Harvard Data Science Review • Issue 4.1, Winter 2022

Data, Design, and Deep Domain Knowledge: Science-Policy Collaboration to Combat Misinformation on Migration and Migrants

Marie McAuliffe^{1,2,3,4} Guy Abel^{5,6} Adrian Kitimbo^{7,8}
Jose Ignacio Martin Galan^{7,9}

¹Migration Research Division, International Organization for Migration, United Nations, Grand-Saconnex, Switzerland,

²School of Demography, Australian National University, Canberra, Australia,

³Global Migration Centre, Graduate Institute, Geneva, Switzerland,

⁴Center for Strategic and International Studies, Washington DC, United States of America.,

⁵International Migration Research Group, Institute for Asian Demographic Research, Shanghai University, Baoshan, Shanghai, China,

⁶School of Sociology and Political Sciences, Shanghai University, Baoshan, Shanghai, China,

⁷International Organization for Migration, United Nations, Grand-Saconnex, Switzerland,

⁸Gordon Institute of Business Science, University of Pretoria, Pretoria, South Africa,

⁹The Interuniversity Institute for Social Development and Peace, United Nations Educational Scientific and Cultural Organization, Jaume I University, Castelló de la Plana, Castelló, Spain

Published on: Jan 27, 2022

DOI: <https://doi.org/10.1162/99608f92.b3353b93>

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

ABSTRACT

In today's data-rich societies there is a strong tendency to assess and analyze complex issues through quantitative methods utilizing new, and rapidly evolving and constantly expanding, user-generated data. While new data and new data science present enormous opportunities for innovation and scholarship, we are also witnessing intensification and expansion of digitalized public discourses that are increasingly enabling misinformation on migration and migrants through the devaluation of accurate data and evidence and proliferation of 'fake news' and inaccurate information. Harnessing 'new' data while utilizing 'traditional' migration data and offering new analytical perspectives underpinned by deep domain knowledge through collaborative science-policy partnerships extends knowledge, fosters inquiry, and promotes accurate understandings of migration. There is the critical role of global reference reports—such as the *World Migration Report*—that collate, present, and analyze data for consumption by general, policy, technical, and educational audiences. Maximizing utility of such reports requires investments in interactive data visualization that support sustainable efforts in countering misinformation on migration and migrants through engaging and appealing design that do not compromise accuracy, but act to promote it in an accessible way.

Keywords: migration data, migrants, international migration, misinformation, disinformation, science-policy collaboration

Media Summary

The number of international migrants has risen dramatically from 153 million in 1990 to 281 million in 2020, making migration a global issue relevant to virtually every aspect of today's societies. Capitalizing on migration and its positive outcomes can, however, be challenging without reliable data and knowledge on migration and its interconnections to other global issues. Further, the use of digital platforms to actively undermine and obscure the benefits of migration has grown in recent years, leading to increased misinformation on migrants globally.

Promoting an evidence-based understanding of migration through combining data, design, and deep knowledge is the basis of the leading reference report on migration—the *World Migration Report*—produced by the United Nations (UN) biennially since 2000. Delving into long-standing migration analytical debates by bringing together longitudinal global data sets, while also drawing upon innovations in data design—static and interactive—build upon the technical fidelity and substantive domain knowledge of migration presented in the report. Through partnering with leaders in their respective fields, the innovative digitalization of migration data analyses on new platforms is enabling the report's contents to reach new audiences, thereby expanding impact by ensuring the report does not remain sitting on a (virtual) shelf. Signs thus far are good, with the latest

edition of the *World Migration Report* receiving a gold award in the 2021 International Annual Report Awards for its new online interactive platform.

1. Introduction

The starting point when it comes to discussions and critical examinations of international migration has always been numbers. Right from the first “Laws of Migration”—outlined by E. G. Ravenstein in 1885 during England’s industrialization involving large-scale internal migration—analysts, academics, and policymakers have sought data on quantitative dimensions of migration to measure, anticipate, and shape change.

Quantitative aspects of volume, proportion, demographic characteristics, and other dimensions are central to understanding longer-term trends and discernable changes in migration patterns. Migration statistics are a key part of the lifeblood of policymaking in a field that can be fast-moving, highly uncertain, and involve multiple migration-related ‘events’ over each lifetime, most especially when compared to the other major aspects of demography accounting for population change (notably births and deaths). Demographers working on global population projections have consistently found ‘international migration’ to be the variable with greatest volatility and hardest to project (United Nations Department of Economic and Social Affairs [UN DESA], 2003).

As part of broader globalization processes, we have seen momentous changes in information communications technology (ICT) as increasingly powerful distance-shrinking technologies elevate individualization of information and publishing en masse. Mobile phone penetration has increased quickly, with some developing countries experiencing great leaps technologically, forgoing older forms of ICT, such as landlines and personal computers to take up smartphones (International Telecommunication Union, 2019). As the volume of data generated globally emanating from the expansion of new forms of ICTs skyrockets, the pressure to quantify every aspect of people’s day-to-day lives intensifies. More data was created in a 2-year period than in the entire previous history of the human race (Marr, 2015). We have well and truly entered a data-driven age. We are also increasingly witnessing the move from evidence-based policy to data-driven policy (van Veenstra & Kotterink, 2017), however, in moving in this direction we risk becoming more and more distant from the underlying meaning of data variables, thereby increasing the risk of ineffective data-driven policy that misunderstands and misinterprets social phenomena such as migration.

