Social memory and the digital domain:
The canonization of digital cultural artefacts


Sub-theme 35: Deconstructing Institutions: Meaning, Technology and Materiality

Attila Marton
London School of Economics and Political Science
a.marton@lse.ac.uk

Abstract

Memory institutions and their practices of canonization are closely tied to the emergence of a social memory based on material artefacts and communication media. As libraries, archives and museums are stepping into the online world of digitality and computational operations, the question arises as to how these institutions and their processes of canonization change. Based on Elena Esposito’s system theoretical concept of memory as an operation of forgetting and on David Weinberger’s three orders of ordering artefacts, the paper analyses the canonization of digital cultural artefacts according to the practices of selection, order and preservation. Against this backdrop, the theme of transversal forgetting is developed as a cyclical process of forgetting-as-data that cuts across the boundaries of libraries, archives and museums and their traditionally separated rationale of what to select and how to order and to preserve the selected. The concept, therefore, is an argument against the still dominating metaphor of social memory being an externalized storage. Thus conceived, transversal forgetting attempts to capture memory institutions as part of the wider information environment of bits and bytes, networks and algorithms.

keywords: social memory, memory institutions, canonization, digital cultural artefacts
Introduction

For centuries, libraries, archives and museums have been at the core of the emergence of a social memory. As memory institutions, they are closely related to the ways society remembers and, more importantly, forgets. Their practices of selecting, ordering and preserving artefacts leads to the formation of a canon – a collection of authoritative artefacts that receive the status of being representative for a scientific discipline, faith, nation, artistic movement, historical period and so forth. These practices of canonization elevate a vast variety of items – be it books, paintings, T-shirts or kitchen utensils - from being cultural artefacts to cultural heritage artefacts deemed worthy of being saved from the ravages of time.

Memory institutions and their practices are closely tied to the peculiarities of physical artefacts and the communication media they are entrusted with (Lanzara 2009). The modern day library, for instance, emerged as an autonomous organizational form in the wake of the rise of print mass media – most notably books. A library takes care of copies of a mass product. In contrast, museums and archives mostly conserve unique or rare items – be it for historical research, administrative documentation or other reasons. As a consequence, memory institutions developed very specific and differentiated rational of selecting, ordering and preserving; that is, museums and archives are focusing more on the conservation and provenance whilst libraries are focusing more on the accessibility and usability of their respective cultural heritage artefacts.

With the rise of the internet and the services it brings forth, a rapidly increasing degree of our cultural heritage is either being migrated or already born into digitality resulting in a new breed of cultural artefact that is ephemeral, networked and dependent on computational operations (Kallinikos, Aaltonen et al. 2010). Woven into a larger information environment, their findability is predominantly provided by algorithmic operations of search engines introducing a new rationale for organizing documents that defies the ex-ante set of fixed categories developed by librarians, archivists and curators (Esposito 2002; Weinberger 2007). It is this digital domain of bits and bytes, networks and algorithmic ordering memory institutions are stepping into. In light of these observations, the question arises, how memory institutions and
their practices of canonization are changing in order for them to remain memory institutions granted the role of caretakers of digital cultural heritage artefacts.

The paper will begin with a brief introduction into social scientific conceptualizations of memory, remembering and forgetting. The second section will explain the institutional role libraries, archives and museums play as memory institutions based on their practices of canonization. Finally, the paper will move on to a more detailed analysis of the practices of selecting, ordering and preserving cultural heritage and how these change in light of contemporary developments of digitality and internet related services.

Social memory
The study of memory from a social scientific perspective can be traced back to the turn of the 19th to the 20th century. A quite exotic concept at that time, memory was not at the centre of attention and only mentioned in passing by leading sociologists, such as Marx, Durkheim, Tönnies and Simmel, or social anthropologists, such as Evans-Prichard (Olick and Robbins 1998). 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.

