Artificial intelligence and the changing costs and benefits of engaging in open and collaborative science

Susanne Beck and Marion Poetz[10]

Rooted in the core values of accessibility, transparency, and inclusivity (Vicente-Sáez & Martínez-Fuentes, 2018), the open science movement advocates for openly sharing scientific knowledge as early in the research process and as widely as possible. Related practices of open access publishing or managing and sharing research data along the lines of the FAIR principles are increasingly adopted by scientific research institutions and scientists across the globe and required by major funding programs such as Horizon Europe. The underlying conceptualization of openness mainly refers to *unidirectional knowledge flows* within science and from science to different levels and actors of an inquiring society. Recent studies on the benefits of openly and widely sharing research outcomes reveal that open access publications significantly broaden citation diversity across institutions, countries, and research fields (Huang et al., 2024), and are more frequently cited in patents (Probst et al., 2023). However, there remains a lack of conclusive evidence on the causal effects on research productivity and societal benefits.

As a remedy to the steady decline in scientific productivity over the past decades (e.g., Park, Leahey & Funk, 2023) and to better align research agendas with increasingly complex societal, health, environmental, cultural, political, or economic issues (e.g., Mazzucato, 2018), conceptualizations and definitions of open science increasingly emphasize openness as a means to foster collaboration (e.g., UNESCO, 2022); indicating a shift towards *bidirectional knowledge flows* for opening up the scientific knowledge production process itself. The concept of Open Innovation in Science (OIS) builds on this and more broadly encompasses inter- and transdisciplinary knowledge flows and collaborations along the entire process of generating and translating scientific research (Beck, Bergenholtz et al., 2022). More specifically, it outlines how and under what conditions practices such as crowd science or citizen science, open data reuse, or open forms of university-industry co-creation can improve the scientific productivity and the societal impact of research projects (Poetz et al., 2024).

Navigating the costs and benefits of engaging in open and collaborative science practices

When applying this framework to study antecedents, boundary conditions and effects of openness and collaboration in science (e.g., Beck et al. 2022, Beck, LaFlamme & Poetz, 2022) or using it to help participants in our Labs for Open Innovation in Science to develop their own OIS projects, we observed a consistent pattern across scientific fields and seniority levels: Researchers frequently view open and collaborative practices as additional efforts that must be undertaken "on top" of their regular duties, largely independent of whether these practices are mandated by institutional or funding requirements or driven by personal motivation. While many researchers had already adopted key open science practices such as pre-registrations of study designs or open access publishing, opening their own knowledge production processes is less common and sometimes viewed as particularly critical with respect to on-top efforts that may not translate into scientific productivity and related career advancements. This is particularly salient when it comes to engaging in transdisciplinary collaborations with companies, citizens, or other societal stakeholders. To put it differently, many scientists we worked with or talked to focus on the costs but often do not see enough benefits for their own projects or careers. Such benefits can, for example, be reflected in increased novelty or relevance of their research questions or hypotheses, improved quality or quantity of their research data, reduced biases in interpreting results, or new pathways to translating their research outcomes into novel applications in business or society. Yet, *researchers' cost-benefit assessments* frequently indicate a disinclination towards engaging in inter- and transdisciplinary collaborations due to potentially higher resource requirements for coordination, increased risk of project failure, reduced chances of securing funding or publication, and potential drawbacks in light of the prevailing approaches for evaluating scientific research impact and individual scientists' performance.

AI's role in shaping costs and benefits of opening the scientific knowledge production process

Considering the swift and transformative rise of artificial intelligence (AI), its widespread accessibility and its profound impact on the workflows in scientific knowledge production (Wang, Fu et al., 2023), we find it compelling to examine the ways in which AI is reshaping researchers' cost-benefit evaluations of opening up their knowledge production processes by engaging with external collaborators. A historical perspective reveals that analytical AI has already been altering scientific practices for decades (Gillies, 1996), particularly in the natural sciences where large amounts of data needed to be processed and analyzed to push the knowledge frontier (Wang & Barabási, 2021). While this change resulted in reduced time and resource investments required for data processing and analysis, and overall accelerated scientific

[10] In the order of appearance: Warwick Business School, Information Systems Management & Analytics Group - Copenhagen Business School, Department of Strategy and Innovation.

discovery, analytical AI also became more adept at handling complex datasets and simulations. As a result, also the costs of engaging in open scientific collaboration decreased, encouraging more scientists to join large collaborative initiatives such as the Human Genome Project (Libbrecht & Noble, 2015).

