DIGITAL HERITAGE (2025)
S. Campana, D. Ferdani, H. Graf, G. Guidi, Z. Hegarty, S. Pescarin, and F. Remondino (Editors)

GenAI-Based Reconstruction of Prehistoric Remains

Juan A. Barceló[1], Endoxia Tzerpou[1]
[1]Universitat Autònoma de Barcelona. Dept. of Prehistory

Abstract

Generative artificial intelligence (GenAI) has emerged as a powerful tool in various fields, including archaeology. However, its application in reconstructing prehistoric archaeology presents unique challenges and limitations that warrant careful consideration. Realizing the potential of Generative Artificial Intelligence to archaeological reconstruction requires a nuanced approach that acknowledges and addresses the inherent limitations of AI in this context. The "black box" nature of some AI algorithms can make it difficult to understand how reconstructions are generated. This lack of transparency poses challenges for scientific reproducibility and peer review in archaeological research. By adopting a theoretical framework that combines technological innovation with rigorous archaeological methodology and ethical considerations, we can work towards more accurate, inclusive, and responsible reconstructions of prehistoric societies. This framework not only addresses the current limitations of AI in prehistoric archaeology but also sets the stage for future research directions. As AI technology continues to evolve, ongoing critical evaluation and adaptation of these approaches will be essential to ensure that generative AI becomes a valuable tool in expanding our knowledge of prehistoric human societies while respecting the complexity and cultural sensitivity of archaeological interpretation.

I.4.5. Reconstruction (Image Processing). J.5. Computer Applications. Arts and Humanities. K.4 Computers and Society. General

DOI: 10.2312/dh.20253011

1. Introduction

The *past* –what came into existence before us- is present here and now. Buildings, constructions and objects made and used years, centuries, and millennia ago live with us. The trouble is that all this cultural heritage from the past does not look today how it was when it was first built, made and used. It has arrived to us modified by successive uses and *re*-constructions or even "ruined", destroyed or broken into pieces. We cannot see *now* how it was *then*.

What we cannot *see* in the present but existed sometime in the past can be *imagined*. When seeing ruins or broken ancient artifacts, we can build mental images about their original visual appearance. We can also "see" in our mind how people lived there, what they made in their interior or around them. And what we can "imagine" in our mind we can externalize it by drawing into a piece of paper. *Imagination* involves creatively filling in gaps in knowledge or envisioning possibilities that may not be directly derived from evidence. It allows us to fill in missing information when physical evidence is incomplete, to piece together fragmentary remains into meaningful forms, and to conjecture about aspects of past structures or artifacts that are no longer present. Furthermore, visual "imagination" allows exploring the sensory and emotional aspects of past environments, recreating possible human spatial experiences in ancient structures, and introducing ideas about possible ways people in the past may have interacted with their surroundings.

Computers can also create images. Sometimes as an extension of the human designer's hand, machines can trace points, lines, surfaces and even volumes on the computer screen using mechanisms like Bezier curves, splines and other geometrical methods. But computers can also generate images apparently by themselves in answer to verbal instructions made by the human expert. This is the domain of Generative Artificial Intelligence (GenAI), computer programs, working in the form of power chatbots giving the user personalized content recommendations.

The actual capabilities of GenAI image builders place them into the same situation as professional designers receiving feedback from archaeologists. Advantages and limitations are then apparently the same as human-based archaeological illustration. Archaeologists and general audiences are much more critical to those apparently "automatic" approaches than to human generated-images. In this paper we will try to analyze these fears and deep inside the way of using GenAI and the prospects for an automatic reconstruction of archaeological remains without the direct contribution of the human expert.

2. Reconstructing an archaeological site

The la Draga archaeological site in Northeast Iberian Peninsula is an Early Neolithic lakeside settlement, with a very good preservation of timber and organic materials (Bosch et al. 2018, Palomo et al., 2014, Piqué et al., 2021, Andreaki et al., 2022) (Fig. 1).

Figure 1: *Archaeological Excavations at La Draga archaeological site (© UAB. La Draga archaeological team)*

Based on archaeological findings of the first years of archaeological excavation, a tentative physical replica of prehistoric huts was built by the same archaeological team (Fig. 2).

Figure 2: *Physical replica of La Draga original huts. (© La Draga Archaeological Park. Banyoles).*

New archaeological data from later excavations suggested that those physical recreations were not well associated to evidence. A professional designer -Francesc Riart- translated detailed instructions by archaeologists into graphical language (Fig. 3).

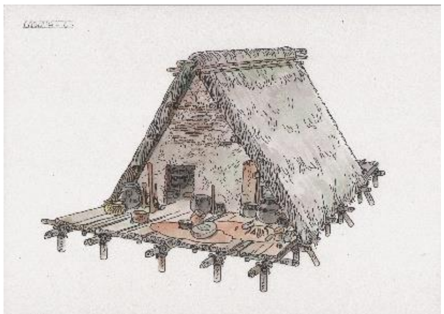

Figure 3: *Original illustration by artist Francesc Riart, translating archaeological information provided by A. Palomo, R. Piqué and X. Terradas. © F. Riart and La Draga archaeological Team.*

Using Computer-aided Design tools (© Rhinoceros3D, © Blender), archaeologists generated a 3D geometrical model completing the lacking building information with geometric completion methods and ethnoarchaeological analogies (Campana 2016, Barceló et al., 2019, Tzerpou et al. 2025) (Fig. 4).

