Bridging Nations: Quantifying the Role of Multilinguals in Communication on Social Media

Julia Mendelsohn
University of Michigan
juliame@umich.edu

Sayan Ghosh
University of Southern California
ghoshsay@usc.edu

David Jurgens
University of Michigan
jurgens@umich.edu

Ceren Budak
University of Michigan
cbudak@umich.edu

- Information on social media can spread quickly across languages and countries. *But we don't really know how.*
- Using NLP, causal inference, and network analysis, we quantify the **role of multilinguals in cross-lingual information exchange** on European Twitter
- Multilinguals play overall important role, with the biggest effects in **spreading otherwise less-accessible information**
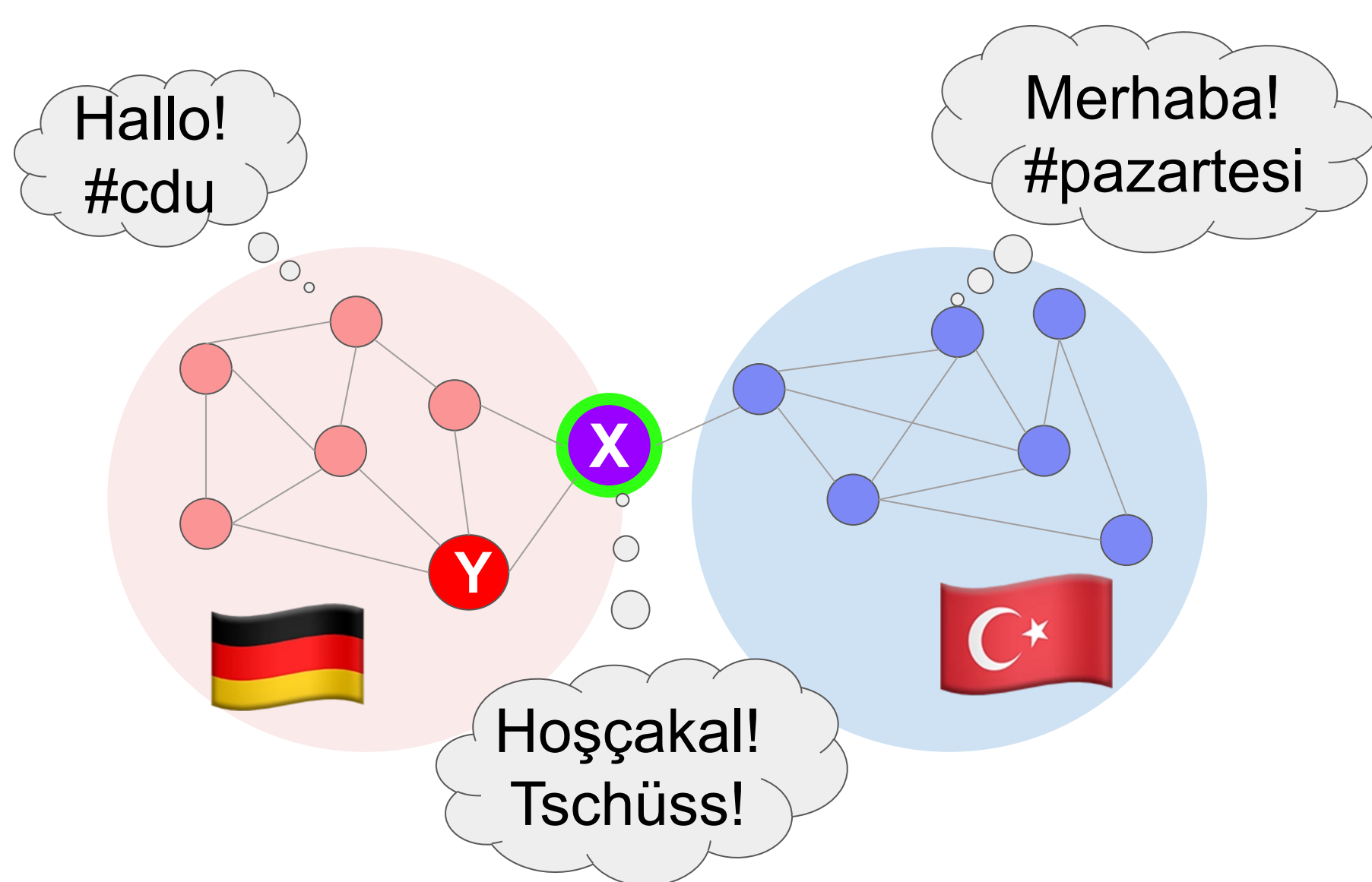

Measures of Influence

- **Structural Role:** To what extent are multilingual users bridges, positioned to share information?
- **Communication Influence:** How does having a multilingual friend impact one's sharing behavior?

