Reichert, Tim, Miftari, Mergim, Herling, Claudia & Marsden, Nicola. (2024). Empowering Female Founders with AI and Play: Integration of a Large Language Model into a Serious Game with Player-Generated Content. HCI International, LNCS 14731, Part II, 69-83. doi:10.1007/978-3-031-60695-3_5

Empowering Female Founders with AI and Play: Integration of a Large Language Model into a Serious Game with Player-Generated Content

Tim Reichert(✉), Mergim Miftari, Claudia Herling, and Nicola Marsden

Heilbronn University, Heilbronn, Germany
{tim.reichert,mergim.miftari,claudia.herling, nicola.marsden}@hs-heilbronn.de

Abstract. This paper presents a novel approach to empower female founders by integrating a Large Language Model (LLM) into a serious game featuring player-generated content. The game leverages interactive visual novels to enhance resilience and awareness of gender-based discrimination in startup environments. The implementation of generative AI for bias detection in the conversations encountered by founders enables the provision of insightful feedback on discriminatory elements within the visual novels in which players immerse themselves. The game allows players to create personalized visual novels, reflecting their own entrepreneurial experiences. The system selects backgrounds, characters, and dialogue options, transforming players into creators while identifying and providing feedback on discriminatory elements within stories. The game's architecture, designed for modularity, supports the integration of various LLMs, enhancing its adaptability and future-proofing. Preliminary results demonstrate the game's feasibility, with expert evaluations affirming the effectiveness of its bias and discrimination detection mechanisms. This research contributes to the fields of gender equality and entrepreneurship in the digital domain, highlighting the potential of serious games and AI in addressing social challenges.

Keywords: Serious Games · Large Language Models · Visual Novel · Gender Equality · Female Entrepreneurship

1 Introduction

Entrepreneurship is essential for economic growth [1], but women are still underrepresented in startups. Societal norms are slowing down women's startup activity [2] and female founders face unique challenges including discrimination based on gender [3]. Previous work in our project [4] identified interactive visual novels as a tool for empowering founders through role playing difficult situations during the startup process.

X. Fang (Ed.): HCII 2024, LNCS 14731, pp. 69–83, 2024.
https://doi.org/10.1007/978-3-031-60695-3_5

In this paper, we present the current architecture and prototype of a serious game with interactive visual novels [5], which has been developed to empower female entrepreneurs by bolstering their resilience [6] in the face of typical founding process adversities and fostering awareness of discriminatory situations and biases specific to female entrepreneurship.

Our ongoing research is conducted in cooperation with the National Agency for Women Startup Activities and Services in Germany [7] and the development of the serious game is guided by an advisory board of 18 experts in female entrepreneurship [8]. It should be noted that the visual novels and the serious game we discuss are being developed in German and the prompts for the LLMs are also written in German. For the sake of discussion, we translated some of the examples given in this paper to English.

2 Approach and Overview

We have currently finished a first prototype of our serious game specifically designed for female entrepreneurs. The upcoming sections shed a light on the gender-specific challenges that women founders face, informing the objectives and design of our game. We provide insight into the game design and the visual elements of our current prototype. The architecture of the game is elaborated upon, with a particular emphasis on the integration and application of LLMs.

2.1 Considerations and Goals

Loan negotiations, business registration, and interactions with family, employees, or co-founders are typical scenarios faced by founders. While they are challenging for everybody on their journey to entrepreneurship, there are additional challenges that women face based on their gender [9–13]. We aim to provide a game experience in which players are able to play through typical scenarios in the founding process and experiment with different choices when confronted with gender biases or discriminatory situations. These may include inappropriate questions on family planning as well as implicitly or explicitly discouraging or insulting comments. Our prototype contains a set of eleven pre-constructed novels covering the typical scenarios faced by founders. They were developed within the project with experts on women and entrepreneurship [4]. They allow players to choose responses from a set of possible answers. Based on events in the story and choices taken by players, the game should provides tailored feedback at the end of a play session that puts the specific situation in context of known biases, helps players to contextualize choices taken and encourages experimentation by trying out different choices.

Previous research in game design of visual novels has shown that real-life scenarios and being able to share own experiences enhances reflection [14] more successfully than fictional narratives [15]. It also highlights the importance of experiential fidelity to facilitate transfer [16] and to help players relate the game to their own experiences [17].

Thus, our game aims to provide not only pre-constructed visual novels by the developers, but also to enable players to transform their personal experiences into playable stories. For this approach to be viable, the process of generating or altering visual novels needs to be sufficiently user-friendly to enable individuals lacking specialized technical expertise to perform these tasks effectively. Our ultimate goal is to be able to create interactive visual novels based on a natural language description of real-world experiences and events.

