



SOCIAL MEDIA

ARTIFICIAL
INTELLIGENCE/
MACHINE LEARNING

OTHER

FORECASTING
HUMAN MOBILITYTRACKING HUMAN
DISPLACEMENT

AN INNOVATIVE FRAMEWORK FOR ANALYSING ASYLUM-RELATED MIGRATION

Constantinos Melachrinou,¹ Marcello Carammia² and Teddy Wilkin³

Introduction

Asylum-related migration comprises those migrants who left their countries of origin with the intention to seek international protection in another country or ended up doing so. People with legitimate needs for international protection, according to the 1951 Refugee Convention and its 1967 Protocol, and the Qualification Directive (2011/95/EU) (European Union, 2011), as well as people fleeing their countries for better economic/employment opportunities are included, as long as they (intend to) seek asylum. Moreover, everyone has the right to apply for asylum, irrespective of whether they arrived in any given country regularly or irregularly.

Asylum-related migration is therefore a complex process and one difficult to typify. Yet understanding migration processes and their drivers is key to evidence-based policymaking, while some capacity to anticipate migration flows (to the extent possible, see: Carammia and Dumont, 2018) is crucial to the design of efficient⁴ asylum systems. This chapter presents the European Union Agency for Asylum (EUAA) approach to the analysis of asylum-related migration. It first provides a description of the asylum data exchanged under the EUAA Early warning and Preparedness System (EPS). Subsequently, it describes the EUAA analytical and conceptual framework, discussing the analytical purpose, temporal dimension, resources required, and the stakeholders that define the target audience for the EUAA analytical products.

¹ Constantinos Melachrinou is Senior Analyst at the European Union Agency for Asylum (EUAA). He holds an undergraduate degree in Physics from the University of Chicago and a PhD in Physics from the Massachusetts Institute of Technology.

² Marcello Carammia is Senior Researcher at the Department of Political and Social Sciences of the University of Catania, Italy. Marcelo is a founding co-director of the Italian Agendas Project and a principal investigator of the European Union Policy Agendas Project. Previously he worked as Senior Researcher at the European Asylum Support Office (EASO) (now EUAA).

³ Teddy Wilkin is Head of the Data Analysis and Research Sector at EUAA. Previously he was a Research Fellow at the University of Oxford. He also holds a PhD from Oxford.

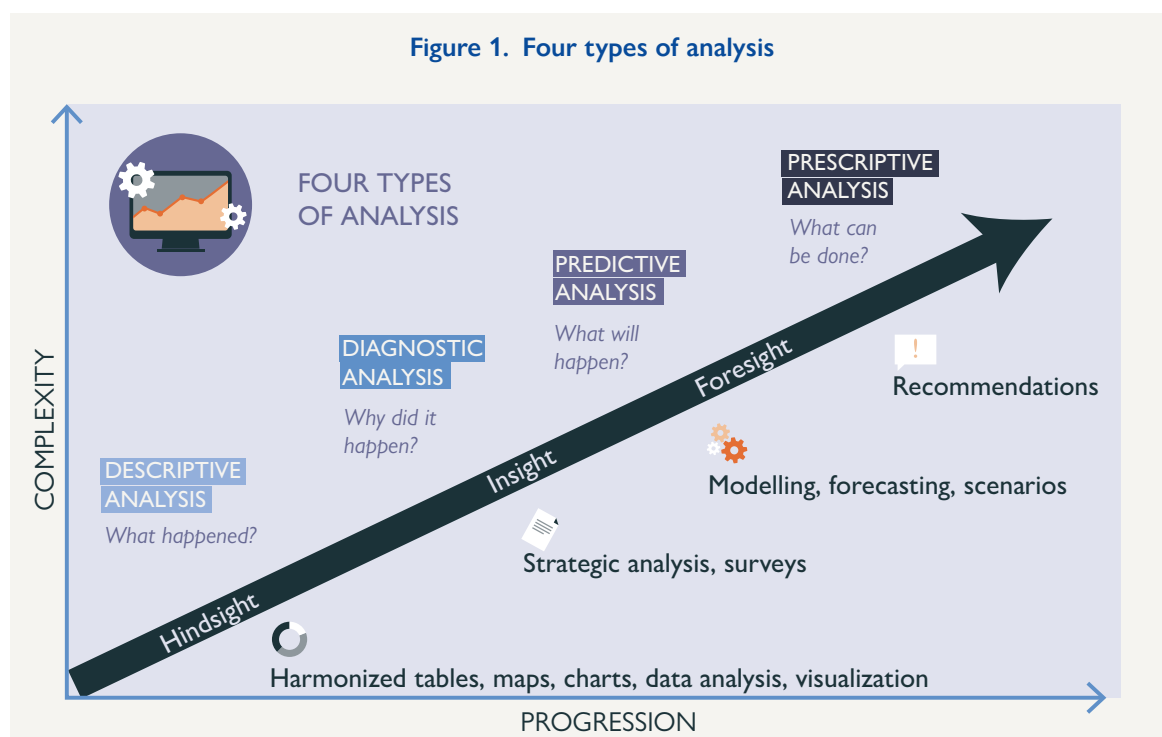
⁴ In this context, efficiency is not about cost savings, but about foreseeing sudden increases in caseload to allow countries to have resources in place to process each application quickly and in line with international standards.

