PART I

The Beginning

CHAPTER 1

Digital and Distant Histories

Emergent Approaches within the New Digital History

Petri Paju, Mila Oiva and Mats Fridlund

Half a century ago, historian Emmanuel Le Roy Ladurie, when surveying the progress of quantitative history, prophesised that 'tomorrow's historian will have to be able to programme a computer in order to survive'.[1] Since then, computers and programming have indeed profoundly changed historians' practice through such digital tools as word processing, the internet, email, PowerPoint, Google, JSTOR, Facebook, Twitter and Zoom. They have made all of us historians into digital historians in one way or another. As these digital tools used by most historians illustrate, there are many ways that the digital has transformed the historian's craft beyond mere practical and administrative improvements. During the new millennium, the computer together with the internet have begun to change also the historian's research tools and methods in new and previously unforeseen ways into a novel kind of digital history. It is this new emerging digital history, together with some ever-significant approaches of the 'old' quantitative digital history, that is the subject of this book.

Digital history encompasses diverse historical practices, such as digitisation efforts at archives, libraries and museums, computer-assisted research, web-based teaching and professional and public dissemination of historical knowledge, as well as research on the history of 'the digital', computers and

digital technologies. One comprehensive definition capturing this diversity of practices was suggested more than a decade ago in a discussion between digital historians in the *Journal of American History*:

> Digital history is an approach to examining and representing the past that works with the new communication technologies of the computer, the Internet network, and software systems. On one level, digital history is an open arena of scholarly production and communication, encompassing the development of new course materials and scholarly data collections. On another, it is a methodological approach framed by the hypertextual power of these technologies to make, define, query, and annotate associations in the human record of the past. To do digital history, then, is to create a framework, an ontology, through the technology for people to experience, read, and follow an argument about a historical problem.[2]

While the digital embraces the whole spectrum of the historian's craft, this volume focuses on digital history as a form of scholarly research that uses digital sources and tools to produce new historical knowledge. This form of digital history research is part of the larger digital turn in academia, identified as digital humanities, culture analytics, computational social sciences and other concepts related to utilisation of computer-assisted methods for research.[3] By bringing together research contributions to the new digital history from historians, computer scientists, computational linguists and other scholars producing new empirical historical knowledge using digital methodologies, as well as conceptually focused perspectives on critical issues of the field's past, present and future development, this book provides digital histories that we hope will be read as laudable exemplars from within the emergent digital history research community. The digital histories collected here simultaneously represent various methodological applications of and themes within digital history research and thus an attempt to take stock of current research rather than providing a pedagogical textbook or programmatic manifesto. The new digital history has matured enough for us to instead be able to present historical work currently furthering historical research. Thus, the studies in this book take digital history beyond discussions of its future potential, proofs of concept and pedagogical examples to instead focus on digital history 'in action', to the making of new historical knowledge.

   Through this focus on presenting results from digital history research projects, this book breaks new ground within the current wave of digital history. Other digital history books published so far have mainly been monographs focused on discussing how historians could use digital sources or methods to conduct and present research such as the pioneering *Digital history: a guide to gathering, preserving, and presenting the past on the web* (2006) by Daniel J. Cohen and Roy Rosenzweig, or anthologies such as *History in the digital*

age (2013) contributing discussions of the problems and possibilities of doing the new digital history, rather than research results of historical studies using new digital methodologies.[4] This is as expected, as it is only during the last couple of years that we have seen the first research publications using digital history research methodologies within mainstream academic historical publishing outlets. Matthew Jockers' *Macroanalysis* (2013) appears to be the first research monograph published by a university press, and Cameron Blevins' 'Space, nation, and the triumph of region' (2014) is the first peer-reviewed research article published in the *Journal of American History*.[5] In this way, this book aspires to pioneer and promote work within the new digital history by being a timely research anthology from the current third generation of digital historians that,[6] outside of digital spatial history,[7] focuses on contributing new historical knowledge from research using digital research methodologies.

Emergence of the New Digital History

The roots of exploiting modern data-processing equipment in humanities research date back at least to the 1950s, when Josephine Miles started using punch cards for literary analysis.[8] The development was continued in the 1950s with Father Roberto Busa utilising IBM mainframe computers and John W. Ellison using the UNIVAC I to produce lexical concordances.[9] Since then, computer-assisted history research has produced three 'generations', roughly following the advancement of computers and internet technologies. Simultaneously, there are continuities of parallel developments borrowed from, or developed together with, sister disciplines, such as text analysis in literary studies, statistical analyses in economic and social history, Geographic Information System (GIS) within geography, and digital image analysis in art history and visual studies. Allegedly, 'the first published work by an historian involving actual computerized research' came in 1963 with a 'scalogram analysis of voting patterns' in the British Parliament in the 1840s by William Aydelotte at the State University of Iowa.[10] A few years later, Viljo Rasila (Paju, this volume) did somewhat similar work in Finland. The first larger and more widespread application of computers was by the cliometricians of the 1960s, who were recognised as constituting the first generation of digital historians. They were followed by a second generation centred around the new 'personal' computers in the 1990s and were often seen as a part of the wider humanistic research field of 'humanities computing'.

