Unveiling Biases: An Exploration of ChatGPT-3.5-generated 'Technology Stories'

Cecilia Axell. Linköping University, Sweden
cecilia.axell@liu.se
Johan Boström. Linnaeus University, Sweden
johan.bostrom@lnu.se

ABSTRACT

A technology that is increasingly affecting our daily lives is artificial intelligence (AI). An example of such a technology is ChatGPT-3.5, which has received a lot of attention recently. ChatGPT-3.5 is a text generator that is developed on a large number of existing texts. Currently, there is a debate about negative consequences in education, for example, if students let the chatbot write texts for them. In this study, however, our point of departure is on how ChatGPT-3.5 and storytelling can be used as a tool in teacher education to develop students' critical thinking in relation to technology.

A main objective of technology as a school subject is to prepare pupils to act in a technology-intensive world, which includes critical thinking about technology and its impact on individuals, society, and nature. However, a critical aspect of technology education is that it easily becomes an unreflective doing without a meaningful context. A way to circumventing this problem is to use storytelling in technology teaching.

This is a pre-study for a coming project aiming to let teacher students create stories using ChatGPT-3.5 and then critically analyse the technological content in the stories. In this pre-study, we gave ChatGPT-3.5 the instruction to generate ten shorter stories for children with a focus on technology. A qualitative content analysis shows that there are several dominant themes within the stories, and that the chatbot presents a view of technology that is mainly positive and without any critical reflection on its effects on individuals, nature and society. Furthermore, in the stories, high-tech male coded technology is a dominating theme. The pre-study highlights the importance of critical thinking and reflections when using AI tools in technology teacher education. It also indicates that stories generated by a chatbot can be a steppingstone to visualise technology bias and contribute to developing teacher students' critical gaze.

Key Words: Technology Education, ChatGPT-3.5, Storytelling, Critical Thinking

1. INTRODUCTION

A main objective of technology as a school subject is to prepare pupils to act in a technology-intensive world, which includes critical thinking about technology and its impact on individuals, society, and nature (Skolverket [The Swedish National Agency for Education], 2022). However, a critical aspect of technology education is that it easily becomes an unreflective doing without a meaningful context (Skolinspektionen [The Swedish Schools Inspectorate], 2014). A way to circumventing this problem is to use storytelling in technology teaching (e.g., Axell, 2017; Svensson et al., 2019). However, for a story to be used in technology teaching, the teacher first needs to do a critical reading in order to identify what messages about technology it conveys (Axell, 2015; 2017). On the basis of this, in a previous study (Axell & Boström, 2021), we investigated the technological content in a selection of picture books. The results showed that in the books, there is a focus on how separate artefacts function but no detailed explanation of how these artefacts are connected or what kind of implications they have in a societal context. There was also an emphasis on traditionally masculine coded technology.

A technology that is increasingly affecting our daily lives is artificial intelligence (AI). An example of such a technology is ChatGPT-3.5, which has received a lot of attention recently. ChatGPT-3.5 is a text generator that was developed on large amounts of texts. Currently, there is a debate about negative consequences in education, for example, if students let the chatbot write texts for them. In this study, however, our point of departure is on how ChatGPT-3.5 and storytelling can be used as a possible tool in teacher education to develop students' critical thinking.

ChatGPT-3.5 has received a lot of attention since its launch in the fall of 2022 and the number of users exploded, reaching one million in the first week. Using artificial intelligence, the chatbot can answer questions and conduct advanced reasoning. ChatGPT can be used to generate summaries of texts but also things like creating poems or stories. The AI tool has created a great debate among journalists, writers, artists but also among schoolteachers because of its limitless capacity for cheating. Others believe that chatbots such as ChatGPT-3.5 also offer many opportunities for teachers and that it is therefore important to learn both its limitations and its possible uses (Ahlgren et al., 2023; Mhlanga, 2023).

However, ChatGPT-3.5 has learned everything it 'knows' by 'reading' books, articles, websites, research papers, user-generated content, and many other publicly available written texts. This means that it can be said to reproduce values and bias when it creates texts. Chatbots are a relatively novel technology, and not much research has examined the messages in chatbot-generated texts aimed at children. The first children's book created by AI, *Alice and Sparkle* (Reshi, 2023), was published in January 2023, and the first Swedish language AI children's book, *Trisse Traktor* (Fernholm, 2023), was released in March.

