

Collection and Use of Data on Equality and Participation

The principle of equal treatment has been codified in various national and international legal instruments. As such the German General Equal Treatment Act (German: Allgemeine Gleichbehandlungsgesetz, AGG) offers access to legal action on the grounds of unequal treatment in order to sanction individual discrimination.

Apart from individual cases of discrimination, inequality may also occur on an institutional level, and therefore warrants equally close attention. For some time now a wide range of data has consistently been collected on the subject of discrimination against women, and people with disabilities. This has allowed for the identification of problem areas, as well as the implementation of targeted measures. Due to Germany's historic background, the collection of data on racial, ethnic, religious, and cultural discrimination is a highly sensitive subject, and continues to cause controversial debate.

„I often get asked how many cases of discrimination against these or those groups there are each year. Unfortunately, I have to say: There are no such numbers, yet. We are working to enable a uniform collection of data.“

2011 speech by Christine Lüders, Head of the German Federal Anti-Discrimination Agency.

However, as the following example illustrates, the collection of data on equality and participation may be beneficial – provided that certain ground rules are followed. In 2003, the PISA study reaffirmed that children with a migrant background have more limited opportunities in the education system when compared to children with no migrant background. This often results in poor learning successes. While this problem area had previously been identified, it was the collected data which helped to pinpoint the cause. The measures which were taken on the basis of these findings continue to show a positive impact, as confirmed by repeated monitoring.

This approach of goal-oriented data collection solely based on voluntary action, and self-categorization in compliance with data protection regulation, may also serve to identify problem areas and adequate measures to be taken in other spheres of life where people are vulnerable to exclusion due to their ethnicity. In concrete cases of discrimination statistics may sometimes even serve as evidence in legal proceedings.

In addition, as signatory to various human rights conventions, Germany has the duty to periodically submit relevant reports. This means that Germany needs to submit information regarding the composition of its population, occurrences of discrimination, as well as the implementation of the principle of equality. As such data is only available to a limited degree, Germany has repetitively been reprimanded.

This dossier provides you with information regarding the collection and use of data on equality and participation.

In this dossier the starting point will be to introduce the concepts of equality and participation data, present previous practice, and elaborate on the possible data uses. Additionally, we will introduce the current legal framework for the collection of so-called 'sensitive data' in Germany. This includes personal data attempting to represent ethnical categories, and which refers to political views, religious affiliation, or sexual orientation. Furthermore, we will introduce universal principles, which must be respected when collecting sensitive data.

Finally, we will present experiences from a number of countries with a tradition in collecting equality data, in which elaborate and pragmatic mechanisms and regulations have already been developed for the purpose of data collection and use.

In addition, we will offer a compilation of relevant materials on the subject.

1. Starting Point

Terminology

This dossier does not use the term ‘ethnic data’. Instead, it speaks of equality and participation data regarding a person’s origins (such as the external appearance, the language of origin, the cultural practices, etc.), their group membership or religious affiliation, and their – or their families – migration history. Given Germany’s heavily loaded history in which ethnical categorization and data collection facilitated a Holocaust, it is hardly appropriate to speak of ‘ethnic data collection’. We therefore made the conscious decision to refrain from using this term.

Using the terminology of equality and participation data attempts to bring the goal of data collection to the forefront. These concepts move the attention away from the categorization of the data object, and towards the aspired results: to dismantle barriers of participation, and equality before the law, as well as de facto equality in everyday life.

For that reason, this dossier will mostly mention the two terms in combination, unless a focus on one of the concepts is intended. The word ‘data’ generally refers to equality and participation data.

Data Protection

Personal data that attempts to represent ethnical categories, and which refers to political views, religious affiliation, or sexual orientation is considered ‘sensitive data’. Such data is subject to specific data protection regulations, which is reflected in Privacy Law on a national and an EU level. This law states that sensitive data may only be collected on a voluntary basis, and that only the interviewed person may make a categorization. Categorization by third parties is therefore legally excluded. In addition, sensitive data must be stored and used in an anonymous form.

The fact that this requirement is sometimes disregarded or circumvented does not principally speak against the collection of data. However, it does speak for the enforcement of existing rules – without compromises or restrictions. A misuse of these rules undermines the legitimate societal intention of putting equality into practice.

Even though there is strong opposition to recording a person’s ‘ethnic origins’ in Germany, the category ‘migration background’ can increasingly be found in surveys and statistics since the 2005 micro census.

„Allowing for statistical data regarding a person’s ethnical or national origin [...] to be used as evidence of discrimination in court proceedings allows for the exemplary demonstration of the possibilities and limits of ethnic monitoring“.

„Ethnic Monitoring als Instrument von Antidiskriminierungspolitik?“ (“Ethnic Monitoring as Anti-Discrimination Policy’s instrument?“), by Mario Puecker

In the following section, this dossier will discuss why, and for what purpose data should be collected and used. Collected data aims at the identification of problem areas in the context of (insufficient) equality. On the basis of these findings, measures to address discrimination mechanisms can be implemented in a targeted manner.

Additionally, quantitative data can sometimes serve as evidence in concrete cases of discrimination.

Furthermore, Germany is under the obligation to periodically submit sociological data, such as data regarding the composition of the population, to international organizations.

1.1. Migration Background as Vicarious Variable

Census and Micro census

A census is the legally prescribed counting of population, buildings and housing. It documents the amount of people that live in a certain area and shows how they live and work.

In contrast, a micro census is an annual survey of one percent of German households, also regulated by law. A micro census aims to document selected structural, economic, and social characteristics of the population.

Initially, both the census and micro census documented the participant's nationality, differentiating between 'German' and 'foreigner'. Since a lot of immigrants and their descendants have taken on German citizenship, and since late repatriates with German citizenship often have similar situations as other immigrants, a mere documentation of nationality was considered sufficient. For that reason, the micro census of 2005 started to indirectly determine the category 'migration backgrounds'.

In the 2004 Micro census Law, it was determined that starting in 2005 1 % of the population is to be questioned annually on their citizenship, and the probable duration of their stay. Also, the "nationality of the parents (shall be recorded) every four years, if they have or have had their permanent residence in Germany since 1960". Furthermore, the "year of arrival and, if naturalized, [the] former nationality" is to be collected.

The Federal Statistical Office of Germany, which is responsible for the execution of the micro census, refers to a person as someone with migration background, if

1. the person was not born in the territory of the present Federal Republic of Germany, or immigrated in 1950 or later.
2. the person does not hold German citizenship, or has not been naturalized.
3. a parent of the person fulfils at least one of the conditions under (1.) or (2.).

The 2009 Census Act established the obligation for the European wide census to declare if people, “themselves or their parents have moved to Germany after the 1 December 1955”. For this purpose the interviewee shall state his or her former country of residence, as well as the year of his/her or his/her parents’ arrival in Germany. 1955 was established as a “demarcation year” because this was when Germany signed its first recruitment agreement for guest workers (from Italy).

Limitations of the ‘Migrant Background’ Category

According to the above definition, repatriates, children of migrant workers, ‘immigrants for other reasons’, as well as adopted children and children born to white German parents abroad are regarded as having a migrant background. Therefore, whether a person is white and German born but raised in Mallorca and then moved to Germany, the grandson of a Turkish guest worker, a woman born in Bangladesh and raised by a white German family, or the son of a black Swede and a white Dane born and raised in Germany they all, irrespective of the (non-German) nationality, fall within the category of ‘migrant background’.

The ‘migrant background’ category – as defined in micro-census and census law, and therefore reproduced in other surveys – does not take account of family language, skin color, ethnic origin, history of discrimination, group affiliation, or identity. Therefore, it is not nuanced enough to provide a sufficient understanding of a person’s prior experience with exclusion and discrimination. Possible specifications of the respective categories will be discussed below.

The survey in question provides important information regarding a person’s migrant background. However, in many respects the survey ignores such factors indispensable for the identification of a person’s experience with discrimination. For example, an Afro-German man has, due to his German father, by birth, the right to a German passport. However, he is no less affected by identity checks through the police than a black South African. Thus, the category of ‘migration background’ compresses and oversimplifies existing diversity.

1.2. Previous Surveys and Categories

It is proactive to collect sensitive data indirectly related to ethnicity. However, this data does not necessarily provide information on the actual immigrant's background.

Surveys frequently request information regarding a person's nationality, birthplace, and the migration history of their parents and grandparents. Below you may find information about where relevant data is collected, and by whom.

Statistical Offices

The regional statistical offices and the Federal Statistical Office are the central actors for collecting data in Germany. They are responsible for carrying out the micro census, as well as for making the collected data accessible in the form of data sets. This enables researchers to retrieve records for scientific studies.

