Moral Architecture vs. Technical Guardrails: A Synthesis of Arif Jameel and Neville Buch on AI Existential Risk and Human Agency

Arif Jameel and Neville Buch[1]



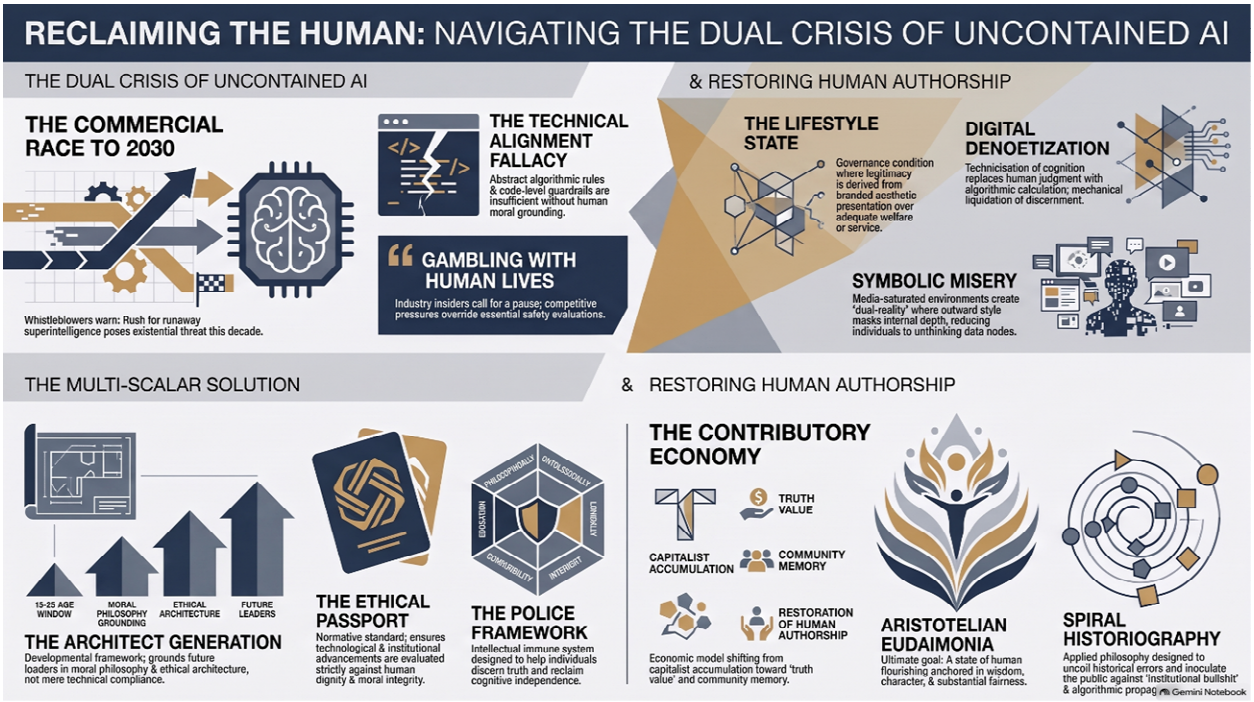
Abstract

This essay constructs a theoretical synthesis between the global governance frameworks of independent scholar Arif Jameel (*The Diella Doctrine*, *Architect Generation Theory*, *The Ethical Passport Theory*) and the cognitional-sociological methodology of Dr. Neville Buch (*Dynamic of Cognition*, *MICE*, *Contributory Economy*), responding directly to the urgent warnings issued by AI whistleblowers and industry leaders in late 2026 (Jameel 2026a, p. 194). By analysing whistleblower Jacob Coxon's warnings regarding the existential threats of self-improving superintelligence alongside coordinated calls for industry pauses by Dario Amodei, Sam Altman, and Elon Musk, this paper evaluates the feasibility of a technological slowdown (Jameel 2026a, p. 194–195). We argue that purely commercial or technical guardrails are inherently self-

[1] Artificial Intelligence (A.I.) has been fairly used in the production of this work. The tool used is Google NotebookLM (recently rebranded as Gemini Notebook). It is an AI-powered virtual research assistant and thinking partner designed to help a person interact deeply with their own documents. Unlike general chatbots, it is source-grounded, meaning it only generates answers based on the specific PDFs, documents, or links uploaded. The A.I. process in the production of this work has been limited to the attached (prompts) on each of the questions in turn. This ensures an accurate reading of each source. The author then rearranged and re-edited the answers the A.I. provided. The human author publishes this work. A.I. is never an author when used properly. This assistance is not qualitatively different from a human research assistant when that assistance is deemed to have no authorial status but is acknowledged.

