

---

# USAID APPROPRIATE USE FRAMEWORK

---

Exploring Fairness in Machine Learning

**Amit Gandhi**

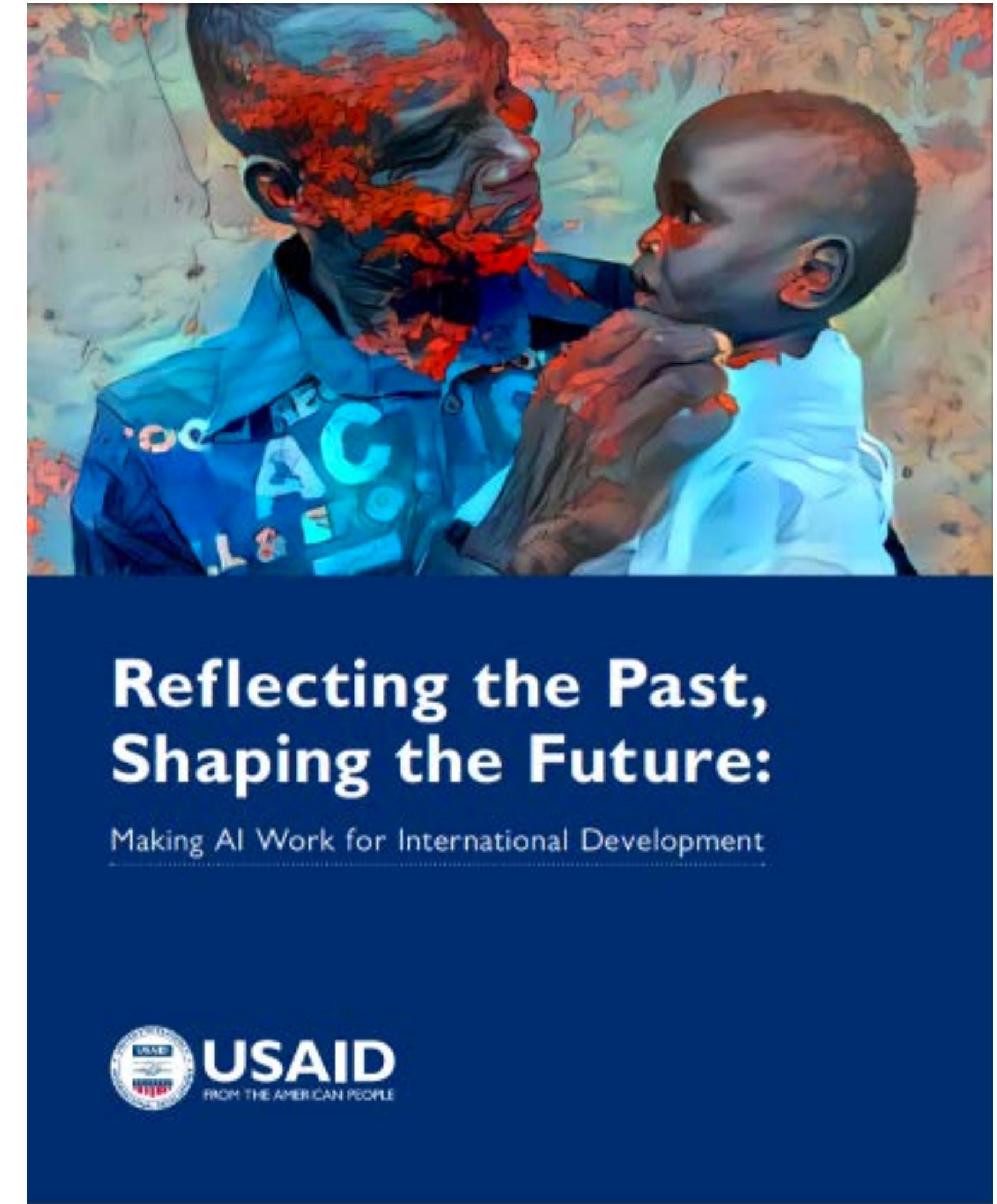
Graduate Researcher, MIT

# Appropriate Use Framework

## Guiding Principles

- Relevance
- Representativeness
- Value
- Explainability
- Auditability
- Equity
- Accountability/Responsibility

US Government work.  
Image is in the public domain.



