European
Commission

Access to external data for the production of official statistics and personal data protection: what can we demand from *Secure Private Computing* technologies?

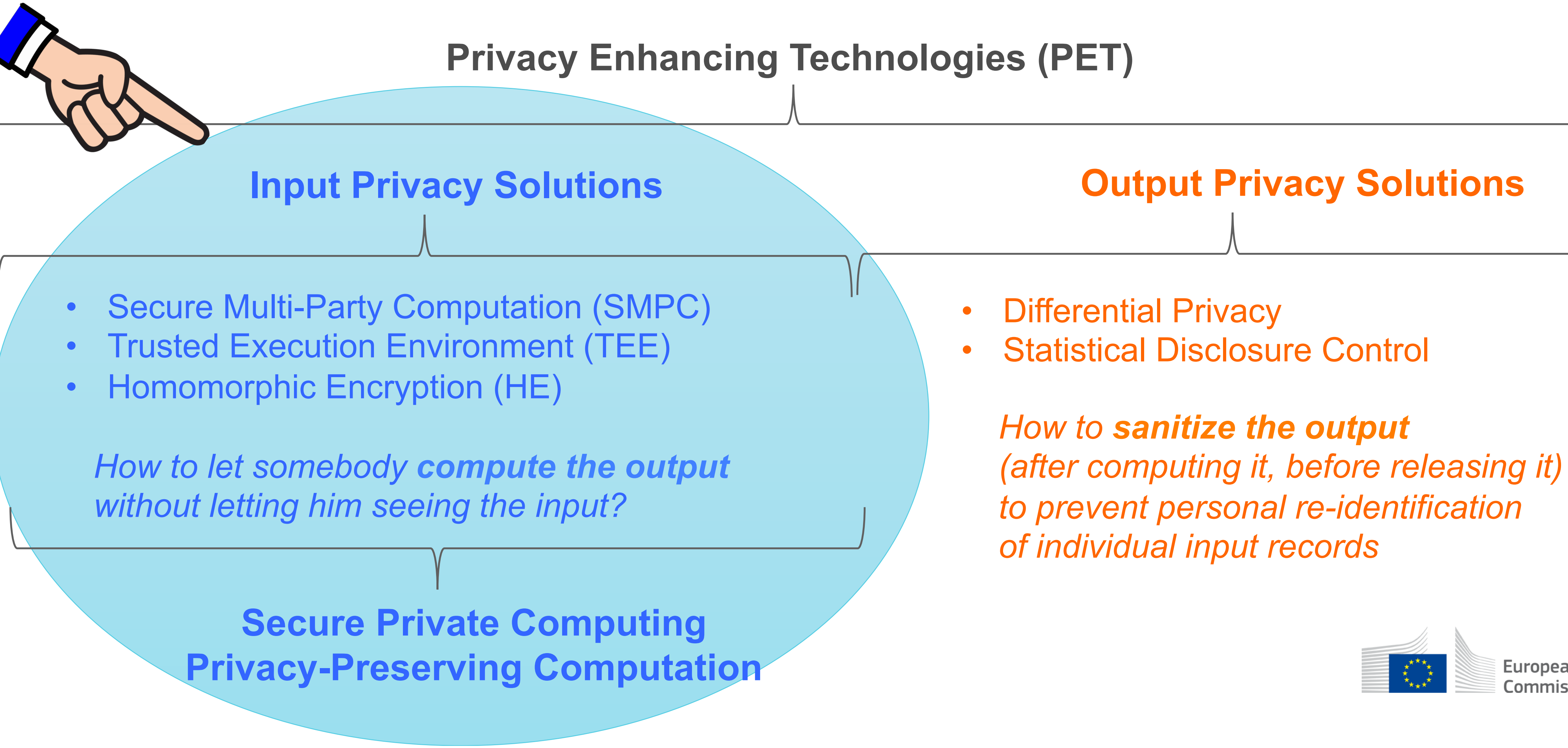
Fabio Ricciato

Eurostat, Unit A5 Methodology; Innovation in official statistics

ISTAT Workshop sulla Protezione dei Dati Personali

23rd June 2021

Terminology



Privacy Enhancing Technologies (PET)

Input Privacy Solutions

- Secure Multi-Party Computation (SMPC)
- Trusted Execution Environment (TEE)
- Homomorphic Encryption (HE)

*How to let somebody **compute the output** without letting him seeing the input?*

