Digital Heritage and Heritagization

Francesca Musiani and Valérie Schafer
In 2016, the celebration of the twentieth anniversary of the Internet Archive foundation brings to our attention an early initiative of digital preservation and heritagization: no later than five years after the birth of the World Wide Web, and at a moment when it had not yet conquered the general public, Internet Archive founder Brewster Kahle was already thinking about its global archiving (Mussou, 2012; Masanès, 2006).

This initiative can be connected to the preoccupation of Internet pioneers, academics and engineers first and foremost, to preserve the traces of their activities, technical ones in particular. Since 1969 and the first steps taken by the ‘network of networks’, Requests for Comments (RFCs, technical specification and standardization documents published by the Internet Engineering Task Force) are the prominent manifestation of this shaping of technical memory. Indeed, as Alexandre Serres reminds us, the Internet was at its origins largely self-referential, and “the network talked about the network” (Serres, 2000).

Brewster Kahle’s initiative is also a reaction to what happened in the case of other media: for example, he cites in 1997 the destruction of movies to recycle the silver contained in early film (Kahle, 1997). Thus, it is the sentiment of volatility and transience of born-digital content that, to a large extent, originates the first initiatives (Lyman & Kahle, 1998).

Some researchers and academics, as well as archivists and librarians, are also quick to acknowledge the important questions raised by the digital turn in terms of preservation. It is the case, for example, of precocious initiatives by Scandinavian countries in the second half of the 1990s, to extend the perimeter of legal deposit, or of the Austrian project AOLA, launched in the early 2000s to develop the archiving of the Austrian Web (AOLA, 2001). All these projects are a reflection of the evolutions experienced by heritage during the last decades. As put by Pierre Nora in the same year the Internet Archive foundation was born: We shifted from a State- and nation-promoted heritage to a society- and community-promoted one, where group identity unfolds; thus, from an inherited
heritage to a claimed heritage. From material and visible, heritage became invisible and symbolic [...] heritage left its historical age to enter its memorial age: ours (Nora, 1996, our translation).

Twenty years after this acknowledgment and the first steps in Web archiving, several initiatives have helped qualify a number of born-digital documents as heritage. Among them are the Archive Team’s rescue of several online communities (e.g. Geocities and Mobileme), the institutional archival of the Web by the National Library (Bibliothèque nationale de France, BnF) and the Audiovisual National Institute (Institut national de l’audiovisuel, INA) since 2006 in France within the frame of the legal deposit, or the archival of Twitter by the United States Library of Congress. They bear witness to the plurality of actors engaged in the process, of incentives and motivations that are often complementary, and especially of an unprecedented documentary proliferation. More than 490 billion pages archived by Internet Archive over twenty years, or the 48.5 billion URL versions preserved by the INA as of July 2016, are the impressive answer to the linear archives of previous periods in time (Schafer & Thierry, 2015).

While it is necessary to take this unprecedented abundance into account (Dougherty et al., 2010), fragmentation, propagation and dematerialization phenomena entail an “epistemological reconfiguration for both the archivist and the researcher, who now think in terms of ‘resources’ and ‘data’ more than of documents” (Mussou, 2012: 264, our translation).

The definition of “digital heritage” promoted by UNESCO in its Charter for the Preservation of Digital Heritage (2003) designates both “digitized” and “born digital” resources (the latter being defined as resources for which “there is no other format but the digital object”). This dual meaning entails a reflection on digital heritage that is able to encompass continuity and long-term processes, as well as ruptures and breakdowns. Indeed, it calls for reflections about the ways in which digital resources can become heritage, but also about the ways in which heritage is ‘appropriated’ by the digital world. It therefore invites to pursue reflections, grounded in the social sciences, such as those on the ‘heritage question’ (Amougou, 2004) and what constitutes today the notion of heritage itself, as well as its ‘reinvention’ (Bourdin, 1984).

What do digital heritage (and the processes of heritagization) tell us about the relationship of our societies to their heritage, and to the digital itself? What happens when it is no longer enough to ‘digitize’ culture, but culture is, at once, digital? What are the consequences on heritage and its uses – social, cultural, political, economic and scientific? This issue presents six articles articulating several scales and spaces: from family memories to collective memory, from micro-history to World War I History, from scientific and technical heritage to memorial issues linked to the history of immigration, from the tensions between the right to be forgotten and the right to memory to those related to copyright and associated geopolitical issues. By means of these contributions, the issue seeks to shed light on the diversity of policies that shape, manage and develop digital heritage, as well as actors and issues underlying heritagization processes (Davallon, 2006).

