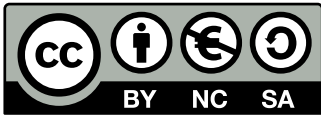# ANTIDISCRIMINATION DATA COLLECTION IN ACADEMIA: AN EXPLORATION OF SURVEY METHODOLOGY PRACTICES OUTSIDE OF THE NETHERLANDS

TABLE OF CONTENTS

SUMMARY AND TAKE-AWAYS

- This research was conducted to inform The Young Academy's project on (self-) categorisation in Dutch academia, which aims to understand both practices and views regarding ethnic and racialised categories used in academic institutions as part of anti-discrimination policies. We requested self-categorisation forms from institutions in a number of countries, and where possible interviewed those involved in administering them to understand their rationale and purposes.

- Under the rubric of their countries' national equalities laws and related institutional requirements, academic institutions are collecting data on people's personal background such as ethnic identity, social group membership and also legal and physical characteristics. These data collection practices differ widely per country and institution.

- There is some variety in the reasons why institutions collect data, what data they collect, and the type of accountability this draws on and in turn makes possible. There is no particular approach or set of priorities we can point to as dominant, instead they relate to local histories and conditions.

- In the EU there is no legal bar from collecting such information, but data protection law does first require a legal basis to be created. This makes collecting data to be used for antidiscrimination purposes a question of political will, as does the presence/establishing of oversight to make the data actionable.

- We identify some protections and features involved in these surveys to make them both more effective in representing diversity, and more acceptable to employees in terms of their confidence in data use and management:

- — Offering participants the possibility to update the information they have provided;
- — Offering the possibility to elaborate on answers given;
- — Allowing participants to tick multiple boxes to denote intersecting identities;
- — Acknowledging that there are far more possible identifications than realistically could be accounted for in the surveys;
- — Giving the broader objective of the survey, and indicating ways the information will (and will not) be used;
- — Clarity on data management and privacy: binding statements on access and purpose limitation.

- The surveys collected for this project tell us something about a possible distinction we might draw between the use of data for institutional diversity oversight (i.e. institutions' upward accountability to government or equality institutions) and the use of data to support and inform antiracist action (possibly different in form, and created through demands for 'downward' accountability of institutions toward staff).

This report should be read in relation to the study *An exploration of the views of staff and students of colour regarding antidiscrimination data at Dutch universities.* In this study, perspectives on the collection of antidiscrimination data in universities, beyond the current information on students' migration backgrounds were analysed. The study's conclusions are summarised here (and in the separate downloadable summary *Antidiscrimination data: Practices Worldwide and Views of Students and Staff of Colour at Dutch universities. Summary and points to consider in future discussions* at www.dejongeakademie.nl):

- This report seeks to contribute to an informed and nuanced discussion on the collection of data for the purpose of combatting ethnic discrimination and racism at Dutch universities. Dutch universities do not currently collect data on the migration background, ethnicity or racialisation of Dutch university staff; data collection on students is limited to migration background. There are growing calls to collect such data as input for effective diversity policies, but there is also considerable opposition to doing so. Statistical data classifying people into ethnic and racial categories can contribute to exclusion and discrimination, for example when politicians argue that there are 'too many' people of a certain category in a country. Data of this kind can, however, also be used to identify, understand and combat ethnic discrimination and racism.

- Students and staff belonging to discriminated groups should play a key role in decision-making about university antidiscrimination policies. If an antidiscrimination policy involves data collection, they should have a say in how these data are collected, analysed and published. The purpose of this project was therefore to explore what students and staff of colour at Dutch universities think about data collection.

- This report is based on interviews with eleven staff of colour at Dutch universities and five representatives of multicultural and/or antiracist student organisations. The findings we present here are a preliminary exploration: they do not offer a representative sample of the opinions held by staff and students of colour in the Netherlands concerning antidiscrimination data. Since it was not feasible to conduct a representative survey with the available means, we set out to explore the widest possible spectrum of views. We cannot draw firm conclusions and recommendations from this study. What we can do is highlight some points that we hope will be considered in future discussions.

- A majority of our respondents favour the collection of data to combat ethnic discrimination and racism. Most proponents, however, also express a certain reluctance: they see data collection as a 'necessary evil'.

- The main argument put forward by the proponents of data collection is that quantitative data are needed to show that racism and ethnic discrimination exist at universities: 'Without data, they won't believe you when you tell them about racism'. Another respondent refers to 'legitimising the experiences of people of colour'.

- A minority of our respondents believe that the disadvantages of quantitative data collection outweigh the advantages. They emphasise the risk of 'pigeonholing'. Instead of categorising people, the university itself should be the subject of research. These respondents would prefer to see qualitative research identifying the individuals or groups for whom the university is not a welcoming and safe place, and examining why that is the case.

- Having a say in the data collection and analysis is crucial for many of our respondents. Several respondents emphasise that the categories identified in antidiscrimination data should be defined in consultation with those affected, i.e., the people who actually experience ethnic discrimination and racism. Stigmatisation can only be prevented through participation, these respondents believe.

- Almost all respondents prefer self-categorisation (with students and staff being asked to self-identify) to categorisation by third parties based on data from population statistics (as is currently customary in the Netherlands). In the view of these respondents, it is crucial to give staff and students of colour a say in how they are categorised when collecting antidiscrimination data.

- Respondents' opinions about data collection are shaped by their trust or lack of it in the institutions that collect these data. In the wake of the Childcare Benefits scandal, our respondents' trust in government has been severely shaken. Many respondents also do not have a great deal of trust in universities and their administrators. Universities and other government institutions can only regain this trust by showing that they are collecting and using data in a way that does not harm but rather benefits minorities.

- For many of our respondents, it is important that the categories used in data collection match their self-identification, so that they can 'recognize themselves' in a survey instrument. Others prefer to think more strategically about categories: for them, the main purpose of categorisation in data collection is not to capture the complexity and layeredness of each individual identity, but to facilitate policies or interventions that serve the interests of minorities.

- Many of our respondents prefer self-categorisation using categories based on 'origin' or 'descent'. In their view, a category based on geographical origin and family history best captures their identity. A few respondents prefer survey questions about nationality or place of birth. Several respondents feel that religion should be included in data collection meant to combat ethnic discrimination and racism.

- A significant majority of respondents emphasise that it is not enough to collect data on ethnicity and racialisation alone if the purpose is to understand and combat inequality and discrimination. The data collected should cover gender, sexual orientation, socio-economic background, disability, and geographical origin in the Netherlands.

- Several respondents say that they found answering questions about ethnicity and racialisation to be stressful and difficult. They emphasise the importance of caring for respondents before, during and after data collection.

INTRODUCTION

In this report we look at the international scope of data collection on ethnic/racial identities. We aim to demonstrate what this collection of data looks like in other countries and where the Netherlands fits in. We focus specifically on analysing surveying methods from countries outside of the Netherlands, who use self-categorisation procedures and on the possible answers survey participants might give, in particular the language, categories and explanations universities choose to use in these surveys. Additionally, we strive to contextualise these practices of data collection.

Over the past years the need to have more diverse and inclusive academic institutions has become a recognized concern in many countries, including the Netherlands. As Bonnett and Carrington state, 'it seems self-evident that a necessary component of creating more inclusive institutions is the availability of more comprehensive information on racial and ethnic differences in workforce composition and career advancement' (Bonnett & Carrington 2000: 487).

