

Dirassat & Abhath
The Arabic Journal of Human
and Social Sciences



مجلة دراسات وابحث
المجلة العربية في العلوم الإنسانية
والاجتماعية

EISSN: 2253-0363
ISSN : 1112-9751

Ethical and Societal Implications of AI in Science

التطبيقات الاخلاقية والاجتماعية للذكاء الاصطناعي في العلوم

Benfareh Aicha

Phd , University centre of Ali Kafi Tindouf ,departement of management, the laboratry for
sustainale local development and entrepreneurship ,Algeria, benfarehaicha@gmail.com

بن فارح عائشة

دكتوراه، المركز الجامعي علي كافي تندوف، قسم علوم التسيير، المخبر الاستدامة التنمية المحلية والمقاولاتية،

benfarehaicha@gmail.com

تاريخ القبول : 11 سبتمبر 2024

تاريخ الاستلام: 28 جوان 2024

Abstract:

The ethical and societal implications of AI in scientific research are vast and will only continue to grow. In this perspective article, we have attempted to outline key challenges and opportunities. Importantly, these are not only challenges for AI itself, if AI is to be used in science, but also challenges for the field of AI in order to make contributions to science and the world in a responsible and successful manner. We emphasize the need not only for the responsible communication of AI results, but also for the development of agreed-upon community standards and metrics for evaluating the ethical implications of our own work. This discussion can be helpful to guide future efforts to bring about mutual benefit in the co-evolution of AI and science.

Keywords: Ethical, AI, scientific research

الملخص:

فالآثار الأخلاقية والمجتمعية التي تترتب على هذه الأنشطة في مجال البحث العلمي هي آثار واسعة النطاق و تستمر في النمو. وفي هذه المادة من هذا المنظور، حاولنا أن نحدد التحديات والفرص الرئيسية. ومن المهم أن هذه التحديات لا تشكل تحديات بالنسبة إلى الذكاء الاصطناعي نفسها فحسب، إذ هذا الأخير يستخدم في العلوم، بل أيضاً تحديات بالنسبة إلى ميدان الذكاء الاصطناعي من أجل الإسهام في العلوم وفي العالم بطريقة مسؤولة وناجحة. ونشدد ليس فقط على ضرورة الإبلاغ المسؤول عن نتائج الأنشطة المنفذة تنفيذاً مشتركاً، بل أيضاً على ضرورة وضع معايير ومقاييس مجتمعية متفق عليها لتقييم الآثار الأخلاقية في استخدام الذكاء الاصطناعي في العلوم والبحث العلمي خاصة. حيث تساعد هذه المناقشة في توجيه الجهود المقبلة لتحقيق المنفعة المتبادلة في التطور المشترك للعلم والتكنولوجيا .

الكلمات المفتاحية : الأخلاقيات، الذكاء الاصطناعي، البحث العلمي

1. INTRODUCTION

What are the societal values that we seek to pursue in our work? We refer to the National Academy of Science, Engineering and Medicine's "Proposed Statement of Principles for the Evaluation of Research Proposals on the Implications of Research," which identifies seven categories of values to As the capabilities of artificial intelligence (AI) in science grow, AI systems are capable of performing innumerable tasks once done by humans, from tasks such as data analysis or visualization on the business end of science to the promotion or review of scientific results, or even to the conduct of experiments and the definition of research programs on the front end. In this paper, we address the ethical implications in support of the pursuit of societal values stemming from funding AI research for science-related effects rather than for commercial gain or national security. Our hope is that by clearly articulating these ethical implications, we can build consensus for action in the form of institutional and organizational best practices, policy and regulation within scientific and funding communities, possibly leading to changes to institutional and organizational missions and the identification and funding of new sources of revenue. (Korteling et al.2021)

Concepts to meet new challenges, 6) enhancing scientific experiments and collections, and 7) other values such as

scientific training, infrastructure, the international character of science, and broadening participation in science.

