GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Pseudonymization as a data protection by design instrument

12 November 2019 - Berlin

Giuseppe D'Acquisto

GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Definition (art. 4.5 GDPR)

The processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person

Mario

Maria

GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

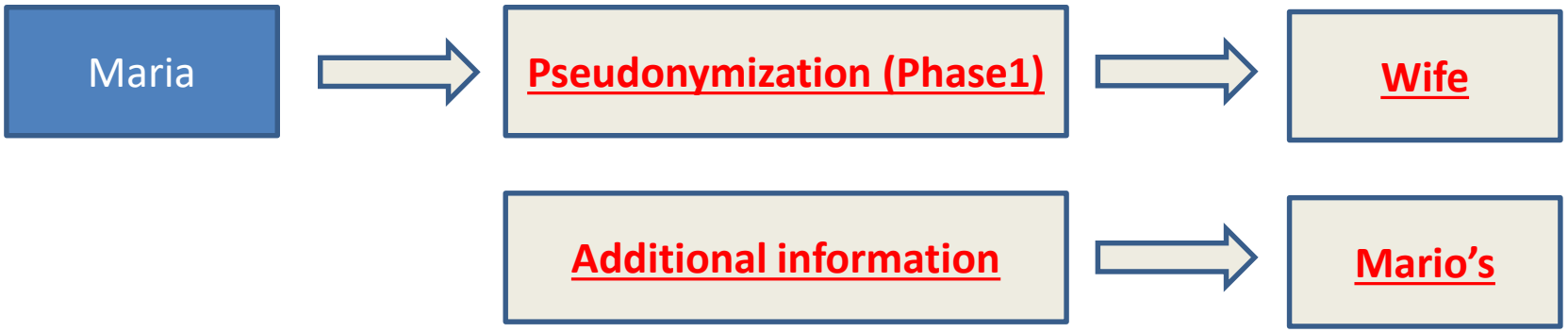
Definition (art. 4.5 GDPR)

The processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person

GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Definition (art. 4.5 GDPR)

The processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person



GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Pseudonimizzazione (art. 4.5)

The processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person

Mario

Maria

Mario's wife

GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

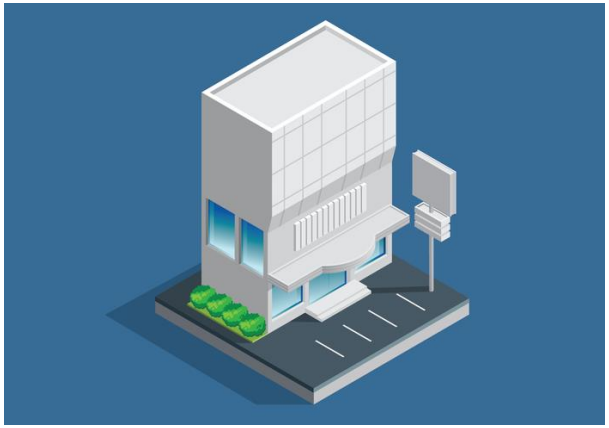
Psudonymization: context matters

Maria's office

Mario's office

Maria

Mario's wife



Maria

Mario's wife

Who wrote this document?

Who is the lady in this picture?

GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Pseudonymization: first considerations

Each of us may have multiple identities (or pseudonyms), depending on the context

Pseudonymization is the reference scheme that we should use to build and process those multiple identities

Pseudonymization introduces a principle of “relativity” of our identities

Pseudonymization is not aimed solely at reducing the power of identification of personal data. It can facilitate identification. As such it is not an anonymization of personal data

The additional information can be held by anyone

Scope of pseudonymization (art. 25.1)

Taking into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing, the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures, such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects.

