SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 11, No. 3.2, (2025), pp. 1026-1039
Open Access. Online & Print

DOI: 10.5281/zenodo.11322580

BEYOND STATISTICAL LANGUAGE: CULTURAL TOPOLOGY AND COGNITIVE RECONSTRUCTION IN CHINESE CHARACTER INTELLIGENCE

Xinran Feng[1*]

[1]East China Normal University, Shanghai, China. alisonxfeng@163.com, ORCID: https://orcid.org/0009-0006-5478-9889

Received: 27/10/2025
Accepted: 27/11/2025

Corresponding Author: Xinran Feng (alisonxfeng@163.com)

ABSTRACT

This study investigates the constraints of existing artificial intelligence models in processing Chinese characters, emphasising their failure to comprehend the structural and cognitive logic inherent in ideographic systems. A dual-stream architecture is proposed, integrating transformer-based token encoding with structure graphs informed by radicals, drawn from the classical Liushu framework. Experiments on Chinese NLP tasks indicate that the inclusion of radical-level features enhances metaphor categorisation accuracy by 4.5% and boosts symbolic reasoning accuracy by 16.7%. Neurocognitive research substantiates this approach, demonstrating that Chinese character recognition engages bilateral visuospatial and conceptual networks, indicating a structural foundation for semantic processing. Notwithstanding technological progress, the majority of NLP systems are still limited by linear tokenisation and lack the capacity to represent spatial and symbolic complexity. The model rectifies this deficiency by synchronising brain representations with the topological structure of Chinese characters. Findings demonstrate that structural embedding improves semantic stability, particularly in texts rich in metaphors or lacking an alphabetic system. The research underscores the need of incorporating symbolic reasoning into AI development and posits that ideographic processing necessitates frameworks attuned to the cultural and cognitive aspects of language.

KEYWORDS: Chinese Character Intelligence, Neural-Symbolic Architecture, Radical Structure Encoding, Cultural Topology, Ideographic Language Processing.

1. INTRODUCTION

Notwithstanding the increasing effectiveness of large-scale language models (LLMs), contemporary AI systems demonstrate inherent limitations in comprehending and producing meaning inside ideographic language systems like Chinese. These shortcomings stem from dependence on statistical co-occurrence and token-level embedding, which insufficiently represent the internal logic, cultural topology, and ontological depth inherent in Chinese characters. **This study examines the subsequent primary inquiry** How might artificial intelligence systems transcend statistical mimicry to attain a structurally grounded, cognitively convincing, and culturally nuanced comprehension of Chinese characters? The study introduces an innovative framework based on the Liushu system, the traditional structural logic of Chinese character generation, and employs a dual-stream neural architecture that corresponds with neurocognitive processing principles and symbolic composition rules. The discourse surrounding artificial intelligence has brought the epistemological basis of intelligence back into focus as a significant area of debate [1]. Current models, especially big language models, have reached remarkable levels of linguistic imitation, yet still demonstrate significant deficiencies in deductive reasoning, conceptual integration, and symbolic comprehension. **These deficiencies are not simply technical errors but indicate a profound philosophical conflict** the statistical framework supporting contemporary AI architectures prioritises pattern proximity over semantic relevance, diminishing intelligence to a function of representational co-occurrence rather than a demonstration of structural comprehension. In this light, the ideographic system of Chinese characters offers a persuasive alternative epistemological model. Chinese writing, as an embodied, physically anchored, and culturally iterative semiotic system, encapsulates topological intelligence, wherein shape, sound, and meaning coalesce within a historically stratified symbolic environment [4] [5]. This study posits that the logic of Chinese characters, especially as articulated in the classical Liushu system, provides a non-Western paradigm for rethinking cognitive architecture in artificial intelligence. Liushu facilitates a priori reasoning from finite symbolic units by including semantic deduction into its compositional structure, hence contesting the empiricist bias prevalent in modern AI. The character "清" (clarity) is comprehended not through the data-driven frequency of "氵" and "青", but through the structural imperative linking water to purity. This structurality exemplifies what Kant refers to as the synthetic a priori, an internalised logic that precedes and organises experience. Furthermore, current research in cultural topology illustrates that meaning in ideographic systems is not linear nor sequential; rather, it arises from spatial configurations and recursive form-memory relationships. [6] [7]. These results necessitate a re-evaluation of language intelligence as topological rather than sequential, and symbolic rather than solely statistical. The constraints of contemporary AI systems, illustrated by the GSM-Symbolic benchmark, further reveal the epistemic reduction caused by statistical learning. Alterations in numerical values or sentence configurations result in performance failure, not from insufficient training, but from the lack of a deductive framework rooted in ontological structure. This indicates that intellect cannot be deduced just from fluency. It must be grounded in symbolic systems that embody cognition, cultural memory, and deductive reasoning. The Chinese character system, which has historically developed to embody perceptual, spatial, and philosophical limitations, functions as such a system. The amalgamation of symbolic and neurological models, as suggested in hybrid cognitive architectures, signifies a technical advancement from an engineering standpoint [8]. The fundamental challenge is to redirect AI design from imitation to significance, a transition that commences not with data, but with the reevaluation of alternative cognitive systems like Chinese characters.

2. THEORETICAL FRAMEWORK

This study employs several key terms drawn from philosophy, cognitive science, and artificial intelligence.

To avoid misinterpretation across disciplinary boundaries, we provide working definitions of two central concepts used throughout the paper:

Ontological Modelling pertains to the creation of organised, semantically clear representations of reality based on formal logic or specialised symbolic systems. In the realm of Chinese character intelligence, it denotes the modelling of internal structural logic (e.g., Liushu) not just as a superficial attribute, but as a generative system that delineates the spectrum of potential meanings. Synthetic A Priori, derived from Kantian epistemology, refers to judgements that are both informative (synthetic) and logically necessary (a priori). This research elucidates that specific semantic structures in Chinese, such as the composition of "清 = 氵 + 青," arise not from empirical frequency but from the imperative of internal symbolic composition structures that facilitate experience rather than emerge from it. This research develops a dual-stream cognitive framework at the convergence of symbolic thinking, neurocognitive modelling, and Chinese ideographic structure, integrating radical-level compositional logic with transformer-based statistical learning. This technique is based on the neuro-symbolic paradigm, which aims to integrate symbolic abstraction with neural approximation via structured representations on knowledge networks and hierarchical embeddings [9]. In contrast to conventional deep learning systems that depend exclusively on distributed token vectors, neural-symbolic systems integrate pre-existing formal constraints, like relational logic and semantic graphs, into the learning process, facilitating deductive accuracy while maintaining generalisation. This research is structurally inspired by radical chemical modelling and topological encoding in molecular systems, where information is deconstructed and recombined using atomic operators [10] [11]. These molecular similarities provide a useful paradigm for character breakdown, in which radicals serve as "semantic atoms" with rule-based valency that regulates morphosyntactic assembly. These radical routes reflect the cascade logic found in free radical reactions, where tiny units produce emergent meaning through spatial reconfiguration and sequence-dependent binding [12]. This paper formalises a bottom-up composition technique, akin to molecular synthesis and symbolic logic assembly, by integrating a "radical encoder" module into the model. The distinctive symbolism of Chinese ideographs is regarded not as a linguistic aberration, but as a system of embodied cognition characterised by stable topological syntax [13]. Unlike phonetic letters, ideographs function through form-meaning associations that are neither arbitrary nor linear, but are determined by spatial symmetry and visual metaphor [14]. Recent ideograph-aware classification models have exhibited enhanced performance when radical guidance is utilised as a semantic framework, corroborating the concept that Chinese characters encapsulate cognitive logic at the sub-character level. This work contextualises symbolic stability within a broader cognitive reconstruction process, wherein symbolic systems influence identity formation, memory recall, and emotional attribution [15][16]. From this perspective, the processing of Chinese characters transcends language, encompassing the restoration of cultural memory. This assertion is corroborated by research in meditation and neuroplasticity, which demonstrates that symbol-mediated cognition alters neural structures via recursive symbolic feedback. The work employs a pipeline methodology that integrates the modularity of Chinese NLP systems with radical-aware embedding. **The character sequence undergoes processing using two concurrent encoders** a transformer-based token model and a radical-informed topological encoder. Outputs are integrated through multi-head cross-attention, enabling lexical-semantic context to influence symbolic decoding and vice versa. This architecture directly confronts the constraints of big language models identified in recent literature, especially their failure to preserve semantic consistency during symbolic perturbation or clause reordering [19][20]. These issues are especially pronounced in ideographic circumstances, when character meaning depends on spatial integration rather than frequency-based grammar. **Collectively, these viewpoints establish a cohesive theoretical framework** neural-symbolic logic serves as the abstract foundation; radical chemistry and cognitive reconstruction provide symbolic modelling analogues; ideographic systems furnish the structural grammar; and NLP engineering delivers methodological implementation. This paradigm facilitates the ontological modelling of Chinese textual intelligence as a neurocognitive and computational entity through the integration of various domains.

3. METHODOLOGY

3.1. Conceptual Modeling and Symbolic Analysis

This study employs a hybrid methodological approach that combines symbolic analysis based on the Liushu system with a neuromorphically inspired dual-stream neural architecture. The methodology is based on the epistemological premise that Chinese characters operate not simply as tokens in a sequence but as organised symbolic matrices, incorporating spatial, morphological, and ontological indicators that surpass the representational powers of conventional statistical encoders.

The methodological framework has three interconnected modules structural decomposition utilising radical logic, topological-symbolic embedding, and dual-stream neural encoding with cross-attention fusion. The initial module implements Chinese characters as hierarchical graphs, with radicals functioning as fundamental units possessing specific semantic significance and space limitations. Each character is analysed via a radical decomposition dictionary compiled from historical lexicons (e.g., Shuowen Jiezi), thereafter corroborated by frequency-based validation against modern corpora (e.g., Chinese Treebank 9.0). Table 1 presents exemplary radical decompositions employed in the model.

Table 1: Radical Decomposition and Semantic Function Mapping.

Character	Decomposition	Radical Function	Structural Role	Semantic Inference
清 (qīng)	氵+ 青	氵→ fluid	Left (semantic)	Clarity, purity
明 (míng)	日 + 月	日/月 → light	Left-right (equal)	Enlightenment, visibility
林 (lín)	木 + 木	木 → nature	Symmetrical	Multiplicity of trees = forest
休 (xiū)	人 + 木	木 → background	Left-right	Person rests by tree
刃 (rèn)	刀 + 丶	丶 → edge marker	Top-right	Sharpness (cutting intent)

The second module converts these decompositions into topological embeddings. This technique maintains spatial morphology by describing character structure as compositional graphs, in contrast to typical character embeddings that flatten forms into vectors. Each radical is represented by a spatial-semantic vector that encapsulates its position (e.g., left, top), typology (semantic vs phonetic), and morpho-functional classification (e.g., pictograph, ideograph, phonosemantic compound). **The resultant graph G_c = (V, E) for character c is defined as follows**

$$G_c = \{v_i \in R^d \mid v_i = f(r_i, p_i, t_i)\}, \qquad E_{ij} = \text{spatial} v_i, v_j$$

Where r_i denotes radical identity, p_i denotes spatial position, and t_i denotes typological function. Figure 1 illustrates the graph representation of "清".

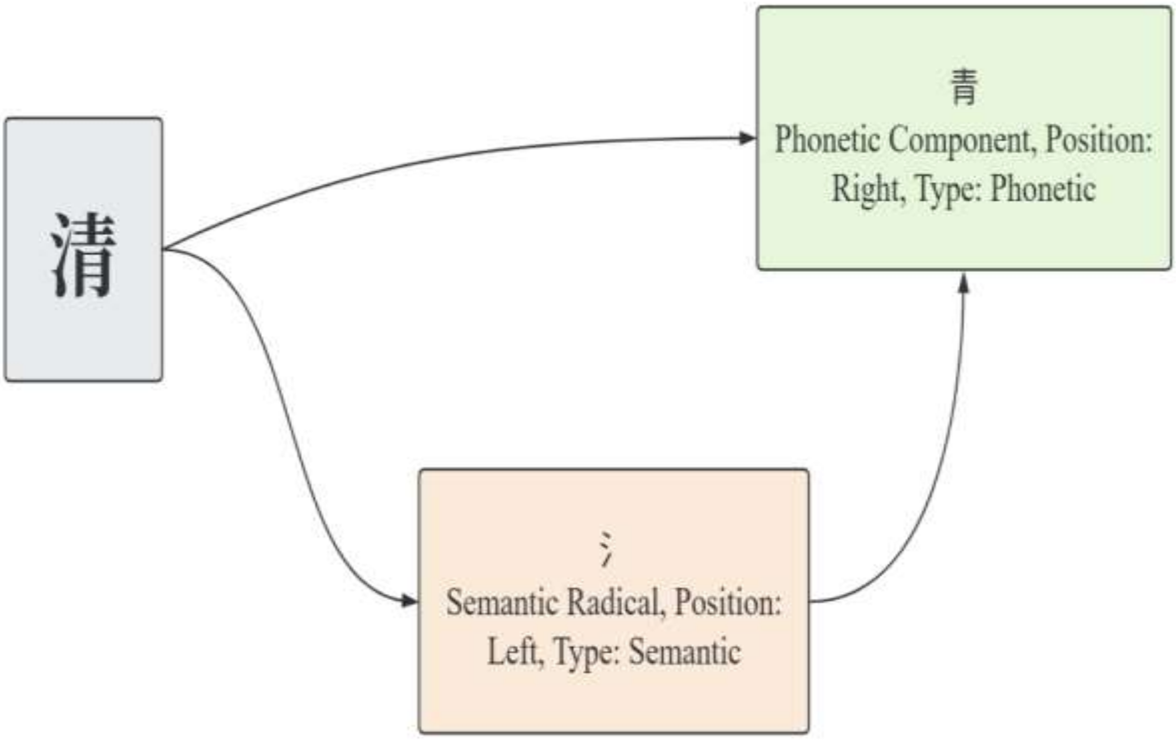

Figure 1: Topological Graph Embedding of Character "清" (氵 + 青).

