Philosophy & Technology (2025) 38:46
https://doi.org/10.1007/s13347-025-00875-8

COMMENTARY

Generative Artificial Intelligence: A Concept in Progress

Francesco Bianchini[1]

Received: 24 March 2025 / Accepted: 25 March 2025

Abstract
Each technology advances at its own pace, often indifferent to theoretical and philosophical-scientific conceptualizations. In the case of technologies like artificial intelligence (AI), and especially generative AI (GenAI), developments are so rapid that conceptual and epistemological reflections struggle to keep up, even at the level of basic definitions. Yet these definitions carry significant non-theoretical implications, including social, legal, and policy-related consequences. In this paper, I offer some reflections on the definition of GenAI proposed by Ronge et al. (2025), using it as an opportunity to highlight the most relevant aspects of the ongoing debate sparked by these new AI systems. In doing so, we will seek to explore the implications of generative AI systems in the context of human interaction, particularly in light of their role as active supports rather than passive tools. This shift, coupled with the increasingly anthropomorphic perception of their capabilities, sets GenAI systems apart from any previous technologies, not only in terms of technical features but also in how they are perceived by everyday users.

Keywords Generative artificial intelligence · Philosophy of artificial intelligence · AI systems · Interaction

Conceptual clarification has long been a central task of philosophy, especially since the first half of the twentieth century, when the ideal of a systematic philosophy gave way, partially, to a view of philosophy as a tool for supporting the creation of objective scientific knowledge and reflecting an effective and unambiguous reality. This legacy, one of the most significant contributions of logical empiricism, finds new expressions today, such as in the practice of conceptual engineering. This emerging field is dedicated to evaluating and restructuring the concepts used across various

✉ Francesco Bianchini
francesco.bianchini@unibo.it

[1] Department of Philosophy, University of Bologna, Bologna, Italy

Published online: 07 April 2025

theoretical domains (Chalmers, 2020), though typically within the broader scope of philosophy. At the same time, this task of clarification and (re)structuring, grounded in its philosophical-scientific roots, aligns well with contemporary philosophical approaches that engage with conceptual challenges arising from the technological world and emerging application technologies. These domains often lack clear constitutive elements and well-defined conceptual boundaries, in part due to their rapid evolution through the continuous interplay between theoretical development and practical application. As a result, they require a form of conceptual "systematization" to help delineate their limits and refine their use.

In their article *Towards a Definition of Generative Artificial Intelligence*, Ronge et al. (2025) undertake such a conceptual operation by examining recent technologies and analyzing how they are discussed both within specialized academic circles and in broader public discourse. Their goal is to offer a definition and conceptualization of generative systems in artificial intelligence (AI). The authors highlight that the very notion of generative AI systems remains unclear and inconsistent, marked by ambiguities, including, and especially, in relation to their technical aspects. These systems are not always equipped with the features that properly characterize generative models in the strict sense; and conversely, not all systems labeled as generative AI exhibit those technical properties. This is despite the widespread use of the "generative" label, which often extends well beyond the technical characteristics of the implemented model in relation to the system being used.

The operation of conceptual clarification is both interesting and valuable, as it helps to fill a conceptual void that extends, often problematically, across the various domains to which technology is connected, beyond the theoretical and epistemic principles guiding its design and implementation. This includes, for example, the social, ethical, political, and legal domains. This need is particularly pressing in the context of AI systems, where the rapid pace of change and the ongoing acquisition of new capabilities compound the already complex theoretical understanding that has characterized these systems since their inception (Cordeschi, 2002). Therefore, if there exists a discrepancy between technical discourse and public discourse, and if in the technical sphere the label "generative" emerges relatively late and is used differently than in the public sphere, this may be for entirely valid reasons. One such reason could be that those who originally developed the techniques that later formed the basis of generative systems did so without any explicit "generative" intent. This observation further strengthens the authors' argument by opening the possibility that *generativity* is something distinct from the techniques employed to achieve it. It may be understood as a (partially) contingent conceptual framework that we apply retrospectively to characterize a significant portion of today's AI systems. At the same time, and in a closely related way, this discrepancy also highlights the difficulty even technical developers may have in fully grasping certain characteristics of AI systems. Such a lack of awareness can carry serious risks, particularly regarding the oversight and control that ought to be exercised over these technologies, beginning with their theorists and implementers, and extending to third-party entities responsible for regulation in political and social contexts.