T4DATA

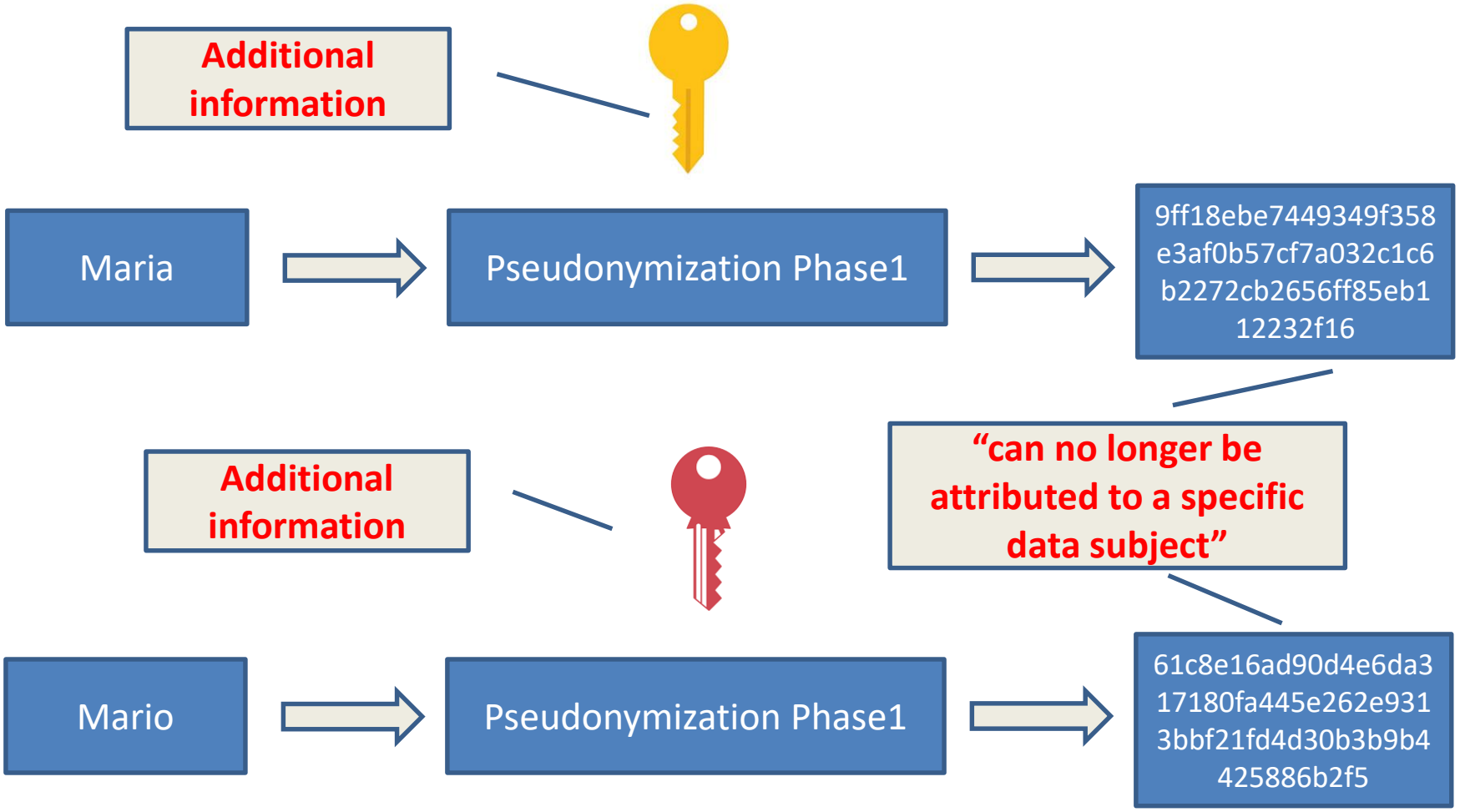
GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Effective implementation of data protection principles

- Lawfulness, fairness and transparency
- Purpose limitation
- Data Minimization
- Accuracy
- Storage limitation
- Integrity and confidentiality

GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Pseudonymization in practice



T4DATA

GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Technical features

Inversion «computationally» impossible

Unintelligible output (pseudo-random) for humans. Not for a machine

Negligible collision probability (there are plenty of pseudonyms)

Output distance insensitive w.r.t input distance (pseudonymization amplifies small distances)

