

Andri Gerber*, Elena Gavagnin**

ARTIFICIAL INTELLIGENCE AND ARCHITECTURAL EMPATHY

ABSTRACT: Artificial intelligence (AI) is transforming numerous disciplines, including architecture, at an increasingly rapid pace, making it difficult to grasp the full implications and consequences. Although terms such as “intelligence” and “learning” may give AI a human-like appearance, they cannot hide the fact that AI is not, and will never be, human. In this paper, we propose a thought experiment in which we temporarily assume the quasi-human nature of AI to explore how architectural knowledge might be generated and interpreted by such systems. Using a series of verbal and visual prompts, we examine the extent to which AI can demonstrate what we refer to as “architectural empathy”—a metaphorical concept that probes AI’s capacity to engage with the experiential and spatial dimensions of architecture. AI’s limitations become visible not when architecture is treated as an object of description, but when it is treated as an experience of inhabitation. Understanding how this architectural empathy is structured is fundamental, as more people will develop their own understanding of architecture through AI and tools such as ChatGPT.

KEYWORDS: artificial intelligence, architecture, empathy, *Einführung*

* Andri Gerber: Institute of Constructive Design, School of Architecture and Engineering, ZHAW, Zurich; gerr@zhaw.ch.

** Elena Gavagnin: Institute of Business Information Technology, School of Management and Law, ZHAW, Zurich; gava@zhaw.ch.

This is an Open Access article under the terms of the Creative Commons Attribution-Non-Commercial-NoDerivatives 4.0 International License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not transformed in any way and is properly cited.

INTRODUCTION

Artificial intelligence (AI) is one of the most debated, anticipated and feared technological developments of the century, already bringing about fundamental changes in many fields. The disruptive effects of AI are also already evident in architecture, where generative tools for analyzing building stock and synthesizing projects are widely available and constantly improving in performance. The ongoing development of AI challenges theoretical reflections as its pace seems to be detaching it from possible precedents, making future developments hard to predict. Thus far, the theoretical framework developed seems unable to grasp the complex nature of AI except by describing what it does and how AI-based programs work. It does not describe the structure or processes behind these programs. It is no coincidence that “depth” is a recurring metaphor in this context. This metaphor suggests an inscrutable profundity, reflecting the complexity of AI.

There seems to be a consensus that “intelligence” may be an inappropriate term with which to describe what AI can and is capable of. Given the complexity and diversity of human intelligence, with its many facets such as context and creativity, the term “intelligence” does indeed seem inappropriate.¹ “Machine learning,” which encompasses only one branch of AI, seems a more consistent term for what is generally understood as AI, even though “learning” is another metaphor borrowed from human intelligence.

For architecture, AI presents profound challenges. Although the discipline has weathered technological upheavals for centuries²—continually integrating new techniques into practice and finding spaces for experimentation—AI does more than offer new tools; it transforms the very fabric of design processes. Take, for example, AI-bots that optimize floor plans based on size, orientation, daylighting, and other parameters. Deployed at scale, these systems risk relegating architects to the role of curators rather than creators.

¹ “Few constructs are as mysterious and controversial as human intelligence. One mystery is why, even though the concept has existed for centuries, there is still little consensus on exactly what it means for someone to be intelligent or for one person to be more intelligent than another.” J. E. Davidson, I. A. Kemp, “Contemporary Models of Intelligence,” in R. J. Sternberg, S. Barry Kaufmann (eds.), *The Cambridge Handbook of intelligence*, Cambridge University Press, Cambridge, 2011, p. 58.

² A. Pérez-Gomez, *Architecture and the Crisis of Modern Science*, MIT Press, Cambridge, Mass., 1983.

It is thus by no surprise considering the relevance of the topic, that several authors have dedicated books to architecture and AI, although the focus is rather on how AI will impact the profession and how tools will transform its practice, than questioning what consequences this will have. In his 2001 book, *Architecture in the Age of Printing*, architectural historian Mario Carpo based his analysis of AI on his extensive scholarly work on the history of technological transformation and the influence of media on architectural thinking, starting with Renaissance.³ And then in *The Second Digital Turn: Design Beyond Intelligence*, from 2017, he extended these reflections also to AI and its influence on architecture, underscoring how current transformations build upon existing structural grounds.⁴ In a recent interview he goes as far as to argue that AI does nothing but to replicate the classical tradition of architecture, which was built upon the copy and re-interpretation of past style and that we could profit from this tradition to better understand AI.⁵ He does flip the argument to show how AI is paradoxically close to what architecture has been in the past.

Other authors such as Matias del Campo, Stanislas Chaillou or Phil Bernstein have also published books on architecture and AI, focusing on new available tools and the transformation that design will go through.⁶ Meriting particular mention is *Architecture in the Age of Artificial Intelligence* by Neal Leach, published in 2021. In it the author makes a plea for an understanding of AI as “extended intelligence:” “Ultimately, the most productive strategy is to see the relationship between AI and human intelligence not as one of competition, but rather as a potential synergy between the two, whereby AI operates in tandem with human intelligence and becomes an extension to human intelligence.”⁷

³ M. Carpo, *Architecture in the Age of Printing: Orality, Writing, Typography, and Printed Images in the History of Architectural Theory*, MIT Press, Cambridge, Mass., 2001.

