SERRC
Social Epistemology Review & Reply Collective
http://social-epistemology.com
ISSN: 2471-9560

Knowledge from AI

S. Orestis Palermos, University of Ioannina, sopalermos@uoi.gr

———————————

Palermos, S. Orestis. 2025. "Knowledge from AI." *Social Epistemology Review and Reply Collective* 14 (11): 54–64. https://wp.me/p1Bfg0-aC6.

SERRC
Social Epistemology Review & Reply Collective
14 (11): 54–64. 2025.
https://wp.me/p1Bfg0-aC6

**Abstract**

I examine whether Generative AI systems such as ChatGPT can transmit knowledge and, if so, how epistemic responsibility for their outputs should be distributed. To address these issues, the discussion proceeds in three stages. First, I situate John Greco's notion of massively shared agency within the recent literature on AI testimony, showing how it challenges the widespread assumption that AI cannot testify due to its lack of intentions. Second, I combine Greco's framework with distributed virtue reliabilism (Palermos 2022a) to argue that Generative AI systems function as large-scale epistemic collaborations whose reliability and responsibility emerge collectively. Third, I address the resulting problem of accountability—what I call the 'problem of the invisible hand,' in which attributions of individual responsibility appear metaphysically inapplicable. I propose resolving this tension by recognising both group-level and individual-level responsibility—the latter associated with those individuals that make up the insitution's 'kernel' of decision-making. Accordingly, while Generative AI may count as an epistemic group agent, epistemic liability ultimately also remains anchored in the individuals that design, maintain, and deploy such systems.

Over the past few years, we have become increasingly reliant on a variety of online socio-technical systems—including Wikipedia, Reddit, and X—for acquiring knowledge. The latest and most remarkable development in this growing trend is Generative AI, including models like Chat-GPT, Gemini, Claude, and DeepSeek.

John Greco (2025) recently offered a novel account of 'massively shared agency' to explain how large-scale institutions such as Wikipedia and Google Search can transmit knowledge through testimony. Toward the end of his paper, Greco also reviews several epistemological models that might capture how (Generative) AI can generate and transmit knowledge, though he stops short of providing a detailed analysis of the epistemics of this technology.

Avoiding further detail is understandable, given the explanatory target of Greco's paper—the transmission of knowledge by large-scale institutions *in general*. Besides, Generative AI represents a particularly peculiar epistemic source. Despite the human-like responses produced by such technologies, the scale and complexity of the socio-technical systems that underly them raise doubts about whether they can genuinely transmit knowledge, rather than merely provide information that users may subsequently employ to acquire knowledge. Nevertheless, Greco's discussion is especially valuable for advancing this emerging debate, as it brings to the fore three key questions concerning the transmission and reception of *information* from Generative AI technologies:

1. Can Generative AI act as a testifier?

2. How does Generative AI *generate* knowledge (if it does)?[1]

3. If Generative AI generates and subsequently transmits knowledge, who bears responsibility for the transmission of this knowledge?

In what follows, I pick up the discussion where Greco leaves off by examining these three questions in turn. I begin by situating Greco's account of 'massively shared agency' within the existing literature on AI testimony. While most authors doubt that Generative AI can act as testifier—primarily because of its apparent lack of intentions—Greco's approach offers a fresh perspective that opens the possibility of AI manifesting the intentions necessary for acting as a source of testimony.

If Greco's account is correct—if AI is indeed the kind of entity that can testify—then the next question arises: how might we account for the hypothesis that Generative AI generates knowledge (which is then transmitted to users of the technology)? Finally, I address a further and more troubling issue concerning epistemic responsibility: how should we understand the responsibility involved in this form of knowledge transmission? To answer this, I examine the mechanics of the underlying socio-technical systems, focusing on Open AI's ChatGPT as a case study, though my points likely extend to other Generative AI models as well.

