# Pseudonymisation Techniques and experience so far

Monir Azraoui – CNIL

*Technology expert*
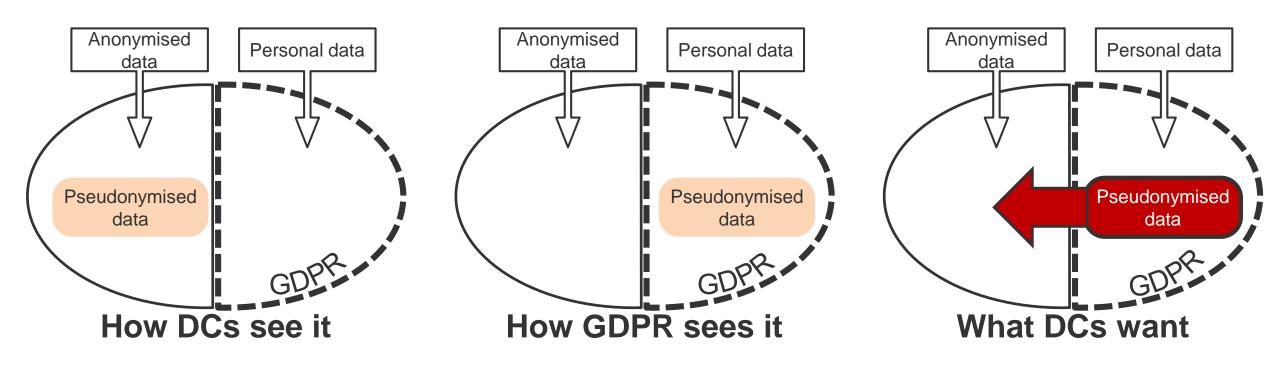
Berlin, November 12th, 2019

CNIL.

# About us: CNIL's missions

- Informs & Protects
- Supports & Guides
- Audits & Sanctions
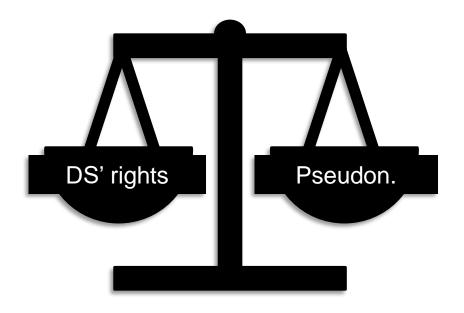- Anticipates & Encourages Innovation

- Team of +200 Law & Technology experts

CNIL.

# EXPERIENCE GAINED SO FAR

# Common mistake:
# Pseudonymisation ≡ Anonymisation



**How DCs see it**      **How GDPR sees it**      **What DCs want**
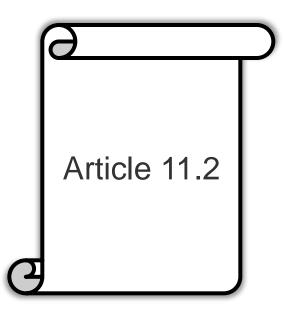
- DCs' grail: anonymise data to escape from GDPR obligations
  - Pseudonymisation often perceived as a "problem"
- GDPR encourages the use of pseudonymisation!

# Challenge with pseudonymisation: Data Subjects' rights

DS' rights
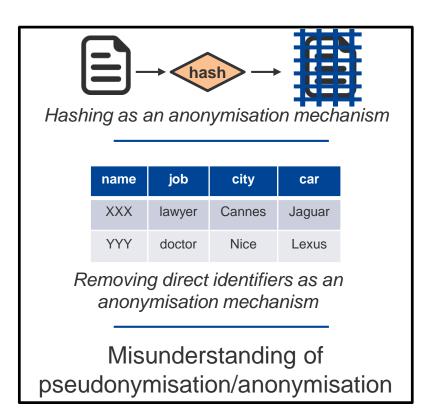
Pseudon.