# Relevance

Is the use of ML in this context solving an appropriate problem?



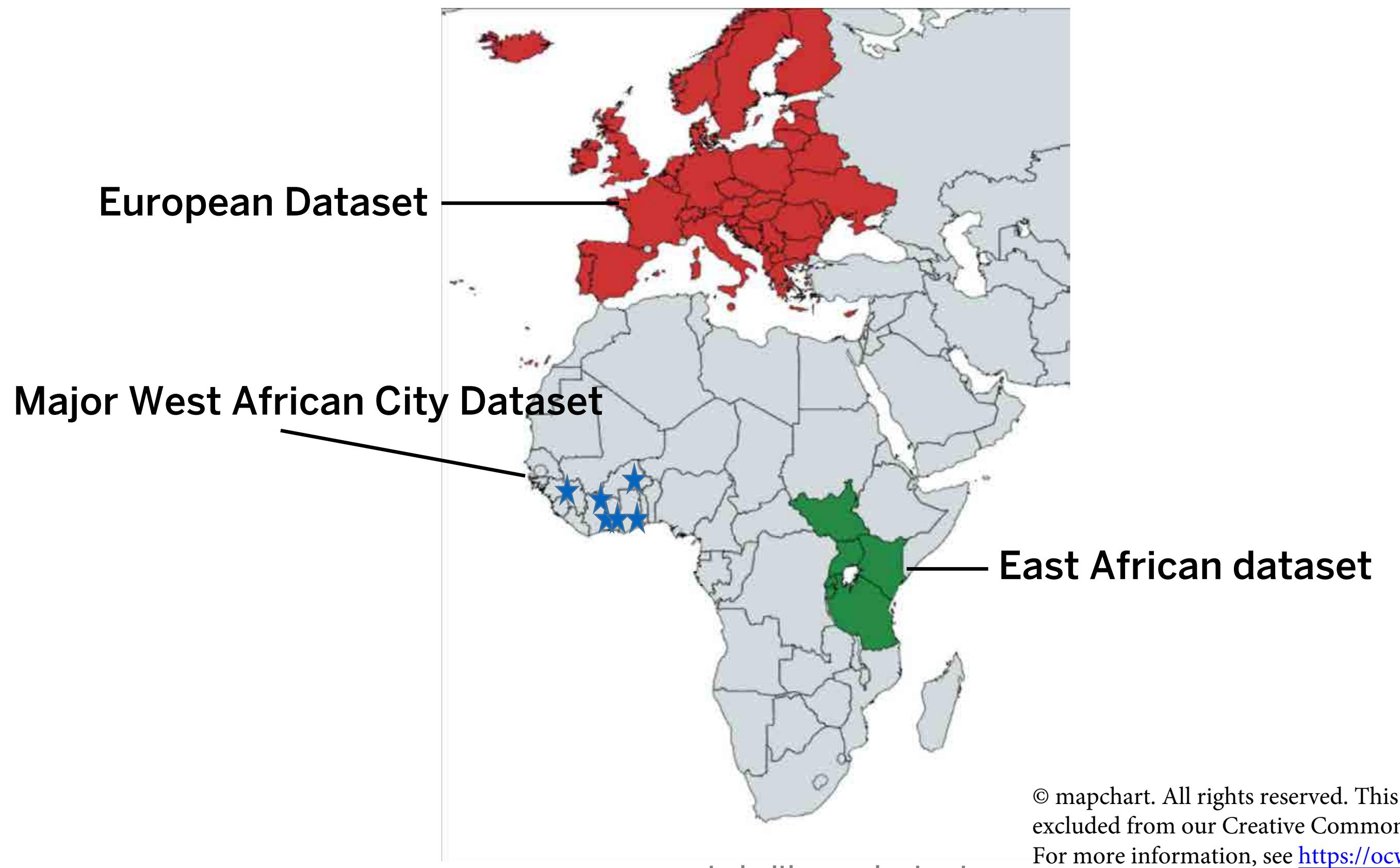
US Government work.  
Image is in the public domain.

# Representativeness

Is the data used to train the ML models appropriately selected?



Courtesy of CDC/ James Gathany. Image is in the public domain.