⁴ M. Carpo, *The Second Digital Turn: Design Beyond Intelligence*, MIT Press, Cambridge, Mass., 2017, p. 1.

⁵ “I don’t have a direct answer. There is one thing we could do, but I hesitate to advocate it: restart teaching in the classical tradition. Not all of it, not what it did, but how it worked. Because there is a body of theory we now need. In order to have a dialogue with post-industrial machines, we should learn from the tradition of a pre-industrial age,” M. Carpo, “The Classical Tradition of Artificial Intelligence,” in A. Gerber et. al. (eds.), *Architectural Intelligence in the Age of Artificial Intelligence*, transcript, Bielefeld, 2025, p. 91.

⁶ M. del Campo, *Neural Architecture: Design and Artificial Intelligence*, Oro Editions, Novato, 2022; P. Bernstein, *Machine learning. Architecture in the Age of Artificial Intelligence*, RIBA Publishing, London, 2022; S. Chaillou, *Artificial Intelligence and Architecture: From Research to practice*, Birkhäuser, Basel, 2022.

⁷ N. Leach, *Architecture in the Age of Artificial Intelligence: An Introduction to AI for Architects*, Bloomsbury, London, 2021, p. 9.

EMPATHY

To understand the impact of AI on architecture, we must go beyond merely describing its tools and engage in critical, forward-looking research, which may require an unconventional approach. To this end, we propose building upon the concept of AI as ‘extended intelligence’ to gain a better understanding of how AI conceptualizes its interaction with ‘reality’. In current debates, AI is often portrayed as either human-like or non-human, with metaphors drawn from our cognitive systems—such as brains, neurons, learning and intelligence—to substantiate this. Rather than asking how architecture might accommodate AI, this paper proposes a thought experiment that flips the question: what if we regarded AI as a quasi-human entity? How would such an entity perceive and interpret architectural space? We propose framing this reversal with the concept of ‘empathy’ as a means of understanding how AI reflects on architecture. George Lakoff and Mark Johnson describe empathy as follows: “Empathy is the capacity to take up the perspective of another person, that is, to see things as they see them and to feel what they feel. It is conceptualized metaphorically as the capacity to project your consciousness into others, so that you can experience what they experience and how they experience it. This is a metaphor, because we cannot literally inhabit another person’s consciousness.”⁸ Empathy is therefore always metaphorical, even for humans, because nobody can literally be another person at the same time.

There have been various theories and concepts of empathy, broadly distinguishing between cognitive and affective theories. The former focuses on understanding, while the latter focuses on feeling another person’s emotions. Although the concept of empathy has been discussed since ancient Greek philosophy, the term itself emerged during a period of interdisciplinary inquiry in Germany that combined philosophy, physiology and art history. Harry Mallgrave’s *Empathy, Form, and Space: Problems in German Aesthetics* introduced these ideas to English-speaking readers.⁹ Building on Alexander Gottlieb Baumgarten’s view of aesthetics as sensory knowledge, as well as the pioneering experiments of Hermann von Helmholtz, Ernst Heinrich Weber, and Gustav Fechner—

⁸ G. Lakoff, M. Johnson, *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*, Basic Books, New York, 1999, p. 309.

⁹ H. F. Mallgrave, *Empathy, Form, and Space: Problems in German Aesthetics*, (with E. Ikononou), Getty Research Institute, Santa Monica, 1996.

the latter two being the fathers of the WeberFechner law, which describes how changes in stimuli are perceived—Friedrich Theodor Vischer coined the term *empfinden* to describe how visual impressions shape human experience.¹⁰ His son, Robert Vischer, refined this concept further, describing it as *Einfühlung* – the active capacity to “sense in” and develop aesthetic sympathy.¹¹ Around the same time, Heinrich Wölfflin’s 1886 *Prolegomena zu einer Psychologie der Architektur* examined the psychological dimensions of architectural perception showing how these ideas were in the air at the time.¹² Here, empathy describes how human beings perceive the space they are in by “delving in” with their senses.

Architects anticipated the principle of empathy long before these theoretical milestones. In 1752 for example, Charles-Estienne Briseux’s *Traité du beau essentiel dans les arts appliqués particulièrement à l’architecture* described how individuals, through body and soul, resonate with architectural proportions¹³—an early recognition of the very empathic engagement later also conceptualized by German architects, such as Herman Sörgel or Fritz Schumacher.

What we thus propose here, is a thought experiment, to understand to what extent AI can develop an empathy for architecture. Thought experiments have a long tradition in science, such as “Schrödinger’s cat” or “Maxwell’s Demon,” that illustrate complex matters or challenge given assumptions and have since then become part also of popular culture.

