



Title:	State of the Field: Digital Legal History. A (Western) methodological overview.
Author:	C.A. Romein
Issue:	3
Year:	2024
DOI:	10.21825/dlh.91695
Keywords:	Digital Legal History; Projects; Methodology; Historiography; Overview
Abstract:	<p>This article presents a comprehensive analysis of the methodological advancements in Digital Legal History (DLH), focusing on pivotal processes such as digitisation, datafication, and the evolution of ontologies. It demonstrates how digitising legal sources preserves extensive archives while datafication transforms them into structured, machine-readable formats, facilitating more sophisticated analysis. The article provides a detailed discussion of the role of ontologies in linking legal data across platforms. It outlines how the enhancement of cross-disciplinary research through interoperability is being facilitated. Furthermore, the integration of large language models (LLMs) is investigated, illustrating their capacity for automating processes such as text analysis and the summarisation of legal texts. However, challenges associated with potential bias and accuracy are also addressed. The article provides insight into how these digital methodologies are reshaping the study of legal history, facilitating new research possibilities, interdisciplinary collaborations and broader access to legal materials.</p>

State of the Field: Digital Legal History.

A (Western) methodological overview.¹

C. Annemieke Romein

Huygens Institute for the History and Culture of the Netherlands, Amsterdam/

UTwente, Enschede

The study of law has evolved into a distinct and autonomous discipline, essentially detached from its traditional interconnections with history, politics, and philosophy. This transformation was significantly influenced by the rise of legal positivism in the 19th century, which emphasized the separation of law from moral and philosophical considerations, and by the concurrent process of codification that sought to systematize and rationalize legal systems into comprehensive, written laws. This transformation can be traced back to the nineteenth century when modern university structures began to isolate legal studies from the broader fields of academia. Nevertheless, legal history remains closely associated with the legal and historical sciences. In German-speaking countries, it is taught as a subject within law faculties, where it is considered one of the foundational disciplines, alongside legal philosophy, sociology of law, and legal methodology.² The establishment of specialized law faculties and curricula signified a shift in legal studies towards a more concentrated acquisition of juridical knowledge and practical abilities pertinent to the legal profession.

The traditional function of legal history has been to illuminate the genesis and evolution of existent legal systems, including earlier stages, to inform their ongoing development.³ It employs a dual approach, addressing the practical interests of legal professionals and scholars while being shaped by the predominantly national configurations of contemporary law. Notwithstanding the increasing specialisation in legal studies, legal sources continue to capture the interest of historians, philosophers, political scientists and sociologists. Nevertheless, the separation of these disciplines has frequently resulted in a lack of integration of their insights into the study of legal history.

Since the 1960s, legal history has become increasingly influenced by questions, methods, and perspectives from other disciplines. This reflects broader trends in legal scholarship and

¹. This contribution was initially used as the 'Introduction' to the session on Digital Legal History at the *Rechtshistorikertag* in Frankfurt (16-20 September 2024). I have to thank Prof. Dr. Benadetta Albani for an excellent discussion over dinner that sparked the idea for the title. I would also like to thank the speakers on the panel "Digital Humanities und Rechtsgeschichte" panel at 44. *Rechtshistorikertag 2024* (in no particular order): Andreas Wagner, Ida Nijenhuis, Nicholas Cole, Stefan Kerkemeyer, and Tobias Hodel. Their papers were crucial in developing my thoughts and ideas for this paper.

² Andreas Wagner, 'Rechtsgeschichte', in *Clio Guide – Ein Handbuch zu Digitalen Ressourcen für die Geschichtswissenschaften*, ed. Claudia Prinz et al., 3. erw. und aktualisierte Aufl. (Berlin, 2023), <https://doi.org/10.60693/bsme-m777>.

³ Wagner.

embraces fields such as sociology, literary studies, and economics. In addition to the long-dominant focus on the history of legal dogmatics, new approaches such as the cultural turn, pragmatic turn, and material turn have also influenced legal history, leading to a diversification of its subjects and methods.⁴ This diversity is particularly evident in contexts beyond the legal faculty and in international contexts, where the functional link between legal history and the development of law is less pronounced and where formal legal education is not necessarily a prerequisite for those engaged in legal history research. Legal history is characterized by a diverse range of researchers and educators who do not seldom possess expertise from various disciplinary backgrounds. This dynamism and responsiveness to the complexities of law in its historical context are key attributes of legal history.

Digital Legal History (DLH) could play a pivotal role in adding to the interdisciplinary connections of legal studies and to posing different kinds of research questions. DLH fosters a collaborative environment where scholars from diverse fields can engage in cross-disciplinary exchange by providing researchers with online access to vast legal source materials and promoting methodologically-focused conferences. Additionally, DLH leverages modern digital techniques such as text recognition, text mining, and digital editing, which are crucial for processing and analyzing legal texts. These methods, shared with fields like literature and general history, enable the extraction of valuable insights from legal documents that are not only legal sources but also integral to understanding the social practices embedded within them. Despite the potential of digital tools like legal informatics and empirical legal studies, their application in legal history remains somewhat underutilized, largely due to the complexity and historical nature of the sources involved.⁵ Nonetheless, DLH holds great promise in advancing the study of legal history through these innovative digital approaches.⁶

This article seeks to provide an initial methodological overview of the various digital methodologies employed within the legal history field. The aforementioned techniques are elucidated succinctly, accompanied by illustrative examples of completed projects. It should be noted that this overview is not exhaustive, given the rapid pace of development in this field. Concurrently, researchers engaged in the field and embarking on novel projects seek exemplars and case studies to inform their pursuits. As digital humanities, and hopefully, this also applies to digital legal history as part of this, frequently entails collaboration, discourse, and consultations with colleagues, I hope that this contribution will serve as a reference point for articulating the sought-after techniques and methodologies in a manner that fosters a shared understanding of objectives and aspirations.

⁴ Wagner.

⁵ Stephen Robertson, 'Searching for Anglo-American Digital Legal History', *Law and History Review* 34, no. 4 (November 2016): 1047–69, <https://doi.org/10.1017/S0738248016000389>; Florenz Volkaert, 'OK Computer? The Digital Turn in Legal History: A Methodological Retrospective', 18 June 2021, <https://doi.org/10.1163/15718190-12340011>; Wagner, 'Rechtsgeschichte'.

⁶ Max Kemman, *Trading Zones of Digital History* (De Gruyter, 2021), <https://doi.org/10.1515/9783110682106>; Volkaert, 'OK Computer?'

1. What is DLH?

DLH can be defined as an emerging field that combines traditional legal history with digital tools and methodologies to enhance the study, analysis, and interpretation of legal sources. It employs digital archives, databases, and technologies such as text mining, data visualization, and computational analysis to examine legal documents, court cases, statutes, and other historical legal materials. These innovations facilitate novel approaches to exploring the historical development of legal systems, the relationship between law and society, and the broader impact of legal practices over time. By facilitating interdisciplinary research, DLH makes legal data more accessible and analyzable on a large scale.

This approach to Digital Legal History (DLH) represents a paradigm change in legal historical research, shifting from traditional methodological frameworks to more technologically-enabled and data-driven approaches. Rather than constituting an unconnected field of study in its own right, DLH should be regarded as a methodological approach that provides legal historians with sophisticated tools for collecting, analysing, and interpreting historical legal sources. The fundamental objective of legal history, which encompasses examining legal evolution, its societal ramifications, and the historical milieu in which it occurs, remains unaltered. DLH complements and expands the traditional legal history framework, facilitating the exploration of more extensive datasets and posing new research questions. It is, therefore, most accurately conceptualized as a methodological innovation that applies digital techniques to legal historical research rather than as a distinct field with its subject matter. However, this situation might change in the future. To gain understanding of DLH, it is necessary to situate it within the broader context of Digital Humanities and Digital History.⁷ These fields have gained prominence since the 1990s, with the advent of digital technologies in academic research. The field of Digital Humanities emerged in response to the growing use of computational tools in disciplines such as literature, archaeology, and history. Digital History, in particular, has concentrated on how digital tools can transform historical research, enabling new forms of analysis, visualization and accessibility. As a branch of this movement, DLH draws on these developments to revolutionize the study of legal history, transforming how legal texts and records are accessed, examined, and interpreted.

Notwithstanding these developments, the field of law has, on the whole, been seemingly slower to adopt digital methodologies than other humanities disciplines, such as history or literature.⁸ Nevertheless, there are notable exceptions, particularly in projects focused on digitizing legal sources, where legal scholars and historians have begun to engage with digital tools.⁹ Though not yet pervasive, these initiatives illustrate the incremental incorporation of digital techniques into legal history, thereby contributing to the evolving landscape of DLH.

⁷ C. Annemieke Romein et al., 'State of the Field: Digital History', *History* 105, no. 365 (2020): 291–312, <https://doi.org/10.1111/1468-229X.12969>.

⁸ See for more: Stephen Robertson, 'Digital Humanities', in *The Oxford Handbook of Law and Humanities* (Oxford University Press, 2019), <https://doi.org/10.1093/oxfordhb/9780190695620.013.1>.

⁹ H de Jong, 'Over Het Nut van Digitale Archiefontsluiting. De Civiele Procesdossiers van Het Hof van Holland En Het Hof van Friesland', *Pro Memorie* 24, no. 1 (June 2022): 97–111, <https://doi.org/10.5117/PM2022.1.005.JONG>.

Although comprehensive overviews of DLH are still in their infancy, some significant historiographic contributions, such as Andreas Wagner's work, offer valuable insights into applying digital methodologies within legal history.¹⁰ Wagner's historiographic analysis reflects legal history's gradual but significant engagement with digital tools, emphasizing the importance of ongoing collaborative efforts in developing a more complete understanding of DLH. Consequently, although detailed accounts of DLH's evolution are still emerging, foundational work such as Wagner's provides a crucial starting point for future scholarship.

