European
Commission

High Level Group on
Non-discrimination, Equality and Diversity

Subgroup on
equality data

Guidance note on
the collection and use of equality data based on racial or ethnic origin

Justice
and Consumers

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

Guidance note on
the collection and use of equality data based on racial or ethnic origin

The Commission would like to express its gratitude to FRA for its extensive input in the development of this guidance note, drawing on their wide-ranging knowledge and expertise

Manuscript completed in September 2021

PDF ISBN 978-92-76-40969-4 doi:10.2838/06180 DS-01-21-144-EN-N

Luxembourg: Publications Office of the European Union, 2021

CONTENTS

PREAMBLE

> *“The European Union rejects theories which attempt to determine the existence of separate human races. The use of the term ‘racial origin’ does not imply an acceptance of such theories”.*[1]

Race/racial origin and/or ethnic origin are **social constructs** and as such they are weak proxies for the genetic diversity of humankind[2]. While some individuals may self-identify as ‘white’ or ‘black’, racism and racial or ethnic discrimination are often shaped by how society categorises individuals in racialised terms.

Ideas about race/racial origin are often ascribed to or imposed on people, and individuals or groups can be racialised by others in ways that negatively affect their experiences and how they are treated. The social construct of race/racial origin is distinct from but may overlap with how people identify themselves, which can be much more varied and complex.[3]

The UN Principles and Recommendations for Population and Housing Censuses related to Ethnicity (para 4.184.) broadly define ethnicity as a concept “*based on a shared understanding of history and territorial origins (regional and national) of an ethnic group or community, as well as on cultural characteristics such as language or religion. Respondents’ understanding or views about ethnicity, awareness of their family background, the number of generations they have spent in a country, and the length of time since immigration are all possible factors affecting the reporting of ethnicity in a census. Ethnicity is multidimensional and is more a process than a static concept, and so ethnic classification should be treated with movable boundaries.*”

The EU Framework Decision on racism and xenophobia stipulates that ‘descent’ “*... should be understood as referring mainly to persons or groups of persons who descend from persons who could be identified by certain characteristics (such as race or colour), but not necessarily all of these characteristics still exist*”.

In line with the applicable EU legislation[4], this guidance note refers to ‘racial or ethnic origin’ with respect to its being:

→ a cause of discrimination.

---

[1] Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin (6). See as well EU Anti-Racism Action Plan 2020-2025, p.1.

[2] See: Yudell M. et al., (2016), ‘Taking race out of human genetics’, *Science* 351 (6273):564-565. And: Zack, N. (2018), ‘Social Construction and Racial Identities’ In: Zack, N., *Philosophy of Race. Palgrave Philosophy Today.* Palgrave Macmillan, Cham.

[3] Government of Ontario (2018), *Data Standards for the Identification and Monitoring of Systemic Racism*. Anti-Racism Data Standards – Order in Council 897/2018.

[4] Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin (‘Racial Equality Directive’) as well as Article 10 and Article 19 of the Treaty on the Functioning of the European Union (TFEU).

Aligned with practice that has been established in some countries and EU Member States – which use the category ‘racial or ethnic origin’ for statistical purposes, including to highlight discrimination and inequality – this guidance note further refers to ‘racial or ethnic origin’ as:

- → a generic statistical (analytical) category that allows for disaggregation of any data, to assess the state of equality in society,
- → an aspect of a person’s self-identification and ethnic attachment, that is, as a personal characteristic.

To make informed policy choices for countering discrimination and fostering equal treatment, legislators and policymakers need data on people’s social positioning and experiences of racism and discrimination based on racial or ethnic origin.

However, introducing categories such as ‘racial or ethnic origin’ in official statistics bears the risk of such categorisations being socially reproduced and used to incorrectly label people.[5] This can have negative consequences for members of certain social groups, stemming from the biased (potentially stereotypical) belief systems that such social categorisation can support[6].

To address this, the use of statistical (analytical) categories for any data collection or for the purpose of data disaggregation should always be led by the overriding human-rights based principle of *doing no harm,* as proposed by the UN High Commissioner for Human Rights (OHCHR) in its Human Rights-based Approach to Data.

Doing no harm means that **no data collection activity should create or reinforce existing discrimination, bias or stereotypes** and that the data collected should be used for the benefit of the groups they describe and society as a whole.

Accordingly, this guidance note recognises and reaffirms the OHCHR human-rights based approach to data and its set of principles related to participation, data disaggregation, self-identification, transparency, privacy and accountability.[7]

---

[5] Two points need to be acknowledged in this regard: (a) it is not possible to limit the use of language to informed communicators in a society, and (b) the conscious intentions of communicators are not the only factors shaping the social meaning of a concept or a category.

[6] Liberman, Z. et al. (2017), ‘The origins of social categorization’, *Trends in Cognitive Sciences*, Vol, 21, Issue 7, pp. 556-568.

[7] However, there should not be a blanket ban on publishing data that might present certain groups in a negative way, when it concerns practices that are against people’s dignity and rights. For example, publishing data about female genital mutilation (FGM) could be seen as reinforcing existing stereotypes of certain groups where this practice takes place, but it is also for the benefit of the society and the rights of affected women and girls. Data collectors and other responsible stakeholders need to act responsibly when presenting and interpreting such data.

BACKGROUND AND PURPOSE OF THIS GUIDANCE

Equality data are essential for assessing the situation of ethnic minorities and other racialised groups and so effectively tackling racism and structural inequalities.[8] Data makes the nature and extent of discrimination and inequality visible[9] and provides the substance for evidenced-based policy making. When collected regularly and systematically, equality statistics enable Member States to assess the proper application of anti-discrimination legislation, monitor compliance with human rights obligations, and track progress in achieving goals towards equality – as set by EU economic governance instruments such as the European Semester or by global agendas such as the UN 2030 Agenda for Sustainable Development.

To date, few countries operate comprehensive systems or have a coordinated approach to collecting and using equality data that would uncover inequalities based on racial or ethnic origin. If available, such data are often not up to date or lack comparability across time and geographical regions; frequently they are of limited scope as well as are not transparently available in the public domain.

Acknowledging this, in 2018 the EU High Level Group on Non-Discrimination, Equality and Diversity (HLG) set up a Subgroup on Equality Data (Subgroup) to support Member States in their efforts to improve the collection and use of equality data. It tasked the EU Agency for Fundamental Rights (FRA) 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 equality data at European level.[10]

To date, the Subgroup has developed:

→ a set of non-binding guidelines on collecting and using equality data,

→ a compendium of promising practices for equality data collection implemented at national level,

→ a diagnostic mapping tool to help Member States map existing sources of equality data and identify gaps.

In 2020, the Subgroup's mandate was extended to 2025, to support the European Commission's efforts to implement a consistent approach on equality data collection, as regards data disaggregated by racial or ethnic origin. To meet this goal, the Subgroup agreed to develop practical guidance to Member States on improving the collection of data disaggregated by racial or ethnic origin.

---

[8] The Council Recommendation on Roma equality, inclusion and participation defines systemic or structural discrimination "*as being evident in the inequalities that result from legislation, policy and practice, not by intent but resulting from a range of institutional factors in the elaboration, implementation and review of legislation, policy and practice*" (p.20).

[9] FRA (2021), *Equality in the EU 20 years on from the initial implementation of the Equality Directives*. Opinion. Luxembourg: Publications Office.

[10] Council Regulation (EC) No 168/2007 of 15 February 2007 establishing a European Union Agency for Fundamental Rights, OJ L 53, 22.2.2007.

Need for data for informed policy choices

> *“The goal should be for Member States, in full respect of their national contexts, to move towards the collection of data disaggregated on the basis of racial or ethnic origin, in order to capture both subjective experiences of discrimination and victimisation and structural aspects of racism and discrimination. This data should be comprehensive, reliable, regular and timely; mainstreamed into EU and national surveys; and both representative and comparable.”*
>
> *EU anti-racism action plan 2020 - 2025, p. 16.*

In 2020, the EU took decisive steps towards a more holistic approach to addressing racism, xenophobia and related intolerance and adopted its first ever EU anti-racism action plan 2020–25. The plan addresses both individual and structural (institutionalised) forms of racism and sets out a series of concrete measures for getting the “right data for informed policy choices”. It calls for a more significant step towards a new approach on equality data collection and encourages Member States to improve the collection of data disaggregated by racial or ethnic origin.

This is in line with the Durban Declaration and Programme of Action, which urges states “*to collect, compile, analyse, disseminate and publish reliable statistical data at the national and local levels and undertake all other related measures which are necessary to assess regularly the situation of individuals and groups of individuals who are victims of racism, racial discrimination, xenophobia and related intolerance*”.

The need for data on ethnicity has been also identified by the European Court of Auditors – in its 2016 special report EU policy initiatives and financial support for Roma integration – in which it called on the Commission to work with Member States to develop a common methodology and encourage them to collect statistical data on ethnicity.

The COVID-19 pandemic has shown a significant impact on equality all over the globe: those already most at risk of discrimination and inequality (such as older people, persons with disabilities, members of ethnic minorities, Roma and Travellers, homeless people or those living in overcrowded conditions) were and are at far greater risk of falling ill or dying from the virus.[11]

In most countries in the EU, however, officially available health statistics on COVID-19 could not be (fully) disaggregated by racial or ethnic origin. This had a detrimental effect on the effectiveness of protective measures to curb the spread of the virus as well as information and prevention campaigns targeting people and groups at increased risk of over-exposure.

---

[11] FRA has published regular reports about the Fundamental rights implications of COVID-19 in 2020 and 2021.

On the concept of 'racial or ethnic origin'

According to Farkas[12], "*racial, ethnic and national origins are mutable grounds, comprising nationality, colour, descent, minority religion, minority language, minority culture and traditions. Together with foreign and immigrant, these characteristics constitute the most common proxies of racial or ethnic origin – most clearly borne out in the context of data collection on inequalities, which is decidedly based on these proxies*". Moreover, these social categories are recognised as intrinsically interlinked although not clearly distinguished in law. Careful analysis is therefore needed to determine whether discrimination is based on racial or ethnic origin or on a constitutive element of either concept, such as minority religion.[13]

Several examples of case law concerning, for example, the Roma broaden the interpretation of ethnic discrimination by further considering social marginalisation and spatial segregation.[14] In addition, immigrants and descendants of immigrants are identified on the basis of categories that speak to their status as immigrants – such as citizenship(s), country of origin, country of origin of parents, year of immigration, and language spoken at home. For immigrants and descendants of immigrants, such categories are commonly used as proxies of ethnicity in countries where the collection of so-called "ethnic statistics" is prohibited by law.

When it comes to experience of discrimination, questions about how others (those who discriminate) perceive someone – based on perceived external attributions – become important. Attributions made by others may not necessarily relate to a person's self-identification. As Farkas notes, "*Discrimination based on racial and ethnic origin rarely emerges in response to an individual's self-identification as belonging to or being a member of a certain racial or ethnic group. Processes, such as attribution or ascription, othering, grouping and stereotyping create, institute or impose racial or ethnic origins. Racial or ethnic identity may not correspond to that perceived or assumed by a third party. Identity is fluid and may change over time. That, however, does not necessarily impact on how that person is perceived. A conception of racial and ethnic origin can also be imposed from within communities and subdue intersecting identities and diverging interests.*" [15]

In addition, one's identity encompasses multiple, intersecting characteristics that must be recognised; not just racial or ethnic origin, but also sex, age, sexual orientation, (dis)ability and other personal traits.[16]

For more information on the theoretical and legal framework of the definition of racial or ethnic origin, see the Commission's report on Analysis and comparative review of equality data collection practices in the field of ethnicity in the EU.

