



Guidance for navigating GDPR and equality data in higher education & research

September 2026, version 0.9.5



Disclaimer

This guidance is intended solely as an informative document and does not constitute legal advice for specific cases. It is the institution's responsibility to ensure that the processing of equality data complies with applicable laws and regulations, including the General Data Protection Regulation (GDPR), and that such processing is carried out in a lawful and responsible manner. For specific legal questions or advice, contact a legal expert.

Table of Contents



Introduction	3
1.1 Purpose	5
1.2 The relevance of inclusivity and data protection in higher education and research	7
Principles of Inclusivity & Privacy	15
2.1 Definition of inclusivity in higher education and research	15
2.2 The importance of privacy protection in an inclusive environment	15
2.3 Ethical considerations	16
2.4 Call to action	17
Legal Framework	18
3.1 The Racial Equality Directive as the European legal basis	18
3.2 Requirements for the processing of equality data in the Netherlands	19
3.3 Cultural Diversity Barometer	20
3.4 Processing of personal data in accordance with the GDPR and the Dutch GDPR Implementation Act (UAVG)	21
Inclusive data processing	33
4.1 Human rights framework for responsible data processing	33
4.2 Overview of use cases for the processing of equality data	35
4.3 Recommendations for improving inclusive data processing	37
4.4 Methodology: practical choices in questionnaire design	40
4.5 Obligation to ensure transparency following surveys or research	43
Conclusions & recommendations	47
5.1 Conclusions	47
5.2 Practical recommendations	48
Appendices	50
A Glossary	50
B. Relevant sources and literature	56

Introduction



To gain a reliable understanding of key aspects of inclusion, diversity, and equality in higher education and research in the Netherlands, processing so-called ‘equality data’ is of great importance.

In this guidance, we use the following definition of the term ‘equality data’: all qualitative and quantitative information that is useful for the purposes of describing and analysing the state of equality.¹

In practice, processing this information raises various objections, with frequent reference being made to the General Data Protection Regulation (GDPR). However, the GDPR states that the right to the protection of personal data is not absolute, ‘but must be considered in relation to its function in society and balanced against other fundamental rights, in accordance with the principle of proportionality’². One of these other fundamental rights is the right to non-discrimination. The right to non-discrimination is enshrined in Article 21(1) of the Charter of Fundamental Rights of the European Union³, which *‘prohibits any discrimination, in particular on grounds of sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or other opinion, membership of a national minority, property, birth, disability, age or sexual orientation’*.

Recital 4 of the GDPR also confirms that the right to the protection of personal data is not absolute and must be balanced against other fundamental rights in accordance with the principle of proportionality.

- 1 See p. 15: European Commission (2016): European handbook on equality data. 2016 revision. Directorate-General for Justice and Consumers. Directorate D - Equality. Online: https://ec.europa.eu/newsroom/just/document.cfm?action=display&doc_id=43205.
- 2 See: GDPR recital 4.
- 3 See: <https://eur-lex.europa.eu/legal-content/NL/TXT/HTML/?uri=CELEX:12012P/TXT>.



In reality, the GDPR sets out conditions under which equality data may be collected and processed, for example, with explicit consent or where substantial public interests are at stake. This may include, amongst other things, identifying and combating discrimination, as required by new regulations, including the EU Corporate Sustainability Reporting Directive (CSRD) from 2024.⁴

Race as a social construct: a legal and statistical concept

In this guidance, the term ‘race’ is understood as a social construct that fulfils both a legal and a statistical function, without thereby implying any biological reality. On this point, the guidance is consistent with the principle set out in Recital 6 of the Racial Equality Directive (Directive 2000/43/EC): ‘The European Union rejects theories which attempt to determine the existence of separate human races. The use of the word “racial origin” in this Directive does not imply an acceptance of such theories.’

‘Race’ and racial origin are social constructs: they reflect how people are categorised by others and how that categorisation influences their experiences and treatment. They bear no relation to humankind’s genetic diversity. The term is used because discrimination on the grounds of (perceived) racial or ethnic background is a social phenomenon that requires legal protection.

As a statistical concept, ‘race’ or racial origin therefore functions solely as a category to make inequality and discrimination visible and measurable, never as a biological classification. When data collection uses categories of origin, it does so with the same caveat. Statistics Netherlands (CBS) does not use a classification by race or ethnicity, but rather a statistical classification by origin, determined by the person’s and their parents’ country of birth. Since 2022, the dichotomy of ‘Western/non-Western migration background’ has been replaced by a geographical classification based on continents and common countries of origin, with the main categories being the Netherlands, Europe (excluding the Netherlands) and outside Europe.⁵

⁴ See: Directive – 2022/2464 – EN – CSRD Directive. See also: Moerel (2023). Workplace Discrimination and Equal Opportunity. Why monitoring cultural diversity in your European workforce is not at odds with the GDPR. Online: <https://fpf.org/blog/workplace-discrimination-and-equal-opportunity/>.

⁵ See: CBS (2022). New classification of population by origin. Replacing classification based on migration background and the concepts western/non-western. Online: <https://www.cbs.nl/en-gb/longread/statistische-trends/2022/new-classification-of-population-by-origin>.

This is also in line with Recital 51 of the GDPR: ‘Those personal data should include personal data revealing racial or ethnic origin, whereby the use of the term ‘racial origin’ in this Regulation does not imply an acceptance by the Union of theories which attempt to determine the existence of separate human races.’

1.1 Purpose

This guidance sets out how and under what conditions organisations can lawfully and responsibly collect and process equality data.

The purpose of processing equality data is to gain insights that can prevent and combat unequal treatment and discrimination. This is because, in cases of suspected discrimination, the burden of proof shifts to the employer⁶. Employers need this data to show that no unequal treatment has occurred.⁷

6 See in this regard: Section 10 of the General Equal Treatment Act (Awgb): ‘If a person who believes that discrimination has been or is being made to their detriment, as referred to in this Act, presents facts before a court that give rise to a presumption of such discrimination, the opposing party must prove that no act has been committed that contravenes this Act.’ Online: <https://wetten.overheid.nl/jci1.3:c:BWBR0006502&hoofdstuk=1¶graaf=5&artikel=10&z=2020-01-01&g=2020-01-01>. The Awgb (2020) has been aligned with the requirements of the Racial Equality Directive (Directive 2000/43/EC) – see: <https://eur-lex.europa.eu/legal-content/nl/TXT/?uri=CELEX:32000L0043>. [in Dutch]

7 See: Directive – 2022/2464 – EN – CSRD Directive. See also: Moerel (2023). Workplace Discrimination and Equal Opportunity. Why monitoring cultural diversity in your European workforce is not at odds with the GDPR. Online: <https://fpf.org/blog/workplace-discrimination-and-equal-opportunity/>.

Although there is no general legal obligation to collect equality data, the regulations do exert indirect pressure to do so. Article 10 of the General Equal Treatment Act (Awgb) stipulates that where an employee or student presents facts giving rise to a presumption of discrimination, the burden of proof shifts to the employer. [Art. 10 Awgb; see also Racial Equality Directive 2000/43/EC, Art. 8.] Those who have not collected data will find it difficult to defend themselves — and the GDPR offers no way out of this. [Moerel, L. (2023), Workplace Discrimination and Equal Opportunity, section 2.] Equality data is therefore not only a tool for combating discrimination, but also for demonstrably complying with equal treatment obligations. For larger educational institutions [The CSRD applies to organisations with more than 250 employees and a balance sheet total > €20 million or net turnover > €40 million. Most Dutch universities and universities of applied sciences fall within this scope.] the EU Corporate Sustainability Reporting Directive (CSRD) also applies, which requires public reporting on policies regarding equal treatment and the prevention of discrimination on grounds including gender and ethnic background. [Directive 2022/2464/EU (CSRD); ESRS S1. See also: Moerel (2023), section 3.] Without the internal collection and monitoring of equality data, it is not possible to comply with this reporting obligation. The CSRD imposes a



For example, when board members (higher education and research) wish to ensure balanced representation within the institution, they must be enabled to steer the organisation in a targeted manner. Structural barriers often underlie a lack of balance, which in turn can lead to discrimination and exclusion. In addition, personal experiences of discrimination, exclusion or inappropriate behaviour may also be a cause of imbalance – and form part of structural barriers. However, efforts to prevent unequal treatment and discrimination can only be effective if governing bodies clearly understand the current situation by mapping facts and figures and linking specific targets to them.

At present, this knowledge is often lacking. Organisations need data to base their policies on concrete insights and results.⁸

To this end, the European Commission published the ‘European Handbook on Equality Data’ in 2007, with a revised version published in 2016. In 2017, it stated: ‘The focus has turned to monitoring and, in turn, the need has arisen for data on (in)equalities based on racial and ethnic origin. Such data is essential to measure the level of implementation and monitor the impact of [anti-discrimination] policies’.⁹

This guidance covers equality data on a broader range of grounds on which unequal treatment occurs, in line with preliminary recommendation 7 of the State Commission against Discrimination and Racism.¹⁰

reporting obligation, but does not itself provide an independent exception to the prohibition on processing set out in Article 9 of the GDPR. The legal basis for processing special categories of personal data must be found in Article 9(2) of the GDPR and its implementation in Dutch law, in particular the exception relating to substantial public interest (Article 9(2)(g) of the GDPR in conjunction with Article 25 of the Dutch Data Protection Act) or explicit consent (Article 9(2)(a) of the GDPR). However, the reporting obligation under the CSRD may substantiate the substantial public interest underlying that exception.

- 8 This is also in line with the second objective of the National Action Plan for Greater Diversity and Inclusion in Higher Education and Research: “Better and broader monitoring of diversity in education and research, including social safety and inclusion”. See: <https://www.dihoo.nl/onderwerpen/n/nationaal-actieplan>. [in Dutch]
- 9 See: p. 4, European Commission (2017): Data collection in the field of ethnicity. Analysis and comparative review of equality data collection practices in the European Union. Online: https://commission.europa.eu/system/files/2021-09/data_collection_in_the_field_of_ethnicity.pdf Directorate-General for Justice and Consumers. Directorate D - Equality. Unit JUST/D1.
- 10 See: <https://www.staatscommissietegendiscriminatieenracisme.nl/documenten/2025/12/10/toekomstvisie---samen-voor-gelijkheid.-een-toekomstvisie-op-het-non-discriminatierecht>.

1.2 The relevance of inclusivity and data protection in higher education and research

Systemic barriers create inequalities within higher education and research. For example, we know that at many universities, half (or sometimes more) of the students and PhD candidates are women, yet men find it easier to progress to senior academic positions. Recent data show that 28.7 per cent of professors in the Netherlands are women.¹¹ Research among PhD graduates indicates that men's theses are awarded the 'cum laude' distinction far more often than those of women. We also know that people from migrant backgrounds face barriers within the education system. This begins as early as primary school, where they are often given lower school recommendations and are more frequently under-recommended.¹² They also face discrimination more frequently when applying for work placements.¹³

Furthermore, recent research among students at Erasmus University Rotterdam shows that women and non-binary students score lower on life satisfaction and other well-being indicators than men.¹⁴ The same applies to students with disabilities (compared with students without disabilities), students who are part of the LGBTIQ+ community (compared with cisgender and heterosexual students) and first-generation students (compared with students whose parents have a university education).

We are learning more and more about inequalities within higher education and research. For example, much is already known about inequalities based on sex and gender, not only regarding under- or over-representation but also, for instance, regarding pay gaps.¹⁵ Regarding diversity dimensions such as non-European background, we are also gaining increasing insights into representation among staff¹⁶

¹¹ See: <https://www.lnvh.nl/a-4071/monitor-vrouwelijke-hoogleraren-2024>. [in Dutch]

¹² See: <https://longreads.cbs.nl/integratie-en-samenleven-2022/schoolloopbanen-van-de-tweede-generatie/#::~:~:text=This%20also%20applies%20to%20the,et%20al.%2C%202021>. However, please also note this clarification: <https://decorrespondent.nl/15358/goed-nieuws-kinderen-met-een-migratieachtergrond-krijgen-niet-structureel-een-te-laag-schooladvies/633e2550-4a59-0827-0074-343d032ab5b9>. [in Dutch]

¹³ See: <https://www.kis.nl/themas/stagediscriminatie>. [in Dutch]

¹⁴ See: Student Wellbeing Programme & IDEA (2024), Student Wellbeing Monitor 2023 (Wave 4).