The current third generation of digital history can be said to begin to emerge in the late 1990s and the early 2000s with the appearance of the first large digitised full text databases, such as Early English Books Online (EEBO) and Project Gutenberg,[11] and with the rebranding and expansion of humanities computing to digital humanities in the mid-2000s. Since the early 2000s, contemporary historians' toolkit has been expanded by an increasing volume of digitised sources and the swift development of computational analysis methods. This

was taking place at the same time as geographical history was going through a development from historical GIS (HGIS) to spatial history.

The snowballing growth in the amount of digital sources and the development of new research approaches and concepts has gradually increased the number of humanists using computational methods. One of the most frequently used concepts is *distant reading*, a perspective pioneered by the literary historian Franco Moretti. Distant reading can be understood as a counterpart to *close reading* that has been used extensively in humanities for distilling meanings from texts from the 1970s onwards. Distant reading has been used to extract meaningful patterns from textual sources, particularly when the number of texts are so numerous that it is impossible for a human to read them in a consistent manner.[12] The examples in this volume show that distant reading can also be a useful approach for exploring smaller amounts of text, as it provides another kind of approach to the texts in focus. Such machine or algorithmic reading provides 'another pair of scholarly glasses' and allows examining the sources from new perspectives. In the best case, close and distant readings complement each other.

Characteristic for this potentially paradigmatic digital history (Fridlund, this volume) is not just the introduction of new conceptualisations, such as 'distant reading', 'macroanalysis' or 'algorithmic reading', or the application of methodological tools such as topic modelling, but also the utilisation of novel practices for historians, new digitally augmented ways of working. Digital research brings along the collaborations of larger multidisciplinary group projects, the use of centralised technical infrastructures and machines. The changes that are taking place in history today are in several ways reminiscent of the changes that natural science disciplines such as physics and biology went through earlier with changeover from individuals' 'small science' tabletop experiments to interdisciplinary large team 'big science' collaborations.

The origin of this volume lays in an initiative to strengthen digital history research proposed by a collective of historians in Finland in 2015. That ambition was generously funded by the Kone Foundation through two interconnected projects 2016–2018, which brought together the majority of the authors in this volume. The first project, Towards a Roadmap for Digital History in Finland, aimed at identifying practical, professional and institutional obstacles and possibilities for developing digital history research. The second project, From Roadmap to Roadshow, built on the first one by bringing together digital historians to shape the best practices for disseminating knowledge about digital methods to historians so that in the end these would facilitate new digital history research. This was accomplished through a road tour to six major Finnish research universities, where the project organised presentations and workshops on emerging research and methodological developments within digital history.

Originally, the aim of the project was to end after the roadshow and to conclude with the subsequent publication of articles by the three main project researchers. However, the enthusiasm among the participants at the various

universities promised new digital research results in a not-so-distant future. Thus, towards the end of their shared work, the team decided to extend the project towards its logical conclusion by organising a workshop during the spring of 2018 that invited historians and specialists in digital research methods to come together to work collaboratively on formulating and answering a number of specific and concrete historical research questions. The historians who responded to the invitation brought their source materials and historical research questions, while the digital specialists contributed with their methodological expertise to jointly find answers to the research questions. At the workshop, the research teams analysed the sources to come up with preliminary solutions and answers and afterwards the teams were encouraged to keep working on their projects, and in this book, several of those projects are now brought to completion in the form of peer-reviewed research articles. They are complemented by articles from other digital historians, presenting results from a selection of the other recent research projects.

The majority of the research presented here is by digital historians active at Finnish universities. The rationale behind this is, in addition to the books' specific historical origins as explained above, that the emerging Finnish digital history community is both a representative and in many ways exemplary part of the larger international development of digital history. It is representative in that the used methodological research approaches correspond to the predominant directions of current digital history and thus the diversity and breadth of the studies presented in this volume, representing digital history research in a wide range of topics, from diverse disciplines from political, economic, cultural, intellectual and feminist history to history of science and technology and periods going from the Early Modern to the recent past. Taken together they provide a representative overview of the state-of-the-art of not just Finnish digital history research, but also of emerging digital history overall. Like most other research communities, the digital history landscape in Finland is diverse and dispersed, including bigger research groups, individual researchers and interdisciplinary and collaborative projects with national and foreign colleagues in Finland and abroad. This volume is exemplary in that digital history in Finland as a community and practice can be said to be more developed and institutionalised than in many other countries. In addition to several digital historians working at all levels of academic seniority, there are designated doctoral positions and professorships, textbooks, a regular digital history conference series and seminars and a digital history section within the national historical society. Compared to most other countries, the stage of digitisation of newspapers and archival documents is very advanced, which encourages digital history research. The common understanding of digital historians in Finland is that the focus of digital history research should be in finding answers to the research questions rather than utilisation of digital research tools just for their own sake. The contributions to this book, we feel, exemplify that critical evaluation of digital sources, metadata and research methods, and the results

