CCIN2P3

cnrs

Centre de Calcul
de l'Institut National de Physique Nucléaire
et de Physique des Particules

GDPR introduction for web developers

Jean-René ROUET <rouet@in2p3.fr>

Preamble

« This a an introduction to GDPR »

I work at CC-IN2P3 in the "applications" team and I work in the following themes:

⇒ Web/Software development

⇒ DevOps

⇒ Support

I am not a legal specialist and I'm not a GDPR specialist.

This presentation includes sections about the GDPR and other sections about best practices in software development and IT security.

The good practices are necessary for a good compliance with GDPR.

I used for this presentation: https://www.cnil.fr/en/gdpr-developers-guide

What is GDPR ?

The GDPR is the European data protection regulation. It came into force in 2018 and impacts all companies processing personal data on European residents (https://gdpr.eu/what-is-gdpr).

The GDPR pursues several ambitious objectives:

⇒ Standardize data protection regulations at European level.

⇒ Empower companies by developing self-control.

⇒ Strengthen the right of people (right to access, right to be forgotten, right to portability, etc.).

It's a legal text, ok, but how do I implement these goals in my software development ?

1-Be compliant with the GDPR

No matter what project you're working on, it's important to keep in mind that users' data and personal information are valuable and need to be protected throughout the lifecycle of the project.

Identify a person responsible for this issue.

⇒ he can be the DPO (Data Protection Officer) of your organization

Its agreement may be necessary in some cases.

⇒ For sensitive data (see slides about that)

Map, Trace and categorize stored data and associated processing

⇒ https://www.cnil.fr/en/gdpr-toolkit/record-processing-activities

⇒ This must be exhaustive

⇒ Personal data may be present anywhere (log files, assets, cache, database, files, …)

1-Be compliant with the GDPR

Prioritize actions to be carried out.

⇒ Identify the actions to be taken upstream of development

- ⇒ Validate that all the data you want to collect is really necessary for the purpose of your project
- ⇒ Legal basis on which your processing is based
- ⇒ Information provided to users
- ⇒ Contractual clauses binding you to contractors
- ⇒ Procedures for exercising rights
- ⇒ Measures to be implemented to secure processing

Manage risks.

⇒ If you identify that a data and/or processing is high-risk

- ⇒ Privacy Impact Assessment (PIA)

I don't think this is a case we encounter.

1-Be compliant with the GDPR

Create internal processes used during development

⇒ They must ensure that aspects impacting data protection are not forgotten

- ⇒ Security breach
- ⇒ Management of user requests (rectification and access)
- ⇒ Modification of the data collected
- ⇒ Change of provider
- ⇒ Data breach, …

⇒ Document GDPR compliance

2-Identify personal data

Some Examples

⇒ The following data is personal data

- ⇒ Last name, First name, Pseudo, Birth date
- ⇒ Photo, Voice record
- ⇒ Phone number, postal address, email address
- ⇒ IP Address, login, cookie id
- ⇒ Fingerprint, retinal impression
- ⇒ Car ID, Social security number, Passport number
- ⇒ App usage data, comments, ...

⇒ You can identify a person

- ⇒ From a single piece of data (first and last name)
- ⇒ Based on a cross-referencing of data (woman living at a particular address, born on a given day and a member of a particular association)

2-Identify personal data

Particularly sensitive data

Their collection is forbidden unless the person has given his or her consent (active, explicit approach, preferably in writing, free, specific and informed).

The following data is sensitive data:

⇒ Health Related

⇒ Relating to sex life and sexual orientation

⇒ Relating to an alleged racial and ethnic origin

⇒ Political opinions, religious or philosophical beliefs, trade union membership

⇒ Genetic data

2-Identify personal data

Anonymization or Pseudonymization.

⇒ You can use anonymization or pseudonymization to keep certain data over a long period of time for statistical purposes.