With regards to the second question above, among the proposals Greco suggests for understanding the epistemics of AI, he refers to a previously developed account of epistemic collaborations (Palermos 2022a), based on what I call 'distributed virtue reliabilism' (DVR). This is an interesting recommendation, as the workings of Generative AI systems suggest they may indeed be fruitfully understood as epistemic collaborations—albeit, as I will explain, *special instances* of them. Due to their large scale and massively distributed structure—which often involves users from around the world—the socio-technical systems underlying Generative AI pose a distinctive challenge for distributed virtue reliabilism: specifically, DVR can yield intuitive results concerning the attribution of epistemic responsibility in Generative AI only if it is combined with a suitably modified version of Greco's take on 'massively shared agency.'

## Can AI Testify?

In recent times, philosophers have increasingly turned their attention to the topic of AI testimony (Heersmink et al. 2024; Freiman 2024, He and Yang 2025), often arguing that AI systems are incapable of genuine testimony. The most common reason for this denial

---

[1] The following takes for granted that Generative AI generates knowledge (as opposed to mere content or information). I make this assumption, because users seem to often treat Generative AI as a source of knowledge, despite been explicitly warned about the possibility of mistakes, such as AI hallucinations. Indeed, the possibility of such mistakes may not be sufficient for claiming that Generative AI cannot generate knowledge. Humans often make mistakes with regards to the information they produce, but this does not disqualify them from counting as sources of knowledge. Being an epistemic agent does not require been absolutely reliable. If so, one way to assess the merits of the hypothesis that Generative AI generates knowledge is to compare its epistemic reliability to the reliability of the average human knower. If AI reliability is close to human reliability then the hypothesis that Generative AI can generate knowledge still stands—though it might well be false for other reasons, such as the lack of genuine metacognitive abilities.

SERRC Social Epistemology Review & Reply Collective
14 (11): 54–64. 2025.
https://wp.me/p1Bfg0-aC6

concerns AI's apparent lack of intentions, which are typically regarded as necessary for a speech act to qualify as testimony.[2] Ori Freiman (2024) observes, for example, that:

> The 'paradigm case of testimony – the intentional transfer of a belief from one agent to another' (Pritchard 2004) […] represents the common view about intentions within testimonial theories of knowledge and justification. Under the assumption that technologies do not have intentions, conversational AIs cannot give testimony – in principle. Therefore, the concept of testimony is not suitable for the analysis of acquiring knowledge from technologies (480).

Similarly, He and Yang (2025) write:

> The difference between testifying and merely presenting a message is that only the former is an intentional act of assertion. Consequentially, although AI systems do not have a capacity to act intentionally, they can be programmed to act according to the designer's intentions. This feature leaves room for including AI systems in the group of possible testifiers of knowledge. As these studies suggest, if the performance of a machine resembles assertoric acts of human agents, the recipient of machine-generated text would and should treat the machine's act as intentionally testifying (section 5.4).

These points suggest that information delivered by AI can, at best, be regarded only *metaphorically* as a form of testimony—a shorthand to convey the idea that human designers have delegated their epistemic intentions to the machine. The machine does not really testify by transmitting knowledge; rather—much like other technologies—it only delivers information, which can be subsequently used to acquire knowledge.

However, this assumption—that AI technologies serve merely as conduits for designers' intention to convey information—not only leaves the sense in which AI can testify somewhat loose, but also gives rise to a deeper problem concerning the allocation of epistemic responsibility. As Freiman (2024) notes:

> [N]ormative assessment is often reduced to the individuals and groups behind the technologies. For example, 'if a given computer yields information that turns out to be false, we will blame the programmer, or our use of the program, . . . but not the computer itself' (Goldberg 2012, 194). Indeed – in the era of AI, and especially with Generative AI, the accountability and responsibility for the truth and the consequences of the output are distributed among many hands (e.g. Slota et al. 2021). While a person can hold the testifier accountable, it is often impossible for a person

[2] Heersmink et al.'s (2024) reasons for denying the transmission of knowledge by AI are different. They argue that the data and algorithmic opacity of AI as well as its phenomenological and informational transparency prevent users from calibrating their trust effectively.