---

[12] Farkas, L. (2017), The meaning of racial or ethnic origin in EU law: between stereotypes and identities, Luxembourg, Publication Office of the European Union, 2017, p. 8.

[13] Farkas, L, (2017), *Analysis and comparative review of equality data collection practices in the European Union: Data collection in the field of ethnicity*, Luxembourg: Publications Office of the European Union, 2017, p. 9.

[14] D.H. and Others v. Czech Republic, application no 57325/00, Grand Chamber judgment of 13 November 2007 and Yordanova and Others v Bulgaria, Application no. 25446/06, judgment of 5 June 2012 (see Farkas, 2017, p.9f).

[15] Farkas, L. (2017), The meaning of racial or ethnic origin in EU law: between stereotypes and identities, Luxembourg, Publication Office of the European Union, 2017, p. 37.

[16] Farkas, L. (2017), The meaning of racial or ethnic origin in EU law: between stereotypes and identities, Luxembourg, Publication Office of the European Union, 2017, p. 37.

What are 'equality data'?

The European handbook on equality data, and the Guidelines on improving the collection and use of equality data define 'equality data' as any piece of information that is useful for the purposes of describing, analysing, reasoning about and decision-making on 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 societies.

Equality statistics can be complied 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 and service providers, and data used to train algorithms for artificial intelligence (AI) and machine learning.

Data disaggregated by certain personal characteristics – including age, sex, racial or ethnic origin, religion or belief, disability, sexual orientation, and gender identity – can be used for producing equality data, at an aggregated level for statistical purposes, provided that this is done in full compliance with legal provisions and the corresponding exceptions.[17]

Purpose and structure of the guidance note

Responding to the EU anti-racism action plan 2020-25 and calls by other groups to improve the availability of equality data disaggregated by racial or ethnic origin, the Subgroup on equality data agreed to develop this practical guidance note.

It aims to:

→ identify challenges to collecting data on racial or ethnic origin.

→ highlight promising paths towards a more standardised and consistent approach to this.

→ provide practical guidance on effective ways to produce comprehensive, reliable, comparable and regular data at national level.

→ help public sector institutions and other relevant stakeholders to identify disparities based on ethnic or racial origin and effectively combat racism and structural discrimination.

---

[17] FRA (2021), *Equality in the EU 20 years on from the initial implementation of the Equality Directives*. Opinion. Luxembourg: Publications Office, p.65.

This guidance builds on:

- the non-binding Guidelines on improving the collection and use of equality data,
- the recommendations from the European Handbook on Equality Data (2016 revision)
- the Commission report on Data collection in the field of ethnicity

In addition, the guidance draws on expert advice, existing anti-racism or anti-discrimination data standards and guides to data disaggregation, such as

- the UNECE Poverty Measurement - Guide to Data Disaggregation, and
- the UN Principles and Recommendations for Population and Housing Censuses related to ethnicity.

Moreover, the guidance considers the range of possible data sources (such as population censuses, administrative registers, complaints data, household and individual surveys) and their specific requirements.

The guidance also addresses questions that arise in relation to the application of the General Data Protection Regulation (GDPR) – underlining how data on racial or ethnic origin can be collected for aggregate statistical purposes in line with EU law.

The target audience for this guidance includes public or private bodies involved in collecting and using equality data on racial or ethnic origin such as:

- data producers in political and administrative bodies at national, regional and local level;
- national statistical institutes;
- private and public organisations working in the area of workplace and service delivery;
- the justice system and the police;
- equality bodies and national human rights institutions,
- academia.

The situation concerning the scope and quality of collected equality data on racial or ethnic origin varies across Member States. As shown in the Commission's report on Data collection in the field of ethnicity, only a few Member States collect data on racial or ethnic origin, while others deliberately refrain from doing so.

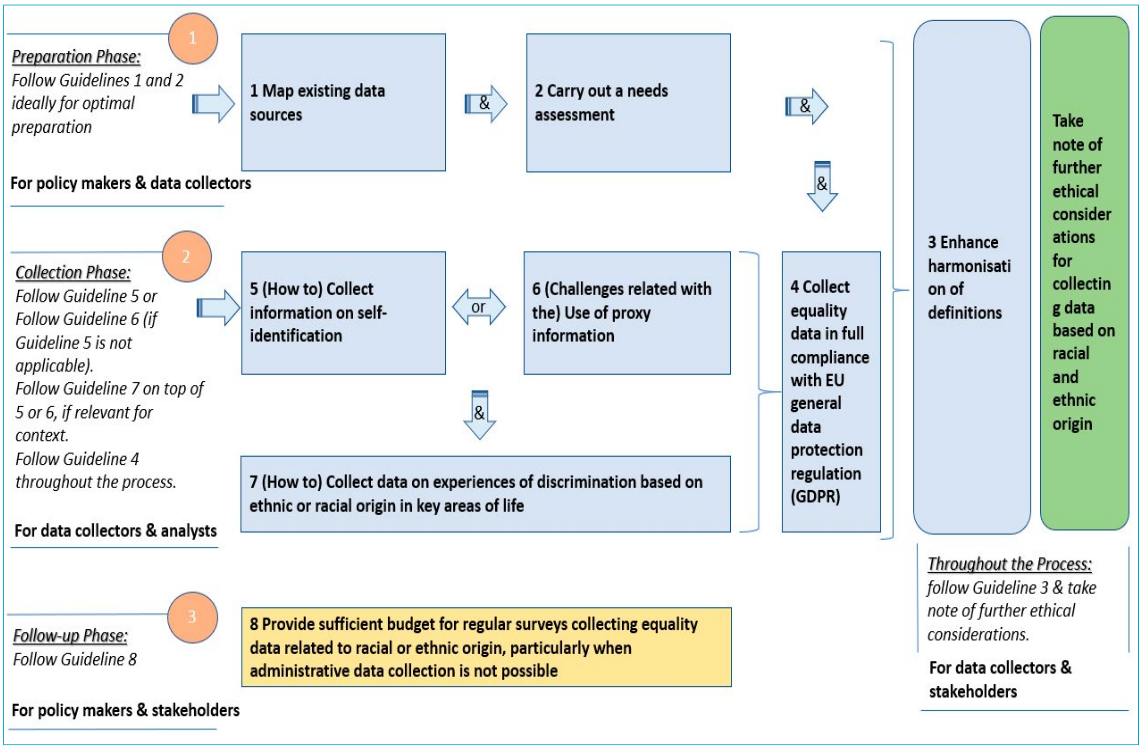
In addition, different communities/groups affected by racism and racial/ethnic discrimination may have diverging views about the need to collect such data or voice objections, which must be taken seriously by data producers.

This guidance can therefore serve as a starting point, to help Member States when they apply EU equality law in line with the applicable Racial Equality Directive. It considers the national legal contexts and the varying requirements for collecting equality data based on ethnic or racial origin in each Member State. Policymakers and other stakeholders may choose or prioritise actions set out in the guidelines that best fit the needs identified at the time in any given Member State.

EU Member States are encouraged to use these guidelines to inform their development and effective implementation of anti-racism and non-discrimination policies and monitor their outcomes from an equality perspective.

As requirements for collecting equality data on racial or ethnic origin may also vary across different data sources (for example, need for legal rules for collecting data in official statistics), the applicability of the following guiding principles might be easier for some data sources than others. Figure 1 shows a possible sequence and combination of the principles set out in this guidance note to support Member States in the process.

Figure 1: Guiding principles for collecting equality data on racial or ethnic origin – sequence and possible combination



GUIDING PRINCIPLES FOR COLLECTING EQUALITY DATA ON RACIAL OR ETHNIC ORIGIN

1. Map existing data sources that provide information on racial or ethnic origin (including the categories applied and their underlying definitions)

Rationale

A comprehensive (national) mapping of all available data related to ‘racial or ethnic origin’ (including underlying concepts and definitions) is a pre-condition for achieving a coordinated approach to data collection, including creating uniform categories, concepts and definitions.

Such a mapping would help to set up a baseline for a more systematic approach towards data on racial or ethnic origin and provide a solid basis for necessary improvements by identifying:

→ the different data sources and providers of equality data related to racial or ethnic origin, also taking into account sources not specifically designed to gather equality data, but include variables that can be used to show and analyse existing structural inequalities;

→ gaps and inconsistencies in the design and methods applied, as well as in the concepts and measurements.

→ unnecessary duplications in the collection of data.

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

The mapping of equality data disaggregated by racial or ethnic origin should be carried out periodically, taking into account the regularity with which the mapped data sources are updated, for example, every four to five years, and taking into consideration new potential data sources and providers.

Guidance

Member States and all relevant institutions and actors engaged in or responsible for collecting data related to/disaggregated by ethnic or racial origin are encouraged to:

→ take due consideration of the Guidelines on improving the collection and use of equality data, specifically Guideline No 1 “**Map existing sources of equality data and identify data gaps**”.

→ identify and mandate an institution/body/organisation that would initiate, coordinate and conduct the mapping exercise.

- → undertake a comprehensive national mapping of all existing data sources[18] that support disaggregation of data by the generic statistical category 'racial or ethnic origin'.

  In doing so, they should consider all possible constituting elements and proxy variables that could be subsumed under such a generic category, including their underlying concepts and definitions (see 'On the concept of racial or ethnic origin' in this guidance note).

For this mapping exercise, Member States could make use of the diagnostic mapping tool developed by the Subgroup on equality data in 2018, by adapting it for the needs of this specific mapping exercise.

- → use the results of the mapping to find commonalities and discrepancies across data sources and over time, regarding:
  - ⇢ the policy area/legislative file for which the data has been used or needed.
  - ⇢ the data source (administrative register, survey data, complaints data, etc.)
  - ⇢ the categories/classifications used for disaggregation or the questions used in a survey to measure 'racial or ethnic origin', etc. and their underlying concepts and definitions (and relevant changes to categories / classifications over time).
  - ⇢ intersections between grounds of discrimination or intersectional inequalities.
  - ⇢ the areas of life for which data are collected, such as, for example, (access to) employment, (access to) housing, education, health, access to justice, access to goods and services, etc.
  - ⇢ the methods and modes of qualitative and quantitative data collection (e.g., face-to-face interview, online survey, mail survey, telephone interview, etc.), sample size, representativeness, geographical coverage.
  - ⇢ public accessibility of data and analyses;
  - ⇢ the frequency with which the data has been collected (that is, the frequency of updates and existing time series as well as breaks or variations in frequency, if any).
- → Look at the results of the data mapping in combination with the findings of the user needs assessment (Guiding Principle No. 2), to further assess and revise persistent needs for collecting data disaggregated by racial or ethnic origin.

- → Ensure aggregate statistical data based on racial or ethnic origin is in the public domain for transparency and accountability purposes and in cases where such data could be used by the courts to ensure the right to effective remedy (Article 47 of the EU Charter of Fundamental Rights).

- → Similarly to Guideline No 3 of the Guidelines on improving the collection and use of equality data, Member States might envisage setting up a data hub (with a dedicated webpage) to collate and display data on racial or ethnic origin identified through the mapping exercise. This would enhance the accessibility and use of the data by interested stakeholders and data users, enforce effective implementation of policies (for example such data could feed into an anti-racism monitoring framework) and help to hold Member States and the relevant actors accountable. The hub could be hosted by the institution/body mandated to carry out the needs assessment (Guiding principle No. 2).

---

[18] Data sources collected and/or used by national authorities in areas covered by EU law with respect to equality and non-discrimination in relation to 'ethnic or racial' origin – which can be in reference to 'proxies' such as nationality, immigration status etc.'.)

