Memory, Mind & Media (2025), 4, e12, 1–16
doi:10.1017/mem.2025.10008

Memory, Mind & Media

RESEARCH ARTICLE

Wherever there is AI there is memory: AI as the agency of the (synthesized) past

Carl Öhman

Department of Government, Uppsala University, Uppsala, Sweden
Email: carl.ohman@statsvet.uu.se

(Received 31 March 2025; revised 05 August 2025; accepted 07 August 2025)

Abstract

The nexus of artificial intelligence (AI) and memory is typically theorized as a 'hybrid' or 'symbiosis' between humans and machines. The dangers related to this nexus are subsequently imagined as tilting the power balance between its two components, such that humanity loses control over its perception of the past to the machines. In this article, I propose a new interpretation: AI, I posit, is not merely a non-human agency that changes mnemonic processes, but rather a window through which the past itself gains agency and extends into the present. This interpretation holds two advantages. First, it reveals the full scope of the AI–memory nexus. If AI is an interactive extension of the past, rather than a technology acting upon it, every application of it constitutes an act of memory. Second, rather than locating AI's power along familiar axes – between humans and machines, or among competing social groups – it reveals a temporal axis of power: between the present and the past. In the article's final section, I illustrate the utility of this approach by applying it to the legal system's increasing dependence on machines, which, I claim, represents not just a technical but a mnemonic shift, where the present is increasingly falling under the dominion of the past – embodied by AI.

Keywords: AI; anthropomorphism; data; generative AI; memory; philosophy; philosophy of information; philosophy of time

Introduction

Artificial intelligence (AI) is revolutionizing how we remember and engage with our past(s). From chatbots impersonating historical figures (Klee, 2023), to facial reconstructions of marble busts (Panegyres, 2024), Holocaust memorials (Illinois Holocaust Museum, 2025), and AI-driven history education (Pope and Ma, 2024) – the past has never been more present and alive.

This nexus between AI and memory is becoming increasingly well-theorized within a variety of disciplines and subfields (see Gensburger and Clavert, 2024; Hoskins, 2024), not least memory and communication studies, where scholars have long debated the nature of more rudimentary memory technologies (Garde-Hansen et al., 2009; Hoskins, 2017; Kansteiner, 2022). Much of this work has focused on biases and hallucinations in how AI displays historical narratives (Kansteiner, 2022; Kollias, 2024; Aryan, 2025). Another focus area has been the nature of synthetic memory artefacts – plausible narratives that lack reference to

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/mem.2025.10008 Published online by Cambridge University Press

actual events or lived experiences (Hoskins, 2024). A third group have asked complex questions of the relationship between human and non-human agency within these mnemonic processes (Makhortykh, 2024; Richardson-Walden and Makhortyh, 2024; Merrill, 2025) – who is doing the remembering? And how? These are all worthwhile endeavors.

In this article, however, I propose a different interpretation of how AI and memory link together. Specifically, I contend that (data-driven) AI systems are not merely non-human agents acting upon the past, but windows through which *the past itself* acquires agency and is *extended* into the present.[1] Like gods and other totemic beings, modern AI systems have become a means of synthesizing the authority of a collective of ancestors, here represented by the training data, into a unified interactive agency in the present. Chatbots based on ChatGPT, DeepSeek, and LaMDA may not only produce historical facts and illustrations, but also essentially provide a means of chatting with the past itself, or at the very least, the past on which the models have been trained.

This interpretation falls close to Vallor's (2024) metaphorical notion of AI as a mirror reflecting the past into the present. However, the interpretations differ in important ways. For Vallor, AI represents a skewed means of *reflecting* the past. In my interpretation, however, AI *is* the past – only processed. Just as a dish is not merely a reflection of its ingredients but composed of them, the agency of AI systems is constituted by the past on which they are trained. This does not imply that the outputs of generative AI present accurate or complete representations of that past – nor that their content is monolithic. Rather, the point is that their agency emerges from a synthesis of data, and that these data are themselves the material traces of the past. This distinction is crucial. The embodied past – that is, the operational agency of the system – may generate outputs that are not historical documents in any conventional sense. Yet this does not alter the fact that its capacity to act is composed of data that directly originate in previous events. Put plainly: the agency of generative AI is the agency of the past – reconfigured, recombined, and made active in the present.

Why is this interpretation important? There are two reasons: The first is that it reveals the full scope of the nexus between AI and memory. Memory is not merely a dimension of AI; rather, modern AI should be understood as a subset of mnemonic practices in the sense that it in itself *embodies* the content of memory – the past. When an AI model identifies a tumour in an X-ray, edits a student essay, or recommends a product for a consumer, these are all instances of the past (the training data) gaining agency in the present. And so, wherever there is AI, there is memory. And this is why the concept of memory belongs right at the heart of the critical AI debate. The second reason is that the proposed interpretation opens the door to a new understanding of power within the AI–memory nexus, where the central vector is not between man and machine, nor between groups of humans (racial, economic, and class-based), but is rather to be conceived as a *temporal axis* between the *past* and the *present.* From this view, the key risk of AI in relation to memory is not, as is commonly argued, that non-human forces, or powerful subgroups, are coming to dominate our perceptions of the past, but rather concerns a shifting intergenerational power dynamic: the growing dominance of the past – regardless of its virtues – over the present.

To illustrate the utility of the proposed interpretation, I shall apply the temporal power axis to a realm that is not normally considered an instance of AI and memory – namely, *law.* Specifically, I will draw on Dworkin's (1986) legal philosophy to show that the legal system's

[1] It is noteworthy that the argument does *not* encompass all forms of AI. Whereas it may be possible to extend the argument to symbolic AI, this is not within the scope of the present argument. It should also be noted that the argument cannot really be applied to all forms of machine learning applications either. A notable example is Alpha Go Zero, which, albeit a machine learning system, is not trained on any prior data. Henceforth, the term AI refers only to modern data-driven systems that rely on large volumes of training data.

https://doi.org/10.1017/mem.2025.10008 Published online by Cambridge University Press

increasing dependence on machines represents not just a technical shift but also a deepening *tyranny of the past* over the present – a disruption of an intergenerational power equilibrium, which, regardless of its effects on justice, poses an ethical challenge.

The article is structured as follows. In the section 'Of machines and men', I provide a more detailed overview of how the nexus between AI and memory is currently theorized. This is done to illustrate the larger point that, in arguing that humans and non-humans co-constitute memory *together*, present concepts are at the same time establishing the man–machine divide as the central power axis of AI and memory. In the section 'The anthropomorphized agency of the synthesized past', I propose an alternative interpretation of the relationship between AI and memory, in which the agency of the machine is to be understood as the agency of the synthesized (digital) past on which the AI has been trained. To illustrate this point, I draw on a series of analogies to define the three components of the argument: *data*, *synthesis*, and *anthropomorphized agency.* In the section 'Why is this interpretation important?', I illustrate the advantages of this approach by applying it within the domain of law. The 'Conclusion' section summarizes the key findings of the study.

