

HistoLens: An LLM-Powered Framework for Multi-Layered Analysis of Historical Texts - A Case Application of Yantie Lun

Yifan Zeng

Sun Yat-sen University, Guangzhou, China
zengyf53@mail2.sysu.edu.cn

Abstract

This paper proposes HistoLens, a multi-layered analysis framework for historical texts based on Large Language Models (LLMs). Using the important Western Han dynasty text "Yantie Lun" as a case study, we demonstrate the framework's potential applications in historical research and education. HistoLens integrates NLP technology (especially LLMs), including named entity recognition, knowledge graph construction, and geographic information visualization. The paper showcases how HistoLens explores Western Han culture in "Yantie Lun" through multi-dimensional, visual, and quantitative methods, focusing particularly on the influence of Confucian and Legalist thoughts on political, economic, military, and ethnic. We also demonstrate how HistoLens constructs a machine teaching scenario using LLMs for explainable analysis, based on a dataset of Confucian and Legalist ideas extracted with LLM assistance. This approach offers novel and diverse perspectives for studying historical texts like "Yantie Lun" and provides new auxiliary tools for history education. The framework aims to equip historians and learners with LLM-assisted tools to facilitate in-depth, multi-layered analysis of historical texts and foster innovation in historical education.

1 Introduction

Traditional historical research often relies on qualitative methods that depend on the subjective judgments and experiences of scholars (Feng, 2023) and requires substantial time for reading historical texts. To address these challenges, this study proposes an interdisciplinary analysis framework, HistoLens, which combines natural language processing technologies, including Large Language Models (LLMs), to integrate qualitative and quantitative methods in historical research (Feng, 2023). This approach not only allows for drawing objective and verifiable conclusions but also enables

rapid processing of large volumes of historical text data, presenting knowledge through diverse charts and graphs to offer new insights and perspectives (Duan et al., 2023).

LLMs, which are trained on vast amounts of textual data, demonstrate powerful capabilities in understanding and generating text across various languages (Litaina et al., 2024), including classical Chinese (Yu et al., 2024; Chen et al., 2024; Zhao et al., 2024). The use of LLMs to assist in the analysis of historical texts has become a novel and effective approach for historians. With well-crafted prompts (Ozdemir, 2023; Qin and Eisner, 2021), historians can adeptly utilize digital tools for rapid and multifaceted textual analysis. The LLM systems utilized in this study include Claude 3.5 Sonnet (Anthropic, 2024), GPT-4o mini (OpenAI, 2024), as well as moonshot-v1¹ and ERNIE 3.5².

The LLM-powered historical text analysis framework includes the following main steps:

Step 1. thematic word frequency analysis: Tokenization, removal of stop words, and statistical analysis of key thematic word frequencies and their distribution in the original text.

Step 2. LLM-based named entity and relationship recognition: Identifying named entities such as individuals and places and extracting relationships between them by using LLM.

Step 3. knowledge graph construction: Building knowledge graphs based on entity and relationship information extracted by LLM.

Step 4. spatiotemporal analysis: Combining Geographic Information System (GIS) technology to visualize geographical distributions and time series in historical texts.

Step 5. ideological school analysis: Constructing text datasets containing historical thoughts, annotated with LLM assistance.

¹<https://kimi.moonshot.cn/>

²<https://yiyao.baidu.com/>

Step 6. machine teaching scenario: Based on the above ideological text datasets, using LLMs' text classification explanation capabilities to teach learners key features of thoughts in historical texts.

In this paper, we use "Yantie Lun" (Gale, 2022) as an example to demonstrate the application of HistoLens in analyzing Confucian and Legalist ideological cultures, which to a certain extent determined the politics, economy, and military affairs of the Western Han period. "Yantie Lun" which records the Salt and Iron Conference held in 81 BCE (Liu, 2021), is a crucial historical source for understanding the politics, economy, and ideologies of the mid-Western Han period. Through our proposed HistoLens, we explore and verify a range of cultural and historical viewpoints and phenomena in a multi-layered, visual, and quantitative manner, and offer new perspectives and vitality to the study of Western Han history, and construct a novel Confucianism and Legalism dataset while making a novel machine teaching scenario via explainable analysis. Details of the Salt and Iron Conference, Legalism and Confucianism are shown in appendix. The code and data are available.



