

DISCOVERING THE ARCHIVE. THE CASES OF GOOGLE AND CANADAHUN¹

Zoltán Dragon

Introduction

A couple of years ago if someone wanted to load up Hungarian ebooks for a newly acquired e-reader, had to head to a buzzing online forum called CanadaHun. The online community was originally founded as a meeting place for Hungarian expats living in Canada destined to help them find resources coming from their home in Hungarian, publications, school textbooks, newspapers, magazines, and literature. One only had to register, get acquainted with members in a couple of sublets in the forum's diverse structure, and find the ebooks desired to read electronically. While CanadaHun was never intended to become the largest online provider of Hungarian literature, it certainly earned its reputation as a dynamically growing on-demand library—an illegal distributor out of the purview of Hungarian copyright law enforcement. As the forum's ebook section had started to grow, it interestingly became a new type of archive that had the potential to shape the literary canon of the day not only for Hungarians in Canada, but for Hungarians at home, as well. Taken together with the examples and practices of many torrent sites around the internet, CanadaHun stands as a grassroots example of how the logic of sharing economy is able to reform and restructure the idea and working of the archive as dictated by the digital platform and new forms of consumption. This essay is an attempt to focus on the long (and sometimes foreclosed) history of the techno-logic of the archive in philosophical and critical theory to better understand the present situation of the digital archive and its role in preserving and, indeed, shaping of our cultural heritage today.

The Archive: The Space of Memory

Ever since the proliferation of textual documents seen as the most manageable and feasible means of the transmission of cultural heritage across generations, the library came to signify not only the collection of human knowledge and culture, but also the intersection of different discourses. The shift to digital platforms and modes of production is not simply a reproduction or recontextualization of the issues inherited from earlier incarnations of archiving systems and their politics. With the new, digital agenda, the *techné* of our age has started to effectively shape the *epistémè* of our cultural inheritance and identities on many levels. Questions of digital authorship and virtual

¹ This essay is based on my "Digitized Spaces of Memories and Cultural Heritage: the Future of Archives" (*AMERICANA-E-Journal of American Studies in Hungary*, Vol. XI., No. 1., Spring 2015; <http://americanajournal.hu/vol11no1/dragon>). The text is reworked and republished here with the consent of the publisher.

property, of data ownership and cyberethics, of the role and rules of digital archiving technologies, and the preservation and enrichment of national cultural inheritance are more pressing issues than ever: it is the ontology of the archive that is at stake in the 21st century. There is a need to discuss the role and possibilities of the digital archive by combining data management logic (database and algorithms) and their integration in theories of culture (via the algorithmic turn as defined by William Uricchio (Uricchio, 2011, 25)). To fully grasp the reorganization of cultural memories today, on the one hand, the humanities need to critically assess the ways information technologies lend themselves to cultural analyses, and on the other, information technologies need to incorporate the critical view distilled from such relevant inquiries. Our versions of cultural heritages are no longer dependent merely on their critical dimensions, but rely heavily on the technologies that needed the digital shift both on a technological and on a cultural level.

We have reached an era where there is a steady, unstoppable shift from the relative material stability of the Gutenberg (i.e. textual) galaxy to the digital (i.e. hypertextual) realm in preserving cultural heritage. Yet we lack the general critical and theoretical, as well as corresponding technological toolset as to how the system of the archive should work and to comprehend how it could move from its closed, fortified, privileged and ultimately institutionalized, paper and media-based structure towards new, open, virtual and augmented, crowd-sourced, open access digital space that eventually defies institutionalization and does away with questions of medium-specificity in the virtual. For this move involves not merely a transmission from materiality to the virtual, but also from the medial to the interface-driven hypermedial, too.

To find potential avenues to channel this transition, on the one hand, we need to revisit some of the most decisive conceptualizations of the internet and digital culture—born also as a transitory project but yet embedded in medial discourse with the help of Vannevar Bush and, later on, Ted Nelson. On the other hand, we need to re-focus our study to establish a dialogue between computation and the philosophical, critical theory tradition. It is this dialogue that can lead the critical discourses on the question of the archive out of the impasse it has recently entered. For this to happen there is an urgent need to rethink the tenets and potentials of the archive as a cultural memory system that is unlike our usual institutions both on the level of politics and on mediality.

