



Making (and de-making) technological risk in banal spaces and speculative language

Suneel Jethani, Faculty of Arts and Social Sciences, University of Technology Sydney, AU, suneel.jethani@uts.edu.au

In this paper I provide a framework for enhancing seemingly objective approaches to AI risk assessment in a way that doesn't just see risk as only scaling vertically, levelling up through thresholds and tolerances(intensities). It draws on Raqs Media Collective's *An Infra-vocabulary for Capital* (2023–2024) and considers how the work can be adapted to imagine horizontal formations of risk (accumulations) by using language and naming resulting in different apertures and resolutions on the relations that hold that risk together as a latent force in operational space.



When you invent the ship, you also invent the shipwreck; when you invent the plane you also invent the plane crash; when you invent electricity, you invent electrocution. Every technology carries its own negativity which is invented at the same time as technical progress.

Paul Virilio

I risk *therefore* I am. I venture therefore I am. I suffer therefore I am [...] The category of risk opens up a world within and beyond clear distinctions between knowledge and non-knowing, truth and falsehood, good and evil [...] it amalgamates knowledge with non-knowing within the semantic horizon of probability.

Ulrich Beck (2009: 5)

The words ‘narcotic clockwork’ stand out in Raqs Media Collective’s *An Infra-vocabulary for Capital* (2023–2024). The work, which has been installed in various configurations and locations over the world draws on the Vishnu Sahasranama, a Hindu text listing the thousand names of Vishnu the devine protector. When I see the work in July of 2024 at 4A Centre for Contemporary Asian Art in Sydney, the black vinyl text on white walls name capitalism’s failing stewardship of systems that organise our access to space, the realm of tools and technologies and others in the form of social relations and power dynamics (**Figure 1**).



Figure 1: Raqs Media Collective – *An Infra Vocabulary for Capital*. Vinyl on Wall. 4A Centre for Contemporary Asian Art, Sydney Australia. July 2024. Image credit: Suneel Jethani.

Free from syntactical organisation, the artists present an open eulogy revealing the nuance of broken systems underpinned by finance, science, technology and industry. Once named, hierarchies, taxonomies, cladistical links, affinity clusters and deep reflection follow. I scan the work, and the following words stand out: *Abeyant* – something that is inactive but capable of being active; *Theorist* – a person concerned with the theoretical aspects of a subject; *Beholden*—having duty to someone or something; *Archon*—ruler; *Panjandrum*—*a person who has or claims to have authority or influence*. My next impulse is to form word pairs: *theorist-stalker*, *potent-core*, *focused-crash*, *transactor-incinerator*, *anthrax-machine*, *innumerable-chaos*. This infra-vocabulary is a generative discourse on risk and has much to offer contemporary debates on Artificial Intelligence. It invites ‘a commitment to a certain way of defining a problem space, rendering problems into thought in a particular way, and establishing and constraining the kinds of explanations’ (Rose, 2023) as a legitimate mode of critical inquiry.

In *Autonomous Technologies: Technics-out of-Control as a Theme in Political Thought* (1977) political theorist, Langdon Winner noted that ‘technology is a source of concern because it changes in itself and because its development brings other kinds of changes in its wake’. Around the same time that Winner was writing about autonomous technologies being ‘engines of change’ in the politics of space, time, embodiment and epistemology (1977: 44–100), computer scientist, Herbert Simon won the 1978 Nobel Prize in Economics for a theory of bounded rationality termed ‘satisficing’ (McCorduck, 2004: xxix). The term, a central principle in the field of Artificial Intelligence (AI), is a portmanteau of what is *satisfactory* and what will *suffice* – a decision making strategy that involves searching through the available alternatives until an acceptability threshold is met. The term, which brings together in language notions of something being *good enough* which tactically functions to de-make its negativity. This notion of a quantifiable acceptability threshold frames much of the discourse on the risks that data-intensive, artificially intelligent systems bring into everyday encounters with automated systems.

Since its emergence as a field in the 1950s AI has been carried by an inflated language of ‘unfulfilled grandiose promises’ (Gebu & Torres, 2024) where deployments of AI systems occur in banal operating spaces where risk is understood, managed negligible and pushed to spatial, temporal zones which are too distant to be of immediate concern. In *Normal Accident Theory* (1984) Charles Perrow called the technological realms of weapons, space exploration and recombinant DNA ‘exotics’. As Franco ‘Bifo’ Berardi notes in *The Uprising: On Poetry and Finance*, such symbolist experiments with language in the early 20th Century have found their deepest expression on circuits of finance and

the quotidian nature of capital injection into these once ‘exotic’ realms by hedge funds, aspiring millionaires and everyday speculators alike (Berardi, 2012).

German sociologist, Ulrich Beck opens *World at Risk* with the assertion that ‘the anticipation of catastrophe is changing the world (2009: 1). A relation of risk, trust, and security underwrites human-technical, ethno-epistemic and political-economic life with an increasing intensity, especially as many parts of the planet pass through a transition where the organising principle for risk management were documented industrial processes with tight coupling *in situ* (Perrow, 1984) to one of loosely bound, opaque operations *in silico*.

