

INTRODUCTION

The OpenData is a vast and destructured world which provides powerful data about a lot of different subjects in a large number of formats.

In this exercise, we will use the platform data.gouv.fr to build a prototype of web application that can retrieve specific datasets from this platform, and show their meta information.

You can use your favorite technology, there is no restriction.

The project should be versioned with Git and stored in a public area (gitlab, github, bitbucket etc...)

SCRAPING

Write a script that :

- Scrap the five next pages from https://www.data.gouv.fr/api/1/datasets/?page=3000&page_size=1 using the `page` query parameter.
- Parse the `title` and the `latest` keys, located in the `resources` arrays and then store them in a database.
- Download the datasets using the `latest` URL and store them in a local storage folder.

API

Write an API with 2 endpoints:

- List files meta data (`title` & `latest`)
- Download a specific file dataset

CLIENT INTERFACE

Using your API, build a web app which displays the list of the downloaded datasets. Each row must show the dataset's title and a button that allow the user to download the dataset.

TESTS & CONTAINER

Write the unit tests covering the code.

Dockerise the project.