

REGIONAL METHODOLOGY ON MAPPING OF ROMA HOUSING

Authors: Boris Žerjav and Nemanja Batrićević

Disclaimer: The views expressed in this document do not necessarily reflect the views of the Regional Cooperation Council or of its participants, nor of the European Union and the Open Society Foundations.

- I Proposal -



Table of Contents

1. Introduction	4
2. Overview of existing information on Roma housing in the region	5
Albania	5
Bosnia and Herzegovina.....	6
Kosovo*	7
Montenegro	8
The Republic of North Macedonia	8
Serbia	9
Summary	10
3. Key issues for the mapping methodology	11
Purpose of the mapping	11
Object of mapping: substandard Roma settlements.....	12
The definition proposed for mapping.....	15
Applicability of the definition in Western Balkans economies.....	16
4. Data on the substandard Roma settlements	18
Volume of mapping and data to be collected	18
Data quality and sources of data	19
Process of data collection	21
5. Data Format and Integration Procedure	25
Format and dimensions of spatial objects.....	25
Dataset integration	27
Software solutions for data analysis.....	29
6. Specific issues to be defined by each individual economy	31
Census methodologies and data by economy.....	31
Satellite images and ortho-photos	32
Criteria for a position of a substandard Roma settlement relative to formal settlement	32





List of acronyms

CSO	Civil Society Organisation
EC	European Commission
EU	European Union
GIS	Geographic Information System
IPA	Instrument for Pre-accession Assistance
MICS	Multiple Indicator Cluster Survey
UNDP	United Nations Development Programme
UN	United Nations
WB	Western Balkans





1. Introduction

The Roma Integration 2020 action is implemented by the Regional Cooperation Council and supported by the European Commission and the Open Society Foundations. It is implemented in the enlargement region (Western Balkans and Turkey), including Albania, Bosnia and Herzegovina, Kosovo*, Montenegro, the Republic of North Macedonia, Serbia and Turkey.

The action contributes to the reduction of the socio-economic gap between the Roma¹ and non-Roma populations in the Western Balkans and Turkey, by increased rate of implementation of and budgetary allocations for the policies for Roma integration, demonstrated through improved monitoring and reporting. The Roma Integration 2020 action assists the governments to maintain coordination and institutional functionality, formulate and implement effective measures, reduce biased attitudes, regionally share knowledge and skills and set common standards. It provides policy guidance and technical support for the policy process and implementation, dialogue on Roma integration at central and regional level, and link to the EU integration process.

The key development within the Roma Integration 2020 action was the adoption of the [Declaration of Western Balkans Partners on Roma Integration within the EU Enlargement Process](#) (Declaration). The Declaration was adopted by all the Prime Ministers from the Western Balkans on the 5 July 2019 in Poznan. Thereby concrete quantified targets for Roma integration are set in the priority areas, including housing, that the economies in the region are expected to achieve by the time they join the EU.

Because of the lack of reliable and relevant data in the area of housing, the target is set to **legalise all the informal settlements where Roma live** wherever possible, or alternatively provide decent, affordable and desegregated housing for Roma living in settlements that cannot be legalised for justified reasons. The Declaration also commits the governments to **use the available data** in order to formulate and implement proper policies, and to establish proper **mechanism for monitoring and reporting**. In the sense of the target in housing, using the available data would mean gathering and systematising the existing data in a way to provide baseline for formulating policy, in a system that would later also provide for proper monitoring.

Gathering the data should be the first step in the process of legalisation, that needs to be followed by **adopting urban plans** for the housing areas where Roma live for all those settlements where legalisation is possible, as well as by **providing support to Roma** in the process of legalisation and finally **actual legalisation** of the housing units of Roma.

Consequently, the reason to propose the present regional methodology for mapping housing of Roma is to make use of the existing data to the maximum extent possible, without creating significant need for additional data collection. The existing data should be gathered and organised in a systematic way. Implementing the methodology would initially provide the needed baseline data, and its regular updating can serve the purpose of monitoring the impact of the policies for Roma integration in the area of housing and the progress towards the achievement of the housing target in the Declaration.

* This designation is without prejudice to positions on status, and is in line with UNSCR 1244/1999 and the ICJ Opinion on the Kosovo declaration of independence.

¹ The term “Roma” within the action is used in line with the EU terminology, as an umbrella term which includes groups of people who have more or less similar cultural characteristics, such as Roma, Ashkali, Egyptians, Dom, Lom, Sinti, Travellers, Kalé, Gens du voyage, etc., whether sedentary or not.





Therefore, the proposed methodology firstly provides an overview of the existing data in the Region, followed by a argumentation on the key aspects regarding the methodology, and detailed elaboration of the proposed methodology, including the technical details required to establish a database, gather all the relevant data and systematise those data in a system.

2. Overview of existing information on Roma housing in the region

Collection and systematisation of data on the economic, social and cultural rights, among which is the right to housing, are perceived by the UN and the EU as a key in the process of improvement of the access to these rights for all, including the Roma population. Therefore the governments are required to collect data regularly in an aggregated manner, among other in relation to ethnicity.

In the *EU Framework for National Roma Integration Strategies up to 2020*, the European Commission recommended to the governments that, when developing national strategies, they should, among other measures, “**map micro-regions or segregated neighbourhoods where the Roma communities are the most deprived, based on the existing socio-economic and geographical indicators**” and “**define methodology for the monitoring and evaluation of the impact of measures aimed at Roma integration, and define mechanisms of Strategy revision**”.

However, data on the socio-economic, and particularly the housing situation of Roma in the Western Balkans region are generally scarce, outdated, unsystematised, collected partially, often on the basis of unstandardized and unverified methodologies. Most of the times, data are neither coming from official sources nor officially recognised, and as such are often disputed, without offering official alternative. Such lack of data in the area of housing is particularly consequential as it has been recognized that resolving the issues in housing, due to its impact on health and education, is a major step towards breaking the vicious cycle of exclusion.²

The Regional Roma Survey implemented periodically,³ is the most reliable source of information on the situation of Roma in the region. But, unlike the other areas (employment, education, health and crosscutting) the data collected in housing are limited. The survey provides valuable data on the rates of Roma dwellings connected to basic infrastructure and on overcrowding of the dwellings. However, it is lacking the crucial data on security of tenure, in the region generally defined by the ownership status and thus the status in the process of legalisation. Legalisation status of the dwellings is crucial for any development programme, because it is unlikely to invest in urban/housing development that is not legal or at least has a perspective to be legalised. The survey does not provide other important data in housing also, such as homelessness rate, safety of the dwelling, basic household equipment, etc.

The following is the overview of existing data by economies in the Western Balkans, from the information made available for this publication.

Albania

² OSCE/ODIHR (2003). *Roma Housing and Settlements in South-Eastern Europe: Profile and Achievements in Serbia in a Comparative Framework – Summary and Recommendations*.

³ Regional Roma Survey 2003, 2011 and 2017 series, implemented by the UNDP, the World Bank and the European Commission. The 2017 Regional Roma Survey is the third survey in a series, implemented by the UNDP, the World Bank and the European Commission. There are fact sheets per economy (by [UNDP](#) and the [World Bank](#)), [raw data and metadata](#), [qualitative study reports](#), and a [full analytical report](#), available online





The most comprehensive source of data on the housing of Roma in Albania is available in the UNDP's *Study of Roma and Egyptian Settlements*⁴ published in 2017. The data collection was conducted in all the municipalities inhabited by Roma population. Although the process of “mapping” Roma settlements did not include use of cadastral or ortho-photo materials in order to design an overarching GIS, this dataset provides data on many aspects of housing and living conditions of the Roma communities, both at the municipal and settlement level.

To the extent it was possible to analyse the studies as part of this regional effort,⁵ the data collection process included four different sources of information: 1) *data from previous studies*, 2) a *survey* focused on Roma and Egyptian families in need of housing, 3) *focus groups* with relevant stakeholders, 4) *on-site verification* of neighbourhood and infrastructure in the municipality.

Due to its focus on assessing and improving living conditions, the dataset consists of numerous useful indicators: 1) *material conditions* – measured as estimated cost needed for constructing new housing units or improving the existing ones; 2) *size of settlement* - exact number of families of Roma and/or Egyptian ethnicity in each settlement, as well as number of families living in substandard housing units; 3) *size of housing units* – measured as number of families per housing units of various size (from studios to +3 bedrooms); 4) *quality of construction* – measured as type of building (e.g. barracks); and 5) *access to infrastructure* – measured as percentage of households connected to the public utility network, as well as condition of road evaluated in a qualitative manner (e.g. “in a good condition”). In terms of geographic location, data clearly locates the settlements in the space, but it does not define (draw) its borders, nor it measures settlement characteristics outside of the Roma households (e.g. total number of housing units in the settlement).

Overall, the dataset is valuable contribution to the efforts to systematise data on Roma settlements. Due to the fact that they contain data at the level of settlement, they can be easily connected with the cadastral and GIS system, as well as census data. On the other hand, its narrow focus on material and living conditions of Roma households requires the list of indicators to be expanded in order include other dimensions of housing and better describe the nature of settlements at large.

In Albania, there is also the RomAlb database in Albania, which is not publically available, thus cannot be described and evaluated here from the perspective of its usability for the mapping purpose. From the available information it is known that it is an online monitoring tool that collects data on the implementation of the Albania's Action Plan for Roma integration.

