



D3.3 European Handbook on Equality Data



**DATA EQUALITY**

**Digital data to identify, prevent and counter intersectional
discrimination in Phygital AI-based environments**

D3.3 EUROPEAN HANDBOOK ON EQUALITY DATA

Lead Beneficiary	AGF
Deliverable Version	V3
Due Date	01/10/2025
Delivery Date	30/09/2025
Dissemination Level	PU (Public)



Project Information

Grant Agreement Number	GAP-101144460
Project Acronym	DATA EQUALITY
Project Name	Digital data to identify, prevent and counter intersectional discrimination in Phygital AI-based environments
Reference	CERV-2023-EQUAL
Type	CERV
Project Starting Date	02/09/2024
Project Duration	24 months
Coordinator	CORTE DI APPELLO DI VENEZIA, Italy

Document Information

Activity	WP3: Methodology Development
Deliverable Name	D3.3 European Handbook on Equality Data
Due Date	01/10/2025
Delivery Date	30/09/2025
Type	R – Document, report
Dissemination Level	Public (PU)
Main Author(s)	Viviana Gullo (AGF), Paula Cano (FUNDEA), José Luis Salido (FUNDEA)
Contributors	All
Document Reviewers	Giuseppe Corasaniti (Ethics and Compliance Committee), Silvia Pellegrini (ANCI-ER), Eleni Lora Daskalou (KMOP)
Security Reviewer	N/A
Ethics Reviewer	Giuseppe Corasaniti (Ethics and Compliance Committee)

Document History

Version	Issue Date	Author(s)	Content and Changes
0.1	12/09/2025	Viviana Gullo (AGF), Paula Cano (FUNDEA), José Luis Salido (FUNDEA)	First draft of the document shared for experts' review
1.0	19/09/2025	Giuseppe Corasaniti (Ethics and Compliance Committee)	Expert's review
1.1	22/09/2025	Viviana Gullo (AGF)	Second draft of D3.3 shared with the partners for review
2.0	29/09/2025	Silvia Pellegrini (ANCI-ER), Eleni Lora Daskalou (KMOP)	Third Draft revised by partners
2.1	29/09/2025	Viviana Gullo (AGF)	Final draft with partners' integration
2.1	30/09/2025	Marina Caneva (COURTAPPVE)	Final review and deliverable Submission



Co-funded by
the European Union

Disclaimer

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 European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

© Copyright in this document remains vested in the Project Partners



Co-funded by
the European Union

Table of Contents

Introduction	7
1. Background and context.....	9
1.1 Definitions	10
1.2 Data Sources for Equality Statistics.....	10
1.3 Bias in AI	11
1.3.1 Importance of Reducing Bias	11
1.4 Project Foundation	11
2. Methodology Development	13
2.1 Phase 1: Literature Review.....	13
2.2 Phase 2: European Survey.....	13
2.3 Phase 3: Semi-Structured Interviews	14
2.4 Phase 4: Synthesis and Analysis	14
2.5 Phase 5: Multi-Stakeholder Workshops.....	14
2.6 Resulting Methodology	15
3. Data collection	16
3.1 A Multi-Actor Model	17
3.2 Defining core data variables.....	18
3.3 Consent, purpose limitation and inclusion	19
3.4 Risk awareness and mitigation.....	20
4. Data management	21
4.1 Visibility, anonymity and trust	21
4.2 Governance, accountability and operational standards.....	22
4.3 From compliance to ethics	24
5. Data Analysis.....	25
5.1 Integrating quantitative and qualitative data	25

5.2	Mitigating bias in interpretation	27
5.3	Towards transparent and verifiable analysis	27
6.	Data exchange and dissemination.....	29
6.1	The institutional ecosystem	29
6.2	Safeguards and conditions	30
6.3	Dissemination of results.....	31
7.	Avoid bias in the use of AI-and-OSINT tools.....	33
7.1	The structural risks and recommended safeguards.....	33
7.1.1	DATA EQUALITY Technical Roadmap	34
8.	Ethical Aspects	36
8.1	Core ethical principles.....	36
8.2	Ethical limits of OSINT	37
9.	Implementation and Evaluation	38
9.1	Preparatory measures.....	38
9.2	Operational protocols and tools	38
9.3	Training and capacity building	39
9.4	Monitoring and evaluation frameworks	39
9.5	Challenges and enabling factors	39
9.5	Interactive cycles.....	40
	Conclusions	41
	Sources.....	43
	Annexes.....	45
	ANNEX I – TEMPLATE FOR THE MEMORANDUM OF UNDERSTANDING (MoU) FOR DATA EXCHANGE.....	45
	ANNEX II - TEMPLATE FOR THE INFORMED CONSENT FORM FOR SENSITIVE DATA COLLECTION	
	47	
	ANNEX III - PROTOCOL AND ETHICAL GUIDELINES FOR THE USE OF AI AND OSINT TOOLS.....	49

ANNEX IV - CHECKLIST FOR HUMAN-IN-THE-LOOP OVERSIGHT IN AI AND DATA ANALYSIS ...	50
ANNEX V - TEMPLATE FOR THE COMMUNITY VALIDATION BRIEF	51

Acronyms and Abbreviations

Abbreviation	Description
AI	Artificial Intelligence
AIA	Algorithm Impact Assessment
CSO	Civil Society Organisation
D	Deliverable
DPIA	Data Protection Impact Assessment
DPO	Data Protection Officer
EIAs	Ethical Impact Assessments
LEA	Law Enforcement Agency
MFA	Multi-Factor Authentication
MoU	Memorandum of Understanding
NLP	Natural Language Processing
OSINT	Open-Source Intelligence
T	Task
VPN	Virtual Private Network
WP	Work Package

Introduction

The **DATA EQUALITY** project, funded by the European Union under the Citizens, Equality, Rights and Values (CERV) programme (Grant Agreement GAP-101144460), represents a decisive step towards strengthening the European capacity to combat discrimination in all its forms¹. It focuses particularly on contexts where physical and digital realities overlap, and where new technologies such as artificial intelligence (AI) and open-source intelligence (OSINT) have created both opportunities and risks in relation to equality. This *European Handbook on Equality Data* is conceived as a cornerstone output of the project, providing stakeholders with a structured, evidence-based, and ethically grounded framework for collecting, managing, analysing, and exchanging equality data across Member States.

Equality data, understood as both quantitative indicators and qualitative evidence, is essential to reveal hidden patterns of discrimination, monitor the effectiveness of laws and policies, and design targeted interventions². In practice, however, gaps persist: data are fragmented, methodologies differ across Member States, and the risk of bias in AI-driven systems threatens to undermine the fairness and trustworthiness of decision-making. The Handbook responds to the growing concern that discrimination in the European Union is evolving into more complex, systemic, and intersectional forms. With increasing reliance on AI and OSINT tools, there is a significant risk of replicating or amplifying societal biases within automated decision-making systems. The Handbook addresses these challenges by offering guidance that is methodologically rigorous, legally sound, and adaptable to diverse institutional and national contexts.

Building on the research and engagement activities carried out in Work Packages 2 and 3 of the project, the Handbook synthesises lessons learned from literature reviews, surveys, interviews, and multi-stakeholder workshops. It provides practical recommendations on data collection, management, analysis, and dissemination, while also addressing the ethical and technical implications of AI and OSINT tools. In doing so, it equips stakeholders—including civil society organisations (CSOs), law enforcement agencies (LEAs), judicial bodies, and policymakers—with

¹ European Commission (2023), *Citizens, Equality, Rights and Values (CERV) programme*, Grant Agreement GAP-101144460.

² European Union Agency for Fundamental Rights (FRA), *Handbook on Equality Data*, FRA, Vienna, 2021.

a shared reference point for improving cooperation, comparability, and accountability in equality data practices.

The Handbook's methodology is designed to be flexible, adaptable to diverse national contexts while fostering a unified European approach. It provides practical tools, including data indicators, formats, and protocols for collaboration, alongside ethical guidelines for bias detection and mitigation. By promoting fairness, transparency, and accountability, it aims to strengthen institutional capacity and cross-sector cooperation. The ultimate goal is to empower stakeholders to leverage data and emerging technologies responsibly, ensuring they serve as instruments of inclusion rather than exclusion, thereby enhancing the EU's ability to combat discrimination effectively.

This document is structured to serve as both a guide and a reference. It begins with the background and context of equality data, followed by the methodology development process, practical implementation strategies, and evaluation mechanisms. Each section incorporates stakeholder insights to ensure practical relevance. By fostering a shared European methodology, the Handbook contributes to a more equitable society, aligning with the EU's commitment to fundamental rights and non-discrimination.