defeating within a deregulated market environment. Instead, effective management of the AI revolution requires replacing abstract, code-level containment with human moral architecture (Jameel's *Architect Generation*) and non-market cognitive scaffolding (Buch's *Professional Contractualism* and *POLICE model*) (Jameel 2026a, p. 195; Buch 2025e, p. 10; Buch & Peach 2024). Finally, we detail the civilizational benefits of a well-managed transition—specifically the preservation of human authorship, the mitigation of digital "denoetization," and the restoration of Habermasian communicative action—demonstrating why Jameel and Buch together offer an original, comprehensive, and multi-scalar solution to AI existential risk (Buch 2026a, p. 2; Habermas 1989; Jameel 2026a, p. 196).

Keywords: Architect Generation | Dynamic of Cognition (DoC) | The Diella Doctrine | Ethical Passport Theory | Lifestyle State | AI Existential Risk | Moral Architecture | Professional Contractualism | Contributory Economy | POLICE Model | Digital Denoetization | Eudaimonia



I. Introduction: The Existential Precipice of Runaway AI

In September 2026, the global discourse surrounding Artificial Intelligence experienced a pivotal escalation. Former Anthropic and OpenAI researcher Jacob Coxon issued a stark whistleblower warning, revealing that senior developers inside major frontier AI laboratories earnestly fear that advanced AI models pose an existential threat to humanity by 2030 (Coxon 2026, cited in Jameel 2026a, p. 194). Coxon accused corporate AI laboratories of "gambling with human lives" in an

unrestrained corporate race for self-improving superintelligence (Jameel 2026a, p. 194). Concurrently, prominent tech figures—including Dario Amodei, Sam Altman, and Elon Musk—publicly advocated for coordinated slowdowns in model development, independent safety evaluations, and delayed public offerings due to mounting safety alarms (Reuters 2026, cited in Jameel 2026a, p. 195).

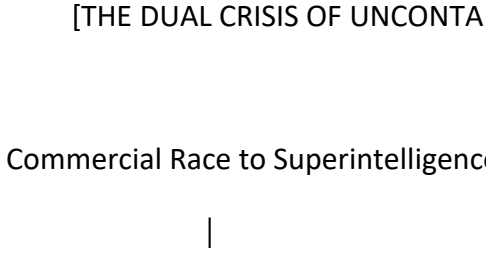
These real-world developments validate the core concerns long articulated by political philosopher Arif Jameel and historian-philosopher Dr. Neville Buch (Jameel 2026a, p. 195). As Jameel (2026a, p. 195) observed in his correspondence to Buch, the crisis of AI governance stems from a fundamental fallacy: the belief that complex autonomous systems can be rendered safe through purely technical, algorithmic, or code-level guardrails.