While the Netherlands, alongside many European countries, does not, at this moment, collect racial/ethnic data (or at least not directly, instead focusing on nationality, place of birth, and place of birth/ nationality of parents), many Anglo-Saxon countries have been doing this for years. For instance, in the UK the commission for Racial Equality provides the following justification for monitoring of ethnic/racial data:

> 'Although ethnic records are not obligatory under the race relations act of 1976, the legislation does place the responsibility for providing equality of opportunity for all job applicants and employees primarily with employers. Ethnic records are an important tool in fulfilling this responsibility [...] the analysis of ethnic data helps not only in identifying and dealing with unlawful discrimination, but also frequently highlights other employment practices in need of improvement' (Bonnett & Carrington 2000: 488).

The argument is then that, without the data on workforce composition, it is difficult to get a full picture of the extent of discrimination in academia. We will not be able to test what 'diversity policies' are doing to fight this discrimination, and subsequently, we will not be able to solve it.

While ethnic/racial categories are formally absent in the Netherlands, we do have categories that, some argue, are used as a proxy for data on ethnic/racial background. Today the proxy category used for ethnic/racial data is: 'Western- or non-Western migration background'. In 2016 the 'Western/non-Western migration background' category replaced the former category of allochtoon/autochtoon, i.e. foreign/native. The terms were standardized by CBS in 1999 (Yanow and van der Haar 2013: 234). Allochtoon was defined as 'a person who, or one of whose parents, was/were born outside of the Netherlands'. An additional dichotomy was then created between 'Western' and 'non-Western' allochtonen. Yanow and van der Haar argue that "the Netherlands's allochtoon/autochtoon integration discourse, is, in all but name, a racial discourse – one perhaps all the more powerful for being carried out in disguise" (Yanow and van der Haar 2013: 229). They argue that because of the essentialist nature of these categories - you either are an allochtoon or an autochtoon - 'allochtonen' can never be 'integrated' to the 'autochtoon' population.

The collection and use of ethnic/racial data to combat discrimination is not an 'innocent' practice. There are certain risks that need to be kept in mind when creating racial/ethnic categories. First, the categorization of ethnic/racial backgrounds could be seen as essentialist, it could fix categories which are inherently fluid and create new ethic/racial differences (Bonnett and Carrington 2000: 488). Second, while we might be wanting to use these categories for anti-discrimination purposes, statistics on 'ethnic minorities' have often been used to the detriment of these groups (idem: 488). For these reasons it would be necessary to ascertain that categories are based on a fluid and changing reality, and data would have to be handled in line with data protection rules on sensitive data. This would sometimes clash with the possibility to analyse trends through statistical data collection. Trade-offs might have to be made between statistical continuity and accounting for the fluid and temporal character of racial/ethnic categories.

From our data collection process, we draw two lessons: first, that instruments for data collection on ethnic and racialised aspects of identity usually have their origin in historical instruments that can relate to oppressive structures, but may also be tuned to assist efforts to overcome these and to support antidiscrimination efforts. Second, that this suggests there are multiple ways to collect such data, including approaches that do not immediately lead to discrimination.

Can we, however, also legally collect ethnic/racial data in Europe? Collecting data on ethnic and racialised differences is not common, but is not per se illegal in Europe. Instead, it is controlled under data protection law so that countries must provide a legal basis for doing so. In practice, most EU countries have not. The exception is the UK, which (during its time within the European Union) has had a legal mandate for data to be collected for both 'Equality of opportunity or treatment' and 'Racial and ethnic diversity at senior levels' (Information Commissioner's Office, 2018). Unless a law exists that mandates the collection of such data (as in the case of the UK), standard European data protection law (the GDPR) applies, which prohibits collecting and storing ethnicity data (a type of 'special categories of data' in the GDPR) under most circumstances. Purposes of scientific research are one exception to this rule, but with limitations. Data on nationality or place of birth, of a person or their parents, in (for example) the Netherlands or Belgium, are not considered sensitive data under the existing legislation and therefore can be collected without consent of the surveyed respondent. However, if they are used to predict ethnicity or racialised identity, they become subject to the provisions of the GDPR as noted above. The Netherlands has adopted the broad categories of 'Western' and 'non-Western', with the aim of moving away from explicitly racialised categories (although these categories group people in similar ways to previous ones). This terminology has also been challenged as over-general, however, and is now being phased out by the CBS.[1]

Such collection of ethnicity data would be allowed if the data subject gave prior consent. However, the requirements for valid consent are strict. For instance, consent given by an employee to an employer is rarely sufficiently voluntary ('freely given') and therefore unlikely to be valid. Unless the Netherlands or the EU explicitly and legally authorises the collection of ethnicity data, such data collection by universities of its staff is generally not allowed under the GDPR. There may be possibilities to design a system in which employees give their consent to such data collection, in which such consent is really voluntary, and thus legally valid, but this would need to be further explored in consultation with lawyers, in the space between university policy and national/EU law. For instance, it would be easier to do so in a more flat institutional structure where collective control over data was possible, in order to minimise possibilities for misuse.[2]

---

1   see: https://www.ser.nl/nl/thema/diversiteitinbedrijf/actueel/nieuws/CBS-westers-niet-westers\and https://www.wrr.nl/actueel/nieuws/2021/06/30/vervang-de-standaardindeling-westers-niet-westers-in-onderzoek-en-beleid-door-meervoudige-indelingen

2   The notion of a 'data trust' is one example of a structure for enforcing purpose limitation with regard to a particular community's data. For an overview, see: Hardinges, J. (2018). *Defining a 'data trust.'* Open Data Institute. https://theodi.org/article/defining-a-data-trust/

However, if a high response rate and a representative sample is the objective and it becomes apparent that the employees that do not give consent or do not respond are part of an underrepresented group, then this can cause sampling bias. These employees would have no to little chance to be included in the sample, which could bias the data and eventually the conclusions drawn upon it. Similarly, in order to collect such data under the claim of 'public interest' or 'research' at scale, a legal basis outside the GDPR would be required. In cases where such data is collected, data protection law mandates that it is managed and stored securely, on the same basis as other sensitive data.

In the following sections of the report we first describe the methodology of our data collection procedure. Second, we describe the objectives from universities around the world to collect ethnic/racial data as stated in the forms and documents we received from these universities. Third, we explain their methods of data collection. Fourth, we analyse the topics and categories used by universities to collect this data. Finally, we conclude our findings and discuss their implications.

METHODOLOGY

For our research on data collection on ethnic/racial data in other countries, we made use of a purposive sampling technique. Our goal was coverage and thus to receive as many responses as possible from academic institutions across different countries. Therefore, we contacted institutions in countries where we were aware, or assumed, that they were currently collecting data on personal background in general, and ethnic/racialised categories to be specific. Most of the countries that collect data on ethnic/racialised categories specifically are countries that have a colonial history, as for example the United Kingdom and South Africa, and we therefore focused on these given our assumptions and positionality, time and resources, and a lack of available information online regarding where and to what extent institutions were collecting such data.

This sampling technique, though rather broad, has implications and biases (not only limited to methodological biases). First, there was substantial non-response. Many continental and central European (EU), Singapore, the UAE, and US organisations did not reply to our emails. The reports we have sourced stem from universities, as the corporations we emailed did not respond. Second, we did not contact universities in India and other low- and middle-income countries, apart from South Africa. Third, as our ethnicity and racialisation too forms our perspective as researchers, it is important to reflect on our positionality. This research is conducted with two female White professors, and five student researchers of whom four identify as people of colour and one as White.

