

Exploring Agency in Chinese Environmental Thought through Human–Machine Co-Creation: Toward a New Methodological Framework

*Jana S. ROŠKER**

*Riccardo POZZO***

*Marko ROBNIK ŠIKONJA****

*Josef VAN GENABITH*****

Abstract

In an age marked by global crises, a transcultural understanding of the philosophical foundations of human thought—especially in epistemology and ethics—has become essential. Contemporary ecological crises unfold on a planetary scale, exceeding both the conceptual limits of nation-states and the historically constituted horizons of dominant Western paradigms. Addressing them requires normative frameworks grounded in global interdependence rather than any single civilizational matrix, emerging through sustained transcultural dialogue that integrates diverse philosophical traditions into a shared ecological perspective. As a major geopolitical actor, China demands serious philosophical engagement in this context. Yet traditional research on Chinese ecological thought has relied on close reading, which limits conceptual analysis. While hermeneutic methods remain indispensable for interpreting individual texts, they cannot systematically identify recurring conceptual patterns across extensive corpora spanning diverse discourses and long, diachronically differentiated periods. The result is a fragmentary interpretation focused on isolated authors, concepts, or epochs rather than reconstructed continuities and ruptures. This paper proposes a new model based on human–machine collaboration, employing self-supervised embeddings and large language models to generate empirical insights that complement qualitative analysis. By integrating theoretical and AI-driven approaches, it expands reflective capacities and

* Jana ROŠKER, Department of Asian Studies, Faculty of Arts, University of Ljubljana, Slovenia.
Email address: jana.rosker@ff.uni-lj.si

** Riccardo POZZO, Department of Literary, Philosophical, and Art History Studies, Tor Vergata University of Rome.
Email address: riccardo.pozzo@uniroma2.it

*** Marko ROBNIK ŠIKONJA, Faculty of Computer and Information Science, University of Ljubljana.
Email address: marko.robnišikonja@fri.uni-lj.si

**** Josef VAN GENABITH, DFKI German Research Center for Artificial Intelligence, Saarland University.
Email address: josef.van_genabith@dfki.de



establishes a framework for evaluating AI-generated philosophical insights, contributing to more robust transcultural ecological frameworks.

Keywords: ecological crisis, transcultural philosophy, Chinese environmental ethics, digital humanities, artificial intelligence, human–machine collaboration, epistemology

Raziskovanje dejavnosti akterja v kitajski okoljski misli skozi soustvarjanje človeka in stroja: oblikovanje novega metodološkega okvira

Izvleček

V dobi, ki jo zaznamujejo globalne krize, postaja transkulturno razumevanje filozofskih temeljev človeške misli nujno, še posebej na področju epistemologije in etike. Sodobne ekološke krize se odvijajo na planetarni ravni in presegajo tako ozke zamejitve nacionalnih držav, kot tudi zgodovinsko pogojena konceptualna obzorja prevladujočih zahodnih paradigem. Njihovo razumevanje in razreševanje zahteva normativne predpostavke, ki temelje na v globalni soodvisnosti, ne pa zgolj na idejnih matricah, vzpostavljenih v posameznih regijah; takšne predpostavke se lahko razvijejo le skozi kontinuiran transkulturni dialog, ki omogoča primerjanje, interakcijo in medsebojno preobrazbo raznolikih filozofskih tradicij znotraj skupnega ekološkega obzorja. Kot pomemben geopolitični akter zahteva Kitajska v tem kontekstu resno filozofsko obravnavo. Vendar pa se je tradicionalno raziskovanje kitajske okoljske misli opiralo predvsem na natančno branje posameznih besedil, kar omejuje možnosti širše konceptualne analize. Čeprav ostajajo hermenevtične metode za interpretacijo posameznih tekstov nujno potrebne, je z njimi težko zajeti in prepoznati ponavljajoče se konceptualne vzorce v obsežnih korpusih, ki segajo čez različne diskurze in dolga, diahrono diferencirana obdobja. Posledica je fragmentarna interpretacija, ki je osredotočena na posamezne avtorje, pojme ali obdobja brez možnosti rekonstrukcije kontinuitet in prelomov. Prispevek predlaga programski raziskovalni okvir, utemeljen na sodelovanju človeka in stroja, ki z uporabo samonadzorovanih vektorskih predstavitev (embeddingov) in velikih jezikovnih modelov omogoča pridobivanje empiričnih uvidov kot pomembne dopolnitve kvalitativnih analiz. Z integracijo teoretskih in z umetno inteligenco podprtih pristopov razširja refleksivne zmožnosti ter vzpostavlja okvir za vrednotenje filozofskih uvidov, ki jih generira umetna inteligenca, s čimer prispeva k oblikovanju zanesljivejših metod transkulturne ekologije.