Bosnia and Herzegovina

Despite improvements made regarding data collection in the last two years, the lack of information remains a serious issue in Bosnia and Herzegovina. In comparison to other economies, the complex nature of the political system in Bosnia and Herzegovina further hinders the efforts to develop systematic methodology for data collection and measurement of policy impact.

⁴ There are three reports published, covering different municipalities: 1) [Tirana, Dures, Kruje, Lezhe and Shkoder](#); 2) [Elbasan, Pogradec, Maliq, Korce and Devol](#); and 3) [Lushnje, Divjake, Fier, Berat, Kucova, Vlora, Gjirokaster, and Delvina](#).

⁵ Report studies were published only in Albanian language.





Currently, no data measuring specific housing conditions are available at the level of settlements. Based on the existing official data, available information rarely goes beyond total number of people/households⁶ and number of newly built objects at the national (or municipal) level⁷. The data from the 2017 Regional Roma Survey,⁸ as in the other economies, offer more specific information on: 1) **overcrowding** – measured as number of rooms per household members,⁹ 2) **access to infrastructure** – measured as share of households with access to piped water, sanitation, and electricity, 3) **sources of energy** (e.g. gas, woods, electricity), and 3) **ownership type** (e.g. private ,public) – measured as self-declared (rather than legal) ownership status.¹⁰

A recommendation by the European Commission in 2015 pointed at the need for *mapping the socially disadvantaged micro-regions or segregated settlements*. General policy activities¹¹ have been based on previously conducted needs assessment exercises, implemented through the local social centres. Such assessments have been done twice since Bosnia and Herzegovina joined the international Roma integration efforts. Up to date information on settlement level, which would measure the housing conditions specific to each settlement relevant for the proposed mapping methodology, has not been done. A mapping exercise to assess the needs of Roma is planned to be performed by the Government and funded through IPA. It would be very important to make sure data can be disaggregated to the level of settlements, rather than only to the level of municipalities.

Kosovo*

There is no available data specific to the housing situation of Roma in Kosovo. The [*“National Strategy for Inclusion of Roma and Ashkali Communities in the Kosovo Society 2017-2021”*](#) states that there are around 24,260 persons or approximately 6,000 families of these communities in Kosovo*, and recognizes the housing as a significant problem, identifying the need for *“registration of the housing needs of the Roma, Ashkali and Egyptian communities”*. It estimates that 80% of the Roma households need to regulate the legal status of their property.

To this point, systematic mapping of Roma settlements has not been conducted in Kosovo*, although it is pointed out that *“there is an idea to address this problem by putting into function the database on housing needs of the Ministry of Environment and Spatial Planning”*. Under the [*Law on Financing Special Housing Programmes*](#) (Art. 25), municipalities should develop three-year programs and projects for housing. However, according to the Strategy only fifteen municipalities drafted them, nine approved, and only four actually worked on specific projects¹².

⁶ The census does not provide information on the number of persons belonging to the Roma national minority. There are estimates (by municipalities and Roma organisations) that there are between 35,000 and 45,000 Roma.

⁷ Information from the National Platform on Roma Integration in 2018 is that there are 2.500 Roma families remaining in need for solving their housing issue.

⁸ See footnote 3.

⁹ The EUROSTAT definition of overcrowding is used: a person is considered as living in an overcrowded household if the household does not have at its disposal a minimum number of rooms equal to: one room for the household; one room per couple in the household; one room for each single person ages 18 or more; one room per pair of single people of the same gender ages 12–17; one room for each single person ages 12–17 and not included in the previous category; one room per pair of children under 12 years of age. Kitchens, corridors, bathrooms or rooms rented out or used by other households are not counted as rooms.

¹⁰ The legal and self-declared ownership most often differ (the household considers the dwelling as its own, but the dwelling is built without a permit, is not registered and very often built on someone else’s – private or public - land).

¹¹ [*Action Plan of Bosnia and Herzegovina for Addressing Roma Issues in the Fields of Employment, Housing and Health Care 2017-2020*](#).

¹² [*The Wall of Antigypsyism, Roma, Ashkali and Egyptians in Kosovo \(Civil Rights Defenders 2017\)*](#).





In comparison to other economies in the region, availability of survey data which could shed more light on the situation at the level of settlement is limiting. UNICEF's [Multiple Indicator Cluster Survey \(2013-2014\)](#), focusing on monitoring the situation of Roma children and women in Kosovo*, includes only few indicators that refer to housing conditions: 1) **quality of construction** – measured as percentage of households with finished floor, roofing, and walls; 2) **size of housing unit** – measured as number of persons per room used for sleeping; and 3) **access to infrastructure** – measured as percentage of household members using improved sources of drinking water.

Montenegro

There is no systematic data being collected/updated on regular basis that specifically target Roma settlements in Montenegro. Nevertheless, the situation in housing of the Roma community is being recognized and addressed.

Currently existing survey-based indicators used to evaluate the quality of living conditions are failing to provide a comprehensive overview of the current situation and serve as a basis for long-term policy strategies. To this point, no attempt at systematic mapping of Roma settlements in Montenegro has been made. Despite of the timely recommendation from OSCE/ODIHR (2003), creation of a database on the situation in housing, which would utilize cadastral and ortho-photo materials, remained a distant idea. While such data is readily available at designated public institutions, it is not being exploited due to lack of integration between various data sources within the same system.

Instead, the vast majority of available data on Roma housing in Montenegro is derived from three survey studies: *Research on Social Position and Social Inclusion of Roma* (2016),¹³ *Regional Roma Survey* (2017),¹⁴ and [MICS Roma Survey](#) (2018). The *Strategy for Social Inclusion of Roma and Egyptians 2016-2020* was formulated based on the mentioned research from 2016. Data was collected at the level of household and included the following relevant variables: 1) **material conditions**, 2) **legal status**, 3) **size of housing units**, 4) **type of construction**, 5) **access to utility networks**, etc. It may be attempted to aggregate household data at the level of settlement in order to use the data for the mapping purpose. However, some settlements are much underrepresented in the survey and 42% of the respondents live outside of Roma settlements, which limit the usability of these data for mapping of settlements. The study reports that 3/4 of Roma and Egyptians in Montenegro live in segregated settlements. Furthermore, a number of settlements referred to as “Roma settlements” in Montenegro, in fact, have majority of non-Roma¹⁵.

The Republic of North Macedonia

Although the need for a systematic evaluation of the results in the social inclusion of Roma has long been identified, modest progress has been made. The situation in the economy is further aggravated because of the lack of recent Census of population (the last was conducted in 2002). With this in mind, in 2018 the European Commission supported social mapping of Roma, which

¹³ A [survey](#) conducted by the Ministry of Human and Minority Rights for the purpose of preparing the [National Strategy](#). According to the 2011 census, there are 6,251 Roma and 2,054 Egyptians in Montenegro.

¹⁴ See footnote 3.

¹⁵ For instance, out of 62 social housing units for internally displaced in Nikšić that are discussed in Strategy, only 13 were assigned to Roma families. Yet, in the public, this settlement is referred to as a Roma settlement.





resulted in the report “[*Thematic evaluation on support for Roma communities and Social Mapping of Roma*](#)”.

To this day, this document represents the most comprehensive data on Roma housing in the economy. It provides an overview of the status of living conditions in the Roma settlements in 14 municipalities¹⁶ with the most numerous Roma population, effectively covering 90% of the Roma. The method of data collection included: 1) qualitative participatory rural assessment; and 2) quantitative surveying based on representative sample.¹⁷ The unit of analysis in the area of housing are households.¹⁸ The study claims it is representative at national and municipal level.

With regards to housing, the following variables have been used: 1) **minimum living space** – measured as the average number of persons in the household, the size of the dwelling in m² (categorical), shared dwelling areas (yard); 2) **equipment** and **access to communal infrastructure** (*electricity, water, asphalted road, and drainage*) – measured as percent of households lacking access to main infrastructure, as well as access to in-house kitchen and toilet with bathroom; 3) **household’s integration into settlement/town**; and 4) **legal certainty** – measured as informal vs. formal residence ownership and percent of legalised housing units.

Although these indicators offer important knowledge about the conditions of Roma households in the substandard settlements, they are insufficient to provide wider contextual description of the settlement which should include coverage by urban plans, state infrastructure, demographic composition, etc.

Serbia

On the basis of 2011 Census, poverty indicators of Roma population are collected and analysed in the study “*Roma in Serbia*”.¹⁹ According to this study Roma generally live in poorer conditions and substandard housing than of the general population described as “*ethno-class status, that is, sub-proletarian status which characterizes discriminated minority groups*”. In average, Roma households live in housing units that are of lower structural quality, smaller, often lacking connections to basic infrastructure (16.7% lacked connection to sewerage and water, 1/3 did not have toilet and bathroom, 3% did not have electricity). Furthermore, the study found that 1/3 of all the homeless people in Serbia are Roma (5,719 homeless Roma). Additionally 1,553 Roma households live in buildings “occupied out of necessity” (not originally built as housing units, such as abandoned storage or office spaces, industry buildings, etc.), considered as secondary homeless. The study stresses that many Roma live in settlements of favela and slum types, especially in Belgrade. However, it does not provide any specific data on the settlements, not even the locations of the settlements, necessary for proper policy planning.