Public Administration

Public administrators have access to data which can allow for conclusions concerning a person's ethnic origin to be made. The Citizens Registry and the Central Registry of Foreigners collects information regarding a person's origin and nationality. In addition, such information is stored by Security services, school authorities and social insurance agencies (e.g. pension scheme).

Police

The police are able to collect data required for the police crime statistics. In the data collection one criterion distinguishes suspects between 'foreigners, and Germans. This unqualified categorization of certain individuals with migrant background as non-citizens, and thus supposedly as criminals, leads to the making of wrong assumptions based on data from the crime statistics. This may mislead some people to draw generalized conclusions about 'crimes committed by foreigners'.

Scientific Research

In the context of scientific research other categories related to 'migrant background' may be surveyed. These may include a person's nationality, native language, or place of birth. Such information itself does not constitute sensitive data. However, it indirectly alludes to a person's possible ethnic origin. Such substitute categories are referred to as 'proxy'. Scientific studies usually aim to analyze or clarify circumstances or problems using these 'proxy'.

Institutions, such as statistical offices, public administration, and the police, collect sensitive data for various reasons. Such data is also collected for scientific research, as well as in the private, and employment sector. This data is mostly compiled for statistical purposes. In some cases its purpose is to demonstrate existing inequalities.

Private Sector

The private sector, including insurance companies and banks, collects data – partly through market research. Therefore, businesses indirectly have access to information regarding the population's ethnic composition.

Employment Sector

The German Employment Agency is responsible for the collection of statistics on employment and unemployment. In this capacity, information relating to nationality, and thus indirectly to ethnic affiliation is gathered. Companies that ascribe themselves to Diversity-Management-Programs usually implement a data collection system in order to measure to what extent their diversity goals have been met.

Due to the diversity of players there is a risk of the insufficient implementation of data protection regulations in practice. These shall be further discussed in the light of the current legal situation, as well as the general principles of data collection.

1.3. Recognizing Problem Areas and Taking Targeted Measures

Sociological data has the ability to display the societal status quo on various levels. Such data can answer questions, including whether people with migrant background obtain a job according to their qualifications and prior experiences; or whether the admission to university of immigrant children raised in Germany stagnates or increases. Such data reveals areas in which targeted interventions may be needed. On the basis of these interventions, after repeated data collection, it is possible to check whether these interventions have had a sufficient impact. This enables legislators and other relevant institutions to steer socio-political developments.

„The insufficient availability of data is still one of the major challenges in anti-discrimination work.“

Speech by Christine Lüders, 2011,
Head of the German Anti-Discrimination Agency

The collection of equality data in a company context allows employers to understand its workforce's needs and composition. This helps employers to retain qualified personnel in the long term. For instance, surveying the demand for kosher, halal, vegetarian or vegan food can offer insights into how a canteen should adapt its offering.

In addition, non-governmental organizations (NGOs) have an interest in incorporating sociological data into their work. Such data can be used to support arguments for or against a certain measure, as well as to indicate the need for policy intervention. For example, demonstrating that the police randomly stop and search young men with dark skin more often than others would offer a clear indication that there is a need for a review of police practices. In order to raise the public's awareness to certain issues, NGOs depend on facts – as made available through data collection.

Science, as a key player in data collection, is the driving force in putting into action specific research interests. As a political consultant science can identify problem areas and suggest appropriate measures. Similarly, through its help, it is possible to detect the potential effects of a measure.

1.4. Proving Discrimination Cases

Proving that there has been a case of discrimination can be difficult. For certain forms of direct and indirect discrimination statistical data may provide evidence, substantiate it, or help to establish proof in court. In cases of direct and indirect discrimination, court proceedings often take recourse to statistics.

However, in Germany this has been done in very few cases, as the required sensitive data (including ethnic origin and religion) is seldom collected, and thus rarely available.

The Labour Tribunal of Berlin-Brandenburg “believes that statistical evidence must be taken into consideration, because there would otherwise be no way of bringing to light hidden discrimination in promotion cases (‘glass ceiling’)”.

Judgement of the Labour Tribunal Berlin-Brandenburg, 26 November 2008, Az. 15 Sa 517/08

Direct Discrimination

We speak of direct discrimination if a person in a concrete situation is treated less favorably than a comparable person. For example, a woman in a senior position filed a complaint in a Berlin court for being denied a verbally promised promotion to a management position. As she had become pregnant before the commencement of her new position, a male colleague was given preference. In 2008, the Employment Appeals Tribunal of Berlin-Brandenburg ruled in her favor, as she was able to provide statistical evidence. The statistics revealed that, even though the vast majority of employees in her company were women, management positions were largely occupied by men.

In another example, parents in the Czech Republic with a Roma background filed a complaint, as their children were placed in a special needs class without having undergone appropriate testing. Statistical evidence that shows that a disproportionate amount of Roma children are placed in special needs classes persuaded the Grand Chamber of the European Court of Human Rights that ethnic discrimination had taken place.

Indirect Discrimination

In cases of indirect discrimination a seemingly neutral regulation disproportionately puts groups with certain characteristics at a disadvantage.

Such discrimination was found in the UK case *Hussein versus Saints Complete House Furniture*. In the late 70s a Liverpool furniture store did not employ anybody from a

particular residential area as a matter of principle. Due to the high unemployment rate in that area it was feared that the employees' unemployed friends would visit the furniture store, thereby driving away potential customers. After it was shown that 50 % of the residential area's inhabitants were black – in comparison to 2 % of the total population of Liverpool – the company was found guilty of indirect racial discrimination.

1.5. Reporting to International Human Rights Bodies

Member States of the United Nations and the Council of Europe have committed themselves to complying with reporting mechanisms by signing important human rights conventions. These reports require background information on the countries' composition. This may include equality and participation data.

Germany is a signatory to the UN Anti-Racism Convention (ICERD). It must therefore submit a comprehensive report on the current state of racism in Germany every five years.

In 2001, the United Nations held its third 'World Conference against Racism, Racial Discrimination, Xenophobia and Related Intolerance' in Durban, South Africa. A declaration was adopted, which called on signatory States to develop national action plans, as well as to implement the recommendations included in the declaration. Without supporting data it would not (or not sufficiently) be possible to monitor the implementation of these commitments.

“While taking note of the State party's reluctance based on its particular history, to categorize its population on ethnicity, the Committee reiterates the concerns which were raised in its previous concluding observations (para. 14) that no adequate and accurate criteria have been put in place to enable the production of reliable statistics on the composition of the population in Germany and recalls the importance of reliable data for identifying and addressing racial discrimination”

CERD/C/DEU/CO/19-22, S. 2.

During its assessment of a country context, the committee also considers 'shadow reports' by civil society organizations. So far Germany has provided insufficient data on the 'ethnic composition' of its population. For that reason, in 2015, Germany was yet again criticized by the UN Committee for the Elimination of Racial Discrimination. Based on the general principles for the collection of sensitive data, as elaborated on further on in this dossier the CERD Committee recommends to the German authorities to collect differentiated equality and participation data.

2. Data Protection Regulations

The possibility to collect, store, and process equality and participation data is regulated through various data protection regulations. On a national and international level efforts are being made to protect personal and particularly ‘sensitive data’. The aim is to prevent abuse, or to be able to sanction it within the existing legal framework.

In 2008, the German Institute for Human Rights held an expert discussion on “Data Collection as Evidence for Ethnic Discrimination”.

In Germany data protection is regulated through the Federal Data Protection Act and through a precedent by the Constitutional Court dealing with ‘informational self-determination’. The state data protection laws are only applicable within the respective state administrations. On the EU level the Data Protection Directive 95/46/EG, and the Charter of Fundamental Rights of the European Union constitutes the main legal framework pertaining to personal and sensitive data.

In international law data protection standards are set by the ‘Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data’ and the Council of Europe’s European Convention on Human Rights. Moreover, the UN International Covenant on Civil and Political Rights includes the protection of privacy.

Personal data, such as ethnic origin and religion are considered ‘sensitive data’, and are therefore subject to special legal protection. The collection of such data may only be carried out on a voluntary basis, and through self-categorization. After its collection, sensitive data must be anonymized. Publishing such data is only allowed if disaggregated and represented through specific characteristics within complex data sets (such as nationality, occupation, marital status, etc.).

2.1. Data Protection in Germany

In Germany, data protection is regulated through the Federal Data Protection Act. A ruling by the Constitutional Court from 1983 on informational self-determination specified the right to data protection.

