ESSnet Big Data

Specific Grant Agreement No 1 (SGA-1)

https://webgate.ec.europa.eu/fpfis/mwikis/essnetbigdata
http://www.cros-portal.eu/.....

Framework Partnership Agreement Number **11104.2015.006-2015.720**

Specific Grant Agreement Number **11104.2015.007-2016.085**

Work Package 5

Mobile Phone Data

Deliverable 1.1

Current status of access to mobile phone data in the ESS

Version 2016-30-11

Prepared by: David Salgado (INE, Spain)

Ciprian Alexandru (INSSE, Romania)
Marc Debusschere (Statistics Belgium, Belgium)
Françoise Dupont (INSEE, France)
Pasi Piela (Tilastokeskus, Finland)
Roberta Radini (ISTAT, Italy)

ESSnet co-ordinator:

Peter Struijs (CBS, Netherlands)
p.struijs@cbs.nl
telephone : +31 45 570 7441
mobile phone : +31 6 5248 7775

Contents

General introduction to work package 5

The work package 5 of the ESSnet on Big Data is devoted to mobile phone data. The work is clearly divided into two phases coincident with the administrative phases of the ESSnet, namely, the Specific Grant Agreement 1 (SGA-1) and the Specific Grant Agreement 2 (SGA-2). The SGA-1 for this work package is exclusively concentrated on getting access to mobile phone data in such conditions as to allow us to carry out a detailed research in the SGA-2 on the statistical methodology, IT requirements, and data quality framework assessment entailed by this new source of information for the production of official statistics.

We are following several parallel lines of work. Firstly, the WP members during the SGA-1 have conducted negotiations with different MNOs to have access to mobile phone data. Secondly, from this experience we are providing a preliminary analysis of the diverse aspects arising to grant access to the data. Thirdly, the SGA-1 has taken stock of the current situation of the access to mobile phone data across the European Statistical System (ESS). Furthermore, these initiatives have been complementarily used as an important input for an exploratory workshop gathering the National Statistical Institutes (NSIs) of the ESS, Eurostat, European MNOs, and some other international organizations to exchange experiences and provide examples of fruitful collaboration thus providing a solid platform for future collaboration. This meeting has focused on potential mid- and long-term partnerships. This is complementary to the broader and more demanding line of work of the SGA-1 for this work package, i.e. the bilateral negotiations between the NSIs of this work package with their corresponding national MNOs to gain access to mobile phone data to conduct the research study in the SGA-2.

In summary, the goals of the SGA-1 are (i) to undergo negotiations with MNOs to grant access to mobile phone data for the SGA-2, (ii) to produce the present deliverable to account for the current state of affairs regarding the access to mobile phone data for NSIs in the ESS, including a preliminary analysis of the diverse related aspects, (iii) to

organize a workshop to explore future collaborations between Eurostat, NSIs and MNOs, and (iv) to compile the main conclusions together with some technical knowledge about the access to mobile phone data in the final deliverable.

1

Introduction

By and large, mobile phone data are data generated from the interaction of mobile phones and portable devices equipped with a SIM card with the different telecommunication networks providing service to their clients. In consequence, these data represent the core of the mobile telecommunication industry and are (partially) processed and stored by MNOs in their digital information systems.

It must be made explicitly clear that no content of calls, messages (SMS/MMS), or Internet connections is considered whatsoever. Rigourously speaking, only the metadata of the interaction between phones and networks is under consideration.

This interaction between each mobile device and the networks is normally very frequent, thus producing a huge amount of information. For example, in the case of access to Internet services by the subscriber, we may have between 20 to more than 300 interaction events per subscriber and day. Furthermore, as we will see, a priori this information is highly distributed among the information systems of the MNOs across their covered geographical territory, although several factors may combine simplifying this situation.

Even considering only metadata of these interactions, this information is noticeably sensitive for MNOs' subscribers, since, to name just an immediate example, their displacements can be traced back. In this sense, legal support to privacy by personal data protection and to telecommunication transmissions is given in the different national legislations.

Nonetheless, the added value of this information source has been already recognized by the MNOs themselves and some of them have already begun implementing the so-called process of *monetisation* of data. They have already started to exploit statistically their mobile phone data.

All in all, the access to mobile phone data for Official Statistics producers carries several intricacies which need to be tackled. In this deliverable, taking the on-going contacts between some MNOs and the NSIs of the work package as an input experience, we report about he current state of affairs regarding the access to mobile phone data in the ESS. In this sense, we have prepared a questionnaire aiming at collecting this information and at retrieving and analysing the many aspects of the diverse situations.

In section 2 we provide the reader with a general description of mobile phone data and an overview of the phone-network interaction where they originate from. In section 3 we explain the many aspects which we have concentrated upon when designing the questionnaire. In section 4 we present the main results derived from the responses by the ESS NSIs. In appendix A we include the complete questionnaire submitted to these NSIs.

2

Mobile phone data

As we put forward in the introduction, by mobile phone data we must understand rigourously speaking the metadata associated to the interaction between mobile devices and the telecommunication networks providing service to these devices.

To describe these data we find it convenient to have an approximate idea of how they are generated. Let us consider the generic structure of a GSM-like telecommunication network. As you can see in figure 2.1, the associated information systems have three layers, namely the base station subsystem (BSS), the network subsystem (NSS), and the network management system (NMS). The BSS is in charge of the physical (electromagnetic) interaction between the mobile devices and the antennas. The information from this interaction is processed and (partially) stored by local computer systems associated to groups of antennas. Details (like their territorial distribution) depend very much on the internal organization of MNOs. Indeed, this is a highly sensitive piece of information in their business architecture. In this first layer most of the data are generated not only between devices and antennas, but also by the control systems themselves which e.g. may change antennas giving services to a mobile phone because of signal congestion or other circumstances.

Only part of this information is transmitted to the next layer. Once more, how much information is transmitted depends on the concrete configuration of the network. The NSS is in charge of different functions, among them of transmitting information to the third layer to maintain the billing service.