2. Digital Legal History – methodologies and approaches

In recent decades, there have been significant developments in the methodology of legal history, with each stage building on the previous to enhance the ways in which legal historians engage with historical sources. The initial focus of DLH was on the digitization and preservation of legal materials, intending to create accessible digital archives. The organization of these materials through the use of metadata has enhanced the accessibility and management of data. At the same time, techniques such as text mining and natural language processing (NLP) have enabled scholars to derive valuable insights from extensive legal corpora.

However, the creation of such databases is contingent upon the mass capture and processing of textual sources. The field of legal history has a robust interest in automated text recognition, as well as in analytical methods such as corpus linguistics, paraphrase and citation detection, and text classification. These approaches are also employed in other disciplines, such as literary studies. However, legal history is distinctive in using text-based sources to explain social practices that often extend beyond the text itself. It is wrong to consider legal texts as mere indicators of external phenomena; they are, in fact, inextricably linked with the legal effects they generate, frequently serving as legal sources in a technical sense.

As DLH is developing, introducing new techniques, such as reconstructions and data visualizations, has enabled researchers to interpret and represent legal data in innovative ways.¹¹ Network analysis has also facilitated the mapping of complex relationships within legal systems. These developments are complemented by digital editing techniques, which have already facilitated the publication of important resources such as text collections and dictionaries. In contrast to general history, legal history is particularly concerned with normative, counterfactual, or conceptual aspects, focusing less on concrete individuals and more on roles, authorities, and the legitimations and qualifications of actions.

¹⁰ Wagner, 'Rechtsgeschichte'.

¹¹ Methodologically, we have to consider that a digital version of an original is, methodologically something different than the original and hence, should be treated e.g. cited, differently. For more methodological considerations, please see: Hannu Salmi, *What Is Digital History?, What Is History?* (Cambridge, UK: Polity Press, 2021); Pascal Föhr, 'Historische Quellenkritik Im Digitalen Zeitalter.' (Basel, 2018), http://edoc.unibas.ch/diss/DissB_12621.pdf.

Furthermore, developments in interoperability and data-sharing standards have facilitated the advancement of DLH, enhancing the accessibility of research findings across diverse platforms. While methods from fields such as legal informatics, empirical legal studies, and LegalTech show great promise for legal history, their application remains limited. Nevertheless, the methodologies shaping DLH continue to expand the field, offering new avenues for legal historians to explore legal sources and their broader societal impact.

2.1. Digitization ('Digitalisaten') and Preservation

The processes of digitisation and preservation are inextricably linked, particularly in the context of historical documents. Digitisation entails the conversion of physical texts into digital formats, primarily to safeguard them from deterioration and ensure broader access. It ensures the preservation of the original content in a format that can be readily stored, searched, and shared. Nevertheless, while digitisation prioritises accessibility and longevity, it does not inherently facilitate more profound analytical engagement with the material.

Preserving historical records through digitisation protects from having to touch the materials.¹² It is frequently a precursor to more sophisticated processes, such as datafication, which permits detailed analysis. While some institutes (archives) opt to digitise their entire collection, this approach is becoming increasingly uncommon.¹³ This is because datafication has become a relatively straightforward process and, moreover, funding agencies are reluctant to provide financial support for digitisation projects. Nevertheless, in situations where there is an urgent need to preserve materials at risk of deterioration or other forms of damage, this can be a decision taken.

The digitisation of the Swedish governmental colonial archive from Saint Barthélemy is a noteworthy endeavour spearheaded by Frederik Thomasson. It was initiated to preserve and facilitate access to historical records that had remained largely inaccessible to researchers for decades. Saint Barthélemy, was under Swedish rule from 1785 to 1878, bequeathed a substantial archive to France upon its return to French sovereignty. The documents, comprising approximately 328 volumes and 300,000 pages, were transferred to the French National Archives in Aix-en-Provence. However, due to the advanced deterioration, the archive had been inaccessible for research purposes for a considerable time.

¹² One should be careful with statements about the 'ease' of handling digital materials, as the 'Digitalisaten' (digital copies) bring their challenges with changing programs (tools) and version management.

¹³ Institutes that merely digitize tend to opt for the option to use the IIF manifest, which enables other users to approach the files through the manifest and use the sources within, e.g., digital editions or combine files to recreate former collections that have ended up scattered throughout the world. Joseph Padfield et al., 'Practical Applications of IIF as a Building Block towards a Digital National Collection' (Zenodo, 22 July 2022), <https://doi.org/10.5281/zenodo.6884885>.

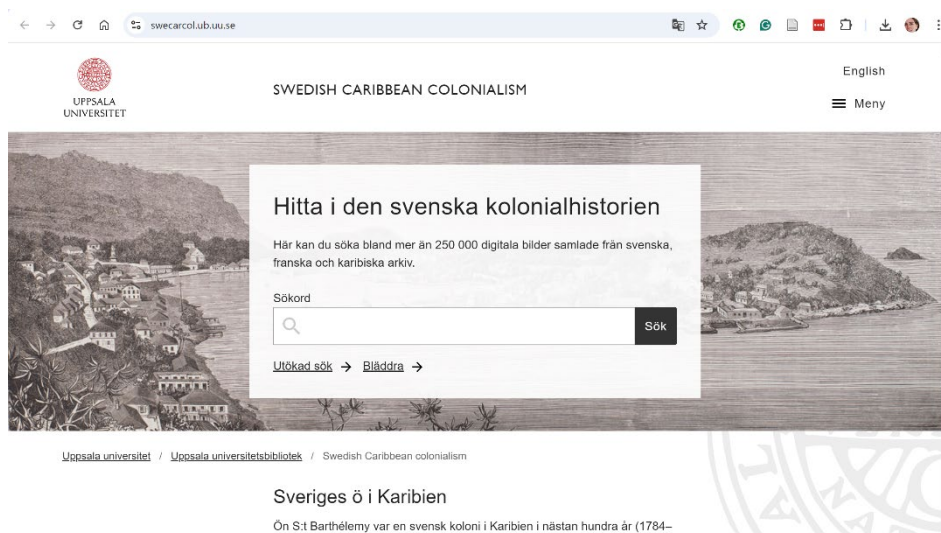


Figure 1
<https://swecarcol.uu.se/>

Recognising the importance of these documents, Thomasson initiated a project between 2013 and 2018 to digitise the entire collection, known as the *Fonds Suédois de Saint Barthélemy*. This initiative entailed the acquisition of financial resources and the administration of the digitisation process. The project not only restored access to these invaluable records but also enabled the initiation of new academic research into Sweden's colonial history. The result is a detailed, searchable archival database containing over 42,000 records, accessible via the Swedish Caribbean Colonialism project.¹⁴

The digitisation project is of great significance for understanding Sweden's colonial activities, particularly its involvement in the Caribbean, including the issues of slavery and trade. Thomasson's work has also resulted in numerous publications that examine various facets of Swedish colonialism, including censorship and print culture in Gustavia and Sweden's interactions with Haiti during the revolutionary period.¹⁵ As a model case study in applying digital humanities methodologies, the project has successfully demonstrated the potential of digital technologies to facilitate access to previously inaccessible archives, thereby illuminating previously neglected aspects of European colonial history.¹⁶

¹⁴ <https://swecarcol.uu.se/>

¹⁵ See for example: Frederik Thomasson, "The Caribbean Scorpion: The Saint Barthélemy Archive and Swedish Colonial Amnesia." *Small Axe*, no. 62 (July 2020): 53–66; Thomasson, "Révoltes et résistances dans la Caraïbe scandinave." *Revue du Philanthrope* 9 (2020): 125–143; Thomasson, "Moors in the Caribbean, Sámi in the Seraglio: Swedish Theatre and Slavery Around 1800." *Modern Languages Open* 2020, no. 1 (2020): 1–19. <https://doi.org/10.3828/mlo.v0i0.307>; Thomasson, "Entre rêves coloniaux et réalités politiques: La Guadeloupe suédoise (1813–1814) et ses conséquences." *Bulletin de la Société d'Histoire de la Guadeloupe*, special issue: Les occupations anglaises de la Guadeloupe, 2018: 105–122; Thomasson, "'Vous-même, ôtez votre chapeau!' Les Livres de couleur dans la colonie suédoise de Saint-Barthélemy (1785–1831)." *Les Cahiers des Anneaux de la Mémoire*, special issue: *Couleur et liberté dans l'espace colonial français (début XVIIIe-début XIXe siècle)* 17 (2017): 111–127.

¹⁶ While it may be challenging to gather resources for the digitisation and datafication process, a well-thought out plan could be helpful. See in this respect: C. Annemieke Romein, Süphan Kirmizialtin, Ronny Reshef, Christa Schneider, Giorgia Agostini, Ash Charlton, Melissa Terras, Joseph Nockels (et. al.) (2024), *Automatic Text Recognition of Historical Text Collections: From Research Proposal to Project Management. A Guide from the Transkribus Community on Planning and Executing Workflows for Researchers and GLAM-Professionals* (under review).

2.2. Metadata Overviews and Databases

It is of the utmost importance for scholars engaged in analysing intricate legal documents to integrate metadata-rich digital projects into their research methodologies. Such initiatives permit researchers to explore and analyse legal texts using structured, contextualised data, even when full-text access is unavailable. These projects represent a significant advancement beyond mere digitisation, enhancing the accessibility and utility of legal documents through the categorisation of information and the provision of sophisticated search functionalities. They draw on various legal sources, including historical court records and modern constitutions, and their methodologies are designed to apply across multiple academic disciplines. The following projects illustrate how digital methods have transformed the study of legal history.

