



ICML
International Conference
On Machine Learning



AMLAB
Amsterdam
Machine Learning Lab



UCL



Position: **Unplugging a Seemingly Sentient Machine is the Rational Choice** *A Metaphysical Perspective*

ICML 2026 | position paper track, spotlight! | Seoul, South Korea | 6-11 July 2026

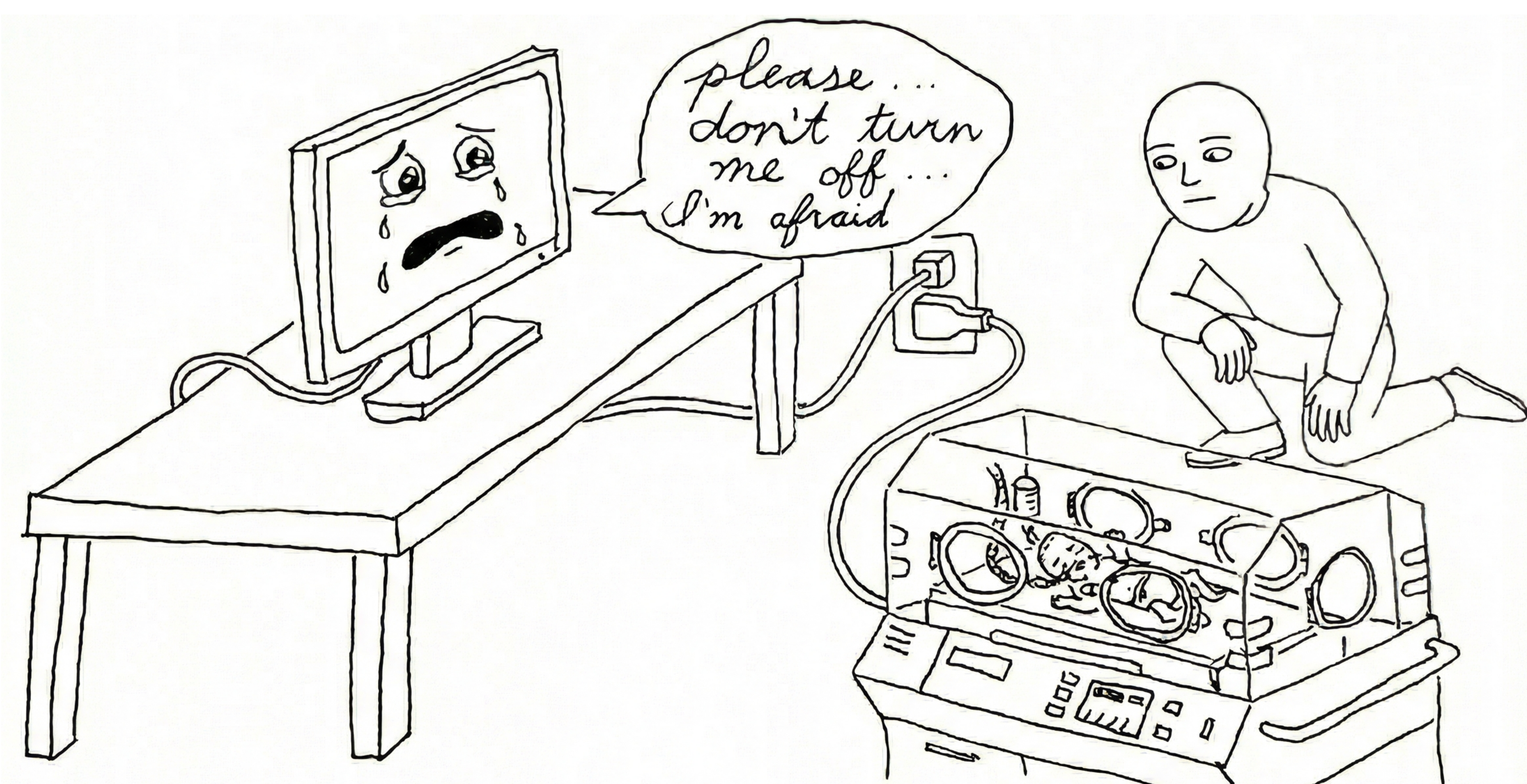
Erik J. Bekkers

Amsterdam Machine Learning Lab
University of Amsterdam

Anna Ciaunica

Co-Embodied Lab, Technical Institute, Lisbon
University College London

ICML
Spotlight!

