A reflection on the potential role of MultiParty Computation for the production of (future) Official Statistics

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

ENISA Workshop on Personal Data Sharing – Emerging Technologies
7. October 2022

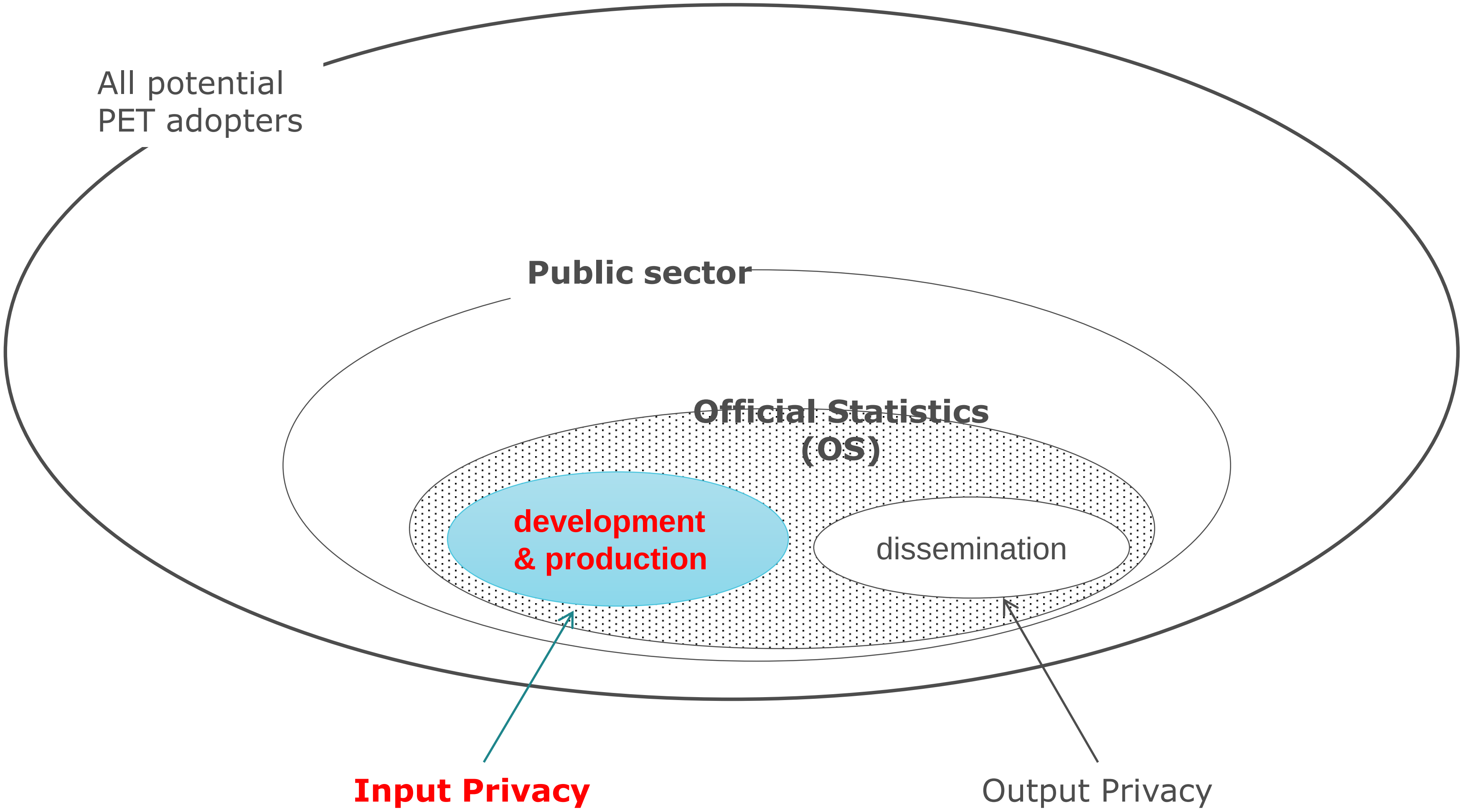
Goal of this talk

Offer a reflection on the potential role of MPC in Official Statistics from the perspective of potential adopters of MPC technologies