In this article we critically examine traditional and ‘new’ migration data in the context of rapidly expanding digitalization of public discourses, which are increasingly enabling misinformation on migration and migrants through the devaluation of accurate data and evidence and proliferation of ‘fake news’ and inaccurate information. We describe recent migration data expansion, linked to global transformations in digital technologies, with reference to the role and responsibilities of global reference reports, such as the *World Migration Report*, that collate, present, and analyze data in collaboration with leading academics for consumption by general, policy, technical, and educational audiences. The article then articulates the use of

data sets to explore new perspectives on international migration in the report, and outlines recent developments in the design, science–policy collaboration, and finalization of digital tools that draw upon key statistical analyses in the report to communicate with nontechnical users through engaging, interactive data platforms.

2. A Short History of Migration Data

On the broad topic of ‘data,’ discussions often proceed with crossed purposes. The definition of what constitutes data is not always agreed upon, and the conceptualization of data utility is frequently contested. The definition of ‘data,’ as defined in the *Cambridge English Dictionary*, is “information, especially facts or numbers, collected to be examined and considered and used to help decision-making, or information in an electronic form that can be stored and used by a computer” (Cambridge University, 2020). Some argue that the term ‘data’ has evolved to the point that it has expanded beyond previous definitions that were based on statistics, quantitative (and qualitative) research, and computer science to the point that it has become potentially problematic at times or even dangerous (Ramanathan, 2016). A significant part of this evolution in the term ‘data’ is due to the substantial changes in the very nature of data, and the emergence of new technologies that produce them (Meng, 2019). In recent years we have heard the head-spinning statistics about how much data are produced globally every year (Desjardins, 2019; Marr, 2015). There has also been the realization that because of the massive amount of data that are being generated, data scientists are needed in order to make sense of it, with some arguing that ‘data’ is increasingly becoming a buzzword designed to attract talented scientists from various disciplines to work for business (Liffreing, 2018; Press, 2013).

Traditionally, in the field of international migration at the global level, statistics have been conceptualized in two main ways: *stocks* of international migrants (measuring the number and proportion of international migrants in a population at a given point in time); and *flows* of international migration events (measuring cross-border movement over a specific period). In compiling global estimates on international migrants (stocks), UN DESA equates these with foreign-born residents, drawing mainly upon national-level data provided via population censuses and supplemented by other relevant data, such as from the UN on refugee populations (UN DESA, 2021a). We know from these global data, for example, that there were around 281 million international migrants living outside their country of birth in mid-2020 (equivalent to 3.6% of the global population), up from 153 million in 1990 (or 2.9%) (UN DESA, 2021b), representing a gradual, modest increase in international migration over the last three decades. The vast majority of people globally remain within their country of birth. The international migrant data set has been built over decades and now encompasses 232 countries and areas, providing the ability to analyze international migration trends globally and regionally (International Organization for Migration [IOM], 2017, 2019). Age and sex disaggregation is also available.

On the other hand, international migration flow statistics are much more limited, with the UN DESA data set comprising statistics on 45 countries, most recently updated in 2015 (UN DESA, 2015). The current data set extends the 2008 edition (29 countries) and the 2005 edition (15 countries), however, the ability to conduct

trend analysis remains limited. Capturing data on migration flows is extremely challenging for several reasons. First, while international migration flows are generally accepted as covering inflows and outflows into and from countries, there has been a greater focus on recording inflows, with many countries only counting entries but not departures (Koser, 2010). Tracking migratory flow events also requires considerable resources, infrastructure, and technology, which poses particular challenges for developing countries, where the ability to collect, administer, analyze, and report migration flow data is often limited (IOM, 2017). Finally, many countries' physical geographies pose tremendous challenges for collecting data on migration flows, such as those with archipelagic and isolated borders, as found in many Southeast Asian countries (Gallagher & McAuliffe, 2016).

Other traditional forms of migration data include administrative data related to migration and visa processing (e.g., visa applications), access to migration-related services (e.g., integration assistance), and humanitarian and related assistance (e.g., victims of human trafficking assistance, displaced persons' assistance). These data are not typically available at the global level, but relate to countries or other administrative units, and they are often not reported due to privacy and other sensitivities, or if they are reported by authorities on an aggregated basis, strict requirements are often in place to ensure privacy is maintained (Yildiz & Abel, 2021). Recent examples of major data breaches in the sphere of migration, however, highlight how susceptible data depositories can be to maladministration or cyberattack (Farrel & Laughland, 2014; Karp, 2020). Other traditional forms of migration data, such as those collected as part of academic research, are often much narrower in scope than national-level censuses, population surveys, and administrative records, but can delve much deeper into areas of critical inquiry, such that ethical considerations are routinely embedded throughout methodological design, development, and fieldwork stages (Bean & Brown, 2015; Bloemraad & Menjívar, 2022).

2.1. Changing Migration Data Landscape: Impacts of Technology

New digital technology has emerged in the last two decades that has changed the way we talk about, as well as use migration data; technology has also enabled new data to be collected. As in other fields, there are concerns that the prerequisites of statistical knowledge, computing skills, and substantive expertise are giving away to a much greater focus on ICT skills without depth of knowledge in statistics or the substantive domain being analyzed (Liffreing, 2018; Ramanathan, 2016). Others critically examine how the emerging field of data science intersects with existing disciplines and the risks associated with the formation of an artificial ecosystem around data studies (Meng, 2019).

This is particularly noticeable in migration. The application of 'new' data science in the study of migration, which at its most fundamental is a highly social phenomenon that is rooted in normative regulatory systems (Castles, 2010; McAuliffe & Goossens, 2018), often fails to take into account the most basic understandings of the substantive topic. Without relevant content knowledge of migration, errors can be made at one or more of the basic steps in data science methods, leading to misspecification and misinterpretation. As noted by Kalev

Leetaru (2019), “for a field populated by statisticians, it is extraordinary that somehow we have accepted the idea of analyzing data [of which] we have no understanding.”

