



# High Level Group on Non-discrimination, Equality and Diversity

2018

## Subgroup on Equality Data

### Guidelines on improving the collection and use of equality data

**EUROPEAN COMMISSION**

Directorate-General for Justice and Consumers

Directorate D – Equal Opportunities Equality and Union citizenship

Unit D1 - Non-discrimination and Roma coordination

*European Commission*

*B-1049 Brussels*

# **Guidelines on improving the collection and use of equality data**

Manuscript completed in July 2018

PDF	ISBN 978-92-76-40995-3	doi:10.2838/9725	DS-01-21-146-EN-N
-----	------------------------	------------------	-------------------

Luxembourg: Publications Office of the European Union, 2021

© European Union, 2021

The reuse policy of European Commission documents is implemented by the Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Except otherwise noted, the reuse of this document is authorised under a Creative Commons Attribution 4.0 International (CC-BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>). This means that reuse is allowed provided appropriate credit is given and any changes are indicated.



## Contents

Section I: Background and purpose of these guidelines.....	3
Identifying gaps in the collection and use of equality data .....	5
How were these guidelines developed? .....	8
Section II: Guidelines on improving the collection and use of equality data .....	10
A. Institutional and structural guidelines .....	10
1. Map existing sources of equality data and identify data gaps .....	10
2. Foster inter-institutional cooperation in the collection and use of equality data.....	11
3. Set up a data hub on equality and non-discrimination.....	12
4. Build institutional capacity to collect robust and reliable equality data .....	12
5. Facilitate effective use of equality data.....	13
B. Operational guidelines.....	14
6. Ensure comprehensiveness of equality data .....	14
7. Mainstream equality data into EU and national surveys.....	14
8. Ensure regular and timely equality data collection .....	15
9. Enhance validity and reliability of equality data.....	15
10. Ensure representativeness of equality data .....	17
11. Improve comparability of equality data .....	17

**This document is not legally binding and is intended for guidance only. It can, therefore, neither provide legal advice on issues of national law, nor an authoritative interpretation of EU law, which remains within the sole remit of the Court of Justice of the European Union.**

## Section I: Background and purpose of these guidelines

The European Union (EU) has in place an advanced legal framework with which to promote equality and non-discrimination. All 28 EU Member States have transposed this legal framework into national law, often going beyond the minimum standards included in the Racial Equality Directive<sup>1</sup> and the Employment Equality Directive.<sup>2</sup>

However, significant proportions of people in the European Union experience discrimination, inequality and social exclusion on a regular basis. This can be based on their gender, ethnic or minority background, skin colour, religious belief, age, sexual orientation, gender identity or disability, or a combination of these, as findings from the EU Agency for Fundamental Rights (FRA) consistently show.<sup>3</sup>

This calls for a reconsideration of how legislation and policies to promote equality and non-discrimination are implemented and how progress on the ground can be monitored and measured. Equality data are a crucial element of this reconsideration and powerful tools to support the fight against discrimination and exclusion.

This includes, in particular, the disbursement of European Union Structural and Investment Funds, where horizontal EU principles such as the aim to combat discrimination based on sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation laid down in article 10 TFEU need to be fully respected. This requires Member States to have arrangements in place, including monitoring through the systematic collection of robust and objective data, that ensure the fulfilment of their fundamental rights obligations.<sup>4</sup>

For historical reasons, the collection of personal data disaggregated by protected characteristics such as racial or ethnic origin is a particularly sensitive issue in many European countries. EU antidiscrimination legislation forbids the use of any such data to the disadvantage of the groups they refer to. Furthermore, such sensitive personal data are severely protected by constitutional norms, the applicable European Union data protection law and the Charter of Fundamental Rights. However, if collected and processed in full respect of this legal framework and the safeguards it sets out, such data are essential for Member States to assess their compliance with human rights obligations and enables policy makers to design evidence-based measures to address discrimination, inequalities and exclusion.

---

<sup>1</sup> [Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin.](#)

<sup>2</sup> [Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation.](#)

<sup>3</sup> See for example, FRA (2017a), [Second European Union Minorities and Discrimination Survey \(EU-MIDIS II\) Muslims – Selected findings](#); FRA (2017b), [Second European Union Minorities and Discrimination Survey – Main results](#); FRA (2016), [Second European Union Minorities and Discrimination Survey \(EU-MIDIS II\) Roma – Selected findings](#); FRA (2014), [EU LGBT survey - European Union lesbian, gay, bisexual and transgender survey - Main results](#); FRA (2013), [Discrimination and hate crime against Jews in EU Member States: experiences and perceptions of antisemitism.](#)

<sup>4</sup> See [Proposal for a Regulation of the European Parliament and of the Council laying down common provisions on the European Regional Development Fund, the European Social Fund Plus, the Cohesion Fund, and the European Maritime and Fisheries Fund and financial rules for those and for the Asylum and Migration Fund, the Internal Security Fund and the Border Management and Visa Instrument](#), COM(2018) 375 final, 2018/0196 (COD), in particular Recital 5 on respect for horizontal EU principles, article 67 on selection of operations by the managing authority based on non-discriminatory procedures and taking account of the Charter of Fundamental Rights of the EU, and Annex III on horizontal enabling conditions.

