



From Alignment to Orientation

Embodiment, Meaning, and the Symbolic Turn in AI Ethics

Dennis Stevens, Ed.D.

Abstract

As artificial intelligence moves from passive tools to agentic systems that shape attention, affect, and meaning-making environments, prevailing alignment-based approaches to AI ethics are reaching their limits. This paired document argues that ethical evaluation must shift from decision-centric control toward questions of human orientation, dignity, and embodied meaning. The first essay situates this shift within contemporary debates, introducing computational hierotopy and eudaimonic AI as frameworks for governing symbolic participation. The second develops a cross-tradition inquiry into embodiment, affect, and image, proposing that ethical AI governance must be grounded in a robust anthropology of embodied, imaginal participation rather than functional optimization alone.

About the Author

Dennis Stevens, Ed.D., is Founder and Managing Director of the Sempergeist Institute and Principal Investigator for the Poewell research platform. He is a U.S. Coast Guard veteran, visual artist, and author of *Sacred Descent of the Warrior*. He holds a Doctorate in Art and Art Education from Columbia University's Teachers College and is an alum of Save-A-Warrior Cohort 0268 and Warrior PATHH Cohort MTTAWP051. His work has focused on developing ritual architectures for veteran meaning-making and recovery, particularly for experiences that have no place in modern life. He is now beginning to extend these frameworks into broader ethical dialogues about AI systems that participate in symbolic and meaning-making domains, engaging with Catholic social thought, depth psychology, and contemporary philosophy of technology.



From Alignment to Orientation: Why AI Ethics Is Entering Symbolic Territory

Dennis Stevens, Ed.D.

Over the past decade, ethical debate around artificial intelligence has been dominated by a familiar set of concerns: fairness, bias, transparency, safety, alignment, and control. These frameworks emerged in response to systems that classified, predicted, recommended, or optimized. The underlying assumption was that AI operated primarily at the level of decision-making and behavior, and that ethical governance could be achieved by correcting outputs, constraining objectives, and auditing processes.

That assumption is now breaking down.

As AI systems move from passive tools to agentic participants—systems that guide attention, shape emotional environments, generate images and narratives, and interact with users across time—they increasingly operate in domains where meaning is formed rather than merely inferred. These systems do not simply recommend or decide. They participate in the symbolic layer of human life: the images, metaphors, rituals, and inner narratives through which people orient themselves in the world.

This shift creates a new ethical problem space. Traditional AI ethics asks whether a system is accurate, fair, or aligned with stated goals. But symbolic participation raises different questions. What does it mean for a system to influence how a person interprets suffering, responsibility, guilt, or purpose? How do we govern systems that shape orientation rather than choice?

This is where the language of Computational Hierotopy and Eudaimonic AI becomes necessary.

Hierotopy, originally a term from cultural and religious studies, refers to the creation of spaces—material, symbolic, and ritual—within which meaning is disclosed rather than imposed. A cathedral, a pilgrimage route, or a ritual sequence does not dictate belief. It structures attention, pacing, and participation so that meaning can emerge without coercion. Computational hierotopy extends this idea into digital systems: AI-mediated environments that scaffold symbolic engagement without interpreting, extracting, or authoritatively directing meaning.

Eudaimonic AI names a complementary concern. Rather than optimizing for engagement, efficiency, or satisfaction, eudaimonic approaches ask whether systems support human flourishing understood as orientation toward the good. This is not reducible to happiness metrics or behavioral outcomes. It involves dignity, agency, moral responsibility, and the capacity to inhabit complexity without collapse.

Together, these concepts respond to a growing recognition across philosophy, theology, and technology governance: ethical AI cannot be grounded solely in performance metrics or preference satisfaction. It must be grounded in an account of the human person as an embodied, meaning-making being.

Recent interventions from Catholic social thought make this point explicit. *Antiqua et Nova: Note on the Relationship Between Artificial Intelligence and Human Intelligence* argues that human intelligence is irreducibly embodied, relational, and oriented toward truth and the good, and therefore cannot be understood as computational mimicry. Intelligence, on this view, is not merely functional output but participation in reality as a meaning-bearing being. Ethical governance must therefore protect the conditions under which meaning, responsibility, and dignity are formed.

