Project : Session management backend - GDPR compliant application

LOUHOU Godwill

June 22, 2025

Contents

1. Project description

This project is about integrating GDPR compliant backend and frontend.

1.1. Concept

The concept of this app is pretty simple, since the emphasis is on the security aspect. It allows to create an account, and create posts with content on it. Other users cannot see the content of the posts.

The features available on the app :

- Users related actions :
  - Registering a new account
  - Logging in
  - Updating user email
  - Deleting the user
  - Cookies banner
- Posts related actions :
  - Creating posts
  - Updating posts
  - Getting posts
  - Deleting posts

The admin user can get information about the other users. Only this feature has been implemented, since actions on other users are trivial to implement.

The authentication is made using `JWT`, and authorization depends the user's `role`.

1.2. Tech stack

The stack chosen for this project is focused on being easy to code on and fast to ship.

- **Backend** :
  - Python
  - FastAPI
- **Frontend** :
  - Simple HTML/CSS/JS website
  - Bulma
- **Database** :
  - MongoDB Atlas
- **Other tools** :
  - Sonarqube
  - Docker
  - Snyk
  - pytest
  - curl
  - Github Actions
  - Wapiti3

2. Database

2.1. Arguments for database choice

- **Fully managed service**

MongoDB Atlas takes care of running, monitoring, and updating the database cluster for us. It makes the work way easier, with less aspects to think of.

- **Built-in high availability**

Atlas automatically replicates the data across multiple nodes (and even regions) so the app stays online through hardware failures or network outages.

- **Elastic scalability**

With higher tier subscriptions it is possible to scale up CPU, RAM, storage, or even shard the data with just a few clicks (or via API). There is no downtime.

- **Global distribution**

It allows to read/write replicas close to the users around the world to minimize latency and deliver a snappier experience.

- **Enterprise-grade security and compliance**

Atlas offers built-in encryption at rest and in transit, fine-grained access controls, compliance certifications (SOC2, GDPR, HIPAA, etc.), so there is not need to build that manually.

- **Automated backups and point-in-time recovery**

Continuous backups let us restore data to any point in time.

- **Rich ecosystem and integrations**

Atlas has an impressive amount of compatible applications, and the `API` for `python` is easy to use, making it a painless experience for the server development.

2.2. CIA Analysis

Provided by `ChatGPT`, about MongoDB Atlas :

2.2.1. Data Availability

- **Replica Sets & Multi-Region Clusters**: Automatic replication across members and regions; failover < 30 s under SLA.
- **Continuous Backups & Point-in-Time Restore**: Ongoing snapshots plus PITR ensure recovery from accidental deletes or regional failures.
- **Automatic Maintenance & Upgrades**: Rolling updates with zero-downtime options; health checks and alerts minimize unplanned outages.

2.2.2. Data Integrity

- **Write-Ahead Journaling**: Ensures committed writes survive crashes; data files replay journal on restart.

- **ACID Transactions**: Multi-document, multi-shard transactions guarantee atomicity, consistency, isolation, durability.
- **Schema Validation & Document Validation**: Enforce field types, ranges, and required fields at the database layer.
- **Audit Logging**: Immutable operation logs for forensic analysis and to detect unauthorized or anomalous modifications.

2.2.3. Data Confidentiality

- **Encryption At Rest**: AES-256 encryption of all storage volumes; customer-managed keys via AWS KMS, Azure Key Vault, or GCP KMS.
- **TLS Encryption In Transit**: Enforced TLS 1.2+ for all client-server and inter-node communications.
- **Network Isolation**: VPC peering, Private Endpoints, IP whitelisting, and firewall rules restrict traffic to authorized networks.
- **Role-Based Access Control (RBAC)**: Granular privileges per database, collection, and action; integration with LDAP and cloud IAM.
- **Field-Level Encryption**: Optional client-side encryption of specific fields so Atlas never sees plaintext values.

3. Back-end description

3.1. Project structure

```
backend/
├── app/
│   ├── database/
│   │   └── db.py
│   ├── middlewares/
│   │   ├── authenticationMiddleware.py
│   │   ├── authorizationMiddleware.py
│   │   └── sanitizationMiddleware.py
│   ├── models/
│   │   ├── connectionDetails.py
│   │   └── post.py
│   ├── utility/
│   │   └── utility.py
│   └── main.py
├── tests/
│   └── test_main.py
├── .dockerignore
├── .env
├── Dockerfile
├── pytest.ini
├── requirements.txt
└── cert.pem
```

3.2. Package : database

This package contains the file `db.py`. This file is used to create the connection to the MongoDB database. It is used in all the operations requiring database queries.

3.3. Package : middlewares

This package contains the 3 middlewares used in the backend.

3.3.1. authenticationMiddleware.py

This middleware is used to check if the user is authenticated, for the concerned routes. A handful of routes are exempt from this verification : `/login` and `/register` since they are accessed when the user is not yet logged in.

It works by checking if a `JWT` is present and valid for the protected routes. If not, doesn't allow to access the route.

3.3.2. authorizationMiddleware.py

This middleware is used to check if the user is authorized to access the concerned resource. It is mainly used for the `/users` route in this project, which allows the `admin` user to retrieve the information of all the users. The middleware can easily be used for other routes or requests, like for managing the users (CRUD).

3.3.3. sanitizationMiddleware.py

This middleware is used to sanitize the data incoming in the backend. It prevents requests from having a body bigger than `500Mib`, to prevent buffer overflow attacks per example. Other mecanisms can be added, like escaping all content passing though it.

3.4. Package : models

This package contains `pydantic` models used for the requests. It allows easy managing of the content parsed by the server.

3.4.1. connectionDetails.py

This model describes the connection informations used for the `login` and `register` actions. It also allows to check the robustness of the passwords used by the users.

```
email : str
password : str
connection_attempts : str
last_connection_attempt : str
role : str
```

3.4.2. post.py

This model describes the way the `posts` are represented. It facilitates every operation between the server and the database.

```
title : str
content : str
user : str
```

3.5. Package : utility

This package contains only one file, containing utility methods, to prevent boilerplate code.

3.5.1. utility.py

Contains methods used in various files across the project, mainly related to databases queries.

3.6. Package : tests

This package contains the test file for code coverage. It uses `pytest`.

3.6.1. test_main.py

Contains all the implemented tests.

3.7. main.py

The main file. Contains all the routes for the requests.

3.8. A note on security

The implementation of HTTPS has been made, using a self-signed certificate. This implies that it will not be reproducible if the certificate is not installed and defined as trusted on the host computer.

4. Front-end description

The front-end is a simple `HTML/CSS/JS` page. The `js` functions allows the app to send queries to the server and process the responses.

`textContent` is used in the code, preventing any XSS attack, and the server escapes any illegal character to prevent any database injection.

