ISSN: 2581-8651
Vol-8, Issue-6, Nov-Dec 2026
Peer-Reviewed Journal

Journal of Humanities and Education Development
(JHED)

Grasping the Core Meaning of Ancient Texts with the Aid of Artificial Intelligence: An Integrative Study of Exegetics and Paleographic Knowledge Bases

Jheng-Yu Yin

Guangzhou Institute of Science and Technology, China
Email: 2952156647@qq.com

Keywords*— *Artificial Intelligence; Ancient Text Interpretation; Exegetics (Xungu); Paleographic Knowledge Base; Excavated Manuscripts.

Received: 03 Oct 2026;

Received in revised form: 31 Oct 2026;

Accepted: 04 Nov 2026,

Published on: 11 Nov 2026

©2026 The Author(s). Published by TheShillonga. This is an open-access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/)

Abstract

Ancient Chinese texts carry the core ideas of Chinese civilization, yet semantic drift across eras and disputes over the decipherment of excavated manuscripts often lead later readers to misinterpretation. Grounded in traditional exegesis (xungu), this paper integrates AI-based paleographic knowledge bases and explores, along three dimensions—the historical evolution of word meanings, the re-interpretation of classical passages, and the textual criticism of excavated manuscripts—how artificial intelligence can be fused with traditional philological research. The study shows that artificial intelligence can assist in constructing knowledge graphs, tracking conceptual evolution, and performing cross-textual intelligent retrieval, thereby preserving the rigor of traditional exegesis while enhancing the efficiency and breadth of ancient-text studies, and providing technical support for the creative transformation and innovative development of traditional Chinese culture.

I. INTRODUCTION

The development of Chinese writing has a long and continuous history. Over millennia of transmission and evolution, the semantic content of the same characters and words has often become far removed from what it was when they were first created. Changes across eras have led characters and words to develop numerous extended meanings; later generations become accustomed to current usage and forget their historical background. If the historical dimension is ignored in interpreting ancient texts, it is very easy to read the text literally and misinterpret it. Traditional exegetics takes graphology, phonology, and exegesis as its core methods and has provided a solid scholarly path for the interpretation of ancient texts; however, it relies on manual retrieval and textual research, so its efficiency is limited, and cross-text comparison is particularly difficult. This paper first uses examples such as “爪牙” (zhaoya; “claws and teeth”), “克己复礼” (keji fuli; “restrain oneself and return to propriety”), the Mencius passage “七十者可以食肉矣” (“those aged seventy may eat meat”), and the Guodian Chu slips 〈五行〉 (Wuxing; “Five Conducts”) and 〈性自命出〉 (Xing Zi Ming Chu; “Human Nature Derives from Mandate”) to demonstrate the interpretive efficacy of exegetical methods; it then explores how artificial intelligence’s knowledge graphs, diachronic analysis, and intelligent retrieval can assist exegetical research, in order to establish a new path for the study of ancient texts that combines “experience-driven” and “data-driven” approaches.

II. THE HISTORICAL EVOLUTION OF CHARACTER AND WORD MEANING AND THE REINTERPRETATION OF EXEGESIS

Take “爪牙” (zhaoya; “claws and teeth”) as an example. People today mostly interpret “爪牙” as “党羽” (partisans) or “帮凶” (henchmen), as in the “Biographies of Cruel Officials” in the Shiji [1]: “深吏多为爪牙用者” (“harsh officials were mostly used as claws and teeth”). However, Mao Shi Zhengyi: Qifu [2] glosses “予王之爪牙” (“I am the king's claws and teeth”) as “爪牙是勇力者也” (“claws and teeth are those of courage and strength”). The same “爪牙” receives different interpretations, which is in fact due to changes over time: in the Shi Jing, Lüshi Chunqiu, and Xunzi, “爪牙” either refers to the sharp teeth and sharp claws of birds and beasts, or is extended to mean “guards” or “martial valor”; by the Shiji, it had shifted to the meaning “partisans” or “henchmen,” while the Hanshu still preserves the interpretation “martial officials.” It can be seen that during the two Han periods, neutral and negative statements of the term “爪牙” coexisted, whereas people today retain only the negative meaning. The author therefore holds that the closer the commentaries and subcommentaries on a classic are to the era in which the text was completed, the closer their habits of wording and understanding are, and the higher their credibility; thus, in interpretive materials, primary importance should be given to the determination of date, with priority placed on those close to the pre-Qin period.