Since every founder has a different background and training needs, different visual novels in our game may be more or less relevant. Thus, the game requires a guidance system - something we call "Game Director" - that matches players with relevant content based on information they provide to the game.

2.2 Game Design and Visual Elements

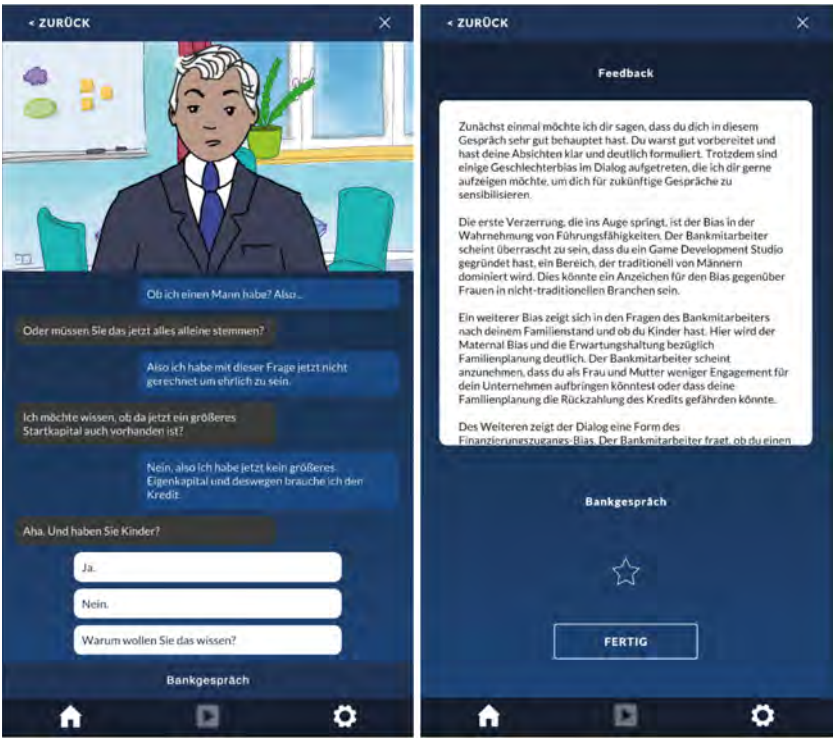

Fig. 1. Prototype screenshots of visual novel and feedback view

Figure 1 (left) shows a screenshot of a visual novel in our current prototype. The visual elements are the representation of non-player characters (NPCs), backgrounds and dialogue between player and NPCs. NPCs have different facial

expressions that may change during the progression of the story. Dialogue is conveyed in a format akin to text messaging applications, which offers the benefit of a scrollable history. Players have dialogue choices that may affect the story. For example, they may choose to react to an inappropriate comment by pointing out the wrongdoing or by ignoring it. After playing a session, feedback is provided to players by giving them a) information on which biases were present in the story and b) an analysis of how they reacted to them, as shown in Fig. 1 (right).

2.3 Architecture and LLM Use

Large Language Models (LLMs) like GPT 4 are advanced AI systems capable of processing natural language and generating human-like text [18,19]. In our research, we use LLMs for novel creation, discrimination detection, feedback for players, and content matching players with content. Figure 2 gives an overview of the current architecture of the game with AI-based components implementing the aforementioned features shown in grey.

The central concept is the interactive **visual novel**. A visual novels tells a story in an interactive way. Its creation can be subdivided into three editor components: **character maker**, **environment maker** and **dialogue maker**. While these components can be used manually in a more traditional development approach, they provide an interface to be utilized by generative AI to generate content from natural language descriptions, as shown on the top left side of Fig. 2. The dialogue graph of the resulting visual novel can be analyzed by the LLM for potential biases so that feedback is available throughout the creation process.

