INTRODUCTION: DIGITAL METHODS AND TOOLS FOR HISTORICAL RESEARCH

DANIEL ALVES

Abstract  As an introduction to a series of articles focused on the exploration of particular tools and/or methods to bring together digital technology and historical research, the aim of this paper is mainly to highlight and discuss in what measure those methodological approaches can contribute to improve analytical and interpretative capabilities available to historians. In a moment when the digital world present us with an ever-increasing variety of tools to perform extraction, analysis and visualization of large amounts of text, we thought it would be relevant to bring the digital closer to the vast historical academic community. More than repeating an idea of digital revolution introduced in the historical research, something recurring in the literature since the 1980s, the aim was to show the validity and usefulness of using digital tools and methods, as another set of highly relevant tools that the historians should consider. For this several case studies were used, combining the exploration of specific themes of historical knowledge and the development or discussion of digital methodologies, in order to highlight some changes and challenges that, in our opinion, are already affecting the historians’ work, such as a greater focus given to interdisciplinarity and collaborative work, and a need for the form of communication of historical knowledge to become more interactive.

Keywords: history, research methods, digital technology, digital texts, interdisciplinary
Daniel Alves

1. Introduction

The articles presented in this special edition are the result of a deeper analysis of ideas and discussions which arose during a seminar on digital methods and tools organized as part of the activities of the research group on digital humanities of the Institute of Contemporary History at Universidade Nova de Lisboa. Via a panel of researchers in the areas of humanities and technology with a highly relevant connection to history, the purpose both of these articles and of the seminar which brought them about was to highlight and discuss in what measure diverse methodological approaches centered on the use of digital technology could contribute to improve analytical and interpretative capabilities available to historians.

The object of this introductory text is to discuss the impact of the application of digital technology on the production and dissemination of historical knowledge, beginning with a brief presentation of some analyses on the subject undertaken since the beginning of the 1990s and highlighting some challenges that historians seem to be facing as a result of the technological evolution of the past decades. Through this approach, I’m also looking at contributing to an integrated reading of the articles published here.

Joining together a group of texts which discuss methodologies and digital tools applied to researching the past is obviously not original. However, the articles presented here seek, on the one hand, to renew the discussion about the interaction between history and information technology at a moment in which the latter provides new tools and methodologies at an ever-increasing velocity, especially those dealing with large amounts of text, spatial analysis, visualization technology, and benefitting from the so-called social web. On the other hand, these articles constitute an attempt to bring digital history closer to a more vast academic community, typically skeptical towards the validity of the use of digital tools and methodologies or not inclined to use them due to the difficulties in understanding their functions and utility or for lack of capacity to spend the time needed to learn these tools and methodologies.

Taking these aspects into account, in the majority of these articles we sought a concrete assessment of historical questions or historical perspectives. Using these problems and perspectives as a base and highlighting the role of a digital approach within them, we could then discuss how far this very approach could go to help resolve those questions or lead to exploring new points of view. The aim here is to go beyond a simple list of digital tools, and, wherever possible, avoid perspectives overly reliant on technology as well as on excessively technological vocabulary whose specialized language is quite often off-putting to the more skeptical or to those less familiar with the digital world.

The target audience of this set of articles falls between digital enthusiasts and digital skeptics, seeking to contribute to the strengthening of a bridge...
Digital Methods and Tools for Historical Research

between the two types of researchers, teachers, and students that respectively represent—even today and despite the recurring promises of revolution—the minority and the majority of the academic world. The point here is not to convince the majority one more time of the supposed marvels of the digital revolution, but to resort to concrete examples of research and discussion of results, increasing the flow on this bridge that still divides historians from information technology. With this statement I wish to make clear, however, that this flow will bring more advantages than disadvantages to building historical knowledge.

2. HISTORY AND THE DIGITAL REVOLUTION(S)

The interaction between history and digital technology has been referred to repeatedly as the catalyst for a revolution, more than as the mere communication, in historical knowledge in general. However, just as in other types of revolution, there are aspects of this knowledge that are being altered and others that, from what we can see, seem unchangeable and persist. In that which concerns historical knowledge, we must ask which part is being affected by the digital movement and which part seems unchanging? To what extent are these alterations contributing to new forms of scientific production and access to historical knowledge? Could it be that digital facilitation will lead to a transformation of the historian’s work or will the difficulties inherent to digital technology resources pile up and restrict the use of these technologies by the academic world?