Ključne besede: ekološka kriza, transkulturna filozofija, kitajska okoljska etika, digitalna humanistika, umetna inteligenca, sodelovanje človeka in stroja, epistemologija

Introduction

The present study is situated at the intersection of three fields: environmental philosophy, traditional Chinese philosophy, and an interdisciplinary methodology that connects humanities research with advanced artificial intelligence (AI) technologies. Its significance lies in the attempt to integrate these domains into a

coherent analytical framework that is empirically grounded, philosophically rigorous, and normatively oriented toward addressing global ecological challenges. The central aim of this paper is therefore twofold: it advances both theoretical and methodological innovation, thereby generating new solutions for systematic, historical, and linguistic research.

At the theoretical level, contemporary ecological crises have reached a degree of systemic complexity that can no longer be adequately addressed within the normative and conceptual frameworks of individual disciplines or philosophical traditions. Climate change, biodiversity loss, and large-scale environmental degradation are the results of interrelated economic, political, and technological processes with global effects. Precisely because these crises are global, they cannot be adequately addressed—let alone resolved—within the narrow confines of individual nation-states. This necessitates a global and egalitarian dialogue that includes the voices of diverse cultural and philosophical traditions.

Non-Western cultures, however, have long been excluded from globally dominant philosophical and political discourses, reflecting persistent forms of Eurocentrism rooted in the historical processes of industrial modernity and its colonial and postcolonial expansion. This expansion entailed not only the global spread of technological and economic structures, but also the dominance of specific theoretical, ethical, and epistemological paradigms. As a result, Western normative and axiological criteria continue to prevail globally, despite the fact that regions such as China and the broader East Asian sphere have significantly reshaped the contemporary geopolitical landscape of technological development and innovation.

Within dominant Euro-American philosophical traditions, ecological thought has often been framed in terms of modern foundational individualism,¹ grounded in a substantial and atomized conception of the subject, as well as within Cartesian dualisms and utilitarian views of nature. Even ecological frameworks that explicitly reject anthropocentrism frequently rely on implicit models of the human self as an autonomous, pre-social individual, whose freedom is defined primarily by the absence of heteronomy and the maximization of benefit and choice. In such frameworks, responsibility toward ecological systems is typically construed as an external moral obligation imposed upon an otherwise self-sufficient subject. This structural separation of freedom, responsibility, and relational embeddedness contributes to tensions between ecological regulation and individual liberty, as well as to a gap between normative ideals and their practical realization.

1 This concept was shaped by Henry Rosemont Jr. in his book *Against Individualism: A Confucian Rethinking of the Foundations of Morality, Politics, Family, and Religion* (2015). In this context, the term *foundational individualism* is linked to the modern Western assumption according to which legal, political and moral principles of a democratic state must necessarily be rooted in the rights and choices of concrete individuals. In this framework, social and interpersonal relations are secondary arrangements among independent, pre-social and fully constituted selves.

In this context, increasing attention has been directed toward non-Western philosophical traditions as potential sources of alternative ecological paradigms (e.g. Nelson 2023, 8). East Asian philosophies, and Chinese thought in particular, occupy a distinctive position in this regard, not only because of their global ecological relevance and geopolitical significance, but also due to the specific paradigms underlying their classical discourses (D'Ambrosio 2023; Li 2023, 95). These include relationality, processual transformation, and the dynamic integration of human beings within social and natural environments (Sernelj 2024a, 35). Classical Chinese thought developed conceptual vocabularies in which ethical cultivation, social order, and cosmological processes are structurally interconnected rather than analytically separated. Concepts such as harmonization, balance, co-feeling, responsive resonance, and transformative process express normative orientations inseparable from underlying cosmological and anthropological assumptions.

At the same time, we distance ourselves from romanticized projections of “Eastern wisdom” as a ready-made solution to contemporary crises. We do not treat classical Chinese philosophies as prescriptive alternatives to modernity, but as historically situated conceptual resources whose dynamic, relational and systemic referential frameworks can critically illuminate the normative assumptions of industrial modernity. Their significance lies not in providing direct answers, but in expanding the conceptual horizon within which ecological responsibility and subjectivity can be rethought² (Zhang 2023, 70).

Another issue of importance in this context is the fact that in currently prevailing interpretations, Chinese philosophy and ecological thought are often interpreted as prioritizing hierarchy and conformity at the expense of individual freedom. Such interpretations are reinforced by contemporary political discourses that selectively instrumentalize classical concepts to legitimize centralized authority and social discipline. Consequently, Chinese thought is frequently characterized as relational yet normatively conservative, and ethically holistic yet lacking a clear conception of free subjectivity (Sernelj 2024b, 145). These readings tend to project Euro-American dichotomies between autonomy and relationality onto traditions grounded in different assumptions about agency and action. More recent research, however, indicates that Chinese philosophy contains internally differentiated notions of subjectivity that cannot be reduced to either collectivism or

2 In this regard, we draw on several comprehensive databases of classical Chinese sources, including the *Chinese Text Project* (<https://ctext.org>), for which a license has already been secured by the library of the Department of Asian Studies at the University of Ljubljana to enable digital analyses and other AI-based procedures; the open-access database *Key Concepts in Chinese Thought and Culture* (<https://www.chinesethought.cn/EN/newslist.aspx?nid=114&pid=131>); and *Shidian guji* (https://www.shidianguji.com/?page_from=reading_page), which contains approximately 5,842 classical texts.

atomistic individualism. Forms of agency grounded in self-cultivation, situational judgment, and empathetic responsiveness point to models of freedom and critical reflection that arise from, rather than stand in opposition to, relational embeddedness. From this perspective, such agents exhibit a distinct internal configuration: their identities are not constituted as substance-bearing entities, but as relational, contextual, and dynamic Selves. These Selves are both capable of and obligated to act in accordance with diverse situations, while simultaneously preserving and further developing a specific, uniquely individual identity. This form of individual identity may be conceptualized as the *Self as continuum*.