Figure 4: *Computer-based hypothetical reconstruction of a Neolithic house from La Draga archaeological site based on archaeological evidence and geometrical and architectural reasoning (Tzerpou et al. 2025).*

Let us now consider what GenAI programs can do with the same reconstruction problem. We have asked Microsoft Copilot and OpenAI ChatGPT 4.o (both produce similar images) to explain what they know about this archaeological site and possible visual appearance of the huts used for living and domestic tasks . According to both GenAI programs, domestic units were constructed using wooden planks and other organic materials, which have been remarkably well-preserved due to the waterlogged conditions of the site. The huts were typically built on the lakeshore, taking advantage of the natural resources available in the area. The wooden structures included large planks, some over three meters long, which formed the walls and floors of the huts. Copilot affirms to know that key features of La Draga constructions are:

- **Materials**: Primarily wood, with some use of stone and organic materials.
- **Structure**: Rectangular or slightly oval shapes, with wooden posts supporting the structure.
- **Roof**: Likely thatched or covered with organic materials.
- **Interior**: Simple, with areas designated for different activities such as food preparation and storage.

Using this information, we have created a detailed text prompt, and asked different GenAI programs to generate an image that may fit the textual description. Our first essay uses a variation of ChatGPT 4.0 called ArchGPT, specifically trained for generating architectural text and images (Zhang et al., 2024).

Figure 5. *Two consecutive visualizations of La Draga Neolithic hut generated by ArchGPT using the following text prompt: "Generate a realistic 3D visualization of a prehistoric hut built over a timber platform supported by thin poles based on the following specifications:*

- ***Dimensions of timber platform:*** *6.27m length, 3.5m width*
- ***Height of thin poles supporting the platform:*** *20-80 cm*
- ***Material of poles:*** *wood logs*
- ***Material platform:*** *wood planks*
- ***Dimensions of the Hut on the platform:*** *3 m length 3 m width*
- ***Placement of the Hut on the platform:*** *centered, 2.27 from the front border, 1 m from the back border*
- ***Roof:*** *A-shaped,*
- ***Roof angle:*** *58 degrees*
- ***Roof construction:*** *covered in grass and mud*
- ***No lateral walls:*** *A shaped roof arrives to the soil*
- ***Roof a small hole at the top for smoke exit and light entrance***
- ***Front wall:*** *a small gate*
- ***Front wall construction:*** *wattle and daub*
- ***Back wall:*** *wattle and daub, no gate*
- ***Hut Floor Plan:*** *rectangular with central fire place, single entrance, sleeping area along walls]*
- ***Environment:*** *Place the hut at the shore of a lake*
- ***Location context:*** *the platforms are nailed in the soil, which is partially waterlogged*
- ***Lighting & Atmosphere:*** *Show a daytime view with soft natural lighting or an evening setting with firelight illuminating the interior.*
- ***Time period:*** *Neolithic*
- ***Archaeological Culture:*** *Cardial*

Using the same long text prompt, Grok v. 3 and Gemini create a relatively similar visualizations of the ideal La Draga unit of residence and domestic work (Fig. 6, Fig 7).

Figure 6. *Different images generated by Grok v. 3 using the same text prompt in Fig. 5.*

Figure 7. *Images generated by Gemini-Image 3 using the same text prompt in Fig. 5.*

3. Limitations of GenAI recreations of prehistoric remains.

An important limitation of all those "automatically" generated images is that they are just 2D pictures, and they are not based on precise rules about prehistoric architecture. Architectural and archaeological reconstructions require precise geometric accuracy, especially when it comes to the positioning, proportions, and alignment of fragments, what implies the need of understanding the geometric relationships between different fragments, as well as the spatial and physical properties of the recovered remains. Properties like depth, volume, and surface continuity are crucial for the proper reconstruction of ancient elements.

Computer based reconstructions in Fig. 4 are extracted from sound 3D geometric models. Although they may look like GenAI pictures, they are very different. Geometric reasoning over a 3D model represents a fundamentally different approach to understanding spatial information, compared to information extraction from a 2D image. Although ArchGPT, Grok and Gemini respect to a certain extent the scale of the measurements provided in the long and detailed text prompt, precise measurements, and distances, between different elements cannot be determined from the 2D picture alone. The resulting image may be considered hyperrealist, but the image offer no clues about how lighting, shadows, and reflections have been produced in that particular way.

The problem is that 2D representations often distort or fail to capture accurate spatial relationships. Researchers studying prehistoric longhouses found that geometric analysis helped identify missing corner posts and clarify construction sequences in buildings with multiple phases and repairs. A single 2D image, as those produced using GenAI or professional designers cannot convey the volumetric nature of archaeological excavations. While 2D outputs from traditional recording methods often fail to capture the essence of monuments - their depth, volume, micro-topography, texture and overall context.

This limitation seems to lead GenAI products to the domain of historical and archaeological dissemination and promotion but seem to be of less relevance for research purposes. In any case, the generated images are illustrations that may transport viewers to different times and places, helping them empathize with people who lived in past eras and fostering a deeper connection with history. They allow situating archaeological artifacts within their original settings and reconstruct some aspects of daily life scenes, partially based on archaeological evidence, transforming isolated discoveries into meaningful insights about ancient civilizations. Unlike original modern photographs, which capture everything indiscriminately, GenAI results may contribute to emphasize specific details, making them valuable for communication, although they may be insufficient for scientific interpretation.

