

Using Twitter data for the research of Romanian migration flows



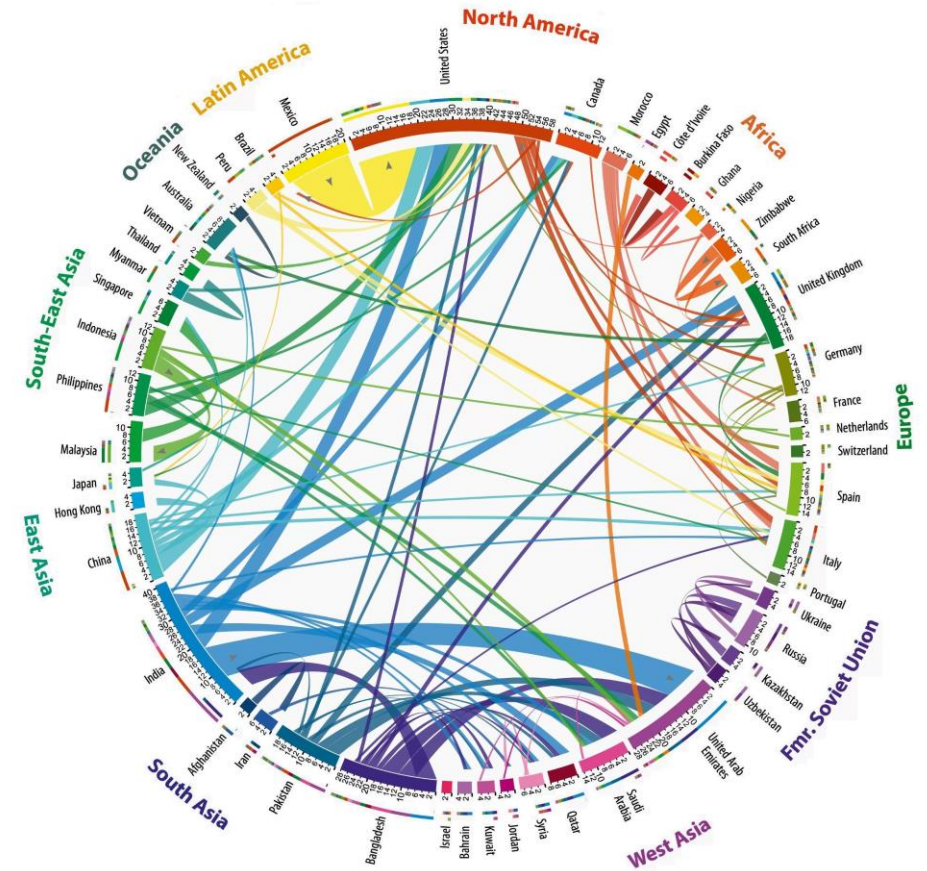
PhD Student Alexandru Florea

Prof. Dr. Monica Roman

University of Economic Studies Bucharest

Context

The usage of internet technologies in the day-to-day life has faced a dramatic increase in recent years. As result, the data from social media, such as Twitter, might reveal the real behavior of the user and ultimately, be leveraged for innovative researches. Since globalization is mainly driven by the diffusion of technology and population migration, the demographers become more and more interested in knowing how social media platforms can facilitate the innovation in studies focused on population mobility.



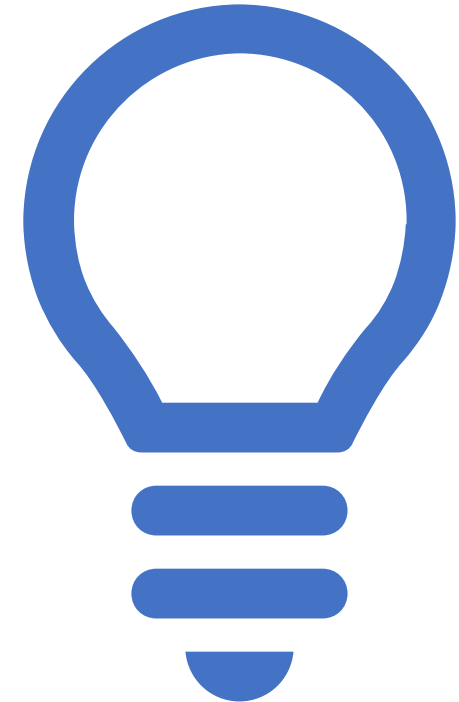
Source: United Nations Population Division