Within the project “We Are Here Together – European Support for Roma Inclusion” funded by IPA 2012, during 2015 and 2016 Serbia mapped all substandard Roma settlements on its territory for the purpose of national-level planning and programming of activities for improvement of housing conditions of Roma. 583 substandard Roma settlements were mapped with exact location and basic information on the physical characteristics and housing conditions

¹⁶ Municipalities included in the study: Tetovo, Prilep, Štip, Vinica, Strumica, Gostivar, Kumanovo, Kočani, Debar, Kavadarci, Šuto Orizari, Bitola, Kičevo, Veleç.

¹⁷ For the analysis at the level of settlement, a combination of quantitative and qualitative methods is used.

¹⁸ Approximately 2200 individuals in 550 Roma households across 14 municipalities were interviewed.

¹⁹ Statistical Office of Republic of Serbia (2014). [*Roma in Serbia*](#). Beograd: RZS.





(quality of housing units, existence of urban plans, state of infrastructure, etc.).²⁰ The data were systematised in a Geographic Information System (GIS) of Substandard Roma Settlements in Serbia, owned by the government and not publically available. This is the first case in the region to systematically map all the areas where Roma live in substandard housing conditions. The GIS database is planned for updating in 2020 through IPA 2014 funding.

Summary

In conclusion, the existing data on Roma housing in the region does not seem to meet the necessary criteria to inform the policy for improving the substandard living conditions, with the exception of the GIS in Serbia, that does not seem to be utilised for this purpose. Certain data relevant for housing already exists, but are not properly systematised. Several reasons for such lack of data are determined.

First, the existing data are scattered across different institutions responsible for different aspects of the housing (legalisation, infrastructure, living conditions, etc.). Depending on the level of decentralization and the complexity of the political system, holistic approach to Roma housing was mostly hindered. Second, data is not being systematically collected or updated. For instance, in certain economies mapping has been conducted for some municipality but not for all. Furthermore datasets on Roma housing need to be updated regularly and development should be tracked over time. Third, even where housing data has been collected at least partially, the choice of indicators, definition of settlement, level of analysis (household, settlement, city, municipality, state), has varied greatly across respective economies.

²⁰ Đorđević, A. (2017). *Substandard Roma settlements in Serbia: An overview of data from the Geographic Information System for 2016*. Belgrade: OSCE. A collection of related publication is available [online](#).





3. Key issues for the mapping methodology

Before discussing the key issues relevant for developing the mapping methodology, it may be useful to note that the methodology proposed with this document is in line with the recommendations from the UN and EU. It may also be useful to note that the example of Serbia's GIS on substandard Roma settlements was used as a model for developing the proposed regional methodology.

There are two key issues to be defined prior to the development of the mapping methodology. These two issues are: **(i) the purpose of mapping** and **(ii) the object of mapping**, or in other words: what and why should be mapped?

By answering these two questions, the inputs that are determining the mapping methodology are provided. Defining the purpose of mapping directly implies the type, volume and sources of data that should be used in the mapping. The definition of the areas to be mapped implies the very process of data collection, processing and matching in order to get the final product – a holistic picture on each area where the Roma population live in “substandard” housing conditions.

Purpose of the mapping

The first purpose of mapping is **to provide a complete overview on the housing conditions of Roma in each economy of the Western Balkans**, in a systematized manner, using standardized methodology, and based on reliable and most recent data, thus filling the currently existing gap in information regarding the housing conditions of Roma.

The second purpose of mapping the housing of Roma is **to provide information necessary for physical improvements of the living conditions in Roma settlements**. This includes enabling pre-conditions for legalisation of individual houses, development of urban plans, construction/improvement of infrastructure, providing access to communal and social services, improvement of individual housing units, solving property/tenure issues, etc. The mapping will primarily serve central and local governments, but also civil society, international and local organizations and other stakeholders to develop well-informed medium and long-term plans and programs for upgrading settlements and improvement of living and housing conditions.

The third purpose of the methodology is to **provide for monitoring of the progress and demonstrate the achievements of the plans and programmes for housing of Roma**. As the first step, the methodology should be used to set the baseline (starting point) for these processes, and afterwards periodically monitor the progress in a comparable way in order to be able to identify trends and evaluate the success of the programmes and strategies for Roma integration in the area of housing.

To this end, the mapping will also serve for informed implementation and monitoring of the progress in **achieving the objective in housing set with the Declaration of Western Balkans Partners on Roma Integration within the EU Enlargement Process**. This objective reads: “wherever possible, legalise all informal settlements where Roma live; or provide permanent, decent, affordable and desegregated housing for Roma currently living in informal settlements that cannot be legalised for justified reasons.” The governments from the region have undertaken the obligation to achieve this objective by the time the respective economy acquires EU membership.





The fourth purpose of the mapping is **to provide comparability of the situation of Roma housing conditions among the economies of the region**. For that reason, data should be compatible and comparable, which means the data should be collected in the same manner, and with the same definition of individual indicator (e.g. definition of “living space” or “adequate sanitation facilities” in the Census in each economy should be the same/compatible). Regional comparability of data will not only help the EU accession process, but may also serve as a basis for regional initiatives in the field of housing of Roma.

Object of mapping: substandard Roma settlements

The data collected by mapping using this proposed methodology should serve primarily as the main input for any long-term strategies, policies and programmes for housing of Roma, many of which are already ongoing or planned. It may be the case that settlements not mapped may be “invisible” for those programmes, and not be included into any improvement projects, which poses a problem. On the other hand, if the settlements that are the object of mapping are not properly defined, it may be the case that certain settlements, regardless of the proportion of Roma inhabitants, that do not need intervention, may end up as mapped and be subject to funding that may be more useful in other locations.

It is therefore necessary to establish a clear definition of the object of mapping, in order to capture exactly those areas where the assistance from international and national organizations and funds is necessary, and which also represent a “spatial issue” – requiring some sort of spatial intervention and improvements before individual houses could be legalised and improved.

Establishing such definition requires answering three main questions:

- What is considered to be a “Roma settlement”?
- What is considered to be a “settlement”?
- How to differentiate general informal/illegal constructions and settlements, widely present in all WB economies, from the Roma settlements that may also be illegal?

a) “Roma” settlement

The term “Roma settlement” is the most widely used in research, programmes and projects for improvement of the housing conditions of Roma, not only in Western Balkans economies, but in Eastern Europe in general. It usually implies a part of the city or village, district or neighbourhood where **Roma population represents majority or a significant percentage** of the total population. However, when planning to map these settlements, two issues appear.

Firstly, there are parts of cities and villages where significant number of Roma live, and are thus considered as “Roma settlements” or “Roma neighbourhoods”, but are **not necessarily below the general standards of adequate housing**. It means that these Roma settlements are equipped with the entire basic infrastructure, have access to all the public services, the dwellings are built from solid materials, providing safety and health for inhabitants, etc. Such areas may be mapped using this methodology, but obviously should not be prioritised for housing improvement programmes for which the mapping is intended.





Picture 1. Examples of Roma neighbourhoods developed in accordance to the adequate standards of housing.

Secondly, there are areas where **Roma have not declared their ethnicity** in the census, thus, although the settlement is often seen from the outside as a “Roma” settlement, and is characterised by living conditions below the generally accepted standards, the data from the Census are not pointing to it as a “Roma” settlement.

On the other hand, there are cases of Roma households living integrated within the mainstream settlements, in housing units in conditions comparable to the rest of the population in the settlement. Often the number of Roma households in such settlements is very small. Since the whole settlement and the individual Roma housing units cannot be characterised as substandard, such settlements should not be the object of the mapping proposed with this methodology.

This mapping methodology proposes the use of the data from the existing official sources, meaning that in order to determine if some area can be considered as a “Roma” settlement, Census data should be consulted. However, if the percentage of Roma population as in the official Census data is put as the main criteria for mapping of an area, there is a risk that some areas that require intervention and assistance may be left out. Therefore, the identification of the settlements that should be the object for mapping should combine other sources of information besides the official Census data. It is crucially important to insist on mapping virtually all the settlements where Roma live.

b) Definition of “settlement”

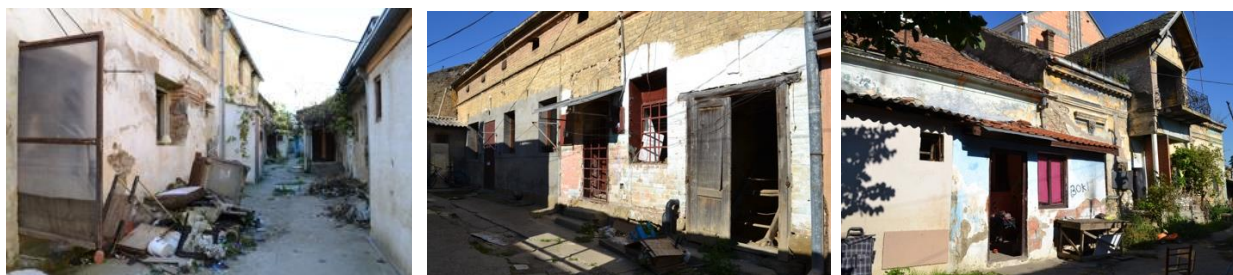
It is also necessary to define, in physical terms, what shall be considered as a “settlement” or “neighbourhood” that shall be mapped with this methodology.

