Pseudonymisation: Feedback from the CNIL

Monir Azraoui – CNIL
Technology expert
December 9th, 2021
About the CNIL: Missions

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

- Team of +200 Law & Technology experts
Pseudonymisation in the GDPR

Recital (26):

Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person.

Article 4 (5):

'pseudonymisation' means 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.
PSEUDONYMISATION VS. ANONYMISATION

HOW DATA CONTROLLERS SEE IT

HOW THE GDPR SEES IT

WHAT DATA CONTROLLERS WANT

Anonymised data

Personal data

Anonymised data

Personal data

Anonymised data

Personal data

Pseudonymised data

Pseudonymised data

Pseudonymised data

Anonymised data

Anonymised data

Anonymised data

GDPR

GDPR

GDPR

Monir Azraoui - CNIL - mazraoui (at) cnil (dot) fr
Audits reveal bad practices...

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

Hashing as an anonymisation mechanism

<table>
<thead>
<tr>
<th>name</th>
<th>job</th>
<th>city</th>
<th>car</th>
</tr>
</thead>
<tbody>
<tr>
<td>XXX</td>
<td>lawyer</td>
<td>Cannes</td>
<td>Jaguar</td>
</tr>
<tr>
<td>YYY</td>
<td>doctor</td>
<td>Nice</td>
<td>Lexus</td>
</tr>
</tbody>
</table>

Removing direct identifiers as an anonymisation mechanism

- Misunderstanding of pseudonymisation/anonymisation

- Weak cryptographic keys
  - SHA-1
  - MD5
  - Obsolete hashing algorithms

- Bad implementation of pseudonymisation mechanisms

- Inadequate measures to protect “additional information”
  - No plan to refresh cryptographic keys
  - No technical measures to protect secrets/additional info
  - No organisational measures to ensure confidentiality of secrets/additional info
... and good practices!

- Ex-post audits to check whether DCs comply with the law
RELEVANT USE CASES
Use Case #1: partitioning data

- Police records transcription
Use Case #2: 2-level pseudonymisation

- French Health Insurance Information System

Sources of data:
- General regime
  - SSN
  - DoB
  - gender
- Special regime
  - SSN
  - DoB
  - gender
- Health care data
  - SSN
  - DoB
  - gender

Level 1 pseudonym:
- Level 1
- Level 1
- Level 1

Level 2 pseudonym:
- Level 2

Database

French Health Insurance Fund

Monir Azraoui - CNIL - mazraoui (at) cnil (dot) fr
Use case #3: hashing and key splitting

- Ticketing data in public transportation systems (fraud)
Main takeaways & Recommendations

- Pseudonymisation often mistaken with anonymisation by DC
  - Misunderstanding of the benefits of pseudonymisation by DC

- Practical pseudonymisation: a variety of techniques
  - Have your pseudonymisation procedure evaluated by experts
  - Implement tailored privacy & security measures to your processing

- Determine risks of re-identification
  - residual
THANK YOU