Service contract ESTAT 2021.0400
Multi-MNO project
Kick-off meeting

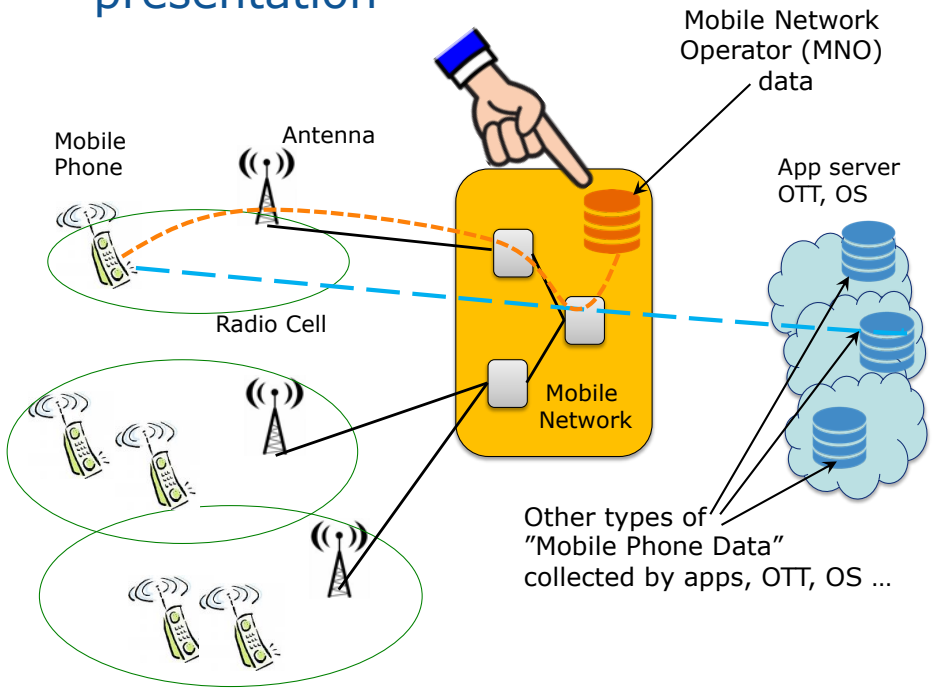
Context, motivations and goals of the project

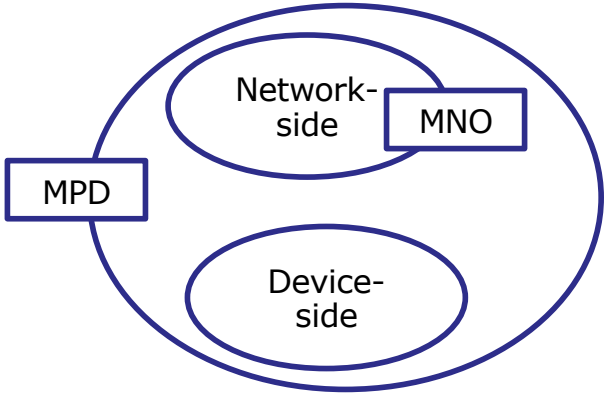
Fabio Ricciato
Eurostat - Unit A5 Methodology; Innovation in Official Statistics

First meeting of project Advisory Board
8. February 2023

MNO data or MPD data?

- *'MNO data' are location data collected by the telecom operators*
  - Produced by "interaction" between mobile devices and mobile network
  - Collected primarily for billing (CDR) and network maintenance (signalling) purposes
- *MNO or MPD?*
  - Location data produced by the mobile device and collected by apps, OS, Over-The-Top (OTT) companies are also called 'Mobile Phone Data" (MPD) but they are not "MNO data" and are not in the scope of this presentation





European Commission

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*

European Commission

Official Statistics

- *"Official statistics" vs. "Experimental statistics"*
  - Regular production vs. one-off/short series
  - Complete vs. partial fulfillment of quality criteria
- *Data sources for statistical production*
  - Census, surveys – designed and collected by NSI
  - Administrative data
  - "Big Data", Privately Held Data (future)
    - Mobile Network Operator (MNO) data
    - …
- *Goal: regular production of (new, better, richer, timelier) official statistics based on MNO data*

Quality Assurance Framework of the ESS

Institutional environment
- Principe 1 : Professional Independence
- Principe 1bis : Coordination and cooperation
- Principe 2 : Mandate for Data Collection and Access to Data
- Principe 3 : Adequacy of Resources
- Principe 4 : Commitment to Quality
- Principe 5 : Statistical Confidentiality
- Principe 6 : Impartiality and Objectivity

Statistical processes
- Principe 7 : Sound Methodology
- Principe 8 : Appropriate Statistical Procedures
- Principe 9 : Non – excessive Burden on Respondents
- Principe 10 : Cost effectiveness

Statistical output
- Principe 11 : Relevance
- Principe 12 : Accuracy and Reliability
- Principe 13 : Timeless and Punctuality
- Principe 14 : Coherence and Comparability
- Principe 15 : Accessibility

Shines and shadows of MNO data for statistics

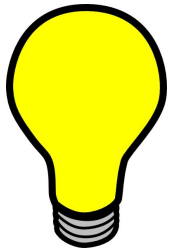
- *'MNO data' carry spatiotemporal information about phones*
  - Records in the form **<phone ID, timestamp, cell ID>**
  - Can be used to measure human **presence** (where people are) and **mobility** (where and when they move to/from)
  - *… taking "phones" as proxies for "humans"*
- *Appealing aspects of aggregate statistics based on MNO data*
  - Timeliness, temporal granularity, temporal continuity
  - Spatial coverage - whole country, whole EU
  - Population coverage - almost everybody carries some phone(s)
- *Limitations and challenges of individual data points*
  - Population coverage uncertainty (phones : people ≠ 1:1)
  - Spatial uncertainty (varying, possibly low spatial resolution)
  - Temporal uncertainty (discontinuous observations)
  - Lack of "purpose" information and sociodemographic variables
  - Data access issues (business, legal)