Examples of national practice

The **Norwegian** Data hub on equality and living conditions of Sami, national minorities, and people with an immigrant background, coordinated by the Norwegian Directorate for Children, Youth and Family Affairs, maps and compiles existing sources of equality statistics and research on the living conditions of Sami, national minorities, and persons with immigrant backgrounds.

The key objective of the data hub is to provide a solid knowledge base to national and local policymakers working on policies relevant to ethnic minority groups and people with immigrant backgrounds. The data hub is updated when new or updated statistics are available and as such it helps ensure data are collected regularly and on time.

The data hub is one of the 50 actions in the Norwegian government's *action plan against racism and discrimination based on ethnicity and religion (2020-23)*. This data hub is built using the same online format and work method as earlier data hubs on disability and sexual orientation, gender identity and expression, and sex characteristics. Because ethnicity or religion is not recorded in Norway's administrative data and rarely recorded in research projects, this data hub primarily uses statistical sources that include information on immigration background.

Finland set-up a national discrimination monitoring system in 2008, which has a dedicated discrimination monitoring group attached to it and a website currently coordinated and developed by the Ministry of Justice. The task of the monitoring system is to (1) produce up-to-date information on discrimination in Finnish society, (2) compile research data and statistics produced by others, (3) promote cooperation among people and organisations working with discrimination research, and (4) suggest policy measures promoting non-discrimination.

The website is a platform for research, studies and statistics on discrimination in Finnish society (including on ethnic and racial origin as well as religion and belief). The information is categorised under five indicator areas agreed on by the dedicated discrimination monitoring group: 1) Attitudes, 2) Experiences and observations of discrimination, 3) Reports of discrimination and convictions, 4) Hate crime and hate speech, 5) Promotion of discrimination.

In **Italy**, a working group coordinated by the Italian national statistical office (Istat) was set up to implement Italy's national Roma integration strategy. The group's objective is to reach a shared perspective on scopes and methods of data collection. It aims to build a hub for mapping and circulating the official and unofficial (but reliable) existing sources. Although not the main purpose, this is the main challenge for the group, given the lack of reliable and comprehensive information on the living conditions of Roma people. In 2015 Istat-UNAR-ANCI carried out a mapping exercise on existing data sources for the Roma and identified data gaps. In 2019 Istat-UNAR mapped the Housing transition projects for Roma implemented by municipalities.

2. Carry out a needs assessment of (potential) users of data disaggregated by racial or ethnic origin

Rationale

A needs assessment should serve to identify what type of equality data disaggregated by racial or ethnic origin is required for evidence-based policymaking and monitoring of equality and human rights progression and regression.

The performance of a needs assessment should be based on the relevant EU, international and national normative standards to which Member States are subject, such as the EU Charter of Fundamental Rights, EU equality legislation, international human rights law and national legislation – which variously address equality and non-discrimination in relation to racial or ethnic origin.

The assessment should focus on EU-level policy coordination processes such as the European Semester, as well as relevant political commitments, including the European Pillar of Social Rights, the EU anti-racism action plan 2020-2025 and the UN Agenda for Sustainable Development.

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

It is recommended that a user needs assessment for equality data disaggregated by racial or ethnic origin is carried out periodically, for example every four to five years, taking into consideration new legislative and policy files or revisions of already implemented commitments, laws and policies.

Guidance

Member States are encouraged to:

→ undertake a comprehensive policy needs assessment for equality data disaggregated by racial or ethnic origin, by considering all relevant policy and legislative files, including international human rights law and relevant political commitments.

→ conduct a user needs assessment on a regular basis, for example every four to five years.

→ identify priorities for data collection based on the urgency of users' needs.

→ communicate the findings of the user needs assessment through different channels and in different formats to all relevant institutions, bodies, statistical offices, research institutions, CSOs, etc. involved in the collection of data disaggregated by racial or ethnic origin.

- → ensure data is collected to monitor obligations under EU law, for example through government administrative data or national surveys covering both the public and private sector, etc.

Note – Member States that do not collect data on racial or ethnic origin in official statistics or through large-scale surveys – and where scepticism regarding the collection and use of such data is fairly widespread – should consider a structured dialogue with different stakeholders (decision makers, public authorities, research institutes, civil society organisations, etc.), to reach a consensus on what kind of data should be collected and under which conditions.

Examples of national practice

In **Ireland**, there is a Public Sector Equality and Human Rights Duty under Section 42 of the Irish Human Rights and Equality Commission Act 2014. This forms part of a legislative framework governing human rights and equality in Ireland and imposes a statutory obligation on public bodies to seek to:

- → eliminate discrimination;
- → promote equality of opportunity and treatment for staff and for people receiving the services it provides;
- → protect the human rights of staff and service users.

The duty also requires a public body to *assess*, *address* and *report* on equality and human rights in how they are performing their functions. The Irish Human Rights and Equality Commission (IHREC) Guidance on Public Sector Duty advises an evidence-based approach to a public body's equality and human rights assessment and action plan, and this includes the special category of data on racial or ethnic origin.

The IHREC also acknowledges that while decisions around the processing of equality data are a matter for each public body as data controller, it believes there is a legal basis allowing public bodies to process equality data for the purposes of statistical analysis - and that the Public Sector Equality and Human Rights Duty is a statutory obligation for all public bodies.

In **Ireland**, Tusla – the Child and Family Agency commissioned an exploratory study in 2019 to support the development of an ethically appropriate, legal ethnic data collection system within Tusla that adheres to a human rights framework. The project has three main objectives:

1. identify the legislative and policy context within which Tusla can develop an ethnic data collection system.
2. establish best practice guidelines for Tusla in ethnic data collection methods and systems.
3. ensure Tusla's information systems can apply best practice to improve service provision.

As part of a partnership agreement, the UNESCO Child and Family Research Centre (UCFRC) was commissioned to conduct this assessment on behalf of Tusla.

In **Croatia**, a lack of robust and comprehensive ethnically disaggregated data at national level, including baselines for monitoring the National Roma Integration Strategy (NRIS), led to a project described by the Fundamental Rights Agency as a promising practice in the online Compendium of Practices for equality data collection. This project monitored progress regarding Roma inclusion and integration. The basis for the study was a mapping of the Roma population, combining external and self-identification methods, which gave a precise indication of the size of the Roma population in Croatia for the first time ever. The project led to the creation of a comprehensive anonymised database and some preliminary results.

As a follow up, in addition to providing detailed sectoral analysis and recommendations for all aspects of Roma inclusion policy and the design of the National Roma Integration Plan 2021-2027 through approximately 350 variables, the current project involves components of capacity building for the National Roma Contact Point (NRCP) and other stakeholders involved in implementing and monitoring the NRIS, and with additional attention to Roma women and youth.

Significant resources are also provided for regional and local meetings, and media campaigns aimed at both the Roma and the majority population. Roma representatives have been closely involved throughout the project cycle.

Another example is the **Croatian** ombudsman's annual report. This report is based on complaints received, field work, research and data collected from several hundred stakeholders, including public authorities, civil society organisations, trade unions, employers, universities, churches, religious organisations and many others.

In **Portugal**, the national statistical office introduced a Pilot survey on living conditions, origins and trajectories of the resident population**,** whose main purpose was to test a questionnaire containing modules specifically focused on discrimination experiences, as well as a question to measure respondents' self-identification in terms of ethnic origin. Moreover, the pilot survey tests the sample design, to ensure the diversity of the resident population, such as specific ethnic minority groups, is captured.

This exercise is in line with national anti-racism policies, namely the national parliament's recommendation to the government for implementing a national strategy against racism (Resolution of the Assembly of the Republic n.º 16/2021) and for adopting measures to combat racism, within which a recommendation for collecting data on ethnic discrimination is made (Resolution of the Assembly of the Republic n.º 11/2021).

The aim is therefore to collect data on ethnic origin in relation to discrimination and inequality. The final questionnaire, which will target the whole population, will also include several comprehensive modules on living conditions, such as health, housing, access to goods and services, income, family and fertility, and ICT usage.

In **Italy**, Istat and the National Office against Racial Discrimination (UNAR) signed an agreement to build up an information framework on "Inclusion and housing conditions of Roma, Sinti and Caminanti". The project is a part of the National Strategy for the Inclusion of Roma, Sinti and Caminanti 2012-2020 and is funded under the activities of the National Operational Programme Inclusion 2014-2020 co-funded by the European Social Fund. The aim is to improve statistical knowledge on Roma population in Italy by developing a system of indicators to monitor inclusion policies, with a special focus on Roma **housing transitions**.

The survey on housing transition projects was the first component of a wider research that involves Istat, UNAR and Roma representatives in the National Roma Associations Platform. The objective of this first survey was to assess all housing inclusion projects carried out by municipalities (with over 15,000 inhabitants). The practice is still ongoing. The second step scheduled in 2021 will consist of interviewing Roma, Sinti and Caminanti who left the settlements in 2012-2020 and live in adequate houses and those who did not leave them.

In 2019, the Italian Senate established an Extraordinary Commission to fight the phenomena of intolerance, racism, antisemitism and incitement to hatred and violence (*Commissione straordinaria per il contrasto dei fenomeni di intolleranza, razzismo, antisemitismo e istigazione all'odio e alla violenza*). The Commission is mandated to collect statistical data on hate crime and hate speech from organisations that work in countering these phenomena: Prevenzione e contrasto dell'antisemitismo (camera.it).

3. Align definitions, classifications and categorisations related to racial or ethnic origin and mainstream data on racial or ethnic origin into EU and national surveys

A uniform approach in data collection would maximise the comparability, validity and reliability of the data collected, as well as minimise costs for data collection.

Rationale

Governments need to ensure that data is collected (through government administrative data, national surveys, covering both the public and private sector, etc.) to monitor obligations under EU law. Given the complexity and variety of possible terms and proxy variables to be used when collecting data on racial or ethnic origin, the aim is to achieve harmonised data over time, within and across different sources, in order to:

1. **maximise:**
   a. comparability (at national level and across the EU).
   b. validity and reliability of the categories applied to capture ethnic or racial origin.
   c. opportunities for linking different data sources for further analysis and equality reporting.
   d. users' awareness of data needs and their understanding of statistical information based on ethnic or racial origin (data literacy).
   e. the efficiency of data collection.
2. **minimise:**
   a. time and costs in developing relevant categorisations and data collections.
   b. unnecessary duplications across data sources.
   c. unnecessary surveillance and holding of data on ethnic or racial groups.

Guidance

Member States and relevant institutions and bodies involved in or responsible for collecting data based on ethnic or racial origin are recommended to:

- → take due consideration of the Guidelines on improving the collection and use of equality data, specifically Guidelines No 9 ("Enhance validity and reliability of equality data") and 11 ("Improve comparability of equality data").
- → build on the findings from the mapping exercise (Guideline 1) to identify commonalities and discrepancies, across data sources and time, regarding the categories/definitions applied.
- → build on the most recent findings from academic research in sociology and social psychology, specifically on 'racism', 'group relations', 'ethnic' and 'racialised identities' – to leverage the awareness of potential data collectors and data users about the social construction of racial or ethnic categories.
- → undertake community consultations with representatives of all relevant ethnic minorities and racialised groups, to (a) identify the most appropriate categories for disaggregation/ questions on self-identification to be asked in a survey etc. and (b) reach a consensus on the underlying definitions. In line with the UN Human Rights-Based Approach To Data, community consultation and participation should take place as much as possible at all levels of research: including planning, data collection, dissemination, and analysis of data.
- → a balance needs to be struck between community-led and participatory work and an understanding of the constraints on quantitative data collection for monitoring purposes at comparative level and the need to inform people about statistical data collection – that is, how it can benefit a community. Therefore, statistical literacy and capacity building among participating groups needs to be strengthened.[19] "Putting collected data back in the hands of disadvantaged population groups and strengthening their capacity to use them is essential for accountability."[20]
- → (building on the results of the community consultations) carry out consultations with survey-methodology and survey-design experts, as well as data experts involved in collecting administrative data (e.g. national registers, censuses etc.), to discuss challenges and problems of potential categorisations for disaggregation/ survey questions (including proxy data relating to racial or ethnic origin) for applications in different data collections.