On the methodological level, our study is defined by the need to overcome the limitations of traditional research methods in the exploration of ecological thought. Existing insights have largely been derived from close readings of selected canonical texts. However, it remains unclear whether the identified concepts represent structurally recurring patterns across broader traditions or merely isolated theoretical postulates. Philosophical influence is shaped not only by canonical texts, but also by institutional practices, socio-political ideologies, and historically specific modes of transmission, including ritual practices, cultural traditions, oral transmission, education, and popularization. These factors determine which conceptual models of relationality, subjectivity, and nature become socially influential and politically relevant.

Traditional methods based on close reading restrict the scope of such analysis. While hermeneutic approaches are indispensable for interpreting individual texts, they cannot systematically address conceptual patterns across large corpora spanning diverse discourses and long, diachronically differentiated periods. The result is a body of fragmentary interpretations focused on individual authors, concepts, or epochs, rather than on systematically reconstructed continuities and ruptures. Claims regarding the presence or absence of free subjectivity in Chinese thought thus often rely on selective evidence rather than empirically grounded reconstructions of conceptual structures.

The increasing availability of digital corpora and advances in artificial intelligence open new possibilities for overcoming these limitations. The central methodological challenge of our study is therefore to develop an approach that integrates AI-based empirical analysis into philosophical research, without reducing normative meaning to statistical regularities.

The considerations outlined above constitute the shared motivation that has brought the authors of this paper into collaboration: to explore ways of advancing philosophical research through AI-based analytical approaches. The contributors represent diverse scientific disciplines and employ heterogeneous methodologies. This study brings together scholars in philosophy and researchers working with advanced digital technologies, affiliated with institutions whose research

traditions are, at first glance, methodologically divergent or even incompatible: the Faculty of Arts and the Faculty of Computer and Information Science at the University of Ljubljana; the German Research Center for Artificial Intelligence at Saarland University; and the Department of Literary, Philosophical, and Art History Studies at Tor Vergata University of Rome. The paper is thus the result of a collaboration that transcends not only national and institutional boundaries, but also disciplinary divides.

This article represents an initial step and first outcome within a broader research trajectory addressing a topic of considerable relevance—not only with regard to its anticipated theoretical contributions, but also in terms of developing a new methodological framework situated at the intersection of conceptual and empirical approaches. More specifically, the aim is to articulate a synthesis of qualitative and quantitative research. Such a synthesis, however, presents substantial challenges. While the social sciences and digital humanities have increasingly adopted machine learning techniques such as topic modelling to address domain-specific questions, this development entails a number of underexamined problems. Existing approaches frequently rely on assumptions that are problematic in relevant respects; they are non-deterministic, difficult to evaluate against a sufficiently wide range of competing models, and often lack a well-defined linguistic interface. Moreover, they do not model topics in a scholarly sense, that is, as coherent thematic or conceptual units. These structural limitations render the interpretation of results vulnerable to apophenia—the attribution of meaning to random patterns – and to confirmation bias, whereby interpretations align with prior expectations. Although certain aspects of the underlying statistical models can be partially validated, any robust conceptual validation would require systematic triangulation with complementary methods and human evaluation, as well as a clarification of whether statistically distinctive patterns of lexical co-occurrence correspond in a reliable manner to conceptually meaningful topics (Shadrova 2021). The present study contributes to this discussion by foregrounding epistemological issues concerning the interface between topic modelling and linguistic concepts, as well as the integration of such results into argumentative structures.

Accordingly, the first phase of the study—the primary focus of this paper—is dedicated to developing a methodological framework grounded in iterative collaboration across distinct procedural stages. Its aim is to facilitate a coherent integration of empirical, AI-based analyses with conceptual and normative interpretation.

A Complicated Attraction: AI Technologies and Normative Research Methods

The present paper thus seeks to clarify both the methodological foundations and the concrete procedural framework underlying the central research question of our study. This question concerns how concepts of human–nature relationality manifest themselves across classical Chinese philosophical texts, and to what extent these relational structures incorporate notions of agency, responsibility, and normative self-cultivation within broader cosmological–ecological frameworks. Within this context, the inquiry is necessarily coupled with a further question: In what ways do different AI technologies—particularly strategies for adapting large language models—shape the structural representation of relational and processual concepts in classical Chinese philosophy? Put differently, how can the paradigm shift from the close reading of individual texts to the distant reading of large corpora be productively integrated into new forms of interdisciplinary research (Pozzo 2021, 21)?