¹⁵ See: <https://www.lnvh.nl/a-1270/lnvh-rapport-financiele-belonging-mv-in-de-wetenschap>. [in Dutch]

¹⁶ See: <https://www.eur.nl/en/news/results-barometer-cultural-diversity-statistics-netherlands-cbs-published>.

and students¹⁷ with the help of Statistics Netherlands (CBS). However, there is still a lot we do not know. What about the representation of other, visible and invisible dimensions of diversity, such as sexual orientation, disability, religion or skin colour? Is there, for example, also a pay gap based on migration background? And what are the experiences of all these groups at university? Do they feel at home? Can they be themselves?

Everyone should have equal opportunities to study and work at a university of applied sciences or a university, and feel included in the process. At the same time, diversity brings a variety of perspectives, greater creativity, and innovation. It is therefore vital to monitor diversity indicators and incorporate them into (policy) research. This will enable us to gain greater insight and answer whether, and in what ways, inequalities exist within our institutions. And if so, how we can then reduce these inequalities. We can also assess whether our interventions to combat discrimination and exclusion and to promote inclusion are actually achieving the desired results. This aligns with the second objective of the National Action Plan for Greater Diversity and Inclusion in Higher Education and Research: *better and broader monitoring of diversity in education and research, including social safety and inclusion*.¹⁸

This is not only important for the institution or for people who identify as belonging to under-represented groups. The ultimate aim of monitoring and research is also to uncover structural barriers and inequalities. Policies can then be developed to tackle these barriers and move towards an inclusive organisation. This benefits people from both under-represented and over-represented groups.

In summary, this guidance aims to:

- provide insight into the preconditions and safeguards that an organisation must observe to ensure the responsible processing of equality data. This enables:
- the more effective (evidence-based; derived from the processing of equality data) removal of structural barriers, based on insights into existing structural barriers and how they operate;

¹⁷ See: <https://www.eur.nl/en/news/research-student-inflow-origin-several-dutch-universities>.

¹⁸ See: OCW (2020). Nationaal actieplan voor meer diversiteit en inclusie in het hoger onderwijs en onderzoek. Online: <https://open.overheid.nl/documenten/ronl-3d4db279-2786-44c8-a1af-5337604eebfd/pdf>.

- the identification of good practices and of the knowledge gaps that limit the effectiveness of current interventions relating to aspects of inequality within educational institutions;
- promoting transparency in policies relating to diversity and inclusion, thereby also facilitating benchmarking between institutions;
- providing institutions with a framework for action, and clarifying roles and responsibilities regarding data collection and research, both within the institution and across the sector.

1.3 Target audience

This guidance is aimed at all key stakeholders within knowledge institutions who can influence policy on processing equality data. The target audience includes the following roles: Chief Diversity Officers, Equality, Diversity and Inclusion Officers, HRM staff, Privacy Officers, Privacy Lawyers, Data Protection Officers (DPOs), Chief Information Security Officers (CISOs), Security Officers, Chief Data Officers, Data Stewards and policy researchers.

Furthermore, the guidance is intended for all parties that have signed the National Action Plan on Diversity in Higher Education and Research¹⁹: ECHO, KNAW, LanDO, LNVH, NWO, OCW, PNN, Postdoc NL, UNL and VH.

The guidance is accessible to non-legal readers, making it applicable beyond a purely legal perspective.

1.4 An overview of common objections to the processing of equality data

Uncertainty about processing equality data leads, in practice, to hesitation in collecting and processing it. For example, at the time of writing this guidance, it is unclear within higher education and research whether the processing of equality data is permitted, whether there is a legal basis for it, what data may be requested and/or used, and whether it is actually possible to ensure that the data cannot be traced back to individuals (anonymisation). See the table at the end of this chapter to illustrate this lack of clarity.

In addition to these legal considerations, institutions are sometimes reluctant to change the status quo and free up budget to actually use the findings to promote diversity, equality and inclusion through policy and resources. Other reasons for hesitation

¹⁹ See: OCW (2020). Nationaal actieplan voor meer diversiteit en inclusie in het hoger onderwijs en onderzoek. Online: <https://open.overheid.nl/documenten/ronl-3d4db279-2786-44c8-a1af-5337604eebfd/pdf>.

include the perception that staff and students already face too many questionnaires, or that the scope of the annual Employee Satisfaction Survey (ESS) is too limited to include valid questions and constructs.²⁰ In addition, reluctance – particularly among data protection officers – and caution on the part of HR can act as barriers.²¹

The selection and clustering of countries of origin and reasons for migration by Statistics Netherlands are also not always meaningful.²² For example, Indonesia is now a third-generation immigration country, and because of the threshold (parents' country, not grandparents'), this yields few significant correlations with outcomes. It is therefore preferable to remove migration motives from the Statistics Netherlands data.²³ At the same time, the assumption is incorrect that, after two generations, there are no significant correlations between migration motive and/or origin and outcomes such as career success.

Finally, those concerned (staff and students) may, for various reasons, decide not to consent to the processing of equality data, even if it is based on self-registration. This may be, for example, because they do not expect that providing this information to the institution will actually lead to more equal treatment for them²⁴, or because they find answering questions about ethnicity and racialisation burdensome and confrontational.²⁵ Another objection raised by those concerned may be that their identity is not limited to a single dimension of diversity, and that it is unclear whether the institution's methodology sufficiently embeds intersectionality.²⁶

20 See: LanDO (diversity officers and policy officers) (2025), Problemen bij uitvragen diversite-itsdata. [in Dutch]

21 See: LanDO (diversity officers and policy officers) (2025), Problemen bij uitvragen diversite-itsdata. [in Dutch]

22 To categorise reasons for migration, Statistics Netherlands (CBS) uses Labour, Education, Asylum and Family (= other): <https://www.cbs.nl/en-gb/our-services/methods/surveys/brief-survey-description/migration-motives-statistic>.

23 <https://www.cbs.nl/en-gb/our-services/methods/surveys/brief-survey-description/migration-motives-statistic>.

24 See: De Jonge Akademie (2022A). Een verkenning van visies van staf en studenten van kleur op antidiscriminatie. Online: <https://www.dejongeakademie.nl/publicaties/2300805.aspx?t=Antidiscriminatiepraktijken-wereldwijd-en-visies-van-studenten-en-staf-van-keur>. [in Dutch]

25 See: De Jonge Akademie (2022A). Een verkenning van visies van staf en studenten van kleur op antidiscriminatie. Online: <https://www.dejongeakademie.nl/publicaties/2300805.aspx?t=Antidiscriminatiepraktijken-wereldwijd-en-visies-van-studenten-en-staf-van-keur>. [in Dutch]

26 See Appendix A, Glossary



Below are some common objections, along with suggestions for addressing them.

Objection

'We cannot or are not allowed to collect diversity data.'

Concern

Many people believe that, following the introduction of the GDPR in 2018, personal data may not be requested or recorded. Even privacy officers within some organisations are reluctant to do so.

Addressing the concern

However, it is both permitted and possible, provided you meet several conditions. You must have a clear purpose for requesting or retaining this personal data; recording it is not an end in itself. This purpose – such as improving policy on diversity, equality and inclusion (DE&I) – must also be communicated to the people whose personal data is being collected or retained. Furthermore, it must be explained exactly what happens to the data collected.

Objection

'The way in which categorisation is carried out is inappropriate.'

Concern

Categorisations are often not inclusive, not appropriate, or do not correspond to people's experiences. In the case of the 'migration background' category, for example, there has been a great deal of criticism of the use of the terms 'Western' and 'non-Western'.

Some people, for instance, prefer not to use the term 'functional impairment' but would rather use 'disability', or vice versa.

It is also worth considering people's experiences: what if, for example, someone has been diagnosed with 'autism' but does not personally experience it as a limitation? In this way, you can ask yourself what the correct categorisation is for each dimension of diversity. Often, there is no single, clear-cut answer.



Addressing the concern

There are various ways to deal with this:

At the very least, provide a disclaimer and explain that the organisation is aware that social identities are more complex than the labels provided, but that these are used for specific reasons – for example, to enable comparability with other datasets. It is also important to involve the target group in research design and work together to find an appropriate categorisation.

Objection

‘What about me?’

Concern

One reason for collecting equality data is to amend existing policies or develop new ones. This often focuses on a single group. This can provoke resistance among other (under-represented) groups, as they may ask: why aren’t they receiving the same attention? People from over-represented groups may also feel disadvantaged: do they have to give something up, or are they missing out on something because of the ‘new’ policy?

Addressing the concern

Clear advance communication is crucial here: (good) policy aims to reduce barriers within the system – which (likely) affect more groups than just those studied – and to improve inclusion.

Objection

‘Putting people into boxes only serves to reinforce stigma.’

Concern

Diversity research often relies on certain assumptions: for example, we investigate our students’ ethnic backgrounds because we assume that, as a group, they face barriers. But some students will always experience things differently. Moreover, general patterns can look very different within a specific context, such as a university. Furthermore, results can reinforce stereotypes: for example, if we see an under-representation of women in professorial posts, some people might interpret this as: ‘See, women simply prefer care roles and aren’t suited to leadership’, whilst others might interpret it as: ‘See, there is a lack of equal opportunities’.



Addressing the concern

This objection highlights how important it is never to present data without context and, for example, to conduct qualitative follow-up research into the causes of certain quantitative patterns.

Objection

‘My identity is not limited to a single dimension of diversity.’

Concern

We all have multiple social identities, but much diversity research focuses on just one dimension of diversity – for example, gender alone, or migration background or skin colour alone. This gives a very limited picture of the situation. A white woman may often have different experiences from a woman of colour, so if we look only at gender, we miss out on a lot of important information.

Addressing the concern

Take intersections into account wherever possible. For example, in the barometer, you can choose to include gender and age category as well. Qualitative research often allows more scope for an intersectional approach.

Objection

‘Will my data really remain anonymous?’

Concern

Many people worry whether data can be traced back to them. For example, if you combine several variables (such as department, gender, age, nationality and ethnic background), it is fairly easy to work out who the data relates to. Even with self-identification, the problem can arise that the data is easily traceable. There are also concerns that data may be used for the wrong purposes.

Addressing the concern

Educational institutions put in place appropriate safeguards and communicate these proactively and in clear language:



Minimum cell size in reporting

Results are never reported if a subgroup consists of fewer than a specified minimum number of respondents (typically $n > 5$ or $n > 10$, depending on the sensitivity). This prevents combinations of characteristics — such as department, gender and ethnic background — from being traceable to individuals.

Generalisation of data

Detailed values are replaced by broader categories: date of birth becomes age group; a specific department becomes a faculty. The less detailed the reporting, the lower the traceability risk.

Aggregation prior to analysis

Raw individual data is aggregated into group statistics as soon as possible. Analyses are carried out exclusively at an aggregated level; individual profiles are not stored or shared.

Separation of data

The equality dataset is stored separately from other staff or student records and is accessible only to a limited group of authorised staff (see section 3.4.2.6 on access restrictions).

No linking of datasets

Equality data is not linked to other files (such as HR data or academic results) without additional safeguards, as combining datasets increases the risk of re-identification of individuals.

Principles of Inclusivity & Privacy



2.1 Definition of inclusivity in higher education and research

In our context, we define inclusivity as a safe learning and working environment in which everyone feels at home and can flourish, and where a diversity of perspectives is engaged with and taken into account.²⁷

Inclusion is not an end in itself, but a prerequisite for a future-proof, high-quality education and research sector that makes the most of talent and diversity. The National Action Plan emphasises the importance of an integrated approach and collaboration: only by working together can educational institutions and the government build an inclusive sector that responds to global challenges.

Inclusion ensures that people from different backgrounds and with diverse perspectives can participate fully, feel safe and develop their potential in higher education and research. Greater diversity among students and staff, and in the content of education and research, can lead to better learning outcomes and innovation.