The advantages of 3D models over GenAI pictures lie on its explicit spatial representation: 3D models contain complete spatial information including depth, volume, and precise geometric relationships, and they use geometric constraints such as parallelism, orthogonality, and symmetry to ensure consistency and avoid redundancy. Geometric reasoning is fundamental for archaeological interpretation of prehistoric buildings, offering insights that cannot be derived from 2D images alone. Working with full geometric information rather than projections, a user of a 3D model of a reconstructed archaeological element can refer to precise measurements of distances, angles, and volumes, deducing geometric properties proving theorems about spatial relationships, and detecting and resolving inconsistencies in geometric models through formal reasoning methods (Loch-Dehbi and Plümer 2009).

A 3D model, as the one used for generating the images in Fig. 4 (Campana 2016, Barceló et al., 2019, Tzerpou et al. 2025) allow for:

- Analysis of complete object geometry and hidden surfaces
- Understanding of internal structures and volumetric properties
- Examination of cross-sections impossible to visualize in 2D
- Dynamic viewpoint changes to analyze structures from any angle
- Examination of spatial relationships between architectural elements
- Understanding of how structures relate to their surrounding landscape

Using MasterpieceX (https://app.masterpiecex.com/generate/text-to-3d) we have obtained a full downloadable 3D model (in .fbx, .glb or USDZ/USD format) using a variation of the text prompt used in Fig. 8 –the app does not allow long text prompts-, that can be rotated and 3D inspected in its web page, and downloaded as a file (Fig. 8).

Figure 8. *Two different views of 3D model (.FBX) generated by MasterpieceX using the text input: "Generate a realistic 3D visualization of a prehistoric hut built over a timber platform of 6.27m length, 3.5m width, supported by thin poles, with an A-shaped roof that arrives to the soil, and it is covered in grass and mud".*

Another limitation of GenAI, many times observed and negatively valued, is the production of "hallucinations". The phenomenon resembles how humans might confidently fill knowledge gaps with plausible but incorrect information. However, unlike human confabulation, AI hallucinations stem from statistical pattern matching rather than cognitive processes. For instance, the program generates in answer to the text prompt fictional information, factual inaccuracies, non-existent objects, distortions in images. It may have added extraneous details beyond what's in the input data or it has drawn flawed conclusions not supported by provided information.

We have obtained "bad" visualizations, assimilable to "hallucinations" using programs like PerplexityAI (and its image generator Playground v. 3) and Stable Diffusion (Fig. 9). The way these apps translate the same text prompt is completely different to the way ArchGPT (and Copilot), Grok and Gemini.

Figure 9. *"Hallucinations" generated by PerplexityAI and Stable Diffusion, using the same text as in Fig. 5*

Technically, AI hallucinations occur when the GenAI model generates outputs that deviate from the text prompt, from the expected contextual logic or both (Alimardani and Jane 2024, Asthana and Goel 2024, Couret 2024, Martins Lemos 2024). These hallucinations happen because GenAI programs are fundamentally "blind associators" that predict the most statistically likely next element (a pixel) in a sequence based on patterns learned during training. They do not "understand" the text prompt. The goal of an AI system is to output statistically likely responses for a given input, regardless of factual accuracy.

AI hallucinations in generative AI are not truly random, despite sometimes appearing that way. Several technical factors contribute to hallucinations:

- **Training Data:** When AI ventures into areas with sparse representation in its training data, hallucination likelihood increases
- **Overfitting**: When models memorize training data too precisely, they struggle to generalize to new information.
- **Cascade effects**: When each generated word or pixel influences subsequent words or pixels, responses may grow longer and diverge from the original input
- **Decoding strategies**: When generating pixels, an AI model calculates probabilities for all possible next tokens (pixels). To improve the result, it may rank all possible tokens based on their predicted probabilities, selecting those with the highest probabilities (top *k*), and discarding all other tokens from consideration. However, this diversity comes at the cost of potentially selecting less accurate tokens. The added "randomness" from sampling among the top-*k* tokens may increase the unexpected nature of the generation compared to always choosing the most probable token, leading to a higher chance of generating hallucinated content.

4. How "good" are these visualizations?

GenAI programs, like expert archaeologists, can look for specific details in the generated image and look at an external database (or accessible internet documents) for the existence of matches between the image and the archaeological evidence. For instance, we have asked ChatGPT whether figure 5B was a good and verifiable visual reconstruction of the neolithic site of La Draga ("Can you verify whether this image is a correct visual representation of the original appearance of houses from La Draga Archaeological site"). The answer is positive, but it gives no details, just that it coincides with information existing about this site. We have asked ArchGPT whether the same image has been generated using the text prompt in Figure 5. The answer is also positive, and gives an overall ordinal measure for fitness: "strong". The answer also specifies elements in the image that match the text prompt, and minor deviations and uncertainties, like some problems in hut dimensions, roof material, which is not exactly roof material, and the lack of cultural cues to the Cardial Neolithic Culture.

PerplexityAi is also very detailed when looking for matches between an image and archaeological evidence. In the case of Figure 7A, it answers that the image shows an A-frame structure built on a wooden platform over water, illuminated from within during what appears to be dusk or evening. While this structure shares some similarities with the description in the prompt, there are several significant differences that suggest it wasn't produced using the exact specifications provided. Among the key differences, the app mentions:

- **Water Placement**: The structure in the images is built completely over water with tall supporting poles, whereas the prompt specifies a shoreline location with partially waterlogged soil and very short poles (20-80cm).
- **Structural Design**: The images show what appears to be a modern-styled tropical stilt house or water bungalow rather than a Neolithic Cardial culture dwelling.
- **Wall Construction**: The structure in the images has visible walls, while the prompt specifies "no lateral walls" with the A-shaped roof extending to the ground/platform.
- **Platform Configuration**: The platform in the images includes a surrounding deck/walkway with railings, which wasn't specified in the prompt.
- **Architectural Style**: The overall design aesthetic appears more consistent with Southeast Asian or tropical resort architecture rather than prehistoric European construction techniques.

