Innovation or Imitation? A Critical Analysis of AI-Authored vs. Human-Authored Scientific Papers

Corinna Hörmann1, Lisa Kuka1, Anneliese Fraser2, and Barbara Sabitzer3

1STEM Education, Johannes Kepler University Linz, Altenbergerstraße 69, Linz, Austria
2Teacher Education Centre, University of Passau, Innsbrucker Straße 41, Passau, Germany

Keywords: Artificial Intelligence, Academic Publishing, Human vs AI Authorship, Digital Transformation.

Abstract: The use of Artificial Intelligence (AI) in research and academic publishing has been a topic of growing interest and debate. While some argue that AI-based systems have the potential to revolutionize the way scientific papers and academic work are generated, others express concerns about the authenticity of AI-authored papers. Several respected organizations have recently developed guidelines regarding the use of AI in scholarly manuscripts and publishing. This critical analysis will examine the advantages and disadvantages of AI-authored scientific papers compared to those authored by humans. The underlying work describes the process of creating different papers solely with the help of ChatGPT and Jenni AI and compares them to human written drafts. Therefore, both AI tools were asked to generate scientific papers about the “History of Digital Education in Austria”, “A History of Women in Computer Science”, and “Modelling of Mental Arithmetic Strategies Using UML”. In conclusion, it is indisputable that AI-driven tools significantly facilitate the drafting of outlines, titles, and the composition of papers. However, the creation of a high-quality scientific academic publication still demands considerable human input, encompassing both creative effort and critical thinking, to ensure depth, originality, and scholarly rigor.

1 INTRODUCTION

During the resulting Emergency Remote Teaching (ERT) due to the COVID-19 pandemic, students tried to find new ways to interact with technology and adapted to new tools (Vargo et al., 2021). Artificial Intelligence (AI) technologies, including language translation, plagiarism detection, grammar and spelling checks, and AI-generated essay outlines, are widely acknowledged as beneficial tools that enhance the writing process and assist scholars and students in their academic endeavors. Furthermore, research shows that AI improves students’ writing skills, sense of self-efficacy, and comprehension of academic integrity, all of which have a favorable impact on academic writing. Still, students seem to be worried about how it will affect their ability to be creative, think critically, and behave ethically (Malik et al., 2023).

With the inclusion of Large Language Models (LLMs) into Digital Writing Assistants (DWAs) like Grammarly or WordTune, students became confused whether it is legitimate to use these tools for academic writing. It is even possible to generate citations alongside the work created by the AI, which are referred to as Automatic Article Generators (AAGs) (El-Sayed Abd-Elaal and Mills, 2022). Researchers, educators, and anti-plagiarism software all face additional work when identifying a piece of work produced by AAGs. More and more universities struggle with the concerns related to the integration of LLMs in student projects. Several authors already tried to answer the question, whether or not using AI technologies is a violation of academic integrity (Perkins, 2023). The International Center of Academic Integrity (ICAI), established in 1992 by notable scholars, provides an authorized definition of academic integrity. Its founder, Don McCabe, is recognized for having popularized the term. The Center first defined the “fundamental values of academic integrity” in 1999 as honesty, trust, fairness, respect, and responsibility. In 2014, they added the virtue of courage as a sixth component. According to the ICAI, academic integrity is
2 AI TOOLS IN THE UNDERLYING EXPERIMENT

For this experiment, two different tools were used to generate content for a scientific paper: ChatGPT-4 and Jenni AI (Ultimate Plan).

In the present study, ChatGPT-4, available at a subscription cost of $20 per month, was selected for its enhanced capabilities over its predecessor ChatGPT-3. These improvements include more precise responses to queries, the ability to generate images through AI (integrating DALL-E), advanced web browsing functionalities, and superior data analysis features.

The premium version of Jenni AI, priced at $20 per month, offers users an unrestricted word count allowance daily. In contrast to the complimentary variant, Jenni AI imposes a limit of 200 AI-generated words per day.

The significant difference in creating scientific papers with ChatGPT to Jenni AI is that the latter makes a whole draft with a single line of input, whereas ChatGPT users have to request every single step and refine them. Still, both tools can be seen as AI co-pilots that assist or augment human capabilities in various tasks or processes (see Figure 1).

