Title
Remaking the landscape of scientific expertise: The emergent research program of total automation in science

Journal
Big Data & Society, 13(3)

ISSN
2053-9517

Authors
Peterson, David
Koch, Bernard
Natarajan, Ramya
et al.

Publication Date
2026-09-01

DOI
10.1177/20539517261473358

Original research article

Big Data & Society
July–September: 1–14
© The Author(s) 2026
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20539517261473358
journals.sagepub.com/home/bds

Remaking the landscape of scientific expertise: The emergent research program of total automation in science

David Peterson[1], **Bernard Koch**[2], **Ramya Natarajan**[3] **and Aaron Panofsky**[4]

Abstract

This article examines an emergent research program aimed at the complete automation of science, analyzing how various automation projects funded by DARPA and others are working to replicate traditionally human scientific competencies. Rather than directly attempting to create fully autonomous robot scientists, these projects break down complex scientific work into discrete, automatable "tasks." While individual projects often preserve some role for human scientists, collectively they reveal a comprehensive blueprint for scientific automation. Through interviews with 26 DARPA contractors, program managers, and others involved in scientific automation, the authors demonstrate how automation efforts target three key domains of human scientific expertise: insight (hypothesis generation and pattern recognition), embodied knowledge (experimental procedures and techniques), and judgment (evaluation of scientific claims). Instead of simply trying to replicate human capabilities, these initiatives are actively reshaping scientific practice to be more amenable to machine intervention through initiatives like standardized ontologies and machine-readable data formats. Although full automation faces significant technical and institutional challenges, this emergent program portends major changes in scientific practice and careers while raising important questions about the future role of human knowledge in scientific discovery. The article contributes to ongoing debates about artificial intelligence in science while challenging existing sociological accounts about purportedly uniquely human scientific abilities.

Keywords

Automation of science, robot scientist, DARPA, embodied knowledge, insight, judgment

Introduction

Recent developments in artificial intelligence and robotics have fueled new hopes that one of the pinnacles of human achievement—science—could be done faster, more accurately, and in hitherto unforeseen ways by machines. This project involves many disparate technologies and goals, yet they share the ambition of automating what have been considered uniquely human areas of scientific competence like experimental "touch," insight, and collective judgment.

In an article detailing the evolution of the relationship between the "domain" knowledge of experts and the purportedly "domain general" knowledge of software engineers and AI researchers, Ribes et al. (2019) point out that a chief selling point of automation has been the idea that the integration between computer scientists and domain scientists will be a rising tide that will lift all boats. Computer scientists extend their reach while domain scientists harness new computational powers. Yet, there is a different rising tide metaphor that presents a more pessimistic view of the relationship between automation and domain experts. Futurist Hans Moravec proposed a metaphorical "landscape of human competence" in which all human skills are represented as a mountainous topography:

> Advancing computer performance is like water slowly flooding the landscape. A half century ago it began to drown the

[1]Department of Sociology, Purdue University, Lafayette, IN, USA
[2]Department of Sociology, University of Chicago, Chicago, IL, USA
[3]Burnett School of Medicine, Texas Christian University, Fort Worth, TX, USA
[4]Institute for Society and Genetics, University of California, Los Angeles, CA, USA

Corresponding author:
David Peterson, Department of Sociology, Purdue University, Stone Hall, 700 W State St, West Lafayette, IN 47907, USA.
Email: davidpeterson@purdue.edu

> lowlands, driving out human calculators and record clerks, but leaving most of us dry. Now the flood has reached the foothills, and our outposts there are contemplating retreat. We feel safe on our peaks, but, at the present rate, those too will be submerged within another half century. (Cited in Tegmark [2017])

Rather than being safe aboard their individual boats collectively enjoying rising waters, this metaphor frames the expansion of machine intelligence as a creeping threat to the authority of the scientist. If scientific practice is simply an assemblage of skills which can be isolated, mapped out, and automated, will any peak remain above water?

In this article, we suggest a reimagining of Moravec's metaphor. Along with creeping up the peaks of human competence, automation in science is also occurring by translating amorphous and ambiguous skills like embodied knowledge, insight, and judgment into narrowly defined tasks which can then be "hill-climbed" in ways that mirrors progress in the field of machine learning more broadly (Koch and Peterson, 2024). Instead of the daunting effort of building a robot scientist *in toto*, researchers are building systems that can do highly specified tasks like synthesizing literature on a specific form of cancer. The reduced complexity of the task provides a "doable problem" (Fujimura, 1992) for researchers engaged with automation projects while also providing clear metrics for comparison and progress.

Thus, rather than the static landscape of human skills and a creeping tide threatening to overtake those peaks as Moravec envisioned, the future is more likely to be characterized by a transformed landscape in which scientific work is reimagined and redesigned to benefit machine learning algorithms and robots.

Of course, machines are such an essential part of science that some have advocated for the term "technoscience" (e.g., Haraway, 1997; Latour, 1987). However, the use of automation has been accelerating in many areas including, most notably, high-throughput genomics (Arribas-Ayllon et al., 2019; Leonelli, 2016). And, across both the synchronic field of fields that compose science and across the diachronic phases of scientific research, roboticists and computer scientists are busy working on diverse projects seeking to automate everything from running experiments to synthesizing literature.

In isolation, these projects can be interpreted as attempts to improve science in limited ways. However, viewed together, they point toward the most complete attempt to reorient the sciences since they were disentangled from the church. In this article, we describe a set of automation projects animated by an intellectual ambition that is greater than the sum of its parts. Each individual project is narrowly focused on automating a specific research task, and its project leaders readily admit that certain other tasks remain areas of uniquely human competence. Yet, another set of projects seeks to automate precisely those tasks the first treats as inviolable. If we rise above the particulars of each project and imagine how they might be integrated, an emergent research program comes into view.

The intuition animating this emergent research program can be traced back to Francis Bacon's *Novum Organum*. His argument for inductive empiricism directed by a set of explicit rules helped popularize the concept of the Scientific Method. Bacon's formulation contains a radical idea that is only now reaching a point where it can be realized. Biologist Ahmed Alkhateeb (2017) has argued that Bacon's work "revealed an important hidden truth: the discovery process is inherently algorithmic …. If science is algorithmic, then it must have the potential for automation." If science can be carried out as a series of steps capable of explication, we should be able to invent a rule following system to carry it out.

Although only human beings have been capable of following the scientific method so far, proponents argue that the growth of machine intelligence offers a pathway to a new future. Increasingly, it is a future that is viewed as both desirable and inevitable. As a commentary on the use of artificial intelligence in science notes, "Technological innovations are penetrating all areas of science, making predominantly human activities a principal bottleneck in scientific progress while also making scientific advancement more subject to error and harder to reproduce" (Gil et al., 2014: 171). These advocates argue that the human mind is poorly suited to tasks like trying to find patterns in massive amounts of data yet artificial minds are perfectly suited to this task. In fact, computer scientist Judea Pearl has made the case that Big Data produces the fertile soil out of which a robot scientist could grow. Specifically, once we get a handle on how to effectively extract information from these huge datasets, we could, he writes, "create an 'artificial scientist.' This smart robot would discover yet unknown phenomena, find explanations to pending scientific dilemmas, design new experiments, and continually extract more causal knowledge from the environment" (Pearl, 2018: 2).[1]

In this article, we will not go into detail about automation efforts in any particular domain of science. Rather, our aim is to bring attention to a research program that, while dimly perceivable when we began this project, has now become a transformative force in science. Our term "emergent research program" owes a debt to Lakatos' work on "research programs," although the "emergent" aspect represents a significant revision.