This endeavour puts us right in the middle of the archive debate with its ethical, ideological, and political dimension reaching into the terrain of the digital, algorithmic, and database world. To talk about the archive, one needs to enter one of the many relevant discourses that operate along massively different sets of rules. We thus need to consult philosophy or ethics to ponder certain types of questions, or select fields of information technology to locate specific problems and issues in programming, the interface, or the algorithms that drive the archives. Unfortunately, whichever track one opts for, the other will be left alone, as the two (or potentially more) discursive lines here do not intersect (not in a meaningful dialogic manner at least). In other words, the *episteme* and the *techné* do not inform each other in contemporary discourses on the questions posed by the digital archive; therefore, none of the participating voices is able to tackle overarching problems and issues that our societies face today

(and in the near future as well) in terms of archiving and accessing cultural heritage. What I propose here is to erase the demarcation lines between humanities and information technology, philosophy and software, critical theories and algorithms in order to lay the foundations of a potentially flexible and fruitful dialogue that leads to a more comprehensive view at the archive.

When pondering the question of cultural heritage or memory organized in a meaningful and systematic manner, we also need to consider whether there is only one way to do it, or there are myriads of avenues opened up by digitized storing capacities. But to answer—or to approach something like a potential answer to—it, one needs to ask the pressing question: what is an archive? Especially: what is an archive in a media saturated age in which content loses its material base, its tenor, as an outcome of massive digitalization? Then comes another question: what is an archive (or what are archives) for and how should we think about its future? What should it contain? Who is it for and who should control it?

Let us start from the beginning; in other words, let us try and archive this particular train of thought in relation to preserving cultural memories. Questions of memory, of intellectual heritage, and of archiving thought and knowledge probably go as far back in the history of humanity as the appearance of literacy which, according to László Ropolyi, can present the case for a kind of digital thinking before the invention of both electricity and the digital proper. He claims that writing is essentially a non-binary, non-electronic digital literacy that later, in the second half of the twentieth century, with the birth of and transition to the digital, turns more and more into a binary, electronic digital literacy (Ropolyi 2014, 12). Curiously, this claim finds parallel with how the issue of mnemotechnics first appeared in philosophical thought only being foreclosed thereafter until the second half of the twentieth century saw a heightened increase in the issue again. As Simon Critchley explains, “Western philosophy arguably begins with the distinction of *techné* and *epistémē*, where Socrates distinguishes himself from the Sophists by showing how the pursuit of knowledge in critical dialogue is superior to and independent of sophistic rhetoric and mnemotechnics” (Critchley 2009, 173). The *techné*, as Critchley argues, comes to be repressed in philosophy and it resurfaces only following Heidegger – and this is where Jacques Derrida, in his insistence on the technicity of memory takes his lead from in his analysis of the Freudian inspired concept of the archive.

Derrida’s analysis of Sigmund Freud’s psychoanalytic ideas in *The Archive Fever* is transformative in many ways: first, it starts out from the subjective realm and enlarges on the issue of memory and the preservation of memory towards a collectively useful and feasible model. Secondly, it uses a logic and terminology that is reminiscent of how contemporary new media theory (especially following the work of Lev Manovich) conceptualizes issues related to how digital composition and data retrieval (two of the major roles of the archive) works in a contemporary, interface-driven media landscape.

As is usual with a deconstructive reading, Derrida sets off by identifying the hidden traces in the etymology of the word *archive*: commencement and commandment, which brings him to claim that the archive is essentially based on a “nomological principle” where the two segments are coordinated (Derrida 1995, 9). It then involves a

place, a special spatial setting, and a topology that governs this special intersection of law and privilege (10): an institution which harbours selected documents and saves them for posterity by exerting its own logic in the method of preservation. At this point, Derrida's language seems to refer to the material, physically established archive, that houses real documents, but it might not be too difficult to dream of an electronic archive whose spatial setting designates a non-place—a virtual entity served by redundant clusters of hardware on the technical side that can, perhaps, go beyond the strict nomological principle and open its boundaries to some extent. This would mean the commencement of a new commandment which might be seen as the implementation of the algorithmic turn, that is, the idea that contemporary representations of virtually any kind are prone to algorithmic manipulation as part of the new media scene of production and consumption (Uricchio 2011, 25).