In May of 2023, the Center for AI Safety (CAIS [pronounced ‘case’]), a San Francisco based research and field-building non-profit organisation which, seemingly, advocates for a reduction of AI-attributable societal risk published a statement declaring that ‘mitigating the risk of [human and planetary] extinction from AI should be a global priority alongside other societal scale risks such as pandemics and nuclear war’.

The Institute of Risk Management (IRM) defines risk as ‘the combination of the probability of an event and its [positive or negative] consequence[s]’ (Hopkin, 2018) in terms of magnitude, size, likelihood and scope. Examples of inflated AI risks and the typical language that it is expressed in include: (1) weaponisation by malicious actors, (2) the facilitation of misinformation spreading through on and offline communication networks; (3) proxy gaming—where AI trained with questionable objectives could find new and unpredictable ways to pursue goals that are at odds with social expectations and human values; (4) enfeeblement—when important tasks are relegated to automated systems and humans lose, to some extent, the ability to have agency and the ability to self-govern; (5) value lock in—where centrally controlled systems give small groups of people a tremendous amount of power (-to, -over, -with) leading to a lock-in of oppressive technocratic regimes; (6) emergent goals—AI models could demonstrate unexpected, qualitatively different behaviour increasing the risk that humans could lose control over and (7) deception—when deception may become part of learned behaviour of a system to achieve its goals or outperform other ‘honest’ systems (Hendrycks & Mazeika, 2022).

But as Beck notes:

From [these kinds] of threat[s], we must distinguish the *semantics of risk* associated since the beginning of the modern period with the increasing importance of decision, uncertainty and probability in the process of modernisation. The semantics of risk refer to the present thematization of future threats that are often a product of the success of civilisation. It also makes possible new, post-utopian mobilisations

of societies [...] as we have seen [...] in] shifting alliances between civic movements, states and companies. (2009: 4)

Yet the statement put forward by CAIS, informed by the inflated risk profile listed above, has been signed by high profile tech industry figures representing Open AI, Microsoft, Google along with other notable figures concerned about ‘severe catastrophic and existential risks’. Similarly, an open letter from the Future of Life Institute (FLI) dated March 2023 warned of ‘ever more powerful digital minds that no one, not even their creators, can understand, predict, or reliably control’. To be sure, AI technologies pose unpredictable and long-term risks but a focus on probabilities and impacts within elongated timescapes (Adam, 2005) deflects from the problem of thinking about long and short term, distal and proximal risk simultaneously as *modalities*. Terms like existential and catastrophic have a transcendental effect on risk discourse that shifts the loci of affects vertically from loci of action.

Calls for precautionary principles to mitigate existential risk (Hendrycks & Mazeika, 2022) have been (Geburu et al., 2024; Bianchi et al, 2023) and should be met with scepticism. This is because: framing risk in this way infers inevitability where in reality existential and catastrophic risk caused by the action of artificially intelligent systems is speculative and uncertain; they divert the energies of those contributing to public discourse and action on AI risk away from real short-term risks and harms that are already occurring; such statements are a form of strategic advocacy to avoid regulation, conduct business as usual and slowly erode existing guardrails and expectations around safety, harm and responsible innovation and signatories are ‘just fuelling counterproductive AI hype’ (Sætra & Danaher, 2023)

New vocabularies for AI risk assessment could allow AI regulation, ethics and design communities to frame risk in ways that don’t see risk as only scaling vertically, levelling up through thresholds and tolerances (intensities). Rather, it accommodates horizontal formations of risk (accumulations) – like lenses in front of each other – resulting in different apertures and resolutions on the relations that hold that risk together as a latent force in ordinary operational spaces. Objectivist representations of space and the authority claims, knowledge systems and practices of technology administration and governance developed from them assume that risks are constituted prior to their confrontation of human subjects (Kinsella, 2010: 268). These are, supposedly, fully describable in ways that allow quantification and the allocations of probabilities, thresholds, tolerances and limits.

For Henri Lefebvre (1991) there are three forms, or moments, of dialectic spatiality, each socially produced and culturally embedded. The first is real space (*espace preçu*),

the product of nature and social labour which transforms land and erects buildings from raw materials. The second is imagined space (*espace conçu*) including the designed and documented spaces of architects, engineers and planners. Finally, there is lived space (*espace vécu*) where the [banal] activities of everyday life take place, including ‘working safely to conserve energy – often steeped in custom and infused with take[n]-for-granted symbolism’ (Bellaby, 1999: 1322). Real and imagined spatialities are imbricated, in the sense that they are different shapes but overlap or interlock to form one surface – lived space (Bellaby, 1999: 1322). Banal operating spaces are produced in chains of “imbrication” over time, and in the resistances and workarounds that emerge:

not simply a linear development and events are not containers with static boundaries, rather they play pivotal roles in both identifying and shaping the discourse–materiality relationship. An emphasis on time also suggests that a process perspective is fundamental to treating the two as situated, dynamically interconnected, and emergent, but not necessarily fused. (Putnam, 2015: 713)

