



Reusing Mobile Network Operator data for Official Statistics: steps towards a common methodological framework for the European Statistical System

Fabio Ricciato

Eurostat - Unit A5 Methodology; Innovation in Official Statistics

First meeting of project Advisory Board

8. February 2023

Eurostat and the European Statistical System (ESS)

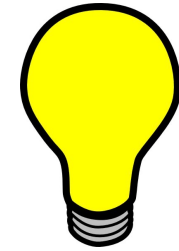


- *Eurostat is ...*
 - **the statistical office of the EU**
 - **a DG of the European Commission**
 - **the coordinator of the ESS**
- *The European Statistical System (ESS) is the partnership between*
 - Eurostat (coordinator)
 - National Statistical Institutes (NSIs) in each EU country
 - Other National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

Past work on MNO data for OS in the ESS

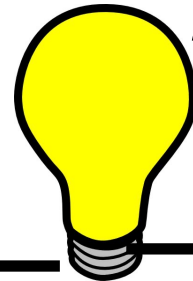
- *Activities at national level in some EU countries, not all*
- *Mostly limited to "research" or "experimental statistics"*
- *Intermittent, heterogeneous in every aspect*
 - Mostly based on purely voluntary agreements NSI-MNO
 - Reflecting **fragmentation of national legislations** and differences in MNO-NSI relationships
 - Mostly based on **proprietary methods** developed by MNOs or their partner companies
 - Rarely involving more than a single MNO within each country
- *Boost of NSI-MNO collaborations during Covid crisis*
- *At European level, the project ESSnet Big Data II (2018-2020) started to lay initial foundational work towards a common EU approach*