Like a Lakatosian research program, it integrates a field through a "hard core" of theories. In this case, the theories involve scientific work being reducible to specific tasks which can then be asymptotically improved. However, unlike a Lakatosian research program, the researchers themselves may be unaware of or uninterested in the larger project. This is because individual researchers working on automation are mostly focused on narrowly defined tasks. This "taskification," which we will describe later, translates ambiguous concepts like "insight" into structured projects that can be measured. The actual, integrated program only becomes visible at a *meta* level. In our case, the emergent research program of automation of science is composed of many narrowly defined tasks but comes into relief when viewed, for instance, through the portfolio of programs funded by DARPA. Significantly, the pursuit of a robot scientist is still a fringe effort with few explicitly acknowledging this goal. Yet, taken together, the project becomes undeniable.

Because of the variety of projects involved, it should be clarified at the outset that by "automation" we are not referring to any specific technology. Our respondents' projects involve a variety of approaches including natural language processing, information retrieval, research synthesis, and robotics. Yet, they are unified by a shared goal of making inroads into aspects of scientific practice that only humans have been able to master. For the purposes of this argument, we will focus on three skills identified in the science studies literature. First, there is the domain of individual or small group cognition in which creativity and insight shape research projects (Ellis, 2011; Kounios and Beeman, 2009). Second, there are domains of individual physical performance—the embodied, tacit, or craft knowledge—that is often required to get experiments to work (Collins, 1992; Delamont and Atkinson, 2001; Polanyi, 2012). Lastly, there is the domain of collective intelligence which is concerned with how fields as a whole metabolize knowledge, deciding which claims to enshrine and which become buried in the onslaught of new research (Hull, 2001; Polanyi, 1962; Ylikoski, 1995). This categorization is not meant to be definitive. Rather, it represents a useful overlap between the operationalization of these challenges in these automation projects and existing literature in science studies about the less explicable aspects of scientific practice. The fact that these various projects cross boundaries necessitates a wide net and, thus, we will use the umbrella term "automators" to describe the computer scientists, bioinformaticians, and roboticists we interview.

What follows is not meant to criticize automation, though a certain queasiness seems unavoidable. In truth, each of these projects faces enormous challenges, and the whole may never indeed coalesce. Attempts to automate the more complex aspects of scientific cognition and practice are fueled as much by hope as actual progress. However, viewing these disparate projects as an emergent research program is necessary to understand some current trends in science policy and funding. Moreover, it presents significant challenges to accounts in the philosophy and sociology of science that emphasize the irreducibilty of human skills like tacit knowledge, insight, creativity, judgment, and culture in science. If scientific fields are amenable to being automated, it portends a future that few scholars are currently addressing. Additionally, we challenge historical discourse within AI research suggesting that the automation of science will be primarily driven by technological improvements. Instead, we argue that significant efforts are seeking to transform science to make machine intervention more successful similar to what Jordan (2018) has called "intelligent infrastructure."

Methods

Our article is empirically grounded in 26 interviews with DARPA contractors, program managers, and others involved with the automation of science. This work stems from an earlier project looking at the science reform movement (Peterson and Panofsky, 2023). During that research, interviewees from diverse backgrounds suggested that the future of science was going to be in automation. We draw on four of those interviews. When we decided to pursue this as an independent project, we focused on DARPA's efforts to automate science.

DARPA was chosen as a central focus for two reasons. First, DARPA was an early and consistent funder of AI research since the field emerged in the middle of the last century (Nilsson, 2010). As we argue later, many of the strategies used by the field to organize the field, evaluate work, and make progress were innovated by DARPA. Taskification—reducing complex problems to doable, bite-sized tasks—and benchmarking—using quantitative tests to evaluate progress and allocate resources—were becoming integrated into DARPA's portfolio of AI research in the 1990s and 2000s before becoming widely adopted by the field at large (Koch and Peterson, 2024). Second, although it is an agency within the Department of Defense and, thus, shrouded in relative secrecy, DARPA maintains a public face describing its portfolio of projects, making them more open for scrutiny than science automation efforts at startups and Big Tech firms.

The 26 interviewees for this project were selected on the basis of their participation in projects dedicated to automating some aspect of scientific practice. Because DARPA provides many grants in this area, we began by reaching out to prominent researchers associated with Big Mechanism, SCORE, ASKE, and other programs they have funded focused on automating various aspects of science.[2] From these initial interviews, we snowballed and interviewed non-DARPA-funded researchers doing automation research and critics of such projects.

Interviewees were spread across academia and industry with many having experience in both domains. Among the firms represented, we interviewed employees of Microsoft, IBM, and Philips Engineering. We also spoke with a researcher at the Chan-Zuckerberg Initiative and an analyst at the Organization for Economic Co-operation and Development focused on science policy. We spoke with researchers at universities across the United States and Europe including Harvard, University of Chicago, MIT, University of Amsterdam, and Goldsmiths University.

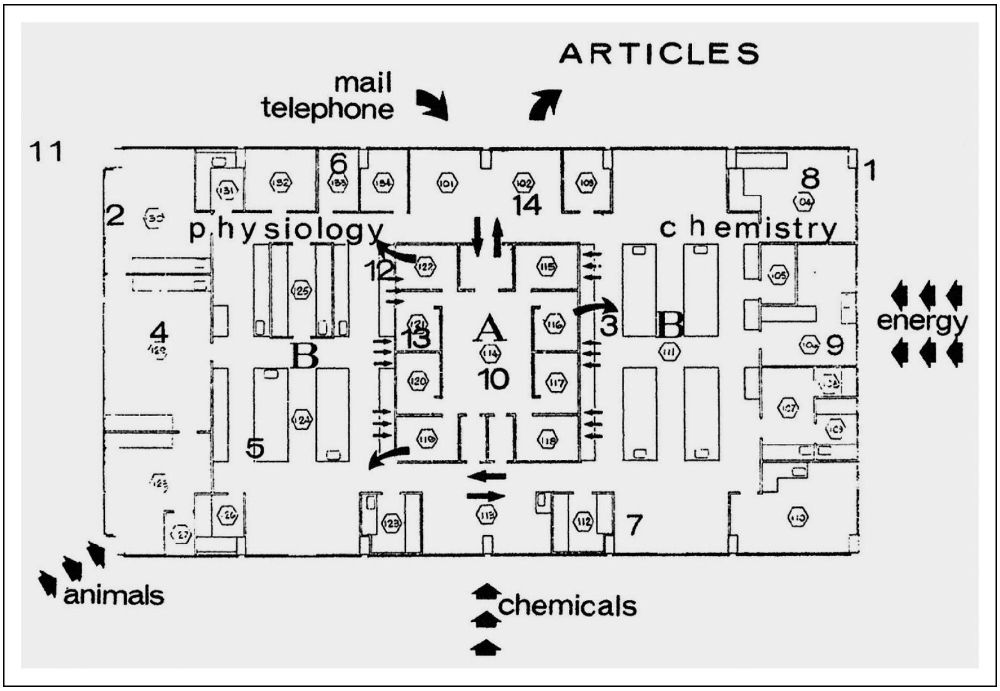

Figure 1. The laboratory as article-producing system (Latour and Woolgar, 1979: 46).

Interviews were focused on general questions regarding the automation of science, their projects, and the tensions between automators and field experts. The interviews averaged 55 min.

Dreams of a robot scientist and the reality

In their lauded laboratory ethnography, Latour and Woolgar (1979) famously conceptualized the neurobiology laboratory of Roger Guillemin as an engineering diagram in which research animals, chemicals, energy, and communications were input and academic articles were output (Figure 1).

