



# MIRREM

Measuring Irregular Migration

[www.irregularmigration.eu](http://www.irregularmigration.eu)

## Ethical Benchmarking for the Measurement of Irregular Migration

Norbert Cyrus (University of Osnabrück)

MirreM Policy Brief No.1, October 2023



Funded by  
the European Union

Co-funded by:



UK Research  
and Innovation



Canada Excellence  
Research Chair in  
Migration & Integration

## SUMMARY

**This policy brief is aimed at everyone interested and engaged in or affected by efforts to measure irregular migration and provides an up-to-date outline of ethical dimensions to be considered for such an endeavour.**

---

Irregular migrants generally are considered as a particular vulnerable group. Therefore, it is imperative to consider the ethical dimensions of addressing irregular migration, including efforts to quantify this phenomenon. These considerations need to be grounded in the values of Responsible & Innovation (RRI) framework and Research Integrity, legal data protection regulations, and specifications outlined in sector-specific Codes of Ethical Conduct. This policy brief uses the term ethical standards to comprise ethical requirements. In the last decade, the understanding of what is deemed to be ethically required has evolved substantially due to updating of regulatory frameworks, the advancements of digital technologies and enhancements of ethical sensitivity. Against this background, the policy brief introduces the current regulatory frameworks consisting of legal data protection regulations and domain-specific codes of ethical conduct that stipulate the scope of what is legally permitted and ethically required. The digitisation of the social world has generated an ubiquity of smart-technologies that capture information of incidents, sequences and individuals. Yet, the emergence of machine-learning and artificial intelligence creates new and unforeseen risks, such as the violation of legal regulations and standards of ethical conduct. Finally, the intensive and ongoing efforts to develop measures and tools that support and secure compliance with ethical standards require formation and further finetuning of context-sensitive ethical toolkits that can account for the different contextual particularities. In a nutshell, the current determination of ethical standards displays features of volatility thus demanding a continuous updating.

### ***ABOUT THE AUTHOR:***

Norbert Cyrus is researcher at the Institute of Migration and Intercultural Studies at Osnabrück University, Germany.

### ***TO CITE:***

Cyrus, N. 2023. Ethical benchmarking for the measurement of irregular migration. MIRreM Policy Brief No.1. Krems: University for Continuing Education Krems (Danube University Krems). DOI: 10.5281/zenodo.10042022

# 1. THE RELEVANCE OF ETHICS

---

Ethical standards are codified in legal regulations and non-binding codes of ethical conduct issued by public and private funding agencies, professional bodies and scientific associations.

## Legal regulations addressing ethics

---

Legal regulations are applied mainly with the objective of protecting the individual right to privacy of living persons whose private data is collected, processed or shared. At European level, since 2016 two bodies of law are regarded as the overall foundations. The General Data Protection Regulation (GDPR) stipulates binding principles that must be respected by all using or reusing private data of living persons, among the principles of lawfulness, fairness and transparency; purpose limitation; data minimisation; accuracy storage limitation, integrity and confidentiality; and proportionality (GDPR Art. 5). In practical terms, GDPR specifies that the collection and subsequent processing of personal data requires the informed consent of the data subject except for a ground explicitly mentioned in the GDPR, such as for law enforcement (Gellert 2023). A complementary body of law regulates the collection, processing and sharing of personal data for the purpose of law enforcement. The Law Enforcement Directive (LED) allows law enforcement agencies to require or order natural persons to comply with requests to access personal data. LED aims to ensure that law enforcement authorities can efficiently do their work using technological means while preserving the fundamental rights of data subjects, such as 'the right to be informed' which data is stored and to ask for correction of incorrect data. LED spells out and constrains law enforcements' scope for data use and re-use and stipulates that data protection continues to matter in a law enforcement context. Related to the ongoing establishment of six separate interoperable systems for processing migration related data, further amendments of legal regulations are in the making, notably the European Artificial Intelligence Act (EAIA)