It was not until the 1980s, however, that memory took a prominent role in academic, political and public life (Assmann 2002:400). Some authors even lament about 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 which present quite a fragmented field of study. 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).
Collective memory

The first seminal contribution to the field was made by Maurice Halbwachs – a student of Durkheim - during the early decades of the 20th century. In his analysis, he develops the concept of a collective memory shared by the members of a group. Remembering is to be seen as a collective practice of reconstructing the past based on the present social framework (cadres sociaux) the group finds itself in (Halbwachs and Coser 1992). The collective memory is located in or, if you will, dispersed among the personal memories of individuals. Although it is the individual member of a group that remembers, remembering can only be practiced in concert with other members of the same group. Hence, memory has a social structure based on individuals and their interactions (Connerton 1990; Olick 1999; Kansteiner 2002).

Social frameworks play a key role as they are the means through which a group accomplishes an image of the past compatible with the dominant views of the very same group. For instance language, the most fundamental framework, does not only enable us to talk about memories but enables us to remember in the first place framed by what can and, more importantly, cannot be expressed with the words that are at our 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 and Coser 1992).

Individual and collective memory are interwoven and even though it is the individual that remembers her/his memory is a social phenomenon. Hence, Halbwachs rejects the separation between the individual and the collective as the loci for memory (Connerton 1990:36). In the Durkheimian tradition, this is to be seen as a direct counterproposal to psychological and psychoanalytical concepts of memory (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 and Coser 1992:38).

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).

Cultural memory
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, 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 conceptual framework is derived from 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 bring some clarity into the field of memory studies, Jan and Aleida Assmann introduced the distinction between communicative and cultural memory. Close to Halbwach’s concept, communicative memory lives in every interaction and is passed on through oral traditions. In contrast, cultural memory can exist in symbolic forms as reminders in need of institutional caretakers for preservation (Assmann 1999; Assmann 2008b). Separated from everyday life, cultural memory is 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). With respect to museums, archives and libraries, the institutional practice is observed as a process of canonization through which cultural artefacts become part of a continuously invoked, thus, remembered cultural heritage (Assmann 2010).
Social memory

Based on 2nd order cybernetics (Foerster 2003) and social system theory (Luhmann 1996; 1997), Esposito (2002) develops an alternative approach to the study of social memory. Memory, in general, is seen as an operation of forgetting by means of constructing abstract categories. In this sense, memory filters singular details. Remembering, on the other hand, is a process of reconstruction following the parameters according to which events have been categorized in the first place. Filtering the singularity of events prevents the system from blocking itself and, therefore, enables a system to remember in the first place (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). This is accomplished by the construction of sameness into difference and, as a result, recursivity and identity. Memory, therefore, is defined as a 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). Conceptualized in this fashion, memory is to be seen as a process or an operation rather than a collection of stored memories taking forms such as neural engrams, mental imagos, cultural artefacts or computational data (Riegler 2005).

In the social domain, 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 make 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, socially constructed classification systems strip that which is to be classified from its individual features according to the parameters 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. This is, however, a communicative accomplishment of social systems, not of a collective of mental memories (Esposito 2008).
Due to developments in communication technologies from writing and printing to relatively recent innovations in tele-media and ICT, social interaction has been enabled to be mediated in increasingly decontextualized ways leading to an increasingly abstract organization of social memory (Esposito 2008). The paradigmatic example is the card catalogue in the library. It does not remember books and their content but rather classifications and instructions on how to find items in the library. Ultimately, social memory turns out to be a self-referential phenomenon; the more elaborate the classification system, the more details can be filtered (Bowker 2005:17-21). However, the more elaborate the classification system, the more can be reconstructed and, thus, remembered as well. The more we forget, the more we remember. The more we remember, the more we forget. Against this backdrop, Esposito (2002) stresses the historical significance of print media that brought about a decline of orality and the repetition of what is known in favour of what is novel and unknown. Hence, remembering, central to Halbwachs’ collective memory, is not the primary function of social memory anymore but rather forgetting (Dimbath and Wehling 2011); an evocative claim that will be further elaborated by taking memory institutions as an example.

**Institutions of forgetting**

The concept of memory basically addresses the relationship between remembering and forgetting. It was Aleida and Jan Assmann’s 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. Esposito 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.

