AI and the scientist

On the fracture of epistemic authority

Kristina Khutsishvili

1. Introduction

The figure of the scientist has long occupied a distinctive place in the epistemic order. From the Enlightenment onward, the scientist was not merely a skilled technician but a disciplined inquirer – a producer and guardian of knowledge – whose legitimacy rested on epistemic virtues such as curiosity, integrity, and humility, as well as on empirical rigour. Today, however, this position appears less secure. The systems we now refer to as 'artificial intelligence', and more concretely large language models (LLMs), produce outputs which, in certain domains, rival human work.

The figure of the scientist is thus threatened with displacement from the centre of knowledge creation. This is not simply a technical redistribution of tasks but a deeper reconfiguration of epistemic authority. The scientist now confronts a figure that resembles an autonomous author: a system generating hypotheses and even theoretical frameworks in ways that appear – symbolically, though not ontologically – independent of human agency. Whether such independence is real or only perceived matters greatly, yet the appearance of autonomy already unsettles established epistemic hierarchies. The scientist risks becoming a steward of results whose genesis lies elsewhere.

Scientific authority has depended on the human capacity to judge, to weigh, to distinguish. What happens when these gestures can be mimicked – and when the mimicry becomes epistemically consequential?

2. Epistemic authority

Charles Taylor has argued that identity is dialogical – formed "always in dialogue with, sometimes in struggle against the things our significant others want to see in us" (Taylor, 1992: 32–33). The identity of the scientist has been no exception. For centuries, scientists have been recognised by peers, institutions, and the public as authoritative knowers capable of exercising inquiry. Their authority rested on a shared understanding that

inquiry relies on a human capacity to weigh, distinguish, and reflect, and that scientific results are answerable to this capacity. Epistemic authority was thus tied not only to the results of inquiry itself, but to the recognisable exercise of judgement that produced them.

When LLMs generate literature reviews, propose theoretical frameworks, or conduct statistical analysis, something in this configuration shifts. These systems are not simply tools in the deterministic sense in which a calculator assists with arithmetic. They increasingly appear as quasi-autonomous epistemic agents: they draft arguments and recombine literatures in ways that are presented, and often perceived, as if they originated in a unitary 'authorial' source. In that sense, LLMs may extend the scientist's reach but they also may be positioned as rivals operating within the same epistemic space.

It is true that humans remain formally responsible: they prompt, read, analyse, correct, and sign their names under the final piece. However, the phenomenology of authorship and authority changes when a substantial part of the argumentative or narrative structure is generated by a system whose internal operations cannot be traced in familiar justificatory terms. The scientist is no longer the primary source of authoritative conclusions and risks becoming a steward of results whose genesis lies elsewhere. The site of authority moves from the act of judging to the act of approving.

In Taylor's terms, this is a risk of misrecognition. Authority has always involved recognition: others see the scientist as a person whose reasons and methods warrant trust. If, however, the 'producer' of the result is a system, the scientist may instead be recognised as the one who 'checked' or 'curated' something rather than 'produced'. The role is reframed, the position is fractured, the identity is unsettled. The scientist's self-understanding as an agent of inquiry becomes intertwined with a system that occupies the generative space of argument.

For researchers in the history, philosophy, and sociology of science (HPSS), this displacement carries particular weight. The discipline does not simply produce facts but interprets the conditions under which facts become meaningful; it investigates how concepts, methods, and institutions confer authority on some claims and not on others (Kitcher, 1993; Kuhn, 2012; Latour, 1987). These interpretive and critical approaches are precisely those that advanced systems can now mimic – producing narratives or arguments with surface plausibility that can obscure the absence of judgement. LLMs excel at re-stating available positions, weaving them into apparently coherent wholes, and generating what looks like 'critical reflection' by combining familiar criticisms.

Thus, there is a double movement. On the one hand, HPSS scholars are exposed to a specific kind of displacement: working in domains where argumentative, narrative, and conceptual work – the very textures of the practice – can be imitated by systems that do not understand what they are doing. In that sense, they encounter AI not at the margins of their work but at its core. On the other hand, precisely because HPSS focuses on the conditions of meaning, there is a possibility that its epistemic relevance increases rather than diminishes. The more central AI systems become to scientific practice, the more urgent the question of how their outputs are authorised, integrated, resisted, or normalised. In other words, the same development that threatens HPSS with mimicry also calls it, perhaps uniquely, to analyse the reconfiguration of epistemic authority itself. Therefore, the question "Who do HPSS scientists become when the ground beneath them

shifts?" cannot be answered simply in terms of replacement or survival. The fracture of epistemic authority is at the same time a fracture of identity.

The novelty of AI-mediated inquiry creates a new demand within the sciences themselves: the need for epistemological orientation. Scientists cannot rely solely on technical expertise when the fundamental processes shaping research no longer fit familiar methodological categories. Questions of what counts as an explanation, how evidence is approached, or how authority is redistributed are no longer abstract philosophical problems but practical conditions of scientific work. Not every discipline is equipped to address these questions from within; epistemic reflexivity is unevenly distributed. HPSS scholars, whose work has long examined methods, frameworks, and justificatory practices, may therefore become indispensable sources of orientation for their non-HPSS colleagues. Their expertise does not only describe the changing epistemic landscape but also provides the vocabulary and ability for navigating it. In this sense, AI does not only challenge HPSS, but it also makes its insights newly necessary.