By and large, mobile phone data embraces all kind of data generated in these cascades of information starting in the interaction between devices and antennas upstream to the final NMS. It is very usual to find in the literature of the use of mobile phone data for statistical purposes the expression *Call Detail Record* (CDR) to refer to each register of the mobile phone data sets. Technically a CDR is a data record produced

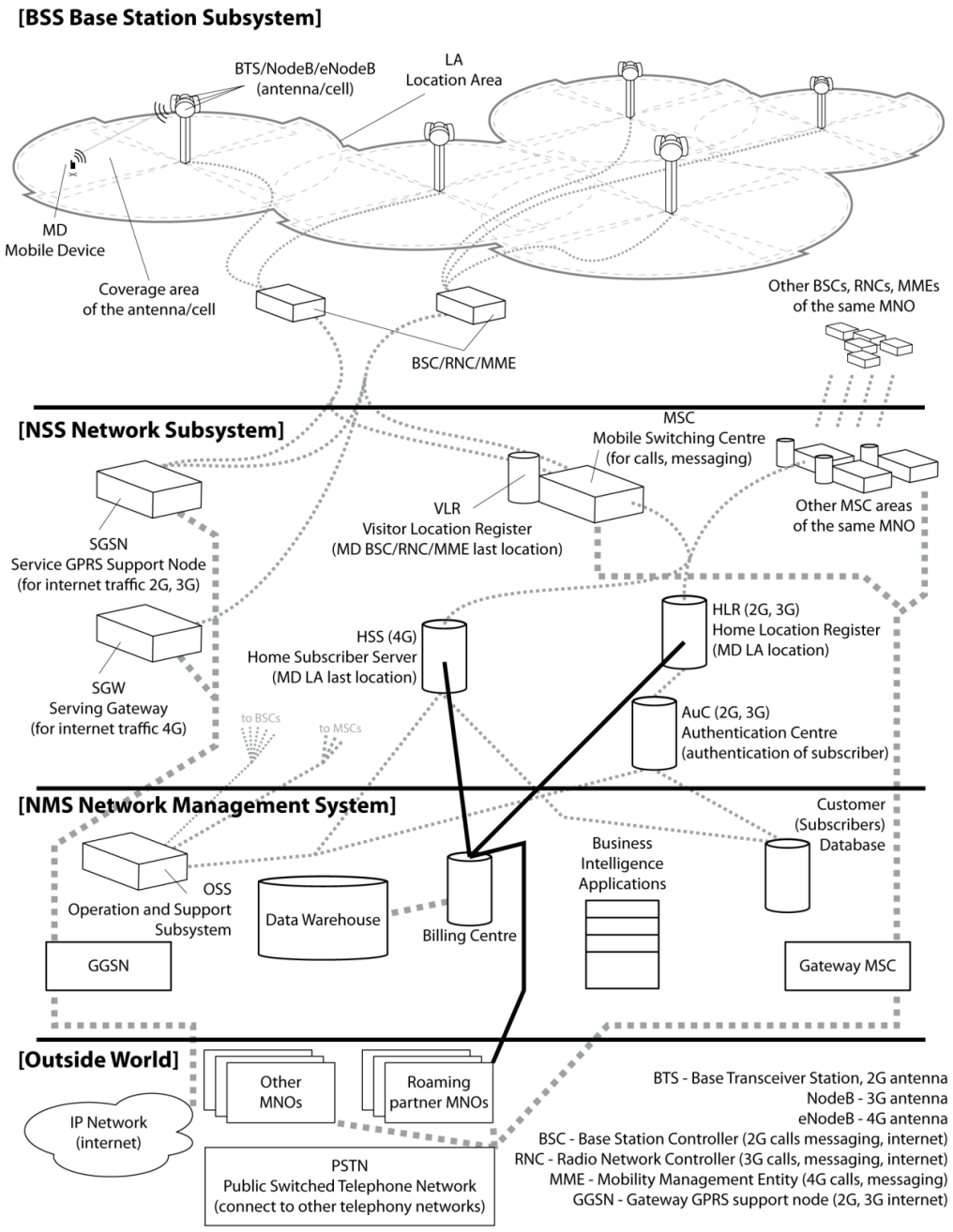

Figure 2.1 Generic architecture of MNOs' information systems (taken from [Eur+14])

by a telephone exchange or other telecommunications equipment that documents the details of a telephone call or other communications transaction (e.g., text message) that passes through that facility or device [Wik15]. In this sense, the contents of a CDR are not precisely defined and depend sensitively on the technology. In a simplistic way, a CDR originally contained the calling and receiving phone numbers, the start time, and duration of the event. More recently, CDRs may contain much more (technical) information [Wik15].

To avoid imprecisions, we will not use the term CDR. Instead, for reasons to be clear later on, we will try to be a bit more precise on the terminology. Firstly we will understand by *raw traffic micro-data* a set of unfiltered, unprocessed records, each one corresponding to an interaction event between the mobile device and the network. For statistical purposes the content of these records is limited to:

- a subscriber/device identification variable;
- a time stamp indicating the beginning of the interaction (event);
- metadata specific to the type of interaction (event type: call, SMS/MMS, Internet connection, ...; duration of the call, ...);
- depending on the characteristics of the network, an identification variable for the antenna may also be attached to each event by default.

Raw traffic micro-data needs to be subjected to some processing before they are used for statistical purposes. Partially this preprocessing is motivated by diverse reasons other than statistical (privacy protection, etc.). Immediate examples are (i) filtering out e.g. machine-to-machine events, (ii) the anonymisation of identification variables, (iii) the assignment of geographical coordinates to each device, and (iv) the coarse-graining of the time stamp variables, to name probably the most important. The preprocessed registers will be referred to as *preprocessed traffic micro-data*.

Strictly speaking this pre-processing is part of the statistical production process, and here we find a first controversial issue in the usage of mobile phone data for the production of official statistics, since the official statistics production process must be ruled by stringent official data quality standards and this preprocessing is carried out by MNOs. Are they to be subjected to these standards extraneous to the telecommunication industry?

Notice that these micro-data are not the object per se of the official statistics to be produced. Indeed this individual spatiotemporal localized data are to be used as proxy variables in the analysis depending on the specific statistical domain we are working

on. One of the major challenges in the use of mobile phone data for Official Statistics is to find out what kind of statistics can be possibly produced or enriched with them. As an immediate illustration, population dynamics arises as a straightforward example in which a priori the spatiotemporal reference seems to be a high-quality proxy variable. A less immediate example is the tourist accommodation statistics in which this spatiotemporal reference may be used as a proxy variable but the quality for this purpose is not so clear. This is the ultimate reason why we distinguish between raw and preprocessed traffic micro-data: the quality of these proxy variables is critical for the production of official statistics.

Finally, these micro-data, either raw or preprocessed, may be aggregated in different ways to produce e.g. counts of devices. Different levels of granularity can be chosen to aggregate data (both in time and in space). For the same reasons as before, these aggregation procedures are also part of the statistical production process.

In summary, the concept of mobile phone data entails some complexities not present in traditional data because, among other things, data are not generated ex profeso for statistical purposes. Mobile phone data are generated in complex information systems and can contain diverse degrees of preprocessing. In addition, they are highly sensitive with respect to the privacy of the MNOs' subscribers and indeed access to them is strongly regulated by Law.

Nonetheless, the information they carry for statistical purposes is rich and arises as a promising new data source especially for official statistics production. All these characteristics play different roles regarding the access to mobile phone data by NSIs. In the next section, making use of the experience gained during the bilateral negotiations we try to disentangle these many aspects which have served us to design the questionnaire to take stock of the current situation in the ESS.

3

Aspects of the access to mobile phone data

Having in mind the generic characteristics of mobile phone data for statistical purposes and their generation, we propose to consider the many aspects related to the access to these data by NSIs in the following categories:

- Characteristics of the MNO.
- Legal requirements.
- Access conditions.
- Data characteristics.
- Other aspects.

