A Theory of Digital Objects

by

Jannis Kallinikos, J.Kallinikos@lse.ac.uk
Aleksi Aaltonen, V.A.Aaltonen@lse.ac.uk
Attila Marton, A.Marton@lse.ac.uk

Abstract

Digital objects are marked by a limited set of variable yet generic attributes such as editability, interactivity, openness and distributedness. As digital objects diffuse throughout the institutional fabric, these attributes and the information-based operations and procedures out of which they are sustained install themselves at the heart of social practice. The entities and processes that constitute the stuff of social practice are thereby rendered increasingly unstable and transfigurable, producing a context of experience in which the certainties of recurring and recognizable objects are on the wane. These claims are supported with reference to 1) the elusive identity of digital documents and the problems of authentication/preservation of records such an identity posits and 2) the operations of search engines and the effects digital search has on the content of the documents it retrieves.

> Each part of the house occurs many times; any particular place is
> another place…The house is as big as the world—or rather, it *is*
> the world.
> Jorge Luis Borges, The House of Asterion[1]

Preamble

In this article, we seek to develop a theory of digital artifacts. The venture assumes that digital technologies of all varieties and breeds share a limited set of qualities that places them apart from other non-digital devices and systems (paper-based) for managing information. This is, no doubt, a contentious claim that many readers may find running counter to a widespread view that portrays the use of artifacts in general and information-based artifacts in particular as highly contingent on the skills and predispositions of human agents. On the other hand, we all know that technologies do count. They enable some things and preclude others while their diffusion is over time associated with the formation of skills, habits and behavioral predispositions. It would thus seem reasonable to expect that the systematic involvement of digital artifacts and technologies in social practice has implications that cannot be traced exclusively to the specific characteristics of the contexts (agents and practices) within which they are encountered. Despite the amazing vari-

ety of studies of information technologies, we currently have either theories of information society of a very general nature (e.g. Castells, 2001) or rich accounts of specific technologies in particular contexts (e.g. Orlikowski, 2000). No middle-range, as it were, theory exists to allow focusing on particular practices or contexts without losing sight of those generic processes that recur across context as the result of the diffusion of information technologies. It seems to us important to seek to conceptualize technological artifacts and their entanglement with human affairs in ways that avoid the improper reification of technology without sacrificing its transformative potential (Jordan, 2009; Marton, 2009a; Pollock and Williams, 2009).

We hope to show in this article that such a project is feasible and that a theory of digital objects is timely and highly relevant. We subsume under the category of digital objects all digital technologies and devices and digital cultural artifacts such as music, video or image. The theory, we contend, provides a useful conceptual grid for studying social practices and identifying the peculiar generativity (Zittrain, 2008) and instability that digital objects introduce across a variety of settings and situations (Kallinikos, 2006). The theory claims that digital objects are marked by a limited set of variable yet generic attributes such as editability, interactivity, openness and distributedness that confer them a distinct functional profile.

We illustrate our claims with reference to 1) the identity of digital documents which the practice of archiving confronts and 2) the operations of search engines and the effects digital search has on the content of the items it retrieves. Digital documents are evasive artifacts that contrast with the solid and self-evident nature of paper-based documents. They occur in many versions that are constantly mutating. Most crucially, they are assembled into units by operations that are technologically driven and frequently far beyond the desktop by which users access or manipulate them. Accordingly, their evasive identity raises problems of authentication and preservation and impinges upon the inherited functions and practices of memory institutions like libraries and archives. Search engines exemplify a different problem. Instead of seeking to fix the peculiar instability of digital objects, search engines plough the constantly changing digital universe the Web is to address the user queries and contribute to the fluid and mutating nature of that universe in a variety of ways. Amongst these figures the imperative of Web findability that feeds back to Web content by exercising a pressure among website owners to (re)arrange website information in ways that make websites identifiable and indexable by search engines.

Both examples are indicative of wider developments that we associate with the functional profile of digital objects and their diffusion across the social fabric.

The paper is structured as follows. In the next section, we provide a detailed account of the distinctive attributes of digital objects and the constitutional texture (i.e. granularity and modularity) of digital technologies from which these attributes derive. We subsequently move to illustrating how these attributes carry over to social practice. We first consider the problem of identification of digital documents within the overall context of social memory which digital records serve. We then shift to examining the logic of search that search engines epitomize and the implications such logic has for the content of digital cultural records. In the final part of the paper we reflect upon our intellectual project and the strengths and limitations of the theory of digital objects we propose.

Digital Objects: Definitions and Attributes

Digital artifacts differ from physical objects and other cultural records (e.g. art objects, paper-based files) of non-digital constitution along a number of dimensions. Taken together, these differences confer digital objects a distinctive functional profile.

To begin with, digital objects are **editable**. In contrast to conventional artifacts, digital objects are pliable and always possible, at least in principle, to act upon and modify continuously and systematically. Editability assumes many forms. It can be achieved by just rearranging the elements by which a digital object is composed (such as items in a digital list or software library), by deleting existing or adding new elements or even by modifying some of the functions an individual element or a group of elements fulfill. In many other cases, editability is built-in the object in the form of regular and continuous updating of items or data fields, as it is the case with digital repositories of various kinds whose utility is closely associated with their steady updating (e.g. databases, transaction or booking systems, currency exchange systems). Indeed, the steady updatability of digital objects suggests that a large group of them have from their very beginning been conceptualized as organized receptacles of change capture (Kallinikos, 2006, 2009b). The editable nature of digital objects contrasts sharply not simply with physical artifacts but also with information contained in cultural records or artifacts of non-digital constitution. Once captured and laid down on a non-digital medium, information, as Borgmann (1999: 167) suggests, becomes "as viscous as molasses and as difficult to manipulate". Pliability and

editability are intrinsic to digital objects and represent crucial dimensions along which they can be distinguished from non-digital artifacts (Manovich, 2001; Weinberger, 2007).

Second, digital objects are **interactive** in the sense of offering alternative pathways along which human agents can activate functions embedded in the object or explore the arrangements of information items underlying it and the services it mediates. While ultimately tied to the pliable nature of digital artifacts, interactivity is here conceived as distinct from editability in that its enactment does not need to result in any change or modification of the digital object. Its key quality is contingent exploration made possible by the responsive and unbundled nature of the digital object rather than change. In this regard, interactivity enables actions of *contingent nature* (depending upon user choice), a condition that sets digital objects apart from the non-contingent, and arrested responses of physical artifacts and the inert nature of paper and other non-digital records or artifacts. To be sure, all artifacts entail some degree of malleability that allows one or another kind of adaptation to contingencies yet interactivity confers digital objects an entirely new spectrum of possibilities. Pre-programmed as it is, interactivity opens up the interior, as it were, of a digital object, unbundling the services it mediates and providing leeway for the exploration of alternative courses of action, as it is often the case with users navigating a website.

