Explorations with Ethically Aligned Stakeholder Elicitation (EASE)

ANNA-KAISA KAILA, PETRA JÄÄSKELÄINEN, and ANDRE HOLZAPFEL, KTH Royal Institute of Technology, Stockholm, Sweden

Development of data-driven tools for artistic creation touches upon a diverse range of stakeholders, but it is insufficiently recognised in the engineering communities working with creative technologies. To support increased awareness and sensitivity to ethical predicaments of the development work, we present an analytical for a structured, power-sensitive stakeholder identification and mapping – Ethically Aligned Stakeholder Elicitation (EASE). As a case study, we test the method in workshops with six groups that develop artificial intelligence in musical contexts (music-AI). The results from the workshops demonstrate that methods like EASE can effectively promote critical self-reflection and expose value tensions in the development processes, thus helping developers move towards ethically aligned research and development of creative-AI.

CCS Concepts: • **Human-centered computing** → **HCI theory, concepts and models**; • **Social and professional topics** → **Socio-technical systems**; • **Applied computing** → *Sound and music computing*.

Additional Key Words and Phrases: computational creativity, music, ethics, stakeholder, Value Sensitive Design

ACM Reference Format:
Anna-Kaisa Kaila, Petra Jääskeläinen, and Andre Holzapfel. 2023. Explorations with Ethically Aligned Stakeholder Elicitation (EASE). In . ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Developing responsible artificial intelligence (AI) applications is a key concern of current machine learning research, but considerations of societal impacts often lag the algorithmic development. Frequent controversies around creative-AI applications for data analysis and generation, such as litigation around the data use legitimacy [10], or dilution of artist portfolios through generative pastiches [34], demonstrate that insufficient consideration is given to the ethical predicaments of such technologies.

Several high-level principle documents for AI and data ethics have been published in recent years [9], but their abstract principles have failed to find any significant traction in the development work. To overcome this gap, new pragmatic methods are needed to support the developer communities in integrating ethical and value reflection into their work on creative-AI tools.

In the wider context of Value Sensitive Design and the people orientation of care ethics, we present a method for structured ethical analysis: Ethically Aligned Stakeholder Elicitation (EASE) [10]. As a case study, we test our method in workshops with six research groups that develop AI in musical contexts. Our results demonstrate that EASE supports critical self-reflection of the research practices, discloses power and value tensions in the development processes, and foregrounds opportunities for stakeholder engagement. These insights can guide developers on their journey towards ethically aligned research and development of creative-AI.

Manuscript submitted to ACM

2 CALL FOR PRACTICAL AI ETHICS

Along with the increased attention to developing AI technologies, a growing number of ethical guidelines for general AI use and development have been published [39] by institutions such as IEEE [16], European Commission [4], and UNESCO [17]. These documents are typically generic, principle-based ethics frameworks that approach their subject matter at a highly conceptual level. Several authors have lamented the mismatch between such high-level policy work and the practical needs of developers who are working with limited resources and under time pressure [11, 12].

There seems to be a need for applied ethics tools that are readily approachable in the narrow development pipelines and accessible as a primer for ethical realignment even with limited prior expertise in ethics (for previous initiatives, see for instance [2, 3, 5, 13, 14, 18]). More specifically, the exploration of ethical predicaments in creative-AI development should start from an analysis of concrete implications of individual research projects on actual people, with a focus on value tensions and power relations involved. For this end, we introduce an analytical approach within the framework of Value Sensitive Design (VSD) [7] and care ethics [8] and test it in the particular context of creative-AI development.

3 EASE METHOD

EASE was designed to be run in a workshop format that centers around a scenario of an application development project. The process runs in two steps: ethically informed stakeholder identification and power-mapping.

3.1 Workshop structure

In the first exercise of the workshop, participants identify and lay out the stakeholders of their scenario on a circular board. The process is facilitated by a list of nine question prompts that guide the participants to extend their consideration of stakeholder groups beyond the core development team and immediate collaborators. The questions are informed by the Data Ethics Canvas [13], but have been reformulated and abridged to ”who”-questions[1] in order to help participants elicit perspectives related specifically to people, in line with the care ethics framework.

