Research Article / Научная статья
УДК 903.27:7.031.1:004.8
https://doi.org/10.14258/tpai(2026)38(3).-06
EDN: KUNAKT

ARTIFICIAL INTELLIGENCE AND THE TRANSFORMATION OF EVIDENCE IN ROCK ART RESEARCH: ETHICAL AND EPISTEMOLOGICAL CHALLENGES

Hugo de Lara López

National University of Distance Education, Madrid, Spain;
hugdelara@calatayud.uned.es, https://orcid.org/0009–0004–0564–7222

Abstract. Artificial intelligence (AI) is increasingly transforming research practices in archaeology, including the study of rock art, by enabling new approaches to data analysis, pattern recognition, and digital documentation. While these developments offer significant methodological opportunities, their ethical and epistemological implications remain insufficiently explored. This paper examines the integration of AI in rock art research, arguing that its impact extends beyond technical innovation to reshape the nature of archaeological evidence. Drawing on recent applications of machine learning and generative AI, it identifies key challenges related to data bias, scarcity, and representativeness, as well as issues of model opacity and the limitations of standard evaluation metrics. Particular attention is given to synthetic imagery as a response to data constraints, highlighting associated risks such as dataset contamination, loss of traceability, and the blurring of boundaries between empirical and generated data. The paper proposes that AI introduces a shift from observation-based evidence to model-mediated evidence, raising critical questions about how knowledge is produced and validated. In response, it outlines a set of practical principles to support the responsible use of AI, emphasizing transparency, traceability, expert supervision, and the clear distinction between empirical and synthetic data.

Keywords: artificial intelligence, rock art, AI ethics, GenAI, epistemology

Acknowledgements: the author would like to express his gratitude to the translator, María Ferre Fernández, for her efforts in working on this article despite its challenges.

For citation: De Lara López Hugo. Artificial Intelligence and the Transformation of Evidence in Rock Art Research: Ethical and Epistemological Challenges. *Teoriya i praktika arheologicheskih issledovanij = Theory and Practice of Archaeological Research*. 2026;38(3):120–134. (In English). https://doi.org/10.14258/tpai(2026)38(3).-06

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ТРАНСФОРМАЦИЯ СВИДЕТЕЛЬСТВ В ИССЛЕДОВАНИЯХ НАСКАЛЬНОЙ ЖИВОПИСИ: ЭТИЧЕСКИЕ И ЭПИСТЕМОЛОГИЧЕСКИЕ ВЫЗОВЫ

Уго де Лара Лопес

Национальный университет дистанционного образования (UNED), Мадрид, Испания;
hugdelara@calatayud.uned.es, https://orcid.org/0009–0004–0564–7222

© Уго де Лара Лопес

Резюме. Искусственный интеллект (ИИ) все больше трансформирует исследовательскую практику в археологии, включая изучение наскальной живописи, открывая новые подходы к анализу данных, распознаванию образов и цифровой документации. Хотя эти разработки открывают значительные методологические возможности, их этические и эпистемологические последствия остаются недостаточно изученными. В данной статье рассматривается интеграция ИИ в исследования наскальной живописи, при этом утверждается, что его влияние выходит за рамки технических инноваций и меняет саму природу археологических свидетельств. С опорой на недавние применения машинного обучения и генеративного ИИ в статье выделяются ключевые проблемы, связанные с предвзятостью, дефицитом и репрезентативностью данных, а также вопросы непрозрачности моделей и ограничений стандартных метрик оценки. Особое внимание уделяется синтетическим изображениям как способу преодоления ограничений, связанных с данными, при этом подчеркиваются сопутствующие риски, такие как загрязнение наборов данных, утрата прослеживаемости и стирание границ между эмпирическими и сгенерированными данными. В статье выдвигается предположение, что ИИ приводит к переходу от свидетельств, основанных на наблюдениях, к свидетельствам, опосредованным моделями, что поднимает важные вопросы о том, как создаются и проверяются знания. В ответ на это в статье излагается набор практических принципов, направленных на поддержку ответственного использования ИИ, с акцентом на прозрачность, прослеживаемость, экспертный надзор и четкое разграничение между эмпирическими и синтетическими данными.

Ключевые слова: искусственный интеллект, наскальная живопись, этика ИИ, GenAI, эпистемология

Благодарности: автор хотел бы выразить признательность переводчику Марии Ферре Фернандес за проделанную работу над данной статьей, несмотря на ее трудоемкость.

Для цитирования: де Лара Лопес Уго. Искусственный интеллект и трансформация свидетельств в исследованиях наскальной живописи: этические и эпистемологические вызовы // Теория и практика археологических исследований. 2026. Т. 38, № 3. С. 120–134. https://doi.org/10.14258/tpai(2026)38(3).-06

Introduction

The rapid development of artificial intelligence (AI) has significantly transformed research practices across multiple scientific disciplines, including archaeology and cultural heritage studies. In recent years, machine learning (ML) and computer vision techniques have been increasingly applied to tasks such as site detection, pattern recognition and digital documentation, offering new possibilities for the analysis of complex archaeological datasets (Bruno et al., 2010; Bevan, 2015; Fiorucci et al., 2020; Casillo et al., 2025). At the same time, numerous scholars have emphasized the need to critically examine the ethical implications of AI systems, particularly regarding issues of bias, transparency, accountability and epistemic reliability (Mittelstadt et al., 2016; Floridi et al., 2018; Jobin, Ienca, Vayena, 2019).