Figure 1: (a) Confucius. (b) A photo of the Yantie Lun.

HistoLens is not limited to analyzing "Yantie Lun" but provides a generalized, universal LLM-assisted analysis framework for the study and learning of other historical texts. HistoLens offers new tools and methods for historical research and learning, enabling scholars to process large volumes of historical documents more rapidly and comprehensively, discovering new historical patterns and insights through multi-layered analysis.

2 Proposed Framework

2.1 Thematic word frequency analysis

Word frequency analysis is a commonly used paradigm in cliometrics (Grajzl and Murrell, 2024). Through word frequency analysis, we identified key terms of six themes: Confucianism, Legalism, agriculture, economy, Huaxia region, and ethnic minorities. Confucian terms outweigh Legalist

ones, as Sang also employs Confucian concepts in debate. This could indicate a synthesis of both two schools, reflecting "outer Confucianism, inner Legalism" (Zhou, 2011).

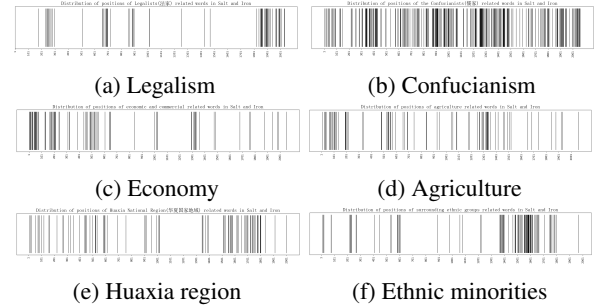


Figure 2: The distributions of topic keywords in text.

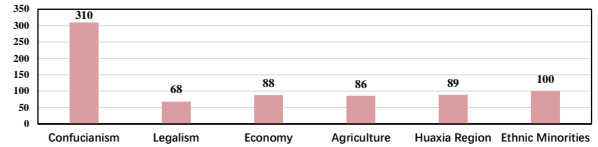


Figure 3: Bar chart of the number of lines.

They emphasized Han Dynasty's agriculture, state-controlled trade and commerce. It also discussed foreign relations, wars with the Xiongnu, and early Chinese nationalism. The language used showed a cultural identity linked to the Han people and their perception of other ethnic groups.

2.2 LLM named entity recognition

LLMs can effectively extract named entities such as individuals and places (Bogdanov et al., 2024). This allows knowledge in historical text to be represented through knowledge graphs (Windhager et al., 2024) and GIS (Giesecking, 2018).

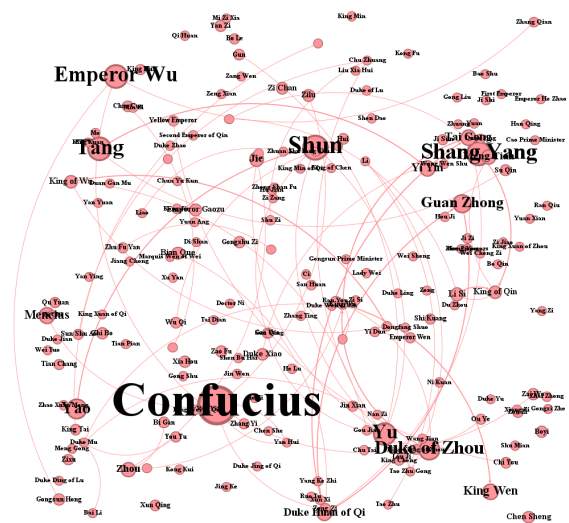


Figure 4: The knowledge graph of the figures mentioned in Yantie Lun.

Su Qin(苏秦) unified the six states with his diplomatic strategies, his achievement was not insignificant; Jie(桀) and Zhou(纣) were mentioned alongside Yao(尧) and Shun(舜), their names have not been forgotten to this day, their fame was not short-lived...

