Policy on the use of GenAI and the PhD-trajectory:

Guideline on legitimate and responsible use of GenAI in PhD research at the EUR

Introduction

The use of artificial intelligence and machine learning (AI/ML) in general and the application of Generative AI (GenAI) has increased tremendously since the launch of ChatGPT in November 2022. The availability of data, algorithms and computer power have contributed to an increase in the application of AI/ML as well as research on the impact of AI/ML. Within seconds these programs can produce texts and images that are hard to distinguish from text created by people.[1] As AI/ML or GenAI are becoming an integrated part of working life the academic world is presented with many new opportunities as well as risks. Every user faces challenges for using GenAI legitimately and responsibly, through this document extra attention is given to PhD candidates and their supervisor as the PhD trajectory focuses on developing the competences needed as an independent researcher. The responsibility for quality and integrity is hierarchically shared between the candidate and supervisor and the department they are embedded in.

National and European codes and policies on, for example, the code of academic integrity and the EU AI Act, are leading and respected in this policy. The EU AI Act for example requires research to be transparent by 'owning up' and discussion the limitations, this policy follows those lines.[2] If national codes or tools are developed for the use of AI, they overrule this policy. The existing codes on scientific integrity provides much guidance on, for example, integrity that are relevant for use of GenAI, however, due to the unique capabilities GenAI offers, these codes do not cover all aspects of responsible use.[3] This policy attempts to fill these gaps. Furthermore, there are policy developments to watch as they can provide insights that will help to sharpen these guidelines. The 'Russell Group Principles on the use of Generative AI Tools in Education' and the NWO's Preliminary Position on Generative AI in the Application and Review Process were, for example, used as input for the policy. This document will be updated in the light of these principles.[4]

The graduate schools, therefore, raised the question on how to deal with GenAI, seeking a policy for PhD trajectories, which is expected to be followed up by faculty specific and/or discipline specific policies and guidelines. The target group of this policy are not the PhD candidates, but the support staff involved in the policy process. The policy should, therefore, be read as a living document that aims to be integrated into an EUR-wide policy. The opportunities and risks that AI currently pose will like change and develop as the software develops. The policy focuses exclusively on the PhD-trajectory after the candidate has started. It does not discuss the assessment of the research proposals. Furthermore, AI-detection tools are not discussed as they are in 2024 not reliable enough to make general guidelines about for examinations of proposals or dissertations. It is, therefore, not recommended to use these tools.

This policy distinguishes three levels within the EUR which need to be reflected on the use of GenAI during the PhD-trajectory:

---

[1] See: https://www.universiteitenvannederland.nl/en/research-integrity

[2] EU AI Act

[3] Nederlandse gedragscode wetenschappelijke integriteit, KNAW et al., 2018.

[4] Russell Group principles on the use of generative AI tools in education, *Russell Group*. And
NWO's preliminary position on generative AI in the Application and Review Process, NWO,
https://www.nwo.nl/en/nwos-preliminary-position-on-generative-ai-in-the-application-and-review-process.

- Firstly, the central level; the focus of this policy. The aim of this document is to provide general principles that give guidance on how to use GenAI legitimately and responsibly. With an emphasis on ensuring that the quality of the PhD dissertation (the test of competence to be an independent scholar) is enhanced with the new opportunities that GenAI offers and not compromised by using GenAI. Furthermore, it focuses on aspects of the trajectory to ensure security (of for example data) and integrity throughout the process.
- Secondly, the faculties all have their specific research fields that come with unique opportunities as well as issues or needs in relation to the use of GenAI during the PhD trajectory. There will be exceptions and finetuning needed to the general guidelines that a faculty can address in their own specific policies and procedures if they promote legitimate and responsible use of GenAI.
- Thirdly, there will be issues that are too specific and/or difficult to fit into this general policy. In that case, this policy refers to the Privacy and Ethics Board at the EUR that through their tailor-made advice and decision will create internal jurisprudence on the use of GenAI.

The policy is developed in cooperation with the AI@EUR team and experts regarding different related risks. Currently, an EUR-wide policy is also being developed. The goal is to incorporate this policy within that EUR wide framework. This policy will start with an overview of general principles that need to be adhered to. Secondly, the document will provide a discussion of the definition of GenAI, the risks related to the use, and the national developments at the Dutch government and other universities. Thirdly, it will discuss the faculties room for specification and the role of the privacy and ethics board.