The vision

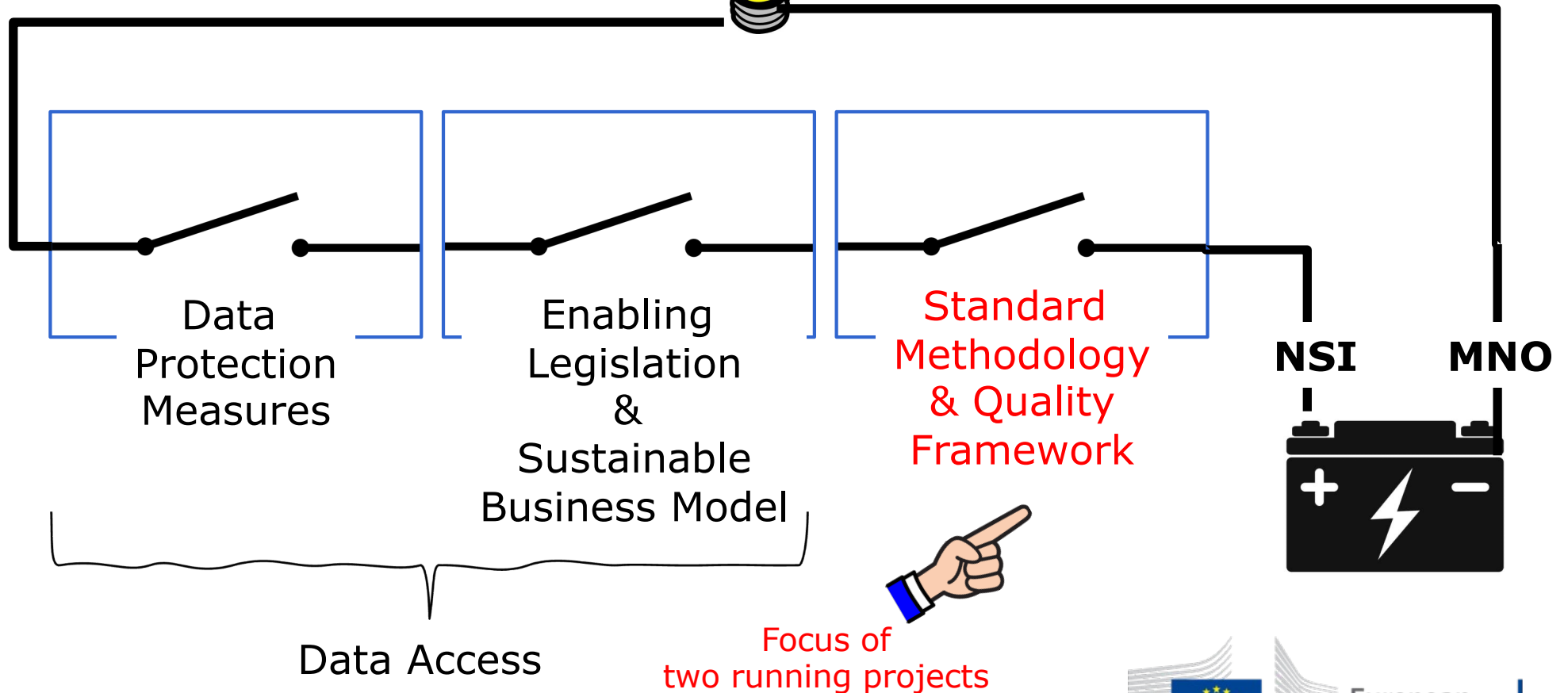


- *In 202x MNO data are (re)used for regular production of **official statistics** (not merely experimental statistics) ...*
 - In all EU countries
 - Based on a sustainable partnerships between NSI and MNO
 - Based on data from multiple MNOs in each country
 - **Processed according to standard and open methodologies and transparent quality criteria defined at EU level**
 - Incorporating strong technical and organisational measures for agreed at the EU level for protecting personal data (privacy) and business sensitive information
 - **Combined with other statistical data for calibration/correction/stabilisation**

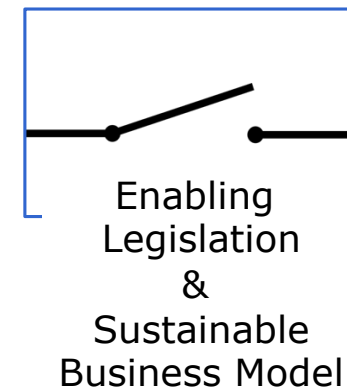
Series of Challenges



Regular production of **Official** Statistics based on MNO data



Data Access – recent actions

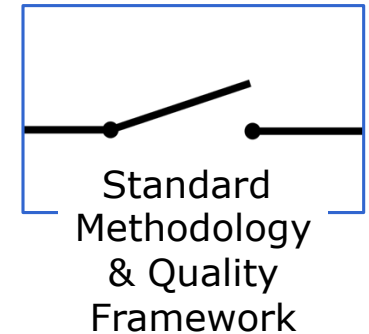


- *Proposal for amendment of Regulation (EC) 223/2009 on European statistics*
 - Relevant for all Privately Held Data
 - **Approved by European Parliament in March'24**
https://www.europarl.europa.eu/doceo/document/TA-9-2024-0152_EN.pdf
 - Article 17b **Obligation of private data holders** to make data available for developing, producing and disseminating European statistics
 - Art. 17c **Requests for data** and arrangements for making data available for development, production and dissemination of European statistics

Article 17c
Requests for data and arrangements for making data available for development,
production and dissemination of European statistics

1. When requesting data in accordance with Article 17b, the NSIs or the Commission (Eurostat) shall:
 - (a) specify what data *and metadata* are required;
 - (b) *specify* the statistical need for which the data are requested in accordance with Article 17b(1);
 - (c) specify the frequency with which and the deadlines by which the data are to be made available;
 - (d) specify the operational arrangements for making the data available.

