Opinion

Dignity, properly used, could be a useful construct in AI ethics

Cait Lamberton,[1,*] Lorenn Ruster,[2] Sakshi Ghai,[3] and Neela Saldanha[4]

[1]The Wharton School, University of Pennsylvania, Philadelphia, PA, USA

[2]School of Cybernetics, ANU College of Systems and Society, Level 3 Birch Building, 35 Science Road, The Australian National University, Canberra, ACT 2600, Australia

[3]Department of Psychological and Behavioural Science, London School of Economics and Political Science, Connaught House, 63-65 Aldwych, London WC2B 4EJ, UK

[4]Yale Research Initiative on Innovation & Scale (Y-RISE), Whitney and Betty MacMillan Center for International and Area Studies, Yale University, 34 Hillhouse Avenue, New Haven, CT 06511, USA

*Correspondence: catlam@wharton.upenn.edu

Rueda et al. argue that the concept of dignity is problematic for AI ethics due to its complexity, ambiguity, and biased usage. While agreeing on many points, we propose that adding the necessary precision to the use of the term is neither difficult nor onerous. Further, doing so may allow understanding of the factors that promote dignity affirmation, match the multifaceted nature of AI systems themselves, and promote pragmatically better design outcomes than will be likely if the idea is avoided in AI ethics discussions.

Rueda et al.[1] present a compelling argument that those discussing artificial intelligence ethics should be skeptical of the use of the concept of dignity. Their argument flows along three axes: first, dignity is an overly complex and ambiguous concept; second, sloppy use has rendered it meaningless; and third, biased employment of the term has tainted it with suspicion. While they stop short of eradicating the concept, their essay suggests that abandoning it may be preferable to continued wrangling.

We applaud many of the article's points but argue that this extreme conclusion would be mistaken. Not only is it possible to raise the precision with which the idea of dignity is deployed, there may be real benefit from doing so. Therefore, we propose some steps that may allow the power of the idea of dignity to be better leveraged, to the benefit of both the AI community and the end user.

Precise use of the concept is possible and not onerous

Rueda et al. are correct that imprecise use of the idea of dignity is unhelpful. As they note, its rhetorical power was manifest in the opening words of the UN's 1948 Universal Declaration of Human Rights, which suggested that signatory nations center dignity as the cornerstone of their policies. However, it was also immediately reasonable to question the serious-

ness of this commitment, given the wide range in respect for human rights demonstrated by the signatory nations.

Scholars have since presented much more precise, context-specific definitions of dignity. Beyond the widely accepted distinctions between inherent and merit-based dignity mentioned in Rueda et al., a rich literature exists related to dignity in the workplace[2] and in healthcare,[3] as well as our own work on the marketplace.[4] Further, research has surfaced valid ways to measure the subjective experience of dignity affirmation, not only in the West[5] but also in the global South.[6] Given this literature, it is reasonable to require that one should specify the specific, context-appropriate conceptualization adopted and use validated measures of that concept—as they would in any scholarly discussion. This is not a difficult request to make or fulfill.

We also call for greater precision. Ethicists should clarify whether they focus on designers' objective actions to convey respect (e.g., are we designing with dignity in mind?) or on users' subjective experiences (e.g., do I feel my worth is respected?). These can diverge. A firm may copy features that affirmed dignity in one market into another, believing it creates a "respectful experience." Yet the new audience may see these choices as inattentive to their needs, treating the original market as more valued. Recog-

nizing such disconnects can deepen understanding of a feature's effects and help designers better fulfill their dignity-related aims.

Ignoring the concept of dignity in favor of its formative grounds may be both logically and practically problematic

Rueda et al. suggest that it may be more productive to advance AI ethics by focusing on concrete values often correlated with dignity, such as autonomy, equality, or (de)humanization. While appealing, this substitution is both conceptually flawed and practically misaligned with goals of promoting well-being.

The conceptual problem arises if factors such as autonomy, equality, fairness, and related values are taken as proxies for dignity. These factors are often philosophically considered "grounds" for dignity. However, they are properly conceptualized as formative, rather than reflective, elements. Formative elements of a construct may be uncorrelated with one another, such that none should be taken as representative of others. Further, the construct being "formed" retains a definition separate from any of its formative elements. In this way, dignity's structure parallels that of socioeconomic status (SES), which is similarly grounded in formative factors like income, education, and occupation yet defined independently and consistently as

"the social standing or class of an individual group." Just as income alone cannot stand in for SES, no single contributor can substitute for dignity itself. Additionally, just as the formative factors of SES in the US may differ from those that constitute SES in other regions, the critical factors that form dignity may differ across contexts.

Practically, privileging one formative factor of dignity while ignoring others can also undermine well-being.[4] For example, a system that maximizes autonomy but neglects equity risks devolving into a market where unchecked competition slides into collusion. This is rarely considered a desirable outcome of a new technology. Conversely, pursuing equality without autonomy produces a uniform but unfree, Orwellian world, which is also not—at least explicitly—a desirable objective for AI systems.