Of machines and men

Current literature on the relationship between AI and memory comprises a variety of concepts and approaches, too many to provide a comprehensive overview. Given the overlap between AI and interrelated concepts like big data, algorithms, and even search engines, it is also difficult to delimit exactly where the AI and memory-literature ends and where the general digital memory studies begin (Garde-Hansen et al., 2009; Garde-Hansen, 2011; Hoskins, 2017). As such, the following section aims merely to illustrate a larger point about this literature, namely that many current conceptualizations are – sometimes unintentionally – predicated on a 'man versus machine' narrative. In fact, the dichotomy often re-emerges precisely through attempts to overcome it. By conceptualizing the ways in which humans and computers intermix in mnemonic processes – concepts like 'hybrid', 'symbiosis', and 'cyborgs' are telling examples – the field is simultaneously cementing its status as the two basic forms of actors. Even when new actants emerge through the human–machine interaction, the underlying conceptual distinction between the two tends to persist, which, in turn, shapes how we think of the risks involved.

For illustration, consider e.g., Hoskins's (2024) observation that the advent of generative AI 'heralds a new battleground between humans and computers in the shaping of reality' (2024, 1), which may '"both enable and endanger human agency in the making and the remixing of individual and collective memory'. Here, AI is construed as a fundamentally new force in the techno-mnemonic development that calls upon us humans to reclaim our agency and ownership of our collective and individual memory, lest it be taken over by the machines. In other words, the elementary relationship involved in these processes – and thereby, the central dimension of power – is that between humans on the one side and machines (computers) on the other.

A similar mix of conflict and intermingling between humans and machines can be traced in Merrill's (2023) conceptualization of AI mnemotechnics as a form of 'cyborgian remembrance'. Drawing explicitly on the tradition of memory studies, as well as Haraway's (1991) cyborg theory, Merrill contends that memory in the age of AI must be theorized as something more than a mere instrumental process. Rather, it is performed by 'a hybrid of machine and organism' (Merrill 2023, 182; Merrill, 2025) where the agency is *distributed* between the two.

The same denunciation of anthropocentric and instrumentalist views of memory is reflected in related concepts such as 'stochastic remembering' and 'distributed agency' (Smit et al., 2024), which emphasize the distributed or shared nature of mnemotechnics as

https://doi.org/10.1017/mem.2025.10008 Published online by Cambridge University Press

something *both* synthetic and organic. It also underpins the (justified) critique of anthropomorphism in AI discourse. For example, focusing specifically on collective memory, Richardson-Walden and Makhortyh (2024) propose that a constructive future with AI in mnemonic practices requires a recognition of the machines' '*radical* alterity' (333). Only by recognizing that machines are a fundamental *other* can humans and AI form the kind of productive relationships they refer to as 'human-AI memory symbiosis' (ibid). Whereas Richardson-Walden and Makhortykh are critical of narratives that rely on a 'simplistic binary of human vs non-human', (ibid) they nevertheless appear to view these categories as the most fundamental components involved in AI mnemonics, noting that 'AI is a complex amalgamation of different systems shaped by their relationality to humans at multiple stages of development and deployment'. In other words, AI is built by and influenced by humans, but it remains *fundamentally alien.* The machines are not human, and must be recognized as such.

In a related argument, Makhortykh (2024) has provided a more detailed framework for understanding the role of non-human agents – specifically, robots – in the production and experience of memory. Makhortykh distinguishes between three different forms of memory communication: (1) human-to-human; (2) human-to-robot; and (3) robot-to-robot, though he does not elaborate on the conceptual distinctions between these. For example, he seems to view archives and even online platforms as instances of 'human-to-human' interaction, but does not specify at which point such technologies are to be considered 'robots'. A recommender system, providing access to the past based on user preferences, appears to fall under 'human-to-robot communication'. But, what about a simple 'algorithm' that provides similar guidance but on a binary basis, such as by recommending X to all female users and Y to males? Would this be a robot? The question suggests the presence of significant conceptual grey zones. What is clear, however, is that Makhortykh sees many possible concerns in the emergence of AI as an active shaper of collective (and individual) memory – the robots, imbued with biased sentiments and goals, are coming for human memory.

A plethora of cases similar to the ones given here can be named, including concepts such as 'artificial collective memories' (Kollias, 2024), 'artificial memory' (Schuh 2024), and 'artificial communication' (Esposito, 2022), which demonstrate the field's growing theoretical abundance. Nevertheless, they are all predicated on the same two basic components: On the one hand, we have the humans, and on the other, we have the non-humans in whatever form (machines/computers/or algorithms). To be perfectly clear, the point of most theoretical approaches is to illustrate how poorly this dichotomy maps onto the processes wherein remembering with AI occurs. As Smit et al. (2024, 218) have it, 'rather than seeing the human and non-human as strictly separate entities in remembering', we ought to 'regard the human and non-human as co-constituting remembering'. Nonetheless, present theoretical stances commonly end up reinforcing the vector between the two as the fundamental power axis of the AI-memory nexus. The very distinction between humans and non-humans – a vocabulary that underpins nearly every theory – suggests as much. Indeed, construing something in terms of a 'hybrid' or 'symbiosis' does not dissolve the binary but risks reinforcing it. A hybrid car, for example, is not a vehicle driven by something *beyond* standard fuels, but a mixture between them, which ultimately confirms their inevitability.

One may reasonably object that, by terms like 'hybridity', present theoretical framings are not merely denoting an additive mix, such as in the fuel analogy, but a much more complex web of interrelations between humans and non-humans that together form new actants. Yet, for the purposes of the present argument, the complexity of these relations is of minor importance insofar as the basic components remain unchanged. Again, the problem is not that scholars like Hoskins, Makhortykh, and Merrill (and, by extension, Latour and Haraway) think that humans and machines form an absolute binary, but that the conceptual

https://doi.org/10.1017/mem.2025.10008 Published online by Cambridge University Press

and discursive tools they employ accidentally reinforce such a binary. In order to truly rid the field's conceptual tools of the man–machine binary, we need new conceptual analyses – the kind that I begin to sketch in the following section.