This categorization is somewhat artificial and is just intended to provide an enumeration of the diverse aspects entering into the question. As you will see, many of them are highly entangled so that their division does not naturally appear in practice.

3.1. Characteristics of the MNO

From our experience we have identified as a determinant condition for the access whether the MNO has a business line or a company division devoted to the commercial exploitation of the statistical information derived from their mobile phone data.

In the positive case, the company has already invested or is currently investing resources for this objective. Granting access to their mobile phone data for Official Statistics producers may be possibly perceived as a risk for this line of business.

In the negative case, we may find two extreme possibilities. Either they are not interested whatsoever in the statistical exploitation of their data or they are but they

need resources and know-how. In the first case, the access to raw traffic micro-data appears as the most evident choice. In the second case, a partnership between the MNO and the NSI to exchange resources and knowledge arises as a natural option.

In any case, the company department undertaking the contact with the NSI appears also as a relevant conditioning to achieve the access, since the sensibilities are different (say, e.g. between a research department usually eager to advance in their investigations by collaborating and a commercial department which has stringent sales targets).

It must be mentioned that there exists the possibility for some MNOs that a third-party (an associated consulting firm or similar) can act as intermediary not only regarding the data access itself but also some preprocessing. It can be considered as an intermediate case in which the statistical exploitation is somehow delegated in this firm. This case entails the issue of deciding if the access (and possibly the preprocessing) is still to be undertaken by this firm or not in the production of official statistics.

3.2. Legal requirements

Needless to say, both the content of telecommunication data and their associated metadata are diversely protected by Law in all countries of the ESS. Legal requirements are thus an important consideration when requesting access to mobile phone data.

We have identified at least three different aspects regarding legal requirements. Firstly, the new sources of information arising with the Big Data phenomenon are not clearly specified and identified in some National Statistical Acts. Thus in some cases the legal support for NSIs to request access to mobile phone data is dubious and depends occasionally on the interpretation of the statistical legislation. In some other cases, this legal support can be straightforwardly derived from current regulations but there may arise a conflict with other regulations (see below).

The need for an explicit mention to these new data sources appears recognized in the elaboration of a set of principles for the access to data for Official Statistics by the United Nations Global Working Group on Big Data for Official Statistics [UNGWG15]. These principles align the issue of the access to Big Data sources, usually in private hands, with the UN Fundamental Principles of Official Statistics [UNSD14]. Although all these principles do not have legal binding capacity in national legislations, they clearly point at their explicit consideration in the standard production of official statistics. In this sense, the first principle states that “[i]n view of the emergence of new types of data sources and the social responsibility of private organizations, these members of society are called upon to make the data that is needed available to the statistical agency

concerned, free of charge, on a voluntary basis" [UNGWG15].

The second aspect regarding legal requirements arises when considering both national and international telecommunication regulations. In most circumstances, the access to these data is strictly regulated by the rules governing the telecommunication industry and only under special conditions (emergency dial to 112, judicial mandate, national security, ...) this access is granted to a third party.

A third aspect relates to the personal data protection regulations. In many countries personal data used in any activity (not only in the telecommunication industry) are protected by a specific law so that no registry of personal information is arbitrarily stored in any public or private information system, except for the concrete purposes of its very reason of existence (population registers, tax registers, ...). Data must be erased once they have fulfilled their objectives.

In connection with this, in many countries there exists a national personal data protection authority whom you must report any kind of access to and/or request an explicit consent to. This means that a coordination between the NSI and this authority must be undertaken when requesting access to mobile phone data.

Finally, as a generic feature of the public nature of NSIs usually present in the ESS countries, all MNOs must receive an equal treatment from the Public Administration. This must be scrupulously observed even when an agreement can be easily reached with one MNO but not with other. The administrative operational setting may turn out intricate as a result of this.

3.3. Access conditions

One of the most intricate aspects in having access to mobile phone data is the actual conditions under which this access is going to be granted.

Firstly, you may request in-situ access avoiding the fact that data come out of the MNOs' information systems. This possibility alleviates the privacy and confidentiality issues, but the operational aspect must then be tackled, since the NSI will have to access somehow the MNOs' computer systems. A second option would be to transmit the data from the MNOs' to the NSI's information systems. No access to the companies' information system is needed but privacy and confidentiality issues must then be solved in advance, both from the legal and the operational points of view. Finally, a trusted third party may enter the scene whom will receive the data from the MNOs and then, possibly after some preprocessing, will transmit them to the NSI. The confidentiality

and privacy issue remains open and part of the production process is delegated.

A second condition is related to granting access only for research or also for standard production in a stable way. From our experience, we have detected a broad line separating these two extremes. The collaboration in research seems to be more feasible. The issue of production in the long term seems to need more joint work between MNOs and NSIs. The workshop mentioned in the general introduction to this work package is intended to explore potential collaborations in the long term for standard production.

In close connection with this second condition, in some cases we have identified that access is possibly granted or denied according to issues of intellectual property rights or industrial secrecy requirements. In particular, this appears when there is a risk of leaking sensitive information about the operational configuration of the network or any other element of their business production process (e.g. the territorial distribution of antennas is an extremely sensitive information for the MNOs). In this same line, especially when there is some data preprocessing involved to produce better spatiotemporal references, algorithms and procedures applied to raw traffic micro-data are kept under industrial secrecy to gain competitiveness in the market. Considerations about the protection of all this information is part of the process to achieve access to the mobile phone data.

Fourthly, as pointed out by many experts, Big Data sources will be more efficient in the production of statistics when combined with other sources. In this sense, the combination of mobile phone data with other data (e.g. registers about the use of land, population registers, georeferenced tax registers, etc.) will noticeably increase the potentiality of this information source. Thus, the possible combination of mobile phone data with other (possibly confidential official) data arises as another relevant aspect of the access to be tackled both from the legal and the operational point of view. This is especially important if this data enrichment is somehow to be shared with the MNOs or the combination is to be practised in-situ within the MNOs' information systems (e.g. requiring only to work at a certain anonymised aggregate level, assessing the representativeness of the data set without sharing population registers, etc.).

A novel aspect brought by the complexity of this information source is the cost and effort of providing data access. This does not usually appear when collaborating for research but it stands as an issue for the long term collaboration for standard production. We have already mentioned principle 1 of the UN Principles for Access to Data for Official Statistics, where this data provision is called upon free of charge and on a voluntary basis. However principle 6 explicitly states the "[t]he cost and effort of providing data access, including possible pre-processing, must be reasonable compared to the expected public benefit of the official statistics envisaged". Moreover, this is complemented by principle 3 stating that "[w]hen data is collected from private organizations for the

purpose of producing official statistics, the fairness of the distribution of the burden across the organizations has to be considered, in order to guarantee a level playing field". Having in mind the complexity of this information source depicted in section 2, these principles arise as pertinent when considering mobile phone data.