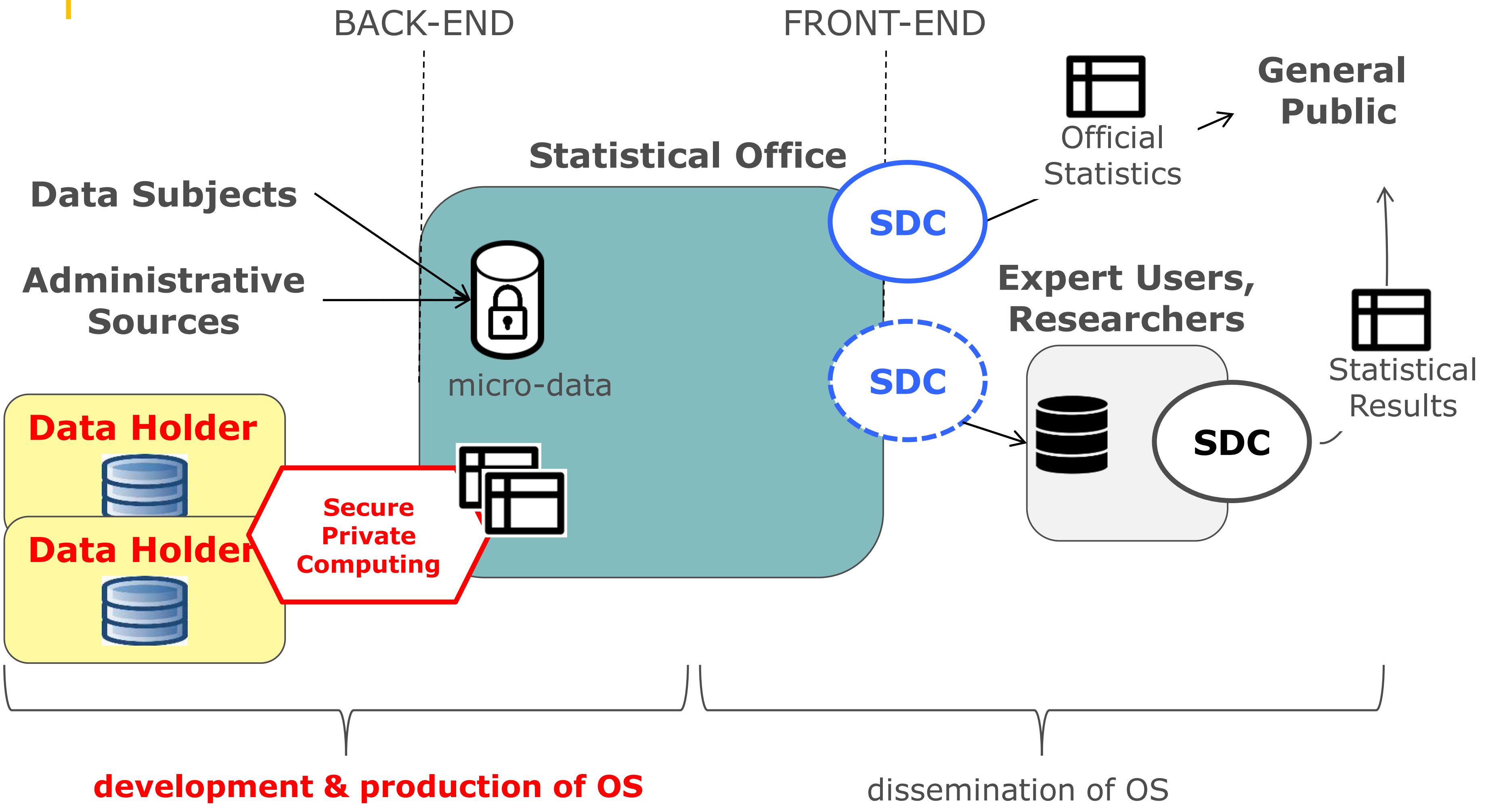
Caveat

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.

Scoping this talk



Scoping this talk

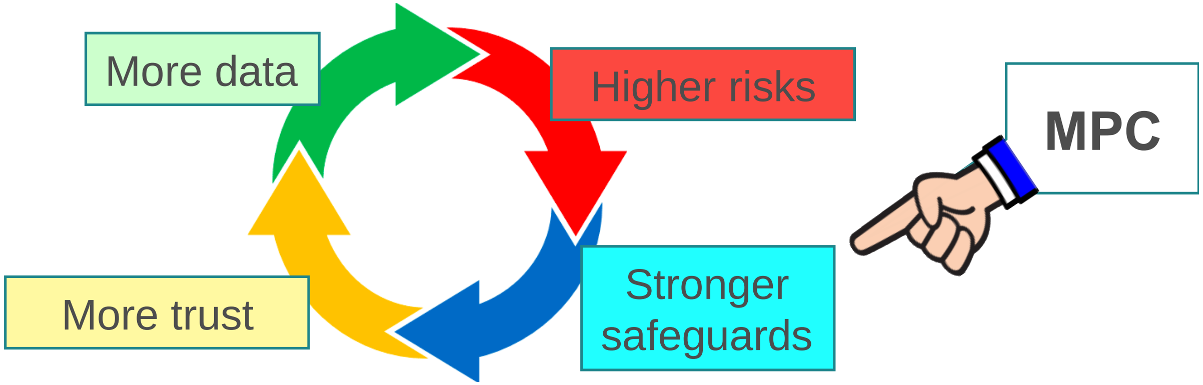


SDC Statistical Disclosure Control (output privacy)
MPC Multi-Party Computation (input privacy)

Why do we care?