The central methodological difficulty consists in formulating an approach that integrates AI-driven analyses into philosophical investigation without reducing the normative dimension of meaning to purely statistical patterns. What is required is a framework in which empirically derived semantic structures and philosophical interpretation interact in a reciprocal and complementary manner, without being subordinated to one another. Within such a configuration, AI-generated outputs are not treated as definitive results, but rather as heuristic resources that guide, problematize, and refine normative–hermeneutic analysis. Conversely, philosophical interpretation supplies the evaluative, interpretative, and normative grounding upon which the final conclusions of our joint investigation are established.

From a methodological perspective, this integration is realized through the systematic interaction of concept clusters, referential frameworks, and their synthesis via the transcultural dialectical method of sublation. Concept clusters are here understood as empirically derived structures produced by AI-based analytical procedures—such as topic modelling and embedding-based clustering—which identify latent semantic patterns in large textual corpora by grouping linguistic units according to distributional similarity and recurrent co-occurrence structures. In computational terms, such methods rely on vector-space representations of language and probabilistic modelling techniques to capture semantic proximity between textual elements (Blei, Ng, and Jordan 2003; Mikolov et al. 2013).

This interpretation is methodologically supported by research in distributional semantics and machine learning approaches to text analysis, including work by Marko Robnik-Šikonja on feature weighting, semantic similarity, and explanatory

modelling in data mining (e.g. Robnik-Šikonja and Kononenko 2008; Vreš and Robnik-Šikonja 2022). Similarly, Josef van Genabith's contributions to multilingual natural language processing, semantic representation, and large-scale language technologies (e.g. Karakanta, Dehdari and van Genabith 2018; Pal et al. 2018) provide important methodological foundations for modelling structured semantic relations across corpora. These approaches establish the computational basis for identifying recurrent patterns of semantic relatedness in textual data. At a higher level of abstraction, such clusters may be interpreted as approximations of recurring semantic relations, while remaining grounded in statistical representations rather than normatively structured meanings. This distinction is crucial for maintaining the methodological boundary between empirical pattern extraction and philosophical conceptual analysis.

In contrast, the notions of referential frameworks and the method of sublation originate in Jana Rošker's prior work in transcultural philosophy. Referential frameworks denote historically sedimented configurations of meaning that organize concepts, categories, norms, and modes of reasoning within specific intellectual traditions. They function as hermeneutic horizons that generate and stabilize distinct philosophical paradigms, including characteristic ontological, epistemological, and axiological orientations (Rošker 2022). Through these paradigmatic configurations, referential frameworks shape the conditions under which concepts acquire meaning, normative force, and relational function within coherent systems of thought.

The transcultural philosophical method of sublation (Rošker 2024, 81) constitutes a process-oriented, dialectical approach developed within the framework of post-comparative philosophy. It facilitates mediation between different philosophical traditions by identifying family resemblances alongside internal divergences, situating these within their respective referential frameworks, critically addressing conceptual limitations, and selectively preserving complementary elements. In doing so, it enables transformative shifts in interpretation. Rather than aiming at a fixed or final synthesis, sublation produces new insights through an ongoing, reflexive engagement grounded in philosophical subjectivity. Within our study, this method is employed to mediate between heterogeneous historical discourses and the conceptual tensions disclosed by AI-based analyses, reconfiguring apparent oppositions into coherent, processual constellations.

Methodologically, this implies that tensions between empirically identified semantic patterns and established interpretative horizons are not addressed through reduction or formal synthesis, but are instead reworked through a process of dialectical reconstruction. On this basis, the philosophical task consists in deriving referential frameworks from the empirical configurations disclosed by concept clusters, and in mobilizing these reconstructed frameworks to interpret,

contextualize, and critically evaluate the semantic structures identified through AI-based analysis.

In this manner, our study aims to respond to the challenges outlined above by embedding AI-based concept cluster analysis within a broader framework of trans-cultural philosophical inquiry. Rather than relying on predetermined conceptual schemes, we need to reconstruct culturally conditioned and context-sensitive structures of meaning from empirically traceable semantic constellations. This approach facilitates a comprehensive examination of the relational embeddedness of humans and nature, ecological normativity, and forms of subjectivity as historically variable and internally differentiated phenomena. By integrating large-scale, diachronic discourse analysis with philosophically grounded interpretation, we seek to move beyond entrenched dichotomies—such as those between freedom and relationality, ecology and autonomy, and empirical analysis and normative theory.

Operational Procedures

In the first phase of the study, reflected in the results presented in this paper, we developed the foundational structure of our interdisciplinary iterative process. The operational core of our study is constructed as an iterative cycle that combines computational and philosophical procedures to identify, classify, validate, and interpret conceptual relations within clearly delineated discursive contexts. The methodology adopts a concept-oriented and discursively organized analytical design, ensuring practical feasibility while maintaining theoretical rigor and hermeneutic depth. It will be implemented in the following eight steps (A-H).