1. Background and context

The European Union has established a robust legal and policy framework to safeguard equality and combat discrimination. Central to this framework are directives such as the Racial Equality Directive (2000/43/EC) and the Employment Equality Directive (2000/78/EC)³, complemented by legislation on gender equality and disability rights. More recently, the EU Anti-Racism Action Plan, the Gender Equality Strategy, the LGBTIQ Equality Strategy, and the EU Strategy on Combating Antisemitism⁴ have reinforced the Union's commitment to advancing equality. Despite these achievements, implementation across Member States remains uneven, particularly with regard to the collection and use of equality data.

Equality data are essential for measuring progress, identifying gaps, and holding institutions accountable. Yet practices differ markedly across the EU. Some countries have robust equality data frameworks embedded in administrative systems and national statistics, while others lack systematic approaches or avoid collecting sensitive data categories for legal or political reasons. The Fundamental Rights Agency (FRA)⁵ has consistently highlighted the lack of reliable, comparable, and disaggregated equality data as a barrier to effective policy design and enforcement. This fragmentation undermines the capacity of the EU and its Member States to address systemic discrimination comprehensively.

The challenges are compounded by the digital transformation of society. Increasingly, discrimination is data-driven, embedded in the algorithms that govern access to jobs, housing, education, healthcare, and justice. AI and OSINT tools promise greater efficiency in monitoring and analysing discrimination but also carry profound risks. Biased training data can lead to discriminatory outcomes, as illustrated by the Dutch childcare benefits scandal (Toeslagenaffaire), in which minority communities were disproportionately targeted⁶. Predictive policing systems, trained on historically skewed datasets, risk reinforcing cycles of over-policing

³ Council Directive 2000/43/EC (implementing the principle of equal treatment between persons irrespective of racial or ethnic origin); Council Directive 2000/78/EC (establishing a general framework for equal treatment in employment and occupation).

⁴ European Commission: *EU Anti-Racism Action Plan* (2020); *Gender Equality Strategy 2020–2025*; *LGBTIQ Equality Strategy 2020–2025*; *EU Strategy on Combating Antisemitism* (2021).

⁵ European Union Agency for Fundamental Rights. (2020). Compendium of practices on equality data. <https://fra.europa.eu/en/themes/equality-non-discrimination-and-racism/about-compendium>.

⁶ European Parliament, (2022), The Dutch childcare benefit scandal, institutional racism and algorithms. Available at: https://www.europarl.europa.eu/doceo/document/O-9-2022-000028_EN.html.

in marginalised neighbourhoods. The opacity of algorithmic systems further exacerbates these challenges, making it difficult for individuals to seek redress and for institutions to ensure accountability.

To effectively address discrimination and promote equality across the European Union, a clear understanding of key concepts and challenges related to equality data and systemic discrimination is essential. The following sections define critical terms, outline data sources, and highlight the implications of bias, particularly in the context of emerging technologies like artificial intelligence (AI). These definitions and insights provide a foundation for understanding the complexities of data-driven approaches to combating discrimination and fostering inclusion.

1.1 Definitions

- **Equality data:** any piece of information that is useful for the purposes of describing and analysing the state of equality. The information may be quantitative or qualitative in nature. It could include aggregate data that reflect inequalities or their causes or effects in society. Sometimes data that are collected primarily for reasons other than equality-related purposes can be used for producing equality data.
- **Equality data analysis:** The collection and analysis of reliable equality data contribute to the fight against discrimination and for equality, by allowing inequalities to be measured, and targeted solutions to be developed. Equality data is also essential for monitoring whether anti-discrimination legislation, positive action measures and social policy efforts are fulfilling their intended objectives.
- **Systemic discrimination** comes under a range of titles in the literature including “structural discrimination”; “institutional discrimination”; and “systematic discrimination”. It is not specifically defined in international or European legislation.

1.2 Data Sources for Equality Statistics

Equality statistics can be compiled from multiple data sources, such as population censuses, administrative registers, household and individual surveys, victimisation surveys, and attitudinal surveys. Other sources could encompass complaints data (including aggregate profiles of

complainants and offenders, for example), criminal justice data (including court statistics and data on outcomes of court cases, as well as compensation offered/sanctions applied, for example), as well as other avenues of data collection, encompassing discrimination testing, diversity monitoring by employers, CSO and service providers, or data related to artificial intelligence, including training data, data analysed by AI and data on AI outputs.

1.3 Bias in AI

Bias is a systematic error that affects the fairness of decisions. Bias occurs when algorithms are trained on skewed data, leading to discrimination and unfair results. Bias in datasets can lead to discriminatory outcomes in AI. Gender, racial, and socioeconomic biases are common biases. Methods for mitigating bias include data augmentation, resampling, and training for fair representation. Creating ethical and unbiased AI is both a moral and technical imperative.

Systemic discrimination involves the procedures, routines and organisational culture of any organisation that, often without intent, contribute to less favourable outcomes for minority groups than for the majority of the population, from the organisation's policies, programmes, employment, and services.

1.3.1 Importance of Reducing Bias

Biased data undermines anti-discrimination efforts by:

1. Masking discrimination, making it invisible to policymakers and practitioners.
2. Misguiding policy priorities, leading to incorrect targeting of groups or issues.
3. Eroding trust in institutions, particularly among marginalized communities.
4. Reinforcing stereotypes, perpetuating harm and inequality.

1.4 Project Foundation

The Handbook is grounded in a multi-layered evidence base:

- **D2.1 Literature Review:** Analysed existing methodologies, legal frameworks, and AI/OSINT tool usage, identifying gaps like underreporting and lack of standardisation. Best

practices included anonymized reporting platforms (e.g., Dutch national discrimination reporting form) and community involvement.

- **D2.2 Needs Assessment:** A survey (141 respondents) and 45 interviews revealed that 46.8% of stakeholders, particularly in Southern Europe, lack formal data management methodologies. Key challenges include fragmented systems, GDPR constraints, limited resources, and low AI literacy (77% unaware of tools like IBM's AI Fairness 360). Best practices emphasised standardised frameworks and secure data-sharing protocols.
- **D3.1 Workshops:** Three online workshops (April-May 2025) engaged 22 LEA/judicial participants, 16 CSO/academia participants, and 25 mixed-group participants from Greece, Italy, Spain, and Albania. Discussions highlighted manual data collection, limited database access, and ethical concerns with AI, alongside successful collaborations like Bologna's Anti-Discrimination Centre.

This Handbook addresses the lack of harmonised methodologies across EU Member States by offering concrete, adaptable tools and guidelines. It supports the implementation of EU directives and national policies through a unified framework for collecting reliable, disaggregated, and ethically sound equality data. By fostering collaboration and responsible use of AI/OSINT tools, it aims to transform data practices into instruments of inclusion and fairness, aligning with the EU's commitment to fundamental rights.

2. Methodology Development

The methodology for this *European Handbook on Equality Data* was developed through a rigorous, participatory, and evidence-based process under Work Packages 2 and 3 (WP2 and WP3) of the DATA EQUALITY project⁷. It aims to provide a robust, flexible, and scalable framework for collecting, managing, analysing, and exchanging equality data across diverse institutional and national contexts. The methodology addresses both technical and ethical dimensions, ensuring alignment with data protection standards (e.g., GDPR⁸) and human rights principles while mitigating biases in AI-driven systems.

2.1 Phase 1: Literature Review⁹

The process began with a comprehensive literature review, which examined existing methodologies, legal frameworks, and debates on equality data management. It explored the use of AI and OSINT tools, identifying gaps such as a lack of standardisation, underreporting, and limited awareness of bias-mitigation techniques. Key findings included the need for integrated approaches to detect systemic discrimination and the importance of ethical AI design. Best practices, such as anonymised reporting platforms (e.g., Dutch national discrimination form) and community-driven data collection, informed the methodology's foundation.

2.2 Phase 2: European Survey

A multilingual survey was deployed across four EU regions (Southern, Northern, Eastern, Western), engaging 141 respondents from LEAs, CSOs, judicial bodies, and public authorities. The survey covered demographics, current methodologies, tools, and stakeholder cooperation.

⁷ DATA EQUALITY project — Deliverable D2.1 (Literature Review) and methodological notes produced under WP2 and WP3 (2025).

⁸ Regulation (EU) 2016/679 — General Data Protection Regulation (GDPR); see also EDPB guidance and European Commission ex-planatory materials.

⁹ Project Deliverable D2.1.