2.2 The importance of privacy protection in an inclusive environment

The National Action Plan emphasises that privacy protection in an inclusive environment is of great importance in enabling people to share personal characteristics – such as gender, cultural background, disability and sexual orientation – voluntarily and safely, without fear of exclusion, discrimination or misuse of that data.

27 See: Ministry of Education, Culture and Science (2020): Nationaal actieplan voor meer diversiteit en inclusie in het hoger onderwijs en onderzoek. Online: <https://open.overheid.nl/documenten/ronl-3d4db279-2786-44c8-a1af-5337604eebfd/pdf>. [in Dutch]



Privacy is therefore a prerequisite for reaping the benefits of a safe learning and working environment. We highlight four of these benefits:

- **Quality improvement:** An inclusive and diverse learning and working environment fosters the highest possible quality of education and research.
- **Future-proofing:** Inclusion keeps the Dutch sector internationally competitive and enables it to contribute to major societal issues, such as sustainability, climate and well-being.
- **Social safety:** Ensuring inclusion is a prerequisite for a socially safe environment, in which discrimination and exclusion are actively combated.
- **Equal access:** Inclusion helps to reduce inequalities and ensures that everyone has equal opportunities to develop, regardless of their background.

This guidance encourages institutions to collect and process equality data based on self-identification, self-categorisation, and self-registration in a secure, privacy-friendly environment. One way to achieve this is to ask staff or students at the start of an academic year which data may be used for surveys carried out later that year. In doing so, an easy-to-understand description of each survey's purpose and practical details is provided proactively. The staff member or student can then specify, for each study, which data may or may not be used. This allows them to retain control and oversight of their data, in line with data protection principles.

2.3 Ethical considerations

Ethical considerations underpin safeguarding inclusion and protecting privacy in higher education and research. These focus on respect for equality, justice, and individual autonomy, and on combating exclusion, discrimination, and bias.

This involves the following ethical principles:

- i Equal treatment and inclusion:** Everyone must be afforded equal access and opportunities, regardless of background, gender, ethnicity, health or beliefs. Diversity and inclusion are seen as a necessary foundation for a genuine meritocracy, in which talent is central without prejudice or exclusion.
- ii Respect for privacy and autonomy:** Individual privacy must be fully respected. The collection and use of sensitive data relating to identity and background may occur only voluntarily and under secure conditions, so that people retain control over their own information.



- iii A safe and equitable environment:** Active efforts must be made to create an environment in which discrimination and inappropriate behaviour are prevented and combated. A 'sense of belonging' is crucial to ensuring that everyone feels safe and valued.
- iv Transparency and accountability:** Policy decisions on diversity, inclusion and privacy must be transparent. Institutions must be accountable for how they safeguard these values.

2.4 Call to action

This guidance is in line with the call to action set out in the 2020 National Action Plan for Greater Diversity and Inclusion in Higher Education and Research²⁸, which calls on all education and research institutions to:

- make inclusion the norm in policy, culture and daily practice, with a structural focus on diversity, equality and social safety;
- actively safeguard privacy in all data registration and monitoring processes, so that participation is safe and voluntary and the GDPR is strictly adhered to;
- promote leadership and dialogue, whereby managers, lecturers and researchers receive training in inclusion, diversity and privacy, and actively engage in dialogue with students and staff on ethical dilemmas;
- to encourage continuous reflection on policy and practice, so that blind spots are identified and addressed, and ethical values remain alive within the organisation.

28 See: Ministry of Education, Culture and Science (2020): Nationaal actieplan voor meer diversiteit en inclusie in het hoger onderwijs en onderzoek. Online: <https://open.overheid.nl/documenten/ronl-3d4db279-2786-44c8-a1af-5337604eebfd/pdf>. [in Dutch]



3.1 The Racial Equality Directive as the European legal basis

Within this Guidance for navigating GDPR and equality data in higher education and research, the diversity dimension of ‘ethnic background’ is frequently cited as an example. This is because the guidance builds on how the legal framework for processing equality data in the Netherlands is rooted in European law. The Racial Equality Directive (Directive 2000/43/EC) constitutes the central legal source in this regard.²⁹ The Racial Equality Directive aims to create a framework for combating discrimination on the grounds of ‘race’ or ethnic background. It obliges Member States to apply the principle of equal treatment. As such, as a directive, it applies in principle to all dimensions of diversity on which discrimination may occur. Its scope is broad: it provides protection not only in employment, but also in social protection, social advantages, education, and access to goods and services.³⁰ This means that the obligation to ensure equal treatment applies to educational institutions in both their capacity as employers and as educational institutions.

The Directive defines two forms of prohibited discrimination, which are also included as such in the glossary of this guidance:

Direct discrimination occurs when a person is treated less favourably on grounds of ‘race’ or ethnic origin than another person in a comparable situation is, has been or would be treated.³¹

Indirect discrimination occurs when an apparently neutral provision, criterion or practice places persons of a particular ‘race’ or ethnic background at a particular disadvantage compared with other persons, unless a legitimate aim objectively justifies that provision and the means are appropriate and necessary.³² Statistics play a crucial role in proving indirect discrimination:

29 Council Directive 2000/43/EC article 1.

30 Directive 2000/43/EC article 3(1).

31 Directive 2000/43/EC article 2(2)(a) and (b).

32 Directive 2000/43/EC article 2(2)(a) and (b).



they can highlight the adverse effects on a specific group resulting from an apparently neutral measure.³³

The Directive also provides for the reversal of the burden of proof: where a person presents facts giving rise to a presumption of discrimination, the defendant - the institution - must prove that the principle of equal treatment has not been breached.³⁴ It is precisely this that makes the collection of equality data not only desirable but, in fact, necessary: without data, an institution will find it difficult to defend itself against an allegation of unequal treatment (see also section 1.1).

The GDPR and the Dutch GDPR Implementation Act (UAVG) apply to the processing of personal data necessary to comply with these obligations. They impose conditions on such processing but do not constitute an obstacle. On the contrary, they provide the legal basis under which equality data may be lawfully processed (see sections 3.2 and 3.3).

3.2 Requirements for the processing of equality data in the Netherlands³⁵

The Netherlands has no specific legal obligation for employers to collect equality data from employees. Nor is there a general prohibition on the collection of such data. However, employers may lawfully process this data for monitoring purposes relating to diversity, equality and inclusion (DE&I) based on two different legal grounds, in order of legal hierarchy:

- 1** Substantial public interest under Dutch law;
- 2** Explicit consent from staff and students.

Organisations with a works council must ensure that the works council approves any policy relating to the processing of equality data.

33 Directive 2000/43/EC and Directive 2000/78/EC, COM (2014).

34 Directive 2000/43/EC article 8(1).

35 See: Moerel (2023). Workplace Discrimination and Equal Opportunity. Why monitoring cultural diversity in your European workforce is not at odds with the GDPR. Online: <https://fpf.org/blog/workplace-discrimination-and-equal-opportunity/>.



3.2.1 Substantial public interest

The conditions set out in Article 9(2) of the GDPR³⁶ have been implemented in the Netherlands via the Dutch GDPR Implementation Act (UAVG). Article 25 of the UAVG stipulates that data relating to ethnic background (limited to country of birth and the country of birth of parents or grandparents) may be processed based on a substantial public interest, provided that:

- the processing is necessary to redress the disadvantage faced by a minority group;
- the data subject has not objected to the processing.

To make use of this legal basis, the employer must, amongst other things:

- demonstrate that certain groups are in a disadvantaged position;
- implement a broader policy aimed at redressing this disadvantage;
- demonstrate that the processing of data relating to diversity dimensions such as ‘race’ and ethnicity is necessary for the implementation of this policy.

3.2.2 Explicit consent

Employers may collect and process data on ethnic background and other diversity dimensions on the legal basis of explicit employee consent and voluntary reporting. The conditions for explicit consent, as set out in Article 9 of the GDPR³⁷, also apply in the Netherlands. This means, amongst other things, that employees must be free to refuse consent without suffering any adverse consequences.

3.3 Cultural Diversity Barometer

Dutch employers with more than 250 employees can make use of the Cultural Diversity Barometer from Statistics Netherlands (CBS). This tool provides employers with an anonymous overview of the cultural diversity within their organisation. Employers provide certain non-sensitive personal data to CBS, which then reports on, for example, the percentage of employees with:

- 1 a Dutch background;
- 2 a background within Europe (excluding the Netherlands);
- 3 a background outside Europe.

36 GDPR article 9(2), exceptions to the prohibition on processing. See: <https://eur-lex.europa.eu/legal-content/NL/TXT/PDF/?uri=CELEX:32016R0679#page=38>. [in Dutch]

37 GDPR article 9, the processing of special categories of personal data. See: <https://eur-lex.europa.eu/legal-content/NL/TXT/PDF/?uri=CELEX:32016R0679#page=38>. [in Dutch]



CBS can provide this information for the entire organisation or for specific departments, provided these have more than 250 employees.

This approach enables employers to monitor diversity without infringing individual privacy, in line with GDPR data protection rules.

3.4 Processing of personal data in accordance with the GDPR and the Dutch GDPR Implementation Act (UAVG)

The GDPR and the Dutch GDPR Implementation Act (UAVG) apply to the fully or partially automated processing of personal data, as well as to the processing of personal data contained in a file or intended to be included therein.³⁸ The processing of personal data encompasses any operation carried out on the data, including non-automated processing, such as providing data vocally.³⁹ Within the Netherlands, the GDPR applies to the processing of personal data by organisations established here, as well as to the activities of establishments of data controllers or data processors within the EU.⁴⁰

All information that can be directly or indirectly traced back to an identified or identifiable individual falls under the term ‘personal data’.⁴¹ The GDPR distinguishes between ordinary and special categories of personal data:

- **Ordinary personal data:** Data such as name, address, place of residence (name and address details), gender, date of birth and email address.
- **Special categories of personal data:** Sensitive data such as health, religion, ethnic origin, ‘race’ or sexual orientation. This data receives extra protection under the GDPR due to the risk of discrimination or exclusion.⁴² In addition, certain data is classified as sensitive, such as financial data, location data and criminal records.

³⁸ GDPR article 2.

³⁹ GDPR article 4(2) and the judgment of the Court of Justice of 7 March 2024, ECLI:EU:C:2024:216, C-740/22, paragraph 39.

⁴⁰ GDPR article 3.

⁴¹ GDPR article 4(1).

⁴² GDPR article 9(1).



3.4.1 Anonymisation and pseudonymisation

Personal data may be anonymised or pseudonymised:

- In the case of **anonymised data**, it is reasonably impossible to identify individuals. Anonymous data does not fall within the definition of personal data and is not subject to the GDPR and the UAVG.⁴³
- In the case of **pseudonymised data**, it is not immediately clear which individuals the data relates to; however, if the data can still be traced back to natural persons when additional data is used, it still qualifies as personal data and is therefore subject to the GDPR and the UAVG.⁴⁴

In the context of equality data, aggregated data may still be traceable to individuals if the source data used for the aggregation still exists. Such data therefore does not qualify as anonymous data. This is also the case where a combination of data could lead to traceability. The GDPR and the Dutch GDPR Implementation Act then apply, and this equality data must therefore be protected appropriately, so as not to jeopardise the rights and freedoms of the individuals to whom this data relates.

Figure 1 on the following page⁴⁵ illustrates, at *data level*, how personal data (*not pseudonymised*) can be transformed, through various stages of increasing pseudonymisation, into anonymous data (*anonymised*) that can no longer be traced back to the individual concerned.

⁴³ GDPR recital 26.

⁴⁴ Pseudonymised data falls outside the scope of the GDPR only if the data can no longer be traced back to an individual, even with additional data. See also GDPR recital 26.

