

# AI-Empowered Oracle Bone Script Classification and Its Academic Implications: Paradigm Reconstruction and Practical Pathways

Xiaogang You, Liang Liang\*

School of Languages and Media, Anhui University of Finance & Economics, Bengbu 233041, Anhui, China

\*Corresponding author: 821818366@qq.com

## Abstract

Oracle Bone Script (henceforth OBS for short), as the core root of Chinese civilization and a precious fossil of early human cognition, has long faced core challenges in its research and teaching, including a vast and complex corpus, intricate glyph structures, a profound knowledge system, and singular modes of inheritance. This paper aims to systematically explore how artificial intelligence (AI) technologies can bring about profound paradigmatic change to this traditional “esoteric scholarship.” In the classification research dimension, this paper constructs a comprehensive enabling framework composed of Data Layer, Algorithm Layer, Knowledge Layer, and Application Layer. It delves into intelligent glyph recognition and periodization based on deep convolutional neural networks and self-supervised learning, deep-level semantic association mining in oracle bone inscription texts using cluster analysis and topic models, and the integration of multi-dimensional entity relationships and intelligent reasoning based on large-scale knowledge graphs, aiming to achieve a paradigm shift in OBS classification research from “experience-driven” to “data-driven.” In the teaching application dimension, based on constructivism and embodied cognition theories, this paper designs an intelligent teaching model following the “Perception-Understanding-Inquiry-Creation” sequence. It elaborates on core application scenarios such as immersive glyph-origin context construction based on AR/VR/MR, adaptive learning path planning based on learner profiles, an intelligent “Digital Diviner” dialogue system based on large language models, and interactive storytelling and gamified creation driven by AIGC. Finally, the paper, from both technical and humanistic perspectives, cautiously analyzes the ethical risks faced by AI applications, including data bottlenecks, algorithmic bias, “black box” decision-making, and the dilution of humanistic spirit, and prospectively proposes a future development path characterized by “Human-Computer Collaboration, Rationale Integration, and Ethics First.” This research aims to provide an integrated solution for the rescue protection, in-depth research, and innovative inheritance of OBS that combines technical forefront, theoretical systematicness, and practical feasibility.

## Keywords

Oracle Bone Script; Artificial Intelligence; Deep Learning; Classification Research; Human-Computer Collaboration.

## 1. Introduction

Since Wang Yirong's discovery of OBS in 1899, its research has spanned a century, forming a traditional paradigm exemplified by the “Four Masters of Oracle Bones” (Luo Zhenyu, Wang Guowei, Dong Zuobin, Guo Moruo). The core of this paradigm relies on scholars' profound paleographical expertise, extensive historical knowledge, acute literary intuition, and long-term practical experience accumulation [1, 2]. This “cottage industry” style individualized

research model, when confronting the massive, complex, and highly unstructured material system of approximately 150,000 fragments and over 4,000 distinct characters (of which only about 1,500 have been deciphered), inevitably reveals efficiency bottlenecks, strong subjectivity, and methodological limitations. Similarly, in terms of inheritance, traditional OBS teaching often relies on static images, rubbings, and abstract textual explanations, struggling to vividly and intuitively present the dynamic character-formation thinking, vivid social life scenes, and the broad historical and cultural context of the Shang Dynasty. This leads to widespread difficulties for learners: “hard to recognize, hard to remember, hard to understand, hard to resonate,” greatly limiting the popularization and revitalization of this cultural heritage. In recent years, disruptive breakthroughs in AI technologies, represented by deep learning, natural language processing, computer vision, and knowledge graphs, have provided an unprecedented set of technical tools for processing massive, high-dimensional, heterogeneous OBS data [3, 4]. The fundamental purpose of AI empowerment is not to replace scholars' deep contemplation and humanistic insight, but to liberate researchers from lots of repetitive and mechanical foundational work, while providing educators and learners with unprecedented immersive, personalized, and interactive teaching and experiential means. This signals that OBS research is moving from the “empirical science” era reliant on individual wisdom towards the “computational humanities” era based on big data analysis; its teaching is also shifting from one-dimensional “knowledge indoctrination” to interactive “cultural inquiry” and “meaning construction.”

