Migration Data Innovation for Evidence-Informed Policymaking

BIGSSS Summer School in Computational Social Science: Data-driven Modelling of Migration

Michele Vespe, Lorenzo Gabrielli - European Commission - Joint Research Centre

Description

Adequate data is key for evidence-informed policymaking and for a better understanding of the rapidly changing scales and dynamics of migration. One of the [Knowledge Centre on Migration and Demography (KCMD)](#) lines of activity is relevant to improving access to migration data, promoting their use across multiple policy areas, as well as exploiting non-traditional data sources to address [data gaps](#). The KCMD and IOM GMDAC launched the [Big Data for Migration (BD4M) Alliance](#), aiming to advance discussions on how to harness the potential of big data sources for the analysis of migration. Concrete research lines using social media, micro census, mobile phones and air traffic data to improve migration data have recently started and this project will further investigate some of them in order to answer pressing policy question.

A first week shall be dedicated to an overview of the data landscape at European and global level and to framing and analysing the relevant migration research question, which will be further explored during the second week by hands-on activities using data-driven modelling techniques.

Research question

- How can non-traditional data fill migration data (stock, flows, skills etc.) gaps (timeliness, granularity), complement official statistics and provide anticipatory capacity?
- What insights can be extracted from data on migrant communities in cities across Europe for integration policies?
- What is the nexus between mobility and migration?

Data

Datasets that to be used for the module/project will be social media advertising platforms (migrant stocks estimates, skilled migration), micro-census maps of migrant communities (residential segregation and integration at local level), air traffic passenger derived data (mobility and migration nexus). One or more datasets will be prepared depending on the selected project (to be discussed), with the aim to complement official migration statistics using non-traditional data. Official statistics will also be used for training/validation, including from Eurostat, OECD, UNDESA and the World Bank.

Team composition

The expertise offered by the project leaders is a mix of migration research, data science, big data and science for policy. In particular:

> *Dr Michele Vespe* is a scientific officer at the European Commission - Joint Research Centre, where he works for the Knowledge Centre on Migration and Demography (KCMD). He coordinates the activities of a team of researchers in transforming migration data into knowledge and scientific evidence for policy makers. This includes research in the fields of

data integration, big data and innovative data sources on migration. He also manages the developments of the KCMD Dynamic Data Hub (https://bluehub.jrc.ec.europa.eu/migration/app/), a platform that allows and facilitates online analysis and time series visualisation of multiple data sets on migration.

Dr Lorenzo Gabrielli is a Data Scientist at the Joint Research Centre for developing innovative tools in the domain of customs risk assessment. Over the last years he has gained experience in the analysis of Big Data with Data Mining and Machine Learning techniques in a national and international context collaborating with several public and private research institutes. His interests concern the study of mobility in order to identify individual and collective patterns of behaviours. He has an excellent experience in managing geospatial databases (PostgreSQL/Postgis) and programming languages (Java, Python and R).

Potential lecture topics

Here below two lectures that can be delivered during the first week of the summer school.

1. *Migration Data Landscape and Perspectives* (Michele Vespe) - The lecture is a hands-on exploration of available migration and population projections data. Practical examples will show current gaps as well as the added value offered by the integration of cross-domain sets of data. Following an introduction to available datasets on migration and demography, the lecture will focus on i) data gaps in terms of fragmented, incomplete or scattered data often undermine their use; ii) combination of data from multiple sources and across domains to provide evidence and elements to better understand the dynamics of migration; iii) innovative data sources, methods and initiatives (e.g. Big Data and Alternative Data Sources on Migration: from Case Studies to Policy Support) to process existing data identify future research strands to a more in-depth or timelier understanding of new forms of migration, their drivers and impacts. Attendants will be guided through case studies and concrete examples by direct access to online resources and tools (such as the Dynamic Data Hub).

2. *International Migration Drivers* (Michele Vespe) - Why do people migrate? Through an extensive analysis of international data, the recent KCMD report International Migration Drivers sheds light into this question by showing which factors are the most important drivers of migration flows around the world. In doing so, the findings take a step towards better understanding what shapes international migration and translating that knowledge into a resource to support policymaking.

3. *Migration Data at City Level, a cornerstone for the local dimension of integration policies* (Michele Vespe) – Recent developments in data creation, collation and analysis provide city authorities with significant opportunities to become more effective and efficient. Through the use of data, authorities can be better placed to see, understand and respond to the challenges they face. But geographically detailed data on migration is not always easily accessible, consistently collected or comparable across EU cities. This lecture will give an overview of the D4I initiative, bringing together data from Censuses from several EU Member States. By processing data from the 2011 Census on the number of migrants by origin (country of birth and/or citizenship) in EU Member States and at the highest possible level of spatial resolution, it shows how diverse datasets can be harmonised to produce valuable insights into the composition of cities. Censuses contain a wealth of information that can be used to calculate

residential segregation and diversity. From these, subsequent analyses can provide insights into the impact of migration and diversity on a range of local issues, from electoral outcomes to housing markets.