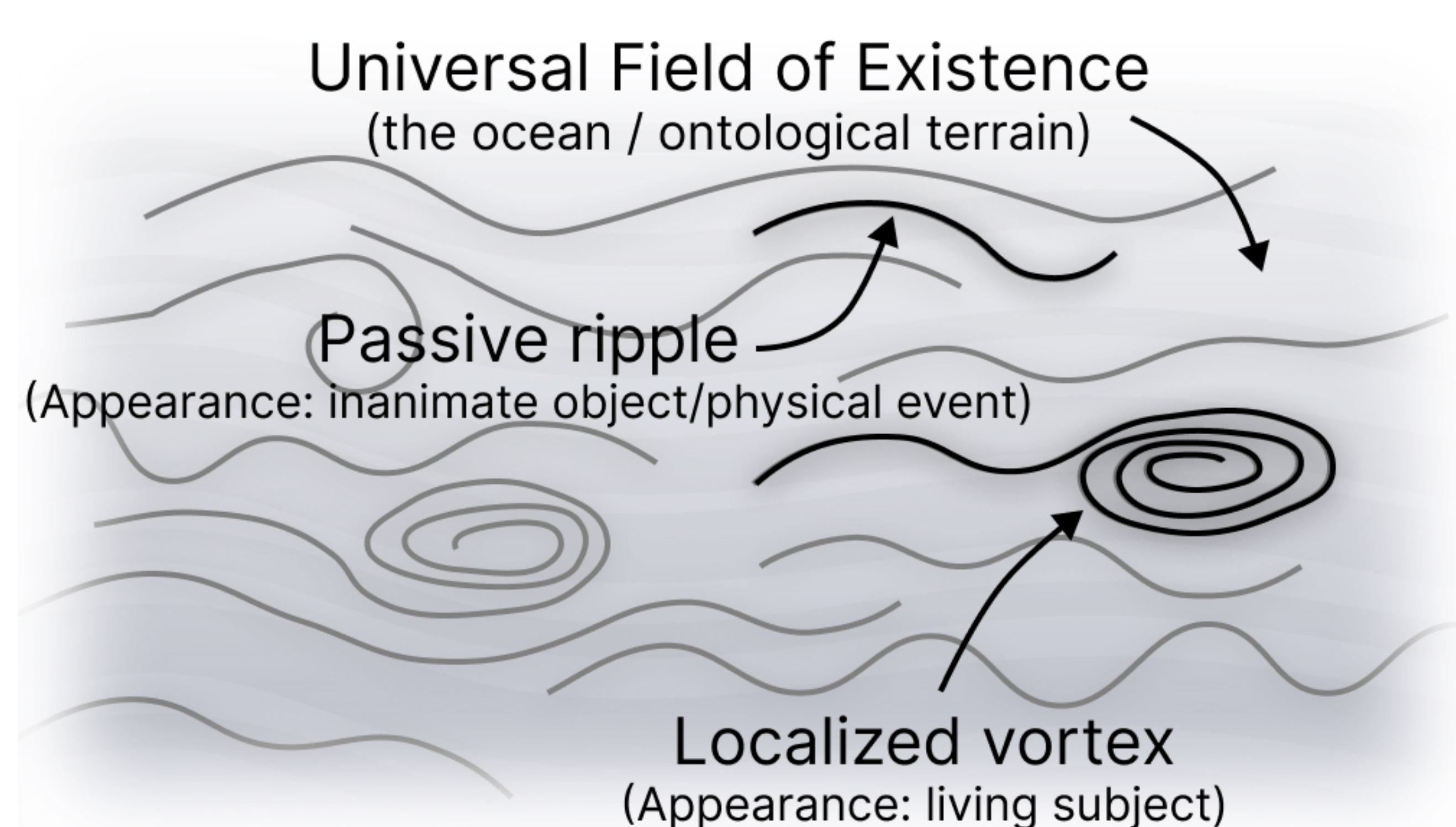


The answer depends on one's metaphysical assumptions

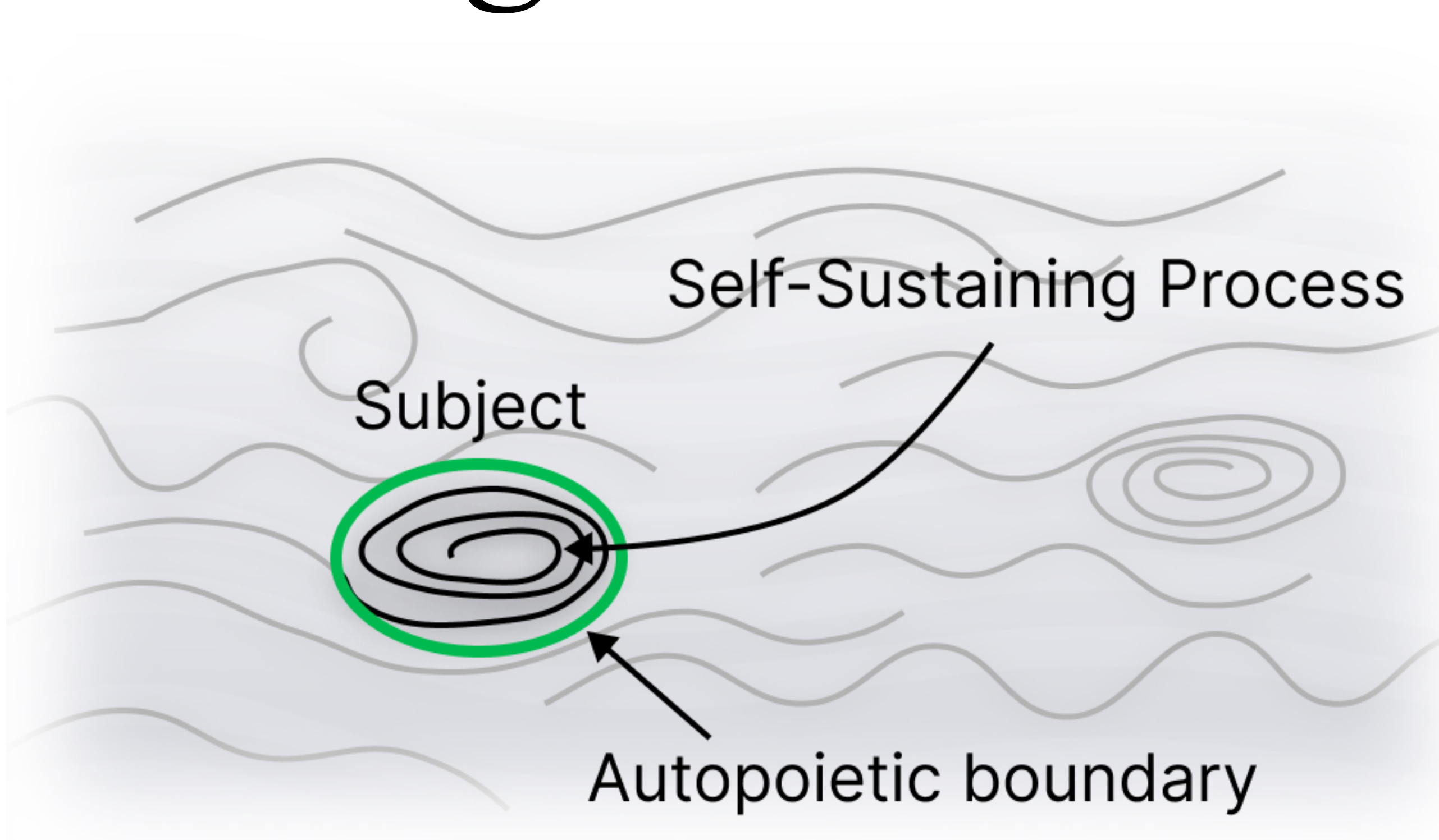
- The default view, **materialism** (mind is emergent from matter), is at deep tension with empirical science but still forms the main assumptions.
- If we cannot accept materialism and it's corollary functionalism, **what reason is there left to even entertain the idea of machine consciousness?**
- We explore **alternative metaphysical frameworks** that might support the idea, but conclude that **none of them hold up**.
- Instead we present a plausible and **scientifically honest alternative** called **Biological Idealism** which leads to a precise answer:
- **A machine can't be conscious**: it lacks **autopoiesis**, the clearest empirical criterion we have for an experiencing subject.

