Westminster Papers in Communication and Culture © 2011 (University of Westminster, London), Vol. 8(2): 154-177. ISSN 1744-6708 (Print); 1744-6716 (Online)

Doing Media History in 2050

Gabriele Balbi
University of Lugano

Abstract
The evolution in both content and storage of information will change the way in which future media historians approach their task: in order to study the history of the mass media, telecommunications and new media of the late twentieth century, scholars will have to interact with digital sources, giving rise to a whole new series of concerns and questions. It is not clear if, how and in what format digital data will be available. What is the relationship between *old* and *new* sources? Who should be responsible for preserving the digital heritage? What should be preserved? How should digital data be preserved? All these questions are crucial for preservation strategies now and, above all, for studying media history in 2050.

Keywords: digitalization, future media historians, late twentieth- and early twenty-first-century communication, media history, preservation

> Now we can say we've been there.
> (Letter from Walter Benjamin to Theodor Adorno, 1940)

Digitalization is one of the most important phenomena to have influenced Western culture in the last decades (Jenkins, 2006; Grant and Wilkinson, 2009). 'Digitalization' is modifying culture and the preservation of our past in at least three different ways. First, there is the gradual and material digitalization of *old* analogue data: paper and magnetic material is being converted into digital format, multiplying, facilitating access to, converting and saving *old* data in new forms (Saksida, 1997). Second, our culture is producing a huge and continuously growing amount of 'born digital' data; contemporary working practices, contemporary communication, contemporary culture is based on the production, use, re-use and storage of binary data. Third, born analogue (and digitalized) and born digital data

are stored in special archives, called digital archives or digital libraries, that introduce a complex set of problems and challenges regarding their use, preservation and maintenance (Borgman, 1999).

The term 'digitalization' therefore also signifies a turning point in the way our culture is stored, recorded and preserved, and this alters and will increasingly alter culture itself, making it 'digital dependent' (Ross, 2000, 3). This cultural change affects a whole range of entities: national governments, private companies and broadcasters are investing money in digital archives, trying to preserve their past and make it available in the future. Future media historians will play an important role in this cultural change (and in this business); they will have to read, listen to, see, select, cut off and re-build our digital past; they will have to interact with digital sources to tell the story of the contemporary mass media, telecommunications and new media .

This article wonders how and why conservation of contemporary media production (especially digital data) will affect the ways in which media historians will study the history of communication in the late twentieth and early twenty-first century. Many crucial topics for media history will be discussed: the impacts of and the paradoxes in the storage and preservation of digital data; the organizations (private or public) that should preserve our mediatic present; which data should be preserved and why; and, finally, the best ways, if any, for preserving digital and analogue materials.

New Sources, Old Sources

Old sources for media history are analogue data that historians have used and will continue to use in the future. They are generally physical objects such as papers (documents, pictures) or magnetic material (films, audiocassettes, recorded radio and TV broadcasts). By *new* (historical) sources we mean digital data that will become digital sources in the near future.

Although old and new sources have defined characteristics to be considered here, their coexistence in contemporary reality should be remembered. First, the old does not disappear. As often happens in the history of communication, new media coexist with the old ones and the advent of digital media did not produce the

death of analogue. In other words printed newspaper, books, letters and broadcasts live together with e-books, SMS, email and any form of podcasting. Archives and analogue sources are still a central part of our present and, probably, they will be part of our future: publication on paper or on tape continues; libraries and archives with paper documents are and will continue to be relevant for humanity and for media historians too. Second, one of the major challenges of the present and probably of the future is the integration between old and new stores, but 'the greatest promise of the world wide web is not that it replaces the need to use the original material, but rather that it supports researchers in becoming aware of and locating archival materials to be used in situ' (Duff et al., 2004, 22). Thus, for media historians, the greatest potential of digital technology could be the ability to see and listen to old broadcasts at home, to receive Excel files with mobile telephony traffic at the end of the twentieth century, to interact remotely with archivists and curators of museums where the old media will be conserved.

Old sources have three distinct characteristics: scarcity, stability and accessibility. Before beginning field research and collecting analogue data, media historians (like all historians) know that they have to fully understand, search and 'burrow' into different collections and, if they are lucky, they find a few interesting documents for their research. *Old* sources, especially those preceding the eighteenth century, are generally scarce for many reasons: they are very old and they did not survive; there was no well-defined institution which had to preserve them; society did not produce a great amount of data (as the later mass society would do). For these reasons, one of the most important skills that a historian working with analogue data must have is a 'nose': he/she must have creativity in order to discover the required sources and, sometimes, he/she has to reconstruct hypotheses and ideas based on a very few data.

Media historians know something else: *old* sources are more stable than *new* ones. They are single, static pieces of data and, when an *old* source is recorded, it is frozen forever and it cannot change. They are stored according to established and defined archiving techniques, according to the paradigm of 'written archives' that was created hundred of years ago and has remained more or less the same ever since.

Another characteristic of *old* sources is their accessibility. In one sense, they are less accessible than digital ones because media historians have to travel, gain access (generally by negotiating with the archivists) and, finally, check if the document is actually in the collection. On the contrary they are also more accessible for two reasons. First, archivists help researchers in their work; they listen to their needs and thus create a kind of archive–researcher interaction, a human-machine that can help us understand and get to grips with the structure of the archive. Second, during their search, media historians may find unexpected documents, sources they were not searching for. This kind of serendipity is more difficult in *new* digital archives or, better, these archives have a different kind of serendipity. It is common when surfing the internet to find documents you were not searching for; but you tend to find documents related to the given keywords and depend on these to give you new ideas, new lines of research.