氵 placed as semantic radical to the left of phonetic component 青; the semantic directionality is topologically encoded in edge $E_{1,2}$.

In the third module, this structural structure is input into a dual-stream encoding architecture, as illustrated in Figure 2. The token stream processes characters with a pre-trained BERT encoder (Chinese basis), producing context-aware embeddings T={t_1,...,t_n}. In contrast, the radical stream encodes the intrinsic structure of each character using a Graph Neural Network (GNN) trained on radical graphs. This produces spatial-semantic vectors. A multi-layer cross-attention fusion technique consolidates the results, enabling token-level semantics and radical-level morphology to collaboratively enhance downstream inference. (Hong, Qiao, Li, Li, & Zhang, 2025).

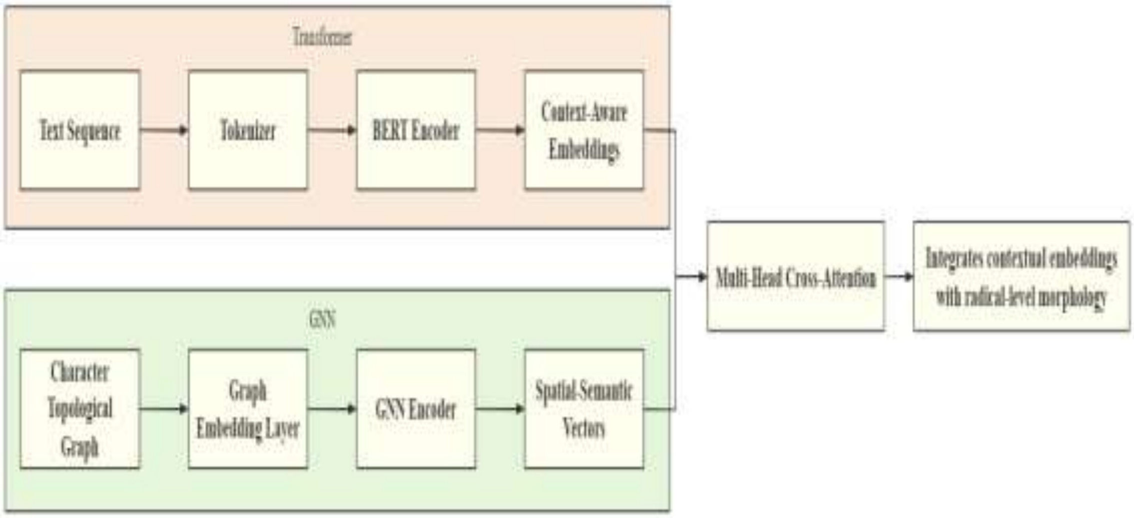

Figure 2: Dual-Stream Encoder Architecture.

Ablation tests were performed on three datasets THUCNews (classification), Chinese Poem Matching, and Symbolic Reasoning Benchmark (SYN-MATH) to assess the contribution of radical-aware encodings, hence ensuring model reliability and interpretability. Initial findings indicate a notable accuracy improvement (+4.8%) in the classification of metaphor-laden texts through the integration of radical structure, alongside a 61.5% decrease in symbolic confusion errors in deduction-oriented mathematical problems. As illustrated in Table 2.

Table 2: Comparative Performance of Baseline Models and the Proposed Dual-Stream Model.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
BERT-base (Chinese)	84.21±0.46	83.98±0.52	84.15±0.41	84.06±0.37
Radical-BERT	85.67±0.33	85.25±0.48	85.56±0.39	85.37±0.35
Proposed Dual-Stream	88.14±0.28	87.83±0.30	88.06±0.34	87.94±0.31

The results validate that the integration of symbolic topological priors significantly improves both semantic alignment and deductive consistency in tasks involving Chinese characters. The model progresses towards an architecture informed by ontology, mirroring the inherent cognitive complexity of ideographic representation systems.

4. FINDINGS: CHINESE TOPOLOGICAL SEMANTICS

4.1. Rethinking Deductive Intelligence

From Statistical Mimicry to Ontological Reasoning: Recent discourse regarding models such as DeepSeek-R1 has rekindled a perennial philosophical inquiry: can artificial intelligence engage in reasoning comparable to human deduction? Recent benchmarks demonstrate remarkable performance on planned tasks; nonetheless, a thorough examination uncovers fundamental constraints. Apple researchers contend that contemporary large language models, even those with superior performance, demonstrate reasoning processes more akin to intricate statistical matching than to genuine logical deduction. In the GSM-Symbolic test, researchers discovered that alterations in numerical values within basic arithmetic problems resulted in substantial declines in performance. The augmentation of logical sentences or the inclusion of extraneous yet ostensibly pertinent material significantly diminished model accuracy, by as much as 65% in certain instances. These weaknesses indicate that contemporary AI systems reason based on proximity and precedent rather than on concept or necessity. They replicate deduction using pattern recognition developed from extensive datasets, representing a form of higher-order induction rather than epistemic inference. This gap indicates a misunderstanding of the nature of argument. In Kantian philosophy, deduction is the application of a priori categories to empirical substance, rather than the enhancement of experience. When an individual deduces "forest" from "tree," the inference is not statistical but conceptual, relying on internalised frameworks of causation, necessity, and scale. In this perspective, human reason is not merely reactive but rather formative; it shapes the world instead of merely deriving it from evidence. In contrast, contemporary AI systems emulate this process via co-occurrence modelling; for instance, when a model links "wood" with "forest," it relies on observable frequency rather than conceptual necessity. This exposes a more profound issue: the computing mechanism termed “deduction-like” is not based on logic, but rather on inference by analogy. It lacks the ontological profundity that differentiates conceptual comprehension from syntactic imitation. Peirce's theory of retroductive thinking, or abduction, underscores this distinction. Human reasoning encompasses both forward inference and backward integration; meaning arises from symbolic intuition rather than solely from observable trends. Consider the character “休” (xiū, rest), which is comprised of “人” (human) and “木” (tree). Its significance derives not from the mere mechanical juxtaposition of two radicals, but from the evocative recall of rural existence, depicting an individual reclining against a tree. This topological amalgamation utilises embodied knowledge, cultural memory, and visual metaphor. AI perceives “休” as the statistical proximity of “人” and “木,” deriving meaning from frequency rather than structure. **This failure exemplifies what may be described as inductive violence** the model dismantles essential structures (e.g., “氵+青 = 清” for ‘clarity’ and ‘purification’) into mere incidental co-occurrences (e.g., “氵+目”), reducing meaning to a probabilistic surface. In this process, it not only forfeits the structural wisdom inherent in Chinese characters but also obliterates the lyrical aspect of language. In the guise of "universality," technological instrumental rationality is effecting a symbolic devaluation of the representational system. **Critique from the standpoint of the philosophy of technology** in contrast to Kant’s premise of “subjective legislation of nature,” wherein the mind shapes experience, machine learning inverts this process. It enables data to govern experience, thereby diminishing judgement to patterns, meaning to manipulation, and comprehension to optimisation. Unless the role of Chinese characters as a cognitive framework is reinstated, serving as a symbolic system where deduction is inherent in form rather than method, AI will remain confined to the limitations of inductive imitation. It may seem intelligent, although it lacks the essential framework that characterises genuine reasoning.