The distinction between public discourse and technical discourse significantly impacts various technological fields, particularly artificial intelligence (AI). This gap

is steadily growing as AI systems become increasingly widespread and accessible to non-experts, even those with no background in the field. One key reason for this discrepancy is the ease of use of generative AI (GenAI), which contributes to widening the divide between technical design, theoretical reflection, and public discussion. This results in a sort of paradox: technical aspects appear to play little role in shaping public discourse about these systems. This trend is continually reinforced by the rapid and repeated adoption of new AI technologies, with user numbers growing exponentially. Yet, this widespread diffusion does not seem to influence technical discourse, which remains focused on different generative modeling approaches. It often overlooks the functional needs of users, or what users themselves consider important in the use and performance of GenAI systems. All of this is unfolding in a context where public discussion about new technologies increasingly takes place not only in non-AI scientific journals but also in the media, which have become key arenas for public debate (Winkel, 2024).

The discrepancy between public discourse and technical discourse around GenAI often stems from users' perceptions, regardless of their level of expertise. This gap is amplified by the fact that one does not need to understand the technical workings of a GenAI system to use its functionalities. As with many complex yet widely adopted technologies, the underlying mechanisms remain largely invisible to the user. What sets GenAI apart from earlier AI systems is that it requires far less theoretical knowledge to operate. Traditional AI applications, such as problem-solving programs, expert systems, or classical neural networks used for classification tasks, often demanded an understanding of their internal logic to interpret their outputs correctly. Without this knowledge, users could easily misjudge the capabilities and limitations of such systems. With GenAI, the paradigm shifts significantly. What matters most is not how the system works internally, but how to interact with it effectively, namely, how to craft appropriate prompts. In many cases, the output itself, such as a natural language text, is immediately accessible as long as it is syntactically correct and semantically coherent. The focus then moves to evaluating the semantic informativeness and epistemic reliability of the output, rather than understanding the system's internal problem space or how it performs categorical clustering (Bock, 1998). While it may be difficult to define what a generative model is in broad technical terms, that question becomes less central. The nature of GenAI's outputs, distinct from those of previous AI systems, shifts the conversation. As a result, the meaning of "generative" in public discourse often drifts away from its technical roots, centering instead on the character and function of the outputs themselves.

If the focus truly shifts to the user's perception this inevitably creates a divide between technical description and public discourse. In this perspective, such a divide brings about some interesting consequences, particularly regarding how much generativity a system possesses, demonstrates, or is perceived to produce. These three dimensions, actual, exhibited, and attributed generativity, have distinct implications and redirect the discussion toward the role of the user in defining a GenAI system. In other words, we are prompted to consider how much responsibility should be assigned to users' decisions and evaluations in relation to GenAI systems. In any case, perceptions of too much or too little generativity tend to be attributed to the system's behavior and output, rather than to its technical components. This leads to

a fascinating and closely related philosophical issue: the attribution of generativity, akin to earlier debates around the attribution of intentionality or intelligence in artificial systems (Dennett, 1987). Moreover, this perspective detaches the description of a GenAI system from its technical characteristics. And this shift can, in fact, be generalized today to other types of AI systems, such as those used for problem-solving, planning, and more.

In their 2025 contribution, Ronge, Maier, and Rathgeber outline, albeit not exhaustively, four dimensions that may serve to define GenAI within non-technical discourse. These dimensions– multimodality, interaction, flexibility, and productivity– are proposed to capture the conceptual richness GenAI has come to embody, surpassing the boundaries of its more narrowly construed modeling counterpart. Crucially, these aspects must be understood in their mutual interdependence, as each, taken in isolation, may already pertain to various forms of AI. For example, numerous AI systems are interactive, and some exhibit a degree of flexibility in task performance. Within this set of features, however, certain elements appear to assume greater definitional weight. Flexibility, in particular, acquires notable significance insofar as it evokes anthropomorphic modes of engagement with users, who, as human agents, tend to ascribe to the system attributes that are intelligible and operative within human interaction. This dynamic is essential to the notion of generativity, understood here not merely as the capacity to produce outputs, but as the ability to sustain an iterative and self-applicative mode of operation: a GenAI system can recursively process its own outputs in an open-ended fashion. This feature arguably constitutes a more distinctive marker of GenAI than interaction alone, which remains a common characteristic of many AI systems. It is important to note that self-application, in this context, should not be conflated with classical forms of self-reference or recursion, as found for instance in certain neural architectures. Rather, it represents a more peculiar property of GenAI systems, insofar as self-application is mediated through user interaction, and is thus, in a sense, externalized, a process that occurs not solely within the internal logic of the system, but within a dialogical or participatory framework.