In the context of a rapidly evolving global data landscape, the very notion of ‘migration data’ has changed (see Table 1). What was once focused more on traditional statistical data collection, reporting and analysis has increasingly expanded to also encompass user-generated data relevant to international migration (Yildiz & Abel, 2021). Examples include user-generated data as part of closed systems, such as digital border processing, but also other user-generated data collection not initially designed for migration-related purposes, such as social media tracking.

Table 1. Migration data—the ‘old’ and the ‘new.’

Term	Characteristics	Examples
'Old' statistical data	<ul style="list-style-type: none"> - frequently curated, managed and in a technical format - presented with assumptions of understanding of quantitative vocabulary 	<ul style="list-style-type: none"> - population surveys and censuses - datasets of United Nations agencies - administrative statistics - academic surveys.
'New' user or event-generated data	<ul style="list-style-type: none"> - limited curation in the collection phase - interaction of someone or something with a system, in an online or offline environment 	<ul style="list-style-type: none"> - border processing data - biometrics capture - online visa application data - online travel/accommodation bookings - smartphone location tracking - social media interactions.

Adapted from McAuliffe & Sawyer, 2021.

As shown in Table 1, the types of migration data that fall under the category of ‘statistics’ are frequently curated and stored in databases with varied degrees of accessibility. By contrast, advancements in communication technology have allowed for exponential growth in user-/event-generated data, resulting in massive data sets with complex structures that update rapidly (Davenport, 2014). Notwithstanding the organizations and individuals who have worked to keep data freely available and transparent, much of what is constituted in user- and event-generated data is of limited accessibility for commercial, proprietary, or security reasons (Davies et al., 2019). However, insights from new data sources, such as call data records and social media postings, are enabling researchers to forge new lines of inquiry involving migration, displacement, and humanitarian response scenarios (Danchev & Porter, 2021; Salah, 2021), as well as to better understand migration in public debates (Culloty & Suiter, 2021; Donato et al., 2022). The shorter timeframes involved in many ‘new,’ digitalized data sources support new areas of research that can more easily advance programmatic rapid response in fast-moving environments, such as those typical of humanitarian crises and media discourse.

2.2. Imprecision and Misinformation

While it goes without saying that migration concepts are central to the formulation of accurate estimates and models produced by data scientists in the area of migration, the complexity of definitions is often glossed over, or not adequately investigated and articulated. Terms that are based on legal definitions tend to be used interchangeably and without precision, causing confusion and eroding credibility of estimates. Misleading estimates can be produced that can be fuel for policy considerations or promote inaccuracies in heavily contested public and political discourses. A recent example of this is Pew Research’s 2019 estimates of irregular migrants in Europe (Connor & Passel, 2019), in which there was a disregard of basic normative settings, resulting in highly inaccurate estimates. In this instance the researchers included asylum seekers whose applications were being processed, despite the fact that this cohort had the lawful right to remain while applications are being processed and so were not irregular migrants (Suro, 2019). The result was an inaccurate and inflated estimate.

Imprecision allows for confusion, which in turn creates the space for misinformation and disinformation to be effectively instrumentalized for disruptive purposes and, ultimately, political and financial gain (Bayer et al., 2019; Simpson & Conner, 2020). The increasing volume of information and data also risks ‘drowning’ rather than enlightening, and both aspects (volume and imprecision) have increasingly been utilized as important tools in overwhelming publics for manipulation and control (Hendricks & Vestergaard, 2019; Zuboff, 2019). This is impacting societies globally, with mega-issues such as climate change, migration and displacement, and COVID-19 facing the devaluation of (accurate) data and evidence in the face of the creation of ever-increasing ‘content’ consuming attention and diverting clear focus on issues (Hendricks & Vestergaard, 2019). Further, in the field of migration, recent studies highlight the increasing role of far-right parties in politicizing migration for political expediency using misinformation (Abou-Chadi & Krause, 2018; Culloty & Suiter, 2021).

It is in this context of increasing volume and the potential for decreasing understanding of underlying definitions of data variables—especially underpinning the ‘new’ or nontraditional data—that scientists are increasingly required to deploy methodological and domain knowledge expertise in critically examining highly salient issues, including international migration. What constitutes signal and noise may vary according to disciplinary constructs across studies (Meng, 2020), however, a potentially more profound challenge goes to the need to double down on data certainty, including through the ongoing analysis of traditional, longitudinal data sets, while simultaneously pushing the boundaries of data-driven communications, such as online interactive data visualization, in order to reach beyond scientific audiences. Science–policy collaborations that bring together deep domain knowledge and new forms of data connected to migration aspects, such as increasingly available COVID-19 mobility data, are increasingly recognized as a cornerstone to robust analysis. This has been central to recent reviews, evaluations, and related enhancements of the *World Migration Report*, which have found that safeguarding rigor, relevance, and accuracy of the data, information and analysis in the report has been achieved, but while also highlighting the need to extend utility through

innovative digital tools (IOM, 2019; Multilateral Organisation Performance Assessment Network, 2019; Office of the Inspector General, 2020).