(A) Corpus Formation and Discursive Structuring

Our study will be initiated through the compilation of a digital corpus comprising classical Chinese philosophical texts. This corpus is organized along major discursive traditions, such as Proto-classical texts, Daoism, Neo-Daoism, Confucianism, Han Confucianism, the Logicians, Eclectics, Buddhism, and Neo-Confucianism, each of which constitutes a distinct analytical segment. At the same time, a number of key concepts undergo internal diachronic transformations within these traditions—for instance, in the transition from early Daoism to Neo-Daoism, or from classical Confucianism to its Han and Neo-Confucian developments. Rather than treating such historical shifts as a separate axis of segmentation, they are incorporated as structurally embedded trajectories within the respective discursive formations. This design allows for systematic comparison of relational configurations across traditions, while preserving sensitivity to their internal historical differentiation and conceptual development.

A defined set of core concepts serves as the principal object of analysis. The analytical focus of the research is organized around an initial cluster of terms central to human–nature relationality and subjectivity, including: 天 (*tian*), 地 (*di*), 人 (*ren*), 身 (*shen*), 己 (*ji*), 心 (*xin*), 自 (*zi*), 性 (*xing*), 情 (*qing*), 体 (*ti*), 用 (*yong*), 理 (*li*), 气 (*qi*), 德 (*de*), 仁 (*ren*), 义 (*yi*), 礼 (*li*), 知 (*zhi*), 行 (*xing*), 诚 (*cheng*), 意 (*yi*), 志 (*zhi*), 欲 (*yu*), 修 (*xiu*), 养 (*yang*), 中 (*zhong*), 和 (*he*), 度 (*du*), 同 (*tong*), 山 (*shan*), 海 (*hai*), 川 (*chuan*), 草 (*cao*), 木 (*mu*), 禽 (*qin*), and 兽 (*shou*). These concepts are introduced into the computational system as lexical units. Through an iterative process combining computational analysis with philosophical assessment, their relational configurations are identified and provisionally categorized into analytically defined types, thereby enabling systematic philosophical validation.

(B) Concept-Oriented Extraction of Relational Structures

For each selected concept, a concept-centred computational procedure is conducted across the entire corpus, which serves as a referential background, as well as within each individual discursive segment. At this stage, the analysis is confined to the extraction of formal patterns. The procedure encompasses the retrieval of local textual contexts (at the level of sentences or paragraphs), the identification of co-occurrence patterns that exceed predefined frequency thresholds, the detection of directional relations (e.g., $A \rightarrow B$ in contrast to $B \rightarrow A$), the recognition of parallel and contrastive structures (such as $A / \text{不}A$, 是 / 非, 有 / 无), and the assessment of distributional stability within each discursive segment. The resulting output consists of structured relational datasets associated with each concept. These results capture empirically observable structural configurations and do not yet constitute interpretive claims.

(C) Philosophically Defined Criteria of Relational Analysis

Prior to computational typification, the philosophical component of the study establishes an orientational taxonomy of relational categories, including:

- normative (A serves as a criterion of validity for B),
- causal (A constitutes the source or condition of B),
- correlative (A and B are mutually co-determining),
- complementary (A and B are interdependent yet non-identical),
- hierarchical (A functions as a superordinate bearer of B),
- contradictory (A and B are mutually exclusive).

For each category, precise operational definitions are formulated. In addition, we will compile a philosophically grounded, curated set of validated examples representing prototypical instances of each relational type. These serve as an initial reference framework to guide the subsequent computational classification.

(D) AI-Supported Classification of Relational Structures

Building on the philosophically specified criteria and curated exemplars, AI-based methods are employed to categorize the extracted relational structures into provisional types. This involves interpreting the detected structural configurations as possible manifestations of particular relational categories, while factoring in their contextual positioning and distributional frequency. Each assigned category is accompanied by a graded confidence measure. At this stage, these classifications are understood as tentative, probabilistic assessments rather than conclusive determinations.

(E) Philosophical Evaluation and Iterative Refinement

The proposed typifications are subsequently subjected to philosophical and sinological evaluation. This stage involves revising the assigned relational categories, identifying ambiguous or borderline instances, and detecting discursively specific variations in relational structures.

On this basis, a validated reference subset is established, which serves as a calibration corpus for the refinement of computational classification. Through iterative cycles, the classification process becomes increasingly consistent, context-sensitive, and attuned to philosophical criteria. In this way, computational modelling and hermeneutic analysis evolve in a relation of reciprocal interaction rather than hierarchical sequencing.

(F) From Patterns of Relations to Configurational Semantics

Following such iterative validation and calibration, the concepts can acquire an analytical role as relational nodes situated within discursively differentiated semantic constellations. Consolidated relational profiles that are weighted by frequency, sensitive to directionality, and anchored in specific contexts, provide the empirical foundation for reconstructing referential frameworks and paradigmatic configurations. Within this methodological architecture, AI operates as an epistemic tool, while philosophical inquiry maintains interpretive and normative primacy.