In the same vein, the United Nations High Commissioner for Human Rights recently published a guidance note to data collection and disaggregation for monitoring progress in achieving the goals of the 2030 Agenda for Sustainable Development.<sup>5</sup> It stresses that data collection activities should always be conducted in accordance with the overriding human rights principle of *doing no harm* and not create or reinforce existing discrimination, bias or stereotypes; data on personal characteristics should be kept safe and used only for the benefit of the groups they describe and society as a whole. “*Do no harm* also means that nothing in the guidance note should be interpreted as an invitation, encouragement or endorsement of any initiative or practice that seeks to discriminate against population groups and expose them to risks of serious human rights violations (or which has this effect)”.<sup>6</sup>

### What are equality data?

The [European Handbook on Equality Data – 2016 Revision](#) (European Commission, 2016, p. 15) defines **data** as any piece of information, whether in numerical or in some other form. The function of data is that they reveal something about some aspect of reality and can therefore be used for analysis, reasoning or decision-making.

**Equality data** is defined as 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.

There are different types of data sources, based on different data collection mechanisms, that can be used to compile equality statistics. Sources of equality data include population censuses, administrative registers, household and individual surveys, victimisation surveys, attitudinal surveys (self-report surveys), complaints data, discrimination testing, diversity monitoring by employers and service providers, as well as qualitative research strategies such as case studies, in-depth and expert interviews.

The collection and use of equality data are indispensable to Member States for a number of different purposes, including:<sup>7</sup>

- enabling proper assessment of the implementation of the relevant EU equality legislation and other relevant national legal and policy frameworks;
- monitoring trends in equality across different areas of life, such as employment rate, percentage of persons in tertiary education, or percentage of early school leavers, to enhance evidence-based policy making;
- monitoring the compliance of Member States with relevant international human rights standards, such as the UN Convention on the Rights of Persons with Disabilities and the UN Convention on the Elimination of All Forms of Racial Discrimination;

<sup>5</sup> Office of the UN High Commissioner for Human Rights (2018), [A Human Rights-based Approach to Data - Leaving No One Behind in the 2030 Agenda for Sustainable Development](#), Guidance note to data collection and disaggregation.

<sup>6</sup> Office of the UN High Commissioner for Human Rights (2018), [A Human Rights-based Approach to Data - Leaving No One Behind in the 2030 Agenda for Sustainable Development](#), Guidance note to data collection and disaggregation, p. 11-12.

<sup>7</sup> For a comprehensive list and broader discussion of the different needs, see European Commission (2016), [European Handbook on Equality Data – 2016 Revision](#), p. 19-22.



- making visible the hidden figure of discrimination, i.e. unreported discrimination, for example, by conducting victimisation surveys or conducting discrimination testing experiments;
- providing reliable evidence in administrative or judicial cases regarding discrimination through data that point to direct or indirect discrimination;
- populating indicators to track progress in the implementation of strategies relevant to fostering equality and non-discrimination to which Member States are committed, such as the European Pillar of Social Rights, the Europe 2020 Strategy or the 2030 Agenda for Sustainable Development;
- supporting advocacy and awareness raising in the field of equality and non-discrimination;
- highlighting evidence gaps where further research is needed.

Acknowledging the importance of equality data for these purposes, the High Level Group on non-discrimination, equality and diversity set up the Subgroup on Equality Data in February 2018,<sup>8</sup> with a view to support Member States in their efforts to improve the collection and use of equality data.<sup>9</sup> The FRA was invited to facilitate the work of the subgroup, in line with the agency's mandate to develop methods and standards to improve the comparability, objectivity and reliability of data at the European level.<sup>10</sup>

### Identifying gaps in the collection and use of equality data

The work of the subgroup revealed much variation in how Member States collect and use equality data. However, the subgroup also agreed that there are common gaps and challenges in current practice, the most important of which are summarised below:

#### 1. *Lack of a coordinated approach for equality data collection and use*

Member States tend not to have a coherent and systematic approach to equality data collection. This often leads to equality data collection being unevenly split between different sources and data collection bodies that are generally not coordinated or connected with each other. As a result, data collected by Member States may not be used to best effect in monitoring the outcomes of legislation, policies and measures they implement to foster equality and promote non-discrimination.

#### 2. *Lack of sufficient resources and of awareness of the importance of collecting equality data*

---

<sup>8</sup> The Subgroup on Equality Data is composed of representatives from the European Commission, FRA, Eurostat, as well as representatives from national administrations, national statistical institutions or equality bodies. National representatives were nominated to take part in the subgroup by Belgium, Bulgaria, Croatia, Hungary, Estonia, Finland, Germany, Greece, Ireland, Italy, Lithuania, Netherlands, Romania, Spain, the United Kingdom, as well as Norway. The participation of Eurostat is of crucial importance to create synergies between the work of this subgroup and the developments at the level of the European Statistical System (ESS). The ESS is the partnership between the statistical authority (Eurostat), and the national statistical institutes (NSIs) and other national authorities responsible for the development, production and dissemination of European statistics in each Member State.

<sup>9</sup> The High Level Group on non-discrimination, equality and diversity (HLG) mandated the Subgroup on Equality Data to prepare a set of non-binding guidelines on improving the collection and use of equality data; develop a diagnostic tool/checklist that EU Member States can use to assess the availability of equality data collected at national level, and compile practices implemented at national level related to the proposed set of non-binding guidelines. The HLG works on the following grounds of discrimination: sexual orientation and gender identity, racial or ethnic origin, age, religion or belief and disability, with gender being taken into consideration in its intersection with these grounds.