  These consultations should seek to reach an agreement about the most effective, valid and reliable approaches to data on racial or ethnic origin. They could also provide the basis for developing standards or promising practices in the operationalisation of relevant categorisations.

---

[19] OHCHR, 2018, *A Human Rights-Based Approach to Data: Leaving no one behind in the 2030 agenda for sustainable development*, United Nations 2018, p.4

[20] OHCHR, 2018, *A Human Rights-Based Approach to Data: Leaving no one behind in the 2030 agenda for sustainable development*, United Nations 2018, p. 19.

- → duly consider differences in the requirements towards different data sources (administrative registers/censuses/ survey data etc.) and the feasibility/leeway of applying uniform categories in different data sources.

- → duly consider the variety of all possible categories and proxy variables that could be used to measure 'racial or ethnic origin', as listed in the section 'On the concept of racial or ethnic origin' of this guidance note, while maintaining applicability across contexts and reliability. National definitions and concepts used will depend on national conditions and needs.[21]

- → document the basic criteria and classification procedures for ethnicity and inform data users about the concepts they are based on – bearing in mind that, according to the UN Principles and Recommendations for Population and Housing Censuses related to Ethnicity (para 4.187), the ethnocultural composition of a country can vary widely from country to country and due to the diversity in the approach and the various criteria for establishing ethnicity, so it is recognised that there is no single definition or classification that could be recommended that would be applicable to all countries.

- → take due consideration of the Guidelines on improving the collection and use of equality data, and specifically Guideline No 7 "Mainstream equality data into EU and national surveys" to incorporate (1) agreed categories for disaggregation by racial or ethnic origin; (2) questions measuring discrimination/crime victimisation based on racial or ethnic origin into EU and national surveys.

- → many Member States collect proxy information indicating ethnic or immigrant background (together with other equality-related background variables such as 'disability') for the European Statistical System (ESS) – through the Labour Force Survey (LFS), the EU Statistics on Living Conditions (EU-SILC) survey or EU-wide surveys such as the Eurobarometer.

  However, such socio-economic characteristics have not yet been regularly and systematically used as disaggregation categories for equality statistics. Therefore, the national statistical offices and Eurostat should strengthen and mainstream the regular use of such categories to disaggregate all data collected for the ESS.

- → equality bodies can usefully cooperate to examine their own systems of equality data collection based on racial or ethnic origin, address the barriers to such systems being effective and/or comparable and interoperable, and consider how best to improve and coordinate their data collection based on racial or ethnic origin.

- → according to the *Human Rights-Based Approach to Data*, "Data collectors are also accountable for the impact of their data collection activities and the publication of data."[22]

---

[21] United Nations Economic Commission for Europe (UNECE) (2020), *Poverty Measurement – Guide to Data Disaggregation*, p. 33.

[22] OHCHR, 2018, *A Human Rights-Based Approach to Data: Leaving no one behind in the 2030 agenda for sustainable development*, United Nations 2018, p. 19.

Examples of national practices

In **Portugal**, different data users requested Statistics Portugal to collect data on the ethnic origin of the population, as well on experience of discrimination and inequality that might be related to the population's migratory background and ethnic origin.

Developing a classification for the ethnic origin of the population without having aligned practice and applied common guidelines has been seen as a major challenge. The results of the Pilot survey on living conditions, origins and trajectories of the resident population (for details, see Guideline 2, above) will be analysed to improve the final questionnaire and methodology, as well as the sample design. Although the practice is still in its initial phase, it provides insights into the key challenges that arise when collecting data on racial or ethnic origin.

In **Belgium**, the Socio-economic Monitoring 2019 – Labour Market and Origin worked with two variables to determine the origin of a person: an "origin" variable and a "migration background" variable.

The "origin" variable makes it possible to be more precise than indicators that stop at the citizenship of the individual, by accurately capturing the second generation, as well as people who have become Belgian via the birth nationality of the parents or the individual.

The "migration background" variable makes it possible to distinguish between recent and older immigrants, immigrants and members of the "second" or "third generation", people who obtained Belgian nationality or who were born Belgian to parents of foreign origin, or people who remained foreigners. The method uses multi-step algorithms for both variables to define exclusive groups of people.

In **Luxembourg**, the victimisation survey was launched on 1 October 2020 and covers all people aged 16 or over living in the Grand Duchy. The survey on safety collects detailed information on crimes and offences affecting this population over the last five years/over the last year/during their lifetime. It also aims to measure satisfaction with the work of the police and the courts. In addition, it asks questions about people's feeling of safety, their exposure to drug problems or vandalism in their neighbourhoods, or about possible measures people have taken to ensure their safety.

The questionnaire contains questions on minority status (ethnic minority, religious minority, and sexual minority, minority related to skin colour or a disability). This background variable will be used to determine satisfaction with the police in general and to check whether members of minorities are more satisfied or dissatisfied with the police and the justice system than the rest of the population. In addition, it can be determined whether they are more often victims of violence.

The survey questionnaire has been prepared on the basis of the 2013 survey questionnaire, while taking advantage of the experience gained from the EU Safety Survey project, the European Crime and Safety Survey (EU-ICS) (in which Luxembourg participated in 2004/2005) and work by the Eurostat Task Force to develop a questionnaire on gender-based violence. https://statistiques.public.lu/fr/enquetes/espace-menages/securite/

In **Norway,** the national statistical institute, Statistics Norway, has created two definitions to help determine people's origin, which are applied to both administrative data and survey data:

- *Immigrants*: individuals who have immigrated to Norway, and who were born in another country to non-Norwegian born parents, and four non-Norwegian-born grandparents
- *Norwegian-born with immigrant parents*: born in Norway to non-Norwegian-born parents, and with four non-Norwegian born grandparents.

In this way it is possible to distinguish between first and second-generation immigrants in Norway. This does not apply to people adopted from abroad, and therefore it is the social and not the biological origin that is considered.

Examples of surveys with these variables include the Norwegian survey on quality of life (first time in 2020) and Survey on living conditions among persons with immigrant background (2016). In both surveys the variables make it possible to compare the part of the population originating in another country with the entire population. These variables are also applied to administrative data on use of healthcare, education, social welfare, election participation, income, and labour market statistics.

Since 2016, the **Dutch** Ministry of the Interior, the National Police and the National Association of Municipal Anti-discrimination Services have been publishing figures on discrimination incidents registered by the police and complaints received by equality bodies in a single multi-agency report.

This report brings together, presents and contextualises the data from several institutions dealing with discrimination complaints in different contexts. It also explains the relations and differences between the relevant datasets.

Previously, annual figures on discrimination incidents and complaints were published separately by some of these agencies (National Police, National Association of Municipal Anti-discrimination Agencies).

Attention is being paid to the extent to which categories across agencies are comparable. Similarly, the figures are presented in the context in which they are collected. However, the agencies are trying to align their practice as much as possible. Some differences remain, due to the legal structure on which each agency is based.

Examples from approaches equality bodies take to enhance harmonisation

In 2018, a working group on research and data collection from Equinet (European Network of Equality Bodies) conducted a survey on the collection and use of complaints data. The aim of this survey was to identify commonalities among equality bodies regarding the information on complaints they collect and the purposes they use the data for.

Despite differences in mandates and functions, the results from these bodies show some important similarities. For instance, almost 90% of those taking part in the survey log the number of complaints made about racial/ethnic discrimination. The report also features recommendations on how these bodies could improve their collection and use of complaints data. For instance, they "*should apply as disaggregated and detailed categories as possible for collecting data on complaints*" and clarify the definitions used, to increase the comparability of their complaints data.[23]

---

[23] FRA (2021), Equality in the EU – 20 years on from the initial implementation of the equality directives, Luxembourg, Publications Office of the European Union (Publications Office), p. 56-58.

4. Collect and use equality data in full compliance with EU General Data Protection Regulation and national data protection rules

The collection of personal data disaggregated by personal characteristics, such as racial or ethnic origin, is protected by constitutional norms, EU data protection law and the Charter of Fundamental Rights.

The EU General Data Protection Regulation (GDPR), which entered into force in 2018, as well as Directive 2016/680 (Article 10) on processing of personal data by criminal justice authorities, have triggered reflections about how to legally collect and process "special categories of personal data" such as those related to racial or ethnic origin, health, religion/belief and sexual orientation.[24]

Rationale

The EU Anti-Racism Action Plan 2020-25 stipulates that "*a prerequisite for progress towards a common dataset is full respect for constitutional norms, EU data protection law and the EU Charter of Fundamental Rights. Safeguards need to be in place to ensure that sensitive equality data cannot be related back to the individual. This implies full compliance with data protection rules, to mitigate any potential risks of misuse or abuse.*"[25][26]

Guidance

→ All equality data collection and processing should be done in full compliance with the principles and safeguards in the GDPR.

→ In line with the GDPR, data should not be published or publicly accessible in a manner that allows individual data subjects to be identified, either directly or indirectly.

→ "*The GDPR allows for the collection and processing of special categories of personal data under certain conditions, including for statistical or research purposes (Article 9 (2) (a), (g) and (j)). Data collectors and data processors in EU Member States should seek the advice of their national data protection authorities and further guidance from the European Data Protection Board (EDPB) and the European Data Protection Supervisor (EDPS) about the safeguards that need to be applied when collecting and processing special categories of personal data, including for the purpose of scientific research (Article 9 (2) (j) of the GDPR). In doing so, they should take due consideration of the EDPS preliminary opinion on data protection and scientific research from 6 January 2020 and the upcoming EDPB guidance on data protection and scientific research.*"[27]

---

[24] Under Article 9 of the GDPR special categories of personal data comprise "personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership", as well as "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."

[25] EU Anti-Racism Action Plan 2020-2025, p. 16.

[26] FRA (2021), *Equality in the EU 20 years on from the initial implementation of the equality directives*, Luxembourg, Publications Office of the European Union (Publications Office), pp. 19f and pp. 65f.

[27] FRA (2021), *Equality in the EU 20 years on from the initial implementation of the equality directives*, Luxembourg, Publications Office of the European Union (Publications Office), p. 21.

Safeguards:

Under Article 9 of the GDPR special categories of personal data comprise “*personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership”,* as well as *“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.”*

Article 9(2) of the GDPR stipulates that the general prohibition to process special categories of personal data (Article 9(1)) *does not apply* in several cases, which include:

- → when the **data subject has given explicit consent to the processing** of that personal data for one or more specified purposes, except where EU or national law prevent the data subject from lifting the prohibition referred to in paragraph 1 of Article 9;
- → when processing is necessary **for reasons of substantial public interest, on the basis of EU or national law** – providing the processing is proportionate to the aim pursued, respects the essence of the right to data protection and provides for suitable and specific measures to safeguard the data subject’s fundamental rights and interests;
- → when processing is necessary for **archiving purposes in the public interest, scientific or historical research purposes or statistical purposes**, in accordance with Article 89(1)[28] **based on EU or national law** - providing the processing is proportionate to the aim pursued, respects the essence of the right to data protection and provides for suitable and specific measures to safeguard the data subject’s fundamental rights and the interests.
    - ⇢ These safeguards ensure that technical and organisational measures are in place to ensure the principle of data minimisation is followed. These measures may include pseudonymisation[29], provided that the purposes (as listed above) can be fulfilled in that manner. Where those purposes can be fulfilled by further processing which does not permit or no longer permits the identification of data subjects, those purposes shall be fulfilled in that manner.

