Journal for the History of Knowledge

KENJI ITO

Who Knows?

▼ **FORUM ARTICLE** in *Decentering the History of Knowledge*
▼ **ABSTRACT** This essay explores some aspects of the future of the history of knowledge by reflecting on my ongoing research concerning the historical study of scientific journals in Japan. It places particular emphasis on the posthumanist approaches and uses of new digital technologies within the history of knowledge.
▼ **KEYWORDS** history of knowledge; posthumanism; flattened ontology; scientific journals; artificial intelligence
▼ **ISSUE** Volume 5 (2024)

Kenji Ito • Kyoto University, JP, ito.kenji.6m@kyoto-u.ac.jp

Cite this article: Kenji Ito, 'Who Knows?', *Journal for the History of Knowledge*, 5 (2024), 179–188
<https://dx.doi.org/10.55283/jhk.19472>

DOI: 10.55283/jhk.19472

BREPOLS

Decentering the history of knowledge is not solely about location but also the types of beings who possess knowledge. After we have witnessed various advocates of the "flattened ontology," and one of the most prominent among them, Bruno Latour, passed away in 2022, considering non-humans as participants in scientific practices should no longer seem novel.[1] Scholars have been deliberating the rejuvenation of the humanities based on posthumanism.[2] It is now high time to incorporate non-human actors into the narrative of the history of knowledge.

Artificial intelligence (AI), especially the recent advancements in large language models, has made it increasingly clear that knowledge is not exclusive to humans. It is likely that this will have a notable impact on how we conceptualize the history of knowledge. It is now easier and more tempting than ever to contemplate how non-human entities can also possess knowledge. The term "machine learning" has become commonplace in our daily vocabulary, and if a machine can learn, it certainly implies that it also possesses knowledge.

Humans and non-humans are undoubtedly distinct, yet a question of strategic importance arises as to whether to commence our exploration from their fundamental disparities or their similarities. It is a mistake to assume complete similarity between humans and non-humans, but the burden of proof now rests with those who claim their differences.

Beyond current trends, it might be beneficial for studies in the history of knowledge to explicitly address the agents of knowledge, given the growing possibility that these agents may not always be human, or at least not purely so (if such a thing as pure humanity exists). When we delve into non-human knowledge, the very comprehension of the history of knowledge becomes accessible to humans and non-humans, which carries certain implications. In this context, I have no intention of speculating about fanciful future research driven by groundbreaking technology, or whether it will come to fruition. Instead, my aim is to reposition my specific (and modest) research project in harmony with non-human cognition. Nonetheless, the consequences may prove to be substantial.

The research project in question is dedicated to the history of scientific journals in Japan, specifically focusing on the twentieth-century Japanese physics community. Within this research, my primary objectives are to investigate the knowledge-making processes employed by Japanese physicists,

---

1 Deleuze and Guattari, *A Thousand Plateaus*; Haraway, *Simians, Cyborgs and Women*; Haraway, *When Species*; Haraway, *Staying with the Trouble*; Latour, *The Pasteurization of France*; Latour, *Reassembling the Social*; Pickering, *Mangle of Practice*; DeLanda, *A New Philosophy*; Barad, *Meeting the Universe Halfway*; Henare, Holbraad, and Wastell, *Thinking Through Things*; Helmreich, *Alien Ocean*; Bennett, *Vibrant Matter*; Kirksey and Helmreich, "Emergence of Multispecies Ethnography"; Harman, *The Quadruple Object*; Shaviro, *The Universe of Things*; Puig de la Bellacasa, *Matters of Care*; Fox and Aldred, *Sociology and the New*.

2 Haraway, *Modest*; Massumi, "Sensing the Virtual"; Braidotti, *The Posthuman*; Braidotti, "A Theoretical Framework."

analyze the knowledge that is produced and published within these journals, trace the evolution of scientific authorship, and gain insight into the prevailing perceptions of scientists who either engage in or abstain from writing scientific papers.