The naming of risk in this way offers a historically specific mode of arraying material forces invested by capital into being, as well as elaborated through the languages of biology, physics, thermodynamics, complexity theory and non-linear rationality (Clough, 2008: 2) that reconfigures bodies, labour, (re)production in the profiling and management of AI risk at the level of banal operational space. We might want to think of the notion of *management* in risk management as Raymond Williams did in *Keywords* which highlights a dual function of “management” (1983/2014: 191). Williams describes management as a bureaucracy, and in the context of this chapter’s argument we could say that management includes technocracy as the work occurring when a select group of elites administer processes of human control in ways that aim to support pre-defined norms that serve managerialism and give it technically mediated precision. The other function of management that Williams refers to is the abstraction of relations that are embedded in processes that automate the internalisation of external forms of power and control in ways that would seem as if they were good for and in the best interests of those being subjected to regimes of management.

Our current moment of technological solutionism sees the semantics of risk as especially important in the languages of technology, economics, politics, design and art. In the linked fields of embodied, data-intensive, sensor-enabled and artificially intelligent technology where the speed of development is rapid and the portability of technology between contexts can proceed without much hinderance, cultural imaginations of risk hinge on dramatized, idiosyncratic or unexpected accounts of technological performance.

Most fears around this class of technologies are directed to opaque and poorly understood processes, timescales that are too far into the future when considered as a function of existing technical capacities and spaces that are far removed from the rhythms and routines of everyday life. Risk is ‘thus a “mediating issue” in terms of which the division of labour between science, politics and the economy in highly innovative societies must be negotiated’ (Beck, 2009: 6). As the embodiment of automated, data-intensive logics becomes increasingly prevalent as a component of one’s lived experience of technology it is not only empirical or ethnographic inquiry that will bring forward the lexicons and frameworks that allow for risk to be understood in more immediate, proximal and phenomenological terms. Anticipatory and speculative engagements with technological risk that are not framed by science fiction tropes inform the critical study of embodied, automated, data-intensive technology in artistic practice has the potential to critically engage with the nuanced aspects and multifaceted implications of living with these systems. Artistic engagements at the level of languages and vocabularies such as Mindy Seu’s *Cyberfeminism Index* (2022), Rosi Braidotti’s *Posthuman Glossary* (2018), and Timothy Neal, Courtney Addison and Thao Phan’s *An Anthropogenic Table of Elements* (2022) all link to ways naming can be used to develop understandings of representations of the self and identity, to issues of power, control and to questions of value and agency to interrogate the kinds of ontologies, relations and communities that are emerging out of the hybrid interweaving of body and technology in the context of datafication and automated decision making and their emerging risk profiles.

Competing interests

The author has no competing interests to declare.

Author info

Suneel Jethani is a Senior Lecturer in the Faculty of Arts and Social Sciences at the University of Technology Sydney. He is the author of *The Politics and Possibilities of Self-Tracking Technology: Data Bodies and Design* (2021, Emerald) and has published in journals including *Continuum*, *Cultural Studies*, *Communication, Politics and Culture*, *M/C Journal*, *International Communication Gazette*, *Persona Studies* and *Conjunctions: Transdisciplinary Journal of Cultural Participation*.

References

Beck, Ulrich. *World at risk*. Polity, 2009.

Bellaby, Paul. "Spatiality, embodiment and hazards encountered in the making of pots." *Social Science & Medicine* 48, no. 10 (1999): 1321–1332.

Berardi, Franco Bifo. *The uprising: on poetry and finance*. MIT Press Books (2012).

Clough, Patricia T. "The affective turn: Political economy, biomedicine and bodies." *Theory, culture & society* 25, no. 1 (2008): 1–22.

Gebru, Timnit, and Émile P. Torres. 2024. "The TESCREAL Bundle: Eugenics and the Promise of Utopia through Artificial General Intelligence". *First Monday* 29 (4). <https://doi.org/10.5210/fm.v29i4.13636>.

Hendrycks, D., & Mazeika, M. (2022). X-risk analysis for AI research. arXiv preprint arXiv:2206.05862.

Hopkin, Paul. *Fundamentals of risk management: understanding, evaluating and implementing effective risk management*. Kogan Page Publishers, 2018.

Kinsella, William J. (2010) Risk communication, phenomenology, and the limits of representation. *Catalan journal of communication & cultural studies* 2, no. 2: 267–276.

Lefebvre, Henri. (1991). *The Production of Space*. Massachusetts: Blackwell.

McCorduck, Pamela. *Machines who think: A personal inquiry into the history and prospects of artificial intelligence*. AK Peters/CRC Press, 2004.

Perrow, Charles. (1984). *Normal Accidents: Living with High-Risk Technologies*. Basic Books.

Putnam, Linda L. "Unpacking the dialectic: Alternative views on the discourse–materiality relationship." *Journal of management studies* 52, no. 5 (2015): 706–716.

Sætra, Henrik Skaug, and John Danaher. "Resolving the battle of short-vs. long-term AI risks." *AI and Ethics* (2023): <https://doi.org/10.1007/s43681-023-00336-y>

Williams, Raymond. *Keywords: A vocabulary of culture and society*. Oxford university press, 2014.