Third, digital objects are possible to access and to modify by means of other digital objects, as when picture-editing software is used to bring changes to digital images. It can also be accomplished in a more profound way, usually by experts or amateur hackers, through accessing the underlying principles or rules of the program that govern the behavior of the digital object or its source code (Jordan, 2009). Digital objects are thus **open** and reprogrammable in the sense of being accessible and modifiable by a program (a digital object) other than the one governing their own behavior (Manovich, 2001; Zittrain, 2008). Thus tied to change and modification, openness or reprogrammability is distinct from interactivity. It also differs from editability, insofar as the latter is confined to the simple reorganization, addition or deletion of the items that make up the digital object or the updating of information (databases) without "external" interference on the logical structure (i.e. the program) that governs the object and the generative mechanisms of information production and processing. Thus conceived, openness is closely tied to the *interoperable* character of digital objects and tends to construct a virtual object universe of a particular kind in which information sources and systems intersect and are brought to

bear upon another (Ciborra, 2007; Ekbia, 2009; Manovich, 2001). It is, of course, a widely diffused social practice to re-edit written information by means of other information. It is also in principle possible to expand, modify, repair or destroy a physical object by means of another or combine two or more physical objects to accomplish a specific task. However, the open character of digital objects and their pliability allow for a much deeper interpenetration of the items and operations by which they are constituted. The open and reprogrammable character of digital objects is, of course, variable and one important attribute of the contemporary digital landscape has been its steady progression towards a deeper interpenetration of codes, systems and artifacts and growing interconnectedness.[2]

Fourth, as the outcomes of interoperability and openness, digital objects are **distributed** and are thus seldom contained within a single source or institution (Haider and Sundin, 2010). In this sense, digital objects are no more than temporary assemblies made up of functions, information items or components spread over information infrastructures and the internet. The hypertext, for instance, underlying many digital documents is just a network of various online media interlinked by a multitude of diverse items, devices and producers. Distributedness confers digital objects some interesting qualities. Digital objects are *borderless*. In comparison to packaged and single media like books, networked media does not have an identifiable border defining it as an obvious entity (Ekbia, 2009; Esposito, 2002: 299). These borders have to be created and maintained technologically. Furthermore, distibutedness makes possible various combinations out of a larger ecology of items, procedures and programs, a condition that renders digital objects fluid and crucially *transfigurable* (Haider and Sundin, 2010). Finally, distributedness accentuates the significance of the links and the *assembly procedures* by which a digital object is brought to being and, at the same time, weakens the importance each item may have as a standalone element.[3]

Compositional Texture

The pliable and generative nature of digital objects sets them apart from artifacts of non-digital constitution and raises the question as to what confers them the qualities of editability, interactivity, openness and distributedness. The issue is certainly related to the prevailing cultural predispositions and attitudes on the basis of which technical artifacts are seen as or made malleable and customizable to the needs of particular individuals, communities or professions (Zittrain, 2008). A web of significations develops around the meaning and use of artifacts and technological systems that Bijker (2001) has subsumed

under the construct of the *technological frame*. As the frame congeals, it furnishes the semantic matrix that reproduces these same attitudes and orientations vis-à-vis an artifact or a family of artifacts. To a certain degree, the enactment of the qualities of digital objects we refer to are associated with a *user-centered technological frame* that sees software and its tokens as the target or medium of individual needs, proclivities or aspirations.

Nevertheless, the qualities of digital objects we list in this paper are not fully accountable by reference to cultural predispositions alone and the meanings developing around the use of an artifact. In addressing this question, it is necessary to distinguish intrinsic from contingent factors. As any object, digital objects are brought to being under varying conditions (institutional settings, resources, skills) and this explains why the basic attributes we ascribe to them exhibit significant empirical variability. But empirical variability does not and cannot address the issue on the basis of which objects are classified as digital and are thus sharply distinguishable from non-digital objects. Only intrinsic and necessary characteristics can accomplish such a task (Sayer, 2000). Let us elaborate.

The attributes we ascribe to digital objects have often been associated with flexible, end-to-end architectures, ultimately resting on the **modular composition** of software and the operations this last enables. Modularity refers to the organization of items and operations that make up a digital object, or an interacting ecology of such objects, in distinct and relatively self-sufficient blocks or units that allow for independence within a wider yet loosely coupled network of functional relationships and dependencies (Benkler, 2006; Kallinikos, 2006; Manovich, 2001). The loose coupling modularity affords allows local manipulation of digital objects without notable effects on the wider system of technical relations into which the object is embedded. It also enables the decomposition of the elements by which digital objects are made and, crucially, the reshuffling and reorganization of these elements to new configurations. In this respect, modularity represents the technical realization of the simple yet powerful idea that *en bloc* objects or operations are hard to act upon and manipulate, a condition that can significantly be altered by conceiving and designing objects as modular.

The breaking up of the *en bloc* character of objects modularity enables is closely associated with the granular constitution of digital objects (Benkler, 2006). As distinct from modularity, **granularity** refers to the minute size and resilience of the elementary units or

items by which a digital object is constituted. Long before the advent of digital technology, verbal alphabetic writing has furnished the archetype of granularity. Verbal alphabetic writing epitomizes a limited number of elementary units (alphabetic characters) whose rule-based combinations enable ascending from the level of individual character through syllables, words and sentences to wider verbal constructions such as texts or discourses. Digital item lists or software libraries provide other instructive examples. A digital list or software library is granular in the sense that the items by which it is made of are separable and clearly differentiated from other items in the list (Weinberger, 2007). For this reason, they can be manipulated (modified, deleted, added) independently of one another and brought to various configurations. The elaborate granular constitution of digital objects is closely associated to their numerical nature and the possibility this furnishes for tracing digital objects deep down to the most minute elements and operations by which they are made (Borgmann, 1999; Manovich, 2001)

Modularity and granularity are inherent to many social practices (e.g. architecture, flexible manufacturing) and the media by which these practices are carried out. However, the making of granularity to a ubiquitous technological principle that relies on binary parsing is a remarkable accomplishment. It confers digital objects a distinct ontological and functional profile and all those qualities we associate with them. In this respect, granularity and the numerical constitution of digital objects furnish the generative matrix, the genetics, as it were, of the properties of editability, interactivity, openness and distributedness. Any system that is made of small, recurrent and identifiable elements that can be decomposed and assembled through a series of operations back to the system again could be defined as granular. As a rule, analogue systems do not obey these principles of organization made of elements in tangled forms that are not decomposable and readily reassembled to the system they were once part of (Goodman, 1976; Kallinikos, 2009a).

Granularity applies to several layers of the digital object but two groups of operations set the concept apart from that of modularity. First, granularity allows the tracing down of the behavior of a digital object to several layers of underlying operations (e.g. a database can be datamined, a video edited by video editing software) by which it is sustained. No matter how difficult this may be in practice, it is always in principle possible thanks to the binary and numerical status of digital objects. Secondly, granularity enables minute and piecemeal intervention, as cases like Wikipedia editing and open source software development reveal. The fine-grained nature of digital objects enables people to contribute to

collective pursuits under widely varying circumstances that fit their time availability, capacity or inclination (Benkler, 2006). To a certain extent, distinguishing granularity from modularity may seem to be a terminological issue. Yet the two concepts direct attention to different facets of digital objects and the operations by which they are constituted.

The construal of digital objects in these terms may be invoked to account for the making of the interconnected information environment in which we live in and the much more flexible interaction between users and artifacts, people and institutions. At the same time, the diffusion of digital objects and the attributes they embody across the social and institutional fabric is associated with new problems and risks that have not always been adequately appreciated. The stability and identifiability of the object world in which human activities are usually embedded are key to the forms of experience such a world sustains. The instrumentation of means-ends sequences, the attribution of cause-effect relationships (March and Olsen, 1989) and sensemaking in general are essentially supported by the stability of the tools and objects on which actors draw upon. In this respect, the malleable and transfigurable character of digital objects undermines basic facts of human experience and may end up constructing a less accountable environment.

