

# Towards Ethical Guidelines for Videogame Character Creation with GenAI

**Jana Hecktor**

University of Tübingen  
International Center for Ethics in the Sciences and Humanities  
Wilhelmstraße 56  
72074 Tübingen, Germany  
jana.hecktor@uni-tuebingen.de

**Theresa Krampe**

University of Tübingen  
International Center for Ethics in the Sciences and Humanities  
Wilhelmstraße 56  
72074 Tübingen, Germany  
theresa.krampe@izew.uni-tuebingen.de

## Keywords

GenAI, game development, conversational AI, videogame characters, ethical values, transparency, privacy, fairness

## EXTENDED ABSTRACT

Generative AI (GenAI) systems such as Large Language Models (LLMs) or Large Multimodal Models (LMMs) have already become pervasive in many areas of our lives. With large firms such as Ubisoft and Xbox currently investing considerable resources, and with highly visible showcases and prototypes such as *NeoNPCs* (Ubisoft 2024) or *AI People* (GoodAI 2024) appearing more and more frequently, it stands to reason that GenAI will also play a major role in game development and gameplay in the not-too-distant future (Shaker et al. 2016; Fuchs and Sudman 2019; Thompson 2024). However, implementing GenAI in videogames also brings new kinds of challenges to games and gaming, not all of which are addressed by existing regulatory frameworks. In addition to optimizing performance and ensuring legal compliance, it is therefore important to consider the ethical implications of using GenAI in videogames. In our paper, we examine present and future uses of GenAI in videogames from an ethical perspective, focusing on the creation of character models and conversational AI as two of the most promising applications at present (Gallotta et al. 2024; Sas 2024; Yannakakis and Togelius 2024; Yang et al. 2024).

The overall goal of our research is to develop a set of ethical guidelines to help game designers make informed decisions about when and how to use GenAI for creating videogame characters. To do so, our analysis proceeds in three steps. First, based on a review of existing ethical frameworks (e.g., Hagendorff 2020; AIEIG 2020; Canca et al. 2024; Melhart et al. 2024), we conceptualize a set of key values to guide a value-

Proceedings of DiGRA 2025

© 2025 Authors & Digital Games Research Association DiGRA. Personal and educational classroom use of this paper is allowed, commercial use requires specific permission from the author.

sensitive approach to creating NPCs with GenAI, with fairness, sustainability, transparency, and privacy protection as our main contenders.

**Fairness** seeks to ensure the just distribution of the costs and benefits of AI. In AI ethics, it has also taken on a more specific meaning as the mitigation of discriminatory biases (Barocas et al. 2023; Buolamwini and Gebru 2018; Mehrabi et al. 2021). Context-sensitive approaches must be attuned to power imbalances—from stereotypes encoded in chatbots and image generators to so-called “click-work” in the Global South (Tacheva and Ramasubramanian 2023). Using fairness as a value for creating videogame NPCs with GenAI thus also means paying attention to the labor conditions within and beyond the gaming industry, to the epistemic injustices that may become encoded in the game, and to the ways the game is distributed to and used by different player communities.

**Sustainability**, here understood mainly in the sense of environmental sustainability, ties in with the increasing research interest in videogames’ relation to planetary health and the ecosystem. As is well known, training and using GenAI requires vast amounts of energy and other resources and fuels forms of extractivism (Crawford 2021). This necessitates careful consideration of the environmental costs of using GenAI (e.g., by choosing more sustainable models where possible), though it also seems worth exploring the potential of AI-powered games to promote environmental awareness in players (Zhang et al. 2025).

**Transparency** refers to the explainability and interpretability of the system by different stakeholders (AIEIG 2020). Game designers, for instance, should receive information about the capabilities of the system and its intended and unintended uses, or the data used to train it, while players must be made aware of *if* and *when* they are interacting with an AI system (rather than, say, another player or a scripted NPC) in a videogame. Key questions thus include which aspects of the AI system and its implementation within the game should be made transparent, and how this could be achieved—with common approaches in other areas of AI design ranging from datasheets and model cards to labelling systems (e.g., Gebru et al. 2021).

**Privacy** typically concerns data protection and the right to control personal information (e.g., Nissenbaum 2010; European Union 2016). Gaming devices already collect or infer an astonishing amount of player data, including sensitive information about age, gender, emotional state, or even health (Kröger et al. 2023; Melhart et al. 2024). The widespread adoption of GenAI will likely increase the gaming industry’s hunger for “Big Data” and exacerbate existing concerns about privacy. Regarding the ethical use of GenAI in videogames, developers must consider questions of informed consent, identifiability, access to data, or future uses of the data—both for the training data and the data collected during gaming sessions (Ostritsch 2019).

Moving on to the second step in our research endeavor, we seek to bridge the gap between theory and practice by supplementing our ethical analysis with an interview study. Based on grounded theory models (Charmaz 2006), we conduct semi-structured interviews with professionals that already use GenAI for NPC design or consider doing so in the near future. The interviews are subsequently transcribed and analyzed using MAXQDA. During the first, exploratory phase of the interview study, we gained insights into the ethical and practical concerns that shaped game designers’ engagement with AI systems. We also used the advantage of semi-structured interviews to ask questions that are more open, which helped us uncover angles

perspectives that did not emerge as clearly from our literature review—such as game designer’s emphasis on solidarity with fellow artists when it comes to questions of copyrighted training data, or the fact that ethical requirements may differ significantly depending on the game genre. This first round of interviews will be followed by a second, refined interview study.