Results revealed that 46.8% of respondents¹⁰, particularly in Southern Europe, reported no formal methodologies for discrimination data management. Limited resources, fragmented systems, and low AI literacy (77% unaware of tools like IBM's AI Fairness 360) were major barriers. Best practices included standardized data formats and secure sharing protocols.

2.3 Phase 3: Semi-Structured Interviews

Forty-five semi-structured interviews with practitioners from LEAs, CSOs, and judicial bodies provided qualitative insights into real-world data practices. Conducted across Greece, Italy, Spain, and Albania, the interviews highlighted cultural and legal variations, such as GDPR constraints and manual data collection in LEAs. They also identified best practices, including community involvement and pseudonymization techniques, which enhance data inclusivity and reduce bias.

2.4 Phase 4: Synthesis and Analysis

The survey and interview data were synthesised to identify methodological gaps, inconsistencies, and harmonisation opportunities. Key challenges included lack of interoperable databases, language barriers, and ethical concerns with sensitive data (e.g., ethnicity). The analysis emphasized the need for standardized indicators, collaborative protocols, and training to address AI biases.

2.5 Phase 5: Multi-Stakeholder Workshops¹¹

Three online workshops (April-May 2025) engaged 63 participants (22 LEAs/judicial, 16 CSOs/academia, 25 mixed) from Greece, Italy, Spain, and Albania. The LEA workshop highlighted

¹⁰ DATA EQUALITY — Deliverable D2.2, *Needs Assessment: Survey results*, May 2025 (survey dataset and analysis).

¹¹ Project deliverable D3.1.

manual, fragmented data collection and limited database access, with GDPR and confidentiality as barriers. The CSO workshop emphasised diverse methods (e.g., surveys, Google Forms) but noted resource constraints and legal inconsistencies. The mixed workshop used case studies (e.g., biased job recruitment algorithms, predictive policing) to co-develop solutions, such as anonymised data and multidisciplinary oversight. Best practices included Bologna's Anti-Discrimination Centre and community-driven data collection.

2.6 Resulting Methodology

The resulting methodology is a flexible framework with:

- **Data Indicators:** Standardized metrics (e.g., hate crime rates by gender, ethnicity) to ensure comparability.
- **Collaboration Protocols:** Secure, GDPR-compliant data-sharing channels and institutional partnerships.
- **Bias Mitigation:** Techniques like data augmentation, resampling, and human oversight to address AI biases.
- **Training Recommendations:** Structured programs on data handling, AI ethics, and cultural competence. This framework is adaptable to national contexts, ensuring practical relevance while promoting a unified European approach. It empowers stakeholders to collect reliable, disaggregated, and ethically sound data, fostering fairness and accountability in combating discrimination.

3. Data collection

Discrimination data is not simply about numbers and statistics: it must capture lived realities, structural patterns, and intersectional vulnerabilities. Flaws in data may arise from defective collection processes. Such defects can stem from ordinary mistakes, which are not morally culpable, or from morally blameworthy practices in data gathering. When the defect is caused by injustice, it constitutes an “accuracy-affecting injustice”¹².

Furthermore, it has to be highlighted that the lack of equality data, deriving from often restrictive equality data collection practices in Europe, raises issues when it comes to identifying inequality in specific domains such as access to housing, education, healthcare, and work, for various protected groups, limiting access to accurate information about structural inequality in societies¹³.

For these reasons, data collection on discrimination requires not only technical rigour but a framework of **legality** and **inclusivity**.

The collection of equality data represents therefore the foundation upon which monitoring systems, analysis, and effective policy responses to discrimination are built. Without **reliable, disaggregated, and ethically collected data**, structural inequalities remain invisible, and discriminatory practices risk going unaddressed. Yet across Europe, significant divergences persist in how data is gathered, which categories are prioritised, and what safeguards are implemented.

This chapter examines the current state of equality data collection in Europe, outlines common challenges identified through the project’s needs assessment and working groups, and proposes methodological recommendations for a harmonised European approach.

The guiding principles are inclusivity, transparency, and respect for fundamental rights, in line with the EU’s General Data Protection Regulation (GDPR), the European Union Agency for Fundamental Rights (FRA) guidelines on equality data, and international human rights standards.

¹² Hellman, D. (2023). Big data and compounding injustice. *Journal of Moral Philosophy*, 21(1-2), 62-83, p. 2.

¹³ Council of Europe Gender Equality Commission (GEC) and the Steering Committee on Anti-Discrimination, Diversity and Inclusion (CDADi), *Study on the impact of artificial intelligence systems, their potential for promoting equality, including gender equality, and the risks they may cause in relation to non-discrimination*, 2023, p. 35.

3.1 A Multi-Actor Model

As suggested in Equinet's recent outputs *Equality by Design, Deliberation and Oversight*¹⁴, a diverse development team and the engagement of CSOs, affected communities and equality bodies, and incorporating stakeholder input are pivotal elements to develop an AI system which provides clear accountability for fairness outcomes, having embedded the principle of equality in its design.

It is already clear that data collection on discrimination and understanding data-driven discrimination requires a set of subjects to be involved.

Effective data collection indeed is enshrined in the cooperation of diverse actors, each holding unique forms of knowledge:

- **Law Enforcement Agencies (LEAs)** gather incident-based data, often focused on suspected perpetrators of hate crimes, through administrative reporting systems.
- **Judicial authorities** document procedural outcomes and final convictions, usually at a later stage and often protected by confidentiality rules.
- **Equality Bodies** collect and process complaints, offer mediation services, and generate analytical reports to identify systemic discrimination.
- **Civil Society Organisations (CSOs)** fill critical gaps by documenting underreported cases through interviews, surveys, shadow reports, and ethnographic research.
- **Academia** contributes with independent research, validation of methodologies, and theoretical frameworks for understanding structural discrimination.
- **Private Platforms**, particularly social media companies, hold vast quantities of behavioural data relevant to online hate and discrimination but remain opaque in terms of access and accountability.

Therefore, a **coordinated multi-actor model** in which these different contributions are not siloed but aligned through shared standards and trust-based exchange.

¹⁴ Ceros, A. (2025). Explaining 'By-Design' Approaches: Principles and Implementation. *Encoding Equality in AI: Design, Deliberation, and Oversight*, p. 7.

3.2 Defining core data variables

To enable comparability and support intersectional analysis, a **minimum set of core variables** should be established in the data gathering process. These variables should always be collected **transparently, with safeguards and consent**, and should include:

- Race or ethnic origin.
- Gender identity (with open self-identification options).
- Religion or belief.
- Sexual orientation.
- Disability status.
- Migratory background.
- Age (expressed in non-identifying ranges).

Adopting these variables as a minimum common denominator would help overcome the invisibility of certain groups and phenomena, while ensuring comparability across jurisdictions. As mentioned in the EC's 2021 Guidelines on improving the collection and use of equality data, among the common gaps and challenges in the current practices, an over-reliance on proxies causes the incomplete identification of population groups at risk of discrimination¹⁵; indeed, although proxies cannot fully capture, for instance, all potential victims of racial or ethnic discrimination, they can offer a reliable basis for tracking changes over time¹⁶.

One important element to bear in mind concerns the nature of these variables, which fall under “special categories of personal data” under **Article 9 of the GDPR**, and therefore require an explicit legal basis and enhanced protection. As the **EDPB's study on the secondary use of personal data in the context of scientific research** noted, such use is not only permitted but encouraged where it serves public interest and fundamental rights, provided that adequate safeguards are in place¹⁷. Data subjects must understand **why such information is being collected, how it will be stored and used, and how to access or withdraw it**.

¹⁵ European Commission: Directorate-General for Justice and Consumers and High Level Group on Non-discrimination Equality and Diversity. Subgroup on Equality Data, *Guidelines on improving the collection and use of equality data*, Publications Office, 2021, p. 10.

¹⁶ Ibidem, p. 11.

¹⁷ European Data Protection Board, Study on the secondary use of personal data in the context of scientific research, 2021, p. 24.

Another relevant element to add concerns the distinction between victims and perpetrators, making it necessary to differentiate between administrative data (suspects, prosecutions, convictions) and victim-based data (complaints, self-reports). Moreover, to ensure consistency and comparability across systems, recording the reporting channel and procedural outcomes is essential.

While other contextual variables, such as location, type of incident, and timeframe, may be equally essential depending on the purpose of the data collection and subsequent analysis.

3.3 Consent, purpose limitation and inclusion

The principle of **purpose limitation** is central. Data should only be collected for clearly defined objectives, whether investigative, monitoring, or research-related. Consent mechanisms must be:

- **freely given**, without coercion or dependency;
- **revocable at any time**;
- **specific to the stated purpose of processing**;
- **accessible and understandable**, using plain language, multilingual formats, and formats inclusive of persons with disabilities.