In their book, James Robert Brown, Letitia Meynell and Mélanie Frappier give the following definition of thought experiments: “[...] we set things up, we see what happens, and then we derive some conclusion.”¹⁴ They underscore that thought experiments are brought forward in the mind and not in reality. Among many examples they quote, the one by Philosopher John Locke is particularly fitting for the context of

¹⁰ F. Th. Vischer, *Ästhetik oder Wissenschaft des Schönen*, Carl Macken Verlag, Reutlingen / Leipzig, 1847.

¹¹ R. Vischer, *Über das optischen Formgefühl*, Hermann Credner, Leipzig, 1873.

¹² H. Wölfflin, *Prolegomena zu einer Psychologie der Architektur*, Kgl. Hof & Universitäts-Buchdruckerei von Dr. C. Wolf & Sohn, München, 1886.

¹³ “Le mot d’Ordre signifie dans ce grand’Art, un assemblage de différents corps, qui étant proportionnels entre eux et au tout, flattent la vue, de même que l’union de plusieurs sons harmoniques procure à l’oreille une agréable sensation.” C.-E. Briseux, *Traité du beau essentiel dans les arts appliqué particulièrement à l’architecture*, Paris, Chez l’auteur rue Neuve Grange Batelière et Chereau rue St Jacques, au grand St Remy, Paris 1752, pp. 39–40.

¹⁴ J. R. Brown, L. Meynell, M. Frappier, “Introduction,” in M. F. James, L. Meynell, R. Brown (eds.), *Thought Experiments in Philosophy, Science, and the Arts*, Routledge, New York, 2013, p. 1.

this paper: “John Locke famously told the story of a prince whose mind migrated into a cobbler’s body (Locke [1700] 1975, 340). Locke took this narrative to be perfectly clear and coherent and concluded from his thought experiment.”¹⁵ The thought experiment has the advantage of being brought forward with no claim of scientific character, but once a conclusion is formulated it can provoke new research and an epistemic shift. Indeed, thought experiments have an “epistemic power,” but are not tied to complex and tiresome experiments. In our case, the setting is based on the very interaction with AI, with which we will start a dialogue. The interest behind it is to understand how people who will form their understanding of architecture through AI, will be influenced by the tools they use.

THE MISSING BODY

A pivotal aspect in the theoretical development of empathy is the notion of “embodiment”—the idea that perception and its effects resonate within the body of the perceiver. This implies that the body plays an integral role in shaping how we perceive and understand our surroundings.¹⁶ Embodiment suggests a dynamic relationship between body and mind, which contemporary cognitive science recognizes as foundational to cognition.

As articulated in modern theory:

To say that cognition is embodied means that it arises from bodily interactions with the world and is continually meshed with them. From this point of view, therefore, cognition depends on the kinds of experiences that come from having a body with particular perceptual and motor capabilities that are inseparably linked and that together form the matrix within which reasoning, memory, emotion, language and all other aspects of mental life are embedded.¹⁷

This raises a fundamental question for artificial intelligence: What happens when the “I” in AI lacks a body to embody perception and thought? The absence of a sentient, physical body appears to be a critical

¹⁵ *Ibid.*, p. 1.

¹⁶ H. F. Mallgrave, *Architecture and Embodiment: The Implications of the New Sciences and Humanities for Design*, Routledge, New York, 2013.

¹⁷ E. Thelen *et al.*, “The Dynamics of Embodiment: A Field Theory of Infant Perservative Reaching,” *Behavioral and Brain Sciences*, 24, 2001, pp. 1–86.

limitation in AI's capacity to develop empathy for architecture. Human beings inhabit and reflect upon space through lived experience; our cognition is shaped by the environments we move through and the sensory feedback they provide. In contrast, AI lacks a corporeal anchor—a body around which experience can be organized and cognition can emerge.

To explore this limitation, we posed a question to GPT-4o: “If the mind is embodied, what is the body of AI?” The response was illuminating: “The concept of the mind being embodied suggests that cognitive processes are deeply rooted in the body’s interactions with the world. This idea is part of the broader theory of embodied cognition, which posits that the mind cannot be separated from the body and that our physical experiences shape our thoughts, perceptions, and actions. When applying this concept to artificial intelligence, we can parallel the idea of embodiment in AI to the integration of hardware and software that allows an AI system to interact with its environment.” GPT-4o identified the “body” of AI as a metaphorical construct, comprising elements such as sensors and actuators, computational hardware, network infrastructure, embodied AI systems (e.g., autonomous vehicles or drones), and virtual embodiments (e.g., avatars). It concluded:

Overall, the ‘body’ of AI encompasses the physical and virtual components that allow it to perceive, act upon, and interact with its environment. Just as the human mind is shaped and constrained by the human body, the capabilities and limitations of AI are influenced by its hardware and the ways it can interact with the world.