## Relevance of sector-specific Codes of Ethical Conduct

---

Further guidance can be found in the Codes of Ethical Conduct adopted by public and private agencies funding research, professional and academic associations featuring distinct levels of liability. In the research sector, the ethical standards adopted by funding agencies and academic associations are relatively binding. Recently the European Commission emphasised that the fact that research is legally

permissible does not necessarily mean that it will be deemed ethical. In addition to legal regulations, EU funded scientific projects like MIRreM are subject to ethical standards specified by European bodies. The Responsible Research & Innovation (RRI) framework requires researchers to respect fundamental rights and the highest ethical standards to guarantee greater social relevance and greater acceptance of the results of research and innovation. Scientific misconduct like fabrication, falsification or plagiarism is prohibited because it undermines scientific progress and can cause harm. A core value put forward is Research Integrity implying the four principles of reliability, honesty, respect and accountability.

In addition, sector-specific codes of ethical conduct established in the public, business and civic sectors are relevant for an ethical assessment when data is used for research purposes that was processed or will be shared with actors from these sectors. As overarching values, the codes of ethical conduct relate to the core principles of respect for persons (autonomy), minimisation of harm (beneficence), and justice (fairness) and to instrumental principles, e.g. consent, privacy or transparency as means to achieve these ends.

The profound awareness and knowledge of legal regulations and ethical standards is a prerequisite for all individuals and institutions collecting, processing or sharing data (European University Institute 2022). Violations of ethical standards do not only cause harm to the persons whose rights and wellbeing are the subject of protection. Also, those violating ethical standards need to be aware of the risk of sanctions and loss of reputation.

## Difficulties in the compliance and implementation of ethical standards

---

However, even with good will, compliance with ethical standards cannot be taken for granted because ethical standards are formulated as principles that require careful appraisal of relevance in an application context. Often, competing principles must be weighed up, such as, for example, the rights of data subjects in relation to the rights of law enforcement agencies or researchers. The appropriate context-sensitive appraisal is a demanding task requiring awareness, expertise and resources in terms of staff and funds. Breaches may occur due to ignorance or misinterpretation of the relevance, unintentional misconception or deliberate side-stepping and conscious avoidance. In order to avoid the risk of being sanctioned for negligent breaches, in particular unsupported small research projects and civil society organisations may refrain from projects or over-comply beyond what is essentially required and thus impair performance.

## Remedies promoting compliance with ethical standards

---

To prevent negligent and deliberate breaches of ethical standards, standard-setting bodies are advised to determine responsibility and accountability and introduce effective oversight mechanisms. In addition, ethical requirements should be communicated in a clear and easy to understand language. Also further guidance needs to be provided on how to assess and weigh up the relevance and essence of ethical requirements in a particular application context. It is helpful to share systematic

considerations of ethical risks and possible remedies, as the next section aims to provide examples for the efforts to measure irregular migration.

## 2. ETHICAL RISKS AND CHALLENGES IN THE CONTEXT OF MEASURING IRREGULAR MIGRATION

**The determination and implementation of ethical standards needs to be substantiated with regard to a particular context. In the context of efforts to measure irregular migration, a set of interrelating factors effectuate the particular risks related to this context.**

Irregular migrants constitute a difficult-to-reach population seeking to elude official registration and scientific observation. Migration as an empirical issue develops in non-linear and disruptive patterns, thus displaying features of complexity. As a political issue, the governance of migration constitutes a wicked problem because many decision makers and stakeholders with conflicting values are involved and available information is fragmented and confusing. Accordingly, the understanding of what exactly constitute the problem that requires a solution is contested and different policy responses preferred:

- (1) suppressing irregular movements (e.g. border control),
- (2) addressing factors conceived as drivers (e.g. poverty, gaps in refugee protection),
- (3) implementation of pragmatic measures that mitigate negative consequences of irregular migration for individuals and societies (e.g. access to basic rights, regularization).