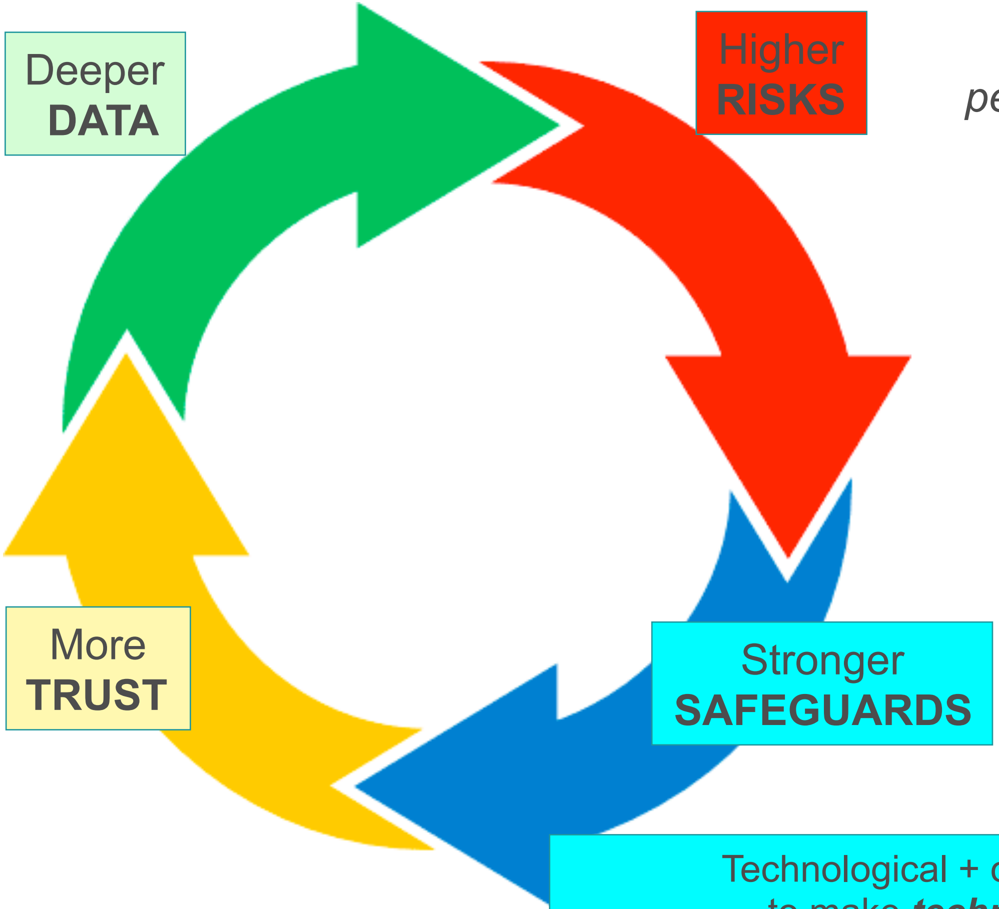
Secure Private Computing
Privacy-Preserving Computation

Output Privacy Solutions

- Differential Privacy
- Statistical Disclosure Control

*How to **sanitize the output** (after computing it, before releasing it) to prevent personal re-identification of individual input records*

Why?



ore detailed
a (nano-data)