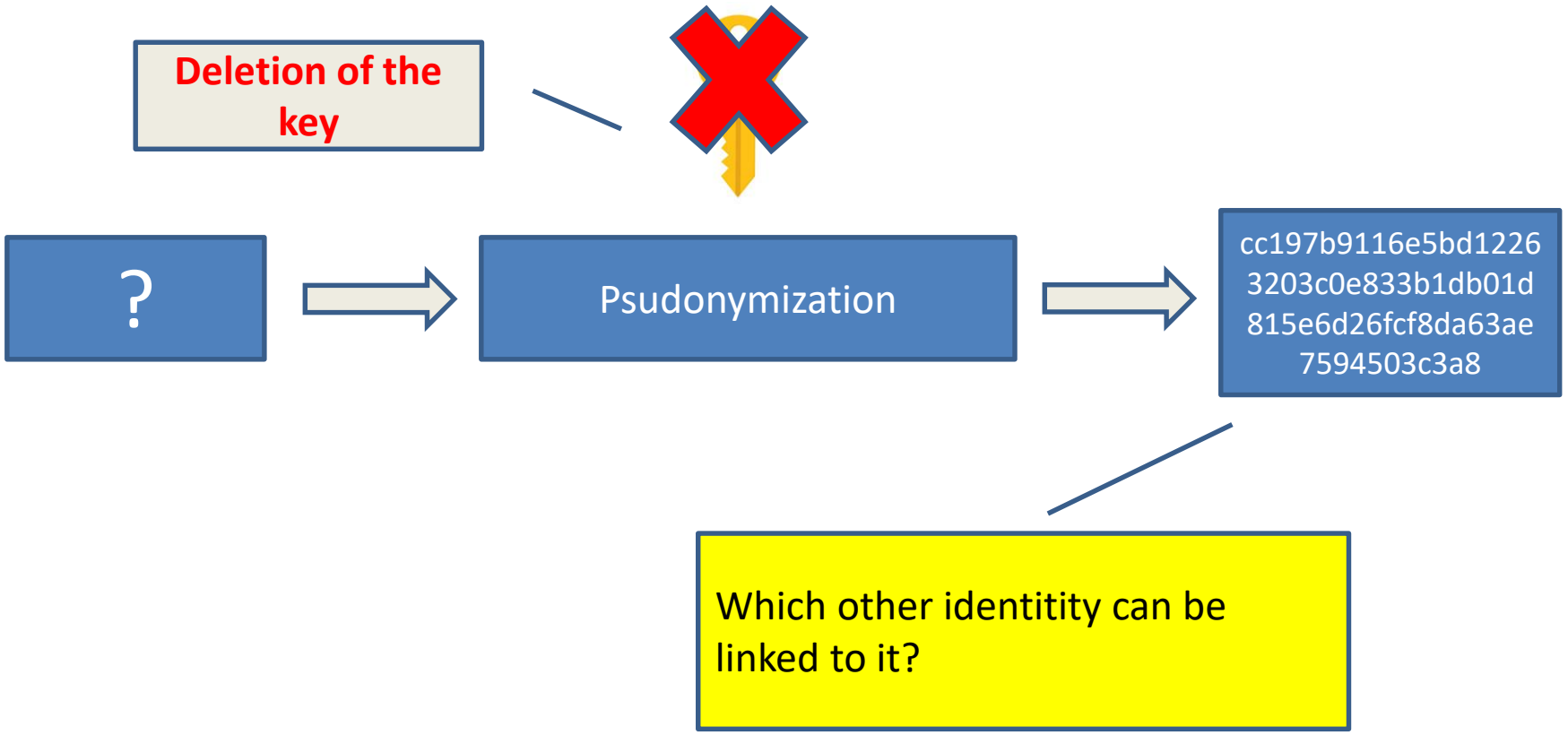
GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Integrity and accuracy

9ff18ebe7449349f358e3af0b57cf7a032c1c6b2272cb2656ff85eb112232f16	9ff18ebe7449349f358e3af0b57cf7a032c1c6b2272cb2656ff85eb112232f16	9ff18ebe7449349f358e3af0b57cf7a032c1c6b2272cb2656ff85eb112232f16	**61c8e16ad90d4e6da317180fa445e262e9313bbf21fd4d30b3b9b4425886b2f5**
Maria	Maria	Maria	Mario
t1	t2	t3	tn

time

Confidentiality



GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Context matters! (on «computational» irreversibility)

cc197b9116e5bd12263203c0e833b1db01d815e6d26fcf8da63ae7594503c3a8 is an argentinian footbal player who scored a goal to England in the quarter final of World Cup 1986 in Mexico. This goal was named «la mano de Dios» (God's hand)

GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Context matters! (on «computational» irreversibility)

9d1e29cea1a72b63ddad8504167f028a22fcdb994d11fd71beb45ac7d8297474 is the SHA 256 Hash of the string «Diego Armando Maradona»

T4DATA

GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Pseudonymization with minimization

cc197b9116e5bd12263203c0e833b1db01d815e6d26fcf8da63ae7594503c3a8 is an argentinian footbal ~~player who scored a goal to England in the quarter final~~ of World Cup 1986 in Mexico. ~~This goal was named «la mano de Dios» (God's hand).~~

Power of re-identification 1/23

cc197b9116e5bd12263203c0e833b1db01d815e6d26fcf8da63ae7594503c3a8 is an argentinian footbal player who scored a goal ~~to England~~ in ~~the quarter final of~~ World Cup 1986 in Mexico~~. This goal was named «la mano de Dios» (God's hand)~~.

