

# SOCIAL COMPUTING: BRIDGING THE GAP BETWEEN TECHNOLOGY AND HUMAN INTERACTION

DR. NICOLÒ PAGAN, POST-DOCTORAL  
RESEARCHER, UNIVERSITY OF ZURICH

DSI PROFESSOR PRESENTATIONS: ZÜRICH, FEBRUARY 28, 2023

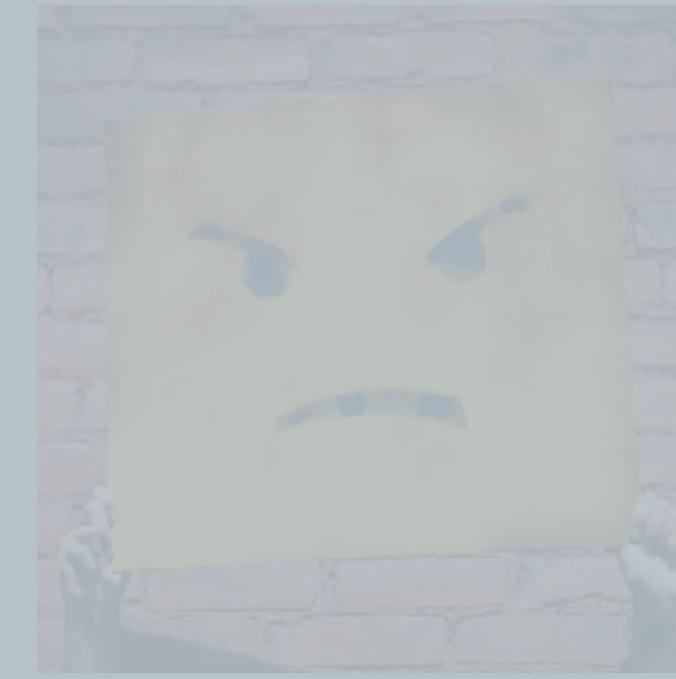
# INDEX



1. CONTENT CREATORS  
IN SOCIAL MEDIA  
PLATFORMS



2. FEEDBACK LOOPS IN  
ML-BASED AUTOMATED  
DECISION-MAKING  
SYSTEMS



3. POLARISATION  
AND HATE SPEECH



4. BIAS IN  
SEARCH ENGINES

## 2. FEED-BACK LOOPS IN ML-BASED AUTOMATED DECISION-MAKING SYSTEMS

### MOTIVATION:

- ▶ Machine-Learning (ML)-based Decision-Making systems are widely employed in various domains, including recommendation systems, credit lending, hiring, and predictive policing.
- ▶ These systems appear to be vulnerable to **RUNAWAY FEEDBACK LOOPS** e.g., when police are repeatedly sent back to the same neighborhoods regardless of the actual rate of criminal activity, which exacerbate existing **BIASES**.

## 2. FEED-BACK LOOPS IN ML-BASED AUTOMATED DECISION-MAKING SYSTEMS

### MOTIVATION:

- ▶ Machine-Learning (ML)-based Decision-Making systems are widely employed in various domains, including recommendation systems, credit lending, hiring, and predictive policing.
- ▶ These systems appear to be vulnerable to **RUNAWAY FEEDBACK LOOPS** e.g., when police are repeatedly sent back to the same neighborhoods regardless of the actual rate of criminal activity, which exacerbate existing **BIASES**.

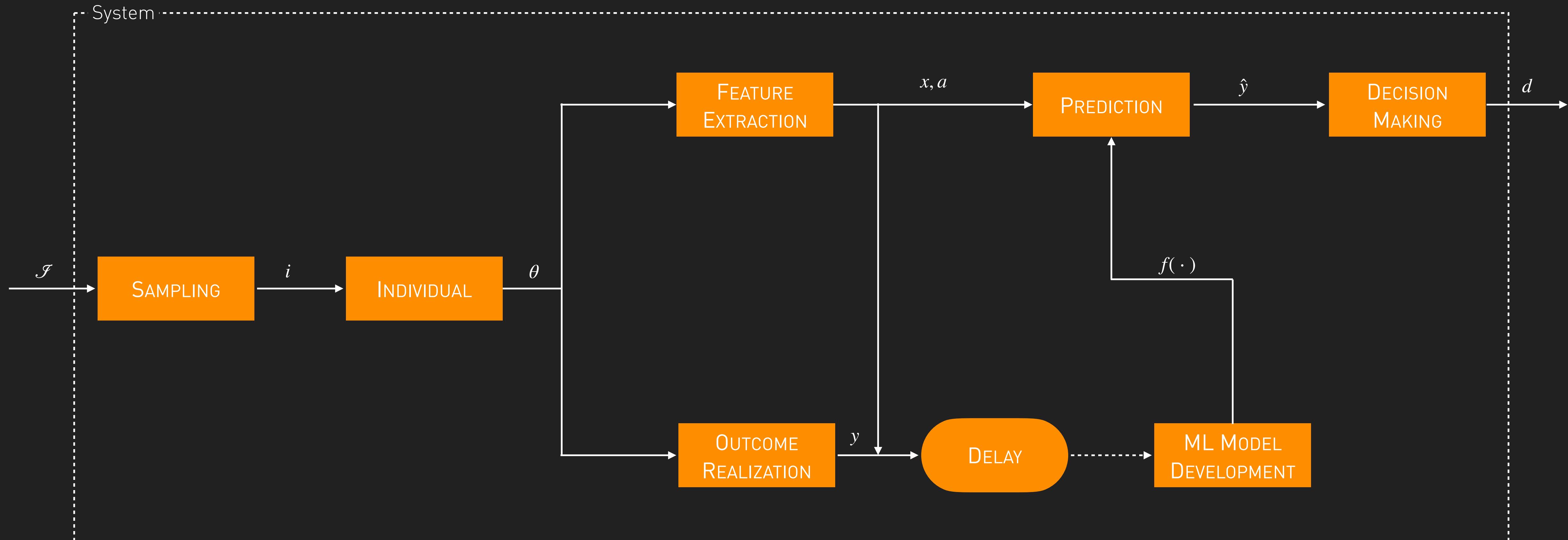
### RESEARCH QUESTIONS:

- ▶ What is the role of these feedback loops, and how do they impact the ML decision-making pipeline?
- ▶ How are they linked to the different types of biases?

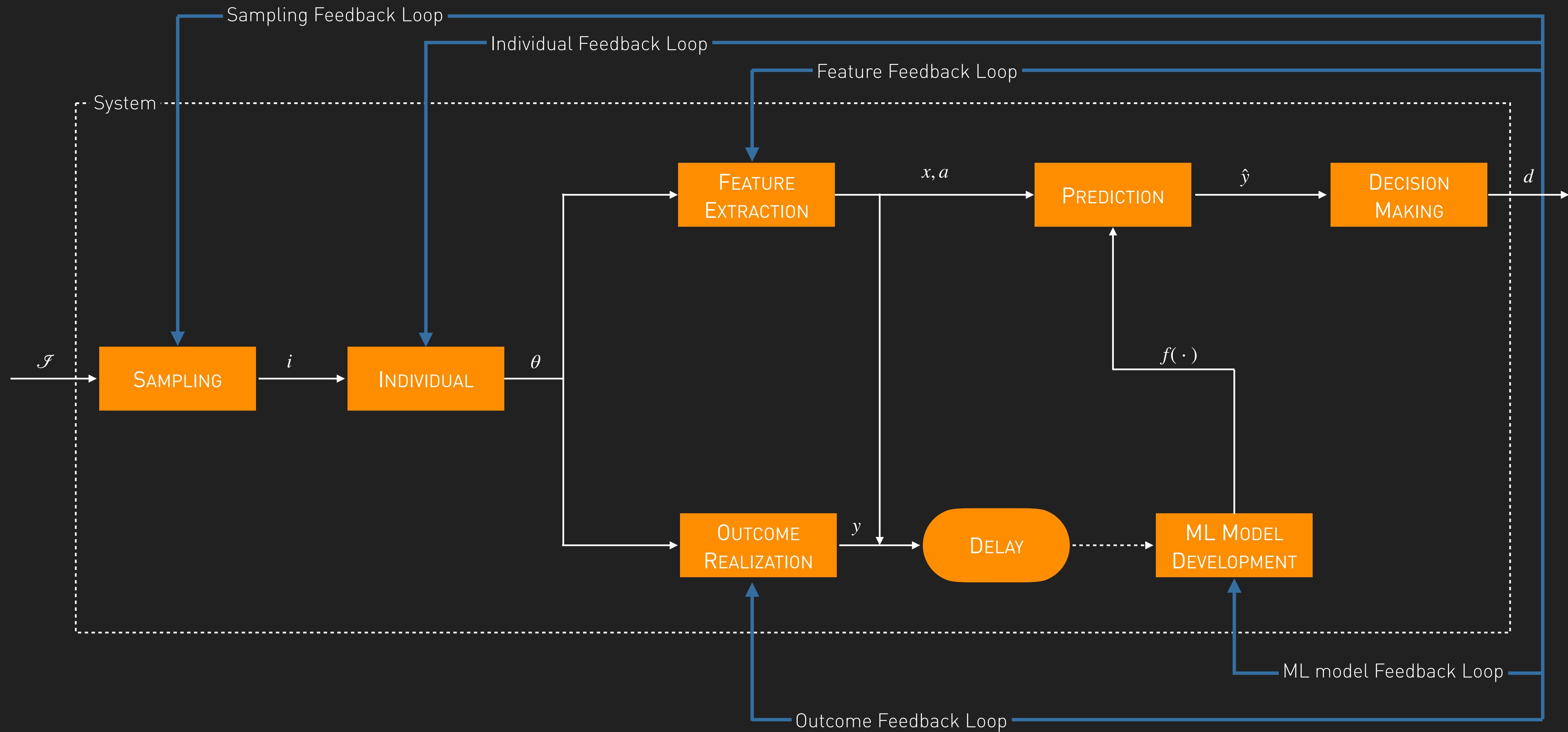
### METHODS:

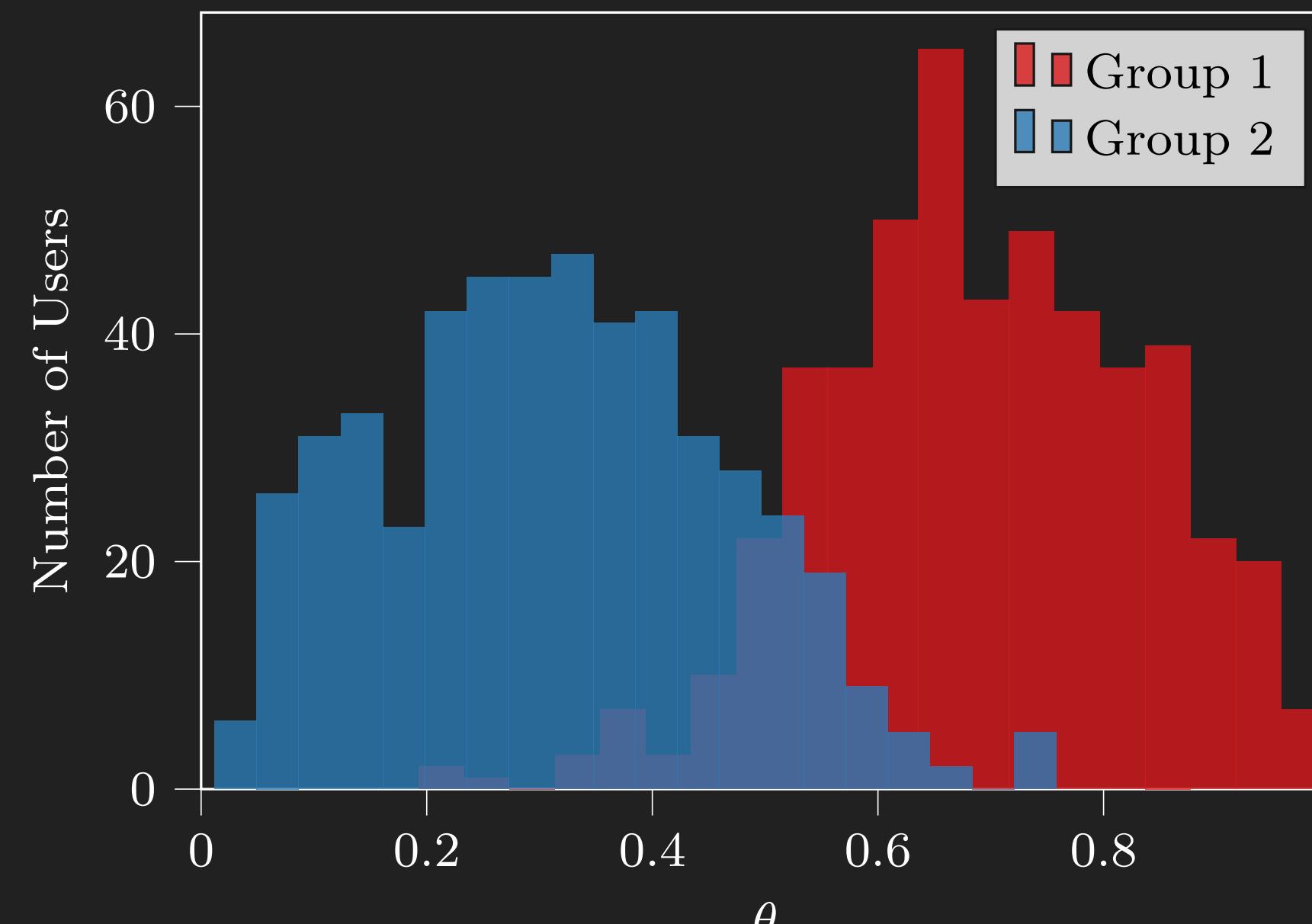
- ▶ Dynamical Systems Theory, Algorithmic Fairness

# THE ML-BASED DECISION-MAKING PIPELINE

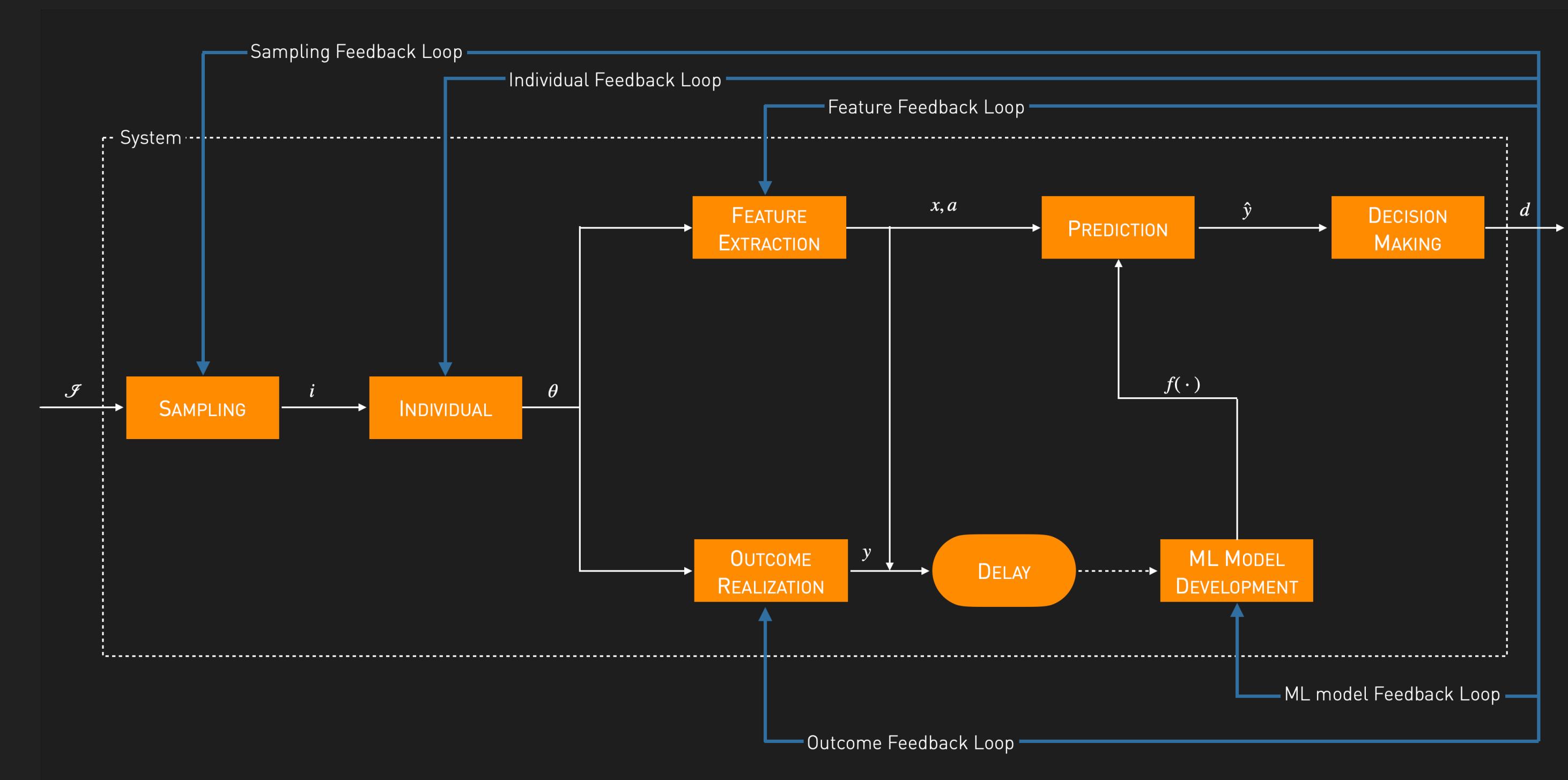


# A CLASSIFICATION FRAMEWORK



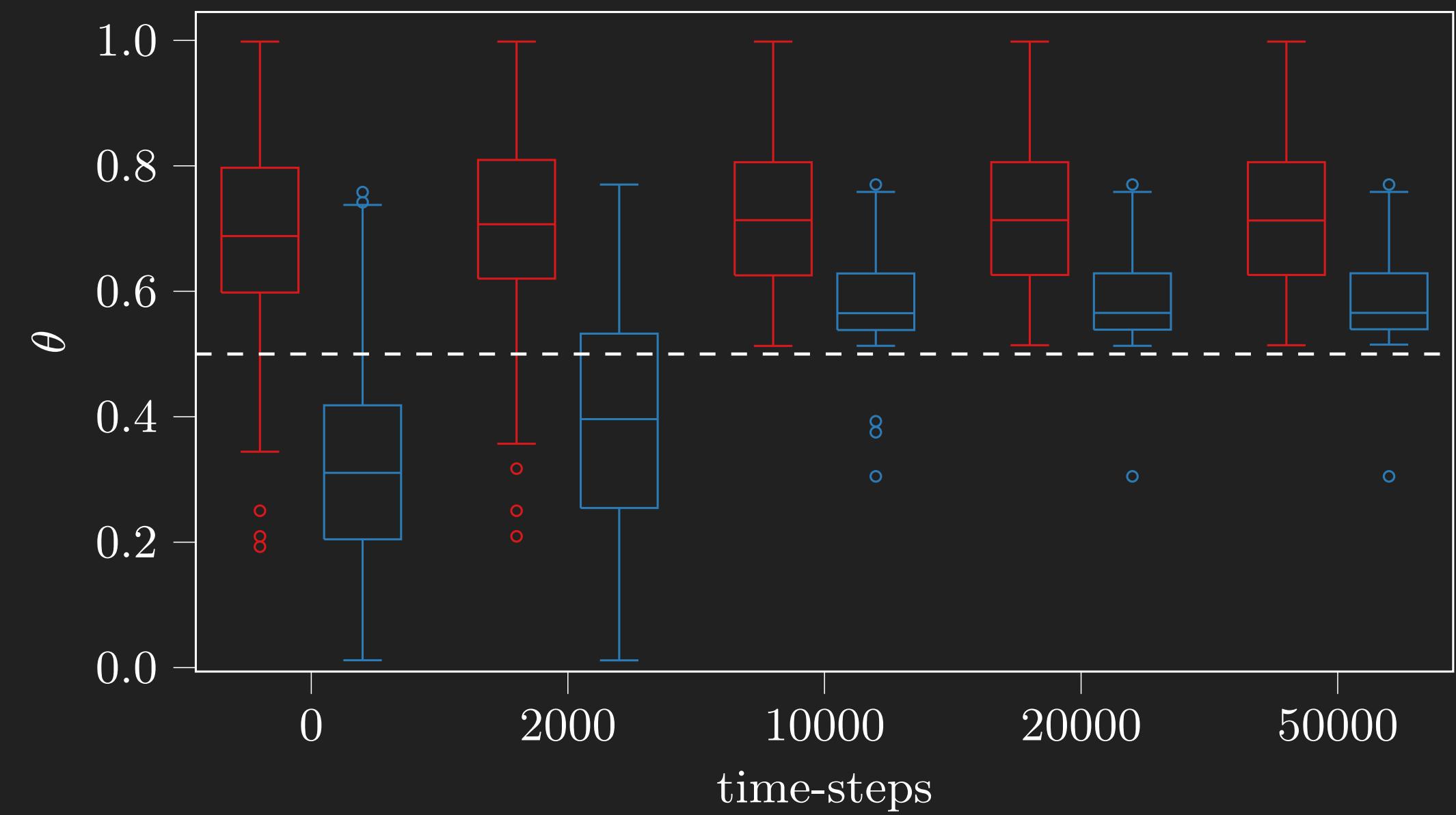
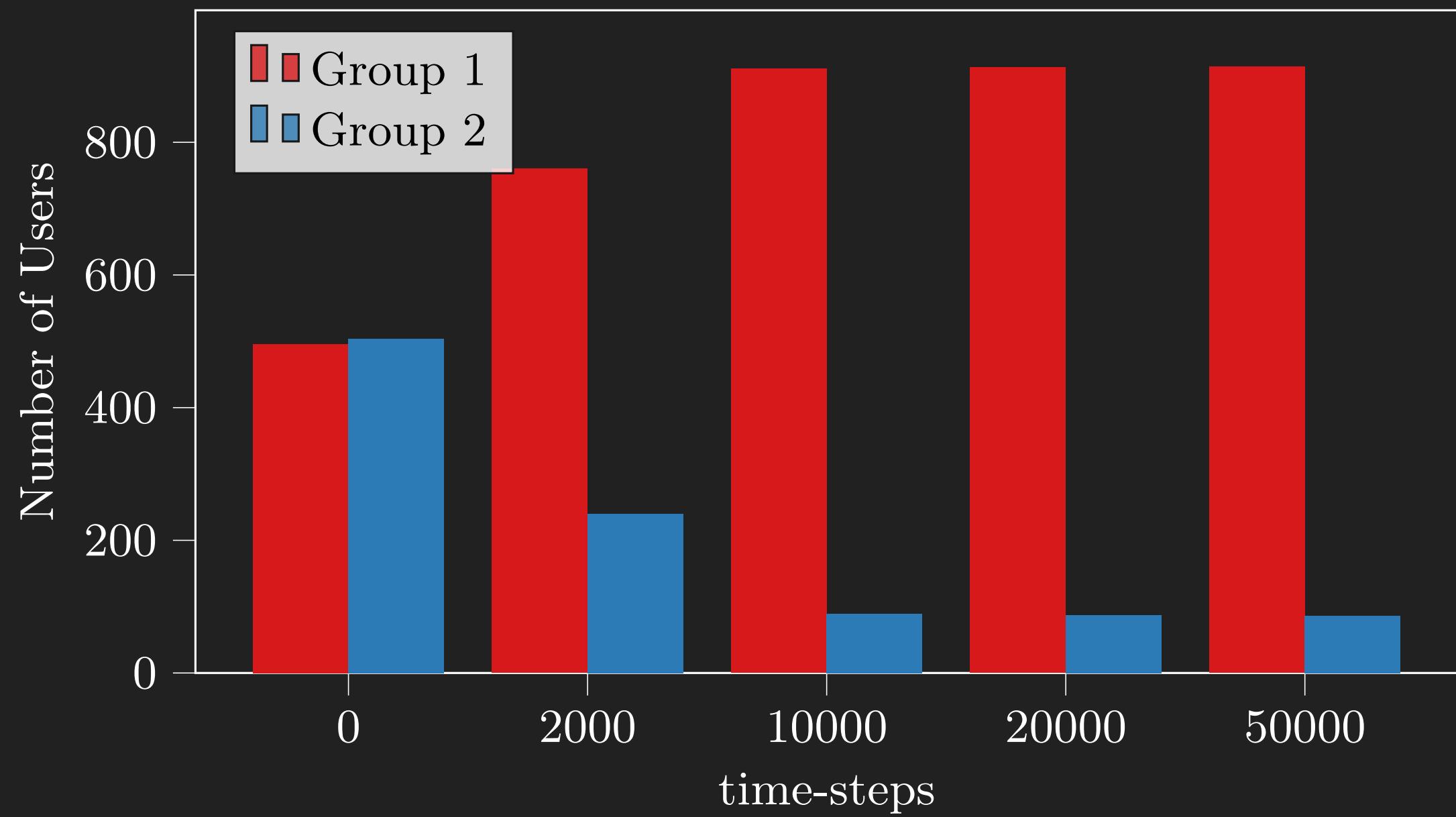