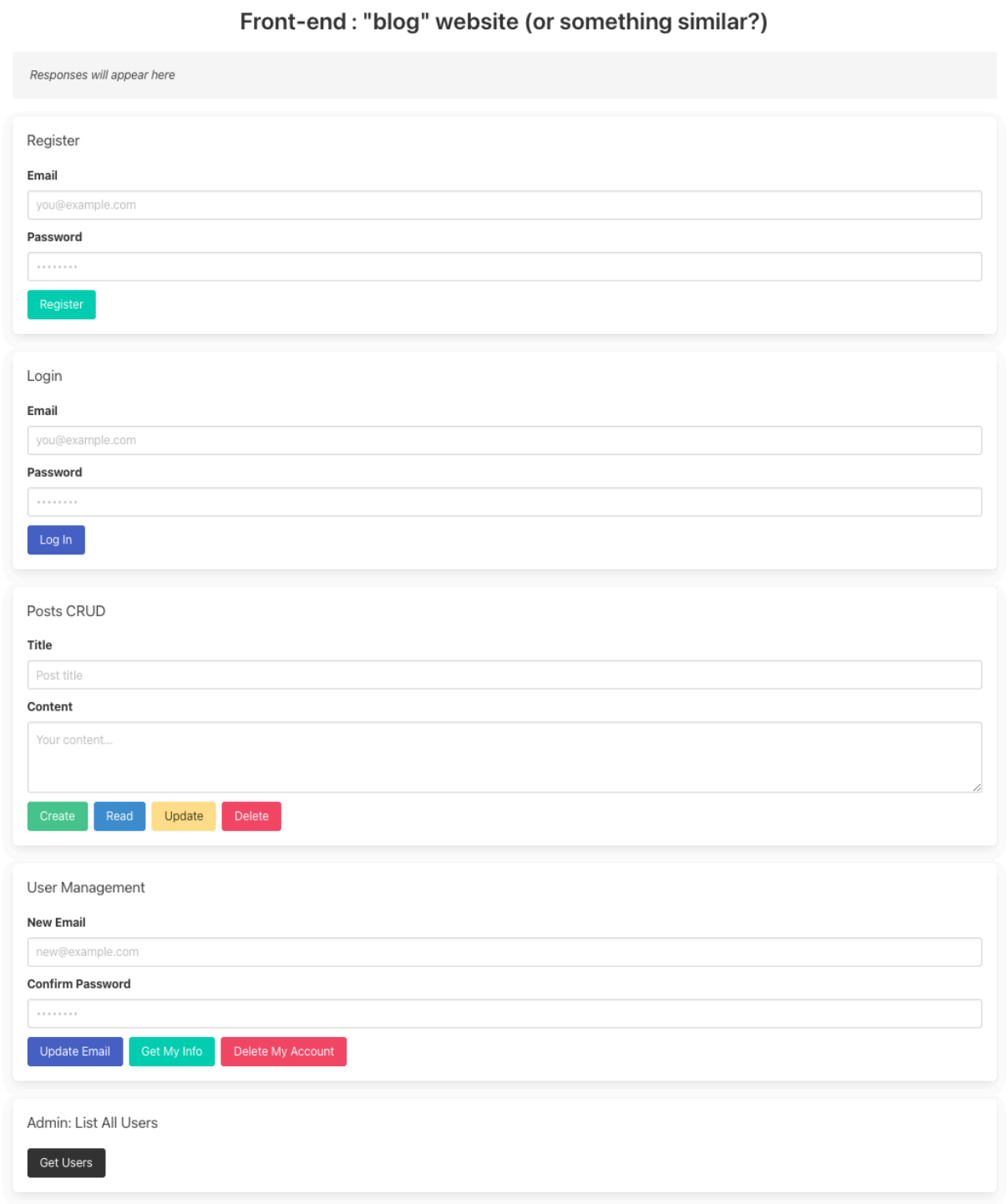


Figure 1: Front-end page

5. Architecture

Here is an overview of the different requests used in the project.

