Research Integrity in the Era of Generative AI – A Perspective

Gideon Ogunniye,
g.ogunniye@abdn.ac.uk
School of Natural and Computing Sciences,
University of Aberdeen,
Aberdeen, AB24 3FX, UK

Amaya Hana,
amaya.hana@ucl.ac.uk
Department of Science, Technology, Engineering
and Public Policy (UCL STEaPP),
University College London, London, WC1E 6JA, UK

Abstract

In the past decade, advances in artificial intelligence (AI) have spurred the creation of AI-based research tools for planning, analysing and reporting results and disseminating knowledge. By harnessing the capabilities of AI models, authors can benefit from automated content generation and analyses for research articles, leading to increased efficiency and effectiveness. However, there is growing debate about how AI might challenge long-established codes and practices of research integrity. There are two camps in academia. The first is early AI adopters who think the technology can enhance the research process – improving data quality, communicating findings more effectively, and publishing results with minimal delays. The second group consists of stakeholders (e.g., academics and users of research) who believe that the accuracy and quality of the outputs produced by AI are not trustworthy (due to factors including bias or limited coverage) and, therefore, may corrupt the integrity of research. Although implementing guidelines or standards for authors on the use of AI tools is currently considered a preferred solution by many, this article argues that on its own, guidelines or standards may not provide sufficient capability to address concerns about the use of AI tools in research. An argument is presented that educational toolkits on how to use AI tools in research and evaluation tools to critically evaluate their outputs, alongside standards and guidelines, are essential for the effective use of AI tools in research and education.

1. Introduction

Advances in Artificial Intelligence (AI) in the past decade have led to the development of Generative AI (GenAI) for use in research (Ganjavi et al., 2024), where AI tools create new information from almost nothing after learning from trained models (Barreto et al., 2023). GenAI can create original work, such as an article, a code, a painting, a poem, or a video. It typically uses Large Language Models (LLM) to generate outputs based on user prompts. LLMs are a type of AI that uses deep learning techniques to analyse and generate natural language. They have become increasingly popular in recent years due to their ability to understand human language, generate human-like text, and complete various tasks, from text classification to language translation. LLMs are essential for natural language processing tasks in various fields, including business, research, and academia. With the ability to analyse human language and generate human-like text, LLMs such as GPT-3[1], BERT[2], LaMDA[3] and XLNet[4] have set new standards for understanding and generating natural language. This is a dynamic and rapidly evolving field, with new tools of increasing sophistication and capability being released almost monthly. However, while LLMs constitute a specific category of GenAI models with a specialised focus on text-based data, the growing impact of non-LLM GenAI models, such as the AlphaFold (Jumper et. al., 2021), which is a neural network-based model used to predict protein structures, should not be disregarded. Generative Adversarial Networks (GANs)[5] is also a type of GenAI that uses two neural networks, a generator and a discriminator, that compete against each other to create new data. GANs can generate realistic images, videos, and voice outputs. Some applications of GANs include image colourisation, increasing image resolution, and turning 2D images into 3D.

AI conveys the broader concept of machines carrying out tasks in ways humans might consider intelligent. It manifests ‘intelligence’ (or reasoning) demonstrated by machines, in contrast to the natural intelligence displayed by humans. Machine Learning (ML) is a subset of AI, focusing on algorithms that can learn from data and make predictions or decisions based on the data. Deep Learning (DL) is a subset of ML that utilises neural networks with many layers (deep neural networks) to learn and make decisions based on data. GenAI represents the next frontier–machines with versatile cognitive abilities to understand, learn and apply knowledge across different domains. In summary, AI sets the stage, ML refines it with data-driven insights, DL adds capability and complexity with deep neural networks, and GenAI emerges as a promising future approach, capable of broader cognitive understanding.

GenAI has gained popularity since the release of Generative Pretrained Transformers (GPT)–namely ChatGPT, launched by the AI research organisation OpenAI in November 2022[6]. Incredibly, within two months, it was estimated that Chat-GPT hit 100 million monthly active users, making it the fastest-growing application in history at the time[7]. ChatGPT is becoming increasingly important in research and scientific writing (Golan et al., 2023; Huang & Tan, 2023).

ChatGPT is a powerful tool that can assist researchers propose novel research ideas (Graf & Bernardi, 2023), write code and novel textual content (Lecler et al., 2023; Macdonald et al., 2023), manage data by analysing and synthesising large volumes of information effectively and more quickly than humans (Bhatia & Kulkarni, 2023), and improve the quality of writing by identifying potential errors, inconsistencies or gaps in an argument (Huang & Tan, 2023). Moreover, it is tremendously helpful to non-native English speakers to generate grammar and sentence structure (Graf & Bernardi, 2023), suggest appropriate vocabulary choices, translate text from one language to another (Lecler et al., 2023), and so on.

Generally, research and AI are becoming increasingly intertwined. As AI advances, researchers will likely continue to embrace its potential or voice concerns about its risks. The rise in interest in the impact of AI as a tool in research has mainly resulted from the abundance of news about dishonest practices in the research community[8]. Some of the ‘acts of dishonesty ’or fraudulent research practices that have been debated within the research community comprise cases such as the creation of false data or manipulating data to generate preferred results, cheating or using others' ideas as own, underserved authorship claims, using AI tools to carry out paper and proposal reviews, to duplicate

[1] https://openai.com/product

[2] https://pub.towardsai.net/understanding-bert-b69ce7ad03c1

[3] https://research.google/pubs/lamda-language-models-for-dialog-applications/

[4] https://towardsdatascience.com/what-is-xlnet-and-why-it-outperforms-bert-8d8fce710335

[5] https://aws.amazon.com/what-is/gan/

[6] https://openai.com/blog/chatgpt

[7] https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

[8]https://www.dailymail.co.uk/sciencetech/article-13211523/ChatGPT-scandal-AI-generated-scientific-papers.

publications, and create ‘predatory publications ’(Dhakal, 2016). These practices pose a serious challenge to the integrity of research.