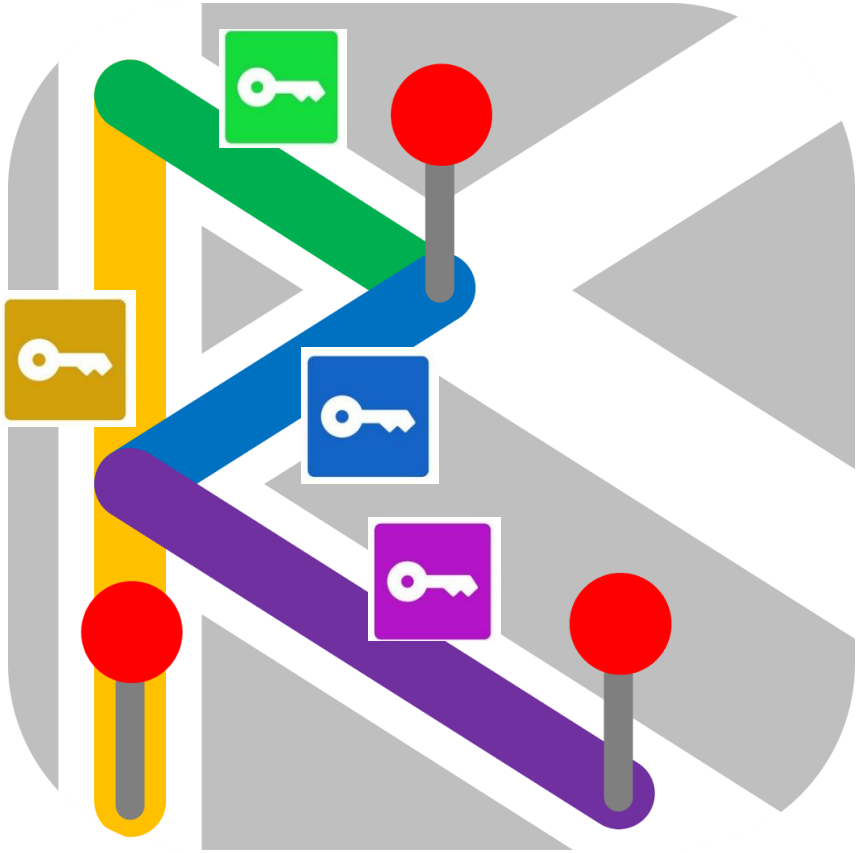
Power of re-identification 1/6

cc197b9116e5bd12263203c0e833b1db01d815e6d26fcf8da63ae7594503c3a8 is an ~~argentinian~~ footbal player who scored a goal ~~to England~~ in the quarter final of World Cup 1986 in Mexico~~. This goal was named «la mano de Dios» (God's hand)~~.

Power of re-identification 1/21

GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Pseudonymization for minimization



GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Fairness

9ff18ebe7449349f358
e3af0b57cf7a032c1c6
b2272cb2656ff85eb1
12232f16

Additional information held by the data subject

Additional information NOT held by the data controller

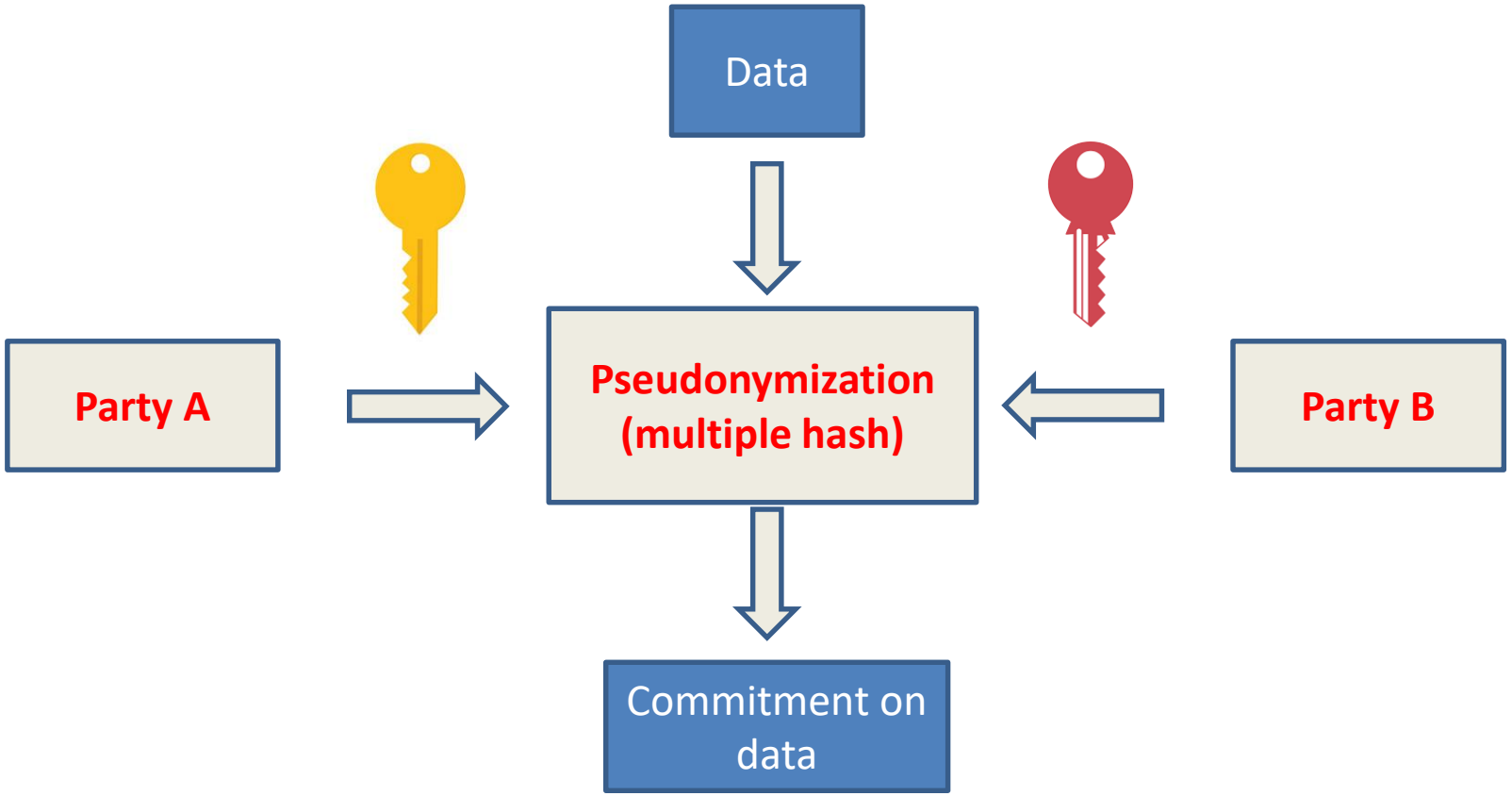
GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Art. 11 (Processing which does not require identification)

1. If the purposes for which a controller processes personal data do not or do no longer require the identification of a data subject by the controller, the controller shall not be obliged to maintain, acquire or process additional information in order to identify the data subject for the sole purpose of complying with this Regulation.

2. Where, in cases referred to in paragraph 1 of this Article, the controller is able to demonstrate that it is not in a position to identify the data subject, the controller shall inform the data subject accordingly, if possible. In such cases, Articles 15 to 20 shall not apply except where the data subject, for the purpose of exercising his or her rights under those articles, provides additional information enabling his or her identification