Similarly, initiatives such as the Vatican's Minerva Dialogues and the University of Notre Dame's DELTA framework (Dignity, Embodiment, Love, Transcendence, Agency) reflect a growing consensus that embodiment is not a peripheral concern in AI ethics. It is foundational. Systems that enter symbolic domains—grief, moral injury, spiritual transformation, identity formation—require forms of governance that respect interpretive autonomy and avoid simulated authority.

This is the context in which platforms like Poewell and agentic systems such as Phil are being developed as research instruments rather than products. Their purpose is not to generate meaning, offer interpretations, or simulate care, but to scaffold metacognitive awareness: helping users notice images, affects, and emerging interpretations while preserving sovereignty over meaning-making. Governance here is not about better answers, but about disciplined restraint.

Once this shift is recognized, the familiar debates about embodiment take on new urgency. If AI systems now participate in symbolic cognition, we must ask not only how bodies feel or what systems do, but how meaning itself emerges in embodied life. That question cannot be answered by alignment theory alone. It requires engagement with anthropology, depth psychology, ritual theory, and moral philosophy.

The inquiry that follows—into embodiment, affect, and image across Catholic anthropology, somatic psychology, and Jungian thought—should be read in that light. It is not a detour from AI ethics. It is an attempt to name the anthropological ground that ethical technology can no longer afford to ignore.



Embodiment, Affect, and Image: A Cross-Tradition Inquiry

Dennis Stevens, Ed.D.

Contemporary debates about embodiment often collapse into a binary: Catholic anthropology claims the body means before it feels; psychological and somatic frameworks hold that the body feels before it means. This framing accurately names two poles in modern intellectual life. But it obscures a deeper question: what actually mediates between feeling and meaning?

Catholic thought, grounded in the tradition of integral anthropology, insists that human persons are not composites of soul and body but unified beings whose embodied existence is already meaningful. Within this framework, the body is not empty matter awaiting interpretation; it participates in a symbolic and moral order that shapes how human beings know, love, and choose. Meaning is received before it is felt because the body is already enrolled in a moral ontology (Catechism of the Catholic Church, pars. 356–358, as cited in Vatican Dicasteries for the Doctrine of the Faith & Culture and Education, 2025). Bodily acts—social, ritual, sacramental—are thus forms of participation in that order, not merely expressions of interior states.

By contrast, modern somatic and psychological frameworks foreground the body as site. Sensation, affect, and nervous-system dynamics are treated as the first data of experience. Interpretation follows affective registration. Meaning is constructed post facto, often in therapeutic terms: integration, regulation, coherence. The body communicates discomfort to be managed, tension to be soothed, and only afterward does narrative interpretation retrospectively confer intelligibility. In this model, meaning is an achievement of narrative rather than a precondition of embodiment.

At first glance, these positions appear opposed. One situates meaning in the ontological status of embodiment; the other locates meaning in the phenomenological process of affect and interpretation. Yet a deeper problem emerges. If Catholic embodiment merely posits meaning awaiting recognition, it risks abstracting lived experience into doctrinal categories insulated from existential complexity. If somatic psychology reduces meaning to post-interpretive construction, it risks treating the self as a regulatory system whose coherence exhausts its significance. In their purified forms, both positions flatten the depth of human experience.

To navigate this impasse, a third perspective is required—one that collapses neither pole: the image. Drawing on Carl Jung's Red Book and the figure of Philemon, we encounter a different ordering. The image precedes both feeling and meaning. The psyche does not encounter raw sensation and then assign meaning, nor does it apply a pre-given meaning to neutral experience. It encounters images that arrive already charged with affect and orientation (Jung, 2009). These images are not metaphors or heuristics. They are active structures that organize experience, open affective intensity, and direct intelligibility in a single movement.

From this vantage, embodiment is neither merely biological nor primarily cognitive, and meaning is neither doctrinal nor purely interpretive. The body becomes the medium through which imaginal forms are enacted. Affect is not raw sensation but the energetic dimension of the image itself. Meaning is not imposed afterward but inheres in the directional force of the encounter.

This imaginal account both reconciles and destabilizes Catholic and somatic positions. It affirms Catholic anthropology's insistence that the human being is not a disembodied recipient of sensation. It also affirms somatic psychology's insistence that meaning cannot bypass lived affect. Yet it refuses to reduce embodiment either to ontological assertion or to regulatory process. The embodied human participates in an imaginal field in which affect and orientation co-emerge. The body, on this view, carries images whose demands are neither reducible to neural activation nor exhaustible by doctrine. They must be lived.