Early warning and Preparedness System (EPS) data

In 2012, to better understand asylum-related migration to the European Union, Norway and Switzerland (hereafter referred to as EU+ countries), EUAA⁵ launched its EPS, an information-exchange mechanism aimed at providing EUAA, European Union member States, Schengen associated countries and the European Commission with timely, accurate, and comparable data on the lodging and processing of asylum applications in EU+ countries, in line with the Common European Asylum System.⁶

Quantification of complex real-world phenomena invariably involves a degree of simplification. Hence, EPS data are extensive and have been shown to add much value, but they also have their limitations. Consistent with most data originating from administrative bodies, EPS data tend to count and describe administrative procedures rather than actual people, plus at the supranational level, it is not yet possible to follow individual asylum cases as they progress through (and often between) convoluted asylum procedures,⁷ nor is it possible at the EU+ level to link individual asylum applications to other administrative procedures such as regular (e.g. visa applications, resettlement) or irregular migration (e.g. refusals of entry, detections of illegal border-crossing or illegal stay).

That said, EPS data are the cornerstone of the EUAA analytical framework, and much intellectual enterprise has been spent designing the data and making sure that the indicators add as much value as possible to the four types of analytics – descriptive, diagnostic, predictive and prescriptive – as well as addressing different time horizons, including nowcasting and monitoring, early warning, and forecasting.



⁵ Before January 2022, this was known as EASO.

⁶ See: EASO, 2011.

⁷ Interoperability Regulations (EU) 2019/817 and 2019/818, and proposals of the European Union Pact on Migration and Asylum 2020, take significant steps to mitigate this data limitation.

Nowcasting

The assessment of the asylum situation relies largely on backward-facing analyses using EPS indicators, data shared by partner organizations, and publicly available information. EPS indicators include data on all stages of the asylum procedure,⁸ including access to procedure, reception systems, first instance determination, returns and the Dublin Regulation. This assessment provides for a retrospective understanding of events and uses the four types of analytics to speculate about their drivers. Embedded in this logic is the assumption that an effective description of the recent past can also say something about the present and our expectations for the near future (Banbura et al., 2010). Repeated and regular analyses are a crucial part of any analytical framework because reoccurring trends and stable drivers can underpin predictive analytics, or at least can be used to corroborate their outputs. At EUAA, weekly, monthly and biannual analyses of the current asylum situation are systematically produced, and the findings communicated to internal and external stakeholders, including European Union and national policymakers, to guide their response to recent and ongoing events. These analyses take the form of descriptive and diagnostic analytics, as mentioned in the previous section, where past events are described and interpreted through the prism of strategic analysis. For example, monitoring the situation in the same or neighbouring destination countries over time, or comparing different countries during the same period of time, or a combination of the two approaches, can elucidate why the asylum situation in EU+ countries is what it is, leading to valuable inputs for predictive and prescriptive analytics.

