

## From 'Human Centered Design' to 'Humanity Centered Design': Towards an AI design methodology for the collective/public good

At the 2021 edition of the Mozilla Festival (on March 9th), around 20 participants took part in a workshop exploring a different approach to designing or developing AI. Rather than centering the human in the design process, the idea would be to center humanity (see: [What's next for design: Towards humanity-centered design](#)).

This document is a reflection of the work that was done in the workshop. After the summary, you will find the complete, anonymised (and unedited) output of the workshop participants.

### Summary of the workshop outcomes

The participants in the workshop could clearly see how AI is often designed and developed with a limited group of stakeholders in mind. This leads to a limited view on the impact of the AI technology, with unexpected results. Here are some of the key insights from the participants:

- Currently, AI designers usually **assumes that everybody has access** to the technology and that everybody has the ability to pay for certain services and lives in dense urban spaces.
- AI design also **doesn't take into account the larger social and natural environment** in which the technology operates. The consequences for the users are usually clear, but what it will do to the rest of society seems to be out of scope for consideration.
- For AI to work well, a certain level of knowledge (i.e. the models) are necessary. Very often, this **knowledge is in private hands**, leading to inequitable access.
- AI that learns from the world will **reflect the bias that is inherent in the world**.

Reflecting on this the workshop participants came up with a set of guidelines for AI designers and developers. Here are some of the key thoughts:

- The most obvious point is to **diversify the stakeholders** and to **make them active participants** in the design process, including their expertise and their needs.
- Furthermore, it is important to **look beyond the AI technology** and see how the technology is embedded in the societal and natural environment. What is the total process the AI is part of and how should that process be designed?

### Case 1: AI in routing traffic

Navigation apps have made travelling much easier. Not only do they help us find our destination, but adaptive navigation apps such as Google Maps and Waze can help us find a faster route based on traffic data. AI is used to predict future traffic based on current traffic (measured by the movements of other users of the app), but also based on historical traffic patterns, the day of the week and the weather.

For the individual owners of a navigation system this is a way to optimize their routes and to avoid the hassle of finding a route on a paper map themselves, although this may come at the expense of a loss of orientation. But what happens if everyone uses adaptive navigation? What is the consequence for people not using these apps?

### Question 1: Whose perspectives were considered when designing the AI based solution?

- Accuracy for maps is easier in cities but in rural areas it is not accurate → with access to infrastructure.

- Ad spend, revenue, shops etc.
- Apps that want to use or connect to the API
- Designers with the perspective of people who want to save time, cars users. It eliminates other options.
- Driver as a default.
- Individual users/Average user (average user not an end user, who is able bodies who has access to internet and phone).
- Keeping flow going, keeping users up to date (i.e. information on jams/roadworks etc.).
- People is not familiar with the place they are going to.
- People who can use smartphone.
- People who have digital knowledge to use it.
- People who trust the app.
- Person who has internet access. - The individual driver, the company that owns the app.
- User centered.

**Question 2: What problems are created (or taken for granted) by only looking at these perspectives?**

- Accuracy of information / Missing local context.
- Delay: on road changes.
- Everyone takes the same routes, the same experience.
- If everybody uses the system, the system becomes a standard. If you do not use the system, you may miss opportunities.
- It Eliminate new discoveries, it converts us to mechanical people.
- It has not in consider disabilities/other needs.
- Lack of accessibility —> exclusion (visually impaired).
- Lack of transparency.
- Loss of orientation skills.
- Loss of spatial orientation, because people over-rely on the apps; is there a need for a critical mass of people using it? Are new traffic patterns being created that disturb reasonable expectations of people who live in the neighborhood?
- Missing specifics of people's needs / curb information and prevents people from getting an overall picture.
- Question of value, historical, monetary, right to existence, who gets to name or erase a name or place.
- Real time tracking.
- The more usage, the better the systems. Hence, more dependent on technology. Tech dominance.
- The neighborhood might have a traffic overdose.

**Question 3: Whose perspectives should also have been considered when designing the solution?**

- Accessibility issues, alternative method of transport.
- Air quality for walkers? Pollen?
- Environmental concerns? Sustainability?
- Historical, “decolonialization” issues (place names)?
- Pedestrians.
- People from various disadvantaged communities (not average users) - older people, language barriers.

- People with accessibility issues.
- People with disabilities.
- People with no digital literacy.
- Places that don't or can't 'pay'.
- The communities involve; People that don't really are going by car/private transport.
- Urban planning and urban development authorities / city planners.
- Walking.
- What are the impacts in public transit efficiency?
- Does it impede how buses work?
- Privately-held data, how will it benefit citizens?
- Pedestrians.
- People who live in neighborhoods where traffic is being diverted.

**Question 4: How would you change the design of the AI (and the process it is embedded in) when considering these other perspectives and taking a more collective view?**

- Collaborative design approach.
- Incorporate different uses or functionalities.
- Neighborhoods, local communities can be involved in the process —> community deliberation on map routing.
- Obligation to analyze public transit and take it into account so that the app doesn't interfere with public transit; access to open data for citizens.
- Open source, adding a user or contributor layer (similar to being able to upload photos).
- Universal/accessibility approach / analogue piece.