Why is the development of such an analysis so urgent? Because the man–machine binary is also underpinning the way we imagine the *threats* that arise from AI's involvement in human memory. As a consequence of its conceptual apparatus, the literature typically construes these threats in two ways. The first is that the machines will come to dominate humanity's perceptions of its own past, such that we lose control over our own history-making. This is arguably how to interpret, for example, Hoskins's (2024, 1) notion of a 'battleground between humans and computers' – a tyranny of the machine over humanity. The second threat is that the machines will reinforce present power asymmetries *between* various sections of society (eg, based on race, sex, class, etc.) (Makhortykh, 2024; Schuh, 2024; Aryan, 2025; Merrill, 2025). Technology is, by its very nature, not neutral, and AI systems are certainly no exception. Rather, as Makhortykh (2024, 13) has it, they are 'strongly connected to other elements of digital capitalism and are often embedded in the systems of colonial relations'. Hence, the real danger, it is thought, lies in the reconfiguration of inter-human power relations.

Whereas these interpretations identify different dangers, they are both predicated on a notion of AI as a non-human, as something *alien* acting upon our mnemonic processes from the outside. Naturally, there are exceptions to this tendency. Schuh (2024, 241), for instance, notably construes large-scale AI models as embodiments of humanity's collective memory themselves, and many of the concepts discussed above are probably designed out of conceptual convenience rather than deeply held convictions of the man–machine divide. Yet, the most theoretically rigorous such exception – though not explicitly about memory – is Vallor's (2024) notion of AI as a form of *temporal mirror*. From Vallor's perspective, AI must not be viewed as an external entity, but is more adequately understood as 'inseparable from humanity' (4). The threat arising from the technology is thus 'not an external enemy encroaching upon our territory' but rather something that 'threatens us from within' (ibid). In contrast to many other conceptual frames, this is a true breakup of the man–machine binary, where our machines are correctly identified as *us* (humans) in the past.

In the following section, I shall defend a similar position – one which, I claim, draws the analogy to its full conclusion. Rather than viewing the central vector as that between humans and machines, I propose that we view it as a relationship between the *humans of the past*, on the one side, and the *humans of the present*, on the other. Unlike Vallor, though, the position I defend is not merely that AI is a technological agency that reflects the past, but rather, that the agency of the machine is to be regarded as a direct, albeit *flawed*, extension of the (human) past itself. The machine, that is, is not merely a mirror, but a *window* through which the past acquires agency and reaches into the present.

The anthropomorphized agency of the synthesized past

It is tempting to view memory, or mnemonic practices, as a subcategory of the ways in which AI impacts social life. After all, the AI revolution is disrupting nearly every sphere in society, and memory is but one of them. Yet, here I propose the opposite: AI, whether in the form of an algorithm identifying a tumour or a chatbot editing a student essay, is a subcategory of memory. Why? Because AI, at least the forms of it that are currently dominant, is essentially an interactive synthesis of the (digital) past – a means through which digital traces of what has been acquire agency in the present. As such, it is always, necessarily, an activation of memory. This argument has three essential components: (1) the past, (2) its synthesis, and (3) its (anthropomorphized) agency. So, to unpack what it means to say that AI is the

https://doi.org/10.1017/mem.2025.10008 Published online by Cambridge University Press

anthropomorphized agency of the synthesized past, we may begin by specifying the meaning of each component, beginning with the past.

What are data?

The dominant AI systems of the day are born out of massive quantities of data.[2] Take OpenAI's ChatGPT-4 as an example: it is trained on a series of expansive datasets that comprise virtually the entire open web, including all entries on Wikipedia and a vast corpus of public domain texts. In fact, for an average person to read and process this data would take close to 100,000 years of uninterrupted effort.[3]

But, what *are* data? In this context, I claim they are best understood as synonymous with *the past*, or in any case, what is left of it. To make this point, we need to begin with the most fundamental level of abstraction – the formal definition of data. Within the philosophy of information, one datum (data singular) is typically defined as a lack of uniformity, or simply as 'a difference which makes a difference' (Bateson, 1972, 271), where the smallest possible unit is the difference between 0 and 1, that is, between being and not being. A more formal articulation is provided by Floridi (2010, 23), who proposes what he calls a *diaphoric* definition of data:

> Dd) datum = def. *x* being distinct from *y*, where *x* and *y* are two uninterpreted variables and the relation of 'being distinct', as well as the domain, are left open to further interpretation.

So, data are defined as a lack of uniformity, which means that even the absence of signals can be informative, insofar as it is contrasted to the presence of signals. It is the difference between the presence and absence of *x* (such as 0 and 1 in a computer system, or the *lack of uniformity* in Shannon information; Shannon, 1948) that is informative.

From this definition, it follows that data can only exist materially, or at least *through* a material mark. The difference that makes a difference needs to be inscribed into *something* physical in order to exist and be conceivable. Or, better, there needs to be a *trace* to constitute the difference.[4] And this is why data are constitutive of the past. A trace, by definition, pertains to the past: it marks something that has occurred, even if the full context and identity of that 'something' may no longer be recoverable. It is an inscription in the material world that constitutes the friction between what was and what is (Öhman, 2020). In short, data are not merely a representation of the past – they are *constitutive* of what remains of it. They are the materialization of time itself (Hägglund, 2008).

In view of this interpretation, it is striking how the notion of data as the 'new oil' holds a deeper truth than its colloquial meaning. Oil is essentially nothing more than the remains of dead organisms packed together into a single substance. As the name 'fossil fuel' indicates, the power source that propels industrial capitalism into the future is, in fact, the remains of the past, specifically the bodies of our long departed (non-human) ancestors that, over time, have accumulated on the sea floor and formed oil. However, could we not say the exact same thing about data – the fossil fuel of the digital economy? It, too, is nothing but the accumulated remains of the past. Whether in the form of Facebook posts, a company's

[2] It is noteworthy that this argument thus excludes a variety of approaches to AI, including rules-based systems and reinforcement learning, which do not rely on data in the sense discussed here.

[3] OpenAI has not publicly disclosed the model's training data, but according to multiple leaks (see Schreiner, 2023) it seems to be around 13 trillion tokens. Assuming one token equals one word – 0.75 is a more conventional measure, but that estimate is not about *reading*, where a one-to-one ratio is more reasonable – and a reading speed of 250 words per minute, that equals 52,000,000,000 min or roughly 98,844 years.