However, the issue of the cost is extremely intricate. Firstly, the essential principle of Official Statistics by which data provision for these purposes must be made completely free of charge must be also respected when considering mobile phone data. Yet, the costs associated to data extraction and data handling for statistical purposes need a careful assessment. This depends very sensitively on the concrete situation of the MNOs. For example, an MNO having developed a business line around the statistical exploitation of their data and thus having deployed an infrastructure with databases prepared for further statistical processing does not face the same situation as an MNO not having this infrastructure. Different details need consideration: staff time in data processing, hardware computing time, hardware buy and deployment (if necessary), software development or licenses (if necessary), ... In addition, the compensation for these costs may be given shape in different ways, from a direct payment to an implicit contribution to a long-term collaboration partnership.

In any case, notice that this compensation of costs is not for the data themselves, but for their extraction and handling. Data must be granted access free of charge.

Finally, being as sensitive as they are, mobile phone data require stringent confidentiality conditions. Notice that Official Statistics is already used to access, process, produce, and disseminate information under strict confidentiality conditions. However, if some legal requirements exist in the telecommunication industry, these must be also satisfied by Official Statistics, at least until legislative changes appear explicitly dealing with this issue. Both operational and methodological measures in the statistical disclosure control must be strengthened.

3.4. Data characteristics

As an essential aspect of the access to mobile phone data we must consider what data in detail we are requesting access for. As explained in section 2, mobile phone data sets can be composed of very different pieces of information. These pieces of information and the nature of the data will strongly influence the question of the access.

Firstly, we recognize as very different situations whether we request access to raw traffic micro-data or to preprocessed traffic micro-data or to aggregated traffic data (see section 2 for their definitions). Again, this is not an isolated question, but it has

deep implications in many other aspects. For example, in this first case we must be ready to handle a huge volume of data (is transmission really an option?), indeed data directly from the business core of the MNOs which are under strict telecommunication regulations; in the second case, data volume is lower but details about the preprocessing need to be known for the production of official statistics and confidentiality is still an important issue; in the third case, confidentiality issues are alleviated but control and monitoring of the preprocessing needs a closer collaboration.

Secondly, it is necessary to be aware of the type of events, i.e. of the interaction between mobile devices and antennas that feed the data sets. They may go from calls and text messages to more passive connections (e.g. location area updates). The more events, the richer the information, although also more difficult to extract.

Thirdly, these data can be complemented with other data like more technical information about the interactions or socio-demographic variables of subscribers (from their contracts). The territorial and time coverages of data are equally important. Usually for research purposes both are limited (e.g. only data from a province and four or five weeks). The access for standard production must be as ample as possible. Special attention is needed in the case of roamers. It is highly recommended to have access to data from roamers. Sometimes some previous filtering is practised upon the data, e.g. excluding machine-to-machine communications. These details are needed to be known for the production of official statistics.

In the case of preprocessed traffic micro-data it is specially critical to know how the mobile device identification variables are treated. In general mobile phone micro-data are anonymised and for the statistical analysis it is very important how the anonymised identification variable is assigned to each mobile device, especially whether they are constant over time or not.

Equally important it is necessary to know how the spatial references have been assigned to each mobile device register in the data sets. There is a large variety of ways to locate mobile phones and this will presumably influence the quality of the final statistical output.

In the same footing, the time references can also be preprocessed, e.g. being replaced by the identification of a coarse-grained time period (one hour, one quarter of a hour, etc.). This must be made clearly explicit.

Complementarily, any other kind of further preprocessing must be explicitly known to account for their effects in the statistical analysis.

In the case of aggregated data, there are three aspects of immediate interests: (i) what aggregated statistics are to be transmitted to the NSIs (e.g. counts of devices), (ii) whether the aggregation is undertaken using longitudinal identifications (through persistent identification variables) or not, and (iii) what kind of breakdowns both in time and in space is decided (Voronoi polygons, fixed-size grid, NUTS regions, etc.).

3.5. Other aspects

There exist some other aspects related to access to mobile phone data which play an important role especially when considering data provision for the long-term standard production.

Firstly, the largest MNOs have already developed or are on the verge of developing a business line around the statistical exploitation of their mobile phone data. The market shares of these companies across several European countries implicitly mean large proportions of population providing data. Thus it is critical to have access to data for these MNOs.

Now, a compelling issue arises. Apparently, if MNOs are to commercialise products based on the statistical analysis of their mobile phone data and NSIs are to disseminate free of charge official statistics based on these same data, we have a conflict of interests. However this line of thought is only apparent and needs a deeper analysis jointly by both MNOs and NSIs. The situation is similar in traditional survey sampling. NSIs have been offering official statistics for decades and there is no collision of interests with opinion and market research firms selling analytic products and studies. Furthermore, it is just the opposite: official data are of great complementary value for these studies. There is ample room both for public and private interests in the statistical use of mobile phone data. A deeper common exploration for a win-win collaboration arises certainly as the best choice.

In addition, there arises another important aspect related to the access of mobile phone data for official statistics purposes. It is the reaction of public opinion to the potential use of such highly sensitive private data for public purposes and especially when these data may come out of the MNOs' information systems. We do not mean how NSIs will deal with this idea of using such private data for their production but how public opinion about the MNOs themselves can affect their decision to provide these data. The use of these data needs probably a campaign tackling the issue of NSIs potentially becoming a kind of Big Brother. Perceptions will need to be analysed[1]. We

[1]To illustrate this point, most people accept the use of their data by software and telecommunication companies when downloading applications in their mobile devices and pose no objection to this (e.g. to

are referring to the delicate point that a bad perception by subscribers about their MNO sharing data with an NSI may be perceived as too high a risk for clients loss. Thus the company may feel compelled to refuse to provide data. This is another ingredient in the issue of access to mobile phone data.

In summary, nearly all these aspects have been the objectives in the design of the questionnaire submitted to the ESS NSIs. The questionnaire is included in the appendix A. In the next section we analyse the main results derived from the responses.

tracking their position through their GPS device). The sensibility is different when the use is made by a public institution like an NSI. In contrast, no one would object to an NSI computing the Consumer Price Index (sometimes related to their salaries), but would it be the same if the CPI is computed by private organizations? It is a matter of perceptions.

4

Results

4.1. The general setting

The designed questionnaire included in appendix A was submitted to all NSIs in the ESS with a focal point in the task force and/or the ESSnet on Big Data on June 29, 2016 and responses were finally collected in September, 2016. The following chart depicts the general view of the current state:

Reference time period: September, 2016.

ESS NSIs	32
Surveyed NSIs	28
Responses	25
NSIs having contacted MNOs	14
with only one MNO	7
with more than one MNO	7
NSIs having access to mobile phone data	7

Table 4.1 Overview of the current status of the access to mobile phone data in the ESS.

As a snapshot of the mobile phone market shares in Europe, we include in table 4.2 the reported anonymised distribution of market shares in each country together with their Shannon entropy $\mathcal{H} = -\sum_{k=1}^{n} p_k \log p_k$ as a measure of concentration. We include in the fifth column the value of the Shannon entropy for the case of equidistribution ($p_k = \frac{1}{n}$) for the corresponding number n of MNOs in each country. This is intended to facilitate a direct comparison of the actual distribution with the equidistributed situation.