This train of thought gets its commencement from Derrida's relegating archival storage to the outside, to the external substrate (Derrida 1995, 12), which commands him to differentiate between two versions of memory as a consequence: *mnēmē* or *anamnēsis*, and *hypomnēma* (14). It is this latter, the *hypomnēma* that is connected to the logic of the *arkhé*, which tells us that the very memory of the archive should be foreclosed, i.e. should be external to it. "The archive is hypomnesic" (14):

Because the archive, if this word or this figure can be stabilized so as to take on a signification, will never be either memory or anamnesis as spontaneous, alive and internal experience. On the contrary: the archive takes place at the place of originary and structural breakdown of the said memory. (14)

Derrida here still thinks of the archive as an institution, a physical establishment, but some passages later, by referring to his reading of Sigmund Freud's psychic archive and problematizing the machinery of archiving logic he ponders the question: "Do these new archival machines change anything?" (15). Later on, his answer is very close to the common trait in the potential answers one could come up with in terms of digital technology today, as it is commonly agreed that "the technical structure of the archiving archive also determines the structure of the archivable content even in its very coming into existence and in its relationship to the future" (17). Its ultimate consequence is that the archive "produces as much as it records the event" (17). In other words, the *techné* puts its constraints on the *epistēmē* it wishes to conserve—further emphasized by Derrida in his insistence on including his computer and its role in the creation of his text within the very text he is writing (archiving in a way)—by claiming that the technical potentials and boundaries do have a well-documented effect in and on the document that eventually becomes an archive of his thoughts.

While it is certainly strange to have specific computer references in a philosophical reading of psychoanalytical thought, this can be the very act of reinscription of the technological into our discourse on cultural memory. The fact that this memory is redefined in terms of "a certain hypomnesic and prosthetic experience of the technical substrate" (22) is a clear reference to the concept and working mechanism of the digital and dynamic model of the archive. This approach fits Derrida's insistence on the role of the archive. He writes: rather than being "a thing of the past," it "should call into question the coming of the future" (26; *emphasis in the original*). Because of the technicality, or the command of the *techné*, archival processes need to work for the future

(for compatibility and for accessibility, among other aspects) instead of modelling themselves on old paradigms (when the *episteme* foreclosed issues of the *techné* even in this regard). Furthermore, the future is not merely an aim, a trajectory shaped by the *techné*, but the realm into which the working of the archive is deferred. Accordingly, the question of archiving and that of the archive *per se* “is a question of the future, the question of the future itself, the question of a response, of a promise and of a responsibility for tomorrow. In short, regarding the archive, if we want to know what this will have meant, we will only know in the times to come” (27).

While some assumptions, especially the deferment to the future, seem highly figurative in discursive terms, there is a very strong technical subtext to Derrida’s conceptualization here. Contemporary archiving is most concerned about the *format* of the digital documents that need to be preserved and there are competing formats supported by different software (and companies, of course) that promise maximal compatibility for future retrieval technologies, yet we do not know which one of these will eventually prove to be consistent enough and widely supported (not only by archivists, archiving policies, but by the user community all around the world) to be the “future proof.” If we widen the scope of the archive beyond written documents (which we obviously should), it is not even evident that we shall have to work with document types that are prevalent formats today: new medial representation forms, technical frames, modes of knowledge transfer, etc. could pop up any day and conquer communication methods the archive should—at least conceptually and in basic technical terms—be ready for all kinds of changes to operate precisely like an archive as Derrida describes it.

I tried to stress the technical relevance of Derrida’s reading of the archive here, but this side may not be so apparent if we read some of the criticism on his discussion. Wolfgang Ernst, for instance, argues that what the philosophical tradition (meaning mainly the continental, French tradition, with Michel Foucault and Derrida taking lead) does to the issue of the archive is nothing but relegating it to the realm of empty metaphors, thus inflating the concept (Ernst 2008, 107). As opposed to what the “metaphorists” say, Ernst proposes an “elaboration of theoretical ideas in digital humanities in medium-specific ways” (Parikka in Ernst 2013, 2), in the sense that he advocates an approach that is informed by specific material, institutional, political, and medial tenets. It is from this position, he criticizes Derrida and Foucault: the former “virtualizes” the concept of the archive, while the latter seems to forget about the complexity of media technology and infrastructure altogether (Ernst 2008, 113). Moreover, while acknowledging that today the archive is driven by a first and foremost “cybernetic” logic (111), Ernst refuses to theorize or comment upon the shift of paradigm in terms of the analogue to digital transition, talking about the system of the archive defined as an apparatus that has both ideological and material dimensions (114). The cybernetic logic of the archive is located in the material foundations of the institution, as it is based on the connection of memory and storage as hardware spaces. To follow this argument one needs to stick to the strictly medial sense of the archive: the different documents and formats retain their characteristics in an analogue model of categorized, catalogued, “shelved” organization structure in which being archived means presence prolonged. While Ernst does not dismiss the growing signifi-