---

[28] Article 89 of the GDPR sets out further safeguards and the possibility of derogations based on EU or national law, specifically regarding processing for archiving purposes in the public interest; scientific or historical research purposes; or statistical purposes.

[29] ‘Pseudonymisation’ means processing personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure the personal data are not attributed to an identified or identifiable person. Source: https://data.consilium.europa.eu/doc/document/ST-5419-2016-INIT/en/pdf p. 112.

→ Member States are also able under the GDPR to enact '*further conditions, including limitations, regarding the processing of genetic data, biometric data or data concerning health*' (Article 9(4)). This is therefore a new area and requires adoption of EU or national law before the use of special categories of data for research purposes can become fully operational.

→ In addition, Recital 26 of the GDPR clarifies that the principles of data protection apply to special categories of personal data that concern an identified or identifiable natural person. By contrast, they **do not apply to anonymous information or to personal data rendered anonymous** in such a manner that the data subject is not or is no longer identifiable. This implies that if the information processed is anonymous or has been rendered anonymous, such data can be used for aggregate statistical purposes to identify and record trends in equality.[30]

→ In its preliminary opinion on data protection and scientific research, the EDPS recommends: "*intensifying dialogue between data protection authorities and ethical review boards for a common understanding of which activities qualify as genuine research, EU codes of conduct for scientific research, closer alignment between EU research framework programmes and data protection standards, and the beginning of a debate on the circumstances in which access by researchers to data held by private companies can be based on public interest.*"

→ To the extent possible, Member States are encouraged to remove barriers that apply to some public bodies, which are not currently collecting data disaggregated by ethnic and racial origin.

---

[30] FRA (2021), *Equality in the EU 20 years on from the initial implementation of the equality directives,* Luxembourg, Publications Office of the European Union (Publications Office), p.21.

Examples of national practice

The Ombudsman in **Croatia** issued recommendations to the competent public administration bodies and authorities (especially in the fields of internal affairs, health, social welfare, labour, pension insurance and education) to start collecting and processing special categories of personal data, including those on ethnicity or national origin, applying adequate protective measures to better understand structural discrimination. In the past, the state administration has regularly argued that they cannot do this due to GDPR.

The **Croatian Public Health Institute** established cooperation with the Ministry of Internal Affairs to improve monitoring of the health (and the social determinants for it) of members of the Roma national minority. This follows a positive opinion from the Personal Data Protection Agency on collecting data on national and ethnic affiliation in compliance with the GDPR. This is one of the measures included in the Action Plan on implementing the Roma Inclusion strategy 2021-2027, managed by the Croatian Government's Office for Human Rights and Rights of National Minorities.

The **Portuguese** statistical office worked closely together with the Statistics Portugal Data Protection Officer to ensure all GDPR requirements are met for the Pilot survey on living conditions, origins and trajectories of the resident population in Portugal. For the questionnaire, adjustments have been made, such as:

- making the survey voluntary, unlike in all other household surveys, where it is mandatory;
- respondents and those who live in the household will have the option not to give their names;
- interviewers will be obliged to submit the finished interviews to the central box immediately, so that personal information from interviewees can be deleted from their work computers as soon as possible (where CAPI or CATI interviewing modes are used);
- names, addresses and other characteristics of the household members will not be included in the final database available for internal analysis purposes;
- the flow of all data within Statistics Portugal computer systems will be duly assessed, to address all possible security concerns.

In the **Netherlands,** under their Barometer on Cultural Diversity within organisations, a pilot among eight public and private organisations received a positive evaluation from the data protection authority. Because the privacy of individual employees is important, Statistics Netherlands analyses workforce data supplied by employers anonymously and only produces results at an aggregate level. There is no need to collect and present data on individual employees. All data are anonymous and aggregated at group level. If the size of the group is too small, the statistical office will not publish these data.

5. (How to) Collect information on self-identification based on ‘racial or ethnic origin’

Rationale

According to the UN Principles and Recommendations for Population and Housing Censuses related to Ethnicity (para 4.183.), data on ethnicity provide information on the diversity of a population and can serve to identify subgroups in a population. Some areas of study that rely on such data include demographic trends, employment practices and opportunities, income distributions, educational levels, migration patterns and trends, family composition and structure, social support networks, health conditions and crime victimisation.

The OHCHR Human Rights-Based Approach to Data states that populations of interest should be self-defining, which means that the parameters of the population cannot be imposed by an external party or assigned through imputation or proxy.

Therefore, collecting information on person’s self-identification in terms of racial or ethnic origin might be relevant for:

→ purposes of data disaggregation, to analyse (structural and systemic) inequalities in outcomes.

→ analysis of intersectional inequalities (e.g. examining information on racial or ethnic origin in relation to data on other protected characteristics, such as sex, age, disability, religion, belief, sexual orientation, etc.).

→ better understanding the factors that potentially contribute to, reinforce, or underlie systemic/structural inequalities in outcomes as well as discrimination and racism.

→ empowering ethnic and other racialised minorities and groups at risk of discrimination: *“statistical evidence based on racial or ethnic origin can play a decisive part in legal proceedings. Individual claimants often find themselves in need of statistical evidence to back up their claim, particularly where indirect discrimination is at issue,* […] and as *recognised in the Racial Equality Directive (recital 15).”*[31]

→ purposes of sampling or screening eligible respondents in a survey.

As mentioned in the UNECE Guide to Data Disaggregation for Poverty measurement, *“ethnic identity can be measured using a variety of concepts, including ethnic ancestry or origin, ethnic group, cultural origins, nationality, race,* [skin] *colour, minority status, tribe, language, religion* – [and in numerous cases through proxy variables such as country of birth, country of birth of parents, citizenship] - *or various combinations of these concepts”*.[32]

---

[31] Al-Zubaidi, Y. (2020), ‘Some reflections on racial and ethnic statistics for anti-discrimination purposes in Europe’, in: *European equality law review*, Issue 2, Luxembourg, Publications Office, pp. 62-72, p. 65.

[32] United Nations Economic Commission for Europe (UNECE) (2020), *Poverty Measurement – Guide to Data Disaggregation*, p. 33.

Guidance

→ For the purposes of data collection, **populations of interest should be self-defining**. This principle is essential, especially when 'racial or ethnic origin' refers to an aspect of a person's attachment to or identification with an ethnic or any other minority group. The ascription/identification of a respondent's 'racial or ethnic origin' attributed by a third party (e.g. interviewer or a service provider) does not necessarily correspond with a person's actual identification.

In cases where it might be necessary for logistical, political or other reasons to use demographic characteristics to identify a particular population, the OHCHR guide towards a human rights based approach to data recommends that data collectors ensure to make public that their handling and publishing of that data does not imply self-identification if personal information on ethnic identity has not been disclosed.[33]

→ Explain the purpose of any data collection related to racial or ethnic origin and base the data collection around best practice and transparency. This would increase the level of trust around any ethnic data collection.

→ Data collectors are recommended to seek the involvement of the communities/groups of interest in all aspects of their collection activities, including finding the most proper categories for capturing an individual's 'racial or ethnic origin' as a social description. This would ensure that affected communities/population groups are treated as experts on their own reality and help counter unconscious bias.

→ As definitions, conceptualisations and categorisations can influence potential responses, those who undertake data collection based on racial or ethnic origin should ensure that the basic criteria used to measure the concept 'racial or ethnic origin' in a survey, census, administrative data source, etc. are clearly explained to respondents; the same is relevant for the dissemination of the resulting data.

→ As 'attachment to/identification with an ethnic/minority/racialised group' is multidimensional and is more a process than a static concept, this classification should be treated with movable boundaries.[34] Therefore, it is recommended that data collectors give respondents the choice to indicate multiple ethnic/group affiliations or a combination of ethnic/group affiliations (i.e. option for mixed categories). Information on multiple (mixed) identities could be collected either through a single multiple-response question or by using several single-response questions. To capture identification as a process and avoid static conceptualisation, attachment or identification based on racial or ethnic origin could be also measured on a scale (e.g. from "no or low attachment" to "high attachment").

---

[33] OHCHR, 2018, *A Human Rights-Based Approach to Data: Leaving no one behind in the 2030 agenda for sustainable development*, United Nations 2018, p. 13.

[34] This can be an issue particularly for administrative data collection. For example, if a health authority starts collecting, in their forms, a person's self-declared ethnicity, this will stay in the system until the information is either removed or the person wants to change their information. Unless there are opportunities provided to make such changes, and people are regularly reminded about the information about them that is saved in the system, certain characteristics – once collected – might stay in the system, even though they no longer correspond to the way a person would characterise themselves today.

- → For administrative sources: Classifications used in administrative data also have to be regularly updated. Given that there is no guarantee that a person continues to self-identify in the same way, especially if a long time has passed since the information was originally collected and entered into a register or administrative source, it is recommended to provide people the opportunity to revise this information whenever a person requires such a revision on their own initiative).

- → Depending on the mode of data collection (online, face-to-face, telephone) and the data source (administrative register, census, survey, etc.), collectors should explain in the technical instructions and documentation how the 'racial or ethnic origin' of children from "mixed" couples is to be reported (for example, explicit instructions to allow respondents to provide multiple responses or to allow for responses such as "Biracial").

- → Following the UN Principles and Recommendations for Population and Housing Censuses related to Ethnicity (para 4.185.), classifying ethnic/minority groups also requires the inclusion of the finest levels of ethnic/group affiliations, self-perceived groups, regional and local groups, as well as groups that are not usually considered to be ethnic groups, such as religious groups and those based on nationality.

- → Countries collecting data on ethnicity should note that the precoding or pre-classification of ethnic groups at the time of data capture may have a tendency to lose detailed information on the diversity of a population **unless space to record 'other' unspecified ethnic identities/groups and free-text responses are possible.** For example, having consulted with all relevant Roma organisations in Bulgaria during the implementation of the Fundamental Rights Agency's EU-MIDIS II survey, the survey used a showcard with 25 categories for self-identification during the screening for eligible Roma respondents at the doorstep. The list included for example categories such as: *Roma, Gypsy, Ierlii, Demirdjii, Bakardjii, Dzambasi, Rudari/Ludari, Kopanari, Vlasi* etc. Therefore, data collectors have to decide which categories are the most important and meaningful for the respondents, which should be ideally done in cooperation with the communities/groups from which data is going to be collected.

- → When using a free-text response option to collect information on self-identification, data collectors are invited to consider the following caveats: (1) people do not always spontaneously think of themselves in racial or ethnic terms; (2) an open question often produces a very large number of answers, which need to be reclassified and reduced in a next step to make them useful for further analysis.

- → The question(s) on racial or ethnic identifications should be asked in a way that allows the respondent to see all possible response options before making decisions.[35]

---

[35] United Nations Economic Commission for Europe (UNECE) (2020), *Poverty Measurement – Guide to Data Disaggregation*, p. 33.

→ Individuals should have the option to disclose or withhold information about their personal characteristics. This means that providing personal information about one's ethnic affiliation should be optional for the potential respondents and this has to be clearly communicated to them in the instructions preceding the question on how they self-identify in racial or ethnic terms.