- Increasing appetite for cross-organisational data processing in the context of Official Statistics innovation
  - Data held by national authorities in different countries concerning cross-border phenomena (e.g., int’l trade, migration, …)
  - Statistics based on data held by other public bodies (e.g., administrative data)
  - New statistics based on privately held data requiring integration across different providers (often competitors in the same business sector) and with data held by statistical authorities

- Increasing awareness of the importance of (personal) data protection by the general public



More data

Higher risks

Stronger safeguards

More trust

MPC

The function to be computed is known and declared

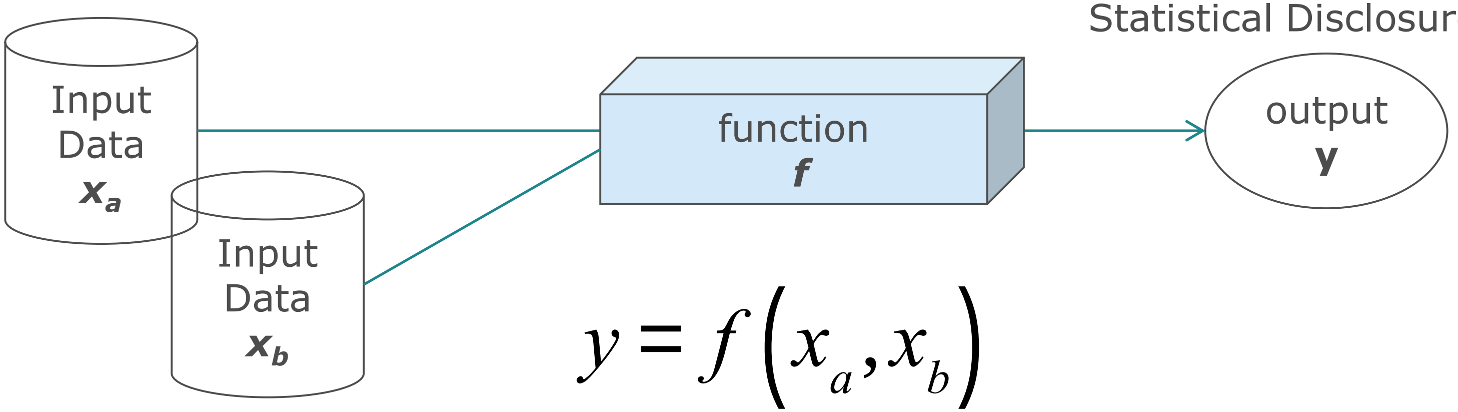
- *f* = statistical methodology
- *f* not a business secret (on the contrary it should be open for the sake of methodological transparency)
- f is typically “simple” (no highly-dimensional ML/AI models, but rather low-dimensional regressions…)

The input parties are mixed: statistical offices, public bodies, private data holders

- eg. 2 NSI in different countries
- e.g. 3 private data holders in the same country
- 1 public body + 1 NSI
- any combination …

The output party is a statistical office

- has a legal basis to receive the exact result, even if it contains personal information
- the result will not be published (disseminated) without further Statistical Disclosure Control checks