3. A Fresh Take on Age-Old Migration Conundrums: Analyzing Data Sets for New Perspectives

In the new edition of the *World Migration Report*, collaboration between migration policy analysts, data experts, and academics has allowed for new analytical perspectives on some of the perennial issues in international migration, including how migration impacts long-term migration patterns around the world. For example, the relationship between economic development and international migration is often characterized using an inverted U-shape, where emigration first increases and then decreases as a country experiences economic development (Åkerman, 1976; Clemens, 2014; Dao et al., 2018; de Haas, 2010; Gould, 1979). Otherwise known as the migration hump model, potential migrants in countries with low levels of economic development are assumed to face unsurmountable barriers that prevent international migration, such as finances for the costs related to a move (Figure 1). In countries further up the development spectrum, barriers to emigration are reduced, leading to larger volumes of emigrants until an inflexion point, whereby higher incomes become a stabilizing influence leading to reduced outward migration.

Figure 1. The migration hump. Adapted from Clemens (2014).

The migration hump model has been widely studied, in particular in the context of impacts of developmental assistance on slowing or increasing emigration and the potential futures of (irregular) migration to high-income countries (Clemens, 2020; de Haas, 2007). Recent research has critically assessed and challenged the migration hump model, including the decline of emigration (or lack thereof) after the peak of the hump (Martin & Taylor, 1996), the timing periods used for analyzing the relationship between emigration and development (Clemens, 2020; de Haas, 2010), and the declining quality and quantity of older migration and economic data. In the remainder of this section, we reexamine the relationship between development and migration using a comprehensive set of global data on migrant population distributions during the past 25 years.

Our study on the relationship of migration with economic opportunity uses the UN DESA bilateral stock data in conjunction with the Human Development Index (HDI) of the United Nations Development Programme (UNDP). The HDI was developed by Pakistani economist Mahbub ul Haq and first used by the UNDP in 1990 as the centerpiece of its 1990 *Human Development Report* in an effort to better encompass the three major human aspects in analysis of development (education, health, and economic living standards) previously dominated by economic indicators (Stanton, 2007). Initially, the HDI covered 130 countries, increasing to 163 in 1995 and progressively reaching a total of 189 countries in the UNDP (2020). We use the HDI in our migration hump model analysis as it draws upon a broader set of indicators relevant to the migration and development process, compared to a single economic indicator such as gross domestic product or the average income of a household.

The UNDP groups countries into one of four categories based on the HDI value (low, medium, high, and very high). In our first analysis, we aggregate the number of international emigrant and immigrants to and from countries in each HDI category. Before doing so, we adjust the UN DESA stock data at each time point for the international forced displacement of migrants using data on refugees and asylum seekers from the United Nations High Commissioner for Refugees (UNHCR). The remainder allows us to study purely the trends and patterns of voluntary migrants and their relationships with respect to economic development.

Figure 2. Immigrants and emigrants by Human Development Index country category, 2020. From UN DESA (2021b) and UNDP (2020).

In Figure 2 we show the share of immigrants and emigrants of the total population in each of the HDI categories. There is a significant prevalence of emigrants from countries with high or very high HDI. In terms of volume (not shown), very high-HDI countries are the source of 76 million migrants, second only to high-HDI countries (86 million); medium-HDI countries account for 11 million and low-HDI countries account for 21 million. The higher share and volume of emigrants in countries with higher HDI than countries with low or medium HDI is counter to the migration hump model and results of previous analyses of older global migrant stock data and HDI, carried out by de Haas (2010). Also notable in Figure 2 is the large share of immigrants of the total population in 2020 in countries with very high HDI. Countries with very high HDI are home to 190 million migrants in 2020, 79.6% of all migrants globally. These numbers are a result of considerable growth in recent years.

The source of the discrepancies between previous global studies on the migration hump model and the results in Figure 2 are predominantly due to two important but distinct processes: 1) there have been significant changes in HDI country classification since 2005 and 2) migration has intensified to as well as from highly developed countries.

In 1995, 23 (or 14% of) countries were classed as having very high levels of HDI, and a further 27 (or 16% of) countries were classed as having high levels of HDI. Almost all countries that have been reclassified over time have moved into a higher category, through improvements in composite measures underlying the HDI. By 2019, 66 countries (or 34%) were classified as very high HDI and a further 53 (or 27%) were classified as high HDI.

Figure 3. Migrants to, between, and from each of the four HDI categories (low, medium, high, and very high), 1995–2020. From UN DESA (2021b) and UNDP (2020).

‘Migration to’ plots refer to migration to that HDI category from the other HDI category countries; ‘Migration from’ plots refer to migration from that HDI category to the other HDI categories. The data points at the 5-year intervals in the color bands reflect the HDI categorization at that time; the black dotted line uses 2020 HDI classifications across all data points (i.e., 1995 through 2020).

The reclassification of countries helps partly explain the different patterns of migrants found in previous studies to those in Figure 2. In order to explore the underlying migration dynamics and their relationship with

HDI, beyond reclassification issues, we looked at the level of migration between 1995 and 2020 based on a fixed set of countries in each HDI group. Figure 3 shows the change in the number of migrants based on aggregations of country-to-country data based on the 1995 HDI classifications throughout. The figure is arranged with panels for four separate HDI categories in each row and three types of migration corridors (migrants moving to, within, or from countries in each HDI classification) in each column. The shading in each panel is based on the HDI category for the origin or destination of migrants. The dotted lines represent the overall level of migrants based on the 2020 HDI classifications.