Consent procedures must be transparent, multilingual and accessible, ensuring that individuals understand why their data is being gathered, how long it will be kept, and who will have access to it. Where consent is not feasible - for example, in judicial records or mandatory reporting systems - alternative legal bases under the GDPR (Articles 6 and 9) must be carefully documented and communicated.

Data collection must be explicitly guided by its purpose. The information gathered by a police authority for investigative purposes cannot be the same as what an equality body requires for monitoring trends or what researchers need for academic analysis. Being clear about the objective allows actors to determine which data are necessary and proportionate.

Moreover, as already mentioned, equality data must capture intersectionality, recognising that individuals often experience discrimination on multiple grounds simultaneously - for instance, a Muslim woman discriminated against both her religion and gender, or a Roma child facing

barriers in education due to ethnicity and socio-economic background. Disaggregation should therefore be carried out in a way that does not erase these overlaps.

Finally, self-identification is important – although it is equally important not to eliminate gender differences, especially when discrimination occurs in healthcare - and must be included in any form used for collecting data: rigid categories risk excluding or misrepresenting individuals, thus, forms should allow respondents to describe their identities in their own terms.

The way forms and surveys are designed has a direct impact on representation (i.e., excluding non-binary options in gender categories not only marginalises certain individuals, but also produces misleading statistics).

Therefore, when gathering data, the following elements should be included:

- open fields for self-identification alongside standardised categories;
- avoidance of binary or exclusionary classifications;
- accessibility for persons with low literacy or limited digital access and people with disabilities;
- use of cultural mediators or community representatives when working with marginalised groups who may mistrust institutions.

3.4 Risk awareness and mitigation

The risks inherent in equality data collection must be explicitly acknowledged and mitigated, including:

- **re-victimisation**, when victims are repeatedly asked to recount traumatic experiences;
- **data misuse**, such as leaks to the media or discriminatory profiling;
- **indirect discrimination**, when seemingly neutral data points combine to reveal sensitive attributes;
- **underreporting**, especially where institutional mistrust persists.

To address these risks, all equality data projects should include a **Data Protection Impact Assessment (DPIA)**, apply pseudonymisation rather than full anonymisation when visibility of groups is essential, and communicate clearly how collected data will lead to tangible change (for example, new policy measures or improved community support services).

4. Data management

Data management lies at the pillar of equality data practices: while collection determines what information enters the system, management governs how it is stored, protected, and made usable. For discrimination-related data, this stage is particularly sensitive because datasets often include highly personal information that falls within the “special categories” of personal data under the GDPR. Mismanagement can therefore produce harm not only in terms of privacy breaches but also in terms of trust, legitimacy, and the very willingness of individuals and communities to engage with reporting systems. Effective data management must therefore balance three objectives: **safeguarding rights**, **preserving analytical value**, and **enabling interoperability across institutions**.

4.1 Visibility, anonymity and trust

One of the dilemmas to be highlighted in the data management process concerns the tension between anonymisation and visibility: **full anonymisation may protect individuals, but can also erase the very categories that make discrimination measurable** - i.e., removing markers of gender identity or ethnicity. As underlined in the already-mentioned EDPB study, serious issues may arise when anonymisation is applied to certain types of data, even though it was originally envisaged as a safeguard, i.e., location data may reveal specific patterns that make the data subject identifiable, thus resisting effective anonymisation¹⁸. CSOs' representative consulted for the draft of the Handbook emphasised how **anonymisation should never lead to the invisibility of marginalised groups**.

Therefore, **pseudonymisation** - which allows records to be linked without direct identifiers - in many contexts is recommended, together with the use of **cell suppression or aggregation to protect small minority groups from re-identification**.

In this regard, Art. 89(1) of the GDPR clarifies specific obligations for data minimisation for the purpose of research as following:

¹⁸ *Supra*, EDPB, Study on the secondary use of personal data in the context of scientific research, p. 28.

- a) *the use of personal data (that allow for the identification of the data subject) is not required, data should be anonymised;*
- b) *the use of personal data (that allow for the re-identification of the data subject) is required, but the use of easily identifiable information including direct identifiers is not required, personal data should be pseudonymised;*
- c) *the use of non-pseudonymised data, including direct identifiers, is required, an appropriate level of security should be achieved using different measures.¹⁹*

In this context, a key question is whether the data subject must be informed about the anonymisation process itself, given that anonymisation also constitutes a processing operation. When data are transferred between controllers, it is generally the first controller who manages the anonymisation. In the case of pseudonymisation, however, the responsibility for upholding the principle of transparency often shifts in practice to the entity collecting the data - though this may not be the party determining the purposes and means of processing²⁰.

Another relevant issue concerns trust, the lack of which will leave victims refraining from reporting and communities disengaging, undermining the cooperative framework established. Building trust requires **transparency** (namely, clear communication on how data is managed and why), **community involvement** (consultation with affected groups on retention and anonymisation practices), and **ethical safeguards** (protocols against secondary victimisation or discriminatory profiling). Indeed, even the assumption of unsafe storage or misuse can be as damaging as actual breaches.

4.2 Governance, accountability and operational standards

Data management is an organisational task: each institution handling equality data must **clearly designate a responsible officer or unit, accountable for compliance with data protection laws and ethical guidelines**. This aligns with GDPR requirements for Data Protection Officers (DPOs) in public bodies, but a similar responsibility framework should be extended to CSO and academic

¹⁹ *Supra*, EDPB, p. 36.

²⁰ *Ivi*.

projects, with **accountability embedded in day-to-day practice rather than treated as a bureaucratic formality**.

Several operational standards emerged as essential to trustworthy data management:

- Data Protection Impact Assessments (DPIAs): mandatory when involving sensitive categories, serving as a forward-looking tool to identify risks and safeguards.
- Retention limits: data should be kept no longer than necessary; different timelines may apply depending on whether the purpose is research, judicial evidence, or statistical monitoring.
- Secure storage: encryption, access restrictions, and multi-factor authentication for all personnel with access to raw data.
- Audit trails: systems must record who accesses the data and for what purpose, ensuring accountability and detectability of misuse.

Fragmentation is a major obstacle: LEAs, equality bodies, CSOs, and researchers often manage data in **incompatible formats and systems**, hindering comparability and exchange. Harmonisation shall not mean imposing uniform software across Europe, but rather agreeing on common standards for structuring, documenting, and coding data. Therefore, adoption of standard taxonomies (for example, using EU-wide codes for hate crime motivations or FRA-aligned equality indicators) would allow national datasets to be compared and aggregated while respecting local specificities.

Finally, when implementing AI tools, data management should be thorough from the design and development stage – especially when considering the quality of data training the AI – as emphasised in Art. 10(2) of the AI Act²¹, in particular:

- (a) the relevant design choices;
- (b) data collection processes and the origin of data, and in the case of personal data, the original purpose of the data collection;
- (c) relevant data-preparation processing operations, such as annotation, labelling, cleaning, updating, enrichment and aggregation;
- (d) the formulation of assumptions, in particular with respect to the information that the data are supposed to measure and represent;

²¹ Raue, B., et al. (2025, August 21). *Artificial intelligence and fundamental rights: The AI Act of the European Union and its implications for global technology regulation* (Schriften des IRDT | Trier Studies on Digital Law), pp. 78-79.

- (e) an assessment of the availability, quantity and suitability of the data sets that are needed;
- (f) examination in view of possible biases that are likely to affect the health and safety of persons, have a negative impact on fundamental rights or lead to discrimination prohibited under Union law, especially where data outputs influence inputs for future operations;
- (g) appropriate measures to detect, prevent and mitigate possible biases identified according to point (f);
- (h) the identification of relevant data gaps or shortcomings that prevent compliance with this Regulation, and how those gaps and shortcomings can be addressed.

4.3 From compliance to ethics

To conclude, data management must be seen not only as a matter of compliance with GDPR but as an ethical practice as well. Therefore, **embedding “privacy-by-design” and “equity-by-design” principles into all systems managing equality data is pivotal**, meaning that systems should be designed from the outset to maximise both protection and fairness, rather than retrofitted with safeguards after the implementation. By embedding robust governance, technical standards, and participatory safeguards, data management can transform from being a bureaucratic requirement into a cornerstone of rights-based equality monitoring.

