



Reusing Mobile Network Operator (MNO) data for Official Statistics: a European perspective

Fabio Ricciato

Eurostat, Unit A.5 Methodology; Innovation in official statistics

Visit to Statistics Sweden (SCB)

3rd September 2024, Örebro

Caveat: *The information and views set out in this presentation are those of the author and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.*

About myself

- Education in Electrical Engineering 1999, PhD in Telecommunications 2023
- 15 years of applied research and teaching, at university and RTOs
 - telecommunication systems, computer networks, mobile networks, signal processing, software-defined radio, radio localisation, ...
- First encounter with Mobile Network monitoring back in 2004
 - METAWIN project in collaboration with A1 Mobilkom in Austria
- Since 2018 with Eurostat as statistical officer
 - Focus on innovation in official statistics, Mobile Network Operator data, Privacy Enhancing Technologies

RTO: Research and Technology Organisation

originalarbeiten

Elektrotechnik & Informationstechnik (2006) 123/7/8: 288–296. DOI 10.1007/s00502-006-0362-y

Traffic monitoring and analysis in 3G networks: lessons learned from the METAWIN project

F. Ricciato, P. Svoboda, J. Motz, W. Fleischer, M. Sedlak, M. Karner, R. Pliz, P. Romirer-Maierhofer, E. Hasenleithner, W. Jäger, P. Krüger, F. Vacirca, M. Rupp

A 3G network is a magnificently complex object embedded in a highly heterogeneous and ever-changing usage environment. It combines the functional complexity of the wireless cellular paradigm with the protocol dynamics of TCP/IP networks. Understanding such an environment is more urgent and at the same time more difficult than for legacy 2G networks. Continuous traffic monitoring by means of an advanced system, coupled with routine expert-driven traffic analysis, provides an in-depth understanding of the status and performances of the network as well as of the statistical behaviour of the user population. Such knowledge allows for a better engineering and operation practice of the whole network, and specifically the early detection of hidden risks and emerging troubles: the exploitation of certain TCP/IP dynamic behaviour, particularly the TCP control-loop, coupled with information extra the 3GPP layers, provides a cost-effective means to monitor the status of the whole network without requiring access to elements. In this article the main lessons are summarized learned from a two-year research activity on traffic monitoring at top of an operational 3G network.

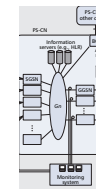
Keywords: traffic monitoring; traffic analysis; 3G; cellular networks; GPRS; UMTS

Analyse und Monitoring von Datenströmen in 3G-Netzwerken: ein Erfahrungsbericht aus dem METAWIN-Projekt

3G/4G/WLAN/WMAN PLANNING AND OPTIMIZATION

TRAFFIC MONITORING AND ANALYSIS FOR THE OPTIMIZATION OF A 3G NETWORK

FABIO RICCIATO, FORSCHUNGSZENTRUM TELEKOMMUNIKATION WIEN



The availability of high-quality traces can greatly empower the measurement-based optimization cycle, with human experts in the loop

ABSTRACT

Recent years have recorded a surge of research activities on IP traffic monitoring, enabled by the availability of monitoring hardware and large-scale storage at accessible costs. More recently, passive monitoring has been applied to operational 3G networks. The passive observation of network traffic, coupled with advanced traffic-analysis methods, can be a powerful and cost-effective means to infer the network status and localize points of performance degradation without requiring complete access to all network elements. Furthermore, the availability of high-quality traces can be exploited to predict the load of the network under hypothetical conditions, variations of the actual network configuration at the capturing time. Both approaches can be useful for some engineering and reoptimization tasks that are commonly encountered in the lifetime of an operational 3G network. In abstract terms, the availability of high-quality traces can greatly empower the measurement-based optimization cycle, with human experts in the loop, thus driving an already operational 3G network toward improved performances. In this article we discuss the contribution that traffic monitoring and analysis (TMA) can provide to the optimization of an operational 3G network.

Since 2004 we have been involved in a research project aimed at exploring traffic monitoring and analysis for the engineering of a GPRS/UMTS network [1]. The initial idea was to capture packet-level traces from the live network and use them to distill synthetic user models fitted to the observed data. The models were meant to feed network simulations and/or mathematical analysis in order to derive dimensioning guidelines for the real network. The conceptual trajectory is represented by the simulation path shown in Fig. 1. As a side value, it was expected to extract fine-grain indicators about the current performances and quality of the underlying network (e.g., packet delay, completion time of signaling procedures, etc.). During the investigation we soon recognized that the availability of high-quality traces (to be exactly defined below) yields a much higher potential for improving the engineering practice of the real network, far beyond the mere opportunity to fit abstract models to observed patterns. More generally, TMA can play a major role in several technical areas within the running of a real 3G network: operation and maintenance, troubleshooting, planning and optimization, design and engineering, security monitoring and fraud detection, and so forth. Its reach can be easily extended to less technical areas like for example marketing, service engineering, billing, and tariff design. However, in the spirit of the current Special Issue we focus here only on the aspects related to the network engineering and optimization.

INTRODUCTION

In this contribution we present a few exemplary applications of TMA to the operation and engineering of 3G network. The goal is to provide an overview of the recent and ongoing work, and to promote further research on TMA in 3G. The referenced works¹ come from a few research groups active in this area and primarily from the METAWIN project [1]. The rest of this article is organized as follows. We define the scope of TMA, and then we discuss the evolution from the classical approach for network monitoring toward large-scale TMA. We introduce the requirements for an advanced traffic monitoring system for 3G networks. We provide examples of the potential role of TMA in the optimization of an operational 3G network. Finally, we draw a summary of the conclusions.

The Cellular Network as a Sensor: From Mobile Phone Data to Real-Time Road Traffic Monitoring

Andreas Janecek, Danilo Valerio, Karin Anna Hummel, Fabio Ricciato, and Helmut Hlavacs

Abstract—Mobile cellular networks can serve as ubiquitous sensors for physical mobility. We propose a method to infer vehicle travel times on highways and to detect road congestion in real-time, based solely on anonymized signaling data collected from a mobile cellular network. Most previous studies have considered data generated from mobile devices active in calls, namely Call Detail Records (CDR), an approach that limits the number of observable devices to a small fraction of the whole population. Our approach overcomes this drawback by exploiting the whole set of signaling events generated by both idle and active devices. While idle devices contribute with a large volume of spatially coarse-grained mobility data, active devices provide finer-grained spatial accuracy for a limited subset of devices. The combined use of data from idle and active devices improves congestion detection performance in terms of coverage, accuracy, and timeliness. We apply our method to real mobile signaling data obtained from an operational network during a one-month period on a sample highway segment in the proximity of a European city, and present an extensive validation study based on ground-truth obtained from a rich set of reference data sources—road sensor data, toll data, taxi floating car data, and radio broadcast messages.

Index Terms—Cellular floating car data, large mobility data sets, travel time estimation, road congestion detection, mobility

representativeness of probes, e.g., when using GPS traces from a taxi fleet or public transport vehicles.

We propose an alternative approach based on the observation of the signaling traffic of a mobile cellular network. Any mobile terminal—including personal phones and tablets, but also navigation devices and on-board units (OBUs)—attached to the cellular network produces signaling messages that can be captured passively on the network side, anonymized, and then processed to derive mobility patterns. We use these messages to infer traffic status and congestion episodes on highways in real-time. Instead of a costly deployment of new sensors, we exploit the legacy cellular network as a large-scale real-time mobility sensor. The traffic information extracted with our approach can serve as a powerful input for ITS applications.

The idea to extract road traffic information from cellular network data has been considered in several other studies. However, in the vast majority of previous work, traffic status reports leverage data only from “active” devices, i.e., devices engaged in a voice call or data connection, based on call details records (CDR) [8]–[11]. Active devices can be tracked at cell level, i.e.,

IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 16, NO. 5, OCTOBER 2015

2551

1536-1284/06/\$20.00 © 2006 IEEE

IEEE Wireless Communications • December 2006



European Commission

Eurostat & the European Statistical System



- The European Statistical System (ESS) is the **partnership** between
 - Eurostat + National Statistical Institutes (NSIs) + Other National Authorities (ONAs) in each EU country
- Eurostat is ...
 - the statistical office of the EU
 - a DG of the European Commission
 - the coordinator of the ESS
- Eurostat does ...
 - produces European statistics
 - contributes to harmonise methodologies, definitions, criteria, etc. within the ESS

ESS mission and principles

*[...] our mission is to provide independent **high quality** statistical information [...].*

*Our principles of quality are: **relevance**, accuracy, **timeliness** and punctuality, accessibility and clarity, as well as comparability and coherence^(*).*

(*) Quality Declaration of the European Statistical System September 2016,
<https://ec.europa.eu/eurostat/documents/4031688/8188985/KS-02-17-428-EN-N.pdf/116f7c85-cd3e-4bff-b695-4a8e71385fd4>

If we were running a long-distance “express” parcel delivery service ...