Figure 1: Co-piloting with AI tools (Revell et al., 2023) (adapted by the authors).

2.1 ChatGPT

ChatGPT by OpenAI was chosen for this experiment because of obvious reasons. It currently is the most well-known AI chatbot and a very potent tool. Neither is ChatGPT the first GPT model, nor the first language model. However, it significantly advanced natural language processing by popularizing huge language models and quickening the use of artificial intelligence. In ChatGPT, the GPT stands for “Generative Pre-Trained Transformer”, a LLM that mimics human speech through deep learning (Yosifova, 2023). ChatGPT can offer several significant benefits in idea generation, outline creation, or writing and editing support.

2.2 Jenni AI

In contrast to ChatGPT, Jenni AI is a specific tool for academic purposes, enabling to correct grammatical mistakes, provide citations in the work (which is not possible with ChatGPT), and create strong arguments. Still, it has to be noted that Jenni AI is also GPT based. With its emphasis on academic writing and integrated plagiarism detector, Jenni AI is ideal for both academic essays and personal statements (Originality.ai, 2023). Jenni AI was created in 2016 by David Park and Henry Mao. In 2020, Park relaunched Jenni AI, and because of some new features and TikTok’s attention, more people started using it. Most AI programs just reply to queries with long text passes...
sages. Jenni AI works on user’s behalf more closely and learns from the input. She will also remind the prompt engineer to write for him- or herself if he or she depends too much on the AI’s help (Giles, 2023).

3 METHODOLOGY

This experiment is designed to critically analyze and compare the quality, authenticity, and overall effectiveness of scientific papers authored by Artificial Intelligence (AI) tools, specifically ChatGPT and Jenni AI, against those authored by human researchers. The experiment focuses on three distinct topics:

1. History of Digital Education in Austria
2. History of Women in Computer Science
3. Modeling of Mental Arithmetic Strategies Using UML

One of the human-authored papers (History of Digital Education in Austria) was written in English by the authors themselves and accepted at a subject relevant international conference (International Conference on Informatics in Schools) (Hörmann et al., 2022). The two other works (History of Women in Computer Science, Modeling of Mental Arithmetic Strategies Using UML) were handed in in German as Bachelor theses at the authors’ department and were both graded with “Excellent”. The human authors were briefed about the scope and expectations of the experiment, ensuring their anonymity. To maintain an unbiased comparison, the human authors independently developed their texts without exposure to the AI-generated content.

The underlying experiment utilizes two AI tools, ChatGPT and Jenni AI. These tools were selected based on their advanced language processing capabilities and relevance in academic research. For each topic, both AI tools were tasked to create comprehensive scientific text. The inputs provided were standardized to ensure consistency in the information available to both AI systems.

The AI-crafted papers were evaluated on multiple dimensions including accuracy of content, depth of analysis, coherence in structure, and adherence to academic writing standards. Furthermore, each work has been tested with the plagiarism detection tool “Turnitin Similarity” which is commonly employed in academic and educational settings to ensure the integrity and originality of written work.

The three human-authored papers, as well as the six AI-created drafts, can be found online on GitHub following the QR code in Figure 2 or the link https://github.com/corinnahoermann/AI-vs-Human_Paper.

4 EXPERIMENT

4.1 Paper 1: History of Digital Education in Austria

4.1.1 Human-Authored Paper

The underlying human-authored paper is entitled “From Non-Existent to Mandatory in Five Years – The Journey of Digital Education in the Austrian School System” and explores the evolution of the subject Digital Education in Austria, from the introduction of Computer Science in schools in 1985 to the implementation of Digital Education as a mandatory subject in the 2022/23 school year, encompassing digital competences, media literacy, and civic education. The manuscript spans a total of ten pages, including about 30,000 characters, and adheres to the conventional structural framework typical of literature-based academic publications:

(1) Introduction
(2) Computer Science in Austrian Schools
(3) Digital Education in Austria
   a. Masterplan for Digitalization
   b. 8-Point-Concept
   c. Introduction of the Subject Digital Education
   d. Compulsory Subject Digital Education in Austria
(4) Conclusion and Outlook
(5) References

4.1.2 AI-Authored Draft – Jenni AI

To initiate the process of composing a paper using Jenni AI, the following steps are required: The first step is to define what “you are writing today?”. To re-create the human-authored paper the prompt a research paper about the history of the school subject “Digital Education” (digitale Grundbildung) in
Austria was typed in. Jenni AI assessed the input as “Great prompt” and the authors also ticked the box “Outline Builder: automatically create document headings”.