Despite this expanding debate, the ethical dimension of AI applications in rock art research remains underdeveloped. Existing studies have primarily focused on methodological innovation and performance evaluation, often overlooking the broader implications that these technologies may have for the interpretation, documentation and conceptualization of the archaeological record (Horn et al., 2022; Suhaimi et al., 2023; Huo D., Yang, Hou M., 2025). This gap is especially significant in the study of rock art, where the material is inherently

© Hugo de Lara López

heterogeneous, fragmentary and context-dependent, posing substantial challenges for data-driven approaches (Fiorucci et al., 2020; Sokolovska, 2024).

In fact, studies on ML and computer vision techniques applied to rock art have demonstrated both their potential and limitations in the field. These techniques can support tasks such as automatic detection, classification and segmentation of painted motifs, but ultimately cannot replace expert interpretation (Kowlessar et al., 2021; Jalandoni, Zhang, Zaidi, 2022; Suhaimi et al., 2023; De Lara, 2025; Huo D., Yang, Hou M., 2025). Model performance seems to be highly dependent on the quality and representativeness of training datasets, which means that model accuracy decreases when working in complex contexts characterized by, for instance, faded pigments or ill-defined superpositions (Fiorucci et al., 2020; Suhaimi et al., 2023; De Lara, 2025; Huo D., Yang, Hou M., 2025).

At the same time, the emergence of generative AI (GenAI) has introduced new possibilities for addressing data scarcity through synthetic imagery (Cioni et al., 2023; Veggi, 2025; De Lara, 2026). Experimental approaches have shown that AI can generate visually plausible variations of documented motifs under different preservation conditions. Nevertheless, these developments also raise important concerns regarding the distinction between empirical and generated material, dataset contamination, and evidential reliability (Cioni et al., 2023; Veggi, 2025; De Lara, 2026).

This paper argues that the integration of AI into rock art research is not merely a technical issue, but an ethical and epistemological one too. Drawing on both AI ethics literature and recent applications to rock art research, this study develops a critical framework to tackle challenges related to data bias, model opacity and synthetic documentation. By situating rock art as a particularly demanding case study, the main objective of this paper is to contribute to broader debates on how AI reshapes the production of scientific knowledge and the boundaries of evidence in the humanities. Certainly, the methodological contributions of AI technologies to our field are evident, but we must ensure an appropriate ethical and epistemological reflection when adopting them.

Artificial Intelligence in Rock Art Research: Current State and Limitations

The application of AI to rock art research is a rapidly growing field within digital archaeology driven by advances in ML, computer vision and the increasing availability of digital imagery (Bevan, 2015; Fiorucci et al., 2020; Casillo et al., 2025). These technologies have introduced new approaches to the documentation, analysis and interpretation of painted motifs, particularly through automated motif detection, classification, segmentation and predictive spatial analysis.

Studies on ML for rock art research have demonstrated the supervised learning and deep learning approaches have potential for the automatic identification and segmentation of rock art motifs (Kowlessar et al., 2021; Jalandoni, Zhang, Zaidi, 2022; Suhaimi et al., 2023; De Lara, 2025; Huo D., Yang, Hou M., 2025). In controlled conditions, these methods can achieve promising results, especially when applied to well-preserved figures and relatively homogeneous datasets. However, their performance decreases significantly when confronted with the variability that characterizes rock art contexts, including irregular

© Уго де Лара Лопес

surfaces, faded pigments, fragmentation and complex superimpositions (Fiorucci et al., 2020; Suhaimi et al., 2023; De Lara, 2025; Huo D., Yang, Hou M., 2025). This limitation has also been confirmed by our latest experimental evaluation, which documented a marked decline in model performance under such conditions (De Lara, 2025). Similarly, predictive and spatial approaches based on environmental and geographical variables are also affected by comparable limitations; as their reliability depends heavily on the quality and representativeness of input data, the resulting outputs experience great changes in performance quality (Bonald et al., 2024).

Other studies have also focused on the use of GenAI for the creation of synthetic datasets to mitigate the structural scarcity of archaeological data. Some scholars argue that, by generating controlled variations of existing imagery, these approaches may improve model robustness and generalization (Cioni et al., 2023; Veggi, 2025; De Lara, 2026). These developments suggest that generative techniques, although still experimental in this discipline, are becoming increasingly relevant within the broader field of digital heritage.

Despite these advances, the field remains characterized by significant methodological and structural limitations, including the scarcity of large annotated datasets, the high variability of rock art imagery, and the lack of standardized protocols for annotation and evaluation (Bevan, 2015; Fiorucci et al., 2020; Horn et al., 2022; Sokolovska, 2024; De Lara, 2025). As a result, many studies continue to rely on small and internally developed datasets, limiting reproducibility and comparability across research contexts.

All in all, the development of AI applications in rock art research has been primarily driven by technical and methodological concerns, with comparatively limited attention to their broader ethical and epistemological implications. While these approaches offer considerable analytical potential, they also reveal structural dependencies and interpretative constraints that extend beyond purely technical considerations. In spite of these limitations, AI has become a valuable tool for rock art research. Applications such as automated detection, classification, segmentation, predictive spatial analysis and synthetic data generation have already expanded the methodological possibilities available to archaeologists and have the potential to improve documentation and analytical workflows. The ethical and epistemological challenges discussed in this paper therefore arise because of the growing influence these technologies have on archaeological research and knowledge production.