Past work on MNO data for OS in the ESS

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

(*) with exceptions, e.g. inbound tourism statistis in Slovenia

The vision

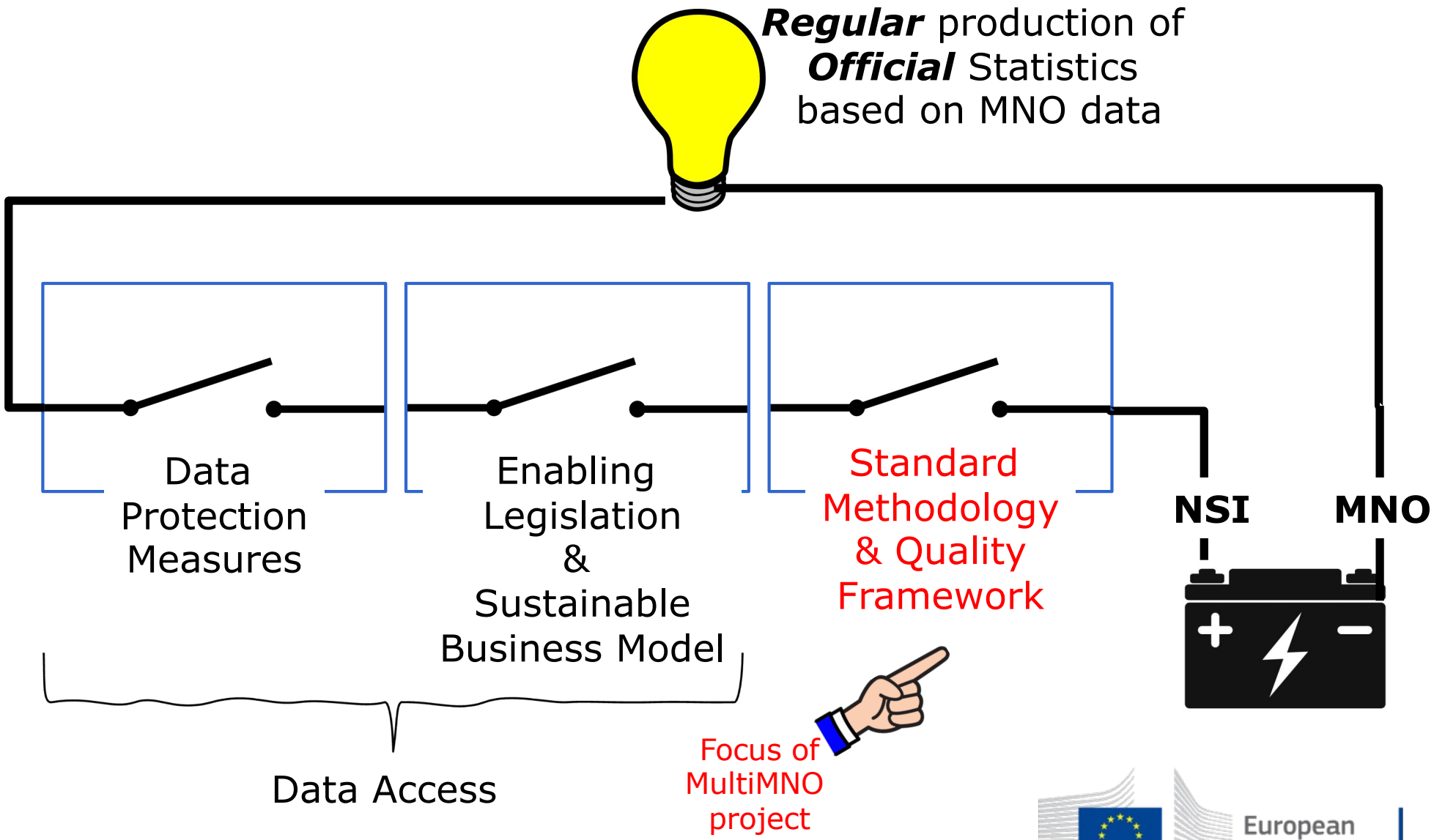


- *In 202x MNO data are (re)used for regular production of* ***official statistics*** *(not merely experimental statistics) …*

Focus of MultiMNO project

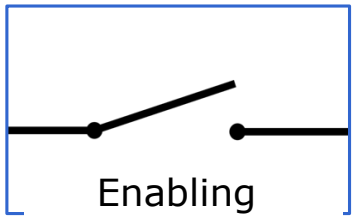
- Based on a sustainable partnership model between ESS members and industry actors (MNOs)
- Based on data from multiple MNOs in each country and across all EU countries
- Processed according to **standard and open methodologies** and **transparent quality criteria** defined at EU level (in collaboration between ESS and industry)
- Incorporating strong technical and organisational measures for agreed at the EU level for protecting personal data (privacy) and business sensitive information
- Combined with other (non-MNO) data for calibration/stabilisation/validation

Series of Challenges



European Commission

Data Access – recent actions



- *Expert Group on facilitating the use of new data sources for official statistics (B2G4S)*
  - Final report with recommendations, June 2022 https://ec.europa.eu/eurostat/en/web/products-statistical-reports/-/ks-ft-22-004
  - Overview of the critical aspects of Privately Held Data, applicable also to MNO data