This paper aims to transcend scattered technical application discussions and systematically construct a holistic theoretical framework and application model for AI-empowered OBS classification research and teaching practice. It will deeply analyze its core principles, key technologies, implementation pathways, and typical scenarios, and, adhering to a critical perspective, cautiously reflect on its potential technical limitations and humanistic ethical risks, ultimately proposing a human-computer collaborative development paradigm for the future, hoping to promote the revitalization of this ancient learning carrying the genes of Chinese civilization in the digital intelligence era.

## **2. AI-Empowered OBS Classification Research: A Paradigm Shift from Experience to Data**

OBS classification is the cornerstone of all research work, traditionally relying mainly on glyph structure analysis (based on the Shuowen Jiezi's “Six Writings”), textual formula comparison, and Mr. Dong Zuobin's “Five Periods” dating method. The introduction of AI enables classification work to evolve towards refinement, quantification, automation, and systematization, thereby achieving a fundamental transformation of the research paradigm.

### **2.1. Data Layer: Construction, Challenges, and Standardization of the OBS Digital Resource System**

High-quality, large-scale, structured data is a prerequisite for the success of any AI model. Building the OBS digital resource system is a foundational project, mainly including the following levels:

**Glyph Image Data:** Acquiring high-definition, multi-modal digital images of oracle bone artifacts and their rubbings through high-precision flatbed scanning, multispectral imaging, 3D laser scanning, and macro photography. This is the fundamental basis for subsequent computer vision processing [5].

**Text Semantic Data:** Performing precise collation and digital transcription of oracle bone inscriptions to form a structured text database containing original text, transcription, modern Chinese translation, punctuation, excavation number, archaeological context, and periodization

information. Examples include the electronic versions of the *Collection of Oracle Bone Inscriptions* and the *Supplement to the Collection of Oracle Bone Inscriptions* [2].

**Domain Knowledge Data:** Systematically sorting, evaluating, and structurally storing the research results of over a century in oracle bone studies, archaeology, and history, such as the history of glyph interpretation and various scholarly views, the trace of semantic extension, the classification system of inscription content (sacrifice, warfare, hunting, agriculture, etc.), Shang dynasty genealogy, historical geography, astronomy and calendar, sacrificial system, and official system, to build an expert knowledge base.

The Core Challenges lie in the high heterogeneity, noise, and academic contentiousness of the data. The quality of oracle bone rubbings and photographs varies due to preservation conditions and production techniques, with fragmentation and erosion being common; transcriptions of the same bone may differ significantly among scholars; academic viewpoints are diverse, making it difficult to establish a single standard. Therefore, data cleaning, format standardization, metadata specification development, and precise annotation based on expert consensus are key to breaking through the “data bottleneck” in AI model training.

## **2.2. Algorithm Layer (I): Intelligent Glyph Recognition, Clustering, and Style Analysis Based on Deep Learning**

Deep learning models like Convolutional Neural Networks have demonstrated performance surpassing traditional methods in image classification, segmentation, and recognition, opening new paths for OBS glyph research.

**Automatic Single Character Recognition and Retrieval:** Building large-scale annotated datasets containing images of both deciphered and undeciphered characters, training end-to-end CNN models (e.g., ResNet, EfficientNet) for glyph classification. The achievable function is: inputting an image of an OBS character, the model can output its corresponding modern Chinese character (for deciphered characters) or provide the most similar candidate characters, calculating confidence levels; for undeciphered characters, it can output a unique identifier for subsequent management. This technology can greatly improve the efficiency of character identification after oracle bone fragment reassembly, assisting scholars in preliminary screening [4, 6].

**Unsupervised Clustering and Undeciphered Character Exploration:** For a large number of uninterpreted OBS characters, unsupervised learning algorithms (e.g., autoencoders, t-SNE, UMAP for dimensionality reduction visualization, combined with K-means, DBSCAN for clustering) can be used for automatic feature learning and grouping. Its core value lies in: morphologically highly similar glyphs are automatically grouped into the same cluster by the algorithm, providing extremely valuable “relational” clues for deciphering undeciphered characters – characters within the same cluster are highly likely to be semantically related, usage-related, or be variant forms or simplified shapes of the same character. This is equivalent to drawing a “potential kinship map of undeciphered characters” for researchers [7].