Even if interpretations from a similar era are adopted, commentators and subcommentators may still distort the meaning of the classics because of subjective standpoints; at this point, one must turn to xiaoxue (traditional philology) and conduct exegesis oneself. For example, in the Analects: Yan Yuan: “Yan Yuan asked about benevolence. The Master said: ‘To restrain oneself and return to propriety is benevolence. Once one restrains oneself and returns to propriety, all under Heaven will return to benevolence.’” Ancient scholars interpreted “克己复礼” (keji fuli; “restrain oneself and return to propriety”) as “restraining one's desires and being able to guide them toward propriety.” Yet how could one person's restraint of private desires cause “all under Heaven to return to benevolence”? The author was perplexed by this and therefore reinterpreted it by exegetical method. The Shuowen Jiezi Zhu glosses “克” (ke) as “肩” (jian; “shoulder”); Duan's commentary states: “To undertake affairs with the shoulder is called 任 (ren), and is also called 克 (ke),” so “克” has the sense of bearing and being competent. “己” (ji) is “中宫也，象万物辟藏诎形也” (“the central palace, depicting the bent form of the myriad things withdrawing and hiding”); paired with “戊” (wu)'s “万物茂盛” (“the myriad things flourishing”), “巳” (si)'s “万物见，成文章” (“the myriad things appear and complete their patterns”), and “中央土” (central earth)'s “地之吐生物者” (“that which earth exhales and produces living things”), it can be extended to mean “万物化生之理” (“the principle of the transformation and generation of the myriad things”). In sum, “克己复礼” means “bearing the principle of the transformation and generation of the myriad things and being able to guide it toward propriety.”

Looking back at the earliest commentaries and subcommentaries on the Analects, He Yan's Collected Explanations quotes Ma Rong's gloss: “克己，约身” (“restraining oneself, restraining the body”) [2]. According to the Hou Hanshu, Ma Rong “达生任性，不拘儒者之节” (“understood life and followed his nature, not bound by the norms of Confucian scholars”) and indulged his passions and desires; He Yan lived in an era when the Cao and Sima clans contended for power, and the Confucian doctrine of names and moral teaching had long become a tool for shackling scholars. Interpreting “克己” as “约身” was in fact consistent with the social ethos of the time. Yet the transmission and study of Confucianism extended from the Son of Heaven above to commoners below; “克己复礼” should be the goal of the practice of moral self-cultivation and the highest realm of “仁” (ren; “benevolence”), and its meaning should not stop merely at restraining desires. From this it can be seen that even though Ma Rong and He Yan lived in an era close to the period in which the Analects circulated, their annotations were still mixed with the intellectual currents and subjective factors of their time, causing the overall coherence of the classic to become obstructed and difficult to understand; therefore, it is necessary to reinterpret it by exegetical method.

III. TEXTUAL AUTHENTICATION AND EXEGETICAL METHODS FOR EXCAVATED TEXTS

The research objects of this paper also include the Guodian Chu slips 〈五行〉 (Wuxing; “Five Conducts”) and 〈性自命出〉 (Xing Zi Ming Chu; “Human Nature Derives from Mandate”). These two texts are contemporary excavated documents, and no authoritative definitive interpretation of their graphs has yet been established. Take the graph “形” in the opening of 〈五行〉, “仁形于内，谓之德之行” (“Benevolence takes form within, and is called the practice of virtue”), as an example: Li Yan, in “Chujian ‘Wuxing’ Zhushi Lunshuo” (“A Discussion of Commentaries on the Chu-

Slip ‘Wuxing’”) [3], adopts the graph “形” and interprets it as “成” (completion); Liu Xinfang, in “Shi ‘Wuxing’ yu ‘Xici’ zhi ‘Xing’” (“Explaining ‘型’ in the ‘Wuxing’ and ‘Xici’”) [4], adopts the graph “型” and interprets it as “模型” (mould/model). The author holds that, as far as graph form is concerned, neither theory is problematic, because in the pre-Qin period the two graphs “形” and “型” were interchangeable: the Shuowen Jiezi Zhu glosses “形，象也” (“形 means image”) [5] and glosses “型，铸器之法也” (“型 is the method of casting vessels”) [5], and Duan's commentary explicitly states that the two graphs are loan characters mutually interchangeable; therefore, whether written as “形” or “型,” this does not affect the interpretation of this graph. This example shows the necessity of textual exegesis for the study of excavated texts—〈性自命出〉 has never appeared in transmitted Confucian literature, and no one over the ages has annotated it; exegesis may be called the most reliable research method.