<sup>10</sup> [Council Regulation \(EC\) No. 168/2007 of 15 February 2007 establishing a European Union Agency for Fundamental Rights.](#)

Member States tend to underestimate the need to ground their anti-discrimination policies on robust and reliable equality data, through which the effectiveness of the relevant legal and policy framework could be assessed. As a result, insufficient resources are invested in improving the collection of such data. In addition, stakeholders who would be willing and competent to collect equality data, such as equality bodies, research institutes and civil society organisations, often do not have the necessary resources to do so, or to do so with the continuity required to allow for consistent monitoring over time.

### *3. Imbalance in equality data collection on different grounds of discrimination*

Where collected, equality data tend to concern some grounds of discrimination more than others. This can lead to imbalances in equality data collection and, therefore, to large knowledge gaps. On the whole, data collection which covers grounds of discrimination such as gender<sup>11</sup> and age is the most developed in Member States. Data collection on disability is gradually improving, partially due to obligations Member States have in that regard under Article 31 of the United Nations Convention on the Rights of Persons with Disabilities (CRPD). Data collection on the discrimination grounds of racial or ethnic origin, religion or belief, sexual orientation and gender identity tends to be the least developed in Member States.<sup>12</sup> Furthermore, such imbalances in equality data collection can lead to a lack of evidence on the extent to which people experience multiple and intersectional discrimination.

### *4. Imbalance in equality data collection across different areas of life*

The Racial Equality Directive prohibits ethnic discrimination in a number of key areas of life, such as employment, social protection, social advantages, education and access to goods and services, including housing, while the Employment Equality Directive prohibits discrimination on the grounds of religion or belief, disability, age or sexual orientation in the fields of employment and occupation. There remains, however, an imbalance in the extent to which Member States collect equality data in those areas covered by the directives, with employment typically being the area where most data is collected. On the whole, this contributes to limited availability of comprehensive data at national level that cover all areas of life where people might face discrimination.

### *5. Insufficient comparability of equality data across different data sources*

Because different data sources often rely on diverging definitions and use various population categorisations, equality data are often not comparable, both within and across Member States. For example, some sources collect data on issues relevant to disability by applying a medical model. Others follow the human rights approach enshrined in the United Nations Convention on the Rights of Persons with Disabilities, which focuses on the interaction between impairments and existing social and environmental barriers. In the absence of a mechanism to coordinate decisions on how to conceptualise and measure different grounds of discrimination and on how to categorise potentially affected persons or groups, data and outcomes will therefore most certainly suffer from incomparability.

### *6. Incomplete identification of population groups at risk of discrimination due to over-reliance on proxies*

---

<sup>11</sup> For sex/gender related statistics see the European Institute for [Gender Equality \(EIGE\)](#), particularly their [Gender Equality Index \(GEI\)](#) that includes the domain of intersecting inequalities. The GEI is a composite indicator that measures the concept of gender equality and, based on the EU policy framework, assists in monitoring progress of gender equality across the EU over time.

<sup>12</sup> See European Commission (2017), [Equality data indicators: Methodological approach. Overview per EU Member State](#).

Self-identification is a key principle of a human rights-based approach to data collection as recommended by the Office of the United Nation High Commissioner for Human Rights.<sup>13</sup> The subgroup also recognises the value of proxy information, which can help provide a consistent basis for monitoring of change over time. Wherever possible, the subgroup suggests that both self-identification and proxy data are collected and analysed to provide a more holistic picture of protected groups. Persons or groups who experience unequal treatment are often captured by means of proxies that might not cover the full range of discrimination experiences. This holds particularly true for personal characteristics such as racial or ethnic origin. Only few data sources enable assignment of personal identities based on the human rights-based principle of self-identification. Instead, it is more common to use imputations or proxies such as citizenship, country of birth, language or place of residence to determine identities related to ethnicity. While proxies may not be sufficient for comprehensively identifying all potential victims of racist or ethnic discrimination, they can provide a consistent picture to enable monitoring of change over time. For example, the variable 'country of birth' cannot reveal how descendants of immigrants who are born in the respective country experience discrimination on the grounds of ethnic origin or skin colour, or other visible characteristics, such as religious clothing. It, however, allows for an analysis of disparities between different population groups.

#### *7. Insufficient consultation with relevant stakeholders*

Only in few Member States attempts are made to involve or consult with persons, groups or communities at risk of discrimination when establishing definitions and categories for data collection. Yet, participation is not only a good practice from the perspective of a human rights-based approach to data collection, but results in more robust data, greater trust in the data collection process and more willingness to engage in data collection exercises. A lack of participation of relevant stakeholders in data collection exercises may adversely impact the response rate and hence the validity, reliability and representativeness of the data collected. For example, if questions on discrimination experienced on the grounds of sexual orientation or gender identity are not consulted with potentially affected population groups, this can lead to a lack of trust in the purpose of data collection and to questioning wordings that do not reflect the self-understanding of the persons under study, thus deterring responses.

#### *8. Inaccurate interpretation of the data protection frameworks*

Sometimes data protection requirements are understood as prohibiting collection of personal data such as a person's ethnic origin, religion or sexual orientation. However, as the explanation below shows the EU General Data Protection Regulation (GDPR),<sup>14</sup> which is directly applicable in all EU Member States since 25 May 2018, establishes conditions under which collection and processing of such data is allowed.<sup>15</sup>

---

<sup>13</sup> Office of the UN High Commissioner for Human Rights (2018), [A Human Rights-based Approach to Data - Leaving No One Behind in the 2030 Agenda for Sustainable Development](#), Guidance note to data collection and disaggregation.