(G) Reconstructing Frameworks of Reference

Once relational classifications have undergone iterative validation and calibration, stable relational profiles emerge for each concept within specific discursive corpus cuts. On this empirically grounded basis, philosophical reconstruction proceeds to a higher analytical level. Through systematic comparison of relational networks across discursive traditions, our investigation can proceed to the reconstruction of discursively specific referential frameworks characterized by distinct ontological

and cosmological orientations, conceptions of relationality, and models of subjectivity and selfhood. On this basis, we will reconstruct and critically assess discursively differentiated configurations of ecological normativity, clarifying how relational embeddedness, agency, and responsibility are expressed within each framework and how these configurations shape broader eco-humanistic paradigms. At this stage, relational patterns will be analysed as structurally coherent constellations that disclose the internal logic of particular traditions.

The transcultural philosophical method of sublation will then be applied to mediate between these reconstructed configurations. We will identify structural resemblances between different discourses and the referential frameworks into which they are embedded. This helps us to expose conceptual divergences within these similarities and to place them into a mutually complementary relationship in order to establish possible alternative models of ecological normativity and relational responsibility, a model that will be capable of creating sustainable forms of human–nature interaction without collapsing autonomy into isolation or embeddedness into subordination.³

(H) Ensuring Methodological Coherence

Across all phases, the study can therefore uphold three core methodological safeguards: (I) statistical regularities are interpreted as evidentiary patterns rather than normative claims; (II) relational classifications remain provisional and open to ongoing philosophical validation; (III) computational abstractions are subjected to continuous epistemic critique and iterative revision.

Within this framework, AI operates as a tool of epistemic support embedded in a philosophically grounded research design, rather than as a source of interpretive authority. The methodology enables scalable semantic analysis while maintaining hermeneutic sensitivity and normative reflexivity. In this way, the study advances from structured relational data toward the reconstruction of paradigmatic configurations, without reducing philosophical meaning to statistical distributions.

Bridging Philosophical Insight and AI Analysis

As we can see from the operational segments described above, the study is organized around a clearly defined distribution of intellectual responsibilities, combined with coordinated leadership of particular sections and sustained iterative

3 Rather than imposing uniform synthesis, sublation generates dynamic and reflexive re-configurations that allow relational embeddedness and free subjectivity to be understood as internally differentiated yet structurally comparable across discursive contexts.

exchange between philosophical–Sinological analysis and AI-based research. In our view, such an organizational framework secures methodological coherence, facilitates an efficient workflow, and ensures the ongoing integration of conceptual reconstruction with computational modelling throughout all stages of the study.

In this way, we aim to overcome the limitations associated with the assumption of a fixed and uniquely defined semantic space, which underlies most topic modelling approaches and stands in significant tension with linguistic and text-theoretical perspectives⁴ (Shadrova 2021, 2). Through the iterative relation between philosophers and AI-experts, we also aim to reduce the danger of oversimplification regarding the complexities of the quantification of meaning, which usually tends to be further complicated by text-linguistic features such as text length, complexity, or genre (*ibid.*). The integration of philosophical interpretation with AI-based linguistic analysis requires a carefully coordinated and continuously reflexive collaboration among the experts from the disciplines involved in the project. This is all the more important because the individual phases of the study, carried out within different disciplinary frameworks, differ significantly in their methodological nature.

The philosophical–Sinological component of the study undertakes philosophical–hermeneutic analysis, including the formulation of relational typification criteria and the selection of representative examples, as well as the reconstruction of referential frameworks and the application of the method of sublation. It ensures the theoretical coherence of the overall approach by embedding computational procedures within a philosophically grounded interpretive framework. In addition, it contributes to the axiological dimension of qualitative normative analysis, engages in the reconstruction of ecological normativity, and provides interpretative expertise in eco-humanistic paradigms and the aesthetic mediation of human–nature relationality. Particular attention is devoted to models of subjectivity and ethical agency within holistic conceptions of human embeddedness in natural and social environments. In this way, philosophical analysis plays a decisive role in reconstructing models of relational selfhood and in carrying out the philosophical validation of relational typifications.

Complementing these dimensions, the perspective of intellectual history—associated in particular with the work of Riccardo Pozzo (e.g. 2014; 2021)—provides a critical layer of contextualization. Intellectual history situates conceptual configurations within their broader historical trajectories, tracing the transformation, transmission, and reinterpretation of ideas across time and cultural settings. It

4 This problem is linked to the fact that topics, themes, or categories of content are not well-defined linguistic concepts. Hence, their boundaries are generally fuzzy and dynamic, and they can only be constructed in an in-depth analytical process, not extracted as basic information, because they are largely derived from implicit meaning and not word co-occurrence *per se*.

thereby enables a historically grounded reconstruction of referential frameworks, ensuring that semantic constellations are not treated as static structures but as dynamically evolving formations. This perspective is essential for integrating diachronic depth into the analysis, for identifying shifts in conceptual usage and normative orientation, and for mediating between empirical pattern recognition and historically informed philosophical interpretation.