**Quantitative Calligraphic Style Analysis and Automatic Periodization Assistance:** Mr. Dong Zuobin’s five-period theory for OBS relies partly on the evolution of calligraphic style, such as the majesty of the first period (Wuding), the neatness and law-abiding nature of the second period (Zugeng, Zu Jia), and the strict refinement of the fifth period (Di Yi, Di Xin). Deep learning models (e.g., style transfer networks, attention mechanism models) can be trained to learn and quantify the calligraphic style features of OBS from different periods (e.g., stroke thickness, structure density, line strength, layout composition), thereby calculating style similarity and predicting the period for unknown or disputed bones, providing objective, quantifiable auxiliary evidence for historical periodization [8, 9].

**Case Analysis:** Assuming a batch of undeciphered OBS characters, after nonlinear dimensionality reduction via an autoencoder, visualization is performed on a 2D scatter plot.

Researchers observe that the data point for undeciphered character “A” clusters closely near the known cluster for “马” (horse), while the data point for undeciphered character “B” is scattered between the clusters for “车” (chariot) and “行” (road, to go). This discovery strongly suggests: “A” is likely related to “horse,” perhaps a horse of a specific color or state; while “B” might be related to the action or scene of “chariot moving on the road.” Such data-driven discoveries can effectively narrow the scope of interpretation and inspire new research ideas.

### **2.3. Algorithm Layer (II): Deep Semantic Analysis and Content Mining of Inscriptions Based on Natural Language Processing**

The inscription texts are an “encyclopedia” for studying Shang Dynasty society, politics, economy, religion, and culture. Natural Language Processing techniques can go beyond traditional keyword search to achieve deep semantic understanding and knowledge discovery from text content.

**Named Entity Recognition and Relation Extraction:** Using sequence labeling algorithms based on pre-trained models (e.g., BERT, ERNIE) to automatically identify and extract entities in the inscriptions, such as person names (e.g., “Fu Hao,” “Wang Cheng”), place names (e.g., “Bo,” “Hu”), state names (e.g., “Tufang,” “Guifang”), official titles (e.g., “Yin,” “Xiao Chen”), and sacrifice names (e.g., “Liao,” “Shan,” “Yi”). On this basis, relationships between entities (e.g., “campaign against,” “sacrifice to,” “hunted at”) can be further extracted, thereby rapidly and automatically constructing the social network, political geographical map, and religious ritual system of the Shang Dynasty [10, 11].

**Topic Models and Automatic Content Classification:** Applying probabilistic topic models (e.g., LDA) or neural topic models for unsupervised clustering analysis of massive inscription texts, automatically discovering latent, stable thematic structures. For example, the model might automatically identify core themes such as “Praying for Harvest and Observing Celestial Phenomena,” “Sacrificing to Former Lords/Kings and Nature Deities,” “External Campaigns and Relations with Border States,” “Royal Hunts, Fishing, and Daily Life,” “Diseases, Childbirth, and Dream Interpretation,” and show the representative vocabulary and their weights under each theme. This achieves automated, fine-grained, quantifiable macro-classification of inscription content, helping to grasp the focus and structure of Shang state affairs as a whole [10].

**Semantic Association Analysis and Word Vector Representation:** Using Word2Vec or context-dependent word vector techniques based on Transformers to represent the vocabulary (characters or words) in OBS as dense vectors in a high-dimensional space. In this semantic space, words that are semantically similar or have similar grammatical functions (e.g., “cow,” “sheep,” “pig” all as sacrifices; “east,” “south,” “west,” “north” all as directions) will have very close vector distances or angles. Calculating the cosine similarity between vectors can reveal potential, unnoticed semantic associations between words, providing a new, data-driven perspective for understanding the language system and conceptual network of the Shang Dynasty [12].

### **2.4. Knowledge Layer and Application Layer: Multi-dimensional Relationship Integration, Intelligent Reasoning, and Research Paradigm Innovation Based on Knowledge Graphs**

The knowledge graph, as a semantic network, can integrate scattered, multi-source, heterogeneous research results into a globally machine-understandable, machine-reasonable knowledge base. It represents the highest level of AI empowerment in OBS research, directly driving research paradigm innovation.

**Large-Scale Domain Knowledge Graph Construction:** Using OBS characters, persons, places, events, time, documents, archaeological remains, etc., as entities, and rich semantic relationships (e.g., “is a variant of,” “participated in event,” “located in area,” “appears in

inscription," "has the meaning of," "belongs to the reign of") as edges, to build an interconnected OBS domain knowledge graph. This is essentially a digital, structural reshaping of the entire knowledge system of Shang civilization [5].