Multimodality is another feature that evokes both the theme of anthropomorphism and the broader question of Artificial General Intelligence (AGI). Although GenAI is typically understood as comprising systems that primarily generate textual, visual, or musical outputs, a more substantial convergence with AGI would require that GenAI systems also be analyzable in terms of problem-solving capabilities and the production of broad, goal-directed behaviors. In this light, the ability to navigate diverse tasks and contexts, hallmarks of AGI, would need to be integrated into the GenAI paradigm. Among the four dimensions proposed, productivity arguably appears to be the most tangential. While it is indeed relevant when considered in conjunction with the other features, it is also true that many traditional AI systems have long contributed to increased productivity across both personal and professional domains. This trend continues with GenAI systems, and what arguably sets them apart is the widespread perception of enhanced productivity, a perception that has quickly become entrenched in public discourse. However, rather than emphasizing productivity per se, it may be more fruitful to consider the autonomy of these systems. Although autonomy is a general characteristic of AI, it has gained particular prominence over the past three decades, especially in the context of robotics, neural networks, and

other forms of so-called black-box AI (Steels, 1995). GenAI systems exhibit a relatively advanced degree of autonomy, insofar as they are often capable of producing outputs that require little to no human intervention or adjustment. This feature is arguably just as relevant to their definition as productivity, if not more so, as it reflects a more intrinsic property of the system, regardless of whether one conceives of autonomy in scalar terms.

Finally, it is worth emphasizing that, in general terms, GenAI refers to a class of systems that share numerous features with earlier forms of AI, features which are now being perceived or reinterpreted in new ways, in what might be described as a kind of *exaptation of the tool*, emerging from the interaction between human users and the system. This shift in perspective does not arise from the fact that these systems are tailor-made for human beings; rather, it likely stems from the opposite, namely, their openness to a wide range of performances of a different, almost adaptive, nature. The GenAI system no longer functions merely as a passive instrument to be used, but as *an interactive agent in active support*. This reconceptualization carries with it significant implications, including the necessity of an ongoing definitional effort. Such an effort will likely extend beyond the current stage of development, evolving in tandem with the systems themselves and with the shifting perceptions of their actual and potential functions. This process unfolds along a continuum of interactive interpenetration between human agents and intelligent artificial systems.

Acknowledgements No.

Author Contributions 100% the only author.

Funding Open access funding provided by Alma Mater Studiorum - Università di Bologna within the CRUI-CARE Agreement.

Data Availability Not applicable.

Declarations

Competing Interests No.

References

Bock, H-H. (1998). Clustering and Neural Networks. In: A. Rizzi, M. Vichi, H-H. Bock (Eds.), *Advances in Data Science and Classification. Studies in Classification, Data Analysis, and Knowledge Organization* (pp. 265–277). Springer. https://doi.org/10.1007/978-3-642-72253-0_37

Cordeschi, R. (2002). *The Discovery of the Artificial. Behavior, Mind and Machines Before and Beyond Cybernetics*, Kluwer Academic Publishers.

Dennett, D. C. (1987), *The intentional stance*, MIT Press.

Chalmers, D. J. (2020). What is conceptual engineering and what should it be? *Inquiry*, 1–18. https://doi.org/10.1080/0020174X.2020.1817141

Ronge, R., Maier, M. & Rathgeber, B. (2025). Towards a Definition of *Generative Artificial Intelligence. Philosophy and Technology* 38, 31. https://doi.org/10.1007/s13347-025-00863-y

Steels, L., (1995). When are robots intelligent autonomous agents? *Robotics and Autonomous Systems*, 15, pp. 3–9,. https://doi.org/10.1016/0921-8890(95)00011-4

Winkel, M. (2024). Controlling The uncontrollable: The public discourse on artificial intelligence between The positions of social and technological determinism. *AI & Society*. https://doi.org/10.1007/s00146-024-01979-z