

Philippe Morel*

BEYOND MIMESIS: FROM IMITATION TO EXTRAPOLATION IN ARCHITECTURAL INTELLIGENCE

ABSTRACT: The progressive encapsulation of architectural knowledge within technical systems has led to increasing automation of design and decision-making processes. Long limited to the execution of explicit rules, this automation is now undergoing a profound transformation with the rise of machine learning systems, characterised by self-regulation, probabilistic decision-making, and increasing opacity of their internal mechanisms. This article analyses this transformation in three stages. First, it revisits Shannon's rigorous quantification of information, a prerequisite for its machine processing. Next, it examines the genesis of computability and the conjecture formulated at the Dartmouth conference, according to which intelligence and creativity can be simulated by formal procedures incorporating controlled randomness. Finally, by comparing Mario Carpo's imitation thesis with recent results in high-dimensional geometry, particularly those of Yann LeCun, the article argues that contemporary AI cannot be described adequately as a simple machine of imitation. Under a strict geometric definition of interpolation, contemporary AI systems are better characterised as operating through extrapolative generalisation than through interpolation in the ordinary sense. This leads to a redefinition of the autonomy of the architectural automaton, understood less as the capacity to reproduce precedents than as the tendency, in high-dimensional spaces, to generate outputs beyond the strict envelope of available data.

* Philippe Morel: The Bartlett School of Architecture, Faculty of the Built Environment, University College London; p.morel@ucl.ac.uk.

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.

KEYWORDS: artificial intelligence, extrapolation, computational architecture, high-dimensional geometry, imitation, information theory, theory of architecture

PART I. ARCHITECTURAL INTELLIGENCE AND THE ADVENT OF MEASUREMENT

Fragmentation of Intelligence and the Persistence of the Qualitative

Each historical period is marked by the emergence, crystallisation, and formalisation of new forms of architectural intelligence. Whether or not one considers form (*morphē*) to be the primary expression of intelligence in a discipline still referred to, for convenience, as “architecture”—even though it now appears increasingly diversified, fragmented, or even dissolved into adjacent fields ranging from computational urbanism to interaction design—the twentieth century can be said to have witnessed the emergence of an unprecedented number of new architectural intelligences. These intelligences, which will be referred to interchangeably as “architectural intelligences” or “intelligences of architecture,”¹ cannot be reduced to a single or monolithic definition. They must be understood in at least two distinct but complementary senses, both of which have structured classical and modern architectural theory. The first sense, which may be described as cognitive and subjective, refers to the human capacity to understand architecture: to grasp its tacit rules, recognise its specificities, and navigate its multiple dimensions—whether material, economic, political, or historical. This form of intelligence is that of the designer who synthesises heterogeneous constraints, of the user who inhabits the space, or of the visitor who experiences it. It falls within the domain of a hermeneutics of space. The second sense, objective and ontological, refers to the intelligence that is inscribed, incorporated, or embedded in a construction, or even in simple graphic documents. A Gothic vault, a Palladian plan, or a detail of modern construction “contains” intelligence—that is, organised information that resolves problems of force, light, or circulation, embodies theoretical positions, and transmits

¹ The expression “the intelligences of architecture” could mean the different understandings of architecture, but this is not the meaning adopted here, although the present text implicitly raises the question of the ability of machines to understand architecture and therefore to process semantic information.

symbolic content. This second meaning of architectural intelligence—intelligence embedded in artifacts—is not a weak metaphor. It designates an operational reality: an organisation of constraints, decisions, and regularities that can be manipulated, transmitted, and, today, reproduced without recourse to conscious human intention. It is precisely this progressive autonomisation of embedded intelligence that makes its mechanical processing possible.

Until very recently—until the advent of generative artificial intelligences behaving as if they understood the meaning of their creations—these two definitions rested on an implicit and insurmountable anthropocentrism. The first definition implied the existence of a human rationality capable of establishing causal relations, hierarchies of value, and symbolic correspondences. The second definition, for its part, always referred to the existence of an additional *je ne sais quoi*, an ineffable quality eluding strict scientific explanation. This romantic idea persisted, inherited in part from the theory of Fine Arts and reinforced by phenomenology, according to which a work of architecture could contain “more”—more know-how, more technical skill, more knowledge, or more emotion—than the simple sum of the information necessary for its conception and execution. The architectural work was perceived as an organic whole, greater than the sum of its technical parts. However, from the strictly quantitative perspective provided by information theory—a theory which, as we shall see, forms the basis of our computational modernity—this hypothesis of an immaterial surplus is untenable, even physically impossible. From the standpoint of digital representation, any object, building, or drawing can be encoded to finite precision and processed as finite data. Faced with this contradiction between informational finitude and the feeling of the inexhaustibility of the sensible, architecture has historically adopted a defensive strategy: postulating the fundamental irreducibility of architectural experience. This irreducibility, however, is never demonstrated; it is asserted as a disciplinary axiom, often due to a lack of formal means to propose an operational reduction. In fact, to circumvent or at least displace the problem, architecture has taken refuge in disciplines external to strict technique, such as aesthetics, phenomenology, or semiotics. The underlying assumption of this resistance is that architecture cannot be reduced to a few megabytes of information stored in a computer’s memory. It is precisely this assumption of the irreducibility of the sensible that the computational age and artificial intelligence radically challenge.

A physical experience of architecture engages one's own personality, culture, and state of mind—in reality, an extremely broad, diverse, and complex general context. At its most extreme, although this experience is described as physical, it also inevitably involves—at least for some people—a worldview (*Weltanschauung*) that artificial intelligence “world models” are attempting to approximate, though these world models focus mostly on their physical dimensions. Trying to circumscribe this context to represent it faithfully as data would lead to immense theoretical difficulties. How can we translate a “state of mind,” a fleeting emotion, or the cultural background that guides an aesthetic judgment into binary variables? Until the advent of AI, these questions were left to the humanities. Architecture, in its practice, was content to quantify the “quantifiable.” During the design phase, the focus was on measuring spaces, evaluating costs (the famous *devis*), and dimensioning structures—areas for which rigorous formalisation had been made necessary, and indeed natural, by the industrial and scientific revolutions. As for the construction phase, the problems of representation and quantification associated with it were *ultimately* just an extension of the former, at least from the Renaissance onward, when graphic design documents truly dictated the building process.

Geometry and Arithmetic: A Historical Division of Labour

This limitation of quantification to physical and economic aspects has historically structured the architectural profession and its relationship with engineering, including before the Renaissance, during the empirical dialogue between drawing (*l'art du trait*) and wood or stone cutting, where quantification and representation went hand in hand within a kind of “geometric and arithmetic certainty.” Ordinary or Euclidean geometry² dealt with surface and volume elements (to simplify, what occupies space and defines form), while arithmetic dealt with elements without direct spatial extension (for example, costs, or more recently, energy losses or carbon footprint). This symbolic division of labour between

² Some might be tempted to invoke descriptive geometry instead of Euclidean geometry, but, as the mathematician Jean Dieudonné pointed out, descriptive geometry is merely a “simple drawing technique:” “I am not speaking of ‘Descriptive Geometry,’ a simple drawing technique which has fortunately been brought back to its proper proportions almost everywhere.” See J. Dieudonné, *Algèbre linéaire et géométrie élémentaire*, Hermann, Paris, 1964, p. 11 (second footnote on p. 11). This footnote, slightly polemical and related to the French context, is absent from the standard English translation *Linear Algebra and Geometry*, Houghton Mifflin, Company Boston, 1969.

geometry and arithmetic has, in the history of architecture, translated remarkably stably into a very concrete division of labour between architects and engineers, definitively established by Jean-Nicolas-Louis Durand³ and Jean-Baptiste Rondelet at the beginning of the nineteenth century. This system, in which architects are responsible for the science of space and engineers and other specialised bodies are responsible for virtually everything else, has endured through the centuries with remarkable stability. While the details became more complex in the twentieth century with the use of new sciences (for example, thermodynamics, which governs thermal engineering), the general structure of this division of labour remained unchanged until the advent of contemporary artificial intelligence systems.

We are witnessing a fundamental epistemological shift today. Current AI systems, by encoding all architectural knowledge (texts, plans, images, historical contexts, critical analyses) within vast neural networks, are achieving an unprecedented fusion of knowledge. They are giving rise to what we propose to call, extending the analogy with the latent space of mathematics, “latent engineering.” This term designates a diffuse, shared, and constantly accessible engineering capability that is no longer limited to load or thermal calculations. The focus of this latent engineering is no longer classical quantification (volumes, material strength, costs) measured in standard units (meter, pascal, euro). The challenge now lies in quantifying and processing what we have never been able to quantify before: complex contextual elements, embodiment (bodily and situated inscription), or even architectural quality itself, invariably reduced in the past to vague notions such as elegance, poetry, beauty, the sensations generated, or famously creativity.⁴ Where aesthetics, psychology,⁵ semiotics,⁶ and cognitive science have failed to provide a universal measure of architectural quality, AI offers a radically different approach: not a direct measure of beauty or creativity, but learned vector representations that

³ Cf. W. Szambien, *J.-N.-L. Durand 1760–1834: de l'imitation à la norme*, Picard, Paris, 1984; J.-B. Rondelet, *Traité théorique et pratique de l'art de bâtir*, published by the author, Paris, 1812–1817; P. V. Aureli, *Architecture and Abstraction*, MIT Press, Cambridge, Mass., 2023.

⁴ See M. Boden, *The Creative Mind: Myths and Mechanisms*, Basic Books, New York, 1991.

⁵ See, for example, R. Arnheim, *Art and Visual Perception: A Psychology of the Creative Eye*, University of California Press, Berkeley/Los Angeles, 1954.

⁶ On this subject, see for example U. Eco, *A Theory of Semiotics*, Indiana University Press, Bloomington/London, 1976.

correlate with patterns of judgment, resemblance, and preference found in data. Automation, as the “progressive encapsulation of human knowledge in technical systems,”⁷ crosses a critical threshold here. It no longer concerns only the execution of repetitive tasks, but the encapsulation of qualitative judgment itself.

Aesthetic Objectivity and Its Limits

It would be wrong to assume that the ambition to rationalise art and measure beauty was born with the computer. Projects to explain and formalise aesthetic experience are concomitant with scientific modernity.⁸ Art history has extensively documented this vast—profound but uncoordinated—project of objectively quantifying artistic qualities that unfolded between the mid-nineteenth century and the 1970s. From theories of pure vision to structuralist analyses, the West has sought to reduce art to equations. The culmination of this quest was reached when the European and American artistic avant-gardes, often in direct or indirect dialogue with logical neo-positivism (the Vienna Circle), attempted to produce an “exact” art. Constructivism and concrete art participated in this desire for rigor, with the “modernist aporias” that conceptual and minimal art would later highlight. However, postmodern relativism brought this dynamic in the cultural sphere to a halt. This scepticism was undoubtedly beneficial, preventing art from becoming a scientistic caricature of science. Nevertheless, as the philosopher of mathematics James Franklin ironically observed, “The postmodernists’ attacks have washed off science like water off a duck’s back.”⁹ Science has not abandoned explaining the qualitative simply because art critics have declared the end of grand narratives; it has merely changed its tools and methods.

⁷ *Khôrein: Journal for Architecture and Philosophy*, “Call for Papers: Volume IV, Issue 3: Automata,” <https://khorein.ifdt.bg.ac.rs/index.php/ch/announcement/view/22>, (accessed 8 April 2026).