In the beginning of the 1990s when Internet browsers took their first steps, the academic community was divided in its interpretation of the future consequences of the then new digital world. Some signs seemingly pointed in the direction of a future paradise, such as the creation in 1987 of The Association for History and Computing, and the journal History and Computing two years later, as well as the growing reliance on computing power for historical studies based on large volumes of sources, mainly of a quantitative type. There was even talk of ‘a qualitative and quantitative revolution in the relationship between history and computing.’\(^5\) However, there were those who observed the availability of an enormous and ever-increasing amount of material in digital format at dizzying speeds and prophesized the slide into the abyss of the age-old system of production and validation of scientific knowledge and academic credibility.\(^6\)

Despite the initial debates, in 1999 the community of historians remained divided between those who resisted the idea of making use of ascendant technology and those who would enthusiastically embrace what the digital world had to offer. The idea remained amongst the skeptic that little or nothing had changed in the traditional way to write history. They had not yet glimpsed significant effects that the digital could introduce into the production and
dissemination of historical knowledge, namely at the level of the historical narrative or in the manner of reaching a much-desired wider audience. The enthusiasts foresaw a democratization of history’s target audience, a stimulus to theme diversity, and a developing interest for new narrative techniques. A strengthening of interdisciplinary collaboration already seemed evident between history and other humanities and social sciences which would make possible a ‘more dynamic, interactive, and reflexive’ history. In this respect, the connection to digital technology seemed to be a natural evolution of history as a subject.\footnote{7}

A few years later a new diagnostic of the interaction between history and digital technology seemed to point towards the same conclusion. Although it did not deny that digital tools had brought changes in terms of society, daily behavior, information seeking, forms of communication, and even changes in how we ‘research, write, present, and teach about the past,’ in 2005 the authors stated that this interaction was still far from a ‘revolution.’ The profusion of web sites about history, the strong growth of publishing sources online, and the fact that a significant number of specialty journals already had a presence on the world wide web were all highlighted as positive aspects. With this the possibility glimmered that historians could, through digital technology’s ‘capacity, accessibility, flexibility, diversity, manipulability, interactivity, and hypertextuality,’ undertake better, richer investigations. However, they pointed to the risks inherent in this digital history, that it could settle for ‘quality, durability, readability (…) and inaccessibility’ of information, and also the ‘passivity’ of the reader.\footnote{8}

Despite everything, the main advantage of digital technology in history seemed to be its storage and processing capacity: to store an enormous amount of information in a very small space, something that had already been noted in the previous approaches.\footnote{9} This aspect, now called ‘Big Data’, is not irrelevant as this characteristic was and continues to be pointed as one of the main justifications for the emergence of some challenges and changes in the work of current historians. These challenges and changes relate to the superabundance or scarcity of sources that, one could predict, would become determinant in future historians’ approach to the past.\footnote{10}

The rhythm of recent evolutions, namely in respect of the so-called web 2.0 or social web, has been significant and has led historians to also suggest a ‘History 2.0.’\footnote{11} Effectively, if we look at the speed of technological changes, a couple of years is a gigantic leap in time, even if we solely concentrate on online tools. In 2005 Google Books, Google Scholar, and Google Maps were all in their infancy and probably no one had any idea how much impact they would have in today’s historical research.\footnote{12} YouTube also debuted in this same year. By comparison, Wikipedia, launched in 2001, was already a senior citizen on the Internet, having arrived at the same time as when the TEI Consortium began operations.\footnote{13} IBM’s Many Eyes arrived on the scene in 2007, symbolizing a greater attention to issues
Digital Methods and Tools for Historical Research

of visualizing data, their connections, hierarchies, and networks that today are central to many research projects in the field of digital humanities. In 2008 it was Academia.edu’s turn to bring the concept of social network into the realm of academic investigation and dissemination. It’s worthy of note that as late as 2005 Twitter had not yet come into existence. Nowadays it is used, amongst other things, for so-called ‘unconferences’ and paved the way for the flow of information that lead up to the so-called Arab Spring. Twitter was characterized in 2010 by the Library of Congress as having an ‘extraordinary potential for research into our contemporary way of life.’