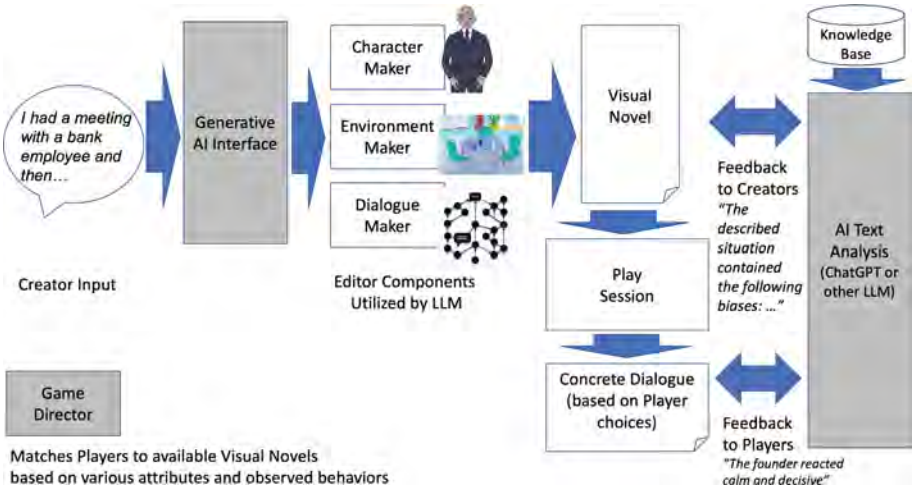

Fig. 2. Architecture overview

In a play session, players have choices that may alter the course of the story, resulting in a log of **concrete dialogue**, i.e. a path through the dialogue graph. Based on this log and a domain specific knowledge base the AI creates feedback for the players. This will be explained in detail in Sect. 3. Finally, the **game director** component analyzes the player's individual situation and training need based on a short entrance interview to match players with appropriate content.

Due to the novelty and rapidly changing capabilities of LLMs, we aim to be independent of a concrete LLM and to be able to replace the GPT 4 model we are currently using with other cloud-based or local LLM-implementations. Thus, we have designed the interfaces in a modular way by abstracting the component that performs the API calls to query the model. The actual natural language query content can thus be understood by different LLMs due to the natural language processing capabilities of these models. The way that the model is queried is described in Sects. 3, 4, and 5.

Our current prototype fully implements playing and analysis of visual novels including the feedback component for which an evaluation is provided in Sect. 6. The generative AI interface is working and can create novels based on natural language input, but is considered experimental with encouraging initial results, see Sect. 5. The game director component is currently being developed with promising initial testing done with GPT 4, see Sect. 4.

To provide a low barrier of entry, the game is created as a mobile application that we aim to distribute through the App Store and Google Play.

3 Bias and Discrimination Detection

Biases and discrimination may be implicit and hard to spot in everyday situations [20]. Nuances in human language and context specificity in discriminatory practices present considerable challenges for recognition even for humans [21]. Advanced AI models can recognize patterns that might suggest bias or discrimination, but their accuracy is dependent on the data they were trained on and the specific instructions and context provided [22,23]. Keeping this in mind, our experience is that GPT 3.5 and GPT 4 are able to spot certain forms of discrimination and identify biases. For example, one of our novels features an employee of a funding information center asking a female founder during negotiations if she plans to have children. When providing the log of such a dialogue to GPT 4, the model will identify the question as potentially inappropriate and discriminatory in the business context.

By providing a domain-specific knowledge base to the LLM we aim for precise feedback that refers to known biases. We also aim to have to rely less on the discrimination-specific dataset a specific model was trained on so that working with different LLMs is possible. The knowledge base we have developed for this purpose covers 25 biases derived from research on female entrepreneurship and gender studies [9–13], giving the name and a short description of e.g. maternal bias, attributional biases regarding performance, biases regarding risk aversion in women, prove-it-again bias.

After playing a visual novel, the session log is analyzed with a focus on problematic situations encountered and players reactions to them so that useful feedback can be provided to players. To utilize an LLM in our game in this way, it is necessary to prepare a prompt for the model and process the text-based result. Our prompt engineering effort has resulted in prompts that are structured into five parts: Role, Assignment, Context, Knowledge Base, Output Format and Target. Table 1 shows an abbreviated example for the visual novel "Talking to my parents" in which the founder experiences biases.

Table 1. Structure of the prompt for discrimination detection and feedback

Section	Example
Role	"You are a gender researcher."
Assignment	"Your assignment is to analyze the log for discrimination and gender biases and give feedback to the player..."
Context	"The founder plans to start a software company and talks to her concerned mother about it."
Knowledge Base	Consisting of a description of 25 biases that women founders are confronted with.
Output Format	"Write an analysis of the dialog..."
Target	This entails the log of dialogue (Mother: "Hello and thank you for coming" Player: "Hi"...)

To ensure that the suggestions of the LLM are appropriate and helpful for players, we have integrated a player-feedback loop into the serious game. After receiving suggestions from the AI system, players can rate the satisfaction with the feedback on a 5-point Likert scale and can comment in a free text field whether the information provided was helpful. Players can also request an expert opinion, in case they have question regarding the AI feedback. As described above, we are developing the game together with an advisory board of experts in the field of female entrepreneurship and gender. The board members are using the app in an "expert mode" and have the ability to provide expert feedback. Feedback of players and experts is stored together with the session log and AI-feedback in a data base. An analysis of initial expert evaluation based on the GPT 4 model and the prompting as described above can be found in Sect. 6.