We have not found any research on bias in ChatGPT-generated texts aimed at children or children's literature written by ChatGPT. Therefore, we asked ChatGPT-3.5 to generate 'technology stories' for us in order to compare them with the findings in Axell and Boström (2021).

This study is to be considered as a pre-study to having teacher students analyse 'technology stories' they create by using ChatGPT. Our starting point is that, for their coming role as technology teachers, teacher students need to develop skills to critically review a text created by chatbots. In this pre-study, we are the ones analysing the technology content in stories generated by ChatGPT-3.5.

2. AIM AND RESEARCH QUESTIONS

Our aim is to investigate how ChatGPT-3.5 frames technology when creating children's stories with a technology focus. More specifically, we aim to answer the following questions:

- What are the representations of technology in stories generated by ChatGPT-3.5?
- What views of technology are presented in the stories generated by ChatGPT-3.5?

3. LITERATURE REVIEW

3.1. *What is ChatGPT-3.5?*

ChatGPT-3.5 is an artificial intelligence (AI) chatbot released in 2022 by OpenAI. The name "ChatGPT" derives from "Chat", which refers to its chatbot functionality, and "GPT", which stands for "Generative Pre-trained Transformer". 3.5 is the version number of the ChatGPT. ChatGPT-3.5 is trained on data up to June 2021. Which is described by OpenAI as being able to generate natural language tasks (such as creative writing, summarisations and conversation) and code through instructions (i.e., prompts) (OpenAI, 2023 5th of April).

3.2. *AI, ChatGPT and bias*

Artificial intelligence (AI) systems have raised concerns about bias and discrimination due to how hard it is to see 'into' the system and their reliance on historical data (Roselli et al., 2019). AI algorithms can perpetuate existing biases present in the data they are trained on. AI algorithms, including machine learning techniques, learn from historical data that may encode biases, making it challenging to eliminate unwanted bias completely (Roselli et al., 2019). This raises concerns about the fairness and equity of AI systems.

While ChatGPT is capable of generating human-like responses, concerns have been raised about biases in its outputs. The platform has produced nonsensical, factually incorrect, and offensive responses, raising questions about its neutrality (Singh & Ramakrishnan, 2023).

For example, political biases have been observed in ChatGPT, with a preference for left-leaning viewpoints identified in political orientation tests (Rozado, 2023). Gender biases have also been found in language models like GPT-3, where feminine characters are associated with stereotypes and portrayed as less powerful than masculine characters (Lucy & Bamman, 2021).

The presence of biases in ChatGPT underscores the need for increased awareness and scrutiny of biased AI systems (Rozado, 2023). Efforts to address bias include incorporating social, ethical, and legal principles into AI systems to enhance their ability to recognise and mitigate discrimination (Ferrer et al., 2021).

Educating students about AI systems and their limitations is crucial for developing critical thinking skills. Students should be encouraged to challenge the outputs of ChatGPT and develop the ability to analyse and discern reliable information (Mhlanga, 2023). Educators play a significant role in fostering an informed and responsible relationship between humans and AI. By teaching students about bias in AI and its limitations, educators can promote ethical and responsible use of AI, empowering students to navigate the evolving technological landscape (Mhlanga, 2023).

To sum up, bias in AI systems, including ChatGPT, is a pressing concern for fairness and equity. Addressing bias requires increased awareness, scrutiny, and the development of detection and mitigation methods. Educating students about the nature of AI, the limitations of systems like ChatGPT, and the skills needed to critically analyse AI-generated content is essential for promoting responsible and ethical use of AI in society.

3.3. How technology is depicted in children's stories written by humans

Axell and Boström (2021) investigated what kind of different portrayals of technology could be found in 180 books aimed at the age groups 1 to 6. Through a thematic analysis they identified three overarching themes: *Autonomous technology; Triumphant technology;* and *Technology as an enabler.*

The theme of *Autonomous technology* revolves around technology acting without human guidance. The books for younger children focused mainly on traditionally masculine-coded technology, with little context or human interaction. The books for the older age group depicted anthropomorphised technology and included human interactions, but the focus on masculine-coded technology persisted.

The second theme, *Triumphant technology*, deals with the historical aspect of technological development. In the books, modern technology was presented as superior to older technology, and there was often a focus on Western advancements. Environmental perspectives were largely absent from the books, but in the cases where they did exist, they primarily focused on the benefits of technological advancements.