⁴⁵ Source: Hrudey et al (2019). Risk Management for Research Data about people. A general matrix to be used by data stewards and research supporters for the assessment of privacy risks with research data and determination of appropriate methods for risk management. Zenodo. <https://doi.org/10.5281/zenodo.3584333>



Figure 1 – From personal data to anonymised data

PS0	PS1	PS2	PS3	ANON
Not pseudonymised	Pseudonymised Level 1	Pseudonymised Level 2	Pseudonymised Level 3	Anonymised
Name Rutger Hauer				
Patient number 09210	Patient number 09210	Study subject 470009	Study subject 470009	
E-mail blade.runner@batman.nl				
Postcode 8911AA	Postcode 8911AA	Region Friesland	Country The Netherlands	Country The Netherlands
City Leeuwarden				
Date of Birth 27-04-1967	Date of Birth 27-04-1967	Year of Birth 1967	Age 51-60	Age 51-60
Income 7,861	Income 7,861	Income 7,500-10,000	Income 5,000-15,000	Income 5,000-15,000
Job Judge	Job Judge	Job Legal	Job Legal	Job Legal
Car DeLorean	Car DeLorean	Car DeLorean	Car Sports Car	Car Sports Car
License plate SN-09-HN	License plate SN-09-HN			

The steps applied at the various levels of pseudonymisation involve techniques such as generalisation, randomisation and deletion. These techniques are explained in detail in the book *Privacyontwerpstrategieën (Het Blauwe Boekje)* by Jaap-Henk Hoepman.⁴⁶

⁴⁶ Hoepman (2020). Privacyontwerpstrategieën. Het Blauwe Boekje. See: <https://www.cs.ru.nl/~jhh/publications/pds-boekje.pdf>. [in Dutch]



3.4.2 Privacy Principles

Where the GDPR applies, processing equality data must demonstrably comply⁴⁷ with the principles of transparency, fairness, and lawfulness.⁴⁸ Process data only for a legitimate purpose⁴⁹ and only to the extent necessary.⁵⁰ Equality data must be accurate⁵¹ and stored on a need-to-know basis.⁵² In addition, take appropriate technical and organisational measures to protect the data adequately.⁵³ These principles are explained below.

3.4.2.1 Transparency, fairness and lawfulness

Transparency & fairness

Basic principle: Transparency and fairness mean the educational institution must clearly explain how and why it collects and processes equality data. Fostering a culture of openness and accountability requires actively encouraging questions and discussion. In an environment where transparency and trust are central, people feel safe expressing themselves. Transparency builds trust and helps ensure data processing is perceived as fair and respectful. Furthermore, processing must be carried out properly and solely for legitimate purposes.

Action: Draw up a detailed privacy statement or information leaflet explaining how equality data is processed. Involve those concerned to ensure the information is understandable and relevant, strengthening trust in the organisation. This can also contribute to a culture of openness and accountability, where those concerned feel safe sharing their equality data.

Lawfulness

Basic principle: Lawfulness means that equality data is processed lawfully. Article 6 of the GDPR sets out six legal bases for processing. Additional requirements apply to processing special categories of personal data, among

47 GDPR article 5(2).

48 GDPR article 5(1)(a).

49 GDPR article 5(1)(b).

50 GDPR article 5(1)(c).

51 GDPR article 5(1)(d).

52 GDPR article 5(1)(e).

53 GDPR article 5(1)(f).

other things. When processing equality data, the questions asked do not normally relate solely to gender, age, or background; several questions are asked at the same time.

Action: Check whether a valid legal basis exists for processing.

We can explain this as follows. The additional legal basis for processing that may be applicable, provided it is carefully structured, is consent under Article 6(1)(a) of the GDPR. As this concerns special categories of personal data, the requirement for explicit consent under Article 9(2)(a) of the GDPR must also be met.

Further processing of equality data under Article 6(4) of the GDPR is not appropriate. That provision permits reusing data for a new purpose compatible with the original purpose. However, for data relating to ‘race’ and ethnicity, this compatibility test is not sufficient: as these are special categories of personal data, a separate ground for exception under Article 9(2) of the GDPR is always required for any processing, including further use.

Article 6(1) of the GDPR (processing necessary to comply with a legal obligation) also does not provide a basis: no legal obligation requires processing this combination of data.

3.4.2.2. Consent

The legal basis of consent must meet the following requirements to be valid:

(i) freely given, (ii) informed, (iii) unambiguous, (iv) active, (v) specific and (vi) demonstrable. When processing special categories of personal data such as ‘race’, ethnicity or religious beliefs, an additional requirement applies: (vii) consent must be given explicitly. These conditions comprise specific elements which are set out in detail below.

i Freely given

Basic principle: Consent must be given without any form of coercion or pressure. Data subjects must have the choice to share their data or not, without facing any negative consequences. Particularly in a hierarchical relationship, such as between an employee and an employer or between a student and an educational institution, there is a relationship of dependency in which the freedom to make a genuine choice may come under pressure. In this situation, it is therefore even more important that the data subject, from whom consent is sought, genuinely feels that they have a sincere choice, free from any form of



coercion. There must be no negative consequences if the data subject does not make the choice or indicates that they do not wish to participate. Furthermore, a realistic alternative must be available.

Practical application

- **Communication:** Communicate clearly and transparently that participation is voluntary and will not result in any negative consequences. Both taking part in a survey and answering specific questions must be voluntary. Offer options such as ‘I choose not to participate’ and ‘I choose not to answer’. Respondents must be able to skip questions, and it must be clear that this freedom exists, which underlines the voluntary nature of participation. A respondent’s willingness to take part in the questionnaire does not mean they also wish to answer every question. A question may not be relevant, or the respondent may not feel comfortable answering it. Therefore, there must be scope to skip questions, and it must be clear that this freedom exists. This emphasises the voluntary nature of participation.
- **Self-identification:** It is essential to enable self-identification. Participants’ trust is undermined if they do not recognise themselves in predefined categories. Effective self-identification means all identity categories are relevant, open-ended responses (including ‘no answer’) are permitted, and participants can identify with multiple identities. Trust is enhanced when the research is carefully designed and has the potential to deliver genuine insight and meaningful change.
- **The method of collecting equality data:** Ensure that the collection of equality data is organised on an organisation-wide basis, rather than by an individual manager. This safeguards the independence and objectivity of the data collection and reduces the pressure on participants when providing sensitive information. Furthermore, those involved should not be individually urged to respond if response rates are low, to safeguard the voluntary nature and integrity of the data collection. This presents a dilemma: the organisation must carefully balance the need to encourage responses against the risk of under-representation in the data.
- **Feedback mechanism:** Implement a mechanism that enables participants to provide feedback on the consent process, so that any concerns about coercion or pressure can be addressed.



ii Informed

Basic principle: Data subjects must be fully aware of what their consent entails. This means that they must be clearly informed in advance about what data is being collected, why it is being collected, how it will be used, how long it will be retained and with whom it may be shared.

Practical application

- **Purpose of the data collection:** Clearly explain why the organisation collects equality data, for example to promote diversity and inclusion within the organisation and to gain an insight into the composition of the workforce.
- **Use of data:** Inform data subjects about how the collected data will be used. For example, state that aggregated data will be used to produce reports, develop policy, or monitor progress towards DE&I objectives.
- **Sharing of data:** Specify with whom the data may be shared, including both the source data from which the aggregated data is derived, the aggregated data itself, and the results of the analysis of the source data. Data may be shared with internal departments such as HR, as well as with external parties such as research agencies. Ensure that those concerned are informed that their data will be treated securely and confidentially.

iii Unambiguous

Basic principle: Consent must be clear; there must be no doubt that the data subject has given their consent. This means that the data subject must make a clear and informed choice. Silence, pre-ticked boxes or inactivity cannot be regarded as valid consent.

Practical application: Word the consent request in such a way that there is only one possible interpretation. Avoid double negatives ('I do not want my data not to be used') and compound questions in which multiple issues are grouped into a single box. Ensure that the data subject knows exactly what their 'yes' applies to: link each consent request to a clearly defined, comprehensible processing activity. Use clear language appropriate to the target audience, without legal jargon that obscures meaning. Separate the consent request visually and textually from other information or terms and conditions, so that consent cannot be inferred from accepting, for example, general terms and conditions. Cases of doubt, such as an incompletely filled-in form or an unclear answer, cannot be regarded as consent.



iv Active

Basic principle: The data subject must take an active step to give consent, such as ticking a box or signing a form. Passive consent is not permitted.

Practical application: Set up the consent mechanism so that the default setting is 'no consent'. Checkboxes are left blank by default, and response options are not pre-selected. The data subject must take an active step to give consent: ticking a box, selecting a button or signing a form. In an online questionnaire, this means a separate consent step before the substantive questions, with an empty box that the data subject ticks themselves. Clicking through, scrolling down, or the passage of time does not constitute valid consent. Record which action was performed at what point in time, so that the active consent can be traced retrospectively. Furthermore, make withdrawing consent just as easy as giving it.

v Specific

Basic principle: Specificity means that consent must be clearly defined for the various processing activities. This implies that, if the data is used for multiple purposes, the data subject must give explicit consent separately for each purpose. The data subject must know exactly how the data will be used.

Practical application: Break consent down into separate purposes and offer the data subject a choice for each purpose. If the organisation wants to use equality data for both internal reporting and external publications, each purpose should have its own tick box so the data subject can consent to one and, if they wish, refuse the other. Avoid blanket consent where various processing activities are grouped under a single 'yes'. Describe each purpose in concrete and recognisable terms, without vague wording such as 'for research and policy purposes', which does not provide the data subject with sufficient clarity as to what consent entails. Limit processing to the purposes for which consent has been given; data collected with consent for diversity monitoring must not be reused for another purpose without new consent. If a purpose changes or is expanded, request consent again for the new purpose.

vi Demonstrable

Basic principle: Although the GDPR does not stipulate that consent must be given in writing, freely given consent must be demonstrated. This makes it necessary to record the consent.⁵⁴

Practical application: Ensure that all consents given are carefully documented, including the date on which consent was given, the specific purposes for which consent was given, and the manner in which consent was obtained. For example, this can be done by logging digital consents or retaining signed consent forms.

vii Explicit consent (when processing special categories of personal data)

Basic principle: In addition to the above conditions that consent must satisfy as a legal basis for the processing of personal data, a prohibition on processing applies in principle to the processing of special categories of personal data, including ‘race’⁵⁵, ethnicity, religious beliefs, sexual orientation and health. This prohibition does not apply where an exception under Article 9(2) of the GDPR applies, in conjunction with specific legislation, such as the Dutch GDPR Implementation Act. These exceptions matter because, when requesting equality data, information is often sought not only on gender but also on background, religion, or a work-related disability. Two exceptions that are relevant in the context of equality data are (i) reasons of substantial public interest under Article 9(2)(g) of the GDPR and (ii) consent under Article 9(2)(a) of the GDPR. For processing based on consent, the requirements described above apply, plus the requirement that this must be done ‘explicitly’. Explicit consent is a stricter form of consent. The term ‘explicit’ refers to the manner in which consent is given, meaning that the data subject, by performing an action, explicitly indicates their agreement to the collection and use of personal data.

Practical application: Make the specific nature of the data clear in the consent request and explicitly state which category is involved, for example, data relating to ethnic background, religious beliefs or a disability. Request consent through an explicit statement, for example, a tick box with the text ‘I expressly consent to the processing of this data relating to my [category] for [purpose]’, rather than a generic

⁵⁴ GDPR article 7(1) and GDPR recital 42.

⁵⁵ The term ‘race’ has the same meaning as in article 1 of the Constitution and must be interpreted broadly. It also encompasses skin colour and national or ethnic background.

pre-ticked box. Keep this consent separate from the standard consent for non-special categories of personal data, so that the data subject makes a separate, informed choice regarding the sensitive data. In addition to the usual details (date, purpose, method), also document that this involved explicit consent for special categories of personal data and what statement the data subject made. Check in advance whether consent is indeed the appropriate legal basis in this context. In a relationship of dependency such as that between employer and employee, the voluntary nature of consent is compromised (see criterion i), meaning that the legal basis of ‘substantial public interest’ under Article 9(2)(g) of the GDPR, in conjunction with the Dutch GDPR Implementation Act, is sometimes more appropriate.

