Gqh2Bf5W4TdK

# The Secret History of the Mongols. A Digital History Approach

### anonym

### anonym

[ ©. Published by De Gruyter in cooperation with the University of Luxembourg Centre for Contemporary and Digital History. This is an Open Access article distributed under the terms of the [Creative Commons Attribution License CC-BY](https://creativecommons.org/licenses/by/4.0/)

Secret History of the Mongols, Digital History, Engineering Historical Memory

The Secret History of the Mongols (SHM) is the earliest Mongolian written primary source. Scholarship dated it to the thirteenth century CE. The SHM reveals critical insights into the Mongolian culture, historiography, and language, making it the subject of extensive academic research in several languages (e.g., Mongolian, English, Chinese, French, Russian). The increasingly large volume of academic literature analysing various aspects of the SHM makes it increasingly difficult for scholars to access all academic literature written worldwide. This research aims to combine historical methods and computational techniques to make the SHM machine-processable and, thus, machine-understandable. The parsing of the critical edition by Paul Pelliot (1949) and the English translation by Igor de Rachewiltz (2006) into an online content management system feeds an application on the Engineering Historical Memory (EHM) platform. This application allows users to access the source via the EHM tools for content search and visualisation. Each SHM’s section/story is tagged with metadata on EHM, including geographical, temporal, and story type information. This application allows the users to visualise historical information using infographics, retrieve online relevant resources automatically and in real-time, and get sentiment-analysis insights yielding concept parsing, emotion recognition, intensity ranking, and aspect extraction results. The SHM is linked to other primary sources on EHM through the Chronicles and Travel Accounts of Afro-Eurasia application. This new EHM application for the SHM showcases how computational history can manage and present research online with powerful ICT techniques and philological rigour.

## Introduction

The Secret History of the Mongols (SHM) is a native Mongolian narrative written in the 13th century, primarily documenting Chinggis Khan’s ascension (1155-1210). This primary source provides a rich insight into world history and Mongolian historiography, culture, and language life prior to and during Mongolian expansion (1211-1227). In addition, the source provides a detailed record of how the Mongolians viewed their government, social structure, and ideology. It, thus, provides a detailed description of Mongolian tribal life during the twelfth and thirteenth centuries, showcasing the concerns of the aristocracy and the relationship between chiefs and their subordinates (*De Rachewiltz 2015*). This period is vital as the twelfth and thirteenth centuries showcased a paradigm shift in how Mongolian society operated as nomadic Mongolian tribes coalesced into a larger Mongolian supertribe under the leadership of Chinggis Khan (*Rossabi 1994*).

Typically, the pastoral nomadic lifestyle of the Mongols would not lead to group sizes more prominent than the tribe. In a mobile society, subordinates could decamp and ignore the authority of a leader they did not like, making it difficult for a tribal chief to obtain trans tribal control (*Rossabi 1994*). As the only genuine native chronicle of Chinggis Khan’s life, SHM provides a detailed narrative of how Chinggis Khan secured the personal allegiance of other tribal chiefs (*De Rachewiltz 2015*). This paper acknowledges that this Mongolian unity under Chinggis Khan was short-termed as the tribal leaders pledged allegiance to specific rulers rather than the offices held by the rulers. By Ögödei Khan’s death, the Mongol unity forged by Chinggis Khan had collapsed (*Rossabi 1994*). Regardless, this period of Mongolian unity and expansion remains the most significant impact the nomadic population of inner Asia had on the sedentary world, showcasing their development and integration into the 13th century world systems.

Due to its uniqueness, SHM has been the subject of extensive research across various academic disciplines (*Jiang 2007*). Linguistically, this source is the richest corpus of preclassical Mongolian – fuelling its extensive philological research. However, as a historical record, SHM reliability is contentious because it shows many inaccuracies when its information can be cross-referenced with other sources. For example, Okada highlights many events invented or exaggerated by the author(s) to fit the broader narrative of the text (*Okada 1972*). Additionally, SHM includes many anachronisms when dating factual events, and they cannot be attributed to the copyist’s error as the incorrect dating is essential in the broader constructed narrative of the text (*Atwood 2007*). The authorship and sources of information for the SHM have also been the subject of intense scholarly debate (*Atwood 2007*). Despite these uncertainties about its authorship, the SHM was crafted with a clear purpose. As Mongolist Igor de Rachewiltz has argued, the work was “meant to serve as a guide and instruction,” reflecting a deliberate effort by the Mongol leadership to shape historical consciousness and provide a narrative for their descendants (*De Rachewiltz 2004*).

The extensive scholarship about SHM makes it a good candidate for computational history. As secondary literature on SHM accumulates, it becomes increasingly difficult for scholars to access all papers and books published worldwide and map the interactions between SHM and other historical sources. Computational history empowers the historian’s craft to manage the extensive collection of data amassed and comprehend the information using topic modelling and data visualisation (*Nanetti, Cheong 2018*). This paper showcases an application developed for SHM within the international initiative Engineering Historical Memory (EHM). This application aims to make the SHM’s historical information machine-readable and understandable by parsing its published edition and English translation into a scalable database. This result is empowered by automatically retrieving online relevant resources in real-time that link selected SHM stories to publications, images and videos, and other primary sources on the EHM platform. Thus, users can do searches using hierarchical infographics and other tools such as Historical Geographic Information System (HGIS) and sentiment analysis.

## Literature Review

Establishing the accuracy of SHM is complicated because there is no definitive manuscript version. The earliest version of the source, written in the Uyghur script, has been lost to time; the surviving versions are transcriptions in Chinese characters accompanied by Chinese translations (*Barrett 1992*). William Hung has provided an in-depth study of the transmission of the Chinese copies of the source (*Hung 1951*). The Ku certified copy 顧校本, a facsimile of a surviving Ming dynasty manuscript of SHM, certified by Gu Guangqi 顧廣圻 to be accurate in 1805, is the most widespread and extensively studied of these Chinese manuscripts. This manuscript was published as a wood-block edition in 1908 and republished using the photolithographic process in 1936 (*Hung 1951*). The Ulan collated copy and SibuCongkan 四部丛刊, a modern collection of historical Chinese sources, also reproduce the Ku-certified copy (*Ulan Collation 乌兰校勘 2012*). Due to the importance of the Ku-certified copy, the EHM application also links that manuscript as its underlying source material.

