/itself or by an observer observing another observer. In the first case, one may speak of introspection or reflection.

> "*To become aware of distinctions, is called observing. To observe oneself as the maker of distinctions, therefore, is no more and no less than to become conscious of oneself*" (Glasersfeld 1991).

However, that does not mean that one simply turns one's eyes inwards to see oneself (Rasmussen 1998) but rather distinguishes oneself based on the distinctions one makes. In other words, the observation observes its own distinctions by applying the very same distinctions it observes. At this point, the paradoxical nature of distinction and observation emerges and, ultimately, leads to questions concerning ontology and epistemology – questions that will be revisited further below once the basics of Luhmann's thinking have been

discussed. For the moment, however, it is sufficient to treat this paradox as a logical or methodological one.

The paradox that appears when one discusses self-observation is referred to as *re-entry* and is a basic characteristic of any autopoietic and complex system (Spencer Brown 1972; Luhmann 1997). The basic distinction with respect to systems is system/environment. However, autopoietic systems are self-referential, that is, whenever they observe themselves or their environment the system/environment difference re-enters the system. The observer observes the distinction between observer and observed based on the very same difference of observer and observed. This is just another way of saying that *the observer is just another observation*. Going back to the example of the eye, the eye cannot see itself. In order to know what one looks like, one needs a mirror to see oneself looking back. In order to know what person one is, one needs to observe "the others" observing back (Mead and Morris 1934; Maturana 1988). This notion of re-entry plays a crucial role in the understanding of complex systems that are not able to distinguish the system/environment difference from the observation of the system/environment difference. Hence, they turn "intransparent" to the outside observer but also to themselves (Luhmann 1997).

Any observation that is capable of looking at both sides of the distinction by crossing from one to the other or, in other words, by marking the previously unmarked side of the distinction (thus, unmarking the previously marked side of the distinction) results in re-entry (Spencer Brown 1972; Baecker 2006a). Self-observation, therefore, is the marking of the observation itself – the marking of the self in the distinction self/other. Observation of the other or rather observing another observer observing is called second-order observation, which is not to be understood as having a superior viewpoint. Second-order observation reveals the blind spot and, therefore, the distinction another observer applies. However, as such it is still an observation with its own blind spot that, again, takes another observer in order to be addressed (Luhmann 1998). As a consequence, social science, for instance, should be seen as a way of observing society of which it is a part of, with its own blind spot based on paradigms (Giddens 1987; Kuhn 1996; Luhmann 2002b; Moser 2004b). In the end, no matter how hard one tries, the

observer finds her/him/itself in the observation (Brier 2009). In other words, observation is an operation of making-sense rather than getting-a-sense of the world based on a distinction. Observation is always self-observation not in a tautological but rather autological sense because;

> "*the knowing system* [the observer] *is itself one of the objects it has to know: when it describes its objects it then also describes itself, and the description modifies the object to be described*" (Esposito 1996:271).

3.3 Observation and Information

The notion of a non-existent difference in combination with the way observation is conceptualized leads directly to the central role information plays in general systems theory and the theory of social systems. As already mentioned above, with the autopoietic turn in systems thinking complex, unpredictable systems are not only conceptualized as self-organizing but self-reproducing – systems whose operations are the result of the very same operations they consist of (Maturana and Varela 1992). This applies also to the distinction between system and environment, which is produced by the system itself (Luhmann and Baecker 2002:92). It is this difference that makes the system observable to itself but also to other observers. In other words, a system is a form (a difference) differentiating an inside (system) from an outside (environment). Thus conceived, a system does not merely maintain its borders; it *is* the difference between system and environment (Luhmann 2006). A system observing itself observes the difference and, therefore, makes a distinction between itself and its environment within the system. The difference re-enters the system as a distinction made by the system. Hence the system/environment difference appears twice – as the difference *produced* by the system and as *observed* inside the system (Luhmann 1998:45).

The, so to speak, copying of the system/environment difference into the system is based on the *operational closure* of autopoietic systems (Maturana and Varela 1992:89; Maturana et al. 2011). The operation that reproduces itself and leads to the emergence of a systemic structure cannot operate outside of itself. The observation of the system or its environment by the system is based on the

system and the distinctions it makes. In other words, the autopoietic system has no direct connection to its environment but rather only to the re-entered version of the system/environment difference. It observes not only itself but also its environment within the system. At a first glance, operational closure may seem to be the exact opposite to the notion of open systems as initially conceptualized by the founding fathers of general systems theory (Bertalanffy 1973). However, operational closure does not refer to thermodynamical closure but rather to the self-reproduction of the system (Luhmann 1998:94).

Being operationally autonomous from its environment, does not mean that the system is autarkic from the environment. In today's terminology, the system, although operationally closed, is structurally coupled to the environment. *Structural coupling* basically refers to recurrent interactions between system and environment where "*the structure of the environment only triggers structural changes in the autopoietic unities (it does not specify or direct them)*" (Maturana and Varela 1992:75). The system is only able to be coupled to the environment in a selective and limited fashion and, thus, reacts only to a selection of events and occurrences. Coming back to the eye as an illustration, the human eye or rather its photo receptors are only irritated by waves within a limited bandwidth. The eye itself is the structural coupling between the brain and the environment. However, seeing is accomplished by the brain based on its internal operations. It is in this sense that autopoietic systems are only *irritated* by environmental perturbations rather than determined through cause-effect chains (as it would be the case with trivial machines) (Luhmann and Baecker 2002:120). In other words, what is observed as an external event by the system is not determined by the environment but rather by the structure of the system itself based on its own operations. Observation is *structurally determined*.

Structural couplings can be imagined to be like channels through which a system reduces environmental complexity into internal, i.e. re-entered selective patterns that are irritable by some events while other events are excluded and, therefore, simply do not exist for the system. Those events fall into the blind spot of observation and can only be addressed by second-order observation (i.e. one system observes another system and the distinctions upon which it operates). It is

those events towards which a system is sensitive that is *information* for the system or rather occurs as information within the system. In other words, selection or rather reduction of environmental complexity through structural coupling is the precursor for information to occur. At this point, the theory refers to the famous definition by Gregory Bateson (2000:459):

> "*Of this infinitude, we select a very limited number, which become information. In fact, what we mean by information - the elementary unit of information - is a difference which makes a difference* […]."

Given above discussed arguments, information occurs based on the observer's self-referential perspective. Information is determined by the observer's structure (Nauta 1972; Artandi 1973; Maturana and Varela 1992). This notion of information is in stark contrast to information seen as something outside waiting to be found (Floridi 2003) or being sent in order to inform (Shannon and Weaver 1998) – a signal, a sign or an artefact in need of being interpreted or as something carrying meaning (Machlup and Mansfield 1983; Boland 1987; Buckland 1991; Capurro and Hjorland 2003; McKinney and Yoos 2010). For instance, Fairthorne (1968:352) defines information as "*an attribute of the receiver's knowledge and interpretation of the signal, not of the sender's, nor some omniscient observer's nor of the signal itself*." At a first glance, the receiver's knowledge is what could be referred to as the structural determinacy of the receiver. In other words, information is the process of an observer (the receiver) observing (interpreting) the observed (signal).

However, the autopoietic approach further radicalizes the observer's structural determinacy by not taking "the signal" as a given. Recognizing a signal as a signal or medium is already an accomplishment from parts of the observer (Lanzara 2009). In other words, the signal is already an attribute of the receiver who differentiates the signal from noise. In more general terms, the first difference – according to Bateson's (2000:459) definition of information as a difference which makes a difference – is an event that occurs through observation by means of distinction and indication. That again depends on the structure of the observer, which refers to the second difference. Information as such only occurs when an event (1st difference) makes a difference for the observer (2nd difference).

The deconstruction of the signal as just another difference of noise/signal leaves nothing behind that actually carries any meaning but rather leads to a difference brought into the world by an observer who actually constructs meaning – the formation of the form, as Derrida (1982) calls it – who makes sense rather than gets a sense. The autological nature of observation and, consequently, information leads to the counter-intuitive conclusion that it is the observer who informs rather than being informed by something "outside" (Esposito 1996). A crucial consequence to be drawn, even the absence of something can be informative or as Bateson (2000:459) puts it: "*The letter which you do not write can get an angry reply*". In other words, when the sender does not send a signal the receiver might still "receive" information (Krippendorff 2009:278-98).

3.4 Communication

Conceptualizing autopoiesis as the key characteristic of every complex, non-trivial system redefines its biological origins to be one possible instantiation or realization of autopoiesis besides others (Sciulli 1994). Biological systems, therefore, come into being through molecular self-reproduction. Other autopoietic systems, according to Luhmann, can be distinguished based on their specific operation of self-reproduction (Luhmann 1986). For mental systems (consciousness or psyche), autopoiesis is thinking that reproduces thinking. For social systems, it is communication that reproduces communication. The following discussion will further explore communication as the distinguishing autopoietic operation of social systems by comparing Luhmann's take on communication as an autopoietic operation with Shannon's information theory as an input-output model.

3.4.1 The Engineering Problem of Information Theory

In terms of observation and information, the question arises as to how one can understand and/or model the exchange of information or rather the mutual information of Ego and Alter-Ego. Such a model is offered by Shannon's (1993)

information theory that has played an immensely influential role in various disciplines as well as in the general understanding of information (see for instance Ashby 1958; Fairthorne 1968; Artandi 1973; Bertalanffy 1973; Introna 1997a; Borgmann 1999; Capurro and Hjorland 2003). Shannon's mathematical theory of communication deals only with the engineering problem of transmitting signals in an optimal way. Although he explicitly states that his notion of information does not apply to semantics (Shannon 1993:1), it has been used for exactly those purposes time and time again – a development Shannon himself regretted (Floridi 2009:24). For instance, Weaver – the second founding father of information theory alongside Shannon – already proposed that this model is applicable to meaning and affecting conduct as well, i.e. semantics and pragmatics (Artandi 1973; Shannon and Weaver 1998:25; Borgmann 1999:132; Lin et al. 2005). Other authors strongly disagree with this approach as they consider equating information with data and communication with signal transmission as a simplistic concept that has led to an objectification of information into thing-like entities (Boland 1987; Lee 2004; Kallinikos 2006a; McKinney and Yoos 2010).

Shannon's model consists of a sender selecting a message out of a range of preselected messages, which is then encoded by a transmitter in order to be mediated through a communication channel. On the other end of the channel, a receiver decodes the message for a given destination. In order to measure the performance of the channel, the basic unit of information as a binary digit (bit) is introduced. Every message can be cut down to a series of the same single difference – be it 0/1, dot/dash, on/off and so forth. The information value of a message does not derive from the content or meaning of the message itself but rather from the number of messages that could have been sent as well. More plainly put, the "informativeness" of a message is not about what has been selected but rather what has not been selected. For instance, the sender sends the result of flipping a fair coin. The information value of that message is 1 bit – 0 for heads, 1 for tails. Correspondingly, the value for sending the result of rolling a six sided dice is 3 bits.[11] A key requirement for a successful transmission is that

[11] E.g.: 000 ... 1; 001 ... 2; 010 ... 3; 100 ... 4; 011 ... 5; 101 ... 6.

both – sender and receiver – have the same list of possible messages, so that the receiver sees the message as the result of throwing a dice rather than flipping a coin three times in a row. Depending on the messages to be sent, the encoding/decoding as well as the channel can be constructed in a way to optimize the performance of the transmission by analysing the probabilities of selection. For instance, the probability of heads or tails is p = 0.5. However, if one takes the English language as an example, the order in which a sequence of letters is transmitted is far from being random. The encoding/decoding can take into account certain attributes of the English language such as that the most used letter is "e" or that a specific letter cannot be followed by a series of other letters (e.g. a "b" is never followed by a "g", "k" and so forth).

A communication channel may be compromised by undesirable disturbances – noise. Taking into consideration that what comes out of the channel may not be what came in, redundancy mechanisms can be built such as sending the same message various times, which of course lowers the performance. Hence, the optimal balance needs to be found between performance and redundancy. As stated above, the information value is not about the content of the message itself but rather about its unpredictability – a notion often interpreted as surprise or novelty (e.g. Luhmann 1996a; Esposito 2004; Kallinikos 2006b). The term "information entropy" basically refers to this relationship of information, noise and unpredictability – redundancy and variety (Artandi 1973; Bertalanffy 1973; Beniger 1986; Simon 1996). Going back to the example of flipping a coin, imagine the channel to be completely noiseless. If one flips a coin with both sides being heads, the information value of the message is 0 bit, since sending the result to the receiver does not make any difference for the receiver. In fact, the receiver does not need to wait for the message at all. In this case, entropy is zero, since flipping the coin does not "deliver" any information. The other extreme example would be flipping a fair coin and sending the result over a channel overridden by noise. Hence, there is no way for the receiver to know which message is the message sent and which message is just noise. In this case, one finds maximum entropy, since flipping the coin on the sender's side, again, does not "deliver" any information at all. In fact, the receiver could start flipping a

coin herself; the probability of finding out the result of the sender would be just as high as getting an undistorted message through the channel by chance.

After the publication of Shannon and Weaver's work, the model was received with much enthusiasm as a model applicable to every level of communication including the level of semantics (Artandi 1973). Ironically, much of the misunderstandings in terms of information theory derive from the meaning of the terms used. In order to avoid further misunderstanding, some propose that the model should be called the "theory of signal transmission" (Littlejohn 1992:61) or "mathematical theory of communication" (Shannon 1993; McKinney and Yoos 2010) rather than "information theory". Boland (1987) even states that the, sometimes very uncritical, work following Shannon's model created a fantasy of "entifying" and therefore making the world controllable (see also Luhmann 1991; Angell and Ilharco 2004) by confusing information with structured data (Lee 2004). As a consequence, the central image of modern times is information without in-formation ignoring communication as a meaningful symbolic action (Boland 1987; Ciborra 2002; Bryant 2008; Kallinikos 2009a; Beynon-Davies 2010; McKinney and Yoos 2010).

The problem with Shannon's model in relation to the social dimension of information – i.e. communication – is the understanding from parts of the destination – the receiver. As already discussed in the section above, it ascribes the signal an objective essence of its own. It assumes that information is, basically, a thing carried from a sender to a receiver and the engineer's problem is to prevent the thing from bouncing into obstacles on the way and, hence, from being distorted (Lin et al. 2005). However, conceiving the sender and receiver as persons, the limitations of this approach become quite obvious. First, when the receiver does not get a message, she still might be informed, if she expected a message (Bertalanffy 1973:41). So she might send a message asking for the reason of the delay or might think that the channel is broken down or might get impatient and move on to do something else and so on and so forth. "*The letter which you do not write can get an angry reply*" (Bateson 2000:459). As already stated above, the question arises as to how it is possible that information – a thing carried to the other end of the channel – still triggers change even when it does

not reach its destination. The, in this case, non-receiver can still interpret the non-signal (see also Krippendorff 2009:282).

Second, the model assumes that the receiver is capable of distinguishing between signal and noise as intended by the sender. If the receiver is only getting noise without being aware of that, she might still be informed as long as the received messages are according to the expectations of the receiver. For instance, the sender flips a fair coin. Sent through a noisy channel, the receiver may get the false message that the result was heads instead of tails. The receiver, however, was still informed since the message makes total sense; it is one of the two expected outcomes. Of course, the receiver might be very careful and send the message back for confirmation. However, that message could also be distorted. The initial sender may turn out to be very careful as well, so the sender may send the receiver's message back for another confirmation. In the end, none of the two can be certain whether the true value of the message was carried through the channel. This scenario illustrates a very important implication in discussing information. It does not really matter what the message is about. As long as it makes sense to the receiver, she is informed. It is the receiver who distinguishes between signal and noise based on her expectations – referred to as structural determinacy further above – which is different from the sender's intentions. Consequently, there is no truth in information, no right or wrong, because it ends up with the receiver of the signal to be informed. Therefore, the notion of disinformation as a way of deliberately manipulating somebody by diffusing biased messages, turns out to be invalid, since it assumes that there is true and false information (Fetzer 2004; Floridi 2004).

Third, distinguishing a sender from a receiver is a second-order observation. Actually, a sender is also a receiver and a receiver is also a sender. Of course, one can define the person sending the message as a sender but that is only of limited help, because then one needs to ask how the sender knows what to send. Hence, the question arises how the communication channel and the rules of encoding were established in the first place. In terms of signal transmission, that is designed by an engineer. But what about, say, spoken language? Clearly, whenever something is being said it is expected to be understood by a more or

less specific person or audience, which brings back the social dimension already introduced above – Alter-Ego is the observation of another Ego (the distinction of Ego/Alter-Ego), another person expected to understand. However, if this argument is followed to the end, a paradox emerges, because Alter-Ego observes back and therefore also expects to be understood, which leads to a loop of infinite regress; i.e. Ego expects that Alter-Ego expects that Ego expects that Alter-Ego expects and so on and so forth. This situation is the very basic problem of any sociology; that is, how is this gap between two Egos bridged in order for something to emerge one may call social order (Littlejohn 1992)?

3.4.2 The Sociological Paradox of Double Contingency

The paradoxical situation of mutual expectations was analysed by Parsons (1968) and Luhmann (1996b) as a situation of *double contingency*. Ego and Alter-Ego observe the other as another Ego as well as the self as being observed. The circularity of mutual observations renders any action indeterminable (Vanderstraeten 2002) and results in a paralysis of anything and, therefore, nothing goes. From the technical perspective of signal transmission the paradox is translated into a technical problem solved according to rules of optimization and efficiency. This is only possible, because the engineer is taking an outside position in relation to the communication system, hence being able to analyse the functionality and efficiency of the design. This notion is mostly ignored when the input/output model of information theory is discussed; the person who draws and builds the coder/decoder and the channel between them is not part of the set-up. In terms of the social dimension of communication, one faces a paradox, which is, per definition, unsolvable. There is no outside view – there is no designer providing lists of possible messages and tools to prevent distortion (Bateson 2000:291). In other words, the "information value" of double contingency, as it is the case with any paradox, is infinite (Baecker 1996:66; Krippendorff 2009:283) because one side of a paradoxical form indicates the other side

resulting in the oscillation of the observation between both sides of the form without a final resolution.[12]

Still, based on the three points raised above, an autopoietic reinterpretation of Shannon's model can further specify what is meant by information as a difference which makes a difference (Bateson 2000:271). First of all, information can be described in relation to a form of redundancy/variety (Ashby 1958; Esposito 2002:264; Baecker 2006a:136). This can be seen, again, as a paradoxical form – redundancy can be observed *because* there is variety and vice versa. After all, how can an observer identify sameness if she is not capable of observing difference as well? Redundancy and variety are two sides of the same coin. However, this is not an attribute of the message itself but rather of an observation performed. For instance, when Borgmann (1999:133) states that, according to information theory, to be told the sun will rise tomorrow is to receive no information at all because it is a totally redundant message, is only valid when redundancy/variety is seen to be carried by the message itself. From the perspective of a structurally determined observer, there is no way of telling how a supposedly receiver may react. She may ask why somebody states something as obvious or interpret it as a message with a hidden meaning. In general terms, even stating that something is totally uninformative is again information (Esposito 2002:264) – anything can be informative, even the statement that the sun will rise tomorrow (Bowker 2005). If anything can be informative than nothing is; information, therefore, is per se unlikely and needs to be seen as problematic rather than as a solution or answer to a question (Wildavsky 1983).

Another possible approach is to observe according to the difference of expectation/surprise. As it was already stated above, information is linked to novelty or surprise. Again, this can be seen as a form, since how can somebody be surprised if nothing is expected. In this sense, expectations are just another expression for the structural determinacy of observation. It is the structural

[12] A telling example is the so-called liar paradox expressed in statements such as: "Believe me, I am a liar!" or "This sentence is false!" Accepting the statement to be true indicates that it is a lie and vice versa. In system theoretical parlour, observing one side of the distinction (e.g. this statement is true) indicates the other side (e.g. this statement is false).

determinacy that allows for the occurrence of information in the first place. In the terminology of autopoiesis, a system is structurally coupled to its environment, which delimits the possibilities for irritation (for being informed) to a manageable degree of potentialities – the observer's (life)world (Maturana and Varela 1992; Luhmann 1996b). In other words, one is surprised because one expects to be surprised – again a paradox. In a situation of double contingency - the paradoxical situation of mutual expectations – the variety of, basically, limitless possibilities is delimited to a manageable degree of potentialities. The expectations are structured, which Luhmann (1998) calls a social system. Now, a social system does not solve the paradox of double contingency, it rather hides and covers it. Consequently, communication – being a very unlikely event due to double contingency – is transformed through the social system into a very likely event because one cannot say whatever one wants if one expects to being understood.

The conclusions drawn from the critique of Shannon's theory in reference to communication are going to be very helpful in the later steps of the analysis in terms of communication technologies and memory institutions. First of all, mutual information is not to be understood as a synchronization of views, opinions or individual life-worlds. Communication is not about maximizing understanding between the involved communicative actors (Littlejohn 1992; Schmidt 2011). For instance in situation theory (Devlin 1999), communication is seen as the establishment and maintenance of a common ground that is increased when new information is added. In a sense, the more one talks, the better people understand each other. In other words, communication is conceptualized as a process with an end-point when all involved parties finally understand one another. However, the question arises again as to how one "knows" one is being understood?

Understanding as well needs to be conceptualized as a communicative act. As already pointed out in terms of information theory; just as a message can be misunderstood, the confirmation that the message is being understood can be misunderstood. There is no absolute proof for mutual understanding. Communicating about communication is communication. Or in other words, the

information that one has been informed is accomplished by an observer. Consequently, if one sees communication as a problem of achieving mutual understanding, one has to face the question of "how do we know that we are being understood"? However, if one sees communication as an operation emerging out of the paradox of double contingency, the question of mutual understanding becomes irrelevant. Communication occurs *because* we do *not* and never will understand each other (Luhmann 1996b). The confirmation of the received message sent back to the receiver is prone to the same disturbances and uncertainties that made the sending of a confirmation necessary in the first place. In the end, one will never know whether the claim of having understood the message is valid or not.

Hence, sociality is bound to communication and to communication only when dealing with the social dimension of information because there is no outsider having an absolute view on society (Luhmann 1996b). That is not an engineering problem but a sociological paradox, which lies at the heart of Luhmann's thinking (Angell 1990). The paradox is the situation of double contingency with an infinite information value - anything and nothing can be informative (Krippendorff 2009). It is a situation very sensitive to random events; something or nothing happens. As a reminder, nothing can also be informative. In other words, due to the indeterminacy of double contingency, literally anything can be observed as informative.

> "*Everything that happens in such a situation, every action, every gesture, every expression, appears as a relevant, meaningful selection. After the first gesture, every subsequent step becomes an action with a contingency reducing effect – be it positive or negative*" (Vanderstraeten 2002:87).

It is a situation without a social memory. Over time, the unlikelihood of communication due to pure chance turns into structured potentialities (Luhmann 1996b:148-90). Consequently, the relationship between Ego and Alter-Ego emerges into a structured order that goes beyond their single contributions. Communication emerges without the need for a designer.

3.4.3 Social Systems as Communicative Systems

As already stated above, the emerging structure organizes the mutual expectations and, therefore, delimits the initially unlimited informativeness of double contingency into delimited potentialities such as, for instance, social roles, which are nothing else than packages of expectations (Dahrendorf 1972; Berger and Luckmann 2003). Communication is the enabler for social order but also its outcome – a self-referential process (Vanderstraeten 2002). Thus conceived, communication cannot be reduced to either Ego or Alter-Ego, since none of the two have the capacity to control the indeterminacy of double contingency. That would mean that one of them is the sole designer – the engineer (Angell 1990). Instead, a social action centred sociology that inevitably leads to a social actor being the foundational element of society (Weber 1976; Giddens 1984) is replaced by a mutualistic perspective of communication as the principle operation of society (Luhmann 1996b; 1998). Therefore, the focus is not on either sender or receiver (on actors) as sources of communication but rather on the emergent and autonomous structure between them – what Luhmann calls a social system. Consequently, Ego and Alter-Ego are not defined by and through themselves but by the social system as communicative partners or persons. It is communication that structures expectations and therefore ascribes social roles. In other words, a social system regulates communicative surprises and thus information. This is a radical anti-anthropocentric approach; it is not human beings that communicate but rather communication that communicates. In light of this argumentation, human beings or rather biological and mental systems are conceived to be in the environment of social systems. Structurally coupled, they can only irritate one another.

In detail, Luhmann (1996b; 1998) incorporates Bateson's (2000) definition of information into his concept of communication as an operation of utterance – information – (mis)understanding. A social system based on the paradox of double contingency does not provide a list of discrete items from which Ego and Alter-Ego select but rather reduces infinite "informativeness" to a manageable degree of complexity by distinguishing between what can and cannot be expected from the self and the other. One person says something, which is not

yet a message (i.e. a signal with a meaning) but rather an utterance. The utterance is a selection amongst the possibilities that is the social system. However, it is not Ego who makes the selection in terms of communication because that selection is based on the expectation to be understood – it is based on what Ego expects that Alter-Ego expects what Ego expects and so forth (Luhmann 1996b; Luhmann and Baecker 2002). In order for Ego to be observed as an Ego (i.e. as a communicative person) she cannot say whatever she wants – communication is "mutualistic". In other words, Luhmann describes the selection of an utterance as a difference between potential and actual utterance (the first difference of Bateson's definition).

Alter-Ego observes the utterance as an utterance and distinguishes between information and utterance – between what is being said and how it is being said. That is (mis)understanding – the difference between information and utterance (the second difference of Bateson's definition) – which is again communicative. If utterance cannot be distinguished from information – the how from the what – it is mere noise. From a social perspective, therefore, it is irrelevant what happens in the head of Alter-Ego. (Mis)understanding is again an utterance that feeds back into the social system. Most importantly, communication only occurs when and only when (mis)understanding occurs. The difference which makes a difference makes another difference and so forth. However, the distinction between information and noise (i.e. which utterance is observed as informative and which is not) is accomplished by the social system. Of course, one can say whatever one wants but if one is not understood it is not communication. The difference does not make a difference.

To come to a conclusion, Shannon's engineering model can be reinterpreted based on a paradoxical take on communication. An utterance is a selection out of a limited set of possibilities. However, there is no menu-type list containing discrete and mutually exclusive messages when it comes to what Bateson (2000:291) calls analogue or iconic communication. In other terms, communication (not signal transmission) or rather the social system distinguishes between informative and non-informative utterances. Consequently, a social system can be seen as an in-formation – as a differance (Derrida 1982; Baecker

1999:60); a pre-selection of potentialities not resulting into a list but rather a continuous horizon. In system theoretical terms, the social system reduces complexity, hence, turning the unlikelihood of communication into a very likely occurrence (Luhmann 1996b).

3.5 Memory/Information

Memory as the faculty of systems, in general, and social systems, in particular, has already been discussed within the context of the research domain of cultural and social memory studies (see Chapter 2.2.3). This chapter will discuss memory within the framework of social systems theory with a particular focus on its relationship to information.

Following the cybernetic tradition, Heinz von Foerster defined memory as a conceptual construction used by an observer to describe a system from the outside (Ashby 1957; Foerster 2003). Although heavily influenced by von Foerster's work, Luhmann had a contrasting view. Based on the notion of re-entry of the system/environment difference into the system, an autopoietic system is capable of observing itself and therefore to identify itself based on an operation that can be called memory (Luhmann 1998:578-60). In more detail, autopoiesis as the self-reproduction of the system requires systemic operations (elements of the system) to connect with previous operations in order for a systemic structure to emerge as the result of these operations (Luhmann 1986; Luhmann 1996b). This principle was already expressed further above with reference to information as a difference which makes a difference, which, in turn, makes another difference. In social systems parlour, communication connects with previous communication and, therefore, sets the stage for communication to follow. The structure of potential linkages can be described as social expectations, which delimit the number of potential linkages in order for the system not to be flooded by complexity (Luhmann 1996b:62). It is in this sense that systems, only capable to operate in the present, require to recur to the past and thus prepare for the future (Dimbath 2011:141). Memory, therefore, is defined as the continuous operation of observing before/after. To be precise,

memory introduces time into the system and, thus, allows it to construct its own past and future. The past, however, is not fixed but is rather constantly modified "*to connect it with a possible future in the present"* (Luhmann 1997:365). Memory does not only recall the past but also calls the future.

Memory enables the system to recognize recurring operations as repetitions; that is, singular events, which are never completely alike, are observed to be the same nevertheless (Schmidt 2008). For instance, it is memory that allows a person to recognize herself as the same person even when her physical characteristics have changed considerably over a lifetime. Memory constructs sameness into difference. This is accomplished by means of distinguishing between forgetting and remembering. In other words, memory is a form of the difference between forgetting and remembering (Esposito 2002). In particular, memory filters observed events as singularities by distinguishing what is remarkable and ignoring the rest. As such, it executes an important operation since it compares what is already known with what is unknown.

If every event was observed as unique by the system (including itself), nothing would be new or surprising. In other words, one needs to expect in order to be surprised. Thus conceived, memory results in redundancy; i.e. the observation of repetition. Information, on the other hand, can only occur if it makes a difference to what is expected. In the end, the concept can be reformulated as a form differentiating between redundancy and variety or rather memory/information (Esposito 2002:25). Since there is an infinite number of singularities to be observed, the main focus of memory is on forgetting. Filtering what makes singularities unique and, thus, making them comparable according to abstract parameters, makes remembering, on the other hand, into an exception. Given these arguments, information only occurs in comparison to what is remembered as well as forgotten. Memory, therefore, is not a stock of recorded events but rather the organization of observing information. Hence, memory is necessary to recognize information and information is necessary for the construction of repetition (Schmidt 2008), since without memory everything and nothing is new and surprising. In order for an event to be observed as novel, it needs to be distinguished from what is repeated (Luhmann 1998:576-94).

Conceived as an active operation of predominantly forgetting, the theory leads to a very different notion of memory in contrast to being a storage of information (Foerster 1967; Riegler 2005). Indeed, the storage metaphor of storing and retrieving information brings the argument back to Shannon's information theory (Shannon 1993). Conceived as a model of coding, transmission and decoding, there are obvious parallels to a model of memory based on storage, retention and retrieval of information (Bussola 2011). In fact, the only difference between these two approaches is that with memory sender and receiver are one and the same. Remembering, therefore, is seen as the successful encoding of information into a record (e.g. neural engrams, mental imagos, cultural artefacts, binary-based data) followed by its decoding (Choi et al. 2010; Wagner Cook et al. 2010). By the same token, distorted remembering is a result of noise while forgetting is the result of unsuccessful encoding or decoding. Both are treated as errors that need to be minimized (Connerton 2008; Hunt et al. 2011). As a consequence, remembering is equated with not-forgetting and forgetting with not-remembering (Singer and Conway 2009).

By contrast, an approach based on autopoietic systems leads to the conceptualization of remembering as a precursor for forgetting and vice versa. Both are necessary for memory to operate. Admittedly, the storage metaphor could be interpreted as having a filtering mechanism by means of the rules of encoding and, thus, as an operation of forgetting. Autopoiesis, however, requires the system to deal with the paradox of re-entry (Bussola 2011). The autopoietic system does not only operate based on a system/environment distinction but also observes itself within itself as a system/environment distinction. In other words, the system represents itself within itself, which relies on memory in order for a system to remember (i.e. re-cognize) itself by forgetting (i.e. filtering) the changes it has gone through over time. It re-cognizes itself as a system with a history. The storage metaphor, on the other hand, is a-temporal since remembering is seen as the retrieval of unchanged records no matter when. This is not meant to be that the storage metaphor is wrong per se but rather it is not viable anymore for the unfolding of the paradox of re-entry and self-recognition. Indeed, the autopoietic system may observes its own memory *as if* it was a storage of recorded events (Dimbath 2011). In light of the empirical case study

discussed further below, however, it is one of the aims of the dissertation to argue for a new metaphor of computation and transversality according to which the paradox of memory can be unfolded.

To conclude, memory can be seen as an ongoing operation that constructs sameness into difference and, thus, redundancy, which in turn allows for the observation of variety, novelty or, in general terms, information. Thus conceived, memory and information can be seen as two sides of the same coin. The main focus of memory is on forgetting, which is defined as the filtering of singularities according to increasingly refined processes of abstraction and categorizations. Remembering, on the other hand, is not a retrieval mechanism of stored records but rather the re-activation of processes that have been activated before (Esposito 2008). In general terms, a system can forget more by remembering more and vice versa. The more refined the processes and categorizations (remembering), the more singularities can be filtered and, counter to intuition, the less needs to be stored. To use computation as an illustration, the more elaborate the algorithm, the less data needs to be actually stored on a harddrive.

At this point, the dissertation reconnects with the paradigm shift of memory from "memory institutions" to "information habitat" discussed in Chapter 2.3.3. Given what has been discussed so far, the paradigm shift can be conceptualized as the emergence of a more refined process of social remembering and forgetting. With the rise of information technology, in general, and algorithmic search engines, in particular, it may be more appropriate to use the metaphor of computation rather than storage to conceptualize contemporary developments of social memory (Esposito 2002; Baecker 2006b). In other words, the shift from one paradigm to another can, indeed, be seen as a shift in the ways the paradox of memory is unfolded rather than as the destruction of memory (Assmann 2006), the forgetting of forgetting (Mayer-Schönberger 2009) or the forgetting of remembering (Brindley 2009). Hence, the chapters on research domain and theory converge at this point and will inform the research question, the case study and, ultimately, the discussion of the findings.

3.6 Epistemological Implications

The operational closure of autopoietic systems as well as the re-entry of an observation into itself leads to a peculiar view on reality and the knowledge of reality. If observation is structurally determined by self-reference to previous observations, reality as a given becomes a questionable category. Thus conceived, the theory of autopoiesis joins a school of thought that can be summarized under the term of *radical constructivism* (Knorr-Cetina 1989). The term originates from the work of Ernst von Glasersfeld (1997; Müller 2011) but is used in this dissertation to refer to equivalent epistemological considerations found in psychology and psychotherapy (e.g. Watzlawick 1977), the study of cognition (e.g. Piaget 1970; Bateson 2000) and the already mentioned theories of general systems theory as open systems (e.g. Bertalanffy 1973:205), second-order cybernetics (Foerster 1993; 2003) and autopoietic systems as conceptualized by Maturana and Varela (Maturana 1988; Maturana and Varela 1992) as well as Niklas Luhmann (1998; 2006). Although coming from quite diverse backgrounds, these and other authors recognized the commonalities between their respective research endeavours and heavily influenced each others' work (Glasersfeld 1991; Brier 2008; Le Moigne 2011; Schmidt 2011).

At the core of this school of thought lies the proposal that the distinction between ontology and epistemology is an epistemological one (Brier 2008). In this sense, radical constructivism can be distinguished from social constructivism (Berger and Luckmann 2003) as well as critical realism (Archer 1995; Bhaskar 1998), which are based on the distinction between a social reality constructed through institutionalization, socialization etc. and a physical reality ruled by natural forces independent of human observation (Sismondo 1993). By contrast, radical constructivism denies the existence of an ontological reality as a given but rather suggests that what one calls reality is an accomplishment of systemic operations, i.e. observations (Boden 2011). Reality is seen as a *viable* model rather than objective truth. Thus conceived, the central tenet can be formulated that *one cannot separate reality from one's knowledge of reality* (Glasersfeld 1997).

Observation was defined as an operation of distinction and indication (marking) that, when it comes to autopoietic systems, always distinguishes itself from its environment and either indicates itself (self-observation) or its environment (second-order observation) within itself. Therefore, the ontological question of "what is?" is radically de-ontologized (hence the term *radical* constructivism) and declared invalid since an observer is not able to distinguish reality from the observation of reality – reality is inseparable from what is known about it (Maturana and Varela 1992; Luhmann 1996a:16-17; Glasersfeld 1997; 1997:362; Brier 2009). "*Radical constructivism simply claims that knowledge in fact lies in the realm of understanding – relational rather than absolute*" (Confrey 2011:279). Radical constructivism, however, is not to be understood as a solipsistic or nihilistic attitude (Martinez-Delgado 2002), for one "[does] *not deny that a 'universal' is out there; however, it asserts that any universal (if there is one) is an unknown and unknowable terrain*" (Angell and Demetis 2010:62). Constructivism refers to a world outside of an observer out of which reality is constantly constructed.

In order to construct, one needs, for the lack of a better word, "something" to construct with. This is also the reason why Maturana (Maturana et al. 2011) chose the term poiesis in order to stress the notion of production (out of something), which stands in opposition to creationism in the sense that an observer creates a reality out of nothing. In this sense, constructivism as a term also refers to reality being *con*structive rather than *in*structive. However, an observer is always bound to her/his/its capabilities to observe (Artandi 1973). The world is only as deep as one's eyes can see. On the other hand, for an observer to be able to observe the world, one needs to be a delimited *part of* the world (Heidegger 1996; Foerster 2003) or - as Maturana and Varela (1992) would put it – a system is structurally coupled to its environment. In different terms, the observer is the difference between system and environment (Luhmann 1996b).

These considerations also have an impact on the way science and scientific knowledge is conceptualized by radical constructivism (Moser 2004a). Transforming reality as a given into a viable model constructed by an observer

refers to her/his/its structural capabilities. Consequently, the information science collects about reality depends on the instruments and measurements science uses (Luhmann 2002b). In more precise terms, information is not an attribute of the world but rather an accomplishment of the discriminatory categories applied (Kallinikos 2006b:103). Thus conceived, radical constructivists are also sceptical towards the scientific endeavour of uncovering absolute truth or objective knowledge, which is reflected in the way science conceptualizes information scientifically. On the one hand, information is defined in the General Definition of Information (GDI) as "data + meaning" (Checkland and Scholes 1990; Devlin 1999; Braf 2001; Floridi 2009). Depending on the ontological stance, meaning may be more or less independent from an observer. To be more specific, data can have semantics independent from their user – data can carry some meaning, which in turn can be further interpreted or "semanticized" (Liebenau and Backhouse 1990; Floridi 2009; Lanzara 2009). According to a stronger realist approach, data can have their own semantics even independent from an intelligent producer; that is, data can carry meaning not only independent of an informee but also an informant (Dretske 1981).

GDI relies on the assumption that there is a structured reality that may be beyond science's (current) capabilities of discovery and research. Information as an ontological concept, therefore, assumes that data can be interpreted in a right or wrong way. As a consequence, truth is in reality and scientific statements about reality may diverge to a more or lesser degree from reality. Even if information is based on pure proto-epistemic data that exists beyond the horizon of human cognition, it is there since it can be inferred from experience (Floridi 2009). Informativeness is (more or less exclusively) an intrinsic attribute of reality. Of course, truth needs to be absolute since there needs to be an absolute point of origin statements may be closer or further away from.

> "*The first step* […] *is to define the concept of "informative content" or intrinsic informativeness of* [a statement] *extensionally and a priori in an ideal context, as a function of the positive or negative degree of "semantic distance" of deviation of* [the statement] *from a fixed point or origin, represented by the given situation w, to which* [the statement] *is supposed to refer*" (Floridi 2004:205).

The closer a statement gets to reality the truer it is and therefore more informative. In other words, the transcendental existence of the first and final referent and referee, which used to be called God, nature, matter, energy, mind, consciousness and so forth (Maturana 1988), receives a new name – information. This is the domain of transcendental metaphysics, logic, abstract mathematics, theology and mysticism.

The second approach is based on Shannon's (1993) model of signal transmission defining information as data communicated. Information is a measure not for what has been said but what else could have been said as well. The central question, then, revolves around the transmission of signals through a communication channel in an as noiseless fashion as possible. Indeed, the model of signal transmission can be seen as a reflection of an epistemological concept of information. Rather than finding or uncovering information and truth, one communicates with reality (Floridi 2003). Knowledge about reality is gained by means of finding the appropriate decoder for the messages sent by reality and the information contained therein. In the form of hypotheses, science asks questions, encodes them and expects reality to answer. The measurement of data, in turn, is the decoding of the message translated into signs and symbols in need of interpretation and "semanticization."

In order for the communication to be undisturbed by perturbations, the communication channel is ideally to be without noise. The laboratory experiment and its setup of a controlled and sterile environment is assumed to come close enough to such ideal conditions as the variables to be measured are separated from the undesired influences of other, noisy variables (Maturana 1988). In this sense, information is not an intrinsic attribute of reality but rather of communication with reality requiring a receiver for "*information cannot exist independently of the receiving person who gives it meaning and somehow acts upon it*" (Liebenau and Backhouse 1990:3). Information, therefore, is akin to knowledge, not to truth. This is the domain of the natural sciences and positivist social sciences.

A radical constructivist approach abolishes this distinction between information as truth and information as knowledge by introducing the notion of observation and the observation of observation (Maturana and Varela 1992; Glasersfeld 1997; Luhmann 1997; Foerster 2003; Brier 2009). Hence, what one knows about reality cannot be separated from reality itself. Information as information about reality and truth, therefore, is a statement made by an observer making the observation that there is an absolute point of origin. However, one has to ask where this observer stands and how she/he/it knows truth. Now, one could remain agnostic towards the existence of absolutes and agree that truth lies in the eye of the beholder. In other words, there is data that is being interpreted by an observer and made semantically meaningful (Liebenau and Backhouse 1990; Braf 2001) based on an ontic reality beyond our cognitive capabilities, which can only be known indirectly and may lead to different interpretations. However, that is, again, based on descriptions and statements made by an observer.

To conclude, radical constructivism argues that the search for an absolute foundation or beginning is a futile endeavour trying to hide the self-referentiality of observations as making a distinction not only between what is observed and what is not but, more importantly, between the observer and the observed that is achieved either by the same observer or a second observer (Luhmann 1996b; 2002a). In the end, one faces a circularity that would require the observer to step out of that circle in order to be able to pinpoint a point of origin or a beginning. Or as Ernst von Glasersfeld puts it;

> "[…] *in thick fog on an Alpine glacier, one places one foot in front of the other without seeing what lies further ahead or further behind one; and as sometimes happens in such a fog, after hours of walking, one realizes that one is walking in one's own footsteps. The fact that one has begun the circle at a specific place could be perceived only from a higher vantage point, if the fog had lifted and made possible a comprehensive view. But the fog that obstructs our view of ontic reality cannot lift*" (Glasersfeld 1991).

In this sense, knowledge is not corresponding to reality but rather a viable structuring of experience. Reality is constructed based on the observer's capabilities to observe, which already delimits what can make a difference; that is, the differences that can make a difference to the observer are brought forth by

the observer. Information, therefore, occurs "inside" the circle – the observer is informationally closed from the world he/she/it lives in (Glasersfeld 1991; Maturana and Varela 1992). To be more precise, the reality the observer describes is based on the internal operations of the observer. Reality, therefore, is an accomplishment from parts of the observer – a reality that can be viable but never true, since the conditions for truth are again based on the internal operations of the observer.

3.7 Research Question

Given the theoretical considerations discussed in this chapter, the research interest can be further enriched leading to the formulation of a research question. Initially, the starting point of the research project was a research interest that can be summarized as the question of *how libraries change in order to remain libraries*. Based on a review of the relevant literature and the framework developed in Chapter 2, the research interest was further specified as the question of how libraries change in order to remain libraries *as they shift from the paradigm of memory institutions to the paradigm of the information habitat*. In the terminology of social systems theory (Luhmann 1998), the initial research interest addressed a distinction between library and its environment and questioned whether that distinction is observable in the information habitat. Conceived as the formation of a form, the research focused on the in-formation of the librarian domain into the information habitat.

As a second step, the first research question was formulated as *how do libraries distinguish themselves from the information habitat as libraries*. Thus conceived, the question directs the research away from studying the essence of libraries towards the difference that makes a library observable as a library in the information habitat. By the same token, the question also defines the librarian domain as the point of reference for the observation. The research is not about whether the environment (e.g. library users) observes a library as a library but rather how the librarian domain observes itself in difference to the information habitat. It is in this sense that the research project is a second-order observation;

i.e. it is the observation of how the librarian domain observes its difference to its environment (see for instance Chapter 3.2). A crucial point since the selection of the point of reference has an impact on the methodological choices available for the project. Selecting the environment as the point of reference would suggest, for instance, a survey among library users as an appropriate research strategy. However, an initial inquiry into the field of so-called digital libraries, conducted at the beginning of the research project, revealed that there are very few comprehensive librarian initiatives with the goal of bringing a critical mass of cultural heritage artefacts into the online world and, thus, to make the artefacts themselves accessible in a digital format (see Chapter 2.4.4). Indeed, the status-quo was to provide only an online catalogue (OPAC), which, as it was argued in Chapter 2.4.2, is still a service firmly embedded in the memory paradigm of memory institutions.

The situation found in the field needed to be taken into account, which, ultimately, resulted in the selection of the librarian point of view as the point of reference. Given the practical limitations of the research project with respect to its projected timeframe and available resources, the European Digital Library project (EDL), as it was called at that time, presented itself as the only promising initiative with the potential to reveal unknown and unexplored aspects of the digitization of cultural heritage artefacts and respective memory practices (see Chapter 4.2). Hence, the research question was further specified as: "H*ow does the European Digital Library distinguish itself from the information habitat as a library?"* In order to evaluate the validity of the assumption that libraries are indeed changing in order to remain libraries as well as the viability of the research project on EDL, an exploratory pilot study was conducted on the Austrian National Library (ANL) – an active member of the EDL project. The results of the pilot study (Marton 2007) confirmed the assumptions and, as a consequence, supported the research question.

During the course of the empirical field work, however, the research focus on libraries turned out to be too narrow in order the capture the studied phenomenon. Indeed, it became increasingly difficult to maintain the category of a library in contrast to museums and archives in the domain of digitized cultural

heritage artefacts. As it will be argued in detail in Chapter 7.2, it is more appropriate to address the broader field of memory institutions, since the distinguishing characteristics between libraries, archives and museums are on the verge of becoming irrelevant. By the same token, the EDL was renamed into Europeana reflecting the goal to provide for a unified framework of libraries, archives and museums. Hence, the research question required adjustment and, finally, was formulated as;

> *How does Europeana distinguish itself from the information habitat as a memory institution?*

Given the definition of memory institutions as institutions of forgetting through the process of canonization (selection, order, preservation) of cultural artefacts into cultural *heritage* artefacts (see Chapter 2.4.1), the research question can be broken down into following sub-questions to be answered in Chapters 6 and 7;

1. *How does Europeana select?*
2. *How does Europeana order?*
3. *How does Europeana preserve?*

The research question plays a pivotal role in any research project as it links the synthesis of the research domain and the theoretical framework with the empirical study (Silverman 2005). Based on the epistemological stance taken in this dissertation, the question plays an even more important role since it does not address an existing phenomenon. It is rather instrumental in the construction of the "object" of study by means of distinguishing between what is relevant and irrelevant for the research project and, therefore, reflects the expectations of the researcher (Esposito 1996). It is against this backdrop of expectations that unknown and surprising events can be observed (Weick 1989; Gaskell and Bauer 2000). As it will be discussed in Chapter 4 on methodology, the research strategy of choice – the case study - will be constructed following the research question and the distinction of relevance/irrelevance the question introduces. If data collected is relevant for answering the research question, it is part of the case. If not, it is part of the context. Thus conceived, the distinction between case and context is not found in the field but rather reflects the distinction between relevance and irrelevance as observed by an observer. The techniques and

procedures applied in order to translate the research question into a research design and empirical study will be discussed in detail in Chapter 4.

3.8 Conclusion

Contemporary society is different from the modern era of industrialization and nationalization. Most notably termed "information society" (Webster 1995; Webster and Puoskari 2003), the differences have been analysed from the perspectives of economy (e.g. Shapiro and Varian 1999), occupation and production (e.g. Bell 1974; Castells 1996), space and time (e.g. Beck 1986; Giddens 1987; Hanseth and Braa 2000), culture (e.g. Borgmann 1992; Baudrillard 1994; Crang et al. 1999) and information technologies (e.g. Masuda 2003). More cautious voices (Winner 1996; Brown and Duguid 2000; Webster 2005) warn that conceptualizing contemporary developments as an emergence of a radically new kind of society is an unfounded leap (Webster 1995:219). Instead of an information society, one should rather think of an increasing "informatization" of society that has been going on for centuries before the invention of the computer (Beniger 1986). Hence, concepts should not claim a total break with the past but rather a radicalization of already existing societal forms and operations – not a post-modernity but rather a hyper-modernity (Esposito 2002) or reflexive modernity (Beck et al. 1994).

The key historical aspect that led to an increasing importance of information is the societal shift from absolute, undisputable knowledge to criticisable claims. Due to the secularization process from Renaissance times on, claims about the nature of the world have been based on human reason rather than faith guided by heavenly forces. The rationalization of human affairs (Weber 1976) does not allow for communication with a god (Habermas 1987:6; Horkheimer and Adorno 1988; Floridi 2003). Basically, European civilization got rid of the last and absolute observer – the ultimate cause (Heidegger 1969). One is reminded of Nietzsche's (1887:154) famous aphorism: "*Gott ist todt! Gott bleibt todt! Und wir haben ihn getödtet!*" - not only is God dead, but it is us who killed him. Hence, knowledge is prone to human limitations and, therefore, unreliable

(Habermas 1987:25) – knowledge has become changeable, the (life)world contingent. This is not to be seen as knowledge having turned into information but rather as the societal shift of focus from the repetition of tradition to the discovery of the unknown, from the past to the future (Nora 1989; Esposito 2002) – a shift very well exemplified by science and technology (Krohn 1989; Marton 2009).

Bereft of absolutes, every statement, fact, description or expression is based on claims of validity criticisable in terms of being true/false, right/wrong, effective/ineffective, efficient/inefficient, aesthetic/unaesthetic, appropriate/inappropriate and so forth (Habermas 1987). By having neither a metaphysical nor theological reference point (Heidegger 1969; Floridi 2003), the search for absolute knowledge is a futile endeavour simply because there is nobody to make the final call – no outside, absolute referent and referee (Angell and Demetis 2010). As a consequence, what is old and known can be seen as being new and improvable. Knowledge has become a historical and contingent phenomenon (Luhmann 2002b:Chapter 3). In light of this argument, the dissertation introduced an equivalent epistemological concept to truth. Following the work of Ernst von Glasersfeld (1997), reality was defined as a viable model (Lanzara 1991:312). Therefore, the ontological question of "what is?" is radically de-ontologized and declared invalid since an observer is not able to distinguish reality from the observation of reality – reality is inseparable from what one knows about it (Maturana and Varela 1992; Luhmann 1996a:16-17; Brier 2009). The world is only as deep as one's eyes can see.

This notion is reflected in the informatization of society (Zuboff 1988; Webster 1995; Kallinikos 2006a). As the discovery of the observer being part of the observed - a notion now also found in the field of quantum-physics (Zeilinger and Griese 2006) - it is the socio-cultural shift from knowing to informing, from the paradigm of control (Beniger 1986) to the paradigm of uncertainty (Esposito 2002; Kallinikos 2006a). Information becomes ubiquitous (Baecker 2006a:119). Thus conceived, the rise of memory studies and the so-called "memory craze" can be conceptualized as the other side of the informatization of society and the shift towards uncertainty. For one, the observation that historiography is just one

way of observing and, therefore, constructing the past in the present seems to have been radicalized according to the argument that the observer is just another observation (Klein 2000). In this case, the observer is the historian caught in the self-referentiality of history being a historic and contingent phenomenon itself, which blurs the distinction to the lay person, collective or community and their practices of remembering. Everybody is her own historian (Nora 1989; Poole 2008). In the end, the distinction between history and memory is based on disciplines rather than on epistemological privilege (Olick and Robbins 1998).