A telling example is the shift from an open to a closed shelf arrangement in libraries that came about as a reaction to the immense increase in the number of printed books roughly from the beginning of the 20th century on (Thompson 1982). The key innovation, in this respect, is the card catalogue that creates an information space representative of the books in the library repository. While the arrangement of the books themselves provide for findability in an open shelf system, it is of no concern in
the closed shelf arrangement since findability is provided by the card catalogue (Weinberger 2007). As a result, a library can forget the books themselves and only needs to remember the catalogue (Esposito 2002). More precisely, a library only remembers the rules of cataloguing and can store the books according to size in order to save precious shelf-space. If the catalogue is lost, however, the collection of books dissolves into mere noise without any inherent structure that would allow a librarian to find anything. The library becomes the catalogue or, in other words, a library remembers more by forgetting more.

With the increasing involvement of online communication and digitality into basically ever aspect of social interaction, social memory has stepped onto an even more abstract level. Findability is provided by search engines relying on algorithmic calculations to bring a momentary order into the online information environment (Kallinikos, Aaltonen et al. 2010). Their logic is not based on a catalogue and its fixed set of ex-ante classification parameters but rather on ex-post ordering. As a consequence, the relationship between remembering and forgetting is structured in a way that requires only for the algorithm to remain stable. In other words, with search engines even the catalogue and its rules are forgotten and only the algorithm is remembered (Sluis 2010). Society is remembering even more by forgetting even more (Esposito 2002).

Weinberger (2007) outlined this development as a shift from the first-order of order, i.e. the order of the things themselves found in an open-shelf arrangement in a library, to a second-order of order, i.e. the order of representations of things such as the card catalogue. Ultimately, a third-order of order is rising based on ex-post performative ordering. The search engine is a paradigmatic example as it constructs the results page as a momentary rendition for each and every user as well as each and every search query only to be lost the moment the results page is closed (Kallinikos, Aaltonen et al. 2010). In light of this argument, the structure of social memory is shifting from the paradigm of information retrieval to a performative paradigm of information construction (Esposito 2002). By now, search engines are the primary service for ordering and finding content in the online world. However, this world is, above all, constantly updated and, thus, ephemeral – its infrastructure lacking any inherent capabilities for persistence and heritage.
Memory institutions, on the other hand, provide for a relatively stable selection of
cultural artefacts which are deemed representative or even authoritative through a
process referred to as canonization. Originated in the practices of the Catholic Church,
the term “canon” has also found a place in literary criticism, arts and history referring
to 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). In more abstract terms, 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, scientific discipline and so forth. The key
aspect of a canon is the continuous sustainment of its relevance through education,
storytelling, broadcasting, exhibiting and so forth. A canon is the active aspect of
cultural memory.

In contrast, Aleida Assmann (2008a) also proposes a passive aspect of cultural
memory – the archive. A storage compartment of historical documents, the archive
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.¹ Thus conceived, memory institutions are a combination of the two aspects
of canon and archive. A circulating library, for instance, offers access to books
relevant within the context of a canon; a museum may exhibit only some artefacts
whilst keeping others secure in its archive. 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.

In this respect, the WWW is not a library nor are search engines librarians (Kuny and
Cleveland 1996). The domain of cultural heritage remains, so far, the sole domain of

¹In this sense, Assmann (2010) refers to a historical archive rather than a political archive which is
very much in use in terms of bureaucratic control and administration.
memory institutions. For instance, it is a library which 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 artefact 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). Hence, the question arises how memory institutions are changing as they are stepping into the online world of momentary and performative ordering.

**Canonization of digital cultural heritage**

Aleida Assmann’s (2010) definition of canonization as a selection of artefacts which are, as a consequence, deemed valuable and preserved is an important starting point but underestimates the importance of forgetting. Following Esposito’s (2002) argument, canonization, however, can be described as a process 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, replacing the aspect of value, as proposed by Assmann, since the observation that cultural heritage artefacts are deemed valuable is quite redundant. In contrast, 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 take an appropriate role in a conceptual framework (Strout 1956; Craig 1992; Svenonius 2000; Alexander and Alexander 2008).

To conclude, canonization entails a first step which is obviously the selection of items according to explicit rules of what comes in and what stays out. The second step entails the ways the selected items are ordered, already discussed above with respect to the introduction of the card catalogue and search engines. Third, the selected and ordered items need to be preserved for future generations – a crucial point that sets
memory institutions apart. In what follows, the paper will analyse each of these three steps and discuss how they are changing in the online world.