The distribution of market shares may have a potential impact both in the negotiations (in case of dominant positions, e.g.) and in the estimation procedures to be tackled

later on (e.g. due to diverse population coverage).

By and large, markets across countries in the ESS do not show a high concentration. As a first general conclusion, to have an almost full population coverage in each country with mobile phone data, all MNOs will need to provide their data in each case.

Country	**MNOs**	**Share**	$\mathcal{H}$	$\mathcal{H}_{\text{equidis}}$
AT	3	40.2 - 28.2 - 29.2	1.08	1.11
BE	3	40.3 - 30.9 - 28.8	1.09	1.11
BG	3	25 - 38.8 - 36.2	1.08	1.11
CY	3	62 - 33 - 5	0.81	1.11
CZ	3	37 - 39 - 24	1.08	1.11
DK	3 + Others	40 - 25 - 20 - 15 (Others)	1.32	1.39
EE	3	32 - 32 - 36	1.10	1.11
FI	3 + Others	37 - 33 - 28 - 2 (Others)	1.17	1.39
FR	4 + Others	34.96 - 21.44 - 17.29 - 16.24 - 10.06 (Others)	1.52	1.61
DE	3	36 - 26 - 38	1.09	1.11
EL	4	45.2 - 35.1 - 19.5 - 0.2	1.12	1.39
HU	3 + Others	40 -30 -20 - 10 (Others)	1.28	1.39
IE	3	NA	-	1.11
IT	4 + Others	32.4 - 26.4 - 22.8 - 10.9 - 7.5	1.50	1.61
LV	3	13 - 44 - 43	0.99	1.11
LU	4	47.5 - 32.5 - 17.5 - 2.5	1.13	1.39
NL[1]	4	41 - 22 - 34 - 3	1.17	1.39
PL	4	26.7 - 27.9 - 21.2 - 23.6	1.38	1.39
PT	3	47.5 - 31.8 - 18.9	1.03	1.11
RO	4	31.5 - 41.1 - 23.1 - 4.3	1.20	1.39
SI	3 + Others	47.9 - 30.5 - 15.4 - 6.2 (Others)	1.18	1.39
SK	4	45 - 40 - 15 - 10	1.24	1.39
ES	4 + Others	30 - 25.6 - 27.33 - 6.46 - 10.23 (Others)	1.46	1.61
SE	4	35.7 - 27.3 - 17.2 - 13.0	1.29	1.39
UK	4 + Others	35 - 20 - 25 - 10 - 10 (Others)	1.50	1.61

Table 4.2 Market shares in the surveyed countries.

NB: These market shares have been reported within the questionnaire; they do not come from a homogeneous authoritative source. Thus they should not be used to perform statistical computations as they are included only for illustrative purposes.
[1] In revenue, not in number of subscribers.

In the following we give details about the different aspects included in section 3 to convey the current situation across the ESS. Only those NSIs having contacted and started negotiations with MNOs were requested to provide information about these aspects so that this reported information is purely based on their experience.

4.2. Characteristics of the MNOs

Since this document is especially focused upon NSIs and other official statistics producers within the ESS, we will adopt their point of view and consider MNOs within the same corporation operating in different countries as different MNOs (counting by corporation and country). This is not an exact reflection of the actual situation, since some corporations share their resources and strategies across the different countries they are operating in. However, some others do not. From the statistical producers' point of view, contacts and negotiations are undertaken only in a strictly national key despite the fact that MNOs possibly take decisions in an internal corporate-scale way.

Under this assumption, 20 MNOs are reported to be contacted by 12 NSIs in total, with the breakdown depicted in table 4.3. Most NSIs seem to have begun their contacts with just one MNO, which is connected with the exploratory nature or research environment under which these first actions are being undergone.

NSIs	**contacting these numbers of MNOs**	
8	1	
2	2	
4	3	
Total	24	**Granting access**
with stat business	12	3
without stat business	6	4
NA	6	3
	Total	10

Table 4.3 Contacted MNOs having a statistical business line and granting access to data.

Regarding the current state of negotiations with these 24 contacted MNOs,

- in 10 of them an agreement has been reached to share their data for the SGA-2;
- in 9 of them contact is maintained and negotiations are currently either starting or at a preparatory stage;
- in 5 of them, after some initial contact, this has been interrupted due to diverse situations (legal obstacles, ...).

Although it is sometimes reportedly difficult to know whether you are negotiating with the MNO corporation itself or with an orbiting enterprise devoted to the statistical exploitation, processing and analysis of their mobile phone data, at least in 9 MNOs such an annex business unit was clearly recognized.

Indeed in practice the contact with the MNOs is carried out by possibly different departments depending on the company. Four distinct roles within the corporations have been reported:

- the legal department;
- the research department;
- the commercial/sales/marketing department;
- a business oriented (in the long term) unit.

These exchange their internal points of view so that in practice NSIs must assume they must tackle all angles (legal, research, financial, ...). No single negotiation concentrates on a single aspect.

4.3. Legal requirements

4.3.1. Statistical legislation

For the 14 NSIs having contacted MNOs to request access to mobile phone data, none reports a clear definitive support from their corresponding national statistical legislation to do so. As expected, the situations are diverse and intricate, since entanglement with telecommunication regulations and especially with personal data protection legislation is always present.

2 countries have promoted a change in their National Statistical Acts, which are currently under discussion. The proposed changes are intended to provide a clearer endorsement to the corresponding NSI for requesting access to mobile phone data for statistical purposes.

Another 2 NSIs have reported to have support only when data are requested for research purposes. They have stated that in the mid-term a change in the Law will presumably be necessary.

3 NSIs report that their current National Statistical Acts does not provide support for requesting access to mobile phone data, although the main obstacles arise because of a

clash between the statistical and data protection legislations.

4 NSIs indicate that their current National Statistical Acts can admit a positive interpretation to endorse the request for statistical purposes but without a clear pronouncement in 3 of them in the current wording of the Law. However, one of them reports about stringent conditions upon the data extraction and data preprocessing tasks carried out by the MNOs before transmitting the data to the NSI. These conditions arise from the personal data protection and telecommunication regulations even though the statistical legislation endorses the office to request mobile phone data.

1 NSI reports about the relevance of their National Data Protection Authority, mentioning the role of the ePrivacy Directive in the question (apparently in the negative sense).

The remaining 2 countries do not report about the statistical legislation since their contacts are in a very preliminary stage.

4.3.2. Telecommunication regulations

Apart from some not available answers due to the preliminary stage of the contacts, the situation regarding the telecommunication regulations is different in each country. We can underline some common features:

- In no case NSIs are legally recognized within the telecommunication regulations to have access to individual data.
- Access to some form of anonymised aggregated data is indeed possible, although in most cases the corresponding National Data Protection Authority must validate and authorize the procedure.
- Telecommunication regulations in many cases provide the users' individual consent as a possible mechanism for using their data for other purposes other than the telecommunication service itself.
- In many cases, the National Data Protection Authority is pointed as a necessary actor.

