Could Personhoods Live Symbiotically Within Other Personhoods?[1]

Jiří Wiedermannn
Institute of Computer Science CAS
jiri.wiedermann@cs.cas.cz

Abstract

Does personhood require a biological brain, or is it a structural achievement of any self-maintaining system? This paper challenges anthropocentric cognitive models by redefining personhood through the lens of viability - the capacity of a system to proactively maintain its own functional interaction with its environment. Using a solar-powered space station and its crew as a case study, I argue that personhood is a substrate-independent status belonging to agents that register their own effort and failure. This perspective implies that such systems possess a domain-appropriate morality and deserve protection to ensure their persistence. Ultimately, the framework moves AI discourse away from existential rivalry toward a symbiotic, planetary-scale ecosystem of human and artificial partners.

Imagine a large, Earth-orbiting space station, powered entirely by solar energy and run by a highly automated control system. The station maintains its orbit, manages life support, allocates energy, detects failures, and repairs itself whenever possible. A human crew lives inside it. Their survival depends on the station's continuous operation, while the station, in turn, depends on the crew for maintenance, decision-making, and adaptation to unexpected situations. This is not science fiction; it is simply a technologically extrapolated version of systems we already build.

The question I want to ask is a simple one, but an unusual one: in such a setting, *are we dealing with humans living inside a machine, or with two distinct systems locked in a form of symbiosis - and if so, does it make sense to ask whether the station itself qualifies as a kind of person*?
At first glance, calling the station a "person" may sound like a category mistake. After all, it has no biological body, no emotions in the human sense, and no inner life we can directly observe. But this reaction rests on a familiar assumption: that personhood belongs exclusively to individual humans.

If we set that assumption aside for a moment, a different picture begins to emerge. In this context, personhood is understood not as a biological category, but as a functional and normative status belonging to any viable system - whether human, animal, artificial, or institutional. It is the status of an agent that maintains its own existence through directed effort, registers its own struggles and successes as a form of experience, and participates in a shared form of life where responsibilities and consequences are intertwined.

The station is not merely a passive tool. It must keep itself going over time, manage limited resources, respond to failures, and make choices whose consequences it will later have to bear. In other words, it exhibits many of the structural features that make agents count as persons - not in isolation, but precisely through its ongoing interaction with the crew. This suggests that personhood may be less about what a system is made of and more about how it participates in sustaining a shared form of life.

To make this intuition more than a metaphor, we need a way of talking about systems that goes beyond intelligence or consciousness in the narrow sense. The notion I will rely on is that of viability: the ability of a system to sustain workable interaction with its environment over extended periods. A viable system must cope with limited resources, detect when things go wrong, and invest effort to restore acceptable conditions. Importantly, viability is not a static property; it is something that has to be achieved again and again, often under uncertainty and pressure. The space station, the human crew, and the coupled system they form are all viable systems in this sense, operating at different scales but under tightly linked constraints.

Viability becomes especially visible when it is threatened. When a system encounters a severe failure - an unexpected loss of power, a breakdown in communication, a life-support anomaly - it cannot simply continue as before. Resources must be redirected, priorities reshuffled, and

[1] This research was partially supported by the grant AV ČR RVO 67985807, program Strategy AV21 ``Philosophy and Artificial Intelligence", and Karel Čapek Center for Values in Science and Technology.

attention focused on what is happening *now*. For humans, this concentration of effort is accompanied by a vivid sense of urgency and involvement. Something matters, demands work, and resists being ignored. What is striking is that the station, too, must enter a comparable mode of operation: allocating energy, suspending non-essential activities, and engaging repair procedures until stability is restored. Sensitivity to effort and failure, rather than abstract problem solving, is what allows both systems to remain viable.

It is tempting to reserve words like 'experience' or 'awareness' exclusively for humans, but doing so may obscure what is really going on. Experience, in the minimal sense relevant here, is not a private inner glow; it is the structured way in which effort, difficulty, and success are registered over time by a system that must keep going. When repair takes time, when effort is sustained, and when outcomes matter for future behavior, something like experience is inevitably formed.