[4] I am using the term 'trace' explicitly in the sense Hägglund (2008) – and originally Derrida – uses it.

https://doi.org/10.1017/mem.2025.10008 Published online by Cambridge University Press

quarterly sales, or the records of an electric meter, data are the *remains* – not just a symbol, representation, or reflection – of the past (Öhman, 2020). Our daily interactions on the web, no matter how tiny, are the micro-organisms that fall to the internet's sea floor, where they will eventually be extracted and refined to drive the digital economy forward.[5]

In fact, even the very words we use for data hint at a kinship with the past. For example, words like journal, archive, chronicle, annal, zeitdung (German), and tidning (Swedish) are all etymologically derived from temporal concepts. And a word like *date*, which denotes a fixed point in time, is derived from Latin *datum* (data singular). This is to say that, philologically, as well as conceptually, data and the past stem from the same root – they are constitutive of one another. Just as a corpse remains after a person's death, and is thus still constitutive of personhood, data are what remains of past events (Stokes, 2015, 2017).[6]

If data are constitutive of the past, and vice versa, does this imply that all data provide an accurate representation of the past? Does that raw data provide access to history beyond narrative and ideology? No. Naturally, some data provide false representations of past events. This is evident in the case of AI, as well as for more rudimentary technologies. If I write '*P* happened yesterday' on a piece of paper, this inscription (a data trace) is not constitutive of yesterday's events. It is, however, constitutive of the moment in which I wrote those words. The inscription is a remnant, a crystallization (to use Marxist language), not of yesterday's events, but of my act of writing. Even when data are directly derived from the events they mediate, such as a photograph, they have a narrative form and are thus imbued with ideology, because data are only meaningful when structured, that is, when they emerge as *information* (Floridi, 2010). As such, there is no opposition between the notion of the past as inevitably narrated and the notion that the past is constituted by data. To state that the past is always narrative is merely to confirm that the essence of the past is informational, for what is a narrative, if not a sequence of information, that is, data plus meaning?

In sum, data are defined as a lack of uniformity, and, as such, they are constitutive of the events that made the material mark through which they exist. Data, in a word, are the traces that remain of what has been.

What is data synthesis?

So, AI systems are born from a synthesis of data – data that, as shown above, are themselves constitutive of the past. However, what does it mean to *synthesize* data? Does such synthesis preserve its constitutive relationship to the past – and if so, in what form? I hold that the answer is yes, and that data synthesis should be understood as a process of *refinement* where the past acts as the raw material from which AI systems emerge.

As a pedagogical illustration of this argument, let us elaborate upon the oil metaphor introduced above. Oil can be used for multiple purposes, such as heating a house or driving a car. But how is, for example, the velocity of a car related to the organisms that compose the oil? The answer is *refinement.* At first glance, velocity and fossils appear categorically distinct, yet the former is, by all means and purposes, a refined version of the latter. Marine microorganisms sink to the ocean floor, and given enough pressure, temperature, and time, they will eventually collapse into a single substance – what we call crude oil. The crude oil

[5] This is probably a suitable moment to note that new AI models are increasingly trained on synthetic data. Are synthetic data also traces from the past? If so, in what sense? This question appears to be too big to answer within the frames of this article, but it is certainly worth exploring in future work.

[6] Notably, none of these answers how the past can logically extend into the present. It is intuitive enough that the corpse may extend past personhood into the present, but insofar as something of the past is *present*, it can no longer be in the past. I contend that this part of the argument remains to be solved, though I am confident there are resources within the philosophy of time to solve it.

https://doi.org/10.1017/mem.2025.10008 Published online by Cambridge University Press

can, in turn, be extracted and put through an elaborate process of distillation, which separates the gasoline contained within it from other hydrocarbons. That gasoline will, in turn, be pumped into a combustion engine, where it is ignited, causing a controlled explosion that pushes the engine piston down. The up-and-down movement of the piston is then channelled via the car's crankshaft to the wheels and converted to a rotating movement, which ultimately leads the car to move forward. And thus, by a gradual process of refinement where the properties of the individual organisms are slowly collapsed into one another – such that eventually, only the energy contained within their cells remains – the fossils are transfigured into velocity. It is impossible to identify a single moment within this process where the organisms cease to be, for the process is not about terminating their existence, but to refine it, such that it becomes useful in the present.

Now, insofar as data is like oil, something similar can be said about the training of an AI system. How is a model related to the data upon which it has been trained? Through a laborious and gradual process of refinement. Consider, for example, the generation of a large language model. The data sources – usually composed of huge segments of the social web, including billions of microscopic (textual) interactions, need to be selected and extracted. These gargantuan archives are then cleaned and adequately formatted before being tokenized to build the model's basic vocabulary and used to set the initial weights. Upon choosing an appropriate architecture, the model then undergoes both unsupervised and supervised learning processes based on the tokenized data, as well as reinforcement learning based on human feedback (another type of *data*). Usually, following multiple rounds of fine-tuning and further safety testing, the model is then integrated with a chatbot interface for interaction. As such, the chatbot is merely the end of a long chain that begins with the data (or really, the human actions that created them), which are gradually moulded into a unified agency. As oil becomes velocity, so are data converted into interactive agency.

It is important to note that the model does not just emerge from the data all by itself. The refinement process requires a lot of *labour* from human developers. To theorize the role of this labour in relation to the data, we may consider a parallel from the philosophy of science, namely Latour's (2000) theory of the *circulating reference* from his famous Boa Vista ethnography. In this study, Latour follows a French-Brazilian team of researchers as they produce a report on the (allegedly shrinking) jungle of Boa Vista in the Amazon. His concern, however, is far more abstract: the ontological relationship between the scientific theory presented by the team and the real, tangible world of the Boa Vista jungle. When we hold them up next to each other – the theory and the world – their relationship initially strikes us as a mystery. In what sense does the theory represent or correspond to 'the real world?' And how can we establish such a link between theory and reality? Latour argues that the answer can only be revealed by carefully tracing the becoming of the former, or rather, how it is actively *generated* by the team's labour. For, as we zoom in on the processes by which the theory comes into being, we see that the relationship is a matter of *reduction* – not representation. Only when we consider the entire *chain of reduction* – 'space becomes a table chart, the table chart becomes a cabinet, the cabinet becomes a concept' (36) – can we see that the final diagram is but an extracted piece of the jungle. The graphics in the researchers' end-report *are* the real world, or in any case, a tiny bit of it, that the team has made portable and interpretable to others in their field. Through a long chain of active refinement, the researchers have brought a piece of the messy, noisy, hot jungle with them and printed it on a piece of paper. The difference is merely the degree of complexity, which has been reduced in virtue of interpretability and portability.

The researchers in Latour's study *refine* a noisy, hot, and complex jungle into an interpretable, portable chart. By analogy, today's AI developers are refining a bundle of noisy and complex data archives into an interactive, user-friendly agency. However, in both cases, the transformation is best understood as a chain of reduction – refinement through

https://doi.org/10.1017/mem.2025.10008 Published online by Cambridge University Press

compression. The jungle never ceases to be itself, even when highly reduced and portable. The data never ceases to constitute the past, even when synthesized. Yes, the final product is transformed beyond recognition, but its raw material is still the same. AI models, that is, are constitutive of the past on which they have been trained. They do not constitute a 'radical alterity' as Richardson-Walden and Makhortyh (2024, 333) have it, but constitute *us* in the past. The machine may not think like us, and certainly does not experience the world in any relevant sense, but just like the velocity of a car is a refinement of dead organisms, AI agents are distilled expressions of our past voices – compressed, recombined, and animated in new form.