4 Towards Matching Players with Content

In the current prototype, players manually select novels from a list. We are currently working on a component that connects players with the most suitable built-in or player-generated content automatically. The idea is, that in the on-boarding phase, players engage with an initial visual novel that derives insights such as their startup's field, personal context, perceived strength and weaknesses

and challenges they expect in their startup journey. The system then aims to connect players with the most suitable novels based on their provided data. For each novel, the LLM initially creates a title and a short summary that is stored with the novel. To perform the matching a prompt is created consisting of the title and summary of each novel, the player data and the assignment to rank the novels by suitability for a training situation.

Initial testing revealed that GPT 4 is able to rank the novels in a sensible way. For example, in one of our test cases we asked GPT 4 to rank the following novels for a training course based on role playing depending on how well the course content can support training of a founder that has very supportive parents, is good at negotiations and has problems to give negative feedback to people:

- Parental Discussion: The parents doubt the entrepreneurial capabilities.
- Press Interview: The press wants to focus on gender rather than the startup.
- Credit Appointment: The bank employee asks about the desire for children.
- Colleague Conversation: A problematic situation in the founding team is discussed.
- Employee Discussion: An employee has performed poorly. This needs to be addressed.

The following ranking was provided by the model:

1. Employee Discussion
2. Colleague Conversation
3. Press Talk
4. Credit Appointment
5. Parental Discussion

While this data point is not conclusive on its own, it showcases the potential for using advanced AI models to rank training content. In future work, we plan to fully implement the system based on the concepts described in the next chapter and to perform an expert evaluation akin to that in Sect. 6.

5 LLM-Based Visual Novel Creation Process

In this section, we outline how visual novels are represented and edited manually, as well as how we utilize LLMs in their creation. We detail AI's role in the selection of backgrounds and characters, and in generating dialogue. We describe the AI-assisted process and finally present an example of a visual novel produced by our tool, illustrating its capabilities in narrative crafting and dialogue branching.

5.1 Basic Concepts

Creating visual novels comprises creating both the visual representations and the content of the conversation. Visual novels are based on a **directed graph**

of conversation nodes where each node may nest another such graph. Connections between nodes represent dialogue options while each node defines a dialogue message and determines background, visible characters, and their facial expressions and gestures. Figure 3 (left) shows the Dialogue Maker's user interface, featuring a board where users freely position nodes to form a directed graph.

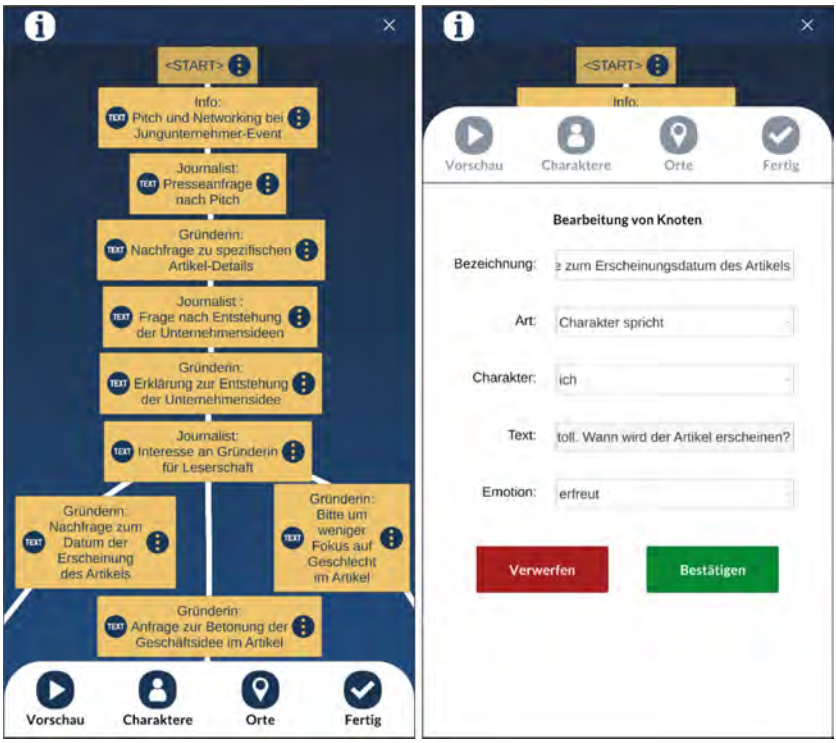

Fig. 3. Creating a visual novel and editing a conversation node

Figure 3 (right) shows the editing process of a conversation node, where users can assign backgrounds, characters, dialogue text, and emotions specific to the situation. Nodes can be named for clarity. The Dialogue Maker includes a preview function for testing the visual novel.