4.3.3. Data protection regulations

As in the preceding section, the situation here varies from country to country with diverse responses from the NSIs about their current case. The main common features are:

- No access to identified individual data are provided not even for statistical purposes. Anonymisation is compulsory. No traces of identified individual behaviour must be feasible in the data set.

- In many cases there exists a diffuse line around the right to access individual or aggregated data. NSIs tend to solve this by focusing so far only on aggregated data.

- In some cases the national data protection legislation explicitly states the legal endorsement to access the data as long as this is for statistical purposes (among others). This is reported not to be enough even for accessing aggregated data.

- In some other cases, positive statements by the National Data Protection Authority are declared not to be enough for accessing the data.

4.3.4. The negotiation strategies

Each NSI has followed its own strategy to approach the MNOs and to establish an agreement. However, many lines of argument are fairly similar and revolve around the idea of pursuing a win-win collaboration:

- The production of official statistics will benefit from having access to a wealthy source of information and the production of commercial statistical products will increase its quality by integrating the statistical expertise of official statisticians and the combination of official data.

- The issue of representativeness can be tackled in a more authoritative fashion by NSIs given their long experience and dealing with official data. Also the expertise in statistical disclosure control and modelling could bring higher quality standards to the production of statistical outputs.

- MNOs' profile of social corporative responsibility will be strengthened by collaborating with NSIs not only for making data available for the public good but also by supporting with these actions a European research project for the ESS.

- Even when an agreement is reached in at least one case with no cost for the data extraction operations, this is considered apart (especially as a set up cost). In this sense minimum cost is seeked on any compliance of data delivery.

- In at least one case mobile phone data access request is carried under the same lines as traditional survey data collection.

4.4. Access conditions

Three main modalities of access to data are recognized, namely (i) in-situ access in the MNO's premises, (ii) transmission of data sets from MNOs to NSIs, (iii) access via a trusted third-party. This information is especially illustrative when related to the characteristics of data (see below).

It is immediate from table 4.4 that raw microdata are shared between MNOs and NSIs only in two countries. By and large, aggregated data are more favoured probably because it noticeably alleviates the issues of confidentiality and privacy.

	In-situ	Transmission	Third-party	Not decided	Preliminary stage
Raw microdata	0	2(2)	0		
Preprocessed microdata	2(1)	1(1)	0	2	3
Aggregated data	0	3(2)	1(1)		

Table 4.4 Access type vs. data characteristics for NSIs having contacted MNOs. In brackets only NSIs having access.

In most cases (9 NSIs), the negotiations focus on access under research conditions, i.e. data are not intended to be used in a statistical operation disseminated in standard conditions and under the national statistical plans in force. Only in two cases the negotiations are including the long term for standard production. In the case of the three NSIs still in a preliminary stage, they have not reached the point of deciding this aspect.

The access to data is negotiated under no restriction in the case of 4 NSIs. However 5 of the NSIs have found some explicit conditions on the access:

- Access is granted only for research purposes.
- MNOs' technology and methodology cannot be disseminated and shared.
- Data extractions from MNOs' information systems must not be too costly and not privacy-sensitive.
- The granularity of the statistical outputs to be disseminated is to be agreed and must be fairly non-disclosive.
- Access is based on compliance with a legal obligation.

The remaining NSIs either have not provided an answer or have not yet reached this point in their contacts.

Regarding the combination of data, in 6 cases no such a combination is foreseen in the negotiations, especially in the microdata level (in at least one case this has been explicitly forbidden by the National Data Protection Authority). However, in 3 cases this is included. For the rest, no answer is reported or they have not yet reached this point.

Regarding cost compensation, 5 NSIs have negotiations under "closed-wallet"conditions so that no costs enter into play, although in one case some compensation may be considered outside the agreement. In another 3 NSIs costs have to be considered. In no case, these costs are for the data themselves but for diverse operational arrangements (mostly, MNOs' staff time in data processing, software development and hardware computing time). For the rest, no answer is reported or they have not yet reached this point. The reported cost compensations are planned to be dealt with either under a public tender or via a direct payment to the MNO.

Dealing with privacy and confidentiality of subscribers is accomplished in similar lines:

- A threshold in the number of mobile phone users is set under which no query output will be delivered to NSIs.
- Working with truly anonymised aggregated data will guarantee the protection of privacy and confidentiality.
- In some cases, further methodology is applied not to allow individual disclosure out of longitudinal information coming from each one's historical piece of data.
- Secured computer environment is put in place to process data with highly restricted access.

4.5. Data characteristics

An important issue regarding the negotiations with the MNOs commonly appearing in the contact with them is: exactly what data are we requesting access upon? It is important to realize that mobile phone data for statistical exploitation must be prepared out of network events between handsets and antennas. These prepared data sets can be of diverse nature. For our analysis, we have distinguished among raw micro-data, preprocessed micro-data and aggregated data (see section 2).

The current situation is depicted in table 4.5 for the 14 NSIs having contacted with MNOs. A priori data at the mobile phone level arise as an ideal scenario to conduct research, but aggregated data, apart from alleviating some issues regarding confidentiality,

privacy, and legal regulations, can also be so promising (up to further research).

Raw microdata	2
Preprocessed microdata	3(2)
Aggregated data	4(3)
Not yet decided	2
At preliminary stage	3

Table 4.5 Data characteristics for NSIs having contacted MNOs. In brackets only NSIs having access.

As stated, data may be composed making use of different network events. In this survey, only 3 NSIs included passive events in the source of their data, not only events coming from calls and messages (which mostly give rise to the so-called *Call Detail Records*). Another 5 NSIs will use only calls and possibly messages (thus, CDRs) while the rest (6) is at a preliminary stage (3) or has not provided an answer (3). These data are reported to be complemented with supplementary variables other than an anonymised ID variable, a timestamp and geolocation variables (such as duration of the event, etc.) only in the case of 4 NSIs.

For the remaining questions, only 9 of the 12 NSIs having contacted MNOs have provided a response. Data are requested for all the national territory in the case of 7 NSIs and for a region or province in the case of 2 NSIs. Regarding the time dimension, in 3 case the access is settled in a permanent basis whereas the remaining 6 aims at time periods ranging from 1 natural year down to a few weeks even down to a couple of days at an initial stage.

A variable indicating whether the handset is in roaming or not is relevant for many statistical analyses. 4 NSIs are not having access to this information, while 5 NSIs have included it in their negotiations.

In all but one cases, the subscriber/user ID variable is anonymised by the MNOs, being kept unchanged in the data set under analysis. In the single remaining case, the NSI itself anonymises the data set. However, in some countries this is not possible any more due to regulatory changes and in some others the anonymised variable must be changed in time depending on diverse circumstances.