2. Definition and Scope of AI

Artificial Intelligence (AI) is a branch of computer science which attempts to create intelligent systems. From a scientific perspective, an "intelligent system" in this context is one that learns from input data by adapting it to produce generally accepted results, connecting this to scientific domains. At present, the vast majority of AI applications employed worldwide are based on data ("data-driven modeling") like that employed by modern statistical methods, such as machine learning and data mining. However, in addition to the use of data, the construction of AI systems also takes into account the manipulation of different elements of knowledge that reinforce the system's general problem-solving competence and common sense, related to knowledge representation and reasoning. Although some of the most important practical achievements in machine learning are related to pattern recognition, there are different tasks that can also be related to it. (Ghahramani et al.2020)

In this report, we attempt to analyze and discuss several sets of challenges that may be relevant to the ethical and societal implications of AI in particular areas of scientific research and the societal systems in which these may be embedded. This is an interdisciplinary question and should be approached as such. We should start by understanding the key concepts of this

synthesis, which are the scientific research process (covered in a few general models of the scientific research process), its development through data management and scientific software, and societal impacts of the science and the research process. Then, the current understandings of ethics, values, and associated ethical concepts are briefly introduced, and we address ethical frameworks that seem most relevant in evaluating the ideas hereto, in particular ethical standards in science. From these established basics, the report then briefly defines AI. After this part of the synthesis, we are ready to analyze the ethical and societal implications of AI taking different dimensions and paths.

3. Integration of AI in Scientific Research

Scientists have already begun integrating AI systems into their research methodologies, using them to conduct powerful data analyses and discover patterns in data that human researchers would be unlikely to notice. For example, in her role as AI for Earth lead at Microsoft, Lucas Joppa is using AI to help researchers turn satellite data into actionable intelligence, enabling scientists to monitor the world more effectively and track global changes across diverse systems such as biodiversity, agriculture, water, and urban systems. As another example, Eric Horvitz, Technical Fellow and Director at Microsoft Research, uses machine learning to mine the scholarly literature and identify important challenges that have not yet been addressed. Together with colleagues from academia and government, he, in turn, tries to stimulate research to address those gaps through seeding efforts and calls to action tailored to the critical knowledge gaps identified. (Xu et al.2021)

4. Ethical Considerations in AI Research

For the societal implications of science, we have many established professional bodies and norms, such as the Committees on Publication Ethics (COPE) and the US Office of Research Integrity, which self-correct with much effectiveness but are often driven by policies and guidelines more than by law. Publication ethics has been a thorny problem for over a century, and questions about fraud, reproducibility, and so on remain, especially in pressing publicism races that accompany the modern economy, but most professionals understand and buy into sustained arguments for promulgating and adhering to sound publication policies. Many of the recent reflections of the scientific community about these central issues of publication and researcher support have also made the importance of creating mechanisms that allow not only for professional self-regulation in staffing honest innovative inquiries but for the choice of research policy experts with a wide variety of experiences, who will likely originate from many distinct groups and institutions. (Armond et al.2021)

The development of artificial intelligence poses substantial challenges for the responsible and ethical conduct of science. Key questions now arise about issues such as research transparency, bias, and the consequences of AI research for society. Because AI systems commonly have a large number of moving parts and elements of randomness, it can be difficult to know from a given AI outcome what caused that result, or how it would generalize across a wide variety of situations. This can be especially problematic for the further development of science. As in science, transparency and

reproducibility will again be issues. And as in science, decisions based on AI matter for individuals and groups in society, and will affect their trust in the results and the methods that produce them.

5. Bias and Fairness in AI Algorithms

The fairness in machine-learning systems that serve an informational or advisory role should be evaluated on how well they guide broader society to help mitigate the societal issues that ultimately cause the bias. The underlying root cause of the algorithm's failure might lie in the inherent bias in society with respect to the predicted target, which would call into question the decision to rely on an advisory AI system. (Mehrabi et al.2021)

The fairness literatures in both AI and science should not only work to reduce the negative effects of AI but also push the envelope to use AI to increase the positive effects, including accessibility and potential benefits to underrepresented groups. In some cases, fairness techniques can correct for societal bias in the underlying AI data. However, fairness sometimes hides more impactful changes in the AI system. (Leavy et al., 2020)

As AI becomes a natural part of scientific discovery, we need to make sure everyone has access to and can participate in the discovery both using AI and through collaboration with AI. One area of focus for AI fairness research in science will be the use of fairness only to mitigate the negative effects of AI systems. In some cases, biased AI may be the only tool available for a particularly difficult problem.