⁸ Beyond classical aesthetics, perhaps the most serious attempt to quantify aesthetics was G. D. Birkhoff, *Aesthetic Measure*, Harvard University Press, Cambridge, Mass., 1933. Although, in my view, it remained inconclusive, the theory of aesthetic measurement found a “second life” with applications of C. Shannon’s information theory to aesthetics; see A. Moles, *Information Theory and Aesthetic Perception*, Denoël, Paris, 1972. More recently, it has reappeared in algorithmic methods (formal grammars, cellular automata, agents, etc.) and in algorithmic information theory. Most current approaches, however, still seem to fall into some form of scientism.

⁹ J. Franklin, “Science Wars: Last Round? (A Review of *Are Science and Mathematics Socially Constructed? A Mathematician Encounters Postmodern Interpretations of Science* by Richard C. Brown),” *The New Criterion*, 29, 9, 2011, p. 68.

Faced with the inconclusive or overly limited results of traditional scientific aesthetics or structural semiotics, Artificial Intelligence has taken up the problem of quantifying the qualitative. From its very origins as a discipline, AI has posited a bold and reductionist hypothesis: qualitative problems (natural language, visual perception, creativity, learning) can be addressed using nothing more than what is strictly measurable, that is, information and computation. This represents a major shift: the search is no longer for the essence of beauty (a philosophical approach), but rather for simulating the effects of intelligence and aesthetic judgment through computational means.

The Dartmouth Programme: A Foundational Conjecture

The institutional birth of this ambition is found in a document of paramount importance to the history of ideas: the *Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*,¹⁰ written in August 1955. The project is not a product of naive positivism but rather a methodological gesture: unable to define intelligence in a substantive way, the Dartmouth researchers chose to make it a working hypothesis falsifiable through simulation. The text of the proposal, signed by the founding fathers John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon, leaves little room for ambiguity regarding the reducible nature of human intelligence. From its preamble, the proposal states a conjecture that remains one of the founding hypotheses informing major strands of AI research today:

The study is to proceed on the basis of the conjecture that every aspect of learning or *any other feature of intelligence* can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find out how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. [Emphasis added]

The document goes on to list seven aspects of the “artificial intelligence problem” that were to be studied during the summer of 1956. The relevance and prescience of this list, seventy years later, are astounding.

¹⁰ J. McCarthy, M. L. Minsky, N. Rochester, C. E. Shannon, “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence,” *AI Magazine*, 27, 4, 2006, p. 12, <https://doi.org/10.1609/aimag.v27i4.1904>.

All the contemporary challenges of generative AI and large language models (LLMs) are foreshadowed: in §1, the “speeds and memory capacity of [then] present computers may be insufficient to simulate many of the higher functions of the human brain,” but, notably, it is immediately clarified that “the major obstacle is not lack of machine capacity, but our inability to write programs taking full advantage of what we have;” §2, the continuation of sentences from a word; §3, “How can a set of (hypothetical) neurons be arranged so as to form concepts;” §4, how to measure the complexity of computations with “a theory of the complexity of functions;” §5, how to enable “self-improvement” (a sign of a “truly intelligent machine”); §6, how to “form abstractions from sensory and other data.” The last aspect of the artificial intelligence problem mentioned is “§7 Randomness and Creativity.” On this last point (§7), crucial for rethinking architecture in the age of automata, the proposal put forward a hypothesis that resonates powerfully with current theories and the idea of an unpredictable automaton:

A fairly attractive and yet clearly incomplete conjecture is that the difference between creative thinking and unimaginative competent thinking lies in the injection of a some [sic] randomness. The randomness must be guided by intuition to be efficient. In other words, the educated guess or the hunch include controlled randomness in otherwise orderly thinking.¹¹

This vision of “guided randomness” at the heart of orderly thinking lays the foundation for a non-deterministic conception of the automaton. Far from the rigid execution of fixed rules (the classical automation mentioned by Marco Vanucci¹²), artificial intelligence, as early as 1955, envisioned creativity as an injection of noise into the signal: a probabilistic deviation (deviating from a predictable trajectory). This suggests that the Dartmouth researchers were willing to treat creativity as at least partly compatible with stochastic search inside a formal system. But that conjecture remains heuristic: contemporary generative systems rely not on randomness alone, but on learned statistical structure plus stochastic sampling. As for intelligence, which includes, among other things, creativity and the capacity for self-learning, it is no longer treated as a

¹¹ *Ibid.*, p. 14. Emphasis added.

¹² *Khôrein: Journal for Architecture and Philosophy*, “Call for Papers: Volume IV, Issue 3: Automata.”

magical property of living beings but as a function that can, in principle, be modelled by sufficiently expressive computational systems, including neural architectures and Turing-equivalent formalisms. This is a crucial insight for understanding the evolution of architectural automata, from simple algorithms to contemporary generative AI. However, even if the theoretical triptych on which artificial intelligence from Dartmouth onwards is based—*quantified information* (Shannon), *universal processing procedure* (Turing), and *generation of novelty by guided randomness* (Dartmouth) still defines the horizon of current machines, the complexity of deep neural networks confronts us with a new challenge: that of extrapolation in high dimension, where controlled randomness is not enough to explain the relevance of the forms generated outside of any known corpus.

The Depsychologisation of Information

To understand how such a conjecture could have been formulated with such confidence in the mid-1950s, one must go back to the genesis of the concepts, notions, and logical and mathematical tools used by the Dartmouth organisers, as well as their backgrounds and knowledge of the state of the art at the time. Of the four organisers, only C. Shannon (born in 1916) possessed a genuine aura, rivalled only by Norbert Wiener and John von Neumann (A. Turing and K. Gödel were brilliant but less widely known theorists). Born in 1919, Nathaniel Rochester's reputation rested on his role at IBM as chief engineer of the IBM 701, the first commercial scientific computer, while the young PhDs John McCarthy and Marvin Minsky (born in 1927), though brilliant, primarily relied on their doctoral degrees from Princeton.¹³ Although only two of the four workshop organisers were senior scientists, all brought with

¹³ It may be noted that McCarthy completed his doctoral thesis, related to differential equations, under the supervision of Donald C. Spencer, a brilliant mathematician who had himself worked under J. E. Littlewood and G. H. Hardy at Cambridge. It is, however, highly unlikely that this influenced either the acceptance of the workshop or the securing of its funding. Moreover, McCarthy's biographical note in the Dartmouth proposal mentions neither of those names. See, J. McCarthy, M. L. Minsky, N. Rochester, C. E. Shannon, "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence," (op. cit.), § "Proposal for Research by John McCarthy" (p. 10) and § "People Interested in The Artificial Intelligence Problem" (pp. 11–13); see J. McCarthy, *Projection Operators and Partial Differential Equations*, PhD Dissertation n°0010958, Princeton University, 1951. <https://www.proquest.com/dissertations-theses/projection-operators-partial-differential/docview/302039611/se-2?accountid=14511>

them the latest and most fruitful scientific concepts. Besides Shannon and Rochester, who were obviously well-versed in the state of the art, McCarthy and Minsky had a precise knowledge of the work of John von Neumann,¹⁴ Norbert Wiener,¹⁵ Alan Turing,¹⁶ Warren McCulloch, Walter Pitts,¹⁷ and of course C. Shannon.¹⁸ This list is obviously not exhaustive, but it is worth noting that all the concepts necessary for the study of “some aspects of the artificial intelligence problem” as defined in the Dartmouth Proposal, had been discussed by the scientists just mentioned. These concepts were, for the most part, perfectly formalised (i.e. mathematically defined), but, even more remarkably, they existed in very limited numbers. The project to process human intelligence was thus developed from the outset as a reductionist project based on a minimal number of mathematically sound hypotheses, as evidenced by the bibliographies of the works of Shannon and Minsky¹⁹ that preceded Dartmouth’s proposal. C. Shannon’s article, “A Symbolic Analysis of Relay and Switching Circuits,” published on recommendation in December 1938 in the year-end supplement of the *Transactions of the American Institute of Electrical Engineers*,²⁰ for example, contains only four bibliographic references. It is indeed a summary of his Master of Science thesis, prepared under the supervision of Vannevar Bush at MIT, but as

¹⁴ Cf. J. von Neumann, *First Draft of a Report on the EDVAC*, Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, 1945; J. von Neumann, *Theory and Techniques for Design of Electronic Digital Computers* [lectures given at the Moore School, 8 July–31 August 1946], Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, 1948.

¹⁵ N. Wiener, *Cybernetics: Or Control and Communication in the Animal and the Machine*, Hermann & Cie, Paris, 1948.

¹⁶ See A. M. Turing, “On Computable Numbers, with an Application to the Entscheidungsproblem,” *Proceedings of the London Mathematical Society*, 2, 42, 1937, pp. 230–265; A. M. Turing, “Computing Machinery and Intelligence,” *Mind*, 59, 236, 1950, pp. 433–460; A. M. Turing, “Intelligent Machinery,” in J. Copeland (ed.), *The Essential Turing*, Clarendon Press, Oxford, 2004, pp. 410–432.

¹⁷ W. S. McCulloch and W. Pitts, “A Logical Calculus of the Ideas Immanent in Nervous Activity,” *Bulletin of Mathematical Biophysics*, 5, 1943, pp. 115–133.

¹⁸ See C. E. Shannon, “A Symbolic Analysis of Relay and Switching Circuits,” *Transactions of the American Institute of Electrical Engineers*, 57, 1938, p. 713; C. E. Shannon, “A Mathematical Theory of Communication,” *The Bell System Technical Journal*, 27, 1948, pp. 379–423, 623–656.

¹⁹ M. L. Minsky, “Theory of Neural-Analog Reinforcement Systems and Its Application to the Brain-Model,” doctoral dissertation (n°0009438), Department of Mathematics, Princeton University, Princeton, 1954, pp. 6–64. ISBN 9781084420847. https://people.csail.mit.edu/brooks/idocs/Marvin_Minsky_PhD_Princeton_1953_1954.pdf

²⁰ C. E. Shannon, “A Symbolic Analysis of Relay and Switching Circuits,” *Transactions of the American Institute of Electrical Engineers*, 57, 1938, p. 713.

Herman H. Goldstine would later say, “This surely must be one of the most important master’s theses ever written. [...] The paper was a landmark in that it helped change digital circuit design from an art to a science.”²¹ By demonstrating the perfect analogy between Boolean algebra (binary propositional logic) and electromechanical circuits (open/closed relays), Shannon paved the way for the physical materialisation of logical thought, which the invention of transistors would usher into a new era.²² He proved that reasoning could be embodied in inert matter. However, it was with his *Mathematical Theory of Communication* (1948) that Shannon made the most radical conceptual break, the one that would make AI theoretically conceivable. This was to answer a question he had already posed in 1939 in a letter to his mentor, V. Bush, outlining some ideas on “the fundamental properties of general systems for the transmission of intelligence:” is there a unified theory of communication?

To achieve this, Shannon had to build upon the work of his predecessors, notably Henry Nyquist and Ralph Hartley. In 1928, Hartley had already suggested, in “Transmission of Information,”²³ replacing the vague²⁴ and anthropomorphic term “intelligence” used by Nyquist (who spoke of the “speed of intelligence transmission”²⁵) with the term “information,” which he precisely defined to avoid any ambiguity. Hartley advocated an “elimination of psychological factors”²⁶ to establish a

²¹ H. H. Goldstine, *The Computer from Pascal to von Neumann*, Princeton University Press, Princeton, 1980, pp. 119–120.

²² Shannon spoke of the invention of the transistor as “[...] quite possibly the greatest invention of the twentieth century.” Claude Shannon, Kyoto Prize speech draft, 1985, Claude Elwood Shannon Papers, Manuscript Division, Library of Congress, Washington, D.C.

