



Soul Trader: Identity in a Digital Age

Mary Harrington

About the Author

Mary Harrington is a British columnist, author, and contributing editor at UnHerd. Her first book, titled *Feminism Against Progress*, argues that the modern understanding of “progress” is turning against all but a tiny elite of women.

This paper was edited by Benedict Rogers, Joshua McNicholas, and Anna Webb.

Identity in a Digital Age

Mary Harrington

Introduction

How does the digital era affect our identities? We are familiar with the warnings about the digital era and mental health, attention, cognitive development, bullying, pornography, and so on. The transgender culture war rages on. Periodically, we hear in the press that some group or other has begun to “identify” as animals, or fictional characters, or perhaps multiple personalities or spirits.

Something odd is going on. But when we pause and zoom in, the territory becomes bewildering. What do we even mean by “identity?” What is an identity made of? Where do our beliefs about this phenomenon come from? It is as disorienting as trying to think about thought. As I will argue, a serious treatment of identity means examining our foundational categories for thinking about what we ourselves are. And what we believe ourselves to be has always also been bound up in what we believe reality to be.

When we explore these phenomena historically, it becomes clear that the term “identity” has emerged only relatively recently. Google’s “Ngram” tool, which allows the user to track the frequency of usage for a term over millions of historic texts, suggests “identity” emerged into mainstream cultural usage after the two twentieth-century World Wars, accelerating rapidly from the 1960s.¹

Why? What preceded it? It is my contention that “identity” has, roughly since the 1960s, come to replace the modern secular “self” as a byproduct of a fundamental change that began at that point, in which we turned our technologies inward and began applying them to our own minds and bodies. And the “digital era” encompasses the technologies with which we did so: an era we also call the “Fourth Industrial Revolution”, or (more dramatically) the transhumanist age. It is a central premise of this essay that we are already over half a century into this transhumanist era.

Meanwhile, the antecedent to “identity”—the secular “self”—emerged in tandem with modernity. As for the predecessor concept to the “self”, this is a term now largely defunct, at least within mainstream public conversation: the soul. Within the Western tradition, these iterations in our ongoing effort to name and make sense of this most distinctively human of phenomena—our capacity for self-reflexive consciousness—are in turn inextricable from the information technologies that have helped to structure our efforts to understand who and what we are. The earliest known references to the

soul emerged in tandem with literacy; the modern “self” cannot be disaggregated from the printing press; and our contemporary understanding of “identity” emerged in tandem with our entry into the digital era—an era that, as this paper will argue, is better understood as a new form of merger with our technologies.

In what follows, this paper will trace the interlocking story of souls and information technologies, from Plato to our transhumanist era.² At each stage, these developments have brought tremendous benefits as well as costs and metaphysical convulsions, and today is no exception. None of what follows is an argument for unplugging the internet. But if we are to survive the latest convulsions, we will need to revisit those ways we have come to apply modernity to ourselves. If we are to master our machines, rather than be mastered by them, we must recognise “identity” for what it is—the dead byproduct of applying our technologies to ourselves. In its place, we can and must restore the older and fuller understanding of ourselves: not as mere bodies, not as plastic self-actualising “identities”, or even as data processors, but souls. And in retrieving this fuller sense of what we are, we may yet manage to escape the epistemological cul-de-sac of modernity, without losing the gifts modernity has afforded us.

Three Singularities

Human-machine convergence was once a niche preoccupation for Silicon Valley tech optimists such as Ray Kurzweil, who forecasted that by 2045 humans and machines would merge, delivering untold, transformative benefits.³ But as I have argued, in a sense this merger has already happened, with hormonal birth control serving as its entry-point.⁴ For the first time, with the Pill, a medical technology was adopted not to restore normal physiological functions—specifically, female fertility—in the name of health, but to interrupt normal functions in the name of individual freedom. This inaugurated a radical transformation in our understanding of what a person is, in which embodied attributes once understood as irreducible givens—beginning with sex difference—have come to be seen as optional, and subject to individual choice and re-engineering. Importantly, inasmuch as women have embraced this technologically-enabled understanding of ourselves—in which female fertility is viewed as an option whose default setting is “off”—we have embraced our own fusion with technology as a precondition for participation in public life. Much of the mainstream social fabric in Western societies is downstream of such technologies. In other words: the default modern Western culture is already a cyborg one.

Whether or not you view this as a good thing—and it has brought many benefits, including greater freedom and opportunity for some women to participate in work, study, and public life—this development also represented a profound shift change in our relationship to our technologies, and in our beliefs about what it means to be a person. This shift in turn accelerated a further transition: from speaking about the

“self” to speaking about “identity”, a move whose logic and profound consequences I will explore below. But it did not stop with the Pill. Its second phase initially appeared in the same decade birth control was legalised, but was universalised much more recently: networked computing.

As the Pill began to revolutionise “identity” throughout the West, the US Department of Defence was sponsoring the Defense Advanced Research Projects Agency (“DARPA”) to explore the first networks of information-processing machines, with the aim of creating a resilient network in the event of a defence crisis. Roughly half a century and a great deal of development later, the second phase of the Singularity began in earnest, when Apple produced the first iPhone. On its launch in 2007, Steve Jobs described this innovation as three products in one: “a revolutionary mobile phone, a widescreen iPod® with touch controls, and a breakthrough Internet communications device with desktop-class email, web browsing, searching and maps”.⁵ By 2021, data traffic on mobile networks increased 300-fold; 2 billion more people gained a smartphone; and 5.6 billion new smartphone subscriptions were opened.⁶

Smartphones proved to be the digital equivalent of the Pill: every bit as intimate, as seemingly indispensable, and as deeply implicated in everyday freedoms and commerce. And as with the Pill, the global transformation that has followed their normalisation has reshaped politics, culture, sex, family, and even consciousness itself. Most centrally, it has reshaped what we mean when we say “identity”. For possessing an always-on, networked, handheld device enables every user to turn away from the arbitrary, immediate, real-world features of our lives as the principal forces shaping “who we are”. Instead, it has grown increasingly possible to curate social groups based on interests and affinities, infinitely adjustable virtual “avatars”, and even erotic and romantic entanglements conducted partly or wholly in a dematerialised, digital realm.