The Federal Data Protection Act came into force in 1977. Since then it has been updated at various instances, and has progressively been adapted to European and international guidelines.

The Federal Data Protection Act applies to both public authorities and private parties, including private law companies, unions, political parties, freelancers or associations. It specifies the conditions, under which federal public authorities may collect data. A legislative act must form the legal basis for a survey. In addition, the surveyed individuals must explicitly agree to the collection of sensitive data, such as religious affiliation or ethnic identity. Private entities are subject to equally strict conditions.

Besides the requirement of explicit consent, there must be a legitimate justification for the collection of data. Thus, the measure must be vital or of important interest. Furthermore, data may only be collected if it has been made public by the surveyed individual. In addition to the collection of data, the processing and use of data is also regulated through similar requirements. Such limitations may influence or determine whether previously collected information may be processed. If collected improperly the further storing or analyzing of such data is prohibited.

In 1983, the Federal Constitutional Court set a precedent concerning ‘the right to informational self-determination’. This right aims to strengthen the protection of privacy, particularly in the light of modern data processing. The judgement confirms people’s self-determination over their personal data, meaning that individuals may decide when, and to what extent their personal information is revealed. This also implies that based on the principle of self-determination, people may refuse to provide information. Even though it does not explicitly mention it, the right to informational self-determination has been derived from the German Basic Law (Grundgesetz, GG). It stems from the right to personality, the freedom of action (art. 2 para. 1 GG), as well as human dignity (art. 1 para. 1 GG). It therefore enjoys constitutional protection. In Germany, the terms ‘data protection’, and ‘right to informational self-determination’ are synonymous.

The State Data Protection Laws apply to the respective state authorities, i.e. state agencies and municipalities. In particular, they stipulate the legal basis for the respective State Data Protection Commissioners.

Although there is robust legal framework in place to protect personal and sensitive data, it can be assumed that is not fully applied in all of Germany. For example, in 2005 it came to light that the police in several states used software which recorded the sexual orientation of suspects without their knowledge.

2.2. Data Protection at EU Level

At the EU level, the handling of personal and sensitive data is mainly regulated through the Data Protection Directive 95/46/EC, and the Charter of Fundamental Rights of the European Union.

Data Protection Directive 95/46/EC

The EU Data Protection Directive was adopted by the EU Council in 1995. By 2006 all new and old Member States had amended their legislation. The directive attempts to regulate a complex field comprising the collection of census data, scientific research, the collection of personal data in public and in private institutions, as well as member databases and police data. For the collection and processing of data in these respective areas there must be specified exceptions and limitations.

On 14 April 2016, the European Parliament voted in favor of the storage of passenger name records (PNR data). A news article in the German newspaper ZEIT describes the objectives and benefits of storing PNR data, as well as summarizing the controversial aspects.

The Directive defines personal data as data that make a person identifiable. The rules regulating the use of such data apply to both the public, and private sector; however, home security, defense, and criminal law are excluded.

It stipulates that data must be collected in a clear and lawful manner. Additionally, Member States are required to specify in greater detail under which conditions such data may be legally processed.

The Directive defines which categories are to be treated as sensitive data. These categories include ethnic origin, political opinion, religious or philosophical convictions; trade union membership, and data on health or sexual preferences.

Personal and sensitive data may only be processed legally if it was collected directly, and with the consent of the concerned person (voluntariness). In some EU countries written approval is required; in others verbal confirmation is sufficient.

The surveyed person has the right to be informed about the processing of their data in an understandable manner and language. This is derived from the principle of 'good faith'. Furthermore, it has been established that any collection of data, such as in the context of government statistics, scientific research, and the identification of discrimination, must serve the general public interest.

Charter of Fundamental Rights

The Charter of Fundamental Rights fulfils the function of a constitution for the European Union. It stands in the tradition of the European Convention on Human Rights of the Council of Europe. The Charter lays down every person's right to respect for private and family life. The protection of personal data is governed in Article 8. A collection of such data is only possible with the person's consent, and a legitimate legal basis. In addition, individuals are granted the right to inspect and correct their data.

2.3. International Data Protection

Internationally, the Council of Europe (CoE) "Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data" provides legal norms on the collection of sensitive data. In addition, the CoE European Convention on Human Rights, and the UN International Covenant on Civil and Political Rights further codify the protection of privacy.

The CoE 'Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data' (ETS Convention No. 108) is the only international agreement specifically dedicated to data protection. It came into force in 1985 and has been signed or ratified by all 47 Member States. It only applies to identifiable persons, and therefore does not include anonymized data. It mainly regulates the processing and storage of data. The collection of data needs to be reasonable and lawful. Article 6 prohibits the automatic processing of sensitive data, unless domestic law provides appropriate safeguards. Article 8 provides that any person may consult, edit, or delete their data.

The European Convention on Human Rights includes in its definition of the right to respect for private and family life the protection of private data. The International Covenant on Civil and Political Rights by the United Nations, also known as the ICCPR, prohibits illegal encroachments into a person's private life, and therein includes data protection.

3. Principles of Collecting Sensitive Data

The collection of sensitive data is highly controversial, in particular with respect to ethnicity, political belief, sexual orientation, and religion. Any objections to disclose such information are understandable and justified.

However, there might be less resistance to the collection and processing of equality and participation data if there were binding principles in place which are strictly adhered to.

Legal requirements, targeted recommendations, analyses, and prior experiences can result in useful, goal-oriented, pragmatic, and to the surveyed individuals, acceptable means of facilitating the collection of equality data.

„ECRI [European Commission against Racism and Intolerance] also consistently stresses that such data shall be collected with due respect to the principles of confidentiality, informed consent, the voluntary self-identification of persons as belonging to a particular group, and in close co-operation with all the relevant actors, including civil society organizations.”

ECRI General Policy Recommendation N°11, 2007

This dossier aims to discuss the ideal design of the principles of self-definition, consent and secure data protection. A solid legal basis is a vital requirement for the collection of equality and participation data. Furthermore, possible, suitable, and targeted categories will be discussed. Finally, the principle of consultation concerned individuals, their self-representative structures, and other relevant actors will be elaborated on.

3.1. Guaranteeing Self-Identification

That people are different and diverse is an accepted fact. Exclusion and discrimination arise when distinctive features are loaded with negative meaning, which is then projected onto individuals bearing such features. Therefore, it is understandable that individuals falling into a certain group reject any categorization to circumvent exclusion and discrimination. This may particularly manifest itself in the context of the collection of data regarding ethnic origin, affiliation, or attribution. This can be explained by looking at German pre-war history, which is dominated by imposed labels such as 'Jews', 'Gypsies' and, 'disabled'.

The principle of self-identification, as part of data collection, attempts the opposite. Interviewees identify themselves as members of a particular group (or not). They have the freedom to assign themselves to a category and remain in complete control. Whether outsiders perceive this as ‘right or wrong’, or ‘appropriate or not’ is irrelevant. No one is entitled to question or change this identification. This principle must remain unchallenged, even if the person chooses to change their identification. It would therefore be possible and understandable if the grand-daughter of a Turkish guest worker were to identify on some occasions in surveys as Turkish, and in others as German.

“... such identification [which identifies individuals as being members of a particular racial or ethnic group] shall, if no justification exists to the contrary, be based upon self-identification by the individual concerned.”

CERD, General Recommendation 8, Thirty-eighth session, 1990

In 1990, the UN Committee on the Elimination of Racial Discrimination (CERD) introduced the principle of self-identification as part of its General Recommendation Number 8. The principle is supported by the ‘European Commission Against Racism and Intolerance’ of the Council of Europe. According to the Framework Convention for the Protection of National Minorities of the Council of Europe, members of national minorities are endowed with the right to self-identification.

The principle of self-determination, as enshrined in these international treaties, must be provided in domestic law, as well as in the context of concrete situations where sensitive data is collected. Only then can the concerned communities be expected to support data collection.

3.2. Guaranteeing Consent

Based on the EU Data Protection Directive 95/46/EC, it is imperative for sensitive data to be given voluntarily. In its 2009 country report, the ‘European Commission Against Racism and Intolerance’ (ECRI) requested Germany to collect data on the ethnic and religious composition of its population, while observing the principle of consent. This means that any surveyed individual must be informed in a clear manner that a refusal to disclose sensitive information will not result in disadvantages or penalties. Similarly, surveyed people must be informed of the purpose of the collection of personal, sensitive information, in order to make a more informed decision on whether to respond or not.

“Member States shall provide that personal data may be processed only if one of those conditions is met:
(a) the data subject has unambiguously given his consent; ...”

