

# Health Data Sharing is Caring Session plan (Remixable version)

This is the remixable version of the session plan presented at Mozfest 2019.

The plan was designed for a 90-minute session with a group of around 30 participants but can be easily be modified for smaller or bigger groups.

The plan was created by session co-facilitators Natalia Norori and Stefano Vrizzi and is licensed CC-BY 4.0. For inquiries please contact [natalianorori@gmail.com](mailto:natalianorori@gmail.com).

## Session Goal

- Raise awareness on the underrepresentation of minorities in health datasets, as the lack of available data about certain communities affects AI algorithm accountability, thus its impact on global health.
- Discuss how balanced healthcare datasets might help us develop more accurate diagnostic tools as well as avoid discrimination and inequality in healthcare.

## Session Plan

### Before the session

1. **Review the session plan and get feedback**
2. **Get informed.** We have curated a list of articles that can help you understand the subject before you start planning your session:
  - [Sharing and utilizing health data for AI applications](#) (this one is a must )
  - [How open data can save the NHS](#)
  - [Good Reddit thread on sharing genetic data](#)
  - [If AI is going to be the world's doctor it needs better text-books](#)
  - [AI-driven dermatology could leave dark skin patients behind](#)
3. **Make sure you have all the materials needed for the session. For our session we used:**
  - A projector (a laptop could also work).
  - Post-its and markers.
  - Session posters. We printed ours in London in 12"x18" sized paper. You can create your own or use the ones created by us. You can download them [here](#).
  - Duck tape.

**4. Get comfortable with the room you'll be running your session in.**

- Put the posters in the walls before the participants arrive, this will ensure you don't spend your session time doing so.
- Make sure the lighting in the room is friendly.
- Make sure the room is accessible for session participants with special needs.

**5. Make sure all your resources are accessible**

- Store all session materials and useful resources online. We recommend using [Zenodo](#) or [Figshare](#). Create a shortened URL to make it easier for participants to access your resources.

## During the session

**1. Introduce the session goals and ice-breaking activity (10 minutes)**

- Introduce yourself (points if you transmit excitement while doing so, participants appreciate knowing they're talking to someone who cares about the subject).
- Invite participants to introduce themselves, share their pronouns, and one word that describes how they are feeling today.
- Know your audience. Try to be as inclusive as possible, make sure you have pen and paper for those not comfortable with verbal communication.
- If you are running the session by yourself, ask a volunteer to take notes.
- After finishing the introductions make sure everyone is familiar with The Matrix Film and briefly explain the red pill blue pill analogy. Understanding this is essential for the session. You can get ideas from the introduction we used for our session:
  - In the matrix, a movie about machines controlling the world, the main character Neo is offered the choice between a red pill and a blue pill. The red pill represented an uncertain future—it would free him from the control of the machine-generated dream world and allow him to escape into the real world, but living the "truth of reality" is harsher and more difficult. On the other hand, the blue pill represented a beautiful prison— Basically red means harsh reality, and blue is equal to the "ignorance is bliss" concept. Please take this into account before we move to the next activity.

**2. Red pill blue pill posters (15 minutes)**

- We pasted 8 posters on the walls of our session room. The posters we used are preserved in Zenodo, and can be downloaded [here](#).
- Participants are invited to interact and reflect on the posters, and write their thoughts, questions, and reactions to each of the statements in post-its and paste them on each of the posters.

**3. Reflect on posters - in groups or as a whole ( 25 minutes)**

- We gave a brief introduction on why the subject is important to us, read some post-it's, and had participants share their thoughts on the posters out loud. Our session was an open space, and conversations were not limited to one subject. We

made sure participants understood everything they read on the walls was taken from recent research studies that prove sometimes algorithms do not work the way they are supposed to. The sources behind the posters were shared on the slides.

- In our session, we explained that we decided to run it because we had read a lot lately about the hype with how artificial intelligence might help us solve public health issues, and it drew our attention to see the inequalities in the datasets behind these algorithms. We also explained the goal of the session is to learn from each other, and brainstorm ways to raise awareness about the inequalities and lack of representation of minorities in healthcare algorithms, how this might affect the world as a whole and what we might do about it.

#### **4. Invite participants to form small groups to brainstorm for solutions to the most frequent challenges (20 minutes).**

You can do two things:

a) Quickly browse through participant posters, identify repetitive challenges and questions, and group them into categories. Share a few categories of post-its with each group participant for them to look for solutions.

or

b) Divide participants into 2, 4 or 8 groups, and invite them to help us find solutions to the challenges we created based on the results of our first workshop at Mozfest 2019. The challenges are all posted as GitHub issues [here](#), and include but are not limited to:

1. How might we gain back people's trust?
2. How might we help researchers identify imbalanced datasets?
3. How might we build a common language between healthcare professionals, researchers, and data scientists?
4. How might we help adapt RRI framework practices to healthcare data management?
5. Data management and sharing are not easy. Human resources and infrastructures are needed. How might we implement global open infrastructures that unlock data's potential to improve global health?
6. How might we contribute to closing feedback loops for better and healthier data management?
7. How might we raise awareness about the need for increasing the presence of non-binary and genderqueer individuals in healthcare datasets to properly identify their specific healthcare needs?
8. How might we help solve healthcare datasets bias problem?

Whichever activity you decide to run, please make sure to take notes of the participant's ideas.

## 5. Wrapping up ( 10 minutes)

- Conclude on your learnings and reflections.
- Make sure participants understand this is not a sad story, and offer them ways to help and get involved.
  - We thought it was a good idea mentioning projects that are already tackling data bias and raising awareness about the issues, such as The Data Nutrition Project, The NHS, Omdena, and Understanding Patient Data.
- Share session notes and other resources.
  - Remember the shortened link we asked you to create before the session? Now is the time to share it! This way, participants will have easy access to session notes and all your resources.
  - Make sure others can reach out to you after the session.

## 6. Optional (If you have time) Explore imbalanced datasets and biased algorithms (15 minutes)

- We shared links to 5 open datasets from the UCI Machine Learning Repository and Kaggle, and explained the analysis of a biased skin algorithm used to detect skin cancer. You can [access our code here](#). If you have time and would like for participants to interact with imbalanced datasets, this might be a good route to take for your session.

## After the session

### 1. Sharing is caring

Share the session outcomes on your social media profile or with anyone who might be interested if you want!

### 2. Email the participants and ask for feedback

### 3. Let us know how it went!

- We would love to learn more about your session. Please contact Natalia at [natalianorori@gmail.com](mailto:natalianorori@gmail.com) to share your learnings.
- [Sign up as a contributor](#) to the Github repository and contribute your learnings in the [issues section](#).

We hope the information above helps you remix the session! If you have any questions or would like to chat, contact Natalia and let us know.