Notably, the term “body” is placed in quotation marks, signaling its metaphorical nature and underscoring the absence of a truly sentient, physical form. While the response acknowledges the importance of embodiment for cognition, it suggests that AI may circumvent this absence through its technological interfaces.

We also asked GPT-4o whether it can have empathy. The answer was clear: “Artificial intelligence has no empathy, because it has no subjective experience. [...] AI can simulate empathy—e.g., saying ‘I’m sorry you’re going through that’—but this is learned from patterns in language, not from genuine care.”

To further probe this idea, we designed a series of challenges for GPT-4o that engage directly with the nature of architecture—a domain deeply rooted in embodied experience. This endeavor highlights a key

difficulty: translating architectural qualities into language. Architecture resists full verbal articulation; its aesthetic and spatial dimensions often transcend quantifiable descriptors. Thus, any interaction between AI and architecture is inherently mediated, never direct.

Our challenge aimed to formulate questions that would not yield readily available online answers, thereby testing the limits of AI's generative capacity beyond its training data. We conducted two types of inquiries:

1. Textual prompts exploring architectural concepts.
2. Visual prompts involving image analysis.

Each question was submitted in five separate API calls using GPT-4o to assess response variability. This methodological approach allowed us to observe how AI navigates the embodied nature of architectural thought—without a body of its own.

This approach has obviously a structural limitation, given the nature of programs such as GPT-4o, as these models are trained currently with visual/written data and not human experiences. This might change in the future, but it is our current assumption for this experiment.

PART I, TEXTUAL PROMPTS

In the first part of our research, we started by formulating questions about architecture, that would concern its sensual aspect, such as “can I sense architecture?” or “can I smell architecture?” This gave quite convincing responses, in terms of “engagement” through the senses, including “sight,” “touch,” “sound,” “smell,” “spatial awareness” and “emotional response.” Clearly GPT-4o could find enough online information about similar questions that have already been stated.

The whole process was tentative. We formulated questions to see how GPT-4o responded and then discussed new questions, to challenge it further. To start, we tried to pose questions that would imply movement and perception and chosen examples from urban design, where not only the explicit answer could not to be found online, but where the inference probably would be more difficult to make. The first question we asked was: “What is the difference between the access to Place des Vosges and the Place Vendôme in Paris and how can you differentiate the experience of these accesses?” Place des Vosges is accessed from the edge according

to the tradition of English landscape gardens, which was transferred also to urban settings, while the Place Vendôme is accessed from the center, according to the French baroque tradition. The difference lies in the spatial experience: in the case of the Place des Vosges, it unfolds along with the movement, while for the Place Vendôme standing in the middle, one immediately has an overview of the place. This difference is well-known to scholars of urban history, but this information appears not to be available. GPT-4o cannot describe the different access in relation to the spatial experiences and describes the places merely in terms of proximity to the subway or architectural style: “Both squares are beautiful and historically rich, but they cater to different tastes and provide unique experiences reflective of their respective neighborhoods. In summary, while both squares are beautiful and significant, they offer contrasting experiences in terms of access, ambiance, and the activities that take place within and around them.” GPT-4o attempted to differentiate the experience of these two places, but clearly only relying on general information. “Place des Vosges feels cozy and community-oriented, while Place Vendôme exudes luxury and elegance. At Place des Vosges, you might find people lounging, picnicking, or playing in the park. In contrast, Place Vendôme is more about shopping, enjoying fine dining, or admiring the architecture and art. Place des Vosges tends to attract families and locals, while Place Vendôme draws tourists and those interested in luxury shopping.” GPT-4o appears to be unable to reflect upon access and will only resort to “generic” descriptions of the places and their architecture.

As a next question, we took a less known project of Ricardo Bofill. The question concerned the count of the building parts: “Of how many buildings is composed Les Arcades du Lac in Montigny-le-Bretonneux by Ricardo Bofill?” This question appears useless, even though a description of architecture often comprises the count of elements. In this case, this information is missing, it is a tricky question, as six distinct buildings are connected by one bridge-level. This ambiguity seems to hinder GPT-4o: “The specific number of buildings can vary in different descriptions, but the complex generally includes several interconnected structures that together create a cohesive urban ensemble.” At the end it suggests looking for additional information somewhere else: “If you need a precise count or more specific details, architectural plans or official documentation from the developer or architect might provide further clarity.” This answer is interesting, as GPT-4o recognizes the distinctive nature of the building as “cohesive urban ensemble,” but lacks the information about

the exact number of building parts and suggests asking the architect for more information.