In recent decades, the study of scientific journals has gained considerable traction within the field of the history of science. Melinda Baldwin's research centered on *Nature* exemplifies the possibility of crafting a biography for a scientific journal while providing valuable insights into the origins of this now excessively prestigious commercial publication. Similarly, Alex Csiszar's examination of scientific journals in England and France highlights the complexities of applying the contemporary definition of a scientific journal, which emerged in the nineteenth century, to periodicals from early modern Europe. This incongruity reflects the evolution of the concept of "science" itself, which has undergone significant transformation since the seventeenth century. Additionally, Aileen Fyfe and her colleagues have significantly contributed to this discourse by compiling a comprehensive volume tracing the 350-year history of publications from the Royal Society. This compilation encompasses *Philosophical Transactions* and the *Proceedings of the Royal Society of London*, enabling the authors to delineate the diachronic changes in scientific publications and offer nuanced insights into the evolving landscape of knowledge dissemination.[3]

The scientific journal presents an excellent subject for the history of knowledge. One apparent justification for this lies in the issue of disciplinary boundaries. The historical exploration of scientific journals cannot be adequately contained within the narrow confines of the history of natural sciences. This is because periodicals relevant to the studies of scientific journals did not always focus exclusively on the natural sciences within the context of modern science. An illustrative example is the French journal *le Journal des sçavans*, which served as one of its origins. This publication did not limit itself to natural knowledge but also encompassed humanistic subjects.[4]

A more fundamental justification for the significance of scientific journals in the history of knowledge is rooted in their inherent function. Ever since the groundbreaking work of Thomas Kuhn, and even preceding that with Ludwik Fleck, the process of knowledge production has been acknowledged as more than an isolated cognitive activity within an individual's mind; it is increasingly seen as a collective effort that engages multiple individuals within a scientific community.[5] Scientific journals are pivotal in this institutionalized and widely dispersed process of knowledge creation within the scientific community. The scientific journal can be considered a kind of knowledge infrastructure, along

3 Baldwin, *Making "Nature"*; Csiszar, *The Scientific Journal*; Fyfe, Moxham, McDougal-Waters and Mørk Røstvik, *A History of Scientific Journals*.

4 Brown, "History and the Learned Journal." See also, Fyfe et al., *A History*, passim; Csiszar, *The Scientific Journal*, 24–32.

5 Kuhn, *Structure of Scientific Revolutions*; Fleck, *Genesis and Development*.

with classification or a global network of climate observation.[6] As we delve into topics such as registration, authentication, communication, and the dissemination of knowledge, it becomes inevitable that we explore the realm of scientific journals. This significance is already reflected in the works of scholars such as Peter Burke.[7] Through the lens of a scientific journal, we gain insights into the dynamics of a community of scientists, as Melinda Baldwin has demonstrated in her research.[8] Additionally, we can explore the emergence of the scientific author, a theme that Alex Csiszar has pursued in his work.[9]

While Fleck and Kuhn considered the human collective the basis of scientific knowledge production, the creation of knowledge in general can be seen as an intricately distributed process involving both human and non-human agents.[10] When we conceptualize knowledge as a collective endeavor rather than an individual psychological process, it becomes natural to view extensive collections of written documents as external "memories" of this collective undertaking. Human knowledge is, to a greater extent, preserved in written records rather than confined to individual minds. In this context, libraries take on the role of collective knowledge repositories, surpassing individual humans, particularly in matters expressible through language. Even before the advent of electronic computers, libraries, and more broadly, any collections of written materials operated as memory devices or "knowledge machines."[11] They effectively possessed knowledge akin to how AI systems do. Since long before the emergence of AI, humanity has relied on external memory systems. This underscores the notion that knowledge involves distributed intelligence, encompassing human and non-human actors.