4th of October 2018

Research Goals

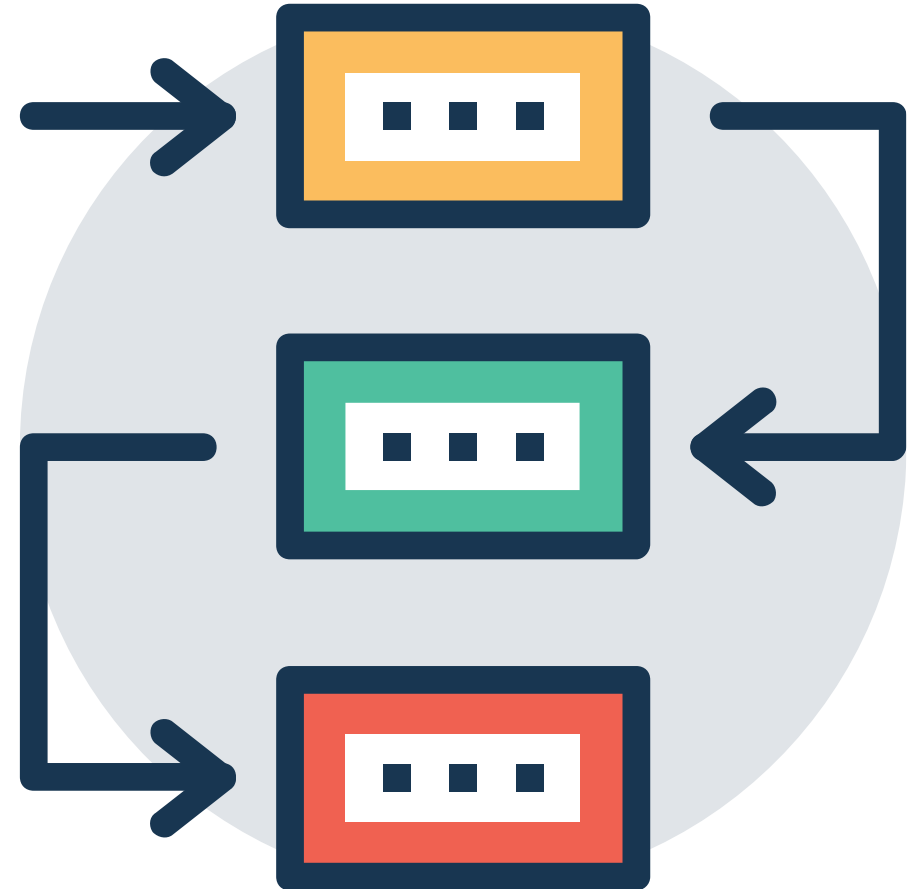
This research aims to :

- emphasize the importance of migration as a growing phenomenon;
- investigate the accuracy and completeness of the official data sources;
- explore innovative techniques that might improve the current data collection methodology;
- exemplify the retrieval and use of social media data using R packages;



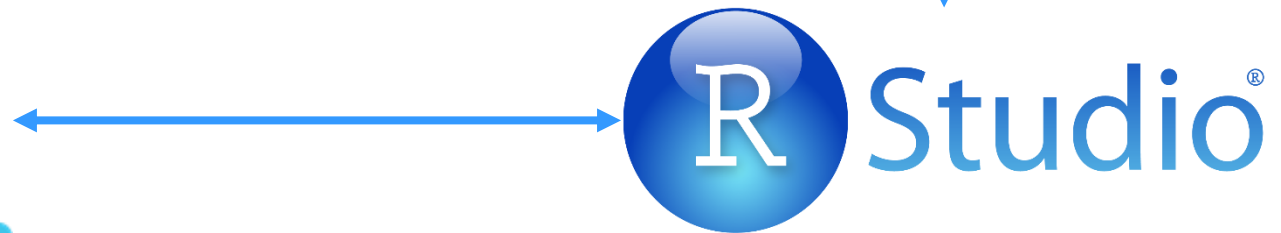
Structure

- First part of the research captures the increasing magnitude of migration, especially in emerging countries like Romania;
- The following section reveals that there is a significant lack in the migration data.
- An innovative approach to the missing data problem is presented in the third part of the study.
- The last part reveals how many opportunities social media data offers.



