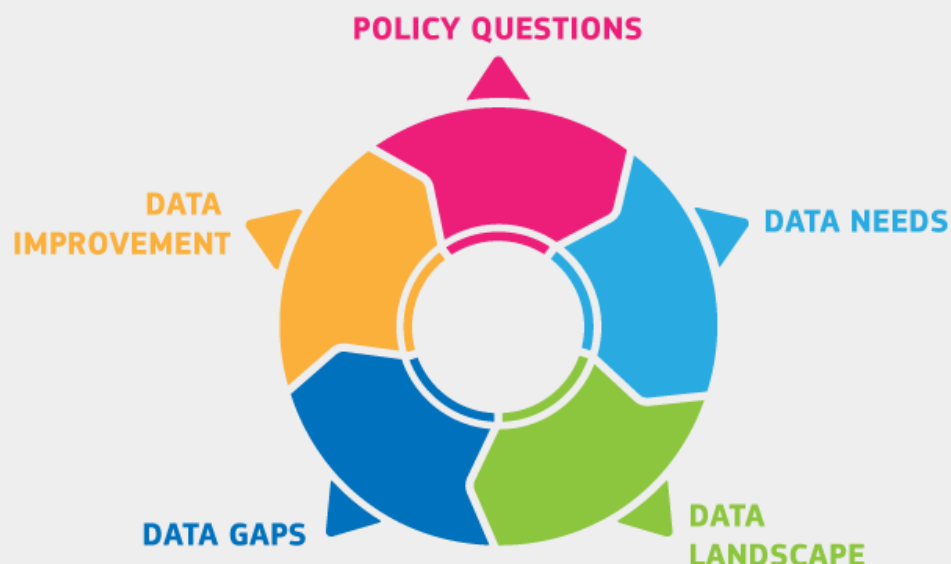European Commission

JRC SCIENCE FOR POLICY REPORT

Towards an EU Policy on Migration Data

Improvements to the EU migration data landscape

Edited by C. Santamaria and M. Vespe

2018

POLICY QUESTIONS

DATA NEEDS

DATA IMPROVEMENT

DATA LANDSCAPE

DATA GAPS

Joint Research Centre

This publication is a Science for Policy report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication.

Towards an EU Policy on Migration Data. Improvements to the EU migration data landscape

Adequate data is key for evidence-informed policymaking. This report describes the initiative coordinated by the KCMD that has brought together relevant EU policy officers (from Commission DGs and services, and EU Agencies) to discuss the current state of migration data for EU policymaking and to coordinate actions intended to address some of the existing data gaps. The report first presents the European migration data landscape and then describes some approaches that are being explored to improve data adequacy.

Contents

Authors

This work has been co-ordinated and edited by Michele Vespe and Carlos Santamaria (JRC), who contributed as authors to all the sections of the report. Alessandra Zampieri (JRC) contributed as author to all the sections of the report. Diana Ivan, Piotr Juchno, Giampaolo Lanzieri, Adam Wronski (Eurostat), Gabriela Senchea-Badea (Sogeti Luxembourg), Frederique Roche, Gemma Amran, Alessandro Sola, Pawel Busiakiewicz, Laurent Aujean (DG HOME) contributed as authors to Section 2. Marzia Rango (IOM GMDAC) contributed as author to Section 3.

Executive summary

This report is intended for anyone interesting in getting an overview of the migration data landscape at EU policymaking level. The target audience is not limited to EU policy-makers, but includes other migration stakeholders, international organisations, non-governmental organization (NGOs), and practitioners. Special emphasis is put on the role science can play to improve the data landscape, so scientists and researchers will also find information relevant to them.

Although the migration data landscape in the EU is better than in many other world regions, the recent migration crisis in Europe has spurred a number of actions to increase the adequacy of the data to policy needs. Given that migration is a global phenomenon, these actions are not limited to European confines. Significant among these international actions is the Global Compact for Safe, Orderly and Regular Migration, which recognises, as one of its objectives, the need to collect and utilise accurate data as a basis for evidence-informed policymaking. Target 17.18 of the Sustainable Development Goals (SDGs) calls for the provision of high-quality, timely and reliable data disaggregated by, among other variables, the migratory status; such disaggregation would make it possible to measure how migrants are faring in many of the Goals (e.g. poverty, health, education, gender equality).

One ongoing initiative that aims to improve the data adequacy for EU policymaking has been coordinated by the European Commission Knowledge Centre on Migration and Demography (KCMD) during the last 24 months. This report documents its main results. As part of this initiative, a series of events (e.g. regular workshops) have stimulated the dialogue among relevant European Commission DGs and services and EU Agencies, with the aim of detailing policy-makers' data needs, mapping data gaps, and working towards improvements in the EU migration data landscape.

This initiative has already brought some benefits, helping to raise awareness about data issues that sometimes affect more than one policy area, facilitating the sharing of information, developments, plans and priorities among policy-makers, and enabling discussions and agreements on measures to address some of the data issues. Some gaps reported by policy-makers have already been filled (e.g. the visibility of existing data on children in migration has been improved) or are undergoing improvements (e.g. Commission proposal to amend the regulation on statistics on migration and international protection - COM(2018) 307[1] -, whose flexibility would allow to take on board findings from this initiative; initial efforts to exploit big data for migration). This initiative has contributed to those efforts by providing background information, consolidating priorities, or promoting them. Ways to involve the EU Member States (MS) and to improve some of the data issues with their support are currently being investigated.

This report first documents the data landscape as stated by EU policy-makers (European Commission DGs and services, and EU Agencies), and presents an analysis and classification of existing data gaps in broad groups. This is followed by a short description of some approaches that are already being explored to address some of the gaps.

The main gaps have been categorised in the following four groups:

- Gaps in existing data: this group encompasses issues related to timing (e.g. data that is received by policy-makers with a significant delay with respect to the phenomenon it describes), quality, or inadequate disaggregation level of existing datasets.