5. Data Analysis

Raw data, however comprehensive, acquires meaning only when interpreted. It only acquires meaning once it is interpreted, contextualised, and made actionable. Today's adoption of AI tools in the analysis leads to different views²²:

1. **AI is perceived as an instrument for data processing, analysis, classification of new data and data visualisation** to help with quick, objective, and robust analysis of large amount of data.
2. **Algorithms used for decision-making support systems**, whereby they not only provide the results of data analysis, but they are also used for decision-making with the recommendation of certain solutions.

Indeed, for equality data, the analytical stage is not a neutral technical exercise and cannot be reduced to a purely statistical exercise but is a process that shapes how discrimination is seen, measured, and addressed. Misinterpretation or bias at this stage can reinforce stereotypes or erase lived experiences, while rigorous, participatory analysis can uncover structural inequalities and strengthen accountability.

5.1 Integrating quantitative and qualitative data

Analysing data on discrimination implies having gathered **both quantitative and qualitative data**, and it is important to underline that the different types of data gathered, and consequently analysed, do not reflect diametrically opposed approaches: **both are essential and need to be considered in a complementary approach**²³. To clarify, quantitative analysis is essential for identifying trends, measuring prevalence, and enabling cross-national comparison, and techniques such as regression analysis, disaggregated indicators, and time-series comparison are indispensable to monitor the evolution of discrimination phenomena across time and geography. However, quantitative methods alone cannot explain why discrimination occurs nor capture the nuances of human experience.

²² GAUDEUL, A., ARRIGONI, O., CHARISI, V., ESCOBAR, P. M., & HUPONT, T. I. (2025). The Impact of Human-AI Interaction on Discrimination, Publications Office of the European Union, Luxembourg, 2025, p. 62.

²³ OHCHR. (2012). *Human Rights Indicators: A Guide for Measurement and Implementation*, p. 16-17.

Qualitative methods, such as including thematic coding, narrative analysis, and ethnographic inquiry, are therefore a necessary complement, allowing researchers and practitioners to uncover the contexts, motivations, and systemic barriers that numbers alone cannot reveal. Practitioners²⁴ emphasise the added value of combining structured software environments (such as R or SPSS for quantitative work) with specialised qualitative analysis tools like Atlas.ti or NVivo. **This mixed approach is also pivotal when analysing the online phenomena:** while OSINT and natural language processing platforms can track large volumes of hate speech, their outputs must always be reviewed by human analysts who can understand context, cultural nuance, and irony.

Deciding upon which kind of data to use and interpret is already part of the analysis process: which forms of discrimination to track, which categories to disaggregate, and which experiences to prioritise inevitably shape public understanding and policy responses. Therefore, selecting the indicators should be a participatory effort, involving CSOs, equality bodies and affected communities' representatives to define and validate the indicators to be used in the equality data analysis to depict which inequalities are visible or remain hidden. Indeed, a two-track design is needed when selecting the indicators:

- a. **structural indicators** — e.g., employment rates by ethnic origin, complaint-to-resolution ratios, conviction rates disaggregated by relevant protected characteristics — provide measurable, comparative evidence;
- b. **experiential indicators** — perceived safety, trust in institutions, experiences of profiling — capture subjective but policy-relevant dimensions that often drive behaviours (such as under-reporting).

International measurement guidance supports this double approach, urging that human-rights indicators combine quantitative and qualitative evidence²⁵, ensuring that institutional records are not detached from lived realities, and that data serves both accountability and empowerment purposes.

²⁴ Precisely LEAs, police forces, CSOs, academics and public authorities such as Municipality representatives.

²⁵ *Supra*, OHCHR, *Human Rights Indicators*.

5.2 Mitigating bias in interpretation²⁶

Every stage of the data process is vulnerable to bias, although the analysis requires more attention. **Bias in the data analysis may arise at several stages:** selection bias (non-representative samples), confirmation bias (interpreting data to fit expectations), algorithmic bias (models reflecting skewed training data), and interpretive bias (analysts' framing). These risks are amplified when analysis is conducted in isolation, without transparency, peer review or human oversight in the case of AI tools. Indeed, focusing on the analysis provided by AI, while it is well-known that human oversight is crucial, it should be highlighted that recent studies demonstrated how *human overseers brought their own biases, predispositions, values, and past experiences*, affecting the outcome of decision-making²⁷. Therefore, although embedding human-in-the-loop principles in all AI-assisted processes, ensuring that no automated output is ever treated as definitive, is pivotal, but **designing an AI in line with the principles of fairness and transparency is not sufficient**: an oversight system preventing also human bias from influencing the AI outcomes should also be foreseen, such as regular audits led by diverse review teams established to scrutinise interpretations from multiple perspectives, guidelines and protocols, and training to the subjects involved in the analysis.

Finally, data analysis in the field of discrimination is not only about accuracy but also about legitimacy; thus, victims and affected communities – as anticipated - must see themselves reflected not as passive data points but as active contributors. Therefore, the analysis activities shall be transparent, inclusive, and anchored in fundamental rights.

5.3 Towards transparent and verifiable analysis

The ultimate objective of data analysis is to ensure that results are not only accurate but also reproducible and verifiable. In this regard, transparency is critical because analysts must clearly document their assumptions, methodological choices, and limitations, with reports communicating not only findings but also uncertainties and caveats, avoiding misleading

²⁶ The issues related to bias will be further discussed in the concerned chapter.

²⁷ *Supra*, GAUDEUL, A. The Impact of Human-AI Interaction on Discrimination.

framings that might distort public or institutional understanding. According to the OHCHR's *Human Rights Indicators: A Guide for Measurement and Implementation*, adopting a structured and transparent approach supports the implementation of policy measures to secure the universal realisation of rights, the streamlining of the reporting, the improvement follow-up to the recommendations and concluding observations of the treaty bodies and other human rights monitoring mechanisms, and exercising more effective oversight of the promotion and protection of human rights²⁸.

By integrating quantitative and qualitative methods, grounding indicators in participatory processes, and actively mitigating bias, data analysis reflects both statistical rigour and social justice, to move data beyond abstract numbers to become a transformative tool: one that reveals hidden inequalities, strengthens institutional accountability, and supports more equitable policies across Europe.

²⁸ *Supra*, OHCHR, *Human Rights Indicators*, p. 104.

6. Data exchange and dissemination

Equality data must circulate to have impact: remaining siloed - locked in law enforcement case management systems, confined to academic repositories, or held only by civil society – data cannot inform cross-sectoral prevention, policy design, or judicial reform. Nonetheless, **exchange must be shaped by rigorous safeguards to protect privacy, prevent re-victimisation, and maintain public trust.**

Indeed, considering the implementation of EU law, primarily the DSA and the AIA, with the requirement for increased transparency and assessments of algorithms in safeguarding against discrimination, companies and public bodies using speech detection should be required to share the information necessary to assess bias with relevant oversight bodies and, when possible, publicly²⁹.

Data sharing on discrimination shall enable collaboration and comparability, while safeguarding sensitive information and the rights of those represented in the data.

6.1 The institutional ecosystem

The institutional ecosystem for exchanging data on discrimination typically encompasses LEAs and judicial authorities (administrative and prosecutorial records), equality bodies (complaint registers and analytical outputs), CSOs (qualitative dossiers, shadow reports), academia (independent evaluation and longitudinal studies), and private platforms (social media datasets and related metadata). Despite this ecosystem, data exchanges are rarely formalised, and, for specific situations, may never happen.

Many cross-institutional data flows rely on personal trust, informal agreements, or opaque arrangements that cannot guarantee continuity or accountability. This undermines both the sustainability of cooperation and the credibility of outputs. Therefore, standardised data-sharing agreements that clearly define roles and responsibilities, **establish the legal basis for processing, set access controls and retention periods, and stipulate re-use limitations** are essential, such as

²⁹ FRA. (2022). *Bias in algorithms - Artificial intelligence and discrimination*, p. 11.

Memoranda of Understanding (MoUs), Data Sharing Agreements (DSAs), and clearly articulated protocols.

Those instruments specify roles (e.g., data controller and processor), legal bases, access rights, retention timelines, re-use limitations, and breach notification procedures. These elements are consistent with GDPR principles and European Commission guidance on cross-sectoral cooperation.

6.2 Safeguards and conditions

Data protection concerns include uncertainty regarding the applicability of data protection laws, especially GDPR, whose compliance is a major concern when researching datasets that potentially contain personal data³⁰. When considering the analysis and sharing, for instance, of data of potential perpetrators of hate and discrimination through online comments, even if in a public context, are to be considered personal data³¹, falling under the purview of the GDPR – although the exceptions under Articles 5 (1) (b) and (e), 6 (1) (e) and 89 (1)³². Moreover, since 2018, when the legal expertise *was scarce in the NLP community*, researchers often adopted an overly cautious interpretation of data protection law, often refraining from collecting sensitive data, and even when datasets were compiled and labelled, they were rarely shared due to concerns about compliance, overall contributing to the persistent scarcity of publicly available labelled datasets in EU languages³³.

