



Delft University of Technology

Tools and techniques to promote reflexivity with the embodied AI community

Zaga, Cristina; Lupetti, M.L.; Cila, N.; Lee, Minha; Huisman, G.; Fosch Villaronga, E.; Arzberger, A.

Publication date

2022

Document Version

Final published version

Published in

Diversity Equity and Inclusion in Embodied AI

Citation (APA)

Zaga, C., Lupetti, M. L., Cila, N., Lee, M., Huisman, G., Fosch Villaronga, E., & Arzberger, A. (2022). Tools and techniques to promote reflexivity with the embodied AI community. In M. L. Lupetti, C. Zaga, & G. Huisman (Eds.), *Diversity Equity and Inclusion in Embodied AI: Reflecting on and Re-imagining our Future with Embodied AI* (pp. 39-48)

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

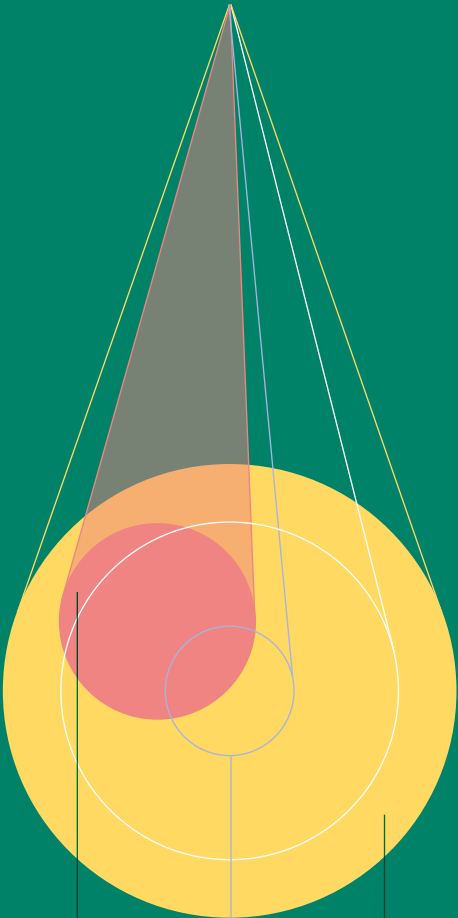
Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Tools and techniques to promote reflexivity with the embodied AI community

— BY CRISTINA ZAGA,
MARIA LUCE LUPETTI,
NAZLI CILA, MINHA LEE,
GIJS HUISMAN, EDUARD
FOSCH VILLARONGA,
ANNE ARZBERGER

NOW



PREFERRED
FUTURES

PROBABLE
FUTURES

POSSIBLE
FUTURES

TEAM 1

TEAM 2

TEAM 3

TEAM 1

TEAM 2

TEAM 3

Reflecting on the implicit assumptions

WHAT IS THE TOOL ABOUT

The tool offers a two-step activity for tangible reflections related to how we design embodied AI and imagine possible, probable, and desirable futures. The first activity of the tool presents current scenarios and embodied AI artifacts that embed specific narratives, biases, and stereotypes. The second activity uses the future cone metaphors to support participants in reflecting on how the future embodied AI will interact with the current infrastructure, who will be involved, and how it might be co-opted. The tool can be used to reflect on implicit assumptions we have about different topics. In our workshops, we used it to reflect on gender and ableist assumptions.

HOW TO USE IT

The tool can be made available online for collaboration, e.g., through Miro, or also printed. Participants are invited to split in groups, read the instructions on the left and work in the assigned space on the right. Teams are asked to note down on post-it their reflection and share them with the other teams. The activity takes 40 to 75 minutes, depending on the number of participants.

WHEN TO USE IT

We suggest using the tool at the very beginning of the design process to map current issues and challenges and kick off a reflection about future visions. It is suitable to establish disciplinary and interdisciplinary understanding.

FOR WHO

Students, researchers, practitioners, and policymakers working in embodied AI.