Second, the crisis of identity, frequently analysed by the community of social memory studies, is directly related to the absence of an absolute point of observation transforming identity into a construction of the self bereft of any metaphysical or theological foundations (Heidegger 1969). As Megill (1998) points out, Nietzsche's famous aphorism about us having killed God is followed by anxiety:

> "*Wie trösten wir uns, die Mörder aller Mörder?* […] *Welche Sühnfeiern, welche heiligen Spiele werden wir erfinden müssen? Ist nicht die Grösse dieser Tat zu gross für uns?*" (Nietzsche 1887:154).

How shall we comfort ourselves, the murderers of all murderers? What festivities of atonement, what holy games will we need to invent? Isn't the magnitude of this deed too great for us? In the end, humanity is left as the sole observer of its own condition fashioning its own identities based on constructed, contingent and contested memories as the only stabilizing factor (Megill 1998). The present is not a consequence of the past; it is the past that is a consequence of the present.

Lastly, Nora (1989) links the rapid acceleration of everyday life to the anxiety of losing the past, hence, the obsessive archiving of its traces. Taken a step further, the present may be seen as shrinking until it remains to be only a difference between the past and the future – an event that passes the moment it occurs; ephemeral and non-foundational (Esposito 2002:276). Constructing the past from the present, therefore, can be seen as the difference the present makes on the past. The present itself being a difference makes a difference. Reminded of Bateson's definition of information, the present only appears as information that in itself is not observable. It is only observable *when* it already passed and *if* it

made a difference on the past. The present is only observable as the past and the future is only observable as not yet having passed. Hence, the new, the surprising, the unexpected – i.e. information – only occurs in contrast to what is old, known and expected - memory. There is an intimate and paradoxical relationship between information and memory; the more one orients oneself towards the new, surprising, innovative, that is, "*the more radical the rejection of anything that came before, the greater the dependence on the past*" (Connerton 1990:61). It comes as no surprise that modernity though breaking with tradition and the past[13] was, at the same time, the harbinger of archives, libraries and museums.

The observation that the observer is just another observation is also valid in terms of social memory, which, as a result, has increasingly become open for questioning and contestation leading to a provisional and temporary view on memory (Wagner-Pacifici 1996; Olick and Robbins 1998). The ubiquity of information is coupled to the memory craze observed by the scholars of social memory. As it was argued further above, information occurs because of memory and memory occurs because of information – a difference of information/memory. One is surprised because one expects and one expects because one is surprised (Connerton 1990:24). Within this context, memory takes a central stage because it is bereft of the idea of an absolute truth and unchangeable past. As a consequence, "*memory gives us a signified whose signifiers appear to be so weighty, so tragic - so monumental - that they will never float free*" (Klein 2000:144). At last, the conceptual journey that began with the "memory craze" in Chapter 2.1 can be brought to a full circle. It is not only the future that is uncertain but also the past, which is the reason why memory is such an important phenomenon to study. Caught in the dynamics of information and memory, it may indeed be more appropriate to address contemporary societal developments as a memory/information society.

[13] For instance, the German term for modernity "Neuzeit" means "new-time."

4. METHODOLOGY

4.1 Introduction

The epistemological issues put forward in the previous chapter have a direct impact on methodological conceptualizations (Moser 2004a). Reality as an accomplishment needs to apply to scientific as well as day-to-day observations, since scientific observation does not have a more privileged epistemological point of view than any other observatory perspective. As already discussed in Chapter 3.2, the observation that the observer/observed distinction is just another observation (Esposito 1996) gives rise to a self-referential approach towards science and its methods. The science of science gave rise to a logos of methods as well as to empirical fields of study such as the sociology of scientific knowledge (e.g. Latour and Woolgar 1986; Sismondo 1993; Kuhn 1996; Knorr Cetina 1999). In what follows, scientific research will be discussed as a paradox of expecting the unexpected that is unfolded into a linear process through empirical methods and techniques. In light of this argument, methods do not reveal certain aspects of a given reality but rather are one of the means to construct a scientific reality. In more abstract terms, scientific methods are structural couplings to the environment of science (Maturana and Varela 1992; Luhmann 2002b). Once methods re-enter science, science observes itself and, thus, turns self-referential giving rise to a methodological discourse on how to define science scientifically.

Against this backdrop, social science and its empirical methods can be defined as second-order observations – as observing observers – bringing along its own distinctions and blind spots (Luhmann 2002b). As a consequence, this dissertation does not claim to be an account of an objective reality or intersubjective truth but rather a viable construction based on specific and structurally determined ways of seeing and not seeing a research domain (Weick 1989). In the end, scientific truth is an accomplishment through and by science itself – a conclusion that leads to the re-evaluation of some methodological concepts (Müller 2008). First of all, the question concerning the quality criteria

for (qualitative) social science research is a means through which social science distinguishes between scientific and non-scientific research in a self-referential way. This notion will be discussed in more detail in Chapter 4.6 on quality criteria. As reality is conceived as a viable model, quality criteria perform as a set of standards according to which the viability of scientific research and, thus, its results can be judged and justified.

Second, data collection and analysis is to be seen as data *construction* since the question what does and does not count as data is related to the distinctions an observation makes. Thus conceived, the various research traditions, strategies and methods are not seen as different lenses that magnify various aspects of the same phenomenon but rather as ways of constructing the phenomenon by means of delimiting a research domain, formulating a research question and research strategy, selecting units of data analysis and choosing methods of data collection and analysis. In other words, it is the (measurement) instrument that defines the phenomenon. "*Method is not* […] *a more or less successful set of procedures for reporting on a given reality. Rather it is performative. It helps to produce realities*" (Law 2004:143).

The collected data is the result of the applied classification systems (Bowker 2005). Indeed, data can be deconstructed in a similar way as "the signal" in terms of Shannon's communication model (see Chapter 3.4.1). Just as the recognition of the signal as a signal (and not as noise) is an accomplishment on the part of the observer, so is the collection of data. In other words, in order for data to be collected, the collector needs to already recognize data as data. In the end, one ends up with the same paradox of distinguishing the observer from the observed as another observation (Esposito 1996). Methodology, therefore, becomes a self-referential observation of science on science – a logos of methods with the aim to narrow down the basically infinite horizon of ways of observing the world to a limited set of scientific options (Müller 2008).

At this point, the form of memory/information, developed in the previous chapter, reappears as a methodological concept. Empirical methods and techniques construct sameness into difference and difference into sameness – a

notion that was referred to as memory above (Bowker 2005). Memory is the organization of the observation of information (Esposito 2002). It is in this sense that scientific research can be seen as the oscillation between the known and the unknown – the unfolding of the paradox to expect the unexpected. These considerations, abstract as they are at this point, will have wide ranging and quite practical implications for the empirical research. In more detail, the rationale for selecting units of data analysis (e.g. documents, interviewees), according to which data about an unknown phenomenon was collected, was developed based on known categories - an approach referred to as corpus construction discussed in Chapter 4.3 (Bauer and Aarts 2000). Second, the data was analysed following the procedure of thematic coding (Flick 2006:307-12), which engages a researcher to switch back and forth between the description and the analysis of the data explained in detail in Chapter 4.4. Both approaches reflect the methodological stance taken in this dissertation, that is, the unkown or suprising (information) can only be observed in contrast to what is known or expected (memory).

Given these arguments, the dissertation will still use the conventional terminology of data collection and analysis despite the fact that they are understood as data construction methods. In fact, research based on a constructivist epistemology does not lead to new methods but rather to a change in attitude (Glasersfeld 1997; Pfeffer 2004). Without any ontological foundation to stand upon, the narrative that is going to be unfolded in the following chapters is to be seen as a scientifically viable description (Czarniawska 2004). In other words, the observations described in this dissertation claim to make a difference to what is already known according to the rules and standards of scientific research.

In the following section, the basic rationale of choosing a revelatory single case study based on a qualitative methodology will be introduced. Defined as a cyclical process of, basically, data collection and analysis, qualitative methodology is approached as a scientific way of unfolding the paradox of studying the unknown based on a technique called corpus construction discussed in the second section. The third section explains how the technique of thematic

coding was used for data analysis, followed by ethical considerations and a discussion of the respective quality criteria applied in the study. Finally, the chapter concludes by revisiting the cyclical process of scientific research summarized in Figure 5 below in order to provide for an overview of the steps followed during the project.

4.2 Research Strategy

Case study research is employed in order to investigate "*a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident*" (Yin 2003:13). By the same token, the phenomenon of interest of this dissertation was defined as the informatization of the field of memory institutions as they step into the information habitat. Conceived as a highly interoperating and networked assemblage (see Chapter 2.3.3), the phenomenon is not only difficult to separate from its context but rather impossible, as the assemblage itself is based on the interoperability of its elements, aspects, procedures and operations. This is reflected by a methodology of observation. Defined as distinction and indication, the phenomenon is brought into a form of case/context according to the research question. For the form to be observable, both sides – case and context - are necessary. As it will be discussed in detail in Chapter 4.3.1, the side of the form that is observed, i.e. marked, will be defined as the case according to a criterion referred to as relevance.

A qualitative approach was chosen due to the fact that the topic of the information habitat and its memory operations (see Chapter 2) are not well grounded empirically and require further in-depth research and theory development rather than theory testing (Benbasat et al. 1987; Eisenhardt 1989; George and Bennett 2005). The approach taken in this dissertation is to define qualitative research procedurally rather than substantially connected to a certain school of thought, tradition or epistemology. In other words, qualitative methods differ substantially (Madill et al. 2000; Reicher 2000:5), however, they are similar in terms of procedure and selection rationale. Qualitative research is a

cyclical process of explaining or rather describing unknown or unknowable phenomena in their variety based on a gradually rather than a-priori determined sampling structure – usually referred to as purposive sampling (Glaser and Strauss 1967; Bauer and Aarts 2000; Teddlie and Tashakkori 2003; Silverman 2005; Flick 2006; Corbin and Strauss 2008). This applies irrespective of the epistemological foundation, methodology, format or style of the research project.

A cyclical approach also reflects the feedback loop of output becoming an input used to describe complex systems (see Chapter 3). Scientific research, in this sense, is an autopoietic operation that leads to viable models stabilized into paradigms, which, in turn, lead to new research questions, empirical studies and theoretical conceptualizations (Moser 2004b). Based on the constructivist epistemology discussed in Chapter 3.6, scientific knowledge (methodology and theory) is self-referential as it studies the unknown based on what is already known scientifically (Luhmann 2002b; Pfeffer 2004).

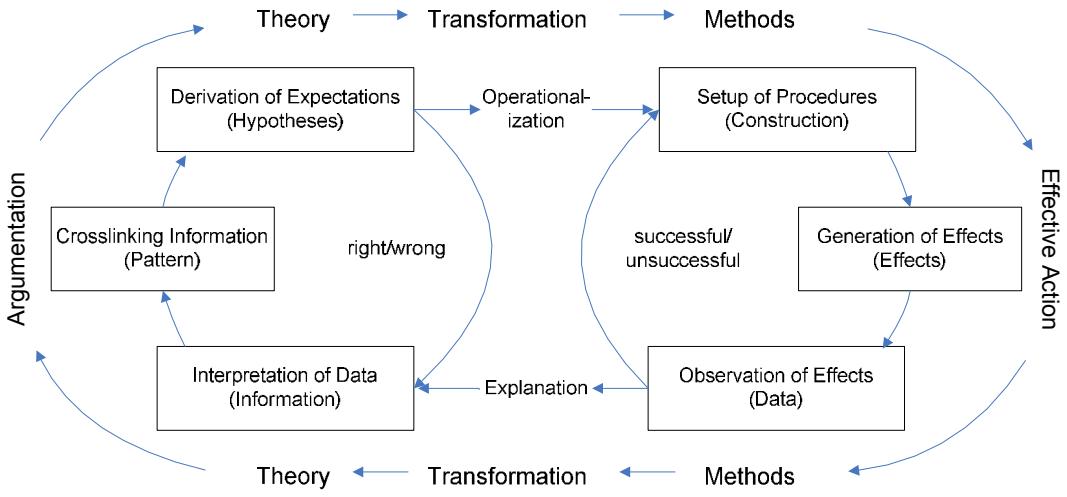


Figure 5: Self-referentiality of scientific research after Krohn and Küppers (1989:58) (translated from German by the author).

The question of methodology, therefore, is to unfold the paradox of studying the unknown into a linear process with a beginning and, more importantly, an end as well as to provide for the criteria and procedures to observe data. Embedded into a case study research strategy, the approach chosen for this research follows the procedures set up by a technique referred to as corpus construction (Bauer and Aarts 2000), while the interpretation of the data is based on the technique of

thematic coding (Flick 2006:307-12). Both issues will be discussed in detail further below.

4.2.1 Research Design and Field Access

In the previous section, the choice for a case study was justified by reference to the interconnectedness of the information habitat. By the same token, the empirically unexplored domain of the memorization of the information habitat suggests the use of a qualitative approach in order to allow for exploration and in-depth analysis leading to theory development rather than theory testing. In detail, the research strategy is a revelatory single case study on the Europeana (www.europeana.eu) initiative launched in July 2007 by the Conference of European National Librarians (CENL), based at the national library of the Netherlands (Koninklijke Bibliotheek) and financed by the European Commission and EU member states. At first called the European Digital Library (EDL), Europeana aims at unifying the digitization efforts of libraries, archives and museums from all over Europe under a single framework by providing a single point of entry to European cultural heritage. After the launch of a prototype online service as a proof of concept in November/December 2008, the project moved on to create an operational service including multilinguality features and browsing capabilities based on semantic web technologies. A more detailed case description can be found in the next chapter.

In order to approach the paradoxical relationship of researching the unknown, the cyclical process of data construction and analysis was unfolded in a three stages case study design; 1) exploratory pilot study, 2) actual single case study and 3) communicative validation.

Exploratory Pilot Study (Feb. – Sept. 2007)

The pilot study was conducted on the Austrian National Library (ANL) – an active partner and contributor to the Europeana project. The ANL was selected due to pragmatic reasons of access. As it turned out during fieldwork, the ANL was running major digitization projects on their newspaper

(http://anno.onb.ac.at) and legal codices (http://alex.onb.ac.at) collections but also was in the process of preparing a pilot project on the preservation of .at domains. All ANL employees involved in Europeana related digitization projects were interviewed amounting to six interviews in total. Supplementary documents on the ANL and Europeana were collected as well. The leading interest at that time was the question how libraries need to change in order to remain libraries; a basic theme that remained for the whole duration of the research project. The main aims of the pilot were 1) to explore the potential for studying the Europeana project in relation to the research interest, 2) to develop a feasible set of expectations and focused research questions, 3) to enhance the author's interactional expertise in terms of librarian and information science practices (Collins et al. 2006) and 4) to gain access to the field (Feldman et al. 2003). [14]

Revelatory Single Case Study (Sept. 2007 – June 2011)

The initial exploration unveiled a fascinating research domain and a very active community of librarians, archivists and curators in terms of the digitization of cultural heritage artefacts as well as respective online services. By the same token, the potential of Europeana as a revelatory case was confirmed as it is quite unique in terms of scale and organizational setup based on an international network across librarian, museum and archival domains. Only comparable to the World Digital Library initiative (www.wdl.org) run by the Library of Congress and UNESCO in terms of scale, Europeana is unique in its cross-domain approach (EX-006:8). However, in light of the available resources, a comparative case study between Europeana and the WDL was simply out of scope for a PhD dissertation. Consequently, the research on Europeana was defined as a revelatory single case study (Yin 2003) on the implications of ICT on the field of libraries at first, later on memory institutions more broadly.

As a result of the pilot study, the theoretical conceptualization and the basic research interest of learning how libraries have to change in order to remain libraries, the following research question was formulated; *How does Europeana*

[14] A report on the preliminary analysis was submitted as an MSc dissertation at the Information Systems and Innovation Group, LSE, and can be downloaded at http://www.tigair.info/docs/inforVivid.pdf (Marton 2007).

differentiate itself from the information habitat as a library? As discussed in Chapter 3.7, the research question was rephrased during the course of the fieldwork in order to address the developments observed in the case, which reach beyond the librarian domain into the field of memory institutions at large. The final research question, therefore, was formulated as;

> *How does Europeana differentiate itself from the information habitat as a memory institution?*

Finally, access to the field was surprisingly easy to achieve as the responsible project members were very welcoming and open towards research projects about Europeana. Granted by Jill Cousins, the project leader of the Europeana project at that time and now Executive Director of the Europeana Foundation, the author had access to the central document server of the project, became member of the e-mail distribution list and was allowed to join conferences and project team meetings. However, there was no real location to be visited for extended fieldwork. During the main phase of data collection and analysis, Europeana was organized into working groups spread across Europe. By the same token, the schedule of meetings, as it was laid out when the Europeana project launched, presented few occasions for data collection. As it will be discussed further below, the data collection, therefore, focused mostly on documents and resorted to interviews in order to check the results with experts from the Europeana team in the last phase of the research design.

Communicative Validation (July 2011)

Communicative validation is a quality criterion recommended for qualitative research (Gaskell and Bauer 2000). Basically, it refers to experts from the research field to give feedback on the results of a qualitative study (Bygstad and Munkvold 2011). In this case, the analysis and interpretation of the single case study was checked with employees of the recently founded Europeana Office. In detail, 4 expert interviews were conducted at the office. Thus conceived, the interviews were not designed to collect new data but rather as discussions between the interviewee and the interviewer. The goal was to evaluate the viability of the results and explanations presented in the empirical study (see Chapter 4.6).

4.3 Corpus Construction

Following Flick's (2006) classification, corpus construction is a technique for defining a sampling structure gradually. In opposition to a priori determined sampling structures, as it is the case with random sampling (Diekmann 2001), a gradually determined sampling structure is developed during the research project as preliminary results and interpretations are used as a basis for selecting further data for analysis. Another term often used is purposive sampling, grounded theory theoretical sampling being the most prominent example (Glaser and Strauss 1967; Corbin and Strauss 1990; Silverman 2005; Corbin and Strauss 2008). However, in contrast to theoretical sampling, corpus construction relies on a two-dimensional unfolding of a social space that is the studying of 1) the variety of an unknown phenomenon based on 2) known categories.

Corpus construction has its origins in linguistics (Atkins et al. 1992; Biber 1993) and was adopted into qualitative social science methodology by Bauer and Aarts (2000). Introducing a terminology of its own, corpus construction helps to clear up some of the misunderstandings with respect to sampling and qualitative research, since terms like purposive or theoretical *sampling* are deemed to be misleading. A sample represents a given population constructed according to the distribution of known variables. As a consequence, the population needs to be known and, therefore, limited (Diekmann 2001). The researcher needs to be confident that the distribution of specific variables used to describe the population is mirrored in the sample.

Language, on the other hand, is an open system that grows and changes and, thus, provides for an unlimited set of possibilities for human communication. The population of language is, therefore, unknowable. In this case, the statistical rationale is neither valid nor viable. This applies to qualitative research as well, since it is impossible to define the population of respective phenomena such as social (inter)-actions, meaning(s), conversations, speech acts and so forth (Bauer and Aarts 2000). The rationale for selecting purposefully, therefore, does not derive from a notion that it is simply infeasible to conduct enough qualitative interviews in order to achieve representativeness of a population (Kemper et al.

2003). It rather derives from the very nature of the phenomena qualitative research usually deals with.

The studied phenomenon may be unknowable like language. Others may require qualitative exploration as its population is yet unknown like actions and situations in a given setting (Bauer and Aarts 2000). As a consequence, the selection rationale needs to take the peculiarities of these kinds of unknown and unknowable populations into account. In other words, proposing a number for how many interviews a researcher should conduct in order to achieve high quality qualitative research is not only arbitrary but defies any methodological basis. It would be an attempt "*to fit the sampling rationale […] like the choice of a false analogy*"(Bauer and Aarts 2000:19). Instead, corpus construction (and also theoretical sampling, for that matter) is an attempt to analyse the variety of a phenomenon and not its distribution.

A corpus is basically a collection of data that may vary from simply collecting text to any kind of symbolic token. Depending on the research interest, a social scientist or linguist would collect the relevant data to represent the full range of variability in the population. Hence, representativeness is not derived from proportional demographic sampling but rather aims for the inclusion of as much variation as possible (Biber 1993). Thus conceived, Bauer and Aarts (2000) propose four basic rules:

Rule 1	Proceed stepwise: select; analyse; select again.
Rule 2	Selection based on strata and function (known variety) precedes variety of representations (unknown variety).
Rule 3	Characterizing the unknown variety of representations has priority over anchoring them in existing, known categories.
Rule 4	Maximize the unknown variety of representations by extending the range of the known variety of strata and function until saturation is achieved.

Table 5: The four rules of corpus construction (Bauer and Aarts 2000).

Rule 1 refers to the very nature of the studied phenomena. Since qualitative researchers deal with phenomena being unknown or unknowable in their variety, the problem is not only where to start but also how to proceed and, above all,

when to stop. A qualitative researcher does not know what is to be found in the field and, therefore, has to expect the unexpected. In this case, corpus construction and traditional qualitative research prescribe a cyclical procedure of selection and analysis to increase the knowledge of a certain phenomenon step-by-step. The collection of data – the corpus – grows over time.

Rule 2 unfolds the social space into two dimensions; the known variety of social strata or functions and the unknown variety of the representations of the phenomenon a researcher wants to study. The researcher starts collecting data based on what is known. In opposition to quota sampling (Teddlie and Tashakkori 2003), however, rule 3 lays out that there does not need to be a correlation between the known and the unknown variety. The known variety of strata or functions is external to the population of the phenomenon and only helps as a systematization device for the selection process. For instance, Biber (1993) suggests a canon of external strata such as primary channel (e.g. written, spoken, scripted), format (e.g. published, unpublished), setting (e.g. public, private) and so forth. By the same token, Bauer and Aarts (2000) suggest the use of gender, education or age groups as known variety. Finally, rule 4 prescribes to continue selecting more strata or functions until saturation is reached, that is, "*one searches for different representations* [unknown variety] *only until the inclusion of new strata* [known variety] *no longer adds anything new*" (Bauer and Aarts 2000:34).

The cyclical procedure of selecting, analysing and selecting again is one way of addressing the paradox of how to study the unknown and the unknowable or, in other words, how to expect the unexpected. The paradox is unfolded in a step-wise procedure that allows for the research to feed back during the empirical study. Corpus construction explicitly requires to apply what is known as a structuring device for observation, thus, acting as a backdrop against which novelty, surprise or, in more general terms, information occurs as the unexpected (Kallinikos 2006b:103). At this point, corpus construction seamlessly ties in with the notion of observation discussed above in Chapter 3.2 (Esposito 1996). Studying the unknown is structurally determined by what is known expressed through the selection rationale or sampling method, which enables to observe by

differentiating between relevant and irrelevant data (Kemper et al. 2003). Consequently, the way corpus construction is implemented into the research project is a logical continuation of the epistemological outset.

Corpus construction also helps to organize the data. Based on Roland Barthes' semiology, Bauer and Aarts (2000) suggest three principles for selection and organization of the data; relevance, homogeneity and synchronicity. Relevance refers to having a thematic focus when collecting data. Each corpus should be relevant for answering the research question from one point of view only. Homogeneity proposes that different communication media or data from different collection methods should be organized into different corpora for comparison. For instance, semi-structured interview data forms one corpus, documents form another. Synchronicity includes the notion of time. The phenomenon of interest will most likely change during the course of the research project, which needs to be taken into account by the researcher by grouping data from different phases or periods of time into different corpora, again, for reasons of comparison.

To conclude, corpus construction formalizes the procedure of selection, not the selection itself. It offers an underlying rationale in order to maximize the variety of an unknown or unknowable population by reference to known social categories (e.g. gender), functions (e.g. occupation) or strata (e.g. social status). Consequently, a researcher is able to outline the external categories expected to guide the data collection. Due to the two-dimensional approach of known and unknown variety, corpus construction differs from theoretical sampling, which requires only a broad theme or research interest as a starting point complemented by a laissez-faire attitude towards the procedural planning of selecting units of data collection. Theoretical sampling, therefore, deals only with unknown variety (Corbin and Strauss 2008:146). If at all, known categories are used as a starting point but not as an incremental aspect of the selection process.

The following sections will outline how corpus construction was implemented into the empirical research by reference to the three principles of relevance, homogeneity and synchronicity. These sections can also be seen as part of the

documentation of the research project and, therefore, apply to the quality criterion of confirmability (see Chapter 4.6).

4.3.1 Relevance

The principle of relevance is a key device in distinguishing between relevant and irrelevant data by focusing the research on answering the research question. In this case; "*How does Europeana differentiate itself from the information habitat as a memory institution?*" The question is very specific as it asks for how Europeana observes itself as a memory institution within the information habitat as its environment. In a most general sense, the question refers to self-descriptions made within the project as well as self-descriptions presented to external parties such as, for instance, the European Commission or the wider communities of librarians, archivists and curators.

Thus conceived, relevance was used as the guideline for the second-order observation of Europeana – that is, relevance/irrelevance was the distinction made by the researcher through which Europeana was observed observing itself in difference to its environment (see Chapter 3). Rather than found in the field, the case was gradually constructed based on the cyclical process of collecting and analysing and collecting more data, as proposed by the four rules of corpus construction (see Table 5). The case is what is relevant for answering the research question while irrelevant data becomes part of the context. As a consequence, the unit of analysis (Yin 2003) was defined as the projects that were engaged in the creation of Europeana as a memory institution of the digital domain. As it will be discussed in Chapter 5.3, these projects were EDLnet/EuropeanaNet as well as Europeana v1.0 and EuropeanaConnect. Based on this approach, a corpus evolved over time with a specific structure that will be discussed in the next section on homogeneity.

The research question also refers to Europeana being a memory institution. Defined as canonizers of cultural artefacts through a process of selection, order and preservation (see Chapter 2.4.1) further sub-questions were derived in order

to specify the analytic gaze. These were 1) "*How does Europeana select?*" 2) "*How does Europeana order?*" and 3) "*How does Europeana preserve?*" Thus conceived, the research studied Europeana as a case for the emergence of memory institutional practices in the information habitat. In other words, the research is about what makes Europeana observable as a memory institution.

By contrast, the research question does not refer to phenomena that may be related to or even indispensable for the existence of Europeana as such. The following two examples may come to mind. For instance, the case study is not about the organizational structure of the Europeana initiative since that is not what distinguishes Europeana as a memory institution. Studying the ways Europeana organizes its own projects would require a different research question such as "How does Europeana distinguish itself from the information habitat as an *organization?*" By the same token, the politics of canonization of what becomes part of a cultural heritage and what does not is, admittedly, an important aspect (Stuurman and Grever 2007). However, structures of power and politics are an important aspect in every decision-making process and are not the distinguishing characteristic for canonization. In this sense, canonization and, thus, Europeana are political but that is not what makes Europeana into Europeana.

4.3.2 Homogeneity

Based on a single case study research design, corpus construction was used to maximize variety within the case rather than between cases. The external categories for organizing the data were based on the distinction between the addressee (author, interviewer) and address (receiver, interviewee) of communication (Biber 1993; Prior 2003; Flick 2006:248). Following the assumption that, as a project, Europeana will have a variety of project stakeholders, two corpora were constructed; one for internal communication (address and addressee are project members) and one for external communication (either addressee or address is a project member). However, during the course of the research it became obvious that the category of "external communication"

was too general. Consequently, a more differentiated set of external categories was developed in order to maximize variety (see Table 5). Table 6 summarizes the external categories based on the various parties related to the Europeana initiative.

		Address				
		Europeana (E)	**External Project Stakeholders (S)**	**Media (M)**	**Public (P)**	**Experts (X)**
Addressee	**Europeana (E)**	**EE** e.g. memos	**ES** e.g. project deliverables	**EM** e.g. press release	**EP** e.g. project website	**EX** e.g. journal publication by Europeana
	External Project Stakeholders (S)	**SE** e.g. project contract	**SS** e.g. project evaluation	**SM** e.g. press release	**SP** e.g. public speech	**SX** e.g. project reports
	Media (M)	**ME** e.g. interview of a project member	**MS** e.g. interview of a EU representative	**MM** n/a	**MP** e.g. newspaper article	**MX** e.g. interview of a librarian expert
	Public (P)	**PE** e.g. user feedback	**PS** e.g. petitions	**PM** e.g. letter to the editor	**PP** e.g. online discussions	**PX** e.g. user feedback
	Experts (X)	**XE** e.g. recommend-ations	**XE** e.g. consultation	**XM** e.g. letter to the editor	**XP** e.g. blogs	**XX** e.g. general guidelines
	Researcher (R)	**RE** e.g. interviews	n/a	n/a	n/a	n/a

Table 6: External categories and examples for possible units of data collection.

The external categories themselves were gradually developed as more data was collected and analysed. The final compilation contained the Europeana project (E), external stakeholders such as the European Commission (S), media outlets such as newspapers (M), the wider public (P), experts within the field of memory institutions and information science (X) and, finally, the researcher himself (R). A cross-tabulation of these categories resulted in a set of homogeneous corpora designed to collect and organize data according to the criterion of address and addressee of communication (Prior 2003; Flick 2006:248) rather than according

to communication genres (e.g. memos, reports, deliverables) (Biber 1993) or data collection methods (Bauer and Aarts 2000). Thus, each cell in Table 6 represents a homogeneous corpus. For instance, a Europeana project memo would be a document authored by project members for project members (codified as EE). A project deliverable would be authored by the Europeana project and received by the European Commission (codified ES). A newspaper report about Europeana would be authored by a reporter (media) and targeted at the wider public (codified as MP).

The examples given in Table 6 are only possibilities for data collection not actually collected or to-be-collected data. The two-dimensional approach according to address and addressee opens up a vast array of sources a researcher could collect data from – be it project internal memos to librarian expert online blogs, press releases, public speeches and news reports. The array of possible corpora, however, is too large to be analysed within the given timeframe requiring a return to the principle of relevance (see above) as a guiding criterion to further focus the data collection. As stated in the research question, what is relevant for the research is defined by what is relevant for the Europeana project, i.e. documents produced by and/or addressing Europeana. Consequently, the field for data collection was delimited to the first column and row of the address and addressee cross-tabulation (see grey marked cells in Table 6); everything else was only supplementary material and is not part of the analysis.

Because of limited resources, the main type of data collected was documents (Feldman et al. 2003; Prior 2003). Since the involved experts of the Europeana initiative were spread all over Europe, it was too costly to visit each of them for interviewing or observation. Interviews would have also demanded too much of the experts' time during their rare meetings. Therefore, the design was split into two phases; in the first phase the focus was on collecting and analysing documents – the main data type of the corpus. The documents were downloaded from the central document server after access has been granted by the project coordinator amounting to a total of 137 documents. The second phase contained four expert interviews used for communicative validation of the interpretation and concepts developed in the first phase. The experts were selected by the

Executive Director of the Europeana Foundation, Jill Cousins, among the employees of the Europeana Office, which was founded during the later stages of the fieldwork (see Chapter 5.3).

4.3.3 Synchronicity

At the beginning of the research project on Europeana, the principle of synchronicity did not seem to be relevant. The initial design distinguished between documents produced before the launch of the Europeana initiative in July 2007, used for describing the historical context, and documents produced during the Europeana project. However, as the prototype online service at www.europeana.eu was launched, the responsible project named EDLnet/EuropeanaNet finished its work. Its successors were two new projects named Europeana v1.0 and EuropeanaConnect with the focus on developing the prototype into an operational service (for a detailed history of Europeana see Chapter 5.3). As the first documents produced by these two new projects became available, it came as a surprise that the operational service was to become a completely different kind of online service in comparison to the prototype. The documents explicitly stated that the operational service was going to break with the traditional paradigms of memory institutions, which had still informed the design of the prototype.

In terms of synchronicity, two cycles can be defined – not counting the pilot study. The crucial event starting a new cycle was the launch of the Europeana prototype in November/December 2008. Data from the start of the main case study until the launch was assigned to one corpus (cycle 1); data from after the launch to another corpus (cycle 2). Thus conceived, the analysis evolved into the comparison of these two cycles as the main approach for answering the research question. As Chapters 6 and 7 will demonstrate, cycle 1 refers to the traditional librarian paradigm of "silos of knowledge", while cycle 2 refers to a paradigm shift towards networking and transversality.

Corpus	Time Period	Units of Analysis	Relevant for…
Pilot Study	Feb. 2007 – Aug. 2007	Digitization projects of ANL (ANNO, ALEX)	the exploration of the research domain and formulation of a research question
Historical Context	until July 2007	n/a	the study of the historical prelude of Europeana
Cycle 1	July 2007 – Dec. 2008	EDLnet EuropeanaNet	the study of the traditional paradigm of memory institutions
Cycle2	Dec. 2008 – July 2011	Europeana v1.0 EuropeanaConnect	the study of the new paradigm of networking and transversality

Table 7: The synchronized corpora of the case study.

To conclude, this section provided for the rationale that guided the unfolding of the paradox every research project faces – how to expect the unexpected. In this spirit, the technique of corpus construction was introduced as a stepwise procedure of selecting, analysing and selecting data based on the two dimensions of known and unknown variety. By the same token, corpus construction was also used to construct relevant, homogeneous and synchronized corpora of, mostly, collected documents used for comparison. The masterlist (Silverman 2005) of all the documents and interviews as well as their classification in terms of relevance, homogeneity and synchronicity can be found in Appendix 9.1.

4.4 Data Analysis

The data analysis was based on a modified version of grounded theory coding techniques (Glaser and Strauss 1967; Corbin and Strauss 2008) developed by Uwe Flick (2006:307-12). Referred to as thematic coding, the method revolves around the notion of developing a theme for every unit of analysis (document, interview) treating it like a case of its own. A description is authored including a central statement or motto of the unit of analysis, followed by information relevant to the research question as well as a summary of the central topics raised in the document or interview (see Appendix 9.2 for an example). During the analysis, the researcher oscillates constantly between the description and the analysis modifying the interpretation accordingly. The analysis itself relies on

open coding for constructing categories and selective coding for developing thematic domains – core categories in grounded theory terminology (Corbin and Strauss 2008). As a result, the researcher analyses one unit of analysis after the other and constantly modifies the thematic domains as the analysis progresses. By doing so, a thematic structure emerges. "*The result of this process is a case oriented display of the way it specifically deals with the issue of the study* […]" (Flick 2006:309). The aim of this specific approach is to increase the comparability of interpretation between the units of analysis by means of a thematic structure while remaining open to the specificities of each individual document or interview at the same time.

The rationale for selecting thematic coding relies on its compatibility with corpus construction since it allows for a priori theoretically developed case selection criteria. Hence, it corresponds with the conceptual elaboration of a research domain (memory institutions in the information habitat) and the selection of a viable case thereof (Europeana). Still, some minor modifications are necessary in order to apply the technique according to the epistemological foundation and design of the research project. First of all, the research is based on a different unit of analysis. Instead of analysing interviews, thematic coding was used to develop thematic domains for the Europeana project. The research strategy was defined as a single case study, since a comparison with other cases was not within the scope of the research project. Second, the selection rationale was based on corpus construction rather than on theoretical sampling as suggested by Flick (2006:311).

Last, the notion of a thematic domain based on a core category needed to be aligned to the basic principle of observation as making a distinction (Esposito 1996). The thematic domains and core categories need to be seen as distinctions as well since they are forms based on observation. It is in this sense that Gibson et al. (2005:10f) suggest to think of primary distinctions rather than of core categories;

> "*The core category is generated on the claim that the writer understands the main concerns of those being observed. This would no longer be appropriate within a truly post ontological tradition. The category in the classical way acts as a kind of container or hold-all concept and this is*

> *more appropriate to a time when the structure of things was believed to contain its object.* […] *Our suggestion therefore would be to replace the notion of the core category with the idea of the primary distinction*".

The primary distinction can, therefore, be seen as that which makes Europeana observable as a memory institution. Thus, thematic coding is applied in order to observe the primary distinction from a second-order perspective (Luhmann 1996b; 2002b). As it will be discussed in Chapter 7, the analysis resulted in the formation of a form referred to as transversality/universality based on the contrast between the two paradigms of institutionalized memory as described by the two cycles of corpus construction (see Chapter 4.3.3).

In detail, the analysis contained the following steps; 1) a document was selected according to the rules of corpus construction, 2) the document was analysed according to its relevance for answering the research question, 3) if the document was relevant, open coding was conducted, 4) during the analysis thematic domains were developed as memos. Last, selective coding was applied to further develop the thematic domains and, ultimately, the primary distinction based on the documents collected. An example for a thematic coding sheet can be found in Appendix 9.2.

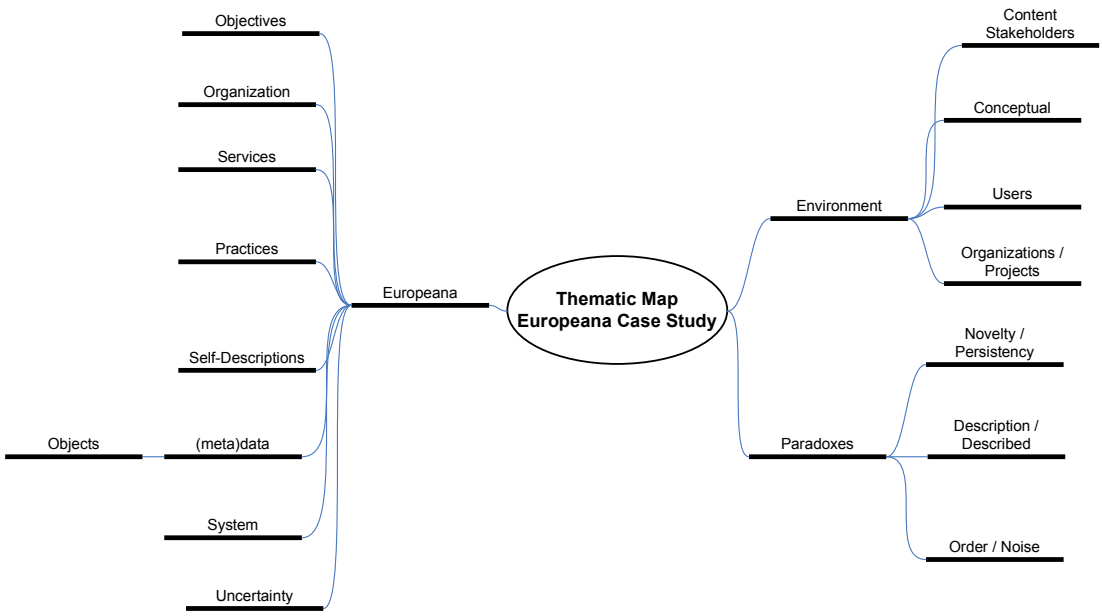


Figure 6: Schematic overview of the thematic structure of the Europeana case study.

Finally, the thematic domains were organized in a mind map in order to construct the thematic structure of the case study. Figure 6 shows a schematic overview in the form of a thematic map. More detailed examples of parts of the thematic structure can be found in Appendix 9.3.

4.5 Ethical Considerations

The research followed the guidelines of the Research and Project Development Division at the London School of Economics and Political Science (LSE)[15] as well as the Research Ethics Framework (REF) by the Economic and Social Research Council (ESRC).[16] The LSE Research Ethics Checklist[17] was filled out and approved by the PhD supervisor, Prof. Jannis Kallinikos. According to the checklist, no further steps were necessary in terms of the LSE research policy since the informants were not to be deceived or caused any kind of pain or stress (see Appendix 9.5 for the research ethics checklist).

Following the LSE ethics guidelines, an informed consent form was created for interviews to be signed by all participants. The form explained the purpose of the research and the rights of the interviewee to stop the interview at any moment, to receive a copy of the digital recording and to have the original recording deleted by the researcher (see Appendix 9.6). Out of the four interviews conducted for communicative validation, none asked to delete the recordings. One interviewee refused to sign the informed consent form. For the PhD dissertation and further publications, the use of the name "Europeana" as well as the quotation of documents that were not declared "confidential" by the Europeana initiative was approved by the Executive Director of the Europeana Foundation Jill Cousins (Cousins 2010).

---

[15] http://www2.lse.ac.uk/intranet/researchAndDevelopment/ethicsGuidanceAndForms/Research_Ethics_Review_Policy_FINAL.pdf, last access: 23 July 2010.

[16] http://www.esrc.ac.uk/ESRCInfoCentre/Images/ESRC_Re_Ethics_Frame_tcm6-11291.pdf, last access: 23 July 2010.

[17] http://www2.lse.ac.uk/intranet/researchAndDevelopment/ethicsGuidanceAndForms/Researh_Ethics_Review_Checklist_FINAL.doc, last access: 23 July 2010.

4.6 Quality Criteria

The quality of research entails two aspects; how to conduct quality research (good practice) and how to enable stakeholders (funding bodies, reviewers, editors, readers and so forth) to assess a research proposal or project (evaluation) (Flick 2007). However, it was not until the 1980s that also the second aspect became a central issue in the methodological discourse of social science research (Packer and Addison 1989; Golafshani 2003; Flick 2007). In their seminal publication, Lincoln and Guba (1985:290) summarized the problem of quality assessment in four questions.

Quality Criterion	Question Asked	Conventional Answer
Truth value	How can one establish confidence in the "truth" of the findings?	Internal validity
Applicability	How can one determine the applicability of the findings in other contexts?	External validity
Consistency	How can one determine whether the findings are replicable in the same or a similar context?	Reliability
Neutrality	How can one determine the degree of bias introduced by the inquirer?	Objectivity

Table 8: The four quality criteria of scientific research in Lincoln and Guba (1985:290).

The conventional answers to these four questions are of course internal validity, external validity, reliability and objectivity. In terms of quantitative social science research, these ideals are approximated by various highly formalized techniques such as measurement-validation in terms of internal validity, representativeness through statistical random sampling in terms of external validity, test-retest in terms of reliability or the correlation-coefficient in terms of objectivity (Diekmann 2001).

Qualitative research, on the other hand, does not seem to lend itself to formalized assessment of validity, reliability and objectivity (Seale 1999; Lee and Baskerville 2003). The underlying reasoning for the scepticism towards quality assessment is based on two basic themes. The first line of argumentation says that the term qualitative research itself is problematic. Being rather an umbrella

term for various methods, it is questionable whether it is even possible to have a common set of standards, since qualitative research traditions differ profoundly in terms of epistemology, methodology, format and even style (Madill et al. 2000; Reicher 2000:5). The second argument addresses the problem of benchmarking (Flick 2007). Since qualitative research does not measure a phenomenon by means of scaled instruments, since it does not even have basic units of measurement, how should qualitative research, in turn, be measured in a standardized way? Consequently, it is impossible to assess the results of qualitative research projects and, therefore, to benchmark their quality (Patton 2002:50). In short, critics doubt the applicability of quality criteria in qualitative research because of a lack of standardization in terms of its epistemological foundation, practice and results.

Agreed, qualitative methods do differ substantially, however they are similar in terms of procedure and selection rationale. As already discussed in Chapter 4.2, qualitative research is a cyclical process of explaining unknown or unknowable phenomena in their variety (Glaser and Strauss 1967; Bauer and Aarts 2000; Flick 2006:97-104). This applies irrespective of the epistemological foundation, methodology, format or style of the research project. Second, the point on benchmarking confuses numerical/non-numerical research with the degree of formalization (Bauer et al. 2000). Quantitative research delivers numerical results, but it takes an additional step of abstraction to formalize the quality of these results. This does not come with the numbers but rather is an accomplishment by the researcher (Jensen 1991:7). Qualitative research may lack formalization but that is not the case because its results do not come in numbers (Lee and Hubona 2009).

Over the years, qualitative methodology provided various answers to Lincoln and Guba's four questions (Flick 2007) ranging from conventional approaches of implementing reliability and validity into qualitative research (Kirk and Miller 1986; Yin 2003) to criteria for specific research traditions and strategies such as grounded theory (Bryant 2002; Charmaz 2006; Urquhart et al. 2009), hermeneutics (Packer and Addison 1989; Walsham 1995; Klein and Myers 1999; Walsham 2006), ethnomethodology (Titscher et al. 2000), case study research

(Benbasat et al. 1987; Eisenhardt 1989; Yin 2003), critical research (Cecez-Kecmanovic 2011; Myers and Klein 2011) or action research (Baskerville and Wood-Harper 1998; Baskerville and Myers 2004). Others propose a more procedural approach that includes standardized documentation (Flick 2007:16-21; Gibbs 2007), qualitative interviewing (Leggewie 1987; Kvale 1994; Myers and Newman 2007) or techniques of convincing the reader of the credibility, trustworthiness or authenticity of the researcher's accounts (Elliott et al. 1999; Golden-Biddle and Locke 2007).

While these two approaches seek to either import or to completely abolish validity, reliability and objectivity, the approach applied in this research discusses criteria as functionally equivalent to the conventional approach. Gaskell and Bauer (2000) trace the quantitative social science tradition of quality assessment back to two basic functions – confidence and relevance. Confidence denotes the researcher's obligation of convincing an informed outsider that the findings are not simply made up. Relevance is about linking the data with the theory (internal relevance) or about reporting novel and unexpected findings (external relevance). Based on these functions, they outline equivalent criteria for achieving confidence and relevance such as triangulation and reflexivity (confidence), thick description (confidence and relevance) or communicative validation (relevance). In a similar fashion, Lincoln and Guba (1985:219) propose the following criteria:

Criterion	Explanation	Equivalent to…
Credibility	Production of credible findings and interpretations achieved through, for instance, prolonged engagement, triangulation and negative case analysis	Internal Validity
Transferability	Provision of the researcher's data-base to enable other researchers to transfer the findings to another setting by means of, for instance, thick description	External Validity
Dependability	Verification of the research process and its product by an "inquiry audit"	Reliability
Confirmability	Scrutiny of raw data, notes, reconstruction and synthesis products, and preliminary developmental information through a "confirmability audit"	Objectivity

Table 9: Quality criteria for qualitative research after Lincoln and Guba (1985).

Following Lincoln and Guba's (1985) criteria as well as Gaskell and Bauer's (2000) suggestions, evidence for the quality of this research was addressed in the following way allowing a reader to evaluate the viability of the research design, empirical study and results.

Credibility

Credibility is addressed through the prolonged engagement with the research domain and case over a period of four years. The explicit discussion of the theoretical expectations, research design, research question and empirical methods should allow a reader to make an informed judgement on the quality of the research. In this spirit, Chapter 4.3 documents corpus construction by outlining the selection rationale in a transparent and systematic way. Chapter 4.4 documents the analysis of the collected data based on thematic coding. As a supplement, the appendix provides the necessary details and examples for a reader to evaluate the quality of the case study.

At the end of the empirical study, the collection of new documents did not deliver any new insights with respect to answering the research question, which is the accepted indicator for the saturation of the developed categories and, thus, functions as the stop criterion for the cycle of data collection and analysis (Glaser and Strauss 1967; Bauer and Aarts 2000). Since there are no conventions or guidelines available on how to report saturation, evidence for saturation was not adduced.

Transferability

Since some of the data is confidential, access to the complete data-base is not possible. In agreement with the informants, the interview transcripts are not published either. However, quotes, figures and diagrams from the data, approved by the Executive Director of the EDL Foundation, were included extensively in order to provide a thick description of the case and the setting.

Dependability

The dependability of the research will be evaluated by the PhD examiners and the supervisor who function as an inquiry audit in this case. The documentation

and justification of each step of the research process should allow them to make an informed judgement based on this quality criterion.

Confirmability

The third and last stage of the empirical research was specifically designed in order to provide for confirmability. Referred to as communicative validation (Gaskell and Bauer 2000) or backtalk (Lanzara 1991; Bygstad and Munkvold 2011), the analysis and interpretation was checked with employees from the Europeana Office by means of expert interviews who, as a consequence, performed a confirmability audit. Based on a topic guide (see Appendix 9.4), the experts were confronted with the transformation of Europeana from a portal-based to a networked service and its interpretation as a paradigm shift from "silos of knowledge" (universality) to transversality. The experts agreed with this interpretation. In addition, the evaluation by the PhD examiners and the supervisor can also be seen as a confirmability audit.

4.7 Conclusion

In this chapter, methodology was discussed as a paradox as it suggests scientific methods for the study of the unknown and, ultimately, techniques on how to expect the unexpected. Going back to the self-referential process of scientific research as depicted by Krohn and Küppers (1989:89) in Figure 5, the paradox was unfolded into a linear process with a beginning and, more importantly, an end the following way;

	Krohn and Küppers	**Case Study**	**Explanation / Outcome**
1	Crosslinking Information (Patterns)	Initial literature review	Exploration of the field of digital libraries
2	Derivation of Expectations (Hypotheses)	Definition of an initial research interest	Expectation that the field of digital libraries is a viable research domain and Europeana is a feasible case study
3	Setup of Procedures (Construction)	Pilot study on ANL	Exploratory case study on digital libraries

	Krohn and Küppers	Case Study	Explanation / Outcome
4	Generation of Effects (Effects)	Accessing the field; Fieldwork	Semi-structured interviews and document collection
5	Observation of Effects (Data)	Corpus construction	Construction of relevant, homogeneous and synchronized corpora
6	Interpretation of Data (Information)	Thematic coding	Data is analysed with regards to the expectations formulated in step 2
7	Crosslinking Information (Patterns)	Preliminary results of the pilot study	Confirmation of the research domain and Europeana as a viable empirical case study
8	Derivation of Expectations (Hypotheses)	Formulation of the research question	How does Europeana differentiate itself from the information habitat as a digital library?
9	Setup of Procedures (Construction)	Main study on Europeana	Revelatory single case study on digital libraries and the information habitat
10	Generation of Effects (Effects)	Accessing the field; Fieldwork	Collection of documents
11	Observation of Effects (Data)	Corpus construction	Construction of relevant, homogeneous and synchronized corpora
12	Interpretation of Data (Information)	Thematic coding	Data is analysed with regards to the research question formulated in step 10
13	Crosslinking Information (Patterns)	Comparison of interpretation and expectations	The focus on digital libraries does not explain the results of the analysis
14	Derivation of Expectations (Hypotheses)	Redefinition of final research interest; Reconceptualization of theoretical framework	The research focus shifts from digital libraries to memory institutions in the information habitat
15	Interpretation of Data (Information)	Thematic coding	Collected documents are analysed again based on the focus on memory institutions
16	Crosslinking Information (Patterns)	Further literature reviewed	Literature on social memory studies and memory institutions reviewed
17	Derivation of Expectations (Hypotheses)	Reformulation of research question	How does Europeana differentiate itself from the information habitat as a memory institution?

	Krohn and Küppers	Case Study	Explanation / Outcome
18	Setup of Procedures (Construction)	Main study on Europeana	Revelatory single case study was continued with a new focus on memory institutions and the information habitat
19	Generation of Effects (Effects)	Fieldwork	More documents were collected as the Europeana project continued its work
20	Observation of Effects (Data)	Corpus construction	Continued construction of relevant, homogeneous and synchronized corpora
21	Interpretation of Data	Thematic coding	Collected documents were analysed; The empirical study was looped back to step 18 until saturation was achieved. Once saturation was achieved, the study moved on to step 22.
22	Interpretation of Data (Information)	Comparison of the two cycles of the Europeana project	The comparison of cycle 1 (traditional library paradigm) and cycle 2 (transversal paradigm) emerged as a viable distinction for interpreting the data
23	Crosslinking Information (Patterns)	Development of the primary distinction	The form of transversality/ universality is developed into the primary distinction
24	n/a	Communicative validation	Interpretation of data is checked with experts from the field

Table 10: Unfolding the self-referentiality of scientific research into a linear case study (see Krohn and Küppers 1989:89).

