

# Forced Migration from Ukraine: Lessons Learned from Organic Data

*Lisa Singh, Katharine Donato, Ali Arab,  
Nathan Wycoff, Elizabeth Jacobs  
Georgetown University, USA*



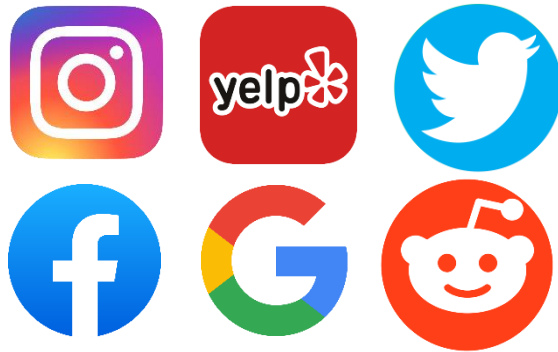
# Broad Goal

What types of organic data can improve our understanding of emerging and/or prolonged forced migration?



# What is Organic Data?

Non-design data generated as part of a person's routine and/or a society's normal functions



# Strengths

Generated in a more natural setting

Offers real-time data for analysis

Promising in difficult-to-access environments, where design data are hard to obtain

# Weaknesses

Lots of it –  
Hard to process

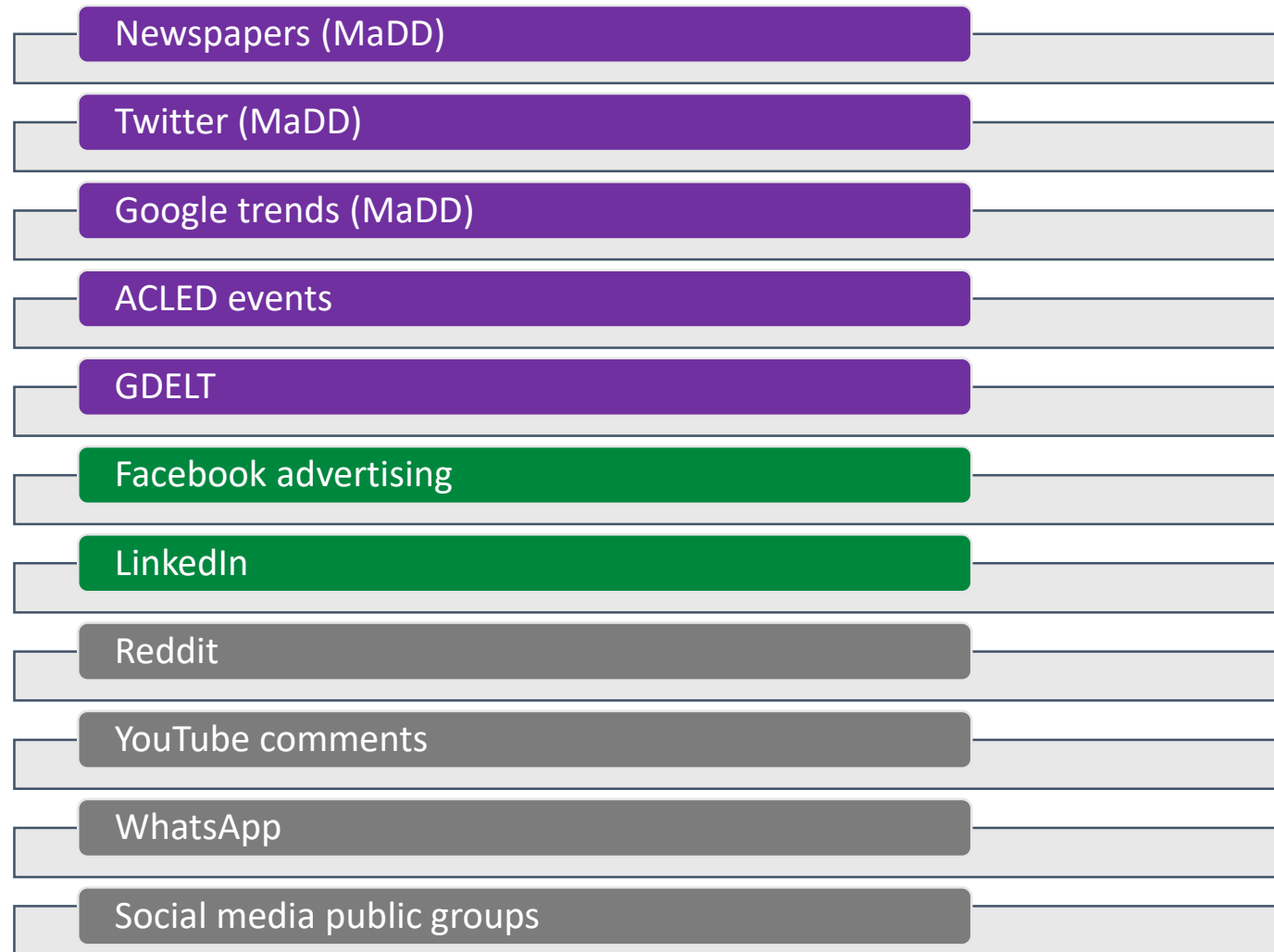
Difficult to  
generate  
variables

