



EUROPEAN ROMA INFORMATION OFFICE

Survey on  
**Ethnic Data Collection**

*Risk or opportunity?*



## **Acknowledgements**

The book and the survey on ethnic data collection were initiated under the 3-year project "Promoting the Empowerment of Roma Communities", supported by the PROGRESS program and supervised by DG Employment and Social Affairs of the European Commission.

The framework and initial draft of the survey findings and analyses was written by Ivan Ivanov. The survey was conducted with the assistance of ERIO's interns Ewa Hajdukonis and Vincenzo Vianelli. Jenny Thomsen provided proof-reading and editing. Matteo Manzonetto undertook the layout and design of the book.

ERIO is grateful to the people interviewed for the survey as well as to its staff and network members who realised that such a survey would be helpful for our future advocacy work at national and European level and supported the publication of this book.

## **Table of contents**

Foreword	3
Introduction	5
What is ethnic data collection?	7
The need for data	8
Risks and concerns	13
Personal data and statistics	15
Survey on data collection among Roma	17
Aim of the survey	18
Survey findings	18
Trends of the survey	25
Recommendations	28
Examples of ethnic data collection	31
Ostrava School Segregation Case (Czech Republic)	31
Integrated plan for the Roma people in Catalonia	32
References	34

## **Foreword**

Data collection is increasingly seen by international and European bodies as a necessary tool in the fight against discrimination. However, despite the political will of states to commit to non-discrimination and the promotion of equality, there is a general lack of ethnic data across Europe. To some extent this can be explained by a lack of awareness of the benefits of equality data. There is also a common misunderstanding that existing data protection laws prohibit all ethnic data collection. Among the groups concerned there are also widespread concerns and fears concerning the collection of ethnic data. These obstacles need to be dealt with in order for ethnic data to function as an instrument in the fight against discrimination.

ERIO is of the opinion that ethnic data collection can be a valuable tool for combating discrimination against Roma, provided that safeguards are in place to avoid misuse of data. Ethnic data is necessary to identify the needs for special policies and to develop and assess the influence of such policies. It can help monitoring discrimination and the implementation of anti-discrimination law. Ethnic data can also be a useful tool in raising awareness about discriminatory practices as well as improving our understanding of discrimination as a phenomenon. However, a large part of the Roma community has not yet realized the need for data. Therefore, a presentation of the role and value of ethnic data as well as an analysis covering limitations and fears of collecting data is essential.

ERIO conducted a survey in 2008 in order to observe the attitudes towards ethnic data collection within the Roma community in Europe. The aim of the survey was to assess the level of awareness among Roma concerning ethnic data collection as well as to identify fears, concerns and

trusts regarding such data. The publication highlights and discusses the results of the survey. It provides a reflection on the different positions concerning ethnic data collection within the Roma communities that can be taken into account when designing Roma policies. The purpose of this publication is to inform stakeholders about ethnic data collection and to promote debate on data collection among Roma communities.



## **Introduction**

The issue of discrimination and the promotion of equality have emerged as a result of long historical processes. All Member States of the EU are today parties to the main human rights conventions prohibiting discrimination. In addition to this, the States have adopted national legislation banning discrimination. The adoption of two EU directives on equal treatment in 2000, the Racial Equality Directive and the Employment Equality Directive, has considerably strengthened the protection from unequal treatment in the EU. The Racial Equality Directive is the first European Union-wide legal prohibition of racial discrimination. One important difference between the two Directives and previous anti-discrimination law is that the emphasis is now placed on addressing institutional and societal patterns rather than individual behaviour and prejudice<sup>1</sup>. Thus the focus is put on group results in order to identify discrimination as well as to determine whether measures to combat discrimination have been effective. This is where data collection becomes crucial since access to relevant data is vital for these purposes. The Directives suggest that in order for discrimination to be established any means may be used “including on the basis of statistical evidence”<sup>2</sup>.

The Directives are not the only documents providing for statistics in combating discrimination. An example of a European Union organ that has highlighted the importance of data is the European Commission against Racism and Intolerance (ECRI), which states that “it is difficult to develop and effectively implement policies in the areas in question without good data” and governments should therefore “collect, in accordance with European laws, regulations and recommendations on data-protection and protection of privacy, where and when appropriate, data which will assist in

---

<sup>1</sup> Makkonen, Timo (2006) Measuring Discrimination - Data Collection and EU Equality Law

<sup>2</sup> EU Race Directive 2000/43/EC and the EU Employment Directive 2000/78/EC

assessing and evaluating the situation and experiences of groups which are particularly vulnerable to racism, xenophobia, antisemitism and intolerance”<sup>3</sup>. A more recent development is the European Union Minorities and Discrimination Survey designed by the European Union Agency for Fundamental Rights (FRA). This is the first ever EU-wide survey focusing on the experiences of discrimination among ethnic minority groups. The aim of the Survey is to address the lack of reliable data on minorities across Europe. FRA points out the need for data “to measure the extent of discriminatory treatment and criminal victimization, including racially motivated crime, experienced by minorities”<sup>4</sup>. It is fair to conclude that among international and European bodies there is a general agreement of the importance of ethnic statistics in anti-discrimination policy, provided that proper safeguards are in place<sup>5</sup>.