**Case 2: AI in recruitment**

Employers that need to choose from many candidates to fill a single vacancy are using AI to screen candidates. The AI is used to automatically create a 'soft skills' profile for each candidate based on the text in their resume, LinkedIn account and motivation letter. This profile is matched to an idealized 'soft skills' profile which is created by analyzing the text in the resumes, LinkedIn accounts and motivation letters of high performing workers who are already in the role. The best matched candidates are invited for the next stage in the recruitment process, the other candidates are dismissed..w

For the employer this is a very efficient way of finding good candidates that fit the role and your company culture. It also removes the potential bias of individual recruiters. But candidates are dismissed without any human intervention and potential biases in the employer's performance ratings can seep into the recruitment process if certain soft skills correlate with, for example, gender or ethnicity.

**Question 1: Whose perspectives were considered when designing the AI based solution?**

- Efficiency-driven.
- Employer / HR / Recruiting.
- Employer perspective (little to no consideration from candidates).
- HR is always the business's needs...
- Industry experts.
- People with privilege.
- Recruiter who want fast answers and to eliminate human bias or outsource bias to the AI.
- The employer, the recruiter.

- The recruiter and enterprises, not candidates.
- Who is this employer? Cis White Men?

**Question 2: What problems are created (or taken for granted) by only looking at these perspectives?**

- Caters to metropolitan urban spaces.
- Cultural discrimination.
- Difficult to create a fair system → Perpetuation of stereotypes.
- Difficult to interpret how do you transfer skills from one job to another.
- Difficult to measure motivation to learn new things.
- Discrimination on different aspects and levels (Race, Disability, Geolocation, Gender, Scholary etc.).
- Diversify the training data – collaborative cloud sources?
- Filter in (and out) the same kinds of candidates.
- Filter in the biases of the employer/recruiter.
- It takes for granted that strong employees will look like the existing employees.
- Keyword silos.
- Language, cultural, translation bias.
- People become resources.
- Prioritizes people with privilege (soft skills are objective).
- Racial and gender bias.
- Reproduce historical bias.

**Question 3: Whose perspectives should also have been considered when designing the solution?**

- AI can follow up / another round before disarding applicant.
- Candidates, people form differente background.
- Candidates' perspectives.
- Diverse range of applicants perspective.
- Human in the loop should be involved (more inclusive) on AI side.
- Multi-cultural contexts / language.
- People that work and research about diversity.
- People who don't want to be tagged.
- People who have to be recruited.
- People working with unemployed.
- The applicant.
- The job seeker.

**Question 4: How would you change the design of the AI (and the process it is embedded in) when considering these other perspectives and taking a more collective view?**

- Add examples of different keywords so that people who use different tones but are still excellent candidates won't be rejected.
- Define what a good candidate is independent from past candidates (to overwrite).
- Incorporate different dialects, different kind of languages, different motivations.
- Recruitment process or portal needs to have a human factor.
- Semi-automation.

## Guidelines

Participants in the workshop were asked to come up with new guidelines for AI developers/designers that would consider public values and would benefit humanity, rather than just the user.

## Highlighted

These were the suggested guidelines that were thought to be the most important:

- Rather than assuming that technology in isolation can solve larger social problems you should reach out to the community outside of your app/website and engage them (especially those who do not use your product) in your solutions.
- Rather than designing for users who might look like you, you should consider diverse audience and include others in the discussion (persons with disabilities, ethnic minority communities, those who are not digital natives).
- Rather than focusing only on the fastest route you should focus also on communities and environmental contexts when routing.
- Rather than focusing only on “successful” applications as envisioned by the AI tool you should include diverse applicant sets who might have different language/words/cultures.
- Rather than only thinking about design aspects you should also consider the knowledge paradigms in which the ‘design’ process is situated.
- Rather than relying only on AI-based rejection in the recruitment tool you should have humans in the loop who could check for biases in the training data set.
- Rather than working with software developers, designers, and managers to come up with the solutions to complex problems you should actively work with people who have experience in the field and who understand the intricacies of the problem you are trying to solve.

## Other

These were the other guidelines that participants in the workshop came up with:

- Rather than ask users for their opinion you should co-create actively with them, give them decision power about the creation of the system/software/solution.
- Rather than collecting as much data as possible you should limit data collection to the minimum needed to solve the problem.
- Rather than considering stakeholders and specialists points of view you should consider them, but look into third parts too (communities affected by the service or product, legal stuff, political situation etc).
- Rather than design a black box where user data feeds global outputs you should allow individual vetting/customization of their own data.
- Rather than determining needs of potential users one time at the beginning of the project you should continually update the needs.
- Rather than extract data from people without their knowledge or consent you should ask permission to use people’s data and make clear how it will be used to their benefit.
- Rather than focus only for people with local contexts you should also design for tourists, or people who might be new to the city.
- Rather than focusing on the use of the individual you should ask to different collectives how they are affected by the system.

- Rather than focusing only certain key words you should allow for more cultural contexts to derive what could be “successful” soft skill applications.
- Rather than focusing only on the driver, you should you should focus also on people with prams, people with mobility devices, non-able-bodied people.
- Rather than just focusing on the user and the app owner you should take into account the perspectives/needs of all relevant stakeholders.
- Rather than looking at product benefits only you should also consider risks to different aspects (human rights, democracy, discrimination, etc.) of the product.
- Rather than privately held data you should organize levels of open data for citizens and researchers.
- Rather than rejecting profiles on the basis of the AI output you should have a loop back system with the applicant to allow for their views and inputs to be considered.
- Rather than thinking about which technology AI practitioners want to use you should focus on the problem.
- Rather than using only one example you should have a broad spectrum to draw an average from.

## **License**

This work by Pascal Wiggers and Hans de Zwart from the Amsterdam University of Applied Sciences (and the 2021 MozFest workshop participants) is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](#).