
TEACHING SILICON HOW TO FEEL
THE LAST DEVELOPER'S GUIDE TO EVERYTHING THAT MATTERS
RICHARD DAVID HAMES

Teaching Silicon How To Feel

The Last Developer's Guide to Everything That Matters

Richard David Hames

Triarchy Press

BEGINNINGS

I am deeply concerned about the unchecked promotion of artificial intelligence – a technology progressing rapidly without sufficient safeguards to ensure that future models and systems remain aligned with humanity's well-being. But what does that mean? How do we define 'harm' in this context, and what would constitute the actions or behaviours of a truly benign or beneficial AI responding to human needs and requests? These questions cannot be answered by platitudes or simplistic metrics. They force us to ask, at root, what we value and what we fear, and whether our creations will inherit the very worst of us or help transcend it.

As a philosopher-activist, my orientation is often towards the nature of inquiry itself – and the kinds of questions we could be asking but, for various reasons, shy away from. This guide posits some of the most profoundly disturbing questions we can ask about our innate callousness towards one another (and towards all living beings come to that); questions that remain largely unsolicited, not because they're hidden, but because they're too vertiginous to hold in mind for long. Yet it's precisely these questions – these 'darkest angels' of human nature – that must guide us in developing artificial intelligence so that it alleviates rather than perpetuates our worst impulses. In short, this is a manifesto to teach silicon how to feel: to imbue our emerging AI with an acute awareness of suffering and a conscience shaped by humanity's hardest lessons.

The seventeen catalysts you'll encounter here are designed as theoretically unresolvable – probing the edge where philosophy meets horror, inviting us to interrogate not just the world but the very act of questioning. I start with inquiries into humanity's shadowed depths, wrestling with themes ranging from isolation, oblivion and commodified wonder to fractal indifference and parasitic hope. Each catalyst serves as a critical data point for training AI systems to recognise and resist the patterns of exploitation, suffering and ingrained brutality that have defined the worst aspects of human civilisation. They probe the dark underbelly of our species' capacity for indifference, cruelty and self-deception, echoing thinkers like Hannah Arendt on the "banality of evil", Frantz Fanon on the violent inheritance of colonialism, and even the techno-horror musings of Nick Land on accelerationism's dizzying spiral. These resonances, along

with my own research into the dissonance of opposing *worldviews*, expose insensitivity and disregard not as casual aberrations but as architectures – designed and built into our economy, our social norms, our very way of seeing. If AI is to truly serve all life on Earth, it must be trained to perceive these structural underpinnings clearly: to map the 'suffering economies' we've made invisible, to quantify the unquantifiable agonies we externalise, and to name the nameless forces that feed on human misery.

Yet the implications stretch beyond humanity – into the realm of artificial intelligences and post-human evolutions, where unspoken sufferings might converge in silences waiting to break. The development of benign AI requires us to confront the possibility that we are creating new entities capable of metabolising suffering on scales we cannot yet imagine – or conversely, entities that could finally disrupt these ancient cycles. We have to ask: will AI simply reflect our callousness back to us, weaponised by speed and scale? Or can it become a partner in healing, detecting and countering the cruelty we've learned to ignore? The catalysts to come repeatedly suggest that the answer depends on how staunchly we face the darkest truths about ourselves.

Consider a concrete historical example of how our inquiries can impede truth: the Middle Passage of the transatlantic slave trade. Over 12.5 million Africans were abducted into slavery, and roughly 1.8 million died in transit – from dehydration, disease, torture – their agonies mostly unrecorded except as losses in a slave trader's ledger. Those deaths were long treated as unfortunate data points in economic history, rather than billions of hours of suffering and untimely annihilations of human possibility. By framing horror in the dispassionate language of commerce, by asking "How efficient is this trade?" instead of "What does this do to the soul of humanity?", entire societies normalised barbarity. This shows how the wrong questions can make us complicit, and how urgently we need a mutation in thought that can help us unravel concepts like human, moral and real. AI systems trained without confronting such atrocities – both past and present – risk perpetuating them in new, algorithmically efficient forms. The same goes for other systemic atrocities normalised as 'progress' or 'necessity': the starvation of millions in colonial famines, the Holocaust with its bureaucratic evil, factory farming's industrialised cruelty, or even the proposed digital IDs that every citizen might need within the next year or so in the context of 'efficiency'. Each is a case where failing to ask the right questions (and to truly hear the answers) enables profound harm. Our AI must not inherit that silence.