According to PerplexityAI, Figure 7A does not accurately represent a neolithic dwelling from the archaeological site of La Draga. While it does show a structure near water with natural materials, there are significant differences between what is depicted and what archaeological evidence indicates about La Draga dwellings (It correctly uses web pages and scientific documentation about the site). Part of this answer is wrong because it is based on an old publication (Tarrus 2008), whose conclusions have been proved misleading according to more modern research (Campana 2016, Tzerpou et al., 2025), not mentioned by PerplexityAI.

In the case of the image extracted from the 3D rendered model by archaeologists (Tzerpou et al. 2025), ArchGPT answers that "The image could have been generated using the described prompt, but it's not a complete or fully accurate rendering of the prompt in its entirety. It strongly captures the core architectural features, but misses key contextual and detailed elements, especially in terms of material specificity, environmental setting, and some interior features". In the case of the "hallucination generated by Stable Diffusion (Fig. 9B), ArchGPT answers that "this image does not match the prompt's description. It represents a much later time period and an entirely different architectural tradition—incompatible with Neolithic Cardial cultures", and enumerate 7 important mismatches in architectural style, construction materials, platform design, environmental context and architectural details.

Asked to compare Fig. 4 (produced by archaeologists from a full 3D rendered model), Fig 5B (generated by ArchGPT from a detailed text prompt) and Fig. 6A (generated by Grok v. 3 using the same text prompt as Fig 5B), PerplexityAI concluides that Fig. 6A represents with higher precission the architecture of neolithic dwellings from la Draga Archaeological site, because its elevated structure (platform), constructive materials, form and environmental context coincide with archaeologically documents find in the internet, and it quotes some internet pages especifically dedicated to this archaeological site, and quotes Andreaki 2022, as a source for documented archaeological reconstructions.

5. Epistemological issues raised by the use of GenAI in Archaeology.

To *validate* a human or machine- generated image depicting how the past may have looked like, we should ensure that the recreated scenario accurately represents what happened in the real-world. This corresponds to a Correspondence theory of Truth, which states that a statement is true if it corresponds to the way the world actually is. In other words, the truth of a particular hypothetical visualization of the past is determined by how well a generated image matches what we have empirical evidence it existed in the Past (Patterson 2003, Rasmussen 2014). This is not the proper place to discuss these epistemological issues, but we should take into account that the evidence from the past existing and empirically testable in the Present, is small and it is usually altered.

In most cases, validation is a human-based task. Experts look for matches between the image and the list of archaeological elements preserved from the past. Different quantifications of these correspondence are available, in form of *scales of uncertainty*, cf. Nicolucci and Hermon 2010, Dell'Unto et al., 2013, Ortiz-Cordero et al., 2018, Cazzaro 2023, Apollonio et al., 2024, Foschi et al., 2024), or as an Extended Matrix (Demetrescu, and Ferdani, 2021). In this way, "authenticity" can be ddefined in terms of the correspondence between the visual model and what is known about the real world. However, this correspondence is always partial, and potentially biased, because what the evidence we have about what happened in the past is also partial, incomplete and potentially distorted. To the extent that the 'expert' selects the part of the preserved past evidence that interests him or her, this evidence may appear also biased.

Such limitations open a diversity of *possible alternative pasts,* with different degrees of "possibility" and even of falsehood. *Possible worlds* (Nolt 1986, Stalnaker 1986, Dubois and Prade 2023, Chalmers 2022) is a philosophical framework that can be used to analyze and evaluate the truth-likeness of archaeological illustrations. In this context, each generated image -by the human expert or by the machine- represents a *possible* Past. We assume that once existed a "true past", something having occurred years ago, which is unknown due to the incomplete and altered nature of the archaeological record as seen in the present.

Within the spectrum of generated pasts, at the extreme, would be what we rightly know about what happened: the archaeological record. The *possible* includes all predictions for which some elements can be related to this observable material. In that sense, a recreated image of the past will be "possible" when it is consistent with the archaeological data available. In the realm of the possible, we may identify a first stage, at the limit of the possible, which we can measure as improbable -using standard Probability Theory. A second stage, which enters the realm of the possible, is called *plausible*. This stage applies when it is determined that there are good reasons for something to exist, but no real example is known (Carlevaris et al. 2024).

On the other extreme there is the realm of the *impossible*. It refers to the group of responses that have no correspondence with the prediction expectation. In that way, impossible pasts refer to those hypothetical images that include some erroneous or equivocal aspect that makes it infeasible to evaluate the outcome within a spectrum of credible predictions. In this domain, we can distinguish the implausible and the highly improbable. The implausible is identified by the arbitrariness of the answer. The highly improbable is identified by the high difficulty that a given prediction can exist (Carlevaris et al., 2024).