As an ethical issue, the parallel pursuance of counter-rotating policies constitutes migration policy-dilemmas because the rights of national collectives to prearrange immigration and the rights of individuals to move freely are in tension. The appraisal of effects, effectivity and efficiency of these policy responses towards irregular migration is uncertain due to the lack of reliable quantitative data. A consequent response is the call for the facilitation of reliable quantitative data and analysis. The MirreM project is expected to provide with a focus on the European area an overview of available data, test innovative methodological approaches and assess reliability of data and methods. An appertaining aspect in effort to measure irregular migration is the requisite context-sensitive consideration of the content and implementation of ethical standards (Salah, Canca, & Bariş 2022). In analytic terms, challenges and risks can be grouped in three areas:

- (1) Concerns related to the protection of personal data,
- (2) concerns related to the quality and analysis of data,
- (3) concerns related to harmful practical impacts.

Issues related to the first area address the violation of data protection regulation impairing not only data subjects' rights but also bearing legal consequences for the wrong-doers. The possibility of de-anonymization and re-identification of data shared with other parties is another threat. Consequently, as possible remedies all processing data have to be aware of data protection regulations and consider thorough implementation. With regard to the risk of de-anonymization and re-identification, data owners have to consider effective technical remedies like the possibility to eliminate communicative graphs in Call Details Records (CDR) datasets, a measure that needs to be weighed up against the ethical research principle of transparency.

Concerns related to the second area address the quality of data and analysis. A general challenge is the problem of dataism, the faith in the objectivity of Big Data and ignorance of uncertainty, emerging from insufficient consideration that the so-called raw data is not objective and neutral but emanates from procedures of interest-guided observation and collection and processing of data. The insufficient regard to particular categories of people (e.g. women, people excluded due to statistical conventions or limited access to digital tools) induce statistical biases. With the application of digital technologies, the flaws can be amplified while hiding them behind a façade of mathematical objectivity. The challenge of algorithmic biases is reinforced with the application of machine learning (ML) and artificial intelligence (AI) and may instead of critically reflecting taken-for-granted perspectives reproduce them. Consequently, as possible remedies researchers are advised to pay special attention to the risk to neglect particular categories of people, by taking up insights from social sciences and humanities regarding risks of harm and injustice, and to be aware of special circumstances of the addressed context. With regard to minimizing algorithmic biases, a useful de-biasing approach is the application of protected attribute suppression system (PASS) that discourages a network from encoding protected attribute information. A further remedy provides the application of models of explainable artificial intelligence that can be understood by humans. With respect to ethically sensitive issues, manual sampling of websites that serve as a data basis for machine learning operations is a further option to enhance explainability.

Concerns related to the third area address harmful practical impacts. Risks relate to the influence of business companies, vested interest groups and political bodies. The unevenness of large-scale data sources, the imbalanced representativeness and the potential for uneven effects when used in policy constitute a central concern for all interested in preventing or minimizing the risk of doing harm. A narrow focus on the micro-level of the individual data processors as wrong-doers may divert attention for the ethical dimension at the macro level of companies and governments. Even the application of properly anonymized datasets raises the concern of group privacy when findings from statistical or algorithmic analysis are used to justify actions that have harmful effects for a group. The dual use of technologies constitutes another severe risk. Mobility tracking application at population level may serve the purpose to control a pandemic but can also be used by an autocratic government to control and suppress oppositionists. The intrinsic uncertainty and the dynamics of complexity make for further risks. Technologies can interact with society in unexpected ways as the introduction of content filtering algorithms illustrate: designed to improve user experiences it ended up in polarizing society. Ignorance of complexity increases difficulties to assess the risks of doing harm. Finally, an attitude of algorithmic hubris may induce a focus on digital technology-based solutions at the expense of other possible solving measures. The risks of harmful practical impacts can be partly addressed by an ethical designing of digital operations that recognizes Big Data analysis as potential agent of change, introduces a control framework where measurements should be continuously checked for drift, and

conceptual tools of complex systems should be used instead of simpler but inadequate cause and effect explanations. At the same time, in order to avoid algorithmic hubris, the relevance and preferability of non-digital solutions should be systematically considered, communicated and supported. Regarding uncertainty, researchers should not contribute to an impression of false certainty but clearly emphasize uncertainty by for example describing findings not as a single average score but as a numerical range (e.g. estimated number of irregular migrants between 180,000 – 520,000).