²³ R. V. L. Hartley, “Transmission of Information,” *The Bell System Technical Journal*, 7, 1928, pp. 535–563. According to *The Bell System Technical Journal*, the article corresponds to a communication delivered in September 1927 at the International Congress of Telegraphy and Telephony on Lake Como. More precisely, it was part of the Volta Congress, a multidisciplinary event held from 11 to 27 September 1927.

²⁴ “When we speak of the capacity of a system to transmit information we imply some sort of quantitative measure of information. As commonly used, information is a very elastic term, and it will first be necessary to set up for it a more specific meaning as applied to the present discussion.” R. V. L. Hartley, “Transmission of Information,” *The Bell System Technical Journal*, 7, 1928, p. 536.

²⁵ H. Nyquist, “Certain Factors Affecting Telegraph Speed,” *The Bell System Technical Journal*, 3, 2, 1924, pp. 324–346; H. Nyquist, “Certain Topics in Telegraph Transmission Theory,” *Transactions of the American Institute of Electrical Engineers*, 47, 2, 1928, pp. 617–644.

²⁶ “By this means the psychological factors and their variations are eliminated and it becomes possible to set up a definite quantitative measure of information based on physical

quantitative measure based solely on physical considerations. Shannon radicalised this “depsychologisation.” In his 1948 text, he laid out the dogma that definitively separated engineering from semantics and enabled universal quantification:

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently *the messages have meaning* [...]. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that *the actual message is one selected from a set of possible messages*.²⁷ [Emphasis added]

This exclusion of meaning is the founding act of the information age. By defining information not as meaning, but as a statistical measure of the improbability of a message within a finite set of possibilities (entropy), Shannon enables its rigorous quantification in bits. The consequence for architecture is deep: architecture, like any other cultural product, becomes translatable into a series of binary choices, independent of its “poetic,” “sensory,” or intentional content. From this perspective, architectural quality can be approached not as an irreducible essence, but as a statistically inflected configuration within a space of possible forms.

The Problem of Defining Intelligence

While information theory resolved the “fundamental problem of communication” by discarding semantics, the “artificial intelligence problem” formulated at Dartmouth reopened Pandora’s box. Intelligence, unlike Shannon’s information, resists a purely quantitative definition. It seems intrinsically linked to meaning, adaptation, and understanding—all

considerations alone.” R. V. L. Hartley, “Transmission of Information,” *The Bell System Technical Journal*, 7, 1928, p. 538.

²⁷ “The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one selected from a set of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design. If the number of messages in the set is finite then this number or any monotonic function of this number can be regarded as a measure of the information produced when one message is chosen from the set, all choices being equally likely.” C. E. Shannon, “A Mathematical Theory of Communication,” pp. 379–380.

things Shannon had set aside “for the sake of engineering.” The definition proposed at Dartmouth—“making a machine behave in ways that would be called intelligent if a human were so behaving”²⁸—is a working definition, but it remains philosophically unsatisfactory because it is circular. It shifts the problem to behavioural imitation without explaining the nature of the phenomenon. We are faced with a logical circularity: intelligence is defined as thinking, but thinking is taken as the defining characteristic of the intelligent agent. Faced with this ontological impasse, and the absence of a unified measure of intelligence (comparable to entropy for information or the joule for energy), the pioneers of AI turned to another rigorous formalisation: that of computation. This is where the guiding figure of Alan Turing comes in. Even before the invention of the physical computer, Turing had, in his 1936 paper (published in 1937)²⁹, formalised the notions of computation and abstract machine. This formalisation made it possible to consider the mind no longer as an ethereal substance (*res cogitans*), but as a logical process that can, in principle, be simulated by a computational machine. The “computational generalisation”—this reductionist yet fruitful thesis that the human brain is an information-processing computer—finds its roots in this historical convergence between information theory (Shannon) and computability theory (Turing, Church, Kleene).³⁰ For contemporary architecture, this means that “the intelligence of architecture” can no longer be conceived as a mystical intuition or an innate genius. It becomes a

²⁸ “For the present purpose the artificial intelligence problem is taken to be that of making a machine behave in ways that would be called intelligent if a human were so behaving.” J. McCarthy, M. L. Minsky, N. Rochester, C. E. Shannon, “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence,” p. 12.

²⁹ Cf. A. M. Turing, “On Computable Numbers, with an Application to the Entscheidungsproblem.”

³⁰ Although the computationalist thesis was massively developed and defended after the Dartmouth workshop, its organisers were clearly not operating in a vacuum. This is evidenced by N. Rochester’s reference to K. J. W. Craik, *The Nature of Explanation*, Cambridge University Press, Cambridge, 1943, and M. L. Minsky’s close knowledge of W. S. McCulloch, W. Pitts, “A Logical Calculus of the Ideas Immanent in Nervous Activity,” *Bulletin of Mathematical Biophysics*, 5, 1943, pp. 115–133; W. S. McCulloch, W. Pitts, “How We Know Universals: The Perception of Auditory and Visual Forms,” *Bulletin of Mathematical Biophysics*, 9, 1947, pp. 127–147; and S. C. Kleene, “Representation of Events in Nerve Nets and Finite Automata,” U. S. Air Force Project Rand, Research Memorandum 704 (RM-704), RAND Corporation, Santa Monica, 15 December 1951. https://www.rand.org/content/dam/rand/pubs/research_memoranda/2008/RM704.pdf On this subject, see also the bibliography of M. L. Minsky, “Theory of Neural-Analog Reinforcement Systems and Its Application to the Brain-Model,” p. 64.

problem of information processing and computational complexity, analysable through the tools forged by science. Artificial intelligence does not solve the mystery of architecture; it shifts the question from the realm of aesthetics to that of the computable. Importantly, the apparent re-emergence of meaning in contemporary AI systems does not imply a reintroduction of semantics at the level of formal operations. Meaning is not explicitly represented in current models, as they manipulate symbols, tokens, and vectors without any explicit human-like grounding procedure built into the computation itself. What users call “meaning” emerges only as an observer-relative effect, produced when statistically coherent outputs are interpreted by human agents embedded in cultural and historical contexts. With this in mind, and in order for computation to become “creative” and autonomous as suggested by the call for *Khôrein*’s special issue, Automata, we need to understand how the simple execution of rules (Turing) can give rise to the unexpected. This is what we will examine in the next section, by analysing the birth of the notion of computability and its limits.

PART 2. THE ARCHITECTURE OF COMPUTATION

From Information to Effective Procedures

While the first part of this essay established how Shannon’s mathematical theory of communication succeeded, as early as 1948, in “depsychologising” information by reducing it to a quantitative measure of uncertainty and selection—thereby eliminating the cumbersome question of meaning for engineering purposes—it nevertheless leaves unresolved a fundamental question for the advent of artificial intelligence: that of processing. Information, as defined by Shannon, is a “physical quantity,” storable and transmissible, comparable to a raw material. But for intelligence—or at least a semblance of architectural intelligence—to exist, this information cannot remain inert; it must be manipulated, transformed, processed. It must enter a logical dynamic.

It is imperative to understand that computational architecture is not based solely on data storage (the Shannon bit), but on the ability to execute logical rules on that data. This is where the second formal revolution of the twentieth century comes in, intimately linked to the first but conceptually distinct: the birth of computability theory. Even before the computer became a physical reality made of vacuum tubes, relays, and

then transistors, it existed as a pure logical necessity. This theoretical existence preceded its material existence by nearly a decade. To think about the intelligence of architecture today, and to understand how systems can autonomously generate forms or solutions, we must return to this epistemological turning point when the intuitive notion of “computation” became a rigorous mathematical object: the formalisation of the universal machine.

The Crisis of Foundations and the Problem of Decision

The deep origin of artificial intelligence is not found in a naive biomimetic attempt to imitate the biological brain,³¹ but in an internal crisis within mathematics and formal logic: the decision problem, or *Entscheidungsproblem*. Formulated by the mathematician David Hilbert in 1928, this problem posed a question of apparent simplicity but of abysmal complexity for the time: is there a general effective procedure—what we would call an algorithm today—that allows us to determine, for any given well-formed statement in a formal system, whether it is derivable or valid within the relevant formal framework? This question was not trivial. It concerned the very possibility of a complete and coherent mathematics, where truth would be accessible through a systematic method, without recourse to the mathematician’s “magical” intuition. To answer this question, it was imperative to rigorously define what was meant by “mechanical procedure,” “effective method,” or “calculus.” Until 1936, these notions remained relatively vague and intuitive. People knew how to calculate, but they did not know how to define calculation in its universality. While initially aiming to provide an answer to the problem of pure mathematics posed by David Hilbert, Turing would in fact be led to formalise these different notions.

It was in this context of a crisis of foundations that Turing provided the definitive formalisation in his 1936 article, “On Computable Numbers, with an Application to the Entscheidungsproblem.” The contribution of this text extends far beyond the realm of pure mathematics to touch upon the very structure of rational thought. By seeking to define what a “human calculator” (a computer, which at the time referred to

³¹ The work of McCulloch and Pitts did not aim merely to mimic the functioning of the brain without understanding it, but to understand as rigorously as possible the processes at work in thought; see W. S. McCulloch, W. Pitts, “A Logical Calculus of the Ideas Immanent in Nervous Activity,” *Bulletin of Mathematical Biophysics*, 5, 1943, pp. 115–133.

a profession and not a machine) does when applying a method, Turing “mechanised” the mind. He proposed the concept of an “automatic machine” (or a-machine), an abstract device with infinite memory (a strip of paper divided into cells) and capable of performing elementary operations on symbols (reading, writing, erasing, moving the read head) according to a finite table describing the so-called “behaviour of the machine.” This device, which history will remember as the Turing Machine, is the absolute formalisation of the automaton: a system that transitions from one state to another according to a pre-established rule, without ambiguity or intuition. By showing that this simple machine could calculate any number definable by a method, Turing closed the debate on the nature of computation and demonstrated the undecidability of the halting problem, thus rejecting Hilbert’s dream.

The Universal Machine and the Separation of Software

Turing’s genius did not stop at describing a machine capable of performing a specific calculation. In the same 1936 paper, he demonstrated an astonishing property: it was possible to design a single, universal machine capable of simulating the behaviour of any other machine. The principle was as follows: if the instruction table of a specific machine (for example, a machine dedicated to addition or geometry generation) could be described in symbolic form, and if this description were written onto the tape of the universal machine, then the universal machine would virtually “become” the described machine. This demonstration marks a decisive conceptual step toward modern computing. It prefigures the separation between a universal machine and the specific instructions it executes. This crucial distinction quickly became the conceptual model for architecture, whose design logic (rules, style, spatial syntax) would then be encoded, transmitted, and executed independently of the hardware and even independently of the rules, style, or spatial syntax involved. Hence the idea of a kind of metalanguage for architecture, associated with a new figure of the architect who no longer designs each project but programs, in a universal language (for example, a Turing-Complete programming language), a similarly universal machine (the computer as we know it today, or the mainframe computers of the 1960s and 70s that followed ENIAC and UNIVAC) which will design each specific project based on changing parameters. Beyond the unimportant question of style, this approach was followed by the pioneers of digital architecture

from the mid-1960s onward (e.g. N. Negroponte,³² C. Price,³³ and later J. Frazer³⁴), by proponents of formal grammars³⁵ (e.g. G. Stiny,³⁶ J. Gero³⁷), and by virtually all subsequent computational architecture. Architects active within what M. Carpo called the “first digital turn”³⁸ of the 1990s held a more ambiguous position regarding programmed and *ultimately* automated architecture. Nevertheless, some, including Kas Oosterhuis, further strengthened the link between programming and architecture by advocating a dynamic architecture whose constant evolution is guided by its own reprogramming.³⁹ Regardless of the “technical details” and other specificities of the various approaches to architectural design, as soon as architects had access to computers, the new reality that emerged was modelled on the shift in “computer science”⁴⁰ from specialised calculators to universal machines for symbolic manipulation, capable of simulating anything, provided the task can be described by finite rules. This universality is the condition of possibility for the “latent engineering” and pervasiveness mentioned above: a ubiquitous computing capacity, ready to receive any architectural instruction.