Directive 95/46/EC of the European Council, Article 7 from 1995

3.3. Ensuring Secure Data Protection

A first step of collecting data is to request information from a specific person. At this point it is still possible to connect the collected data to the surveyed person. The person's data is then transferred to a collection system. In case the data has not yet been anonymized, this will be done in this step. This means that it will most likely not be possible to link the information back to the person. Subsequently, the data from the entire survey is combined into a comprehensive data set, called aggregated data. If this information is then grouped according to certain characteristics (for example age), it is spoken of as disaggregated data. In the context of aggregated and disaggregated data it is no longer possible to single out individuals, or to draw conclusions regarding their affiliation to a certain group.

„In order for the fundamental right to informed self-determination to be respected in an information society as a citizens' right to freedom, data protection aims to limit the processing of personal data through legal regulation and technical measures.”

State Commissioner for Data Protection and Freedom of Information Nordrhein-Westfalen, Germany

The described process is subject to data protection regulations. When collecting individual data, subjects are entitled to refuse the disclosure of sensitive data without fearing sanctions. Should a person choose to disclose personal information, they may also choose how to categorize themselves. In addition, data protection includes the requirement that personal data must be stored anonymously. Only disaggregated data may be made accessible to the public.

The protection and confidentiality of personal information is one of the key principles in data collection. Nevertheless, critics of data collection fear that these principles are circumvented or ignored. Thus, data protection must be upheld at all costs, especially by government institutions.

3.4. Legal Basis

Public, scientific, and economic interests in information regarding social contexts result in large amounts of data collection. This collection of data is not always necessary, or socially responsible. It is the responsibility of the legislature to specify the framework in which the collection of, often sensitive, equality and participation data is allowed. Thereby misuse can be prevented.

§ 13 of the Federal Data Protection provides that “categories of personal data (§3 para. 9) are only allowed to the extent that regulation intends this or requires it for substantial public interest

“(2) Such data shall be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to rectify them.”

Article 8 Paragraph 2 of the Charter of the Fundamental Rights of the European Union. 2000/c 364/0

reasons.”

3.5. Possible Survey Categories

Data collection which is aimed at the documentation of discrimination or the recognition of social problems requires the collection of sensitive data. Individuals and their affiliation(s) are not monolithic, but a combination of many facets. A person has not only a gender, an origin, an external appearance, a political opinion, a religion, a nationality, a native language, but also other facets that make up their identity. This illustrates that people are made up of a combination of diverse and sometimes complex traits, affiliations and backgrounds which shape their respective lives, as well as social dynamics in general. Surveys must adequately reflect these complexities.

In a person’s various stages or areas of life these facets may become more or less dominant or relevant. Self-perception may be different from other people’s perception (external perception).

To represent this multifaceted character, unspecific and broad categories, such as migration background, should be avoided. Thereby, findings from research, and studies on experiences with discrimination would provide more detailed information. This would allow for the execution of targeted social interventions. In its 2009 country report, the ECRI advised Germany to collect the data on its population composition according to religion, language, nationality, and national or ethnic origin.

Furthermore, it must be weighed on how many categories can be surveyed, as financial and time limitations usually play an important role. However, a determining factor should be the objective of the study. Any categories should be determined in line with the objective.

Therefore, a study on ‘racial profiling’ would predominantly contain categories on visible minorities. For data collection regarding success in schools, the category of e.g. native language would be more relevant.

In the following section, it will be discussed which possible categories should be considered in the course of specific data collection.

“Each survey requires sensitive handling when establishing definitions and the education category, in order to prevent any further stigmatization of the discriminated against group.”

Quote from the protocol of the discussion of the Anti-Discrimination Agency in 2009.

a) **General Personal Data:**

Citizenship

Citizenship often plays a central role as a survey category. Respondents are questioned on their nationality, indicating whether they possess a German passport, whether they have 'EU citizenship', or whether they are 'third country nationals'. However, it is equally important to survey whether a person has changed their nationality. This could mean that they have changed their citizenship, or that they have a dual-citizenship.

Residential Status

For people without a German passport the immigration status may have a profound impact on their possibilities and quality of life. People with a temporary or permanent residence permit have different rights and obligations than those with a tolerance permit, or those in an asylum procedure. This can affect the possibility of employment, the right of establishment, and the right to travel. The residential status category is an integral part of surveys regarding demographic compositions. However, this category is usually not considered in the context of statistical surveys on the detection of discrimination. Surveying this category would offer insights on the potential exclusion of segments of the population.

Gender

At first glance the category 'gender' seems neither sensitive, nor controversial, as it is surveyed almost always, and practically everywhere. Even though a certain level of awareness was achieved with respect to trans- and intersexuality in Germany, the categories '3rd gender', 'intersexual', 'trans', or 'other gender' have not yet been included in the relevant questionnaires. This means that the problems of an entire population group are being ignored.

Migrant Background

It has become common practice to survey a person's 'migrant background'. However, this means that only a limited variety of circumstances can be included. Beyond the prevailing definition, it should also be considered to include inquiries on individual migration experience (1st generation), parents' migration experience (2nd generation), and grandparents' migration experience (3rd generation). In certain surveys, the respective unilateral or bilateral migration experience of parents or grandparents could provide valuable insights.

(Family) Language

The collection of data on 'family language' can provide information on whether children are educated in their first, or second (or third/fourth/etc) language at school, as well as how this affects their performance.

b) Sensitive personal data:

Ethnic Origin

In Germany, regional and cultural groups such as Swabians, Franks, Pomeranians, Saxonians, Friesians, or Rhinelanders are not defined as an “ethnic category”. In numerous countries in Asia and Africa the affiliation to a certain cultural and linguistic group is an integral part of everyday life. In cases where colonialism has not poisoned the groups’ relationships, it is still possible for the groups to exchange and coexistence without tension. People from Africa or Asia are identified as belonging to a certain ethnic origin while, Germans are not. This phenomenon causes unidentified imbalance.

Affiliation to an ethnic group is not represented in a person’s nationality. Often, countries may host more than one ethnic group. Therefore, the category ‘ethnic origin’ adds informational value. As a result, Kurds can have a passport from Iraq, Iran, Turkey, or Syria. South African citizens can be Zulu, Xhosa, Ndbele, Khoisan, or, Afrikaans.

It must also be considered however, that asking for ethnic origin may have its shortcomings. An example of this would be if the aforementioned ethnic origin were to be used as grounds for discrimination, as provided for in the General Equal Treatment Act (AGG). In such situations, other aspects such as language (Schwäbisch, Basque, Ibgo), appearance, cultural clothing, styles, hairstyles, etc. may not be adequately addressed. Therefore these should be surveyed separately wherever it is appropriate.

Language

As previously noted, language is often closely connected to the category of ‘ethnic origin’.

‘People of Color’/ Skin Color / Appearance

Whether a person is white, black or identifies as a person of color, skin color is a factor which is directly noticeable. It has been known to often trigger racist behavior. Therefore, this category may be of particular relevance for the collection of data relating to experience with racial discrimination. In this capacity, rather than the actual skin tone being the problem, it appears that it is the perception of an individual’s belonging to a particular ethnic group that can cause issues.

Other Changeable Features

While cultural and/or religious features such as clothing, hairstyles, headgear, etc. may be chosen freely, and can be changed, they may also constitute essential attributes of a cultural/ethnic/religious group. In cases where such features are clearly visible (e.g. side locks, turbans, headscarves), they often invoke a feeling of “otherness” in people, similarly to skin color. Therefore, in surveys on discrimination experience, the category ‘other changeable characteristics’ would be relevant.

Religious Affiliation

In official questionnaires in Germany, religious denomination (Catholic, Protestant, no religious belief) are required. In 2011, the micro census interrogated people on their religious community membership (for instance Roman Catholic, Orthodox, Jewish), as well as on their religious confessions, such as Islam, Hinduism or Christianity. Moreover, information provided in the category ‘active or passive religion’ could offer important information on how to organize schools, public holidays, and public work spaces according to specific needs.

The categories presented here give each for themselves only one small insight into the complicated identities and dynamics with regard to integration, exclusion, discrimination experience, and educational success. These must be considered and analyzed in an overall view so that it is possible to recognize differentiated problems, and be able tackle them accordingly. The survey divided into small categorical pieces as suggested here allows, in the context of investigations concerning discrimination, for recognition of complicated characteristic values. It would for example, be more recognizable that Muslim women who wear a headscarf are more likely to be discriminated against than Muslim women who do not wear a headscarf. Equally, it would provide the evidence to prove that young men with dark skin are much more frequently controlled by the police than young men with light skin.

