



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

Fabio Ricciato, EUROSTAT Big Data Task Force
Freddy De Meersman, Proximus

DGINS 2018

Bucharest, 10-11 October 2018



The views expressed in this presentation 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.

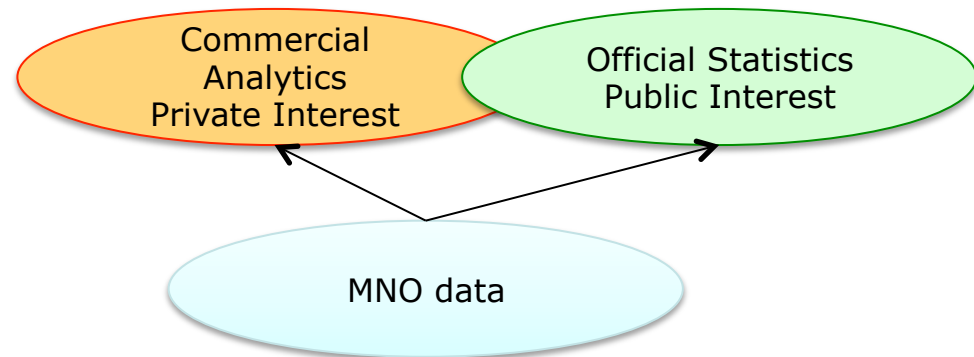


Opportunities

*MNO data embed information about human mobility
(where you are, where you go)*

*Information about human mobility is relevant for Official Statistics:
present population, tourism flows, etc.*

*"Analytics" services and products based on MNO data increasingly seen
by MNOs as additional business branch*



MNO : mobile network operators

Questions

What relationship between MNO and Statistical Offices (SO)?

- *Customer-provider?*
- ***Partnership?***
- *...*

How?

*A Reference Architecture &
Methodological Framework for
MNO data processing for
Official Statistics*

Why?

*Overview of MNO-SO
partnership gains*

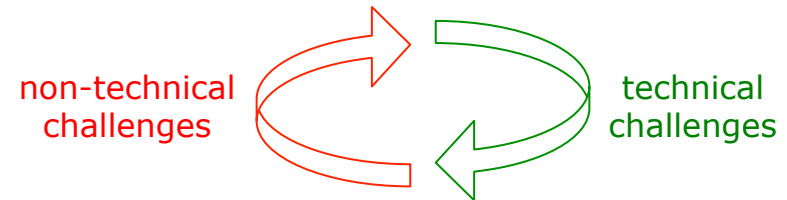
SO : Statistical Offices

MNO : mobile network operators





Our Goals



Develop a unified methodological view:

Reference Architecture & Methodological Framework [RAMF] for processing MNO data for Official Statistics

in order to:

- *facilitate interworking MNO-ESS at technical & organisational level*
- *ensure consistency, reproducibility, **evolvability** and portability of processing methods (between MNOs and SOs)*
- *provide concrete basis to clarify legal aspects (→ GDPR)*
- *enable multi-MNO analysis (fusion of data from different MNO)*





Design Principles

Processing methods *design* based on layered structure, hourglass model, uniformed data semantic, *see next slide*

Processing methods fully transparent to (possibly co-developed by) MNO and SO

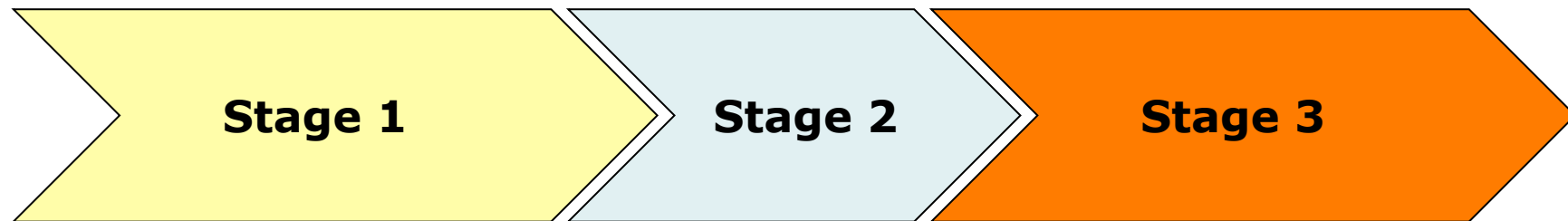
Ref. other presentation in the afternoon session

Processing *execution*: exchange computation, not input data!

...



Defining a Reference Architecture & Methodological Framework: staged approach



Single MNO

- definition of a reference layered architecture (hourglass model) and common data structures (C-layer)
- clarification of GDPR aspects

Multiple MNOs with *output* data fusion on NSI (silos model)

- testing and possible refinement of reference layered architecture across heterogeneous network operation settings

Multiple MNOs with *input* data fusion (via SMPC)

- definition of reference architecture for Secure Multi-Party Computation
- clarification of GDPR aspects related to SMPC