Given this and more (the list is far from exhaustive), it would be safe to say that if history and historians didn’t embark upon a supposed epistemological revolution, it’s clear that they would not be immune to these changes. Relational databases or Geographic Information Systems (GIS), software considered revolutionary when applied to history two decades ago, are nowadays quite common work tools for a lot of historians, albeit often used in an indirect manner and with the help of computer technicians. Perhaps the current trend for these two tools would be to try and find systems or methodologies that would allow to better deal with the narrative aspect of history. At the same time, this attention to textual sources (which are the main sources for historians and other humanities researchers) results as well from the big undertakings of digitization and from the democratization of the Web which puts ever more texts at the disposal of databases and GIS and their capacity to extract and analyze, allowing historians to posit new problems to sources which hardly could have entered into the equation before.

This has turned back the spotlight on the text in the relationship between history and digital technology, in what we could call a textual turn in the field of digital history (following the fashion for turns which has become popular in epistemological discussions since the 1970s). Some projects’ concern with the publishing of sources has been to publish highly structured, annotated, and searchable editions which enable them to be shared in the future in an interoperable, reusable manner that would presumably be compatible with the technological evolution.

At the same time, the evolution of digital technology has facilitated the relatively successful integration of new collaborative methodologies amongst researchers, as well as amongst themselves and a wider community, as in crowdsourcing. And it has facilitated management and data sharing via cloud services such as Zotero or DropBox, for instance. Furthermore, the growing critical approach to texts and the use of international standards (such as TEI) in text preparation and dissemination can be considered an attempt to overcome the problems connected with the durability and usability of digital sources in the future. At the core, these efforts demonstrate a concern for the preservation of the cultural heritage that, increasingly, is to be found in digital format.
The very process of publishing and disseminating history has undergone changes, being in some ways facilitated, for example, by the profusion of digital journals and by the transposition to the digital format of journals with paper editions. Having empowered information sharing and democratized access to it, even if we don’t take it in to account the current movement towards Open Access, this digital era has created new demands on historians who now see their work scrutinized by a much wider audience. These changes challenge their capacity to produce original material and at the same time unleash very significant problems regarding the protection of their intellectual rights and the frequency and extent of plagiarism, for example.

Despite all this, something is changing, perhaps in a silent or indirect manner as mentioned above, without historians taking note and without the ability to even speak with the certainty of reliable indicators. I believe two of these changes have to do with the construction of historical narrative and the development of collaborative and interdisciplinary work.

As for the historical narrative, there are those who defend that it is changing, in large part, due to external pressures, the most significant being that which is exercised by digital communication channels and, primarily, the social web. Resorting to Marshall McLuhan and his notion of the culture of the printed book as something that made the typical interactivity of oral culture difficult, it is argued that the digital medium can help the historical narrative (which is still to a large extent dependent on printed culture) recover that interactivity. The changes of the past decade or so have brought implications not only to publishing, passing from the printed to the digital form, but also in social and cultural terms inasmuch as new media have the capacity to ‘generate ideas about information and newsworthiness, steer information flows, create networks, and shape attitudes’ in relation to aspects perpetually defended by academia such as ‘authorship, authority, and reliability’. In this sense, I think that it is necessary to affirm that the historian effectively needs to improve the above-mentioned capacity or he/she runs the risk of seeing his/her discourse ignored or undermined by easier forms of memory construction.

In the interdisciplinary field as well, this evolution has more than ever reinforced what we could classify as silent change in history and in the humanities. In the case of GIS, for example, various approaches to linguistics and literature have been explored. This can be seen, e.g., in the recent works of Ian Gregory, a specialist in historical GIS, in collaboration with colleagues in those areas. Archaeology, classical studies, and computational science have equally worked together as in the case of Google Ancient Places project, for instance. Other projects gather specialists from areas as diverse as geography, history, economics, computational science, and physics as with the project Water, Road & Rail. This crescendo of interdisciplinary
work—whether due to research projects crossing knowledge from various disciplines, whether due to the use of certain methodologies or in order to reach certain results, the historian is now more aware that it is necessary to resort to specific collaborations with other specialists—is today an undeniable reality. Perhaps attentive observation to the indexes of scientific magazines, an accounting of the number of authors per article, and the identification of their particular curriculums would allow us to confirm that the work of the historian is no longer that of a tireless archive researcher who continues, in an insular way and over many years, on a single-themed investigation, or the work of someone who will singlehandedly take credit for the results. Whether we like it or not, this change, silently, has established itself as a revolutionary (it is difficult to escape the term!) element of the historian’s work and, in my opinion, it represents added value to the diversification, democratization, and deepening of historical knowledge.