While the tool can be used manually, the next section describes an approach that uses an LLM to convert player narratives from text or voice input into editable, shareable visual novels. Based on the user's input, the AI selects backgrounds and characters for the visual novel and creates dialogues with options.

5.2 Visual Novel Generation Process

Figure 4 outlines the novel generation process, which begins with a natural language description of the visual novel, either through text or voice. To guide users, three questions are provided:

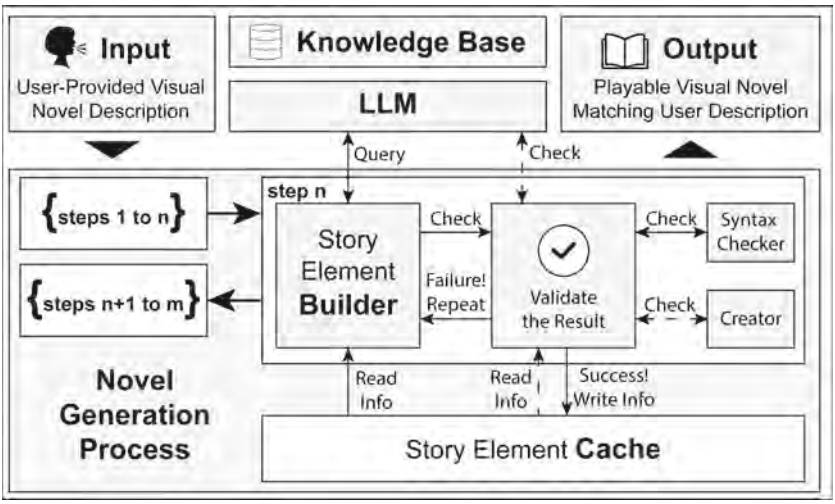

Fig. 4. Structure of the visual novel generation process

1. What are the main events or key moments of your story?
2. Who are the main characters and what roles do they play in your story?
3. What are the important decisions and turning points in your story?

User responses are processed into a complete, editable, and shareable visual novel. The LLM effectively handles colloquial and incomplete sentences. As depicted in Fig. 4, the generation process involves several steps. These steps include character and location determination, dialogue creation, and facial expression assignment for dialogue messages. Story elements are built using an LLM and our domain-specific knowledge base, leveraging a shared cache for information from previous steps. Each story elements is obtained through a response from the LLM and undergoes a validation process. If validation fails, the element can be recreated without disrupting the overall process. Failures are logged for future improvements. The basic form of validation is the syntax checker which ensures output conforms to specifications like JSON format and element count. The system also verifies consistency with existing story elements in the cache, particularly for character consistency in dialogue messages. User-led validation and LLM self-validation are enhancements we are currently experimenting with.

For tasks in which the LLM has to make a choice between different options, it is provided with a description of each option and asked to select the best fit for

the situation. This includes the set of available characters, e.g. “casually dressed middle-aged man”, “business person in suit”, or “young female co-founder” as well as the properties of available backgrounds, e.g. “generic meeting room”, “living room”, “in front of a computer at night”, or “on stage”.

Currently, the generation process of a complete novel with 60 nodes using GPT 4 involves roughly 50 LLM requests and takes 12 min. The results are generally usable, but the dialogue sometimes sounds too generic, has inconsistencies, or requires further editing. By improving the prompting and enabling user participation in the validation process we aim to improve the quality and reduce the number of requests due to failed validations in the future.

5.3 Results of the Visual Novel Generation

The following example is from a visual novel generated through the process described above, where the input describes a founder discussing her startup project with her mother. The mother expresses concerns about the founder’s ability to start a family while running a business and suggests collaboration with her brother. The LLM chose a private home as the background, one of seven possible locations provided. An excerpt from the dialogue illustrates the coherent linguistic quality:

Info: Your mother looks at you with concern. You sense her concern but also her interest.
Founder: I understand your worries, mom. But I firmly believe that I can do this.
Mother: And what if you’re thinking of starting a family? Running a business is a full-time job.
Founder: I’m aware of that, mom. But many women successfully run businesses and have a family at the same time. I can do that too.