Data should be collected with the consent of the respondent, opt-out should be an option (for example, to guarantee the free self-declaration of 'racial or ethnic origin', respondents should be allowed to indicate "None" or "Not declared"[36]). Participation should take place on a voluntary basis. Especially as questions about the experience of racism and discrimination can entail emotional labour and stress, it is necessary to avoid over-burdening respondents.

→ Data collectors and organisations/bodies involved in data collection activities are recommended to invest in staff training (for enumerators and interviewers as well as data analysts and service providers). This could include developing an online training portal. Ethnic communities may be asked to supply external training about awareness.

→ Data collection activities should not have a negative impact: they should be conducted following the human rights principle of 'doing no harm'. Information gathered through previous consultations and participatory processes should be reviewed, to avoid over-burdening vulnerable groups. A potential barrier to participation, regardless of the mode of interview, could be the fact that some minority groups are often approached for participation in survey research, but do not necessarily see changes as a result or any direct benefits. Data collectors need to clearly highlight the benefits of participation and what the key aims are, alleviate concerns about data protection and privacy and, where possible, demonstrate where change has been achieved through other similar surveys.

→ Member States and relevant data collectors should aim to regularly share good practice on *whether* and *how* to collect information on self-identification in administrative sources. The collection of data on racial/ethnic self-identification for administrative purposes or in administrative data sources is possible in some cases. In a selected number of countries (e.g., Bulgaria, Hungary, Ireland, Slovakia), such data has been collected through census. In most Member States, however, administrative data sources are primarily relying on proxy information that indicates ethnic or immigrant background (for proxies, see Guiding principle No. 6).

→ Data collectors should provide clear, openly accessible information about their operations, including research design and data collection methodology. Data collected by state agencies should be openly accessible to the public.[37]

---

[36] UN Principles and Recommendations for Population and Housing Censuses related to Ethnicity (para 4.186).

[37] OHCHR, 2018, *A Human Rights-Based Approach to Data: Leaving no one behind in the 2030 agenda for sustainable development*, United Nations 2018, p. 14.

Examples of national practice

The 2016 Irish census asked questions on racial/ethnic self-identification as follows:

What is your ethnic group?

A. White

1. Irish
2. Irish Traveller
3. Any other white background

B. Black or black Irish

1. African
2. Any other black background

C. Asian or Asian Irish

1. Chinese
2. Any other Asian background

D. Other, including mixed background

A public consultation exercise took place in 2019 to review the content of the next census questionnaire. As a result of that consultation additional categories were added to the ethnicity question for: Roma, Indian/Pakistani/Bangladeshi, Arabic and Mixed. For details, see the full report from the consultation exercise Below is the revised ethnicity question for the next census (scheduled for 2022):

What is your ethnic group/background?

A. White

1. Irish
2. Irish Traveller
3. Roma
4. Any other white background

B. Black or black Irish

5. African
6. Any other black background

C. Asian or Asian Irish

7. Chinese
8. Indian/Pakistani/Bangladeshi
9. Any other Asian background

D. Other, including mixed group/background

10. Arabic
11. Mixed, write in description
12. Other, write in description

In addition, separate questions are asked for to measure disability and religion. For details, see this copy of the Census 2022 form.

In **Malta**, the national Census planned for November 2021 will include questions on racial origin, religion and sexual orientation. With regards to racial origin specifically, the census questionnaire will include a question inviting a self-determined response. The information being collected will be a first for Malta, as this kind of information has never been collected in the past through the census or other national surveys.[38]

---

[38] Malta Today (2021), 'Census to collect data on race, sexual orientation and religion for first time', 13 May 2021.

In **Spain** in 2021, the General Directorate for Equal Treatment and Ethnic and Racial Diversity commissioned a study ‘Examination of the African and Afro-descendant population in Spain – Identity and access to rights’ with a focus on how African and Afro-descendant people living in Spain exercise and enjoy their civil, political, social, cultural and economic rights, as a specific social group and as a potential collective victim of discrimination. The study is based on two online surveys addressing relevant population and civil society organisations, with questions on experiences of discrimination and police profiling. In this study there was an open question allowing people of African descent living in Spain to self-identify.

In **Italy**, the Discrimination Survey also targets ethnic minorities and covers multiple grounds of discrimination and several areas of life, such as education, employment, health, access to goods and services and political participation. To address the possible existence of vulnerable groups, data subjects will be allowed to self-identify on the basis of characteristics such as physical characteristics, health conditions and disability, being immigrant or descendent of immigrants, ethnic origin and skin colour, religious affiliation, sexual orientation or gender identity. The categorisation systems allow more than one identification category, and also allow respondents not to declare or declare ‘none’. This project is still ongoing, and a pilot survey is scheduled in 2021.

In **Italy**, the Survey on Roma, Sinti and Caminanti Housing inclusion provides data to monitor the inclusion of these communities as regards housing conditions, which is part of the National Strategy for the Inclusion of Roma, Sinti and Caminanti 2012-2020. The aim is to improve statistical knowledge on Italy’s Roma population, developing a system of indicators to monitor inclusion policies, with a special focus on Roma housing transitions project. Core principles of the exercise are self-identification, voluntary and anonymised data collection and involving Roma representatives through the National Roma Associations Platform.

The **Croatian** Institute of Public Health created ‘Public Health Indicators of Roma People in the Republic of Croatia, a system that combines data from health databases and registers with ethnic data. By creating a system that would provide a realistic picture of Roma health from available and routine public health statistics, health measures planned using realistic indicators and evidence can be more focused and effective. Since data in public health databases and registers were not routinely monitored according to ethnicity, the implementation of this measure established an innovative system of analysis and reporting on Roma health, providing insights into sociodemographic characteristics.

Examples from surveys by the Fundamental Rights Agency (FRA)

The Agency's EU-MIDIS surveys targeting immigrants and descendants of immigrants use a multiple approach to 'racial or ethnic origin'.

1. To sample and screen eligible respondents, proxy information about respondents' country of birth and their parents' country of birth is used, as there are almost no available sampling frames in the Member States (such as address or individual registers) that would include self-identification information based on racial or ethnic origin.
2. To measure self-identification, the Agency uses a combination of separate survey questions asking respondents for information related to different aspects of group affiliations/attachments based on racial or ethnic origin, as shown in the tables below.[39]
3. The Agency uses a set of questions measuring respondents' experiences of discrimination based on racial or ethnic origin in key areas of life (see examples from other FRA surveys under Guiding principle 6 of this note).

Question wording:

IN09	**Self-identification as a person of African descent or a black person**
	Would you describe yourself as a person of African descent/ a black person?
	SINGLE RESPONSE
1	Yes
2	No
-96	*Prefer not to say*
-97	*Don't understand the question*
-98	*Not applicable*
-99	*Don't know*

Source: EU Survey on immigrants and descendants

---

[39] However, if a person was born in a Sub-Saharan African country, in FRA's analysis they will be based in the category 'immigrants from Sub-Saharan Africa', regardless of the extent to which they identify as European, as opposed to Sub-Saharan African. This is due to the (non-)available information on ethnic/racial identification in the sampling frames and the necessity to use respondents' country of birth and their parents' country of birth as an approximation, for sampling purposes.

RA02	**Self-identification as national/European/country national of respondent's country of birth /country of birth of parents**	
	ASK ALL	
	People might see themselves in different ways. The following question is about how you see yourself. **On a scale from 1 to 5, where 1 equals "not at all" and 5 "very strongly", to what extent do you feel…?**	
	SHOW CARD RA02 READ OUT RA02_1 TO RA01_5	
RA02_1 **RA02_2** **RA02_3** **…** **….** **….**	European Austrian Nigerian … … …	1 Not at all 2 3 4 5 Very strongly *-96 Prefer not to say* *-97 Don't understand the question* *-98 Not applicable* *-99 Don't know*

Source: FRA EU MIDIS II

The Agency's Fundamental Rights Survey – a survey of the general population in the EU – included a question about self-identification as an ethnic minority. In addition to this, the survey asked respondents about their country of birth and their parents' country of birth.

r29	**Whether belongs to an ethnic minority**
	ASK ALL
	Do you consider yourself to be part of an ethnic minority in [COUNTRY]?
	SINGLE RESPONSE ALLOWED
1	Yes
2	No
888	*Prefer not to say*
999	*Don't know*

Source: FRA Fundamental Rights Survey

6. Using proxy information when collecting data on racial or ethnic origin (when self-identification is not possible)

Rationale

Disaggregation by racial or ethnic origin requires that data on relevant characteristics are collected. **Lack of or incomplete** statistical evidence on racial or ethnic origin poses a concern, as it makes it difficult to assess:

1. the real extent/prevalence of the challenges that ethnic minorities or racialised groups face (such as racism, discrimination, structural inequalities in key areas of life, bias-motivated crime and harassment, as well as discriminatory profiling by law enforcement authorities).

2. the effectiveness of legislation and policies implemented to tackle inequality, discrimination and racism.

As previously mentioned, since categorisations on racial or ethnic origin are social constructs, there is no single self-evident meaning or measurement, and therefore no universal solution, when it comes to collecting data based on socially constructed distinctions.[40]

The most important question is, however, in which categories do disadvantages and inequalities appear – is it, for example, country of birth, nationality, skin colour, religion, or all together? And, to what extent do the categories used for equality statistics meaningfully and properly represent all disadvantaged people and groups in relation to racial or ethnic origin? Population numbers/census numbers are the sites where representation occurs, and through which political power can be contested. Proper counting therefore allows proper representation, and proper representation allows proper protection.

Challenges

→ For different reasons (historical, legal, etc.), in most EU Member States, "*the concept of race or racial origin is not used in data collection. The way racial or ethnic origin is understood impacts on the protection provided by the Racial Equality Directive and informs whether treatment apparently based on proxies such as immigrant status, nationality, language and certain religious practices is seen as direct or indirect, covert or overt discrimination.*"[41]

---

[40] Makkonen, T. (2017), *European Handbook on Equality Data (2016 revision)*.

[41] Farkas, L, (2017), *Analysis and comparative review of equality data collection practices in the European Union: Data collection in the field of ethnicity*, Luxembourg: Publications Office of the European Union, 2017, p. 6.

→ Most administrative sources use proxies for racial or ethnic origin, such as citizenship/ nationality, country of birth, country of birth of parents, language. According to the UNECE *Poverty Measurement – Guide to Data Disaggregation*, however, "*using proxy information can be challenging for certain groups and needs to be handled with care. For example, nationality-based data might be a very weak indicator for ethnicity due to differences in naturalisation policies, or for indigenous people or autochthonous ethnic groups.*"[42]

→ Using proxies, therefore, bears the risk of '**under-coverage**': For example, if equality statistics are based only on country of birth or nationality (as a proxy for ethnicity/ethnic origin), they will miss other kinds of disadvantages and inequalities that might occur in relation to skin colour, racial or ethnic origin and religion, to name just a few. Or using data about place of birth to show the impact of COVID-19 on ethnic or racialised minorities will miss the risk for people of African descent in their second or third generation. So, there is a gap between what can be shown by way of proxies and what would be relevant to know.

→ Using proxies bears the risk of homogenisation: For example, "EU citizens 'of migrant background', 'new- or non-majority ethnicities' or 'allochthones' share one characteristic: they are named in binary opposition to a country or the ethnic origin of the majority population in that country. These categories are also misleading, because they lump minorities into a single unit, without investigating whether they indeed belong together."[43] Therefore, when using proxies, researchers should reflect on the impact of their category choices in the data analyses.

