Registered in Germany Europa Revue Journal
ISSN: 1862-6564 Vol. 6 No. 10 (2026) Published October 11, 2026

DIGITAL HISTORY AND ITS APPLICATION IN CONTEMPORARY HISTORICAL RESEARCH

SHARIPOVA GUZAL USMANBEKOVNA

Doctor of Philosophy (PhD) in Historical Sciences Department of History and Source Studies of Islam – IRCICA, International Academy of Islamic Studies of Uzbekistan11, A. Kadriy, 100011, Tashkent, Uzbekistan

g.sharipova@iiau.uz

Abstract

This research paper systematises the theoretical and methodological foundations and practical applications of digital history and historical informatics. The paper explores the conceptual differentiation of the sub-discipline, the genesis of the digital humanities, the principles of digital hermeneutics, and the classification of historical sources according to types of information modelling. Based on an analysis of the sources, the study describes the use of intelligent data analysis algorithms, machine handwriting and optical character recognition technologies, text attribution methods such as BERT and Ithaca, and search-augmented generation architectures. Particular attention is paid to spatial analysis in geoinformation systems, 3D and 4D virtual reconstruction technologies for cultural heritage objects, the functioning of digital libraries and archives, as well as the specifics of working with search engines and the phenomenon of the deep web.

Keywords

Digital history, historical informatics, digital hermeneutics, remote reading, the RAG approach, artificial intelligence, geographic information systems, virtual reconstruction, digital archive, the deep web.

In the present day, the digital transformation of research practices is a defining factor in the evolution of modern historical science. The development of information and communication technologies and the introduction of algorithmic tools have transformed traditional methods of working with source material, giving rise to an interdisciplinary field at the intersection of source studies, computer science, mathematical modelling and linguistic analysis.

The subject of digital history is the patterns governing the creation, analysis, preservation and representation of digital and digitised historical sources. Within the conceptual framework of the discipline, it is essential to draw a distinction between the institutionalised sub-discipline of Digital History which has its own academic publications and educational programmes and the broader process of digital transformation, which affects the day-to-day practice of every researcher.

The aim of this study is to systematise the theoretical and methodological foundations of digital history and to conduct a comprehensive analysis of its practical applications in source studies, field archaeology, archival science, spatial analysis and the research and education sectors, based on the analysed materials. To achieve this aim, the following research objectives are addressed: to elucidate the origins and interdisciplinary links of digital history and historical informatics; to investigate the principles of digital hermeneutics and approaches to the formalisation of sources; to analyse methods of intelligent data analysis and artificial intelligence technologies in source studies; to characterise the role of geographic information systems and 3D and 4D modelling in the reconstruction of cultural heritage; to assess the infrastructural potential of electronic libraries, digital archives and web search algorithms.

The origins of digital history can be traced back to the digital humanities movement of the 1950s and 1960s, which developed primarily within the fields of computational literary studies and the Text Encoding Initiative. The historical development of this field comprises several key stages: the era of quantitative history and cliometrics in the 1960s and 1970s, which relied on the use of large mainframe computers for the statistical processing of mass sources such as population censuses and tax registers; the "microcomputer revolution" of the 1980s to the early 1990s, linked to the spread of personal computers, database management systems and graphical web browsers; and the phase of institutionalisation from the mid-1990s, marked by the founding of the Centre for History and New Media by R. Rosenzweig's Centre for History and New Media (1994) and the establishment by E. Ayers and W. Thomas of the Centre for Digital History at the University of Virginia with the project "The Valley of the Shadow" (1997). Two main strands can be identified within the structure of contemporary digital

history. The public-communication strand builds on the traditions of social and public history, utilising new digital media to democratise historical knowledge. The research and computational strand builds on the legacy of quantitative history and historical demography, utilising algorithms for the automated collection of document corpora, the shift from “close reading” to “distant reading”, the extraction of structured features, and multi-faceted data visualisation. Within the Russian academic tradition, the status of a fundamental analytical discipline is retained by historical informatics, the institutionalisation of which is linked to the work of the Commission on the Application of Mathematical Methods and Computers at the History Department of the USSR Academy of Sciences, led by Academician I. D. Kovalchenko, and to the founding in 1991 of the “History and Computer” Association.

The development of digital history is intertwined with the emergence of the principle of critical digital literacy. Digital databases and archival platforms are not neutral repositories, as the selective nature of digitisation is driven by commercial interests and the geographical imbalance of institutions in the Global North. This gives rise to the “Matthew effect”, whereby digitised collections are cited more frequently and become increasingly enriched, whilst non-digitised collections fall out of scholarly circulation, leading to the homogenisation of historiography.