<sup>14</sup> [Regulation \(EU\) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC.](#)

<sup>15</sup> In case of uncertainty, data collectors could seek the advice of their national data protection authorities in order to confirm that their procedures do comply with the GDPR. The European Commission has published general guidance under the following link: [https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules\\_en](https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en)

*Processing of special categories of personal data under the General Data Protection Regulation*

Article 9 of the General Data Protection Regulation (EU) 2016/679 (GDPR) prohibits the processing of personal data revealing characteristics such as racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, as well as the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health, or data concerning a natural person's sex life or sexual orientation. However, the GDPR clarifies that this prohibition does not apply, among others, when:

- the data subject has given explicit consent to the processing of those personal data for one or more specified purposes, except where Union or Member State law provide that the prohibition may not be lifted by the data subject (Article 9.2 (a));
- processing is necessary for reasons of substantial public interest, statistical purposes, scientific or historical research purposes, or for archiving purposes in the public interest, on the basis of Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject (Article 9.2 (g) and (j)).

Recital 26 of the GDPR further states the principles of data protection should not apply to anonymous information, that is, information which does not relate to an identified or identifiable natural person, or which relates to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. Recital 26 further states that the GDPR does not concern the processing of such anonymous information, including for statistical or research purposes.

Finally, the way personal data are processed has to comply with the GDPR and in particular with the general principles of data protection as set out in Article 5: lawfulness, fairness, transparency, purpose limitation, data minimisation, accuracy, storage limitation, integrity, confidentiality and accountability.

### How were these guidelines developed?

These guidelines were developed by the Subgroup on Equality Data. They are intended to provide practical guidance to Member States on how to gradually improve the collection and use of equality data, with a view to assist them in monitoring the implementation of relevant legislation, policies and measures they devise to that effect.

The guidelines were finalised on the basis of systematic exchanges of practices and discussions held in the framework of both the High Level Group and the subgroup, including on how to address the gaps and shortcomings identified in current national practices. The guidelines further build on recommendations from the updated *European Handbook on Equality Data*.<sup>16</sup> These include utilising multiple sources of equality data, promoting coherent definitions and statistical categories, improving the validity and reliability of existing data collection exercises, and ensuring that sufficient funding is available for equality data collection.

They are divided into two groups:

- A. **Institutional and structural guidelines** provide guidance to Member States in setting up structures that enable a systematic, long-term and cooperative approach to collection and use of equality data; they mainly target policy-makers and public administrations that have both a mandate and the potential to trigger change at institutional level.

---

<sup>16</sup> European Commission (2016), [European Handbook on Equality Data – 2016 Revision](#).

- B. **Operational guidelines** provide guidance to Member States in ensuring comprehensiveness, timeliness, validity, reliability and representativeness of equality data and in enhancing their comparability; they target in particular institutions/actors engaged in the collection, analysis and dissemination of equality data, such as national statistical offices, ministries or other national or local administrative bodies, research institutes, equality bodies and civil society organisations.

All institutional/structural and operational guidelines are interconnected, they build upon one another, and in many ways complement and reinforce each other. The guidelines should be read and applied as a whole. Yet, as the discussions within the High Level Group and the subgroup have shown, the situation concerning the scope and quality of collected equality data varies across Member States. Therefore, policy makers and other stakeholders may also choose or prioritise those actions set out in the guidelines which best fit the needs identified at that time in the relevant Member State. It is possible to start by applying selected elements of the guidelines and then, with the aim of building up a comprehensive system of high quality equality data, expand by implementing further elements. In this sense, the implementation of these guidelines is seen as an ongoing process and the extent to which they are applied a matter of degree.

EU Member States are encouraged to use these guidelines to inform the development of policies and the evaluation of their outcomes from an equality perspective.

## Section II: Guidelines on improving the collection and use of equality data

### A. Institutional and structural guidelines

#### 1. Map existing sources of equality data and identify data gaps

A national mapping exercise should be conducted to identify existing sources of equality data in Member States, avoid duplication of data collection efforts and establish a baseline for a more systematic approach towards equality data collection.<sup>17</sup>

Such an initiative could be facilitated or coordinated by a relevant ministry or governmental department. Alternatively, a specific mandate and dedicated resources could be provided to the national statistical office, an equality body or a research institute to carry out such a mapping.

In the first instance, the mapping should help Member States to:

- identify the different sources and providers of equality data, taking also into account data sources that are not specifically designed to gather equality data, but include variables that can be used to show and analyse existing inequalities;
- identify for which grounds of discrimination data are collected;
- take account of intersections between grounds of discrimination;
- identify for which areas of life data are collected (for example, (access to) employment, (access to) housing, education, health, access to justice, access to goods and services, etc.);

---

<sup>17</sup> For an overview of Member States' practices in this regard, see European Commission (2017), [Legal framework and practice in the EU Member States](#), Analysis and comparative review of equality data collection practices in the European Union.

- establish with which frequency and based on which definitions and categorisations data are collected;
- collect, where available, information on sample size, representativeness (including whether some groups are systematically under-represented or excluded), identification of target groups via self-identification or by other means, mode of data collection, geographical coverage, frequency of updates and existing time series (if any);
- include reference or links to existing outputs of the data sources, such as reports or latest publications, data visualisations, web pages and other relevant meta data.