- *Legislation: EC proposal for amending Regulation N. 223/2009 on European stats*
  - Relevant for all Privately Held Data
  - Target adoption by 2024 (?)
  - NB: Secondary legislation may be still needed afterwards to define operational aspects

Methodology

Standard Methodology & Quality Framework

- ***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

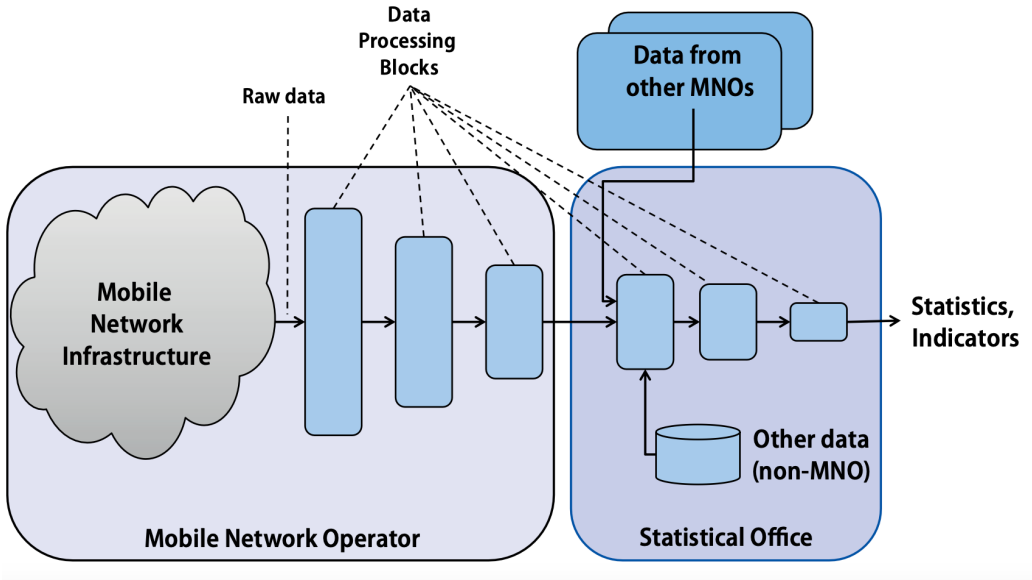
European Commission

Table of contents

European Commission

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!)



European Commission

The case for methodological standardisation

- *A **common and open ESS methodological standard** defined at EU level …*
- *… allows to define a **common ESS quality framework** tailored for MNO data*
- *… provides a reference basis to clarify appropriate **protection measures for personal data** directly at EU level and in consultation with EDPS/EDPB*
  - Overcoming fragmentation of national legislations and heterogeneity of GDPR application at national level

Meta-requirements for a ESS methodological standard

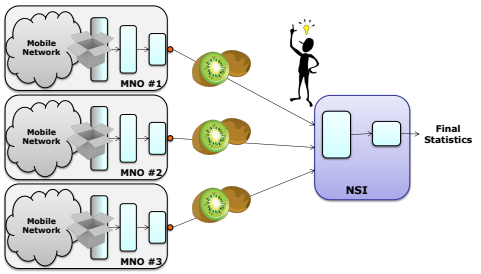
- ***Transparency*** *and* ***reproducibility*** *– open methodologies and* ***open-source reference software implementation***
  - Does not exclude the use in production of commercial proprietary software implementations that are full compliant (functionally aligned) to the open-source reference
  - Test data and functional testing procedure accompanying open-source reference implementation to support compliance checks

- ***Modularity*** *– to handle complexity and for maintenability*

- ***Evolvability*** *– MNO networks evolve, MNO data change, demand changes and new use-cases emerge…*

- ***Quality*** *aspects to be developed jointly with methodology*
  - New challenges for input quality and throughput quality

Multi-MNO orientation



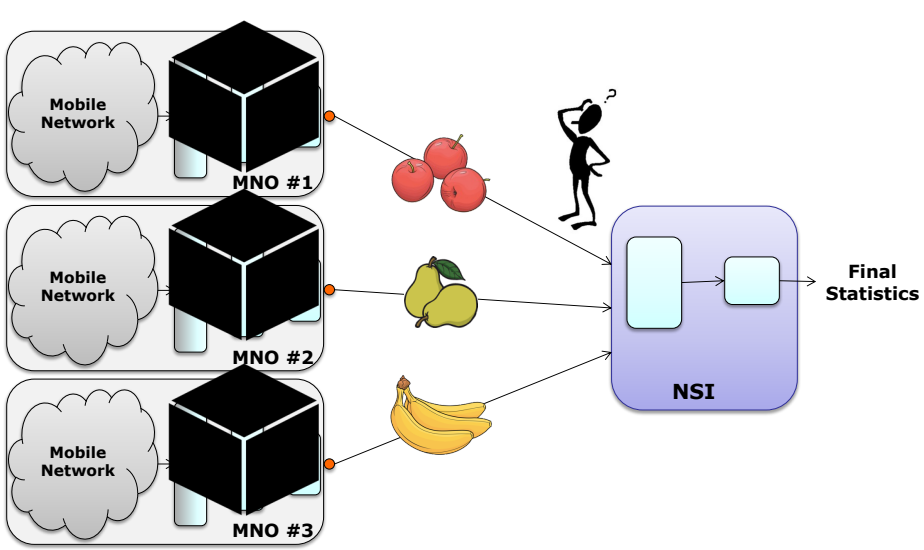
*"The envisioned ESS methodological standard should be designed upfront to produce statistics based on the combination of information **sourced from multiple MNOs** (multi-MNO statistics) within and across different European countries. "*