In this context, “identity” has begun to work loose from embodiment, family, and other interpersonal relationships, and even in some cases our self-understanding as human beings. But even when unmoored from such familiar anchors, self-curated identities require validation in order to feel real: a dynamic that, in turn, has increased the salience of *recognition* as a political desideratum. It is not a coincidence that “identity politics” came to mainstream public prominence in the wake of the smartphone revolution. The so-called “Great Awakening” is generally dated from around 2010, with many researchers linking its emergence directly to mass adoption of social media.⁷

Now, as I write, we are in the early stages of a third convergence with our machines: the onrushing development of the programmes colloquially known as “artificial intelligence” or AI. These technologies draw on the patterns discernible across vast datasets, granting the models themselves eerie-seeming ability to identify or even predict future patterns on the basis of existing ones. Such systems have been employed in areas as diverse as image generation, medical diagnosis, biometric security, and even criminal law. Most famously, “large language models” (or “LLMs”) such as Chat-GPT respond as if in conversation with human interlocutors, answering questions,

systematising information, and interacting in often surprising ways with questioners. Elsewhere, rapid and astonishing AI-powered advances in automation are proceeding everywhere from manufacturing to transport and defence.

LLM usage has already become part of mainstream culture, from management consultancy to students cheating at homework, and even one innovator employing it to approach prospective partners on a dating website.⁸ In the developed world, most of us already interact with such technologies regularly. For the most part, the response to such developments is wonder, surprise, and sometimes unease at the range of human functions that might potentially be automated by their advance. Behind the unease lies a far deeper question: are there any limits to what these machines can do? Is there anything at all about humans that cannot be replicated, and perhaps outclassed, by one of our own inventions?

AI maximalists imagine that there is not. The endpoint of AI development will, they imagine, be “artificial general intelligence”: a moment at which such models leap from prediction within a specific dataset to a general power to think, after which they will become recursively self-improving. For the optimists, this represents a transformative accelerant of everything we value in human life, or even the means whereby humans create our own successor species.⁹ Conversely, pessimists warn that the risks of such a super-intelligent machine destroying us are so grave that we should halt all research.¹⁰

But the most contested question, for boosters and doomers alike, is whether or not human thought and computation will ever amount to the same thing. This is illustrated by the way those existing applications of AI that cause the deepest unease tend to be those that act as if this was so, or nearly so. For example, the Character AI app enables users to create “roleplaying” interlocutors, prompting outrage when AI “character” versions of murder victims were found to have been created.¹¹ Such simulated “people” can have still more troubling sequelae: for example, in February 2024, Sewell Setzer III, a 14-year-old boy from Orlando, Florida, took his own life after forming an intense and highly sexualised relationship with a chatbot he created using Character AI. Elsewhere, common descriptors used online for the “Moxie” AI-powered “companion robot”, launched in 2021 to teach “emotional intelligence” to children,¹² include “creepy”,¹³ or even “nightmare inducing”.¹⁴

“We are”, claims tech investor Marc Andreessen, “literally making sand think”.¹⁵ Is this, as Andreessen says, “literally” true? If so, it surely follows that we just need to make machine computation better, faster, and more “aligned” or otherwise complete, in order for artificial intelligence to produce better and more effective “thinkers” than have emerged as yet from the human species. But as I will argue, this is a profound category error that, in the end, will not aid but impede scientific understanding, and with it, real human progress.

The First Singularity is Literacy

Three technology transitions define our era. As we will see, all three represent radical and transformative applications of technology to the human soul. But while this revelation may sound alarming, in truth it exists in continuity with a far longer civilisational trajectory. For our shared and nested understandings of human consciousness, and of reality, have always interlocked with our information technologies. More plainly: this is not our first Singularity event.

The long arc of Western culture tracks a slowly evolving view of our inner lives, that developed in tandem with beliefs about the nature of reality and the divine. And this evolution has always been both shaped and transmitted by our information technologies. Of these, reading and writing was always the technology most profoundly linked to the soul. In particular, literacy appears to be the critical catalyst for a sense of *having* inner lives, whose activity and topography is distinct from our engagement with the world around us.¹⁶ The historian Adam Garfinkle argues that what Karl Jaspers called the “Axial Age” (within classical antiquity) emerges downstream of “the spread of deep literacy in a still-small but critical share of the population”—and, with it, of “a new sense of interiority”.¹⁷ The first recorded reference to the soul—*psyche*—in the Western tradition appears in Plato, one of the most profoundly influential philosophers of the Axial Age.

For Plato, and later (in modified form) Aristotle, the soul was understood as having three parts. Plato sets out his three-part understanding of the soul in book four of *The Republic*: appetite, spirit, and reason. The first of these, *eros*, concerns urges and desires such as food and sex. The second, *thumos*, concerns spirit, courage, passion, and the drive for recognition. The third, *logos*, Plato situates in the head. Its role is governance of the whole, oriented toward comprehension of the divine order as expressed in the world. And this divine order reflects a crucial corollary: from these earliest conceptions of the soul, it is clear that such self-reflection cannot be separated from how we understand the rest of reality. For both Plato and Aristotle, to be “rational” as such did not connote, as it does for moderns, unemotional detachment. Rather, “rationality” as expressed by *logos* denotes a capacity to discern and order one’s life and behaviour to “rationality” at the cosmic scale.