Mission

*... deliver reliably and **quickly**...*

**Expectation
by users**

weeks

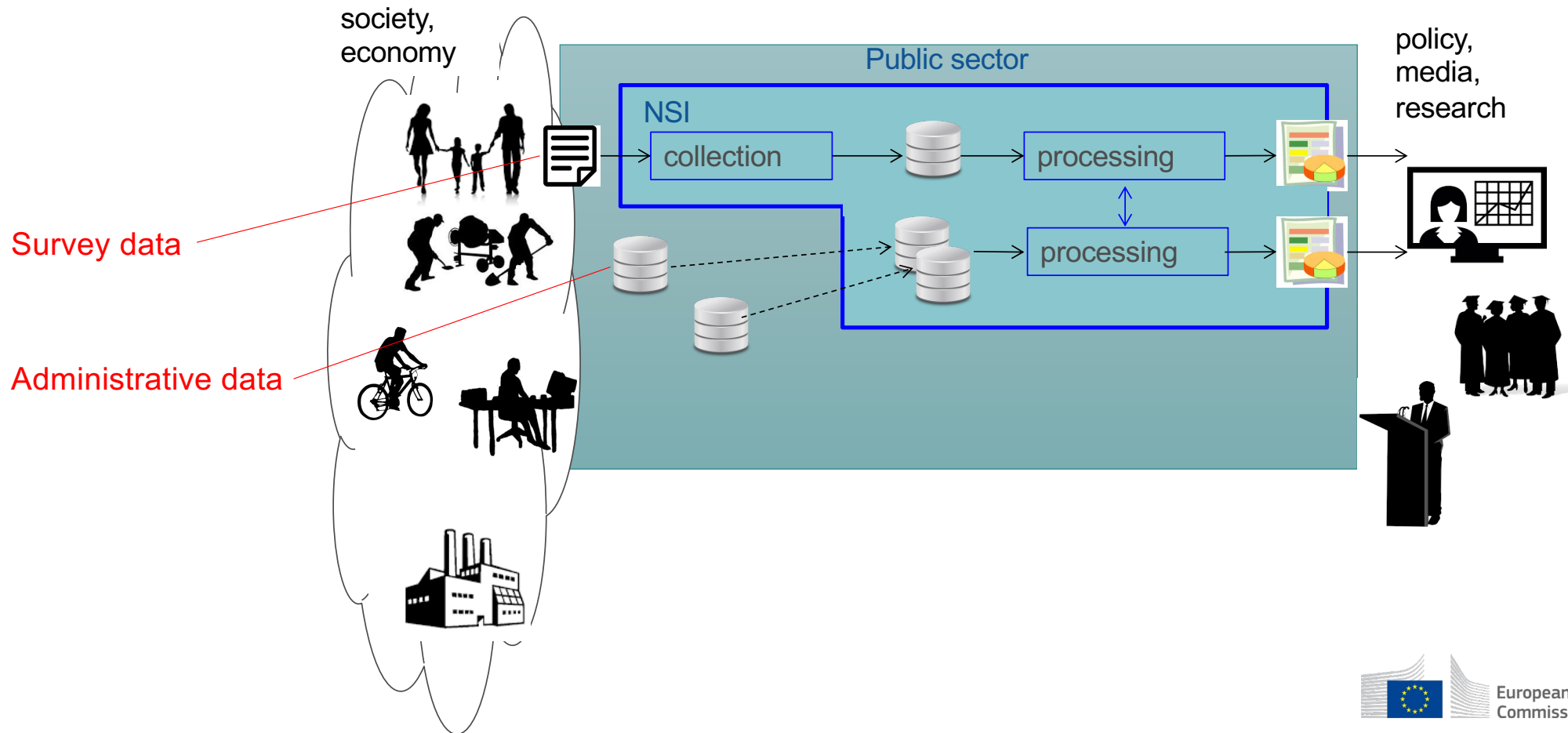
days

hours

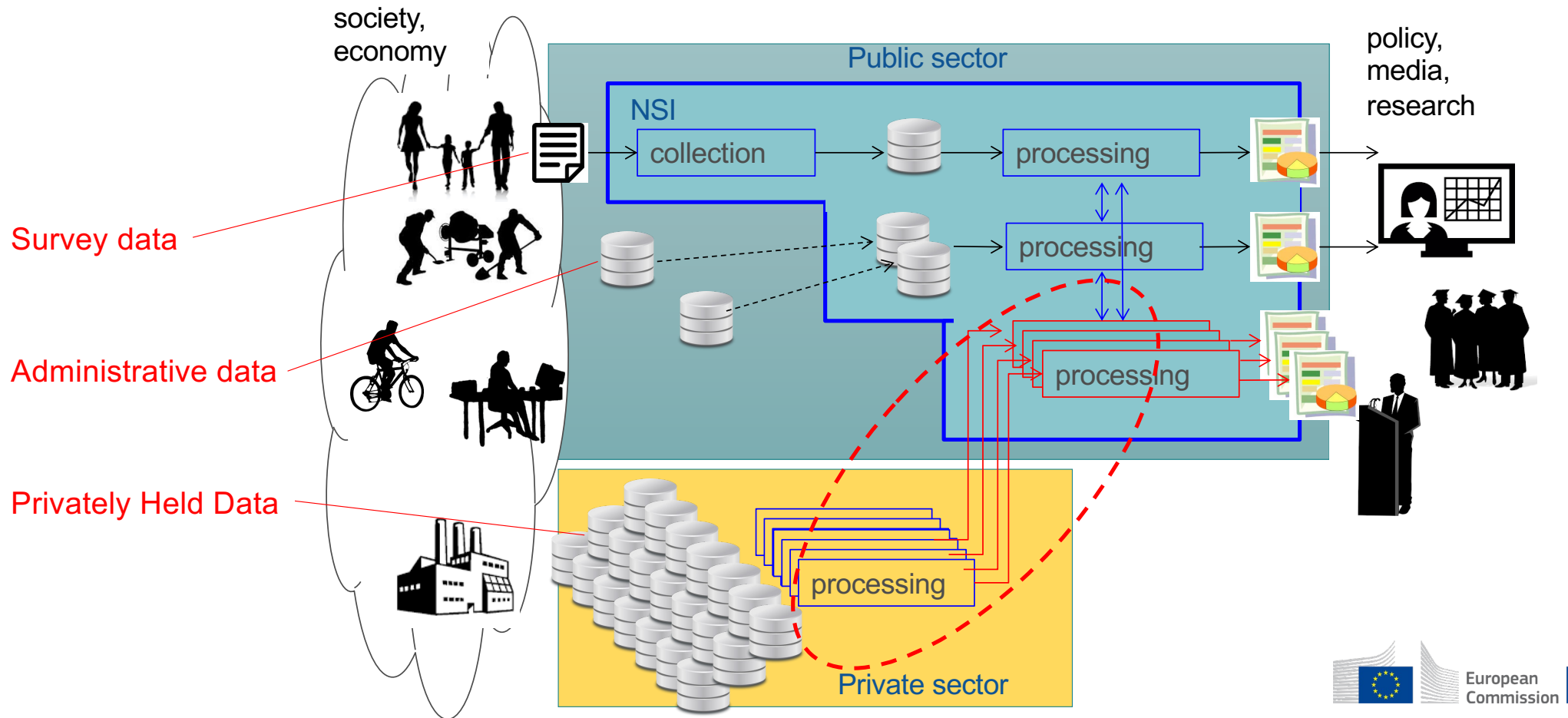
**Operation
model**



Official Statistics based on traditional data



Official Statistics “augmented” by privately held data





Over-skepticism

Over-enthusiasm

Official Statistics

“augmented” by privately held data

- Official statistics are based on “traditional” data sources
 - Survey data – collected by NSI primarily for statistical purposes
 - Administrative data – collected by public administrations, (re)used for statistical purposes
- Future perspective: (re)using *also* Privately Held Data (PHD)
 - MNO data as one of most interesting types of PHD
 - Some NSIs produced “experimental statistics” based on MNO data based on different arrangements with local MNOs, but no “official statistics” yet
 - NB: “official statistics” is much more than “experimental statistics”

Quality in Official Statistics (OS)

- OS influence opinions/decisions/actions by policymakers, citizens, businesses, researchers ...
- **Quality** as differentiator of Official Statistics vs. other statistics
 - OS ≠ commercial statistics, other public statistics
 - **OS ≠ “experimental statistics”**
 - One-off case study or short series vs. **no commitment** to continued regular production
 - Partial fulfillment vs. full compliance to Quality criteria
- **Official Statistics = Quality Statistics**
- Quality has a **cost!**



Institutional environment	
Principle 1 : Professional Independence	
Principle 1bis : Coordination and cooperation	
Principle 2 : Mandate for Data Collection and Access to Data	
Principle 3 : Adequacy of Resources	
Principle 4 : Commitment to Quality	
Principle 5 : Statistical Confidentiality	
Principle 6 : Impartiality and Objectivity	
Statistical processes	
Principle 7 : Sound Methodology	
Principle 8 : Appropriate Statistical Procedures	
Principle 9 : Non – excessive Burden on Respondents	
Principle 10 : Cost effectiveness	
Statistical output	
Principle 11 : Relevance	
Principle 12 : Accuracy and Reliability	
Principle 13 : Timeless and Punctuality	
Principle 14 : Coherence and Comparability	
Principle 15 : Accessibility	

Recent milestones

- **B2G4S Expert Group** on facilitating the use of new data sources for official statistics (2021-2022)
 - [Final report](#) with several recommendations, central concept of “fair and effective” **partnerships between NSI and data holders** 
- Revision of Regulation (EC) 223/2009 on European Statistics
 - Text approved by European Parliament in March 2024 ([link](#))
 - Introduces **legal enablers** for the sustainable reuse of privately held data for official statistics – Art. 17b, 17c
 - balanced mix of **obligations** and **safeguards** for data holders



European Parliament
2019-2024

STATISTICS REPORTS | eurostat

TEXTS ADOPTED

P9_TA(2024)0152
Amending Regulation (EC) No 223/2009 on European statistics
European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council amending Regulation (EC) No 223/2009 on European statistics (COM(2023)0402 – C9-0246/2023 – 2023/0237(COD))
(Ordinary legislative procedure: first reading)

The European Parliament,

- having regard to the Commission proposal to Parliament and the Council (COM(2023)0402),
- having regard to Article 294(2) and Article 338(1) of the Treaty on the Functioning of the European Union, pursuant to which the Commission submitted the proposal to Parliament (C9-0246/2023),
- having regard to Article 294(3) of the Treaty on the Functioning of the European Union,
- having regard to the opinion of the European Central Bank of 28 September 2023¹,
- having regard to the provisional agreement approved by the responsible committee under Rule 74(4) of the Rules of Procedure and the undertaking given by the Council

Composing a partnership is like composing a (big) puzzle



Business



Methodology

Data Protection

Legal

European Parliament
2019-2024



TEXTS ADOPTED

P9_TA(2024)0152

Amending Regulation (EC) No 223/2009 on European statistics

European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council amending Regulation (EC) No 223/2009 on European statistics (COM(2023)0402 – C9-0246/2023 – 2023/0237(COD))

(Ordinary legislative procedure: first reading)

The European Parliament,

- having regard to the Commission proposal to Parliament and the Council (COM(2023)0402),
- having regard to Article 294(2) and Article 338(1) of the Treaty on the Functioning of the European Union, pursuant to which the Commission submitted the proposal to Parliament (C9-0246/2023),
- having regard to Article 294(3) of the Treaty on the Functioning of the European Union,
- having regard to the opinion of the European Central Bank of 28 September 2023¹,
- having regard to the provisional agreement approved by the responsible committee under Rule 74(4) of the Rules of Procedure and the undertaking given by the Council

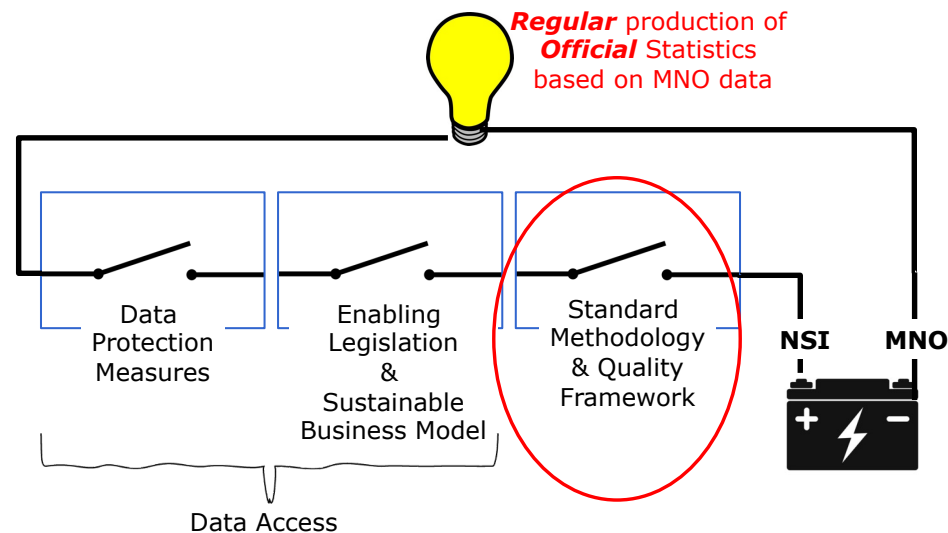


ESS Task Force on MNO data (TF-MNO)

- Established in 2021, representatives from 19 countries, focus on **methodological aspects**
 - sharing of experiences from national projects; networking of the relevant experts across ESS members
 - **identify key methodological directions to be prioritized** for further development
 - **proposing activities to further develop methods and tools** for the (re)use of MNO data in the ESS.