- Dissemination issues: this group covers various aspects related to the way data is presented to policy-makers and that may prevent a full exploitation of the data, e.g. unclear assumptions or methodologies, and reduced visibility of key data.

---

[1] COM(2018) 307 final, Proposal for a Regulation of the European Parliament and of the Council amending Regulation (EC) No 862/2007 of the European Parliament and of the Council on Community statistics on migration and international protection

- <u>Data not collected</u>: this group includes topics for which data is not always collected or available, like the educational attainment of refugees, more detailed information on children, or intra-EU mobility of third country nationals.

- <u>Useful data that is currently inaccessible</u>: this comprises data that, even if it exists, is not easily accessible to EU policy-makers. This could be because of legal or commercial terms governing data access, (e.g. some MS data held by EU Agencies, new data sources like mobile phone data or social media), privacy concerns, or the need of innovative or laborious analysis of the data (e.g. new data sources, surveys).

Approaches that are already underway to address some of those gaps are classified as follows:

- <u>Collection of new data or production of new statistics</u>: this requires long-term planning and dialogue with MS, but it guarantees the results will have high quality and harmonisation across the EU. The Commission has already adopted the abovementioned proposal COM(2018) 307, whose overall aim is to provide more detailed and timely statistics on asylum and managed migration. Eurostat is also in active discussion with MS on the 2021 census round, which will include some variables geocoded at 1 km^2 grid, and, for the longer term (i.e. from 2024/5), on the collection of annual census-type population data. From 2021 the EU Labour Force Survey (LFS) will also include additional migration variables. In parallel, work is ongoing to provide finer and more regionally disaggregated migration-related estimates.

- <u>Better utilisation of existing data</u>: this refers mostly to the use of administrative data to produce more frequent and timelier migration statistics. The transition towards an increased use of administrative data sources is happening across the EU, although not all countries are equally prepared to take advantage of this shift, i.e. not all countries collect the same administrative data or have developed methods to extract statistics from those data sources.

- <u>Exploitation of new data sources</u>: this is a nascent line of research that aims to capture migration-relevant insights from new data sources, e.g. social media. Significant obstacles need to be overcome before valid information can be extracted from these sources: access to the data is normally restricted to the service providers, user confidentiality needs to be guaranteed, sampling bias. However, the pervasiveness of these new data sources means that they are likely to play an increasingly relevant role in the efforts to address the existing data gaps, complementing traditional sources.

Important efforts are being directed at addressing migration data gaps. These efforts take the form of various complementary approaches. All of these approaches are likely to contribute to improving the data landscape, although this process will be gradual and it will extend for a long period of time.

While migration policy can be driven by a broad variety of interests, sometimes even conflicting ones, in this report it is argued that adequate data (i.e. accurate, frequent, timely) is needed to inform all the policy areas that deal with various aspects of migration and that strive to manage migration better. Adequate data is also necessary to inform public discourse on this topic.

1 Introduction

The work presented in this report is motivated by the need to support evidence-informed EU policymaking in the area of migration with adequate data, and is intended to recognise new data needs, to strengthen data governance by facilitating a wider use and sharing of existing datasets and by ensuring that the datasets are made accessible and searchable, and to avoid unnecessary work for re-users of this data and for the Commission services alike. This document is a contribution towards the objectives of the Communication to the Commission on Data, Information and Knowledge at the European Commission[2], which identifies the strategic use of data, information and knowledge as an essential element to harness synergies across policy domains.

More specifically, the aim of this work is to identify shortcomings in migration data and to present a brief analysis of the challenges that explain the existence of such shortcomings or the difficulties in closing them. The initiative is driven by the need to better answer information and analysis requirements of EU policy-makers, and it intends to make positive steps towards that goal by signalling pathways to improve the data. This process of data improvement can also lead to the recognition of new opportunities to formulate relevant, targeted policy questions. As a case in point, efforts will also be devoted to help identifying new and more transient forms of migration[3] policy-makers are interested in, as a preliminary step towards the identification of the data gaps preventing their appropriate characterisation.

---

[2] C(2016) 6626 final, Communication to the Commission. Data, Information and Knowledge Management at the European Commission

[3] According to the standard definition of migrant recommended by the UN (see Glossary in the Annex), a migrant is a person who changes his or her country of 'usual residence'. For this change to take effect, a migrant has to move to the new country for a period of at least 12 months. There are various population groups for which some uncertainty may arise about their inclusion in the usually resident population of a country, but their classification is clarified in the UN international recommendations and/or in Eurostat technical notes. While this strict classification is needed to guarantee comparability and harmonization of the data across countries, it also results in the loss of information that may be of interest for policy-makers, e.g. persons who spend several months (but less than one year) in a country that is not their country of usual residence.