A scientific journal can be regarded as an illustrative instance of external memory. When a journal serves as the official publication of a scholarly society, it frequently forms a substantial part of the institutional memory within that academic institution. If we envision the scholarly society as an entity endowed with the ability to accumulate knowledge, analyzing a scientific journal can be perceived as a form of interspecies communication. It represents an endeavor to examine the cognitive entity that is non-human—the scholarly society—through its manifestations, which are the scientific journals.

This form of communication is naturally beset with challenges stemming from disparities between human and non-human entities, akin to the difficulties encountered in intercultural communications due to cross-cultural differences. The emergence of these challenges has already been elucidated by ex-

6 Star and Bowker, *Sorting Things Out*; Edwards, *A Vast Machine*.
7 Burke, *Gutenberg to Diderot*; Burke, *Encyclopédie to Wikipedia*.
8 Baldwin, *Making "Nature."*
9 Csiszar, *The Scientific Journal.*
10 Hutchins, *Cognition in the Wild*.
11 Krajewski, *Paper Machines*; Blair, *Too Much to Know*; Bittel, Leong, and Von Oertzen, *Working with Paper*; Duncan, *Index*.

amining communications between humans and machines.[12] Similar challenges can also be observed in debates concerning distinctions between human and non-human animal cultures.[13]

These are the challenges frequently encountered by researchers delving into scientific journals. It is inevitable that they grapple with a sense of overwhelming magnitude when confronted by an extensive collection of volumes demanding their scrutiny. Collectively, the volumes of scientific journals represent a distinct, almost alien, form of knowledge, diverging from what researchers typically engage with. Consequently, conducting a research project encompassing numerous scientific journals presents a tangible difficulty. In many instances, perusing every article published within a substantial journal series is simply not humanly feasible. Nevertheless, this dilemma raises the question: How can one claim to provide a comprehensive analysis of a journal without meticulously reviewing its entire content?

A research project aiming to comprehensively review all issues of a scientific journal would require an individual with an inexhaustible capacity beyond human capabilities. Consequently, my approach involved using a method called "topic modeling" to analyze scientific journals. Topic modeling is a natural language processing technique to uncover document patterns. In this model, a "topic" represents a list of words, each assigned a specific statistical weight. Each document is treated as a "bag of words" and assigned a predefined number for the topics. Each topic has its own associated statistical weights. If a document is predominantly about a specific "topic" to a certain percentage, and that topic predominantly comprises specific words to a certain percentage, the product of these percentages offers a measure of the word's frequency within the document.[14] Various methods can be employed to calculate these statistical weights, with one approach being iterative calculations through Bayesian inference, assuming a Dirichlet distribution (referred to as latent Dirichlet allocation, or LDA).[15] Upon completion of numerous iterative processes, the eventual compilation of topics bestows unique statistical weights upon each document. Essentially, this methodology equips a machine with the capacity to proficiently "comprehend" all the papers contained within a journal and relay their content to us. Remarkably, it holds an advantage over certain text-mining methods, as it does not mandate the prior specification of keywords. Rather, the analysis of texts inherently generates appropriate keywords.

With the emergence of ChatGPT and other computer applications based on large language models, my research project has required a redirection. Nonetheless, the fundamental premise remains unaltered. I continue to rely on

12 Suchman, *Plans and Situated Actions*.
13 Laland and Galef, *Question of Animal Culture*.
14 Jockers, *Text Analysis*, 135–36.
15 Blei, Ng, and Jordan, "Latent Dirichlet Allocation."

computers to assist in the exhaustive examination of articles within the journals I intend to investigate. More broadly, I require a posthuman approach to the history of knowledge. In this rendition of the history of knowledge, agents of knowledge encompass both the human and non-human realms, affecting both its content and its practice. In essence, this version of the history of knowledge portrays a narrative where non-human entities assume pivotal roles in the creation, dissemination, elimination, and other actions about knowledge. Simultaneously, these non-human entities also play significant roles in the creation of this rendition of the history of knowledge. These dual facets appear to be essential for maintaining consistency once we embrace the tenets of posthumanism.