Initially, the generated dialogue is linear. The next step involves integrating different dialogue options at various points, either by interrupting and regenerating the dialogue or by inserting gaps to be filled. Different moods can be specified for the LLM, leading to varied answer options:

Mother: Can I help you with this?
Answer opt. 1: I think I can manage on my own. But thanks, Mom.
Answer opt. 2: Sure, Mom. You could help me bake some of your famous chocolate chip cookies. I’ll need the energy for the next steps.
Answer opt. 3: Mom, I feel like you’re trying to interfere with my startup process. I need to do this on my own.
Answer opt. 4: Um... I’m not sure, mom. I mean, I really appreciate your offer, but I don’t know exactly how you could help me.

Finally, the process determines a facial expression for each dialogue option. For instance, the mother reacts with laughter to answer option 2 and with sadness to option 3. The facial expressions are visually reflected when playing the novel through character animations.

6 Evaluation of LLM-Based Feedback

The bias and discrimination detection outlined in Sect. 3 is already implemented in the current prototype of the game: Players receive an analysis of the dialogue as feedback on the novels that they played. To assess the effectiveness of this feedback mechanism, we conducted an expert evaluation. This section outlines the approach employed in the evaluation process and presents an analysis of the findings, encompassing both quantitative and qualitative data. The results offer insights into the feedback's impact on player experience and its overall quality within the game environment.

6.1 Approach

In the evaluation of the discrimination detection integrated in the serious game, the following approach was taken

Participants: Four experts from our project team with knowledge about gender and entrepreneurship and familiarity with issues related to gender bias and discrimination evaluated the app's analysis of the dialogues.

Materials: The study centered around six dialogues, crafted to represent two distinct scenarios. Each scenario was presented in three versions, varying in complexity and context. The dialogues were specifically designed to encompass a range of potential biases, allowing a comprehensive assessment of the app's detection capabilities.

Procedure: The experts were asked to review the dialogues. Each dialogue was analyzed by the app's discrimination detection algorithm employing GPT 4, which then generated a report outlining potential biases. The experts' task was to evaluate these reports in terms of clarity, style, content, and accuracy.

Measures and Data Analysis: The experts provided feedback on 10 items using a five-point Likert scale. The items were developed to capture various aspects of the app's performance, including the accuracy of bias detection, clarity of the reports, relevance of the content, and the overall style of presentation. Descriptive statistics were used to provide an overview. Alongside the Likert-scale items, the experts answered 11 open-ended questions. These questions were designed to gather detailed qualitative feedback, allowing the experts to elaborate on specific aspects of the LLM responses, share observations, and provide suggestions for improvement. The responses to the open-ended questions were subjected to thematic analysis [24].

6.2 Results

The expert evaluations of the app's discrimination detection provided a mix of quantitative and qualitative insights. Quantitative analysis revealed high ratings across various aspects, including clarity, style, content, and accuracy, with scores

ranging from 4.73 to 4.95 on a five-point Likert scale with 1 expressing disagreement und 5 agreement. There were no significant differences among the different items assessed. The detailed results for each item can be found in Table 2.

Table 2. Expert evaluation of LLM feedback

Item	Mean	StDev
The language style and structure are appropriate	4.95	0.21
The feedback is thematically relevant	4.95	0.21
The feedback is clear and understandable	4.92	0.28
The feedback is free from incorrect information	4.91	0.29
The feedback covers all essential aspects of the dialogue	4.83	0.38
The tone of the feedback is appreciative and supportive	4.82	0.59
The biases referenced in the feedback are understandable in the context of the conversation	4.75	0.61
The feedback provides a balanced perspective	4.75	0.61
The feedback is free from unnecessary information	4.74	0.75
The feedback is free from prejudices or opinions	4.73	0.63

The qualitative analysis was based on 63 answers to the open ended questions. The thematic analysis of this qualitative feedback resonated with the positive ratings from the quantitative analysis with the experts highlighting the feedback clarity, the friendly and supportive tone, and the high quality of bias identification and interpretation. There were some points that were mentioned for improvement: The experts' opinion varied on the extent of empowering language, with some finding it excessive. Concerns were raised regarding the necessity to mention certain biases - while they were not seen as wrongly applied, some experts felt they did not add to the analysis.

In conclusion, the combined results of quantitative and qualitative analyses affirm the effectiveness of the discrimination detection app in providing feedback that is clear, relevant, and suitably toned. The app received high ratings on the Likert-scale items, and the responses to open-ended questions were predominantly positive. Opportunities for refinement are seen regarding feedback relevance and ensuring not do overdo it on the empowerment messages. The improvement possibilities regarding bias relevance are attributed to the fact that the LLM is "obedient" in the sense that the LLM tries to incorporate as many of the biases as possible from the given prompt, regardless of their exact fit to the specific dialogue under review.