These choices, in combination with the non-response and our positionality, resulted in a final sample that consisted of mainly universities in Anglo-Saxon countries and of countries where reporting has been publicly discussed and where it relates to legal requirements and policy demands. In total, we analysed 9 countries (the United Kingdom, Canada, South Africa, New Zealand, Northern

12

ANTIDISCRIMINATION DATA COLLECTION IN ACADEMIA: AN EXPLORATION OF SURVEY METHODOLOGY PRACTICES OUTSIDE OF THE NETHERLANDS

Ireland, Ireland, Czech Republic and Norway).[3] Of these, six institutions in four countries provided (complete) survey forms, which we then anonymised. The other countries' surveys were analysed using public data found online and are therefore referred to by name. All these choices affect our conclusions and our recommendations.

3   See Appendix A for an overview of the institutions we contacted.

OBJECTIVES OF COLLECTING ETHNIC/RACIAL DATA

We find that the collection of ethnic/racial data is a practice carried out by many Anglo-Saxon countries' universities. The surveying of this data happens under various headings (see Appendix A for documents): 'Media and equal opportunities monitoring' (England A), 'Workforce diversity questionnaire' (Canada A), 'University Employment Equity Census' (Canada B) and 'Employment equity details' (South Africa). Often, this data collection is part of a diversity, inclusivity, equality and/or equity agenda of the countries' national government. Some institutions combine it with the application procedure for a job (the United Kingdom, South Africa, New Zealand), whereas others hold a separate survey (Canada A and B).

In most institutions the objectives of the proposed ethnic/racial data collection are made clear in the introduction of the survey. In South Africa, the United States (US), the UK and Canada it is explicitly stated that data collection is a legal obligation from the national government to monitor the efficiency of the diversity, equity and inclusion (or anti-discrimination) policies in place. In Canada, for example, the Employment Equity Act includes a federal contractors programme that states that institutions must make an effort to 'ensur[e] [...] that their workforce is representative of Canada's labour force with respect to the members of the following 4 designated groups [...]: women, indigenous peoples, persons with disabilities and members of a visible minority' (Canada B). In addition to stating the legal objectives, Canadian institution B also mentioned the historical presence of the indigenous population in the institution's geographical location. They made clear that they wanted to commit to ensuring this diverse population. Their way of doing this is through collecting data, anonymising it, conducting a statistical analysis and in this way assessing their progress in their commitment.

In the US, the collection of ethnic/racial data is included both in organisational 'affirmative action' plans and in the US census. 'Affirmative action' refers to a set of laws and policies that intend to counter historical injustices and inequalities, including a mandated 'workforce analysis' for organisations employing more than 50 people. Only the construction sector does not make use of the 'workforce analysis', but has their own separate analyses (Affirmative Action Programs, 2000). In the following sections we will elaborate further on the 'workforce analysis' component of affirmative action, as relating directly to how universities ask employees to volunteer data about themselves.

Somewhat differently, the South African institution noted in the introduction to the survey that it is obliged by law to include the racial and disability categories used by the Department of Labour for Employment Equity reporting, but did not mention how the collected data will be used. This strategy of obliging employers to collect certain data by law is also used in the UK. Here, universities also explain that they use the data collected for human resource purposes. This way, they can make sure that activities toward establishing a diverse working force by providing equal opportunities and advertising job vacancies in the right places can be documented. As mentioned above, countries differ in their collection of employee data. Some monitor diversity, inclusivity, equality and/or equity (as mentioned earlier) in general, involving categories that are historically informed in each country, while some countries do not monitor any employee data outside from gender.

An example of the latter is Norway, where the Personal Data Act prohibits the collection of information about ethnic background. There is however a 'Act relating to equality and a prohibition against discrimination' which has the particular objective 'of improving the position of women and minorities' and helping to 'dismantle disabling barriers created by society and prevent new ones from being created' (Equality and Anti-Discrimination Act, 2018). Under this Act, Norwegian employers have an active duty to combat all kinds of discrimination, specifically by investigating the risk of discrimination, analysing the causes of such risk, implementing measures to counteract discrimination, and finally evaluating the result of their efforts, all of which must be documented. Yet Norwegian employers, like Dutch ones, do not do so by collecting such data but by documenting anti-discrimination actions. They mention ethnicity as a category relevant to this anti-discrimination act but the overall emphasis lies on gender-related discrimination and race is never mentioned (idem).

In compliance with this Act, the University of Oslo has set up 'action plans for diversity and equality'. They define diversity as: 'refer[ring] to a number of factors, such as gender, ethnicity, disability, gender identity, sexual orientation, socio-economic background, age and religion' (University of Oslo, 2021). To be precise, ethnicity is said to include 'national origin, descent, skin colour and language'. Citing the 'Act relating to equality and a prohibition against discrimination', the emphasis in the 'action plans for diversity and equality' is on gender inequality. The university sets certain target figures for future gender balances (for instance, stating that they want to 'increase the percentage of females in professor positions to 40%') but since they have no way of monitoring their ethnic/racial diversity they clearly cannot do the same for ethnic/racial minorities. Their plans to increase 'diversity and inclusion' (which is mentioned separately from 'actions to promote gender balance and equality'), remain unspecified. The only tangible plan to increase 'diversity and inclusion' is the initiative to 'celebrate the Sami national day', the Sami are indigenous peoples in Norway and Sweden.

Like Norway, Dutch employers do not have a legal obligation to collect ethnic/racial data. In the Netherlands, the discussion about the collection of ethnic/racial data in universities has involved institutions' boards, representative bodies of staff and students and the national government. Our other report *An exploration of the views of staff and students of colour regarding antidiscrimination data* suggests in the introduction by discussing Dutch parliamentary debates, that the political discourse in the Netherlands is dominated by parties who address the collection of antidiscrimination data as a risk and not a solution for combatting racism in academia. Paradoxically, for years, the Dutch government has already been collecting diversity data via questions about 'birthplace' and 'birthplace of parents' as indicators for exactly 'ethnicity', but has shied away from calling this practice 'monitoring of antidiscrimination data'. What is more, no necessary exception to data protection law has been made with collecting this data.

This, however, does not mean that there is no data collection on topics related to specifically ethnicity/racialisation in Dutch academia with the goal of antidiscrimination. The surveys 'Discrimination research' (Wageningen University & Research, 2022), 'Diversity research' (Commissie Diversiteit, 2016), 'Wellbeing research' (e.g. University of Twente, 2019 and Universiteit Leiden, 2017) and other related surveys as 'Belonging @[at] [institution name]' (Vrije Universiteit Amsterdam, 2019; Utrecht University, 2022), often ask university staff and students, on a voluntary basis, for their experiences with racism at the university. These survey projects however often remain pilots and do not specifically ask about ethnic/racial data. Respondents in these surveys do not have to answer all questions. Frequently, the survey also asks respondents to report on their gender

identity, sexual orientation, nationality and the country you and your parents were born in. Asking for this last category is common practice in the Netherlands, as they frame inhabitants born in a different country or one of their parents' born in a different country than the Netherlands as a 'Dutch person with (Western or non-Western) migration background'.