**Intelligent Semantic Query and Complex Association Discovery:** Researchers can perform highly complex graph queries, for example: "Query all inscriptions related to the 'Huan River', involving 'sacrificial' activities, and where 'Qiang' captives appear as sacrifices, meanwhile list the approximate time (reign) of these events and other related places." The knowledge graph can instantly present results and visually display the complex association paths between all entities in a graph format, revealing hidden historical connections.

**Relationship Reasoning and Automatic Hypothesis Generation:** Based on existing entities and relationships in the knowledge graph, graph reasoning algorithms (e.g., path ranking, graph neural networks) can be used for logical reasoning to discover implicit knowledge. For example, if the graph shows that "Location A" frequently co-occurs with "Clan B" in "campaign" events, and the settlement area of "Clan B" is known, the system can infer that "Location A" might be located within the sphere of influence or along the campaign route of "Clan B," thereby automatically generating a verifiable spatial-historical hypothesis, providing researchers with new directions for exploration.

**Case Analysis:** In the constructed knowledge graph, "Shang Jia" is connected via "sacrificed to" relationships with "Bao Yi," "Bao Bing," "Bao Ding," forming a lineage of pre-dynastic lords; while "Da Yi" (Cheng Tang) is connected via "campaign against" relationship with "Renfang"; simultaneously, the geographical entity "Muye" is connected via "occurred at" relationship to the "King Wu's Conquest of King Zhou" event (though early Zhou, it serves as background). When a researcher queries the "geographical patterns of military activities in the early Shang," the graph can, through spatiotemporal path analysis, visually show the geographical focus shift of campaign activities across different reigns and suggest potential correlations with resource control and the rise and fall of border states, thereby integrating fragmented historical information into a dynamic, interpretable model of historical process.

### **3. Construction and Practice of an AI-Empowered Intelligent OBS Teaching Model**

AI technology can fundamentally change the mode of inheriting OBS knowledge, transforming profound and abstruse learning into vivid, interactive, personalized, and inquiry-based learning experiences, effectively stimulating learners' intrinsic motivation and cultural identity.

#### **3.1. Theoretical Foundation of Teaching: Constructivism, Embodied Cognition, and Situated Learning**

The design of this model is deeply rooted in modern learning science theories:

**Constructivism:** Emphasizes that learners are active constructors of knowledge, not passive receivers. Teaching should create environments conducive to learners' active exploration and meaning construction [13].

**Embodied Cognition:** Emphasizes that cognitive processes depend on the interaction between the body and the environment. Allowing learners to "personally" experience the Shang Dynasty context through immersive technology can deepen understanding and memory [14].

**Situated Learning:** Emphasizes that knowledge is generated and applied in specific contexts. Learning OBS within its authentic historical and cultural background facilitates knowledge transfer and application [15].

### **3.2. Intelligent Teaching Model: The “Perception-Understanding-Inquiry-Creation” Four-Stage Spiral Model**

We construct a learner-centered, four-stage spiraling intelligent teaching model:

**Perception Stage:** Using cutting-edge immersive technology to create realistic historical and cultural contexts of the Shang Dynasty, fully engaging learners' senses, establishing initial impressions, and stimulating strong interest and curiosity.

**Understanding Stage:** Through dynamic visualization, adaptive feedback, and interactive explanations, deeply analyze the structural principles of OBS, the origin and evolution of meanings, and the underlying social and cultural connotations, promoting conceptual understanding.

**Inquiry Stage:** Providing professional intelligent research tools and data platforms, guiding learners to ask questions, gather evidence, analyze materials, and form viewpoints like scholars do, cultivating critical thinking and academic inquiry skills.

**Creation Stage:** Encouraging learners to engage in personalized, artistic expression, storytelling, and creation based on the knowledge acquired, assisted by AIGC and other creative technologies, achieving deep internalization of knowledge and cultural innovation.

### **3.3. Core Technical Application Scenarios and Practical Cases**

#### **3.3.1. Perception Stage: Immersive Glyph-Origin Context Construction Based on AR/VR/MR**

**Technical Realization:** Constructing complete virtual environments through VR headsets, or superimposing virtual information onto the real world via AR glasses/MR devices. For example, reconstructing the palace-temple area of Yinxu in VR, where learners can stroll around, watch the process of diviners heating turtle shells and inscribing characters in a “sacrificial” scene, while the corresponding OBS character “祀” (sacrifice) appears suspended and dynamically evolves into Bronze Script, Small Seal Script, and finally Regular Script.