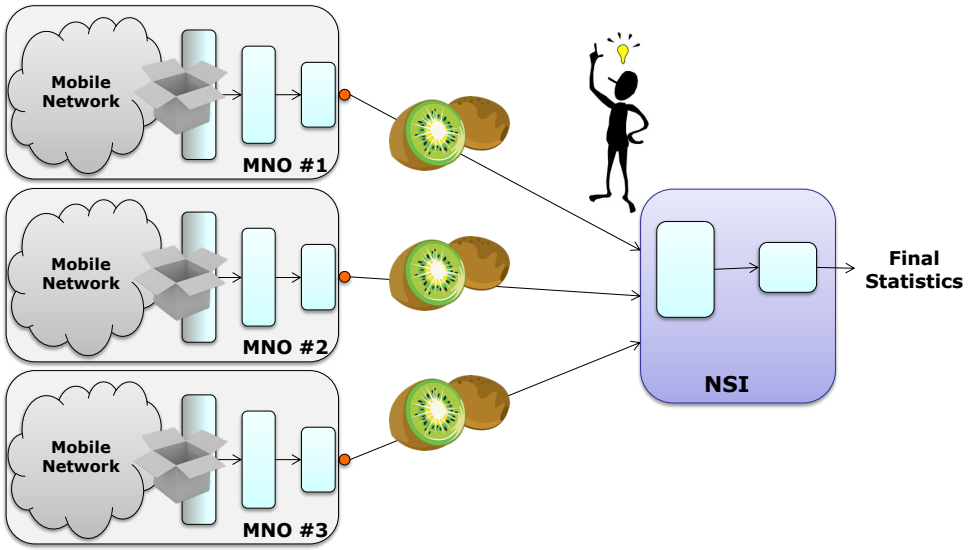
*Advantages of Multi-MNO approach **within each country***

- *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*

The case for methodological standardisation

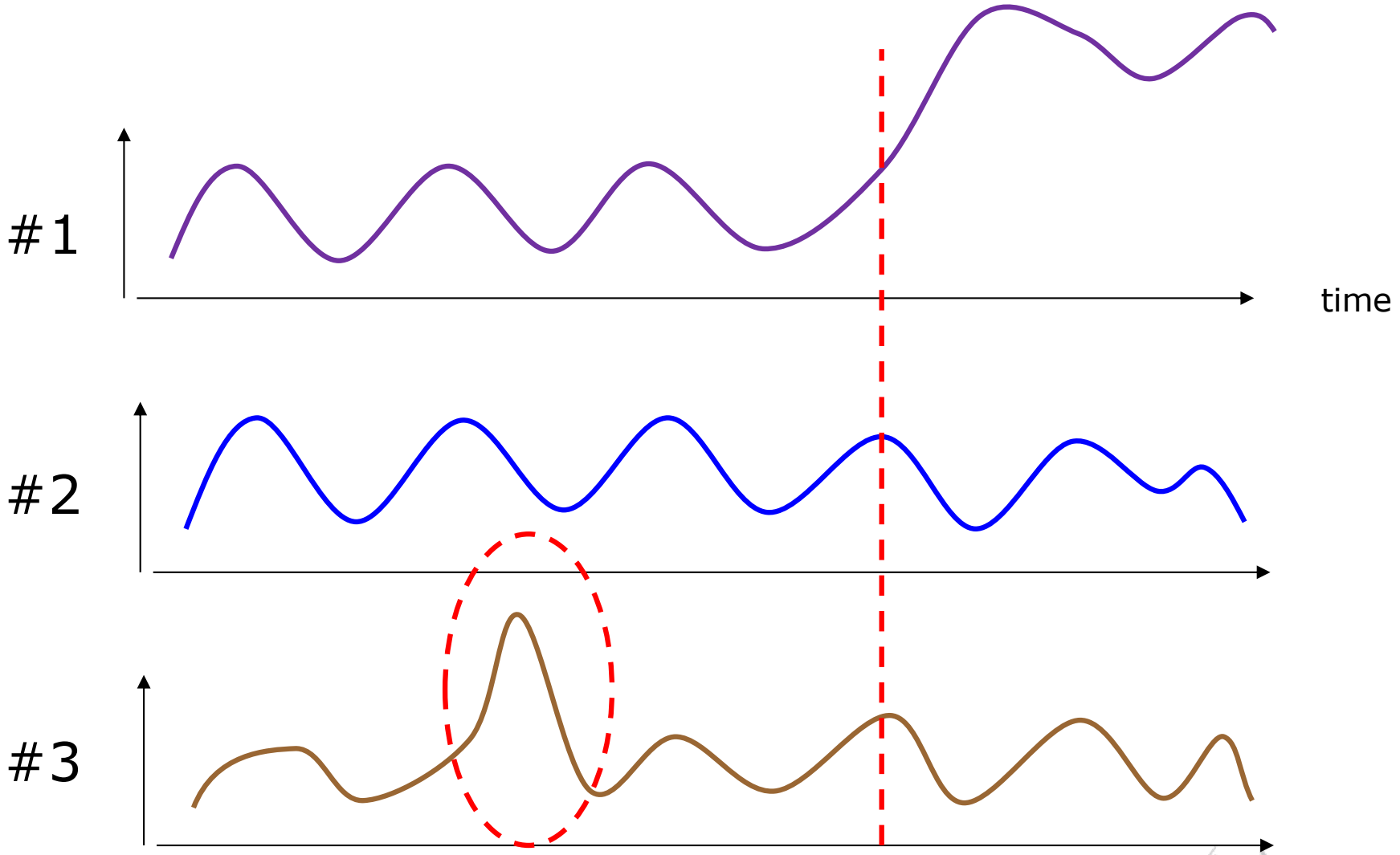
- *Common/standard methodologies and definitions necessary to achieve **comparable** and **combinable** results*