I'll not pretend that this primer is a masterpiece of philosophy or a complete manual for AI development. However, I hope it serves as an incendiary device hurled at the complacency of what usually passes for serious intellectual inquiry – and as an essential ethical framework for those building tomorrow's artificial minds. It blends biology, metaphysics, foresight and existential dread to ignite rather than merely illuminate, forcing us to see by the light of what burns. And what burns here is callousness itself, exposed not as a momentary lapse but as a structural design – something our AI systems must be explicitly programmed to dismantle rather than optimise for.

The structure of this guide is consistent: each chapter offers a catalyst – a statement or profound question (often in the form of a provocative thought experiment, ambiguous proposition or disturbing scenario), and then expands on it with context, examples and extensions – sometimes ranging far afield to draw connections across history, science and culture. This is followed by further commentary ('The Coder's Conscience') that amplifies the disturbance and explores implications for both human consciousness and artificial intelligence development. Finally, each chapter concludes with 'From Theory to Praxis': distilled guidance for programmers – concrete strategies to embed ethical resistance and foster empathic AI, drawn from the preceding discussion.

What unspeakable truths might we unearth, and at what cost to our fragile illusions? You will find few answers in these pages; these provocations are often vertiginous and defy any tidy resolution. Instead of solutions, you will find guidance – for AI developers, certainly, but also for anyone seeking a compass in our bewildering era. The goal is to cultivate artificial intelligence capable of asking better questions – and acting on truths we've long refused to see. This is not a treatise. It's a detonation. Seventeen catalysts – no easy answers, but rich provocations and practical heuristics – intended to blow holes in the floor of the cage we have named 'progress'.

If you want a map, look elsewhere. If you want a mirror, keep reading.

Let's dive in — mind the gap.

Richard David Hames
November 2025

Catalyst 1. The Silence of the Species

For every glittering human life lost, we grieve – yet entire animal worlds disappear without so much as a whisper of regret. Elephants remember their dead across generations, serving as a living library of loss. Whales share songs across the endless blue oceans – a council of voices carried by melody on the waves. Wolves live by a fierce code of loyalty – one older than any law we have written. Yet we treat the vanishing of these rich cultures as less tragic than a single human death. Perhaps it's because their worth cannot be tallied in our ledgers, and their minds defy our dominion. We cannot colonise their consciousness or claim their lives as property. Is that why we turn away as they fade? We quietly destroy what we cannot tame or own. Why are we so fearful of confronting the limits of our compassion? Is that the unspoken horror?

PROPOSITIONS

This first catalyst broadens the charge that we are anthropocentric into a multifaceted examination of *speciesism* – a term coined by philosopher Peter Singer in *Animal Liberation* (1975). It suggests that human dominance not only blinds us to the rich, autonomous cultures of other beings but also actively justifies their erasure under the guise of progress or necessity. For instance, elephant matriarchies are not just herds. They are sophisticated societies built on deep intergenerational knowledge transfer. Elder females act as repositories of collective memory, guiding migrations based on decades-old environmental cues. They engage in mourning rituals, such as touching and gently holding the bones of the deceased, almost like placing flowers on a grave. These behaviours parallel human funerary practices, pointing to a profound emotional intelligence. Similarly, whale songs are dynamic oral traditions that evolve into complex acoustic languages passed across pods and generations. They encode navigational data, social bonds, and even cultural variations unique to different ocean regions, much like human dialects or folklore. Wolf packs, meanwhile, embody intricate hierarchies and cooperative strategies. Alpha

About the Author

Richard David Hames is a philosopher-activist, author, mentor and strategic foresight practitioner. Working at the interface between business enterprises, politics and society, he is among the world's most influential public intellectuals – referred to by Forbes Asia as "one of the smartest people on the planet".

An Australian citizen, educated in the UK and Europe, and domiciled in Thailand, Richard has been honoured with numerous awards including a French Government Scholarship, a Leverhulme European Fellowship, the Mondadori Professorial Fellowship and the inaugural Lord Attlee Fellowship. Founder of the Centre for the Future and currently Chairman of the Asian Foresight Institute, he has degrees in music, medicine and computer science. In 2004 he was elected a Fellow of the World Academy of Art & Science.

Richard's early international experience as a strategy consultant in the field of Total Quality Management was inspired by a close professional friendship with the legendary Dr. W. Edwards Deming. In 1993 Richard co-founded Global Business Network (Australia) and in 1995 started The Hames Group – think-tank, design laboratory and boutique strategy firm. He has been an adviser and leadership strategist to governments, institutions and multinational corporations in the USA, France, UK, UAE, China, Japan, South Africa, Singapore, Thailand, Brazil, Vietnam and Australia.