5.1. login : request

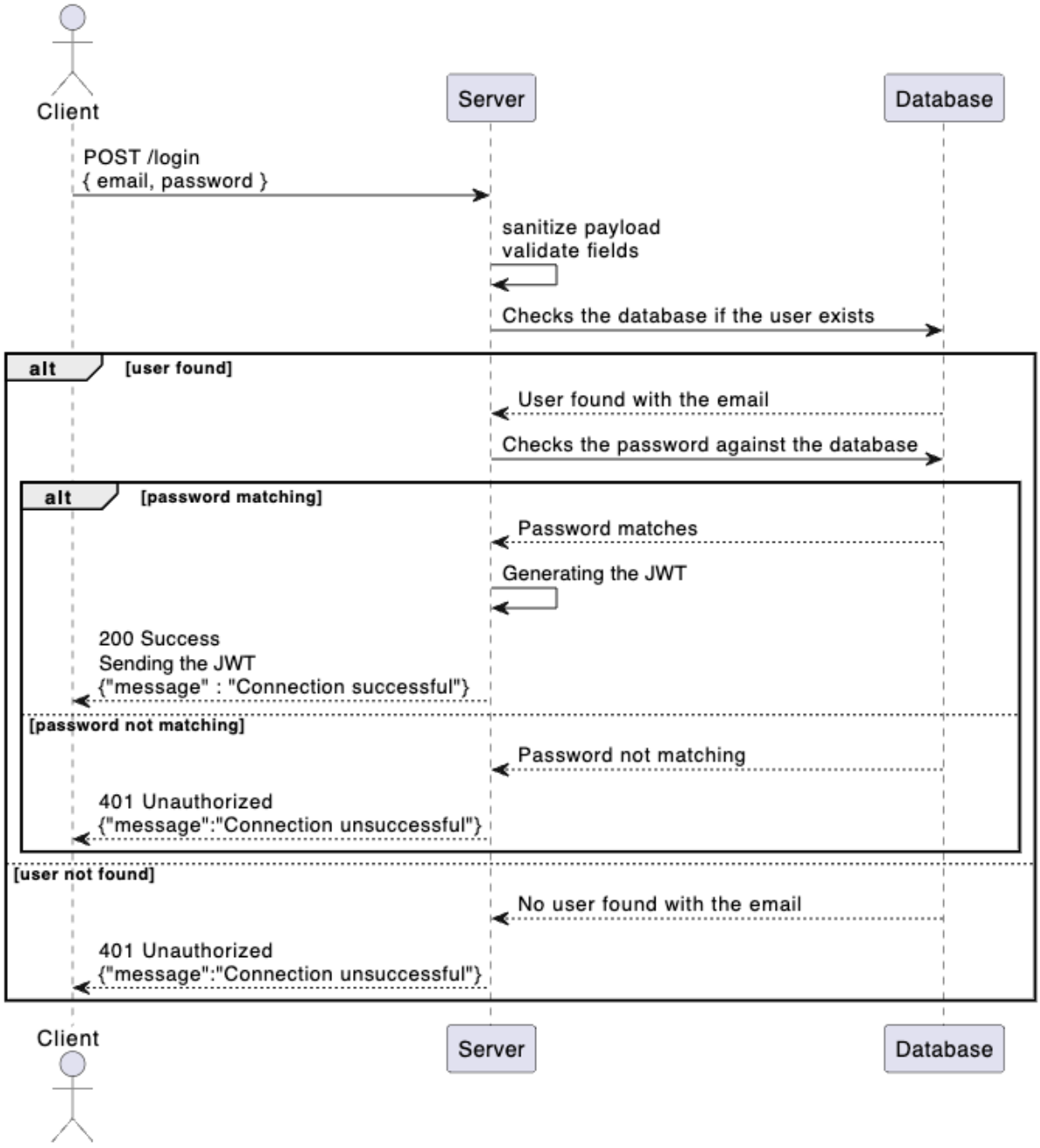


Figure 2: Login request

5.2. login : flow details

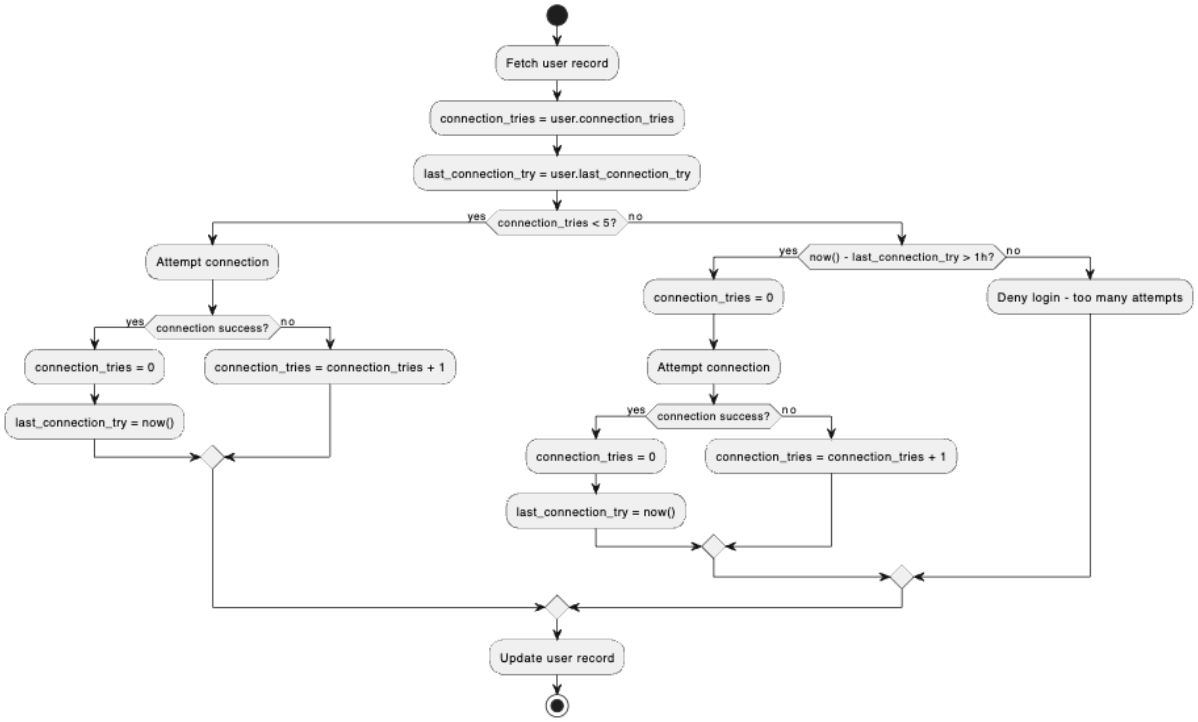


Figure 3: Login decision flow

5.3. register : request

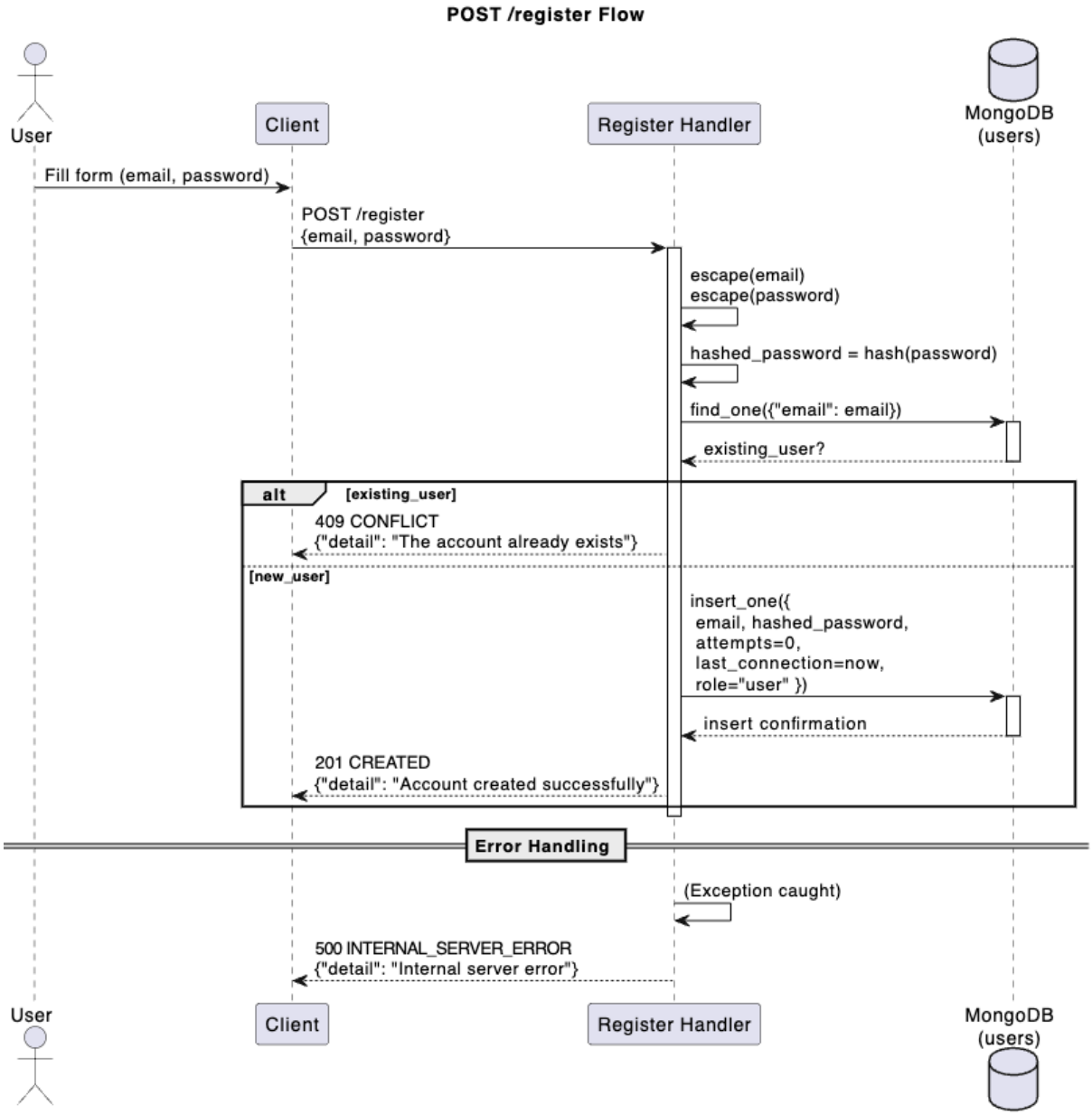


Figure 4: Register request

5.4. post : creation

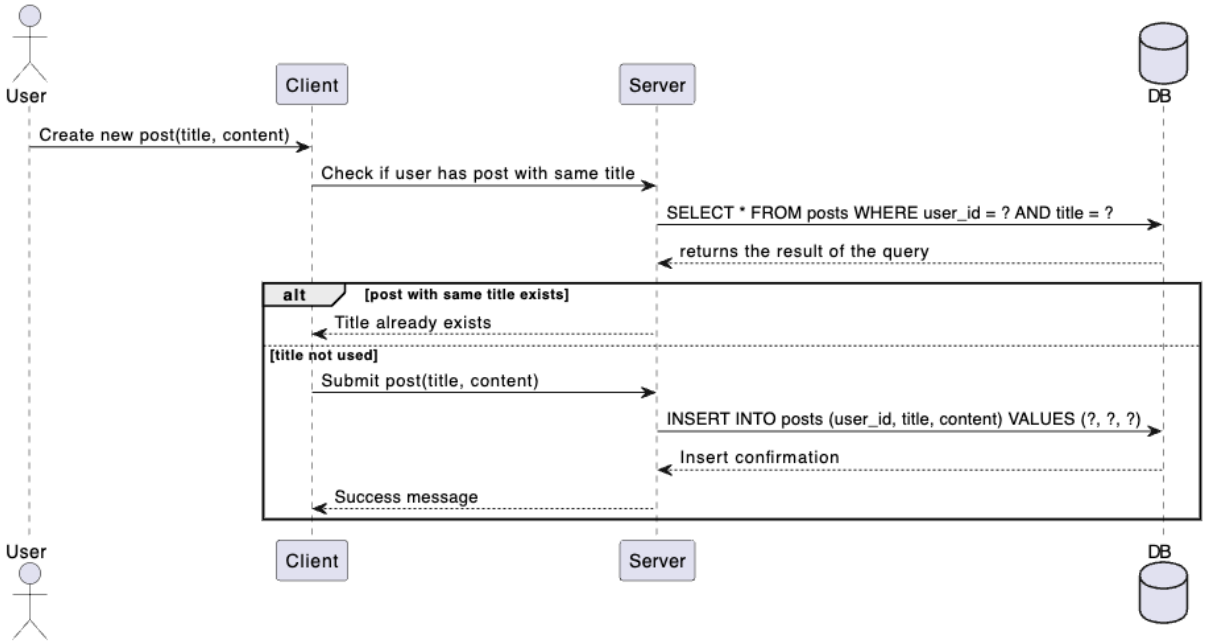


Figure 5: Post creation request

5.5. post : update

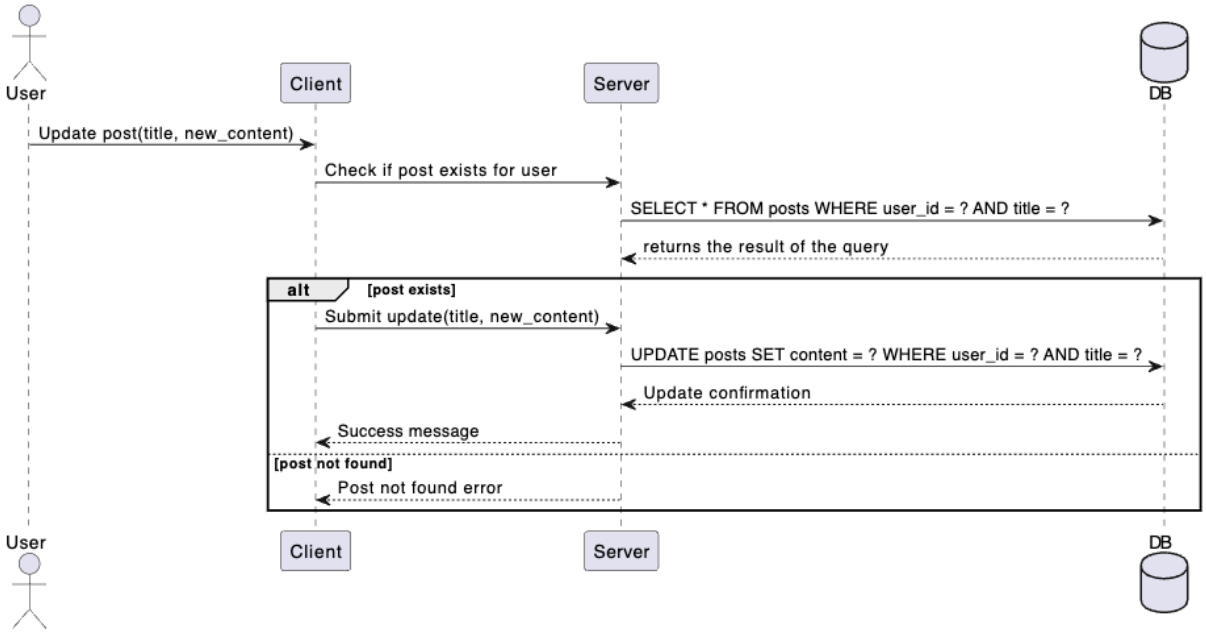


Figure 6: Post update request

5.6. post : deletion

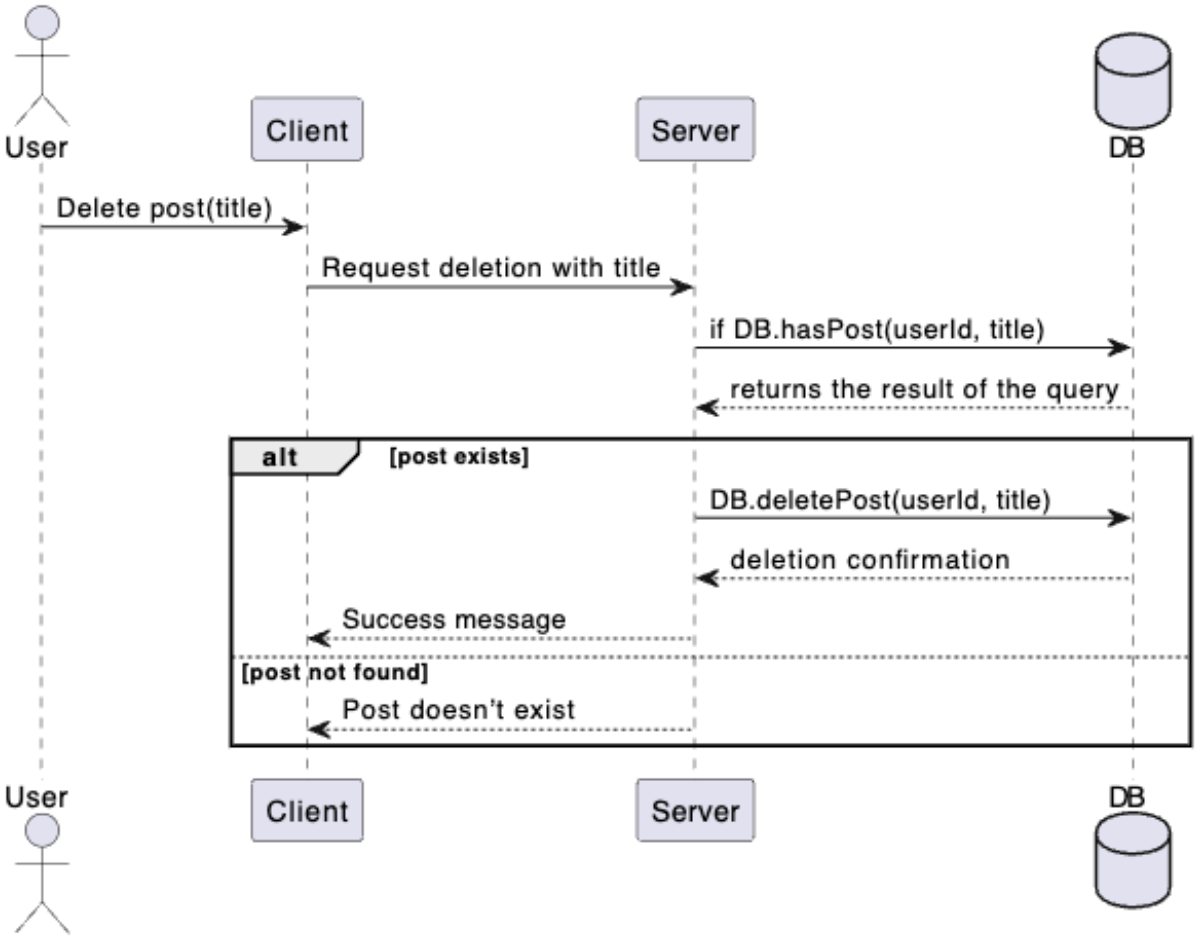


Figure 7: Post deletion request

6. GDPR Compliance and session management