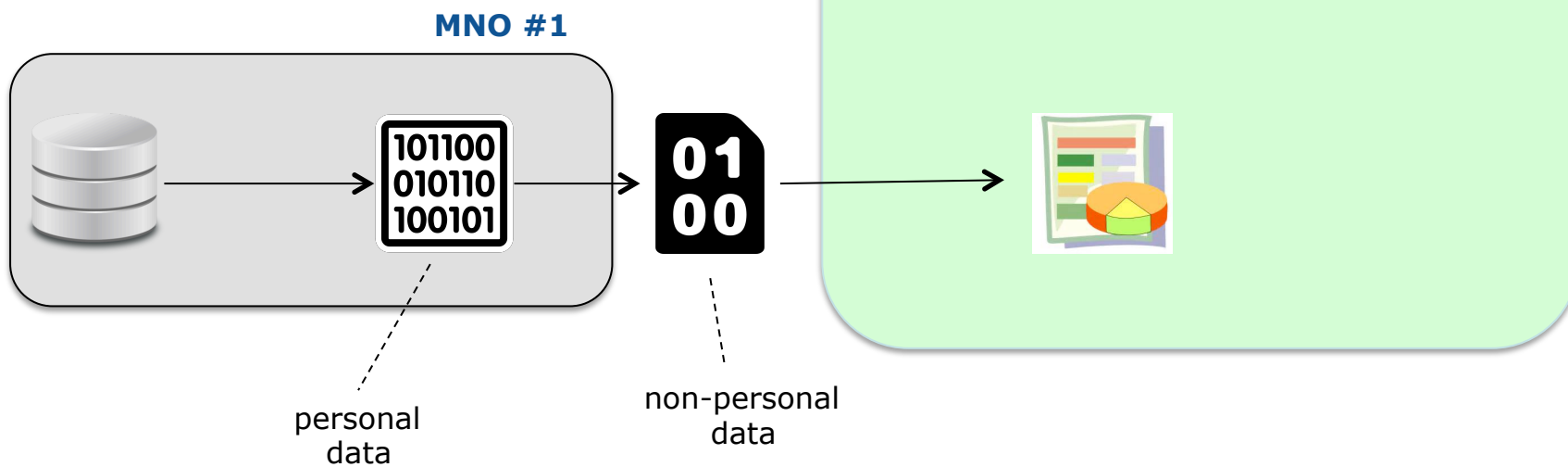
Stage 1 scenario

Raw
micro-data
(D-layer)

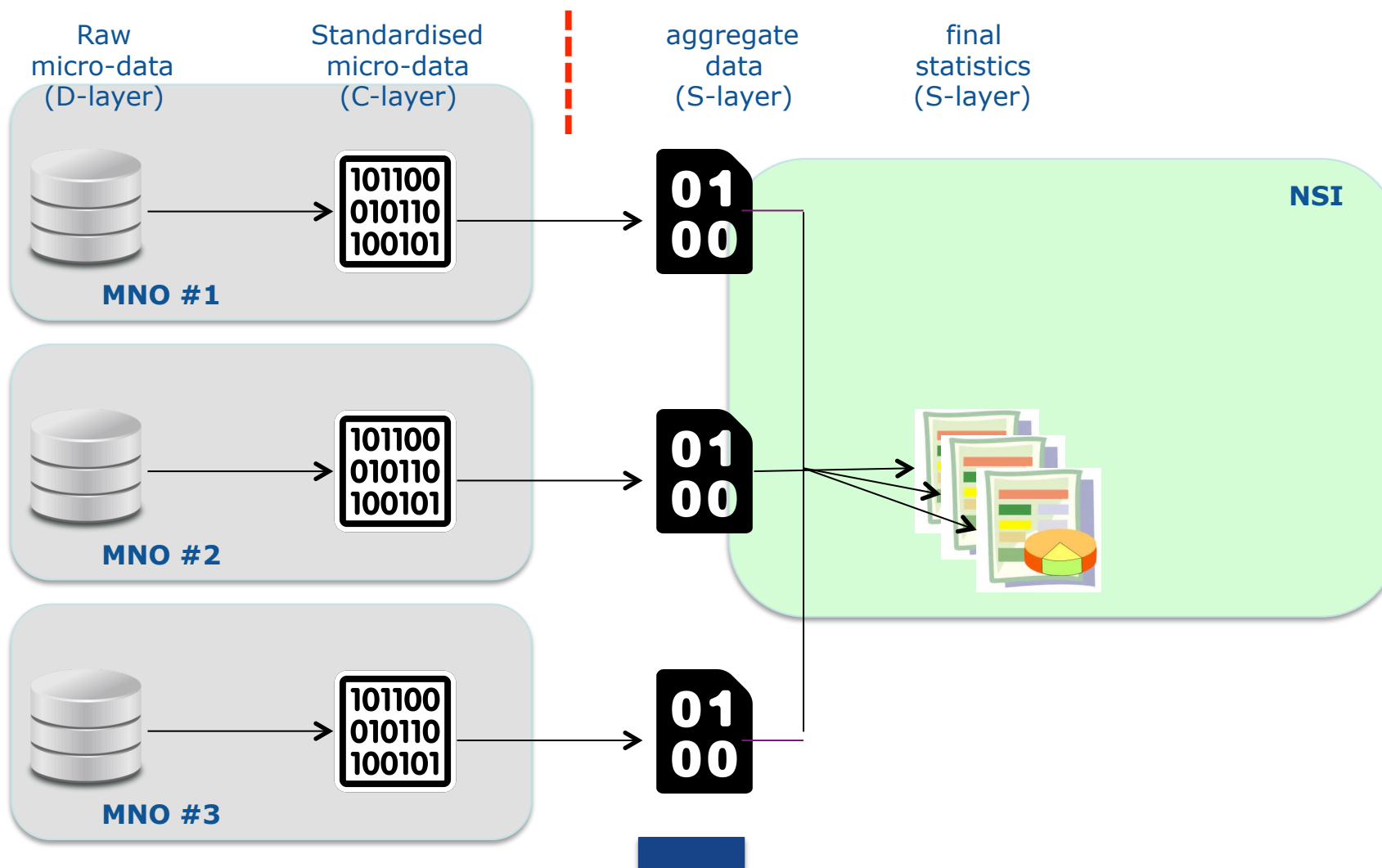
Standardised
micro-data
(C-layer)

aggregate
data
(S-layer)