© mapchart. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

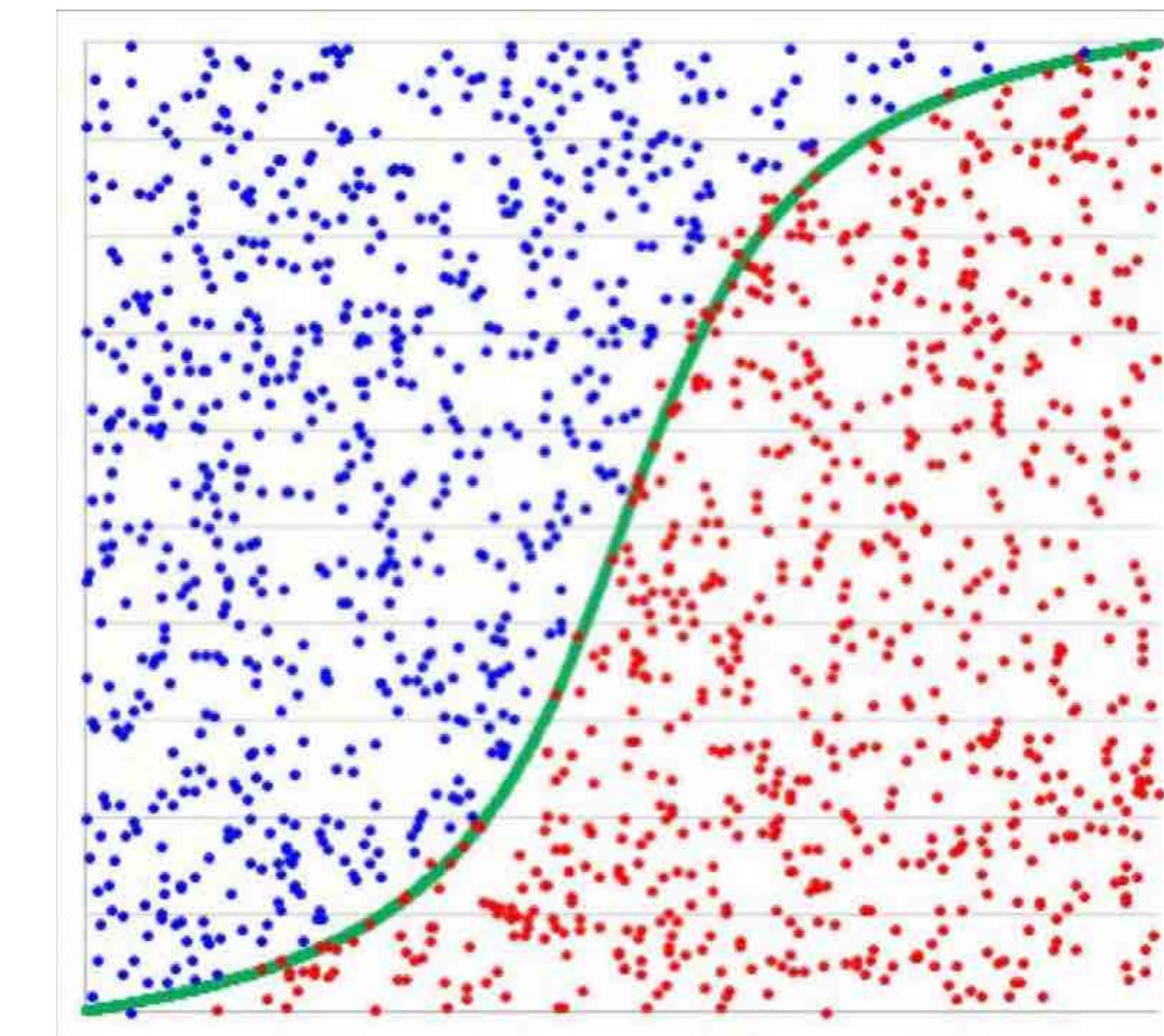
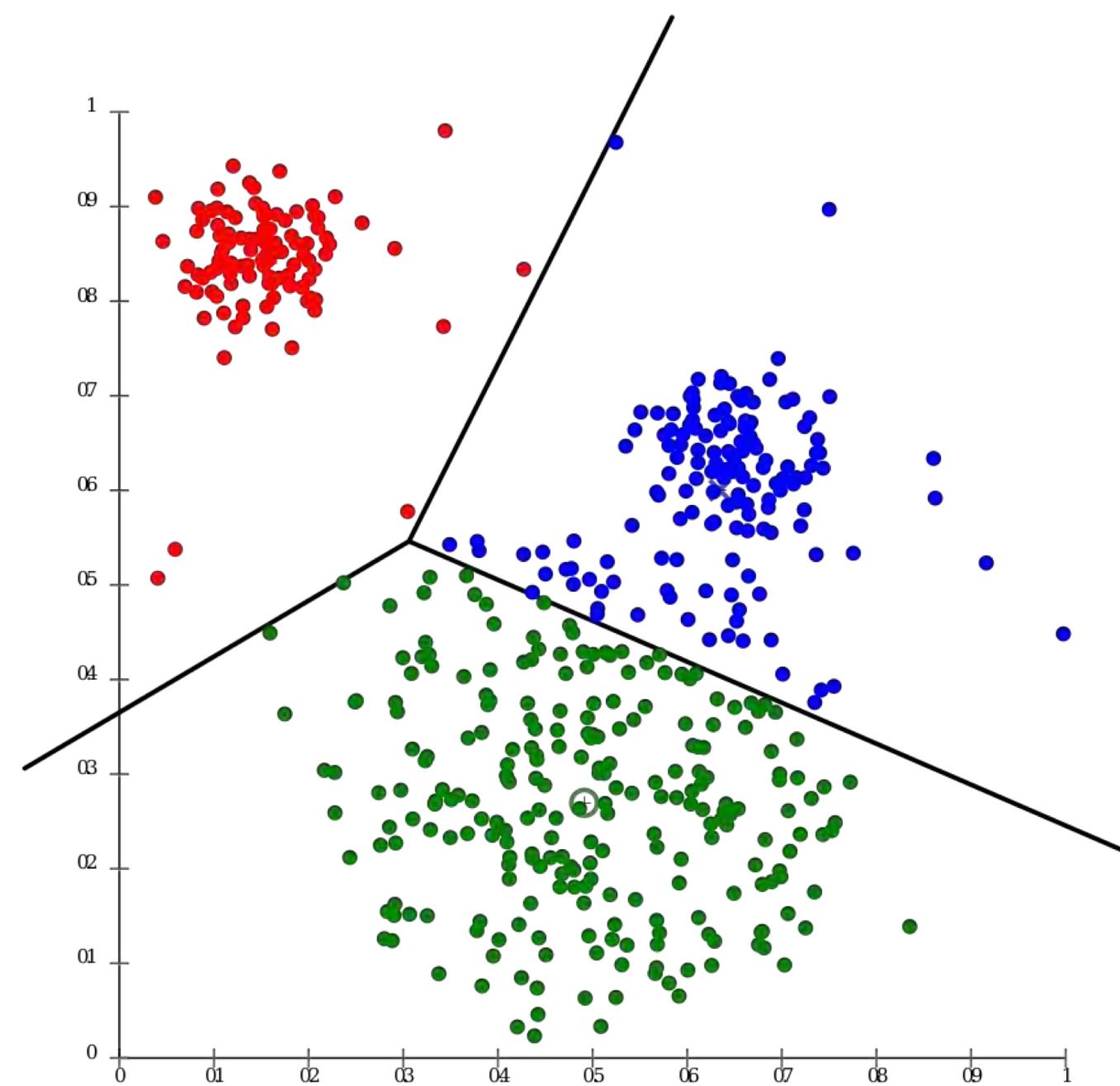
---