& higher
perception of risks

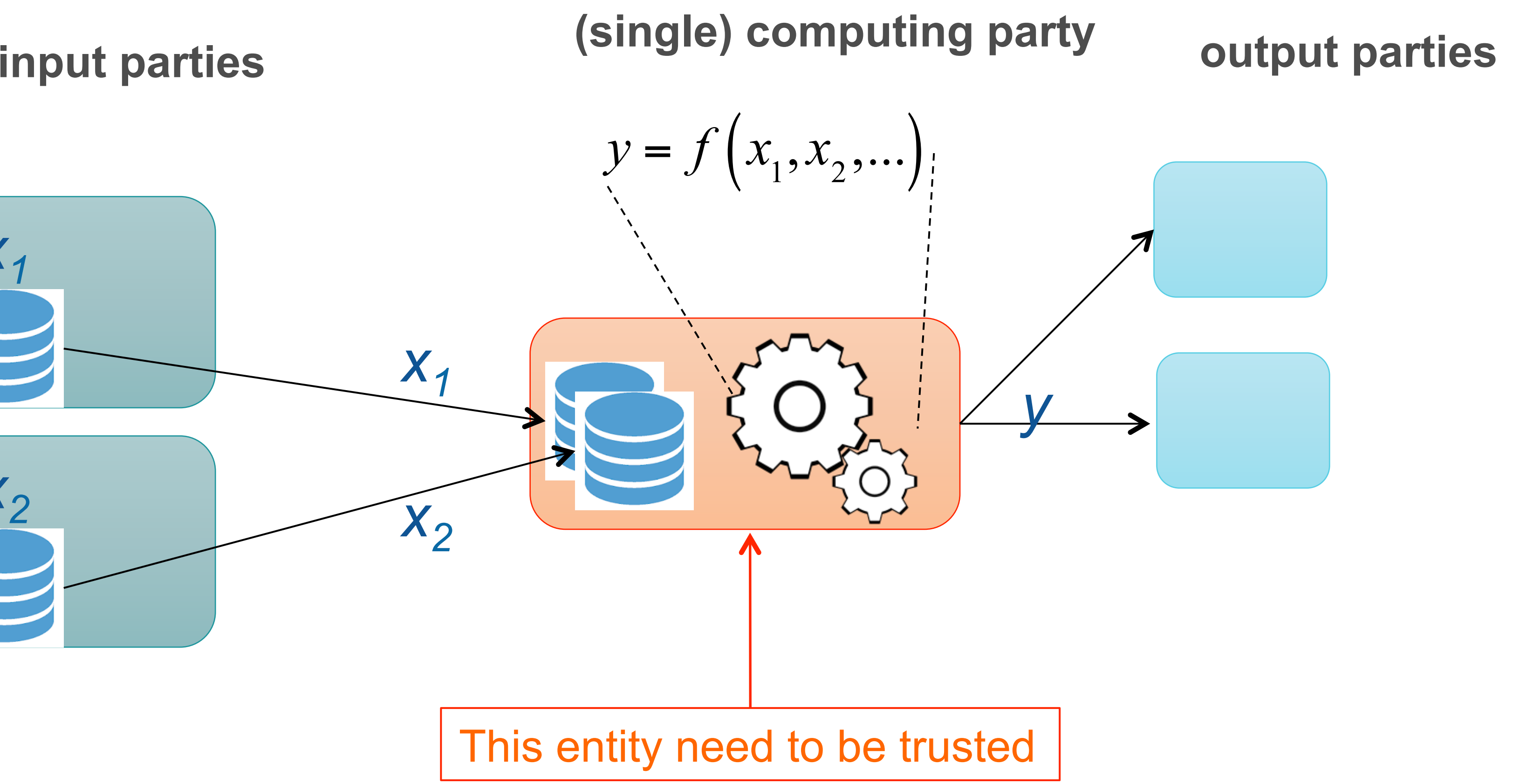
to be put in place
+
to be clearly communicated

blic acceptance,
public trust

Technological + organisational means
to make ***technically unfeasible***
what is **legally prohibited**
and **transparently communicated**

In traditional computing models data are moved → data get **centralised**

→ all players must trust the single computing party (delegation of control)