⇒ See https://www.cnil.fr/en/sheet-ndeg1-identify-personal-data for more information

3-Prepare you work

You need to incorporate GDPR principles into software development.

Methodologic choice

⇒ Privacy by design https://www.edpb.europa.eu/our-work-tools/documents/public-consultations/2019/guidelines-42019-article-25-data-protection_en

⇒ You are agile ? https://cyber.gouv.fr/sites/default/files/2018/11/guide-securite-numerique-agile-anssi-pa-v1.pdf (in French)

⇒ Conduct a Privacy Impact Assessment (PIA) if mandatory

Technologic choice

⇒ Data protection may impact your architecture (e.g. decentralized or not) or functionality (anonymization and data minimization)

⇒ Minimum security requirements (e.g. password complexity)

3-Prepare you work

Technologic choice

- ⇒ Keep control
  - ⇒ Keep simple
  - ⇒ Increase complexity by little steps
- ⇒ Have multiple lines of defense
  - ⇒ Form data validation
  - ⇒ But data query protection also
- ⇒ Tools and practice
  - ⇒ Use programming language standards
  - ⇒ OWASP for the web development
  - ⇒ Use frameworks and libraries backed by a large community

4-Secure your development environment

Evaluate the risks about tools you use

⇒ Particularly for the SAAS services

Secure your servers and workstations

⇒ Consistently and reproducibly

⇒ Upgrade your servers and workstations

Emphasize access management and traceability of operations

⇒ SSH keys

⇒ Strong authentication

⇒ Trace machine access

⇒ Don’t use a generic account

5-Manage your source code

Use Git (or whatever)

⇒ With secure access

⇒ With differentiated permissions depending on the developer's roles

⇒ With a backup

⇒ With procedures to manage work in //

Keep an eye on your source code

⇒ Don't store secrets in your source code

6-Make an informed choice of architecture

Identify the journey of personal data within your architecture

The data can be transmitted through online services or managed by a service provider

For more information: https://www.cnil.fr/en/sheet-ndeg5-make-informed-choice-architecture

7-Securing your web site, your application and your servers

Securing the communications

⇒ Use TLS 1.2 TLS 1.3

⇒ bye to old programming language version

⇒ TLS is mandatory

⇒ You need server certificate (don't manage it manually)

Securing authentications

⇒ Follow recommendations on passwords

⇒ Don't store then in clear text (you don't ?)

⇒ If cookie is used

⇒ Force HTTPS via HSTS, secure flag and HttpOnly flag

⇒ Disable obsolete cryptographic suites on your systems

⇒ For administrators, enforce authentication

⇒ Limit access to administrative tools

⇒ Use VPN

Securing infrastructure

- ⇒ Make backup (encrypted and regularly verified)
- ⇒ Limit the size of software stack
  - ⇒ Critical updates is not an option
  - ⇒ Automate a vulnerability watch (NVD)
- ⇒ Use vulnerability detection tools
  - ⇒ On servers
  - ⇒ In continuous integration
- ⇒ Restrict physical access to servers
- ⇒ Protect access to databases by network filtering for example
  - ⇒ Use one account for one database
  - ⇒ Revoke administrative rights if possible
  - ⇒ Injection SQL
- ⇒ Use isolation and non root containers

8-Minimize the data collected

Think ahead about the data collected and try to narrow it down to what is strictly necessary

⇒ Document this thinking

⇒ If specific data is not needed for a category of people, do not collect it

⇒ Reduce accuracy if possible (birth year instead of birth date)

⇒ For particularly sensitive data (health), the easiest way is to do without it

⇒ Put the minimum amount of data in the log files

  ⇒ IP Address rather than IP Address and Username

⇒ Some data can improve the user experience but is not strictly necessary.

The user must be able to object to geolocation for example,

and geolocation must only be kept for as long as is strictly necessary for its use.

