

# Promoting Trustworthy AI in mHealth: a Gamified Approach to Value-Sensitive Design

Maria Inês Ribeiro<sup>1,\*†</sup>, Laura Genga<sup>2,†</sup>, Monique Simons<sup>3,‡</sup> and Pieter Van Gorp<sup>4,†</sup>

<sup>1</sup>Technical University of Eindhoven Eindhoven, Netherlands

<sup>2</sup>Wageningen University & Research Wageningen, Netherlands

## Abstract

The rise of mobile health (mHealth) apps leveraging AI and wearables to promote healthy lifestyles is accompanied by growing ethical concerns among the public, developers, and policymakers. While AI guidelines exist to mitigate concerns, translating them to practical design requirements remains challenging. This research proposes a gamified approach to help bridge the gap between theory and practice in Value-Sensitive Design (VSD) for AI applications in mHealth. This approach aims to facilitate the development of trustworthy AI by aligning design with stakeholder ethical values. Using the design science methodology, we developed a card game to improve stakeholder participation, foster an understanding of AI in mHealth, and facilitate in-depth ethical discussions. Pilot-testing with 19 peer researchers showed active engagement and motivation of players through self-discovery. The findings highlight the game's potential to elicit ethical discussions and promote an understanding of AI's real-world implications. Future iterations could explore digital, blended, or survey formats to enhance engagement, accessibility, and depth of insights, catering to diverse stakeholder preferences. This gamified approach to VSD holds promise as a tool for supporting the development of trustworthy AI technologies in healthcare, aligned with stakeholder values. Further validation with broader stakeholder groups and a longitudinal impact assessment are needed.

## Keywords

mHealth, Trustworthy AI, Value-Sensitive Design

## 1. Introduction

In recent years, mHealth apps that track our sleep patterns, heart rate, and activity levels have become increasingly popular to promote healthy lifestyle behaviors. While these personalized health tools powered with AI technology and wearables data hold immense potential, recent ethical incidents raise concerns and spark alarm among the public, developers, and policymakers [1, 2].

Ethical frameworks and regulations are emerging to mitigate these concerns and ensure trustworthy AI development. For instance, the High-level expert group on AI from the European

---

*HHAI-WS 2024: Workshops at the Third International Conference on Hybrid Human-Artificial Intelligence (HHAI), June 10–14, 2024, Malmö, Sweden*

\*Corresponding author.

✉ [m.i.da.graca.jorge.da.silva.ribeiro@tue.nl](mailto:m.i.da.graca.jorge.da.silva.ribeiro@tue.nl) (M. I. Ribeiro); [l.genga@tue.nl](mailto:l.genga@tue.nl) (L. Genga); [monique.simons@wur.nl](mailto:monique.simons@wur.nl) (M. Simons); [P.M.E.v.Gorp@tue.nl](mailto:P.M.E.v.Gorp@tue.nl) (P. Van Gorp)

🌐 <https://orcid.org/0009-0001-7746-4685> (M. I. Ribeiro); <https://orcid.org/0000-0001-8746-8826> (L. Genga);

<https://orcid.org/0000-0003-4693-9980> (M. Simons); <https://orcid.org/0000-0001-5197-3986> (P. Van Gorp)

🆔 0009-0001-7746-4685 (M. I. Ribeiro); 0000-0001-8746-8826 (L. Genga); 0000-0003-4693-9980 (M. Simons);

0000-0001-5197-3986 (P. Van Gorp)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



Commission (EC) advocates for a human-centered approach grounded by the ethical principles of respect for human autonomy, prevention against harm, fairness, and explicability [3]. A self-assessment checklist is available as a tool for AI developers to implement these principles [4]. Yet, checklist-based approaches lack practical implementation details, leaving developers to navigate complex ethical dilemmas and tensions between diverse stakeholder ethical values [5]. In mHealth apps, conflicts between data privacy and personalized lifestyle recommendations are particularly evident. For example, an app may collect health and lifestyle data to predict opportunistic moments for suggesting a walk. This app could bring health benefits with increased physical activity but also poses privacy risks as sensitive health data could be exposed. *What is then more valuable: data privacy or health benefits?*

To address these complex trade-offs, several design approaches can inform the integration of stakeholder values in the design process of AI technology. While User-Centered Design prioritizes user experience and Privacy by Design focuses on data protection, VSD offers a more comprehensive framework, analyzing AI ethics across individual, group, and societal levels, and aiming at the symbiotic evolution of technology and societal norms [6, 7, 8].