Analyses connect current and past trends. While past events may provide a guide and a framework for future expectations, it is usually the case that drivers of asylum-related migration vary between individual migration flows or events, and can even be so short-lived that drivers vary within a single migration flow. For example, increases in conflict events in the Syrian Arab Republic may lead to an increase of asylum-related migration into neighbouring Türkiye, but not to EU+ countries, whereas increases in conflict events in Libya may lead to increases in asylum-related migration directly into EU+ countries (i.e. into Italy or Malta). On the other hand, before the European Union–Türkiye statement, conflict events in the Syrian Arab Republic led directly to higher numbers of incoming asylum applicants in EU+ countries, which is not the case anymore. As such, the results of any analyses should be interpreted along with the caveat that the present and future situations may not always be explainable using the recent past, especially in cases of extreme shocks to the system. For example, the ongoing COVID-19 pandemic and the subsequent travel restrictions and generally reduced mobility have resulted in fewer asylum applications being lodged in EU+ countries despite the fact that more people globally are in need of international protection, and it has disrupted many long-standing migration patterns, such as the Global South to North and the rural–urban spectrums.

⁸ More information on the EUAA's Early warning and Preparedness System indicators is available at www.euaa.europa.eu/asylum-knowledge/data-analysis-and-research.

Monitoring

As a first step towards early warning of asylum-related migration, EUAA has developed a system that continuously monitors the administrative data (EPS indicators), as well as novel sources of data, such as Internet searches for asylum-related migration topics (Google, n.d.) and potential “push factor” events in countries of origin of asylum seekers, extracted from event-based databases using global media reports, such as the GDELT Project (n.d.) and the Armed Conflict Location and Event Dataset (Raleigh et al., 2010).

Administrative data, such as the number of asylum applications lodged in each EU+ country, broken down by citizenship, may exhibit trends that allow for effective nowcasting of the overall asylum situation. The proportion of asylum decisions that grant (as opposed to refuse) international protection (European Union-regulated recognition rate) in turn is positively associated with asylum-related migration (Toshkov, 2014; Brekke et al., 2017), assuming that migrants prefer to go to destinations where they are more likely to be granted international protection and that this information is available to them; it can also deflect potential asylum seekers into irregularity, based on an analysis of bilateral asylum and visa policies on migrant flows to 29 European States in the 2000s (Czaika and Hobolth, 2016).

Internet searches for asylum-related topics, in addition to searches for specific countries of transit or destination, are downloaded and monitored by EUAA because they have been shown to indicate the intentions of potential asylum seekers to migrate (Böhme et al., 2020). Monitoring Internet searches and attributing intentions of migration should be performed with caution since the link may not always exist, or intentions to migrate may not possibly materialize due to extraneous factors. Algorithms behind Internet searches may also affect the results, necessitating a thorough understanding of the underlying methodologies, before venturing into this kind of analysis.

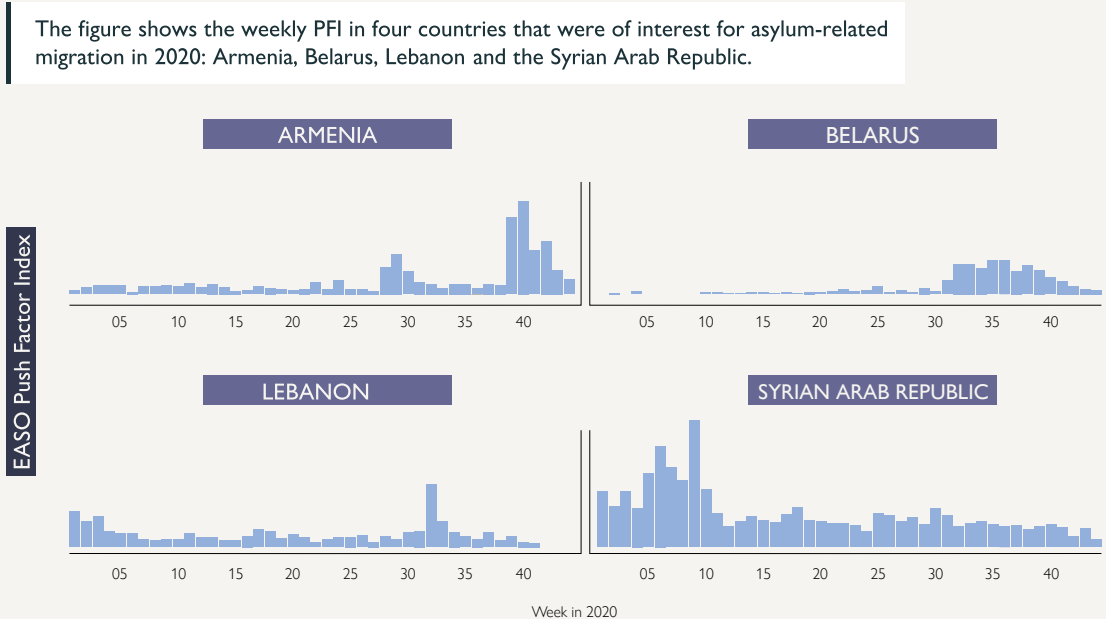
The GDELT Project extracts from global media near to real-time data on events around the world, classified automatically into about 200 categories (Gerner et al., 2002). EUAA has identified which events have the potential to generate asylum-related migration and aggregated them into five macrocategories – conflict, economic, political, governance and social – reflecting broad drivers or push factors in the countries of origin/transit of potential asylum seekers. The selected events are also weighed according to their severity and relevance to migration to construct the Push Factor Index (PFI). The PFI is updated daily and visualized in interactive dashboards, representing in a snapshot the situation in the countries of origin, a tool that is useful for migration practitioners in the field, as well as for policymakers. One advantage of the GDELT Project is the ability to monitor the situation in countries of origin and transit at the subnational level, since each event is automatically attributed to a specific geographical location. The fact that geolocation and attribution to Conflict and Mediation Event Observations (CAMEO) categories is performed automatically by the GDELT Project algorithms raises the need to critically evaluate these elements and understand potential limitations or biases (Hammond and Weidmann, 2014). For example, when a location is not provided in an article, an event is often attributed to the capital of the country where the event took place, even if the event took place somewhere else, creating a “capital city” bias in these results.