In addition to the suggested features, the option ‘other’ should also always be available within each category in an attempt to include ‘forgotten’ features.

3.6. Integration of Relevant Actors in the Development and Collection of Data

When collecting sensitive data on ethnicity, religion, and group affiliation the principle of involvement of affected communities, their representative structures in the development process of studies, as well as in the definition and selection of categories to be used are recommended by the European Commission against Racism and Intolerance (ECRI).

“ECRI also consistently stresses that such data should be collected (...) in close co-operation with all the relevant actors, including civil society organizations.”

ECRI General Policy Recommendation N°11, 2007

Conceptualizing Studies

Although this is so far only considered in exceptional cases, it is advised that during the development phases of a survey it is important to involve socio-political and sociologically oriented communities in the direction of these surveys. In research design, data collection methods can therefore be tailored for the target group, questions made culturally sensitive, and appropriate terminology found. This would strengthen the acceptance, and cooperation of respondents towards a survey.

Creation of categories in accordance with communities

Category names are not static and depend on the evolution of societies. Whereas it seemed appropriate in the 70s to describe immigrants as ‘foreigners’ in Germany, this term is now considered inappropriate. A social change has taken place here. Therefore, categories that appear politically correct, adequate, and proper today may become outdated in 10 years. To reflect such social changes in statistics, it is important to be in dialogue with relevant communities, and where necessary adjust terms for categories.

4. International Examples of the Collection and Use of Sensitive Data

An overview, as well as suggestions for the FRG will now be presented using examples of different countries in the context of collection of equality and participation data. These examples are intended to show how the principles presented in the collection of sensitive data are already being implemented in specific areas of life in other countries. It presents the UK, the USA, and South Africa.

4.1. Great Britain

Great Britain has had a diverse population for many centuries. This began as early as the Roman Empire, and further developed later in parallel with the history of the so-called British World Empire. The British World Empire contained many colonies, and otherwise occupied areas. This diversity has also partially resulted from complex migration flows.

Not all members of ethnic minorities are immigrants, and not all immigrants are members of ethnic minorities.

Black people reached Britain early on through British participation in the enslavement of West African people, as well as through the Atlantic triangular trade with British goods. In the 16th century, Sinti people of the European mainland immigrated. In the 17th century people were recruited as seafarers from occupied territories in South Asia and prevented from returning, this led to East Asian population growth in Europe. During the 19th century, many German, and Russian Jews immigrated.

On the battlegrounds of World War I soldiers from South Asia fought on the British side. After the war some of these soldiers stayed in Britain, but most returned to their country of origin. After the Second World War, some colonies and British occupied territories gained independence. These colonies however, still remained organized among the commonwealth. Citizens of the Commonwealth Countries had special rights to immigration, especially during periods in which immigrant labor force was economically needed. Many people came from India, Bangladesh, Pakistan, the Caribbean, South Africa, Kenya, and Hong Kong. Later however, the immigration policy was tightened, and this complicated immigration.

According to the UK census of 2011, 87.1 % of the population is composed of a white majority, and 12.9 % of an ethnic minority. Approximately 7 % of the population is Asian or Asian-British, 3 % Black or Black British, and 0.9 % belongs to a different

ethnic group. In the 2011 census, the category ‘mixed’ was newly introduced, which was chosen by 2 % of the population. The census also showed that 11.9 % of the population was born abroad. In 2010, most of the immigrants came from India, Poland, Pakistan, Ireland and Germany.

A few decades ago racial discrimination was a part of everyday life in the UK. In 1965 a law forbade discrimination based on skin color, but only in 1976 did the circumstances really improve. In 2001, public institutions were obliged to check their policies to see whether they were conducive to equal treatment, or not. Positive measures are being increasingly introduced, and they should continue to do so. Therefore, data will be collected according to general principles. For the census, and labor market equality and participation data are collected. Reporting forms now available for ‘stop & search’ should make the practice of racial profiling visible through the inclusion of a field on ethnic origin, such as ‘race’ of the controlled person.

4.1.1. General Data Collection

Data that belongs to the spectrum of equality and participation data are collected in different areas in the UK. One of the general principles of this involvement is that it is always voluntary. Under the UK law no one is obliged to give answers to questions on ethnicity, gender, sexuality, disability, religion, or age.

Data on ethnicity is collected in censuses that are carried out every ten years; in school censuses since 2003, in the four national surveys on ethnic minorities from the Policy Studies Institute since the 1960s, and for national databases for university access.

The question of ethnic origin appeared for the first time in national statistics in 1976. Before that, information on country of birth, and nationality were collected. Since then, survey categories have been changed and revised. The challenge is on the one hand to track developments of the years, and on the other hand to adapt and observe the changing realities. For example, the category ‘Bangladeshi’ remained constant, whereas the category ‘West Indian’ changed to “Black Caribbean”, which is a result of the word ‘Black’ being considered insulting. The category

Gypsy/Roma/Traveler was previously only collected in the school census, and was only introduced in the national census in 2011. A further challenge remains that the categories

The voluntary nature of surveys is very important. Voluntary surveys within the school system develop as follows. Parents or guardians specify the ethnic origin of their children up to 11 years old. Children from 11 to 15 decide themselves with the support of their parents or guardians. For example, an answer sheet to determine the ethnic origin of a school’s students is brought home by a child. If this sheet is not completed, and the parents or guardians have expressed no objection to answering this category, the school may and will determine the category to the best knowledge of the school. Parents should then be informed of the decision of the school, and should see this change and have the ability to remove it. In a formal objection the school may not provide any information on ethnicity. Young people aged 16 and older are allowed to make their own decision and concerning which category to fill out.

must be able to reflect the diversity of the population and also the diversity within the groups.

In 1965 and 1976 discrimination was made illegal through laws which also laid the foundation for 'ethnic monitoring', and the 'Commission of Racial Equality'. It was through this which the Equality and Human Rights Commission (ECHR) opened, observed, and supported the progress on equal treatment. 'Ethnic monitoring' is a consequence of the McPherson investigation from 1999, following the murder of Stephen Lawrence in 1993, in which the British police were accused of racism.

The 'Race Relations Act' of 2000 required public institutions to practice 'ethnic monitoring'. They are also required to make equality a central part of their work, as well as adjust policies procedures, services, and working conditions. Public institutions are required to have anti-discrimination directives, and show that they have no adverse effects on equality. Private institutions, such as companies and parties, have no obligation to 'ethnic monitoring'. Nevertheless, many do it in order to verify the effectiveness of their equality policies. For this purpose, the collection of equality and participation data is essential.

4.1.2. Census

In the UK, census data is collected every 10 years in order to provide information on the population. This data is important government, commerce, and industry. Since 1841, respondents provide information about their country of birth, and nationality, this information on is important to determine the makeup of country. In the 1970s, anti-discrimination laws were concretized, and it became clear that questions on the country of birth and nationality were not sufficient to capture the ethnic and cultural diversity of the country. Some members of ethnic minorities can trace their ancestors in the UK for several generations, and still belong as a part of the visible minority, but yet could still not be recognized as being part of a group with the potential for discrimination experience.

The question of ethnic origin could have been introduced in the 1981 census, but was postponed after some discussions. The debate focused on the possibility, and legitimacy of classifying people in racialized and ethnic categories. Some argued that it would be politically, and morally reprehensible, as the concept of race would be revived, and surveys could lead to further disadvantage minorities. The 'Commission for Racial Equality' (CRE) could not support and examine the implementation of anti-discrimination guidelines without data. Some tests on surveys were carried out between 1975 and 1989. These showed that minorities had generally no objections regarding the questions being surveyed, but were against the interrogation methods.

In the 1991 census, the question of ethnic origin was incorporated into the census. Over time, the categories available were modified. The affiliation to an ethnic group is seen as a personal identity, instead of as an objective fact. Therefore, self-identification is a central component of any survey. Respondents are given the possibility to choose

between a selection of predetermined categories, 'other', or the option to specify more precisely.

In the 1991 census, respondents could choose between the following categories: White, Black Caribbean, Black African, Black, other Black: please describe, Indian, Pakistani Bangladeshi, Chinese, or other ethnic group with a possibility to add on. People coming from different groups were instructed to select the category which they most associated with, or to specify further.

