

Author(s): Seraya Maouche

AI-enhanced coding in bioinformatics: legal and ethical considerations

Abstract:

In this article, we describe a case study that explores the ethical and legal issues relating to the introduction of Artificial Intelligence (AI)-enhanced coding in bioinformatics and related life science research. Using this case study, we highlight the potential dangers posed by the introduction of AI-assisted coding in programming and analysis of health data. The aim is to understand and consider the potential harms it poses and to help students and young researchers on how to use AI responsibly in their work. Recent developments in generative artificial intelligence (Gen AI) and the emergence of Large Language Models (LLMs)-based chatbots such as Chat Generative Pre-Trained Transformer (ChatGPT) launched by OpenAI on November 30, 2022 are currently matter of many debates mainly about AI-generated scientific research and publications. Several scientists, editors, and publishers disapproved that ChatGPT made its way into scientific production by being listed as a co-author. Programming is another domain where LLMs-based chatbots have proven immense potential. These AI systems have the ability to assist human programmers at different stages, including writing and debugging code. The rapid development of AI and related emerging technologies and its wide deployment in different domains, including life-science research gives rise to multiple ethical, legal and technical considerations. We designed the present case study to describe a plausible situation in biomedical research, and to elucidate some legal and ethical issues resulting from the introduction of AI in life-science research.

Agenda:

Case description	2
Questions	2
Exercises	3
The principles of AI ethics	3
Nonmaleficence.....	4
Transparency & explainability.....	4
Accountability	4
Analysis and reflections	5

Author(s):

Seraya Maouche^{1,2}, researcher in translational bioinformatics & engineer in computer sciences.

- ¹ Translational bioinformatics and emerging technologies in healthcare R&D. TBi Scientific®. 15 rue Lauriston Paris 75016, France
- ² Ethique & Intégrité. 6 rue de Douai 75009, Paris. France
- ✉ seraya.maouche@tbiscientific.fr

Case description

In 2022, the bioinformatic researcher, Ada¹, joined the MatildaGC², a large scale European-funded life-science project dedicated to the genetics of complex diseases. MatildaGC is a multicentre project in which patients' data is collected in hospitals based in five EU countries. Bioinformatics is an interdisciplinary field of science that involves developing algorithms, methods, and software tools to analyze biomedical data, particularly answering key questions in the field of life-science research. Recent developments in different emerging technologies (*Maouche, 2019, p. 447*) such as Big Data, Cloud computing and Artificial Intelligence (AI) are upsetting the field of bioinformatics and the whole of life-science research. Ada's tasks include writing code for the preprocessing and analysis of data of multiple natures: genetics, transcriptomics, proteomics, demographic such as gender and age, family medical history, laboratory test results, diagnostics, treatments, medications, and electrocardiograms. As a co-author on the publications resulting from this project, Ada must participate in the writing of the manuscripts, in particular detailing its data analysis method in the "materials and methods" section. Patients' data is anonymized, safeguarded, and stored in hospitals. For analysis purposes, this data is remotely accessed by Ada and transferred to a private cloud where the code is generated and run.

Without informing the project coordinator, the five principal investigators (PIs), or the ethical authorities, Ada decided to adopt an AI-enhanced coding strategy to accelerate code generation and data analysis. Integrating AI-based tools into the process of software development is often referred to as "AI-enhanced coding" or "AI-assisted coding". Trained using a vast corpus of texts and code, LLMs-based chatbots have the ability to engage in text-based conversations and provide accurate answers, including generating code. An AI-enhanced coding system aims to assist human programmers at different stages, including writing and debugging code. Without doubt, code translation is the most powerful ability of such AI tools. Programmers could get assistance to translate code between different programming languages to accelerate or facilitate cross-platform development. There are several prominent AI-assisted coding tools which are available today. These tools offer significant benefits in terms of productivity and efficiency. A few weeks later, one of the hospitals involved in the MatildaGC project was a target of ransomware cyberattack and had to pay hackers two million euros in Bitcoin to unlock the patients' data blocked by the attack. Surgery operations and patient care were delayed as the medical staff was unable to access the Hospital Information Systems (HIS). Was there a connection? Could Ada's AI-assisted code could have included malicious code?