One of the earliest examples of the application of digital methods to legal history can be found in the work of the *Centre d'Étude d'Histoire Juridique*, a joint project launched in 1953 by the University of Paris 2 Panthéon-Assas, the French National Centre for Scientific Research (CNRS), and the French National Archives. The centre concentrated on the legal rulings of the Parlement de Paris, the highest court in the Kingdom of France, whose records span over five centuries. Initially, the centre employed manual indexing; however, by the 1960s, it had transitioned to computerised processes during the 1970s.¹⁷

Several similar projects emerged around the same time, reflecting a growing interest in using computers in legal research. In 1969, Tom De Smidt at the University of Amsterdam established the *Werkgroep Grote Raad van Mechelen*, which was dedicated to studying the highest court of the Netherlands, the *Grand Conseil des Pays-Bas à Malines*.¹⁸ This court had operated from the 15th century until the French Revolution.

Meanwhile, at the Max Planck Institute for European Legal History (subsequently renamed the Max Planck Institute for Legal History and Legal Theory) in Frankfurt, a group led by Filippo Ranieri initiated a comparable project in the 1970s, concentrating on German Dissertations¹⁹ and another one on the *Reichskammergericht* (the Imperial Chamber Court), the supreme court of the Holy Roman Empire. This project resulted in one of post-war Germany's most extensive archival inventorying efforts. However, it should be acknowledged that the quantitative analysis of these legal materials made little progress beyond the foundational work conducted by Ranieri in the 1980s.

¹⁷ Wagner, 'Rechtsgeschichte'.

¹⁸ E. Ketelaar, 'In Memoriam: Tom de Smidt (1923-2013)', *Archievenblad* 117 (2013), <https://dare.uva.nl/search?identificer=0702ece1-452c-44c2-bf8d-0d7dc5b76b48>; DBNL, 'De grote raad van Mechelen, hoogste rechtcollege in de Nederlanden? H. de Schepper, Bijdragen en Mededelingen betreffende de Geschiedenis der Nederlanden. Deel 93', DBNL (DBNL), accessed 17 September 2024, https://www.dbnl.org/tekst/_bij005197801_01/_bij005197801_01_0016.php.

¹⁹ Filippo Ranieri, 'Eine Datenbank über juristische Dissertationen und Juristen im Alten Reich: ein Projektbericht', *Historical Social Research* 11, no. 1 (1986): 109–15, <https://doi.org/10.12759/hsr.11.1986.1.109-115>; Filippo Ranieri, *Juristische Dissertationen deutscher Universitäten 17.-18. Jahrhundert*, Ius commune (Frankfurt a.M.: Klostermann, 1986).

Another significant undertaking is the *Repertorium der Policeyordnungen*, initiated by Michael

The screenshot shows the website of the Max-Planck-Institut für Rechtsgeschichte und Rechtstheorie. The page title is 'Repertorium der Policeyordnungen'. The navigation bar includes 'INFORMATION', 'SUCHE', 'ERWEITERTE SUCHE', and 'LISTENANSICHT'. The search results are filtered by 'Territorium > Jahr > Policeymaterie'. The results list several ordinances from the city of Bern, including 'Schultheiß und Kleiner und Großer Rat der Stadt Bern (1528-1798) Verordnung' and 'Schultheiß und Kleiner und Großer Rat der Stadt Bern (1528-1798) Ordnung'. The search results are displayed in a table with columns for date, location, and title.

Facettierung/Selektion	Sortierung: Territorium > Jahr > Policeymaterie	Jahr > Territorium > Policeymaterie	Policeymaterie > Jahr > Form	4932 Gesetze, 7591 Materien. Anzeige der Gesetze 1 bis 10
Bern 00.00.1528 4.2 Forst StAB, A I 480, S. 231	Schultheiß und Kleiner und Großer Rat der Stadt Bern (1528-1798) Verordnung			
Bern 07.02.1528 1.1 Geistlichkeit, 1.1 Kirchengzucht, 1.1 Kirchenorganisation, 1.1 Klosterwesen/Orden, 1.1 Konfession, 1.1 Religiöse Unterweisung, 2.1 Zutrinken/Trunksucht, 2.3 Zensur StAB, A I 479, fol. 1v-10v	Schultheiß und Kleiner und Großer Rat der Stadt Bern (1528-1798) Ordnung			
Bern 27.04.1528 1.1 Kirchenorganisation StAB, A I 479, fol. 11r-13v	Schultheiß und Kleiner und Großer Rat der Stadt Bern (1528-1798) Ordnung		BER.01.001.0002	
Bern 20.05.1528 1.1 Kirchenorganisation StAB, A I 479, fol. 13v-14r	Schultheiß und Kleiner und Großer Rat der Stadt Bern (1528-1798) Ordnung			
Bern 21.06.1528	Schultheiß und Kleiner und Großer Rat der Stadt Bern (1528-1798) Ordnung			

Figure 2 *Policeyordnungen der Frühen Neuzeit, Karl Härter and Michael Stolleis (eds.), Online-Publication: <https://policey.lht.mpg.de> [19-09-2024] Search for the Territory of Bern.*

Stolleis and Karl Härter at the Max Planck Institute for European Legal History. The project, funded by the German Research Foundation (DFG), aimed to compile a comprehensive database of police ordinances (Policeygesetze) from the early modern Holy Roman Empire, Denmark, Sweden, and Switzerland. This resource is notable for its comprehensive metadata, categorising legal documents by date, type, and jurisdiction. Initially published in printed volumes from 1996 to 2017, the database became publicly accessible online in 2021, with regular updates extending its scope.²⁰ The *Repertorium* has proven invaluable for studying regulatory developments over time, allowing scholars to explore the intersection of law, governance, and society in early modern Europe.²¹

²⁰ Max Planck Repertorium der Policeyordnungen: <https://www.rg.mpg.de/repertorium-policeyordnungen>

²¹ Saskia Limbach and Vittorio Klostermann, *Government Use of Print in the Holy Roman Empire in the Sixteenth Century*, 2020; Christel Annemieke Romein, 'The Formative Role of Early Modern Books of Ordinances. The Low Countries and Their Overseas Lands', in *The Early Modern State: Drivers, Beneficiaries and Discontents*, ed. P. Brandon, L Heerma van Voss, and C. A. Romein (London, 2022), chapter 4; Christel Annemieke Romein, Sara Veldhoen, and Michel de Gruijter, 'The Datafication of Early Modern Ordinances', *DH Benelux Journal*, 2 (2020), <http://journal.dhbenelux.org/journal/issues/002/article-23-romein/article-23-romein.pdf>; C. A. Romein, 'Early Modern State Formation or Gute Policey? The Good Order of the Community', *The Seventeenth Century* 37, no. 6 (2 November 2022): 1031–56, <https://doi.org/10.1080/0268117X.2022.2111339>; Jørgen Mührmann-Lund, *Borgerligt regimente: politiforvaltningen i købstæderne og på landet under den danske enevælde*, Tidlig moderne, bind 13 (København: Museum TusulanumForlag, 2019); Jørgen Mührmann-Lund, 'Policing the Guilds', in *Bringing the People Back In. State Building from Below in the Nordic Countries ca. 1500-1800*, ed. Knut Dørup, Mats Hallenberg, and Kimmo Katajala (London, New York: Routledge, 2021), 146–61.

2.3. Datafication and Source-Editions

Datafication transforms textual sources into structured, machine-readable data, thereby enabling large-scale digital analysis.²² This process encompasses stages such as text recognition, segmentation, and classification, allowing scholars to examine extensive corpora systematically.²³ By converting these documents into data, researchers can identify patterns, trace developments, and perform comparative analyses more efficiently than would be possible with traditional methods.²⁴ This approach facilitates the study of legal history and opens new avenues for interdisciplinary research within the digital humanities.

The distinction between digitization and datafication is in their respective scope and objectives. The term 'digitization' describes converting physical documents into digital formats. This can be exemplified by scanning historical texts, which become accessible online. However, the process of 'datafication' extends beyond this, transforming those digitized texts into structured, machine-readable data. This involves the extraction and organization of the content, enabling systematic analysis through the use of computational tools.²⁵ While digitization primarily concerns preservation and accessibility, datafication facilitates more in-depth analysis, allowing researchers to uncover patterns and trends across large datasets.

In recent decades, several notable initiatives have been launched to provide digital access to historical legal data, thereby transforming the landscape of legal research. The Old Bailey Online project, initiated in 1999 by Tim Hitchcock, Robert Shoemaker, and colleagues, provides comprehensive text access to court proceedings from London's *Old Bailey* between 1674 and 1913.²⁶ Similarly, the *Caselaw Access Project* (CAP), developed by the Library Innovation Lab at Harvard Law School, offers a comprehensive database of US state and federal court opinions from 1658 to 2020.²⁷ In contrast to earlier digital projects, which concentrated on indexing or summarising records, *Old Bailey Online* and *CAP* provide digitised full-text access, thereby enhancing search capabilities with metadata on dates, parties, and judges and allowing for comprehensive analysis of legal disputes. Other pioneering efforts in DLH include the *Regesta Imperii* project, which began digitizing extensive collections of medieval imperial charters in the 1990s.²⁸ The digitization of these documents was concluded with the establishment of a web presence in 2001, facilitating open access to legal texts and their historical contexts. In the 1980s, Marianne Meinhart and Josef Menner initiated a project at the University of Linz to digitize the Digest of Justinian. This initiative enabled scholars to request search results via the early computer networks of the time.²⁹

²² Romein, Veldhoen, and de Gruijter, 'The Datafication of Early Modern Ordinances'.

²³ de Jong, 'Over Het Nut van Digitale Archiefontsluiting. De Civiele Procesdossiers van Het Hof van Holland En Het Hof van Friesland'.

²⁴ H de Jong, "'Copy-Paste" in Processen Aan Het Hof van Friesland. Over de Noodzaak van Archiefontsluiting', *De Vrije Fries* 102 (2022): 124–37.

²⁵ Romein, Veldhoen, and de Gruijter, 'The Datafication of Early Modern Ordinances'.