Another example is the graph “性” in 〈性自命出〉. Its calligraphic form differs from the seal-script form of the ordinary graph “生,” and the two graphs are clearly written separately. “性” can originally be interchangeable with “生.” In the Xunzi: Rich State, “同欲而异知，生也” (“the same desires but different knowledge—this is 生”), Wang Niansun states that “生读为性” (“生 is read as 性”); its basis is precisely the phonological principles “声义同源” (“sound and meaning share the same origin”) and “凡同声多同义” (“where sounds are the same, meanings are mostly the same”) [6]. The author made a rough statistical survey: the graph “性” appears only three times in the Shi Jing, five times in the Shangshu, six times in the Zhouyi, and nine times in the Zuozhuan; there is no “性” graph in the Zhouli or Laozi; it appears three times in the Mozi, eighteen times in the Han Feizi, and twenty-five times in the Liji; although it reaches eighty-five occurrences in the Zhuangzi, all are in the “Outer Chapters” and “Miscellaneous Chapters,” and it does not appear in the “Inner Chapters,” which are generally recognized as authored by Zhuang Zhou himself. It can be seen that the concept “性” was not an important concept in the pre-Qin period; discussion of it gradually flourished with the Warring States as the dividing line. 〈性自命出〉 is the earliest pre-Qin document devoted specifically to “性”; that “性” and “生” are written separately may be precisely because the concept had just separated and its boundaries were not yet clear. If one does not adopt exegetical methods to distinguish and analyze the graphs when studying such documents, one has no way to judge whether the concept passed through a fuzzy stage in which writing was differentiated while intension had not yet been clearly divided. This is not an isolated case in the history of writing development: for example, in the evolution of “乞,” “气,” “气,” and “饩,” in ancient texts “乞” and “气” were interchangeable and their graph forms easily confused; later “饩” was borrowed and then abbreviated to “气” to express “云气” (cloud vapor), and the “cloud vapor” meaning of “乞” thus fell into disuse.

More than half of the author's research scope involves newly excavated texts; therefore, exegesis is primary, with textual analysis also adopted, and interpretations far removed in date from the excavated texts are excluded—the later the era, the more easily key characters and words acquire extended meanings appended by later generations, even far exceeding the original meaning. Thus it is still necessary to supplement with the historical-source method and theoretical analysis of key concepts, in order to test whether later scholars' interpretations accord with the overall theory of the entire text. Dai Zhen's “Gujing Jie Goushen Xu” (“Preface to A Recovery of Lost Explanations of the Ancient Classics”) states: “所以明道者，其词也；所以成词者，未有能外小学文字者也。由文字以通乎语言，由语言以通乎圣贤之心志” (“That by which the Way is made clear is its words; that by which words are formed has never been able to go beyond the graphs of xiaoxue. Through graphs one communicates with language; through language one communicates with the minds and intentions of the sages”) [7]. “Ti Hui Dingyu Xiansheng Shoujing Tu” (“Inscription on Mr. Hui Dingyu's Diagram of Transmitting the Classics”) also says: “故训明则古经明，古经明则贤人圣人之理义明” (“When the ancient glosses are clear, the ancient classics are clear; when the ancient classics are clear, the principles and righteousness of worthies and sages are clear”) [7]. The author greatly agrees with what Dai Zhen says.

Let us give an example. In the Mengzi: King Hui of Liang I, “七十者可以食肉矣” (“those aged seventy may eat meat”), both the Mengzi Zhushu and the Mengzi Jizhu interpret it as “七十非肉不饱” (“a person of seventy cannot be satiated without meat”); their basis lies in the “老” of “老者衣帛食肉” (“the aged wear silk and eat meat”)—the Shuowen glosses “老” as “言须发变白也” (“it means that the beard and hair turn white”), and the Liji: Quli also says “七十曰老” (“seventy is called lao”). However, the author considers this far from correct. This passage focuses on the economy and the storage of national strength: in a five-mu homestead planted with mulberry, the silk and cloth obtained could serve as currency in transactions in the pre-Qin and two Han periods; this is in fact economic wealth and not merely

clothing. Only when the number of mu of fields reaches seventy can one develop breeding, for breeding must be founded on a surplus of rice; therefore “五十” and “七十” are both units of field area, not ages. Moreover, among the seven states of the Warring States, only Qin implemented private land ownership; King Hui of Liang, to whom Mencius was replying, still practiced the well-field system, and the people's production gains had to be turned over to the lord, with only the surplus becoming their own. The commentary on the Yili: Pinli states “老，宾之臣” (“lao are the ministers of guests”), and the subcommentary says “大夫家臣称老” (“the retainers of a grand master are called lao”). Therefore the “老” in “老者衣帛食肉” refers to retainers of a grand master; such an interpretation both conforms to the Zhou ritual system and fits the historical background. What Dai Zhen criticized specifically refers to the Song Confucians' groundless fabrication; yet even the Han Confucians, who were skilled in ancient glosses, were not able to be comprehensive, let alone those who do not understand graphs and are not proficient in character meaning, who are even more prone to read the text literally and generate forced meanings.