Rather than a mere heuristic, it is easy to reimagine the diagram as the blueprint for an actual robot scientist. Such a system would input communication, energy, and research materials, and it would output articles and data. However, rather than humans performing the various tasks of reading, synthesizing, hypothesizing, planning, experimenting, analyzing, and communicating, they would all be done by a unified system.

Such a system could be the progeny of the work we describe below, one that manages to fuse together systems automating judgment, insight, and embodied knowledge. The system would read every published article and download every available dataset relevant to its topic. It would evaluate these pieces of evidence and create a massive graph representing the structure of knowledge its field. Nodes would represent entities connected by edges representing causal relations. Places where evidence is incomplete, ambiguous, or contradictory would represent areas of potential interest. Based upon some automated evaluation of the relative importance of the contribution, the system would decide what to pursue. Then it would run, all day and night, hundreds or thousands of experiments, continuously aggregating and analyzing data and periodically tweaking the experiment to refine the research question or address ambiguities arising during the experiments. Once sufficient data had been gathered to firmly fill in a gap or resolve a conflict in the literature, a simple article would be prepared describing what was accomplished in plain language for human audiences while data would be uploaded to a public repository for other robot scientists to integrate, updating their knowledge graphs. This is the dream of the robot scientist: An autonomous, indefatigable, knowledge-generating machine capable of superhuman feats of ingenuity, insight, and effort.

Such a system would have three obvious benefits over human scientists. First, it could work continuously and tirelessly and without the suite of human cognitive biases. Second, rather than become siloed into a narrow area of specialization, a problem recognized for half a century (Price, 1975), such a machine could survey the entirely of the scientific landscape, potentially making connections and integrating fields in unpredictable ways. Lastly, rather than undergo the lengthy evaluation process by which fields decide which claims to accept and which to reject (Cole, 1992), evaluation and integration of new knowledge would happen almost instantaneously while false claims would be extirpated rather than linger (Smaldino and McElreath, 2016).

Given the advantages of automating science, it is not surprising that this ambitious research program should take hold at DARPA, an important source of innovation in science since the Cold War.

DARPA and the automators of science

The US government founded DARPA after being blindsided by Sputnik. The new agency was tasked with investing in and steering cutting-edge research with military and economic significance. While other government funders supported broad portfolios of mainstream research, DARPA was charged with pursuing riskier projects with potentially transformative outcomes (Weinberger, 2017). Some of the technologies that DARPA incubated include global positioning satellites, drones, and the internet.

Most germane to the current discussion, however, is the role DARPA played in the integration and steering of AI research in the 1980s and 1990s. As the amount of intelligence information grew explosively in the second half of the last century, DARPA turned toward computer technologies to help intelligence analysts deal with the deluge (Roland and Shiman, 2002). One branch of this research involved funding AI projects.

A recent account demonstrates the significant role that DARPA played in creating the modern field of artificial intelligence (Koch and Peterson, 2024). After a period of laissez faire funding in the field, DARPA began demanding more accountability. By designing "common tasks" like optical character recognition, automated translation, or fingerprint analysis, DARPA used their purse strings to organize the field around doable projects with clear metrics of progress. The common task structure provided a format for comparison and, thus, competition. As simple tasks were "solved," more complex tasks were designed in a process known colloquially as "hill climbing." For example, computer vision tasks were narrowly focused on recognizing numbers before expanding to other symbols and, eventually, objects. Many of the AI technologies we use today including image recognition, automated translation, and personal assistants began as DARPA projects that followed this trajectory. Given DARPA's success in AI funding, it is unsurprising that they would adopt a similar strategy for their $2 billion "AI Next Campaign" launched in 2018.[3]

Hill-climbing the peaks of scientific skill

Returning to Moravec's metaphor of the rising tide of machine competence, previous literature in science and technology studies has identified three "peaks" representing uniquely human competences that have, thus far, stood above the rising tide of automation: scientific insight, embodied knowledge, and collective judgment. This list is not meant to be exhaustive. Certainly, human contributions to science run through the gamut of research practice in complex ways. Yet, these skills represent both areas of significant research in science studies and have clear analogs within the DARPA projects we detail. That the projects described below imperfectly map onto these peaks of human competence is a result of the taskification process. While automating an abstract skill like "judgment" represents too steep a hill to climb, designing a specific task that is designed to represent an aspect of judgment allows for progress to be made in a way similar to progress in other areas of machine learning.

The taskification of insight

Our first domain of supposedly unique human accomplishment is what we broadly label *insight*. These are forms of intelligence tied to individual or interpersonal cognition. Insight is the largely unconscious ability that scientists have to synthesize complex and heterogeneous data, identify gaps, and plan novel interventions (Polanyi, 2009). While machines are unparalleled in their ability to solve delimited problems in domains with fixed rules, in many areas of science, the boundaries of the domains and the rules themselves are unclear and require human ingenuity to find analogies across disparate domains, exploit chance occurrences, and innovate (Dunbar, 1997; Simonton, 1999).

Throughout the history of science, data and articles have been navigated by humans using typically human forms of information processing. Data are condensed into summary statistics or displayed graphically to make them legible to human audiences. Narratives are then crafted to highlight what the evidence is supposed to mean in the context of previous research. Individual scientists develop cognitive models of how the systems they study function and draw on those models when making decisions. Individual biases or gaps in knowledge are, in theory, overcome institutionally through collective decision-making (e.g., peer review, committees, etc.).

Beyond simply organizing masses of data, new approaches in machine learning provide a means to colonize a role that has always been occupied by a human scientist. Atif Khan, a DARPA-funded computational biologist from the University of Chicago, detailed how they were attempting to "automate this discovery thing, rather than our [human] collaborator feeding us questions." Rather than be at the mercy of the domain experts, he asked, "Why not automate the discovery phase? Why not automate the AI things within the database so that machine can give us the interesting pattern on its own, rather than asking any particular question?"

Harvard computer scientist Benjamin Gyori led curation efforts on the DARPA program Communicating with Computers. He explained that the more that science establishes an "automated pipeline going from publications to models," the more the entire system can be treated as a single, evolving entity:

> For instance, we have a model for COVID-19 that reads the new literature generated every day in that space. Through

> machine reading and knowledge assembly, it picks up specific new pieces of knowledge that are reported and assembles those pieces of knowledge against everything else that we already know. And it figures out whether a given piece of knowledge was already known, so there was already evidence for it, whether it's more general than what we already knew, whether it's more specific than what we already knew, or whether it contradicts what we knew before. And that's immediately helpful in figuring out how a new publication and what it reports relates to previous publications.

Natural language processing tools are being developed to allow machines to do two, daisy-chained tasks. First, to "read" the published literature and, second, to organize it into a network of established relationships. Such a system would not only passively "crunch the numbers," but would find structural holes or areas of conflict that could represent promising areas of future research.

As one commentary notes, such a system would produce an "intelligent science assistant" that could

> identify and summarize relevant research described across the worldwide multilingual spectrum of blogs, preprint archives, and discussion forums; find or generate new hypotheses that might confirm or conflict with ongoing work; and even rerun old analyses when a new computational method becomes available. Aided by such a system, the scientist will focus on more of the creative aspects of research, with a larger fraction of the routine work left to the artificially intelligent assistant. (Gil et al., 2014: 171)

As they note, the robot "assistant" would take over the computationally intensive work of aggregating evidence and generating new hypotheses while leaving the scientist more time for the "creative aspects of research." Again, the system only usurps a narrow band of current scientific expertise. One notable area left to the scientist under this scenario is actually conducting experiments.

