

The Enlightenment Reborn

The Honorable Xiąże John McClellan Marshall*¹, Dr. Roger F. Malina²

¹Senior Judge, Fourteenth Judicial District of Texas; Honorary Professor of the University, UMCS, Lublin, Poland, Academician, International Academy of Astronautics.

²Distinguished Professor of Arts and Technology and Physics; University of Texas at Dallas; Academician, International Academy of Astronautics.

***Correspondence:** The Honorable Xiąże John McClellan Marshall, Senior Judge, Fourteenth Judicial District of Texas; Honorary Professor of the University, UMCS, Lublin, Poland, Academician, International Academy of Astronautics.

Received: September 19, 2026 | **Published:** October 09, 2026

How to cite this article:

Marshall, J.M. and Malina, R.F. (2026) 'The Enlightenment Reborn', *Journal of Artificial Intelligence and AI Ethics*, 1(4), pp.1–4. <https://doi.org/10.64978/jaiae.2026.1009008>

“We are not meant to view the world through machines” ~ “Surrogates”

Abstract

In this article the authors recognize that the rapid adoption of artificial intelligence and virtual reality has created an epistemic crisis. This requires a renewal of the Enlightenment commitment to empirical inquiry, critical reason and human responsibility. An historical-conceptual analysis, supported by illustrative cases from law and medicine, demonstrates how technological efficiency is increasingly confused with intelligence, judgment and wisdom. In common parlance, the terms such as “artificial intelligence” and “virtual reality” obscure the dependence of these technologies on human-designed mathematical systems, selected data, embedded biases, and potentially vulnerable infrastructures. The central problem, however, lies not solely in the machine but in the man-machine interface, i.e. users may abdicate judgment without adequately examining the origins, limitations or accuracy, and possible hallucination, of the information produced. While computational systems can search extensive bodies of data with remarkable speed, they lack the embodied human experience, empathy, moral intuition, and practical wisdom required for reliable consequential decisions. Only a human can make the final decisions, and it is in the retention of the human factor in this man-machine relationship that the Enlightenment concept of the application of human intellect to the analytical process comes to the fore. The beneficial future of artificial intelligence will therefore depend upon preserving human critical agency and treating machines as instruments of inquiry rather than substitutes for judgment.

Keywords: wisdom, cyberefficiency, 18th Century Enlightenment, moral intuition

1. Introduction

The terms “artificial intelligence” and “virtual reality”, when properly defined are, in truth, oxymorons. As unflattering as it may seem to some, the term “oxymoron” first appeared in English during the 17th Century, as the Enlightenment was gaining momentum in the West. Literally, it translates from the Greek as “sharp, but foolish”, and it describes self-contradictory words that are paired. In the context of the modern day, two such pairings are frequently heard and accepted as if they had an underlying reality that validates their existence: “artificial intelligence” and “virtual reality”. In the former case, the definition of “artificial intelligence” necessarily implies that it does not relate to any reality. Indeed, something either is intelligent or it is not, and an inanimate object does not possess inherent intelligence. Similarly, “virtual reality” is not only self-contradictory, but is a definitional denial of reality. “Virtual” removes its object from being authentic to being merely a construct that imitates some aspect of the real world. In both cases, these concepts in modern parlance are widely accepted and used as if they were objectively expressive of some

reality. It is the popular perception of modern cyberefficiency as if it equates to some level of wisdom that is at the root of the problem. In fact, this view of AI and virtual reality is antithetical to the bases of Western thought as formed in the 17th and 18th Centuries’ Enlightenment because they deny the creative and analytical actions of the human intellect in observation of the world around us.

Dealing with these concepts as mere wordplay simply misses the point that they are an Orwellian newspeak substitute for an analytical thought process. As such, the very vagueness of the terms as applied in almost any context dilutes the ability of the hearer to grasp what is being conveyed. In order to remedy this situation, it is incumbent upon the listener who might wish to utilize some platform labeled “AI” to engage in a protracted analysis of the origins of the platform. This would include such things as the mathematical precision, security provisions, and, at its root, the biases of the author of the platform. Recent surveys indicate, for example, that analyses of some articles in *Wikipedia* and Google searches disclose a left-of-center orientation that tends

to skew the “factual” content of the article. The researcher who might use those sources would need to be careful about any bias that is revealed.