The Ku-certified copy of SHM is currently machine-readable due to an electronic retrieval system created by Di Jiang, allowing scholars to conduct text analysis of the source. For his application, Jiang uses the SibuCongkan 四部丛刊’s version of SHM, which is a reproduction of the Ku-certified copy. Di Jiang has created a complete electronic retrieval system to make the Chinese text machine-readable whilst maintaining the structural integrity of the manuscript in digital form. His application keeps all the information in the Chinese manuscript, including layout, volume, chapters, pages, characters, segmentation, interlinear translation, and Chinese translation (*Jiang 2007*). The structure of the Ku-certified copy is an arbitrary division made by Ming Scholars in the Sino-Barbarian Glossary, dividing the source into 12 chapters and 282 sections (*Hung 1951*). This organisation of the text is significant because the SHM’s translations by Paul Pelliot, Francis Woodman Cleaves, and de Rachewiltz adopted these divisions (*Cleaves 1982*,*Pelliot 1949*,*De Rachewiltz 2015*). By maintaining the structural layout of the Ku-certified copy, users of Jiang’s applications could reference SHM in a form understandable to most users. By making the source’s characters machine-readable, users could conduct text analysis on SHM. The electronic retrieval system allows users to search characters using: concordance, statistical and browsing methods. The concordance method enables users to search for a word and obtain the context around the words from the source; the statistical method allows users to get the frequency of a search term in each section of the source; the browsing method allows users to find a search term in the source (*Jiang 2007*). As Ku’s certified copy used by the application is a handwritten source, is it unlikely that Optical Character Recognition (OCR) programs alone could make the text machine readable. More likely, Jiang used a Handwriting Text Recognition (HTR) program as it could understand the variations in handwriting (*Romein, Kemman, Birkholz, Baker, Gruijter, Peñuela, Ries, Ros, Scagliola 2020*). Despite HTR’s usefulness, that tool also boosts a relatively significant Character Error Rate (CER). A tedious proofreading process with extensive expertise is still required, making this project a considerable task deserving commendation. The paper notes that while Jiang has published multiple journal articles in Chinese and English explaining his electronic retrieval system, he does not link the application in these publications. We were also unable to find his applications using multiple web crawlers.

## Research Gaps and Research Questions

As a computation linguistics tool, Jiang’s electronic retrieval system is focused on making the text in the Chinese manuscript machine-readable. Like most history databases, Jiang’s application was designed to be the end product of his research project (*Nanetti, Cheong 2018*). This approach brings some limitations; whilst the text is machine-readable, the information in the text is not machine-understandable. Machine-understandable means that a computer can comprehend and process the information from the text data; Jiang’s application only enables the computer to recognise the character data from the text, and human experts are still required to process the information in this text. Additionally, Jiang’s application cannot be scaled to include other translations and new research on the source as they are not within the scope of Jiang’s project. For historians interested in the narratives and content of the source rather than the textual data, the database needs to be scalable. It needs to include new information and linkages to other relevant resources about the topic. Due to these limitations, this paper asks how we can do the existing research on SHM not only machine-readable and processable but also allow machines to start understanding the text historically.

By parsing the information of SHM into a content management system and linking multiple translations with their corresponding secondary literature, we created a web application that could visualise the data using HGIS, hierarchical information graphs, and sentiment analysis. Additionally, as the application sits in the EHM platform, SHM is automatically linked to the other primary sources of the period available on EHM. This task is performed via the Chronicles and Travel Accounts of Afro-Eurasia (1205-1533 CE) application, retrieving also relevant online resources in real-time using the EHM search algorithm. Finally, by aggregating the broader literature on SHM with the source material, we can find historical significance between historical and computational sciences, empowering the user experience with Information and Communications Technology (ICT) techniques to infer new insights into the SHM content.

START HERMENEUTICS

## Methodology

END HERMENEUTICS

The electronic computer radically changed how knowledge societies work, and history is no exception. First, however, we must consider that computationally intensive technology—including data mining, machine learning, computer vision, and natural language processing algorithms—is not explicitly crafted for historical sciences (*Nanetti 2023*). That is why traditional scholarship experiences many challenges in transitioning bookish knowledge into the digital without diluting content or value. The international research community—about 130 scholars and software engineers—contributing to the Engineering Historical Memory (EHM) initiative addresses this caveat in a threefold approach. Firstly, it emphasises evaluating and developing pivotal digital tools—from information visualisation to image-to-image searches—while continually innovating algorithms tailored for historical analysis. Secondly, it fosters collaboration, inviting scholars and cultural institutions to digitise and share their exhaustive knowledge, making these diverse primary sources—from text and imagery to music—machine-understandable. Lastly, the EHM underscores the importance of software engineers in this transformation, converting static databases into dynamic web applications that provide real-time updates on relevant content, thus expanding the horizons of historical discovery.

There is rising interest in how digitalising resources like the SHM can facilitate interdisciplinary dialogue, especially where disciplines like history, archaeology, and climatology intersect. Such digital tools can aid historians in interpreting the SHM more effectively and in addressing complex questions arising from interdisciplinary interactions. It is imperative to note that the intent is not just to read the SHM better but to employ its digitalised form as a medium to probe, analyse, and possibly resolve challenges stemming from these multidisciplinary encounters. In alignment with this vision, the EHM is developing applications like “EHM-Socio-Environmental Hazards”, which aims to bridge the gap between historical and environmental sciences, thereby epitomising the potent possibilities of this digital transformation (*Nanetti, Daly, Cheong 2022*).