> to trace the accountability for the outputs of AI-systems (Alvarado 2023; Burrell 2016; Pasquale et al. 2015) (480-481).

When people, instruments, algorithms and other things affect the production of an epistemic outcome, the problem of assigning individual epistemic responsibility arises. This is the epistemic problem of 'many hands' (482).[3]

Thus, while these approaches to AI 'testimony' are intuitively plausible—framing the acquisition of knowledge from AI as the reception of information that reflects the intentions of the creators of AI—they ultimately fail to account for where epistemic responsibility lies. To avoid this difficulty, in the following sections, I propose an account of testimonial knowledge from Generative AI. I suggest that when distributed virtue reliabilism is combined with a suitably modified version of Greco's 'massively shared agency', it can yield a *literal* account of AI testimony—whereby AI transmits knowledge, not just information—and which justly allocates epistemic responsibility to both the whole socio-technical system that realises the Generative AI technology *and* the individuals that make executive decisions surrounding its design and development.

## Generative AI Systems as Epistemic Collaborations

Leaving aside, for the moment, the question of AI intentions and responsibility, I will here focus on whether Generative AI systems can be understood as epistemic collaborations capable of *generating* knowledge. In Palermos (2022a) I argue that certain groups, such as scientific research teams, often exemplify this phenomenon: their individual members *self-organise* and *self-regulate*, morphing into epistemic socio-technical systems capable of *generating*—reliably and responsibly—true beliefs. In other words, certain groups can collectively generate knowledge in a collaborative fashion.

For this to occur, members must integrate through continuous and bidirectional interactions—both with one another and with their tools or equipment. When this integration condition is met, an epistemic group agent emerges, encompassing all the interacting participants and their instruments. The knowledge and justification produced by such collaboration do not reside in the individuals separately or in their mere sum, but in the entirety of a corresponding *epistemic group agent*. The reason for claiming this is that the reliability and responsibility—i.e., the justification—of the true beliefs generated are *collective properties* that can be attributed only to this higher-level epistemic group agent.

The question, then, is whether the working of Generative AI justify the claim that they similarly qualify as epistemic collaborations that collectively generate knowledge.

Consider the case of OpenAI's ChatGPT. Examining how it is built and operates, ChatGPT appears to satisfy the criteria for a collaborative epistemic group agent. It is a large-scale

---

[3] As Freiman (2025, 482) further notes, according to Helen Nissenbaum (1996):
[W]here a mishap is the work of 'many hands', it may not be obvious who is to blame because frequently its most salient and immediate causal antecedents do not converge with its locus of decision-making. The conditions for blame, therefore, are not satisfied in a way normally satisfied when a single individual is held blameworthy for a harm (29).

14 (11): 54–64. 2025.
https://wp.me/p1Bfg0-aC6

socio-technical system, comprising its engineers, users, and the underlying hardware that embodies a vast neural network capable of processing prompts and delivering appropriate outputs.

This neural network consists of multiple layers of simulated nodes (artificial neurons), each connected to others with adjustable weights that determine how signals are propagated through the network. The final weighted structure of the neural net is not explicitly programmed by humans. Instead, it is *automatically* learned through extensive training. The first phase—called *pre-training*—involves exposing the model to massive amounts of text data (largely from the internet) and having it learn how to generate the next token (usually a word) in a sentence. This is done by letting the model produce predictions of what the next token in various sentences is supposed to be, compare it with the actual token in the data, calculate the error, and then adjust its weights according to its learning algorithm.[4] This produces the so-called *base model*.

This base model is then *fine-tuned*, often using methods like *reinforcement learning* from human feedback. During this phase, human evaluators rate different outputs, helping the system to learn how to generate responses that are more useful, accurate, and aligned with human values. Moreover, Chat-GPT's structure continues to evolve by means of aggregated user content, feedback, and usage patterns, which guide future updates and system refinements. As OpenAI explains on its website:

> When you share your content with us, it helps our models become more accurate and better at solving your specific problems and it also helps improve their general capabilities and safety. […] ChatGPT, for instance, improves by further training on the conversations people have with it, unless you opt out (OpenAI n.d.).