Ethical Challenges of AI Systems: Data, Models and Interpretability

One of the main ethical challenges associated with the use of AI in rock art research concerns the nature and quality of the data used to train ML models. Unlike domains where large, standardized and well-annotated datasets are available, rock art research is characterized by structural data scarcity and a high degree of heterogeneity in terms of typology, style, technique, preservation and documentation conditions (Bevan, 2015; Fiorucci et al., 2020; Horn et al., 2022; Sokolovska, 2024; De Lara, 2025). From an ethical perspective, this limitation is not simply a technical constraint, but a source of potential bias that directly shapes the behavior and outputs of AI systems. As widely discussed in AI ethics, ML models inherit the

© Hugo de Lara López

limitations, omissions and distortions embedded within their training data (Mittelstadt et al., 2016; Gebru et al., 2021; Suresh & Guttag, 2021).

In the case of rock art, this dependency becomes particularly problematic because available documentation tends to privilege well-preserved, visually clear and frequently studied motifs over more deteriorated or ambiguous contexts. Empirical applications reinforce this concern: detection models generally perform better when motifs are clearly defined and the visual signal is strong, while accuracy decreases significantly in the presence of faded pigments, irregular surfaces, fragmentation or complex superimpositions (Fiorucci et al., 2020; Suhaimi et al., 2023; De Lara, 2025; Huo D., Yang, Hou M., 2025). This asymmetry reveals a fundamental bias: AI systems are more effective at recognizing what is already visible and well documented, and less capable of identifying motifs that require interpretative reconstruction.

Additional ethical challenges emerge from annotation practices and the use of transfer learning. Image labeling through bounding boxes, segmentation masks or categorical classifications depends on expert interpretation and is therefore shaped by theoretical assumptions, prior knowledge and subjective criteria. Consequently, annotations are not neutral representations of reality. In complex cases, thusly, annotations become constructed interpretations that are encoded into the model itself. Similarly, pre-trained models adapted from large-scale non-archaeological datasets may prioritize visual patterns that are irrelevant or misleading in rock art contexts, introducing external biases that further complicate interpretability and reliability.

Beyond data-related issues, important ethical concerns also arise from the opacity of many ML systems. Deep learning models are frequently characterized as "black boxes" due to the difficulty of understanding how specific predictions or classifications are generated (Doshi-Velez & Kim, 2017; Samek, Wiegand, Müller, 2017; Lipton, 2018; Rudin, 2019). In archaeological contexts, this lack of transparency is particularly problematic because AI outputs are often presented as objective or data-driven results, despite depending on complex interactions between training data, model architecture and optimization processes.

This issue becomes especially evident in the use of evaluation metrics such as precision, recall or mean average precision (mAP). Although these indicators provide useful quantitative measures, they do not necessarily correspond to the epistemological requirements of archaeological interpretation. A model optimized for recall may generate numerous false positives, whereas one optimized for precision may overlook subtle or deteriorated motifs that require contextual and interpretative expertise. Previous studies have shown that models capable of producing acceptable results under controlled conditions may nevertheless generate inconsistent or misleading outputs when applied to degraded or heterogeneous contexts (Fiorucci et al., 2020; Suhaimi et al., 2023; De Lara, 2025; Huo D., Yang, Hou M., 2025) (Fig. 1).

Overall, these issues call into question the assumption that AI-based analyses provide neutral or objective representations of the archaeological record. Instead, they suggest considering AI outputs as situated and model-mediated results shaped by both computational design and human interpretation. Recognizing this condition is essential to avoid over-reliance on automated analyses and to keep expert judgment at the core of rock art research.

© Уго де Лара Лопес

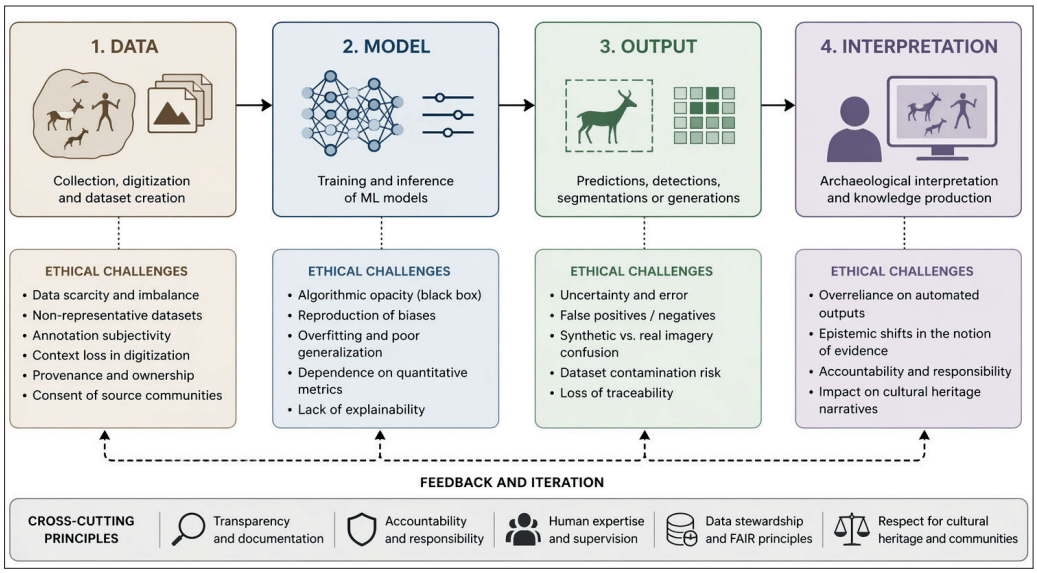


Fig. 1. Ethical pipeline of AI in rock art research. The diagram represents the main stages of an AI workflow (data, model, output and interpretation) and the key ethical challenges associated with each stage. Cross-cutting principles such as transparency, accountability, expert supervision and data stewardship are essential to ensure responsible and epistemologically grounded research practices

Рис. 1. Этические аспекты применения ИИ в работе с наскальным искусством. На схеме представлены основные этапы рабочего процесса с использованием ИИ (данные, модель, результат и интерпретация), а также ключевые этические проблемы, связанные с каждым из них. Сквозные принципы, такие как прозрачность, подотчетность, экспертный надзор и ответственное управление данными обоснованных методов исследования

Generative AI and Synthetic Evidence in Rock Art Research

The structural limitations associated with data availability in rock art research have encouraged the exploration of alternative strategies aimed at increasing the quantity and diversity of training datasets. One of them is the use of AI-generated synthetic data, a promising approach particularly in contexts where empirical documentation is limited, heterogeneous or difficult to standardize.