Mapping privileges

WHAT IS THE TOOL ABOUT

This tool is designed to let people reflect on personal position of privilege, or lack of. The tool is based on reknown models of privileges and power, such as the one by Erete et al. (2018) that shows intersecting axes of privilege, domination, and oppression. Our tool invites participants to position themselves into binary axes generally associated with privilege, i.e., white color of the skin vs. black. The tool is purposefully forcing people into binary categories that hardly represent the complexity of reality. By doing so, it aims at provoking discussions and reflections on how we personally experience privilege and how also we attribute it to others.

HOW TO USE IT

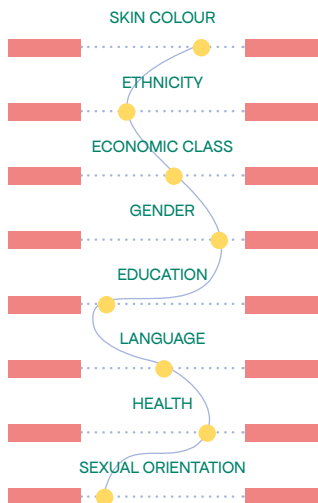
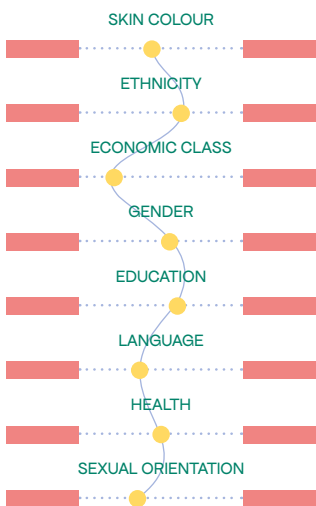
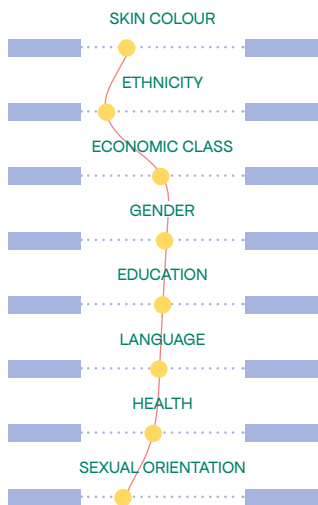
The tool can be prepared as a card, either printed or set up in a digital collaboration environment, i.e., Miro. The activity is organised into three steps. First participants need to split into groups of maximum four people and each person has to fill in one card individually. After that, people share how their 'curve' looks like (see in the example image) and discuss how this matches or not with their felt experience of privilege or exclusion. Last, all participants share highlights of the discussion emerged in their group. The whole activity is guided by a moderator who can, optionally, decide to turn the activity also into a sort of satiric competition about who holds more privilege.

WHEN TO USE IT

The tool is particularly suited for workshop settings where people need to familiarize with the concepts of diversity, privilege and power. The activity is best suited at the beginning as it facilitates jumping into the topic while also functioning as an icebreaker.

FOR WHO

Students, researchers, practitioners, policymakers with an interest in diversity, equity and inclusion. It is not necessary to be an expert in the field nor to have any technical knowledge.





Punkbots collages against status quos

WHAT IS THE TOOL ABOUT

The tool is meant to support the overturn of the status quo of robot design using déorname^{nt} activities inspired by punk techniques by the Letterist international. These activities help participants to grow awareness and implement critical reflection in their design process. Participants are going to be presented with one HRI scenario (photograph and description) and two bots: a StatusQuo Bot and Punkchabot. The bots are designed with DALL-E and presented to the participants through a power-point presentation. The bots are prompting the participants to first describe the status quo of the images and then overturn it. To do so, they use a collage technique developed by Letterist International to reverse storylines and images, collating and pasting images that change their meaning.

HOW TO USE IT

The activity is better suited for in-person workshops, but can easily be organized online. Materials are prints of status quo pictures, post-its and magazines and scissors for collages. The activity is designed for 45–75 minutes but can be shortened if necessary. To facilitate the activity, moderators should be ready with questions to prompt participants and have ready examples of détournement from Letterist international. Please beware of accessibility issues for disabled participants.

WHEN TO USE IT

This tool is meant to support transdisciplinary discussion and to create common ground and consciousness-raising among participants.