---

<sup>3</sup> ECRI General Policy Recommendation n°1: Combating Racism, Xenophobia, Antisemitism and Intolerance

<sup>4</sup> European Union Agency for Fundamental Rights (2009) EU-MIDIS at a Glance

<sup>5</sup> Goldston, James A. (2001) Race and Ethnic Data: A Missing Resource in the Fight against Discrimination

## **What is ethnic data collection?**

Data can be understood as any piece of information revealing some parts of reality and can thus be used for analysis or decision-making. The term data collection refers to “mechanisms for gathering information on the situation of disadvantaged groups”<sup>6</sup>. Equality data thus refers to information that can be used to analyze the state of equality. The purpose of equality data is to show information about a certain population or phenomenon and to be able to analyze or make decisions at a general, collective level. An example is where data reveals under-representation of a certain group in educational institutions and thus can be used in an indirect discrimination case or as a basis for positive action measures. This illustrates the fact that ethnic data do not aim to draw conclusions at an individual level. The information should be anonymous and not able to be linked to a specific individual<sup>7</sup>. In cases where the data can be linked to an identifiable person, it is called personal data. When making use of personal data, certain safeguards needs to be taken into account and these are specified in international and European data protection laws. However, anonymous data, originally related to an identifiable person, which has been processed in a way so that it can not be linked to the person, is not covered by these laws<sup>8</sup>.

Furthermore, taking a human rights approach to data collection, the individual shall be in control of data relating to her, including the personal choice to reveal one’s ethnicity. International human rights law is supporting the principle of self-identification, meaning that it is up to the individual to

---

<sup>6</sup> European Commission (2009) International Perspectives on Positive Action Measures

<sup>7</sup> Makkonen, Timo (2006) Measuring Discrimination - Data Collection and EU Equality Law

<sup>8</sup> European Commission (2006) European Handbook on Equality Data

decide which ethnic group to be identified with<sup>9</sup>. When considering ethnic data from a human rights perspective, it can be seen as an element of the right to be free from discrimination. If ethnic data were to be crucial as evidence proving discrimination, the victim should have the right to access such data. There is also the duty of governments to ensure equality, which should involve the collection and processing of ethnic data<sup>10</sup>.

There are a number of methods that can be used when collecting ethnic data. Official statistics, including censuses and surveys, are collected by states in order to get a picture of the demographic and socio-economic situation of their populations. Complaints data is another method of data collection which refers to data gathered by bodies handling discrimination complaints. Information includes numbers and types of complaints recorded at such bodies or organizations. Moreover, organizations and companies can make use of diversity monitoring in order to ensure that they comply with equality laws. This method allows organizations to examine how different groups are influenced by policies and practices<sup>11</sup>.

### **The need for data**

Despite the fact that legal and political commitment to the fight against discrimination is stronger than ever, discrimination continues at disturbing levels. There are many factors contributing to this state of affairs, one of them being a general absence of accurate data relating to discrimination. Discrimination is one of the most difficult rights abuses to prove, since

---

<sup>9</sup> Makkonen, Timo (2006) *Measuring Discrimination - Data Collection and EU Equality Law*

<sup>10</sup> Dimitrina Petrova (2004) *Ethnic Statistics*

<sup>11</sup> European Commission (2006) *European Handbook on Equality Data*

modern forms of discrimination are often subtle and hidden. Without proof the right to equality of treatment can not be justified, which suggests the importance of statistical evidence in identifying and challenging discrimination. International bodies have continually insisted on governments to present employment, housing, health, educational and other data broken down by ethnicity in their reports and yet few governments have provided sufficiently detailed data<sup>12</sup>.

It can be concluded that in order to battle discrimination a variety of measures, besides legal frameworks and political commitment, is needed. Measures such as identification and analysis of discriminatory patterns, monitoring of compliance with equal treatment laws, adoption of awareness-raising programmes and positive action measures, are additionally required in the struggle against discrimination<sup>13</sup>. Empirical evidence of discrimination is needed in order for such activities to fulfill their purpose. Policies and practices in political, administrative and business life have to be based on reliable data in order not to arrive at the wrong decisions. Exact documentation of the inferior position of ethnic minorities is fundamental to promoting non-discrimination throughout Europe. The present lack of reliable data leaves non-discrimination advocates limited in terms of proving and combating discrimination<sup>14</sup>.

Regarding **policy development, implementation and evaluation**, equality data is needed both at the national and European levels. Decision-makers need reliable information in order to make the right decisions. Without accurate information there is a risk that adopted programmes are

---

<sup>12</sup> Dimitrina Petrova (2001) in Kriszan A. (Ed.) Ethnic Monitoring and Data Protection

<sup>13</sup> European Commission (2006) European Handbook on Equality Data

<sup>14</sup> Goldston, James A. (2001) Race and Ethnic Data

irrelevant or damaging to the society and the minority<sup>15</sup>. To identify and conquer inequities in the fields of employment, education, health care and social services, data can be crucial<sup>16</sup>. Statistics can support policy making in terms of identifying needs and analyzing issues, offering guidance on means to deal with difficulties and evaluating whether implemented policies have been successful. With accurate analyses at the start, suitable policies can be developed and the rights of individuals will thus be better secured<sup>17</sup>.

Statistical evidence can play a vital role in the **legal assessment** of whether discrimination has occurred. Claimants are often in need of statistical evidence to support their claim before courts and other competent bodies, especially in cases of indirect discrimination. Indirect discrimination occurs “where an apparently neutral provision, criterion or practice would put persons of a racial or ethnic origin...at a particular disadvantage compared with other persons, unless that provision, criterion or practice is objectively justified by a legitimate aim and the means of achieving that aim are appropriate and necessary.”<sup>18</sup> The effect of the measure or practice will have to be established by a comparison on a group level rather than an individual level<sup>19</sup>. Since proving a disadvantage in comparison with other persons entails information of how different groups have been treated, the very idea of indirect discrimination involves a need for data. This is recognized in the two Directives, where it is provided that “rules of national law or practice...may provide in particular for indirect discrimination to be

---

<sup>15</sup> Katy Negrin, *Collecting Ethnic Data: An Old Dilemma, The New Challenges*

<sup>16</sup> Susanne Milcher and Andrey Ivanov (2004) *The United Nations Development Programme's*

*Vulnerability Projects: Roma and Ethnic Data*

<sup>17</sup> European Commission (2006) *European Handbook on Equality Data*

<sup>18</sup> EU Race Directive 2000/43/EC, Article 2(2)(b)

<sup>19</sup> Makkonen, Timo (2006) *Measuring Discrimination - Data Collection and EU Equality Law*

established by any means including on the basis of statistical evidence.”<sup>20</sup> Due to the fact that discrimination has become more hidden and subtle it is also less easy to prove. In some cases it is nearly impossible to establish discrimination without statistical evidence. As a result of the lack of information documenting discriminatory behaviour, victims of racial discrimination are extremely disadvantaged<sup>21</sup>.