3.4.2.3 Purpose limitation and data minimisation

The principles of purpose limitation and data minimisation mean that personal data may be processed only for specific, explicitly defined, and legitimate purposes. This means that, at every stage of the processing, it must be clear which personal data is required for which purpose. Furthermore, personal data obtained on the basis of consent must not be processed for other purposes. In short, you may collect and process no more data than is strictly necessary.

3.4.2.4 Accuracy

In accordance with Article 5.1(d) of the GDPR, only accurate personal data should be processed, and it should be kept accurate to prevent misuse. There is a risk, however, that incorrect and/or out-of-date information may create a misleading impression. This also implies limited storage.

3.4.2.5 Storage limitation

Under Article 5(1)(e) of the GDPR, personal data must not be retained longer than strictly necessary. The organisation must set out this duration internally. Distinguish between raw data and aggregated data. Once the retention period expires, the organisation must delete the personal data unless it has explicitly stated interests that justify retaining it longer.

3.4.2.6 Appropriate technical and organisational measures

Under Article 5(1)(f) of the GDPR, personal data must be secured through appropriate technical and organisational measures. This includes, amongst other things:



Encryption: Ensure that all communications involving the transmission of equality data, such as emails, are encrypted. This prevents unauthorised parties from accessing the data during transmission.

Secure storage: Store equality data in secure environments, such as encrypted databases or secure servers. This protects the data against unauthorised access and potential data breaches.

Access restrictions: Restrict access to the raw data to a select group of staff who actually require this data for their work. This can be achieved through Role-based Access Control (RBAC), where only authorised individuals can access specific datasets.

Logging: Implement detailed logging mechanisms to record all access to and operations on equality data. This helps monitor who accessed the data and what actions were taken, which is essential for detecting unauthorised access or misuse.

Rapid aggregation: To minimise the exposure of raw data, equality data must be aggregated as quickly as possible. This involves combining individual data points into anonymised or pseudonymised statistics that cannot be traced back to individual persons. Rapid aggregation reduces the risk of identifiability and better protects data subjects' privacy.

3.4.3 Processing special categories of personal data on grounds of substantial public interest

Exceptions to the prohibition on processing special categories of personal data may be permitted on grounds of substantial public interest. The Netherlands has used this option in Articles 25 to 30 of the Dutch GDPR Implementation Act (UAVG), with each special category having its own provision and conditions. For data revealing 'race' or ethnic origin, Article 25 of the UAVG is the relevant provision. This article waives the prohibition on processing, provided that such data is processed to eliminate or reduce actual disadvantages linked to the 'race' or ethnic origin of members of ethnic or cultural minority groups (affirmative action). Separate regimes apply to other diversity characteristics: religious or philosophical beliefs (Article 27 of the UAVG) and data relating to health, including a disability (Article 30 of the UAVG); each has its own, more limited grounds for exemption. Where no appropriate statutory exemption exists, explicit consent (Article 9(2)(a) of the GDPR) remains the appropriate legal basis. However, the exemption applies only subject to the following three conditions:



- 1 Necessity for the purpose of the processing:** Processing is only permitted to the extent that it is necessary for the purpose of the processing, namely the implementation of the affirmative action policy. The affirmative action policy is only justified as long as a particular minority group remains in a disadvantaged position. If the affirmative action policy changes, the processing of data relating to cultural diversity must also be reconsidered.
- 2 Restriction to specific data:** Processing must be limited to data concerning a person's country of birth, the country of birth of their parents or their grandparents. This means that data other than that which reveals 'race' or ethnic background may not be processed under such a preferential policy. During the parliamentary history of the predecessor to Article 25 of the UAVG, namely Article 18 of the Personal Data Protection Act (Wbp), it was emphasised that even the data subject's consent to process additional data cannot override the prohibition. The conditions set out in Article 18 of the former Wbp and Article 25 of the UAVG are intended to be exhaustive, leaving no scope for the grounds for consent under Article 23 of the former Wbp and Article 9(2)(a) of the GDPR.⁵⁶
- 3 Right to object:** The processing of data revealing a person's 'race' or ethnic origin in the context of affirmative action policies is not permitted if the data subject objects to this in writing. This applies regardless of whether such an objection would also have effect under Article 21 of the GDPR, which governs the right to object.

Given these specific conditions, the collection of equality data cannot be carried out under the substantial public interest exception. This is also because the collection was not initially intended to introduce preferential treatment, but to identify diversity figures and determine what follow-up steps are required. Furthermore, this data collection focuses not only on the place of residence of parents or grandparents, but also on other aspects, such as religious beliefs, ethnicity, 'race', sexual orientation and health.

56 See, in this context, Parliamentary Papers II 1997/98, 25 892, No. 3, p. 106. [in Dutch]

Inclusive data processing →

This guidance follows the human rights-based approach to personal data set out by the Office of the United Nations High Commissioner for Human Rights (OHCHR).⁵⁷ The same applies to the principles governing the processing of equality data relating to participation, data disaggregation, self-identification, transparency, privacy and accountability. This guidance does not address general methodological principles.

4.1 Human rights framework for responsible data processing

Equality data processing must be embedded within a human rights framework. This guidance therefore follows the OHCHR's approach to data.⁵⁸ This framework identifies six principles which, taken together, form the framework for responsible equality data processing. These six principles are set out below and explicitly linked to the recommendations in sections 4 and 5 of this guidance.

- 1 Stakeholder participation** - Staff and students must be actively involved in all stages of the data collection process: from defining the research question and selecting categories, to interpreting results and formulating follow-up policies. Participation enhances data quality and credibility and strengthens trust in the institution (see section 4.3, recommendations 1 and 2).

57 See p. 9: High Level Group on Non-discrimination, Equality and Diversity, Subgroup on equality data (2020). Guidance note on the collection and use of equality data based on racial or ethnic origin. Directorate-General for Justice and Consumers. Directorate D – Equal Opportunities Equality and Union citizenship. Online: https://commission.europa.eu/system/files/2022-02/guidance_note_on_the_collection_and_use_of_equality_data_based_on_racial_or_ethnic_origin_final.pdf

58 OHCHR (2018). A Human Rights-Based Approach to Data - Leaving No One Behind in the 2030 Agenda for Sustainable Development. Guidance note to data collection and disaggregation. Online: <https://www.ohchr.org/Documents/Issues/HRIndicators/GuidanceNoteonApproachtoData.pdf> The six principles are also reaffirmed in: High Level Group on Non-discrimination (2021). Equality and Diversity, Subgroup on equality data, Guidance note on the collection and use of equality data based on racial or ethnic origin, p. 8 (Preamble), and Moerel (2023), section 4.



- 2 Data disaggregation** - Data must be disaggregated according to the relevant diversity dimensions, so that inequalities affecting individual groups become visible and are not obscured by averages for the total population. This is also a prerequisite for meaningful intersectional analyses (see section 4.4, Methodology).
- 3 Self-identification** - Data subjects themselves determine whether, how and to what extent they wish to share personal data about their background. Categories are not imposed externally, and questionnaires always offer the option to skip questions, tick multiple options and provide free-form responses. This principle is elaborated further in the glossary (Appendix A) and in section 3.4.2.2 on consent.
- 4 Transparency** - Data subjects must be fully informed about the purpose, use, retention period and recipients of their data, in understandable language and before consent is sought. Transparency is also an obligation under the GDPR (see section 3.4.2.1) and a prerequisite for fulfilling the transparency obligation following research (see section 4.6).
- 5 Privacy** - Personal data is protected by appropriate technical and organisational measures, including encryption, access restrictions and rapid aggregation of raw data. Privacy is not only a legal obligation under the GDPR, but also a prerequisite for the trust of data subjects that is necessary for them to participate in research voluntarily and in good faith (see section 3.4.2.6 and section 4.7 on the personal data vault).
- 6 Accountability** - Institutions are accountable for how equality data is collected, used and secured, and for the policy measures taken based on the findings. Transparency about both the results and how they translate into action is essential to this principle (see section 4.5 on the transparency obligation following surveys or studies).

Together, these six principles form the benchmark for every specific data collection activity described in sections 4.1 to 4.6 of this guidance. They also form the basis for the safeguards that organisations must put in place when collecting equality data based on explicit consent (see section 3.2.2 of this guidance), and are consistent with the requirements arising from the Racial Equality Directive and the GDPR (see chapter 3).



4.2 Overview of use cases for the processing of equality data

We identify several common use cases⁵⁹ for the processing of equality data in higher education and research in the Netherlands.

4.2.1 ‘For’ the institution

4.2.1.1. Diversity of new staff

To understand the diversity of applicants and those subsequently recruited, the institution collects applicants’ background details anonymously and securely.

4.2.2 ‘Within’ the institution (two ‘types’, at both staff and student level)

4.2.2.1 Monitoring the diversity of staff and/or students

The aim is to obtain insights into representation within the population (of the Netherlands or the region). Within this representation, the aim may then be to understand, at group level, aspects such as intake, progression, and attrition.

4.2.2.2 Employee Satisfaction Survey

An Employee Satisfaction Survey provides an institution with valuable insights. The data collected is used to make strategic decisions based on facts rather than assumptions, and to retain talent:

- 1 Insight into staff behaviour and needs:** It helps universities and universities of applied sciences understand what motivates staff and the challenges they face.
- 2 Improving satisfaction and engagement:** By identifying and addressing issues in the workplace, universities and universities of applied sciences can develop action plans to improve staff satisfaction and the quality of teaching and research. This involves setting measurable targets, assigning responsibilities and planning the proposed activities. These plans are often developed at team or departmental level to tackle specific issues. Measures such as targeted training programmes, adjustments to terms and conditions of employment or improvements in leadership can also be implemented.

⁵⁹ We can either (re)use existing data or collect new data. In this context, the Cultural Diversity Barometer is, in effect, the use case for the (re)use of existing data.

By involving staff in formulating the questions, determining the types of answers, interpreting the results and devising solutions, the quality – and therefore also the reliability – of the outcomes is improved. It also fosters a culture of open communication and collaboration. This strengthens staff engagement and motivation. Furthermore, regularly repeating the surveys helps evaluate the effectiveness of implemented strategies. This supports continuous improvement and helps identify new trends or problems on time. Finally, results can be compared with those of other organisations within the same sector (benchmarking). This helps measure performance against similar institutions and identify best practices.

4.2.2.3 Student Wellbeing Monitor

In addition to Employee Satisfaction Surveys, many higher education institutions conduct research into well-being and into the extent to which students feel at home at the institution. Equality data is increasingly being incorporated into this research. One example is the biennial Student Wellbeing Monitor at Erasmus University Rotterdam: a comprehensive survey that, in addition to wellbeing, also focuses on experiences of inclusion and discrimination, and covers a wide range of diversity dimensions (including ethnic background, sexual orientation, and parents' educational attainment). Beyond general reporting, attention increasingly focuses on statistically significant differences between groups. Similar to the Employee Satisfaction Survey, such research aims to gain insights that help in making strategic decisions to prevent student dropout:

- 1 Gaining an understanding of students' wellbeing and needs.** This helps universities and universities of applied sciences understand how well students are doing overall and within specific subgroups.
- 2 Improving wellbeing.** By identifying problems (both for the student body as a whole and for specific subgroups), universities and universities of applied sciences can develop targeted interventions to improve the situation.

Here too, it is important to involve the target group and to repeat the survey regularly. Furthermore, this survey also offers the possibility of benchmarking.

4.2.2.4 Exit surveys

The results of exit surveys can help institutions improve their staff policies and working culture:

1 Insight into reasons for leaving

Exit surveys provide insight into the positive and negative factors influencing staff decisions to leave the institution, such as terms and conditions of employment, workload, management or career prospects. Exit surveys among students provide insight into reasons for dropout.

2 Improving staff retention

By analysing reasons for leaving, institutions can develop strategies to reduce staff turnover and retain talent.

3 Optimising work culture and management

Feedback from departing staff can reveal issues in organisational culture or leadership styles, enabling targeted improvements.

4 Retaining institutional knowledge

Exit surveys offer an opportunity to capture valuable knowledge and experiences from departing staff, which is useful for succession planning and continuity.