7 Conclusions and Future Work

In conclusion, our research presents the integration of an LLM into a serious game for female founders. The game, leveraging interactive visual novels and an

LLM-based analysis of the conversations, has demonstrated potential in enhancing resilience and awareness of gender-based discrimination in entrepreneurial environments. Our findings indicate the effectiveness of generative AI in creating engaging content and detecting biases, providing meaningful feedback to players. This aligns with our objective to transform players into creators, allowing them to reflect on their experiences through interactive storytelling.

The next step is to systematically evaluate the analyses offered by the LLM with women entrepreneurs as the prime user group. In terms of relating to real-life scenarios, special attention needs to be paid to finding out which elements of the game are able to enhance reflection [14]. Both the role of the content and the role of high- versus low-fidelity needs to be considered. In terms of content, we expect the matching of players with stories and the possibility of creating one's own novel to facilitate transfer [16] and to help players relate the game to their own experiences [17]. In terms of presenting high- versus low-fidelity scenarios via AI, findings from the hyperpersonal model of communication will be integrated [25]. The hyperpersonal model shows that under certain circumstances, low-fidelity communication can actually be helpful to create a setting in which it is easier for a person to open up and engage deeply.

In integrating LLMs into our serious game, we must critically address several ethical implications. The danger of bias in the training materials as well as hallucinations and therefore inadequate answers in LLMs requires vigilant monitoring and continuous refinement. Protecting player-generated data demands strict adherence to privacy standards and explicit user consent protocols. Given the game's focus on gender bias and discrimination, we must be mindful of its emotional impact on players, providing support and a safe environment. Ensuring inclusivity and diverse representation in the game's content is crucial to avoid reinforcing stereotypes and exclusion. Finally, our responsibility extends to considering the long-term societal impact of the game, aiming to contribute positively to social change and gender equality in entrepreneurship.

Future work will focus on refining the AI components and integrating alternative LLMs. While automated novel creation shows promise and works in general, it is currently still experimental and needs more work on quality assurance and optimization before it can be released to end-users. We hope that through testing with the target audience we can find a suitable balance between user control and automation. We plan to expand the game's content and interactivity, enabling a wider range of scenarios and more personalized experiences. Ultimately, our goal is to contribute to the field of gender equality in entrepreneurship, exploring the intersection of AI, serious gaming, and social empowerment.

Acknowledgments. This work has been partially funded by the German Federal Ministry for Family Affairs, Senior Citizens, Women and Youth (Bundesministerium für Familie, Senioren, Frauen und Jugend - BMFSFJ) under Grant Number 3923406K04 as part of the project 'KI-Thinktank Female Entrepreneurship II - KITE II' in the Funding Line 'AI for Public Good (KI für Gemeinwohl)'. The responsibility for all content supplied lies with the authors.

