HANNU SALMI

What is Digital History?

Contents

- Series title
- Title page
- Copyright page
- Introduction
  - Notes
- 1 The Digital Past: Sources and Problems
  - Digitization of cultural heritage
  - Biases of the digital past
  - Born-digital era
  - Virality
  - Toward the epistemology of the born digital
  - Notes
- 2 Reading and Textuality in Digital History
  - Reading as a research method
  - Reading with machines
  - Distant reading
  - The challenge of big data
  - Texts for historical research
  - Distant reading techniques
  - Notes
- 3 Mapping and Viewing History
  - Maps as digital sources
  - Maps as historical interpretations
  - From visual evidence to distant viewing
  - Sound and vision

Visuality, textuality and metadata
Notes
4 Interdisciplinarity: Challenges for Research
Beyond disciplinarity
Trading across borders
Towards interdisciplinary practices
Notes
5 Presenting the Past in the Digital Age
Digital history for the public
Visualizing data
Presenting change over time
Augmented and mixed realities
Travelling in time
Notes
Conclusion
Notes
Further Reading
Index
End User License Agreement

Series title

What is History? series

John H. Arnold, *What is Medieval History?* 2nd edition

Peter Burke, *What is Cultural History?* 3rd edition

Peter Burke, *What is the History of Knowledge?*

John C. Burnham, *What is Medical History?*

Pamela Kyle Crossley, *What is Global History?*

Pero Gaglo Dagbovie, *What is African American History?*

Shane Ewen, *What is Urban History?*

Christiane Harzig and Dirk Hoerder, with Donna Gabaccia, *What is Migration History?*

J. Donald Hughes, *What is Environmental History?* 2nd edition

Andrew Leach, *What is Architectural History?*

Stephen Morillo with Michael F. Pavkovic, *What is Military History?* 3rd edition

James Raven, *What is the History of the Book?*

Sonya O. Rose, *What is Gender History?*

Barbara H. Rosenwein and Riccardo Cristiani, *What is the History of Emotions?*

Brenda E. Stevenson, *What is Slavery?*

Jeffrey Weeks, *What is Sexual History?*

Richard Whatmore, *What is Intellectual History?*

Merry E. Wiesner-Hanks, *What is Early Modern History?*

What is Digital History?

Hannu Salmi

polity

Copyright page

First published in 2021 by Polity Press

Polity Press

65 Bridge Street

Cambridge CB2 1UR, UK

Polity Press

101 Station Landing

Suite 300

Medford, MA 02155, USA

ISBN-13: 978-1-5095-3701-3

ISBN-13: 978-1-5095-3702-0 (pb)

A catalogue record for this book is available from the British Library.

Library of Congress Cataloging-in-Publication Data

Names: Salmi, Hannu, author.

Title: What is digital history? / Hannu Salmi.

Description: Medford : Polity Press, 2020. | Includes bibliographical references and index. | Summary: “The first student handbook on digital methods and sources for historians”-- Provided by publisher.

Identifiers: LCCN 2020014313 (print) | LCCN 2020014314 (ebook) | ISBN 9781509537013 (hardback) | ISBN 9781509537020 (paperback) | ISBN 9781509537037 (epub)

Subjects: LCSH: History--Computer network resources. | History--Methodology. | History--Research.

Classification: LCC D16.117 .S25 2020 (print) | LCC D16.117 (ebook) | DDC 907.2--dc23

LC record available at https://lccn.loc.gov/2020014313

LC ebook record available at https://lccn.loc.gov/2020014314

Typeset in 10.5 on 12pt Sabon

by Fakenham Prepress Solutions, Fakenham, Norfolk NR21 8NL

Printed and bound in the UK by TJ International Limited

For further information on Polity, visit our website: politybooks.com

Introduction

Our world is saturated with the digital. ‘To speak of the digital’, wrote Charlie Gere in his 2002 book *Digital Culture*, ‘is to call up, metonymically, the whole panoply of virtual simulacra, instantaneous communication, ubiquitous media and global connectivity that constitutes much of our contemporary experience.’[1] Today, even more than in 2002, the digital is all around us. When we speak of the digital culture, we are often referring to computers, mobile devices, information flows and the inescapable pervasiveness of social media.