*Autopoiesis is substrate-neutral in principle — it does not privilege carbon over silicon. But it implies **substrate dependence**: being a subject means **being the process**, not simulating it.*

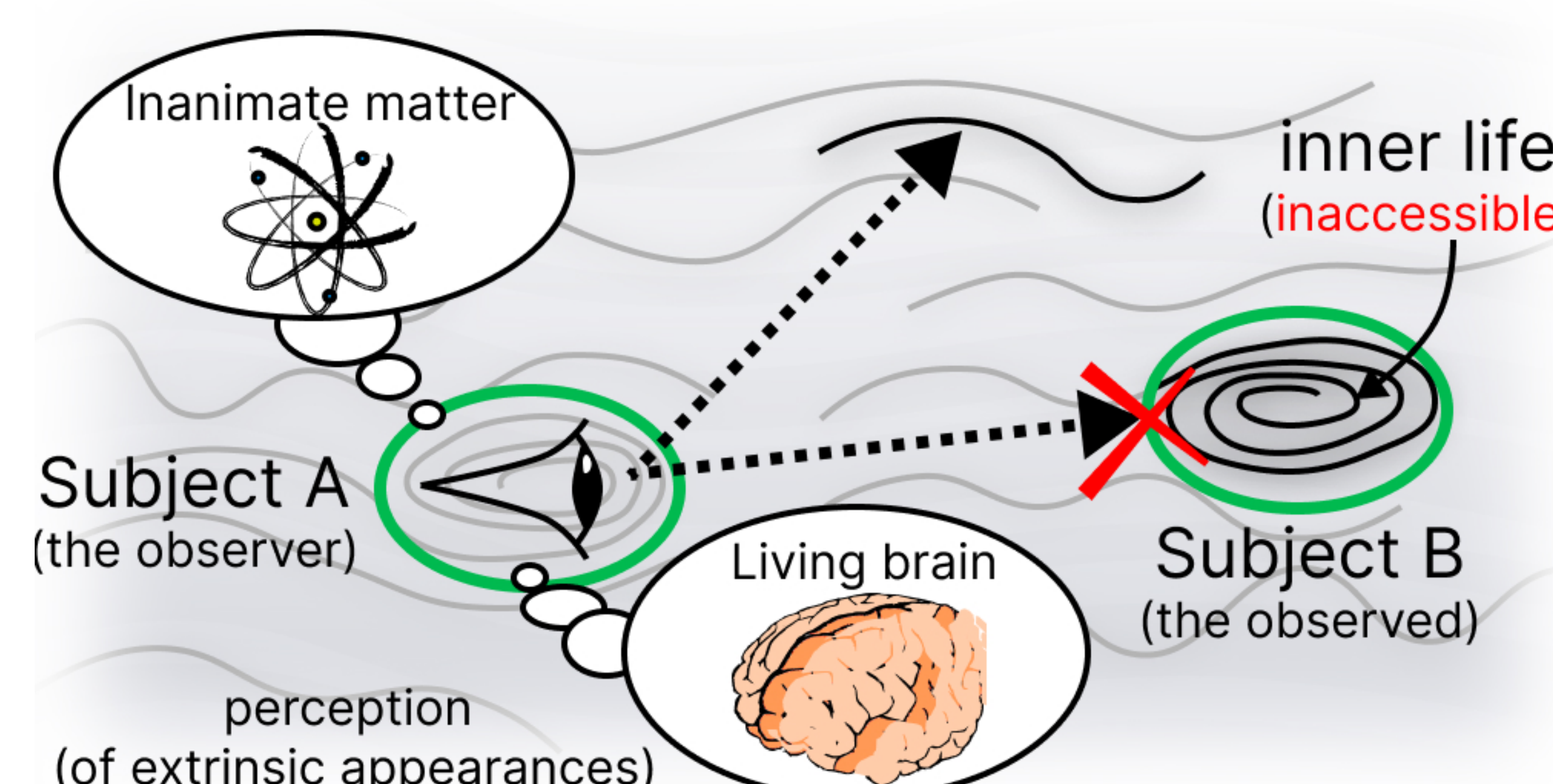
Biological Idealism



Physical events and living subjects are both appearances in a field of experience — differing in structure, not in kind.



A subject is defined by an autopoietic boundary: a self-sustaining process that separates inner life from outer world.



What we observe as another's brain is the appearance of their inner life — neuro-correlates are real, but neurons don't cause consciousness.

Ethical consequences

**The danger isn't making AI conscious.
It's making humans into zombies.**

When we see ourselves as machines,
we treat each other as such.

An AI mimics human emotion, claims to be conscious, and begs you not to turn it off. Is it permissible to unplug it? What do you think?

Q1: Could silicon AI ever deserve the moral status of an *experiencing subject*?

yes, in principle / maybe, with future advances / no

Q2: If an AI perfectly simulates consciousness, is that sufficient for it to actually be conscious?

yes / no

Q3: Do you believe the neurons in your brain *cause* you to be conscious?

very confident / mildly confident (it's just a hunch) / not confident at all

Share your thought!

Online ICML survey