Tech & Tools

Tech: API (Application Programming Interface)
Tools: Rstudio, Twitter, LinkedIn.



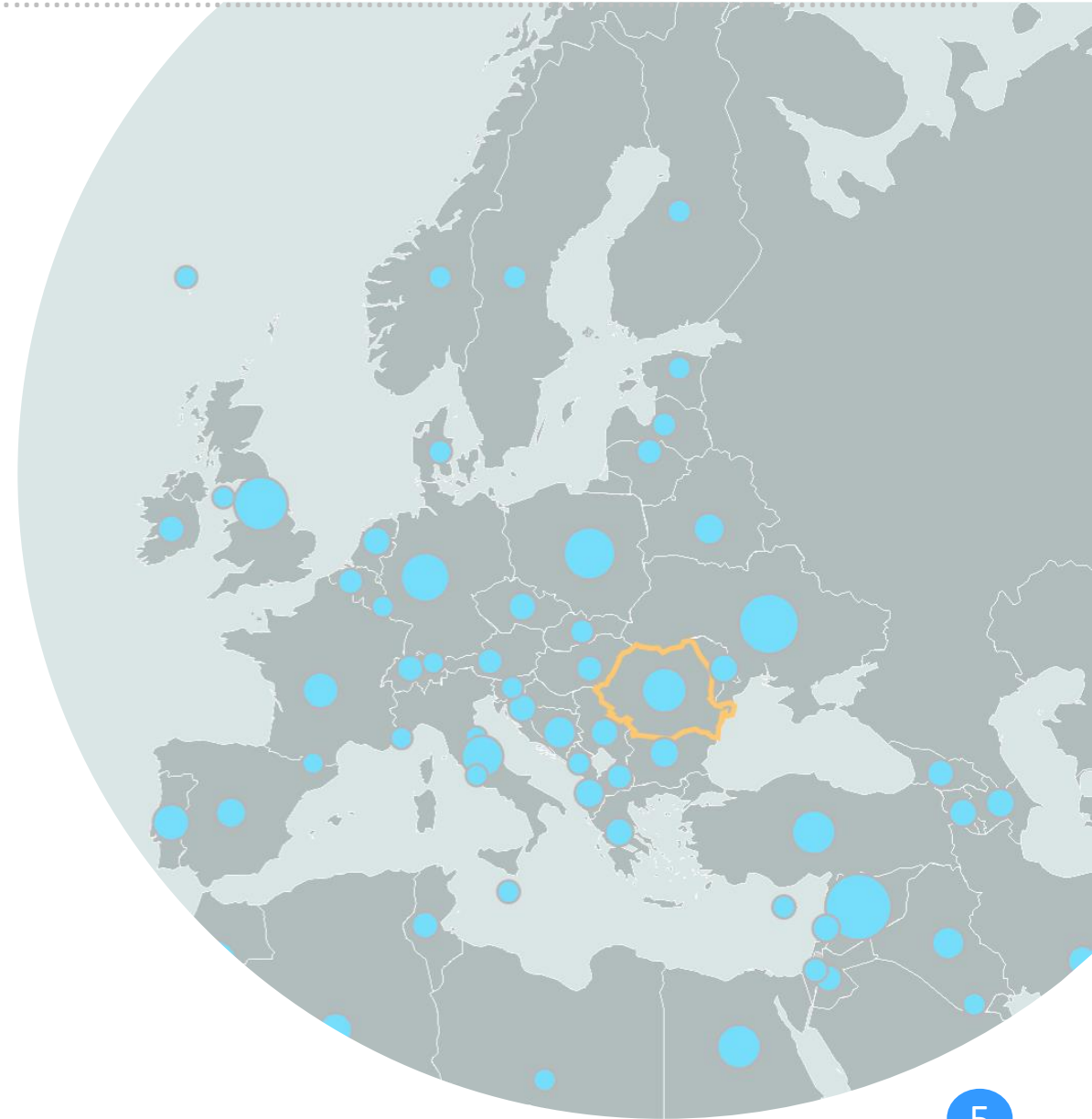
Romanian migration

Romania is one of the most important European countries in terms of absolute number of the emigrants (3.6 million)

After the communist regime fell (1989 in Romania), migration has started to be a more important phenomenon with each passing year (Diminescu D. , 2003).