The Church-Turing Thesis

Turing’s formalisation of computability did not occur in isolation. It was part of an exceptional intellectual convergence that mobilised the greatest logicians of the time: Alonzo Church, Kurt Gödel, Emil Post, and Stephen Kleene. While Turing was imagining tape machines, Alonzo

³² See N. Negroponte, *The Architecture Machine: Toward a More Human Environment*, MIT Press, Cambridge, Mass., 1970; N. Negroponte, *Soft Architecture Machines*, MIT Press, Cambridge, Mass., 1975.

³³ S. Hardingham (ed.), *Cedric Price Works 1952–2003: A Forward-Minded Retrospective*, 2 vols., Architectural Association / Canadian Centre for Architecture, London and Montreal, 2016.

³⁴ J. Frazer, *An Evolutionary Architecture*, Architectural Association, London, 1995.

³⁵ On this subject, see S. Kotsopoulos (ed.), *Shape Computation: Fifty Years, 1972–2022*, Birkhäuser, Cham, 2025, <https://link.springer.com/book/10.1007/978-3-031-81623-9>.

³⁶ G. Stiny and J. Gips, “Shape Grammars and the Generative Specification of Painting and Sculpture,” in C. V. Freiman (ed.), *Information Processing 71*, North-Holland, Amsterdam, 1972, pp. 1460–1465; G. Stiny, J. Gips, *Algorithmic Aesthetics: Computer Models for Criticism and Design in the Arts*, University of California Press, Berkeley, 1978.

³⁷ J. S. Gero (ed.), *Design Optimization*, Academic Press, Orlando, 1985.

³⁸ M. Carpo, *The Alphabet and the Algorithm*, MIT Press, Cambridge, Mass., 2011.

³⁹ K. Oosterhuis, *Programmable Architecture*, L’Arca Edizioni, Milan, 2002; K. Oosterhuis, *Hyperbodies: Towards an E-motive Architecture*, Birkhäuser, Basel, 2003.

⁴⁰ It goes without saying that, in the period from Babbage to ENIAC, people were not yet speaking of “computer science.”

Church was developing a different formal system at Princeton, lambda calculus, and Stephen Kleene was working on recursive functions, a work he would pursue in relation to living organisms notably in his research paper, “Representation of Events in Nerve Nets and Finite Automata,”⁴¹ in which he claims to have “[...] demonstrated that McCulloch-Pitts neurons can govern any kind of behavior which any other kind of digital automaton at all can govern.”⁴²

It emerged, almost miraculously, that all these formal definitions of “computable,” despite their very different approaches (mechanical for Turing, algebraic for Church), were mathematically equivalent. Everything computable by lambda calculus is computable by a Turing machine, and vice versa. This observation led to the formulation of the famous “Church-Turing Thesis.”⁴³ It is not a proved theorem but a foundational and widely accepted thesis about effective computation, establishing that every physically realizable or mentally possible process is Turing-computable, or in other words that there is an equivalence between the intuitive notion of “effective computability” (what a human can calculate with paper, a pencil, infinite discipline, and without any particular genius) and computability by a Turing machine (itself equivalent to being λ -definable). The philosophical implications of this thesis for artificial intelligence—and for the intelligence of architecture—are, of course, crucial. If, as the proponents of computationalism would later do, we accept that human cognitive processes (logical reasoning, pattern recognition, architectural decision-making) are information processing processes, then, within a computationalist framework the Church–Turing thesis supports the claim that formalizable cognitive processes can, in principle, be simulated by a machine. The thesis establishes a kind of “epistemological closure,” which does not mean that everything computable is computed by existing machines, nor that every intelligent phenomenon is trivially programmable. It simply means that no formalizable cognitive process requires, in principle, a non-computational substrate. Thus, there is no magical computation that the brain could perform that would inherently elude the universal machine.

⁴¹ S. C. Kleene, “Representation of Events in Nerve Nets and Finite Automata.”

⁴² “In particular, we have thus demonstrated that McCulloch-Pitts neurons can govern any kind of behavior which any other kind of digital automaton at all can govern.” S. C. Kleene, “Representation of Events in Nerve Nets and Finite Automata,” p. 72.

⁴³ J. Copeland, “The Church-Turing Thesis,” *NeuroQuantology*, 2, 2, 2004, pp. 101–115, <https://doi.org/10.14704/NQ.2004.2.2.40>.

Consequently, intelligence becomes, *in principle*, an object of engineering. It was this theoretical certainty that later allowed the organisers of the Dartmouth conference to assert with almost unwavering confidence that “[...] every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.” Indeed, Turing’s 1936 formalisation and the equivalent formalisations of the concept of computation by Church, Gödel, Post, and Kleene transformed potential intelligence into a programming problem.

Intelligence as Performance

If the 1936 machine was a mathematical abstraction intended to resolve a problem of pure logic (and to answer Hilbert in the negative: no, not everything is decidable, since the halting problem is not), the postwar period saw the question of machine intelligence emerge on an unmistakably concrete level. In 1950, Alan Turing published in the philosophical journal *Mind* the article that is considered the symbolic birth of AI: *Computing Machinery and Intelligence*.⁴⁴ This text is fundamental because it introduces a strategic displacement similar to the one carried out by Shannon two years earlier. Where Shannon eliminated meaning to measure information, Turing eliminates consciousness to measure intelligence. From 1950 onward, two problems haunted research: that of testing and that of definition. Traditionally, intelligence is associated with the capacity to think, but the definition is, as mentioned above, circular. Faced with this difficulty and the “dangerous attitude” of referring to the common usage of the terms “machine” and “think,”⁴⁵ Turing proposes in his article to consider the problem from the perspective of a game—the “imitation game”:

It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the

⁴⁴ A. M. Turing, “Computing Machinery and Intelligence.”

⁴⁵ “The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words ‘machine’ and ‘think’ are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, ‘Can machines think?’ is to be sought in a statistical survey such as a Gallup poll. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.” *Ibid.*

interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either “X is A and Y is B” or “X is B and Y is A.” The interrogator is allowed to put questions to A and B. [...]

We now ask the question, “What will happen when a machine takes the part of A in this game?” Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, “Can machines think?”⁴⁶

We know that, although it is anthropomorphic and therefore problematic, this exact device described by Turing contains the seeds of the entire philosophy of AI that would follow, whether that AI is symbolic (“GOFAI—Good Old Fashioned AI”) or connectionist (the “current” AI of deep neural networks). Like Shannon, who succeeded in depsychoanalysing information theory, Turing sought to remove the problem of intelligence from the qualitative human approach in favour of a quantitative statistical one. By substituting the machine for the human in a social simulation game, Turing defined intelligence not by an internal substance (the soul, consciousness, biology), but by an external performance indistinguishable from that of a human. It is a statistical definition: if the machine succeeds in deceiving the interrogator as often as a human, it is intelligent, regardless of its internal mechanisms. Intelligence ceases to be an essence and becomes a capacity for successful simulation. For architecture, this redefinition has major consequences. It suggests that the intelligence of a project lies not in the author’s intention, but in the effect produced by the artifact. It matters little whether the machine “feels” the emotion of a space or understands the history of architecture. Provided it can generate plans, facades, or structures—in reality, any elements traditionally associated with architecture, and which in fact represent a genuine limit of the discipline that must precisely be overcome—deemed relevant, coherent, and culturally valid by human experts (the interrogators), it demonstrates architectural intelligence. The automaton becomes a producer of valid signs.

⁴⁶ *Ibid.*

Probabilistic Autonomy and Opacity

The combined legacy of Turing and Dartmouth allows us to understand the precise nature of the “new architectural intelligences” discussed in the first part. We are no longer dealing with simple optimisation tools (e.g. cost optimisation, structural element sizing, quantification of light inputs, or energy balance), but with systems capable of simulating complex behaviours. Recent technological developments, however, have introduced a form of complexity that was not anticipated, at least not in its full phenomenological scope: opacity. The classical Turing machine is deterministic and legible: a given state corresponds to a single action. One can unroll the tape and follow the computation step by step. Contemporary artificial intelligence systems, by contrast, especially those based on deep learning in high-dimensional spaces, operate differently. They rely on nonlinear and probabilistic logics applied to data spaces so vast that their internal functioning becomes only partially intelligible to us. This opacity does not contradict the computational framework inherited from Turing and Dartmouth, but it does show that computability does not entail transparency. If intelligence is a statistical simulation (Turing) and if creativity is guided randomness (Dartmouth), then it is natural that the most efficient systems are those that exploit these statistical dimensions to the fullest, to the point where causal relations become difficult for the human mind to trace. The contemporary architectural automaton no longer simply calculates optimal solutions in the engineer’s sense; it explores spaces of possibilities by simulating qualitative judgments. This transition from deterministic interpolation⁴⁷ to an opaque probabilistic output establishes the basis of the central problem addressed in the third part of this article.

If the automaton no longer simply computes what it has been taught, where does it find its solutions? How can intelligence emerge from raw information without simply regurgitating it, without being a mere “stochastic parrot” responding with probable solutions? Solutions which, moreover, are generally the most probable, while the least probable are often—wrongly—associated with “creativity” obtained by a greater injection of randomness, materialised for example by a high value for the

⁴⁷ Deterministic interpolation is a known deterministic transformation of data and is therefore, in principle, simulable by a Turing machine. It can be understood as the search for a stable deterministic point in a value space, equivalent to a specific final state of the tape of a Turing machine. In practice, however, the operation would be far from simple.

“temperature” parameter of generative AI interfaces. This false creativity—in fact what we could call “semantic hallucination” which results from failures of alignment or contextual grounding, should be distinguished from what we refer to, for analytical purposes, as “geometric hallucination,” which is a structural consequence of extrapolation in high-dimensional spaces. The latter is not a malfunction to be eliminated but the necessary condition under which novel and coherent forms can be generated. Returning to the question “how can intelligence emerge from raw information,” for some vehement critics, including N. Chomsky, the answer is quite simple: no genuine intelligence (i.e. at least equal to human intelligence) can be produced by machines.

The human mind is not, like ChatGPT and its ilk, a lumbering statistical engine for pattern matching, gorging on hundreds of terabytes of data and *extrapolating the most likely conversational response or most probable answer* to a scientific question. On the contrary, the human mind is a surprisingly efficient and even elegant system that operates with small amounts of information; it seeks not to infer brute correlations among data points but to create explanations.⁴⁸ [Emphasis added]

For the American linguist, in addition to the ability to create explanations (on which we agree), “True intelligence is demonstrated in the ability to think and express improbable but insightful things.”⁴⁹ While it is true that one can expect true intelligence to possess the capacity to think improbable but insightful things, it is worth dwelling a little longer on the notion of probability, and on an important mathematical distinction between interpolation and extrapolation, which, in our view, becomes a decisive critical tool for understanding the nature of artificial creativity. Faithful to the Dartmouth programme, the contemporary automaton does not simply calculate the known; most of the time—almost by default, so to speak—it “hallucinates” the possible to varying degrees. The real difficulty, therefore, is not so much how to make AI systems creative by

⁴⁸ N. Chomsky, I. Roberts, J. Watumull, “The False Promise of ChatGPT,” *The New York Times*, <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>, (accessed 8 April 2026).

