



ETHICS GUIDE OF GENERATIVE ARTIFICIAL INTELLIGENCE USE IN THE SCIENTIFIC RESEARCH AND PUBLICATION PROCESS OF HIGHER EDUCATION INSTITUTIONS

MAY 2024

Purpose

Within the scope of protecting scientific integrity and trust in science throughout the processes of higher education, it is of utmost importance to determine ethical values governing and to establish ethical rules based on these values in the integration of Generative Artificial Intelligence (GAI), which offers multidimensional and qualified contributions to education and training, scientific research, and publication activities.

Higher education institutions and scientists have a responsibility toward society to protect and develop ethical standards in academic studies and scientific production. Science, technology, and ethics must be balanced in decisions regarding the use of GAI in scientific research and publications. There is a significant difference between using the possibilities of technology to the fullest and being aware of these possibilities, but preferring not to use some of them due to high risks they pose (until the risks are eliminated) in terms of ethics and scientific integrity.

The use of GAI in higher education involves both opportunities and risks. This guide has been prepared to inform higher education institutions to contribute to both understanding and evaluating the risks and benefits as well as, taking precautions against risks in the rapidly developing field of GAI.

1. The Ethical Dimension of GAI Use in Scientific Research and Publications

1.1. Purpose and Scope of GAI Use

If any tasks such as research hypothesis, method, sample size determination/power analysis, data analysis, data collection, data storage and sharing (including federated data sharing/use systems), literature review, source editing, grammar and spelling control, or translation is outsourced to GAI, the assigned tasks must be carefully and diligently reviewed by the researchers to correct any material errors, to identify and eliminate bias or factual/interpretive errors, to ensure contextual consistency, and to take legal and ethical responsibility for the



resulting product (text, image, sound, image, code, decision, conclusion). Scientific research and publication is a process that gains value from the original ideas and original findings of researchers. For this reason, the purpose and scope of the use of GAI should not include stages that require high-level skills, experience, and expertise, such as hypothesis development, discussion, interpretation, and application. Conclusions should be drawn by examining whether the use of GAI will eliminate the researcher's contribution in terms of the extent to which it should be used in scientific studies. It should be recognized that ethical problems are likely to arise if the process of using GAI eliminates the researcher's contribution.

1.2. Ethical Values Underpinned/Threatened by the Use of GAI

The desire of the researcher to do time consuming and labor-intensive tasks more quickly and easily with the help of technology is one of the main motivations that encourage the effective use of GAI in research and publication. However, this motivation and the fact that GAI provides the expected convenience are not sufficient to ethically justify the use of GAI. This is because the scientific research and publication process requires not only basic functions such as observing, measuring, analyzing, synthesizing, interpreting, inferring, classifying, and predicting but also a variety of integrated scientific process skills such as hypothesizing, defining variables with concrete indicators, designing, and communication. In addition, the data collected and used in this process have a unique value due to the nature of the source from which they are obtained (humans, other living beings, society, or nature). Protecting the confidentiality and privacy, and respecting the autonomy of the data subject is thus important. In addition, ethical obligations which arise from scientific research's aim to benefit humanity and the related charge to share the results of such research in a way that contributes to collective scientific knowledge exist. In light of this information, it can be said that the scientific research and publication process is a factual, social, and ethical process that requires care, experience, and high-level cognitive skills. It should not be forgotten that GAI can be used as a tool in the scientific research and publication process, but it cannot be held accountable and responsible for the determination, execution, accuracy, and supervision of the above-mentioned main components of the process.



1.3. Risks to Scientific Accuracy and Integrity of the Use of GAI

Although they were initially designed as open-source systems, the data sources used in the training of GAI systems, the models they are trained on, and the logic of their operation are not completely open to observation or control. Therefore, experts state that the information produced by GAI may be inaccurate or biased. For this reason, the information provided as the outputs of GAI needs to be tested in terms of scientific validity. In addition, the outputs also need to be confirmed in terms of their presence in scientific sources. Moreover, regarding the scientific accuracy of outputs, the second major problem area is scientific bias. GAI can only be as unbiased as the data it has access to. In other words, if the data on which GAI is biased, inevitably, the outputs will also be biased. The data sources of GAI are not an exact reflection of the real world, but only a filtered representation of a certain part of it. It is important to keep in mind that today's GAI is not yet at a sufficient level in terms of scientific accuracy and objectivity, and therefore its use in scientific research and publication should be approached with caution.

2. Core Ethical Values in the Use of GAI

The prioritized and fundamental values in the use of GAI in the scientific research and publication process and the principles produced on their basis will provide the necessary ethical climate for the use of GAI in the scientific field.

Transparency

Transparency is the open declaration of which functions of the GAI are used in which stages of the research and to what extent. The transparency approach, which can also be defined as the principle of openness, is valuable in terms of maintaining trust in science as a requirement of scientific integrity. In addition, transparency can help to take precautions against potential problems such as bias, breaches in data reliability, confidentiality and privacy as well as factual and interpretative errors, which are known as artificial intelligence hallucinations, in the phases in which GAI is involved.