Nowadays, the situation has not changed much, the excessive caution and avoidance of data sharing, especially by NLP researchers, often due to a lack of knowledge of data protection rules, leading to difficulties in assessing and documenting bias in AI models, also because of the absence of full documentation and available tests for identifying bias and in obtaining datasets³⁴.

Focusing on the **exchange of special category data**, which is legally constrained under the GDPR (requiring a clear legal basis, stricter safeguards, and proportionality), in the anti-discrimination context, public-interest justifications (where permitted by national law) and **explicit consent**

³⁰ *Supra*, FRA, *Bias in algorithm*, p. 74.

³¹ Since it can be easy to trace the author of a comment.

³² Exceptions for the conduct and reproducibility of scientific research in the public interest.

³³ *Ibidem*, FRA, *Bias in algorithm*, p. 74.

³⁴ *Ivi*, p. 15.

mechanisms are common bases, but both **require tightly defined scope, documented proportionality tests, and clear withdrawal options for data subjects**. Institutional exchanges should therefore include prior **Data Protection Impact Assessments (DPIAs)** and **well-documented legal rationales to demonstrate compliance with Articles 6 and 9 of the GDPR**. When considering the use of AI tools, according to the FRA's study on bias in algorithms, EU institutions should address issues of correctly implementing data protection law in relation to data sharing with respect to sensitive data for the purpose of researching and monitoring discrimination through clearer guidance, to avoid misinterpretation of data protection law³⁵.

Cross-border exchange introduces further complexity due to divergent national interpretations of GDPR, differing equality frameworks, data localisation rules, and inconsistent taxonomies, which prevent interoperability. In this regard, **three possible solutions** are recommended:

- developing a shared European framework for describing and classifying discrimination-related data to allow information to be compared and aggregated across countries, without requiring identical national systems;
- promoting secure, decentralised models of cooperation, where data can remain under national or institutional control but be shared in a harmonised way for joint analysis;
- strengthening the capacity of smaller CSOs and equality bodies so they can take part in cross-border exchanges on an equal footing, including through training, resources, and technical support.

6.3 Dissemination of results

Dissemination strategies must reconcile transparency with the protection of individuals and communities. Publishing aggregated trends, dashboards with appropriate aggregation levels, plain-language summaries, and community briefings helps ensure results are accessible without exposing identifiable information. **Publishing practices should include risk-mitigated protocols:** suppressing small cells, aggregating by broader geographies, and contextualising findings to avoid stigmatising communities. **Ethical issues must also be considered in disseminating equality data and analysis results because it requires risk awareness:** publishing patterns or aggregate statistics can unintentionally expose communities when small numbers enable indirect

³⁵ *Supra*, FRA, *Bias in algorithm*, p. 15.

identification. For instance, reporting a rise in hate crimes against a specific religious group in a small locality may risk stigmatising individuals or making them identifiable.

Furthermore, **accountability** is a key element in the dissemination because **communities that contribute to equality data should receive accessible feedback and evidence of how data leads to concrete policy or support actions**³⁶. Therefore, dissemination is not limited to institutional reporting but also plays a crucial role in empowering communities and informing the public.

³⁶ *Supra*, OHCHR, *Human Rights Indicators*, p. 137.

7. Avoid bias in the use of AI-and-OSINT tools

AI and OSINT technologies are increasingly being deployed also by law enforcement agencies, judicial authorities, and other public institutions to monitor online content, identify hate, predict risk, or support decision-making. While these tools hold considerable promise in terms of efficiency and scale, they also carry **profound risks of amplifying structural discrimination if not carefully designed, governed, and monitored**. Moreover, higher consideration shall be given to the **adoption of AI and OSINT tools by public authorities, especially police**, as emphasised in the UNICRI Report on public perceptions of AI in law enforcement, whereby AI's implementation may lead to discrimination and exacerbate unfair treatment, although its general application is perceived as useful or even essential for police work in some cases³⁷.

Data-driven technologies are not neutral, they embody social assumptions, replicate historical patterns of exclusion, and often conceal their inner workings under a veneer of objectivity.

7.1 The structural risks and recommended safeguards

Algorithms trained on historical records tend to reproduce and amplify existing inequities rather than challenge them. For example, **predictive policing systems trained on biased arrest data concentrate surveillance in already over-policed neighbourhoods, thereby reinforcing cycles of criminalisation**. Furthermore, **AI models are often opaque**, their outputs presented as neutral or authoritative despite the absence of transparency or interpretability, and the lack of clear legal accountability means that when automated tools cause harm, responsibility becomes diffuse, and victims face barriers in seeking redress. In addition, OSINT practices raise additional dangers concerning the surveillance creep whereby minority groups, activists, or vulnerable communities are disproportionately monitored, under the justification that data are public.

³⁷ UNICRI. (2024). *Not Just Another Tool - Public Perceptions on Police Use of AI*, pp. 42, 51.

These risks are specifically addressed by the European Data Protection Board regarding **profiling and automated decision-making tools**, which are largely adopted in both private and public sectors, namely finance, healthcare, insurance, etc., significantly impacting individuals' lives.³⁸

To mitigate these risks, technical safeguards and accountability mechanisms shall be adopted, considering that avoiding bias in AI and OSINT requires a comprehensive architecture of technical, ethical, and institutional safeguards. First, all datasets intended for model training must undergo **systematic bias auditing and representativity checks, with fairness validations both before and after deployment**. In addition, all sensitive decisions informed by AI or OSINT must be subject to **human-in-the-loop review**, although, as anticipated, this is not sufficient and requires an **interdisciplinary team and audit mechanism**, with responsibility which must be clearly designated: institutions should appoint **AI Ethics Officers or establish independent advisory boards with the power to review, question, and halt deployments**. Finally, transparency must extend beyond technical documentation, and individuals and affected communities must be offered plain-language explanations of how outputs are generated and how they are used in practice.

7.1.1 DATA EQUALITY Technical Roadmap³⁹

Creating an AI-based OSINT tool that is entirely “bias-free” is complex, but the process can start with robust data collection that avoids subjective interpretation. Such a tool must strictly adhere to the GDPR, ensuring privacy and data protection.

Key compliance elements include: **respecting data subject rights** (access, rectification, erasure), **clear consent management with accessible terms and policies**, **a breach notification system**, **continuous monitoring through audits**, and **thorough due diligence on third-party providers**.

The tool must uphold the highest standards of data security and privacy, aligning with GDPR and international standards such as ISO/IEC 27001:2013. Moreover, providers should have clear

³⁸ EDPB. (2018). Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679.

³⁹ The following section is delineated in D1.2 Technical Roadmap, which presents in more detail the strategic steps needed to develop and implement AI-and OSINT-based solutions to address data-driven discrimination, including the technical adaptation, the evaluation of the existing AI solutions for predictive justice, and the existence of pre-existing bias-free OSINT platforms to gather and elaborate data.

policies for processing, storing, and deleting data within a multi-tenant infrastructure, ensuring strict role-based access controls. Data must be logically and physically isolated, with strong segmentation and encryption during transmission to prevent unauthorised access and protect confidentiality.

Generally speaking, it is crucial to implement an ethical approach in deploying AI and OSINT tools, identifying and mitigating harmful and unjustified disparities produced by these tools⁴⁰.

⁴⁰ EUROPOL. (2025). *AI bias in law enforcement – a practical guide*, p. 33.

8. Ethical Aspects

Data collection, management and exchange must be founded in ethical principles, especially in the field of discrimination and data-driven discrimination, and **subjected to rigorous control and monitored by all the subjects involved**. This approach embeds a human-rights approach grounded in the EU Charter of Fundamental Rights, the GDPR, and international standards such as the OHCHR's guidance on human-rights indicators and the EUROPOL, UNICRI and FRA documents on AI and bias. Ethical governance is therefore both legal and moral: it demands compliance with binding rules and continuous ethical reflection about how data practices shape social realities.

8.1 Core ethical principles

Legitimacy is the ultimate goal of ethical data practice, meaning that collecting, managing and exchanging data on discrimination under precise ethical principles guarantees reliability, effectiveness, and reducing rather than reproduces discriminatory behaviours.