Some quantitative criteria may be applied to define the minimum size of what is considered a “settlement”, in terms of minimal number of objects, inhabitants, land surface area or similar, in order to avoid mapping as “settlement” a single dwelling or uninhabited areas. Some researches in the Western Balkans focused on mapping of Roma settlements using benchmarks such as, for example, clusters of at least 15 houses, or 100 inhabitants, etc.²¹ These usually represent a very good insight into the current state of living and housing conditions of Roma living below the poverty line, and are an excellent basis for planning interventions in existing settlements.

²¹ Jakšić, B. and Bašić, G. (2002), *The art of surviving – where and how Roma in Serbia live*. Belgrade: Institute of Philosophy and Social Theory.



However, the purpose of the proposed mapping methodology is not only to provide reliable insight of the current situation, but also information and inputs for future interventions. Experiences from Roma housing projects in the region show that there are many cases of poor Roma families living in abandoned warehouses, industrial, office or public buildings, very often several dozens of families under the same roof, and it is clear that such situations represent a “spatial” problem in the sense that they demand the provision of adequate housing, which means provision of land, development of urban plans, construction of infrastructure, etc. If any criteria for the minimal size of “settlement” are applied, there is a risk that such cases will be left out of mapping since in physical terms, there is only one object, or “house”.



Picture 2. Examples of Roma families living in buildings not constructed for housing: 18 families living in abandoned warehouse (left), 12 families also living in abandoned warehouse (middle) and 14 families living in former City House (right), all in Obrenovac municipality, Serbia. Source: OSCE (2014).

Furthermore, it is also hard to determine universal spatial criteria that would be applicable in all the economies of the region, but also in all the municipalities within the same economy. There is a big difference in what is considered to be a group or cluster of houses that requires some spatial intervention in a village with few hundred inhabitants, and in big cities of the Western Balkans region. In some economies, Roma prevalingly live in urban areas (such is the case of the Republic of North Macedonia), while in others they live in urban and rural areas in approximately equal percentages (Serbia, Montenegro, etc.).

Therefore, in physical terms the mapping of settlements should cover all the spatial areas with one or group of objects where any number of Roma households live.

c) Differentiation from the mainstream illegal construction

It is very often the case that the term “informal Roma settlements” is used in research, programmes, projects, and even strategic and legal documents in the Western Balkans region. However, this definition is also not precise enough to be used for mapping at the national and regional level.

In all the economies of the region, informal construction represents a serious issue, with huge numbers of buildings constructed without building permits and not respecting urban regulations, reaching significant percentage of the total building stock.²² Researches on informal construction in the region show that although it is a very complex phenomenon which has a social, economic, legal, cultural and other aspects, it cannot be related directly and only to poverty of informal builders - houses in these settlements are usually built of solid materials, have large surface

²² NALAS (2011). [Challenges for regulation of informal settlements in Southeast Europe](#). Skopje: NALAS.





areas, and on large land plots.²³ Also, very often not only housing but commercial or business buildings can be found in these settlements, and some of the housing units are luxurious.²⁴ Furthermore, many of the illegal buildings for housing are not necessarily the primary and/or only housing option for the owners, while among the Roma their illegal dwellings are usually their only housing option.

These mainstream illegal buildings and settlements are mainly the target of various laws that define the process of legalisation, and as such are not targeted with the housing programmes.

Areas inhabited by Roma families are very often informal (illegal), and although some of them fit into the description of the mainstream informal construction as described above, the majority of them are out of the official regulations due to the poverty and general exclusion. A great number of Roma families is living below the poverty line, and cannot afford any legal housing option or any legal way to obtain land plot. They often occupy land that “nobody else wants” - land exposed to natural disasters and hazards, land zoned for industry, even land next to the city landfill sites, etc. The illegal status of these settlements is therefore a direct consequence of poverty and social exclusion of their inhabitants. In great majority of cases, housing units in these settlements cannot be legalised because they do not fulfil the criteria provided by the laws on legalisation or construction (adequate location, adequate construction standards, etc.).

Although very few, there are also Roma settlements that are completely legal, some of them existing as so-called Roma “mahala” for centuries,²⁵ but still having housing conditions below the standards (e.g. dilapidated houses due to poverty and consequential lack of proper maintenance, thus not safe for habitation, lack of infrastructure, etc.). There are also already mentioned cases where families live in buildings that are legal, but not constructed for housing purpose, but are abandoned objects for other purposes (warehouses, industrial buildings, etc.) and usually dilapidated.

Therefore, it is important to make a clear distinction between the mainstream informal construction and settlements that are to be mapped on one hand, and on the other hand to capture in the mapping the settlements that are legal but still with housing conditions below the standards.

The definition proposed for mapping

For the purpose of this mapping methodology, the areas proposed as the object of the mapping are called “**Substandard Roma Settlements**”.

The term “**substandard**” implies areas in which living and **housing conditions are not at the level of generally accepted standards of adequate housing**.

The term “**Roma**” implies presence of Roma in a given area in **any percentage**, primarily in accordance to the data in the official Census, and secondary as pointed out by key representatives

²³ For example, Grubović, Lj. (2006). *Belgrade in Transition: an analysis of illegal building in a post-socialist city*. PhD. London School of Economics.

²⁴ Žerjav, B. (2009). *Incorporating Informal Construction: Urban Planning in Belgrade and Proposals for Changes*. MSc. IHS Institute, Erasmus University Rotterdam.

²⁵ In most of the economies in the region, objects built before the first regulations on legal building are automatically recognized as legal. The issue in such cases is often proving the age of the building.





of the Roma community. A criterion in terms of minimum percentage or number of Roma people/families living in certain area should not be applied in order to cover as many settlements in need for housing development as possible. The mapping should be done primarily based on the physical parameters, and percentage or number of Roma in an area should serve only as an indicator, rather than criterion for mapping.

The term “**settlement**” relates to **any cluster of housing units that fulfils the criteria of “substandard”** and that requires any form of spatial intervention before improvement of individual housing units and their subsequent legalisation is possible (resettlement, development of urban plan, construction of infrastructure, etc.). It is not limited by quantitative criteria, in order to allow even for a single dwelling to be mapped if necessary to be considered for a future development policy. Therefore, this methodology proposes that a “settlement” is understood as **one or a group of housing dwellings that make a distinguishing whole (from the neighbouring settlements)**.

The proposed definition of settlements to be mapped is in line with the **operational UN definition of slums**. Since a term “slum” is not widely used in the WB region, the term “substandard settlement” is used instead.

According to this definition, **a substandard settlement represents any cluster of housing units that is characterised by at least one of the following:**

- Lack of access to improved *water source*;
- Lack of access to improved *sanitation facilities*;
- Lack of access to *basic services* (sewerage, electricity, access to road network, access to other communal services, public transport, health and education facilities, etc.);
- Lack of sufficient *living area* (overcrowding in terms of housing space and/or settlement space, against the number of inhabitants);
- Lack of *structural durability* and *safety of housing* (houses built from inadequate materials, or by inadequate building techniques, or dilapidated, not safe for human habitation, not providing shelter from environmental conditions, or spaces not suitable for human habitation such as warehouses, abandoned commercial buildings, etc.);
- Lack of *security of tenure* (uncertainty of legal status of houses and land use, unresolved property and land use/ownership issues).

Therefore, the definition of “**substandard Roma settlements**” combines the three aspects: the definition of a “settlement” as one or a group of dwellings, the definition of “Roma settlement” as a settlement with any proportion of Roma inhabitants (according to the Census and/or as identified by the local community/Roma representatives), and the definition of “substandard” as a physical condition of the settlement and dwellings thereof characterised by one or more of the listed criteria. The mapping, if so agreed at regional level, may include settlements without any portion of Roma inhabitants, if the settlement is severely underdeveloped (characterised by a combination of at least two/three of the above listed criteria).

Applicability of the definition in Western Balkans economies





Although there are cases of other ethnic groups that live in housing conditions fitting the proposed definition, practical and research experiences in the region show that the great majority of such settlements are inhabited with Roma. For instance, based on the data from the 2017 Regional Roma Survey, 23 per cent of the Roma households are not connected to piped water, 30 per cent do not have sanitary facilities and 30 per cent are not connected to sewerage.²⁶ Obviously, this number becomes significantly higher when other indicators are added, such as the size of living area and security of tenure. Therefore, just by mapping the settlements (or clusters of houses) based on the criteria of underdevelopment, the largest proportion of Roma dwellings with inadequate housing conditions would be included in the mapping.

The experience from the mapping in Serbia that resulted in the GIS of Substandard Roma Settlements showed that, when solely physical parameters given above are used, about 90% of the population in such settlements are Roma. At the consultative meeting for the development of this methodology in the Republic of North Macedonia it was estimated that at least 80% or more of the whole Roma population would be included in the mapping of such settlements. In Kosovo*, almost exactly the same definition is already in place for “informal settlements” within the Spatial Planning Technical Norms by-law of the Law on Spatial Planning. The same criteria as those proposed with the definition of substandard Roma settlements are applied for informal settlements, which is distinguished from the general “illegal construction”. The relevant stakeholders in Kosovo* confirm that Roma mostly live in such settlements. In the other economies of the region consultative meetings regarding the proposed methodology were not held, but the communication with some of the relevant stakeholders indicated that the situation is the same, implying that the proposed definition for the object of mapping, namely “substandard Roma settlements” is applicable throughout the region.