There are a number of notable trends in the levels of migration shown in Figure 3. First, there is a marked increase in the levels of ‘migration to’ countries in higher HDI categories (moving down the left-hand panels of Figure 3), with very few migrants in low-HDI countries, more migrants in medium-HDI countries, more again in high-HDI countries, and the largest number of migrants in very high-HDI countries. Migrants to very high-HDI countries have become increasingly composed of populations born in high-HDI countries (bottom left panel). Second, ‘migration from’ (in the right-hand panels) follow a ‘stepladder’ principle, whereby migrants are residing in countries with higher HDI than their country of birth. The reclassification of HDI categories have had a notable impact in the stepladder pattern, resulting in a pronounced number of migrants residing in very high-HDI countries. Third, there are sizeable differences in the number of migrants residing in a country with the same HDI category as their place of birth (middle panels), in particular between low-HDI countries and between very high-HDI countries. The levels of migrants moving within medium-HDI countries and within high-HDI countries is comparatively smaller when compared to the number of migrants to or from the country groups. Finally, the biggest increases over the period are in the number of migrants from high-HDI countries residing in very high-HDI countries and migrants from very high-HDI countries residing in very high-HDI countries.

In the context of the migration hump model, it is notable that the number of migrants from low- and medium-HDI countries increased between 1995 and 2020, but only slightly. The combination of migration aspirations and migration infrastructure (or lack thereof) did not result in high growth rates of international migration from low- and medium-HDI countries, even when accounting for recategorization of HDI ratings over time. This is consistent with past macroeconomic analyses, which show that international migration from low-income countries has historically been very limited. However, counter to the migration hump model, the number of migrants from high- and very high-HDI countries have risen consistently over the period, with emigration increasingly associated with highly developed countries.

The use of international migrant stock data has provided the solid foundation for a new perspective on international migration patterns through a human development lens. Through analyzing empirical estimates, we can see that, over time, there appears to have been a shift in the migratory practices of people born in high- and very high-HDI countries, highlighting the need to further examine the efficacy of the migration hump model. The correlation between HDI level and international migrant stock changes indicates that there are

likely to be policy-related reasons underpinning these changes. The need to reduce global inequality through the implementation of the UN Sustainable Development Goals—and target 10.7 on safe, orderly, and regular migration—further underpins the significant implications of better understanding shifts in global migration patterns (UN, 2015).

4. Taking Migration Data to the Next Level: Designing Impactful Tools

Data visualization, defined as “the visual representation and presentation of data to facilitate understanding” (Kirk, 2019, p. 15) or “drawing graphic displays to show data” (Unwin, 2020), has become an important means of communication in a range of fields, including migration. The increasing visualization of data, including migration data, reflects the effectiveness of this form of communication. Unwin (2019) argues that graphics not only show aspects that models and statistics may miss, but that they also foster questions “that stimulate research and suggest ideas.” Li (2020) points to several advantages of data visualization, such as allowing users to identify emergent patterns and instantly show large amounts of data while boosting the understanding of both small-scale and large-scale data. Others have pointed to how visualizations aid in decision-making, particularly in this day and age, when decision makers increasingly rely on data (Berinato, 2016).

With several aspects that can be conveyed visually as well as being both a politically and publicly relevant topic, international migration is especially well-suited for visualization (Allen, 2018, 2021). Migration entails several dimensions and features that can be visualized, and key data have recently been captured on the *World Migration Report*’s new interactive platform. In addition to presenting interactive components on international migrant stock and international remittances data, the report microsite extends the above analysis of global migration patterns by bringing together the three data sets—HDI, the Passport Index, and the Fragile States Index (Fund for Peace, 2021; Henley & Partners, 2021). Through interactive data visualization, users can select specific countries to explore correlations across the indices, revealing patterns or connections that may not have previously been as clear when looking at raw data or static visuals. This visualization is particularly revealing when it comes to the availability (or lack thereof) of migration options for some people as opposed to others. For example, nationals from countries with very high levels of human development can largely travel visa-free to other countries, while those from fragile countries with low levels of human development are much more restricted, with fewer migration options.

Designed to further enhance both the utility and accessibility of the *World Migration Report*, its new digital platform presents a selection of key data visualizations from the report in a way that is accessible, interactive, and visually engaging for readers and users of migration data. Extending robust outputs in this way expands access to evidence-based information about migration, providing the potential to support constructive debates to help transform the polarization triggered by misinformation and fake news that influence the perceptions of reality in the current “network society” (Castells, 2009).

Figure 4. Static screenshot of the interactive platform. From IOM (2021).

The data visualizations presented on the platform are based on the data and analysis in the *World Migration Report*, whose chapters are coauthored with, and peer-reviewed by, some of the leading migration academics as well as IOM experts (see static screenshot in Figure 4). The short narratives that accompany the visualizations support user interaction, making key facts more accessible to a wide variety of users, including policymakers, researchers, and media professionals (such as journalists, fact-checkers, and social media/community managers), as well as the general public. As Engebretsen (2020) explains, the purposes of data visualization include making complicated issues more accessible and allowing readers or users to navigate large data sets on their own.

Further, the microsite contains global and regional international migrant estimates and presents this data—over a 30-year period—for all UN world regions, including Northern America, Asia, Europe, Oceania, Africa, Latin America, and the Caribbean. By bringing these data to life, users seeking to explore how the number of international migrants has evolved over time can instantly compare these dynamics across regions. Using an interactive map, the platform also provides important migration statistics at the country level. And by visualizing migration corridors, users can get a snapshot of how migration patterns have resulted in substantial foreign-born populations in some destinations. Another key element of the interactive is to highlight international remittance flows; these data are visualized for both top recipient and source countries globally, going as far back as 1995. Visualizing how international remittances have shifted over time—in terms of both sending and receiving countries—is particularly useful given the heightened focus, in both policy and academic spheres, on the role that the money migrants send to their families and friends plays in areas such as development, including toward achieving the Sustainable Development Goals.