The efficacy of exegesis can also be seen in contemporary debates. Liao Mingchun's “Lunyu ‘Fuzi Huyin’ Zhang Xinzheng” (“A New Proof of the ‘Father and Son Conceal Each Other’ Chapter in the Analects”) [8] uses exegesis to examine various commentaries on “㯳梧” in the Xunzi, the “Wood Section” of the Shuowen, and texts such as the Hanshi Waizhuan and Kongzi Jiayu, arguing that “隐” and “㯳” gloss each other, thereby opening a new interpretation of “父子互隐” (“father and son conceal each other”); Guo Qiyong and Ran Shijun's “‘Men Nei’ de Rujia Lunli—Jian yu Liao Mingchun Xiansheng Shangque Lunyu ‘Fuzi Huyin’ Zhang zhi Lijie” (“Confucian Ethics ‘Within the Gate’—Also a Discussion with Mr. Liao Mingchun on the Understanding of the ‘Father and Son Conceal Each Other’ Chapter in the Analects”) [9] likewise refutes this by exegesis, pointing out that in the Shuowen “隐” is classified under the “阜” radical and “㯳” under the “木” radical, and their graph origins differ, and additionally citing the Yupian and Guangyun for textual research. Both sides contended with graphology and phonology, which suffices to show that xiaoxue exegesis can both accurately grasp the definitions of characters and words and further reach interpretation of principles and meaning, greatly benefiting research in Chinese literature and philosophy. However, the author's primary research method is in fact motivated by “方法的怀疑” (“methodological doubt”): first, doubts are raised about existing interpretations (such as the divergent views of Han and Song Confucians on “性善” and the positioning of the Xunzi: Xing'e); then key concepts are examined one by one by exegesis; subsequently, logical judgments are made about the meaning of the preceding and following text, concluding with the most appropriate interpretation; and from this the similarities and differences among the theoretical systems of the various texts are induced. Therefore, in addition to “methodological doubt,” logical analysis and induction are also supplemented, together constituting the research method of this paper.

IV. APPLICATION AND ASSISTANCE OF ARTIFICIAL INTELLIGENCE TECHNOLOGY

Beyond traditional exegetical methods, contemporary artificial intelligence provides new tools for the study of ancient texts and can realize a paradigm shift from “experience-driven” to “data-driven”; its auxiliary role is mainly manifested in three aspects.

First, knowledge graphs can eliminate exegetical discrepancies [10-12]. As in the above-cited graph “老” in the Mengzi, which is interpreted as “old in age” or “retainer of a grand master,” the root of the discrepancy lies in different grasps of the historical context and institutional background. A knowledge graph can establish semantic associations between “老” and concepts such as “周礼” (Zhouli; “Rites of Zhou”), “井田制” (jingtianzhi; “well-field system”), “大夫” (dafu; “grand master”), and “家臣” (jiachen; “retainer”), forming a multidimensional knowledge network and visually presenting its multiple interpretations and associated evidence in different texts and under different institutional backgrounds. Through association analysis, it can be seen that when “老” is interpreted as “大夫家臣” (“retainer of a grand master”), its logical association with economic production units such as “五亩之宅” (“a five-mu homestead”) and “百亩之田” (“a hundred-mu field”) is closer, and it conforms to the land-system background from the Western Zhou to the Warring States, providing an objective and comprehensive chain of evidence for judging the merits of exegesis and effectively eliminating subjective conjecture.

Second, diachronic analysis can reconstruct conceptual evolution. For example, “爪牙” shifted from “勇力” (“courage and strength”) in the Shi Jing to “党羽” (“partisans”) in the Shiji; the traditional historical-source method relies on scholars' individual combing of texts and inevitably involves omissions. Using natural language processing technology [13,14], one can automatically scan massive texts from the pre-Qin to the two Han periods, accurately count the frequency of occurrence, contextual environment, and sentiment polarity of “爪牙” in each

period and each classic, and generate a semantic evolution map, marking the key nodes of its shift from neutral to negative. This not only verifies the paper's assertion that "during the two Han periods neutral and negative meanings coexisted," but also moves historical-source analysis from qualitative description toward precise verification combining quantitative and qualitative methods, greatly enhancing the persuasive force of the argument.