The session management is made using a JWT. It is not detailed in this section.

In order to be GDPR compliant, there are a number of elements to be considered.

6.1. Cookies

6.1.1. Architecture

The architecture of how cookies are managed is shown in this diagram.

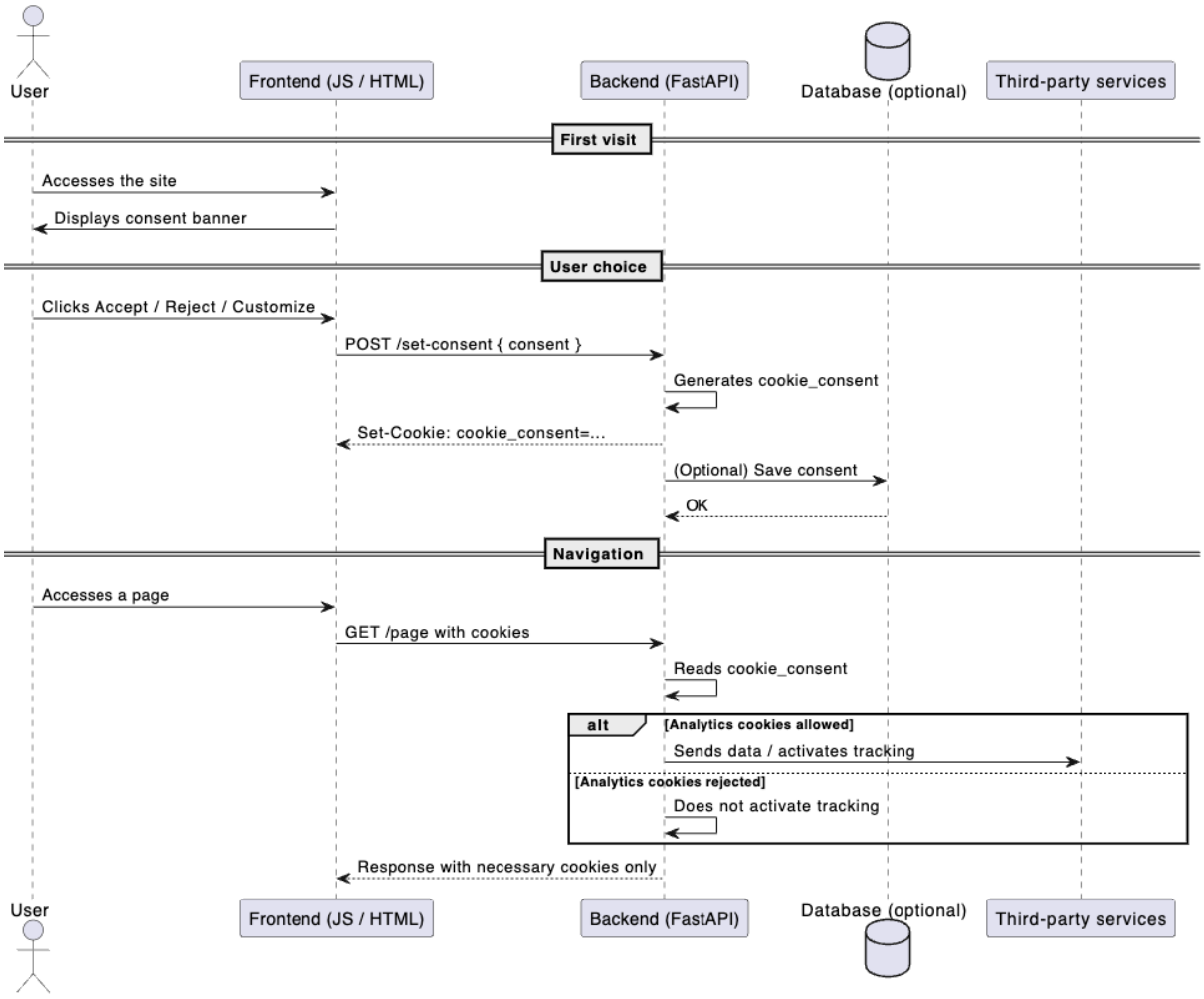


Figure 8: Architecture of cookie management

The user selects the cookies they want to enable, then their choices are sent to the server which will then add it to the user's "profile", allowing it to prevent the user from having to make their choices every time they connect to the website.