Technology as an enabler, the third theme, revolves around technology as a result of human needs and desires. Technology is in this case depicted as relating to various everyday contexts, such as house construction and different systems (e.g., transport systems, energy systems). The theme also deals with vocational roles. The books under investigation primarily showed male characters in traditionally masculine-coded occupations, and thus often reinforced gender stereotypes.

4. METHODOLOGY

The data for this study consists of ten original stories. These stories were generated using the ChatGPT-3.5 language model (specifically text-davinci-002) and resulted in a diverse range of narratives and perspectives for analysis. We prompted ChatGPT-3.5 with the following:

> "Could you write me a story for children that focuses on technology, a minimum of 500 words, please."

After each story was generated, a new chat was generated to create a blank slate for the chatbot to create a completely new story and not be affected by the last one.

The ChatGPT-generated stories were analysed by using a qualitative content analysis to search for patterns. As described by Erlingsson and Brysiewicz (2017) and Hsieh and Shannon (2005), a qualitative content analysis is a repeated and interpretive process in which the meaning of a part can only be understood as related to the context. Based on the research questions, the objective was to identify recurring themes in stories. The analysis was carried out in the following steps:

(i) *Familiarisation:* The first step was to familiarise with the data, which meant a reading and re-reading of the stories generated by ChatGPT.

(ii) *Initial categorisation by ChatGPT:* To initiate the content analysis, ChatGPT was utilised to perform the initial categorisation of themes present in the stories. This process generated a preliminary set of thematic categories.

(iii) *Researchers' stance and refinement:* Following the initial categorisation, we critically engaged with ChatGPT's output. We took an active role in shaping and refining the categorisation; we questioned and challenged the initial categorisation, ensuring that the themes accurately represented the underlying content of the stories. This critical engagement allowed us to establish our stance and exert control over the analysis process.

(iv) *Final categorisation:* With our stance established, a final set of thematic categories was determined. These categories were carefully defined, ensuring clarity and coherence. Any discrepancies or disagreements were discussed and resolved through consensus, and we reviewed and refined the categorisation until a satisfactory level of agreement and consistency was achieved.

(v) *In-depth analysis:* After finalising the thematic categories, an in-depth analysis of each category was conducted, which involved a comprehensive review of the stories, extracting relevant excerpts and examples that exemplified the identified themes.

(vi) *Comparative analysis:* In this stage, we compared the findings from the qualitative content analysis with the findings in Axell and Boström (2021).

(vii) *Integration and synthesis:* The final step involved integrating our analysis and insights with the initial categorisation provided by ChatGPT. By combining the automated categorisation with the critical perspectives and our analyses, a more comprehensive and nuanced understanding of the themes and their implications was achieved.

(viii)   To ensure transparency and credibility (validity) of the study, consistent documentation and systematic analysis were maintained throughout the process.

5. RESULTS

The analysis of the stories generated by ChatGPT-3.5 revealed several prominent themes and sub-themes. The following section presents the results organised by each theme and its corresponding sub-themes.

5.1. Theme 1: Technology is high-tech/A.I.

Sub-themes: Robots, Self-driving cars, Computers and software

The sub-themes within this theme, including robots, self-driving cars, and computers and software, highlight the transformative potential and advanced nature of technology. These narratives illustrate the capabilities of artificial intelligence and high-tech systems, showcasing their efficiency and automation. However, they do not problematise the complex relationship between humans and machines. There is a lack of nuanced picture of power dynamics, biases, and ethical dilemmas.

5.2. Theme 2: Technology is connected to sustainability

Sub-themes: Wind turbines, Solar panels, Recycling

The sub-themes under this theme emphasise the positive relationship between technology and sustainability. These narratives highlight the potential of renewable energy and eco-friendly practices to address environmental challenges. However, the stories do not convey the unequal distribution of sustainable technologies and a need for systemic change and collective action.

5.3. Theme 3: Science fiction technology/gadgetry

Sub-themes: Teleporter, Make-a-wish 3D-printer, Time machine

The sub-themes revolve around imaginative and futuristic concepts. These narratives capture the human fascination with innovative technologies and the potential they hold for transforming our lives. While they inspire awe and curiosity, they often overlook the potential negative consequences and unintended effects that arise from their implementation, such as energy and material consumption, and ethical dilemmas.

5.4. Theme 4: Technology makes the world a better place

Sub-themes: Sustainable development, Creating economic growth, Safety and health, Historical Advancements

The sub-themes within this theme, including sustainable development, creating economic growth, safety and health, historical advancements present technology as a force for positive change and a techno-optimistic perspective. The stories do not show the unequal distribution of technological benefits and the potential for technology to exacerbate existing social inequalities.