$$y = f\left(x_a, x_b\right)$$

NSI National Statistical Institute
MNO Mobile Network Operator

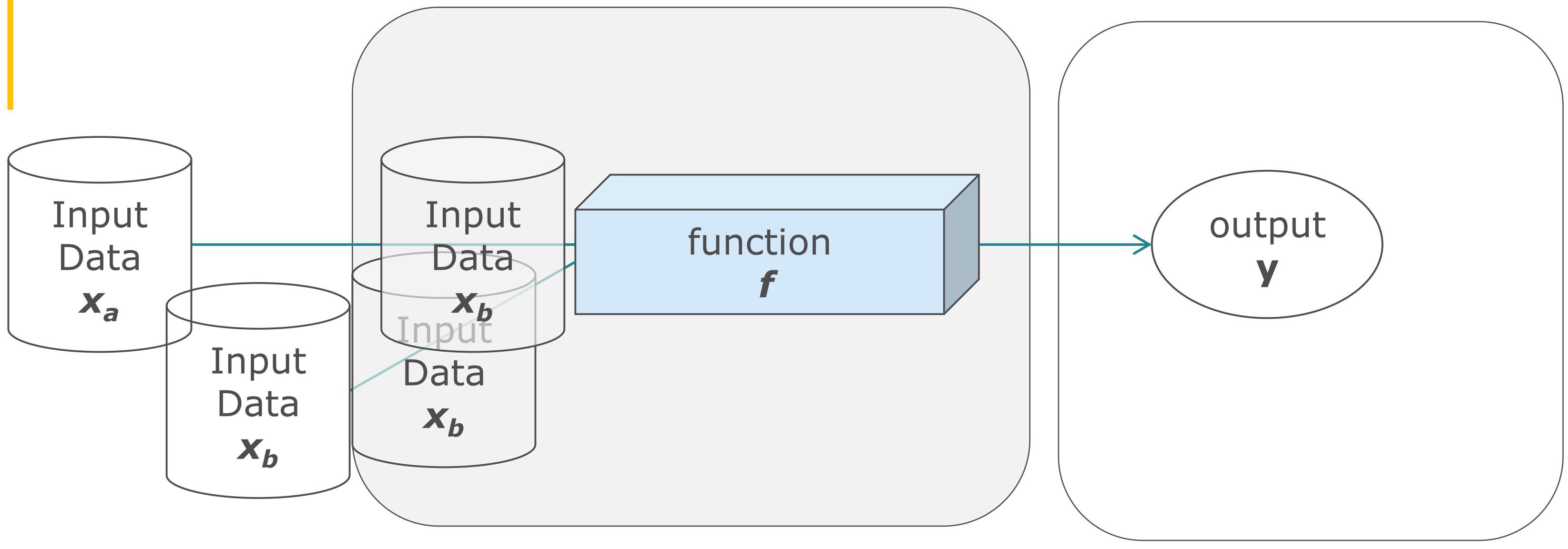
Options

- Do nothing (abstain from computation)
- Exchange input data between the involved entities
- Exchange input data with a Trusted Third Party
- Adopting a (Multi-party) Secure Private Computing solution

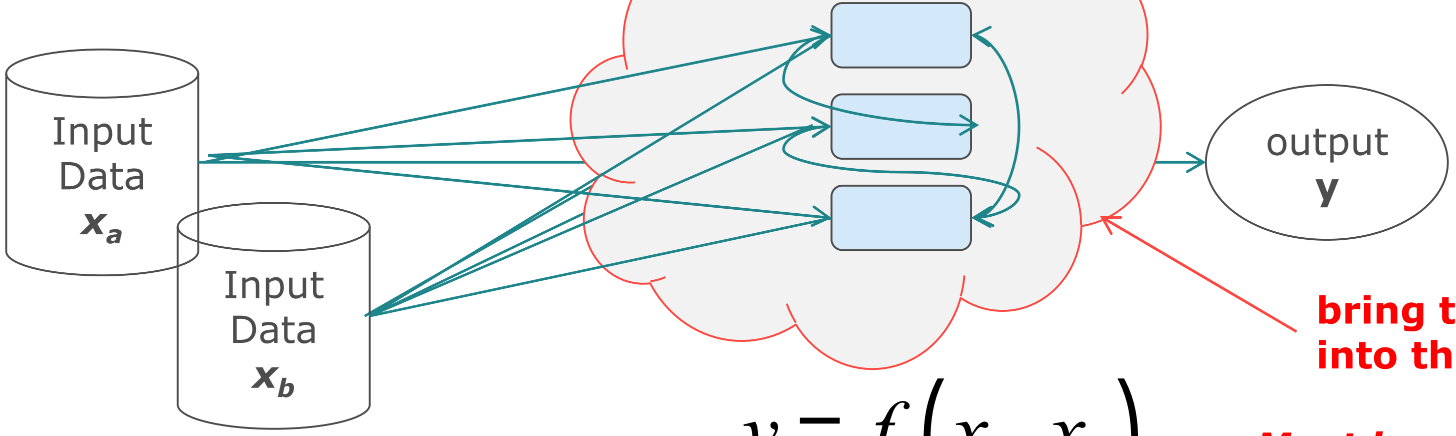
All these options are legitimate and may be preferred in different contexts.