**Digital heritage and social legacy**

The practice and processes that contribute to the shaping and legitimation of digital heritage imply choices, reconfigurations, tests – a social work. This work is conducted
by a variety of actors, and at several scales. Beyond the role of professionals, of amateurs, sometimes grouped into associations, of the public at large, libraries, museums and research groups assume responsibility of some archiving processes and preservation policies. Private companies, such as Google – the ‘giant’ gets involved in the preservation of some Usenet newsgroups, and created the Google Cultural Institute – are also part of this movement, as well as international institutions such as UNESCO or the International Internet Preservation Consortium (IIPC).

These different actors appear as the issue unfolds; sometimes, this happens in an isolated manner, as illustrated by one of the articles exploring a family-originated heritagization process. However, more often than not, actors are shown as they interact, complement each other, or enact their divergent or concurrent views about digital heritage. Thus, two articles in the issue demonstrate how memorial practices find their expression online, either by highlighting a still-conflictual and vital memory, or rallying around a more “soothed” memory – both endowed with a heavy political charge. Institutions and their heavy weight are balanced with more local, individual or community-led initiatives, as well as projects by professionals, of history, documentation, archive or communication, and by amateurs. The authors in this issue avoid what would be an artificial and blunt separation of memorial expressions from heritagization attempts, and from the reappropriation/diffusion of already constituted heritage; instead, they focus their (and our) attention on the role of new actors of heritagization too, be they associations, amateurs and/or activists. Their presence is not necessarily perceived in a positive light by historical actors of heritagization, cautious about the possible confusion between the memorable and the ‘remember-able’ (see Hoog in Coutant & Stenger, 2010), and wary of the illusory idea that the appropriation of heritage by the public would be enabled merely by making content available. Després-Lonnet underlined this aspect in 2009, when discussing the Joconde database, then the efforts towards making European cultural heritage available online: “We indulge in the confusion between the technical capacity of ‘making data accessible’ via their online presence, and true access to knowledge, i.e. the actual possibility given to an individual to appropriate new knowledge (Després-Lonnet, 2009, citing Miège, 1997: 46, our translation).

Thus, the tendencies described above invite us to interrogate our relationship to technology, temporality, individual and collective memory, oblivion and the past, as well as the risk of being too ‘presentist’ in our appreciation of history. In addition, they invite us to reconsider important questions already raised in the study of heritage, and demonstrate their relevance for the present. “Heritage has become an umbrella term, producing anaplasia; however it is, in several respects, dissociated from the relation to the past. By questioning the modes of representation of the past in procedures of heritagization, preservation, collection and mediation, we need in the first place to investigate heritage anew, as a mode of relation to the past and to the future” (Les patrimoines en recherche d’avenir symposium, 2015, our translation). Investigating these modes of relation to the past, the analyses included in this issue invite us to go back to the very notion of heritage, similar in this instance to the approaches proposed by Laurajane Smith and Brian Graham; they consider, respectively, that heritage can be considered as an engagement process, more than a condition (Smith, 2006), or as a communication vector, a way to transmit ideas and values, as well as a form of knowledge including material, intangible and virtual features (Graham, 2002: 1006).
The “making of” digital heritage

12 The necessity of taking into account the emergence and the development of digital heritage was first identified in 2003, enshrined in UNESCO’s Charter for the Preservation of Digital Heritage, and supported most notably by its Information and Communication division within the programme *Fostering equitable access to information and knowledge for development* and the sub-programme *Fostering actions to reduce digital divide and promote social inclusion* (UNESCO, 2003; Lusenet, 2007). This emerging field calls for research addressing issues and strategies of appropriation and legitimation of digital heritage, its perimeters of preservation, comparative studies of preservation policies and tools among different countries and/or actors (Schafer, Musiani & Borelli, 2016).

13 What is perhaps the most striking aspect revealed by this issue is the particular attention authors pay to the role of amateurs, citizens, initiatives born outside the walls of institutions traditionally dedicated to heritage. Their respective fieldworks, and their thorough observation of the “making of” digital heritage, help to shed light on an aspect that was still considered under-explored a few years ago:

Present virtual heritage projects are mostly focused either on ‘process’ or ‘product’ but rarely consider ‘users’ [...] for better interpretation and experience of a digital heritage site, a comprehensive interpretation method is required, which should address end-users with various background, overcome the linearity in narrative level and subjectiveness in content creation (Rahaman & Tan, 2011).