Article 11.2

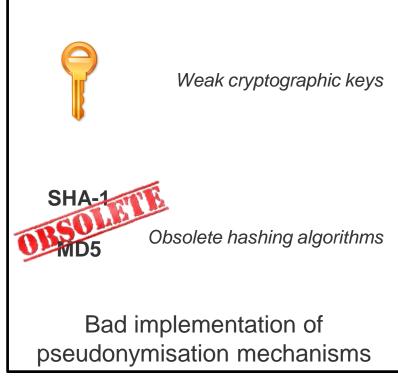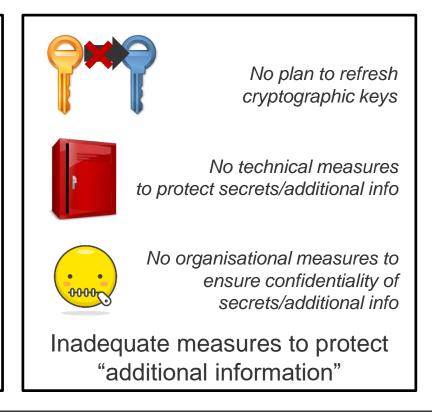Find a balance between DS' rights and pseudonymisation

DCs invoking Article 11.2 to be exempted from the Art. 15-20 obligations

# Audits reveal bad practices…

- Ex-post audits to check whether DCs comply with the law



*Hashing as an anonymisation mechanism*

| name | job | city | car |
|------|-----|------|-----|
| XXX | lawyer | Cannes | Jaguar |
| YYY | doctor | Nice | Lexus |

*Removing direct identifiers as an anonymisation mechanism*

Misunderstanding of pseudonymisation/anonymisation

*Weak cryptographic keys*

**SHA-1**
**OBSOLETE**
**MD5**  *Obsolete hashing algorithms*

Bad implementation of pseudonymisation mechanisms

*No plan to refresh cryptographic keys*

*No technical measures to protect secrets/additional info*

*No organisational measures to ensure confidentiality of secrets/additional info*

Inadequate measures to protect "additional information"

CNIL.

# … and good practices!

- Ex-post audits to check whether DCs comply with the law

| | | |
|---|---|---|
| Risks / Data (balance scale) | Secure storage of *cryptographic keys or secrets* | Encryption of *pseudonymised files* |
| | Plan to refresh *cryptographic keys* | |

**Re-identification risk analysis performed**

**Protection of "additional information"**

| name | job | city | car |
|------|--------|--------|--------|
| XXX | lawyer | France | Jaguar |
| YYY | doctor | France | Lexus |

$k$-anonymity
$l$-diversity

**Combination of pseudonymisation with other security techniques**

CNIL.

# Sanctions when pseudonymisation failed



- Warning
- Formal notice
- Sanction

3 levels of sanctions



**Example 1**
(public) **Formal Notice** against the French Health Insurance System for **poor pseudonymisation** of health insurance data



**Example 2**
(confidential) **Injunction under a coercive fine** against a French public institution for using **an obsolete algorithm for pseudonymisation**
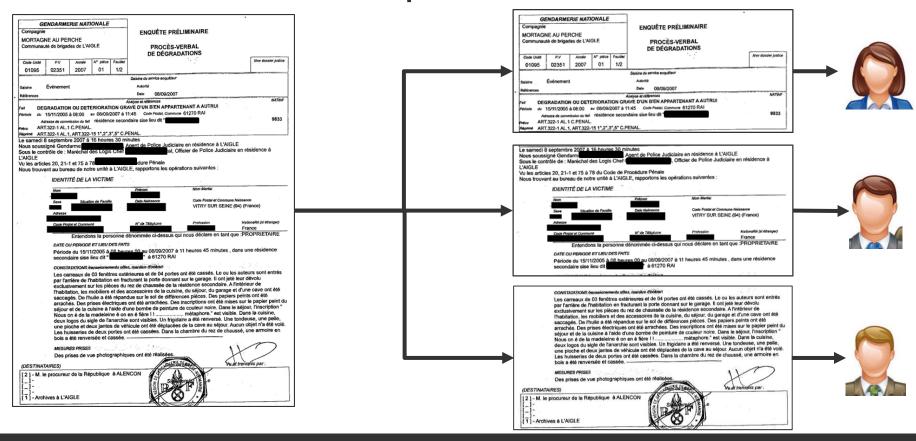
# RELEVANT BEST PRACTICES