However, this does not imply that we must attribute a complete spectrum of agency to machines and other non-human entities, whether in terms of content or practices within this framework of the history of knowledge. Topic models or any other computational techniques enabled by non-human entities can aid human researchers in comprehending the body of knowledge under examination. Yet the posthumanist approach to the history of knowledge neither mandates nor rules out the possibility of machines attaining a degree of self-understanding regarding the content.

Nor is it necessary for human researchers to possess a flawless comprehension of machines. Scholars engaged in topic modeling research acknowledge that the outcomes of topic modeling are not straightforward. These results necessitate interpretation, and the interpretability of a topic model itself can become a subject of investigation.[16] The level of mutual understanding between machines and humans remains inherently imperfect, echoed by the imperfect understanding between scientific journals and machines. This situation resembles the one scrutinized by Peter L. Galison in the context of CERN, a large-scale, high-energy physics laboratory. In such an environment, researchers from diverse subcultures can coexist within a "trading zone." Without sharing a common framework, which Kuhn might call a "paradigm," they can still communicate sufficiently, if not perfectly, using a kind of pidgin language and engaging in collaborative efforts based on partial understanding.[17]

A potential consequence of embracing a posthuman history of knowledge lies in the ongoing dissolution of boundaries separating the humanities from other branches of knowledge, notably the natural sciences and engineering. Biologists investigating the cognitive processes of specific animal species, such as insects, birds, or fish, and engineers delving into information technology, may discover a greater degree of common ground with historians of knowledge than previously envisioned. These disciplines can be viewed as continuous rather than isolated by distinct boundaries.

16 Chang et al., "Reading Tea Leaves."
17 Galison, *Image and Logic*.

Another implication pertains to the format of research output in the history of knowledge. Research findings need not be presented solely in written form. The visual representation of scientific results has long been a significant concern for natural scientists.[18] With the emergence of digital humanities and the increasing availability of publication, visualization, and other digital tools, it has become increasingly common for researchers to create digital products, including interactive maps, chronologies, and networks involving both human and non-human actors.[19] These formats need not be considered mere byproducts of research papers or monographs. When necessary, they can undergo the same formal procedures as conventional formats, including peer review. While assessing them may be challenging due to their novelty and limited precedents, this landscape is poised to evolve.

However, it is worth noting that these digital formats, like interactive maps, are primarily designed for human consumption. This need not always be the case. If historians of knowledge start considering non-human knowledge agents, then the audience for their research findings need not be limited to humans. Their work could be structured in a manner that is accessible to machines as well. This would likely entail the establishment of specific protocols for digital content to be "read" by machines. We already possess machine-readable data formats, including markup languages such as HTML, textual data formats such as JSON, and library data formats such as MARC. Machines can also communicate with each other using codes or media primarily designed for machine-to-machine interactions, such as barcodes (or QR codes) and neural networks.

If research output continues to take a written form, the standard for research publications would then shift from merely emulating native English to writing in a manner comprehensible to machines. Establishing protocols for research papers within a specific discipline may not be as daunting as it seems. While this notion might appear ambitious, it remains a reasonable pursuit for the sake of consistency.

The practice of adapting content for non-human readers can have profound implications. Overcoming human exceptionalism might provide a basis for more appropriate environmental norms. Another ethical justification for transitioning to machine-readable research outputs instead of exclusively relying on native English is linguistic fairness and equity. One of the long-standing norms in academic writing, specifically the insistence on conforming to English as defined by native speakers, now appears narrow-minded compared to the attempts to incorporate non-human actors.

The entrenched dominance of the English language in academia is a remnant of colonial legacies. Native English-speakers enjoy unfair advantages over

18 Lynch and Woolgar, *Representation in Scientific Practice*; Coopmans et al., *Representation in Scientific Practice*.

19 Burdick et al., *Digital_humanities*.

non-native counterparts in international scientific competitions. Non-natives require much more time and other resources to read and write in English. Despite the inherent injustices and ethnocentrism in this system, the English language remains an inescapable platform for many. Those who benefit from this linguistic inequality also constitute the mainstream of the ruling faction in the international scientific community. Established scholars who write in English are reluctant to forfeit their advantage, as they are vested in upholding the dominance of English in scholarly communication. Technological innovation is one of the few possibilities that might alter this situation.