The digital revolution has completely changed these aspects typical of *old* sources. First, with the digital revolution, there is no longer a scarcity of sources. Indeed, this has now been replaced by an abundance of material and the contemporary digital age is characterized by the continuous production of data that could represent future historical sources. If truth be told, this is a pre-digital phenomenon, which began with the advent of radio and television that produce a never-ending flow of sounds and images, along with a series of problems for their archiving. Digitalization is radicalizing this tendency and propelling Western societies from a culture of scarcity to a culture of abundance (Rosenzweig, 2003).

Perhaps the greatest difference between *old* and *new* sources is their volatility (Schloman, 2003). We have said that *old* data and *old* historical sources are static while new digital materials are unstable, volatile and fragile. First of all, old sources have a physical unit: every paper, cassette, etc. is a single piece of a document linked to others in the same folder, but with its own materiality and unity. New digital data have no singularity: they are heterogeneous and fragmented; they are interrelated and fragmented because 'the user experience is not confined to a single object at a time, but is constructed out of multiple fragments from different sources and of different types' (Mackenzie Owen, 2007, 48). If you preserve a single Facebook profile without considering its connections

with, at least, all the friends' profiles, the posted links, and many other applications, you cannot understand the networked nature of Facebook, a medium of personality in which interrelation is the central character.

The volatile nature of digital data is not to do with their textuality but their instability, especially when published. For example, a website is a publication, but it is different from a paper one because it is undefined, it is a temporary intellectual product that could disappear and can always be changed. Indeed, the quality of a website is measured also through its continuous updating (Ortoleva, 1999). This is another reason why new media have been defined as underdetermined (Poster, 1999), always modifiable by their designers and, increasingly, also by their users. This unstable nature of digital data affects the way in which these materials can be stored and preserved. First, many documents can be and often are already lost forever (Kuny, 1997), without the possibility of ever finding them again; if a website changes its structure, content, pictures, *old digital* data disappear without any possibility of retrieval. This means that they will not be able to reach future media historians because they disappear forever and scholars will be unable to understand the web-style, language, cultural significance and historical development of websites and the internet. Second, the volatility and changeability of digital products forces historians and archivists to rethink their preservation philosophy. There is an intrinsic paradox involved here: we want to permanently preserve (and aim to study) objects and data that are ephemeral by nature – that were born to die quickly (Byerly, 2009). Indeed the specificity of digital data is interaction and the continuous intervention of users or what Henry Jenkins called the 'participatory culture'. They are fluid/dynamic objects that change over time because they are interactive and collaborative. For memory institutions it is extremely difficult to save this 'infinite heritage' because digital data 'are dynamic and always in progress rather than having a final and complete state' (Uricchio, 2007, 22). We cannot store and preserve them in a fixed manner and then study them in this new form because this fixed and frozen source is something completely different from the underdetermined datum. Indeed, what we should try to preserve are the uses and processes, not the materials and artefacts (MacKenzie Owen, 2007). For example, let us consider a page of

Wikipedia: the page can potentially always be changed by users with consequences for the present and future history (Rosenzweig, 2006) and, indeed, the power and popularity of this project is the fact that definitions can be refined, improved day by day by a 'collective (and connective) intelligence' (Lévy, 1997). So what Wikipedia page should we preserve? Which ones and at which point of the timeline will show future historians how a concept was considered in the past?

In *The Bias of Communication*, Harold Innis (1951) distinguished between durable and portable media and, paraphrasing him, we could suggest a history of stability. It seems that, in the evolution of communication history, data and documents have become increasingly volatile, passing from stone, to paper, to magnetic, to digital. These supports are evermore portable, inexpensive and easy to access, but can be more easily lost and erased and, above all, they are more difficult to store. This is the process of dematerialization and 'lightening' of culture, but it is also a process of structural weakening of culture, because the technical support has become more and more fragile and breakable. Stone, paper and magnetic are successively easier to damage. For example, you had to use brute force with stone or fire with paper, while a digital collection can be completely destroyed by damaging just one bit: just one letter or one unit of a repository could make the repository unusable.

Digitalization has also changed the availability of data and, compared to the past, sources appear more available but with an inherent risk. Online digital documents are accessible to a huge number of people and, in this sense, digitalization has democratized access to historical sources. With the internet, not only the source, but also historical papers, articles and books are digitalized and always available, anywhere in the world (Cohen, 2004, 294); you can download them – maybe with password access – when and where you need them and historians no longer have to spend money to physically reach the archives. Moreover, access to the source is direct, an old dream of historiography which has finally come true: accessing the past directly, without mediation.

On the other hand, the absence of mediation, generally represented by the archivist, risks leaving historians out on their own and, above all, without basic information. For media historians, the relationship with the archivist is

fundamental because he/she provides valuable information on the research object, on new and innovative keys of research, on the structure of the archive and, consequently, on the past archival mentality.