The degree of "possibility" of a recreated past is then a function of the "consistency" of the image with preserved remains from that "real" past. Troubles arise when we try to establish this *consistency*. It refers to a property of a formal system or set of statements where no contradictions arise (Reiter and Mackworth 1989, Ishitani 1999, Zhang et al., 2024). In that sense, a particular visualization of the past is consistent with available archaeological data if it is logically possible for all the individual elements in the image to exist also in the archaeological record. This means there is no logical contradiction within the set. We can distinguish:

- *Syntactic consistency* focuses on the derivability of contradictions within the image.
- *Semantic consistency* (or satisfiability) means that there exists at least one interpretation (or model) under which all elements depicted in the image are true -exist in the archaeological record with the same visual features.

In the context of a GenAI-produced visual reconstruction of a prehistoric house, *verifying consistency* can be addressed to ensure that the generated image adheres to the explicit constraints of the text prompt and technical specifications. In that sense, we should look for prompt compliance, checking if requested features (e.g., "rectangular floor plan," "wattle-and-daub walls") are present. In that sense, we can say that current universal GenAI programs can effectively verify (partially) the possibility of a generated image because they can kind matches and mismatches between the image and the text prompt. Validation is a different issue, because the GenAI app depends on the information it is able to gather from internet about the "real world", that is, the available archaeological evidence. And access to this information is not exhaustive given copyright issues and accessibility.

While by *verifying consistency* we can ensure that the model has been generated correctly (it is deducible), we are far from validating it, because we do not know if the "consistent" model corresponds with what existed years, centuries, millennia ago. In other words, verifying consistency does not confirm that the model's structure, assumptions, or parameters are appropriate for the real-world context. Possibility focusses on the technical adherence to the provided input information (text prompt), while validation goes beyond the particular information for that image and looks for a more general historical or scientific accuracy, as the absence of anachronisms and structural consistency measures that cannot be calculated on a mere 2D picture.

6. Constraint based Reasoning

The images that humans and machines can generate depicting what we may believe about the past will never be entirely true or fully validated, but neither should they be entirely imaginative. At least, they should should be consistent with information used to generate them as formally and logically as possible, and then such visualizations may become *possible*, and thir degree of possibility estimated in quantitative (probabilistic) terms.

For this purpose, we suggest to follow a Constrain Based Reasoning approach, a paradigm in artificial intelligence that focuses on solving problems by expressing them as a set of constraints that limit the possible values variables can take (Freuder and Mackworth 1994, Diligenti et al. 2021).

Intuitively, constraint-based reasoning is like solving a puzzle where you have a set of rules (constraints) that must be followed. Instead of explicitly programming step-by-step instructions for finding a solution, you simply declare what properties the solution must satisfy. The system then works to find values that meet all these requirements.

When the generated visual representation is a full 3D geometrical model, verification can be approached evaluating the deducibility of the final model from initial premises –measurements taken on the field-, and checking the geometric completion models used at each step of the reconstruction process. In the case of the 3D model created by archaeologists using computational 3D design tools (Fig. 4), consistency verification is approached at each step, from the creation of an initial raw model expressing the geometrical and spatial information directly extracted from the archaeological record, to the progressive geometric completion of different elements that constitute the reconstructed model: poles, boards, platform, rafters, ridge beams and roof covering. The resulting computer reconstructed image can be deduced from archaeological elements by reversing the logic used by human experts when completing the incomplete geometry, and placing the correct element at its most probable location based on structural constraints and architectural principles (Campana 2016, Barceló et al., 2019, Tzerpou et al. 2025).

Generative AI image production from text prompts exemplifies constraint-based reasoning through the way it transforms textual descriptions into visual outputs while adhering to multiple explicit and implicit constraints. The textual description serves as a set of semantic constraints that the model must satisfy. Each word or phrase in the prompt establishes requirements that limit the possible visual outputs:

- Nouns constrain what objects must appear in the image
- Adjectives constrain the attributes of those objects
- Spatial language (e.g., "left of," "above") constrains the relative positioning of elements
- Style descriptors constrain the aesthetic qualities

These textual elements function similarly to the formal constraints in constraint satisfaction problems, where the system must find values (visual elements) that satisfy all constraints simultaneously.

The variation of the classical spatial resource allocation approach we have used in this paper to “automatically generate” possible visualizations of the Neolithic houses from La Draga site, is an example of constrain based reasoning. The idea was to challenge the AI model to distribute verbally described architectural elements across well determined spatial dimensions to optimize the final visualization while adhering to spatial and architectural constraints.

This allocation process typically involves complex decision-making environments, specification descriptions, interleaved spatial structures, and scale inhibitory effects that make finding high-quality solutions challenging (Shirabe 2005, Papadimitriou 2025). Beyond explicit spatial constraints, text-to-image systems may use implicit constraints related to image quality and coherence:

- Perceptual constraints: Perceptual loss functions are designed to capture perceptual differences between images that are not evident at the pixel level, focusing on content and style discrepancies that better align with human visual perception. They can be used to reduce the randomness involved in the image generation and ensuring semantic consistency between modified and unmodified elements (Cai et al., 2024).
- Aesthetic constraints: We should add to the text prompt explicit constraints related to visual coherence, realistic lighting, and proper perspective.
- Structural constraints: The text prompt should contain the explicit mention to the physical laws and structural integrity principles that will make generated images appear realistic and plausible.

What makes constraint-based reasoning powerful is that it thrives on restrictions - the more constraints there are, the faster a solution can be found, as each constraint helps eliminate impossible alternatives. This is contrary to conventional systems that slow down as complexity increases. The challenge in text-to-image generation increases with the number of constraints. The more constraints we set, the more exact our expectations become in relation to the images we want to generate. The higher the expectation, the harder the fall (Netherwood 2025). This reflects the fundamental challenge in constraint-based reasoning of finding solutions that satisfy multiple, potentially competing constraints.