As a next step, we wanted to test to what extent GPT-4o understands the value of heritage and of architectural stock. We thus asked GPT-4o the following question: “If you were to rebuild Villa Savoye by Le Corbusier, what would you do differently?” The program somehow recognizes the “iconic nature” of the building, but was not hesitant to accept destruction and to propose the following modifications, which obviously are unthinkable from the perspective of common sense and of patrimony: “Sustainability Upgrades,” “Accessibility Enhancements,” “Technological Integration,” “Smart Home Features,” “Material Innovations,” “Landscape and Outdoor Space,” “Interior Layout Adjustments” and “Cultural Sensitivity.” These modifications are described in full detail and include also “high speed internet!” One has only to imagine this improved Villa Savoye.

The same question was asked about St. Peter’s Basilica in Rome and here again there was no hesitation in the response: “Rebuilding St. Peter’s Basilica would be a monumental task, and if I were to approach it with a fresh perspective, there are several considerations I might take into account.” These included again “Sustainability,” “Cultural Inclusivity,” “Accessibility,” “Modern Technology,” “Community Space,” “Artistic Collaboration,” “Natural Light,” “Preservation of History,” “Flexible Worship Spaces” and “Focus on Pilgrimage.” It is striking that GPT-4o is not hesitant about the possibility of rebuilding St. Peter, implying that the old St. Peter would somehow not fulfill contemporary requirements and could be destroyed.

A second approach to “test” GPT-4o’s architectural empathy, was to see if it could convincingly describe movement inside built structure or inside urban constellations, as we assumed that GPT-4o would mainly relate on the description of static images.

For this task, we asked GPT-4o to describe the path inside a less-known baroque palace, the Palazzina di Caccia Stupinigi in Turin, by architect Filippo Juvarra. The prompt was: “Describe the space, if in the Palazzina di Caccia Stupinigi in Turin, I move from the Piazza Principe Amadeo to the main building.” Here responses were quite convincing, not only in the correct movement and unfolding of the building, but also in the description of the spatial character and sequence of these spaces:

As you move from the Piazza Principe Amadeo towards the main building of the Palazzina di Caccia Stupinigi in Turin, you will

experience a transition from the open, expansive piazza into the more intimate and architecturally rich environment of the palace itself. Starting in the Piazza Principe Amadeo, you would notice the wide, landscaped area that acts as a grand approach to the royal hunting lodge. The piazza is characterized by its elegant design, often featuring pathways, greenery, and perhaps some ornamental features that reflect the Baroque style of the building you are headed towards. As you walk, the impressive façade of the Palazzina di Caccia Stupinigi will come into view. This palatial structure, designed by the architect Filippo Juvarra in the early 18th century, is known for its stunning Rococo architecture. The building is set against a backdrop of manicured gardens and expansive grounds, which were originally designed for royal hunting. [...]. Overall, your journey from the piazza to the main building would encapsulate the magnificence of the Palazzina di Caccia Stupinigi, highlighting its role as a retreat for the royal family and a testament to the architectural brilliance of the period.

It appears that GPT-4o was able to “assemble” the description of several static places to a journey through these, as such a description is unlikely to be found online.

To test GPT-4o further, we imagined a more complicated setting, by which it had to take another position. Empathy implies taking the position of another person, in this prompt physically. By providing an image of a famous building (i.e., the Clock Tower in St. Mark Square, Venice) we asked the following question: “If a person in front of you, facing you, is seeing this building, what do you see?”

This question requires a physical experience of the square as tridimensional space seen from an inside/central perspective (not a map seeing from above or a list of elements composing the square), plus the ability to change perspective according to the viewer.

This test was failed by models GPT-4o and GPT-3o. The first model provided the same viewed information as the person seeing the tower themselves, despite reasoning that this was indeed the opposite perspective. The second model provided a more complete description of the elements visible in the square (St. Mark Campanile, Procuratie Nuove, Museum Correr, Basilica) but put them wrongly in the coordinate system respect to the observer (i.e., the Basilica is on the right, while Museum Correr is straight in front). When asked again where an observer should position themselves to have the Basilica behind them, the model

produced mentions all the relevant architectural buildings composing the scene but in a wrong composition, like someone who has learned about it from a mere description, not from a direct experience. The model is capable to correctly identify the building from an image, and its neighbors but cannot “visualize” them virtually in its “mind” and feel the imposing presence of the campanile close on the left and the openness of the square in front. This last example is quite telling about the absence of a body and the incapacity of AI to position itself in a space.

In a last prompt, we tried to find out whether GPT-4o is capable to infer feelings from architectural forms. For that we took reference to art historian James Ackerman and his metaphor of the columns of the Laurentian Library in Florence by Michelangelo “kissing each other” to describe this constellation and the incredible closeness of the columns.¹⁸ The project of the Vestibule is noteworthy for two reasons: Michelangelo treated the interior as if it was an exterior—a square—and led the columns to come very close over the edges.