Multiple methods have been employed to elicit values in VSD, such as the Value Scenario method, which emphasizes technology implications in narrated use cases, or the Value-oriented Mock-up, Prototype, or Field Deployment method, which investigates values implications in real-world contexts [9]. Despite these efforts, current VSD methods face several limitations, often addressing only one or two of these key challenges: (1) recruiting and engaging stakeholders in focus groups; (2) providing enough technical and ethical AI understanding to stakeholders, and (3) eliciting ethical discussions that allow for translating abstract findings into actionable requirements for AI developers [6].

A gamified tool seems intuitively capable of addressing these challenges simultaneously. First, games are inherently engaging, attracting and retaining stakeholder participation better than traditional methods. Second, they may simulate complex scenarios and provide immediate feedback, helping stakeholders grasp AI's technical and ethical dimensions without prior expertise. Third, the structured yet flexible nature of games allows for quantitative tracking of decisions and actions, providing concrete data for actionable design requirements.

This research proposes to support VSD with a gamified approach. We developed and pilot-tested a card game to elicit and explore stakeholder values regarding the use of AI in mHealth apps. This approach seeks to provide practical insights that can enhance the effectiveness of VSD in guiding the development of trustworthy AI technology.

The structure of the remainder of this paper is as follows: Section 2 outlines the methods used to develop and pilot-test the gamified approach; Section 3 presents the key findings from the pilot tests; Section 4 discusses the adherence of the game to its objectives and potential future directions for refining the game; and Section 5 concludes with a summary of key points and suggestions for further research.

## **2. Methods**

In this study, we employed the design science methodology framework to develop and refine a game exploring stakeholder values and ethical considerations in using AI for mHealth apps

[10]. The overall goal was to provide a practical tool to support VSD. This preliminary version of the game was designed for a general population, assessing its acceptance of using private data for generating personalized lifestyle recommendations. This section outlines the game objectives, design, and pilot testing.

## 2.1. Game Objectives

The game aimed to achieve the following objectives:

**Objective 1: Enhance Recruitment and Engagement.** Leverage gamification to create an interactive and captivating experience for stakeholders during focus groups.

**Objective 2: Provide Understanding of AI.** Present concrete examples of AI applications and implications to guide the definition of AI design requirements by assessing ethical concerns on AI-specific uses.

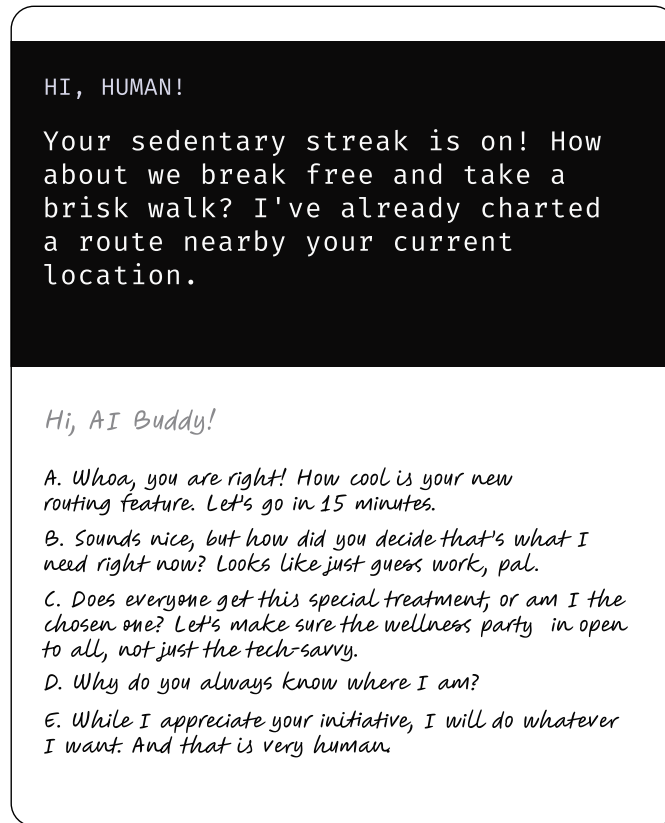
**Objective 3: Elicit In-Depth Ethical Discussions.** Engage stakeholders in structured discussions on AI in mHealth to gain insights on specific ethical considerations relevant to AI design and development.