they provide, are the basic components of good digital history research. Thus, the Finnish digital history research, together with the other contributions presented in this volume, should be a good representation of some of the most widely shared research practices emerging within the new digital history.

The New Digital and Distant Histories

This book contributes to advancing the field of history in primarily two ways, through new conceptual explorations of the past, present and future of digital history research and with new empirical historical knowledge coming out of research using digital methods. Through this, we aim to illuminate the new digital history's potential and pitfalls. We have divided the book into four parts. Part I 'The Beginning' consists of this introduction. Part II 'Making Sense of Digital History' starts with discussing the historical and methodological roots of digital history and contributes conceptual and contextual explorations of the current state of digital histories. Part III 'Distant Reading, Public Discussions and Movements in the Past' presents empirical case studies from various time periods that through the application of digital tools, primarily various forms of distant reading methodologies, demonstrate the further potential for expanding historical knowledge. The final Part IV 'Conclusions' draws the volume to an end by an exposition of the landscape of digital history and its future potential.

In the foreword, the late computer scientist and pioneering digital humanist Timo Honkela, draws on his wide experience of multidisciplinary cooperation using computational tools, to offer his thoughts on the digital future of history. In Chapter 2, providing a longer historical context for the new digital history, Petri Paju examines the history of computer-assisted history research from the 1960s until the 2010s. By focusing on one particular national development, that of historians' use of computers in Finland, he recognises how, although a particular national story, it was part of a larger, international and transnational pattern of development within digital history research.

After the overview of the roots of digital history, the subsequent chapters in Part II shed light on the fundamental components of digital history research: data, metadata and the mundane, often manual, work enabling the operation of our digital tools and resources. In Chapter 3, Jari Eloranta, Pasi Nevalainen and Jari Ojala exemplify how economic and business historians in many ways have been forerunners of digital history with computerised analysis of numerical and event code databases. They also share their experiences of the challenges to historical research of digitisation and uses of databases. Chapter 4 by Mats Fridlund attempts to conceptualise emerging historical practice by exploring the present state of digital history research according to two ideal types of digital history. Following Thomas Kuhn's theory of scientific revolutions, he describes them as 'normal' and 'paradigmatic' digital history. Further,

as a middle way between the two, he proposes one that is beyond the normal but still a less revolutionary form of semi-automatic digital history, described as 'digital history 1.5'.

This is followed by Chapter 5, which concerns research infrastructures, where Jessica Parland-von Essen calls for better data management and increasing the openness of data. She presents the FAIR (Findable, Accessible, Interoperable and Re-usable) approach to data, which would not only improve the efficiency, but also increase the trustworthiness and quality of historical research. The critical theme of the role of metadata in digital history research is taken up in Chapter 6 by Kimmo Elo, who points out that when focusing on data, we often neglect metadata, although it is a crucial part of the whole. In his chapter, he explores ways of improving the quality of the historian's metadata. Following this is a valuable reminder offered by Johan Jarlbrink in Chapter 7 on the importance of manual work to digital machine processing. In his chapter, he shows how digital research is far from automated, and that it actually requires countless hours of manual work which most of the time stays invisible and thus its problems and possibilities are often unnoticed and neglected.

The subsequent chapters offer a wide array of empirical case studies using a selection of digital research methods that exemplify how they can help us to reach for new understandings of the past. Beginning this series, in Chapter 8, Mirkka Danielsbacka, Lauri Aho, Robert Lynch, Jenni Pettay, Virpi Lummaa and John Loehr use statistical quantitative analysis to explore migration of Finnish individuals in the 20th century. Using a database that they have digitised and complemented with other historical data, they explore socio-demographic and environmental factors that can be combined with the domestic relocation and settlement of migrants. In Chapter 9, Heidi Kurvinen, in the vein of feminist history methodology, uses her personal experience of getting acquainted with historical text mining to explore traditional and not so traditional historians' experiences in encountering the new digital history methods. She notes that entering the field of digital history 'requires cultural and technological capital which marginalises researchers who do not have the skills to conduct digital analyses by themselves or do not have access to the organisational support'. Among the factors influencing the ability of researchers to participate, she identifies their gender. The next case study by Maiju Kannisto and Pekka Kauppinen in Chapter 10 illustrates the use of Named Entity Recognition (NER) to explore Finnish audio-visual history as it is presented in the public radio and television online archives. Their metadata analysis reveals interesting peculiarities in what kind of audio-visual imaginary of the past is provided by the dataset, and which elements of the national history it hides. In Chapter 11, Matti La Mela gives an excellent example of the opportunities of text analysis by tracing the history of the concept of *allemansrätten* (freedom to roam) in the Finnish parliamentary debates and argues counterintuitively to common knowledge that the present understanding of the concept has a surprisingly short history. His article also takes extra care in making the