It is important to define the concept of ‘integrity ’in research. According to Macfarlane (2010), research integrity is the consolidation of a researcher’s personal values, norms, and identity with their professional practice as a researcher. However, research integrity goes beyond the values or practices of an individual researcher. It touches on the values and practices of the research institution and the community they are part of. Undoubtedly, research must be of high quality to build knowledge that guides and informs the academic community of practice and applies outside of the academic community – contributing to social and global development. There is a growing understanding that the credibility of research can be debilitated if there is room for doubt concerning common ethical norms such as reliability, reproducibility, rigour, and accountability (Daniel, 2018; Mebane et al., 2019). Therefore, a research institution should implement and apply appropriate guidelines to ensure research integrity, while it is incumbent on a researcher to adhere to them.

Furthermore, there is a pressing need to understand the impacts of GenAI on research integrity and the perceptions of different stakeholders on these impacts. Will it be considered cheating if a researcher gives a chunk of text to a GenAI-based tool and asks it to correct grammar and spelling, or if the tool is asked to rewrite that chunk of text in line with its feedback, or if it is asked to provide feedback in terms of clarity or strength of arguments, or to synthesise information from a large volume of data? Given these scenarios, a legitimate question will be, at what point does the use of GenAI in research constitute cheating? To address these questions, it is important to understand the long-established codes and practices of research integrity in different fields of study. For example, Arts and Humanities may have different requirements to Science and Engineering.

It should be noted that guidelines for using GenAI in research are lacking. Among the institutions that have provided guidelines, the allowable uses of GenAI and how they should be disclosed vary substantially (Ganjavi et al. 2024). In their survey, Ganjavi et al. (2024) revealed that among the top 100 largest academic publishers, only 24% provided guidance on the use of GenAI. There is also a lack of standards, making it difficult to evaluate GenAI tools for research. Standardised guidelines are required to ultimately define the responsible use of GenAI within the general definition of research integrity. However, there is a consensus that AI cannot be cited as an author. In February 2023, the Committee on Publication Ethics (COPE), an organisation comprised of editors, publishers, universities, and research institutes that helps inform publication ethics across all academic disciplines[9] released a position statement on AI tools in research publications, highlighting that “AI tools cannot meet the authorship requirements, as they cannot take responsibility for the submitted work”, while also suggesting ways to disclose AI use and emphasising that authors are responsible for the work produced by AI tools.

Although implementing guidelines or standards for the use of GenAI in research is an important and preferred solution, this article argues that, on their own, guidelines or standards may not provide sufficient capability. To achieve the fast and agile approach required to deal with GenAI in research, this article proposes leveraging educational toolkits on how to use them, and evaluation tools to evaluate their outputs alongside standards for responsible use.

2. The Most Typical Uses of AI Tools in Research

AI has the potential to enhance the efficiency of the research process. Despite some obvious ethical concerns, AI is rapidly reshaping the research landscape. Researchers in

[9] https://publicationethics.org/about/ourorganisation.

universities, industry, and government institutions increasingly rely on the power of using AI tools to save time, write more effectively, disseminate their work more widely, and measure the impact of their work more accurately.

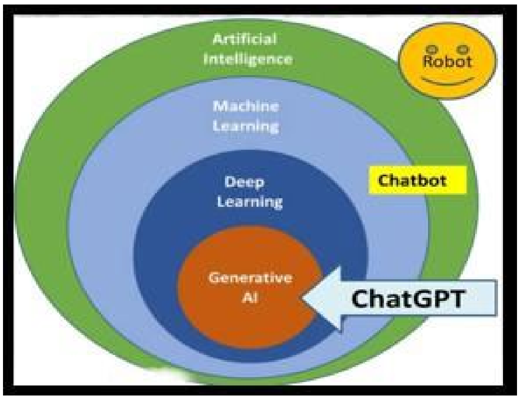

Figure 1: Relationships between AI, ML, DL and Gen AI

AI has already left a mark on every stage of the research process: from hypothesis generation (Amato & Coronato, 2018), to planning experiments, collecting and analysing data (Crawford et al., 2023), writing and editing scholarly manuscripts, citing relevant sources and identifying potential venues best suited for disseminating studies (Hosseini et al., 2023). However, it is important to ensure that the vital human intelligence element of research is not overlooked and that a balance between GenAI output and human intelligence is maintained. This is especially important because GenAI has areas for improvement, including outdated training data (Algahtani, 2024), training data with many inaccuracies and biases, misinterpreting prompts that are adversarial attacks and producing hallucinations (Templin et al., 2024). GenAI can also be exploited to create fake news, misinformation and 'deep fakes'.

2.1. AI Tools for Literature Search and Review

In this digital era, the availability, through the Internet, of large amounts of information at the researchers ’fingertips makes finding relevant materials for literature reviews challenging. In such a scenario, automating the task of identifying reliable articles can simplify literature reviews. Numerous GenAI tools such as ChatGPT, Elicit[10], Scholarcy[11] and Consensus[12], among others, can help researchers find and collate information sources, identify knowledge gaps that need to be bridged, thus helping generate potential project ideas, or summarise, paraphrase or write scholarly texts (Khalifa & Albadawy, 2024).

These tools help researchers identify relevant literature quickly and efficiently, even in large and complex datasets. Some can analyse research papers at a ‘superhuman ’ speed. They achieve this by automating time-consuming tasks like summarising research papers and extracting figures, tables, and other data. Some tools also visualise networks of related papers and authors, allowing researchers to share their collections with colleagues and improving collaboration. The pros and cons of such tools when incorporated into a research workflow are listed below.

---

[10] https://elicit.org/
[11] https://www.scholarcy.com/
[12] https://consensus.app/

Table 1: Pros and Cons of AI Tools for Literature Review

Pros	Cons
Time and resource efficiency	Concerns about the originality of AI-generated content
Higher writing standards due to assistance in organising data	Ethical concerns, especially regarding intellectual property rights
Higher-quality research insights	Lack of nuanced understanding by AI algorithms
Personalised assistance	Risk of hallucination

2.2. AI Tools for Planning and Study Design

"What are the most suitable research methods to answer my questions? What is the best approach to validate my results? And what issues should I care about when designing this or that specific type of inquiry?" These are some examples of prompts that researchers could use to gain insight while designing their research procedures.