In the next two sections we provide an account of two fields of contemporary social practice that demonstrate the double-edged processes into which digital objects become embedded and the promise of the theoretical ideas outlined above. The first one is the *archive*. This age-old institution wrestles with novel dilemmas of organizing and preserving digitally born culture. The second concerns the *document search apparatus* deeply embedded into the digital online environment, and the significance findability is acquiring in shaping cultural records as search engines become the primary mechanism of mediating access to these records. The case illustrations discuss the contrasting sets of problems and opportunities these two fields of practice confront and the strategies they deploy to cope with these. An archive of digital objects is an attempt to freeze the inherent fluidity of digital objects. It is a reinvention of the archival function that – as a cultural memory – seeks to maintain the identifiability of cultural artefacts over time. Search engines employ a rather different approach. They are disinterested in the objects themselves but make their findability a paramount concern. Ultimately, these two strategies of *freezing* and *finding* become ways of constructing some of the units of culture and perception that populate our interconnected information space.

The Making of a Memorable Web

A broad range of societal practices rely on enduring and persistent artifacts (e.g. books, imagery, documents) authenticated, canonized and collected by dedicated authorities for reasons of documentation, reference and identity (Assmann, 2008a). Cultural heritage institutions such as museums, libraries and archives are a case in point. In the form of digital objects, however, cultural artifacts undergo constant change that renders their identification over time a problem (Coyle, 2008). While the fluidity of digital cultural artifacts may be of little concern in day-to-day social interaction, it is nevertheless bound to have a significant impact on the ability of future generations to access historical documents, a vital and pervasive social practice. In its ever increasing capability to store data, contemporary society may paradoxically end up with a much weaker institutional memory (Young, 1996; Brindley, 2009). Already, early artifacts of the digital age are inaccessible due to the disintegration of the medium they are stored on or because the respective hard- and software standards used to create and access these artifacts are obsolete. Now that an increasing degree of communication is conveyed via online services, even more of our cultural heritage runs the risk of disappearing into ephemera due to the absence of an institutionalized archival trustee and dedicated custodian.

Archives have played a crucial role in this context (Cox, 2007). Committed to providing persistent access to reliable testimonies of social facts, archives have been entrusted with the key tasks of collecting unique cultural items, documenting their provenance and preserving their integrity. The ways these tasks have been carried out reflect a longstanding process of institutionalization whereby skills, values, materials and technologies have coalesced around the formation of a practice centering on the authenticity of a document (Dreyfus and Spinosa, 1997: 180). From this perspective, cultural artifacts are not stumbled upon but collected because they are deemed worthy of being selected, catalogued and preserved as evidence of past social facts. At the core of archival practice lies the identifiability of a specific document as the very same document not only today but also in a year, a decade or even a century. Archives maintain the identifiability of cultural artifacts over time (Assmann, 2008b). In order to provide reliable testimony of social facts, cultural artifacts are, after being selected, turned into archived documents by being indexed, catalogued and preserved. In this processes, the cultural artifact that is the object of traditional archiving practice does not change. It is just cut off from its context and brought into an archival setting. The intact character of the culture item being archived is a core principle of archiving that is crucially associated with the quality of the evidence it

mediates. This core principle is currently challenged by the compositional texture and attributes of digital objects.

A digital document is not simply designed and developed digitally, nor is it just a digital record. Understood as digital objects, digital documents have a double mode of existence, being composed by the content arrangement they mediate plus the operations by which this content is assembled and maintained. In the digital, networked environment, digital documents gain their functionality thanks to these operations. Hence, the preservation of a digital document needs to take into account the attributes and compositional texture its functionality is based on. Since these qualities tend to violate standard principles deployed by memory institutions, the question arises how to preserve digital objects for future generations (Lyman and Kahle, 1998; Marcum, 2003). In what follows we will show some of the operations of rendering digital objects persistent and hence identifiable for future reference and the challenges this entails in the contemporary digital and networked environment (Greenstein, 2000; Cox, 2007; Schnapp, 2008).

The Internet Archive (www.archive.org) is an attempt to prevent digital objects from disappearing into a traceless past (Green, 2002). Amongst its various activities, it is mostly known for archiving the Web in order to "change the content of the Internet from ephemera to enduring artifacts"[4]. Since its foundation in 1996, it has managed to build a collection of over 150 billion pages[5] making it, by now, the biggest database in the world with a growth rate of roughly 20 terabytes a month[6]. A user can access and browse through the collection by means of the, so called, WayBackMachine, a service that allows to search for previous versions of a webpage based on a URL query. The database is also openly accessible to researchers for data mining which makes this collection an archival source for research on the Web rather than a library service that stresses the accessibility and usability of its library collection (Marton, 2009b).

Given the nature of the internet and the digital objects it brings forth, the Internet Archive developed new practices in order to document the evolution of the Web. The first notable difference is the way the immensely growing amount of online content is being selected for archiving based on algorithmic search engine harvesting. The major part of the collection is provided by the for-profit search engine Alexa (www.alexa.com). Like most contemporary search engines, Alexa copies web content into a database for indexing. After the commercial value of the data has expired, the copies are donated to the Internet Ar-

chive for preservation. Given the fact that the Alexa crawl is based on the popularity of a website among the Alexa user community[7], some sites are extensively documented by the Internet Archive, while some are not at all (Howell, 2006). Taking into account that an estimated 16% of indexable webpages actually are harvested and indexed by search engines, only a relatively small proportion of the online world is processed and archived (Thelwall and Vaughan, 2004). The important difference in terms of archival practice is that the selection process is not based on professional experts but on the popularity of content derived from online user behavior. What is being archived and therefore selected to represent the Web of the past is following the very same rationale that makes the Web navigable – the rationale of search engines. In other words, history will be remembered, if ever, through the algorithmic eyes of search engine technology.

The central issue in relation to our argumentation, however, is the practice of preserving online content – that is preserving digital objects as digital objects. Online content shows a high degree of editability (e.g. wiki pages), interactivity (e.g. discussion boards), openness (e.g. crawled by search engines), and distributedness (e.g. dynamic webpages drawing content from various databases and image collections). As a consequence, online content is highly transfigurable, it does not present itself in the form of clear cut, easily identifiable documents with distinct borders (Hjorland, 2000). The preservation efforts of the Internet Archive need to counter some of these attributes in order to cast digital objects into persistent cultural forms. As we demonstrate in the following pages, archived cultural artifacts are constructed rather than collected by ways of snapshots taken of online digital objects *freezing* their actual content into a fixed and preservable entity. The actual archival function of guaranteeing the provenance, integrity and authenticity (Seadle and Greifeneder, 2008) of a cultural item is in this context redefined by preserving a version of the cultural item delimited in terms of its attributes.