- Milestones

- **Position paper** ([link](#))
- **Multi-MNO** project
- Research grant (ESSnet) **MNO-MINDS**



TF-MNO position paper

Executive Summary

1. Introduction
2. What are MNO data and why they are relevant for Official Statistics
3. Methodological challenges
4. MNO data analytics in the private sector: current state of play
5. MNO data and official statistics: current state of play
6. The case for methodological standardisation
7. An open methodological standard for official statistics based on MNO data
8. The importance of taking a multi-MNO perspective
9. Fusion of MNO data with other data sources
10. Data access
11. The way forward

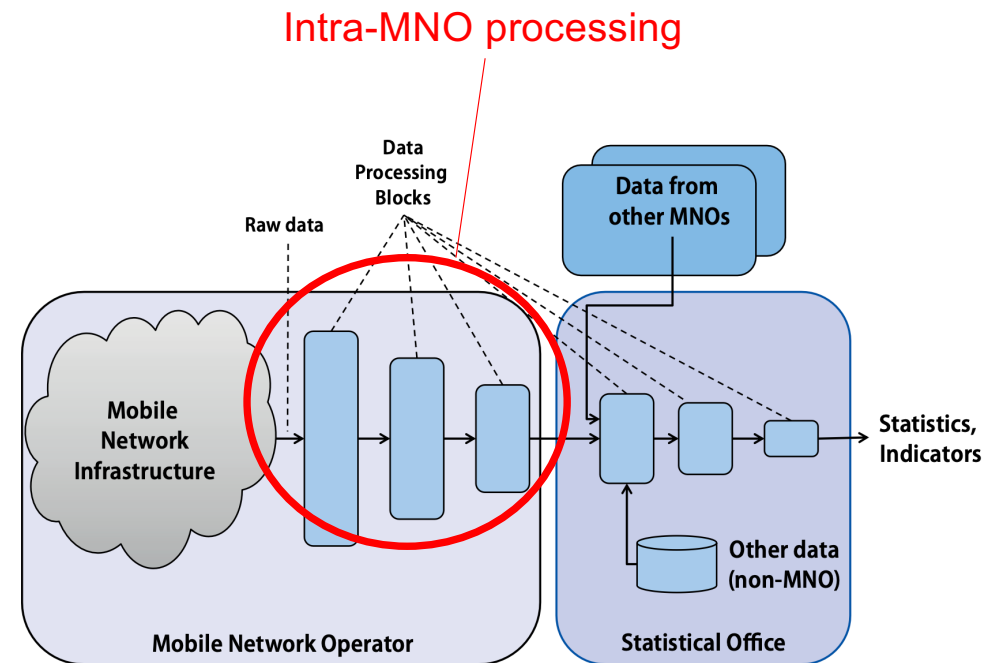
Focus of
MNO-MINDS
research grant

Focus of
Multi-MNO project



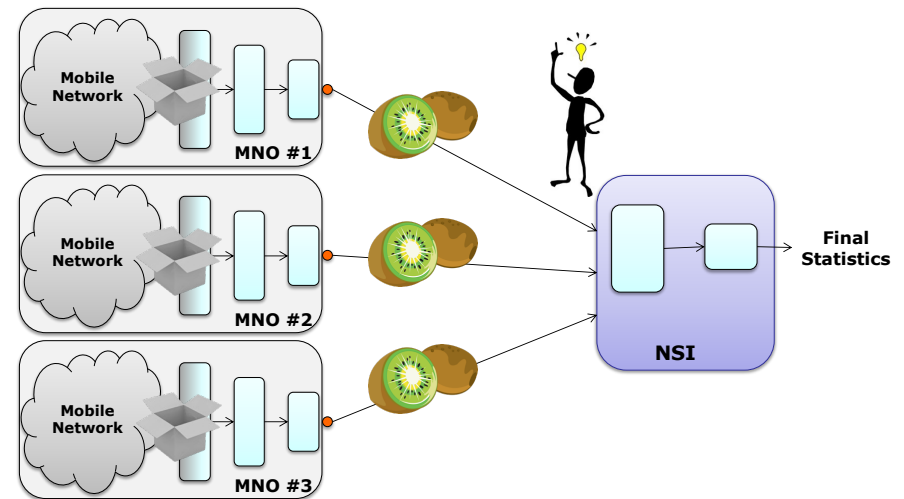
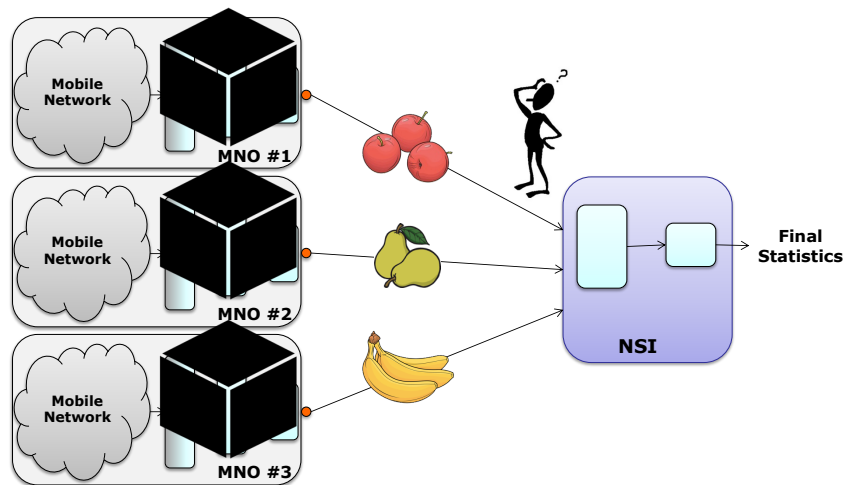
Methodological standardisation

- **Reference methodological pipeline:**
complete, detailed, non-ambiguous definition of whole end-to-end data processing workflow
- All operations and data structures described in formal language → *softwarisation* of statistical methodology
- **Reference open-source implementation & reference test data**
 - enable development and deployment of alternative, possibly proprietary implementations compliant to the open methodological standard
- Developed by the ESS for OS, could be reused by industry also for commercial statistics (more on that later)



Methodological standardisation

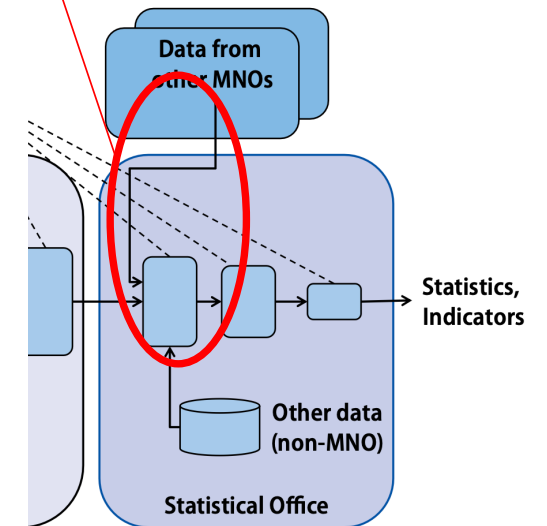
- Common/standard methodologies and definitions necessary to achieve **comparable** and **combinable** results across different MNOs



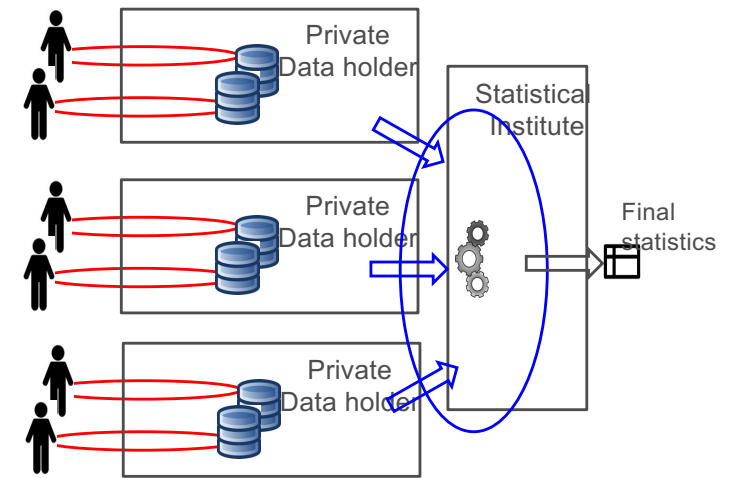
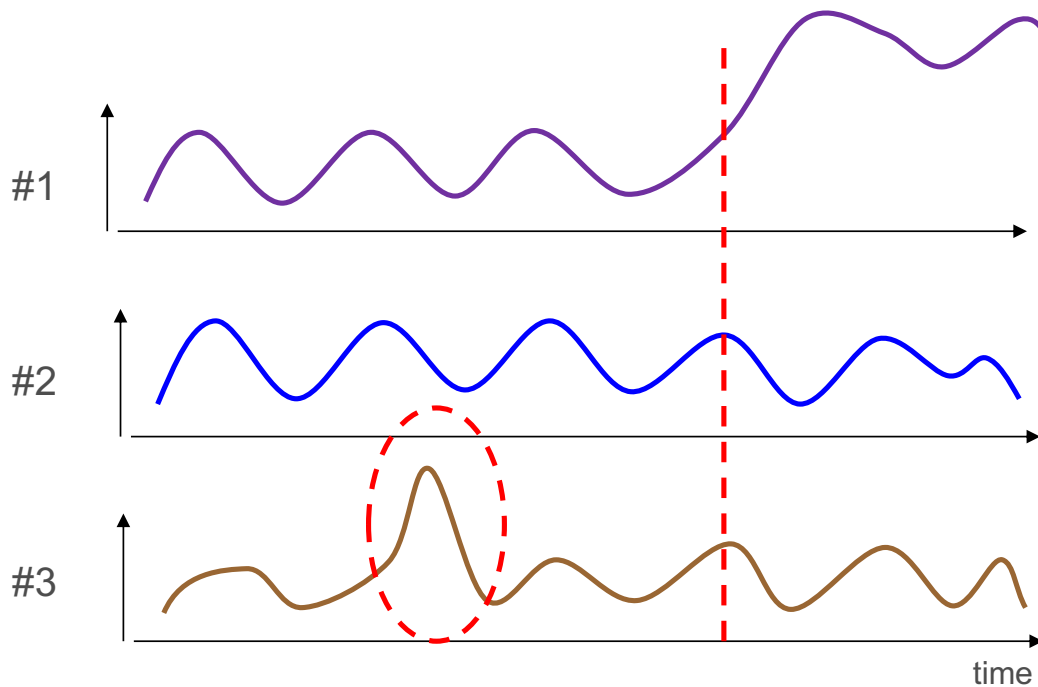
Integration of data from multiple MNOs

- **Integration of data from multiple MNOs improves quality**
- Better representativeness of the total population, mitigation of selectivity and coverage bias in the final statistics
- Improved temporal stability, mitigation of customer churning
- Mitigate sensitivity to operator-specific aspects (technical or business) of data generation
- Improved robustness to anomalies, outages, disruptions of data provision, enables cross-operator plausibility and consistency checks
- Equal treatment of all MNOs in the country (“level playing field”)
- Easier protection of business-sensitive information from each operator in the final statistics

Inter-MNO processing



Cross-MNO consistency checks

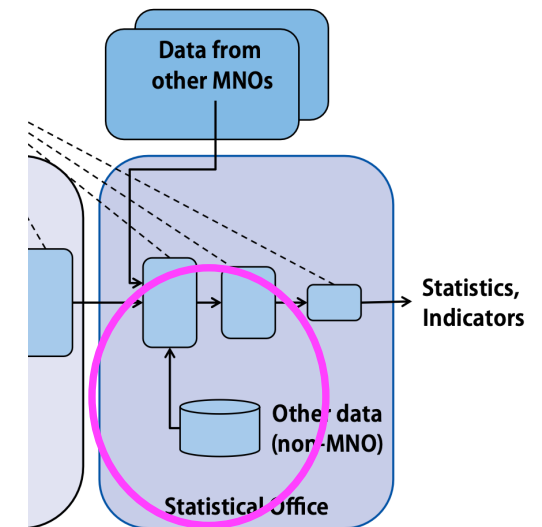


Reporting back may be beneficial for the affected MNO (more on that later)



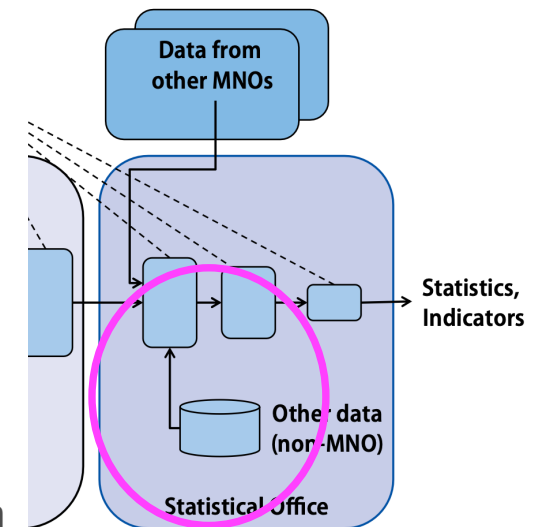
Integration of MNO data and statistical data

