frontiers | Frontiers in Human Dynamics

TYPE Review
PUBLISHED 22 August 2025
DOI 10.3389/fhumd.2025.1540827

OPEN ACCESS

EDITED BY
Lisa Singh,
Georgetown University, United States

REVIEWED BY
Goran Miladinov,
Center for Research and Policy Making, North Macedonia
Avijit Mistri,
Manipur University, India

*CORRESPONDENCE
Ahmad Wali Ahmad Yar
✉ ahmad.wali@vub.be

RECEIVED 06 December 2024
ACCEPTED 28 July 2025
PUBLISHED 22 August 2025

CITATION
Yar AWA and Bircan T (2025) New frontiers in migration statistics: a narrative review on big data's role.
Front. Hum. Dyn. 7:1540827.
doi: 10.3389/fhumd.2025.1540827

New frontiers in migration statistics: a narrative review on big data's role

Ahmad Wali Ahmad Yar* and Tuba Bircan

AIMS-Lab, Brussels Institute for Social and Population Studies (BRISPO), Department of Sociology, Vrije Universiteit Brussel, Brussels, Belgium

Understanding and managing international migration relies heavily on the availability of timely, accurate, and detailed data. Traditional migration statistics, though foundational, often suffer from delays, limited granularity, and inconsistencies across countries. In response, researchers and institutions have increasingly turned to big data sources, including social media activity, mobile phone records, satellite imagery, and web-scraped content, to address these gaps and offer new insights into migration dynamics. This narrative review critically examines the role of big data in migration research. Drawing on peer-reviewed literature and grey sources, the study maps how big data has been used to track real-time migration flows, predict emerging trends, and analyse patterns of integration, while also identifying the ethical, methodological, and technical challenges involved. Findings reveal that big data offers significant potential to complement traditional statistics, especially in crisis contexts or for underreported migration flows. However, its use remains largely experimental, with key concerns around data access, representativeness, privacy, and the integration with official statistics. The review concludes that big data can support more responsive, evidence-based migration governance if its limitations are acknowledged and addressed. To that end, it recommends greater interdisciplinary collaboration, the adoption of ethical-by-design frameworks, and hybrid methodological approaches that combine big data with traditional and qualitative methods. Addressing digital inequality and fostering inclusive data practices will be critical to ensuring that big data serves as a tool for empowerment rather than exclusion.

KEYWORDS

big data, migration, data innovation, ethical challenges, migration statistics

1 Introduction

International migration has emerged as a defining issue of the 21st century, shaping global discourse through its social, economic, and political ramifications (Adamson and Tsourapas, 2019). Migration flows, particularly from economically disadvantaged or conflict-affected regions to high-income countries, have intensified and become highly politicised (Van der Brug et al., 2015). The movement of refugees and asylum seekers from war-torn areas, in particular, has sparked contentious debates around national security, cultural integration, and resource allocation, turning migration into a focal point of policy tension in many countries (Jumana, 2024). Simultaneously, migration exerts transformative effects on labour markets, demographic structures, and long-term economic growth, highlighting its complexity as both a challenge and an opportunity for governance (Martin, 2022).

Within these dynamics, the role of highly skilled migrants has gained strategic importance. Many governments view this group as vital to innovation, productivity, and global competitiveness. Consequently, immigration policies in various countries have been reformed

to facilitate the attraction and retention of high-skilled workers, framing human capital as essential to national economic development strategies (Přívara et al., 2020). Effective migration management, however, depends critically on the availability of accurate, timely, and comprehensive data (Kraler and Reichel, 2022). Traditional data sources form the backbone of migration statistics but suffer from several shortcomings, such as publication delays, definitional inconsistencies, and limited granularity across space and time (Ahmad-Yar and Bircan, 2021). In many contexts, restricted access to migration data further hampers robust analysis and policy development (IOM, 2021).

In response to these limitations, scholars and institutions have increasingly turned to alternative data sources to complement and enhance migration statistics (Ashton et al., 2016). A prominent development in this regard is the integration of big data, defined as large-scale, high-frequency, and digitally generated data streams derived from sources such as mobile phone usage, social media, satellite imagery, and internet search behaviour (Bircan et al., 2023). These data offer the potential to overcome many of the temporal and spatial constraints of traditional sources, providing more immediate and granular insights into migration flows (Aydogdu et al., 2023). As a result, national statistical institutions (NSIs), researchers, and international organisations have begun exploring the feasibility of incorporating big data into migration analysis and policy (Ahmad Yar and Bircan, 2025).

While prior studies have investigated reliability, validity, scope, access, and ethics (Tjaden, 2021), fundamental questions remain. To what extent can big data meaningfully complement or improve upon traditional migration data? This narrative review addresses these questions by synthesising the current academic and grey literature on big data applications in migration research. It maps where and how big data has been used, evaluates the strengths and limitations of these approaches, and explores unresolved methodological and governance issues. The paper aims to advance ongoing discussions on the potential and pitfalls of data innovation in migration studies and offers recommendations for improving data quality, ethical safeguards, and interdisciplinary collaboration. The remainder of the paper outlines the review's objectives and methods, presents a thematic synthesis of the findings, and concludes with implications for research, policy, and practice.

2 Methodology

The methodology employed in this study is centred on a comprehensive narrative review, a qualitative approach that synthesises and analyses existing literature to provide a thorough understanding of the current state of knowledge on the use of big data in migration studies (Ferrari, 2015). This type of review is particularly well suited to complex and interdisciplinary topics, as it allows for the integration of diverse sources, the identification of research gaps, and the formulation of future research directions (Green et al., 2006). Such flexibility is essential in the context of big data and migration research, where the literature encompasses theoretical debates, exploratory case studies, and emerging applications.

Unlike systematic reviews, narrative reviews are not bound by rigid protocols. This allows for a more nuanced engagement with literature across migration studies, data science, and public policy, enabling the identification of overarching patterns, unresolved issues, and emerging opportunities (Collins, 2005). Moreover, a narrative review facilitates critical engagement with both academic and grey literature, allowing for a balanced assessment of methodological strengths, limitations, and ethical challenges.

The choice of a narrative review is further justified by its alignment with established practices for examining complex social phenomena, where conceptual clarity and the synthesis of heterogeneous evidence are prioritised. This approach directly supports the study's aim to produce actionable insights and recommendations for researchers and policymakers navigating the evolving landscape of big data in migration research (Greenhalgh et al., 2018).

2.1 Scope and objectives of the review

The primary objective of this narrative review is to synthesise and critically evaluate the existing academic and grey literature on the use of big data in migration studies. Specifically, the review seeks to address several research questions, including:

- How has big data been used to complement or replace traditional migration data sources?
- In which specific areas of migration research has big data been most effectively applied?
- What are the ethical, methodological, and technical challenges associated with the use of big data in migration studies?
- What are the potential future directions for research in this field?

To maintain focus and coherence, the review deliberately avoids studies that centre on human mobility, commuting, and other forms of temporary movement unrelated to international migration. Likewise, while topics such as climate-induced migration, irregular migration, and border crossings are important, they fall outside the scope of this study in order to keep the analysis targeted and manageable.

The review covers a broad range of sources, including peer-reviewed journal articles, conference papers, and reports from national statistical institutes (NSIs), policy institutions, and international organisations. Including both academic and grey literature ensures that the review captures a comprehensive and policy-relevant spectrum of research, including influential studies that may not have undergone formal peer review.

To identify relevant academic and grey literature, we searched the following databases: Scopus, Web of Science, Google Scholar, and SSRN. We used a combination of keywords related to big data and migration, including: "big data," "migration," "digital data," "mobility," "ethics," "statistics," and "governance." Boolean operators were applied to refine the searches. For example, queries included combinations such as "big data" AND migration, "big data" AND (ethics OR privacy) AND migration, and "digital data" AND ("forced migration" OR refugees). Each query was adapted to the specific syntax and filtering options of the database. The search was limited to publications between 2012 and 2024, with an emphasis on peer-reviewed articles and policy reports. The table below summarises the main database-keyword combinations used.

2.2 Literature search strategy

The literature search was conducted systematically across several leading academic databases, including Web of Science, Scopus, and Google Scholar, and the official websites of relevant organisations. These sources were selected for their comprehensive coverage of multidisciplinary research, ensuring that relevant studies from various fields such as migration studies, data science, and public policy were included.

The search strategy utilised a combination of keywords and Boolean operators to identify relevant literature. Key search terms included "big data," "migration," "international migration," "migration data," "data integration," "ethical challenges," and "policy implications." To capture the most recent developments in the field, the search was limited to literature published within the last decade, reflecting the rapid advancements in digital technologies and their application to migration studies. Seminal works predating this period, which are foundational to the field, were also included.