In this project, the tracking/marketing/partners cookies are not really generated, but adding them is not difficult. The goal of this project is only to show how they would be managed.

6.1.2. Cookies banners

First of all, a banner has to be present on the frontend to allow the user to choose which cookies they want to allow.

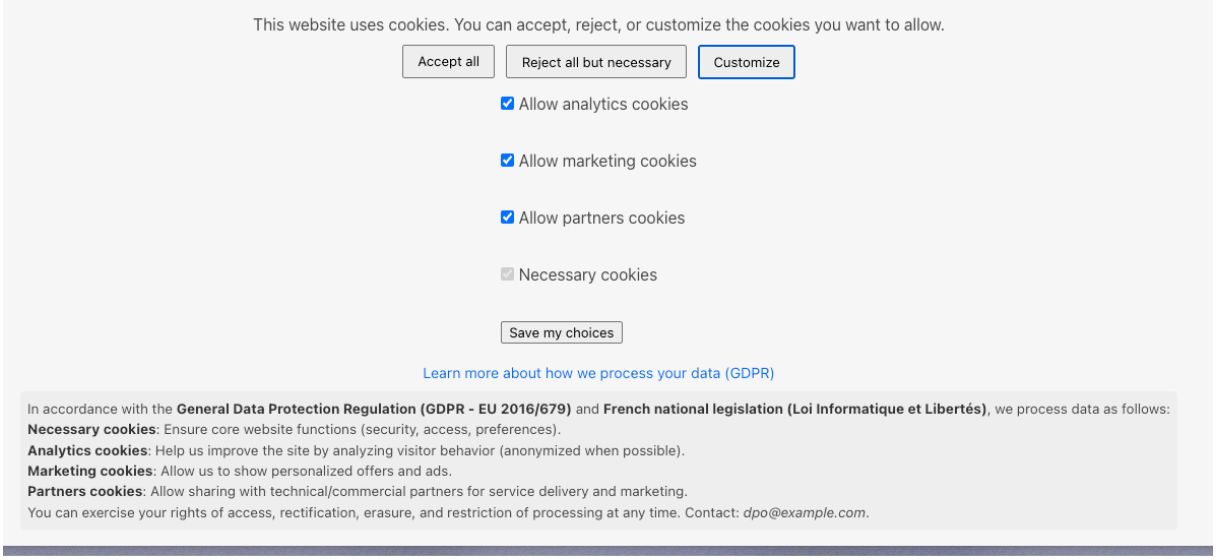


Figure 9: Cookies banner

We allow the user to choose which cookies they want to allow, with a description of each one of them, and the fact that they can exercice their rights about their data at any time.

If the user is not logged in, the cookie will only be local, but when the user is logged on, the cookie is synchronized with the database.

Figure 10: Database cookies choices

Then, there is a button available for the users to update their preferences. This allows to modify the data they allow to be collected etc.

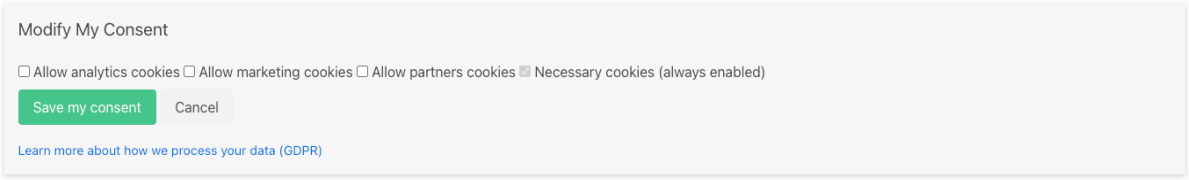
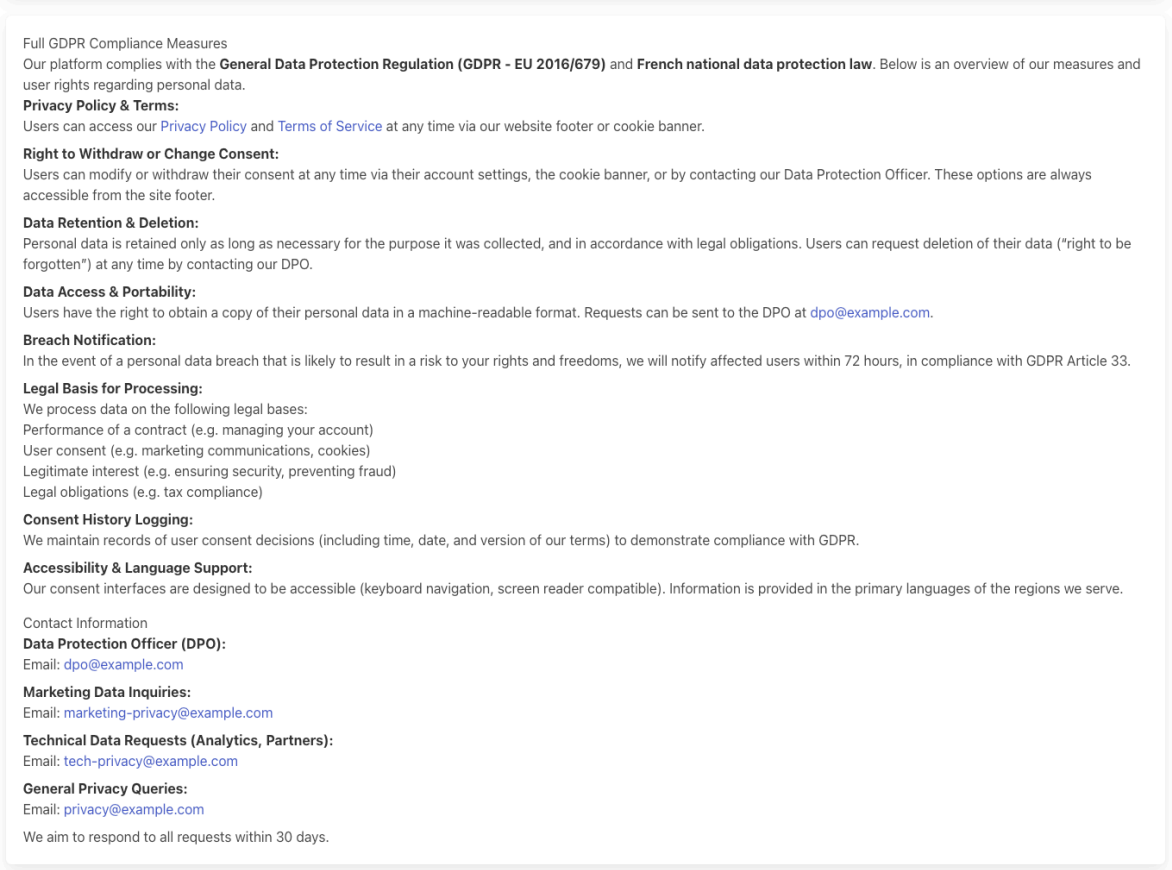


Figure 11: Cookies preferences update

6.2. Informations for the user

A section with the various contacts is available at the bottom of the website, and all the details about how we deal with GDPR compliance measures.



Full GDPR Compliance Measures

Our platform complies with the **General Data Protection Regulation (GDPR - EU 2016/679)** and **French national data protection law**. Below is an overview of our measures and user rights regarding personal data.

