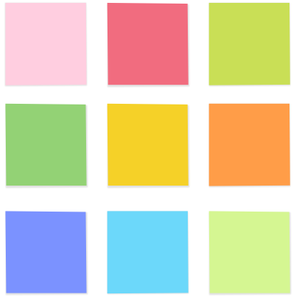
## 1.

| INTENTION | DESIRED OUTCOME | AGENDA | ROLES | RULES | TIME |
|---|---|---|---|---|---|
| Equip the nonprofit sector with the information it needs in order to implement AI responsibly and ethically. | Understand the potential risks of designing and using AI in the international development contexts.<br><br>Learn how to practically apply ethical considerations and implement AI responsibly. | (1) Introduction to the breakout<br>(2) Participant introduction<br>(3) Participants read the case study and ask for any clarifications<br>(4) Participants respond to the questions<br>(5) Discussion<br>(6) Consolidate feedback and get ready to present | We are the facilitators and you are the explorers.<br><br>We lead you through the process, and you actively participate - reflect, share, listen, and learn together. | The more you lean in, the better the experience, for you and everyone else.<br><br>Be proactive, open, respectful, and collaborative.<br><br>Be here now - no email or phones. | 75min |

## 2.

### INTRODUCTIONS

Your name, organization, role, country, and favorite food.



## 3.

### HUMANITARIAN RESPONSE

In humanitarian crises, many people go missing because of conflict, disasters, or during migrations. Currently, international humanitarian laws require accounting for missing persons and providing information to family members. Image processing and facial recognition technologies can be used to uniquely identify individuals in order to reconnect families that have been separated. Initiatives such as ICRC's Trace the Face have been using manual detection as a way of finding missing persons, and automating the process with AI/ML techniques can help identify individuals more quickly.

Consider an NGO that is working with a national government to implement facial recognition techniques on finding missing persons as a result of an ongoing conflict. They plan to use digital photographs submitted by the family members of missing persons and scraping photographs from public social media sources to identify missing persons.

(1) What are some of the fairness concerns that we must think about at different steps in a machine learning project development process?
(2) How might you address some of these concerns?
(3) What are some of the other concerns or additional considerations that will be relevant when deciding whether or how this ML system should be integrated into decision-making processes?

## 4.

| What are some of the fairness concerns? | How might you address some of these concerns? | What are some of the other concerns? |
|---|---|---|
| *Consider this: How might data and ML model implementation cause disproportionate harm? How well do we understand how ML models work? What happens when things go wrong?* | *For example: Asking about bias in data, including protected attributes, diversifying development teams, etc.* | *For example: Organizational capacity to implement, privacy and under-regulation, whether ML is the right approach, etc.* |

**What are some of the fairness concerns?**
Fairness refers to how benefits and harms are distributed across different groups.

How might data and ML model implementation cause disproportionate harm?
- **Equity** refers to the extent to which an ML model may disproportionately benefit or harm some individuals or groups more than others.
- **Representativeness** refers to whether the data used to develop AI/ML models are representative of the regions, communities, and contexts that will be affected by their use.
- **Bias** refers to systematically favoring or disfavoring different groups based on erroneous assumptions. Bias is defined in terms of attributes such as gender, economic standing, or ethnicity, among others. Consider different types of bias that may be present and how they affect the equity of ML/AI outcomes. Often bias will be embedded in data unintentionally as an artifact of the power dynamics that exist in the world.

How well do we understand how ML models work?
- **Explainability** refers to the extent to which individual predictions made by an ML model can be communicated in terms non-technical experts can understand.
- **Auditability** refers to the extent to which an AI/ML model's decision-making processes and recommendations can be queried by external actors or made transparent to a broader community of actors. Audits can sometimes help to identify concerns about equity, representativeness, and explainability.

What happens when things go wrong?
- **Accountability** refers to whether there are mechanisms in place to ensure that someone will be responsible for responding to feedback and redressing harms if necessary.

**How might you address some of these concerns?**
- Asking the right question
- Defining the protected attributes, making sure that data and outcomes are not correlated with them
- Identifying sources of bias (historical biases, individual biases, biases in data)
- Technical approaches to testing for bias in data
- Reviewing/strengthening data for representativeness
- Implementing fairness algorithmically
- Diversifying team of people working on AI/ML solutions
- Auditing model outcomes

**What are some of the other concerns?**
- What are the ethical concerns outside of fairness?
- Is AI/ML a good fit for my development problem?
- How is AI/ML better than existing approaches?
- How will the AI/ML approach align with organization structure and value?
- What other resources and capabilities may be needed to effectively implement AI/ML?

## 5.

### REPORT

NETHOPE