This perspective also links design factors to their consequences in ways that can guide decision-making. First, by capturing both a measure of dignity affirmation and, separately, formative factors such as agency, representation, fairness, and humanization, we can see how these elements interact to shape experience across contexts. Second, an overall measure of dignity affirmation allows us to assess its consequences—and how they differ from those of any single formative factor.

Dignity is practically useful

If more precisely specified, the concept of dignity can also be of practical use. Recent practice suggests that dignity-driven design processes or audits[7] can be executed as a concrete activation of the concept. For example, consider efforts to embed children's rights in AI governance. UNICEF's AI for Children 2020 guidance highlights that children must be treated not merely as "passive objects of care" but as rights-holders whose inherent dignity must be affirmed and protected. Translating to practice, a recommendation is clear: safety-by-design practices would create not only stronger safeguards against manipulation or harm but also age-appropriate consent that recognizes children's capacities, allows them voice, and preserves agency—thus affirming dignity in context-appropriate ways.

As another example, Ruster et al.[8] identify different ways of thinking about dignity in algorithmic development processes,

with a clearly specified concept of dignity culminating in a Dignity Lens. The meaningfulness of this framework is evidenced through its ongoing implementation into the everyday operations of a data science team. Concepts of dignity have also informed a preliminary reflective practice prototype for technologists who build AI systems, working in early-stage start-ups.[9] Initial benefits identified from enacting dignity-centered reflective practice included value to the individual cofounder and the organization as well as evidence of integration of the practice into their ongoing routines.

Dignity's complexity is appropriate for AI

Lissack[10] argues that complex systems prompt us to either decrease what we pay attention to or increase the range of our control system to match the complexity of the environment. It appears that Rueda et al.[1] advocate for the former path—decreasing scope of attention—claiming that because dignity is complex, its impact on AI is not "self-evident" and thereby should not be kept in focus.

We advocate for the second path. Because dignity is complex, it has the potential to match the complexity of the environment in which it may be deployed as a "control system" or set of ethical guidelines. We root this argument in Ashby's[11] law of requisite variety, which calls for a match between the complexity of a system and the complexity of responses available to regulate, manage, or control that system. Following this principle, any engineering problem might be explored for its dignity-related implications. However, some technologies are simpler than others and thus do not need the complexity offered by dignity. For example, consider a well-functioning stapler as a technology at the heart of a paper-organizing system: this system does not usually involve fairness and does not imply different degrees of dehumanization or trigger a need to express voice. As such, considering dignity as part of managing this paper-organizing system would likely engage with more complexity than the stapler needs.

By contrast, an AI system is far more complex, involving factors including the technology itself, its users and those with whom they interact, stakeholders participating in the flow of data and capabilities, environmental impacts, and inves-

tors. A single construct, like autonomy, would not capture the range of effects created across the system. Thus, it is necessary to engage concepts that are at least as nuanced as the problems faced—that can be both decomposed into their constituent parts and combined in ways that make their effects on shared human experiences legible and comparable. Dignity, in its complexity and variety, is a good match to this requirement.

Antao et al.[12] provide one example. In their work, dignity guides ethical development precisely because it applies across public health and AI. Variation in dignity affirmation can be tracked across interactions with an AI-enabled health system, AI's uses can be assessed for their impact on patterns of dignity denial, and its capacity to affirm or deny dignity can be examined for implications for inclusion and equity.

Where to go from here?

Thus, while we resonate strongly with the calls to remove sloppy uses of "dignity" from AI ethics, we would like to prompt the AI community to resist giving the construct up altogether. Rather, we propose the following guidelines.

(1) "Dignity" is precisely specified when used. This includes references to specific dignity types as defined in prior literature, as well as a speaker's locus of concern (e.g., the actor's intention to affirm dignity or the target's subjective experience).

(2) The AI community resists tendencies to avoid the discussion or measurement of dignity in favor of factors that may be its formative elements. Rather, both potential formative elements and dignity itself may be discussed, measured, and analyzed for their connections in different contexts and with different goals.

(3) The AI community embraces complex and adaptive systems orientations toward AI, such that the match between the complexity of AI and the complexity of dignity as a concept can be leveraged.

(4) To avoid the use of "dignity" as a conversation-stopper, the AI community acknowledges heterogeneity in the relative value placed on dignity affirmation across contexts

and uses. In some cases, it may rank lower than priorities such as scientific discovery or enhanced efficiency. Recognizing such differences may ground more productive debate, while also highlighting opportunities to develop exogenous safeguards for dignity affirmation in contexts where it is not designers' primary focus.