The platform also features several useful and specialized resources, including educational and other tools for key audiences. The educators' toolkit, for example, draws on the data, research, and analysis in the *World Migration Report* series to provide key tools for use in the classroom. With the toolkit, educators can stimulate the active engagement of their students in global and local conversations about migration, including on themes such as demographic change and the drivers of migration, displacement, and mobility. Further, to ensure that audiences across various geographies can access the interactive, this digital platform is available in multiple languages, including English, Spanish, and French.

5. Discussion

The data landscape has been changing rapidly in recent years as the volume of user-generated data has increased to reach previously unimaginable levels. Responding to, and making sense of, the virtual mountain of data in many spheres, including international migration, have created challenges in academia, as well as in industry and policy. Increasingly, we are also seeing transformations of digital systems that facilitate the mass dissemination of misinformation, creating 'thoughtless' bubbles that are expanding in reach and influence. Now, more than ever before, data collaborations between scientists and other experts are needed to ensure that meaningful and robust data analysis based on deep domain knowledge can be communicated with impact in digital landscapes.

The *World Migration Report* seeks to do this by offering new perspectives on long-term migration trends to show that, contrary to previous understandings on the migration of people from high-income countries, namely, that as country income levels increase above a threshold, international migration rates decline, the scale and proportion of outward migration from high- and very high-HDI countries has increased significantly. In fact, this bivariate analysis of migration stock across the last quarter century indicates that there has been a 'polarizing' effect, with migration activity increasingly being associated with highly developed countries. This correlation raises the key issue of visa access and related migration policies, especially in the context of migration aspirations held by potential migrants around the world who may wish to realize opportunities through international migration but are unable to do so.

The report incorporates this new analysis in open source 'traditional' settings through co-authorship collaboration with some of the world's leading migration data experts and academics, while also building upon partnerships with data visualization experts to realize cocurated, interactive platforms for use by wide audiences, including policymakers, media, students, and researchers. Maintaining technical fidelity *and* expanding online data design offers an important way to help in the battle against misinformation on international migration and migrants, providing accessible tools to foster informed and horizontal conversations on migration, and to better understand its crucial role in the success of the UN 2030 Agenda for Sustainable Development (UN, 2015). Migration has indeed enriched societies all around the world over hundreds of years, despite what those behind the social media bots spreading fake news may have people believe. While it is too early to assess the effectiveness of these new collaborations in bringing digital life to a

long-standing report, the signs are good with the *World Migration Report*'s new online platform being awarded gold in the 2021 International Annual Report Design Awards (IADA, 2021).

Acknowledgments

The authors are grateful to the organizers of the May 2021 World Migration and Displacement Symposium—*Harvard Data Science Review*, USA for IOM, and USA for UNHCR—for the opportunity to situate analysis that formed the early basis of this paper. Please note that views expressed are the authors' and do not necessarily reflect those of IOM or its member states.

Disclosure Statement

Marie McAuliffe, Guy Abel, Adrian Kitimbo, and Jose Ignacio Martin Galan have no financial or non-financial disclosures to share for this article.

References

- Abou-Chadi, T., & Krause, W. (2018). The causal effect of radical right success on mainstream parties' policy positions: A regression discontinuity approach. *British Journal of Political Science*, 50(3), 829–847. <https://doi.org/10.1017/S0007123418000029>
- Åkerman, S. (1976). *Theories and methods of migration research*. In H. Runblom, & H. Norman (Eds.), *From Sweden to America* (pp. 19–75). University of Minnesota Press.
- Allen, L. W. (2018). *Representing freedom and force: How data visualisations convey the complex realities of migration*. Dahrendorf Programme for the Study of Freedom, University of Oxford. <https://www.sant.ox.ac.uk/sites/default/files/related-documents/2017-11-06 - dahrendorf essay - w allen.pdf>
- Allen, L. W. (2021). The conventions and politics of migration data visualizations. *New Media & Society*. Advance online publication. <https://doi.org/10.1177/14614448211019300>
- Bayer, J., Bitiukova, N., Bard, P., Szakács, J., Alemanno, A., & Uszkiewicz, E. (2019). *Disinformation and propaganda – Impact on the functioning of the rule of law in the EU and its member states*. European Parliament, LIBE Committee, Policy Department for Citizens' Rights and Constitutional Affairs, HEC Paris Research Paper No. LAW-2019-1341. <http://doi.org/10.2139/ssrn.3409279>
- Bean, F., & Brown, S. (2015). Demographic analyses of immigration. In C. Brettell, & J. Hollifield (Eds.), *Migration theory: Talking across disciplines* (pp. 67–89). Routledge.
- Berinato, S. (2016, June). Visualizations that really work. *Harvard Business Review*. <https://hbr.org/2016/06/visualizations-that-really-work>

Bloemraad, I., & Menjívar, C. (2022). Precarious times, professional tensions: The ethics of migration research and the drive for scientific accountability. *International Migration Review*, 56(1), 4–32.

<https://doi.org/10.1177/01979183211014455>

Cambridge University (2020) *Cambridge English dictionary*. Cambridge University Press.

<https://dictionary.cambridge.org/dictionary/english/data>

Castells, M. (2009). *Communication power*. Oxford University Press.

Castles, S. (2010) Understanding global migration: A social transformation perspective. *Journal of Ethnic and Migration Studies*, 36(19), 1565–1586. <http://doi.org/10.1080/1369183X.2010.489381>

Clemens, M. A. (2014). Does development reduce migration? In R. Lucas (Ed.), *International Handbook on Migration and Economic Development* (pp. 152–185). Edward Elgar.

<https://doi.org/10.4337/9781782548072.00010>

Clemens, M. A. (2020). *The emigration life cycle: How development shapes emigration from poor countries*. IZA Discussion Paper No. 13614. IZA Institute of Labor Economics.