However, we shouldn’t let all these characteristics of the digital age make us forget two essential components in the historian’s work which should be ever-present whether working with information in analog or digital format. Despite the advantages of going digital—be they the sheer potential of storage, increased capacities of visualization and analysis, the democratization of access, or the development of sharing and collaborative work—it is fundamental that the historian, enthusiast or merely proficient in the use of technology, not let himself/herself become dazzled by the technology itself and center his/her work on, at times, endless experimentation with tools that are each time more powerful and more available. Digital technology should be just one more tool available to the historian to which he/she resorts directly or in collaboration with others who are more proficient, so that it is later possible to get back to what really matters, the production of new historical knowledge. The second component is the critical mind, the detailed analysis of sources, permanently questioning whether and in what way digital technology could help the historian’s work and turn it into better and/or different work. This critical reasoning should cause the historian to look at the advantages and disadvantages of resorting to digital methods and tools for historical research and at each step be able to, as mentioned in 2005, “maximize the former while minimizing the latter.”

It is not the objective of this text to trigger a discussion over whether digital technology is calling these factors into question in either a good or bad sense. But it is undeniable that these issues are having an impact on the historian’s work and become more relevant as digital technology becomes all the more important in the everyday life of the individual and in framing the historian’s ways of accessing and participating in culture and in the circulation of information in general. Among other things, this means that the changes will be even more evident to a generation of young and future historians that were born in the digital age than to a generation which grew up professionally in the shadow of other paradigms.
Daniel Alves

It was with this concern of starting in history, taking advantage of all that digital technology has to offer, to once again return to history, that we gathered the following contributions, from which resulted a significant part of the previously presented reflections and ideas about the challenges and opportunities accessible to the historian in the digital age.

3. DIGITAL METHODS AND TOOLS FOR HISTORICAL RESEARCH

The articles in this special issue deal with different facets of doing history in a digital age, ranging from research tools and methods to web archiving, as well as data collection and text edition. John Bradley presents an article in which he discusses the ways of using and exploring models of structured data, giving particular attention to questions related to the treatment of texts and not quantitative sources. He questions the reasons for the resistance to the use of databases in history and whether this resistance is not related to the narrative nature of the subject itself. Luís Silveira presents an analysis of the use of GIS in historical research, highlighting some of the results and possible developments achieved, and discussing the contributions and limitations of historical GIS. He discusses the idea that the technology has caused a revolution in the production and dissemination of knowledge in history, and questions the idea that it provoked a spatial turn in this subject. With Tim Causer and Melissa Terras we are in the domain of data collection and mark-up of a considerable manuscript corpus through crowdsourcing. These authors present and assess an example of how, using web 2.0 tools, volunteers may be involved not only in the relatively simple task of transcribing a vast collection of manuscripts, but also in the more demanding job of marking up texts, making them more easily and rapidly available for historical research. Rita Marquilhas and Iris Hendrickx explain a method for automated spelling normalization of a body of historical letters from a linguistics point of view, a prior step to the analytic corpus linguistic methods. This approach, a kind of distant reading process, presents an excellent prospect for historians, usually accustomed to dealing with large amounts of texts and not always blessed with the time to read all the sources that they would like to. Malte Rehbein, also discussing the combination of distant and close reading, considers the scholarly edition of historical texts and discusses the potential of visualization techniques in the analysis of a large data set. Finally, Daniel Gomes and Miguel Costa reflect on the need to preserve digital information available on the Internet which will constitute tomorrow’s resources for historical research while at the same time presenting some tools currently available which facilitate the researcher’s task to dig into the (near) past.

The importance of textual information in historical research is one of the features highlighted by all of the authors but especially in the contributions
Digital Methods and Tools for Historical Research

from Marquilhas and Hendrickx, Causer and Terras, and Rehbein. At the same time, Bradley’s and Silveira’s articles reflect on the problems posed by the unstructured condition of this type of source, and occasionally by the uncertain and fuzzy character of historical information in the application of tools based, as a norm, on binary logic, and frequently conceived to deal with structured and precise data. For decades, this has been one of the major challenges facing the dissemination of databases and GIS, for instance, in history and in other humanities subjects.