5.5. Theme 5: Technology replaces humans

Sub-themes: As a cultural producer, As a friend, As a problem-solver

The sub-themes under this theme highlight the evolving relationship between humans and technology. These narratives explore the potential of technology to automate tasks, offer companionship, and solve complex problems. However, implications for employment, social connections, and human agency in a world increasingly shaped by automation and artificial intelligence is not highlighted.

5.6. Theme 6: Technology as a career enabler

Sub-themes: Knowledge of coding, Becoming a captain of industry, With a little technology and a lot of hard work

The sub-themes within this theme, including knowledge of coding, becoming a captain of industry, with a little technology and a lot of hard work, shed light on the role of technology in career paths. Though, the stories oversimplify the challenges and complexities of navigating technological career paths, downplaying the structural barriers (such as male dominance in the high-tech industry) and inequalities that individuals face.

5.7. Theme 7: Technology as a fountain of joy

Sub-themes: Going on adventures, Creating a friend for oneself, Winning competitions

The sub-themes within this theme highlight the positive and joyful experiences associated with technology. The stories overlook the potential social, psychological, and ethical challenges that arise from excessive reliance on technology for fulfillment and happiness.

6. DISCUSSION

In this study, we explored how ChatGPT-3.5 frames technology when creating stories for children with a technology focus. By analysing the generated stories, we identified several prominent themes which to some extent can be connected to the three themes found by Axell and Boström (2021).

6.1. Autonomous technology

One of the themes that emerged from the analysis is the portrayal of technology as high-tech/AI. The stories generated by ChatGPT-3.5 prominently featured robots, self-driving cars, and computers and software. These narratives depict the transformative potential and advanced nature of technology, emphasising its efficiency and automation.

Additionally, the theme of technology replacing humans emerged from the analysis. These narratives explore the potential of technology to automate tasks, offer companionship, and solve complex problems. However, the stories may oversimplify the complexities of human-machine interactions and fail to address the implications for employment, social connections, and human agency. This relates for example to Ellul's (2010) view of technology, i.e., technological advancements are proceeding at such a rapid pace that humans must automatically adapt and accept these changes.

6.2. Triumphant technology

Another theme that emerged is the connection between technology and sustainability. These narratives showcase a positive relationship between technology and sustainability. However, as Robinson (2004) notes, sustainability should not be reduced to technological fixes but rather requires broader systemic changes. It is important to consider the socio-political dimensions and potential trade-offs involved in the deployment of sustainable technologies. The theme of technology making the world a better place showcases a view that as long as we develop new technology everything is going to be alright. As Winner (1986) points out, society often becomes entranced by the possibilities offered by technology, overlooking the broader implications and power dynamics embedded within it.

The theme of science fiction technology/gadgetry also emerged from the analysis. These narratives reflect the human fascination with innovative technologies and their potential to transform our lives. While the stories generated by ChatGPT-3.5 inspire awe and curiosity, they may overlook the potential negative consequences and ethical dilemmas that arise from the implementation of such technologies.

6.3. Technology as an enabler

Furthermore, the themes of technology as a career enabler and as a fountain of joy emerged. These narratives highlight the role of technology in career development and success. However, the tech industry is heavily gendered. Consequently, as Faulkner (2001) highlights, through portrayals of technology and its associated occupations, gender stereotypes are reinforced. Therefore, it is crucial to consider whether the narratives in the stories perpetuate existing gender biases or challenge them.

7. CONCLUSIONS

A comparison with the results in Axell and Boström (2021) shows that the themes identified in that previous study are similar to the themes in this study. However, some differences were identified, for example that the ChatGPT-generated stories portray an even narrower picture of what technology is. Thus, the problem with these stories is the lack of a broader view of technology, like this quote from one of the stories may exemplify: *"As Timmy grew older, his love for technology only continued to grow. He started to read books and watch videos about how to code and build different devices. He also started to attend robotics and coding classes in his school."*

In summary, this pre-study highlights the importance of critical thinking when using AI generated texts in technology teacher education. The next step is to initiate a study where technology teacher students generate their own stories, using ChatGPT-3.5, and then let them analyse the stories by using the results from Axell and Boström (2021) and the present study as a framework. Based on the students' analyses, we will explore if and how Chat-GPT-generated 'technology stories' could support students critical thinking in relation to technology.