The initial search yielded a large number of potential sources, which were then screened for relevance based on their titles and abstracts. This preliminary stage allowed for the removal of studies that did not directly address the integration of big data in migration research, ensuring that the final selection of literature was focused and relevant.

2.3 Inclusion and exclusion criteria

To refine the selection of literature for in-depth analysis, specific inclusion and exclusion criteria were applied. Studies were included if they focused on the use of big data in the context of migration studies, provided empirical evidence or theoretical insights into the integration of big data with traditional migration data sources, addressed ethical, methodological, or technical challenges related to the use of big data, and were published in peer-reviewed journals, conference proceedings, or as reports by reputable institutions.

Conversely, studies were excluded if they did not directly address the intersection of big data and migration studies, or if they focused primarily on other aspects of migration, such as legal frameworks or sociocultural integration, without reference to data usage. Also excluded were opinion pieces lacking substantial empirical or theoretical contributions, non-English sources, and research centred on commuting or intracity mobility. The emphasis was placed not on the quantity of studies but on their quality and relevance. Rather than aiming for exhaustive coverage, the review focused on works that aligned directly with the research questions and scope. This careful selection process ensured that the final body of literature was both relevant and high-quality, forming a robust foundation for the narrative synthesis.

2.4 Data extraction and synthesis

Once the relevant studies were identified, data extraction was conducted to systematically collect key information from each source. Extracted data included details on study objectives, methodology, data sources, findings, conclusions, and any discussion of ethical, methodological, or technical challenges. This information was organised into a structured framework to facilitate comparison and synthesis across studies.

The synthesis process involved a thematic analysis approach, where the data were grouped into key themes and sub-themes that emerged from the literature. These themes then structured the review, ensuring coherence and alignment with the research questions. As shown in Figure 1, the identification phase of the PRISMA 2020 (Haddaway et al., 2022) flow chart began with the collection of relevant studies from multiple sources, specifically three databases and 230 registers. In total, 230 records were identified. After removing 150 duplicates, five records flagged as ineligible by automation tools, and 20 excluded for other reasons, 75 records remained for screening.

In the screening phase, 40 records were excluded, resulting in 35 studies sought for retrieval. All 35 were successfully retrieved and deemed suitable for inclusion in the narrative review. This systematic selection ensured that the final body of literature was focused, comprehensive, and relevant to the study's objectives.

2.5 Critical evaluation of the literature

In addition to synthesising the findings from the literature, the narrative review also included a critical evaluation of the strengths and limitations of the existing research. This evaluation considered the robustness of methodologies, the generalisability of findings, and how effectively studies addressed ethical and practical challenges related to big data in migration research. Particular attention was given to the ethical implications of using big data, such as issues of privacy, consent, and the potential for bias in data collection and analysis.

The review also examined methodological issues, including the integration of big data with traditional sources, questions of representativeness, and technical difficulties related to processing large datasets. Through this critical engagement, the review identifies key gaps in the literature and outlines opportunities for strengthening future research, both empirically and ethically.

2.6 Comparison with existing reviews and related work

Over the past decade, the literature on the use of big data in migration-related research has expanded considerably. Several reviews and empirical contributions have addressed the topic, although their aims, methodologies, and thematic focus often differ from those of the present study. This review aims to clarify these distinctions in order to delineate its specific contribution to the field.

A key point of reference is the review by Sîrbu et al. (2020), which provides a comprehensive mapping of big data sources applied to various stages of migration, including flows, integration, and return. Their work offers a detailed typology and classification of data types, including mobile phone data, social media, and satellite imagery, and highlights their potential in complementing traditional data sources. However, the review primarily adopts a technical and data-centric approach, with limited engagement in the ethical and governance dimensions that are central to the present analysis.

The *Practitioners' Guide on Harnessing Data Innovation for Migration Policy* published by the International Organization for Migration and Joint Research Centre (2023) similarly constitutes a

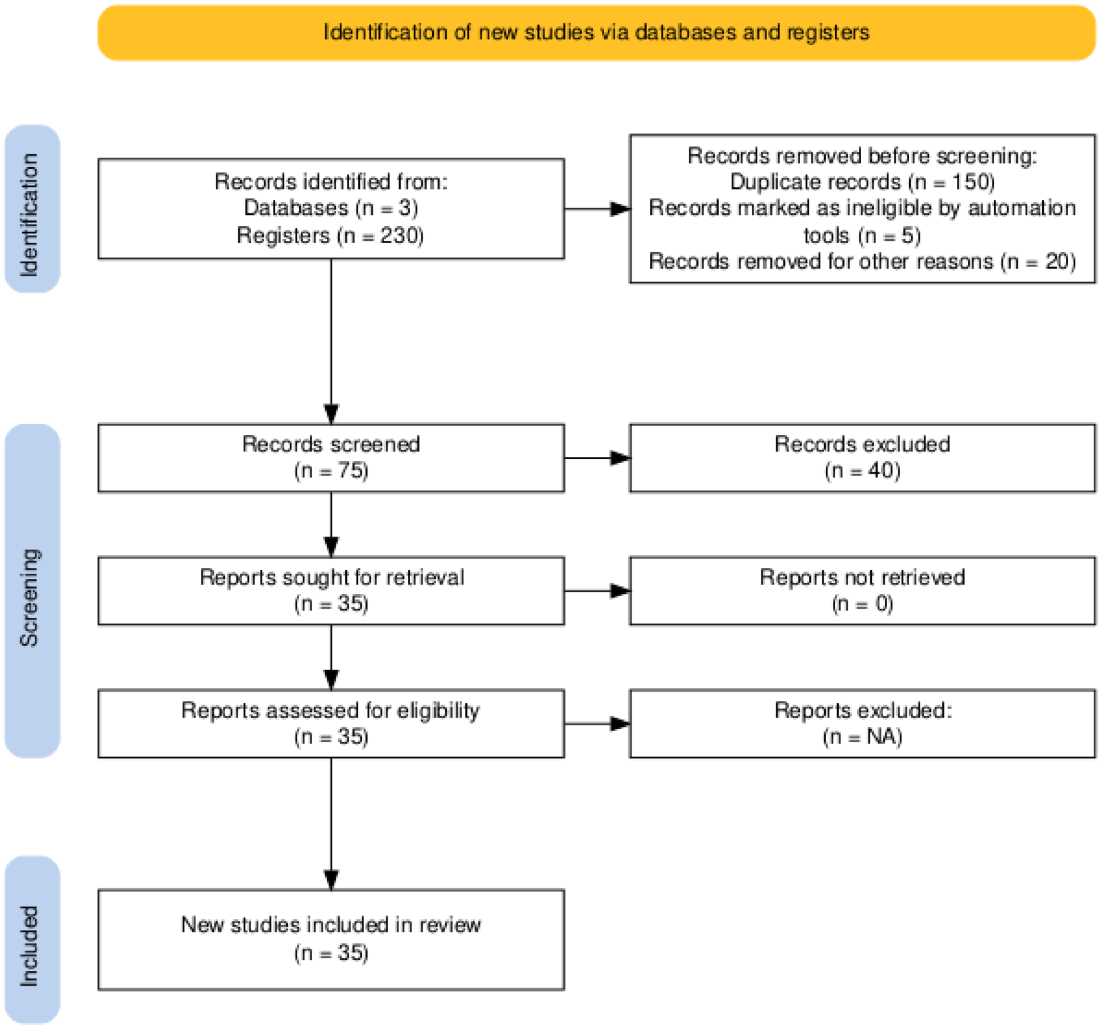

FIGURE 1
PRISMA flow chart for study selection.

foundational contribution. The guide includes chapters that examine novel data applications in migration monitoring, public sentiment analysis, and integration studies. While the volume incorporates both conceptual reflections and applied case studies, it is primarily oriented towards operational use and policy implementation. The present review differs in scope by focusing on academic and grey literature through a critical lens, particularly concerning the integration of big data into migration statistics, methodological challenges, and ethical concerns.

Several studies have advanced the application of big data in the prediction of forced displacement and refugee flows. Martin and Singh (2018), for instance, explore the use of big data and ICTs to forecast large-scale movements of displaced populations. Singh et al. (2019) combine Twitter data with conventional conflict indicators to model forced migration in Iraq, while Wycoff et al. (2025) employ social media activity and conflict data to predict refugee flows from Ukraine. Similarly, Juric (2022) uses Google Trends and Facebook data to monitor migration intentions during the 2022 Ukrainian crisis. Henningsen (2023) offers a comprehensive review of such predictive approaches, focusing on their accuracy and limitations in contexts of forced displacement. These studies demonstrate significant methodological innovation, yet they primarily concern crisis-induced migration and humanitarian forecasting topics that are explicitly excluded from the scope of the present review. Although excluded from the scope of this review, these studies demonstrate the methodological breadth of big data applications and suggest fruitful avenues for future integrative reviews that span both emergency and routine migration contexts.