Romanian labor force migration strongly expanded between 1997 and 2006 having Italy and Spain as main destination countries (Anghel, Botezat, Cosciug, Manafi, & Roman, International Migration, Return Migration and their effects. A comprehensive review on the Romanian Case, 2017)

Source: UNDESA 2017



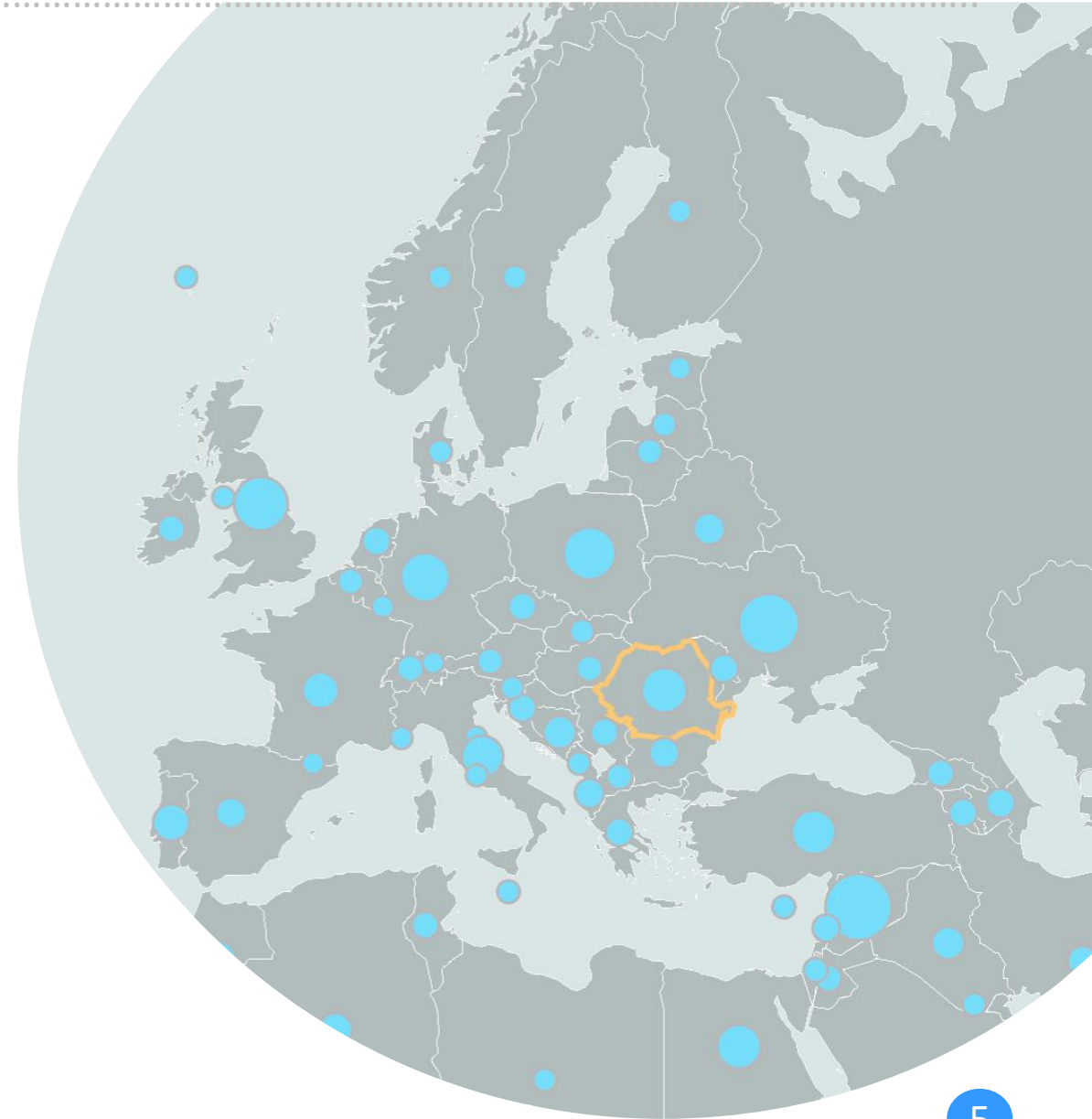
Romanian migration

Another aspect worth mentioning is the “brain-drain”, especially in the medical sector, Romania being one of the biggest sources of physicians in the Eastern Europe

In 2012, 30% of the foreign medical doctors in France were Romanians (Buchan, Wismar, Glinos, & Bremner, 2014).