Pre-processing methods transform the data such that the fairness of the classifier can be

increased without changing the classifier. In-processing methods alter the learning procedure such that the output classifier is fair with respect to some fairness definition. Post-processing methods learn a machine-learning model using any empirical loss and then apply a fairness intervention on the learned model without changing it.

The first step toward building fair AI is to measure fairness in terms of equal treatment, equal outcomes, or equal opportunity. Based on definitions of fairness, many algorithmic fairness definitions have been proposed. These can be broken down into three groups: pre-processing methods, in-processing methods, and post-processing methods.

Bias and fairness in AI algorithms are crucial concepts, especially when AI algorithms are used for decision making. A biased decision-making system could have life-changing consequences. In the case of science and research, biased AI algorithms can slow down collaborations, impede scientific discovery, and, in the worst-case scenario, exclude primary contributors from the research team

6. Privacy Concerns in Data Collection and Analysis

The privacy and associated implications in modeling of AI based methods, say for example, in probabilistic graphical models, for individual-level predictions or for individual risk can be quite daunting. For instance, consider a model being trained to predict a disease using EMR data. There are three plausible scenarios. One, wherein all the EMR data and the disease labels are available for all patients. Two, wherein

disease labels are not available for a few patients (in a big data context, small in number). Three, wherein the existence of the patient, and his/her demographic features is known, but the patient's EMR features are missing. The use of AI in the former two scenarios is fairly straightforward and acceptable and does not inherently have issues on privacy of the patient. However, ethical and moral implications arise when the model considers only demographic data of the patient and says something about the trained model for such individuals as a function of the diagnostic model. Given the collected EMR observations, how does one shield from the risk of violation of privacy and the moral harm that the individual (or maybe his/her family) may face? The potential risk of being identified (by any means) from the population given the observed disease probabilities, especially in the case of rare diseases, could cause huge distress to the patients and such distress can largely compromise the patient trust on the healthcare system. (Clark et al.2020)

The privacy of the participants in any research is of utmost importance and has been a matter of significant debate recently in various domains. Especially with the kind of data being collected about individuals and its usage for analysis using AI techniques, the issue has grown more ominous. In this context, the personal healthcare data that are collected as a part of epidemiological studies need explicit mention. Anonymization and selective dissemination of sensitive but important healthcare data collected in epidemiological studies and stored in university and public databases, while ensuring that privacy of the study participants is maintained, thus represents an important challenge. Liability protection and public trust is also crucial. The futures of

such databases in the age of AI cannot be underestimated, considering the enormous benefits that healthcare stakeholders (public) and researchers can derive from these databases – for example, the two-way learning from rare complications of a disease that would be beneficial to both the society and researchers. (Hensen et al.2021)

7. Societal Impact of AI in Science

3.3 How Ethical Principles Apply to Science
Belonging Science, technology, and innovation operate in complete isolation from societal norms. In convergent multidisciplinary and trans disciplinary research, interrelations become more evident. Maintenance of quality with respect to cultural traditions is an attribution to membership requirements of the concerned scientific agencies. Liberties, equality, and human rights combine with ideas of justice for the governance of technological infrastructures. Direct and indirect obligations of protection but also empowerment of civil rights has to coexist. The laws governing science directly and indirectly that address these needs should comply with societal expectations concerning fundamental ethical principles. Political acceptance opens debate, and scholarly activists can bring this concern into informed discourse. (Beck et al.2022)