General principles of legitimate and responsible use of AI for PhD research

In a PhD-trajectory responsibilities are shared and/or divided between the candidate, supervisor, and graduate school, with a more indirect role for the Dean and Beadles' office. Research on the roles and responsibilities in stem cell research demonstrates that late career researchers (in this case supervisors) often consider themselves accountable for the research and have a broader conception of ethics, at the same time they have limited time and sometimes lack technological knowledge that the PhD-candidates have.[5] These contradictions in knowledge and time show that every stakeholder in the process has their own role and obligation in ensuring legitimate and responsible use of GenAI. This also includes other stakeholders such as the graduate schools and the Beadles Office.

The guidelines below start with general principles followed by an overview per stakeholder with end responsibility (while keeping in might that the realization is often a team effort). The aim is that the guidelines provide an overview for legitimate and responsible use, by setting boundaries as well as providing a design that promotes appropriate use as GenAI offers many possibilities for PhD-candidates. Even though it is hard to operationalize these principles due to the newness and ever-changing field of GenAI, it is important to have principles and adhere to them as they help us to navigate through this new landscape.

---

[5] Assen LS, Jongsma KR, Isasi R, Tryfonidou MA, Bredenoord AL. Roles and responsibilities in stem cell research: a focus group study with stem cell researchers and patients. Regen. Med. 17(7), 445–459 (2022).

General:

- If the other principles are adhered to it is allowed to use (Gen)AI systems in the PhD-trajectory for research purposes or as a support tool.
- Existing rules, regulations, and codes on, for example, academic integrity or privacy need to be respected when using GenAI systems.
- There should be clarity about the use of GenAI; work performed by GenAI should not be claimed as their own work (see EU AI Act)
- Any forms of plagiarism in the output of the GenAI system should be detected by the user of the GenAI system (in this case the PhD candidate), they are responsible for the legitimacy of the information (this is regarding text as well as other forms such as video's, images, or audio).
- It is highly recommended that either the licensed GenAI system or the own EUR language model is used. If another program is used it is advised to consult with the privacy officer.

Graduate School:

- Responsible and legitimate use starts with knowledge of both the possibilities and risks of GenAI. The graduate schools are in the lead of educating candidates and supervisors about the societal, ethical, legal, and environmental implications of GenAI systems.
- Together with Academic Affairs and the Library, the graduate schools provide information about rules, regulations, and guidelines. For example, information that the use by the EUR developed or licensed GenAI are preferred should be shared by the Graduate schools.
- Every course should establish whether the use of GenAI is allowed and in what way. These rules about (Gen)AI use should be clearly communicated to the PhD-candidates.
- Misuse of data breach should be reported to the graduate schools and privacy officers.

Supervisor:

- A PhD-trajectory is a form of guided research, the supervisor's role is to ensure that the candidate adheres to the code academic integrity also in relation to the use of (Gen)AI. To do this the supervisor should be knowledgeable about the opportunities and risks related to GenAI.
- As it is known that GenAI sometimes plagiarizes other sources in their output or come up with incorrect results, it is the supervisor's responsibility to make the candidate aware of need to have an end product that is free of plagiarism and other incorrect information created by GenAI also in the cases of picture, tables, etc.
  - This means that there is a need for schooling supervisors on the risks and possibilities of GenAI.

PhD candidate:

- Transparency is an essential aspect of responsible use of GenAI during the PhD-trajectory. There is still much uncertainty regarding GenAI and how it works or what it means for society, but it is

becoming clear that it opens a lot of opportunities for researchers. Being transparent about how GenAI is used will help the academic community to explore all ways in which GenAI can be an asset, as well as provide clearly or at least context if in the future errors in text, picture, or data are found. Transparency can be provided in different ways:

- o Discuss the use of GenAI in the Data Management Plan
    - This will also help ensure the GDPR laws are always adhered to.
- o Reference the use of GenAI (text, images, etc.) in research, just as other information sources are referenced. Use either the reference rules of the used reference style or faculty specific rules on references GenAI.
- o Add a note to a research product about the contribution of GenAI. This is often already done with a multi-author article and a similar addition could be added about how GenAI was used during the research (this makes it possible to adhere to EU AI Act rules of owning up and critical reflection on the use).
- o In cooperation with other researchers (internationally) it is advised to make clear agreements about the responsible use of GenAI and one's role in realizing this.
- Reflection on the use of GenAI during the process is important for the personal development of the candidate to become an independent researcher.
    - o It is, therefore, advised that candidates add a methodologic reflection on the role of GenAI in their dissertation.
- Security of data and participants is considered meaning that personal data or research sensitive information is not shared with GenAI system unless the privacy and ethics board is consulted. Both to protect the information about the people involved and the work of the candidate.