In addition, recent contributions have examined the role of algorithmic governance and artificial intelligence in migration management systems. Iazzolino (2025) analyses the deployment of automated decision-making tools in asylum procedures and digital border infrastructures, raising critical concerns about surveillance and algorithmic opacity. Rönsch (2024) provides a systematic review of AI applications in refugee processing, highlighting the risks associated with bias, exclusion, and lack of accountability. Although these works offer valuable insights into the political and ethical implications of digital technologies in migration governance, they fall outside the thematic scope of this review, which does not include the digitalisation of border control or asylum determination processes.

In contrast to the above-mentioned studies, the present review is specifically concerned with the use of big data in the analysis of international migration that is not driven by humanitarian emergencies or border enforcement. Its primary focus lies in evaluating the methodological, ethical, and technical aspects of integrating big data into migration statistics and research. By synthesising academic and grey literature, this review contributes to an emerging field that critically examines how digital data sources may enhance the analysis of migration trends, integration, and mobility, while also interrogating the limitations and risks associated with such approaches.

3 Analysis and findings

The application of big data in migration studies has introduced significant shifts in how migration is tracked, analysed, and understood. While still in the experimental phase, big data presents novel opportunities that go beyond the limitations of traditional migration statistics. This section explores the transformative potential of big data, highlighting its effectiveness in key research areas, its contributions to social integration studies, and its application across different geographical regions. The section concludes by examining the experimental uses of big data in migration studies and the challenges that lie ahead.

3.1 The transformative potential of big data in migration studies

Big data has introduced a paradigm shift in migration research, offering novel opportunities to overcome the inherent limitations of traditional data sources. Conventional migration statistics, which often rely on government censuses and administrative data, have long been critiqued for their delayed availability, lack of coverage, and inability to capture short-term or undocumented migration (Ahmad-Yar and Bircan, 2021). In contrast, big data enables researchers to gather real-time insights into migration patterns by drawing from a wide range of digital sources, including social media platforms, satellite imagery, internet search queries and mobile phone records.

Platforms like Twitter and Facebook, for instance, offer geolocated data that has been utilised to monitor migrant movements and interactions in real time (Kim et al., 2020; Zagheni et al., 2017). Mobile phone data, on the other hand, has provided detailed insights into climate-induced displacement and internal mobility (Aydoğdu et al., 2022)., enabling researchers to understand migration dynamics that would otherwise remain invisible in traditional datasets.

Beyond granularity, the ability of big data to capture migration dynamics in real time has led to new possibilities for predictive modelling. Analysing digital footprints, such as social media activity or Google search trends, researchers have begun to forecast migration flows with unprecedented accuracy and timeliness (Salah et al., 2022a,b; Turper, 2023). These capabilities mark a departure from the slower, static nature of traditional migration data.

Nevertheless, while big data presents numerous advantages, its application remains experimental, often used to complement rather traditional data sources rather than replacing them (Ahmad Yar and Bircan, 2023). Concerns about data quality, representativeness, and ethical issues still limit its widespread adoption. However, the transformative potential of big data is undeniable, offering migration researchers new tools to explore mobility trends and dynamics that were previously difficult to capture.

Table 1 offers a comprehensive overview of the diverse types of big data applied in migration studies, highlighting their specific uses and limitations across various research areas.

3.2 Key areas of application: migration flows, stocks, and predictive models

Big data has proven particularly effective in tracking migration flows and stocks, especially in capturing mobility patterns that traditional data sources may overlook. Mobile phone metadata, for example, has been widely used to map migration flows and create indicators of wealth, social integration, and internal mobility (Salah, 2022). By capturing real-time data, mobile phone records allow researchers to identify temporary or seasonal migration patterns, which are often underreported in government-collected data (Ahmad Yar and Bircan, 2023).

Predictive models have also benefited from the integration of big data. Social media platforms like Twitter and LinkedIn have been employed to forecast migration flows and estimate migration stocks, providing timely insights into migration dynamics. Google Trends, for instance, has been used to predict migration intentions based on changes in search behaviour (Turper, 2023). These innovations have not only improved the timeliness of migration models but have also opened up new possibilities for more accurate forecasts.

However, the integration of big data into predictive migration models remains fraught with challenges related to data quality and representativeness. The over-reliance on digital platforms risks overlooking segments of the migrant population who are less digitally connected, thus raising concerns about bias in data collection and analysis (Salah et al., 2022a,b). Despite these limitations, big data has clearly expanded the horizons of migration research, offering new ways to monitor, model, and predict migration dynamics.

3.2.1 Social integration and refugee studies

In addition to tracking migration flows, big data has contributed significantly to understanding social integration, particularly within refugee populations. Mobile phone data and social media interactions have been used to infer socio-economic indicators, helping researchers assess how well migrants integrate into host societies (Salah et al., 2022a, 2022b). For instance, call detail records (CDRs) have been employed to monitor the social integration of refugees in Rwanda, providing insights into their mobility and settlement patterns (Blumenstock et al., 2015).

Through these digital footprints, researchers can gather more comprehensive data on refugee integration than would be possible through traditional surveys. Big data offers a level of granularity that captures day-to-day interactions and movements, helping policymakers better understand how refugees adapt to new environments and what factors contribute to successful integration (Sîrbu et al., 2024; Smirnov, 2022). However, ethical concerns about privacy and surveillance, particularly in vulnerable populations like

TABLE 1 Overview of big data applications and limitations in migration studies.

Data type	Applications	Strengths	Limitations
Social media	• Tracking real-time migrant flows through geolocated posts (Kim et al., 2020). • Forecasting migration trends using Twitter and Facebook data (Carammia et al., 2022; Yildiz et al., 2024). • Analysing migrant networks and sentiments (Ahmad Yar and Bircan, 2023).	• Timely and granular insights. • Rich in behavioural data. • Facilitates understanding of social integration.	• Data quality and representativeness issues. • Ethical concerns: privacy and consent (Ahmad Yar and Bircan, 2023).
Mobile Phone Data	• Estimating migration flows and socio-economic integration (Salah et al., 2022b). • Tracking internal displacement (Isaacman et al., 2017).	• Real-time tracking capabilities. • High granularity for mobility patterns.	• High cost and technical expertise required. • Data coverage limited to phone users.
Remote Sensing	• Monitoring cross-border migration (Bircan, 2022). • Mapping internal displacement via satellite imagery (Momeni et al., 2024).	• Visual evidence for migration studies. • Useful in regions with limited traditional data sources.	• Access often restricted by private ownership. • Requires substantial expertise to process.
Google Trends	• Predicting migration intentions via search behaviour (Juric, 2022). • Enhancing nowcasting models with search analytics (Avramescu and Wisniowski, 2021).	• Cost-effective and accessible. • Real-time indicators of migration motivations.	• Methodological validation needed. • Non-representative of offline populations (Salah et al., 2022a,b).
Web Scraped Data	• Enriching migration models with online content (Chehbi-Gamoura et al., 2020).	• Supplementary data for traditional approaches.	• Ownership issues. • High processing demands.
Multi-source Integration	• Blending traditional statistics with digital traces for improved migration modelling (Singh et al., 2019). • Using hybrid AI-statistical methods for integration analysis (Bircan and Korkmaz, 2021). • Demonstrating applied use cases for policy (International Organization for Migration and Joint Research Centre, 2023). • Exploring ethical and representational limits of digital demography (Cesare et al., 2018).	• Enhances robustness by combining data types. • Improves temporal and spatial resolution. • Supports policy-relevant migration analysis.	• High methodological complexity. • Risk of compounding biases across sources. • Ethical concerns persist.

refugees, remain a significant challenge to scaling such approaches (Ahmad Yar and Bircan, 2023).

3.2.2 Geographical variations in big data application

The application of big data in migration research varies significantly across geographical regions, largely depending on available infrastructure, technical expertise, and legal frameworks. Countries like Belgium, the Netherlands, Estonia, Germany, Hungary, Italy, and Latvia have pioneered projects that utilise big data to track mobility patterns and transportation behaviours (Ahmad Yar and Bircan, 2023, 2025). In Estonia, mobile phone data has been used to analyse daily travel patterns in both urban and rural areas, contributing to more efficient infrastructure planning and migration management.