In the second step, the identified stakeholders are evaluated in terms of their relative power and interest towards the project on a 2x2 grid [1, 6]]. Through this process, disparities in the power distribution between stakeholder groups become apparent and the resulting grid is then used as a basis for discussing potential action points for the participants.

Current workshop format was tested in April-May 2022 with six volunteering teams (17 participants total) active in music-AI research in academic research institutions in Australia, Austria, Finland, France, and Great Britain. Each workshop lasted approximately 90 minutes, with an equal amount of time devoted to stakeholder identification and mapping, respectively.

3.2 Insights from EASE workshops

Workshop participants identified between 9 and 30 stakeholders or stakeholder groups, including typically artists, data providers, users and listeners, research communities, and various companies and employees in the music industry. Meaningful differentiations, however, emerged within these broad categories in the course of the workshops. For example, participants distinguished between musicians and composers actively involved in development, testing or use of their system from musicians whose works are part of the used training data sets.

Previously undetected use cases were often discovered and explored in the workshops, as were the conditions of legitimate and fair data use in AI research, value tensions, biases, and structural inequalities within music industry.

[1]For instance, "Who provides data for the project?" and "Who does not use your product but could still be positively impacted by it? Are you making things better for the society? For whom (individuals, groups, demographics, or organisations?"

Some of the teams recognised their own ability to foreground user and other stakeholder engagement in their work and influence the power dynamics in order to protect and empower weaker stakeholders or alleviate harms towards them. As a condition for these efforts, teams expressed they would first need to critically reflect on and define their internal values. Furthermore, it was observed that the ethical stance of the development project may change over time. The exploration initiated in the workshops should therefore fuel a continuous process of proactive ethical reflection including, but also extending beyond the stakeholder elicitation.

Overall, the workshops succeeded in stimulating conversation and debate about a wide range of ethical predicaments, opening up perspectives towards more speculative scenarios of societal and ethical impacts in the current time or in the foreseeable future. Participants themselves described the process as ”revealing” and ”holistic”, reporting they had gained ”a better view” of the roles different stakeholders play in the project, while the process had helped "to elicit values”.

4 FURTHER DEVELOPMENT OF EASE

Besides insights on the ethical impacts of the individual music-AI projects themselves, the workshops shed light on the benefits and limitations of EASE in facilitating structured ethical reflection in the creative-AI development.

Sufficient size of the participant group and its internal diversity, as well as individual access of each team member to the workshop board, proved to be critical factors for the success of the workshops in providing a variety of insights. Testing also exposed certain blind spots in the Data Ethics Canvas framework used for the prompt questions, such as the role of funding institutions, the structures of ownership, access, and economic power of commercial platforms or service providers and more-than-human agents, such as ecosystems. While developing the method further, we will look into reformulating some of the question prompts to provide wider topic coverage, better clarity and inspiration.

To complement the current case study, the testing of EASE could be extended to wider geographic areas, to adjacent fields of creative-AI (e.g. visual, performative), and possibly to commercial developers of creative-AI applications. Furthermore, we aim to accompany some of the participating teams in longitudinal studies to evaluate how action points developed using EASE are implemented in long term, observe, and analyse the current AI-development culture and practices in greater detail, and collaboratively develop new solutions from bottom-up rather than top-down. These efforts will add to the insights obtained by [15] and others, and provide critical domain-specific context both for the developing further creative-AI specific ethical analysis methods, as well as for drafting of policies and regulatory frameworks for creative-AI.

5 CONCLUSION

The speed of research and development of creative-AI applications is currently far outpacing the reflection on social impacts and ethics. To support the creative-AI community in finding ways to proactively contribute to the development of ethically aligned and responsible applications, we have introduced an analytic method, Ethically Aligned Stakeholder Elicitation (EASE).