Integrity

Integrity is a value that includes transparency but has a broader content beyond that. As one of the basic building blocks of science, integrity binds not only scientists but also all institutions



involved in the production and use of scientific knowledge. It includes principles such as the use of science for the benefit of humanity, the use of the knowledge produced for purposes that will benefit all humanity, the removal of obstacles to access to scientific knowledge, and the production and publication process of scientific knowledge to be free from bias and conscious errors. In this context, integrity is the fundamental ethical value in planning, conducting, evaluating, and reporting research. The use of GAI has the potential to threaten these principles in many ways.

Diligence

Diligence is an obligation that includes transparency and honesty. Diligence can be defined as all individuals and institutions that produce science being aware of the responsibility they undertake and displaying the seriousness and ethical sensitivity required by this responsibility in all their attitudes in all stages of scientific knowledge production and academic publications. In this context, the principle of due diligence is effective in a wide range of issues, from sloppy research, which is classified as unintentional scientific misrepresentation, to plagiarism, data fraud or manipulation, conflict of interest, divisive publishing, publication of the same study in more than one place, which are among the intentional ethical violations, and even in protecting the rights and benefits of research volunteers in authorship contributions and ranking. The increasing use of GAI increases the vigilance required of scientists and individuals and institutions involved in any stage of scientific production. Within the scope of this responsibility, researchers should aim to benefit the individual, society, and all of humanity. In addition, diligent care should be taken to ensure that research processes and results do not harm humans, the environment, and other living things.

Justice and respect

It is an ethical responsibility to be fair and respectful to other stakeholders in the scientific research and publication process, the individuals/societies from which the research data are obtained, other living beings, and the environment. It is mandatory to show the sources of the information obtained from the outputs of GAI, primarily as a matter of honesty and respect for the labor of other scientists. In scientific research and publication writing stages, the sections where GAI is used should be explained in the method section. Ignoring these issues will lead to disciplinary responsibility within the framework of its conditions.



Privacy and protection of confidentiality

Data privacy protection has both ethical and legal dimensions. It is important to protect the privacy of data providers, to respect their autonomy regarding the use of the data they provide, to carefully carry out the informed consent process, which is legally and ethically mandatory, to take precautions by anticipating possible violations during data collection, storage, and transfer, and to submit all these stages to the approval of the ethics committee when necessary.

Accountability and responsibility

In the process of scientific research and publication, researchers should be aware of the legal and ethical responsibility of using content developed by GAI. The fact that the decision-making/reasoning process of the GAI is unknowable does not eliminate the legal and ethical responsibility of the researcher. The researcher should not be responsible for the reasoning process of the GAI, which is unknowable due to its nature, but should take responsibility for using the results produced by that process and be accountable for these results. Testing the objectivity, reliability, and accuracy of the outputs provided by the GAI is among the legal and ethical responsibilities of the researcher.

Contributing to ethical climate

It is the responsibility of all scientists to identify, publicize, and plan steps to resolve potential ethical issues related to the use of technologies with high-level cognitive abilities such as GAI and similar technologies in scientific research and publication. It is necessary to create an institutional climate to develop and support the necessary ethical sensitivity and to facilitate ethically correct choices. Developing a human-centered approach to the use of GAI in scientific research and publication is of utmost importance, and ensuring that the use of GAI proceeds for the benefit of humanity in a way that does not deepen existing inequalities, by observing the principles of equality and equity is among the priority elements of this ethical climate. Organizing trainings and workshops by relevant institutions, providing online access to informative visuals, developing an institutional culture environment to ensure that ethically correct choices are the easiest choices, and creating ethical guidelines and ethical compliance checklists that are updated in line with the developments in technology are among the important steps that will contribute to this ethical climate to raise awareness on the ethical use of GAI in scientific research and publication and to ensure that the right steps are taken.



3. Some Important Risks and Ethical Issues in the Use of GAI

GAI encompasses natural language processing and large language models; applications that can generate new content in the form of text, images, video, audio, music, or software code; and artificial intelligence applications related to the combination of their integrated forms. The ability to create new content makes GAI stand out. While traditional artificial intelligence draws some conclusions by analyzing a data set, the ability of GAI to generate new content (text, picture, sound, image, code) makes it functional in high-level functions similar to human intelligence. On the other hand, the ability to generate new content does not necessarily mean that GAI understands the real world and social relations, norms, or the contexts of the problem it is dealing with. The validity, accuracy, meaningfulness, and objectivity of the content produced by GAI does not depend on its human-like understanding of the world, but on how well the data provided to it represents the real world, its objectivity, timeliness, and size.

Although GAI offers groundbreaking solutions in many areas, it also brings many risks in its use. These risks arise from the vulnerability of the GAI to threats such as the manipulation of training data, models, system decisions, algorithms, or the entire process. All these risks should be taken into account when using GAIs.

It is necessary for the studies that are conducted, ideas that are developed, and works that are produced to be determined and implemented within the framework of scientific research techniques and ethics. In addition, it is important to establish rules based on ethical values in this regard to protect scientific integrity and trust in science. Otherwise, ethical violations caused by the user will bring disciplinary responsibility according to their type and severity.