In the past ... from the Chanyu(单于) downwards, all were closely aligned with the Han, coming and going beneath the Great Wall(长城). Afterwards, Wang Hui(王恢)'s plot for Ma Yi(马邑) was mistaken, the Xiongnu(匈奴) severed their peaceful relations, attacked the strategic passes ...

In ancient times, Duke Huan of Qi(齐桓公) crossed Yan(燕) to campaign against the Shanrong, destroyed the Guzhu, and devastated Lingzhi. King Wu Ling of Zhao(赵武灵王) surpassed the Juzhu(句注), passed through the Daigu(代谷), ... crossed the Eastern Liao(辽东), and attacked Korea(朝鲜).

Figure 5: LLMs extract entities of figures and places (entities of people and places are represented in yellow and green, respectively).

We used LLMs to extract historical figures from the Yantie Lun, noting their frequency and relationships to created a graph using Gephi. Confucius, as the Confucian figurehead, was a central point of xianliang wenxue’s discussion. Shang Yang, representing Legalism, was both praised by Sang and criticized by Confucians, becoming a focal point of disagreement. After Dong Zhongshu’s promotion of Confucianism, figures such as Confucius, Mencius, King Wen, Yao, Shun, and Yu, who serve as symbols of Confucianism, received positive mentions from both sides. Even Sang cited Confucian exemplars, such as Shun, Yu and Tang. This indicates some convergence of Legalism and Confucian thought (Liu, 2024). Confucian ideas like the golden age of the Three Dynasties, became mainstream or political correctness in mid-Western Han. Moreover, Qin Shihuang and Shang Yang were frequently mentioned, highlighting the lessons drawn from the Qin Empire, which firstly achieved unification but quickly collapsed (Ouellette, 2010).

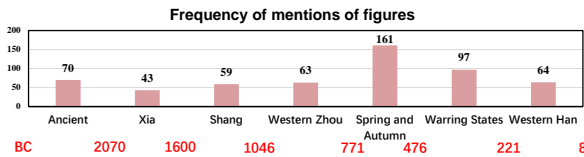


Figure 6: Historical period figure mention frequency.

Further analysis of figures mentioned more than twice by era (shown in Figure 6) reveals a tendency to cite ancient examples. This reflects a widespread tendency among Western Han intellectuals to look back to the past, venerate antiquity, and hold a historical view that favors the past over the present ("ancient" appears 121 times in the full text), which later influenced Wang Mang’s extensive restoration of ancient systems (Shi, 2007). This can also verify the Axial Age theory proposed by the Jaspers (Smith, 2015) and show that the Spring and Autumn philosophical diversity marked a creative peak in Chinese civilization.

We used LLMs to extract place name entities from the Yantie Lun and plotted them on a map using a geographic information system (GIS). This allows us to analyze the geographical and spatial understandings of the Western Han intellectual elite during that period, including their perception of the world layout and the relationship between the Huaxia region and surrounding ethnic minorities.

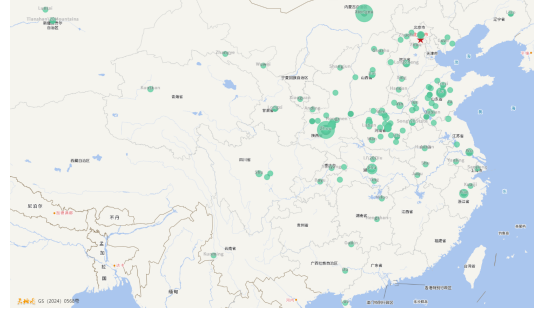


Figure 7: The map of the places mentioned in text.

The text discusses the historical concentration of place names along the Yellow River, emphasizing the region’s cultural and economic importance during China’s early dynasties, such as Xia, Shang, and Han. It contrasts this with the later southward shift of China’s centers during the Jin, Tang, and Song dynasties (Démurger et al., 2002). Notable areas include ancient states like Jin, Zheng, and Yan, as well as the Qi and Lu regions where Confucianism began. The Xiongnu’s Great Desert region and the Hexi Corridor, which led to the Western Regions and was pivotal for the Silk Road under Emperor Wu, are also highlighted (Zhang, 2021). Additionally, the text mentions the Yue people in the southeast, indicating their significance in Qin and Han times (Meacham, 1996).