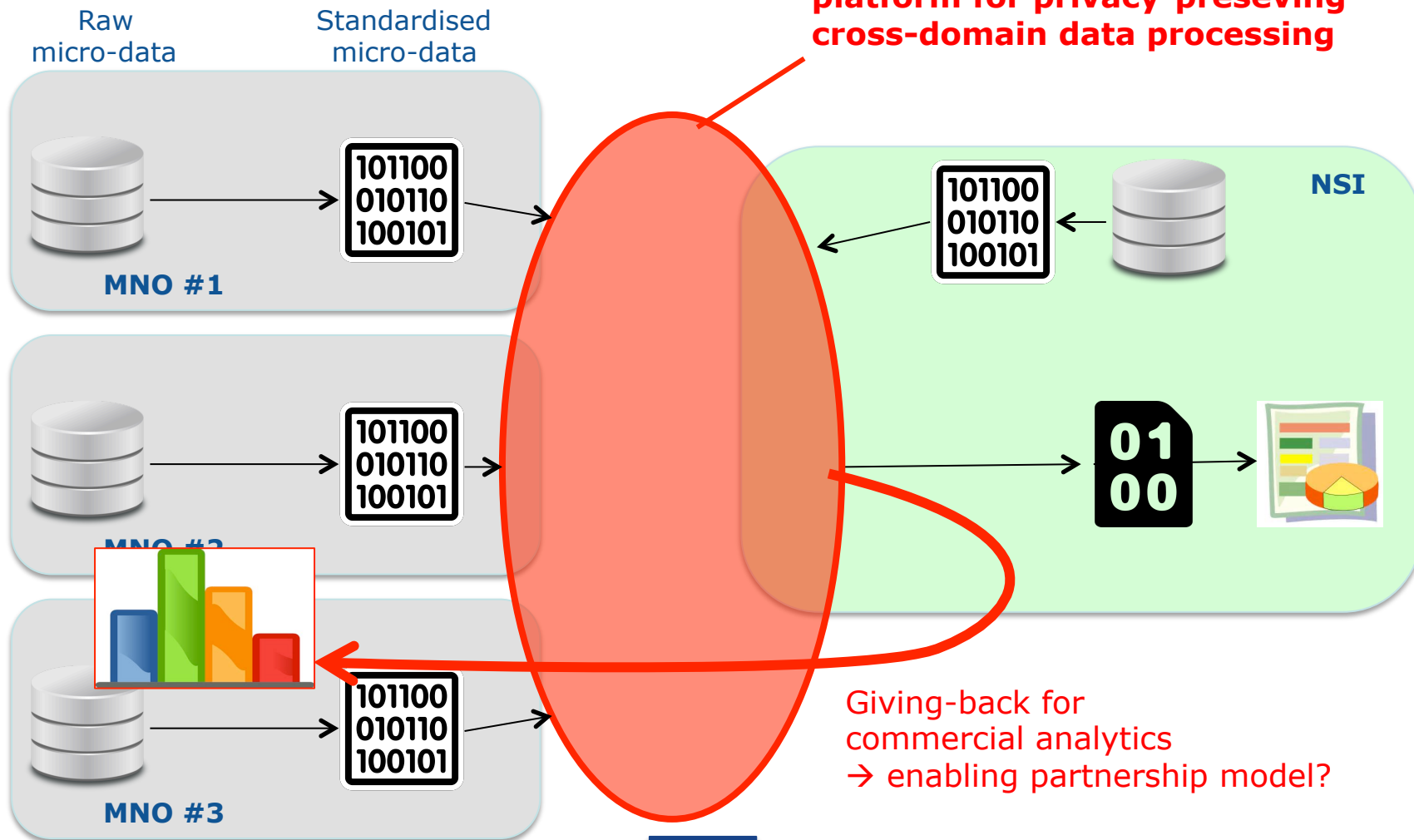
final
statistics
(S-layer)



Stage 2 scenario

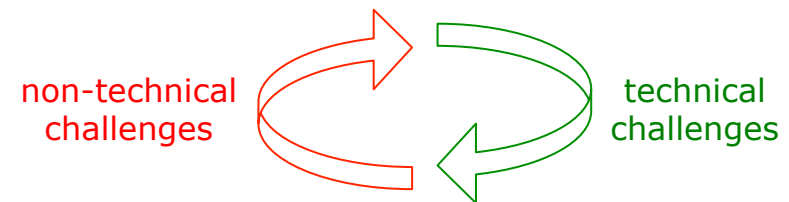
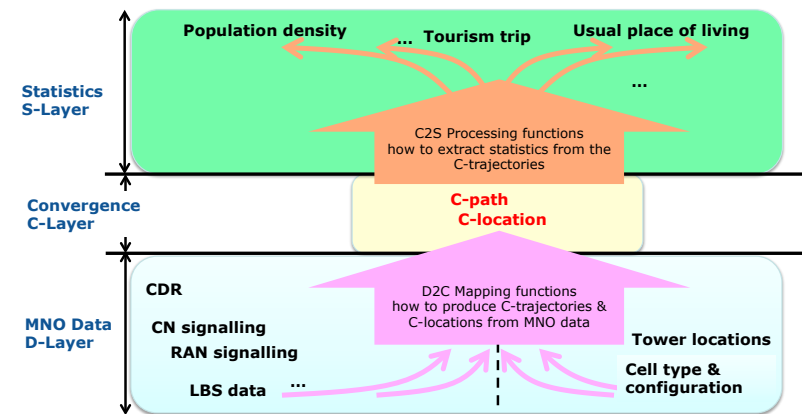