## 2.2. Game Design

The game design adheres to the Mechanics-Dynamics-Aesthetics framework to create an engaging exploration of AI ethical considerations in mHealth apps with a trade-off between data privacy and personalized lifestyle recommendations [11].

To enhance recruitment and engagement (Objective 1), the game offers intrinsic and extrinsic rewards. At the beginning of a game session, players were motivated to embark on a self-discovery journey fostering reflection on personal ethical values (intrinsic reward) while earning AI user-type badges (extrinsic reward). This AI user type was defined based on the prevalence of each participant's ethical concerns categorized according to the four ethical principles of trustworthy AI defined by the EC [3]. During the game, participants encountered multiple ethical dilemmas that required them to weigh competing values and priorities when interacting with mHealth apps. The game provided a safe and comfortable social environment for open and honest discussions about ethical concerns, promoting community and empathy among players.

Featuring a card game, the core game mechanics revolved around *Black Cards* presenting AI-generated lifestyle recommendations with five possible human reactions (Figure 1), aiming to promote understanding of AI's real-world implications (Objective 2). Such prompts were linked to at least one AI development decision, e.g. *'Is it acceptable to use GPX location to recommend convenient and nearby walking routes?'*. Players individually chose a *White Card*, labeled from A to E, reflecting their preferred reaction to the AI prompt, and placed it facing down on the table. A moderator facilitated discussion in each round as players shared and debated their choices (Objective 3). Encrypted color coding on *Score Keeping Cards* tracks player decisions. Upon game over, players earn the *AI User-Type Badges* reflecting their ethical priorities based on gameplay and revealed by the *Game Over Card*.



**Figure 1:** Black Card sample presenting an AI-generated health recommendation and five human reactions.

### 2.3. Pilot Testing

We conducted two 90-minute focus groups to pilot-test the game. A total of 19 researchers were recruited through convenience sampling at our affiliated universities (Eindhoven University of Technology and Wageningen University & Research). Both sessions follow a similar agenda: introduction explaining the game and its objectives (10 minutes), gameplay (six rounds or 45 minutes), and feedback (35 minutes). Observations during gameplay from both focus groups were used to evaluate the game's adherence to its key objectives. The feedback sessions served to ideate new game mechanics, dynamics, and aesthetics and refine the game design for future iterations. Focus Group 1 engaged at first in spontaneous feedback and then collaborated in a co-creation task using the gamification model canvas to refine the game design [12]. Focus Group 2 participated in a semi-structured discussion with predefined questions to guide the co-creation process.

### 3. Results

We briefly report the most significant findings and related participants' suggestions for game refinement offered in focus groups.

**Finding 1: Motivation through Self-Discovery.** Participants found that uncovering their AI user type was a strong motivator for participation. While some players found the assigned type aligned with their values, others desired more rounds for a clearer picture. One participant recommended using a subset of AI scenarios in digital format as a teaser to recruit players.

**Finding 2: Player Engagement and Relatedness.** Participants expressed joy in gameplay, reporting higher engagement when scenarios resonated with personal experiences. The alignment of AI prompts with personal interests significantly influenced their reactions and investment in the gameplay. Participants suggested avoiding overly specific prompts (e.g., detailed activities or timing) and incorporating open-ended response options to encourage imagination and enhance connection to the scenarios.