Based on a revelatory single case strategy, the research was designed according to the qualitative methodology of corpus construction to unfold the paradox in a step-wise, cyclical fashion of selecting, analysing and selecting data (Bauer and Aarts 2000). Complemented by the analysis technique of thematic coding (Flick 2006:307-12), a thematic structure about the case on the Europeana project was developed. Two distinct phases of the project were compared as a means to find answers to the research question of *how Europeana distinguishes itself from the information habitat as a memory institution*. The cycle of collecting, analysing and collecting more data (see steps 18-21 in Table 10) was followed until saturation was achieved. Finally, the primary distinction of transversality/universality was developed as the form that answers the research question

The following chapters will discuss in depth the case study and its results. In detail, Chapters 5 and 6 will describe Europeana as a case based on the thematical structure constructed through thematic coding. Chapter 7 will then focus on the notion of transversality as a key concept for the understanding of Europeana and the memorization of the information habitat.

5. CASE STUDY – CONTEXT

5.1 Introduction

The following chapters will exclusively focus on Europeana as a case study of memory institutions that are stepping into the emerging social memory paradigm of the information habitat (Primary Research Group 2008). To begin with, this chapter will briefly introduce the Europeana initiative and outline its history and main contextual aspects.

As it was already discussed in the previous chapter, the distinction between the case and its context or environment is a tricky and, to some extent, strategic task. It is strategic inasmuch as the distinction serves the exploration and construction of answers to the research question rather than the capturing of a given truth of an instructive reality. Therefore, the case is delimited according to the criterion of relevance as it is proposed by corpus construction (see Chapter 4.3.1). The case is a methodological artefact or construction rather than an object discovered in the field. Everything else that is not directly related to answering the research question is, therefore, defined as being contextual. Although still an aspect of the case study, the context is what remains out of focus in order to sharpen the view on the aspects addressed by the research question. Thus conceived, the history of Europeana, wider related issues such as private-public-partnerships as well as the organizational structure of Europeana are seen as context for the analysis.

As a final note, the case study chapters reference the collected documents that serve as evidence for the description, analysis and interpretation. The chapters also rely heavily on quotations taken from the data corpus in order to provide for thick description (Bauer and Aarts 2000). The referenced documents are identified according to the corpus they have been assigned to (see Chapter 4.3.2), followed by a unique number and a page number, in case of a quotation. For instance, “EE-001:3” is the reference for page 3 of document number 001 from the corpus “EE” – the corpus containing Europeana project internal documents.

A complete list of the whole corpus including document titles, document types and other details can be found in Appendix 9.1.

5.2 Europeana – An Overview

Europeana is an umbrella term for a variety of projects coordinating the efforts of European memory institutions in terms of digitization, meta-data enrichment and accessibility of cultural heritage artefacts financed by the European Commission (EC) and EU member states. The most visible part of the project is certainly the online portal www.europeana.eu. A single point of entry, it is not only a unified meta-catalogue of European memory institutions but also offers direct access to the digitized cultural heritage artefacts themselves provided by Europeana member libraries, archives and museums.

Figure 7: Screenshot of the online portal www.europeana.eu.

Since the launch of the prototype in November/December 2008, Europeana has remained a work in progress during the authoring of this dissertation and will remain as such for years to come. To be launched late 2011, the first operational

release – Europeana v1.0 - will mark an important milestone of incremental implementations of advanced functionalities such as semantic search, multilingual accessibility, multimedia annotation systems and an Application Programming Interface (API), to name a few. Although the project has not come to an end until the submission of this dissertation, it has made enough progress to provide strong evidence for the comparative analysis and discussion to follow.

The digital cultural heritage artefacts to which Europeana grants access are hosted by the respective libraries, archives and museums that hold the original cultural artefact and its digitized version. For mostly legal reasons, Europeana merely hyperlinks to the digitized artefact rather than stores them at its site. According to latest available statistics from November 2010, Europeana grants access to over 14 million digital artefacts – text, image, audio and video (EM-003) – a goal that was initially expected to be reached by 2014 (ES-011:9). The following milestone is set to be over 30 million items by 2015 and ultimately "*all of Europe's digitized cultural heritage*" by 2025 (EP-004:5).

The term digital artefact is to be understood in the broadest sense and will be used interchangeably with the terms digital objects, digital items, digital content, digital cultural artefacts and digital cultural heritage artefacts throughout the case study. Within Europeana, every item, object or artefact is a cultural heritage artefact; hence there is no need to keep in mind the distinction between cultural artefact and cultural heritage artefact as it was discussed in Chapter 2.4.1. These terms basically denote anything that is digital. One can find, for instance, high quality scans of books but also thumbnail size images of artefacts, which hardly qualify as usable digital objects in their own rights. While Europeana is keen on introducing minimum standards in terms of digitization (as well as the description of digital objects for that matter), the project obviously has accepted anything a library, archive or museum has been willing to provide in order to reach the targeted number of objects accessible through Europeana (EE-071:11).

As Table 11 shows, the vast majority of objects are images such as digital photography of a painting or an image scan of a handwritten letter. However, the categorization of image, text and audio/video needs to be handled with care. For

instance, a quick search for "Mozart" results in OCR scanned books (digital objects consisting of mostly computational text) but also image scanned handwritten artefacts, both being categorized as text. In other words, it is not clear what counts as text and what as image. Does the digitized artefact count as text when the material original is text-based or when the digitized artefact itself is text-based? The separation between the original artefact and its digitized version has become a central point and is explicitly addressed in later developments of the metadata scheme to be implemented in future releases of the online service (see Chapter 6.5).

Image	64%
Text	34%
Audio/Video	2%

Table 11: Percentage of digital media types accessible via Europeana - November 2010 (EM-003; EE-071:8).

The content is provided by more than 1500 European memory institutions either directly or through so-called aggregators (EM-003). Aggregators are basically associations representing domain specific or cross-domain interests on a national or European level. Members of these aggregators usually share standards or even a union catalogue, which makes the ingestion of their metadata into the Europeana framework easier to manage. For instance, museums are represented by the International Council of Museums Europe (ICOM-Europe), archives by the European Regional Branch of the International Council on Archives (EURB-ICA) and national libraries by the Conference of European National Librarians (CENL). On a national level, one can find, for instance, the Museum, Libraries and Archives Council UK (MLA) or the German based Arbeitsgruppe zu Europäischen Angelegenheiten für Bibliotheken, Archive, Museen und Denkmalpflege (EUBAM). [18] Direct contributors are usually large individual memory institutions, such as the British Library, who have the necessary know-how and financial resources to deliver their digitized items directly to Europeana.

[18] Working Group on European Affairs of Libraries, Archives, Museums and Historic Preservation (my translation).

With respect to countries, the vast bulk of content contributions are made by France and Germany (roughly 17% each). 19 of the 27 member states, however, provide individually for less than 3% each. Quite a tremendous bias in terms of the representativity of European cultural heritage (see Table 12), underrepresented countries are given a higher priority in support of the digitization and ingestion of their cultural heritage artefacts (EP-002:10). Considering the fact that, up until June 2010, France amounted for 30% of the digital objects (EE-063:5), a very important issue.

EU Member States	Percentage
France	17.98%
Germany	17.10%
Sweden	9.69%
Spain	8.85%
The Netherlands	7.89%
Italy	7.03%
Ireland	6.47%
United Kingdom	6.14%
Poland	2.81%
Belgium	1.45%
Greece	1.40%
Finland	1.31%
Slovenia	0.98%
Austria	0.32%

EU Member States	Percentage
Estonia	0.27%
Romania	0.17%
Slovakia	0.15%
Denmark	0.11%
Portugal	0.11%
Bulgaria	0.08%
Czech Republic	0.08%
Hungary	0.07%
Lithuania	0.05%
Luxembourg	0.04%
Latvia	0.01%
Cyprus	<0.01%
Malta	<0.01%

European collections without an attribution per Member State, e.g. from European projects	2.91%
Non-EU countries	
Norway	5.76%
Others	0.85%

Table 12: Europeana content by country. Percentage of the total number of objects in Europeana contributed by country – November 2010 (EM-003; EP-005:8).

Although Europeana's inception is clearly rooted in the domain of libraries, the need to step beyond established boundaries has become clear over time – be it the boundaries between libraries, archives and museums or the external boundaries with respect to users and the wider environment of memory

institutions. The first point was symbolically addressed by rebranding the initiative from "European Digital Library" (EDL) to "Europeana" thus conveying the message that the initiative reaches beyond the librarian domain.

The latter point has been repeatedly called for in terms of bringing Europeana to the users rather than to expect them to come to the portal as it was traditionally the case.

> "*Europeana will enable citizens across Europe to access this valuable material in ways unimaginable before; through a well designed portal but also at places where they are already active, such as social networks. The reuse of this material in turn will foster a true dialogue and create new value through mash-ups and companies able to create new business out of the aggregation of this material, such as semantic operators*" (ES-011:9).

Europeana presents itself as a revelatory and, to some extent, extreme case study for the digitization of cultural heritage and social memory. The leading project members see it as a new breed of memory institution breaking with traditional boundaries – be it institutional, technological or cultural. Europeana is expected to go beyond the traditional setup of isolated silos of knowledge by bringing the content to the users and, most importantly, by enabling users to re-use, recycle and mash-up Europeana's content as they see fit (EX-002:8-9).

Europeana seems to release the tight control over the collection of cultural heritage artefacts and its organization traditionally assigned to the institutions of libraries, archives and museums. In other words, Europeana marks a transition for memory institutions from being (meta-)aggregators to also becoming information brokers of the online world (EE-034) – a transition that will be analysed in depth in the next chapter. In order to give an overview and understanding of the scope and structure of the project, this chapter will continue with a focus on a detailed description of Europeana and its history.

5.3 Historical and Political Context[19]

The birth of Europeana can be traced back to Google's announcement of its "Print Project" on 14th December 2004 (ES-016:6). Later renamed into "Google Books" (http://books.google.com), Google stepped into the world of print media by mass digitizing books as well as making them findable and, to some extent, accessible via its index and search engine services. As discussed in Chapter 2.4.4, the scanning of material from the public domain but also of out-of-print editions caused a lot of turmoil in the print world leading to protests from librarians and court cases by publishers and authors (Bjorner 2006).

It was the protest of Jean-Noel Jeanneney, Director of the National Library of France from 2002 until 2007, which reached the farthest into the sphere of European politics. In his article "Quand Google défie l'Europe", published in Le Monde 22nd January 2005, Jeanneney strongly criticized Google's approach in terms of the quality of its scans and metadata and accused Google of being disrespectful of the public domain (Bearman 2006). His article also called for a European initiative to counter the tendencies towards privatization and Anglo-American domination of the public domain of knowledge and culture in the online world – a call directed at President Jacques Chirac and Chancellor Gerhard Schröder, at that time the heads of the states of France and Germany respectively (ES-016:6; EX-001:1-2).

Jacques Chirac and Gerhard Schröder responded by sending a letter[20] to the President of the European Commission, José Manuel Barroso, on 28th April 2005. Joined by the leaders of Poland (Aleksander Kwasnieski), Italy (Silvio Berlusconi), Spain (José Luis Rodriguez Zapatero) and Hungary (Ferenc Gyurcsány) (ES-016:6), the letter praised the unequalled richness and diversity of heritage held in Europe's libraries and advocated the foundation of a virtual European library.

[19] A broad overview can be found at http://ec.europa.eu/information_society/activities/digital_libraries/timeline/index_en.htm, last access: 20th January 2011.
[20] The scanned version of the letter can be found at http://ec.europa.eu/information_society/activities/digital_libraries/doc/letter_1/index_en.htm, last access: 18th January 2011.

> "[If Europe's cultural heritage] *is not digitised and made accessible online, this heritage will not occupy its rightful place in the future knowledge landscape*" (translation in Edwards 2009:4).

The initiative was supported by José Manuel Barroso in a reply from 7th July 2005[21] shortly after the European Union launched its i2010 agenda in June 2005 (ES-016:6, EX-001:1-2). Under the leadership of Viviane Reding, Information Society and Media Commissioner, the i2010 agenda outlined a five year plan for the formation of a European Information Society for growth and employment by focusing on research and development as well as on the implementation and diffusion of ICT services and appliances. In a communication issued on 1st June 2005, the European Commission highlighted the central role of digital libraries in the European Information Society by declaring it as one of three flagship projects of the i2010 agenda; the others being technologies for the wellbeing of the elderly and intelligent cars (European Commission 2005a:11).

From those early days on, it was already obvious that a common European digital library would not only be a response to the rise of digitality and the internet but also a driver for the emergence of a common European identity. As Viviane Reding was quoted in a press release from 30th September 2005;

> "*Without a collective memory, we are nothing, and can achieve nothing. It defines our identity and we use it continuously for education, work and leisure*."[22]

The press release heralded the launch of the "i2010: Digital Libraries" initiative by the Commission's Directorate for Information Society and Media outlining its vision for a future European digital library, which was subsequently supported by the Ministers of Culture of the Member States (European Commission 2006:4). A core document for Europeana, it defined digital libraries as "*organised collections of digital content made available to the public*" (European Commission 2005b:3) involving digitized as well as born-digital material.

---

[21] The scanned version of the letter can be found at http://ec.europa.eu/information_society/activities/digital_libraries/doc/letter_2/index_en.htm, last access: 18th January 2011.

[22] The original press release can be found at http://europa.eu/rapid/pressReleasesAction.do?reference=IP/05/1202&format=HTML&aged=0&language=en&guiLanguage=en, last access: 18th January 2011.

As broad goals, the vision set up three main pillars for a future digital library;

- Online accessibility;
- Digitization of analogue collections;
- Preservation and storage.

The communication also pointed out a lack of collaboration among the EU Member States and unsolved challenges of costing and organization as well as technological problems and copyright issues. As already shown in Chapter 2.4.4, digitization initiatives of cultural heritage were already implemented by various nation states and their leading institutions such as the French National Library or the British Library (Meyyappan et al. 2000; Raitt 2000; Griffin et al. 2005). Since these and many more projects were isolated efforts, the European Commission pushed towards a pan-European approach in order to create synergies by providing research funding through various research programme frameworks.

As a first step, an online consultation platform was set up until 20th January 2006 to get feedback from the relevant communities and experts (European Commission 2006:4). The replies were included into the Commission's "Recommendation on the Digitisation and Online Accessibility of Cultural Material, and Digital Preservation" published on 24th August 2006 (EE-009). The document recommended investing into the digitization of cultural heritage for online access but also for re-use by educational institutions, touristic service providers and creative industries. More importantly, it suggested creating a common multilingual access point.

> "*A common multilingual access point would make it possible to search Europe's distributed - that is to say, held in different places by different organisations – digital cultural heritage online. Such an access point would increase its visibility and underline common features*" (European Commission 2006:9).

The recommendation also encouraged to build on already existing know-how mentioning specifically the Conference of European National Librarians (CENL) and its online portal The European Library (TEL) – an online service that grants access to bibliographic records and, to some extent, to digitized items from all

European national libraries hosted at the National Library of the Netherlands (www.theeuropeanlibrary.org) (Collier 2004; Woldering 2004; Bjorner 2006). Endorsed by the European Council on 13th November 2006 (EP-001:25-30; ES-016:8), the European Parliament voted in favour of a European digital library on 27th September 2007.[23] Finally, the vision was set to build;

> "[a] *digital library that is a single, direct and multilingual access point to the European cultural heritage*" (ES-016: 3).

5.3.1 From EDL to the Europeana Prototype (2007-2009)

Since the CENL received an honorary mentioning in the EC's recommendation, it is hardly surprising that the following EU call for building a European digital library was secured by the CENL. As a first step, the CENL was expected to make the necessary preparations for building an online service. In July 2007, a so-called Thematic Network was initiated named EDLnet with a running time until March 2009 (ES-016:8) and later extended until June 2009 (EX-001:5; EE-052:3; EE-058:7; EE-008). The goal of EDLnet was;

> "[…] *to tackle the fragmented cultural heritage map of Europe by bringing on board the key European stakeholders to build consensus on creating the European Digital Library. The Thematic Network will encompass representatives from the following four cultural domains across Europe:* [libraries], *museums, archives and the audiovisual*" (EE-009:1).

The objectives were clearly defined as community building and the authoring of recommendations based on a common consensus among the traditional domains of libraries, archives and museums. The key outcome was to build a working prototype as a proof of concept (EE-050:6; ES-004:5);

> "*a multi-lingual, single portal for searching across Europe's distributed and varied cultural resources*" (EE-009:1).

The project was organized in five work groups comprising volunteering professionals from memory institutions and led by experts designated by the

[23] The documentation of the vote can be found at http://www.europarl.europa.eu/sides/getDoc.do?type=PV&reference=20070927&secondRef=ITEM-009-01&language=EN&ring=A6-2007-0296, last access: 19th January 2011.

CENL (EE-032). Due to its central role for TEL, the National Library of the Netherlands (Koninklijke Bibliotheek - KB) provided resources for project management and dedicated members of its staff. Project leadership was assigned to Jill Cousins, the Director of TEL, who later became the Executive Director of the Europeana Foundation (see Chapter 5.5.1). The project started formally with the signing of the grant agreement between the European Commission and the KB on 1st July 2007 (ES-010:3).

The second major task of EDLnet was the foundation of a legal entity with a formal organizational governance structure to represent the European digital library. Shortly after the launch of the EDLnet Thematic Network, the Stichting European Digital Library (EDL Foundation in short) was incorporated under Dutch law on 8th November 2007 (EE-051:1; ES-002). Led by an executive committee comprising representatives from major associations of the four memory domains (libraries, museums, archives, A/V archives) (ES-001:4-7), its main role has been to formalize agreements with Europeana's members as well as to build the capacity to run the online service in a sustainable fashion (EE-058:7).

After an online voting, "Europeana" was selected as the name for the digital library on 18th December 2007 – the other option being "Click2Culture" (EE-044:1; EE-046:7;21). In the following weeks, key functions were rebranded accordingly. It was a symbolic gesture to express that the initiative is truly cross-domain in nature and not only focused on libraries as an adherence to the term "European Digital Library" would have suggested. For instance, EDLnet was later renamed into EuropeanaNet and the Stichting European Digital Library has been referred to as the Europeana Foundation in short. In the meantime, work on the prototype continued and Europeana began with the ingestion of metadata from contributing libraries, archives and museums. In June 2008, the number of items reached the 500,000 target comprising material from Slovenia, France, the Netherlands, Germany, Sweden and Austria (EE-010:1). When the prototype online portal was launched on 20th November 2008, the number had grown to 4.2 million provided by over 1000 contributing organizations from every European country – more than twice as much as the initial target (ES-016:17).

The inauguration of www.europeana.eu – the domain of the online portal - took place in a very festive environment at the Palais de Charles de Lorraine in the National Royal Library of Brussels in the presence of Commission President José Manuel Barroso, Commissioner Viviane Reding, the EU Council of Culture Ministers and representatives of Europe's cultural institutions. Naturally, the representatives were very enthusiastic. Commission President Barroso, for instance, likened Europeana with a dynamo for a 21st century Renaissance while Elisabeth Niggemann, chair of the Europeana Foundation, saw European culture moving towards the Web 2.0 generation.[24] The launch was also supported by a joint conclusion of the Council of Culture Ministers expressing their strong support for Europeana not only as guardian and gatekeeper of cultural heritage but as a catalyst for innovation, economy and democratization;

> "[T]*he creation of the European digital library EUROPEANA, provides an excellent opportunity to showcase the cultural heritage of the Member States and to provide access for everyone to that heritage*, […]. [D]*igitisation and online accessibility of cultural material are essential to highlight cultural heritage, to inspire the creation of new content and to encourage new online services to emerge. They help to democratise access to culture and knowledge and to develop the information society and the knowledge-based economy*".[25]

To the surprise of system designers and engineers, Europeana turned out to be immensely popular on the first day. As it was reported extensively in the news, up to 10 million hits per hour brought the site even to crash. It was more than twice the number expected by the Europeana team. As an anonymous spokesperson was quoted by the BBC, the crash was perceived as a reassuring sign;

> "*Thousands of users were searching for the words 'Mona Lisa' at the same time.* […] *It confirms it's worth doing, European culture is more popular than we had anticipated in our wildest dreams*" (BBC 2008).

After the necessary upgrades of the infrastructure, the site was reopened to the public on 17th December 2008 and has remained online since then (ES-019:7-8).

---

[24] The official press release can be found at http://europa.eu/rapid/pressReleasesAction.do?reference=IP/08/1747, last access: 19th January 2011.

[25] The original Council Conclusion can be found at http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2008:319:0018:0019:EN:PDF, last access: 20th January 2011.

Although the prototype was praised as a success hitting or even exceeding most of its targets, it did not achieve all of them. Especially multilinguality, the translation of queries and search results (EE-002:1-2), turned out to be too big a challenge,

> "*since the resources required for seriously tackling* [multilinguality] *are substantial and certainly not available for building the first public prototype*" (ES-006:14).

5.3.2 Europeana v1.0 (2009 – 2011)

After the successful launch of the prototype, the EU continued to support the initiative financially and politically. As the next step, the prototype was to be developed into an operational and, ultimately, sustainable service. In this spirit, the European Commission called the Member States' attention to the need to increase their efforts in terms of digitization for accessibility and long-term preservation as well as to facilitate the accessibility of out-of-print and orphan works through legislative means (European Commission 2008a:9). The motion was supported by the European Parliament on 5th May 2009[26] and the Council of Education, Youth and Culture on 10th May 2009.[27] Since the conflict between copyright and digitality within the public domain, in general, and within the institutional field of memory, in particular, has already been discussed in Chapter 2.4.4, this chapter will not go into further details at this point.

The issue of digitization, on the other hand, addressed the necessity to increase the sheer volume of accessible digitized items in order to reach a critical mass that would attract online users. By the same token, the European Commission also complained about the representativeness of the Europeana collection since the majority of the Member States were underrepresented – a problem that has not been resolved until today (see Table 12). Hence, the need for better

[26] The documentation of the parliamentary session can be found at http://ec.europa.eu/information_society/activities/digital_libraries/doc/parliament/resolution_europeana.pdf, last access: 21st January 2011

[27] The press release can be found at http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/educ/114361.pdf, last access: 21st January 2011.

coordination among the Member States was stressed in order to prevent situations;

> "*where classics from Europe's literature are available through Europeana in a range of languages, but not in the language of origin. For example, you will find the works of Goethe in French, Polish and Hungarian, but not in German*" (European Commission 2008b:6).

All these aspects, technological and organizational, were to be addressed in the phase following EDLnet. As a direct successor, a new Thematic Network – called Europeana v1.0 - was launched in February 2009 with a running time of 30 months (EP-002:1). Its main targets were to develop back-end processes, a business model, end-user awareness and services allowing re-use and re-purposing of Europeana's data (EE-058:4; EE-050:6-7; EX-001) – aspects deemed necessary for a sustainable service. In addition, the number of accessible digital artefacts was set to reach 25 million items by 2014 (ES-016:21).

Organized in six work packages, Europeana v1.0 planned two releases each implementing incremental updates to the online portal www.europeana.eu. The first release, called Rhine release, was soft launched over summer and autumn of 2010. The most visible changes were the introduction of new browsing functionalities such as curated virtual exhibitions, a timeline or new search functionalities such as the auto-completion of search terms and phrase searching (EE-064; EP-002:14-15). The second release, called Danube release, was implemented in April 2011 (EP-002:1; ES-022:8; EP-005:20; ES-025). It brought improvements in terms of search results ranking, browse functionalities via map- and timeline-based interfaces, iconography, interoperability with translation services such as Google Translate, accessibility features for visually impaired users, social tagging, re-use of Europeana data, data enrichment, semantic contextualization and many more (ES-022).

As a complementary project, Europeana also launched a so-called Best Practice Network named EuropeanaConnect coordinated by the Austrian National Library running from May 2009 until October 2011 (ES-023). Defined as the main technology provider to Europeana, it is developing further enhancements but also future core functionalities (ES-024:2). Organized in seven work packages,

EuropeanaConnect focuses on multilingual search and browsing, semantic enrichment of digital content, provision of an interface for mobile devices, integration of annotation tools, inclusion of audio and its metadata into the Europeana information space and the deployment of infrastructural components such as large-scale metadata harvesting or a resolution service for unique resource identification (ES-023:2). As EuropeanaConnect was still running during the authoring of this dissertation, most of the features and improvements were still under development.

Although Europeana is still under construction, it has already received considerable acclaim and, hence, expectations are quite high. In October 2009, for instance, Europeana received the "Erasmus Award for Networking Europe" from the European Society for Education and Communication – an organization dedicated to the promotion of a common European cultural area. Based on the accomplishments at that time, the laudation highlighted Europeana's contribution towards "*the development of a common European consciousness*."[28] Europeana is not only about keeping cultural heritage and its institutions relevant in the 21st century; it is recognized as a key instrument towards the creation of a European identity.

The continuing support from the Council of Ministers, the European Parliament and Neelie Kroes, the new Commissioner for the Digital Agenda, indicates that the future of Europeana is secure for years to come. Europeana is slowly maturing and its potentials for further developments are being recognized by non-experts as well. In particular, the European Commission will continue to finance Europeana until 2015 with over 30 million items made accessible via the online service (EP-004). In 2025, it is to be "*all of Europe's digitized cultural heritage*" (EP-004:5). However, in order for Europeana to become a sustainable service, it is also expected to generate revenue by charging 3rd party industries (search engines, semantic operators, advertisers etc.) for access to its data and users. The business model is not decided yet but, since it is not relevant for answering the research question, it is not of concern for this research.

[28] The laudation can be found at http://www.euromediaawards.eu/europeanaeu-internet-site/, last access: 24th January 2011

To conclude, Europeana has become more than a mere prestige project of showcasing the treasures of European heritage. As discussed in Chapter 4.3.3 and demonstrated in this historical section, the development of Europeana can be split into two cycles. Cycle 1 was still in line with the "traditional" digital library paradigm and culminated into the launch of the prototype. Cycle 2, lasting from the launch of the prototype until roughly the end of 2011, has been introducing a lot of changes to that paradigm – some of them quite fundamental. In the next chapter, a detailed analysis will illuminate this shift based on the comparison of these two cycles. In order to provide for a clear framework, the resulting online service of cycle 1 – the prototype – will be referred to as EDL (the European Digital Library) in order to highlight its affiliation to the "traditional" digital library paradigm. As a contrast, the resulting online service of cycle 2 will be referred to as Europeana subsuming the developments achieved through the Thematic Network Europeana v1.0 as well as the Best Practice Network EuropeanaConnect. Although cycle 2 has not come to its completion during the authoring of this dissertation, the long duration of the Europeana initiative as a whole as well as the concrete and precise definitions of the services yet to be implemented provide for strong evidence as to how Europeana will evolve into a new breed of digital memory institution.

5.4 Private-Public-Partnerships

Besides the actual implementation of a "European digital library" covered by the various projects outlined above, the European Commission also addressed issues relevant for the institutional field of memory and cultural heritage at large. In this respect, the so-called High Level Expert Group (HLEG) on digital libraries was a key platform. Chaired by Commissioner Viviane Reding, it brought together representatives of cultural institutions, publishers, technology firms and academia in order to discuss the controversial topic of copyright and the related issue of private-public-partnerships (HLEG on Digital Libraries 2009). From its

inception on 27th February 2006[29] to its conclusion in December 2009, the platform invited and collaborated with associations of memory institutions but also with copyright holders such as the International Federation of Reproduction Rights Organisations (IFRRO), newspaper associations such as the European Newspaper Publishers' Association (ENPA), academia and research such as the European Science Foundation (ESF) and online service providers such as Google and the Internet Archive (for a complete list see HLEG on Digital Libraries 2009:11-15). As this selection of members shows, the HLEG aimed at reaching a broad consensus among key players involved in the public domain and digital cultural heritage.

As the key outcome, the HLEG proposed a set of due diligence guidelines when searching for copyright holders in the "Memorandum of Understanding on Orphan Works"[30] expressing the willingness to clarify the copyright status of a work first, before making it accessible online. A clear statement against Google's "digitize first, ask later" policy (see Chapter 2.4.4). With respect to out-of-print works, the expert group supported full online access. However, the clarification of the copyright status is a very costly process. To give a rough idea, it took a Dutch digitization project of history handbooks five months to clear 50 books (HLEG on Digital Libraries 2009). Hence, the due diligence approach has clear limitations when it comes to the mass digitization of millions of items – a limitation recognized by the HLEG. In order to enable synergies across nation borders and memory institutions, the implementation of "data-base and rights clearance centres" for orphaned and out-of-print works was suggested. In this spirit, the European Commission financed the Accessible Registries of Rights Information and Orphan Works (ARROW, www.arrow-net.eu).

It is not only the clearance of works that is expensive but also the digitization thereof. Hence, the HLEG supported the formation of private-public-partnerships (PPPs) in order to cut costs and to build on the know-how and expertise of

[29] The original decision of the European Commission can be found at http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:063:0025:0027:EN:PDF, last access: 24th January 2011.
[30] The original memorandum can be found at http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/orphan/memorandum.pdf, last access: 25th January 2011.

technology companies. Still, it seems that the expert group also felt the need to protect the public domain from privatization tendencies and to prevent companies from claiming the rights for digitized copies of public domain works. The goal needs to be to grant non-exclusive access. If that is not possible, exclusive access needs to be limited in time at least.

> "[P]*ublic domain content in the analogue world should remain in the public domain in the digital environment*" (HLEG on Digital Libraries 2009:8).

To continue the work of the HLEG, the European Commission installed a reflection group, a Comité de Sages, in April 2010. Comprising Elisabeth Niggemann (Director General of the German National Library and Chair of the Europeana Foundation), Jacques De Decker (author and Permanent Secretary of Belgium's Royal Academy of French Language and Literature) and Maurice Lévy (Chairman and Chief Executive Officer of the advertising and communications company Publicis), the reflection group set out to formulate further recommendations with respect to the issues of copyright, sustainability and private-public-partnerships. In this spirit, the group organized public consultations and hearings as well as commissioned studies and audits over a period of 8 months. The basic tone of their final report, published 10th January 2011, reflected very much the attitude of the HLEG.

> "*Culture is following the digital path and "memory institutions" are adapting the way in which they communicate with their public.* […] *We are of the opinion that the public sector has the primary responsibility for making our cultural heritage accessible and preserving it for future generations. This responsibility for and control over Europe's heritage cannot be left to one or a few market players, although we strongly encourage the idea of bringing more private investments and companies into the digitisation arena through a fair and balanced partnership*" (Niggemann et al. 2011:5).

The report also stresses that the digitization of an item does not change its copyright status. Hence, material from the public domain should remain in the public domain when digitized. The group specifically laments that online access to the collections of some publicly funded memory institutions is not free of charge. As a response, the group strongly stated the following;

> "[P]*ublic domain material digitised with public money should be freely accessible for all. This should be part of the funding conditions for the*

> *digitisation of public domain material across Europe*" (Niggemann et al. 2011:13).

If a memory institution wants funding from the European Union, it needs to grant non-exclusive accessibility. However, to create an incentive for private enterprises, the group also recommends a 7 year limit for preferential use in order to allow for a return of investment. In contrast, Google's standard agreements signed with the first libraries joining Google Books granted exclusive rights to the digitized material for 15 years (see Chapter 2.4.4).

A telling example for the potential of PPPs is the cooperation between the Austrian National Library (ANL) and Google Books announced in June 2010. The project will digitize and grant online accessibility to ANL's collection of 400,000 books from between the 16th and 19th century, all belonging to the public domain.[31] The collection contains treasures such as the first translation of the Old and the New Testament by Martin Luther and the works of the alchemist Paracelsus. Costs and workload are split between the two partners. While Google covers the digitization and OCR scan of the books, the ANL is responsible for the preparation of the books and its metadata as well as for long-term preservation. Access will be granted via Google Books and Europeana in a non-exclusive fashion. Hence, the digitized books will remain in the public domain.[32]

In this complex situation of copyright and digitization, public domain and profit interests, Europeana itself starts to play an influential role as a lobbyist for the member organizations and its users (ES-005:6). In the "Europeana Public Domain Charter" from April 2010, Europeana declared that it;

> "*belongs to the public and must represent the public interest*" (EP-003:1) and, therefore, needs "*to ensure that works in the Public Domain are accessible to all of society, by making them available as widely as possible*" (EP-003:4).

[31] For further information visit http://www.onb.ac.at/bibliothek/austrianbooksonline.htm, last access: 26th January 2011.
[32] For further information visit http://www.onb.ac.at/austrianbooksonline/18689.htm, last access: 26th January 2011.

In this spirit, Europeana implemented the creative commons licensing scheme[33] to mark the extent to which cultural heritage artefacts found on its portal can be used, re-used and re-purposed. It was also the first major adopter of the Public Domain mark under creative commons licensing, which was proudly announced jointly by Europeana[34] and the Creative Commons[35] in October 2010.

To conclude, the HLEG as well as the Comité de Sages convey a message of urgency. They call upon the EU Member States and the European Commission to drastically increase their efforts in digitizing cultural heritage artefacts as well as to harmonize the respective legal frameworks in terms of in-copyright, orphan and out-of-print works. Most of in-copyright works will be or already are out of distribution and/or orphaned. If these issues are not addressed and the procedure to clarify the copyright status streamlined, the majority of last century's and contemporary cultural heritage may fall into "*a black hole*" (Niggemann et al. 2011:9). That is, most of contemporary cultural heritage, if not published online, may never be seen by large parts of the population.

5.5 Organizational Structure

The vision of building a "European virtual library", as it was initially referred to by the European Commission, has lead to a variety of different projects that are either directly involved in or contribute towards the building of Europeana. Hence, the term Europeana, in general, refers to a whole group of projects.

As Figure 8 shows, the two main projects are Europeana v1.0 (the successor of EDLnet/EuropeanaNet) and EuropeanaConnect discussed above. They are responsible for the creation of the operational service accessible via www.europeana.eu. In addition, they also coordinate the research and

[33] The creative commons licensing scheme is an alternative to the "all rights reserved" copyright framework deemed to be more appropriate to the online culture of sharing, mash-ups and universal access (www.creativecommons.org).
[34] See, for instance, http://group.europeana.eu/web/guest/news/-/blogs/europeana-and-creative-commons-launch-new-public-domain-mark, last access: 26th January 2011.
[35] See the press release at http://creativecommons.org/press-releases/entry/23755, last access: 26th January 2011.

development efforts of the group and formulate standards and models. In this sense, they are the core projects and define the case of the case study analysed in the next chapter.

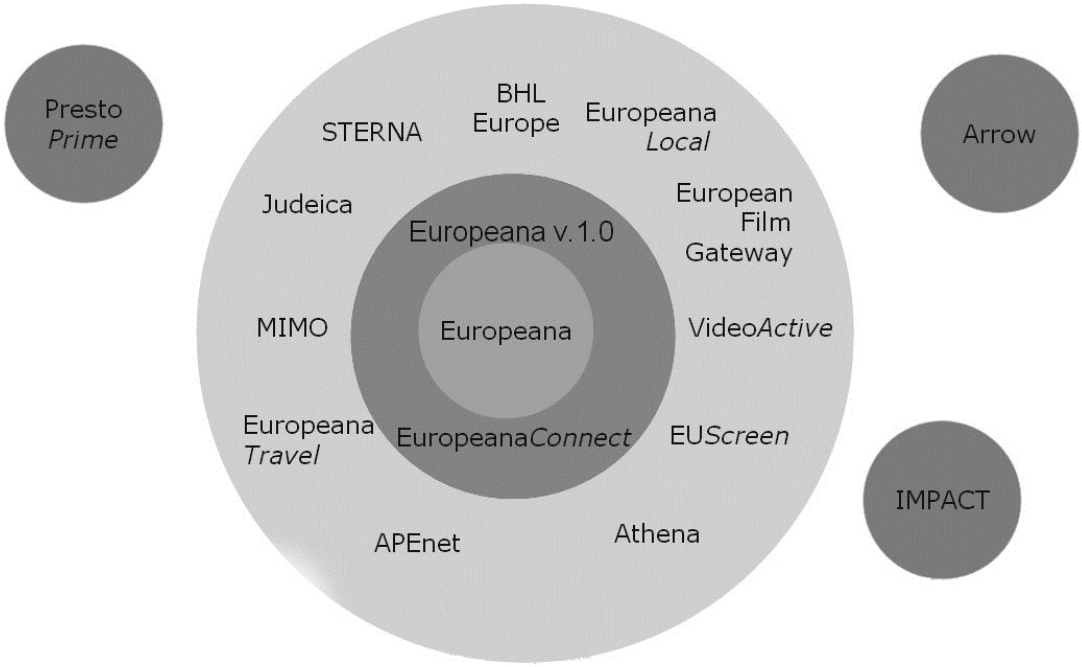


Figure 8: Selection of projects part of or complementary to the Europeana Group of Projects (EP-068:5).

The other projects, of which a selection is displayed in the outer circle of Figure 8, are mostly platforms for collaboration based on specific themes or domains. Most of them also function as aggregators collecting material from their respective domain to be ingested into the Europeana repository (EP-002:12). They are involved as content providers. For instance, APEnet is the Archival Portal Europe focusing on national archives (www.apenet.eu). ATHENA focuses on museums (www.athenaeurope.org). Further projects range from television (www.euscreen.eu) and cinema (www.europeanfilmgateway.eu) to musical instruments (www.mimo-project.eu), travel and tourism (www.europeanatravel.eu) or biodiversity (www.bhl-europe.eu) – i.e. natural history collections and botanical gardens. The list is getting longer and longer as Europeana moves forward.[36]

[36] For an up-to-date list visit http://version1.europeana.eu/web/guest/home, last access: 27th January 2011.

Finally, ARROW, PrestoPrime and IMPACT do not provide content but rather support the development of complementary functionalities and applications. These three projects are referred to as content enablers as opposed to content providers. ARROW is the project towards Accessible Registries of Rights Information and Orphan Works that should streamline the clearance of the copyright status of works already discussed in the previous section. PrestoPrime (www.prestoprime.org) is an excellence centre for the long-term preservation of A/V archives of mostly broadcasting companies. IMPACT (www.impact-project.eu) is a research project to improve access to text through mass-digitization and advanced OCR scanning.

The Europeana initiative covers a huge spectrum of topics and issues related to the field of digital memory institutions at large. The content contributors named above play a key role as aggregators – that is, as interfaces between individual organizations and Europeana (EE-003:1). The aggregators as well as the organizations being aggregated can be categorized into four domains; libraries, archives, museums and A/V archives (e.g. broadcasting companies like the BBC). For instance, The European Library (TEL) is an aggregator for European national libraries. Some aggregators, however, are oriented towards a specific nation and, thus, cross domains. Examples are Heritage Malta (www.heritagemalta.org) or www.culture.fr run by the French Ministry of Culture and Communication.

Most libraries, archives, museums and A/V archives contribute their content to an aggregator, which, in turn, adapts the material and its description (descriptive metadata) to pass it on to Europeana (ES-001:4-6). In this sense, aggregators standardize or rather normalize the content of their respective members and perform as a filter in order to make the ingestion of the content into the Europeana repository more effective and efficient (EE-033:5; ES-001:15). Large organizations, such as the British Museum, have the necessary resources in terms of know-how and personnel to contribute directly to Europeana without the mediation of an aggregator. Smaller and local organizations, neither being part of an aggregator nor having the resources to join Europeana directly, are supported

by a project called EuropeanaLocal (www.europeanalocal.eu) in adopting Europeana standards for interoperability.

Due to legal reasons and limitations in storage capacity, Europeana itself does not host copies of the digitized cultural heritage artefacts themselves but only metadata about it. However, as it will be argued further below, the distinction between description and the described, metadata and artefact is fuzzy and difficult to draw. Still, from this perspective, Europeana can be seen as a meta-aggregator (ES-020:8) providing an "*online end-user service that operates on a European scale*" (EP-002:14) as well as a coordinator of mostly EU sponsored projects (ES-005:7). The content is delivered based on a hybrid organizational model of domain and cross-domain aggregators as well as individual organizations (EE-033:5; ES-001:13-14).

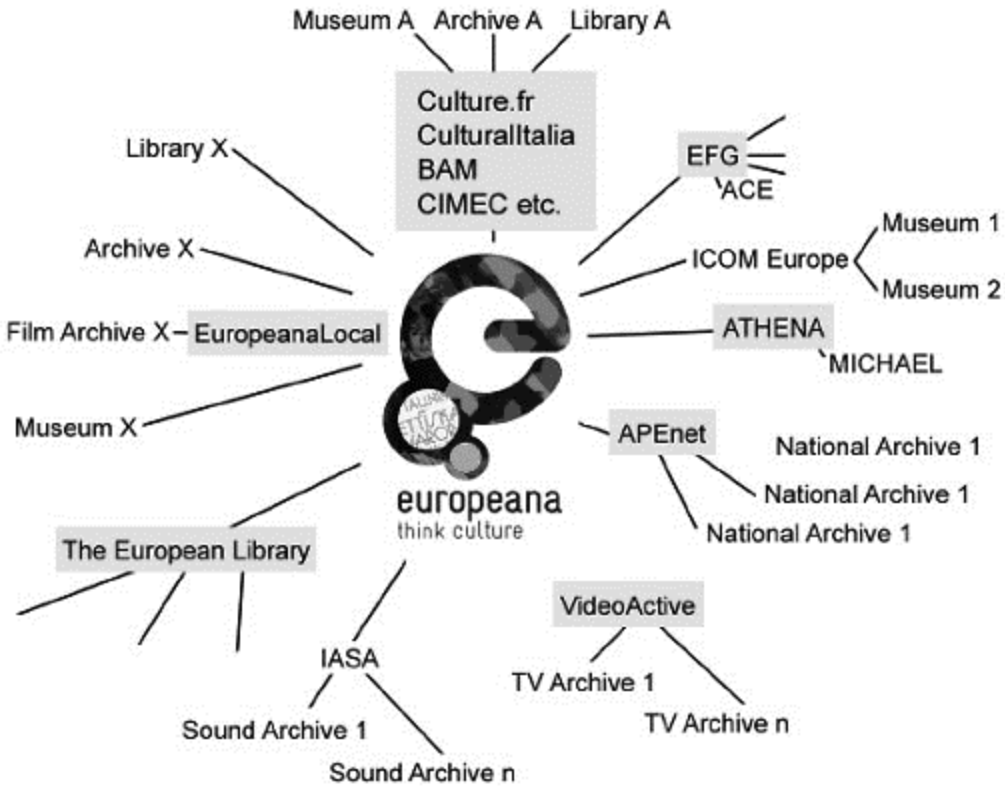


Figure 9: Organization of content providers (ES-011:24).

5.5.1 The Europeana Foundation

The Europeana Foundation[37] or EDL Foundation is the legal entity responsible for the Europeana services (ES-002). Incorporated under Dutch law and hosted by the Koninklijke Bibliotheek in The Hague, it is the governance and strategic decision-making entity bringing together representatives from major European associations of the memory institutional field and leading organizations as stakeholders. Hence, it is the Europeana Foundation that signs the formal agreements with the contributing libraries, archives and museums or their respective associations (ES-002:1-2).

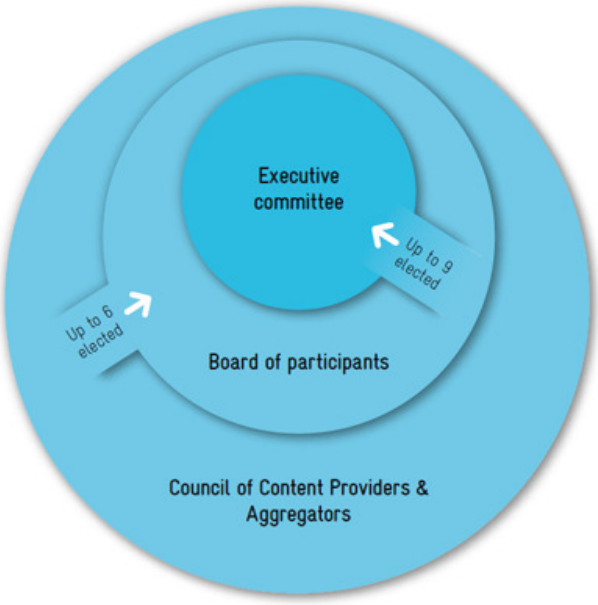


Figure 10: Europeana Foundation organizational structure (EP-005:5).

The foundation is led by an executive committee responsible for representing, budgeting, accounting and strategy while maintaining the right to change the articles as well as to wind up and liquidate the foundation (ES-002:7-8; EE-058:11). Chaired by Elisabeth Niggemann, Director General of the German National Library and representative of the Conference of European National Librarians (CENL), the committee also comprises the National Library of the Netherlands, the European Branch of the International Council on Archives (EURB-ICA), the International Federation of Television Archives (IFTA/FIAF),

[37] The website can be found at http://group.europeana.eu/web/europeana-foundation/introduction, last access: 28th January 2011. The statutes can be found at http://group.europeana.eu/c/document_library/get_file?uuid=8493d624-7b32-4a6a-8a41-0c765922874e&groupId=10602, last access: 28th January 2011.

the Association Cinémathèques Européennes (ACE), the European Museum Academy (EMA), the Ligue des Bibliothèques Européennes de Recherche (LIBER) and the Multilingual Inventory of Cultural Heritage in Europe (MICHAEL) (ES-011:8; ES-005:7).[38]

As Figure 10 shows, the Executive Committee is elected from the Board of Participants; those are European associations or the European branches of international associations of the four domains, individual organizations that were founding members of the Europeana Foundation and the six officers of the Council of Content Providers and Aggregators.[39] The Board of Participants' role is to deliberate by majority vote on resolutions formulated by the Executive Committee. At the bottom, the Council of Content Providers and Aggregators represents the member base. Any organization that wants to provide content and/or metadata to Europeana is welcome to join (EE-069:2). At this moment, there are 162 members ranging from aggregators and individual memory institutions to universities and research institutions (EP-005:5).[40] However, if one would include the organizations represented by aggregators, the count would go beyond 1500 (EP-004:5).

5.5.2 The Europeana Office

The last part of the organizational structure is the Europeana Office, which is basically responsible for the day-to-day operation of the online services including marketing, business development, administration, technical operations, data ingestion and development.

Headed by the Executive Director Jill Cousins, the office hired 28 full-time employees and 7 freelancers (see Figure 11). The employees were also asked for interviews in the third phase of the research in order to accomplish

[38] The up-to-date list can be found at http://group.europeana.eu/web/europeana-foundation/executive-committee, last access: 28th January 2011.
[39] The up-to-date list can be found at http://group.europeana.eu/web/europeana-foundation/board, last access: 28th January 2011.
[40] The up-to-date list can be found at http://group.europeana.eu/web/europeana-foundation/content-council-members, last access: 28th January 2011.

communicative validation of the research results and conclusions (see Chapter 4.2).

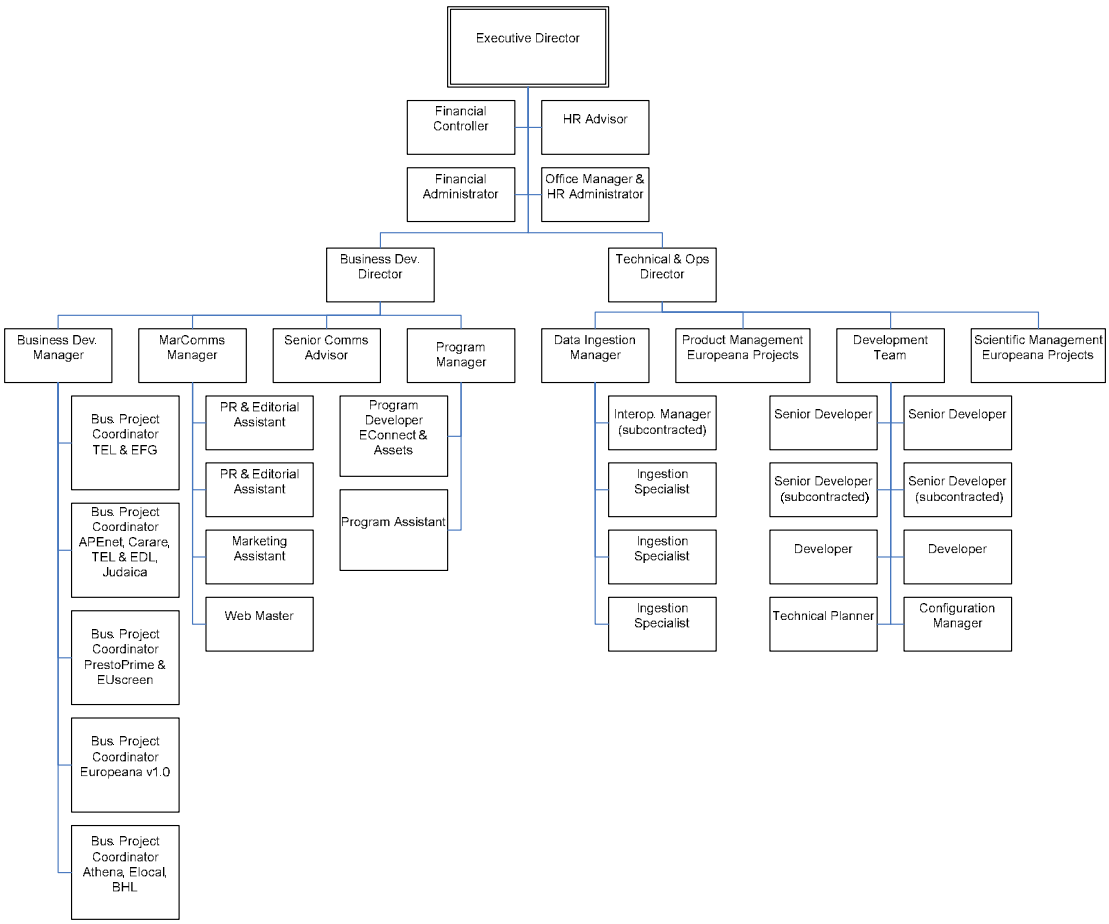


Figure 11: Europeana Office organization chart (EE-061).

5.6 Conclusion

The vision of a European digital library was born out of a political reaction against Google's mass-digitization of books. However, as will be discussed in more detail in the following chapter, the resulting online service turned out to be quite different from the initial goal to build a "traditional" digital library. From a historical perspective, the breaking point was the launch of the prototype as proof that Europeana can indeed build a working, though rudimentary, online portal for European cultural heritage. During the years that followed, Europeana has become a consolidated organization with its own resources and structures as well

as an influential player within the digital domain as exemplified by Europeana lobbying for the public domain.

Still, the developments described in this chapter do not help in answering the research question; that is, how does Europeana become a memory institution within the information habitat? The key to exploring this question lies within the paradigm shift in the ways cultural heritage artefacts are being ordered - a shift from catalogue to metadata, from categorization to contextualization. This will be the topic of the following chapter.