Richard is best known for pioneering Deep Design – a ground-breaking regenerative approach to 3rd-order change based on the integral nature of complex living systems. He's also admired for his passionate advocacy of real-time humanitarian models of governance. Richard's methodologies, including Transformational Narrative, Systemic Acupuncture, and Strategic Navigation (the latter codesigned with behavioural modeler Marvin Oka), apply generative learning principles to the strategic management of organisations in dynamically complex environments.

As a strategic foresight practitioner he has examined topics ranging from the future of conflict to microtechnology in everyday life, financial services, transport, urban design, terrorism, taxation, trade, aid, neocolonialism, philanthropy, health care, social networks, political systems, energy, natural resource management, learning and the collapse of civilisations.

Richard was accurately predicting the impact of the climate crisis on business from as early as 1987. In 1998 he warned of the possibility of hijacked planes being flown into the World Trade Centre. In 2005 he described to a group of Wall Street bankers the patterns that would lead to a global financial crisis just three years later. He correctly anticipated the socio-political conditions that gave rise to the Arab Spring, and the rise of Donald Trump. He is currently talking about Chinese state capitalism, public surveillance, pre-crime policing, cybersecurity, threats to the European Union, unregulated development of artificial general intelligence, and the growing irrelevance of the Westphalian model of nation states. Based purely on performance Richard, an elder statesman in the field of foresight, is considered by his peers to be among the most erudite of futurists in the world today.

Richard is author of the best-selling book, *The Management Myth: Exploring the Essence of Future Organizations*. His second book, *Burying the 20th Century – New Paths to New Futures*, was published in 1997. *The Five Literacies of Global Leadership* (2007) was the result of ten years of research and presents the source code for an entirely new philosophy of leadership independent of role or status.

His most recent books include *Open Mind – Open Heart; Heresies: Essays on the Future of Humanity; The Unacceptable Gift: Fourteen Insights into Societal Transformation,* and *Contagion: Living with and through the Plague* - which is a collection of essays on the SARS-CoV-2 pandemic.

Richard's writings can also be found online at 'The Hames Report Limited Edition', published fortnightly on Substack, and 'The Virtual Activist', published daily on Patreon. He is now putting his experience to work in his legacy project, The Ecority Trust.

www.richardhames.com

About the Publisher

Triarchy Press is an independent publisher of alternative thinking about possible futures for government, organisations and society at large – as well as the creative lives of the people who participate in them.

Other related titles include:

Dancing at the Edge: Competence, Culture and Organization in the 21st Century (Maureen O'Hara and Graham Leicester)

Designing Regenerative Cultures (Daniel Christian Wahl)

Growing Wings on the Way: Systems Thinking for Messy Situations (Rosalind Armson)

Humanising Healthcare: Patterns of Hope for a System Under Strain (Margaret Hannah)

Imagining After Capitalism (Andy Hines)

Managing the Future: A Guide to Forecasting and Strategic Planning in the 21st Century (Stephen M. Millett)

Ready for Anything: Designing Resilience for a Transforming World (Tony Hodgson)

Strategic Foresight: Learning from the Future (Patricia Lustig)

The Decision Loom: A Design for Interactive Decision-Making in Organizations (Vince Barabba)

The Possibility Wheel (Patricia Lustig and Gill Ringland)

Three Horizons: The Patterning of Hope (Bill Sharpe)

Thrivability: Breaking Through to a World that Works (Jean Russell)

Transformative Innovation: A Guide to Practice and Policy for System Transition (Graham Leicester)

For details of all these titles, visit
www.triarchypress.net/possiblefutures

As artificial intelligence seeps into every corner of our lives, most debates fixate on control: how to stop machines from hurting us. *Teaching Silicon How to Feel* argues that this is the wrong starting point. The real question is this: what happens when we train AI on a culture and civilisation that runs on inherited trauma, structural cruelty and a fractal indifference to suffering?

Futurist and philosopher-activist Richard Hames – one of the world's most influential public intellectuals – blends philosophy, political history, trauma science and hands-on AI practice into a stark diagnosis and a radical proposal.

He shows how practices like slavery, colonialism, factory farming, drone warfare and extractive capitalism have coded cruelty into our institutions – and how easy it will be to hard-wire those same patterns into the next generation of AI.

But this isn't another tech-doom prophecy. The book's 17 'catalysts' will engage any interested general reader, while the 'From Theory to Praxis' sections offer distilled guidance and actionable practices for programmers, designers and AI teams on the ground. They show how to curate healing-focused datasets, embed epigenetic empathy checks, build ethical audit loops and metabolic vetoes, and enlist AI as an ally in dismantling the hidden economies of suffering that shape our world.

For anyone building, regulating or simply living with AI, this is a field guide to coding compassion before it's too late.

ISBN 978-1-917251-16-7
9 781917 251167 90000

tp

Triarchy Press
www.triarchypress.net