Processing of Mobile Network Operator data for Official Statistics: the case for public-private partnerships

Fabio Ricciato[(*)], Freddy De Meersman[(+)], Albrecht Wirthmann[(*)], Gerdy Seynaeve[(+)], Michail Skaliotis [(*)]

[(*)] EUROSTAT Task Force on Big Data, 5, rue Alphonse Weicker, L 2721 Luxembourg

[(+)] Proximus, Boulevard du Roi Albert II, 27 - B-1030 Brussels.

Abstract. This paper discusses various aspects related to the potential partnership between Statistical Offices (SO) and Mobile Network Operators (MNO) to leverage MNO data for the computation of official statistics. MNO data are complementary to other data sources that are already available to SOs (e.g., survey data, administrative registers) and their combination can lead to a new generation of statistical products, delivered more timely and with better spatio-temporal resolution than traditional statistics. This enables statisticians to gain more accurate and up-to-date insight into various aspects of human mobility and related socio-economic phenomena (e.g., tourism flows, presence and residence, commuting patterns, use of transportation means among others) with clear advantages for the process of policy design and evaluation based on such statistics. The cooperation between SO and MNO can be designed to prevent potential conflicts between the public and private interests, e.g. by the provision of adequate protection for business confidentiality, methodological quality and process transparency. We argue that partnering with SO brings direct and indirect benefits also to the MNOs, particularly in terms of empowering the portfolio of commercial analytic products they can offer to business customers. Synergies between the production of official statistics and commercial analytic products can be positively leveraged within the framework of a well-designed partnership model. By doing so, the SO-MNO partnership does not represent as a risk to the MNO business nor a diminution of the role and independency of SO, but rather as an additional opportunity for both sides. While the focus of this paper is on partnership models between SOs and MNOs, many elements of the discussion apply as well to private data holders from other sectors, and may contribute to advance the future vision of public-private partnerships for joint data analytics[1].

1. A new context: the datafied world

A combination of technological developments introduced during the last two decades (online platforms, Smart Devices, Internet-of-Things to name just a few) has led to pervasive "*datafication*" of almost any aspect of our lives [1]. The terms "data revolution" and "data economy" have been popularized to represent the impact on business and society caused by the availability of "data" about everything, everybody and everywhere. New business models are emerging based on the information (and value) that can be extracted from such new data. The telecommunications sector is no exception to this trend: an increasing number of Mobile Network Operators (MNO) are extending their business offerings towards **data analytics services and products** built upon the continuous flow of data about location and activity status of the mobile devices – hence of mobile users – served by their networks. Such data, produced through the network infrastructure operations, include Call Detail Records and more sophisticated signalling data [2]. Several MNOs are developing new organizational branches, laboratories or departments, dedicated to develop commercial analytic products

1 The views expressed in this paper are those of the authors and do not necessarily reflect the official views of the European Commission. Any potential errors, omissions and inconsistencies are the sole responsibility of the authors.

based on such data. Examples of business cases in this field include e.g. support of urban planning, road traffic operation [3] and tourism flow analysis [4] among others.

In the new *datafied* world, the data sources traditionally accessed by SOs, namely survey data and administrative data, constitute a small fraction of the globally available data stock. Large parts of such "new data" are held by private companies (e.g. MNOs) and constitute valuable business assets, but at the same time they embed information that is relevant for public interest, and specifically for official statistics. In this new context, SOs seek ways to leverage new data held by the private sector to enrich and complement its statistical products [5]. As far as MNO data are concerned, they can be leveraged by SO to produce statistics that better capture phenomena related to human presence and mobility, and do so more timely (possibly quasi real-time), with higher population coverage and better spatio-temporal resolution than is possible with traditional survey methods. Implementing this vision requires clarifying the following key questions:

- What information is to be elicited by SO from MNO data? And how?
- What are the motivations for MNO to cooperate with the SO?

Both items are addressed in this contribution. Before discussing the MNO motivations to engage into partnerships with SO, it is convenient to first delve into the technical and methodological aspects related to the object of such potential partnership.