The snapshot is basically a copy of what is rendered as html in a browser. A dynamic webpage, for instance, mostly consists of instructions how to generate an actual webpage to be displayed to a user. Processed by a web server, these instructions compile the assemblage of various parts found in various sources. The snapshot taken is not of these instructions, nor can it be, but of the resulting page temporarily assembled and rendered as a html page for a given user. The archived digital object is not a dynamic webpage anymore but a static one, its constitution not depending on the access to original sources for up-to-date information. By the same token, if something cannot be harvested by the

search engine, it cannot be archived. In the final step of the process, the snapshot is tagged with a timestamp in order to document the date and time of its harvest.

The central question however is how two archived snapshots are identified as different versions of the same webpage? It is neither, nor can it be, the content or the title of the webpage, since these may have changed in the meantime. Instead the combination of the webpage's URL and its file name is used as the identity marker. Hence, the very same service that allows for the finding of pages or rather locations on the Web is also applied in the Internet Archive to identify a series of archived items as snapshots of the same webpage. It is the webpage that performs as the basic unit of the collection – as the definition of the document. This has far reaching implications, since the archive does not simply collect already bounded entities themselves but rather seeks to construct the boundaries that demarcate and therefore make an archival document. It is the Internet Archive that produces persistent artifacts relying on rules and procedures that change the societal role memory institutions have been entitled with. The provenance and authenticity of historical documents are not merely recorded and preserved together with the document itself but rather actively created in order to transform a transfigurable digital object into a clear-cut and identifiable archival item bereft of the degree of editability, interactivity, openness and distributedness the original affords.

This is only possible since webpages are defined as computer files rather than as cultural records and content carriers of social events. An online webpage may only consist of instructions on how it is to be assembled and is thus left devoid of any content until it is presented in a browser. By contrast, its URL and filename are uniquely identifiable by the automated computational procedures of the archive. Therefore, if the URL or the filename of a webpage changes, it is a new webpage in the eyes of the Internet Archive even when the content stays the same. In comparison, hypertext ends up being even fuzzier than on the Web. The reason lies in the modification of hyperlinks to point to the temporally closest archived version of the target webpage based on the timestamp of the snapshot. Thus, a user of the WayBackMachine can surf through a very popular and, therefore, thoroughly harvested website or even follow a link to another domain if the target of the link is also part of the collection. A user may, for instance, access the Microsoft website from October 1996 in its totality, follow a hyperlink to another website from January 1998 or end up on the Web if the target page of the hyperlink was never archived. In other words, the transfigurability of hypertext is increased within the collection since a user

does not only jump from webpage to webpage but also from one moment in time to another moment in time. Cut off from the ongoing digital environment from which it derives the hypertext ends up being even less coherent, fuzzier with even weaker edges in the Internet Archive. Obviously, hypertext – though a digital object – does not offer itself as a documentary unit the way a webpage does since it is not defined technologically. Hypertext does not provide for the equivalent of a filename in terms of its identifiability nor does it allow for a coherent temporal fixation. Bits and pieces are archived from different moments in time while some bits and pieces are not archived at all. In short, the elementary archival unit is based on technological not semantic (profession-based) considerations.

The issue, however, is not simply whether the snapshot retains the characteristics of the digital objects described above but rather what kind of relationship it maintains with the ongoing character of the online digital environment from which it is harvested and, by extension, the degree to which it can be invoked as a reliable testimony of the social facts it refers to. Our observations suggest various problems in this regard. Based on the technologically delimited possibilities of what and how to preserve, it constructs rather than collects identifiable digital objects by freezing their mutability and, therefore, making each snapshot *re*cognizable. The memorization of the Web is not a mere copying process of bits and bytes from one server to another but rather a transformation of digital objects into a different type of digital object that is made to fit into the archival world of provenance and authenticity. In light of our argumentation, this transformational process goes beyond the traditional practices of collection, documentation and preservation leading not merely to a change of the context in which the object is embedded but to a change of the object itself. The digital object is not documented by recourse to external, professionally produced rules but reflexively produces its own documentation. No matter which change documentation may bring to the object it seeks to archive it always encounters a version of that object. In the example we described, the digital object owes its ontology to the computer-based operations by which it is brought to being. What is being archived is not the unique, original artifact but a transformed version of it.

The Making of a Navigable Web

The Web has grown from an impenetrable morass into a surprisingly pliable source of information thanks to dramatic innovations in document search. In contrast to libraries and archives, search engines dissociate the mechanism for accessing information from the

practices of preserving cultural records. The companies maintaining major search engines have generally no stake in the preservation of the items they make accessible (Brindley, 2009). Projects such as Alexa's cooperation with the Internet Archive or Google Books are mere exceptions to this rule, since the business model of search engines is by and large satisfied, and so are search engine users, as long as something useful comes up in the search results. This superficial indifference obscures, however, the intriguing issue of the influence web document search has upon what is being found. The matter is not how well algorithmic search engines represent what is "out there" – as this has already been discussed in numerous studies (cf., Introna and Nissenbaum, 2000; Shaker, 2006; Zick, 2000; Waller, 2009; Fortunato, et al., 2006). Rather, what is "out there", we suggest, is to some degree shaped by the exogenous pressure on the actual content of cultural records search imposes by making findability an intrinsic consideration of the very production of such records (Morville, 2005).

In order to shed light on this matter, we distinguish between the search engine results page, and the relationship the results page has with the respective target webpages. The former is the familiar display of search results listed according to their assumed relevance while the latter refers to the items matching the search query. One might argue that the search engine results page merely provides access to digital cultural records and it does not need to be considered other than a simple display of such records. Such a belief is re-inforced by the fact that individual page instances are discarded after the user clicks one of the links provided. However, taken together the results pages furnish the means that mediate our relationship with the objects of knowledge in non-trivial ways. The search engine results page is not a mere list of digital cultural records, but a nexus of important and evolving mechanisms shaping our informated environment. Let us elaborate.

Document search has evolved in less than twenty years from an obscure operation preferably left to trained librarians into a crucial cognitive style needed to cope with the exigencies of the contemporary networked environment of information affluence. The usefulness of the Web rests to a significant degree on search engines that have progressively taken over the practice of following hyperlinks from a website to another as the dominant solution for navigating the Web (Evans, 2007; Fortunato, et al., 2006; Sen, 2005). The utility of search stems from its superficial straightforwardness or, to be more precise, from the immense reduction of complexity it affords. A search engine orders the Web in real-time to fit the user's queries. Early directory-based systems relied on human editorial

labor and professional practices such as librarianship to rank individual pages into the categories of the search database. In 1998, the founders of Google successfully adapted a method used for scientific citation indexes to rank webpages automatically thus paving the way for second-generation algorithmic search engines. A key business model innovation was to start inferring people's ongoing interests from the submitted search keywords and to accompany organic search results with context-sensitive advertising (Bermejo, 2009). People are more likely to click higher than lower ranking items in search results and, task-specific variation notwithstanding, usually settle for the links listed on the first page of results even if there would be thousands of choices available just a click away on the consecutive results pages (Höchstötter and Lewandowski, 2009; Malaga, 2008). Given our increasing reliance on search, this behavioral pattern makes it desirable for a wide range of organizations, groups, as well as individuals to attain a high ranking among relevant search results.

