Qeios

Peer Review

Review of: "Artificial Intelligence (AI), Extended Phenotypes, and the Bio-Evolutionary Anthropocene Hypothesis"

Felipe Arocena[1]

1. Department of Sociology, Universidad de la República, Uruguay

1. Review of the Commentary: *Artificial Intelligence (AI), Extended Phenotypes, and the Bio-Evolutionary Anthropocene Hypothesis*, by Pablo José Francisco Pena Rodrigues

Felipe Arocena

Let us begin by clarifying the title of the article, which may not be immediately accessible to many readers, even to those who have reflected on the implications of the rapid evolution of artificial intelligence systems. The title brings together several dense theoretical concepts, and understanding them is essential in order to grasp the author's argument.

The first concept to consider is that of the **extended phenotype**, introduced by Richard Dawkins in *The Extended Phenotype* (1982). In classical evolutionary biology, the phenotype refers to the observable traits of an organism—its morphology, physiology, and behavior—produced through the interaction between genes and environment. Dawkins proposed that the influence of genes does not stop at the boundaries of the organism's body. Instead, genes can generate effects that extend outward into the environment (Dawkins 1982).

Structures such as a beaver's dam, a bird's nest, or a spider's web are therefore not merely environmental modifications. They can be understood as expressions of the organism's genetic program. In this sense, the phenotype extends beyond the physical organism and into the structures that organisms systematically create in their surroundings.

When this idea is applied to human beings and technology, it leads to the notion of **extended technological-human phenotypes**. The central claim is that human technologies can be interpreted as externalized manifestations of human biological and cognitive capacities. Just as a beaver's dam reflects

 doi.org/10.32388/P4WR07

the activity of beaver genes, technological artifacts can be understood as expressions of human cognitive, social, and cultural capabilities. Technologies such as writing systems, smartphones, the internet, and artificial intelligence systems therefore function not simply as tools external to humanity but as extensions of human capacities themselves.

Within the context of contemporary artificial intelligence, this perspective becomes particularly significant. AI systems are no longer merely neutral instruments occasionally employed by humans. Increasingly, they participate in processes of perception, classification, prediction, and decision-making across many domains of social life. In this sense, they become components of a broader technological-human phenotype.

Because AI systems are trained on human-generated data and shaped by institutional, political, and cultural contexts, they inevitably reflect human assumptions, values, and biases (Crawford 2021). The cultural diversity—or lack thereof—in the design and governance of AI systems therefore has consequences for the evolution of this extended technological phenotype. Technological artifacts are not simply external devices; they become part of the ample expression of human cognitive, social, and cultural capacities, extending the human phenotype beyond the biological organism into technological and institutional environments.

The second concept embedded in the title is the **Anthropocene**, a term popularized by the atmospheric chemist Paul Crutzen to describe a new geological epoch in which human activity has become a dominant force shaping the Earth's systems—its climate, oceans, atmosphere, and sedimentary layers (Crutzen 2002). The **Bio-Evolutionary Anthropocene Hypothesis** pushes this idea further. According to this view, humans are not only transforming the planet's surface; they are increasingly redirecting evolutionary processes themselves.

Evolution, under this perspective, is no longer shaped primarily by natural selection acting within relatively untouched ecosystems. Instead, it is increasingly influenced by anthropogenic selection pressures generated by human technological, economic, and ecological interventions. The Bio-Evolutionary Anthropocene Hypothesis therefore proposes that the Anthropocene represents not only a geological epoch defined by human planetary impact, but also a new evolutionary phase in which human technological and cultural systems actively reshape biological evolution at a planetary scale.

Having unpacked the conceptual density of the article's title, we can turn to the central claims of the commentary itself, particularly its argument concerning the role of artificial intelligence in this new evolutionary context. AI systems are advancing at extraordinary speed. Computational power and

algorithmic capability continue to expand rapidly, following a dynamic historically associated with Moore's Law, whereby computing capacity roughly doubles every two years. AI systems evolve even quicker.

Within this accelerating technological landscape, the author advances several key propositions. He argues that:

"AI systems are a key anthropogenic mechanism and an expression of the extended human phenotype."

"AI systems, humans, and novel biological and technological entities will interact within an evolutionary bias, leading to unexpected outcomes for Earth's evolution."

"AI is us," meaning that these systems should be understood as expressions of our extended phenotype and therefore as extensions of human evolutionary processes."

"Moreover, AGI (Artificial General Intelligence) and ASI (Artificial Superintelligence) may eventually possess the capacity to learn, reason, and make decisions in ways that could trigger even more complex transformations in human societies and in Earth's biological systems—for example through developments such as gene editing, new molecules, or novel pharmaceutical discoveries."

That present and future AI systems will become powerful forces shaping the evolution of both biological and non-biological systems seems increasingly evident. Evolution must now be considered within a radically transformed scenario—perhaps even, as one might tentatively suggest, within a **post-Darwinian landscape**. This does not mean that Darwinian evolution disappears, but rather that human technological agency increasingly becomes one of the dominant forces influencing evolutionary trajectories through artificial selection. Artificial selection—now mediated through technological systems—may in many contexts become as influential as natural selection itself.

The most stimulating aspect of the article lies in its discussion of the interaction between AI systems, human societies, and the environment. The author maintains that contemporary AI systems remain fundamentally human artifacts. They do not yet possess genuine autonomy or independent agency. Their functioning depends entirely on human infrastructures, data flows, and institutional frameworks. Claims that present-day AI systems have already achieved autonomous agency belong, for now, more to the realm of science fiction than scientific reality.

At the same time, the author acknowledges that future developments could alter this situation. As he notes, it is conceivable that future AI systems—perhaps through forms of biological or technological

hybridization—might eventually acquire properties resembling those of living organisms: reproduction, regeneration, persistence over time, and even forms of evolution.

Such systems could profoundly shape the transformations that humans impose on the planet, influencing human decisions and technological trajectories. At present, however, no AI system possesses such characteristics. They remain, in the strict sense, expressions of our extended phenotype. For the moment, humans remain in control. Yet it is not impossible to imagine that this situation could change in the future.

Indeed, the chance that autonomous superintelligent systems might eventually emerge represents one of the major existential risks of our time, alongside climate change and the continuing menace of nuclear destruction. A superintelligent AI capable of independent strategic reasoning might render humans obsolete, marginalize human decision-making, or generate technological dynamics so complex that human societies would struggle to govern them (Bostrom 2014).

The technological powers we have created may therefore become central drivers of planetary evolution—either in interaction with humanity or perhaps, at some point, independently of it. What we know with certainty is that evolution never stops. Geological evolution continues. Biological evolution continues. And now the convergence of bio-technological evolution unfolds alongside them.

The question that remains is a simple but unsettling one: **does anyone still believe that we can simply press the brakes?**

In this sense, the commentary by Pena Rodrigues contributes to an emerging line of thought that seeks to integrate evolutionary theory, technological development, and planetary-scale transformations. By framing artificial intelligence as part of humanity's extended phenotype within the evolutionary dynamics of the Anthropocene, the article invites readers to reconsider the relationship between technological innovation and biological evolution. Whether one agrees with all of its implications or not, the argument usefully situates AI within a broader evolutionary narrative—one in which the future of human societies and the future of life on Earth will become increasingly intertwined with the technological systems we create.

Declarations

Potential competing interests: No potential competing interests to declare.