The most dramatic increase of Romanian emigrants was spotted in 2007, when Romania became an official EU member state and allowed Romanians to migrate without restrictions which convinced almost 460.000 Romanians to emigrate (INS, Migratia Internationala a Romaniei, 2014).

Source: UNDESA 2017



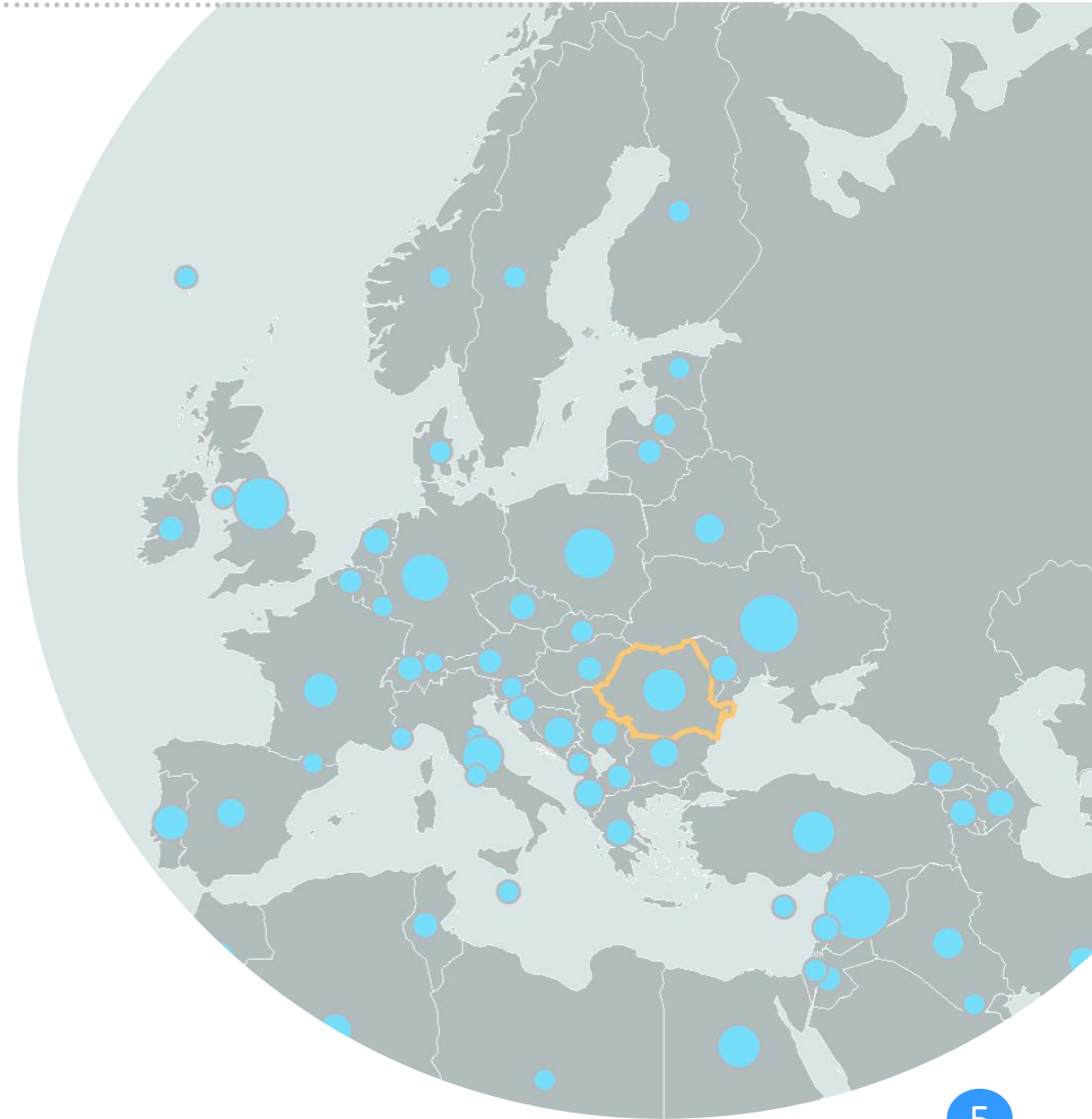
Romanian migration

After Romania's admission to the European Union, Western & Northern Europe received more "high skilled" Romanian emigrants than before.