3. The scientist as member of an imagined community

Benedict Anderson's (2006) metaphor of "imagined communities" is also suitable for scientific disciplines: they are sustained not predominantly by face-to-face contact but by shared languages, symbols, and rituals. Belonging depends on mutual recognition; to be counted as part of the community means having one's contributions integrated into its ongoing dialogue. Scientific communities define themselves through journals, conferences, textbooks, and informal networks – an entire symbolic system in which membership is conferred and denied.

At first sight, this might suggest that scientists retain complete control over the boundaries of their community. If mutual recognition is the criterion, one might argue that AI systems can be kept outside simply by refusing to cite AI-generated work, by not attributing authorship to non-humans, and by normatively enforcing human responsibility in publication and review. In this view, the imagined community is guarded by its members; they decide who counts as a knower, and under which conditions.

However, this picture underestimates how indirect the mechanisms of inclusion are. LLMs disrupt the fabric of the scientific imaginary not only when they are explicitly named, but also when their outputs circulate through human names, embedded in texts that pass peer review, enter databases, and are cited as if they were straightforward products of human inquiry. If a scientist relies heavily on AI-generated literature reviews, the resulting article may formally satisfy all existing norms of authorship while being largely materially shaped by a system that is not a member of the community and cannot be held to account.

The imagined community begins to shift not when it officially admits non-human members, but when its criteria for recognising human members no longer presuppose that the work of judgement and composition was carried out by those members themselves. The line between internal and external contributions becomes blurred. Belonging is shared – without explicit consent – with entities alien to the processes of recognition, which nevertheless shape what is recognised.

The HPSS area of studies is particularly exposed to this development. The imagined community is compact, and its cohesion rests on slow, reflective work. The incorporation of LLMs into these domains threatens to conflate the product of interpretive intellectual work with an accumulation of textual patterns. A careful historical analysis or a nuanced conceptual distinction may be seen as 'just another kind of text' that a system could plausibly generate. In the long run, this may alter how marginal or central such work appears to be within broader scientific imaginaries that increasingly equate value with speed and volume of output.

4. Identity, belonging, and (mis)recognition

Identities are never static; they are forged both through belonging and through refusal – through 'who we are' and 'who we are not'. Jean-Luc Nancy captures this complexity when he writes that identity "is not a figure", "identity is something more subtle, more delicate, and more evasive", and "its force lies in displacing, in changing figures" (Nancy, 2015: 10–11). The scientist's identity has traditionally been defined not only by what it is – a disciplined inquirer – but also by what it is not: speculation, baseless claim, the work of a dilettante. The 'non-scientific' has been a constitutive outside for the self-image of science.

This distinction becomes harder to maintain when LLMs generate work that appears indistinguishable from human scholarship. The boundary between the 'scientific' and the 'non-scientific' becomes uncertain. Without that line, the grounds for recognition shift, and with them, the self-understanding of the scientist. Gönül Pultar's reflections on rapid structural change that brings "instability, distortion of meaning, and volatile identities" are relevant here (Pultar, 2014: 155). When a scientist's contribution is reduced to approving or disapproving the outputs of systems whose operations are beyond human comprehension, the possibility of this scientist's unique epistemic authority diminishes. Authority that rests on results alone – without the shared understanding of how those results came to be – can no longer be sustained through the same dialogical relation that underpins recognition. For HPSS scholars – whose expertise lies in critically interrogating the conditions of knowledge 'production' – this transformation challenges the very core of their professional identity. Yet it is precisely here that their irreplaceability can also be articulated.

If epistemic authority is being fractured, the loss is not only in how scientists are recognised by others, but also in how they recognise each other and themselves. As Taylor (1992) reminds us, identity is formed in dialogue, and without a community capable of mutual recognition, the professional self is affected. Anderson (2006) would add that this community has always been, to some extent, imagined, and LLMs unsettle these shared imaginaries. For scientists, and especially for those in HPSS, the task is unlikely to be protecting a role in the old sense, but nurturing the community's life: ensuring that its narratives continue to speak of curiosity, human inquiry, responsibility, and the bonds that make recognition possible.

In this light, the future of science will not be determined solely by the distribution of tasks between human and machine. Rather, it will depend on whether the imagined

community of inquirers can survive the arrival of a participant that cannot imagine itself. The answer will depend on our capacity to renew the stories we tell about what it means to know and to belong – in an age when knowing is no longer ours alone.[1]

References

Anderson B (2006) *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso.

Kitcher P (1993) *The Advancement of Science: Science without Legend, Objectivity without Illusions*. New York: Oxford University Press.

Kuhn TS (2012) *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Latour B (1987) *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge, MA: Harvard University Press

Nancy J-L (2015) *Identities: Fragments, Frankness*. New York: Fordham University Press.

Pultar G (ed.) (2014) *Imagined Identities: Identity Formation in the Age of Globalization*. Syracuse, NY: Syracuse University Press.

Taylor C (1992) 'The Politics of Recognition'. In: Gutmann A (ed) *Multiculturalism: Examining the Politics of Recognition*. Princeton: Princeton University Press, 25–73.

1 The author of this chapter declares the use of a large language model (LLM) to obtain critical feedback on ideas (as a form of intellectual sparring) and to assist with final copy-editing.