This intuitive sense of an inextricable, ordered, and transcendent correspondence between what we are as souls, what and how we perceive, and what the world *is* at the most fundamental level, is characterised by the historian Charles Taylor as “ontic logos”: an experience of reality in which mind and world cannot be disentangled.¹⁸ And it is clear that these earliest Western theories of the soul—themselves inextricable from theories of reality—are both profoundly tied to the capacity for “deep reading” in whose wake they became possible. Seen thus, we might argue that the entire history of Western Civilisation follows from this first Singularity.

Rationality versus Freedom

Though a brief paper cannot do justice to the detail, there existed considerable continuity between Plato's divine order and the version of "ontic logos" discernible in influential early Christian thinkers. Where Plato posited a world of Forms, for example, Saint Augustine reconceived these as ideas in the mind of God. Importantly, though, Christianity brought an inward turn: as Charles Taylor shows, while Saint Augustine retained the ontic logos—the suffusion of divinity in nature—orientation toward the divine was now understood to be achieved via the discovery of God within the soul, aided by divine grace. This crucial inward turn would have momentous consequences over time—as would the Christian emphasis on divine and human freedom.¹⁹ From the perspective of these interlocking views of the soul and the world, one of the most seismic inflection points in that story came in the thirteenth century AD, with William of Ockham's challenge to the "ontic logos". For this thirteenth-century Franciscan friar, such a static order implied constraints on God's freedom to act in history. How could it be the case that God's divine freedom might be constrained by the limits of God's own created order? Ockham argued that since it could not be proved that pre-existing divine templates existed for forms in nature, such entities could not be assumed to exist.²⁰

Downstream of this challenge, the immanent and self-evident nature of the divine order came increasingly into question. In turn, as the world's orderliness grew less self-evident, so individual human freedom to act *upon* that world came also to seem not just legitimate but morally right, in order to grasp the fulness of God's creation and realise its potential for good. In this context, the first proto-scientific enquiries began, and the pace of mechanical innovation slowly to increase. Of these innovations, Gutenberg's movable type was to prove the most shattering of all.

Moving from hand-lettered to printed books accelerated their production, which widened access to the soul technology of literacy. And with this change, all the cognitive transformations associated with literacy also began to spread, including the experience of interiority. Public willingness to rely on clerical authority for spiritual guidance waned. Central to the Protestant Reformation that followed—bringing political, religious, and intellectual convulsions in its wake—was a new relationship between individual souls and the divine: one less mediated by liturgy or religious authority, but rather as between God and individual readers of Scripture.

In turn, wherever Protestantism took hold, so too in literacy followed, often via state-mandated education. And, inevitably, as the relation between individual believers and the divine shed its previous anchors in ontic logos and social or liturgical mediation, so too the relation between soul and world also transformed. The modern era that has followed in the wake of this revolution could not have flourished without, as Ong describes it, this distinctly Protestant-flavoured literacy: "a world dominated by writing with its drive toward carefully itemised introspection and elaborately worked out

analyses of inner states of soul and of their inwardly structured sequential relationships.”²¹

The Literate Self

In very reductive summary, the printing press both arose from, and in turn accelerated, a transformation in thinking derived from Ockham’s war on the notion of divine forms: a nominalist revolution. In time, this shift would deliver a radically new perspective on the natural world, as well as a new kind of soul: the (increasingly) secular “self”. Francis Bacon was a pioneer of this modern formulation, in *Novum Organum* (1620), where he detailed a new method of inductive reasoning that, he argued, was superior to the scholastic method of reasoning by syllogism and that would allow us to force nature to give up her secrets. The violence of the analogy—Bacon speaks of putting Nature “to the question”, which is to say, torture—is a far cry from the sense of ambient divinity expressed by the ontic logos.

Not long after, René Descartes would write *Discourse on the Method* (1637) in which he dismissed “that speculative philosophy which is taught in the Schools” in favour of “a practical philosophy” that will convey knowledge of the natural, external world “as distinctly as we know the different crafts of our artisans”. Thus, Descartes reasoned, we could employ discoveries usefully to our own benefit, “and thus render ourselves the masters and possessors of nature”.²²

As Taylor puts it, this emerging consensus on human selfhood presupposed “the utter separation of mind from a mechanistic universe of matter”, understood as “most emphatically not a medium of thought” but rather “expressively dead”.²³ Beyond us, this expressively dead world is hypothesised as entirely knowable via the scientific method. In other words: as enchantment drained from the world with modernity, the inner, perceiving self would take on ever greater centrality. In a world ever more bare of God’s presence, it follows that value and meaning can only come from within.²⁴ Meanwhile, that which lay beyond our minds was now understood as inert and mechanistic, awaiting our study and mastery.

Within the “secular” modern world that developed in the wake of the resulting scientific advancements, God came increasingly to be understood as superfluous or even as an obstacle to further advancements. Though Bacon and Descartes were Christians, they anticipated how, in order to be a legitimate object for scientific study, the world must be purged of divine aura. In their wake God thus retreated by stages, first to the remote status of “divine clockmaker” as in William Paley’s 1802 formulation,²⁵ and, over time, to a kind of “tolerated” status as optional extra.