In ML, data augmentation techniques have long been used to improve model robustness through controlled variations of existing datasets, such as rotations, scaling or changes in illumination (Shorten & Khoshgoftaar, 2019). More recently, advances in generative models, including generative adversarial networks and diffusion-based systems, have extended these possibilities by enabling the creation of entirely new images that preserve key characteristics of the original data while introducing controlled variability (Goodfellow et al., 2014; Ho, Jain, Abbeel, 2020; Bommasani et al., 2021).

In the context of rock art research, these developments have opened the door to the experimental generation of synthetic documentation to simulate different conditions of preservation, visibility and recording (Cioni et al., 2023; Veggi, 2025; De Lara, 2026). In point of fact, our findings from our latest experimental work have shown that GenAI

© Hugo de Lara López

can produce visually plausible variations of documented motifs. We have simulated processes such as pigment fading, fragmentation and changes in lighting conditions while maintaining a direct relationship with the original archaeological record, achieving variable degrees of success yet consistently reaching the same conclusion: synthetic documentation can indeed support ML models in rock art research (De Lara, 2026). From a methodological perspective, these approaches may improve model robustness and generalization, facilitate the creation of more balanced datasets, and allow the systematic testing of models under varying visual conditions, thereby expanding the range of experimental tools available for rock art research.

At the same time, the use of synthetic imagery introduces significant ethical and epistemological challenges. As generative models become increasingly sophisticated, the distinction between empirical documentation and artificially generated material becomes more difficult to establish. Synthetic images may appear visually coherent and archaeologically plausible while lacking a direct correspondence to observed material reality. In rock art research, where visual evidence plays a central role in documentation and interpretation, this ambiguity is certainly problematic (Chesney & Citron, 2019; Crawford, 2021).

A further concern involves the potential contamination of datasets. If synthetic images are incorporated into training data without clear differentiation, they may reinforce recursive biases and amplify patterns that are not grounded in empirical evidence (Bender et al., 2021). This risk is especially significant in a field already characterized by limited and heterogeneous datasets. Closely related to this issue is the problem of traceability. The use of synthetic material requires transparent documentation regarding how images are generated, transformed and linked to their empirical sources. Without such transparency, it becomes difficult to assess data validity, reproduce results or critically evaluate the role of synthetic imagery in the analytical process.

Beyond these practical concerns, GenAI raises deeper questions regarding the status of evidence in archaeological research. Synthetic images are not direct observations of the past, but constructed representations derived from probabilistic models trained on existing datasets. As such, they occupy an ambiguous position between empirical documentation and computational inference, introducing a new category of visual material that challenges traditional distinctions between observation, interpretation and representation.

Hence, the growing adoption of GenAI reflects a broader shift towards hybrid forms of knowledge production that combine empirical observation with computational modeling. While these approaches offer important opportunities for addressing long-standing methodological limitations, they also require a critical framework capable of ensuring that synthetic representations remain transparent, traceable and grounded in archaeological reasoning (Fig. 2).

The increasing integration of synthetic and model-generated material into archaeological workflows therefore challenges both methodological practices and the epistemological foundations upon which archaeological evidence has traditionally been defined. As AI systems progressively mediate the production, transformation and interpretation of visual data, the distinction between observation, inference and representation becomes increasingly unstable.

© Уго де Лара Лопес

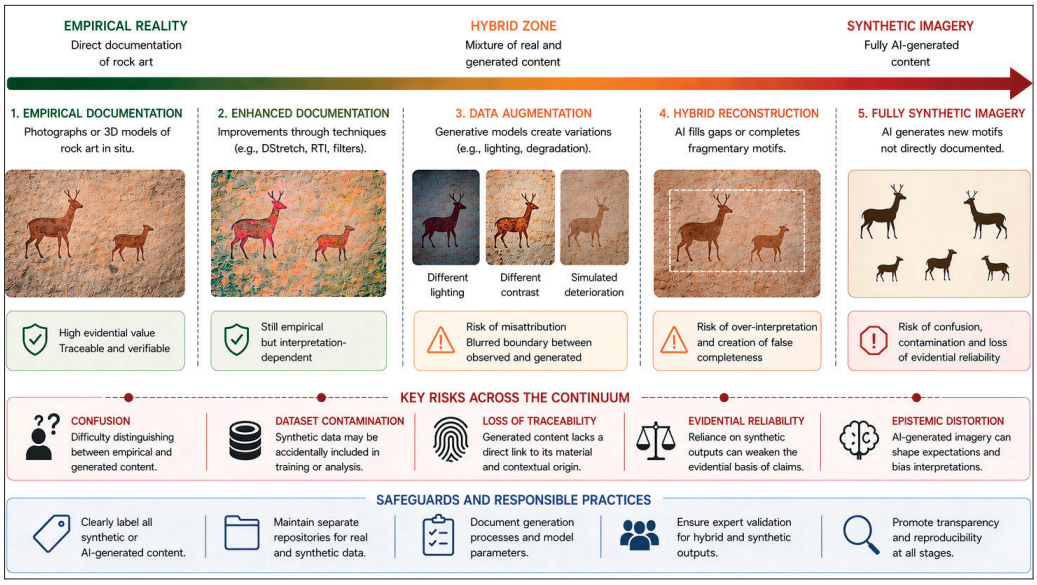


Fig. 2. Real to synthetic continuum and associated risks in rock art research. From empirical documentation to fully synthetic imagery, highlighting the hybrid zone where generative AI (GenAI) introduces both opportunities and ethical challenges. Key risks include confusion between real and generated data, dataset contamination, loss of traceability, and the weakening of evidential reliability. Responsible practices such as clear labeling, transparency and expert validation are essential to maintain the integrity of archaeological research