²⁶ Regional Cooperation Council (2017). *Targeting Roma in Housing Policies of the Western Balkans: A Legal Overview*. Regional Cooperation Council Roma Integration 2020 Action Team. Belgrade. Serbia.





4. Data on the substandard Roma settlements

Volume of mapping and data to be collected

From the purposes of the mapping described in the previous chapter, the key criteria for the technical aspects of the mapping could be drawn:

- The mapping should focus on those **areas requiring assistance** and state intervention with regards to improvement of living and housing conditions (excluding mainstream settlements with accidental cases of integrated Roma households).
- The mapping should be holistic for each economy, covering **whole territory** and **determining locations** of all the areas where Roma population lives in substandard housing conditions.
- The mapping should provide decision-makers at the national level with enough data for evidence-based planning and programming of improvement of living conditions in those areas, which means collecting all the relevant data about **physical characteristics on Roma housing (both for the housing units and the settlement)**.
- The mapping should provide monitoring of the progress of upgrading/improvement projects and programmes, which implies that the mapping process should be **replicable in regular time periods (e.g. 5 to 10 years)**, to provide update on the situation.
- The mapping should provide comparability of the state of Roma housing across the economies of the Western Balkans, which means that it should be done with the same methodology, using same definition of areas that should be mapped, and data collected in the same or comparable manner, from the similar sources.

This implies that the mapping should focus on collecting alphanumeric and spatial data on **physical characteristics of areas where Roma population lives** - data related to housing conditions. It is important therefore to collect both qualitative and quantitative data on the housing conditions in the settlements (quality of housing units, infrastructure, services, urban plans, land ownership, etc.). While it has been recognized that efforts to tackle housing issues should not be disconnected from the wider socio-economic integration (e.g. in employment, education, etc.)²⁷, such data are of secondary importance for the national-level decision-making with regards to improvement of housing conditions of Roma.

Furthermore, for the purpose of replicability of the mapping, as well as maintenance and regular update of the database, it is important to collect only the necessary data, and not to overload the system with data that is not of key importance, since it can create additional burden to the central and local institutions that will be responsible for maintaining the database, and can create difficulties in its maintenance and periodical update. Such a database would be at risk to soon become outdated and thus useless for Roma housing improvement programmes.

Data on the socio-economic situation of Roma population in healthcare, employment and education, are sometimes collected in some of the economies of the region by the responsible institutions. Any such administrative register or statistical data may be combined with the data on housing, in a separate database. However, it is not the direct aim of this mapping.

²⁷ Ibid.





Data quality and sources of data

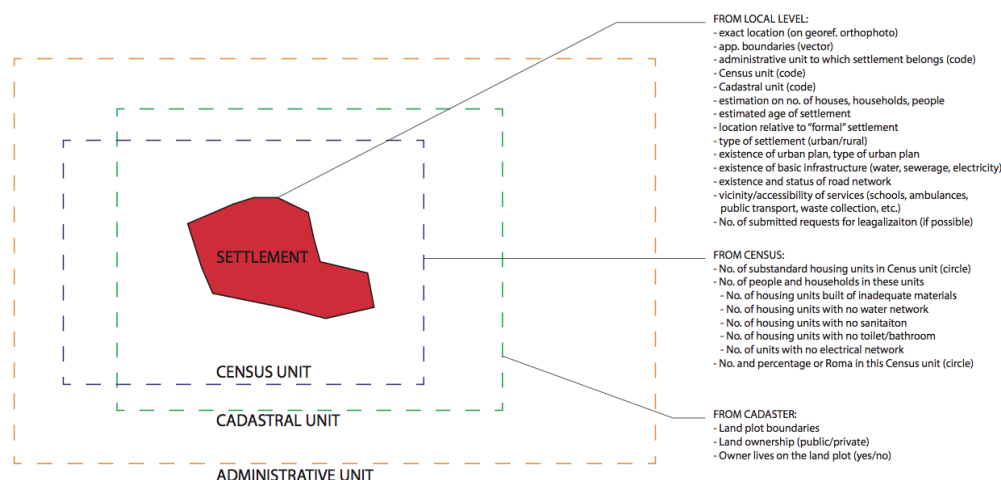
The list of criteria provided above implies that **only data from official sources, central and local, should be used, in order to provide reliability of the mapping, compatibility and comparability across economies**. This means data should be collected in a comparable manner, with the same definitions of individual indicators, etc. Additional sources of information should be consulted as much as possible. The mapping should be done at the level of “settlement”, for the purpose of statistical analysis, so no personal or data on individual households are necessary and should not be collected.

Data from 3 types of sources should be used:

- Central-level government;
- Local-level government;
- Independent state institutions such as agencies and public enterprises from central and local levels.

There are three groups of data that should be collected for each settlement:

- **Data related to the location of the settlement:** exact location of a settlement, approximate boundaries, administrative, cadastral and census units in which the settlement is located, rural/urban area, position relative to the “mainstream” settlements (integrated, at the edge, disconnected), etc.
- **Data on the physical characteristics of settlement:** age of the settlement, number of housing units, number of people, quality of the houses, quality and existence of basic infrastructure, access to services, etc.
- **Data on the legal status of the settlement:** existence of urban/spatial plans, land zoning, land ownership, is the location suitable/safe for human habitation, number of households in the different stages of the legalisation process, etc.



Picture 3. Diagram showing combining the data from different levels and sources, and for the different spatial/administrative units (census circle, cadastral unit, administrative unit).

The first step in the mapping is to actually agree, according to the proposed definition of substandard Roma settlements above, on a list of settlements that should be included in the





mapping exercise. As highlighted above, the list should contain virtually all the settlements that fall under the definition, throughout the territory of the economy, for each municipality. This should be done in cooperation among the leading central institutions responsible for Roma issues, the relevant local institutions for each municipality, the representatives of the Roma community recognised in the respective economy, including civil society, political, and any other relevant representatives, both at central and particularly at local level.

To determine the exact locations of the settlements, geo-referenced satellite images (ortho-photos) from the national Cadastres should be used. These are produced regularly in all economies, cover the whole territory of each economy, and are done in the same manner. The fact that they are geo-referenced will enable precise location of each individual settlement, and compatibility with the official maps, municipal, cadastral and census unit boundaries, for later intersection of data. The satellite images are not done in the same year in each economy, but this will not affect the reliability and the possibility for comparison, as the situation on the ground is generally stable over long periods of time.

Data on substandard housing units, as well as data on number of households and persons, and number of Roma people in each settlement, can be obtained from the official Census. The Census is done using the same methodology in all the economies. Data in the Census cover all municipalities, all individual houses and housing units, and all the specific indicators are the same. For the purpose of this methodology, it is important that the census unit is the smallest spatial unit in which the census survey is done, and that in each Census, data on housing units that do not fulfil criteria of adequate housing are collected (that is: housing units build of inadequate materials, dilapidated housing units, units inhabited from necessity, units with no connections to water, sewerage, electricity, no toilets or bathrooms, etc.). The last Census was done in 2011 in all the economies except in the Republic of North Macedonia (where the latest Census was done in 2002). However, at the moment, the next Census cycle is under preparation and will be done in 2021 in all the economies of the Western Balkans. The new Census will contain all the data necessary for this mapping, and thus it is recommended that the mapping is done after the Census takes place, in order to use the most up to date information.

For information on the general quality of housing, existence of urban plans, infrastructure, access to services and other data describing the settlement as a whole, data should be collected from the local level administrations - cities, municipalities, or regions where applicable. In each economy, local level governments have departments responsible for urban planning which can provide data on existence and type of urban plan and land zoning for the whole territory of municipality. Similarly, departments responsible for infrastructure, as well as local public utility companies, possess data on existence and status of infrastructure for the whole territory of a municipality. Local cadastral offices are able to locate each cluster of housing units that fulfils the criteria from the proposed definition, and information about the type of ownership of land on which the settlement is located (public or private). In some cases, local departments responsible for legalisation will be able to provide data on the number of applications for legalisation submitted from a mapped settlement, but it is rarely the case since such data are usually cumulative, for the whole municipality, and are not disaggregated by census or cadastral units, or addresses. In this sense, preparatory work should be done in order to be able to obtain relevant data on the legalisation status of the housing units in the mapped settlements.

Data from unofficial sources, such as surveys, data collected based on sampling, estimations from Roma national minority organizations, CSOs etc. should be used only as a reference for





cross-checking the mapping results. These data can be useful as an estimation, to have an approximate idea on the number of settlements and their characteristics that could be expected as a final result of mapping. Such data are collected using different methodologies. Therefore it is difficult to use such data for the mapping. Nevertheless, if the responsible institutions can accept such estimations as relevant, when the needed official data are missing, such estimations may be included in the mapping exercise with a special note on their reliability and a recommendation to provide an official data source for the next mapping exercise.

Process of data collection

As the first step of the mapping, a geo-referenced ortho-photo should be provided by the national Cadastres to each local self-government, covering the whole territory of the municipality or city.

As a second step, each local self-government should provide the exact locations for each settlement fitting the described criteria in the definition above on its territory, and mark its approximate boundaries in a vector format. Local urban planning departments, or local cadastral offices, should be the most suitable instances to locate the settlements on the territory of their municipalities, since they have a good insight in the situation on the field.