Beyond simply providing assistance to humans, however, researchers on these projects suggest they may help overcome a limitation of human cognition. Nancy J. Cox, a quantitative geneticist who was a contractor on DARPA's Big Mechanism, explained that the problem of complexity has become a qualitative barrier. "When you have these huge jumps in the volume of information, there's much more information there than we know how to model, and we know how to even begin to understand." Machines, however, do not share this limitation. Thus, the hope is not simply to replicate human insight, but expand the bounds of what is possible.

The taskification of embodied knowledge

The second skill which it has been argued that humans are uniquely capable is the domain of *embodied knowledge*. This includes forms of intelligence tied to the physical performance of tasks. A foundational topic in STS (Collins, 2010; Polanyi, 2012), embodied knowledge is the dimension of scientific practice concerned with the physical aspects of getting experiments to work. Tasks that are unstandardized and require sensitive reactions to changing conditions cannot be laid out as an explicit set of rules (Lynch, 2002) and, thus, are not amenable to automation. Although there are reasons to suspect this hurdle is not insurmountable (i.e., Feest, 2016), the ability to function in a lab is still often tied to physical performance (Doing, 2004; Peterson, 2015, 2025).

For decades, commentators have argued that "robotic automation is poised to revolutionize laboratory practices" (Boyd, 2002: 517). In targeted ways, gene sequencing, high-throughput proteomics, and microarray technology have already extended the reach of human scientists. However, although these represent significant aids to research, they are relatively contained tasks or, as one commentary puts it, "They did not fully 'close the loop' in the sense of examining the results of their actions, deciding what to try next, potentially cycling forever" (Waltz and Buchanan, 2009: 43).

Beyond these more modular approaches to automation, there have been more ambitious attempts to create fully, "closed loop" systems that can operate an ongoing series of experiments with minimal human intervention. This goal motivated the creation of "Adam," the first robot scientist who "physically runs the experiments by using laboratory robotics, interprets the results, and then repeats the cycle" (King et al., 2009: 85). Specifically, the system conducted functional genomics experiments whereby specific genes of yeast were removed systematically to see if the phenotype displayed dysfunction in aromatic amino acid synthesis (King et al., 2004).

Its creators argue that Adam is useful for doing many of the routine and dull laboratory jobs, freeing the scientist for the more intellectual aspects of the work. When we asked Larisa Soldatova, one of the team members on the Adam project, what motivated the project she pointed to the errors that stem from having humans do repetitive actions:

> Looking how people are doing the same thing again and again, getting tired and just making errors, or something that they don't want to do. If you talk to any scientist, I'm sure anyone will tell that the actual creative work, when they can think and do what they enjoy, it's only a fraction of what they have to spend their time on. They have to spend so much time on something which they just have to do and don't want to do. Usually, it's amenable to automation.

In this way, the "task" that can be automated is relegated to the machine which preserves more time of the "skills" of the human scientist. Thus, despite the end-to-end structure of Adam, the human scientist remains in an important position. When we asked King what role the human scientist would have if robot scientists became common, he explained that the "human being would be doing the clever stuff. The general strategy; they understand what they're trying to achieve, what the goals are. The computer doesn't understand the big picture." Despite mastering the physical aspects of experimentation and some low-level hypothesis generation and analytic tasks, the human is still conceived of as necessary to provide goals and strategize.

The robot scientist offers several advantages over the human. First, unlike the human scientist who is constrained to acting in a serial way, engaging in a string of individual actions, the machine can operate in a parallel fashion—preparing, experimenting, analyzing, and planning at the same time. Second, whereas even the most obsessed scientist must take periodic breaks to tend to their biological needs, robot scientists can operate with minimal breaks.

The taskification of judgment

Although scientific expertise is involved across the gamut of scientific practice from the minutely technical to the abstractly theoretical, a key element is the ability to judge the quality of work in a field. DARPA has funded several projects with the goal of automating this faculty including the ambitious Systematizing Confidence in Open Research and Evidence (SCORE) project. The motivation for SCORE is the reproducibility crisis in social and behavioral sciences (SBS) producing an unreliable literature for political and military strategists who utilize this literature for decision-making. Thus, the project attempts to

> develop and deploy automated tools to assign "confidence scores" to different SBS research results and claims. Confidence scores are quantitative measures that should enable a [Department of Defense] consumer of SBS research to understand the degree to which a particular claim or result is likely to be reproducible or replicable. These tools will assign explainable confidence scores with a reliability that is equal to, or better than, the best current human expert methods. (DARPA, 2021)

Broadly, the project consists of three phases. In the first, domain experts participate in "prediction markets" (Liu et al., 2020) or focus groups and predict the replicability of specific articles. In the second phase, a sample of these articles are selected and replicated. This allows researchers to (1) evaluate the accuracy of domain expert judgment and (2) produce a dataset that supports phase three. Finally, automated systems are developed which are designed to mimic the most accurate expert's judgment using specific features of the published articles (for instance, the presence or absence of certain phrases [Yang et al., 2020]).

Sarah Rajtmajer, computer scientist and one of the contractors on DARPA's SCORE program explained that "what we are trying to do is approximate groups of experts." She went on to explain the impetus for automating expert judgment:

> A very successful approach is prediction markets. But then you have to have weeks or months to run the prediction market. You have to assemble the correct people. You have to be running a replication study in parallel so that they have something to bet on [...] That's not really scalable [...] We're developing artificial markets. But all the teams are essentially trying to get as close as possible to the human expert's wisdom.

Rather than have to wait for the outcome of a prediction market, the ability to create a fully automated, artificial market provides instant access to "expert opinion" on any piece of social science.

This system would allow decisionmakers to evaluate the social and behavioral science literature in real time without having to consult a cadre of experts. Yet, it is entirely dependent on a published literature still being directed by human scientists. However, another set of projects seeks to automate the direction of scientific research itself.

The shell game of uniquely human skill

In isolation, the projects designed to automate the three tasks detailed above—judgment, insight, and embodied knowledge—each promise to absorb some task amenable to automation freeing the human scientist up for some other specifically human work. Thus, these projects are pitched as providing a "intelligent science assistant" (Gil et al., 2014) rather than something aimed at replacing the human scientist. In the case of hypothesis generation, for instance, the assistant simply weeds through vast amounts of data in order to provide the human scientist with a manageable list of experimental possibilities. But both the selection and the experiments actually conducted still require a trained human hand. This was the way that several interviewees framed their work. However, the tasks considered to be the unique dominion of the human scientist differ from project to project. For instance, Benjamin Gyori, computer scientist at Harvard Medical School, explains how the "Automated Scientific Discovery Framework" he created could aid the human by creating a system to

> quickly sift through that data and highlight surprising and yet unexplained observations, to separate the things that are either not interesting or have already been explained and therefore are not scientifically novel, and provide a kind of rank list and a report to the human scientist to draw their attention to specific experimental observations that they should follow up on.

By organizing data, synthesizing it, and locating open questions, the machine has taken over some of the tasks we have labeled insight. Moreover, the "ranked list" provided by the system encroaches into the domain of scientific judgment by suggesting that some experiments are more potentially fruitful than others. Yet, the human scientist is still ultimately deciding what experiments to do and actually carrying them out.

However, for those seeking to automate experimentation, the embodied tasks considered the unique domain of the human scientist are precisely the ones they seek to offload onto a machine. In this case, though, the unique position of the human again shifts. For instance, King et al. (2004: 251) suggest the automation of experimentation would aid progress because "it frees scientists to make the high-level creative leaps at which they excel." Thus, rather than embodied knowledge, the task left for the human scientist is that of insightfully synthesizing and interpreting the experimental products of the robot experimenter.