input parties

(single) computing party

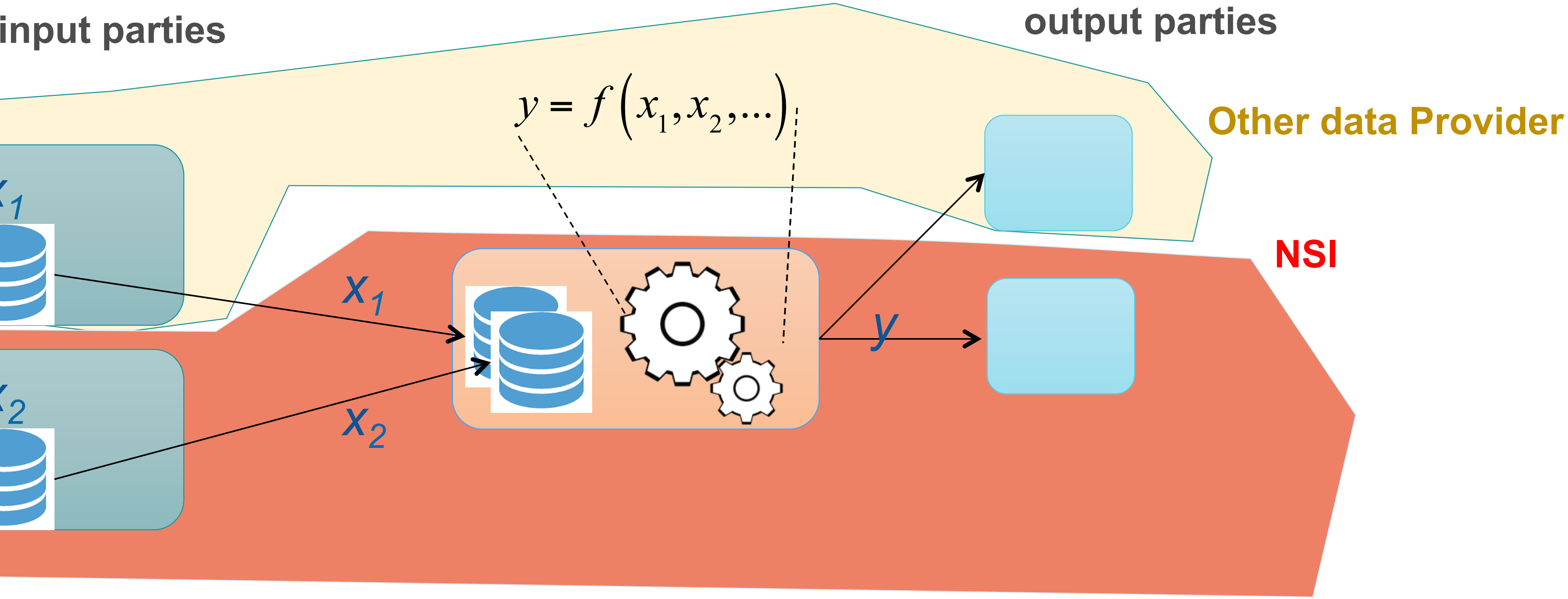
output parties

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

This entity need to be trusted

In traditional computing models data are moved → data get **centralised**

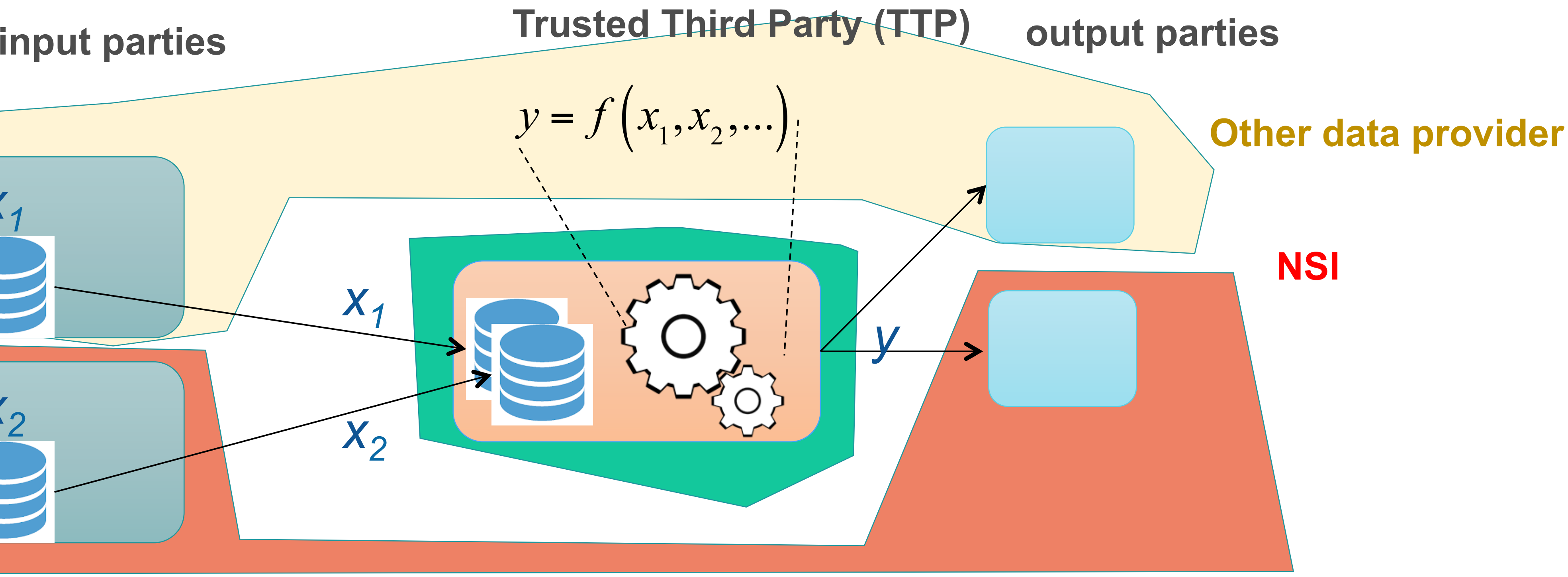
→ all players must trust the single computing party (delegation of control)