<https://www.iza.org/publications/dp/13614/the-emigration-life-cycle-how-development-shapes-emigration-from-poor-countries>

Connor, P., & Passel, J. S. (2019, November 13) Europe’s unauthorized immigrant population peaks in 2016, then levels off. *Pew Research Center*. <https://www.pewresearch.org/global/2019/11/13/europes-unauthorized-immigrant-population-peaks-in-2016-then-levels-off/>

Culloty, E., & Suiter, J. (2021). *Disinformation and manipulation in digital media: Information pathologies*. Routledge. <https://doi.org/10.4324/9781003054252>

Danchev, V., & Porter, M. (2021). Migration networks: Applications of network analysis to macroscale migration patterns. In M. McAuliffe (Ed.), *Research handbook on international migration and digital technology* (pp. 70–89). Edward Elgar.

Dao, T. H., Docquier, F., Parsons, C., & Peri, G. (2018). Migration and development: Dissecting the anatomy of the mobility transition. *Journal of Development Economics*, 132, 88–101.

<https://doi.org/10.1016/j.jdeveco.2017.12.003>

Davenport, T. (2014, November 5) Stop using the term “big data.” *Deloitte Insights*.

<https://www2.deloitte.com/us/en/insights/topics/analytics/big-data-buzzword.html>

Davies, T., Walker, S.B., Rubinstein, M., & Perini, F. (Eds.), (2019) *State of open data: Histories and horizons*. African Minds. www.africanminds.co.za/state-of-open-data/

de Haas, H. (2010). Migration and development: A theoretical perspective. *International Migration Review*, 44(1), 227–264. <https://doi.org/10.1111/j.1747-7379.2009.00804.x>

de Haas, H. (2007). Turning the tide? Why development will not stop migration. *Development and Change*, 38(5), 819–841. <https://doi.org/10.1111/j.1467-7660.2007.00435.x>

Desjardins, J. (2019, April 23). How much data is generated each day? *eNewsWithoutBorders*. www.enewswithoutborders.com/2019/04/23/how-much-data-is-generated-each-day/

Donato, K., Singh, L., Arab, A., Jacobs, E., & Post, D. (2022). Misinformation about COVID-19 and Venezuelan migration: Trends in Twitter conversation during a pandemic, *Harvard Data Science Review*, 4(1). <https://doi.org/10.1162/99608f92.a4d9a7c7>

Engelbrechtsen, M. (2020). From decoding a graph to processing a multimodal message: Interacting with data visualisation in the news media. *Nordicom Review*, 41(1), 33–50. <https://doi.org/10.2478/nor-2020-0004>

Farrell, P., & Laughland, O. (2014, June 12). Review blames immigration for data breach exposing 10,000 detainees. *The Guardian*. <https://www.theguardian.com/world/2014/jun/12/review-blames-immigration-data-breach-detainees>

Fund for Peace. (2021). *Fragile states index 2020*. <https://fragilestatesindex.org/data/>

Gallagher, A., & McAuliffe, M. (2016) South-East Asia and Australia. In M. McAuliffe, & F. Laczko (Eds.), *Migrant smuggling data and research: A global review of the emerging evidence base* (pp. 211–241). International Organization for Migration.

Gould, J. (1979). European inter-continental emigration 1815-1914: Patterns and causes. *Journal of European Economic History*, 8(3), 593–679.

Hendricks, V. F., & Vestergaard, M. (2019). *Reality lost: Markets of attention, misinformation and manipulation*. Springer. <https://doi.org/10.1007/978-3-030-00813-0>

Henley & Partners. (2021). *Passport Index 2021, Q2*. Henley Passport Index (2021: Q1 Update). www.henleyglobal.com/passport-index/ranking

International Annual Report Design Awards. (2021). *2021 International Annual Report Design Award winners*. <https://www.iada-award.co.uk/winner.php>

International Organization for Migration. (2017). *World migration report 2018*. <https://publications.iom.int/books/world-migration-report-2018>

International Organization for Migration. (2019). *World migration report 2020*.

<https://publications.iom.int/books/world-migration-report-2020>

International Organization for Migration. (2021). *World migration report 2022, interactive platform*.

<https://worldmigrationreport.iom.int>

International Telecommunication Union (2019) *Digital inclusion of all*.

www.itu.int/en/mediacentre/backgrounders/Pages/digital-inclusion-of-all.aspx

Karp, P. (2020, May 4). Government investigates data breach revealing details of 774,000 migrants. *The Guardian*. <https://www.theguardian.com/australia-news/2020/may/05/government-investigates-data-breach-revealing-details-of-774000-migrants>

Kirk, A. (2019). *Data visualisation: A handbook for data driven design* (2nd ed.). SAGE.

Koser, K. (2010). Dimensions and dynamics of irregular migration. *Population, Space and Place*, 16(3), 181–193. <https://doi.org/10.1002/psp.587>

Leetaru, K. (2019, March 7) How data scientists turned against statistics. *Forbes*.

www.forbes.com/sites/kalevleetaru/2019/03/07/how-data-scientists-turned-against-statistics/#471afdd257c8

Li, Q. (2020). Overview of data visualization. In *Embodying data* (pp. 17–47). Springer. Advance online publication. https://doi.org/10.1007/978-981-15-5069-0_2

Liffreing, I. (2018, November 19) Confessions of a data scientist: “Marketers don’t know what they’re asking for.” *Digiday*. www.digiday.com/marketing/confessions-data-scientist-marketers-dont-know-theyre-asking/

Marr, B. (2015, September 30). Big data: 20 mind-boggling facts everyone must read. *Forbes*.

www.forbes.com/sites/bernardmarr/2015/09/30/big-data-20-mind-boggling-facts-everyone-must-read/#6719e28917b1

Martin, P. L., & Taylor, J. E. (1996). *The anatomy of a migration hump*. In J. E. Taylor (Ed.), *Development strategy, employment, and migration: Insights from models* (pp. 43–62). OECD Development Centre.