- Combining MNO data with statistical data to get the best of both
 - From MNO data: timeliness (near real-time), spatio/temporal detail, temporal continuity and spatial coverage, variables derived from “objective” observations
 - From statistical data: correct projection to target population (mitigation of bias, multiple counting, coverage gaps), additional variables of interest not observed by MNO data
- **Improved quality of final statistics**



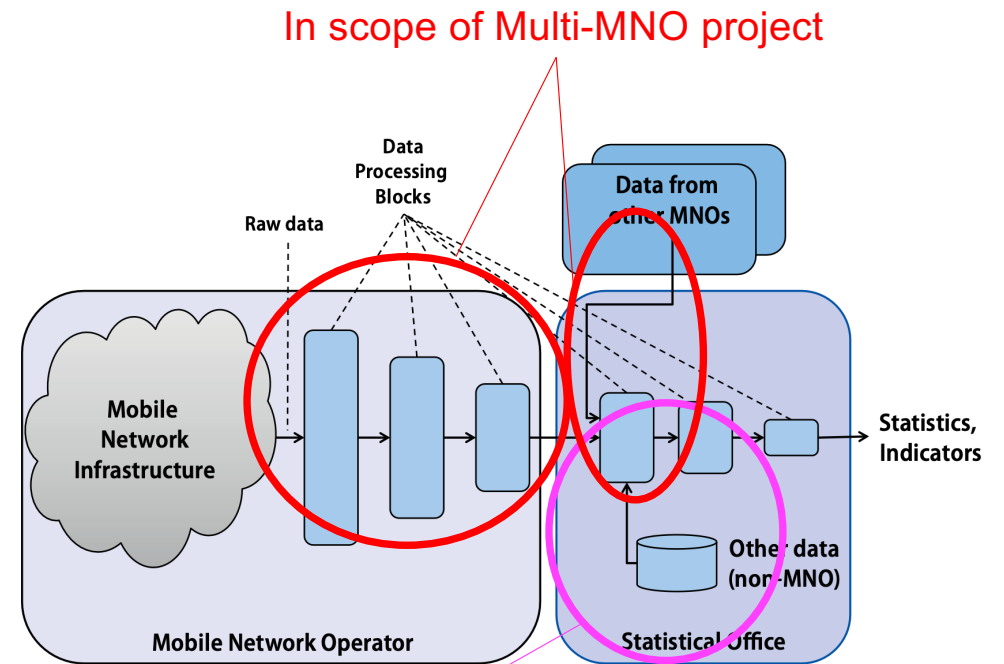
Integration of MNO data and statistical data: strategic considerations

- Reassert the role of NSIs and statistical data in the new data-rich ecosystem
 - Survey still needed, but smaller and less burdensome if combined with MNO data, for better/richer/timelier final statistics
- More balanced NSI-MNO relationship: not provider-consumer but win-win partnership
- Final statistics integrating data from **multiple providers and statistical data** produces an authoritative final statistics that may serve as “calibration reference” also for individual providers
 - Intermediate aggregate non-personal data may be shareable back with individual providers (no derogations to statistical confidentiality principle prohibiting to use statistical data for non-statistical purposes)
 - Privacy-Enhancing Technologies (PET) may help protecting data confidentiality (more on that later)



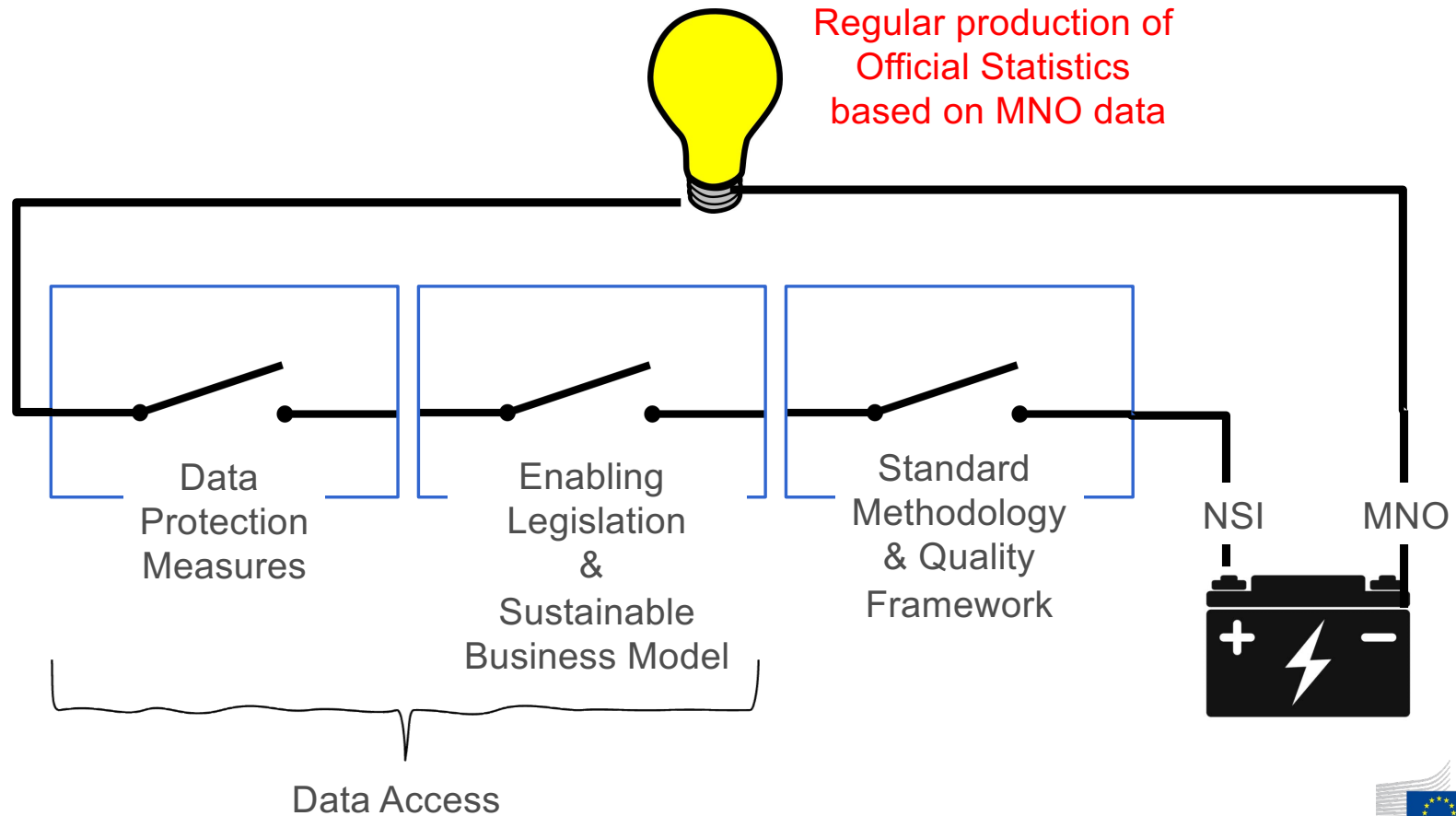
Ongoing projects

- Multi-MNO project
 - Focus on data processing pipeline: intra-MNO processing plus initial steps towards inter-MNO integration
 - Development of open-source implementation, testing on real data from 5 MNO in 4 different countries
 - NSI + Industry experts
- ESSnet MNO-MINDS
 - Research project, focus on integration of MNO and statistical data sources
 - experts from 10 NSI



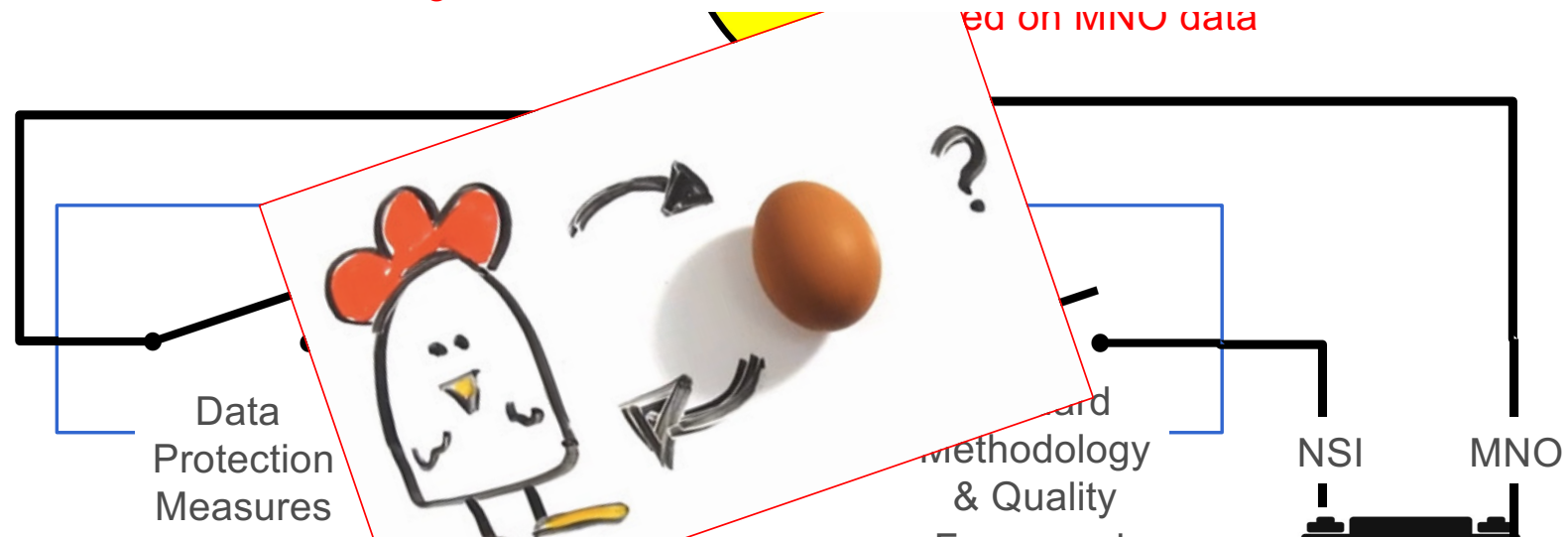
In scope of MNO-MINDS

Series of challenges ...



... with interdependencies

But there are inter-dependencies, e.g. between methodological choices and data protection aspects, and between methodological choices and data access modalities ...