⁴⁹ “True intelligence is demonstrated in the ability to think and express improbable but insightful things.” *Ibid.*

forcing them out of an “interpolation regime”⁵⁰ in order to prevent them from repeating what is already known—whether highly probable or, on the contrary, extremely unlikely—but rather how to control their hallucinations, insofar as their operation appears to involve a significant extrapolative component rather than, as most critics suggest, mere interpolation.

PART 3. BEYOND IMITATION: EXTRAPOLATION AND AUTONOMY

Generative AI as a Machine for Imitation?

While the first two parts of this essay traced the historical and epistemological arc from Claude Shannon’s quantification of information to Alan Turing’s formalisation of the effective procedure, we now encounter a dominant critical interpretation that threatens to reduce artificial intelligence to a mere echo chamber of the past. Faced with the increasing opacity of so-called, and in some respects mythical, “black box” systems, a strong theoretical temptation is to reduce the unknown—the internal functioning of the neural network⁵¹—to the known: the history of art, the history of architecture, and more generally the accumulated archive of past human productions. On this view, AI systems would merely recombine what already exists. This position was recently articulated with clarity by the architectural historian Mario Carpo. In his recent work, and particularly in his essay “Generative AI, Imitation Learning, and the Automation of Tacit Knowledge,”⁵² Carpo argues that generative artificial intelligence—whether based on generative adversarial networks (GANs)

⁵⁰ The expression “interpolation regime” is used loosely here. Strictly speaking, the argument follows the definition proposed by R. Balestrieri, J. Pesenti, Y. LeCun: “Before going further, we insist that throughout this manuscript, interpolation is to be understood as characterizing the data geometry as per Def. 1. This is not to be mistaken with the often employed interpolation regime of models which occur whenever the latter has 0 training loss on the data [...]”. R. Balestrieri, J. Pesenti, Y. LeCun, “Learning in High Dimension Always Amounts to Extrapolation,” *arXiv*, 2021, p. 1. <https://doi.org/10.48550/arXiv.2110.09485>, (accessed 8 April 2026). On the distinction from the model regime defined by zero training loss, see N. S. Chatterji, P. M. Long, P. L. Bartlett, “When Does Gradient Descent with Logistic Loss Find Interpolating Two-Layer Networks?,” *Journal of Machine Learning Research*, 22, 159, 2021, pp. 1–48.

⁵¹ Although, mathematically speaking, these processes are not strictly unknown, the depth of the networks and the number of parameters make them particularly opaque to most human users.

⁵² M. Carpo, “Generative AI, Imitation Learning, and the Automation of Tacit Knowledge,” *A+U*, 646, 2024, p. 106.

or large language models (LLMs)—is, “in epistemological terms, a machine that automates imitation.”⁵³ For Carpo, these technologies remind us of a “[...] universal condition of human creativity: namely, that nothing is created out of nothing, but all is generated out of something that already exists.”⁵⁴ From this perspective, AI does not constitute a rupture but rather a return to the long European tradition of mimesis, understood not as mechanical copying but as a process of “creative imitation” based on “the awareness, acknowledgement, election, selection, assimilation, and intelligent transformation, of a corpus of models.”⁵⁵

This interpretation is compelling because it situates AI within a reassuring cultural continuity, that of the European humanist and classical tradition, and more recently that of the variations⁵⁶ of the 1990’s “first digital turn.”⁵⁷ However, it rests on a geometric intuition derived from human-scale experience, an intuition that collapses when transposed into the high-dimensional spaces in which contemporary AI systems operate. While Carpo’s thesis remains valid as an anthropology of human creativity, it becomes difficult to sustain when elevated from an anthropology of human creativity to a strict ontology of machine generation. To translate imitation into operational terms, one must abandon metaphor and adopt a mathematical definition. In learning theory, imitation corresponds to interpolation: the generation of new samples that lie within the convex hull of the training data. Yet, as Balestrieri, Pesenti, and LeCun⁵⁸ demonstrate, under a strict convex-hull definition, interpo-

⁵³ “Generative AI, either based on legacy GANs or on more recent large language models, is, in epistemological terms, a machine that automates imitation.” *Ibid.*, p. 106.

⁵⁴ The full passage is: “Generative AI, either based on legacy GANs or on more recent large language models, is, in epistemological terms, a machine that automates imitation; it is a powerful—albeit purely technical—reminder of a basic, universal condition of human creativity: namely, that nothing is created out of nothing, but all is generated out of something that already exists. Just like Generative AI can only invent new things based on—but in fact derived from—a dataset of precedents, so all human invention entails a component of creative imitation: whereby “imitation” means, as in the Greek and Roman, humanistic and classical tradition, the awareness, acknowledgment, election, selection, assimilation, and intelligent transformation, of a corpus of models.” *Ibid.*, p. 106.

⁵⁵ *Ibid.*, p. 106.

⁵⁶ We can recall here how the theme of variation was developed during the first digital turn, particularly with Greg Lynn and Lars Spuybroek. See G. Lynn, “New Variations on the Rowe Complex,” *ANY: Architecture New York*, 7–8, 1994, pp. 38–43; L. Spuybroek, *Research & Design: The Architecture of Variation*, Thames & Hudson, London, 2009.

⁵⁷ Cf. M. Carpo, *The Alphabet and the Algorithm*; M. Carpo (ed.), *The Digital Turn in Architecture 1992–2012*, Wiley, Chichester, 2012.

⁵⁸ R. Balestrieri, J. Pesenti, Y. LeCun, “Learning in High Dimension Always Amounts to Extrapolation.”

lation in high-dimensional spaces becomes vanishingly unlikely as dimensionality increases. As dimensionality increases, the probability that any new point lies within the convex hull of existing data converges toward zero, unless the dataset grows exponentially with dimension—a condition that no finite corpus can satisfy. Consequently, under this geometric definition, contemporary generative AI systems cannot be described primarily as navigating within the envelope of precedents; they frequently generate outputs that lie beyond it. What appears as stylistic continuity is not the result of interpolation but of structural constraints shared by the data themselves, which occupy an extraordinarily sparse space. Generation is therefore not a recombination of known forms but an extrapolative projection into regions of space that have never been sampled. From a geometric standpoint, strict interpolation-based imitation becomes highly implausible in high-dimensional settings.

A Geometric Definition of Imitation

To carry out this comparison, we must, beyond metaphors, translate the literary and artistic concept of imitation into an operational mathematical concept. If, as Carpo suggests, to imitate means to produce a plausible variation from a corpus of precedents, geometrically this would correspond to the operation of interpolation.⁵⁹

In function approximation theory, as in machine learning, interpolation, in its strictest definition, consists of calculating a new value or a new point that lies between known values. If we imagine the entirety of existing architectural knowledge (all plans, all images, all texts, and even, for our thought experiment, all fine-grained 3D scans of every building ever constructed) as a point cloud $X = \{x_1, x_2, \dots, x_N\}$ in a vector space, “imitating” would consist of generating a new project x that lies geometrically within the envelope formed by these points. More precisely, we can hypothesise, and even use as a definition of imitation, that it is nothing other than an interpolation, or possibly an infinitesimal extrapolation, of a point existing in a space of any dimension. This geometric definition aligns with our intuition, common practices, and examples taken by Carpo: if we copy a style, we remain close to examples of that style; if we blend two styles, we position ourselves “between”

⁵⁹ Although the mathematical problems of interpolation and extrapolation in AI models are complex, the present argument remains valid.

them. If AI were an imitation machine in Carpo's sense, it would navigate cautiously within the territory defined by its training data. It would produce weighted averages of Le Corbusier and Mies van der Rohe. It would remain, to use the precise topological term, within the convex hull of the training data.

In machine learning, according to R. Balestrieri, J. Pesenti, and Y. LeCun, the common adage has long been, "as an algorithm transitions from interpolation to extrapolation, as its performance decreases." [*sic*] In this classical view, interpolation is the domain of mastery and faithful reproduction (successful imitation); extrapolation is the domain of uncertainty, error, and dangerous hallucination. However, recent discoveries in deep learning theory contradict this reassuring model. As Randall Balestrieri, Jérôme Pesenti, and Yann LeCun demonstrate in their 2021 article "Learning in High Dimension Always Amounts to Extrapolation" that the intuition that neural networks interpolate is not only false but mathematically almost impossible when dealing with complex data.

The Curse of Dimensionality

The core of the argument rests on the counterintuitive behaviour of geometry in high dimensions. An architectural image, even a low-resolution one, or a text describing a project, exists in mathematical spaces with thousands, even millions, of dimensions ($d > 100$); moreover, the more multimodal the description, the larger the space will be. Balestrieri, Pesenti, and LeCun first recall the canonical definition of interpolation that we use: "Interpolation occurs for a sample x whenever this sample belongs to the convex hull of a set of samples $X \triangleq \{x_1, \dots, x_N\}$, if not, extrapolation occurs."⁶⁰ This definition does not depend on the local metric or the intuitive notion of perceptual resemblance; it depends solely on the global affine geometry of the dataset, which renders any analogy with human imitation based on local semantic proximities mathematically irrelevant. They then rely on a result from stochastic geometry—the Bárány and Füredi theorem (1988)—to conclude that in high dimensions, interpolation "almost surely never occurs." Their proof revolves around an asymptotic theorem that describes the behaviour of the interpolation probability as the dimension d tends to infinity. The conclusion is that to remain within the framework of interpolation, a virtually

⁶⁰ *Ibid.*, p. 1. <https://doi.org/10.48550/arXiv.2110.09485>, (accessed 8 April 2026).

infinite amount of training data would be required, with the amount of data needed increasing exponentially with the dimension.⁶¹

The general result proposed by Balestrieri *et al.* is, beyond learning theory, conceptually very interesting for addressing architecture in the age of artificial intelligence, and particularly the imitation theory proposed by M. Carpo. By stating that in order to maintain a non-zero chance that a new point will fall inside the convex hull of the existing data (i.e. for imitation/interpolation to occur), the size of the dataset must grow exponentially with the dimension d , the Balestrieri *et al.* result amounts to asserting that no amount of data, however large, will ever be sufficient to make the absence of hallucinations near-certain or, with regard to architecture, to secure the kind of humanist rationality characteristic of the theory of creative imitation. While we humans are overwhelmed by data, contrary to the common intuition that leads us to project our perceptual and computational abilities onto machines, for machines, the entirety of the data ever produced represents only an infinitesimal fraction of what they might need and potentially use. Despite now possessing trillions of parameters, AI models are, to borrow a famous epistemological thesis, still “underdetermined” by reality.⁶² Moreover, even if we were to produce an immense quantity of synthetic data, this quantity would likely remain negligible compared to the results highlighted by Balestrieri *et al.* Indeed, as soon as the spatial dimension exceeds a few dozen (remember that a simple “style vector” in a latent space often has 512 or 1024 dimensions), the amount of training data—in architecture, for example, projects, buildings, construction data, etc.—required to make convex-hull interpolation common grows so rapidly with dimension that, for realistic datasets, it becomes practically unattainable. The consequence, in our view, is that for any real, finite dataset that can be stored on current servers (however massive they may be, such as the petabytes used to train GPT-4), the probability that a newly generated project will fall within the strict convex hull of the training data becomes extremely small in high-dimensional settings. In that sense, the interior coverage of the space is sparse, and new samples are likely to fall outside the hull. The question we can nevertheless ask is this: why do we get this feeling of constant resemblance when we ask generative AI to design this

⁶¹ “As such, in the presence of a nonlinear manifold, we can see that the exponential requirement from Thm. 1 [theorem 1] in the number of samples required to preserve a constant probability to be in interpolation grows exponentially with d .” Ibid., p. 2.