**Teaching Value:** Achieves the deepest level of embodied cognition and situated learning, deeply integrating abstract characters with vivid historical scenes, allowing learners to “travel” through time, intuitively understanding the creative origin and practical context of the script, and deeply feeling the wisdom of Chinese character formation and the atmosphere of Shang civilization.

#### **3.3.2. Understanding Stage: Adaptive Learning Based on Learner Profiles and Dynamic Glyph Evolution System**

Technical Realization will be achieved with the following elements.

**Adaptive Learning System:** Constructing dynamic learner profiles (including knowledge state, cognitive style, interest preferences) through pre-tests and analysis of learning behavior data (e.g., clickstream, dwell time, error patterns). AI algorithms are adopted to design unique learning paths for each learner, dynamically recommending the most suitable learning resources (e.g., providing more images and animations for visual learners, more character structure analysis diagrams for logical learners) and practice questions [16].

**Interactive Dynamic Evolution Animation:** Developing delicate interactive animations that not only show the linear evolution of an OBS character from graphic to symbol but also allow learners to click on different components to understand their meanings and roles in the structure (e.g., clicking the “戈” and “止” in “武” to explain their meanings respectively), thereby deeply understanding the “Six Writings” character formation principles.

**Teaching Value:** Truly achieves large-scale personalized teaching, improving learning efficiency and satisfaction; transforms the history of Chinese character evolution into an interactive,

explorable cognitive process, helping learners establish a systematic and profound view of script development.

### **3.3.3. Inquiry Stage: Intelligent Assisted Inquiry Platform and “Digital Diviner” Dialogue System Based on Large Language Models**

Technical Realization will be made up of the following elements.

**Intelligent Assisted Inquiry Platform:** Building an online inquiry-based learning platform integrating OBS digital libraries, knowledge graphs, GIS maps, and data analysis tools (e.g., clustering, topic models). Learners can autonomously pose research questions (e.g., “What was the climate like in the Shang Dynasty?” “What was the status of women in Shang society?”), then use the platform's tools to search relevant inscriptions, analyze word frequency distribution, visualize geographical information, and form their own small research reports.

**“Digital Diviner” Intelligent Dialogue Agent:** Building a knowledgeable, linguistically stylized (fitting the Shang context, or switchable to modern scholar mode) AI dialogue system based on fine-tuned large language models. Learners can engage in open-domain, heuristic Q&A, for example: “Why does the character '王' (king) resemble a battle-axe? What concept of power does this reflect?” “If I want to learn about Shang agriculture, which OBS characters should I focus on?” The AI can generate responses that are not only accurate but also guide learners to think further.

**Teaching Value:** Cultivates learners' information literacy, critical thinking, and active inquiry ability, transforming learning from passive reception to active knowledge discovery and construction, personally experiencing the logic and joy of academic research.

### **3.3.4. Creation Stage: Interactive Storytelling, Gamified Creation, and Cultural Innovation Supported by AIGC**

Technical Realization is composed of the following parts.

**AIGC-Assisted Creation:** Learners input one or a set of OBS characters (e.g., “舟” boat, “河” river, “渔” fish), and the AI, based on its understanding of Shang history and culture, can automatically generate a reasonable micro-story, a poem full of ancient style, an illustration of a scene consistent with archaeological findings, or a short animation script.

**Gamified Learning Mechanisms:** Designing serious games of genres like role-playing, puzzle, strategy. For example, in the game “Secret Codes of the Shang,” learners play the role of a “diviner” who needs to correctly interpret OBS to receive “divine messages” (quests), decide the tribe's actions by interpreting inscriptions, thereby influencing the game's plot development. In-game achievement systems, leaderboards, and instant feedback mechanisms continuously motivate learning engagement.

**Teaching Value:** Greatly stimulates learners' imagination, creativity, and cultural innovation awareness; gamification mechanisms effectively enhance learning motivation, engagement, and persistence, making OBS learning a pleasant, fulfilling cultural experience.