It was criticized that categories were partly built on ethnic and racialized elements, and as a result lead to the racialization of ethnic groups. Other groups started campaigns for the inclusion of additional categories that reflect their collective identity. Since the 2001 census, some amendments have been implemented. The category 'White' was further split into 'British', 'Irish', and 'other white backgrounds' with supplementary resources. This was done in order to make diversity within this category visible. The category 'Asian or Asian-British' contained the specifications 'Indian', 'Pakistani', 'Bangladeshi' and 'other Asian backgrounds'. The category 'Black or Black British' included 'Caribbean', 'African' or 'other black backgrounds'. There were other options of 'Chinese' and 'other ethnic groups'. The possibility for the further breakdown of 'mixed race' was provided in the options 'white and black Caribbean', 'white and black African', 'white and Asian', or 'other mixed backgrounds: please specify'.

The 2011 census introduced the category 'Traveller/Gypsy'. Questionnaires varied depending on the location of the survey, distinguishing specific questions in England, Wales, Scotland and Northern Ireland. Only the sub-categories can be harmonized.

4.1.3. Labor Market

The Commission for Racial Equality (CRE) published the first issue on data collection in 1978 regarding 'ethnic monitoring' (the collection of equality and participation data for the purpose of detecting (un)equal treatment of equal opportunity policies in the work place, entitled 'Monitoring of Equal Opportunity Policy: A Guide for Employers.' This handbook stressed the importance of monitoring for effective equal opportunity policies. With the collected data selection, choice criteria, and processes that are examined it is easier to get an overview of whether individuals are selected according to their performances and skills, or whether preferences (usually for White British) play a role. All employees shall enjoy equal opportunities, and if necessary receive training opportunities in order to be able to develop within the company. Such directives shall be statistically monitored in order to check their effects. The following recommendation was issued for data collection: employees and applicants are to specify their ethnic origin which will be stored in a personal file. During the application process it should be reviewed how application and employment patterns for specific positions, job groups, and departments look, are broken down by criteria of ethnic origin. This would demonstrate whether equality is respected or not.

Public institutions must comply with “ethnic monitoring”, unless if they are explicitly told not to. According to the British labor law, employers should promote equal opportunities, and store information about employee benefits, as well as pay for this purpose as well as for application, promotion and training on ethnicity. Public institutions with more than 150 employees must also examine the distribution of further training, benefits, time limits, complaints, and procedures, and publish the findings annually.

The CRE also stressed that it is not only essential for data to be collected, but it is also the way it is used afterwards that is important. This implies collecting data, counting employees, and examining whether general obligations are being fulfilled. It is therefore important that public bodies check the differences in treatment between members of different ethnic groups, and investigate the reasons which might have led to those differences. This would then be used to tackle disadvantages and possible discrimination.

As a result, discrimination shall be reduced and equal treatment explicitly encouraged. Public authorities shall not only monitor the company of employees through data collection, but also extend the monitoring to the overall scope of obligations and competence of an authority.

Categories from respective previous censuses shall be used for monitoring within institutions in order to provide comparability. Classification is carried out by the person itself and is subject to the principle of self-identification. In case it is chosen by a third party, who shall be considered as a last resort, the opportunity of confirmation or correction must be given. ‘Ethnic monitoring’ is permitted to be carried out in case of classifications by third parties. However, any other use of this data is unlawful. Private institutions are not subject to ‘ethnic monitoring’ duty. Many however, do this anyway to verify the implementation of its internal equal opportunities policies.

4.1.4. 'Racial Profiling'

Racial profiling describes practices of police, and security forces to use criteria such as ascribed ethnic affiliation, religion, and national origin to serve as basis for control, and thereby indicate the racist patterns of the executive forces. Therefore this practice is considered institutional racism. According to Section 60 of the Criminal

The following recommendation was issued for data collection: Employees and applicants are to specify their ethnic origin, which will be stored in a personal file. During the application process it should be reviewed how application and employment patterns for specific positions, job groups, and departments look, are broken down by criteria of ethnic origin. With this, reality of equal opportunities will be reviewed.

"Sir William Macpherson's groundbreaking investigation into the murder of Stephen Lawrence marks a turning point for the police in dealing with racism within their own ranks and the treatment of the population."
Jason Bennetto

Justice and Public Order Act of 1994, the police are allowed to stop suspects anywhere and search them, if the person is suspected to be involved, to have committed, or to be committing a crime. This permission is called “stop and search”.

Due to the increasing public pressure after a long absence of clarification for the race motivated murder of Stephen Lawrence in 1993, a commission of inquiry was convened. The ‘Macpherson Report’ of 1999 contains 70 recommendations, most of which were addressed to the British police in order to reduce the identified institutional racism that causes the practice of ‘racial profiling’.

The British Home Office urged the implementation of the recommendations, which covered all areas of police work. After this, polices involving responsibility to citizens were more thought out, structures changed, and procedures were adapted. Officials were increasingly recruited from ethnic minorities, and there was improved reporting of statistical data on racist crime.

New guidelines for ‘stop and search’ identity checks were issued. Officially, personal, and vehicle controls shall be done on the principles of equal treatment and non-discrimination. Suspicion may never be founded on the sole basis of appearance, and exterior signs, for example, hair color, skin color, clothes, age, and gender. In order to examine, and verify this, documentation of personal controls were reworked, and an independent complaint commission was founded. This is called “The Independent Police Complaints Commission” (IPCC).

Since then, ‘stop and search’ measures set up a protocol where the affected person can determine their own ethnic origin. The arrested person gets a copy of the recorded data with the number of the inspecting officer, and can thus turn to the independent supervisory body if they feel they were treated, and controlled unjustly. The information found in identity checks also includes the reason for the check. The data collected by the police is fed into a database that centrally holds all checks performed by the police.

The personal and sensitive information of the controlled person are kept anonymous, and are mainly used for internal purposes only. This is to ensure that this data will not reach the public. The recording of ethnic origin is meant to illustrate which groups of people are controlled most often. This allows the method of ‘racial profiling’ to be made visible. Statistics from the last few years show that black people and those from ethnic minorities were controlled at disproportionately high rates.

For example, from April 2007 to March 2008, Black people were controlled 7.4 times more and Asian people 2.3 times more than White British people. It is thus possible for NGOs, parties, and associations to denounce such practice with concrete numbers, and promote changes in policing.

The Interior Ministry urged the creation of measures to stop racial profiling. However, the situation evolved through legislative changes. Meanwhile, the indication of ethnic origin is increasingly voluntary, making documentation of racial profiling difficult.

4.2. USA

In this part of the dossier, an overview of the collection of equality and participation data in the USA will be given.

Knowledge of the population's demography is essential for the discussion and implementation of social changes. Without information on categorizations such as ethnic origin, gender, sexuality, age, as well as many others, the protection of the rights of marginalized groups within these categories is considerably more difficult.

To ensure that the collection of sensitive data is carried out legitimately and appropriately the Office of Management and Budget (OMB) in 1977 fixed standards for the classification of federal data on ethnic origin, and ethnicity. These regulations include rules for survey, such as the listing of a number of categories with respect to ethnic origin. They also include strict rules in case foreign identification of ethnic origin or ethnicity of a person is necessary, and the ability to select more than one option when a person may identify as more than one ethnicity.

Before the OMB implemented these standards the category of ethnic origin was broadened, this prompted people to choose incorrectly. Since 1977, people who participate identify themselves with an ethnic origin in the survey. Before that, those who carried out the survey noted the ethnic origin of the participant. That undermined the right to self-identification, and therefore led to the discrimination of many minorities. An important aspect of the standards of OMB classification is that race and ethnicity are viewed as evolving and changing concepts.

In the US issues of improvements and regulations are publicly discussed. These guidelines shall be adhered to in the implementation of data collection for the US government, education and the labor market. Provisions requiring police to adhere to these rules do not exist.

4.2.1. Data Collection in Education

Different organizations carry out research in education. The US Department of Education is the main representative of data collection in the school system.

A division of the Office of Education, the Civil Rights Data Collection is exclusively dedicated to the collection of information on minorities in schools. School districts are required to abide by the guidelines of the OMB, and to allow students or students' guardians to define their own ethnic origin. This ensures that the district data, which is later used by the CRDC, is properly collected.

All data is evaluated under the categories of ethnic origin/ethnicity, gender, disability and English language skills. Information such as students' grades, courses taken by students, access to advisers and advanced courses, as well as other criteria are analyzed according to ethnic origin, and other sensitive data.

In the previous 2011 CRDC survey, a striking imbalance in quality of education was revealed for children with various ethnic backgrounds. It was found that children from minority groups are more often punished than white children, even though they account for a smaller proportion of pupils. These tendencies already appeared in preschools, and continue through high school. Black students represent 18 % of the preschool children in total, but represent 42 % of school children expelled from school.

