Journal of Ethics and Emerging Technologies

IEET

Book Review

Review of *Cyborg Rights. Extending Cognition, Ethics, and the Law*

S. Orestis Palermos, *Cyborg Rights. Extending Cognition, Ethics, and the Law*. (Milton Park: UK: Routledge, 2025)

Paolo Capriati [1,*]

[1] University of Palermo; paolo.capriati@unipa.it
* paolo.capriati@unipa.it

Abstract: What if the Hypothesis of Extended Cognition and the Extended Mind Thesis are true? According to S. Orestis Palermos, in *Cybor Rights*, the acceptance of these theses implies a reconsideration of the way in which we understand mental privacy, mental integrity, and personal integrity. However, the reconsideration of these concepts could generate an undue assimilation between our thoughts and our cognitive extensions. The consequences of such an assimilation allow us to see the Hypothesis of Extended Cognition and the Extended Mind Thesis in a different light.

Keywords: extended mind; extended cognition; mental privacy; mental integrity; personal integrity.

Citation: Capriati Paolo. 2026. Review of Cyborg Rights. Extending Cognition, Ethics, and the Law. *Journal of Ethics and Emerging Technologies* 36: 2. https://doi.org/10.55613/jeet.v36i2.246

Received: 02/06/2026
Accepted: 11/06/2026
Published: 01/07/2026

Publisher's Note: IEET stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. About the Hypothesis of Extended Cognition and the Extended Mind Thesis

If the Hypothesis of Extended Cognition (henceforth, HEC) and the Extended Mind Thesis (henceforth, EMT) are true, then we must reconsider the way we understand mental privacy, mental integrity, and personal integrity. This is what S. Orestis Palermos' *Cybor Rights* suggests.

In the first part of the text, Palermos addresses the issue of HEC and EMT. His starting point is, unsurprisingly, *The Extended Mind* by Clark and Chalmers (1998). In the text, the two authors, in order to overcome internalism — that is, the thesis according to which the mind and cognition reside entirely within the individual's brain —, propose "an active externalism, based on the active role of the environment in driving cognitive processes" (Clark, Chalmers, 1998, 7). In this sense, cognition and the mind may even extend beyond the organism itself, also incorporating artefacts. It is for this reason that the suggestive and postmodern image of cyborgs is introduced.

The well-known thought experiment employed by Clark and Chalmers is that of Otto. In order to counteract the negative effects of memory deterioration caused by Alzheimer's disease, Otto always carries a notebook with him. Whenever he needs to record new information, Otto writes it down in his notebook; thus, whenever he needs to retrieve information, he consults the notebook. In this sense, the information stored in the notebook functions in the same way as the information stored in the biological memory of any other person. Otto's example — and, more generally, the extended mind thesis — is vulnerable to two common criticisms: (1) the 'cognitive bloat' worry; (2) the claim that extended beliefs are not genuine beliefs.

Concerning (1), if we admit that mind and cognition transcend the boundaries of the organism, we risk falling into a *reductio ad absurdum*: the entire surrounding environment

could be considered an extension of cognition and mind. The crucial question, therefore, becomes the following: what exactly can count as an extension of cognition and mind? The criterion invoked by Palermos is the one suggested by Beer (1995), who proposes conceptualizing extended cognitive systems in terms of Dynamical System Theory. According to this approach, the criterion for integration is the presence of continuous bidirectional interactions — also known as ongoing feedback loops — between artefacts and agents during the cognitive task (Kaplan 2012).

Criticism (2), by contrast, would concern only EMT and not also HEC. Indeed, in order to justify EMT, Clark and Chalmers suggest the existence of extended beliefs. But can we really consider the notes in Otto's notebook as his dispositional beliefs? According to Palermos, even if we admit that the notes in the notebook are not dispositional beliefs — which are instead necessarily characterized by their informational integration and by being dynamic — we could still conclude that "Otto's mind extends because his process of storing, encoding and retrieving information – his overall process of memory – extends" (Palermos, 2025, 31).

2. Mind-reading and the nature of thought

Having overcome these preliminary problems, the analysis focuses on the following question: what would happen to mental privacy, mental integrity, and personal integrity if Chalmers and Clark were right? These are, in fact, three distinct issues, which are addressed separately in chapters 3 (mental privacy), 4 (mental integrity), and 5 (personal integrity).

According to Palermos, these issues are not merely speculative because, regardless of the theoretical soundness of Clark and Chalmers' theses — and of their profound influence across many fields of knowledge — "cognitive and mental extensions already exist – or will emerge soon" (Palermos, 2025, 18). In any case, according to him, no one doubts that "the minds can – metaphysically speaking – extend" (Palermos, 2025, 18). Therefore, even if it were unlikely that such issues would arise in the present or in the immediate future, it would nonetheless not be absolutely impossible for them to occur.

To demonstrate the urgency and importance of the issues he raises, Palermos, in a rather impressionistic way, introduces the case of mind-reading: in May 2023, functional Magnetic Resonance Imaging (fMRI) was used in combination with AI techniques to read and interpret subjects' brain activity (Tang et al., 2023). This study showed that the technology employed for this kind of mind-reading, although unable to decode thought word by word, nevertheless produced highly accurate results. Further studies subsequently downsized the significance of these discoveries: the exact decoding of a person's thoughts through fMRI techniques would face structural limitations (Brown, 2024, 34). If the idea that someone could read our thoughts understandably disturbs us, then, following Palermos, the possibility that someone could access our digital memories should terrify us even more. Indeed, whereas in the case of mind-reading the technologies currently available (as well as those of the near future) are unable to finely read the content of our thoughts — in addition to the fact that the subject undergoing mind-reading would in any case have to consent to this activity, since it is a lengthy procedure — in the case of autobiographical data stored, for example, on our smartphones, a third party could potentially gain complete access to our memories.