The transfer of a researcher’s observations from historical sources into machine-readable, formalised datasets constitutes a procedure of digital hermeneutics a methodological act of abstraction and source-critical interpretation. In accordance with the principle of information modelling, historical sources are classified into five groups:

Firstly, statistical sources, which are processed in spreadsheets using mathematical operations, summarisation and the creation of charts;

Secondly, structured sources, systems of diverse data types modelled in relational database management systems “Access, KLEIO” through the establishment of relationships and cross-queries;

Thirdly, narrative sources: materials characterised by a high degree of incompleteness and ambiguity, analysed using the KLEIO expert system and the TACT full-text system with the Label Markup, COCOA and BYU markup languages;

Fourthly, cartographic sources, which are modelled in geographic information systems by linking spatial observations to coordinates;

Fifthly, illustrative sources, which are processed using 3D modelling, digital restoration and multispectral imaging, requiring the historian to possess multimodal literacy.

Data formalisation draws on M. Taller's source-oriented approach, which aims to transfer the structure of the original text in its entirety without prior reduction, and the problem-oriented approach, which focuses on the rigorous selection of parameters tailored to a specific hypothesis. When working with weakly structured texts, the academic standard for formalisation is based on the OHCO concept and the declarative markup languages SGML, TEI and XML. Digital texts are divided into documents that have traditional paper equivalents, such as letters, books and diaries, and primary digital objects without paper prototypes.

Since the early 2000s, digital history has integrated data science methods for working with Big Data, as described by the 6V conceptual framework. To extract hidden patterns, intelligent analysis methods are applied: data mining, text mining, web mining and convolutional neural networks. In source criticism, specialised algorithms have been developed to solve information-related problems: the three-component COEL system by C.Harvey and E.Green for identifying individuals and linking records; logistic regression using King and Zeng's method for analysing rare discrete events; Gary King's ecological inference method for overcoming the "ecological fallacy" when working with aggregated data; and the method of attributive temporal boundary alignment in historical GIS.

Artificial intelligence, a term coined by D.Mc Carthy in 1955 – acts as a methodological assistant to the historian, combining the capabilities of "close" and "distant reading". The key areas of application for artificial intelligence in working with historical and cultural heritage are:

-Optical OCR and neural network-based HTR;

-Text attribution and source dating;

-3D reconstruction of monuments and image processing using generative adversarial networks.

The RAG approach is used to eliminate factual errors and "hallucinations" in generative language models. In this architecture, the neural network extracts data strictly from a verified, digitised corpus of documents, grounding its conclusions and providing direct links to the original sources.

Geographic Information Systems are a class of software designed for the collection, storage, analysis and photorealistic visualisation of spatially localised geodata. The conceptual precursor to spatial analysis is John Snow's cholera outbreak map, whilst the concept of GIS itself was developed under the leadership of Roger Tomlinson in the 1960s. In historical and archaeological research, GIS and services such as Google Earth Pro enable the construction of archaeological information systems, the modelling of palaeoreliefs and the spatial triangulation of monument protection zones based on the universal "in situ" principle.

In the field of virtual reconstruction, 3D modelling introduced by I. Sutherland in 1963 and further developed by P. Reilly within the concept of "virtual archaeology" is evolving from simple visualisation towards spatio-temporal analysis. The model serves as an independent object of research, enabling the calculation of structural strength characteristics within the ArchiCAD and AutoCAD environments. The addition of a time axis to the three-dimensional model forms a 4D reconstruction, illustrating the stages of construction, alterations and demolition of the structure over time. The practical algorithm for 3D modelling of monuments comprises four sequential steps, the collection and analysis of source data; the drafting of a two-dimensional vector plan in AutoCAD; the construction of a three-dimensional polygonal mesh and texturing in 3ds Max; finalising the 3D object and carrying out a virtual reconstruction of the historical landscape. The Agisoft Metashape software suite is used to generate point clouds from digital photographs, whilst the MatchPhoto function in SketchUp is used to reconstruct three-dimensional geometry from archives of historical photographs.