### 3. PROPOSAL FOR AN ETHICAL BENCHMARKING TOOLKIT

**Compliance and implementation of ethical standards is both a practice and the outcome of the practice.**

A guidance for the practices of complying and implementing ethical standards provides an ethical benchmarking toolkit that assembles components developed and made available by researchers and organizations. In the context of Ethics Assessment, the concept of benchmarking is conceived as a tool primarily used for diagnostic purposes to help to identify areas of intervention with regard to the completeness and coherence of compliance with ethical standards. The approach does not define static benchmarks for imposing limits or bans on certain approaches. Rather, its application serves as a criterion (or benchmark) that ethical standards are properly regarded. Ethical Benchmarking is a structured practice approach that does not aim to judge and classify those performing ethical benchmarking as good or bad but aims to support the reflection of good research (in scientific and ethical terms) and its implementation in practical terms thus helping to navigate the efforts to comply with ethical standards.

The introduced ethical benchmarking toolkit synthesizes components of already developed frameworks for an ethical appraisal of research with reference to the Reliable Research & Innovation (RRI) agenda, including the FAIR data principles and the frameworks of AREA and CARE & ACT. Following Leslie (2020, 2023), the ethical benchmarking toolkit is a sequential approach of five steps with an added transversal component. As a basic method, the toolkit designates for each step themes and formulates a structured set of questions thus systematically guiding ethical appraisal.

The first step raises awareness that responsible research should correspond with claims of Open Science and Open Research as means to build public trust. Properly managed accessibility to research findings and maximal data integrity provide the basis to share data according to the FAIR principle (findable, accessible, interoperable and reusable).

The second step addresses the responsibility of researchers to abstain from harmful and damaging research. Referring to the CARE & ACT framework, researchers are supported to consider context; anticipate impacts; reflect on purposes, positionality and power; engage inclusively; and act

responsibly and transparently. The third step addresses the societal constitution characterized by diversity of values, interests and power. Responsible data practices cannot be achieved through merely prescribing sets of core values, providing checklists or even certification processes or audits. The context-sensitive implementation of ethical principles requires the inclusion of various stakeholders and their different perspectives, for example in workshops utilizing the method of Data Ethics Decision Aid (DEDA) in order to gain local and context-sensitive knowledge. The fourth step calls for a reflection of societal responses to research outcomes and impacts. Researchers are reminded to consider and adjust transparency, to clarify and safeguard responsibility and accountability, and to take care that these regimes of transparency and accountability facilitate informed communities and individual consent. The fifth step adheres to the RRI requirements with the particular focus to avoid practices that negatively impact vulnerable and disadvantaged people. The application of tools like the bias audit toolkit for machine learning developed as open source by the University of Chicago or the TU Berlin's datasets and software for detecting algorithmic discrimination help to identify and avoid ethical risks.

Finally, as a transversal component the iterative application of Critique Guided Designing supports and guides the identification of determination of research objectives that contribute to preferable and feasible practical measures promising to contribute to the overcoming of harmful and unjust constellations.