FOR WHO

The tool is meant for professionals, academics, and non-professional communities. It is not necessary to be an expert in the field nor to have any technical knowledge.

Exploring spaces between categories — A biased classifier

WHAT IS THE TOOL ABOUT

Personal values, interests and experiences can impede designers' intentions in creating fair and inclusive solutions for people of all gender. Counteracting bias perpetuation and augmenting limitations of human cognition, this tool aims to break with stereotypical expectations and thinking into binary categories. By surfacing our unconscious associations and the narrow of ways of categorizing things, a classification algorithm (i.e., Teachable Machine) can be used to help us challenge gender norms and stereotypes.

HOW TO USE IT

Step 1 (for groups) — Visualize yourself — Each participant creates 3–4 cards that represent personal interests, personality and identity. This can be done by cutting out images from magazines and gluing them on paper cards.

Step 2 — Curate the Dataset — Each person or team has to create its own dataset. For this activity, either the facilitator or the individual participant can setup a board on Miro (see example online). First two categories are set, for instance “male” and “female”. Next the participant can intuitively classify images from magazines or images found online in either one of the binary categories. This can be done as a team of max 4, or individually.

Step 3 Train the classifier — Download the collected images, keeping them separated into categories, and pass them on to the classification algorithm. You can use any classifier for this experiment, however we recommend Teachable Machine. Once the images are uploaded you can start the training.

Step 4 Explore, challenge and learn — Try out any kind of object, design, prototype, or photo, including the identity cards, to explore how the “biased” classifier we created sees the world. For this, simply upload the images of interest or hold your object in front of the camera and observe to what extent the classification result reflects your view of the object. Question why the algorithms reacts in a certain way and how can you influence that.

**MATERIALS
(FOR
GROUPS)**

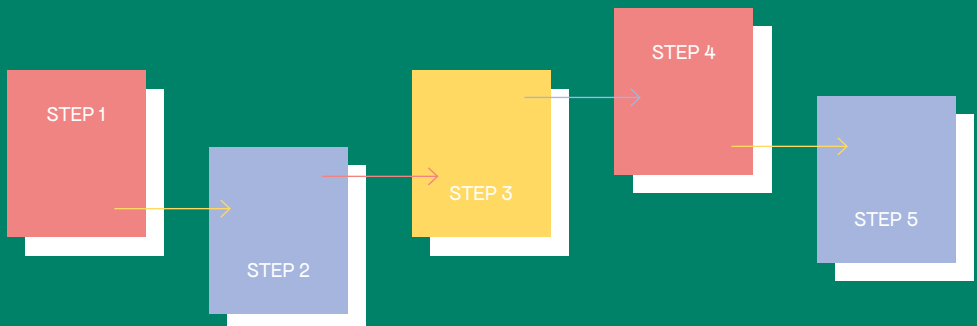
Magazines (e.g. science, cars, nature, baking, etc.), Scissors, Glue, Paper; mandatory: Laptop (with camera), internet to access Miro, for material preparation, and Teachable Machine, for running a classification algorithm.

**WHEN TO
USE IT**

This tool can be used when designing artifacts to be mindful and reflect on the unconscious biases that can be projected into objects. From early design process doodles, to more elaborate design sketches or even prototypes, looking through the eyes of the algorithm you trained can be a powerful moment of confrontation. The same reflexive potential can be used as a way of introducing people with diverse background and expertise to issues of diversity and biases into automated systems, in collaborative work settings, i.e. multistakeholders workshops.

FOR WHO

The tool can either be used in a workshop with multiple people, best grouped in teams of 4, or applied in individual sessions. For using the tool in a workshop include the “group” marked activities. Depending on the amount of groups, more than one facilitator might be necessary. If used by individuals, no facilitator is necessary.



To conclude this book, we want to offer a set of reflections distilled from our final internal event in which we looked back at the experiences of the workshops. What have we learned with students, researchers, policymakers, and communities? In this section, we argue for collective actions focused on re-imagining our futures with embodied AI and conclude with a vision of desirable future(s) horizons.