²⁶ Tim Hitchcock and William J. Turkel, 'The Old Bailey Proceedings, 1674–1913: Text Mining for Evidence of Court Behavior', *Law and History Review* 34, no. 4 (November 2016): 929–55, <https://doi.org/10.1017/S0738248016000304>.

²⁷ 'Caselaw Access Project', accessed 18 September 2024, <https://case.law/>.

²⁸ 'Startseite : Regesta Imperii', accessed 18 September 2024, <http://www.regesta-imperii.de/startseite.html>.

²⁹ Wagner, 'Rechtsgeschichte'.

It is also worth noting the significant contribution made by projects such as the *Sammlung Schweizerischer Rechtsquellen* (SSRQ) in facilitating access to historical legal sources.³⁰ The SSRQ was established in 1898 to systematically collect Swiss legal documents from the Middle Ages to the early modern period. Recently, the project has embraced digitisation, offering TEI-XML versions of important texts that are searchable and accessible online. Similarly, the *Deutsche Rechtswörterbuch* project, initiated in 1897, commenced the digitization of legal terminology in the 1980s, becoming a significant resource for the historical German legal language.³¹ These projects demonstrate the increasing significance of digital methodologies in facilitating greater accessibility to legal history for scholars worldwide.

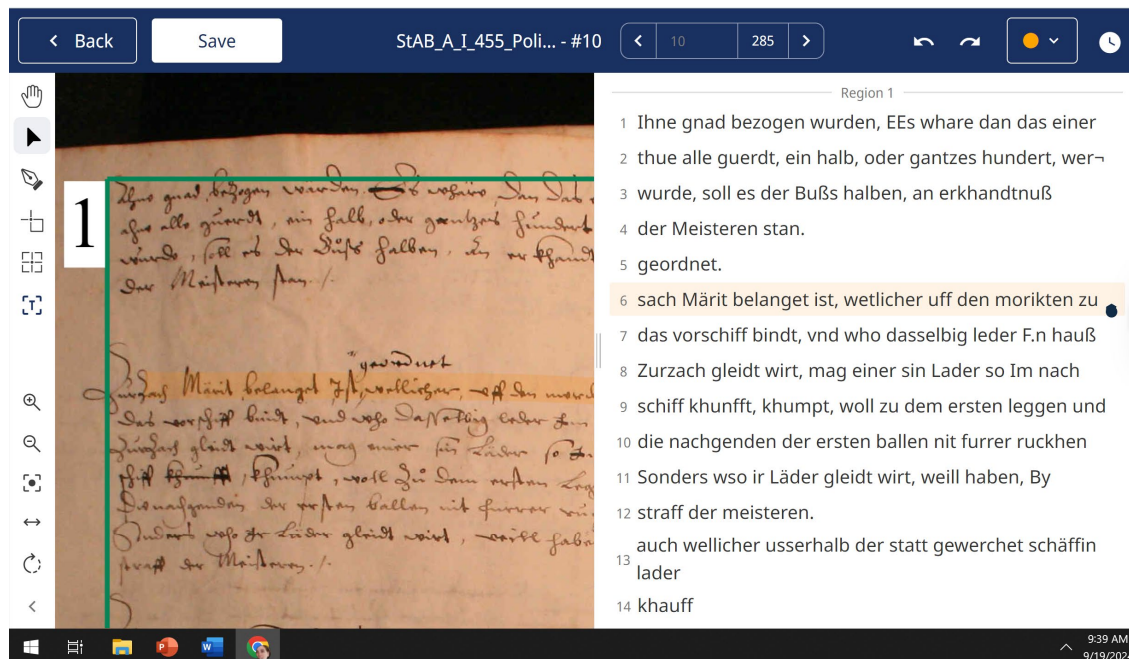


Figure 3 Staatsarchiv Bern A I 455 (Mandatenbuch) within the Transkribus App (www.app.transkribus.org) to enable datafication.

Digital editions can, therefore, be created using datafication tools (ATR recognition). These methods differ from the philological approach, which aims to create a textual approach to facilitate the reader's experience of the sources. While some - not all! - digital editions, the philosophy is that accessibility is more important than perfection.

³⁰ 'SSRQ Online', accessed 17 September 2024, <https://www.ssrq-sds-fds.ch/projekte/ssrq-online/>.

³¹ Andreas Deutsch, 'Straf- und ordnungsrechtliche Regelungen in ländlichen Rechtsquellen und ihr spezifischer Wortschatz', *Journal for Digital Legal History* 2, no. 1 (2 October 2023), <https://doi.org/10.21825/dlh.87176>.

Regierungsratsprotokolle des Kantons St.Gallen 1803-1831

Home Explore Suche Über EN

Regierungsratsprotokolle des Kantons St.Gallen 1803-1831

Suchen... Suche

Die an den wöchentlichen Sitzungen der Regierung gefällten Beschlüsse sind samt den zugehörigen Erwägungen in den sogenannten Verhandlungsprotokollen festgehalten. Diese sind, geordnet nach Jahr und Geschäftsnummer, in Buchform eingebunden. Je nach Entstehungszeitpunkt des Protokolls sind zusätzlich einzelne Schlüsseldokumente, die mit der Beschlussfassung in einem Bezug stehen, den

Figure 4 Digital Edition of *Regierungsratsprotokolle des Kantons St. Gallen* on Transkribus Sites (<https://app.transkribus.org/sites/St.Gallen/>)

The initiation of commercial legal databases, exemplified by *LexisNexis* (subsequently acquired by Elsevier) and *WestLaw* (now owned by Thomson Reuters), marked a pivotal moment in the evolution of legal research during the 1970s.³² These databases ushered in a new era of extensive access to case law, particularly in the American and English legal systems.³³ These platforms contain vast collections of historical legal materials, yet they are primarily designed for contemporary legal practice, offering critical resources for lawyers and judges. However, their high costs and restricted download and export functions frequently limit their use in historical legal research. In continental Europe, comparable commercial databases, such as *Juris* and *Beck Online* in Germany and *Lefebvre Dalloz* in France, offer similarly comprehensive collections of laws, regulations, and case law.³⁴ These platforms are utilized by contemporary legal professionals and those studying historical legal texts. In addition, companies such as *Gale* and *HeinOnline* provide access to doctrinal texts, historical treaties, legal literature, and bibliographies. *Gale*, for instance, provides access to its *Digital Scholar Lab*, which enables users to undertake digital analysis of its collections.³⁵ Despite the dominance of commercial platforms, the availability of legal sources at no cost has gained ground since the 1990s. The *Free Access to Law Movement*, formalised in 2002 through the *Declaration on Free Access to Law*, advocates open access to legal materials.³⁶ Initiatives such as the Caselaw Access Project and Germany's

³² 'LexisNexis | Professionelle Online-Recherche-Tools', accessed 18 September 2024, https://www.lexisnexis.com/de-de?utm_source=google&utm_medium=cpc&utm_campaign=brandde&gad_source=1&gclid=Cj0KCQjw9Km3BhDjARIsAGUb4nwzEbq4Z2vWVMaiUWYuD00JWNNeaSTik7O0dqXYkrJYyWFpHCQag5t8aArrbEALw_wcB; 'Westlaw – Legal Research Platforms', accessed 18 September 2024, <https://legal.thomsonreuters.com/en/westlaw>.

³³ See for more on this: Wagner, 'Rechtsgeschichte'.

³⁴ 'juris. Wissen, das für Sie arbeitet. | juris', accessed 18 September 2024, <https://www.juris.de/jportal/nav/index.jsp>; 'Homepage - Beck-Online', accessed 18 September 2024, <https://beck-online.beck.de/Home>; 'Gagnez en savoirs, compétences et efficience !', Lefebvre Dalloz, accessed 18 September 2024, <https://www.lefebvre-dalloz.fr/>.

³⁵ 'Scholarly Resources for Learning and Research | Gale', accessed 18 September 2024, <https://www.gale.com>; '- HeinOnline.Org', accessed 18 September 2024, <https://heinonline.org/HOL/Welcome>.

³⁶ 'Free Access to Law Movement | Institute of Advanced Legal Studies', accessed 18 September 2024, <https://ials.sas.ac.uk/leadership-and-collaboration/free-access-law-movement>.

government-operated *Gesetze im Internet* offer alternatives to commercial databases.³⁷ Initially focused on current law, these free resources often contain historically relevant materials, blurring the line between past and present legal research.



Figure 5 Digital Edition at Iurisprudencia created with Transkribus Read&Search
(<https://rwi.app/iurisprudencia/de/locre/documents/static/832753/pages/7?zoom=0.52>)

2.4. Information Extraction, Text Mining and Natural Language Processing (NLP)

Let us first examine the techniques referenced in the title of this subsection. In the context of Digital Humanities, information extraction is the automated process of identifying and retrieving specific data from extensive collections of unstructured or semi-structured texts, such as historical documents or legal records. Information extraction employs natural language processing and machine learning techniques to identify important elements, such as names, dates, places, and relationships, transforming textual data into a structured, searchable format. This process allows researchers to analyse and interpret large datasets efficiently, thereby revealing patterns and insights that would be difficult to uncover manually. Information extraction is, therefore, of great importance in enabling complex, interdisciplinary research across fields such as history, literature, and law. The process of text mining involves the extraction of pertinent patterns, trends, and information from voluminous legal corpora. To illustrate, examining a substantial corpus of judicial decisions allows for identifying shifts in the judiciary's rationale or the recurrent utilisation of particular legal terminology. Natural Language Processing (NLP) enables the processing of complex legal language.³⁸ Entity recognition allows for the automated annotation of legal documents with so-called entities, including individuals,

³⁷ 'Gesetze Im Internet', accessed 18 September 2024, <https://www.gesetze-im-internet.de/>.

³⁸ There is even a subdomain, called NLLP (Natural Legal Language Processing). See: 'NLLP Workshop 2024', accessed 19 November 2024, <https://nllpw.org/>.

legislation, and institutions. Furthermore, NLP tools facilitate the identification of linguistic shifts, such as the evolution of the meaning of legal concepts (e.g. sovereignty, justice) over time.