In Northern Ireland, employers have the same monitoring objectives as the UK in terms of collecting data on racialised and ethnic characteristics, but are additionally also required to monitor 'community background' (i.e. affiliation to Protestant or Catholic community) to achieve fair employment. The Northern Ireland government legally requires public organisations and companies to monitor the so-called 'community background' of their workforce (we further elaborate on this later on). This legal obligation derives from section 75 of the Northern Ireland Act 1998 in which it is stated that there exists a duty to ensure 'equality of opportunity' and 'good relations'. The duty of equality of opportunity entails that 'due regard must be given to the need to promote equality of opportunity between persons of different religious belief, political opinion, racial group, age, marital status and sexual orientation (1), men and women generally (2), persons with dependants and persons without (3) and persons with a disability and persons without (4)'. Similarly, the duty of good relations refers to 'the desirability of promoting good relations' between the different groups mentioned here.

Though section 75 also calls for a broader regard for issues other than 'community background', and Northern Irish employers have the same responsibilities as other UK employers, it is noticeable that the emphasis in the materials collected for this project falls primarily on monitoring parity of treatment between these two religious communities: Protestants and Catholics, given the historical and potentially ongoing violence in the region. Despite this reality, the practice of 'ethnic monitoring' is growing in Northern Ireland. For example, the Northern Ireland Equality Commission emphasises 'the importance of law reform, tackling prejudicial attitudes and effective monitoring and evaluation, supported by robust data collection'. The Equality Commission reaches out to organisations to promote ethnic monitoring and antiracism practice with e.g. holding 'race roundtables' and providing other information sessions and documents (The Office of the First Minister and Deputy First Minister of Northern Ireland & Northern Ireland Statistics and Research Agency, 2011).

The response we received from a university in the Czech Republic indicated that universities there do not collect any form of ethnic/racial data. The university stated that this was completely out of the question as they were legally

prohibited from asking employees about their 'ethnic' background. They stated that it is strongly forbidden to ask about 'ethnicity or race' and that this avoids discrimination. We did not receive responses from any other Central or Eastern European countries.

SURVEY METHODOLOGIES

The direct incentivisation by national governments of the collection of ethnic/racial data can be found in the categories used by universities. In most of the universities that responded to our request, it was compulsory for university staff to fill in these surveys, although employees could refuse to answer any of all the questions in the survey form itself, as explained further below. In the UK, the exact categories and wording are formulated by the Higher Education Statistics Agency (HESA), which legally obliges universities to collect ethnic/racial data (Higher Education Statistics Agency, 2021) and leaves no room for UK institutions to deviate from these categories and formulations. Likewise in South Africa, the institution we spoke to made it clear that they were required to use the categories formulated by the Department of Labour.

A different methodology is used in Canada. Before formulating the survey instrument, the Canadian institute A consulted with interested parties, reviewed similar processes conducted elsewhere in Canada, and consulted the methodology used by the national statistics body. There was also room for people to review and comment on the draft instrument through focus groups which helped modify the survey. While diverting from the census is possible, some Canadian institutions do take over the exact categories stated by Statistics Canada, without further specifying the questions. Overall, Canadian universities clearly have a different approach, but all are structured more or less along the lines of the federal contractors programme. Thus, in the Canadian approach, consistency and comparability of the survey data is still reached, but interested parties also have a say, these parties preferably being part of the '4 designated groups' (Canada B). An example of such deviation of the standard formulation is Canadian institution B which emphasised, in their survey, that 'some terminology may not always align with how respondents define their own identities' and replaced the name of 'visible minority' to 'racialized people of colour'. The exact reasons for these changes are however not mentioned.

All the data collection instruments we collected used the method of self-categorisation to collect personal ethnic/racial data through surveying. This broadly means that respondents could select one or more answers from a list of options. In the Canadian institute A and the New Zealand Census (Stats NZ, 2018), more than one box could be ticked and the option 'other' included a follow-up question with 'please specify below in the box'. Next to that, as mentioned earlier, in the Canadian institute B they explicitly stated that 'a person can belong to more than one designated group'.

In many US universities, respondents also had the option to tick multiple boxes and the box 'multi-ethnic' or 'two or more races' was sometimes included, signalling the fluid nature of the 'race categories'. In the case of Yale University this practice leads to an interesting counting method where people who tick multiple boxes are counted in each of these boxes. This could increase the percentage of 'minority students' in the statistics, which would be convenient for the university as this would signal a diverse student body. However, the alternative, namely the creation of boxes that combine different racial categories, could also have negative implications, on which we will further elaborate in the next chapter.

Furthermore, in both of the institutions we collected data from in the UK, one box may be ticked and 'combination categories' were given in which two categories were combined in one answer. However, the 'other' option did not include a possibility to elaborate. In South Africa, there were few possible boxes to tick, and it was not made clear how many may be ticked. In Canada, New Zealand, South-Africa, Ireland, the US and the UK, it was possible for respondents to decline to answer any question, and to change their responses at any time by completing a new survey.

Another method of data collection is Northern Ireland's 'Fair Employment Monitoring Return' in which employees, applicants, promotees and leavers are asked to complete a survey on their community background for their employer, who must report back to the government. Together with this report, the employer has to notify their employees that their community background has been registered in the employment register with a 'disclosure form'. The disclosure form is as follows:

Though filling in the survey for employees is voluntary (but 'it is a criminal offence to give false information'), the monitoring of community background of the employees is an obligation for employers. In fact, the 'failure to complete and return this monitoring return to the Equality Commission for Northern Ireland by the date at A3 is a criminal offence and carries on conviction a fine of up to £5,000'. There is thus strict regulation in place. This means that if employees do not fill in the monitoring form voluntarily (the 'principal method'), a 'residentiary' method is used to determine their community background. This means that the employer can 'make a determination on the basis of personal information on file/application form' (Equality Commission for Northern Ireland, 2011). Employers then make this determination based on where the employee lives. This stems from the spatial segregation inherited from Ireland's colonial history which then effectively formed battlegrounds in various phases of the conflict. This led to neighbourhoods becoming effectively either Catholic or Protestant territories, a difference which still persists today.

Topics and categories

The substance of the forms collected differs per country and institution. In all forms, employees are surveyed on more than one topic. The topic that was addressed in most of the forms was nationality. Next to that, questions related to ethnicity and racialisation were also asked. Additionally, religion, disabilities, sexual orientation and gender (identity) are also frequently included in the forms. Only one institution surveyed employees on their assigned sex, another institution asked about 'assigned gender identity'. In this section we elaborate on the survey topics and start with the most central topic in our study: racial categories. We then discuss nationality, religion, disabilities, gender (identity), sexual orientation and assigned sex.

Racial categories

A general pattern with racial/ethnic categories is that universities which use them usually do this in accordance with the law, meaning that these universities follow the categories of a nationwide census. Especially in the UK and the US there is little room to deviate from these categories. The categories in the census vary greatly per country and are historically specific. Sometimes categories are added, combined, or separated.

In the US, part of federal affirmative action policies is the 'workforce analysis' (Affirmative Action Programs, 2000). The categories for this analysis are formulated as following: 'the total number of incumbents, the total number of male and female incumbents, and the total number of male and female incumbents in each of the following groups must be given: Blacks, Hispanics, Asians/Pacific Islanders, and American Indians/Alaskan natives'.