The network infrastructure of historical science is based on Internet technologies, which evolved from the ARPANET and the World Wide Web, created by T. Berners-

Lee at CERN in 1989–1990. The long-term preservation of historical copies of network segments is ensured by the non-profit Internet Archive, founded by B. Kayle in 1996. The degree to which information is machine-readable is described by Tim Berners-Lee's "5-Star Linked Open Data" system (ranging from a scanned PDF at 1 star to RDF semantic triples and URIs at 5 stars; author's note).

V. Arms defines an electronic library as a managed collection of information stored in digital formats and accessible via a network. Storing data in digital form reduces the costs of maintaining collections and ensures the preservation of the originals. An electronic archive is a system that supports a multilevel structure for describing documents according to the "Collection, Inventory, File, Document" scheme. The real-time assessment of the effectiveness of scientific publications relies on alternative metrics that emerged in 2010 and track item-level statistics on downloads, views, and citations in the online environment.

Search engines are used for navigation in the digital environment. The architecture of a search engine includes indexing and query processing, which are carried out by agents, spiders, and crawlers. No single system covers more than 30% of the World Wide Web, since a significant amount of information resides on the Deep Web, which is inaccessible to standard indexing. The structuring of search queries relies on Boolean logic operators such as AND, OR, and NOT, proximity operators such as NEAR, W/n, phrase search in quotation marks, pattern truncation ("*****"), and document field restrictions. In particular, Google Scholar is used to search for scientific literature; it ranks materials based on full-text analysis, authority, and the PageRank algorithm. At the same time, commercial search engines are unsuitable for direct retrospective studies due to the regular updating of indexes, the removal of outdated pages, and the systematic shifting of web documents' timestamps from the past to the present, which leads to the erasure of the historical trace of the online environment.

In conclusion, it should be noted that the following conclusions have been drawn on the basis of the analysis of the sources. Firstly, digital history is an evolving interdisciplinary field that is transforming academic standards of evidence and requires researchers to possess critical digital literacy whilst rejecting the perception of

algorithmic platforms and databases as ‘black boxes’; secondly, the procedure of digital hermeneutics and varied approaches to the formalisation of sources enable statistical, structured, narrative, cartographic and illustrative artefacts to be effectively converted into machine-readable datasets without losing their original context; thirdly, the integration of mathematical and statistical models, OCR technologies, neural network architectures and the grounding of linguistic model inferences based on the RAG approach ensures data verification and prevents the spread of factual errors; fourthly, the use of geographic information systems, photogrammetry and 4D modelling elevates the study of archaeological and architectural monuments to the level of spatio-temporal simulation analysis, ensuring the digital conservation of cultural heritage; fifthly, network infrastructure, digital libraries, institutional digital archives and academic search engines form a unified environment for scientific communication, breaking down geographical barriers and ensuring access to digitised collections.

LIST OF SOURCES

1.Бородкин Л. И., Жеребятьев Д. И. Технологии 3D-моделирования в исторических исследованиях: от визуализации к аналитике //Историческая информатика. Информационные технологии и математические методы в исторических исследованиях и образовании. – 2012. – №. 2. – С. 049-063.

2. Boot P. Distant Reading. Franco Moretti. – 2015.

3. Володин А. Ю. Цифровая герменевтика исторического источника: формализация как толкование //Вестник Пермского университета. Серия: История. – 2025. – №. 2 (69). – С. 87-100.

4. Гарскова И. М. Методика сетевого анализа в исторических исследованиях: специфика проблематики и источниковой базы исследования //Исторические исследования в контексте науки о данных: информационные ресурсы, аналитические методы и цифровые технологии. – 2020. – С. 217-229.

5.Бизенков Е. А. Практическое применение поисковой и наукометрической платформы GOOGLE SCHOLAR (АКАДЕМИЯ GOOGLE) //Международный журнал прикладных и фундаментальных исследований. – 2017. – №. 10-1. – С. 9-15.5.

6.Калмыков М. А., Медникова О. В. История развития поисковых систем и алгоритмы их работы //Научные известия. – 2022. – №. 28. – С. 33-35.7.

7. Шарипова Г. Цифровая история. Учебное пособие / под ред.— Ташкент: GRAND KONDOR PRINT, 2026. — 539 с. [1–538]

8. Salmi H. What is digital history?. – John Wiley & Sons, 2020.

9.Croft W. B. et al. Search engines: Information retrieval in practice. – Reading : Addison-Wesley, 2010. – Т. 520. – С. 131-141.

10. Haider J., Sundin O. Invisible search and online search engines: The ubiquity of search in everyday life. – Taylor & Francis, 2019. – С. 160.