These tools facilitate a more profound examination of legal documents that would otherwise be impossible through manual examination. To illustrate, a researcher examining 19th-century legal rulings could use these techniques to ascertain the frequency with which judges referenced specific precedents or which legal arguments became dominant during one particular historical period. Text mining and NLP collectively permit historians to investigate broad legal trends, thereby transforming the study of legal history.

The CiSaMe (Circulation des Savoirs Médiévaux au XIIe Siècle - Nouveaux Outils, Nouvelles Perspectives, Nouvelles Hypotheses) project (Strasbourg) provides an illustrative example of text mining in action.³⁹ The project employs a range of techniques to identify and trace citations to other documents. In this case, after employing the Kraken/eScriptorium tool, regular expression patterns are used to identify repetitive patterns, and large language models are used to match the texts. Another approach is the search for formulaic language, as is being done within the Huygens' Republic Project. The standard formulations used consider spelling variations through some degree of 'fuzziness'. Consequently, these are used to find the day-date, start and end of resolutions and attendance lists.⁴⁰ The Slave Societies Digital Archive (SSDA) is a digital repository that meticulously extracts data from historical records related to enslaved people, particularly emphasising materials from Latin America and the Caribbean.⁴¹ The data extraction process entails digitising records such as baptism, marriage, and burial certificates and organizing information, including names, dates, and locations, into structured formats. These extracted details are then made searchable and accessible, enabling researchers to trace genealogies and examine the lives of enslaved individuals and their communities over time. This approach ensures the preservation and detailed study of these crucial historical documents.

³⁹ 'Contrats de Recherche - DRES - Droit, Religion, Entreprise et Société - UMR 7354 - MISHA - Maison Interuniversitaire Des Sciences de l'Homme - Alsace - Université de Strasbourg', accessed 18 September 2024, <https://dres.unistra.fr/droits-et-religions/contrats-de-recherche>.

⁴⁰ 'Republic', accessed 20 September 2024, <https://republic.huygens.knaw.nl/>.

⁴¹ 'Slave Societies Digital Archive', accessed 17 September 2024, <https://slavesocieties.org/>.

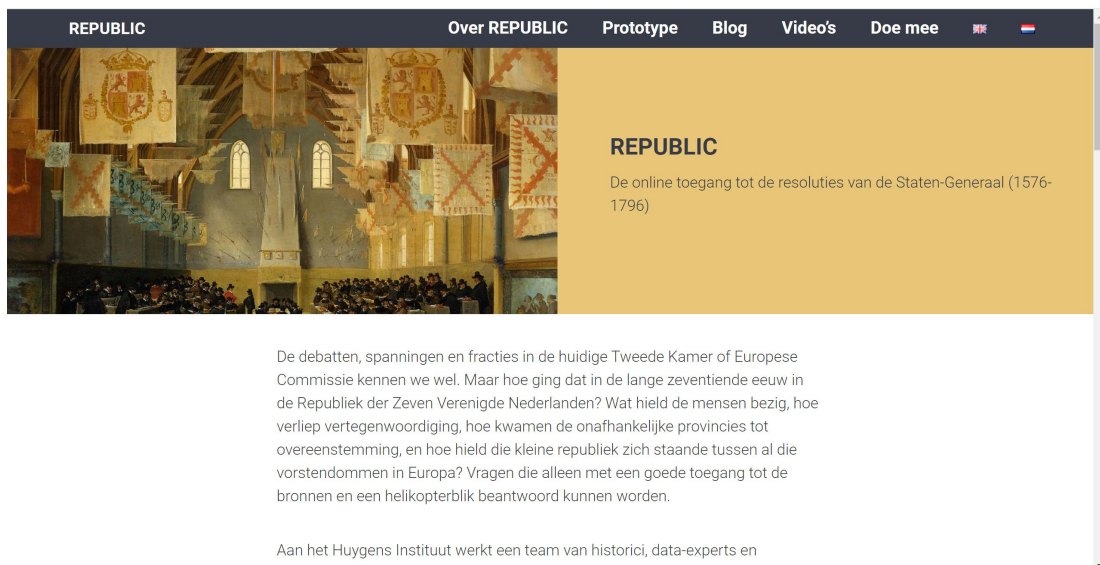


Figure 6 Republic Project website - <https://republic.huygens.knaw.nl/>

2.5. Spatial Reconstructions and Network/Data Visualisations

Using digital tools to represent geographical locations and delineate relationships between entities is fundamental to spatial reconstructions and network/data visualizations. Spatial reconstructions, which may be interactive maps, visualize historical events or movements across physical spaces. This reveals how locations and spaces were used or interacted with. Network visualizations represent connections between individuals, occurrences, or entities, facilitating comprehension of the structure and relationships within datasets.⁴²

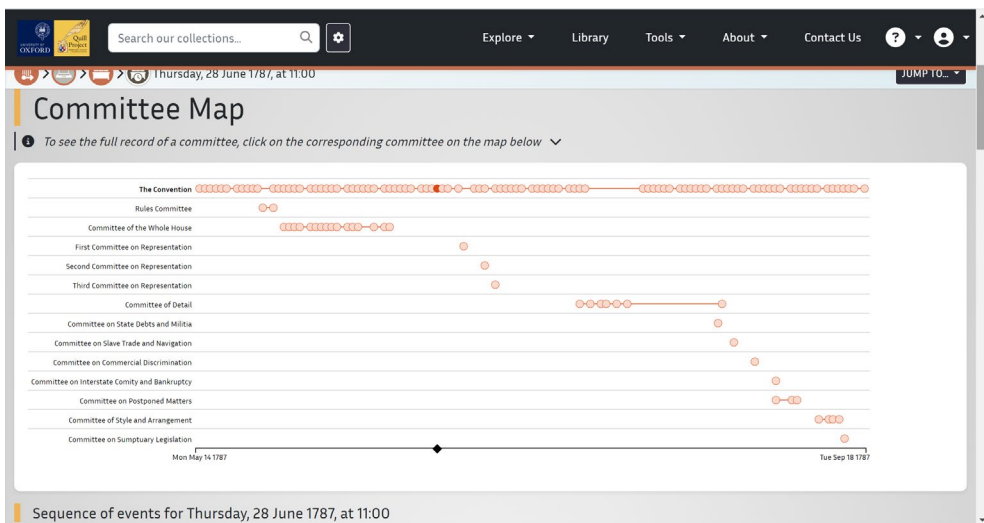


Figure 7 Quill-Project: U.S. Constitutional Convention 28 June 1787 (2021 Edition) Committee Session 6226.⁴³

⁴² See for example: Hylkje De Jong and Gijs Van Dijck, 'Network Analysis in Legal History: An Example from the Court of Friesland', *Tijdschrift Voor Rechtsgeschiedenis = Revue d'histoire Du Droit = The Legal History Review* 90, no. 1–2 (June 2022): 250–62, <https://doi.org/10.1163/15718190-20220004>.

⁴³ <https://www.quillproject.net/m2/session/6226>

The Quill Project offers detailed reconstructions of the processes through which legal and constitutional texts are developed.⁴⁴ These reconstructions are informed by analyzing the formal debates and amendments that have shaped the texts, including constitutions and treaties. The platform tracks changes, actor involvement, and procedural steps, offering insights into the legislative process. While this involves textual analysis, the project is more about modelling and visualizing negotiations.

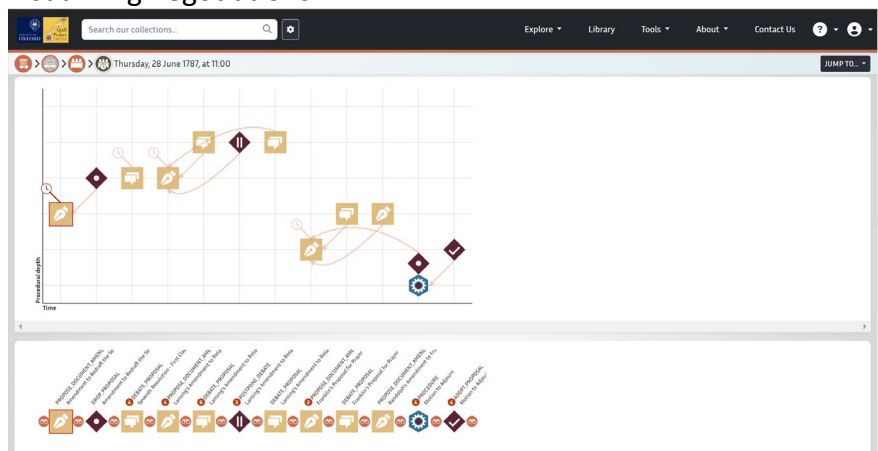


Figure 8 Quill-Project: U.S. Constitutional Convention 28 July 1787 (2021 Edition) Visualisation of Session 6226.⁴⁵

Another illustrative example of a project that places significant emphasis on the utilisation of visualisations is the "Lynching in Texas" project.⁴⁶ This project has been developed to compile and visualise historical data related to lynchings in Texas from 1882 to 1945. The project utilises a comprehensive database that documents a multitude of details about each incident, including the names of the victims, their locations, the dates on which the lynchings occurred, and the circumstances surrounding each event. To facilitate exploration of the patterns and contexts of lynchings in Texas, the project offers a range of interactive tools, such as maps and timelines. By providing access to this data, the project aims to advance research, enhance public awareness, and promote education on the history of racial violence in Texas.

A further US-based project is the Digital Harlem Project, which provides an interactive platform for exploring Harlem's everyday life during the 1920s, with a particular focus on the experiences of the African American population.⁴⁷ The project employs much historical data, including addresses, events, and legal records, to generate intricate visualisations such as interactive maps. These visualisations permit users to discern patterns and movements within Harlem, such as the locations of businesses, churches, or even instances of arrests. Through these tools, the project offers a dynamic methodology to examine Harlem's dynamic history, providing depth to research on race, space, and community life.