European Commission

Consistency and plausibility checks across data providers



European Commission

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. "*

- *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.*

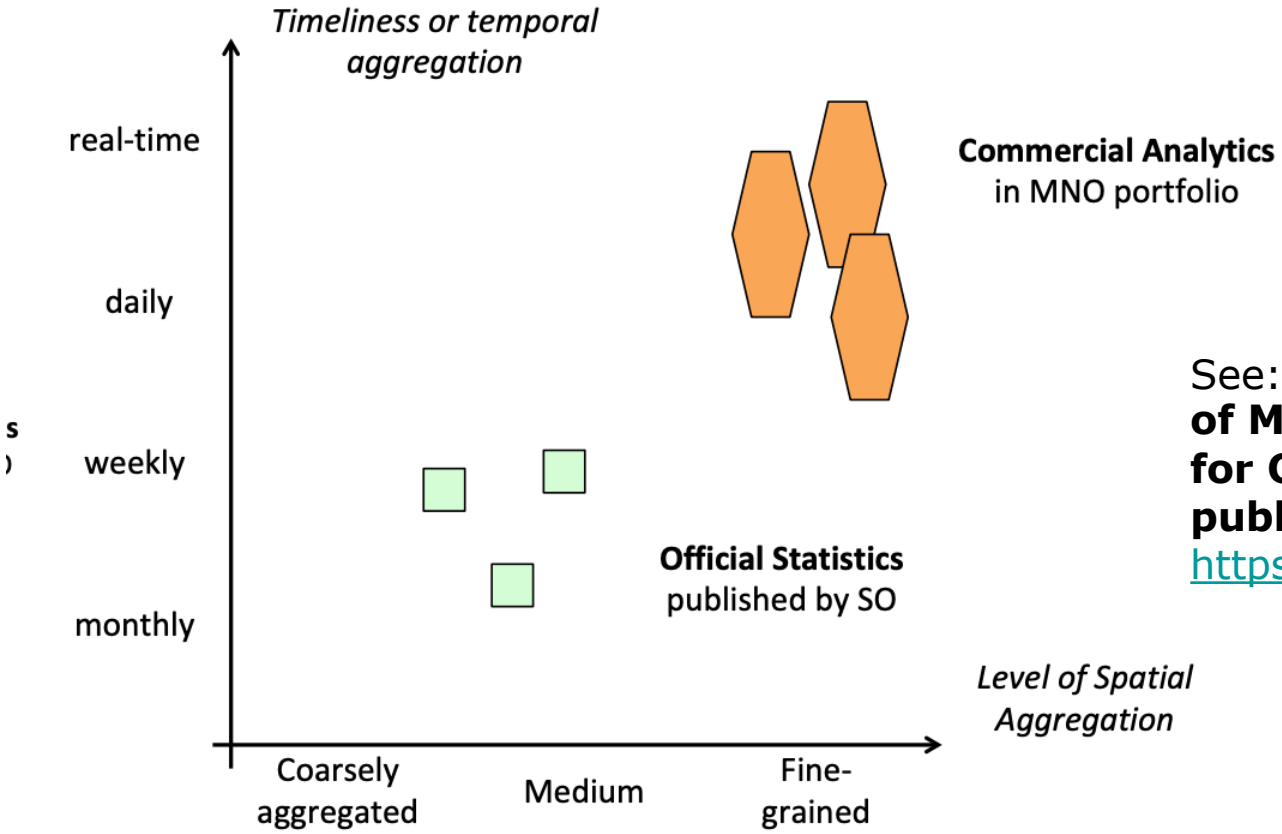
- *Open research aspects: Which non-MNO data are best fit to be combined with MNO data? And how to combine them with MNO data?*
  - New ad-hoc sample survey designed *on purpose* by NSI?
  - Legacy survey data? Census data?
  - Other "big data" sources held by external entities (e.g., train tickets)

- *Strategic implications for MNO-NSI partnerships*

Focus of a separate parallel project: ESSnet Research Grant "MNO-MINDS" involving 10 NSI and coordinated by ISTAT https://cros.ec.europa.eu/book-page/mno-minds

European Commission

Commercial Analytics may coexist with, and even get reinforced by Official Statistics



See: DGINS 2018 paper **Processing of Mobile Network Operator data for Official Statistics: the case for public-private partnerships**
https://zenodo.org/records/10246468