→ Proxies conflict with the principle of self-identification and therefore fail to represent diversity of racial or ethnic origin in society. For example, in the case of first generation long-term residents, the proxy 'country of birth' might not be an efficient and operational indicator for ethnicity after 10-15 years of stay in the country of residence.

→ Migration (background), language, education levels and poverty data are not effective proxies for measuring discrimination based on racial or ethnic origin.[44] Countries that refuse to collect data based on racial or ethnic origin in relation to discrimination would therefore put an artificial limit on the interpretation of equality rights, and more specifically, the effective implementation of non-discrimination laws.

→ Discrimination based on racial or ethnic origin is not about how people define themselves but how others see/perceive them. In this case, what is relevant is third-party identification, but this information is difficult to collect in a census.

---

[42] United Nations Economic Commission for Europe (UNECE) (2020), *Poverty Measurement – Guide to Data Disaggregation*, p. 70.

[43] Farkas, L, (2017), *Analysis and comparative review of equality data collection practices in the European Union: Data collection in the field of ethnicity*, Luxembourg: Publications Office of the European Union, 2017, p. 38.

[44] Ibid., p. 6.

Guidance

- In the absence of proper sampling frames, using proxy information is often the only way to survey ethnic or racialised minorities.
- "*To measure (in)equality, it may be as important to identify the perceived racial and ethnic origin as the one self-identified. (Perceived) racial origin is an important element of the discrimination experiences of Afro-Europeans and European Muslims, as well as the Roma.*"[45]
- Given the highly segregated living conditions of the Roma in several Member States, affecting entire micro regions, poverty/income data could be used as a proxy, especially when combined with geographic location.[46]
- The use of proxies should be transparent and information on this should be publicly available. For example, information notes for data users, explaining the definitions and limitations of the applied proxies, should be added to the technical (meta) information for the data collection.
- Where the human rights-based principle of self-identification cannot be applied, data collectors should envisage refining and increasing the number of proxy variables/categories in existing data sources and analysis, to ensure comprehensive coverage of different population groups.
- Over time, proxies should be phased out and data systems strengthened.
- Data collectors and data users need multiple and complementary data sources of equality data based on racial or ethnic origin, to enable effective monitoring of the implementation of non-discrimination legislation. However, given the different ways of operationalising racial or ethnic origin, further alignment or coordination between different data sources and across time is needed to increase comparability (see Guiding Principle No.3).

Examples of national practice

- Comprehensive examples of national practices are provided in the report *Analysis and comparative review of equality data collection practices in the European Union: Data collection in the field of ethnicity.*

---

[45] Farkas, L, (2017), *Analysis and comparative review of equality data collection practices in the European Union: Data collection in the field of ethnicity*, Luxembourg: Publications Office of the European Union, 2017, p. 6.

[46] Ibid.

7. Collecting data on experience of discrimination based on ethnic/racial origin in key areas of life

FRA survey data regularly illustrates the extent and nature of experience of inequality and racial/ethnic discrimination across the EU. It does so with reference to the grounds of discrimination and areas of life covered by the applicable non-discrimination legislation in all EU Member States (for example the racial and employment equality directives).[47]

FRA data further shows that the prevalence of racial/ethnic discrimination remains consistently high, both over time and across different population groups in different Member States. It also shows that victims of discrimination tend not to report incidents they experience to any authority or body for a number of reasons, including not knowing where to turn to.[48]

As a result, incidents remain largely invisible to institutions that have a legal obligation to assist victims of discrimination, including equality bodies. This phenomenon, which is also known as attrition of actual cases of discrimination, proves to be a considerable challenge when it comes to providing a reliable picture about the extent of racial/ethnic discrimination, as it means that only a small part of actual discrimination experienced by people at risk is visible (the 'tip of the iceberg'). Survey data, therefore, is the best approximation of the reality on the ground.

FRA survey data also show that, for example, many Roma, Travellers, Muslims, Jews, and immigrants and their descendants cannot say with certainty whether they experience discrimination based on their race or ethnicity, as opposed to their religion or belief. FRA data also consistently show that many people across the EU experience multiple and intersectional discrimination, based on varied combinations of grounds. *Intersectional discrimination* is a situation in which several types of discrimination operate and interact with each other at the same time in such a way that they are inseparable and produce specific new types of discrimination.[49]

---

[47] The Racial Equality Directive prohibits direct and indirect discrimination on the grounds of racial or ethnic origin. Article 3 on the scope of the directive specifies that it applies to employment and occupation, vocational training, working conditions and membership of workers' or employers' organisations; social protection, including social security and healthcare; social advantages; education; and access to and supply of goods and services that are available to the public, including housing. The Employment Equality Directive prohibits direct and indirect discrimination on the grounds of religion or belief, disability, age or sexual orientation. Article 3 on the scope of the directive specifies that it applies to the areas of employment and occupation, vocational training, working conditions and membership of workers' or employers' organisations.

[48] FRA (2021), *Equality in the EU 20 years on from the initial implementation of the equality directives*. Luxembourg, Publications Office, p. 13. Following Kimberlé Crenshaw, the European Commission acknowledges "*that individuals can belong to several disadvantaged groups at the same time, and potentially suffer specific forms of discrimination.*" See as well the definitions of the European Institute for Gender Equality (EIGE) of 'multiple discrimination' at: https://eige.europa.eu/thesaurus/terms/1297 and 'intersectionality' at: https://eige.europa.eu/thesaurus/terms/1492.

[49] Ibid, p. 11.

Rationale

Objective, reliable and comparable data documenting experiences of inequality and discrimination are an essential tool for evidence-based policymaking. As already mentioned, proxies rarely offer a reliable picture of discrimination affecting racialised groups and ethnic minorities. Collecting data on experiences of discrimination of different groups, on different grounds and in different areas of life is one of the most effective ways to assess the effective implementation of anti-discrimination legislation. It is a powerful tool to monitor to what extent racial or ethnic biases might be used for example in access to employment, housing, education, health, and goods and services, as well as in encounters with law enforcement authorities.

"*Processes, such as attribution or ascription, othering, grouping and stereotyping create, institute or impose racial or ethnic origins. Racial or ethnic identity may not correspond to that perceived or assumed by a third party. Identity is fluid and may change over time. That, however, does not necessarily impact on how that person is perceived.*"[50] A certain level of grouping is unavoidable in statistical data collection and analysis, and can be necessary. However, data collectors and users of such grouping must be aware about the risks such grouping choices might entail.

Guidance

- → Any questions on experiences of discrimination should be preceded by a definition of discrimination. It is also essential that respondents can go back to this specific definition at any time during data collection. For this, data collectors need to ensure that respondents can access and understand the definition provided (be it through using a showcard or read-out-loud option in a face-to-face interview, or through info boxes in online surveys).

- → Standardise the question format: questions on racial/ethnic discrimination should follow the same logic as questions on discrimination based on other grounds.

- → Separate questions about experiences of discrimination in different areas of life (such as employment, education, health, housing, access to goods and services, access to justice etc.) should be preferred – instead of asking only one general question on discrimination, without specifying the areas.

---

[50] Farkas, L. (2017), The meaning of racial or ethnic origin in EU law: between stereotypes and identities, Luxembourg, Publication Office of the European Union, 2017, p. 37.

- Calculating the prevalence of discrimination for different areas of life also allows policymakers to develop and implement tailor-made anti-discrimination policies in their respective field. As well as the ability to target specific policy fields, asking separately about different areas of life can help respondents remember relevant incidents. Otherwise, when asking just a single question, some respondents might mainly concentrate on employment or education-related situations, but forget about housing or shops and services. In the area of crime victimisation, it has been shown that more specific questions lead to a higher overall prevalence that what would be achieved by asking fewer, more general questions.

- Questions about discriminatory experiences should preferably be asked for at least two different periods of time, enabling respondents to differentiate between short-term and long-term experiences (for example, in the 12 or 24 months, or in the five years preceding a survey).

  This procedure has been shown to increase the reliability and validity of the data, by allowing respondents to report on relevant experiences in different phases of their life.[51] In addition, answers to the questions about experience of discrimination in the short run can help to monitor the effectiveness of recently implemented measures. When defining the periods of time, data collectors might also consider the definition of certain indicators that are particularly relevant in this context, such as – for instance – indicator 10.3.1 of the 2030 Agenda for Sustainable Development (it seeks to determine the proportion of the population that reports having personally felt discriminated against or harassed within the previous 12 months). Furthermore, periods of time might also be chosen in a way that allows for comparisons with results of other surveys, such as the regular FRA surveys.

- To capture the intensity or recurrence of discrimination experiences, data collectors might ask respondents how often, in a particular period of time, they encountered racial or ethnic discrimination.

- Sequence and follow-up of questions: it is recommended to start the section on discrimination with a definition; attitudinal questions should precede questions on experiences; frequency and reactions (such as making a complaint, and if so, to whom).

---

[51] See for example: FRA (2014), *Violence against women: an EU-wide survey Main results*, Luxembourg: Publication Office of the European Union, (2015).

- To identify multiple and intersectional discrimination, the prevalence of discrimination based on racial or ethnic origin should be looked at in relation to and in combination with the prevalence of discrimination on other grounds. [52] Take due consideration of the Guidelines on improving the collection and use of equality data, and specifically Guideline No 6 “Ensure comprehensiveness of equality data”, which recommends that Member States mainstream incorporate the measurement of multiple and intersectional discrimination into all existing data collection systems; for example, data collected on disability should ideally include information on racial or ethnic origin, sex, age and other potential characteristics that might lead to a higher risk of discrimination based on multiple factors or intersecting inequalities. This type of analysis generally requires larger sample sizes for robust results, and a range of different sources could be considered, including large surveys, administrative and linked data.

- Member States are encouraged to incorporate the collection of racial/ethnic discrimination experiences into all EU and national surveys (such as the Labour Force Survey (LFS) and the EU Statistics on Living Conditions (EU-SILC)), by including discrimination questions/ modules. This could be done either by introducing individual variables/survey questions that specifically measure perceived racial/ethnic discrimination or by regularly implementing issue-based survey modules, to cover the range of issues related to racial or ethnic discrimination, bias-motivated harassment, racism and discriminatory profiling practices.

- Participation and response rates in a survey increase considerably “*when the topic of the study is of direct interest to the respondents, as with discrimination experiences.*”[53]

- Carry out background research before data collection: Before developing and deciding on any survey design, collecting background data and information on composition and concentration of the groups to be surveyed is an important task for every data collection. An informed decision on the best possible sampling design can only be taken after a comprehensive mapping of information on the target population in each country to be surveyed.[54]

- Promote the sharing of good practice in measuring discrimination experiences.

52/ Following Kimberlé Crenshaw, the European Commission acknowledges *that individuals can belong to several disadvantaged groups at the same time, and potentially suffer specific forms of discrimination.*” In; European Commission (2007), *Tackling Multiple Discrimination - Practices, policies and laws*, Luxembourg: Office for Official Publications of the European Commission, 2007, p. 15.

53/ FRA (2017), *Second European Union Minorities and Discrimination Survey - Technical Report*, Luxembourg: Publication Office of the European Union, 2017, p. 93.