Based on the observations and feedback from the workshops, EASE stimulated a structured critical self-reflection, inspired teams to discover previously unidentified stakeholders and other new aspects of their work, as well as articulate a shared understanding of value tensions and power asymmetries that require their attention. As a low-barrier analytical method, EASE paves the way towards sector-specifically relevant ethical standards that can guide the progress of the creative-AI development in the future. As we continue developing and testing of EASE, we hope this work to be inspiration for more collaborative and interdisciplinary efforts of co-designing a greater diversity of AI ethics analysis approaches.

REFERENCES

[1] John M Bryson. 2004. What to do when Stakeholders matter. *Public Management Review* 6, 1 (2004), 21–53.

[2] Digital Catapult. 2021. Ethics Framework. https://aiblindspot.media.mit.edu. Last accessed Dec 07, 2022.

[3] European Commission and Directorate General for Communications Networks, Content and Technology. 2020. *The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment.* Publications Office. https://doi.org/doi/10.2759/002360

[4] European Commission. Directorate General for Communications Networks, Content and Technology. and High Level Expert Group on Artificial Intelligence. 2019. *Ethics guidelines for trustworthy AI.* Publications Office, LU. https://data.europa.eu/doi/10.2759/346720

[5] Aline Shakti Franzke, Iris Muis, and Mirko Tobias Schäfer. 2021. Data Ethics Decision Aid (DEDA): a dialogical framework for ethical inquiry of AI and data projects in the Netherlands. *Ethics and Information Technology* 23 (2021), 551–567.

[6] R Edward Freeman. 1984. *Strategic management: A stakeholder approach.* Pitman series in Business and Public Policy.

[7] Batya Friedman and David G Hendry. 2019. *Value sensitive design: Shaping technology with moral imagination.* MIT Press, Cambridge, MA.

[8] Virginia Held. 2006. *The ethics of care: Personal, political, and global.* Oxford University Press, Oxford.

[9] Anna Jobin, Marcello Ienca, and Effy Vayena. 2019. The global landscape of AI ethics guidelines. *Nature Machine Intelligence* 1, 9 (2019), 389–399.

[10] Anna-Kaisa Kaila, Petra Jääskeläinen, and Andre Holzapfel. 2023. Ethically Aligned Stakeholder Elicitation (EASE): Case Study in Music-AI. In *Proceedings of NIME.*

[11] Jessica Morley, Luciano Floridi, Libby Kinsey, and Anat Elhalal. 2020. From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. *Science and Engineering Ethics* 26 (2020), 2141–2168. Issue 4. https://doi.org/10.1007/s11948-019-00165-5

[12] Luke Munn. 2022. The uselessness of AI ethics. *AI and Ethics* (2022). https://doi.org/10.1007/s43681-022-00209-w

[13] Open Data Institute. 2021. Data Ethics Canvas. https://theodi.org/article/the-data-ethics-canvas-2021/ Last accessed 18 May 2022.

[14] Cathy O'Neil and Hanna Gunn. 2020. Near-Term Artificial Intelligence and the Ethical Matrix. In *Ethics of Artificial Intelligence*, S. Matthew Liao (Ed.). Oxford University Press, 235–69.

[15] Nick Seaver. 2022. *Computing Taste: Algorithms and the Makers of Music Recommendation.* University of Chicago Press, Chicago.

[16] The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. 2019. *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems.* Technical Report First edition. IEEE. https://standards.ieee.org/content/ieee-standards/en/industry-connections/ec/autonomous-systems.html

[17] UNESCO. 2022. Recommendation on the Ethics of Artificial Intelligence. https://unesdoc.unesco.org/ark:/48223/pf0000381137 Last accessed 18 May 2022.

[18] Ville Vakkuri, Kai-Kristian Kemell, Marianna Jantunen, Erika Halme, and Pekka Abrahamsson. 2021. ECCOLA — A method for implementing ethically aligned AI systems. *Journal of Systems and Software* 182 (2021), 111067.

Received 23 February 2023; accepted 10 March 2023