In the second instance, the mapping should serve to identify the needs for equality data in Member States. The performance of such a needs assessment should be based on relevant EU, international and national normative standards to which Member States are bound such as the Charter of Fundamental Rights of the EU, the EU equality legislation, international human rights law and national legislation. The needs assessment should also respond to EU-level policy coordination processes such as the European Semester, as well as to relevant political commitments, including the European Pillar of Social Rights or the UN Agenda for Sustainable Development. As a whole, the needs assessment should serve to identify what type of equality data Member States require for evidence-based policy making.

Possible outcomes of the mapping would consist of:

- a comprehensive and systematic description of all available sources of equality data in a Member State;
- identification of all providers of equality data;
- an equality data needs assessment;
- suggestions or recommendations for possible ways or instruments to address the identified gaps or shortcomings in equality data collection.

Taken together, these outcomes would allow for the identification of existing data gaps, and bring to light areas where improvements are needed.

The mapping of equality data could be designed so that it is carried out periodically, dependent on the regular updates of the mapped data sources, for example, every five years, taking into consideration new potential data sources and providers.

## 2. Foster inter-institutional cooperation in the collection and use of equality data

An inter-institutional working group should be established to ensure coordination and monitor progress in addressing the gaps and shortcomings identified through the mapping exercise. The activities of the working group could be laid down in a multiannual action plan or roadmap, with an appropriate budget and clearly defined priorities and allocation of tasks. This would contribute to attaining timely and fit-for-purpose data, relevant to the needs identified, and increase efficiency.

The working group should be composed of relevant public institutions with competencies (and, where relevant, a mandate) in the field of equality data collection. This could include ministries, national statistical offices, equality bodies, national human rights institutions, research institutions and the scientific community, as well as other relevant actors and data providers such as representatives of relevant local authorities, the judiciary, the police, etc. One organisation should be mandated to coordinate the working group and follow up on its decisions. However, improving the collection and use of equality data is a collective endeavour and should be based on joint ownership and shared responsibilities. At its best, inter-institutional cooperation is a mutual learning process that gradually builds up capacity, expertise and trust among different stakeholders, and fosters an increasing awareness of the need for equality data.

The work of the inter-institutional working group should be complemented through regular consultations with data collectors and data users who are not represented in the working group. These include representatives from other relevant national and local authorities, statisticians, survey research experts and academics, as well as community and civil society organisations. The involvement of these organisations would ensure that the perspective of those most at risk of experiencing unequal treatment and discrimination is duly taken into consideration.<sup>18</sup>

Such regular consultations would enhance transparency, help build up trust in equality data collection initiatives and ensure that relevant expertise feeds into the activities of the working group. Overall, regular consultations with relevant stakeholders would ensure that the priorities set out in the action plan or roadmap respond to the most important needs for equality data in Member States.

### 3. Set up a data hub on equality and non-discrimination

A national data hub on equality and non-discrimination should be set up to collate and display the data and related information identified in the mapping exercise and, thereby, enhance the accessibility and use of equality data by different stakeholders. The inter-institutional working group could be tasked with setting up the data hub, for example, in the form of a dedicated webpage and coordinating its further development. The hub could be hosted by the institution coordinating the inter-institutional working group or another suitable institution with the capacity to guarantee the continuity and periodic update of the hub.

In the first instance, the data hub would provide an overview of existing sources of equality data, referencing where the information can be accessed. The hub would further include available publications and the presentation of existing indicators that relate to different population groups, grounds of discrimination and areas of life, as well as technical information on the different data, including:

- source;
- responsible institution;
- type of data (administrative data, complaints data, survey data, etc.);
- population group;
- definitions and systems of categorisation applied;
- ground(s) of discrimination/protected characteristics;
- areas of life/domains covered;
- mode of data collection (face-to-face, online, telephone, mixed mode etc.);
- sampling frame and sampling method (where applicable);
- periodicity of data collection;
- geographical coverage and granularity;
- representativeness.

To facilitate the proper understanding and interpretation of the equality data gathered in the data hub and the establishment of meaningful comparisons, the data hub should include, where possible,

---

<sup>18</sup> As stated in the guidance note to data collection and disaggregation prepared by the Office of the UN High Commissioner for Human Rights (2018), [A Human Rights-based Approach to Data - Leaving No One Behind in the 2030 Agenda for Sustainable Development](#), participation is central to a human rights-based approach. Therefore, “all data collection exercises should include means for free, active and meaningful participation of relevant stakeholders, in particular the most marginalized population groups. Participation should be considered in relation to the entire data collection process: from strategic planning through identification of data needs; selecting and testing an appropriate collection methodology, data collection (for instance, hiring interviewers from particular communities to improve response rates); and to data storage, dissemination, analysis and interpretation”.

information notes for users, synthesizing the basic facts about the data on display.<sup>19</sup> These information notes should also make clear when comparisons between data cannot be drawn due to differences in definitions or other factors.

In the second instance, the data hub should be further developed and refined, following, for example, the implementation of measures foreseen under the action plan to address gaps and shortcomings in equality data collection. Furthermore, (new) rights-based indicators relevant to monitoring equality in outcomes and non-discrimination could be added to the data hub. Such indicators could be jointly developed to this end at the initiative of the inter-institutional working group.

