Subject Access Request and Proof of Ownership

Cédric Lauradoux

October 25, 2019

Laws on Personal Data

in France and EU

- 1976: Loi Informatique et Liberté.
- 1995: Directive 95/46/EC.
- 2002: Directive 2002/22/EC.
- 2016: EU-US Privacy Shield and Umbrella.
- **2016: Regulation (EU) 2016/679 or GDPR.**
- **But what are personal data?**

Personal data

GDPR definition

‘personal data’ means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;

What is a personal data?

- A personal data can be **anything**. The definition is intended to be very broad. It has lot of consequences.
- A personal data can be a objective or subjective (comments or evaluations).
- Notice that the information can be true or not. Subjectivity is a source of false/incorrect data but poor data quality or errors also.

What is the relation between data and person?

- An information can be bind to a person by its content or purpose or result.
  - ▷ **Content:** the information is given directly about the person.
  - ▷ **Purpose:** the information is used to evaluate the behavior of a person.
  - ▷ **Result:** the exploitation of the information has an impact on the person.

How does the data identify the person?

- An identified person can be distinguished from a group of persons.
- **Direct identification** provides the true identity of a person: his/her real name and any additional information that can remove any ambiguity (possible namesake).
- **Indirect identification** is related to the concept of identifier.

Identifiers

- An identifier is a value that identifies an element within an identification scheme. Uniqueness is an important question.
- **Absolute uniqueness:** an identifier is assigned only to one element.
- **Relative uniqueness:** at a given time t, an identifier is assigned only to one element.
  (re-allocation is authorized)

Quasi-identifiers

- A quasi-identifier is well correlated with an element but several elements can share it.
- A quasi-identifier can identify uniquely an element in a small group (no collision).
- Several quasi-identifiers can create a profile which is considered as a (absolute) unique identifier!

Who is concerned?

- It applies to **a natural person**: adults or children.
- It does not apply to the data of dead persons neither legal persons.
- Now you know how to define personal data!

Example

Data	Perso. or not ?
My tax number	✓
INRIA Logo	✗
Color of my tshirt	✓
URL of my webpage	✓
IP address	?

- **What is the status of an IP address?**

Are IP addresses personal data or not?

Address type	Perso or not ?
Static IP address	✓ (??)
VPN IP address	✗
Free Proxy IP address	✗
Dynamic IP address	✗ or ✓

- What is the opinion of the CJEU?
  - Case C-131/12 – Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD)
  - Case C-582/14 – Breyer v Bundesrepublik Deutschland

Can we create a test?

- **Solution 1:** ask the CNIL.
  Too long. . .
- **Solution 2:** ask the CJEU.
  Even longer. . .
- **Solution 3:** do it yourself!
  Collect data from a sample. . . (INSEE)
- **Solution 4:** Personal data $\leftrightarrow$ rights
  Try to exercise your rights! (Lauradoux)

Rights of the data subject (I)

Article 15-22

- **To be informed:** know what's going on.
- **Access:** get your data.
- **Rectification:** correct your data.
- **Opposition:** stop any processing on your data.
- **Right to be forgotten:** bury results of search engines.

Rights of the data subject (II)

- **Portability:** get your data in readable format.
- **Objection against:**
  - Marketing
  - Profiling
  - Automated decision making
- **Class actions for reparations**

Right to be informed

- **A data controller must inform you:**
  - what data?
  - how?
  - why?
  - for how long. . .
- In this before the data collection/processing starts!

Right to access

- You can request your data and how they have been used to any data controller.
- The controller has 1 month* to answer your request.
- Charter of Fundamental Rights of The European Union (Article 8).
- It can be direct or indirect.

Subject Access Request (SAR)

Hello,

I have visited the website XXX using the IP address 194.199.28.40.

I want a copy of the data you collected during my visit of XXX according to Articles 15-21 of the GDPR.

Thank you
Cedric Lauradoux

Are IP addresses personal data or not?

- **I contacted 20 random websites. . .**
- **I was denied 20 times!**
  Really 20 times. . .
- **IP addresses are not personal data!**
  It is a bit more complicated than that. . .
- I became addicted to SAR. . .

Decathlon

July 2018

- I needed data so I submit a SAR at Decathlon.
  I was preparing a quiz for my MOOC on Privacy.

- Email to `vosdonneespersonnelles@decathlon.com`

  Bonjour,
  Je souhaiterai faire valoir mon droit d'accès aux données me concernant via la carte Decathlon numéro XZYZZZYVWZW et cela en respect des Articles 15-21 du RGPD.
  Merci

Answer I

- **First answer of the DPO:** I am on vacation!
- But I know I was dealing with a professional: Privacy Is Good For Business (link in his signature) https://sites.google.com/a/decathlon.com/personal-data-protection-quality/
- **Second answer of somebody else (named Marie).**

Answer II

- *Bonjour,*
  Afin que je puisse la traiter au mieux, j'ai besoin de savoir que c'est bien vous qui êtes à l'origine de la demande.