2.3 Confucianism and Legalism dataset

Utilizing deep learning and natural language processing (NLP) algorithms, we aim to gain a deeper understanding of the Confucianism and Legalism ideologies in the Yantie Lun. By processing the text, extracting debates, and labeling passages that represent Confucian or Legalism views, we construct a dataset of these philosophical perspectives. The dataset can be utilized for various tasks such as text classification, sentiment analysis, creating educational tools, and text generation using NLP algorithms. Such an approach offers new perspectives and tools for studying and researching ancient Chinese philosophy and history of Western Han. The dataset contains 116 statements representing Le-

1	Prevent the source of debauchery, expand the beginning of morality, suppress the end benefits and ...	Confucian
2	The Xiongnu betrayed the subject, and repeatedly made raids and violence on the border, and prepa...	Legalist
3	Confucius said: "Those who have a country and a family, do not worry about poverty but worry abou..."	Confucian
4	The Xiongnu are rebellious and enter the border at will, violating China, killing and killing counties, ...	Legalist
5	In ancient times, it was noble to use virtue and despise the use of troops. Confucius said: "If the dista..."	Confucian
6	In ancient times, when establishing a country, it was necessary to open the way of the beginning an...	Legalist
7	Leading the people with virtue, the people will return to simplicity; showing the people benefits, the...	Confucian
8	Guan Zi said: "The country has fertile fields and the people are not enough to eat, because the equipm..."	Legalist
9	The country has fertile fields and the people are not enough to eat, and the industry and commerce ...	Confucian
10	In the past, the princes and vassals each paid tribute to their local products, and the coming and goi...	Legalist
11	In the past, when Shang Jun was the prime minister of Qin, he established the system and laws insid...	Legalist
12	In the past, when Emperor Wen was in power, there was no profit from salt and iron, but the people ...	Confucian
13	Shang Yang's steep law and long-term profit, the people of Qin could not live, and they cried togethe...	Confucian

Figure 8: Part of Confucianism and Legalism dataset.

galism views (from Sang Hongyang) and 123 representing Confucianism (from xianliang wenxue), totaling 239. Each statement contains an average of 188 Chinese characters, amounting to approximately 45,000 characters in total. The number of statements from both sides is roughly equal, showing balanced arguments. It can be seen that the author Huan Kuan, a Confucian scholar, did not show particular bias towards the Confucianism side due to his educational background. And this can avoid algorithmic bias caused by imbalance. The dataset contains complete sentences to avoid taking things out of context. The dataset has removed the identities of speakers to prevent the algorithm from guessing based on the speaker's identity. Data is in classical Chinese, with some rare characters used in ancient times, which poses difficulties for humans who have not received professional training, and also presents a challenge for NLP algorithms.

2.4 Interpretability analysis and machine teaching

Humanities scholars without a background in computer science can achieve surprisingly good results by using simple LLM prompting techniques to process historical texts. Numerous studies and practices have shown that some prompting tricks can make LLMs perform even beyond expectations. In a novel approach, we have employed prompting engineering to classify and analyze the interpretability of the dataset proposed in the previous section. The prompting engineering techniques we used include role-playing (having the model act as an expert), chain-of-thought (CoT) (Feng et al., 2024), and few-shot learning (Brown, 2020).

Interpretability analysis (Belinkov et al., 2020) enables models to provide the basis for text classification, allowing humans to understand the internal mechanisms behind the model's decision-making. This not only fosters trust in the decisions made by the model (or enables the rapid identification of errors in the model's decision basis), but also facilitates AI models in teaching humans

how to capture the decisive and key characteristics and differences between Confucian and Legalist thought. It achieves machine teaching while performing machine learning. Analysis can teach beginners that a text belongs to Confucianism as it contains terms like "benevolence", "morality" and "education", which are typical characteristics of Confucian thought. If machine learning involves humans giving vast amounts of data for models to learn from, machine teaching imparts knowledge acquired by models to humans (for example, helping beginners easily grasp the distinctions between Confucianism and Legalism). It can also offer humanities scholars novel perspectives after learning a large corpus of texts, presenting an innovative approach to learning and research.