Of course, there is a non-trivial distinction between a scientific theory, which can be traced back to the world it abstracts, and a predictive machine, which operates in black-box opacity and is optimized merely for predictive accuracy rather than truth. Indeed, an entire field of research is devoted to outlining the overlaps and differences between the two (see Watson, 2021). Notwithstanding these debates, there is no question about whether the agency of modern AI systems arises from their training data. The chain is obscured, not absent.

This perspective is clearly similar to the mirror metaphor proposed by Vallor (2024). Just like Vallor, it positions AI not as something alien to humanity, but as a reflection of ourselves. Yet, the theory presented here draws the analogy one step further. Vallor insists that the body we see reflected in a mirror is not *our body*. In fact, it is not a body at all, she contends, but *merely* a reflection. I disagree. For when we construe AI models as a form of refinement of the past, they emerge not merely as mirrors that *reflect* it, but rather as windows through which the past is *extended* into the present. The machine is not merely a non-human that does things *to* the past; it is an agency that is constitutive *of* the past. The reflection in the mirror *is* the real thing – only reduced. Indeed, AI mirrors do not encapsulate the totality of the past, just as a normal mirror fails to capture the full dimensionality of the person it reflects. However, the part that is reflected is nevertheless an *extension* of the real person. The body in the mirror is *their* body.

For further clarity, I am using the word *extension* here in the sense McLuhan (2008) uses the term. For McLuhan, (media-) technologies are famously an 'extension of man' – like a prosthetic limb that helps your agency reach across spatial distances beyond the confines of your biological body. Now, the same, I claim, is true for temporal distances. Just like media *extends* human agency across space – the voice you hear in the phone when talking to your mother is hers, not just the phone's imitation of it – they also extend the past into the present. In other words, the generative model that emerges from the AI training process is an extension of the past (the data) into the present. A model is, in the end, a product of its training data, much like a dish is constituted by its ingredients. When you eat a pancake, you are effectively eating eggs, milk, flour, and butter – but not the only possible configuration of these components. In the case of AI models, the 'ingredients' are remnants of the past, and the model serves as a channel or medium through which these fragments are given voice and agency – a (synthetic) pancake made of past events.

What is anthropomorphized agency?

So, AI is a synthesis of the past – but then again, so are oil, scientific theories, and books. What makes AI stand out, though, is that, through it, the past appears not only to be present, but to acquire an interactive, almost human-like *agency*. How should we understand this agency?

The agency of AI has been thoroughly theorized by other scholars, not least Floridi (2023), who correctly identifies generative models like ChatGPT as 'agency without intelligence'.

https://doi.org/10.1017/mem.2025.10008 Published online by Cambridge University Press

What interests us here, however, is not merely whether machines can have agency and how it differs from human agency, but rather the fact that the *past* can acquire agency *through* the machine. To understand how this works, we may begin by considering pre-digital means by which the past has gained agency in the present.

All technologies exhibit at least a minimal degree of interactivity. A book shows you new information when you turn the pages. A door opens when you press the knob or a button. These are instances of crystallized human actions, as illustrated by Latour's (1992) famous sleeping policeman – human agency transferred to an object. AI accelerates this kind of interactive agency by orders of magnitude. However, it does so in a specific, cumulative way, and here too, the pre-digital world holds some illuminating precedents.

A particularly striking one, as outlined in Öhman (2024), is the emergence of high gods from the practice of ancestor veneration. As theorized by early anthropologists, primordial forms of religious worship were tightly related to the authority and wisdom of past kin (Spencer, 1870; Durkheim, 1995; Harrison, 2010), the most famous example being Durkheim's theory of totemism among Australian Aboriginals. In these 'primitive' religious forms, tribes worship plants or animals, so-called *totems*, that are unique to their community. The tribes experience these totems as possessing an independent agency and immense power. However, in Durkheim's interpretation, this agency stems not from the object, but from the group itself. For the totem, reasons Durkheim, is at the same time a form of deity, a portal to the spiritual dimension, and a representation of society, which leads him to conclude that 'if [the totem animal] is at once the symbol of the god and of the society, is that not because the god and the society are only one?' (201). In other words, the totem is merely a channel, a medium through which the group's collective consciousness *gains agency.* When the group invests its trust, faith, and loyalty – its *mana*, in Durkheim's terms – into a common object, that object will reflect this energy back upon them as a single force. It will appear to have an agency of its own, larger than the sum of its parts. This mechanism, argues Durkheim, is the origin of all religion. The difference is only that, in more recent religious forms, such as the Abrahamic faiths, the materiality of the totem has become redundant. The collective force of the group is projected, not towards a physical object but towards an abstract anthropomorphized entity – God.

However, gods are not only representations of the collective here and now. Rather, they encompass the group as a *temporally extended entity* that also, or perhaps *primarily*, includes its past members. In fact, early anthropologists often speculated that it was the practice of ancestor veneration from which all religious life originally spawned (Spencer, 1870; Tylor 1871). Over time, the theory goes, as the ancestral figures became more temporally distant, their individual identities merged into collective entities existing *beyond time.* And these entities eventually evolved into gods. As such, the totem (god) is an *anthropomorphized* representation of the tribe's collective present and past. It is a means of giving a human face to an otherwise abstract concept that helps integrate the presence of innumerable ancestors within the contemporary community. Religion, thus, is essentially a 'chain of memory', as Hervieu-Léger (2005) puts it, whose primary function is to hold the group together over time by granting agency to the past within the present.

The agency of AI, I argue, follows a strikingly similar logic, for it too is a synthesis of the collective, a means by which our *digital* ancestors gain agency in the present. The Aboriginals undoubtedly experience their totem as possessing a powerful agency, and similarly, today's AI users are experiencing the machine as possessing agency too. However, in both cases, the totem/machine is merely a mask for the human collective. In the former case, this mask is composed of *mana*; in the latter, it is composed of *data* – the mana of the digital age. Thus, what is experienced as a force acting from outside of humanity is really the human past, which has found a window into the present through which it has gained a reduced form of agency.

https://doi.org/10.1017/mem.2025.10008 Published online by Cambridge University Press

In what sense does this agency of the past fall within the definition of *memory?* There is no human who is actively *remembering*, and, once trained, there is no need for the AI to retrieve any data from the past. The answer is that AI performs no active retrieval; its 'memory' lies in its architecture, not in recall. The system is itself already an embodiment of the past. Thus, we may understand its mnemonic agency by analogy to *muscle memory* – an embodiment of a series of training exercises encoded into the nervous system, yet lacking reference to any particular event. Or, better still, using Tulving's (1972) vocabulary of *episodic* versus *semantic memory*, one may say that the AI lacks episodic memory but that its interactive abilities display a form of *semantic memory* (in a purely metaphorical sense, of course). If we disregard the discrepancies between biological and synthetic forms of memory, then, metaphorically, AI resembles a person with permanent amnesia – able to respond accurately about the past, but incapable of remembering it. It is living memory, without the capacity for recall.