Multilingual Country Pair (MCP) Networks

- Country pairs C_x & C_y with dominant languages L_x & L_y
- C_x & C_y must have single, distinct dominant languages
- Undirected network of mutual mentions from Decahose
- ~250 MCPs from Europe
- Multilingual users have ≥10% of tweets in L_x & L_y (Twitter's LangID). This captures language *performances* on Twitter.
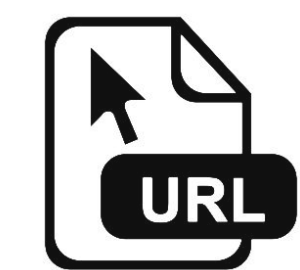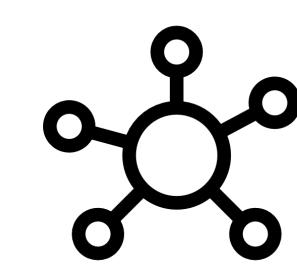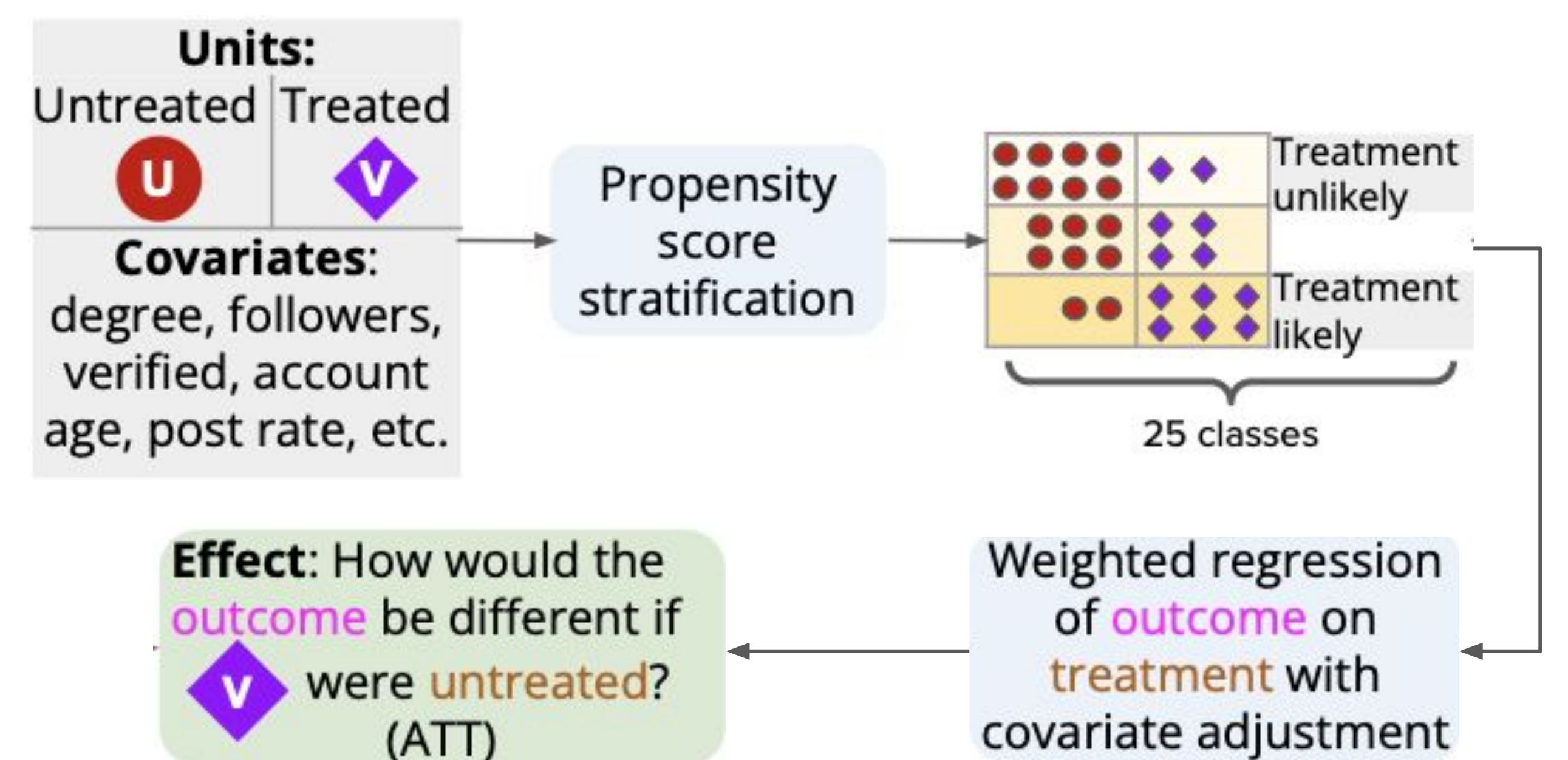
Effects of multilinguals vary across MCPs