Stage 3 scenario



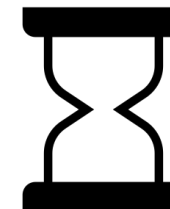
Stage 1 goals

- *Define the Reference Architecture & Methodological Framework for a single MNO data stream*
- *Proof-of-concept application on selected use-case*
 - population density (ongoing work 2018)
 - tourism (next year)
- *Clarify GDPR aspects*
 - started dialogue with European Data Protection Supervisor



- **Collaboration EUROSTAT-Proximus**
- **Dedicated WP in future ESSnet on Trusted Smart Statistics**

Hourglass model



Statistics S-Layer

Heterogeneity of applications & use-cases
Diversity of statistical definitions
Complexity of statistical objects
Multiple NSIs

Domain of Expertise
Statisticians, NSI

Convergence C-layer

**Few common
definitions**

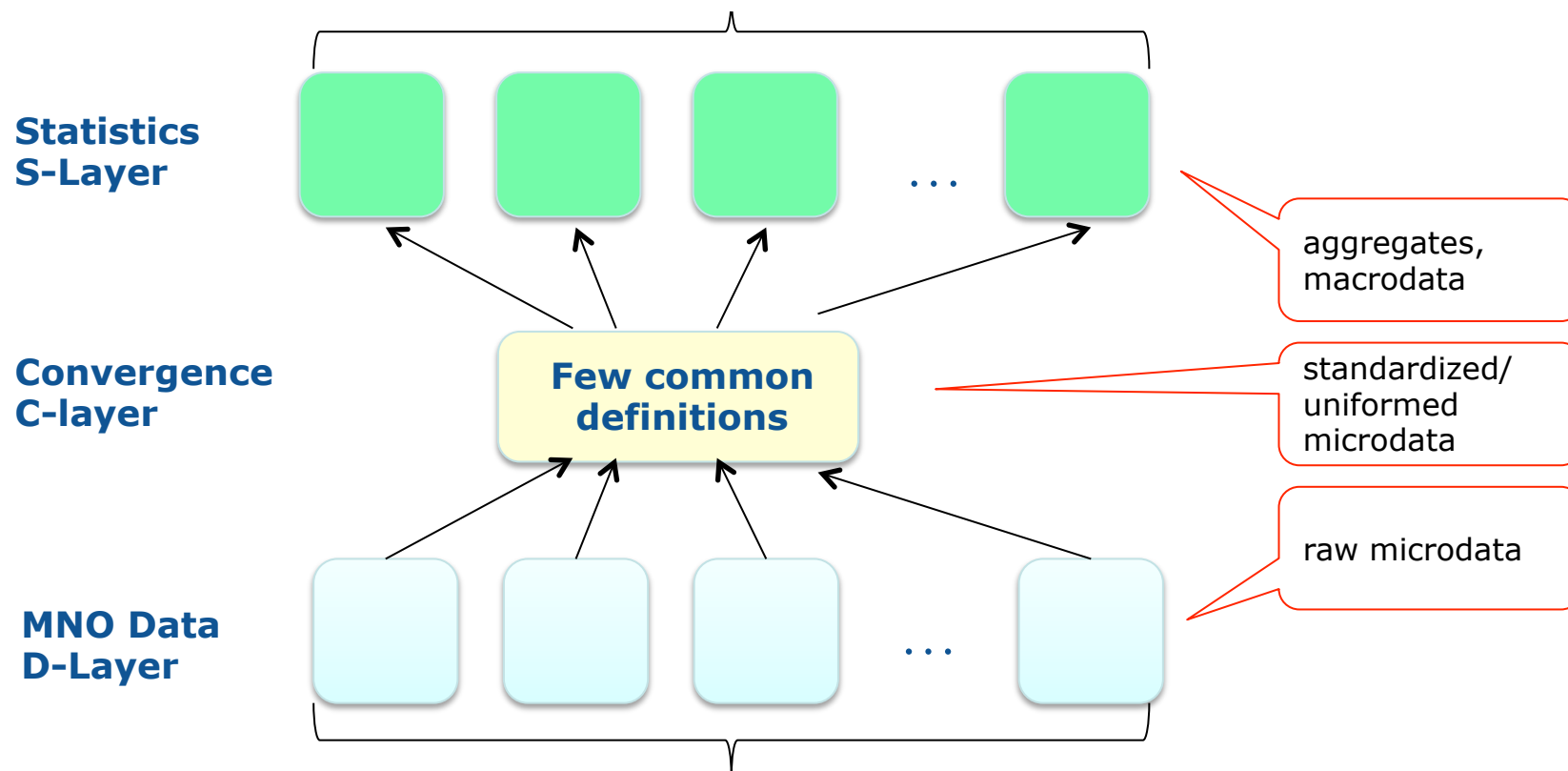
MNO Data D-Layer

Data Heterogeneity
Diversity of data collection methods
Complexity of data semantics
Multiple MNOs

Domain of Expertise
Telco Engineers, MNO



Multiple data consumers: ESTAT, NSI#1, NSI#2...
Different subject matter experts & use-cases:
tourism, population, transport, ...



Multiple data sources: MNO#1, MNO#2...
Different data types: CDR, signalling data, RAN data, LBS, ...