cance of the internet and its archival potentials, he nonetheless argues that the emphasis on this technology, which is a “transarchival” approach, “dissimulates the ongoing existence of material memory agencies, both hardware and institutions, which still govern the power of deciding what can be stored legally and technically” (Ernst 2013, 97).

Ernst’s approach, rooted in media technology, has not put the metaphoric rhetoric off the point, however: Derrida’s insistence on the hypomnemic nature of the archive is an unambiguous reference to the *techné*. What Ernst’s reading refuses, but Derrida’s thought embraces is precisely what George P. Landow noticed in the development of metaphorical rhetorics in literary theory and philosophy simultaneously with discourses on computation:

When designers of computer software examine the pages of *Glas* and *Of Grammatology*, they encounter a digitalized, hypertextual Derrida; and when literary theorists examine *Literary Machines*, they encounter a deconstructionist or post-structuralist Nelson. These shocks of recognition can occur because over the past several decades literary theory and computer hypertext, apparently unconnected areas of inquiry, have increasingly converged. Statements by theorists concerned with literature, like those by theorists concerned with computing, show a remarkable convergence. (Landow 2006, 1)

What Landow describes here is applicable for the internet as well, not only computer software; after all, the internet is run by software of many kinds. This approach resonates with the idea of transcoding, as defined by Lev Manovich. This is the most important principle out of those five that together characterize what a new media object is as opposed to an old, analogue media object (Manovich 2001, 45–46). Transcoding does not refer to the digitalization process of analogue formats and objects: it rather describes the dialogic interaction of what Manovich calls the “cultural layer” and the “computer layer” (46). This ongoing interaction between realms of cultural objects along with their interpretative framework and their computational metadata as used by the computer forms the basis for the shift from medial representation to the logic of the interface (277). The interface is no longer connected to any specific medium: it is a new, flexible, digital delivery format that is capable of organizing and managing previously separate media formats and the users’ interaction with them.

I think that with the current emphasis on the interface and on the practice of transcoding as cultural and digital exchange, we need to free the idea of the archive from its techno-determinist anchorage—the more so as the archive is not for the past but for the future—and it is still unclear what technological model is going to be the most proper and suitable solution to storage and retrieval in ten or fifty years. In other words, while today I claim that it is the *techné* that determines our *epistémè* and the discourses surrounding the issue of the archive, it would clearly be a mistake to take the *techné* of our era for granted when thinking about new ways of archival *epistémès* of the future.

Erasure as Formation of the Archive: The Right to be Forgotten

To talk about the archive in a contemporary, digital cultural setting, we need to re-think it in a manner that reshapes the conceptual basis modelled on the traditional library/catalogue setting. If we accept the proposition that the Google is far from being *the* internet, but it might be technically (and culturally) seen as an operating,

forming, dynamic archive, we should also assess the ethical-political dimension of its working mechanism. While probably the best known archiving projects are proper archivist jobs at Google—the Google Books scanning project or the beautifully executed Google Art gallery—perhaps the truly archivist dimension of the company is lesser discussed in these terms. For what Google is doing as a primary operation, i.e. its practice of indexing websites and different formats of content via its bots is a meticulous, neverending archiving project *per se*. However, it should be clear that what Google sees of the entire internet is but one percent of what is thought to comprise its totality.