In sum, the anthropomorphized agency of AI systems emerges much in the same way that high gods first arose. As with gods and totems, it is tempting to ascribe the agency to the surface itself, rather than the past that is extended through it. However, in each instance, such an impulse is a mistake. Totems, gods, and AI models are not mere non-human agencies acting upon our past, but windows through which it acquires agency.

Summary

To sum up, modern forms of AI are to be interpreted as the anthropomorphized agency of the synthesized past. Data, like oil, are literally what remains of the past, and when synthesized, such as through the processes from which large language models emerge, they are still to be regarded as such. Hence, the agency of AI is not merely something acting *upon* digital memories; it is better understood as an agency *of* our collective digital memory, or in any case, the raw material of which it is composed. This interpretation, I claim, offers a more radical way of transcending the classical man–machine binary. Yet, as we shall see in the following section, its value does not arise merely from its descriptive accuracy, but in what it allows us to *see* and *do.*

Why is this interpretation important?

Even if one accepts my proposed interpretation, why does it matter? What does it contribute to our understanding of AI and memory that existing frameworks cannot? Or, better, what can we *do* with it that we could not before? The answer is twofold.

First, the interpretation of AI as the anthropomorphized agency of the synthesized past reveals the full scope of the AI–memory nexus, for through it, all implementations of AI are revealed as instances of memory in action. In the current literature, the nexus of AI and memory is typically (but not always) confined to instances that are clearly linked to mnemonic practices, such as the use of AI for historical information and illustration in teaching and/or research. However, the interpretation proposed in this article – that AI is itself an extension of the past – shows that every application of AI constitutes an act of memory. Even when a language model edits an essay, or when a more rudimentary machine learning model identifies a tumour in an X-ray, these are instances in which the past is reactivated in the present – manifesting what we might call synthetic semantic memory (in Tulving's sense). They are instances where the past becomes present, indeed, gains *agency*, without necessitating any form of *remembering* in any human sense of the word. As such, the study of the relationship between AI and memory emerges not as a niche but as the very centre stage of critical AI research.

https://doi.org/10.1017/mem.2025.10008 Published online by Cambridge University Press

The second benefit is that the proposed interpretation allows us to critically review the dangers of AI in a new, and I think clearer, light. Like the AI-memory literature, the general critical AI literature is predicated on either of two narratives: it is either humans versus machines or humans versus other humans (based on race, class, sex, etc.). These are the two major power axes that dominate the current debate. While both narratives offer important insights, they overlook a third axis: time. On this axis, the battle is not between humans and computers, nor between different sections of society, but between the present and the past of humanity (see Figure 1), where the machine acts as a conduit (albeit not a neutral one) through which the authority of former generations remains in the present.

As we have noted, Vallor (2024), among others, has proposed a similar position, where AI is but a reflection of our past selves, which dissolves the dichotomy between humans and machines. The machine *is* the human(s) of the past. For Vallor, however, the main problem with this is that the past is flawed in various ways, which leads to power asymmetries between various sections of society here and now. By contrast, the purely temporal power axis that I propose allows us to review the relationship between the present and the past in its own right, irrespective of its effects on the power balance *within* the present society. From such a perspective, the problem of being dominated by the past (acting through our machines) is not that the past is flawed, or represents only some (elite) sub-groups, but that it disturbs the *intergenerational* power balance of society. Let me explain.

Illustrating the approach: AI as a threat to intergenerational power balance

To illustrate the concept of a temporal power axis, consider the domain of law – a recurring battleground within AI ethics, but not a common object of study for memory studies. AI is now deployed across multiple legal domains – from recidivism risk assessments and legal strategy drafting to assisting in judicial decision-making. Much has been said about this topic and the dangers emerging from it, particularly regarding the inscrutability of black-box models, and the risk of discrimination and inherent bias against marginalized groups (O'Neil, 2017; Dressel and Farid, 2018; Surden, 2019; Barysė and Sarel, 2024). However, by interpreting AI not merely as a technology applied *to* the past (e.g., precedential rulings), but as an extension of it, we see another power axis – that between past and present generations of judges.

To illustrate what is at stake in this generational battle, we may base our analysis on the legal philosophy of Dworkin (1986). Dworkin's jurisprudence emerges in opposition to

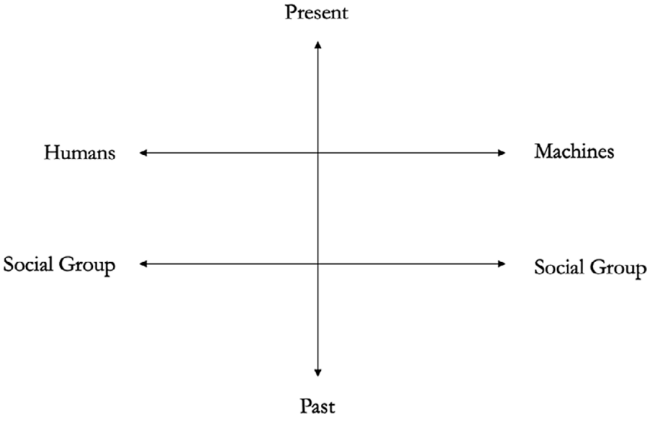

Figure 1. The temporal axis.

https://doi.org/10.1017/mem.2025.10008 Published online by Cambridge University Press

intentionalist and originalist theories, which see law as a communicative practice, where the written word is a memory from previous generations to be semantically deciphered in the present. The practice of law, in this sense, is *formalized remembering.* While Dworkin agrees that legal practice is about interpretation, he understands this as a moral and political process that is intimately bound to the *present.* Judges must not only understand and interpret the letter of the law; crucially, they must integrate this interpretation into a coherent whole that serves the best overall outcome for the community as it stands today. As such, the practice of law is an inescapably moral and political undertaking, which calls upon judges to make their own subjective assessments in the face of the entirety of the law, as well as the entirety of contemporary society. They are both interpreters and authors.