Satellite data has also been critical in tracking migration flows in regions experiencing humanitarian crises, such as Hungary and Italy. Satellite imagery enables near real-time monitoring of border crossings, allowing governments to respond swiftly to sudden shifts in migration flows. While big data applications in these contexts have proven invaluable, they also raise issues regarding the reliance on privately owned data sources, such as satellite and mobile phone data, which can limit access and raise costs (Salah et al., 2022a,b).

Overall, the regional differences in the use of big data reflect varying levels of adoption and experimentation. Countries with more advanced digital infrastructure and established data-sharing agreements are better positioned to leverage big data for migration research, while others lag behind due to technological or legal barriers.

3.2.3 Experimental uses of big data in migration studies

In addition to its established uses, big data is increasingly being tested in experimental contexts to explore new dimensions of migration research. Countries like Australia, Canada, and the United Kingdom have initiated pilot projects that combine big data techniques with traditional statistics to enhance migration models. Australia, for example, has developed a hybrid model that merges census data with mobile phone network data to provide real-time

statistics on both internal and international migration flows (Fosso Wamba et al., 2014).

These experimental projects represent the next frontier for big data in migration studies, offering policymakers new tools for more responsive and informed decision-making. However, these applications remain at the pilot stage, with many challenges still to be addressed. The cost of processing vast volumes of data, concerns about data privacy, and the reluctance of private companies to share proprietary data continue to impede the full integration of big data into official migration statistics (Salah et al., 2022b). Despite these challenges, experimental uses of big data have shown great promise in improving the accuracy, granularity, and timeliness of migration insights climate change-induced (Salah et al., 2022b). For example, big data techniques have been successfully employed to track population movement trends in countries like Colombia, while in Europe, countries such as Italy, Germany, and Estonia, have utilised big data to develop internal migration statistics (Ahmad Yar and Bircan, 2023). Mobile phone data, in particular, has been explored in France to map internal migration flows, offering a novel approach to understanding how populations move within national borders.

However, the application of big data remains fragmented and varies significantly by region, with its use still limited in scope.

The utilisation of big data in migration research, while promising, is accompanied by a series of ethical, methodological, and technical challenges. A key issue is the proprietary nature of many big data sources, which are often controlled by private corporations, creating significant barriers to data accessibility (Reed-Berendt et al., 2022). Additionally, National Statistical Institutes (NSIs) have highlighted three primary concerns: the absence of comprehensive legal and ethical guidelines, a lack of technical expertise and resources, particularly in less economically developed countries, and questions regarding data coverage and representativeness (Ahmad Yar and Bircan, 2023).

Privacy concerns are a substantial obstacle in the deployment of big data for migration studies, as highlighted in the reviewed literature. Existing legal frameworks, primarily designed for conventional data collection methods, are often inadequate to address the nuances of big data (Beduschi, 2017). Countries such as the Netherlands and Belgium have encountered challenges related to privacy, underscoring the need for more robust legal structures to ensure the ethical use of such data (Utts, 2021). Methodological challenges also persist, particularly concerning the quality and representativeness of big data. While big data offers the potential for real-time insights, questions remain about its accuracy, biases, and the extent to which it can reliably reflect broader population trends (Ahmad Yar and Bircan, 2023).

The future of big data in migration research hinges upon overcoming these ethical, legal, and methodological hurdles. The studies suggest that a concerted effort is needed to improve data accessibility, establish clearer and more comprehensive legal frameworks, and build technical capacity within NSIs (Braaksma et al., 2020). Furthermore, future research should explore pathways for deeper integration of big data into migration statistics, moving beyond its current experimental phase. Addressing privacy concerns, enhancing data quality, and fostering international collaboration between NSIs and private companies will be essential to realising the full potential of big data in migration research (Tjaden, 2021). There is also an increasing need for research into the potential of big data to shed light on emerging migration trends, such as those driven by climate change and short-term mobility patterns.

In general, the reviewed body of academic work highlights that the advent of big data, characterised by the large-scale aggregation of individual behavioural data through online platforms, mobile services, and other digital means (Huberty, 2015), has significantly impacted various fields, including medicine (Batko and Ślęzak, 2022), public policy (Wong and Hinnant, 2023), and governance (Zhu, 2019). Migration studies, traditionally reliant on government-collected data, are increasingly looking to big data as a solution to address the gaps in reliable and timely migration statistics at both global and national levels (Ahmad-Yar and Bircan, 2021). The growing complexity of global migration patterns, illustrated by the increase in the migrant population from 161 million in 1995 to 281 million in 2020 (IOM UN Migration, 2022), underscores the urgent need for innovative data solutions to effectively inform migration policy and management.

3.3 Areas of migration research that benefit from big data

The analysis of the literature reveals that several key areas of migration research are particularly well-suited to benefit from the inclusion of big data. Firstly, big data enhances the granularity and timeliness of insights into migration flows and stocks, capturing real-time mobility patterns that traditional statistics often miss (Ahmad Yar and Bircan, 2023). For example, mobile phone metadata has been used to map migration flows and create indicators of wealth or social integration (Salah, 2022). A significant advantage of big data lies in its ability to track temporary migration, which is frequently underreported in traditional sources (Ahmad Yar and Bircan, 2023).

Moreover, big data has demonstrated its value in improving predictive migration models (Hsiao et al., 2023; Juric, 2022; Sîrbu et al., 2024; Smirnov, 2022; Tsapenko and Yurevich, 2022). By analysing user-generated content on social media, language patterns, and Google Trends search behaviours, researchers can forecast migration flows. For instance, social media platforms like Twitter and LinkedIn have been used to nowcast migration stocks, while Google Trends data has been employed to predict migration intentions (Salah et al., 2022b; Turper, 2023). This not only improves the timeliness but also the potential accuracy of migration models, although questions remain regarding the quality of such predictions due to limitations in data representativeness.

Big data also has the potential to deepen the understanding of social integration, particularly in refugee populations. Through the analysis of call detail records (CDR) and social media interactions, researchers can infer socio-economic indicators, which aid in assessing how well migrants integrate into host societies (Salah et al., 2022a,b). For example, mobile phone data has been used to monitor the social integration of refugees in Rwanda, offering insights into their mobility and settlement patterns (Blumenstock et al., 2015).

3.4 Big data across geographies

The results of the review indicate that the use of big data in migration studies has become a valuable tool for analysing patterns,

statistics, and trends across various geographical regions (Avramescu and Wisniowski, 2021; Juric, 2022; Sîrbu et al., 2024). The integration of mobile phone data, web scraping, social media analysis, and sensor data has allowed countries to gain a better understanding of human mobility and migration trends. This section explores how different nations employ big data for migration-related purposes, while also highlighting the opportunities and challenges encountered in these contexts.

In the countries such as Belgium, the Netherlands, Estonia, Germany, Hungary, Italy, and Latvia, among others, experts and academics have implemented or piloted projects using big data to track mobility patterns and transportation behaviours, employing data derived from mobile phones, web scraping, and sensor networks. These sources provide real-time insights into human movement, offering a detailed understanding of transportation trends. In Estonia, for instance, mobile phone data has been utilised to analyse daily travel patterns in urban and rural areas, contributing to more efficient infrastructure planning (Ahmad Yar and Bircan, 2023). By leveraging big data, these countries can address mobility challenges more effectively, fostering improved transportation networks and migration management.

Additionally, several studies have emphasised the significance of satellite images in understanding migration and mobility, particularly in countries like Hungary and Italy. Satellite data helps monitor migration flows across borders in near real-time, allowing governments to respond more swiftly to migration crises or sudden shifts in mobility patterns, an essential factor in humanitarian response efforts. However, reliance on privately owned data, such as mobile phone and satellite data, presents access and legal challenges, which will be further explored (Salah et al., 2022a,b).

3.5 Big data for migration statistics

The review highlights that big data plays a crucial role in generating comprehensive migration and mobility statistics. Countries across Europe and North America have explored using data from mobile phones, social media, and satellite images to enhance migration studies. Mobile phone data, in particular, allows for the tracking of individuals' movements over extended periods, providing detailed insights into migration patterns. For example, in Germany and the United Kingdom, mobile phone data has been instrumental in mapping the journeys of migrants and understanding their settlement and integration processes (Salah et al., 2022a). Additionally, big data helps distinguish between short-term, temporary migrants and those who settle permanently.

Social media platforms have also become valuable sources of data, especially in situations where official statistics are lacking. Platforms like Facebook and Twitter generate large datasets that can be mined to track migration-related conversations, conduct sentiment analysis, and map migration networks. This is particularly useful in countries with large diasporas, where social media activity may reflect migration trends more accurately than traditional surveys or censuses (Ahmad Yar and Bircan, 2025). However, concerns about privacy and the representativeness of these data sources have limited their use in official statistics, sparking debates about their validity.