The 17th and 18th Centuries intellectual and philosophical growth commonly referred to as “The Enlightenment” developed as a clear reaction to the medieval pattern of a dogmatic approach to analysis of the world. The shift in emphasis was from a “traditional” mode of thinking to a more “empirical” examination of scientific phenomena (Gay, 1966). This was not merely a matter of a philosophical shift. Intellectually it moved from a focus on the Aristotelian view of the universe to the humanistic context characterized by Hobbes and Locke in England and Galileo in Italy. Perhaps the “bridge” in this process was Copernicus who led the shift away from the Ptolemaic model to the heliocentric model that he articulated in *De revolutionibus orbium coelestium*. This work was the product of his extensive observations and, while started in 1532, was not published in its final form until 1543, following his death in 1542. Nevertheless, his thought impacted the scientific approach of Galileo and similarly fostered a philosophical expansion into the social sciences and politics that ran into the late 17th Century.

At the heart of The Enlightenment was the notion that abstract philosophy alone, such as Aristotle, would not suffice as a basis for scientific or practical humanistic endeavors. No matter what “science” might say, it was up to human beings to analyze the premises in order to reach valid conclusions about what is “real” and what is not. This was clearly depicted in the progression in political thought from Locke to Montesquieu to Rousseau to Jefferson. Their work not only changed the concept of government, but the relationship of citizens to their government as well, ushering in the development of “representative democracy”. Indeed, it can be posited that the clearest example of The Enlightenment as a political phenomenon is the Constitution of the United States which set a model for many successor changes in governmental polity elsewhere (May, 1976; Reck, 1991).

Similarly, the period was characterized by an expansion of the information base through the mechanism of translation from one language to another of basic materials. A typical example of this was the work of Émilie du Châtelets (1706-1749). She was a French mathematician and philosopher/physicist whose 1740 treatise *Institutions de Physique* [The Foundations of Physics], became one of the leading texts on the subject, so much so that it was translated into German and Italian. On her own, she translated *Newton's Philosophiæ Naturalis Principia Mathematica* into French with such skill that it remains the definitive French text to the modern day. Of course, her relationship with Voltaire was also characteristic of the major figures of the 18th Century Enlightenment. The point is that during this period there was, relatively speaking, an expansion both of horizon and expression of thought that transcended nationalities and marked a definitive break with the pre-Renaissance way of thinking.

Beginning about two decades ago, a similar philosophical break began, first as a way of defining “reality” and then, more recently with the aid of technology, restructuring that process. The verbal descriptive mechanism for this transition generally is described as “artificial intelligence” or “virtual reality” with little critical thought as to what those terms imply. Chronologically, the transition becomes most pronounced in the period during and following the COVID pandemic in part because of the breakdown of a society composed of integral units. For example, in the legal community it became frequent, if not a norm, to have remote hearings and even trials. Of necessity, some evidence that normally would have

been produced in “hard copy” was offered digitally for the sake of convenience. In the case of expert testimony, exhibits created in “virtual reality” to assist the judge or jury, were sometimes created with little or no deep attention to detail. As a result, the outcome of the hearing or trial was potentially skewed by inaccuracy. Similarly, the creation of ChatGPT provided lawyers with an enormously expanded dataverse to be searched by the chatbot. This combined sometimes with human laziness to produce completely factitious results. For example, in one case in New York federal court, the lawyers presented a brief that contained totally fictitious case citations. When the judge confronted the lawyers and gave them a chance to review the brief, the lawyers simply resubmitted the same documents. The result was that the judge levied a fine on the lawyers and required them to write letters to the judges who had been falsely cited apologizing (Mata v. Avianca, 2023, cited in Law360, 2023). In a recent case in Texas, an attorney informed the court that the parties to the lawsuit had been using AI and ChatGPT in the preparation of some of their materials. He referred to this situation as an “abomination” (Letter to the author). This reaction is telling in terms of identifying limits to the “benefits” of AI-bots in the work of some professionals. At a recent judicial conference, three chief justices, two former and one sitting, of the Supreme Court of Texas expressed caution in the use of AI-bots in the preparation of pleadings before the court. One of the comments was to the effect that, if lawyers do not review the product of the AI-bot, then the law is at risk of becoming interpreted by machines instead of people. Of equal importance, of course, is the impact on the uncritical use of generative AI in the judicial process simply to get an answer. Beyond the issue of possible “hallucination” by the AI-bot that creates the appearance of authenticity, there is the possibility that its product is in fact fictitious and, therefore, logically not relevant to the ability of the court to adjudicate issues and grant relief to the litigants before it. Such generative AI pleadings might as well be a Christmas card signed by Donald Duck and, therefore, of no significance to the court.