Suppose you are an expert on ancient Chinese history and the history of philosophical thought (**Role Prompting**). Please analyze the given example and data step by step (**CoT**):

Step 1: Highlight the terms in the text data that are associated with Confucianism and Legalism.

Step 2: Infer whether the text belongs to Confucian or Legalist thought based on the highlighted terms.

Example: Input: ... (original text data); Analysis: ... (terms in the text data that are associated with Confucianism and Legalism); Label: ... (Confucian or Legalism); (**Few-shot**)

Input: In ancient times, **valuing virtue** and **despising the use of arms**. Confucius said: "If distant people do not submit, then cultivate **cultural virtue** to attract them ...". Now, abandoning **morality** and relying on military reforms, **raising armies** to attack, **stationing troops** to prepare, **exposing troops** to the elements ... and the **common people to toil and suffer** within ... beginning with **profit officials** to provide for them, is not a long-term strategy. Therefore, it is convenient to **abolish them**.

Analysis: "Confucius" represents a central figure in Confucian thought. "virtue", "morality" and "cultural virtue" highlighting the moral education. "common people toil and suffer" shows Confucianism's concern for the populace. "despising the use of arms", "raising armies", "stationing troops", and "exposing troops" reflects a critique of military reliance. "profit officials" and "abolishing them" indicates views on economic policy. Overall, it reflects Confucianism's advocacy for governance through virtue and opposition to over-reliance on military force.

Label: Confucian

Figure 9: Applying LLM prompt engineering for text classification and interpretability analysis on the dataset.

3 Conclusion

This paper proposes HistoLens, an LLM-powered framework for multi-layered analysis of historical texts, demonstrated through a case of the Western Han text "Yantie Lun". HistoLens offers a novel approach to historical research and education and goes beyond traditional qualitative analysis, providing a comprehensive exploration of historical time, space, figures, and ideologies. Through HistoLens, we've investigated key ideological features of the Western Han, such as the debate and integration of Confucianism and Legalism (as "outer Confucianism, inner Legalism"), historical decline mindset, the impact of the Axial Age, and the Huaxia national consciousness. HistoLens offers new tools for historians and learners to efficiently analyze historical documents. It not only offers fresh insights into "Yantie Lun", but also presents a generalizable approach for analyzing other historical texts across various periods and cultures, and uncover insights through multi-layered analysis. In conclusion, HistoLens represents a step forward in the application of AI to historical research and education.

Limitations

While HistoLens offers numerous advantages, it's important to acknowledge its limitations and areas for future improvement:

LLM Limitations: The effectiveness and accuracy of LLM on classical historical texts might be limited due to lack of specialized fine-tuning. Future work could involve developing or fine-tuning LLM specifically for historical texts' ancient language processing, attempting to further mitigate LLM's hallucinations regarding historical facts.

Scope of Application: Due to resource constraints, this study primarily applied HistoLens on "Yantie Lun," which represents a specific historical period (Chinese Western Han Dynasty). Expanding the application of HistoLens to a broader range of historical texts from different periods and cultures, such as Plutarch's "Parallel Lives" from the Roman period, would enhance its generalizability.

Comprehensiveness: While our proposed HistoLens's analysis covers multiple dimensions, such as historical figures and their relationships, locations and ideologies, it may not encompass all relevant historical phenomena, such as genre analysis of historical works from different periods. Further interdisciplinary collaboration could help improve HistoLens to identify and incorporate additional relevant factors.

Validation: The accuracy of LLM-assisted analysis needs rigorous validation against traditional historical research methods. Developing standardized evaluation metrics for computational historical analysis would be beneficial.

Ethical Considerations: The use of AI in historical interpretation raises questions about potential biases, such as implicit value judgments in historical events, which may be Eurocentric or reflect great-power chauvinism, potentially offending certain national sentiments. Future work should address these ethical implications and develop guidelines for responsible use of AI in historical research and learning.

User Interface and Accessibility: While not addressed in this paper, developing a user-friendly interface for HistoLens would be crucial for its widespread adoption among historians and educators who may not have extensive computer technical expertise.