Methodological aspects



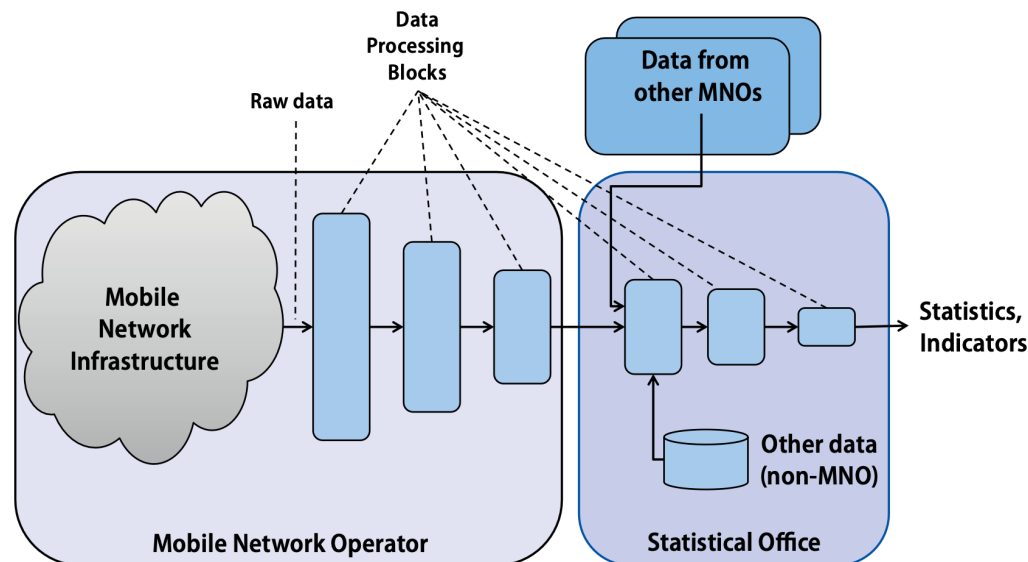
- ***ESS Task Force on MNO data for Official Statistics (TF MNO)***
 - Established in 2021 with the official mandate to steer methodological development in the field
 - NSI representatives from 19 EU countries
- ***Position paper by the TF MNO clarifying the methodological approach (2023)***