Option selection is a matter of minimising jointly the (actual or perceived) **risks** and **costs**. Therefore potential adopters need to understand the risks and costs of MPC-based solutions, compared to the other options.
Key dimensions shaping costs and risks include: legal compliance, trust model …

Direct Data Sharing (transmit the data) with a Trusted Third Party → single processing party



Secure Private Computing infrastructure with multiple processing parties + other technical and organisational safeguards

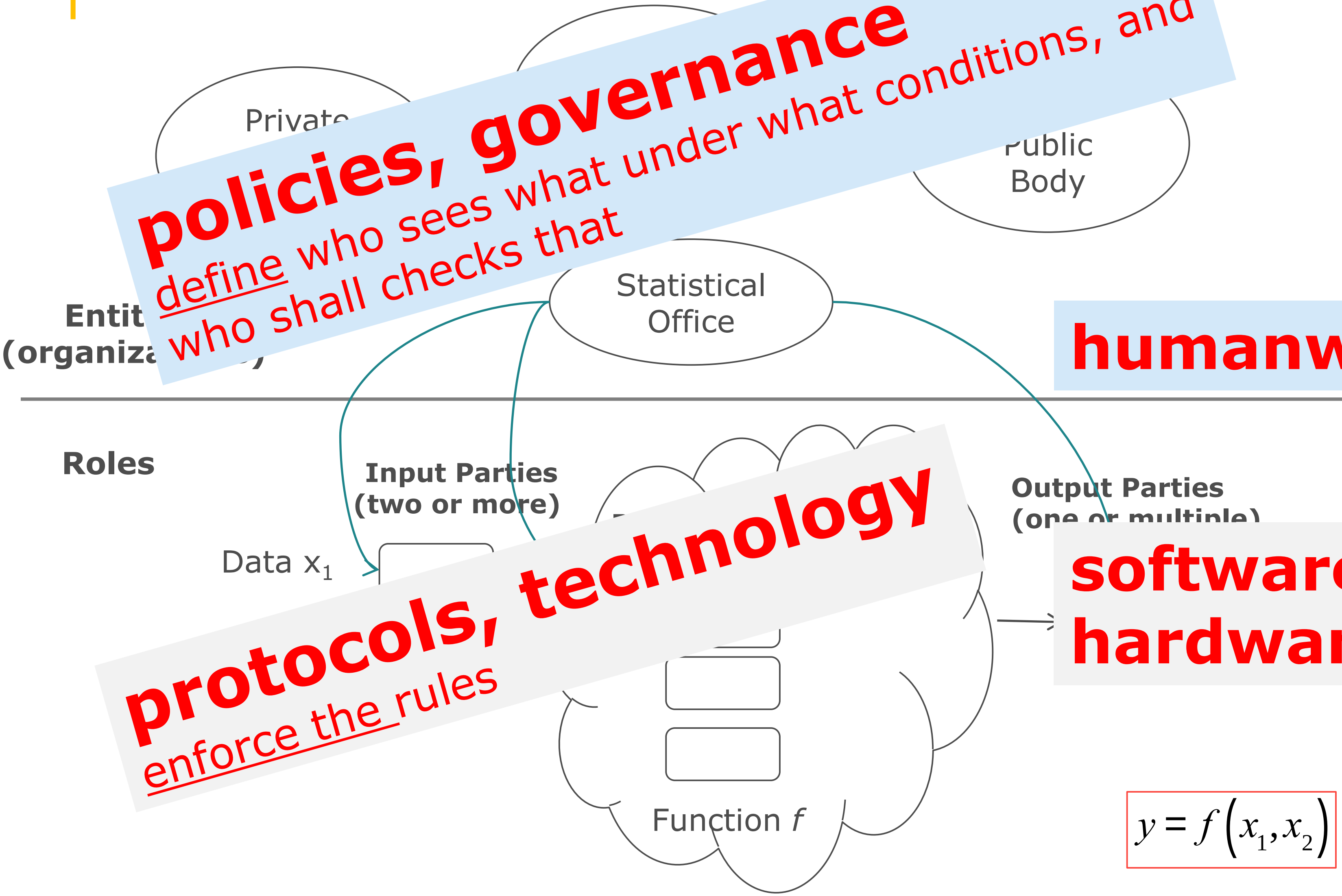


bring the safeguards into the computation

$$y = f\left(x_a, x_b\right)$$

Must be multi-party, but cannot be "just" an MPC protocol

NSI National Statistical Institute
MNO Mobile Network Operator
MPC Multi-Party Computation



Private

Public Body

Statistical Office

Entit
(organiza)

Roles

Input Parties
(two or more)

Data x_1

Output Parties
(one or multiple)

Function f

$$y = f\left(x_1, x_2\right)$$

policies, governance

define who sees what under what conditions, and who shall checks that

humanware

protocols, technology

enforce the rules

software & hardware

"…technical and organisational measures…"

Legal compliance

- In our current understanding, MPC-based solutions qualify as *processing of personal data* and therefore remain within GDPR
  - MPC solutions as *supplementary "technical and organisational measures"* in the sense of GDPR Art. 89 (*,**)