In this context, it has grown increasingly difficult to speak of “souls”, with the spiritual presuppositions this term would imply. And yet that *something* persists that does the

perceiving, thinking, understanding, and innovating—only now it is more commonly referred to not as the soul, but its secularised analogue: the “self”. All that is distinctive about modernity is held in this formulation for the relation between our “selves” and the reality we inhabit. On the one hand, a radically subjective inner “self” capable of thought but not (or not necessarily) spiritual in nature, and on the other an inert, comprehensible, mechanistic outer world ordered to our instrumental mastery.

Out of this—and out of reactions to it—came high European culture, industrial modernity, the scramble for empire, and much else that is commonly gestured at, when we use the phrase “Western Civilisation” today. At its core lies the postulate of an inert external world, fundamentally predictable and mechanistic in nature, ultimately intelligible to and predictable within the laws of physics, and amenable to instrumental management by a disengaged, rational, but also radically subjective “self”.

This model of self and reality has, since the two World Wars, seemingly given way to the postmodern construct of “identity”. But it is a mistake to think of that segue as the next stage of our civilisational evolution after modernity. Rather, it represents modernity’s last stand, against the germ of an intellectual and metaphysical revolution which emerged from the bosom of science itself more than a century ago. This revolution that has proved so indigestible since, that we have spent the ensuing decades doing everything in our power to avoid its implications, including attempting to become our own machines.

The Great Weirding

“I can safely say”, declared theoretical physicist and Nobel Prize winner Richard Feynman in 1965, “that nobody understands quantum mechanics”.²⁶ First postulated in 1900 by Max Planck, quantum physics set out to explain inconsistencies that had proved irresolvable within the mechanistic model of physics. Its postulates proved astonishingly effective at resolving the paradoxes they addressed. But the catch is that, if true, its precepts invalidate everything we have believed, more or less since Francis Bacon, about what the world actually is.

According to the principles of quantum mechanics, the world is predictable except when it is not. It gets worse: these developments suggest that the smallest components of matter are not, as Newton suggested in 1704, “solid, massy, hard, impenetrable, movable particles”.²⁷ Rather, as the semiconductor inventor Federico Faggin puts it, they comprise “the collective behaviour of [...] invisible and interacting probability waves that do not exist as such in our world until they manifest as particles in spacetime.”²⁸ These waves behave differently depending on whether or not they are being observed. There are more than four dimensions. Things a long way apart in physical space can affect one another through “entanglement”. And no one really knows what quantum mechanics is describing.

Reality, it turns out, is *weird*. And no one, not even theoretical physicists, really understand how any of this works. As one physicist put it recently, most do not even want to: “Physicists don’t understand their own theory any better than a typical smartphone user understands what’s going on inside the device.”²⁹

All of this ought, if we take it seriously, at least to invite questions about whether the modernist model of reality as “expressively dead” adequately describes the world we move through. And indeed, very recently, thinkers have begun to ask these questions: the cognitive scientist Donald Hoffman argued in 2019 that the “reality” we take for granted every day is a simplified metaphor, analogical to the “icon” on a computer “desktop” relative to the programme it represents.³⁰ It is less, Hoffman suggests, that we see what is there than that we see what evolution has confirmed, over time, to be *useful* to our continued survival. And if this is so, it stands to reason that new conclusions about the nature of reality ought to rebound on our conceptions of what we ourselves are.

So far, though, we have done everything in our power to avoid drawing this conclusion—despite (or perhaps because of) having now lived through several large-scale object lessons in the ambivalent power of our own innovations. The two World Wars were perhaps the climactic event of high modernity; in their wake, public faith in the self-evident goodness of modernity and its model of selfhood and reality was—to put it mildly—badly damaged. So, too, was any prospect of returning to religious faith: as the Catholic writer Ronald Knox documents, the events of Hiroshima and Nagasaki cemented many in a certainty that with horrors like this in the world there *could* be no God.³¹ After a brief revival in the 1950s, Church of England membership almost halved from 9.9 million in 1960 to 5.5 million in 2010,³² and lost another million by 2020: a precipitous collapse in observant faith.³³

Over the twentieth century, a consensus seems to have emerged that the problem with modernity was not the paradigm itself; rather, the weak point in the project of utopian mastery was *us*. We were, as evidenced by the destructiveness of two World Wars, not virtuous enough to be trusted with our own inventions. Thereafter, the terrain for utopian re-engineering was, inescapably, humans ourselves: with this, the transition began from modernity’s secular, modern “self” toward postmodernity’s pliant, self-creating, and far more nebulous idea of “identity”. This has since comprised a concerted effort to study, control, and ultimately remodel all three of those aspects of the soul first mapped by Plato: the physical desires he called *eros*, the spirit or wish for recognition he called *thumos*, and the capacity to comprehend he called *logos*. But this time round, these are considered not as aspects of our inner life, to be governed by *logos* to the good of the whole. Now, they are raw material, to be captured by the three Singularities that define the “identity” era.

Who Am I?

What the ancients would have called *eros* came to seem amenable to mastery via contraceptive technologies. But in this, the first Singularity, it was less a project of mastering *eros* as wielding technology to unchain it without cost or consequence. These interventions, it was hoped, might unshackle sexual desire from those social asymmetries baked into differential male and female reproductive roles, and enable a broader liberation and utopian re-engineering of desire as such. In its wake, too, a new tech-enabled feminism flourished whose most radical visionaries painted a utopian picture of a world in which embodiment is entirely subordinate to desire.³⁴

Downstream of the moral, behavioural, and cultural revolution enabled by the Pill, whole new industries have flourished: in sex, in pornography, in large and small gametes, even gestation itself. And this logic has extended more broadly, too: the Pill's technologisation of *eros* could stand as metonym for what David Courtwright calls "limbic capitalism": an economy grounded in the intentional hijacking and commercial exploitation of human desires and pleasure centres, via super-stimuli such as junk food and pornography.³⁵ In other words: what was presented as a liberation of desire has proved, in practice, to represent its capture by industries calibrated to hacking our most atavistic pleasure and reward systems. *Eros* has not been freed so much as captured and strip-mined.