The changes in the materiality of sources have influenced and will influence the work of media historians in several respects (Manoff 2006). First, as already mentioned, digital sources will be more abundant than analogue ones; this means that researchers will have more difficulty in selecting *relevant* materials and in not being fetishists of the source. They will have to endeavour to maintain the critical distance between the object of research and the source. Second, due to the volatility and instability of digital data, significant sources may be lost to historians trying to understand the digital past of contemporary society. Losing some YouTube videos, for example, might mean losing a key element of the present communication environment (but the crucial question is always: which videos have to be saved and which not?). Third, considering that the most relevant characteristic of digital data is that they are continuously changed and updated, future media historians could find frozen digital documents that do not correspond to the original changeable and participatory ones. Finally, digital documents are more readily available and historians can easily find them and search them (e.g. with Google Scholar); on the other hand, the relationship with the archivist changes, giving rise to new methods, new forms and a new loneliness for media historians. How many historians have had (voluntary or involuntary) ideas and suggestions from archivists? Is it really possible to reduce or remove this figure?

Who Preserves?

Analogue data have usually been preserved by specific subjects that have two principal characteristics. First, they are structured at a national level and so they generally preserve documents useful for remembering the nation's past and for recreating national unity. Second, especially in the twentieth century, these subjects have either been public institutions or private companies. Public institutions preserve the past for three main reasons: first, because their aim is to have a monopoly over national representation; second, because national libraries, archives and museums have the specific public service function of preserving data;

finally, a public service corporation like the BBC 'is legally required to record its TV and radio output off-air to enable the corporation to answer complaints from the listening and viewing public' for 90 days and 42 days following the initial broadcast respectively (C. Smith, 2005, 17). Private companies have conserved data and created archives for the re-use of data. For example, professional firms, private broadcasting companies and universities have archived radio and TV shows in order to sell them, re-use them and re-broadcast them for educational purposes.

With digital data, many of these assumptions no longer apply. First, digital data nationality is an issue that does not make sense. A website has no nationality, it sometimes has a national language but English-language websites are frequently not even produced in Anglophone countries, because English is the lingua franca of the web. Digital data and digital historical sources erode national boundaries and establish new international relationships. The second aspect is more pertinent and raises the question of who has the duty to conserve the digital heritage. Unlike the case of analogue data, public companies are not interested in conserving data that are not national and that cannot be re-used for national purposes. Private companies have launched many projects concerning the archiving of the digital past. The Internet Archive (http://www.archive.org), for example, is a project that began archiving the web in 1996 with the aim of preserving the internet, and especially text, audio and audiovisual images appearing on the web. There are important questions that inevitably affect the future media historian: Is it right to leave the past in private hands, even if it is still seen as a public asset? Will the politics of these two private companies change in the future? Will the Internet Archives decide to delete archives older than x years? Who can stop them? 'Can they safeguard in the long term materials that have lost their commercial value?' (de Jong and Wintermans, 2007, 6). How much can a private company charge a user or historian to access old sources? Who should set the price? This is a crucial issue and it is not easily resolved because private companies have to make profits, public institutions have to act for the public good and, to become a reliable subject,

private companies that deal with digital preservation should act at the boundary between business and public good.

There are possible solutions to these potential digital catastrophes. First, digital heritage preservation could be set as a policy goal (B. Smith, 2005, 7), not only at a national level, but also with international recognition. What is more, national legislation is ineffectual in preserving copyright and this is another issue in understanding who should have the duty of preserving data. Another solution is a public–private partnership: public companies have the duty of preservation for humanity and private companies are specialized in providing services and could contribute their know-how to creating user-friendly solutions.

With digital data there is also a new and unexpected subject who could be involved in the preservation process: the user. Users now produce content and store it in specific repositories: for example YouTube is an audiovisual archive of private content, videos taken from the web and, frequently, snippets of broadcast and recorded shows. This is another meaning of the so-called Web 2.0; the web becomes a medium of preservation of old sources that form part of popular mentality. The user, from this point of view, can be considered a sort of archivist because he/she store videos from his/her personal collection and shares them with other unknown people. Users are a crucial link in the chain of digital preservation, because the 'preservation of digital content must be a collaborative effort that involves the professional archivist, the technology expert, the user, and creating and producing entity' (Ide et al., 2002, 12).

Users are pivotal also in the view of Bearman, the author of the most interesting solution to identifying the subjects tasked with digital preservation. Bearman maintains that museums, libraries and archives have different traditions, have saved different things and have been interested in different aspects of the same sources. This model could be applied and improved in the digital environment as well. In other words, the same digital datum could be preserved by different subjects (including the users) for different purposes, and each of these subjects preserves a particular aspect of the same data that could be of interest from different historical perspectives. To do this, 'we will need to create and maintain multiple representation of digital objects to satisfy different aspects of

reliability and authenticity' (Bearman, 2007, 30). This is a collective model that assigns to the entire society the duty of preserving pieces of different digital pasts; the collection will not be complete but at least it will reflect different interests and instances of the society.