3.2 Science Justice
As values, intentions, behaviors, capabilities, concerns, and outcomes in science, technology, and innovation are ethical matters, ethical belonging to the concept of Artificial Intelligence with applications in scientific research are considered here. The scientific research collaborations that we commit to, or become shareholders, designers, customers, employers, financiers, incubators, operators, or users of, have roles

to be filled based on the application of the principles of prudent scientific research. Positive and negative aspects of formulated policies and codes of conduct reveal the impact of science on society. The underlying context leading to society's caution is unknown. Ethical issues are considered here and a science justice position is put to establish control, privacy, security, confidentiality, integrity, transmissibility, adequacy, retention, disposal, and availability. AI technologies should not be misused; the world's public bodies also have ethical obligations to embody and achieve standards for social responsibility within research, regulation, and risk management processes affecting those affected by artificial intelligence's capabilities. Crescent and Njie describe human rights, dignity, and solidarity as means of achieving International Law's responsibility for assessing the AI risk-reduction capabilities of public and private bodies responsible for policing 'governing' AI. (Orr & Davis)

3.1 Sociology of Scientific Research with AI
Europol has identified the potential misuse of digital and physical robots, or agents acting on behalf of human handlers, to conduct illegal activities. Waste, through failed scientific projects, loss of confidence or acceptance, and misapplication of knowledge can occur if AI is used to conduct scientific research. With the formalization and standardization of the scientific research process, there has been consistent criticism of its over application. The use of heuristics, or instruments of data collection that are debatable, have been identified. However, their purpose is to conduct the scientific research more cost effectively. Only when working with reproducible knowledge, such as empirically derived engineering know-how, can the

lessons of the scientific method be restricted. The prospects of utilizing AI in scientific research, alongside established models of scientific research, are discussed here. Societal concerns are fundamental; they guide how much work should be done by experts, or lay members acting for the common good, and in what context. As organizational, ethical, legal, and social considerations, the English National Grid and Smart Meter policies suggest lay member roles to support energy needs can arise. Timo Kivimaki has already identified that our surveillance institution—the 'State'—needs cooperation to conduct legitimate surveillance. (Baltrūnienė, 2022)

8. Job Displacement and Economic Inequality

Many human employment activities have the potential for machine automation. The process of implementing new AI systems begins with computational analysis, progresses to understanding and addressing targeted tasks, and then applies task-interpreted improvements in various healthcare, agricultural, service, and manufacturing industries at the levels of devices, processes, and systems for utilization by companies, entrepreneurs, individuals, and various tiers of civil society. The power of technological disruption is beginning to strain the ability of societies to react and recover with minimal impacts. Impact on local and global economies causes disruption and displacement that need to be harmonized in some way such that perceived fairness, social benefits including renewed social balance and retraining, and efficiency outweigh potential harm to those affected. The deployment of machines and the effects of job substitution are relatively well understood, but the pace

of technological change is often underestimated, leading to scenarios where integrated solutions are lacking, resulting in infrastructural, economic, and societal inadequacies that stall society's advancement. (Tschang & Almirall, 2021)

The global race to maintain scientific and technological leadership is continuing to gain steam and will likely intensify over upcoming decades. Individuals must equip themselves with both the skills of the future as well as an understanding of how and where AI tools can be deployed to maintain employability. Countries that fall behind will likely face a reduced level of collective well-being, much like the perils confronting the industrialized middle class in the latter decades of the twentieth century. The stakes are high and are likely to become even higher by the conclusion of the twenty-first century. In developing countries, the absence of these capabilities will likely continue to reinforce dependence on external assistance, a situation that runs parallel to many humanitarian organizations, where AI tools can be used to break the cycle of dependence and enhance governance, agriculture, and political stability. Strategic focus is necessary to calibrate progress on tools that match in

9. Access to AI Technologies in Developing Countries

So, how could this be addressed? Should aid be given? Aid can be used for the purchase of advanced technologies developed in the north but aid donors would rather "invent" something suited to the local needs or create indigenous expertise able to participate in this invention. Yet, while architectural principles of real relevance to critical sectors such as education, emergency intervention, public health, or agricultural development

are well within the scope of present-day AI technology analyses and design, little activity is observed in this direction. Not coincidentally, these are also sectors where public-sector adoption is highly desirable because motivations do not tend to be "money first" or to provide services only to the most affluent classes in the population. Furthermore, regardless of the sector concerned, these are also precisely the sectors in developing countries where affordability tends to be a major objection; substantial public subsidy will be a must.