⁶² Cf. W. V. O. Quine, *Word and Object*, MIT Press, Cambridge, Mass., 1960.

or that object (object in the generic sense of “output”)? In our view, and limiting ourselves to architecture, it stems primarily from the fact that the essential characteristics of all architectural projects, quite simply *of all architectures ever produced*, built or not, are *very limited in number*. This property is precisely what enables the dimensionality reduction widely used in AI models during training, and which allows for relatively densely populated latent spaces within their convex hull. Indeed, the characteristics common to two houses, as well as to two cats or two pairs of shoes, are far more numerous than those that differentiate them.

The Limits of Latent Space

The significance of everything said thus far aligns with Mr. Carpo’s assertion that current generative AI reminds us of the “universal condition of human creativity: namely, that nothing is created out of nothing, but all is generated out of something that already exists.” Humans, in fact, most often create imitations augmented with minor changes, which explains the appallingly slow pace of architectural evolution and the recurring similarities, across one or even several centuries, between buildings—cathedrals, palaces, towers—that differ only in a few details of no importance except for their hermeneutic value. The only truly significant changes occur when new parameters external to architecture come into play. For example, when new construction techniques or new materials (steel, reinforced concrete) are sufficiently mature to enter the field of architecture, or when a few brilliant individuals effect a shift large enough to be considered “out-of-distribution” (OOD). While the theory of creative imitation mentioned by Mr. Carpo seems to be valid for architecture and even beyond this discipline, in our view, it is only valid *on a human level*. While creative imitation applies without difficulty to all kinds of past architectures, we believe—beyond the obvious logical impossibility of theorizing what does not yet exist—that the situation is very different with *artificial architecture*. By limiting ourselves to analysing what already exists, including generative AIs themselves, we believe, on the other hand, that Mr. Carpo is wrong when he asserts that “*generative AI can only invent new things based on—but in fact derived from—a dataset of precedents*.”⁶³ [emphasis added]. While the apparent

⁶³ “[...] Generative AI, either based on legacy GANs or on more recent large language models, is, in epistemological terms, a machine that automates imitation; it is a powerful—albeit purely technical—reminder of a basic, universal condition of human creativity: namely,

functioning of current generative AI suggests they operate through creative imitation, aptly recalling a distinctly human mechanism of creativity, attributing this same mechanism to the quasi-essence of current connectionist AIs does not withstand scrutiny. Indeed: 1) we saw earlier that datasets will never be large enough to guarantee a kind of imitative derivation characteristic of interpolation; 2) it is not the models themselves but the data on which these models are trained that exhibits similarities and statistical regularities, largely because these are the product of either human cultural history (effectively based on creative imitation) or nature, both of which are remarkably homogeneous, continuous, and even “smooth;” 3) in high dimensions, in almost all cases, generative AI does not perform interpolations but extrapolations, therefore, the quality of the results or what is more precisely called the ability to generalise, does not come from interpolations that can be likened to “creative imitations,” but from other factors.

A common objection to the view that generative AI *is not* the stochastic parrot it is wrongly labelled as, is that architectural data are not random (unlike the uniform points of the theorem invoked by Balestriero *et al.*). They are based on an underlying structure, a lower-dimensional manifold. The intelligence of architecture would then consist precisely in finding this hidden structure, made up of a sum of the characteristics common to most architectures ever produced and those specific to each one. We mentioned earlier that the very limited number of these characteristics allows for dimensionality reduction. By doing so, we could operate in a compressed “latent space” in which interpolation (and therefore imitation) would once again become possible. However, Balestriero *et al.* strongly challenge this hypothesis, both theoretically and empirically. They argue that reducing intrinsic dimensionality does not, by itself, restore interpolation in the strict convex-hull sense.⁶⁴ What matters for the probability of interpolation is not the dimension of the manifold (the actual complexity of the object, for example, the “dimensionality of style”), but the dimension of the smallest affine subspace containing the

that nothing is created out of nothing, but all is generated out of something that already exists. Just like Generative AI can only invent new things based on—but in fact derived from—a dataset of precedents [...].” M. Carpo, “Generative AI, Imitation Learning, and the Automation of Tacit Knowledge,” p. 106.

⁶⁴ See the section “The Role of the Intrinsic, Ambient and Convex Hull Dimensions” in R. Balestriero, J. Pesenti, Y. LeCun, “Learning in High Dimension Always Amounts to Extrapolation,” pp. 2–3. <https://doi.org/10.48550/arXiv.2110.09485>, (accessed 8 April 2026).

data (the convex hull). Their experiments on real-world image databases (MNIST, CIFAR-10, ImageNet) show that the proportion of test samples in the interpolation regime drops to zero as soon as more than 30 or 40 dimensions are considered. Even when projecting the data into latent spaces optimised by deep networks, extrapolation remains the rule.⁶⁵ We can therefore, both technically and conceptually, oppose Mario Carpo's position by arguing that contemporary AI systems cannot straightforwardly be described as machines of imitation if imitation is defined in the strict geometric sense of interpolation. The mere fact that they operate in high-dimensional spaces constrains them to move beyond the envelope of precedents. While the argument that everything is generated from something that already exists remains true at the dataset level, it becomes false at the generation topology level. One cannot describe generation solely as navigation within existing knowledge; a substantial part of it must also be understood as extending beyond previously sampled regions of the space.

In fact, we believe that AI's capacity to generate outputs that are not reducible to straightforward recombinations may be greater than many of its critics allow, sometimes through the "stochastic parrot" label, a term that might be more fitting for humans. Indeed, in most cases, we think, conceive, and act in surprisingly similar ways within frameworks of thought that we believe we are challenging, even though, in most cases, we have not even managed to identify those frameworks; a behaviour that Alfred North Whitehead identified very well.

When you are criticizing the philosophy of an epoch, do not chiefly direct your attention to those intellectual positions which its exponents feel it necessary explicitly to defend. There will be some fundamental assumptions which adherents to all the variant systems within the epoch unconsciously presuppose. Such assumptions appear so obvious that people do not know what they are assuming because no other way of putting things has ever occurred to them. With these assumptions a certain limited number of types of philosophic systems are possible, and this group of systems constitutes the philosophy of the epoch.⁶⁶

⁶⁵ *Ibid.* Table 1 shows that (in a latent space of dimension 512), proportion of the test set that falls into the interpolation regime for datasets of dimensions >30 can be as small as 0 per cent.

⁶⁶ A. N. Whitehead, *Science and the Modern World*, Cambridge University Press, London, 1925, p. 61.

One advantage of rarely changing our thought patterns is that it preserves the historical and cultural continuity to which humans are accustomed. Humans experience immense difficulty—being biologically “wired” in a certain way—thinking, conceiving, and acting in radically different ways, ways that would differ between individuals and differ from each person’s own prior ways of thinking. The human world, like the natural world of which we are a part and to which we are accustomed, is made up of slow evolution through limited differentiations in which, to borrow A. N. Whitehead’s words, we believe that only certain systems of thought are possible. Applied to the potential hierarchy between human and machine intelligence, this observation leads us to recognise that it is primarily because no different organisation of things has yet emerged—an organisation in which machine intelligence is, *for everyone, resolutely and visibly superior to human intelligence*—that the still-dominant system of thought presupposes the superiority of human intelligence. Although it seems natural, we believe this assumption is false, for at least three reasons: 1) numerous recent experiments suggest that AI systems can exceed human performance on a growing number of bounded benchmarks and specialised tasks. 2) Combining all specialised AIs and their extraordinary ability to exchange information at the speed of light leads to performances comparable to those of AGI, keeping in mind that there is no proof of a settled transition to AGI and that AGI itself remains ill-defined. 3) Drawing definitive conclusions about inherent limitations to machine intelligence on the basis of current performances is no less intellectually risky than asserting the opposite—namely, the necessary emergence of future superhuman capabilities. Yet opponents of what are commonly (and somewhat simplistically) labelled “strong AI” theses commit, in most cases, the very same fallacy of reasoning of which they accuse their adversaries. Having recalled these points, we will not elaborate on them further, as they have already been discussed in detail by leading authors.⁶⁷

Extrapolation and Autonomous Creativity

The geometric reversal just described leads us to redefine what we mean by creativity and autonomy. If imitation is an impossible interpolation, then AI is by default an extrapolating machine. Extrapolating often

⁶⁷ See, in particular, N. Bostrom, *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, Oxford, 2014.

means, in everyday language, making a risky assumption from insufficient data, but in the context of high dimensions, extrapolation becomes the very condition for the production of form. The word “form” is understood here either in the strict sense or in the generic sense of a result that can be mathematically formalised. Since the training data (the architectural past) is scattered throughout the vastness of the space of possibilities, any response provided by a generative AI—be it a Midjourney image, a ChatGPT text, or any other mode—appears as a projection into the void, beyond the convex hull of human knowledge, though generation remains guided by learned regularities, objective functions, and sampling procedures. LLMs, diffusion models, and other generative AI systems do not simply repeat the past but are rather engines of exploration that, beyond the apparent information overload, build precarious bridges over the “information void.” If human creativity is defined not as creation *ex nihilo* (which no one supports) but as a structured generalisation beyond known experience, then AI is also creative insofar as it extrapolates. It generates responses to situations (most commonly prompts) it has never seen, located in regions of mathematical space that have never been mapped by training data. The prompts may, of course, already be seen if prompt generation is part of the training, or simply because several users ask the same questions in the same terms (which is common). AI is not creative despite being derived from data, but because the geometry of its space forces it to invent in order to connect disparate points in its memory. Since interpolation is forbidden by the curse of dimensionality, innovation (extrapolation) becomes a statistical necessity. It is autonomy through geometric constraint.

This perspective allows us to return to the surrealist intuition evoked in the call for contributions: “Surrealist automatism, for instance, aimed to bypass rational constraints and tap into unconscious creativity. There, technical systems become instruments of deviation, surprise and poetic rupture.”⁶⁸ Mathematically, this “deviation” is extrapolation. When the algorithm generates a shape that is not within the convex hull, it “hallucinates” a reality. The paradigm of classical engineering sees hallucination as a malfunction or an error to be corrected. In high-dimensional extrapolation, hallucination is the default mode of existence for the digital object. To generate, for example, a coherent image (or other digital file) of

⁶⁸ *Khōrein: Journal for Architecture and Philosophy*, “Call for Papers: Volume IV, Issue 3: Automata.”

a building that does not exist, AI must “hallucinate” plausible details, textures, and structures—consistent with the learned distribution—but not present in any specific example. The automaton no longer calculates the known; it hallucinates the possible. In less metaphorical terms, the automaton no longer operates only as a calculator of predefined solutions but as a generator of candidate outputs under uncertainty, often beyond any exact precedent, while remaining statistically constrained by what it has learned. The guided randomness postulated by Dartmouth researchers in 1955 as a condition of creativity is therefore indeed present in generative AI; but it is anything but a simple noise term added to simulate variety: it is a vectorial component that allows movement in the empty areas of latent space without collapsing.

Although we disagree with M. Carpo that, 1) “generative AI is an imitation machine;” 2) that we “need to learn again at least the basic [sic] of the arts and sciences of imitation, so we know what Generative AI does—and what we do, when we use it;”⁶⁹ and 3) that we ultimately need “to come up with a new theory of mimesis,” assuming that this new theory were proposed, we should first emphasise that “mimesis in the age of AI” is not a reproduction of models (imitation/interpolation) but an extrapolative generation from scattered precedents. It is not a derivation along known paths but a navigation through a topology barely sketched by previous models. Since the very term *mimesis* is, in fact, inappropriate, it seems preferable not only to avoid it but, even more so, to escape entirely from the design logics it encompasses, as Peter Eisenman already urged us to do 40 years ago.