14 Digital heritage must be embedded in its long-term dynamics - the link of societies to their heritage - to interrogate the continuities (Oury, 2012), the disruptions, the documentarization, the transformation of “traces in communities of memory” and the “ways to account for the self and the ‘us’” (Le Deuff, 2010). However, it also shapes “archives of a novel kind, at once completely dependent upon traceability techniques (e.g. buttons, automatic imports, timelines, metadata) and perfectly arbitrary, thus, unpredictable” (Merzeau, 2012: 11, our translation).

15 Deserving our scholarly attention are also the ‘heritage’ value that communities (Derrot et al., 2012) may attribute, for example, to the preservation of discussion groups (such as Usenet or Geocities) and the willingness to transmit and ‘heritagize’ new forms of expression and communication. Beyond the heritage value of documents and the social status of objects (Davallon, 2006), it is appropriate to address the notion of immaterial heritage (Bortolotto, 2011; Jadé, 2012), and that of common good. Mélanie Dulong de Rosnay (2012:143-144) has shown, drawing on Elinor Ostrom’s work, how the notion of common good is at the core of “very structured collective creations, whose modes of governance and appropriation also belong to peers”.

16 As already stressed by Maurice Halbwachs (1950), every organized group creates a memory of its own, and individual memory is anchored in remembering and localization processes specific to collectives and communities. When the Archive Team preserves Geocities’ personal pages, it is actually protecting an ‘ordinary heritage’, directly stemmed from experience, individual and collective at once (Paloque-Berges & Schafer, 2015: 258), confronted to the threat of Yahoo!’s intended closing of the service in 2009.
Along these lines, the conversion of ‘user generated archives’ into a common good finds its conclusion in the institutional archiving of the Web itself (Merzeau, 2012).

Public memory, “[d]e-personalized and dissociated from the commercial paradigms surrounding and predicting us [...] re-creates here the link between an individual and a political community, in the open time of a collective reflexivity” (Merzeau, 2012). Thus, it can reconcile community-driven (Derrot et al., 2012) and institutional initiatives of different natures and perimeters. However, the articulation of these issues with a ‘data market’ in a large sense, which is currently particularly thriving, remains necessary.

Public initiatives should not lead us to forget that digital firms can put at risk of disappearance entire sectors of collective digital memory (such as Yahoo! and Geocities, previously mentioned). However, the private sector shows, in parallel, its important role in the preservation of digital data and traces. Examples of this are Twitter archives, the role of Google in protecting digitized heritage with Google Books and the Google Art project, or in preserving born-digital heritage such as newsgroups.

From browsing to sedimentation

If we wish to preserve technical contents, we should keep them as intact as possible, respecting their physical identity and perhaps restoring them, as we see in the classic museographic disciplines. But [...] we also need to transform them to enable their accessibility (Bachimont, 2008: 2, our translation).

The “born-digital heritage” we just mentioned has been underexplored by researchers as opposed to heritage converted into digitized form (Dufrène et al., 2013; Bachimont, 2014). However, these different facets of heritage converge in a process of digital heritagization, inviting the researcher to consider not only the making of heritage as human organization, but also the ‘hybrid agentivities’ (Abbate, 2012), both human and technical, at work throughout this process. These agentivities mix social and technical mediations in a multi-layered system prompting us to explore black boxes, especially those of Web archiving, and its modes of governance. Previous work has sought to demonstrate the diversity of actors engaged in heritagization processes and of their motivations, and the socio-technical devices they mobilize to serve their objectives. Digital heritage becomes the boundary object (Star and Griesemer, 1989; Bowker et al., 2016) reclaimed by multiple communities whose logics and aspirations are as diverse as those mobilized for the “living” Web (Schafer, Musiani & Borelli, 2016). The entanglement between human and technical processes that shape born-digital heritage can be perceived as a destabilization factor within the world of archives, of organizations or of research. However, it also suggests we should explore the “unorganized” and “raw” in the digital (Chabin, 2016) and find new usefulness in approaches derived from long-standing disciplines such as diplomacy (Chabin, 2011) or philology (Brügger, 2012).

Moreover, digitized as well as digital heritage are today inciting researchers and archiving actors to turn their attention on the tools promoted by digital humanities (Andrews, 2015; Le Deuff, 2014; Rogers, 2013), whether in the context of creation of a corpus, of its documentation or its exploitation. Several articles in this issue tackle the issue of metadata and indexability, at the heart of heritage digitization practices as well as born-digital heritage. Louise Merzeau underlined the importance of this issue a few years back: not only do metadata associated with heritage content describe it, but they...
enable their segmentation and recomposition. Thus, “by its tendency to auto-reference, starting from the moment it is emitted, digital information carries within itself its archival and accessibility” (Merzeau, 2012: 4, our translation). Today, the issues of accessibility and interoperability imply a variety of actors, from researchers and research engineers to preservation professionals and companies.