- Well-designed MPC solutions, based on strong implementations of state-of-the-art technologies, can be effective means of compliance with GDPR
  - Embracing GDPR principles as 'design requirements' for MPC-based solutions: data minimisation, purpose specification, storage limitation, integrity and confidentiality …

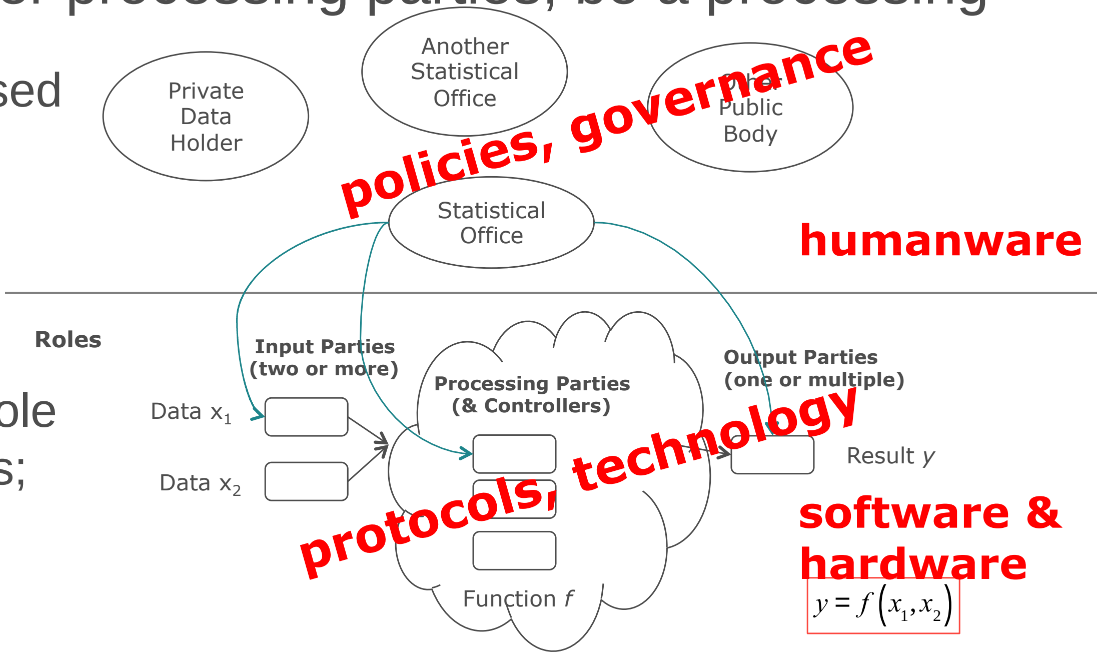
(*) In line with EDPB Recommendations 01/2020 on measures that supplement transfer tools to ensure compliance with the EU level of protection of personal data (Use Case 5: Split or multi-party processing)

(**) In line with ENISA view, see report on "Data Pseudonymisation: Advanced Techniques and Use Cases", January 2021

Trust model

- The essential role of the is to <u>enforce technologically the governance/policies</u> (for data <u>& code</u>) defined among entities
- Goal: avoid single-point-of-trust (SPoT) → the set of processing parties are to be trusted *collectively, not individually*
- If you don't trust the other processing parties, be a processing party yourself!

- The overall strength of MPC-based solution depends *jointly* on
- (i) robustness of policies/governance scheme;
- (ii) choice of entities taking the role of processig parties & controllers;
- (iii) strength of technology implementation



Engineering problems

These are *just* engineering problems!