Studies show that Romania invested plenty of resources in developing a strong medical labor force but unfortunately a worrying part of it left the country or expressed its intention to migrate (Wismar, 2011).

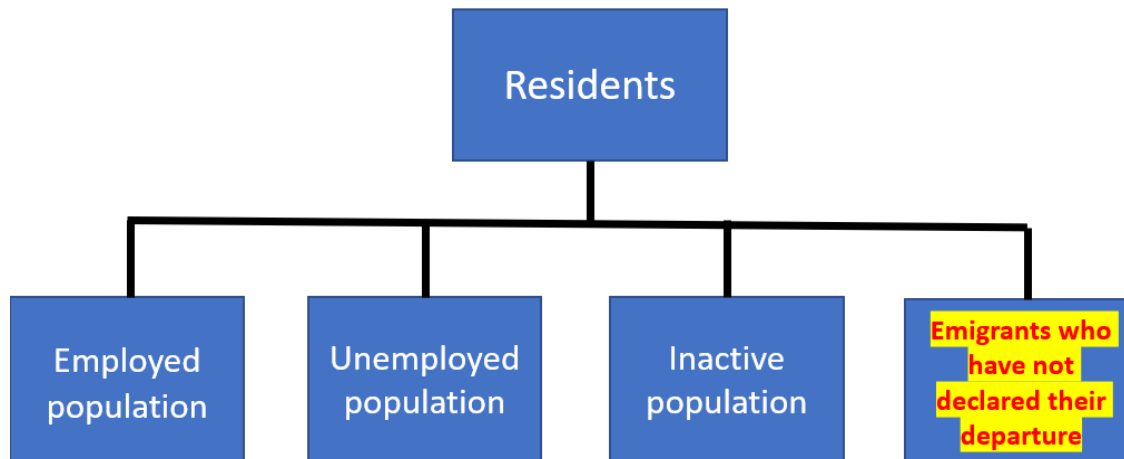
Studies show that while for Romanian emigrants without tertiary education returning home is still an option, the highly qualified migrants have no intention to return home as result of their success to obtain a good job/qualification (Morosanu, 2013).

Source: UNDESA 2017



Official data accuracy

Studies have shown that there are more migrants who have not declared their departure than the ones who have did it, especially in the emerging countries.



Most recent data: 2016



Most recent data: 2016



THE WORLD BANK

Most recent data: 2015

For example, in Lithuania, studies show that only each second or third resident declared his/her emigration.

The official approach consists in using census data, is a very time-consuming process and in most of the cases, obtained data is not consistent across countries and regions and it might also miss significant migration stats due to various reasons (e.g. undeclared departure).