This capture of *eros* in turn loosened those aspects of secular "selfhood" that had previously seemed given and grounded in the body, and the normative social and family roles previously associated with the sexes. It represented a significant unmooring from the givens of our bodies, that left many asking: who am I really? This question that looms large over the twentieth century, and the Catholic social critic Mary Eberstadt argues that it gained such potency downstream of the solvent impact of the contraceptive paradigm.³⁶ For in the absence of external givens robust enough to provide answers to that question, it becomes more difficult to experience oneself as a "self" at all, except in the eyes of those around us—or, failing that, of political and bureaucratic structures. In this context, as Francis Fukuyama has argued recently, the desire for recognition—*thumos*—has grown in political salience over the twentieth century.³⁷

With faith imploding, ethnic and political nationhood rendered taboo by the wars, and the givens of embodiment itself increasingly reframed as optional and subject to individual control, a new set of cultural imperatives emerged to cater to this desire. Post-structuralist thinkers were prescient in anticipating this dynamic, and the hunger it encodes. The French psychoanalytic thinker Jacques Lacan, for example, theorised our inner lives as characterised by a "mirror stage" in which we come to "identify" with the version of ourselves reflected back to us in the gaze of the other—usually the mother—and derive the energy for psychic activity from the ever-present gap between this reflection and our full experience of ourselves.³⁸ Lacan's theory of "identification" was then further elaborated by queer theorist Judith Butler, who applied its principles

to our bodies, arguing in *Gender Trouble* (1990) that both sex and gender are socially constructed, with their seemingly immutable traits fixed through the coercive actions of ideology and power.³⁹

This anxious territory was turbo-charged around the turn of the millennium, by the digital revolution—one that became near-universal with smartphones, a few years later. For while, as Martha Nussbaum pointed out, Butler’s 1990 arguments offered little substantive scope for escaping these coercive ideologies, save a kind of defeatist “disruption” at the margins.⁴⁰ But the scope for radical self-reinvention seemed to grow limitless in the internet’s radically de-materialised social spaces. As social interaction diverged from in-person presence, the question “who am I?” took on both new urgency and formlessness—while also admitting of new answers.

Disembodied interaction appears to grant each individual the power to craft an “identity” for themselves based solely on inner desire. In the virtual realm, I no longer need to *be* whatever I wish to be recognised as; I can simply say I am that thing. Who is anyone to disagree? This was captured in a prescient 1993 *New Yorker* cartoon, in which one dog tells another: “On the internet, nobody knows you’re a dog.”⁴¹ The most familiar and politically well-established example of this phenomenon is transgender activism, but in theory the scope for disembodied identity-formation is limitless.

In the wake of this new era, following the digital revolution, a kind of pop-Butlerism has proliferated within digital subcultures, according to which all one need do is “identify as” a thing and one *is* that thing, irrespective of how plausibly this meshes with physical appearance or real-world circumstances. The incongruous juxtapositions that result, such as middle-aged men “identifying” as six-year-old girls,⁴² became bread-and-butter culture-war discourse in the immediate 2010s aftermath of the smartphone Singularity.

A central powerhouse for the multi-billion-dollar digital economy that emerges from this ferment of online self-creation is the capture and technologisation of *thumos*, the yearning for recognition. This arises from the way the digital revolution de-materialises our social identities. For in practice, it is not enough to just say I *am* something; to feel real and true, my claim requires the assent of those around me. In the bodiless world of the internet, this feels relatively easy to achieve. As it becomes easier to create such virtual identities, based supposedly on pure “identification”, so too it becomes easier to curate an online social group based not on accommodation with those in my physical proximity, such as family or neighbours, but rather with those selected based on interests or affinities. In other words, I can “find my people” anywhere in the world, and associate exclusively with them, while dismissing those in my immediate environment as secondary. I can then receive the recognition I crave, for the de-materialised “identity” I have crafted myself, from a de-materialised social group that has no reason not to grant that recognition.

Of course, this runs into difficulties once attempts are made to map this dynamic onto interactions in the physical world. This has, in turn, produced new ecosystems of power, money, and political activism ordered to demanding, disputing, or distributing resources in accordance with such abstract forms of “identity”: the now crowded and bitterly contested terrain of “wokeness” and contemporary identity politics. I will leave a treatment of that terrain to others; the key point, for our purposes here, is firstly to note that this economy of self-creation and recognition is every bit as parasitic on *thumos* as “limbic capitalism” is on *eros*, and is also profoundly inimical to practical politics and solidaristic movement-building of every kind. The swift collapse of the 2016 Women’s March movement into identitarian in-fighting is just one notorious case in point.⁴³ And secondly, to highlight what “identity” and “identification” themselves imply. Namely: that which could be said to have an “identity”, or to be “identifying” as one sex or the other, or as a dog, or mermaid, or who knows what else, is itself distinct from any of those attributes. The mirror’s reflection is not what it reflects. Even as the field in which “identity” is negotiated has grown ever more minutely analysed, politicised, and commodified, whatever it is that *identifies* has receded to some, as yet indeterminate, vantage-point.

Who is Asking?