Even if you've opted out of training, you can still choose to provide feedback to us about your interactions with our products (for instance, by selecting thumbs up or thumbs down on a model response). If you choose to provide feedback, the entire conversation associated with that feedback may be used to train our models. (OpenAI n.d.)

The above demonstrates that ChatGPT's outputs results from an intricate web of ongoing, bidirectional interactions between numerous individuals and the associated equipment. The causal loops involved, can be summarised as follows:

*During Pre-Training*:

- Between OpenAI engineers while designing and selecting training algorithms;

- Between OpenAI engineers while curating training data;

[4] Though it should be noted that reducing Large Language Models to mere 'next token predictors' may not do justice to their true nature. See, for example, Grzankowski et al. (2024).

- Between internet users and the neural network through the data internet users generate and make publicly available.

*During Fine-Tuning and Ongoing Evolution*:

- Between OpenAI engineers and the neural net, as the former provide evaluative feedback to the latter;

- Between the neural network, users, and OpenAI engineers, where user interactions provide data and feedback that engineers curate and use to refine the model further.

Given these multiple interconnected layers of interactions, it is tempting to treat ChatGPT as a distinctive kind of collaborative group agent, whose epistemic performance is distributed across its various components.

However, this characterisation introduces a troubling implication. As noted earlier, in epistemic collaborations, the associated epistemic responsibility is an emergent property—it arises from the ongoing interactions between the various components of the socio-technical system and it can only be attributed to the system as a whole. This means that epistemic responsibility, in such cases, belongs not to any individual member or their sum but to the group agent—in this instance ChatGPT—*as a whole*. The reason this is concerning is that it leads to a problem that is similar to the 'problem of many hands'—what I call 'the problem of the invisible hand.'

While the well-known 'problem of many hands' arises when individual responsibility is *difficult* to trace among many contributors, due to the sheer number of the individuals involved in the performance of the epistemic action, the 'problem of the invisible hand' is deeper: it concerns that *inapplicability* of individual responsibility altogether. If responsibility in epistemic collaborations is genuinely emergent, then, metaphysically speaking, attributing epistemic responsibility to individuals may no longer make sense. *Unless*, of course, there is some other rationale for attributing responsibility to individual members of epistemic collaborations.

Like the 'problem of many hands,' the 'problem of the invisible hand' is deeply troubling—especially when it concerns technologies that influence millions worldwide. When ChatGPT produces epistemic errors, it seems unsatisfactory to simply assign blame to an abstract entity that stands above the network of engineers, algorithms, and users of the model. Intuitively, there must also be a well-defined target—some individual or set of individuals—of individual epistemic liability.

Fortunately, this concern can be mitigated. The solution relies on demonstrating that while ChatGPT qualifies as a collaborative epistemic group agent, it represents a special instance of one. In the next sections, I draw on Greco's account of 'massively shared agency' to clarify this claim.

14 (11): 54–64. 2025.
https://wp.me/p1Bfg0-aC6

**AI Intentions**

As noted above, the standard assumption in the epistemology of testimony holds that an entity can act as testifier only if it can manifest intentions—specifically, the intention of transmitting information to a hearer. In his paper, Greco (2025) explains how large-scale institutions might satisfy this condition by drawing on elements from Bratman's account of shared agency.

According to Bratman (e.g., 1999), two or more individuals manifest shared agency, when they possess common knowledge of their *interlocking* shared intentions to perform a certain act together. Each participant must intend to perform the act *with* the others, who must also intend to do so, and this shared intention must be known by everyone involved. Additionally, they must have *meshing subplans*—subplans that enable coordination toward the shared goal. Finally, they must also be mutually responsive in both action and intention. That is, they must be capable and willing to respond to changes in others' intentions and actions while performing the act so as to *remain* coordinated throughout.