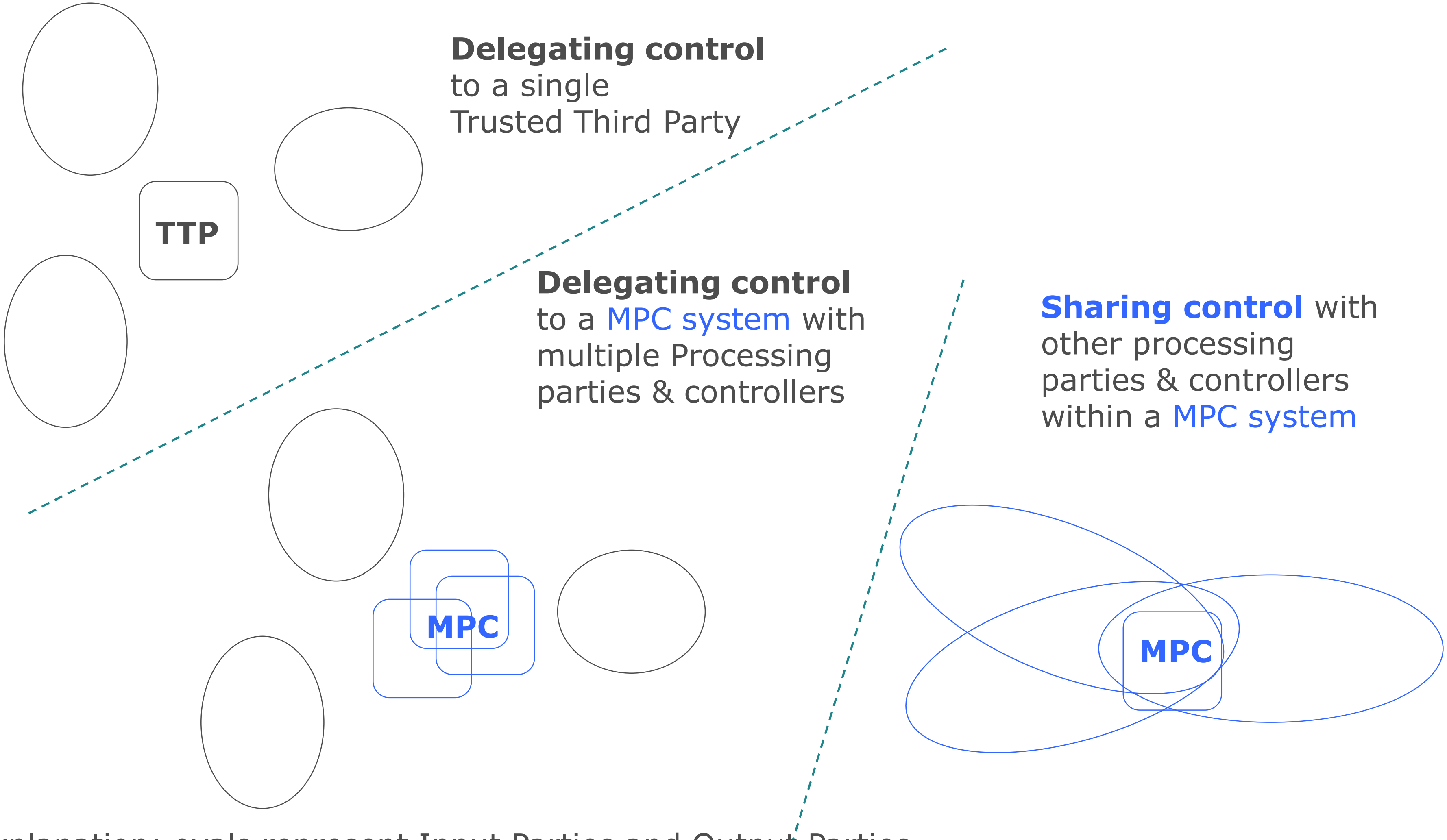
- The overall strength of MPC-based solution depends *jointly* on
- (i) robustness of policies/governance scheme;
- (ii) choice of entities taking the role of processig parties & controllers;

e.g., mutual independence, (partly) antagonist goals,…

- (iii) strength of technology implementation

e.g., combine technologies with complementary guarantees, overlay multiple security layers

From delegation to sharing (of processing control)



Delegating control to a single Trusted Third Party

Delegating control to a MPC system with multiple Processing parties & controllers

Sharing control with other processing parties & controllers within a MPC system

Explanation: ovals represent Input Parties and Output Parties.
Rectangles represent processing parties & controllers

Cost-Risk trade-offs

Number of involved processing parties (& controllers) ↑ Þ

↓ Residual risks (collusion, intrusion,…)

↑ Costs (organisational, operational, …)

Strength of technological components ↑ Þ

↓ Residual risks (breaches, attacks,…)