There is a final issue related to the subject tasked with preserving the data: power (Schwartz and Cook, 2002). With analogue data and sources specific institutions were set up to preserve the past heritage. Considering schematically the history of preservation, until the invention of the printing press in the sixteenth century there was a link between preservation and religious power: from the pyramids in ancient Egypt to the Roman churches and monasteries, it was always religious institutions that had a monopoly over the preservation of the past. This power became more secular in the eighteenth and nineteenth century: public museums and archives were created to re-use the national past. In the twentieth century, the power of memory was shared by public institutions and private companies, for example broadcasting companies. With digital data, the preservation power seems to be at stake again and the very concept of power afforded by preservation will continue to change. The term 'archives' comes from the Greek *arkhein* (to command) and, until now, conserving and archiving have always been synonymous with power (Sorlin, 2007, 16). When information is scarce, as in the past, its value increases; when it is abundant and easily available as in our digital era, and the only problem is that of selecting what is relevant, its value inevitably decreases (Esposito, 2001). The transition from analogue to digital data and sources could mask a similar change in power and, in an era when producing information is an inexpensive and widespread habit, the task of preserving old digital information could be assigned to a less powerful institution.

Who preserves the data, an issue that already existed in the analogue word, is important for media historians because it could affect their work in at least three different ways. First, it could affect source accessibility. Public and private companies have different responsibilities but private companies do not have the duty to preserve data even if they do it. What claims could a historian make of private companies? If the *New York Times* decides to erase its digital past how will media historians be able to study the way the newspaper saw and reported, for

example, 9/11? If the Internet Archive decides to end its project tomorrow, who will take it over and, above all, who will guarantee that all the gathered sources will be available for historians? These questions deal with the possibility of understanding contemporary society. A second issue is what could be called the risk of a one-dimensional past. Private companies could decide to preserve sources they deem important (and probably those that put them in a good light). For a history of television scholar, for example, studying programmes with scant audience success is as important as analysing those that are really popular because a media historian need to think about the political, economic and social reasons that determined the failure. Finally, if historical sources became a kind of merchandise to sell on the market (and this is one of the media firms' purposes), access to these documents could only be given to the richer, more famous or more influential historians. If broadcasting companies, for example, decide to make people pay for access their digital archives, which universities, which young researchers, which media historians can afford it?

Preserving What?

A central question in understanding digital data and in transforming them into historical sources for future historians is what should be preserved. There are two main approaches, two main schools of thought in what to preserve in the digital era (Phillips, 2005, 58). The first is the *old* one. In the past, national libraries and archives were typically comprehensive and their aim was to collect at least one copy of everything published. Supporters of this approach want to continue this tradition of conserving the entire (digital) heritage. The second school of thought maintains that online publishing is completely different from traditional publishing and so requires a selective methodology. Using poker jargon, the first could be called the 'all-in approach' and the second the 'slow-play approach'.

If a poker player goes 'all-in', he/she bets all the chips left on the table. Translating this into our language and interests, this approach aims to preserve everything produced in the digital era. There are three main advantages in the all-in approach: first, with this approach the entire digital domain will be available for future researchers and humanity; second, this full back-up is performed

automatically and so human intervention (human errors and choices) is minimized; finally, due to the decreasing costs of memory hardware, for the first time it is actually possible and quite cheap to store huge amounts of data, maybe everything.

The all-in approach is also criticized for four main reasons. First, the main disadvantage is that it appears to be a utopia. Storage back-up is performed at intervals (for example every two months) and during these intervals new digital data originate and die because the production of digital data is incessant. Furthermore, due to its volatile and connective nature, the digital heritage is potentially infinite and it is extremely difficult, almost impossible, to store each and every connection. A second disadvantage is that quality control is impossible because, for example, inaccessible websites (with password) are automatically rejected and are not saved. Third, one crucial aspect is that of metadata. The quality, accuracy and richness of metadata are essential when retrieving born digital data (Wactlar and Christel 2002) and the all-in approach does not allow complete and efficient control of their quality. Finally, this approach implies a critical change in the filing clerk's work methodology because, if we let the machine archive everything, the duty of selection vanishes.

The opposite school of thought could be named the 'slow-play approach'. A poker player plays slow, with small bets, in order to get to know his opponents and elaborate a strategy. Similarly, due to its increasing volume, the digital heritage could be conserved in small, representative pieces, selecting them from the entire digital heritage. This approach has two advantages. First, accuracy, because, focusing on a small sample, you can check the completeness (with all the links), the evolution of the website pages considered and the quality of the metadata. In other words, the stored data are fully catalogued. Second, access for users is generally easier because, if you store a small amount of data, you can negotiate access permission with the publishers. This is true also for inaccessible data that could be identified and negotiated with publishers.

There are three main disadvantages to this approach. The most important one is the risk of rejecting what *could be useful* for future media historians. What are the selection criteria? Is it right to define these criteria *now* for future

historians? 'Do we really know what will be important for future researchers?' (Phillips, 2005, 61). For example, gossip blogs and websites could appear to be a superfluous and rejectable part of our digital heritage, but these sources could be crucial for a cultural and social historian of the late twenty-first century. Another example is spam, that it is seen as rubbish or, better, digital rubbish, but maybe future historians will see it as an important source, a significant product of contemporary society and therefore something to study. Cohen and Rosenzweig think that it is possible to assess the long-term value of a document by asking questions such as: Is the information held unique? And how significant is the source for research? (Cohen and Rosenzweig, 2006, 226). But, on the one hand the concept of uniqueness contrasts with that of digital reproducibility and, on the other, who knows the significance of a datum for future researchers? Furthermore, and above all, who has the duty to choose what is relevant and what is not? The second disadvantage of the slow-play approach is also one of its positive aspects: the limited amount of material preserved allows better and complete storage but it also represents a limited part of the huge volume of digital data produced. The final disadvantage of this approach is the risk of including data out of context 'and often not including other resources to which it is linked' (Phillips, 2005, 61); for example, if you preserve a set of websites, you could perhaps exclude a semantic link that you are not aware of.