Third, intelligent retrieval can strengthen "methodological doubt." When doubts arise about the traditional interpretation of "克己复礼" (keji fuli; "restrain oneself and return to propriety"), researchers can use artificial intelligence systems to quickly retrieve all chapters in the Analects related to "克" (ke), "己" (ji), "礼" (li), and "仁" (ren), and juxtapose them for comparison with texts such as the Xunzi and Liji. Intelligent analysis can assist in discovering that if "克己" is interpreted merely as "restraining one's own desires," its logical chain with "天下归仁" ("all under Heaven return to benevolence") appears weak; conversely, the interpretation obtained by this paper's exegesis, "bearing the principle of the transformation and generation of the myriad things," is more consonant with the overall thought system of the Confucian "内圣外王" ("inner sage and outer king"). The rapid retrieval and association capabilities of artificial intelligence enable "methodological doubt" to be established on a more comprehensive and solid textual foundation, thereby enhancing the rigor and comprehensiveness of logical inference.

V. CONCLUSION

In summary, to grasp the core connotations of ancient texts, exegesis is of primary importance: through graphs one communicates with language, and through language one communicates with the minds and intentions of the sages; only when this is supplemented by "methodological doubt," the historical-source method, and logical analysis can one avoid reading the text literally and misinterpreting it, and avoid deviation in interpretation. Artificial intelligence is not intended to replace traditional exegesis, but serves as a powerful enhancement tool: through knowledge graph construction, conceptual evolution tracking, and cross-text intelligent retrieval, it provides data-based, visual, and systematic support for eliminating exegetical discrepancies, restoring historical context, and verifying logical inference. The integration of the two both preserves the rigor of xiaoxue and expands the efficiency and breadth of research, and can provide solid support for the creative transformation and innovative development of fine traditional Chinese culture.

REFERENCES

[1] Q. Sima, *Records of the Grand Historian*. Beijing, China: Zhonghua Book Company, Aug. 2011.

[2] Y. Ruan, Ed., *Commentaries and Explanations on the Thirteen Classics*. Beijing, China: Zhonghua Book Company, 2009.

[3] Y. Ding, Ed., *Commentaries and Discourses on Newly Excavated Bamboo and Silk Manuscripts*. Taipei, Taiwan: Taiwan Shufang Publishing Co., Ltd., May 2008.

[4] X. Liu, P. Pang, et al., *Guodian Chu Slips and Early Confucianism*. Taipei, Taiwan: Taiwan Guji Publishing Co., Ltd., May 2002.

[5] S. Xu, *Annotated Shuowen Jiezi*, with annotations by Y. Duan. Taipei, Taiwan: Tiangong Book Company, Sep. 1996.

[6] J. Zhu, *Phonology*. Taipei, Taiwan: Wunan Book Publishing Co., Ltd., 2009.

[7] Z. Dai, *Collected Works of Dai Zhen*. Taipei, Taiwan: Liren Book Company, 1980.

[8] M. Liao, "A new interpretation of the chapter 'Father and Son Conceal for Each Other' in the Analects," *Journal of Hunan University (Social Sciences)*, Mar. 2013.

[9] Q. Guo and S. Ran, "Confucian ethics 'within the family gate': A discussion with Mr. Liao Mingchun on the understanding of the chapter 'Father and Son Conceal for Each Other' in the Analects," *Journal of South China Normal University (Social Sciences)*, Feb. 2014.

[10] A. Hogan, E. Blomqvist, M. Cochez, et al., "Knowledge graphs," *ACM Computing Surveys*, vol. 54, no. 4, Art. no. 71, 2021, doi: 10.1145/3447772.

[11] J. Ouyang, Z. Liang, and S. Ren, "Research on the construction of a knowledge graph of large-scale extant Chinese historical classics," *Library and Information Service*, pp. 126–135, Apr. 2021.

[12] T. Chen, X. Zhao, X. Yang, and L. Lin, "Research and design of a general knowledge model for ancient Chinese documents," *Journal of Information Resources Management*, pp. 139–153, Jan. 2025.

[13] A. Kutuzov, L. Øvrelid, T. Szymanski, and E. Velldal, "Diachronic word embeddings and semantic shifts: A survey," in *Proc. 27th Int. Conf. Computational Linguistics*, 2018, pp. 1384–1397. [Online]. Available: https://aclanthology.org/C18-1117

[14] N. Tahmasebi, L. Borin, and A. Jatowt, "Survey of computational approaches to lexical semantic change," arXiv:1811.06278, 2018. [Online]. Available: https://arxiv.org/abs/1811.06278