Questions

1. Regardless of how much anonymization and pseudonymization of patients' personal data are performed with respect to ethical standards, can we predict all the potential security and data privacy harms that may occur when code is generated using an AI-based tool?
2. Can AI-enhanced coding tools write malicious code that can leak sensitive data or include a block of harmful code which could create vulnerabilities that attackers can exploit to perform ransomware cyberattacks, denial-of-service (DoS) or delete critical data? What might programmers do to avoid these risks?

¹ The choice of the first name "Ada" is a tribute to Ada Lovelace, the first programmer in history.

² This case study is semi-fictional, it was inspired by the author's work within the EU-funded Cardiogenics project (Maouche & Schunkert, *Arterioscler Thromb Vasc Biol.* 2012 & Schunkert et al. *Nat Genet.*, 2011).

3. Should an area of research which involves the use of sensitive data exclude any use of AI-enhanced coding systems?
4. Who should be held accountable? Should this accountability be extended to the individuals or the firms who develop AI systems such as OpenAI?
5. Who is responsible when things go wrong with AI systems within a large-scale multicentre project?
6. How can hospitals and other health stakeholders reassure people about the protection of their family's medical privacy given such incidents?
7. Is there a need for new privacy regulation and ethical standards that can balance privacy against concerns of efficient research for public benefit?
8. Should the AI-based system which assisted bioinformaticians in code writing and debugging process be cited as a co-author or at least acknowledged on scientific publications in which the code has been used?

Exercises

1. Imagine that you are a PhD candidate or a bioinformatics researcher in a multidisciplinary team. Establish a list of the main ethical issues related to the use of AI-enhanced coding in biomedical research that you will face.
2. Suppose that you are a coordinator of an EU-funded large scale biomedical research project. Develop a guideline specifying how privacy and data protection must be promoted and establishing a framework to avoid vulnerabilities to cyberattacks resulting from AI-enhanced coding tools used by bioinformaticians.
3. Imagine that you are an editor and you are tasked by a publisher of a scientific journal to update the editorial policies. Add a section to the "Guide to authors" about AI authorship and the use of Large Language Models (LLMs) such as ChatGPT in writing, and the use of generative AI to produce code, figures, images, videos or graphics. Include instructions on how the use of AI-enhanced coding should be properly and transparently documented in the Methods section. Provide explanations and ethical restrictions related to responsibility and accountability for the work submitted to a scientific journal to explain why ChatGPT or similar tools can or cannot currently satisfy the journal authorship criteria.

The principles of AI ethics

The Principles of AI ethics are intended to guide how we design, develop, build, deploy, use, and manage AI-based systems and related technologies. Drawing from UNESCO's Ethics of AI recommendation (UNESCO, 2021, p. 7-10), important principles include transparency, accountability, justice, explicability, respect for autonomy, beneficence, nonmaleficence, fairness, privacy protection and security, contestability, and human-centered values.

"While AI-assisted programming tools have been validated for their positive effects on learning and applying programming, societal concerns have emerged regarding the use of such tools. These include complex ethical issues encountered during the use of AI-assisted programming tools, problems related to users' over-reliance, and the sustainability of programming education." (Zelin, 2024, p. 2)

At least three principles of AI ethics are at issue in this case: nonmaleficence, transparency & explainability (T&E), and accountability:

Nonmaleficence

By using an AI-enhanced coding strategy, it is impossible to exclude that Ada will allow harm to be caused, not only to the patients recruited within the MatilgaCG project (disclosing sensitive patients' data,..) but also to the whole healthcare system in the five hospitals involved in this project. It is inevitable to imagine a scenario when data could be corrupted by an external party, including by cybercriminals. A ransomware attack on one of the five hospitals is indeed a violation of privacy in several ways. Patient data, including medical records, personal information, and financial details, may be exposed to the attackers. These last gain unauthorized access to the hospital's systems, which often contain sensitive patient information. It breaches the trust patients place in healthcare institutions to keep their information private and secure. These privacy violations can have serious consequences for both patients and the healthcare institution. In addition, this situation breaches the trust patients place in healthcare institutions to keep their information secure and private. It is important that AI-based systems can be deployed in a manner to ensure data protection from cybercrimes, data breaches and other corruption threats.