Integration with Multimodal Historical Data: Despite language being the primary form of historical narration, future iterations of HistoLens could

leveraging Multimodal Large Language Model (MLLM) to enhance its analytical capabilities by integrating multimodal historical data. This includes the vast amount of historical photographs and even videos that have been preserved since the invention of photography.

4 Statements

We used LLM systems for academic research purposes and provided appropriate citations. We cited the sources appropriately when using "Yantie Lun" for academic research.

References

- Anthropic. 2024. Claude 3.5 Sonnet. <https://www.anthropic.com/news/claude-3-5-sonnet>. [Online; accessed 2024-10-03].
- Yonatan Belinkov, Sebastian Gehrmann, and Ellie Pavlick. 2020. Interpretability and analysis in neural nlp. In *Proceedings of the 58th annual meeting of the association for computational linguistics: tutorial abstracts*, pages 1–5.
- Sergei Bogdanov, Alexandre Constantin, Timothée Bernard, Benoit Crabbé, and Etienne Bernard. 2024. Nuner: Entity recognition encoder pre-training via llm-annotated data. *arXiv preprint arXiv:2402.15343*.
- Tom B Brown. 2020. Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*.
- Andong Chen, Lianzhang Lou, Kehai Chen, Xuefeng Bai, Yang Xiang, Muyun Yang, Tiejun Zhao, and Min Zhang. 2024. Benchmarking llms for translating classical chinese poetry: Evaluating adequacy, fluency, and elegance. *arXiv preprint arXiv:2408.09945*.
- Sylvie Démurger, Jeffrey D Sachs, Wing Thye Woo, Shuming Bao, Gene Chang, and Andrew Mellinger. 2002. Geography, economic policy, and regional development in china. *Asian Economic Papers*, 1(1):146–197.
- Siyu Duan, Jun Wang, Hao Yang, and Qi Su. 2023. Disentangling the cultural evolution of ancient china: a digital humanities perspective. *Humanities and Social Sciences Communications*, 10(1):1–15.
- Guhao Feng, Bohang Zhang, Yuntian Gu, Haotian Ye, Di He, and Liwei Wang. 2024. Towards revealing the mystery behind chain of thought: a theoretical perspective. *Advances in Neural Information Processing Systems*, 36.
- Zhiwei Feng. 2023. The four levels of digital humanities research. *Journal of Nanjing Normal University (Social Sciences Edition)*, (3):1–9. In Chinese.