It remains to be seen whether machine-mediated communication and the resulting linguistic equity could liberate those who suffer under the tyranny of English. How AI has been developed and managed so far does not allow us to dream of a technological Utopia. On the contrary, if profit-driven capitalist principles continue to develop these new technologies, the worst technological nightmare and a different kind of colonialism might be what we should anticipate.[20] Still, we may or should keep thinking about possible positive outcomes of new technologies. It is not because these outcomes are promised to happen automatically, but precisely because they will not come without much effort to realize them.

About the Author

Kenji Ito is a historian at Kyoto University who explores the global history of knowledge, with a particular focus on Japan. He recently published a two-volume biography of Nishina Yoshio. His current research involves two major projects: the history of scientific journals and the global development of quantum physics.

Bibliography

Baldwin, Melinda. *Making "Nature": The History of a Scientific Journal*. Chicago, IL: University of Chicago Press, 2015.

Barad, Karen. *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Durham, NC: Duke University Press, 2007.

Bennett, Jane. *Vibrant Matter: A Political Ecology of Things*. Durham, NC: Duke University Press, 2010.

Bittel, Carla, Elaine Leong, and Christine von Oertzen, eds. *Working with Paper: Gendered Practices in the History of Knowledge*. Pittsburgh, PA: University of Pittsburgh Press, 2019.

Blair, Ann M. *Too Much to Know: Managing Scholarly Information before the Modern Age*. New Haven, CT: Yale University Press, 2011.

20 For example, Muldoon, Graham, and Cant, *Feeding the Machine*.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3 (2003): 993–1022.

Bowker, Geoffrey C., and Susan Leigh Star. *Sorting Things Out: Classification and Its Consequences*, Cambridge, MA: MIT Press, 2000.

Braidotti, Rosi. *The Posthuman*. Cambridge: Polity Press, 2013.

———. "A Theoretical Framework for the Critical Posthumanities." *Theory, Culture & Society* 36, no. 6 (2019): 31–61.

Brown, Harcourt. "History and the Learned Journal." *Journal of the History of Ideas* 33, no. 3 (1972): 365–78.

Burdick, Anne, Johanna Drucker, Peter Lunenfeld, Todd Presner, and Jeffrey Schnapp. *Digital_Humanities*. Cambridge, MA: MIT Press, 2012.

Burke, Peter. *A Social History of Knowledge: From Gutenberg to Diderot*. Cambridge: Polity Press, 2000.

———. *A Social History of Knowledge: From the Encyclopédie to Wikipedia*. Vol. 2. Cambridge: Polity Press, 2012.

Chang, Jonathan, Jordan Boyd-Graber, Sean Gerrish, Chong Wang, and David M. Blei. "Reading Tea Leaves: How Humans Interpret Topic Models." In *NIPS'09: Proceedings of the 22nd International Conference on Neural Information Processing Systems*, edited by Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, and A. Culotta, 288–96. Red Hook, NY: Curran Associates, 2009.

Csiszar, Alex. *The Scientific Journal: Authorship and the Politics of Knowledge in the Nineteenth Century*. Chicago, IL: University of Chicago Press, 2018.

Coopmans, Catelijne, Janet Vertesi, Michael E. Lynch, and Steve Woolgar, eds. *Representation in Scientific Practice Revisited*. Cambridge, MA: MIT Press, 2014.

Duncan, Dennis. *Index, A History of the: A Bookish Adventure*. London: Penguin, 2022.

DeLanda, Manuel. *A New Philosophy of Society: Assemblage Theory and Social Complexity*. London: Continuum, 2006.