Transparency & explainability

Transparency & explainability (T&E) requirements could not be fully respected in this case. Ada should commit to transparency and responsible disclosure regarding the AI systems she used. In January 2024, the International Committee of Medical Journal Editors (ICMJE) updated its recommendations to include a new section about Artificial Intelligence (AI)-Assisted Technology:

"At submission, the journal should require authors to disclose whether they used Artificial Intelligence (AI)-assisted technologies (such as Large Language Models [LLMs], chatbots, or image creators) in the production of submitted work. Authors who use such technology should describe, in both the cover letter and the submitted work in the appropriate section if applicable, how they used it." (ICMJE, 2024, p.3)

It is important to identify all people responsible for the different phases of the AI system lifecycle. They should be accountable for the outcomes of the AI systems.

Accountability

When using AI-assisted coding tools, accountability remains an important consideration. We can distinguish legal accountability, ethical accountability, societal accountability and technical accountability (*Singhal et al. 2024, p. 7*). The AI-based system which assisted Ada is not auditable or traceable. We cannot exclude that the outcomes from this AI system is not in conflict with data sharing, research integrity and copyright rules. As a developer for the MathildaCG project, Ada is ultimately responsible for the generated code and its consequences. She main person responsible if the generated code introduces security vulnerabilities or exposes sensitive information. In addition, for future maintenance and transparency, it is important that Ada should clearly document which parts of the code were AI-assisted.

"For the Science journals, the word "original" is enough to signal that text written by ChatGPT is not acceptable: It is, after all, plagiarized from ChatGPT. Further, our authors certify that they themselves are accountable for the research in the paper. Still, to make matters explicit, we are now updating our license and Editorial Policies to specify that text generated by ChatGPT (or any other AI tools) cannot be used in the work, nor can figures, images, or graphics be the products of such tools. And an AI program cannot be an author. A violation

of these policies will constitute scientific misconduct no different from altered images or plagiarism of existing works." (Thorp, 2024, p. 313)

Analysis and reflections

AI-based systems hold tremendous promise in a variety of domains, including life-science research. However, it is important to admit, understand and consider the potential harms they pose. Recent advances in AI and other emerging technologies pose legitimate philosophical questions about ethics and moral values. From the deontological perspective of fixed rules, one should focus on the intention of Ada rather than the results. The rightness or wrongness of Ada's acts need to be judged by their conformity to her duties and to the research rules. Deontological ethics (from the Greek "deon", which means "duty"), often associated with the philosophy of Immanuel Kant, is an ethical theory that is based on rules and universal moral laws to distinguish between right and wrong actions regardless of their consequences. Clearly, sharing patients' personal data or introducing a vulnerability in the code generated by an AI-enhanced coding system is against the rules. In addition, the EU General Data Protection Regulation (GDPR), adopted in 2016 and entered into effect on May 25th, 2018, governs how the personal data of individuals in the EU may be processed, shared, and transferred. Patients who gave consent as part of the recruitment process did not give a clear consent to the processing of their personal data using an AI-based system or transferring them to non-EU countries and international organizations.

Applying Kantian ethics to AI, particularly in the context of AI-enhanced coding, provides a human-centric approach to formulating moral rules and an interesting framework for ethical decision-making. According to the concept of the "categorical imperative" which is a central point of Kantian ethics, it is useful to consider whether the use of AI-enhanced coding could be universalized. We need to evaluate if it would be beneficial if all developers and programmers are assisted by AI. We also need to evaluate whether the AI systems themselves follow universal rules.

"Act only on that maxim through which you can at the same time will that it should become a universal law" (Kant, 1948, p. 421)

Moral rules should be capable of universalisation as illustrated by the above Kant's formulation. In the context of this case study, Ada needs to ask herself if the principle of her action could become a universal rule for everyone to follow.

Kant's moral theory is based on human dignity and humanity as an end in itself. This includes considering the impact on privacy, human labor and autonomy. It is important to ensure that AI systems enhance human capabilities rather than replace human judgment entirely. By using an AI-enhanced coding strategy, the aim of Ada is to maximize efficiency. However, according to Kantian ethical theory, AI systems should be designed and used with the goal of benefiting humanity, not just maximizing efficiency, productivity or profit. Human dignity and humanity as an end in itself is expressed in Kant's maxim:

"Act in such a way that you always treat humanity, whether in your own person or in the person of any other, never simply as a means but always at the same time as an end". (Kant, 1948, p. 429)

It is based on Kant's human-centric approach, and it is the basis of all moral conduct. The question is whether this AI-based coding system is intended to completely replace human functions and rational thinking or to complement and enhance such human characteristics. This question is motivated by the field of "augmented intelligence" also known as "intelligence amplification".