"pull data to the NSI" → NSI as single point of trust

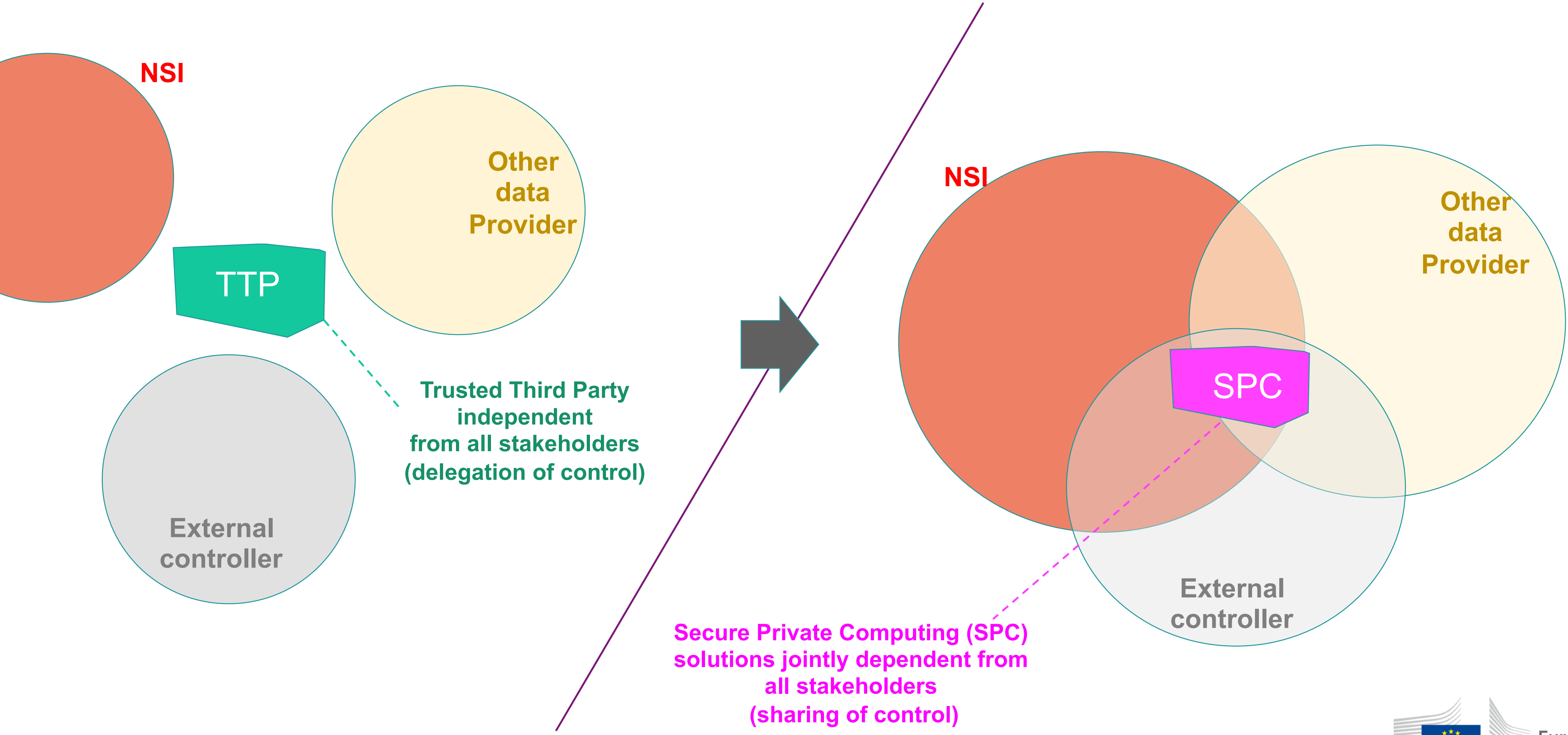
In traditional computing models data are moved → data get **centralised**

→ all players must trust the single computing party (delegation of control)