#### 4. Build institutional capacity to collect robust and reliable equality data

The capacity of relevant stakeholders to collect reliable and robust equality data should be enhanced to effectively address the areas for improvement identified by the inter-institutional working group. This would also require providing adequate funding to those bodies and organisations that collect equality data, which includes national statistical offices and research institutes carrying out large scale surveys, as well as equality bodies, other public institutions and civil society organisations that carry out targeted surveys on discrimination experiences, conduct discrimination testing experiments or collect complaints data.

In addition, it is also important to ensure that the staff from different stakeholders regularly engaging in equality data collection have the relevant skills, expertise and awareness of best practice, including the use of harmonised definitions, to adequately design the collection of equality data and to comply with safeguards set out in the relevant legislation and UN principles of human rights-based approaches to data collection and analysis. This is, for example, the case of employers and service providers who collect data to assess their compliance with specific equality duties set out in legislation, or of civil society organisations working with groups at risk of discrimination. Therefore, tailored training should be developed, drawing on the expertise of different stakeholders and covering the range of activities relevant to equality data collection, i.e. design, collection, processing, analysis, use and dissemination as well as the provisions of the data protection legislation governing all steps of the life cycle of personal data. More generally, inter-institutional cooperation, as foreseen under Guideline 2, is enhanced when representatives of the different institutions engaging in equality data collection have the necessary statistical literacy and knowledge about equality issues and share a common language.

Overall, an institutional culture acknowledging the need for robust and reliable equality data to ensure evidence-based policy making should be promoted across public institutions and all relevant stakeholders.

#### 5. Facilitate effective use of equality data

To facilitate its effective use, equality data should be disseminated as quickly as possible after collection and be presented in an accessible language and format, taking into consideration that it should be understandable to the greater public.<sup>20</sup>

---

<sup>19</sup> This would be in line with Principle 15, on ‘Accessibility and Clarity’, of the [European Statistics Code of Practice-revised edition 2017](#), adopted on 16 November 2017 by the European Statistical System Committee.


<sup>20</sup> In line with the recommendations under the ‘Transparency’ section of the Guidance note to data collection and disaggregation, Office of the United Nations High Commissioner for Human Rights (2018), [A Human Rights-based Approach to Data - Leaving No One Behind in the 2030 Agenda for Sustainable Development](#), p. 15.



The inter-institutional working group could promote an enhanced use of equality data by contributing to the dissemination of the data collected and by providing explanations and interpretations where appropriate. It could also issue policy briefs that present outcomes from recent surveys or research and provide easy-to-use equality data for policy makers, experts and media related to the relevant policy. These policy briefs could also be used for awareness raising and advocacy.

The information displayed in the data hub foreseen under Guideline 3 could be used to inform public debate on equality and non-discrimination. An enhanced use of equality data to discuss and to justify legislative and policy measures would also, as a by-product, strengthen public support for investing in the collection of robust and reliable equality data by showing its utility.

Ultimately, equality data should feed into the policy cycle to enhance evidence-based policy making in all relevant policy areas, across different discrimination grounds and areas of life. Here, the term 'policy cycle' is meant to include public debate, the initial design of measures and policies and the evaluation of their impact. Effectively incorporating equality data in the policy cycle will depend on three inter-related factors: (1) a will to develop evidence-based policy making; (2) the timely availability of equality data for policy makers; and (3) the accessibility of equality data for the media and the general public.



## B. Operational guidelines

### 6. Ensure comprehensiveness of equality data

Data collectors and data users need multiple and complementary sources of equality data to enable evidence-based policy making, as well as to get a full picture of the reality of equality and non-discrimination on the ground. In addition, an adequate impact assessment of policies in the field of equality and non-discrimination requires a regular data collection on all grounds of discrimination and for all areas of life covered by the relevant legislation. To ensure the comprehensiveness of equality data, Member States should:

- use a broad set of equality data sources; this would include combining and linking different sources of data, such as large scale surveys, attitudinal surveys, victimisation surveys, discrimination testing, administrative data, complaints data (and their outcomes), robust and reliable data from civil society organisations, data collected by employers and service providers and other quantitative and qualitative research;
- ensure that the full range of key areas of life where discrimination can occur is captured by equality data sources; these areas of life include employment, education, access to goods and services, housing, health, poverty/social exclusion, political and civic participation, etc.;
- use existing data to their full potential, including data collected for purposes other than equality and non-discrimination policies;
- establish or further develop an efficient system for collating and regularly publishing statistical data on reported cases of discrimination and their administrative or judicial outcomes;<sup>21</sup>
- mainstream multiple and intersectional discrimination in data collection and analysis; for example, data collected in the area of disability should ideally include information on gender, ethnic or racial origin, age and other potential characteristics that might lead to a higher risk of discrimination based on multiple factors (this type of analysis generally requires larger sample sizes for robust analysis, and a range of different data sources could be considered including large surveys, administrative and linked data);
- foster where possible data linkage between different data sources;
- increase the amount of regional equality data, geo-referenced to the smallest possible grid.