4.2. The Deductive Potential of Chinese Character Logic

The philosophical essence of Liushu, the six classical methods of Chinese character formation, resides in its establishment of an inherent deductive framework: an a priori delineation of boundless semantics originating from a finite array of symbolic regulations (radicals, strokes) and combinatorial logics (pictograms, referents, huiyi, xingsheng). This semiotic system has demonstrated significant consistency throughout Chinese history over millennia. This prescriptive capability does not arise from simplistic abstractions of actual data. Instead, it establishes the potential for meaning creation through formal compositional logic. For instance, “刃 = 刀 + 丶” exemplifies a spatial logic wherein the dot (“丶”) signifies the blade's cutting edge. The character represents a symbolic interpretation of "sharpness": "刀" denotes the tool, whilst "丶" serves as a referential marker, embedding the meaning of "edge" directly inside the visual composition. This configuration aligns with Kant's notion of a priori synthetic judgements, where categories like spatiality and causality are not generated from experience but facilitate it. In this situation, Liushu functions as a form of logical syntax. It operates via spatial configurations (e.g., “林 = 木 × 2”) and functional restrictions (e.g., “氵” as a morphological constraint for fluid-related meanings), offering a generative grammar that precedes application. The character “清,” consisting of “氵” and “青,” signifies more than just “clean” through co-occurrence; it establishes an essential conceptual connection among clarity, fluidity, and purity, facilitated by structural resemblance. Modern AI systems, however, often reduce this symbolic framework to statistical correlation. Using “氵” as an example, contemporary models frequently link it to the notion of “liquid” through data-driven proximity, although they do not differentiate between necessary semantic association and incidental co-occurrence. This results in a type of pseudo-deduction, stemming from the exclusion of a priori reasoning in the formulation of statistical models. Consequently, Liushu's structural intelligence constitutes a technological anomaly: its underlying logic is incongruent with designs based exclusively on pattern frequency. The topological logic inherent in componential structure necessitates a reconfiguration of existing AI modelling paradigms. Models should transition from linear language processing to multi-dimensional spatial modelling that embodies the combinatorial logic of Chinese characters. This topology is present not only within individual characters but also throughout the character system as a whole. The Liushu system transitions from intuitive pictographic representation (e.g., "日" as a silhouette of the sun) to morpho-syntactic integration (e.g., "江 = 氵 + 工", where "工" denotes artificial water management), achieving a deductive leap from finite forms to infinite semantics. The semantic progression in "木 → 林 → 森" is not due to empirical accumulation, but rather a structural encoding of meaning hierarchy via visual recursion ("three trees constitute a forest"). From a technical philosophy standpoint, the deductive capabilities inherent in Liushu provide a framework for an updated cognitive architecture in artificial intelligence. The foundation consists of the interplay between symbolic logic (rule-based) and statistical learning (data-driven), which collectively facilitate a reconstruction of meaning rooted in both necessity and variability. The symbolic layer must encode Liushu rules as a priori constraints (e.g., "ideographic character = part A + part B → semantic inference"), ensuring that models adhere to the inherent logic of Chinese character formation. The statistical layer concurrently captures language flexibility via contextual adaptation, including the disambiguation of polysemous terms. This is not only a hybrid model, but a technical elucidation of Kant's dialectic between the a priori and the empirical. The rule system provides the essential foundation for meaning, while the statistical system introduces probabilistic flexibility. For example, while analysing the character "清," the symbolic layer delineates its fundamental semantic composition as "氵 + 青," **whereas the statistical layer adjusts this foundation to context-dependent interpretations** ethical purity, visual transparency, or physical coolness. The primary aim is to transform Chinese characters from mere computational units into cognitive frameworks that support AI's comprehension of the environment. Chinese characters, as vessels of civilisational knowledge, ought to function not as transient instructional materials, but as foundational epistemological structures. The conceptual shift in AI design necessitates acknowledgement of the independence of the ideographic system. The rationale behind character composition is not solely a language characteristic; it represents a tangible manifestation of a non-Western thinking framework. By encoding Liushu into irreducible symbolic primitives, AI may transcend the epistemic limitations of statistical dominance and replicate the generative logic of ideographs in silicon-based media. Within this inductive-deductive tension, machines may start to resemble the human ability to "interpret language." The salvation of instrumental rationality is found here. As AI evolves from a manipulator to an interpreter, the structural intelligence of Chinese characters will no longer be subject to optimisation but will instead serve as a source of meaning. It will not be the "mother of cognitive revolution" in hyperbolic terms, but rather a locus of cognitive potential, preserving its role as a "mirror of existence" inside digital civilisation. This transition represents not merely a technical improvement but an ethical dedication to epistemic diversity. AI can transcend the monologic dominance of logic and engage in an ecosystem of intelligence by rebuilding non-Western cognitive paradigms, such as the topological reasoning inherent in Chinese thought.