This trend explains clearly that students that are more often punished are more likely to repeat classes, leave school, or come into contact with the law. Not only was it found that Black and Hispanic students were more often punished, but also that their teachers were less experienced, and less paid than their colleagues in schools with smaller populations of minority students.

Like the case with experienced teachers, advanced courses are often less accessible to non-White students, with only 57 % of Black students, and 67 % of Latin American students offered all math and science courses. This number, compared to 81 % of Asian-American students, and 71 % of White students, points to a gap in access to adequate education.

These statistics on education are particularly important in the consideration of labor statistics. As marginalized groups are already discriminated against in schools, it is likely that this trend continues in life after school.

4.2.2. Police and Racial Profiling

Information on the collection of police statistics is hardly available. The US lacks a uniform system of nationwide statistic collection, in particular since the recording of crimes is arranged differently in each state.

Statistics on crime are compiled on a national level by the FBI in Uniform Crime Reports, by the Bureau of Justice Statistics, and in the National Crime Victimization Survey. The lack of standardization of data offers few reliable statistics.

Information collected for the UCR is given either by an office, or by the police. Before 2013, most information was dispatched by mail, leading to a long and expensive process of analyzing data. Today, data is collected online in order to accelerate the process.

Although the survey process is very complex, statistics provide an insightful view of crime in the United States. The UCR provides statistics on the number of murders, attacks on police men, general crime, and hate crimes. In this publication, the FBI documents on the number of incidents in which a partial or complete connection with ethnic origin, religion, gender, sexuality, etc. could be detected.

“During a Chicago event with law enforcement officials, the president [Barack Obama] recalled times when he’d been pulled over in his car by police. ‘Most of the time I got a ticket, I deserved it. I knew why I was pulled over,’ Obama said. ‘But there were times when I didn’t’”

Huffington Post, 2015

In the National Crime Victimization Survey, data was collected by the Bureau of Justice. Samples of 160,000 people were interviewed twice a year. Here information about the rate, characteristics, and the effects of crime in the United States is collected. Participants are asked, for example, how likely it would be for them to be a victim of rape, robbery, assault, theft, and other crimes. They shall indicate their sex, give their ethnic origin, their age, and say if they live in an urban environment or not. This question focuses on victims of violence rather than on ethnic majority, as well as on how environmental factors can contribute.

Although there is a nationwide collection of data on crimes; there are no rules on the way crimes are to be specified. This means that police officers can distort characteristics of crimes by not reporting every incident. Even information considered 'urgent to report' varies from one state to another.

As with the discrepancies in reporting, there is also no mandatory and complete collection of data on abuse and misconduct of the police.

4.2.3. Census

Collecting sensitive data through the census is the basis for understanding the political and social climate in the US. A central aspect of the US census is the absolute transparency of questions asked. The US Census Bureau's website provides an explanation on the necessity of each question. In the context of ethnicity, the Bureau states that statistics on this characteristic are helpful for the monitoring of political framework, government programs, and racial discrimination. The census is not only based on monitoring discrimination in society.

The OMB provides these ethnic categories for the census in order to create a unified data collection. These categories include 'White', 'Black' or 'African-American', 'American Indian' or 'Alaska-Native', and 'Local-Hawaiian' or 'Pacific-Islander'.

The most recent amendment in the census is the question on ethnic origin in relation with Latin and Central Americans (Hispanic People). This aspect is clarified before the question on ethnic origin, as this category is not classified as an ethnicity ('race') but rather as a nationality.

Years of discussion and constant developments of the social treatments of ethnic origins were needed to implement these standard categories. The US Census Office consulted organizations representing many minority groups. This allows minority groups to gain more control over questions that concern them. Minority groups can act as representatives for their own interests, and rights, and call for different methods of data collection in the census.

There are several different methods of data collection which apply during the census. The most common methods are survey questionnaires and interviews. First, a questionnaire is sent to the residents of a household. If not returned telephone calls and private surveys are conducted. Authorized staffs from the US Census Office complete the calls.

If the respondent refuses to provide information on ethnic origin, the interviewer shall decide on the basis of the interviewee's external characteristics. The process of foreign identification is however, complex. All previously provided information shall be included, and then family members' files related to identification are incorporated.

4.2.5. Labour Market

The US Census Bureau provides the most statistics on employment in the US. There are many different surveys which the agency uses in order to collect data; a few of these are the American Community Survey, The Survey of Business Owners, as well as The Current Population Survey. All of these surveys are an integral part of policy making, and of the distributing of state and federal public funds. Each survey includes a question on ethnic origin; this provides an illustration on the situation of members of various minorities within the labor market in the US.

The collection of this information can help identify whether certain ethnic groups hold which professions more often, and whether they are more often represented in certain positions than others. This is done in order to discover unequal treatment, and discrimination in career advancement opportunities, as well as in the hiring process. The most important point in collecting sensitive data from US labor statistics is the list from the Equal Employment Opportunity Tabulation. A business that is committed to the standards of Equal Employment Business provides positions solely on the basis of their suitability and qualifications. Characteristics such as ethnicity, gender, religion, and sexuality are perceived as unimportant criteria. By interviewing such 'Equal Employment Businesses', data can be used to compare the employment of minorities to that of the majority.

This data shows the development of the number of representatives of certain ethnic groups in employment. For example, the share of Central American (Hispanic) workers has increased by 53 % since 2000. The use of this data can be used to show the proportions of members from different ethnic origins in regard to employment. African Americans make up 11 % of the American work force, and 13 % of the population. This proportion of unemployed African-Americans in relation to the population is much higher when compared to the White population.

If a discrepancy in comparison to the majority is recognized, the law can be used to prevent such discriminations in advance. In construction, for example, some states require entrepreneurs to meet a certain quota of employees with different ethnic origins or gender. If this quota is not met, the company will be taxed higher. If the quota is met, taxes are remitted. Initiatives like this make it impossible to ignore discrimination in the labour market.

4.3. South Africa

South Africa's history is marked by colonization, enslavement, and the separation of racialized populations through numerous practices and laws since the arrival of White colonists. The discovery of diamonds and gold led to the consolidation of White settlers' privileges, and of the legal exclusion and discrimination of Black people.

In 1948, apartheid was institutionalized, and with the 1950 Population Registration Act the population was divided into three categories: White, Black (additional non-African, indigenous Bantu) or Colored (this included South African minorities with 'Cape Malay' and 'Griqua' but also 'Indian', 'Chines', and 'Cape Colored', people of mixed ancestry).

This categorization permitted structural discrimination that produced and consolidated a fall in income, power, and skills. In 1990, after many years of resistance the apartheid system finally broke down, and prisoners, such as Nelson Mandela, were released from political imprisonment.

Through the centuries of racism a huge imbalance in resources and skills in families, and communities arose. When apartheid was dismantled some measures were developed, such as housing projects, anti-discrimination directives, and positive actions in the labor and education markets. These were designed to tackle the lasting effects of apartheid.

In order to observe the progress of these measures, data was regularly collected. For this purpose this dossier first refers to the general collection of data. Data is collected through the census, crime statistics, and in business.

4.3.1. General Information on Data Collection

The collection of personal and sensitive data in South Africa is protected by the human rights privacy, and the 'Protection of Personal Information Act'. This law is seen as a fundamental right, but not as an absolute right. This law is not seen as an absolute right because deviations of data protection in certain situations are legitimate. Therefore, decency, justice, and effectiveness shall be guaranteed in the collection of data, and in accordance with 'Principles of Information Protection'.

This specifies the limitation of the processing and dissemination of data, the indication of the purpose of data collection, the quality of information about what data should be collected, transparency, and insurance on the handling of data, voluntary participation, and accountability. The collection of sensitive data may in this case take place for essential purposes, as well as in an appropriate form.

„The time will come when our nation will honor the memory of all the sons, the daughters, the mothers, the fathers, the youth and the children who, by their thoughts and deeds, gave us the right to assert with pride that we are South Africans, that we are Africans, and that we are citizens of the world.“

Nelson Mandela

The South African Law Reform Institution refers in its discussion paper to the ‘protection of privacy and data’ in existing policies, such as the EU-Directive, or the US Federal Trade Commission. The definition of sensitive data is also broad, and covers the following areas: ethnic origin, political opinions, religious/philosophical beliefs, trade union membership, information about health and sex life, information on debt, financial situation, criminal history and social benefits.

The processing of such sensitive data is subject to prohibition, unless the person concerned has given consent to further processing. Processing should only be carried out in the context made known to the concerned person. In case of inability of the subject to agree to the processing of data (for example, in the collection of blood, and investigations during an accident) the analyzing of data should only be carried out in the interest of the subject, and of a compelling magnitude.