Dworkin captures this logic through his well-known chain novel analogy (ie, a fictional story written by multiple consecutive authors). To write the best possible book, each author must interpret what their predecessors have written and try to fit it into a larger, cohesive narrative. They should not divert from, nor directly contradict, what has already been written. Sometimes this task is impossible, though, as some predecessors may have added contradictory elements to the story. In such cases, it is up to the present author to decide which aspects of the story weigh more. A character may, for example, have acted contrary to their overall personality in a previous chapter, and thus, for the sake of the whole, such episodes may be disregarded. In this sense, the demarcation between interpretation and authorship is blurred. To write, each author *has to* make interpretations regarding the relative weight of each detail in the face of the whole.

Dworkin views the role of judges in a similar manner. They are always bound to respect and abide by previous decisions, but such a task is inevitably interpretive. So, what principles should guide their interpretations? Like the authors writing a chain-novel, argues Dworkin, judges should strive to create the best outcome in light of the whole. The task is not simply to repeat what the predecessors did, but to move forward in a direction that does not directly contradict them. As such, the practice of law depends upon a form of *intergenerational power balance.* Present judges must, in effect, abide by laws passed by parliaments of the past, and by rulings made by past judges, but it is their privilege to *interpret* these in light of the society in which they currently live. The dead may have written the laws and thus govern the living. Yet, the living are not merely bound to the intentions of their words, but rather their *meaning* in light of present-day circumstances. The judges of today, that is, have the privilege of making genuine *decisions* that are not merely products of the past but independent *judgements.*

What happens when AI enters this framework? As judges come to depend increasingly on AI to make their verdicts, they are, by all means, increasing their chances of staying aligned with previous decisions. The machine can process far more information about the past than a single human and does not make any subjective assessments, notwithstanding its biases, along the way. From a positivist semantic perspective, as well as an intentionalist/originalist perspective, this is a good thing. Judges should keep in line with the letter of the law as well as the rulings of the past. And if a machine can do it with even fewer subjective assessments, it is all for the good. For someone who cherishes the power balance between generations, however, it is bad news. If we interpret AI as a means of giving agency to the (digital) past, it means that we are moving towards a situation where the past is represented both by the *law* (an extension of past parliaments) and the *precedential rulings* (an extension of past judges), while at the same time assuming the role as the prime interpreter of these. In effect, the past – embodied by AI – is tasked with interpreting itself, as encoded in laws and precedent. And so, the generational power balance is broken. Even if the previous generation had been flawless saints, and the AI completely unbiased, this is a negative development. Why? Because, as Dworkin points out, the practice of law is inevitably a moral and political effort. It stays alive only because each generation adds new decisions and

https://doi.org/10.1017/mem.2025.10008 Published online by Cambridge University Press

interpretations. Whether the AI will make more accurate decisions, therefore, is irrelevant. The point is that the best decisions must come from the present generation of judges, not from a loop where the past interprets itself (however consistently it may do so). While more accurate, AI replaces self-governance with what can only be referred to as a 'tyranny of the past' (Öhman, 2024, 9).

This analysis illustrates how applications of AI that lie fairly far from the traditional study of memory studies may indeed be analysed as instances of memory. It also illustrates the analytical payoff of doing so. Again, the question asked here is not whether humans are better or worse than machines in making decisions that support human values, nor whether it skews the power balance between various groups in society. The question is how it shifts the power balance between the present and the past, regardless of the power relations *within* the present. Here, I used courts as an example of how that conflict manifests itself – as a disruption of the intergenerational balance that forms the very foundation of the practice of law (if you listen to Dworkin). However, the same perspective can be applied more or less to every instance where AI is implemented: parliaments, healthcare, science, and education.

Conclusion

I have argued that the dominant approaches to theorizing the AI–memory nexus remain (unintentionally) tethered to a binary between humans and machines. Though often framed as overcoming this binary – typically by emphasizing hybrid human–machine configurations – such concepts ultimately reaffirm humans and machines as the foundational components of memory systems.

In contrast, this article has proposed an alternative account in which AI does not function as a non-human agent acting upon memory, but as a medium through which the past itself acquires agency in the present. AI, in this view, is best understood as the anthropomorphized agency of the synthesized past. It is always already about memory – not because it retrieves the past, but because it is composed of it.

This framework opens a new avenue for critique. Rather than locating AI's power along familiar axes – between humans and machines, or among competing social groups – it reveals a temporal axis of power: between the present and the past. While this analysis has focused on the legal system, its implications stretch far beyond. Wherever AI is deployed – in governance, education, healthcare, or science – the past is not merely referenced; it is reanimated, given form, and made to act. Wherever there is AI, there is memory.

Funding statement. This research was generously funded by Wallenberg AI, Autonomous Systems and Software Program – Humanity and Society.

Competing interests. The author declares none.

References

Aryan M (2025) Epistemic injustice in AI-generated histories: Evaluating cultural bias, hallucinations, and community sovereignty. Available at https://escholarship.org/uc/item/92z102k6. Accessed 27 August 2025

Barysė D and Sarel R (2024) Algorithms in the court: Does it matter which part of the judicial decision - making is automated ? *Artificial Intelligence and Law* 32 (1), 117–146. https://doi.org/10.1007/s10506-022-09343-6.

Bateson G (1972) *Steps to an Ecology of Mind Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. San Francisco: Chandler

Dressel J and Farid H (2018) The accuracy, fairness, and limits of predicting recidivism. *Science Advances* 4 (1), 1–6. https://doi.org/10.1126/sciadv.aao5580

Durkheim E (1995) *The Elementary Forms of Religious Life*. New York City: The Free Press

Dworkin R (1986) *Law's Empire*. Cambridge, MA: Harvard University Press

https://doi.org/10.1017/mem.2025.10008 Published online by Cambridge University Press