Digitality tends to emphasize the present tense. What is happening now can be communicated faster than ever, and we seem to live in an extended now. During the last decades, however, the role of the digital in everyday life, which is called *digitalization*, has also affected our sense of the past, including the ways we explore history and communicate our findings. The aim of this book is to discuss this emerging field and its ramifications for historical research. As a concept, ‘digital history’ has been around for decades. However, during the 2010s, it has become a distinct branch of the study of the past, with many interdisciplinary ties through which it is continually positioned and re-positioned. Besides these considerations, throughout the forthcoming chapters, this book offers an introduction to a wide array of digital archives and digital history projects, drawing examples from various countries and continents.

Before discussing the concept ‘digital history’ and its implications further, it is necessary to provide an overview of digital culture and the history of computerization. Etymologically, ‘digital’ comes from the Latin word

‘digitalis’, which is derived from ‘digitus’, meaning finger or toe.[2] *Digitalis* has meant something done with the fingers. Eventually, this became the background for the meaning of ‘digit’, a number that is usually less than ten because it can be counted on the fingers. After the Second World War, ‘digit’ began to be used in reference to computers, which were essentially calculating machines that processed information in the form of digits. This was and still is called binary code. Only two digits are involved: zeros and ones. The word stem for the era of information technology, *digitalis* is echoed in many languages. However, there are also languages in which numbers are referred to despite the lack of a Latin connection. The French word for digital is *numérique*. In Russian, digital history is Цифровая история (tsifrovaya istoriya); Цифр (tsifr) means numbers. In Chinese, the notion is 數位史學 (shùwèi shǐxué); 數位 (shùwèi) refers to both a number (數) and a bit (位). When used together, they mean ‘digit’.

The emergence and expansion of ‘the digital’ is part of a historical process: the history of computerization, which began after the Second World War. If ‘digit’ originally referred to counting with our fingers, it soon began to be employed in more sophisticated calculation processes. Computers were machines for performing calculations when fingers, pens and paper proved insufficient. The early history of computerization was characterized by mainframe computers that universities and companies used to perform scientific and insurance calculations and to maintain large databases.[3] Society’s need for computing was estimated to be limited, but new technology opened new ways of thinking.

The words attributed to IBM chairman Thomas J. Watson in 1943 epitomize the difficulty in anticipating the future: ‘I think there is a world market for maybe five computers.’[4]

Whether Watson ever said this is questionable, but the statement has been employed to symbolize the rapid changes in not only computers but the mindset regarding information technology.[5] Many of Watson's contemporaries in the 1950s and 1960s tried to estimate the number of mainframe computers that would be needed for calculating processes in companies and state offices. The conclusion often reached was that the number of machines required would be small. Soon everything would change.

In the 1970s, the introduction of the microchip enabled the miniaturization of information technology.[6] In the 1980s, microcomputers became ubiquitous in offices and homes, and computers became part of everyday life.[7] Minitel was a French pioneering effort to provide homes with online services through telephone lines. It was launched in 1982 and became very popular with millions of people in France.[8] Soon, online connections would become the backbone of knowledge production and, thus, also a precondition for the breakthroughs in digital history. The end of the Cold War in the late 1980s to early 1990s opened the door for a global economy and transregional information flows. This was accompanied by advances in the communication technologies, including satellite and broadcasting networks and the Internet. These changes paved the way for many digital projects in the 1990s and early 2000s: from the digitization of historical materials to the provision of curatorial services, many of which are still in use. These will be discussed in greater detail in chapter 1.

The idea of 'the digital' was at the centre of these changes in the 1990s. Many economists believed that traditional industry, often referred to as the 'chimney industry', had come to an end. The future would be neither industrial nor post-industrial; rather, it would be knowledge-based.