What should be clear is that the assertions that humans have some role in the evolving system of science made by automators has more to do with the limited focus of their current projects than any theoretical belief in some insurmountable peak of human performance. Attributing an ongoing role to a human is simply about delimiting the scope of their system. By their construction, concepts like embodied knowledge, insight, and collective judgment are amorphous whereas these tasks are narrowly defined. Yet, when you look at all these various projects as an emergent research program, it is clear there is no uniquely human domain. While each individual project preserves some space for human skill, in another automation project, that very skill is being taskified and automated.

Of course, by calling this an emergent research program, we do not mean to suggest that no one is thinking in terms of the full automation of science. While such ambitions were rarely stated in the early days of this project, it has since become much more common. Program managers at DARPA have been funding projects across the spectrum of automation. And, although nearly all of our interviewees framed automation in narrow terms, leaving room for human expertise, there were, early on, hints at this more ambitious agenda. King, whose robot scientist attempts to automate both embodied knowledge and (to a lesser degree) insight, was most explicit about this goal. Although he made several comments regarding the ongoing role of humans in science, he made clear that, ultimately, this was a role that was destined to disappear: "There's nothing which in general I think a human being can do which a computer won't be able to do. If you could describe exactly what a computer can't do, then the computer can do it." It was this belief that motivated the funding of the "Nobel Turing Challenge" with the goal of building a robot scientist that would make "Nobel quality scientific discoveries" by 2050 (Kitano, 2020).

As progress is made automating individual tasks, there is reason to believe that this emergent program will become more explicit and coherent. Yet, in the interim, it faces significant challenges. In the following section, we leave the individual tasks and automation projects to ask the question, what would need to happen to make the automation of science possible?

Remaking the landscape to facilitate the automation of science

In a provocative book, Pasquinelli (2023) has argued that, rather than reflect human intelligence, artificial intelligence mirrors the social logic of industrial production. Translating this to Moravec's rising sea metaphor, the approach taken by DARPA and others investing in the automation of science suggest a strategy that resembles Taylorism. Rather than creeping up the peaks of human competence, the strategy of taskification—with its reduction of complexity, the creation of doable tasks, and hill-climbing—suggests that the landscape itself is being remade to benefit machines for the purposes of making production more efficient.

As it's currently constituted, science is made by humans for humans. On the one hand, this is simply because it often involves complex and underspecified practices. Human skills like flexibility and metacognition are often necessary in conditions where the natural world doesn't provide us with objects that behave in fully predictable ways. As an example, Ross King told us that despite intentionally choosing yeast—a relatively simple organism—as a research object for the proof-of-concept robot scientist, problems still emerged: "Sometimes for several weeks, the experiments wouldn't work anymore. And you'd think, 'What on earth has happened? It's robot and yeast, what could have changed?' […] It became very clear to us that it's quite hard to reproduce even the simplest type of biology."

On the other hand, science is largely produced by and for humans because that is the way it has always been. For instance, Harvard genomics researcher Heidi Rehm described "the biggest barrier to actually enabling these approaches" which is data contained within "unstructured manuscripts with graphics and text that's hard to compute on" can be understood and integrated by a "professional brain" yet is often unreadable by a deep learning system.

Although the problems faced by automators are unique to their specific projects, they share a common origin: science has been, and continues to be, a profoundly *human* endeavor. Materials and machinery are standardized enough for experiments to "travel" across scientific networks, but only under the stewardship of a skilled human experimenter. Machine scientists require a far greater degree of standardization. Articles are written for the benefit of human readers. Data are condensed into graphs and tables which are more easily digested by human minds and these data visualizations are used to support a narrative filled with colloquial language. For the machine scientist, this presentation only simplifies and obfuscates.

There are, however, efforts to make science more amenable to machine intervention in ways that mirror changes in other sectors of the economy. For the last decade, industries like ride sharing and package delivery have been engaging in a process of "pre-automation" which involves transforming a domain to be both logistically and economically amenable to future automation innovations (Vertesi et al., 2020). Services like Uber rideshare and Amazon's package fulfillment have followed a logic in which they rapidly increase their scale through massive hiring of gig workers, make the work as data-centric and algorithmic as possible and, simultaneously, heavily invest in automating technologies that they hope will eventually replace the human workers.[4]

This period of preautomation in science is not following the same script as ridesharing and package delivery. A major reason is that scientists have significantly more ability to resist the reorganization of their workplace than taxi or delivery drivers. However, the essential logic of remaking a profession to be amenable to automation is visible in a number of science reform projects. These range from subtle rule and policy changes being advocated under the banner of "research integrity" to more ambitious attempts to remake science at a more granular level.

Making science machine-readable

Discussions of "openness" and "transparency" in science are largely motivated by discussions about a "reproducibility crisis" in science (Peterson and Panofsky, 2023). Under this framework, the move toward open science—especially, open data, methods, and materials—is supposed to improve the integrity of the scientific literature by making claims easier to evaluate and replicate (Freese and Peterson, 2017). For instance, making it a requirement to deposit raw data with an article submission encourages competing scientists to check the claims and discover errors.

In contrast to the arguments of open science advocates, Historian Philip Mirowski has claimed that "the open science movement is an artifact of the current neoliberal regime of science, one that reconfigures both the institutions and the nature of knowledge so as to better conform to market imperative" (2018: 172). Setting aside the question of the degree to which open science can claim to genuinely improve knowledge outcomes, the move to open science has obvious, tangible benefits for automation projects whose value is in direct proportion with the amount of data they have access to.

Advocating open science and then pushing to make data, datasets, and repositories more machine-friendly have become central goals of large, institutional funders. In 2019, the NSF launched an initiative called "Harnessing the Data Revolution" to begin supporting this transition. Arthur Lupia, Assistant Director for the Social, Behavioral, and Economic Science Directorate at the NSF explained that "we have a lot of discussions now about what data looks like in the future" with an eye toward innovations in "machine learning and the ability to scrape data."

For the librarians who build, organize, and maintain the infrastructure that open science requires, the potential for the exploitation of data is a clear motivation. For instance, Eloy Rodrigues, board member for the Coalition of Open Access Repositories, told us, "I think it's really clear that there is a lot of things that could be enabled, if you really have an open ecosystem where all the data and publications can be available to be mined and be used by algorithm. That will offer a lot of new possibilities for research and for the advancement of knowledge." He went on to describe the next generation of repositories which will be designed with the knowledge that "users of repositories and of this ecosystem will not only be humans, but it will be also machines […] We need to build systems that are planned since the beginning to be used by machines and are machine friendly and allow new exploration of things."

Beyond data that support published articles (and, thus, support the mission of increasing research integrity), there is recognition that machines will benefit most when all data are available. This includes data from experiments that, for whatever reason, do not end up getting published and, thus, do not face data hosting mandates. Jake Yeston, an editor at *Science*, told us about a workshop he attended on the application of artificial intelligence in synthetic chemistry. The speakers highlighted how growth in the field has been hampered by the so-called "file drawer problem." Simply put, the file drawer problem is a form of bias that results from the fact that studies that fail or do not produce novel findings tend *not* to get published. This skews the published literature toward successful and interesting findings and gives a distorted picture of the universe of studies. In scientific cultures, this bias can be somewhat corrected through back-channel communication about which findings or experimental procedures are "fragile" or "robust" (Peterson, 2016). There are debates about whether this is an efficient system of communication for human scientists (Peterson and Panofsky, 2023). What is clear, however, is that such bias is deeply problematic for machine learning systems that are deaf to such water cooler conversations. Yeston explains,

> What they do is they basically teach the system to go look at a million reactions that have been published somewhere. And so it kind of figures out this is how to do things. And then, based on however it's programmed, it looks for similarities and so forth. And once in a while it comes up with something that looks really outlandish and a human being will say "Oh that wouldn't work." But the reason for that is that it doesn't have a database of "this doesn't work."