4.3 Recommendations for improving inclusive data processing

The objectives of the use cases mentioned above relating to the processing of equality data are best served by ensuring that staff and/or students are properly represented in the surveys. It is also important to provide the correct context to prevent stigmatisation. Responsible processing of equality data requires several measures. The key recommendations⁶⁰ in this regard are:

1 Indicator development

Develop indicators that measure the extent to which equality data is regulated, valid, reliable and usable. Apply these indicators to all relevant grounds, such as ethnicity, religion, age, disability, and sexual orientation. Self-identification plays an important role. This can prevent assumptions about specific indicators, such as religion based on origin or field of study based on disability.

60 See: European Commission (2017): Data collection in the field of ethnicity. Analysis and comparative review of equality data collection practices in the European Union. Online: https://commission.europa.eu/system/files/2021-09/data_collection_in_the_field_of_ethnicity.pdf Directorate-General for Justice and Consumers. Directorate D - Equality. Unit JUST/D1.



2 Harmonisation of definitions and categories

Harmonising definitions, classifications, and categorisations related to ethnic background and other diversity characteristics is essential (see 4.3.1 and 4.3.2).

3 The necessity for a needs analysis

A needs analysis among potential users of data disaggregated by ethnic background helps to determine what data is required to tackle the negative effects of inequality in a targeted manner.

4 Use of different data sources

Using a variety of data sources can compile a broad range of equality data.

Additional recommendations based on best practice⁶¹:

5 Equality

Equality is a fundamental requirement when conducting research. Only then can participants contribute meaningfully. Therefore, ensure that materials and procedures are accessible and understandable to everyone.

6 Preparation and aftercare

Taking part in research can affect respondents. For example, it may raise expectations or bring negative experiences to the fore. Thorough preparation is therefore essential, including plans for aftercare and identifying who will be available to provide it. Good aftercare also involves providing feedback on what ultimately happens with the results.

4.3.1 Harmonisation of categories as a prerequisite for comparability

One aim of this guidance is to enable benchmarking between institutions (see section 1.2). However, this aim is achievable only if institutions use the same or mutually comparable categories when collecting equality data. At present, such harmonisation is largely lacking, both in the Netherlands and across Europe. Because institutions use different definitions, classifications, and categories, their survey results are not comparable.⁶²

⁶¹ See: Pharos, Veelgestelde vragen over inclusief onderzoek. Online: <https://www.pharos.nl/veelgestelde-vragen-over-inclusief-onderzoek/>. [in Dutch]

⁶² EC Guidelines (2018), p. 10 (gap 5): “Because different data sources often rely on diverging definitions and use various population categorisations, equality data are often not comparable, both within and across Member States. In the absence of a mechanism to coordinate

The EC Guidelines (2018) identify this as one of the main bottlenecks in current equality data collection practice: without a mechanism to coordinate how different grounds for discrimination are conceptualised and measured, data and findings will almost certainly be incomparable.⁶³ The recommendation is therefore to harmonise definitions, classifications and categorisations - preferably at sectoral level - so that higher education and research institutions can meaningfully compare their findings with one another.⁶⁴ For the dimensions of ethnicity and ‘race’, the choice of categories deserves particular attention. The transition by Statistics Netherlands (CBS) from the categories ‘Western’/‘non-Western’ to a classification based on regions of origin is relevant in this regard; institutions that align with the current CBS classification increase the possibility of linking their data to national statistics via the Cultural Diversity Barometer.⁶⁵

4.3.2 Comparability and local validity

A recent example of a tool that meets the need for comparability and harmonisation is the so-called ‘Diversity Minimal Item Set’ (DiMIS).⁶⁶ The DiMIS comprises nine diversity dimensions on which unequal treatment occurs: gender & sex, age, socio-economic status, care work, sexual identity, ethnic-racial identity, religion & world view, mental health, and physical health & disability.⁶⁷ The DiMIS has been deliberately designed as a minimum set: it is concise, widely applicable, and provides the basis for intersectional analyses by collecting data across multiple diversity dimensions simultaneously. Using such an item set as a common framework for questionnaires and monitoring tools enables comparisons between institutions

decisions on how to conceptualise and measure different grounds of discrimination and on how to categorise potentially affected persons or groups, data and outcomes will therefore most certainly suffer from incomparability.”

63 EC Guidelines (2018), p. 10 (gap 5).

64 European Commission (2018). Guidelines on improving the collection and use of equality data. Guideline 11 “Improve comparability of equality data”, p. 17-18.

65 See: Statistics Netherlands (2022). New classification of population by origin. Replacing classification based on migration background and the concepts western/non-western. Online: <https://www.cbs.nl/en-gb/longread/statistische-trends/2022/new-classification-of-population-by-origin>. See also section 3.2 of this guidance on the Cultural Diversity Barometer.

66 Stadler et al. (2023). Diversified innovations in the health sciences: Proposal for a Diversity Minimal Item Set (DiMIS), Sustainable Chemistry and Pharmacy, vol. 33. Online: <https://doi.org/10.1016/j.scp.2023.101072>

67 Gültzow et al. (2026). The Dutch version of the Diversity Minimal Item Set (DiMIS-NL). OSF. <https://doi.org/10.17605/OSF.IO/HWDPJ>.

and increases the likelihood that findings will be recognisable and reliable for all stakeholders.

However, flexibility is also important so that categories can be tailored locally to the target group and context. This ensures local validity, but may hinder comparability. The choice between local validity and comparability should depend on the research aim. If, for example, the aim is to understand how people of non-European origin are represented in senior positions within Dutch universities, it is obviously important to use the same categorisations. If the aim is to gain insight into experiences of inclusion and exclusion among different groups within the organisation, tailor the choice and formulation of categories as closely as possible to the target group(s) to increase their recognisability.

For all dimensions of diversity, the target group must be involved in selecting and formulating categories (see also section 4.3, recommendation 1 on indicator development and self-identification). Only in this way can categories be developed that are both statistically comparable and socially credible.

4.4 Methodology: practical choices in questionnaire design

When setting up equality data research, institutions face a series of methodological choices that determine both data quality and the trust those involved place in the research. Based on international comparative research into questionnaire methods at universities⁶⁸ and an exploration of the views of staff and students of colour at Dutch universities⁶⁹, the most relevant choices are outlined below.

Self-categorisation as a starting point

Virtually all respondents in the De Jonge Akademie study (2022) prefer self-categorisation - whereby staff and students are asked to categorise themselves. They prefer this to third-party categorisation based on administrative data, such as

⁶⁸ DJA (2022B). Dataverzameling ter bestrijding van discriminatie aan universiteiten: een verkenning van buiten Nederland toegepaste vragenlijstmethoden. Online: <https://www.dejongeakademie.nl/publicaties/2300805.aspx>. [in Dutch]

⁶⁹ DJA (2022A). Een verkenning van visies van staf en studenten van kleur op antidiscriminatiegegevens. Online: <https://www.dejongeakademie.nl/publicaties/2300805.aspx?t=Antidiscriminatiegegevens-praktijken-wereldwijd-en-visies-van-studenten-en-staf-van-kleur>. [in Dutch]

population statistics.⁷⁰ This principle aligns with the human rights standard of self-identification (see section 4.1) and enhances the validity and credibility of the data. In practical terms, self-categorisation means that questionnaires always offer the following options:

- ticking multiple boxes to indicate multiple identities;
- skipping questions or refusing to answer, without this having any consequences;
- entering free text when the categories provided do not apply;
- amending the information provided at a later date.⁷¹

Choice of category: geographical origin over administrative divisions

An important methodological choice concerns the categories used for the various dimensions of diversity. Research among staff and students of colour at Dutch universities shows, for example, that a large majority prefer categories based on geographical origin and family history over administrative classifications such as the former Statistics Netherlands (CBS) categories ‘Western’/‘non-Western’, which CBS has since phased out.⁷² The reasoning is that a category based on geographical origin better reflects one’s own identity.⁷³

For institutions wishing to use the CBS Cultural Diversity Barometer, this means that the new CBS classification based on regions of origin (Europe excluding the Netherlands, outside Europe) provides a better basis than the old ‘Western’/‘non-Western’ dichotomy, and enables comparison with national statistics (see section 3.3).

Recognising complexity

Categories in questionnaires can never fully capture the complexity of social identities. It is therefore important to explicitly acknowledge this to those involved.

70 DJA (2022A). Een verkenning van visies van staf en studenten van kleur op antidiscriminatiegegevens. Online: <https://www.dejongeakademie.nl/publicaties/2300805.aspx?t=Antidiscriminatiegegevens-praktijken-wereldwijd-en-visies-van-studenten-en-staf-van-kleur>. [in Dutch]

71 DJA (2022B). Dataverzameling ter bestrijding van discriminatie aan universiteiten: een verkenning van buiten Nederland toegepaste vragenlijstmethoden. Online: <https://www.dejongeakademie.nl/publicaties/2300805.aspx>. [in Dutch]

72 DJA (2022A). Een verkenning van visies van staf en studenten van kleur op antidiscriminatiegegevens. Online: <https://www.dejongeakademie.nl/publicaties/2300805.aspx?t=Antidiscriminatiegegevens-praktijken-wereldwijd-en-visies-van-studenten-en-staf-van-kleur>. [in Dutch]

73 DJA (2022A). Een verkenning van visies van staf en studenten van kleur op antidiscriminatiegegevens. Online: <https://www.dejongeakademie.nl/publicaties/2300805.aspx?t=Antidiscriminatiegegevens-praktijken-wereldwijd-en-visies-van-studenten-en-staf-van-kleur>. [in Dutch]

A disclaimer accompanying the questionnaire, explaining why certain categories are used and what their limitations are, increases willingness to participate and prevents respondents from feeling misunderstood.⁷⁴ A tried-and-tested solution is a statement indicating that more possible identifications exist than can be included in the questionnaire, and explaining why the chosen categories were used nonetheless, for example to enable comparison with other datasets.

Choice of terminology and open-ended response options

Political and academic considerations influence terminology choices.⁷⁵ The term ‘migration background’ is not suitable for measuring experiences of discrimination, as a person’s place of birth or their parents’ origin does not determine whether they will face ethnic discrimination. The terms ‘Western’/‘non-Western’ lack scientific basis and have now also been abandoned by Statistics Netherlands. For organisations collecting equality data to combat discrimination, terms such as ‘ethnicity’, ‘racial or ethnic background’ or ‘geographical origin’ are preferable, although even here no uncontroversial choice is possible. Response options such as ‘other, namely ...’ are also unsuitable, as they have an exclusionary effect. A better option is ‘none of the above, namely ...’. Involve the target group in choosing the wording (see also section 1.1, the first principle on participation).

Care for respondents

Taking part in research on diversity and inclusion can be stressful and confrontational for those involved.⁷⁶ A carefully designed study accounts for this by planning aftercare in advance, communicating clearly about what will happen to the results, and providing feedback on the findings to those involved (see also section 4.5 on the obligation to ensure transparency following surveys).

- 74** DJA (2022B). Dataverzameling ter bestrijding van discriminatie aan universiteiten: een verkenning van buiten Nederland toegepaste vragenlijstmethoden. Online: <https://www.dejongeakademie.nl/publicaties/2300805.aspx>. [in Dutch]
- 75** DJA (2022A). Een verkenning van visies van staf en studenten van kleur op antidiscriminatie-data. Online: <https://www.dejongeakademie.nl/publicaties/2300805.aspx?t=Antidiscriminatie-data-praktijken-wereldwijd-en-visies-van-studenten-en-staf-van-kleur>. [in Dutch]
- 76** DJA (2022A). Een verkenning van visies van staf en studenten van kleur op antidiscriminatie-data. Online: <https://www.dejongeakademie.nl/publicaties/2300805.aspx?t=Antidiscriminatie-data-praktijken-wereldwijd-en-visies-van-studenten-en-staf-van-kleur>. [in Dutch]

4.5 Obligation to ensure transparency following surveys or research

Universities and universities of applied sciences have an important obligation to ensure transparency towards their staff and students after conducting research such as Employee Satisfaction Surveys. This obligation involves sharing the results and the resulting actions in a fair, accessible and comprehensible manner.