Noisy, partial  
and biased

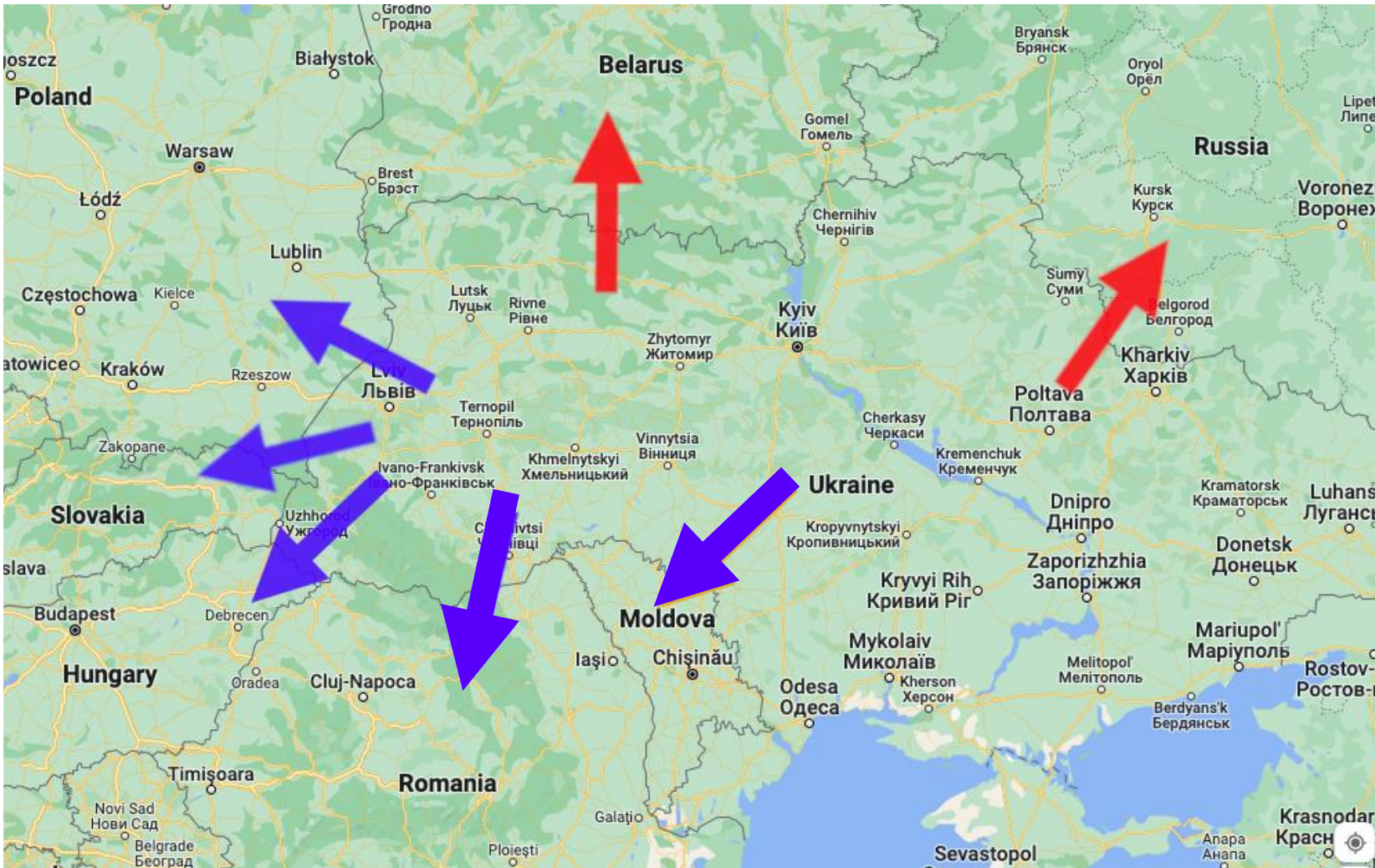
Possible ethical  
considerations



# Organic (Big) Data Sources Used in Research



# Predicting International Migration Flow from Ukraine

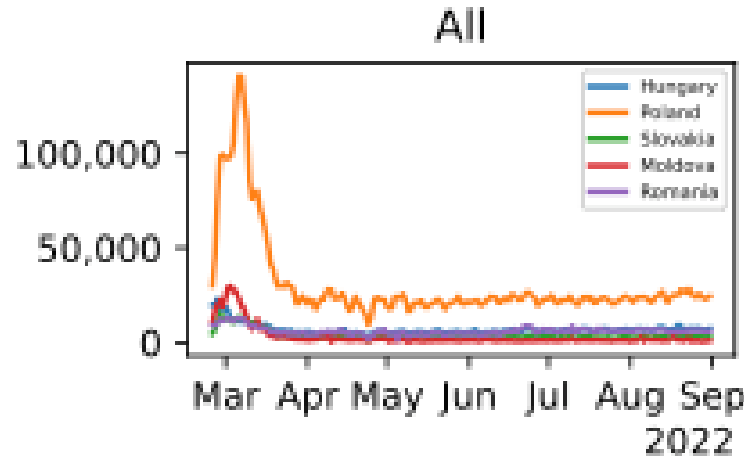
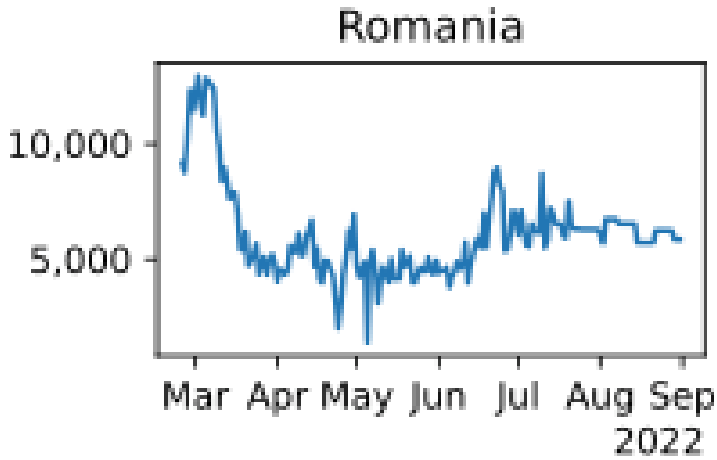
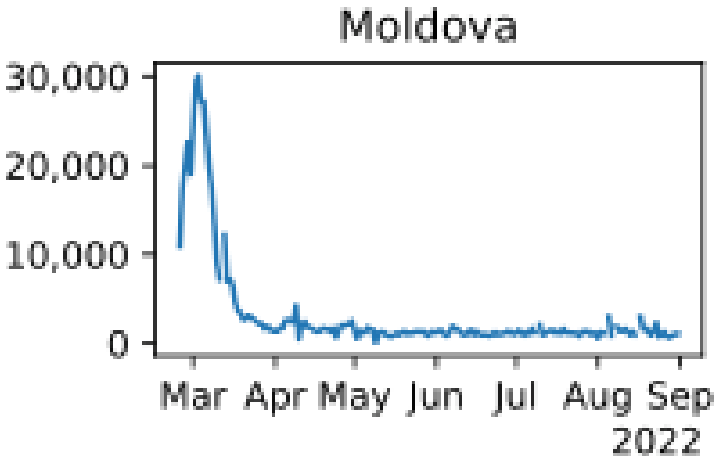
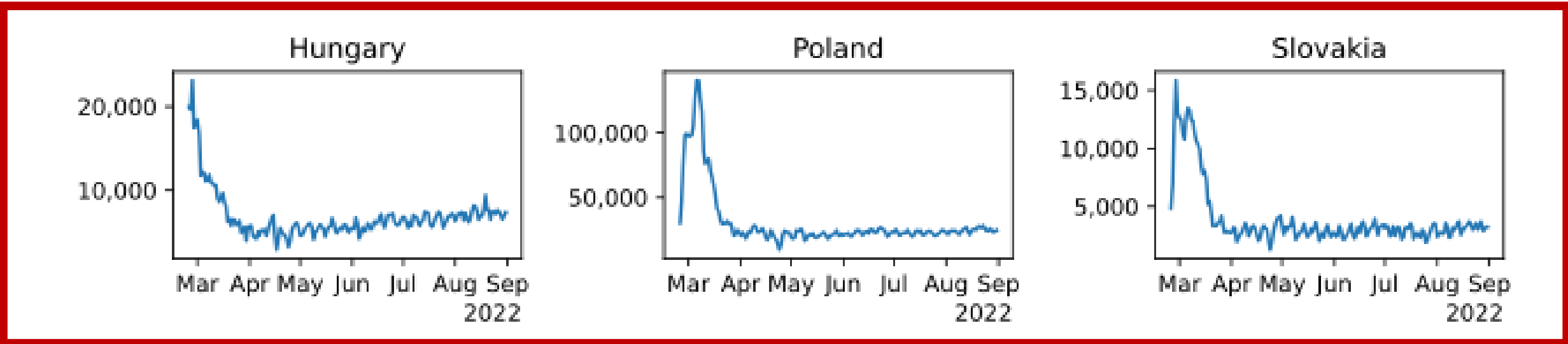