Рис. 2. Континуум «от реального к синтетическому» и сопряжённые с ним риски в исследованиях наскального искусства. От эмпирической документации к полностью синтетическим изображениям: фокус на гибридной зоне, в рамках которой применение генеративного ИИ (GenAI) открывает новые возможности, но одновременно порождает ряд этических проблем. К числу ключевых рисков относятся смешение реальных и сгенерированных данных, контаминация наборов данных, утрата прослеживаемости происхождения материалов, а также снижение доказательной надёжности результатов исследования. Соблюдение ответственных практик — в том числе чёткой маркировки материалов, обеспечения прозрачности методологии и обязательной экспертной валидации — является необходимым условием сохранения научной достоверности и целостности археологических исследований

Rethinking evidence: epistemological implications of AI in rock art research

The challenges identified in the previous sections point towards a broader issue that extends beyond methodological and ethical concerns: the need to reconsider the nature of evidence in rock art research in the context of AI. The integration of AI systems into archaeological workflows does not simply introduce new analytical tools, but also transforms the ways in which visual data are produced, interpreted and validated, as already noted in previous studies on the impact of data-driven approaches in archaeology (Huggett, 2020).

Traditionally, archaeological evidence has been grounded in direct observation, contextual analysis, and expert interpretation. In the case of rock art, this involves the careful examination of pigment traces, surface morphology, superimpositions, typology, style and technique all of which are evaluated within their archaeological and environmental context. This process

© Hugo de Lara López

is inherently interpretative, requiring the integration of multiple lines of evidence and the application of domain-specific knowledge.

In contrast, ML models are designed to detect statistical patterns in visual data without the benefit of contextual or theoretical frameworks. While they can identify regularities and correlations that may not be immediately apparent to human observers, they do not possess an understanding of the cultural, historical, or material dimensions of the record. As a result, their outputs reflect patterns in the data instead of interpretations grounded in archaeological reasoning (Boyd & Crawford, 2012; Leonelli, 2016).

This distinction becomes particularly significant when considering the use of AI-generated outputs as forms of evidence. In the case of detection models, the identification of a motif is based on learned visual features, which may or may not correspond to meaningful archaeological categories. In the case of generative models, the situation becomes even more complex, as the images produced are not observations of existing material, but synthetic constructions derived from probabilistic processes. These outputs introduce a new type of visual material that challenges traditional distinctions between data, interpretation and representation.

Empirical observations from recent applications reinforce this tension. It has been shown that ML models struggle to reconstruct fragmented or deteriorated motifs that archaeologists can interpret through contextual reasoning, highlighting a fundamental gap between computational detection and human inference (De Lara, 2025). At the same time, generative models are capable of producing visually coherent representations that simulate plausible states of preservation, despite not being grounded in direct observation (De Lara, 2026). Together, these findings illustrate a paradox: AI systems are limited in their ability to interpret incomplete evidence, yet capable of generating convincing representations in the absence of empirical data.

This paradox has important epistemological implications. It suggests that the use of AI in rock art research introduces a shift from evidence based on observation to evidence mediated by models. Therefore, the reliability of a result depends not only on the quality of the original data, but also on the assumptions embedded in the model and the processes through which outputs are generated. Consequently, the boundary between empirical evidence and computational inference becomes increasingly blurred (Fig. 3).

Recognizing this shift requires a critical reassessment of how evidence is defined and validated in AI-assisted research. Rather than treating model outputs as direct representations of reality, they should be understood as interpretative constructs that depend on both data and algorithmic processes. This perspective does not diminish the value of AI as a research tool, but situates it within a framework that acknowledges its limitations and its role in shaping knowledge production.

In this sense, the integration of AI into rock art research can be seen as part of a broader transformation in the epistemology of archaeology, in which digital methods play an increasingly central role. Addressing this transformation requires technical expertise and, most importantly, a reflexive approach that critically examines how new technologies influence the ways in which the past is reconstructed and understood.