The global south can benefit from AI in several ways: the development of appropriate AI technologies which would harness the creativity of local software developers, appropriate technological capacity building strategies, bridging the yawning AI infrastructure gap, global ethical collaboration involving both the north and the south, and finally embracing the public sector use of AI systems by ensuring their ethical acquisition and deployment, therefore avoiding the pernicious use of AI technologies in the south. Developing countries continue to lack access to existing AI technologies notably because of lack of internet and digital infrastructure. Many commercial AI systems are not tailored to local needs, and small-scale innovative and necessary applications of AI technologies are unaffordable. A vicious circle is initiated, technique required expertise becomes rare because of lack of access to modern AI technologies. The youth in developing countries get deprioritized in the relevant curricula and educational systems as they get swallowed up by the masses of more traditional technology users. (Mhlanga, 2021)

10. AI Regulation and Governance

Fortunately, the intersection of these challenges has gathered the attention of some US regulatory agencies, researchers, and the broader international community. They already house many of the experts and decision-makers with the ability to guide the improvement and responsible realization of AI. Nevertheless, we believe a deeper sustained effort is necessary at the governance level, and we support the call for the establishment of a critical body of experts, representative and transparent in order to clarify, frame, and ensure the responsible and convergent development of AI. In the global effort towards coherent AI development, collaboration and alignment of efforts will be crucial. This international coordination should not preclude room for adaptation, growth, and differentiation towards achieving objectives that are shared by a majority of actors in countries and regions and that could contribute to the development of AI not simply as a local product of UX but as a global common good. (Smuha)

The novelty of AI is confronting developers and regulators with challenges such as those we have just described - robustness, transparency, privacy, etc. - that are common to any new technology development. Although it shares these challenges with prior technology developments, the nature of the AI and its presence in the virtual and physical space will deepen their reach and urgency

10.1. Current Regulatory Frameworks in AI

Apart from that, existing laws often do not fit AI well, as they may not even recognize non-human agents as possible transgressors, thus making it next to impossible to assign liability. The current lack of regulation

poses three major issues. Firstly, these tools can be used to exploit individuals, to reduce privacy of individuals or to impair the social fabric. Secondly, these tools can coerce or severely manipulate individuals. Thirdly, these tools at the same time do not have the ability to augment society's democratic mechanisms or institutions.

In the first place, regulators often do not have sufficient information to assess or control the potential harms resulting from the widespread deployment of AI systems. Our societies currently do not have the tools to monitor and evaluate the proliferation and widely ranging impacts of AI systems. Even experts have difficulty providing a comprehensive overview at the present time. Moreover, steering sociotechnical systems with fast-evolving, virtually autonomous components is a challenge by itself. Closer observation and control may not be possible.

10.2. Proposed Guidelines for Ethical AI Development

1. Do not harm: Artifact development in and of itself poses no direct threat to humanity or the research area as applied generally, only the implementation of the technology within research applications can. Any technology developed following the above guidelines cannot have malicious properties. In fact, promoting research that meets these guidelines can lead to the development of systems designed to detect and prevent those threats that are malicious or being used to issue harm. Before any implementation, such systems should undergo the standard verification and validation process for application in the area of study.

As the impact of AI technologies has grown, so too have questions about the ethical implications of their prevalence and use. In

the research realm, AI technologies are increasingly deployed in sensitive application areas such as criminal risk assessments, loan approval processes, and future labor market predictions. Developing and deploying these systems responsibly and ethically is essential. In this work, we discuss how some of these AI technologies can be used to support good practices in science, but their effects must be carefully considered. Similar principles can be applied to many fields of research, as similar effects arise from pervasive technologies. Thus, the proposed guidelines for AI research can be extended to a wide range of research topics. (Dwivedi et al.2021)

Conclusion and Future Directions

The AI applications for scientific research are an exciting and powerful area of development, promising major breakthroughs across a broad range of problems. These applications raise issues compelling the AI community to engage the ethical and societal implications of our work. However, ensuring that broad benefits result from AI applications in science will require explicit discussions of these issues and deliberate actions. We believe much progress can be made by developing agreed-upon community standards and metrics for evaluating the social and ethical implications of our work. In this paper, we summarize some of the ethical and societal implications of AI in the context of scientific research.