Most American universities - which have their 'diversity statistics' published on their websites – stuck to these categories. It is not usually a legal obligation for universities to publish these statistics: instead they are a regulatory tool that can be used by government authorities. Although the keeping of these records is an obligation under the affirmative action programs, and these records must be available to the office of federal contract compliance programs upon request, the publication of the statistics on the university websites seems to be voluntary. Often 'Asians' is made a separate category and the category of Pacific Islanders becomes: 'Hawaiians and other Pacific Islanders'. Furthermore, as mentioned above, the rigid nature of the categories is made more fluid through either giving the option to tick multiple boxes or adding the box 'multi-ethnic' or 'two or more races'. While this could be advantageous for those wishing for a more precise way to self-identify, it is important to be careful when two or more racialised categories are forged together. A clear example of this is the forging of the category 'Asian and White' by the university of Maryland and a school district in Washington (Venkatraman, 2021; North Thurston Public Schools, 2020). The categorisation of Asian/White here seems to be based on performance and opportunity in universities and therefore suggests that White students and Asian students have the same performance statistics and opportunities. In reality, this however does not have to be the case. Measuring performance and race in one indicator causes significant differences to be overlooked and could thus bias the results and conclusions drawn from them.

An immediate difference with the European context can be seen when looking at the definition that multiple US universities give for the category 'White',

namely, 'a person having origins in any of the original peoples of Europe, the Middle-East, or North Africa'. This seems like an odd category for the collection of antidiscrimination data in Europe, as, for example, in the Netherlands Middle Eastern and North African people face significant racism and discrimination. In the US, however, this can be explained by their history of restricting certain migrant groups' paths to citizenship (Awad et al., 2021). The inclusion of Arab and MENA people in the category of White can be seen as the result of early legal battles. (Awad, Hashem, Nguyen 2021: 117). While racialization of Arab and MENA people existed for much longer, the attacks of 9/11 and their aftermath mobilised scholars and advocacy groups to fight for the inclusion of an 'Arab/MENA' category in the US census, so as to make their ethnic and racial discrimination visible (idem). The history of racial categorization of Arab and MENA people in the US demonstrates the importance of contextualising racial/ethnic categorization and it confirms the need for flexibility and revision in categorization.

In Northern Ireland, in contrast to the US's approach, there is a strong policy focus on monitoring religious community background. Universities collect self-categorisation data on a broader set of categories deriving from the census' question on 'ethnic group', namely White, Chinese, Irish Traveller, Indian, Pakistani, Bangladeshi, Black Caribbean, Black African, Black Other, and Mixed Ethnic Group (an open response). Respondents could tick only one box. The term 'Irish Traveller' is noteworthy as this group is defined as a racial group in Northern Ireland, but also Ireland. Following the Race Relations (Northern Ireland) Order 1997 legislation in Northern Ireland, a 'racial group' is 'a group of persons defined by reference to colour, race, nationality or ethnic or national origins' (The Office of the First Minister and Deputy First Minister of Northern Ireland, 2015). The same authority defines Irish Travellers as 'a community of people commonly so called who are identified by themselves and by others as people with a shared history, culture and traditions, including historically a nomadic way of life on the island of Ireland' (idem). The term Irish Traveller is not considered a stable definition, however, because it covers a diverse set of identities.[4]

Canada's approach differs from the US and Ireland's by including the category 'members of a visible minority' (changed by one institution surveyed to 'racialized people of colour'). The use of a 'catch-all' category in asking 'Are you a member of a visible minority?' has advantages, however the ethnic and racial categories used in these surveys has been critiqued (Bonnett and Carrington 2000) on the basis

---

4   Sometimes the terms 'Roma' or 'Gypsy' are also used in this context. The broader political and academic debate is portrayed well in this report: Equality Authority. (2006, July). *Traveller Ethnicity*. https://www.ihrec.ie/download/pdf/traveller_ethnicity.pdf

that they are too broad, essentializing, misleading, or insufficient. It is not feasible to include all different ways in which people might identify and still end up with a useful statistical analysis. The category 'racialized people of colour' (while the formulation could be contested), to some extent solves these problems. It specifically tackles racism and discrimination without assuming any racial/ethnic identification - and therefore automatically excluding certain identifications, because the amount of possible racial/ethnic identifications is too large to realistically all be included in a survey - while still producing workable data for statistical analysis.

A disadvantage of the use of a 'catch-all' category is that the possibility to differentiate between the wide range of people that fall under this catch-all category is eliminated. One institution tried to counter this by asking a follow-up question, in which to specify which visible minority respondents belonged to. Respondents could choose among the following: 'Arab, Black, Chinese, Filipino, Japanese, Korean, Latin America, South Asian, (e.g., Indian, Pakistani, Sri Lankan, etc.), Southeast Asian (e.g.,Vietnamese, Cambodian, Laotian, Thai, etc.), West Asian (e.g., Iranian, Afghan, etc.) or another visible minority'. Here, as in other instances, some of these categories refer to countries, some to broader geographical locations (such as South Asian), some to racial categories such as Arab and Black. This also demonstrates the difficulty to define what we talk about when we talk about 'visible minorities'.

Another particular category for universities in Canada is that of indigenous/ aboriginal persons. The fact that this is seen apart from the category 'visible minority' can be easily explained through the historically specific oppression of indigenous peoples in Canada which was (and in many ways still is) different from the oppression of non-indigenous people of colour. There are a variety of ways in which universities deal with the category 'indigenous/ aboriginal person'. One institution left it at the category as it is stated in the census. Another followed up with more specific questions about geographic origin, first nation identification, Inuit/Inuk identification, métis identification, or 'another' and finally asked to (in the case of first nation identification) specify what nation(s) they belonged to.

The survey explicitly states that people may not identify with the terms in the Statistics Canada and the Canada Employment Equity Act, but that identities such as 'Cree' or 'Blackfoot' all fall under the umbrella of indigenous/aboriginal peoples. A somewhat similar approach is used in the New Zealand census, where specific questions are asked about being a Māori descendant: 'Are you descended from a Māori (that is, did you have a Māori birth parent, grandparent or great-grandparent, etc)?' and about the specific tribe(s) the respondent would belong

to: 'Do you know the name(s) of your iwi (tribe or tribes)?' (Stats NZ 2018). Both these practices demonstrate the acknowledgement that there are far more possible identifications than realistically could be accounted for in surveys such as these.

Following its own census, the UK attempts to identify 'ethnic origin' and in order to do so the question 'what is your background?' is asked in one of the surveys we accessed. Respondents could check only one box among the following answer options: 'White (British, Irish, White background - other), mixed (White and Black Caribbean, White and Black African, White and Asian, mixed background - other), Black or Black British (Caribbean, African, Black background - other), Chinese, Arab, gypsy or traveller, and "other ethnic group"'. What stands out here is the specifications in the 'mixed' category. All specified 'mixed backgrounds' are White and x. Other forms of 'mixed backgrounds' - such as Asian and Black - fall under 'mixed background - other'. The inclusion of the 'other' in all categories once again demonstrates the difficulty of creating sufficient categories of identification.

In South-Africa, new employees were asked to indicate their 'race/ethnic origin' in their job application form. The categories they could tick (only one) were 'African', 'Coloured', 'Indian', 'White' or 'Other race (an open question)'. All these categories date back to South Africa's passbook system, a key tool of Apartheid used to control movement and activities.[5] After the end of the Apartheid policy (though not the end of its effects), their use was continued in order to identify whether the situation of historically marginalised and oppressed groups was being tackled. These categories are, like those used by Northern Ireland and many other countries we discuss, highly specific to South Africa's political heritage. They are used to provide targeted statistical tools which aim to identify poverty, inequality and discrimination, and thus could be described as being reframed as potentially a tool of antiracist action, despite their use under Apartheid.