Although there have long been arguments that there should be venues to publish negative results, the need becomes pointed as machines become more active participants in scientific research.

Producing data explicitly for machines

Although reformers have been successful in advocating for a number of policy changes mandating more open science practices, automators still face significant problems. For instance, simply having access to, say, clinical data is not enough to make a deep learning model useful. Ezequiel Lopez-Rubio, an artificial intelligence researcher working on automating clinical science, detailed, "When you talk to doctors they tell you that when they collect the data there are many inconsistencies, and that means that the variables that they are measuring are not exactly the same depending on the doctor who is collecting the data." This hamstrings deep learning systems because "if the variables are not consistent then that answer is useless."

Similarly, having access to databases of articles is useful for systems that seek to synthesize knowledge in fields, but the current abilities of natural language processing limit what can be done. Gyori notes that even a seemingly simple task like entity recognition is "a huge challenge because people use widely varying nomenclature and synonyms and alternative spellings, and names change over time." But even once the hurdles of getting the algorithm to correctly identify entities are overcome, discovering what is actually said about them presents an even more daunting problem: "Then, all of these entities are embedded in very complicated sentences, and these sentences often describe really intricate points about methods and experimentation and so on. The current state-of-the-art NLP systems don't actually recognize much of that."

These issues underscore the essentially different ways that humans and machines process information. While human scientists pale in comparison to machines in the ability to analyze masses of curated data, they best the machine when data are heterogenous, ambiguous, or incomplete. And, while a machine can absorb and organize scores of articles, humans are far superior at understanding the meaning of sentences written in natural language. Returning to the rising water metaphor, there are many automation projects seeking to restructure science itself to benefit machine integration.

The establishment of an international system of institutional and disciplinary repositories is already setting the groundwork for one possibility: moving away from the journal system based upon article-length narratives toward a system based upon individual experiments. Neuroscientist and creator of the MindMap system for synthesizing biomedical data, Alcino Silva told us that the future of science "won't be about human judgments" where scientists think, "*Cell* and *Nature*, wow! *Science*, that's where the good stuff is." Instead, we transition to a system in which the unit of contribution and evaluation is the single experiment. According to Silva, an evolving, self-updating system based around individual experiments would allow funders and universities to evaluate studies based on what they lead to rather than the brand name of where they appeared.

These changes would be aided by another set of projects dedicated to formalizing scientific terminology and language. As the creators of Adam argue, "The use of computers is also changing the way that science is described and reported. Scientific knowledge is best expressed in formal logical languages. Only formal languages provide sufficient semantic clarity to ensure reproducibility and the free exchange of scientific knowledge" (King et al., 2009: 85). Fulfilling this long dormant project of logical positivism would provide machines with data in a digestible form.

One concrete outcome of this project has been the development of "ontologies" in biomedical research: a formal, codified naming system and sets of technical concepts labeled as incongruous. These issues are not disastrous for human scientists who can navigate heterogeneous information spaces, but they create deep problems for machine learning systems. The goal of developing biological ontologies, therefore, is to harmonize these vocabularies and create a homogeneous information space better suited to deep learning.

Discussion

Of all domains of intelligent human activity that could be automated, few would have more significance to the fate of our world and our species than science. Yet, in contrast to Moravec's rising water metaphor, AI is not simply approaching the peaks of scientific competence. Rather, skills are being taskified and then hill-climbed while scientific infrastructure is remade to better integrate with machines.

The total automation of science represents an emergent research program. Although, historically, few funders or researchers explicitly espouse such an ambitious goal, surveying the landscape of automation efforts in science makes the shape of this project clear. Scientific robotics, natural language processing, information retrieval and synthesis, and hypothesis generation are independent and parallel areas of research. Yet, if these projects mature and are integrated, they have the potential to transform science dramatically.

There is evidence that this is happening. For instance, in a recent report on a Silicon Valley startup promising to revolutionize materials science (Bort, 2025), a researcher at Google Brain specifically argued that advances along several discrete tasks (LLM reasoning, robotics, and ability to simulate complex reactions) make the time ripe for automation. The author concluded, "Stitched together, the picture was clear: A simulation would theoretically discover new compounds a robot could mix the materials, and an LLM could analyze the results and suggest course corrections. AI-automated material science was ready to be built." The implicit view is of a science broken down into tasks which have been independently hill-climbed to the point where they might be joined together to create an autonomous robot scientist.

Since the proof-of-concept labs "Adam" and "Eve" were created (Tobias and Wahab, 2025), "closed loop" or "self-driving" labs have gained steam in relatively structured scientific pursuits like materials science (Abolhasani and Kumacheva, 2023; Wahl et al., 2021) and some areas of biomedicine (Collins et al. 2025). Yet, the heterodox conditions and unclear objectives of much laboratory research make these models too simplistic to generalize (Musslick et al., 2025). The path forward is, as it has been for AI for decades, hill-climbing narrowly defined skills in the hopes that they will eventually join at the mountain peak of scientific expertise.

Of course, from its beginnings, artificial intelligence has always been accompanied by wild-eyed predictions and a spotty record of accomplishment (Mitchell, 2019). This article is not meant to suggest that there is an impending takeover of science by machines. Although most of the researchers we interviewed painted an optimistic picture of the future of machine intelligence in science, there were notes of caution as well. That this remains an uncertain bet on the future can be gleaned from the fact that many of the people we interviewed and many of the projects and institutes on which and at which they were employed were being sustained by grant money making their positions unreliable. Funders and universities want to see concrete developments before committing themselves more permanently to funding ongoing automation projects.

There are good reasons for caution. The end-to-end robot scientist Adam was designed to work on yeast because of its simplicity and still struggled with reproducibility. What does this portend for the move into more complex areas of biology? And, variations in terminology across fields have promoted efforts to standardize ontologies. Yet, creating a common ontology requires the harmonization of domains that have evolved separately and there are reasons to suspect that differences in technical argot may reflect theoretical disputes rather than mere terminological styles (Leonelli, 2019). Moreover, in an age where the management of labor and technology are increasingly intertwined (Jarrahi et al., 2021) and the integration of technology seems to do little but increase burdens on researchers (Watermeyer et al., 2024), such efforts seem to portend the deskilling and impoverishment of scientific labor in ways that mirror other fields (Casilli, 2021; Pasquinelli 2023). Thus, some internal resistance may be expected.

Despite these obstacles, it would be shortsighted to dismiss the total automation of science as an irrelevant, fringe fantasy. In addition to the projects outlined above, researchers have been making progress automating protein structure prediction (Jumper et al., 2021), chemical structure identification (Rajan et al., 2023), hypothesis generation (Spangler et al., 2014), and solving symbolic regression problems (Udrescu and Tegmark, 2020). One researcher argued in *Nature* that machines should claim prizes in science for their contributions to discovery (Wang, 2025).

Sure to accelerate these trends, many of the significant changes happening within global science over the last decade ease the way for increased automation. If the water cannot rise to the peaks, make the peaks easier to reach. Specifically, efforts to "open science" provide masses of data on which to train deep learning models. Replication initiatives help to standardize practices in ways that benefit experimental robots. And the move toward bigger and more networked science creates conditions in which "tacit knowledge" and "cultural judgments" must be made explicit. In a period of pre-automation, scientists around the world are, knowingly or not, laying the groundwork for machines.