 - Reusing Mobile Network Operator data for Official Statistics: the case for a common methodological framework for the European Statistical System

<https://ec.europa.eu/eurostat/en/web/products-statistical-reports/w/ks-ft-23-001>



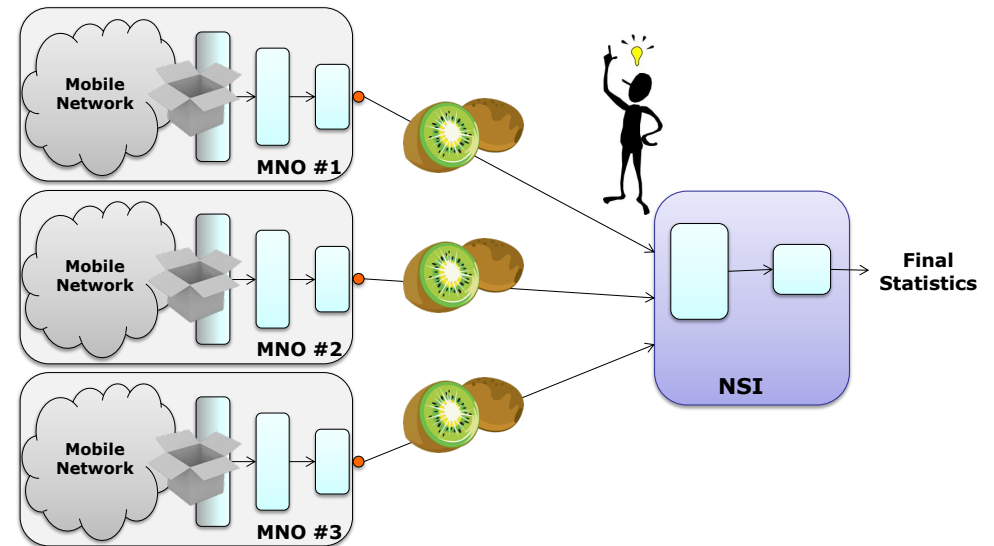
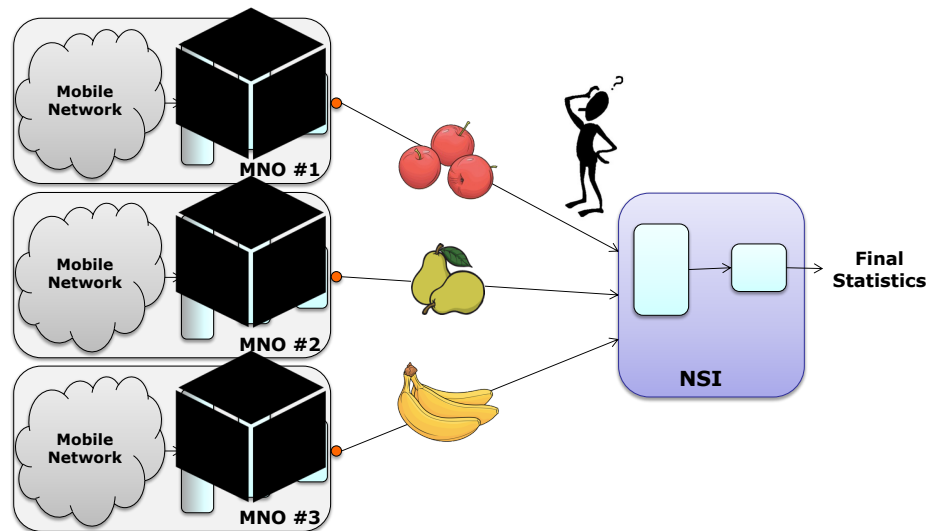
The case for methodological standardisation