The core ethical principles to be followed are:

- **Transparency:** data subjects must know what data are collected, why, and by whom. Practices must be intelligible to the public.
- **Accountability:** institutions must assign clear responsibility, especially in AI-assisted or sensitive processes.
- **Fairness and non-discrimination:** systems must be designed to reveal, not reproduce, inequalities; datasets and outputs must be regularly audited.
- **Explainability:** results from automated systems must be presented in plain language. Opacity itself is a form of inequality.
- **Proportionality and purpose limitation:** data use must be justified by clear benefits that outweigh risks. Each purpose must be documented.
- **Participation:** communities must be active partners in shaping methodologies, not passive data sources.

Ethics must go beyond compliance: GDPR and national law constitute the baseline, but institutions must adopt practices that **anticipate harms and embed responsibility into**

institutional culture, including conducting **Ethical Impact Assessments (EIAs)** in parallel with **DPIAs**, **designating ethics advisors or establishing standing ethics committees**, and **involving CSOs and community** actors in governance roles not limited to consultation but extending to oversight and co-decision when feasible.

Embedding ethics in institutions means designing processes that enable revision - where indicators, tools, or definitions perpetuate harm, institutions must be prepared to halt, correct, and re-engage affected communities.

Ethical data governance, in this sense, is an ongoing process between rights, risks, and responsibilities.

8.2 Ethical limits of OSINT

OSINT also demands particular attention to purpose limitation and proportionality. The public availability of content does not, by default, justify broad harvesting or retention: the intended use, the degree of identifiability, and foreseeable harms must be assessed before collection. Where possible, **personal identifiers should be removed or pseudonymised and retention minimised**. Ethical oversight should involve community stakeholders and CSOs, **particularly when monitoring content from marginalised populations, so that monitoring practices do not replicate surveillance inequalities**.

9. Implementation and Evaluation

The implementation of the European Handbook on Equality Data must be conceived as a **dynamic and iterative process**, rather than a static transfer of guidelines from the European to the national or institutional level. The Handbook itself has been designed as both a **normative reference** and a **practical toolkit**; its successful roll-out requires that stakeholders not only **understand its principles but also adapt them** to the diverse legal, cultural, and technological landscapes that exist across Europe. From the outset, implementation should be guided by the **principle of harmonisation without homogenisation**: ensuring comparability of practices and outcomes across Member States while leaving sufficient room for national specificities, cultural traditions, and institutional capacities.

9.1 Preparatory measures

The first stage of implementation is preparatory. Institutions at national and local levels must **internalise the key concepts** introduced in the Handbook, such as the distinction between equality data and general administrative statistics, or the risks and opportunities associated with AI and OSINT tools. This requires **awareness-raising campaigns**, introductory training modules, and the **designation of focal points** within organisations who can translate European-level recommendations into internal protocols. Only once this conceptual foundation has been laid can institutions begin to embed the Handbook's recommendations into their working routines.

9.2 Operational protocols and tools

Once institutions are familiar with the basic principles, the second stage involves the **development of operational protocols**. Here, the role of **memoranda of understanding**, **standardised templates**, and **secure communication channels** becomes central. Implementation cannot be reduced to the isolated adoption of indicators or forms; it must be embedded in everyday practices. For example, law enforcement agencies need to integrate equality data

variables into incident-reporting software, while CSOs may require adaptable digital platforms for anonymous self-reporting. The process is **iterative**: initial pilot schemes must be reviewed and adapted on the basis of feedback from both staff and affected communities.

9.3 Training and capacity building

The third stage of implementation focuses on **training and professional development**. As highlighted in both the survey and interview results, stakeholders consistently reported insufficient resources and methodological fragmentation. The Handbook proposes a modular training system that combines technical instruction on **GDPR compliance** and **bias-mitigation techniques** with broader reflections on **ethical principles** and **intersectionality**. Evaluation metrics at this stage include not only the number of professionals trained but also the extent to which training translates into tangible improvements in the quality of data collected and in the trust communities place in reporting systems.

9.4 Monitoring and evaluation frameworks

Monitoring and evaluation mechanisms must be embedded from the outset rather than added as an afterthought. In line with GDPR principles of **accountability** and **transparency**, institutions should be required to **conduct regular Data Protection Impact Assessments (DPIAs), Algorithmic Impact Assessments (AIAs), and Ethical Impact Assessments (EIAs)**. These instruments allow not only for compliance monitoring but also for the **proactive identification of risks** before they materialise. Evaluation should combine **quantitative indicators**—such as the number of datasets shared across borders or the frequency of joint reports—with **qualitative measures**, including perceptions of fairness, community satisfaction, and improvements in institutional trust.

9.5 Challenges and enabling factors

Despite these aspirations, implementation faces several foreseeable challenges. **Legal fragmentation** remains a significant obstacle: while some Member States have robust frameworks for equality data, others restrict or prohibit the collection of sensitive variables. **Resource disparities** are equally problematic; smaller CSOs and local authorities often lack the technological infrastructure to engage with advanced tools. **Cultural barriers** must also be recognised: in certain contexts, discussing race, religion, or sexual orientation remains politically sensitive or even taboo, making data collection an inherently contested practice. Finally, the **rapid evolution of AI and OSINT technologies** means that implementation must remain flexible and ready to adapt to new risks as they emerge.

9.5 Interactive cycles

Ultimately, implementation is not a linear trajectory but an **iterative cycle** of testing, feedback, and refinement. The Handbook should be understood as a living document whose success depends on its **capacity to evolve**. Evaluation reports, case studies, and comparative analyses across Member States should be fed back into subsequent editions of the Handbook, creating a virtuous circle between practice and theory. The implementation and evaluation of this Handbook will therefore be measured not only by the number of datasets collected or policies informed but by its contribution to strengthening trust, transparency, and fairness in European societies.

Conclusions

The European Handbook on Equality Data brings together the insights, challenges, and opportunities identified throughout the DATA EQUALITY project and translates them into a coherent methodology for practice. At its core, the Handbook responds to a pressing need: despite the EU's robust legal framework against discrimination, the absence of harmonised and reliable data has long undermined the capacity of institutions to act effectively. By systematising both technical standards and ethical guidelines, the Handbook offers a concrete roadmap towards more inclusive, transparent, and accountable practices.

The Needs Assessment (D2.2) made it clear that fragmentation is the defining feature of current practices. Almost half of surveyed stakeholders reported the absence of any formal methodologies for data collection or management, while more than three quarters admitted to lacking awareness of AI tools capable of detecting or mitigating bias. This baseline reflects not only institutional weaknesses but also broader structural inequalities across regions: Southern Europe, for example, reported particularly acute challenges, including insufficient coordination and limited technical capacity. By contrast, Northern and Western Europe revealed somewhat stronger frameworks but also highlighted the persistent lack of harmonisation at the transnational level.

The contributions of this Handbook are several. First, it consolidates a multi-actor model that affirms the complementary roles of LEAs, CSOs, equality bodies, and judicial authorities. No single actor can adequately capture the complexity of discrimination, but together, these institutions can produce a more accurate and comprehensive picture. Second, it establishes a minimum set of core variables—race or ethnic origin, gender identity, religion or belief, sexual orientation, disability, migratory background, and age—that can serve as a baseline for comparable analysis. Third, it provides methodological guidance for embedding safeguards into data collection, including purpose limitation, consent mechanisms, and pseudonymisation techniques. These contributions strengthen not only the technical rigour but also the ethical legitimacy of equality data practices across Europe.

Yet, the Handbook also acknowledges its limits. The political sensitivity of equality data, particularly regarding ethnicity and religion, means that implementation will continue to face

resistance in certain Member States. Legal fragmentation, including divergent interpretations of GDPR, creates uncertainty that no methodological handbook can resolve on its own. Resource disparities remain acute, and without dedicated funding, the recommendations risk being aspirational rather than operational. Moreover, the rapid development of AI and OSINT tools introduces new risks that require continuous monitoring and adaptation.

Looking forward, the Handbook should serve not as an endpoint but as a foundation. Its recommendations need to be tested, refined, and expanded through practice. EU institutions have a role to play in ensuring that Member States adopt the core standards while respecting national specificities. Funding streams should be directed towards capacity-building, particularly in regions and sectors where resources are most scarce. Multi-stakeholder platforms must be strengthened to foster trust and collaboration, while evaluation cycles must remain transparent and participatory. Only through continuous engagement and adaptation will the Handbook fulfil its potential.

Ultimately, the significance of this Handbook lies in its transformative potential. By providing a structured yet adaptable methodology, it equips stakeholders to convert fragmented practices into coordinated strategies; by embedding ethical safeguards, it ensures that data collection becomes a tool for inclusion rather than exclusion; and by emphasising transparency and accountability, it strengthens the democratic legitimacy of institutions across Europe. Its success will be measured not merely by the quantity of data collected, but by the extent to which that data empowers communities, informs fairer policies, and prevents discrimination in both physical and digital environments.