Trusted Third Party (TTP) external to NSI → TTP single point of trust

From **delegation** to **sharing** of control

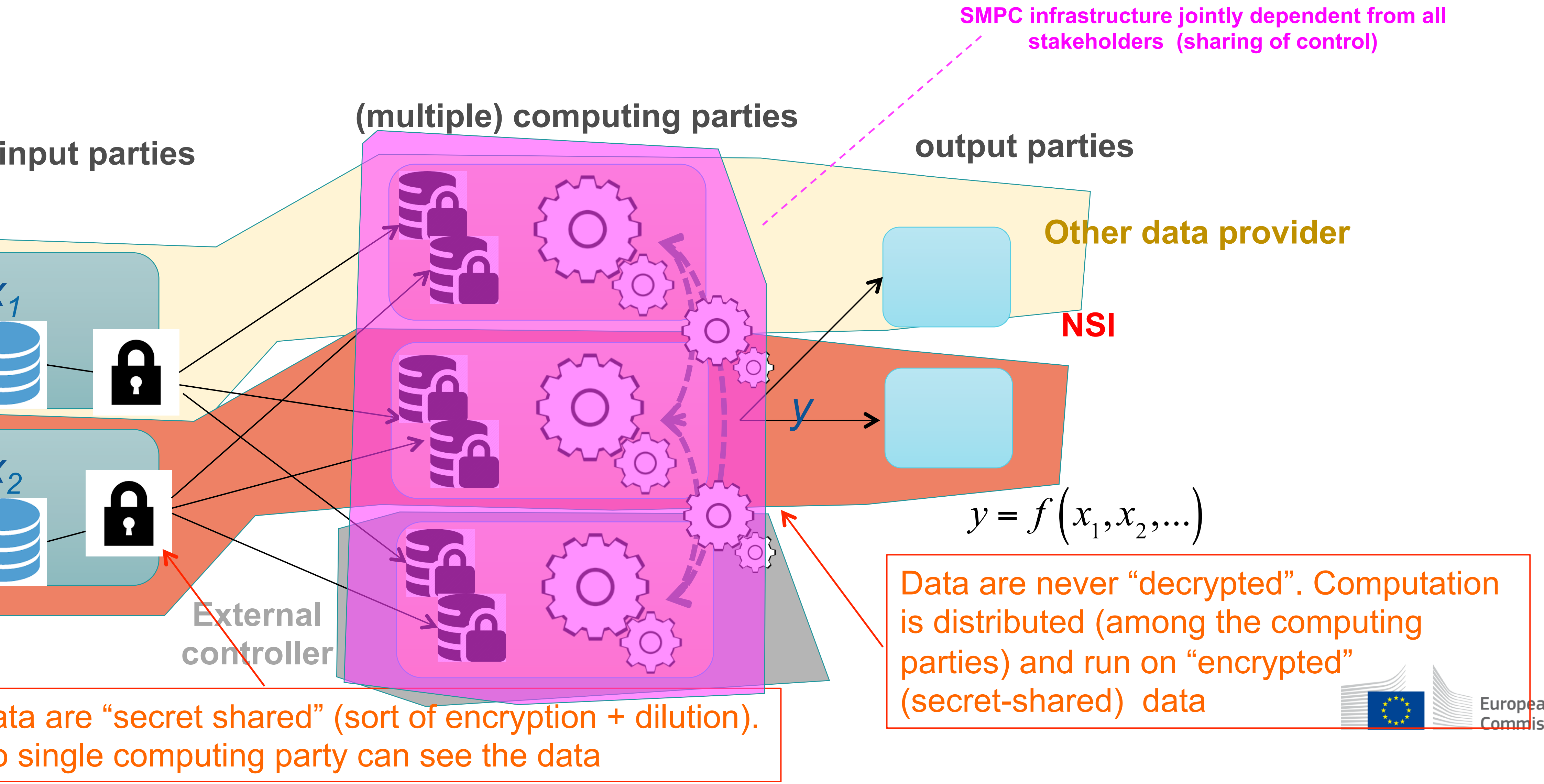


Trusted Third Party independent from all stakeholders (delegation of control)

Secure Private Computing (SPC) solutions jointly dependent from all stakeholders (sharing of control)

e: Trusted smart statistics: Motivations and principles. Statistical Journal of the IAOS 35 (2019) 589–603

Secure Multi-Party Computation (SMPC)