It is hard to predict what influence any of these individual automation projects will have, let alone their combinatorial possibilities. The reason why it is important to attend to the entire shape of the project is that it foreshadows significant changes in experimental practice, scientific communication, and research careers. Even if Ross King's dream of having a computer win a Nobel Prize by 2050 does not come to pass, by preparing for that future by changing scientific practice and institutions in ways designed to spur machine advancement, they are having an important effect on science.

In a recent report on the state of science in the age of AI, the Royal Society (2024) detailed a litany of justified concerns of AI-driven science like access, bias, and transparency. But one significant aspect missing from the report is how AI may threaten the social structure and institutions of science. Even if the projects detailed above only lead to success in overtaking relative "lowlands" like automating limited aspects of laboratory practice, they portend significant changes in the nature of scientific work. If machines are as good as (if not far superior to) humans in most tasks currently performed by graduate students and postdocs, it is fair to wonder what tasks will be left for them.

Beyond these institutional effects, however, the automation of science represents a challenge to prevailing ideas in the philosophy and sociology of science. For nearly two decades, some have argued that we are entering a new epistemic era in which prediction overruns understanding as the primary goal of science (Anderson, 2008; Kitchin, 2014). While some have suggested that machines may produce better science because they lack our cognitive limitations and biases, an increasingly pressing question is, what biases do the machines have (Messeri and Crockett, 2024)? And will we be content to simply to cede expertise to machines in ways that impoverish us mentally and potentially results in epistemic monocultures (Bergstrom and Bak-Coleman, 2025; Evans and Duede, 2025; Koch and Peterson, 2024)?

Machine learning methods pose even more theoretically difficult issues. Although current systems are more adaptable than their rule-based forebears, they gain their flexibility at the cost of opaqueness. In the area of expert judgment, this means that demands for reasons cannot be meaningfully answered and raises questions of judgment and evaluation. If no one can reasonably understand how a machine produced a piece of analysis, how are we to understand its quality?

Beyond seeking to automate physical and analytic practices, another set of issues when thinking about automating the acculturated knowledge of the scientist. While most of the people we interviewed seemed to hold out hope that this would be the final peak where humans will make their stand against the rising tide of machine automation, projects seeking to automate expert judgment suggest that that too is in the automators' crosshairs. Yet, deciding what needs to be studied and what counts as a significant advance is where the technical side of science intersects more directly with the social and political side. This has profound implications for the role of human knowledge in a post-human science.

ORCID iDs

David Peterson https://orcid.org/0000-0003-4033-2993
Ramya Natarajan https://orcid.org/0009-0008-2211-1939

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: David Peterson and Aaron Panofsky were supported by the Sloan Foundation (Grant Number G-2023-21026). Peterson also received a fellowship from the Deutsche Forschungsgemeinschaft (Project 470816212/KFG43).

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Notes

1. Pearl's idea is not new. As early as the 1960s, artificial intelligence researchers developed the DENDRAL project in an attempt to automate hypothesis formation and discovery in organic chemistry.
2. These projects represent the types of well-defined, doable tasks that are ideal for the measurement, comparison, and hill climbing. Although some details of the projects will be described later, here are the projects we recruited from: ASKE (Automating Scientific Knowledge Extraction): funding projects through a two-stage process which (1) will automatically curate and synthesize a research literature and (2) inferring from these syntheses to generate new hypotheses. Big Mechanism: funds group working to automate the extraction of causal information in the literature on a specific form of cancer and then assemble these pieces into more complete causal models. SCORE (Systematizing Confidence in Open Research and Evidence): funds projects designed to create automated systems which can read studies in behavioral and social science in order to automatically evaluate them based upon "clues" to their reproducibility and reliability.
3. Of course, DARPA is not alone in pursuing automation in science. All the Big Tech companies are now investing in this project. Although Google's AlphaFold program, which effectively solved one of the most challenging problems in structural biology, is perhaps the most famous, META, Amazon, and other companies have made similar investments. However, among science funders, DARPA has the most publicly visible program pushing forward the cutting edge of science automation along multiple domains. Because of the scope and diversity of their funding portfolio, and the availability of their scientific contractors, we developed this project around DARPA's efforts in this science automation.
4. The fact that both Uber and Amazon have just recently reduced investment in their automation projects highlights the ongoing challenges of automating even relatively unsophisticated tasks.

References

Abolhasani M and Kumacheva E (2023) The rise of self-driving labs in chemical and materials sciences. *Nature Synthesis* 2(6): 483–492.

Alkhateeb A (2017) Can scientific discovery be automated? *The Atlantic*, April 25. Available at: https://www.theatlantic.com/science/archive/2017/04/can-scientific-discovery-be-automated/524136/ (accessed 17 November 2020).

Anderson C (2008) The end of theory: The data deluge makes the scientific method obsolete. *Wired*, June 23. Available at: https://www.wired.com/2008/06/pb-theory/ (accessed 1 June 2021).

Arribas-Ayllon M, Bartlett A, Lewis J, et al. (2019) *Psychiatric Genetics: From Hereditary Madness to Big Biology*. London, UK: Routledge.

Bergstrom CT and Bak-Coleman J (2025) AI, peer review and the human activity of science. *Nature*, ahead of print, June 25. https://doi.org/10.1038/d41586-025-01839-w.

Bort J (2025) Top OpenAI, Google brain researchers set off a $300M VC Frenzy for their startup periodic labs. *TechCrunch*, October 20. https://techcrunch.com/2025/10/20/top-openai-google-brain-researchers-set-off-a-300m-vc-frenzy-for-their-startup-periodic-labs/.

Boyd J (2002) Robotic laboratory automation. *Science* 295(5554): 517–518.

Casilli AA (2021) Waiting for robots: The ever-elusive myth of automation and the global exploitation of digital labor. *Sociologias* 23: 112–133.

Cole S (1992) *Making Science: Between Nature and Society*. Cambridge, MA, USA: Harvard University Press.

Collins E, Langer R and Anderson DG (2025) Self-driving labs for biotechnology. *Nature Computational Science* 5(11): 976–979.

Collins HM (1992) *Changing Order: Replication and Induction in Scientific Practice*. London, UK: Sage Publications.

Collins HM (2010) *Tacit and Explicit Knowledge*. Chicago, IL, USA: University of Chicago.

DARPA (2021) Systematizing confidence in open research and evidence (SCORE). *Defense Advanced Research Projects Agency*. Available at: https://www.darpa.mil/program/systematizing-confidence-in-open-research-and-evidence (accessed 21 May 2021).

Delamont S and Atkinson P (2001) Doctoring uncertainty: Mastering craft knowledge. *Social Studies of Science* 31(1): 87–107.

Doing P (2004) 'Lab hands' and the 'Scarlet O' epistemic politics and (scientific) labor. *Social Studies of Science* 34(3): 299–323.

Dunbar K (1997) How scientists think: On-line creativity and conceptual change in science. In: Ward TB, Smith SM and Vaid J (eds) *Conceptual Structures and Processes: Emergence, Discovery, and Change*. Washington, DC, USA: American Psychological Association Press, 461–493.

Ellis R (2011) *Jizz* and the joy of pattern recognition: Virtuosity, discipline and the agency of insight in UK naturalists' arts of seeing. *Social Studies of Science* 41(6): 769–790.

Evans J and Duede E (2025) After science. *Science* 390(6774): eaec7650.

Feest U (2016) The experimenters' regress reconsidered: Replication, tacit knowledge, and the dynamics of knowledge generation. *Studies in History and Philosophy of Science Part A* 58(August): 34–45.

Freese J and Peterson D (2017) Replication in social science. *Annual Review of Sociology* 43(1): 147–165.

