

**The Collection of Ethnic Data:  
Outputs of Working Seminar on  
Sharing Good Practices in Ethnic Data Collection**

Ministry of Education, Youth and Sport  
Karmelitska 7, Prague, Czech Republic  
Tuesday 5 October 2010



**LIGA LIDSKÝCH PRÁV**

*In 2007 the League of Human Rights (LIGA) elaborated a detailed analysis on the possibilities of collecting ethnic data in the Czech Republic<sup>1</sup>. Regarding the fact that we do not witness almost any improvement of debate on this issue in our country or other countries participating at the international initiative Decade of Roma Inclusion, we have prepared an English translation of our still up-to-date Policy paper from 2007 in addition to more concise material. The purpose of this material is to present a possibility to initiate the collection of ethnic data in the Czech Republic.*

*Seeing the persistent difficulties with the data collection in majority of the countries of our region<sup>2</sup>, we also organized an international workshop called "Sharing good practices in ethnic data collection". The workshop took place on the 5th October 2010 in Prague with the support from the foundation Open Society Fund.*

*These outputs were written by Michaela Tetrevova, in cooperation with Iva Pikalova.*

**League of Human Rights' office:**

Liga lidských práv  
Burešova 6  
602 00 Brno  
Czech Republic

Tel. +420 545 210 446

Fax +420 545 240 012

[brno@llp.cz](mailto:brno@llp.cz)

[www.llp.cz](http://www.llp.cz)

*Supported by the Open Society Fund, Prague  
through East East Program – Partnership Beyond Borders.*

Nadace

OPEN SOCIETY FUND PRAHA

---

<sup>1</sup> Policy paper of the League of the Human Rights no. 4: Ethnic Data Collection as a Tool for Discovering the Extent of Discrimination Against Roma Children. Available on [http://www.llp.cz/files/file/sber\\_etnickych\\_dat.pdf%20](http://www.llp.cz/files/file/sber_etnickych_dat.pdf%20) (27 September 2010).

<sup>2</sup> More information on the unsatisfactory practice of the collection of the ethnic data in other countries participating on the Decade of the Roma Inclusion in the publication „No Data – No Progress“ – available on [http://www.soros.org/initiatives/roma/articles\\_publications/publications/no-data-no-progress-20100628/no-data-no-progress-20100628.pdf%20](http://www.soros.org/initiatives/roma/articles_publications/publications/no-data-no-progress-20100628/no-data-no-progress-20100628.pdf%20) (27 September 2010).

Working Seminar on Ethnic Data Collection was held on 5 October in premises of the Ministry of Education, Youth and Sports. The seminar was organized by League of Human Rights (LIGA) within the international initiative Decade of Roma Inclusion 2005 – 2015 and with the support of the Open Society Fund – Prague.

In the seminar actively participated representatives from Albania, Bosnia and Herzegovina, Romania, Hungary, Montenegro, Ukraine, Macedonia and the Czech Republic.

The seminar was split into two parts. In the first part, several presentations have been taken and some questions were asked. The second part was organized as a workshop in order to enable participants to actively join the discussion.

The purpose of the seminar was sharing good practises in ethnic data collection. Most of the countries, which are part of the Decade of Roma Inclusion 2005 – 2015, do not collect ethnic data. Lack of data makes it impossible to measure effectiveness of anti-discrimination policies. Therefore, it is not possible to say whether states fulfil their obligations to combat discrimination against ethnic minorities. Thus, the main aim of the seminar was to find a consensus on the way how the data should be collected in practice.

Participants were welcomed by **Mr. Viktor Hartos** from the Ministry of Education, Youth and Sports. After his opening speech a presentation part of the seminar commenced.

## **I. Presentations**

The goal of the series of presentations was to introduce to participants experiences of ethnic data collection in the Czech Republic and in Romania.

First speaker was **Michaela Tetrevo**va from the hosting organization, League of Human Rights. She provided an overview of current legal status of ethnic data collection and disproved the myth of legal barriers which allegedly hinder the collection. In the second part of the presentation she proposed how the data could be collected at schools and referred to Policy paper<sup>3</sup> prepared by LIGA, where could be found more details. She also emphasized that an immediate start with data collection is necessary because in order to gain any results, the data must be collected for several years. Therefore, time shouldn't be wasted.