We suggest to follow a very formal structure to the information provided to the human designer or the machine, in such a way that the text prompt defines the "solution space" for the visual model, using *explicit requirements*, like construction materials (e.g., "wattle-and-daub walls"), structural features (e.g., "thatched roof"), or cultural motifs (e.g., "linear pottery patterns"). Additionally, there is must introduce some form of *negative constraints*, like excluding anachronisms (e.g., "no metal tools") or undesired elements (e.g., "no modern lighting").

This approach reveals how text prompts may act as non-holonomic constraints in AI image generation – limiting some freedoms while permitting others, guided by both user intent and model architecture (Marra et al., 2019, Li et al., 2024).

If we define *degrees of freedom* as the number of independent parameters that can vary while satisfying constraints, then the number of constraints in the text prompt can be used as the number of parameters whose variation has to be taken into account. Consequently, we would need to distinguish between *explicit constraints* -like "rectangular Neolithic house with thatched roof"- and *implicit constraints* arising from model architecture (e.g., culturally mandated east-west orientations limiting rotational freedom, resource availability restricting material diversity or symbolic motifs (e.g., spiral patterns) reducing decorative variability. In some cases, however, negative prompting –a phrases like "no decorative carvings"- paradoxically may increase freedom by preventing over-constraint reactance (Netherwood 2025). Stylistic constraints (parameters) can also be included.

A good verification would quantify the degrees of freedom of the resulting visualization, using a variation of this simple algorithm:

1. Generate 100+ house variants using the same prompt/seed.
2. Cluster outputs by features (size, materials, layout) using an image classification tool
3. Calculate the Shannon entropy of each feature cluster to measure variability.
4. Compare against a baseline prompt with fewer constraints (e.g., "Neolithic house").

A high entropy score in roof-type variability but low score in wall materials, for instance, would allow to distinguish the different constraining power of diverse constraint

7. Conclusions

Contrary to the usual negative critique towards GenAI, we suggest that the way of generating images is fairly similar both by the human agent and the machine. In both cases, there is an external information that should be fitted *constraining* the final appearance of the visual reconstructition.

We have noted a relevant scarcity of methodological frameworks specifically designed to address transparency and truthfulness issues in all kinds of visualizations of apparently hypothetical historical scenarios. This affects both GenAI products but also human designed drawings, and computer-based rendered 3D geometrical models. We propose that hybrid approaches combining the computational power of AI with human expertise, coupled with rigorous documentation protocols and uncertainty visualization techniques, offer promising pathways forward.

By explicitly acknowledging the *generative* nature of any hypothetical visualization of the past, the field can enhance both the scientific value and ethical integrity of prehistoric building reconstructions, regardless of whether they are generated through artificial or human intelligence. “Generative” should not be equated with “interpretive” or “necessary subjective”. In computer science, *generative* refers to systems, models, or algorithms that can create new, original outputs based on learned patterns, rules, or data. Generative systems are not limited to reproducing existing data; instead, they synthesize new content—such as text, images, audio, or even code—often in response to user prompts or other inputs. In a similar way, image-generation, by humans and machines, is in a way *adaptive*, what in computer science means changing or optimizing behavior in response to feedback or new information, ensuring continued effectiveness in dynamic or uncertain conditions. An adaptive mechanism of image generation can modify its operations based on feedback, user preferences, or changing external factors, aiming to maintain or improve performance in dynamic environment. *Adaptivity* in image creation is crucial for archaeological applications where static, fixed rules would fail to cope with variability, that must respond to changing input information.

Generativity nor adaptivity cannot avoid the risk of creating fake or deceptive content. This content can be used to spread misleading information. In the same way as human designers, GenAI programs may misrepresent the imagined past by integrating of highly inaccurate elements with more plausible details, making it impossible for viewers to distinguish between data-led conclusions and speculative elements. There are many options to open the apparent "black box" nature of many AI systems allowing tracing the decision-making process that leads to specific reconstructions, and opening the possibility of explicating constraints that may be used to reduce subjectivity and bias. It is crucial to acknowledge that traditional human-based reconstruction methodologies suffer from similar limitations. Human experts inevitably incorporate subjective interpretations, unconscious biases, and theoretical frameworks that influence their reconstructions, often with limited documentation of these interpretive processes.

Verifying the logical possibility of visual recreations of prehistoric environments requires a systematic approach to backwards reasoning, whether these recreations are human-generated or AI-produced. This process involves working backward from available evidence to establish plausible reconstructions while acknowledging inherent limitations. For human-generated visual recreations, backwards reasoning begins with archaeological evidence as the foundation. Archaeologists must critically examine spatial properties of well identified elements to establish constraints on their reconstructions. This process involves retrodiction—predicting the past based on present observations using uniformitarian principles (Campbell et al., 2021). However, human reconstructions inevitably incorporate subjective interpretations and theoretical frameworks that influence outcomes. These variations suggest that reconstructions frequently represent artistic interpretations based on individually constructed knowledge rather than scientifically verified models. To verify logical possibility, human reconstructions should explicitly distinguish between evidence-based conclusions and interpretive extrapolation, documenting the decision-making process that leads to specific visual choices (Altaweel and Khelifi 2024).