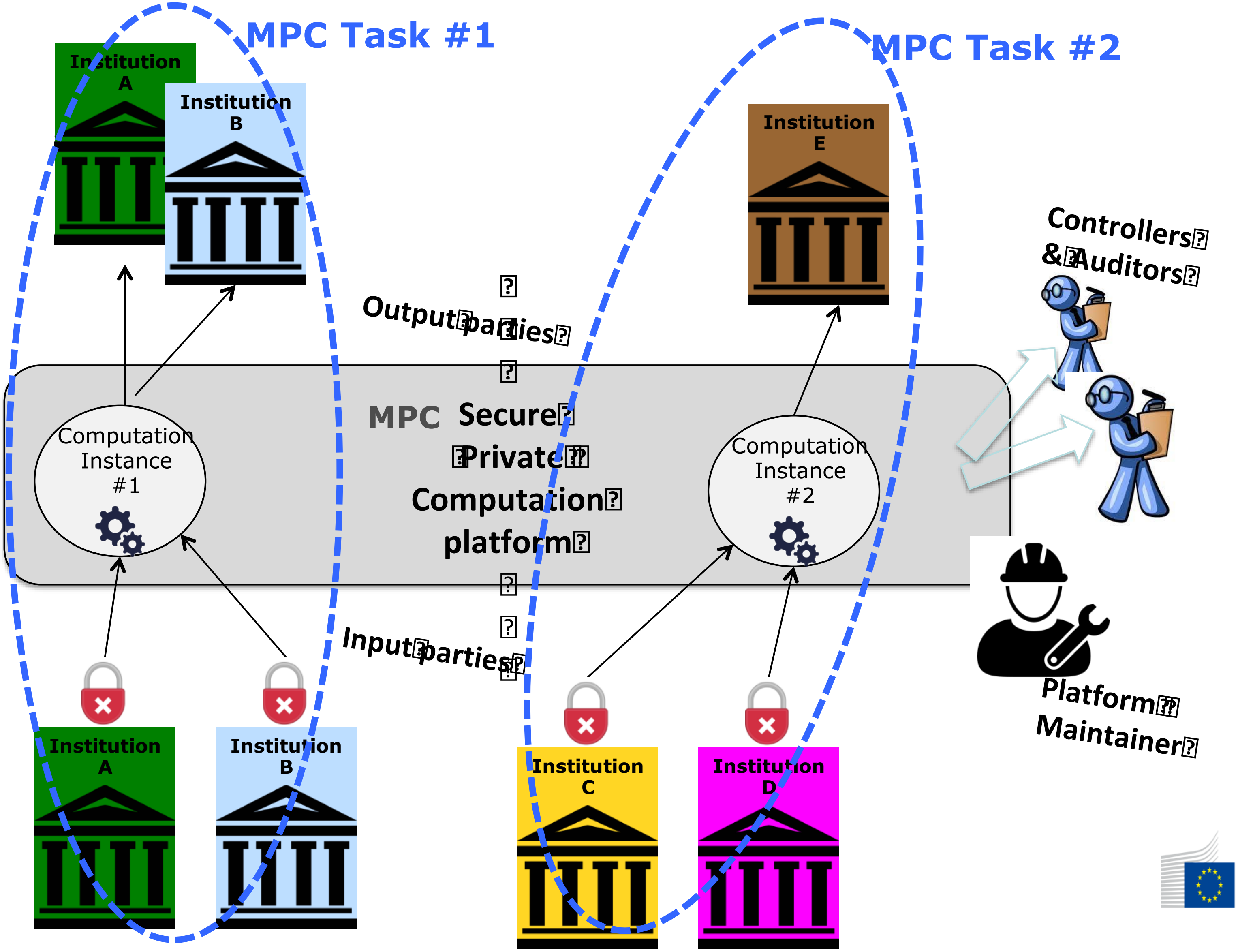
↑ Costs (design, implementation, hw resources, …)

Joining forces among potential adopters

- Q. How to make the strongest possible MultiParty Secure Private Computing (MPSPC) solution affordable *for the adopters*?
  - Lowest risk at low cost

Shared MPSPC platform → MPSPC-as-a-service

MPC Secure Private Computing-as-a-service



MPC Task #1

MPC Task #2

Institution A

Institution B

Institution E

Output parties

MPC Secure Private Computation platform

Computation Instance #1

Computation Instance #2

Input parties

Institution A

Institution B

Institution C

Institution D

Controllers & Auditors

Platform Maintainer

MPC Secure Private Computing-as-a-service

- Built and operated by a consortium/network *of* public institutions *for* public institutions (+ their private partners)
  - E.g. European Statistical System (ESS)

- Team-up with specialised technology providers for co-design of all-round solution (policies & protocols)

- Consultation with Data Protection Authorities already at design phase to ensure legal compliance
  - *Embrace GDPR: take GDPR principles as design requirements*

Take-home message

- MPC-based solutions have an important role to play (also) in the public sector as alternative to direct data exchange.
  - Technology for embracing GDPR, not eluding it
- Shared (Multi-Party) Secure Privacy Computing-as-a-service platform as possible way to facilitate adoption in Official Statistics
  - Can serve as a lighthouse and showcase for other sectors
- Co-design of all-round solutions between technology providers and potential adopters & consultation with Data Protection Authorities as key success factors
  - Constructive viewpoint: GDPR principles as design requirements

Thank you for your attention

More about the work done at Eurostat on Privacy Enhancing Technologies for Official Statistics:

https://ec.europa.eu/eurostat/cros/content/privacy-enhancing-technologies-official-statistics-pet4os_en