... so we need to

- (1) take plausible assumptions about the future state ("reference scenario");
- (2) develop a solution based on those assumptions and reference scenario;
- (3) interact with the relevant stakeholders to discuss the proposed solution and viability of reference scenario
- (4) If needed, go back, correct assumptions and adjust methodological choices accordingly

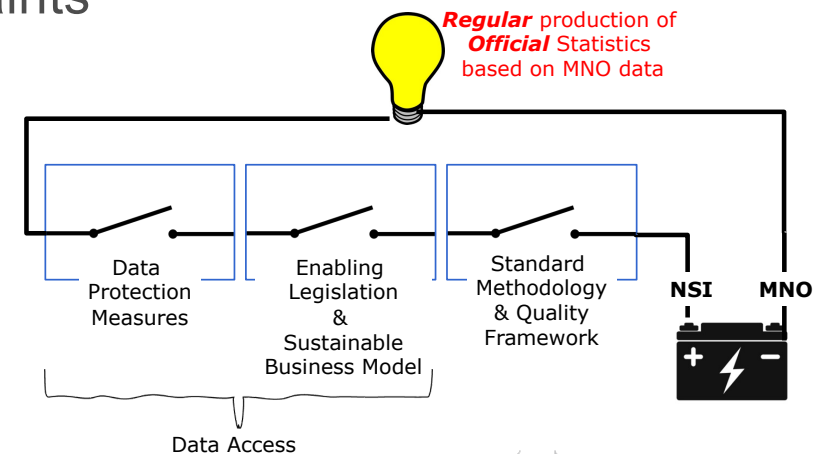
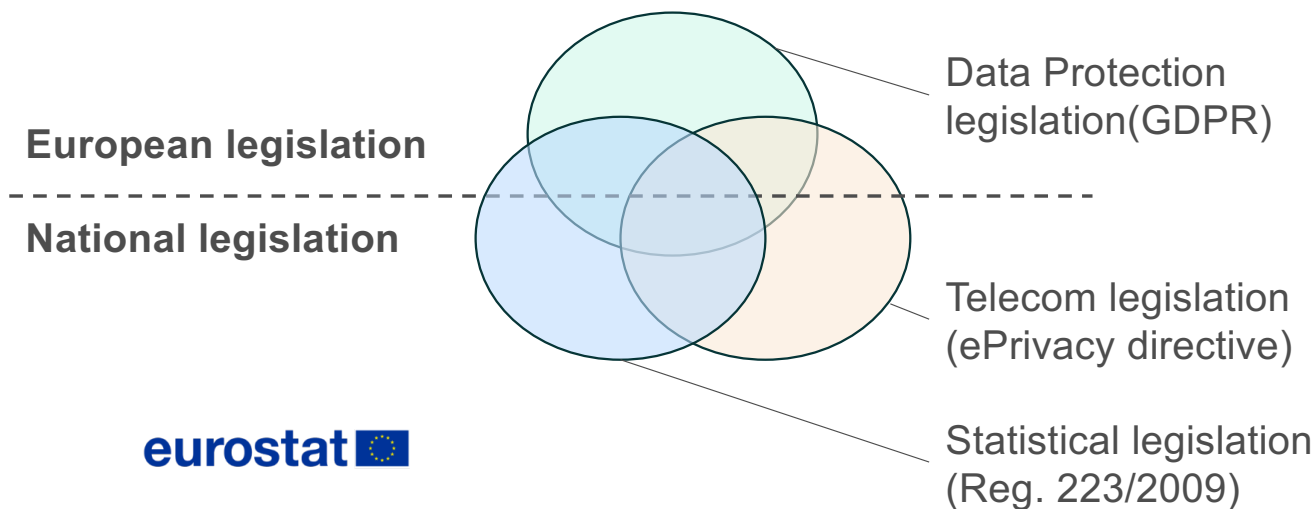
Multi-MNO project Overview

- First **co-development partnership** between NSI experts + industry experts
 - Service contract ESTAT 2021.0400 based on open call for tenders, budget 1.2 Mio
 - Started in January 2023 for 2.5 years until mid 2025
- Project Objectives
 - Develop a first proposal for an open end-to-end methodological framework (including quality aspects) with focus on an initial selection of use-cases
 - Develop and release open-source reference software pipeline implementing the proposed methodological framework
 - Practical demonstration of processing pipeline on real data from 5 MNOs in 4 countries
- Consortium: GOPA (DE); NOMMON (ES), POSITIUM (EE), ISTAT (IT), CBS (NL) + 5 MNOs (Orange Spain, Vodafone Spain, Vodafone Italy, A1 Slovenia, POST Luxembourg).
- Advisory Board of 14 external experts

Multi-MNO

Reference scenario vs. testing scenario

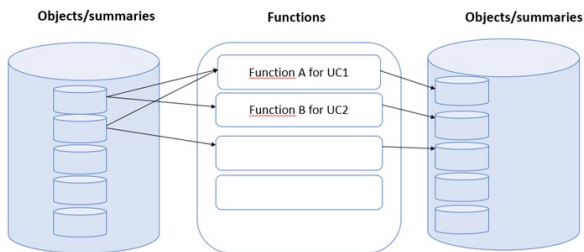
- Legal framework: fragmented (across MS) and evolving (223/2009 revision)
- Implications for data protection, data availability, access conditions
- Decoupling “*what is aimed for*” tomorrow from “*what is possible*” today
- Reference scenario: plausibly optimistic assumptions about future state
- Testing scenario: current limitations and constraints



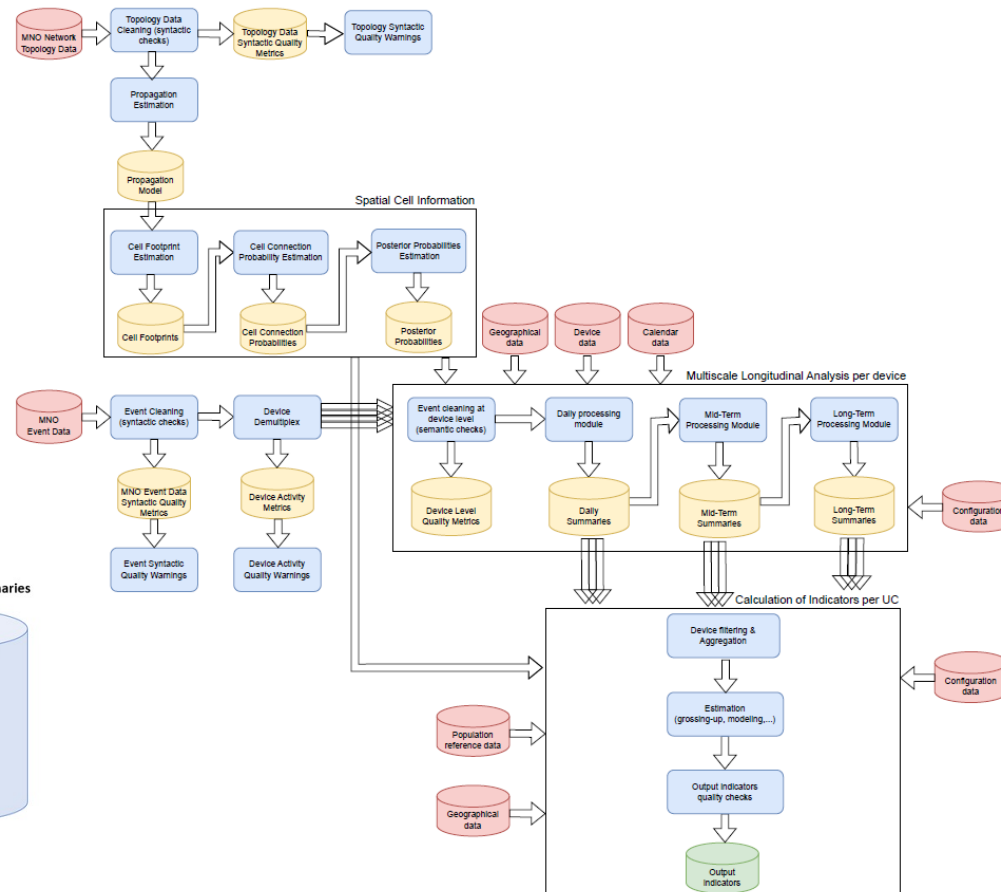
Reference methodological pipeline: Architecture

Modularity, flexibility,
and evolvability

- Data (sub)objects
- Processing (sub)modules

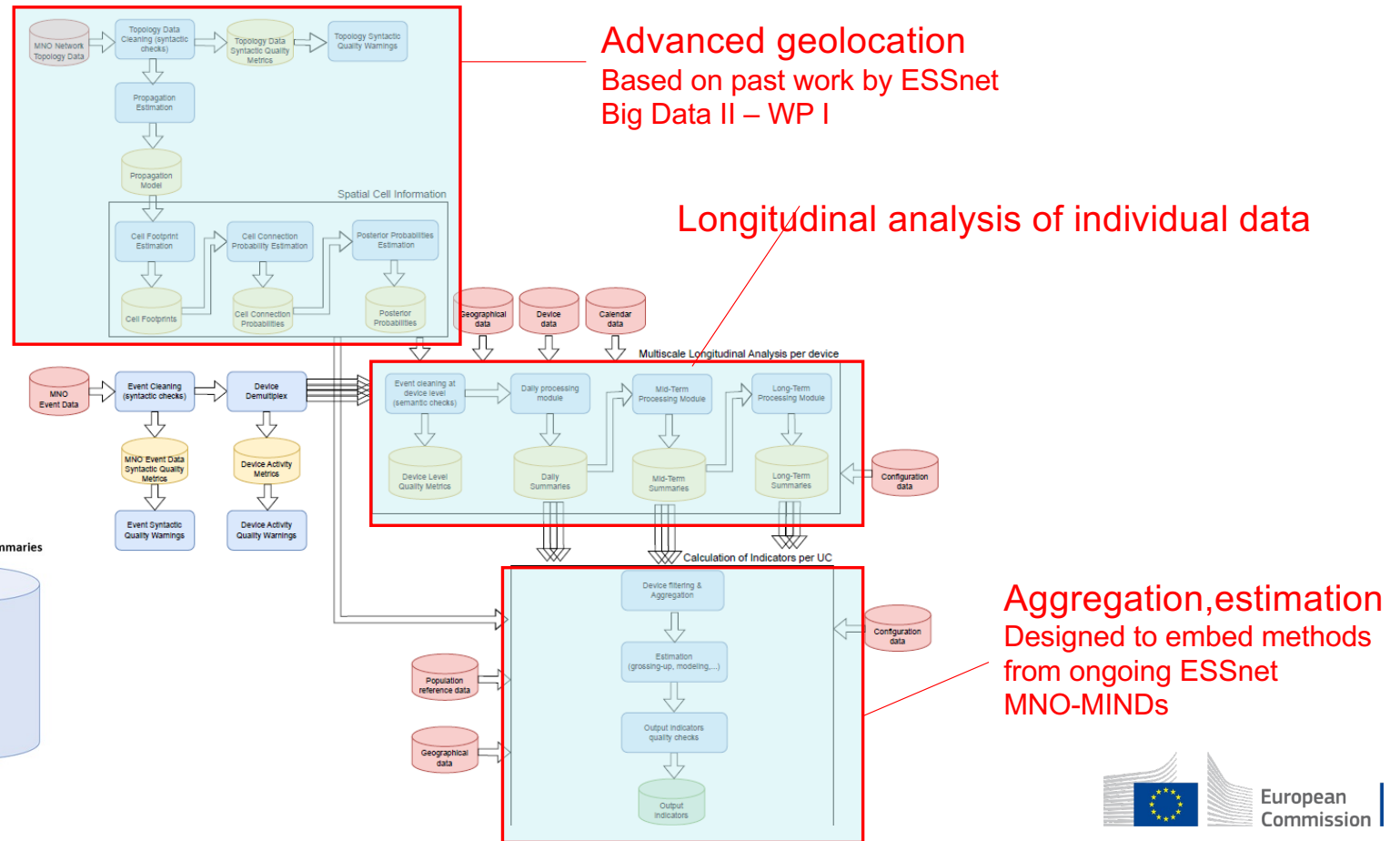
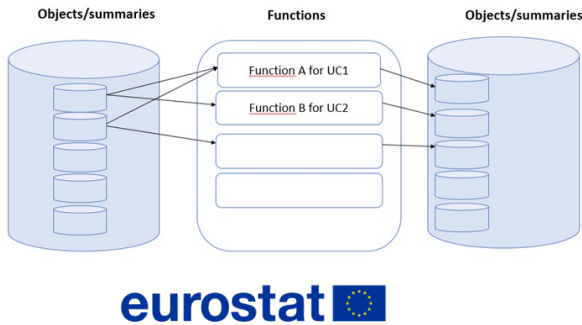