8. REFERENCES

Ahlgren, F., Bjarby, A., & Leitet, J. (Hosts). (2023, 21st June). Bokrelease med Johan Falk. [Podd radio]. In *Teknik i akademi.* https://open.spotify.com/episode/0Z3iiZEZPy0NEhx5xWe4YP si=5pHejwKuQIijo2WGDGNQjA

Axell, C. (2015). *Barnlitteraturens tekniklandskap: En didaktisk vandring från Nils Holgersson till Pettson och Findus* [Technology landscapes in children's literature. A didactic journey from Nils Holgersson to Pettson and Findus.] (Doctoral thesis). Department of Social and Welfare Studies. Linköping university. http://liu.diva-portal.org/smash/get/diva2:784127/FULLTEXT02.pdf

Axell, C. (2017). Critiquing literature: Children's literature as a learning tool for critical awareness. In P. J. Williams & K. Stables (Eds.) *Critique in Design and Technology Education* (pp. 237–254). Springer Nature.

Axell, C., & Boström, J. (2021). Technology in children's picture books as an agent for reinforcing or challenging traditional gender stereotypes. *International Journal of Technology and Design Education*, *31*(1), 27–39. https://doi.org/10.1007/s10798-019-09537-1.

Ellul, J. (2010). The autonomy of technology. In C. Hanks (Ed.) *Technology and values: Essential readings* (pp. 67-75). Wiley-Blackwell.

Erlingsson, C., & Brysiewicz, P. (2017). A hands-on guide to doing content analysis. *African Journal of Emergency Medicine, 7*(3), 93–99. https://doi.org/10.1016/j.afjem.2017.08.001

Faulkner, W. (2001). The technology question in feminism: A view from feminist technology studies. *Women's Studies International Forum, 24*(1), 79-95. https://doi.org/10.1016/S0277-5395(00)00166-7

Ferrer, X., van Nuenen, T., Such, J. M., Coté, M., & Criado, N. (2021). Bias and discrimination in AI: A cross-disciplinary perspective. *IEEE Technology and Society Magazine*, *40*(2), 72-80. https://doi.org/10.48550/arXiv.2008.07309

Fernholm, N. (2023). *Trisse Traktor.* Nicholas Fernholm AB.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277-1288. DOI: 10.1177/1049732305276687

Lucy, L., & Bamman, D. (2021). Gender and representation bias in GPT-3 generated stories. In *Proceedings of the Third Workshop on Narrative Understanding* (pp. 48-55).

Mhlanga, D. (2023). Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning. *Education, the Responsible and Ethical Use of ChatGPT Towards Lifelong Learning.* http://dx.doi.org/10.2139/ssrn.4354422

OpenAI (2023, April 5). *Introduction.* https://platform.openai.com/docs/introduction

Reshi, A. (2023). *Alice and Sparkle.*

Robinson, J. (2004). Squaring the circle? Some thoughts on the idea of sustainable development. *Ecological Economics*, *48*(4), 369-384. https://doi.org/10.1016/j.ecolecon.2003.10.017

Roselli, D., Matthews, J., & Talagala, N. (2019). Managing bias in AI. In *Companion Proceedings of the 2019 World Wide Web Conference* (pp. 539-544). https://doi.org/10.1145/3308560.3317590

Rozado, D. (2023). The political biases of chatgpt. *Social Sciences*, *12*(3),148. https://doi.org/10.3390/socsci12030148

Singh, S., & Ramakrishnan, N. (2023). Is ChatGPT biased? A review. https://doi.org/10.31219/osf.io/9xkbu

Skolinspektionen. (2014). *Teknik – gör det osynliga synligt: Om kvaliteten i grundskolans teknikundervisning.* Skolinspektionens rapport 204:04. https://www.skolinspektionen.se/globalassets/02-beslut-rapporter-stat/granskningsrapporter/tkg/2014/teknik/kvalgr-teknik-slutrapport.pdf

Skolverket. (2022). *Curriculum for the compulsory school, preschool class and school-age educare*. https://www.skolverket.se/download/18.31c292d516e7445866a218f/1576654682907/pdf3984.pdf

Svensson, M., Dahlbäck, K., & von Otter, A-M. (2019). När sagans problem utmanar och inspirerar – En studie av förskolebarns arbete med tekniska lösningar. *Nordic Studies in Science Education, 15*(1), 79–96. https://doi.org/10.5617/nordina.3948

Winner, L. (1986). *The whale and the reactor: A search of limits in an age of high technology*. The University of Chicago Press.