6. CASE STUDY – ANALYSIS

6.1 Introduction

The following analysis will explore in detail the paradigmatic shift in the way Europeana orders the collection of cultural heritage artefacts – a shift from an ex-ante static order to an ex-post ordering based on contextualization instead of categorization. As it was discussed in Chapter 2.4.1, ordering is one of the key social practices of forgetting especially within the context of canonization and institutionalized social memory. The clearest approach to demonstrating that transition is the comparison of the two cycles defined in Chapter 4.3.3 on corpus construction. The first cycle refers to the EDLnet/EuropeanaNet project dedicated to the launch of a prototype that was still firmly established within the traditional, portal-based digital library paradigm. By contrast, the second cycle refers to the Europeana v1.0 and EuropeanaConnect projects driving the development towards a semantic web user interface interoperating with other information service providers of linked data through APIs. In order to prevent unnecessary confusion, the term EDL will be used to refer to the time period of cycle 1 and Europeana to refer to the time period of cycle 2.

The analysis begins with a short excursion into the scenarios of use and the user classifications applied by Europeana. The way Europeana observes and constructs its users is closely related to the ways it orders its cultural heritage artefacts since they are both linked to the topics of navigability and accessibility. The first section on scenarios of use, therefore, is an attempt to tell the whole story by referring to the changes in the ways users are perceived and categorized by Europeana in order to give an overview over the thematic map that emerged as a result of thematic coding.

The sections that follow will retell the story in a much more detailed fashion starting with the analysis of how cultural heritage artefacts change when they are digitized. From this starting point, the analysis will move on and describe how those digital cultural heritage artefacts are represented by so-called Digital

Surrogate Objects (DSO), which form the basic building blocks of the Europeana information space. Finally, the analysis will demonstrate the immersion of the Europeana information space into the wider information habitat through Linked Open Data (LOD) (see also Chapter 2.4.3). In the conclusion, the comparison between EDL and Europeana will be revisited in order to provide an analytical foundation for Chapter 7 on interpretation and discussion.

6.2 The Transition in Scenarios of Use

Probably the best way to describe the transition from EDL to Europeana is to begin with the project's classification of its user groups and, thus, the construction of scenarios of use for Europeana's repository. EDLnet's main target was to create the prototype platform www.europeana.eu as a single point of access to the cultural heritage artefacts of the content providers – libraries, archives and museums providing digitized items and metadata for indexing. From this perspective, the users can be grouped into end-users and re-users. While end-users are basically individuals using the services provided at the online portal, re-users work with the underlying data and metadata structure in order to provide services of their own.

User category	Search skills	Domain knowledge	Motivation to use Europeana	Attraction
General user	Basic	None	Browsing	Large volume of content, easy to use
School child	Basic	None	Exercises, homework	Visually attractive, easy to use
Academic user	Basic/Advanced	Field/Discipline of study	Studying, assignments	Trusted material, referencable, representative
Expert researcher	Advanced	Field/Discipline of research	Research	System functionality
Professional user (librarians, curators, archivists)	Advanced	Memory institutions	Improve own service	Metadata

Table 13: User profiles (EE-011:6; ES-007:15; ES-012:7-8; ES-020:5-6).

EDLnet was focused on the creation of an online interface tailored to the perceived needs of the users of the services. This was done very much like the way one would expect a library to manage the usability of its repository or any online service provider for that matter, that is, by creating user groups based on the expected motivations and skills in terms of information retrieval (Arms 2001).

A complementary approach towards analysing end-users was to categorize them according to the objectives of using the system. These were defined as;

The user wants to be entertained; a. Users who have time available to browse around the Internet; b. A structural or incidental interest in cultural heritage; c. Expect a lot of interesting content; d. For these users it is not important what they find as long as it is interesting and entertaining.	**The user wants to know more about a cultural or historic subject or person;** a. Users have a specific reason for their interest (study, work, current news etc.); b. Users looking for the most relevant results and would not want to see lots of results that are not relevant to them; c. Able to determine what is relevant to them, information about the specific objective of the user is necessary.
The user wants to know the current whereabouts of cultural heritage; a. Users planning to see the original objects for research purposes; b. Users that are about to undertake a trip and would like to know what cultural heritage they can visit during a touristic trip or other type of stay; c. These persons will also be interested in getting more information on interesting events and collections in the area, as well as local services such as guided tours.	**The user wants to be part of a community of interest;** a. Users want to share their knowledge via an online environment such as a social platform with a cultural focus; b. They may want to present their opinions and ratings of cultural heritage resources to their kin as well as share personal items (photographs, documents etc.).

Table 14: User motivations (ES-020:5-6).

The combination of the user profiles (Table 13) and the user objectives (Table 14) were then used to define scenarios of use of the online portal.

> "*One thing all end-users have in common is that they want access to the Europeana fullcontent* [!] *through search and browse and direct surrogate addressing and access options. Which option (or configuration of that option) they will choose depends on the user profile (general user, school child, researcher) and objective (entertainment, research, community building)*" (ES-020:6).

In other words, EDL was closely linked to the notion of a portal offering discovery services for and accessibility to digitized cultural artefacts. Just as one would expect from an online library catalogue, the scenarios of use were primarily defined as discovery and access. As already discussed in Chapter 2.4.1, these are basically the objectives of a library catalogue;

> "*1. find a known object* [and] *2) browse surrogates having a known property (e.g. subject)*" (ES-012:36-37).

The fourth motivation of Table 14 – community building – reflects the involvement of Web 2.0 services as a way to accommodate internet users who expect to find similar functionalities they are accustomed to in the online world.

> "*It is not only about making contents accessible, it is also about making them accessible in a Web 2.0 environment which opens completely new possibilities of use than those to which users are accustomed* [with regards to a library, archive or museum]" (ES-014:5).

As a user survey conducted in May 2009 (EP-002:5-7) shows, internet users expect interactive and personalized features, which can be shared with other users – features such as social tagging, user oriented annotations or create-your-own-virtual-exhibition. However, this requires the opening up of the closed silos of cultural heritage traditionally administered by dedicated professionals. In detail, the EDLnet project addressed the question of openness through three options for end-user involvement (ES-014:11);

1. **Keeping Europeana closed but more personalized;** a. Easy to use, setting of preferences (for instance how search results are presented); b. Enabling services such as ordering a reproduction and making reservations for a service set in a physical environment (rerouted to the Europeana partner concerned).
2. Opening up Europeana; a. Users helping in building and maintaining Europeana; b. Enriching the content by importing or suggesting related items or sites; c. Suggestions for improvement (having spotted an error for instance in an OCR scan or provided links); d. Social tagging; e. Enriching the content by recommending, commenting, annotating etc.; f. Enriching the content by shared editorial working processes (wiki).
3. **Selective;** a. Reserve some personalized services to experts or other particular groups of users.

Table 15: Options for end-user involvement (ES-014:11).

As one can see, the first two options present two extremes of a continuum unfolded between a traditional service managed by professionals and users fully involved as amateur librarians or curators reminiscent of open platforms such as Wikipedia. The decision was to go for the middle way by creating a personalized space for the users – called MyEuropeana – allowing the saving of items and their sharing, for instance, via Facebook, Twitter or email, complemented by the tagging of these items as well as the saving of search results. Hence, Europeana provided a total of four scenarios of use;

1) Discovery (search and browse);
2) Access to the items themselves;
3) Personalization;
4) User oriented editorial and sharing (social tagging).

Further portal-based scenarios of use, to be implemented in future versions of the online service, include features such as collaborative workspaces and linking between users (ES-012:8-9). An unresolved issue is still the topic of user generated content (UGC);

> "*An open question still remains the possibility for users to upload content, either in certified contexts or in specific publication spaces. This possibility is far from being possible today, however it is an important orientation for the future since there is a very important demand from the part of users to contribute either to enrich existing contents* [or] *to add new contents in relation to existing ones. In any case, in order to keep the high cultural value of Europeana, incoming contents and annotations should be validated and checked concerning their ownership. Europeana is based on certified digital repositories from the four domains; certification has to be kept as a high ambition to prevent inappropriate content to be added to the certified ones*" (ES-014:11).[41]

Be as it may, all the scenarios of use introduced or planned during EDLnet were based on the notion of a user coming to the portal to use the offered services and functionalities. In other words, even the Web 2.0 based scenarios were to a considerable degree under the control of Europeana since most of them are hosted at www.europeana.eu.

[41] For a demonstration and minimal prototype visit http://dme.arcs.ac.at/annotation/, last access: 1st March 2011.

Complementing the generic end-user, the second user group can be referred to as the "re-user". The last row of Table 13 – the professional user – already addresses the possibility for librarians, curators and archivists to access the repository of descriptive metadata directly in order to use them for services of their own. From this perspective, the content providers (be they content aggregators or individual organizations) do not only ingest their content into the Europeana information space for metadata enrichment but also re-use the enriched metadata for their own discovery and access functionalities.

By the same token, search engines can be seen as another type of re-user. Since Europeana is working on the search engine optimization of the metadata created for each cultural heritage artefact, search engines will be able to crawl the metadata directly and rank these according to their respective algorithms (EE-064:1-4). From this perspective, search engines are an alternative scenario of use to the portal's search and browse functionalities for the discovery of cultural heritage artefacts hosted by Europeana. Instead of conducting a search at the online portal, a user may just as well stumble upon a cultural heritage artefact via Google or Yahoo! The search engine offers a path to discovery equivalent to Europeana's search functionalities.

With the Europeana v1.0 project, the possibilities for re-use, which were only open to professional users, are in the process of being made available to practically everybody. This is to be achieved through the provision of application programming interfaces (APIs), which basically give everybody with the necessary IT skills access to the Europeana data-bases in order to create new scenarios of use beyond the Europeana information space. Initially considered to be of a lesser concern, APIs turned out to be the core topic that would change the way Europeana defines its own role as an online service provider;

> "*Application programming interfaces (APIs) represent one of the initially unsuspected issues for Europeana. The Portal is indeed a huge collection of contents, in different formats, with different access authorizations and from different collections. At the same time it is a unique possibility to have so many different contents in the same location, permitting a transversal approach to contents never before available. This will permit institutional users initially and the general users later on to bring together contents in new applications on different platforms. A Museum*

> [!] *in one country may make a European exhibition on prehistoric objects from all the available collections on their website, based on the Europeana portal and respecting the access rights of every content owner*" (ES-014:12).

The aim is, indeed, to bring the huge mountain of content to the user rather than to expect the user to come to Europeana. Hence, it should be possible for a variety of online service providers (e.g. social network sites, educational platforms or touristic portals) to embed Europeana's collection into their websites. Europeana's underlying vision is to achieve a paradigm shift in terms of what services a memory institution is supposed to deliver as well as how they are to be delivered.

> "[W]*e suppose that instead of trying to sustain the digital information silos of the past, cultural heritage communities are ready for an information paradigm of linked data and thus for sharing as much semantic context as possible. Only in such a mental setting does the shift from the portal paradigm to the vision of an API as Europeana's primary incarnation truly make sense*" (EX-002:8-9).

A truly remarkable development, since Europeana is opening up the silos of memory institutions and, as a result, looses quite a degree of control over the ways cultural heritage will be approached, processed and accessed as well as, most importantly, contextualized and categorized.

> "*This implies giving up some autonomy: the very idea of 'control' becomes obsolete to some extent that way – but the gain in functionality and rich context will be considerable and – above all – this step makes Europeana part of a much larger community and in a way simply an integrated part of the WWW, the biggest interoperability framework the world has ever seen*" (EX-002:9).

In addition to the scenarios of use offered on the portal www.europeana.eu, APIs will enable the creation of new scenarios of use beyond the control of Europeana.

> "*The idea thus is not to pre-aggregate information in fixed structures for basically static reuse, but to make it available together with functional primitives for usage scenarios not exclusively defined by Europeana*" (EX-002:7).

Hence, third parties from non-heritage sectors will be able to create mash-ups for innovative services (EE-034:2). TripAdvisor, for instance, would be able to offer virtual tours by mashing up information about and images of museum artefacts

from Europeana's repository with Google Maps. Educational platforms could create interactive learning materials. Users could link the metadata with a Wikipedia article and, thus, use it as authority files (EE-046:17-19). The potential for re-use is enormous and far-reaching.

To conclude, the distinction between end-user and re-user is becoming increasingly fuzzy. Admittedly an already well known phenomenon, it also reaches the quite strictly regulated domain of institutionalized social memory. The notion of an end-user within the domain of memory institutions makes only sense in reference to a catalogue. The user uses a discovery system based on professional practices and conventions allowing for the identification, collocation and evaluation of cultural heritage artefacts – the objectives of a library catalogue (see Chapter 2.4.1). The procedures to be followed are pre-defined and fixed. In the terminology used in this section, the scenarios of use concerning searching and browsing for cultural heritage artefacts are defined by the memory institution holding the catalogue.

The portal mostly developed by the EDLnet project is very much embedded within this paradigm. It is a meta-catalogue with the objective of mediating access across the traditionally separated domains of libraries, archives and museums. More precisely, it identifies, collocates and evaluates cultural heritage artefacts based on a standardized way of describing books, paintings, documents, newspapers, films and so forth. Neither the ability to use advanced search parameters, such as Boolean Operators, nor the ability to access and use the digitized version of a cultural heritage artefact itself, for that matter, changes the fact that the scenarios of use these functionalities afford are defined by Europeana.

However, with the introduction of APIs the situation changes from a focus on using the discovery system and the digitized object to a focus on using metadata – descriptions of cultural heritage artefacts. From this perspective, Europeana is not the exclusive provider of scenarios of use anymore. There is no predefined order according to which identification, collocation and evaluation is offered to a user. In other words, there is no underlying rationale – no "kata logos" (see

Chapter 2.4.1). The rationale is brought by the user who arranges the metadata according to her own order. As will be discussed in the following sections, the disassembling of the catalogue relies on a redefinition of what is referred to as surrogate objects in the domain of memory institutions. These surrogate objects or functional primitives, as it was called in a quote above, are basically digital objects based on relationships between metadata elements. Hence, the transition from EDL to Europeana is a transition from the paradigm of a meta-catalogue to the paradigm of linked metadata.

This section provided an overview of the thematic map (see Appendix 9.3) developed throughout the research from the perspective of the user. The outline of the changes the scenarios of use went through addresses the core theme of the dissertation and should be seen as a means to convey the main points of the analysis. However, due to its introductory nature, it was superficial and imprecise in terms of the terminology used and the details provided. This imprecision will be remedied in the following sections, which will focus on what digital cultural heritage artefacts are and what Europeana is actually doing with them.

6.3 Digital Cultural Heritage Artefacts

In order to provide for persistent accessibility a memory institution does not only have to tend to the usability of its discovery system but also to the integrity and, above all, authenticity of its cultural artefacts. In terms of digital objects, accessibility depends just as much on the software standards used to create a document as on the integrity of the data stored. For a computer file to be displayed in the correct way, instructions on how the binary code is to be interpreted by software need to be provided. This is not the case with, say, books that can be read as they are. These instructions blur the initially clear distinction between the description and the described.

A case in point is the digitization of complex cultural heritage artefacts. The following example is from a metadata enrichment project at the Heidelberg University Library using an xml-based metadata standard called METS

(Metadata Encoding and Transmission Standard).[42] It shows the structural metadata used to bind the various aspects of a digitized manuscript into a complex cultural artefact.

The digitization of the manuscript resulted in image files of various qualities for each scanned page. Usually, the high resolution images are used for preservatory reasons while the low resolution versions are used for access via the internet or as thumbnails for navigation. In this example, five different quality levels of the scanned pages are provided;

Minimum Quality;	**Standard Quality;**	**Thumbnail Image;**
1) MIN; 2) MINplus.	3) DEFAULT; 4) DEFAULTplus.	5) THUMB.

Table 16: Five quality levels of image-scans as used by the Heidelberg University Library.

As a first step, the scans are grouped according to their quality. In Figure 12, a group is set for the lowest quality scans (USE="MIN"). Each image file is assigned a unique identifier starting with the cover of the manuscript ("Vorderdeckel") as ID="filemin00001" accompanied by the definition of the file type as a .jpg compressed image file. Finally, the ID is linked to the actual computer file via a URL. This is done for all the minimum quality scans forming a File Group. Medium and high quality scans as well as thumbnails are grouped into their respective File Groups accordingly.

[42] The report is available at http://enrich.manuscriptorium.com/files/enrich/ENRICH_WP5_D_5_2_final.pdf, last access: 17th Feb 2011.

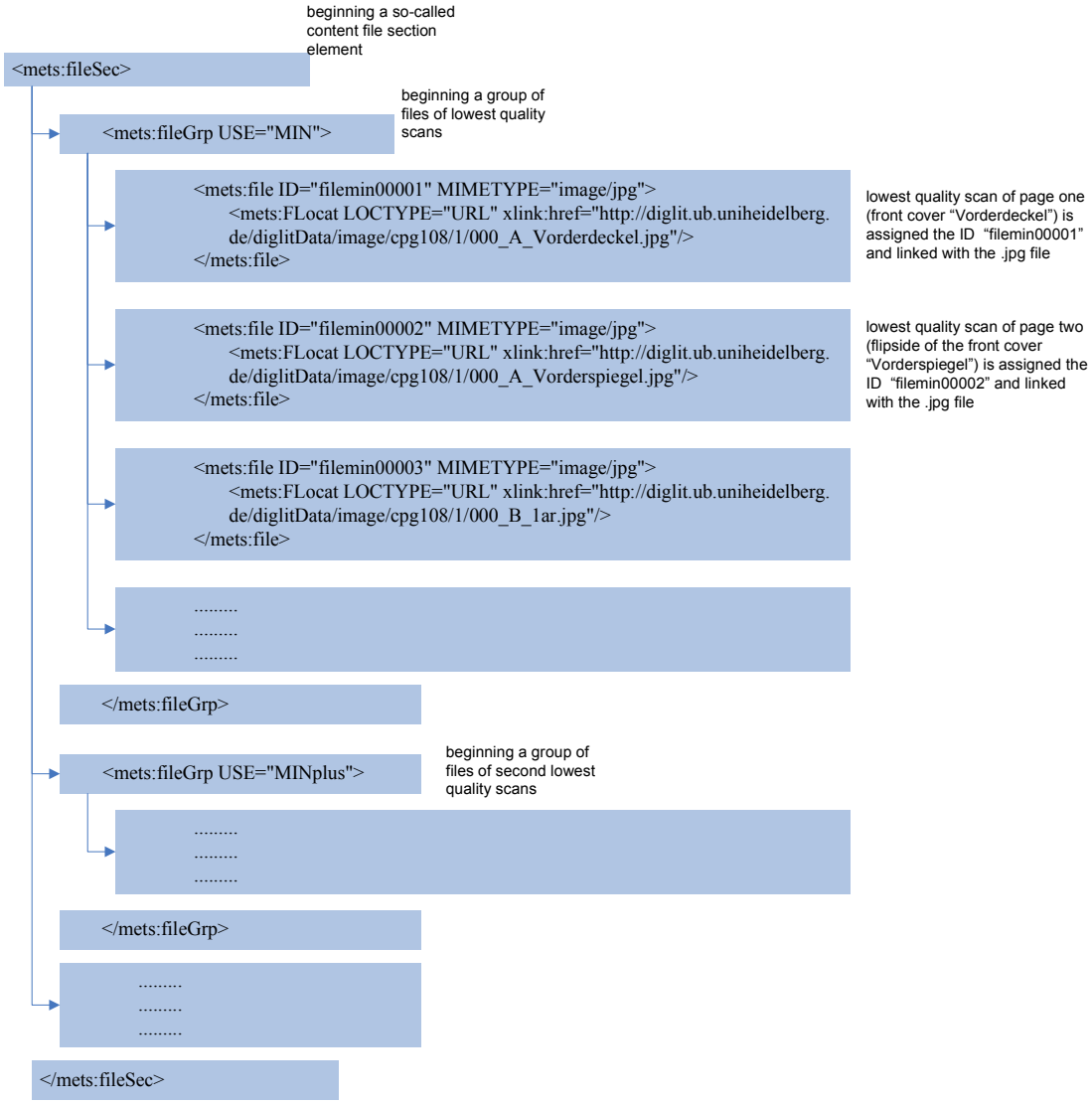


Figure 12: File grouping of scanned images according to image quality.

As a second step, a so called Physical Structural Map is applied that basically reflects the material make-up of the artefact, in this case the page sequence of the manuscript (see Figure 13). The page, used as the basic unit of the manuscript, is merely a conceptual container holding the various versions of the scanned pages and ordering them based on the sequence of the pages of the original. Again starting with the manuscript cover ("Vorderdeckel"), all the digitized versions are linked to the cover as the first page (ID="phys00001") based on the File Group IDs assigned as shown in Figure 12. In this case, there are five different versions per page starting with the lowest quality scans (filemin and fileminplus) to the standard (filedefault), above standard (filedefaultplus) and thumbnail

quality scans (filethumb). This is repeated for every page in the order of the physical manuscript's page sequence.

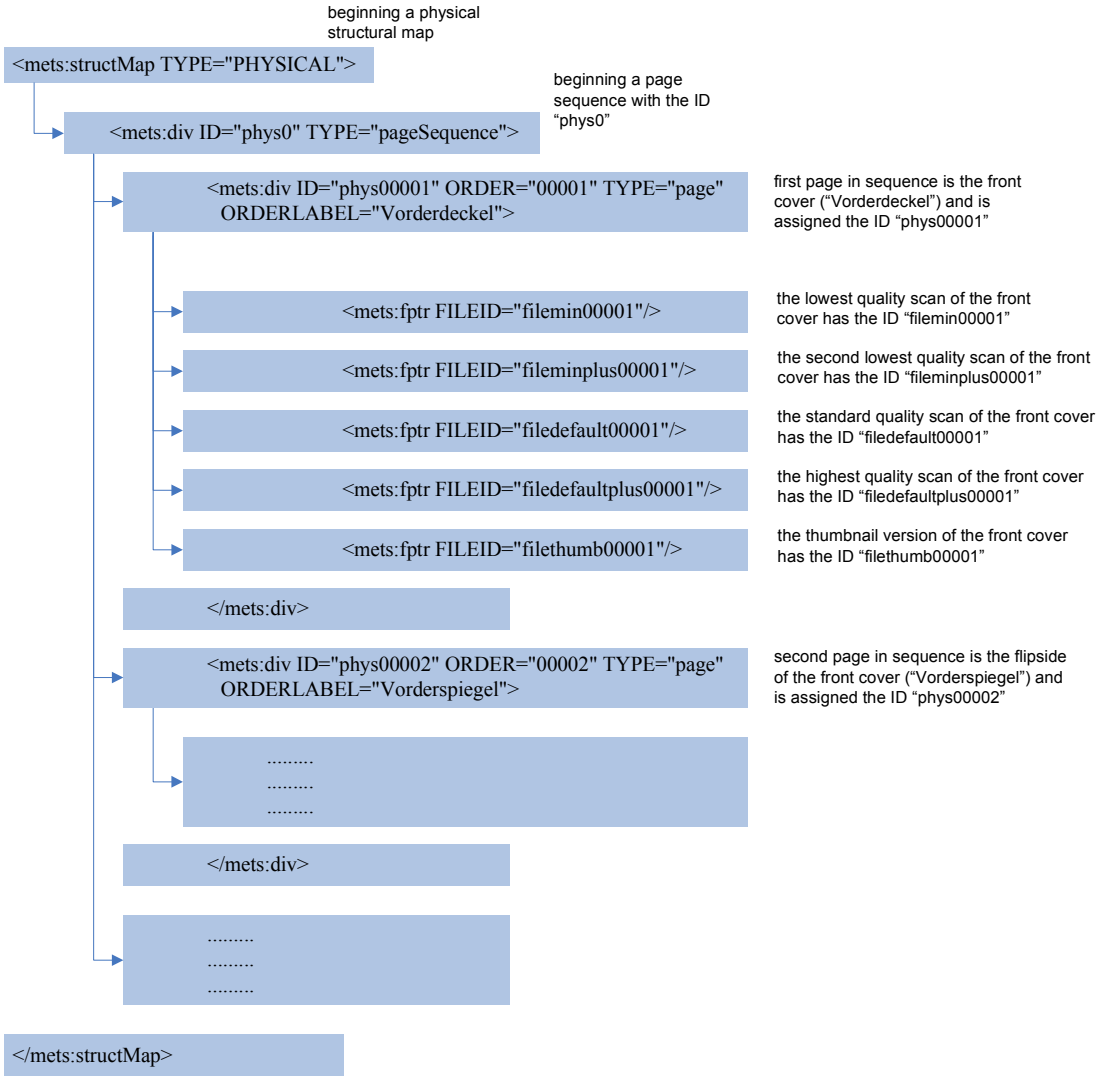


Figure 13: Physical Structural Map.

In addition to the order of the pages, however, there is also a logical structure in books based, in this case, on chapters. This is reflected in Logical Maps (see Figure 14) breaking the content of the manuscript down into meaningful units rather than into data-files as it was the case in the steps above. Each chapter of the manuscript is assigned a unique ID and labelled, starting with the front cover ("Einband vorne") that is treated as if it was a chapter of its own containing the front cover and flipside of the front cover.

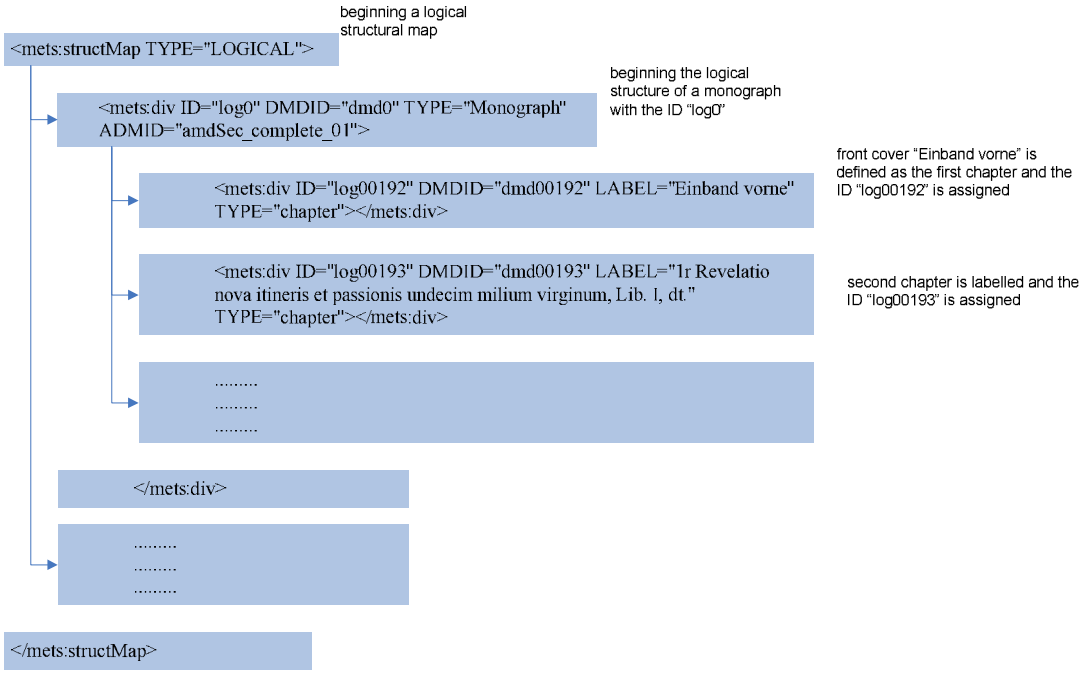


Figure 14: Logical Map.

Finally, the physical and logical maps are linked together in a Structural Links Map. Figure 15 shows how the first two pages (phys00001- phys00002), each of them linked to the actual versions of the scanned page, are linked to the logical unit front cover (log00192 is the ID for "Einband vorne" – the front cover) defined in Figure 14.

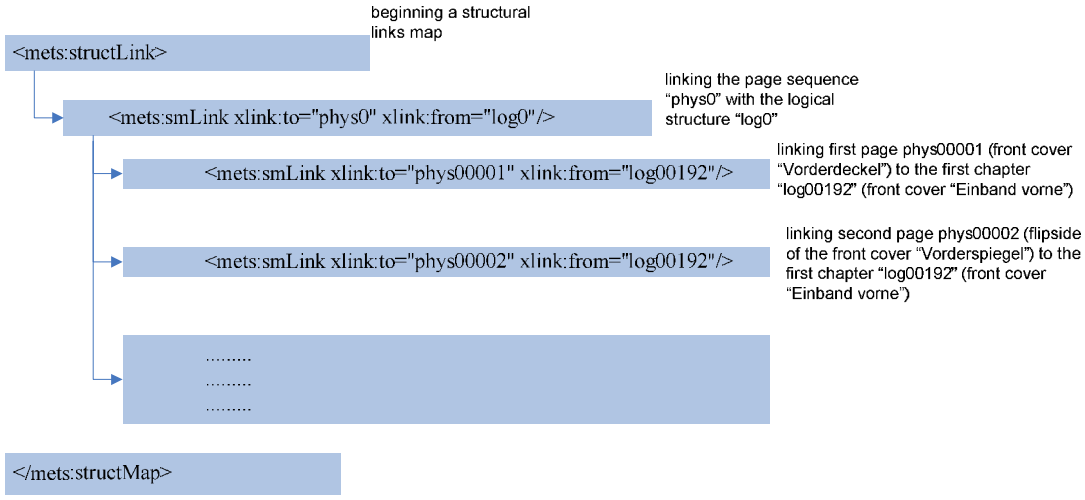


Figure 15: Structural Links Map.

All these various maps combined result in the structural metadata of the manuscript. In more abstract terms, the assemblage of the digitized manuscript

can be depicted as shown in Figure 16. In fact, the figure shows the functional equivalent to the binding of the manuscript – the glue that holds the pages together. The book as a packaged entity actually does not exist anymore but rather becomes a logical entity defined by a set of instructions on how the various parts are to be assembled the moment a user accesses the item.[43] The object that was given by the material make-up of the physical artefact has to be emulated computationally in order to be usable for a human being. Metadata, traditionally used to describe and to make an item findable by means of a catalogue, is in addition now used to actually assemble the item.

The elementary units used to assemble a cultural artefact can vary to some degree. While an image scanned manuscript lends itself to be decomposed into single pages, an OCR scanned print-book, for instance, could be decomposed into paragraphs, sentences or words. The textual elements could be stored in files separate from the pictorial elements. The structural metadata would then need to contain information related to layout, page breaks and so forth. The digital cultural heritage artefact, the item to be organized by a memory institution, is only a set of instructions on how it is supposed to be assembled and displayed. In other words, the artefact describes its own construction. It is what linguists call autological.[44]

The structural metadata *is* the manuscript. It defines the logical entity "manuscript" by instructing software applications on the emulation of itself. As a consequence, a distinction can be made between descriptive metadata used for the documentation and discovery of the information object and structural metadata used for the assemblage of the digital object. The crucial point, however, is that it is not only the descriptive metadata that needs to be preserved but also the structural metadata. If one finds an item in the catalogue, without the structural metadata intact, it will not be accessible; hence, the manuscript ceases to exist.

[43] In principle, the same can be said about the scanned images that form the elementary unit – the page – of this complex information object. Image files – in this case .jpg files – contain metadata instructions as well that tell a software application what to do with the 0s and 1s it processes.
[44] An example of an autological term is the word "English", which, as an adjective, signifies itself.

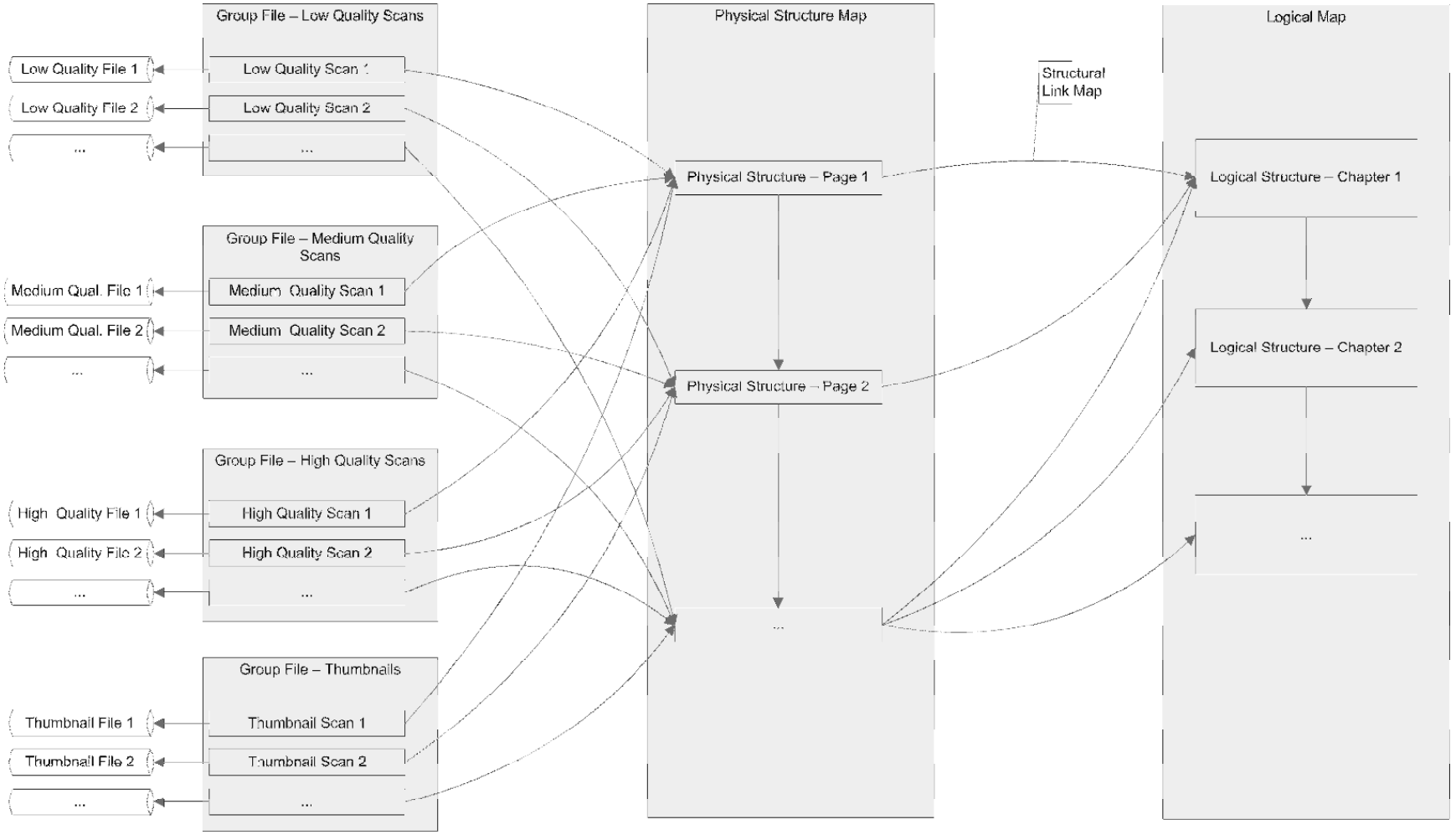


Figure 16: Schematic depiction of the structural metadata of a complex digital cultural artefact.

A case in point is the eLuxemburgensia online portal run by the Bibliothèque Nationale de Luxembourg (www.eluxemburgensia.lu) offering access to digitized newspapers. Besides the obvious search and browse functionalities, it provides highly enriched digital objects. Going back to as early as the mid 19th century, the newspapers have been image as well as OCR scanned. In terms of the image scan, the basic unit, however, is not the scanned newspaper page but paragraphs. By contrast, the basic unit of the computational text is the word allowing for full-text search of the whole collection.

Older versions of the newspapers were originally printed in the German typeface "Fraktur", which was broadly used in German-speaking countries until the early 20th century. Since most contemporary users are not familiar with Fraktur, the result of the OCR scan is displayed next to the image scan in a standard font.

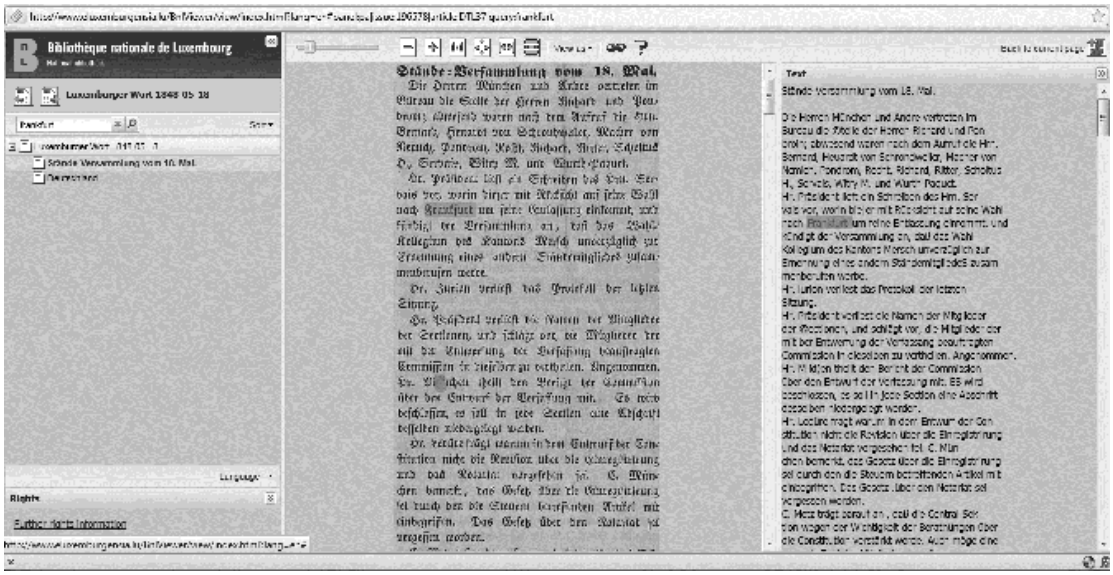


Figure 17: Screenshot of an emulation of a newspaper article at eLuxemburgensia.

Hence, an issue of a newspaper is in fact a container or a category, a complex digital object comprising articles, paragraphs, images and computational text assembled according to the instructions encoded as structural metadata. The images and the computational text are linked allowing for advanced search features. For instance, if one searches for a specific word that matches a word in the index the relevant article is shown with the queried term highlighted in the

image as well as in the text.[45] Crucially, there is no such thing as *the* newspaper or *the* article. There is only data that is assigned to the category of a specific paragraph, assigned to the category of a specific article, which, in turn, is assigned to a specific page and, finally, to the broader category of a specific issue of the newspaper. The digital object a user interacts with is an emulation based on the assemblage of data via structural metadata – a momentary rendition to be disassembled the moment it is abandoned by the user.

6.4 Digital Surrogate Objects

The traditional way of memory institutions to provide for discovery is based on very specific procedures. In a library, for instance, the paper-based card catalogue usually allows only a search according to author names or keywords referring to the location of the actual artefact (a book, a newspaper issue, a CD-ROM) via a shelf mark.

Figure 18: Catalogue discovery system.

The basic set-up shown in Figure 18 can be easily translated into a data-based model as it is the case with Open Public Access Catalogues (OPAC) by copying the metadata from the card to a data-base.

Digital objects are made to fit into this paradigm by being treated like books or newspapers. The crucial point is that descriptive metadata and described object are separated (see Figure 19). The user navigates through the index in order to discover artefacts. Although data-base technology allows for additional functionalities such as searching for titles or the usage of Boolean operators, the underlying concept remains the same. Hence, digital libraries can turn out to be nothing more than online portals allowing users to search through the catalogue

[45] Google Books, for instance, works basically the same way.

data-bases similar to searching through a card catalogue. This was the case with the EDL prototype launched in November 2008.

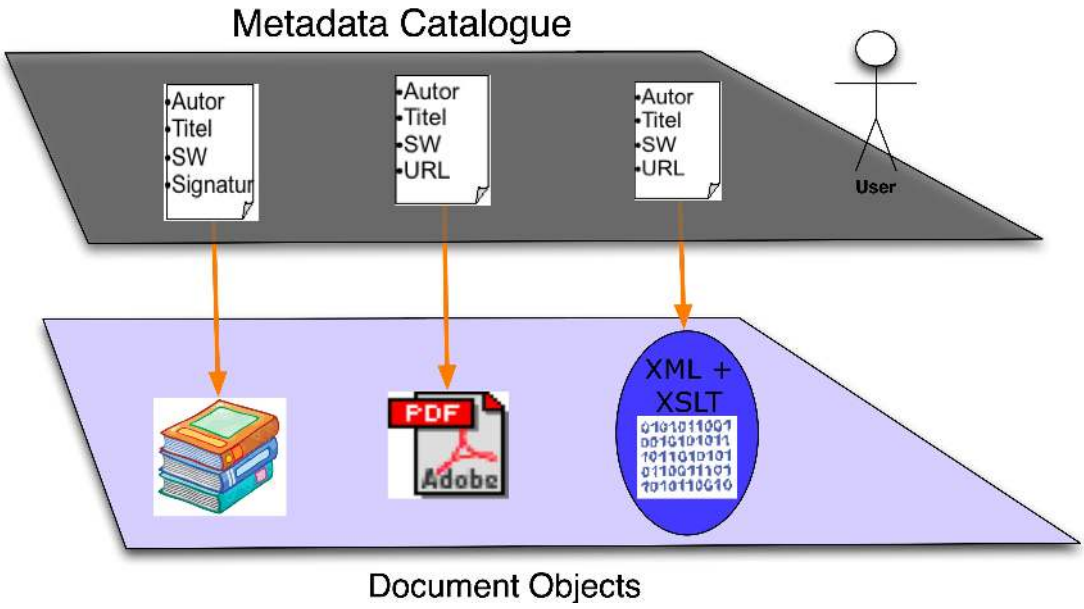


Figure 19: The "traditional" digital library discovery system (ES-020:14).

The metadata schema used for the prototype was a subset of the Dublin Core (DC) schema discussed in Chapter 2.4.3. The metadata provided by the member libraries, archives and museums was basically mapped according to EDL's Dublin Core standards. The rationale of DC as the lowest common denominator for describing electronic content helped in standardizing the way digitized cultural heritage artefacts are described within the Europeana information space and across the domains of libraries, archives and museums traditionally using their own sets of schemas and standards.

> "*Our metadata standard builds on the Dublin Core standardisation efforts. The metadata that we are supplied with is created using different standards and local variants, and we are compiling guidelines to help contributors make their metadata Europeana-compliant. This usually simply means mapping the institution's metadata to Europeana's broad schema*" (EP-001:4-5).

The DC standard set was extended by twelve Europeana specific elements into a schema referred to as the Europeana Semantic Elements (ESE) specifications, which, for instance, contain additional information such as the link to the digitized cultural heritage artefact at the provider's site, user tags or the categorization of the artefact into text, image, sound or video (EX-003:15).

Figure 20 illustrates the prototype data-model showing a selection of DC descriptive metadata categories and their linkage to a digital object stored at the content provider's site. Similar to a catalogue card, the metadata record combines descriptions and categorizations as well as a pointer to the location of the described.

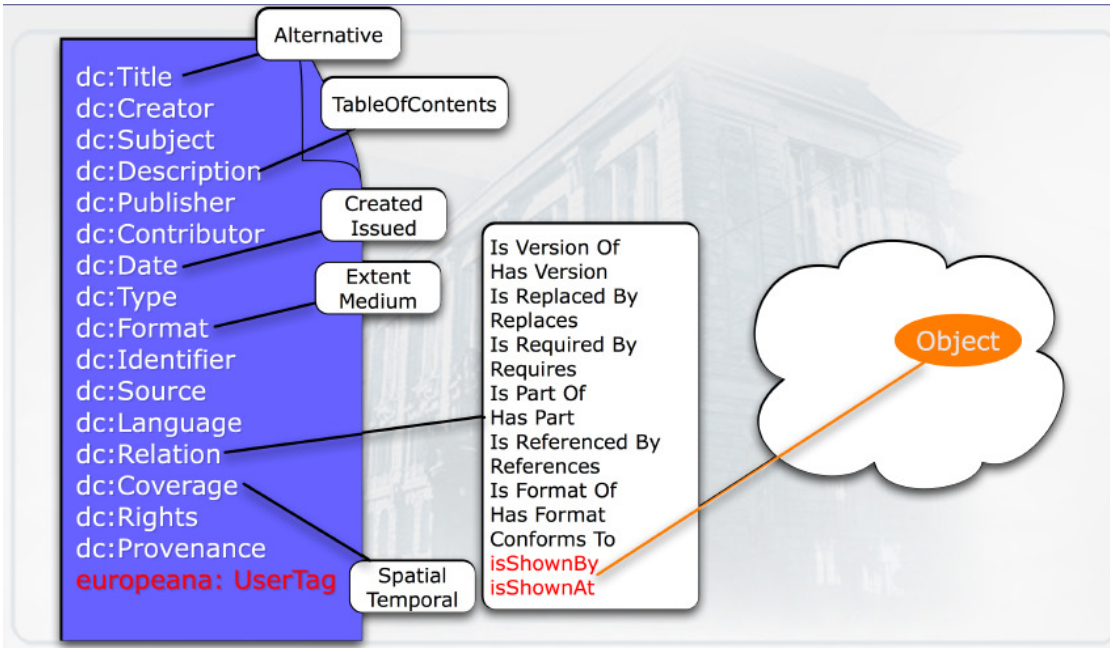


Figure 20: Prototype data-model based on ESE (EX-001:9).

The metadata collected and processed according to the ESE standards fit the traditional digital library paradigm as shown in Figure 19. It addresses the core functionalities of the portal that is discovery via search or browse interfaces. In this sense, calling the schema "semantic" is misleading as it may invoke the image of semantic web functionalities and "*probably shouldn't have been called "semantic"*"" in the first place (EX-004:14). However, the standardization of the description of all the digital cultural heritage artefacts available through Europeana was the first step towards interoperability and, thus, the modelling of rich digital surrogate objects in the Europeana v1.0 phase.

> "[T]*he ESE metadata are just another form of object description, but their specific value lies in the fact that all surrogates have a description of this format, thus they enable interoperability*" (ES-020:24).

Hence, Europeana can be described as a meta-aggregator of metadata about digitized cultural heritage artefacts hosted by member libraries, archives and

museums – the content providers. However, the metadata, harvested and ingested into Europeana's system, is brought into a form referred to as a Digital Surrogate Object (DSO). The term surrogate is a basic concept of library and information science (LIS) referring to anything that represents and stands for an actual object – a catalogue card being the best known example. In the case of Europeana v1.0 however, the surrogate gains a new role within the Europeana information space.

First of all, the DSO is a representation of a representation. The surrogate stands for mostly digitized cultural heritage artefacts. Due to issues of copyright and storage space, Europeana does not copy full digital objects into their repository but rather links to them through its surrogates. As a result, Europeana aggregates surrogate objects only. As Figure 21 shows, a surrogate must contain a persistent link (the Root Component) to the digitized object (Digital Representation Object - DRO) and descriptive metadata about the object the surrogate stands for. Abstractions, such as a thumbnail, or annotations may be part of a surrogate object but are not necessary for its existence.

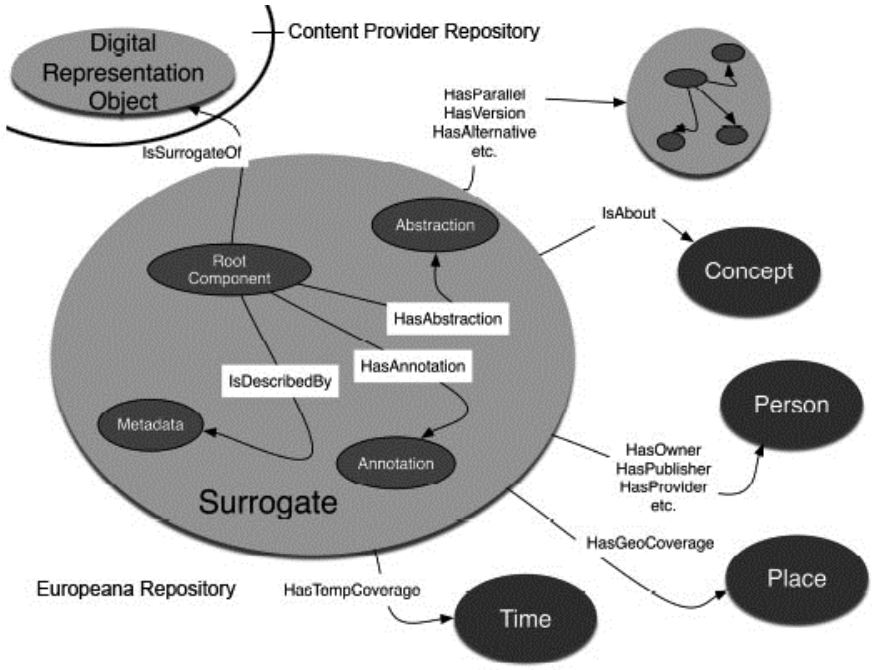


Figure 21: Europeana's DSO model (ES-006:17; ES-012:21).

Second, the collection of surrogates is supposed to correspond to the content provider's collection of Digital Representation Objects one-to-one.

> "*Europeana will create a parallel data space inside the system that is a representation of the content providers' object space*" (ES-020:15).

This also applies to complex digitized objects composed of smaller modules as described in the previous section. Hence, each of these smaller parts, say each page, has a corresponding surrogate object. Within Europeana's repository, those surrogate objects are then assembled into a complex surrogate object that stands for the complex compound digitized object – the book as a whole in this example. Hence, Europeana mirrors the internal structure of the digitized objects through its surrogates. In this sense, surrogates do not only stand for specific objects but are partial replications of them. The structure of the digitized object is translated by means of qualified links holding the various parts of the artefact together. This approach will be discussed in more detail in Chapter 6.5.

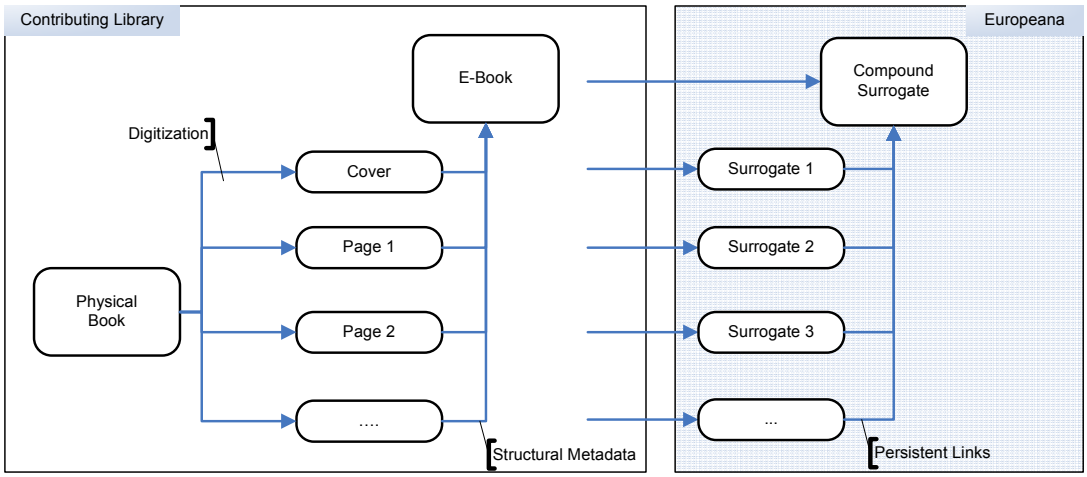


Figure 22: Structural mirroring of a digitized book via simple and compound surrogates.

The key element in this setup is metadata. Be it the complex e-book itself or the single scanned page files, the compound surrogate or the single surrogates (see Figure 22), each of these obtain their functionality within the corresponding information system because they are described as such through metadata. For instance, a specific e-book is in fact only an abstract category – a container for various files, which, in turn, are identified, combined, brought into the appropriate page order and displayed by means of processing computational instructions – structural metadata.