2 EU migration data landscape: gaps and challenges

2.1 Migration Data Issues and Opportunities

As in many statistical domains, an increased policy relevance of a topic is accompanied by a (much) larger data need. This is the case for migration as well, to which it must be added the intrinsic difficulty to monitor human mobility, especially in the contemporary world. For decades, statistical offices have been striving to improve availability and quality of migration data, particularly in the European Union where a set of EU statistical regulations adopted in the first decade of the new century has set a new benchmark in migration statistics. Since then, the data situation has constantly improved, thanks to coordinated efforts in the European Statistical Systems. Nonetheless, the increased relevance of migration in the policy agenda calls for further efforts, to expand both data offer and their quality. These needs are even stronger at global level. For instance, the New York Declaration for Refugees and Migrants of 19 September 2016[4] underlines the importance of improved data collection and capacity building. The Declaration has been followed up by the Global Compact for Safe, Orderly and Regular Migration[5], prepared under the auspices of the United Nations, which, although not legally binding, has a list of objectives and aims its participating states politically commit to. The Compact, which will be adopted in December 2018, recognises, as one of its objectives, the need to collect and utilise accurate data as a basis for evidence-based policymaking. Adequate data will also be necessary to measure progress towards migration-related targets of the Sustainable Development Goals. The International Organization for Migration (IOM) established in 2015 the Global Migration Data Analysis Centre (GMDAC) to compile, analyse and share data on international migration. In December 2017 GMDAC launched the Migration Data Portal[6], with the aim to serve as an access point for migration data globally. Still at the global level, the Organisation for Economic Co-operation and Development (OECD), IOM, and the United Nations Department of Economic and Social Affairs (UN DESA) jointly organised in January 2018 the first International Forum on Migration Statistics[7]. At the EU level, the European Commission launched the Knowledge Centre on Migration and Demography (KCMD) in 2016. Among many other actions, the Commission has also highlighted the need for more detailed and disaggregated data in, for example, its Communication on the protection of children in asylum migration[8]. Eurostat has also continued in its actions to improve the statistical concepts as well as data availability and quality; additionally, by stimulating with various initiatives (such as the 'Power from Statistics'[9] or the Budapest Memorandum[10]) and contributing to the overall strive towards this aim.

Appropriate data on migration is essential to inform policy-makers during all phases of the policy cycle. More detailed data and improved uses of the existing ones are needed to **better manage migration** (irregular and regular migration, asylum, protection, integration, border management, visa and return policies, Asylum Migration and Integration Funds).

Of key importance among the policy and operational needs of policy-makers is the understanding and awareness of movements of asylum seekers, regular and irregular migrants to and within the EU. Timely information on flows helps policy-makers in many areas, for instance to ensure rapid response and preparedness for emergencies, to preserve orderly and safe migration while securing borders, to improve recruitment and selection of future labour migrants, and to better monitor emigration. Richer socio-

---

[4] http://refugeesmigrants.un.org/declaration
[5] https://refugeesmigrants.un.org/migration-compact
[6] https://migrationdataportal.org/
[7] http://www.oecd.org/migration/forum-migration-statistics/
[8] COM(2017) 211 final, Communication from the Commission to the European Parliament and the Council. The protection of children in migration
[9] https://powerfromstatistics.eu/home
[10] https://circabc.europa.eu/sd/a/da7b77d3-395d-4d2f-ad6f-e20d65231c62/CENS(2017)05%20Budapest%20Memorandum.pdf

economic data and demographic indicators are also needed to improve education, social cohesion, welfare and labour integration policies at local, national, and, for some of those aspects, supranational level. More detailed data on countries of origin are also key to improve the understanding of the drivers of migration, and to better target development, humanitarian aid and neighbourhood policies to reduce migratory pressure to the EU. Better data and evidence are also of utmost importance to redress the narrative and public debate on migration.

To serve the aforementioned policy information needs, datasets on migration, asylum, and integration exist at national, regional and worldwide scales and comprise censuses, population registers, surveys, administrative and operational[11] data (for a comprehensive list see the KCMD Data Catalogue[12]). Nevertheless, gaps in terms of availability, quality, accessibility, discoverability[13] and harmonisation currently prevent unleashing the full potential of the data and the effective identification of areas where data is non-existent or incomplete.

Data unavailability and other limitations related to its inadequate geographic and temporal coverage, frequency of update or attributes (no information about sex or insufficient age disaggregation levels, for example) can be observed in many datasets, preventing the analysis and understanding of the new forms that migration is taking nowadays. For some policy needs, datasets with breakdown of migrants by age, sex, families with children, unaccompanied minors; with population and socio-economic projections for origin and transit countries with clearly documented migration assumptions; with information on effective timeframes to access healthcare, housing or employment in destination countries; to name a few, are either non-existent, non-accessible, fragmented or hard to find and use by the policy-makers.

The incomplete knowledge of data limitations may hinder the assessment of the quality of the data, necessary in order to associate the right level of uncertainty and reliability to the derived information. Accurate understanding of the data gaps is also essential to compare data collected in different ways and tackle harmonisation issues.

Data on migration are scattered across international and national institutions, EU Agencies, Commission services. Difficulties in discovering and accessing the data can be explained by restrictions of use (the data that may have been collected or provided by EU Member States under specific terms of reference), confidentiality or sensitivity. Nevertheless, in some cases, anonymization might help in using the data in a more integrated policy perspective when operational and/or political sensitivity of the data does not constitute an insurmountable obstacle to accessing the data.

There is a pressing need to translate policy information needs into data requirements, and to transform data into usable and actionable knowledge through data analysis. The latter is strongly restricted by the complexity and limitations of putting together datasets of different policy areas. A first–level solution to address this issue is offered by the KCMD Dynamic Data Hub[14].

---

[11] "Censuses" are official, comprehensive and periodic collections of data and information on a given population. "Population registers" are updatable accounts of persons within a given area, updated according to registration or de-registration events and maintained by local and/or national authorities. "Administrative data" are collected by government departments and other organizations for the purpose of implementing administrative procedures (e.g. residence permits, asylum or healthcare databases). "Operational data" are collected to manage and support operations in the field (e.g. Irregular Border Crossing data collected by Frontex, IOM Displacement Tracking Matrix etc.).

[12] https://bluehub.jrc.ec.europa.eu/catalogue

[13] Discoverability is the quality of a dataset to be found by the people who want to use it. It is related to, but differs from, accessibility, a term that refers either to the removal of barriers that prevent access to the dataset or to the facilitation of the access, for example via user-friendly interfaces.

[14] https://bluehub.jrc.ec.europa.eu/migration/app/index.html

2.2 Detailed analysis

During the second half of 2017, the KCMD, following consultation with relevant EU policy-makers (European Commission DGs, EU Agencies, EEAS, and EPSC), collected and analysed gaps in EU migration data. This section presents the main findings of this analysis.