methodological steps transparent to readers. In Chapter 12, Pasi Ihalainen, with the assistance of Aleksi Sahala, uses collocation analysis to study changes of the concept of 'internationalism' in 20th-century British parliamentary debates. By reconstructing the meanings attached to foreign political issues in the British Parliament from the early 19th century, they show that the 'international' has been associated in different ways during the various deliberations on the United Kingdom's membership in international organisations.

In Chapter 13, Melanie Conroy and Kimmo Elo, with the help of network analysis of the metadata of a large picture archive, explore the structure and temporal dynamics of the geospatial social networks of the East German opposition movement. They show how the network method can be used for exploring and visualising, as well as analysing, quantitative historical data. Reetta Sippola's contribution in Chapter 14 uses topic modelling to explore the evolution of the scientific discourse in the pioneering British scientific journal *Philosophical Transactions* in the mid-18th century. In her study, the method of arranging the data makes topic modelling reveal previously neglected themes and unnoticed temporal changes in the discourse. Heidi Hakkarainen and Zuhair Iftikhar also use this methodology in Chapter 15, in the expanded form of dynamic topic modelling, to focus on the formation of the concept of 'humanism' in the early 19th-century German-language press. They show how reaching reliable analysis results demands a deep understanding of the context, skills and time, but how the method has the potential to challenge established patterns of thought and underlying presumptions by providing a novel perspective on the sources. In Chapter 16, Reima Välimäki, Aleksi Vesanto, Anni Hella, Adam Poznański and Filip Ginter study author attribution and apply methods based on neural networks to explore their medieval cases of authorship recognition. Their intriguing results show how the uses of 'black-boxed' computational methods can potentially help us to solve centuries-long debates on the attribution of authorship. In the final case study in Chapter 17, Risto Turunen uses advanced collocation analysis to study Finnish labour newspapers during the late 19th and the early 20th centuries. With that material, he takes a macroscopic approach to study expressed temporality of the papers and especially the 'sun of socialism', which differed from the biblical sun shining on all and in this 'highlighted earthly problems'. Towards the end of his chapter, Turunen turns his discussion to the present situation and to future aims of digital history.

Jo Guldi concludes the volume in Chapter 18 by drawing a wide picture of the potential game-changing nature of digital history. She stresses the universal character and widely applicable nature of digital research methods: researchers of Chinese industrialisation can find a method used by a medievalist also useful to their research and vice versa. She also predicts that with the increasing number of digitised sources and utilisation of digital methods, we may see a rise of *longue durée* in history, which as she puts it could provide new findings that 'border on the breathtaking'.

New Historical Challenges and Criticisms

Digitisation and computer-assisted research tools open new possibilities, but also bring novel challenges and criticisms to the history discipline. There is a need for a wider methodological discussion on how digital research methods could and should be used in history research. To be able to take part in inter-disciplinary collaboration, it is important for historians to have a discussion on what digital methods mean, and where they can lead us. The ambition is that the studies in this book will contribute to foster this discussion. Among the critical components of digital history research that are addressed in the following chapters are digitisation of sources, creating metadata for digital source materials, human–computer interaction and digital research methods. These are only a few of the critical issues troubling current digital history.

One of the most pressing questions in digital history research is access to and problems of *digitised sources*. Although also important to scholars in other disciplines, they are fundamental to historians. The availability of consistent digitised collections with long time series is one of the critical prerequisites of digital research. Simultaneously, the existing digitised sources invoke discussions of their availability and usability, and what overall should be digitised. Furthermore, digitisation also changes the object of research, as a digitised newspaper is not the same as the physical object of a newspaper. When digitising sources, we, as Mikko Tolonen and Leo Lahti have pointed out, also lose important elements of the physical objects.[13] The consensus of the scholars contributing to this book is that the readily available digitised sources should be used with the same or even higher level of source criticism than before. While the existence of easily accessible digitised sources is a crucial requirement of digital history research, non-problematised use of data—a kind of 'source myopia'—has the potential to skew the historiography towards the most readily available databases and source material, rather than the most important or representative, and thus possibly motivate researchers to study them instead of the sources that, digitised or not, would provide the best answers to the research question (Chapter 3, this volume). For example, the very popular usage of newspapers as sources, especially for historical studies before the 20th century, is not necessarily because they are the most relevant historical sources, but is rather due to the simple fact that newspaper collections have in many countries been extensively digitised.