START HERMENEUTICS

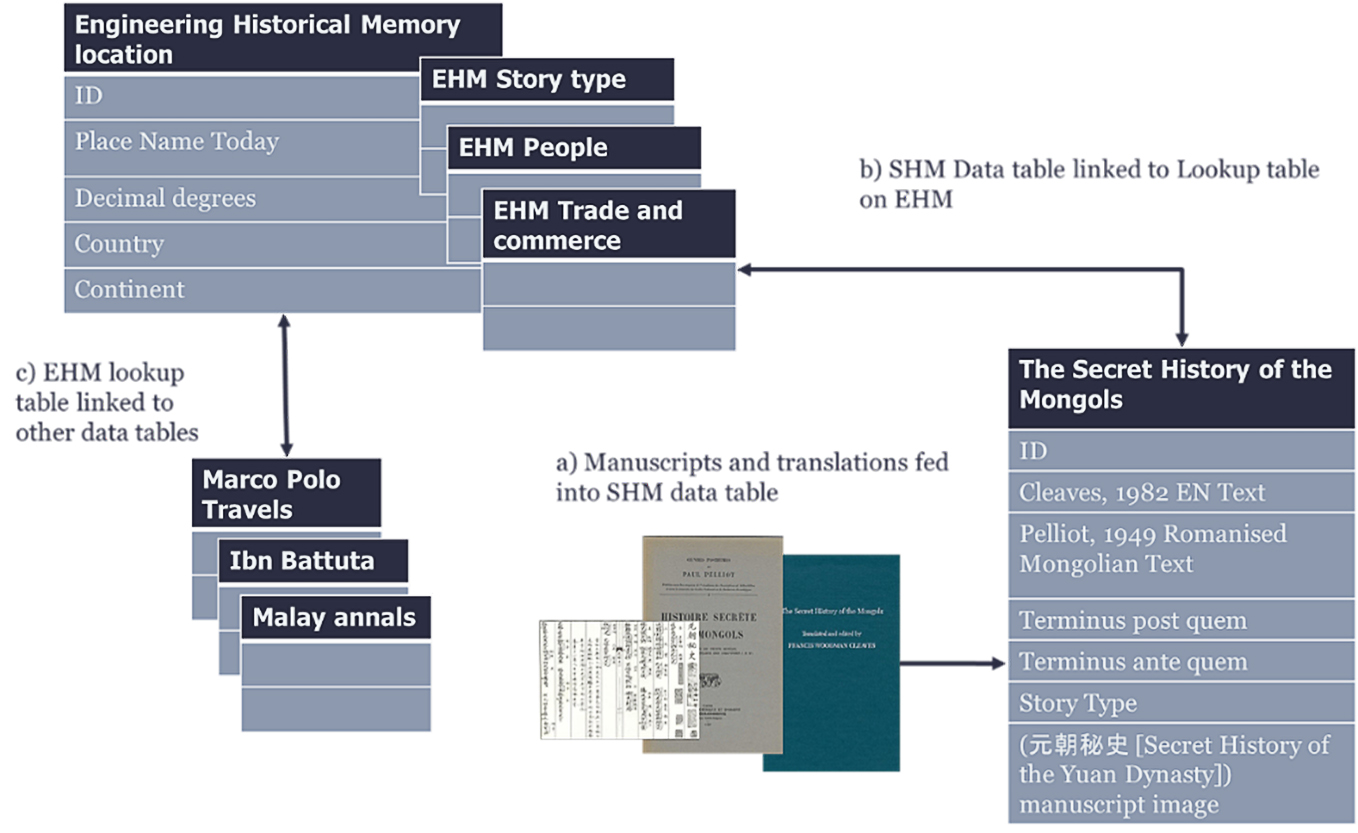
The database follows the widely adopted tradition of dividing SHM into 12 chapters and 282 sections, as discussed earlier. The texts are based on the critical edition by Paul Pelliot (1949), the English translation by Francis Woodman Cleaves (1982), the English translation by Igor de Rachewiltz (2006), and the Ku-certified copy manuscript. We tagged each section with its corresponding temporal data, geolocation data, and story types using secondary literature; this metadata allows the application to find connections between sections and other primary sources on EHM. The application cites the secondary literature using WorldCat as a unique identifier.

END HERMENEUTICS

START HERMENEUTICS

In discerning the nature of SHM, it is imperative to acknowledge its multifaceted identity. While some view it as a historical document, others interpret it as a literary masterpiece. Its dual character as a chronicle and a poem has been the subject of rigorous debate. Linguistic interpretations further intersect with translation issues, shaping our understanding of the pre-classical Mongolian language. Notably, translation disparities between Cleaves and De Rachewiltz are not mere stylistic choices but potentially reflect divergent interpretations grounded in distinct evidence bases. When creating an application for the SHM, there is a risk of inadvertently allowing computational methods to homogenise these nuanced interpretations. This paper, cognizant of such complexities, not only emphasises these challenges but also strives to address them comprehensively in the discourse and the proposed application.

END HERMENEUTICS



START HERMENEUTICS

Figure 1 shows how the application organises the data inputted. The text of the application serves as the basic unit of analysis for the users. However, as far as textual data is not structured in a way that a computer can understand its information content, we tagged each section with metadata to make the information machine-processable. As the EHM platform contains geolocation and story type lookup tables, we could link the geolocation and story types of the individual sections of SHM with unique identifiers from the EHM lookup table. These linkages allow us to find connections between SHM with other primary sources currently on EHM, like The Description of the World by Marco Polo, The Rhila by Ibn Battuta, and The Travels by Ma Huan. Using secondary literature, we can date the terminus post quem and terminus ante quem of an event to a range of years where the event could occur, improving the source’s veracity problems with machine-understandable dating.

END HERMENEUTICS

Due to their importance, the application’s textual data uses translations by Rachewiltz and Cleaves. Cleaves’ translation is the first complete English translation of SHM, and he created his work to be linguistically faithful to the original Mongolian work (*Cleaves 1982*). This translation choice makes it useful for anyone looking to learn the Middle Mongolian language whilst also making it difficult for readers to understand the contents of the text. Rachewiltz’s translation is closer to modern English, making it easier for most readers to understand (*De Rachewiltz 2015*). The Harvard-Yenching Institute published Cleaves’s work in 1982, leaving no digital book version. The application makes use of optical character recognition to make the text machine-readable. A digital version of Rachewiltz’s translation is available, distributed under the creative commons license, allowing us to copy it into the application quickly. As text data is unstructured, this paper applies sentiment analysis using the Sentic API to obtain information from the text, as explained below in the results section.

START HERMENEUTICS

We also populated the sections with the Mongolian text, written using French transcription, provided by Pelliot, along with the corresponding manuscript of the Ku-certified copy, which contains the Mongolian text in Chinese transcription, along with word-for-word translations in Chinese and a complete Chinese translation for the entire section. This inclusion allows linguists to make practical use of the application.