Thinking about what to preserve of our cultural heritage is a key issue for media historians of the future because on this decision depends the variety and quality of sources that scholars have access to. First, will the needs of a contemporary historian and those of 2060 will be the same? Will 2050 media historians need all the websites or all digital television broadcasts of today in order to study the contemporary? Perhaps media historians of 2050 will not require all the series of *Big Brother* to understand the social and cultural meaning of the programme in the late twentieth century. Maybe one or two episodes will be enough. But the real question is: which of them? Episodes broadcast on TV, saved in YouTube, commented on by the audience, etc.? Second, historians can now try to look at *all* the sources they think are relevant for their studies, but in this new era of abundance, they will probably select relevant sources after seeing only a part of

the total. On the one side, preserving the entire digital heritage means preserving volumes of sources that future media historians will not be able to consult. On the other side, each media historian will look at a small, and probably different, piece of digital data and so preserving everything means preserving many possible lines of research. An historian of late twentieth-century telecommunications, for example, can view only part of the billions of text messages exchanged every day in the world (assuming someone to be responsible for saving SMS, apart from their users), but every historian of the telephone will probably study SMS with a personal view, searching for personal interest. And so, again: which messages are more relevant than others?

The problem of selecting or rejecting material has always been a crucial topic for historians, but the digital era increases the scale of the problem. This is a methodological and technical change too. Even if only a small part of the digital heritage is preserved, historians in any case will have to learn how to make searches in digital repositories, for example using automatic tools, and in general they will have to select and reject much more than in the analogue past.

How to Preserve?

Last but not least is the problem of how to preserve the new data and make them available for future researchers. As already mentioned, digital data are more difficult to preserve than analogue data, especially because of their instability and volatility. One of the most important problems in this respect is the longevity of the digital heritage. Media longevity problems exist both for analogue and digital, but the scale is different: stone remains available and readable longer than printed books, and printed books last longer than magnetic equipment (Conway, 1996). Digital data and records age (much) faster than all other technological supports because of both the hardware and the software. Digital data become rapidly obsolete because of the hardware or, better, the support on which they are stored. Digital data may age rapidly also due to their organization. They are structured in a format designed for an application program that may no longer exist; they are designed and organized for a specific software that could become obsolete in a few years or even just a few months.

Howard Besser (2000) has identified at least five key factors that pose digital longevity problems. The first is the so-called viewing problem. 'When we discover older analogue works, at least we can view them and their structure even if we had lost the ability to decode their language' (Besser, 2000). Cuneiform tables are readable even if we do not know what they mean, but it is more difficult to read digital information because it requires specific software and hardware. Let's take the example of 8-inch floppy disks, that cannot be read because personal computers are not equipped to read them and, even if appropriate hardware could be found, the problem of deciphering the file format still remains (Brand 2003). The second difficulty is the so-called scrambling problem, which involves three different concepts. There are problems created by compression, because the long-term effects of compression are still not fully understood and what it is compressed today may make it impossible to use of these digital data in the future. There are also problems related to the encoding of a file and it may be difficult for future archaeologists and media historians to decipher these digital sources. Finally, there are problems related to the ways in which digital files are protected and encrypted, and so it could be difficult for future historians to decrypt the digital heritage. Another difficulty is created by interrelational problems. As already mentioned, a central characteristic of digital data is their interrelation, their potentially infinite linkability, and in order to better preserve a digital product it is important to try to establish the boundaries of the work, but this is difficult for the digital heritage. Besser has identified a fourth problem, the custodial one. This is what was discussed in the second section: while we wonder how our digital past should be preserved, it is also crucial to understand which organization(s) should take responsibility for their long-term preservation. The fifth and last problem is translation. Translation was necessary in the past as well, because analogue data had to be translated into other forms (also digital ones), but this is even more crucial in the digital work. We could believe that, because we can make identical copies of digital files, then the digital-digital translation will not create problems. This is a mistake because:

> though the bits in the file's contents may be identical, the application environment used to view the file most certainly will be different. In fact the very reason for converting the file is because we are unable to successfully sustain that application's environment over time. (Besser, 2000)

Media historians have to be aware that, if translated into other support and technical environments, digital sources change, but at the same time translation is indispensable in preserving documents and giving future historians the possibility of reading, seeing, listening to them.

Caroline Arms (2000) has developed a conceptual framework that shows at least five different preservation methods, five avenues for saving our digital present. Media historians have to know these methods in order to fully understand how digital sources are stored now and how they will probably be available in the future. The first long-term preservation method is the so-called better media. Longevity and technology independence of the digital heritage is ensured also by the medium on which it is stored; this does not mean that a single medium is *the* best one for all digital sources, but that there are media that are more suited than others to preserve specific data. It entails an ability to transfer the information to new media (Abdelaziz, 2007). A second and related avenue in long-term digital preservation is the migration of content, which implies converting digital data into different and more evolved digital formats. The first two methods for long-term digital preservation are time-consuming and expensive and are also paradoxical. As noted by Su-Shing, digital preservation is subject to a fundamental paradox: 'On the one hand, we want to maintain digital information intact as it was created; on the other, we want to access this information dynamically and with the most advanced tools' (Su-Shing, 2001, 3). In other words – and this aspect is of fundamental importance for media historians – we aim at maintaining the originality of the digital source that implies either the originality of the content or the originality of the environment, but we also have to deal with hardware and software obsolescence and, if we maintain digital info without modification, accessing them will eventually become impossible. What's more, we want to use the most advanced machinery for optimal search in these sources.