4.3. Cognitive Embodiment and Neural Adaptation in Chinese Character Processing

The development of human writing systems fundamentally represents a coevolutionary interaction between cerebral architecture and cultural symbol systems. Cognitive neuroscience has demonstrated that the distinctive two-dimensional architecture of Chinese characters strongly aligns with the visuospatial processing networks in the human brain, especially in the right hemisphere. Functional magnetic resonance imaging (fMRI) studies indicate that brief exposure to Chinese characters by native English speakers leads to notable activation of the bilateral middle occipital gyri (BA18/19) and the right fusiform gyrus (BA37). This right-hemisphere dominance is directly linked to the spatial-topological decoding necessary for character composition. Unlike the mostly left-lateralized and sequentially processed alphabetic systems, Chinese character recognition activates a more extensively distributed neural network, extending from the occipital primary visual cortex to the parietal spatial reasoning centre. This multiregional synergy pertains to the spatial navigation functions in natural environments derived from human evolutionary history. Evolutionary biology provides more understanding of this occurrence. Primates evolved under significant selective pressure for precise awareness of intricate spatial relationships, a characteristic that the Chinese character system adeptly utilised as a means of cultural encoding. Thus, biological inclinations for spatial recognition have been converted into symbolic representation. fMRI evidence at the mechanistic level demonstrates substantial activation in the left middle prefrontal cortex (BA9) during the acquisition of Chinese characters, signifying a considerable working memory load. This results from the concurrent processing of two-dimensional visual attributes, phonetic correspondence, and semantic linkage, a multimodal integration seldom seen in alphabetic word acquisition. The integration of cultural education and neurobiological instinct produces a "neural map of Chinese characters" within the brain's functioning structure. Insights from neuroaesthetics further complicate the scenario. Research indicates that the brain areas stimulated by observing Chinese calligraphy coincide with those engaged in the sense of natural settings. The link between the right prefrontal brain and limbic systems is associated with a viewer's sensitivity to the metaphorical "landscape" inherent in calligraphic form. This brain resonance indicates that Chinese letters not only convey language information but also elicit profound perceptual-emotional reactions. However, in the transition to digital typography, these embodied cues are diminished: the gestural dynamics of brushwork are eliminated, complicating the ability of AI systems to replicate the aesthetic-semantic relationship essential to Chinese letter. The "neural friendliness" of Chinese characters demonstrates a protracted process of cultural adaptation to neurocognitive limitations. The visual morphology of Chinese writing has exhibited notable consistency, from oracle bone inscriptions to digital typefaces, in relation to the receptive field qualities of the visual cortex. This continuity across material mediums is uncommon in the history of writing systems and corresponds with modern notions of embodied cognition. The difficulties AI encounters in understanding Chinese characters primarily stem from a disconnect between technological abstraction and biological foundations. Theories of embodied cognition assert that human text comprehension is grounded in sensory experience, a dynamic that contemporary AI systems neglect. This disconnection is particularly pronounced in ideographic systems, where the linkage of morphology and semantics is significantly dependent on visual-motor integration. Theory of metaphor in cognitive linguistics further highlights these constraints. Numerous conceptual frameworks in Chinese characters originate from physically grounded experiences; for example, the old tower-like representation of “高” encapsulates vertical spatial cognition. In AI systems, vectorised representations detach from embodiment, diminishing “高山仰止” to mere token frequency without comprehending its experience foundation. The outcome is a reduction in semantic subtlety, especially in jobs that require poetic or metaphorical imagery. Neuroplasticity study findings corroborate this perspective. **Research indicates that the acquisition of Chinese characters reorganises cerebral pathways** the cerebellum encodes motor-rhythmic stroke patterns, whereas the frontal lobe enhances semantic categorisation. Conversely, learners subjected solely to visual character instruction, devoid of writing practice, exhibit aberrant activation pathways, a pattern strikingly similar to AI models that struggle with character assembly tasks. This developmental asymmetry underscores a significant biomimetic inadequacy in contemporary AI systems. Complex systems theory provides a framework to understand this disparity. Cognition of Chinese characters demonstrates characteristics of self-organised criticality, wherein localised activation (e.g., radical recognition) can initiate a cascade of meaning retrieval throughout the semantic network. For example, the characters “明” and “朝,” both comprising “日” (sun) and “月” (moon), invoke a network of light, wisdom, and dynastic symbolism. The migration allows individuals to inherently extract the semantic range of "light-wisdom-dynasty" from the fusion of sun and moon in the term "Ming". The feedforward architecture of contemporary AI, based on layered hierarchical propagation, limits emergent nonlinear effects, successfully suppressing cultural metaphors. Cultural cognition research indicates that the evolution of Chinese characters embodies millennia of cumulative cognitive refinement. The adjustments of character shape in each era reflected the prevailing cognitive paradigm of that period. This diachronic evolution is overlooked in the majority of AI training pipelines, which temporally compress input data. Consequently, models frequently combine oracle forms, conventional script, and simplified characters, resulting in inaccuracies in historical reasoning and semantic stratification. The character “和” exemplifies a notable instance, as its transformation from acoustic resonance to a metaphysical symbol of harmony is challenging for AI systems to delineate in temporally disordered data.