© Уго де Лара Лопес

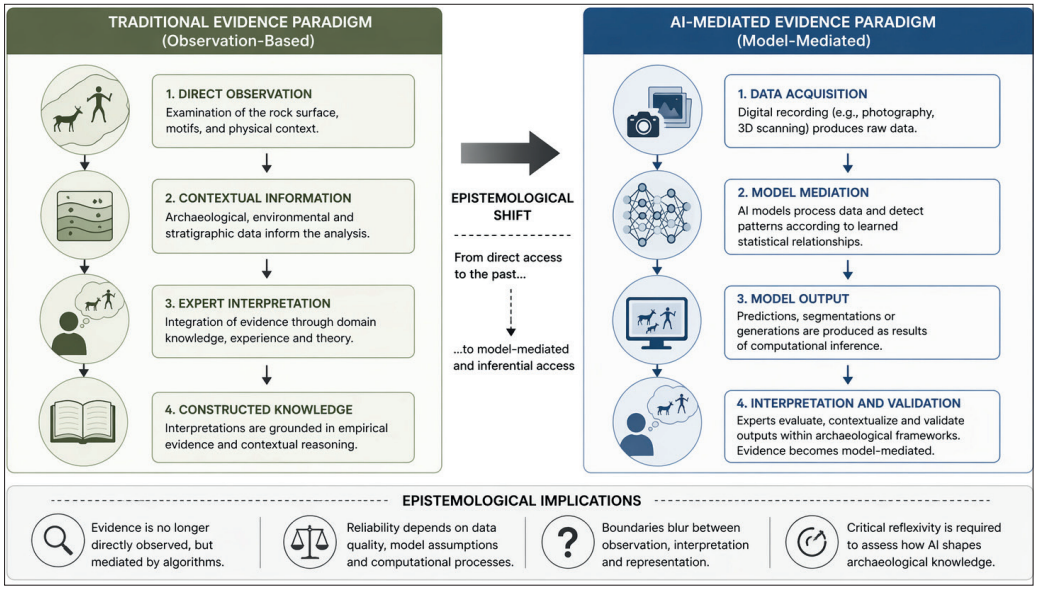


Fig. 3. From observation to model-mediated evidence: an epistemological shift in rock art research. Contrast between the traditional observation-based paradigm and the AI-mediated paradigm: AI transforms the production, mediation and validation of archaeological evidence. Particular attention is placed on the growing role of models in shaping knowledge and blurring the boundaries between observation, interpretation and representation

Рис. 3. От наблюдения к доказательствам, опосредованным моделями: эпистемологический сдвиг в исследованиях наскального искусства. Контраст между традиционной парадигмой, основанной на наблюдении, и парадигмой, опосредованной ИИ: искусственный интеллект трансформирует процессы производства, трансляции и верификации археологических данных. Особое внимание уделяется возрастающей роли моделей в формировании знания, а также размыванию границ между наблюдением, интерпретацией и репрезентацией

Towards an Ethical Framework for AI in Rock Art Research

The challenges outlined in previous sections highlight the need for a structured ethical framework to guide the use of AI in this discipline. Such a framework should not be conceived as a set of abstract principles, but as a series of practical guidelines that address the specific conditions of archaeological data, interpretation and documentation.

The first fundamental principle concerns the clear distinction between empirical and synthetic data. As GenAI becomes increasingly integrated into research workflows, it is essential to ensure that all artificially generated material is explicitly identified and documented as such. This distinction is technical, epistemological, and aims to preserve the integrity of the archaeological record and prevent the unintentional conflation of observation and simulation.

The second principle relates to data transparency and reproducibility. In line with the FAIR principles for scientific data management (Wilkinson et al., 2016), datasets used for training and evaluating ML models should be openly documented, including information on their origin, composition, annotation criteria, and potential limitations. This transparency is

© Hugo de Lara López

particularly important in rock art research, where datasets are often small, heterogeneous and depend on the context.

The third principle involves expert interpretation. Empirical supervision must remain central to all stages of the process: while AI systems can support tasks such as detection, classification, or data augmentation, their outputs should always be interpreted within an archaeological framework and validated by specialists before being incorporated into archaeological interpretation or publication. The aim is for expert review to ensure that automated outputs are critically assessed, contextualized and examined for potential errors or misinterpretations.

The fourth principle focuses on the responsible use of evaluation metrics. Quantitative indicators such as precision, recall, or mAP should not be interpreted in isolation. All indicators must be interpreted in relation to the specific research objectives and the nature of the archaeological material, which, therefore, guides methodological decisions by interpretative relevance rather than by the optimization of abstract performance values. For example, maximizing recall may be appropriate for exploratory surveys, whereas higher precision may be preferable for systematic documentation and cataloguing.

Particular attention should be given to the traceability of computational processes, which will be considered as the fifth principle. This includes documenting model architectures, training procedures, parameter choices and data transformations, especially when synthetic data are involved. Where GenAI is used, traceability should also include preserving prompts, source images, model versions and image-generation workflows whenever possible. Ensuring traceability ensures, at the same time, the replication of results and allows for the critical assessment of computational processes by other researchers.

Finally, the sixth principle, the implementation of AI in rock art research should be accompanied by a reflexive awareness of its epistemological implications. Researchers must remain attentive to how these technologies shape the production of knowledge, influence interpretative practices, and redefine the boundaries of what is considered evidence. This reflexive dimension is essential to prevent the uncritical adoption of AI tools and to maintain a balanced relationship between technological innovation and archaeological reasoning.

Ultimately, these six principles outline a framework that seeks to integrate AI into rock art research in a responsible, transparent and critically informed manner. This approach aims to maximize the potential of AI in rock art research while ensuring that its application remains aligned with the methodological and epistemological standards of the discipline.

Conclusions

The integration of artificial intelligence into rock art research represents a significant methodological advance, offering new possibilities for the documentation, analysis and interpretation of complex visual datasets. However, as this study has shown, these developments are accompanied by ethical and epistemological challenges that cannot be reduced to purely technical considerations.

The analysis of recent applications demonstrates that AI systems are deeply conditioned by the quality, structure and representativeness of the data on which they depend. In rock art research, where datasets are inherently limited and heterogeneous, these constraints introduce

© Уго де Лара Лопес

systematic biases that affect both the detection of motifs and the interpretation of results. At the same time, the opacity of many machine learning models and the reliance on quantitative evaluation metrics may create an illusion of objectivity that obscures the interpretative nature of archaeological analysis.