Background; definition and risks

GenAI was a trending topic last year with messages varying from dystopic warnings to cheerful laudations. Even though GenAI has been available since 2018, the free access through an online interface in the form of ChatGPT made the software recently available to anyone with access to the internet (and a basic understanding of mainly English), thus making the technology widely known and used across the globe. To define the technology, this policy adopts the definition of the Organisation for Economic Co-operation and Development's (OECD) as this is the definition to be used in the AI Act as well. The definition:

"An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment." [6]

---

[6] Recommendation of the Council on Artificial Intelligence, OECD, 08-11-2023, https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449.

This means that GenAI systems receive input in the case of ChatGPT from the user which is broken down into smaller units (called tokens).[7] One the basis of these tokens, statistical patterns are developed to predict what content (found within the information that is fed to the system) are likely to form a coherent response and these outcomes are converted into a readable text, a picture, or a video. That output is filtered through guardrails which aim to remove any (by the system marked as) unwanted content. Furthermore, when looking beyond the model, guardrails can also be found in the form of human feedback on the system.[8]

Many opportunities arise with these new technologies, but there are also some vulnerabilities and/or ethical questions, especially when using such a system in academia. This policy will continue discussing these issues in relation to the PhD-trajectory. This means larger fundamental ethical or political questions or issues will not be discussed as they go beyond this document's scope. The following risks are identified:

Impact	Risk
Vulnerabilities	Bias: Biases or specific cultural or commercial values cannot be traced back from the information that is provided by an GenAI system. The dominance of the English language in the creation and training of the systems is for example a predictor of bias in the system.
	Plagiarism: GenAI systems are known to paraphrase or quote existing work without citation this goes against the code of academic integrity.
	Lack of transparency: It is difficult (if not impossible) to know where content is coming from due to the technical design of the programme.
	Environmental costs: Large GenAI models are energy intensive in their inference process. Over usage may lead to damaging environmental consequences.
Undesirable situations	Outsourcing the métier: The PhD end product submitted for publication and judgement should be the researchers own work as the PhD trajectory is in essence a test of competence. GenAI as tool can be helpful in drafting, creating, or methodological purposes but should never be used as a final product in itself.

---

[7] This could include text as well as images, sounds and other types of inputs.
[8] UNESCO, Guidance for Generative AI in Education and Research

	Accuracy: GenAI systems can fabricate information creating misinformation or disinformation. As a result, a growing body of false information will circulate online. The methodological reflection on how one uses data will, therefore, become even more important.
	Ownership of product: If the GenAI system is used to create texts and original combinations of information that has meaning, the PhD candidate is not the sole creator of the work, which makes it difficult to protect the IPR.
	Data security: In the case of unlicensed programs, sensitive information fed to the program can be spread by GenAI as the systems use information to answer the input of other users.
Criterial/High Risk	Scientific integrity: Users of GenAI systems should always follow the principles of Scientific integrity: Five principles form the basis for integrity in research[9]: honesty, scrupulousness, transparency, independence, responsibility.
	Input become training material: All input fed to GenAI systems is kept and re-used by the program to train itself and can be used to produce output in reaction to input from other users. When a PhD candidate feeds the GenAI data, they no longer have control over how this data is used by the system, which could lead, for instance, to a scoop by another researcher who is able to reconstruct methods, questions, and outcomes, and use this for their own work. This could have serious implications for the use of such a data set and the uniqueness of the dissertation.
	Intellectual property: Great ideas about prototypes or ideas that lead to valorization, should not be fed to the GenAI system, because one loses control over the confidentiality of the idea and the product, code, or service that can be created based on this, which practically results in giving away your intellectual property rights.