A presentation from **Lucie Obrovsk**a, from the Office of Public Defender of Rights, followed. She presented the role of Public Defender of Rights in ethnic data collection. Firstly, she introduced the competences of ombudsman and highlighted new powers, which stem from the new Antidiscrimination Act. She raised the question of legal admissibility of ethnic data collection. She also mentioned that the Office for Personal Data Protection does not support collection of this kind of sensitive data. Nevertheless, the new ombudsperson feels that collection of ethnic data is crucial in terms of combating discrimination against ethnic minorities. Nevertheless, the new ombudsman thinks that his office is not the right authority which should be responsible for data collection due to the fact that the department for equality treatment is consisting just from lawyers.

**Pavla Polechova** from the Ministry of Education, Youth and Sports, firstly emphasized the necessity of the collection of ethnic data at schools due to the persisting discrimination of Roma children in

---

<sup>3</sup> Accessible at: <http://www.llp.cz/en/policy-papers/policy-paper-1-p1>

education. She demonstrated the discrimination on the Strasbourg's case of D.H. and others against the Czech Republic, which regarded 18 Roma children who were placed without proper reason into special schools. She also mentioned the recent open letter of James Goldstone, advocate of D.H., to the Czech government, which criticizes that nothing has been changed in this field so far. Afterwards, she expressed her opinion, that the disaggregated data should be collected instead of aggregated, as LIGA suggests in its Policy paper<sup>4</sup>. She added that if disaggregated data are collected it is possible to monitor anything by ethnicity, e.g. the wellbeing of the child. Moreover, *'it is possible to make from the disaggregated data aggregated data later on, but not conversely'*, she added.

Subsequently, **Karel Cada**, from GAC s.r.o., focused his speech on his practical experiences with gathering and collecting ethnic data. He explained the method of self-identification and external classification. He stated that from several reasons the method of external classification is largely sufficient for collection of ethnic data at schools. However, Cada emphasized that not even this methodology is unproblematic. In practice, sometimes there is no clear distinction between Roma and non-Roma. Therefore, he recommended using some additional characteristics as control variables. Finally, he stressed that for successful research it is necessary to explain the importance of the research to participating schools.

**Jana Strakova** from ISEA (Institut for Social and Economic Analysis) was talking about the ethnic research at Czech schools. She stressed that there is no national assessment in the Czech Republic and thus the sources comes from international comparative studies only. ISEA was inspired by OECD study called PISA 2009 dealing with 15 year olds at schools. They did their own study at 15 Czech schools with different representation of Roma children. For identification of Roma pupils they used 2 methods: 1) identification by teachers and 2) questionnaires on the language spoken at home. According to this research, very low amount of pupils declared that they speak Roma language at home. Therefore, this method proved to be unreliable for collecting the ethnic data, at least in the Czech Republic.

The first part of the working seminar was closed by **Adrian Dan**, from the University of Bucharest. He presented the research which was aimed on determining the number of Roma people in Romania, where is living the biggest Roma population in Europe. The private researchers used an interesting methodology for collecting the ethnic data. They contacted formal and also informal leaders of Roma communities in 2700 locations, whose aim was to ascertain who is deemed to be Roma in the particular location. According to their experiences, especially the cooperation with the informal leaders was very fruitful. Due to this cooperation 90% of pre-identified Roma people identified themselves as Roma. Nevertheless, later on in national census in 2002 less than half of this people declared themselves as being Roma, which shows that this minority is still afraid of collecting the ethnic data due to the events of WW2.

## II. Workshop and final debate

After a short break, participants were divided into three working groups. In those groups they discussed the topic of ethnic data collection and tried to find answers on laid questions. At the end of the workshops the answers were put together and examined in final debate.