## First 6 Months of Conflict

	Total Flow	Share(%)
Poland	5,820,189	60.63
Hungary	1,300,034	13.54
Romania	1,147,112	11.95
Slovakia	739,635	7.70
Moldova	593,045	6.18

Based on UNHCR flow data

# UNHCR Flow Data from Ukraine



# Constructing Variables: Twitter Example

## Ukrainian keywords used to construct conversation buzz variables

- **Flee measures**
  - Flee: I am leaving; going to; taking train to; arrived at
- **Insecurity measures**
  - Physical: Weapons; soldiers; rockets; bombs; explosion; attack; deaths
  - Food: Hunger; food shortage; rationing, drinking water
  - Health: COVID; corona; omicron; pandemic; hospitals; medical supplies
- **Contextual measures**
  - Political: Zelensky; Putin; negotiations; declaration; protests; war
  - Economic: Economy; exchange rate; gas; oil; sanctions; exports; money



# Modeling: Using Organic Variable to Measure Flow

$$\mathbb{E}[\log(y_t)] = z_t' \beta + \delta_{d(t)}$$

- $\log(y_t)$  - Order of Magnitude of Outflow to Slovakia, Hungary and Poland.
- $z_t = \frac{1}{2W} \sum_{\omega=-W}^W x_{t+\tau+\omega}$  - Lag and Aggregation (Laggregated) Organic Variable:
  - $W$  - Window radius
  - $\tau$  - Lag
- $\beta$  - Regression coefficient vector
- $\delta_{d(t)}$  - Day of week effect.

*Gaussian, Poisson, Negative Binomial likelihood give qualitatively similar results*

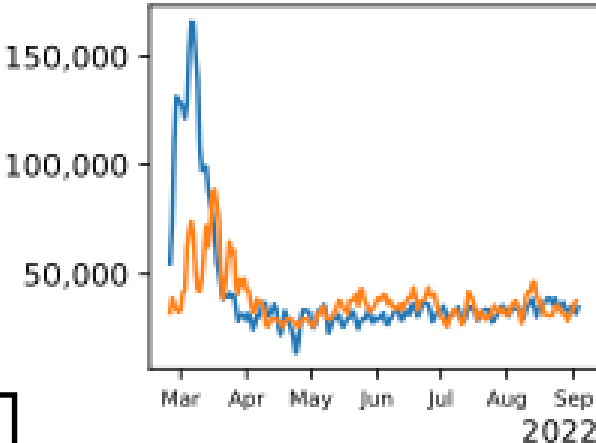
# Comparing Data Sources Relationship to Flow