Moreover, national specialized bodies and international monitoring bodies need equality statistics to be able to carry out their **monitoring** functions. All EU Member States are parties to the core human right conventions and are hence obliged to draft periodic country reports, and to include statistical information in relation to discrimination. It is difficult for governments to meet obligations to eradicate discrimination without data showing the impact of policies in the fields of housing, employment, education and criminal justice. Businesses, government agencies and other organizations can also make use of equality statistics to monitor compliance with equal treatment laws internally<sup>22</sup>.

Statistical data can also be vital in the development of **positive measures**. The term refers to the implementation of specific measures to prevent or compensate for disadvantages linked to discrimination. In order for Member States to take positive action there needs to be reliable information on discriminatory treatment<sup>23</sup>. An assessment of whether disadvantages exist that would justify positive action is necessary, which can be detected by the use of relevant data. Data also plays a crucial role in

---

<sup>20</sup> EU Race Directive 2000/43/EC and the EU Employment Directive 2000/78/EC

<sup>21</sup> Dimitrina Petrova (2001) in Kriszan A. (Ed.) Ethnic Monitoring and Data Protection – The European Context

<sup>22</sup> European Commission (2006) European Handbook on Equality Data

<sup>23</sup> Simon, Patrick (2007) Statistics for positive action: more than a tool, a duty

designing and performing special measures with the purpose of reaching equality<sup>24</sup>.

Finally, statistical evidence can significantly increase awareness among governments and the majority populations by providing convincing facts that can be used in national and local discussions on discrimination. Governments and NGOs can make use of the data in **awareness-raising**, advocacy and education. By making discrimination visible, inequality can be turned into a societal matter rather than an individual concern of the victims<sup>25</sup>.



---

<sup>24</sup> Makkonen, Timo (2006) Measuring Discrimination - Data Collection and EU Equality

<sup>25</sup> European Commission (2006) European Handbook on Equality Data

## **Risks and concerns**

Considering the need for ethnic data in addressing discrimination, how can the lack of data across Europe be explained? A lack of awareness about ways of collecting data as well as the benefits of equality data can to some extent explain this situation. Misunderstandings of what data collection involves in practice and what influence data protection laws have on data collection are other aspects to the present lack of data<sup>26</sup>. However, there are also other factors that are more difficult to address. Concerns and fears are raised by the very idea of collecting and processing data.

A common concern about ethnic data collection is based on the fact that data has historically been misused on several occasions. Situations where data systems have been used to target individuals or groups are used as arguments for not collecting data on ethnicity. Examples in Europe include the abuse of ethnic data during the Second World War where Jews, Roma and other groups were exterminated as well as abuse by some Communist regimes before 1989 in Central and Eastern Europe<sup>27</sup>. Experiences like this have understandably resulted in a fear among some members of minorities concerning the gathering of data. However, it should be kept in mind that in such cases we are referring to personal data, which can be linked to an identifiable person. Many forms of data collection are based on anonymous data and can thus not be used to target individuals. Furthermore, the misuse of data during the Second World War led to the creation of the contemporary human rights system with safeguards against such abuse. The situation has changed a great deal since then with the implementation of data protection laws and modern data protection

---

<sup>26</sup> Ibid

<sup>27</sup> Goldston, James A. (2001) Race and Ethnic Data

technology that can protect data from unlawful access and render data useless for unauthorized parties<sup>28</sup>.

These concerns are related to the identification of individuals which possibly could lead to further discrimination. Another common concern focuses on the use of statistics to stigmatize groups. Examples include intentionally confirming stereotypes by using statistics or publishing numbers without explaining and analyzing elements behind the differences. The best example is illustrated by the use of crime statistics where figures show that certain groups more often commit crimes. This spreads the stereotype that these groups have a more natural tendency to commit crimes. Ethnicity should be seen as an indicator to other features, such as social exclusion, low income etc. associated with crime and not as a threat in itself. Interpreted in the wrong way statistics can lead people to view social problems as a community burden rather than as difficulties for the persons concerned<sup>29</sup>. Nevertheless, ethnic statistics are not necessary in order to promote negative stereotypes or as a requirement for discrimination and prejudice. Authors can, for instance, publish the ethnicity of persons arrested for criminal activities without statistics. Political leaders do not need statistics in order to promote racial prejudice. It is rather the absence of accurate data that provides actors with possibilities to make use of data that does not accurately portray the situation of vulnerable groups<sup>30</sup>.

There are also fears concerning ethnic data among governments. Reluctance to engage in collection of data can be based on a fear that it will reveal uncomfortable facts and reflect badly on the governmental policies.

---

<sup>28</sup> Makkonen, Timo (2006) Measuring Discrimination - Data Collection and EU Equality Law

<sup>29</sup> Simon, Patrick (2007) "Ethnic" Statistics and Data Protection in the Council of Europe Countries

<sup>30</sup> Goldston, James A. (2001) Race and Ethnic Data

Such data will also demand improvement of the situation, which involves distribution of financial resources. This fear of data collection causes governments to avoid dealing with discrimination<sup>31</sup>.