The main situations that may cause possible ethical problems (plagiarism, data fabrication, data falsification, violation of confidentiality and privacy, etc.) that may be encountered in the use of GAI in scientific research and publications are as follows:

1. Failure to disclose in the work that GAI was used in content production
2. Unauthorized use of content produced by someone else
3. Inappropriate citation of existing information in the literature without attribution
4. The production of false or misleading data by the GAI and the use of this data by the researcher



5. The inclusion of data and results produced by an irreproducible and unexplained research method into the academic literature
 6. Deepening of discrimination that vulnerable groups may face due to biased and limited data
 7. Collection, storage, transfer, use, and reuse of personal data in violation of legislation
4. **Some Important Recommendations for the Use of GAI Systems**
1. Before starting to use GAI systems, information about how they work and their potential risks should be obtained.
 2. GAI systems, which have great potential in education and research, can significantly improve the quality and effectiveness of education and research when they are used with ethical principles and with scientific responsibility.
 3. It should be recognized that users are held accountable for any ethical violations that may arise from the use of GAI systems.
 4. It should be recognized that GAI systems may not always be trained with high-quality data. Be aware of the potential risks that systems may produce incomplete, inaccurate, or misleading information and may lead to attitudes of bias and discrimination. Products that apply responsible GAI principles should be preferred.
 5. The scope of service, terms of use, or policy documents of the systems used must be read. Entering original ideas, methods, or documents into or sharing them through the GAI systems may lead to unintended consequences.
 6. Personal data should not be entered into these systems unless they are anonymized or masked outside the GAI systems. Before starting to use the GAI systems, the user should have detailed information about how the systems work and their potential risks.
 7. Users who want to improve themselves on the subject can benefit from resources such as the "Artificial Intelligence Risk Management Framework" documents, which have been developed to minimize the potential negativity and maximize the positive effects that may occur in GAI systems and cover the steps to ensure the reliability of the systems and manage the risks that occur in the system.



Frequently asked questions and answers to these questions regarding some of the issues encountered in the use of GAI are given below.

1. Can GAI be used in scientific studies?

Since the use of GAI in scientific studies involves ethical issues such as authorship, copyright, and plagiarism, special attention should be paid to citation principles. In this context, since it is not conceivable to dictate an entire article to GAI, the administrative and legal sanctions that may arise from non-compliance with the citation principles, even anonymous information, should not be ignored.

The use of text, visuals, audio, etc. produced by GAI may pose ethical risks, even if it is clearly stated in scientific studies. Since GAI cannot take responsibility for the final version of a study as a researcher, it cannot take part as an author in scientific studies. These studies should consist of the original opinions and findings of the author(s).

GAI may be used in a literature review. However, it should be kept in mind that the GAI may present fabricated sources as real, may be far from evaluating the scientific quality of the sources, and may exclude current sources from the search. Additionally, problems such as bias are common.

2. Is it ethically appropriate to use GAI in data analysis?

It is ethically appropriate to use the self-trained model in data analysis. However, if a ready-made GAI system is used, it should be taken into account that the methods used in data analysis may create ethical problems.

3. Can GAI be used for translation or language control of scientific studies?

Yes. However, the final version of the translated content or the language checked with GAI must be also be controlled and approved by the user. The final responsibility for the resulting text lies with the author(s).

4. Is it right to use GAI in surveys?

It is not correct to use GAI instead of real participants.



5. Can GAI be used for data labeling?

GAI can be used for text and image labeling. However, it should be noted that the accuracy of labeling depends on the objectivity and quality of the dataset used to train the GAI. As in all other stages of using the GAI, scientific accuracy and accountability are the responsibility of the user.

6. Can the GAI be used for data quality assessment, error/inconsistency detection, and bias detection?

Yes, but it may be more appropriate to position the use of GAI as part of a comprehensive data evaluation strategy, rather than as the sole tool for this purpose.

7. What should be considered in terms of data ethics in the use of GAI in scientific research and publication?

1. Data source and permission to use the data
2. Data privacy and confidentiality
3. Data bias
4. GAI hallucinations
5. Data timeliness
6. Data reliability
7. Intellectual property rights

8. Is it necessary to provide information on the use of GAI in the Ethics Committee application?

The researchers should inform the Ethics Committee about the purpose, scope, and nature of GAI use in the proposal.

9. What should be considered when reporting the content obtained from GAI systems?

1. The version of the GAI, and the phase of the study at which it is used should be specified.
2. The contents of GAI should be reviewed by academic rigor and ethical rules, and the content should be checked and reported.
3. The accuracy of the content is the responsibility of the user.



4. The contents should be critically reviewed, and potential biases arising from the GAI should be assessed and confirmed by the researchers. The responsibility for the report content, results, and arguments lies with the researchers.

10. Is there legislation on GAI in Türkiye?

There is no legislation directly related to artificial intelligence in Türkiye yet. However, the following legislations are recommended to be taken into consideration by researchers in the use of GAI:

- Law No. 6698 on the Protection of Personal Data
- Law No. 2547 on Higher Education
- Law No. 5846 on Intellectual and Artistic Works
- Graduate Education and Examination Regulations
- Higher Education Institutions Scientific Research and Publication Ethics Directive