While promising, Greco observes that Bratman's account does not scale to large institutions, for at least three interrelated reasons:

> 1) *The Problem of Authority*: Bratman's framework presupposes equality between participants, yet large institutions often depend on hierarchical structures (Shapiro 2014).
>
> 2) *The Problem of Alienation*: Many institutional participants may not share intentions toward the overall goal; they might simply be there to do their part of the job (i.e., complete their sub-task) without intending to coordinate with others (Shapiro 2014).
>
> 3) *The Problem of Opacity*: In large institutions, participants often work in isolation—especially across departments—and thus cannot know each other's intentions or remain mutually responsive (Shieber 2022).

To address these challenges for scaling up his account, Bratman (2022) introduces two important modifications. First, he proposes emphasising the role of 'social rules' (Hart, 1961): publicly endorsed and observed behavioural norms that guide action and justify criticism, demands, and coordination within social structures. These rules—often imposed top-down within hierarchical systems—promote the institution's overarching goals by shaping the intentions and actions of participants, ensuring that every member will act accordingly, even when, due to the problems of alienation and opacity, not all members directly meet Bratman's criteria for shared agency.

Second, Bratman adopts a 'kernel/penumbra' model of institutional agency. According to this distinction, only the *kernel*—a subset of participants, typically at higher hierarchical levels— directly satisfies his conditions for shared agency, while the *penumbra* consists of members whose participation is coordinated indirectly through social rules and authority structures.

Greco, however, is critical of this sharp division. While he appreciates Bratman's appeal to Hart's notion of social rules, Greco worries that emphasising the kernel undermines the idea of *massively shared* agency. If the kernel is too small, we are no longer dealing with a *massively* shared agency, but with a *standard* kind of shared agency—that of the kernel. And if the kernel is too large, the problem of opacity resurfaces.

To overcome this tension, Greco proposes shifting the focus away from the need for explicit shared intentions and directing it instead toward Hart's notion of 'social rules.' His central claim is that instead of invoking direct shared intentions, instances of massively shared agency can be explained through rule-based coordination. The first step in his argument draws on Shapiro's (2014) insight that authority functions as a *mesh-creating mechanism*. The larger a group becomes, the more costly and impractical it is for its members to directly coordinate their subplans. Thus, in such cases, it is rational and often preferable for participants to submit to some authority that will guide their team effort. But, Greco notes, just as authority can play this coordinative role, so too can social rules—including formal institutional rules, protocols, and policies, as well as informal norms or even etiquette. If that's correct—if social rules can likewise serve as coordinating forces—it is possible to downplay the importance of authority and focus instead on the coordinative power of social rules:

> One consequence of the resulting account is that it alleviates the need for Bratman's distinction between kernel and penumbra participants in massively shared agency. When someone accepts some set of institutional rules, or internalizes some set of social norms, or engages in some norm-governed social practice, they thereby become participants in shared agency (7).
>
> In effect, the proposal is that social rules and practices in general function so as to offload the hallmarks of shared agency onto the social environment. That is, the requirements of shared intentions, common understanding, and mutual responsiveness are satisfied in virtue of adopting institutional policies, internalizing social norms, and participating in social practices (7).

## AI Responsibility

While I agree with Greco that social rules can play this coordinating role, I wish to remain neutral on whether the kernel/penumbra distinction is necessary for explaining massively shared agency. Although the presence of a Bratman-style kernel with shared intentions may not be necessary for *massively shared* agency to occur, in most hierarchical institutions—including AI systems such as ChatGPT—there typically exists a decision-making kernel. Recognising this feature is crucial, as it provides a way to address the 'problem of the invisible hand' in an intuitive way.