- Esson McDowell Gale. 2022. *Discourses on Salt and Iron: A Debate on State Control of Commerce and Industry in Ancient China: Chapters I-XIX translated from the chinese of Huan K'uan with introduction and notes*, volume 2. Brill.
- Jen Jack Gieseeking. 2018. Where are we? the method of mapping with gis in digital humanities. *American Quarterly*, 70(3):641–648.
- Peter Grajzl and Peter Murrell. 2024. How machine learning will change cliometrics. In *Handbook of Cliometrics*, pages 2721–2750. Springer.
- Tsung-shun Lin. 2018. New interpretation on the discussion on salt and iron—the political intention of huoguang and the unperceived situation of sang hongyang. *Humanitas Taiwanica*, (89).
- Tania Litaina, Andreas Soularidis, Georgios Bouchouras, Konstantinos Kotis, and Evangelia Kavakli. 2024. Towards llm-based semantic analysis of historical legal documents.
- Dongwang Liu. 2024. Differences and integration of political thought between ancient chinese confucianism and legalism. *Trans/Form/Ação*, 47(4):e0240042.
- Qunyi Liu. 2021. Yantie lun in the pro-legalist and anti-confucian campaign. In *European and Chinese Histories of Economic Thought*, pages 202–214. Routledge.
- William Meacham. 1996. Defining the hundred yue. *Bulletin of the Indo-Pacific Prehistory Association*, 15:93–100.
- OpenAI. 2024. GPT-4O Mini: Advancing Cost-Efficient Intelligence. <https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/>. Accessed: 2024-10-4.
- Patrick Ouellette. 2010. *Power in the Qin dynasty: Legalism and external influence over the decisions and legacy of the first emperor of China*. Ph.D. thesis.
- Sinan Ozdemir. 2023. *Quick start guide to large language models: strategies and best practices for using ChatGPT and other LLMs*. Addison-Wesley Professional.
- Guanghai Qin and Jason Eisner. 2021. Learning how to ask: Querying lms with mixtures of soft prompts. *arXiv preprint arXiv:2104.06599*.
- Bertram Schefold. 2019. A western perspective on the yantie lun. In *The political economy of the Han dynasty and its legacy*, pages 153–174. Routledge.
- Feng Shi. 2007. A study on wang mang’s jade tablet for the feng-shan ceremony of the xin dynasty. *Chinese Archaeology*, 7(1):163–169.
- Andrew Smith. 2015. Between facts and myth: Karl jaspers and the actuality of the axial age. *International Journal of Philosophy and Theology*, 76(4):315–334.
- Ming Wan. 2012. Discourses on salt and iron: A first century bc chinese debate over the political economy of empire. *Journal of Chinese Political Science*, 17:143–163.
- Florian Windhager, Saminu Salisu, Johannes Liem, and Eva Mayr. 2024. The knowledge graph as a data sculpture: Visualising arts and humanities data with maps, graphs, and sets over time. *Geographical Research in the Digital Humanities; Bielefeld University Press: Bielefeld, Germany*, pages 113–134.
- Chengyue Yu, Lei Zang, Jiaotuan Wang, Chenyi Zhuang, and Jinjie Gu. 2024. Charpoet: A chinese classical poetry generation system based on token-free llm. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations)*, pages 315–325.
- Defang Zhang. 2021. Using excavated slips to look at effective governance of the northern frontier during the han dynasty—the lelang commandery in han slips. *Bamboo and Silk*, 4(2):336–364.
- Cheng Zhao, Bin Wang, and Zhen Wang. 2024. Understanding literary texts by llms: A case study of ancient chinese poetry. *arXiv preprint arXiv:2409.00060*.
- Haiwen Zhou. 2011. Confucianism and the legalism: A model of the national strategy of governance in ancient china. *Frontiers of Economics in China*, 6(4):616–637.

A Data Availability

The open-sourced code and data can be found here: <https://anonymous.4open.science/r/YantieLunData-8E57/> (anonymous git repository for double-blind review). The full text and book images of Yantie Lun can be found here: <http://www.xueheng.net/>.

B Detail background of the Salt and Iron Conference

"Yantie Lun" (Discourses on Salt and Iron) is an important political treatise and historical document from ancient China. It serves as a crucial historical source for understanding the politics, economy, and ideologies of the mid-Western Han period. The Yantie Lun compiled by Huan Kuan records the proceedings of a debate held in 81 BCE (the 6th year of Emperor Zhao’s Shiyuan era), which is known as the Salt and Iron Conference (Liu, 2021). The conference was organized under the influence of Huo Guang, and was presided over by Imperial Censor (yushi dafu) Sang Hongyang (Lin, 2018). More than 60 scholars (xianliang wenxue) from various regions were invited to participate in the

The conference (debate) and Yantie Lun appears to discuss economic policies (Schefold, 2019) on salt and iron, as well as warfare against the Xiongnu (Wan, 2012), but it is fundamentally a confrontation and clash between Legalist and Confucian ideologies and political philosophies represented by Sang Hongyang and the scholars (xianliang wenxue), respectively (Liu, 2021). This is the aspect that this paper focuses on exploring and illustrating.

In response, Sang Hongyang implemented a series of political and economic measures. These included state monopolies on essential goods like salt and iron production, alcohol taxation (jiuque), equitable transportation tax (junshu), price stabilization (pingzhun), and wealth control and taxation of merchants (suanmingaomin). These policies amassed vast wealth to support the empire's enormous military expenditures. However, these measures also exhausted the people's resources, making life unbearable for many. In his later years, Emperor Wu issued the "Luntai Edict of Self-Criticism", reflecting on his excessive use of the people's strength, a mistake reminiscent of the Qin dynasty's downfall (Lin, 2018). Consequently, Huo Guang, a powerful minister, sought to restore the Han Empire's strength, which emphasized allowing the people and the country to recuperate (Lin, 2018). Huo had risen to political prominence through his family connection to Emperor Wu's favored Consort Wei. It was under Huo Guang's initiative that the Salt and Iron Conference was convened. His aim may have been to use the criticisms of the Confucian scholars (xianliang wenxue) gathered from among the people to publicly reflect the popular opposition to the salt and iron policies. This would undermine Sang Hongyang's political capital and allow Huo to wrest control of financial and economic power from Sang Hongyang (Lin, 2018).