Another common activity, a frequent task for empirical researchers, is designing research instruments such as survey questionnaires or interview and observation scripts. GenAI can also be very helpful in this regard. For example, if a researcher provides context about research objectives, ChatGPT-like tools could suggest some survey questions. AI-powered experimental design tools use machine learning algorithms to optimise parameters. They can save a lot of time and effort by automating experiment design. These tools can reduce human errors and R&D costs. It is noteworthy that they vary for specific disciplines or types of studies.

However, according to a paper published in Nature, Messeri and Crockett (2024), the authors argue that some future AI approaches could restrict or constrain the questions researchers ask, the experiments they perform, and the perspectives that come into play concerning scientific data and theories. These factors could leave people vulnerable to 'illusions of understanding 'in which they believe they understand the world better than they do. According to them, "there is a risk that scientists will use AI to produce more while understanding less". While they support the use of AI tools by researchers, they also advocate for a conversation about how they should be used and suggest that the assumption that all uses of the technology, or the ubiquitous use of it, will benefit research is wrong.

2.3. AI for Data Collection, Visualisation, Interpretation and Reporting

Collecting usable data can be a challenge – data may be noisy, or collection may be too costly. As such, designing data collection workflows to capture high-quality data is important. This is particularly relevant for industry researchers to understand market dynamics, stay ahead of the competition, and provide value to their stakeholders. Machine learning researchers must also continuously collect high-quality data to update and train AI models. AI tools[13] offer solutions to collect, manage, store, and access data in such cases. This saves researchers 'efforts and time, increasing the output by automating the tedious multiple steps of identifying, profiling, sourcing, and preparing relevant data.

After collecting usable data, AI-assisted optimisation tools[14] have been used to present data as an image or graph, making it easier to identify patterns, interpret and analyse

[13] https://www.appen.com/ai-data

[14] https://www.heavy.ai/learn/data-visualization, https://research-ai.io/

data, obtain insights and monitor trends without biases.
However, researchers evaluating the risks of AI in research and society have recognised a variety of ethical concerns in the use of AI for data collection, including algorithmic bias (Benjamin, 2019; Rahman, 2020), errors and hallucinations **(**Kidd & Birhane, 2023), failures of reproducibility (Kapoor & Narayanan, 2023), and lack of interpretability (Rudin, 2019). Scholars recognise that more than technical approaches are needed to address these ethical concerns.

2.4. AI for Manuscript Preparation

AI-based tools have been used in manuscript preparation, including those that can help write and proofread articles, track references, cite sources, and detect plagiarism. Several AI writing assistance tools[15] edit text in real time, proofread and fix spelling, punctuation, and grammar and can suggest alternate words to diversify the vocabulary. AI-based note-taking systems can track source information and avoid plagiarism.

Ethical principles, including openness, honesty, transparency, efficient use of resources, and fair allocation of credits, are essential to drive the disclosure of AI in manuscript preparation (Hosseini et al., 2023; Khalifa & Albadawy, 2024). According to Hosseini et al. (2023), these principles are paramount to fostering integrity, reproducibility, and rigour in research. They argue that banning the use of AI in research is not a reasonable response to the moral conundrums created by its use and that bans are unenforceable and would encourage undisclosed use of AI. Further, they argued that naming AI as authors or mentioning them in the acknowledgements are inappropriate forms of recognition because AI do not have free will and, therefore, cannot be held morally or legally responsible for what they do. They recommended that i) the use of AI should be disclosed in the introduction or methods section to describe details such as the query prompts used transparently and notes on which parts of the text are affected, ii) the use of in-text citations and references (to cite their applications and to improve findability and indexing), and (3) the recording and submission of researchers 'interactions with AI as supplementary material or appendices.

2.5. AI Tools for Peer Review Assistance

The volume of peer review submissions is constantly growing. Reduced screening and reviewing time can save millions of working hours and boost academic productivity. AI-powered peer review tools can create the potential for semi-automated peer review systems where low-quality or controversial studies could be flagged, and reviewers could be matched with manuscripts according to their subject-matter expertise.

However, it is not recommended that AI tools be used to automate peer review. Rather, these tools can be effectively used in the peer review process to suggest appropriate journals and venues for an article, perform initial quality control for submitted manuscripts, and find reviewers.

[15] https://paperpal.com/paperpal-for-researchers, https://app.grammarly.com/, https://www.writefull.com/