Benefits of layering

Decouples the two domains

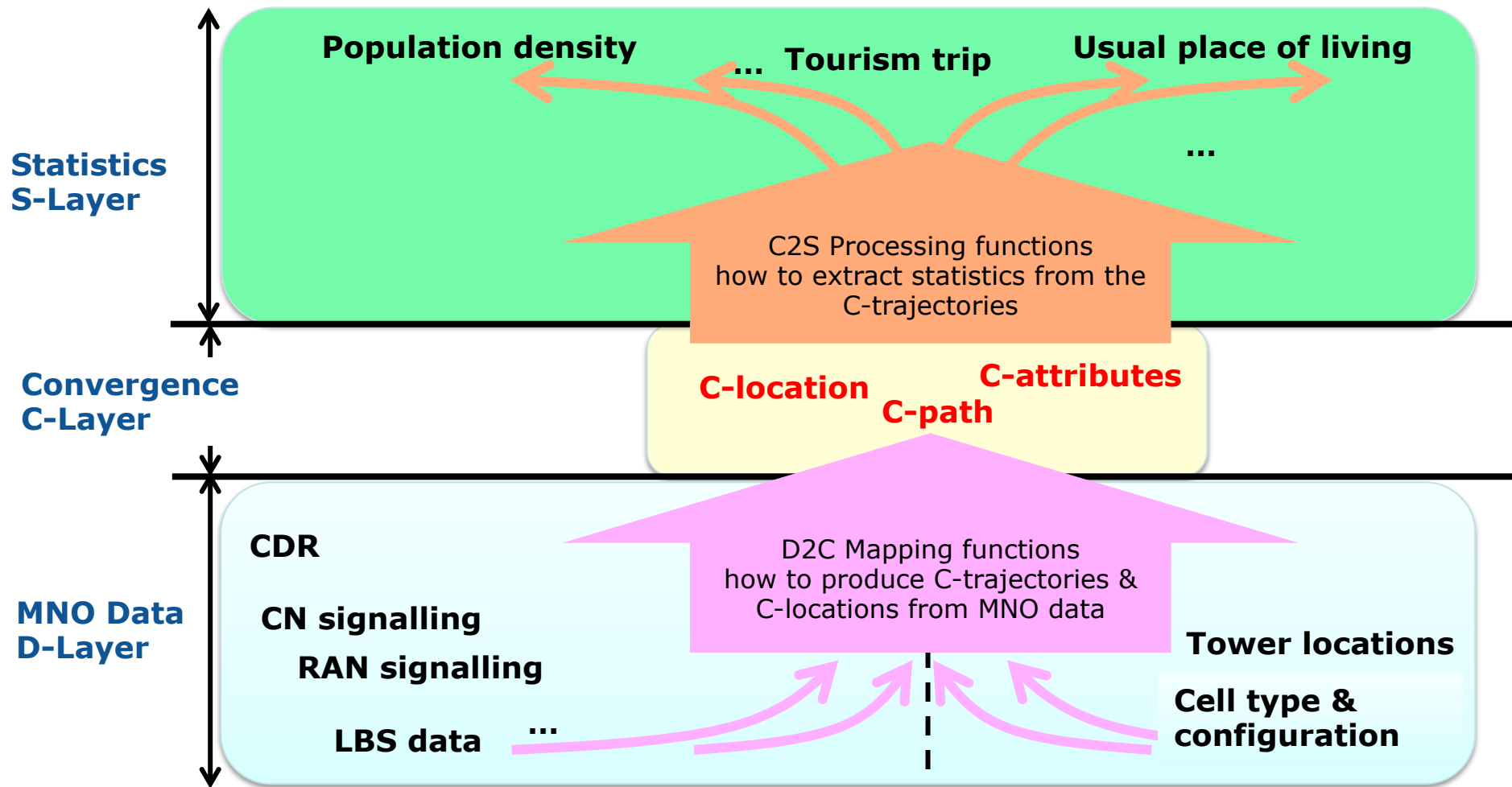
- Hides complexity & heterogeneity of MNO data to statisticians
- Hides complexity & heterogeneity of statistical concepts to telco engineers

*Decoupling enables independent **development, adoption & evolution** at each domain*