In conclusion, the European Handbook on Equality Data is both a product of collective research and a catalyst for institutional change. It synthesises the lessons of the past, addresses the challenges of the present, and anticipates the risks of the future. It is offered as a living document that will require continuous engagement, revision, and adaptation. If implemented with care, responsibility, and determination, it has the capacity to reshape how Europe understands and addresses discrimination, making equality data a cornerstone of democratic governance and social justice.

Sources

- Ceros, A. (2025). *Explaining ‘By-Design’ Approaches: Principles and Implementation. Encoding Equality in AI: Design, Deliberation, and Oversight.*
- Council Directive 2000/43/EC (implementing the principle of equal treatment between persons irrespective of racial or ethnic origin).
- Council Directive 2000/78/EC (establishing a general framework for equal treatment in employment and occupation).
- Council of Europe Gender Equality Commission (GEC) and the Steering Committee on Anti-Discrimination, Diversity and Inclusion (CDADi). (2023). *Study on the impact of artificial intelligence systems, their potential for promoting equality, including gender equality, and the risks they may cause in relation to non-discrimination.*
- DATA EQUALITY project — Deliverable D2.1 (Literature Review) and methodological notes produced under WP2 and WP3 (2025).
- DATA EQUALITY — Deliverable D2.2, *Needs Assessment: Survey results.*
- EDPB. (2018). *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679.*
- European Commission (2023). *Citizens, Equality, Rights and Values (CERV) programme, Grant Agreement GAP-101144460.*
- European Commission. (2020). *EU Anti-Racism Action Plan.*
- European Commission. (2020). *Gender Equality Strategy 2020–2025.*
- European Commission. (2020). *LGBTIQ Equality Strategy 2020–2025.*
- European Commission. (2021). *EU Strategy on Combating Antisemitism.*
- European Commission: Directorate-General for Justice and Consumers and High Level Group on Non-discrimination Equality and Diversity. (2021). *Subgroup on Equality Data, Guidelines on improving the collection and use of equality data.* Publications Office.
- European Data Protection Board. (2021). *Study on the secondary use of personal data in the context of scientific research.*
- European Parliament, (2022), The Dutch childcare benefit scandal, institutional racism and algorithms. Available at: https://www.europarl.europa.eu/doceo/document/O-9-2022-000028_EN.html.
- European Union Agency for Fundamental Rights (FRA). (2021). *Handbook on Equality Data.* FRA, Vienna.
- European Union Agency for Fundamental Rights (FRA). (2020). *Compendium of practices on equality data.* <https://fra.europa.eu/en/themes/equality-non-discrimination-and-racism/about-compendium>
- FRA. (2022). *Bias in algorithms - Artificial intelligence and discrimination.*
- GAUDEUL, A., Arrigoni, O., Charisi, V., Escobar, P. M., & Hupont, T. I. (2025). *The Impact of Human-AI Interaction on Discrimination.* Publications Office of the European Union, Luxembourg.

- Hellman, D. (2023). *Big data and compounding injustice*. Journal of Moral Philosophy, 21(1-2), 62-83.
- OHCHR. (2012). *Human Rights Indicators: A Guide for Measurement and Implementation*.
- Raue, B., et al. (2025, August 21). *Artificial intelligence and fundamental rights: The AI Act of the European Union and its implications for global technology regulation (Schriften des IRDT | Trier Studies on Digital Law)*.
- Regulation (EU) 2016/679 — *General Data Protection Regulation (GDPR)*.
- UNICRI. (2024). *Not Just Another Tool - Public Perceptions on Police Use of AI*.
- EUROPOL. (2025). *AI bias in law enforcement – a practical guide*.

Annexes

ANNEX I – TEMPLATE FOR THE MEMORANDUM OF UNDERSTANDING (MoU) FOR DATA EXCHANGE

This Memorandum of Understanding (MoU) is entered into by and between:

[Institution A Name] and [Institution B Name]

1. Purpose

The purpose of this MoU is to establish a cooperative framework for the ethical and secure exchange of data related to [discrimination, hate crime, or hate speech], in accordance with the General Data Protection Regulation (GDPR) and applicable national laws.

2. Scope

This MoU covers the sharing of [specify type of data: e.g. anonymised incident records, survey results, complaints, etc.], for the purposes of [e.g. joint research, monitoring, policy development].

3. Roles and Responsibilities

Each party agrees to:

- Appoint a data protection lead for all data handling activities.
- Ensure legal compliance, including valid legal basis for processing (Articles 6 and 9 GDPR).
- Apply data minimisation and purpose limitation principles.
- Maintain secure data transmission and storage systems.

4. Data Protection Measures

- Use encryption or pseudonymisation where appropriate.
- Restrict access to authorised personnel only.
- Conduct and document Data Protection Impact Assessments (DPIAs) where required.

5. Duration and Termination

This MoU shall be valid for [insert period], and may be terminated by either party with [insert notice period].

6. Dispute Resolution

Disputes shall be resolved through negotiation between parties or, failing that, through a mutually agreed mediation mechanism.

Date and place,

Signature

ANNEX II – TEMPLATE FOR THE INFORMED CONSENT FORM FOR SENSITIVE DATA COLLECTION

[Title of the document]

[Organisation responsible]

Contact Person:

We are collecting information for the purpose of [state purpose: research, monitoring, policy development] on discrimination-related issues. Your participation is entirely voluntary.

1. What data will be collected?

We will collect [e.g. race or ethnicity, gender identity, religion, disability, etc.]. You may decline to answer any question.

2. How will your data be used?

Your data will be anonymised or pseudonymised and used for [scope of use: analysis, reporting, sharing with XX].

3. Who will have access to your data?

Only the project team and authorised partners will have access, under strict confidentiality agreements.

4. How long will your data be stored?

5. Data will be stored securely for a maximum of [insert duration], after which it will be deleted.

6. What are your rights?

You have the right to access, correct, or request deletion of your data at any time.

7. Withdrawal

You may withdraw your consent at any time without consequences.

☐ I have read and understood the information provided.

☐ I voluntarily consent to the collection and use of my data for the stated purposes.

Name:

Date and place

Signature

ANNEX III – PROTOCOL AND ETHICAL GUIDELINES FOR THE USE OF AI AND OSINT TOOLS

This protocol establishes minimum ethical conditions for the deployment of AI and OSINT tools in the context of discrimination-related data.

1. Risk Assessment
 - Conduct a risk and impact assessment before deployment.
 - Identify potential harms to affected communities.
2. Human Oversight
 - Ensure human-in-the-loop decision-making at all stages.
 - Assign legal responsibility to a designated officer or team.
3. Bias Prevention
 - Audit datasets for representativity and completeness.
 - Apply fairness metrics to model outputs.
4. Transparency
 - Document system logic and assumptions.
 - Provide accessible explanations of system outputs to users.
5. Community Engagement
 - Validate indicators and analytical outcomes with community stakeholders where feasible.
6. Proportionality
 - Ensure the use of automated tools is proportionate to the intended purpose and avoids overreach.

ANNEX IV – CHECKLIST FOR HUMAN-IN-THE-LOOP OVERSIGHT IN AI AND DATA ANALYSIS

This checklist supports organisations in identifying where and how human oversight is required.

- ☐ Are humans involved in the selection of training data?
- ☐ Has the dataset been reviewed for potential bias?
- ☐ Are model parameters and logic documented and peer-reviewed?
- ☐ Is there a named person responsible for outputs of automated systems?
- ☐ Are users required to manually review system recommendations before action?
- ☐ Are communities consulted on the interpretation of outputs?
- ☐ Are override or appeal mechanisms in place?
- ☐ Has the ethical impact of automation been evaluated?

ANNEX V – TEMPLATE FOR THE COMMUNITY VALIDATION BRIEF

[Title of the document]

Validation: [Community Organisation Name(s)]

Objective: To involve affected communities in reviewing, validating, or shaping data indicators, interpretation, and use.

Scope

- Review of indicators and definitions
- Interpretation of preliminary findings
- Ethical considerations and data sensitivity

Process

- Participatory workshop / online consultation/survey [choose one or more]
- Facilitated by: [Organisation or Independent Expert]
- Summary of feedback will be incorporated into final outputs

By signing below, the undersigned confirm their involvement in the validation process.

Date and place

Signature



www.project.website.eu



Co-funded by
the European Union