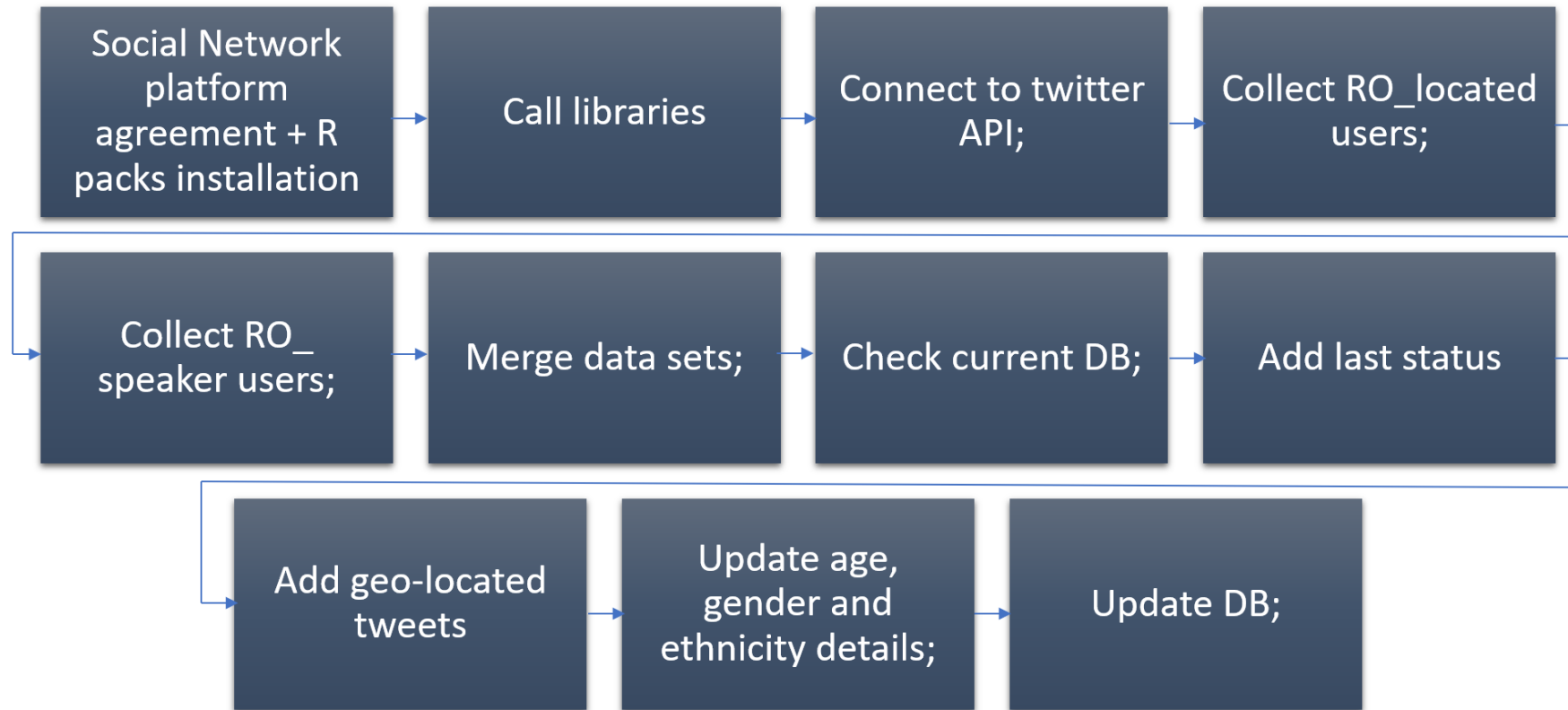
Migrants: people changing their residence to or from a given area (usually a country) during a given time period (usually one year).

Social Media data

- Over the past 15 years social media has become a more and more important part of our “every day” life;
- It contributes to migration;
- It captures the migration phenomenon;
- It can provide to statisticians a better understanding of the “brain-drain” phenomenon;
- It constantly improves in terms of accuracy, completeness, reliability and representability;




Algorithm



Face Recognition Accuracy

The famous artist “Justin Bieber” was estimated to be 25 years old, based on his twitter profile picture while his real age is 24. The same applies for the other characteristics.



| | |
|-----------------------------|----------|
| chin_tip_x | 137 |
| chin_tip_y | 338 |
| confidence | 0.99937 |
| eye_distance | 81 |
| face_id | 1 |
| left_eye_center_x | 176 |
| left_eye_center_y | 164 |
| pitch | 14 |
| quality | -1.61715 |
| right_eye_center_x | 96 |
| right_eye_center_y | 175 |
| roll | 0 |
| top_left_x | 28 |
| top_left_y | 113 |
| yaw | 32 |
| face_age | 25 |
| face_asian | 5.00E-05 |
| face_black | 0.00068 |
| face_glasses | None |
| face_hispanic | 0.02126 |
| face_lips | Together |
| face_other | 0.00128 |
| face_white | 0.97674 |
| face_gender_female_confiden | 1.00E-05 |
| face_gender_male_confidence | 0.99999 |
| face_gender_type | M |