During the same period, the expansion of technology began to overtake what otherwise might have been a more orderly evolution of AI and virtual reality. Software engineers and creators experimented with the notion that AI, because of its ability to reach out to the dataverse, potentially could help solve a wide range of problems to the benefit of mankind. Further, the speed with which the dataverse was accessed made the AI-bot even more important as a research tool. In particular, the medical community saw the opportunity for quicker, and therefore more “accurate” diagnoses of injuries and illnesses, especially in overcrowded ER environments, leading to the development of “tele-medicine”. The results, however, has not been as uniformly helpful as anticipated. In April 2026, a young dental student was checked into a “tele-ICU” where he was diagnosed by an off-site nurse using AI, based upon his symptoms, with “pancreatitis, dehydration, metabolic acidosis, and alcohol withdrawal.” Despite a worsening condition over the next 18 hours, he was never examined in person by an on-site physician or nurse. He was pronounced dead by a remote practitioner (Shultxm, 2026).

Further, with patients having increased access to medical AI-bots, self-diagnosis can be at an increased risk of inaccuracy. For example, a middle-aged patient in Seattle used a generative AI-bot called “Perplexity” to research various treatments for leukemia. The AI-bot provided erroneous information, upon which the patient relied to refuse the treatment suggested by his oncologist. Predictably, he died in late 2025. This episode exposes a flaw in generative AI, namely its ability to produce answers that sound authoritative, but which are in fact unreliable when

examined closely (Gordon, 2026). An interesting exposé as to the machine-man interface arose in a Brazilian hospital patient death in 2026. The critically ill patient, herself a psychologist, was awaiting transfer to the ICU was reviewed and downgraded over the protests of the hospital staff by the algorithms in the hospital administration AI-bot. As a result of the five-day delay she died (Patel, 2026).

Part of the problem, particularly in the medical context, is that the AI-bot does not have a body and, therefore, cannot relate directly to such things as pain, anxiety, or “feelings”, leading to valid response to an inquiry by a trained physician. For example, in the recent case of a surrogate mother who refused to obtain an abortion sought by the biological parents of the child, there were both legal and moral issues at work in the decision process relative to whether the child would be born or aborted. While there was no question that the child had a genetic defect that would likely lead to problems later in life, the moral position of the surrogate mother overrode her contractual relationship with the biological parents. She simply moved to a state where “abortion on demand” did not exist and gave birth normally, making arrangements for the child’s condition to be taken care of post-delivery. Similarly, the recent case of a mother who murdered her three children presents a mixture of legal and psychological issues which an AI-bot is unable to address because it has no ability to assess the factors that impact human mentality. Yet, in both cases, human-generated decisions, whether by a jury or by an individual are essential to the ability of society to “move on.” Without the ability to relate directly to the human condition, the scope of validity of the product of AI-bots, and therefore its utility, may be much narrower than might be at first imagined.

What is significant about these illustrations of the machine-man interface in the “new world” of AI-bots is not that the AI-bot is a proactive agent of the change. Instead, they reveal that it is the action or inaction of the human user/observer that is the active agent of error. While there is little question that the AI-bot, whatever the platform, has the ability to search the dataverse both more broadly and more rapidly than a human being, there is a latent self-deception that human beings can, or should, accept the response uncritically. This fallacy lies in the difference between the machine and the human who is using it. At its most basic, the machine is a creation of human invention that is a prisoner of mathematics. In reality, it can only respond with what the mathematics that define its programming allows, regardless of the questions that may be asked.

That said, it is important to understand that mathematics as a creation of the human mind impacts the perception of reality (Emmer, 1987). The burden on the user, therefore, is to frame the inquiry as narrowly as possible to fit the ability of the platform to obtain a meaningful response. A failure to do so may result in a response that is part fact and part hallucination, leaving the user with the task of sorting out which is which. “Time saving” in such a context is ephemeral at best. Similarly, the machine by its very nature lacks empathy, a rather human characteristic, and its ability to respond to inquiries that require that component is, by definition, flawed, perhaps in the extreme.