4.3.2. Census

Three censuses (1996, 2001, and 2011) have taken place since the first democratic elections in South Africa. Even during the apartheid era, census surveys were conducted. That information however, is irregular and unreliable, especially in the context of the African population. Laws such as the Native Urban Areas Act, and the Group Areas Act, which govern the spatial separation of Africans and the white minority, distorted official statistics, and underestimated the number of people in cities.

The first census after apartheid, in 1996 was crucial for South Africa’s history, as it was the first in which the entire population was counted. Former censuses estimated the numerical size of each population group as they were defined in ‘Population Registration Act’ of 1950. This categorized South Africans as ‘White’, ‘Colored’, ‘Indian’ and ‘Native’ (Black). Earlier censuses contained neither accurate, nor demographic socio-economic data.

In the 2011 census 120 000 workers were employed to interview each member of individual households. Information about age, gender, work, education, income, births, and deaths in the household were collected. Moreover, questions on access to water, electricity, sanitation, and communication technologies were surveyed. The 2011 census showed that 79.2 % of the population classified as 'African', 8.9 % as 'People of Color', 8.9 % as White, 2.5 % Indian / Asian', and 0.5 % other / unspecified.

„The poor, the rich, the homeless, those in transit at hotels, the young and the old, boys and girls, men and women, the disabled, the educated and the schooled, the wise, the blue collar, the white collar, those on the edge, the sick, the healthy, the unemployed and the employed, those in educational institutions, those in the country under whatever conditions, legal or not, it matters less, there is no difference among yourselves, you are all important and all your importance is equal before the census and you shall thus be counted“

Census 2011, Statistics South Africa

4.3.3. Crime Statistics

As with crime statistics in Germany, statistics about criminal activity do not represent the actual number of crimes. Not all crimes are recorded by the police, and the reporting behavior of victims varies widely depending on the crimes. Nevertheless, every year in September since 1994 the 'National Crime Statistics' issues the results of crime statistics. The offenses are defined in the opening of a case to a particular system of codes; the information is then entered from police lists, where the nature of the crime is reported. This information is found in the South African Police Service. It is collected and compiled from the 1130 police stations found in the country.

The amount of crime is measured by the total number of crimes per category, as well as the number of crimes per every 100,000 people. The latter allows the calculation of the risk from becoming a crime victim for individuals. A general statement about the increase or decrease of crime is not provided.

Problematic and unreflective interpretations, as well as the developments of crime statistics are often carried out. The SAPS in the preparation of its data, for example, didn't refer to updated population figures from the 2011 census, which therefore distorted the figures.

The use of crime statistics shall always be compared to the relative context. These include the relation to time, or global, local and regional comparisons.

4.3.4. Labour Market

Through centuries of racism a huge imbalance in resources and skills in families and communities arose. By the end of apartheid, management positions were mostly occupied by White people, property was also largely in white hands. There are measures that can fight the existing inequality in the distribution of resources, and opportunities, and historically disadvantaged individuals can be offered better opportunities.

'African' or 'Colored' people were categorized as 'historically disadvantaged individuals' under the apartheid system. These categorizations of people did not even hold citizenship before 1993. Today they all fall under the category 'Black'. In addition, incentives are given for employing and promoting multi-disadvantaged people.

Individuals categorized as 'African' or 'Colored' under the apartheid system and those as South African citizens before 1993 fall under the category 'historically disadvantaged Individuals'. Today they all fall under the category 'Black'. Additional incentives for the employment and promotion of multi-disadvantaged people, such as Black women or Black people with disabilities are set.

People can voluntarily classify themselves according to the South African legislation. If they refuse their employers will rely on existing information, and grant the concerned person the right to approve the choice. If the person is considered 'mixed race', they can freely choose one of the relevant categories. If the employer takes over this decision, it is

recommended to choose the category 'Black'. Disability shall also be recognized by voluntary self-assessment.

Categorization is however, still criticized. This is perpetrated on one hand by people against continued classification, and preferential treatment, as well as discrimination of people due to group membership. On the other hand, opposition also comes from White people fearing disadvantage, and who would indirectly be advantaged if these classifications did not exist. Since categorization can have great consequences, but is based on self-assessment, 'People Against Race Discrimination' call out to be registered as 'Black African' when applying for a job in the public sector.

Requirements in the labour market, which also use data on categorisation, include Employment Equity. This obliges employers to put members of historically disadvantaged groups preferentially in higher positions. These directives fall within the category of positive action. Companies with more than 50 employees or turnovers with a profit margin superior to 5 million must submit personnel statistics broken down by census categories. In addition, private business enterprises including NGOs, trade unions, foundations, cooperatives and churches must adhere to this.

A related but separate concept is that of Black Economic Empowerment (BEE). Guidelines are set in the 'Codes of Good Practice' of 2007. Whether a company shall participate depends on their sales and business partners. The participants shall be measured on a 'score card' with up to seven criteria, after which they will be assigned a BEE certificate indicating a participation level that is important for trade and cooperation. The seven columns of BEE include ownership, management, positive action in personnel, training, preferred sources and suppliers, corporate development of smaller companies (mainly owned by black people), and support of socio-economic (charity) projects. In each category a company can collect 5 to 25 points. Companies with high scores reach high levels – level 8 is the lowest and level 1 is the highest, and those with higher levels can then enjoy certain advantages. Companies with a majority of Black employees can only collect points if some of them work in management. Black driving staff or cleaning staff do not grant more points to the company on the BEE score card.

To determine the obligations of each company, companies are divided into three categories based on. 'Exempted Micro Enterprise' are smaller companies with a profit margin up to 5 million (EUR 1 ≈ 17 border, April 2016). They shall not undergo verification, and automatically reach level 4 if more than 50 % of the company is in White possession. They can reach Level 3 when the company is owned by Black owners. The companies participate voluntarily, and only with the application for a public contract. 'Qualifying Small Enterprises' are companies with a profit margin from 5 to 35 million. This must be measured with 4 of the 7 criteria. The category they belong to can be specified by the company. Under 'Generic Enterprises' falls companies with a profit margin superior to 35 million, and they must earn points for each of the seven pillars.

In 2012 the Codes of Good Practice were revised, and the Broad-Based Black Economic Empowerment Amendment Act was adopted. The seven pillars were reduced to five by

merging positive measures in with management and personnel, as well as preferred sources and suppliers, and corporate development of small firms. This raised the threshold for qualifying participation. Exempted Micro Enterprises were now considered to have a turnover inferior to 10 million, Qualifying Small enterprises were then considered companies with a turnover between 10 to 50 million, and Generic enterprises began with a turnover of 50 million. Small businesses make up 95 % of all enterprises, 3 % are medium sized, while only 2 % are large enterprises.

In the public sector, participation in BEE is a duty. Thereafter, the supplier is selected, licenses are given, and State sales of possessions from public-private partnerships are decided. With the assignment of public orders, general criteria - like price, time, and quality – are all considered on the one hand, but on the other hand the BEE score of an enterprise is also considered.

Some small businesses feel compelled to hire black employees in order to not be excluded from potential customers, business partners, donors, and supply services. Since the BEE's core values also affect business partners, many opt for a voluntary verification and participation.

5. Material

Here is a list of further material on collection of equality and participation data. Some links to materials and documents are only available in German.

Publications

Simon, Patrick (2007): "Ethnic" statistics and data protection in the Council of Europe countries study report.

Makkonen, Timo (2007): Measuring Discrimination: Data Collection and EU Equality Law

Ringelsheim, Julie and De Schutter, Olivier (2008): "Ethnic Profiling: A Rising Challenge for European Human Rights Law"

Pücker, Mario (2009): "Ethnic monitoring as an instrument of anti-discrimination policy?"

Heinrich Boell Foundation (2009): Ethnic monitoring: data collection with or about minorities?

Peucker, Mario and Lechner, Claudia (2010): Feasibility study: "Standardized data collection to detect discrimination!?" - Review and Outlook

Press

Tagesspiegel.de (2005): "Police have software with gay's data".

Conferences, briefings, reports, etc. to data collection

Anti-Discrimination Agency (2009): Meeting documentation of technical discussion of the Anti-Discrimination Agency's "Standardized data collection to detect discrimination!?" - Review and Outlook - on 26/11/2009 in Berlin

Open Society European Policy Institute (2014): Background paper on "Equality Data Initiative".

German Institute for Human Rights (2008): Protocol of the technical discussion on "data collection was to demonstrate ethnic discrimination".