Yet, most of the legislative procedures concerning human rights, rights to privacy, and of course, copyrights cases leads to the deletion of certain metadata and indices from the vast collection that Google works on. This cumulated to the ruling of the European Court on 13 May, 2014, which coined the term “The Right to Be Forgotten” (C-131/12), arguing that “individuals have the right—under certain conditions—to ask search engines to remove links with personal information about them. This applies where information is inaccurate, inadequate, irrelevant or excessive for the purposes of the data processing” (“Factsheet on the ‘Right to Be Forgotten’ ruling”). In 2010, a Spanish citizen lodged a complaint against a Spanish newspaper with the national Data Protection Agency and against Google Spain and Google, Inc. The citizen complained that an auction notice of his repossessed home on Google’s search results infringed his privacy rights because the proceedings concerning him had been fully resolved for a number of years and hence the reference to these was entirely irrelevant. He requested, first, that the newspaper be required either to remove or alter the pages in question so that the personal data relating to him no longer appeared; and second, that Google Spain or Google Inc. be required to remove the personal data relating to him, so that it no longer appeared in the search results. (“Factsheet on the ‘Right to Be Forgotten’ ruling”)

Commenting on the decision, Viviane Reding welcomed the urge to tighten privacy and data protection legislation—a view that raises more questions than settling some. This case puts forward questions about the relation of data, or content, and archiving processes: the court decision makes us uncertain about the legitimacy of connecting archived data to the particular subject, the authorship of records and the rights and mandates it forges, and finally the archivist project, the process of archiving itself. Whose property is the archive—and what kind of an archive is in question at all?

Database + Algorithm = Archive

While it is clear that the EU decision concerns the World Wide Web, it is important to emphasize that it goes beyond that layer of the internet and directly questions the position of the archive—not merely from an information technological aspect, but in terms of cultural memory, history, and the subject. Those in favour of the decision hail the “right to be forgotten,” claiming the victory of individual sovereignty. However, what this case brings to the forefront is not a spontaneous gesture of forgetting, but a forceful attempt at repressing something embarrassing or simply outdated record concerning an individual. Is this not a replay of the infamous case of Paul de Man, whose personal archive spat out the forgotten traces of a young pro-Nazi com-

menter (see Ernst 2008, 150–154)? Seen in this light, would de Man be right to ask Google to “forget” his past to mask his youth? And once digital technology is involved: is there a possibility to implement the right to be forgotten the way EU legislation envisages it? In other words, can digital mnemonic traces, algorithmic mnemotechnics deleted or altered in a way that the record in question leaves no trace?

The answer to that should be an absolute no. No in terms of the philosophical aspect, and no in terms of the technological side. The digital archive, in this context, is nothing else but an indexed database that serves as a pool of data from which algorithms carry pieces of information or records to the site of the human-computer interface (in the above case, to the search result page of a Google query). It means that the data (or collection of records), the search algorithms, and the interface that appears for the user are different levels of the archival constellation, where these technically separate levels can be programmed differently and behave according to different logics. This is the trick of the archive in the age of the internet: while an EU Court decision can make Google erase the hyperlinks pointing at the unwanted content, it obviously does not mean that the references should completely disappear from either the database or the algorithms at work. On the contrary: the hyperlinks do not show up on Google’s interface precisely because the algorithm is taught to foreclose it; in other words, there needs to be a record kept of the specific content or data, and the order to evade its listing.

This is the logic of the trace in Derridean thinking: “[t]he word ‘trace’ is a metaphor for the effect of the opposite concept, which is no longer present but has left its mark on the concept we are now considering” (Balkin 1998a). This is how deconstructive logic proceeds, as it untangles the repressed, forgotten, or erased concepts in another (usually opposite) one, and lays their codependence bare to investigate (Derrida 1976, 46–47). The lack (the apparent non-presence, the spectral quality) of the missing link the record presents in the database of the archive is thus more emphatically remembered than ever before: the algorithm needs to identify its trace in order to effectively put it under erasure (*sous rature*). The enforcement of cases based on the right to be forgotten are examples of what in Derridean parlance should be thought of *l’avenir*: the shadowy, phantomatic, blurred figure or image of the Other intervenes in unexpected places and domains to unthinkable ends. Especially so if one is to assess the rationale behind the false pretense of “being forgotten:” while in legislative terms the well-founded request would mean that fellow netizens would not be able to locate data that the plaintiff would not want them to see, in practice it means that a shell is thought to protect the information or data posted on the net, and it is calculated to be enough to erase the very hypertextual trace of (in effect, the link to) the contested content to protect an individual’s right to privacy. However, this only means that the act of erasing, or in Derridean terms to put under erasure (*sous rature*) is a *trace* born, a mnemonic signpost right at the center of what should be foreclosed from the system. As is clear in Derrida (and in the working mechanism of digital computing), the tricky thing with terms under erasure is that they do not simply disappear, or get out of the system. Rather, they signify the position of “inadequate yet necessary” (Sarup 1999, 33) status.