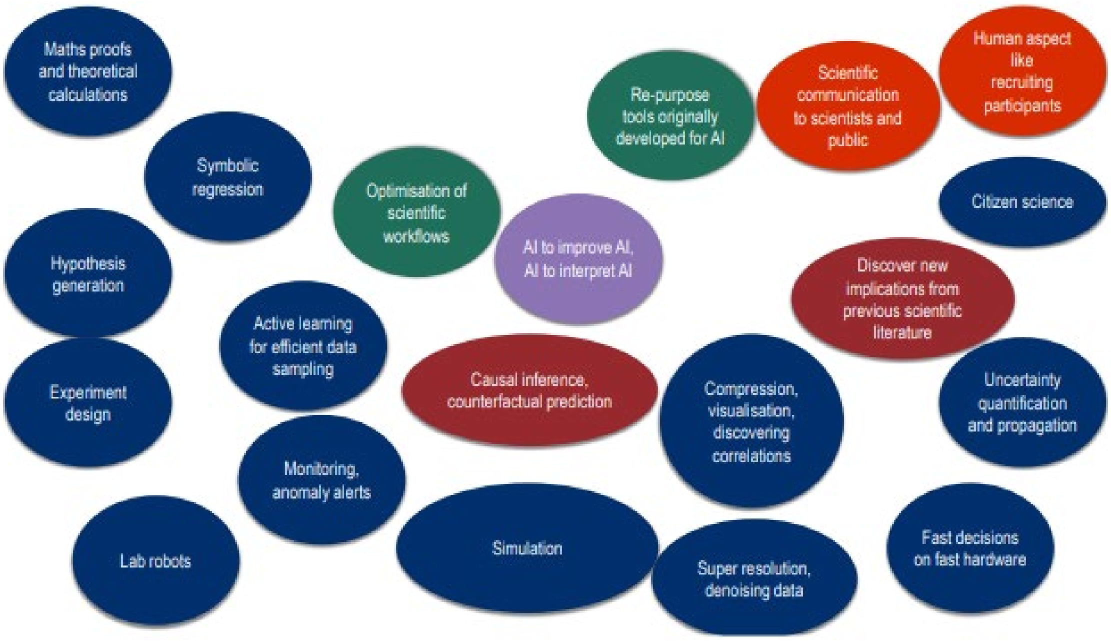

Figure 2: A (non-exhaustive) plethora of AI uses in science. Source: https://www.oecd-ilibrary.org/sites/a8e6c3b6-en/index.html?itemId=/content/component/a8e6c3b6-en
Note: Blue examples show where AI is directly used to improve a core aspect of the scientific progress; light red examples show uses that help set up studies or communicate results to peers or the public. Green bubbles represent the benefit to science not from AI directly but from the software and hardware infrastructure developed primarily for AI uses. Dark red bubbles refer to the frontiers of AI for science. Violet signifies AI for AI research

3. Principles of Research Integrity

According to the UK Committee on Research Integrity (UKCORI), research integrity entails conducting research in a trustworthy, ethical and responsible way[16]. It refers to factors underpinning good research practice and promoting trust and confidence in the methods used and the findings in results[17]. It is a set of requirements for individual researchers, their institutions, and the wider research community. Research Integrity applies to the whole research lifecycle, from the initial idea and design of a research project through the conduct of the research and the publication and dissemination of findings. To achieve research integrity, the environments and systems for research (often described as 'research culture') must safeguard and enhance good research practice rather than hinder it.

Various codes of practice are developed globally to govern the responsible conduct of research (Smith et al., 2024). These codes highlight several principles to ensure the credibility and reproducibility of research findings, forming the bedrock upon which public trust in research is built. Table 2 lists some international codes and highlights their common principles[18].

In our context, theConcordat to Support Research Integrity 2025 (see Figure 3) and the UKRIO Code of Practice for Research establish eleven principles with which researchers and research institutions are expected to comply, irrespective of the technological approaches and tools utilised in the research process. Researchers, readers, research organisations, funders and publishers are expected to contribute to a research system that describes how these broad principles should be applied to govern research associated with GenAI. Examples of critical questions will be:

---

[16] https://ukcori.org/what-research-integrity-is/

[17] https://ukrio.org/research-integrity/what-is-research-integrity/

[18] More international codes: https://new.nsf.gov/policies/responsible-research-conduct/international-contexts

- How can researchers apply the principle of transparency and open communication in disclosing their use of GenAI-based tools?
- How can researchers take steps to maintain the repeatability, reproducibility and replicability of their research when relying on data generated or analysed by GenAI-based tools?
- What steps should researchers take to ensure rigour in research processes that involve GenAI-based tools, and ensure consideration of risks around consent regarding data collection and usage.

It is noteworthy that GenAI-based tools are used in different fields of study, including Science, Technology, Engineering, and Mathematics (STEM), Humanities, Law, Business, and Social Sciences. Given the variation in the use of these tools in these fields, the requirements for responsible use may differ. Table 3 presents case studies of the use of GenAI in different disciplines.

Table 2: Examples of codes for the responsible conduct of research.

Codes	**Principles**
Australian Code for the Responsible Conduct of Research (the 2018 Code)	Honesty, Rigour, Transparency, Fairness, Respect, Recognition, Accountability, Promotion
European Code of Conduct for Research Integrity	Reliability, Honesty, Respect, Accountability
Concordat to Support Research Integrity	Honesty, Rigour, Transparency and Open Communication, Care and Respect, Accountability
UKRIO Code of Practice for Research	Repeatability, Reproducibility, Replicability, Trustworthiness, Credibility, Authenticity and Meta-research
Enhancing the Security and Integrity of America's Research Enterprise (The White House Office of Science and Technology Policy)	Openness and transparency, Accountability and Honesty, Impartiality and Objectivity, Respect, Freedom of Inquiry, Reciprocity and Merit-based competition
WHO Code of Conduct for Responsible Research	Integrity, Accountability, Independence and Impartiality, Respect for Persons and Communities and Professional Commitment
Asian-Pacific Economic Corporation (APEC) Guiding Principles for Research Integrity	Honesty, Responsibility, Rigour, Transparency, Respect, Fairness, Diversity
French Charter for Research Integrity	Compliance, Reliability of research work, Communication, Responsibility in collective work, Impartiality and independence in assessment and expertise, Collaborative work and plurality of activities, Training