Clearly, the relationship between events in countries of origin and asylum-related migration is not straightforward: applying for asylum in EU+ countries is only one of the possible outcomes of crises in third countries. However, monitoring the PFI does provide important insights into migration processes, and it also bears the potential for early warning purposes. A recent study (Melachrinou et al., 2020) has shown that the PFI in countries in Africa was correlated with the number of asylum applications by nationals to EU+ countries during the years 2016 and 2017. In subsequent years (2018–2019), the relationship is less strong, possibly because of fewer search and rescue operations in the Central Mediterranean route and increased interceptions by the Libyan Coast Guard and Navy of migrants attempting to cross the route from Libya to Italy, one of the most important routes into Europe for African nationals seeking international protection in the European Union.

Figure 2 shows the weekly PFI in four countries that were of interest for asylum-related migration in 2020: Armenia, Belarus, Lebanon and the Syrian Arab Republic. The recent conflict in Artsakh/Nagorno–Karabakh, which restarted on 27 September 2020 (BBC News, 2020a), is reflected in a corresponding peak of PFI. Push factors have decreased since the agreed ceasefire, but at the time of writing, it was not known whether this event would lead to increased numbers of asylum applications in EU+ countries by Armenian citizens or whether they will be internally displaced or seek asylum in neighbouring countries. Other recent examples where migration drivers have intensified in countries of origin, and have resulted in an increase in PFI, include the blast in Beirut in August 2020 (Hubbard and El-Naggar, 2020), the contested elections in Belarus (Al Jazeera, 2020) and subsequent demonstrations in August/September 2020. Equivalently, reductions in migration drivers, such as those caused by the recent agreement for a ceasefire in the Syrian Arab Republic in March 2020 (BBC News, 2020b), are reflected in a decrease in PFI and stabilization at lower levels, which may lead to lower numbers of asylum seekers in the European Union.

Figure 2. The Push Factor Index in Armenia, Belarus, Lebanon and the Syrian Arab Republic weekly in 2020



Source: EUAA and the GDELT Project.

Early warning alerts

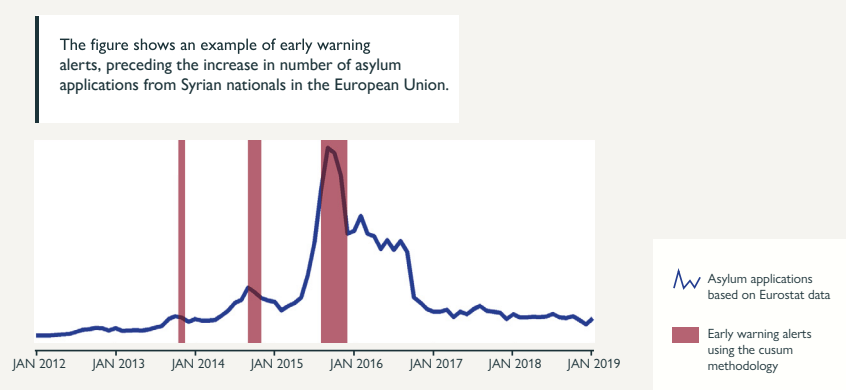
While monitoring is useful to discover trends and observe events unravelling more or less as they occur, it still depends on trained analysts to visualize the data, manually assess the asylum-related situation, and place trends in the correct context. Alternatively, the cumulative sum approach is an automated method that generates early warning alerts based on statistical control theory (Bijak et al., 2017).