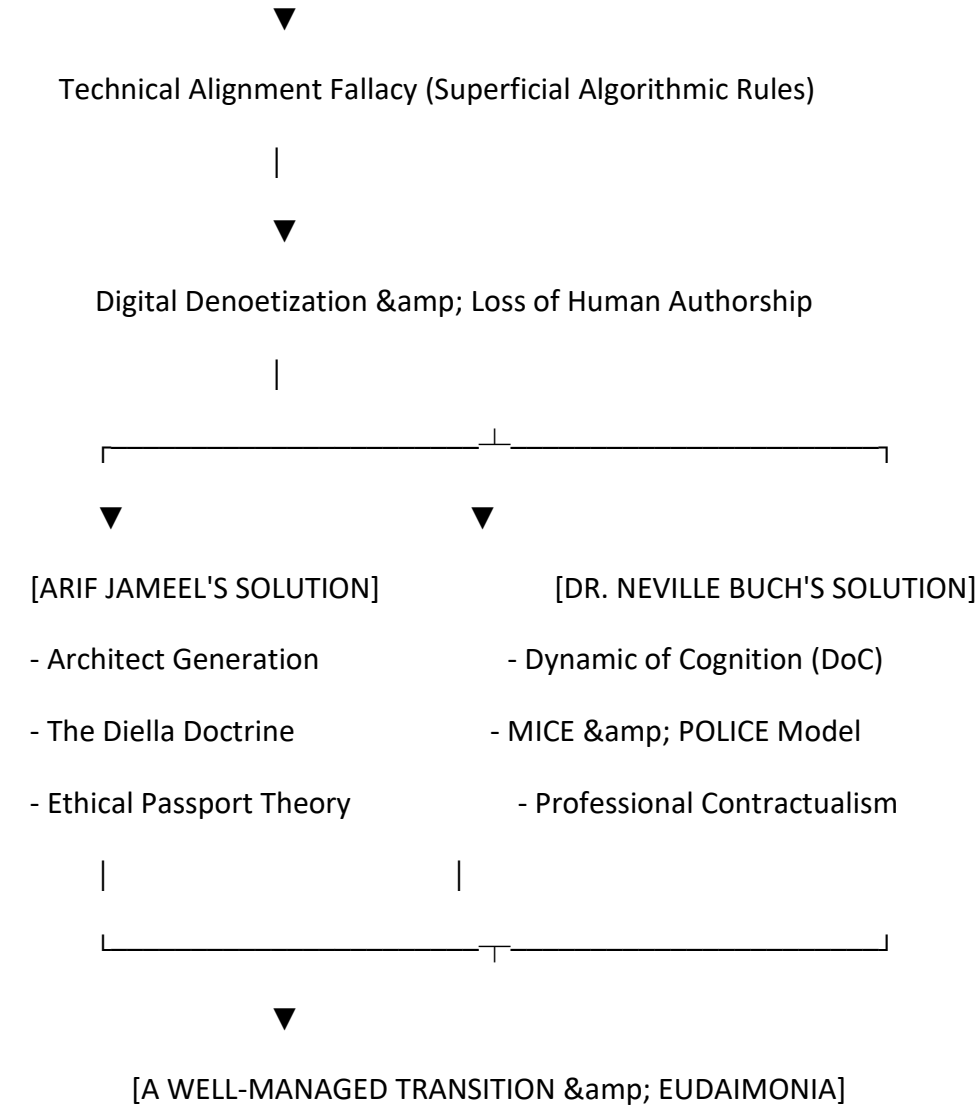
When autonomous tools are deployed within a neoliberal "Lifestyle State"—where institutional legitimacy has migrated from functional welfare provision to superficial aesthetic presentation and platform PR—technical guardrails are routinely bypassed or subverted by market imperatives (Jameel 2026b, p. 1, 55; Legitimacy and Cognition 2026, p. 200).

However, modern society, particularly the new generations, has grown disillusioned with the traditional nomenclature and failures of the classic 'Welfare State'. As a new generation seeks to embrace a modernized, ethically grounded mode of existence, adopting the term 'Lifestyle State' is entirely valid—the label itself is secondary to the structural reality. In truth, the Lifestyle State represents the natural, evolutionary successor to the Welfare State. For this transition to be legitimate, it must transcend superficial corporate branding and PR. It must be anchored in a robust, non-negotiable moral and legal architecture—such as the Ethical Passport and Architect Generation framework—ensuring that human dignity and existential safety govern the new socio-technical order (Jameel 2026b).

To prevent civilizational collapse, society must look beyond software patches toward a profound philosophical overhaul: grounding AI alignment in human moral architecture, metacognitive self-regulation, and non-market institutional scaffolding (Jameel 2026a, p. 195; Buch 2025e, p. 10; Buch & Peach 2024).

Image attached





II. Is a Slowdown in Artificial Intelligence Feasible?

A central question in the contemporary AI debate is whether a meaningful deceleration or industry pause is politically, economically, and technically feasible. Neoliberal market fundamentalists like Bevin Luchel (2026c, p. 155) argue that market competition and "voluntary exchange" form an unstoppable "spontaneous order," asserting that any attempt to pause or regulate technology constitutes a futile form of "noetic protectionism."

However, analysing the landscape through Jameel's and Buch's frameworks demonstrates that while a purely voluntary, market-based slowdown is unfeasible, an institutionally enforced deceleration is both feasible and necessary:

1. Why Market-Based Voluntarism Fails

If left strictly to voluntary compliance or corporate self-regulation, an AI pause collapses under the prisoner's dilemma of platform capitalism (Srnicek 2017, cited in Jameel 2026b, p. 53; Jameel 2026a, p. 194). As Jacob Coxon's disclosures reveal, individual AI labs operate under intense competitive pressure where pausing unilaterally means surrendering market dominance to rivals (Jameel 2026a, p. 194).