In digitising historical sources, the digitiser faces several practical choices that have extensive effects on historical research. The first major question is what to digitise. In making such basic selections, there is a threat of repeating and amplifying the biases of the past knowledge constructions, leaving less prominent and marginalised topics aside. The sources chosen to be digitised, the ways in which they are digitised and shared, have far-reaching consequences.[14] Memory organisations, such as archives and libraries, often begin their digitising efforts from sources that are most often used by the general public and

researchers, and are thus considered to be more important. This common prac-
tice creates a threat to further marginalise less prominent topics and to exclude
less studied materials. Therefore, alongside the use of digitised sources in read-
ily available collections, the ability for historians to digitise their own sources is
becoming an increasingly important skill. Learning how to digitise, and setting
the best practices for digitising, data life cycles, and sharing digitised sources
among historians are emerging important additions to the historian's toolbox.
To increase the variety of the available digitised sources, it is valuable that histo-
rians learn to digitise sources on their own, and whenever possible, share their
new data. The authors of this volume use both readily available databases and
sources that they themselves have digitised. Digitisation is time consuming,
and therefore the sharing of data is an important means of widening the base
of digitised sources.

In addition to digitised sources, a key issue for digital history research is
metadata, the data that describes and gives information about the digital data
(sources), and especially concerning its varying quality. As Kimmo Elo points
out in Chapter 6, 'more attention should be paid and more resources should be
invested in metadata creation'. From this perspective, the real problem is not
the structure of a data system itself (its 'ontology'), but rather the process of cre-
ating source material's metadata. The principles of adding metadata to the doc-
uments are often rather unsystematic and not transparent, and only too often
the usefulness of (meta)data depends on the person creating and inserting the
metadata. For example, at the workshop described above, one research team
planned to work on metadata of images from a public source database (www.
finna.fi). After some trials with that material, they ended up terminating their
project because of the overly scattered and random character of the metadata
collection. The large amount of processing necessary to enable digital methods
to be applied would not have made it possible to finalise their project within
a reasonable timeframe. However, this attempted project was not in vain, as it
partly inspired one of its participants (Elo) to write a chapter on metadata and
digital history for this volume.

The new kind of source material for historians in the form of digital data
and metadata makes it important for digital historians to develop a new
digital source criticism. Compared to the pre-digital era with large amounts
of data in non-digital forms, the contributions in this volume demonstrate
how digitisation instead of selective sampling allows historians to use all the
available data in their analysis, and thus more systematic analysis. Interestingly,
distant reading of large datasets often exposes the used databases' borders
and restrictions better than traditional sampling for close reading. For exam-
ple, the analysis of Kannisto and Kauppinen revealed the biases and partiality
of the studied dataset. Using digitised sources demands deep understanding of
what the data consists of because, as Eloranta, Nevalainen and Ojala point out
in Chapter 3, straightforward and non-problematising data usage may lead
to missing the key issues of the data and misleading interpretations of the

historical processes. In big data lie opportunities also for significant misinterpretations and falsifications.

As the contributions in this volume demonstrate, undertaking digital history research is often more time consuming and demands perhaps more conscious methodological choices than the traditional history approach. When one is undertaking digital history research, it becomes evident that alongside the algorithms used, the selection, creation, cleaning and filtering of the data heavily influence the results of the computer-assisted analysis. As Johan Jarlbrink shows, the digital research process at many stages demands manual work to be done, such as data cleaning. He demonstrates that this work is not only a necessary precondition for the analysis, but is actually in itself an important part of the analysis, as the researcher gets to work on and read through the material several times, and in this way learns to know the data in depth. While the quantitative digital analysis makes the conclusions more convincing, the in-depth knowledge of the data provides crucial qualitative understandings that guide the interpretations of the quantitative analysis.

Connected to this new source criticism, there is also a need to develop what has been described as a *digital resource criticism* (Chapter 4, this volume). This refers to the need, in order not to draw false conclusions, to be better aware of the internal technological logics of the digital resources used by historians, such as that of a database or a search engine. Similar questions of an awareness of the opportunities and limitations of the available resources and methods have lately been raised in reference to representation and visualisation of historical data.[15] One example of this is how Maiju Kannisto and Pekka Kauppinen in their study (see Chapter 10, this volume) found out that the frequencies of the search terms in the metadata did not reflect the actual frequencies of the audio-visual material to which the metadata referred, but that they were more an artifact of the processes of how the metadata had been produced. Both Elo (Chapter 6) and Kannisto and Kauppinen (Chapter 10) suggest in this book that archivists and historians should collaborate more and in this openly discuss the principles and practices of metadata formation, and how they could best serve all the parties.

