Test data preparation for GDPR compliance

Presented by Florence ROLLAND-SOULIER
GDPR core principles

Data subject rights
GDPR core principles – data subject rights

- Breach Notification (within 72 hours)
- Right to Access by the Data Subject
- Right to be Forgotten / Right to Erasure
- Right to Object Processing & Right to Restriction of Processing
- Data Portability
- Privacy by Design

→ How does it impact sample datasets used for tests?
→ What can be considered as “Personal data”?
GDPR core principles

Right to be Forgotten
GDPR core principles – Right to be Forgotten

• How do we know when we should remove data from a dataset?

• How long should we keep our datasets?

• Conditions for erasure?
GDPR core principles

Right to Object Processing
& Right to Restriction of Processing
GDPR core principles – Right to be Object

• Articles 18 and 21 are related to each other

• Using data for testing purposes can be considered processing

• The data subject can oppose the erasure of the personal data and request the restriction of their use instead
GDPR core principles

Privacy by Design
GDPR core principles – Privacy by Design

- Data protection through technology design
- Pseudonymisation is not enough
- Basic data obfuscation is certainly not enough
- Why do we use production data, or data derived from production data, for testing?
GDPR core principles – Privacy by Design

• What can we use to identify an individual?

Think about a few examples and write them down.
GDPR core principles – Privacy by Design

• Think beyond social security numbers, names and dates of birth.

• Can you identify an individual with combination of « insensitive » data types?

• In this dataset, what can be used to identify an individual?
### GDPR core principles – Privacy by Design

<table>
<thead>
<tr>
<th>A</th>
<th>Social security number</th>
<th>Name</th>
<th>Surname</th>
<th>Address</th>
<th>Postcode</th>
<th>City</th>
<th>Birth date</th>
<th>Car plate</th>
<th>Invoice number</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>29021500044266</td>
<td>Laurence</td>
<td>Joly</td>
<td>292 passage Michael Henry</td>
<td>84570</td>
<td>MORMOIRON</td>
<td>1956-07-28</td>
<td>ZS-795</td>
<td>390</td>
</tr>
<tr>
<td>1</td>
<td>23711894582877</td>
<td>Laura</td>
<td>Vincent</td>
<td>184 avenue Annie Garcia</td>
<td>14700</td>
<td>ST PIERRE CANIVET</td>
<td>2009-04-06</td>
<td>OV-818</td>
<td>393</td>
</tr>
<tr>
<td>2</td>
<td>14506405780899</td>
<td>Frédéric</td>
<td>Martin</td>
<td>333 impasse Adrien Aubry</td>
<td>38008</td>
<td>ST MARCEL BEL ACCUEIL</td>
<td>1902-05-18</td>
<td>UB-898</td>
<td>393</td>
</tr>
<tr>
<td>3</td>
<td>17109843507232</td>
<td>Guichard</td>
<td>Joel</td>
<td>242 avenue Joel Menard</td>
<td>18400</td>
<td>LINERY</td>
<td>1974-11-10</td>
<td>XD-733</td>
<td>393</td>
</tr>
<tr>
<td>4</td>
<td>13595638086957</td>
<td>Benoît</td>
<td>Barthelemy</td>
<td>54 impasse Cécile Rodriguez</td>
<td>28170</td>
<td>THIEMERT GATELLES</td>
<td>1981-02-16</td>
<td>ZP-660</td>
<td>393</td>
</tr>
<tr>
<td>5</td>
<td>19009142940099</td>
<td>Jean-Paul</td>
<td>Vasseur</td>
<td>323 passage Michel Baron</td>
<td>63000</td>
<td>PERONNET SUR ALLIER</td>
<td>1952-11-02</td>
<td>XD-730</td>
<td>393</td>
</tr>
<tr>
<td>6</td>
<td>18103039272730</td>
<td>Michael</td>
<td>Aubry</td>
<td>304 route Jean-François Roussel</td>
<td>38310</td>
<td>BELLOCHE</td>
<td>2005-02-21</td>
<td>YY-384</td>
<td>393</td>
</tr>
<tr>
<td>7</td>
<td>185054737876</td>
<td>Laurent</td>
<td>Vasseur</td>
<td>231 boulevard Raphaëlle Lambert</td>
<td>45310</td>
<td>LA CHAPELLE OZERAIN</td>
<td>2010-03-05</td>
<td>MA-520</td>
<td>393</td>
</tr>
<tr>
<td>8</td>
<td>14208323533488</td>
<td>Laurent</td>
<td>Mounin</td>
<td>56 impasse Laura Klein</td>
<td>52170</td>
<td>FONTAINES SUR MARNE</td>
<td>1939-03-21</td>
<td>FF-719</td>
<td>393</td>
</tr>
<tr>
<td>9</td>
<td>12505870045020</td>
<td>Georges</td>
<td>Leroux</td>
<td>204 avenue Gilles Barre</td>
<td>14480</td>
<td>LANTHEUIL</td>
<td>1974-08-15</td>
<td>IB-641</td>
<td>393</td>
</tr>
<tr>
<td>10</td>
<td>24103961284389</td>
<td>Charlotte</td>
<td>Baron</td>
<td>259 passage Claude Guerin</td>
<td>67140</td>
<td>ITTERSWILLER</td>
<td>1941-08-21</td>
<td>AB-150</td>
<td>393</td>
</tr>
<tr>
<td>11</td>
<td>12901579815024</td>
<td>Laurent</td>
<td>Tessier</td>
<td>114 impasse Jocelyne Bouvier</td>
<td>56820</td>
<td>CROIXANVEC</td>
<td>1900-08-10</td>
<td>ZJ-448</td>
<td>393</td>
</tr>
<tr>
<td>12</td>
<td>21091295342857</td>
<td>Sophie</td>
<td>Miller</td>
<td>186 place Nadine Blanc</td>
<td>69440</td>
<td>TAUYNERS</td>
<td>1971-03-31</td>
<td>PS-429</td>
<td>393</td>
</tr>
<tr>
<td>13</td>
<td>22807584702931</td>
<td>Sylvie</td>
<td>Laporte</td>
<td>102 passage Charlotte Bernard</td>
<td>85320</td>
<td>CORNE</td>
<td>1961-09-27</td>
<td>RB-552</td>
<td>393</td>
</tr>
<tr>
<td>14</td>
<td>23802121042466</td>
<td>Aurélie</td>
<td>Noel</td>
<td>318 boulevard Alexandre Remy</td>
<td>63490</td>
<td>SUGERES</td>
<td>1948-12-29</td>
<td>FA-529</td>
<td>393</td>
</tr>
<tr>
<td>15</td>
<td>29203833766132</td>
<td>Christiane</td>
<td>Gauthier</td>
<td>171 impasse Isabelle Leandrat</td>
<td>30300</td>
<td>BOURG DE SIROD</td>
<td>1981-08-18</td>
<td>NO-230</td>
<td>393</td>
</tr>
</tbody>
</table>

© All rights reserved
How to solve the privacy issue

Put on your white hat
How to solve the privacy issue when you test with production data

• Analyze, clean and organize your data

• Minimize your datasets

• Do not store datasets that are bigger than necessary

• When it is possible, use randomly generated data (NOT production data)
Advocate for Privacy by Design

Why does it matter?
Advocate for privacy by design