One of the tenets of legal interpretation and argumentation for the right to be forgotten is apparently that Google is the one to create and to manage data that might be prone to be erased from the list of searches; even if the text of the Court decision is not overtly specific in this regard, as it designates search engines (in the plural) to be “controllers” of personal data (Suhajda 2014). While other search engines are clearly left alone to produce search results where links to the contested contents appear, it is also problematic to control data which is not in one’s possession. It should then be(come) evident from the working mechanism of Google that they do not control or manage content on the different websites where they appear—they only harvest metadata and *map* the content that is produced and served elsewhere. Hence the primary slippage of the legal claim in terms of technicalities: while some legal professionals (like Zoltán Ormós in Hungary; Suhajda 2014) proclaim that if something cannot be found with the help of a Google search, it certainly does not exist, reality seems to know otherwise. Firstly, it is only the accessible part of World Wide Web that the search engine can visit and map, which is a shockingly small portion of all the sites, documents and other objects present on the internet; secondly, just because one search engine does not map a site on the net, another one still can.

Google equaling the Other in the Lacanian sense is thus a fundamentally false idea. It is rather Google’s Other, the construed, spectral dimension or image of the slippery and volatile totality of the internet that should be addressed Google equals the Other in the Lacanian sense a legal impossibility of course, though the only true ethical option. It is the Other’s gaze that is responsible for the presence of objects on the internet: to turn the legal argumentation inside out one might claim that once something is put on the net (any corner of the net, not necessarily the widely used protocol of the World Wide Web), even Google might be able to “see” it—and there is no way that object (document, image or any content, for that matter) can be erased. It is evident that technically speaking if a content disappears, the links (the mnemonic, or rather hypomnemic traces) will remember it for us—and vice versa, if the links are removed, the object will turn into a digital memorabilia of itself. In other words, it is not up to Google (or any other search engine) if a content is available or not on the internet.

But what is this vast digital network that can be thought of as the Other of Google? None other than the momentary sum of computers connected in decentralised network organized by nodal points that provide continuous information flow among the peers of the system. It escapes categorization or finality, changing by the milliseconds, deferring its completion endlessly, playing on the potentials of the practice of Derridean *différance* in a technological context. It is precisely in this shoreless pool that forms the rumbling underside of the dynamic, digital archive (or rather, potentially, archives, in the plural). Ever since the paradigm of Web 2.0, it should be clear that it is on this amoeba-like, ever-changing, ever-growing document space that means the future of the archives for the new generations. This virtual space that is governed and managed by digital interfaces is by no means a unified system but a dialogic space of different segments of culture connected and nurtured by hypermedial networks in real time (rather than retrograde as is characteristic of traditional, curated archives). This type of dynamism cannot be regulated: it is technically not possible

because only some of its nodes and their corresponding interfaces might or can be muted or controlled.

The González vs Google case raises further questions, as well. To finish the legal case study, one has to look at the curiously vague and non-liberating, essentially self-contradictory subclauses of the ruling of the EU Court. While González asked the Court to make all media completely erase the data in question, the deletion part concerns only Google's search algorithm: nor the original publisher of González's story, *La Vanguardia*, the printed newspaper, neither its online edition was made to perform the erasing. The court's claim was that the newspaper and its online version had all the right to publish the problematic content, so there is no legal way of forcing them to withdraw it. That is, Google has no right whatsoever to direct searches on González to that rightfully communicated content. In other words, the archivist of the data should deny the existence of the very data, while the publisher of the same data can continue to serve that information for anyone interested (searching for it in alternative search engines).

The moral of this legal intervention is perhaps that one may think of the archival database in terms of being the vessel of our *episteme*, or that Google is *par excellence* the interface for all the cultural products that enter the digital realm as a result of participatory culture. However, this approach neglects the technicity of the underlying network, as it is precisely this *techné* that, at the present, gives a framework for that *episteme* and not vice versa. For that matter, any legal intervention in terms of data protection and the archive should take its start from this pretext. And, finally, apart from the ethical, political, economical and legal dimensions of the above case, one may ponder the question what the right to be forgotten means in terms of an ever-growing archive.