### 7. Mainstream equality data into EU and national surveys

Large-scale surveys are a powerful tool, which enables the collection of a comprehensive set of equality data. When developing such surveys or when collecting data within the framework of the European Statistical System (ESS) through the Labour Force Survey (LFS), the EU Statistics on Living Conditions (EU-SILC), or through other EU-wide surveys such as Eurobarometer and the FRA surveys, Member States should, where relevant, systematically incorporate information referring to equality and non-discrimination through, among others:

---

<sup>21</sup> This could be linked to the implementation of the Recommendation on standards for equality bodies adopted by the European Commission in June 2018 to set out measures that Member States may apply to improve the equality bodies' independence and effectiveness. See European Commission (2018), [Commission Recommendation of 22 June 2018 on standards for equality bodies](#), C(2018) 3850 final. In particular, its recommended measure No. 9, on independent reports, states that "for the purpose of obtaining independent reports of high quality, Member States should enable equality bodies to conduct independent research. This could include collecting data in particular on the number of complaints per discrimination ground; the duration of administrative proceedings from submission of the complaint to the closure of the case; the outcome of administrative proceedings; and the number, duration and outcome of judicial cases in which the equality bodies are involved".

- introducing individual variables and survey questions that specifically relate to measuring inequalities and perceived discrimination on a number of grounds;
- developing and running additional thematic survey modules on a regular basis to cover the range of issues of relevance to equality and non-discrimination (for example, rights awareness, rates of reporting experiences of discrimination, experiences of victimisation, etc.);
- disaggregating the collected data as much as, and where possible, on the basis of age, gender, ethnic or racial origin, religion, migration background, sexual orientation and disability;
- increasing the number of people who otherwise wouldn't be adequately represented in the survey samples due to small numbers in general population (booster samples);
- undertaking efforts to develop, test and apply precise and sensitive categories for identification of persons with a high risk of discrimination (for example, variables for racial or ethnic origin, sexual orientation or gender identity);
- increasing the number of proxy variables/categories in existing data sources where the human rights-based principle of self-identification cannot not be applied;
- increasing the number of variables measuring perceived discrimination on different grounds and in different areas of life;
- supporting the development and implementation of innovative survey methodologies that can capture highly dispersed or hard-to-reach population groups or persons at risk of discrimination;
- identifying geographical areas where inequalities are more often observed and where data should be collected on regular basis to allow monitoring of trends over time.

## 8. Ensure regular and timely equality data collection

Equality data should be collected, analysed and reported on on a regular basis. The data therefore need to be made available for the intended use within a reasonable time frame and with a regular frequency, if they are to meet information needs and feed into the policy cycle.

To ensure regular and timely equality data collection, Members States should:

- systematically and on a regular basis, collect data on all grounds of discrimination and for all areas of life/domains covered by equality and non-discrimination legislation (for example, every three to five years);
- systematically and on a regular basis, update the data hub on equality and non-discrimination;
- promote the inclusion of equality data in longitudinal surveys, i.e panels, to monitor impact on individuals;
- systematically identify and monitor trends when collecting equality data;
- use information from existing equality data sources available at regular intervals to avoid costly duplications, particularly when no targeted data collection exercises are taking place;
- develop a methodological guidance for the systematic collection of equality data over time, to be used by the different data collectors.

## 9. Enhance validity and reliability of equality data

Validity and reliability need to be built in the design of any equality data collection exercise. Validity<sup>22</sup> relates to the correct measurement of a variable of interest, which is itself observable in

---

<sup>22</sup> An indicator, a test, a survey question or a system of categorisations used to classify person's characteristics in administrative or in other data sources is valid: (1) if it correctly measures what it is supposed to measure, and (2) if it renders phenomena or characteristics that are directly observable or have at least been perceived to exist in the 'real world'. Validity is for example linked to the extent to which

the ‘real world’. Reliability<sup>23</sup> relates to the stability or consistency of the measurement applied to the variable.

To enhance the validity and reliability of equality data, Member States should:

- apply the human rights-based principles of participation, self-identification, transparency, privacy and accountability in the design, collection and use of equality data to retain or boost the trust of relevant stakeholders;<sup>24</sup>
- apply a participatory approach to the data collection design and process, for example, when developing definitions, systems of classification or categorisation, or when making decisions on the use of proxies; such an approach could include hiring interviewers or mediators drawn from particular population groups to increase response rates;
- carry out community consultations with representatives of the targeted population groups with a view to reaching a consensus on the definitions and categories to be used in equality data collection;
- encourage cooperation with the targeted population groups and stakeholders by informing them that data are collected and used to support programmes designed to promote equal opportunities or fight discrimination, with a view to ensuring a higher response rate;
- improve response rates by adapting and redrafting (ambiguous) questions/categories where necessary; this could be based on insights from the community consultations;
- where possible, enable self-identification by individuals who are the subject of data collection on the basis of characteristics such as racial and ethnic origin, religion, disability, sexual orientation or gender identity; such categorisation systems should allow for the selection of more than one identification category, as well as allow respondents not to declare or declare ‘none’;
- refine and/or increase the number of proxies applied to ensure comprehensive coverage of different population groups for which direct (self-) identification is not possible;
- thoroughly test any categories, variables and questions developed for data collection;
- ensure the use of consistent procedures and definitions over time;
- apply the ESS statistical code of practice<sup>25</sup> and the quality reports produced by the data producers where relevant;
- where relevant, apply the following general quality criteria to available indicators and those to be developed in the future: consistency, clarity, timeliness, punctuality, accuracy, continuity, objectivity, relevance, comparability and transparency;
- develop a protocol summarising explanations for continuing differences in definitions and categorisations across different data sources in cases where no development of new or

---

respondents misinterpret survey questions or response categories and to the extent that they deliberately do not reveal the truth (often the case when the personal information asked is perceived as sensitive). A lack of validity can also be observed if, for example, the extent of discrimination is only measured through the number of incidents reported to the competent bodies as there might be a high number of non-reporting. There are no perfectly valid measures, but some measures are more valid than others.