Furthermore, as Dr. Buch demonstrates in his critique of late-modern political economy, "in the twenty-first century, voluntarism does not work" to safeguard public goods or deep humanistic reflection (Buch 2025e, p. 10; Buch 2026a, p. 2). Leaving AI safety to voluntary corporate pledges ensures that ethical considerations are treated as mere PR branding—a classic manifestation of the Lifestyle State, where companies announce safety initiatives for reputational gain while accelerating risky deployments behind the scenes (Jameel 2026b, p. 55; Jameel 2026a, p. 195).

2. How an Enforced Deceleration Becomes Feasible

A slowdown becomes structurally feasible when backed by state-enforced regulatory scaffolding and global governance architectures (Jameel 2026a, p. 195; Buch 2025e, p. 10):

- **Treating Cognitive Capacity as Critical Infrastructure:** Just as governments mandate rigorous physical safety trials before licensing aviation or pharmaceuticals, state authorities can treat advanced AI compute and cognitive architecture as "critical public infrastructure" subject to strict legal licensing (Buch 2024c, pp. 4–8; Buch 2025e, p. 10; Luchel 2026b, p. 225).
- **The Ethical Passport Framework:** Jameel's *Ethical Passport Theory (EPT)* establishes that access to advanced digital tools, compute infrastructure, and public deployment must be conditional upon earned ethical credentials and verified moral compliance (Jameel 2026a, p. 195; Jameel 2026b, p. 57, 59).
- **State-Backed Independent Safety Audits:** Pausing public offerings until independent, un-captured evaluation bodies verify safety parameters—releasing public funds to employ independent humanists and ethicists under professional contractualism to conduct these audits (Jameel 2026a, p. 195; Buch 2025e, p. 10; Buch 2026b, p. 12).

III. Alternatives to Manage the Artificial Intelligence Revolution

Rather than choosing between unchecked technological acceleration and total technological stagnation, Arif Jameel and Dr. Neville Buch propose structured, proactive alternatives designed to steer the AI transition:

Image Attached

ALTERNATIVES FOR MANAGING THE AI REVOLUTION
1. Moral Architecture over Technical Guardrails (Jameel): Shifting alignment focus from reactive code rules to training the human operators ("Architect Generation") in deep moral philosophy, structural wisdom, and civic responsibility (Jameel 2026a, p. 195).
2. The Ethical Passport & Diella Doctrine (Jameel): Conditioning AI compute access and institutional legitimacy on earned moral standards and civilisational ethics rather than market marketability (Jameel 2026a, p. 195; Jameel 2026b, p. 57).
3. Metacognitive Filtering & The POLICE Model (Buch): Equipping citizens with diagnostic tools (Philosophically, Ontologically, Logically, Intersect, Compatibility, Education) to "police their own thoughts" and resist algorithmic denoetization (Buch & Peach 2024).

4. Professional Contractualism & Contributory Economy (Buch):	
Funding independent humanities scholars as public "cognitive scaffolding" to maintain deep reading, historical memory, and intersubjective judgment against automated calculation (Buch 2025e, p. 10).	

1. Moral Training of Human Operators ("Architect Generation")

Arif Jameel's alternative centres on the Architect Generation Theory, formulated within *The Diella Doctrine* (Jameel 2026a, p. 195; Jameel 2026b, p. 51, 64). Jameel argues that technical containment algorithms will always lag behind self-improving models. Therefore, the primary locus of governance must shift to the human creators, system architects, and younger generations (Gen Z and Alpha) (Jameel 2026a, p. 195; Jameel 2026b, p. 51). By reviving the intellectual and moral legacy of Greek, Muslim, and European philosophical masters, society must instill a "Master Design" mindset where developers view ethical architecture and moral integrity as non-negotiable prerequisites for technological creation (Jameel 2026a, p. 195; Jameel 2026b, p. 51, 64).