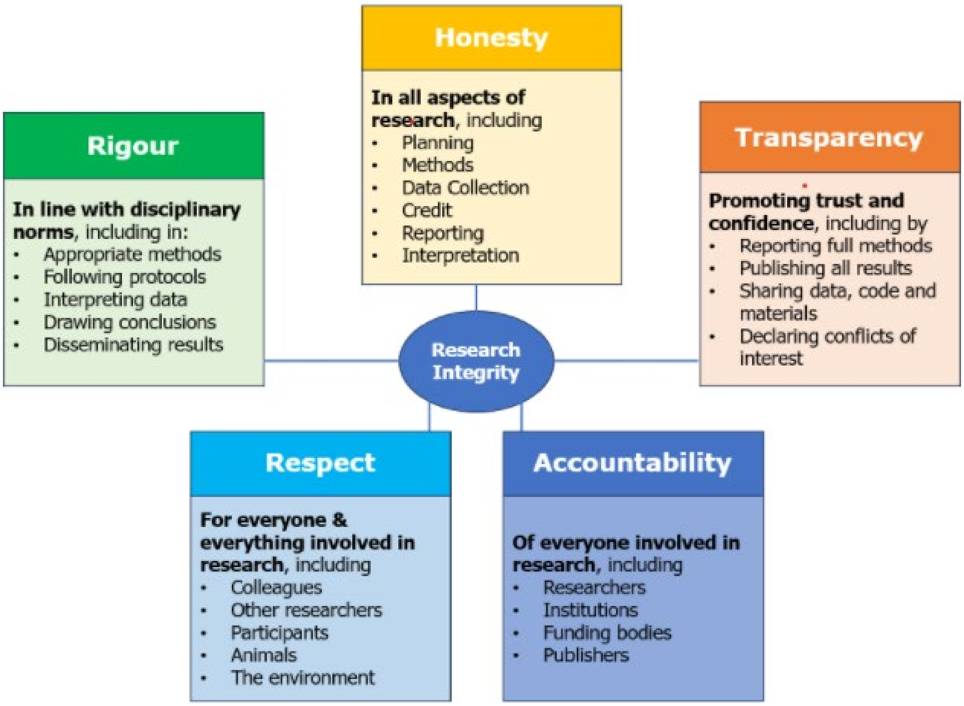

Figure 3: Graphical representation of research integrity based on the core areas described in The Concordat to Support Research Integrity 2025, created by UKRIO. Source: https://ukrio.org/research-integrity/what-is-research-integrity/

Table 3: Some case studies

Case Study 1: Mauritshuis Museum Case[19] **Subject Area: Art and Design**
Case Description: “One of the most famous paintings in history is literally being replaced by one of my A.I. images,” A user wrote on Instagram. After sending ‘Girl With a Pearl Earring’ out on loan, the Mauritshuis hung an AI facsimile in its place. The controversial portrait was submitted through an open call for ‘Pearl Earring’ substitutes launched by the museum. The museum initially selected five temporary artworks by contemporary artists in an open competition, out of which one was AI-generated. But Mauritshuis’ decision to select the AI-generated one has been controversial: critics condemn the choice to elevate machine-created images over the manual creation of real human artists.
What are the moral and ethical issues of this case? ● How do we define human originality when a GenAI-based tool is used in the process of creatorship? What is the trade-off between the credits that can be given to the humans and the GenAI? ● Can an AI-generated art be registered under the name of a human creator? ● Should AI-generated art compete with human-created art in an exhibition? ● When a GenAI-based tool is used in the process of creatorship, who owns the copyright of a piece of art between the tech company that owns the tool and the human creator who used it? ● What are the roles of the artist, institution (e.g., museum), other artists (chosen or outcompeted) and art curators, among others, in detecting and disclosing AI-generated art, and how do we define these responsibilities?

[19] Adapted from: Paschke, M. (2024). Cases for Research Integrity: Mauritshuis museum case. Originality and copyright when using AI-based tools. Zurich-Basel Plant Science Center: ETH Zurich.
DOI: 10.3929/ethz-b-000664648

What are the requirements for the principles of research integrity in this case?

- There is a need to define the concepts of creatorship, originality and copyright when a GenAI-based tool is used in the process of creatorship
- There is a need to identify the current limitations of existing standards and guidelines for the responsible use of GenAI in Art and Design

Case Study 2: The Ethics of Scientific Authorship in the Era of AI[20]
Subject Area: STEM

Case Description:

A doctoral student at a renowned university, was working on her cumulative dissertation. She needed to finish her PhD within a year, so she wanted to publish her two remaining papers as quickly as possible. She heard about ChatGPT from colleagues and started working with this new tool. She realised ChatGPT's potential for the scientific writing process and drafted her paper with ChatGPT.

She sent the first draft to her three co-authors without mentioning the use of this tool. Later, she received their feedback on the draft and implemented the changes, which all the co-authors approved. Before they submitted the manuscript, her supervisor and co-author shared the requirements for the ethical use of AI-based tools with her. As a result, the studentdeclared the use of ChatGPT for the draft. The authors checked online whether the target journal allows AI-based tools to generate content, and they found in the authors' guidelines that ChatGPT and similar tools are allowed when properly acknowledged in the paper. They prepared a table listing the exact model of the tool used, the use case, the location of the material involved, and other details. The supervisor carried out human supervision to ensure that the AI-generated content was not biased, unsafe or hallucinated.

After submitting the paper, they received a rejection from the journal with a stem note from the editor reminding all authors of the severity of plagiarism. The paragraphs generated by ChatGPT recreated passages verbatim from an earlier paper by one of the anonymous reviewers without citation. The reviewer flagged the plagiarism to the editor. All authors were mortified because they were unaware of the plagiarism, and all highly valued their scientific integrity.

What are the moral and ethical issues of this case?

- What does the transparency and accountability of scientific authorship entail when GenAI-based tools are used for manuscript preparation?
- How do we define scientific originality when GenAI-based is used in manuscript preparation? Should there be a threshold for how much text, data or imagery the tool should produce?

What are the requirements for the principles of research integrity in this case?

- The general consensus is that AI tools cannot be listed as paper authors. They are non-legal entities and, as such, cannot meet the requirements for authorship, including the responsibility for asserting the presence or absence of conflicts of interest, managing funding disclosure, copyright and licence agreements, carrying out fact-checking and source-checking, relying on trusted primary sources, adding citations, and avoiding plagiarism. Therefore, the guidelines for using AI in research should describe the responsibilities of humans for these activities to ensure research integrity.
- There is a need to clearly define when using AI-based tools does not jeopardise research integrity and the researchers' credibility.
- Beyond the responsibilities of researchers, the developers of GenAI-based research tools need to be aware of the principles of research integrity and how their tools should be trained to incorporate them.
- Publishers and funders should be able to identify and fact-check AI-generated content to flag inaccurate, misleading, or false information, hallucination, verbosity, copyright infringement, and non-citation of primary sources.