Confucianism, represented by Confucius, emphasizes moral cultivation, advocating for moral education to guide people towards goodness and social harmony. Legalism, as represented by Han Fei and Shang Yang, believes in the inherent selfishness of human nature and advocates for strict laws and punishments to maintain order. It bolsters monarchical authority and uses legal systems to centralize power and improve state efficiency.

We used LLMs to extract place name entities from the Yantie Lun, and mapped these ancient Chinese locations to their modern counterparts. Some required manual correction, such as "Kunshan" referring to the Kunlun Mountains, not the modern city in Jiangsu province today, and "Zhongshan" denoting the Zhao kingdom's city, not the Guangdong province city, and "Penglai" referring to a mythical offshore mountain, not Penglai in Shandong Province today.

王侯将相宁有种乎
 中国史
 春秋战国
 秦朝
 汉朝
 三国
 晋朝
 南北朝
 隋朝
 唐朝
 宋朝
 元朝
 明朝
 清朝
 民国
 新中国

Yanfe Lun (盐铁论) word frequency

Character	Frequency
天下	120
得	82
得	68
君子	62
得	58
得	52
得	48
得	42
得	40
得	38
得	35
得	32
得	30
得	28
得	25
得	22
得	20
得	18
得	15
得	12
得	10
得	8

7

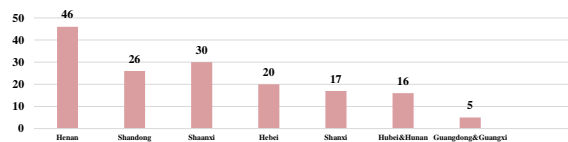


Figure 12: The geographical distribution of place names mentioned in the text.

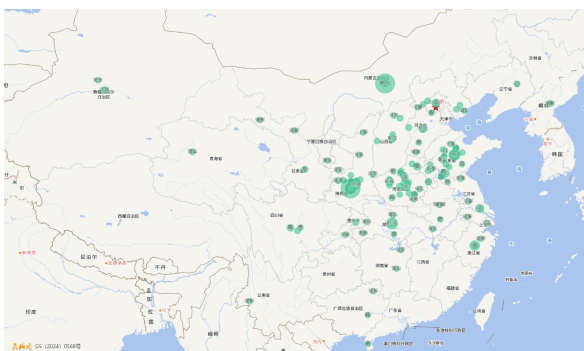


Figure 13: The map of places mentioned in the Yantie Lun (Chinese version).

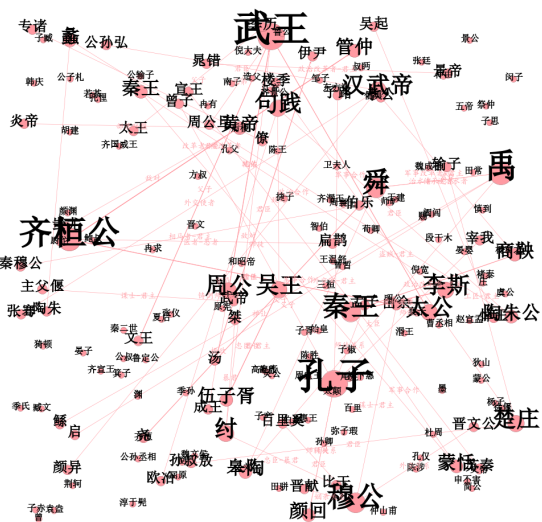


Figure 14: The knowledge graph of the figures mentioned in the Chinese version of Yantie Lun.

[illegible]

Figure 15: The Confucianism and Legalism dataset (Chinese version).