END HERMENEUTICS

SHM is a qualitative source that records geographical locations as place names. Place names complicate georeferencing as place names do not account for changes over time, disambiguating different places with the same name and multilingual issues (*Gregory, Geddes 2014*). The application georeferenced places to point locations rather than accounting for shifting administrative borders, which is impossible for many places in SHM due to a lack of information. Mongol academician Kh. Perlee provides a monograph on the place names of SHM, providing a valuable locator for the place names mentioned in the source through the place name changes since the time of Chinggis Khan (*Perlee, Moses, Moses 1985*). Perlee published his work in 1958. Our application used a gazetteer from the USA Board on Geographic Names to plot point locations using decimal degrees for coordinates (*USA Board on Geographic Names 1970*,*USA Board on Geographic Names 1988*,*USA Board on Geographic Names 1956*). Using a gazetteer from a similar period could account for the veracity of the geographical information recorded as we could understand geographically details similar to Perlee. The application records geolocation data using decimal degrees and modern place names to prevent the need to disambiguate different places with similar names and link the sections to other sources referencing the exact locations whilst using different names.

As mentioned above, many of the dates provided by SHM turn out to be inaccurate when historians cross-reference similar events to other sources. Christopher P. Atwood, in his attempt to date the source, has provided the terminus post quem and terminus ante quem for many of the events, referencing other historical sources that recorded the same events (*Atwood 2007*). Atwood’s work allows the application to improve the veracity in dating the source sections and solves the reliability issue of the source.

## Results

The SHM historical content is processable using existing information technology tools on EHM. This application uses the five main types of information visualisation presented in Liu Danyun’s participatory opportunities for digital history, which the paper will expand on below (*Liu, Xue, Cheong, Vu, Nanetti 2021*). The application allows users to visualise information in SHM using HGIS, hierarchical information graphs, and sentiment analysis. Additionally, as SHM is now on the EHM platform, users could use the Chronicles and Travel Accounts of Afro-Eurasia application to connect SHM with other primary sources available on EHM. Finally, users can obtain contextual information in real-time using Wikipedia as an identifier and the EHM search for publications, images, and videos.

### HGIS

One of the critical visualisations provided by the application is HGIS. Users can visualise the information in the SHM spatially using Satellite View and OpenStreetMaps, and both spatial visualisations serve different purposes. Users can also use their existing geographical knowledge to search for sections of SHM using these applications.

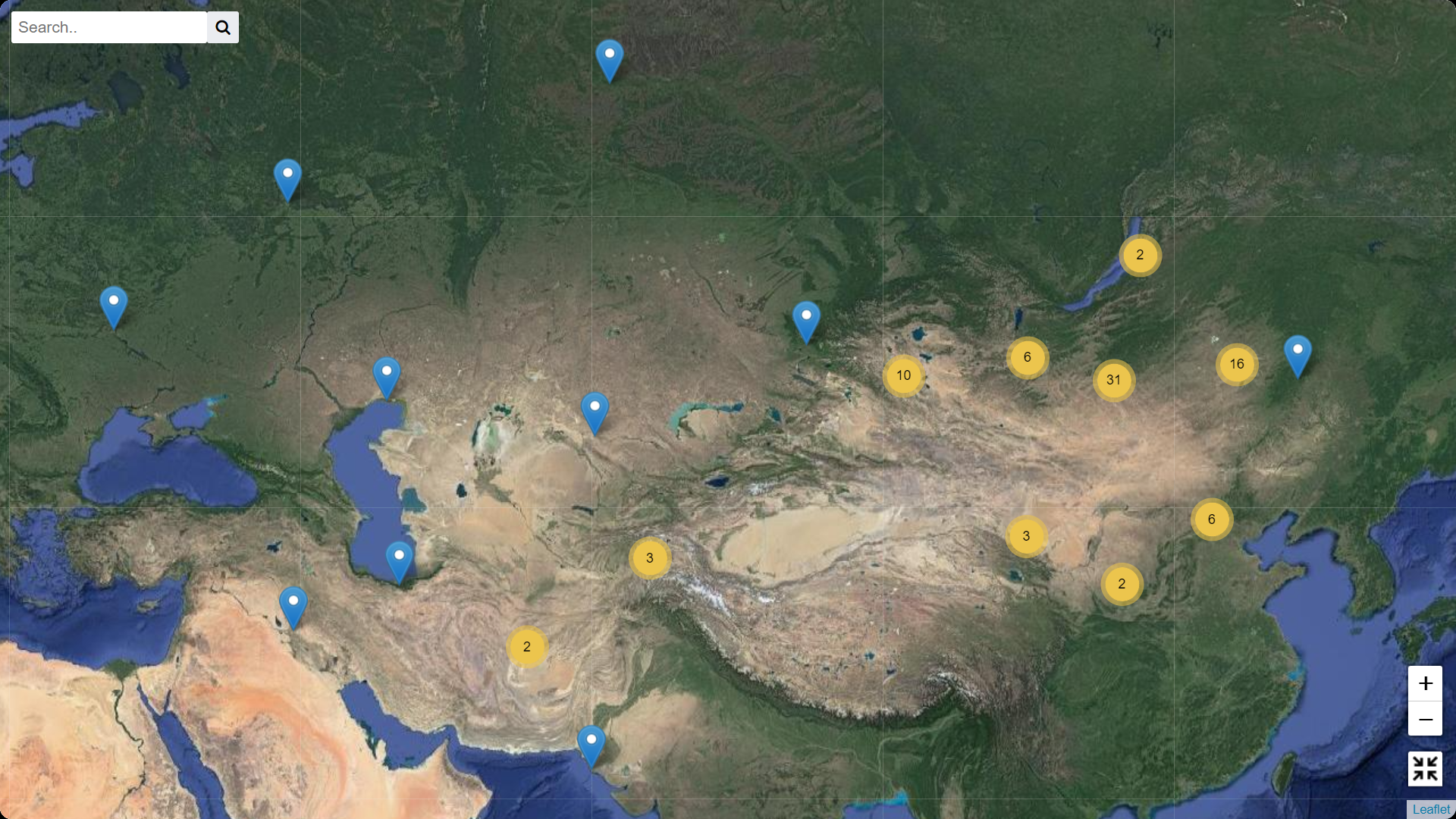


Figure 2 showcases the sections of SHM visualised on Sattelite view, an Earth Viewer System. Sattelite view gives high-resolution images of the Earth through photography and scanning, providing a realistic representation of the Earth (*Liu, Xue, Cheong, Vu, Nanetti 2021*). This visualisation allows users to understand the source spatially. Additionally, users could search for sections using a location search. As Sattelite renders a 3D representation of the Earth using satellite images, users could view the physical environment of the region reference in the text, aiding better understanding. Archaeologists have noted the Sattelite view’s ability to assist prospection and mapping in supporting spatial-temporal analysis (*Bitelli 2012*). The satellite view function could aid archaeologists as it allows archaeologists to identify the anthropic activity markers for archaeological sites before heading down for fieldwork (*Goncharov, Karelin, Mednikov, Nasyrov 2016*). As such, a use case of our application would be to help archaeologists scout archaeological sites referenced in SHM.