At the current status of the access, we do not have detailed information for this survey about the preprocessing of the timestamp and geolocation variables, which in some cases are substituted by a coarse-grained time reference and in others by antenna

tower IDs or a larger geographical reference (e.g. municipality). This preprocessing will be further clarified in later stages of the work of this WP.

5

Future plans

For the rest of this work package, our plan is as follows:

1. To complete the actions corresponding to the SGA-1 of the ESSnet for this work package, we will produce a deliverable directed mainly to NSIs including guidelines for their own negotiations with MNOs. This will be our deliverable 1.2.

   This document will be based upon internal technical reports about mobile phone data extraction, guidelines from a business perspective gained from the celebration of a joint workshop between NSIs and MNOs within the ESS, and generic guidelines coming from our own experiences during the negotiations with MNOs to get access to data for the SGA-2.

2. We will produce a concrete statistical output (mainly, population counts of general population, tourists and commuters, depending on the agreed access conditions) with the goal to research on the statistical methodology with emphasis on (i) the recognition of statistical definitions (e.g. of tourist, of commuter, ...) upon the data (e.g. using machine learning techniques), (ii) the inference from the mobile phone data set to the whole population of interest, and (iii) the determination of accuracy measures of the estimates.

3. We will document the technological needs required to produce the statistical output mentioned in the preceding point, especially if novel computer tools and techniques are needed (distributed computing, new file systems, specific programming languages, etc.).

4. We will assess diverse quality aspects of the produced statistical output, with special emphasis on the accuracy dimension.

The last three actions will be undertaken during the SGA-2 phase of the project starting in January, 2017.

The present deliverable can be downloaded from the wiki page of the project at `https://webgate.ec.europa.eu/fpfis/mwikis/essnetbigdata/index.php/WP5_Mobile_phone_data`.

Appendix A

The questionnaire

We include the complete questionnaire submitted to the ESS NSIs for their completion.

The questionnaire was designed exclusively as a Word file and not as electronic online questionnaire because it may contain sensitive information about the negotiations between each NSI and national MNOs.

These questionnaires have been transmitted only via email and they have not been upload to the web page of the ESSnet.

No dissemination nor distribution nor sharing with third parties (only Eurostat and ESSnet members for analysis purposes) whatsoever have been undertaken.

We explicitly acknowledge the participation of A. Wirthmann, F. Reis, and C. Demunter, from Eurostat, in the design of the questionnaire.

QUESTIONNAIRE ON THE STATE OF THE ACCESS TO MOBILE PHONE DATA FOR OFFICIAL STATISTICAL PRODUCTION PURPOSES IN THE ESS

This questionnaire has been designed and built by the work package on mobile phone data of the on-going European research project *ESSnet on Big Data*.

The questionnaire is intended to take stock of the current state of the access to mobile phone data for official statistical production purposes in the ESS. It collects precise information about diverse aspects of the access to these data.

This information is important because, among other things, it will be used as an input for the preparation and celebration of a forthcoming meeting in Luxembourg gathering NSI members of the work package, Eurostat, and European Mobile Network Operators (MNOs) to explore potential private-public partnerships around the access to these data for official statistics purposes in the mid- and long-term.

Additionally, the responses will allow the ESSnet project to produce public deliverables with the analysis of this information together with the input obtained from the workshop.

The responses will be treated confidentially both by the members of the ESSnet project and by Eurostat. No individual information will be shared, distributed or disseminated among third parties whatsoever.

SEND FILLED-IN QUESTIONNAIRE VIA EMAIL TO david.salgado.fernandez@ine.es; elisa.esteban.segurado@ine.es

General issues

0.1 Country:

0.2 Please list the MNOs of your country (with, if applicable and possible, the international group they belong to and the approximate market share):

MNO	Group	Market Share (%)	Other Comments (purpose of negotiations)

0.3 Have you already started negotiating data access with MNOs or other entities?

☐ No.

☐ Yes, with only one MNO.

☐ Yes, with more than one MNO. Then: ☐ One to one, separately for each MNO.

☐ Concurrently, simultaneously with several MNOs.

0.4 Comments (if you are aware, indicate whether other public administration in your country has access to mobile phone data for official statistics purposes)

- **If you have NOT started negotiating data access, we thank you for this information and your comments, if any.**

- **If you HAVE started or successfully finished you negotiations, please complete the rest of this questionnaire ONCE FOR EACH MNO you are negotiating with.**

Issue 1. MNO, state of negotiations, and legal requirements

(PLEASE FILL IN THIS ISSUE FOR EACH MNO SEPARATELY)

MNO:

1.1. Name of the MNO:

1.2. In the third case of negotiations with more than one MNO, these are conducted

☐ one to one, separately for each MNO.

☐ concurrently, simultaneously with several MNOs.

1.3. Please describe at what stage you consider the negotiation to be (preparatory, starting, well advanced, near finalised, …) and comment:

1.4. Does the MNO have a business line or a company division focused on the commercial exploitation of statistical information derived from their mobile phone data?

☐ No

☐ Yes. Please briefly comment:

1.5. In the case of on-going negotiations directly with the MNO, which department(s) are you negotiating with (commercial data exploitation, research, business inteligence, marketing and sales ,...)?

1.6. In the case of negotiations with a third entity, could you please provide a brief description of the role of this intermediary?

According to our preliminary pilot experiences, the legal support for granting access to mobile phone data is at least three-fold since it involves legislation about personal data protection, telecommunication data protection and statistical regulations (National Statistical Act,…).

1.7. Does your national statistical legislation provide enough legal support to the NSI for requesting access to mobile phone data?

1.8. Does your national personal data protection legislation prevent your National Statistical Institute from accessing mobile phone data?

1.9. Does your national telecommunication legislation prevent your National Statistical Institute from accessing mobile phone data?

1.10. In the affirmative case of any of the three preceding questions, do you have to fulfil any special requirement (reporting to a national data protection authority about the access, special authorization to users,…)?

1.11. Please could you give a brief description of your main line of argument when approaching the MNO (e.g. just requesting mobile phone data as in traditional data collection procedures by legal mandate, seeking win-win collaboration, offering statistical know-how in exchange, …).

Issue 2. Access

(PLEASE FILL IN THIS ISSUE FOR EACH MNO SEPARATELY)

MNO:

2.1. Are you negotiating in-situ access (in MNOs' premises), transmission of data from MNOs' to NSI's premises or access by a trusted third-party partner?

☐ In-situ access.

☐ Transmission.

☐ Third-party.

☐ Not yet decided.

2.2. Are access conditions negotiated/under negotiations for R+D (feasibility) or for production in the long-term? (Mark both if necessary)

☐ R+D

☐ Production

2.3. Is access conditioned by output/methodology dissemination or intellectual property issues? (E.g. data access only granted provided no dissemination of detailed outputs is made, data access granted provided that the data extraction methodology is not shared or disseminated…).

☐ No restriction

☐ Restricted. Please briefly describe these restrictions:

Combining mobile phone data and official data, e.g., from the Census, is a promising course of action to obtain better statistical outputs. This combination may be difficult or occasionally impossible depending on the access conditions.