To an extent, this is a modern reflection of the dichotomy between objective reality and intuition discussed by philosophers in the 18th Century Enlightenment (Kant, 1781). In relation to the utility of AI-bots, it is crucial to recognize this limitation because it defines the man-machine interface in the modern context. For example, in the case of a medical issue such as whether to recommend the termination of a problematic pregnancy, it would make little

sense to ask an AI-bot for its advice, knowing of its ability to make only mathematically sufficient responses, not morally complex, or intuitive, decisions. Even so, there are examples of just that in which young people ask the machine for advice relative to major life choices, and reliance upon the response can be inaccurate at best, or catastrophically cold at worst. Put another way, while the machine may be able to provide opinions that may be based on possibly plausible facts, it cannot reply with wisdom. Unfortunately, it is too easy to confuse the mechanical abilities of the machine with the creation of wisdom, so that the uncritical user is left with inaccurate, if not “hallucinatory”, responses on which to base decisions that may involve life or death. Put another way, as George Santayana said, “It is wisdom to believe the heart” (Santayana, 1979).

Another problem that has developed in tandem with the broader capabilities of AI-bots and virtual reality is the stripping away of security, both internal and external, as to what is generated and by whom. Internally, the AI-bot, because of its use of the LLM platforms, leaves “footprints” of where it has gone to acquire its data and reveals the intentions of the user. Externally, the use of AI-bots to gather information exposes the user’s platform and machine to teenage hackers and cyberthreats to the platform that is being used and, therefore, its vulnerabilities. The result is that the notion of “secure” communications with outsiders is illusory at best without some proactive access to appropriate encryption programs. Further, for professionals, such as physicians and lawyers, the use of encryption for communications with patients, including simply a hospital’s maintenance of patient information [PHI], or clients becomes essential in order to retain the confidentiality expected by the community. That is the “bad” news.

Fortunately, there is “good” news that is beginning to modify, if not slow, the pace of AI-bot utilization in the modern day. The first sign of this is that the AI-bots are becoming more specialized in their structure and use. For example, an AI platform dedicated to the management of business accounts is not likely to be used for medical diagnosis. The point is that the specialization of AI-bots will likely tend to remove them from “daily” use by unskilled and uncritical users. This means that professionals will need to be trained in those AI-bots that are peculiar to their professions and the security that might be required. They can then check the responses both as to the accuracy of the facts that are presented and as to conformity with the learning and experience of the user. In other words, the proactive insertion of the “human perception and experience” of the user, as manifested in “checking the AI-bot”, hopefully will determine the ultimate utility, or “wisdom”, of the AI-bot product. It is this 21st Century synthesis, combining the human analytical foundation of The Enlightenment and modern technology, that has the potential to expand the usefulness of AI-bots and virtual reality tools to make life for human beings healthier, more productive and more secure.

Disclaimer: This paper was not created, directly or indirectly, by the use of any AI platform, nor is it financially sponsored by any entity other than the authors.

References

- Gay, P. (1966) *The Enlightenment: An interpretation. Vol. 1: The rise of modern paganism*. Knopf.
- May, H.F. (1976) *The Enlightenment in America*. Oxford University Press.
- Reck, A.J. (1991) ‘The Enlightenment in American law II: The Constitution’, *Review of Metaphysics*, 44, p. 723.

-
- Law360 (2023) 'Mata v. Avianca, Case No. 1:22-cv-01461', 23 June.
- Shultxm, C.L. (2026) 'Dental student dies in "fake ICU" . . .', *People*, 8 April.
- Gordon, L. (2026) 'This is where real doctors actually stand . . .', *Parade*, 21 April.
- Patel, V. (2026) 'Brazilian woman dies after AI flags her . . .', *International Business Times*, 15 June.
- Emmer, M. (1987) 'Soap bubbles in art and science: From the past to the future of math art', *Leonardo*, 20(4), pp. 327–334.
- Kant, I. (1781) *Critique of pure reason*.
- Santayana, G. (1979) *The complete poems of George Santayana: A critical edition*. Associated University Press.



Copyright: Copyright © 2026 by John McClellan Marshall and Roger F. Malina. All Rights Reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. To view a copy of this licence, visit <https://creativecommons.org/licenses/by/4.0/>