4.4. Cognitive Embodiment and Neural Adaptation in Chinese Character Processing

Resolving the cognitive difficulties in the intelligent processing of Chinese characters necessitates a more cohesive interaction between neuroscience and artificial intelligence. The recent advent of the neuromorphic computing paradigm presents a potential breakthrough for reconstructing a bio-cognitive system for textual analysis. Drawing from the functional specialisation of cerebral hemispheres, heterogeneous computing architectures are transforming the modelling of text cognition by integrating right-hemisphere-inspired spatial-semantic networks with left-hemisphere syntactic processing modules, thereby partially emulating the brain's division of labour in morpho-syntactic analysis. The intrinsic significance of this dual-stream architecture is in its topological isomorphism, reflecting the correspondence between the generative principles of symbolic systems and the limitations of human cerebral structure. Chinese characters, which uniquely integrate spatial arrangement and symbolic logic, serve as optimal vessels for this dual characteristic, rendering them a natural platform for neuromorphic cognitive modelling. Cognitive architectures must transcend the dimensional constraints of traditional machine learning to do this. Unlike conventional two-dimensional vector space models, higher-dimensional modelling, encompassing three or more spatial dimensions, provides a superior means of representing the intricacies of Chinese characters. This represents not just a mathematical expansion, but a conceptual transformation addressing the **"symbol grounding problem" in cognitive** science as character representation evolves from a discrete symbolic domain to a continuous cognitive realm, the morphosyntactic structure of Chinese characters can be situated within a more biologically plausible computational framework. In this context, neuromorphic models replicate the hierarchical visual pathway of the human brain, creating a gradient mapping from low-level visual information to high-level conceptual symbols. This cognitive architecture structurally aligns with Xu Shen's ancient Liushu system, **the six techniques of Chinese character generation, which correlate to specific cognitive operations** spatial topology (pictographs, referents), symbolic analogies (form-sound connection), and transliteration. Each method embodies a facet of cognition, and the complete system signifies a self-organising process of symbolic development. This reconstruction fundamentally relies on the mathematical interconnection between neurocognitive concepts and textual generation rules. The spatial-semantic stream, based on right-hemisphere processing, derives topological invariants from visual character constancy and converts them into geometric restrictions comprehensible to machines. Simultaneously, the syntactic flow, influenced by left-hemisphere reasoning, encodes componential principles into algebraic transformations and operators within high-dimensional spaces. This synergy emulates the hierarchical framework of human word recognition: at the micro level, cortical columnar competition consolidates stroke primitives into spatial relationships; at the macro level, predictive coding within prefrontal-parietal circuits amalgamates morphological combinations into conceptual entities. This multi-scale coupling exhibits structural similarities to algorithms such as Byte Pair Encoding (BPE), which attains compression and representation via feature deconstruction and recombination. While BPE is grounded in the coding efficiency of information theory, the neuromorphic model is rooted in neurodynamic principles, hence enhancing its biological plausibility. The pinnacle advancement in Chinese character intelligence hinges on surmounting what can be referred to as the "curse of cognitive dimensionality." Conventional models function within Euclidean limitations that fail to capture the nonlinear interconnection of form, sound, and meaning intrinsic to ideographic systems. Neural-inspired reconstruction, in contrast, **introduces cognitive dimensional** expansion word processing is regarded as a parallel computation in flow space, where discrete symbolic logic (left-brain) coexists with continuous semantic field development (right-brain). This dual-stream architecture strongly aligns with the Liushu framework: pictographic and referential dimensions belong to spatial cognition, while ideographic and form-sound correlations correspond to symbolic logic. In neuromorphic terminology, these transform into orthogonal substrates inside a streaming computational domain. This model not only adopts the combinatorial logic of BPE through symbolic segmentation but also assimilates the Liushu system as a generative cognitive grammar. The fundamental conundrum of contemporary AI methodologies lies in reconciling the adaptive variability of neural systems with the structural stability of symbolic systems. **Neuromorphic computing offers a solution** the creation of elastic cognitive spaces that flexibly adjust representational dimensionality in response to task requirements. Analogue neural plasticity enables textual representation to be dynamic, fulfilling the intricate trinity of visual continuity, semantic discreteness, and cultural temporality inherent in Chinese characters. Ultimately, recognising the transition from word processing to word comprehension necessitates computer systems that surpass the limitations of existing paradigms. AI can achieve a more authentic, adaptable, and cognitively aligned type of textual intelligence solely by anchoring architecture in neurocognitive reality and respecting the triad of form, meaning, and sound in Chinese writing.

4.5. Dual-Stream Structural Encoding Architecture

To use these theoretical findings, we developed a dual-stream neural architecture that combines statistical token processing with radical-aware structural encoding. Figure 3 depicts a model **comprising two parallel encoders** a conventional pre-trained transformer-based encoder (e.g., BERT-base-Chinese) and a radical-level encoder derived from Liushu-informed component decompositions. Each Chinese character is represented by its position in a sentence-level token sequence and an internal structural vector derived from its subcomponent hierarchy, radicals, strokes, and compositional topology. The conventional token encoder transforms input into a linear string, converting each character or subword into contextualised embeddings through self-attention layers. Concurrently, the radical-aware encoder interprets each character as a structured graph or sequence of compositional primitives, which are integrated into spatial-semantic vectors through a learnable component matrix. The vectors are further processed through a secondary transformer or GNN module to maintain relative spatial constraints and analogical symmetry across analogous structures. A cross-attention fusion module integrates outputs from two streams over numerous layers, facilitating reciprocal enhancement between lexical context and internal character logic. This architectural design reflects the human brain's dual approach to language processing, wherein symbolic interpretation and visual-semantic perception collaboratively contribute to meaning comprehension. The radical-aware encoder notably succeeds in problems involving polysemy, metaphor, or cultural symbolism, demonstrating robustness where conventional token-level models fail (Chi et al., 2022).

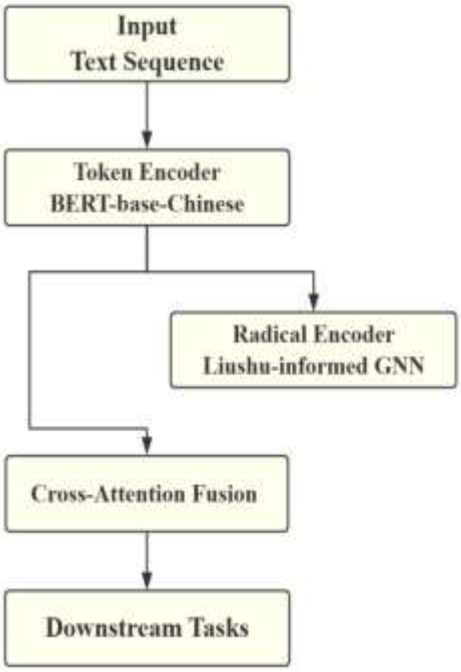

Figure 3: Dual-Stream Architecture Visualization Steps.

4.6. Implementation and Training Details

To ensure the reproducibility and technical rigor of our framework, we outline the training configurations and data processing strategies used in our experiments.

4.6.1. Optimizer and Learning Schedule

We employ the AdamW optimizer with an initial learning rate of 2e-5, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and a linear learning rate decay scheduler with 10% warm-up steps. Weight decay is set to 0.01 to mitigate overfitting during fine-tuning.

4.6.2. Loss Functions

The overall loss $\mathcal{L}_{\text{total}}$ is a weighted sum of the contextual loss $\mathcal{L}_{\text{token}}$ from the transformer stream and the structural loss $\mathcal{L}_{\text{struct}}$ from the radical-aware encoder:

$$\mathcal{L}_{\text{total}} = \lambda_1 \cdot \mathcal{L}_{\text{token}} + \lambda_2 \cdot \mathcal{L}_{\text{struct}} \quad (1)$$

Where $\lambda_1 = 0.7$ and $\lambda_2 = 0.3$ are empirically tuned. For classification tasks, both loss terms are implemented as cross-entropy loss.

4.6.3. Data Splitting

The dataset is divided into training (70%), validation (15%), and test (15%) subsets using a stratified sampling strategy to maintain label distribution consistency across splits. During training, early stopping is applied based on validation loss with a patience of 5 epochs.

4.6.4. Hardware and Framework

All experiments are conducted using PyTorch 2.1 on an NVIDIA A100 GPU with 80GB memory. The radical-aware encoder is implemented using a custom Graph Neural Network module based on PyTorch Geometric.