eurostat 

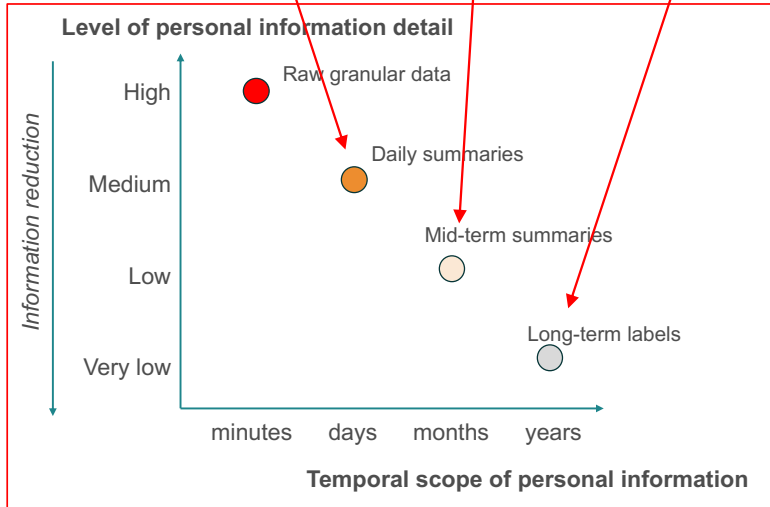
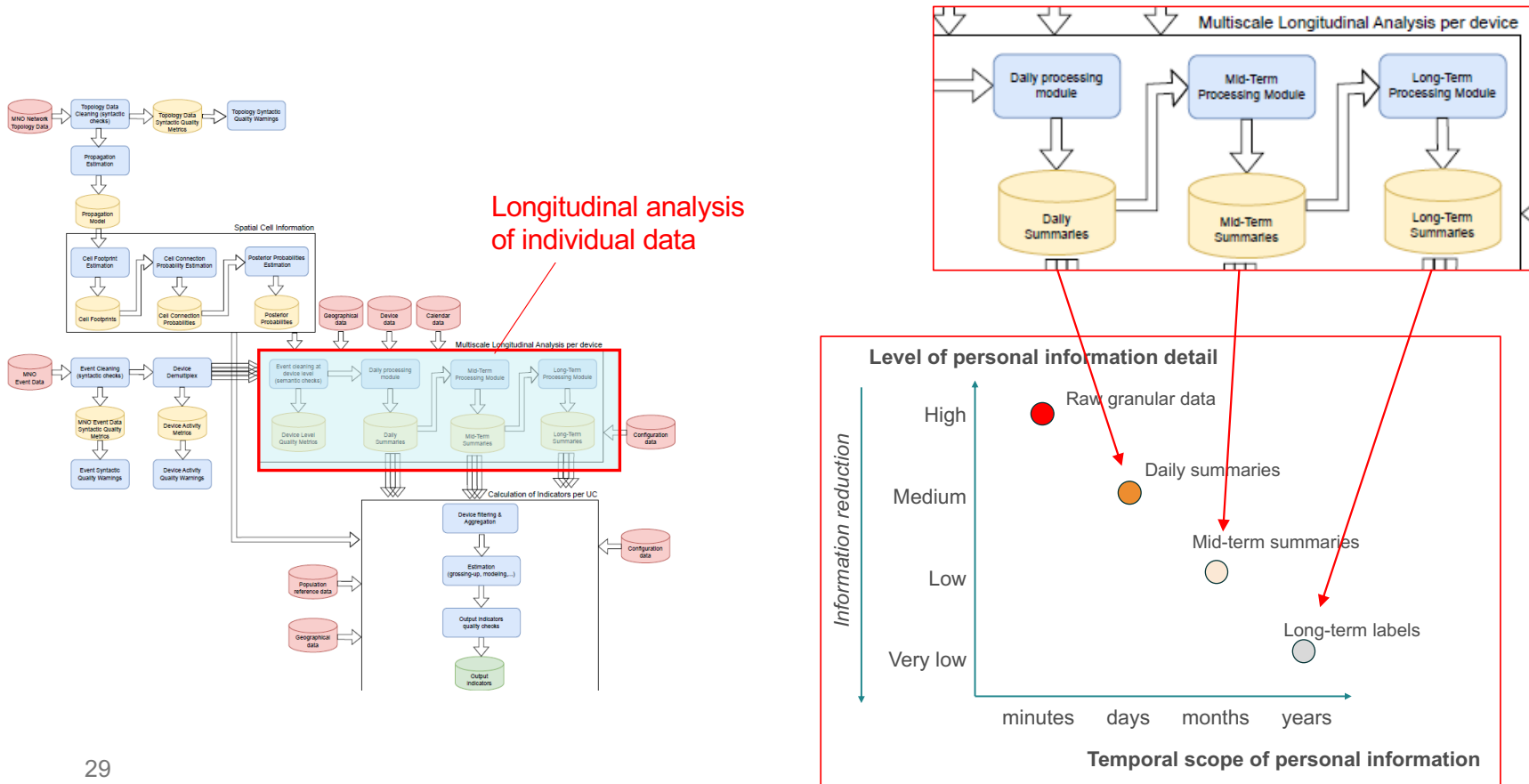


Reference methodological pipeline: Highlights

- Modularity, flexibility, and evolvability
- Data (sub)objects
 - Processing (sub)modules

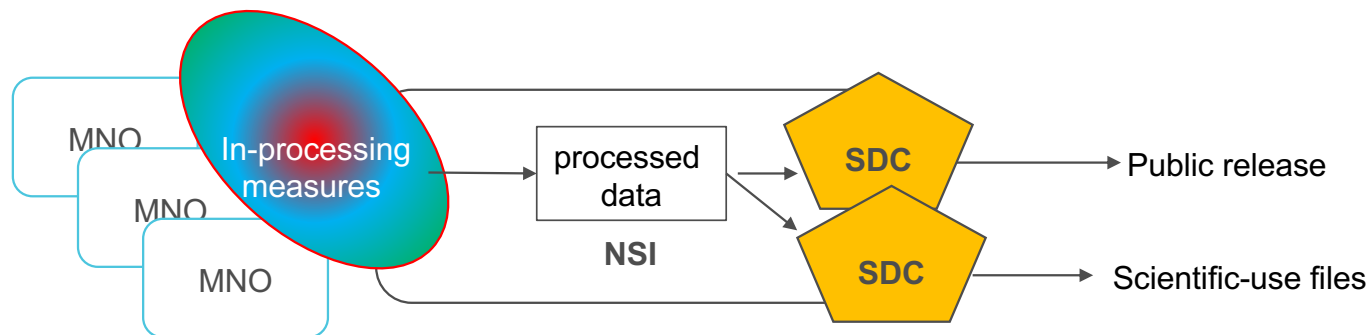


Data minimisation & Storage minimisation by [methodological] design

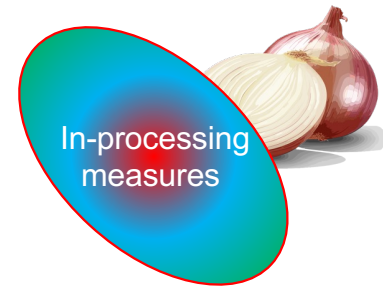


Data protection

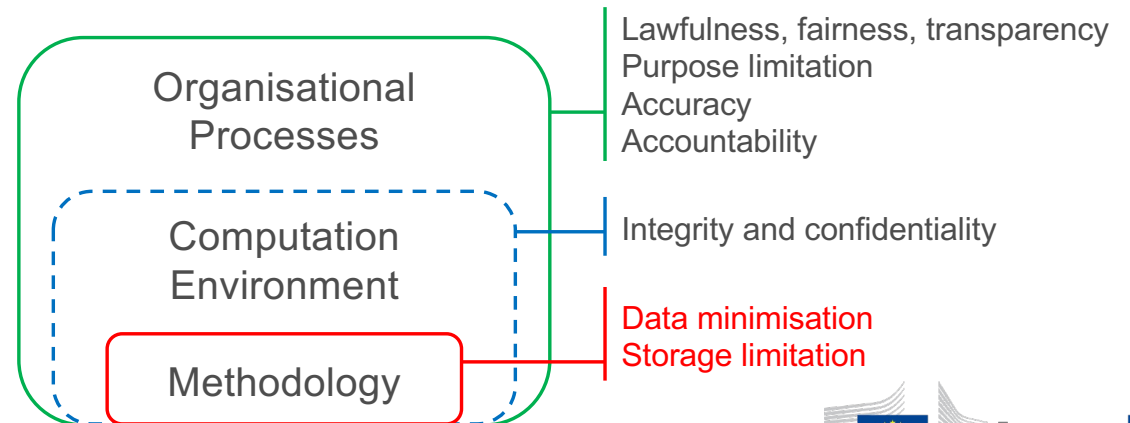
- Current practice very heterogeneous across MNOs
 - Persistence of individual pseudonyms: 24 hours, 1 day, 3 months, several years ...
 - Suppression of small aggregate values based on k-anonymity: different values of k, different points of application along the workflow, ...
- Establishing a common methodology enables **establishing a common set of data protection measures**, consistent with the proposed methodology.
 - Distinguish measures at **processing** stage (data protection engineering) and Statistical Disclosure Control (SDC) before **dissemination**



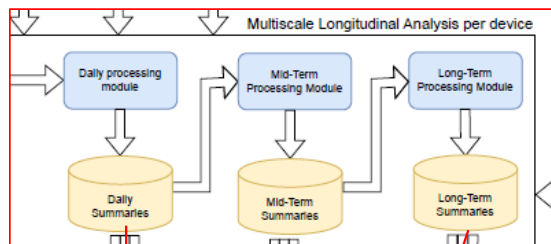
Data protection



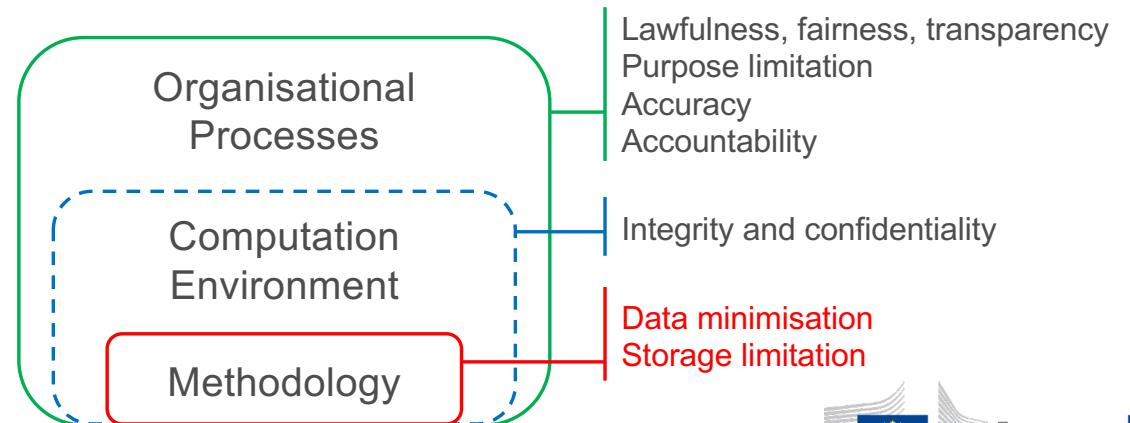
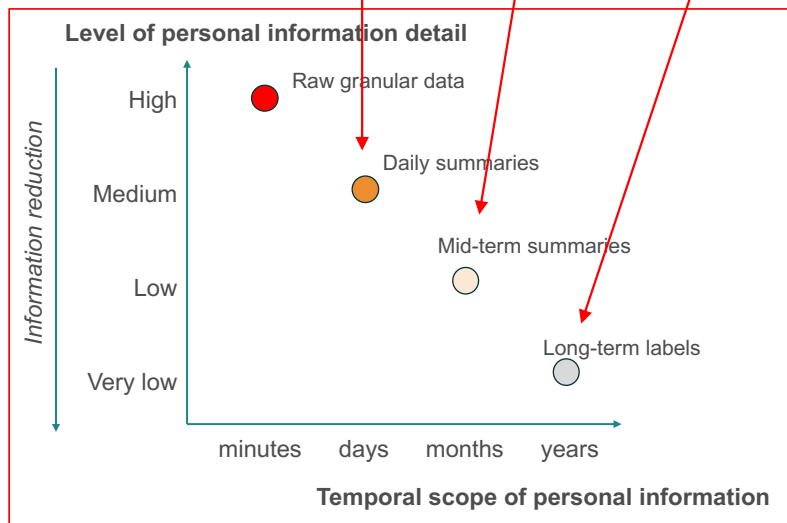
- SDC methods for dissemination of final indicators: not in the scope of current project
- In-processing data protection measures based on **onion-layered approach**: natural mapping of GDPR principles to each layer
 - Organisational environment
 - Computation environment
 - **Methodology**
- Methodology layer is in the scope of the current project