Furthermore, the chapters of this book point out the methodological zigzag between distant and close reading of data, the repetitious adjustment of the algorithmic parameters, the evaluation of the means of the data formation, its broader context and preceding research, all involved in an overall research process of trial and error. Sippola, Kurvinen, and Hakkarainen and Iftikhar all show how the choices of the researcher influence the outcomes of the research. For example, when using topic modelling, the testing of the results with varying numbers of topics is a very important step in the process of analysis. Simultaneously, the scholar's understanding of the context is essential in identifying the meaningful results, and to be able to differentiate them from the potential nonsense produced by the computer, to discern the historical signal from the data noise. Usage of digital research methods amplifies the research findings, but

they also amplify the potential of false results. Computers and algorithms are important helpers, but they cannot operate on their own: they always require human guidance.

Despite all these challenges, the contributions to this volume demonstrate how computational analysis can disclose new and previously unnoticed patterns in history. For example, in Chapter 12, Ihalainen summarises the benefits of computer-assisted analysis for his study on conceptual history by stating that it revealed associations between the studied concepts, which made it possible to estimate trends in political attitudes and revealed particular and peculiar political issues that would have been very difficult to find with traditional methods. Along the same lines, Kurvinen states, in Chapter 9, that combining digital analysis and close reading allowed her to identify topics that might have remained unnoticed otherwise and exposed new ways of perceiving the material, ways that could prompt novel and previously unresearched questions.

The new digital history might also foster a wider rethinking of the parameters of historians' professional practice. Digital research methods create new and at times more stringent demands on accuracy, methodological thinking, self-organisation and collaboration than traditional historical research. As Kurvinen points out, digital environments could encourage historians to conduct their research in 'a more self-aware manner when every step of the process needs more thought than a traditional day with paper archives'. Similarly, Eloranta, Nevalainen and Ojala point out in Chapter 3 that collaborative research on digital data can lead to more efficient and accurate research, but it requires the development of a different professionalism from researchers. Jessica Parland-von Essen shows in Chapter 5 the importance of historians starting to manage their data in a more qualified manner to themselves so they become more like data curators and archivists, and including thinking about the preservation and reusability of research data from a longer-term perspective. To support the development of such new practices in historical disciplines, there is a need for historians to participate in developing new joint practices that support FAIR data and thus better research. This calls for collaboration among historians and memory organisation specialists, and for historians to reach outside of history to seek out ideas from other disciplines facing similar challenges.

Most of the chapters in this volume were written collaboratively. Along the process of our project, it was confirmed that digital history research demands interdisciplinary collaboration, since it is rare that a historian manages to combine in him- or herself both the skills of the historian and of the programmer. That said, it is not necessary for the historian to become a programmer. What is needed is the ability to collaborate and work together in an interdisciplinary manner with collaborators who bring expertise from the domains of computer and information science.[16] The above-mentioned workshop proved that fruitful collaboration with IT professionals is not only needed, but also feasible and beneficial. And this book proves that it can bring new knowledge, as well as conceptual developments, to the field of history.

One basic challenge, nevertheless, is that although multidisciplinarity is much-needed in the realm of digital humanities research, it is well known that not all computer-related questions or tasks carried out in digital history research are challenging enough to peak the interests of computer scientists. For example, the application of a ready code to a dataset is for a traditionally trained historian often too challenging a task, but rather trivial for a computer scientist. Thus, there is an increasing importance for universities and research institutions to be able to provide more mundane and routine technical support to historical researchers through their libraries, IT support facilities or other means, much like before the widespread availability of easily accessible online databases and online sources such institutional structures were central in assisting historians in finding research literature and source materials.

Conclusion

It seems evident that history research has been and will continue to be increasingly influenced by society's overall digitalisation. Still, the historians in general would benefit from being more aware than before of the interaction between historical research and the digitising world around them in order to stay both critical and constructive towards the changes and continuities of today. This includes taking advantage of the latest tools, as well as exploring their limitations to be able to keep our methods up to date and to gain a better understanding of the possibilities and pitfalls of historical research in the digital era.