Rereading Eisenman

On the occasion of a Mies van der Rohe exhibition in Chicago,⁷⁰ Eisenman revisited the importance of escaping mimesis and reading architecture as a text: “Text emerges in Mies’s architecture when the symbolic

⁶⁹ “And, in so far as Generative AI is an imitation machine, we urgently need to learn again at least the basic [sic] of the arts and sciences of imitation, so we know what Generative AI does—and what we do, when we use it. In short, we need to engage with imitation again, in critical and creative terms; after a century of abeyance and modernist mandated oblivion, we need to come up with a new theory of mimesis.” M. Carpo, “Generative AI, Imitation Learning, and the Automation of Tacit Knowledge,” p. 106.

⁷⁰ Exhibition *Mies van der Rohe and His Disciples* (director: John Zukowsky), organised at the Art Institute of Chicago in 1986 and presented from 1 April to 15 June 1987 in the exhibition hall of the Information Center at the Centre national d’art et de culture Georges Pompidou, Paris.

connection between man and object (and hence symbol and object) loses its relevance and thus can be taken apart. Instead of simulating the vertebrate structure of man, Mies's architecture simulates a textual structure. The separation of an architecture from the mimesis of man is the production of a simulation; this is the nature of a text."⁷¹ By emphasizing the text, or more precisely the text as a set of symbols dissociated from traditional architectural symbolism—that is, from representation—Eisenman severs architecture from the primacy of meaning and the anthropomorphism it embodies. While classical text and its alphabet—a finite alphabet—aim to construct and transmit human thoughts, Eisenman's text does not aim to communicate meaning but simply "to be." In doing so, it is comparable to 1) what a text is for the information theory we have discussed: a message whose meaning is irrelevant and for which we can compute an amount of information; 2) the purely formal text that culminates in Hilbert's axiomatic project.

[For Hilbert] Certainty will come from the formal manipulation of the graphic characters that compose statements about infinity and space, and not from space itself or from geometric forms within space. To achieve this, it is necessary to return to the very theory of meaning by questioning the idea that spacetime is the locus of its elaboration. It then becomes a matter of reducing mathematical statements to their written characters, which are always finite in nature, and of setting aside their potential meaning, which can encompass infinity and its dangers [...]. [For Hilbert] there is only formal calculus, the locus of certainty, because it is, precisely, "potentially mechanizable." Certainty lies in the mechanical, far removed from the ambiguities of meaning, which are human, all too human.⁷²

Very early on—as early as his 1963 Cambridge doctoral thesis, "The Formal Basis of Modern Architecture"—Eisenman perceived the profound conceptual shifts brought about by twentieth-century science and with which architecture would be confronted. This led him, beyond

⁷¹ P. Eisenman, "miMISes READING: it does not mean A THING," in *Eisenman Inside Out: Selected Writings, 1963–1988*, Yale University Press, New Haven, 2004, p. 191. The text was originally published for the exhibition *Mies van der Rohe and His Disciples* (ed. J. Zukowsky) held at the Art Institute of Chicago in 1986; see note 70 above.

⁷² J. Lassègue, G. Longo, *L'empire numérique*, Humensis, Paris, 2024, pp. 124–125. Translation modified.

the succession of styles characteristic of the five centuries separating the birth of modern science from its “closure,” to seek the common characteristics of the different architectural styles that had succeeded one another. In 1984, he identified three such characteristics: Representation, Reason, and History, which he called fictions and grouped under a single term: “the classical.”

Architecture from the fifteenth century to the present has been under the influence of three “fictions.” Notwithstanding the apparent succession of architectural styles, each with its own label—classicism, neoclassicism, romanticism, modernism, postmodernism, and so on into the future—these three fictions have persisted in one form or another for five hundred years. They are representation, reason, and history. Each of the fictions had an underlying purpose: representation was to embody the idea of meaning; reason was to codify the idea of truth; history was to recover the idea of the timeless from the idea of change. Because of the persistence of these categories, it will be necessary to consider this period as manifesting a continuity in architectural thought. This continuous mode of thought can be referred to as the classical.⁷³

While Eisenman, in our view, describes with great accuracy the persistence of these anthropic fictions which he associates with Meaning, Truth and Time, he does not, however, go so far as to postulate the possibility of an artificial architecture completely external to the human figure of an author-reader:

[...] a “not-classical” architecture begins actively to involve an idea of a reader conscious of his own identity as a reader rather than as a user or observer. It proposes a new reader distanced from any external value system (particularly an architectural-historical system). Such a reader brings no a priori competence to the act of reading other than an identity as a reader. That is, such a reader has no preconceived knowledge of what architecture should be (in terms of its proportions, textures, scale, and the like); nor does a “not-classical” architecture aspire to make itself understandable through these preconceptions.⁷⁴

⁷³ P. Eisenman, “The End of the Classical: The End of the Beginning, the End of the End,” *Perspecta*, 21, 1984, pp. 155–173.

⁷⁴ *Ibid.*, p. 171.

The new “author-reader” must, as far as possible, discard any pre-conceptions about what architecture should be, but by remaining a human author-reader, they can only operate within their own biological limitations, which are, moreover, the very reason for the existence and persistence of the fictions identified by Eisenman. The fact that the new (architectural) “object” the American architect calls for “must have the capacity to reveal itself first of all as a text, as a reading event,” and that the whole requires “an other ‘timeless’ space of invention,” [...] “a ‘timeless’ space in the present, without a determining relation to an ideal future or to an idealised past,”⁷⁵ does not change the question. Eisenman’s new *author-reader*, *text-object*, and *timeless space* remain humanly conceived, so that the design of (architectural) *text-objects by author-readers* will, in any case, only be realised within a space of human thought—that is to say, ultimately, the very space of classical thought that Eisenman aimed to transcend. Whatever deconstruction project is put forward—and it seems to me that we now have evidence of its limitations, even its absurdity in some extreme contexts—what endures and what Eisenman fails to dismiss with his critique of the classical is, first and foremost, the *primacy of human intelligence*. The supposed superiority of human intelligence over all others—for example, over that of animals or, as far as we are concerned here, over machine intelligence—thus manifests itself, albeit for a very short time, as the true “metafiction” that makes the former possible and even inevitable. Eisenman rightly judges that “[...] the period of time dominated by the classical [the last five centuries] can be seen as an ‘episteme’, to employ Foucault’s term—a continuous period of knowledge that includes the early twentieth century,”⁷⁶ but he fails to theorise the profoundly mechanistic nature of the ongoing epistemological bifurcation. Moreover, this bifurcation, materialised by artificial intelligence whose success has been actively prepared by modern science, transcends epistemology. We are no longer faced with different spheres of knowledge within a single human space, but rather with different spaces of knowledge operating side by side. While these spaces may continue to interact for some time, there is no doubt that the knowledge space associated with AI is already expanding faster than individual human comprehension can track, and in some domains its internal procedures remain only partially

⁷⁵ *Ibid.*, p. 172.

⁷⁶ *Ibid.*, p. 155.

interpretable.⁷⁷ It is therefore plausible that highly capable systems will become increasingly difficult for unaided humans to understand in detail, even when their outputs remain useful and empirically successful. While we may acknowledge undeniable results, the modes of thought, the reasoning, —in short, the entire “logic” inherent in these results computed within new kinds of knowledge spaces—will elude us.

P. Eisenman’s author, as previously noted, is first and foremost a *reader*, hence the author-reader expression we used. In his human dimension, he remains the very embodiment of the humanism that underpins “the classical.” Like a Renaissance scholar, he only “writes” after extensive reading, and creates through creative imitation. At the end of his life, his output, however vast, represents only a fraction of his reading—human writing abilities being far inferior to reading abilities, especially if we take the term “reading” in its broadest sense of understanding or perceiving the world. For all the architecture that precedes the “computational turn”—or, if we prefer to refer to a computational paradigm precisely at work in recent AI models, the “connectionist turn”—“the primary function of alphabetic writing, which is to reproduce human speech,”⁷⁸ is preserved. The Eisenmanian author-reader can, alternatively, be either author or reader insofar as these roles coexist within a resolutely organic unity. If, as Turing asserted, “Mechanism and writing are [from his point of view] almost synonymous,”⁷⁹ it is, as Jean Lassègue and Giuseppe Longo point out, “the ‘almost’ that should hold our attention [...]” If, according to them, this “almost” “contains in itself the entire philosophical question of the limitation of formalisms,”⁸⁰ it also symmetrically contains the entire superiority of the mechanical, which can exploit (binary) coding to produce knowledge on a scale unattainable by any human means whatsoever. By its very existence “Coding completely breaks with the primary function of alphabetic writing, which is to reproduce human speech.”⁸¹ Moreover,

⁷⁷ This autonomisation of knowledge may or may not coincide with the Singularity, and may or may not belong to its very definition. My own view is that the autonomisation of knowledge is already deeply ingrained in our civilisation and that, once complete, the very word knowledge—still deeply associated with human knowledge—will become irrelevant.

⁷⁸ J. Lassègue, G. Longo, *L’empire numérique*, p. 150.

⁷⁹ A. M. Turing, “Computing Machinery and Intelligence”; cited in J. Lassègue, G. Longo, *L’empire numérique*, p. 134.

⁸⁰ *Ibid.*, p. 134.

⁸¹ *Ibid.*, pp. 132–133.

The emergence of coding in the 20th century in the form of programming languages is innovative in that it aims to automate the writing of formal deduction, and no longer just the reading as was the case with the Greek alphabet, and this is a major achievement in the history of writing. The concatenation of characters then takes place according to logical rules that require no interpretation to be applied and can therefore be executed automatically without recourse to phonemes and the voice through which the meaning of alphabetic writing is reconstructed. Every computer programme is a formal deduction that executes without relation to the voice of a speaking body.⁸²

Lassègue and Longo note, “from geometry to the living body, by way of speech as the collective guarantee of meaning, the entire phenomenology specific to the relationship between humans and living beings and the world is devalued and evacuated, encoded in sequences of 0s and 1s [...]”⁸³. However, by only partially evacuating phenomenology in order to preserve a human form of thought—thought whose no less human “reading” is the sign—P. Eisenman, and even more so any reading of architecture “in the Eisenman style,” fails to identify the full radicality of the computational. The computational dimension certainly does not appear as directly in connectionist AI as it did in first-generation AI, but it has not disappeared. The meaning that seems to emerge from neural networks as if by magic has not, in fact, been reintroduced “in the old way.” During the learning phases, no semantic network or definition is provided for each term encountered. The performance of deep neural networks does not stem from having discarded formalisation or computation to reintroduce meaning. On the contrary, it stems from the fact that we have not hesitated to discard meaning, perhaps even more radically than in symbolic AI, by exploiting the geometry of vectorised data spaces in an extraordinarily intelligent way—for example, words transformed into tokens, which are themselves transformed into vectors carrying complex context information (the embedding).

Singularity of the Work and Convex Position

In current generative AI, architecture no longer exists strictly as an Eisenman-like text, but as a set of (digital) vectors in a high-dimensional space.

⁸² *Ibid.*

⁸³ *Ibid.*, p. 127.