Initial Distribution



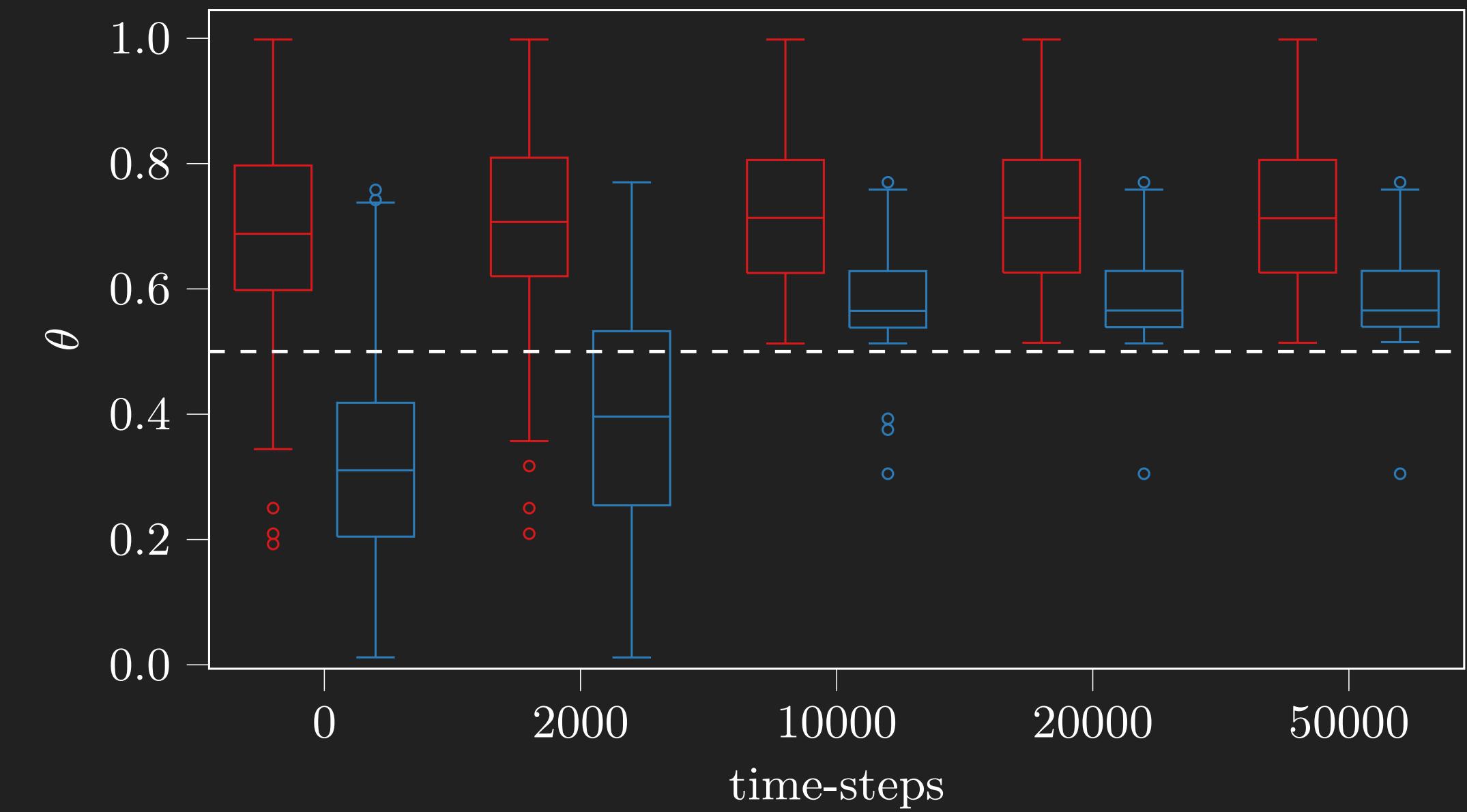
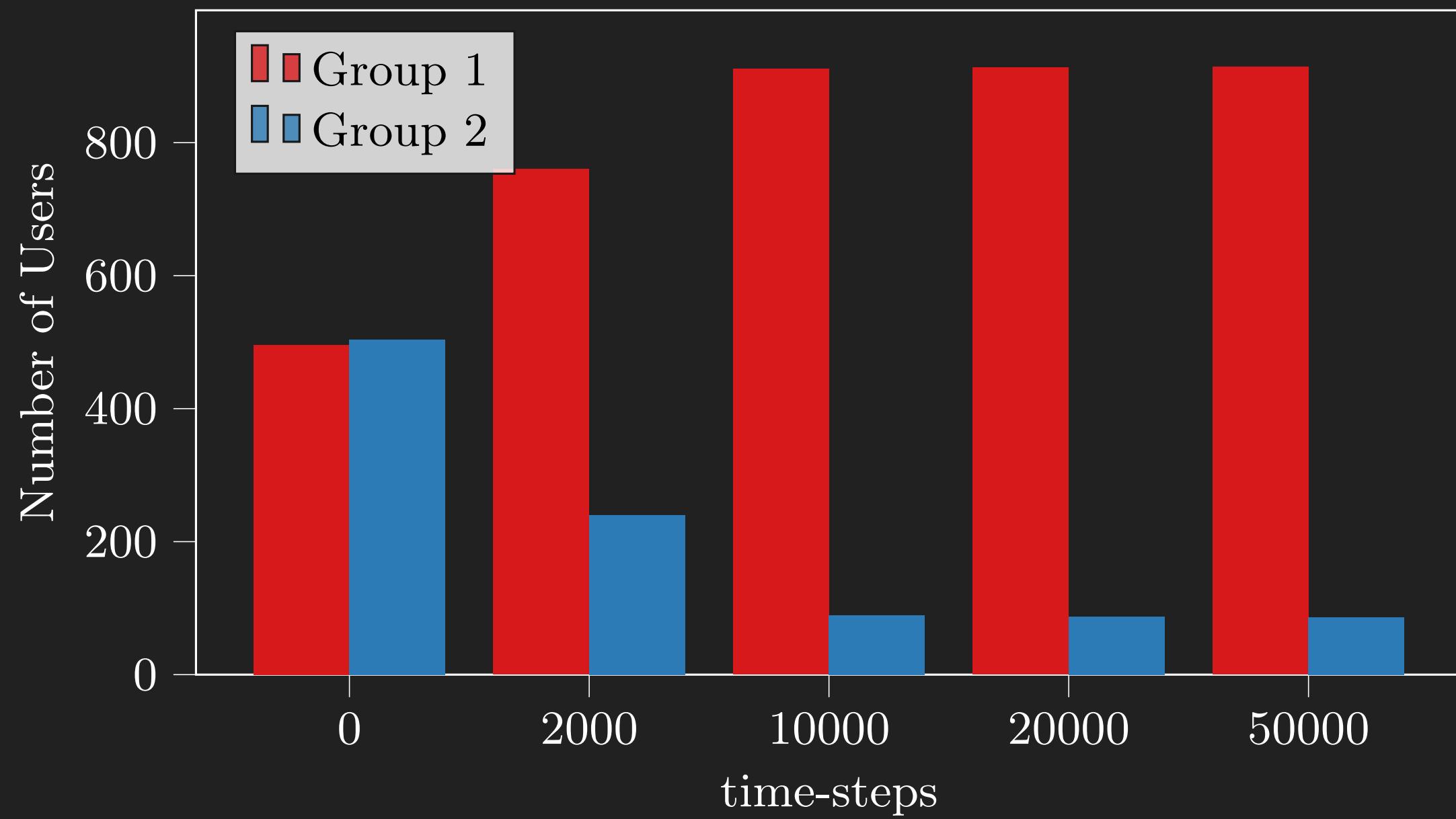
## 1. SAMPLING FEEDBACK LOOP:

The platform is right-skewed and users that do not see content aligned with their view leave the platform)

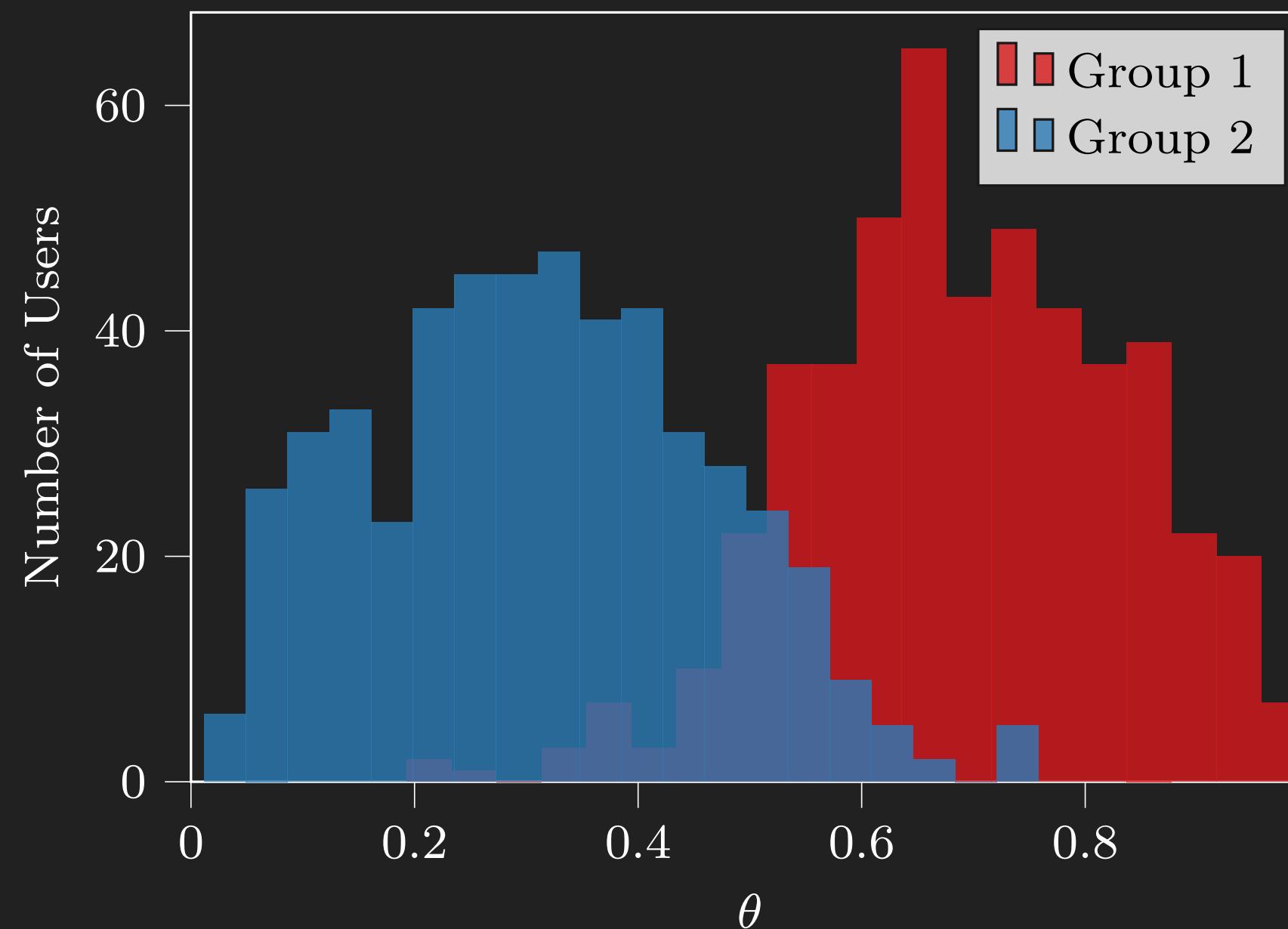


## 1. SAMPLING FEEDBACK LOOP:

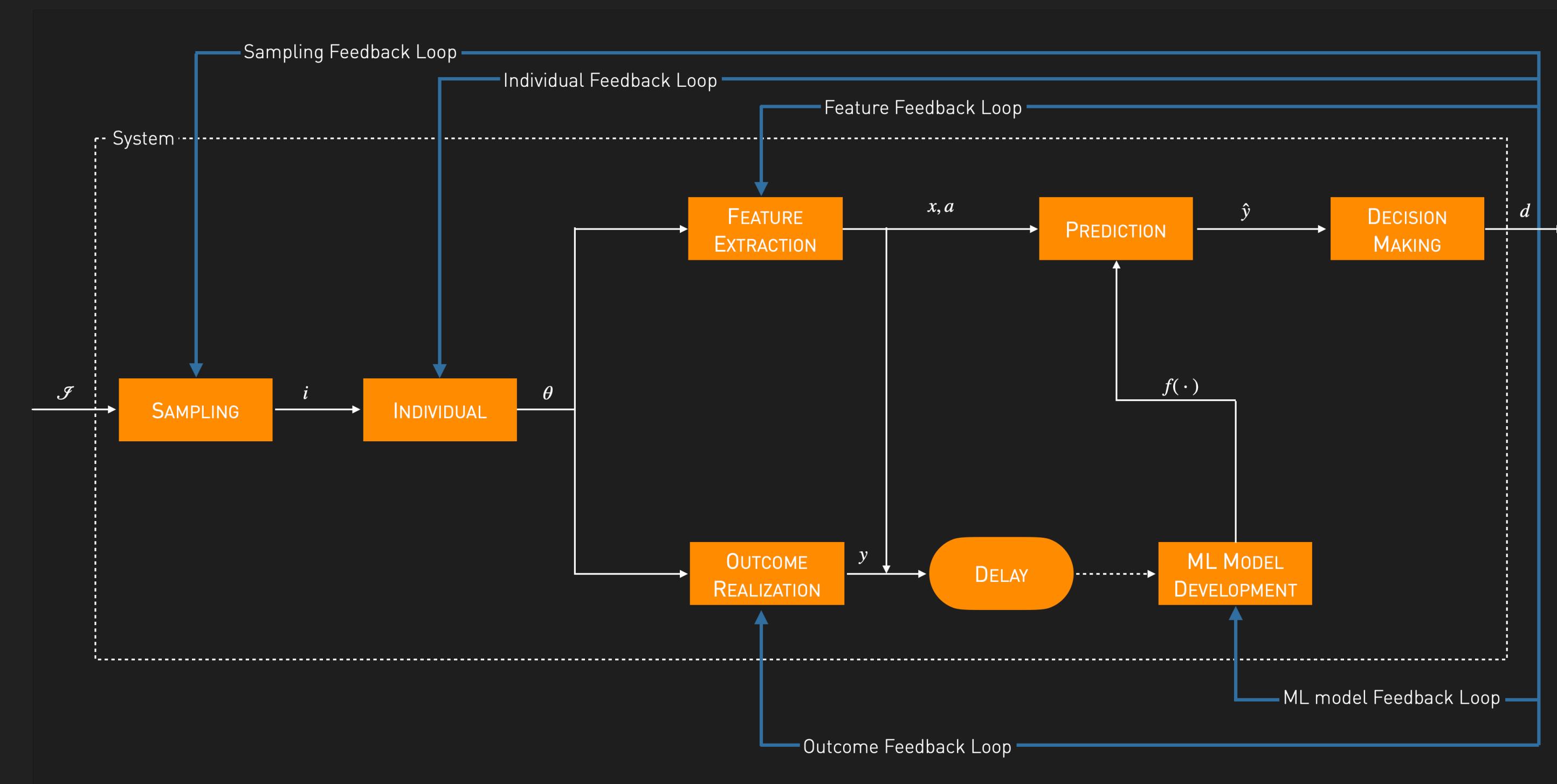
The platform is right-skewed and users that do not see content aligned with their view leave the platform)