Data protection: privacy-by-[methodological]-design



- **Data minimisation and Storage minimisation by methodological design** within longitudinal analysis module
- Individual data never exported outside the safe computation environment at MNO premises
- Further supplementary Technical and Organisational Measures (TOM) may be added^(*) preferably defined at EU level – definition of methodological workflow *enables* discussion with EDPS/EDPB



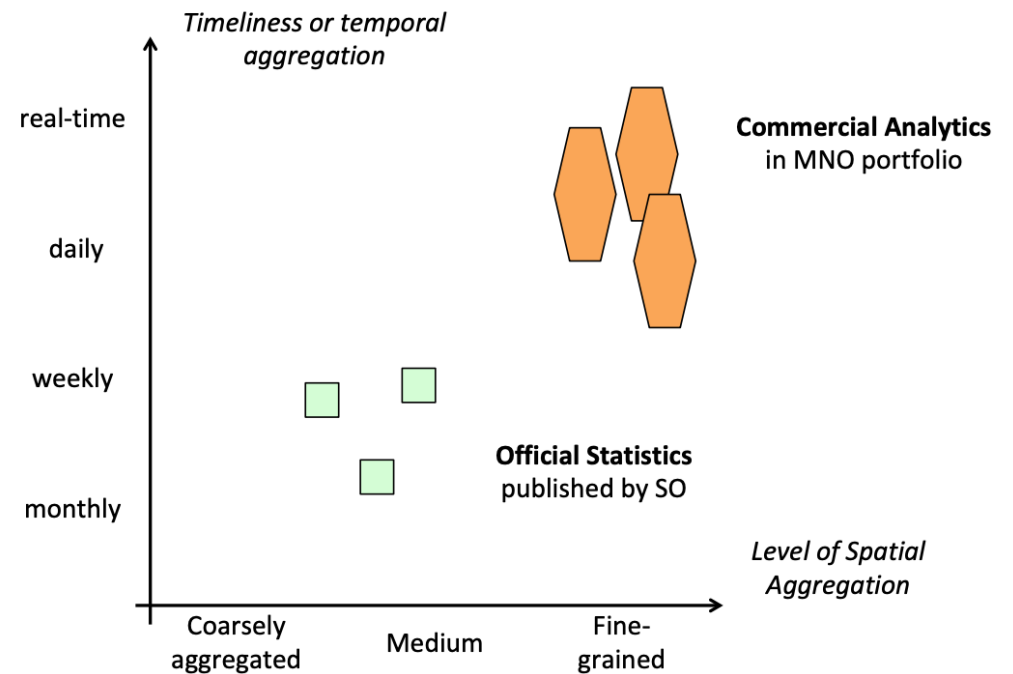
Data protection: supplementary measures

- Data minimisation and Storage minimisation *by-[methodological]-design* within longitudinal analysis module
- Individual data remain inside **safe computation environment** at MNO premise
 - *If* micro-data linkage will be needed for some use-cases, it may be implemented with support of advanced Privacy Enhancing Technologies (PET), see e.g. the parallel Eurostat project JOCONDE <https://cros.ec.europa.eu/joconde>
- Further supplementary Technical and Organisational Measures (TOM) may be added, **preferably defined at EU level**: the specification of (proposed) methodological workflow enables discussion with EU data protection authorities, going beyond current heterogeneous practices, with a view to increase both **usability** and **protection** of data
- Potential synergies with the industry if the “package” of reference methodological pipeline and associated PET/TOM is reused by industry for commercial statistics



Official Statistics and Commercial Statistics: friends, not foes

- Methodological standardisation implies also a clear specification of the statistical indicators to be eventually released by NSI
- Official Statistics released for-free by NSIs can coexist with (*and even boost the appetite by potential customers for*) commercial statistics offered for-a-fee by industry players
 - Analogy with the “Freemium” model (*)
 - Differentiation is possible along multiple dimensions



(*) see: DGINS 2018 paper Processing of Mobile Network Operator data for Official Statistics: the case for public-private partnerships

<https://zenodo.org/records/10246468>

Quality assurance requires communication between NSI and MNO

- The data generation process is “dynamic”: factors that induce a “change in the data”
 - Changes in business and/or technology, change in customer behaviour
 - Planned changes, e.g. system upgrade, new devices, new capabilities, new tariffs
 - Unplanned changes, e.g., system outage, network problems, errors, anomalies ...
- All these “facts” and “events” affect input data and their quality, therefore must be identified and reported proactively (when possible) or at least detected and reported reactively
- This requires a continuous and bidirectional exchange of information (*operational metadata, or paradata*) between NSI and MNOs

OPERATIONAL METADATA

Operational metadata are metadata that describe the expected or actual outcomes of a process using evaluable and operational metrics.”

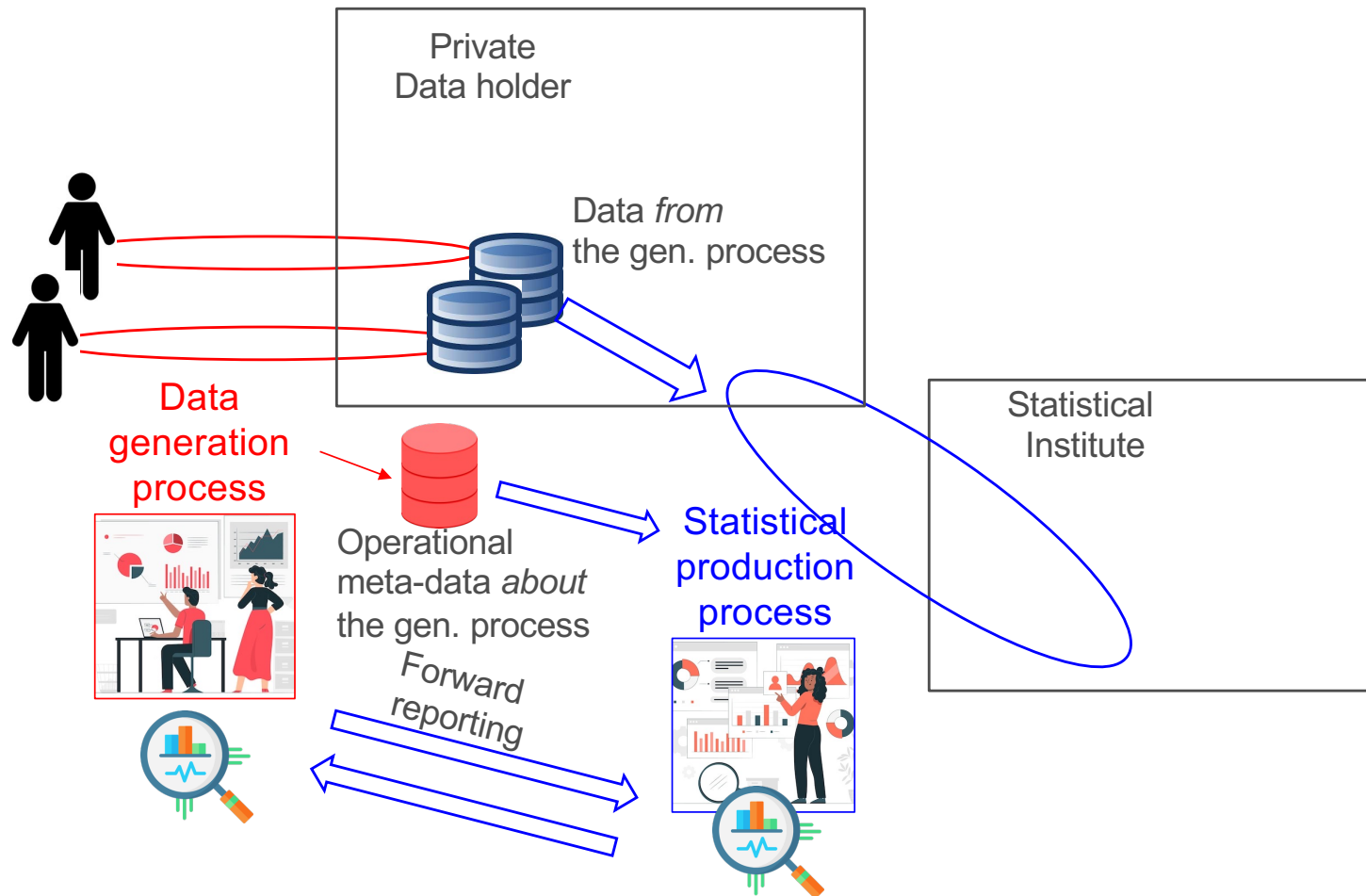
Operational metadata are a type of *reference metadata*. They include *quality metadata* and metadata measuring performance.

An alternative name is paradata.

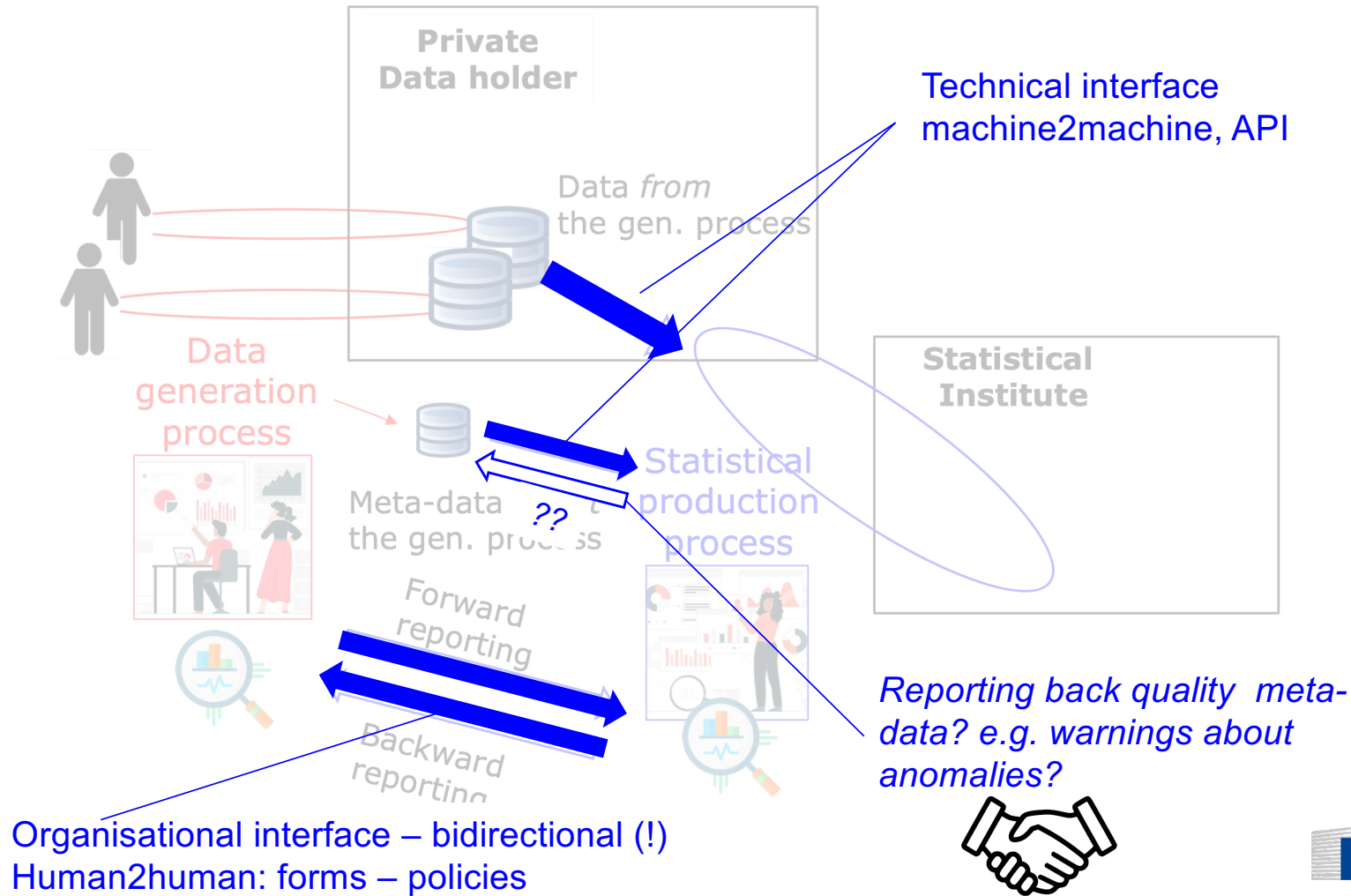
Source: ESS Handbook for quality and metadata reports