Crowdsourcing the Archive: Archive-on-Demand

As opposed to the traditional concept of the archive, rooted in the paradigm of the physical library, new media technologies have brought about a new generation of systematic memory building possibilities. While World Wide Web interfaces to immense object databases like Europeana.eu (which is trying to collect a vaguely defined European cultural heritage), or in Hungary, the Manda (Hungarian Digital Archives, aiming to make Hungarian cultural heritage freely accessible for everyone, a claim left unfulfilled, and put under erasure; in 2016, the Hungarian government decided to restructure and reorganize the public factor and basically finish Manda's operation in its present form), alongside university and scholarly repositories continue to build on the tradition of strict and meticulous categorization and shelving policies that they hope to transfer to the digital realm. But these systems and policies essentially contradict the logic of memory in digital computing. The library paradigm is a closed, ideologically constrained, centralized and hierarchically structured unit; the new media paradigm is open, without constraints (as it is dynamic), is decentralized and—because of the hypertextual logic of linking—is flat, in structure.

If the library is the place and logic of the archive, it can also be regarded as the space of the canon. However, with the advent of digital archiving technologies, traditional libraries cannot keep the pace with the ongoing reformations of the various canons, literary or scientific. Curiously, the step beyond Google's quasi-archivist efforts

might be seen as the future for the archive in general and in particular. In general, as the completely new model of collecting and preserving cultural knowledge takes its cue from that intersection of theory and digital computing Landow noticed earlier; in particular, as the digital, algorithmic logic behind the platform creates a futuristic, yet contemporary model deriving from a special mix of hypertextual logic and open access policies.

CanadaHun presents in this context a curious example: it had not been intended as a library or archive. It was merely a place for forging expat identity through the internet, perhaps the most viable and easily accessible platform for communication and connecting with the culture of the country of origin. Technologically speaking, CanadaHun remains what it had attempted to be: a forum, a relatively open space where items of commerce can be exchanged, and not an archiving engine. Indeed, there is a network system, a technology that is more adept at doing the job of community archiving and sharing effectively without reproducing the system and catalogue structure of the traditional archive. The technology in question is the torrent, a type of P2P (peer-to-peer) file sharing solution.

Thought to be the facilitator of copyright infringement and all the evil that there is on the internet, the torrent is one of the most efficient sharing and storing technology. It utilizes the file sharing users' connections and creates a loose, participatory subnetwork of digital objects that can be shared faster than any other technology at the present could do. It also is capable of acting like an archive proper; from its start up to 2012, Library.nu in fact served as *the* academic archive and library for a majority of Central and Eastern European scholars, when university libraries drastically cut their subscriptions to and acquisitions of scholarly journals and monographs in the region. The lack of current academic literature in the libraries soon gave way to an abundance of the torrent server's offering and actually began to transform the academic canon for some disciplines. Although Library.nu was taken down in 2012, its legacy (and technology) was taken over by others, and by now various reincarnations serve eager readers with a million academic books, a million of literary works, and almost ten million academic studies and articles (Bodó 2014).

The way the internet copyright researcher Balázs Bodó traces the development of not only the technology but the activists behind these new, pirate-type archives, shows us a new model for the digital dynamic archive: a crowdsourced archive, without (virtual or material) walls, working on demand, with scanning and making those pieces of academic literature available for sharing which most users wish to read. Also, volunteers may contribute and enrich the database so that potentially all fields of academic research get the latest scholarly input. The system at first looks as if it were a somewhat structured anarchy or a mass of lazily categorized collection of traces of cultural objects, yet the technology behind is clearly one that is reminiscent of the dream of hypertext and docuverse put forth by Ted Nelson in the 1960s, or that of the Derridean system of nodes, traces, and *différences*. As Landow notes on Derrida's vision of a quasi hypertextual network, it is described by the terms "*link (liaisons), web (toile), network (réseau), and interwoven (s'y tissent)*" (Landow 2006, 53; *emphasis in the original*): the terms get their most perfect technological realizations in terms of the torrent logic.