**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. In the librarian domain, selection has been codified in the Functional Requirements for Bibliographic Records (FRBR) that distinguishes between work, expression, manifestation and item (IFLA 1998). For instance, Shakespeare’s Macbeth would be a work. However, the work can be expressed in many ways; it can take a written form, it can be performed on stage or made into a film. These expressions manifest themselves in, for instance, various editions of books, recordings of theatre performances or feature films depicting the play according to the vision of the respective director. The item, at last, is the actual physical object (paperback or hardcover print, DVDs or BlueRay discs) that ends up on a shelf of a library. Crucially, the work is defined as an abstract concept – a category filled with fitting instances - items. A library selects and, as a consequence, orders both (Svenonius 2000).

In contrast, the distinction between work and item is unnecessary in museums and most archives since the selected items are rare, if not unique. The first folio print of Macbeth, therefore, belongs into a book museum rather than a library. By the same token, one of John Lennon’s T-Shirts becomes a unique cultural heritage artefact because of its context in pop culture. An exception to these selection criteria are, obviously, objects used as examples such as a rock of salt exhibited in a natural history museum. In this case, the rock would be an instance for the category “rock salt” – similar to the work/item distinction above.

In the digital domain, these distinctions do not make any sense anymore. While CD-ROMs, for instance, may still be seen as mass media due to being produced as packages, the digital files themselves defy the logic of rarity, copy and original (Benkler 2006; Ekbia 2009; Faulkner and Runde 2010; Kallinikos, Aaltonen et al.)
A telling example is a digitized cultural heritage artefact; that is an already canonized artefact digitized for the library, archive or museum that holds it.

Figure 1: A Digitized cultural heritage artefact using Shakespeare’s Macbeth as an example.

As the example shown in Figure 1 illustrates, the work Macbeth – the abstract category – still exists but finds its textual manifestation in digitized artefacts. Being 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 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 of copy, rarity or uniqueness. Instead, they gain a new status as a digital cultural artefact described by so-called structural metadata - instructions for complementary software applications on how the artefact is to be emulated as a digitized book, painting and so forth (Arms 2001; Marton 2010a). In this sense, items have two modes of “existence”: codified instructions and emulations. For instance, a mass printed contemporary edition of Macbeth can be emulated in different ways – be it a high quality image scan version for a computer screen or a low quality version for the mobile phone - although it remains the same manifestation.
This framework cuts across the traditionally separated domains of libraries, archives and museums – an attribute of what I refer to as transversal. By the same token, it is quite misleading to call institutions selecting digital cultural heritage artefacts digital libraries, digital archives and digital museums respectively. They are transversal memory institutions because their artefacts are binary-based and, as data, cut across all types and genres of communication media and types of artefacts alike (Tilson, Lyytinen et al. 2010; Yoo, Henfridsson et al. 2010). They remain transversal because those types and genres still exist as emulations. That is, memory organizations still take care of digitized books, paintings and so forth. They all share the same transversal foundation of bits and bytes. However, they are also maintained and emulated as recognizable cultural artefacts within the context of a canon.

![Figure 2: Transversality of digital cultural artefacts.](image)

The cultural heritage artefacts selected for digitization, however, are already part of a canon. Hence, selection for digitization cannot be seen as part of a canonization process as long as libraries, archives and museums continue to digitize only what they already hold in their repositories. The situation is completely different when it comes to born-digital artefacts and, more importantly, online content. 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 selection 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 (Kallinikos 2006). So far, the task has been addressed mostly on a national level under the term web-archiving. 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 Austria² - 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 2010.³

The pioneer in web-archiving is, however, a non-governmental and non-for-profit organization – the Internet Archive (www.archive.org). Since 1996, it has continuously archived the WWW amounting to more than 150 billion web pages today. The selection process is based on the for-profit search engine Alexa that donates its web harvest to the Internet Archive once the commercial value of the data has expired. The Alexa crawler, however, prioritizes web sites according to their popularity within the Alexa user community. Thus, the Internet Archive does not select according to professional criteria and procedures but popularity scores and search engine algorithms (Lyman and Kahle 1998). As a result, some web sites are being archived regularly and in depth, whilst others are not archived at all. In other words, the selection of web pages meant to represent the entire WWW follows the very same rationale that makes the WWW navigable – the rationale of search engines (Kallinikos, Aaltonen et al. 2010). Ironically, search engine algorithms, developed to provide for relevance by ranking everything ex-post, are used to provide for relevance by selecting something ex-ante.