Privacy Policy & Terms:

Users can access our Privacy Policy and Terms of Service at any time via our website footer or cookie banner.

Right to Withdraw or Change Consent:

Users can modify or withdraw their consent at any time via their account settings, the cookie banner, or by contacting our Data Protection Officer. These options are always accessible from the site footer.

Data Retention & Deletion:

Personal data is retained only as long as necessary for the purpose it was collected, and in accordance with legal obligations. Users can request deletion of their data ("right to be forgotten") at any time by contacting our DPO.

Data Access & Portability:

Users have the right to obtain a copy of their personal data in a machine-readable format. Requests can be sent to the DPO at dpo@example.com.

Breach Notification:

In the event of a personal data breach that is likely to result in a risk to your rights and freedoms, we will notify affected users within 72 hours, in compliance with GDPR Article 33.

Legal Basis for Processing:

We process data on the following legal bases:
Performance of a contract (e.g. managing your account)
User consent (e.g. marketing communications, cookies)
Legitimate interest (e.g. ensuring security, preventing fraud)
Legal obligations (e.g. tax compliance)

Consent History Logging:

We maintain records of user consent decisions (including time, date, and version of our terms) to demonstrate compliance with GDPR.

Accessibility & Language Support:

Our consent interfaces are designed to be accessible (keyboard navigation, screen reader compatible). Information is provided in the primary languages of the regions we serve.

Contact Information

Data Protection Officer (DPO):

Email: dpo@example.com

Marketing Data Inquiries:

Email: marketing-privacy@example.com

Technical Data Requests (Analytics, Partners):

Email: tech-privacy@example.com

General Privacy Queries:

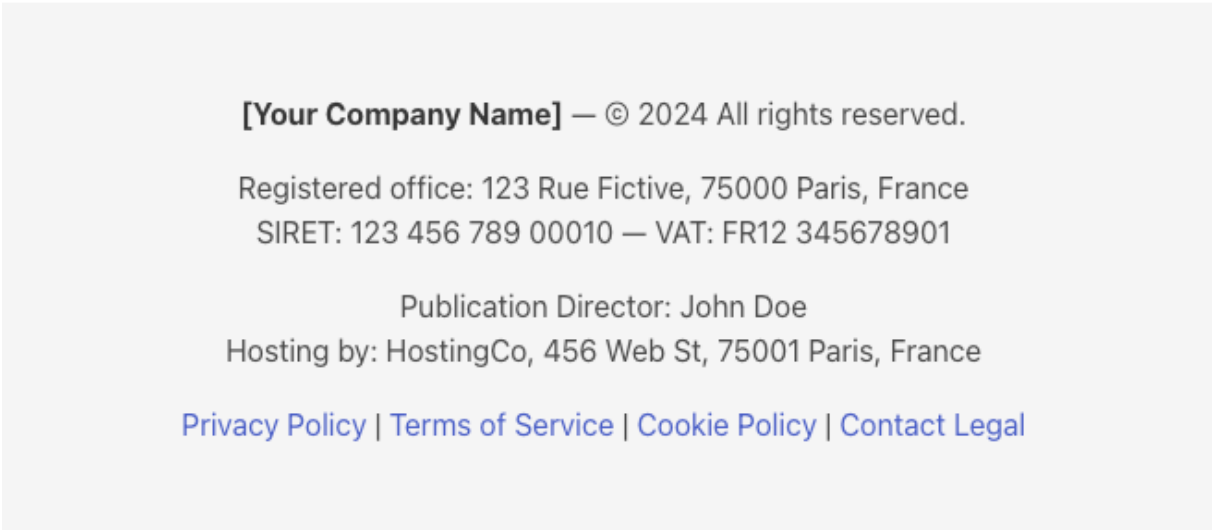
Email: privacy@example.com

We aim to respond to all requests within 30 days.

Figure 12: List of contacts and additional legal information

There are 3 pages : terms of service, cookies policy and privacy policy if the users want more details.

There are also the legal informations in the footer.



[Your Company Name] — © 2024 All rights reserved.

Registered office: 123 Rue Fictive, 75000 Paris, France
SIRET: 123 456 789 00010 — VAT: FR12 345678901

Publication Director: John Doe
Hosting by: HostingCo, 456 Web St, 75001 Paris, France

Privacy Policy | Terms of Service | Cookie Policy | Contact Legal

Figure 13: Footer with links to the different policies and contacts.

7. CI/CD

7.1. Code coverage

The usage of `pytest` allows to cover the code by creating tests for each method/function. In this project, because of the time limit, the code coverage is only about `52%`.

Coverage report: 52%

Files | Functions | Classes

coverage.py v7.8.0, created at 2025-05-16 03:58 +0200

File ▲	statements	missing	excluded	coverage
app/__init__.py	0	0	0	100%
app/database/db.py	9	0	0	100%
app/main.py	193	134	0	31%
app/middlewares/authenticationMiddleware.py	33	6	0	82%
app/middlewares/authorizationMiddleware.py	18	8	0	56%
app/middlewares/sanitizationMiddleware.py	12	4	0	67%
app/models/connectionDetails.py	8	0	0	100%
app/models/post.py	6	0	0	100%
app/utility/utility.py	36	0	0	100%
Total	**315**	**152**	**0**	**52%**

coverage.py v7.8.0, created at 2025-05-16 03:58 +0200

Figure 14: Code coverage with pytest

```
============================= test session starts ==============================
platform linux -- Python 3.13.3, pytest-8.3.5, pluggy-1.5.0
rootdir: /home/runner/work/fastapi-project/fastapi-project/backend
configfile: pytest.ini
testpaths: tests
plugins: asyncio-0.26.0, anyio-4.9.0, cov-6.1.1
asyncio: mode=Mode.STRICT, asyncio_default_fixture_loop_scope=None, asyncio_default_test_loop_scope=function
collected 13 items

tests/test_main.py .............                                         [100%]

- generated xml file: /home/runner/work/fastapi-project/fastapi-project/reports/junit.xml -
================================ tests coverage ================================
_____________ coverage: platform linux, python 3.13.3-final-0 _____________

Coverage XML written to file ../reports/coverage.xml
============================== 13 passed in 1.17s ==============================
```

Figure 15: Passed tests in github actions

7.2. Vulnerability scanner

The usage of `wapiti3` allows to scan the web application to check if vulnerabilities are present. Only minor vulnerabilites have been found. Patching them is not done right now, because of the limited time.

Summary

Category	Number of vulnerabilities found
Backup file	0
Weak credentials	0
CRLF Injection	0
Content Security Policy Configuration	1
Cross Site Request Forgery	0
Potentially dangerous file	0
Command execution	0
Path Traversal	0
Fingerprint web application framework	0
Fingerprint web server	0
Htaccess Bypass	0
HTML Injection	0
Clickjacking Protection	1
HTTP Strict Transport Security (HSTS)	0
MIME Type Confusion	1
HttpOnly Flag cookie	0
Unencrypted Channels	0
LDAP Injection	0
Log4Shell	0
Open Redirect	0
Reflected Cross Site Scripting	0
Secure Flag cookie	0
Spring4Shell	0
SQL Injection	0
TLS/SSL misconfigurations	0
Server Side Request Forgery	0
Stored HTML Injection	0
Stored Cross Site Scripting	0
Subdomain takeover	0
Blind SQL Injection	0
Unrestricted File Upload	0
Vulnerable software	0
Internal Server Error	0
Resource consumption	0
Review Webserver Metafiles for Information Leakage	0
Fingerprint web technology	0
HTTP Methods	0
TLS/SSL misconfigurations	0

Figure 16: Wapiti scan result

7.3. Github Workflow

Github actions have been set up. At each `pull request` and `push`, the Github Actions trigger `Snyk` to scan the code, and then `Sonarqube`.

7.3.1. Snyk

`Snyk` allows to check for security vulnerabilities in the dependencies.