If it is not our physical desires, or our yearning for recognition, what is at the core of our supposedly authentic selves? Who or what even asks the question? Plato would have said this is surely *logos*—our rational capacity for comprehension. We might note that deep reading as an habitual practice has declined significantly since the digital revolution has fostered. We might speculate, with theorists of literacy such as Walter Ong and Adam Garfinkle, that this phase shift in our use of information technologies has likely also altered our experience of interiority. Perhaps without this change, we might not believe it possible for a machine to “think”. In any case, this is the metaphysical, cultural, and political question that now confronts us: whether or not we are willing to accept, collectively, that *logos* is something a machine might one day achieve.

Is there any difference between the kind of pattern recognition and computation performed by AI, and the thinking performed by a human intelligence? Are computation and comprehension the same thing? It should be clear by now that I do not think this to be true. But it remains difficult to explain the nature of the error, from inside the modernist account of “reality” that has reached its pinnacle in the effort to mechanise *logos* itself. We can, though, draw some inferences from the results of extending the modernist paradigm to *eros* and *thumos*, on the likely outcome of applying the same principles to *logos*. For in none of these cases so far has the effort at instrumental control of our souls provided solutions, or even much engineering leverage, to the “problem” of human nature.

As I have argued, our efforts to re-engineer *eros* have not delivered a utopia of pleasure, but rather its deadening and flattening across the board: one in which the underlying asymmetries of embodiment and desire nonetheless persist, just less constrained by social guardrails and subject mainly to the impersonal laws of supply and demand.⁴⁴ Similarly, the expansion of commerce into *thumos* has not accorded perfect self-realisation and recognition for all. Rather, it has delivered a dead and empty proliferation of hollow identities: flags, parades, merchandise, personal branding, and behind it the same alienation as ever. Meanwhile, the power of *thumos* to orient individuals or societies toward honourable achievement, or meaningful action, has been captured and transformed into “clicktivism”, influencer status, and fissiparous political networks more oriented toward infighting than institution-building. In this bleak and hyper-mediated hall of mirrors, even groups with genuine political grievances can end up politically crippled, unable to organise purposefully at scale toward long-term ends.

What, then, will be the result of attempting to extend the mechanistic modern paradigm to *logos*, our capacity for comprehension? If the pattern holds, the result will not be thought, let alone god-like artificial general intelligence. Rather, it will—in however sophisticated a form—remain dead calculation, without understanding. Such calculation may have innumerable useful applications, but it cannot, by definition, ever cross the line to creative intelligence. Dead sexuality parasitises on reproductivity: phenomena such as pornography and hookup culture map onto falling rates of family formation. Dead identity politics parasitises on our capacity to form ourselves in relationship with (and service to) those around us, crippling our capacity to function as polities and social groups. And following a similar logic, we can predict that dead thinking—that is to say, the kind of probabilistic calculation based on datasets, weightings, and prompts—will parasitise on our capacity to comprehend our own world, and beyond this, to understand what distinguishes mechanical pattern recognition and the human capacity for comprehension.

The cost will be felt in countless small acts of impersonal cruelty, performed by computer decision-making in the absence of understanding; in pseudo-relationships that end in inhuman disaster, like the calamity that befell Sewell Setzer III; and, more pervasively, in a slow attrition of our capacity to comprehend, and with it to ascribe comprehension to humans, having accepted the personhood of machines in its absence. Indeed, at the fringes of transhumanist thinking stand those who argue simply that transhumanists are “not human”,⁴⁵ or that in the future digital selves should be accorded the same rights as humans.⁴⁶

There is a Place to Stand

If this prophecy seems bleak, there is hope. In his famous essay on technology, Martin Heidegger quoted the poet Hölderlin to the effect that the only way out is through: “where the danger lies, there also grows the saving power”. As we have seen, throughout our civilisational history our ideas concerning the nature and topography of the soul itself have developed in tandem with our understanding of the nature of reality more broadly, while both have been profoundly shaped by the information technologies at our disposal. Today the same dynamic applies. And the crucial epistemological direction—the one that both deprives us of a place to stand, but also returns that ground to us—has already been set, from within the quantum discoveries of science itself. And this foundation in turn hints at the most tantalising prospect of all: that we may yet find a way forward, without needing to discard the remarkable achievements of modernity to do so.

Most of what has ensued since 1900 has been a continuation of modernity by other means: re-orienting inward to the “self” the instrumental modernist mindset. The result has been a steady dismantling of the “self” into the malleable stuff of “identity”, now opened to commodification. But what we have, throughout the history of the West, referred to variously as the “soul” then, latterly, the “self” remains elusive. As Lacan noticed, it never becomes wholly identified with its identities. But this conundrum is not, as Lacan imagined, just an endlessly frustrating source of psychic hunger; it is the way out.

Chasing that elusive extra *something* has become the quixotic quest of AI research. A great many remarkable, powerful, and culture-changing discoveries have already been made as a byproduct of such efforts, and we should expect many more. But much as the long-term post-Pill feminist consequence of attempting to abolish sex difference was a new wave of feminist sex realism,⁴⁷ in attempting to mechanise *logos* we will learn the hard way *why* “artificial intelligence” is not *logos*.