## **4. Challenges, Reflections, and Future Pathways: Towards Deep Empowerment through Human-Computer Collaboration**

While enthusiastically embracing the enormous potential of AI technology, we must maintain a clear critical mind, deeply recognizing its current limitations and potential ethical risks, to ensure that technological development proceeds on a healthy, sustainable, and humanistically aligned track.

### **4.1. Core Challenges at the Technical and Data Level**

**Data Quality and Annotation Bottleneck:** AI models are “data-hungry.” The scarcity of high-quality OBS data, the high dependence of annotation work on expert knowledge, and the

inevitable subjectivity in the annotation process are the primary obstacles constraining model performance improvement. Building large-scale, high-quality, consensual annotated datasets requires sustained, interdisciplinary collaborative investment.

**Algorithmic Bias and Domain Adaptability:** The algorithms used to train AI models may themselves contain designers' biases. More importantly, general-domain AI models (e.g., pre-trained language models), if not sufficiently domain-adapted, are highly prone to producing outputs with "anachronisms" or "cultural misinterpretations" when understanding highly specialized, context-specific ancient texts like OBS.

**Model "Black Box" and Decision Interpretability:** The decision-making process of deep learning models is often an opaque "black box." When a CNN model classifies an undeciphered character, or a topic model generates a new topic, scholars urgently need to understand the rationale behind the decision (e.g., which key stroke features led to the classification? Which co-occurring words define the topic?). Lack of interpretability severely hinders the acceptance and credibility of research results within the academic community.

## **4.2. Deep Reflections at the Humanistic and Ethical Level**

**Academic Authority and the Changing Mode of Knowledge Production:** When AI systems can automatically discover glyph associations, generate research hypotheses, or even write preliminary analysis reports, how should the priority, originality, and interpretative authority of academic discoveries be defined? This challenges the traditional human-scholar-centric mode of knowledge production.

**Risk of Weakening Humanistic Spirit and Intuitive Thinking:** Over-reliance on data-driven approaches and algorithmic recommendations may lead the new generation of researchers to neglect the basic training of "intensive reading and refined thinking" and "immersive contemplation" of original materials, weakening the "insight" and "intuition" cultivated through long-term academic practice, which are based on holistic grasp and inspiration.

**Concerns about Cultural Diversity and Homogenization of Historical Narrative:** The quality of AIGC-generated content is highly dependent on the distribution of training data. If the training data primarily stems from one mainstream academic viewpoint or historical compilation, the AI may inadvertently reinforce a single historical narrative while marginalizing other valuable, non-mainstream scholarly interpretations, causing a "digital divide" in cultural expression and a narrowing of historical understanding.

## **4.3. Future Pathway: A New Paradigm of "Human-Computer Collaboration, Rationale Integration, and Ethics First"**

Facing the challenges, the future direction is certainly not AI replacing humans, but constructing a new paradigm of deep human-computer collaboration where strengths complement each other. Its core principle is "Rationale Integration" – the deep integration of the algorithm's "data-driven" approach with the scholar's "theory-driven" and "intuitive insight" approaches.

**Vigorously Develop Explainable AI (XAI):** Actively research and apply XAI technologies in the OBS field, for instance, visualizing CNN attention heatmaps to highlight the glyph regions most critical for classification decisions; providing interactive explanation interfaces for document-topic-word distributions in topic models. Allow scholars to "see through" the algorithm's operation mechanism, enabling effective dialogue and bidirectional calibration between humans and computers.

**Build a "Human-in-the-Loop" Collaborative Research Workflow:** Establish AI's role as a "super research assistant." Form a closed loop of "AI preliminary processing and intelligent recommendation → Scholar professional screening, deep interpretation, and value judgment → Scholar injects feedback and domain knowledge into the system → AI model continuous

learning and iterative optimization.” In this workflow, human wisdom always controls the research direction and holds the final decision-making power.

Adhere to Ethics First and Humanistic Leadership: Embed ethical considerations throughout the entire process of technology R&D and application. Clearly define the auxiliary positioning of technical tools, always emphasizing that humanistic questions are the starting point and ultimate goal of research. All intelligent systems should serve the fundamental objectives of deepening the understanding of the origins and characteristics of Chinese civilization and promoting the permanent preservation and innovative inheritance of OBS. Strengthen interdisciplinary dialogue and cultivate new talents possessing both digital literacy and profound humanistic heritage.