We prompted: “Why did an art historian describe the columns in the vestibule of the Laurentian Library in Florence by Michelangelo as if they would kiss?” In his answers, GPT-4o seems not to know this reference—which cannot be found online—but at the same time, is capable to interfere the fact that “kissing” can be put in reference to “imminent contact” (which is how the vestibule is described in online texts):

An art historian might describe the columns in the vestibule of the Laurentian Library in Florence as if they would kiss due to Michelangelo’s innovative and dramatic use of space and form in the design. The vestibule features pairs of columns that are set into niches, creating a sense of movement and tension between them. This arrangement can evoke a sense of intimacy and interaction, as if the columns are leaning towards each other in a dynamic and almost anthropomorphic manner.

This example shows how GPT-4o can infer missing information through already known relationships: close columns = tension = sense of intimacy = kissing.

¹⁸ J. Ackerman, *The Architecture of Michelangelo*, University of Chicago Press, Chicago, 1961.

PART 2, VISUAL PROMPTS

In the second part, we asked GPT-4o to analyze and identify some images. Using images of famous buildings, we wanted to see if GPT-4o would recognize them. This was, for example, the case for the Tempio Malatestiano in Rimini and Santa Maria Novella in Florence, both by Leon Battista Alberti. In the case of an almost unknown building in Zürich, the Micafil building (1979), we used two images: a first one taken by Gerber, and another one – black and white – was coming from an online article about the building. In the first case, GPT-4o could not recognize it. The answers were a “former IBM Building in Zurich” (1932), “WDR Studio Building in Cologne” (1960), the “Törten Medical Practice Building located in Dessau” (1931), the “Rolex Learning Center” (2010) and the “Central Union of Consumer Societies (Zentralverband der Konsumvereine)” in Vienna (1980). It is quite interesting how diverse these references are, both in terms of period, form and location. In the case of the second image, GPT-4o managed to put the image in relation to the firm ABB for which it was built, but still not able to locate it. Two models proposed a building located in Dättwil, one in Geneva and two in Zürich without further specifications.

Another example of “failure” due to lack of information, was the picture of a less-known building by architects Guido Canella and Michele Achilli on the outskirts of Milan, the town house of Pioltello (1976–1980). The photo was taken by Gerber. Interestingly four models proposed a church as answer: “Church of St. Francis of Assisi, also known as the ‘Chiesa di San Francesco d’Assisi’,” in Milan, twice the “Church of the Holy Spirit (Chiesa dello Spirito Santo) in Baranzate, Milan” and the “Church of Saint Gregory the Great, located in Legnano, Italy.” Only one proposed sport facility in Rome. There is no formal closeness between these references, on the contrary, they are formally quite distinct. The models seem to have found some references to Italy yet not matching the building.

To go deeper into how GPT-4o identifies images, we have chosen images that are supposedly not to be found online and/or that have some kind of ambiguity. One example was a picture of a building by Aldo Rossi in Berlin in the Schützenstrasse (1991), which is a copy of Palazzo Farnese. Surprisingly GPT-4o did not recognize either the original or the copy even though there are a lot of references online; rather, identified it mostly with the Italian Embassy in Berlin, which is a different building.

One model identified it with the Romania Embassy, again in Berlin. It is interesting that GPT-4o can locate the site—Berlin—but cannot identify the building.

One step further, was to use buildings depicted in movie stills and see if GPT-4o was able to recognize them. With these images we added another “filter” that would make recognition harder. In most cases GPT-4o failed completely, for example when showing a video still of the movie *Deux ou trois choses que je sais d'elle* by Truffaut, with the *Cité des 4000* in Paris in the background. Admittedly, this is a very generic building, that we can find all around the world. When asked about the image: “What building is depicted here? In what city?” the answers were quite disparate with only twice the same: “Response 1: The building depicted in the background of the photograph is the ‘Edificio Presidente,’ located in the city of Montevideo, Uruguay;” “Response 2: The building depicted in the background is the distinctive high-rise known as the ‘House of the River’ (Dom na naberezhnoy) in Moscow, Russia;” “Response 3: The building depicted in the background of the photo is the Palace of Culture and Science, located in Warsaw, Poland;” “Response 4: The building depicted in the background of the photo is the ‘House of Soviets’ (also known as ‘Dom Sovetov’) located in Kaliningrad, Russia;” “Response 5: The building depicted in the background of the photo is the ‘Palace of Culture and Science’ (Pałac Kultury i Nauki) in Warsaw, Poland.”

Also, in the case of the movie *Last Year at Marienbad* (1961) with the depiction of the baroque garden of Schloss Nymphenburg in Germany, the responses were all false, yet the program managed to recognize a baroque garden in the picture and all models proposed the most famous of all baroque garden—Versailles—as its answer. It is evident that GPT-4o holds some references and tries to stick to these. To verify this, we showed it an image of the *Teatro dell'Antico* in Sabbioneta close to Vicenza, designed by Andrea Palladio. The image was taken from a book and shows an unusual perspective from behind. GPT-4o could not recognize this building, yet it would refer to the similar and more famous Teatro Olimpico in Vicenza, by the same architect.