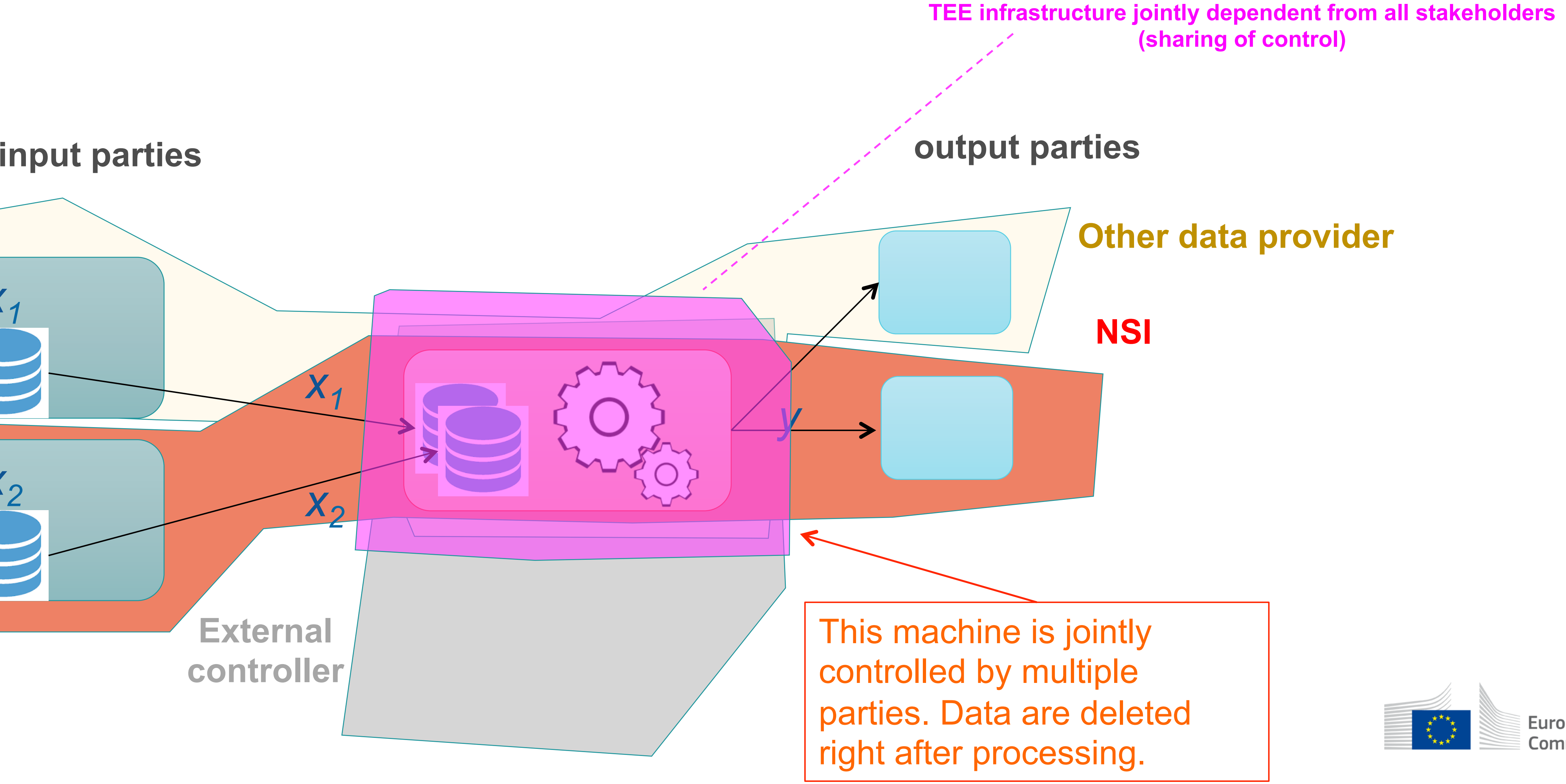

SMPC infrastructure jointly dependent from all stakeholders (sharing of control)
(multiple) computing parties
input parties
output parties
Other data provider
NSI
x_1
x_2
y
External controller

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

Data are never “decrypted”. Computation is distributed (among the computing parties) and run on “encrypted” (secret-shared) data

ata are “secret shared” (sort of encryption + dilution). o single computing party can see the data

Trusted Execution Environment (TEE)



This machine is jointly controlled by multiple parties. Data are deleted right after processing.

From delegation to sharing of control… via Secure Private Computing (SPC)

The SPC process is designed so as to avoid "**single point of trust**"