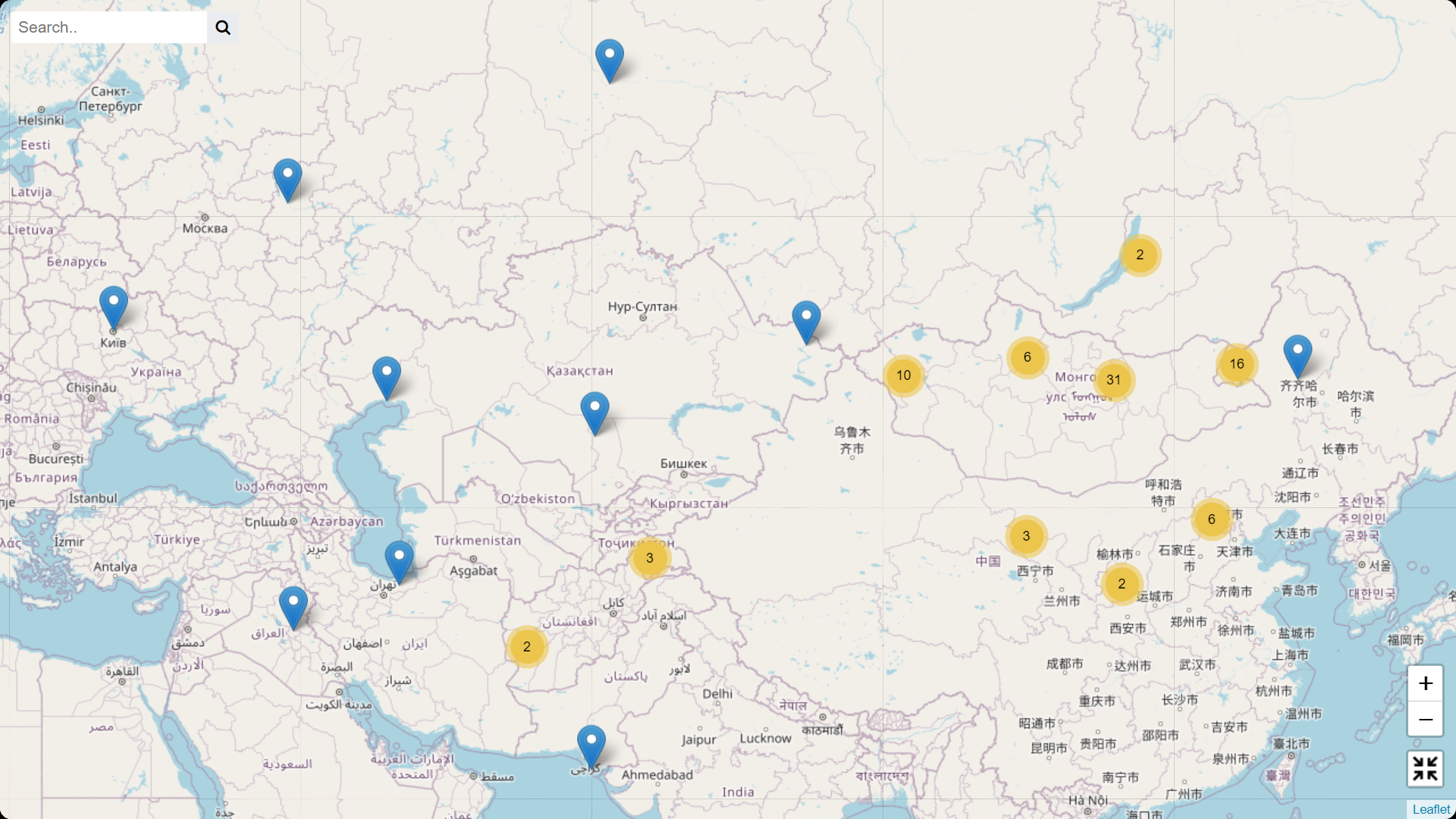


Figure 3 showcases the sections of SHM visualised on OpenStreetMap. Unlike Satellite view, OpenStreetMap is a conventional GIS application that showcases the city layout in flat Map format produced through user-generated street maps that can be freely editable (*Liu, Xue, Cheong, Vu, Nanetti 2021*). This application gives users a flat representation of Earth, presenting administrative districts and modern borders, which would benefit specific research questions.

### Geographical Visualisation of the Movements of Chinggis Khan

The geographical visualisation of the movements of Chinggis Khan and others would help investigate possible correlations with Mongolia’s environmental and climate maps. As the SHM narrates the movements of Chinggis Khan, we were able to make an itinerary of his movements. To ensure the veracity of the data, we have provided notes from other scholars who found discrepancies with the dating of SHM’s events along with Chinese and Iranian sources. Scholars have innovatively constructed a climate map of Mongolia during the 12th and 13th centuries, using tree-ring data and sediment analysis to supplement the gap in historical records (*Pederson, Hessl, Baatarbileg, Anchukaitis, Di Cosmo 2014*). This climate data has been applied to understand the Mongolian military developments in Europe and the Middle East (*Büntgen, Di Cosmo 2016*,*Di Cosmo, Wagner, Büntgen 2021*).

### Hierarchical Information Graphs

Aside from HGIS, hierarchical information graphs are another critical visualisation provided by the application. Users can visualise the information in the SHM using information visualisations like Treemap Chart, Force-directed graph, and Radial Tree.