The emergence of generative AI (GenAI), however, could transform the benefits and costs of engaging in open and collaborative science practices at a much greater scale (Beck, Poetz & Sauermann, 2022). First, because its transformative potential applies to many more stages in the scientific knowledge production process than data processing and analysis. This includes ideation, literature and theory work, the development of research questions, hypotheses or proposals, the research design and the development of methods and materials, codification and writing processes, dissemination and – under certain conditions – even the data collection process itself (Wang, Lin & Shao, 2023). Second, GenAI is comparably easy to access to every scholar with an internet connection. And third, these changes in the scientific practice have expanded beyond the natural sciences and strongly also affect the social sciences and the humanities (e.g., Dell'Acqua et al., 2023).

Considering the different roles AI can play in scientific research projects may be a helpful starting point to discuss changes in the costs and benefits of engaging in open and collaborative science (Agrawal, Gans & Goldfarb, 2023; Kellogg, Valentine & Christin, 2020; Koehler & Sauermann, 2023). First, AI can take over tasks across different stages of the research process that are traditionally performed by scientists and/or their external collaborators (role of AI: *automation*). This reduces the benefits of collaborating with others on such tasks, for example, when AI assumes roles like image classification or protein structure prediction that were previously carried out by citizen scientists on crowd science platforms like Zooniverse or Foldit (Franzoni, Poetz & Sauermann, 2022; Boussioux et al., 2023). AI might also reduce the costs of filtering external knowledge and preference inputs for setting a research project's agenda, for example, when scientists engage in crowdsourcing research questions among citizens, patients or other societal stakeholders (Beck et al 2022a). On the other hand, automating simpler tasks may free-up researchers' capacities to intensively engage with external collaborators for addressing highly complex tasks, where recombining human intelligence still outperforms the capabilities of AI. Automating research tasks can, however, also increase the costs of open and collaborative science: As AI can produce new insights or data itself, determining the ownership and proper credit for AI-generated contributions as well as considering confidentiality issues might, for example, become more complex, introducing new challenges in collaborative projects and related costs for mitigating them. This can particularly be an issue when the collaborators come from different institutions or countries with varying regulatory and legal frameworks.

Second, AI can support researchers and/or their external collaborators in performing their tasks by decreasing the effort needed, enhancing the quality of outcomes, or accelerating the completion of tasks (role of AI: *augmentation*). Providing access to vast amounts of existing bodies of knowledge in diverse fields of research and practice as well as insights into and interpretations of data that might not be immediately apparent to human researchers can, for example, enrich discussions and inspire novel hypotheses in existing collaborative projects. Furthermore, it may facilitate the recombination of knowledge across disciplines both within and beyond academia, fostering research ideas that might not only be more innovative but also of greater societal relevance. AI can additionally empower a wider range of scientists, including researchers from resource-limited settings and citizen scientists, to contribute more meaningfully to scientific collaborations, and assist researchers in finding and reusing knowledge and data from diverse and distant sources and identifying collaborators from different fields within and across academia more efficiently. Moreover, AI can increase the benefits and reduce the costs of engaging in science communication activities as it aids in more effectively disseminating knowledge to non-academic audiences, for instance, by crafting easily comprehensible summaries of scientific studies for citizens, policymakers, or other stakeholder groups. In addition to supporting different tasks in the scientific research process, AI can help researchers with facilitating larger-scale collaborations more efficiently by, for example, synthesizing, integrating, and sharing distributed findings from diverse collaborators within and across academia and allocating tasks to those with the required skills or motivations (algorithmic management).

Although these factors highlight AI's capacity to amplify the benefits and minimize the costs of openness and collaboration in the scientific knowledge production process, they also hint at potential additional costs related to maintaining data quality and preventing a skills gap in collaborative projects. Collaborative science projects that rely on shared datasets might, for example, face challenges when exposing the data to AI for exploratory analysis. As ensuring the quality and reliability of AI-generated insights is crucial, collaborators might be required to establish consensus on verifying and validating AI contributions, which could introduce additional steps and complexities in the collaboration process. Furthermore, there's a risk that not all researchers will have equal access to or familiarity with AI tools, potentially creating a skills gap that could either hinder some researchers from fully participating or require substantially bigger efforts to do so.

Independently of whether AI automates or augments task of collaborators in scientific knowledge production, researchers may need to consider potential costs arising from the way AI works: When trained on biased data and powered by opaque algorithms, AI systems risk reinforcing or exacerbating biases in both, performing specific research tasks, and algorithmically managing collaborative projects. Ensuring ethical use and addressing biases in AI becomes an additional responsibility for collaborative teams, requiring vigilance and potentially more resources.