Immediately, the AI-generated document, entitled “Austrian Digital Education History”, was presented, exhibiting an organizational framework. The structure of the paper was delineated as follows:

1. Introduction to Digital Education in Austria
2. Historical Development of Digitale Grundbildung
3. Evolution of Digital Education Curriculum in Austrian Schools
4. Key Milestones in Austria’s Digital Education Journey
5. Impact of Digital Education on Austrian School System
6. Challenges and Solutions in Implementing Digitale Grundbildung
7. Case Studies: Successful Implementations of Digital Education in Austria
8. Future Outlook: Prospects for Digital Education in Austria
9. Conclusion: Reflections on the History of Digital Education in Austria
10. References

Next, the authors just let the tool fill the sections with text without editing it any further and Jenni AI came up with seven pages in the first version. As there was no abstract, the authors added the headline manually and again let the tool fill the section with text on its own. In total, the AI drafted seven pages with about 24,000 characters, as well as eight valid references. Thirty percent of the information was flagged by the “Turnitin” software as possibly copied, which was the highest of all papers.

4.1.3 AI-Authored Draft – ChatGPT

For creating a paper in ChatGPT the authors used a similar prompt: *I want to write a research paper about the history of the school subject “Digital Education” (digitale Grundbildung) in Austria* and ChatGPT provided tips to structure the work:

1. Introduction
2. Historical Overview
3. Conclusion

There was no abstract and no references provided. Still, the authors separately asked for an abstract.

When prompting: *Can you create a title for this paper for me?*, ChatGPT provided “Navigating the Digital Wave: A Historical Analysis of Digital Education in Austria”. The prompt engineers then had to separately tell ChatGPT to write the beforehand defined sections, like: *Can you write the introduction for me concerning this paper?*

It has to be mentioned that in this case ChatGPT was not able to provide valid references as it cannot access or verify some sources, such as the latest scientific journals, books, or articles. Still, ChatGPT drafted ten pages with about 30,000 characters, while the “Turnitin” software detected 6% of potential plagiarism.

Moreover, the authors instructed ChatGPT to produce an image that visually conveys the paper in order to illustrate it (see Figure 3).

![Image created by ChatGPT visually representing “Navigating the Digital Wave: A Historical Analysis of Digital Education in Austria”](image_url)

Figure 3: Image created by ChatGPT visually representing “Navigating the Digital Wave: A Historical Analysis of Digital Education in Austria”.

4.2 Paper 2: History of Women in Computer Science

4.2.1 Human-Authored Paper

The bachelor thesis with the title “The Female Side of Computer Science” comprises 46 pages (about 85,500 characters) in total and highlights the often overlooked contributions of women in the field of Computer Science, despite their active and ongoing involvement in its rapid advancements. It aims to showcase a selection of pioneering women, detailing their lives and achievements, and examining how their significant research has shaped and continues to influence the field. It is outlined like the following:

1. Introduction
4.2.2 AI-Authored Draft – Jenni AI

Again the authors prompted a single line: *I want to write a thesis about the most important women in Computer Science* and confirmed the outline builder. The work was entitled “Influential Women in Computer Science” and had the following structure:

1. Introduction to Women in Computer Science
2. Historical Context of Women’s Contributions to Computing
3. Challenges Faced by Women in Computer Science
4. Key Female Figures in the Development of Computer Science
5. The Impact of Women on Computer Science Evolution
6. Challenges and Triumphs of Women in the Tech Industry
7. Case Studies: Influential Women in Computer Science
8. Future Outlook: Encouraging Female Participation in Computing
9. Conclusion: Recognizing Women’s Role in Shaping Computer Science

The abstract was again missing but prompted by the authors.