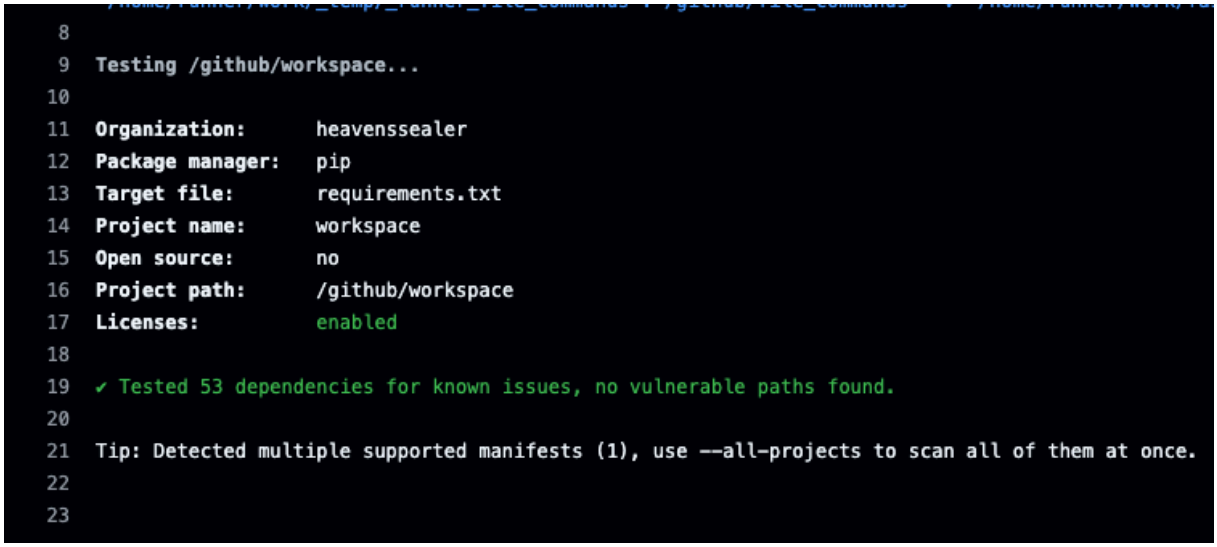


Figure 17: Snyk test

7.3.2. Sonarqube

`Sonarqube` checks for multiple indicators :

- Reliability flaws (bugs)
- Security vulnerabilities
- Maintainability issues (code smells, technical debt)
- Quality metrics (coverage, duplications)

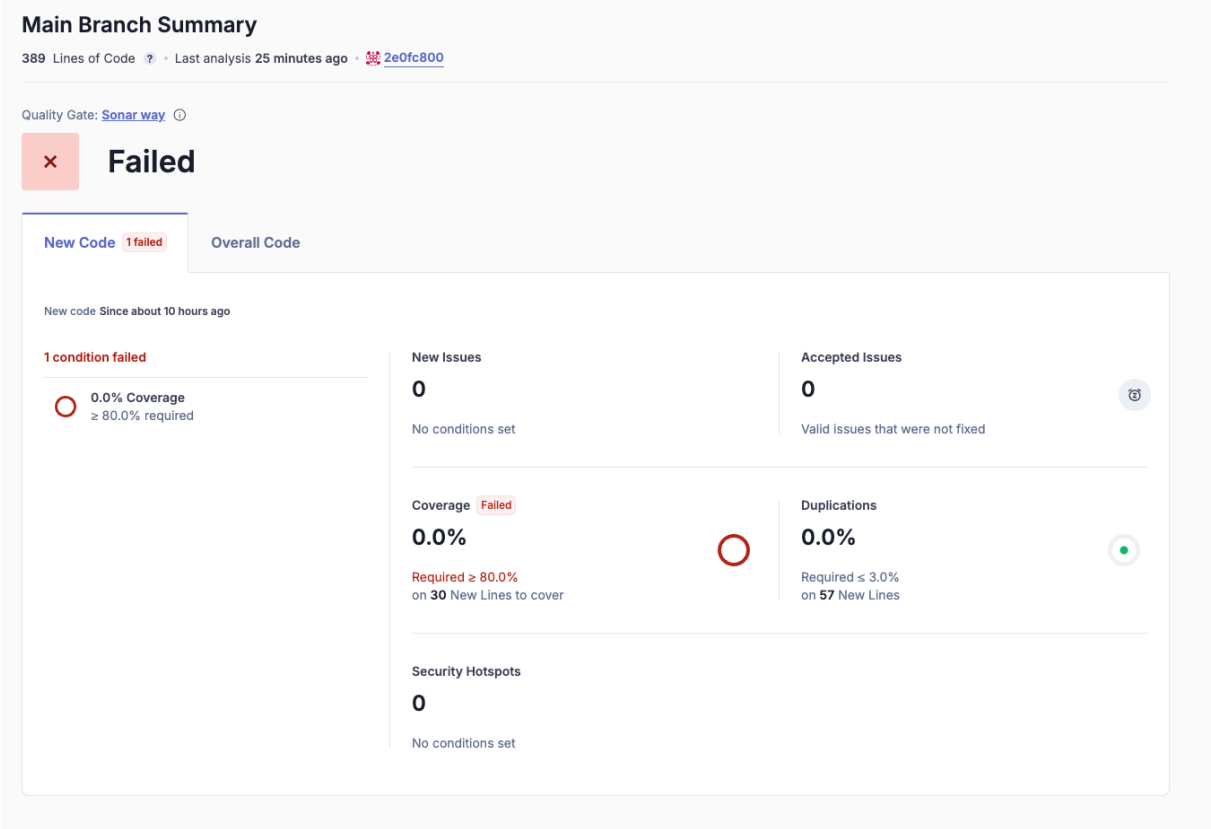


Figure 18: Sonarqube analysis

The coverage is done using pytests, so it is not accounted for here. Despite it showing Failed, the conditions that we check are Duplications and Security Hotspots.

7.4. Dockerization

The server is dockerized, allowing it to be executed from another computer easily. It can also help to host it on cloud services.

The command used to dockerize :

- to build :
  - `docker build -f backend/Dockerfile -t my-fastapi-app:latest backend/`
- to launch the container :
  - `docker run -d -p 8000:8000 --env-file backend/.env my-fastapi-app:latest`

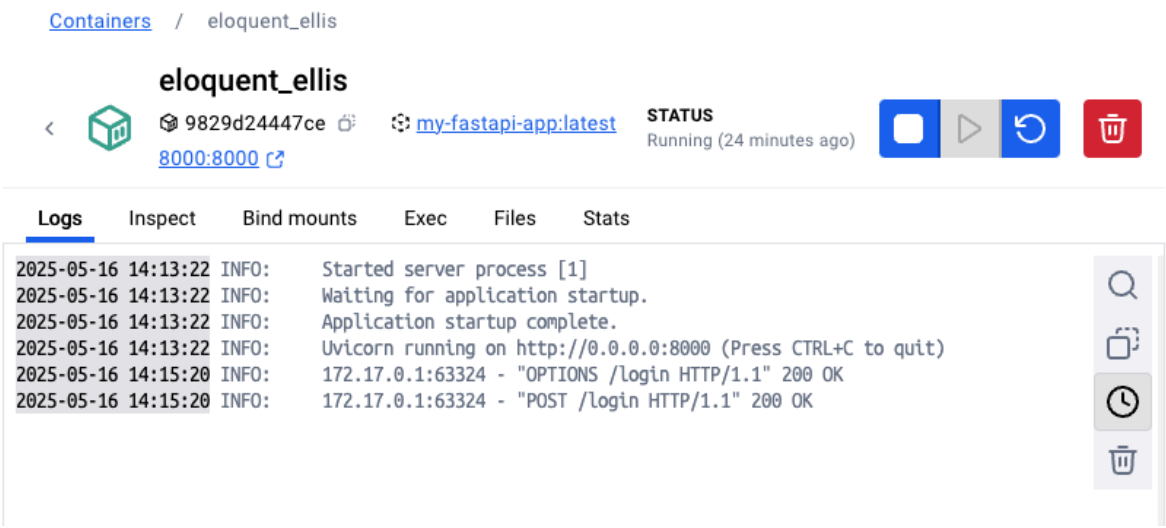


Figure 19: Docker running

8. Project management

`Notion` was used for `agility`, helping with organizing and managing tasks.