A series of emulation techniques has been created to respect the original environment of utilization. This third long-term preservation method implies that new technologies could simulate the ways a previous generation of technologies function and reconstruct their ambience and so digital sources can be read, seen, listened to in their original environment, even if this originality is an artifice. The fourth method of long-term preservation is called refreshing bits, and 'includes any operation that simply involves copying a stream of bits from one location to another, whether the physical medium is the same or not' (Arms, 2000). This means that if you frequently and automatically copy *old digital* sources into new repositories and with new formats you will probably contribute to their preservation. The final method considered by Caroline Arms (2000) is digital archaeology and is used only if 'all else fails'. If a future historian is unable to access a digital past in any way, he/she will have to pick up pieces, traces, fossils of this digital past as an archaeologist does with dinosaurs.

I would add a sixth avenue, a sixth way of preserving data that could ensure the long-term availability of digital data: the multiple (potentially infinite) copy. As pointed out in the second section, users will increasingly become a *fragmented institute of preservation*, because they have personal repositories and sometimes share them. A decentralized memory with redundancy and continuous multiplication in copies could ensure long-term digital preservation. Millions of decentralized personal computers own small or large pieces of cultural heritage and, if they were linked, a network-repository would be created and this network would enhance the long-term availability of digital sources. This is a fairly old preservation strategy. Above all in the analogue world the existence of multiple copies of a work stored at different geographic locations helped preserve this work and, for example, copies of the books stored in the Alexandria library helped save the cultural heritage after the library was destroyed. The multiplication of copies could be an analogue strategy, maybe *the* analogue strategy, best suited to the digital case as well.

The way in which digital sources are preserved obviously affects future media historians' work. Historians of 2050 studying our present media history will have to use different methodologies and approaches depending on how data are

(and were) stored. For example, to study digital forms they will have to abandon the idea of reading on paper or, at the very least, they will have to follow the inextricable connection between paper and bits. In fact, you can print a website and read it on paper but you automatically lose its interactive and hypertextual character. Another major problem in studying digital sources on paper is the impossibility of the automatic search typical of digital data. When you have to search within a huge number of emails, you need to use automatic functions, and if you print them out you miss out on a large part of the potential and particularity of the digital heritage. This is the greatest danger for historians: sacrificing 'the original form, which may be of unique historical, contextual, or evidential interest' (Rothenberg, 1999).

Abandoning paper is a small example of a necessary change in mentality for historians approaching digital media history. This change implies that historians will stop using the web like a printed repository guide or the telephone book (Tibbo, 2002, 9) or printing resources rather than consulting them online (thus maintaining the better medium and the better original context in which the source was created). This change of mentality implies that historians will have to understand new search methods and learn new skills. Huge amounts of data could be stored in a non-readable format in the near future, and this is a danger, but it also implies that future media historians will be digital archaeologists as well. This change in mentality implies that historians will consider – in McLuhan's words – either the message or the medium; in other words, they will have to be aware of and understand how the message is and was conveyed, and how the hardware affects the way in which the message is and was perceived (evaluating the migration and emulation techniques).

The way in which digital sources are preserved and will be used by historians also entails new dangers; for example, the danger of de-contextualization. In *old* analogue archives, documents were collected in folders and so many papers were *physically* close to each other. In other words, each document was part of a more complex sphere and this singularity was influenced by multiplicity (Vitali, 2000). In digital repositories – due to the huge amount of data stored, the immateriality of digital data and the structure of digital archives –

media historians risk decontextualizing documents because the automatic search produces singular data, singular digital documents not linked with other (maybe linked) digital objects.

Conclusion

On digital preservation, there are still more open questions than solved ones. Compared to the long-standing experience of preserving the *old* analogue heritage, it is still unclear who has responsibility for preservation, what should be preserved and how it should be preserved. Responsibility for the preservation of culture, which once was attributable to specific institutions, is and will be increasingly difficult to assign; the crisis of national public institutions and the interests of private companies (which of course archive to make profits), for example, show the need for others to be involved in this operation. In this aspect users could play a decisive role, because they are appropriating and reworking the media messages and, increasingly, they are producing contents (this is the logic of the so-called prosumer). It is even more difficult to ask what should be preserved. Storing all the contemporary communication that is produced is utopian and there is much 'noise' that it would probably make no sense to save; at the same time, however, future media historians' interests cannot be forecast and so the more data that are saved, the more lines of research are preserved. How to save the heritage is the third and final question, increasingly important in an interconnected world of communications, and, to fully understand our contemporary media system, media historians of 2050 will have to recreate the present idea of networking and connection among different cultures.