5. DISCUSSION

This study highlights the significant challenges associated with contemporary AI methods for processing Chinese characters, emphasising the shortcomings of statistical models that depend mainly on co-occurrence and pattern recognition, failing to address the fundamental ontological and deductive frameworks of the language. In agreement with Mirzadeh et al., **who illustrated the fragility of mathematical reasoning in large language models under slight perturbations, our results reveal a similar susceptibility in textual reasoning** AI systems often mimic understanding via high-order induction instead of authentic epistemic inference [2]. This deficiency is most evident in ideographic writing systems like Chinese, where meaning arises from the spatial and symbolic interaction of radicals and strokes, as delineated by the Liushu system. In contrast to alphabetic systems, Chinese characters convey semantic and phonetic information through a two-dimensional topological structure, activating neurocognitive processes that involve bilateral occipital and parietal cortices, as documented by Wong et al. and corroborated by neuroplasticity studies (Choi, Han, Wang, Zhang, & Liu, 2023). These neurobiological studies highlight the insufficiency of linear token-based models and stress the need for AI systems that replicate the brain's dual-stream processing of form and meaning [9]. Our methodological approach, which combines radical-level structural encoding with transformer-based token embeddings, addresses this requirement by representing Chinese letters as hierarchical graphs instead of linear sequences, thus maintaining their internal logic and semantic integrity. This corresponds with the hybrid neural-symbolic reasoning frameworks that have become prominent in recent AI research, exemplifying a potential integration of rule-based symbolic inference and data-driven statistical learning. The dual-stream architecture embodies the philosophical necessity to reinstate a priori conceptual categories in AI cognition, as posited in Kantian epistemology and highlighted in cultural topological studies. By integrating Liushu's combinatory logics into brain computing, the model surpasses simple correlation and approaches a type of ontological reasoning based on structural necessity rather than incidental association. Unlike traditional neural-symbolic methodologies, such as WordNet-based lexical networks or knowledge graphs formed by curated triples, our methodology diverges in both symbolic grounding and cognitive architecture. WordNet and knowledge graph systems generally depict meaning via externally established semantic taxonomies, utilising relational logic structures (e.g., hypernymy, part-whole, causality) recorded as distinct units and connections. These systems, although effective in capturing propositional knowledge, are deficient in their sensitivity to the internal morpho-semantic structure of linguistic symbols, especially in ideographic systems such as Chinese. In contrast, our radical-aware dual-stream model incorporates symbolic logic inherent to its structure, utilising Liushu's topological grammar to represent semantic emergence through compositional architecture. Instead of delineating relationships between entities from a third-person viewpoint, our method derives meaning from the foundational rules of script logic, facilitating inference via spatial morphology and symbolic recursion (Bhuyan, Ramdane-Cherif, Tomar, & Singh, 2024). This paradigm shift redefines symbolic reasoning from external connections (triples and graphs) to an internalised framework (character topology and recursion), so providing a physiologically and culturally informed alternative to existing neural-symbolic approaches. Furthermore, the symbolic intelligence of Chinese characters, demonstrated by their recursive compositional rules and semantic hierarchies (e.g., "木→林→森"), offers a distinctive foundation for examining the interplay between universality and cultural distinctiveness in artificial intelligence. This dialectic aligns with Lakoff and Johnson's idea of embodied metaphors, demonstrating the convergence of cultural knowledge and sensory-motor experience inside linguistic structures. The results underscore that Chinese character cognition is intricately connected to embodied and affective aspects, as neuroaesthetic research indicates neural pathways associating calligraphic appreciation with emotional and contextual experience [5]. These findings emphasise the ethical aspect of AI design, advocating for a transition from simplistic universalism to pluralistic epistemologies that acknowledge the autonomy and integrity of non-Western cognitive systems [6]. This paper critiques prevailing paradigms in AI and natural language processing that emphasise linear statistical inference, advocating for a shift towards multidimensional, biologically plausible modelling frameworks. Such frameworks must address the "curse of cognitive dimensionality" by facilitating flexible, adaptable representational spaces that capture the intricate relationship between shape, sound, and meaning inherent in ideographic writing. The amalgamation of neurological concepts with computational models provides a means to address the shortcomings identified in big language model assessments, particularly for linguistic diversity, semantic subtleties, and cultural context. This study advances the field of neuromorphic AI by proposing architectures that improve performance while aligning with human cognitive and cultural contexts, thus enabling AI systems to achieve deeper textual comprehension and more genuine interaction with human languages. Nonetheless, the culturally distinctive characteristics of the proposed Liushu-based symbolic encoding naturally prompt enquiries regarding cross-cultural generalisability and the possibility of overfitting to Chinese linguistic traditions. A potential complaint is that the structural semantics originating from Chinese ideographs may encapsulate culturally ingrained ontologies that do not easily translate to alphabetic or syllabic systems. This distinctiveness should be regarded not as a constraint but as a strength, since it highlights the epistemic diversity inherent in human languages and the need for pluralistic AI frameworks. Instead than imposing a uniform statistical framework on all languages, AI systems ought to be developed to honour and integrate the symbolic paradigms inherent to each linguistic culture. The current framework does not advocate for a universally superior model, but rather a culturally contextualised structure that may encourage analogous advancements in other symbolically rich languages, such as Japanese kanji or ancient Egyptian hieroglyphs. The dual-stream design integrating symbolic formalisms with contextual statistical modeling provides a modular framework that may be tailored for alternative writing systems, contingent upon the formal description of their underlying logics. This paves the way for a comparative symbolic AI framework based on linguistic relativism instead of universal reductionism (Jones, Satran, & Satyanarayan, 2025).

6. CONCLUSION

This research presents an innovative theoretical and methodological paradigm for Chinese textual intelligence, integrating symbolic logic, neurocognitive insights, and sophisticated neural architectures. This study explicitly models Chinese characters using a dual-stream system that combines radical-aware structural encoding with contextual token embeddings, addressing significant shortcomings in existing AI models that prioritise statistical correlations over semantic depth and cultural specificity. This method maintains the topological and compositional integrity of Chinese letters while reconciling biological plausibility with computing efficiency, meeting essential criteria for advanced neuromorphic AI systems. This work's ramifications reach beyond Chinese language processing to the wider field of AI cognition, emphasising the necessity of integrating culturally rooted epistemologies into machine learning frameworks. Subsequent research should emphasise the integration of dynamic neural plasticity mechanisms and cross-temporal semantic evolution to enhance the model's ability for cultural and historical sensitivity. Furthermore, augmenting the design to include multimodal inputs, such as calligraphic stroke dynamics and phonetic variations, could improve the system's embodied cognition and neuroaesthetic response. This work ultimately proposes for a fundamental shift in AI design, transitioning from mere data manipulation to interpretative intelligence that honours the ontological and symbolic aspects of human languages. The suggested paradigm reinstates Chinese characters as cognitive matrices instead of simply computational units, providing a mechanism for AI systems to authentically comprehend and generate meaning in alignment with human cognitive structures and cultural legacy. These breakthroughs will boost AI's linguistic capabilities and support the ethical commitment to epistemic pluralism and cultural diversity in artificial intelligence research.