The questions were following:

---

<sup>4</sup> Accessible at: <http://www.llp.cz/en/policy-papers/policy-paper-1-p1>

- 1) Suggest any particular tools for ethnic data collection.
- 2) How to secure the ethnic data collection?
- 3) How would you prepare informative campaigns?
- 4) Can you find any obstacles of ethnic data collection? How would you overcome them?
- 5) What kind of legislative changes would you suggest?

In each group the discussion was firstly about the topic itself and general comments were made. After a slower start, more heated and deepened discussion about several issues occurred.

First and foremost, participants agreed that in many countries, including the Czech Republic, there is **no political will** to implement the collection of ethnic data in practice and that this topic would deserved more attention.

Subsequently, participants tried to find a consensus about the way how the data should be collected. There are two possible ways – data could be collected aggregated (i.e. statistical data) or disaggregated (i.e. personal data). Both ways have its pros and cons.

Although disaggregated data offer greater details about their subjects and therefore enable monitoring influence of ethnicity on school success, there is an enormous fear of misuse of the data, particularly amongst Roma people. On the other hand aggregated data does not face such problem. It is also easier to collect and process them. Moreover, in some countries there is even no need to change the legislation, because current law does not prevent the collection of aggregated ethnic data. Collecting statistical data enjoys also better political support than collection of disaggregated data.

Therefore, it was concluded by the majority of participants that **collecting statistical data should be sufficient**.

Consequently, the issue of methodology has been broadly discussed. In general, there are two methods how the ethnic data could be collected. These are methods of self-identification and identification by others. In this field the views of participants differed most. Whereas the biggest asset of the method of self-identification is the fact that it is on voluntary basis (i.e. it is up to the subject if he/she declare his/her ethnic background), it usually fails on providing real numbers. For example, method of self-identification was used in the Czech National Census in 2001, when less than 12 thousands persons declared themselves as Roma. However, the government estimated that there were from 150 to 300 thousand Roma living in Czech territory.

Therefore the method of identification by others seems to be more suitable. However, this method is somehow controversial, because it does not provide free choice of ethnic affiliation to the subject. On the other hand, discrimination on ethnic grounds is usually based on other people's perception of the subject. In other words, if the child is by his bullies perceived as a Roma, he is bullied even if he is not Roma.

Therefore, it was concluded by the majority that an **emphasis should be put on self-identification**. But this method should be **accomplished by method of identification by others** at least at first years of the data collection in order to get reliable numbers.

The issue of methodology was very tightly connected to the topic of **awareness and informative campaigns**. Almost all of the participants agreed, that campaign should be aimed at several target

groups. First group should be Roma people themselves, because they fear the ethnic data collection at most. Other group should be public in general and government and other state bodies in particular. The main goal of the campaign should be to convince people about the importance of ethnic data collection and its security.

Last but not least, participants discussed **which state body should be responsible** for ethnic data collection. This issue differs in several countries. Representatives of LIGA suggested that in the Czech Republic, Public Defender of Rights might be responsible for ethnic data collection. However, this proposal has been met with great disapproval on the side of representative from the Ombudsman's Office. Nevertheless, she expressed support for ethnic data collection and offered to help with promoting this policy issue. Also other participants mostly agreed that Office of Ombudsman is not the right state body for ethnic data collection. It was suggested that authority responsible for collecting the data should be relevant ministry. Therefore, if the data will be collected at schools, the Ministry of Education would be responsible for processing and administering it.

### **III. Conclusions**

The seminar could be considered as successful. A debate about some practical aspects of ethnic data collection has been triggered and good practices have been shared. Although it has shown that finding unanimous consensus is impossible due to the controversies of this topic, at least majority of participants agreed on following points:

- without collecting ethnic data we cannot reach any progress and the question about effectiveness of various affirmative action stays unanswered
- it is necessary to raise political motivation for ethnic data collection, because so far it is at the periphery of political interest
- collection of statistical ethnic data is sufficient
- the emphasis should be put on method of self-identification, but at least at first years this method should be accomplished by method of identification by others
- awareness campaigns must target on several groups and they should convince the respondents about importance and security of ethnic data collection
- body responsible for ethnic data collection should be relevant ministry