This argument, however, only makes sense under one condition — the very condition upon which Palermos grounds his entire analysis — namely, that we accept EMT and

HEC. Why, however, do these two situations not seem even remotely comparable to us? Why does intrusion into the memories stored in our devices not appear to have the same gravity as mind-reading? Although Palermos attempts to assimilate so-called mental data — by which he means "any dynamic and informationally integrated data that are stored in applications of cognitively integrated technologies" (Palermos, 2025, 44) — to our thoughts, a possible violation of mental data still appears to us as a more acceptable intrusion into the personal sphere than mind-reading.

It is precisely from this differing perception of violation that one can test the solidity of the premises underlying Palermos' argument. On what is this differing perception grounded? Is it perhaps a consequence of the fact that thoughts and mental data possess different kinds of value? But what exactly do we mean by value? There is no doubt that certain mental data may, in terms of utility, possess greater value than some thoughts. Imagine that we have climbed to the top of a mountain and, suddenly, night falls and an intense cold sets in: we absolutely must return to the valley. On the one hand, there is our memory of the route back — a memory that, due to the absence of landmarks in the darkness, may be confused — and, on the other hand, there is the route stored on our smartphone (a precise route that has recorded every step). In such a situation, would you rather lose your memory of the path or find yourself with a dead phone?

However, the question concerning the value of our thoughts is not about utility, but about something else: perhaps their value lies in the fact that our thoughts are not perfectly decipherable. In some way, our attachment to our thoughts is linked to their presumed reluctance to fit perfectly within the narrow boundaries of language. Probably what most disturbs us about mind-reading (beyond the invasion of privacy) is the possibility that our thoughts could be translated exactly into language. This is, in some sense, an ontological limit that we do not want to see overcome. What would happen, indeed, if our thoughts were perfectly expressible through some form of language?

3. The consequences of accepting the Extended Mind Thesis

At present, there seems to be no technique capable of perfectly decoding and reading our thoughts. Not only does such a technique not exist, but perhaps we strongly desire and are convinced that such a technique can never be invented. This perhaps means that we are persuaded that no one can or will ever appropriate our thoughts in the same way that one may appropriate the data stored on our smartphone. The idea that thought exceeds language and cannot be perfectly grasped by it – besides probably constituting the strongest safeguard in support of the irreducibility of the human being to the machine – cannot be applied to mental data: they can be transferred from one device to another precisely because they have already been translated into a language (videos, images, text, music).

But what does this tell us about our initial theses? Although it is undeniable that certain extracranial tools are necessary for cognition, it is clear that mental data and our thoughts belong to two different categories. Even though Palermos does take this fact into account by distinguishing the content of intracranial mental activity and calling it 'neural data' (2025, 6), he nevertheless ends up claiming that "if third parties gain access to technologies of mental extension, they could, for the first time in history, introspect the detailed contents of our minds" (Palermos, 2025, 107). Not only that: he also argues that "if cognitive and mental extension is imminent or already a reality, then access to others' extensions would allow third parties to manipulate the detailed contents of individuals' minds" (Palermos, 2025, 107). The final comparison suggested by Palermos — while also

problematizing it — concerns assaults: "if artefacts extend minds, they also extend persons. Thus (…) intentionally damaging mental and cognitive extensions constitutes assault" (Palermos, 2025, 101). But what does all this imply? Are we really prepared to accept that the same legal consequences should follow both when a subject's mental integrity is manipulated through torture and when such manipulation occurs through actions directed at their mental extensions? Would it not be wiser to keep these two categories distinct? Mind-reading (as well as mental manipulation through techniques of torture) proves to be infinitely more pervasive than the appropriation of data from cognitive extensions (or manipulation through intervention upon cognitive extensions), because it concerns an object (the so-called 'neural data') whose nature is not comparable to that of the content of cognitive extensions.

Ultimately, the merit of Palermos' work lies in the fact that it allows us, retrospectively, to assess the value of two theses — EMT and HEC — on the basis of their consequences. Palermos' genuine and reasonable concerns — namely, how to protect and guarantee mental privacy, mental integrity, and personal integrity in an age in which we have externalized mental activity — risk overshadowing, or failing adequately to consider, a very simple truth: we are not what we write, paint, or play; we are not the information that has been collected about us.

References

(Beer 1995) Beer, Randall D. 1995. A dynamical systems perspective on agent-environment interaction. *Artificial intelligence* 72(1–2): 173–215.

(Brown 2024) Brown, Cohen M. L. 2024. Neurorights, mental privacy, and mind reading. *Neuroethics* 17(2): 1–19.

(Clark and Chalmers 1998) Clark, Andy, and Chalmers, David. 1998. The extended mind. *Analysis* 58(1): 7–19.

(Kaplan 2012) Kaplan, David M. 2012. How to demarcate the boundaries of cognition. *Biology & Philosophy* 27(4): 545–570.

(Palermos 2025) Palermos, S. Orestis. 2025. *Cyborg Rights: Extending Cognition, Ethics, and the Law*. New York: Taylor & Francis.

(Tang et al. 2023) Tang, Jerry et al. 2023. Semantic reconstruction of continuous language from non-invasive brain recordings. *Nature Neuroscience* 26(5): 858–866.