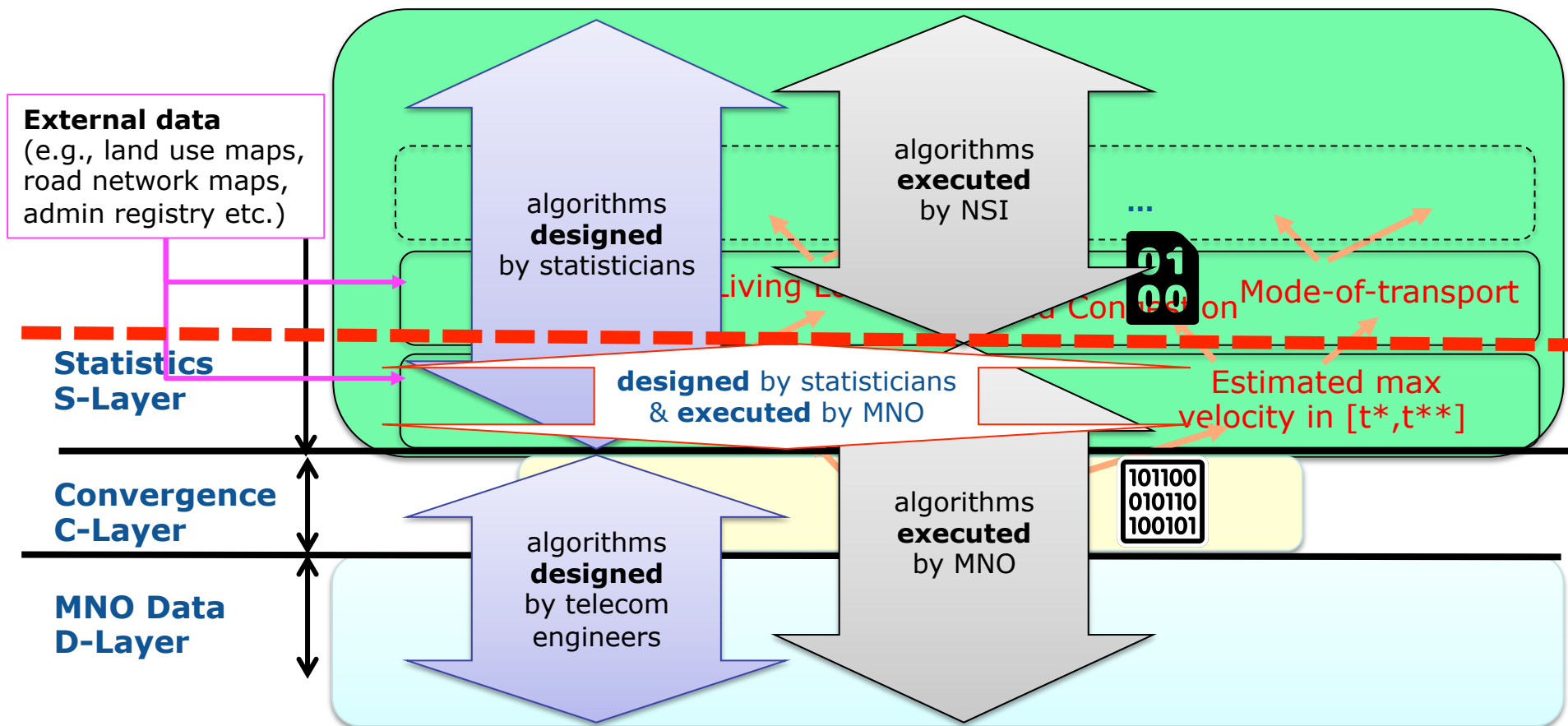
*The C-layer is abstract "knowledge interface" between domains
→ relevant at **design** stage of processing methodology*

*Within the S-layer is the physical interface for data export
→ relevant at **execution** stage*





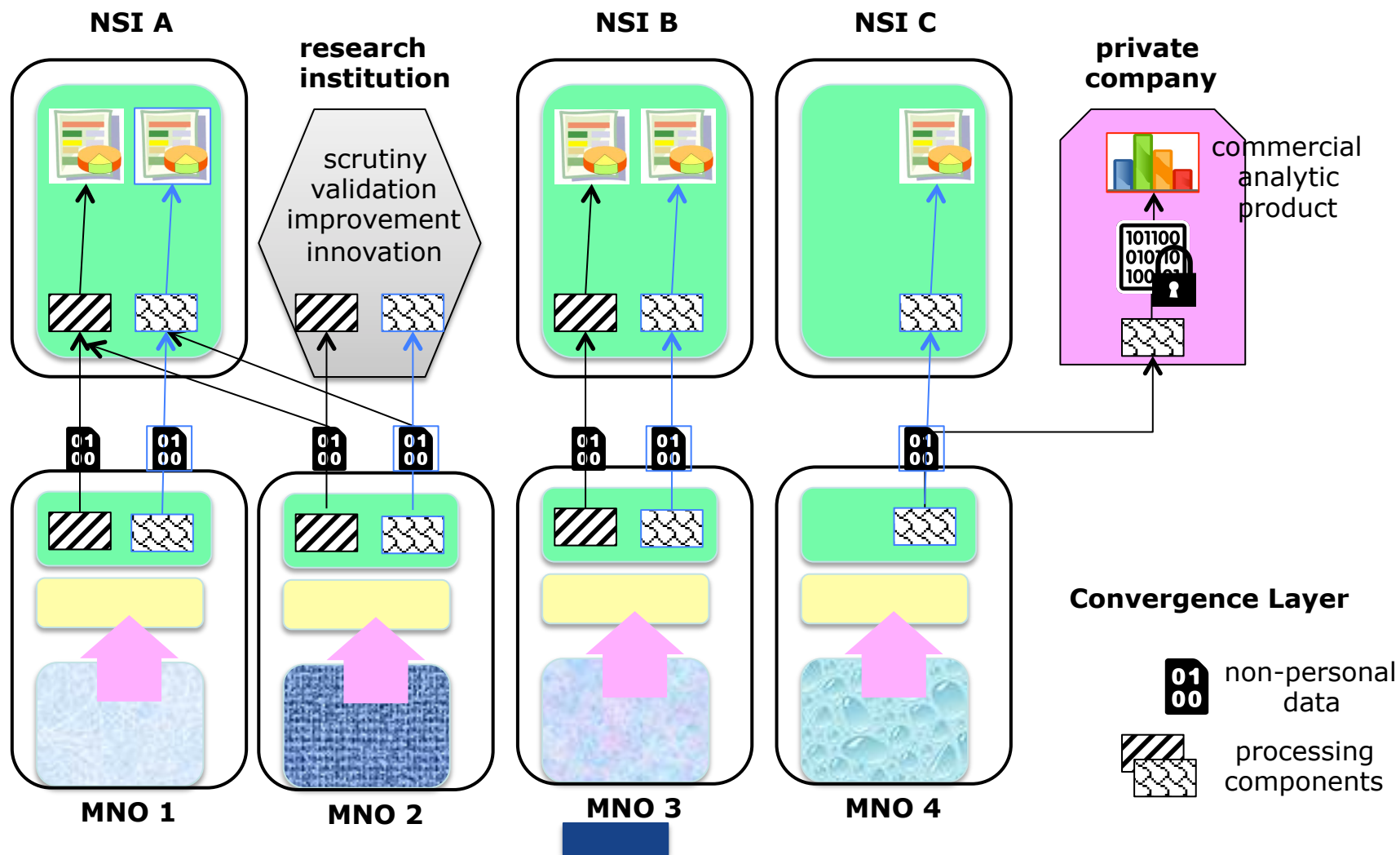
Processing method (algorithm) design vs execution





