



BRIEF ON ETHNIC DATA COLLECTION

Introduction

The aim of this brief is to start up a reflection inside the EURoma network about the issue of Roma data collection.

With a view to develop the network activity it will be necessary to progress in relation to our capacity to provide with more accurate data on how the EU Structural Funds are being effectively used in practice to address the situation of the Roma.

This first move aims at opening a debate to explore what each country can do in its particular environment as to this matter. From previous discussion it was clear that it is a general concern to improve the methods to identify where and how the interventions are taking place and what are their impacts on the Roma. At the same time, the controversy about the possibility of collecting data on a specific ethnic group still remains and it is also behind some of those concerns.

This first brief intends to highlight the current growing consensus about the need to collect ethnic data, to underline the need of progressing in this matter in order to improve policy effectiveness, to dispel doubts about the legality of this practice, and to open and stimulate an internal debate to explore ways forward for the future.

1. Background

There has been a growing debate around the need to collect data disaggregating ethnicity in the last years. Generally, activists and international organizations are clearly supporting and recommending to national governments to establish the grounds for collecting ethnic data. Nonetheless, governments have been often reluctant to this and have argued a number of reasons related predominantly with legal and moral considerations. The rationale for this has been generally the need to preserve the privacy of data against potential abuses, which have unfortunately occurred in the past under totalitarian regimes. Constitutional provisions and data protection laws have thus been claimed for not articulating data collection on Roma.

In general, there is a misperception that personal data protection laws prohibit the collection of ethnic data while there is still an insufficient awareness of the importance of ethnic monitoring for fighting against discrimination and social exclusion. Furthermore, it has been pointed out that in the case of some ethnic minorities there is sometimes reluctance from governments to carry out ethnic monitoring for avoiding shedding light on complex problems.



Beyond the possibility for abuse one can also see the advantages of having ethnic statistics. The surveys disaggregating ethnicity produce a wealth of statistics regarding such important questions as income, age, education, housing situation, and in some cases comparisons with non-Roma population. Policies must be based on robust data on the target group's situation and measuring the impact of policies can only be effectively done on the basis of accurate information.

In the case of Roma, a group severely affected by discrimination and social exclusion, there is a lack of relevant statistical data which prevents from monitoring performance and measuring the impact of both policies Roma-targeted and mainstream programmes. Reliable data on Roma is indispensable for policy making purposes and for monitoring performance as regards social inclusion and discrimination. In the case ESF Programmes which seek to have an impact on the conditions of Roma this problem is thus central.

2. EU legislative framework on personal data protection

Generally, international texts on data protection, freedom of information and respect for privacy refer as "sensitive" categories such as ethnicity, race, religion, etc., given that historically the recording of this characteristics has had negative effects on the individuals, especially under totalitarian regimes.

Two main types of regulations concern the issue of ethnic data and ethnic monitoring:

- The ones related to data protection which preserve privacy and regulate the dissemination of information, etc.;
- The ones related to statistics which regulate the compiling of statistics and collection of data.

The main EU instrument regulating this is the **Council Directive 95/46** on the protection of individuals with regard to the processing of personal data.

Article 8.1 of the Directive sets that 'Member States shall prohibit the processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life'.

Exemptions to this prohibition are also provided by the Directive. This prohibition shall not apply, amongst others where:

- **the data subject has given his explicit consent to the processing of those data**, (except where the laws of the Member State provide that the prohibition referred to in paragraph 1 may not be lifted by the data subject's giving his consent) (art. 8.2a);



Structural Funds: Investing in Roma

- processing is necessary for the purposes of carrying out the obligations and specific rights of the controller **in the field of employment law** insofar as it is **authorised by national law providing for adequate safeguards** (article 8.2b);
- processing is necessary **to protect the vital interests of the data subject** or of another person where the data subject is physically or legally incapable of giving his consent (Article 8.2c);
- processing is carried out in the course of its legitimate activities with appropriate guarantees **by a foundation, association or any other non-profit seeking body** with a political, philosophical, religious or trade-union aim and on condition that the processing relates solely to the members of the body or to persons who have regular contact with it in connection with its purposes and that the **data are not disclosed to a third party without the consent of the data subjects** (Article 8.2d);
- the processing relates to data which are manifestly made public by the data subject or is **necessary for the establishment, exercise or defense of legal claims** (Article 8.2e);
- processing of the data is required **for the purposes of preventive medicine, medical diagnosis, the provision of care or treatment or the management of health-care services**, and where those data are processed by a health professional subject under national law or rules established by national competent bodies subject to the obligation of professional secrecy or by another person also subject to an equivalent obligation of secrecy. Article 8.3.
- **Subject to the provision of suitable safeguards, Member States may, for reasons of substantial public interest, lay down exemptions in addition to those laid down in paragraph 2 either by national law or by decision of the supervisory authority.** Article 8.4.
- Processing of **data relating to offences, criminal convictions or security measures** may be carried out only **under the control of official authority**, or if suitable specific safeguards are provided under national law, subject to derogations which may be granted by the Member State under national provisions providing suitable specific safeguards. However, a complete register of criminal convictions may be kept only under the control of official authority. Article 8.5.

Therefore, there are three crucial issues:

- there must be a explicit consent of the subject to the processing of the data;
- there must be adequate safeguards:
- reasons of public interest are a basis for deciding to collect sensitive data.

The European Commission Communication on the application of Directive 2000/43/EC (COM (2006) 643) highlights the crucial role played by statistics in activating anti-discrimination policies and increasing their capacity to ensure social cohesion and promote diversity and equality. Furthermore it takes position as to the controversy of the prohibition of the collection of this data:



“There have been objections to the collection of such data on the grounds that it would breach the provisions of the EU Data Protection Directive. This does not reflect the true situation. The Directive generally prohibits the processing of sensitive personal data. However, certain exemptions to this rule are provided for, including where “the data subject has given his explicit consent to the processing of those data”; or “processing is necessary for the purposes of carrying out the obligations and specific rights of the controller in the field of employment”.