While many social rules arise informally—emerging bottom-up through repeated public patterns of behaviour that gradually become entrenched as norms—such rules are rarely the only ones at play. In hierarchical institutions, there are also top-down rules: formal protocols

SERRC
Social Epistemology
Review & Reply Collective
14 (11): 54–64. 2025.
https://wp.me/p1Bfg0-aC6

organisational workflows, distributions of labour *embodied* in the institution's socio-technical structure, all of which guide and constrain participants' activities. These top-down rules of conduct are often particularly salient in the case of complex socio-technical systems.

For instance, in the development and operation of ChatGPT, decisions concerning the choice and design of learning algorithms, the collection and curation of training data, the fine-tuning of the model, and choices regarding workflows necessary for coordinating this widely distributed effort are all presumably determined by OpenAI's teams of engineers and managers. Furthermore, although user-generated content, feedback, and usage patterns influence ChatGPT's ongoing updates and refinements, these fine-tuning processes are likely mediated and controlled by OpenAI's protocols and design features of its platform, all of which are, again, most likely centrally decided by OpenAI's employees.

Therefore, if something like the above is true of OpenAI's organisational structure, when something goes wrong epistemically—when the model produces unreliable or misleading outputs—it seems appropriate to direct responsibility not only toward the entire distributed system *qua* an epistemic group agent, but also toward this kernel-like subset of decision-makes that is likely involved in the design of the organisation's structure. Whether further responsibility can be directed to specific individuals within the company—rather than the *sum* of individuals making up this kernel—will often depend on the specifics of its internal decision-making process and the nature of the epistemic shortcoming in question. Nevertheless, the upshot is that, most likely, there *is* a specific subset of the epistemic collaboration—its *operational backbone*, so to speak—which can be held responsible at the individual level for the group's overall performance. Importantly, staying cognizant of this point offers a way out of the 'problem of the invisible hand.' Specifically, it provides a principled way to locate epistemic responsibility at the individual level without collapsing it entirely into the emergent properties of the system as a whole.

Granted, when ChatGPT fails epistemically, there is a sense in which it is itself epistemically responsible—after all its integrated operation fell short of preventing the generation of the false output. This is something that is always worth noting, because, to prevent similar problems in the future, solutions should also be directed to addressing *structural* issues concerning the integration of the system's distributed structure.[5] At the same time, however, in hierarchical socio-technical systems like ChatGPT, most likely, the shape of this integrated structure almost entirely depends on a specific set of individuals making executive decisions. Thus, rather than directing all epistemic liability to the group agent itself, a significant amount of individual responsibility likely also lies with and should be attributed to individuals comprising the corresponding kernel of decision-making.[6] In other words, both levels of responsibility—the emergent and the individual—must be acknowledged.

This conclusion supports the intuitive view that, while epistemic responsibility in distributed AI systems may partly emerge collectively, it also remains anchored in identifiable human agents and organizational structures. It is for this reason that I resist dismissing the

[5] See also Palermos (2022b).
[6] See also Palermos (2022b).

importance of Bratman-style kernels in epistemic collaborations that exhibit hierarchical features. Although some large scientific collaborations can function without a kernel,[7] private and for-profit endeavors such as ChatGPT are likely to heavily rely on an explicitly hierarchical structure designed and run by a kernel of decision-making. The importance of this kernel in directing our attributions of responsibility regarding the transmission of knowledge by such systems (or their failure to do so) should not be overshadowed by their otherwise densely collaborative nature and features of massively shared agency.

When Generative AIs like ChatGPT fail to generate and transmit knowledge, epistemic responsibility likely lies not only with the emergent system but also, and more directly, with the company—or at least its executive and technical leadership—that was likely involved in its design, structure, and deployment.

## References


Alvarado, Ramón. 2023. "AI as an Epistemic Technology." *Science and Engineering Ethics* 29 (32). doi: 10.1007/s11948-023-00451-3.

Bratman, Michael. 2022. *Shared and Institutional Agency: Toward a Planning Theory of Human Practical Organization.* United Kingdom: Oxford University Press.