By the same token, the Europeana surrogate object becomes a surrogate object only when descriptive metadata is a part of it (see Figure 21). As it was already said above, the minimum requirement for the formation of a surrogate object is the Root Component and descriptive metadata. As a consequence, the descriptive metadata (that which used to be on the card) identifies the surrogate (that which used to be the card) as a surrogate of which it is a part of. While the catalogue card was identifiable as a surrogate by means of its physical format and due to being filed in a catalogue, the surrogate object describes *itself* as a surrogate through metadata. It is autological, just like the digital cultural heritage artefact it represents. The Europeana surrogate object comes into being through the extraction of specific elements and the recombination of these elements – their relationships encoded into structural metadata. In other words, Europeana constructs an entirely new digital object that only exists through the links between its elements. The Root Component (the link to the Digital Representation Object) is only findable and, thus, only exists when it is linked to metadata (see Figure 21). By the same token, descriptive metadata achieves its functionality only by informing a user about the location of the item it describes, that is, if it is linked to the Root Component (which is only a link itself).

The disintegration of the catalogue into digital surrogate objects results in a different view on Europeana as a meta-aggregator. In a way, collecting metadata from libraries, archives and museums is similar to providing a unified card catalogue. In contrast to the prototype's mapping of the providers' metadata, Europeana's digital surrogate object is not a mere copy of descriptive metadata provided by the member organizations but a different kind of surrogate altogether. Europeana harvests the provider's metadata but lumps it together with a partial replication of the "original" digital object. To some extent, it is a mix between the metadata object (e.g. the catalogue card) and the object itself (e.g. the book).

To conclude, as catalogue cards, surrogates contain descriptive metadata about cultural heritage artefacts. However, as Europeana Digital Surrogate Objects, the surrogate becomes a partial replication of the artefacts it stands for. Depending

on the level of granularity the digitization results in, the surrogate can become a very complex digital object itself. The complexity can vary from simply being a low-resolution thumbnail of a picture taken from a famous painting to compound surrogates containing other surrogates. A surrogate for a highly granular digitized newspaper issue, as exemplified by the eLuxemburgensia described above, combines surrogates for each paragraph, article and page, maybe even in various degrees of quality, with the result of the OCR scan into a complex surrogate that stands for the newspaper issue as such. The relationships between the actual computer files, the simple surrogates, the compound surrogate and descriptive metadata are then described through qualified links between them, as will be shown in the next section.

Accessing a cultural heritage artefact online requires a lot of effort in contrast to enabling a user to simply find the description and documentation of an artefact. In this context, surrogates are not only the primary interface for discovery, as is the case with catalogue cards and the prototype portal, but also the primary and, for most users, only artefact to interact with.

6.5 Networked Contextualization

During the prototype phase, the Europeana data-base was only accessible through its online portal, which allows for search queries according to traditional categories such as name of creator, title of item or subject matter. However, developments are on the way to make the Europeana surrogate space "*an integral part of the overall information architecture of the WWW*" (ES-020:15). This is expected to be accomplished in various ways. For one, the surrogate objects will be made crawlable for search engine harvesting. Hence, the navigability through the Europeana repository is, to some extent, passed on to search engines and their algorithmic operations of indexing and ranking.

Second, a navigational layer is implemented on top of the surrogates based on semantic web technologies to enable a user to browse through the collection based on semantic concepts (see Figure 23).[46]

> "*The user now primarily interacts with the semantic network to explore the Europeana surrogate space which now has the metadata as parts of the surrogates and surrogate aggregations*" (ES-020:15).

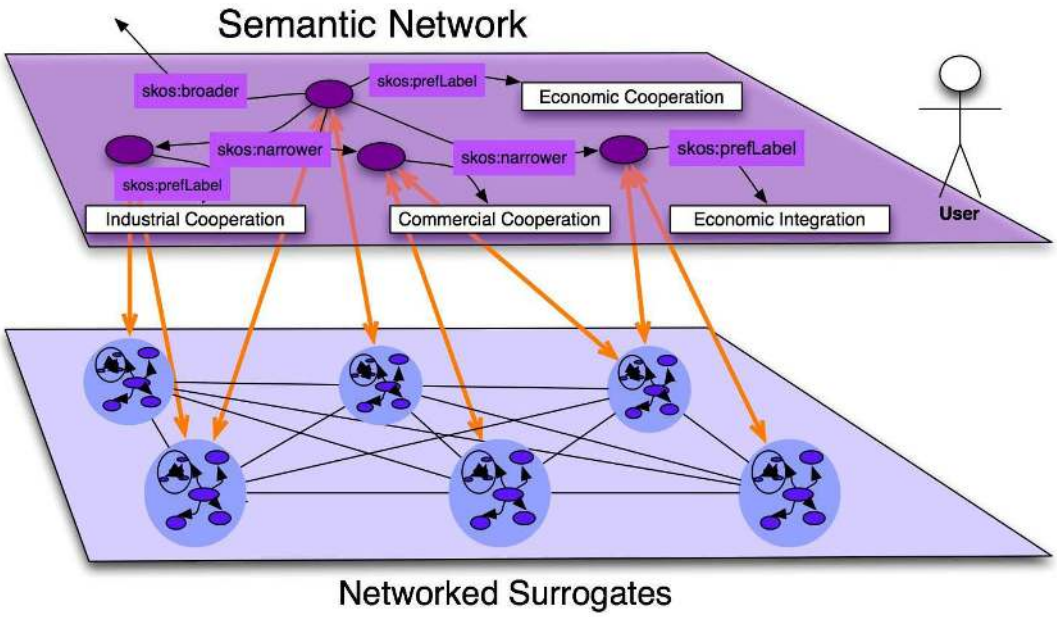


Figure 23: The two layers of the Europeana Information Space (ES-020:15).

Combined, these two layers form what is referred to as the Europeana information space – a metaphorical landscape of surrogate objects made navigable by the second layer – the conceptual map of the surrogate layer. The actual digitized artefacts hosted by the content providers form a third layer but are of a lesser concern at this point (EE-071:2). A major step towards the accomplishment of advanced semantic web based browsing functionalities is the ongoing development of the Europeana Semantic Elements (ESE) schema into the Europeana Data Model (EDM) (EE-065:28). The EDM is based on the separation between the actual object – the physical painting, book, document and so forth – and its digital representations – the digital surrogate object. In contrast to ESE, which allows only one set of descriptions (see Figure 20), EDM allows for multiple sets of descriptions, so called proxies.

[46] The prototype of the semantic search functionality can be found at http://eculture.cs.vu.nl/europeana/session/search, last access: 28th Feb. 2011.

> "[A] *proxy is specific to a given provider, and is used to represent the description of the provided object, as seen from the perspective of that specific provider. With proxies it is possible to represent different, possibly conflicting pieces* [of] *information on provided objects, while still keeping track of the provenance of this information. For instance, the title of Mona Lisa for Joconde could be "Portrait de Mona Lisa" while for Louvre it could be "Portrait de Lisa Ghirardini""* (EE-070:12).

In other words, the same physical object can be represented by many surrogates even containing contradictory metadata. As a result, the metadata provided by a memory institution is kept, in order to preserve the provenance of the artefact, rather than normalized as was the case with ESE.

Encoded as RDF Triplets (see Chapter 2.4.3), each element of the description is a web resource; that is, each has a unique URI that can be linked to other resources into increasingly complex aggregations of descriptions and digital representations. Figure 24 shows the example of a proxy of the Mona Lisa portrait based on the title and the creator metadata elements as provided by the Direction des Musées de France (DMF). For this and the following figures, the actual RDF codes and qualifiers have been simplified by exchanging them with more meaningful and easy-to-understand terms.

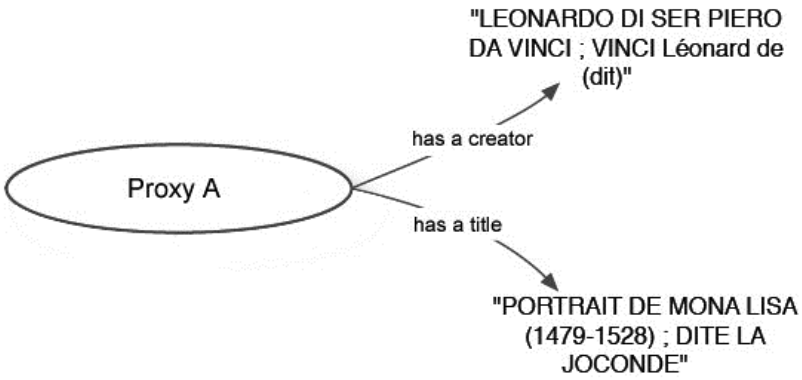


Figure 24: Simplified schematic depiction of a proxy (EE-070:14).

The proxy, as depicted in Figure 24, contains only descriptive metadata. However, it can be linked to an aggregation with its own set of descriptive metadata. The aggregation can be linked to further elements such as a digital placeholder representing the actual material object or to the various digitized

versions of the painting in different resolutions to be viewed online (see for instance the example of the manuscript scanned for the Heidelberg University Library in Chapter 6.3). Figure 25 shows an extended aggregation of the Mona Lisa with further metadata records. It should be noted in Figure 25 that the aggregation has its own creator, which is not Leonardo da Vinci (he is the creator of the actual painting), but the provider of the descriptive metadata – the DMF in this case. This is based on the strict distinction between the actual painting and its digital surrogate representation mentioned above.

> "*One of the most important pieces of information expected from providers is the distinction between the metadata that applies to the object itself, and the metadata that applies to the digital representations*" (EE-070:18-19).

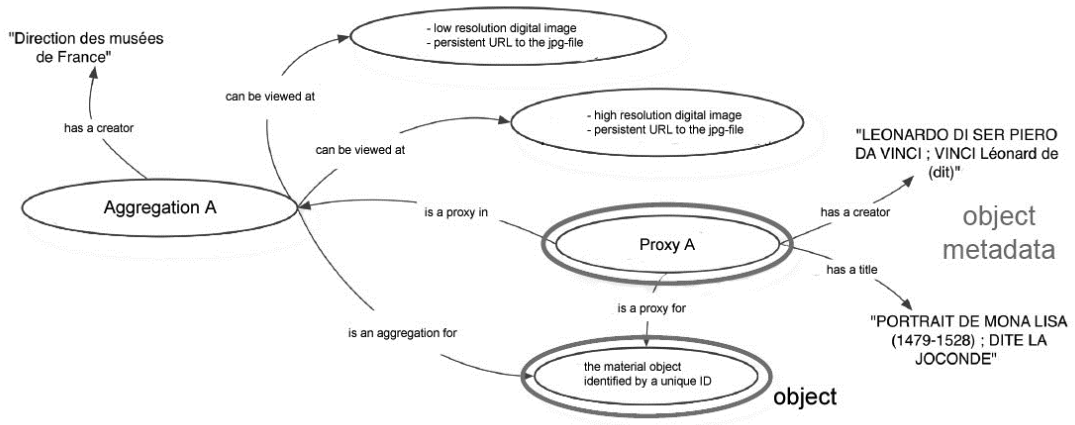


Figure 25: Simplified schematic depiction of a metadata record (EE-066:12).

The digitized painting is an object of its own; it was not created by Leonardo da Vinci but by the DMF. The same applies to other descriptions. For instance, the dimensions of the material painting are different than the dimensions and resolutions of the digital images and are documented separately. Hence, Proxy A could have a link to a metadata element containing the information on the dimensions of the painting. However, these are not the dimensions of the digital representation, which may vary according to the level of the scan.

Now, the way the DMF describes the Mona Lisa and the digital representations (the digital images) it provides differs from the data provided by the holder of the painting – the Louvre. Normally, these two different descriptions would have

needed to be normalized. With the Europeana Data Model (EDM) both can be implemented into the Europeana information space as two separate aggregations (see Figure 26). The same material object – the portrait of Mona Lisa in this case – can be described and, therefore, represented in multiple ways.

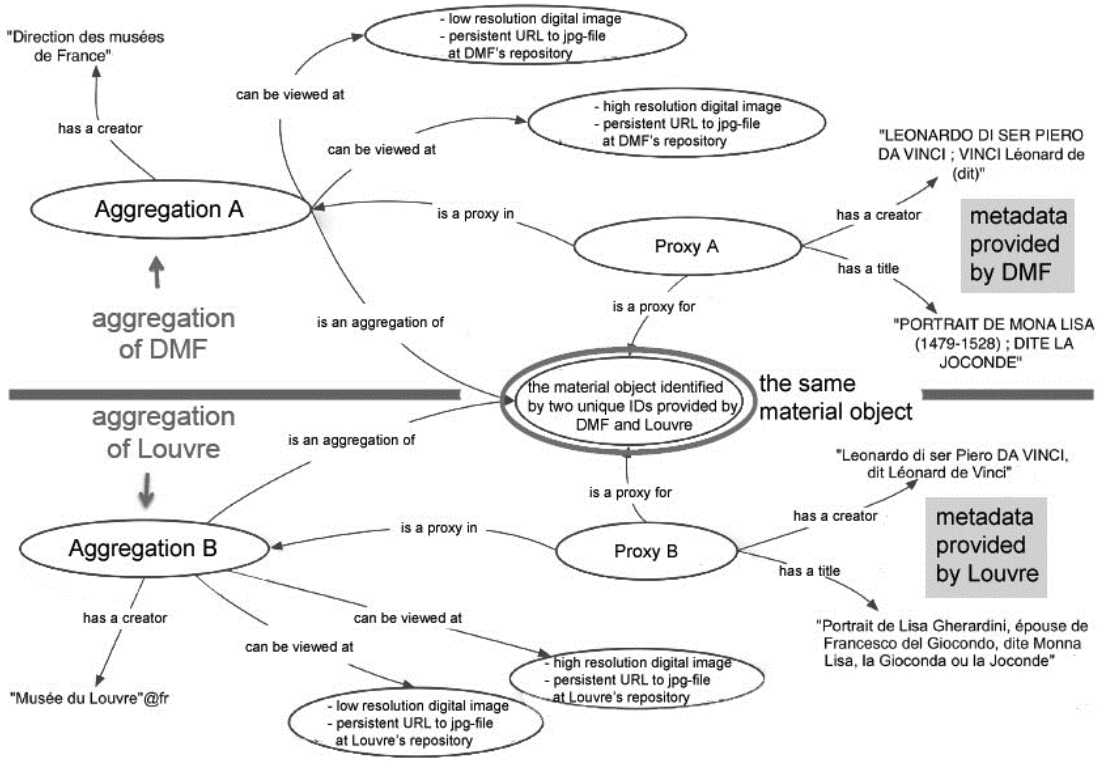


Figure 26: Simplified schematic depiction of two metadata records (aggregations) (EE-066:13).

Finally, Europeana adds its own aggregation aggregating the aggregations provided by contributing libraries, archives and museums (see Figure 27). A meta-aggregation enriched with additional metadata, the Europeana aggregation forms the foundation for further functionalities and services such as the semantic navigational layer depicted in Figure 23. Another example is the so-called "landing page" created by Europeana for every cultural heritage artefact – a basic web-page containing imagery and information about the respective artefact, which are also crawlable by search engines (ES-012:18).

> "*On a very abstract level, Europeana can be seen as a large collection of representations of born digital or digitised cultural heritage objects which themselves remain outside the Europeana data space. In this abstract vision, the representations are linked to each other and additionally are contextualised with links to nodes of a semantic network that forms the second data layer in Europeana. These two links together*

are used to create rich functionality that is offered on the user interface giving the choice to the user of navigating on either of these levels" (EX-002:10).

Going back to the example of Mona Lisa's portrait, the creator, for instance, can be linked to further resources such as authority files containing the various ways Leonardo da Vinci's name is spelled in various languages or to his place of birth, a historic period or artistic tradition, which, in turn, can lead to other cultural heritage artefacts.

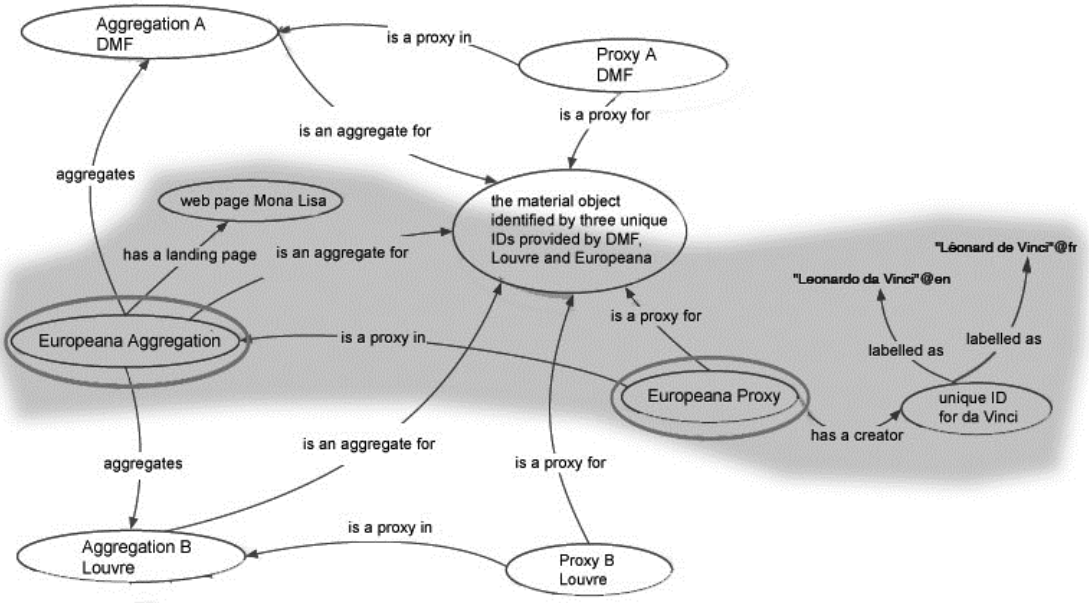


Figure 27: Simplified schematic depiction of two metadata records and a Europeana aggregation (EE-070:16).

The same model can also be applied for versioning such as aggregating different editions of the same book as separate aggregates. Hierarchical relationships are another possibility; that is, for instance, a book can be an aggregation of pages (EX-004:22). Represented by proxies and aggregations themselves, the pages can be linked to one another through sequencing. Hence the proxy of page 2 is next in sequence after the proxy of page 1. The aggregation representing the book would then aggregate all the pages or rather the aggregations of the pages (EE-066:27).

As one can see, the quite simple surrogate model described in the previous section (see Figure 21) becomes a very complex aggregation of aggregations and relationships. In this sense, the surrogates are not ordered in a fixed way

according to a fixed rationale but rather contextualized in a network of descriptions, labels and classification schemes, which can be contradictory.

At the most general level, the contextualization can be accomplished according to concepts, time, agents and places (depicted as Time, Space, Person and Concept in Figure 21) basically answering the questions of what, when, who and where (EE-065: 5). As a consequence, new associations can be offered that lead to improved search and, above all, browse functionalities, which would have been impossible with card or even data-based catalogues. As the following example from the Europeana website illustrates, the avenues to discovery are not pre-defined any longer.

> "*For example, a search for "Paris" intuitively leads to connected concepts like items in the Louvre, a Paris-based institution, or paintings portraying people named Paris, such as the Greek prince, who abducted* […] *Helen of Troy. From there, more links point the way to topics like the "myth of Paris" or the mythical Apple of Discord - a golden apple that sparked a dispute and eventually led to the Trojan War - and then on to the forbidden apple eaten by Adam and Eve*".[47]

The truly remarkable outcome of this development is that the contextual information does not need to come from Europeana or from other Europeana content providers. The semantic layer itself is embedded into a wider information landscape coupled to external resources provided by the Linked Open Data (LOD) community (www.linkeddata.org), which is expected to be joined by Europeana sometime in 2011 (EE-067).

LOD is an initiative that connects, for instance, geolocation data with data from Eurostats, Flickr or Wikipedia's "dbpedia." Hence, the contextualization of the surrogate object is partly passed on to external resources and services. In return, Europeana combines data from various external providers into mash-ups, which would, for instance, include translations of names of personalities or localities. The semantic network layer is not exclusively created and maintained by Europeana but rather in cooperation with other (meta)data providers, which do

[47] http://www.europeana.eu/portal/thoughtlab_semanticsearching.html, last access: 28th Feb. 2011.

not need to be memory institutions. In other words, data from other providers is used as metadata for contextualization.

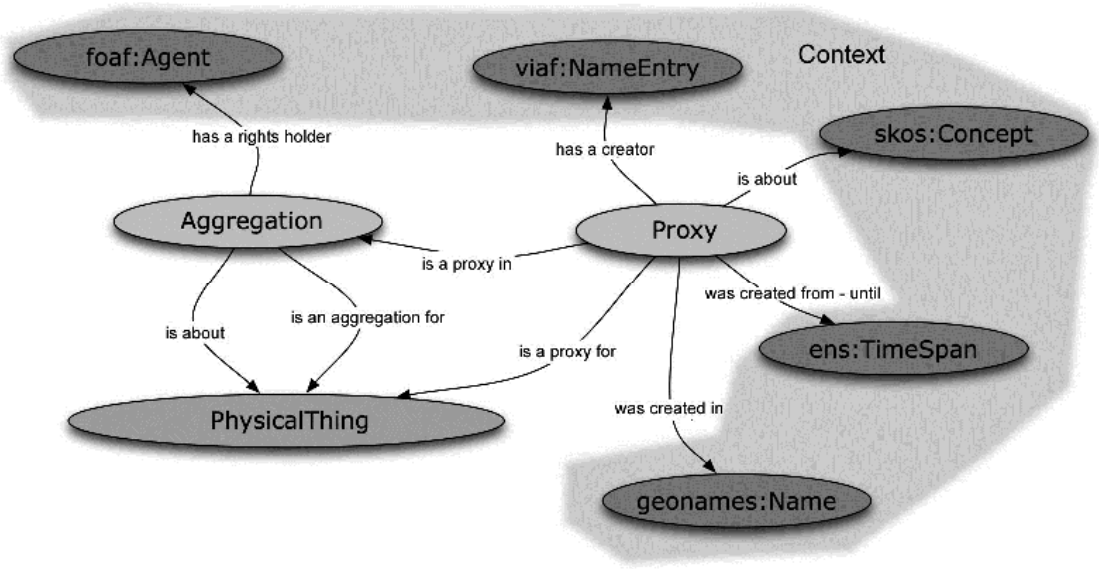


Figure 28: Simplified schematic depiction of a contextualized metadata record (EX-004:27).

Figure 28 shows an example for the contextualization of aggregations and proxies. For instance, the aggregation can be linked to a rights holder whose identity is stored in the Friend Of A Friend (foaf) data-base – an open service allowing to connect social network sites.[48] The Virtual International Authority File (viaf) brings together names of persons and institutions provided by memory institutions from around the world hosted by the Online Computer Library Center (OCLC, see also Chapter 2.4.3).[49] For instance, it provides the various ways Leonardo da Vinci's name can be spelled in different languages - ranging from Léonard de Vinci, Leonard Davincy, Leonaďo da Vinchi to the name written in Hebrew and Arabic. Each of them is made identifiable as different names of the same person. The same applies to geonames – an open data-base of geographical names.[50] In this case, a search for Istanbul, for instance, delivers the latitude/longitude of the city (displayed via Google Maps) and also a list of current and historical names in a range of languages (e.g. Byzantium in Latin and Constantinople in English). Another noteworthy possibility not depicted in Figure 28 would be a link to Wikipedia's open data-base dbpedia providing

[48] For further information visit http://www.foaf-project.org/, last access: 28th Feb. 2011.
[49] For further information visit http://viaf.org/, last access: 28th Feb. 2011.
[50] For further information visit http://www.geonames.org, last access: 28th Feb. 2011.

annotations, abstracts and other information related to a cultural heritage artefact, its creator, the time period it was created in and so forth.

Of course, these data are linked to other data. For instance, being the owner of the portrait of Mona Lisa, the Louvre can be linked to latitude/longitude data, to members of its staff or dbpedia. Similar to a hypertext, the context and, therefore the surrogate, has no boundaries as such. In metaphorical terms, the catalogue has been disassembled and its cards are stretching beyond the boundaries of memory institutions through networks of linked and open data. By the same token, Europeana's metadata is open to be accessed by others and used as they see fit.

> "*It should be clear by now that a view of Europeana as a huge agglomeration of data would be terribly inappropriate. However, viewing Europeana as a huge information repository would be almost as inadequate. Instead of such views, we have described the intended characteristics of Europeana as part of what we called a "cultural commonwealth"*"" (EX-002:8-9).

> "*In the perspective of this approach, Europeana can be thought of as a network of inter-operating contextualised object surrogates enabling semantics based object discovery and use. This network in turn is an integral part of the overall information architecture of the WWW: especially in the case of context data Europeana will use Open Linked Data from the WWW wherever possible instead of creating specific context nodes within the Europeana data space. Likewise, the semantic layer of Europeana itself is served to the outside as linked data* […]" (ES-020:15).

Given the transition from a rationale based on the catalogue to a rationale based on contextualization through linked data and metadata, it becomes clear why the notion of APIs takes such a central role in the way Europeana defines its services and, ultimately, itself. Initially planned as an interface "*that learning sites and partners can use to integrate Europeana content into their own content*" (EP-002:14-15), it will permit "*the general users later on to bring together contents in new applications on different platforms*" (ES-014:12). The definition of scenarios of use, being in the hands of experts for the portal, is extended to the general user making the distinction between re-user and end-user obsolete.

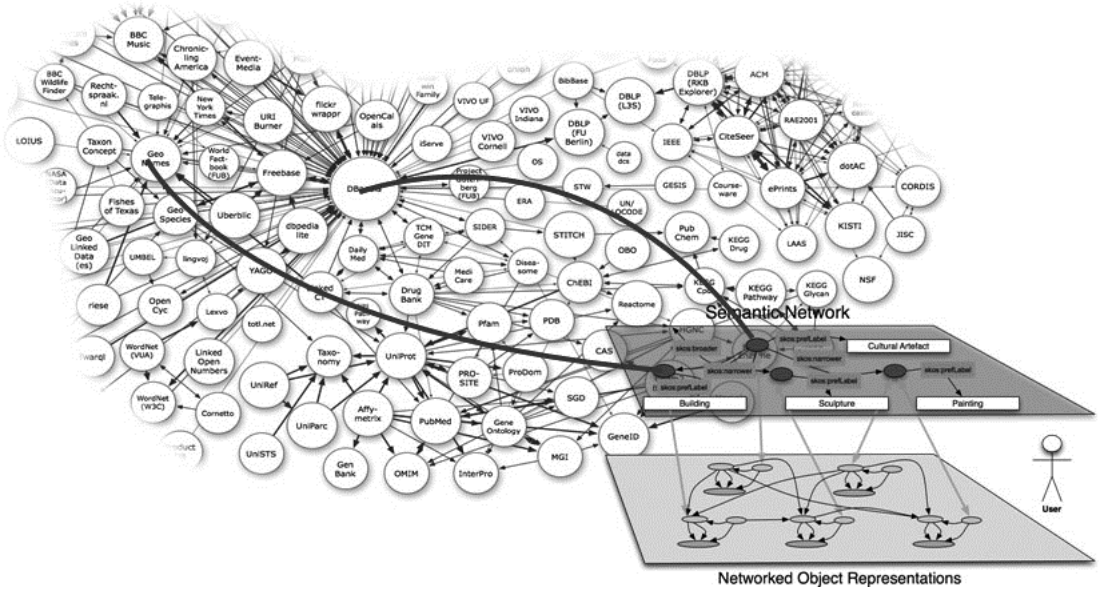


Figure 29: Integration of Europeana into the LOD information space (EX-004:26).

More importantly, however, the API has replaced the notion of a "portal" as the central and defining characteristic of Europeana, as "*Europeana's primary incarnation*" (EX-002:9). The API is not only another interface allowing access to Europeana's repository complementing the portal functionalities found on www.europeana.eu. More fundamentally, the portal or rather the scenarios of use Europeana offers are built on top of the API (see Figure 30).

A "*transversal approach*" (ES-014:12), APIs cut across content, the domains of memory institutions and accessibility. This is made possible due to the transition from a metadata data-base to Digital Surrogate Objects, which perform as functional primitives that can be linked to other data or used for the development of further services or mash-ups within or outside of Europeana. Two telling examples are the Europeana search widget and the Europeana API geo-wrapper. The first is a widget developed by Europeana that can be easily implemented into a website allowing a user to directly query the Europeana data-base. The widget, however, is controlled by Europeana and simply brings some Europeana portal functionalities to another website. Initially, the widget could only be used by Europeana partners but was made accessible to everybody allowing lay users, for

instance, to showcase their favourite artefact from the Europeana collection on their personal website or blog (ES-025:2).[51]

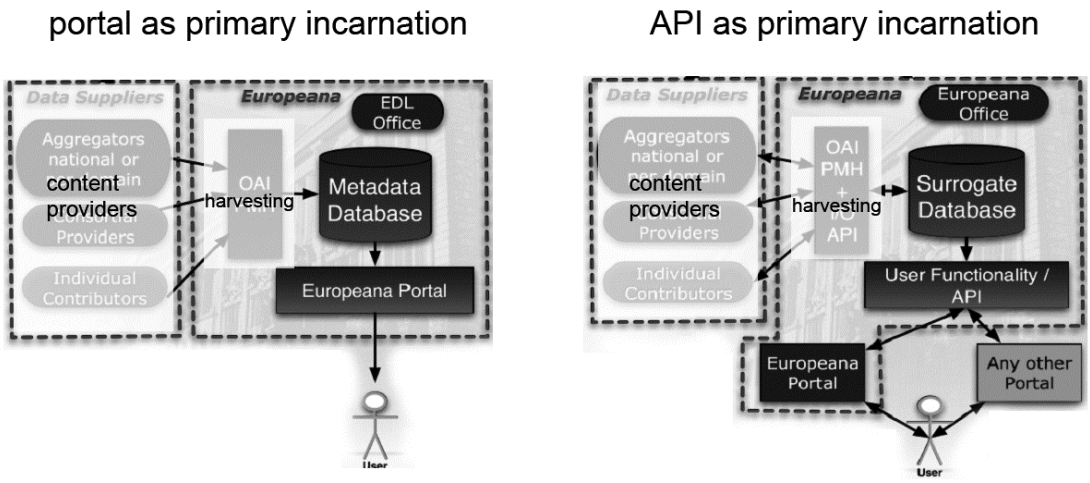


Figure 30: Portal and API as Europeana's primary incarnation (EX-001:17-18).

The geo-wrapper, on the other hand, was developed by an independent programmer.[52] A mash-up of openstreetmap.org – an open source online map service – and Europeana data, the wrapper shows the location where cultural artefacts from a given period of time can be found. In contrast to the search widget, this and other API based apps[53] provide for alternative scenarios of use independent of Europeana. Users approach data about cultural artefacts (rather than cultural artefacts themselves) and make them accessible in various ways. In other words, users also create cultural artefacts out of data by making them findable and usable – a task that was exclusive to libraries, archives and museums in the domain of cultural heritage.

As a consequence of the networked contextualization, Europeana is not able to exercise the degree of control over its own collection as memory institutions were used to in the paradigm of silos of knowledge. For instance, Europeana was forced to renegotiate the contractual agreement with its own content providers regarding a clause that Europeana content may only be re-used in non-

[51] An example can be found at http://www.athenaeurope.org/index.php?en/187/europeana-search-widget, last access: 16th May 2011.
[52] A working demo can be found at http://amercader.net/dev/geoeuropeana, last access: 16th May 2011.
[53] For a selection of promoted APIs visit http://version1.europeana.eu/web/api/application-gallery, 16th May 2011.

commercial contexts (EE-071:2). "*Once metadata is made available as Linked Open Data to the community of application developers, it is no longer possible to have absolute control over it, and data elements may be re-used in commercial applications or websites* [e.g. websites with commercial elements such as adverts, online shops etc.]" (EP-005:13).

6.6 Conclusion

Europeana does not provide an ex-ante order nor does it exclusively provide the means to navigate its repository. Instead, it provides potentialities for scenarios of use that can be actualized by, in principle, anybody. In this sense, Europeana uses the potentialities offered by APIs in relation to Linked Open Data by contextualizing its surrogates through metadata provided by other institutions. Europeana does not only offer an API but also accesses the data of others through their respective APIs, hence, using the data of others as metadata in ways unintended by those data providers. In other words, Europeana provides for surrogates of cultural heritage artefacts to be used in ways unimagined before but also provides for services and functionalities unimagined before.

All the developments outlined above are made possible only through the disintegration of the catalogue and the construction of metadata enriched and interoperable surrogate objects. It is a development from being a mere aggregator and cataloguer to becoming an information broker of cultural heritage. One can of course say that memory institutions have always been information brokers. However, that role was performed in a very different way. On the most basic level, Europeana manages Digital Surrogate Objects or, more generally, data; not books, paintings or documents. As will be argued in more detail in the following chapter, Europeana needs to be seen as a transversal memory institution rather than as a digital library and/or digital archive and/or digital museum.

Due to the disintegration of the catalogue, the descriptive metadata is not the index anymore but rather is being indexed in order to support various discovery

services. For instance, in the card catalogue, the descriptive metadata is filed in a fixed order separated from the object it describes. Europeana, on the other hand, moves beyond this traditional view by abandoning this separation. It constructs a new type of surrogate that does not merely stand for something. It is not a representative placeholder filed in a fixed order but a partial replication of it. As a consequence, Europeana does not become an information broker by means of its exclusive and non-ambiguous classification system but rather by doing the exact opposite; implementing interfaces for users (individuals, touristic service providers etc.) to order and discover cultural heritage artefacts according to their own rationales.

	EDL – Prototype (Cycle 1)	**Europeana (Cycle 2)**
Social Memory Paradigm	Memory Institution	Information Habitat
Paradigmatic Metaphor	Silos of Knowledge	Linked Open Data
Main Attributes	- Bound - Delimited	- Transversal - Networked
Main Role	Meta-Catalogue	Information Broker
Primary Incarnation	Portal	API
Scenarios of Use	Controlled Scenarios • Search and Browse • Access and Use • Personalization • Social Tagging	Uncontrolled Scenarios Examples: • Mash-Ups • Widgets • Etc.
User Groups	• End-Users • Re-Users	End-/Re-Users
Metadata Scheme	Europeana Semantic Elements (ESE) Collection Centred	Europeana Data Model (EDM) Object Centred
Surrogates	Descriptive Metadata	Partial Replications (Digital Surrogate Objects)
Description	• Classification • Controlled Vocabulary • Provided by EDL	• Contextualization • Not exclusively provided by Europeana
Discovery System	• Boolean Operators • Relevance Ranking	• Semantic Network • Mash-Ups beyond Europeana's Control

Table 17: Comparing the EDL-prototype and Europeana.

In more abstract terms, Europeana introduces contingency. The semantic layer shown in Figure 23 is a case in point, since it does not prescribe how users are supposed to find an item but rather allows them to browse according to concepts, which, to a significant degree, will be provided by external sources. This navigational semantic layer is not an enhanced catalogue but completely replaces the logic of the catalogue with networked ordering - contextualization instead of classification. It performs as the layer on top of the surrogate layer, which, in turn, represents the digitized cultural artefacts through partial replication.

Networked Contextualization
Digital Surrogate Object
Digitized Cultural Artefact

Table 18: Three layers of the Europeana information space.

Summarized in Table 17, it is this fundamental transition that allows for the changes analysed in this section - be it in terms of scenarios of use, discovery systems or the role of surrogates. It is the differences between the initial Europeana prototype (referred to as EDL) and the transversal Europeana built on the notion of the API that will allow Europeana to;

> "*provide a framework for Linked Open Data developments, seen as a key tool in digital innovation and the exploitation of knowledge* [as well as] *to move from a centralised model in which we* [Europeana] *gather data, to a distributed model in which we collaborate within a sustainable European information space. Enriching data and returning it to providers is one aspect of this; another is the distribution of Europeana data through our API and search widget, for use by partners and in educational and cultural sites* […]" (EP-005:4).

7. DISCUSSION

7.1 Introduction

In this chapter, the notion of transversality will be further conceptualized as a framework distinct from the paradigm of silos of knowledge and the storage metaphor. Europeana describes itself as an online service provider "*permitting a transversal approach to contents never before available"* (ES-014:12). Transversality will be observed as the primary distinction that makes Europeana observable as a phenomenon. Discussed in Chapter 4.4, thematic coding was applied as a means to develop a primary distinction rather than a core category that guides the analytical lens of the empirical research (Gibson et al. 2005). As the case study analysis illustrated, Europeana is in a process of transformation from the paradigm of;

> "*the digital information silos of the past* [to] *an information paradigm of linked data and thus for sharing as much semantic context as possible. Only in such a mental setting does the shift from the portal paradigm to the vision of an API as Europeana's primary incarnation truly make sense*" (EX-002:8-9).

The primary distinction emerged during the analysis as a temporal distinction of before/after following the two cycles of corpus construction – a distinction expressed as portal/API or silo/network.

According to the Oxford Online Dictionary, the adjective "transverse" refers to "*situated or extending across something*."[54] By the same token, a transversal line denotes a line "*cutting a system of lines*"[55] in geometry – a term also used in multidimensional topology (see Figure 31). Based on this understanding, Europeana will be discussed as an instance for transversal social memory that cuts across traditional organizational and institutional boundaries based on a shared layer of data. However, just like the set of lines continue after being crossed by the transversal line, so do the processes of canonization (see Chapter 2.4.1). In-formed into the binary-based medium of data (the transversal line or

[54] http://oxforddictionaries.com/definition/transverse, last access: 30th June 2011.
[55] http://oxforddictionaries.com/definition/transversal, last access: 30th June 2011.

layer), the “old” artefacts and approaches do not simply disappear but are rather reintroduced. Books, for instance, continue to exist as digital books (selection, preservation), just like exhibitions as virtual exhibitions (order), complementing the new, digital born artefacts and methods of algorithmic ordering and social tagging. In other words, the “old” is re-membered (i.e. transformed and made fit) as emulations in the “new”.

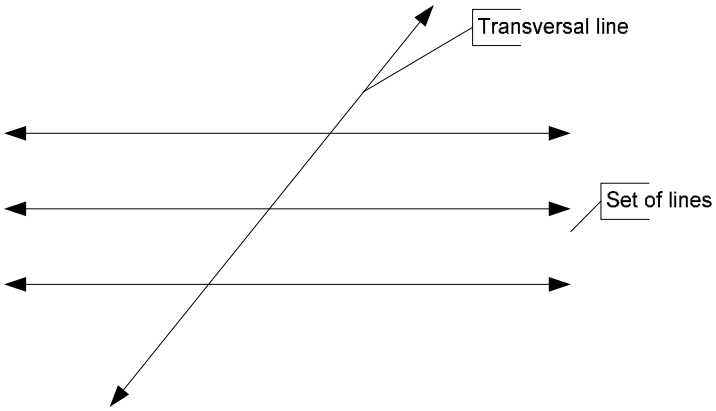


Figure 31: Example of a transversal line.

Binary based data, in this respect, is a foundational layer that does not only provide for possibilities to come up with innovative ways of content presentation, accessibility and navigability such as mash-ups or widgets but also for the reinvention of the traditional ways of discovery. The difference between the portal-based EDL and the API-based Europeana does not derive from the breaking down of the boundaries between libraries, archives and museums alone. As it will be discussed in this chapter, the difference rather derives from the construction of a platform that goes beyond the domain of memory institutions as such. Transversal, therefore, can be contrasted with universal in the sense that memory institutions used to be differentiated into separate, closed universes. For instance, the domain of librarianship used to be a distinct universe of practices, standards and specific types of collectibles in contrast to the universes of museums and archives. Europeana also breaks down the borders of memory institutions in toto by situating and extending its services into the WWW.

Thus perceived, transversality is not to be seen as convergence. True, the digitization of cultural heritage artefacts ultimately converges into bits and bytes,

which are not distinguishable from any other bits and bytes as such. Convergence, however, is followed by the divergence of bits and bytes into recognizable emulations of the original cultural heritage artefacts by virtue of descriptive and structural metadata (Herzhoff 2009). The same applies to the emulation of the old ways of first- and second-order of order, which continue to exist alongside the third-order of order of search engines and social tagging (Esposito 2002; Weinberger 2007).

In what follows, the concept of transversality will be discussed with respect to digitized cultural heritage artefacts first. The second section will extend the concept to all three layers of the Europeana information space, which share the common transversal layer of data. As a result, social memory in the information habitat will be defined as "forgetting-as-data". Equipped with the concepts of transversal memory and forgetting-as-data, the institutionalized practices of canonization will be revisited followed by a re-evaluation and critique of the storage metaphor and the distinctions it affords as being inadequate for the observation of the contemporary dynamics of forgetting and memory.

7.2 Transversal Memory Organization

One of the defining characteristics of the paradigm of memory institutions (see Chapter 2.4.1) is the internal differentiation into the organizational forms of libraries, archives and museums. Rooted in the rise of mass media, libraries have emerged as caretakers of mass produced media, while archives have remained dedicated to the documentation of evidence and museums to the caretaking of unique or rare artefacts. Suffice to say, these analytical distinctions become less clear when applied to actual organizations. For instance, the Austrian National Library (ANL) is partly a research library for the humanities, a reference library and archive for Austrian literature as well as a museum of valuable globes and manuscripts (Marton 2007).

However, these distinctions remain valid even within a single organizational entity. For instance, this can be observed in the ways the artefacts are treated.

While a book of ANL's humanities collection can be borrowed without a problem, it may be more difficult to borrow a valuable globe from the ANL. By the same token, a book is treated differently in terms of preservation compared to the conservation of a globe. Therefore, even when one organization is partly library and partly museum those parts are compartmentalized silos within the organization. In other words, the organization is split due to the differences between the types of cultural artefacts collected; that is the distinction between copy and original.

Going back to the FRBR categories of work, expression, manifestation and item (see Chapter 2.4.1), the question arises as to how to categorize digital artefacts within this framework. Is a digitized e-book an expression or an item? Does it even fit into this scheme?

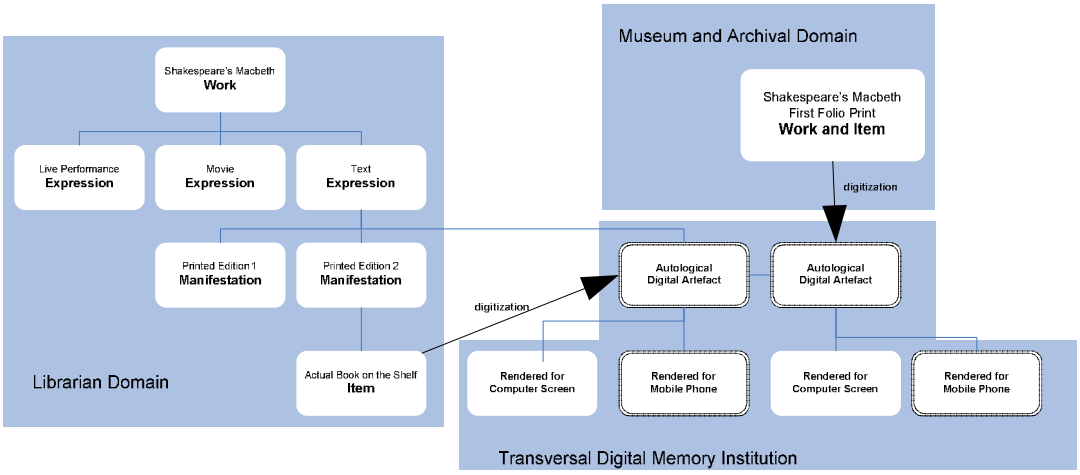


Figure 32: Digital cultural artefacts in memory institutions.

In a sense, digitized cultural heritage artefacts can be seen as a bridge between the traditionally separated domains of libraries, archives and museums. The librarian FRBR framework emerged out of the distinction of the original and its copy – a distinction that became clear through the rise of mass produced print media and the caretaking of one copy or few copies out of many. By contrast, a museum never had to make a distinction between work and item because of the status of its artefacts as being unique or rare. In the digital domain, however, the distinction between original and copy does not make any sense anymore since digital items defy the logic of distinguishing between unique or rare items,

collected by archives and museums, and mass produced communication media collected by libraries. The reasons are widely known and lie in the binary codification allowing for perfect copies and non-rival goods (Benkler 2006; Kallinikos et al. 2010a; Faulkner and Runde 2011).

However, the distinction between work and item is still useful. Separated from the original/copy distinction, it regains a more central role within the whole field of digitized cultural heritage. The example of Figure 32 shows how one copy of a mass print edition and the rare first folio print of Shakespeare's Macbeth end up being digitized autological cultural artefacts. In this scenario, the work Macbeth – the abstract category – still exists but finds its textual manifestation in digitized artefacts, which are nothing more than instructions on how data is to be interpreted – the textual expression of the work Macbeth is manifested in code. The same applies to the print folio digitized by a book museum or archive. The digitized folio is manifested in binary code the same way as its mass printed counterpart. In other words, the museum artefact used to be work and item in one. That is not the case with a digitized museum artefact anymore. Since the artefact can be used in many different ways (e.g. high quality and low quality scans for computer screens or mobile phones), it becomes an abstract category for all those renditions. It becomes a work separated from its items or rather the computational equivalences to items.

The librarian as well as the museum artefact lose their respective status as either being one copy of many or being valuable due to being rare and gain a new status as an autological digital cultural artefact – an artefact that instructs complementary software applications by means of describing itself as being a digitized book, a digital painting, an e-document and so forth. On the level of the item, the changes reach even further. The digitization of the cultural heritage artefact results in the dissolution of the item as an entity and replaces it with emulations of items according to the instructions encoded as structural metadata and the end-device used for access. For instance, a mass printed contemporary edition of Macbeth is the same manifestation emulated in different ways – be it a high quality image scan version for a computer screen or a low quality OCR

version for the mobile phone. Again, the same applies for the rare print folio or rather its digitized version.

Since the item is being dissolved into binary-based code and only "exists" as a momentary rendition when the code is being executed by applications/devices, the work/item relationship needs to be reconceptualised. As discussed in Chapter 2.4.1, a traditional library orders works and items (Svenonius 2000). Based on what has been discussed in this chapter so far, it would be a simple conclusion to just exchange "item" with "emulation" and declare that digital memory institutions order works and emulations. However, the analysis showed that digital cultural heritage artefacts are abstract containers defined and autologically described by metadata. They do not really fit into the FRBR framework comfortably since they do not follow the logic of mass production of copies. As a consequence, the categories closely linked to the mass production of media – the categories of manifestation and item – need to be replaced by a level of digital cultural artefacts.

In this sense, order or rather ordering refers to works and digital artefacts, which, in turn, have two modes of "existence"; codified instructions and emulations. All of these categories - works, expressions and digital cultural heritage artefacts (instructions and emulations) - are mediated as binary-based data and metadata. However, this framework cuts across the traditionally separated organizational forms of libraries, archives and museums. In other words, the relationship between works and digital cultural heritage artefacts is transversal, since, in the digital domain, libraries, archives and museums order the same kind of "thing". In fact, it is quite misleading to call these emerging online services digital libraries, digital archives and digital museums respectively. These distinctions are based on the modern concepts of original and copy and do not fit into the digital domain. In this respect, Europeana is a revelatory case study as it does not follow a specific trajectory rooted in one of these domains anymore (ES-029:6). Europeana does not order either books or paintings but digital artefacts.

It is these organizational forms, which are based on the order or rather ordering of digital artefacts for long-term access, usage and preservation, which can be

called transversal memory organization. They are transversal because the artefacts they order cut across all types and genres of communication media and types of artefacts alike. They remain transversal because those types and genres still exist as emulations. That is, memory institutions still take care of digitized books, digital paintings, software packages, web-pages and so forth. However, these are all based on data autologically described as digital cultural heritage artefacts. Sharing the same transversal foundation of bits and bytes, they are also maintained as *cultural* artefacts.

The hierarchy of work, expression, manifestation and item is replaced by a new hierarchy of work, expression and digital cultural artefact, which is applicable to all domains – be it library, archive or museum – due to the mediation in bits and bytes. However, cultural artefacts are much more than just data – just like a book is more than just mere text. The emulation of that data as a recognizable cultural artefact is, therefore, an important aspect and relies exclusively on the structural metadata of the cultural artefact it describes. It is the structural metadata that instructs the emulation of Macbeth as an e-book for an e-book reader, a mobile phone or a regular computer screen. The book that used to be the item, the actual book on the shelf, has become another abstract category, a container for computer files, descriptions and instructions. In more general terms, the distinctions in terms of expressions (text, movie, recording of a live performance etc.) or artefact types (book, painting etc.) are reintroduced as emulations after the cultural artefact has been digitized into a series of 0s and 1s.

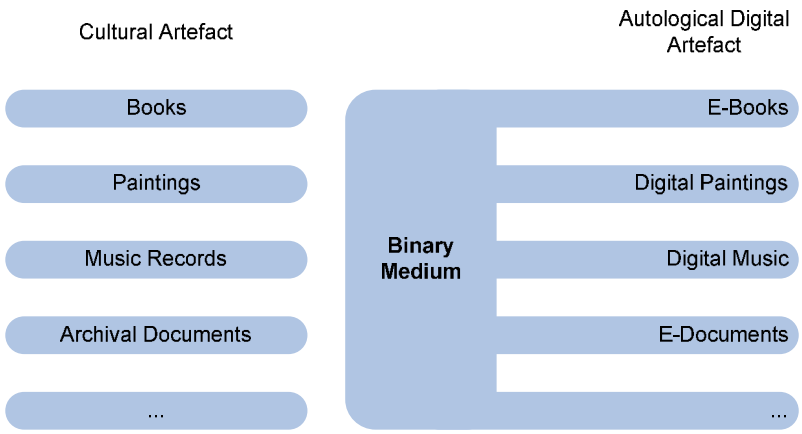


Figure 33: Transversality of digital cultural artefacts.

Europeana is a case in point as it, first, strictly separates the "real" material heritage artefact from the digitized version – the Digital Representation Object (DRO) – and, second, only deals with the DROs and their partial replications through Europeana's own Digital Surrogate Objects (DSO) (see Chapter 6.4). In this respect, the organization of social memory is transversal – it cuts across the former silos of knowledge and institutional boundaries.

Transversal is not to be understood as convergence, since books, paintings, documents, photographs and so forth do not converge into a universal type of artefact – at least not in terms of cultural heritage artefacts and institutionalized memory (Earnshaw and Vince 2008; Herzhoff 2009). The reason lies in the fact that memory institutions do not simply store things but cultural forms such as works and items. In an equivalent fashion, digital memory institutions do not store data but rather works and digital cultural heritage artefacts, which need to be accessible; their potentiality for emulation as e-books, digital paintings and so forth needs to be sustained. That is a very different task compared to the maintenance of a data-base. Memory institutions are not warehouses storing things or the computational equivalent of storing data. As discussed in Chapter 2.4, they are institutions of forgetting.

To conclude, Europeana orders digital artefacts, which cut across existing categories of expressions, genres and artefact types. It is only in this sense, that Europeana can be described as a transversal social memory institution rather than as a digital library and/or digital archive and/or digital museum. However, the question remains whether Europeana performs as an institution of forgetting through a process of canonization defined in Chapter 2.4.1. The following sections will be dedicated to the exploration of that question. First, the discussion will focus on the dynamics of remembering and forgetting within Europeana, followed by a discussion of whether digitized cultural artefacts are in fact digitized cultural *heritage* artefacts.

