https://doi.org/10.46938/tv.2026.730

////// editorial //

AI IN SCIENCE: LESSONS FROM THE ONGOING REVOLUTION

This special issue of *Teorie vědy / Theory of Science* asks how artificial intelligence (AI) is transforming science. The question is no longer whether AI will affect scientific research, but rather how deeply it will alter scientific methods, epistemic practices, institutions, and the self-understanding of scientific inquiry. Computers have long been central to scientific work: they have transformed modelling, simulation, statistical analysis, visualisation, and large-scale computation. Yet the recent emergence of large language models (LLMs), foundation models, and agentic AI systems marks a watershed moment in this development, as AI systems no longer function only as instruments for calculation or data processing. They increasingly participate in the production, interpretation, and communication of knowledge.

This transformation is of direct relevance to philosophy of science, as it concerns, for instance, the nature of scientific discovery and the relation between explanation and prediction, while also bringing serious epistemic and social consequences, such as the changing division of cognitive labour between human researchers and artificial systems. It also raises practical questions concerning scientific institutions: how should scientists use AI responsibly; how should AI-generated results be evaluated; how should accountability and authorship be understood; and how can scientific communities preserve human expertise while taking advantage of powerful new technologies?

The contributions collected in this issue approach these questions from several complementary directions. They do not attempt to offer a complete overview of the AI transformation, which would be impossible in a single issue. Instead, they provide a set of focused perspectives on AI as a tool of discovery, as an educational and methodological instrument, as a technology reshaping historical research, and as a challenge to theories of meaning and communication.

Petr Jedlička's article, *Quo Vadis? Scientific Discovery in the Age of Artificial Intelligence*, offers the broadest perspective. It examines the growing role of AI in scientific discovery, beginning with the rapid increase in AI capabilities and the diffusion of AI across scientific disciplines. Jedlička documents this shift through the sharp rise in scientific publications that either employ AI methods or study AI itself, particularly after the public

release of ChatGPT and other widely accessible LLMs in 2022. AI, he argues, has become an increasingly important participant in scientific practice.

A central contribution of the article is its typology of AI systems currently used in research. Jedlička distinguishes specialised scientific AI, scientific AI assistants, scientific AI agents, and hybrid AI experimental systems that combine digital computation with physical experimentation in laboratory settings. This typology classifies AI systems not primarily by their technical architecture, but by their role within scientific work. Some systems remain narrowly specialised tools; others assist researchers in writing, coding, searching, or data analysis; still others perform more autonomous tasks, such as hypothesis generation, experiment design, evaluation, and iterative refinement. Hybrid systems go further by connecting AI reasoning with robotic experimentation and self-driving laboratories.

Jedlička then offers a selective overview of AI-assisted achievements in mathematics and computer science, physics, chemistry, the life sciences, and the behavioural and social sciences. The examples range from AI-assisted work on mathematical problems and algorithmic discovery to protein-structure prediction, drug-discovery workflows, and systems capable of analysing or influencing human behaviour. The article is careful, however, not to present these developments as an uncomplicated triumph, since it points to well-known drawbacks, such as hallucinations, epistemic opacity, limited robustness, dependence on human oversight, and the difficulty of evaluating machine-generated outputs. It also raises wider institutional concerns. If AI systems increasingly perform tasks traditionally assigned to doctoral students and early-career scientists, the training pipeline through which scientific judgement, tacit knowledge, and disciplinary norms are reproduced may be weakened.

Jedlička's article also addresses the dual-use character of AI-driven science. Like earlier forms of scientific and technological knowledge, AI may generate extraordinary benefits while also amplifying risks. This applies not only to current problems such as unreliable outputs, fabricated citations, biosecurity, cybersecurity, and misuse, but also to longer-term alignment concerns associated with autonomy or self-improvement, which could lead to disastrous consequences if the delicate balance of human-AI relations and control were to tip in favour of artificial systems.

The issue then turns from advanced research systems to one of the preconditions of science itself: curiosity. Roman Lyach and his coauthors, in their contribution *How Can AI Help to Research and Potentially Enhance the Curiosity of Elementary School Children?* show that the AI transformation

of science begins not only in laboratories or research institutes, but also in educational environments where future scientific dispositions are formed. They examine how AI might support the study and cultivation of curiosity in educational settings. Their focus is methodological. They present an AI-supported tool designed to analyse children's questions, classify them, provide adaptive feedback, and offer teachers insights into pupils' engagement and curiosity.