⇒ Associate a retention time with each data category

8-Minimize the data collected

Think about automatic deletion process

⇒ Automatic purge

⇒ Physically erase data with specialized tools

⇒ If the data is still useful, you can reduce its sensitivity by pseudonymizing it or even anonymizing it

⇒ Log the automatic erasure procedures. The corresponding logs can be used as proof of deletion of data

9-Manage user profiles

Good practices

⇒ One id for one user (developer or user)

⇒ Authentication needed before data access

⇒ Access management by user, he must only access the data he actually needs

⇒ Logging can include activity trace, anomaly detection, or security-related event detection

⇒ Code auditing and penetration testing are also a good idea

Managing Entitlement Profiles

⇒ Document and/or automate user movement, registration and deregistration procedures

⇒ Regular review of granted rights

⇒ Don't use root/admin access for simple operation

10-Control your libraries and SDKs

Make an informed choice

⇒ Assess the value of adding each dependency

⇒ Choose maintained software, libraries and SDKs

- ⇒ For open source, choose project with an active community, regular updates and good documentation
- ⇒ For commercial solutions, contractually ensure that the code will be maintained and updated for the life of your project
- ⇒ Take privacy into account. Some SDKs or libraries pay for themselves by using personal data collected from the applications or sites on which they are integrated. Make sure that such third parties comply with applicable laws regarding personal data, including a mechanism for obtaining user consent.
- ⇒ If you use cryptographic mechanisms, it is strongly discouraged to implement cryptographic algorithms or protocols yourself

10-Control your libraries and SDKs

Evaluate the selected elements

⇒ Read the documentation and change the default configurations

⇒ Audit your libraries and SDKs

⇒ Map your dependencies

⇒ Beware of typosquatting and other malicious techniques

Maintain libraries and SDKs

⇒ Use dependency management systems

⇒ Manage updates to your dependencies with a documented procedure

⇒ Be aware of the versions of libraries and SDKs at the end of support

⇒ Check the status of open-source projects,

11-Ensure quality of the code and its documentation

Document code and architecture.

Check the quality of your code and its documentation.

For more information: https://www.cnil.fr/en/sheet-ndeg10-ensure-quality-code-and-its-documentation

12-Test your applications

Automate testing.

Watch out for your test data !!!

⇒ Real production data should not be used during the development and testing phase

⇒ This is a change in the destination of the data collected

⇒ You should have a dummy test data

⇒ If you need to do performance tests that could require a volume of data comparable to production data, anonymize the data. Anonymization must be done on the production infrastructure, we don't copy the data outside before anonymizing it.

For more information: https://www.cnil.fr/en/sheet-ndeg11-test-your-applications.

13-Inform users

The transparency principle of the GDPR requires that any information or communication relating to the processing of personal data should be concise, transparent, comprehensible and easily accessible in plain and simple language.

Who should be informed and when should it be done?

⇒ The persons concerned must be informed
  ⇒ both in the case of direct data collection or when it is collected via devices or technologies for observing people's activity
  ⇒ only in the case of indirect collection of personal data, when the data is not collected directly from individuals

⇒ When this information is needed:
  ⇒ at the time of data collection in the case of direct collection;
  ⇒ as soon as possible in the case of indirect collection (in particular during the first contact with the person) and at the latest, within one month (with some exceptions);
  ⇒ in the event of a substantial change or special event. For example: new purpose, new recipients, change in the way rights are exercised, data breach.

13-Inform users

What information do I need to give ?

⇒ The identity and contact details of the organization collecting the data

⇒ Purposes

⇒ The legal basis, see later

⇒ The mandatory or optional nature of the collection

⇒ The recipients or categories of recipients of the data

⇒ Data Retention duration

⇒ The existence of the rights of the persons concerned and the means of exercising them. The rights of access, rectification, erasure and limitation are applicable for all processing

⇒ Contact details of the Data Protection Officer, if there is

⇒ The right to lodge a complaint with the CNIL if the processing is in France

⇒ Transfer outside the EU

⇒ …

13-Inform users

In what form should I provide this information?