Additionally, a notable trend emerging from the reviewed literature is the integration of multiple data sources, particularly the fusion of traditional statistics with big data such as social media and mobile phone records, to enhance the reliability and policy relevance of migration research. Singh et al. (2019), for instance, propose a Bayesian framework that combines structured movement variables with social media data to predict migration flows, offering methodological insights that extend beyond the context of displacement. Similarly, Bircan and Korkmaz (2021) examine hybrid AI-statistical models for measuring migrant integration, highlighting governance and ethical challenges associated with algorithmic use. Cesare et al. (2018) further underscore the representational and technical limits of digital traces in demographic research, calling for critical scrutiny when combining them with official sources. The IOM–JRC Practitioners' Guide (2023) also presents applied cases where hybrid data practices have improved migration analysis and planning. Together, these contributions illustrate the growing importance of multi-source integration as a methodological innovation in international migration research.

3.6 Experimental usage of big data

Countries such as Australia, Canada, and the United Kingdom have initiated pilot projects that use big data techniques specifically for migration studies. These initiatives seek to experiment with new methodologies, such as combining satellite data with web scraping and traditional statistics, to enhance the understanding of migration dynamics (Tjaden, 2023). For instance, Australia has launched pilot projects that merge traditional census data with big data from mobile networks, providing real-time migration statistics that account for both internal mobility and international migration simultaneously (Fosso Wamba et al., 2014). Such experimentation is crucial for generating timely, policy-relevant insights, which are essential for governments dealing with dynamic migration flows.

However, while these countries have explored the potential of big data, its integration into official statistical systems has been slow. One of the primary limitations is the cost and infrastructure needed to process and store the vast volumes of data generated. Moreover, access to privately owned data from companies, such as telecom providers, remains a major barrier. These companies may be unwilling or unable to share their data due to privacy regulations, which vary significantly across countries (Salah et al., 2022b). Consequently, while the use of big data in migration studies shows great promise, it has so far remained mostly at the experimental or pilot stage.

4 Discussion: can big data thrive without traditional data sources?

While big data has shown promise in complementing traditional migration data, it remains far from being a standalone resource. The use of big data to address gaps in conventional data sources has yet to be fully incorporated into official decision-making processes. Traditional migration data, such as administrative records and census data, provide broad coverage but often suffer from inconsistencies, delays, and varying definitions across countries (Ahmad Yar and Bircan, 2023; Salah, 2022). This is especially problematic when it comes to capturing temporary or undocumented migration.

In contrast to conventional migration statistics, big data sourced from mobile phone metadata, social media activity, and satellite imagery, offers granular, high-frequency and real-time insights into population mobility. These digital traces enable the tracking of migration flows and integration dynamics in near real-time, providing

a responsiveness that traditional datasets often lack (Salah, 2022; Ahmad Yar and Bircan, 2023, 2025; Turper, 2023). For example, mobile call detail records (CDRs) have been used to infer proxy indicators of socio-economic status, spatial mobility and social integration levels, particularly among refugee populations in Rwanda and Turkey (Blumenstock et al., 2015; Aydoğdu et al., 2022). Social media, such as Twitter geotags or Facebook activity, have proven effective in estimating migrant stocks and predicting migration intentions, enhancing both the granularity and temporal resolution of migration statistics (Kim et al., 2020; Yildiz et al., 2024). Machine learning techniques further extend the analytical potential of big data by uncovering complex, non-linear migration patterns, often integrating multiple sources such as Google Trends and mobile data for predictive modelling (Alexander et al., 2022; Avramescu and Wisniowski, 2021). However, despite these advancements, significant challenges remain. Big data often lacks demographic granularity, representativeness, and standardised methodologies, raising concerns about selection bias, especially for digitally excluded groups such as older migrants, children, or those in precarious legal statuses (Hsiao et al., 2023; Meissner and Taylor, 2024). These omissions risk reinforcing existing inequalities and misrepresenting migration dynamics.

To mitigate these limitations, scholars increasingly advocate for hybrid approaches that combine big data analytics with qualitative and participatory methods, enabling a more inclusive and context-sensitive understanding of mobility (Sandberg and Rossi, 2022). By embedding critical data perspectives and involving migrant communities, researchers can ensure that innovation enhances rather than distorts the evidence base for migration governance.

4.1 Contradictory perspectives on big data in migration research

The reviewed literature presents contradictory perspectives on the role of big data in migration research, reflecting both optimism about its potential and concerns about its limitations. Some studies, such as those by Claudio et al. (2022) and Spyratos et al. (2018), emphasize the value of big data, particularly mobile phone and social media data, as valuable complements to traditional migration statistics. These sources are seen as filling gaps in demographic insights and improving policymaking, with some findings correlating closely with traditional data sources like Eurostat. However, Ahmad Yar and Bircan (2023) take a more critical stance, arguing that big data's role remains experimental due to significant legal, methodological, and accessibility barriers. The lack of standardization and practical guidelines complicates the integration of big data into migration research, revealing a substantial gap between its potential and actual implementation.

Despite its innovative potential, big data cannot fully supplant the foundational contributions of traditional migration statistics. Traditional data sources such as population censuses, (household) surveys, and administrative registers remain indispensable due to their standardised definitions, legal frameworks and demographic completeness. These instruments capture core migration dimensions, such as the place of birth, citizenship, migration history, and legal status, enabling longitudinal and comparative analyses that are crucial for policymaking and academic research (UN DESA, 2020; Kraler and Reichel, 2022) gal status, enabling longitudinal and comparative analyses that are crucial for policymaking and academic research (UN DESA, 2020; Kraler and Reichel, 2022). In contrast, big data is often typically a by-product of digital transactions or behavioural traces (e.g., GPS data, mobile phone metadata, or social media interactions) and lacks consistent demographic tagging or metadata standards (Struijs et al., 2014; Hsiao et al., 2023).

While big data excels in timeliness, scale, and granularity, it often falls short in terms of representativeness, comparability, and conceptual clarity. For example, social media data may provide insight into mobility trends among digitally active populations but systematically exclude vulnerable groups such as undocumented migrants, older adults, or those in low-connectivity regions (Meissner and Taylor, 2024; Tjaden, 2021). Furthermore, big data rarely captures critical socio-political attributes, such as reason for migration, legal pathways, remittance behaviours, or indicators of discrimination and integration, which are vital for evidence-based policy design (Ahmad Yar and Bircan, 2023).

Consequently, big data should be conceptualised not as a replacement but as a complementary asset to traditional sources. Its strength lies in filling spatio-temporal gaps and offering real-time monitoring in contexts where conventional data are unavailable or delayed, such as in humanitarian crises or fast-changing mobility trends (Gendronneau et al., 2019; Yildiz et al., 2024). Realising this potential, however, requires investment in hybrid methodological frameworks that bridge the statistical rigour of traditional data with the dynamic responsiveness of digital sources. These hybrid approaches can enhance the robustness, inclusivity, and policy relevance of migration research if designed with ethical safeguards and cross-sectoral collaboration.

Ethical concerns also add to the complexity of big data's role. While some scholars, such as Sandberg and Rossi (2022) and Hayes (2017), highlight the risks of privacy violations, mass surveillance, and the need for stronger data governance, others focus primarily on the technical and legal aspects, often overlooking ethical dimensions (Claudio et al., 2022; Ahmad Yar and Bircan, 2025). The inconsistencies in findings between Spyratos et al. (2018) and Gendronneau et al. (2019) concerning the reliability of big data estimates further demonstrate the variability across different platforms and methods. In some cases, big data aligns well with official statistics, while in others, large discrepancies exist. These contradictions suggest that, while big data holds potential for migration research, it is not yet ready for widespread policy application, and ethical issues must be addressed before it can be fully integrated (Alexander et al., 2022; Hsiao et al., 2023; Meissner and Taylor, 2024).

The tension between big data's promise and its challenges is also evident in the field of artificial intelligence (AI) and machine learning in migration research. While AI and machine learning models offer the potential for predictive analytics, early warning systems, and enhanced migration management, they also raise concerns about their reliability and ethical implications. Predictive models, for instance, struggle with the inherent complexity and unpredictability of migration patterns, making it difficult to rely solely on these systems for policy decisions. This is compounded by challenges of selectivity, replicability, and research ethics, especially when dealing with vulnerable populations (Salah et al., 2022a).

Furthermore, the contradictions between the benefits and limitations of big data in migration studies reflect deeper issues related

to data quality and algorithmic accountability (Meissner and Taylor, 2024). Claims about AI's effectiveness is undermined by unresolved challenges such as data noise, privacy concerns, and the representativeness of data sources. While interdisciplinary collaboration is often called for, existing technical, legal, and political hurdles hinder these efforts, creating additional barriers to the successful implementation of big data in migration research.