- *Methodologies for transforming raw MNO data into official statistics to be as standardised as possible across EU*
 - One open ESS methodology applied to all MNOs instead of many proprietary/closed methodologies
- *Standardisation to cover whole **end-to-end data flow***
 - Oblivious to where processing takes place, MNO premises vs NSI
 - Standard cannot be too rigid: some degree of flexibility is needed to adapt to different MNO settings (raw data are NOT standardised!)

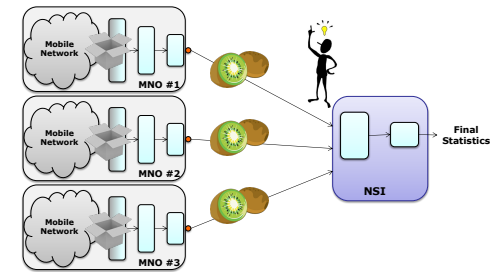


The case for methodological standardisation

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



Multi-MNO orientation



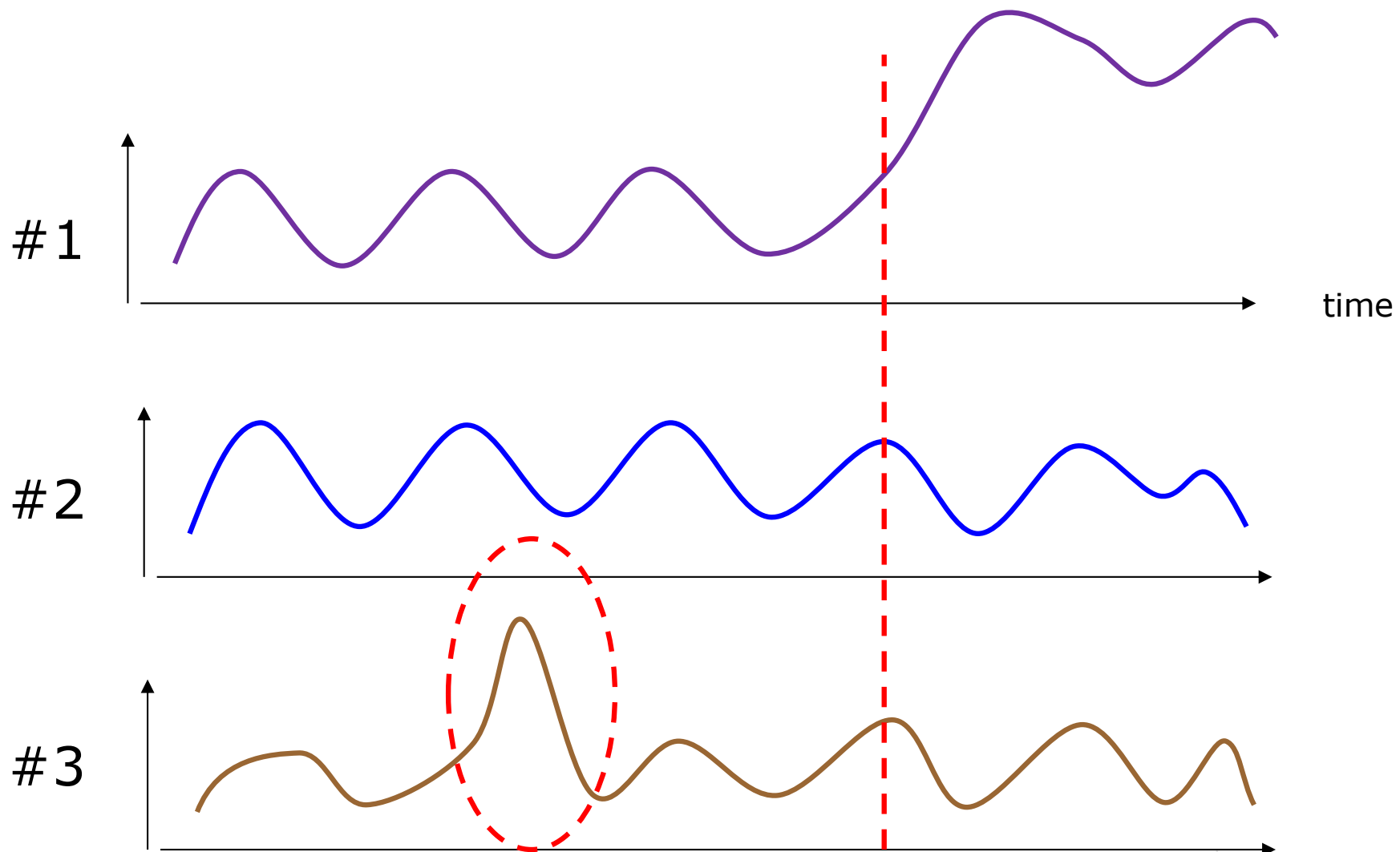
*"The envisioned ESS methodological standard should be designed upfront to produce statistics based on the combination of information **sourced from multiple MNOs** (multi-MNO statistics) within and across different European countries. "*

From <https://ec.europa.eu/eurostat/en/web/products-statistical-reports/w/ks-ft-23-001>

*Advantages of Multi-MNO approach **within each country***

- *Better representativeness of the total population, lower exposure of the final statistics to population coverage bias*
- *Improved temporal stability, mitigation of churning in mobile subscriptions*
- *Improved robustness to anomalies and disruptions of data provision, lower sensitivity to MNO-specific network conditions*
- *Equal treatment of all MNOs ("level playing field")*
- *Easier protection of business-sensitive of individual MNOs information in the final statistics*