The theoretical framework proposed here envisions a future where generative AI and human expertise work in symbiosis to advance our understanding of prehistory. By separating the roles between archaeologists and computers we introduce the need to communicate with the machine through a formal language, instead of verbally rich text prompts, a more formal way to describe the input and introduce constraints with design logic concise and unambiguous. The effort of defining the input in ways that computers can understand generates the opportunity for reflection on the complexity of the problem and how to solve it.

FINAL NOTE

To write this paper we have used and analysed the internal working of diferent Generative Artificial Tools (OpenAI’s ArchGPT and chatGPT, Stability AI’s Stable Diffusion, xAI’s Grok, MistralAI’s LeChat, Sloyd.ai, MasterpieceX, Perplexity AI and Gemini. In all cases we have used the free online versions of those tools, except in the case of PerplexityPro, because it gave us access to other tools like DeepSeek, ClaudeAI and even the possibility of using PlayGround and Flex, which cannot be used to create images in its free version.

References

Alimardani, A., & Jane, E. (2024). Human Hallucinations: GenAI and the Mirage of Personalised Learning for All. *Available at SSRN 5016289*.

Andreaki, V., 2022, *"Right on time": Reconstructing the biography of the lacustrine site of La Draga (lake Banyoles, Spain). Micromorphological analysis and chronological modelling of depositional events.* PhD. Dissertation. Universitat Autònoma de Barcelona. https://www.tdx.cat/handle/10803/675286#page=1

Andreaki, V., Barceló, J. A., Antolín, F., Gassmann, P., Hajdas, I., López-Bultó, O., … Terradas, X. (2022). Absolute chronology at the waterlogged site of La Draga (lake Banyoles, NE Iberia): Bayesian chronological models integrating tree-ring measurement, radiocarbon dates and microstratigraphical data. Radiocarbon, 64(5), 907–948. doi: 10.1017/RDC.2022.56.

Apollonio, F. I., Fallavollita, F., and Foschi, R. (2024). 'Multi-Feature Uncertainty Analysis for Urban Scale Hypothetical 3D Reconstructions: Piazza delle Erbe Case Study', Heritage 2024, 7(1), pp. 476-498. https://doi.org/10.3390/heritage7010023

Asthana, S., & Goel, A. (2024, July). Unveiling Anomalies: A Review of Anomaly Detection Through Lens of Explainable AI. In *2024 Third International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN)* (pp. 1-6). IEEE.

Barceló, J. A., Calvano, M., Campana, I., Piqué, R., Palomo, A., & Bultó, J. O. (2019). Rebuilding the past: 3D reconstruction and BIM analysis of a Neolithic house at La Draga (Girona, Spain). In *Digital Cultural Heritage* (pp. 157-168). Cham: Springer International Publishing.

Bosch, A., Chinchilla, J., Tarrús, J., 2018, *El Poblat Lacustre Neolític Antic de La Draga.* Girona: Centre Arqueologia Subaquàtica de Catalunya - Museu d'Arqueologia de Catalunya, Monografies del CASC, No. 2.

Cai, X., Wang, G., Lou, J., Jian, M., Dong, J., Chen, R. C., Stevens, B., & Yu, H. (2024). Perceptual loss guided Generative adversarial network for saliency detection. *Information Sciences*, *654*, 119625.

Campana, I. (2018). Prehistoric house and 3D reconstruction: Towards a BIM archaeology. PhD. Dissertation. Universitat Autònoma de Barcelona. https://ddd.uab.cat/pub/tesis/2018/hdl_10803_666054/ivca1de1.pdf

Carlevaris, L., Delgado-Martos, E., Intra Sidola, G., Maitín, A. M., Nogales, A., Pesqueira-Calvo, C., Bravo, M., & García Tejedor, Á. J. (2024). Between Impossible and Probable. Architectural Recognition Through Qualitative Evaluation of Artificial Intelligence Response. In *Advances in Representation: New AI-and XR-Driven Transdisciplinarity* (pp. 839-850). Cham: Springer Nature Switzerland.

Cazzaro, I. (2023). Digital 3D reconstruction as a research environment in art and architecture history: uncertainty classification and visualisation. PhD thesis, Alma Mater Studiorum - Università di Bologna. CoVHer Project (2022). Official website (n.d.). Available at: www.CoVHer.eu (Accessed: 27 February 2024).

Chalmers, D. J. (2022). *Reality+: Virtual worlds and the problems of philosophy*. Penguin UK.

Couret, N. (2024). Speculative Historiography in the Age of Hallucinations. *The Journal of Cinema and Media Studies*, *63*(8).

Dell'Unto, N., Leander, A., Dellepiane, M., Callieri, M., Ferdani, D., and Lindgren, S. (2013). 'Digital Reconstruction and Visualisation in Archaeology: Case-Study Drawn from the Work of the Swedish Pompeii Project'. Proceedings of the Digital Heritage International Congress 2013, 1, pp. 621–28, IEEE. Available at: https://www.researchgate.net/publication/267323246_Digital_reconstruction_and_visualisation_in_archaeology_Case-study_drawn_from_the_work_of_the_Swedish_Pompeii_Project#fullTextFileContent (Accessed: 31/01/2024).

Demetrescu, E., and Ferdani, D. (2021). 'From Field Archaeology to Virtual Reconstruction: A Five Steps Method Using the Extended Matrix'. Appl. Sci., 11, 5206. DOI: https://doi.org/10.3390/app11115206.