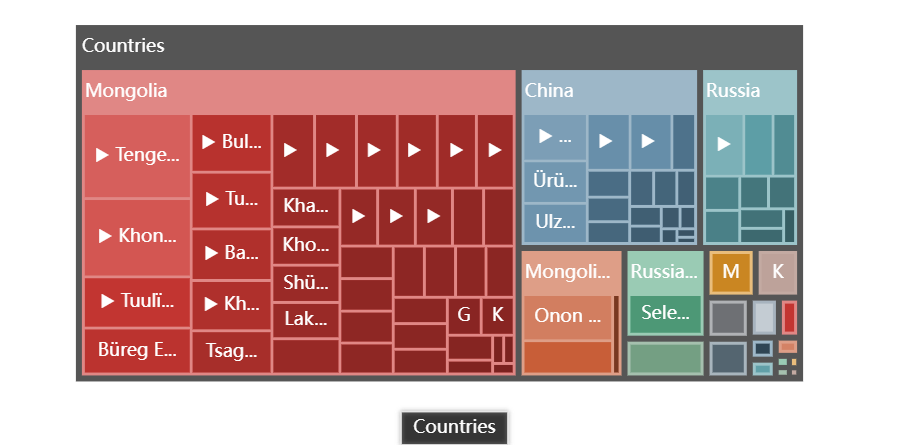


Figure 4 showcases the sections of SHM visualised on Treemap Chart. Treemap Charts is a hierarchical information graph. Treemapping displays extensive hierarchical data as a set of nested rectangles. As history is continuous and connected, hierarchical data allows users to identify the relationship between different objects (*Liu, Xue, Cheong, Vu, Nanetti 2021*). However, as a tool to represent spatial data, the node location does not correspond to its known geographical location, preventing users from locating areas using their geography knowledge (*Ghoniem, Cornil, Broeksema, Stefas, Otjacques 2015*). This functionality is covered by both the satellite view and OpenStreetMap instead.



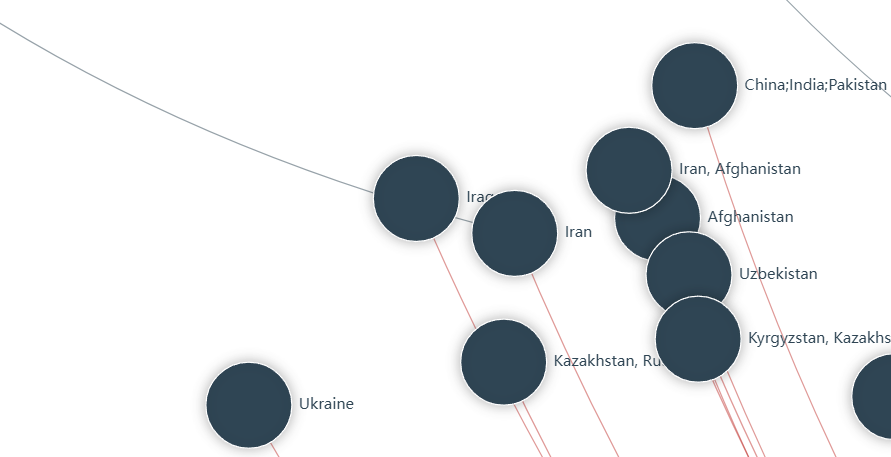


Figure 5 showcases the sections of SHM visualised as a Force-directed graph. This tool displays data nodes in a 2-dimensional plane, dynamically displaying the layout converging and stabilising – allowing users to understand large-scale datasets quickly and easily (*Liu, Xue, Cheong, Vu, Nanetti 2021*).

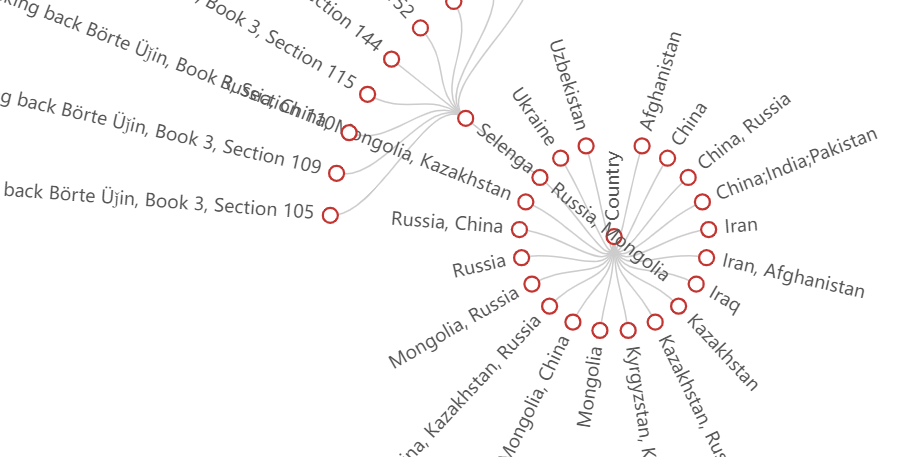


Figure 6 showcases the sections of SHM visualised as a Radial tree. A radial Tree is an information graphic that displays a tree structure that expands outward (*Liu, Xue, Cheong, Vu, Nanetti 2021*). In this application, we use hierarchical information graphs to present spatial data.

### Links to other primary sources

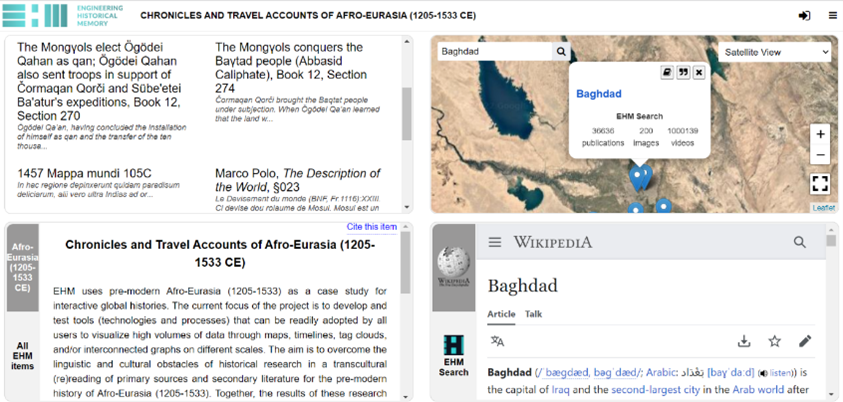


Figure 7 showcases the Chronicles and Travel Accounts of Afro-Eurasia application. The figure shows that the application links multiple sources with the same geographical data. This use case allows users to observe change over time through the ethnographic sources available on EHM. For example, the terminus post quem and the terminus ante quem of the Travels of Marco Polo are dated between 1269-1298. Marco Polo only left Venice after the death of Pope Clement IV, and he was in a Genoese prison with Rustichello da Pisa, the author of the source, at 1298, representing the end of his travels (*Latham 1958*). This dating places his ethnographic description of Baghdad 46 years after the terminus ante quem of SHM, proposed to be 1252 by Atwood (*Atwood 2007*). By viewing multiple sources describing Baghdad, users could observe the impacts of Mongolian rule on society.