Consistency and plausibility checks across data providers



Fusion of MNO data and non-MNO data

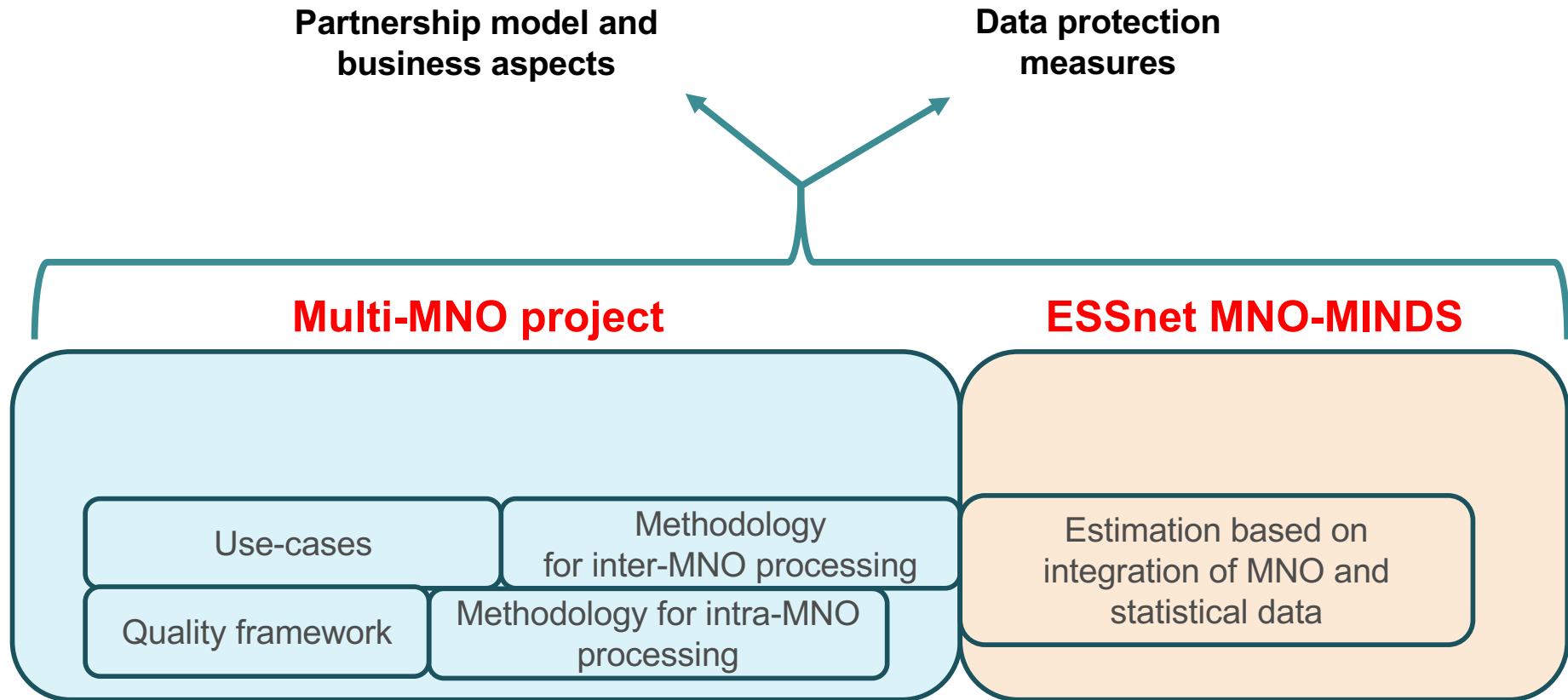


*"Experience with past experimental and research projects has shown the necessity to **combine MNO data with other kinds of non-MNO data to stabilise, correct or calibrate** the final statistics. "*

From <https://ec.europa.eu/eurostat/en/web/products-statistical-reports/w/ks-ft-23-001>

- *Methodological and quality need: non-MNO data can help to assess/correct/mitigate issues related to population coverage uncertainty (under- and over-representation), churning of mobile users, lack of socio-demographic variables, etc.*
- *Open research aspects: Which non-MNO data are best fit to be combined with MNO data? And how to combine them with MNO data?*
 - New ad-hoc sample survey designed *on purpose* by NSI?
 - Legacy survey data? Census data?
 - Other "big data" sources held by external entities (e.g., train tickets)