Generative AI introduces an additional layer of complexity. While synthetic data offer promising strategies for addressing data scarcity, they also raise important concerns regarding authenticity, traceability and dataset contamination. More fundamentally, generative models challenge traditional distinctions between empirical observation and constructed representation, introducing new forms of visual material that occupy an ambiguous position between documentation and computational inference.

When taken together, these developments point towards a broader transformation in the epistemology of rock art research. Artificial intelligence does not simply provide new analytical tools; it reshapes the conditions under which archaeological evidence is produced, mediated and validated. Recognizing this shift is essential to ensure that technological innovation does not outpace critical reflection.

In response to these challenges, this paper has proposed a framework based on the clear differentiation between empirical and synthetic data, transparency, expert supervision, traceability, and the reflexive awareness of the epistemological implications of AI. These principles seek to integrate the use of AI within a research model that remains both methodologically rigorous and epistemologically grounded.

Future research should focus on the development of standardized datasets, shared annotation protocols and hybrid approaches that combine computational methods with expert-driven interpretation. Concurrently, continued dialogue between archaeology, computer science and AI ethics will be essential to address the evolving implications of these technologies.

A further step in this trajectory may involve the progressive automation of the analytical pipeline, including image enhancement, preprocessing and synthetic data generation. In such scenarios, AI systems could evolve from assisting researchers to autonomously producing and evaluating visual data in ways that optimize their own performance. While this may offer advantages in terms of scalability and efficiency, it also raises profound ethical and epistemological concerns, particularly regarding opacity, accountability and the displacement of expert interpretation by increasingly self-referential computational processes. For this reason, full automation should not be seen as an inevitable objective, but as a critical horizon that requires sustained scrutiny and human control.

In essence, the value of AI in rock art research lies not in its capacity to replace human interpretation, but in its potential to support and extend it. Ensuring that this potential is developed in a responsible and critically informed manner remains one of the central challenges for the field.

REFERENCES

Bender E.M., Gebru T., McMillan-Major A., Shmitchell S. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In: Proceedings of the 2021 ACM Conference on

© Hugo de Lara López

Fairness, Accountability, and Transparency. New York: Association for Computer Machinery — ACM, 2021. Pp. 610–623. DOI: 10.1145/3442188.3445922

Bevan A. The Data Deluge. *Antiquity*. 2015;89(348):1473–1484. DOI: 10.15184/aqy.2015.102.

Bommasani R., Hudson D.A., Adeli E., Altman R., Arora S., von Arx S., Bernstein M.S., Bohg J., Bosselut A., Brunskill E., Brynjolfsson E., Buch S., Card D., Castellon R., Chatterji N., Chen A., Creel K., Davis J.Q., Demszky D., et al. On the Opportunities and Risks of Foundation Models. arXiv. 2021. DOI: 10.48550/arXiv.2108.07258

Bonald L., Mützenberg D., Krempser E., Verhagen P. Predicting Rock Art Sites in the Pajeú Watershed, Brazil. *Digital Applications in Archaeology and Cultural Heritage*. 2024;35(e00372). DOI: 10.1016/j.daach.2024.e00372

Boyd D., Crawford K. Critical Questions for Big Data. *Information, Communication & Society*. 2012;15(5):662–679. DOI: 10.1080/1369118X.2012.678878

Bruno F., Bruno S., De Sensi G., Luchi M.L., Mancuso S., Muzzupappa M. From 3D reconstruction to virtual reality: A complete methodology for digital archaeological exhibition. *Journal of Cultural Heritage*. 2010;11(1):42–49. DOI: 10.1016/j.culher.2009.02.006.

Casillo M., Colace F., Gaeta R., Lorusso A., Pellegrino M. Artificial Intelligence in Archaeological Site Conservation: Trends, Challenges, and Future Directions. *Journal of Computer Applications in Archaeology*. 2025;8(1):224–241. DOI: 10.5334/jcaa.207

Chesney R., Citron D.K. Deep fakes: A Looming Challenge for Privacy, Democracy, and National Security. *California Law Review*. 2019;107(6):1753–1820

Cioni D., Berlincioni L., Becattini F., Del Bimbo A. Diffusion Based Augmentation for Captioning and Retrieval in Cultural Heritage. In: 2023 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). IEEE, 2023. Pp. 1699–1708. DOI: 10.1109/ICCVW60793.2023.00186

Crawford K. Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. New Haven : Yale University Press, 2021. 288 p.

De Lara H. Inteligencia artificial en el arte rupestre: límites, posibilidades y desafíos de una evolución incierta. *Salduie*. 2025;25(2):33–46. DOI: 10.26754/ojs_salduie/sald.2025212652

De Lara H. Inteligencia artificial generativa y documentación sintética en el estudio del arte rupestre. *Cuadernos de Arqueología de la Universidad de Navarra*. 2026;34:297–314. DOI: 10.15581/012.34.027

Doshi-Velez F., Kim B. Towards A Rigorous Science of Interpretable Machine Learning. Statistics. Machine Learning. *arXiv*. 2017. DOI: 10.48550/arXiv.1702.08608

Fiorucci M., Khoroshiltseva M., Pontil M., Traviglia A., Del Bue A., James S. Machine Learning for Cultural Heritage: A Survey. *Pattern Recognition Letters*. 2020;133:102–108. DOI: 10.1016/j.patrec.2020.02.017

Floridi L., Cowls J., Beltrametti M., Chatila R., Chazerand P., Dignum V., Luetge C., Madelin R., Pagallo U., Rossi F., Schafer B., Valcke P., Vayena E. AI4People — An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*. 2018;28(4):689–707. DOI: 10.1007/s11023–018–9482–5

Gebru T., Morgenstern J., Vecchione B., Vaughan J.W., Wallach H., Daumé III H., Crawford K. Datasheets for Datasets. *Communications of the ACM*. 2021;64(12):86–92. DOI: 10.1145/3458723

© Уго де Лара Лопес

Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative Adversarial Nets. In: Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge : MIT Press, 2014. Vol. 2. Pp. 2672–2680.