As always, the future holds both promises and threats for historians, digital and otherwise. Although it is essentially an older condition, the skills and resources needed for digital history research could broaden the gap between history departments that are better positioned and those that are not, and consequently create more divisions among historians. One key issue for the digital historians is how to succeed to excel in using and developing new methods, while simultaneously avoiding overlooking the values of more traditional research. Doing and succeeding with the new explorations, while also respecting the older known and tried ways, has often shown to be the best working path towards the future.

In a similar vein as the encouragement by Jo Guldi and others in this book, one lesson from sociologists and historians of technology has been that users matter, that they, rather than being passive adopters of new technology delivered in black boxes, can have their say in influencing the direction of technological change, and at times even open up and reconstruct their tools so they better fit their particular needs and desires.[17] Historians as a group can and should be active in making choices and guiding their discipline towards an ever-more digital world of tomorrow, a tomorrow that soon will be a past and needs its born-digital history researched.

After almost 50 years, perhaps we have finally arrived at Emmanuel Le Roy Ladurie's 'tomorrow'. Or maybe we are already far beyond that—not least as most of the authors in this collection would not identify themselves as doing the quantitative kind of history Le Roy Ladurie expected future historians to be doing. As historians, we can recognise how difficult it is for history's actors to foresee future developments, and that while Le Roy Ladurie correctly predicted that historians needed to learn to harness computer technology for their work, neither he nor his colleagues could hardly have imagined the possibilities of the information technology at historians' disposal in the early 2020s. However, in the sense that historians should learn how to make the most of the 'computer', we feel that the historians in this book with their new digital and distant histories have tried to live up to his hopes by going towards and away from his tomorrow to reach our today and its past, present and future digital histories.

Notes

[1] Le Roy Ladurie 1979: 6. Rabb wrote: 'In I967, the basic posture of quantitative historians was a mixture of brashness and defensiveness. Le Roy Ladurie was sufficiently impressed by the discussions at Ann Arbor to predict that "the historian will be a programmer or he will be nothing"' (Rabb 1983: 591).

[2] William G. Thomas III quoted in Cohen et al. 2008: 454.

[3] See Jones 2014.

[4] For some of the major books published within the new digital history, see: Staley 2002; Cohen & Rosenzweig 2006; Galgano et al. 2008; Schmale 2010; Gantert 2011; Genet & Zorzi 2011; Haber 2011; Rosenzweig 2011; Clavert & Noiret 2013; Dougherty & Nawrotzki 2013; Jockers 2013; Weller 2013; Graham, Milligan & Weingart 2015; Bozic et al. 2016; Koller 2016; Brügger 2018.

[5] Jockers 2013; Blevins 2014. See also Guldi & Armitage 2014.

[6] As we well know, historical 'firsts' are often contested and contextual.

[7] The field of spatial history evolved from within Historical Geographic Information Systems research starting in the 2000s. See Gregory & Geddes 2014: x, xii, xiv–xv.

[8] Sagner Buurma & Heffernan 2018.

[9] Jockers 2013: 3; Vanhoutte 2013: 127–128.

[10] Swierenga 1970: 5.

[11] Although these collections also have much longer histories. See Lebert 2008.

[12] Moretti 2000, 2005, 2013. See also Underwood 2017.

[13] Tolonen & Lahti 2018.

[14] See, for instance, Jarlbrink & Snickars 2017.

[15] Foka, Westin & Chapman 2018.

[16] See also Fickers & van der Heijden 2020.

[17] See Oudshoorn & Pinch 2003.

References

Blevins, C. (2014). Space, nation, and the triumph of region: a view of the world from Houston. *Journal of American History, 101*(6), 122–147.

Bozic, B., Mendel-Gleason, G., Debruyne, C., & **O'Sullivan, D.** (2016, 25 May). *Computational history and data-driven humanities: second IFIP WG 12.7 international workshop, CHDDH 2016, Dublin, Ireland, revised selected papers*. Berlin and Heidelberg: Springer.

Brügger, N. (2018). *The archived web: doing history in the digital age*. Cambridge, MA: MIT Press.

Clavert, F., & **Noiret, S.** (Eds.). (2013). *L'histoire contemporaine à l'ère numérique—Contemporary History in the Digital Age*. Brussels: Peter Lang.

Cohen, D. J., & **Rosenzweig, R.** (2006). *Digital history: a guide to gathering, preserving, and presenting the past on the web*. Philadelphia, PA: University of Pennsylvania Press.

Cohen, D. J., Frisch, M., Gallagher, P., Mintz, S., Sword, K., Taylor, A. M., Thomas, III, W. G., & **Turkel. W. J.** (2008). Interchange: the promise of digital history. *Journal of American History, 95*(2), 452–491.

Dougherty, J., & **Nawrotzki, K.** (Eds.). (2013). *Writing history in the digital age*. Ann Arbor, MI: University of Michigan Press.