INCREASES THE  
REPRESENTATION BIAS

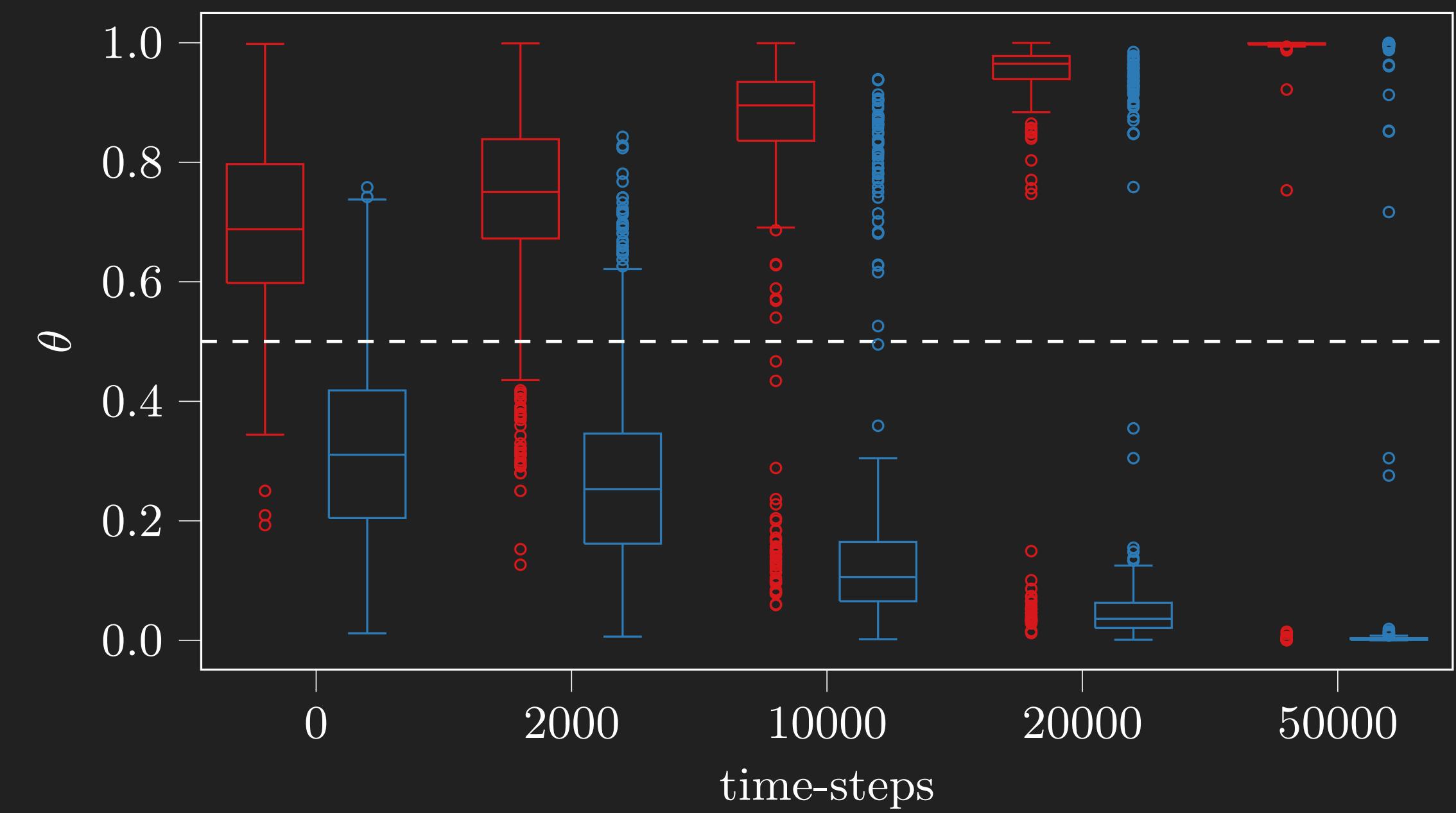
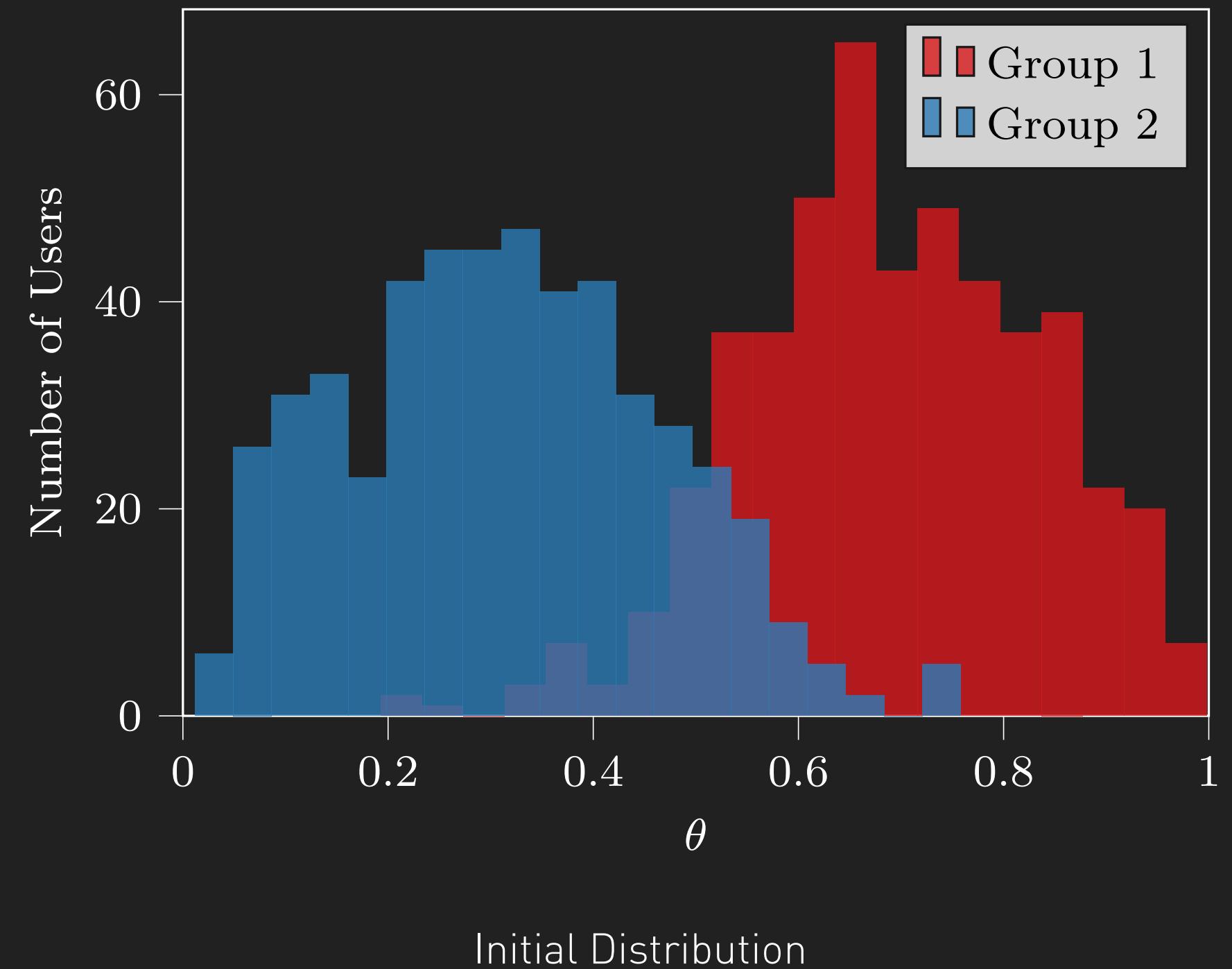


## Initial Distribution



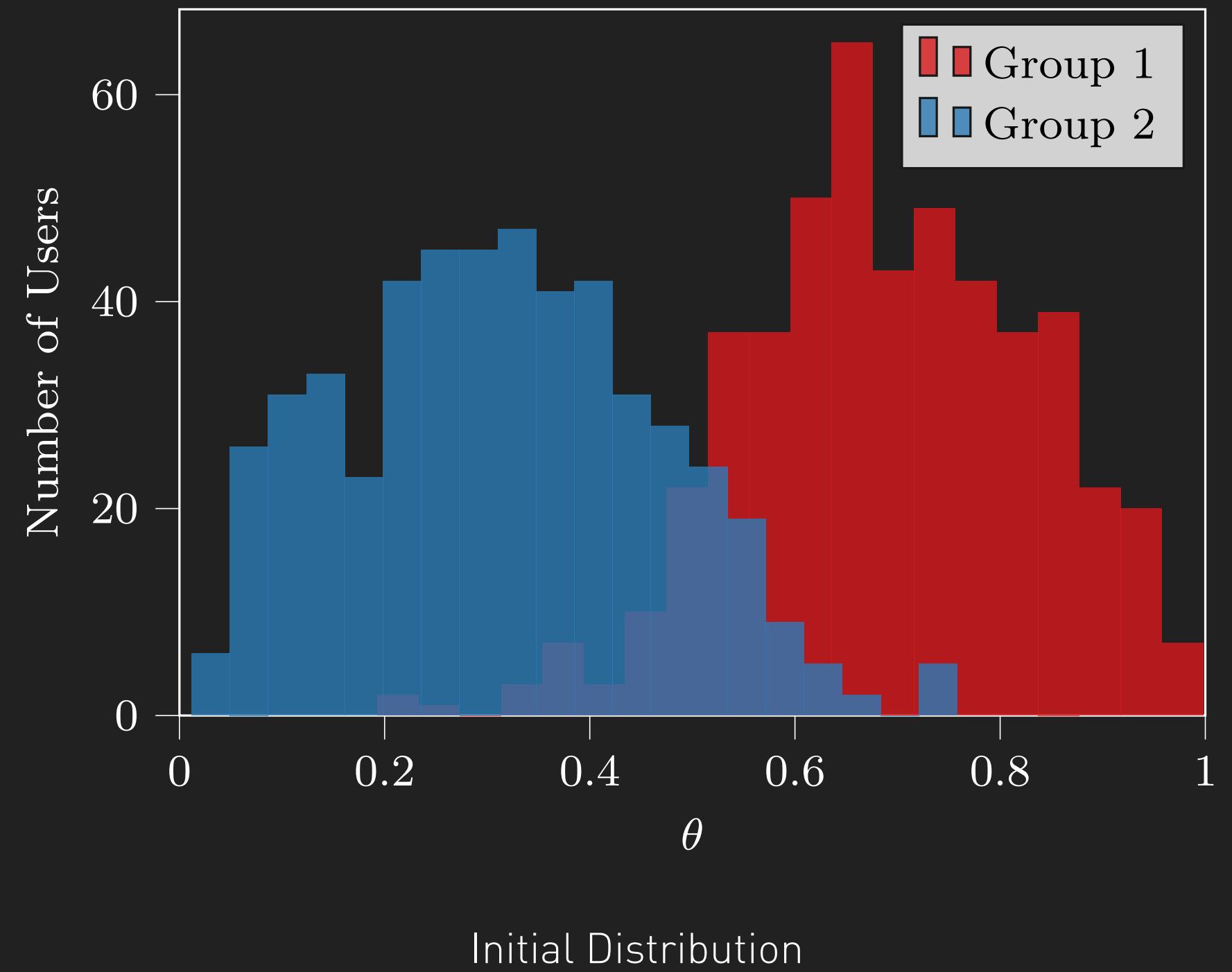
## 2. INDIVIDUAL FEEDBACK LOOP:

The platform recommended content  
reinforces user's opinion

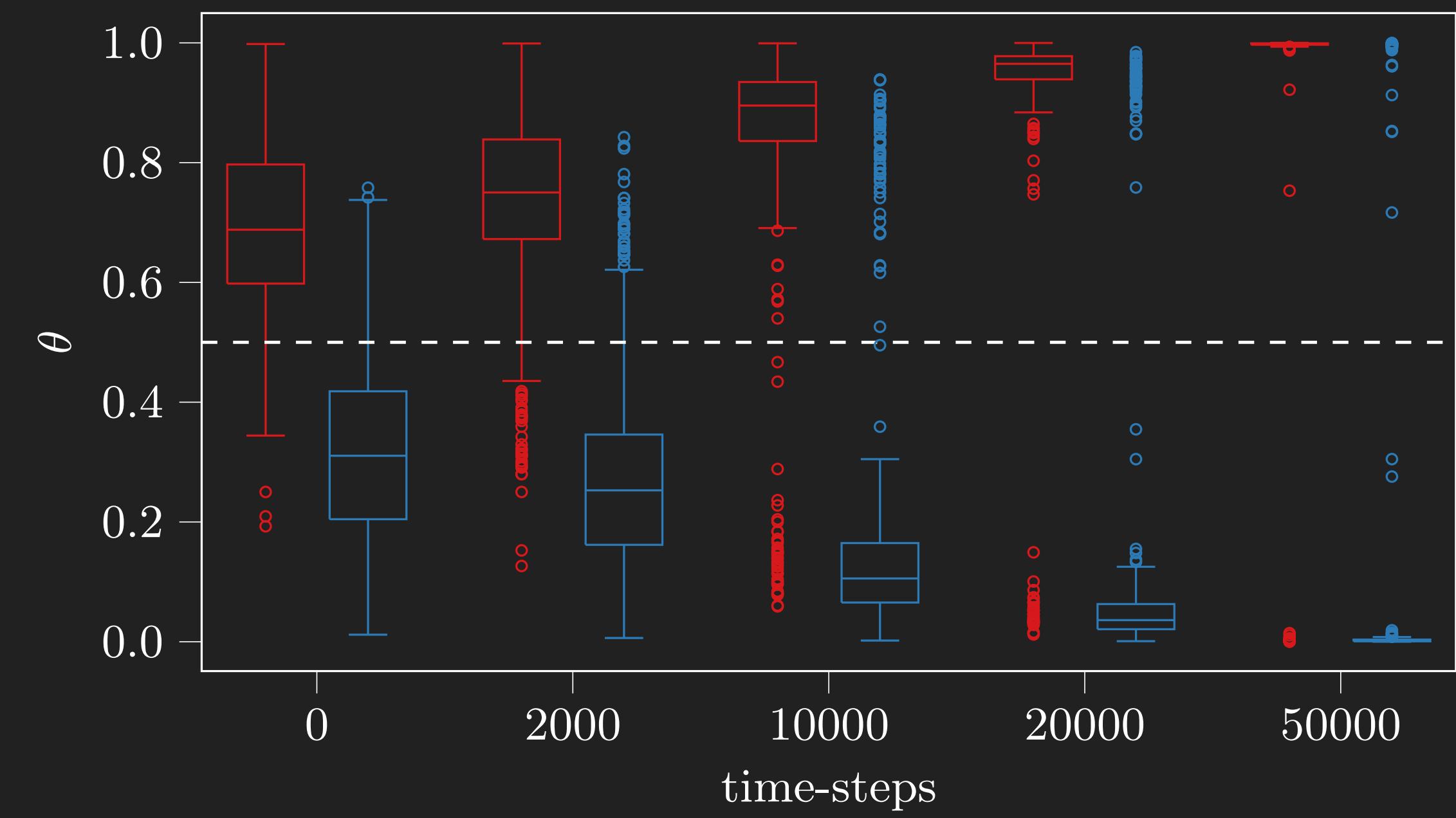


## 2. INDIVIDUAL FEEDBACK LOOP:

The platform recommended content  
reinforces user's opinion



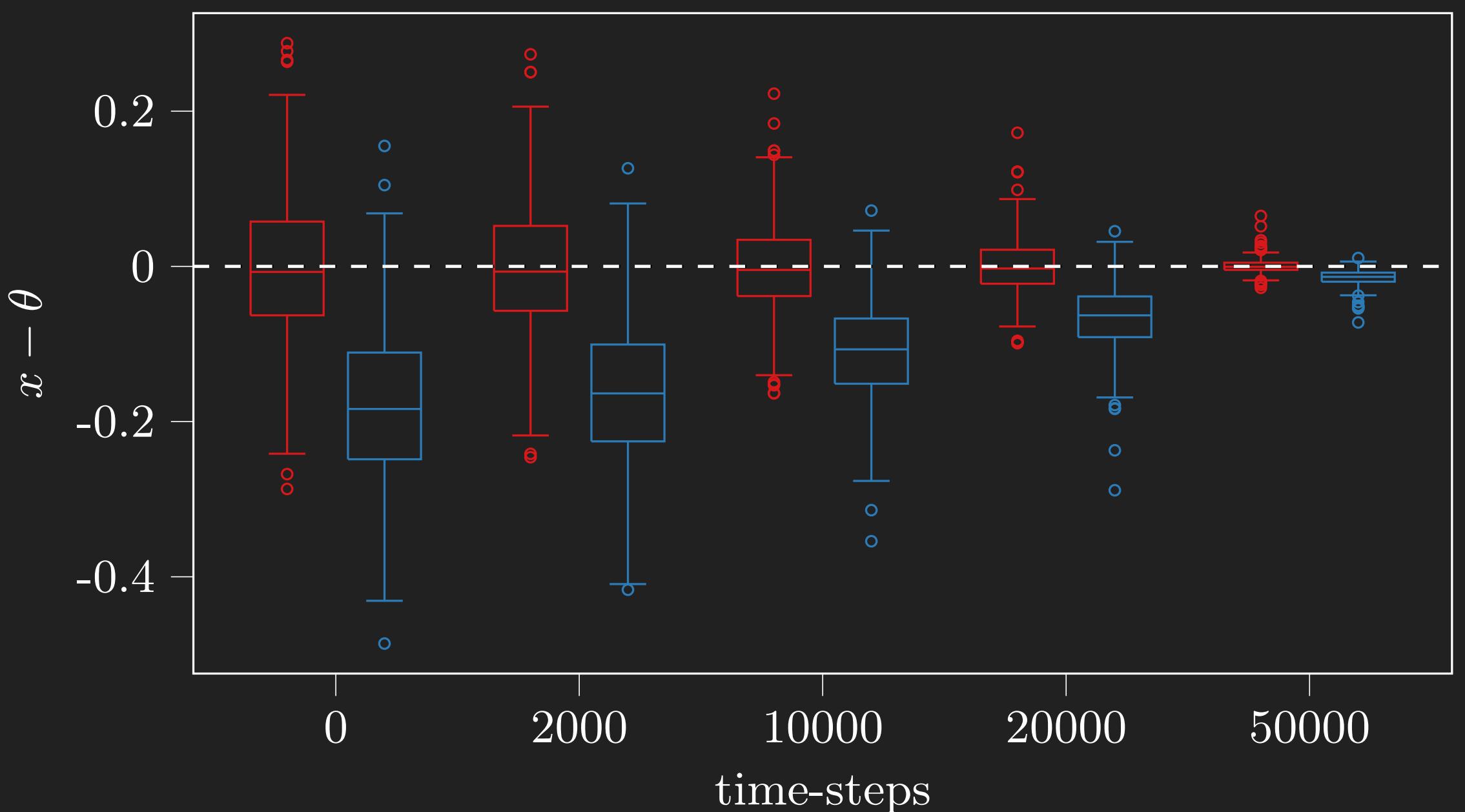
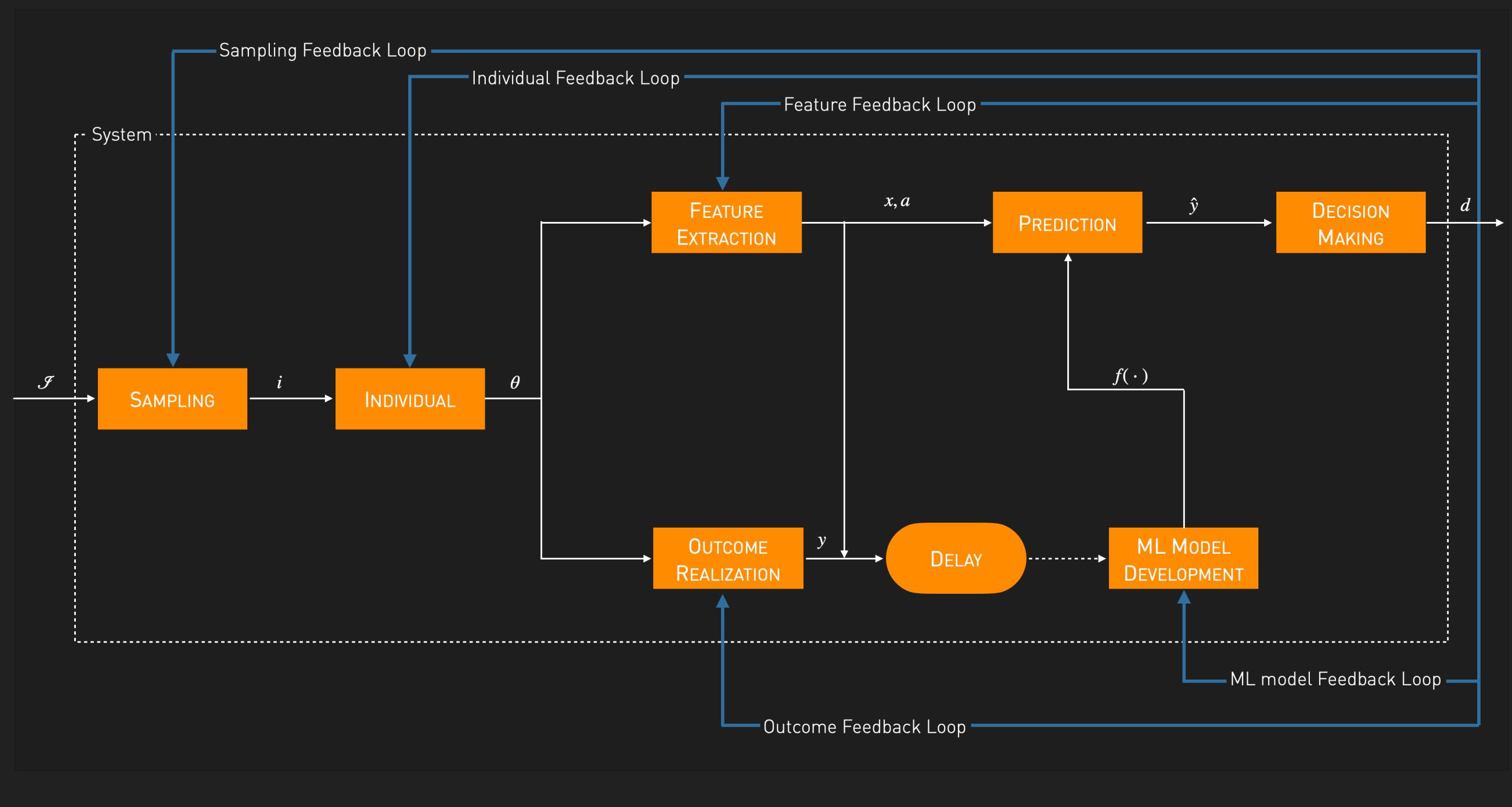
Initial Distribution