McAuliffe, M., & Goossens, A. M. (2018) Regulating international migration in an era of increasing interconnectedness. In A. Triandafyllidou (Ed.) *Handbook of Migration and Globalisation* (pp. 86–104).

Edward Elgar. <https://doi.org/10.4337/9781785367519.00012>

McAuliffe, M., & Sawyer, A. (2021). The roles and limitations of data science in understanding international migration flows and human mobility. In M. McAuliffe (Ed.), *Research handbook on international migration and digital technology* (pp. 42–57). Edward Elgar. <https://doi.org/10.4337/9781839100611.00012>

Meng, X.-L. (2019). Data science: An artificial ecosystem. *Harvard Data Science Review*, 1(1).
<https://doi.org/10.1162/99608f92.ba20f892>

Meng, X.-L. (2020). Information and uncertainty: Two sides of the same coin. *Harvard Data Science Review*, 2(2). <https://doi.org/10.1162/99608f92.c108a25b>

Multilateral Organisation Performance Assessment Network. (2019). *2017–18 Assessments: International Organization for Migration*. <https://www.mopanonline.org/assessments/iom2017-18/IOM%20Report.pdf>

Office of the Inspector General. (2020). *Evaluation of the World Migration Report*. Independent Office of the Inspector General, International Organization for Migration. <https://evaluation.iom.int/report/evaluation-world-migration-report>

Press, G. (2013, August 19) Data science: What's the half-life of a buzzword. *Forbes*.
www.forbes.com/sites/gilpress/2013/08/19/data-science-whats-the-half-life-of-a-buzzword/

Ramanathan, A. (2016, November 18) The data science delusion. *Medium*. www.medium.com/@anandr42/the-data-science-delusion-7759f4eaac8e

Salah, A. A. (2021). Mobile data challenges for human mobility analysis and humanitarian response. In M. McAuliffe (Ed.), *Research handbook on international migration and digital technology* (pp. 106–122). Edward Elgar. <https://doi.org/10.4337/9781839100611.00017>

Simpson, E., & Conner, A. (2020, August 18). *Fighting coronavirus misinformation and disinformation preventive product recommendations for social media platforms*. Center for American Progress. <https://www.americanprogress.org/article/fighting-coronavirus-misinformation-disinformation/>

Stanton, E. A. (2007). *Human development index: A history*. Political Economy Research Institute Working Paper Series (Vol. 127). https://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1101&context=peri_workingpapers

Suro, R. (2019, November 14) The trouble with Pew's estimates of the "unauthorized" migrant population in Europe. *Medium*. www.medium.com/@suro26975/the-trouble-with-pews-estimates-of-the-unauthorized-migrant-population-in-europe-8df6be3972e9

United Nations (2015). *Transforming our world: The 2030 agenda for sustainable development*. <https://sdgs.un.org/publications/transforming-our-world-2030-agenda-sustainable-development-17981>

United Nations Department of Economic and Social Affairs. (2003). *Long-range population projections: Proceedings of the United Nations Technical Working Group on Long-Range Population Projections*. UN

DESA/Population Division.

www.un.org/en/development/desa/population/publications/pdf/trends/populationprojections.pdf

United Nations Department of Economic and Social Affairs. (2015). *International migration flows to and from selected countries: The 2015 revision*. UN DESA/Population Division.

www.un.org/en/development/desa/population/migration/data/empirical2/migrationflows.shtml

United Nations Department of Economic and Social Affairs. (2021a). *International migrant stock 2020 documentation*. UN DESA/Population Division.

https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/undesa_pd_2020_international_migrant_stock_documentation.pdf

United Nations Department of Economic and Social Affairs. (2021b). *International migrant stock 2020* (United Nations database, POP/DB/MIG/Stock/Rev.2020). UN DESA/Population Division.

<https://doi.org/10.18356/b4899381-en>

United Nations Development Programme. (2020). *Human development report 2020: The next frontier human development and the Anthropocene*. RR Donnelley. <http://hdr.undp.org/en/2020-report>

Unwin, A. (2020). Why is data visualization important? What is important in data visualization? *Harvard Data Science Review*, 2(1). <https://doi.org/10.1162/99608f92.8ae4d525>

van Veenstra, A. F., & Kotterink, B. (2017) Data-driven policy making: the policy lab approach. In P. Parycek, Y. Charalabidis, A. V. Chugunov, P. Panagiotopoulos, T. A. Pardo, Ø. Sæbø, & E. Tambouris (Eds.), *Lecture notes in computer science: Vol. 10429. Electronic Participation* (pp. 100–111). Springer.

https://doi.org/10.1007/978-3-319-64322-9_9

Yildiz, D., & Abel, G. (2021). Migration stocks and flows: Data concepts, availability and comparability. In M. McAuliffe (Ed.), *Research handbook on international migration and digital technology* (pp. 29–41). Edward Elgar. <https://doi.org/10.4337/9781839100611.00011>

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

©2022 Marie McAuliffe, Guy Abel, Adrian Kitimbo, and Jose Ignacio Martin Galan. This article is licensed under a Creative Commons Attribution (CC BY 4.0) [International license](https://creativecommons.org/licenses/by/4.0/), except where otherwise indicated with respect to particular material included in the article.