### **Personal data and statistics**

In order to meet the legitimate fears expressed by the members of the groups concerned, international texts on data protection have been established. In 1981 the Council of Europe's Convention 108 for the protection of individuals with regard to automatic processing of personal data was adopted. Its main principles were later implemented into national legislation. The EU Data Protection Directive was adopted in 1995 and has been transposed into the EU member states legislation<sup>32</sup>. Since there is a concern that individuals can be identified and hence possibly discriminated against, the distinction between personal data and statistical use of data is crucial to make. The data protection laws are only applied in cases where personal data is involved. According to Article 2(a) of the EU Data Protection Directive, personal data is "any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity". Once personal data have been made anonymous in order to be used in statistics it does not constitute personal data anymore and are thus not covered by data protection laws. Statistical results used to document and prove discrimination do not constitute personal data. Most forms of discrimination

---

<sup>31</sup> Ramsay, Kathryn (2006) Disaggregated Data Collection

<sup>32</sup> Simon, Patrick (2007) "Ethnic" Statistics and Data Protection in the Council of Europe Countries

testing involve anonymous data and can thus not be linked to an identifiable person<sup>33</sup>.

In cases where personal data is involved, certain rules to protect the data must be observed. The right to privacy is guaranteed in several international human rights conventions and must be protected and interfered with only where truly necessary. In addition, the Council of Europe Convention 108 and the EU Data Protection Directive specify certain conditions for the processing of personal data and in order for ethnic data to be legitimate these requirements have to be fulfilled. The principles include data having to be obtained “fairly”, “lawfully” and “for specified and legitimate purposes”. Data must be “accurate” and “adequate, relevant and not excessive in relation to the purposes for which they are stored”. Furthermore, the data has to be rendered anonymous as soon as it has fulfilled its purposes so that it can no longer be linked to an identifiable person. All collection of personal data needs to fulfill these criteria as well as comply with certain restrictions such as notification of persons concerned, confidentiality and security of processing<sup>34</sup>. The aim of these laws is to protect individuals against data collection violating the right to privacy or used with wrong intentions.

There is as broad conviction that data protection laws, particularly the EU Data Protection Directive, prohibit ethnic data collection. The Directive provides a general prohibition of the processing of sensitive personal data, revealing racial or ethnic origin, but certain exceptions are specified to this rule. Conditions include where “the data subject has given his explicit consent to the processing of those data”; or “processing is

---

<sup>33</sup> Makkonen, Timo (2006) Measuring Discrimination - Data Collection and EU Equality Law

<sup>34</sup> Simon, Patrick (2007) “Ethnic” Statistics and Data Protection in the Council of Europe Countries

necessary for the purposes of carrying out the obligations and specific rights of the controller in the field of employment”. Furthermore, “subject to the provision of suitable safeguards, Member States may, for reasons of substantial public interest, lay down exemptions”. This gives the Member States the responsibility to decide whether ethnic data should be collected in order to be used as a tool in the fight against discrimination, given that the safety measures are guaranteed<sup>35</sup>. To conclude, the Directive does not forbid all ethnic data, but rather provides credible safeguards to protect individuals from data collection that could be used for the wrong purposes.

### **Survey on data collection among Roma**

In 2008, ERIO conducted a survey on the Roma in Europe in relation to ethnic data collection. The Roma community exemplifies an ethnic minority facing discrimination in all spheres of society. According to the 2009 EU-MIDIS report, the Roma report the highest levels of being discriminated against of the groups surveyed<sup>36</sup>. In recent years governments have paid increased attention to the minority by initiating special programs related to it. Such programs require accurate statistics in order to be successful. However, there is a general lack of data concerning Roma and where data do exist, it is often inaccurate<sup>37</sup>.

Attitudes among Roma towards ethnic data collection are divergent, therefore ERIO decided to conduct a survey on the attitudes of Roma organizations and individuals towards ethnic data collection in order to map

---

<sup>35</sup> Simon, Patrick (2007) “Ethnic” Statistics and Data Protection in the Council of Europe Countries

<sup>36</sup> FRA (2009) EU-MIDIS:Roma

<sup>37</sup> Council of Europe (2000) Roma and Statistics

out their experiences and opinions. It can be concluded that there are two contrasting positions among Roma on the topic. One position is the part of the Roma who are against any form of ethnic data collection, arguing that data have been used against the Roma in the past. Arguments based on the reinforcement of stereotypes are also common among persons in this position. The other position constitutes persons who find ethnic data necessary in the fight against discrimination<sup>38</sup>.

#### *Aim of the survey*

The aim of the survey is to assess the level of awareness among Roma concerning ethnic data collection as well as to identify fears, concerns and trusts regarding such data. A parallel objective is to bring these results to the attention of institutions and other stakeholders to be taken into consideration while designing Roma policies where data collection has an important role.

#### *Survey findings*

The survey was addressed to a large number of people. It was initially disseminated through the Roma information network which reaches more than 20 000 recipients. More than 300 Roma were directly approached and 102 answers were received. The questions were asked via e-mail, by phone and face to face interviews.

The survey was directed at Roma represented in different sectors including governmental and intergovernmental institutions and NGOs as well as persons outside of these sectors, such as journalists and freelancers. Roma from most of the countries in Europe were reached and mainly from

---

<sup>38</sup> Ibid

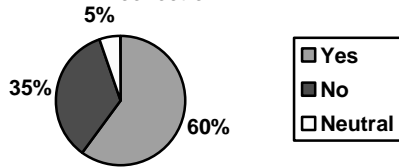
those with larger Roma population. Effort was made to keep the proportionality of Western and East European countries.