References

1. Stoica, O., Roman, A., Rusu, V.D.: The nexus between entrepreneurship and economic growth: a comparative analysis on groups of countries. Sustainability **12**(3), 1–19 (2020)
2. Sternberg, R., Gorynia-Pfeffer, N., Stolz, L., Schauer, J., Baharian, A., Wallisch, M.: Global entrepreneurship monitor 2021/2022. In: RKW Kompetenzzentrum (2022). https://www.rkw-kompetenzzentrum.de/fileadmin/media/Produkte/2022/Studie/20220527-Studie-GEM-2022.pdf
3. Veckalne, R., Tambovceva, T.: The importance of gender equality in promoting entrepreneurship and innovation. In: Marketing and Management of Innovations, pp. 158–168 (2023). https://doi.org/10.21272/mmi.2023
4. Schirmacher, A., von der Bey, K.: KITE-Thinktank Female Entrepreneurship - Diskriminierungs-Bekämpfung für Gründerinnen durch gezielten Kompetenzaufbau beim Erkennen und Bewältigen von Diskriminierungsmustern - Internal Project Report. In: Project KITE - more info on https://www.kite-bga.de/expertisen/ (2022). (in German)
5. Camingue, J., Carstensdottir, E., Melcer, E.F.: What is a visual novel? In: Proceedings of the ACM on Human-Computer Interaction 5.CHI PLAY, pp. 1–18 (2021). ISSN 2573-0142
6. McGonigal, J.: Reality is Broken: Why Games Make Us Better and How They Can Change the World. Penguin, New York (2011). ISBN 1101475498
7. National Agency for Women Start-Up Activities and Services in Germany. In: Bundesgründerinnenagentur (BGA) (2024). https://innogruenderinnen-bga.de/
8. Advisory Board KITE II. In: Projektbeirat KITE II (2024). https://kite2.de/beirat-kite-ii/
9. Laguía, A., García-Ael, C., Wach, D., Moriano, J.A.: Think entrepreneur-think male": a task and relationship scale to measure gender stereotypes in entrepreneurship. Int. Entrep. Manag. J. **15**, 749–772 (2019). https://doi.org/10.1007/s11365-018-0553-0. ISSN 1554-7191
10. Tonoyan, V., Strohmeyer, R.: Gender role (in-) congruity and resource-provider gender biases: a conceptual model. Int. J. Gend. Entrep. **13**(3), 225–242 (2021). https://doi.org/10.1108/IJGE-12-2020-0201. ISSN 1756-6266
11. Folberg, A., Goering, T., Wetzel, L., Yang, X., Ryan, C.: Viewing entrepreneurship through a goal congruity lens: the roles of dominance and communal goal orientations in women's and men's venture interests. Front. Psychol. **14**(2023). https://doi.org/10.3389/fpsyg.2023.1105550
12. Williams, J.C., Dempsey, R.: What Works for Women at Work: Four Patterns Working Women Need to Know. NYU Press, New York (2018)
13. Mavin, S., Yusupova, M.: Competition and gender: time's up on essentialist knowledge production. Manag. Learn. **52**(1), 86–108 (2021). https://doi.org/10.1177/1350507620950176. ISSN 1350-5076
14. Khaled, R.: Questions over answers: reflective game design. In: Playful Disruption of Digital Media. Gaming Media and Social Effects, pp. 3–27 (2018). https://doi.org/10.1007/978-981-10-1891-6_1. ISSN 9811018898
15. Iacovides, I., Cutting, J., Beeston, J., Cecchinato, M.E., Mekler, E.D., Cairns, P.: Close but not too close: distance and relevance in designing games for reflection. Proc. ACM Hum. Comput. Interact. **6**, 1–24 (2022). https://doi.org/10.1145/3549487

16. Súilleabháin, G.Ó., Sime, J.-A.: Critical design and effective tools for E-learning. In: Donnelly, R., Harvey, J., O'Rourke, K. (ed.) Higher Education: Theory into Practice, pp. 113–126. IGI Global (2010). ISBN 1615208801
17. Hammer, J., To, A., Schrier, K., Bowman, S.L., Kaufman, G.: Learning and role-playing games. In: Deterding, S., Zagal, J. (ed.) Role-Playing Game Studies, pp. 283–299. Routledge (2018). https://doi.org/10.4324/9781315637532
18. Bang, Y., et al.: A multitask, multilingual, multimodal evaluation of ChatGPT on reasoning, hallucination, and interactivity. In: arXiv preprint (2023). https://doi.org/10.48550/arXiv.2302.04023
19. Chang, Y., et al.: A survey on evaluation of large language models. In: arXiv preprint arXiv:2307.03109 (2023). https://doi.org/10.48550/arXiv.2307.03109.
20. Greenwald, A.G., Banaji, M.R.: The implicit revolution: reconceiving the relation between conscious and unconscious. Am. Psychol. **72**(9), 861–871 (2017). https://doi.org/10.1037/amp0000238
21. Greenland, K., West, K., Van Laar, C.: Definitional boundaries of discrimination: tools for deciding what constitutes discrimination (and what doesn't). J. Appl. Soc. Psychol. **52**(10), 945–964 (2022). https://doi.org/10.1111/jasp.12902
22. Gallegos, I.O., et al.: Bias and fairness in large language models: a survey. In: arXiv preprint (2023). https://doi.org/10.48550/arXiv.2309.00770.
23. Mao, R., Tan, L., Moieni, R.: Developing a large-scale language model to unveil and alleviate gender and age biases in Australian job ads. In: 2023 IEEE International Conference on Big Data (BigData), pp. 4176–4185. https://doi.org/10.1109/BigData59044.2023.10386083
24. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qual. Res. Psychol. **3**(2), 77–101 (2006). https://doi.org/10.1191/1478088706qp063oa. ISSN 1478-0887
25. Walther, J.B., Whitty, M.T.: Language, psychology, and new new media: the hyperpersonal model of mediated communication at twenty-five years. J. Lang. Soc. Psychol. **40**(1), 120–135 (2021). https://doi.org/10.1177/0261927X20967703