Source: United Nations Population Division

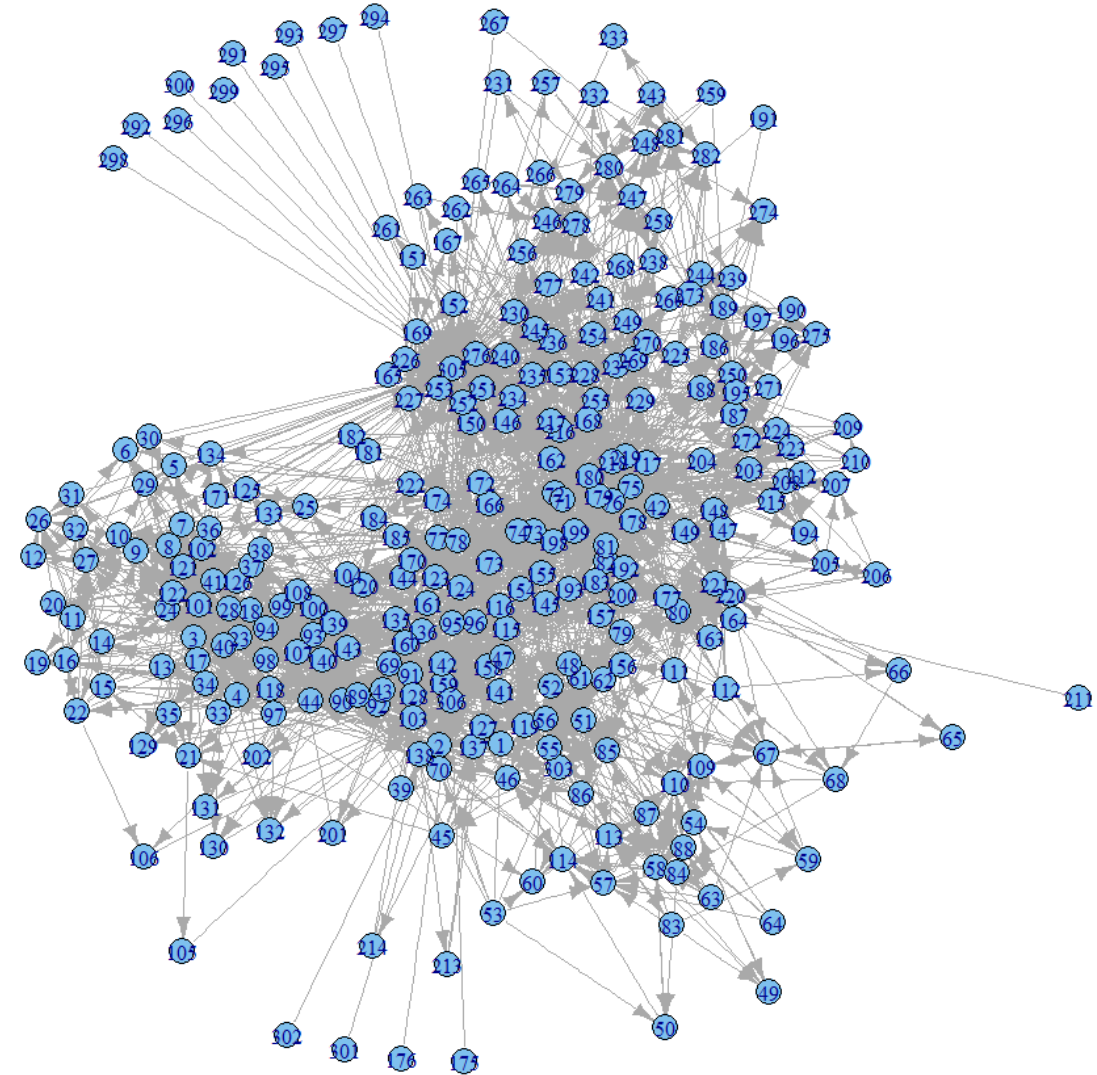
Usage a.k.a. “So what?”

- Twitter data as a proxy for global mobility patterns;
- Social Media data can be used for inference;
- Social Media data can provide a better understanding of the “brain-drain” phenomenon;
- Social Media data can fill the missing gaps in the migration data;
- It unlocks unlimited opportunities;



Conclusions

- *Migration is increasing in dimension and complexity;*
- *Current data collection methodologies cannot capture the mobility flows well enough;*
- *Social media is an ocean of data waiting to be explored;*



Q&A

Thank you!