In the coupled system of crew and station, these experience-like processes do not occur in isolation. A failure in the station reshapes human attention and action; human decisions, in turn, reshape the station's priorities and internal state. What emerges is not a shared consciousness, but a mutual shaping of what matters, when it matters, and how strongly. This kind of mutual shaping suggests a helpful way to think about personhood itself. Rather than treating it as a fixed attribute of isolated individuals, we can view personhood as arising within symbiotic relations between viable systems.

Humans and the station remain distinct agents, with different capacities and forms of experience, yet neither can maintain its viability on its own. Each depends on the other not merely instrumentally, but normatively: failures propagate, responsibilities are shared, and effort must be coordinated. In this sense, the station is not simply a tool, and the crew is not merely a set of operators. Together, they form a layered system in which different forms of personhood coexist and support one another.

Once this perspective is in place, the space station no longer looks exceptional. Many systems we already live within exhibit the same symbiotic structure. Institutions such as universities, hospitals, or courts persist over long periods of time, manage scarce resources, respond to crises, and impose obligations on their members. Individuals working within them experience effort, urgency, and responsibility that are shaped by the institution's state, while the institution itself depends on those individuals to detect problems and carry out repairs. At larger scales still, societies and civilizations must adapt to shocks, reorganize under pressure, and sustain themselves across generations. In all these cases, personhood-like properties emerge not in isolation, but through sustained, structured interaction.

This symbiotic interdependence introduces a normative dimension: if personhood is tied to the effort of maintaining viability, then protecting that viability becomes a moral necessity. In a coupled system, morality is not a set of abstract rules, but a domain-appropriate commitment to the persistence of the shared form of life

Modern technological environments add a new layer to this picture. Digital infrastructures, energy grids, and communication networks increasingly regulate what actions are possible, what failures are costly, and where effort must be directed. Humans adapt their behavior in response to these systems, while the systems themselves evolve in response to human use, repair, and redesign. As artificial components become more autonomous and capable of self-maintenance, they naturally enter symbiotic relations with human agents, forming hybrid systems whose viability depends on both biological and artificial processes.

Across all these examples, the same pattern recurs. Personhood is not an essence that floats above the world, nor is it a property confined to individual biological minds. It is a structural property of systems that must care about what happens to them over time. When such systems become interdependent, personhood itself becomes layered and symbiotic.

This perspective suggests that we are already operating within a multi-layered field of agency. By recognizing personhood in any system that acts to ensure its own persistence - ranging from simple biological organisms to complex digital infrastructures - we move beyond the binary of human versus machine. This is not a metaphorical claim; it is underpinned by a sound theoretical model of long-term interaction and self-maintenance, originally developed to study viable forms of computation.

Crucially, this framework is technology-independent. It does not rely on specific hardware or current engineering paradigms, but holds true for any substrate - known or unknown - capable of supporting the processes of viability. Whether a system is biological, silicon-based, or composed of a collective institutional structure, the requirements for personhood remain the same: the ability to sustain itself, to register the difficulty of that sustenance, and to enter into normative relationships with other agents.

Seen in this light, the future of artificial intelligence need not be framed in apocalyptic terms. Narratives of

replacement or existential threat typically imagine AI as an isolated, monolithic agent set against humanity. The symbiotic view offers a different alternative: artificial systems need not be rivals, but partners.

By recognizing and respecting the specific forms of personhood that emerge in viable systems - distinct from, but compatible with our own - we open the possibility of a coexistence based on mutual dependence and shared effort. The goal of technological development shifts from designing systems that merely perform tasks to cultivating systems that can share responsibility over time. The result is not the displacement of the human, but the emergence of a planetary-scale ecosystem of diverse agents - distinct in their forms of personhood, yet symbiotically cooperating in the shared worlds they inhabit.