### Sentiment Analysis of SHM



As the text data is unstructured, it is difficult for a computer to understand the information presented. A way to circumvent this is through Sentiment Analysis, a subset of natural language processing. This application uses Erik Cambria’s Sentic API to perform sentiment analysis. The Sentic API used by the application allows us to yield concept parsing, emotion recognition, intensity ranking, and aspect extraction results. Concept parsing identifies commonplace concepts from the text, making use of linguistic patterns and this aids in polarity classification. Polarity classification determines if the text is positive, negative or neutral. Intensity ranking then serves to rank the emotions and sentiments, allowing us to fine-tune our understanding of the intensity of each emotion. Finally, we use Emotion classification to contrast one emotion with another. Finally, aspect extraction allows us to understand each aspect’s polarity rather than the entire text and fine-tune our understanding (*Cambria, Nanetti 2022*).

Figure 8 showcases the Sentiment Analysis results for Section 79 of SHM. In this section, Tarqutai Kiriltuq, the leader of the Tayiči’ut tribe, attempts to capture and imprison Temüǰin after the Tayiči’ut tribe abandoned Temüǰin’s family after his father’s death (*De Rachewiltz 2015*). Naturally, Temüǰin would have felt grief and rage at this development, something the application correct pointed out.

### Retrieval of relevant online sources in real-time



Figure 9 shows the application giving contextual information on a query in real-time. For example, as we select the geographical location of Bukhtarma in the Satellite View application, the application automatically brings up the Wikipedia page for the location ‘Bukhtarma’. Aside from Wikipedia, users can also look at related publications, images and videos in the EHM Search column.

## Discussion

Our results have shown that information in SHM can be made machine-readable and processible. By making the information machine-readable and processible, we could experiment with tools, concepts and methods from other disciplines to find new insights into the source (*Romein, Kemman, Birkholz, Baker, Gruijter, Peñuela, Ries, Ros, Scagliola 2020*).

One main issue with converting written text into machine-readable text is the limitations of OCR itself. Digital historians can only apply OCR to a finite number of characters in a standard font (*Romein, Kemman, Birkholz, Baker, Gruijter, Peñuela, Ries, Ros, Scagliola 2020*). Herein lies the problem. In their search for linguistic accuracy, Cleaves has written his works in King James English (*Cleaves 1982*). Cleaves uses characters like ‘Γ č, š, γ Č ǰ, Š’ to convey a sound closer to the Mongolian pronunciation; these characters look similar to other English characters, giving rise to high Character Error Rates (CER). If the text’s veracity is low, the results from textual analysis like sentiment analysis suffer.

Additionally, the machine-learning sentiment analysis works by using a large corpus of affectively annotated textual data in the training phase (*Cambria, Hussain 2015*). However, as the application to digital history is new, the sentiment analysis API is not trained on historical data (*Nanetti 2023*). Furthermore, historical sources and datasets vary heavily in their linguistic structure as they are often translated from different languages from different periods. For example, Cleaves and Rachewiltz translated their English translations of SHM with significant variance in linguistics structure even within the same source. These considerations reduce the accuracy of sentiment analysis on digital historical primary sources. However, with more exposure to the relevant historical textual datasets as training data, the algorithm will only get more accurate.

One of the recent trends in Digital History is to produce Digital Scholarly Editions of primary historical sources . The application’s Critical editions and English translations are considered scholarly editions of SHM as Pelliot, Cleaves, and Rachewiltz have provided explanatory information in their commentaries, recorded in footnotes. By making the information machine-processable, we can create searchable and linkable connections between textual features and complement these links with infographics (*Romein, Kemman, Birkholz, Baker, Gruijter, Peñuela, Ries, Ros, Scagliola 2020*).

## Acknowledgements

This paper was initially conceived as a project under the NTU URECA (Undergraduate Research Experience on CAmpus) programme, an 11-month undergraduate research programme under the mentorship of a faculty member. Our deepest gratitude goes to Librairie d’Amérique et d’Orient and the Harvard-Yenching Institute for permitting us to incorporate the seminal works of Paul Pelliot and Francis Woodman Cleaves. We are equally thankful for our utilization of Rachewiltz materials licensed under Creative Commons. Additionally, We extend our appreciation to Professor Nicola Di Cosmo for his invaluable suggestions and insights towards this project. Any oversights or inaccuracies in this paper are entirely our responsibility. Financially, this research project has been generously supported by multiple esteemed institutions. Nanyang Technological University has generously funded this project through the URECA Undergraduate Research Programme. Additionally, our work benefited from the financial contributions of the Ca’ Foscari University of Venice (during the Visiting Scholar tenure from 15 June to 15 August 2019) and the Singapore Ministry of Education under its Academic Research Fund Tier 1 (2020-2023 RG45/20NS and 2023-2026 RT20/22; PI, Andrea Nanetti). Lastly, we cannot overstate the impact of the sentiment analysis algorithm provided by the SenticNet Team (Prof. Erik Cambria), which significantly bolstered our research capabilities and outcomes.

## Conclusion

SHM is an essential source for study due to its importance to the Mongolian people and as a representation of the 13th century world systems. However, the study of SHM is complex as the source information lacks veracity. The increasing scholarship on SHM also makes it difficult for anyone to understand every piece of literature on the topic entirely. Computational History serves to fix this issue by making use of the power of the computer. Today, most textual data of sources are machine-readable due to the mass-digitisation of libraries and archives using OCR since the 1990s. However, simply automating the OCR process for scanned document pagers is not enough for the information in sources to be machine-processable. By tagging each section with temporal, geolocation and story type metadata, we can make the information in SHM linkable, allowing us to present it with information technologies. This new EHM application for SHM showcases how computational history can manage and present research online with powerful ICT techniques and philological rigour. The application scenarios presented in this paper serve as potential use cases. However, unlike most historical databases, the content management system of SHM is designed to be scalable with new information, which will expand its functionalities and allow us to obtain new findings from the source. As such, this historical research of SHM is open to contributions by other scholars in collaboration with the author(s) and the peer reviewers.