- Western European multilinguals who post in Eastern European languages have an especially big influence
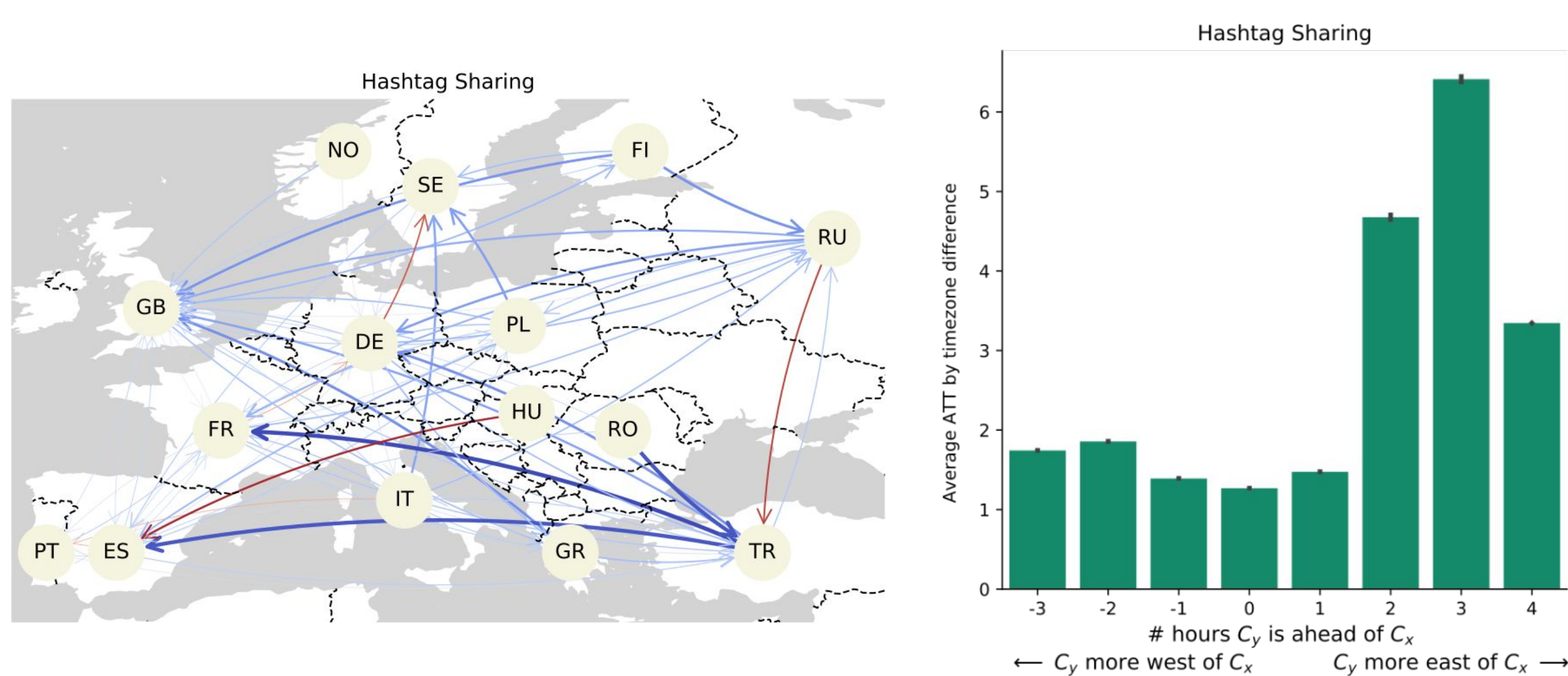
Scan the QR code for our ICWSM 2023 paper!

Causal Inference Design

	Structural role (Study 1)	**Communication Influence** (Studies 2 & 3)
Unit	Users from C_x who post in L_x	Monolinguals from C_x who use L_x
Treatment	Posting in L_x and L_y	Having a multilingual friend who posts in L_x and L_y
Outcome	Betweenness centrality (log)	Sharing a URL or hashtag associated with L_y (based on log-odds)

For a user in C_x, posting in both L_x & L_y increases:
- Betweenness centrality by **13.5%**
- Odds of a L_x friend sharing L_y hashtags **4-fold**
- Odds of a L_x friend sharing L_y domains **16-fold**

Effects of multilinguals vary across hashtag topic

- Multilingual contextualized topic model [Bianchi et al., 2021]
- k=50, evaluation with topic-intrusion in 5 languages
- Largest effect of multilinguals for:
  - **Political content** over **entertainment**
  - **Newer (health related) social movements**
  - Searching for and offering **job opportunities**