7.3 Transversal Forgetting

The increased involvement of internet services and resources into the practices of memory institutions as well as the mass-digitization of cultural heritage artefacts unfolds the relationship between remembering and forgetting on three levels; the cultural artefacts themselves, the surrogates they are represented by and the networked data community they are a part of (see Table 18). On the level of the items, the cultural heritage artefacts are forgotten as data, since an item is only a logical entity - a category in itself. As discussed in Chapter 2.2.3, categorization is indeed a way of social forgetting as the details and singularities of an event or instance, which do not fit into a category, are filtered and, thus, forgotten. Categorization constructs sameness into different events or instances by making them comparable along defined dimensions and, thus, assigning them to the same class or category. With digital cultural artefacts, the parameters are not only encoded as descriptive metadata categorizing the artefacts as, for instance, paintings or monographs, belonging to a subject area or as being created by the same creator in a certain time period. In addition, the parameters also entail structural metadata according to which items are disassembled into data and, ultimately, the binary classification of 0 and 1 only to be reassembled into a recognizable emulation of a cultural artefact. Hence, digital cultural artefacts are not only forgotten as instances of various classification systems but also as data. What is remembered are only object models – blueprints of instructions encoded as metadata on how to assemble an identifiable cultural heritage artefact out of data. In the digital domain, memory institutions do not collect cultural heritage artefacts but actively construct them. After all, a library user wants to find a book – even if it is only an emulation of a book – and not data. Lose the object models and instructions and the data turns into mere noise. The data loses its potentiality to inform.

As discussed further above, the card catalogue emerged as a second-order of order of things (Weinberger 2007). In opposition to its predecessor – the book catalogue – the card catalogue is able to include metadata, at least in principle, from an unlimited number of items. The order of the cultural artefacts shifts from an order of the objects themselves to an order of the representations or surrogates

of the objects – the metadata that fit on a single card. In simpler terms, institutionalized social memory becomes the catalogue, with the exception of open-shelf public libraries and museum exhibitions. The catalogue exemplifies an increased level of abstraction in terms of the operation of selecting what is remembered and forgetting the rest. The artefacts themselves are forgotten, that is, they are not arranged in a way to allow for their findability but rather to save storage space. What is being remembered is what fits on a catalogue card forming an abstraction of the repository with very specific avenues for discovery.

The WWW, being on the verge of becoming the next primary communication technology, is memorized in an even more abstract way. Its navigability is not provided by catalogues but rather by search engines and increasingly by social tagging (Weinberger 2007). Especially with search engines one can witness a restructuring of social memory. The navigation through the online information space is guided by the search results page, which is created, based on algorithmic calculations, for a specific user every time a search query is processed. Hence, with search engines social memory does forget fixed categorizations and the selective, persistent avenues paved for information discovery. Instead, it is enough if the algorithms are remembered (Esposito 2002).

The abstraction of the repository based on representational properties of a fixed order is exchanged by a higher degree of abstraction based on performative ordering only to be forgotten the moment the ordering is abandoned – the moment the results page is closed. This is not to say that the ordering is lost but rather it is stored as data. In terms of memory, the contemporary challenge is to make petabytes of data informative by means of second-order technologies (Gantz, Chute et al. 2008). Thus conceived, data mining tools or online search engine services are, in fact, technologies of remembering. They reconstruct events stripped of their singularity by being categorized and stored as data based on the very parameters according to which they were collected and stored in the first place.

A book comes as an object with physical structures and borders enabling its usability by means of its material make-up. Digitized books, on the other hand,

are only logical constructs, their borders and structure in need of being created and maintained by means of information technology. In more abstract terms, cultural artefacts are actively constructed and "entified" only when accessed. Europeana is a case in point as it also disassembles the catalogue into Digital Surrogate Objects. Hence it can organize the immense amount of digitized cultural artefacts stored at the providers' site by simply filtering most of their attributes as noise. Europeana radicalized this notion by making the descriptive metadata into a partial replication of the artefact it describes. This is the second level on which the relationship between remembering and forgetting is unfolded. It is not only the digitized cultural artefact (e.g. the digitized manuscript) that is forgotten as data but also its surrogate used to describe the artefact it represents.

Metadata – be it descriptive metadata linked to the Root Component or structural metadata of complex digital objects represented by compound surrogates – is therefore forgotten as data as well. A part of the surrogate does not have any meaning on its own but rather gains its functionality due to its relationship to other nodes. Being a logical entity rather than a copy of another artefact, the surrogate only identifies certain aspects such as the location and mirrors the structural make-up of what it represents qualified by a limited set of categories. The rest is forgotten as data. This notion very much fits into the performative and momentary ordering via online search engines described above. From the perspective of navigating the WWW, it is not the content of a web-page that matters but rather how it is indexed and ranked according to the algorithmic rules and procedures of the service provider.

By the same token, the instructions that actually make up a web-page are left for browsers to be interpreted and displayed. In a similar fashion, the cultural artefact is made up of instructions that need to be processed by software first in order to be presented to a user. Given the autological nature of digital artefacts, the memory institution needs to take appropriate steps in order to guarantee the integrity of its data but also of its logical objects it is dedicated to preserve. What is actually preserved by Europeana is, first of all, the surrogate model (see Figure 21) that is the blueprint of how the various parts of the surrogate are to be linked.

It is this schematic model that provides the parameters according to which surrogates are being constructed out of data and, therefore, remembered.

Complementing the level of the artefact and Europeana's surrogate object, the third level addresses the notion of findability and discovery of cultural heritage artefacts traditionally provided by a catalogue. In the case of Europeana, the catalogue is disintegrated into surrogate objects and replaced by a navigational layer based on semantic web technologies. Providing a context for the surrogates, that layer is, however, not under the total control of Europeana but rather integrated into the wider information habitat of linked data providers by means of APIs. As a result, the concepts according to which users will navigate through the Europeana information space are provided by a variety of sources not necessarily coming from memory institutions. This is a remarkable development, since Europeana gives up a significant degree of control over how its surrogates are being found and used by individual users as well as third parties such as social network sites or touristic portals. The semantic layer shown in Figure 23 does not prescribe how users are to find an item but rather allows them to browse according to concepts, which, to a significant degree, will be provided by external sources. This navigational semantic layer is not an enhanced catalogue but completely replaces the logic of the catalogue with networked ordering.

Europeana actively pursues the abolishment of the traditional card catalogue – a notion already accomplished by algorithmic search engines with respect to the WWW. True, a user may still search through the Europeana repository via the portal www.europeana.eu. However, the primary focus is on providing potentialities for re-use and mash-ups by means of APIs. Be it the semantic network layer or the embedding of a surrogate in a, say, social network site; all these potentialities rely on APIs. This is, again, a remarkable step, since third parties will have access to the data and not to cultural heritage artefacts. In other words, Europeana forgets the fixed order of the catalogue and only remembers the links and relations to other data providers, which help to contextualize the cultural artefact rather than to categorize it once and for all (see Figure 28).

As Figure 30 illustrated, Europeana's efforts revolve around objects rather than collections constructing context for its surrogates. With respect to remembering and forgetting, contextualization of Europeana's DSOs is indeed the equivalent to categorization. Categorization according to a classification system makes sense on the aggregate level of the collection. A classification system describes the collection in terms of specific parameters (e.g. name of the creator, genre, type or subject keywords) according to which the specific items are categorized. However, the way Europeana is going about its collection, those parameters are translated into links between each and every surrogate and to additional descriptive elements. In turn, these elements are linked to even further descriptive elements leading to a network without an underlying rationale – without a "kata logos". Instead of remembering the "kata logos" what is remembered are the qualified links between the elements that ultimately make up the surrogate object, the proxies, the aggregations and the linked data community.

Just as data and metadata are networked, so is remembering and forgetting. Europeana forgets the cultural heritage artefact as well as the fixed rules of cataloguing as data. What is remembered, on the one hand, is the object model – the computational instructions of emulating cultural heritage artefacts. Lose the object model and all the data dissolves into a sea of noise. On the other hand, the organization of the data and, therefore, the findability of items are provided by a consortium of interoperable information service providers. Lose the links to other nodes in the network and the discovery system dissolves into mere noise as well. Since Europeana does not solely decide what keywords are to be used to describe, say, the content of a book, it does not classify its items in the strict sense of the word but rather coordinates the contextualization of its surrogates by seamlessly integrating the Europeana information space into the wider information habitat. Hence, it is crucial which data provider the surrogate space is linked with since Europeana does not exclusively control the vocabulary used for description. Findability and discovery is, therefore, distributed throughout the internet. Forgetting through cataloguing is replaced by networked contextualization.

Unit of Analysis	Europeana Case Study	Forgotten as…	What is remembered?
Digital Cultural Artefact	Content Provider	Data	Object Models
Digital Surrogate Object	Europeana	Data	Surrogate Object Models
Networked Contextualization	Linked Open Data	Data	Qualified Links

Table 19: Three levels of forgetting and remembering.

The three levels, summarized in Table 19, come together in terms of binary-based data or, to be precise, in terms of the forgetting of the respective unit of analysis as data. What they all share is the mode of forgetting-as-data, what differentiates them is what is remembered on each level. In this sense, data transverses across all three levels presenting forgetting-as-data as a transversal phenomenon, which builds a new foundation upon which operations of remembering can be built. Those operations of remembering comprise a vast variety of possibilities reaching from the traditional OPAC scenario of search and access, relevance ranking based on search engine algorithms, social tagging and folksonomies, semantic browsing, widgets, mash-ups and so forth. The variety is possible *because* data is transversal and provides potentialities for these kinds of operations but also for operations yet to be invented. In this sense, transversal forgetting can be contrasted with the traditional ways of forgetting bounded to the "silos of knowledge" and delimited by the respective standards of the domains of libraries, archives and museums forming separated universes of knowledge.

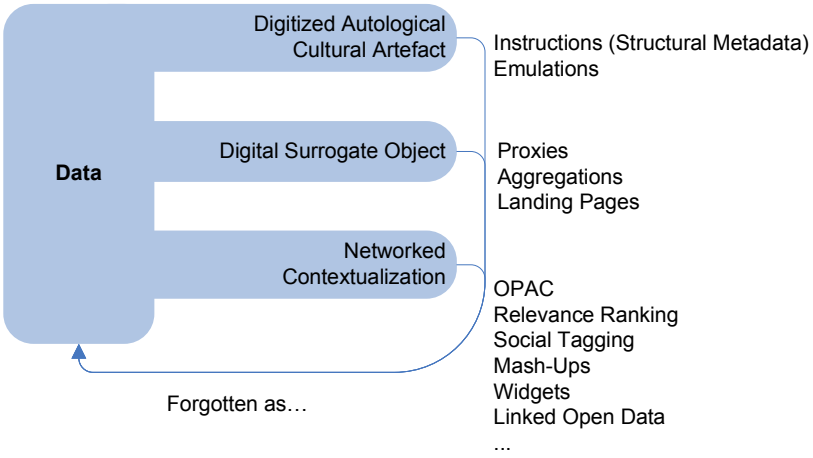


Figure 34: Transversal forgetting and remembering.

To conclude, Europeana does develop a discovery system but it is not the exclusive system according to which cultural artefacts are categorized. Instead, it contextualizes its Digital Surrogate Objects (DSO) in order to allow third-order of orderings – be it performed by Europeana itself, commercial search engines or by independent apps developers. To a considerable degree, the context is linked to external providers whose data is used as metadata. Therefore, the distinction between data and metadata or, on this level, DSO and context comes forth due to the relationship the one has to the other. For instance, data provided by geonames does not become metadata by virtue of some essential attribute of the data. Geonames does not provide metadata. It is Europeana that links to geonames and defines their data as context and therefore descriptive metadata. By the same token, another data provider could use Europeana's data as metadata. In this sense, it is indeed more accurate to use the term metadata, as it has been the case in Library and Information Science for some time now. The term cataloguing, in comparison, does not invoke this relationship that one's data can be another's metadata and vice versa.

7.4 Memory Institutions Revisited

The ideas put forward above suggest that society shifts towards a memory paradigm of ordering rather than static order. In this sense, search engines present a momentary catalogue on-the-fly only to forget it as data the moment the results page is closed. This notion is complemented by social tagging, which results in online folksonomies bearing the potential of becoming a complementary alternative to professional taxonomies. As Weinberger (2007:102) puts it, the filter is not on the way in but on the way out.

Memory institutions are stepping into this online world of digitality and ex-post performative ordering. However, as the case of Europeana shows, libraries, archives and museums do not simply become search engines or mere platforms for social tagging themselves. The situation presents itself as more complex since digitization initiatives within the domain of memory institutions deal with cultural heritage artefacts rather than just data. From this perspective, Europeana

is an extreme case as it highlights how far initiatives of this kind can go and how far the binary-based mediation of data reaches in terms of social memory, remembering and forgetting. In the previous section, the argumentation proposed a transversal concept of forgetting based on bits and bytes and data. However, the question arises whether Europeana or rather the potentialities it provides can be seen as canonization?

To begin with, the initial research question – how do libraries change in order to remain libraries – turns out to be of little help. As the analysis showed, libraries do not simply turn into digital libraries but rather transverse the domains of libraries, archives and museums due to the digital nature of the data and the emulations of cultural artefacts they take care of. They do not deal with books, paintings, photographs and so forth but with digital cultural artefacts that are logical objects of instructions and emulated as e-books, digital paintings, digital photographs and so forth.

Hence, the research question changed into "how do memory institutions change in order to remain memory institutions?" In other words, does Europeana perform the distinguishing operation of memory institutions; does Europeana canonize? As it was discussed in Chapter 2.4, canonization was defined as the institutionalized process of forgetting by means of selection, order and preservation, a process that turns cultural artefacts into cultural *heritage* artefacts. In what follows, each of the three processes will be addressed and discussed with respect to the case study. Section 7.5 will then discuss the notion of canon and the storage metaphor (Assmann 2008a).

7.4.1 Selection

The selection of artefacts to become part of a canon is the most obvious operation of forgetting. In case of Europeana, however, one can hardly speak of selection because of two reasons. First of all, the artefact that ends up being described and represented in Europeana's information space is already a cultural heritage artefact – it has already been selected. In contrast to Google's Book

Project, for instance, artefacts are being digitized *because* they have already been canonized, because they are already part of a librarian, archival or museum collection. Second, the decision process of what to digitize and in what order, which is also a selection process, is not up to Europeana but rather to the content providers bounded by their respective budgetary and logistical constraints. The filter is, therefore, with the content providers and with the aggregators specifically and, from their perspectives, remains on the way in rather than on the way out (EE-037:4; ES-001:14).

Europeana itself does not have a filtering function by means of selection. Quite the opposite, it tries to represent as many cultural heritage artefacts from as many European countries as possible. From Europeana's perspective, Weinberger's (2007) notion that, in the digital domain, with the rise of the third-order of order the filter moved from the way in to the way out, can be supported to some degree. In other words, the selection of what comes in and what stays out loses the key role it had played as a mnemonic process in the social memory paradigm of second-order of order. Instead, the key role of filtering and, therefore, forgetting is taken over by ordering; by momentary renditions of order, to be precise.

However, Europeana is a special case, since it deals exclusively with digitized cultural heritage artefacts. It does not address the question of canonizing born-digital and, specifically, online documents or artefacts. As preliminary research on the Internet Archive (www.archive.org) reported elsewhere (Kallinikos et al. 2010a) indicates, efforts to archive the WWW indeed have a selection mechanism separating websites to be archived from the ones which are not. In case of the Internet Archive, the selection is based on algorithmic rankings performed by the Alexa search engine (see Chapter 2.4.5). The reason for a selection of this kind lies in the notion of preservation – the third canonizing process. While ordering "everything" is computationally feasible, preserving "everything" is not. As discussed in more detail further below, the preservation of a cultural heritage artefact is not the same as storing data as bits and bytes. Hence, memory institutions, in contrast to search engines for instance, still need to select when it comes to the collection of born-digital cultural artefacts.

As a consequence, selection is still an issue in terms of the digitization of cultural artefacts and the preservation of born-digital cultural artefacts. Europeana, on the other hand, is not a case for either of these two issues since selection through digitization is left to the content providers and the preservation of digital artefacts is left to external projects such as PrestoPrime (see Figure 8). However, selection through digitization does not mean that the "old" media are replaced. As the analysis illustrated (see Figure 33 for instance), the old media are re-established by means of their emulation. Books, paintings, films and so forth still exist but are, as Bolter and Grusin (2001) put it, *remediated* by binary-based digital media based on the intricacies and peculiarities of data-based operations and computational processes. In other words, books do not disappear but are dis-membered into bits and bytes in order to be re-membered as emulations of books (Garde-Hansen et al. 2009a). Once they are digitized or born digital, they are forgotten as data.

7.4.2 Order

Obviously, Europeana is exclusively focused on the notion of bringing order to the millions and millions of digitized cultural heritage artefacts enabling a wide variety of discovery and access possibilities. In this sense, order is fundamentally linked to finding and accessing items and, ultimately, using them. By the same token, the analysis also showed the tremendous changes that memory institutions are going through in order to accommodate the information habitat. Europeana has proven itself to be revelatory in this respect since it is explicitly promoting a paradigmatic shift from the silos of knowledge and ex-ante order to transversal memory organization and ex-post ordering.

As discussed in Chapter 2.4.1, the aim of a bibliographic discovery system is to find, identify, select and obtain entities or artefacts (IFLA 1998). The basic question is how to associate one artefact with another in order to enable a user to discover what is in the collection. This fundamental notion applies to an open shelf arrangement in a library as well as an exhibition in a museum discovered by means of walking through the collection. It applies to the arrangement of the

representations of things – the surrogates – which allows, for instance, a researcher to discover historical evidence in an archive or an ancient artefact conserved by a museum. In a sense, libraries, archives and museums can be seen as functional equivalences in terms of how the aims of finding, identifying, selecting and obtaining are achieved.

Clearly, these objectives are not only valid in the digital domain but more important than ever before. To some degree, the equivalences are even being reconstructed and emulated. Based on the "same" pool of data and computational operations, virtual exhibitions are created following the curatorial tradition of museums by offering a very different option for discovery than the search functionality of a digital library portal. A telling example is Google's recent arts project (www.googleartproject.com). In cooperation with leading arts museums such as Tate Britain, the Van Gogh Museum and many more, Google offers a virtual tour through the partner museums via Google's streetview technology and access to digitized versions of the exhibited paintings in incredibly high resolution. In this scenario, the associations between the individual paintings are constructed in a very different way than, say, through the standard Google search engine index algorithm. However, both rely on binary-based data that function as a transversal layer on a fundamental level.

The associations established in a first-order of order – the order of the things themselves – rely on the spatial proximity between what is to be ordered. By contrast, the second-order of order – the order of surrogates – establishes associations based on abstract categorization rules and descriptions, which apply for the whole collection and, ultimately, culminate into a catalogue (rather than just a mere inventory list). With the third-order of ordering the focus shifts from the collection to the objects that make up the collection. In case of Europeana, those are the Digital Surrogate Objects. Embedded into contextualizations[!] instead of categorizations, order presents itself only as a potentiality that needs to be actualized into a momentary ordering be it through Europeana's own portal-based search, ranking, semantic browsing and social tagging functionalities, external search engine harvesting or discovery services developed by external developers based on Europeana's API.

The associations between the artefacts or, as it is the case with Europeana, its representations are not "hard-wired" into the discovery system; they exist as potentialities, which, due to the high level of granularity afforded by binary-based data (Kallinikos et al. 2010a; Yoo et al. 2010), can be actualized in a variety of ways and forms reaching from traditional digital library OPACs to experimental and uncontrolled mash-ups. The key novelty is the rise of contextualization and the decline of exclusive classification systems. For instance, social tagging, one of the paradigmatic examples for the third-order of ordering (Weinberger 2007), is indeed only one way of contextualizing cultural artefacts conducted by lay users and amateurs. It is, however, not the only possibility. Thus conceived, the often invoked primacy of search engines and social tagging turns out to be only two alternatives among numerous other alternatives of discovery in memory institutions. Similar to the digitization of cultural heritage artefacts, which are re-membered as emulations of recognizable artefacts, the "old" ways of ordering are not replaced or abandoned. Forgotten as (meta)data, they are re-membered according to the present intricacies and peculiarities of the information habitat and computational operations. In other words, the third-order of order does not play an exclusive role in the information habitat but rather allows for the re-membering of first- and second-order of order as well.

Based on these observations, Europeana is indeed ordering. However, the kind of order it provides is different from a catalogue but equivalent in terms of achieving associations between the Digital Surrogate Objects. Notably, Europeana is not only ordering data but also cultural heritage artefacts. It provides for the potentiality for ordering works and digital artefacts, which is not the same as, say, Google indexing web sites or Amazon recommending books to a customer. In more general terms, a digital memory institution may use search engine algorithms or recommendation systems, it does not necessarily mean that it turns into a search engine.

The case study analysis shows in detail that the transition from cataloguing to contextualizing as the primary way of ordering is a wide-reaching change in terms of social memory. Being a transition from fixed order to potential

orderings, this change is very closely aligned with the new paradigm of the information habitat. However, as Europeana illustrates, this transition is not necessarily the end of libraries, archives and museums. Quite the opposite, these institutions re-incarnate as information spaces by means of emulating familiar discovery systems such as virtual exhibitions or new and innovative discovery systems such as semantic browsing. This increase in variety is based on the transversality of the digital artefacts, their surrogates and networked contextualization, which cut across the silos of knowledge once guarded by libraries, archives and museums but also across institutional boundaries. Instead of exclusive ex-ante order, the primary focus turns towards *transversal* ex-post ordering that re-members libraries, archives and museums by means of emulating their respective methods of creating associations between cultural artefacts.

7.4.3 Preservation

Long-term preservation or even conservation of digital cultural artefacts is not addressed by the Europeana initiative directly. The bigger question of how future generations will remember the past, based on what kind of documents, is indeed too big for one initiative to tackle. At this point, the dissertation can only point to further research conducted on this issue by projects such as the already mentioned PrestoPrime (www.prestoprime.org) or Planets (www.planets-project.eu), CASPAR (www.casparpreserves.eu), Digital Preservation Europe (DPE, www.digitalpreservationeurope.eu), the Internet Archive (www.archive.org) and many more. Europeana will benefit from these research programs but, up until today, is not directly involved.

As discussed in Chapter 2.4.5, long term preservation faces a variety of challenges including the maintenance of accessibility to digital content and the obsolescence of hard- and software standards – challenges that go beyond the mundane task of copying bits and bytes from one storage medium to another. Preservation is a problem that comes with information technology. Binary-based digitality is a fragile medium that requires various layers in order to be accessible and usable. At the most basic level, it is, after all, the problem of keeping data

potentially informative, which is a transversal problem just like binary-based data itself. Preservation indeed illustrates the notion of transversal forgetting more clearly than selection or order since it is approached as a "networked" problem. Entailing data and accessibility, preservation is necessary on all three layers of the memory institutional information space. The first two layers of artefacts and surrogates could be preserved "in-house" but the third layer of networked contextualization distributes preservation across the network of the LOD community, as it is the case with Europeana. Thus conceived, the preservation of contextualizations[!] cuts across not only the previous silos of knowledge and the domains of memory institutions but every involved data and metadata provider.

To conclude this section, canonization as the process of selection, order and preservation is indeed occurring. However, a canon is not exclusively constructed by memory institutions but rather in concert with other organizations, companies and research projects from beyond the field of memory institutions. Canonization, as a result, is shared across various stakeholders, networked through computational means and founded on the common layer of data. With respect to the case study, selection is conducted by the content providers. Ordering is only partially conducted by Europeana complemented by external search engines, metadata providers and so forth. Finally, preservation requires the collaboration of organizations within and beyond the field of memory institutions in order to guarantee the persistency of the cultural heritage artefacts as well as their findability. Thus conceived, canonization receives a different meaning with respect to its traditional counterparts – the canon and the historical archive, as referred to by Assmann (2008a).

7.5 Beyond the Storage Metaphor

The notion of memory is inherently paradoxical since it is directly related to self-observation (Esposito 2002). In an abstract sense, remembering and forgetting

delimit identities of self, nation, generation and so forth from the alien, foreign or, simply put, the others. In the sociological tradition, Halbwachs (1992) already pointed out that groups or collectives only remember what unites its members rather than what separates them. Through continuous reconstruction of the past in the present, an identity is remembered each and every moment in time, which also includes the member of a group re-membering her/himself into the very same group that enables remembering in the first place. By the same token, Jan and Aleida Assmann (Assmann 1995; Assmann 2008a) describe cultural heritage as a way for a society to observe itself. Institutionalized through libraries, archives and museums, artefacts go through a process of canonization receiving the status of cultural *heritage* artefacts being representative for a specific social formation.

Differences put aside, both approaches seem to be aware of the paradox of social memory - the faculty to remember already presupposes the faculty to remember. Or in more general terms, the faculty to observe and, thus, recognize oneself already presupposes the faculty to observe and, thus, recognize oneself. Thus conceived, memory is closely associated with self-awareness. However, it is also self-referential since the recognition of the self by the self also means that the self recognizes that it recognizes itself. An exhibition of national treasures by a national museum, for instance, is self-referential since a nation represents itself within itself. Discussed as re-entry in Chapter 3, the inevitable conclusion is an infinite regress of representations within representations within representations ad infinitum – a paradox (Bussola 2011).

The paradox of memory has been addressed in quite different ways by the conceptual frameworks discussed during the course of this dissertation. For Halbwachs, there is no difference between memory and remembering as individuals continuously reconstruct the past in the present. As a consequence, memory does not have an inherent persistence since no present moment in time is quite like the other. Memory, therefore, is conceptualized as a process of continuous remembering and change, which is the reason why social frameworks play such a central role in Halbwachs' collective memory as a structuring and stabilizing factor that, ultimately, constructs recursivity (Wetzel 2011). In this

sense, the collective does not represent itself within itself but rather in its members' individual memories. The paradox of social memory is transformed into and, thus, hidden as the distinction between individual and collective memory. The paradox, however, resurfaces as the conceptual inability to decide whether collective memory is a collected memory of individual memories or a collective phenomenon sui generis, as already discussed in Chapter 2 (Olick 1999).

Jan Assmann (2002) criticized Halbwachs' account of memory as being overtly "presentist" turning a blind eye towards the past as the past. By contrast, the concept of cultural memory was explicitly constructed in order to introduce persistence into social memory. Taking the form of physical artefacts and institutional practices, cultural memory is, therefore, not limited to a momentary present but is mediated over space and time reaching beyond the immediacy of Halbwachs' collective memory that only requires face-to-face interaction (Brockmeier 2002; Dijck 2007). Cultural memory becomes a storage of artefacts or rather a resource of reminders activated by rituals, exhibitions, documentaries and so forth (Luhmann 1998). The paradox of memory is transformed into and, thus, hidden as a process of information retrieval. It is a memory of objects that finds its clearest expression in Aleida Assmann's (2008a) distinction between canon and archive or, in more general terms, relevance and irrelevance. While the archive operates as the repository of irrelevant, hence, forgotten artefacts, the canon is what is relevant, hence, remembered - the active part of cultural memory. One cannot help but associate this distinction with the computational set-up of data being stored on a hard drive and loaded into the Random Access Memory for processing. As discussed in Chapter 3.5, this approach resembles Shannon's model of encoding, transmitting and decoding with the only exception that sender and receiver are one and the same (Bussola 2011).

The transformation of the paradox of memory – the representation of oneself within oneself – into the distinction of relevance/irrelevance requires institutional guardians entrusted with the authority to decide what is relevant as cultural heritage (canon) and what is not (archive). As a conceptual framework, this distinction is very useful with respect to mass media since institutions of

education (e.g. textbooks used for teaching), of diffusion (creators, publishers, broadcasting) and memory (libraries, archives, museums) can be pinpointed as gatekeepers producing or sustaining relevance (Seixas 2007). Especially the rise of television proves to be a fruitful research area since broadcasting fits well into the category of a canon (Hoskins 2009b). Television is the example par excellence of continuously reminding its viewers of what is relevant as canonical knowledge.

Canon and archive, however, do not sit well with digital media and the way they circulate through the internet. At this point, the following conclusion drawn by Aleida Assmann (2006:18), already quoted in Chapter 2.3.3, is worth repeating;

> "*In the culture of the new media, memory is more likely to be destroyed than constructed. The internet, as we all know, is a medium that provides an unlimited plethora of information without actually storing it.*"

This conclusion is very problematic for numerous reasons. For one, privacy and data protection advocates would most definitely disagree with the observation that the internet does not allow for the storage of data (Mayer-Schönberger 2009). Assmann's point, however, may not be about the mere storage of data but rather about online information not being archived as cultural heritage. The internet lacks the institutional formations to distinguish between relevance and irrelevance and thus canon and archive. As the dissertation illustrated, the distinction between canon and archive indeed disappears in the online world, but that does not necessarily mean that memory is being destroyed. With the rise of algorithmic search engines, relevance is created through ranking applied to basically everything published onto the internet and made available through the services the internet affords. Every web-page, tweet, blog, wiki and so forth being indexed by a search engine is potentially relevant, which, in turn, means that nothing is potentially irrelevant as such and therefore nothing is forgotten in an archive.

While for Halbwachs it is the individual that remembers and the social frameworks that forget, for Aleida and Jan Assmann it is the canon that remembers and the archive that forgets. Thus conceived, the concept of memory refers to remembering as that which is not forgotten and vice versa. In other

words, the paradox of memory is transformed into an either-or concept – one either remembers or forgets. However, as the analysis of the case study illustrated, memory is a cyclical operation emerging into a self-referential form - the more society forgets the more it remembers. In this sense, the self-referentiality of memory is reflected in the self-referentiality of remembering and forgetting. The storage metaphor of canon and archive is, hence, not a viable concept anymore since its translation into the digital domain as storing data and retrieving information as if they were things does not capture the intricacies of social memory based on momentary ordering and the construction of potentialities and emulations.

The concept of forgetting-as-data goes beyond the mere storage of data. As the case study illustrated, social forgetting and remembering requires a wide array of object models, structural and descriptive metadata schema, methods of providing for orderings and discovery as well as concerted efforts with regards to long-term preservation. The feedback loop between forgotten-as-data and, say, finding and reading an e-book is a dynamical process that needs to be maintained as such rather than only the raw material of bits and bytes. Data may be stored but, without the respective potentialities for their emulation as findable and usable cultural artefacts, they remain data. The continuous sustainment of data or rather of its potentiality to be rendered as cultural heritage artefacts feeds back into the data. Be it the description of the content or the structure of the cultural artefact, it is all, in the end, data that performs as metadata by virtue of how the various elements (digital artefact, surrogate, contextualization) are related to one another.

Thus conceived, canonization does not result in a relatively stable collection of classics ordered according to a stable rationale. Indeed, as Europeana explicitly states, the ordering of its items and therefore the collection of its items is partly passed on to third parties. Depending on the service used for discovery, users will find different orderings of the items and, therefore, find different items. This is the main point of the third-order of order; one does not have direct access to the collection as such but only to momentary orderings of it or rather instantiations of the collection. The collection and its ordering are separate. In this sense, persistence does not refer to the persistence of *what* is and is not in the

canon but rather to canonization as a continuous process of selection, ordering and preservation. Instead of seeing memory institutions as guardians of a canon, they should be seen as continuously canonizing without ever reaching a point of "something" being canonized. Europeana or any digital library, archive or museum is part of the canonizing process rather than the exclusive holder of a canon – a silo of canonized knowledge. Europeana, therefore, distinguishes itself from the information habitat as a persistent canonizer executing the continuous cycle of forgetting-as-data and remembering through emulations of cultural heritage artefacts as well as through providing potentialities for momentary renditions of order. In other words, Europeana does not hold cultural heritage but rather the ingredients for the (re-)computation and (re-)construction of cultural heritage.

Social memory is becoming a transversal operation that is continuously constructed instead of being retrieved. The construction is distributed across networks of computational operations and interoperable data-bases. Thus conceived, it is not libraries, archives and museums that *have* cultural memories. It is the information habitat that continuously remembers and forgets. Memory institutions, in turn, are being re-membered as computational operations as they are becoming a part of the information habitat. Memory is empty in the sense that there is nothing *in* memory as such since remembering is based on computational emulations and instantiations of cultural artefacts and orderings. One does not have cultural heritage artefacts first and orders them afterwards; rather one emulates and orders first in order to have cultural heritage artefacts.

In light of this argumentation, the memory paradox unfolds in a different manner in comparison to Halbwachs' individual/collective as well as Jan and Aleida Assmann's canon/archive distinction. Supported by system theoretical conceptualizations, the dissertation strongly argues that social memory is indeed a process of forgetting-as-data, be it canonization or the daily chatter of social networks. In other words, contemporary society observes itself through the continuous construction of itself in-formed into computational data. The paradox of memory is, therefore, not hidden by the either-or separation that one either remembers or forgets. The paradox unfolds as a new distinction of

data/information. Hence, contemporary society can indeed be seen as a memory/information society since it represents itself within the information constructed out of data – data as a difference which makes a difference. Social memory, therefore, cannot be treated as if it was a storage of knowledge or information as things. It is not a resource one can draw upon to reconstruct the past. Just as data needs to be constructed into a digital cultural artefact so does social memory in general. In other words, social memory needs to be informative - it is not a storage but a process of information transversing throughout the fabric of the information habitat and, ultimately, cutting through the distinct universes once separated into institutions and day-to-day living, canon and archive. Social memory is forgotten *as* data and remembered *through* information.

8. CONCLUSIONS

The present chapter concludes the research project and seeks to offer a platform for reflections and thoughts about the past and the future. To begin with, the next section summarizes the dissertation. The second section addresses the contributions made by the study with respect to 1) theory within the two domains of information systems research and social memory studies, 2) methodology and 3) practice. Followed by reflections on the limitations of the research design, the chapter closes with an exploration of possibilities for future research.

8.1 Summary of the Dissertation

The introductory chapter set out to explain the factors that motivated the study of the digitization of memory institutions. Finding its point of departure in the initial question as to what is happening to one of the oldest information-based institutions still in existence today – the library – the research has led to a systematic investigation of the evolving domain of digital cultural heritage and the role ICT plays in the ways society, through its memory institutions, will remember the past in the future – a field of study that has, so far, not been adequately investigated in information systems research and even social theory at large.

Chapter 2 critically reviewed the literature on the three major concepts of memory as a social phenomenon – collective, cultural and social memory. While the first two frameworks were deemed to be problematic with respect to the question of who or what the "carriers" of a societal memory are, an approach based on the theory of social systems was introduced as an alternative defining memory as a social operation of primarily forgetting and, only secondarily, remembering. Followed by the structuring of social memory concepts into the social memory paradigms of 1) a memory of things, 2) places for musing, 3) institutionalized memory and 4) the information habitat, the dissertation explored the various notions of digital memories in relation to the interconnected and

binary-based information habitat as a research domain. The chapter reviewed libraries, archives and museums as memory institutions engaged with the canonization (selecting, ordering, preserving) of cultural artefacts into cultural *heritage* artefacts. The initial exploration of social memory and the information habitat was complemented by the discussion of contemporary challenges and issues faced by memory institutions with a focus on digital libraries, metadata, copyright and digital preservation.

Finally, the review led to the formulation of the research interest addressing the change memory institutions are going through as they step from the paradigm of institutionalized memory to the paradigm of the information habitat. This approach was based on the following gaps identified in the literature; 1) the conceptual primacy of the storage metaphor in the field of cultural or social memory studies underestimates memory as a continuous social operation. 2) As a consequence, a reliance on IT and its capabilities for data storage and retrieval, seen as a remedy against forgetting, is a too simplistic view in order to address pressing questions about the future of cultural heritage. 3) A perspective declaring an end to libraries, archives and museums in the age of the internet ignores the transformation process memory institutions are going through in order to remain memory institutions.

In Chapter 3, the theoretical framework of the study was discussed based on the theory of social systems by Niklas Luhmann. Defined as making a distinction, the basic operation of observation was introduced with reference to information and communication. In this spirit, the foundation of a philosophy of difference (rather than essence) was laid leading to the concept of autopoietic systems, which are observable by others as well as by themselves based on their difference to their respective environments. Thus conceived, information occurs as a self-referential and systemic phenomenon against a backdrop of what is already known and expected. Hence, the form of memory/information was developed in contrast to the metaphor of information storage and retrieval. Informed by the theoretical framework, the research interests were translated into research questions focusing the empirical study on how memory institutions differentiate themselves from the new environment of the information habitat.

The concept of observation was further developed into a methodological approach in Chapter 4. The case was defined as being the difference to its context constructed by the researcher in order to collect relevant data addressing the research question. By the same token, the Europeana initiative was briefly introduced as a revelatory single case study on the digitization of cultural heritage artefacts by European libraries, archives and museums as well as on the innovative online services these digital artefacts afford. Based on a qualitative methodology, the three principles of corpus construction (relevance, homogeneity and synchronicity) were applied to guide the data collection. As a complementary method, thematic coding was used to develop a thematic map of the case and, ultimately, a primary distinction that makes the phenomenon of interest – Europeana - observable in difference to its environment – the information habitat. Finally, the chapter explained how issues concerning the quality criteria of qualitative research were approached as well as how the empirical study was documented in order to allow the reader of the dissertation to evaluate the research project.

Based on the methodological distinction of case/context, Chapter 5 introduced the context of Europeana with respect to its history, European Union politics, private-public-partnerships and organizational structure. By the same token, Chapter 6 presented a detailed analysis of the Europeana information space consisting of digitized cultural heritage artefacts, their representation through so-called Digital Surrogate Objects and their linkage to a wider network of contextualizations[!]. The analytical gaze was guided by the comparison of two distinct phases of the Europeana initiative. The first phase resulted in the creation of a prototype online portal, which was still based on the traditional paradigm of institutionalized memory as silos of knowledge. The second phase, however, explicitly broke with this traditional approach by abandoning the notions of catalogue and ex-ante classification systems in favour of providing potentialities for the re-use of Europeana's data by external service providers not exclusively from the field of memory institutions.

Finally, Chapter 7 discussed the analysis as examples for a transversal social memory – a social memory that is no longer structured into silos of knowledge.

In this sense, transversality refers to a basic layer of data that cuts across the previously separated universes of libraries, archives and museums. Forgotten as data, the canonization of cultural artefacts is re-membered into the information habitat by means of emulating recognizable cultural *heritage* artefacts. In this sense, the three aspects of canonization (selection, ordering, preservation) were discussed as transversal phenomena resulting in a cyclical, continuous process of remembering through emulations of cultural artefacts and their ordering, on the one hand, and the forgetting of these emulations and orderings as data. Thus conceived, the dissertation argued against the storage metaphor as a viable depiction of contemporary social memory. As an alternative, social memory was conceptualized as an operation of forgetting-as-data. In other words, society remembers itself in-formed into computational data.

8.2 Contributions

8.2.1 Theoretical Contributions

In the field of *information systems* (IS) research, the dissertation contributes the notion of memory as a relevant conceptual framework for the study of contemporary information and communication technologies. In the form of cultural and social memory studies, a new reference discipline is introduced to the community of IS researchers. In detail, the dissertation raises attention that information systems are also memory systems. As the research argues, however, memory is to be seen as more than the storage of binary-based data. Defined as the organization of the observation of information, memory complements information leading to forms such as redundancy/variety, expectation/surprise and, ultimately, memory/information. In this sense, these forms are not discussed as contradictions or as mutually exclusive categories. Conceived as two sides of the same coin, memory is necessary for information to occur and vice versa.

Against this backdrop, the conceptual framework of memory/information was introduced on various levels in order to demonstrate its explanatory power. For one, the brief excursion into the discourse on the information society in Chapter

3.8 was enriched by the discourse on the memory craze found in contemporary society. Thus conceived, the dissertation proposes the notion of a memory/information society. By the same token, the concept can be applied on the level of techniques, technology and technological artefacts. In other words, information technologies can be seen as memory/information technologies computing or constructing information out of data based on the very same categorization parameters according to which data were collected and stored in the first place. It is in this sense that singularities are forgotten as data and remembered through information.

A similar contribution can be formulated with respect to the field of *social memory* studies. In contrast to the often invoked storage metaphor, the dissertation proposes a metaphor that is based on computation and construction. In detail, the image of emulations is helpful as it refers to computational operations rather than packaged items. Be it the emulation of cultural heritage artefacts or the ordering thereof through virtual exhibitions, Open Public Access Catalogues (OPACs) or mash-ups, cultural heritage is presented as momentary renditions to be forgotten as data the moment the artefact or the ordering of artefacts is abandoned by a user.

Taken together, all the points raised above result in the central contribution of the dissertation. Termed *transversality*, the dissertation strongly argues against viewpoints that interpret contemporary developments related to the rise of the internet as 1) the destruction of memory, 2) the forgetting of forgetting or 3) the forgetting of remembering. Based on the analysis of the changes memory institutions are going through in order to remain memory institutions, the dissertation claims that the self-referential relationship between remembering and forgetting is restructured. In detail, the paradigm of "silos of knowledge", reflecting the separated and, to a significant degree, isolated universes of libraries, archives and museums, is contrasted with a paradigm of networking and transversality. In other words, social memory is mediated by a fundamental layer of binary-based data that transverses across previously established classificatory boundaries (e.g. communication genres or artefact types), traditional memory instutional arrangements (e.g. librarian cataloguing or

museum exhibitions) and, finally, beyond the field of memory institutions at large. These genres, types, institutional arrangements and so forth are forgotten as data; however, they are re-membered as emulations – be it emulations of recognizable cultural heritage artefacts, of museum exhibitions and so forth.

In other words, memory institutions and their practices of canonization (selection, order, preservation) are not lost or abandoned. They are re-membered into a complex assemblage of data, metadata, computational operations and networked contextualizations. The dissertation demonstrated that the digitization of libraries, archives and museums is indeed a dismemberment of their traditional objects and operations into bits and bytes, which are not recognizable as cultural heritage as such. It requires the continuous computation and processing of object models, surrogate models and momentary orderings in order for cultural heritage to be constructed out of data taking forms such as cultural heritage artefacts, collections of artefacts, exhibitions, relevance rankings, mash-ups and so forth. As the research question was answered during the course of the analysis and discussion of the case study, Europeana distinguishes itself from the information habitat as a persistent canonizer executing the continuous cycle of forgetting-as-data and remembering through emulations of cultural heritage artefacts as well as through providing potentialities for momentary renditions of order. In other words, Europeana does not hold cultural heritage but rather the ingredients for the (re-)computation and (re-)construction of cultural heritage. Thus conceived, the dissertation contributes an in-depth analysis of a major change in an important and venerable institution – an analysis of the shift from canonized silos of knowledge to platforms for canonizing.

8.2.2 Methodological Contributions

Within the field of IS research, qualitative methodology relies mostly on a canon of "classics" for guidance and reference with respect to good practice. Tied to epistemological schools of thought, the field is fragmented accordingly and, thus, produces methodological accounts and standards based on phenomenology, ethnomethodology, critical theory and so forth. In Chapter 4, this approach was

referred to as a methodology introducing substantial differences with respect to methods, techniques and quality criteria. By contrast, the dissertation argues that qualitative research is a *cyclical procedure* to explore unkown or unknowable populations. Thus defined, the qualitative methods used throughout the research project apply irrespective of the epistemological stance taken.

Against this backdrop, the technique of *corpus construction* was introduced as an alternative to theoretical sampling for selecting units of data analysis. Based on a two-dimensional unfolding of the field of study into what is known and what is unknown, corpus construction offers explicit rules for conducting as well as documenting the specific procedure followed by a researcher for selecting units of data analysis. Thus conceived, corpus construction is not only a compilation of good practices but also allows readers to evaluate the quality of the research project. Since a researcher is forced to make the initial expectations explicit by defining the known categories used for selection, the rationale for the subsequent selections is made explicit as well. In other words, corpus construction allows a reader to make an informed judgement concerning *why* and *how* units of data analysis were selected. Therefore the quality of the procedure can be evaluated by a reader irrespective of her knowledge on the epistemological foundation of the research project. In this sense, the dissertation contributes a procedural approach towards the selection of units of data analysis affording the comparison of qualitative research projects across the field.

The same observation can be made with respect to quality criteria in general. Based on an approach that is functionally equivalent to established quantitative quality criteria, the research does neither attempt to mimick validity, reliability and objectivity nor does it apply to criteria only valid within a specific epistemological tradition or school of thought upon which a specific qualitative methodology is based. The dissertation contributes an approach to research but also a rationale for documenting procedures that should not only allow for a comparison with other qualitative empirical studies according to a set of shared parameters but also with quantitative empirical studies according to a set of functionally equivalent parameters. Thus, the dissertation can be seen as a

contribution towards the methodological bridging of epistemological divides as well as the divide between qualitative and quantitative research.

8.2.3 Practical Contributions

Throughout the course of the research project, preliminary and final results were presented at conferences relevant for the field of information systems research. Notably, the research was also presented to the communities of librarians, archivists and curators at the "1st Symposium on Medialibraries and Archives in the 21st century" leading to a publication in a special issue of Artnodes – a journal on digital arts and heritage. Further publications, addressing the theoretical but also methodological contributions of the dissertation, are either already reviewed by leading academic journals or being prepared for submission (see Appendix 9.7).

By the same token, the research was also introduced to the community of digital memory studies through networking activities. While the contact with the London South Bank University resulted in a lively exchange of ideas and papers, the Oxford Internet Institute expressed its willingness to collaborate on future research projects focusing on the role of forgetting and remembering in the digital age.

As discussed in Chapter 4.6, the research was also discussed with expert employees of the Europeana Office. The results presented led to interesting and lively discussions about the future of social memory in the digital age indicating the relevance of the empirical study for the field of inquiry.

8.3 Limitations

The dissertation studies libraries, archives and museums as memory institutions. As a consequence, other mnemonic practices, which can be seen as part of the field of memory institutions, were disregarded in order to accomplish analytical

focus on a specific phenomenon of interest. Monuments, commemorations, rituals and so forth, which form a considerable part of the study of social and cultural memory (e.g. Erll and Nünning 2008), are cast into digital formats as well. A telling example for this field of research is the commemoration of the 9/11 terror attacks on the Twin Towers on websites such as www.911day.org allowing people to upload personal video messages about the event.

A second area of study not addressed by the research, online privacy and data protection is a complementary field of study addressing the other side of the preservation of online content (e.g. Introna 1997b; Mayer-Schönberger 2009). In this sense, the interests of the communities of librarians, archivists and curators to preserve contemporary online content for future generations contradicts the interests of privacy advocates campaigning against the long-term archiving of online personal data and the so-called digital footprint collected for mostly commercial reasons (Kapadia et al. 2007). These contradictory interests were illustrated recently when the Library of Congress received the whole archive of public tweets as a donation from twitter in 2010. While the LoC regards tweets as part of our cultural heritage worthy of being preserved, concerns are being raised with respect to the data being potentially usable for purposes other than historic research (Lohr 2010). By the same token, the European Commission is considering a "right to be forgotten" allowing European Union citizens to revoke their consent to their personal data being processed as well as data to be deleted once it passed an expiry date (European Commission 2010b:8).

The last area of study to be discussed in this chapter is the field of organizational memory and recent developments with respect to organizational forgetting (Stein 1995; Bowker 1997; Ciuk and Kostera 2010). At a first glance, organizational memory may seem to be an obvious reference to be addressed. However, the dissertation does not make any direct contributions to this particular field of study since the unit of analysis is not focused on the organizational memory of memory institutions but rather on libraries, archives and museums as memory institutions of the digital age. The dissertation, therefore, joins the discourse within the IS community discussing ICT as a phenomenon not confined by organizational boundaries (e.g. Kallinikos et al. 2010a; Yoo 2010). Still, future

research that bears on the lively discourse on organizational memory and oblivion could be developed.

From a methodological perspective, the single case study design leads to limitations with respect to the generalizability of the findings. Due to the available resources as well as to the quite unique characteristics of the Europeana initiative, the dissertation aims at what is referred to as Type 2 generalizability by Lee and Baskerville (2003) or analytical generalization by Yin (2003) – the generalization from empirical to theoretical statements. In this sense, the dissertation does not claim representativity (see also Chapter 4.3) but rather explores a relatively unknown phenomenon within the research domain of digital memories.

Finally, the focus on documentary evidence was a necessary tradeoff between the practical limitations of the actual research project and an ideal qualitative research design. In detail, semi-structured expert interviews were only used for communicative validation (see Chapter 4.6) rather than as the primary method of data collection as is traditionally the case in qualitative research. The author is aware that the analysis of documents lacks the richness provided by personal accounts and opinions usually generated in an interview situation. However, document analysis clearly compensates for this weak point by the fact that the documents were not produced for the researcher and, therefore, tailored to fit his expectations.

8.4 Future Research

The PhD is part of The Information Growth and Internet Research (www.TIGAIR.info) project focusing on the implications of the internet and the online services it affords on the wider institutional landscape of contemporary society. The author will continue to be part of the team after completion of the PhD programme. Since the research presented in the dissertation studied the digitization of cultural heritage artefacts, the next logical step is to study born-digital cultural artefacts and their transformation into born-digital cultural

heritage artefacts. Briefly addressed in the dissertation, first results were already published (see Kallinikos et al. 2010a; Marton 2010) and more publications are being prepared for submission (see Appendix 9.7).

During the course of the empirical study, a variety of possibilities for future research were discovered. For one, the topic of memory and identity in the digital age is of interest to institutions other than memory institutions such as libraries, archives and museums. It is in this sense that the author was approached by the Austrian Broadcasting Company to write a report on the future of public broadcasting and the internet from the perspective of national identity and cultural memory. In collaboration with Viktor Mayer-Schönberger from the Oxford Internet Institute, the project will launch in October 2011.

As discussed in the previous section, digital memory and privacy are complementary areas of study with sometimes contradictory approaches towards questions as to how and what contemporary society should remember and forget. The author is convinced that to find a balance between the right to be forgotten and the necessity to archive documents will be of central interest to academics, policy makers and the public. Hence, the author will continue to research this topic as a post-doctoral researcher.

To conclude, the topic of social memory and digital cultural heritage has proven to be an interesting area of study with a lot of potential for future research. It can be argued that memory may be as central a concept as information has been for quite some time now. Encouraged by the positive reception of the research project from parts of the academic community, librarians, archivists and curators, the author will continue his journey of exploring the future of the past.