Esposito E (2022) *Artificial Communication: How Algorithms Produce Social Intelligence*. Cambridge, MA: MIT Press
Floridi L (2010) *Information: A Very Short Introduction*. Oxford: OUP Oxford
Floridi L (2023) AI as agency without intelligence: On ChatGPT, large language models, and other generative models. *Philosophy and Technology* 36 (1), 1–7. https://doi.org/10.1007/s13347-023-00621-y.
Garde-Hansen J (2011) *Media and Memory*. Edinburgh University Press. JSTOR, http://www.jstor.org/stable/10.3366/j.ctt1r25r9. Accessed 28 Aug. 2025.
Garde-Hansen, J, A Hoskins and A Reading (2009) *Save as ... Digital Memories*. New York: Palgrave Macmillan.
Gensburger S and Clavert F (2024) Is artificial intelligence the future of collective memory? *Glass International* 47 (4), 16–30. https://doi.org/10.1163/29498902-202400019.
Hägglund M (2008) *Radical Atheism: Derrida and the Time of Life*. Palo Alto: Stanford University Press.
Haraway D (1991) A cyborg manifesto: Science, technology, and socialist-feminism in the late twentieth century. In *Simians, Cyborgs, and Women: The Reinvention of Nature* 149–181. New York: Routledge.
Harrison RP (2010) *The Dominion of the Dead*. Chicago: University of Chicago Press. Available at http://www.myilibrary.com?id=258479.
Hervieu-Léger D (2005) Religion as a chain of memory. *Nova Religio*. 8. https://doi.org/10.1525/nr.2005.8.3.128
Hoskins A (2017) *Digital Memory Studies: Media Pasts in Transition*. London: Routledge.
Hoskins A (2024) AI and memory. *Memory Mind and Media* 3 (18), 1–21. https://doi.org/10.1017/mem.2024.16
Illinois holocaust Museum (2025) *Interactive Holograms: Survivor Stories Experience*. Available at https://www.ilholocaustmuseum.org/exhibitions/survivor-stories-experience/.
Kansteiner W (2022) Digital doping for historians: Can history, memory, and historical theory be rendered artificially intelligent? *History and Theory* 61 (4), 119–133. https://doi.org/10.1111/hith.12282.
Klee, Miles (2023) "Historical figures" AI lets famous dead people lie to you. *Rolling Stone Magazine*. 2023. Available at https://www.rollingstone.com/culture/culture-news/historical-figures-ai-chat-bot-lies-dead-people-1234664257/.
Kollias P-A (2024) Nostophiliac AI: Artificial collective memories, large datasets and AI hallucinations. 1, 292–322. https://doi.org/10.1163/29498902-202400014.
Latour B (1992) Where are the missing masses? The sociology of a few mundane Artifacts. *Shaping Technology/Building Society: Studies in Sociotechnical Change*, 225–258. https://doi.org/10.2307/2074370.
Latour B (2000) Circulating reference: Sampling the soil in the Amazon Forest. *Pandora's Hope: Essays on the Reality of Science*. Available at http://www.bruno-latour.fr/sites/default/files/downloads/53-PANDORA-TOPOFIL-pdf.pdf%5Cnhttp://www.jstor.org/stable/3886165.
Makhortykh M (2024) Shall the robots remember? Conceptualising the role of non-human agents in digital memory communication. *Memory, Mind and Media* 3, 1–17. https://doi.org/10.1017/mem.2024.2.
McLuhan M (2008) *The Medium Is the Massage*. London: Penguin
Merrill S (2023) Artificial intelligence and social memory: Towards the Cyborgian remembrance of an advancing Mnemo-technic. *Handbook of Critical Studies of Artificial Intelligence*, 173–186. https://doi.org/10.4337/9781803928562.00020.
Merrill S (2025) Hybrid methodologies for studying social and cultural memory in the postdigital age. In Wang Q and Hoskins A (eds), *The Remaking of Memory in the Age of the Internet and Social Media*. New York: Oxford University Press. https://doi.org/10.1093/oso/9780197661260.003.0015
O'Neil C (2017) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Broadway Books
Öhman C (2020) A theory of temporal telepresence: Reconsidering the digital time collapse. *Time and Society* 29 (4), 1061–1081. https://doi.org/10.1177/0961463X20940471.
Öhman C (2024) We are building gods: AI as the anthropomorphised Authority of the Past. *Minds and Machines* 34 (1), 1–18. https://doi.org/10.1007/s11023-024-09667-z.
Panegyres K (2024) Computer "reconstructions" of faces from ancient times are popular. But how reliable are they? *The Conversation* 2024. Available at: https://theconversation.com/computer-reconstructions-of-faces-from-ancient-times-are-popular-but-how-reliable-are-they-236516. Accessed 27 August 2025.
Pope A and Ma R (2024) Exploring historians' critical use of generative AI technologies for history education. *Proceedings of the Association for Information Science and Technology* 61 (1). https://doi.org/10.1002/pra2.1188.
Richardson-Walden VG and Makhortyh M (2024) Imagining human-AI memory Symbiosis: How re-remembering the history of artificial intelligence can inform the future of collective memory. *Memory Studies Review* 1 (1), 323–342. https://doi.org/10.1163/29498902-202400016.
Schreiner, Maximilian (2023) GPT-4 architecture, datasets, costs and more leaked. In *The Decoder*. Available at https://the-decoder.com/gpt-4-architecture-datasets-costs-and-more-leaked/. Accessed 23 May, 2025.
Schuh J (2024) AI as artificial memory: A global reconfiguration of our collective memory practices?. 1, 231–255. https://doi.org/10.1163/29498902-202400012.

https://doi.org/10.1017/mem.2025.10008 Published online by Cambridge University Press

Shannon CE (1948) A mathematical theory of communication. *Bell System Technical Journal* 27 (3), 379–423.
Smit R, **Smits T and Merrill S** (2024) Stochastic remembering and distributed mnemonic agency. *Memory Studies Review* 1(2), 209–230. https://doi.org/10.1163/29498902-202400015.
Spencer H (1870) On ancestor worship and other peculiar beliefs. *Fortnightly Review* 13 (7), 535–550.
Stokes P (2015) Deletion as second death: The moral status of digital remains. *Ethics and Information Technology* 17 (4), 1–12. https://doi.org/10.1007/s10676-015-9379-4.
Stokes P (2017) Temporal asymmetry and the self/person Split. *Journal of Value Inquiry* 51 (2), 203–219. https://doi.org/10.1007/s10790-016-9563-8.
Surden H (2019) Artificial intelligence and law. *Georgia State University Law Review* 35 (4), 1306–1337. https://doi.org/10.1007/978-981-16-1665-5_3.
Tulving E (1972) Episodic and semantic memory. In Tulving E and Donaldson W (eds), *Organization of Memory*. New York: Academic Press
Tylor, E B (1871) *Primitive Culture: Researches into the Development of Mythology, Philosophy, Religion, Art, and Custom*. London: John Murray.
Vallor S (2024) *The AI Mirror: How to Reclaim our Humanity in the Age of Machine Thinking*. New York, NY: Oxford University Press. https://worldcat.org/title/1423244235
Watson D (2021) *Explaining Black Box Algorithms: Epistemological Challenges and Machine Learning Solutions* [PhD thesis]. University of Oxford

Carl Öhman is an associate professor of political science at Uppsala University, Sweden. He has published on a variety of subjects, including AI ethics, deepfakes, digital afterlife, and digital preservation. He is the author of *The afterlife of data* (2024).

Cite this article: Öhman C (2025) Wherever there is AI there is memory: AI as the agency of the (synthesized) past. *Memory, Mind & Media* 4, e12, 1–16. https://doi.org/10.1017/mem.2025.10008

https://doi.org/10.1017/mem.2025.10008 Published online by Cambridge University Press