Although this discussion on the way AI potentially changes cost-benefit assessment of engaging in open and collaborative science is far from comprehensive, we hope it serves as starting point to more systematically think about how and at what stages in the process of generating and translating new scientific insight AI can increase the benefits and reduce the costs of inter- and transdisciplinary knowledge flows and collaborations, either by means of automating tasks or by augmenting human contributions (Beck, Poetz & Sauermann, 2022). Additionally, it may help understand when and why AI may even be a more effective knowledge actor than human collaborators or, on the other hand, potentially reduces the likelihood of achieving outlier creativity (Dell'Acqua et al., 2023).

Boundary conditions for AI's cost-benefit optimization in open and collaborative science

Following the preceding discussion, it is important to think about necessary boundary conditions to leverage the benefits and mitigate potential costs. To what extent researchers will be able to experience a better or worse cost-benefit ratio may depend on boundary conditions on the individual, organizational, and system level. On the *individual level*, scientists' ability to take advantage of the outputs generated by AI and to integrate them with their own knowledge might depend on their "cognitive complexity", i.e., their individual ability to understand the world in more complex ways, to internalize knowledge from multiple fields of science, and to observe and understand the connections between phenomena in different fields (Hollingsworth, 2007). Scientists with higher levels of cognitive complexity might be more likely to be able to connect to and internalize diverse and potentially distant AI outputs (Jia et al., 2023). Also, it is likely that scientists with a proficiency in an AI's operational intricacies and foundational mechanisms can be more critical towards AI-generated outputs (Wang, Fu et al., 2023), thus preventing them from falling for GenAI's "hallucinations" or pursuing paths that are based on flawed or incomplete data (Lebovitz, Lifshitz-Assaf & Levina, 2022).

Furthermore, the adoption of new practices critically depends on being considered legitimate (Bitektine & Haack, 2015). Whether or not individual researchers judge a new practice to be legitimate in a given setting (e.g., using AI for open and collaborative research) depends on their perceived propriety, i.e., the strategic importance and value complementarity of a new practice for achieving their goals, as well as the perceived validity, i.e., the perception that key social referents (e.g., funding organizations, peers, policymakers) regard the new practice as desirable by interpreting validity cues (Jacqueminet & Durand, 2020). The lack of either of the two legitimacy dimensions will increase the (perceived) costs and decrease the (perceived) benefits. For instance, scientists may fear reputation or even career threats if using GenAI for collaboratively working with citizens on developing hypotheses for a research project, if this application of GenAI is not considered desirable by their promotion committee or funding organization.

On the *organizational level*, access to resources and support structures might be particularly relevant boundary conditions (Beck, LaFlamme & Poetz). To utilize AI effectively, resources are needed to provide appropriate training for researchers, access to licenses, or potential investments in adapting easily accessible AI tools such as ChatGPT or using more specialized ones as well as the computational power required to operate them. Researchers with access to these resources will be able to leverage the capabilities of AI and may embrace the possibilities for improving the cost-benefit ratio of engaging in open and collaborative practices better than those without access, as their (perceived) costs decrease. In a similar vein, organizational support structures have the potential to alleviate the associated costs of using AI for engaging in open and collaborative research. For example, the use of AI tools for inter-organizational collaborations likely requires formal agreements for data usage, etc., increasing the costs of such a collaborative endeavor. If a dedicated support service is empowered to handle the necessary formalities, researchers will be able to increase their (perceived) net benefit.

Finally, researchers across disciplines are discussing how regulatory frameworks on the *systems level* may influence the use of AI in science (Birhane et al., 2023). Such frameworks related to, for example, intellectual property rights, data protection and data reuse, the availability of human-generated training data and the pace of technological advancements will influence whether the use of AI increases or reduces the benefits and costs of engaging in open and collaborative science practices.

References

Agrawal, A., Gans, J. S., & Goldfarb, A. (2023). Do we want less automation? *Science*, 381(6654), 155-158.

Beck, S., Bergenholtz, C., Bogers, M., Brasseur, T. M., Conradsen, M. L., Di Marco, D., ... & Xu, S. M. (2022). The Open Innovation in Science research field: a collaborative conceptualisation approach. *Industry and Innovation*, *29*(2), 136-185.