2. Sharing computation, not data

We start considering a simple reference scenario where one MNO holds data that are relevant (possibly in conjunction with other data sources) for some new statistical product elaborated by the SO. Generally speaking, only *part of the information* embedded in the MNO *data* is of interest for SO. We are using the term "information of interest" in a rather abstract way here, and we remark the distinction from "data". To illustrate by a simple example, if the final statistical product is set to report the (estimated) number of mobile users that travelled between two cities within a specific temporal interval, then the information of interest is represented exclusively by the value of this aggregate indicator, and does *not* include e.g. the collection of all individual mobile user trajectories. In other words, strictly speaking the *SO is not interested in the MNO data as such, but only in particular pieces of information* embedded therein. We call ``computation" the process of extracting the desired (output) information from the input data.

In the simplest scenario, one party (e.g., the SO) is interested in the output of the computation while another party (the MNO) holds the input data. The traditional approach to cope with this scenario is to move the entire input dataset from MNO to SO, and then execute the computation entirely within the SO premises. This is a simple and intuitive approach, but not the only one. On the opposite extreme, the computation might take place entirely on the side of the MNO, and then only the desired output is passed to the SO. In between these two extremes there are intermediate solutions where the computation process is split into distinct parts that are executed by the two parties, and intermediate data are passed from MNO to SO.

In a more complex scenario the desired output information lies in the fusion of data held by different MNOs, operating in the same country or in different ones depending on the particular use-case. In certain cases, the global computation process can be factorized into *parallel components* that are run independently by each MNO, with their respective outputs being combined at the SO. A possible example includes the computation of geographical density maps of mobile users observed by the MNO infrastructure at a particular reference time, or the origin-destination matrix of tourism flows between different cities: in such cases the individual views computed by different MNO can be superimposed to obtain a more complete view of the whole population[2]. In other cases, the computation can be factorized into *sequential components* that are run independently by different MNOs, each component taking in input the output of another component. A possible example would be e.g. the process of training a neural network over a superset of records held by different MNOs. Finally, in the more sophisticated scenario the computation process cannot be factorized and require the *joint processing of input data* across multiple MNOs. This case is encountered for example when the regression variables for the same data subjects are held by different MNOs, or when regression must be run on the intersection elements between different MNO datasets. Also in this case technological solutions exist that allow eliciting only the desired output information, without requiring the disclosure of input data across different parties. One prominent solution to this class of problems is provided by Secure Multi-Party Computation (SMPC) methods. In a nutshell, SMPC allows to process confidential input data across administrative domains (e.g., MNOs and SOs) without disclosing the input data nor leaking any related information other than the desired output. We refer the reader to the specialised literature for a more rigorous explanation of SMPC methods (see e.g.[6]). What is important

[2] The superposition of aggregate estimates obtained from different MNO suffers from the problem of double counting users with multiple subscriptions. Similarly, double-counting errors may occur for international roaming users attaching to different networks in the visited countries. Such errors might be prevented, in principle, by adopting more sophisticated strategies for the fusion of input micro-data, rather than simple superposition of output aggregates.

to remark here is that such technology has considerably matured in the last decade, making his way out of academic laboratories into commercial products [7]. Pilot projects and early deployments based on SMPC technology are being carried out [8]. Besides the technical aspects, such pioneering activities are contributing to clarify the legal aspects around the use of this technology for the processing of personal data and GDPR compliancy.

Let us assume that a trusted SMPC platform is put in place in order to allow the SO eliciting information from input data held by one or multiple MNOs. In this scenario, SO is playing the role of an output party, while each MNO plays the role of an input party. Now, the same SMPC platform can be used to fuse MNO data with confidential input data held by SO: in this way, the SO can play *also* the role of input party. Furthermore, the same SMPC infrastructure can be used in principle to deliver specific computation output to the participating MNOs, letting them play *also* the role of output party. In other words, the SMPC platform allows the joint processing of confidential data across MNO and SOs (and possibly other organizations) and may support configurations where each party *gives and takes information* (not data!) to/from other parties. In this sense, SMPC may be seen as a technological enabler for partnership models between public institutions (e.g., SO) and private companies (e.g., MNOs) based on the *exchange of (non-personal) information,* in full respect of input data confidentiality and privacy.

In summary, technological solutions are available today allowing one institution to elicit (compute) the desired and agreed-upon output information from the input data held by one or more other institutions with no need to disclose the input data. Such technologies represent a major paradigm shift from “sharing data” to “sharing computation”. A change of terminology is needed to reflect the fact that input data can be “used” across different organizations and administrative domains without being “shared”.