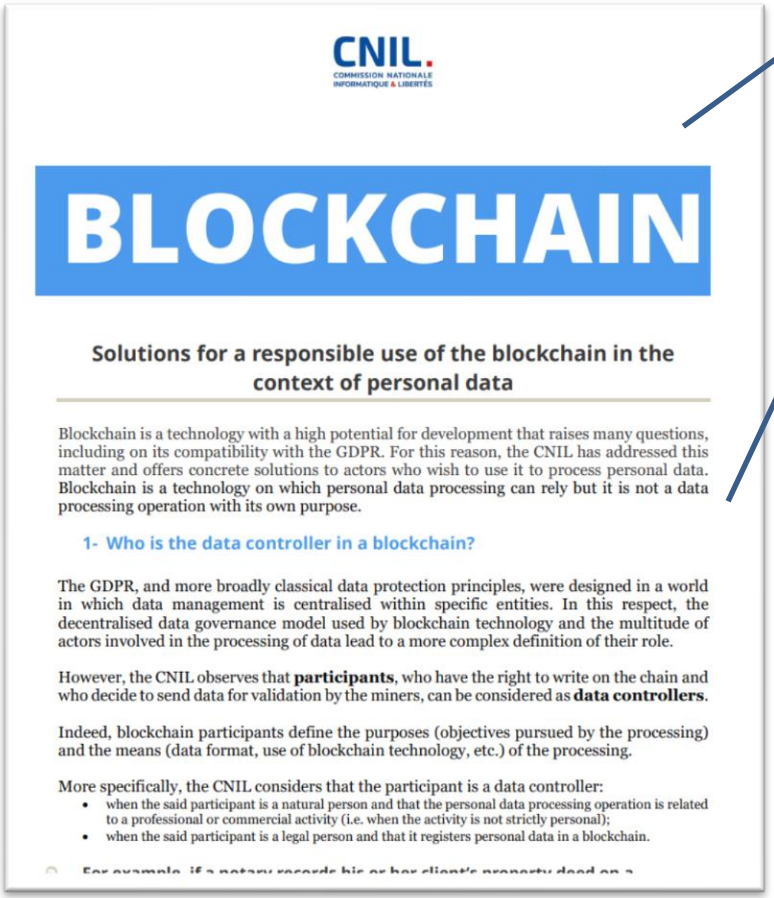
Purpose and storage limitation



Zero knowledge proof (Oded Goldreich, Silvio Micali, Avi Wigderson. Proofs that yield nothing but their validity. Journal of the ACM, volume 38, issue 3, p.690-728. July 1991)

GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Purpose and storage limitation



CNIL.
COMMISSION NATIONALE INFORMATIQUE & LIBERTÉS

BLOCKCHAIN

Solutions for a responsible use of the blockchain in the context of personal data

Blockchain is a technology with a high potential for development that raises many questions, including on its compatibility with the GDPR. For this reason, the CNIL has addressed this matter and offers concrete solutions to actors who wish to use it to process personal data. Blockchain is a technology on which personal data processing can rely but it is not a data processing operation with its own purpose.

1- Who is the data controller in a blockchain?

The GDPR, and more broadly classical data protection principles, were designed in a world in which data management is centralised within specific entities. In this respect, the decentralised data governance model used by blockchain technology and the multitude of actors involved in the processing of data lead to a more complex definition of their role.

However, the CNIL observes that **participants**, who have the right to write on the chain and who decide to send data for validation by the miners, can be considered as **data controllers**.

Indeed, blockchain participants define the purposes (objectives pursued by the processing) and the means (data format, use of blockchain technology, etc.) of the processing.

More specifically, the CNIL considers that the participant is a data controller:

- when the said participant is a natural person and that the personal data processing operation is related to a professional or commercial activity (i.e. when the activity is not strictly personal);
- when the said participant is a legal person and that it registers personal data in a blockchain.

For example, if a notary records his or her client's property deed on a

*…The CNIL considers that personal data should be registered on the blockchain **preferably in the form of a commitment**…*

*…With respect to additional personal data, **in order to ensure compliance with data protection by design** and by default and data minimisation obligations, the CNIL recommends solutions in which data is processed outside of the blockchain or, in which the following are stored on the blockchain, **in order of preference**:*

- *a **commitment** of the data;*
- *a hash generated by a keyed hash function on the data;*
- *a ciphertext of the data.*

GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Lawfulness

Art. 2-quater (Regole deontologiche)

4. Il rispetto delle disposizioni contenute nelle regole deontologiche … costituisce condizione essenziale per la liceità e la correttezza del trattamento dei dati personali.

Il Garante promuove, ai sensi dell'articolo 2-quater, l'adozione di regole deontologiche per il trattamento dei dati personali provenienti da archivi, registri, elenchi, atti o documenti tenuti da soggetti pubblici, anche individuando i casi in cui deve essere indicata la fonte di acquisizione dei dati e prevedendo garanzie appropriate (art. 61.1)

le misure di garanzia …comprese quelle tecniche di cifratura e di pseudonomizzazione, le misure di minimizzazione …(art. 2 septies)

GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Conclusions

Pseudonymization is not just a security measure

Pseudonymization is not an anonymization technique

Pseudonymization must be used to implement data protection principles

There is still a lot of work to do on the legal side (to what extent pseudonymization interprets the principles) and on the technical side (zero knowledge, identity management) to foster its adoption

GARANTE
PER LA PROTEZIONE
DEI DATI PERSONALI

Thank you!