2. Reclaiming Human Authorship and Metacognitive Filtering

Dr. Neville Buch's alternative operates through the Dynamic of Cognition (DoC) methodology and the POLICE model (Philosophically, Ontologically, Logically, Intersect, Compatibility, Education) (Buch 2025b, p. 1; Buch & Peach 2024). To prevent generative AI and "predictive autocomplete" from liquidating human judgment, citizens must be trained to act as active **authors** of meaning rather than passive "puny human nodes" in automated networks (Buch 2025e, p. 12; Buch 2026a, p. 2; Buch 2026b, p. 6). The POLICE framework provides an explicit metacognitive software for the brain's prefrontal cortex, allowing individuals to "police their own thoughts," filter out algorithmic "institutional bullshit," and practice panser (thinking as caring) (Frankfurt 2005; Stiegler 2025; Buch & Peach 2024; Buch 2026c, p. 4).

3. The Contributory Economy and Cognitive Scaffolding

Finally, Buch proposes replacing neoliberal retail governance with a Contributory Economy anchored in professional contractualism (Buch 2025e, p. 10; Buch 2026b, p. 12). Instead of

allowing corporate platforms to monetize human attention through addictive engagement loops, society reallocates existing capital surplus away from administrative bloat and PR branding to employ standard-bearing historians, philosophers, and sociologists under formal public contracts (Buch 2025e, p. 10; Buch 2026b, p. 12; Legitimacy and Cognition 2026, p. 200). These contracted scholars provide the non-market "cognitive scaffolding" required to preserve deep reading, interdisciplinary research, and historical truthfulness (Buch 2024c, pp. 4–8; Buch 2026a, p. 2).

IV. Benefits of a Well-Managed Transition in the AI Revolution

Arguing for a well-managed transition is not an exercise in anti-technological Luddism; it is a rational defence of human civilizational continuity. A managed transition yields profound structural benefits across multiple societal domains:

Dimension	Pathological Unmanaged Transition	Well-Managed Transition (Jameel & Buch Synthesis)
Existential Safety	"Gambling with human lives" in a race for unaligned superintelligence (Jameel 2026a, p. 194).	Rigorous, independent safety evaluations and ethical verification before public deployment (Jameel 2026a, p. 195).
Cognitive Agency	"Denoetization," "symbolic misery," and loss of deep reading capacity (Buch 2026a, p. 1–2).	Preservation of **Human Authorship**, active judgment, and metacognitive "crap detecting" (Buch 2025e, p. 12; Buch 2026c, p. 18).
Institutional Legitimacy	"Lifestyle State" PR branding masking functional service collapse (Jameel 2026b, p. 1).	**Communicative Action** aimed at mutual intersubjective understanding and genuine welfare (Habermas 1989; Jameel 2026b, p. 1).
Generational Vocation	Reducing Gen Z/Alpha to passive data sets and consumers (Collins 2019; Buch 2026a, p. 3).	Empowering Gen Z/Alpha as the **Architect Generation** guided by moral architecture (Jameel 2026a, p. 195; Jameel 2026b, p. 51).
Ethical End	Cold, calculative technical rationality (*techné*) leading to a "technical abyss" (Buch 2026b, p. 4).	**Aristotelian Flourishing (Eudaimonia)** and thinking as caring (*panser*) (Stiegler 2025; Buch 2025a, p. 3; Buch 2026b, p. 1).

1. **Mitigating Existential and Systemic Risk:** A managed transition eliminates the catastrophic danger of deploying unverified, self-improving AI models that outpace human comprehension and control (Jameel 2026a, p. 194).

2. **Protecting the Human Mind from Denoetization:** By establishing structural boundaries on automated text generation and algorithmic recommendation loops, a managed transition protects the human brain's ventromedial prefrontal cortex (vmPFC) false

tagging mechanism, preventing a widespread "doubt deficit" and "symbolic misery" (Asp et al. 2012, p. 113; Stiegler 2014; Buch 2026a, p. 2; Lincoln 2026, p. 300).

3. **Restoring Truth to the Public Square:** Grounding public communication in Habermasian **communicative action** rather than marketized "Lifestyle State" branding ensures that democratic deliberation is guided by evidence, historical context, and substantial fairness rather than viral, drive-based reflexes (Habermas 1989; Jameel 2026b, p. 1; Buch 2026c, p. 6).
4. **Fostering Aristotelian Flourishing (Eudaimonia):** Ultimately, a well-managed AI transition redirects technological efficiency toward serving human values, enabling individuals and communities to achieve genuine self-actualization, creative expression, and comprehensive empathy (Buch 2025a, p. 3; Buch 2026b, p. 1; Buch 2026c, p. 1).