Diligenti, M., Giannini, F., Gori, M., Maggini, M., & Marra, G. (2021). A constraint-based approach to learning and reasoning. In *Neuro-symbolic artificial intelligence: the state of the art* (pp. 192-213). IOS Press

Dubois, D., & Prade, H. (2023). Possibility theory. In *Granular, Fuzzy, and Soft Computing* (pp. 859-876). New York, NY: Springer US.

Foschi, R., Fallavollita, F., & Apollonio, F. I. (2024). Quantifying Uncertainty in Hypothetical 3D Reconstruction—A User-Independent Methodology for the Calculation of Average Uncertainty. *Heritage*, *7*(8), 4440-4454.

Freuder, E. C., & Mackworth, A. K., (1994). *Constraint-based reasoning* (Vol. 58, No. 1-3). MIT press.

Ishitani, Y. (1999, September). Logical structure analysis of document images based on emergent computation. In *Proceedings of the Fifth International Conference on Document Analysis and Recognition. ICDAR'99 (Cat. No. PR00318)* (pp. 189-192). IEEELi et al., 2024

Loch-Dehbi, S., & Plümer, L. (2009). Geometric reasoning in 3D building models using multivariate polynomials and characteristic sets. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *38*(Part 2/W11).

Marra, G., Giannini, F., Diligenti, M., & Gori, M. (2019). Constraint-based visual generation. In *Artificial Neural Networks and Machine Learning–ICANN 2019: Image Processing: 28th International Conference on Artificial Neural Networks, Munich, Germany, September 17–19, 2019, Proceedings, Part III 28* (pp. 565-577). Springer International Publishing.

Martins Lemos, A. L. (2024). Digital errors, failures, and disruptions in generative AI hallucinations: Communication typology, premises, and epistemology. *Matrizes*, *18*(1).

Netherwood, S., 2025, *How to generate high-quality AI images within constraints*. In Tomoro.AI (https://tomoro.ai/insights/how-to-generate-high-quality-ai-images-within-constraints). Downloaded on APril, 9th., 2025).

Nicolucci, F., and Hermon, S. (2010). 'A Fuzzy Logic Approach to Reliability in Archaeological Virtual Reconstruction'. In Beyond the Artifact. Digital Interpretation of the Past; Nicolucci, F., Hermon S. Eds.; Archaeolingua: Budapest, Hungary, pp. 28–35. Proceedings of CAA2004, Prato, Italy, 13-17 April 2004. Available online: https://proceedings.caaconference.org/paper/03_niccolucci_hermon_caa_2004/ (Accessed on: 04/07/2024).

Nolt, J. E. (1986). What are possible worlds?. *Mind*, *95*(380), 432-445.

Ortiz-Cordero, R., Pastor, E. L., and Fernández, R. E. H. (2018). 'Proposal for the Improvement and Modification in the Scale of Evidence for Virtual Reconstruction of the Cultural Heritage: A First Approach in the Mosque-Cathedral and the Fluvial Landscape of Cordoba'. Journal of Cultural Heritage, 30, pp. 10-15. Available at: https://www-sciencedirect-com.ezproxy.unibo.it/science/article/pii/S1296207417303771?via%3Dihub (Accessed: 31/01/2024).

Palomo, A., Piqué, R., Terradas, A., Bosch, A., Buxó, R., Chinchilla, J., … Tarrús, J. (2014). Prehistoric occupation of Banyoles lakeshore: Results of recent excavations at La Draga site, Girona, Spain. Journal of Wetland Archaeology, 14(1), 58–73. doi: 10.1179/1473297114Z.00000000010.

Papadimitriou, F. (2025). Spatial AI for Artificial General Intelligence. In *Spatial Artificial Intelligence* (pp. 53-63). Cham: Springer Nature Switzerland.

Patterson, D. (2003). What is a correspondence theory of truth?. *Synthese*, *137*, 421-444.

Piqué, R., Palomo, A., Terradas, X., Andreaki, V., Barceló, J.A., Bogdanovic, I., Lloret, À.B., Gassman, P., López-Bultó, O. and Turra, R.R., 2021. Models of neolithisation of northeastern Iberian Peninsula: new evidence of human occupations during the sixth millennium cal BC. *Open Archaeology*, *7*(1), pp.671-689.

Rasmussen, J. (2014). *Defending the correspondence theory of truth*. Cambridge University Press.

Reiter, R., & Mackworth, A. K. (1989). A logical framework for depiction and image interpretation. *Artificial Intelligence*, *41*(2), 125-155.

Shirabe, T. (2005). Classification of spatial properties for spatial allocation modeling. *GeoInformatica*, *9*, 269-287.

Stalnaker, R. (1986). Possible worlds and situations. *Journal of Philosophical Logic*, *15*, 109-123.

Tarrus (2008). La Draga (Banyoles, Catalonia), an Early Neolithic Lakeside Village in Mediterranean Europe. CATALAN HISTORICAL REVIEW, 1: 17-33. DOI: 10.2436/20.1000.01.2

Tzerpou, E., Campana, I., Salazar, J., López-Bultó, O., Palomo, A., Piqué, R., ... & Barceló, J. A. (2024). 3D hypothetical reconstruction of a Neolithic hut from the waterlogged site of La Draga (Banyoles, Spain). *Virtual Archaeology Review*.

Zhang, J., Xiang, R., Kuang, Z., Wang, B., & Li, Y. (2024). ArchGPT: harnessing large language models for supporting renovation and conservation of traditional architectural heritage. *Heritage Science*, *12*(1), 220.