3. Sharing control over processing methods.

Regardless of where the computation is physically *executed*, all involved parties can exercise full (non-exclusive) control over the *design* of the computation procedure. This is to say that the processing algorithm – from the very first stages of data cleaning, imputation etc. – can be developed in cooperation, or at least agreed upon, by experts of both MNO and SO. Technical means (e.g. Trusted Execution Environment or Certified Hardware) can be adopted to guarantee that what is executed (binary code) corresponds to what was agreed upon (source code). In this way, each party can directly safeguard the legitimate interests connected to its mission: the SO will ensure that the overall processing workflow meets the

target level of methodological quality and soundness (including a proper handling of errors and bias in the input data) while MNOs can ensure that the (final or intermediate) computation results delivered to the SO do not jeopardize their business. Additionally, if personal data are involved along the process, both parties are jointly responsible of ensuring compliance towards the applicable privacy regulations, *in primis* GDPR[3].

We remark the importance of setting a fully transparent processing workflow where all parties, SO and MNOs, have full visibility of how the data are handled and transformed at each processing stage, regardless of where it is physically executed. Each party should maintain the right to inspect the source code of all processing components[4]. This is in stark contrast with the alternative "black box" approach, where MNO passes to SO the pre-processed intermediate data, or even final statistics, merely with a high-level synthetic description of the implemented processing methods[5].

The vision outlined above implies that domain experts from different fields and with different background knowledge, i.e., statisticians from SO and telecom engineers from MNO, co-operate towards the definition of a processing workflow that is understood and accepted by both sides. Furthermore, extending our view from a single MNO-SO cooperation towards a larger "ecosystem" involving multiple MNOs, multiple SOs (at national or supranational level) and possibly other institutions (e.g., universities, researchers, private companies), it is clearly desirable to develop processing components that can be reused across different organizations. In order to achieve these goals, EUROSTAT is cooperating with Proximus (the incumbent MNO in Belgium) and other members of the European Statistical System (ESS) to develop a Reference Methodological Framework (RMF) for the processing of MNO data for official statistics [9].

[3] In order to strengthen compliance to legal and ethical regulations, an independent authority or accredited agency might be called to inspect and certify the processing algorithm (source code) before it is executed on actual data. The certification authority would then logically share control over the processing design with the other players, namely SO and MNOs. The benefits of this option are twofold. First, it would remove any residual legal risks from MNO and SO. Second, *ex ante* compliancy validation of the whole workflow would strengthen the protection of citizen privacy and overall transparency.

[4] Source code inspection can be arranged under strict non-disclosure agreements Alternatively, if closed-source components cannot be avoided, a process for software qualification can be put in place based on extensive testing with benchmark reference data (possibly synthetic).

[5] With such "black box" model it remains unclear how the SO would be able to ensure the appropriate level of methodological quality and soundness, that is a fundamental mandate of its institutional mission.

The design of the RMF is inspired by the principles of "layering" and "hourglass model" that lie at the foundation of the Internet [10][11]. A key component of RMF is the conceptualization of a unified "convergence layer" (C-layer) between the lower "data layer" (D-layer) and the upper "statistics layer" (S-layer), as depicted in Fig. 1. The role of the C-layer is analogous to that of the Internet Protocol in the Internet: by providing a *logical interface* between the upper and lower layers it effectively decouples them, enabling their independent development and evolution. A parsimonious number of common data structures are defined in the C-layer whose semantic should be (i) simple enough to be understood by experts from both knowledge domains; and (ii) general enough to capture the information relevant for SO across different kinds of MNO data (CDR and signalling data), different technologies (2G, 3G, 4G and in the future also 5G) and different network-specific configurations. With the C-layer in place, statisticians can then focus on the development of statistical methodologies in the S-layer (including inference, aggregation, record selection etc.) taking in input C-layer structures, while telecom engineers can focus on the transformation of their raw network data into C-layer structures[6]. Thanks to the common format of C-layer structures, algorithms and processing components that are developed at the S-layer by one institution (e.g., a national statistical institute or university research tram) can be adopted by other institutions and/or run on data from other MNOs. In this way, the RMF will contribute to grow an "ecosystem" for the development, exchange, validation and adoption of software components for processing MNO data across different organizations.