9. APPENDIX

9.1 Corpus Construction - Masterlist

Code/ Corpus	Type	Title	created	modified/ published	collected/ last access	relevance	cycle
EE-001	Agenda	Agenda for EDL Foundation Executive Committee Meeting	n/a	14/03/08	05/06/08	yes	1
EE-002	Paper	Work Packages for a Best Practice Network to support the development of an Operational Europeana	n/a	14/03/08	05/06/08	yes	1
EE-003	Minutes	EDLnet Kick-off Meeting, The Hague, 12-14 September 2007	n/a	27/03/08	19/05/08	yes	1
EE-004	Slides	EDLnet KICK-OFF meeting WG 2.3: Technical Interoperability	13/09/07	n/a	19/05/08	yes	1
EE-005	Slides	EDLnet KICK-OFF meeting WP2: Technical & Semantic Interoperability	12-14/09/07	n/a	02/06/08	yes	1
EE-006	Slides	EDLnet KICK-OFF meeting WG2.1: Standards & Interoperability	13/09/07	n/a	05/06/08	yes	1
EE-007	Slides	EDLnet KICK-OFF meeting WG2.1: Standards & Interoperability	13/09/07	n/a	05/06/08	yes	1
EE-008	Project Plan	Outline Planning EDLnet v1	17/08/07	n/a	05/06/08	yes	1
EE-009	Memo	EDLnet Thematic Network-Summary of the project's Description of Work	25/04/08	25/04/08	10/10/08	yes	1
EE-010	Slides	Europeana Prototype 1	23/06/08	n/a	10/10/08	yes	1
EE-011	Slides	Your Europeana - Delivering User Expectations	23/06/08	n/a	10/10/08	yes	1

Code/ Corpus	Type	Title	created	modified/ published	collected/ last access	relevance	cycle
EE-012	Slides	Panel: Archives and Interoperability	23/06/08	n/a	10/10/08	yes	1
EE-013	Slides	Usability, Interoperability & Europeana	23/06/08	n/a	13/10/08	yes	1
EE-014	Slides	Feedback to Prototype 1 by WG1	23/06/08	n/a	13/10/08	yes	1
EE-015	Slides	Feedback to Prototype 1 by WG2.1	23/06/08	n/a	13/10/08	yes	1
EE-016	Slides	Feedback to Prototype 1 by WG2.2	23/06/08	n/a	13/10/08	yes	1
EE-017	Slides	Feedback to Prototype 1 by WG2.3	23/06/08	n/a	13/10/08	yes	1
EE-018	Slides	Round up of day 1 (23.06.08)	24/06/08	n/a	13/10/08	yes	1
EE-019	Slides	Top 5 priorities WG1	24/06/08	n/a	13/10/08	yes	1
EE-020	Slides	Top 5 priorities WG2.1	24/06/08	n/a	13/10/08	yes	1
EE-021	Slides	Top 5 priorities WG2.2	24/06/08	n/a	13/10/08	yes	1
EE-022	Slides	Top 5 priorities WG2.3	24/06/08	n/a	13/10/08	yes	1
EE-023	Slides	Priorities WG3	24/06/08	n/a	13/10/08	yes	1

Code/ Corpus	Type	Title	created	modified/ published	collected/ last access	relevance	cycle
EE-024	Minutes	EDLnet WP Leaders Project Management meeting	07/11/07	n/a	14/10/08	yes	1
EE-025	Summary	Collated actions EDLnet WP meetings & Foundation meeting	Nov/Dec- 07	n/a	14/10/08	yes	1
EE-026	Action	Content Themes	10/11/07	n/a	14/10/08	yes	1
EE-027	Action	Surveys to be carried out in next 6 months	07/11/07	n/a	14/10/08	yes	1
EE-028	Slides	Europeana maquette Where to click	29/01/08	n/a	14/10/08	yes	1
EE-029	Slides	Project Clusters for future Europeana	17-18/03/08	n/a	14/10/08	yes	1
EE-030	Paper	Work Packages for a Best Practice Network to support the development of an Operational Europeana	n/a	17/03/08	14/10/08	yes	1
EE-031	Slides	Maquette Feedback	17-18/03/08	n/a	14/10/08	yes	1
EE-032	List	Work Groups	03/10/07	n/a	14/10/08	yes	1
EE-033	Slides	WP1 recommendations for WP2	1-2/11/07	n/a	14/10/08	yes	1
EE-034	Paper	Business Model Europeana 2011-2015	Aug-08	n/a	16/10/08	yes	1
EE-035	Agenda	Agenda for EDLnet - WP1 Human & Political Interoperability Work Group meeting	30-31/10/08	n/a	17/10/08	yes	1

Code/ Corpus	Type	Title	created	modified/ published	collected/ last access	relevance	cycle
EE-036	Slides	EDLnet: Dissemination and advocacy activity for content	30-31/10/08	n/a	17/10/08	yes	1
EE-037	Minutes	EDLnet WP1 Human & Political Interoperability Work Group meeting	30-31/10/08	n/a	17/10/08	yes	1
EE-038	Image	Flipchart WP?	17/12/07	n/a	17/10/08	no	1
EE-039	Image	Flipchart WP?	17/12/07	n/a	17/10/08	no	1
EE-040	Image	Flipchart WP?	17/12/07	n/a	17/10/08	no	1
EE-041	Image	Flipchart WP?	17/12/07	n/a	17/10/08	no	1
EE-042	Image	Flipchart WP?	17/12/07	n/a	17/10/08	no	1
EE-043	Slides	WP2 - Access & Licensing	05/12/07	n/a	17/10/08	yes	1
EE-044	Agenda	WP1 Human & Political Interoperability Work Group 3rd meeting	17-18/12/07	n/a	17/10/08	yes	1
EE-045	Action	WP1 3rd Meeting	17-18/12/07	n/a	17/10/08	yes	1
EE-046	Minutes	WP1 3rd Meeting	17-18/12/07	n/a	17/10/08	yes	1
EE-047	Slides	Group Discussion about Europeana	17-18/12/07	n/a	17/10/08	yes	1

Code/ Corpus	Type	Title	created	modified/ published	collected/ last access	relevance	cycle
EE-048	Agenda	Agenda for EDLnet - WP1 Human & Political Interoperability Work Group 4th meeting	21/05/08	n/a	17/10/08	yes	1
EE-049	Paper	Partner Branding in Europeana	21/05/08	n/a	17/10/08	yes	1
EE-050	Paper	Europeana - a sustainable service	10/05/08	n/a	17/10/08	yes	1
EE-051	Paper	Overview of EDL Foundation and Europeana funding trajectory	21/05/08	n/a	17/10/08	yes	1
EE-052	Paper	EDLnet progress report on overall planning	14/05/08	n/a	17/10/08	yes	1
EE-053	Paper	Review & Report on Maquette and Recommendations to WP1	21/05/08	n/a	17/10/08	yes	1
EE-054	Paper	Prototype development & interoperability	21/05/08	n/a	17/10/08	yes	1
EE-055	Paper	Communications plan for launch of Europeana Prototype Nov 2008	08/05/08	n/a	17/10/08	yes	1
EE-056	Paper	Content acquisition for Europeana prototypes	14/05/08	n/a	17/10/08	yes	1
EE-057	Paper	EDLNet Work Package 1 Discussion Paper: Review of Governance and Organisation for Europeana	19/05/08	n/a	17/10/08	yes	1
EE-058	Paper	Europeana v1.0	21/05/08	n/a	17/10/08	yes	2
EE-059	Paper	Business Planning scenarios and models for Europeana	08/05/08	n/a	17/10/08	yes	1

Code/ Corpus	Type	Title	created	modified/ published	collected/ last access	relevance	cycle
EE-060	Slides	Europeana	21/05/08	n/a	17/10/08	yes	1
EM-001	Press Release	European Digital Library Foundation welcomed by the Commissioner	28/08/07	n/a	10/10/08	yes	1
EP-001	Article	Web 2.0 ideas enrich Europe's digital heritage	Aug/Sept-08	n/a	16/10/08	yes	1
ES-001	Deliverable	Report Detailing Organisational Structure to be used by WP's 2 & 3	19/12/07	n/a	16/10/08	yes	1
ES-002	Deliverable	Deliverable 1.2 Statutes of EDL Legal Entity	27/11/07	n/a	16/10/08	yes	1
ES-003	Deliverable	D1.4 Project proposal for an operational service	11/06/08	n/a	16/10/08	yes	1
ES-004	Deliverable	EDLnet Project work plan	03/12/07	n/a	16/10/08	yes	1
ES-005	Deliverable	Europeana - a sustainable service - Outline Business Plan for Europeana as a service of the EDL Foundation	30/09/08	n/a	16/10/08	yes	2
ES-006	Deliverable	Initial Semantic and Technical Interoperability Requirements	17/12/07	n/a	16/10/08	yes	1
ES-007	Deliverable	User Use Cases - Functional requirements for EDL Maquette	06/02/08	n/a	16/10/08	yes	1
ES-008	Deliverable	European Digital Library - Thematic Network	Nov-07	n/a	16/10/08	yes	1
ES-009	Deliverable	Interim Report on the first 6-months period	28/01/08	n/a	16/10/08	yes	1

Code/ Corpus	Type	Title	created	modified/ published	collected/ last access	relevance	cycle
ES-010	Deliverable	Deliverable 5.5 Network Agreement	01/09/07	n/a	16/10/08	yes	1
ME-001	Interview	Europas digitale Bibliothek	10/06/08	n/a	16/10/08	yes	1
MP-001	Article	Europeana: Europäisches Kulturerbe bald online	17/08/08	n/a	16/10/08	no	1
SE-001	Slides	It's a New Day	23/06/08	n/a	10/10/08	yes	1
XE-001	Slides	How CIDOC-CRM supports interoperability	23/06/08	n/a	10/10/08	yes	1
XE-002	Slides	Archives and Interoperability - OAI ORE promotes serendipitous interoperability	23/06/08	n/a	10/10/08	yes	1
XE-003	Slides	Interoperability and museums	23/06/08	n/a	10/10/08	yes	1
XE-004	Slides	Interoperability of Museum Information: “Objects rarely talk”	23/06/08	n/a	10/10/08	yes	1
XE-005	Slides	Interoperability in Museum environment	23/06/08	n/a	13/10/08	yes	1
XE-006	Slides	Key Results from Maquette Surveys	03/03/08	n/a	14/10/08	no	1
XX-001	Manifesto	The DELOS Digital Library Reference Model	n/a	Dec-07	06/06/07	no	1
XX-002	Guideline	The Digital Library Reference Model in a Nutshell	n/a	Oct-07	10/06/08	no	1

Code/ Corpus	Type	Title	created	modified/ published	collected/ last access	relevance	cycle
XX-003	Manifesto	The Digital Library Manifesto	Jan-06	n/a	06/06/07	no	1
XX-004	Questionnaire	Questionnaire about Cultural Heritage institutions and the European Directive on the re-use of Public Sector Information (PSI):	27/04/08	n/a	n/a	no	1
XX-005	Guideline	D3.1.8 - BRICKS API Reference Guide	01/01/04	31/12/06	13/10/08	no	1
XX-006	Report	Digital Repository Infrastructure Vision for European Research (DRIVER)	01/06/06	15/11/06	13/10/08	no	1
XX-007	Specification	Potential services for the European Library	n/a	08/11/07	13/10/08	no	1
XX-008	Deliverable	Video Active Content Selection Strategy	17/08/07	n/a	14/10/08	no	1
ES-011	Deliverable	Outline Business Plan for Europeana as a service of the EDL Foundation	n/a	14/11/08	13/08/09	yes	1
ES-012	Deliverable	Europeana Outline Functional Specification	10/02/08	20/08/08	13/08/09	yes	1
ES-013	Deliverable	Report on User perspectives on prototypes	29/09/08	Nov-08	13/08/09	yes	1
ES-014	Deliverable	Final Recommendations of Users for Usability Work Group	03/10/08	Nov-08	13/08/09	yes	1
ES-015	Deliverable	European Digital Library Thematic Network	n/a	Apr-09	13/08/09	yes	2
ES-016	Deliverable	Project power point presentation - Results	n/a	30/01/09	13/08/09	yes	2

Code/ Corpus	Type	Title	created	modified/ published	collected/ last access	relevance	cycle
ES-017	Deliverable	First Annual Report to the Commission	01/07/07	30/07/08	13/08/09	yes	1
ES-018	Deliverable	Interim Report on the third 6-months period	01/07/07	02/02/09	13/08/09	yes	2
ES-019	Deliverable	Final report	01/07/07	08/05/09	13/08/09	yes	2
ES-020	Deliverable	Europeana Outline Functional Specification For development of an operational European Digital Library	n/a	01/03/09	30/04/10	yes	2
EX-001	Presentation	Functionality, Data and Infrastructure. Evolution from Prototype to Europeana 1.0	18/09/08	n/a	18/05/10	yes	1
EX-002	White Paper	Knowledge = Information in Context: On the Importance of Semantic Contextualisation in Europe	Apr-10		01/06/10	yes	2
XE-007	Web Design Draft	n/a	n/a	n/a	05/06/10	no	1
XE-008	Web Design Draft	n/a	n/a	n/a	05/06/10	no	1
XE-009	Web Design Draft	n/a	n/a	n/a	05/06/10	no	1
EE-061	Org. Chart	Organization Chart EDL Foundation	24/03/10	n/a	14/07/10	yes	2
SP-001	Press Release	Digital Agenda: Reflection Group on digitisation seeks views on boosting cultural heritage online	18/08/10	n/a	20/08/10	no	2
EP-002	Report	Annual Report	26/02/10	n/a	01/03/10	yes	2

Code/ Corpus	Type	Title	created	modified/ published	collected/ last access	relevance	cycle
EP-003	Charter	The Europeana Public Domain Charter	n/a	Apr-10	14/04/10	yes	2
EM-002	Press Release	Boosting cultural heritage online: the European Commission sets up a Reflection Group on digitisation	n/a	21/04/10	22/04/10	yes	2
EE-062	Email	Results of Council elections and User testing report	n/a	27/04/10	27/04/10	yes	2
EE-063	Report	Highlights of Europeana v1.0 and other projects in the Europeana Group	n/a	Jul-10	s	yes	2
EE-064	Report	Rhine features in scope for July 2010	n/a	Jul-10	30/07/2010	yes	2
EE-065	Report	Definition of the Europeana Data Model elements	n/a	30/07/10	14/09/10	yes	2
EE-066	Slides	EDM	n/a	01/09/10	14/09/10	yes	2
EM-003	Press Release	Digital Agenda: Europeana gives online access to over 14 million examples of Europe's cultural heritage	18/11/10	n/a	19/11/10	yes	2
ES-021	Handbook	Europeana Aggregators' Handbook	01/05/10	n/a	14/12/10	yes	2
EE-067	Memo	Open Linked Data and Europeana	n/a	07/12/10	07/12/10	yes	2
ES-022	Deliverable	Functional specification for the Europeana Danube release	31/08/10	n/a	21/01/11	yes	2
ES-023	Deliverable	EuropeanaConnect First Annual Progress Report. May 2009 – April 2010	01/04/10	n/a	24/01/11	yes	2

Code/ Corpus	Type	Title	created	modified/ published	collected/ last access	relevance	cycle
ES-024	Deliverable	EuropeanaConnect General Presentation	11/08/10	n/a	24/01/11	yes	2
EP-004	Report	Europeana Strategic Plan 2011-2015	n/a	n/a	24/01/11	yes	2
EE-068	Slides	Europeana Connect	02/04/09	n/a	27/01/11	yes	2
EE-069	Statutes	EDL Foundation Council of Content Providers and Aggregators	11/11/09	n/a	28/01/11	yes	2
EX-003	Slides	Integration of Heterogeneous Metadata in Europeana	n/a	25-30/05/09	18/02/11	yes	2
EE-070	Report	Europeana Data Model Primer	n/a	05/08/10	24/02/11	yes	2
XE-010	Slides	EUROPEANA - ONLINE VISITOR SURVEY -2011	n/a	26/07/11	07/09/11	yes	2
EP-005	Report	Networking - Annual Report and Accounts 2010	n/a	09/06/11	07/09/11	yes	2
ES-025	Deliverable	Europeana v1.3 Danube Major Release	30/04/11	30/04/11	07/09/11	yes	2
ES-026	Report	Business Plan 2011	n/a	28/04/11	07/09/11	yes	2
ES-027	Deliverable	User tests report	31/03/11	31/03/11	07/09/11	yes	2
EE-071	Report	Content Development Strategy	Dec-10	14/12/10	07/09/11	yes	2

Code/ Corpus	Type	Title	created	modified/ published	collected/ last access	relevance	cycle
ES-028	Report	Progress report. 1 August 2009 - 31 May 2010	06/07/10	01/10/10	07/09/11	yes	2
ES-029	Deliverable	Initial Technical & Logical Architecture and future work recommendations	30/07/10	01/10/10	07/09/11	yes	2
EX-004	Slides	Culturer Creativity Growth. Business Strategy 2011-2015	n/a	02/12/10	07/09/11	yes	2
EX-005	Slides	Europeana - The book. Europeana - The future	n/a	12/10/10	07/09/11	yes	2
EX-006	Slides	Welcome to Europeana	25/03/10	25/03/10	07/09/11	yes	2
EX-007	Slides	Europeana	17/03/10	17/03/10	07/09/11	yes	2
RE-001	Interview		n/a	n/a	14/07/11	yes	2
RE-002	Interview		n/a	n/a	14/07/11	yes	2
RE-003	Interview		n/a	n/a	14/07/11	yes	2
RE-004	Interview		n/a	n/a	14/07/11	yes	2

9.2 Thematic Coding - Example Thematic Coding Sheet

Document EE-006 **Type**: Powerpoint Slides

Title: EDLnet KICK-OFF meeting WG2.1: Standards & Interoperability

Context: presented by Makx Dekkers, 13th Sept. 2007, The Hague

Theme: Mapping the information environment
If Europeana works as an information environment for the user, this document deals with the technicalities of constructing = mapping it.

Quote from page 1, slide 3

- • Agree on
  - common **metadata standards and models**
  - in relation to the **operational models** (centralized and distributed),
- • Identify potential
  - common **sets of vocabularies and classification systems**
  - for various **descriptive properties** (object types, subjects, audiences, rights etc.)
- • and consider **persistent identifiers** approaches for objects and collections.

Open Coding
The metadata standards and models (metadata models?) are related to the operational model of Europeana. The operational model seems like a technical expression regarding where the content is actually stored. (memoModel) So, I could say that the description of the object or rather the standards according to which description is done is related to technical issues. Where the objects are stored (in one place or in many places) has an influence on the metadata standards and vice versa.

"descriptive properties" are what? Properties of description? Properties of the metadata? No… properties of the described! But these properties are actually not properties in themselves but become properties only when observed, ie. captured by a classification system. For instance, a book is not of a certain object type in itself, but is classified as such! This refers to ways of describing an object, in other words making it into a librarian object.

Category: descriptive properties

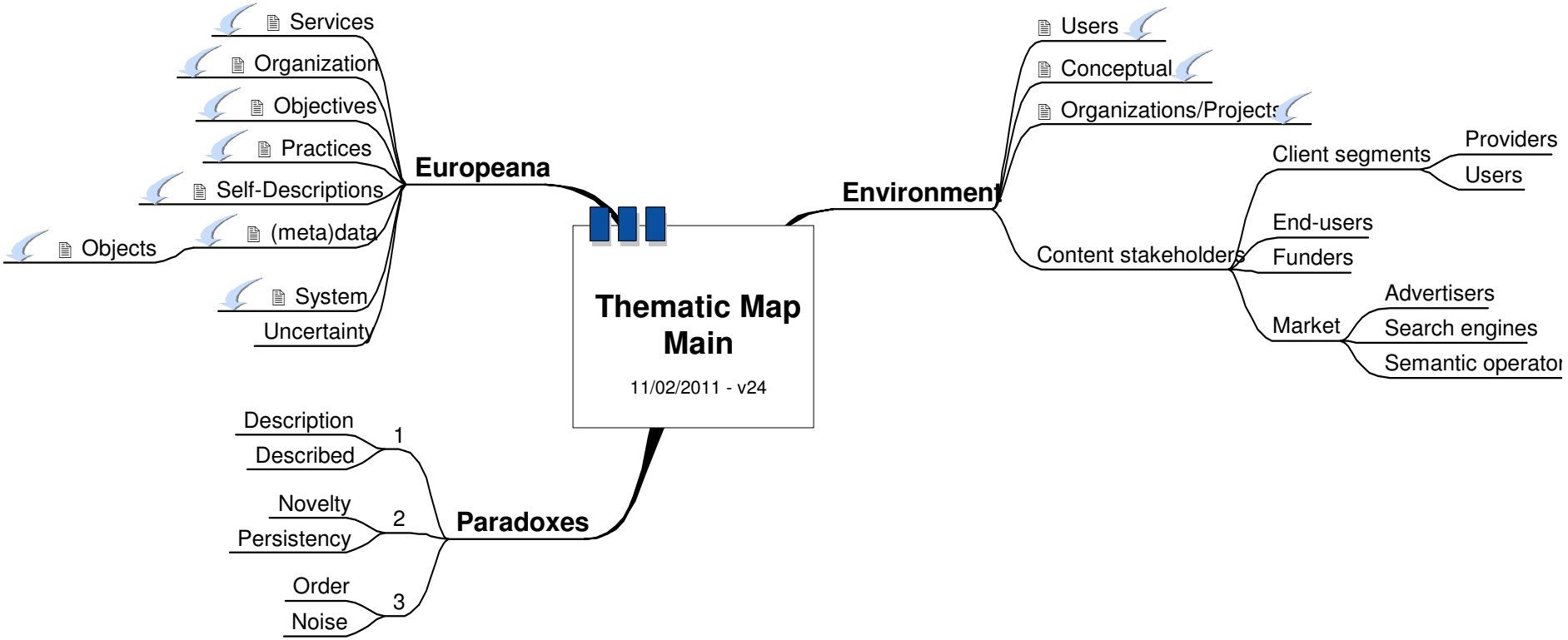
- object types
- subjects
- audiences
- rights
- etc.

"persistent identifiers"… again persistence not only for objects but also for collections. Could it be that a collection is a metaobject (a librarian object made up of other librarian objects? Themes could be metaobjects as well?) The question remains whether PIDs are assigned to collections as well. Does a collection become a metaobject as soon as one library includes another collection? When the library is differentiated from the collection it harbours.

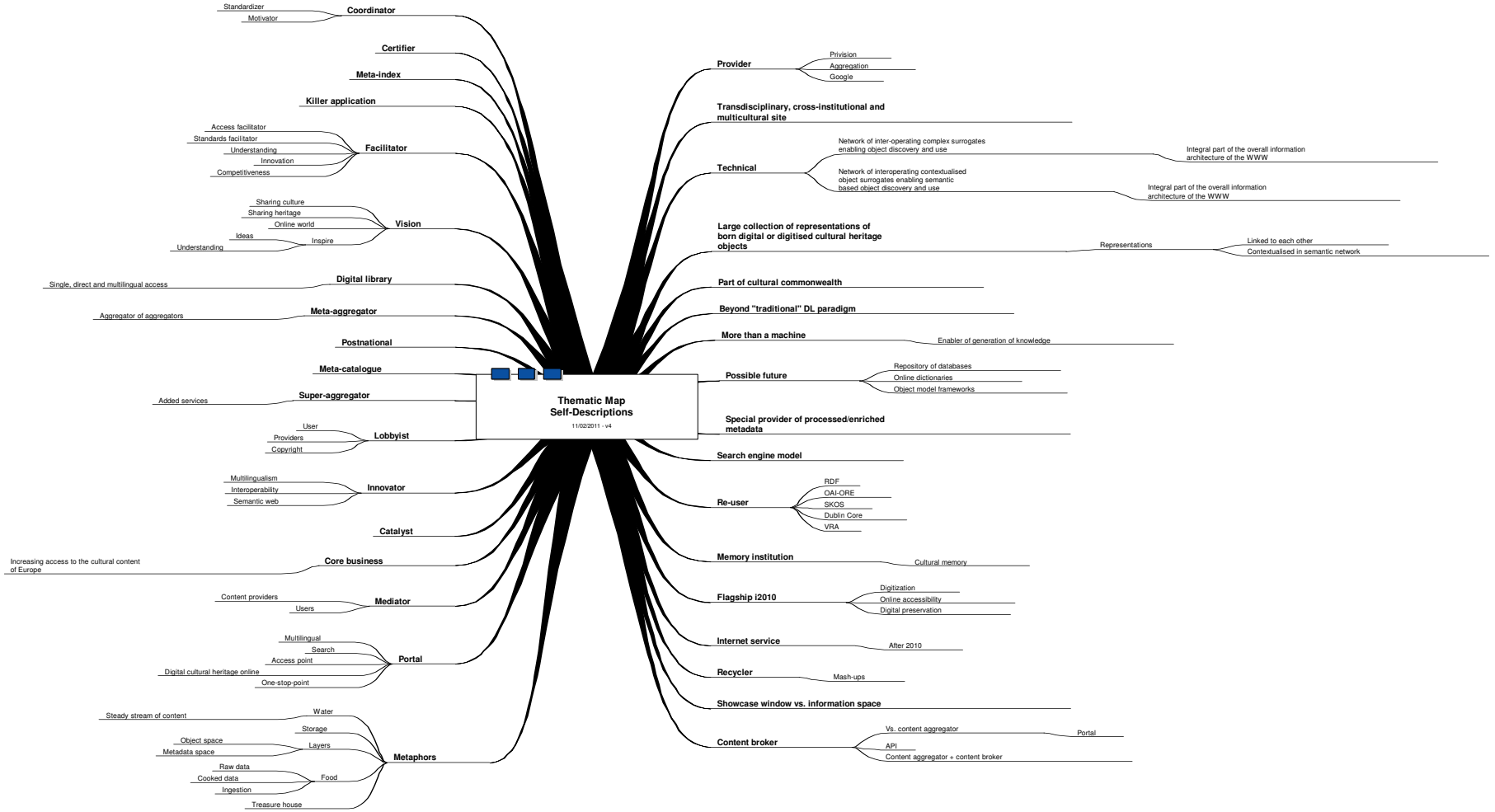
[…]

9.3 Thematic Coding – Example Thematic Maps

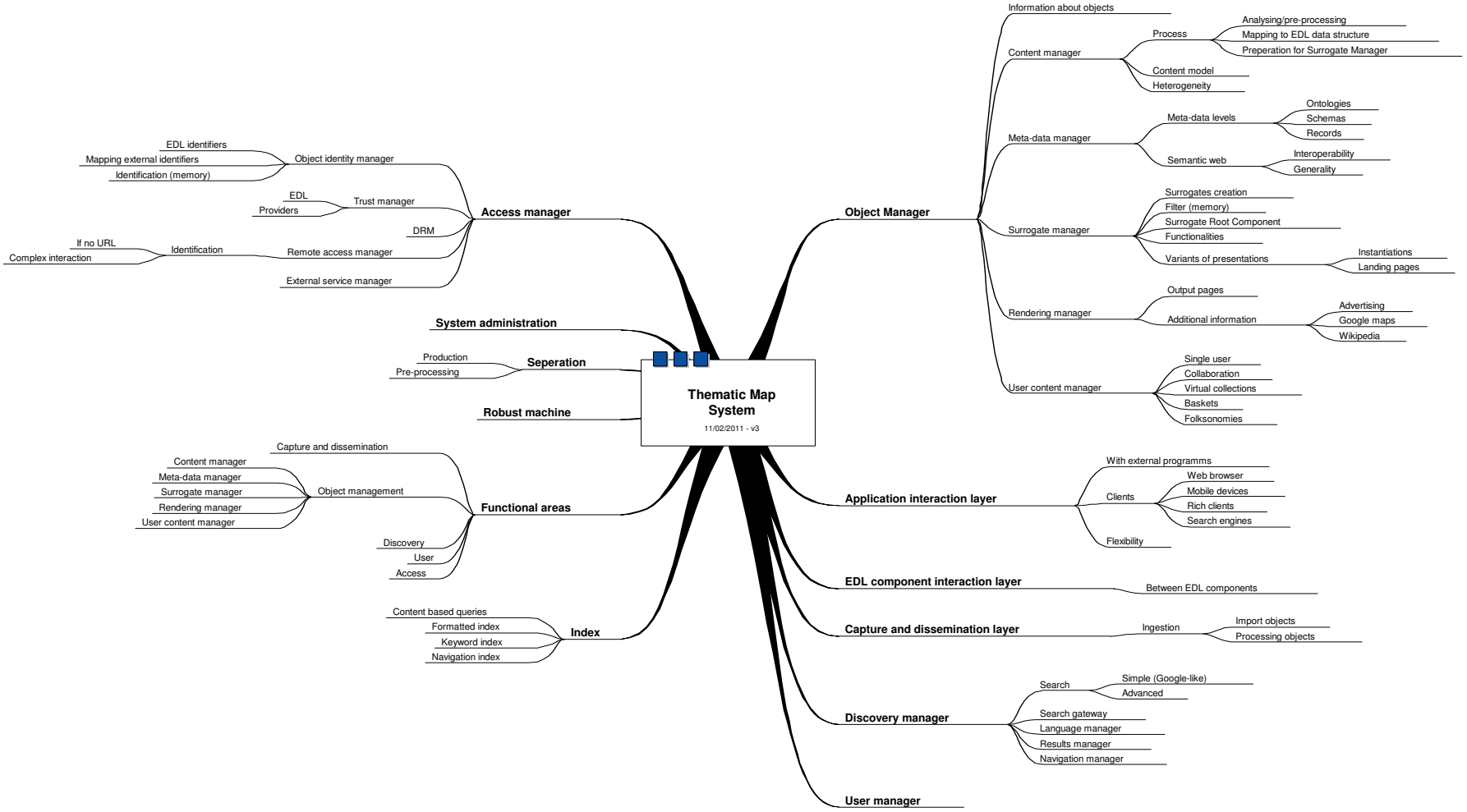
9.3.1 *Example Thematic Map – Main*



9.3.2 *Example Thematic Map – Self-Descriptions*



9.3.3 *Example Thematic Map - System*

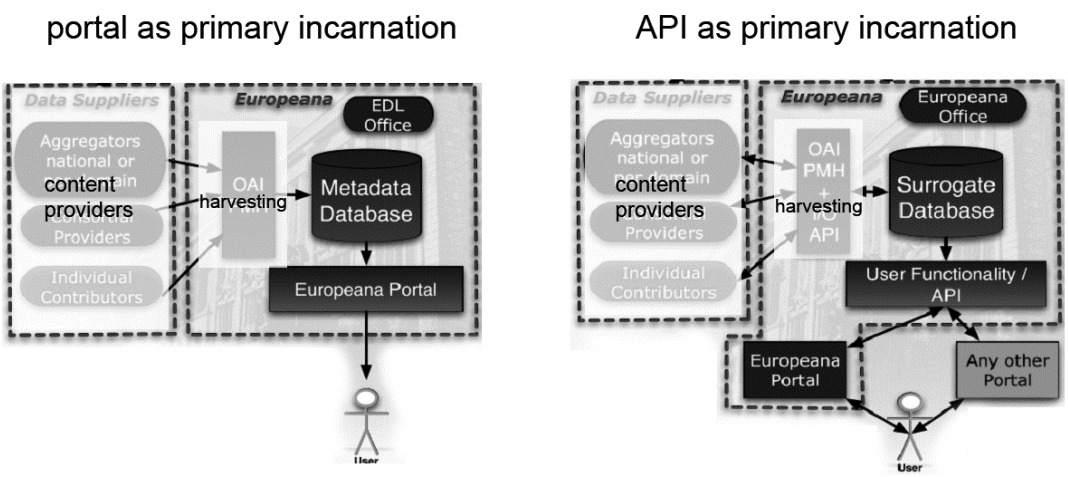


9.4 Communicative Validation – Interview Topic Guide

1. **Introduction**
- What is your current position in the organization? What is your job about?
- What is Europeana currently working on?

2. **Europeana – Comparison with Memory Institutions**
- How would you describe Europeana?
- Why is a project like Europeana necessary?
- What do you think are the differences and similarities between Europeana and traditional libraries, archives and museums?
- What do you think are the differences and similarities between Europeana and *digital* libraries, archives and museums?

3. **Europeana – Internal Comparison**
- The following, slightly modified figure was presented by Stefan Gradmann at a workshop in 2008. Do you think this depiction is still valid?



- How would you characterize the differences between the prototype portal and Europeana as it is implemented today but also envisioned for the future?
- The following quote is taken from a Europeana project document.

  "*Application programming interfaces (APIs) represent one of the initially unsuspected issues for Europeana. The Portal is indeed a huge collection of contents, in different formats, with different access authorizations and from different collections. At the same time it is a unique possibility to have so many different contents in the same location, permitting a transversal approach to contents never before available. This will permit institutional users initially and the general users later on to bring together contents in new applications on different platforms. A Museum in one country may make a European exhibition on prehistoric objects from all the available collections on their website, based on the Europeana portal and respecting the access rights of every content owner*"
- Could you explain to me what this quote refers to? What is meant by "a transversal approach"?

4. **Conclusion**
- What do you think is expected from Europeana?
- What would happen, if Europeana was shut down today?

9.5 Research Ethics Checklist

	Yes	No	Not certain
Does the study involve participants who are in any way vulnerable or may have any difficulty giving consent?		**X**	
Will the study require the co-operation of a gatekeeper for initial access to the groups or individuals to be recruited? (e.g. students at school, members of a self-help group, residents of a nursing home).		**X**	
Will it be necessary for participants to take part in the study without their knowledge and consent at the time? (e.g. covert observation of people in on-public places)		**X**	
Is the subject of the study such that it might cause participants undue stress or anxiety?		X	
Are drugs, placebos or other substances to be administered to the study participants or will the study involve invasive, intrusive or potentially harmful procedures of any kind?		X	
Is pain or more than mild discomfort likely to result from the study?		X	
Could the study induce unacceptable psychological stress or anxiety or cause harm or negative consequences beyond the risks encountered in normal life?		X	
Will the study involve prolonged or repetitive testing?		X	
Will financial inducements (other than reasonable expenses and compensation for time) be offered to participants?		X	
Does the research methodology use deception?		X	
Will the study involve recruitment of patients or staff through the NHS or Councils with Social Services responsibility (CSSRs)?		X	

9.6 Informed Consent Form

TITLE OF RESEARCH PROJECT:
Memory Institutions and the Interconnected Information Environment

DATE: 04/07/2011

You have been asked to take part in a research project which is part of a PhD thesis at the London School of Economics. The purpose of the research is to explore how memory institutions change their practices when they are stepping into the online world.

PARTICIPATION AND CONFIDENTIALITY
Your participation in this research project involves a 1 hour interview with Attila Marton, the researcher on the project. The interview will be taped by a digital audio recorder. All interviews will be kept confidential and will be fully anonymised. This means that personal information such as your name will not be used in any work that results from this research.

Results of this research project will be written up as part of a PhD thesis at the London School of Economics. In addition, results may be published in academic journals and discussed at conferences.

You have the right to stop the interview at any time. You also have the right to ask the researcher for a copy of the computer file of the recording and/or for the deletion of the original recording.

CONSENT
I understand the purpose of this research project and all my questions have been answered. I understand that my interview will be kept confidential and will be fully anonymized. I understand that I have the right to stop the interview at any time.

I give my consent to be interviewed.

---------- ----------
Participant's Signature and Date Participant's Printed Name and Date

---------- ----------
Interviewer's Signature (witness)/ Date Interviewer's Printed Name/Date

Contact: Attila Marton
Information Systems and Innovation Group
Department of Management
The London School of Economics and Political Science
Houghton Street
London WC2A 2AE, UK
a.marton@lse.ac.uk

9.7 List of Publications, Conference Papers and Presentations

Publications

Marton A. (2011) Editorial: The black quasi-box. iSChannel **6**(1): 4-5.

Marton A. (2010) The transfigurability of digital objects. Artnodes **10**. http://bit.ly/hJgFnb. Also available in Spanish: http://bit.ly/i0VW1Y.

Marton A. (2010) Editorial: The requisite variety of IS research. iSChannel **5**(1): 3.

Kallinikos J., Aaltonen A. and Marton A. (2010) A theory of digital objects. First Monday **15**(6). http://bit.ly/1stmon.

Marton A. (2009) Self-referential technology and the growth of information. From techniques to technology to the technology of technology. Soziale Systeme **15**(1): 137-159.

Marton A. (2008) Organisierte Kommunikation ohne Gesicht: Die Bewältigung zeitlich und sachlich entkoppelter Kommunikation in Virtuellen Organisationen am Beispiel einer MMORPG-Gilde. Saarbrücken:VDM.

Conference Papers and Invited Presentations

Marton A. (2011) Social memory and the digital domain: The canonization of digital cultural artefacts. EGOS: Gothenburg, Sweden.

Marton A. (2011) Social memory and IT – Forgetting or remembering. SSIT-ORF: London, UK.

Kallinikos J., Hasselbladh H. and Marton A. (2010) Governing social practice: Technology versus institutions. EGOS: Lisbon, Portugal.

Marton A. (2010) The transfigurability of digital objects. 1st Symposium of Medialibraries and Archives for the 21st Century: Gijon, Spain.

Marton A. (2010) The autological constitution of digital cultural artefacts. An analysis of the implications of ICT on memory organizations. LAEMOS: Buenos Aires, Argentina. Finalist for Best Paper Award.

Marton A. (2009) Digital libraries as information organizations: The re-unfolding of the memory/information paradox. ECIS: Verona, Italy.

Marton A. (2009) Things have to change in order to be the same: Functional equivalence as a method for analysing “the new” in “the old” and “the old” in “the new”. Quadrangular Conference in Technology, Organization & Society: Cambridge, UK.

Marton A. (2007) The European Digital Library (EDL). SSIT-ORF: London,UK.

Works In Progress

Kallinikos J., Aaltonen A. and Marton A. “A conceptual framework for studying digital objects: A new frontier for IS research”.

Kallinikos J., Aaltonen A. and Marton A. “Information generativity and logics of innovation” under review at the Journal of Information Technology.

Kallinikos J., Hasselbladh H. and Marton A. “Governing social practice: Technology versus institutions”, under review at Theory and Society.

Marton A. “Qualitative selection as a quality criterion. Corpus construction and the quality of qualitative research”, in preparation.

Marton A. “The IT-turn in social memory”, in preparation for Organization Studies.

Kallinikos J. and Marton A. “Social memory and the archiving of the web: Re-establishing the link between objects and institutions” in preparation for Organization Studies.

10. BIBLIOGRAPHY

Aaltonen, A. and G. F. Lanzara (2010). Unpacking Wikipedia governance: The emergence of a bureaucracy of peers? 1st LAEMOS Colloquium, Buenos Aires, Argentina.

Alexander, E. P. and M. Alexander (2008). Museums in motion: An introduction to the history and functions of museums. Lanham, AltaMira Press.

Allison, A., J. Curall, M. Moss and S. Stuart (2005). "Digital identity matters." Journal of the American Society for Information Science and Technology **56**(4): 364-72.

Ames, M. M. (1992). Cannibal tours and glass boxes: The anthropology of museums. Vancouver, UBC Press.

Angell, I. O. (1990). "Systems thinking about information systems and strategies." Journal of Information Technology **5**(3): 168-74.

Angell, I. O. and D. S. Demetis (2010). Science's first mistake. Delusions in pursuit of theory. London, Bloomsbury.

Angell, I. O. and F. M. Ilharco (2004). "Solution is the problem: A story of transitions and opportunities", in: C. Avgerou, C. Ciborra and F. Land (Eds.), The social study of information and communication technology: Innovation, actors and contexts. Oxford, Oxford University Press: 38-61.

Archer, M. S. (1995). Realist social theory: The morphogenetic approach. Cambridge, MA, Cambridge University Press.

Arms, W. Y. (2001). Digital libraries. Cambridge, MA, MIT Press.

Arnoldi, J. (2001). "Niklas Luhmann - An introduction." Theory, Culture & Society **18**(1): 1-13.

Artandi, S. (1973). "Information concepts and their utility." Journal of the American Society for Information Science **24**: 242-45.

Ashby, W. R. (1957). An introduction to cybernetics. London, Chapman & Hall.

Ashby, W. R. (1958). "Requisite variety and its implications for the control of complex systems." Cybernetica **1**(2): 83-99.

Assmann, A. (1999). Erinnerungsräume: Formen und Wandlungen des kulturellen Gedächtnisses. Munich, C.H. Beck.

Assmann, A. (2006). "The printing press and the internet: From a culture of memory to a culture of attention", in: N. Gentz and S. Kramer (Eds.), Globalization, cultural identities, and media representations. New York, State University of New York Press: 11-23.

Assmann, A. (2008a). "Canon and archive", in: A. Erll and A. Nünning (Eds.), Cultural memory studies: An international and interdisciplinary handbook. Berlin, New York, Walter de Gruyter: 97-107.

Assmann, A. (2010). "Archive und Bibliotheken", in: C. Gudehus, A. Eichenberg and H. Welzer (Eds.), Gedächtnis und Erinnerung: Ein interdisziplinäres Handbuch. Stuttgart, Weimar, J.B. Metzler: 165-70.

Assmann, J. (1995). "Collective memory and cultural identity." New German Critique **65**: 125-33.

Assmann, J. (2002). "Nachwort", in: E. Esposito (Ed.), Soziales Vergessen: Formen und Medien des Gedächtnisses der Gesellschaft. Frankfurt am Main, Suhrkamp: 400-14.

Assmann, J. (2008b). "Communicative and cultural memory", in: A. Erll and A. Nünning (Eds.), Cultural memory studies: An international and interdisciplinary handbook. Berlin, New York, Walter de Gruyter: 109-18.

Atkins, S., J. Clear and N. Ostler (1992). "Corpus design criteria." Literary and Linguistic Computing **7**(1): 1-16.

Avram, H. D. (1975). MARC: Its history and implications. Washington, D.C., Library of Congress http://eric.ed.gov/ERICWebPortal/recordDetail?accno=ED127954

Baecker, D. (1996). "Wenn es im System rauscht." gdi_IMPULS **1**: 65-74.

Baecker, D. (1999). Organisation als System. Frankfurt am Main, Suhrkamp.

Baecker, D. (2006a). "The form of the firm." Organization **13**(1): 109-42.

Baecker, D. (2006b). "Niklas Luhmann in the society of the computer." Cybernetics & Human Knowing **13**(2): 25-40.

Baker, D. (2006). "Digital library futures: A UK HE and FE perspective." Interlending & Document Supply **34**(1): 4-8.

Baker, T. (2000). "A grammar of Dublin Core." D-Lib **6**(10). http://www.dlib.org/dlib/october00/baker/10baker.html

Barreau, D. (2001). "Information systems and collection development in public libraries." Library Collections, Acquisitions, and Technical Services **25**(3): 263-79.

Barron, S. (2011). Google can't be trusted with our books. Guardian. London http://www.guardian.co.uk/commentisfree/2011/apr/26/google-books-videos

Baskerville, R. L. and M. D. Myers (2004). "Special issue on action research in information systems: Making IS research relevant to practice." MIS Quarterly **28**(3): 329-35.

Baskerville, R. L. and A. T. Wood-Harper (1998). "Diversity in information systems action research methods." European Journal of Information Systems **7**(2): 90-107.

Basu, L. (2011). "Memory dispositifs and national identities: The case of Ned Kelly." Memory Studies **4**(1): 33-41.

Bateson, G. (2000). Steps to an ecology of mind. Chicago, University of Chicago Press.

Baudrillard, J. (1994). Simulacra and simulation. Ann Arbor, University of Michigan Press.

Bauer, M. and B. Aarts (2000). "Corpus construction: A principle for qualitative data collection", in: M. Bauer and G. Gaskell (Eds.), Qualitative researching with text, image and sound: A practical handbook. London, Sage: 19-37.

Bauer, M., G. Gaskell and N. C. Allum (2000). "Quality, Quantity and Knowledge Interests", in: M. Bauer and G. Gaskell (Eds.), Qualitative researching with text, image and sound: A practical handbook. London, Sage: 3-17.

Bausch, K. C. (1997). "The Haberman/Luhmann debate and subsequent Habermasian perspectives on systems theory." Systems Research and Behavioral Science **14**(5): 315-30.

Baxter, C., Ed. (1999). The business of memory. The art of remembering in an age of forgetting. Saint Paul, MN, Greywolf Press.

BBC (2002). Digital Domesday Book unlocked. London http://news.bbc.co.uk/1/hi/technology/2534391.stm

BBC (2008). European online library crashes. London http://news.bbc.co.uk/1/hi/7742390.stm

Bearman, D. (2006). "Jean-Noel Jeanneney's critique of Google: Private sector book digitization and digital library policy." D-Lib **12**(12). http://www.dlib.org/dlib/december06/bearman/12bearman.html

Bearman, D. and J. Trant (2005). "Converting scanned images of the print history of the world to knowledge: A reference model and research strategy." Russian Digital Libraries Journal **8**(5). http://www.elbib.ru/index.phtml?page=elbib/eng/journal/2005/part5/BT

Beck, U. (1986). Risikogesellschaft auf dem Weg in eine andere Moderne. Frankfurt am Main, Suhrkamp.

Beck, U., A. Giddens and S. Lash (1994). Reflexive modernization: Politics, tradition and aesthetics in the modern social order. Stanford, Stanford University Press.

Becker, C. and A. Rauber (2011). "Decision criteria in digital preservation: What to measure and how." Journal of the American Society for Information Science and Technology **62**(6): 1009-28.

Bedini, S. A. (1965). "The evolution of science museums." Technology and Culture **6**(1): 1-29.

Bell, C. G. and J. Gemmell (2009). Total recall: How the e-memory revolution will change everything. New York, Dutton.

Bell, D. (1974). The coming of post-industrial society: A venture in social forecasting. London, Heinemann.

Benbasat, I., D. K. Goldstein and M. Mead (1987). "The case research strategy in studies of information systems." MIS Quarterly **11**(3): 369-86.

Beniger, J. R. (1986). The control revolution: Technological and economic origins of the information society. Cambridge, MA, Harvard University Press.

Benkler, Y. (2006). The wealth of networks. How social production transforms markets and freedom. New Haven, London, Yale University Press.

Bennett, S. (2001). "The golden age of libraries." The Journal of Academic Librarianship **27**(4): 256-59.

Berger, P. L. and T. Luckmann (2003). Die gesellschaftliche Konstruktion der Wirklichkeit. Eine Theorie der Wissenssoziologie. Frankfurt am Main, Fischer.

Berliner, D. (2005). "The abuses of memory: Reflections on the memory boom in anthropology." Anthropological Quarterly **78**(1): 197-211. http://muse.jhu.edu/journals/anthropological_quarterly/v078/78.1berliner.html

Berners-Lee, T. (2006). Linked data. Design Issues http://www.w3.org/DesignIssues/LinkedData.html

Bertalanffy, L. v. (1973). General system theory: Foundations, development, applications. Harmondsworth, Penguin.

Beynon-Davies, P. (2010). "The enactment of significance: A unified conception of information, systems and technology." European Journal of Information Systems **19**(4): 389-408.

Bhaskar, R. (1998). "Philosophy and scientific realism", in: M. S. Archer (Ed.), Critical realism: Essential readings. London, Routledge: 16-48.

Biber, D. (1993). "Representativeness in corpus design." Literary and Linguistic Computing **8**(4): 243-57.

Bishop, A. P., N. A. Van House and B. P. Buttenfield (2003). "Introduction: Digital libraries as sociotechnical systems", in: A. P. Bishop, N. A. Van House and B. P. Buttenfield (Eds.), Digital library use: Social practice in design and evaluation. Cambridge, MA, MIT Press: 1-21.

Bjorner, S. (2006). "Building a European digital library: A challenge in the culture wars." Searcher **14**(3): 27-34.

Blanchette, J.-F. (2011). "A material history of bits." Journal of the American Society for Information Science and Technology **62**(6): 1042-57.

Boden, M. A. (2011). "Against Constructivism." Constructivist Foundations **6**(1): 84-89.

Boland, R. (1987). "The in-formation of information systems", in: R. Boland and R. A. Hirschheim (Eds.), Critical issues in information systems research. Chichester, Wiley: 363-79.

Bollen, J., H. V. d. Sompel, J. A. Smith and R. Luce (2005). "Toward alternative metrics of journal impact: A comparison of download and citation data." Information Processing and Management **41**(6): 1419–40.

Bolter, J. and D. Grusin (2001). Remediation: Understanding new media. Cambridge, MA, MIT Press.

Borgmann, A. (1992). Crossing the postmodern divide. Chicago, University of Chicago Press.

Borgmann, A. (1999). Holding on to reality: The nature of information at the turn of the millennium. Chicago, University of Chicago Press.

Borgmann, A. (2010). "Reality and technology." Cambridge Journal of Economics **34**(1): 27-35.

Borneman, J. (1997). Settling accounts: Violence, justice, and accountability in Postsocialist Europe. Princeton, Princeton University Press.

Bowker, G. C. (1997). "Lest we remember: Organizational forgetting and the production of knowledge." Accounting, Management and Information Technologies **7**(3): 113-38.

Bowker, G. C. (2005). Memory practices in the sciences. Cambridge, MA, MIT Press.

Bowker, G. C. and S. L. Star (1999). Sorting things out: Classification and its consequences. Cambridge, MA, MIT Press.

Boyden, M. (2003). "The rhetoric of forgetting: Elena Esposito on social memory." Image [&] Narrative - Online Magazine of the Visual Narrative **6**. http://www.imageandnarrative.be/mediumtheory/michaelboyden.htm

Braf, E. (2001). Knowledge or information - What makes the difference? IFIP TC8/WG8.1, Montreal, Canada, Kluwer Academic.

Branigan, T. (2010). Google to end censorship in China over cyber attack. Guardian. London http://www.guardian.co.uk/technology/2010/jan/12/google-china-ends-censorship

Brier, S. (2008). Cybersemiotics: Why information is not enough! Toronto, University of Toronto Press.

Brier, S. (2009). "Cybersemiotic pragmaticism and constructivism." Constructivist Foundations **5**(1): 19-38.

Brindley, L. (2009). We're in danger of losing our memories. The Observer. London http://www.guardian.co.uk/technology/2009/jan/25/internet-heritage

Brockmeier, J. (2002). "Remembering and forgetting: Narrative as cultural memory." Culture & Psychology **8**(1): 15-43.

Brown, J. S. and P. Duguid (2000). The social life of information. Boston, Harvard Business School Press.

Bryant, A. (2002). "Re-grounding grounded theory." Journal of Information Technology **4**(1). http://aisel.aisnet.org/jitta/vol4/iss1/7

Bryant, A. (2008). "The future of information systems - Thinking informatically." European Journal of Information Systems **17**(6): 695-98.

Buckland, M. K. (1991). "Information as thing." American Society for Information Science **42**(5): 351-60.

Buckland, M. K. (1997). "What is a "document"?" Journal of the American Society for Information Science **48**(9): 804-09.

Bussola, E. (2011). "What is a memory, that it may comprehend itself? Self-referential implications of the phenomenology of remembering", in: G. Campardo, F. Tiziani and M. Iaculo (Eds.), Memory mass storage. Berlin, Heidelberg, Springer: 1-57.

Bygstad, B. and B. E. Munkvold (2011). "Exploring the role of informants in interpretive case study research in IS." Journal of Information Technology **26**(1): 32-45.

Byrne, G. and L. Goddard (2010). "The strongest link: Libraries and linked data." D-Lib **16**(11/12). http://www.dlib.org/dlib/november10/byrne/11byrne.html

Campbell, D. G. (2006). "Metadata, metaphor, and metonymy", in: R. P. Smiraglia (Ed.), Metadata: A cataloger's primer. New York, Haworth Press: 57-73.

Capurro, R. and B. Hjorland (2003). "The concept of information." Annual Review of Information Science and Technology **37**(1): 343-411.

Castells, M. (1996). The rise of the network society. Oxford, Blackwell.

Cecez-Kecmanovic, D. (2011). "Doing critical information systems research – Arguments for a critical research methodology." European Journal of Information Systems **20**(4): 440-55.

Charmaz, K. (2006). Constructing grounded theory: A practical guide through qualitative analysis. London, Sage.

Checkland, P. B. and J. Scholes (1990). Soft systems methodology in action. New York, John Wiley & Sons.