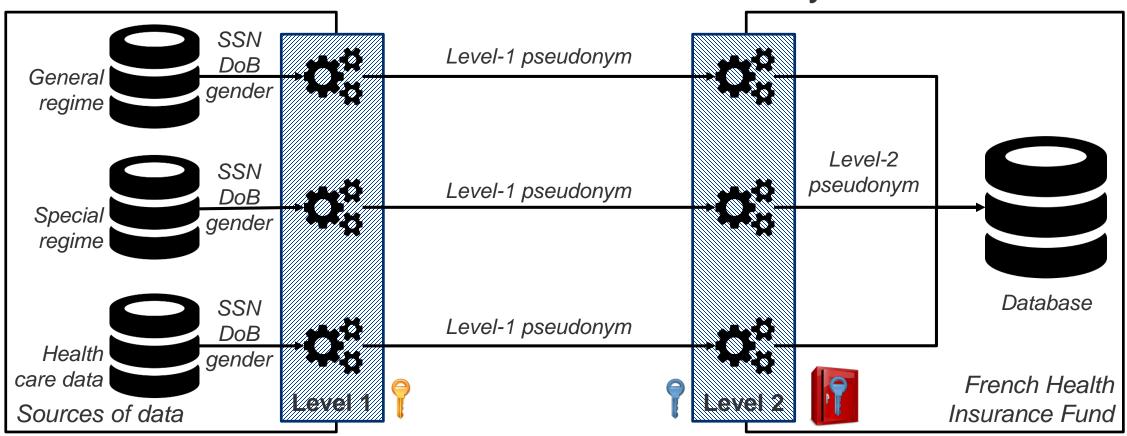# Use Case #1: partitioning data
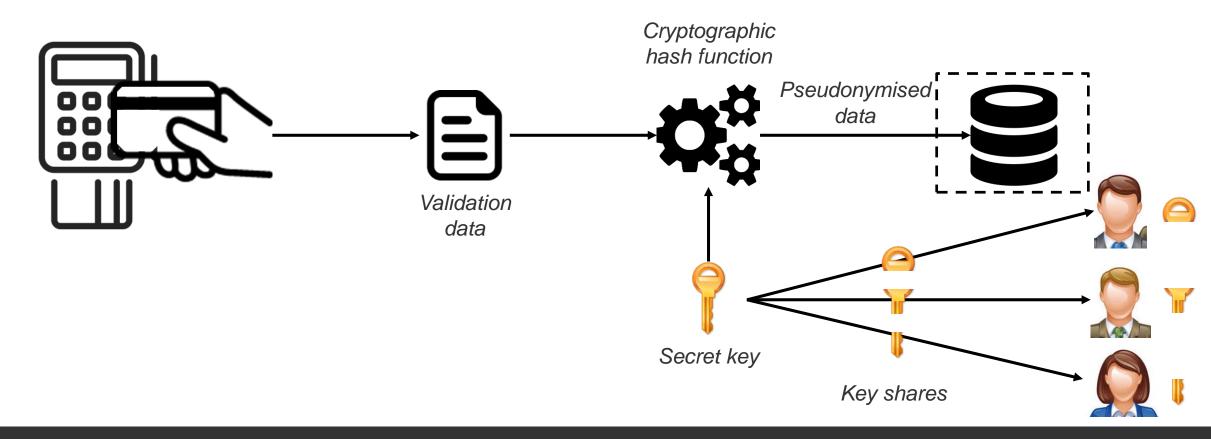
- Police records transcription

# Use Case #2: 2-level pseudonymisation

- French Health Insurance Information System

# Use case #3: hashing and key splitting

- Ticketing data in public transportation systems (fraud)



Cryptographic hash function

Pseudonymised data

*Validation data*

*Secret key*

*Key shares*

# CONCLUDING REMARKS

# Main takeaways

- Pseudonymisation, long-term experience of the CNIL
- Pseudonymisation often mistaken with anonymisation by DC
- Misunderstanding of the benefits of pseudonymisation by DC
- Audits by the CNIL reveal good and bad practices among DC
  - Some bad practices resulted in sanctions by the CNIL
- Practical pseudonymisation: a variety of techniques

# CNIL.

# QUESTIONS?

✉ mazraoui (at) cnil (dot) fr