Referrals And References:

Korteling, J. H., van de Boer-Visschedijk, G. C., Blankendaal, R. A., Boonekamp, R. C., & Eikelboom, A. R. (2021). Human-versus artificial intelligence. *Frontiers in artificial intelligence*, 4, 622364. [frontiersin.org](https://www.frontiersin.org)

Ghahramani, M., Qiao, Y., Zhou, M. C., O'Hagan, A., & Sweeney, J. (2020). AI-based modeling and data-driven evaluation for smart manufacturing processes. *IEEE/CAA Journal of Automatica Sinica*, 7(4), 1026-1037. [\[PDF\]](#)

Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., ... & Zhang, J. (2021). Artificial intelligence: A powerful paradigm for scientific research. *The Innovation*, 2(4). [cell.com](https://www.cell.com)

Armond, A. C. V., Gordijn, B., Lewis, J., Hosseini, M., Bodnár, J. K., Holm, S., & Kakuk, P. (2021). A scoping review of the literature featuring research ethics and research integrity cases. *BMC Medical Ethics*, 22(1), 50. [springer.com](https://www.springer.com)

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM computing surveys (CSUR)*, 54(6), 1-35. [\[PDF\]](#)

Leavy, S., O'Sullivan, B., & Siaper, E. (2020). Data, power and bias in artificial intelligence. *arXiv preprint arXiv:2008.07341*. [\[PDF\]](#)

Clark, A., Jit, M., Warren-Gash, C., Guthrie, B., Wang, H. H., Mercer, S. W., ... & Jarvis, C. I. (2020). Global, regional, and national estimates of the population at increased risk of severe COVID-19 due to underlying health conditions in 2020: a modelling study. *The Lancet Global Health*, 8(8), e1003-e1017. [thelancet.com](https://www.thelancet.com)

Hensen, B., Mackworth-Young, C. R. S., Simwinga, M., Abdelmagid, N., Banda, J., Mavodza, C., ... & Weiss, H. A. (2021). Remote data collection for public health research in a COVID-19 era: ethical implications, challenges

and opportunities. Health policy and planning, 36(3), 360-368. [nih.gov](https://pubmed.ncbi.nlm.nih.gov/)

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. [tandfonline.com](https://www.tandfonline.com)

Orr, W. & Davis, J. L. (). Attributions of ethical responsibility by Artificial Intelligence practitioners. *Information*. [academia.edu](https://www.academia.edu)

Baltrūnienė, B. (2022). Place of artificial intelligence in the detection and investigation of crime: the present state and future perspectives. *Problemy Współczesnej Kryminalistyki*. [academicon.pl](https://www.academicon.pl)

Tschang, F. T. & Almirall, E. (2021). Artificial intelligence as augmenting automation: Implications for employment. *Academy of Management Perspectives*. [smu.edu.sg](https://www.smu.edu.sg)

Bongomin, O., Nganyi, E. O., Abswaidi, M. R., Hitiyise, E., & Tumusiime, G. (2020). Sustainable and dynamic competitiveness towards technological leadership of industry 4.0: implications for East african community. *Journal*

of Engineering, 2020(1), 8545281. [wiley.com](https://www.wiley.com)

Mhlanga, D. (2021). Artificial intelligence in the industry 4.0, and its impact on poverty, innovation, infrastructure development, and the sustainable development goals: Lessons from Sustainability. [mdpi.com](https://www.mdpi.com)

Smuha, N. A. (). From a 'race to AI' to a 'race to AI regulation': regulatory competition for artificial intelligence. *Law*. [kuleuven.be](https://www.kuleuven.be)

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... & Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57, 101994. [openrepository.com](https://www.openrepository.com)