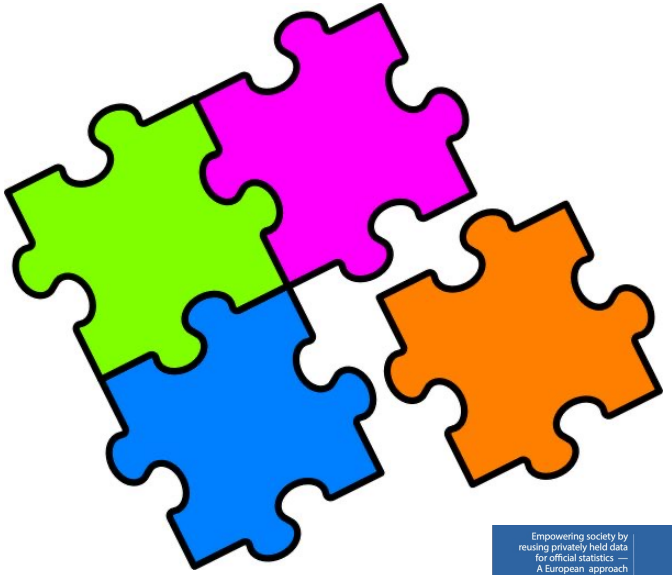
European Commission

Composing the puzzle

ESSnet Big Data II – WPI & WPK

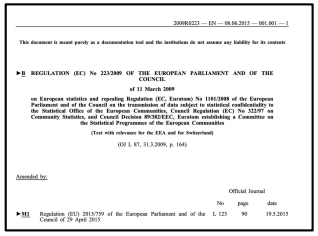
Multi-MNO project

TF MNO



Other projects on Privacy-Enhancing Technologies

Reserch Grant ESSnet MNO-MINDS



Rev. 223/2009 Regulation on EU Statistics

EG B2G4S

European Commission

MultiMNO - Project Factsheet

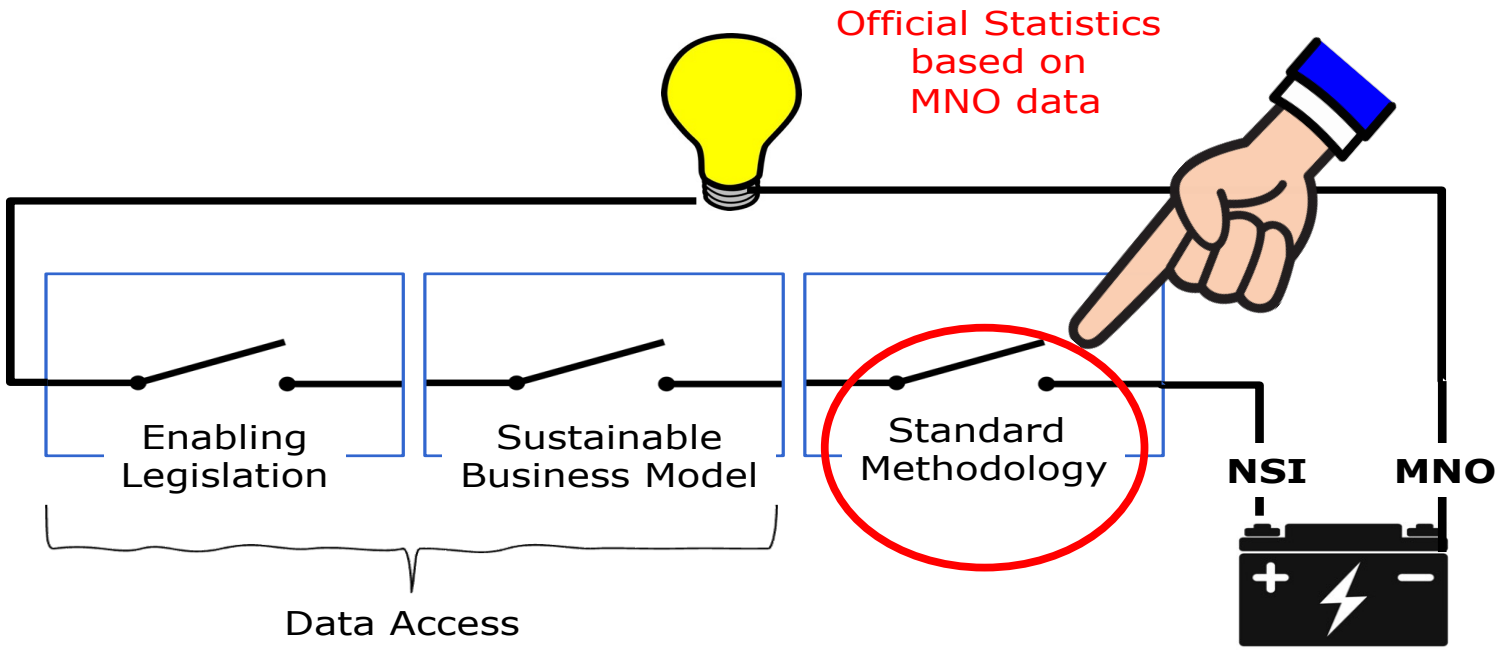
- *Multi-MNO project – service contract ESTAT 2021.0400*
  - following open call for tender https://etendering.ted.europa.eu/cft/cft-display.html?cftId=10149
- *Total duration 2.5 years - termination by mid-2025*

- *Consortium*
  - **GOPA Germany (consortium leader)**
  - **POSITIUM Estonia (consortium partner)**
  - **NOMMON Spain (consortium partner)**
  - **ISTAT Italy (subcontractor)**
  - **CBS Netherland (subcontractor)**
  - **5x MNOs (subcontractors) - ORANGE Spain, VODAFONE Spain, VODAFONE Italy, A1 Slovenia, POST Luxembourg**

- *Advisory Board – composed of individual independent experts acting in their personal capacity*

European Commission

Focus of project is methodology



*Project Goal: develop a **proposal** for a **methodological standard** and associated quality aspects…*

*… taking **plausibly optimistic assumptions** about (future state of) data access: EU legislation, business models, etc.*

- *Not necessarily reflecting the current situation, but anticipating **desirable and plausible** future situation ("Reference Scenario")*

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 (Task 5)

European perspective

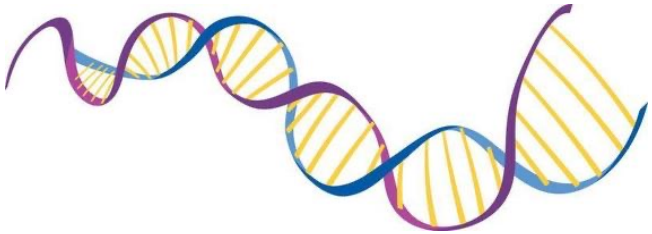
1. *Assuming (prospectivey) comparable data access situation across EU countries (legal, business)*

2. *Developing methodology that may work (prospectivey) 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