## 5. Summary

The systematic empowerment by artificial intelligence is infusing unprecedented, powerful digital vitality into the ancient learning of Oracle Bone Script, which carries the code of a three-thousand-year-old civilization. At the research level, through intelligent recognition, semantic mining, knowledge graphs, and intelligent reasoning, it is pushing OBS classification research from traditional empirical induction towards a new computational humanities paradigm that is data-driven, hypothesis-led, and combines macro and micro perspectives, giving birth to the interdisciplinary field of “Computational Oracle Bone Studies.” At the teaching and inheritance level, through creating immersive contexts, planning personalized paths, empowering deep inquiry, and stimulating innovative creation, it has constructed an intelligent teaching model centered on learner cognition and development, completely transforming OBS education from static, one-way knowledge transmission into a dynamic, interactive, immersive process of cultural experience and meaning construction.

However, we must always remember that the dazzling light of technology should never, and cannot, overshadow the core of humanistic spirit. The most desirable future vision is not that of relentless machine intelligence alone solving all the millennium-old puzzles, but rather a continuous and profound creative dialogue among the field insight of archaeologists, the textual erudition of paleographers, the grand narratives of historians, and the ingenious algorithms of computer scientists. In this dialogue, AI is a powerful catalyst, multiplier, and source of inspiration, while the ultimate light of wisdom, the profound interpretation of civilizational values, and the innovative continuation of cultural heritage, their source and destination, always lie in humanity's persistent, warm, and respectful quest for and guardianship of the source of its own civilization. We should unswervingly uphold the concept of “Human-Computer Collaboration, Rationale Integration, and Ethics First,” prudently and actively promoting this historic process, allowing the oracle bones slumbering underground to radiate a more dazzling light of civilization in the new era of digital intelligence.

## Acknowledgements

The authors gratefully acknowledge the financial support from the Key Project of Scientific Research in Universities of Anhui Province (Grant No. 2022AH050564), and the University-level Projects of Anhui University of Finance & Economics (Grant No. ACKYC23075).

## References

- [1] Qiu, Xigui. Chinese Writing. Beijing: The Commercial Press, 1988.
- [2] Guo, Moruo. Collection of Oracle Bone Inscriptions [M]. Beijing: Zhonghua Book Company, 1978-1982.
- [3] Zhou, Zhihua. Machine Learning. Beijing: Tsinghua University Press, 2016.

- [4] Liu, Yongge, et al. A Survey of Oracle Bone Character Recognition Based on Deep Learning. *Acta Automatica Sinica*, 2021, 47(5): 1009-1022.
- [5] Li, Yizhan, et al. Research on the Construction and Application of the Yinxu Oracle Bone Inscriptions Knowledge Graph. *Data Analysis and Knowledge Discovery*, 2022, 6(5): 1-15.
- [6] Feng, Xiang. Ancient Script Research in the AI Era: Opportunities and Challenges. *Documents*, 2023, (2): 45-56.
- [7] Wang, J., & Li, S. Unsupervised Oracle Bone Script Character Discovery via Deep Convolutional Autoencoder. *Proceedings of the ACM on Multimedia*, 2020.
- [8] Dong, Zuobin. Examples of Periodization Research in Oracle Bone Inscriptions. // *Bulletin of the Institute of History and Philology, Academia Sinica*. 1933.
- [9] Zhao, Ping'an. *Theory and Practice of Computer Recognition of Oracle Bone Script*. Shanghai: Shanghai Ancient Books Publishing House, 2021.
- [10] Li, Bin, et al. Content Analysis of Oracle Bone Inscriptions Based on LDA Topic Model. *Journal of Chinese Information Processing*, 2019, 33(8): 76-83.
- [11] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of NAACL-HLT*, 2019.
- [12] Pennington, J., Socher, R., & Manning, C. D. GloVe: Global Vectors for Word Representation. *Proceedings of EMNLP*, 2014.
- [13] Piaget, J. *The Principles of Genetic Epistemology*. Translated by Wang Xiantian, et al. Beijing: The Commercial Press, 1981.
- [14] Lakoff, G., & Johnson, M. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. Basic Books, 1999.
- [15] Lave, J., & Wenger, E. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, 1991.
- [16] Li, Mang. *Design of Technology-Supported Learning Environments*. Beijing: Beijing Normal University Press, 2018.