Bratman, Michael. 1999. *Faces of Intention: Selected Essays on Intention and Agency*. United Kingdom: Cambridge University Press.

Burrell, Jenna. 2016. "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms." *Big Data & Society* 3 (1). doi: 10.1177/2053951715622512.

Freiman, Ori. 2024. "AI-Testimony, Conversational AIs and Our Anthropocentric Theory of Testimony, *Social Epistemology* 38 (4): 476–490. doi: 10.1080/ 02691728.2024.2316622.

Goldberg, Sanford C. 2012. "Epistemic Extendedness, Testimony, and the Epistemology of Instrument-Based Belief." *Philosophical Explorations: An International Journal for the Philosophy of Mind and Action* 15 (2): 181–197. doi: 10.1080/ 13869795.2012.670719.

Grzankowski, Alex, Stephen M. Downes and Partick Forber. 2025. "LLMs are not Just Next Token Predictors." *Inquiry* 1–11. doi: 10.1080/0020174X.2024.2446240.

Greco, John. 2025. "The Transmission of Knowledge via Large-Scale Technology: A Shared Agency Account." *Social Epistemology* 1–14. doi: 10.1080/ 02691728.2025.2463054.

Hart, H. L. A. 1961. *The Concept of Law*. Oxford University Press.

He, Jinhua and Chen Yang. 2025. "Testimony by LLMs." *AI & SOCIETY* 40: 6201–6213.

Heersmink, Richard, Barend de Rooij, María Jimena Clavel Vázquez and Matteo Colombo. 2024. "A Phenomenology and Epistemology of Large Language Models: Transparency, Trust, and Trustworthiness." *Ethics and Information Technology* 26 (41): 1–15.

Knorr Cetina, Karin. 1999. *Epistemic Cultures: How the Sciences Make Knowledge*. Harvard University Press.

Nissenbaum, Helen. 1996. "Accountability in a Computerized Society." *Science and Engineering Ethics* 2 (1): 25–42. doi: 10.1007/BF02639315 .


[7] See Knorr Cetina (1999).

SERRC Social Epistemology Review & Reply Collective
14 (11): 54–64. 2025.
https://wp.me/p1Bfg0-aC6

OpenAI. 2025. "How Your Data is Used to Improve Model Performance." Retrieved November 1. https://help.openai.com/en/articles/5722486-how-your-data-is-used-to-improve-model-performance.

Palermos, S. Orestis. 2022a. "Epistemic Collaborations: Distributed Cognition and Virtue Reliabilism." *Erkenntnis* 87 1481–1500. doi: 10.1007/s10670-020-00258-9.

Palermos, S. Orestis. 2022b. "Responsibility in Epistemic Collaborations: Is it Me, is it the Group or are We All to Blame?" *Philosophical Issues* 32 (1): 335–350.

Pasquale, Frank. 2015. *The Black Box Society: The Secrgorithms That Control Money and Information*. Cambridge, MA: Harvard University Press.

Pritchard, Duncan. 2004. "The Epistemology of Testimony." *Philosophical Issues* 14 (1): 326–348. doi: 10.1111/j.1533- 6077.2004.00033.x.

Shapiro, Scott. 2014. "Massively Shared Agency." In *Rational and Social Agency: The Philosophy of Michael Bratman*, edited by M. Vargas and G. Yaffe, 257–293. United Kingdom: Oxford University Press.

Shieber, Joseph. 2022. "Challenging the Ability Intuition: From Personal to Extended to Distributed Belief-Forming Processes." *Philosophical Issues* 32 (1): 351–366. doi: 10.1111/phis.12222

Slota, Stephen C., Kenneth R. Fleischmann, Sherri Greenberg, Nitin Verma, Brenna Cummings, Lan Li and Chris Shenefiel. 2021. "Many Hands Make Many Fingers to Point: Challenges in Creating Accountable AI." *AI & Society* 38 (4): 1–13. doi: 10.1007/ s00146-021-01302-0.