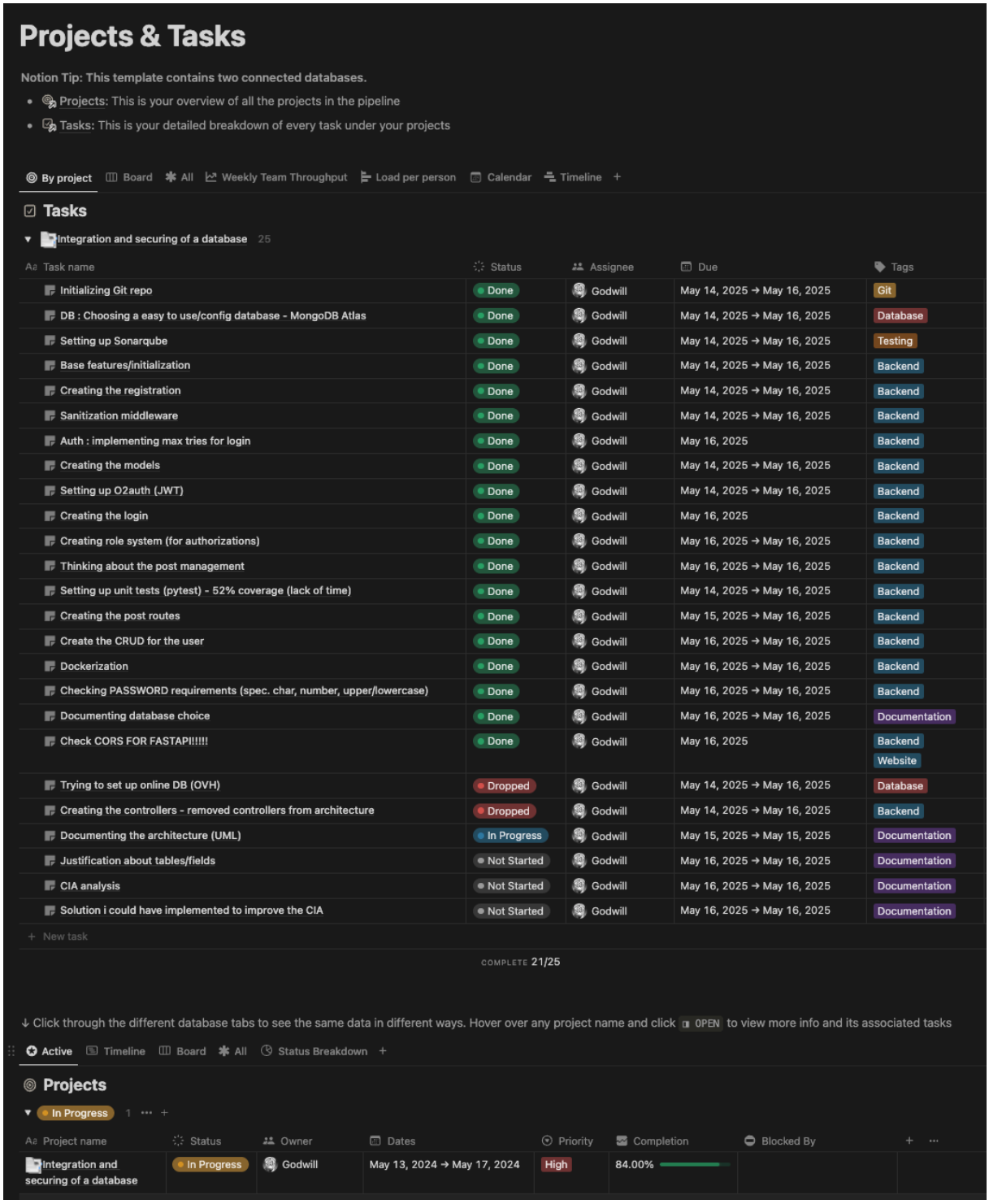


Figure 20: Project management with Notion

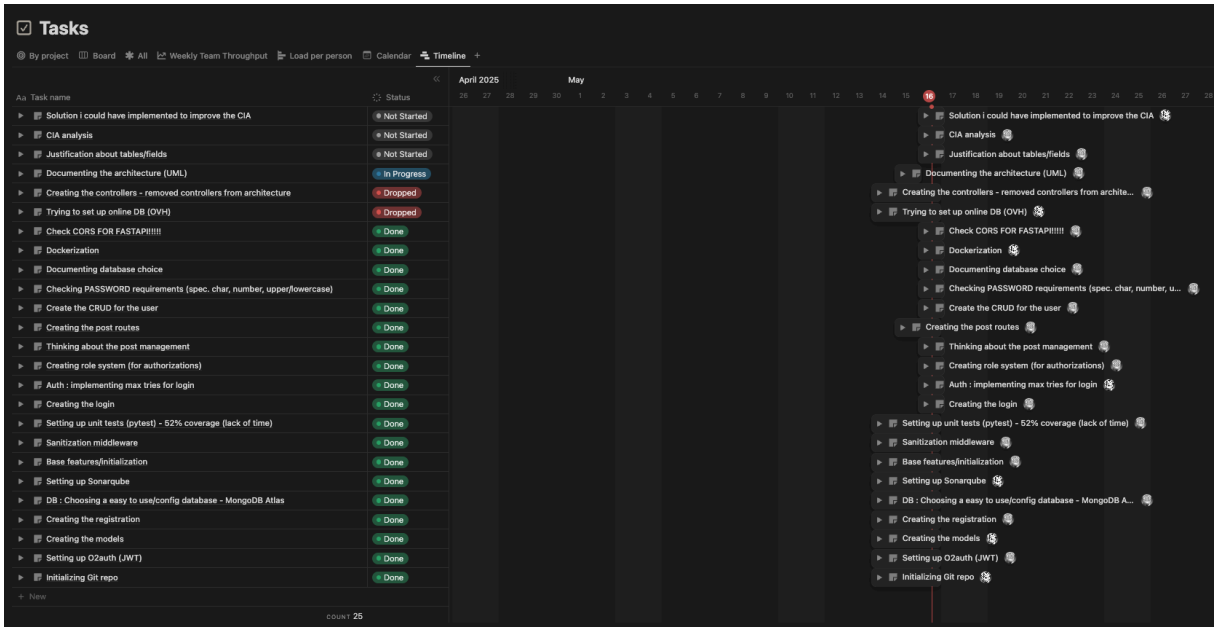


Figure 21: Tasks with Notion

9. Improvements

To improve the project, there is a certain number of elements that could have been improved/implemented:

- TLS/SSL : generating a real certificate instead of a self signed one.
- Improved code coverage : code coverage for more than 80% is the industry norm.
- Responsive/Reactive design : the frontend could be made using React or other libraries/frameworks, to enhance user experience.
- More robust input sanitization: although the inputs are sanitized in the backend, there might be better ways to improve the robustness of the server.
- Add real analytics.
- Add real/mock advertisement to simulate cookies choices.

10. Installation

10.1. Installation details

To make the project work on your machine, because of the certificate constraints, you will have to generate your own.

First, create a `san.cnf` file and add the following lines to it :

```
[req]
req_extensions = v3_req
distinguished_name = req_distinguished_name
prompt = no

[req_distinguished_name]
CN = localhost

[v3_req]
subjectAltName = @alt_names

[alt_names]
DNS.1 = localhost
IP.1  = 127.0.0.1
IP.2  = ::1
```

Then, execute this command to generate the certificate and the key :

```
openssl req -x509 -nodes -days 365 \
  -newkey rsa:2048 \
  -keyout local_key.pem \
  -out local_cert.pem \
  -config san.cnf \
  -extensions v3_req
```

Then, go to `backend/certs/` and add the generated certificates `local_key.pem` and `local_cert.pem`.

Then, you can just build the docker image and run it.

Finally, to run the frontend, go to the `frontend` directory and run :

`npm install -g http-server`

Install the `mkcert` utility, and generate a pair of certificates.

```
┌[godwill@macbookpro.home]-(~/Documents/ipssi/bdd/fastapi-project/frontend)-[git://backend x]
└> mkcert localhost
Created a new local CA 💥
Note: the local CA is not installed in the system trust store.
Note: the local CA is not installed in the Firefox trust store.
Run "mkcert -install" for certificates to be trusted automatically ⚠️

Created a new certificate valid for the following names 📜
 - "localhost"

The certificate is at "./localhost.pem" and the key at "./localhost-key.pem" ✅

It will expire on 22 September 2027 📅
```

Figure 22: Certs generation

Then, launch the server with the following command :

```
http-server -S -C localhost.pem -K localhost-key.pem -p 8080
```

This will allow you to access the web application from `https://127.0.0.1:8080/index.html`.

You should then be good to go and test.

10.2. Link to the project

Link to the project : https://github.com/heavenssealer/GDPR-compliant-project°

Index of Figures