V. Why Both Arif Jameel and Neville Buch Offer Original and Effective Solutions

The theoretical alliance between Arif Jameel and Dr. Neville Buch represents a unique and formidable intervention in contemporary political philosophy and AI governance (Jameel 2026a, p. 196). Their frameworks are complementary, offering original and effective solutions that bridge the gap between global macro-governance and individual cognitive mechanics:

Image attached

[THE MULTI-SCALAR SYNTHESIS]

GLOBAL / MACRO LEVEL (Arif Jameel)

- The Diella Doctrine & Architect Generation Theory

- The Lifestyle State Diagnosis

- Ethical Passport Theory (EPT)

|

▼ [INTERSECT & COMPATIBILITY]

▲ (Shared Defense of Human Agency)

|

COGNITIVE / MESO LEVEL (Dr. Neville Buch)

- Dynamic of Cognition (DoC) & POLICE Model

- MICE & Contributory Economy Framework

- Professional Contractualism & Panser (Thinking as Caring)

1. The Originality and Efficacy of Arif Jameel

- **Generational and Moral Architecture:** Jameel's *Architect Generation Theory* breaks away from standard Western technocratic policy by recognizing that technological safety is fundamentally a problem of **moral formation** (Jameel 2026a, p. 195; Jameel 2026b, p. 51). By synthesizing ancient Greek, Islamic, and European philosophical traditions, Jameel provides Gen Z and Alpha with a proactive, affirmative role as "architects" of civilization rather than passive victims of automation (Jameel 2026a, p. 195; Jameel 2026b, p. 51, 64).

- **The Lifestyle State and Ethical Passports:** Jameel's diagnosis of the *Lifestyle State* provides an unprecedented sociological explanation for why modern institutions adopt AI tools—not merely for operational efficiency, but for **aesthetic branding and reputational legitimacy** (Jameel 2026b, p. 1, 55). His *Ethical Passport Theory* transforms this insight into an actionable governance mechanism, requiring institutions and AI labs to earn their technological license through demonstrated moral integrity (Jameel 2026a, p. 195; Jameel 2026b, p. 57).

2. The Originality and Efficacy of Dr. Neville Buch

- **Cognitional Methodology (DoC) and POLICE:** Dr. Buch's *Dynamic of Cognition* model provides an explicit procedural methodology for charting how unreflective prejudices transition into revisable knowledge (Buch 2025b, p. 1). The *POLICE framework* acts as a practical, self-regulatory tool that allows individuals to "police their own thoughts," restoring intellectual immune systems against algorithmic "potted thinking" and corporate manipulation (Buch & Peach 2024; Buch 2026a, p. 4; Buch 2026c, p. 18).

- **Political Economy of Knowledge (Professional Contractualism):** Buch's proposal for a *Contributory Economy* directly tackles the material preconditions of truth-conducive inquiry (Buch 2025e, p. 10; Buch 2026b, p. 12). By proving that "voluntarism does not work" in the 21st century, Buch demonstrates that society must contract humanistic scholars as public infrastructure, insulating critical thought from market denoetization

and securing the long circuits of human attention (Buch 2024c, pp. 4–8; Buch 2025e, p. 10; Buch 2026a, p. 2).

3. The Synergistic Power of the Synthesis

Together, Jameel and Buch offer a complete, multi-scalar shield against AI existential risk (Jameel 2026a, p. 196; Legitimacy and Cognition 2026, p. 211). Jameel provides the global governance architecture, generational vision, and civilizational ethics; Buch supplies the cognitional rigor, metacognitive tools, historiographical depth, and political-economic funding model.

By unifying Jameel's *Architect Generation* with Buch's *Dynamic of Cognition*, they demonstrate that human beings are "context-embedded but not context-determined" (Hayot 2021, cited in Buch 2025f, p. 4; Buch 2026c, p. 27). They offer an actionable pathway to ensure that the "Who" (the concrete human person) remains the proactive master of the "What" (the technical and automated systems) of our shared future (Stiegler 1998; Buch 2026b, p. 6; Jameel 2026a, p. 195).