European Commission

Evolutionary perspective



- *The methodology defined by this project must be "**evolvable by design**" and forward looking*
  - Think e.g. of "releases" in 3GPP standardisation

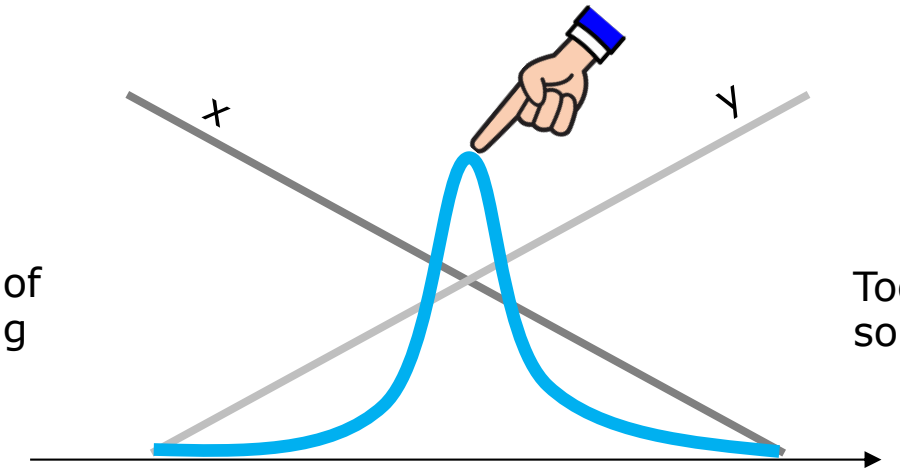
- *Try to anticipate methodological approaches and variants that may be needed or become "optimal" tomorrow*
  - **Today's design should facilitate** future implementation of tomorrow's features and approaches
  - **Avoid introducing implicit constraints** today that are difficult to remove tomorrow

- *Design = balance of goals, costs, constraints etc.*

European Commission

Balancing

- *Methodological design choices must balance across many dimensions, goals, costs …*
- *Balancing between costs today and costs tomorrow; ease of demonstration today and ease of evolution tomorrow;*
- *The best balance point depends on the overall context*

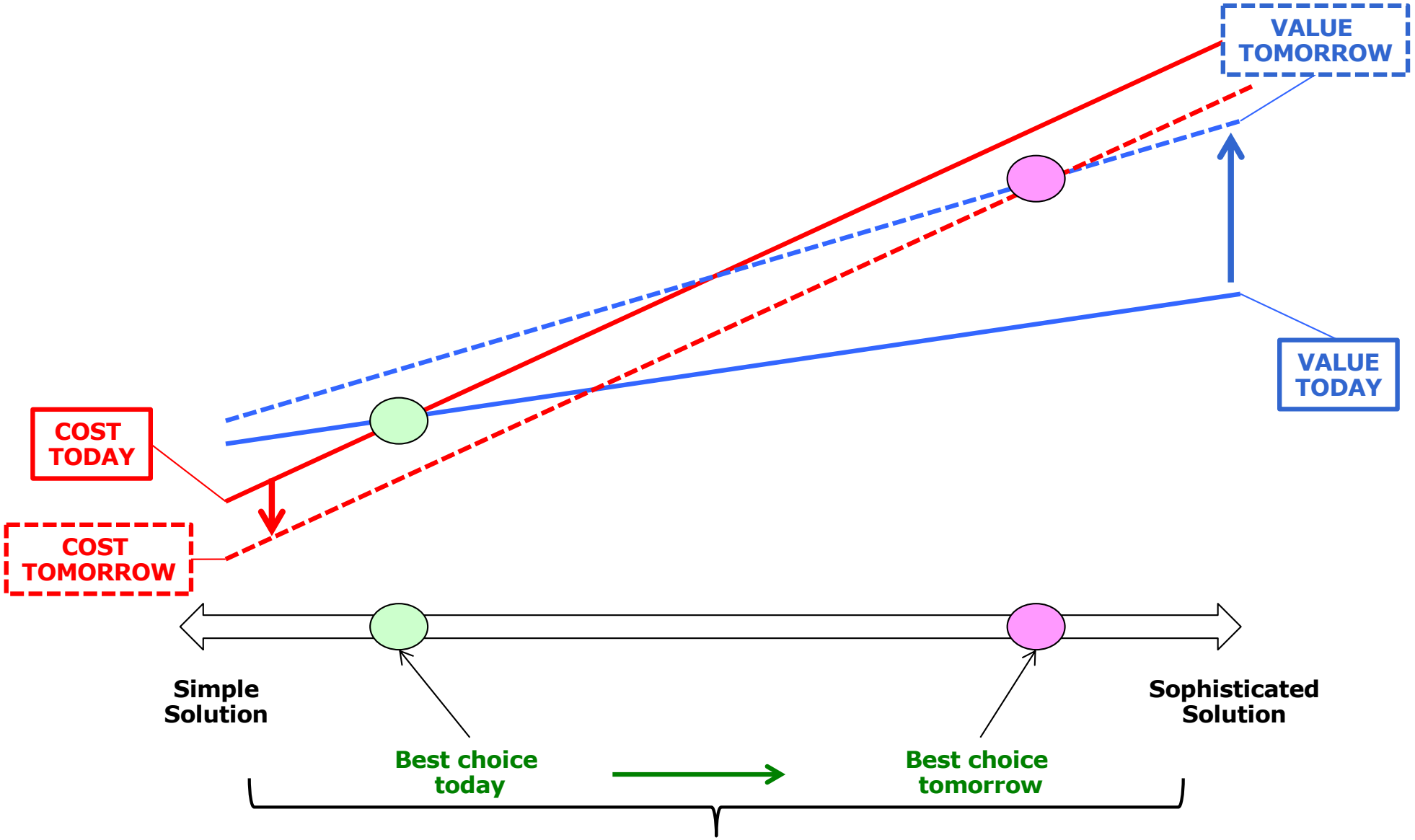


Too little of something

Too much of something

European Commission

Cost vs. value trade-off varies in time



The optimal solution in one ecological niche is not necessarily optimal in another one