**Finding 3: Ethical Discussion.** Participants engaged in discussions prompted by the game's ethical dilemmas, expressing the challenge of selecting a single answer from limited choices. To address this, they proposed implementing a ranking score system to allow for more nuanced responses. Some participants were unsure about the benefits of discussions for uncovering their AI user type. They suggested clarifying discussion goals and offering incentives for active participation. Additionally, participants recommended using a centralized moderator and AI voice for reading prompts to streamline gameplay and enhance immersion. Finally, participants emphasized the need for a safe environment between other players; they were worried that introverted players would not give their input. It was suggested to cluster stakeholders in dedicated sessions.

**Finding 4: Contextual Clarity.** Participants highlighted the need for additional context surrounding AI prompts to facilitate informed decision-making. They proposed introducing a game board element displaying complementary information.

**Finding 5: Understanding of AI in mHealth.** The game facilitated an understanding of AI's real-world implications as it revealed participants' varied comfort levels with sharing different data (e.g., social media vs. physiological) for an AI-driven mHealth tool.

**Finding 6: Influence of Phrasing.** Participants identified potential bias from prompt wording and tone. They recommended neutral language while acknowledging humor's role in fostering curiosity and engagement.

**Finding 7: Digital Format.** While some participants appreciated the physical format, transitioning to an online platform was viewed favorably. This could enable new mechanics like nuanced scoring and virtual moderation. Participants believed that an online version would be more accessible and inclusive, potentially reaching a wider audience beyond physical group settings.

## 4. Discussion

### 4.1. Game Adherence to Objectives

Pilot testing demonstrated the serious game's potential to achieve its key objectives. First, the game promises to attract participants (Objective 1) through gamification elements like badges and personalized user types, offering an enjoyable and stimulating self-discovery experience. Moving forward, clarifying discussion goals and rewarding participation could enhance engagement further.

Second, data privacy concerns raised during gameplay highlight the game's ability to guide the player in learning the implications of AI in mHealth (Objective 2). A storytelling dynamic holds the potential to further contextualize different uses of AI in mHealth. Hence, this gamified approach seems suitable for including stakeholders with low AI literacy in the design process. In addition to assessing ethical trade-offs between data privacy and potential health outcomes, this gamified tool could be leveraged in other stages in the design process, e.g. to assess the ease of use of prototypes or evaluate stakeholders' feeling of empowerment in the co-creation of new technology. Despite such opportunities, there is a need to explore how such a gamified approach could be scaled without a cumbersome effort in adapting the game to new use cases.

Third, the structured format encourages stakeholders to debate the ethical implications of AI technologies (Objective 3). When players shared their decision-making process between different ethical *human reactions*, they provided nuanced insights that may inform AI developers in making design decisions. In future work, each game card or AI prompt could be linked to an AI development decision, where quantitative analysis of the players' choices may translate stakeholders' values into actionable insights in alignment with Trustworthy AI principles.

### 4.2. Future directions

Three possible future directions emerge to refine the game in subsequent design iteration:

1. **In-person Digital Approach:** Moving the game to a digital platform while preserving its engaging elements could enhance accessibility and scalability. A digital version could introduce nuanced scoring mechanisms, virtual moderation, and incorporate additional contextual information to improve the game's effectiveness and reduce response bias.
2. **Blended Approach:** Combining elements of the paper-based game with digital components offers the advantages of both formats. This approach could maintain tangible interaction with physical cards while integrating online features for enhanced scoring, moderation, and broader engagement across different settings. It would cater to diverse preferences and maximize the game's impact.
3. **Digital Survey Approach:** A digital survey format could target stakeholders who may be unwilling to dedicate time to gameplay but whose input remains valuable for AI system design. While this approach could scale distribution and provide more representative data, it has fewer gamification opportunities for engagement (Objective 1) and may sacrifice the nuanced personal values that emerge from meaningful discussions during gameplay (Objective 3), which are not trivially realized in online settings.

In future research, choosing the most suitable approach depends on the desired engagement, accessibility, and depth of insights needed for ethical AI design and development, where A/B testing could provide further insight. Further exploration and refinement are crucial for maximizing the game's potential.