Information would be a society's principal product. These were the essential thoughts behind projects such as the National Information Infrastructure promoted by the Bill Clinton Administration in the United States in 1991 under the leadership of Vice President Al Gore.[9] It popularized the idea of the high-speed data transfer process as the information superhighway. The hope that information would become the economic engine was soon adopted and promoted in other countries. In Europe, this was particularly the case after the 1994 Bangemann report, which proposed 'an action plan of concrete initiatives based on a partnership between the private and public sectors to carry Europe forward into the information society'.[10] During the 1990s, the change became global, with regional differences.[11]

The period after the early 1990s was characterized by the rapid expansion of online communication, especially after the adoption of the World Wide Web, which facilitated information access via graphical browsers, such as Mosaic, Netscape and, later, Explorer and Firefox. Graphic interfaces meant that more bandwidth was needed for downloads and uploads. The telephone modems of the 1980s were replaced by high-speed Internet access in the 1990s and wireless and mobile broadband and optical fibre connections in the 2000s.

With the emergence of the World Wide Web, navigating the Internet became much easier. Graphical browsers enabled the presentation of images and visualizations over the Internet, thus offering new ways for providing historical material. Soon, historians and other humanities researchers would stress the need to make the past accessible online. Returning to the background of 'digital history' as a concept, the following two examples illustrate the developments in Finland and the United States.

In Finland in the early 1990s, Internet-mediated history was described as *sähköinen historia* or *elektroninen historia*. Both words mean electric history. It was by no means in the mainstream of historical practice. Rather, it was an emerging trend that stressed the need to acknowledge online resources, to curate 'electric' services for historians and to use virtual platforms for teaching. In English, the notion of 'digital history' was increasingly used in the same sense. In 1996, the Finnish efforts led to the launch of the Agricola portal, a curatorial service addressing both scholars and students of history with the support of national memory organizations. In the United States, Edward L. Ayers and William G. Thomas used the term 'digital history' in 1997 in suggestions for the founding of a centre dedicated to the emerging field. The next year, it became known as the Virginia Center for Digital History (VCDH). On its original webpage, it defined its task as follows:

> VCDH's mission is to develop high-quality, well-researched, and reliable history materials for the World Wide Web and deliver them to schools, colleges, libraries, historical societies, and the general public. Our goal is to make history in a digital format, make it widely accessible, appealing, and useful.[12]

VCDH became one of the first centres for exploring the 'hypertextual power' of information technology and the ramifications of the advent of the Internet for historical practice.[13] Another pioneering centre was the Roy Rosenzweig Center for History and New Media (RRCHNM), which was founded in 1994 to support digital historians, classrooms and the wider public in their use of online historical content. Rosenzweig presented the most well-known definition of digital history:

> [It] is an approach to examining and representing the past that takes advantage of new communication technologies such as computers and the Web. It draws on essential features of the digital realm, such as databases, hypertextualization, and networks, to create and share historical knowledge.[14]

As Stephen Robertson has pointed out, the RRCHNM's mission regarding the use of digital technology and media was 'to democratize the past – to incorporate multiple voices, reach diverse audiences, and encourage popular participation in presenting and preserving the past'.[15]

Rosenzweig's agenda has influenced later definitions of digital history. In 2009, Douglas Seefeldt and William G. Thomas stated:

> On one level, digital history is an open arena of scholarly production and communication, encompassing the development of new course materials and scholarly data collection efforts. On another level, digital history 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.[16]

In the United States, digital history has been closely connected with public history: the effort to use the new media technologies to communicate with wider audiences and to facilitate oral history and folklore studies. The efforts to 'gather, preserve and present the past', to refer to the title of Daniel J. Cohen and Roy Rosenzweig's 2006 book, were a transnational phenomenon for the researchers and educators who tried to expand the practice of history.

There are other storylines, too, that must be considered in the definition of digital history. In the 2000s, the digital humanities became an amalgamation of approaches for

using information technology in the humanities. If the term ‘digital history’, which was coined in the 1990s, emphasized the nature of the discipline as an open platform for exploration and dissemination, the breakthrough in the digital humanities shifted attention to the earlier humanities researchers’ computational methods, such as ‘humanities computing’ in English or *historische Informationsverarbeitung* in German.