The survey was based on a short questionnaire with the following questions:

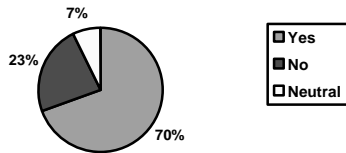
- 1) Do you support data collection on ethnic ground? Why?
- 2) Do you think ethnic data would be a good instrument to fight discrimination against Roma?
- 3) Would you agree to disclose your ethnicity for statistics and surveys in your country in relation to education, employment, health care, housing and social services?

Questions	1)			2)			3)		
	YES	NO	Neutral	YES	NO	Neutral	YES	NO	Neutral
Answers	58	34	5	59	19	6	59	12	8
Percentages	60%	35%	5%	70%	23%	7%	75%	15%	10%
Total amount of answers	97			84			79		

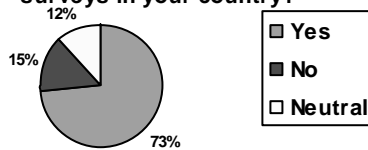
**Do you support ethnic data collection?**



**Do you think ethnic data would be a good instrument to fight discrimination against Roma?**



**Would you agree to disclose your ethnicity for statistics and surveys in your country?**



As can be seen in the table above, 60% of the responding informants to the **first question** answered that they support data collection, 35% that they do not support it and 5% took a neutral stand to the question. A high number of informants responded to the question with conditions. They argued that the answer is depending on who is in charge of the data collection, its purpose and how it is conducted. Several stated that the data collection has to be based on confidentiality and voluntary self-identification in order to make sure that it will not be used for the wrong purposes. Others expressed a wish for NGOs to be in charge of the collection rather than the state due to the concern that the data could be manipulated at the state-level. One example of this was expressed by an informant from Kosovo who stated that “ethnic data collection by the police or other state authorities such as housing departments, social services, etc. raises substantial security concerns and should be done by an independent unit”. Moreover, a couple of informants argued that Roma and Roma organizations need to have timely access to the collected data as well as an informatory role in the analysis of the data. In some cases, this insecurity about the purpose and processing of the data appears to be based on a lack of awareness among the informants concerning existing international, European and national rules on protection of privacy and data.

Several informants expressed a fear that the data could be used against them by authorities in terms of denying them access to work and public services. One informant from France stated that “data collection on ethnic ground can be extremely misleading and/or manipulated, in particular by governmental institutions”. Some based their arguments on own experiences where statistics have been used in a discriminatory manner. One informant exemplified this by explaining that in Denmark “the police use

the database on Roma in solving crimes". Others mentioned that Roma are frequently being discriminated against by employers.

There was a concern among some of the informants that ethnic data will be misused to harm Roma respondents. A couple of informants meant that even if the data is collected with good objectives it can be used later as a tool for discrimination. An example of this is where an informant from Austria stated that "even if someone does it with good intentions, no one knows who will use that data in the future (let imagine a government change after a few years)". The fear was expressed as well in relation to the increasing power of right-wing parties in some countries across Europe. Some said that ethnic data can be a positive instrument, but since nobody knows what the future will bring it is not worth taking the risk. This unwillingness to self-identify as Roma due to a fear of discriminatory practices is one of the most significant factors to the under sampling of Roma in surveys.

The concern that data can be used to support negative stereotypes of Roma was also raised. One informant gave the example of how a high unemployment rate can be used to argue that Roma are lazy and abuse the social system. In contrast, others mentioned that the data could be used positively to show the real crime rate in the Roma community which often is presented in a higher degree than what is reflected in reality. This illustrates the fact that ethnic data in itself is not the problem; it is rather the context of its use that raises concerns. Data can, by itself, be used for both legitimate and illegitimate purposes, which highlights the importance of creating a collection process which prevents the abuse of data rather than its gathering.

70% of the informants who answered the **second question** believed that ethnic data could be a good instrument in the fight against discrimination of Roma, 23% did not agree with this and 7% remained neutral to the question. It should be noted that the percentage of informants positive towards data collection is higher than in the first question. This can be explained by the number of informants who are skeptical towards data collection, but at the same time recognize its potential ability to serve as an instrument in the fight against discrimination.

Several informants expressed insecurity in relation to the role of ethnic data. On the one hand they said that it can function as an important tool, but on the other hand that it can be dangerous if used incorrectly. As in the first question, informants answered the question with conditions, pointing out that the answer depends on who is controlling the data, the aim of its collection and how it is performed. One informant from Romania answered the question regarding data collection as a tool for fighting discrimination by arguing that “If the data collection is used only to prove that Roma are criminals then the answer is negative”. This statement among others indicates that several of the informants were not accurately informed about data collection and its purpose. A large number appears to have misunderstood the concept and its capacity to function as a tool in the fight against discrimination.

There was a fear among a number of the informants that ethnic data will have opposite effects. Some informants argued that it will increase discrimination rather than fighting it. Others even claimed data collection in itself to be a sort of discrimination, which could lead to further stigmatization and segregation. An example of this was expressed by an informant from Italy who claimed that data collection “is very dangerous and useless. Data collection is against Roma and Sinti communities. It represents an act of

racism". Again, the reason behind the use of ethnic data collection and the fact that its aim is to fight discrimination is not always realized.

On the other hand, a number of informants noted the positive effects of data collection in terms of its ability to improve monitoring, implementation and evaluation of anti-discrimination measures. It is believed that its use will bring clarity to the number of Roma in need of housing, jobs, education, health care etc. Data collection is seen as a tool to measure progression and to analyze the impact of specific anti-discrimination policies. This was, for example, expressed by an informant from Bulgaria who stated that "data collection on ethnic ground is essential for better understanding the situation of the minorities - the cornerstone for creating a sound anti-discrimination policy". Furthermore, without comprehensive data, policies can not be adequately matched with financial resources. Some informants also mentioned that data collection is not only important at a governmental level, but also for the organizations working with Roma issues in terms of being able to better organize their activities. Additionally, one informant from Bulgaria stated that "I believe that the appropriate use of ethnic data will ensure the involvement of minorities themselves, a very important step towards an effective anti-discrimination policy".