4.1.1 Ethical and practical challenges of big data in migration research

Big data offers powerful tools for migration tracking and policy support, but it also raises serious ethical dilemmas. Migrants, particularly refugees and undocumented populations, are often vulnerable to exploitation and misuse of their data.

Furthermore, the tension between the need for data-driven insights and the protection of human rights complicates the role of big data in migration governance (Meissner and Taylor, 2024). While predictive models and early warning systems may enhance government preparedness for migration flows, the ethical risks of deploying these technologies, particularly in relation to privacy and discrimination, must be carefully managed. The drive for efficiency in migration management often clashes with the necessity of safeguarding human rights, suggesting a conflict between technological innovation and ethical governance.

In conclusion, the reviewed literature reveals both the potential and the limitations of big data in migration research. While big data can offer timely, granular insights that complement traditional migration statistics, its integration into migration studies is fraught with challenges, including ethical concerns, data quality issues, and the complexity of migration dynamics. As the field moves forward, greater collaboration between policymakers, researchers, and data providers will be essential to address these challenges and unlock the full potential of big data in migration research. The application of big data in migration research is often misunderstood, particularly in public discourse where it is frequently conflated with AI-driven border management technologies. These technologies, including drones and surveillance systems aimed at preventing irregular migration, differ significantly from the role big data plays in academic migration studies (Leese et al., 2022; Tazzioli, 2019). While border management tools are geared towards migration control, big data in research settings enhances migration statistics, offering more accurate estimations of migrant flows, trends, and the socio-economic impacts of migration policies (Sandberg and Rossi, 2022). This confusion blurs the distinction between scholarly efforts to improve migration data accuracy and the more restrictive, security-driven applications of big data. To avoid perpetuating misconceptions, it is crucial to clearly delineate between the academic use of big data for evidence-based policymaking and its employment in border control technologies.

Although big data presents significant opportunities for advancing migration research, it also introduces ethical, methodological, and technical challenges. As discussed in the previous section, ethical challenges such as privacy and informed consent remain central to the debate on big data's role in migration research. The use of mobile phone metadata, social media activity, and other personal data sources raises serious issues regarding consent and the potential for misuse, especially in the context of marginalised communities such as refugees and undocumented migrants (Salah et al., 2022a,b). These populations are particularly vulnerable to biased analyses or surveillance, which can exacerbate existing inequalities and further marginalise them (Hayes, 2017; Leese et al., 2022).

From a methodological perspective, ensuring the quality and representativeness of big data is another pressing challenge. Big data sources, such as social media or mobile phone records, do not encompass all migrant populations, especially those without access to digital platforms or technologies (Ahmad Yar and Bircan, 2023). Additionally, the distinction between migrants and non-migrants within these datasets is not always clear, as many of these data sources do not explicitly capture demographic or migration-specific information. Therefore, validating the results from big data against traditional migration statistics is critical to avoid inaccuracies or biased findings, though this process is often complex and resource-intensive (Ahmad Yar and Bircan, 2023).

Big data in migration research often consists of digital traces generated passively through individuals' interactions with digital systems, including mobile phone metadata, GPS logs, social media content, satellite images, and web-scraped records. These datasets are diverse in format ranging from highly structured call detail records (CDRs) to unstructured text or image data and differ in their potential utility depending on the research objective. The acquisition and use of such data raise significant legal and regulatory concerns. Many sources are owned by private companies, and data sharing is constrained by national and international privacy laws, inconsistent governance frameworks, and concerns about consent and misuse. In terms of analytical tools, machine learning (ML) has emerged as a powerful method for processing and interpreting these large, complex datasets. Techniques such as natural language processing (NLP), clustering, and predictive modelling are used to classify migrant statuses, detect mobility trends, and forecast flows. However, ML models require substantial volumes of high-quality training data and carry the risk of amplifying existing biases if not properly validated. As such, their application in migration research demands both technical sophistication and strong ethical oversight.

Technically, the integration of big data into migration studies requires considerable expertise and computational capacity, which are often lacking in national statistical institutions (NSIs) (Salah et al., 2022a,b). Furthermore, the absence of standardised methodologies for processing and analysing big data in migration contexts exacerbates the challenges of ensuring consistency and reliability across different studies. The sheer volume of data and the sophisticated machine learning techniques needed to analyse it require advanced computational resources that may not be readily available to many institutions.

4.2 Ethical considerations

Ethical safeguards such as 'Ethics by Design' and the FQIR principles have been proposed to mitigate concerns around consent, surveillance, and data transparency. These include concerns about the privacy and confidentiality of individuals whose data are collected and analysed, the potential for data misuse or exploitation, and the need for informed consent in the use of digital data sources. Ethical challenges such as these call for the adoption of frameworks like "Ethics by Design" (Reed-Berendt et al., 2022), FQIR data principles,

and the development of regional guidelines to address legal and cultural differences in data governance. These frameworks advocate embedding ethical safeguards, ensuring data transparency, and protecting the rights of vulnerable populations throughout the research process. Such principles are especially vital when balancing the drive for data-driven insights with the imperative to safeguard human rights.

In addition to privacy concerns, the review also considers the broader ethical implications of relying on big data for migration policymaking, such as the risk of reinforcing existing inequalities or biases in the data. The assertion that "migrants lack control over their data usage" highlights the potential for big data to exacerbate existing inequalities, as these populations are often subject to surveillance without their consent.

4.2.1 Challenges and limitations of big data

Despite the promising opportunities big data offers for improving migration statistics, its adoption remains limited due to several persistent challenges. One of the primary obstacles is data access, especially regarding privately owned datasets. National statistical institutes (NSIs) often rely on third-party entities, such as mobile network operators and satellite data providers, to access critical data sources, and these relationships are governed by complex legal and ethical frameworks (Ahmad Yar and Bircan, 2023, 2025). Differences in privacy laws across countries further complicate data sharing for migration studies, as no universal standards exist for how data should be collected, anonymised, and used responsibly.

Technical challenges also hinder the effective use of big data in migration studies. NSIs often lack the technical expertise and computational capacity required to process large datasets or implement complex machine learning techniques.

In light of these challenges, it is evident that while big data has the potential to revolutionise migration studies, its application remains in the early stages. Addressing legal, ethical, and infrastructural issues will be crucial to enabling the broader use of big data in tracking migration patterns and mobility. The development of standardised methodologies for data collection, sharing, and analysis will also be essential for ensuring that big data can be effectively integrated into official migration statistics (Salah et al., 2022a).

4.2.2 Limitations of the review

While this narrative review offers a broad and comprehensive analysis of the use of big data in migration studies, several limitations should be acknowledged. As a narrative review, the findings are based on a qualitative synthesis of existing literature rather than a systematic, quantitative approach such as a meta-analysis. Consequently, the conclusions drawn are interpretative and should be viewed as providing general insights rather than definitive evidence.

Furthermore, the review is constrained by the availability and accessibility of the literature. Despite efforts to include grey literature, some relevant studies may not have been identified or accessible, especially those published in non-English languages. Additionally, the rapidly evolving nature of big data and migration research means that the review may not fully capture the latest developments or emerging trends in the field. Continued research and updated reviews will be necessary to monitor the progress and new applications of big data in migration studies.

5 Conclusion

This narrative review has explored the emerging role of big data in migration studies, revealing both its transformative potential and its inherent limitations. Big data has introduced a paradigm shift in how migration patterns are monitored and analysed, offering timely, granular insights that complement, but do not replace, traditional data sources. Its capacity to generate near real-time information makes it particularly valuable in fast-changing contexts, such as refugee movements or climate-induced displacement. Yet, its integration into migration research remains constrained by issues of representativeness, digital exclusion, ethical uncertainty, and technical complexity.

The review emphasises that big data's promise lies in its ability to fill temporal and spatial gaps in official statistics, especially in contexts where conventional data are delayed or incomplete. However, this promise is tempered by structural limitations, including the systematic underrepresentation of digitally disconnected groups and the risks of reproducing biases through algorithmic analysis. In particular, the use of personal data from mobile devices or social platforms raises urgent concerns about consent, privacy, and potential misuse, especially for vulnerable populations such as refugees, stateless individuals, or undocumented migrants.

Regional disparities in digital infrastructure and governance further shape the applicability of big data across contexts. High-income countries with advanced data ecosystems, are leading innovation in this area, while lower-income settings often face barriers to access, interoperability, and technical capacity. These inequalities raise pressing questions about global data justice and equitable access to the benefits of data innovation.