Responses were received from officials working in the following **policy areas**: regular migration; integration of migrants in the job market, in the educational system, in rural areas; children in asylum migration; irregular migration; monitoring and analysis of migratory flows (to Europe and within it), including an assessment of the situation at the EU external border; asylum; forced displacement; development and migration governance; fundamental rights aspects related to migration; migration statistics; operational management of migration-related databases.

A wide variety of **data sources** are being used by EU policy-makers. The main categories are as follows: Eurostat data (statistics on migration flows and stocks, population, asylum, residence permits, irregular migration, return, integration indicators; EU LFS[15] and EU SILC[16] surveys); other EU data (Frontex, EASO, eu-LISA, IPCR platform); operational and administrative data collected by MS, as disseminated by the MS or by the relevant EU Agency; data from international organisations (UN Population Division, IOM, UNHCR, OCHA, UNICEF, World Bank, OECD, IDMC, Save the Children, HRW); own data from targeted surveys; data from existing networks (research networks, field networks); qualitative data (legislation, reports); open source data.

It was also noted that some relevant administrative data is currently being collected by MS but not used by EU policy-makers. Measurements or studies on integration were mentioned as examples. In most cases, these data are not made available by national authorities, and if they are, they tend to be un-harmonised and difficult to work with.

In terms of the **types of migration phenomena** being studied, both long-term international migration (as defined by the UN) as well as other more temporary forms of migration (e.g., circular[17], short-term, seasonal) are of interest. Regular, irregular and asylum flows to and within Europe receive policy attention, as do global forced displacement and migration crises. Significant emphasis is placed in countries or regions presently suffering from conflict or weak institutional governance and the migrants arriving to the EU from those countries.

The analysis of the **data gaps** was structured in four groups: gaps in existing data, dissemination issues, data not collected, and useful data that are currently inaccessible. These groups are presented next. For each group, the data gaps are described in more detail, alongside the challenges that need to be overcome to close the gaps.

2.2.1 Gaps in existing data

A number of gaps in existing data were identified. Such gaps prevent the full exploitation of the datasets and broadly relate to four categories: (1) timing; (2) data quality or inconsistencies, including in the definitions employed; (3) methodological issues; (4) inadequate aggregation level.

The most commonly reported **timing problem** was the long time lag between some events (or the collection of data on the events) and the publication of the data (e.g. on regular migration), which renders the data old by the time it is received by the policy-makers. This is despite the efficient data processing by the statistical agencies. In the specific case of statistical data on regular migration, the long time lag is due to the

---

[15] European Union Labour Force Survey
[16] European Union Statistics on Income and Living Conditions
[17] An internationally agreed definition of circular migration has recently been endorsed by the highest body of the European official statistics (the Conference of European Statisticians). Eurostat has contributed to this achievement and it is now working with volunteering countries to produce the very first internationally comparable data on circular migration.

current definition of migration (based on duration of stay of 12 months) and on the EU regulations in force, which result in the migration data being available at international level (e.g., Eurostat) one to two years after the migration happens. Different time lags for different data sources – the examples stated were for the number of arrivals, as given directly by some MS and by Frontex; and for humanitarian reports issued at different time periods - and time series that are too short – this relates for instance to EU residence permits statistics, where Eurostat's harmonised data only starts from around 2007-08 - were also mentioned as significant issues.

Doubts about the **quality of the data** or clear inconsistencies between sources were identified as transversal issues, affecting different policy areas. Some steps towards correcting these issues have already started. Agencies declared their efforts, in collaboration with MS, to tackle quality issues in their domains (e.g. double counting in EU aggregates of data on asylum applications). Specific remarks on the **quality of emigration flows** were given: reported figures often underestimate the magnitude of the real out-flows, thus making it difficult, among other things, to obtain accurate net migration figures as the difference of immigration and emigration. In this context, the need to understand the socio-political conditioning of the collection and reporting of migration data was expressed. This could explain some of the inconsistencies (e.g. low emigration figures).

Insufficient methodological clarification on the definitions (for data that are not collected according to EU regulation on migration[18]) constitutes the third category of data gaps, which also comprises comparability problems when data collection entities employ different procedures (e.g. differences in asylum procedures among MS in aspects like the distinction between 'making' and 'lodging' applications, as experienced by EASO: some MS make such distinction, others don't).

Finally, the **disaggregation level** of some variables in some datasets is sometimes seen as too coarse. Many of the remarks in this category refer to the impossibility to disassemble data for children. The option to disaggregate socio-economic variables by reason of migration would be valuable, but this information was only collected in the 2008 and 2014 ad-hoc modules of the EU LFS survey (this information will also be included in the 2021 survey and every two years thereafter). Breakdown by individual country of citizenship or country of birth in EU official immigration statistics is also missing for some Member States, as there is no legal obligation to provide such detail. A related issue is when variables are not crossed in a manner that would produce richer information for policy-makers. This comment referred to the variables citizenship (or country of birth) and previous/next country of residence in EU immigration or emigration data: these variables are not crossed and, therefore, it is only possible to understand the nature of the migratory flow by one of the variables at a time but not by several variables simultaneously. Again, this is the outcome of the discussion of the EU Member States at the time of the EU regulation on migration statistics and there is currently no legal obligation for providing such cross-tabulations. Thanks to the continued efforts of Eurostat, tables with broad tabulations by country of birth and citizenship are however available (still provided on voluntary basis).

A challenge common to many of these data gaps is the need to actively involve the Member States, be it because new data (or new categories in existing datasets) would need to be collected at national level to fill the gaps in insufficient disaggregation levels, because extra effort from MSs would be needed to meet the required timeliness, or because further harmonisation of their practices would be necessary to improve comparability even more. The implementation of these corrective measures is likely to involve high costs.