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

² http://www.onb.ac.at/about/webarchivierung.htm. Last access 01 June 2011.
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 with respect to professional but also lay users. 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 collected items. Representing the librarian domain, the International Federation of Library Associations and Institutions (IFLA 1998) defines the goals of ordering as enabling a user to 1) find, 2) identify, 3) select and 4) acquire entities matching the user’s search criteria (Svenonius 2000). Traditionally, these goals were achieved through associations by creator and subject classification. The introduction of data-based catalogues did not fundamentally change this approach. They 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 card catalogue it replaced. Even the early search engines approached the WWW in a similar fashion as they manually categorized web sites to offer online services very much like subject catalogues (Introna and Nissenbaum 2000).

With the digitization of cultural heritage artefacts, however, memory institutions are in a process of redefining the way items are being ordered and, therefore, made accessible. A revelatory example is the Europeana project. Funded by the European Union, the initiative coordinates the digitization efforts of European libraries, archives and museums and makes them accessible via the portal www.europeana.eu. One of the main contributions of Europeana is the creation of digital surrogate objects for each of the digitized artefact provided by its members. The surrogates, basically, consists of descriptive metadata about and a link to the digital artefact it represents. In turn, the surrogates serve as the platform for further services and functionalities. For the moment, the Europeana repository can only be searched and browsed through an interface found on the portal. The goal, however, is not to arrange the surrogates only
by means provided by Europeana but rather to create potentialities for others to arrange the surrogates as they fit.

The catalogue is being disassembled and the surrogates are made part of the WWW as web resources allowing for a variety of ways to navigate through Europeana’s repository. For one, the surrogate objects will be made crawlable for search engine harvesting. Hence, the discovery of artefacts 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 enabling a user to browse through the collection based on semantic concepts.⁴

![Semantic Network](image)

**Figure 3: The two layers of the Europeana Information Space (Europeana Documentation).**

The two layers depicted in Figure 3 form what is referred to as the Europeana Information space – a metaphorical landscape of surrogate objects made navigable by the second layer – the metaphorical map of the surrogate layer. The actual digitized artefacts hosted by the content providers form a third layer that is represented by the surrogate layer one-to-one. The semantic layer, however, does not function as a fixed set of classifications used to categorize the digitized artefacts. It is used to contextualize the surrogate by linking it to other web resources. Notably, the

---

⁴ The prototype of the semantic search functionality can be found at [http://eculture.cs.vu.nl/europeana/session/search](http://eculture.cs.vu.nl/europeana/session/search), last access 28th Feb. 2011.
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) - an initiative that connects databases across the internet such as 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.

The semantic network layer is not exclusively created and maintained by Europeana but rather in cooperation with other (meta)data providers. For instance, the surrogate can be linked to a rights holder whose identity is stored in the Friends Of A Friends (foaf) database – an open service allowing to connect social network sites.\(^5\) Another possibility is the Virtual International Authority File (viaf) service which brings together names of persons and institutions provided by trusted sources around the world.\(^6\) It allows a user to search for names spelled in different ways and in different languages. By the same token, the surrogate could be further enriched by linking it to Wikipedia’s open database dbpedia for annotations, abstracts and other information related to an artist, a cultural heritage artefact, the time period it was created in and so forth. Of course, these data are linked to other data, ultimately, resulting in a context without boundaries.