Deleuze, Gilles, and Félix Guattari. *A Thousand Plateaus: Capitalism and Schizophrenia*, translated by Brian Massumi. Minneapolis, MN: University of Minnesota Press, 1987 [originally published in French in 1980].

Edwards, Paul N. *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. Cambridge, MA: MIT Press, 2010.

Fleck, Ludwik. *Genesis and Development of a Scientific Fact*. Edited by Thaddeus J. Trenn and Robert K. Merton, translated by Fred Bradley and Thaddeus J. Trenn, foreword by Thomas S. Kuhn. Chicago, IL: University of Chicago Press, 1979.

Fox, Nick J., and Pam Aldred. *Sociology and New Materialism: Theory, Research, Action*. London: Sage Publications, 2017.

Fyfe, Aileen, Noah Moxham, Julie McDougal-Waters and Camilla Mørk Røstvik. *A History of Scientific Journals: Publishing at the Royal Society, 1665–2015*. London: UCL Press, 2022.

Galison, Peter L. *Image and Logic: A Material Culture of Microphysics*. Chicago, IL: University of Chicago Press, 1997.

Haraway, Donna J. *Simians, Cyborgs and Women: The Reinvention of Nature*. New York: Routledge, 1991.

———. *Modest_Witness@Second_Millennium. FemaleMan© _Meets_Oncomouse™*. New York and London: Routledge, 1997.

———. *When Species Meet*. Minneapolis, MN: University of Minnesota Press, 2007.

———. *Staying with the Trouble: Making Kin in the Chthulucene*. Durham, NC: Duke University Press, 2016.

Harman, Graham. *The Quadruple Object*. Winchester and Washington, DC: Zero Books 2011.

Helmreich, Stefan. *Alien Ocean: Anthropological Voyages in Microbial Seas*. Berkeley, CA: University of California Press, 2009.

Henare, Amira, Martin Holbraad, and Sari Wastell, eds. *Thinking Through Things: Theorising Artefacts Ethnographically*. London: Routledge, 2006.

Hutchins, Edwin. *Cognition in the Wild*. Cambridge, MA: MIT Press, 1996.

Jockers, Matthew L. *Text Analysis with R for Student of Literature*. Heidelberg: Springer, 2014.

Kirksey, Eben, and Stefan Helmreich. "The Emergence of Multispecies Ethnography." *Cultural Anthropology* 25, no. 4 (2010): 545-76.

Krajewski, Markus. *Paper Machines: About Cards & Catalogs, 1548–1929*, translated by Peter Krapp. Cambridge, MA: MIT Press, 2011.

Kuhn, Thomas S. *The Structure of Scientific Revolutions*. With an Introductory Essay by Ian Hacking. Chicago, IL: University of Chicago Press, 2012.

Laland, Kevin N., and Bennett G. Galef, eds. *The Question of Animal Culture*. Cambridge, MA: Harvard University Press, 2009.

Latour, Bruno. *The Pasteurization of France*. Cambridge, MA: Harvard University Press, 1993.

———. *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press, 2005.

Lynch, Michael E., and Steve Woolgar. *Representation in Scientific Practice*. Cambridge, MA: MIT Press, 1990.

Massumi, Brian. "Sensing the Virtual, Building the Insensible." *Architectural Design* 68, no. 5–6 (1998): 16–24.

Muldoon, James, Mark Graham, and Callum Cant. *Feeding the Machine: The Hidden Human Labor Powering A. I*. New York, NY: Bloomsbury Publishing, 2024.

Pickering, Andrew. *Mangle of Practice: Time, Agency, and Science*. Chicago, IL: University of Chicago Press, 1995.

Puig de la Bellacasa, María. *Matters of Care: Speculative Ethics in More than Human Worlds*. Minneapolis, MN: University of Minnesota Press, 2017.

Steven Shaviro. *The Universe of Things: On Speculative Realism*. Minneapolis, MN: University of Minnesota Press, 2014.

Suchman, Lucy A. *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge: Cambridge University Press, 1987.