As a third and final step, we will consolidate the findings of our ethical analysis and the interview results and use them to develop a set of ethical guidelines for a value-sensitive approach to using GenAI in videogames. By combining theory with insights from practice, we hope to not only provide inspiration for future research but also to create a blueprint for ethics guidelines that can be circulated among stakeholders and form the basis for guidelines, policies, and frameworks that help implement ethical values in game design.

## REFERENCES

- AIEIG (AI Ethics Impact Group). 2020. *From Principles to Practice: An Interdisciplinary Framework to Operationalise AI Ethics*. Gütersloh, Germany: Bertelsmann Stiftung.
- Barocas, S., Hardt, M. and Narayanan, A. 2023. *Fairness and Machine Learning*. Cambridge, MA, USA: MIT Press.
- Buolamwini, J., and Gebru, T. 2018. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification." In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 77–91. <https://proceedings.mlr.press/v81/buolamwini18a.html>.
- Canca, C., Haaber Ihle, L. and Schoene, A. M. 2024. "Why the Gaming Industry Needs Responsible AI." *ACM Games* 2(2): Article 17. [doi.org/10.1145/3675803](https://doi.org/10.1145/3675803).
- Charmaz, K. 2006. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. London, UK: Sage.
- Crawford, K. 2021. *Atlas of AI: Power, Politics, and the Planetary Cost of Artificial Intelligence*. New Haven, CT, USA: Yale University Press.
- Fuchs, M. and Sudmann, A. 2019. "Games and AI: Paths, Challenges, Critique." *Eludamos: Journal for Computer Game Culture* 10 (1): 1–7. [doi.org/10.7557/23.6168](https://doi.org/10.7557/23.6168).
- Gallotta, R., Todd, G., Zammit, M., Earle, A., Liapis, A., Togelius, J. and Yannakakis, G. N. 2024. "Large Language Models and Games: A Survey and Roadmap." In *IEEE Transactions on Games*. [doi.org/10.1109/TG.2024.3461510](https://doi.org/10.1109/TG.2024.3461510).
- Gebru, T., Morgenstern, J. Vecchione, B., Wortman Vaughan, J., Wallach, H., Daumé III, H., & Crawford, K. (2021). "Datasheets for Datasets." *Communications of the ACM* 64 (12): 86–92. <https://doi.org/10.1145/3458723>.
- European Union. 2016. "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation, GDPR)." <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679>.
- GoodAI. 2024. *AI People*. PC game (alpha version). Good AI. <https://www.aipeoplegame.com/game/>.

- Hagendorff, T. 2020. "The Ethics of AI Ethics. An Evaluation of Guidelines." *Minds and Machines*: 1–22. <https://arxiv.org/pdf/1903.03425.pdf>.
- Kröger, J., Philip Raschke, Jessica Percy Campbell, und Stefan Ullrich. "Surveilling the gamers: Privacy impacts of the video game industry." *Entertainment Computing* 44: 100537. [doi.org/10.1016/j.entcom.2022.100537](https://doi.org/10.1016/j.entcom.2022.100537).
- Mehrabi, N., Morstatter, F. Saxena, N. Lerman K. and A. Galstyan. 2021. "A Survey on Bias and Fairness in Machine Learning." *ACM Computing Surveys* 54 (6). [doi.org/10.1145/3457607](https://doi.org/10.1145/3457607).
- Melhart, D., Togelius, J., Mikkelsen, B., Holmgård, C. and Yannakakis, G. N. 2024. "The Ethics of AI in Games." *IEEE Transactions on Affective Computing* 15 (1): 79–92. [doi.org/10.1109/TAFFC.2023.3276425](https://doi.org/10.1109/TAFFC.2023.3276425).
- Nissenbaum, H. 2010. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford, CA, USA: Stanford University Press.
- Ostrisch, S. 2019. "Computerspiele und Privatheit." *Privatsphäre 4.0: Eine Neuverortung des Privaten im Zeitalter der Digitalisierung*, edited by H. Behrendt, W. Loh, T. Matzner and C. Misselhorn, 231-244. Stuttgart, Germany, JB Metzler.
- Sas, M. 2024. "Unleashing Generative Non-Player Characters in Video Games: An AI Act Perspective." *IEEE Gaming, Entertainment, and Media Conference*, Turin, Italy, 5–7 June 2024. [doi.org/10.1109/GEM61861.2024.10585442](https://doi.org/10.1109/GEM61861.2024.10585442).
- Shaker N., Togelius, J. and Nelson, M. J. 2016. *Procedural Content Generation in Games*. Cham, Switzerland: Springer.
- Tacheva, J. and Ramasubramanian, S. 2023. "AI Empire: Unraveling the interlocking systems of oppression in generative AI's global order." *Big Data & Society* 10 (2). <https://doi.org/10.1177/20539517231219241>.
- Thompson, T. 2024. "The Changing Landscape of AI for Game Development." In *Game AI Uncovered*, edited by P. Roberts, 1–11. New York, NY, USA: Routledge.
- Ubisoft. *Neo NPCs*. Game prototype presented at GDC 2024. Ubisoft.
- Yang, D., Kleinman, E. and Hartevel, C., 2024. "GPT for Games: An Updated Scoping Review (2020–2024)." *arXiv*. arXiv:2411.00308v1 [cs.AI].
- Yannakakis, G. N. and Togelius, J. Forthcoming/2025. *Artificial Intelligence and Games*. 2nd ed. Cham: Springer. <https://gameaibook.org/>. (early access).
- Zhang, Q., Wen, R., Hendra, L. B., Ding, Z. and Ray LC. 2025. "Can AI Prompt Humans? Multimodal Agents Prompt Players? Game Actions and Show Consequences to Raise Sustainability Awareness." In *CHI '25: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* 1044: 1–29. <https://doi.org/10.48550/arXiv.2409.08486>.