<sup>23</sup> Reliability implies stability or consistency of the measurement/test applied. A measure of discrimination for example is reliable to the extent to which the measuring procedure yields the same results in repeated trials. For example, if we measure the number of reported incidents per year, but the number of competent (equality) bodies changes over time, this will have a negative effect on the reliability of this measure and will also impact on its comparability over time. No measure is absolutely reliable; reliability is therefore always a matter of degree. There is a direct relation between validity and reliability in the way that tests that are valid are also reliable (repeatable). Tests that are reliable, however, are not always valid.

<sup>24</sup> See Office of the United Nations High Commissioner for Human Rights (2018), [A Human Rights-based Approach to Data - Leaving No One Behind in the 2030 Agenda for Sustainable Development](#), Guidance note to data collection and disaggregation, 2018.

<sup>25</sup> European Statistical System Committee (2017), [European Statistics Code of Practice - revised edition 2017](#).

adaptation of existing categories is possible (for example, when the main objective of existing data collection is the monitoring of trends).

#### 10. Ensure representativeness of equality data

Representativeness shows the extent to which characteristics of a sample, such as composition along gender, age, ethnic and racial origin, etc., correspond to the characteristics of the population from which it has been taken, in accordance with established rules. The term 'representativeness' is often used to indicate that a sample mirrors a population group, reflecting all its essential properties in a correct way.

Representativeness can be negatively affected if the sample size is too small, when it does not include a sufficient number of persons belonging to a targeted population group, or if specific subgroups of the population are systematically excluded from data collection in the first place (because of language difficulties, impairments etc.). These challenges apply for many different surveys and are especially problematic when it comes to the collection of equality data, as groups with a high risk of discrimination are disregarded and cannot be considered while analysing the data. Representativeness can also be affected if the sample is large enough, but biased, for example, when those persons belonging to a target population who experience more discrimination are less or more inclined to take part in a survey. To address such difficulties, Member States should apply an inclusive data collection approach, undertaking efforts to:

- reach out to under-represented groups and achieve robust samples and effective sample sizes to the extent possible;
- conduct targeted surveys;
- include members of minority groups and hard-to-reach groups in existing surveys, such as LFS and EU-SILC, through boosting samples where relevant;
- include persons living in institutional settings such as care homes, prisons, asylum reception centres, nomadic people, homeless people and people living in temporary accommodation in equality data collection, where possible;
- foster qualitative and mixed-methods research especially when target groups are too small or highly dispersed;
- ensure effective insights into a full range of protected characteristics or groups at risk of discrimination who might be left out from existing sampling frames through conducting qualitative research; qualitative methods can be used throughout the research process to complement statistical insights and to enhance representativeness and validity.

#### 11. Improve comparability of equality data

Comparability is defined as the extent to which differences between findings, statistics or outcomes from different equality data sources, countries, regions, cultures, life domains or time periods can be attributable to differences in target population's true values. Comparability is strongly affected when concepts, definitions or categories of study design vary across data sources, points in time or countries. Improving comparability implies that error due to sample or questionnaire design, mode of data collection, translation etc. is minimised. To help determine comparability, the decisions involving the design, definitions and categories to be applied should be well-documented.

To improve comparability of equality data, Member States should:

- identify differences in definitions, systems of classification and categorisation across different data sources and across time; this could be done through the mapping exercise

(Guideline No. 1) and during the consultation processes with relevant stakeholders (Guidelines No. 2 and 9);

- to the extent possible, harmonise definitions, categorisations and the data collection design;
- develop and issue guidance for a harmonised administration of data collection;
- where comparability cannot be achieved (for example, when data are gathered for different purposes), develop and issue guidance for the interpretation of existing differences between data sources, languages or time points regarding definitions, design and systems of classification and categorisation; this guidance should make clear that while results might differ due to lack of comparability, they are still valuable and feed into the overall assessment of the state of equality and non-discrimination in a country;
- where relevant, make efforts to improve comparability of equality data on the EU level, namely of the comparability of data collected through the standard EU data collection systems (comparative-by-design);
- where relevant, and to the extent possible, improve comparability of equality data with regard to available OECD, UN and World Bank data in the area of equality and non-discrimination; in this regard, Member States could make efforts to utilise existing national/local equality data for comparative purposes by using rights-based indicators as the basis for comparison;
- promote exchange of information among countries relating to good practices to enhance the comparability and compatibility of equality data at national level.

## GETTING IN TOUCH WITH THE EU

### In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: [https://europa.eu/european-union/contact\\_en](https://europa.eu/european-union/contact_en)

### On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696, or
- by email via: [https://europa.eu/european-union/contact\\_en](https://europa.eu/european-union/contact_en)

## FINDING INFORMATION ABOUT THE EU

### Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: [https://europa.eu/european-union/index\\_en](https://europa.eu/european-union/index_en)

### EU publications

You can download or order free and priced EU publications from: <https://op.europa.eu/en/publications>. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see [https://europa.eu/european-union/contact\\_en](https://europa.eu/european-union/contact_en)).

### EU law and related documents

For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: <http://eur-lex.europa.eu>

### Open data from the EU

The EU Open Data Portal (<http://data.europa.eu/euodp/en>) provides access to datasets from the EU. Data can be downloaded and reused for free, for both commercial and non-commercial purposes.



■ Publications Office  
of the European Union