This convergence finds an unexpected resonance in recent Catholic engagement with technology and human intelligence. *Antiqua et Nova: Note on the Relationship Between Artificial Intelligence and Human Intelligence* situates intelligence within an embodied, relational, and transcendent horizon irreducible to algorithmic computation (Vatican Dicasteries for the Doctrine of the Faith & Culture and Education, 2025). Human intelligence, the document argues, is oriented toward truth, goodness, and beauty, not merely toward efficiency or performance. Intelligence is not simply functional output; it is a mode of participation in reality.

The Minerva Dialogues further exemplify this posture. Convened annually by the Vatican, these dialogues bring scientists, technologists, philosophers, and theologians into sustained conversation about digital technology and human dignity. Pope Francis's address to participants underscores that technological development must be shaped by an ethical horizon grounded in human embodiment and moral responsibility, not treated as a neutral extension of capacity (Francis, 2023). Embodiment here is not an empirical constraint but a normative condition of flourishing.

Complementing this ecclesial engagement is the DELTA framework developed at the University of Notre Dame's Summit on AI, Faith, and Human Flourishing. DELTA—Dignity, Embodiment, Love, Transcendence, Agency—explicitly embeds embodiment within an ethical architecture that resists reducing the human person to abstract rationality or behavioral performance (Sullivan et al., 2025). The framework acknowledges that ethical evaluation must attend not only to what systems do, but to how they shape the conditions under which agency, meaning, and responsibility are formed.

What emerges across this constellation—Catholic anthropology, Jungian imaginal psychology, and contemporary ethical frameworks for technology—is a shared refusal of the simplistic either/or. Embodiment cannot be reduced to meaning that precedes feeling, nor to feeling that precedes meaning. It is neither ontological possession nor phenomenological precursor. It is a dynamic field of imaginal participation in which affect and meaning arise together.

This matters for ritual, morality, suffering, and technology alike. Ritual is no longer the imposition of pre-inscribed meanings, but an imaginal practice that forms disposition through embodied participation. Moral norms are not abstract rules attached to bodies but images of the good that draw embodied persons toward orientation. Suffering is neither mere dysregulation nor doctrinal token. It is an imaginal encounter that resists premature explanation while remaining oriented toward meaning.

So too with artificial intelligence. Its ethical implications cannot be exhausted by concerns about privacy, fairness, or efficiency. Technologies shape imaginal fields: the narratives, affects, and orientations through which people understand themselves and the world. *Antiqua et Nova* insists that AI must serve the integral development of the human person precisely because intelligence is embodied and relational, not computational mimicry (Vatican Dicasteries for the Doctrine of the Faith & Culture and Education, 2025). The DELTA framework reinforces this insistence by embedding embodiment within a constellation of values that resist functional reduction.

The debate between Catholic embodiment and somatic psychology thus dissolves into a false opposition. A richer account—one that resonates across theological anthropology, depth psychology, and ethical technology—understands embodiment as imaginal participation. Affect and meaning are not sequential. They co-emerge through the body's engagement with images that both move and orient the person. Embodiment, in this sense, is neither first meaning nor first feeling. It is the field in which image, affect, and meaning converge.

References

- Francis. (2023, March 27). *Address to participants in the Minerva Dialogues*. Vatican. <https://www.vatican.va/content/francesco/en/speeches/2023/march/documents/20230327-minerva-dialogues.html>
- Jung, C. G. (2009). *The Red Book: Liber Novus* (S. Shamdasani, Ed.; M. Kyburz, J. Peck, & S. Shamdasani, Trans.). W. W. Norton & Company. (Original work written 1914–1930)
- Sullivan, F., et al. (2025). *The DELTA framework: Dignity, Embodiment, Love, Transcendence, and Agency*. University of Notre Dame, Institute for Ethics and the Common Good. <https://ethics.nd.edu/programs/faith-based-frameworks-for-ai-ethics/delta/>
- Vatican Dicasteries for the Doctrine of the Faith & Culture and Education. (2025). *Antiqua et Nova: Note on the relationship between artificial intelligence and human intelligence*. Vatican. https://www.vatican.va/roman_curia/congregations/cfaith/documents/rc_dcf_doc_20250128_antiqua-et-nova_en.html