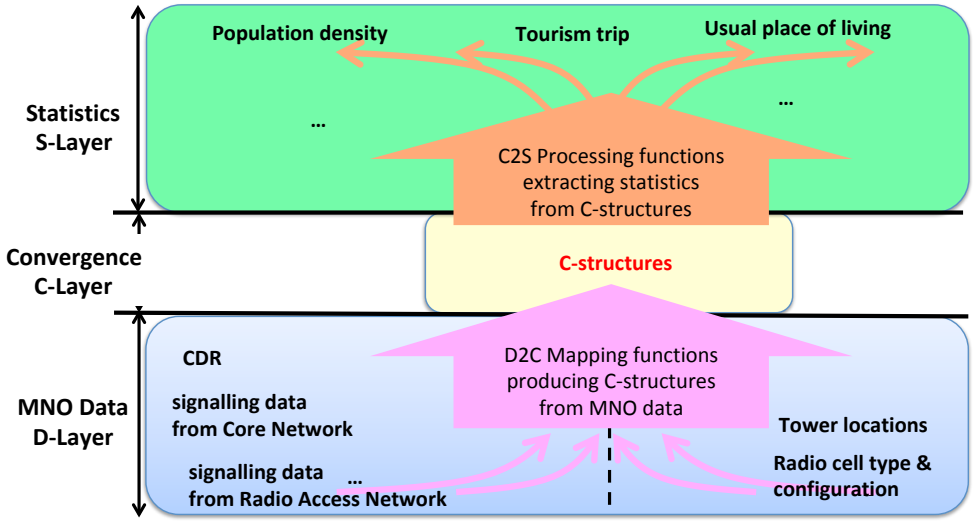

Figure 1 – A layered view of the Reference Methodological Framework for the processing of MNO data under development by EUROSTAT

[6] Note that C-layer structures still represent personal data, and therefore should be processed within the MNO domain. Through the processing workflow logically placed in the S-layer, aggregate non-personal data are generated that can be passed to the SO for subsequent processing.

4. Scope and benefits of MNO-SO partnerships

The RMF and SMPC introduced in the previous sections are *technological components* that, together with other components, form the possible architecture of a partnership model on the technical side. In this section we shift our focus from the *What* to the *Why*, and move to discuss the potential motivations for establishing MNO-SO partnerships.

4.1 Clearly identifying the scope of the output statistics

Before entering the discussion it is convenient to make a distinction between information (or analytic products) of "public interest" and of "commercial interest" with regard to the information that can be potentially elicited by MNO source data. By analytic products of "commercial interest" we refer to information (processed data) that is valuable for, hence can be purchased by, business customers on a commercial basis. Among many illustrative examples we can mention e.g. the daily inflow of customers towards a particular shopping mall area, or the peak-hour intensity of commuting flows between two city districts. Potential customers for such analytic products would be, respectively, companies offering business intelligence services and urban planners. The latter type might be interesting as well for the public office in charge of public transport planning, threfore it represents an example at the intersection between "commercial interest" and "public interest" (ref. "B" in Fig. 2).

The information of interest for official statistics is contained within the "public interest" category. One element of differentiation versus other categories is often the coarser level of temporal and/or spatial aggregation (ref. Fig. 3): to stay with the above examples, official statistics would be concerned with the total amount of customer inflow to all shopping areas throughout a larger region (e.g., an entire province) over a longer period, and to aggregate indicators of commuting intensity for the whole city.