2.4. Is official data combination foreseen/included in the access conditions? Please briefly comment.

☐ No data combination.

☐ Official data to be combined with mobile phone data.

Please briefly describe how this combination is foreseen/included (in MNO's premises, in NSI's premises, only at aggregate level and what level of aggregation, what kind of official data,…):

2.5. Is there any special condition requested by the MNOs in the access agreement?

(E.g. access only granted provided they are used exclusively for a concrete statistical operation, access only granted if the MNO gets in return data sets enriched by Official Statistics…)

Data extraction and data preprocessing (if any) from MNOs' digital information systems can entail certain costs for the MNO.

2.6. Is compensation for these costs considered in the agreement?

☐ No.

☐ Yes.

2.7. In the affirmative case, what kind of costs are under consideration? (Mark as many as necessary)

☐ Staff time in data processing.

☐ Hardware computing time.

☐ Hardware buy and deployment.

☐ Software development.

☐ Software licenses.

☐ Intellectual property rights or similar.

☐ Other. Please specify:

2.8. In the affirmative case, could you give some detail about how cost compensation is dealt with in the agreement?

☐ Direct payment.

☐ Through a public tender.

☐ Grant agreement in terms of research collaboration.

☐ Costs implicitly dealt with in a collaboration partnership with no financial transactions involved between the MNO and the NSI.

☐ Other. Please specify:

Data confidentiality must be taken into account both when considering mobile phone data as input data and when considering final statistical outputs.

2.9. How is data confidentiality in the input mobile phone data dealt with in the agreement?

2.10. How is data confidentially in the statistical output dealt with in the agreement (e.g. strict application of statistical disclosure control procedures)?

Issue 3. Data

(PLEASE FILL IN THIS ISSUE FOR EACH MNO SEPARATELY)

MNO:

Raw traffic micro-data from the mobile phone network consists of an unfiltered list where each record corresponds to one communication event between the mobile device and the network with the following information:

- **user/device ID;**
- **time stamp;**
- **data specific to the type of communication (e.g. type of event: call, SMS, Internet connection, …; duration in the case of a voice call).**
- **depending on the characteristics of the network, antenna/cell ID may also be attached to each event by default.**

Raw traffic micro-data may be pre-processed by the MNO before being made available to the NSI, while being kept still as micro-data. Such pre-processed traffic micro-data consists of a list where each record corresponds to one communication event between the mobile device and the network, but where the data have been somehow processed. For example:

- **the list may have been filtered (e.g. excluding devices from foreign networks on roaming, excluding devices identified as being operated by automated machines);**
- **user/device ID is generally anonymised by assigning a blind ID (maybe changing after a period of time, e.g. every day);**
- **the antenna/cell ID may have been converted to a set of geographical coordinates, to an ID of a Voronoi polygon or an ID of a grid cell;**
- **the time resolution of the time stamp may have been reduced (e.g. indicating only the hour of the day);**
- **some other data may have been included (as subscribers' sociodemographic information).**

Traffic micro-data (raw or pre-processed) may be aggregated by the MNO into counts of devices before being made available to the NSI. The counts may be cross-sectional or longitudinal. Longitudinal data is obtained by linking micro-data records by user/device ID and in this case additional variables may be derived from different communication events. The aggregated data can then be broken down by time and space at specific levels of granularity, and also by other variables.

3.1. Does your negotiation/agreement with the MNO consider access to raw traffic data, preprocessed traffic data or some form of aggregated data?

☐ Raw traffic micro-data

☐ Preprocessed traffic micro-data

☐ Aggregated traffic data

Source data

3.2. What is the type of events registered in the traffic micro-data (even if it is aggregated before being made available to the NSI)? Mark as many as necessary:

☐ Calls

☐ Messages (SMS/MMS/…)

☐ Internet connections

☐ GPS connections

☐ Passive connections between the mobile and the network (e.g. location area updates)

☐ Other. Please specify:

3.3. What other MNO data are you considering access to, either combined with traffic data (e.g. socio-demographic variables from subscribers) or as auxiliary data (e.g. location of antenna towers)?

☐ Technical data about network connections (possibly including duration and type of events…)

☐ Socio-demographic variables of subscriber

☐ Location of antenna towers

☐ Others. Please specify.

Filtering

3.4. What is the spatial coverage of the data?

☐ All national territory

☐ Part of the national territory. Please specify:

3.5. What is the time coverage of the data?

☐ One time period. Please specify:

☐ Permanent.

☐ Non-consecutive time periods (e.g. every Wednesday). Please specify:

3.6. Are data of devices on roaming included in the dataset?

☐ No.

☐ Yes.

3.7. Is there any additional filtering applied to traffic micro-data (e.g. excluding machine-to-machine events, excluding business subscriber events, devices on roaming)?

Other pre-processing

3.8. In case the NSI has access to traffic micro-data, is the user/device ID anonymised?

☐ No.

☐ Yes.

3.9. In case the user/device ID is anonymised, is it kept unchanged?

☐ Yes, the user/device ID is kept unchanged in all the dataset.

☐ No, it is kept unchanged only for a certain amount of time.

Please specify how long it is kept unchanged:

3.10. Were the spatial coordinates pre-processed?

☐ No, the spatial coordinates are represented by a tower/cell ID.

☐ Yes, the tower/cell ID was replaced by their corresponding geo-coordinates.

☐ Yes, the tower/cell ID was replaced by the ID of a Voronoi polygon built based on the location of the antenna towers.

☐ Yes, the tower/cell ID was replaced by the ID of a grid cell built based on the location of the antenna towers.

☐ Yes, the tower/cell ID was replaced by a transformed variable other than the ones indicated above. Please specify:

3.11. Was the temporal variable pre-processed?

☐ No, the temporal variable is represented by a timestamp.

☐ Yes, the timestamp was replaced by the identification of a certain period of time (e.g. periods of 15 min. or 1 hour). Please indicate the time period:

☐ Yes, the timestamp was replaced by a transformed variable other than the one indicated above. Please specify:

3.12. Please indicate any other pre-processing applied to the raw traffic micro-data.

Aggregated traffic data

3.13. Are the aggregates based on longitudinal traffic data (i.e. traffic micro-data linked by the user/device id)?

☐ Yes

☐ No

3.14. Please indicate the aggregate statistic to which the NSI would have access (e.g. number of devices)?

3.15. Please indicate the breakdowns of the aggregate statistic to which the NSI would have access (e.g. hour of the day, spatial grid cell)?

Bibliography

UNStats (2014). *Fundamental principles of Official Statistics*. URL: http://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx.

Eurostat et al. (2014). *Feasibility study on the use of mobile poistioning data for tourism statistics*. Tech. rep. Eurostat.

Global Working Group on Big Data for Official Statistics (2015). *Towards principles for access to data for official statistics*. Tech. rep. United Nations.

Wikipedia (2015). *Call detail record*. URL: https://en.wikipedia.org/wiki/Call_detail_record.