---

[18] Regulation (EC) No 862/2007 of the European Parliament and of the Council of 11 July 2007 on Community statistics on migration and international protection.

Another challenge relates to the fact that some of these data are collected by international organisations (e.g. IOM's DTM[19], UNHCR's and IDMC's data on forced displacement). The main purpose of much of this data is to answer local/national situations. Data collection methodology is therefore often tailored to local circumstances, which restricts data comparability across countries. Organisations are nonetheless continuously working on the improvement of their data (e.g. harmonisation), which would make the data more easily actionable for external actors (e.g. EU policy-makers).

2.2.2 Dissemination issues

Fewer issues were identified in the way the data is disseminated to the policy-makers. **Assumptions need to be clearly explained** in the dissemination material. Definitions/coverage/metadata are missing or incomplete in some datasets, for instance in Eurostat's residence permits issued to highly skilled workers (under national schemes), where there is no information on the schemes/definitions the permits correspond to. Additionally, some relevant data is disseminated by international organisations following country-dependent collection methodologies. This makes the interpretation of the data more difficult, although ongoing harmonisation efforts should improve the situation.

Computer file format in which the data is disseminated can also make a difference in the data's usability. In this regard, PDF is the least preferred format. New presentation of the data would be needed to increase its usability.

Some comments related to the visibility of certain information in Eurostat's Eurobase[20] dissemination platform were received: **summary tables for children** were requested, and Eurostat has already ensured their availability; the need to present some datasets with a higher level of aggregation was also indicated as a way to improve the user-friendliness of the platform and was exemplified by the aggregation in age groups (below 15 years old, 15-64, and over 65; or by 5-year age groups) for those datasets that contain age year by year; some relevant sub-categories are deemed hard to find or hidden, in particular the asylum-related residence permits, which are included in the generic 'other' category alongside some types of non-asylum-related permits (leading to a potential misinterpretation of the figures in 'other' as entirely asylum-related) and are not separable from the main table. On the other side, these needs for *ad-hoc* tabulations must be confronted against the criteria for an efficient dissemination. The proliferation of published tables with the same content presented in different aggregations - or, even more, the same aggregation under different headings - is not a viable solution when applied widely, especially in response to occasional needs.

Importantly, **dissemination policies of some of the Agencies** are under review or new policies have been recently agreed. Access to data collected by the MSs and held by the Agencies is of high importance for many Commission services and other Agencies. However, sensitivity challenges linked to legal terms, purpose of the data, etc. prevent the dissemination of the data to all interested EU policy-makers.

2.2.3 Data not collected

Stakeholders expressed policy interest in some datasets that are currently not always collected. Significant among these missing datasets is information on the **educational attainment of migrants and refugees**, with special emphasis on recently arrived asylum seekers/refugees. Educational attainment of migrants is collected and published based on EU LFS collection. Second generation migrants are currently covered in EU LFS ad-hoc module and will in the medium term (possibly in 2021) be collected in core EU LFS annually. For the longer-term (2024/5 onwards), Eurostat is developing a strategy

---

[19] Displacement Tracking Matrix http://www.globaldtm.info/
[20] http://ec.europa.eu/eurostat/data/database

for the collection of annual census-type population data. The collection of information on migration-related topics is a priority. This is expected to provide comprehensive information on the differing educational and employment situation of migrant populations (defined by country of birth, citizenship, duration of residence) as well as regional dimension. Important efforts have been already made by Eurostat to collect these data on a wider scale, but the MSs repeatedly express their concerns about the difficulty and quality related to cost, technical issues and environment of their collection. This knowledge would have impact on integration, education and employment policies. Collection of this data would necessarily be carried out by MS, which would need to provide the required training to their relevant staff (e.g. border operations). Additionally, efforts to define guidelines for harmonisation would be needed. All of these factors would certainly lead to high costs being incurred. In this respect, Eurostat involvement in the preparation of the International Recommendations on Refugee and Internally Displaced People (IDP) statistics is a good example of such efforts undertaken recently.

Data on minors (e.g. children in detention, previous school records, etc.) was also seen as insufficient and necessary to guarantee the protection of children in migration. Similarly, there is no systematic harmonised data collection on violence against (female) asylum seekers and refugees. The collection of this data would also need to be conducted by MS, leading to even greater challenges than for the education attainment data.

Data on **intra-EU mobility of third-country nationals** (TCN) is largely missing. Some data exists: statistics for mobile EU Blue Card holders and their families, a 2013 European Migration Network (EMN) report[21] with some quantitative data on the subject, and estimates based on EU LFS extractions. However, the existing data offers limited knowledge about the number of TCN moving between Member States. As mentioned in section 2.2.1, according to the EU regulation on migration statistics there is no legal obligation for MS to provide emigration and immigration statistics broken down by both citizenship and next/previous country of residence.

The importance of the (sub-national) **regional dimension** of migration, both in urban and rural areas, was also underlined, and the lack of data and analyses at that level, compared with at national level, was recognised. Closing this gap would require to put the existing data together and to carry out a notable analytical effort to exploit the existing data. Eurostat is working on the estimation of bilateral intra-EU migration flows at NUTS3 level by sex, and on collection of limited annual population census migration data on the 1 km^2 grid level, which would contribute to closing this gap. Also recently, Eurostat started to release infra-national statistics based on EU LFS on migrant integration and the work will be continued over 2018. For the long term, the annual census component mentioned above will close this gap. The OECD has also published a study[22] on the regional dimension of migration, mostly based on the Eurostat EU LFS survey. These developments stem from the Urban Agenda for the EU (Partnership on Inclusion of Migrants and Refugees) and efforts by REGIO, HOME, and Eurostat, together with OECD.

The publication of the single **country of birth information in the EU SILC & EU LFS surveys** (i.e. beyond the categories nationals, EU citizens and others) is not possible due to confidentiality reasons. The collection of this information is, however, undergoing improvement.