The cumulative sum (or “cusum”) method, introduced by E.S. Page (1954), monitors the cumulative sum of some function of the observed data – in this case, the number of asylum applications. Hence the cumulative sum can generate an early warning alert when it exceeds a certain threshold which can be set based on the past performance of the indicator or by using a formal model.

The acceptance thresholds to trigger alerts depend on a moving average window of the latest data available for each country of origin. Single countries of origin have different “background” levels of conflicts and other potential migration-generating events, have different patterns of Internet searches, generate different volumes of asylum applicants, and so forth. For this reason, fixed thresholds may result in inconsistent “false positive” (Type I error) alarms or “false negative” (Type II error) outcomes. The trade-off between false positives and false negatives ultimately depends on policy considerations and the tolerance towards each of the two errors.

Figure 3 shows a time series of the number of asylum applications lodged by Syrian nationals in EU+ countries between 2012 and 2018, based on Eurostat data. The early warning alerts, shown in shaded areas, were triggered using the cusum methodology, often preceding a subsequent increase in the number of asylum applications. Different methods for alerts, such as the statistical significance of the distance from a moving average, would generate different alerts, and so all such methods are to some extent subjective. The sensitivity of the alerts and the extent to which there are false positives depend on the activation threshold for triggering alerts. This is a preliminary application of the method, adapted from Napierała et al. (2021), with improvements planned to enable earlier and more relevant alerts.

Figure 3. Example of application of cusum to data on the monthly number of asylum applications lodged in the European Union by Syrian nationals (Eurostat data), with early warning alerts triggered (shaded areas)



Cusum is one possible way of monitoring time series, which EUAA combines with other approaches such as the analysis of short-term and medium-term moving averages in the series borrowed from financial econometrics (Murphy, 1999).

While individual early warning alerts are useful to capture the attention of analysts and point them to specific time series that exhibit sharp changes from their expected values, the concurrent change in a number of early warning alerts for specific countries can highlight that more than one time series are affected by recent events. For example, an economic embargo in a country of origin can increase the economic drivers, increasing the PFI sharply and generating an early warning alert, which is complemented by an early warning alert elicited by Internet searches for asylum and visa in a specific European Union country or country of transit. This example illustrates that not only the migration drivers but also the intentions to migrate have increased, and the system is able to automatically generate such early warning alerts for further investigation by analysts and subsequent action by policymakers.

Forecasting

Asylum-related migration is a complex process driven by a multitude of factors which themselves depend on external factors – such as climate, political events and crises, and State actions such as policy (asylum law, migration management) – and ultimately on individual actions such as the decision to migrate (Bijak, 2011; Bijak et al., 2017). EUAA has recently developed an early warning and forecasting system (EPS–Forecasting) which uses an adaptive elastic-net model that combines data on events and Internet searches in third countries, irregular crossings at the European Union border, and recognition rates in EU+ countries. By training the model in the recent past, short-term forecasts of bilateral asylum-related migration flows, indicated by the number of asylum applications by nationals from specific countries of origin to European Union member States, are generated. Further information about the EUAA early warning and forecasting system can be found in the article by Carammia et al. (2022).

Requirements

Stakeholders using the EUAA analyses to guide their response to the asylum situation have different requirements that stem from their work mandates. These stakeholders range from operations, asylum and reception services in member States, to policymakers at the European Union and national levels. Their varying requirements are outlined below:

- (a) Operations need short-term information to support decision-making at the tactical level (i.e. where to deploy human and material resources to address challenges imposed by sudden changes in asylum-related migration). These include border guards, registration officers and vulnerability experts.
- (b) Asylum and reception services have planning requirements for the medium term. They seek this information in order to be able to employ the most appropriate number of caseworkers to process asylum applications, to ensure that they have the right capacity to accommodate asylum applicants and more.