The intrinsic interdisciplinary nature of digital research methods is an aspect that also clearly emerges from all of the texts: databases arose in computer science; GIS was developed in engineering and landscape architecture before being picked up by geography; text encoding and procedures for normalizing spelling grew from linguistic and literary studies; and the link with computer science emerges as well in the context of web archiving. Interdisciplinary encouragement is an important result of the application of digital methods to history, but poses a major challenge to the historians’ training as it requires them to be prepared to enter into dialogue with people from other areas and to engage in teamwork. This aspect is pointed out or is implicit in the research projects described in almost all articles.

The impact of the Internet on the field of history is also reflected in some texts. Silveira focuses especially on web GIS, on its ability to combine time and space, to transmit a sense of place, and to superimpose past and current events. He also underscores the ability of this type of web site to engage users. This latter feature is central to the article by Causer and Terras, whose project relies on the development of the collaborative Internet. The Web and the preservation of the historical memory is the very subject of the contribution by Gomes and Costa.

Visualization has recently become an area of great development within the larger context of digital humanities. The need for new ways to represent spatial-temporal information is increasingly felt, as shown by Silveira. Its ability to represent networks of knowledge is exemplified by Rehbein.

Although not an explicitly discussed factor, it’s possible to glean from some articles the potential that digital technology possesses to develop a facet which preserves and values cultural heritage, whether this heritage was born in the digital format or converted from the analog. This seems evident in the example of Web archives and publishing of critical editions of historical sources that arise from the contributions of Gomes and Costa, Marquilhas and Hendrickx, Causer and Terras and Rehbein.

Finally, the various projects mentioned in the articles implicitly confirm the efficiency of digital means to process large amounts of data, but also to try and focus attention on the fact that their application is, often, generally associated with expensive projects requiring extensive human resources with diverse skills.
Daniel Alves

and adequate funding, and this is one of the challenges of the interaction between history and information technology that should always be taken into consideration in the planning of research work.

4. CONCLUSION

The different perspectives on how digital technology fits into the historian’s work which were approached during the seminar in Lisbon and compiled in the articles presented here contain, in my point of view, a significant potential to improve the connections between history and the digital world. The apparent textual turn within the interconnection between history and the digital world over the past decade or so can greatly contribute for this, I believe, taking into account that just a few years ago the lack of flexibility of the available digital tools enabled only the effective utilization and analysis of quantitative sources or sources easily transformed into a quantitative format, and almost always highly structured. This textual focus of the digital tools and methods is enabling new ways of exploring old sources, and the formulation of different questions, difficult to think of and/or resolve in an efficient way in a recent past: be it the discovery of inter-textual relationships hidden in tens of thousands of letters between writers of past centuries, that can give us a deeper insight about their social and intellectual networks; or the possibility of searching for changes in the concepts and perceptions about poverty and inequality, for instance, through the exploration of millions of books in digital format spanning three centuries. This capabilities and the fact that all articles in this special issue called attention to or were exclusively dedicated to text sources seems to be a positive omen for a reduction of the gap between the majority of academia and the recourse to digital technology in historical research, based both on the textual nature of the historical narrative as well as on the textual nature of the vast majority of its sources.

END NOTES

1 In the course of events that lead to publishing these articles, I should thank all that contributed to the results, the positive ones of course, as any others are solely my responsibility. I thank the attentive and participative public that attended the seminar in Lisbon; all the colleagues who presented their work and ideas there, even if they’ve not been able to accompany the rest of this process; the authors of the works presented here for having accepted the challenge and having had the patience for my e-mails; Paul Ell and David Bodenhamer for having responded warmly to the idea of this publication; and to Luís Silveira for his encouragement from the initial draft all the way to the seminar.


3 See, for example, M. Greengrass and L. Hughes, eds., The Virtual Representation of the Past (Surrey, 2008); J. Genet and A. Zorzi, eds., Les historiens et l’informatique: Un métier à
réinventer (Roma, 2011); F. Clavert and S. Noiret, eds., L’histoire contemporaine a l’ère numérique. Contemporary History in the Digital Age (Bruxelles, 2013).


Daniel Alves


19 A. Rigney, ‘When the Monograph is no Longer the Medium: Historical Narrative in the Online Age’, *History and Theory* 49, 4 (December 2010), 100–117. Cited here at 100–104.

20 Rigney, ‘When the Monograph is no Longer the Medium’, 105.


22 Barker, ‘Taking a GAP Year’.

23 Alves and Queiroz, ‘Studying Urban Space and Literary Representations Using GIS’.


26 Cohen and Rosenzweig, *Digital History*, Introduction.