Theoretically, big data enables a move from static, retrospective analysis to dynamic, predictive models that can capture the fluid and multidimensional nature of migration. It also opens opportunities for intersectional analysis, revealing how migration trajectories are shaped by overlapping factors such as gender, age, legal status, and socio-economic position. Practically, big data can support more agile policy responses in crisis scenarios, improving the ability of governments and humanitarian actors to act on emerging mobility trends. However, such applications must be embedded within robust ethical frameworks and technical safeguards to avoid unintended harms. To that end, ethical considerations must extend beyond general principles to include actionable, context-specific frameworks. Approaches like "Ethics by Design" (Reed-Berendt et al., 2022) and the FQIR principles offer pathways to embed rights-based protections throughout the research lifecycle. Participatory methods, where affected communities are actively involved in decisions about data use, can further strengthen transparency and accountability, building trust between data producers and subjects.

This review also acknowledges its limitations. Despite efforts to include a wide range of sources, gaps remain, particularly in comparative assessments of the effectiveness and biases of different big data types. Furthermore, the rapidly evolving landscape of data technologies means that some recent developments may not yet be reflected in the literature. Ongoing monitoring and iterative review will be essential.

Future research should prioritise comparative studies that evaluate how big data performs relative to traditional sources in diverse socio-political contexts. Special attention should be given to examining the representational biases in big data and developing methodologies for

their mitigation. Hybrid methodologies that integrate qualitative and quantitative data will be particularly valuable for n capturing lived experiences and ensuring the inclusion of marginalised voices. Qualitative approaches can illuminate how data practices intersect with power, identity, and vulnerability, while quantitative approaches can test the scalability, replicability, and robustness of new tools. Together, such mixed methods can support the development of more inclusive, ethical, and scientifically sound migration data infrastructures.

Finally, the ethical and equitable use of big data has far-reaching implications for global migration governance. If developed and applied responsibly, big data can support more anticipatory, evidence-informed decision-making by states and international organisations. Yet this potential will only be realised if innovation is matched by critical reflection, inclusive governance, and ethical integrity.

In conclusion, while the relevance of big data in migration studies is undeniable, its responsible application requires more than technological advancement. t demands sustained investment in ethical design, interdisciplinary collaboration, institutional capacity, and inclusive participation. Only by addressing these challenges can big data be leveraged not as a tool of exclusion, but as an instrument for empowering diverse migrant populations and advancing a more equitable migration research and policy landscape.

Author contributions

AY: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. TB: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This paper was supported by funding from the MIrreM and HumMingBird projects, which are part of the European Union's Horizon 2020 Research and Innovation Programme under grant agreements No. 101061314 and No. 870661, respectively.

Acknowledgments

We express our gratitude to the European Union for enabling this research.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

References

Adamson, F. B., and Tsourapas, G. (2019). Migration diplomacy in world politics. *Int. Stud. Perspect.* 20, 113–128. doi: 10.1093/isp/eky015

Ahmad Yar, A. W., and Bircan, T. (2023). Big data for official migration statistics: evidence from 29 national statistical institutions. *Big Data Soc.* 10:20539517231210244. doi: 10.1177/20539517231210244

Ahmad Yar, A. W., and Bircan, T. (2025). Challenges with international migration data: an analysis of the experience of national statistical institutions. *Int. Migr. Rev.* 59, 1332–1366. doi: 10.1177/01979183231205564

Ahmad-Yar, A. W., and Bircan, T. (2021). Anatomy of a misfit: international migration statistics. *Sustainability* 13:4032. doi: 10.3390/su13074032

Alexander, M., Polimis, K., and Zagheni, E. (2022). Combining social media and survey data to Nowcast migrant stocks in the United States. *Popul. Res. Policy Rev.* 41, 1–28. doi: 10.1007/s11113-020-09599-3

Ashton, W., Bhattacharyya, P., Galatsanou, E., Ogoe, S., and Wilkinson, L. (2016). Emerging uses of big data in immigration research. Brandon, Canada: Rural Development Institute, Brandon University.

Avramescu, A., and Wisniowski, A. (2021). Now-casting Romanian migration into the United Kingdom by using Google search engine data. *Demogr. Res.* 45, 1219–1254. doi: 10.4054/DemRes.2021.45.40

Aydoğdu, B., Ahat, B., Salah, A. A., and Bircan, T. (2022). Temporal and spatial analysis of indicators on segregation of Syrian refugees in Turkey with mobile phone data. *IEEE.* 1–4.

Aydogdu, B., Yar, A. W. A., Bircan, T., and Salah, A. A. (2023) White paper on policy implications of using mobile phone data for migration. HumMingBird.

Batko, K., and Ślęzak, A. (2022). The use of big data analytics in healthcare. *J. Big Data* 9:3. doi: 10.1186/s40537-021-00553-4

Beduschi, A. (2017). The big data of international migration: opportunities and challenges for states under international human rights law. *Geo. J. Int'l L.* 49:981.

Bircan, T. (2022). "Remote sensing data for migration research" in Data science for migration and mobility studies. eds. A. A. Salah, E. E. Korkmaz and T. Bircan Oxford University Press, 121–148.

Bircan, T., and Korkmaz, E. E. (2021). "Trading efficiency for control: the AI conundrum in migration management" in Research methodologies and ethical challenges in digital migration studies. (Oxford University Press, Springer), 123–140.

Bircan, T., Salah, A. A., and Sîrbu, A. (2023). How can big data analytics help understand migrant integration? In Practitioners' Guide on Harnessing Data Innovation for Migration Policy, ed. M. Rango (In press).

Blumenstock, J., Cadamuro, G., and On, R. (2015). Predicting poverty and wealth from mobile phone metadata. *Science* 350, 1073–1076. doi: 10.1126/science.aac4420

Braaksma, B., Zeelenberg, K., and De Broe, S. (2020). "Big data in official statistics: a perspective from statistics Netherlands" in Big data meets survey science: a collection of innovative methods, eds. C. A. Hill, P. P. Biemer, T. D. Buskirk, L. Japec, A. Kirchner, S. Kolenikov, et al. 303–338.

Carammia, M., Iacus, S., and Wilkin, T. (2022). Forecasting asylum-related migration flows with machine learning and data at scale. *Sci. Rep.* 12:1457. doi: 10.1038/s41598-022-05241-8

Cesare, N., Lee, H., McCormick, T. H., Spiro, E. S., and Zagheni, E. (2018). Promises and pitfalls of using digital trace data for demographic research. *Demography* 55, 1979–1999. doi: 10.1007/s13524-018-0715-2

Chehbi-Gamoura, S., Derrouiche, R., Damand, D., and Barth, M. (2020). Insights from big data analytics in supply chain management: an all-inclusive literature review using the SCOR model. *Prod. Plan. Control* 31, 355–382. doi: 10.1080/09537287.2019.1639839

Claudio, B., Sara, G.-B., Stefano, I., Umberto, M., Francesco, S., and Spyridon, S. Data innovation in demography, migration and human mobility. (2022).

Collins, J. A. (2005). Balancing the strengths of systematic and narrative reviews: Oxford University Press.

Ferrari, R. (2015). Writing narrative style literature reviews. *Med. Writing* 24, 230–235. doi: 10.1179/2047480615Z.000000000329

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., and Gnanzou, D. (2014). How 'big data' can make big impact: findings from a systematic review and a longitudinal case study. *Int. J. Prod. Econ.* 165, 234–246. doi: 10.1016/j.ijpe.2014.12.031

Gendronneau, C., Wiśniowski, A., Yildiz, D., Zagheni, E., Fiorio, L., Hsiao, Y., et al. (2019). Measuring labour mobility and migration using big data. Brussels: European Commission.

Green, B. N., Johnson, C. D., and Adams, A. (2006). Writing narrative literature reviews for peer-reviewed journals: secrets of the trade. *J. Chiropr. Med.* 5, 101–117. doi: 10.1016/S0899-3467(07)60142-6

Greenhalgh, T., Thorne, S., and Malterud, K. (2018). Time to challenge the spurious hierarchy of systematic over narrative reviews? *Eur. J. Clin. Investig.* 48:e12931. doi: 10.1111/eci.12931

Haddaway, N. R., Page, M. J., Pritchard, C. C., and McGuinness, L. A. (2022). PRISMA2020: an R package and shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and open synthesis. *Campbell Syst. Rev.* 18:e1230. doi: 10.1002/cl2.1230

Hayes, B. (2017). Migration and data protection: doing no harm in an age of mass displacement, mass surveillance and "big data". *Int. Rev. Red Cross* 99, 179–209. doi: 10.1017/S1816383117000637

Henningsen, G. (2023). Predicting refugee flows with big data: a new opportunity or a pipe dream? UNHCR Blogs. Available online at: https://www.unhcr.org/blogs/predicting-refugee-flows-with-big-data-a-new-opportunity-or-a-pipe-dream/

Hsiao, Y., Fiorio, L., Wakefield, J., and Zagheni, E. (2023). Modeling the bias of digital data: an approach to combining digital with official statistics to estimate and predict migration trends. *Sociol. Methods Res.* 53:00491241221140144. doi: 10.1177/00491241221140144

Huberty, M. (2015). Awaiting the second big data revolution: from digital noise to value creation. *J. Ind. Compet. Trade* 15, 35–47. doi: 10.1007/s10842-014-0190-4

International Organization for Migration and Joint Research Centre (2023). Harnessing data innovation for migration policy: a handbook for practitioners IOM Publications. Available online at: https://publications.iom.int/books/harnessing-data-innovation-migration-policy-handbook-practitioners

IOM. (2021). Using "big data" to forecast migration. Medium. Available online at: https://medium.com/@UNmigration/using-big-data-to-forecast-migration-8c8e64703559

IOM UN Migration (2022). World Migration Report 2022. NA, Jumana: CUNY Academic Works.