Counts of  
Individuals  
Leaving  
Ukraine

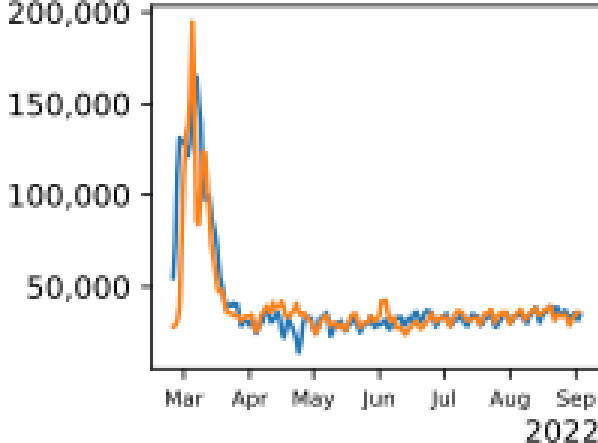
**LEGEND**

- UNHCR Flow Estimate
- Organic Data Estimate

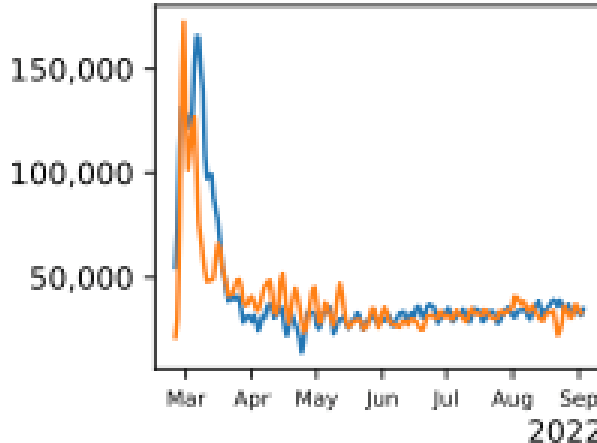
Twitter-Economic



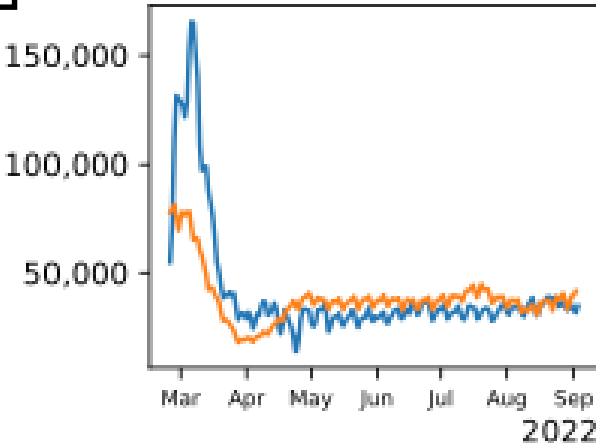
Trends-Travel



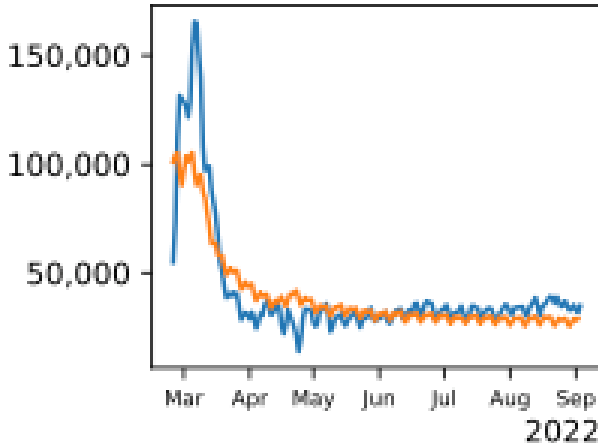
Trends-Physical



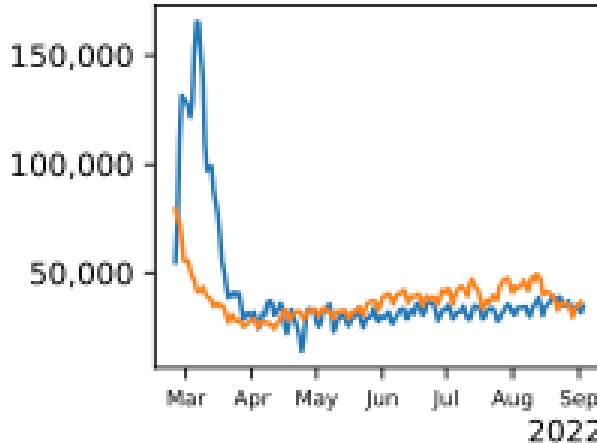
ACLED-Fatalities



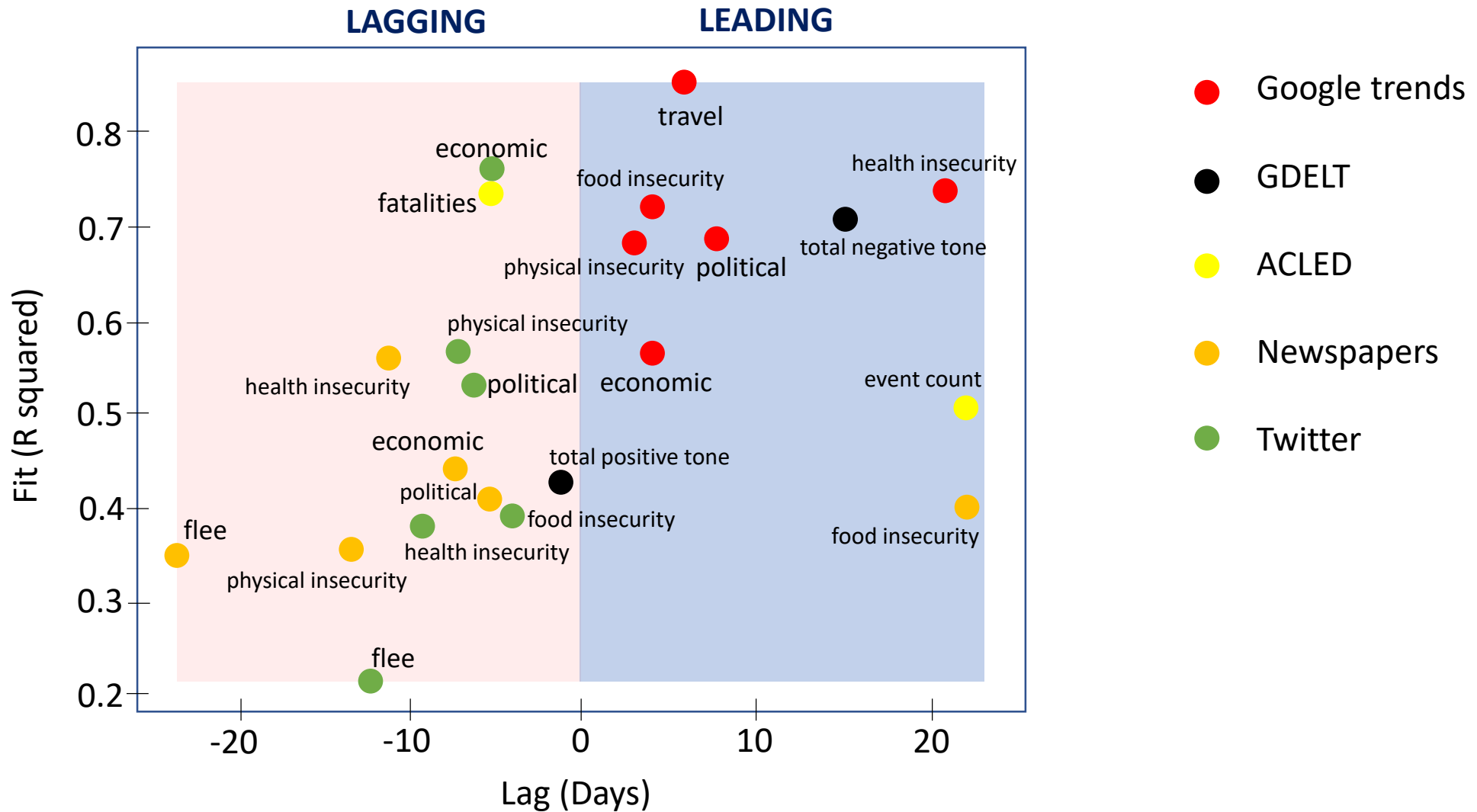
GDELT-Negative Tone



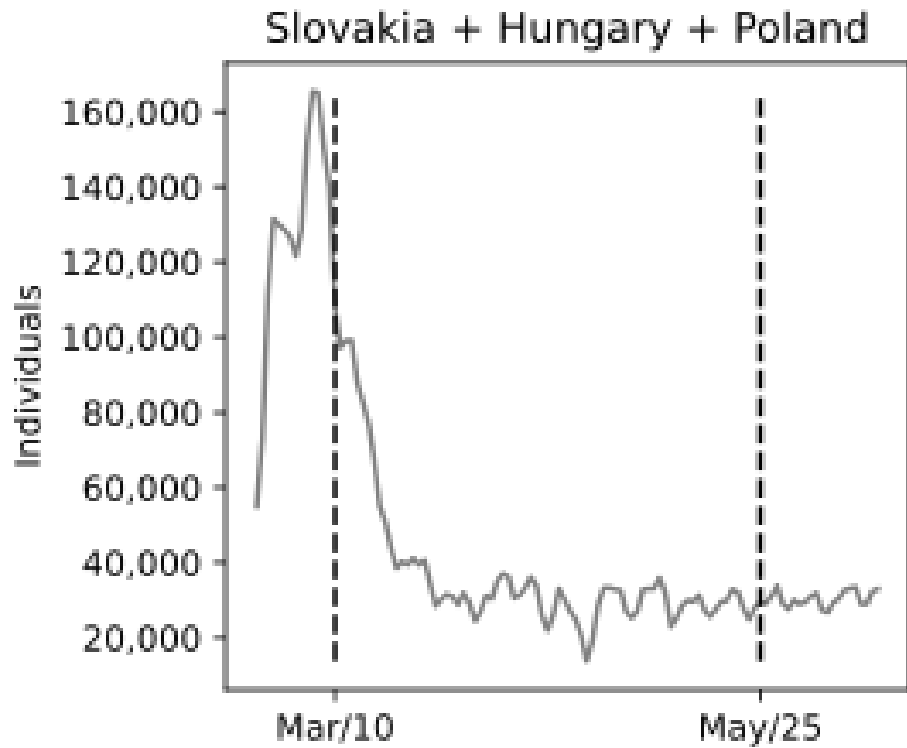
News-Physical




# Explainability vs Timeliness



# 'Prediction Error' at Two Time Points



# Lessons Learned



Because Ukraine has more granular flow data, using a simple model that