  Pourriez-vous me faire parvenir une copie de votre pièce d'identité ?

- Do you also want one kidney and my blood with it ?

Why I was mad!

- I never gave Decathlon my true identity for the fidelity program and I always pay with banknotes!
- There is **no reason** to give Decathlon my true identity to exercise my rights!
- **Their request is:**
  - abusive,
  - irrelevant!

Can we achieve something useful out of this?

- **What are the consequences for SAR?**
- **Why data controllers enforce bad policy?**
  DPAs are to blame for that. . .
- **What are doing others data controllers?**
  - We submit SARs to Top 50 Alexa websites
  - We submit SARs to 30 popular third parties
- **Can we do better?**
  We need real proof of data ownership!

What are the consequences for SAR?

- **Data Breach:** anybody can get your data.
  Impersonation or Incorrect disclosure.
- **Privacy Invasion:** you have to expose yourself.
  Abusive identity check.
- **Denial of Access:** you cannot get your data.
  Impossibility to authenticate.

Why data controllers enforce bad policy?

- **We are lazy!** We often obey to their request.
- Data controllers are used to receive copies of ID.
  Why?
- Data Protection Authorities (DPAs like CNIL) provides template to write SAR!
  Providing copies of ID is recommended by DPAs?

Investigations on DPAs

Methodology

- In December 2018, we review the websites of 28 DPAs in EU with the help of our colleagues.
- **Step 1:** we explain them the access right, and what the DPA of their home country.
- **Step 2:** they got 30 min to find some guidelines to write a SAR on the DPA website.

Investigations on DPAs

Results

- 17/28 DPAs provide guidelines/templates for SAR
- 4 require to provide systematically ID copy.
  6 are safe recommendations.
- About CNIL:
  - provide ID copy (before September 2018)
  - proportionality must apply! (after September 2018)

Investigations on Top 50 Alexa

- **We benchmark websites with SAR:**
  - to check if they respect proportionality,
  - to measure data accessibility.
- **We observe:**
  - 7 sites do abusive identity check
  - 7 sites uses emails to process SAR.
  - The GAFAM have the best procedures

Investigations on third parties

Results

- Same methodology for third parties.
- **We observe:**
  - 4 no answer at all,
  - 3 denials of access,
  - 4 incorrect disclosures,
  - 8 abusive identity check.

Can we do better?

- **Common misconception:**
  - Personal data $\rightarrow$ identity,
  - SAR $\rightarrow$ copy of an ID.

- GDPR applies when I use the pseudonym "doudou".
  - What if controller asks for my ID?
  - Cédric Lauradoux = doudou or,
  - Nataliia Bielova = doudou.
  - **True identity is irrelevant here!**

Who is doudou? or. . .

- **Bad question:** *Who are you?*
  Not relevant and abusive in this case.

- **Good question:** *Can you prove you are doudou?*
  We need to establish **ownership**.

How to prove ownership

- The proof must be proportional to the knowledge of the controller on the subject.
- **The proof can be:**
  - the success to an authentication protocol,
  - an element provided by a trusted third party
    - ◇ verifiable attestation,
    - ◇ the copy of an ID.
  - anything else. . .
- The request must identify uniquely the subject in the information system (**unambiguity**).

Why do my IP-based SAR fail?

- The controllers often say that they cannot verify if an IP address is associated to my identity.
- I try to contact ISPs or IP providers (VPN and free proxy) to get an attestation. . . no answer at all.
- IP addresses are believed to be personal data but you cannot exercise your rights on them because you cannot claim ownership.

What about cookies?

- You cannot claim ownership on your cookies!
  Cookie not personal data. . . IAB is so happy!

- Industrial cookies are bad for you! (too much sugar!)
  privacy-unfriendly by finality and by design.

- How do you get a cookie on your browser? (RFC 6265)
  - Subject $\leftarrow$ Controller
  - `Set-Cookie: ID=1234`
  - `ID=1234`

How to make homemade cookie?

- **Use a commitment scheme:**
  - Subject $\leftarrow$ Controller
  - `Set-Cookie: ID=1234`
  - Subject $\rightarrow$ Controller
  - ID=$\mathrm{Commit}(\texttt{value})$
- **Subject Access Request**: Open the commitment!

- GDPR is an obligation of mean.
- GDPR need to be an obligation of result.

Want more details. . .

- **Security Analysis of Subject Access Request Procedures**. C. Boniface , I. Fouad , N. Bielova , C. Lauradoux and C. Santos. Rome, Annual Privacy Forum 2019.

- **IP addresses are not personal data! (but they should be)**. C. Lauradoux. Submitted at IMC 2020.

- **Designing GDPR compliant identifiers.** N. Bielova, C. Castelluccia and C. Lauradoux. In preparation.