The AI-based empirical component, by contrast, focuses on the adaptation of large language models (such as Gemma 3, Mistral, and potentially OpenEuroLLM, alongside selected proprietary models), building on prior developments in modelling for Slovene and other low-resource languages—an approach that is equally applicable to specialized philosophical corpora. Among other procedures, Direct Preference Optimization (DPO) is employed to improve alignment and reasoning consistency. This component also draws on recent datasets and methods developed for explanation-oriented modelling, including approaches from common-sense reasoning and natural language inference. The objective is not mechanistic interpretability in a narrow sense, but the training of models capable of generating structured, context-sensitive explanations of their outputs (cf. Žagar et al. 2024; Klemen et al. 2024).

In addition, computational expertise contributes advanced knowledge in AI modelling and language technologies, particularly in relation to large language model-based decision support for the extraction and classification of relational concepts from classical Chinese textual data. Such decision-support systems are of particular relevance insofar as they enable natural language-based interaction between textual data and philosophical analysis. Model transparency is addressed through approaches in explainable AI and mechanistic interpretability, ensuring that outputs can be understood by both developers and interpreters. Evaluation procedures combine automated and human assessment, integrating quantitative and qualitative methods. This component also encompasses data preprocessing and corpus preparation. Across all stages, AI-based analytical procedures remain closely integrated with philosophical interpretation, ensuring continuous interaction between empirical modelling and conceptual analysis.

Taken together, these elements establish a methodological foundation that is both structurally coherent and conceptually creative. The first phase of the study has thus been devoted to constructing an integrated framework in which computational procedures and philosophical analysis are not merely combined, but systematically interrelated through iterative feedback and mutual calibration. This groundwork represents a decisive step toward a new mode of interdisciplinary inquiry, in which empirical semantic patterning and hermeneutic interpretation are brought into a productive, non-reductive relation.

The novelty of this approach lies in its refusal to subordinate philosophical meaning to computational abstraction, while at the same time fully engaging the analytical potential of AI-based methods. By embedding large-scale semantic analysis within a philosophically grounded architecture—supported by sinological expertise and intellectual-historical contextualization—the study establishes a model for integrating quantitative and qualitative research without collapsing their respective epistemic functions. In doing so, it not only addresses the methodological limitations of existing approaches, but also opens a pathway toward a more reflexive, context-sensitive, and conceptually robust form of AI-assisted philosophical analysis.

The importance of this interdisciplinary methodological work lies precisely in its capacity to overcome the entrenched limitations of both methodologies employed in our research. On the one hand, it addresses the insufficiency of purely hermeneutic approaches when confronted with large-scale, diachronically differentiated corpora; on the other, it critically repositions AI-based methods, preventing the reduction of meaning to statistical regularities and reasserting the necessity of philosophical interpretation as a normative and evaluative practice. By establishing iterative procedures of validation, calibration, and reconstruction, the study demonstrates how computational analysis can be transformed into an epistemic instrument that expands, rather than replaces, philosophical inquiry.

At a broader level, the significance of this approach extends beyond state-of-the-art research methodologies. By enabling a more comprehensive and structurally grounded analysis of human–nature relationality, subjectivity, and ecological normativity in Chinese philosophy, the study as a whole aims to contribute to the ongoing effort to rethink the conceptual foundations of global ecological ethics. It opens the possibility of identifying alternative configurations of agency, responsibility, and relational embeddedness that do not reproduce dominant dichotomies between autonomy and ecology, or between individual freedom and normative constraint.

Conclusion

The present study has sought to establish a methodological foundation for integrating philosophical–hermeneutic analysis with AI-based linguistic modelling in the investigation of classical Chinese environmental thought. As demonstrated throughout, such an integration is neither a straightforward technical extension of existing methods nor a simple juxtaposition of disciplines. Rather, it requires a carefully constructed framework in which empirical pattern extraction and conceptual interpretation enter into a sustained, reflexive, and non-hierarchical relation.

The first phase presented in this paper is therefore not merely preparatory, but constitutive. It provides the conceptual architecture, procedural design, and epistemological safeguards required for a new form of transcultural research. In particular, the integration of concept clusters, referential frameworks, and the method of sublation—supported by sinological expertise and intellectual-historical contextualization—creates the conditions for a systematic reconstruction of relational semantic structures across traditions. This groundwork enables the transition from isolated textual interpretations to empirically informed, philosophically validated models of conceptual constellations.

In this sense, the methodological framework developed here constitutes an initial but decisive step toward the construction of new paradigms of ecological thought. These paradigms are not derived from any single tradition, but emerge through a transcultural process of reconstruction, comparison, and transformation. By bringing philosophical insight and AI-based analysis into a productive and critically mediated relationship, our methodology possibly lays the foundation for a more inclusive, reflexive, and globally attuned ecological ethics—one capable of addressing the complexity of contemporary environmental crises at the level of both empirical knowledge and normative understanding.

Acknowledgements

The authors acknowledge the financial support from the Slovenian Research and Innovation Agency (ARIS) as part of the research core funding *Asian Languages and Cultures* (P6-0243) and in the framework of the research project *New Theoretical Approaches to Comparative Transcultural Philosophy and the Method of Sublation* (J6-50208). The research for the present paper has also been funded by the University of Ljubljana, in the framework of the Start-up programme (RSF_ARIS.ZRP.26-TS) *Ecology and Relational Ethics in the East Asian Past and Present* and by Tor Vergata University of Rome (Uniroma2_ALEF_RSA24_Pozzo) *Nature and Life from Plotinus to Kant*.