Media historians, as well as all other historians, must be willing to accept significant changes in their profession: they will have to approach the digital heritage differently than the analogue one, because digital data are volatile, interconnected, unstable and abundant; they will seek new tools to manage this great mass of data; they will use technological systems because 'Archives of the future will be different and researchers will adopt new, and more technology dependent, ways of working' (Ross, 2000, 11); they will interact with new institutions and they will learn new methods of access to sources; they will have to

learn new methods to reconstruct the past from these digital sources; and, finally, they will face new difficulties, such as accessibility, ownership, fragility, originality and contextualization of digital sources (Poster, 2007).

As with any change, however, there will be strong continuities with the past. Just as they did in the past, for example, media historians will continue to choose between sources; maybe they will have to learn new research strategies, but they will still interpret the data based on personal and specific research questions. In other words, the digital revolution has nothing to do with the *interpretation* of historical sources: doing media history is and will continue to be a far more complicated job than selecting and combining the available sources to reconstruct a media story; historians interpret the documents according to a theory of reference – the sources are only a sign of the past and significance of this past cannot be explained only by those contingencies. Media history remains a discipline that deals with the reconstruction and the interpretation of reality: there have been and will be different readings of the same historical and media event and in future the profession will continue to be concerned with the interpretation of the (many) data available.

What is certain is that it is time to ask questions and to act in order to preserve the fragile memory of our times (Donk, 2009) so that media historians of 2050 have sources available to interpret the contemporary media system. Deciding now which institutions have the responsibility for preservation, what should be preserved and how, is the only way to give future media historians traces of our digital media present. To let them study and try to understand late twentieth-century memory and, finally, say 'Now we can say we've been there' (Adorno, 1999, 326).

References

Abdelaziz, A. (2007) 'Safeguarding our digital heritage: a new preservation paradigm', in Y. de Lusenet and V. Wintermans (eds.) *Preserving the Digital Heritage: Principles and Policies*, selected papers of the international conference organized by Netherlands National Commission for UNESCO, 4

November 2005, The Hague, The Netherlands, available at www.knaw.nl/ecpa/publ/pdf/2735.pdf (accessed 17 April 2009), pp. 7–14.

Adorno, T.W. (1999) *The Complete Correspondence, 1928–1940 / Theodor W. Adorno and Walter Benjamin*. Cambridge, MA: Harvard University Press.

Arms, C.R. (2000) 'Keeping memory alive: practices for preserving digital content at the National Digital Library Program of the Library of Congress', *RLG DigiNews* 4(3), June, available at http://webdoc.gwdg.de/edoc/aw/rlgdn/preserv/diginews/diginews4-3.html (accessed 17 April 2009).

Bearman, D. (2007) 'Addressing selection and digital preservation as systemic problems', in Y. de Lusenet and V. Wintermans (eds.) *Preserving the Digital Heritage: Principles and Policies*, selected papers of the international conference organized by Netherlands National Commission for UNESCO, 4 November 2005, The Hague, The Netherlands, available at www.knaw.nl/ecpa/publ/pdf/2735.pdf (accessed 17 April 2009), pp. 26–54.

Besser, H. (2000) 'Digital longevity', in M. Sitts (ed.) *Handbook for Digital Projects: A Management Tool for Preservation and Access*, Andover, MA: Northeast Document Conservation Center, 155–66.

Borgman, C.L. (1999) 'What are digital libraries? Competing visions', *Information Processing and Management: An International Journal* 35(3): 227–43.

Brand, S. (2003) 'Escaping the digital dark age', *Library Journal* 124(2): 46–9.

Byerly, A. (2009) '*What Not to Save: The Future of Ephemera*'. Paper presented at MiT 6, Boston, MA, 24–6 April.

Cohen, D.J. (2004) 'History and the second decade of the Web', *Rethinking History* 8(2): 293–301.

Cohen, D.J. and R. Rosenzweig (2006) *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web*. Philadelphia: University of Pennsylvania Press.

Conway, P. (1996) *Preservation in the Digital World*. Washington, DC: Commission on Preservation and Access.

de Jong, A. and V. Wintermans (2007) 'Introduction', in Y. de Lusenet and V. Wintermans (eds.) *Preserving the Digital Heritage: Principles and Policies*, selected papers of the international conference organized by Netherlands National Commission for UNESCO, 4 November 2005, The Hague, The Netherlands, available at www.knaw.nl/ecpa/publ/pdf/2735.pdf (accessed 17 April 2009), pp. 1–6.

Donk, A. (2009) '*The Digitization of Memory: Blessing or Curse? A Communication Science Perspective*'. Paper presented at MiT 6, Boston, MA, 24–6 April.

Duff, W., B. Craig, and J. Cherry (2004) 'Historians' use of archival sources: promises and pitfalls of the digital age', *The Public Historian* 26(2): 7–22.

Esposito, E. (2001) *La memoria sociale: mezzi per comunicare e modi di dimenticare*. Roma: Laterza.

Grant, A.E. and J.S. Wilkinson (eds.) (2009) *Understanding Media Convergence: The State of the Field*. New York: Oxford University Press.

Ide, M., D. MacCarn, T. Shepard and L. Weisse (2002) *Understanding the Preservation Challenge of Digital Television*, Council on Library and Information Services report, available at www.clir.org/pubs/reports/pub106/television.html (accessed 17 April 2009).

Innis, H.A. (1951) *The Bias of Communication*. Toronto: University of Toronto Press.