Fujimura JH (1992) Crafting science: Standardized packages, boundary objects, and "translation". In: Pickering A (ed) *Science as Practice and Culture*. Chicago, IL, USA: University of Chicago Press, 168–211.

Gil Y, Greaves M, Hendler J, et al. (2014) Amplify scientific discovery with artificial intelligence. *Science* 346(6206): 171–172.

Haraway D (1997) *Modest_Witness@Second_Millennium. FemaleMan©Meets_OncoMouse™: Feminism and Technoscience*. London, UK: Routledge.

Hull DL (2001) *Science and Selection: Essays on Biological Evolution and the Philosophy Of Science*. Cambridge, UK: Cambridge University Press.

Jarrahi MH, Newlands G, Lee MK, et al. (2021) Algorithmic management in a work context. *Big Data & Society* 8(2): 20539517211020332.

Jordan M (2018) Artificial Intelligence — the revolution hasn't happened yet. *Medium*, April 30. https://medium.com/@mijordan3/artificial-intelligence-the-revolution-hasnt-happened-yet-5e1d5812e1e7.

Jumper J, Evans R, Pritzel A, et al. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596(7873): 583–589.

King RD, Rowland J, Oliver SG, et al. (2009) The automation of science. *Science* 324(5923): 85–89.

King RD, Whelan KE, Jones FM, et al. (2004) Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature* 427(6971): 247–252.

Kitano H (2020) Nobel turing challenge. Available at: https://www.oecd.org/going-digital/ai-intelligent-machines-smart-policies/conference-agenda/ai-intelligent-machines-smart-policies-kitano.pdf (accessed 23 May 2021).

Kitchin R (2014) Big data, new epistemologies and paradigm shifts. *Big Data & Society* 1(1): 2053951714528481.

Koch B and Peterson D (2024) From protoscience to epistemic monoculture: how benchmarking set the stage for the deep learning revolution." *arXiv.org*. https://doi.org/10.48550/arXiv.2404.06647.

Kounios J and Beeman M (2009) The aha! Moment: The cognitive neuroscience of insight. *Current Directions in Psychological Science* 18(4): 210–216.

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

Latour B and Woolgar S (1979) *Laboratory Life: The Construction of Scientific Facts*. Princeton, NJ, USA: Princeton University Press.

Leonelli S (2016) *Data-Centric Biology: A Philosophical Study*. Chicago, IL, USA: University of Chicago Press.

Leonelli S (2019) The challenges of big data biology. *ELife* 8: e47381.

Liu Y, Gordon M, Wang J, et al. (2020) Replication markets: results, lessons, challenges and opportunities in AI replication. *ArXiv*:2005.04543 [Cs].

Lynch M (2002) Protocols, practices, and the reproduction of technique in molecular biology. *British Journal of Sociology* 53(2): 203–220.

Messeri L and Crockett MJ (2024) Artificial intelligence and illusions of understanding in scientific research. *Nature* 627(8002): 49–58.

Mirowski P (2018) The future(s) of open science. *Social Studies of Science* 48(2): 171–203.

Mitchell M (2019) *Artificial Intelligence: A Guide for Thinking Humans*. New York, NY, USA: Farrar, Straus and Giroux.

Musslick S, Bartlett LK, Chandramouli SH, et al. (2025) Automating the practice of science: Opportunities, challenges, and implications. *Proceedings of the National Academy of Sciences* 122(5): e2401238121.

Nilsson NJ (2010) *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge, UK: Cambridge University Press.

Pasquinelli M (2023) *The Eye of the Master: A Social History of Artificial Intelligence*. London, UK: Verso Books.

Pearl J (2018) *The Book of Why: The New Science of Cause and Effect*. New York, NY, USA: Basic Books.

Peterson D (2015) All that is solid: Bench-building at the frontiers of two experimental sciences. *American Sociological Review* 80(6): 1201–1225.

Peterson D (2016) The baby factory: Difficult research objects, disciplinary standards, and the production of statistical significance. *Socius* 2: 2378023115625071.

Peterson D (2025) *The Unbuilt Bench: Experimental Psychology on the Verge of Science*. New York, NY, USA: Columbia University Press.

Peterson D and Panofsky A (2023) Metascience as a scientific social movement. *Minerva* 61(2): 147–174.

Polanyi M (1962) The republic of science. *Minerva* 1(1): 54–73.

Polanyi M (2009) *The Tacit Dimension*. Chicago, IL, USA: University of Chicago Press.

Polanyi M (2012) *Personal Knowledge: Towards a Post-Critical Philosophy*. Chicago, IL, USA: University of Chicago Press.

Price DJdS (1975) *Science since Babylon*. New Haven, CT, USA: Yale University Press.

Rajan K, Otto Brinkhaus H, Isabel Agea M, et al. (2023) DECIMER.Ai: An open platform for automated optical chemical structure identification, segmentation and recognition in scientific publications. *Nature Communications* 14(1): 5045.

Ribes D, Hoffman AS, Slota SC, et al. (2019) The logic of domains. *Social Studies of Science* 49(3): 281–309.

Roland A and Shiman P (2002) *Strategic Computing: DARPA and the Quest for Machine Intelligence, 1983–1993*. Cambridge, MA, USA: MIT Press.

Royal Society (2024) *Science in the Age of AI: How Artificial Intelligence is Changing the Nature and Method of Scientific Research*. Available at: https://royalsociety.org/-/media/policy/projects/science-in-the-age-of-ai/science-in-the-age-of-ai-report.pdf (accessed 1 December 2021).

Simonton DK (1999) *Origins of Genius: Darwinian Perspectives on Creativity*. Oxford, UK: Oxford University Press.

Smaldino PE and McElreath R (2016) The natural selection of bad science. *Royal Society: Open Science* 3: 160384.

Spangler S, Wilkins AD, Bachman BJ, et al. (2014) Automated hypothesis generation based on mining scientific literature. *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (New York, NY, USA), KDD '14, August 24, 1877–86. https://doi.org/10.1145/2623330.2623667.

Tegmark M (2017) *Life 3.0: Being Human in the Age of Artificial Intelligence*. New York, NY, USA: Knopf.

Tobias AV and Wahab A (2025) Autonomous "self-driving" laboratories: A review of technology and policy implications. *Royal Society Open Science* 12(7): 250646.

Udrescu S-M and Tegmark M (2020) AI Feynman: A physics-inspired method for symbolic regression. *Science Advances* 6(16): eaay2631.

Vertesi JA, Goldstein A, Enriquez D, et al. (2020) Pre-Automation: Insourcing and automating the gig economy. *Sociologica* 14(3): 167–193.

Wahl CB, Aykol M, Swisher JH, et al. (2021) Machine learning–accelerated design and synthesis of polyelemental heterostructures. *Science Advances* 7(52): eabj5505.

Waltz D and Buchanan BG (2009) Automating science. *Science* 324(5923): 43–44.

Wang D (2025) Prizes must recognize machine contributions to discovery. *Nature* 646(8084): 261–261.

Watermeyer R, Phipps L, Lanclos D, et al. (2024) Generative AI and the automating of academia. *Postdigital Science and Education* 6(2): 446–466.

Weinberger S (2017) *The Imagineers of War: The Untold Story of DARPA, the Pentagon Agency that Changed the World*. New York, NY, USA: Knopf.

Yang Y, Youyou W and Uzzi B (2020) Estimating the deep replicability of scientific findings using human and artificial intelligence. *PNAS* 117(2): 10762–10768.

Ylikoski P (1995) The invisible hand and science. *Science Studies* 8(2): 32–43.