There neither is nor can be a single best or neutral way to order internet resources (Höchstötter and Lewandowski, 2009; Introna and Nissenbaum, 2000). Search not only makes some things more likely to be found than others but also breaks away from stable classifications and the normative importance of categories as the basis of such order. Scholars have also identified a number of potential biases. It has been argued that search engines index only a fraction of the whole Web, occasionally promote undesirable material and amplify the popularity of already popular pages as a result of the way their algorithms rank pages (Waller, 2009). Furthermore, the value of organic ranking has created a market for consulting companies offering search engine optimization services for website owners to increase their likelihood of achieving high ranking. More important than any specific bias is, however, how users relate to the potential prejudices. Introna and Nissenbaum (2000: 176) duly note "not only are most users unaware of these particular biases, they seem also to be unaware that they are unaware".

The evolving Web search apparatus has turned out to be a moving target for spot studies tied to particular, timely datasets. The ranking algorithms and optimization methods are in constant flux rendering some of the details reported in previous studies already obsolete. For instance, metadata, which underpins most professional cataloguing practices, is today generally ignored as a ranking criterion by search engines that attempt to base their ranking on the actual text and other user visible features of webpages to avoid being tricked by website owners (Introna and Nissenbaum, 2000; Malaga, 2008; Zhang and

Dimitro, 2005). This is just one example of a bigger game of mutual adjustment between the ranking methods and website owners catalyzed by the search engine optimization consultants who make inferences about undisclosed ranking algorithms and adjust their suggestions for website owners accordingly. Scholarly studies chasing the changing nature of search engines have produced a number of useful observations yet the nature of the wider institutional transition remains largely unexplored. Apart from some anecdotal evidence, we do not know how search shapes the constitution of cultural records in the interconnected information space and the operations of the actors involved. Theorizing digital objects the way we do helps to pin down some of these dynamics that have evaded most attempts to understand search engines.

The results page is an interactive, radically open and distributed artifact that sits between human actors and the cultural records they wish to access. Just like its traditional counterpart, the library catalogue, the results page both orders and locates cultural records. Despite being a temporary achievement contingent upon the user-specified keywords and the continuously evolving search index, search results impose order that can reinforce identity and support authority by granting publicity for those who are able to achieve a high ranking (Bowker and Star, 1999; Morville, 2005). The mediation provided by the search results is not, however, controlled by a single institution but distributed between the search engine companies, webpage editors, optimization consultants and advertisers who all influence what counts as a relevant search result in the open anatomy of the search engine results page. The inherent distributedness of the search engine results page runs from the ways its open constitution brings together snippets of target pages across a myriad of institutional and cultural boundaries in real time. Even if these two objects, the search results and the matching Web resources, are wrapped as separate webpages it is difficult to draw a definitive line between them. If the content of the target pages changes, so does the search engine results page that is, in a sense, nothing but a temporary assembly of other objects. Placed in this light, the idea that Google could control the world's information through its admittedly powerful search engine (Waller, 2009) seems exaggerated and misses more evasive institutional dynamics that are currently taking shape.

Even if a particular cultural record can be located through search today, this may not be the case tomorrow. The search engine results page is an unpredictable form of mediation in contrast to the library catalogue firmly embedded in professional frameworks often active within a national culture. By virtue of its interactive, open and distributed constitu-

tion, the search engine results page introduces a peculiar instability into information access and retrieval. This instability is starting to resonate with the equal mutability of target pages opening up mutually constituting relationships between the search index and what is being indexed. The items available through searches are to a varying degree editable ranging from somewhat laborious modifications to their underlying text and media files to wiki systems affording equally the reading and writing of cultural records. Version control and permalink systems provide some stability for digital cultural records and their addresses in the cloud, but cannot protect the cultural records from the contingent findability encroaching on target pages. Given that the inclusion of a page in algorithmic search results is a transient achievement, editability makes it possible to try to maintain and enhance this possibility by constantly tweaking the page so that it serves better the search engine ranking algorithms. The rise of search engine optimization consultants is a testimony to the fact that website owners are increasingly tapping into this possibility.

The way the search engine results page mediates cultural records is not (yet) controlled by an established institutional setting and is therefore inherently unstable. This instability has generated a market opportunity that feeds back into the editable nature of digital cultural records due to tangible benefits that can be reaped by manipulating webpages to enhance their findability through search. Unlike library cataloguing that attaches metadata to documents, optimizing cultural records for search engines subtly but unavoidably shapes the objects we look for. These operations range from a simple on-page optimization method of repeating certain keywords in the page title and body text (Zhang and Dimitro, 2005) to more complex off-page approaches altering the network topology around the content node for example by stimulating hyperlinking from other pages. Even if the changes required to make a document rank higher in the search index are relatively small, they introduce nonetheless a new source of variation that makes, for instance, the preservation of cultural records more difficult. Preservation requires freezing the digital object, while maintaining its findability provokes constant rewriting. Technologies such as version control systems may provide some stability for individual versions of the cultural artifact while laying ground at the same time for the endless proliferation of updated editions. Wikis are a prime example of this. Finally, the importance of being found is also influencing the rationale of content creation as Morville (2005: 142) points out: "Articles, books and blogs are not simply destinations, for they often serve as inverse queries that draw users to authors. We write, not just to communicate, but to enhance our own personal findability." While not a probably completely new phenomenon, the minimal pro-

duction costs and the chance of attracting unexpected readership have certainly helped to unleash a wave of egocentric publishing in the Web.

Conducting a Web search is today a matter of simple routine, a condition that renders search engines increasingly part of the invisible equipment with which we encounter the everyday world (Dreyfus, 1991). The internet document search apparatus puts a premium on the findability of digital objects. It has probably never been irrelevant for authors to think of how to make their pieces available to the audience, but before this game was played against the thousands of different outlets and their local rules. If a publisher put unreasonable demands on content for the sake of accessibility and market appeal, it was, at least in principle, possible to try another outlet. The observations made in this section can be seen as a gradual disembedment of this game from such local contexts and its transition to another generic context marked by the interplay between few major search engines and search engine optimization consultants. If, as it has been argued here, making the interconnected information space navigable entails tailoring its objects to be findable by few major search engines, this could develop into a new kind of global isomorphism. The argument is aligned with neo-institutional theory stating that entities under similar environmental pressures tend to adopt similar structures (Scott, 2001; DiMaggio and Powell, 1983). The short history of the search engine industry has seen a rapid concentration to a point where few major search engines provide over 90 percent of Web searches[8].

Discussion: Digital Objects and Social Practice

The Internet Archive and the document search apparatus presented above are certainly snapshots of a far more complex institutional and technological change. And yet, these examples provide incisive and useful illustrations of the challenges raised by the deepening involvement of digital objects in social life.

The case of the Internet Archive suggests that the category of the document, a pillar of the practice of archiving, is not any longer a clear-cut and evident object of social practice. Constantly mutating bits and pieces of content distributed over the web are harvested and assembled to cultural units that are frozen and stored as distal documents. In this process, the definition of what is to count as a document to be archived is embedded into and performed by means of software. The bits and pieces that are assembled into a digital object are themselves an assemblage of modules. In other words, we find digital objects within digital objects within digital objects and so forth. On the Web, hypertext is a fuzzy and

continuously shifting assemblage of hyperlinked webpages while each webpage is itself an assemblage of text, imagery, databased information, etc. However, which assemblage ends up being preserved as an archival document is not based on professional semantics and rules but on technological considerations. As far as the Internet Archive is concerned, the focus is on the webpage that is rendered the elementary unit of collection. Now, a webpage is technologically defined as the html-file or rather what renders as html in the browser. Hypertext, on the other hand, though meaningful, defies a technological delimitation. This lack of delimitation is further aggravated by the changes which the bits and pieces of the hypertext are undergone over time. It comes therefore as no surprise that the archival documentation of digital objects is based on a procedural rationale that derives from information technology and the operations performed by software. What is to be preserved is selected by means of search engine algorithms, identified by the combination of URL and file name, and preserved as a snapshot.