Beck, S., Brasseur, T., Poetz, M., & Sauermann, H. (2022a). Crowdsourcing Research Questions in Science. *Research Policy*, *51*(4), 104491.

Beck, S., LaFlamme, M., & Poetz, M. (2022). Organizing for inter- and transdisciplinary collaboration in science: The role of autonomy and control. *Academy of Management Proceedings 82*. Beck, S., Poetz, M., & Sauermann, H. (2022). How will Artificial Intelligence (AI) influence openness and collaboration in science?. *Industry and Innovation*, *29*(2), 136-185.

Boussioux, L., N Lane, J., Zhang, M., Jacimovic, V., & Lakhani, K. R. (2023). The Crowdless Future? How Generative AI Is Shaping the Future of Human Crowdsourcing. *HBS Working Paper*.

Birhane, A., Kasirzadeh, A., Leslie, D., & Wachter, S. (2023). Science in the age of large language models. *Nature Reviews Physics*, *5*(5), 277-280.

Bitektine, A., & Haack, P. (2015). The "macro" and the "micro" of legitimacy: Toward a multilevel theory of the legitimacy process. *Academy of Management Review*, *40*(1), 49-75.

Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., ... & Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*, (24-013).

Franzoni, C., Poetz, M., & Sauermann, H. (2022). Crowds, citizens, and science: a multi-dimensional framework and agenda for future research. *Industry and Innovation*, *29*(2), 251-284.

Gillies, D. (1996). Artificial intelligence and scientific method. Oxford University Press.

Huang, C. K., Neylon, C., Montgomery, L., Hosking, R., Diprose, J. P., Handcock, R. N., & Wilson, K. (2024). Open access research outputs receive more diverse citations. *Scientometrics*, 1-21.

Hollingsworth, J. (Ed.) (2007). High cognitive complexity and the making of major scientific discoveries. SAGE Publications Ltd.

Jacqueminet, A., & Durand, R. (2020). Ups and downs: The role of legitimacy judgment cues in practice implementation. *Academy of Management Journal*, *63*(5), 1485-1507.

Jia, N., & Luo, X., Fang, Z., & Liao, C. (2023). When and How Artificial Intelligence Augments Employee Creativity. *Academy of Management Journal*. 10.5465/amj.2022.0426.

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, *14*(1), 366-410.

Kennedy, A. (2006). Interstellar travel: The wait calculation and the incentive trap of progress. *Journal of the British Interplanetary Society*, *59*(7), 239-246.

Koehler, M., Sauermann, H. (2023). Algorithmic Management in Scientific Research. SSRN. *ESMT Berlin Working Paper*, available at SSRN: https://ssrn.com/abstract=4497871.

Lebovitz, S., Lifshitz-Assaf, H., & Levina, N. (2022). To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. *Organization Science*, *33*(1), 126-148.

Libbrecht, M. W., & Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nature Reviews Genetics*, *16*(6), 321-332.

Mazzucato, M. (2018). Mission-oriented innovation policies: challenges and opportunities. *Industrial and corporate change*, *27*(5), 803-815.

Park, M., E. Leahey, R.J. Funk. 2023. Papers and patents are becoming less disruptive over time. *Nature*, *613*(7942), 138–144.

Poetz, M., Beck, S., Grimpe, C., & Sauermann, H. (2024). *Chapter 27: Open Innovation in Science*. In Chesbrough, H., Radziwon, A., Vanhaverbeke, W., & West, J. (Eds.). The Oxford Handbook of Open Innovation. Oxford University Press.

Probst, B., Lohmann, P. M., Kontoleon, A., & Anadón, L. D. (2023). The impact of open access mandates on scientific research and technological development in the US. *Iscience*, *26*(10).

UNESCO (2022). Understanding open science. URL (accessed March 2024). https://unesdoc.unesco.org/ark:/48223/pf0000383323

Vicente-Saez, R., & Martinez-Fuentes, C. (2018). Open Science now: A systematic literature review for an integrated definition. *Journal of Business Research*, *88*, 428-436.

Wang, D., & Barabási, A.-L. 2021. *The Science of Science*. Cambridge: Cambridge University Press.

Wang, H., Fu, T., Du, Y., Gao, W., Huang, K., Liu, Z., ... & Zitnik, M. (2023). Scientific discovery in the age of artificial intelligence. *Nature*, *620*(7972), 47-60.

Wang, X., Lin, X., & Shao, B. (2023). Artificial intelligence changes the way we work: A close look at innovating with chatbots. *Journal of the Association for Information Science and Technology*, *74*(3), 339-353.