considers the mean flow from the previous week is reasonable (after the initial hump) for international migration.

When a crisis emerges, public organic data sources are a viable option for modeling the changing dynamics of flow.

For this crisis, Google trends data (generally) is the best organic data signal for retrospective analysis and nowcasting. All organic sources captured the two phases of the crisis.

For longer term forecasting in countries with more sparse flow data, more variables are needed and models that quantify uncertainty and allow for more variation in temporal and spatial resolution are important (hierarchical Bayesian).

# MaDD Core Team

## Social Scientists

- Faculty

- Katharine Donato
- Elizabeth Ferris
- Susan Martin

- MDI/ISIM Fellow

- Nathan Wycoff
- Elizabeth Jacobs\*

## Computational Scientists

- Faculty

- Lisa Singh
- Ali Arab
- Ameeta Agrawal

- MDI Technical Team

- Colton Padden\*
- Yiqing Ren
- Maanasa Vatsavayi

## Interdisciplinary Students

- Yanchen Wang (PhD)
- Rob Churchill (PhD)\*
- Yaguang Liu (PhD)
- Ken Kawintiranon (PhD)\*
- Didier Akilimali (MS)
- Qihang Wang (MS)
- Sonali Rathinam (MS)
- Aidan Pizzo (undergrad)
- Jenny Park (undergrad)

\* Recent alumni

*Thanks to McCourt Institute, Massive Data Institute,  
& Institute for the Study of International Migration  
for their funding.*