Like before, the text was generated automatically and the AI tool came up with a total of seven pages, consisting of approximately 24,000 characters. Despite the activation of the “auto cite from new sources - external sources will be considered” feature, Jenni AI did not include any external references in its output. Moreover, the content was found to contain 24% instances of plagiarism according to the “Turnitin” software.

4.2.3 AI-Authored Draft – ChatGPT

Like before, the authors asked ChatGPT to write a bachelor thesis about the most important women in computer science and to create a title. The AI tool came up with “Pioneering Code: The Untold Stories of Women in Computer Science”. The outline that was provided, looked like the following:

1. Introduction
2. Historical Context and Gender Dynamics in STEM
3. The Pioneers of Computer Science
4. The ENIAC Programmers and Post-War Computing
5. Modern Trailblazers in Computer Science
6. Contemporary Issues and Progress
7. Conclusion

Fourteen pages, or about 43,000 characters, were produced by ChatGPT in all, while the “Turnitin” software identified 17% of the content as potentially plagiarized.

To represent the thesis visually, the authors told ChatGPT to create an image that represents this thesis (see Figure 4).

![Figure 4: Image created by ChatGPT visually representing “Pioneering Code: The Untold Stories of Women in Computer Science”.](image-url)
4.3.2 AI-Authored Draft – Jenni AI

When Jenni AI was asked to create a bachelor thesis about the integration of Unified Modeling Language (UML) into Mathematics education, specifically in the context of mental arithmetic, it produced six pages with approximately 17,000 characters and four reliable references. Additionally, the “Turnitin” program revealed that there were 13% instances of plagiarism in the content. The following outline for the thesis entitled “UML Integration in Mental Arithmetic: Bachelor Thesis” was created by Jenni AI:

(1) Introduction
(2) Exploring the Role of UML in Enhancing Mental Arithmetic
(3) Theoretical Framework: UML Meets Mathematics Pedagogy
(4) Case Studies: UML Application in Learning Mental Arithmetic
(5) Methodology for Integrating UML into Math Curricula
(6) Assessing the Impact of UML on Mathematical Cognitive Development
(7) Challenges and Solutions in the Implementation of UML Tools
(8) Conclusion and Outlook
(9) References

The prompt engineers also told Jenni AI to create an example using an UML diagram how mental arithmetic could be displayed, but the tool did not understand the instruction.

4.3.3 AI-Authored Draft – ChatGPT

A total of 18 pages with about 63,000 characters was generated by ChatGPT when asked to create a bachelor thesis (several sub-prompts were typed in of course). However, the “Turnitin” program found 5% of possible plagiarism. The tool drafted the title “Integrating Unified Modeling Language (UML) into Mathematics Education: A Theoretical Approach to Enhancing Mental Arithmetic Skills” and outlined the thesis the following way:

(1) Introduction
(2) Literature Review
(3) Theoretical Framework for UML Integration in Mathematics Education
(4) Hypothetical Application in Education
(5) Potential Impact and Hypothetical Outcomes
(6) Conclusion

Figure 5: UML diagram outlining the process of solving $(5 + 3) \times 2$ created by ChatGPT.

5 DISCUSSION

Jenni AI does not draft an abstract by default, despite being one of the most popular text editors for creating research papers or other academic work. However, when this AI tool creates one, the output is valid. Moreover, the text of Jenni AI was somewhat superficial, characterized by a tendency to reiterate points without substantial depth. Jenni AI takes a lot of work to produce long texts, as one must repeatedly motivate the AI to draft additional text. Furthermore, it can be observed that the titles Jenni AI creates are simple and could be more creative. Still, the option to utilize valid references is a huge help.