Only a couple of informants mentioned the fact that data collection could be a useful instrument to effectively implement positive action measures. As well in terms of indirect discrimination only a small number pointed out the possibility to combat it with the use of ethnic data. As these are important factors behind data collection this indicates that there is a lack of awareness of its potential functions among the respondents.

Some of those supporting data collection explained the failure of the state to understand the importance of ethnic data for the fight against

discrimination. A few of them stated that they have made efforts to encourage officials to collect data, but that they have been met with reluctance towards initiating such a project.

Another obstacle in the survey is the misperception that personal data protection laws prohibit the gathering of all ethnic data. There is a common misinterpretation in some countries that international, European and national rules on protection of privacy and data categorically prohibit any kind of ethnic data collection, whereas this legislation merely calls for guarantees concerning the management of such data. One informant from Denmark stated that “Ethnic data collection is not allowed in Denmark, but local police uses it anyway, municipalities use it illegally, and insurance companies use it illegally”. Here a distinction has to be made between personal and anonymous data. The informant is referring to personal data which falls under the jurisdiction of data protection laws, meaning that such processing has to follow certain requirements. The failure to understand the difference between personal and anonymous data as well as the misinterpretation of data collection laws illustrate the importance of informing people as well as governments about the concept and laws protecting the data through its processing. Emphasis must also be placed on the creation of a secure and legal processing of ethnic data.

75% of the informants who answered the **third question** said that they would agree to disclose their ethnicity for statistics and surveys, 15% stated that they would not and 10% were neutral to the question. While the first and the second question required a position on a general issue, here the informants had to take a personal decision, whether they would identify themselves as Roma in statistics and surveys or not. Although most of the answers were explicit, in some cases informants elaborated on the answers with specific conditions as to when they would disclose their identity. Again,

main issues were whether the state authorities were to be in control of the data, whether it was to be used with good intentions and how the collection was going to be performed. The percentage of informants answering the question positively is higher than in the two other questions. This can be explained by the number of informants who argued that although they disclose their identities to surveys, they know and respect others who do not due to a fear of discrimination.

Some informants argued that hiding their identities gives them better chances for employment. Several said that even though they are proud of their identity, they do not want to disclose it due to the possible consequential discrimination. Their arguments were based on cases where Roma have revealed their identity with negative outcomes. A number of informants argued that they are themselves disclosing their identity, but they know that other members of the Roma community hide theirs out of fear of discrimination. As an example, one informant from Italy stated that “I would agree because I am proud of it, but I respect those who do not disclose it because that is a direct consequence of the high level of discrimination”.

In contrast, a few informants stressed the importance of revealing their identity. They argued that hiding it will only lead to more discrimination in the long run. Some stated that they would disclose their identity hoping that it would improve the situation of the Roma. Others argued that the reality of the problem is often denied, something that data collection would put an end to.

#### *Trends of the survey*

One observation that can be made from the survey is that a relatively large number of the informants appear to have misunderstood the

concept of ethnic data collection. One example is when a couple of informants stated that ethnic data in itself is a form of discrimination and that people should not be labeled, at the same time saying that they always disclose their identity because they are proud of it. Some mentioned that discrimination should not be fought by figures, but by policies aimed at counteracting discrimination. This indicates a lack of understanding of the function of data collection and the fact that one of its aims is to better adjust and monitor such policies.

It can be concluded that the negative answers dominate among the Western European Roma in comparison to those living in the Central and Eastern European countries. Almost 50% of the informants from the Western European countries are negative towards data collection as opposed to approximately 20% of the informants from Central and Eastern Europe. In general the negative stand is based on a fear of discrimination in case the data were to be misused.

In most cases the number of informants from each country is too small for a country-based generalization. However, there are some countries with a more substantial body of cases and hence some trends can be observed. The French informants constitute an interesting case where four out of five took a negative position to data collection. There is a general suspicion against the collection of data and who is in charge of it. One informant stated that "Most people in France are against ethnic statistics; Roma and Travellers tend to have the same opinions for good reasons". Also the Italian respondents were generally negative towards data collection. To the question if informants support ethnic data one informant from Italy stated "No, absolutely not. I consider it a form of discrimination". Other statements were based on a concern that data collection will lead to discrimination.

The correlation between country and attitudes was exceptionally strong in the Bulgarian case, which makes an interesting example of a country with overwhelmingly positive respondents. Out of 17 informants only one took a negative stand to data collection. A number of them stressed the importance of ethnic data as an instrument to combat discrimination. Moreover, all of the Macedonian respondents had a positive attitude towards ethnic data, seeing it as a necessity to reach an equal society. Also the Romanian informants constitute a larger part of the total amount of participants to the survey. In this case, however, 50% of the informants stated that they were negative towards data collection.

It is also interesting to note that some Roma immigrants from Eastern European countries living in Western Europe do not support data collection. The perception that they will be more discriminated if state administration, emigration authorities or employers would learn about their ethnic background makes them hide their ethnicity. One informant currently living in the Netherlands stated that "If I am not recognized as Roma I have better chances for employment and business as immigrant in the Netherlands. I have been living here for more than 30 years and we usually hide our identity".