We can capture this succinctly in the distinction between computation and thinking or—if you prefer—between AI and living souls. And we need not rewind even as far back as Ockham, or St Augustine, let alone Plato, for conceptual tools with which to frame this. We need only step back a little over a century, to Max Planck and the great weirding of reality that never quite reached the mainstream. The central insight of this missed revolution in our metaphysics is that the universe of classical physics—while an adequate account of reality for everyday purposes—is itself a byproduct of forms of reality, and of consciousness, that are not reducible to mechanism. For quantum physics, consciousness is both irreducible, and also coextensive with the world outside our heads. If this is so, then as the philosopher and technologist Bernardo Kastrup puts it, “the substrate of mind itself is not stuff: it is the subject, not an object.”⁴⁸

If this interpenetration of reality and idea is accepted in principle, so too accepting its entanglement with our own consciousness becomes a logical inevitability. And from

this it follows that we can only answer the question posed by AI by exiting the modernist cul-de-sac, aided by the discoveries of Planck, Feynman and others. And in doing so, we retrieve a twenty-first-century version of the “ontic logos”, that premodern intuition of the infinitely complex interpenetration of that *something* within us that perceives, and that which we perceive.

The heart of this understanding of reality is our willingness to insist on a conceptual distinction between that *something*—the soul—and those soul simulacra we are able to engineer. For if, as Kastrup puts it, “mind” is not stuff but “the substrate” of what we perceive as stuff, then it follows that “all of reality—the entire universe—is in mind”.⁴⁹ But it also follows that while this encompasses our mechanical inventions, the converse is not true. Once we grasp that, it is easy to understand that computation and comprehension are not the same thing. And then it becomes easier to see why—for example—there is a qualitative difference between giving your child an AI-powered language learning tool, and giving your child an AI-powered digital “friend”. It also becomes clear why “identity” is not the self, or the soul, but merely the dead byproducts of our failed attempts to commodify the soul.

There is a great deal more work to be done in distinguishing the proper use of digital calculation, from mistaken efforts to conflate such calculation with comprehension. That work is beyond my scope here. My point is simple, and optimistic: once we accept that the human soul is irreducible to mechanical calculation, we need neither reject the digital era (and its Singularities) out of hand, nor fear being subsumed by them. Instead, we can once again orient ourselves toward genuine progress, in the sense of committing our collective endeavours to pursuing human flourishing across all three aspects of our nature: *eros*, *thumos*, and *logos*, understood as integrated parts of a whole.

From a vantage-point that retrieves the soul as a necessary predicate for reality as such, the notion that humans might be conflated with our own mechanical inventions is as self-evidently absurd as the eighteenth-century fantasy of engineering clockwork humans. It is likely, too, that from this vantage-point we will rediscover the value of pre-modern theorists of the “ontic logos” and the human soul, to a postmodern physics and technology capable of encompassing these complexities. Enough, perhaps, to free us from the quagmire of dead “identity” and rise once again, as the souls we actually are.