Avoid centralised control over the **data**

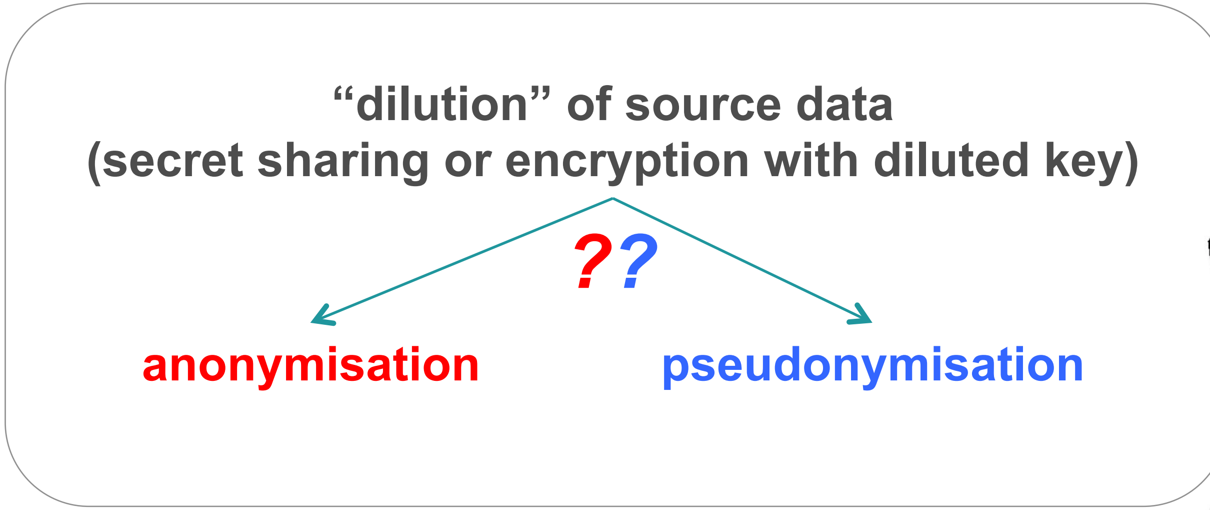
- Either data are "secret shared" (sort of encryption where the **cipher-text is diluted** among multiple parties; computation run without de-ciphering the input data → SMPC
- … or data are encrypted (with some traditional scheme) and the **cipher-key is diluted** amo multiple parties; data are provably deleted after computation → TEE

Share control over the **code** - involve as many (external) controllers as needed

- Trust **collectively** the set of controllers & the whole process
- Ex-ante controls - e.g. preliminary code approval to prevent mis-use
- Ex-post controls (e.g. detailed non-modifiable logging) to enable forensic audits → deterrer

Europea
Commis

SPC and GDPR


“dilution” of source data
(secret sharing or encryption with diluted key)
??
anonymisation
pseudonymisation

- SPC to “escape” GDPR ?
    - If “dilution” of input data is considered “*anonymisation*”
      → “diluted data” *are not* personal data → GDPR *does not* apply
    - [endorsement of this view by DPAs unlikely - anyway not our view]

- **SPC to strengthen GDPR implementation !**
    - If “dilution” of input is considered “*pseudonymisation*”
      → “diluted data” *are* personal data → GDPR does apply
    - [more conservative approach, endorsement by DPAs more likely - our view!]

C: Secure Private Computing

SPC to strengthen GDPR implementation

- GDPR requires
  - Legal basis to process the data
  - A set of appropriate **technical and organisational safeguards** to protect the data

SPC

- GDPR principles relevant to SPC
  - **purpose limitation** → in a well-designed SPC solution only approved code can be executed tightest possible form of purpose specification (purpose = code)
  - **data minimisation** → in a well-designed SPC solution only the very final result is disclosed, no other information can be leaked – tightest possible form of data minimisation
  - **storage limitation** → a well-designed SPC solution shall include automatic deletion of secre shared data or encrypted data and related leys – tightest possible form of storage limitation
  - **integrity and confidentiality** → a well-designed SPC solution comes with state-of-the art security functions
  - **[privacy by design]** → inherent to SPC!

rticle 89(1) - Processing for archiving purposes in the public interest, scientific or historical research purposes or **statistical purposes**, shall be **subject opriate safeguards** [...]. Those safeguards shall ensure that **technical and organisational measures are in place** in particular in order to ensure or the principle of **data minimisation**. Those measures may include **pseudonymisation** provided that those purposes can be fulfilled in that manner.

Wrap-up

Ethical duty + legal obligation
to protect personal data with *appropriate safeguards (technical + organisational*

- (i) proportional to the risks: more detailed input data → higher risks → stronger safeguards
- (ii) taking into account state-of-the-art technologies

Well-designed solutions based on SPC technologies as "appropriate safeguards

- Key ingredients of SPC: sharing of control (over the code, over the data), transparency, auditability
- Reminiscent of the "checks and balances" principle underlying the democratic system

SPC technologies are the bricks, not a magic stick --- you still need to engineer
whole solution (hardware, software and … humanware)

Europea
Commis

Thank you