"Augmented intelligence (AI), also known as intelligence augmentation (IA) or cognitive augmentation, is next level in artificial intelligence. The word "augmented" means "to improve." AI software will simply improve products and services, not replace the humans that use them." (Sadiku et al., 2021, p. 772)

Another fundamental principle of the Kantian ethical theory is the “autonomy of the will”. Kant emphasized rational autonomy in moral decision-making. We need to use AI-based systems to enhance human understanding not to replace human reasoning.

Within a large-scale biomedical research project, Ada is dealing with a huge amount of patients’ data. According to the utilitarian approach, which is a form of consequentialism, the main idea to achieve the greatest good for society. Consequentialist reasoning, often associated with the philosophers Jeremy Bentham (1789), John Stuart Mill (1861), and Henry Sidgwick (1907), focuses on the consequences or the utility of an action. According to this ethical approach, an action can be judged right or wrong on the basis of its outcomes.

AI-enhanced coding aims at empowering programmers and introducing more efficiency in the development process. However, the introduction of an AI-based system which can lead to an unethical usage of data still could be punished by the laws. The data breaches behind the ransomware cyberattack on one of the hospitals involved in the MatildaCG project cannot be considered as a good consequence for society. Delayed healthcare and surgery operations could have caused the deaths of patients. AI is an emerging technology which continues to evolve rapidly, it is likely that deontological ethics with fixed rules and consequentialism will have difficulty quickly adapting ethical principles to help people make appropriate ethical decisions.

Does a virtue ethical approach then make sense in this situation? It appears that this theory, which takes its inspiration from Aristotle and deals with the morality and honesty of a person, is a more flexible approach to addressing ethical considerations related to this case with the ultimate goal to develop a framework for enhancing “practical wisdom”. A virtuous person respecting research integrity rules and patients’ rights would not request assistance from an AI-enhanced coding system without considering all risks of harm associated mainly with data security and protection.

On March 13, 2024, the European Parliament adopted the Artificial Intelligence Act (AI Act) which is considered as the world’s first legal framework for AI. This long-awaited legislation was under discussion since April 2021. Based on a risk-based approach which classifies AI applications by their risk of causing harm, it aims to “address risks to health, safety and fundamental rights” resulting from the applications of AI.

"The regulation of new technologies is an unenviable, but essential, task. Governments need to support innovation, but they also have a duty to protect citizens from harm and ensure that people’s rights are not violated." (There are holes in Europe’s AI Act, 2024, p. 216)

References

- Kant, Immanuel. 1948. *Moral Law: Groundwork of the Metaphysics of Morals*: New York: Routledge.
- ICMJE. 2024. *Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals*. <https://www.icmje.org/icmje-recommendations.pdf>
- Maouche, Seraya. 2019. "Google AI: Opportunities, Risks, and Ethical Challenges". *Journal of Contemporary French and Francophone Studies* 23:4, 447-455.
- Sadiku, Matthew N. O.; Ashaolu, Tolulope J.; Ajayi-Majebi, Majebi; Musa, Sarhan M. 2021. "Augmented Intelligence". *International Journal of Scientific Advances (IJSCIA)*, 2/Issue 5: 772-776.
- Singhal, Aditya, Neveditsin, Nikita, Tanveer, Hasnaat, Mago, Vijay. 2024. "Toward Fairness, Accountability, Transparency, and Ethics in AI for Social Media and Health Care: Scoping Review". *JMIR Med Inform.* 12:e50048.
- "There are holes in Europe's AI Act - and researchers can help to fill them". 2024. *Nature*. 625(7994):216.
- Thorp, Holden. 2023. "ChatGPT is fun, but not an author". *Science*. 379(6630):313.
- UNESCO's *Ethics of AI Recommendation*. 2021. <https://unesdoc.unesco.org/ark:/48223/pf0000380455>
- Zelin, Pan; Xie, Zhendong; Liu, Tingting; Xia, Tiansheng. 2024. "Exploring the Key Factors Influencing College Students' Willingness to Use AI Coding Assistant Tools: An Expanded Technology Acceptance Model". *Systems*, 2024, 12, 176.