<https://europa.eu/!3cWkFk>

Communication between NSI and MNOs



Communication between NSI and MNOs



Communication between NSI and MNOs

- Communicate in both directions !!
 - MNO-to-NSI: “Next week we plan a large system upgrade, will cause outages or generation of spurious data from time X to Y in region Z”
 - NSI-to-MNO “We detected an anomalous data pattern starting around time X that apparently affects region Z, can you please help us determine what is happening”?
- Benefits for both NSI and MNO
- Attention to confidentiality of business-sensitive information (only reported to the affected MNOs – neither to its competitor nor to the public.
- All these organisational processes need to be defined in the partnership agreement, and require **resources (IT, human) on both NSI and MNO sides**
- **Balance** between under-reporting (lower quality) and over-reporting (higher burden)

Pillars of NSI-MNO partnership



- Legislative obligations
- Cost compensation
- Non-financial incentives

First, **fair and effective partnerships between businesses and statistical authorities** must be promoted on a systematic and regular basis. Such partnerships have to be based on a mutual recognition that it is legitimate for the different parties to have different roles and interests; they also need to be based on trust, social responsibility, and the premise that the reuse of privately held data for official statistics will benefit the whole of society – including the partners themselves. This presupposes a balanced regulatory framework that makes it possible to reuse privately held data for statistical purposes in a sustainable way.

- **Statistical authorities and private data holders should develop a partnership approach to maximise business incentives and minimise risk, based on mutually agreed operational modalities of data reuse.** Statistical authorities should provide incentives and make it easier for private data holders to set up successful partnerships with them. They could, for example, help private data holders to enhance their public reputation and provide feedback advice on ways to improve and better use their data.
- The legal framework should set out a clear set of requirements and safeguards for private data holders. Where it is not possible to establish voluntary partnerships based on incentives, there should be a **legal requirement to ensure the sustainable production of official statistics – subject to a clear set of limitations, conditions and safeguards** (proportionate costs and benefits; mechanisms to deal in a fair way with costs incurred by data holders; protection of business interests and limitation of business liability; and adequate protection for data confidentiality). While statistical authorities should not pay for data, **viable cost compensation schemes** should be developed, including for services offered by data holders that the statistical authorities decide to use. Dispute resolution mechanisms should also be put in place.



According to the B2G4S
Expert Group
<https://europa.eu/!JGR3Gx>

Legislative obligations



- **Obligation** already included in the revised text of the amended Regulation (EC) 223/2009 on European Statistics, Article 17b
- Balanced by a set of **conditions** (for requesting authority) and **safeguards** (for data holders)

Article 17b

Obligation of private data holders to make data available for developing, *producing* and *disseminating* European statistics

1. Without prejudice to reporting obligations, *data collections or any data access* laid down in sectoral statistical legislation of the Union *or* to the obligation for data holders to make data available on the basis of an exceptional need in accordance with *Regulation (EU) 2023/2854 of the European Parliament and the Council**, an NSI or the Commission (Eurostat) may request a private data holder to make data and the relevant metadata available *free of charge where the data requested are strictly necessary* for the development, *production and dissemination* of European statistics *and cannot be obtained by other means* or their reuse will result in a considerable reduction in the response burden on data holders and other businesses. *Such data collections or data access may be*

Financial Compensation?



- Purchase of data is excluded: NSI don't pay for the data as such
 - Official Statistics do not create any new “data monetisation” opportunity.
- However, the revised text of the amended Regulation (EC) 223/2009 on European Statistics, Article 17b, point 5, foresees the **possibility** (for Member States or the Commission) to provide financial compensation to data holders for “**specific [data] processing services**”

5. *Where data requested by an NSI under paragraph 1 require a specific processing service, Member States may provide compensation to the private data holder for that specific processing service, except where national law prevents NSIs or other national authorities responsible for the production of statistics from compensating data holders. Where data are requested by the Commission (Eurostat) for efficiency reasons pursuant to paragraph 3, and a specific processing service is needed, the Commission (Eurostat) shall propose reasonable compensation to the private data holder for that specific processing service.*

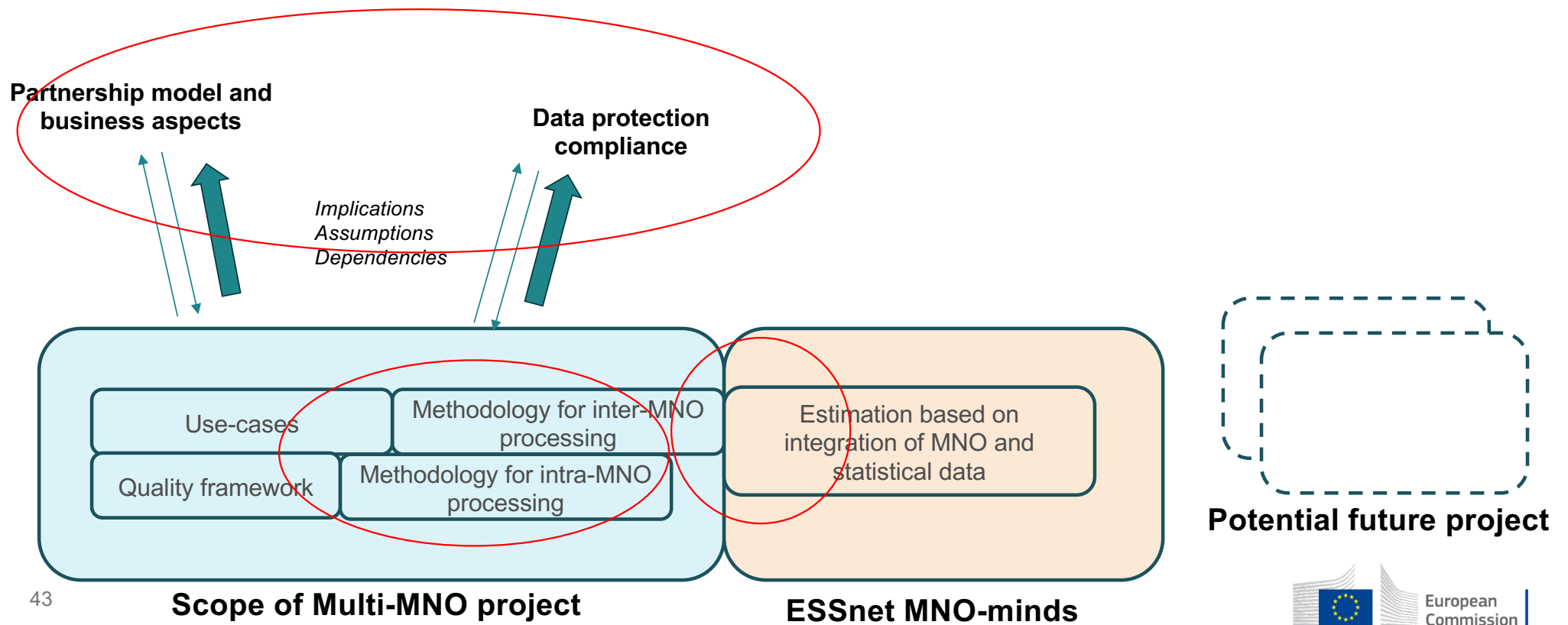
Financial Compensation? – open issues



business liability; and adequate protection for data confidentiality). While statistical authorities should not pay for data, viable cost compensation schemes should be developed, including for services offered by data holders that the statistical authorities decide to use. Dispute resolution mechanisms should also be put in place.

- *In what form?*
 - Tax deductions or cash?
 - *Who* should provide it? NSIs or other public bodies? Public programs?
- *How much?* On the basis of what “cost model”?
 - Are new investment needed to comply with NSI requests?
 - How to incentivise continuous improvements?
 - What financial terms are “fair” given heterogeneity of technical capabilities and past investments across MNOs?
 - The reference methodology (and quality) framework specifies what exactly is requested/expected by MNOs and therefore enables this discussion

Reference methodology and quality framework enable addressing business-related aspects



Non-financial incentives

- Reputation and Corporate Social Responsibility: Official Stat. as public good
- Benefit on Commercial Statistics
 - Final statistical indicators (integrating data from all MNOs and adjusted with statistical data) serving as “calibration reference” for commercial statistics
 - Quality monitoring for Official Statistics “feedback” and improves commercial statistics
 - Increase appeal among potential customers – as in freemium model
- Additional benefits if the Reference Methodological Framework defined by the ESS for Official Statistics is taken as base for Commercial Statistics
 - Sound definitions and concepts, increase transparency and credibility of commercial statistics without jeopardizing competition across MNOs
 - Clear and homogeneous Data Protection measures

Outlook

- Multi-MNO project ends in June 2025. MNO-MIND ends in October 2025.
 - Project deliverables will be made public and serve as basis for next step
- Adoption of amended Regulation (EC) 223/2009 expected in early 2025
- Next step: Eurostat dialogue with EU Data Protection Supervisor (EDPS)
- Future actions (to be planned)
 - Follow-up projects on methodological specifications, piloting, testing, ...?
 - Dialogue with industry representatives?
 - Dialogue with civil society?

Final words



- Official Statistics based on MNO must be based on fair and sustainable **partnership** between NSI and MNOs
- Developing a partnership model is **possible** if based on an holistic view combining **legal obligations, financial compensation** and a set of **non-financial incentives**
- The clear specification of **open reference methodological framework** is a **precondition** for designing such a partnership model
- Multi-MNO and MNO-MINDS to provide a **sound methodological basis**, to be further consolidated in future projects and methodological actions
- No compromise on protection of personal data and business confidentiality.
 - Advanced Protection measures, possibly based on state-of-the-art Privacy-Enhancing Technologies (PET), see e.g. the project JOCONDE (Joint On-demand Computation with No Data Exchange)
<https://cros.ec.europa.eu/joconde>

For follow-up

- Email to
`Fabio.Ricciato@ec.europa.eu`
- MNOdata4OS page on CROS portal <https://cros.ec.europa.eu/MNOdata4OS>
 - With links to relevant resources: Position Paper by TF MNO, B2G4S final report, approved text of amended regulation etc. 223, project pages, etc.

Thank you



© European Union 2024

Unless otherwise noted the reuse of this presentation is authorised under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license. For any use or reproduction of elements that are not owned by the EU, permission may need to be sought directly from the respective right holders.