[20] Adapted from: Mihálka, R. and Paschke, M. (2024). Cases for Research Integrity: The ethics of scientific authorship in the era of AI. Zurich-Basel Plant Science Center: ETH Zurich. DOI: 10.3929/ethz-b000664648

Case Study 3: IBM Watson for Oncology,Subject Area: Medicine and Diagnostics[21][22]

Case Description:

The GenAI-powered IBM Watson for Oncology system was developed by IBM Corporation (USA) with the help of top oncologists from Memorial Sloan Kettering Cancer Center (MSK). It took more than 4 years of training, based on National Comprehensive Cancer Network (NCCN) cancer treatment guidelines and more than 100 years of clinical cancer treatment experience in the United States. It analyses a large amount of medical literature, patient data, and therapy recommendations using natural language processing, machine learning, and big data analytics methods and can recommend appropriate chemotherapy regimens for specific cancer patients.

IBM Watson for Oncology aids oncologists in quickly accessing pertinent medical information by processing and comprehending unstructured clinical material. It also offers recommendations for treatments that are supported by evidence. Oncologists typically use IBM Watson for Oncology as a decision support tool in the healthcare sector. It can be connected with electronic health records (EHRs) and other data to analyse patient data, including medical history, test results, pathology reports, and treatment recommendation sources. With the system, Oncologists can access a lot of clinical information with evidence-based treatment options by providing access to the most up-to-date research and clinical recommendations.

A significant additional benefit is the tailored therapeutic recommendations offered by IBM Watson for Oncology. The algorithm considers patient-specific characteristics like health history, genetic data, and treatment response to create personalised therapeutic options.

What are the moral and ethical issues of this case?

- How do we define informed consent to use data and confidentiality when GenAI models collect and analyse significant patient data and therapy recommendations?
- How do we define the requirements for algorithmic transparency vis-à-vis confidentiality, copyrights and intellectual property?

What are the requirements for the principles of research integrity in this case?

- It is important to note that despite being educated on a vast body of medical research, IBM Watson for Oncology’s suggestions could not always coincide with those of particular oncologists or established institutional policies. In health research contexts, human-in-the-loop criteria are essential for safety. A GenAI-based tool’s suggestions need extensive clinical confirmation and should be viewed as an additional tool to aid clinical judgment rather than a complete answer.
- When GenAI-based tools are used in the health research ecosystem, the requirements for transparency around the design of data sharing, interoperability, confidentiality, standards, accuracy, and explainability need to be defined to engender trust in the tools' outputs.
- In the health research ecosystem, research institutions and funders must define the criteria for evaluating and validating outputs from GenAI models.

4. Strategic Framework for Responsible Use of GenAI in Research

The scenarios in Table 3 show that ethical considerations are of the utmost importance in adopting GenAI in research. Ethical principles such as transparency, accountability, and honesty are vital to maintaining research ethical standards. Additionally, the scenarios reveal that the increasing use of GenAI in research may introduce new forms of research dishonesty and misconduct, such as data fabrication, hallucination, automated texts and artwork generation without appropriate references, which could jeopardise research

[21] Adapted from:S. Sai, A. Gaur, R. Sai, V. Chamola, M. Guizani and J. J. P. C. Rodrigues, "Generative AI for Transformative Healthcare: A Comprehensive Study of Emerging Models, Applications, Case Studies, and Limitations," in IEEE Access, vol. 12, pp. 31078-31106, 2024, doi: 10.1109/ACCESS.2024.3367715.

[22] Adapted from: Jie, Z., Zhiying, Z. & Li, L. A meta-analysis of Watson for Oncology in clinical application. Sci Rep 11, 5792 (2021). https://doi.org/10.1038/s41598-021-84973-5.

integrity, mislead research direction and, in turn, negatively impact the integrity and credibility of the researchers involved and the institution they are part of. Importantly, these scenarios reveal the different roles of researchers, readers, funders, publishers and research institutions in preserving research integrity. Researchers must adhere to the highest ethical standards irrespective of the technology used in their research. Research institutions and funders play a key role in establishing a research culture that supports research integrity, which includes providing guidelines, instruction and assistance to researchers. Publishers (including reviewers and editors) have a role in upholding ethical standards in the publishing and dissemination of research results. Readers are important in detecting and reporting research dishonesty and fraudulent activity. Furthermore, the scenarios reveal that more than standards and guidelines may be needed to address the ethical challenges posed by GenAI use in research (see Table 3, Case Study 2). Researchers must be educated on adhering to relevant guidelines and standards while using GenAI-based tools. There is also a need to empower publishers, funders and research institutions with the capabilities to uphold research ethical standards when GenAI-based tools are used.

To ensure responsible use of GenAI in research, we advocate a strategic framework consisting of three components, including i) structured guidelines and standards, ii) educational toolkits and iii) evaluation tools. Structured guidelines and technical standards are key to specifying the requirements for the responsible use of GenAI in research. Educational toolkits with inputs from various stakeholders are important to guide researchers in understanding the guidelines and standards and how to adhere to them. The toolkits can address skill gaps in applying GenAI guidelines or standards by providing a fast and straightforward approach that can rapidly be taken up across multiple domains. Evaluation tools will equip funders, publishers and research institutions with the capabilities to identify AI-generated content and evaluate researchers ’adherence to relevant guidelines and standards.

For the framework to be effective and adaptable across disciplines, an interdisciplinary approach is crucial to incorporate ethics and best practices for cross-disciplinary research environments and recognise the nuances in different disciplines. This requires continuous engagement and collaboration between science, technology, social sciences and humanities experts, as the requirements for responsible use of GenAI in research need to be elicited across disciplines. In addition, the framework must explicitly address critical questions (see examples in Section 3) that institutions and relevant stakeholders must note as a starting point to guide appropriate, effective and responsible use of GenAI in research and address its challenges. Furthermore, there is a need for direct engagement between the policymakers, industry, AI developers and AI researchers. This engagement is crucial to investigating and identifying the requirements for the responsible use of GenAI in research by directly exploring public opinions and expectations and communicating best practices, existing guidelines and standards to the relevant stakeholders. The engagement needs to be embedded into the existing AI ecosystems and knowledge exchange frameworks to make the outcomes widely accessible to the stakeholders involved.