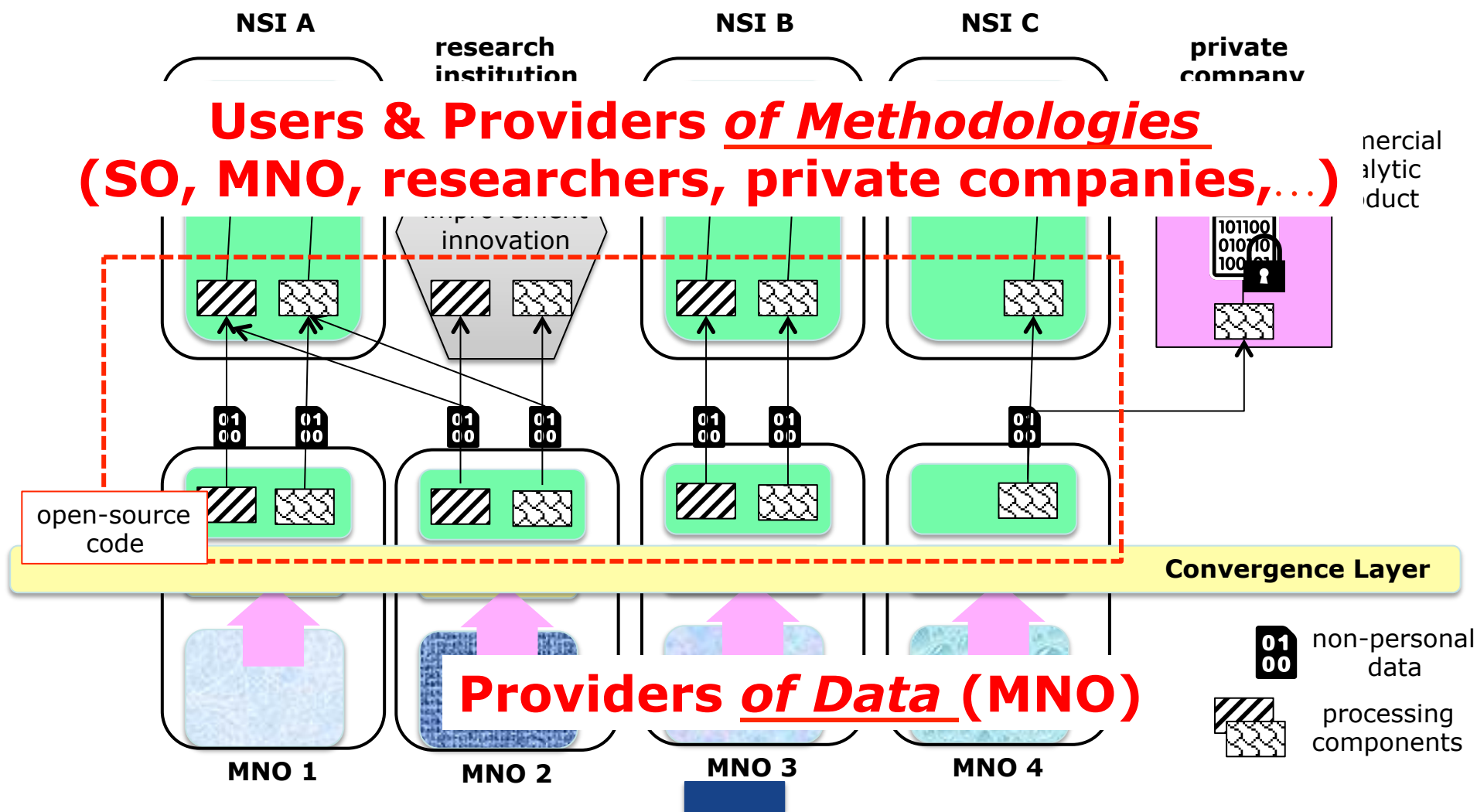
European
Commission

C-layer as a common substratum for MNO data users





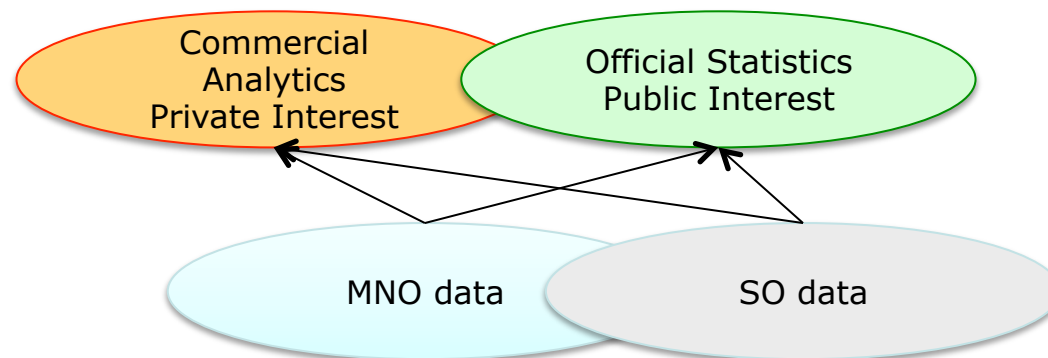
C-layer as a common substratum for MNO data users