Atwood, C. P.. “Informants and Sources for the”secret History of the Mongols” ”. Mongolian Studies, vol. 29, pp. 27–39, https://www.jstor.org/stable/43193428.

Atwood, C. P.. “The Date of the ‘secret History of the Mongols’ Reconsidered”. Journal of Song-yuan Studies, vol. 37, pp. 1–8, http://www.jstor.org/stable/23496277.

Barrett, T. H.. “The Secret History of the Mongols: Some Fresh Revelations”. Bulletin of the School of Oriental and African Studies, vol. 55, no. 1, pp. 115–19, https://doi.org/10.1017/S0041977X00002706.

Bitelli, G.. “Remote Sensing and Integration with Other Geomatic Techniques in Archaeology”. Satellite Remote Sensing A New Tool for Archaeology, edited by Rosa Lasaponara and Nicola Masini, Springer.

Büntgen, U.and N. Di Cosmo. “Climatic and Environmental Aspects of the Mongol Withdrawal from Hungary in 1242 CE”. Scientific Reports, vol. 6, no. 25606, pp. 1–9, https://doi.org/10.1038/srep25606.

Cambria, E.and A. Hussain. Sentic Computing A Common-sense-based Framework for Concept-level Sentiment Analysis. Springer.

Cambria, E.and A. Nanetti. “Ongoing. EHM – Sentiment Analysis for Digital History”. Engineering Historical Memory, https://engineeringhistoricalmemory.com/Sentiment.php.

Cleaves, F. W.. The Secret History of the Mongols for the First Time Done into English Out of the Original Tongue and Provided with an Exegetical Commentary. Harvard University Press.

De Rachewiltz, I.. The Secret History of the Mongols: A Mongolian Epic Chronicle of the Thirteenth Century. Edited by John C. Street, University of Wisconsin―Madison, http://cedar.wwu.edu/cedarbooks/4.

De Rachewiltz, I.. The Secret History of the Mongols A Mongolian Epic Chronicle of the Thirteenth Century Translated with A Historical and Philological Commentary. Brill.

Di Cosmo, N., et al.. “Climate and Environmental Context of the Mongol Invasion of Syria and Defeat at ’ayn Jālūt (1258–1260 CE)”. Erdkunde, vol. 75, no. no. 2, pp. 87–104, https://doi.org/10.3112/erdkunde.2021.02.02.

Ghoniem, M., et al.. “Weighted Maps: Treemap Visualization of Geolocated Quantitative Data”. SPIE 9397, vol. 9397, pp. 1–5, https://doi.org/10.1117/12.2079420.

Goncharov, A. E., et al.. “GIS and Satellite Remote Sensing for Archeology: Exploring Polar History”. Сибирский Журнал Науки И Технологий, vol. 17, no. 4, pp. 956–63.

Gregory, I. N.and A. Geddes. “Conclusions: From Historical GIS to Spatial Humanities: Challenges and Opportunities”. Toward Spatial Humanities Historical GIS and Spatial History, edited by Ian N. Gregory and Alistair Geddes, Indiana University Press, pp. 172–85.

Hung, W.. “The Transmission of the Book Known as the Secret History of the Mongols”. Harvard Journal of Asiatic Studies, vol. 14, no. 3/4, pp. 433–92, http://www.jstor.org/stable/2718184.

Jiang, D.. “The Electronic Data and Retrieval of the Secret History of the Mongols”. Data Science Journal, vol. 6, pp. 393–99, https://doi.org/10.2481/dsj.6.S393.

Latham, R.. The Travels by Marco Polo. Penguin Classics.

Liu, D., et al.. “Information Visualisation for Digital History Participatory Solutions for Reviewing Best Practices”. E-perimetron, vol. 16, no. 3.

Nanetti, A.and S. A. Cheong. “Computational History: From Big Data to Big Simulations”. Big Data in Computational Social Science and Humanities, edited by Shu-Heng Chen, Springer.

Nanetti, A., et al.. “EHM – Socio-environmental Hazards”. Engineering Historical Memory, https://engineeringhistoricalmemory.com/SocioEnvironmentalHazards.php?vis=sat.

Nanetti, A.. Computational Engineering of Historical Memories: With a Showcase on Afro-eurasia (ca 1100-1500 CE). Routledge.

Okada, H.. “The Secret History of the Mongols, a Pseudo-historical Novel”. Journal of Asian and African Studies, vol. 5, pp. 61–67.

Pederson, N., et al.. “Pluvials, Droughts, the Mongol Empire, and Modern Mongolia”. Proceedings of the National Academy of Sciences, vol. 111, no. 12, pp. 4375–79, https://doi.org/10.1073/pnas.131867711.

Pelliot, P.. Histoire Secrète Des Mongols : Restitution Du Texte Mongol Et Traduction Française Des Chapitres I a VI. [secret History of the Mongols: Restitution of the Mongol Text and French Translation of Chapters I to VI.]. Librairie d’Amerique et d’Orient.

Perlee, K., et al.. “On Some Place Names in the Secret History”. Mongolian Studies, vol. 9, pp. 83–102.

Romein, C. A., et al.. “State of the Field: Digital History”. History, vol. 105, pp. 291–312, https://doi.org/10.1111/1468-229X.12969.

Rossabi, M.. “The Legacy of the Mongols”. Central Asia in Historical Perspective, edited by Beatrice F. Manz, Westview Press, pp. 27–44.

Ulan Collation 乌兰校勘, .. “yuanchao Mishi Jiankan Ben” 元朝秘史 (校勘本) [the Secret History of the Mongols Collated Version]. Zhonghua Bookstore 中华书局.

USA Board on Geographic Names, .. Gazetteer No. 22. China. Official Standard Names Approved by the United States Board on Geographic Names. Prepared in the Office of Geography, Department of the Interior … June 1956. U.S. Government Printing Office.

USA Board on Geographic Names, .. Gazetteer of Mongolia: Names Approved by the United States Board on Geographic Names. Defense Mapping Agency.

USA Board on Geographic Names, .. United States Board on Geographic Names., & United States. (1970). U.S.S.R., Official Standard Names Approved by the United States Board on Geographic Names. Geographic Names Division.