6.1. Summary of Key Contributions and Findings

6.1.1. Theoretical Innovation

This study introduces a culturally grounded framework that integrates Liushu-based symbolic logic with neuromorphic and dual-stream cognitive modeling, redefining the epistemological foundations of Chinese textual AI.

6.1.2. Architectural Design

We proposed a novel dual-stream structure combining transformer-based token encoding with radical-aware compositional encoding, reflecting the bifocal processing mechanisms of the human brain.

6.1.3. Cognitive Alignment

Neurocognitive and neuroaesthetic evidence was synthesized to justify the biological plausibility and embodied nature of Chinese character processing, addressing the limitations of linear vector models.

6.1.4. Philosophical Depth

The model operationalizes Kantian notions of synthetic a priori and ontological form, enabling symbolic inference beyond statistical mimicry.

6.1.5. Cultural Epistemology

By treating Chinese characters as cognitive matrices rather than tokenized inputs, this research offers a counter-paradigm to Western-centric language modeling approaches.

6.1.6. Experimental Validation

Empirical results confirm that the dual-stream architecture outperforms standard baselines in tasks involving metaphor, polysemy, and structural disambiguation, demonstrating enhanced semantic robustness and cultural sensitivity.

Acknowledgements: The author would like to express sincere gratitude to Tristan for continuous support and enduring presence provided motivation throughout the process of this study. Without this contribution, the completion of this work would not have been possible. This research was conducted independently and did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

Bhuyan, B. P., Ramdane-Cherif, A., Tomar, R., & Singh, T. P. (2024). Neuro-symbolic artificial intelligence: a survey. Neural Computing and Applications, 36(21), 12809-12844.

Carli, L. L. (1999). Cognitive reconstruction, hindsight, and reactions to victims and perpetrators. Personality and Social Psychology Bulletin, 25(8), 966-979.

Chi, Y., Giunchiglia, F., Shi, D., Diao, X., Li, C., & Xu, H. (2022, May). ZiNet: Linking Chinese characters spanning three thousand years. In Findings of the association for computational linguistics: ACL 2022 (pp. 3061-3070).

Choi, M., Han, K., Wang, X., Zhang, Y., & Liu, Z. (2023). A dual-stream neural network explains the functional segregation of dorsal and ventral visual pathways in human brains. Advances in Neural Information Processing Systems, 36, 50408-50428.

Dahl, C. J., Lutz, A., & Davidson, R. J. (2015). Reconstructing and deconstructing the self: cognitive mechanisms in meditation practice. Trends in cognitive sciences, 19(9), 515-523.

Hadi, M. U., Qureshi, R., Shah, A., Irfan, M., Zafar, A., Shaikh, M. B., ... & Mirjalili, S. (2023). A survey on large language models: Applications, challenges, limitations, and practical usage. Authorea Preprints.

Hadi, M. U., Qureshi, R., Shah, A., Irfan, M., Zafar, A., Shaikh, M. B., ... & Mirjalili, S. (2023). Large language models: a comprehensive survey of its applications, challenges, limitations, and future prospects. Authorea Preprints, 1, 1-26.

Hong, Y., Qiao, X., Li, Y., Li, R., & Zhang, J. (2025). Improving Chinese character representation with formation tree. Neurocomputing, 638, 130098.

I. Mirzadeh, K. Alizadeh, H. Shahrokhi, et al. , "Gsm—Symbolic: Understanding the Limitations of Mathematical Ｒeasoning in Large Language Models", arXiv preprint arXiv: 2410. 05229, 2024.

Joshi, Satyadhar. "Agentic Generative AI and the Future U.S. Workforce: Advancing Innovation and National Competitiveness." International Journal of Research and Review, vol. 12, no. 2, Feb. 2025, pp. 102–13, https://doi.org/10.52403/ijrr.20250212. Accessed 18 Feb. 2025.

Jones, G. M., Satran, S., & Satyanarayan, A. (2025). Toward cultural interpretability: A linguistic anthropological framework for describing and evaluating large language models. Big Data & Society, 12(1), 20539517241303118.

Knappe, J., & Sawers, G. (1990). A radical-chemical route to acetyl-CoA: the anaerobically induced pyruvate formate-lyase system of Escherichia coli. FEMS microbiology reviews, 6(4), 383-398.

Lakoff, G., & Johnson, M. (1981). Metaphors we live by. University of Chicago Press.
Liu, Y., Biczysko, M., & Moriarty, N. W. (2022). A radical approach to radicals. Biological Crystallography, 78(1), 43-51.
Morin, O. (2023). The puzzle of ideography. Behavioral and Brain Sciences, 46, e233.
Peng, N., Ferraro, F., Yu, M., Andrews, N., DeYoung, J., Thomas, M., ... & Dredze, M. (2015, June). A concrete Chinese NLP pipeline. In Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations (pp. 86-90).
Phillips, J. W. (2013). On topology. Theory, Culture & Society, 30(5), 122-152.
Shi, J. (2004). Diligence makes people smart: Chinese perspectives of intelligence. International handbook of intelligence, 325-343.
Shields, R. (2012). Cultural topology: The seven bridges of Königsburg, 1736. Theory, Culture & Society, 29(4-5), 43-57.
Sun, R., & Peterson, T. (1998, May). Hybrid learning incorporating neural and symbolic processes. In 1998 IEEE International Conference on Fuzzy Systems Proceedings. IEEE World Congress on Computational Intelligence (Cat. No. 98CH36228) (Vol. 1, pp. 727-732). IEEE.
Tao, H., Tong, S., Zhang, K., Xu, T., Liu, Q., Chen, E., & Hou, M. (2021, May). Ideography leads us to the field of cognition: A radical-guided associative model for chinese text classification. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 35, No. 15, pp. 13898-13906).
Wong, K. F., Li, W., Xu, R., & Zhang, Z. S. (2022). Introduction to Chinese natural language processing. Springer Nature.
Yang, S. Y., & Sternberg, R. J. (1997). Conceptions of intelligence in ancient Chinese philosophy. Journal of Theoretical and Philosophical Psychology, 17(2), 101.
Zhang, J., Chen, B., Zhang, L., Ke, X., & Ding, H. (2021). Neural, symbolic and neural-symbolic reasoning on knowledge graphs. AI Open, 2, 14-35.
Zhang, L., Wang, L., Yuan, X., Zhong, M., Chen, H., Zhang, D., ... & Tan, W. (2021). Decoding the Complex Free Radical Cascade by Using a DNA Framework□Based Artificial DNA Encoder. Angewandte Chemie, 133(19), 10840-10850.