In contrary to Jenni AI, ChatGPT does not draft a whole paper by prompting a single sentence. The user has to specify many prompt lines and copy the results into a text-processing application on their own. The
The plagiarism checker’s response to AI-generated texts showed an interesting pattern. In texts drafted by ChatGPT, approximately 9% on average (5 - 17%) of the content was flagged for potential plagiarism. This relatively low percentage suggests a degree of originality in AI-generated content. However, a notable increase in potential plagiarism was observed in Jenni AI generated texts, where the detection rate was significantly higher with about 22% on average (13 - 30%). The discrepancy in plagiarism detection between texts generated by Jenni AI and ChatGPT raises important questions concerning the data sources and methods employed by different AI tools. It implies that Jenni AI may depend more on pre-existing text sources, which then raise the rate of content matching. On the other hand, texts produced by models like ChatGPT have a lower detection rate, which suggests an advanced approach to content generation that may involve more innovative information combinations.

5.1 Discussion of Topic 1: History of Digital Education in Austria

Each paper follows a traditional academic structure but focuses on different aspects of digital education in Austria. The ones by ChatGPT and Jenni AI are more focused on historical and policy perspectives, while the human-authored paper includes a comparative analysis and specific case studies.

5.2 Discussion of Topic 2: History of Women in Computer Science

Every paper adopts a somewhat similar structure, starting with an introduction and background, moving into detailed discussions of key female figures, and concluding with future outlooks and recommendations. However, the focus and depth of content differ, with some emphasizing biographical case studies and others more on the overall impact and challenges. In the assessment of the auto-generated content of both AI tools, it could be observed that the curriculum vitae of the selected female scientists could be used without modification, demonstrating a satisfactory level of detail and accuracy. However, Jenni AI produced a somehow confusing outline, by creating two different chapters dealing with challenges of women in Computer Science (“Challenges Faced by Women in Computer Science”, “Challenges and Triumphs of Women in the Tech Industry”) and another two for describing female figures (“Key Female Figures in the Development of Computer Science”, “Case Studies: Influential Women in Computer Science”).

5.3 Discussion of Topic 3: Modeling of Mental Arithmetic Strategies Using UML

Each paper adopts a structured academic approach, starting with an introduction and theoretical background, moving into the application and analysis of UML in educational contexts, and concluding with implications and future outlooks. However, they vary in their focus areas, with two papers concentrating more on mental arithmetic, while the third paper by ChatGPT takes a broader view of UML’s integration in mathematics education. However, the human-authored one is the only one that includes valid UML diagrams.

6 CONCLUSION AND OUTLOOK

The exploration of AI-authored versus human-authored scientific texts in this experiment highlights the evolving landscape of academic writing. AI tools, represented by ChatGPT and Jenni AI, demonstrate proficiency in generating structurally sound and coherent academic texts. However, their capabilities are currently best utilized as augmentative tools rather than replacements for human intellect and creativity.

The experiment’s investigation into the capacities and limitations of AI in academic writing demonstrates the current state and potential of AI tools in scholarly work. While AI tools like ChatGPT and Jenni AI demonstrate remarkable abilities in drafting structured, coherent papers and theses, they still require human oversight for depth, originality, and academic rigor. Particularly, the underlying work reveals that AI can efficiently generate outlines, titles, and even complete drafts, but these outputs often lack in understanding and critical analysis that human expertise brings. Moreover, the experiment highlights a significant difference in the approach of AI tools in creating content. For instance, Jenni AI’s ability to
draft entire papers with minimal input contrasts with ChatGPT’s requirement for detailed prompts at each step. This distinction underscores the diverse methodologies inherent in AI tools and their varied applications in academic writing.

The ethical implications of AI in academic authorship also form a critical part of future discussion, as there is a vast need for clear guidelines regarding the use of AI in scholarly work. This is especially important as AI tools become more and more accessible.

To further grasp the consequences of AI in educational contexts, more empirical research is necessary, particularly in fostering creativity, critical thinking, and maintaining academic integrity. Additionally, the exploration of AI’s potential in aiding diverse academic tasks, such as data analysis and literature review, is recommended as a future research direction. In conclusion, while AI represents a significant advancement in academic writing, its role remains supplementary to human intellect and creativity. The findings of this experiment lay a foundation for future exploration into the evolving relationship between AI and human authorship in academia, emphasizing a need for an ethical balanced approach to integrate AI tools in scholarly practices.

REFERENCES


International Center for Academic Integrity (2021). The fundamental values of academic integrity third edition.