The article treats question-asking as an observable expression of epistemic curiosity. Traditional instruments for measuring curiosity, such as self-report questionnaires, are often limited, especially when applied to children – they may capture only static snapshots, depend on self-assessment, and be vulnerable to social desirability bias. By contrast, the AI-supported tool proposed by Lyach et al. offers a different approach: it integrates curiosity measurement into classroom practice by analysing the questions students ask in a concrete learning situation. At the same time, it not only observes curiosity – it is also intended to stimulate it. This dual role illustrates how AI can change not only the measurement of educational phenomena, but also the phenomena being measured. The authors also address ethical issues, including privacy, the responsible use of student data, and the risk of bias in AI-generated feedback.

Marie Michlová's survey article, *Current Role of AI in the Historical Sciences and Research*, extends the issue into the humanities. It shows that AI's importance is not confined to the natural sciences or to data-intensive experimental fields. In the historical sciences, AI is already impacting access to sources, methods of analysis, and practices of scholarly writing. Michlová maps selected AI-based projects, programs, and applications used in historical research, focusing especially on developments from 2022 to mid-2025, while also including earlier projects of methodological significance.

The article identifies four principal areas in which AI is currently affecting historical research: digitization and recognition of sources, extraction of structured data from texts, heuristic support and hypothesis formulation, and academic writing. This article thus demonstrates that AI enters historical work at several levels: it can assist with optical character recognition (OCR) and handwritten text recognition (HTR), the transcription of printed and handwritten materials, the processing of oral-history recordings, the classification of maps and visual sources, and the extraction of data from large corpora. It can also support the formulation of research questions, the identification of patterns, and the reconstruction of damaged or incomplete sources. Michlová also discusses concrete examples from Czech and inter-

national contexts, for example systems such as Ithaca and Aeneas, which support the reconstruction and interpretation of ancient inscriptions. Such examples show that current AI systems in the historical sciences often function less as autonomous discoverers than as tools of access, reconstruction, and interpretation.

At the same time, Michlová stresses that historical research remains dependent on human source criticism, contextual interpretation, and methodological judgement. AI systems may hallucinate, reproduce biases, misinterpret context, or privilege dominant narratives contained in training data and archival structures. For the historical sciences, this is not a peripheral problem. These questions are central to the discipline, and the use of AI therefore requires transparency, AI literacy, and sustained critical oversight.

Finally, Vít Gvoždiak's *Pragmatics beyond Humans: Meaning, Communication, and LLMs* approaches AI from the perspective of language and communication. Gvoždiak asks how LLMs affect our conceptual understanding of meaning itself. His article argues that the rise of LLMs in communicative contexts requires pragmatics to reconsider some of its basic assumptions. Pragmatics should not be treated simply as a subordinate layer added after syntax and semantics, but as an interface through which language has social functions.

Gvoždiak argues that LLMs destabilize established hierarchies of meaning and require a framework better suited to communication involving artificial systems. He therefore turns to Human–Machine Communication (HMC) as a more appropriate perspective. The article also criticizes substitutionalism: the tendency to treat LLMs as substitutes for human speakers, or to evaluate them through assumptions derived from ordinary human communication. Such assumptions can lead to anthropomorphic bias and obscure the actual role of human users in constructing the communicative situation.

One of the suggestive concepts introduced in the paper is "context frustration." LLMs can process large amounts of contextual input, yet this does not mean that they share context in the human sense: the user is often compelled to construct or repair the pragmatic conditions of communication both for the model and for themselves. This point has wider implications for AI in science. Scientific use of LLMs depends not only on their factual accuracy, but also on the possibility of establishing reliable communicative and epistemic contexts. If human researchers cannot determine what kind of context the system is using, how it interprets instructions, or why

it produces a given answer, then this lack of transparency also becomes an epistemic problem.

The contributions in this issue show that AI is not a single phenomenon with a single impact on science – it appears in the sciences in many roles, and it affects discovery, education, interpretation, and authorship. The issue also shows that the current transformation cannot be captured by simple optimism or simple alarm. AI systems can accelerate research, open new methodological possibilities, and make previously inaccessible materials available for analysis. At the same time, they introduce new forms of bias, error, epistemic opacity, dependence, and institutional vulnerability. For philosophy of science, this is an important moment, as AI compels us to revisit questions that have always been central to the field: What counts as a discovery? What kind of understanding is required for explanation? How should evidence produced by AI systems be evaluated? What is the role of human judgement in scientific practice? And how should scientific institutions govern technologies that are both powerful and difficult to control?

The present issue offers only a partial view of an ongoing revolution. Yet because the transformation is unfinished, such philosophical reflection is urgently needed. The future of AI in science will not be determined by technical capability alone; it will also depend on how these systems are integrated into scientific life and on how existing epistemic standards and institutional environments are maintained.

Petr O. Jedlička