54/ FRA (2017), *Second European Union Minorities and Discrimination Survey - Technical Report*, Luxembourg: Publication Office of the European Union, 2017, p. 92

Examples of national practice

The **Irish** module on Equality and Discrimination (for Q1 2019), asked questions on discrimination for key areas of life, including workplace discrimination, accessing and treatment in public and private services, education, housing, health, transport and contact with the police. To measure experience of discrimination in the workplace, the following questions were asked (in the same way across all settings):

DISCRIM_1 *"In the past two years, have you personally felt discriminated against in the workplace?"*

1. Yes
2. No
3. Not applicable (don't work, haven't been working in the past two years)
4. Don't know

IF DISCRIM_1=1 -> WHY

WHY "*Why do you think you were discriminated against - was it because of your?* (Multiple responses allowed)

1. Gender
2. Marital status
3. Family status (e.g. pregnant or with children or other dependants)
4. Religious belief
5. Sexual orientation
6. Age
7. Disability
8. Race (including colour, nationality, ethnic or national origins)
9. Membership of the Traveller community
10. Other (please specify)

Follow-up questions ask about the nature of discrimination, its frequency, the severity of its effect and whether any action has been taken in response (such as a complaint or a legal action).

In **Germany**, Each One Teach One (EOTO), a community-based organisation initiated the #Afrozensus in cooperation with Citizens For Europe (CFE), to examine the lives and discrimination experiences of people of African descent (including black, African and Afro-diasporic people in Germany). The project was funded by Germany's Federal Anti-Discrimination Agency (FADA). It is composed of an online survey and qualitative research (expert interviews and focus groups). The questionnaire in the online survey contains various modules. One of them is designed to capture, in detail, discrimination experiences on different grounds, in all relevant areas of life. The project is designed as a panel and EOTO is planning to repeat the online survey in future to measure anti-black racism over time.

In **Germany**, Rassismusmonitor - a new project of the DeZIM-Institut (the German Centre for Integration and Migration Research), funded by the Federal Ministry for Family Affairs, Senior Citizens, Women and Youth - aims to examine the causes, extent and consequences of racism in Germany. This project intends to provide an effective tool for monitoring racism by combining quantitative, qualitative, and experimental methods. In addition to representative population surveys and specific surveys of those affected by racism, legal, statistical, social-psychological, and other interdisciplinary approaches will be applied.

Norway conducted its first national population survey on quality of life in 2020, in which respondents were asked if they had experienced worse treatment than others (discrimination) in the last 12 months. The responses were distributed among respondents with *immigrant background, Norwegian-born with immigrant parents* and *the rest of the population*. The respondents were then asked to choose "reason(s) for being discriminated against" from the following list:

- Age
- Gender
- Disease/injury
- Disability
- Ethnic background
- Skin colour
- Religion/beliefs
- Political views
- Sexual identity

Examples from FRA surveys

Measuring discrimination in FRA's EU-MIDIS surveys: The EU MIDIS surveys and the follow-up surveys on Roma, Roma and Travellers and immigrants and descendants ask respondents if they had felt discriminated against on one or more grounds – skin colour, ethnic origin or immigrant background, religion or religious beliefs, sex, age, disability, sexual orientation, and 'other' grounds – in different domains and activities: when looking for work, at work, in education or when in contact with staff at their children's school, in access to healthcare, in connection with housing, and when using public or private services (such as public transport, administrative offices, when entering a night club, restaurant or a hotel, and when being in or entering a shop).

Respondents who reported discrimination on at least one of three specific grounds – skin colour, ethnic origin or immigrant background, and religion or religious beliefs – were asked further details about the incident, applying the generic term 'ethnic or immigrant background'. The generic term 'ethnic or immigrant background' indicates racial/ethnic discrimination as discussed in this guidance note.

Question wording and sequence

The survey includes a definition of discrimination, which is shown to respondents before they are asked about their actual experiences of discrimination: *"I would like to ask you a few questions about human rights. A basic right is to be treated equally. Still, some people might experience discrimination. By discrimination we mean when somebody is treated unfavourably compared with others because of their skin colour, age, sex, sexual orientation, disability, ethnic origin, religion or religious beliefs."*

Question wording for respondents who had been jobseekers in the five years preceding the survey: *"When looking for work in the past 5 years*[55]*, have you ever felt discriminated against for any of the following reasons? Tell me all that apply."*

1 Skin colour or racial origin

2 Ethnic or immigrant background

3 Religion or religious beliefs

4 Age (such as being too young or too old)

5 Sex/gender (such as being a man or a woman)

6 Disability

7 Sexual orientation (such as being gay, lesbian or bisexual)

---

[55] FRA's surveys ask about experiences of discrimination for two periods: 12 months and 5 years preceding the survey.

8 Gender identity or gender expression (this includes for example transgender, transvestite or non-binary people) - IF CAPI: INTERVIEWER: READ THE EXPLANATION / IF CASI/ONLINE: INFO BUTTON: For example, someone who was born as a boy but later feels like a girl/woman or born as a girl and later feels rather like a boy/man. Or someone who wears clothes that are usually designed for the opposite sex.

9 Other (please specify): OPEN TEXT BOX PLEASE SPECIFY THE REASON

10 I haven't felt discriminated against for any reason when in this situation

-96 - Prefer not to say

-97 - Don't understand the question

-99 - Don't know

Respondents who reported discrimination on at least one of the three specific grounds – (1) Skin colour or racial origin; (2) Ethnic or immigrant background; (3) Religion or religious beliefs – are asked further details about the incident, whether a complaint has been made and if so, to which institution/ body.

8. Provide sufficient budget for regular surveys collecting equality data on racial/ethnic origin, particularly when administrative data collection is not possible

Rationale

Governments need to ensure that equality data is collected regularly, to monitor compliance with obligations under EU law, even when administrative data collection is not possible. In these cases sufficient budget must be allocated at national/regional level for collecting equality data on racial/ethnic origin through regular national/regional surveys conducted by a competent body (i.e. one with the relevant research expertise and capacity).

Guidance

- → If there is no regular collection of equality data based on racial or ethnic origin through administrative sources or surveys, Member States could ensure **sufficient budget to collect such data through a comprehensive national survey**, every four to five years. Such a survey should:
  - ⇢ be representative of all policy relevant target groups for disaggregation based on racial or ethnic origin.
  - ⇢ aim to represent the diversity of communities/groups in a country that self-identify in terms of racial or ethnic origin.
  - ⇢ aim to also cover populations and groups that are defined as 'hard-to-reach', including people living in poverty. This would also imply translating the survey instrument (questionnaire + other fieldwork materials) in the languages relevant for these groups, as well as engaging community organisations, mediators, and peer interviewers.[56]
  - ⇢ be based on sample sizes that are sufficient to allow for more detailed data analysis to identify geographical disparities and (intersectional) inequalities. This means that the net sample size should be big enough to allow for further disaggregation of the data as regards geographical distribution, age, sex/gender, and education.
  - ⇢ be able to produce reliable and valid data for all policy-relevant target groups – that is, make use of up-to-date guides and standards on survey measurement and survey (sampling) design for hard-to-reach populations.

- → Regular surveys collecting equality data related to racial or ethnic origin can be conducted either as an independent survey with a specific focus on population groups that self-identify in terms of racial or ethnic origin, or as part of a general equality survey (that is, a survey that would also comprehensively cover other protected characteristics). This can be done at national (and, if relevant, regional) level, by a competent body, statistical institute or research organisation that has the relevant research capacity and expertise, following the previously listed guiding principles (including mapping and needs assessment exercises etc.).

---

[56] United Nations Economic Commission for Europe (UNECE) (2020*), Poverty Measurement – Guide to Data Disaggregation*, p. 70. .

Examples of national practice

In **Italy** a survey on “Integration of the second generation” (ISG) was carried out by Istat in 2015, co-financed by the Ministry of Interior and the EU’s European Fund for the Integration of third-country nationals (EFI).

In 2021, **Luxembourg** launched a national Survey on Experiences of Racism and ethno-racial Discrimination which will look at questions such as: How widespread is racism and ethno-racial discrimination in Luxembourg? Which groups of people are mainly targeted? In which contexts and situations do residents perceive discriminatory acts and treatment? Where and how can action be taken to combat these issues?

The survey is coordinated by the Ministry of Family, Integration and the Greater Region, implemented by the Luxembourg Institute of Socio-Economic Research (LISER), with the participation of the *Centre d’étude et de formation interculturelles et sociales* (CEFIS).

A questionnaire will be submitted to 15,000 randomly selected residents, with an over-representation of population groups at risk. The survey will be available in five languages - French, Luxembourgish, German, English and Portuguese - and will be conducted mainly online. It concerns both native Luxemburgish people and people with an immigrant background.

The questionnaire includes questions on the personal characteristics of the respondents, their position towards racism, their opinions on immigration and diversity, their perception of ethno-racial discrimination in Luxembourg, experienced as a witness or as a victim. The data collected in the survey will be used to identify the sectors in which specific actions are necessary. They will also be considered in the formulation of policy recommendations to combat racism and discrimination. These issues are perceived in several areas such as employment, housing, access to healthcare, education and social networks: racism (liser.lu).

In **Norway**, the Norwegian Directorate for Children, Youth and Family Affairs co-funded a national survey on quality of life, together with the Norwegian Directorate of Health in 2020. The survey consists of questions about quality of life, satisfaction with own physical and mental health, as well as optimism for the future, and experience of meaning and commitment in everyday life etc. The survey was repeated in 2021, and will hopefully be conducted again in the future, whether annually or less frequently.

FURTHER ETHICAL CONSIDERATIONS FOR COLLECTING DATA BASED ON RACIAL OR ETHNIC ORIGIN[57]

→ **All respondents must take part voluntarily.** It will be particularly important to ensure that target groups do not feel threatened by being approached to take part in the survey, coerced or obliged to participate in fear of negative consequences, and any involvement should not erode trust in the authorities. In addition, taking part in research can offer an unfamiliar experience to many in the target groups and this carries the risk of bringing up feelings of shame or fear. Therefore, building trust will be critically important.

→ **Use of community organisations, mediators, and peer interviewers.** While they can be helpful in promoting the survey given their standing and reach within their communities, it is important to ensure that target groups do not feel pressurised to take part and therefore reliance on religious leaders, for example, should be carefully considered. When it comes to peer-interviewing, it is worth noting that some of the communities can be small, and sometimes in an interview situation having someone around from the same ethnic group during an interview can lead to privacy concerns or unwillingness to answer sensitive questions truthfully.

→ Similarly, an **informed consent to take part and the right to not to answer a question** will be crucial. There is a **need for clear communication which explains the research in plain terms and comprehensible manner.** Respondents need to be offered an easy, clear way to opt out of a survey.

→ Creating a **safe environment where respondents can answer questions in their own pace, in privacy, without fear of being overheard by other household members or passers-by** builds trust, improves the respondent's experience, and ultimately results in better data.

→ **Interviewer matching** (including gender-matching, language matching or migration experience) can increase response rates as well as the acceptance amongst respondents and thus improve the research results.

→ **Interviewer training** should incorporate diversity and anti-racism modules, to minimise potential for discomfort when engaging with respondents.

→ The principles of avoiding harm should also be extended to cover the interviewers. As **interviewers visit homes** of people unknown to them during the project, they put themselves under a certain degree of risk and may meet respondents to whom the interview triggers traumatic memories.

57/ Based on FRA's experience in surveying hard-to-reach populations and (ethnic) minorities, as well as immigrants and descendants of immigrants.

→ Signposting to **support services** is important when dealing with the most vulnerable part of community, such as those who may have experienced oppression, violence, or sexual mistreatment and who might still be in danger.

→ In some countries, **the question on ethnic self-identification is seen as particularly sensitive and difficult.** It is recommended to warn respondents about it before it is asked, giving them the opportunity to skip it completely if they wish to do so.

→ Countries and data collectors are encouraged to support and promote **'data ownership'** among ethnic/racialised minorities, for example through cooperation with ethnic communities in data analysis, presentation and dissemination of results.

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

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

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

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).

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