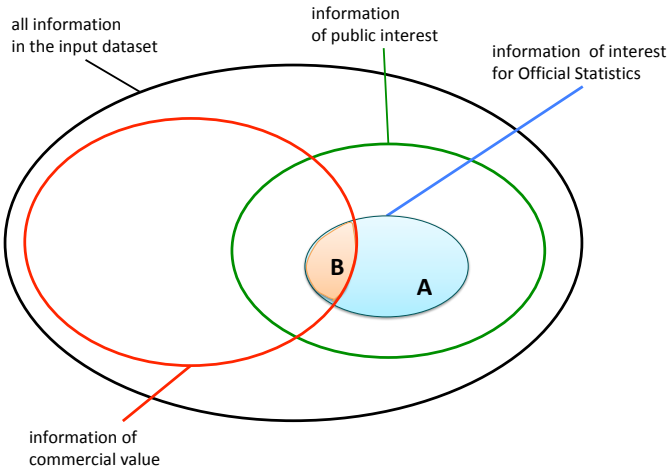

Figure 2

The dialogue between MNO and SO should start from the identification of analytical products that have a clear interest for official statistics, not quite so for commercial customers (ref. "A" in Fig. 2). These products are in the focus of the following discussion. In many cases, it is not difficult to draw a reasonable separation line from commercial analytics products, also in terms of different spatial and/or temporal aggregation levels (ref. Fig. 3). Through such open dialogue, the SO and MNO can reduce the intersection area between the two categories ("B" in Fig. 2). There might still be analytical products that have both a "statistics interest" and "commercial interest" that may require special agreements. However, addressing these cases can be seen as a future extension and evolution of the MNO-SO partnership model that is addressed in this contribution.

4.2 Benefits for MNO

In one extreme scenario, MNOs might be obliged by law to *provide access* to their data to SOs. A public consultation was recently carried out on this topic[7] and some national legislation initiatives already go in this direction[8]. At the other extreme, the SO might merely purchase *access to data* from the MNO (possibly through SMPC platform) in a plain customer vs. provider business relationship.

Both scenarios can be legitimately taken into consideration, but we do not elaborate them here. The goal of this contribution is to offer additional input to the public debate by elaborating on alternative forms of MNO-SO cooperation, namely *partnerships*. A necessary pre-condition for the deployment of a partnership relation is the "win-win" factor, i.e., both parties must recognize a mutual benefit. From the side of SO the benefits are clear: the perspective of delivering more advanced, timely and accurate statistical products based on MNO data help the SO to better fulfil its statutory mission and strengthen its role. Hereafter we elaborate on the potential "partnership gain" on the side of MNOs.

Let us consider the case of a generic MNO that is developing a line of business analytics services or products, based on its own network data, to be offered to potential customers on a commercial basis. In parallel, this MNO is cooperating with the SO towards the development of a new line of official statistics, to be officially published by SO. The partnership with SO brings direct and indirect benefits to the business analytic line of the partner MNO, as summarised in the following.

[7] https://ec.europa.eu/digital-single-market/en/news/summary-report-public-consultation-building-european-data-economy

[8] https://www.legifrance.gouv.fr/eli/loi/2016/10/7/2016-1321/jo/texte

- **Access to additional information**. If a SMPC platform is put in place that enables the extraction of aggregate information without requiring sharing of input data, the SO may make its internal data available to the partnering MNO. In fact, SO holds a wide range of legacy datasets, including micro-data about individuals, companies, places etc. that are gathered to accomplish the SO statutory mission. Such data include variables and dimensions that are complementary to MNO data. SO micro-data are treated confidentially and the SO has the legal responsibility to protect them. However, with appropriate privacy-preserving computation technologies (e.g., SMPC), SO and MNO data could be fused together for the computation of aggregate (non-personal) analytic products that could not be obtained from either source individually. The SO and MNO might negotiate to *use jointly* (not share!) their input data to produce a combination of statistical products for public interest as well as for commercial interest: the latter would enrich the portfolio of business analytic products offered by the MNO to its customers.
- **Calibration**. In a variant of the previous scenario, the SO data are not used to add additional variables to micro-records, but rather to correctly calibrate the aggregate values computed from MNO data with "ground truth" data obtained from SO data sources (e.g., the official number of residents in a given area).
- **Statistical Knowledge**. The partnering MNO can "pull" methodological knowledge from cooperating with SO experts, particularly for what concerns the correct handling of data errors, bias etc. Such knowledge is often complementary to the skills already present in the MNO tram (e.g., data scientists, telecom engineers).
- **Business reputation**. This is tightly connected to the previous point. As discussed earlier, SOs have developed a formal quality framework that is regularly applied to their statistical production process, and a solid reputation in terms of quality and reliability of the delivered statistical products. These are "soft assets" developed by SOs through several decades. By collaborating with SO on the definition of sound processing methodologies, the partner MNO must adhere to such best practices and therefore "inherits" part of SO reputation in terms of methodological quality and reliability. Likewise other "soft assets", such reputation can be leveraged to further promote its business analytics product line.
- **Public reputation**. The mission of SOs is to deliver veracious statistics serving the general public interest. Besides fuelling the process of policy design, policy evaluation and academic research in various fields, it informs the view that society has of itself. The quote "*Knowledge is power; [official] statistics is democracy*" summarizes effectively the

fundamental role of SO in modern democratic systems [12]. Although the mission of SO was shaped in the previous century where data were scarce, the fundamental role of SO to provide *veracious statistics* is not less important and compelling now that data are abundant[9]. By partnering with SO, the MNO would directly contribute to its mission and share the same positive ethical values. With proper branding and communication, the MNO might "inherit" part of the SO ethical reputation and leverage such "soft asset" for the promotion of its own brand.