Partnership gains for MNO 1/3

1. Access to additional **information** held by SO

- Additional dimensions in SO micro-data
- SO data as “ground truth” for **calibration**





Partnership gains for MNO 2/3

1. Access to additional **information** held by SO

- Additional dimensions in SO micro-data
- SO data as "ground truth" for **calibration**

2. Access to statistical **knowledge**

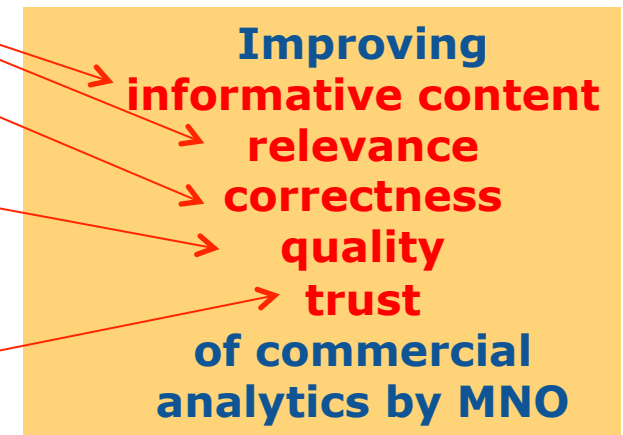
- SO experts complementing MNO experts (telco engineers, data scientists)

3. Inherit **reputation**

- To business customers for commercial analytics
- To the public – SO working for the public interest

4. Stimulate the **market**

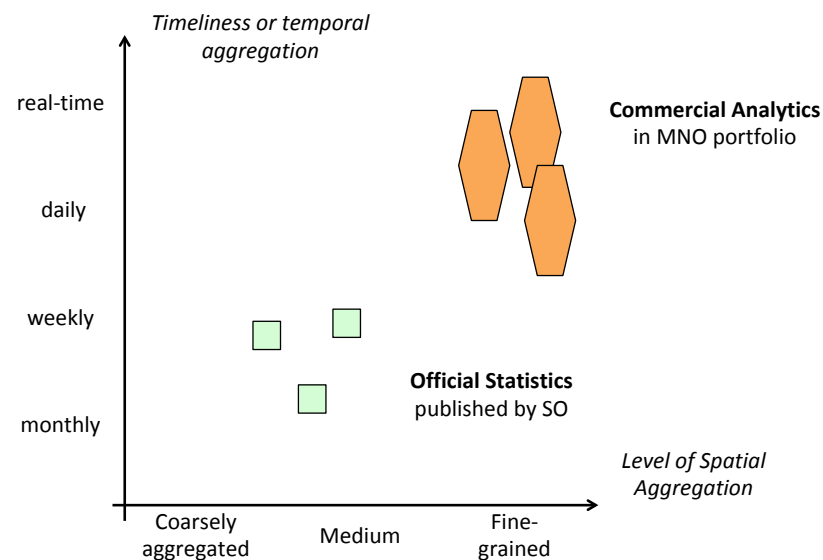
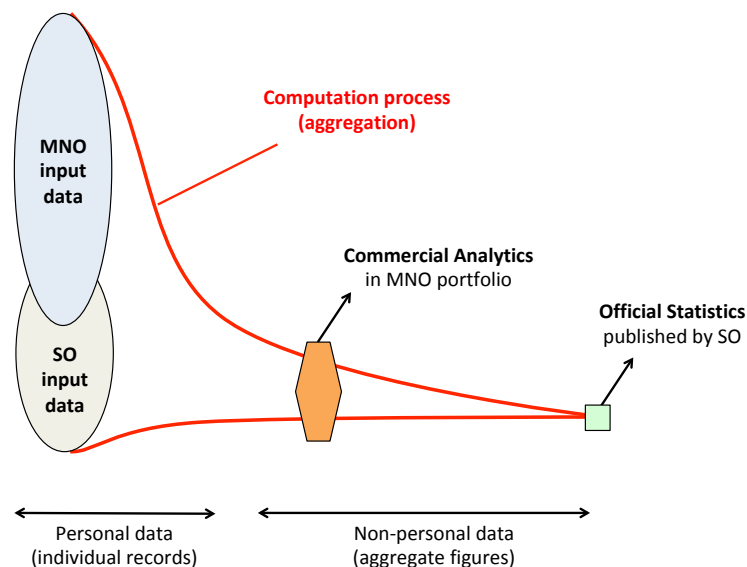
- Like in "**freemium**" models: public official statistics as "basic version" of more detailed, fine-grained, timely delivered commercial analytics



Partnership gains for MNO 3/3

4. Stimulate the **market**

- Like in “**freemium**” models: public official statistics as “basic version” of more detailed, fine-grained, timely delivered commercial analytics





Outlook

Ongoing collaboration between Eurostat and Proximus

- on the definition of methodological aspects (Reference Architecture and Methodological Framework)
- on the identification of concrete use-case for SO-MNO partnerships

Coordinated work with new ESSnet on Trusted Smart Statistis

Seeking to involve other MNOs (also via GSMA, ETIS)





Thanks for your attention

For follow-up:

freddy.demeersman@proximus.com

fabio.ricciato@ec.europa.eu