Isaacman, S., Frias-Martinez, V., Hong, L., and Frias-Martinez, E. (2017). Climate change induced migrations from a cell phone perspective. *Poster* 46.

Jumana, R. N. (2024). Is the European Union (EU) a normative power?: An in-depth look at its actions and policies on refugees, asylum seekers and forced migrants.

Juric, T. (2022). *Facebook* and *Google* as an empirical basis for the development of a method for monitoring external migration of Croatian citizens—web of science Core collection. https://www.webofscience.com/wos/woscc/full-record/WOS:000782912400002

Kim, J., Sîrbu, A., Giannotti, F., and Gabrielli, L. (2020). Digital footprints of international migration on twitter. Springer, Cham. 274–286.

Kraler, A., and Reichel, D. (2022). "Migration statistics" in Introduction to migration studies. ed. P. Scholten (Springer, Cham: Springer), 439–462.

Iazzolino, G. (2025). Trading efficiency for control: The AI conundrum in migration management. *Cosmopolitan Civil Societies. Interdiscip. J.* 17, 35–46.

Leese, M., Noori, S., and Scheel, S. (2022). Data matters: the politics and practices of digital border and migration management. *Geopolitics* 27, 5–25. doi: 10.1080/14650045.2021.1940538

Martin, P. L. (2022). Migration and economic development: understanding global migration. Stanford University Press.

Martin, S., and Singh, L. (2018). Data analytics and displacement: using big data to forecast mass movements of people. Georgetown University. Available online at: https://isim.georgetown.edu/publication/data-analytics-and-displacement-using-big-data-to-forecast-mass-movements-of-people/

Meissner, F., and Taylor, L. (2024). Migration information infrastructures: power, control and responsibility at a new frontier of migration research. *J. Ethn. Migr. Stud.* 50, 2227–2246. doi: 10.1080/1369183X.2024.2307772

Momeni, R., Bircan, T., King, R., and Santos, E. Z. (2024). Deciphering climate-induced displacement in Somalia: a remote sensing perspective. *PLoS One* 19:e0304202. doi: 10.1371/journal.pone.0304202

Přívara, A., Rievajová, E., and Barbulescu, A. (2020). Attracting high skilled individuals in the EU: the Finnish experience. *Migrat. Lett.* 17, 369–377. doi: 10.33182/ml.v17i2.927

Reed-Berendt, R., Dove, E. S., and Pareek, M.UK-REACH Study Collaborative Group (2022). The ethical implications of big data research in public health: "big data ethics by design" in the UK-REACH study. *Ethics Hum. Res.* 44, 2–17. doi: 10.1002/eahr.500111

Rönsch, A. (2024). The use of artificial intelligence in migration management-a systematic literature review.

Salah, A. A. (2022). Can big data deliver its promises in migration research? *Intern. Migra.* 60, 252–255. doi: 10.1111/imig.12984

Salah, A. A., Bircan, T., and Korkmaz, E. E. (2022a). "New data sources and computational approaches on migration and human mobility" in Data science for migration and mobility studies. eds. A. A. Salah, T. Bircan and E. E. Korkmaz (Oxford University Press), 3–23.

Salah, A. A., Korkmaz, E. E., and Bircan, T. (2022b). Data science for migration and mobility studies. British Academy Scholarship Online. eds. A. A. Salah, E. E. Korkmaz and T. Bircan Oxford University Press.

Sandberg, M., and Rossi, L. (2022). "Caring for (big) data: an introduction to research methodologies and ethical challenges in digital migration studies" in Research methodologies and ethical challenges in digital migration studies: caring for (big) data? eds. M. Sandberg, L. Rossi, V. Galis and M. Bak Jørgensen (British Academy Scholarship Online, Spyratos: Publications Office of the European Union), 1–21.

Singh, L., Florio, T., Xu, K., Srinivasan, A., Taneja, J., and Martin, S. (2019). Blending noisy social media signals with traditional movement variables to predict forced migration. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (1975–1983). Association for Computing Machinery.

Sîrbu, A., Andrienko, G., Andrienko, N., Boldrini, C., Conti, M., Giannotti, F., et al. (2020). Human migration: The big data perspective. *Intern. J. Data Sci. Anal.* 1–20.

Sîrbu, A., Goglia, D., Kim, J., Magos, P. M., Pollacci, L., Spyratos, S., et al. (2024). International mobility between the UK and Europe around Brexit: a data-driven study. *J. Comput. Soc. Sci.* 7, 1451–1482. doi: 10.1007/s42001-024-00277-4

Smirnov, A. V. (2022). Digital demography methods for forecasting migration processes. *Econ. Reg.* 18, 133–145. doi: 10.17059/ekon.reg.2022-1-10

Spyratos, S., Vespe, M., Natale, F., Weber, I., Zagheni, E., and Rango, M. (2018) Migration data using social media: a European perspective. Publications Office of the European Union.

Struijs, P., Braaksma, B., and Daas, P. J. (2014). Official statistics and big data. *Big Data Soc.* 1:2053951714538417. doi: 10.1177/2053951714538417

Tazzioli, M. (2019). The making of migration: the biopolitics of mobility at Europe's borders. Available online at: https://www.torrossa.com/gs/resourceProxy?an=5018362&publisher=FZ7200

Tjaden, J. (2021). Measuring migration 2.0: a review of digital data sources. *Comp. Migr. Stud.* 9, 1–20. doi: 10.1186/s40878-021-00273-x

Tjaden, J. (2023). Web scraping for migration, mobility, and migrant integration studies: introduction, application, and potential use cases. *Int. Migr. Rev.* 59:01979183231208428. doi: 10.1177/01979183231208428

Tsapenko, I., and Yurevich, M. (2022). Nowcasting migration using statistics of online queries. *Econ. Soc. Chang. Facts Trends Forecast* 15, 74–89. doi: 10.15838/esc.2022.1.79.4

Turper, S. (2023). Big data in migration research: the ethical considerations. *Reflektif J. Soc. Sci.* 4, 795–807. doi: 10.47613/reflektif.2023.139

UN DESA. World Population Prospects 2022. (2020). New York: United Nations Department of Economic and Social Affairs, Population Division.

Utts, J. (2021). Enhancing data science ethics through statistical education and practice. *Int. Stat. Rev.* 89, 1–17. doi: 10.1111/insr.12446

Van der Brug, W., D'Amato, G., Ruedin, D., and Berkhout, J. (2015). The politicisation of migration: Routledge.

Wong, W., and Hinnant, C. C. (2023). Competing perspectives on the big data revolution: a typology of applications in public policy. *J. Econ. Policy Reform* 26, 268–282. doi: 10.1080/17487870.2022.2103701

Wycoff, N., Arab, A., Donato, K., Singh, L., Kawintiranon, K., Liu, Y., et al. (2025). Forecasting Ukrainian refugee flows with organic data sources. *Int. Migr. Rev.* 59, 37–60. doi: 10.1177/01979183231203931

Yildiz, D., Wiśniowski, A., Abel, G. J., Weber, I., Zagheni, E., Gendronneau, C., et al. (2024). Integrating traditional and social media data to predict bilateral migrant stocks in the European Union. *Int. Migr. Rev.* 59:01979183241249969. doi: 10.1177/01979183241249969

Zagheni, E., Weber, I., and Gummadi, K. (2017). Leveraging Facebook's advertising platform to monitor stocks of migrants. *Popul. Dev. Rev.* 43, 721–734. doi: 10.1111/padr.12102

Zhu, C. (2019). Big data as a governance mechanism. *Rev. Financ. Stud.* 32, 2021–2061. doi: 10.1093/rfs/hhy081