Migration status is not available in EU SILC (country of birth and citizenship are) and, therefore, income-related measures cannot be disaggregated by that status. It was suggested that large national and international surveys (e.g. EU LFS) collect data on **legal status**. This is however not planned, since it would not allow disseminating data comparable across countries. Collection, in line with national laws and guaranteeing data protection, of the **ethnicity** of respondents was also suggested. These last measures

---

[21] EMN Synthesis Report – Intra-EU mobility of third-country nationals
[22] OECD (2018), Working Together for Local Integration of Migrants and Refugees, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264085350-en

would allow a more detailed analysis of EU policies in the area of migration and discrimination. However several EU Member States cannot legally collect ethnicity information.

Need for more detailed data on the **external dimension of migration** (e.g. flows in transit countries, reintegration, limited geographical coverage of data on IDPs, insufficient data to understand the duration of displacements, incomplete data on displacements caused by slow-onset disasters, etc.) was also indicated.

2.2.4 Useful data that is currently inaccessible

This group includes the **data collected or originated by MS, held by EU Agencies**, and disseminated in aggregated form or as statistics only. Access to more disaggregated data was requested by several Commission services and Agencies. The dissemination of this data faces sensitivity obstacles linked to legal terms, purpose of the data, etc. Tackling these issues and building an environment of trust will be necessary to grant the dissemination of the data to EU policy-makers.

Similarly, some **administrative data collected by some MS** (e.g. data on recently arrived migrants, percentage of immigrants who have taken part in integration initiatives, longitudinal studies of integration measures) exists but is not available to EU policy-makers[23]. Access to national data on ethnicity, migration background, registers of asylum seekers, etc., valuable to reach a representative sample when conducting surveys targeted at the migrant population, is often restricted. The value of these targeted **surveys** was recognised, but the difficult access to the surveys was regretted. To try to mitigate this shortcoming, the KCMD, EASO, and FRA are members of the COST Action 16111 ETHMIGSURVEYDATA[24], whose aim is to improve the access, usability, dissemination and standards of the multiple and scattered survey data that exist on the economic, social and political integration of ethnic and migrant minorities.

The increasing importance for policy-making of **data on public opinion and perceptions** was highlighted. Some such opinion data exists (e.g. Gallup), but it demands intricate analysis. During 2017, the Migration Policy Centre (MPC) of the European University Institute (EUI) established the Observatory of Public Attitudes to Migration (OPAM), with the aim to produce comprehensive, pan-EU data and accounts of public attitudes to migration in Europe[25]. In October 2017, the KCMD, in collaboration with MPC and the International Centre for Migration Policy Development (ICMPD), organised the seminar 'Public attitudes on migration'.

Finally, **new sources of data and Big Data** (e.g. mobile phone records, social media data, etc.) exist and research suggests it could be used to fill some of the gaps of the traditional data. Access to this new data would need to be granted (there are commercial and/or privacy issues to solve) and substantial research and analysis are required before it can meet policy requirements. EMPL is currently conducting a study on the use of Big Data to estimate labour mobility and migration in the EU. The KCMD and Eurostat are also taking steps on the Big Data front.

---

[23] However the article 17a of the REGULATION (EU) 2015/759 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 29 April 2015 amending Regulation (EC) No 223/2009 on European statistics give legal basis to the NSIs, other national authorities and Eurostat the right to access and use, promptly and free of charge, all administrative records and to integrate those administrative records with statistics, to the extent necessary for the development, production and dissemination of European statistics.
[24] http://www.cost.eu/COST_Actions/ca/CA16111
[25] http://www.migrationpolicycentre.eu/opam/

3 Addressing the gaps and Science contribution

The previous section has presented an overview of the migration data landscape in Europe, detailing gaps and challenges from the perspective of EU policy-makers.

This section briefly describes some of the main courses of action to improve data on migration, with a special focus on how science and research can contribute to addressing some of the data gaps.

The first approach involves the **collection of new data and production of new statistics**. This is a time-consuming effort that usually requires long-term planning and careful evaluation and prioritisation of the data gaps to address based on a cost-benefit analysis, as well as consideration of existing challenges and policy-makers' needs. At EU level, data provision or reporting to the Commission (Eurostat) enshrined in EU legislation guarantees the highest possible quality, harmonisation and compliance among Member States. This process involves consultations with MS to make sure they can accommodate the new data requests. A current initiative to improve statistics on international protection using this approach is the Commission proposal to amend the regulation on statistics on migration and international protection[1]. This proposal is based on existing voluntary data collections, and aims to provide more detailed and timely statistics on asylum and managed migration. Among other aspects, the proposal intends to provide more information on unaccompanied minors applying for international protection; more frequent, timelier, and more detailed information on returns; and greater disaggregation (by age and sex) of the number of residence permits issued to third-country nationals. More importantly looking beyond these specific and immediate needs, and to support future developments in migration policy in general, the amendment will provide an embedded flexibility for further improvements to migration statistics. This will enable new and emerging data needs to be addressed more quickly.

Another initiative, currently in the drafting stage, is for the 2021 EU population census round to disseminate statistics on selected topics (including some migration-relevant variables) geocoded to a 1 km^2 grid[26]. Again, significant dialogue and preparatory work with MS is ongoing and results will be available after the 2021 census. Nevertheless this data will provide valuable information at the (sub-national) regional dimension in a harmonised way across the EU and allow other organisations to link with a variety of geospatial data. Eurostat, based on an endorsed strategy, is working with Member States on the new framework for population and migration statistics that includes the collection of annual census-type population data, with 2024 as first reference year. In particular, Eurostat is analysing the possibility of modifying the population base definition because the current 12 months usual residence requirement impacts negatively on the timelines of migration statistics.