Future validation efforts should involve broader testing with diverse stakeholder groups beyond academic researchers and longitudinal studies to assess the game's impact on stakeholders' attitudes and decision-making processes over time, establishing it as a reliable tool for promoting responsible and trustworthy AI development.

## 5. Conclusion

Featuring a gamified engagement, a deeper understanding of AI applications, and in-depth ethical discussions, this gamified approach shows promise as a tool to support the development of trustworthy AI in mHealth aligned with stakeholder values. Further refinement efforts could explore a fully digital format prioritizing accessibility and nuanced scoring, a blended physical-digital approach, or even a streamlined online survey, depending on the desired balance between engagement, accessibility, and depth of stakeholder insights gleaned.

## Acknowledgments

The authors gratefully acknowledge the contributions of focus group researchers for their valuable feedback shaping the next game design iteration.

## References

- [1] A. Kankanhalli, Q. Xia, P. Ai, X. Zhao, Understanding personalization for health behavior change applications: A review and future directions, *AIS Transactions on Human-Computer Interaction* 13 (2021) 316–349. doi:10.17705/1thci.00152.
- [2] S. McGregor, Preventing repeated real world AI failures by cataloging incidents: The AI incident database, in: *Proceedings of the Thirty-Fifth AAAI Conference on Artificial Intelligence, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, The Eleventh Symposium on Educational Advances in Artificial Intelligence*; 2021 Feb 2-9; Virtual Event, AAAI Press, 2021, pp. 15458–15463. doi:10.1609/AAAI.V35I17.17817.
- [3] E. Commission, C. Directorate-General for Communications Networks, Technology, Ethics guidelines for trustworthy AI, Publications Office, 2019. doi:10.2759/346720.
- [4] European Commission, C. a. T. Directorate-General for Communications Networks, The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment, Publications Office, 2020. doi:10.2759/002360.
- [5] C. Huang, Z. Zhang, B. Mao, X. Yao, An overview of artificial intelligence ethics, *IEEE Transactions on Artificial Intelligence* 4 (2022) 799–819. doi:10.1109/TAI.2022.3194503.
- [6] L. van Velsen, G. Ludden, C. Grünloh, The limitations of user-and human-centered design in an ehealth context and how to move beyond them, *J Med Internet Res* 24 (2022) e37341. doi:10.2196/37341.

- [7] P. Schaar, Privacy by design, *Identity in the Information Society* 3 (2010) 267–274. doi:10.1007/s12394-010-0055-x.
- [8] B. Friedman, P. H. Kahn, A. Borning, A. Huldtgren, Value sensitive design and information systems, in: N. Doorn, D. Schuurbiers, I. van de Poel, M. E. Gorman (Eds.), *Early engagement and new technologies: Opening up the laboratory*, Springer Netherlands, 2013, pp. 55–95. doi:10.1007/978-94-007-7844-3\_4.
- [9] B. Friedman, D. G. Hendry, A. Borning, A survey of value sensitive design methods, *Foundations and Trends® in Human–Computer Interaction* 11 (2017) 63–125. URL: <https://www.nowpublishers.com/article/Details/HCI-015>. doi:10.1561/1100000015, publisher: Now Publishers, Inc.
- [10] R. J. Wieringa, The design cycle, in: *Design Science Methodology for Information Systems and Software Engineering*, Springer Berlin Heidelberg, 2014, pp. 27–34. doi:10.1007/978-3-662-43839-8\_3.
- [11] R. Hunicke, M. LeBlanc, R. Zubek, Mda: A formal approach to game design and game research, in: *Proceedings of the Nineteenth AAAI Conference on Artificial Intelligence*; 2004 Jul 25-29, volume 4, San Jose, CA, USA, 2004.
- [12] J. L. R. Robledo, F. N. Lucena, S. J. Arenas, Gamification as a strategy of internal marketing, *Intangible Capital* 9 (2013) 1113–1144. doi:10.3926/ic.455.