As the third option, Europeana’s surrogates are open to be accessed by others and used as they see fit. This is made possible by implementing an Application Programming Interface (API) which allows for access to the data itself. This signifies a paradigmatic change in the way a memory institution defines its basic role since the portal www.europeana.eu will not be the main incarnation of Europeana but rather the API. As a consequence, the portal or rather the services it offers will be only one possibility of associating cultural heritage artefacts among a vast variety of possibilities ranging from social network sites, online educational services and tourism portals to mash-ups, widgets and apps. The experimental “geo-wrapper” app is one of the very first examples for these possibilities. Developed by an independent programmer, it is a mash-up of openstreetmap.org – an open source online map

\(^5\) For further information visit http://www.foaf-project.org/, last access: 28\(^{th}\) Feb. 2011.
\(^6\) For further information visit http://viaf.org/, last access: 28\(^{th}\) Feb. 2011.
service – and Europeana data. Based on the user’s input in terms of geographical area and time period, it shows the location where cultural artefacts were found and excavated, thus, constructing an association among them.

In this respect, the ordering of the surrogates can be described as transversal as well as it cuts across the former silos of knowledge and institutional boundaries. In other words, the artefact is neither described nor ordered by Europeana alone. There is no universal rationale of ex-ante order or ex-post ordering. The descriptions are only contextualized in a networked fashion based on a shared layer of linked data open to be ordered and combined with other data as one sees fit.

In contrast to Weinberger’s (2007), the association between artefacts is not bound to a primacy algorithmic relevance ranking or social tagging. An association can be based on a traditional, ex-ante selected and fixed virtual exhibition just as much. 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 partnering 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 the very same way as in the actual exhibition. Be it the

---

Figure 4: Comparison of the portal and the API as Europeana’s primary incarnation (Europeana Documentation).

---

7 A working demo can be found at http://amercader.net/dev/geoeuropeana, last access: 16th May 2011.
curated virtual exhibition or the search engine results page, the ordering relies on binary-based data that serves as a transversal layer on a fundamental level. Going back to Weinberger’s (2007) terminology, the order of cultural artefacts is transversal as well since first-, second- and third-order of order can be used for constructing associations and discovery.

![Figure 5: Transversal order.](image)

**Preservation**

The preservation of cultural heritage artefacts is a key aspect of memory institutions and canonization (Assmann 2008a). 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, conservation or preservation does not necessarily mean the prevention of forgetting but the prevention of disintegration and loss. The card catalogue, for instance, forgets the items which are, nevertheless, conserved or preserved. By the same token, the catalogue needs to be preserved as well in order to provide for persistent findability and accessibility.

The preservation of digitized as well as born-digital artefacts and their persistent accessibility are probably the most challenging tasks for memory institutions (Russell, Weinberger et al. 1999; Bennett 2001). The reasons lie in the makeup of digital artefacts consisting of bits, standards and their correct interpretation by the
corresponding soft- and hardware (Faulkner and Runde 2010; Kallinikos, Aaltonen et al. 2010). Some digital artefacts, such as hypertext, may not even have boundaries or may be constantly updated (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).

In light of these observations, Russel, Weinberger and Stone (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.

Digital artefacts are, in fact, computational operations based on data and their correct interpretation (Faulkner and Runde 2010; Kallinikos, Aaltonen et al. 2010). In 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 (Marton 2010b). A telling example is, again, the Internet Archive. 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 by the Alexa crawler, 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. In its original meaning, preservation refers to the content of, in the end, written or printed documents. 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.

To conclude, 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. Considering the fact that data as well as metadata cut across institutional boundaries, communication genres and artefact types, so must also their preservation. Taking Europeana’s networked contextualization of its surrogates as an example, preservation can be conceptualized as transversal as well. While Europeana’s contributing member libraries, archives and museums preserve the actual digitized cultural heritage artefacts in their repositories, Europeana can only preserve the surrogates it created including the links to external data providers but not the networked contextualization they provide. 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. Since data and meta-data are transversal, so must be their preservation.

Transversal Forgetting

The increased involvement of born-digital artefacts in social and cultural life 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 contextualization they are a part of. On the level of the items, the cultural heritage artefacts are forgotten as data. What is remembered are only object models – blueprints of instructions encoded as metadata on how to assemble recognizable cultural heritage artefacts. After all, a librarian 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.

A similar argument can be made in terms of the order of the cultural heritage artefacts. Performative ordering by search engines, for instance, is only a momentary rendition of an order only to be forgotten the moment it 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. By the same token, as Europeana disassembles the catalogue into surrogate objects, these are stored as meta-data, as data about data, as well. This is the second level on which the relationship between remembering and forgetting is unfolded. It is not only the digitized cultural artefact that is forgotten as data but also its surrogate. Europeana, therefore needs to only remember the surrogate model which, equivalent to the object model, is a template for the assemblage of the actual surrogate object. It is this schematic model that provides the parameters according to which surrogates are being constructed out of data and, therefore, remembered.