Another way of collecting new data is through surveys targeted at population groups of interest. These surveys tend to be highly tailored to a specific task/population group. Although access to and analysis of the data can be laborious, these surveys provide a wealth of information about the population group. A good example is the special

---

[26] https://ec.europa.eu/info/law/better-regulation/initiatives/ares-2018-3255714_en

migration module of EU LFS that will be repeated for reference year 2021. Also the core EU LFS improves the quality of migration statistics as gradually more statistics are available broken down by citizenship and/or country of birth on regional level. Furthermore, from reference year 2021 the core EU LFS will include country of birth of mother / father and reason for migrating.

A second approach is to **better utilise existing data**, which may or may not have been collected for the specific purpose of informing policy-makers in the migration domain. An example of this is administrative data (e.g. data from population registers), i.e. data collected by government departments for various administrative procedures (e.g. residence permits, asylum, healthcare, or tax databases). Even if such data are not collected with the explicit aim of providing statistical information on the migration situation, they are increasingly being used also for statistical purposes. If appropriate administrative sources exist and suitable processes and methodologies are developed to obtain quality migration statistics from those sources, then this approach can result in more frequent and timelier migration statistics, with the added benefit of a reduced cost over alternative methods (e.g. large-scale surveys). The development of suitable administrative sources and methods could be however a long-term effort that may need substantial development timeframes before it can be completed. Across the EU, there are indeed large differences in the level of preparedness to this approach. Still, the trend towards an increased use of administrative data sources for the collection of population census data is clear. A specific example of existing administrative data that is currently underutilised is the information on region of origin (i.e. finer geographical specification than the country of origin) of the applicants for international protection. Even if this information is collected by some MS during the application, the data is not aggregated at EU level. A more detailed knowledge of the specific places of origin of the applicants could, for instance, be used to design more targeted development actions in those regions from which the highest number of applicants originate.

Better utilisation of existing data can also involve simple analytical processing of the data to produce new indicators (e.g. ongoing work by Eurostat to produce new indicators on children, like the share of children in the asylum applicants) or to facilitate the use of the data by improving its visibility (e.g. Eurostat summary tables on children in migration, already available in the cross cutting topics sections of Eurobase[20], or the data catalogue on children in KCMD's Dynamic Data Hub[14]).

A third approach, and where science and research can have the biggest impact, is in the **exploitation of new data sources and innovative methodologies** to extract insights on human mobility and migration. These sources provide information about users of commercial digital services like mobile phone communication services and social media networks. Even if those services have not been designed to provide any information about migration, the data generated by their users (so-called "digital breadcrumbs"), which often include geo-located information, coupled with new tailored ways to analyse the data, have been shown to deliver valuable knowledge on various migration aspects. For instance, a recent JRC report[27] presents a methodology to estimate the number of foreign-born individuals based on data from the Facebook advertising platform. Use of new data sources for migration is in its infancy. Major obstacles need to be overcome before this approach can be fully and systematically exploited: access to the data is not always granted as the data are normally in the hands of the service provider; confidentiality and protection of the information have to be guaranteed, and use of data for research or policy purposes needs to follow ethical practices; robust analytical methods to understand the data and to deal with existing biases in the data (e.g. sampling bias) need to be developed, including their evolution over time. The existence of these obstacles probably means that the utilisation of these data sources for migration will be gradual, that significant research effort will be required to capitalise on the

---

[27] Spyratos, S., Vespe, M., Natale, F., Ingmar, W., Zagheni, E. and Rango, M., Migration Data using Social Media: a European Perspective, EUR 29273 EN, Publications Office of the European Union, Luxembourg, 2018, ISBN 978-92-79-87989-0, doi:10.2760/964282, JRC112310.

potential of these sources, and that they will be complementary to the official statistics, rather than replace them. On the other hand, obvious advantages of new data sources, such as the availability of information on large segments of the population, including on hard-to-reach groups, or the timely insights that can be derived from such data, mean that the interest in and potential of these data in addressing existing migration data gaps are likely to increase in the future.

Some significant steps have already been taken along this approach. The KCMD and IOM's GMDAC co-organized the expert workshop "Big Data and alternative data sources on migration: from case-studies to policy support" in November 2017[28]. This was followed by the launch in June 2018 of the Big Data for Migration Alliance (BD4M)[29], jointly convened by KCMD and GMDAC, with the aim to promote dialogue and collaboration among policy-makers, international organisations, NGOs, data providers, national statistical offices, and researchers, thereby creating the conditions for the successful exploitation of big data sources for migration. The three main areas of work of the Alliance will be[30] a) awareness-raising and knowledge-sharing, b) capacity-building, and c) policy-oriented analysis.

Science can also provide new ways of elaborating and combining existing datasets. An example of this is a study by G. Abel and N. Sander[31], where the authors produced indirect estimates of global bilateral migration flows from UN data on migrant stocks. The role of science here is not only to provide the estimates, but to explain the methodology, highlighting caveats to consider when using the results, so that users of such estimates, including policy-makers, are well informed about their limitations.

---

[28] https://bluehub.jrc.ec.europa.eu/bigdata4migration/
[29] https://bluehub.jrc.ec.europa.eu/bigdata4migration/bd4m
[30] Big Data for Migration Alliance (BD4M). Harnessing the potential of new data sources and innovative methodologies for migration. Joint concept note. IOM's GMDAC and KCMD. Available at: https://bluehub.jrc.ec.europa.eu/bigdata4migration/uploads/attachments/cjio8ha96019f99zvr9xq6025-big-data-for-migration-alliance-concept-note.pdf
[31] Abel, G. and Sander, N. (2014). Quantifying Global International Migration Flows. Science, 343 (6178).