⇒ Easy to access

⇒ Clear and comprehensible manner (accessibility)

⇒ In a concise manner (avoid pitfall of a flood of information)

⇒ Distinguishable from information that is not specifically related to privacy (contractual clauses, general terms, …)

13-Inform users

What communication should be made when data security is compromised?

An organization may mistakenly or negligently suffer, accidentally or maliciously, a personal data breach, i.e. the destruction, loss, alteration or unauthorized disclosure of data.

⇒ In this case, the organization must report the violation to the local data protection agency within 72 hours if it is likely to pose a risk to the rights and freedoms of individuals

⇒ If these risks are high, the organization must also inform the persons concerned as soon as possible and provide them with advice on how to protect their data (e.g. cancellation of a compromised bank card, modification of a password, modification of privacy settings, etc.)

⇒ Notification of the violation to the CNIL must be made via the CNIL website (in France)

13-Inform users (a simple example)

The information collected on this form is recorded in a computerized file by [identity and contact details of the data controller] for [purposes of the processing]. The legal basis for processing is [legal basis for processing].

The data collected will be communicated only to the following recipients: [recipients of the data].

The data is kept for [data retention period provided for by the controller or criteria for determining it].

You can access your data, rectify it, request its deletion or exercise your right to restrict the processing of your data. (depending on the legal basis for the processing, also mention: You can withdraw your consent to the processing of your data at any time; You can also object to the processing of your data; You can also exercise your right to data portability)

Visit the cnil.fr website for more information about your rights.

To exercise these rights or for any questions about the processing of your data in this system, you can contact (if applicable, our data protection officer or the department responsible for exercising these rights): [email address, postal address, telephone number, etc.]

If, after contacting us, you believe that your "Data Protection Freedoms" rights have not been respected, you can file a complaint with the CNIL.

14-Prepare for the exercise of people's rights

The persons whose data you process have rights on his or her data: right of access, to rectification, to object, to erasure, to data portability and to restriction of processing.

You must give them the means to effectively exercise their rights and provide in your computer systems the technical tools that will allow their rights to be properly taken into account.

Preparing in advance how they will contact you and how you will deal with their requests will enable you to manage the exercise of these rights effectively.

For more information: https://www.cnil.fr/en/sheet-ndeg13-prepare-exercise-peoples-rights.

15-Define a data retention period

Personal data cannot be kept for an indefinite period of time: this must be defined according to the purposes of the processing.

Once this purpose has been achieved, the data should be archived, deleted or made anonymous (e.g. in order to produce statistics).

For more information: https://www.cnil.fr/en/sheet-ndeg14-define-data-retention-period.

16-Take into account the legal basis in the technical implementation

Processing of personal data must be based on one of the “legal basis” mentioned in Article 6 of the GDPR.

The legal basis of a processing operation is in a way the justification of the existence of the processing operation.

The choice of a legal basis has a direct impact on the conditions for implementing the processing operation and the rights of individuals.

Thus, anticipating the legal basis of the processing operations prior to any development will help you integrating the necessary functions to ensure that these processing operations comply with the law and respect the individuals' rights.

16-Take into account the legal basis in the technical implementation

Definition of the legal bases in the RGPD

In the context of a development for a private organization:

⇒ The contract: the processing is necessary for the performance or preparation of a contract between the data subject and the body carrying out the processing operation;

⇒ The legitimate interest: the organization has a “legitimate” interest in carrying out the processing and it is not likely to adversely affect the rights and freedoms of the data subjects;

⇒ Consent: the data subject has given his or her explicit consent to the processing.

If you are a public authority or a body pursuing tasks in the public interest, other legal bases may also be used:

⇒ The legal obligation: the processing is imposed by regulatory texts;

⇒ The public-interest mission: the processing is necessary for the performance of a task carried out in the public interest.