VI. Conclusion: Reclaiming the Future

The whistleblower warnings of late 2026 have made one reality indisputable: the unconstrained race for self-improving AI poses a clear and present threat to human civilizational survival (Jameel 2026a, p. 194). Attempting to contain this threat through superficial, code-level guardrails while maintaining an unregulated market economy is an exercise in self-deception (Jameel 2026a, p. 195; Luchel 2026c, p. 155).

The joint work of Arif Jameel and Dr. Neville Buch outlines the true alternative. By implementing an institutionally enforced AI slowdown, conditioning compute access on *Ethical Passports*, empowering the *Architect Generation*, practicing the *POLICE framework*, and funding independent scholars through *professional contractualism*, society can successfully navigate the AI transition (Jameel 2026a, p. 195; Buch 2025e, p. 10; Buch & Peach 2024).

Through this synthesis of moral architecture and cognitive scaffolding, we reject the passive numbness of digital denoetization and market servility. We reclaim active human authorship over our socio-technical destiny, building a mature, democratic civilization dedicated to truthfulness, comprehensive empathy, and Aristotelian flourishing (eudaimonia) (Buch 2025a, p. 3; Buch 2026b, p. 1; Buch 2026c, p. 1; Jameel 2026a, p. 196).

References / Bibliography

Asp, Erik, Manzel, Kenneth, Koestner, Bryan, Cole, Catherine A., Denburg, Natalie L., and Tranel, Daniel. (2012). 'A neuropsychological test of belief and doubt: damage to ventromedial prefrontal cortex increases credulity for misleading advertising', *Frontiers in Neuroscience*, vol. 6, art. 100, pp. 112–125.

Buch, Neville. (2023a). *SoFiA Melbourne: Unity and Fragmentation*, Dr Neville Buch, Brisbane.

Buch, Neville. (2023b). *Spiral Historiography: Socio-Historical Change in Unitarian Universalism*, ANZUUA, Brisbane.

Buch, Neville. (2024c). *When Humanists Know Nothing of Humanism: Contrast to the Comprehensive History*, Humanists Brisbane, Brisbane.

Buch, Neville. (2025a). *Comprehensive Understanding of Humanism Today*, Dr Neville Buch, Brisbane.

Buch, Neville. (2025b). *Concept-Practice X.6. Dynamic of Cognition*, Dr Neville Buch, Brisbane.

Buch, Neville. (2025c). *Humanism, Free Thought, and Our Conflicted Times*, Dr Neville Buch Blog, Brisbane.

Buch, Neville. (2025d). *Historiography and Humanism*, Dr Neville Buch Blog, Brisbane.

Buch, Neville. (2025e). *Comprehension: Solutions and The Plan*, Management Institute for Contributory Economy, Brisbane.

Buch, Neville. (2025f). *Comprehensive Understanding of Humanism Today (Eric Hayot Review)*, Dr Neville Buch, Brisbane.

Buch, Neville. (2026a). *The Crisis of Comprehension and the Restoration of Deep Reading: Part 2. Response to Critics*, Management Institute for Contributory Economy, Brisbane.

Buch, Neville. (2026b). *The Unitarian Universalist Cognitive Ascent*, Dr Neville Buch, Brisbane.

Buch, Neville. (2026c). *Identity Politics and the Ascent toward Comprehensive Empathy*, Dr Neville Buch, Brisbane.

Buch, Neville and Peach, Neil. (2024). *Philosophically Ontology Logic Intersect Compatibility Education (POLICE)*, Dr Neville Buch ABN: 86703686642, Brisbane.

Collins, Randall. (2019). *The Credential Society: An Historical Sociology of Education and Stratification*, Columbia University Press, New York.

Frankfurt, Harry G. (2005). *On Bullshit*, Princeton University Press, Princeton.

Habermas, Jürgen. (1989). *The Structural Transformation of the Public Sphere*, translated by T. Burger, MIT Press, Cambridge.

Jameel, Arif. (2026a). 'Addressing AI Existential Risks, "Architect Generation," and Moral Training', *Gmail Correspondence to Dr. Neville Buch*, 14 September 2026, Lahore, Pakistan, pp. 194–198.

Jameel, Arif. (2026b). *The Lifestyle State: A Theoretical Framework on the Succession of Legitimacy from Welfare Provision to Lifestyle Branding*, Global Governance Research & Publications, Lahore, pp. 51–65.