Ho J., Jain A., Abbeel P. Denoising Diffusion Probabilistic Models. In: Proceedings of the 34th International Conference on Neural Information Processing Systems. New York : Curran Associates Inc., 2020. Art. 574. Pp. 6840–6851.

Horn C., Ivarsson O., Lindhé C., Potter R., Ashley G., Ling J. Artificial Intelligence, 3D Documentation, and Rock Art — Approaching and Reflecting on the Automation of Identification and Classification of Rock Art Images. *Journal of Archaeological Method and Theory*. 2022;29:188–213. DOI: 10.1007/s10816–021–09518–6

Huggett J. Is Big Digital Data Different? Towards a New Archaeological Paradigm. *Journal of Field Archaeology*. 2020;45(1): S8–S17. DOI: 10.1080/00934690.2020.1713281

Huo D., Yang S., Hou M. Using the Improved YOLOv7-Seg Model to Segment Symbols from Rock Art Images. *npj Heritage Science*. 2025;13(16). DOI: 10.1038/s40494–025–01620–2

Jalandoni A., Zhang Y., Zaidi N.A. On the Use of Machine Learning Methods in Rock Art Research with Application to Automatic Painted Rock Art Identification. *Journal of Archaeological Science*. 2022;144(105629). DOI: 10.1016/j.jas.2022.105629

Jobin A., Ienca M., Vayena E. The Global Landscape of AI Ethics Guidelines. *Nature Machine Intelligence*. 2019;1(9):389–399. DOI: 10.1038/s42256–019–0088–2

Kowlessar J., Keal J., Wesley D., Moffat I., Lawrence D., Weson A., Nayinggul A., Mimal Land Management Aboriginal Corporation. Reconstructing Rock Art Chronology with Transfer Learning: A Case Study from Arnhem Land, Australia. *Australian Archaeology*. 2021;87(2):115–126. DOI: 10.1080/03122417.2021.1895481

Leonelli S. Data-centric Biology: A Philosophical Study. Chicago : University of Chicago Press, 2016. 288 p.

Lipton Z.C. The Mythos of Model Interpretability. *Communications of the ACM*. 2018;61(10):36–43. DOI: 10.1145/3233231

Mittelstadt B.D., Allo P., Taddeo M., Wachter S., Floridi L. The Ethics of Algorithms: Mapping the Debate. *Big Data and Society*. 2016;3(2). DOI: 10.1177/2053951716679679

Rudin C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *Nature Machine Intelligence*. 2019;1(5):206–215. DOI: 10.1038/s42256–019–0048-x

Samek W., Wiegand T., Müller K.R. Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models. *ITU Journal: ICT Discoveries*. 2017;1:39–48.

Shorten C., Khoshgoftaar T.M. A Survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*. 2019;6(60). DOI: 10.1186/s40537–019–0197–0

Sokolovska A.A. Artificial Intelligence and Prospects for Its Use in the Study of Upper Palaeolithic Rock Art. *VITA ANTIQUA*. 2024;15(Methods for Monitoring and Research of Archaeological Heritage Objects):127–135. DOI: 10.37098/VA-2024–15–127–135

Suhaimi M.S., Zainuddin K., Ghazali M.D., Marzukhi F., Samad A.M., Majid Z. Comparison of One-stage and Two-stage Strategies of Machine Learning Model for Rock Art Object Detection. In: 2023 IEEE 13th International Conference on System Engineering and Technology (ICSET). IEEE, 2023. Pp. 215–220. DOI: 10.1109/ICSET59111.2023.10295089

© Hugo de Lara López

Suresh H., Guttag J.V. A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle. In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. New York : ACM Press, 2021. Pp. 172–182. DOI: 10.1145/3465416.3483305

Veggi M. State of the Art on Artificial Intelligence Resources for Interaction Media Design in Digital Cultural Heritage. *arXiv.* 2025. DOI: 10.48550/arXiv.2504.13894

Wilkinson M.D., Dumontier M., Aalbersberg I.J., Appleton G., Axton M., Baak A., Blomberg N., Boiten J.W., da Silva L.B., Bourne P.E., Bouwman J., Brookes A.J., Clark T., Crosas M., Dillo I., Dumon O., Edmunds S., Evelo C.T., Finkers R., et al. The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Scientific Data*. 2016;3:160018. DOI: 10.1038/sdata.2016.18

INFORMATION ABOUT THE AUTHOR / ИНФОРМАЦИЯ ОБ АВТОРЕ

Hugo de Lara López, PhD in Prehistory, Researcher, Professor-Tutor and External Lecturer, Department of Prehistory and Archaeology, National University of Distance Education (UNED), Madrid, Spain.

Уго де Лара Лопес, доктор философии по доистории (PhD in Prehistory), исследователь, профессор-консультант и внешний преподаватель кафедры доистории и археологии Национального университета дистанционного образования (UNED), Мадрид, Испания.

Статья поступила в редакцию 30.05.2026;
одобрена после рецензирования 17.07.2026;
принята к публикации 31.08.2026.
The article was submitted 30.05.2026;
approved after reviewing 17.07.2026;
accepted for publication 31.08.2026.

© Уго де Лара Лопес