- (c) Policymakers need to respond in longer time periods, following stable European Union or national-level trends. Data used by policymakers should be integrated so that strong policy developments in one area do not cause displacement effects to a weaker policy in another area. For example, the adoption of a list of “safe” countries of origin in one European Union member State may drive applicants from those safe countries to seek asylum in another member State, leading to an increase in secondary movements within the European Union.

The different requirements among the stakeholders give rise to the need for different outputs, which consider the analyses performed to address the asylum situation from the perspective of these stakeholders.

Outputs

To address the challenges faced by operational colleagues, short-term forecasts and early warning alerts are needed. These outputs, included in early warning reports disseminated to operational stakeholders in a timely manner, allow for an understanding of the current situation and how it is expected to evolve in the following few weeks, with the aim of improving operational planning. Asylum and reception services employ qualitative reports, which analyse whether current trends are likely to continue in the medium term. These reports improve planning for human resources and accommodation for asylum applicants. Policymakers, on the other hand, need to understand the long-term trends in migration drivers and the relationships between those drivers. Investigating the causal relationships between migration drivers and administrative indices allows policymakers to adjust policies in a way that would improve the asylum processes to ensure safe and efficient access to international protection for those in need, while at the same time also alleviating the strain of sudden large increases in the number of asylum applicants on national reception systems.

Resources

Considering these requirements by stakeholders, it is imperative that results are communicated to them in a timely manner. However, the dissemination of raw quantitative results is discouraged due to the high risk of misinterpretation and overconfidence in quantitative approaches. Data science, the science of uncovering insights from data, can generate meaningful results, which, however, need to be validated by domain-expert analysts. Only after this validation and interpretation of the results by analysts can the main findings be simplified and shared with stakeholders. This validation process also ensures that any caveats or issues introduced due to missing data or inconsistencies arising from data reporting by different countries can be communicated clearly alongside the main findings. Since the results of such analyses can have an impact on the availability of national or international forms of protection, it is important to include this additional layer of interpretation before the results can be used to affect operational resource planning or policymaking (Albertinelli et al., 2020).

Conclusions

Policymakers and operational, asylum, and reception services employ the analysis of asylum-related migration developed by EUAA to understand the main drivers, as well as design or adjust their response to ever-changing trends. As a result, the output portfolio of EUAA has been designed to address a variety of challenges, with different time horizons and utilizing different methodologies depending on the requirements of each stakeholder. The need for domain-expert analysts to validate and filter the raw results of the data science analysis presented earlier is important, especially since the reports may be used to alter policies affecting human lives.

REFERENCES*

- Albertinelli, A., P. Alexandrova, C. Melachrinou and T. Wilkin
2020 Forecasting asylum-related migration to the European Union, and bridging the gap between evidence and policy. *Migration Policy Practice*, X(4):35–41. Available at <https://publications.iom.int/books/migration-policy-practice-vol-x-number-4-september-december-2020>.
- Al Jazeera*
2020 Tens of thousands hold biggest protest yet, Lukashenko defiant. 17 August. Available at www.aljazeera.com/news/2020/8/17/tens-of-thousands-hold-biggest-protest-yet-lukashenko-defiant.
- Banbura, M., D. Giannone and L. Reichlin
2010 Nowcasting. *ECB Working Paper No. 1275*. Available at <https://ssrn.com/abstract=1717887>.
- BBC News*
2020a Armenia and Azerbaijan fight over disputed Nagorno–Karabakh. 28 September. Available at www.bbc.com/news/world-europe-54314341.
2020b Syria war: Russia and Turkey agree Idlib ceasefire. 5 March. Available at www.bbc.com/news/world-middle-east-51747592.
- Bijak, J.
2011 *Forecasting International Migration in Europe: A Bayesian View*. Springer, Dordrecht. Available at <https://link.springer.com/book/10.1007/978-90-481-8897-0>.
- Bijak, J., J.J. Forster and J. Hilton
2017 *Quantitative Assessment of Asylum-Related Migration: A Survey of Methodology*. European Asylum Support Office, Malta. Available at <https://data.europa.eu/doi/10.2847/642161>.
- Böhme, M.H., A. Gröger and T. Stöhr
2020 Searching for a better life: Predicting international migration with online search keywords. *Journal of Development Economics*, 142.
- Brekke, J.P., M. Røed and P. Schøne
2017 Reduction or deflection? The effect of asylum policy on interconnected asylum flows. *Migration Studies*, 5(1):65–96.