- ¹ Google Books Ngram Viewer, "Identity", https://books.google.com/ngrams/graph?content=identity&year_start=1800&year_end=2022&corpus=en&smoothin g=3, Retrieved 28 October 2024.
- ² It is of course true that literacies, technologies, and theories of the soul have flourished outside the West over recorded history, and interwoven with the West. My writing space is limited, and the Western tradition remains the dominant cultural and technological driving force today, and must therefore be my focus in this paper. This focus should not be taken to imply erasure of other traditions or their contributions to world culture.
- ³ Ray Kurzweil, *The Singularity Is Near: When Humans Transcend Biology*, London: Penguin (2006).
- ⁴ Mary Harrington, "Transhumanism is Already Here", lecture delivered 20 January 2023. <https://unherd.com/newsroom/mary-harrington-transhumanism-is-already-here/>, Retrieved 1 November 2024.
- ⁵ Apple Newsroom, "Apple Reinvents the Phone with iPhone", 9 January 2007, <https://www.apple.com/newsroom/2007/01/09Apple-Reinvents-the-Phone-with-iPhone/>.
- ⁶ Ericsson, "Ten years of the Ericsson Mobility Report", <https://www.ericsson.com/en/reports-and-papers/mobility-report/mobility-report-journey/10-years>.
- ⁷ See for example, David Rozado, "Where did the Great Awakening come from?", *Unherd*, 8 September 2022, <https://unherd.com/newsroom/where-did-the-great-awakening-come-from/>.
- ⁸ Maxwell Zeff, "This Guy Used ChatGPT to Talk to 5,000 Women on Tinder and Met His Wife", *Gizmodo*, 7 February 2024, <https://gizmodo.com/guy-used-chatgpt-talk-5-000-women-tinder-met-his-wife-1851228179>.
- ⁹ See for example, Elise Bohan, *Future Superhuman*, 2022.
- ¹⁰ Eliezer Yudkowsky, "Pausing AI Developments Isn't Enough. We Need to Shut it All Down", *TIME*, 29 March 2023, <https://time.com/6266923/ai-eliezer-yudkowsky-open-letter-not-enough/>. Retrieved 30 October 2024
- ¹¹ Chris Vallance, "'Sickening' Molly Russell chatbots found on Character.ai", *BBC News*, 30 October 2024, <https://www.bbc.co.uk/news/articles/cg57yd0jr0go>, Retrieved 30 October 2024.
- ¹² Devindra Hardawar, "Living with Moxie, the robot companion for kids", *Engadget*, 22 March 2024, <https://www.engadget.com/moxie-robot-kids-hands-on-embodied-120031341.html>, Retrieved 30 October 2024.
- ¹³ See the comments on this YouTube short about Moxie, https://www.youtube.com/shorts/ZmET2_Ta-aY?app=desktop Retrieved 30 October 2024.
- ¹⁴ See the comments on this TikTok about Moxie, <https://www.tiktok.com/@reviewed/video/7133757171065670958?lang=en>, Retrieved 30 October 2024.
- ¹⁵ Marc Andreessen, Ben Horowitz, *The Techno-Optimist Manifesto*, October 16, 2023, <https://a16z.com/the-techno-optimist-manifesto/>, Retrieved 30 October 2024
- ¹⁶ W.J. Ong, *Orality and Literacy*, Routledge (1982).
- ¹⁷ A. Garfinkle, "The Erosion of Deep Literacy", *National Affairs*, 2020, <https://www.nationalaffairs.com/publications/detail/the-erosion-of-deep-literacy>.
- ¹⁸ Charles Taylor, *Sources of the Self*, Cambridge: Cambridge University Press, (1989), p. 144
- ¹⁹ See Larry Siedentop, *Inventing the Individual*, London: Penguin (2014).
- ²⁰ Siedentop, p. 308.
- ²¹ Ong, p. 149.
- ²² "Discours de la Méthode", in *Oeuvres de Descartes*, ed. Charles Adam and Paul Tannery (Paris: Vrin, 1973) VI, pp. 61-62. Cited in Taylor, p. 149.
- ²³ Taylor, p. 148.
- ²⁴ Taylor, p. 174.
- ²⁵ William Paley, *Natural Theology*, (1802).
- ²⁶ Feynman, "Nobody understands Quantum Mechanics", <https://youtu.be/w3ZRLlIWgHI?si=XXKyBjlb7yMrWef0>, Retrieved 11 November 2024.
- ²⁷ Isaac Newton, *Opticks*, (1704).
- ²⁸ Federico Faggin, *Irreducible: Consciousness, Life, Computers, and Human Nature*, Washington: Essentia, (2024) p. 40.
- ²⁹ Sean Carroll, "Even Physicists Don't Understand Quantum Mechanics", *New York Times*, 7 September 2019, <https://www.nytimes.com/2019/09/07/opinion/sunday/quantum-physics.html>.
- ³⁰ Donald Hoffman, *The Case Against Reality: how evolution hid the truth from our eyes*, Penguin (2019).
- ³¹ Ronald Knox, *God and the Atom* (1945).
- ³² John Hayward, "The Decline of the Church of England", *Church Growth Modelling*, 9 October 2013, <https://churchmodel.org.uk/2013/10/09/the-decline-of-the-church-of-england/>.
- ³³ Faith Survey, "UK Christian Statistics 2: 2010-2020: Introduction", <https://faithsurvey.co.uk/download/csintro2.pdf>.
- ³⁴ See for example Shulamith Firestone (1970) *The Dialectic of Sex*. This edition Verso, 2015, p. 187.
- ³⁵ David Courtwright, *The Age of Addiction: How Bad Habits Became Big Business*. Cambridge, MA: Harvard University Press, (2019).
- ³⁶ Mary Eberstadt, *Adam and Eve after the Pill: Paradoxes of the Sexual Revolution*, San Francisco: Ignatius Press, (2013).
- ³⁷ Francis Fukuyama, *Identity: the Demand for Dignity and the Politics of Resentment*, New York: Farrar, Strauss and Giroux Books (2018).
- ³⁸ Jacques Lacan, *The mirror stage as formative of the function of the I as revealed in psychoanalytic experience, Ecrits: A Selection*, New York: Norton, (1977), pp. 1-7.
- ³⁹ Judith Butler, *Gender Trouble: Feminism and the Subversion of Identity*, New York: Routledge, 1999 (1990), pp. 23-25.

⁴⁰ Martha C. Nussbaum, "The Professor of Parody", *The New Republic*, 22 February 1999, <https://newrepublic.com/article/150687/professor-parody>.

⁴¹ Thomas Germain, "'On the Internet, Nobody Knows You're a Dog:' The Web's Most Iconic Cartoon Is for Sale.", *Gizmodo*, 15 September 2023, <https://gizmodo.com/new-yorker-dog-internet-cyberdog-cartoon-auction-1850840101>.

⁴² Emily James, "'I've gone back to being a child': Husband and father-of-seven, 52, leaves his wife and kids to live as a transgender SIX-YEAR-OLD girl named Stefunknee", *Daily Mail*, 11 December 2025, <https://www.dailymail.co.uk/femail/article-3356084/I-ve-gone-child-Husband-father-seven-52-leaves-wife-kids-live-transgender-SIX-YEAR-OLD-girl-named-Stefonknee.html>.

⁴³ Shikha Dalmia, "Why the Women's March on Washington has already failed", *The Week*, 2 January 2017, <https://theweek.com/articles/667163/why-womens-march-washington-already-failed> Retrieved 6 November 2024.

⁴⁴ Harrington, *Feminism Against Progress*, (2023) pp. 81-103.

⁴⁵ Rachel Haywire, John Schoenecker, *A Transhuman Separatist Manifesto* (2010) <https://web.archive.org/web/20110202103153/http://hplusmagazine.com/2010/12/09/transhuman-separatist-manifestotop-5-reasons-transhumanist-militia/> Retrieved 1 November 2024.

⁴⁶ See for example Martine Rothblatt, *From Transgender to Transhuman: A Manifesto for Freedom of Form*, 2011, Kindle edition, loc. 2848.

⁴⁷ See "Fairer Disputations: Sex-Realist Feminism for the 21st Century", <https://fairerdisputations.org/>, for more.

⁴⁸ Bernardo Kastrup, *Why Materialism Is Baloney*, Alresford: ifff books (2014), p. 67.

⁴⁹ Kastrup, p. 59.



This research has been produced and published by ARC Research, a not-for-profit company limited by guarantee registered at Companies House with number 14739317, which exists to advance education and promote research.

ISBN: 978-1-916948-46-4

www.arcforum.com

February 2025