When settlements are located, the local self-governments (different departments, agencies and local public utility companies) should provide all the necessary data from their domain for the territory of each settlement. This information should be gathered through a questionnaire that is part of this methodology, and should be sent to each local self-government. There should be an institution and a person coordinating this process at each local self-government.

The questionnaire contains closed-ended questions, and provides a collection of standardized alphanumeric data ready to be used for the development of a GIS or digital databases. A sample of the questionnaire is provided in Annex I of this document.

It will probably be necessary to develop guidelines for the municipalities and deliver a series of trainings for the local administrations officials on how to properly fill-in the questionnaire, prior to sending the questionnaire to all the municipalities.

Four groups of information should be collected for each substandard Roma settlement through the questionnaires:

1. Information on the exact location of the substandard Roma settlement:

When the location of a settlement is determined and geo-referenced in a vector format on the ortho-photo, it will be possible to determine to which administrative units it belongs:

- Municipality to which a substandard Roma settlement belongs;
- Settlement (village, town, city district, etc.);
- Census circle(s);
- Cadastral unit(s);

Codes/numbers of each unit can be provided from the official Registers of spatial units that exist at the central level in each economy.

2. General data on the substandard Roma settlement:





The questionnaire contains questions on an estimated number of housing units, households and people in each settlement. Assessment on the prevailing quality of housing units in the settlement should also be provided (quality of construction, structural quality of housing units, materials used, etc.). It also contains questions about the estimated age of a settlement (to determine the time when the settlement was established) and its position in relation to the mainstream settlements (integrated, at the edge or disconnected). These estimations can be provided by local departments responsible for urban planning.

3. Information on the status of infrastructure in the substandard Roma settlement:

Information on existence, coverage and quality of sewerage, water, electricity, road network, street lightning, heating system, access to other public communal services like waste collection, health and education facilities, etc. can be provided from the respective local administration departments and utility companies for the territory of each settlement.

4. Information on urban/spatial planning documents covering the substandard Roma settlement:

This part of the questionnaire contains questions on the existing adopted spatial/urban plans that cover the settlement, land zoning, type of land ownership, and estimated percentage of legalised objects in the settlement.

The questions about the spatial and urban plans should be filled-in by the local urban planning departments. All the plans that cover a settlement should be listed, from the top-level plans like spatial plan of municipality/city, to the lower-level plans such as plans of detailed regulation. Important information from this section is whether a plan that covers the settlement (if any plan exists) provides preconditions for direct issuing of construction permits, or it is necessary to produce some additional, lower-level plan(s) in order that construction in the settlement is legally possible. Information on the land zoning, that is, the purpose of the land on which the settlement is located (zoned for housing or another purpose, such as industry, commerce, transport, etc., or not zoned for construction at all).

Information on the form of land ownership should be provided by the local cadastral offices. The key information from this part of the questionnaire is whether the land is in public or private ownership, since it usually critically determines the process of improvement/regularization of the settlement. The estimated number or percentage of legalised objects in the settlement, or number of submitted requests for legalisation, should be obtained from the local administration's department responsible for legalisation. As explained above, in most of the cases data in these departments are cumulative, and it will be hard to get a precise number of legalisation requests from a settlement, but a relatively precise estimation should be possible across the region. The information should include the various levels of legalisation (at the minimum: not legalised and not in the process of legalisation, not legalised and in the process of legalisation, and legalised).

As a next step, answers obtained through the questionnaire from the municipalities will need to be checked and quality of collected data assessed and verified by a group of experts (e.g. GIS experts, Roma housing experts, etc.). It is necessary to determine the completeness of data (percentage of completeness of each questionnaire), consistency of data (are data collected in accordance to the guidelines provided), accuracy (precision of answers, accuracy of spatial data about the boundaries of a settlement, etc.), and credibility of the data (this can be assessed based





on the experience of experts, e.g. if number of registered settlements in a municipality is logical and expected or not).

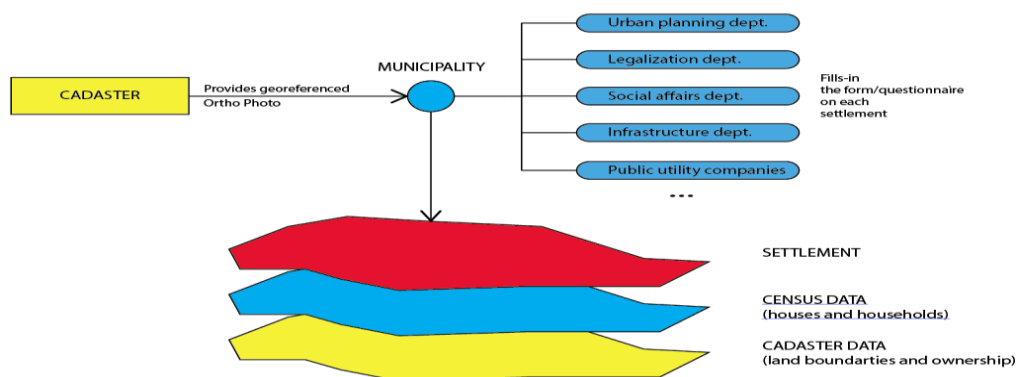
When the data from the questionnaires are verified, they can be matched with the data from the Census and the National Cadastre. Based on the codes of the census circles obtained through the mapping at the local level, data on “substandard” housing units that exist in each census circle in which a substandard Roma settlement was registered can be obtained from the central statistical authorities.

The central statistical office should provide Census data for the statistical units (Census circles) in which the mapped settlements are located (numbers/names collected through mapping), on:

- total number of housing units, households and people in each census unit,
- number and percentage of Roma people, households and housing units,
- number of housing units constructed with inappropriate materials and/or inadequate construction techniques,
- number of non-housing units in which people live (units populated from necessity),
- number of units with no connection to water network,
- number of units with no connection to sewerage network,
- number of units with no toilets/bathrooms,
- number of units with no connection to electricity grid,
- number of overcrowded housing units.

This way, a precise number of housing units that are not in accordance with the official standards of adequate housing will be obtained, as well as a number of households and people living in them. The number and percentage of Roma population within the same census circle will also be obtained from the statistical office. Obviously, this information will include only those persons that voluntarily self-declared as Roma.

Using the same method, land plot boundaries can be obtained from the **National Cadastres**, by using the codes of cadastral units in which the substandard Roma settlements are located. This will provide the exact boundaries of the land plots on which the settlement is located, on top of the information about the form of land ownership, if such information exists.





Picture 4. Diagram of types of data provided by different sources.





5. Data Format and Integration Procedure

This proposition of the regional methodology for mapping of Roma settlements relies on integration of already existing sources of data. These data, currently scattered across multiple public institutions, hold information crucial for development of comprehensive and systematic dataset of Roma housing in the Western Balkans region. In order to fulfil its purpose, the respective institutions should be required to deliver distinct pieces of data in a format that would guarantee smooth data integration process. Below, the technical nature of datasets is briefly explain and suggestions of software solution for both integration and analysis are offered.

Format and dimensions of spatial objects

Without a doubt, the crucial piece of the data comes from the cadastral services (or their geographic information system (GIS) departments). Once borders of Roma settlements are identified by local authorities and then mapped, geospatial data serve as a “vehicle” in which other information about the settlement are being stored.

Geospatial data should be provided in the shapefile (.shp) format, which is a vector data format for geographic information system (GIS). The shapefile format is used to spatially describe vector features, such as: points, lines and polygons. Each unit in the shapefile has additional dimension used to describe the object – attributes. For example, marking the territory of Roma settlement as a polygon would provide geospatial information about the settlement. Its ability to attach attributes allows for other sources of data to broaden the description of spatial object with non-spatial characteristics. Under this methodology, majority of settlements’ attributes are provided by the local authorities.

The shapefile is composed out of three separate components: Main file, Index file, and dBASE table. In order to properly function, each of these three components have to have the same prefix²⁸ (the name of the file) and different suffix (extension). In addition, all three mandatory pieces of shapefile need to be stored in the same directory (folder) (for organization of directory and naming see **Figure 6**):

- **Main file (.shp)** – shape format: file containing geometry data in which each record describes a shape with a list of its vertices.
- **Index file (.shx)** – shape index format: where each record contains the offset of the corresponding main file record from the beginning of the main file. Index file allows for seeking backwards and forwards.
- **dBASE table (.dbf)** – file storing features attributes with one record per feature (organized in columns).

Figure 5 shows an example of geospatial mapping of the Roma settlement Konik Camp, in Podgorica (with and without ortho-photo background), as well as the wider census circle.²⁹ The shape of the Roma settlement in the *Main file* will be described as the list of vertices marking the edges of the settlement (represented in the dots K1-K6). Each of these vertices holds record of geolocation represented by two coordination system integers (longitude, latitude). List of

²⁸ Based on ESRI (Environmental Systems Research Institute) naming convention prefix (name) must start with an alphanumerical character (a-Z, 0-9) followed by zero or seven characters (a-Z, 0-9,_,-).