- **Stimulating the market for "premium" commercial analytics.** The so-called "*freemium*" approach has proved successful in certain business sectors, where the availability of basic service or content for-free has the effect of *expanding, rather than compressing* the market for "premium" for-fee services or content. Along a similar reasoning, the public availability of coarsely aggregated official statistics, produced through standardized and statistically approved methods in partnership between MNO and SO, will potentially increase the appetite by potential customers for more detailed and fine-grain commercial analytics for business purposes, possibly delivered in real-time, offered by the MNO based on the same data sources and methodologies (ref. Fig. 3).

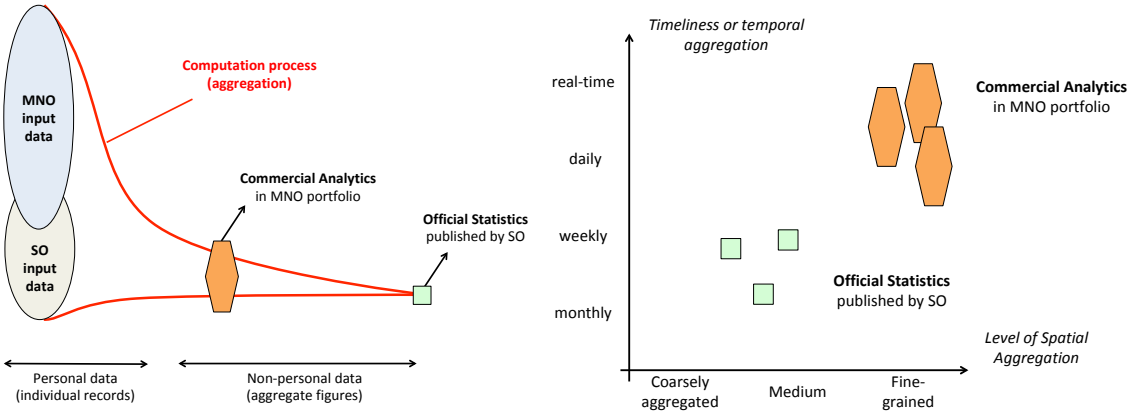

Figure 3 –Left: intermediate data generated durig the computation of Official Statistics enrich the portfolio of Commercial Analytical products. Right: analogy with "freemium" service model.

4.3 Partnership risks

In the previous section we have highlighted some potential benefits associated to MNO-SO partnerships. Here we elaborate on the potential risks and ways to mitigate it.

The first obvious risk is reputational: the MNO-SO partnership might be misinterpreted as a way to tighten government control on individuals through data that are considerably more

[9] We can make an analogy between "data" as "water", and "veracious statistics" as a source of "drinkable water": it has a clear vital role (for life, for democracy) in a dry desert where water is scarce (the pre-datafied world of XX century) as well as in a swamp or after a flood (the current datafied world in the post-truth era), where water is everywhere around but not necessarily drinkable (veracious).

pervasive and informative about individual behaviour than traditional sources. Two obvious means to mitigate this risk are *increased transparency* and *better communication*. The positive mission of SO, including the legally established independence from other governmental institutions, must be clearly communicated to the general public. Also, detailed and transparent information should be given about what parts of the MNO data are used, in which way and for which purpose. Ideally the processing components (algorithms) involved in the computation of final statistics would be made publicly available as open-source in order to allow public scrutiny by independent experts. At the end of the day, an increased level of process transparency is due in order to balance the increased level of pervasiveness of such new data sources.

The other risk is related to business. The statistical products developed by the SO based (also) on MNO data will be made publicly available for free, and this could be perceived as a cannibalization of the potential market for commercial analytic services based on the same MNO data. This risk can be mitigated by careful design of the statistic products for SO, with particular attention to the degree of spatial and temporal aggregation, level of details and timeliness of the publication. As indicated in the previous section with reference to the "*freemium*" model, the coexistence with free statistical products might promote, rather than jeopardize, the potential market for "premium" commercial analytics.