4.1. Controlling GenAI: Guidelines and Standards

The development of GenAI so far mirrors the ethos of the early Internet, championed by Tim Berners-Lee (the father of the Internet), as a space of openness and freedom. However, this approach to GenAI development suggests the need for governance frameworks that include guidelines and standards.

GenAI guidelines and standards should address data quality and robustness concerns, ethical considerations (such as privacy, transparency, fair and efficient use, reproducibility, honesty, originality and accountability), and security requirements for safe, secure, and equitable data transfers. Other essential requirements include the responsible use of GenAI and due diligence, the provenance and quality of training data, and the evaluation of the reliability and trustworthiness of the outputs of GenAI-based tools.

While many institutions still need to implement guidelines and standards for using GenAI in research, several initiatives, guidelines and standards have been developed. These include the guidelines published by COPE, which several journals have endorsed to ensure that GenAI tools are accountable, transparent, and consistent with ethical principles for their use in research. In this article, we argue that a number of important questions remain concerning ethics in GenAI. The scope of when and how the use of GenAI falls within the definition of research dishonesty has yet to be clearly defined. A comprehensive and functional set of guidelines and standards is lacking to guide the responsible use of GenAI in research. Third, funders and publishers need help identifying and evaluating AI-generated content due to the natural language of GenAI outputs. This is essential to evaluate researchers' adherence to guidelines and standards on the responsible use of AI. To ensure authors/researchers take full accountability for using AI in research, there should be better training and education regarding its responsible use – particularly pertinent to early career researchers but applicable to all. There should be a standardised reporting process for using GenAI in research.

Although GenAI guidelines and standards can ensure trustworthiness, transparency, and common definitions and frameworks, limitations should be recognised. Some guidelines or standards may reduce flexibility and innovation, and the time involved in developing them may result in their redundancy.

4.2. Explaining GenAI

Although researchers are incorporating GenAI-based tools in different aspects of research to supplement traditional methods, their understanding of how to use the tools responsibly while maintaining research integrity is limited. According to Liu and Jagadish (2024), research institutions are not set up to be agile in the face of rapidly advancing technologies; adopting new technologies usually falls on individual researchers. By implication, researchers with various technical expertise adopt GenAI-based tools without an adequate understanding of their dynamics and implications for research integrity. Liu and Jagadish (2024) argue that the current norm of relying on individual researchers for new technology adoption is no longer adequate. There is a need for research organisations to develop new mechanisms to help researchers adopt new technologies, especially those that cause major seismic shifts, such as GenAI.

This article suggests the need for researchers to understand GenAI and their impacts on research integrity. There is a need for educational toolkits on using GenAI to enhance critical thinking and problem-solving skills, identifying research gaps and generating original ideas, among other uses, without compromising research integrity. This includes i) an explainer of the technology, its impacts on research and relevant guidelines and standards, and ii) a step-by-step guide on the ethical principles and practices to be considered when using GenAI-based tools in a particular context. To help researchers fully understand the implications of GenAI on research integrity, the toolkits need to incorporate various explanation styles, including textual and graphical explanations, to describe various concepts and address researchers 'questions. Previous research in Explainable AI has shown that the explanation of AI systems significantly impacts people's perception of trust in the technology (Chen et al., 2023; Ogunniye et al, 2021).

4.3. Evaluating GenAI

Critical challenges surround the question of authenticity, originality, and honesty in research in the era of GenAI when machine-generated content can be indistinguishable from human content. For example, Case Study 1, presented in Table 3, revealed an AI-generated artwork selected in an exhibition over human-created ones. This raises concerns about whether AI-generated content should be selected over human-created ones. A more critical concern surrounds the question of responsibility. Who should be responsible for AI-generated content? For example, if a GenAI algorithm generates false, biased or offensive content, who takes the responsibility: the user, the algorithm creator, the algorithm itself, the data provided to the algorithm, or the person curating training data and/or supervising the learning process?

In the research context, GenAI's ability to generate content can alleviate researchers ' workload, but it could also blur the line between technology-assisted research and research dishonesty. Ultimately, GenAI places a critical responsibility on publishers, readers, researchers, funders, and research institutions to evaluate the impact of technology on research integrity. It is not just about adopting GenAI-based tools but critically evaluating how they can change the research culture. Importantly, there is a need to foster responsible and ethical development of GenAI-based tools and their use in research. This involves understanding the risks and challenges of GenAI in research and developing appropriate tools to mitigate them.

To ensure that ethical research principles are strictly adhered to, this article suggests equipping publishers, funders, and research organisations with an evaluation tool to identify, fact-check, source-check, and evaluate AI-generated content and ensure it is appropriately referenced and disclosed. In addition, the tool should be robust enough to detect and, if required, suggest amendments when ethical research principles are violated.

5. Call to Action

The transformative impact of GenAI on research has come to stay. To improve understanding of the impact of these tools, enhance their effectiveness, and ensure that they are used responsibly, support and collaboration between members of the GenAI development and research user ecosystems are needed. Cooperation is essential to help develop the body of knowledge to train GenAI models and evaluation tools to enhance the reliability and trustworthiness of their outputs. Given the scale and dynamics of GenAI, performing a reliability and trustworthiness check is a daunting task that exceeds human capabilities. What is required are ideas beyond simple retrospective checks, along with adequate time and space to consider and concentrate on the more significant issues and problems, not least imagining the possible avenues of further GenAI development and the consequent emerging capabilities and risks.