ANNEX - Glossary

Migrant: A person who moves to a country other than that of his or her usual residence for a period of at least a year (12 months), so that the country of destination effectively becomes his or her new country of usual residence. [32]

Source: United Nations United Nations Recommendations on International Migration Statistics, Rev. 1

Short-term migrant: A person who moves to a country other than that of his or her usual residence for a period of at least 3 months but less than a year (12 months) except in cases where the movement to that country is for purposes of recreation, holiday, visits to friends and relatives, business, medical treatment or religious pilgrimage[33].

Source: United Nations United Nations Recommendations on International Migration Statistics, Rev. 1

Additional definitions in Regulation (EC) No 862/2007:

(a) '**usual residence**' means the place at which a person normally spends the daily period of rest, regardless of temporary absences for purposes of recreation, holiday, visits to friends and relatives, business, medical treatment or religious pilgrimage or, in default, the place of legal or registered residence;

(b) '**immigration**' means the action by which a person establishes his or her usual residence in the territory of a Member State for a period that is, or is expected to be, of at least 12 months, having previously been usually resident in another Member State or a third country[34];

(c) '**emigration**' means the action by which a person, having previously been usually resident in the territory of a Member State, ceases to have his or her usual residence in that Member State for a period that is, or is expected to be, of at least 12 months;

(d) '**citizenship**' means the particular legal bond between an individual and his or her State, acquired by birth or naturalisation, whether by declaration, choice, marriage or other means according to national legislation;

(e) '**country of birth**' means the country of residence (in its current borders, if the information is available) of the mother at the time of the birth or, in default, the country (in its current borders, if the information is available) in which the birth took place;

(f) '**immigrant**' means a person undertaking an immigration;

(g) '**emigrant**' means a person undertaking an emigration;

(h) '**long-term resident**' means long-term resident as defined in Article 2(b) of Council Directive 2003/109/EC of 25 November 2003 concerning the status of third-country nationals who are long-term residents;

Source: Regulation (EC) No 862/2007 of the European Parliament and of the Council of 11 July 2007 on Community statistics on migration and international protection.

---

[32] This is the definition currently recommended by the UN, although it may be subject to changes, in the long term, as part of broader efforts to redefine the architecture for population statistics.

[33] According to the UN recommendations, these types of temporary travel abroad are excluded from the definition of migrant because they do not entail a change in country of usual residence.

[34] It has to be noted that this definition includes intra-EU migration between Member States

Irregular migrant: In the global context, a person who, owing to irregular entry, breach of a condition of entry or the expiry of their legal basis for entering and residing, lacks legal status in a transit or host country. In the EU context, a third-country national present on the territory of a Schengen State who does not fulfil, or no longer fulfils, the conditions of entry as set out in the Schengen Borders Code, or other conditions for entry, stay or residence in that Member State.

Source: European Migration Network

Refugee: means a third country national who, owing to a well-founded fear of being persecuted for reasons of race, religion, nationality, political opinion or membership of a particular social group, is outside the country of nationality and is unable or, owing to such fear, is unwilling to avail himself or herself of the protection of that country, or a stateless person, who, being outside of the country of former habitual residence for the same reasons as mentioned above, is unable or, owing to such fear, unwilling to return to it, and to whom Article 12 (of the Directive 2011/95/EU) does not apply[35].

Source: Article 2 (d) of Directive 2011/95/EU on standards for the qualification of third-country or stateless persons as beneficiaries of intentional protection, for the uniform status for refugees or for persons eligible for subsidiary protection, and for the content of the protection granted (Qualification Directive).

---

[35] Article 12 details the people who are excluded from being a refugee. In brief, they are people for whom there are serious reasons for considering that they have committed, incited, or participated in the commission of a crime against peace, a war crime, a crime against humanity, a serious non-political crime outside the country of refuge prior to his or her admission as a refugee, or acts contrary to the purposes and principles of the United Nations It also excludes people who are receiving protection or assistance from organs or agencies of the UN other than UNHCR, and people who are recognized by the competent authorities of the country in which they have taken residence as having the rights and obligations which are attached to the possession of the nationality of that country.

List of abbreviations

COST	European Cooperation in Science and Technology
DTM	Displacement Tracking Matrix
EASO	European Asylum Support Office
EEAS	European External Action Service
EMN	European Migration Network
EPSC	European Political Strategy Centre
ETHMIGSURVEYDATA	International Ethnic and Immigrant Minorities' Survey Data Network (a COST action)
EU LFS	European Union Labour Force Survey
EU SILC	European Union Statistics on Income and Living Conditions
EUI	European University Institute
eu-LISA	European Agency for the operational management of large-scale IT systems in the area of freedom, security and justice
FRA	European Union Agency for Fundamental Rights
HRW	Human Rights Watch
ICMPD	International Centre for Migration Policy Development
IDMC	Internal Displacement Monitoring Centre
IDP	Internally Displaced People
IOM	International Organization for Migration
IOM's GMDAC	IOM's Global Migration Data Analysis Centre
IPCR	EU's Integrated Political Crisis Response
JRC	Joint Research Centre
KCMD	Knowledge Centre on Migration and Demography
MPC	Migration Policy Centre
MS	Member State (used in this report to refer to EU Member States only)
NGO	Non-governmental organization
NUTS	Nomenclature of Territorial Units for Statistics
OCHA	United Nations Office for the Coordination of Humanitarian Affairs
OECD	Organisation for Economic Co-operation and Development
OPAM	Observatory of Public Attitudes to Migration
PDF	Portable Document Format
TCN	Third-country national
UN DESA	United Nations Department of Economic and Social Affairs
UNHCR	United Nations High Commissioner for Refugees
UNICEF	United Nations Children's Fund

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: http://europea.eu/contact

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

- at the following standard number: +32 22999696, or

- by electronic mail via: http://europa.eu/contact

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: http://europa.eu

EU publications
You can download or order free and priced EU publications from EU Bookshop at: http://bookshop.europa.eu. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see http://europa.eu/contact).

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

Joint Research Centre

EU Science Hub

Publications Office