In this regard, the generativity of digital objects and the opportunities it creates are offset by the evasive character of digital documents, and the tasks and operations that pivot around the stabilization and management of these documents. A limited sample of the population of digital documents circulating in the Web is selected and stabilized by means of unobtrusive, technologically driven, processes with the view of rendering them identifiable and archivable. This is one alternative a digitally born culture offers to the practices of memory institutions that over a considerable period of time have been shaped by professional rules and an elaborate set of organizational routines and procedures. Little wonder that the identification and selection of documents to archives performed by traditional means entailed cultural biases of various kinds. After all, a document that is rendered the target of archiving is a cultural artifact defined and partly brought to being by the knowledge, skills and practices of memory institutions and the professions they host. However, the solution provided by the Internet Archive is of an altogether different nature. The whole arrangement could be seen as a good approximation of the kind of problems the construction of cultural memory engenders as digital objects diffuse throughout the institutional fabric of modern societies.

Some of these issues are further reinforced while new ones emerge as search engines become the primary medium for identifying documents and information in a navigable Web. The document search apparatus provides an inherently unstable form of mediation as compared to memory institutions. By shifting away from stable categorizations em-

bedded in the professional practice of memory institutions, it promotes a formal and context-free logic of search that makes the findability of websites a driving force in the development of the web content. To some degree, this may seem as a sheer technical matter in which target pages seek to adapt to the way the results page is identified and displayed by search engines. But it is more than this. As we have been at pains to show above, an exogenous pressure is exerted upon producers of information to present and arrange information provided by websites in ways that make that information findable and accessible. Such a task in turn promotes opportunism manifested in the drive to be as visible as possible by deploying a technologically supported practice called search engine optimization. In this regard, findability works at cross purposes with the logic of authentication, identification and preservation underlying the traditional practices of archiving and institutional memory building. Table 1 offers a schematic summary of the key blocks of findings derived from the study of the Internet Archive and the document search apparatus:

	The Making of a Memorable Web	**The Making of a Navigable Web**
Institutional setting	Memory institutions struggling to preserve digital objects for future generations	The institutionalization of a web search apparatus drawing together search engines, optimization consultants and website owners
Function	Focus on preservation: Create a memory that societal practices can draw on	Focus on findability: Enable immediate information access and retrieval
Technologically-induced tension	The impossibility of archiving hypertext	The destabilization of document retrieval
Challenge	The provision of persistent and recognizable cultural artifacts	The provision of relevant search results
Counter-mechanism	Freezing the fluidity of digital cultural records	The mutual constitution of the search index and what is being indexed
Emerging practices	Reliance on search engine technologies for the selection of archivable records, creation of new kinds of digital objects by taking snapshots of webpages	Tailoring digital objects for search engines by constantly rewriting them (and thus escalating the number of editions)

Table 1: Web Memorization versus Web Navigability

The examples of the Internet Archive and the network and mechanisms underlying digital search suggest that the theory of digital objects presented in this article is a useful concep-

tual grid for analyzing the growing diffusion of information technologies and artifacts in social and economic life. The theory helps focusing on the **dynamic processes** and **interactions** by means of which the aspects of the world we work and live in is assembled to a transient order (Weinberger, 2007). In this respect, the attributes of editability, interactivity, openness and distributedness we ascribe to digital objects and their modular and granular constitution pierces deep down into the processes, media and interactions by which they (digital objects) are sustained. It thus helps to disentangle particular facets of the composite reality into which digital objects are embedded and subject it to analytic scrutiny and examination.

The attributes we ascribe to digital objects and their modular and granular constitution are no doubt generic qualities that do not confront the specific nature of particular technologies and the functionalities they embody. But functionality is too conspicuous an attribute that offers itself to observation rather straightforwardly. There is no way to discuss any particular technology without confronting the primary functional task it addresses (as we did here with the Internet Archive and web document search). Functionality is also very closely tied to the user-centered technological frame mentioned earlier in this article. Our account of digital objects reveals a complex and double-edged process of user involvement that considerably qualifies the constructs of "functionality" and "user". On the one hand, the generativity of digital objects enables and empowers users. The editable, interactive, open and distributed properties of digital objects circumscribe a space of possibilities in which users can assimilate the use of digital technologies and artifacts to the specific projects they pursue and the needs or feelings they wish to express. On the other hand, the theory of digital objects discloses a vast space of processes and mechanism beyond the discretion and perception of users and the straightforward functionality digital objects embody. In a sense, the theory de-centers the user and reveals a complex and distributed apparatus of data-driven operations in which digital object functionality and user are no more than nodes on the steadily mutating and displacing information universe (Kallinikos, 2006).

In this respect, it would be interesting to observe that digital objects are objects only in an elusive and perhaps euphemistic way. For, the steady transfiguration and the permeable boundaries underlying them suggest that they are no more than **operations** by means of which they are assembled to proxies of objects (Ekbia, 2009; Manovich, 2001) only to be unpacked, edited, reprogrammed and reassembled again. The theory of digital objects

thus draws attention away from fixed entities and the role they are supposed to serve. Instead, it focuses on the data driven operations (editability, interactivity, openness) by which digital perceptions and practices are assembled and dissolve, and the complex ecology of relations (distributedness, openness) within which these operations are embedded. This is an important insight and a key contribution this paper claims it makes to the literature.

Do these observations justify reifying the attributes of digital objects the way we do? Are not these attributes, as the motto goes, socially constructed and therefore shaped by the choices of human actors? Are not editability and interactivity the outcome of human design choices that determine whether and to which degree an artifact should be made editable or interactive? Are not distributedness and openness ultimately sustained by political choices concerning the standards underlying computer networks and the interactive patterns, modes of use and access prevailing in the internet? Certainly! All artifacts on earth are human made and technology is no exception. But from this does not follow that technical artifacts are simple conventions or agreements ready to dissolve by volition and preference invocation. Even in the case that the attributes we ascribe to digital objects are the outcome of cultural predispositions that congeal, as Bijker (2001) wants, to a stable technological frame they merit consideration. For, being part of that frame these attributes shape the ways human actors approach and use them.

The distributed online environment in which digital objects are deployed in roughly the way we outline in this paper Zittrain (2008) calls "generative internet". The internet as we have experienced it so far owes its freedom to the end-to-end architectures and the concomitant capacity to use computers as open, editable and reprogrammable machines. Such a state, Zittrain claims, is facing the unfortunate prospect of becoming a closing circuit of controlled online interactions under the growing security risks and the relentless profit pursuit that haunt the internet. If realized, this unfortunate prospect would lead to considerably restricting the open, editable and reprogrammable nature of computers transforming them in what he calls "information appliances". The prospect of closing up the open avenues of interactivity, creativity and freedom that prevail in the present regime may seem to provide another example of the social construction of technology and how experience of artifacts ultimately rests on human values, regulation and power. But whether in the current form or the new forms that may result through the renegotiation and closure of the internet, the key implications we ascribe to digital objects will persist.