Finally, it is important to remark that a successful partnership between SO and a particular MNO represents a competitive advantage for the latter, and might raise question of fairness by other competing MNOs. To preserve impartiality, the opportunity to engage into the same partnership model should be offered to all other MNOs. In other words, no exclusivity clause should be in place in MNO-SO partnership agreements. On the contrary, successful initial partnerships between SO and forerunner MNOs might represent a stimulus for other MNO (in the same country or in others) to join similar partnerships, in a positive virtuous cycle that would ultimately lead to better coverage of new statistic products.

5. Conclusions and outlook

In this paper we have outlined the main components of potential partnership models between SO and MNOs. On the technical side, key components of the proposed view are (i) the adoption of privacy-preserving computation models to transfer the desired information across organizational domains without sharing the confidential input data; and (ii) the adoption of a Reference Methodological Framework to enable the joint design of processing methods by SO and MNO experts. Building upon such components, MNO and SO can define partnership

models to mutually empower their portfolios of official statistical products and commercial analytics services, respectively.

EUROSTAT and Proximus are actively cooperating to work out a general model of MNO-SO partnership, in close coordination with the ESS Task Force on Big Data and the ESS network on Big Data II, and the present contribution is part of this collective effort. Our near-future perspective is to implement a concrete proof-of-concept instance of MNO-SO partnership by end of 2019 with focus on a small set of clearly identified statistical products. Furthermore, we aim to extend the discussion also outside the ESS, involving the relevant industry forums (e.g., GSMA[10] and ETIS[11]) as well as the UN Global Working Group on Big Data for official statistics. Our mid-term goal is to launch an ESS-driven pilot project involving multiple MNOs by end of 2020.

REFERENCES

[1] K. Cukier and V. Mayer-Schoenberger. The rise of big data. Foreign Affairs, May/June 2013. https://www.foreignaffairs.com/articles/2013-04-03/rise-big-data

[2] F Ricciato, P Widhalm, F Pantisano, M Craglia. Beyond the "single-operator, CDR-only" paradigm: An interoperable framework for mobile phone network data analyses and population density estimation. Pervasive and Mobile Computing, vol. 35, February 2017.

[3] A. Janecek et al., The Cellular Network as a Sensor: From Mobile Phone Data to Real-Time Road Traffic Monitoring, IEEE Transactions on Intelligent Transportation Systems 16(5), 2015.

[4] Rein Ahas et al. Feasibility Study on the Use of Mobile Positioning Data for Tourism Statistics, Consolidated Report, June 2014. https://ec.europa.eu/eurostat/documents/747990/6225717/MP-Consolidated-report.pdf

[5] F. De Meersman et al. Assessing the quality of mobile phone data as a source of statistics. European Conference on Quality in Official Statistics, June 2016

[6] R. Cramer et al., Secure Multiparty Computation and Secret Sharing, Cambridge University Press, 2015.

[7] D.W. Archer, D. Bogdanov, B. Pinkas, P. Pullonen. Maturity and Performance of Programmable Secure Computation. In IEEE Security & Privacy, 14(5), September 2016. doi: 10.1109/MSP.2016.97

[8] D. Bogdanov et al. Students and Taxes: a Privacy-Preserving Social Study Using Secure Computation. In Proceedings on Privacy Enhancing Technologies, PoPETs, 2016 (3), 2016.

[9] Towards a Reference Methodological Framework for the processing of mobile network operator data for official statistics, presentation at Mobile Tartu 2018. https://ec.europa.eu/eurostat/cros/system/files/rmf_mobiletartu2018_ricciato_printout.pdf

[10] S. Akhshabi and C Dovrolis. The evolution of layered protocol stacks leads to an hourglass-shaped architecture. ACM SIGCOMM 2011.

[11] M. Chiang, S. Low, A. Calderbank, and J. Doyle. Layering as optimization decomposition. Proceedings of the IEEE, 2007.

[12] E. Letouzé, J. Jütting. Official Statistics, Big Data, and Human Development. White Paper. March 2015. http://datapopalliance.org/wp-content/uploads/2015/11/Big-Data-Official-Statistics-FINAL.pdf

[10] https://www.gsma.com
[11] https://www.etis.org