This specificity should lead us to study the properties of these spaces in an attempt to understand how meaning emerges and what we readily call creativity; meaning being obviously inseparable from all creativity, which, without it, would be nothing more than the mere production of random objects. To better understand why we maintain that the emergence of meaning/creativity is not reducible to the creative imitation discussed earlier, but is always, in one way or another, a production of radical novelty, we can invoke the “Convex Position Problem” (defined by P. Valtr, 1995) cited by Balestrieri *et al.* Before doing so, however, let us clarify that by radical novelty we are not referring to what appears new to us humans, but to this question from the perspective of a generative AI. Statistically, a simple answer drawn at random from a massive dataset would very likely seem new to us. It would almost certainly be more likely to seem new than a sensible answer to the original question, but it is precisely because it would also have a very high chance of being nonsensical that it would be an inadequate answer. However, to an outside observer, there would be no easily identifiable difference between this same (bad) answer drawn at random and that provided by an AI only capable of hallucinations.

Returning to the Convex Position Problem, it calculates the probability that points are in convex position, meaning that all points lie outside the convex hull formed by the others. The theorem cited by Balestrieri *et al.* (based on the work of C. Buchta, 1986) shows that in a high-dimensional hypersphere, as the number of samples grows linearly with the dimension, this probability tends towards 1. This means that each point is an extreme vertex, an exception that defines the boundary of the hull. If we interpret this result through the lens of the question “what happens inside current generative AI models when we query them?” (for example, by asking them to design a specific architecture), the answer would be that, contrary to intuition, far from producing a mass of conformist copies—as critics of automation fear—AI tends to produce objects that are all, mathematically, exceptions to the training corpus. In fact, each project generated by an AI appears as a singularity that defines its own territory, refusing to be the average of others, without (we agree on this) this singularity necessarily being equivalent to what a human would deem relevant or of high quality. What matters here is that with current generative AI, artificial architecture shifts from a logic of parametric variation—which consists of deforming a known form while remaining within certain bounds—to a logic of stochastic

exception—which amounts to producing an out-of-the-ordinary form through extrapolation. If we take current generative AIs as the state of the art and the models of the contemporary automaton, then, as the call for papers for the special issue of *Khōrein* suggests,⁸⁴ the latter functions as an “enabling partner,” “unknowable,” and “unruly,” because it operates in dimensions where the order of imitation is mathematically untenable.

CONCLUSION: AUTONOMY OF AUTOMATA AND THE END OF IMITATION

This essay has traced a theoretical trajectory from the quantification of information to the autonomy of contemporary architectural automata. Claude Shannon’s information theory neutralised meaning by reducing communication to probabilistic selection, thereby rendering cultural artifacts computable. Alan Turing’s theory of computation subsequently neutralised inspiration by demonstrating that any effective procedure could be executed by a universal machine. Together, these developments desacralised intelligence while simultaneously making it accessible to technical systems. A third and more recent transformation emerges with deep learning and high-dimensional geometry. Here, repetition itself is neutralised. As the work of Balestrieri, Pesenti, and LeCun shows, machine intelligence cannot be constructed through the accumulation and interpolation of past examples. The curse of dimensionality acts as a paradoxical constraint: mathematically it renders imitation extremely rare and forces extrapolation as the default mode of generation.

The contemporary architectural automaton must therefore be understood neither as a mechanical executor of predefined rules nor as a stochastic imitator of historical precedents. It functions instead as a constrained geometric explorer, navigating spaces so vast and sparse that novelty becomes an almost structural necessity rather than an intentional choice. Autonomy does not stem from consciousness, intention, or will, but from the necessity of bridging gaps between scattered data points. Architectural intelligence emerges from the mathematical topology of the spaces in which the machine operates, as it fills the immense void of latent space through extrapolation. The intelligence of the automaton is thus neither human nor post-human; it is geometric. Condemned to operate where imitation fails, it is structurally driven to invent.

⁸⁴ *Khōrein: Journal for Architecture and Philosophy*, “Call for Papers: Volume IV, Issue 3: Automata.”

REFERENCES

- Arnheim, Rudolf (1954), *Art and Visual Perception: A Psychology of the Creative Eye*, Berkeley and Los Angeles: University of California Press.
- Aureli, Pier Vittorio (2023), *Architecture and Abstraction*, Cambridge, Mass.: MIT Press.
- Balestriero, Randall, Jérôme Pesenti, and Yann LeCun (2021), "Learning in High Dimension Always Amounts to Extrapolation," *arXiv*, <https://doi.org/10.48550/arXiv.2110.09485> (accessed April 8, 2026).
- Birkhoff, George D. (1933), *Aesthetic Measure*, Cambridge, Mass.: Harvard University Press.
- Boden, Margaret (1991), *The Creative Mind: Myths and Mechanisms*, New York: Basic Books.
- Bostrom, Nick (2014), *Superintelligence: Paths, Dangers, Strategies*, Oxford: Oxford University Press.
- Carmo, Mario (2011), *The Alphabet and the Algorithm*, Cambridge, Mass.: MIT Press.
- Carmo, Mario (2024), "Generative AI, Imitation Learning, and the Automation of Tacit Knowledge," *A+U*, 646, pp. 104–109.
- Carmo, Mario (ed.) (2012), *The Digital Turn in Architecture 1992–2012*, Chichester: Wiley.
- Chatterji, Niladri S., Philip M. Long, Peter L. Bartlett (2021), "When Does Gradient Descent with Logistic Loss Find Interpolating Two-Layer Networks?," *Journal of Machine Learning Research*, 22, 159, pp. 1–48.
- Chomsky, Noam, Ian Roberts, Jeffrey Watumull (2023), "The False Promise of ChatGPT," *The New York Times*, March 8, <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html> (accessed April 8, 2026).
- Copeland, B. Jack (2004), "The Church-Turing Thesis," *NeuroQuantology*, 2, 2, pp. 101–115.
- Craik, K. J. W. (1943), *The Nature of Explanation*, Cambridge: Cambridge University Press.
- Dieudonné, Jean (1964), *Algèbre linéaire et géométrie élémentaire*, Paris: Hermann.
- Eco, Umberto (1976), *A Theory of Semiotics*, Bloomington/London: Indiana University Press.
- Eisenman, Peter (2004), "miMISes READING: It Does Not Mean a Thing," in *Eisenman Inside Out: Selected Writings, 1963–1988*, New Haven: Yale University Press, pp. 189–201.
- Eisenman, Peter (1984), "The End of the Classical: The End of the Beginning, the End of the End," *Perspecta*, 21, pp. 155–173.
- Franklin, James (2011), "Science Wars: Last Round?," *The New Criterion*, 29, 9, p. 68.
- Frazer, John (1995), *An Evolutionary Architecture*, London: Architectural Association.
- Gero, John S. (ed.) (1985), *Design Optimization*, Orlando: Academic Press.
- Goldstine, Herman H. (1980), *The Computer from Pascal to von Neumann*, Princeton: Princeton University Press.

- Hardingham, Samantha (ed.) (2016), *Cedric Price Works 1952–2003: A Forward-Minded Retrospective*, 2 vols., London and Montreal: Architectural Association and Canadian Centre for Architecture.
- Hartley, Ralph V. L. (1928), “Transmission of Information,” *The Bell System Technical Journal*, 7, pp. 535–563.
- Khôrein: Journal for Architecture and Philosophy* (2026), “Call for Papers: Volume IV, Issue 3: Automata,” <https://khorein.ifdt.bg.ac.rs/index.php/ch/announcement/view/22> (accessed April 8, 2026).
- Kleene, Stephen C. (1951), “Representation of Events in Nerve Nets and Finite Automata,” *RAND Research Memorandum RM-704*, Santa Monica: RAND Corporation.
- Kotsopoulos, Sotirios (ed.) (2025), *Shape Computation: Fifty Years, 1972–2022*, Cham: Birkhäuser.
- Lassègue, Jean, Giuseppe Longo (2024), *L’empire numérique*, Paris: Humensis.
- Lynn, Greg (1994), “New Variations on the Rowe Complex,” *ANY: Architecture*, New York, 7–8, pp. 38–43.
- McCarthy, John (1951), *Projection Operators and Partial Differential Equations*, PhD dissertation, Princeton University.
- McCarthy, John, Marvin L. Minsky, Nathaniel Rochester, Claude E. Shannon (2006), “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence,” *AI Magazine*, 27, 4, pp. 12–14.
- McCulloch, Warren S., Walter Pitts (1943), “A Logical Calculus of the Ideas Immanent in Nervous Activity,” *Bulletin of Mathematical Biophysics*, 5, pp. 115–133.
- McCulloch, Warren S., Walter Pitts (1947), “How We Know Universals: The Perception of Auditory and Visual Forms,” *Bulletin of Mathematical Biophysics*, 9, pp. 127–147.
- Minsky, Marvin L. (1954), *Theory of Neural-Analog Reinforcement Systems and Its Application to the Brain-Model*, PhD dissertation, Princeton University.
- Moles, Abraham A. (1972), *Information Theory and Aesthetic Perception*, Paris: Denoël.
- Negroponte, Nicholas (1970), *The Architecture Machine: Toward a More Human Environment*, Cambridge, Mass.: MIT Press.
- Negroponte, Nicholas (1975), *Soft Architecture Machines*, Cambridge, Mass.: MIT Press.
- Nyquist, Harry (1924), “Certain Factors Affecting Telegraph Speed,” *The Bell System Technical Journal*, 3, 2, pp. 324–346.
- Nyquist, Harry (1928), “Certain Topics in Telegraph Transmission Theory,” *Transactions of the American Institute of Electrical Engineers*, 47, 2, pp. 617–644.
- Oosterhuis, Kas (2002), *Programmable Architecture*, Milan: L’Arca Edizioni.
- Oosterhuis, Kas (2003), *Hyperbodies: Towards an E-motive Architecture*, Basel: Birkhäuser.
- Quine, W. V. O. (1960), *Word and Object*, Cambridge, Mass.: MIT Press.
- Rondelet, Jean-Baptiste (1812–1817), *Traité théorique et pratique de l’art de bâtir*, Paris: Published by the author.

- Shannon, Claude E. (1938), "A Symbolic Analysis of Relay and Switching Circuits," *Transactions of the American Institute of Electrical Engineers*, 57, pp. 713–723.
- Shannon, Claude E. (1948), "A Mathematical Theory of Communication," *The Bell System Technical Journal*, 27, pp. 379–423, 623–656.
- Spuybroek, Lars (2009), *Research & Design: The Architecture of Variation*, London: Thames & Hudson.
- Stiny, George, James Gips (1972), "Shape Grammars and the Generative Specification of Painting and Sculpture," in C. V. Freiman (ed.), *Information Processing 71*, Amsterdam: North-Holland, pp. 1460–1465.
- Stiny, George, James Gips (1978), *Algorithmic Aesthetics: Computer Models for Criticism and Design in the Arts*, Berkeley: University of California Press.
- Szambien, Werner (1984), *Jean-Nicolas-Louis Durand 1760–1834: De l'imitation à la norme*, Paris: Picard.
- Turing, Alan M. (1937), "On Computable Numbers, with an Application to the Entscheidungsproblem," *Proceedings of the London Mathematical Society*, 2, 42, pp. 230–265.
- Turing, Alan M. (1950), "Computing Machinery and Intelligence," *Mind*, 59, 236, pp. 433–460.
- Turing, Alan M. (2004), "Intelligent Machinery," in B. Jack Copeland (ed.), *The Essential Turing*, Oxford: Clarendon Press, pp. 410–432.
- von Neumann, John (1945), *First Draft of a Report on the EDVAC*, Philadelphia: University of Pennsylvania.
- von Neumann, John (1948), *Theory and Techniques for Design of Electronic Digital Computers*, Philadelphia: University of Pennsylvania.
- Whitehead, Alfred North (1925), *Science and the Modern World*, London: Cambridge University Press.
- Wiener, Norbert (1948), *Cybernetics: Or Control and Communication in the Animal and the Machine*, Paris: Hermann.

