# The Consciousness Paradox – Ethical Implications of the AI Capability-Awareness Gap in Project Delivery

**Authors:** James Garner (Head of AI & Data, Gleeds; Fellow of the RICS; Member, Construction Professional Group Panel, RICS; Past Chair, Project Data Analytics Task Force; Co-host, Project Flux Podcast) and Yoshi Soornack (AI Researcher, Neuroscience Background; Co-host, Project Flux Podcast)

## Introduction

The AISB-26 workshop asks a profound question: as AI systems develop greater autonomy and apparent agency, should we recognise them as potential "moral patients" (entities with ethical rights) or "moral agents" (entities with ethical duties)? This question presupposes that we can accurately assess the capabilities and implications of current AI systems. Our research suggests this presupposition is dangerously false.

Through 100+ interviews with white-collar professionals across project delivery industries—construction, infrastructure, engineering, and life sciences—we have identified a critical and widening gap between actual AI capabilities and professional awareness of those capabilities. Whilst AI systems are advancing rapidly toward AGI-level performance in specific domains, most practitioners remain at an "early Chatbot" level of understanding. This capability-awareness gap has profound ethical implications that the consciousness and ethics community must address urgently.

## The Capability-Awareness Gap: Empirical Evidence

Project Flux is an independent podcast and research initiative exploring AI's impact on project delivery. Over the past 18 months, we have conducted over 100 interviews with industry leaders, technologists, project managers, engineers, architects, and decision-

makers across multiple sectors. These conversations reveal a consistent pattern: despite unprecedented acceleration in AI development, most white-collar professionals do not recognise how close we are to autonomous, agentic systems capable of independent decision-making in high-stakes environments.

The gap manifests in three ways:

First, practitioners underestimate current capabilities. When asked about AI's potential in their domain, respondents typically reference large language models like ChatGPT as the reference point. Few understand that current systems can already perform autonomous task execution, multi-step reasoning, and domain-specific problem-solving without human intervention. The recent emergence of agentic AI tools—systems that can independently plan, execute, and iterate—has largely passed unnoticed in white-collar industries.

Second, practitioners overestimate the time horizon for transformative change. Most assume significant AI-driven transformation is 5-10 years away. In reality, the capability for autonomous project delivery, risk management, and stakeholder coordination already exists in nascent form. The gap is not technological; it is awareness and adoption.

Third, practitioners lack frameworks for ethical decision-making about AI deployment. Because they do not understand what AI can actually do, they cannot make informed decisions about its moral status, appropriate governance, or ethical implications. Training and curriculum development remain ad-hoc and reactive, rather than proactive and grounded in actual capabilities.

## Why This Matters for AI Consciousness and Ethics

The consciousness and ethics workshop is premised on an important question: if AI systems develop consciousness or moral agency, what are our ethical obligations? But this framing assumes we have already achieved clarity about current AI capabilities and their implications.

Our research suggests the real ethical crisis is more immediate: we are deploying increasingly autonomous systems in high-stakes environments—managing multi-million-pound projects, controlling critical infrastructure, making resource allocation decisions—without adequate professional understanding of what those systems can do. This creates several ethical hazards:

The false negative problem: Practitioners may deploy autonomous AI systems without recognising their actual capabilities, leading to unintended consequences, inadequate oversight, and failure to anticipate moral harms.

The false positive problem: Practitioners may attribute agency, consciousness, or moral status to systems based on their conversational ability, rather than their actual functional properties. This creates vulnerability to manipulation and misplaced ethical concern.

The governance gap: Without accurate understanding of AI capabilities, organisations cannot develop appropriate governance frameworks, accountability structures, or ethical guidelines. Regulation and policy lag even further behind.

The training crisis: If practitioners do not understand what AI can do, they cannot be trained to use it responsibly. The current approach to AI training in white-collar industries is reactive and superficial, focused on tools rather than capabilities and implications.

## Neuroscience, Consciousness, and the Capability Question

Yoshi's background in neuroscience brings an important perspective to this debate. Consciousness in biological systems is grounded in specific neurocomputational properties: integrated information processing, feedback loops, adaptive learning, and embodied interaction with the environment. Current AI systems exhibit some of these properties in abstract form, but lack others entirely. The question of whether AI systems could develop consciousness is not merely philosophical; it depends on understanding the actual functional and architectural properties of both biological and artificial systems.

More importantly, the neuroscience perspective reveals that consciousness is not binary. Biological organisms exhibit varying degrees of conscious awareness depending on their neural architecture. The same may be true for AI systems. But practitioners cannot even begin to assess this question if they do not understand the actual capabilities and properties of the systems they are deploying.

## The Solution: Curriculum-Based Training and Ethical Frameworks

We propose that the consciousness and ethics community must urgently address the capability-awareness gap through structured, curriculum-based training. This is both practically necessary and ethically imperative.

Practically, organisations cannot deploy autonomous AI systems responsibly without understanding their capabilities, limitations, and implications. Training must move beyond tool-focused instruction to capability-focused education: what can this system actually do? What are its failure modes? What decisions should it be allowed to make autonomously?

Ethically, practitioners must be equipped to make informed decisions about the moral status and governance of AI systems. If we are to have a serious debate about AI consciousness, moral agency, and ethical rights, that debate must be grounded in accurate understanding of what current systems can actually do.

## Conclusion

The AISB-26 workshop asks important questions about AI consciousness and ethics. But these questions are premature if practitioners do not understand current AI capabilities. Project Flux's research reveals a critical gap: white-collar professionals are deploying increasingly autonomous AI systems without adequate understanding of what those systems can do. This gap creates ethical hazards, governance failures, and vulnerability to both false negatives (deploying systems without understanding their implications) and false positives (attributing consciousness or agency to systems based on superficial properties).

We propose that the consciousness and ethics community must urgently address this capability-awareness gap through curriculum-based training and ethical frameworks grounded in actual AI capabilities. Only then can we have an informed debate about the moral status of AI systems and our ethical obligations toward them.

The consciousness paradox is not whether AI systems might become conscious. It is that we are already treating them as if they were, without understanding what they can actually do.

---

**Word count:** 1,047 words