Use case name and description:

Consider: *How do algorithms account for geopolitical issues, socio-economic changes and infrastructure concerns in a rapidly changing landscape? Which groups might be affected by system decisions? Which fairness notion and metric can support monitoring the impact of decisions effectively? How can we achieve fairness and eventually trust with transparency?*
What should be the level of transparency?

Potential harms: *Describe potential harms of "unfair" state of the system*

Disparate Impact: *Unintentionally disadvantage members of certain groups*
Disparate Treatment: *Explicitly treat certain groups differently, even developer's intention is not.*
Loss of Rights or Agency: *It includes invading people's privacy, or other human rights.*

Allocational Harms: *Unequal distribution of resources or opportunities, such as jobs, loans, and healthcare.*
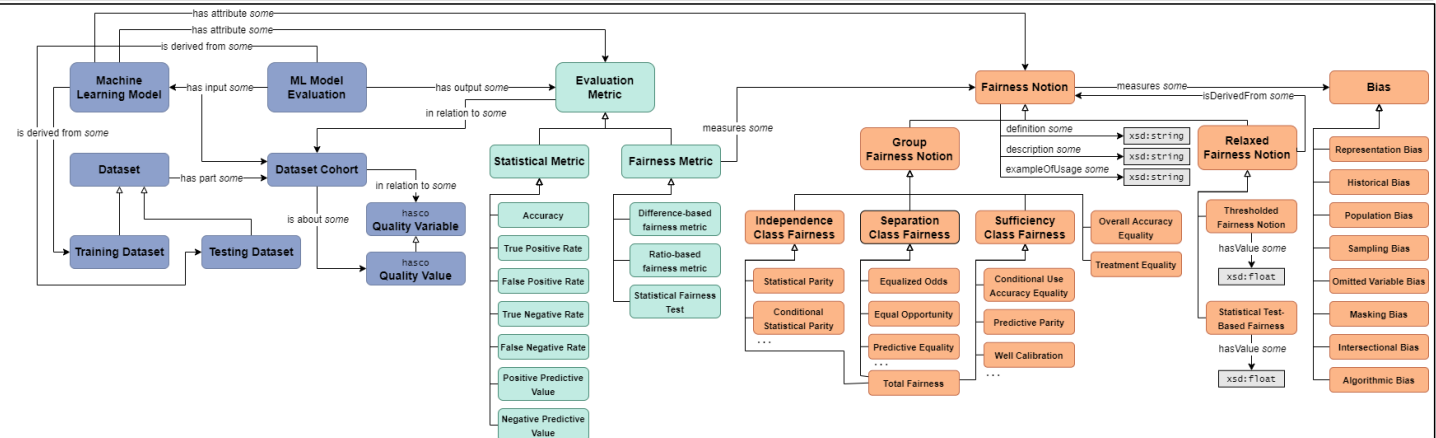*Quality of Service Harms:* *Disparities in the quality of service that people receive.*
Representational Harms: *Perpetuating or amplifying harmful stereotypes about certain groups of people.*
Harms Affecting People's Well-being: *It includes physical, mental, or emotional well-being.*

Fairness Evaluation and Mitigation Stages:
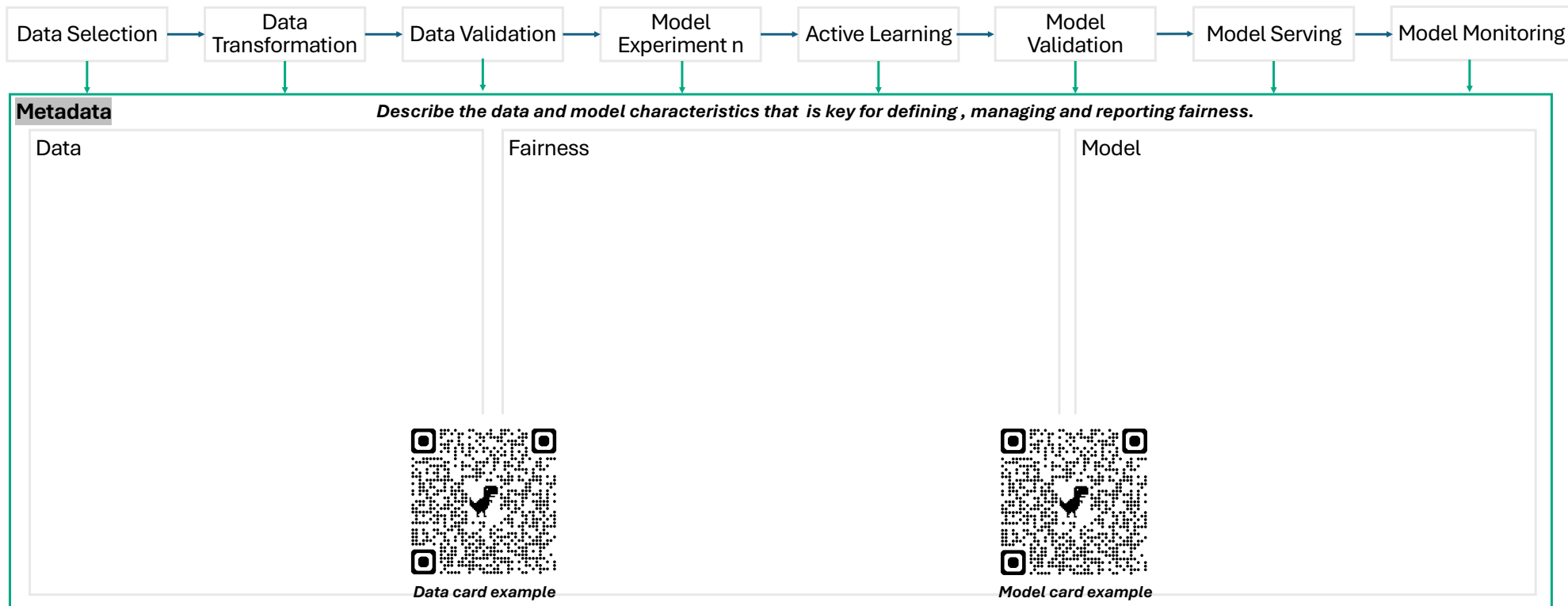*Select the right metrics and techniques to evaluate and mitigate biases*

- *Evaluation techniques*
- *Potential mitigation approaches*
- *Unit tests for continuous monitoring (e.g. ratio of GS/GN countries ~= 1, len(potential bias-indicator words for each country) > 0, etc.)*

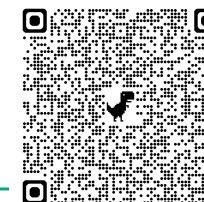**Due to harms:** _____, **we selected fairness notion:** _____. **We can monitor this notion with the fairness metric:** _____.

Remastering the model towards this measure will ensure that _____.

Now, we will define the required (1) data and (2) model characteristics to monitor all aspects related to fairness.

Data Selection → Data Transformation → Data Validation → Model Experiment n → Active Learning → Model Validation → Model Serving → Model Monitoring

Metadata — *Describe the data and model characteristics that is key for defining , managing and reporting fairness.*

Data

Fairness

Data card example

Model

Model card example

Reporting

Algorithmic Transparency Record Format

Use case name and description:

Potential harms:

Describe potential harms of "unfair" state of the system

Dataset Characteristics:

Describe the data characteristics that is key for defining reporting fairness

Model Characteristics:

Fairness Evaluation and Mitigation Stages:

Select the right metrics and techniques to evaluate and mitigate biases

Fairness Overview

Reporting

Transparency Report