²⁹ Census circle borders are drawn arbitrary for the purpose of this example.



attributes attached to this spatial object are stored as separate variables collected from the other institutions, such as:

- **Local authorities:** *Settlement's unique code, type of settlement, existence of urban plan, access to utility infrastructure, vicinity/accessibility of services (schools, ambulances, transport), number of submitted requests for legalisation etc.*
- **Census:** *Number of people living in the Census unit (circle), number of substandard housing units in the circle, number and percentage of Roma in the circle, etc.*
- **Cadaastre:** *Land plot ownership (public/private), etc.*



Red lines
represent
borders of
census circle
**Light-green
lines**
represent
borders of
substandard
Roma
settlement.

Figure 5A. Konik Camp, with ortho-photo background

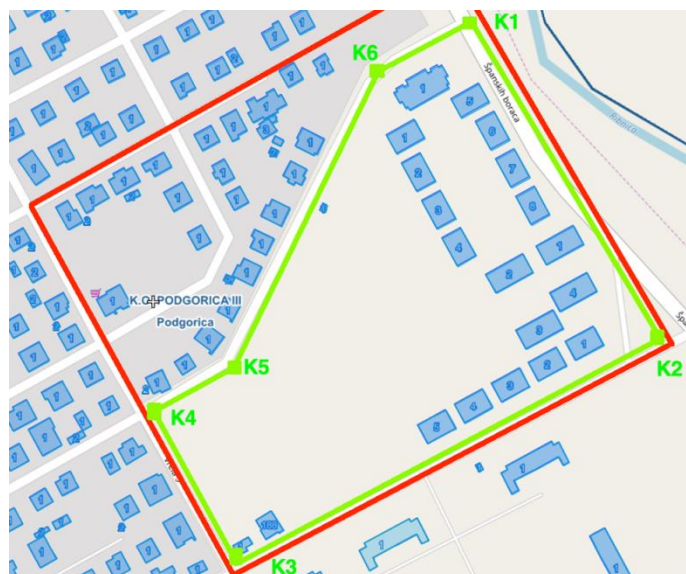


Figure 5B. Konik Camp, without orthophoto background



Dataset integration

It is often the case in practice that data about the same object (substandard Roma settlement) from different sources need to be combined. As previously explained, one data source includes shapefile with geometries and an identifying index variable, while other data sources include the same index variable with additional attributes. In simple terms, that means that each dataset in the combination has to have the same **matching variable (ID variable)** indicating the unique code of the spatial object (settlement). Thus, unless each data point can be clearly tied to the particular settlement it describes, data integration will **not be possible**.³⁰

Besides matching ID variables, data preparation would, ideally, require the institutions to provide the data on settlements in **comma separated value (.csv)** format. This is simply a plain excel file that consists of data (numeric or string) and does not allow excel file to hold any additional complex qualities, such as formulas, drop-down menu, colouring, etc. While many programs are capable of handling such more complex files, simple formats such as .csv guarantee readability and smooth transition between different software platforms.

Below, an example with simulated dataset in which spatial data is being integrated and prepared for the analysis is provided.

There are four files in the working directory. Three of them are components of the previously discussed shapefile (*Main file, Index file and dBase table*) and one .CSV file comprised of non-spatial data on Roma settlements. “*RomaSettlements.shp*” is *Main file* which contains spatial information. This file has multiple dimensions, the attributes of the spatial object being of interest for this explanation, as this is the place where integration of multiple datasets occurs.

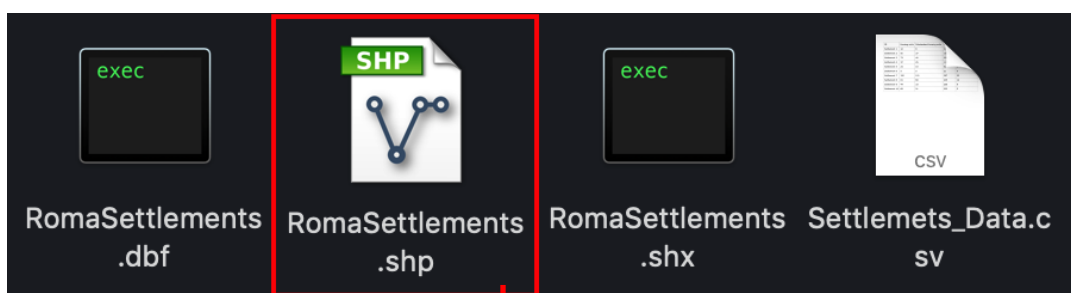


Figure 6. Organization of working directory.

A	B	C	D	E	F	G	H
ID	Municipality	Cadastral Municipality	Area Size	Coord X Min	Coord X Max	Coord Y Min	Coord Y Max
Settlement 1	Podgorica	Podgorica III	150	679253.6	680805.2	234881.6	237763.6
Settlement 2	Bar	Stari Bar	75	679127.6	681154.1	234561.2	237443.6
Settlement 3	Podgorica	Podgorica III	88	678388.9	681134.5	234918.8	237053.1
Settlement 4	Podgorica	Podgorica III	298	678237.2	680926.7	235234.9	236883.6
Settlement 5	Berane	Rujišta	11	678439.9	680954.7	235358.4	236506.9
Settlement 6	Nikšić	Glibavac	25	678559.8	680604.1	235856.4	236344.3
Settlement 7	Berane	Berane	52	678206.1	680393.3	236272.8	236419.1
Settlement 8	Nikšić	Nikšić	78	678199.5	680583.4	235439.2	237021.9
Settlement 9	Nikšić	Drenoštica	39	678268.4	680491.2	235035.4	237268.2
Settlement 10	Tivat	Donja Lastva	7	678556.5	680379.2	235028.9	237264.5

Figure 7. Attributes in the shapefile.

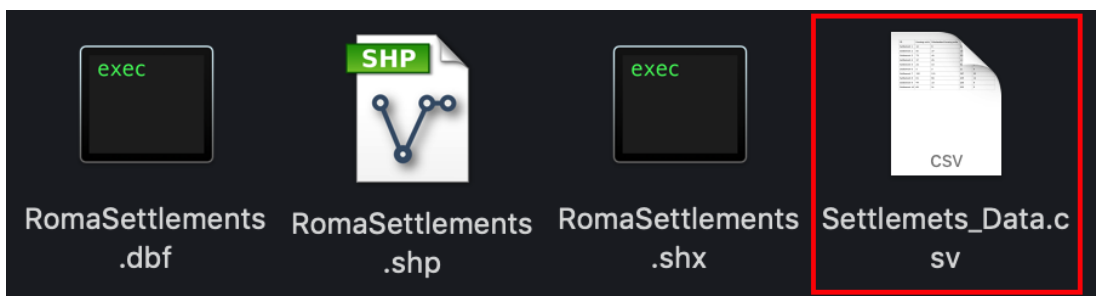
³⁰ Matching variables (settlement ID) is a string variable that is character sensitive, meaning that IDs need to match across datasets, even with regards to capital letters, spaces, etc.





The hypothetical dataset in the example, which is the file with the attributes of the mapped settlements presented in the Figure 7 above, has 8 variables. The matching variable (which is the ID) has 10 distinct settlement codes, for 10 distinct settlements. Besides name, the spatial object (the raw for each settlement) provides information on which political and cadastral municipality the settlement belongs to. The last four variables represent information on the location of settlement by showing lowest (min) and the highest (max) value of **latitude** (Coord X) and **longitude** (Coord Y). These values serve the purpose of giving broad estimation of where the settlement is located. However, the separate layer of data within the shapefile holds precise coordinates for each point in the settlement borders. In other words, the coordinates within the shapefile draw the borders of the settlement by storing information about each point between vertices K1 – K6 (Figure 5.), while the location related variables in the mock dataset represent one point in the settlement around which the settlement is spread within its precise borders.

When the mock .CSV file (“*Settlements_Data.csv*”) containing data on the mapped settlements is opened, there is the ID variable for each settlement, together with four more numeric, non-spatial variables, indicating number of housing units, substandard housing units, inhabitants, and requests for legalisation. This is shown in the Figure 8 below.



A	B	C	D	E
ID	Housing units	Substandard housing units	Inhabitants	Legalization requests
Settlement 1	12	8	36	2
Settlement 2	50	37	158	6
Settlement 3	73	40	228	8
Settlement 4	37	25	110	4
Settlement 5	24	18	82	1
Settlement 6	5	5	21	0
Settlement 7	120	101	397	23
Settlement 8	81	68	235	12
Settlement 9	44	33	256	9
Settlement 10	68	21	222	5

Figure 8. Attributes in the .CSV file.

It is clear that the matching variable has the same name (ID) as the one in the shapefile. In addition, each settlement code corresponds to the codes provided by the spatial object. Now the datasets are ready to be merged into a single file, containing both spatial and non-spatial





attributes of the settlement. Integrated dataset should match data based on the ID variable, maintain the content from the shapefile, and append the data from other sources to the corresponding settlement.