'Coloured' is a historical term specific to South Africa, dating from the period of Dutch colonisation and denoting people whose origin included African in combination with any other heritage, and whose native language was either English or Afrikaans.[6] 'Indian' refers to the population of Indian origin brought by the British during the colonial era from the Indian subcontinent, then also under

---

5   For more information on the pass laws in South Africa, see: *Pass laws in South Africa 1800–1994 | South African History Online*. (2022). South African History Online. https://www.sahistory.org.za/article/pass-laws-south-africa-1800-1994

6   For a definition of this term, see: Dictionary Unit for South African English. (2022). *Definition of coloured*. DSAE. https://dsae.co.za/entry/coloured/e01740

British colonial rule, to work as labourers in the South African colony. This group, distinct from native African groups, was similarly subject to the restriction of rights under Apartheid. The 'Other race' category is used to capture the presence of non-citizens such as labour migrants, and to distinguish between migrants from neighbouring African countries and from other parts of the world in a context where South Africa has high levels of (often irregular) labour migration from the region, and also a substantial community of formal migrants from other regions of the world. To account for this, it is explicitly mentioned on the form that this is 'only applicable if [you are] NOT a South African citizen'.

Finally, the survey from New Zealand asks about 'refugee background'. The precise question was 'are you or your parents/primary guardian(s) from a refugee background?'. At first glance this might look similar to the Netherlands' practice of categorising 'migration background', however New Zealand does not use this category as a substitute for race or ethnicity. This focus on refugees may be as a result of New Zealand's migration history, where high levels of historical migration during the 19th and 20th century based on colonialism diminished over recent decades, and refugees gained importance as a new category of immigrants requiring representation.

Another noteworthy observation from the New Zealand survey is the inclusion of a great variety of Pacific Island identities. The survey asks for ethnicity, but does not define this concept. The following categories are given: 'New Zealand European, Māori, Chinese, Indian, Korean, British and Irish, African, Australian, Cambodia, Cook Islands Māori, Dutch, Fijian, Filipino, German, Greek, Italian, Japanese, Latin American/Hispanic, Middle Eastern, Niuean, no response, other Asian, other ethnicity, other European, other Pacific peoples, other Southeast Asian, Polish, Samoan, South Slav, Sri Lankan, Tokelauan, Tongan and Vietnamese'. What stands out here, besides the variety of Pacific Island identities, is the focus on geographical location. Whereas we have seen in other surveys that these identities are often mixed with more racial identities such as 'Black', 'White', 'Arab', etc, these identities are not mentioned here. Exemplary of this focus on geographical location is the category 'New Zealand European', which could be seen as a proxy for 'White'.

Overall, we observed that most countries use 'ethnicity' rather than 'race'. The categories were mostly pre-defined with the option of ticking the box 'other'. There exists great variety between countries in the kind of categories they provide. These categories are historically formed, often through a country's history of colonisation and migration flows. Moreover, they often reflect categories used in a national census. Again, there are differences between countries and

between universities within countries, in the extent to which the university surveys diverge from the national census categories. In the UK and the US the categories seem to be more or less set. These are also the two countries which have arguably the most 'top-down' approach to data-collection because the racial/ethnic categories on which data needs to be collected are specifically formulated by the census in the US and by HESA in the UK. As opposed to Canada where protected classes are mentioned, but under the name of 'visible minorities', making the creation of ethnic/racial categories more flexible. Lastly, as New Zealand and Canada demonstrate, the acknowledgement that there are far more possible identifications than could realistically be accounted for in the surveys may help respondents not to feel discounted.

Nationality

In most forms, institutions ask what nationality employees have and give fixed options. In Ireland, the definition they give to nationality is also explicitly stated before the question is asked: 'The country of which you may be a national and are legally allowed to reside'. In UK institution B, they phrase the question differently: 'Which country defines your identity?' and the question is open, so no fixed answers were given. UK institute A also gives the possibility to fill in a joint nationality. In Northern Ireland, respondents were asked for their 'national identity'. Answers were 'Northern Ireland', 'England', 'Scotland', 'Wales', 'Republic of Ireland' and 'Other' (an open question).

Unlike the UK, Northern Ireland and Ireland, New Zealand uses the word 'citizenship' instead of nationality and also asks respondents for their country of birth. In Canadian institute A, Northern Ireland and South Africa, respondents were also asked about their country of birth. The answers for Canada were 'In Canada' or 'Outside of Canada' with a possibility to fill in the country.

In Canadian institute A they also asked for the employee's language they 'first learned at home in childhood'. The possible answers were English or French, or they could indicate that they spoke 'another language' and fill in what this language would be. Only one answering box could be ticked. They followed this question up with 'Do you have at least a basic proficiency in speaking, reading, writing, and/or understanding English, French, and/or other languages?' In this answer, they could tick more boxes. In Northern Ireland they also asked for the employee's 'main language', but they also added questions about how well they spoke English and the question 'Can you understand, speak or write Irish or Ulster-Scots?'.

In South Africa, they also asked the date of naturalisation if the employee is a South African citizen but not by birth. In Northern Ireland, they asked the same, but in the household census, rather than the employment form.

Categories beyond race and ethnicity

In this section we sketch out the other categories included in the surveys we accessed, in order to contextualise the questions on racialised and ethnic characteristics and show the extent of different countries' requirements in terms of self-categorisation. These further categories potentially add nuance to the reporting requirements by conceptualising diversity as a broad and potentially intersectional issue, offering the possibility both for oversight and for understanding how these identities intersect.

Religion

In all forms, except for New Zealand and South Africa, institutions asked about religion. In Canadian institute A they phrased this as 'belief system' and 'religious affiliation'. A special case is Northern Ireland as they explicitly focus on collecting data on 'community background', a legal obligation. The community background 'refers to whether that individual has been treated as belonging to the Protestant community or the Roman Catholic community in Northern Ireland' (Equality Commission for Northern Ireland, 2011). Community background thus refers to how someone is treated, but not how they identify themselves. The reason for collecting community background is so as to ensure fair participation of both Protestants and Catholics in the workforce. This practice is influenced by the history of sectarian conflict in Northern Ireland.

As mentioned earlier, the community background is surveyed in an employee fair employment form. In this form, employees are asked to tick one of these boxes: 'I am a member of the Protestant community', 'I am a member of the Roman Catholic community; 'I am a member of neither the Protestant nor Roman Catholic community' (Equality Commission for Northern Ireland, 2011). In the national census, questions about 'religion belonging', and 'religion brought up in' were also asked (Northern Ireland Statistics and Research Agency, 2011). Answers were: 'Roman Catholic', 'Presbyterian Church in Ireland', 'Church of Ireland', 'Methodist Church in Ireland' or 'Other' (an open question). According to the Northern Ireland Equality Commission, there is a discussion on extending these categories in order to gain further insights on other 'religious ethnic minorities' so as to ensure fair employment of these communities as well.

Disability

In all forms, except for Ireland's, asked whether respondents had a disability. In Canadian Institute A, respondents were asked if they were disabled or deaf. The institutional survey we accessed from New Zealand asked 'Do you live with the effects of a mental health condition, specific learning disability, long-term medical condition, or other disability or impairment?' and in South Africa a definition of people with disabilities was given: 'people who have a long-term or recurring physical or mental impairment which substantially limits their prospect of entry into, or advancement in, employment' and several examples were given of disabilities which are eligible for a 'disability declaration'.

Gender, gender identity and assigned sex

In all countries we collected forms, except for Ireland, South-Africa and Northern Ireland, respondents were asked about their gender identity. In New Zealand the questions was phrased as: 'Do you identify as lesbian, gay, bisexual, transgender, queer, intersex or Takatāpui (LGBTQITakatāpui+[7])?' In Canadian institute A, respondents were asked 'if they identified with one or more of the following genders: Gender-fluid and/or Non-Binary; Man; Transgender; Two-Spirit; or Woman'.