We close by also affirming Rueda et al.'s point that it is important to examine the ways in which the concept of dignity is employed in service of ideological agendas. Paralleling this, we suggest that it is equally important to reflect on the assumptions or ideologies underlying calls for its dismissal. What we believe the authors' arguments make clear is the need for more discourse about dignity—but a discourse that is more precise and careful. Precisely because of its deep philosophical roots, ongoing conceptual and methodological development, and nearly universal capacity to elicit support, the construct may be worthy of the work necessary for refinement rather than obviation.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the authors used GPT-5 in order to tighten and clarify language to meet word count requirements. This resulted in dropping approximately 50 words. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

REFERENCES

1. Rueda, J., Ausín, T., Coeckelbergh, M., del Valle, J.I., Lara, F., Liedo, B., Albareda, J.L., Mertes, H., Ranisch, R., Raposo, V.L., et al. (2025). Why dignity is a troubling concept for AI ethics. Patterns *6*, 101207. https://doi.org/10.1016/j.patter.2025.101207.

2. Hodson, R. (2001). Dignity at Work (Cambridge University Press). https://doi.org/10.1017/CBO9780511499333.

3. Chochinov, H.M. (2022). Dignity in Care: The Human Side of Medicine, First Edition (Oxford University Press). https://doi.org/10.1093/med/9780199380428.001.0001.

4. Lamberton, C., Wein, T., Morningstar, A., and Ghai, S. (2024). Marketing's role in promoting dignity and human rights: A conceptualization for assessment and future research. J. of the Acad. Mark. Sci. *52*, 1391–1411. https://doi.org/10.1007/s11747-024-01008-x.

5. Pirson, M., Piedmont, R.L., Nagy, N., and Hicks, D. (2023). Establishing a Dignity Scale -Measuring Intrinsic Value within Social Contexts. Humanist. Manag. J. *8*, 97–112. https://doi.org/10.1007/s41463-023-00147-7.

6. Wein, T., Khatry, P., and Bhimani, R. (2022). A five-item survey measure of Felt Respect for Dignity. OSF. https://doi.org/10.17605/OSF.IO/VG2AC.

7. IDinsight Upholding Dignity. https://www.idinsight.org/services/upholding-dignity/.

8. Ruster, L.P., Oliva-Altamirano, P., and Daniell, K.A. (2025). The dignity lens: advancing human-centred protective and proactive algorithmic responsibility. Inf. Syst. J. https://doi.org/10.1111/isj.12601.

9. Ruster, L.P. (2023). Dignity-centred reflective practice for responsible action in technology startups: a preliminary approach. In Australasian Conference of Information Systems (ACIS) 2023 Proceedings.

10. Lissack, M. (2019). Understanding Is a Design Problem: Cognizing from a Designerly Thinking Perspective. Part 1. She Ji *5*, 231–246. https://doi.org/10.1016/j.sheji.2019.07.002.

11. Ashby, W.R. (1968). Variety, Constraint, and the Law of Requisite Variety. In Systems Research for Behavioral Science, W. Buckley, ed. (Routledge), pp. 129–136.

12. Antao, E.-M., Rasheed, A., Näher, A.-F., and Wieler, L.H. (2025). Reason and responsibility as a path toward ethical AI for (global) public health. npj Digit. Med. *8*, 329. https://doi.org/10.1038/s41746-025-01707-x.

About the authors

Cait Lamberton is Vice Dean of the Undergraduate Division and Alberto I. Duran Presidential Distinguished Professor of Marketing at the University of Pennsylvania's Wharton School. Her research and writing focus on consumer behavior, with concentrations in areas such as retail size and assortment, technology in marketing, the sharing economy, attitudes toward public policy, and the relationship between dignity and human rights. She is cochair of the American Heart Association's Health Care by Food behavioral science task force, past editor of the *Journal of Marketing* and *The Cambridge Handbook of Consumer Psychology*, and an active consultant in the pharmaceutical industry.

Lorenn Ruster is a Responsible AI scholar working at the intersection of academia and industry. She combines over a decade of experience working in professional services specializing in systems thinking, community-led tech development, strategy, human rights, and governance with her current research in responsible AI practices. She is an interdisciplinary scholar, and her PhD focuses on participatory, engaged research with organizations building (or thinking about building and using) AI-enabled products, where she codesigns and embeds organizational processes and practices to enable responsible AI. One prototype developed throughout her PhD is the Dignity Lens for algorithmic development.

Sakshi Ghai is an Assistant Professor of Psychological and Behavioural Science at the London School of Economics and Political Science. She earned her PhD in psychology from the University of Cambridge. Her research spans two strands: applied work on how digital technologies such as social media, gaming, and AI shape young people's well-being and online safety in the Global South and meta-scientific work on diversity in behavioral science. She examines issues of sample representativeness, generalizability, and open science, with a focus on culturally relevant, rigorous approaches that broaden behavioral research to include populations in the Global South.

Neela Saldanha is a behavioral scientist and the executive director of the Yale Research Initiative on Innovation & Scale (Y-RISE). She spent 15+ years in the private sector leading teams at Nestlé, Unilever, Accenture, and PepsiCo. She has been mentioned in *Forbes* magazine's "10 Behavioral Scientists You Should Know." Her book *Marketplace Dignity*, coauthored with Cait Lamberton and Tom Wein, was published by Penn Press in June 2024. Neela is a board member of The Life You Can Save and the Busara Center for Behavioral Economics. She has a PhD in marketing from the Wharton School, University of Pennsylvania.