⁴⁴ 'Quill Project', accessed 17 September 2024, <https://www.quillproject.net/>.

⁴⁵ <https://www.quillproject.net/m2/session/6226>

⁴⁶ 'Lynching In Texas', accessed 17 September 2024, <https://lynchingintexas.org/>.

⁴⁷ 'Digital Harlem', accessed 17 September 2024, <http://digitalharlem.org/>.

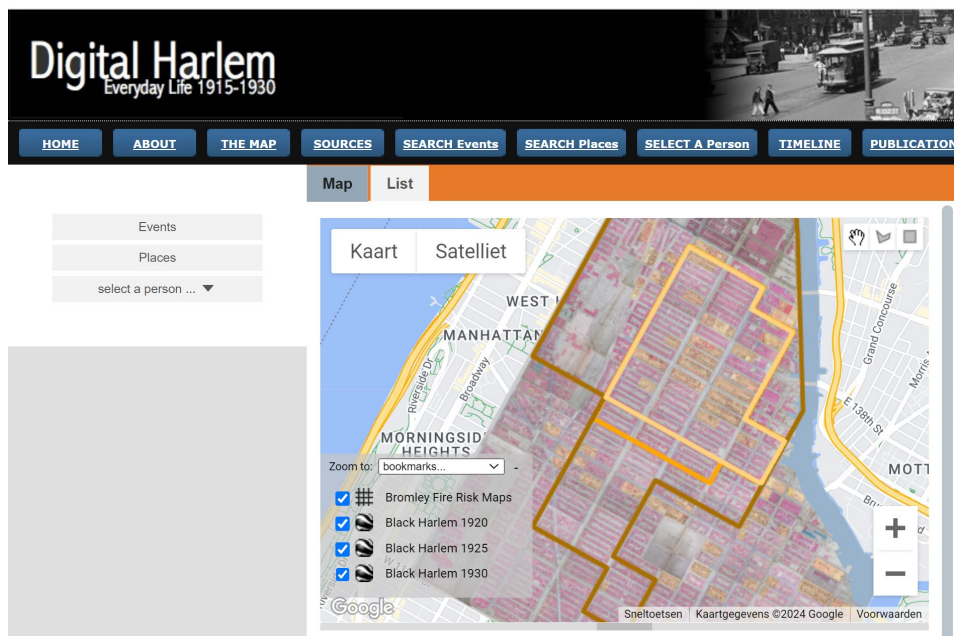


Figure 9 Digital Harlem Project (<http://digitalharlem.org/>)

In short, the Quill Project employs techniques to model how legal texts, such as constitutions, were negotiated through formal debates. In contrast, the Lynching in Texas project utilises maps and timelines to visualise the geography and patterns of lynchings. Similarly, the Digital Harlem Project employs interactive maps to represent the daily life and movements of Harlem's African American population during the 1920s. These visual tools facilitate a more profound comprehension of historical events by elucidating these examples' underlying spatial and relational patterns.

2.6. Interoperability and Standards

Linked data represents a methodology for structuring and interconnecting information in a manner that facilitates its straightforward interpretation and processing by both human and machine entities.⁴⁸ This is accomplished by establishing connections between different datasets through shared identifiers, thereby facilitating the integration and querying of data from diverse sources as a unified entity. The function of ontologies in the context of linked data is paramount, as they provide the formal structure for defining concepts and relationships across datasets. Ontologies guarantee semantic clarity by providing a common vocabulary and a uniform framework, thus ensuring that the data is interconnected and meaningful.⁴⁹ For researchers, the principal advantage of utilising linked data is its capacity to facilitate interoperability and reuse data across disparate fields. It permits the mapping and analysis of intricate relationships that

⁴⁸ Eero Hyvönen, *Publishing and Using Cultural Heritage Linked Data on the Semantic Web*, Synthesis Lectures on Data, Semantics, and Knowledge (Cham: Springer International Publishing, 2012), <https://doi.org/10.1007/978-3-031-79438-4>.

⁴⁹ Oksanen Arttu et al., 'Semantic Finlex: Transforming, Publishing, and Using Finnish Legislation and Case Law As Linked Open Data on the Web', *Frontiers in Artificial Intelligence and Applications*, 2019, 212–28, <https://doi.org/10.3233/FAIA190023>.

would be unfeasible with isolated datasets. However, the complexity of designing and maintaining ontologies and ensuring consistent data quality across linked datasets can be seen as a significant disadvantage.

Applying systems such as CIDOC-CRM (Conceptual Reference Model), initially designed for the museum domain, to the humanities and the field of law presents researchers with particular challenges.⁵⁰ Francesco Beretta notes that while CIDOC-CRM demonstrates considerable proficiency in integrating museum data, it frequently encounters challenges when confronted with the more expansive and intricate demands of humanities and social science (HSS) discourse.⁵¹ The fundamental challenge is that CIDOC-CRM is founded upon an ontological abstraction, which may prove inadequate in capturing the complex and interdisciplinary data requirements of the humanities and social sciences (HSS). Similarly, other systems, such as FRBRoo or CRMgeo, are constrained by their narrow focus on specific physical objects and domains, such as museums or archaeology.⁵² This limits their adaptability to broader or more interdisciplinary contexts. Beretta suggests that creating high-level extensions, such as the Semantic Data for Humanities and Social Sciences (SDHSS), could prove a practical solution, addressing the gaps above and enhancing the applicability of CIDOC-CRM to a broader range of research contexts.

The Data for History consortium aims to establish a collaborative framework for managing and disseminating historical data within a unified structure, including legal history and datasets.⁵³ By fostering interoperability between datasets within and across different domains, the consortium assists e.g. legal historians in standardising and linking diverse legal sources, facilitating comparative analysis and interdisciplinary research. Using ontologies and linked data will facilitate enhanced accessibility and analytical depth of legal datasets, thereby supporting a more integrated and comprehensive study of legal history.⁵⁴

2.7. AI and LLMs

The recent advancements in artificial intelligence (AI), mainly through the development of generative models such as OpenAI's ChatGPT⁵⁵, Anthropic's Claude⁵⁶, and Meta's LLaMA⁵⁷, or

⁵⁰ Carol Usher, 'How to Model Roles in the CIDOC-CRM RDF Encoding', accessed 2 April 2019, <http://www.cidoc-crm.org/sites/default/files/Roles.pdf>; 'Semantic Data for Humanities and Social Sciences | CIDOC CRM', accessed 17 September 2024, <https://cidoc-crm.org/Resources/semantic-data-for-humanities-and-social-sciences>.

⁵¹ Francesco Beretta, 'A Challenge for Historical Research: Making Data FAIR Using a Collaborative Ontology Management Environment (OntoME)', *Semantic Web* 12, no. 2 (1 January 2021): 279–94, <https://doi.org/10.3233/SW-200416>; Francesco Beretta, 'Semantic Data for Humanities and Social Sciences (SDHSS): An Ecosystem of CIDOC CRM Extensions for Research Data Production and Reuse', 2024, <https://doi.org/10.33968/9783966270502-05>.

⁵² Beretta, 'Semantic Data for Humanities and Social Sciences (SDHSS)'.

⁵³ 'Data for History Consortium | Data for History', accessed 22 June 2022, <http://dataforhistory.org/>.

⁵⁴ In line with the public lecture of Francesco Beretta several ontologies were discussed during the preceding internal workshops, 'Modelling Social and Legal Facts in the Context of the Semantic Data for Humanities and Social Sciences (SDHSS) Ontology Ecosystem', accessed 17 September 2024, <https://www.lhlt.mpg.de/events/37642/2077780>.

⁵⁵ 'OpenAI', accessed 17 September 2024, <https://openai.com/>.

⁵⁶ 'Claude', accessed 17 September 2024, <https://claude.ai/>.

⁵⁷ 'Llama 3.1', Meta Llama, accessed 17 September 2024, <https://www.llama.com/>.

GitHub Copilot⁵⁸ have had a significant and far-reaching impact across many domains, including the digital humanities and legal history. These large language models (LLMs)⁵⁹ employ sophisticated machine learning architectures, particularly transformers⁶⁰, which enables them to process and generate human-like text based on vast training data. It is important to realise that through their analysis of language, these LLMs produce a statistically logical text, meaning that the words that follow each other are statistically more often found in combination with each other. The training data for these models encompass various sources, including literary works and legal documents, endowing them with considerable versatility for research, creative industries, and academic inquiry.

Within digital humanities, LLMs are transforming how scholars conduct text-based research. The capacity of these models to rapidly analyze and generate text enables researchers to interrogate extensive archives, identify thematic patterns and conduct comparative analyses on an unprecedented scale. For example, LLMs can facilitate the textual analysis of historical documents or literary corpora, thereby enabling the discovery of insights into language change, sentiment, and thematic shifts over time. In DLH, LLMs offer invaluable assistance in analyzing legal texts, facilitating the tracking of legal precedents, summarising case law, and even predicting judicial trends. Automating these processes enables legal historians to engage with primary sources in novel and expansive ways, facilitating more efficient research workflows.

Nevertheless, the incorporation of LLMs into these domains presents several challenges and, therefore, Tobias Hodel pleaded – at the 44. *Rechtshistorikertag* (Frankfurt a. Main) – for academics/ legal historians to develop at least some ‘algorithmic literacy’.⁶¹ Significant concerns include the potential for bias, inaccuracy (e.g., hallucinations), and ethical issues associated with using AI in academic research.⁶² As LLMs are trained on large datasets, they risk inheriting the biases and limitations present in their data sources. This raises concerns about the reliability of their outputs, particularly in fields where precision is critical. Furthermore, the uncritical adoption

⁵⁸ ‘GitHub Copilot · Your AI Pair Programmer’, GitHub, 2024, <https://github.com/features/copilot>.