Jenkins, H. (2006) *Convergence Culture: Where Old and New Media Collide*. New York: New York University Press.

Kuny, T. (1997) '*A Digital Dark Ages? Challenges in the Preservation of Electronic Information*'. Paper presented at the 63rd IFLA Council and General Conference, Copenhagen, Denmark, 31 August –5 September.

Lévy, P. (1997) *Collective Intelligence: Mankind's Emerging World in Cyberspace*, trans. R. Bononno. New York: Plenum Trade.

Mackenzie Owen, J. (2007) 'Preserving the digital heritage: roles and responsibilities for heritage repositories', in Y. de Lusenet and V. Wintermans (eds.) *Preserving the Digital Heritage: Principles and Policies*, selected papers of the international conference organized by Netherlands National Commission for UNESCO, 4 November 2005, The Hague, The Netherlands, available at www.knaw.nl/ecpa/publ/pdf/2735.pdf (accessed 17 April 2009), pp. 45–49.

Manoff, M. (2006) 'The materiality of digital collections: theoretical and historical perspectives', *Libraries and the Academy* 6(3): 311–25.

Ortoleva, P. (1999) 'La rete e la catena. Mestiere di storico al tempo di Internet', *Memoria e ricerca* 31(3), available at www.fondazionecasadioriani.it/modules.php?name=MR&op=body&id=76 (accessed April 17, 2009).

Phillips, M.E. (2005) 'What should we preserve? The question for heritage libraries in a digital world', *Library Trends* 54(1): 57–71.

Poster, M. (2007) 'Manifestos for a history of the media', in K. Jenkins, S. Morgan, A. Munslow (eds.), *Manifestos for History*. New York and Abingdon: Routledge.

Poster, M. (1999) 'Underdetermination', *New Media & Society* 1(1): 12–17.

Rosenzweig, R. (2003) 'Scarcity or abundance? Preserving the past in a digital era', *The American Historical Review* 108(3): 735–62.

Rosenzweig, R. (2006) 'Can history be open source? Wikipedia and the future of the past', *Journal of American History* 93(1): 117–46.

Ross, S. (2000) *Changing Trains at Wigan: Digital Preservation and the Future of Scholarship*. National Preservation Office Occasional Papers, available at http://www.bl.uk/npo/pdf/wigan.pdf (accessed 17 April 2009).

Rothenberg, J. (1999) *Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation: A Report to the Council on Library and*

Information Resources. Washington, DC: Council on Library and Information Resources, available at http://www.clir.org/pubs/reports/rothenberg/contents.html (accessed 17 April 2009).

Saksida, M. (1997) 'The information society in the 21st century: converting from analogue to digital', *International Information & Library Review* 29(3–4): 261–7.

Schloman, B.F. (2003) 'Information resources: now you see it, now you don't: the ephemeral nature of digital information', *Online Journal of Issues in Nursing*, 8(2), available at www.nursingworld.org/MainMenuCategories/ANAMarketplace/ANAPeriodicals/OJIN/TableofContents/Volume82003/No2May2003/EphemeralNatureDigitalInformation.aspx (accessed 17 April 2009).

Schwartz, J.M. and T. Cook (2002) 'Archives, records, and power: the making of modern memory', *Archival Science* 2(1–2): 1–19.

Smith, B. (2005) 'Preserving tomorrow's memory: preserving digital content for future generations', *International Preservation News* 29: 4–10.

Smith, C. (2005) 'Building an internet archive system for the British Broadcasting Corporation', *Library Trends* 54(1): 16–32.

Sorlin, P. (2007) 'Historians at the crossroads: what can we do with visual archives?', in H. von Kurt Deggeller, U. Ganz-Blättler and R. Hungerbühler (eds.) *Heard-Seen: The Uses of Digitised Archives for the Sciences*, Baden and Lugano: Hier und Jetzt, pp. 16–23.

Su-Shing, C. (2001) 'The paradox of digital preservation', *Computer* 34(3): 24–8.

Tibbo, H.R. (2002) 'Primarily history: historians and the search for primary source materials'. In *Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries.* Portland and Oregon, available at http://portal.acm.org/citation.cfm?doid=544220.544222 (accessed 17 April 2009), pp. 1–10.

Uricchio, W. (2007) 'Moving beyond the artifact: lessons from participatory culture', in Y. de Lusenet and V. Wintermans (eds.) *Preserving the Digital Heritage: Principles and Policies*, selected papers of the international conference organized by Netherlands National Commission for UNESCO, 4 November 2005, The Hague, The Netherlands, available at www.knaw.nl/ecpa/publ/pdf/2735.pdf (accessed 17 April 2009), pp. 15–25.

Vitali, S. (2000) ''*Archivi Online': Qualche Riflessione Metodologica'*. Paper presented at the Workshop Archivi storici e archivi digitali tra ricerca e comunicazione, Florence, Italy, 20–21 October, available at www.dssg.unifi.it/_storinforma/Ws/archivi/vitali.rtf (accessed 17 April 2009).

Wactlar, H.D. and M.G. Christel (2002) 'Digital Video Archives: Managing Through Metadata', in *Building a National Strategy for Digital Preservation: Issues in Digital Media Archiving* April: 84–99, available at www.ri.cmu.edu/publication_view.html?pub_id=3964 (accessed 17 April 2009).