1 Sharing results

- **Staff:** Universities and universities of applied sciences are obliged to provide clear feedback to staff on the outcomes of research such as an Employee Satisfaction Survey. This can be done through formal consultation with staff representative bodies and via internal communication channels such as emails, the intranet or presentations at meetings. In doing so, provide a clear context and interpretation of the data.
- **Students:** Although students are not directly involved in Employee Satisfaction Surveys, transparency regarding general organisational improvements may be indirectly relevant, for example, if these affect teaching processes or facilities.

2 Accessibility and comprehensibility

- The results must be communicated in plain language, avoiding jargon. This increases the engagement and trust of staff and students.
- The use of visual aids such as graphs or infographics can help to make complex data more accessible.

3 Accountability

- Universities and universities of applied sciences must be accountable for how the results are used to improve policies and strategies. This can be included in annual reports or policy plans, as required by the Higher Education and Research Act.
- Transparency also helps to build trust among external stakeholders, such as accreditation bodies and government agencies.

4 Communicating action plans

- In addition to presenting results, it is essential to share concrete action plans arising from the findings. This demonstrates that feedback is taken seriously and leads to improvements.
- Monitoring progress and providing regular updates on actions reinforces credibility.

5 Benchmarking and external transparency

- Universities can compare their results with those of other institutions (benchmarking) to evaluate performance and identify best practices.
- Transparency towards external parties, such as the Ministry of Education, Culture and Science or collaborative partners, can be provided through reports or public documents.

4.6 Equality data and the paradox of algorithmic bias

Educational institutions are increasingly using automated systems in their decision-making processes. Examples include recruitment selection tools, ranking algorithms for course allocation, and performance-monitoring systems. Such systems present a paradox directly relevant to the issue of equality data: the less diversity data is available, the harder it is to detect whether an algorithm is producing discriminatory outcomes.

Self-learning algorithms do not discriminate because of code errors, but because the training data they rely on contains bias. Only when organisations know which individuals belong to vulnerable groups can they make data bias visible and train algorithms correctly.⁷⁷ The lack of equality data therefore offers no protection against algorithmic discrimination; instead, it prevents its detection. As Moerel (2023) puts it: ‘Race-blind is not race-neutral, and unawareness does not equal fairness.’⁷⁸

This insight also has legal implications. The EU AI Act (Regulation 2024/1689), which came into force in 2024, explicitly obliges providers of high-risk AI systems to examine their training, validation and test data for potential biases that could lead to prohibited discrimination, and to take corrective measures.⁷⁹ To this end, providers may — to the extent strictly necessary — also process special categories of personal

77 See: Moerel (2023). Workplace Discrimination and Equal Opportunity. Why monitoring cultural diversity in your European workforce is not at odds with the GDPR. Online: <https://fpf.org/blog/workplace-discrimination-and-equal-opportunity/>.

78 See: Moerel (2023). Workplace Discrimination and Equal Opportunity. Why monitoring cultural diversity in your European workforce is not at odds with the GDPR. Online: <https://fpf.org/blog/workplace-discrimination-and-equal-opportunity/>.

79 Regulation (EU) 2024/1689 (AI Act) article 10.2(f) and (g).



data, including data relating to ethnic background.⁸⁰ The EU AI Act thus confirms that the processing of equality data for bias detection is expressly permitted. Furthermore, the AI Act obliges providers of high-risk AI systems to examine their datasets for bias (Article 10.2). In this context, collecting and processing equality data are necessary to comply with this obligation to investigate.

For educational institutions that use automated systems — in recruitment, assessment or monitoring — this means, in practical terms, that the absence of equality data poses a risk: not only in terms of non-discrimination, but also in terms of GDPR compliance and compliance with the AI Act. In this context, collecting equality data is not merely a policy choice, but a necessary condition for the responsible use of data-driven systems.

4.7 Personal data vault

The concept of a ‘personal data vault’ enables staff and students to retain control over the equality data relating to themselves. This ensures self-reporting of equality data for specific purposes (such as research or the use of the Cultural Diversity Barometer), while allowing the individual concerned to amend this data at any time, as required, for each specific purpose (Employee Satisfaction Survey, academic research, etc.).

Furthermore, permission could be sought annually from staff and students as to whether this data may be used for specific surveys or research. The respondent could then, for example, distinguish between consenting to the survey(s) and consenting to linking the data to personnel records.

4.7.1 Safeguards for confidentiality in a personal data vault

Based on best practices, specific safeguards ensure the confidentiality of data in a personal data vault, particularly regarding the role of the institution’s administrator(s) and the logging of data access.

80 AI Act article 10(5): To the extent that it is strictly necessary for the purpose of ensuring bias detection and correction in relation to the high-risk AI systems, providers of such systems may exceptionally process special categories of personal data as referred to in Article 9(1) of the GDPR, subject to appropriate safeguards for the fundamental rights and freedoms of natural persons.



1 Role of the organisation's administrators

- **Restricted access:** Administrators should only have access to data that is strictly necessary for their duties ('principle of least privilege'). This prevents unnecessary access.
- **Authentication and authorisation:** Access for administrators is secured using strong authentication methods, such as two-factor authentication (2FA), to prevent unauthorised access.
- **Separation of duties:** Data management and system administration are kept separate to avoid conflicts of interest and minimise the risk of misuse.
- **Anonymisation and pseudonymisation:** Identifiable data is anonymised or pseudonymised, so that administrators in the IT department do not have direct access to personal data without the user's explicit consent. This is because administrators do not have access to the personal data or the pseudonyms used.

2 Logging of data access

- **Detailed logging:** All access requests and operations are recorded in log files. This includes who accessed the data, what data was viewed or modified, and when access occurred.
- **User auditability:** Users can view log files to check who has accessed or processed their data. This increases transparency and trust in the system.
- **Regular audits:** Logs are periodically reviewed by independent parties to detect misuse or unauthorised access. This may form part of a wider security audit.
- **Limited retention period:** Log files are only retained temporarily, for example, for thirty days, unless they are required for a specific investigation into security incidents.

3 Technical security

- **Secure-by-design software:** The data vault is designed with built-in security mechanisms, such as encryption of data at rest and in transit (end-to-end encryption).
- **Independent testing:** Regular penetration tests carried out by external parties ensure that vulnerabilities are detected and rectified at an early stage.

4 User authorisation and control

- **Users themselves determine which organisations or individuals are granted access to specific datasets in the data vault.** These authorisations can be amended or revoked at any time.
- **Access is restricted to the purpose for which authorisation has been granted,** and organisations are obliged to delete the data after use.

Conclusions & recommendations



This Guidance for navigating GDPR and equality data in higher education and research closes with several conclusions and recommendations. Monitoring developments relating to diversity and inclusion is vital to this sector. Still, it faces a range of objections concerning the methodology to be used, the preconditions, and the manner of implementation. Below is a concise summary of what is explained in greater detail, with a view to addressing these objections and answering questions about processing equality data.

5.1 Conclusions

- **Privacy and inclusivity are complementary**

Privacy protection is a prerequisite for the trust needed to share equality data voluntarily and securely, which is essential to ensure that inclusion and equal treatment policies are effective.

- **Equality data is crucial for tackling discrimination**

Without reliable data, it is impossible to tackle structural inequalities and discrimination within higher education and research, or to underpin and evaluate policy measures.

- **The collection of equality data can be carried out lawfully and responsibly within the framework of the GDPR**

This is subject to compliance with strict conditions, such as explicit consent, a balancing of interests, and technical and organisational measures.

- **Transparency, voluntariness and self-identification are core principles**

Data subjects must be fully informed, given the opportunity to decide for themselves what data they wish to share, and always retain control over their personal data.



- **Risks of discrimination, stigmatisation and bias require a careful approach**

Preventing bias in data analysis, providing context alongside the figures, and treating social identities with respect are essential for the ethically responsible use of equality data.

- **There is a need for continuous monitoring, review and transparency**

Results must be shared within the organisation and with those involved, including concrete action plans and a culture of open dialogue.

5.2 Practical recommendations

- **Implement a *privacy by design* and *privacy by default* approach**

Develop procedures and IT solutions that protect privacy from the design stage, such as pseudonymisation, restricted access and data encryption.

Use clear and understandable communication

Explain on time and in clear terms why data is being collected, how it is used, what rights data subjects have and how data is protected. Specifically: explain the broader purpose of the questionnaire and specify what the information will and will not be used for.

- **Request explicit, voluntary and informed consent whenever equality data is processed**

Ensure active consent mechanisms are in place, allowing data subjects to skip questions or withdraw consent without facing any adverse consequences.

- **Provide options for self-identification and multiple identities**

Ensure that respondents can express themselves in a way that suits them, taking into account complexity and intersectionality. Specifically: offer participants the option⁸¹ to

- 1 amend the information provided;
- 2 provide anonymous explanations for their answers;
- 3 tick multiple boxes to indicate multiple identities.

⁸¹ DJA (2022A). Een verkenning van visies van staf en studenten van kleur op antidiscriminatie-data. Online: <https://www.dejongeakademie.nl/publicaties/2300805.aspx?t=Antidiscriminatie-data-praktijken-wereldwijd-en-visies-van-studenten-en-staf-van-kleur>. [in Dutch]



- **Carry out regular audits and evaluations**

Monitor compliance with the GDPR and the effectiveness of privacy and inclusion measures.

- **Promote training and awareness amongst staff regarding the processing of equality data**

Train staff on diversity, inclusion, ethics, and privacy to foster shared responsibility.

- **Develop a feedback mechanism for data subjects**

Involve respondents in the process and take their concerns seriously, for example via a personal data vault with control functions.

- **Foster a culture of accountability, openness and transparency**

Ensure that (1) findings relating to the equality data collected and (2) action plans to prevent and combat unequal treatment and discrimination are widely shared within the organisation, and use benchmarking to measure progress.

Appendices



A Glossary

The terms below are taken from the European Handbook on Equality Data⁸², unless otherwise stated.

Anonymisation The conversion of personal data into data that can no longer be used to identify a natural person, taking into account ‘all means that could reasonably be used for this purpose’ by both a controller and a third party. A key factor here is that the processing must be irreversible.⁸³

Data Any piece of information, in numerical or other form, that reveals something about an aspect of reality and can be used for analysis, reasoning or decision-making.

Data subject (Data subject) The person to whom the personal data relates. (Article 4(1) of the GDPR)

Direct discrimination This occurs when a person is treated less favourably than another person is, has been, or would be treated in a comparable situation on the grounds of ‘race’ or ethnic origin, religion or belief, disability, age or sexual orientation.

Data disaggregation⁸⁴ The breakdown of data by protected characteristics to reveal inequalities.

⁸² See: European Commission (2016). European handbook on equality data. 2016 revision. Directorate-General for Justice and Consumers. Directorate D - Equality. Online: https://ec.europa.eu/newsroom/just/document.cfm?action=display&doc_id=43205

⁸³ See: WP29, Opinion 05/2014 on Anonymisation Techniques. Online: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

⁸⁴ See, for example, p. 19: European Commission (2018). Guidelines on improving the collection and use of equality data. High Level Group on Non-discrimination, Equality and Diversity (Subgroup on Equality Data): “disaggregating the collected data as much as, and where possible, on the basis of age, gender, ethnic or racial origin, religion, migration background, sexual orientation and disability”.

Discrimination testing⁸⁵ A form of social experiment conducted in a real-life situation, involving two or more individuals who are identical in all relevant characteristics except for one: the protected characteristic based on which discrimination is expected, such as ethnic background or a disability. The testers apply for a job, respond to a housing advertisement or request a service, usually on a large scale, after which the outcomes and the treatment they receive are carefully compared. Because all other variables are kept constant, any difference in treatment can only be explained by discrimination based on the characteristic in question. Situation tests measure discrimination in areas such as access to employment, housing, the hospitality sector, and other services, and the results are increasingly recognised as evidence in legal proceedings.

Diversity dimensions⁸⁶ Also known as ‘grounds of discrimination’ or ‘characteristics’. These reflect the areas in which individuals may differ from one another, and the grounds on which distinctions may be made. They can be used to analyse the extent to which teams, departments, and organisations are diverse, and where they may wish to become more diverse.