In some of the cases where informants do not support data collection, arguments based on historical grievances were used. Informants referred to the Roma's experiences during the Holocaust and the fact that ethnic data was used in order to send Roma to internment camps. One example of this was given by an informant from Austria who stated that data collection "is a method well-known in the Roma community since 1940s and many are still afraid of it. I think it is immoral with any minority group". An informant from France also expressed the fear of data collection referring to

the experience during the Second World War, arguing that “during the Second World War the French collaborators of the Nazis used our ethnical data to send us to the 30 internment camps which were established all over France and North Africa”. The misuse of data during the Second World War is the most commonly mentioned historical occasion where sets of data have been used to target vulnerable groups, and this has resulted in a fear among the respondents of revealing their ethnic identity in surveys and censuses.

The results also show that the Roma involved at national and European level appears to have a better understanding of the use of data collection than those working at a grassroots level. The informants at a national and European level also tend to have a positive view on data collection, referring to its importance as a tool for fighting discrimination.

It can also be noted that informants working outside of the NGO sector are generally more reluctant to accept data collection as a useful instrument to fight discrimination. A number of them perceived data collection as discrimination in itself. Among these informants there was also a tendency to misunderstand the concept of data collection and its purposes.

### *Recommendations*

The results of the survey show that there is a broad misunderstanding of the concept of ethnic data and its purposes among the informants. Several of the informants failed to understand its capacity to function as an instrument in the fight against discrimination. Hence, if governments seriously want to make use of data collection they need to accurately inform people about the concept and explain clearly how the data will be used and stored in order to avoid misinterpretations and consequently disapproval of such data.

A large number of informants expressed insecurity about the purpose and processing of the data. First of all, informants seemed unaware of the use of anonymous data in statistics making it impossible for the information to be traced back to the individual. This suggests that the distinction between the collection of personal data and statistical use of data needs to be clarified. In some cases, the insecurity appears to be based on a lack of awareness about existing international and European rules on protection of privacy and data. The concerns brought up in relation to personal data processing are covered by the existing data protection laws, something that several informants appear to be unaware of. There is also a common misunderstanding in some countries that all ethnic data is prohibited by data protection laws, whereas these laws merely provide certain conditions under which information can be collected, registered, used and disseminated. The misinterpretation of the laws illustrates the importance of informing people about the concept and laws protecting the data through its processing.

There is a fear among the respondents that ethnic data can lead to even further discrimination of Roma. The percentage of informants who were unwilling to self-identify as Roma because of a fear of discriminatory practices is relatively high. This unwillingness to reveal one's minority background due to a fear of discrimination constitutes one of the most significant factors of the under sampling of Roma in surveys. Informants pointed out that data can be used for both legitimate and illegitimate purposes and there is a common skepticism towards data carried out on a state-level. The relatively high percentage of informants who expressed a fear of negative consequences of ethnic data shows that the concerns need to be taken seriously. It highlights the importance of ensuring that ethnic data can not be misused and that people are accurately informed about its

aim. The reduction of the risk of abuse needs to be prioritized and taken into account in all data collection.

It can also be observed that Roma involved at national and European level appears to have a better understanding of the use of data collection than those working at a grassroots level. They also tend to take a positive stand towards ethnic data as a tool to fight discrimination. In contrast, informants working outside of the NGO sector are generally more reluctant to accept data collection as a useful instrument to fight discrimination. Among these informants there was also a tendency to misunderstand the concept of data collection and its purposes. Cooperation with NGOs in relation to ethnic data can be a help in avoiding the mistrust among the Roma and it is therefore crucial for NGOs to be accurately informed about data collection.

In order for ethnic data to be successful in the fight against discrimination, accurate statistics need to be gathered and this can only be done by ensuring that people feel safe enough to reveal their ethnic identity. The results of this survey show that a large number of Roma do not. Special programs related to Roma that are based on unreliable labour, health, education and housing statistics broken down by ethnicity fail to reach their goal of addressing discrimination. Furthermore, where no reliable data is available, political actors can make use of estimations that do not correctly portray the situation of Roma in order to better fit the data into a particular political context. Governments need to redress this by explaining and demonstrating in practice the full benefits of ethnic data in ensuring policy efficacy, guaranteeing equality as well as raising awareness among the majority population.

### **Examples of ethnic data collection**

#### *Ostrava School Segregation Case (Czech Republic)*

On April 18, 2000, an application was filed with the European Court of Human Rights on behalf of eighteen Roma children claiming that their assignment to “special schools” for children with learning disabilities breached the European Convention. Data supplied by the applicants revealed that 56% of the total amount of the pupils in Ostrava that were placed in “remedial special schools” for the mentally disabled were Roma. This should be seen in light of the fact that only 2,26% of the total amount of the primary-school pupils were represented by Roma. Moreover, only 1,8% of non-Roma pupils were placed in special schools, in comparison to 50,3% of the Roma pupils. Thus, a Roma child is more than 27 times more likely to be placed in schools for the learning disabled than a similarly situated non-Romani child<sup>39</sup>.

Several independent specialized intergovernmental bodies also submitted information on the situation of the Roma in schools in the Czech Republic. According to data from the European Monitoring Centre for Racism and Xenophobia (now the European Union Agency for Fundamental Rights), more than half of the Roma children in the Czech Republic attend special schools. The Advisory Committee on the Framework Convention for the Protection of National Minorities observed in its report of 26 October 2005 that, according to unofficial estimates, the Roma represent up to 70% of pupils enrolled in special schools. ECRI noted that, in 2000, Roma children were “vastly overrepresented” in special schools.