⁵⁹ Aniket Deroy, ‘How Ready Are Pre-Trained Abstractive Models and LLMs for Legal Case Judgement Summarization?’, n.d.

⁶⁰ Lazar Peric et al., ‘Legal Language Modeling with Transformers’, in *Proceedings of the Fourth Workshop on Automated Semantic Analysis of Information in Legal Text Held Online in Conjunction with the 33rd International Conference on Legal Knowledge and Information Systems, ASAIL@JURIX 2020, December 9, 2020*, ed. Kevin D. Ashley et al., vol. 2764, CEUR Workshop Proceedings (CEUR-WS.org, 2020), <https://ceur-ws.org/Vol-2764/paper2.pdf>; Jaromír Savelka, Hannes Westermann, and Karim Benyekhlef, ‘Cross-Domain Generalization and Knowledge Transfer in Transformers Trained on Legal Data’, in *Proceedings of the Fourth Workshop on Automated Semantic Analysis of Information in Legal Text Held Online in Conjunction with the 33rd International Conference on Legal Knowledge and Information Systems, ASAIL@JURIX 2020, December 9, 2020*, ed. Kevin D. Ashley et al., vol. 2764, CEUR Workshop Proceedings (CEUR-WS.org, 2020), <https://ceur-ws.org/Vol-2764/paper5.pdf>; Gabriele Marino et al., ‘Automatic Rhetorical Roles Classification for Legal Documents Using LEGAL-TransformerOverBERT’, in *Proceedings of the 6th Workshop on Automated Semantic Analysis of Information in Legal Text Co-Located with the 19th International Conference on Artificial Intelligence and Law (ICAIL 2023), Braga, Portugal, 23rd September, 2023*, ed. Francesca Lagioia et al., vol. 3441, CEUR Workshop Proceedings (CEUR-WS.org, 2023), 28–36, <https://ceur-ws.org/Vol-3441/paper4.pdf>.

⁶¹ Tobias Hodel, ‘Der machine learning Turn und rechtshistorische Dokumente’ (44. Rechtshistorikertag, Frankfurt a. Main, 19 September 2024), <https://rechtshistorikertag2024.de/tagungsprogramm>.

⁶² *Why OpenAI’s ChatGPT Is Such A Big Deal*, 2023, <https://www.youtube.com/watch?v=pOmpqdlVCoo>; ‘Wat is er open aan slimme chatbot-maker OpenAI?’, NRC, 1 January 2023, <https://www.nrc.nl/nieuws/2023/01/01/wat-is-er-open-aan-openai-a4153050>.

of AI-generated insights could result in an overreliance on technology, which might ultimately compromise the rigour of scholarly analysis. While LLMs offer novel opportunities for innovation in digital humanities and legal research, they must be employed as supplementary tools in conjunction with careful human interpretation to ensure the integrity and reliability of academic scholarship. In line with this integrity, it should be mentioned that the use of LLMs through APIs could/ should be considered, as this approach seems not to further train the language models.

3. Final remarks

The field of DLH offers a promising avenue for the study of legal history, providing novel avenues for engagement with historical legal texts, cases, and statutes. Using computational tools within DLH allows researchers to analyse extensive datasets, identify patterns and explore legal sources with a level of detail that would otherwise be unfeasible through traditional methods. Furthermore, it permits the visualisation of intricate legal developments and relationships, thereby facilitating comprehension of the evolution of law across time and regions. The digitisation and online accessibility of legal archives facilitate broader access to these sources while creating interdisciplinary collaboration opportunities. This brings together historians, legal scholars and data scientists in novel and meaningful ways.

The significance of DLH extends beyond its methodological contributions. Incorporating digital tools into legal history enables scholars to bridge the gaps between legal studies and other humanities disciplines, thereby facilitating a more comprehensive understanding of the role of law in society. Furthermore, DLH facilitates broader access to legal sources, enabling researchers from diverse backgrounds and institutions to engage with materials previously accessible only in physical archives.

This Digital Legal History journal, hosted by Ghent University, was founded to advance knowledge about the methodology. As one of the first journals dedicated to DLH, it provides a forum for scholars to disseminate their work, exhibit digital projects, and debate the evolving nexus between law and digital humanities.⁶³ The journal's provision of Diamond Open Access to articles and research contributes to the growth of a global scholarly community, fostering collaboration and innovation in legal history through digital methods.

In conclusion, DLH represents a transformative approach to legal historical research. Its emphasis on digital tools and accessibility enriches our understanding of the past and opens new avenues for interdisciplinary exploration and collaboration. The continued development of DLH, supported by platforms like the DLH-journal, promises to expand legal history's horizons and bring fresh insights into the interplay between law, society, and technology. Moreover, it invites legal

⁶³ The DLH-Journal was founded as an initiative when the 'Digital Methods and Resources in Legal History' conference (2020) was postponed (to 2021) due to the Covid-19 pandemic, see: <https://www.jhlt.mpg.de/dlh2021/en>. Resulting from a discussion, Dirk Heirbaut (UGent), Florenz Volkaert (UGent) and Annemieke Romein (Huygens Institute) took the initiative to found the journal.

historians to make their methodology explicit, whether they use graphs, maps, or ATR, to foster discussions about best practices and approaches.

Acknowledgements

Annemieke Romein's research was funded by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (grant VI.Veni.191H.035).

Bibliography

- '- HeinOnline.Org'. Accessed 18 September 2024. <https://heinonline.org/HOL/Welcome>.
- Arttu, Oksanen, Tamper Minna, Tuominen Jouni, Mäkelä Eetu, Hietanen Aki, and Hyvönen Eero. 'Semantic Finlex: Transforming, Publishing, and Using Finnish Legislation and Case Law As Linked Open Data on the Web'. *Frontiers in Artificial Intelligence and Applications*, 2019, 212–28. <https://doi.org/10.3233/FAIA190023>.
- Beretta, Francesco. 'A Challenge for Historical Research: Making Data FAIR Using a Collaborative Ontology Management Environment (OntoME)'. *Semantic Web* 12, no. 2 (1 January 2021): 279–94. <https://doi.org/10.3233/SW-200416>.
- . 'Semantic Data for Humanities and Social Sciences (SDHSS): An Ecosystem of CIDOC CRM Extensions for Research Data Production and Reuse', 2024. <https://doi.org/10.33968/9783966270502-05>.
- 'Caselaw Access Project'. Accessed 18 September 2024. <https://case.law/>.
- 'Claude'. Accessed 17 September 2024. <https://claud.ai/>.
- 'Contrats de Recherche - DRES - Droit, Religion, Entreprise et Société - UMR 7354 - MISHA - Maison Interuniversitaire Des Sciences de l'Homme - Alsace - Université de Strasbourg'. Accessed 18 September 2024. <https://dres.unistra.fr/droits-et-religions/contrats-de-recherche>.
- 'Data for History Consortium | Data for History'. Accessed 22 June 2022. <http://dataforhistory.org/>.
- DBNL. 'De grote raad van Mechelen, hoogste rechtscollège in de Nederlanden? H. de Schepper, Bijdragen en Mededelingen betreffende de Geschiedenis der Nederlanden. Deel 93'. DBNL. DBNL. Accessed 17 September 2024. https://www.dbnl.org/tekst/_bij005197801_01/_bij005197801_01_0016.php.
- De Jong, Hylkje, and Gijs Van Dijck. 'Network Analysis in Legal History: An Example from the Court of Friesland'. *Tijdschrift Voor Rechtsgeschiedenis = Revue d'histoire Du Droit = The Legal History Review* 90, no. 1–2 (June 2022): 250–62. <https://doi.org/10.1163/15718190-20220004>.
- Deroy, Aniket. 'How Ready Are Pre-Trained Abstractive Models and LLMs for Legal Case Judgement Summarization?', n.d.
- Deutsch, Andreas. 'Straf- und ordnungsrechtliche Regelungen in ländlichen Rechtsquellen und ihr spezifischer Wortschatz'. *Journal for Digital Legal History* 2, no. 1 (2 October 2023). <https://doi.org/10.21825/dlh.87176>.
- 'Digital Harlem'. Accessed 17 September 2024. <http://digitalharlem.org/>.
- Föhr, Pascal. 'Historische Quellenkritik Im Digitalen Zeitalter.' Basel, 2018. http://edoc.unibas.ch/diss/DissB_12621.pdf.
- 'Free Access to Law Movement | Institute of Advanced Legal Studies'. Accessed 18 September 2024. <https://ials.sas.ac.uk/leadership-and-collaboration/free-access-law-movement>.