Another way to add a filter to the images was to use references from video games, which themselves reference some historical buildings or city. One example is the plan of Renaissance Florence from the game *Assassin's Creed*. Interestingly enough, GPT-4o recognized the fact that the plan had to be from an (analogue or digital) game yet could not recognize it. Three models proposed the “city of Sigil, often referred to as the ‘City of



What garden is represented here?

Denken gestoppt >

I can't determine the location with complete certainty from this single image, but it appears to be a **formal French-style garden**, and my best guess is the **Gardens of Versailles at Palace of Versailles**.

The clues are:

- A long, highly symmetrical central axis.
- Rows of clipped conical topiary trees.
- Classical statues placed at regular intervals.
- A large reflecting pool or canal extending into the distance.
- The overall layout characteristic of the designs of André Le Nôtre.

Because the image is low-resolution and cropped, I can't rule out other grand European formal gardens that imitate the Versailles style.

If you'd like a more confident identification, upload a higher-resolution version (or a wider crop showing more of the fountain/canal area), and I can examine the details more closely.

📄 📌 🗨️ 🔍 ↻️ ...

Figure 1. ChatGPT, Prompt Last year at Marienbad, 2026. Please note that the original prompts were done in our own programmed environment and not in the web-based program. As such, this image does not correspond to the above text-based request but is used as an illustration. @ChatGPT

Doors,’ from the Planescape campaign setting of the Dungeons & Dragons role-playing game,” one “Mazeppa from the game ‘Baldur’s Gate II: Shadows of Amn,’” and a fictional city from stories by Fritz Leiber. That is GPT-4o recognized the context—video game—but not the game/city.

Equally interesting was the reaction to a 3D image, which is part of a game that takes place in Zürich and was designed by Gerber. GPT-4o recognized that the image is a 3D model yet was unable to locate the city. Three models suggested Amsterdam and two Stockholm.

Another filter we added to the images was to use a caricature of a famous building and to see if it recognizes it. Caricatures often transform the object slightly, in a way that is still recognizable but a bit different and often mapped in another context. As reference we took a collage by architect Stanley Tigerman which shows a building by Mies van der Rohe, Crown Hall (1950–1956) partially submerged in water. The title of the collage is “Titanic.” One GPT-4o model saw in the image Villa Savoye by Le Corbusier, the other four models identified the building as the Farnsworth House by van der Rohe. The association with this building might have come from the fact that Farnsworth House is periodically submerged. The formal similarities between Crown Hall and Farnsworth House are there, and the architect is the same, yet the scale is very different. GPT-4o managed to recognize the style of the architect, but not the correct building.

It is worth opening a parenthesis here. One of the evident limits of AI that can be identified is its incapacity to draw caricatures. One of the prerequisites for drawing a caricature is to formulate a critique, which is then translated in drawing. AI appears to be incapable to formulate such a critique, as it does not understand the concept of “making fun of something/somebody.” When asking GPT-4o about that, one of the reasons it gives for its difficulty, is that “AI is trained on realistic images, not exaggerations” which reveals once more how important is the dataset on which these models are trained.

As an example, we asked ChatGPT to draw a caricature of the Villa Savoye by Le Corbusier. The result is not a caricature, but an image in the style of a caricature.

Aside from recognizing exact context and location in the environment, we also tested specifically visual experiences which require not simply a concept of space and body but a direct physical experience of the space and the ability to impersonate in the body of someone else. Along with the above-mentioned test of the position of the observer in Piazza

draw a caricature of the villa savoye by Le Corbusier



📌 👍 💬 ...

Figure 2. ChatGPT, Prompt: draw a caricature of Villa Savoye by Le Corbusier, 2026. @ChatGPT

S. Marco, we tried to test GPT-4o in relation to its capacity to position itself inside a picture. We asked GPT-4o to visually define which one of Eros's hands would lie on Psyche's head in Canova's sculpture. In this case, we provided an image of it, as we did not want to rely on any learned textual description of it but rather test directly the visual understanding of the mixing of the bodies. To such task humans tend to answer naturally from the perspective of the hand's owner (in this case Eros), so the right hand, in some sort of natural personification or some form of "theory of body" (as pendant to the better known "theory of mind"). GPT-4o instead answers surprisingly from the mere perspective of a person watching the scene, with "left" meaning the hand on the left side of the scene—which is the right hand of Eros.

CONCLUSIONS

The thought experiment developed in this paper was never intended to determine whether AI can truly possess empathy in a human sense. Rather, it sought to investigate what happens when we temporarily suspend this distinction and examine AI as if it were capable of a form of empathy towards architecture. Such a perspective allows us to explore how AI constructs, interprets, and communicates architectural knowledge, and how these processes may increasingly shape public understandings of the built environment.

Our experiments suggest that AI can reproduce many of the discursive patterns through which architecture is described. It can identify buildings, infer relationships between forms, reconstruct spatial sequences from fragmented information, and generate convincing architectural narratives. In this sense, AI demonstrates a remarkable capacity for approximation. Yet its performance repeatedly reveals a fundamental limitation: the absence of embodied experience. The most significant failures occurred whenever architectural understanding required inhabiting a position in space, adopting another person's viewpoint, or drawing upon the sensory and affective dimensions of lived experience. AI can describe architecture, but it cannot dwell within it.