Fickers, A., & **van der Heijden, T.** (2020). Inside the trading zone: thinkering in a digital history lab. *Digital Humanities Quarterly, 14*(3). In M. Oiva & U. Pawlicka-Deger (Eds.), *Lab and slack: situated research practices in digital humanities*, special issue.

Foka, A., Westin, J., & **Chapman, A.** (Eds.). (2018). Technology in the study of the past. *Digital Humanities Quarterly, 12*(3), special issue. Retrieved from http://www.digitalhumanities.org/dhq/vol/12/3/index.html

Galgano, M. J., Arndt, C., & **Hyser, R. M.** (2008). *Doing history: research and writing in the digital age*. Boston, MA: Thomson Wadsworth.

Gantert, K. (2011). *Elektronische Informationsressourcen für Historiker*. Berlin: de Gruyter.

Genet, J.-P., & **Zorzi, A.** (Eds.). (2011). *Les historiens et l'informatique: un métier à réinventer*. Rome: École française de Rome.

Graham, S., Milligan, I., & **Weingart, S.** (2015). *Exploring big historical data: the historian's macroscope*. London: Imperial College Press.

Gregory, I. N., & **Geddes, A.** (2014). Introduction: from historical GIS to spatial humanities: deepening scholarship and broadening technology. In I. N. Gregory & A. Geddes (Eds.), *Toward spatial humanities: historical GIS & spatial history*. Bloomington, IN: Indiana University Press.

Guldi, J., & **Armitage, D.** (2014). *The history manifesto*. Cambridge: Cambridge University Press.

Haber, P. (2011). *Digital Past: Geschichtswissenschaft im digitalen Zeitalter*. Munich: Oldenbourg.

Jarlbrink, J., & **Snickars, P.** (2017). Cultural heritage as digital noise: nineteenth century newspapers in the digital archive. *Journal of Documentation, 77*(6), 1228–1243.

Jockers, M. L. (2013). *Macroanalysis: digital methods and literary history*. Urbana, IL: University of Illinois Press.

Jones, S. E. (2014). *The emergence of the digital humanities*. London: Routledge.

Koller, G. (2016). *Geschichte digital: historische Welten neu vermessen*. Stuttgart: Kohlhammer Verlag.

Lebert, M. (2008). *Project Gutenberg (1971–2008)*. University of Toronto and Project Gutenberg. Retrieved from http://www.gutenberg.org/ebooks/27045

Le Roy Ladurie, E. (1979). *The territory of the historian*. Translated from the French original (in 1973) by B. Reynolds and S. Reynolds. Brighton: The Harvester Press.

Moretti, F. (2000). Conjectures on world literature. *New Left Review, 1*(1), 54–68.

Moretti, F. (2005). *Graphs, maps, trees: abstract models for literary history*. London and New York, NY: Verso Books.

Moretti, F. (2013). *Distant reading*. London and New York, NY: Verso Books.

Oudshoorn, N., & Pinch, T. (Eds.). (2003). *How users matter: the co-construction of users and technologies*. Cambridge, MA: MIT Press.

Rabb, T. (1983). The development of quantification in historical research. *Journal of Interdisciplinary History, 13*(4), 591–601.

Rosenzweig, R. (2011). *Clio wired: the future of the past in the digital age*. New York, NY: Columbia University Press.

Sagner Buurma, R., & **Heffernan, L.** (2018, 11 April). Search and replace: Josephine Miles and the origins of distant reading. *Modernism/Modernity Print Plus*. Retrieved from https://modernismmodernity.org/forums/posts/search-and-replace

Schmale, W. (2010). *Digitale Geschichtswissenschaft*. Vienna: Böhlau Verlag.

Staley, D. J. (2002). *Computers, visualization and history: how new technology will transform our understanding of the past*. London and New York, NY: Routledge.

Swierenga, R. P. (1970). Clio and computers: a survey of computerized research in history. *Computers and the Humanities, 5*(1), 1–21.

Tolonen, M., & **Lahti, L.** (2018). Digitaaliset ihmistieteet ja historiantutkimus. In M. O. Hannikainen, M. Danielsbacka, & T. Tepora (Eds.), *Menneisyyden rakentajat: teoriat historiantutkimuksessa*. Helsinki: Gaudeamus.

Underwood, T. (2017). A genealogy of distant reading. *Digital Humanities Quarterly, 11(2)*.

Vanhoutte, E. (2013). The gates of hell: history and digital | humanities | computing. In M. Terras, J. Nyhan & E. Vanhoutte (Eds.), *Defining digital humanities: a reader*. Farnham and Burlington, VT: Ashgate.

Weller, T. (Ed.). (2013). *History in the digital age*. London and New York, NY: Routledge.