6. Concluding Remarks

In this article, we have outlined a number of issues to be addressed for the ethical and responsible use of GenAI in research. First, the scope of when and how the use of GenAI falls within the definition of research dishonesty needs to be clearly defined across different fields of study. Second, a comprehensive and functional set of guidelines and standards is needed to guide the responsible use of GenAI in research. Third, researchers need to be equipped with educational toolkits to understand GenAI and their impacts on research integrity. Lastly, funders and publishers need to be equipped with evaluation tools to identify and evaluate AI-generated content.

Alongside the exponential increase in the capabilities of GenAI models, an audience split between unbridled trust and acceptance and those who call for a ban; questions of trust are crucial. This article recommends the need for educational toolkits and evaluation tools to supplement guidelines and standards for the use of GenAI toward ensuring that GenAI improves the efficiency of some repetitive research tasks and mitigates the risks of inaccuracy, dishonesty, and machine hallucinations.

References

Amato, A., & Coronato, A. (2018). Supporting hypothesis generation by machine learning in smart health. In Innovative Mobile and Internet Services in Ubiquitous Computing: Proceedings of the 11th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS-2017), pages 401–410. Springer.

Barreto, F., Moharkar, L., Shirodkar, M., Sarode, V., Gonsalves, S., & Johns, A. (2023). Generative artificial intelligence: Opportunities and challenges of large language models. In International Conference on Intelligent Computing and Networking, pages 545–553. Springer.

Benjamin, R. (2019). Race after technology: Abolitionist tools for the new Jim code. John Wiley & Sons.

Bhatia, G., & Kulkarni, A. (2023). Chatgpt as co-author: Are researchers impressed or distressed? Asian journal of psychiatry, 84:103564.

Chen, C., Tian, A. D., & Jiang, R. (2023). When post hoc explanation knocks: Consumer responses to explainable AI recommendations. Journal of Interactive Marketing, 10949968231200221.

Crawford, J., Cowling, M., Ashton-Hay, S., Kelder, J.-A., Middleton, R., & Wilson, G. S. (2023). Artificial intelligence and authorship editor policy: Chatgpt, bard bing ai, and beyond. Journal of University Teaching & Learning Practice, 20(5):1.

Daniel, B. K. (2018). Empirical verification of the “TACT” framework for teaching rigour in qualitative research methodology. Qualitative Research Journal, 18(3), 262-275.

Dhakal, R. K. (2016). Responsible practice of research: Safeguarding research integrity and publication ethics. Journal of education and research, 6(2):1–11.

Ganjavi, C., Eppler, M. B., Pekcan, A., Biedermann, B., Abreu, A., Collins, G. S., Gill, I. S., & Cacciamani, G. E. (2024). Publishers ’and journals ’instructions to authors on use of generative artificial intelligence in academic and scientific publishing: bibliometric analysis. bmj, 384.

Golan, R., Reddy, R., Muthigi, A., & Ramasamy, R. (2023). Artificial intelligence in academic writing: a paradigm-shifting technological advance. Nature reviews urology, 20(6):327–328.

Graf, A., & Bernardi, R. E. (2023). Chatgpt in research: balancing ethics, transparency and advancement. Neuroscience, 515:71–73.

Hosseini, M., Resnik, D. B., & Holmes, K. (2023). The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts. Research Ethics, 19(4):449–465.

Huang, J., & Tan, M. (2023). The role of chatgpt in scientific communication: writing better scientific review articles. American journal of cancer research, 13(4):1148.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596(7873), 583-589.

Kapoor, S., & Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science.

Patterns, 4(9).

Khalifa, M., & Albadawy, M. (2024). Using artificial intelligence in academic writing and research: An essential productivity tool. Computer Methods and Programs in Biomedicine Update, 100145.

Kidd, C., & Birhane, A. (2023). How ai can distort human beliefs. Science, 380(6651):1222–1223.

Lecler, A., Duron, L., & Soyer, P. (2023). Revolutionizing radiology with gpt-based models: current applications, future possibilities and limitations of chatgpt. Diagnostic and Interventional Imaging, 104(6):269–274.

Liu, J., & Jagadish, H. V. (2024). Institutional Efforts to Help Academic Researchers Implement Generative AI in Research. Harvard Data Science Review, (Special Issue 5). https://doi.org/10.1162/99608f92.2c8e7e81

Macdonald, C., Adeloye, D., Sheikh, A., & Rudan, I. (2023). Can chatgpt draft a research article? an example of population-level vaccine effectiveness analysis. Journal of global health, 13.

Macfarlane, B. (2010). Researching with integrity: The ethics of academic enquiry. Routledge.

Mebane, C. A., Sumpter, J. P., Fairbrother, A., Augspurger, T. P., Canfield, T. J., Goodfellow, W. L., ... & Verslycke, T. A. (2019). Scientific integrity issues in environmental toxicology and chemistry: improving research reproducibility, credibility, and transparency. Integrated environmental assessment and management, 15(3), 320-344.

Messeri, L., & Crockett, M. (2024). Artificial intelligence and illusions of understanding in scientific research. Nature, 627(8002):49–58.

Ogunniye, G., Legastelois, B., Rovatsos, M., Dowthwaite, L., Portillo, V., Vallejos, E. P., ... & Jirotka, M. (2021). Understanding user perceptions of trustworthiness in e-recruitment systems. IEEE Internet Computing, 25(6), 23-32.

Rahman, A. (2020). Algorithms of oppression: How search engines reinforce racism. New Media & Society, 22(3), 575–577. https://doi.org/10.1177/1461444819876115

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature machine intelligence, 1(5):206–215.

Smith, S., Tate, M., Freeman, K., Walsh, A., Ballsun-Stanton, B., Hooper, M., & Lane, M. (2024). A University Framework for the Responsible use of Generative AI in Research. arXiv preprint arXiv:2404.19244.

Templin, T., Perez, M. W., Sylvia, S., Leek, J., & Sinnott-Armstrong, N. (2024). Addressing 6 challenges in generative AI for digital health: A scoping review. PLOS Digital Health, 3(5), e0000503.