3. National Practices

There are differences between national practices. Generally laws protect individuals against the collection and processing of data which may be used for wrong purposes or violate their privacy. In all countries legal provisions restrict the processing of this sensitive data but national legislation does not generally prohibit the collection of this type of data, but restricts it and makes it conditional on the respect of certain safeguards. With some minor differences of content all countries include exceptions to the prohibition of ethnic data processing in line with the provisions of the Directive, and all countries have bodies safeguarding personal data protection.

As it was pointed out above, it has been common that governments oppose practices of ethnic monitoring with the justification of the contents of the domestic legislation, while others (e.g. actors in the fight against discrimination, NGOs, etc.) state that this is actually a narrow interpretation of the law or an excuse for not doing something that for historical or political reasons appears as controversial.

Main sources of ethnic data are thus:

- official socio-economic statistics, (administrative records, census data and data from official surveys);
- research: sociological surveys, attitudes surveys;
- Complaints data: NGO reports, Ombudsman or Equality bodies reports.

Census data

Some countries provide with census data with disaggregated data on ‘ethnic group’ and census forms include fields related to sensitive personal data such as race, ethnicity, religion or language. Poland, Romania, Slovakia, Bulgaria, Czech Republic, and Hungary include in different forms the question about ‘ethnic group’. Generally the answer to these questions is optional. Sometimes these are open fields (e.g. Czech Republic) while in other cases there is a list of options given where the subject must tick in a box (e.g. Bulgaria, Romania). Nonetheless, the completion of these fields is normally optional.

The collection of data through a national census is thus accepted by several governments. However, in the case of Roma it is widely accepted that the figures provided by census data in the countries this information is gathered are much lower than the actual figures.



Official surveys

Amongst surveys which allow a comparison of the condition of Roma and the rest of the population are the Households Budget Surveys and Labour Force Surveys. However, these surveys often fail to include a representative sample of minorities among the respondents, so that even if there is a question on ethnicity so that the data can be disaggregated, the results will not be representative for the minority. More representative samples of Roma are thus necessary to make these surveys more accurate for this group.

When the disaggregation of Roma in the sample of national surveys is not possible another experience (i.e. Spain) has consisted in replicating labour or households surveys on a sample of Roma which results in indicative gaps between the group studied and the overall population.

Other methods for reaching potential data subjects

Several methods can be used by researchers or administrative bodies for reaching potential data subjects: using name and address lists; using information from intermediaries (community leaders, schoolteachers, etc.); or using personal knowledge and contacts of NGOs, Associations. Lists of non-ethnic categories, which might contain an important representation of Roma, can be used (people with street vendor licenses for example) and personal contacts are used to remove non-Roma names from the lists.

As far as guarantees for protecting information privacy and ensuring data protection are concerned, throughout the processing of ethnic data in the sociological studies data from the survey forms are incorporated in computer programs eliminating names and addresses. The original survey forms are stored for a period of time and then destroyed.

In addition, the collection of ethnic data poses methodological problems as to the identification of the subjects.

4. The position of other international and European actors

The international and European institutions have taken position as regards data protection and the possibility to produce ethnic statistics pointing to a growing consensus on the need to produce such data.

Furthermore, “subject to the provision of suitable safeguards, Member States may, for reasons of substantial public interest, lay down exemptions”. Thus, it is for the Member States to decide whether or not ethnic data should be collected to produce statistics for combating discrimination, provided that the safeguards set out in the Data Protection Directive are respected.”



Structural Funds: Investing in Roma

In 2007 the European Committee against Racism and Intolerance (ECRI) of the Council of Europe issued the special report “Ethnic Statistics and data protection in the Council of Europe Countries” which advocates for the production and use of ethnic data in order to combat discrimination.

The United Nations Development Programme (UNDP) has made a contribution to the collection of ethnic data on Roma in the context of the Millenium Development Goals. UNDP collected data on the status of vulnerable groups in the region, comparing the situation of Roma, non-Roma living near Roma and IDPs and refugees for use by different bodies, including governments.³⁰ The UNDP has also coordinated an ‘Experts Group on Data and Measurements’ within the Decade of Roma Inclusion to identify gaps in existing data and suggest ways of improving the situation

5. Concluding remark. Implications for EU Structural Funds implementation.

It is clear that processing of ethnic data is not unlawful. It is sensitive information that must be subjected to safeguards, but there is a wide consensus between relevant actors that making progress in this field is one of the biggest challenges to improve the policy-making and effectiveness of implementation of Roma-related policies and programmes.

In the context of the implementation of EU Structural Funds for the benefit of Roma all this debate have important implications in regards to improve monitoring and evaluation mechanism. There is an urgent need to verify whether the Operational Programmes activities are actually implemented in accordance with the original plan and what are the actual results and outcomes, which will allow redesigning approaches, reseting priorities, and reallocating resources when necessary.

In the context of the implementation of the EU Structural Funds there is a need to collect better data on:

- Roma-targeted interventions contained in the relevant Operational Programmes where the Roma are explicitly defined as a target group and as the intended final beneficiaries.
- other interventions where Roma are not the only specific target group amongst other vulnerable groups
- the impact of other activities (specially infrastructure, urban regeneration, etc) on the conditions on Roma.

In line with the aims of the EURoma network, it is crucial to stimulate debate, exchange and improve practices and remove obstacles in order to produce more and better data on the performance and implementation of the actions for Roma.