



# Consilience and AI as technological prostheses

Jeffrey B. White<sup>1</sup>

Accepted: 5 August 2024

© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2024

Edward Wilson wrote in *Consilience* that “Human history can be viewed through the lens of ecology as the accumulation of environmental prostheses” (1999 p 316), with technologies mediating our collective habitation of the Earth and its complex, interdependent ecosystems. Wilson emphasized the defining characteristic of complex systems, that they undergo transformations which are irreversible. His view is now standard, and his central point bears repeated emphasis, today: natural systems can be broken, species—including us—can disappear, ecosystems can fail, and technological prostheses potentiate rather than ameliorate disaster when the fragilities that they introduce into our evolutionary co-dependencies with natural systems, and with each other, are under-appreciated.

Given the lessons of our technological inheritance from the present endpoint of human history, what might Wilson recommend be done with AI, virtual reality and the charade of computer-thought? He placed his greatest hopes in the human capacity to form long-term principled associations for cooperative action prior to any market economy and over-ruling its billionaire champions. Contrary to human “exemptionalism”, consilience involves an integrative and “holistic” approach to problems that cut across disciplines and affect us all, to mitigate expanding exposures to risks due to increasing dependencies on increasingly complicated technologies, in order to provide not only practical solutions to inherited problems, but also to guide forward progress for our collective betterment. Wilson recommended that we adopt such an integrative approach to technological development. This journal has championed a consistent strategy sensitive to the diverse needs of the international community in the context of AI since 1986, and continues, stalwart in this mission, today.

Yong-In Park (this issue) resonates deeply with this forward view. Drawing from established successes in collective

agreements to address global environmental concerns, Park analogizes our situation in a way that invites comparison with Wilson’s *Consilience*. Park as well as Wilson stress that different disciplines recognize the same apparent problems in different ways, focus on different dimensions as salient and recommend different solutions to address these. Both emphasize that problems such as with migrating pollution and exploitation of shared natural systems are unmanageable without a collective overview. Consistent with *Consilience*, Park leverages recent world-wide environmental initiatives to emphasize that any way forward begins with a coming together of specialists into a holistic manifold, and in this way frames ongoing efforts with AI. Differently from Park, Wilson might stress that consensus is potentiated as specialists insulated in disciplinary silos can understand complex problems in common terms, and that striving for consensus introduces a tendency to reduce such problems to digestible dimensions, such as with an exclusive focus on carbon dioxide. Wilson might argue that defenestrating infelicities of such a reductive approach include that proposed remediations cause other problems in other dimensions that make things worse, requiring further actions with similar side-effects, and that industries that might be throttled directly with more accurate accounting are spared correction, such as with e-waste from the West dumped in Africa, and heat pollution of regional waterways by massive server farms which are more pressing concerns in the context of AI, independent of carbon dioxide, and with more specific solutions.

Consilience involves not only bridging disciplines, but also cultures. Consensus in addressing such complex problems as impacts of AI must be potentiated without defenestrating over-simplification, for example in the uncritical imposition of big-tech monoculture on different social and ecological contexts. *AI & Society* is exemplary in this capacity, as a forum that brings together diverse interests in communication of representative values. Other contributions to this issue illustrate the diversity of impacts of technologies on different cultures and their natural environments. For example, Biju and Gayathri consider the Indian approach to AI through an analysis of regional policy discussions in the

---

✉ Jeffrey B. White  
jeffreywhitephd@gmail.com

<sup>1</sup> Onna, Okinawa, Japan

context of the Indian culture and constitution. They expose Indian policy discussions relying too heavily on examples from other regions that do not fit Indian values. Indian “societal needs, aspirations, and ethical concerns” should shape Indian AI policy, rather than apply one-size-fits-all economic incentives to motivate development. From Sweden, Wennerström and Foka consider some obstacles to culturally sensitive AI development. For example, in language processing, there are no resources available to develop applications for smaller linguistic groups, again, for economic reasons, with developers left to start from scratch. Ratana, Sharifzadeh, and Krishnan consider natural language processing tools for mental health care in the Māori population of New Zealand. They reveal obstacles both technological and institutional, as the Maori are a small and culturally specific sub-group covered under a larger, national system, exposing risks of misdiagnoses and mistreatment due to culturally insensitive applications of ubiquitous AI tools.