An internet archive will need to confront some of the key issues we have outlined here event though some of the social forms and technical means this will be sought may differ. And findability will still be a problem and an opportunity in the information affluence of the contemporary digital world even though it may be shaped or arrived by different social and technical routes.

There is more to the construct of digital objects than just a solidification of attitudes and predispositions vis-à-vis technologies that Bijker (2001) and company (cf., Bijker, et al., 1987) and, to a certain degree, Zittrain (2008) advocate. For, technology as a means in the service of human ends exemplify a particular mode of being which is that of objectification, that is, the embodiment of functions onto specifically designed objects. Thus objectified functions retrieve from the human interface (Borgmann, 1984) and contact with the machine gets thus simplified so that it can be summoned and enacted by anyone in possession of the relevant skills (e.g. biking, driving, computer-based text editing). As a strategy, objectification stands at the opposite end of norm, value and skill interiorization that underlie the construction of agency forms and occupational identity by cultural means (Kallinikos and Hasselbladh, 2009; Lessig, 1999[2006]). It is the implications of this strategy that we sought to retrace and analyze by focusing on the generic attributes and constitution of digital technologies.

Concluding Remarks

In this paper we have sought to outline the elements of a theory of digital objects and spell out the implications such a theory may have for our understanding of social practices. Digital objects are editable, interactive, open or reprogrammable and distributed. Rather than being simply the contingent outcome of design, these attributes derive from the constitutional texture of digital technologies, most notably the modular and granular make-up of digital objects and their numerical nature. Taken together the attributes of digital objects and the operations by which they are sustained mingle with social practices redefining the scope, the object of work and the modes of conduct underlying them. We have provided evidence for our claims by analyzing the logic and the technologically driven procedures by which the memory of the fluid and shifting nature of information in the Web is frozen down to identifiable documents. We further considered and analyzed the digital search apparatus that functionally dominates the Web and the significance the formation of this complex has for the content of these documents and ultimately the Web itself.

The interaction of technology with social practices and the way technological artifacts shape and are shaped by human agency is one of the most vexed issues in the social and administrative sciences (Sismondo, 1993). We have here sought to navigate between the many divides that afflict the social study of technology and provide the conceptual space upon which the latter can be viewed as a potent agent of social change. Technology matters, not in the simple and unambiguous ways technological determinism depicts the issue, but it matters nevertheless. If this is the case then the "how it matters" needs imaginative theorizing and careful empirical documentation. This paper has responded to this quest by providing the conceptual scaffold on which to craft a theory of digital objects and by supplying empirical evidence that indicates the timely and relevant character of such a venture. It goes almost without saying that much more reflection and research are needed in this direction.

---

[1] From the collection *Aleph* in Jorge Luis Borges, *Collected Fictions*, New York: Penguin, 1998. The italicized "is" occurs in the original.

[2] These characteristics are currently renegotiated under the security threats an open internet posits and the dilution of copyright it allows (e.g., Lessig, 1999[2006]; Zittrain, 2008).

[3] This is surely the case with complex systems and natural language. The characters of the alphabet itself do not make sense as standalone marks but as elements of the collective nature of the marks that make up the alphabet as a signifying medium (e.g., Borgmann, 1999).

[4] www.archive.org/about/about.php, accessed on 6th Jan. 2010.

[5] www.archive.org/index.php, accessed on 6th Jan. 2010.

[6] www.archive.org/about/faqs.php#8, accessed on 6th Jan. 2010.

[7] See the top 500 sites at http://www.alexa.com/topsites, accessed on 13th Jan. 2010.

[8] Search Engine Watch (15 September 2009): Top Search Providers for August 2009. http://searchenginewatch.com/3634991. After recent Microsoft-Yahoo! deal the vast majority of search results are provided by Google and Microsoft. There are few countries that are exceptions to this with their local search engines, but the global trend over the last ten years is clear.

References

Aleida Assmann, 2008a. "Canon and Archive," In: A. Erll and A. Nünning (editors). *Cultural memory studies: an international and interdisciplinary handbook*. New York: Walter de Gruyter, pp. 97-107.

Jan Assmann, 2008b. "Communicative and Cultural Memory," In: A. Erll and A. Nünning (editors). *Cultural memory studies: an international and interdisciplinary handbook*. New York: Walter de Gruyter, pp. 109–118.

Yochai Benkler, 2006. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven: Yale University Press.

Fernando Bermejo, 2009. Audience manufacture in historical perspective: from broadcasting to Google, *New Media & Society*, volume 11, number 1&2, pp. 133–154.

Wiebe E. Bijker, 2001. "Understanding Technological Culture through a Constructivist View of Science, Technology, and Society," In: Stephen. H. Cutcliffe, and Carl Mitcham (editors). *Visions of STS: Counterpoints in Science, Technology and Society.* Albany: State University of New York Press, pp. 19–34.

Wiebe E. Bijker, Thomas P. Hughes, and Trevor Pinch (editors), 1987. *The Social Construction of Technological Systems.* Cambridge, MA: The MIT Press.

Lynne Brindley, 2009. "We're in danger of losing our memories," at http://www.guardian.co.uk/technology/2009/jan/25/internet-heritage, accessed 13 January 2010.

Albert Borgmann, 1984. *Technology and the Character of Contemporary Life.* Chicago: The University of Chicago Press.

Albert Borgmann, 1999. *Holding on to Reality: The Nature of Information at the Turn of the Millennium*. Chicago: The University of Chicago Press.

Geoffrey C. Bowker and Susan Leigh Star, 1999. *Sorting Things Out: Classification and Its Consequences*. Cambridge, Massachusetts: The MIT Press.

Manuel Castells, 2001. *The Internet Galaxy: Reflections on the Internet, Business, and Society*. Oxford: Oxford University Press.

Claudio Ciborra (editor), 2000. *From Control to Drift: The Dynamics of Corporate Information Infrastructures.* Oxford: Oxford University Press.

Claudio Ciborra, 2007. "Digital technologies and risk: a critical review," In O. Hanseth, and C. Ciborra (editors). *Risk, Complexity and ICT*. Cheltenham, UK: Edward Elgar, pp. 23–45.

Richard J. Cox, 2007. "Machines in the archives: Technology and the coming transformation of archival reference," *First Monday*, volume 12, number 11 (November), at http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2029/1894, accessed 13 January 2010.

Karen Coyle, 2008. "Managing Sameness," *The Journal of Academic Librarianship* volume 34, number 5, pp. 452–53.

Paul J. DiMaggio and Walter W. Powell, 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields," *American Sociological Review*, volume 48, number 2, pp. 147–160.

Hubert L. Dreyfus, 1991. *Being-in-the-World: A Commentary on Heidegger's Being and Time*. Cambridge, Massachusetts: The MIT Press.

Hubert L. Dreyfus and Charles Spinosa, 1997. "Highway bridges and feasts: Heidegger and Borgmann on how to affirm technology," *Man and World*, volume 30, pp. 159–177.

Hamid R. Ekbia, 2009. "Digital artifacts as quasi-objects: Qualification, mediation, and materiality," *Journal of the American Society for Information Science and Technology*, volume 60, number 12, pp. 2554–2566.