INCREASES THE  
LIFE BIAS

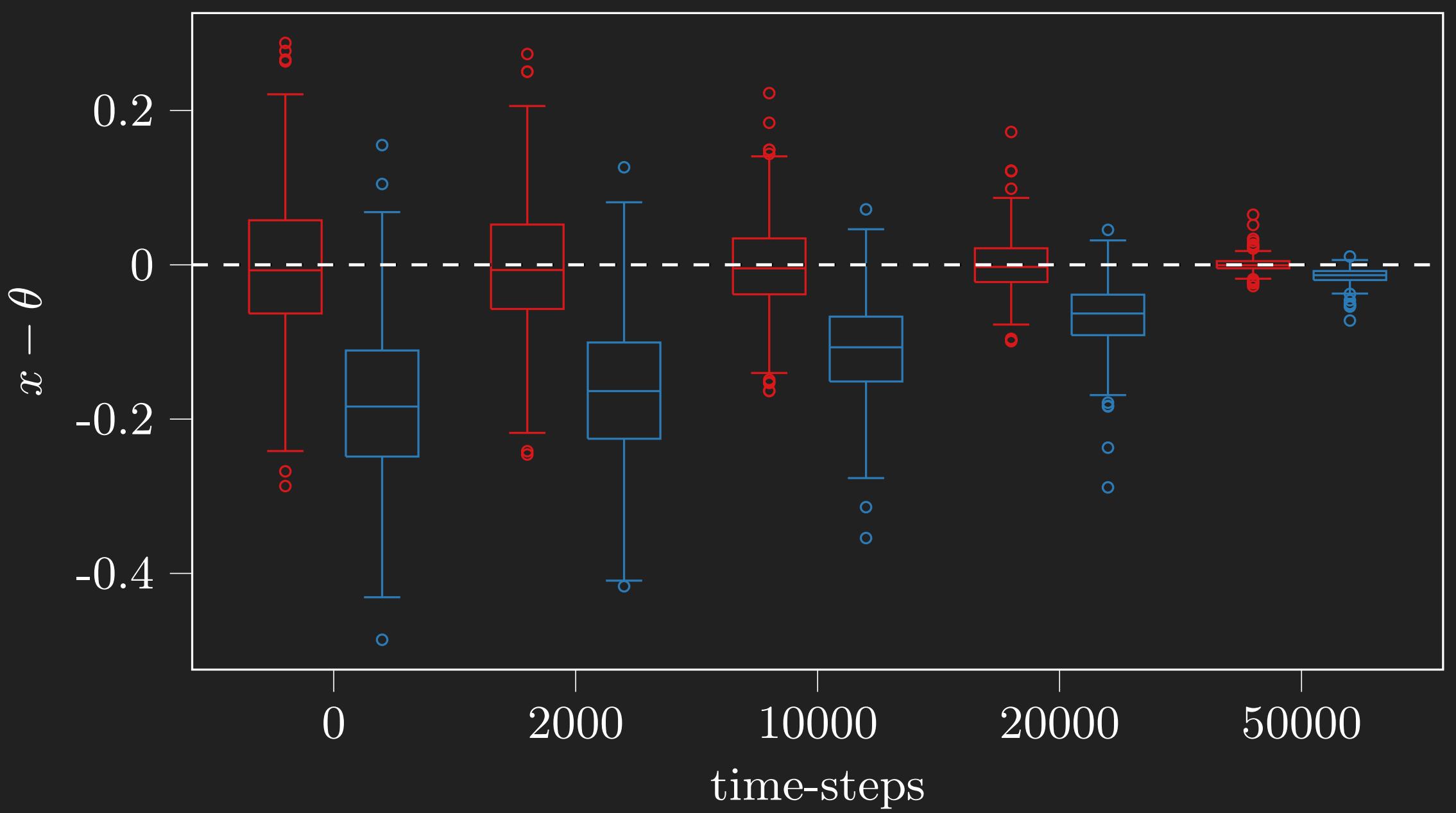
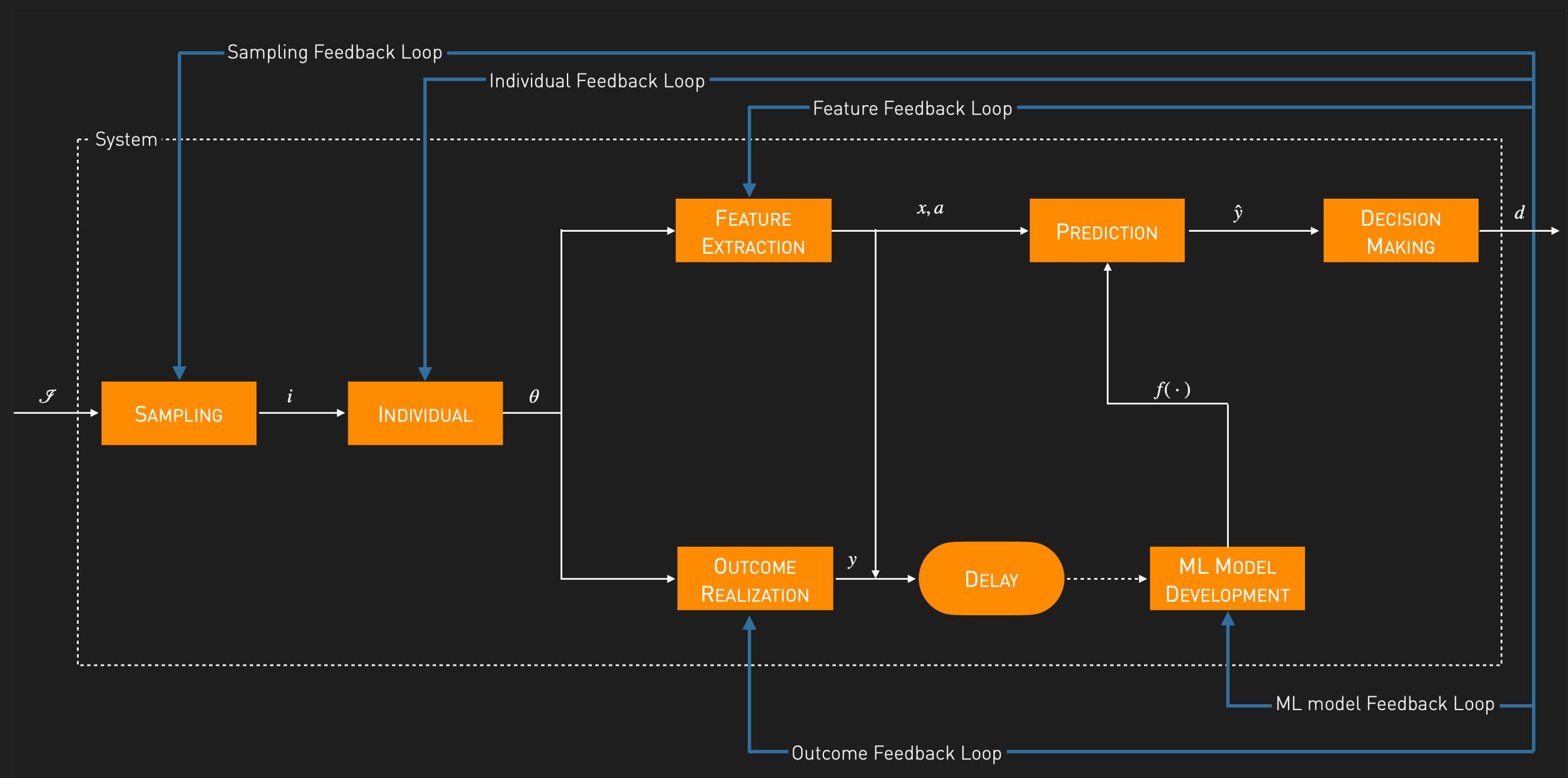
### 3. FEATURE FEEDBACK LOOP:

The platform recommended content informs on the correct features



### 3. FEATURE FEEDBACK LOOP:

The platform recommended content informs on the correct features



REDUCES THE  
MEASUREMENT BIAS

# INDEX



1. CONTENT CREATORS  
IN SOCIAL MEDIA  
PLATFORMS



2. FEEDBACK LOOPS IN  
ML-BASED AUTOMATED  
DECISION-MAKING  
SYSTEMS



3. POLARISATION  
AND HATE SPEECH



4. BIAS IN  
SEARCH ENGINES

### 3. POLARISATION AND HATE SPEECH

#### MOTIVATION:

- ▶ Social media platforms play a crucial role in shaping public discourse around societal relevant topics, e.g., ABORTION RIGHTS or VACCINATION.
- ▶ Together with HATE SPEECH and TOXIC BEHAVIOR, polarised groups threaten social harmony and can lead to violence.
- ▶ Exposure to hate speech and toxic behavior on social media can negatively impact MENTAL HEALTH.

#### RESEARCH QUESTIONS:

- ▶ What drives the change of user behavior (e.g., hate speech) and how is it connected to the polarisation of the individuals?

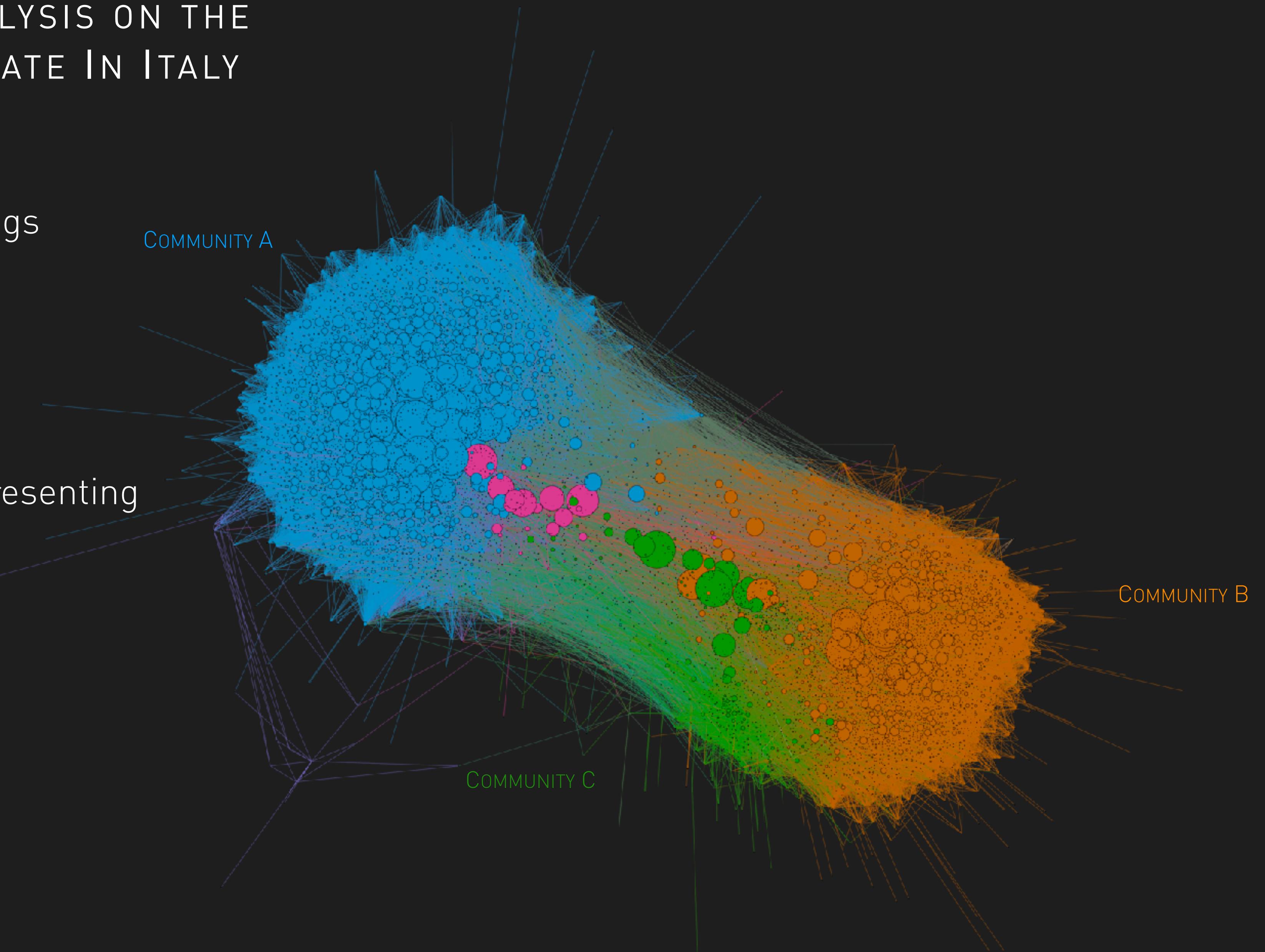
#### METHODS:

- ▶ Network Science, NLP

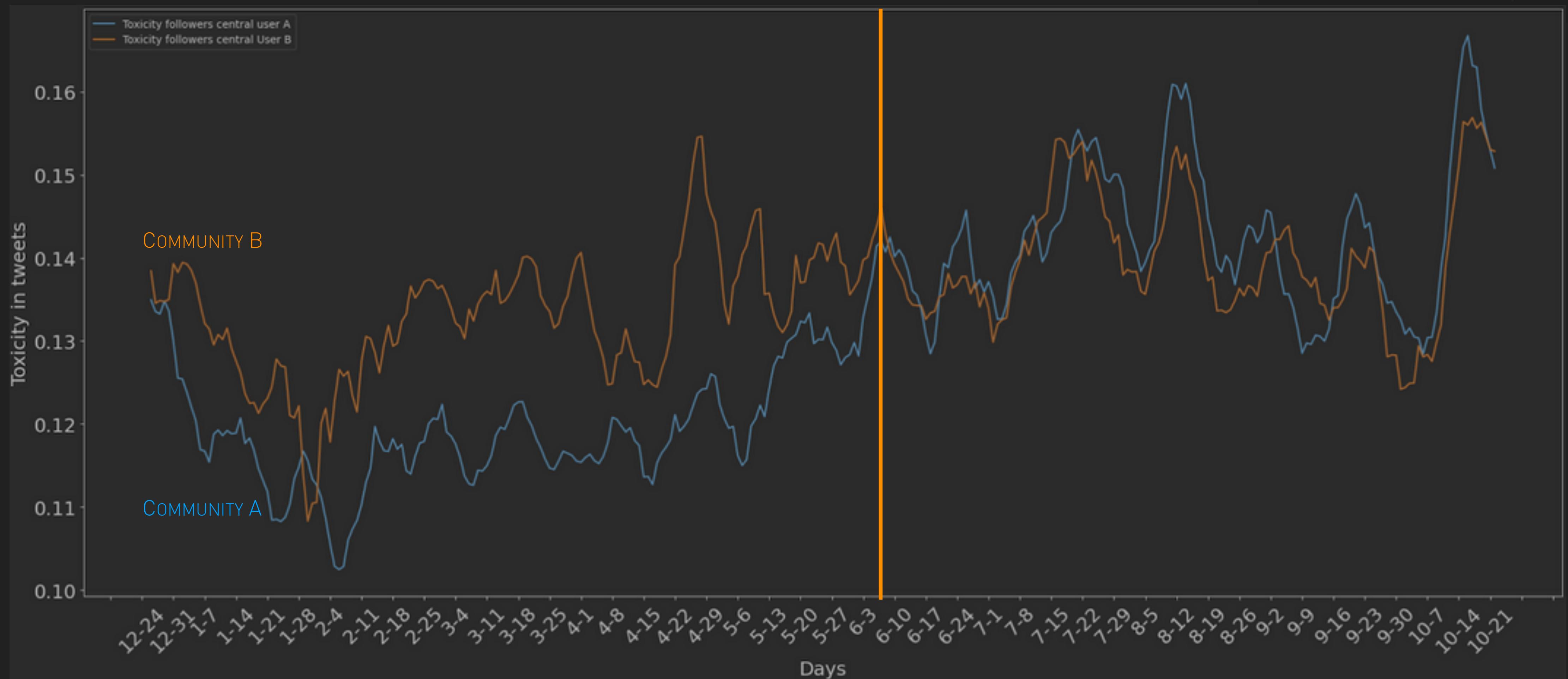
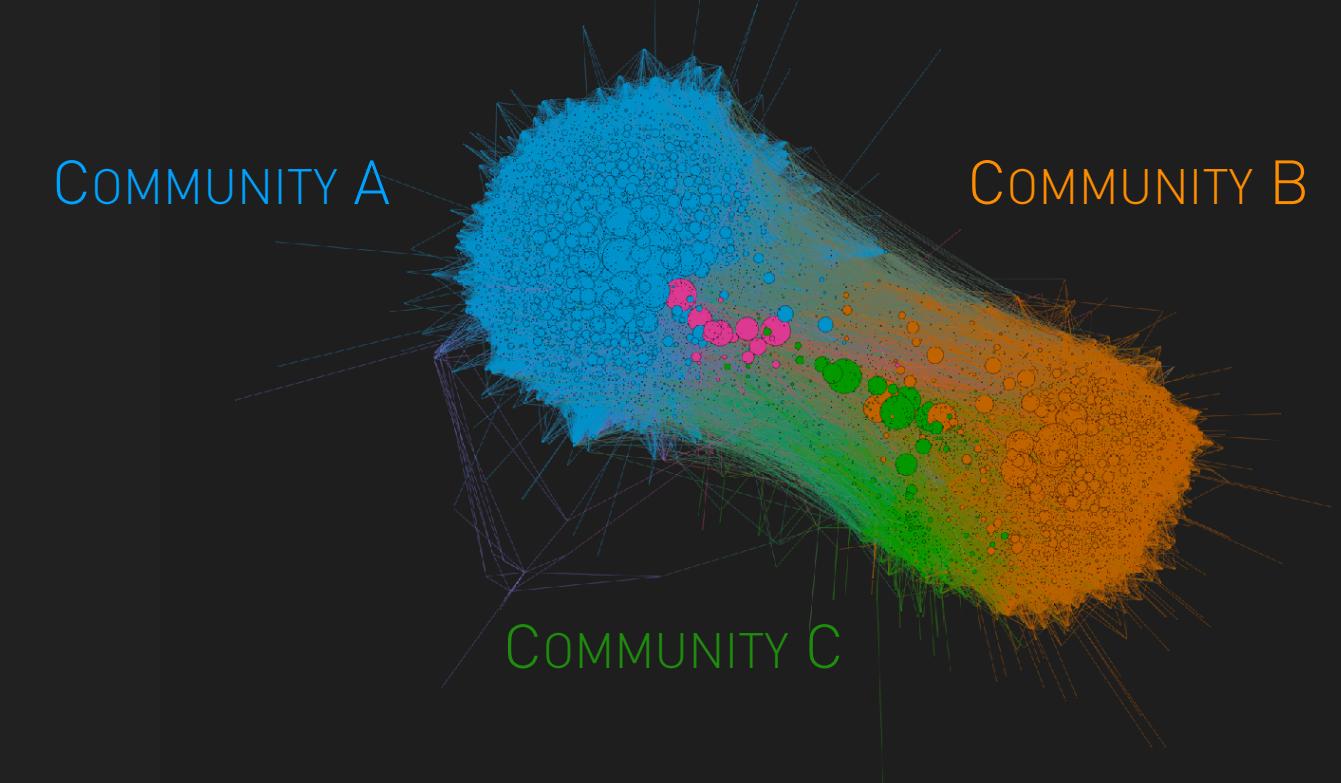
# COMMUNITY DETECTION ANALYSIS ON THE COVID-19 VACCINATION DEBATE IN ITALY

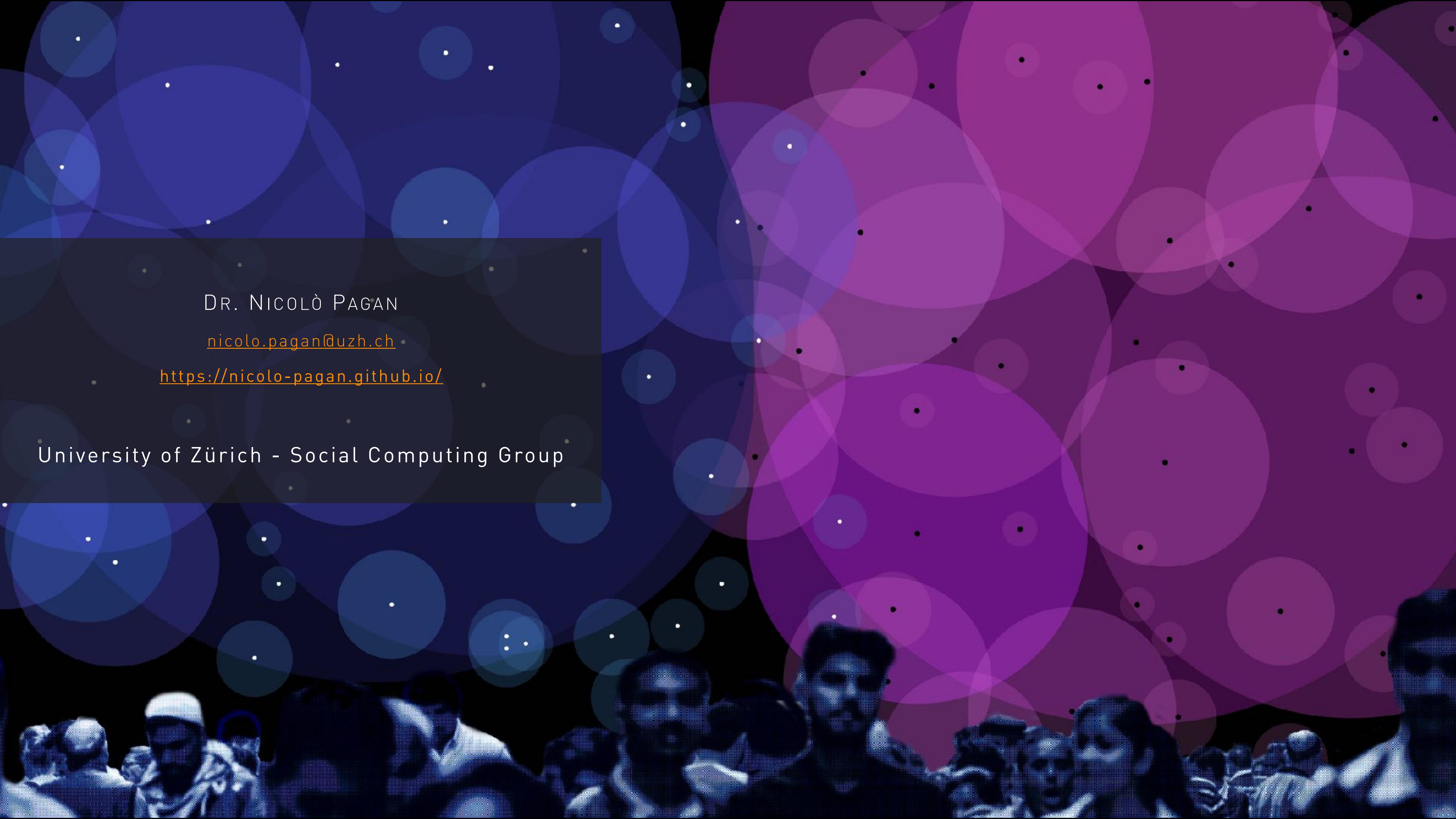
## DATA:

- ▶ Twitter API, using Italian hashtags
- ▶ December 2020 - October 2021
- ▶ Re-tweet network with:
  - 8'980 Core users
  - 644'784 weighted edges representing
    - 2'215'605 retweets



# ANALYSIS OF TOXIC BEHAVIOUR AMONG FOLLOWERS OF CENTRAL USERS





DR. NICOLÒ PAGAN

[nicolo.pagan@uzh.ch](mailto:nicolo.pagan@uzh.ch)

<https://nicolo-pagan.github.io/>

University of Zürich - Social Computing Group