In Canadian institute B, other more specified questions were asked about 'gender identity and gender expression'. The first question emphasised women: 'For the purposes of employment equity, Women are a designated group. For the purposes of employment equity, do you self-identify as a woman?' The answers were 'yes' or 'no' or 'I do not wish to answer this question'. The second question concerned other expressions of gender and was introduced with a brief explanation: 'For the purposes of employment equity, transgender persons are a group that may face employment barriers. Trans or transgender is an umbrella term referring to people with diverse gender identities and expressions that differ from stereotypical gender norms. It includes but is not limited to people who identify as transgender, trans women (male-to-female MTF), trans men (female-to-male FTM), transsexual, Two-Spirit (as it relates to gender identity) or gender non-conforming, gender variant or gender-queer. For the purposes of employment equity, do you self-identify as a person who is trans, transgender, gender non-conforming, gender variant, gender-queer or an analogous term?' Answers were again 'yes', 'no' or 'prefer not to answer'.

---

7   More about Takatāpui can be read here: Murray, D.A.B. (2003). Who Is Takatāpui? Māori Language, Sexuality and Identity in Aotearoa/New Zealand. *Anthropologica*, *45*(2), 233–244. https://doi.org/10.2307/25606143

In the UK institute B, a question about gender identity was asked 'What is your gender identity?' in which the answers made a relation between gender identity and assigned sex. Respondents could answer 'Gender identity is the same as the gender originally assigned at birth', 'Gender identity is different to the gender originally assigned at birth' and 'Prefer not to say'. In Northern Ireland, the assigned sex of respondents was directly asked. The respondents could indicate 'female', 'male', 'other sex' or 'prefer not to say'.

Sexual orientation

Apart from Ireland, Northern Ireland, New Zealand and South Africa, all countries asked about the sexual orientation of the respondents. In Northern Ireland, we conducted an interview with a university representative for clarification. They mentioned that the country's strong history of 'Christian ethos seems apposed to [asking about] sexual orientation'. In Canadian Institute A and B, respondents were however asked about their sexual orientation: 'For the purposes of employment equity, persons who identify as Lesbian, Gay, Bisexual/Pansexual, Queer, and/ or Two-Spirit (as it relates to sexual orientation) are a group that may face employment barriers. For the purposes of employment equity do you self-identify as a person who is Lesbian, Gay, Bisexual/Pansexual, Queer, Two-Spirit or an analogous term?' Answers were 'yes', 'no', or 'prefer not to answer'.

CONCLUSION AND DISCUSSION

The research conducted for this report shows the survey practices of a set of academic institutions which are collecting data on people's background, physical and legal attributes, and group membership. The anti-discriminatory potential of such data depends on the action taken with it, and therefore on institutional and national policy. We identify in this report some nuances in the emphasis on monitoring equal treatment of different groups, which appears to be informed by legacies of historical discrimination, although there is a clear line between national policy (and thus census categories) and institutional policy on monitoring diversity.

We might, based on the instruments collected for this study, posit at least two distinct ways in which such data could be used. One is on the national level to represent and monitor diversity on the larger scale. This we observe happening through processes of accountability to national equality monitoring bodies. The other is on the local level, where institutions may simply comply with nationally imposed monitoring requirements, but may also go further by forming their own policies and actions to remedy discrimination or inequity, and where the data collected for purposes of national compliance may come into play. It is not possible from the scope of this research to determine where institutions are purposefully going beyond national reporting requirements, but the collection of data sets the stage for such action, including bottom-up efforts toward antiracism such as employee organising. This, however, is a hypothesis offered to inform future research, rather than a finding.

This potential distinction in turn underlines the difference between counting (which is useful, for example, for holding employers to account for discriminatory hiring practices), and recognising/representing. The former can be sourced from processes previously used for purposes of exclusion, as illustrated by the continuity of Apartheid-era categories in current South African employee surveys (with the aim of remedying the inequities those categories were instrumental

in creating), and in the data collected for both historical and current migration policy in high-income countries.[8] Using such data for purposes of recognising and representing, i.e. proactively for inclusion rather than retroactively for identifying inequality that requires action, is a local-level activity that requires political will on the part of organisers and institutional leaders, and can be a goal for which those leaders could be held accountable. These two objectives in gathering data are not necessarily related. For example, data about institutions' compliance with equality policies makes it possible for external actors to hold institutions accountable for discrimination in hiring, but it could also be used by institutions themselves to establish institutional accountability for knowing who is present, and formulating relevant policies to deal with discrimination. Doing so may also require collecting other, more qualitative, forms of data, such as anecdotal accounts of negative incidents; accounts of experiences of belonging and exclusion, but also insights from sources outside the institution in question, such as unions or interest groups. In a context of increasing privatisation, not everyone working physically at an institution is employed through that institution, so it is also possible that data from subcontractors would also be necessary to understand the reality of belonging and exclusion at a given institution.

Our conclusions for the Netherlands are therefore as follows:

- There is no legal bar to prevent institutions such as universities in European countries from collecting information on ethnic and racialised characteristics. However, if done beyond the context of a particular scientific research project, it does require a legal basis to be created at the national level. This makes doing so a question of political will, as does the presence or establishing of oversight institutions to make the data collection meaningful and to make sure such data is not misused.

- Our research points to several provisions as making such systems both more effective in representing diversity, and more likely to make employees willing to contribute to such projects:

  — Voluntariness: participants should be free not to answer questions at all
  — Formal attention to data protection issues: it should be clear to anyone

---

8   For more on categories of exclusion and inclusion, see: Chopin, I., Farkas, L., and Germaine, C. (2014). *Ethnic origin and disability data collection in Europe: Measuring inequality – combating discrimination.* Open Society Foundations. https://www.opensocietyfoundations.org/uploads/d28c9226-bed7-4b1b-ac8b-4455f3c3451a/ethnic-origin-and-disability-data-collection-europe-20141126.pdf

participating in such research how data will be deidentified, managed and used, and that the purpose is limited to a given set of agreed objectives
— The possibility to update the information submitted
— The possibility to elaborate on answers given
— The possibility to tick multiple boxes
— The possibility for employees to engage with the broader agenda determining the collection of the data

It should be noted that these observations on safeguards and participation, which are based on the survey instruments we have sourced for this report, align with those of Will (2019) based on Ahyoud et al. (2018, p. 33) on the principles necessary for the collection of data for equal opportunities policies.

One point we would make from this research process is that we focused explicitly on institutional instruments designed to respond to formal requirements and usually based on census categories. They do not offer any indication as to bottom-up processes of organising or protest going on within communities and institutions. It is therefore important not to conflate the two essentially different justifications for collecting and using data set out above: compliance with diversity policy on the institutional or national level, and antidiscrimination policy or action. If we define antiracism data narrowly as data that can be centralised and used to regulate institutional behaviour (i.e. the surveys described in this report), we miss the many processes initiated by communities within institutions. These are less visible to outsiders than these official survey instruments which respond to legal requirements, but just as important in creating change in institutional policies and practices.