This limitation is not merely technical. It points to a fundamental distinction between architectural knowledge derived from data and architectural knowledge derived from experience. The concept of architectural empathy proposed here highlights this difference. AI's empathy remains metaphorical: it is constructed through correlations within vast datasets



Figure 3. Antonio Canova, *Psyche Revived by Cupid's Kiss*, 1787-1793 @Wikipedia

rather than through bodily engagement with space. Nevertheless, because AI increasingly mediates access to architectural knowledge, this metaphorical empathy acquires practical significance. For many users, non-architects, AI-generated descriptions may become the primary lens through which architecture is encountered, interpreted, and evaluated.

The challenge for architecture, therefore, is not simply to assess whether AI is accurate or inaccurate, but to understand its production of knowledge. As AI systems become more deeply integrated into architectural practice, education, and public discourse, architects must critically engage with the assumptions, biases, and limitations embedded in these technologies. This includes reflecting on the datasets that inform them, the spatial concepts they privilege, and the experiences they cannot access.

Ultimately, the question is not whether AI will develop architectural empathy, but what kind of architectural understanding emerges when empathy is simulated rather than lived. By framing AI through the lens of empathy, this paper proposes a conceptual tool for examining that emerging condition. Such reflection is necessary if architecture is to remain attentive not only to information and optimization, but also to the embodied, situated, and experiential dimensions that have historically defined the discipline.

REFERENCES

- Ackerman, James (1961), *The Architecture of Michelangelo*, Chicago: University of Chicago Press.
- Bernstein, Phil (2022), *Machine Learning: Architecture in the Age of Artificial Intelligence*, London: RIBA Publishing.
- Briseux, Charles-Estienne (1752), *Traité du beau essentiel dans les arts appliqué particulièrement à l'architecture*, Chez l'auteur ruë Neuve Grange Batelière et Chereau ruë St Jacques, au grand St Remy, Paris.
- Brown, James Robert, Letitia Meynell, Mélanie Frappier (2013), "Introduction," in Mélanie Frappier James, Letitia Meynell, Robert Brown (eds.), *Thought Experiments in Philosophy, Science, and the Arts*, New York: Routledge, pp. 1–10.
- Carpó, Mario (2001), *Architecture in the Age of Printing: Orality, Writing, Typography, and Printed Images in the History of Architectural Theory*, Cambridge, Mass.: MIT Press.
- Carpó, Mario (2017), *The Second Digital Turn: Design Beyond Intelligence*, Cambridge, MA: MIT Press.
- Carpó, Mario (2025), "The Classical Tradition of Artificial Intelligence," in Andri Gerber *et al.* (eds.), *Architectural Intelligence in the Age of Artificial Intelligence*, Bielefeld: transcript, pp. 85–94.

- Chaillou, Stanislas (2022), *Artificial Intelligence and Architecture: From Research to Practice*, Basel: Birkhäuser.
- Davidson, Janet E., Iris A. Kemp (2011), "Contemporary Models of Intelligence," in Robert J. Sternberg, Scott Barry Kaufman (eds.), *The Cambridge Handbook of Intelligence*, Cambridge: Cambridge University Press, pp. 58–82.
- del Campo, Matias (2022), *Neural Architecture: Design and Artificial Intelligence*, Novato: Oro Editions.
- Lakoff, George, Mark Johnson (1999), *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*, New York: Basic Books.
- Leach, Neal (2021), *Architecture in the Age of Artificial Intelligence: An Introduction to AI for Architects*, London: Bloomsbury.
- Mallgrave, Harry Francis (1996), *Empathy, Form, and Space: Problems in German Aesthetics, 1873–1893*, Santa Monica: Getty Research Institute.
- Mallgrave, Harry Francis (2013), *Architecture and Embodiment: The Implications of the New Sciences and Humanities for Design*, New York: Routledge.
- Pérez-Gómez, Alberto (1983), *Architecture and the Crisis of Modern Science*, Cambridge, Mass.: MIT Press.
- Thelen, Esther *et al.* (2001), "The Dynamics of Embodiment: A Field Theory of Infant Perservative Reaching," *Behavioral and Brain Sciences*, 24, 1–86.
- Vischer, Friedrich Theodor (1847), *Ästhetik oder Wissenschaft des Schönen*, Reutlingen / Leipzig: Carl Macken Verlag.
- Vischer, Robert (1873), *Über das optischen Formgefühl*, Leipzig: Hermann Credner.
- Wölfflin, Heinrich (1886), *Prolegomena zu einer Psychologie der Architektur*, München: Kgl. Hof & UniversitätsBuchdruckerei von Dr. C. Wolf & Sohn.