Two projects



Multi-MNO - Project Factsheet

<https://cros.ec.europa.eu/multi-mno-project>

- *Multi-MNO project – service contract ESTAT 2021.0400*
 - (public procurement based on open call for tenders)
- *Total duration 2.5 years - termination by mid-2025*
- *Goals*
 - Define reference methodological pipeline
 - Develop an open-source reference implementation (in python)
 - Test it on real-world data from 5 MNOs in 4 EU countries
- *Consortium*
 - GOPA Germany (consortium leader)
 - POSITIUM Estonia and NOMMON Spain (industry partners)
 - ISTAT Italy and CBS Netherland (NSIs)
 - 5x MNOs (subcontractors) - ORANGE Spain, VODAFONE Spain, VODAFONE Italy, A1 Slovenia, POST Luxembourg
- *Advisory Board composed of independent individual experts*
<https://cros.ec.europa.eu/multi-mno-project>

Note on ownership and market relevance

- *The standard methodology developed by the ESS will be fully open. It should be accompanied by a open-source reference implementation owned by EC and **released under EUPL license***
- *The reference pipeline will be completely **open-source**, thoroughly documented, and with everything that is needed to enable **independent reproduction** and alternative (possibly proprietary) implementations (e.g., functional test scenarios, benchmark data). The **reference pipeline** is also owned by EC.*
- *Private companies may compete in the future to offer, possibly on a commercial basis: proprietary pipeline implementations, supplementary out-of-standard features, consultancy services for deployment and operation of the standard pipeline, etc.*
- *Parts of this standard may be freely reused by private companies and MNOs for commercial statistics*

Reference Scenario vs. Demonstrator Scenario

- *Decoupling "what we aim for tomorrow" (RS) from "what is possible today" (DS)*
 - **Don't let today's constraints limit the design for tomorrow**
 - **To some extent, ambitious design for tomorrow helps waiving today's constraints**
- *Reference Scenario (RS)*
 - Future desirable state, based on plausibly optimistic assumptions
 - Mainly relevant for **Specifications of methodologies and quality** aspects (Task 2, Task 3)
- *Demonstrator Scenario (DS)*
 - What is possible today based on existing legislation (eg., data protection) and contractual agreement
 - Relevant only for **specification of demonstration/tests** on real-data

MNO-MINDS - Project Factsheet

<https://cros.ec.europa.eu/mno-minds>

- *Research grant reserved for NSI (internal ESS call)*
- *Total duration 2 years - termination by end-2025*
- *Focus on the **integration of MNO and non-MNO data***
 - Landscaping non-MNO data and identify best candidates
 - Developing methods and tools for their integration
 - Design ad-hoc survey for calibrating MNO data
- *Consortium by 10 NSIs coordinated by ISTAT (IT)*

ISTAT	ISTITUTO NAZIONALE DI STATISTICA	IT
STAT	BUNDESANSTALT STATISTIK OESTERREICH	AT
DESTATIS	STATISTISCHES BUNDESAMT	DE
INE	INSTITUTO NACIONAL DE ESTADISTICA	ES
INSEE	MINISTERE DE L'ECONOMIE DES FINANCES ET DE LA RELANCE	FR
CBS	CENTRAAL BUREAU VOOR DE STATISTIEK	NL
SSB	STATISTISK SENTRALBYRAA	NO
INS	NATIONAL INSTITUTE FOR STATISTICS	RO
SCB	STATISTISKA CENTRALBYRAN	SE
INE-PT	INSTITUTO NACIONAL DE ESTATISTICA PORTUGAL	PT

European perspective



- 1. Assuming (prospectively) comparable data access situation across EU countries (legal, business)*
- 2. Developing methodology that may work (prospectively) for all MNOs in all EU countries*
 - Built-in flexibility to accommodate differences about mobile infrastructures and business processes
 - Setting reasonably ambitious minimum standards (neither bound to current “worst case” nor unrealistic “ideal case”)
- 3. Tailored to European state-of-play in mobile technology, services, usage and statistical needs*
 - May be different from non-European region

Thanks for your attention!



More info on Eurostat activities in the field:
<https://cros.ec.europa.eu/MNOdata4OS>