Elena Esposito, 2002. *Soziales Vergessen: Formen und Medien des Gedächtnisses der Gesellschaft*. Frankfurt am Main: Suhrkamp.

Michael P. Evans, 2007. "Analysing Google rankings through search engine optimization data," *Internet Research,* volume 17, number 1, pp. 21–37.

S. Fortunato, A. Flammini, F. Mencer, and A. Vespignani, 2006. "Topical interests and the mitigation of search engine bias," *Proceedings of the National Academy of Science of the United States of America*, volume 103, number 34, pp. 12684–12689.

Nelson Goodman, 1976. *Languages of Art.* Indianapolis: Hackett.

Heather Green, 2002. "A Library as Big as the World," at http://www.businessweek.com/technology/content/feb2002/tc20020228_1080.htm, accessed 13 January 2010.

Dan Greenstein, 2000. "DLF draft strategy and business plan," Digital Library Foundation, at http://www.diglib.org/about/strategic.htm, accessed 23 June 2009.

Jutta Haider and Olof Sundin, 2010. "Beyond the legacy of the Enlightenment? Online encyclopaedias as digital heterotopias," *First Monday*, volume 15, number 1–4 (January), at http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2744/2428, accessed 14 January 2010.

Birger Hjorland, 2000. "Documents, Memory Institutions and Information Science," *Journal of Documentation*, volume 56, number 1, pp. 27–41.

Beryl A. Howell, 2006. "Proving Web History: How to use the Internet Archive," *Journal of Internet Law*, volume 9, number 8, pp. 3–9.

Thomas P. Hughes, 2004. *Human-Built World: How to Think about Technology and Culture*. Chicago: The University of Chicago Press.

Nadine Höchstötter and Dirk Lewandowski, 2009. "What users see – Structures in search engine results pages," *Information Sciences,* volume 179, pp. 1796–1812.

Lucas D. Introna and Helen Nissenbaum, 2000. "Shaping the Web: Why the Politics of Search Engines Matters," *The Information Society*, volume 16, pp. 169–185.

Tim Jordan, 2009. "Hacking and power: Social and technological determinism in the digital

age," *First Monday*, volume 15, number 1–4 (January), at http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2417/2240, accessed 14 January 2010.

Jannis Kallinikos, 2006. *The Consequences of Information: Institutional Implications of Technological Change*. Cheltenham, UK: Edward Elgar.

Jannis Kallinikos, 2009a. "On the Computational Rendition of Reality: Artefacts and Human Agency," *Organization*, volume 16, number 2, pp. 183–202.

Jannis Kallinikos, 2009b. "The Making of Ephemeria: On the Shortening Life Spans of Information," *The International Journal of Interdisciplinary Social Sciences*, volume 4, number 3, pp. 227–236.

Jannis Kallinikos and Hans Hasselbladh, 2009. "Work, Control and Computation: Rethinking the Legacy of Neo-Institutionalism," *Research in the Sociology of Organizations*, volume 27, pp. 257–282.

Giovan Francesco Lanzara, 2009. "Reshaping Practice Across Media: Material Mediation, Medium Specificity and Practical Knowledge in Judicial Work," *Organization Studies*, volume 30, number 12, pp. 1369–1390.

Lawrence Lessig, 1999[2006]. *Code and other laws of cyberspace.* New York: Basic Books.

Peter Lyman and Brewster Kahle, 1998. "Archiving Digital Cultural Artifacts," *D-Lib Magazine*, volume July/August, at http://www.dlib.org/dlib/july98/07lyman.html, accessed 13 January 2010.

Ross A. Malaga 2008. "Worst Practices in Search Engine Optimization," *Communications of the ACM*, volume 51, number 12, pp. 147–150.

James G. March and Johan P. Olsen, 1989. *Rediscovering institutions: the organizational basis of politics*. New York: The Free Press.

Lev Manovich, 2001. *The Language of New Media*. Cambridge, Massachusetts: The MIT Press.

Deanna Marcum, 2003. "Requirements for the future digital library," *The Journal of Academic Librarianship*, volume 29, number 5, pp. 276–79.

Attila Marton, 2009a. "Self-Referential Technology and the Growth of Information. From Techniques to Technology to the Technology of Technology," Soziale Systeme, volume 15, number 1, pp. 137-159.

Attila Marton, 2009b. "Digital Libraries as Information Organizations. The Re-Unfolding of the Memory/Information Paradox," *European Conference on Information Systems*, Verona, Italy, at http://is2.lse.ac.uk/asp/aspecis/20090224.pdf, accessed 13 January 2010.

Peter Morville, 2005. *Ambient Findability: What We Find Changes Who We Become*. Sebastopol, California: O'Reilly Media.

Wanda, J. Orlikowski, 2000, "Using Technology and Constituting Structures: A Practice Lens for Studying Technology in Organizations', *Organization Science*, 11, number 4, pp. 404-28.

Neil Pollock and Robin Williams, 2009. *Software and Organisations: The biography of the enterprise-wide system or how SAP conquered the world*. London: Routledge.

Andrew Sayer, 2000. *Realism and Social Science*. London: SAGE Publications.

Jeffrey Schnapp, 2008. "Animating the archive," *First Monday*, volume 13, number 8 (August), at http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2218/2020, accessed 13 January 2010.

W. Richard Scott, 2001. *Institutions and Organizations*. 2nd ed. Thousand Oaks: SAGE.

Michael Seadle and Elke Greifeneder, 2008. "In archiving we trust: Results from a workshop at Humboldt University in Berlin," *First Monday*, volume 13, number 1 (January), at http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2089/1923, accessed 13 January 2010.

Ravi Sen, 2005. "Optimal Search Engine Marketing Strategy," *International Journal of Electronic Commerce*, volume 10, number 1, pp. 9–25.

Lee Shaker, 2006. "In Google we trust: Information integrity in the digital age," *First Monday*, volume 11, number 4 (April), at http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1320/1240, accessed 14 January 2010.

Sergio Sismondo, 1993, "Some Social Constructions", *Social Studies of Science*, volume 23, number 3, pp. 515-593.

Mike Thelwall and Liwen Vaughan, 2004. "A fair history of the Web? Examining country balance in the Internet Archive," *Library & Information Science Research*, volume 26, number 2, pp. 162-76.

Vivienne Waller, 2009. "The relationship between public libraries and Google: Too much information," *First Monday*, volume 14, number 9 (September), at http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2477/2279, accessed 14 January 2010.

David Weinberger, 2007. *Everything Is Miscellaneous: The Power of the New Digital Disorder*. New York: Times Books.

Jin Zhang and Alexandra Dimitro, 2005. "The impact of webpage content characteristics on webpage visibility in search engine results (Part I)," *Information Processing and Management*, volume 41, pp. 665–690.

Laura Zick, 2000. "The Work of Information Mediators: A Comparison of Librarians and Intelligent Software Agents," *First Monday*, volume 5, number 5 (May), at http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/748/65, accessed 14 January 2010.

Jonathan Zittrain, 2008. *The future of the Internet – and how to stop it*. Harrisonburg, Virginia: R. R. Donnelley.

Arthur P. Young, 1996. "Libraries and digital communication: Collision or convergence?" *Journal of Academic Librarianship*, volume 22, number 1, pp. 11–13.