Equality data Any information that is useful for describing and analysing the current state of affairs about equality. This may be qualitative or quantitative.

Inclusion/inclusivity⁸⁷ A prerequisite for a safe learning and working environment in which everyone feels at home and can flourish, and where a diversity of perspectives is taken into account and considered.

85 See p. 77: European Commission (2016). European handbook on equality data. 2016 revision.

86 See: Social and Economic Council of the Netherlands (SER). Under the banner of ‘Diversity in the Workplace’, the SER publishes reports and charters that are in line with European initiatives, such as the EU Platform of Diversity Charters. The SER draws on academic research and the practical experience of organisations, and is recognised by the government as the leading advisory body in the socio-economic field. Online: <https://www.ser.nl/-/media/ser/downloads/thema/diversiteitinbedrijf/2024/charter-etnisch-culturele-diversiteit.pdf> (in Dutch)

87 See: Ministry of Education, Culture & Science (2020). Nationaal actieplan voor meer diversiteit en inclusie in het hoger onderwijs en onderzoek. Online: <https://open.overheid.nl/documenten/ronl-3d4db279-2786-44c8-a1af-5337604eebfd/pdf> (in Dutch)



Indirect discrimination occurs when a rule or practice that, at first glance, appears to be the same for everyone, in practice affects a protected group more severely than others. Protected characteristics include ‘race’ or ethnic origin, religion or belief, disability, age and sexual orientation. Such a rule is only permitted if it serves a legitimate aim and the means chosen to achieve that aim are appropriate and necessary. For people with disabilities, discrimination is not unlawful if the employer is legally obliged to make reasonable adjustments that remove the disadvantage.

Intersectionality (also known as intersectional thinking) was introduced by Kimberlé Crenshaw⁸⁸ to highlight that different forms of injustice and inequality can be intertwined, meaning that individuals may face multiple forms of discrimination or privilege simultaneously.

No harm to those concerned⁸⁹ This means that no data collection activity should cause discrimination, bias or stereotypes, or reinforce existing discrimination, bias or stereotypes. Furthermore, the data collected must benefit the groups it describes and society as a whole.

Personal data Any information relating to an identified or identifiable individual. (Article 4(1) of the GDPR)

88 Crenshaw (1991). Mapping the margins: Intersectionality, identity politics, and violence against women of color. Stanford Law Review 43(6),1241-1299. Online: <https://doi.org/10.2307/1229039>.

89 See p. 9: High Level Group on Non-discrimination, Equality and Diversity, Subgroup on equality data (2021). Guidance note on the collection and use of equality data based on racial or ethnic origin. Directorate-General for Justice and Consumers. Directorate D – Equal Opportunities Equality and Union citizenship. Online: https://commission.europa.eu/system/files/2022-02/guidance_note_on_the_collection_and_use_of_equality_data_based_on_racial_or_ethnic_origin_final.pdf



Proxy data⁹⁰ Data that does not directly relate to a characteristic such as ethnic or racial background, but which is used as an indirect indicator of this when direct recording based on self-identification is not possible or not permitted. Examples include nationality, country of birth (of the person themselves or of their parents/grandparents), length of residence and language spoken at home. Proxy data can provide a consistent picture for monitoring over time, but they have limitations: they do not capture all experiences of discrimination — particularly those based on skin colour or perceived identity — and may exclude groups that have lived in a country for generations.

Pseudonymisation means processing personal data so it can no longer be traced back to an individual without additional data, while that additional data is stored separately and protected by technical and organisational measures, so the link to the individual cannot be easily established.

Racialisation⁹¹ Being regarded by others as belonging to a particular ‘race’.

90 See p 8: Farkas, (2017). The meaning of racial or ethnic origin in EU law: between stereotypes and identities. Online: <https://ec.europa.eu/newsroom/just/items/54924/en>. “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.”

91 See p. 12: European Commission (2018). Guidelines on improving the collection and use of equality data. High Level Group on Non-discrimination, Equality and Diversity (Subgroup on Equality Data). “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.”

See also: p. 37: 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. “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.”

Processing Any operation or set of operations performed on personal data, such as collection, recording, use, disclosure and erasure. (Article 4(2) of the GDPR)

Race as a social construct⁹² (as a legal and statistical concept) The EU rejects theories based on the existence of distinct human races. The use of the term ‘race’ in legislation and regulations — including the GDPR and the UAVG — therefore does not imply any acceptance of such theories. ‘Race’ and racial background are social constructs: they do not reflect the genetic diversity of humanity, but rather how people are categorised by others and how this influences their experiences and treatment. The term is used because discrimination on the grounds of (perceived) racial or ethnic background is a social phenomenon that requires legal protection.

Self-categorisation Refers to the process whereby individuals situate themselves within ethnic, religious or other social categories. This process matters for how people perceive themselves and interact with others. This principle is based on voluntariness and self-identification.

Self-identification⁹³ Refers to the principle that individuals determine how and to what extent they wish to share personal data relating to characteristics such as ethnicity, religion, gender identity or other relevant indicators. External parties must not impose this process, and it must be entirely voluntary. It enables individuals to define their identity in ways that reflect their own perceptions and preferences, without proxies or third-party assumptions.

92 See also the Racial Equality Directive (Directive 2000/43/EG) article 6: “The European Union rejects theories which attempt to determine the existence of separate human races. The use of the term “racial origin” in this Directive does not imply an acceptance of such theories.” And GDPR recital 51: “Those personal data should include personal data revealing racial or ethnic origin, whereby the use of the term ‘racial origin’ in this Regulation does not imply an acceptance by the Union of theories which attempt to determine the existence of separate human races.”

93 See p. 9: High Level Group on Non-discrimination, Equality and Diversity, Subgroup on equality data (2021). Guidance note on the collection and use of equality data based on racial or ethnic origin. Directorate-General for Justice and Consumers. Directorate D – Equal Opportunities Equality and Union citizenship.



Self-registration⁹⁴ This involves employees voluntarily providing information about themselves, with explicit consent, for example regarding their ethnic background, religion, or sexual orientation. This may take place, for instance, via surveys or staff satisfaction surveys, whilst complying with the requirements of the General Data Protection Regulation (GDPR).

Special categories of personal data Personal data relating to matters such as racial or ethnic origin, religion or belief, disability or sexual orientation. In the context of EU legislation, these are not referred to as ‘sensitive data’ but as ‘special categories’ (Article 9(1) of the GDPR)

⁹⁴ See: SER Diversiteit in Bedrijf (2021). Meten Is Weten. Zicht Op Effecten Van Diversiteits- en Inclusiebeleid. Charterdocument. Online: https://www.ser.nl/-/media/ser/downloads/thema/diversiteitinbedrijf/publicaties/2021/Charterdocument_meten-is-weten.pdf. (in Dutch)

B Relevant sources and literature

Commissie Meijers (2023). *Achtergrondnotitie Etnische Dataregistratie: Wat Zijn De Effecten Voor Discriminatie?* Online: <https://www.commissie-meijers.nl/wp-content/uploads/2023/10/Achtergrondnotitie-Conferentie-‘Etnische-dataregistratie-wat-zijn-de-effecten-voor-discriminatie---20-oktober-Amersfoort-3.pdf>. [in Dutch]

De Jonge Akademie (2022A). *Een verkenning van visies van staf en studenten van kleur op antidiscriminatie data.* Online: <https://www.dejongeakademie.nl/publicaties/2300805.aspx?t=Antidiscriminatiepraktijken-wereldwijd-en-visies-van-studenten-en-staf-van-kleur> [in Dutch]

De Jonge Akademie (2022B). *Dataverzameling ter bestrijding van discriminatie aan universiteiten: een verkenning van buiten Nederland toegepaste vragenlijstmethoden.* Online: <https://www.dejongeakademie.nl/publicaties/2300805.aspx?t=Antidiscriminatiepraktijken-wereldwijd-en-visies-van-studenten-en-staf-van-kleur>. [in Dutch]

European Commission (2017). *Data collection in the field of ethnicity. Analysis and comparative review of equality data collection practices in the European Union.* Online: https://commission.europa.eu/system/files/2021-09/data_collection_in_the_field_of_ethnicity.pdf Directorate-General for Justice and Consumers. Directorate D – Equality. Unit JUST/D1.

European Commission (2016). *European handbook on equality data.* 2016 revision. Directorate-General for Justice and Consumers. Directorate D - Equality. Online: https://ec.europa.eu/newsroom/just/document.cfm?action=display&doc_id=43205

European Commission (2018). *Guidelines on improving the collection and use of equality data.* High Level Group on Non-discrimination, Equality and Diversity (Subgroup on Equality Data). Online: https://commission.europa.eu/document/download/8731d6c1-3716-4c9e-a179-6db6de6064a6_en?filename=final_guidelines_4-10-18_without_date_july.pdf

Farkas (2017). *The meaning of racial or ethnic origin in EU law: between stereotypes and identities*, Luxembourg. Publication Office of the European Union. European network

of legal experts in gender equality and non-discrimination. Directorate-General for Justice and Consumers Directorate D — Equality. Online: <https://ec.europa.eu/newsroom/just/items/54924>

High Level Group on Non-discrimination, Equality and Diversity, Subgroup on equality data (2021). *Guidance note on the collection and use of equality data based on racial or ethnic origin.* Directorate-General for Justice and Consumers. Directorate D – Equal Opportunities Equality and Union citizenship. Online: https://commission.europa.eu/system/files/2022-02/guidance_note_on_the_collection_and_use_of_equality_data_based_on_racial_or_ethnic_origin_final.pdf.

Hoepman (2020). *Privacyontwerpstrategieën. Het Blauwe Boekje.* Online: <https://www.cs.ru.nl/~jhh/publications/pds-boekje.pdf> [in Dutch]

Moerel (2023). *Workplace Discrimination and Equal Opportunity. Why monitoring cultural diversity in your European workforce is not at odds with GDPR.* Online: <https://fpf.org/blog/workplace-discrimination-and-equal-opportunity/>

Racial Equality Directive (Richtlijn 2000/43/EG). Online: <https://eur-lex.europa.eu/legal-content/nl/TXT/?uri=CELEX:32000L0043>

Schermer (2013). *Risks of Profiling and the Limits of Data Protection Law.* In: Custers, B., Calders, T., Schermer, B., Zarsky, T. (eds) *Discrimination and Privacy in the Information Society. Studies in Applied Philosophy, Epistemology and Rational Ethics*, vol 3. Springer, Berlin, Heidelberg. Online: https://doi.org/10.1007/978-3-642-30487-3_7.

SER Diversiteit in Bedrijf (2021). *Meten Is Weten. Zicht Op Effecten Van Diversiteits- en Inclusiebeleid. Charterdocument.* Online: https://www.ser.nl/-/media/ser/downloads/thema/diversiteitinbedrijf/publicaties/2021/Charterdocument_meten-is-weten.pdf [in Dutch]

Van den Hoven (2008). *Information Technology, Privacy, and the Protection of Personal Data.* In: *Information Technology and Moral Philosophy.* Online: http://ndl.ethernet.edu.et/bitstream/123456789/47638/1/Jeroen%20Van%20Den%20Hoven_2008.pdf



Guidance for navigating GDPR and equality data in higher education & research

September 2026, version 0.9.5



This guidance was developed with contributions from members of an ad hoc working group chaired by Marlon Domingus (Erasmus University Rotterdam). The working group also included: Aya Ezawa (Leiden University), Jule van den Berg (Erasmus University Rotterdam), Helen Tibboel (Erasmus University Rotterdam), Nellie Konijnendijk (Royal Netherlands Academy of Arts and Sciences, National Network of Diversity Officers), Jan van den Berg (Amsterdam University of Applied Sciences), Floor May-Smit (NHL Stenden), Dick Houtzager (Office of the National Coordinator against Discrimination and Racism), Onur Şahin (State Commission against Discrimination and Racism), Danique Knibbeler (NautaDutilh), and representatives from the Advisory Committee Diverse and Inclusive Higher Education and Research.