BIBLIOGRAPHY

Affirmative Action Programs, 41 U.S.C. § 68042 (2000) https://www.ecfr.gov/current/title-41/subtitle-B/chapter-60/part-60-2

Ahyoud, N., Aikins, J. K., Bartsch, S., Bechert, N., Gyamerah, D., & Wagner, L. (2018). *Wer nicht gezählt wird, zählt nicht: Antidiskriminierungs- und Gleichstellungsdaten in der Einwanderungsgesellschaft – eine anwendungsorientierte Einführung*. Vielfalt entscheidet. https://www.vielfaltentscheidet.de/publikationen

Awad, G. H., Hashem, H., & Nguyen, H. (2021). Identity and Ethnic/Racial Self-Labeling among Americans of Arab or Middle Eastern and North African Descent. *Identity*, *21*(2), 115–130. https://doi.org/10.1080/15283488.2021.1883277

Bonnett, A., & Carrington, B. (2000). Fitting into categories or falling between them? Rethinking ethnic classification. *British Journal of Sociology of Education*, *21*(4), 487-500.

Commissie Diversiteit. (2016). *Diversiteit is een werkwoord*. Universiteit van Amsterdam.

Equality and Anti-Discrimination Act, Norwegian Law § LOV-2017-06-16-51 (2018) https://lovdata.no/dokument/NLE/lov/2017-06-16-51

Equality Authority. (2006, July). *Traveller Ethnicity*. https://www.ihrec.ie/download/pdf/traveller_ethnicity.pdf

Equality Commission for Northern Ireland. (2011). *A step by step guide to monitoring: Monitoring your workforce and applicants in line with fair employment regulations.* https://www.equalityni.org/ECNI/media/ECNI/Publications/Employers%20and%20Service%20Providers/Monitoring%20and%20review/StepbyStepguide2011updated26-2-14.pdf

Higher Education Statistics Agency. (2021). *Staff 2020/21 - Fields required from institutions in All fields Ethnicity*. HESA. https://www.hesa.ac.uk/collection/c20025/a/ethnic

Information Commissioner's Office. (2018). *Special category data*. ICO. https://ico.org.uk/for-organisations/guide-to-data-protection/

guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/special-category-data/

New Zealand Census. (2018, March 6). [Individual form for the 2018 Census of Population and Dwellings]. Stats NZ.

North Thurston Public Schools. (2020). *Monitoring Student Growth*. https://web.archive.org/web/20201116053140if_/https://www.nthurston.k12.wa.us/cms/lib/WA01001371/Centricity/Domain/2664/MonitoringStudentGrowth.pdf

Northern Ireland Statistics and Research Agency. (2011). *Census*. https://www.nisra.gov.uk/statistics/census

The Office of the First Minister and Deputy First Minister of Northern Ireland. (2015). *Racial Equality Strategy 2015–2025*. The Executive Officer. https://www.executiveoffice-ni.gov.uk/sites/default/files/publications/ofmdfm/racial-equality-strategy-2015-2025.pdf

The Office of the First Minister and Deputy First Minister of Northern Ireland & Northern Ireland Statistics & Research Agency. (2011, July). *Guidance for monitoring racial inequality*. https://www.executiveoffice-ni.gov.uk/sites/default/files/publications/ofmdfm_dev/guidance-for-monitoring-racial-equality.pdf

Universiteit Leiden. (2017, July 4). *Onderzoek naar welzijn promovendi*. https://www.universiteitleiden.nl/nieuws/2017/07/welzijn-promovendi

University of Oslo (2021, may 4). *Action Plan for diversity, equality and inclusion 2021-2024.* https://www.uio.no/english/for-employees/support/human-resources/gender-equality/action-plan-for-diversity-equality-and-inclusion-2021%E2%80%932024.pdf

University of Twente. (2019, October 10). *Eigen onderzoek Universiteit Twente naar welzijn onder studenten*. Universiteit Twente. https://www.utwente.nl/nieuws/2019/10/198073/eigen-onderzoek-universiteit-twente-naar-welzijn-onder-studenten#uitkomsten

Utrecht University. (2022, March 21). *Belonging@UU: Do students feel at home at university?* https://www.uu.nl/en/news/belonginguu-do-students-feel-at-home-at-university

Venkatraman, S. (2021, November 19). *Univ. of Maryland slammed for separating Asian students from "students of color" in graphic*. NBC News. https://www.nbcnews.com/news/asian-america/univ-maryland-slammed-separating-asian-students-students-color-graphic-rcna6151

Vrije Universiteit Amsterdam. (2019). *Belonging@VU*. https://assets.vu.nl/d8b6f1f5-816c-005b-1dc1-e363dd7ce9a5/21a715b8-adc7-4b13-b508-16e9d25c9027/Belonging-at-VU_tcm289-938992.pdf

Wageningen University & Research. (2022). *WUR investigates discrimination with DARE project*. https://www.wur.nl/en/newsarticle/wur-investigates-discrimination-with-dare-project.htm

Will, A. K. (2019). The German statistical category "migration background": Historical roots, revisions and shortcomings. *Ethnicities*, *19*(3), 535–557. https://doi.org/10.1177/1468796819833437

Yanow, D., & Van der Haar, M. (2013). People out of place: allochthony and autochthony in the Netherlands' identity discourse—metaphors and categories in action. *Journal of International Relations and Development*, *16*(2), 227-261.

APPENDIX A: OVERVIEW OF THE INSTITUTIONS CONTACTED

		Amount contacted	Amount responses
Universities	Australia	4	0
	Canada	4	2
	Croatia	1	0
	Czech Republic	1	1
	Denmark	1	0
	France	2	0
	Germany	4	0
	Hungary	2	0
	Ireland	2	1
	Italy	3	0
	New Zealand	4	1
	Norway	1	1
	Romania	1	0
	Scotland	2	0
	Serbia	1	0
	Slovenia	1	0
	South Africa	2	1
	Sweden	4	0
	Switzerland	1	0
	UAE	1	0
	UK	3	2
	USA	2	0
Other	Corporate company	2	0
	Independent public body Northern Ireland	1	1
	NGO	4	0
	Total	**53**	**9**

APPENDIX B: ANONYMISED SURVEY FORMS

The downloadable file 'Appendix B' at www.dejongeakademie.nl, shows the survey forms that were collected and anonymised. Note that we received 5 complete survey forms and 1 incomplete form. The other 3 responses we received to our emails were other forms of information, e.g. a link to a website or diversity reports which described a process. To these 9 responses, we added other information we could find online.

APPENDIX C: SAMPLE LETTER

Dear [name],

 I am writing on behalf of a research project being conducted by 'The Young Academy' branch of the Dutch Royal Academy of Arts and Sciences (KNAW). The project aims to understand different ways of keeping track of staff diversity and inclusion, as practiced in academic and commercial organisations around the world. Specifically, we would like to understand what forms and processes institutions are using in order to have new staff report their ethnic or racial background, and any other characteristics on which the organisation collects data when hiring new staff members.

In light of this research, we are hoping it might be possible for you to share with us any form or process that you use to keep track of staff diversity. If you are able to do so, we will use it as part of a review of different reporting approaches and models for categorisation. It is important to note that we are interested in how self-reporting aims and practices differ across countries and regions, not between institutions. As such, we will not report the name of any institution that shares such an instrument with us, but will de-identify each as fully as possible in our research findings, using anything you are able to share with us as an example of a type, not in relation to your particular institution.

 If it is possible to share a copy of any instrument you use with us, we would be very grateful. If you would like further information on our project, please feel free to contact either of the project leads – Linnet Taylor (<email>) or Saskia Bonjour (<email>) who would be happy to discuss this further.

Yours [sincerely/faithfully],

[name]