# Value

- Does the machine learning model produce predictions that are more accurate than alternative methods?
- Do the predicted values inform human decisions in a meaningful way?
- Does it explain variation more completely than alternative models?

# Explainability

How effectively is the use of machine learning communicated?  
Including: data selection, model selection, (un)intended consequences



© sources unknown. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

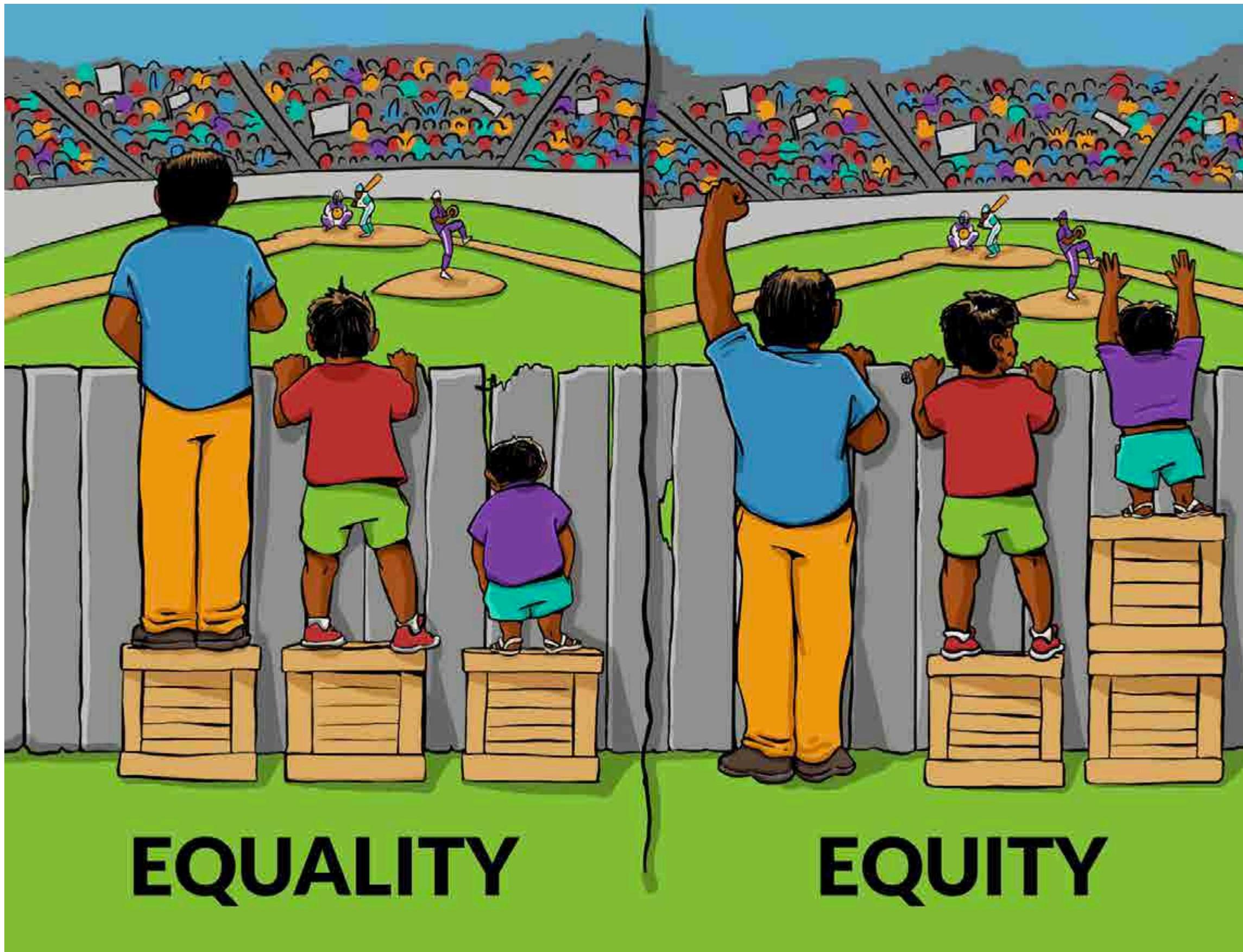
# Auditability

Can the model's decision-making processes be queried or monitored by external actors?

- Ensure model is fair, unbiased, and does not harm users
- Models are often hard to understand
- May require institutional or legal infrastructure to require audits, provide access to proprietary data, and act on findings

# Equity

Has the ML model been tested to determine whether it disproportionately benefits or harms some individuals or groups more than others?



© Interaction Institute for Social Change, Angus Maguire. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>

# Accountability/Responsibility

Are there mechanisms in place to ensure that someone will be responsible for responding to feedback and redressing harms, if necessary?



---

# Key Takeaways

## Guiding Principles

- Relevance
- Representativeness
- Value
- Explainability
- Auditability
- Equity
- Accountability/Responsibility

---

# Thank you

**Amit Gandhi**  
Graduate Researcher, MIT

[amitg@mit.edu](mailto:amitg@mit.edu)

MIT OpenCourseWare  
<https://ocw.mit.edu>

RES. EC-001 Exploring Fairness in Machine Learning  
Spring 2019

For information about citing these materials or our Terms of Use, visit: <https://ocw.mit.edu/terms>.