In “Trust, understanding, and machine translation: the task of translation and the responsibility of the translator”, Melvin Chen critically compares different methods of machine translation, and the responsibility of the human being in ensuring translation accuracy. Chen picks up on Bar-Hillel’s contention that fully automated machine translation is impossible, and advances this discussion through direct analysis of existing methods. Ultimately, the argument is one of trust, leaving open ideas pursued by other contributions to this issue involving human-centeredness and empathy as mediated by AI and associated technologies. Murphy, Carew and Stapleton review AI and virtual reality powered initiatives affording immersive, personalized exposure in the context of cultural heritage preservation, to social conditions potentially inaccessible otherwise. They find that education encouraging respect and empathy are main driving factors behind such programs. In addition, on the issue of perceived empathy in machines, Concannon and Tomalin note that different researchers consider empathy in human–computer interactions in different ways. They propose a novel assessment framework for perceived empathy in human interactions with dialogic AI systems based on a review of how empathy is assessed between human beings. Browning considers personhood from autonomy and moral responsibility, concluding that contemporary generative AI cannot be considered persons by these criteria, in this way arguing that current models cannot understand what it is to be human. Similarly, Harding, D’Alessandro, Laskowski and Long argue that machines cannot stand in for human subjects in scientific research. They note that AI models represent values expressed in specific contexts, that these values change quickly, and that automated judgements from prior value orientations cannot predict future directions. Each of these contributions treat in different ways the mediation of AI in

social interactions with foci on technological limitations and the irreplaceable human interests central to system design and implementation characteristic of this journal, and reflective of Wilson’s view, as well.

Taylor, O’Dell and Murphy treat human–AI value alignment in a way deeply resonant with Wilson’s condemnation of the “exceptionalist” strategy. They show that AI systems are already human-centered, and aligned with human values, but to the wrong ones. They consider Ubuntu and Maximal Feasible Participation as two alternative strategies for decentralization of AI development and dislocation from perverse top-down economic incentives commensurate with the exceptionalist worldview yet strictly contrary to broad human interests. They argue that the common understanding at work in system design and implementation separates “human” interests from broader—holistic—concerns. Human-centeredness is reconsidered as community-centeredness in comparison, and Taylor and colleagues close convincingly that issues of AI value alignment must begin with the values, themselves, bottom-up rather than attempt to reform what is already misconceived. One theme binding this with other contributions to this issue involves the pervasive encroachment of profit incentives into public interests and representative institutions in the context of AI. Specifically pressing this point, Bartłomiej Chomanski in “Pauses, parrots, and poor arguments” starkly exposes our current situation in terms closely aligned with those of Taylor and colleagues, and consistent with Wilson’s critique. Chomanski lays it bare, that we are foolish to be suspicious of market-driven exploitation while feeling confident that public institutions administered by people pursuing the same (exceptionalist) strategies will counter private interests and “achieve goals of safety and equity”.

As Wilson ended *Consilience*, he emphasized the promise of human potential to self-organize around mutually beneficial moral precepts into cooperative, self-regulating systems of mutual concern that carry humanity forward over evolutionary time. “In the course of it all” he wrote, “we are learning the fundamental principle that ethics is everything.” (1999 p 325) As we learn more, and understand better, we can choose more wisely how to proceed. However, his closing words came in the form of a warning, as a reminder of the crux of choice at which we each find ourselves at such moments such as now, that we can go the wrong way:

“And if we should surrender our genetic nature to machine-aided ratiocination, and our ethics and art and our very meaning to a habit of careless discursion in the name of progress, imagining ourselves godlike and absolved from our ancient heritage, we will become nothing.” (1999 p 326).

## Reference

Wilson EO (1999) Consilience: the unity of knowledge. Vintage, New York

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.