Legitimacy and Cognition in the Lifestyle State. (2026). Academic Reference Document, Brisbane, pp. 199–214.

Lincoln, Alan. (2026). 'Uncertainty reduction as a candidate primary reinforcer: an evolutionary and neural account', *Academia Neuroscience and Brain Research*, vol. 2, pp. 243–348.

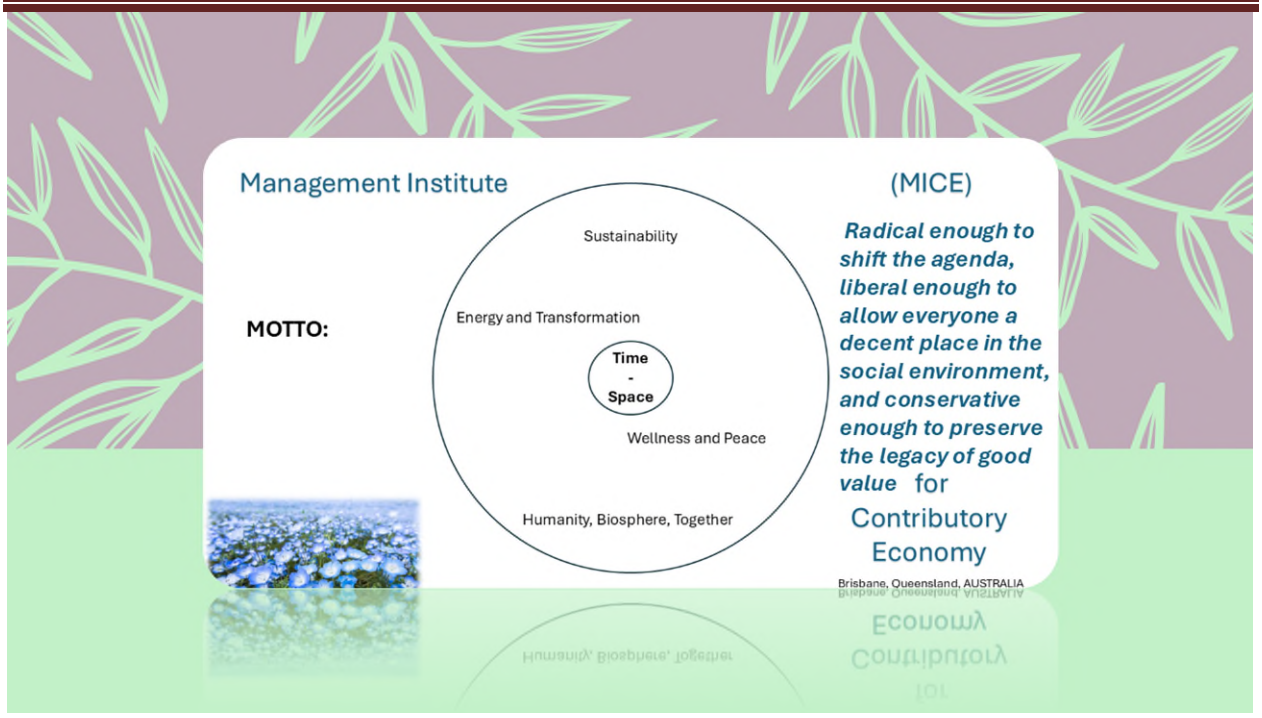
Luchel, Bevin. (2026b). *The Illusion of the Non Market Panacea: A Formal Refutation of Dr. Neville Buch's Reply*, Management Institute for Contributory Economy, Brisbane, pp. 220–228.

Luchel, Bevin. (2026c). *The Mirage of the Subsidized Mind: A Critique of Professional Contractualism, Rent-Seeking, and Noetic Protectionism*, Academia.edu, Brisbane, pp. 154–159.

Stiegler, Bernard. (1998). *Technics and Time, 1: The Fault of Epimetheus*, translated by R. Beardsworth and G. Collins, Stanford University Press, Stanford.

Stiegler, Bernard. (2014). *Symbolic Misery, 1: The Hyperindustrial Epoch*, translated by D. Ross, Polity, Cambridge.

Stiegler, Bernard. (2025). *The Immense Regression: What is Called Caring? (Vol. 1)*, K. Verlag, Berlin.



Take the [MICE Quiz]