Finally, a reflection on another type of heritage which could be labelled as “heritage of the digital” (e.g. hardware, external devices, supporting documents) needs to be conducted to retrace those complex networked and digital environments where content is only the tip of the iceberg. This reflection appears as necessary as that conducted on content, especially as the two are rarely spearheaded by the same services or actors (with a few exceptions, such as the preservation of videogames and their manuals and the material devices supporting them).

**Presentation of the issue**

The six articles and the introduction composing this issue fully situate themselves within the interdisciplinary dimension of digital heritage analyses, including perspectives from history, information and communication sciences, sociology of innovation, digital humanities or juridical sciences. Authors bring us to different universes: Fannie Valois-Nadeau explores, with anthropological thoroughness, the online conversion and heritagization of Canadian professional hockey player Leo Gravelle’s family archives; Enrico Natale analyses the European dimension of World War I commemorations and the corpus associated with its analysis, while Sophie Gebeil addresses Web sites dedicated to the memory of Maghreb immigrants and immigration dynamics; Mélanie Dulong de Rosnay & Andrés Guadamuz and Rolf Weber & Lennart Chrobak introduce us respectively to the transnational questions posed by the right to be forgotten and by copyright. Furthermore, Camille Paloque-Berges examines the role of communities in heritagization, by analyzing networking pioneers and how they have left us a rich scientific and technical heritage via their online communication. While digitized heritage is at the core of Valois-Nadeau’s and Weber & Chrobak’s studies, born-digital heritage takes center stage in Gebeil’s, Natale’s and Paloque-Berges’s work.

Articulating reflections on data, traces, archives, heritage, history and memory, national and international levels, several articles merge scientific reflections with epistemological and methodological ones (it is especially the case in Paloque-Berges’ and Natale’s contributions) and enlighten us on evolutions in both heritage and digital processes.

While publics, uses or the economic dimension are less present in the analyses included in this issue, they appear discreetly, taking the shape of an “amateur culture” that leads to “creative reception” practices, built on the acquisition of innovative competences, forms of engagement and mobilizations (Flichy, 2010). Indeed, amateur initiatives and hybridization of amateur and professional practices – via acquisition, imitation, motivation – are often present throughout the issue.

Archiving and museum professionals, as well as researchers, will not however be absent from these different universes entangled in a common, albeit heterogeneous, world. In this respect, Dominique Boullier once suggested that politics of heritage and memory should be able to draw upon communities’ capacity to produce their memory – as does tradition –, on the capacity to revisit those memories and capitalize on them – as does
scientific activity –, and finally the capacity, and prerogative of media, to foster the emergence of new centers of interest and references (Boullier, 2008).

From the making of family heritage to the renewal of Maghreb immigration or World War I memories, through the scientific and technical heritage of pioneer Internet users, the study of forms, productions, practices, discourses and ideologies accompanying these initiatives enables us to more accurately apprehend the contours of a digital heritage – and a heritage of the digital – that is currently taking shape, as well as their social, political and cultural dimensions.

BIBLIOGRAPHY


COLLOQUE « LES PATRIMOINES EN RECHERCHE D’AVENIR », [En ligne], URL : http://colloquepatrimoines.passes-present.eu/


NOTES

1. International, informal and (in principle) open to every individual, the IETF is an organization communicating mostly online, authoring most Internet standards since its inception.
5. Expression derived from ‘user generated content’, frequently used in relation to Web 2.0.
6. In France, the institutional archive of the Web was programmed in 2006 by the DADVSI law, putting the BnF and the INA in charge of the legal deposit of online publications within national borders.
7. The AOLA project notes the existence of this expression since 2001: http://www ifs.tuwien.ac.at/~aola/publications/trans10
8. We think, for example, about the specific tools, functions and interfaces that the BnF and the INA developed to explore data and metadata in the corpuses constituted at the moment of the Parisian terrorist attacks of January and November 2015. We have followed these dynamics within the frame of the ASAP project (Archives Sauvegardes Attentats Paris), conducted in cooperation with the two archiving institutions and financed by the French National Research Council (CNRS), (https://asap.hypotheses.org, site visited September 21, 2016). Within the Web90 project (ANR-14-CE29-0012), we have also explored how the BnF has implemented full text research in Web archives of the 90s.