According to the Court, “when it comes to assessing the impact of a measure or practice on an individual or group, statistics which appear on

---

<sup>39</sup> Case of D.H and others v. the Czech Republic

critical examination to be reliable and significant will be sufficient to constitute the prima facie evidence the applicant is required to produce". The Court stated that, although it is difficult to establish an exact number of Roma children in special schools, it is clear that the number was disproportionately high. The Roma pupils also constituted the majority of the pupils in special schools. The evidence was thus regarded as "sufficiently reliable and significant to give rise to a strong presumption of indirect discrimination"<sup>40</sup>. Consequently, the Court judged in favour of the applicants stating that discrimination has occurred. This case illustrates that, although the Court claimed that it is not necessary with statistical evidence to build an indirect discrimination case, without data it would be difficult to prove that persons of a particular ethnic origin are disadvantaged in comparison with other persons.

#### *Integrated plan for the Roma people in Catalonia*

Another example where data has been used was in the process of the formation of the integrated plan for Roma in Catalonia. The aim of the Plan is to combat the unequal situation of Roma in Catalonia as well as to recognize their specific culture. The proposals were based on the conclusions of the Study on the Roma Population in Catalonia, drawn up by the Department of Welfare and Family of the Generalitat de Catalunya in 2003. A particular difficulty for this study was the lack of reliable quantitative data. The population census contained no information relating to ethnic group and thus the study had to be based on statistics from previous research or by Roma organizations. However, it was impossible to give an exact number of Roma living in Catalonia using the existing data. Figures

---

<sup>40</sup> Ibid

varied as much as between thirty and sixty thousand. Some of the data from different organizations revealed a degree of coherence whereas in other cases there was a great distance between the figures. Generalitat de Catalunya stated that data can complement each other, but that there can be no guarantee as to their accuracy. Consequently, they stressed the importance of having access to reliable figures and that serious efforts need to be made in order to produce such data<sup>41</sup>.

This illustrates the importance of ethnic data when developing policies with the aim to combat discrimination. Since the proposals have been generated on the basis of the conclusions of the study it is of great importance that reliable data is available. Decision-makers need reliable information in order to make the right decisions. Statistics can show the real needs of a minority group and hence the right projects can be initiated. Without accurate data there is a risk that projects will not fulfill its aims.



---

<sup>41</sup> Generalitat De Catalunya (2006) Integrated Plan for the Roma People In Catalonia

## References

- Council of Europe (2000) *Roma and Statistics*, MG-S-ROM (2000) 13
- Dimitrina Petrova (2001) in Kriszan A. (Ed.) *Ethnic Monitoring and Data Protection – The European Context*, Budapest, CEU Press, p. 9-11
- Dimitrina Petrova (2004) *Ethnic Statistics*, Roma Rights, p. 7-13
- European Commission (2006) *European Handbook on Equality Data*, Directorate-General for Employment, Social Affairs and Equal Opportunities
- European Commission (2009) *International Perspectives on Positive Action Measures* Directorate-General for Employment, Social Affairs and Equal Opportunities
- European Commission against Racism and Intolerance (1996) *ECRI General Policy Recommendation n°1: Combating Racism, Xenophobia, Antisemitism and Intolerance*, Strasbourg
- European Union Agency for Fundamental Rights (2009) *Data in Focus Report – The Roma*
- European Union Agency for Fundamental Rights (2009) *EU-MIDIS at a Glance: Introduction to the FRA's EU-wide Discrimination Survey*
- Generalitat de Catalunya (2006) *Integrated Plan for the Roma People in Catalonia*
- Goldston, James A. (2001) "Race and Ethnic Data: A Missing Resource in the Fight against Discrimination" in Kriszan A. (Ed.) *Ethnic Monitoring and Data Protection – The European Context*, Budapest, CEU Press, p.19-41
- Makkonen T. (2007) Report for the European Network of Legal Experts in the non-discrimination field, Luxembourg, European Commission
- Milcher, Susanne & Andrey Ivanov (2004) "The United Nations Development Programme's Vulnerability Projects: Roma and Ethnic Data" in *Roma Rights*, Budapest, the European Roma Rights Center
- Negrin, Katy *Collecting Ethnic Data: An Old Dilemma, The New Challenges*, (<http://www.eumap.org/journal/features/2003/april/olddilemma>) Open Society Institute

Ramsay, Kathryn (2006) *Disaggregated Data Collection: A Precondition for Effective Protection of Minority Rights in South East Europe*, Minority Rights Group International

Simon, Patrick (2007) *"Ethnic" Statistics and Data Protection in the Council of Europe Countries*, European Commission against Racism and Intolerance (ECRI)

Simon, Patrick (2007) "Statistics for positive action: more than a tool, a duty" in *Putting Equality into Action: What Role for Positive Action?* European Commission - Directorate-General for Employment, Social Affairs and Equal Opportunities

#### Legal sources

Council Directive 2000/43/EC (Race Directive)

Council Directive 2000/78/EC (Employment Framework)

Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (Convention 108), 1981

Case of D.H and others v. the Czech Republic (2007) European Court of Human Rights *Application no. 57325/00*, Strasbourg







*The European Roma Information Office (ERIO) is supported by the European Community program for Employment and Social solidarity PROGRESS (2007-2013). The sole responsibility lies with ERIO and the Commission is not responsible for any use that may be made of the information contained therein.*



European Roma Information Office

***The European Roma Information Office (ERIO) is an international advocacy organization, which promotes political and public discussion on Roma issues by providing factual and in-depth information on a range of policy issues to the European Union institutions, Roma civil organizations, governmental authorities and intergovernmental bodies. ERIO cooperates with a network of a large number of organizations and acts to combat racial discrimination and social exclusion through awareness raising, lobbying and policy development.***

**ERIO**  
**Av. Edouard Lacomblé 17**  
**Brussels 1040**  
**Belgium**

**Tel: 0032(0)27333462**

**Fax: 0032(0)27333875**

**For more information: [press@erionet.org](mailto:press@erionet.org) / [www.erionet.org](http://www.erionet.org)**