The formation of a new type of canon through the use and flourishing of the new archive-on-demand is now imminent: most of the students in Hungary working with academic sources in English use the latest releases in their disciplines to the amazement of supervisors—especially those who still try to convince their librarians to acquire at least one of the myriads of important new releases for which the library has less and less financial backing. As students, members of the new, digital native generation, turn toward this new type of archive not only because of its price advantage (it is free) but primarily because of its availability and ease of use, academic discourse starts to be shaped by the new and rumbling, developing, non-fixed canon of the dynamic digital archive.

Conclusion

The archive-on-demand, the most perfect rendition of dynamic digital archives ever envisioned by philosophers and new media theorists is on its way to become the new cultural software (Balkin 1998b, 5) for our societies in preserving and, perhaps even more importantly, sharing cultural heritage. Based on principles in the intersection of Derridean philosophy, new media theory, and computer science, this new format of the cultural software comes with a new ideology, a new know-how, and an ultimately new model that defies almost all elements of traditional document archives and the model of the library. Whether it is Google-type of metadata harvesting or Library.nu's (and its heirs'), or CanadaHun's slightly different logic of crowdsourcing and on demand serving, the new archive is no longer just a dream. It is here to shape our present to prepare us for the future—but first we need to unthink the archive to build the archive for the future and not for an unfortunate *l'avenir*.

This issue gets more pressing by the day with the appearance—and disappearance—of new forms of cultural production that defy the old systems of archival practice and require more flexible, dynamic registration in cultural memory—for it is not the content and its structural pattern that should be adjusted to the requirements of an archival framework in the digital era, but vice versa, the archive should be dynamic enough to be able to operate in the future without disruption. Today's techne is in formation and is effectively shaping future document formats. Now it is the turn of our archival *episteme* to take the step toward the new digital paradigm and embrace its logic, as it is only in the intersection of these two dimensions that our archive for the future could be secured.

Works Cited

- Balkin, J. M. (1998a) *Deconstructive Practice and Legal*. Yale University. <http://www.yale.edu/lawweb/jbalkin/articles/decprac1.htm>. Accessed: 24 February, 2015.
- Balkin, J. M. (1998b) *Cultural Software*. New Haven: Yale University Press.
- Bodó, B. (2014) "Kalózok mentik meg az irodalmat". *Könyves blog*, November 11, 2014. http://konyves.blog.hu/2014/11/11/kalozok_mentik_meg_az_irodalmat Accessed: 27 February, 2015.
- Critchley, S. (2009) *Ethics-Politics-Subjectivity. Essays on Derrida, Levinas, and Contemporary French Thought*. London: Verso.
- Derrida, J. (1995) "Archive Fever. A Freudian Impression." *Diacritics*, Vol. 25, No. 2 (Summer, 1995), 9–63.

- Derrida, J. (1976) *Of Grammatology*. (Trans. Gayatri Chakravorty Spivak). Baltimore: The Johns Hopkins University Press.
- Ernst, W. (2008) *Archívumok morajlása. Rend a rendtelenségéből*. [Das Rumoren der Archive. Ordnung und Unordnung] (Trans. Tamás Lénárt). Budapest: Kijárat Kiadó.
- Ernst, W. (2013) *Digital Memory and the Archive*. (Ed. by Jussi Parikka). Minneapolis: University of Minnesota Press.
- "Factsheet on the 'Right to Be Forgotten' Ruling" (C-131/12). (2014). European Commission. http://ec.europa.eu/justice/data-protection/files/factsheets/factsheet_data_protection_en.pdf Accessed: 27 February, 2015.
- Landow, G. P. (2006) *Hypertext 3.0. Critical Theory and New Media in the Era of Globalization*. Baltimore: The Johns Hopkins University Press
- Manovich, L. (2011) *The Language of New Media*. Cambridge, MA: MIT Press.
- Ropolyi, L. (2014) "Digitális írásbeliségek." *Korunk*, XXV/10, October 2014, 8–14.
- Sarup, M. (1993) *An Introductory Guide to Post-structuralism and Postmodernism*. Athens, Georgia: The University of Georgia Press
- Suhajda, Z. (2014) "Felejtésre kényszerül a Google". *Metropol*, May 22, 2014. <http://www.metropol.hu/cikk/1187475-felejtésre-kenyszerul-a-google>. Accessed: February 27, 2015.
- Uricchio, W. (2011) "The Algorithmic Turn: Photosynth, Augmented Reality and the Changing Implications of the Image." *Visual Studies*, Vol. 26., No. 1., March 2011.