---

[9] See: https://www.universiteitenvannederland.nl/en/research-integrity

	GDPR: If the GenAI systems are fed personal data, the control over the data and the protection of the data is lost, which conflicts with GDPR. It is the PhD candidate's responsibility to ensure that the GDPR is adhered to.
Yet to determine	Unfair reuse: There are several court cases currently taking place on whether ChatGPT violates copyright laws by infringing upon other creators' work. Once there is a verdict this policy will be updated with copy rights principles/rules.

The principles in this policy aim to tackle these themes while raising awareness among people involved in the PhD-trajectory.

Exceptions and Faculty specific issues or unique challenges within PhD research

The above discussed principles provide a useful overview of typical use of GenAI. However, the variety of research fields and methodologies might require more specific and tailor-made guidelines. Furthermore, the above discussed guidelines need to be translated into practical information per graduate school. These initiatives are already taking place regarding education; RSM, for example, create guidelines for evaluating usage of GenAI in the master thesis trajectory. Such tools targeted at the PhD trajectory will help candidates and their supervisors to use GenAI in a responsible way. Therefore, the schools and/or graduate schools are encouraged to translate these guidelines into faculty specific tools with attention to the unique situations they encounter.

When working with the guidelines, there will be cases which do not fit this policy and/or are too specific to make general guidelines about. For these cases, the privacy and ethics board at the EUR should be consulted for advice. This board can evaluate specific cases and has the expertise and network to reflect on all the above discussed risks. Furthermore, their advice will create internal jurisprudence on the use of GenAI.

In the case of a breach of the general principles, the graduate school should be contacted. The graduate school can determine the nature and severity of the situation and contact the involved parties. In the case of deliberate wrongful use, it should be treated as a suspected case of fraud. If it is less clear either the privacy and ethics board or the coordinator scientific integrity should be contacted depending on the case.

Implementation plan

When the general principles are adopted, they will be added to the PhD policy. To implement the policy, Academic Affairs will be in the lead and work with different stakeholders to ensure that the guidelines

become part of the PhD trajectory. All stakeholders will be informed, however, with some groups we will discuss the implementation further:

- The graduate schools: they will play a crucial role in implementing the policy and guiding candidates and supervisors towards responsible use. They will be asked to incorporate the principles and to finetune them to their specific context and research methods. We will discuss how the principles can be integrated into the training program for the PhD-candidates. Furthermore, we will investigate how to reach supervisors through, for example, information sessions.
- Together with the University Library (and/or ERS) academic affairs will investigate way to translate the guidelines into tools/information documents for PhD-candidates with the "The Five Pitfalls of Using ChatGPT for You Research Assignment" (created by the University Library) as a great example of how to operationalize principles for a specific user group, in this case students.
- The principles will be shared with AI@EUR and other groups that work on AI policy/projects, for example in the field of education. They will be asked to incorporate the principles into their work.
- Together with CLI it will be explored if and how we can spread information about the use of GenAI.

Living document and the integration into EUR wide policy

The developments regarding GenAI are taking place at a fast pace and the PhD trajectory is not the only process that is affected by such models. AI@EUR is, therefore, working on an EUR wide policy in which these principles can be incorporated. Until the adoption of the EUR wide policy on GenAI, this policy framework can be incorporated into the PhD policy. Furthermore, policy developments on GenAI are also taking place in education as they are closely linked with policy and will be updated to follow the same line of policy and/or reasoning.

Both the technological possibilities and the legislative reactions at the national and European level are rapidly developing. To keep up with them, this policy is a living document that will adapt to the new technological possibilities and adopted rules and regulations from the government and the EU.

Literature:

- EU AI Act, 2024
- Nederlandse gedragscode wetenschappelijke integriteit, KNAW et al., 2018.
- Russell Group principles on the use of generative AI tools in education, *Russell Group.*
- Assen LS, Jongsma KR, Isasi R, Tryfonidou MA, Bredenoord AL. Roles and responsibilities in stem cell research: a focus group study with stem cell researchers and patients. Regen. Med. 17(7), 445–459 (2022).
- Recommendation of the Council on Artificial Intelligence, OECD, 08-11-2023, https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449.
- UNESCO, Guidance for Generative AI in Education and Research.