In the aftermath of the Second World War, the idea of using computers in humanities research began to emerge. Roberto Busa, an Italian Jesuit priest, initiated the project to create a searchable database for St Thomas Aquinas’ writings. He succeeded in persuading IBM to support his initiative, which started in 1949 and lasted until the 1970s.[17] Another early example of ‘humanities computing’ was cliometrics in the 1960s and 1970s. The technique was originally an effort to modernize economic history by using computational tools and drawing on vast amounts of data.[18] These two cases, Busa and cliometrics, have become almost canonized examples of the pioneering efforts in computer-assisted history, whose multiple roots have often been either forgotten or obscured by language barriers.

These divergent paths in early digital history are still waiting to be written about. As pointed out by historian Petri Paju, for example, in Sweden in 1966, Carl Göran Andræ wrote about the benefits of computers for historians. In Finland, Viljo Rasila used computer-based factor analysis in his 1968 monograph on the Finnish Civil War. In Estonia in 1971, Juhan Kahk and Enn Tarvel discussed the possibilities of computerized historical analysis.[19] These examples indicate that computers were increasingly being used in historical research from as early as the 1960s; however, only a few of these efforts have received international attention.

In the 2000s, the continuation of the pioneering work of ‘humanities computing’ is vital. An ever-expanding amount of information is available in digital formats. Researchers must gather, organize and manage large amounts of data; develop new data analysis methods and draw conclusions; and present their results through the use of digital tools and platforms. Researchers can produce the data and preserve these corpora for replication by future scholars. They can also explore the origins of existing collections and the information that can be learned about and through them.

The breakthrough in the digital humanities in the 2000s has also changed the landscape of digital history. Melissa M. Terras has described the digital humanities as being ‘at the intersection of digital technologies and humanities’ and aiming to ‘produce and use applications and models that make possible new kinds of teaching and research, both in the humanities and in computer science (and its allied technologies)’. As pointed out by Terras, there has been a huge expansion of this interdisciplinary research cluster over the last ten to fifteen years.[20]

Digital history also lies at the intersection of disciplines. In emphasizing the study of the past and focusing on historical problems, digital history is more discipline-based than the digital humanities, which are a wider cluster of research settings and paradigms. Digital history acknowledges its origins in historians’ serious efforts to engage with the Internet, digital tools and information technology. It is also a repository for the computational methods that were developed in the digital humanities and can be applied in and refined for solving historical problems. The definition of digital history can today be reformulated as follows: digital history is an approach to examining and representing the past; it uses new communication technologies and media applications and

experiments with computational methods for the analysis, production and dissemination of historical knowledge.

Digital history is currently a vibrant field of historical practice with a plethora of approaches, projects, publications, services and sources. It also connects with the wider, more general questions that are relevant for any field of inquiry, including the quality and critical assessment of research and the issues related to open access. The five chapters in this book concentrate on five areas that characterize digital history today. The first chapter, ‘The Digital Past: Sources and Problems’, focuses on the digitization of history. It discusses several digitization projects and the outcomes that have influenced the ways in which we conceive the past. Born-digital history, which refers to the ontological problem of having the history of our own time exist primarily in the digital realm, is addressed. The second chapter, ‘Reading and Textuality in Digital History’, deals with texts and textuality, which have always been important for historians and historical writing. The chapter foregrounds the question of reading for two reasons. On the one hand, close reading and interpretation have been at the heart of history as a discipline. On the other hand, precisely this, the ‘closeness’ of reading, has been challenged by the idea of distant reading, which began to influence the digital humanities in the 2000s. Chapter 3, ‘Mapping and Viewing History’, creates some distance from textuality and focuses on the visual in digital history. In the 1990s, maps and mapping were groundbreaking in the field. The sophisticated techniques that draw upon geospatial data and cartographic applications are evidence of their continuing significance. The chapter also explores methods for incorporating the visual and audio-visual sources that are increasingly employed by digital historians.