References

- Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 2003 (3): 993–1022.
- D'Ambrosio, Paul. 2023. "Boundary of the Sky: Environmentalism, Daoism, and the Logic of Increase." *Asian Studies* 11 (2): 41–67. <https://doi.org/10.4312/as.2023.11.2.41-67>.

- Karakanta, Alina, Jon Dehdari, and Josef van Genabith. 2018. "Neural Machine Translation for Low-Resource Languages Without Parallel Corpora." *Machine Translation* 32 (1): 167–89.
- Klemen, Matej, Aleš Žagar, Jaka Čibej, and Marko Robnik Šikonja. 2024. "SINLI: A Slovene Natural Language Inference Dataset and its Evaluation." In *The 2024 Joint Intern. Conference on Computational Linguistics, Language Resources and Evaluation*, edited by Nicoletta Calzolari. (LREC-COLING 2024, Torino: 14859-14870). <https://aclanthology.org/2024.lrec-main.1294.pdf>.
- Li, Manhua. 2023. "The Biopolitics of Nourishing Life: Daoism as Environmental Philosophy." *Asian Studies* 11 (2): 95–116. <https://doi.org/10.4312/as.2023.11.2.95-116>.
- Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. "Efficient Estimation of Word Representations in Vector Space." *arXiv preprint arXiv:1301.3781*. <https://arxiv.org/pdf/1301.3781>.
- Nelson, Eric S. 2023. "Perspectives on Chinese, Buddhist, and Environmental Philosophy: Introduction." *Asian Studies* 11 (2): 7–11. <https://doi.org/10.4312/as.2023.11.2.7-11>.
- Pal, Santanu, Nico Herbig, Antonio Krüger, and Josef van Genabith. 2018. "A Transformer-Based Multi-Source Automatic Post-Editing System." *Proceedings of the Third Conference on Machine Translation (WMT)*. Volume 2. *Shared Task Papers*: 827–35. <https://aclanthology.org/W18-6468.pdf>.
- Pozzo, Riccardo. 2014. "Storia storica e storia filosofica della filosofia nel XX e nel XXI secolo." *Archivio di Storia della Cultura* 27 (1): 361–372.
- . 2021. *History of Philosophy and the Reflective Society*. Walter de Gruyter GmbH & Co KG.
- Robnik-Šikonja, Marko, and Igor Kononenko. 2008. "Explaining Classifications for Individual Instances." *IEEE Transactions on Knowledge and Data Engineering* 20 (5): 589–600.
- Robnik-Šikonja, Marko, et al. 2014. "Explaining Data Mining Models." In *Encyclopedia of Machine Learning and Data Mining*.
- Rosemont, Henry Jr. 2015. *Against Individualism: A Confucian Rethinking of the Foundations of Morality, Politics, Family, and Religion*. Lexington Books.
- Rošker, Jana S. 2022. "Sublating Sinic Relationism: On a Winding Path from Transcultural to Global Ethics." *Asian Studies* 10 (3): 81–104. <https://doi.org/10.4312/as.2022.10.3.81-104>.
- . 2024. "Re-Inventing Chinese Philosophy through the Method of Sublation: Introducing a New Methodology for Research in Chinese Philosophy from the Perspective of Transcultural Comparisons." *Asian Studies* 12 (2): 245–70. <https://doi.org/10.4312/as.2024.12.2.245-270>.

- Sernelj, Téa. 2024a. "The Rise of Individual Personhood in Early Medieval China." *Asian Studies* 12 (3): 131–49. <https://doi.org/10.4312/as.2024.12.3.131-149>.
- . 2024b. "In Defence of Subjectivity and Autonomy: Shitao's Aesthetic Theory and His Critique of the Mainstream School of Painting in the Early Qing Dynasty." *Asian Studies* 12 (2): 145–69. <https://doi.org/10.4312/as.2024.12.2.145-169>.
- Shadrova, Anna. 2021. "Topic Models do not Model Topics: Epistemological Remarks and Steps Towards Best Practices." *Journal of Data Mining & Digital Humanities* 2021. <https://jdmdh.episciences.org/8608>.
- Vreš, Domen, and Marko Robnik-Šikonja. 2022. "Preventing Deception with Explanation Methods using Focused Sampling." *Data Min Knowledge* 2024 (38): 3262–3307. <https://doi.org/10.1007/s10618-022-00900-w>.
- Zhang, Ellen Y. 2023. "The 'Greening' of Daoism: Potential and Limits." *Asian Studies* 11 (2): 69–94. <https://doi.org/10.4312/as.2023.11.2.69-94>.
- Žagar, Aleš, Katja Dobrovoljc, Tina Munda, Mojca Brglez, and Marko Robnik Šikonja. 2024. "Knowledge-Enhanced Winograd Schema Challenge KEWSC 1.0." *Slovenian Language Resource Repository CLARIN.SI*, 2024. CLARIN.SI data & tools. <http://hdl.handle.net/11356/1988>.