A	B	C	D	E	F	G	H	I	J	K	L
ID	Municipality	Cadastral Municipality	Area Size	Housing units	Substandard housing	Inhabitants	Legalization requests	Coord X Min	Coord X Max	Coord Y Min	Coord Y Max
Settlement 1	Podgorica	Podgorica III	150	12	8	36	2	679253.6	680805.2	234881.6	237763.6
Settlement 2	Bar	Star Bar	75	50	37	158	6	679127.6	681154.1	234561.2	237443.6
Settlement 3	Podgorica	Podgorica III	88	73	40	228	8	678388.9	681134.5	234918.8	237053.1
Settlement 4	Podgorica	Podgorica III	298	37	25	110	4	678237.2	680926.7	235234.9	236883.6
Settlement 5	Berane	Rujišta	11	24	18	82	1	678439.9	680954.7	235358.4	236506.9
Settlement 6	Nikšić	Glibavac	25	5	5	21	0	678559.8	680604.1	235856.4	236344.3
Settlement 7	Berane	Berane	52	120	101	397	23	678206.1	680393.3	236272.8	236419.1
Settlement 8	Nikšić	Nikšić	78	81	68	235	12	678199.5	680583.4	235439.2	237021.9
Settlement 9	Nikšić	Drenoštica	39	44	33	256	9	678268.4	680491.2	235035.4	237268.2
Settlement 10	Tivat	Donja Lastva	7	68	21	222	5	678556.5	680379.2	235028.9	237264.5

Figure 9. Integrated dataset.

Software solutions for data analysis

Although simplified, this mock example conveys the logic behind the process of data integration. However, manipulating shapefiles in practice is somewhat harder and requires using software platforms that are specifically designed to create, edit and visualize geospatial information. In addition to these, there are number of general programming/statistical languages that handle spatial data, but these require more advanced programming skills.³¹ Each of the alternatives mentioned are open-source software, and therefore, completely free for use.

a) QGIS

QGIS³² (previously Quantum GIS) functions as a geographic information system (GIS), allowing users to analyze and edit spatial information. QGIS is an open-source program with a graphical user interface that can be navigated using the mouse and keyboard. It supports both raster and vector layers, and numerous formats of spatial data, shapefile included. In addition to this, it integrates open-source packages and plugins, which can perform functions outside of QGIS itself, such as: geocoding (using Google Geocoding API), geoprocessing (similar to ArcGIS), and interface with MySQL databases.

³¹ In case more advanced statistical analyses are expected to be performed on integrated dataset later on, it is suggested that the software *R* is used. *R* is a programming language for statistical computing, as it provides wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, etc.) and graphical techniques. *R* community has developed a number of packages specialized for handling spatial data, such as: „sp“, „raster“, „rspatial“, „rgdal“, „rgeos“, „maptools“, etc.

³² The latest version QGIS 3.8 Zanzibar is available since September 2019: <https://www.qgis.org/en/site/>



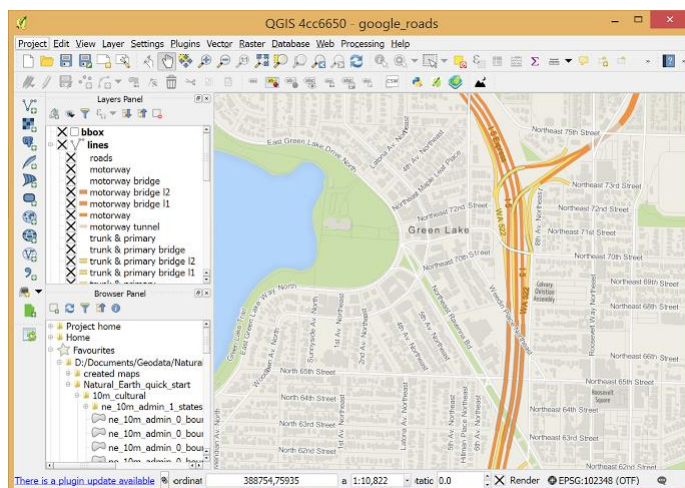


Figure 10. QGIS interface (source: <https://qgis.org/en/site/about/index.html>)

More important to this task, QGIS can be used for manipulation of both spatial (.shp) and non-spatial data (.csv). The process of matching these two types of data (known as **Table Join**), proceeds in following order:

1. Load Roma settlements **shapefile** in QGIS and save it as *Vector Layer*.
2. Create sidecar file with **.csvt** extension to tell QGIS which variable to import as *string* (text) and which as *integer* (number). This file has only one row specifying data types for each column. Save this file in the same directory as the original .csv file.
3. Import **.csv** file with non-spatial data of Roma settlements to QGIS (via *Add Delimited Text Layer*).
4. Specify that non-spatial data contains **no geometry** and select *Attribute only table*.
5. In the *Add vector join* dialog (properties) select **matching variables** with unique IDs of Roma settlements.
6. Spatial and non-spatial data on Roma settlements are now **integrated**.





6. Specific issues to be defined by each individual economy

Although this methodology for mapping of substandard Roma settlements aims to be universal for all the Western Balkans economies, there are still some specific issues that are not possible to be determined at this stage, and need to be clarified by each economy before the mapping process takes place. These mainly relate to the specific type of data that should be collected by the local administrations, provided in the sample questionnaire given in Appendix I.

Census methodologies and data by economy

With the exception of the Republic of North Macedonia, where the last census took place in 2002, all other economies have conducted most recent censuses on population and dwellings in 2011 (Bosnia and Herzegovina in 2013). Following EUROSTAT's recommendation, census methodology relied on updated spatial registry with clear definition of census circles (or enumeration areas) and their mapping in a GIS system. It is usually left to the governments to decide on the definition of census circles, but without an exception, they represent small geographical units that include cluster of buildings that can reasonably be considered living in the same locality.³³

Two types of census information are of crucial importance for this methodology – **demographic data including ethnicity** and **quality of housing units**. In all the economies covered by this methodology, census allowed for questions on ethnicity.³⁴ However, the proposed methodology is not fully applicable in Bosnia and Herzegovina, due to the fact that all ethnic minorities are merged into one category “Others”. Therefore, using census data to identify localities with Roma population is impossible. As the proportion of ethnic minorities is rather small (below 3%), this measure could be used to identify minorities and then local authorities should be able to provide more precise information on the dominant ethnicity of particular groups/settlements in Bosnia and Herzegovina.

With respect to dwellings, the label of what is here called “substandard settlements” differs across the economies. For instance, the census in Albania differentiates between “conventional” and “non-conventional” dwellings, with latter being defined as *“mobile, semi-permanent or improvised, or housing units not designed for permanent human habitation, but which were nevertheless being used at the census moment as the usual residence of one or more persons”*. This definition, substantively speaking, puts households of interest into the same category of substandard housing units/settlements. Hence, these labelling differences have no repercussions for the proposed methodology as data is being collected using uniformed methodology, including each of the characteristics for substandard settlement.

The new round of censuses is scheduled for 2021. Thus, the methodology proposed here is likely to use newly provided data to map the substandard Roma settlements. While almost all economies are finishing their preparation, not all provided the public with sufficient information regarding the potential changes in the methodology. However, as the methodology relies on essential data for every society, it is almost completely certain that the proposed methodology

³³ For example, Montenegro was divided in 3,740 census circles, while Albania had 11,712 enumeration areas.

³⁴ As a voluntary self-declared ethnicity.





shall be applicable in the future (notably after the next Census cycle). With regards to this, Bosnia and Herzegovina is the only European economy where the next census will not be conducted in 2021. At this point it is uncertain whether the new census will allow for the desired differentiation between Roma people and other ethnic minorities.

Satellite images and ortho-photos

For the most part, geo-referenced data in the respective economies is based in ortho-photos which have been recently updated: 2017 - Montenegro and Bosnia and Herzegovina and 2018 - Serbia, Albania, and Kosovo. The only exception to this rule is the Republic of North Macedonia, whose latest available orthophoto is from 2004.³⁵ In terms of geographic coverage, based on the conducted interviews, the national geoportals cover the largest portions of the territories of the economies. For instance, it has been reported that in Montenegro above 90% of the territory is covered (the percentage is higher for urban settlements), although matching with cadastral data is still in process. However, it is worth noting that the new *General plan of regulation* (together with *Detailed plan of regulation*) in Montenegro is expected to be put in use during 2020. The situation is most probably similar across the region, although more precise information on this issue has not been obtained from the respective governments.

Criteria for a position of a substandard Roma settlement relative to formal settlement

For each settlement mapped, local administration should provide information on its location relative to a formal settlement – if it is integrated or disconnected from city, town or village to which it administratively belongs. This information is important for determining the list of priorities for interventions (improvement of living conditions) when designing programmes at the national level, but also determines the intervention project on the field – e.g. should the settlement be kept in place and upgraded or resettlement is a better option for the inhabitants.

Prior to determining if a settlement is integrated, it is necessary to define the criteria for integration or disconnection. However, it is hardly possible to make one universal criteria on this issue that would be equally applicable in all the economies in the region, due to the differences with regards to the substandard Roma settlements. For example, in Serbia about half of the substandard Roma settlements are located in urban areas and the other half in rural. In the Republic of North Macedonia great majority of Roma live in urban areas. Therefore, it is clear that criteria for integration into formal settlements would be different for Serbia and the Republic of North Macedonia. Similar differences can be found in all the economies involved. During the mapping implemented in Serbia, the criteria of 1km distance was used as a parameter for disconnection – if a settlement is located 1km or more from a formal settlement, it was considered disconnected. However, no clear explanation is provided on how this parameter was defined, and it seems it is an arbitrary criterion. There is a necessity that this criterion is precisely defined at the level of individual economies and at regional level, in accordance with the experiences from the ground related to Roma housing, before the mapping process starts.

³⁵ As stated on the official website of the central Cadastre of this economy: <http://www.katastar.gov.mk/en/data/cartographic-products/orthophoto-maps/#>