*This policy brief is based on:*

Cyrus, N. (2023). *Ethical benchmarking in irregular migration research*. Zenodo.  
<https://doi.org/10.5281/zenodo.8389189>

## 4. REFERENCES

European University Institute (2022) Guide on Good Data Protection Practice in Research. Fiesole: European University Institute  
<https://www.eui.eu/documents/servicesadmin/deanofstudies/researchethics/guide-data-protection-research.pdf>

Gellert, R. (2023). Data Protection Law and Responsible Data Science. In W. Liebrechts, W.-J. Van Den Heuvel, & A. Van Den Born (Eds.), *Data Science for Entrepreneurship* (2023, pp. 413–439). Springer International Publishing. [https://doi.org/10.1007/978-3-031-19554-9\\_17](https://doi.org/10.1007/978-3-031-19554-9_17)

Leslie, D. (2023). The Ethics of Computational Social Science. In E. Bertoni et al. (2023, pp. 57–104). [https://doi.org/10.1007/978-3-031-16624-2\\_4](https://doi.org/10.1007/978-3-031-16624-2_4)

Leslie, D. (2020) Tackling Covid-19 through Responsible AI Innovation: Five Steps in the Right Direction. *Harvard Data Science Review* (Special Issue 1)  
<https://doi.org/10.1162/99608f92.4bb9d7a7>

Salah, A. A., Canca, C., & Barış, E. (2022). Ethical and Legal Concerns on Data Science for Large-Scale Human Mobility. In Salah et al. (2022, pp. 25–50).



<https://doi.org/10.5871/bacad/9780197267103.003.0002>

---

## Suggested

## Readings

Bertoni, E., Fontana, M., Gabrielli, L., Signorelli, S., & Vespe, M. (Eds.). (2023). Handbook of Computational Social Science for Policy. Springer International Publishing.

<https://doi.org/10.1007/978-3-031-16624-2>

*The handbook describes foundational issues, methodological approaches and examples on how to analyse and model data using Computational Social Science (CSS) for policy support.*

---

Salah, A. A., Korkmaz, E. E., & Bircan, T. (Eds.). (2022). Data Science for Migration and Mobility (1st ed.). British Academy. <https://doi.org/10.5871/bacad/9780197267103.001.0001>

*The collection provides an overview of the major data sources and links them to migration and mobility from multiple aspects in a way accessible to both migration scholars and data scientists.*

---

Sandberg, M., Rossi, L., Galis, V., & Bak Jørgensen, M. (Eds.). (2022). Research Methodologies and Ethical Challenges in Digital Migration Studies: Caring For (Big) Data? Springer International Publishing. <https://doi.org/10.1007/978-3-030-81226-3>

*The volume focuses on the role and responsibility of a researcher and investigates the methodological and ethical dilemmas when working with digital technologies and large-scale datasets.*

---

## THE MIRREM PROJECT

**MirreM examines estimates and statistical indicators on the irregular migrant population in Europe as well as related policies, including the regularisation of migrants in irregular situations.**

MirreM analyses policies defining migrant irregularity, stakeholders' data needs and usage, and assesses existing estimates and statistical indicators on irregular migration in the countries under study and at the EU level. Using several coordinated pilots, the project develops new and innovative methods for measuring irregular migration and explores if and how these instruments can be applied in other socio-economic or institutional contexts. Based on a broad mapping of regularisation practices in the EU as well as detailed case studies, MirreM will develop 'regularisation scenarios' to better understand conditions under which regularisation should be considered as a policy option. Together with expert groups that will be set up on irregular migration data and regularisation, respectively, the project will synthesise findings into a Handbook on data on irregular migration and a Handbook on pathways out of irregularity. The project's research covers 20 countries, including 12 EU countries and the United Kingdom.

More information on the project is available at <http://irregularmigration.eu>.

Funded by the European Union. Views and opinions expressed are however those of the author(s)

---

only and do not necessarily reflect those of the European Union or the Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

In addition, MirreM benefit from funding provided by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee. The Canadian research component of this project is undertaken, in part, thanks to funding from the Canada Excellence Research Chairs Program of the Government of Canada.

## COPYRIGHT CLAUSE



This work is licensed by the MirreM Consortium under a Creative Commons Attribution-ShareAlike 4.0 International License, 2020. For details, see <http://creativecommons.org/licenses/by-sa/4.0/>