Within the domain of memory institutions, the third level of contextualization complements the level of the artefact and its surrogate object. The context, however, is not under the total control of a single organization but rather integrated into the wider information environment of linked data providers. As exemplified by Europeana, the association between artefacts will be provided by a network of data sources and information service providers replacing the logic of the catalogue with networked contextualization.
Europeana is a revealing example as it actively pursues the abolishment of cataloguing – a notion already accomplished by algorithmic search engines. The primary focus is on providing potentialities for re-use and mash-ups. 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.

Just as the cultural heritage artefact is forgotten as data so are the fixed rules of cataloguing. 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 is 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.

<table>
<thead>
<tr>
<th>Unit of Analysis</th>
<th>Example</th>
<th>Forgotten as...</th>
<th>What is remembered?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital Cultural Artefact</td>
<td>- Digitized Book</td>
<td>Data</td>
<td>Object Models</td>
</tr>
<tr>
<td></td>
<td>- Archived Web Page</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digital Surrogate Object</td>
<td>Europeana</td>
<td>Data</td>
<td>Surrogate Object Models</td>
</tr>
<tr>
<td>Networked Contextualization</td>
<td>Linked Open Data</td>
<td>Data</td>
<td>Qualified Links</td>
</tr>
</tbody>
</table>

Table 1: Three levels of forgetting and remembering.

The three levels, summarized in Table 1, 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 differentiates them is what is remembered on each level. In this sense, data cuts 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, mesh-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 bound by the “silos of knowledge” and delimited by the respective standards of the domains of libraries, archives and museums.

![Diagram](image)

**Figure 6: Transversal forgetting and remembering.**

As Figure 6 illustrates, the relationship between forgetting and remembering is caught up in a cycle. All three levels of memory institutions analysed in this paper, ultimately, end up being forgotten as data. Be it the digital cultural heritage artefact, the digital surrogate object or the networked contextualization, the fundamental layer of data cuts across all of them turning canonization into a transversal process of forgetting. Selection is transversal due to the data-based emulation of the selected. Order is transversal because first-, second- and third-order of order can be used alike. Finally, preservation is transversal since it is not enough to preserve just the data but the cycle depicted in its entirety; i.e. the long-term storage of data, the preservation of the cultural heritage artefact and its surrogate as well as the sustainment of persistent findability and accessibility.

**Conclusion**

The analysis of the transformation of memory institutions, as they step into the online world of digitality and computational processes, reveals a necessity to reconsider the conceptualization of social memory and the distinctions applied by its field of study. For one, concepts of social memory are still dominated by an image of externalized, recorded knowledge stored in libraries or on computer hard disks (Riegler 2005). The
distinction between canon and archive, as the active and passive aspects of cultural memory, reflects this image (Assmann 2008a). Applied to the internet and related online services, the distinction is, however, of little help since it is based on the assumption that if books are cultural heritage artefacts, so must be bits and bytes; if books are canonized or archived, so must be bits and bytes. Hence, Aleida Assmann (2006:18) concludes that “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.” A conclusion privacy advocates, for instance, would strongly disagree with (Mayer-Schönberger 2009). Storing information or rather data conveyed via the internet is, actually, hardly a problem (Kallinikos 2006).

The storage metaphor and, therefore, the distinction between canon and archive is abandoned in favour of a cyclical operation of transversal forgetting based on the notion that information is forgotten as data. Hence, the paper developed an alternative version of canonization as a process of forgetting consisting of the selection, order and preservation of digital cultural artefacts. Based on a layer of binary data, transversal forgetting cuts across all traditional boundaries and “silos of knowledge” which used to be separated knowledge universes with their own sets of standards, rules, procedures, types of artefacts and so forth. In other words, the canonization of digital cultural artefacts is described as transversal, since cultural heritage is forgotten as binary data irrespective of what is selected and how it is ordered and preserved.

References