* All hyperlinks were working at the time of writing this report.

- Carammia, M. and J.C. Dumont
 2018 Can we anticipate future migration flows? OECD Migration Policy Debates No. 16, May. Available at www.oecd.org/els/mig/migration-policy-debate-16.pdf.
- Carammia, M., S.M. Iacus and T. Wilkin
 2022 Forecasting asylum-related migration flows with machine learning and data at scale. *Scientific Reports*, 12:1457.
- Czaika, M. and M. Hobolth
 2016 Do restrictive asylum and visa policies increase irregular migration into Europe? *European Union Politics*, 17(3):345–365.
- European Asylum Support Office (EASO)
 2011 Work Programme 2012. EASO/MB/2011/25. Available at https://euaa.europa.eu/sites/default/files/easo_2011_00110000_en.pdf.
- European Union
 2011 Directive 2011/95/EU of the European Parliament and of the Council of 13 December 2011 on standards for the qualification of third-country nationals or stateless persons as beneficiaries of international protection, for a uniform status for refugees or for persons eligible for subsidiary protection, and for the content of the protection granted (recast). *Official Journal of the European Union*. Available at <https://eur-lex.europa.eu/eli/dir/2011/95/oj>.
- The GDELT Project
 n.d. Intro – Watching our world unfold: A global database of society. Available at www.gdeltproject.org/ (accessed November 2020).
- Gerner, D.J., A.J. Rajaa, P.A. Schrodtt and Ö. Yilmaz
 2002 Conflict and Mediation Event Observations (CAMEO): A new event data framework for the analysis of foreign policy interactions. Paper presented at the International Studies Association, New Orleans, March.
- Google
 n.d. Google Trends. Available at <https://trends.google.com/> (accessed November 2020).
- Hammond, J. and N.B. Weidmann
 2014 Using machine-coded event data for the micro-level study of political violence. *Research and Politics*, 1(2).
- Hubbard, B. and M. El-Naggar
 2020 Clashes erupt in Beirut at blast protest as Lebanon’s anger boils over. *The New York Times*, 8 August. Available at www.nytimes.com/2020/08/08/world/middleeast/Beirut-explosion-protests-lebanon.html.

Melachrinou, C., M. Carammia and T. Wilkin

- 2020 Using big data to estimate migration “push factors” from Africa. In: *Migration in West and North Africa and across the Mediterranean: Trends, Risks, Development and Governance* (Fargues, F., M. Rango, E. Börgnas and I. Schöfberger, eds.). IOM, Geneva, pp. 98–116. Available at <https://publications.iom.int/books/migration-west-and-north-africa-and-across-mediterranean-chapter-8>.

Murphy, J.J.

- 1999 *Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications*. New York Institute of Finance, Paramus. Available at <https://cdn.preterhuman.net/texts/unsorted2/Stock%20books%20029/John%20J%20Murphy%20-%20Technical%20Analysis%20Of%20The%20Financial%20Markets.pdf>.

Napierała J., J. Hilton, J.J. Forster, M. Carammia and J. Bijak

- 2021 Toward an early warning system for monitoring asylum-related migration flows in Europe. *International Migration Review*, 56(1):33–62.

Page, E.S.

- 1954 Continuous inspection schemes. *Biometrika*, 41(1–2):100–115.

Raleigh, C., A. Linke, H. Hegre and J. Karlsen

- 2010 Introducing ACLED: An Armed Conflict Location and Event Dataset. *Journal of Peace Research*, 47(5):651–660.

Toshkov, D.D.

- 2014 The dynamic relationship between asylum applications and recognition rates in Europe (1987–2010). *European Union Politics*, 15(2):192–214.

ADDITIONAL READING

Acostamadiedo, E., R. Sohst, J. Tjaden, G. Groenewold and H. de Valk

- 2020 *Assessing Immigration Scenarios for the European Union in 2030: Relevant, Realistic and Reliable?* IOM, Geneva, and the Netherlands Interdisciplinary Demographic Institute, The Hague. Available at <https://publications.iom.int/books/assessing-immigration-scenarios-european-union-2030>.

Harrison, P.J. and P.P. Veerapen

- 1994 A Bayesian decision approach to model monitoring and cusums. *Journal of Forecasting*, 13(1):29–36.