Finally, in very specific cases, protect of vital interests may be used as a legal basis, for example when processing is necessary to monitor the spread of epidemics or in cases of humanitarian emergency.

For more information: https://www.cnil.fr/en/sheet-ndeg15-take-account-legal-basis-technical-implementation.

17-A recent CC-IN2P3 example: Canyon development

https://canyon.cc.in2p3.fr

Canyon is a tool to organize the visit of the CC-IN2P3 during major events (such as “Journées du Patrimoine” in France)

Visitor register with an email and a name and then choose a visit slot.

The data is collected on the basis of legitimate interest and legal obligation.

17-A recent CC-IN2P3 example: Canyon development

Here is the GDPR text translated into English:

Data protection (GDPR)

Information:

The IN2P3/CNRS Computing Centre, in its capacity as data controller, informs you that the data collected directly or indirectly on this web service are subject to processing. When certain information is mandatory to access specific features of the site, this mandatory nature is indicated at the time of data entry. If you refuse to provide mandatory information, you may not have access to some of the services offered.

Collected data:

Personal data may be collected: surname, first name, email address, IP address of network connection.

Data destination:

The data collected on this website is used to create schedules for visits to the facilities of the IN2P3/CNRS Computing Centre.

Data recipients:

This data is accessible by the administrators of the website and by the agents of the IN2P3/CNRS Computing Centre who use the service and work in connection with the activity of visiting the facilities.

…

17-A recent CC-IN2P3 example: Canyon development

Here is the GDPR text translated into English:

…
Data retention period:

The data required for registration for visits is kept for one month after the event is finished and is then irretrievably deleted.
Data concerning CC-IN2P3 agents who use the service are kept for the duration of their activity in connection with the service.
Network connection data is kept for 1 year in accordance with the CNRS Trace Retention Policy.

Rights of individuals:

In accordance with the applicable legal and regulatory provisions, in particular Law No. 78-17 of 6 January as amended relating to information technology, files and freedoms and European Regulation No. 2016/679/EU of 27 April 2016 (applicable since 25 May 2018),

you have the following rights:
. Right of access, to know the personal data concerning you
. Updating your data if it is inaccurate
. Portability or deletion of your data
. Restriction of the processing of your data
. Objection, for legitimate reasons, to the processing of your data
. Withdrawal of your consent to the processing of your data
…

17-A recent CC-IN2P3 example: Canyon development

Here is the GDPR text translated into English:

…

You can exercise these rights by contacting:

Mr. Director,
IN2P3/CNRS Computing Center
21, avenue Pierre de Coubertin
CS 70202
69627 Villeurbanne

CNRS
Data protection service
2, rue Jean Zay
54519 Vandœuvre-lès-Nancy

If, after contacting us, you believe that your data protection rights have not been respected, you have the option of sending a complaint online or by post to the CNIL.

Resume

Personal data aspects:

⇒ Identify Personal data

⇒ Choice of architecture

⇒ Minimize the data collected

⇒ User profiles

⇒ Inform users

⇒ Data retention

⇒ Legal basis

⇒ Exercise of people's rights

Development & Security aspects:

⇒ Prepare your work

⇒ Secure your development environment

⇒ Manage your source code

⇒ Securing your web site, your application and your servers

⇒ Control your libraries and SDKs

⇒ Ensure quality of the code and its documentation

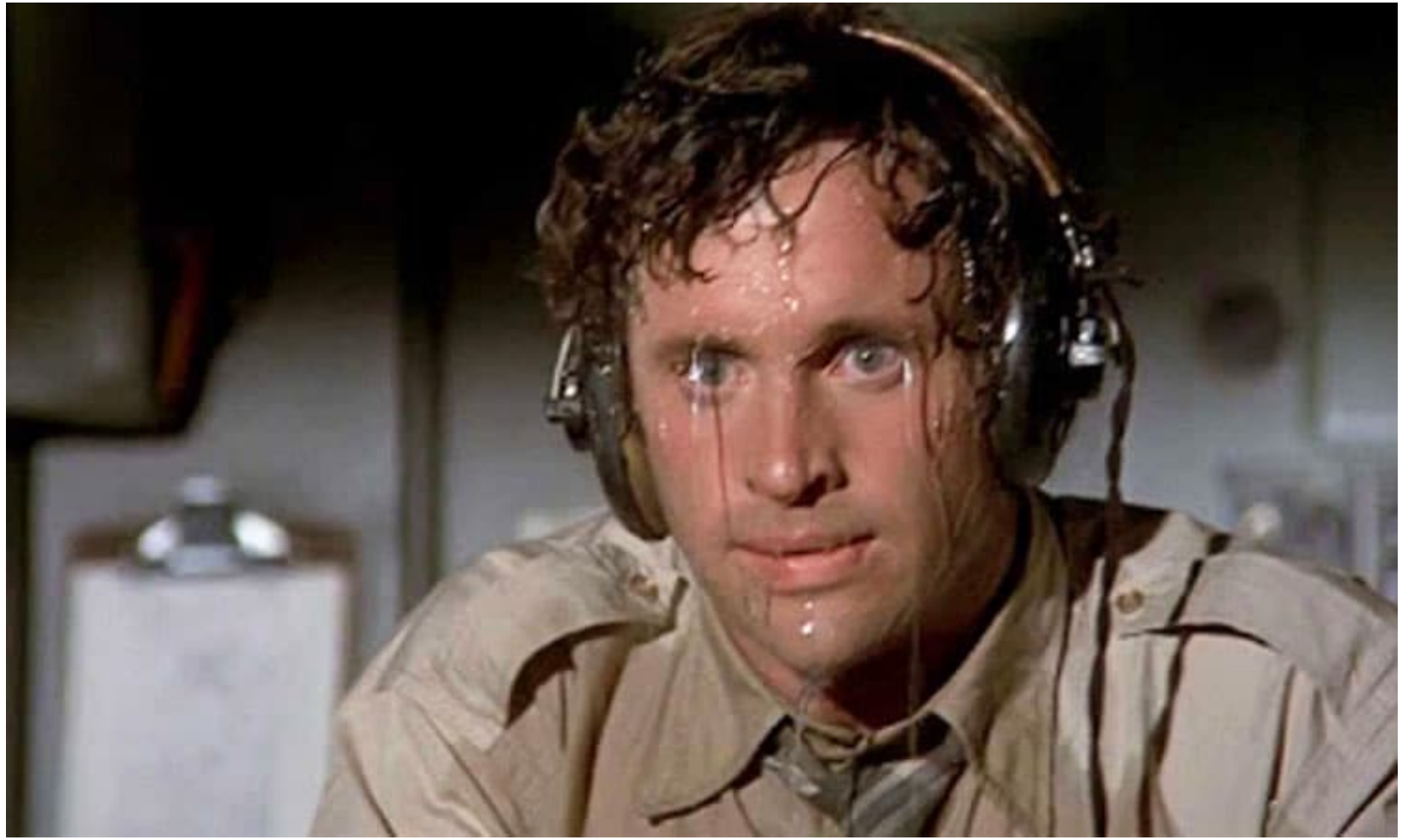
⇒ Test your applications

In case of data breach

You may have to prove that you have addressed the obligations of the GDPR in a relevant manner.

Questions

Thank you



References

- ⇒ https://gdpr.eu/what-is-gdpr
- ⇒ https://www.custup.com/introduction-gdpr-rgpd/ (in French)
- ⇒ https://www.cnil.fr/en/gdpr-developers-guide
- ⇒ https://www.cnil.fr/fr/la-cnil-publie-un-guide-rgpd-pour-les-developpeurs (in French)