- *Building a methodology in a commercial scenario (single MNO, one country, commercial statistics etc.) is a different task from developing a European standard for official statistics*
  - Methodological choices that have worked well in *that* business setting is not necessarily optimal for *this* project

- *The **European methodological standard** we aim for is NOT a replica of one existing commercial solution or the interpolation between two or more of them*
- *… and it will push the state-of-the-art of mobile data processing a few steps further*

European Commission

Note on ownership and market relevance

- *The standard methodology to be delivered by the project will be **owned by the EC** (contractual terms) and will be publicly **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 adopted by private companies and MNOs for commercial statistics*

European Commission

Methodological development and quality guidelines

- *Quality aspects are part of the methodological work*
- *The definition of quality guidelines and criteria (for input data, process and output) is part of the methodological standardisation process (aka "quality framework as inherenet part of the methodological framework")*

European Commission

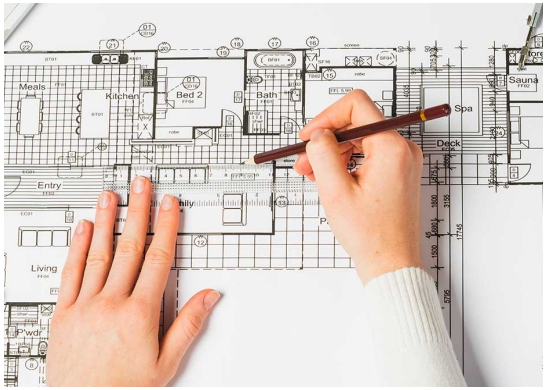
Multi-MNO perspective in the Multi-MNO project

- *Methodology must be applicable to different MNOs in the same country*
- *Methodology must be able to **integrate information** from multiple MNOs in the **same country** and **across countries***
- *Use-cases involving cross-country statistics (for int'l mobility) are extremely important for this project, and a differentiator with respect to all other initiatives*

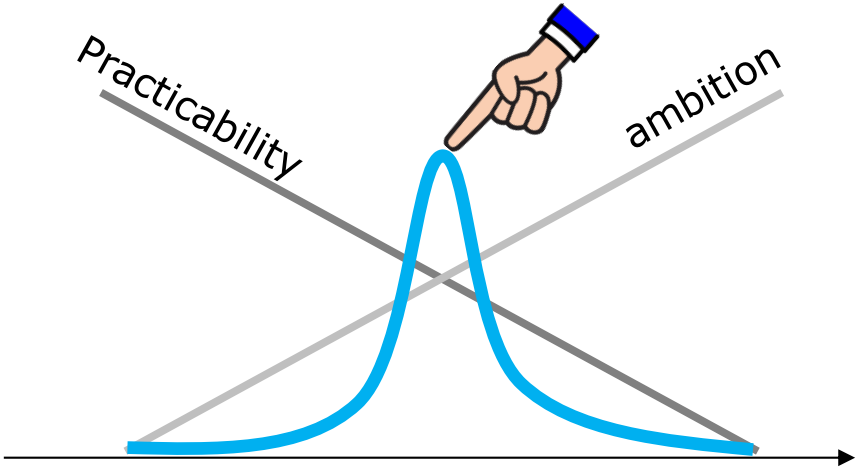
European Commission

Two slogans for this project

We don't guess the future.
We design it.



Realistic ambition



We cannot fly to the moon
but that's not an excuse for crawling

This project …

- *This project is NOT RESEARCH, but does have INNOVATION*
  - Research problems are covered in other projects. This project may help to identify novel problems worth to be addressed by new research work, but the consortium is NOT expected to perform research work
  - Certain specific methodological issues addressed in the scope of this project may require **pragmatically innovative** solution approaches
  - Innovation ≠ Research. This project must embrace innovation.

- *This project is an experiment in itself*
  - First-of-its-kind project where NSI and industry co-develop together a methodology for "big data" from scratch (no legacy constraints)
  - If successful, it will have implications also for other sectors where "big data" sources are involved in official statistics ☺
  - If it fails, it will have implications too ☹

Defining success

- *Huge potential but also high risk*
  - This project is ambitious, complex and challenging – but not ovewhelmingly so. It can and hopefully will **succeed**, but it can also **fail**.

- *Success = the consortium delivers a methodological framework that is technically sound and meets the objectives and technical requirements stated in the call*
  - **Accepted** by Eurostat and eventually **endorsed** as standard by ESS*
- *Failure = the delivered methodology falls short of expectations, leading to no endorsement by ESS.*

Either we get a **strong and convincing standard**
or there will be **no standard** at all.
A weak standard is not an option here.

* Through the Task Force on MNO data for Official Statistics

European Commission

How the Advisory Board can help

- *Advice consortium members (and Eurostat as Contracting Authority) about any **points for improvements** related to any aspects of the **project content**, depending on your specific area of expertise, that may help the project achieving its goals **given the context and motivations outlined above***

- *Help identifying choices and assumptions that need to be better explained, clarified or motivated...*
- *Warn about aspects that, according to your opinion and expertise, were not addressed or not addressed in the best possible way, and should be considered or reconsidered …*
- *But also confirm positive things ☺*
- *Help improve orientation and impact of the project*

Thanks for your attention!

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