Choi, S. Y., H. Lee and Y. Yoo (2010). "The impact of information technology and transactive memory systems on knowledge sharing, application, and team performance: A field study." MIS Quarterly **34**(4): 855-70.

Ciborra, C. (2002). The labyrinths of information: Challenging the wisdom of systems. Oxford, Oxford University Press.

Ciuk, S. and M. Kostera (2010). "Drinking from the waters of Lethe: A tale of organizational oblivion." Management Learning **41**(2): 187-204.

CLIR and Library of Congress (2002). Building a national strategy for digital preservation: Issues in digital media archiving. Washington D.C.

Coleman, A. S. (2006). "From cataloging to metadata: Dublin Core records for the library catalog", in: R. P. Smiraglia (Ed.), Metadata: A cataloger's primer. New York, Haworth Press: 153-81.

Collier, M. (2004). "Development of a business plan for an international co-operative digital library - The European Library (TEL)." Program: Electronic Library & Information Systems **38**(4): 225-31.

Collins, H., R. Evans, R. Ribeiro and M. Hall (2006). "Experiments with interactional expertise." Studies in History and Philosophy of Science **37**(4): 656-74.

Confino, A. (1997). "Collective memory and cultural history: Problems of method." The American Historical Review **102**(5): 1386-403.

Confrey, J. (2011). "The transformational epistemology of radical constructivism: A tribute to Ernst von Glasersfeld." Constructivist Foundations **6**(2): 177-82.

Connerton, P. (1990). How societies remember. Cambridge, Cambridge University Press.

Connerton, P. (2008). "Seven types of forgetting." Memory Studies **1**(1): 59-71.

Cook, T. (1997). "What is past is prologue: A history of archival ideas since 1898, and the future paradigm shift." Archivaria **43**: 17-63.

Cooper, R. (2005). "Relationality." Organization Studies **26**(11): 1688-710.

Corbin, J. M. and A. L. Strauss (1990). "Grounded theory research: Procedures, canons, and evaluative criteria." Qualitative Sociology **13**(1): 3-21.

Corbin, J. M. and A. L. Strauss (2008). Basics of qualitative research: Techniques and procedures for developing grounded theory. Los Angeles, Sage.

Cousins, J. (2010). LSE research on Europeana. A. Marton. The Hague

Covi, L. and R. Kling (1996). "Organizational dimensions of effective digital library use: Closed rational and open natural systems models." Journal of the American Society for Information Science **47**(9): 672-89.

Cox, R. J. (2007). "Machines in the archives: Technology and the coming transformation of archival reference." First Monday **12**(11). http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2029/1894

Coyle, K. (2006). "Mass digitization of books." The Journal of Academic Librarianship **32**(6): 641-45.

Craig, B. L. (1992). "Outward visions, inward glance: Archives history and professional identity." Archival Issues **17**(2): 113-24.

Crane, S. A. (1997). "Memory, distortion, and history in the museum." History and Theory **36**(4): 44-63.

Crang, M., P. Crang and J. May, Eds. (1999). Virtual geographies: Bodies, space, and relations. London, Routledge.

Czarniawska, B. (2004). Narratives in social science research. London, Sage.

Dahl, S. (1958). History of the book. New York, Scarecrow Press.

Dahrendorf, R. (1972). Homo Sociologicus. Ein Versuch zur Geschichte, Bedeutung und Kritik der Kategorie der sozialen Rolle. Opladen, Westdeutscher Verlag.

Davis, J. R. and C. Lagoze (2000). "NCSTRL: Design and deployment of a globally distributed digital library." Journal of the American Society for Information Science **51**(3): 273-80.

Dekkers, M. (2001). "Application profiles, or how to mix and match metadata schemas." Cultivate Interactive **3**. http://www.cultivate-int.org/issue3/schemas/

Dempsey, L. (1999). "Scientific, industrial, and cultural heritage: A shared approach. A research framework for digital libraries, museums and archives." Ariadne **22**. http://www.ariadne.ac.uk/issue22/dempsey/

Derrida, J. (1976). Grammatology. Baltimore, John Hopkins.

Derrida, J. (1982). "Differance", in: J. Derrida (Ed.), Margins of philosophy. Chicago, University of Chicago Press: 3-27.

Devlin, K. J. (1999). Infosense: Turning information into knowledge. New York, W.H. Freeman.

Diekmann, A. (2001). Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen. Hamburg, Rohwolt.

Dijck, J. v. (2007). Mediated memories in the digital age. Stanford, Stanford University Press.

Dimbath, O. (2011). "Systemvergessen. Zur Vergesslichkeit sozialer Systeme", in: O. Dimbath and P. Wehling (Eds.), Soziologie des Vergessens. Theoretische Zugänge und empirische Forschungsfelder. Konstanz, UVK: 139-66.

Dosi, G. (1982). "Technological paradigms and technological trajectories." Research Policy **11**(3): 147-62.

Dretske, F. I. (1981). Knowledge and the flow of information. Oxford, Blackwell.

Dreyfus, H. L. and C. Spinosa (1997). "Highway bridges and feasts: Heidegger and Borgmann on how to affirm technology." Man and World **30**: 159-77.

Duguid, P. (2007). "Inheritance and loss? A brief survey of Google Books." First Monday **12**(8). http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/1972/1847

Earnshaw, R. and J. Vince, Eds. (2008). Digital convergence - Libraries of the future. London, Springer.

Eden, A. H. (2007). "Three paradigms of computer science." Minds and Machines **17**(2): 135-67.

Edwards, L. (2009). Cultural heritage and the social sciences: The European perspective. IFLA General Conference, Milan, Italy.

Eisenhardt, K. M. (1989). "Building theories from case study research." The Academy of Management Review **14**(4): 532-50.

Ekbia, H. R. (2009a). "Digital artifacts as quasi-objects: Qualification, mediation, and materiality." Journal of the American Society for Information Science and Technology **60**(12): 2554-66.

Ekbia, H. R. (2009b). "Information in action: A situated view." Proceedings of the American Society for Information Science and Technology **46**(1): 1-11.

Elings, M. W. and G. Waibel (2007). "Metadata for all: Descriptive standards and metadata sharing across libraries, archives and museums." First Monday **12**(3). http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1628/1543

Elliott, R., C. T. Fischer and D. L. Rennie (1999). "Evolving guidelines for publication of qualitative research studies in psychology and related fields." British Journal of Clinical Psychology **38**(3): 215-29.

Ercegovac, Z. (1997). "The interpretations of library use in the age of digital libraries: Virtualizing the name." Library and Information Science Research **19**(1): 35-51.

Erll, A. and A. Nünning, Eds. (2008). Cultural memory studies: An international and interdisciplinary handbook. Berlin, New York, Walter de Gruyter.

Eschenfelder, K. R. and G. Agnew (2010). "Technologies employed to control access to or use of digital cultural collections: Controlled online collections." D-Lib **16**(1/2). http://www.dlib.org/dlib/january10/eschenfelder/01eschenfelder.html

Esposito, E. (1996). "From self-reference to autology: How to operationalize a circular approach." Social Science Information **35**(2): 269-81.

Esposito, E. (2002). Soziales Vergessen: Formen und Medien des Gedächtnisses der Gesellschaft. Frankfurt am Main, Suhrkamp.

Esposito, E. (2004). "The arts of contingency." Critical Inquiry **31**(1). http://www.uchicago.edu/research/jnl-crit-inq/features/artsstatements/arts.esposito.htm

Esposito, E. (2008). "Social forgetting: A systems-theory approach", in: A. Erll and A. Nünning (Eds.), Cultural memory studies: An international and interdisciplinary handbook. Berlin, New York, Walter de Gruyter: 181-89.

European Commission (2005a). i2010 – A European information society for growth and employment. Brussels, European Union

European Commission (2005b). i2010 – Digital libraries. Brussels, European Union

European Commission (2006). Commission recommendation on the digitisation and online accessibility of cultural material and digital preservation. Brussels, European Union

European Commission (2008a). Europe's cultural heritage at the click of a mouse. Progress on the digitisation and online accessibility of cultural material and digital preservation across the EU. Brussels, European Union

European Commission (2008b). Europeana - Next steps. Brussels, European Union

European Commission (2010a). Assessment of the orphan works issue and costs for rights clearance. Brussels, European Union

European Commission (2010b). A comprehensive approach on personal data protection in the European Union. Brussels, European Union

Factor, M., E. Henis, D. Naor, S. Rabinovici-Cohen, P. Reshef, S. Ronen, G. Michetti and M. Guercio (2009). "Authenticity and provenance in long term digital preservation: Modeling and implementation in preservation aware storage", in: 1st Workshop on Theory and Practice of Provenance. San Francisco, CA, USENIX Association: 1-10.

Fairthorne, R. A. (1968). Towards information retrieval. Hamden, Archon Books.

Faulkner, P. and J. Runde (2011). The social, the material, and the ontology of non-material technological objects. 27th EGOS Colloquium, Gothenburg, Sweden.

Feldman, M. S., J. Bell and M. T. Berger (2003). Gaining access: A practical and theoretical guide for qualitative researchers. Walnut Creek, AltaMira.

Fentress, J. and C. Wickham (1992). Social memory. Cambridge, MA, Blackwell.

Fetzer, J. H. (2004). "Information: Does it have to be true?" Minds and Machines **14**: 223-29.

Flick, U. (2006). An introduction to qualitative research. London, Sage.

Flick, U. (2007). Managing the quality of qualitative research. London, Sage.

Floridi, L. (2003). "Two approaches to the philosophy of information." Minds and Machines **13**: 459-69.

Floridi, L. (2004). "Outline of a theory of strongly semantic information." Minds and Machines **14**: 197-221.

Floridi, L. (2009). "Philosophical conceptions of information", in: G. Sommaruga (Ed.), Formal theories of information: From Shannon to semantic information theory and general concepts of information. Berlin, New York, Springer: 13-53.

Flusser, V. (2002). Die Schrift. Hat Schreiben Zukunft? Göttingen, European Photography.

Foerster, H. v. (1967). "Biological principles of information storage and retrieval", in: A. Kent, O. Taulbee, J. Beltzer and G. Goldstein (Eds.), Electronic handling of information: Testing and evaluation. London, Academic Press: 123-47.

Foerster, H. v. (1993). KybernEthik. Berlin, Merve.

Foerster, H. v. (2003). Understanding understanding: Essays on cybernetics and cognition. New York, Springer.

Franke, K. M. (2003). "On discipline and canon." Columbia Journal of Gender and Law **12**(3): 639-43.

Froese, T. (2010). "From cybernetics to second-order cybernetics: A comparative analysis of their central ideas." Constructivist Foundations **5**(2): 75-85.

Gantz, J. F., C. Chute, A. Manfrediz, S. Minton, D. Reinsel, W. Schlichting and A. Toncheva (2008). The diverse and exploding digital universe: An updated forecast of worldwide information growth through 2011. Framingham, MA, IDC http://www.emc.com/collateral/analyst-reports/diverse-exploding-idc-exec-summary.pdf

Garde-Hansen, J., A. Hoskins and A. Reading (2009a). "Introduction", in: J. Garde-Hansen, A. Hoskins and A. Reading (Eds.), Save as... digital memories. New York, Palgrave Macmillan: 1-21.

Garde-Hansen, J., A. Hoskins and A. Reading, Eds. (2009b). Save as... digital memories. New York, Palgrave Macmillan.

Gaskell, G. and M. Bauer (2000). "Towards public accountability: Beyond sampling, reliability and validity", in: M. Bauer and G. Gaskell (Eds.), Qualitative researching with text, image and sound: A practical handbook. London, SAGE: 336-50.

George, A. L. and A. Bennett (2005). Case studies and theory development in the social sciences. Cambridge, MA, MIT Press.

Gerolimos, M. and R. Konsta (2011). "Services for academic libraries in the new era." D-Lib **17**(7/8). http://www.dlib.org/dlib/july11/gerolimos/07gerolimos.html

Gherardi, S. and D. Nicolini (2002). "Learning in a constellation of interconnected practices: Canon or dissonance?" Journal of Management Studies **39**(4): 419-36.

Gibbs, G. R. (2007). Analyzing qualitative data. London, Sage.

Gibson, B., J. Gregory and P. G. Robinson (2005). "The intersection between systems theory and grounded theory: The emergence of the grounded systems observer." Qualitative Sociology Review **1**(2): 3-21.

Giddens, A. (1984). The constitution of society: Outline of the theory of structuration. Cambridge, Polity.

Giddens, A. (1987). Social theory and modern sociology. Cambridge, Polity.

Gladney, H. M. (2006). "Principles for digital preservation." Communications of the ACM **49**(2): 111-16.

Gladney, H. M. (2007). "Digital preservation in a national context. Questions and views of an outsider." D-Lib **13**(1/2). http://www.dlib.org/dlib/january07/gladney/01gladney.html

Glaser, B. G. and A. L. Strauss (1967). The discovery of grounded theory; Strategies for qualitative research. Chicago, Aldine.

Glasersfeld, E. v. (1991). Distinguishing the observer: An attempt at interpreting Maturana http://www.oikos.org/vonobserv.htm

Glasersfeld, E. v. (1997). Radikaler Konstruktivismus. Ideen, Ergebnisse, Probleme. Frankfurt am Main, Suhrkamp.

Golafshani, N. (2003). "Understanding reliability and validity in qualitative research." The Qualitative Report **8**(4): 597-607.

Golden-Biddle, K. and K. Locke (2007). Composing qualitative research. Thousand Oaks, Sage.

Gook, B. (2011). "Being there is everything!" Memory Studies **4**(1): 13-22.

Graham, G. (2005). "Strange bedfellows? Information systems and the concept of a library." Interdisciplinary Science Reviews **30**(2): 137-44.

Gray, P. H., S. Parise and B. Iyer (2011). "Innovation impacts of using social bookmarking systems." MIS Quarterly **35**(3): 613-27.

Green, H. (2002). A library as big as the world. Businessweek. New York http://www.businessweek.com/technology/content/feb2002/tc20020228_1080.htm

Greenberg, J. (2006). "Understanding metadata and metadata schemes", in: R. P. Smiraglia (Ed.), Metadata: A cataloger's primer. New York, Haworth Press: 17-36.

Greene, S., G. Marchionini, C. Plaisant and B. Shneiderman (2000). "Previews and overviews in digital libraries: Designing surrogates to support visual

information seeking." Journal of the American Society for Information Science **51**(4): 380-93.

Gregor, S. (2006). "The nature of theory in information systems." MIS Quarterly **30**(3): 611-42.

Griffin, S., C. Peters and C. Thanos (2005). "Towards the new-generation digital libraries: Recommendations of the NSF/EU-DELOS working groups." International Journal of Digital Libraries **5**(4): 253-54.

Gudehus, C., A. Eichenberg and H. Welzer, Eds. (2010). Gedächtnis und Erinnerung: Ein interdisziplinäres Handbuch. Stuttgart, Weimar, J.B. Metzler.

Habermas, J. (1987). Theorie des kommunikativen Handelns. Frankfurt am Main, Suhrkamp.

Habermas, J. and N. Luhmann (1971). Theorie der Gesellschaft oder Sozialtechnologie - Was leistet die Systemforschung? Frankfurt am Main, Suhrkamp.

Hagen, W. (2011). "Medienvergessenheit. Über Gedächtnis und Erinnerung in massenmedial orientierten Netzwerken", in: O. Dimbath and P. Wehling (Eds.), Soziologie des Vergessens. Theoretische Zugänge und empirische Forschungsfelder. Konstanz, UVK: 243-73.

Hahn, T. B. (1998). "Text retrieval online: Historical perspective on web search engines." ASIS(Apr/May). http://www.asis.org/Bulletin/Apr-98/hahn.html

Halbwachs, M. (1992). On collective memory. Chicago, University of Chicago Press.

Hanseth, O. and K. Braa (2000). "Globalization and 'risk society'", in: C. Ciborra (Ed.), From control to drift: The dynamics of corporate information infrastructures. Oxford, Oxford University Press: 41-55.

Hanseth, O. and K. Lyytinen (2010). "Design theory for dynamic complexity in information infrastructures: The case of building internet." Journal of Information Technology **25**(1): 1-19.

Heery, R. (1996). "Review of metadata formats." Program: Electronic Library & Information Systems **30**(4): 345 - 73.

Heidegger, M. (1969). Identity and difference. New York, Harper & Row.

Heidegger, M. (1996). Being and time. Albany, NY, State University of New York Press.

Herzhoff, J. (2009). The ICT convergence discourse in the information systems literature - A second-order observation. 17th ECIS, Verona, Italy.

Himanen, P. (2001). The hacker ethic and the spirit of the information age. London, Vintage.

Hirtle, P. B. (2000). Archival authenticity in a digital age, in: Building a national strategy for digital preservation: Issues in digital media archiving. Report commissioned by the CLIR and Library of Congress. Washington D.C., pp. 8-23.

Hirtle, P. B. (2008). "Copyright renewal, copyright restoration, and the difficulty of determining copyright status." D-Lib **14**(7/8). http://www.dlib.org/dlib/july08/hirtle/07hirtle.html

Hitchcock, S., T. Brody, J. M. N. Hey and L. Carr (2007). "Digital preservation service provider models for institutional repositories." D-Lib **13**(5/6). http://www.dlib.org/dlib/may07/hitchcock/05hitchcock.html

Hjorland, B. (2000). "Documents, memory institutions and information science." Journal of Documentation **56**(1): 27-41.

HLEG on Digital Libraries (2009). Digital libraries: Recommendations and challenges for the future. Final report. Brussels, European Commission

Holley, R. (2010). "Crowdsourcing: How and why should libraries do it?" D-Lib **16**(3/4). http://www.dlib.org/dlib/march10/holley/03holley.html

Horkheimer, M. and T. W. Adorno (1988). Dialektik der Aufklärung. Philosophische Fragmente. Frankfurt am Main, Fischer.

Hoskins, A. (2009a). "Flashbulb memories, psychology and media studies: Fertile ground for interdisciplinarity?" Memory Studies **2**(2): 147-50.

Hoskins, A. (2009b). "The mediatisation of memory", in: J. Garde-Hansen, A. Hoskins and A. Reading (Eds.), Save as... digital memories. New York, Palgrave Macmillan: 27-43.

House, N. v. (2009). "Technologies of memory: Key issues and critical perspectives." Memory Studies **1**(3): 295-310.

Howarth, L. C. (2006). "Metadata and bibliographic control: Soul-mates or two solitudes?" in: R. P. Smiraglia (Ed.), Metadata: A cataloger's primer. New York, Haworth Press: 37-56.

Howell, B. A. (2006). "Proving web history: How to use the Internet Archive." Journal of Internet Law **9**(8): 3-9.

Hunt, R. R., R. E. Smith and K. R. Dunlap (2011). "How does distinctive processing reduce false recall?" Memory and Language **65**(4): 378-89.

IFLA (1998). Functional requirements for bibliographic records. Munich, K.G. Sauer Verlag.

Introna, L. D. (1997a). Management, information and power: A narrative of the involved manager. London, Macmillan.

Introna, L. D. (1997b). "Privacy and the computer: Why we need privacy in the information society." Metaphilosophy **28**(3): 259-75.

Introna, L. D. and H. Nissenbaum (2000). "Shaping the web: Why the politics of search engines matter." The Information Society **16**(3): 169-85.

Irwin-Zarecka, I. (1993). Frames of remembrance: Social and cultural dynamics of collective memory. New Brunswick, Transaction Publishers.

Jensen, K. B. (1991). "Introduction: The qualitative turn", in: K. B. Jensen and N. Jankowski (Eds.), A handbook of qualitative methodologies for mass communication research. London, Routledge: 1-11.

Jochum, U. (1999). "The Alexandrian Library and its aftermath." Library History **15**: 5-12.

Jones, A. L. (1993). "Exploding canons: The anthropology of museums." Annual Review of Anthropology **22**: 201-20.

Kallinikos, J. (2006a). The consequences of information: Institutional implications of technological change. Northampton, MA, Edward Elgar.

Kallinikos, J. (2006b). "Information out of information: On the self-referential dynamics of information growth." Information Technology and People **19**(1): 98-115.

Kallinikos, J. (2009a). "The making of ephemeria: On the shortening life spans of information." The International Journal of Interdisciplinary Social Sciences **4**(3): 227-36.

Kallinikos, J. (2009b). "On the computational rendition of reality: Artefacts and human agency." Organization **16**(2): 183-202.

Kallinikos, J. (2010). Living without lives. Experience and memory in the internet age. Telos http://www.telos-eu.com/fr/article/living-without-lives-experience-and-memory-in-th

Kallinikos, J. (2011). Governing through technology: Information artefacts and social practice. Basingstoke, Palgrave Macmillan.

Kallinikos, J., A. Aaltonen and A. Marton (2010a). "A theory of digital objects." First Monday **15**(6). http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3033/2564

Kallinikos, J., A. Aaltonen and A. Marton (2011). "Information generativity and logics of innovation." Working Paper.

Kallinikos, J., H. Hasselbladh and A. Marton (2010b). Governing social practice: technology versus institutions. EGOS, Lisbon.

Kallinikos, J. and J.-C. Mariategui (2011). "Video as digital object: Production and distribution of video content in the internet media ecosystem." The Information Society(forthcoming).

Kansteiner, W. (2002). "Finding meaning in memory: A methodological critique of collective memory studies." History and Theory **41**(2): 179-97.

Kansteiner, W. (2010). "Memory, media and Menschen: Where is the individual in collective memory studies?" Memory Studies **3**(1): 3-4.

Kapadia, A., T. Henderson, J. J. Fielding and D. Kotz (2007). Virtual walls: Protecting digital privacy in pervasive environments? 5th International Conference on Pervasive Computing, Toronto, Canada.

Kemper, E. A., S. Stringfield and C. Teddlie (2003). "Mixed methods sampling strategies in social science research", in: A. Tashakkori and C. Teddlie (Eds.), Handbook of mixed methods in social and behavioral research. Thousand Oaks, Sage: 3-50.

King, M. (1993). "The "truth" about autopoiesis." Journal of Law & Society **20**(2): 218-36.

King, M. and C. Thornhill (2003). ""Will the real Niklas Luhmann stand up, please". A reply to John Mingers." The Sociological Review **51**(2): 276–85.

Kirk, J. and M. L. Miller (1986). Reliability and validity in qualitative research. Beverly Hills, Sage Publications.

Kjaer, P. (2006). "Systems in context. On the outcome of the Habermas/Luhmann-debate." Ancilla Iuris **66**: 66-77.

Klein, H. K. and M. D. Myers (1999). "A set of principles for conducting and evaluating interpretive field studies in information systems." MIS Quarterly **23**(1): 67-93.

Klein, K. L. (2000). "On the emergence of memory in historical discourse." Representations **69**: 127-50.

Klump, J. (2011). "Criteria for the trustworthiness of data centres." D-Lib **17**(1/2). http://www.dlib.org/dlib/january11/klump/01klump.html

Knell, S. J. (2003). "The shape of things to come: Museums in the technological landscape." Museum and Society **1**(3): 132-46.

Knorr Cetina, K. (1999). Epistemic cultures: How the sciences make knowledge. Cambridge, MA, Harvard University Press.

Knorr-Cetina, K. (1989). "Spielarten des Konstruktivismus." Soziale Welt **40**(1): 86-96.

Kreps, D. (2010). "My social networking profile: Copy, resemblance, or simulacrum? A poststructuralist interpretation of social information systems." European Journal of Information Systems **19**(1): 104-15.

Krippendorff, K. (2009). On communicating: Otherness, meaning, and information. New York, Routledge.

Krohn, W. (1989). "Die Verschiedenheit der Technik und die Einheit der Techniksoziologie", in: P. Weingart (Ed.), Technik als sozialer Prozess. Frankfurt am Main, Suhrkamp: 15-43.

Krohn, W. and G. Küppers (1989). Die Selbstorganisation der Wissenschaft. Frankfurt am Main, Suhrkamp.

Kuhn, A. (2010). "Memory texts and memory work: Performances of memory in and with visual media." Memory Studies **3**(4): 298-313.

Kuhn, T. S. (1996). The structure of scientific revolutions. Chicago, University of Chicago Press.

Kuny, T. and G. Cleveland (1996). The digital library: Myths and challenges. IFLA General Conference, Beijing, China. http://www.ifla.org/IV/ifla62/62-kuny.pdf

Kvale, S. (1994). "Ten standard objections to qualitative research interviews." Journal of Phenomenological Psychology **25**(2).

Landheer, B. (1957). Social functions of libraries. New York, Scarecrow.

Landon, G. V. (2009). "Toward digitizing all forms of documentation." D-Lib **15**(3/4). http://www.dlib.org/dlib/march09/landon/03landon.html

Lanzara, G. F. (1991). "Shifting stories. Learning from a reflective experiment in a design process", in: D. A. Schon (Ed.), The reflective turn: Case studies in and on educational practice. New York, Teachers College Press of Columbia University: pp. 285-320.

Lanzara, G. F. (2009). "Reshaping practice across media: Material mediation, medium specificity and practical knowledge in judicial work." Organization Studies **30**(12): 1369-90.

Latour, B. and S. Woolgar (1986). Laboratory life: The construction of scientific facts. Princeton, Princeton University Press.

Lavoie, B. and L. Dempsey (2009). "Beyon 1923. Characteristics of potentially in-copyright print books in library collection." D-Lib **15**(11/12). http://www.dlib.org/dlib/november09/lavoie/11lavoie.html

Law, D. (2008). "Beyond the hybrid library: Libraries in a Web 2.0 world", in: R. Earnshaw and J. Vince (Eds.), Digital convergence - Libraries of the future. London, Springer: 107-18.

Law, J. (2004). After method: Mess in social science research. London, New York, Routledge.

Le Moigne, J.-L. (2011). "From Jean Piaget to Ernst von Glasersfeld: An epistemological itinerary in review." Constructivist Foundations **6**(2): 152-56.

Lee, A. S. (2004). "Thinking about social theory and philosophy for information systems", in: L. Willcocks and J. Mingers (Eds.), Social theory and philosophy for information systems. Chichester, John Wiley & Sons: 1-26.

Lee, A. S. and R. L. Baskerville (2003). "Generalizing generalizability in information systems research." Information Systems Research **14**(3): 221-43.

Lee, A. S. and G. S. Hubona (2009). "A scientific basis for rigor in information systems research." MIS Quarterly **33**(2): 237-62.

Lee, D. (2000a). "The society of society: The grand finale of Niklas Luhmann." Sociological Theory **18**(2): 320-30.

Lee, H.-L. (2000b). "What is a collection?" Journal of the American Society for Information Science **51**(12): 1106-13.

Leetaru, K. (2008). "Mass book digitization: The deeper story of Google Books and the Open Content Alliance." First Monday **13**(10). http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2101/2037

Leggewie, H. (1987). "Interpretation und Validierung biographischer Interviews", in: G. Jüttemann and H. Thomae (Eds.), Biographie und Psychologie. Berlin, Springer: 138-50.

Lessig, L. (2006). Is Google Book Search "Fair Use"? http://www.youtube.com/watch?v=TmU2i1hQiN0

Li, L. (2006). Building the ubiquitous library in the 21st century. IFLA General Conference, Seoul, Korea. http://www.ifla.org/IV/ifla72/papers/140-Li-en.pdf

Liebenau, J. and J. Backhouse (1990). Understanding information: An introduction. London, Macmillan.

Lin, L., X. Geng and A. B. Whinston (2005). "A sender-receiver framework for knowledge transfer." MIS Quarterly **29**(9): 197-219.

Lincoln, Y. S. and E. G. Guba (1985). Naturalistic inquiry. Beverly Hills, Sage.

Littlejohn, S. W. (1992). Theories of human communication. Belmont, Wadsworth.

Locke, C. (2000). "Digital memory and the problem of forgetting", in: S. Radstone (Ed.), Memory and methodology. Oxford, Berg: 25-36.

Lohr, S. (2010). Library of Congress will save tweets. The New York Times. New York http://www.nytimes.com/2010/04/15/technology/15twitter.html

Luhmann, N. (1986). "The autopoiesis of social systems", in: F. Geyer and J. v. d. Zouwen (Eds.), Sociocybernetic paradoxes: Observation, control and evolution of self-steering systems. London, Sage: 171-92.

Luhmann, N. (1991). Soziologie des Risikos. Berlin, New York, Walter de Gruyter.

Luhmann, N. (1996a). Die Realität der Massenmedien. Opladen, Westdeutscher Verlag.

Luhmann, N. (1996b). Soziale Systeme: Grundriss einer allgemeinen Theorie. Frankfurt am Main, Suhrkamp.

Luhmann, N. (1997). "The control of intransparency." Systems Research and Behavioral Science **14**(6): 359-71.

Luhmann, N. (1998). Die Gesellschaft der Gesellschaft. Frankfurt am Main, Suhrkamp.

Luhmann, N. (2002a). "Deconstruction as second-order observing", in: W. Rasch (Ed.), Theories of distinction: Redescribing the description of modernity. Stanford, Stanford University Press: 94-112.

Luhmann, N. (2002b). Die Wissenschaft der Gesellschaft. Frankfurt am Main, Suhrkamp.

Luhmann, N. (2006). "System as difference." Organization **13**(1): 37-57.

Luhmann, N. and D. Baecker, Eds. (2002). Einführung in die Systemtheorie. Heidelberg, Carl-Auer-Systeme Verlag.

Lutz, C. E. (1978). "The oldest library motto." The Library Quarterly **48**(1): 36-39.

Luyt, B. (2011). "The nature of historical representation on Wikipedia: Dominant or alterative historiography?" Journal of the American Society for Information Science and Technology **62**(6): 1058-65.

Lyman, P. (2002). Archiving the World Wide Web, in: Building a national strategy for digital preservation: Issues in digital media archiving. Report commissioned by the CLIR and Library of Congress. Washington D.C., pp. 38-51.

Lyman, P. and B. Kahle (1998). "Archiving digital cultural artifacts." D-Lib(July/August). http://www.dlib.org/dlib/july98/07lyman.html

Machlup, F. and U. Mansfield, Eds. (1983). The study of information: Interdisciplinary messages. New York, Wiley.

Madill, A., A. Jordan and C. Shirley (2000). "Objectivity and reliability in qualitative analysis: Realist, contextualist and radical constructionist epistemologies." British Journal of Psychology **91**(1): 1-20.

Maier, C. S. (1988). The unmasterable past: History, holocaust, and German national identity. Cambridge, MA, Harvard University Press.

Manovich, L. (2001). The language of new media. Cambridge, MA, MIT Press.

Marcum, J. W. (2001). "From information center to discovery system: Next step for libraries?" The Journal of Academic Librarianship **27**(2): 97-106.

Martinez-Delgado, A. (2002). "Radical constructivism: Between realism and solipsism." Science Education **86**(6): 840-55.

Marton, A. (2007). Keeping information vivid: The implications of information growth on librarian sense-making. An exploratory case study on the Austrian National Library. Information Systems and Innovations Group, Department of Management. London, LSE

Marton, A. (2009). "Self-referential technology and the growth of information. From techniques to technology to the technology of technology." Soziale Systeme **15**(1): 137-59.

Marton, A. (2010). "The transfigurability of digital objects." Artnodes **10**. http://artnodes.uoc.edu/ojs/index.php/artnodes/article/view/n10-marton/n10-marton-eng

Marty, P. F. (2008). "Museum websites and museum visitors: Digital museum resources and their use." Museum Management and Curatorship **23**(1): 81-89.

Massart, D., E. Shulman, N. Nicholas, N. Ward and F. Bergeron (2010). "Taming the metadata beast: ILOX." D-Lib **16**(11/12). http://www.dlib.org/dlib/november10/massart/11massart.html

Masuda, Y. (2003). "Image of the future information society", in: F. Webster and E. Puoskari (Eds.), The information society reader. London, Routledge: 15-20.

Maturana, H. R. (1988). "Reality: The search for objectivity or the quest for a compelling argument." The Irish Journal of Psychology **9**(1): 25-82.

Maturana, H. R., A. Paucar-Caceres and R. Harnden (2011). "Origins and implications of autopoiesis. Preface to the second edition of De Maquinas y Seres Vivos." Constructivist Foundations **6**(3): 293-306.

Maturana, H. R. and F. J. Varela (1992). The tree of knowledge: The biological roots of human understanding. Boston, New York, Shambhala.

Mayer-Schönberger, V. (2009). Delete: The virtue of forgetting in the digital age. Princeton, Princeton University Press.

McKie, R. and V. Thorpe (2002). Digital Domesday Book lasts 15 years not 1000. The Observer http://www.guardian.co.uk/uk/2002/mar/03/research.elearning

McKinney, E. H. and C. J. Yoos (2010). "Information about information: A taxonomy of views." MIS Quarterly **34**(2): 329-44.

Mead, G. H. and C. W. Morris (1934). Mind, self, and society: From the standpoint of a social behaviorist. Chicago, University of Chicago Press.

Megill, A. (1998). "History, memory, identity." History of the Human Sciences **11**(3): 37-62.

Merali, Y. (2006). "Complexity and information systems: The emergent domain." Journal of Information Technology **21**(4): 216-28.

Meyyappan, N., G. G. Chowdhury and S. Foo (2000). "A review of the status of 20 digital libraries." Journal of Information Science **26**(5): 337-55.

Mingers, J. (2002). "Can social systems be autopoietic? Assessing Luhmann's social theory." The Sociological Review **50**(2): 278-99.

Moser, S., Ed. (2004a). Konstruktivistisch Forschen: Methodologie, Methoden, Beispiele. Wiesbaden, VS Verlag für Sozialwissenschaften.

Moser, S. (2004b). "Konstruktivistisch Forschen? Prämissen und Probleme einer konstruktivistischen Methodologie", in: S. Moser (Ed.), Konstruktivistisch Forschen: Methodologie, Methoden, Beispiele. Wiesbaden, VS Verlag für Sozialwissenschaften: 9-42.

Müller, K. H. (2008). "Methodologizing radical constructivism: Recipes for RC-designs in the social sciences." Constructivist Foundations **4**(1): 50-61.

Müller, K. H. (2011). "The two epistemologies of Ernst von Glasersfeld." Constructivist Foundations **6**(2): 220-26.

Myers, M. D. and H. K. Klein (2011). "A set of principles for conducting critical research in information systems." MIS Quarterly **35**(1): 17-36.

Myers, M. D. and M. Newman (2007). "The qualitative interview in IS research: Examining the craft." Information and Organization **17**(1): 2-26.

Nauta, D. (1972). The meaning of information. The Hague, Mouton.

Nietzsche, F. (1887). Die Fröhliche Wissenschaft. Leipzig, E. W. Fritzsch.

Niggemann, E., J. De Decker and M. Lévy (2011). The new Renaissance. Report of the 'comité des sages'. Reflection group on bringing Europe's cultural heritage online. Brussels, European Commission

Nora, P. (1989). "Between memory and history: Les lieux-de-mémoire." Representations **26**: 7-25.

OCLC (2003). Summaries DDC - Dewey Decimal Classification http://www.oclc.org/dewey/resources/summaries/deweysummaries.pdf

Olick, J. K. (1999). "Collective memory: The two cultures." Sociological Theory **17**(3): 333-48.

Olick, J. K. (2008). ""Collective memory": A memoir and prospect." Memory Studies **1**(1): 23-29.

Olick, J. K. (2010). "Das soziale Gedächtnis", in: C. Gudehus, A. Eichenberg and H. Welzer (Eds.), Gedächtnis und Erinnerung: Ein interdisziplinäres Handbuch. Stuttgart, Weimar, J.B. Metzler: 109-14.

Olick, J. K. and J. Robbins (1998). "Social memory studies: From "collective memory" to the historical sociology of mnemonic practices." Annual Review of Sociology **24**(1): 105-40.

Oppenheim, C. and D. Smithson (1999). "What is the hybrid library?" Journal of Information Science **25**(2): 97-112.

Orr, J. M. (1977). Libraries as communication systems. Westport, Greenwood.

Packer, M. J. and R. B. Addison (1989). "Evaluating an interpretive account", in: M. J. Packer and R. B. Addison (Eds.), Entering the circle: Hermeneutic investigation in psychology. Albany, State University of New York Press: 275-92.

Page, B. (2011). New York judge rules against Google books settlement. Guardian. London http://www.guardian.co.uk/books/2011/mar/23/google-books-settlement-ruling

Parsons, T. (1968). "Social interaction", in: D. L. Sills (Ed.), International Encyclopedia of the Social Sciences. New York, Macmillan: 429-40.

Patton, M. Q. (2002). Qualitative research and evaluation methods. Thousand Oaks, Sage.

Pentzold, C. (2009). "Fixing the floating gap: The online encyclopedia Wikipedia as a global memory place." Memory Studies **2**(2): 255-72.

Perra, E. (2010). "Ligitimizing fascism through the Holocaust? The reception of the miniseries Perlasca: un eroe italiano in Italy." Memory Studies **3**(2): 95-109.

Peterson Bishop, A. (1999). "Document structure and digital libraries: How researchers mobilize information in journal articles." Information Processing and Management **35**(3): 255-79.

Peterson Bishop, A., L. J. Neumann, S. L. Star, C. Merkel, E. Ignacio and R. J. Sandusky (2000). "Digital libraries: Situating use in changing information infrastructure." Journal of the American Society for Information Science **51**(4): 394-413.

Petschar, H. (1997). "Die Nationalbibliothek in der Informationsgesellschaft." Relation - Wissenschaft und Öffentlichkeit **4**(2): 169-78.

Petschar, H. (2002). "Vom Zettel- zum Bildkatalog: Das Projekt "Digitales Bildarchiv" an der Österreichischen Nationalbibliothek - Ein Werkstattbericht." Zeitschrift für Bibliothekswesen und Bibliographie **49**(2): 59-68.

Pfeffer, T. (2004). "Die (Re-)Konstruktion sozialer Phänomene durch 'zirkuläres Fragen'", in: S. Moser (Ed.), Konstruktivistisch Forschen: Methodologie, Methoden, Beispiele. Wiesbaden, VS Verlag für Sozialwissenschaften: 67-92.

Piaget, J. (1970). Genetic epistemology. New York, London, Columbia University Press.

Polastron, L. X. (2007). Books on fire. The tumultuous story of the world's great libraries. London, Thames & Hudson.

Poole, R. (2008). "Memory, history and the claims of the past." Memory Studies **1**(2): 149-66.

Powell, J. E., L. Marks Collins and M. L. B. Martinez (2010). "Semantically enhancing collections of library and non-library content." D-Lib **16**(7/8). http://www.dlib.org/dlib/july10/powell/07powell.html

Primary Research Group (2008). Libraries & the mega-internet sites: A survey of how libraries use and relate to Google, Yahoo, Wikipedia, eBay, Amazon, Facebook, YouTube & other mega internet sites. New York, Primary Research Group Inc.

Prior, L. (2003). Using documents in social research. London, Sage Publications.

Radstone, S., Ed. (2000a). Memory and methodology. Oxford, Berg.

Radstone, S. (2000b). "Working with memory: An introduction", in: S. Radstone (Ed.), Memory and methodology. Oxford, Berg: 1-22.

Raitt, D. (2000). "Digital library initiatives across Europe." Computers in Libraries **20**(10): 26-34.

Ranganathan, S. R. (1931). The five laws of library science. London, Edward Goldston.

Rasmussen, J. (1998). "Constructivism and phenomenology. What do they have in common, and how can they be told apart?" Cybernetics and Systems **29**(6): 553-76.

Reading, A. (2009). "Memobilia: The mobile phone and the emergence of wearable memories", in: J. Garde-Hansen, A. Hoskins and A. Reading (Eds.), Save as... digital memories. New York, Palgrave Macmillan: 81-95.

Reicher, S. (2000). "Against methodolatry: Some comments on Elliott, Fischer, and Rennie." British Journal of Clinical Psychology **39**(1): 1-6.

Riegler, A. (2005). "Constructive memory." Kybernetes **34**(1/2): 89-104.

Ristow, W. W. and D. K. Carrington (1971). "Machine-readable map cataloging in the Library of Congress." Special Libraries **62**(9): 343-52.

Roediger, H. L. and J. V. Wertsch (2008). "Creating a new discipline of memory studies." Memory Studies **1**(1): 9-22.

Rusbridge, C. (1998). "Towards the hybrid library." D-Lib(July/August). http://www.dlib.org/dlib/july98/rusbridge/07rusbridge.html

Russell, K., E. Weinberger and A. Stone (1999). "Preserving digital scholarship: The future is now." Learned Publishing **12**(4): 271-80.

Saidis, K. and A. Delis (2007). "Type-consistent digital objects." D-Lib **13**(5/6). http://www.dlib.org/dlib/may07/saidis/05saidis.html

Samuelson, P. (2010). "Google Book Search and the future of books in cyberspace." Minnesota Law Review **94**(5): 1308-74.

Sandell, R. (1998). "Museums as agents of social inclusion." Museum Management and Curatorship **17**(4): 401-18.

Sassen, S. (2004). "Towards a sociology of information technology", in: C. Avgerou, C. Ciborra and F. Land (Eds.), The social study of information and communication technology: Innovation, actors and contexts. Oxford, Oxford University Press: 77-99.

Schmidt, S. J. (2008). "Memory and remembrance: A constructivist approach", in: A. Erll and A. Nünning (Eds.), Cultural memory studies: An international and interdisciplinary handbook. Berlin, New York, Walter de Gruyter: 191-201.

Schmidt, S. J. (2011). "Radical constructivism: A tool, not a super theory!" Constructivist Foundations **6**(1): 6-11.

Schnapp, J. (2008). "Animating the archive." First Monday **13**(8). http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2218/2020

Schwartz, C. (2000). "Digital libraries: An overview." The Journal of Academic Librarianship **26**(6): 385-93.

Sciulli, D. (1994). "An interview with Niklas Luhmann." Theory, Culture & Society **11**(37): 37-68.

Seale, C. (1999). The quality of qualitative research. London, Sage.

Seixas, P. (2007). "Who needs a canon?" in: S. Stuurman (Ed.), Beyond the canon: History for the twenty-first century. Basingstoke ; New York, Palgrave Macmillan: 19-30.

Shadbolt, N., T. Berners-Lee and W. Hall (2006). "The semantic web revisited." Intelligent Systems **21**(3): 96-101.

Shaker, L. (2006). "In Google we trust: Information integrity in the digital age." First Monday **11**(4). http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1320/1240

Shannon, C. E. (1993). "A mathematical theory of communication", in: C. E. Shannon (Ed.), Claude Elwood Shannon: Collected papers. New York, IEEE Press: 1-55.

Shannon, C. E. and W. Weaver (1998). The mathematical theory of communication. Urbana, University of Illinois Press.

Shapiro, C. and H. R. Varian (1999). Information rules: A strategic guide to the network economy. Boston, Harvard Business School Press.

Shirky, C. (2008). Here comes everybody: The power of organizing without organizations. New York, Penguin Press.

Silverman, D. (2005). Doing qualitative research: A practical handbook. London, Thousand Oaks, Sage.

Simon, H. A. (1996). The sciences of the artificial. Cambridge, MA, MIT Press.

Singer, J. A. and M. A. Conway (2009). "Should we forget forgetting?" Memory Studies **1**(3): 279-85.

Sismondo, S. (1993). "Some social constructions." Social Studies of Science **23**(3): 515-53.

Sluis, K. (2010). "Algorithmic memory? Machinic vision and database culture", in: A. Mousoutzanis and D. Riha (Eds.), New media and the politics of online communities. Oxford, Inter-Disciplinary Press: 227-35.

Smiraglia, R. P., Ed. (2006). Metadata: A cataloger's primer. New York, Haworth Press.

Sompel, H. V. d. and C. Lagoze (2000). "The Santa Fe Convention of the Open Archives Initiative." D-Lib **6**(2). http://www.dlib.org/dlib/february00/vandesompel-oai/02vandesompel-oai.html

Sorensen, C. (2010). "Cultivating interaction ubiquity at work." The Information Society **26**(4): 276-87.

Spencer Brown, G. (1972). Laws of form. New York, Julian Press.

Stein, E. W. (1995). "Organization memory: Review of concepts and recommendations for management." International Journal of Information Management **15**(1): 17-32. http://www.sciencedirect.com/science/article/B6VB4-3Y5FM94-11/2/c2d8d63e1836161d8bfd07a60d3f8325

Stichweh, R. (2000). "Systems theory as an alternative to action theory? The rise of "communication" as a theoretical concept." Acta Sociologica **43**(5): 5-13.

Strout, R. F. (1956). "The development of the catalog and cataloging codes." The Library Quarterly **26**(4): 254-75.

Stuurman, S. and M. Grever (2007). "Introduction: Old canons and new histories", in: M. Grever and S. Stuurman (Eds.), Beyond the canon: History for the twenty-first century. Basingstoke, New York, Palgrave Macmillan: 1-16.

Sugimoto, S., J. Adachi, T. Baker and S. L. Weibel (2002). Report from International Conference on Dublin Core and Metadata Applications 2001. IFLA General Conference, Glasgow, Scotland.

Suleman, H. and E. A. Fox (2001). "A framework for building open digital libraries." D-Lib **7**(12). http://www.dlib.org/dlib/december01/suleman/12suleman.html

Suton, J. (2009). "Looking beyond memory studies: Comparisons and integrations." Memory Studies **2**(3): 299-302.

Svenonius, E. (2000). The intellectual foundation of information organization. Cambridge, MA, MIT Press.

Tanner, S., T. Munoz and P. H. Ros (2009). "Measuring mass text digitization quality and usefulness. Lessons learned from assessing OCR accuracy of the British Library's 19th Century Online Newspaper Archive." D-Lib **15**(7/8). http://www.dlib.org/dlib/july09/munoz/07munoz.html

Teddlie, C. and A. Tashakkori (2003). "Major issues and controversies in the use of mixed methods in the social and behavioral sciences", in: A. Tashakkori and C. Teddlie (Eds.), Handbook of mixed methods in social and behavioral research. Thousand Oaks, Sage: 3-50.

Thompson, J. (1982). The end of libraries. London, C. Bingley.

Thong, J. Y. L., W. Hong and K.-Y. Tam (2002). "Understanding user acceptance of digital libraries: What are the roles of interface characteristics, organizational context, and individual differences?" International Journal of Human-Computer Studies **57**(3): 215-42.

Tilson, D., K. Lyytinen and C. Sorensen (2010). "Digital infrastructures: The missing IS research agenda." Information Systems Research **21**(4): 748-59.

Titscher, S., M. Meyer, R. Wodak and E. Vetter (2000). Methods of text and discourse analysis. London, Sage.

Truran, M., J.-F. Schmakeit and H. Ashman (2011). "The effect of user intent on the stability of search engine results." Journal of the American Society for Information Science and Technology **62**(7): 1257-75.

Tuominen, K., S. Talja and R. Savolainen (2003). "Multiperspective digital libraries: The implications of constructionism for the development of digital libraries." Journal of the American Society for Information Science and Technology **54**(6): 561-69.

Urquhart, C., H. Lehmann and M. D. Myers (2009). "Putting the 'theory' back into grounded theory: Guidelines for grounded theory studies in information systems." Information Systems Journal.

Urrestarazu, H. (2011). "Autopoietic systems: A generalized explanatory approach - Part 1." Constructivist Foundations **6**(3): 307-24.

Vanderstraeten, R. (2002). "Parsons, Luhmann and the theorem of double contingency." Journal of Classical Sociology **2**(1): 77-92.

Varela, F. J. (1981). "Describing the logic of the living: The adequacy and limitations of the idea of autopoiesis", in: M. Zeleny (Ed.), Autopoiesis: A theory of living organization. New York, North Holland: 51-62.

Varian, H. R. (2010). "Computer mediated transactions." The American Economic Review **100**(2): 1-20.

Wagner Cook, S., T. K. Yip and S. Goldin-Meadow (2010). "Gesturing makes memories that last." Memory and Language **63**(4): 465-75.

Wagner-Pacifici, R. (1996). "Memories in the making: The shapes of things that went." Qualitative Sociology **19**(3): 301-21.

Wagner-Pacifici, R. and B. Schwartz (1991). "The Vietnam Veterans Memorial: Commemorating a difficult past." The American Journal of Sociology **97**(2): 376-420.

Waibel, G. and R. Erway (2009). "Think globally, act locally: Library, archive, and museum collaboration." Museum Management and Curatorship **24**(4): 323-35.

Waibel, G., R. LeVan and B. Washburn (2010). "Museum Data Exchange: Learning how to share." D-Lib **16**(3/4). http://www.dlib.org/dlib/march10/waibel/03waibel.html

Waller, V. (2009). "The relationship between public libraries and Google: Too much information." First Monday **14**(9). http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2477/2279

Walsham, G. (1995). "Interpretive case studies in IS research: Nature and method." European Journal of Information Systems **4**: 71-81.

Walsham, G. (2006). "Doing interpretive research." European Journal of Information Systems **15**(3): 320-30.

Watzlawick, P. (1977). How real is real? Confusion, disinformation, communication. New York, Vintage Books.

Weber, M. (1976). Wirtschaft und Gesellschaft. Grundriss der verstehenden Soziologie. Tübingen, Mohr.

Webster, F. (1995). Theories of the information society. London, Routledge.

Webster, F. (2005). "Making sense of the information age." Information, Communication and Society **8**(4): 439-58.

Webster, F. and E. Puoskari, Eds. (2003). The information society reader. London, Routledge.

Weick, K. E. (1989). "Theory construction as disciplined imagination." Academy of Management Review **14**(4): 516-31.

Weinberger, D. (2007). Everything is miscellaneous: The power of the new digital disorder. New York, Times Books.

Wetzel, D. J. (2011). "Maurice Halbwachs - Vergessen und kollektives Gedächtnis", in: P. Wehling (Ed.), Soziologie des Vergessens. Theoretische Zugänge und empirische Forschungsfelder. Konstanz, UVK: 37-55.

Wiener, N. (1948). Cybernetics: Or control and communication in the animal and the machine. Cambridge, MA, MIT Press.

Wiggins, R. (1996). "The mysterious disappearance of the White House Speech Archive." First Monday **1**(2). http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/472/828

Wildavsky, A. (1983). "Information as an organizational problem." Journal of Management Studies **20**(1): 29-40.

Wilensky, R. (2000). "Digital library resources as a basis for collaborative work." Journal of the American Society for Information Science **51**(3): 228-45.

Winner, L. (1996). "Who will we be in cyberspace?" The Information Society **12**(1): 63-72.

Woldering, B. (2004). "The European Library: Integrated access to the national libraries of Europe." Ariadne **38**. http://www.ariadne.ac.uk/issue38/woldering/intro.html

Yates, F. A. (2001). The art of memory. Chicago, University of Chicago Press.

Yin, R. K. (2003). Case study research: Design and methods. Thousand Oaks, Sage Publications.

Yoo, Y. (2010). "Computing in everyday life: A call for research on experiential computing." MIS Quarterly **34**(2): 213-31.

Yoo, Y., O. Henfridsson and K. Lyytinen (2010). "The new organizing logic of digital innovation: An agenda for information systems research." Information Systems Research **21**(4): 724-35.

Young, A. P. (1996). "Libraries and digital communication: Collision or convergence?" Journal of Academic Librarianship **22**(1): 11-13.

Zeilinger, A. and F. Griese (2006). Einsteins Spuk: Teleportation und weitere Mysterien der Quantenphysik. München, Bertelsmann.

Zittrain, J. (2008). The future of the internet and how to stop it. New Haven, Yale University Press.

Zuboff, S. (1988). In the age of the smart machine: The future of work and power. Oxford, Heinemann Professional.