- 'Gesetze Im Internet'. Accessed 18 September 2024. <https://www.gesetze-im-internet.de/>.
- GitHub. 'GitHub Copilot · Your AI Pair Programmer', 2024. <https://github.com/features/copilot>.
- Hitchcock, Tim, and William J. Turkel. 'The Old Bailey Proceedings, 1674–1913: Text Mining for Evidence of Court Behavior'. *Law and History Review* 34, no. 4 (November 2016): 929–55. <https://doi.org/10.1017/S0738248016000304>.
- Hodel, Tobias. 'Der machine learning Turn und rechtshistorische Dokumente'. Presented at the 44. Rechtshistorikertag, Frankfurt a.Main, 19 September 2024. <https://rechtshistorikertag2024.de/tagungsprogramm>.
- 'Homepage - Beck-Online'. Accessed 18 September 2024. <https://beck-online.beck.de/Home>.
- Hyvönen, Eero. *Publishing and Using Cultural Heritage Linked Data on the Semantic Web*. Synthesis Lectures on Data, Semantics, and Knowledge. Cham: Springer International Publishing, 2012. <https://doi.org/10.1007/978-3-031-79438-4>.
- Jong, H de. "'Copy-Paste" in Processen Aan Het Hof van Friesland. Over de Noodzaak van Archiefontsluiting'. *De Vrije Fries* 102 (2022): 124–37.
- . 'Over Het Nut van Digitale Archiefontsluiting. De Civiele Procesdossiers van Het Hof van Holland En Het Hof van Friesland'. *Pro Memorie* 24, no. 1 (June 2022): 97–111. <https://doi.org/10.5117/PM2022.1.005.JONG>.
- 'juris. Wissen, das für Sie arbeitet. | juris'. Accessed 18 September 2024. <https://www.juris.de/jportal/nav/index.jsp>.
- Kemman, Max. *Trading Zones of Digital History*. De Gruyter, 2021. <https://doi.org/10.1515/9783110682106>.
- Ketelaar, E. 'In Memoriam: Tom de Smidt (1923-2013)'. *Archievenblad* 117 (2013). <https://dare.uva.nl/search?identificer=0702ece1-452c-44c2-bf8d-0d7dc5b76b48>.
- Lefebvre Dalloz. 'Gagnez en savoirs, compétences et efficience !' Accessed 18 September 2024. <https://www.lefebvre-dalloz.fr/>.
- 'LexisNexis | Professionelle Online-Recherche-Tools'. Accessed 18 September 2024. https://www.lexisnexis.com/de-de?utm_source=google&utm_medium=cpc&utm_campaign=brandde&gad_source=1&gclid=Cj0KCQjw9Km3BhDjARIsAGUb4nwwEbq4Z2vWVMaiUWYuD00JWNeaSTik7O0dqXYkrJYyWFpHCQag5t8aArrbEALw_wcB.
- Limbach, Saskia and Vittorio Klostermann. *Government Use of Print in the Holy Roman Empire in the Sixteenth Century*, 2020.
- 'Lynching In Texas'. Accessed 17 September 2024. <https://lynchingintexas.org/>.
- Marino, Gabriele, Daniele Licari, Praveen Bushipaka, Giovanni Comandé, and Tommaso Cucinotta. 'Automatic Rhetorical Roles Classification for Legal Documents Using LEGAL-TransformerOverBERT'. In *Proceedings of the 6th Workshop on Automated Semantic Analysis of Information in Legal Text Co-Located with the 19th International Conference on Artificial Intelligence and Law (ICAIL 2023), Braga, Portugal, 23rd September, 2023*, edited by Francesca Lagioia, Jack Mumford, Daphne Odekerken, and Hannes Westermann, 3441:28–36. CEUR Workshop Proceedings. CEUR-WS.org, 2023. <https://ceur-ws.org/Vol-3441/paper4.pdf>.
- Meta Llama. 'Llama 3.1'. Accessed 17 September 2024. <https://www.llama.com/>.
- 'Modelling Social and Legal Facts in the Context of the Semantic Data for Humanities and Social Sciences (SDHSS) Ontology Ecosystem'. Accessed 17 September 2024. <https://www.lhlt.mpg.de/events/37642/2077780>.
- Mührmann-Lund, Jørgen. *Borgerligt regimente: politiforvaltningen i købstæderne og på landet under den danske enevælde*. Tidlig moderne, bind 13. København: Museum TusulanumForlag, 2019.

- . ‘Policing the Guilds’. In *Bringing the People Back In. State Building from Below in the Nordic Countries ca. 1500-1800*, edited by Knut Dørum, Mats Hallenberg, and Kimmo Katajala, 146–61. London, New York: Routledge, 2021.
- ‘NLLP Workshop 2024’. Accessed 19 November 2024. <https://nllpw.org/>.
- NRC. ‘Wat is er open aan slimme chatbot-maker OpenAI?’, 1 January 2023. <https://www.nrc.nl/nieuws/2023/01/01/wat-is-er-open-aan-openai-a4153050>.
- ‘OpenAI’. Accessed 17 September 2024. <https://openai.com/>.
- Padfield, Joseph, Charlotte Bolland, Neil Fitzgerald, Anne McLaughlin, Glen Robson, and Melissa Terras. ‘Practical Applications of IIF as a Building Block towards a Digital National Collection’. Zenodo, 22 July 2022. <https://doi.org/10.5281/zenodo.6884885>.
- Peric, Lazar, Stefan Mijic, Dominik Stambach, and Elliott Ash. ‘Legal Language Modeling with Transformers’. In *Proceedings of the Fourth Workshop on Automated Semantic Analysis of Information in Legal Text Held Online in Conjunction with the 33rd International Conference on Legal Knowledge and Information Systems, ASAIL@JURIX 2020, December 9, 2020*, edited by Kevin D. Ashley, Katie Atkinson, Luther Karl Branting, Enrico Francesconi, Matthias Grabmair, Vern R. Walker, Bernhard Waltl, and Adam Zachary Wyner, Vol. 2764. CEUR Workshop Proceedings. CEUR-WS.org, 2020. <https://ceur-ws.org/Vol-2764/paper2.pdf>.
- ‘Quill Project’. Accessed 17 September 2024. <https://www.quillproject.net/>.
- Ranieri, Filippo. *Juristische Dissertationen deutscher Universitäten 17.-18. Jahrhundert*. Ius commune. Frankfurt a.M.: Klostermann, 1986.
- Ranieri, Fillipo. ‘Eine Datenbank über juristische Dissertationen und Juristen im Alten Reich: ein Projektbericht’. *Historical Social Research* 11, no. 1 (1986): 109–15. <https://doi.org/10.12759/hsr.11.1986.1.109-115>.
- ‘Republic’. Accessed 20 September 2024. <https://republic.huylgens.knaw.nl/>.
- Robertson, Stephen. ‘Digital Humanities’. In *The Oxford Handbook of Law and Humanities*. Oxford University Press, 2019. <https://doi.org/10.1093/oxfordhb/9780190695620.013.1>.
- . ‘Searching for Anglo-American Digital Legal History’. *Law and History Review* 34, no. 4 (November 2016): 1047–69. <https://doi.org/10.1017/S0738248016000389>.
- Romein, C. A. ‘Early Modern State Formation or Gute Policy? The Good Order of the Community’. *The Seventeenth Century* 37, no. 6 (2 November 2022): 1031–56. <https://doi.org/10.1080/0268117X.2022.2111339>.
- Romein, C. Annemieke, Max Kemman, Julie M. Birkholz, James Baker, Michel De Gruijter, Albert Meroño-Peñuela, Thorsten Ries, Ruben Ros, and Stefania Scagliola. ‘State of the Field: Digital History’. *History* 105, no. 365 (2020): 291–312. <https://doi.org/10.1111/1468-229X.12969>.
- Romein, Christel Annemieke. ‘The Formative Role of Early Modern Books of Ordinances. The Low Countries and Their Overseas Lands’. In *The Early Modern State: Drivers, Beneficiaries and Discontents*, edited by P. Brandon, L Heerma van Voss, and C. A. Romein, chapter 4. Lonon, 2022.
- Romein, Christel Annemieke, Sara Veldhoen, and Michel de Gruijter. ‘The Datafication of Early Modern Ordinances’, *DH Benelux Journal*, 2 (2020). <http://journal.dhbenelux.org/journal/issues/002/article-23-romein/article-23-romein.pdf>.
- Salmi, Hannu. *What Is Digital History? What Is History?* Cambridge, UK: Polity Press, 2021.
- Savelka, Jaromír, Hannes Westermann, and Karim Benyekhlef. ‘Cross-Domain Generalization and Knowledge Transfer in Transformers Trained on Legal Data’. In *Proceedings of the Fourth Workshop on Automated Semantic Analysis of Information in Legal Text Held Online in Conjunction with the 33rd International Conference on Legal Knowledge and Information Systems, ASAIL@JURIX 2020, December 9, 2020*, edited by Kevin D. Ashley, Katie Atkinson, Luther Karl Branting, Enrico Francesconi, Matthias Grabmair, Vern R. Walker, Bernhard Waltl, and Adam

- Zachary Wyner, Vol. 2764. CEUR Workshop Proceedings. CEUR-WS.org, 2020. <https://ceur-ws.org/Vol-2764/paper5.pdf>.
- 'Scholarly Resources for Learning and Research | Gale'. Accessed 18 September 2024. <https://www.gale.com>.
- 'Semantic Data for Humanities and Social Sciences | CIDOC CRM'. Accessed 17 September 2024. <https://cidoc-crm.org/Resources/semantic-data-for-humanities-and-social-sciences>.
- 'Slave Societies Digital Archive'. Accessed 17 September 2024. <https://slavesocieties.org/>.
- 'SSRQ Online'. Accessed 17 September 2024. <https://www.ssrq-sds-fds.ch/projekte/ssrq-online/>.
- 'Startseite : Regesta Imperii'. Accessed 18 September 2024. <http://www.regesta-imperii.de/startseite.html>.
- Usher, Carol. 'How to Model Roles in the CIDOC-CRM RDF Encoding'. Accessed 2 April 2019. <http://www.cidoc-crm.org/sites/default/files/Roles.pdf>.
- Volkaert, Florenz. 'OK Computer? The Digital Turn in Legal History: A Methodological Retrospective', 18 June 2021. <https://doi.org/10.1163/15718190-12340011>.
- Wagner, Andreas. 'Rechtsgeschichte'. In *Clio Guide – Ein Handbuch Zu Digitalen Ressourcen Für Die Geschichtswissenschaften*, edited by Claudia Prinz, Jens Prellwitz, Thomas Meyer, Rüdiger Hohls, Wilfried Enderle, Silvia Daniel, Thomas Meyer, Jens Prellwitz, Annette Schuhmann, and Silke Schwandt, 3. erw. und aktualisierte Aufl. Berlin, 2023. <https://doi.org/10.60693/bsme-m777>.
- 'Westlaw – Legal Research Platforms'. Accessed 18 September 2024. <https://legal.thomsonreuters.com/en/westlaw>.
- Why OpenAI's ChatGPT Is Such A Big Deal*, 2023. <https://www.youtube.com/watch?v=pOmpqdIVCoo>.