BPM-Conseil

Running AI-LLM Application on Kata Containers

Patrick Beaucamp

Mail : patrick.beaucamp@bpm-conseil.com

kata containers

Data4Citizen

Presentation Agenda

- **Legal Context in Europe (and France)**
- **Migration to Kata Container**
- **Running AI Platform powered by LLM in a Kata cluster**
  - **Data4Citizen Platform**
  - **Data4Citizen Architecture**
  - **Data4Citizen & AI-LLM**
  - **Data4Citizen & DM-LLM**

Part 1 : Legal Context in Europe

When too many rules and security kills the security !!!

Legal Context – NIS2

- As the cyber threat increases and information systems remain partly vulnerable, the NIS 2 (Network and Information Security) directive, published in the Official Journal of the European Union in December 2022, represents a unique opportunity.

- NIS2 implementation will enable thousands of entities that affect the daily lives of citizens to better protect themselves.

Legal Context – GDRP

- The EU general data protection regulation (GDPR) is the **strongest privacy and security law in the world**.

- This regulation updated and modernised the principles of the 1995 data protection directive. It was adopted in 2016 and entered into application on 25 May 2018.

- The GDPR defines:
  - individuals' fundamental rights in the digital age
  - the obligations of those processing data
  - methods for ensuring compliance
  - sanctions for those in breach of the rules

Some immediate Consequences

When recommandations turn as requests :

- Recommandation/Need to encrypt Virtual Machine
  - Impact on performance for standard déployment
- Recommandation/Need to encrypt Database storage
  - Compliance : PostGreSql encryption with pgcrypto using AES256
- Recommandation/Need to Secure Docker Image
  - Compliance : running images in Kata Container

Part 2 : Migration and Running Kata

A short procedure to run your image on Kata

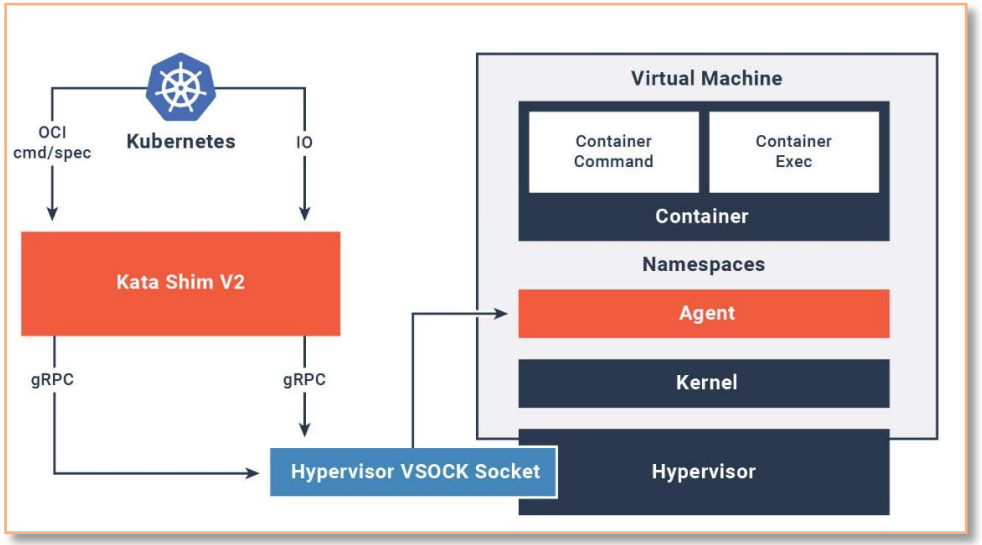
Install Kata Containers (kata-runtime, kata-proxy, kata-shim).
Configure Docker (daemon.json) or containerd.
Run a Kata container (docker run --runtime=kata-runtime).

Kata Containers - Overview

- **Kata Containers** combine the benefits of containers and virtual machines (VMs) to provide enhanced isolation. Here's how to build and run a Kata container on a Linux system.

Why Use Kata Containers?

✓ **Enhanced security** via VM isolation
✓ **Compatible** with Docker and containerd
✓ **Ideal for sensitive workloads** (cloud, edge computing, etc.)



kata
containers

Install Kata Containers

Prerequisites

- A Linux host (Ubuntu, Debian, CentOS, etc.)
- A compatible hypervisor (QEMU/KVM)
- A container runtime (Docker or containerd)

Exemple of Installation on Ubuntu/Debian

```
sudo apt-get update
sudo apt-get install -y software-properties-common
sudo add-apt-repository -y ppa:kata-containers/ppa
sudo apt-get update
sudo apt-get install -y kata-runtime kata-proxy kata-shim
```

```
kata-runtime --version
```

Configure Docker to Use Kata Containers

If you want to run Kata Containers with Docker, you need to add Kata as an **alternative runtime**.

Modify the Docker configuration file:

```
sudo mkdir -p /etc/docker
cat <<EOF | sudo tee /etc/docker/daemon.json
{
  "runtimes": {
    "kata-runtime": {
      "path": "/usr/bin/kata-runtime"
    }
  }
}
EOF
```

Restart Docker:

```
sudo systemctl restart docker
```

Run a Kata Container with Docker

Once everything is set up, run a Kata container with:

```
docker run --rm -it --runtime=kata-runtime ubuntu bash
```

This will start an **Ubuntu container** but with **enhanced isolation** via a **Kata micro-VM**.

Use Kata with containerd (Optional)

If you are using containerd instead of Docker, edit its configuration file:

```
sudo nano /etc/containerd/config.toml
```

Add Kata as a runtime:

```
[plugins."io.containerd.grpc.v1.cri".containerd.runtimes.kata]
  runtime_type = "io.containerd.kata.v2"
```

Restart containerd

```
sudo systemctl restart containerd
```

Then run a container:

```
ctr run --runtime io.containerd.kata.v2 -t
docker.io/library/alpine:latest kata-test sh
```

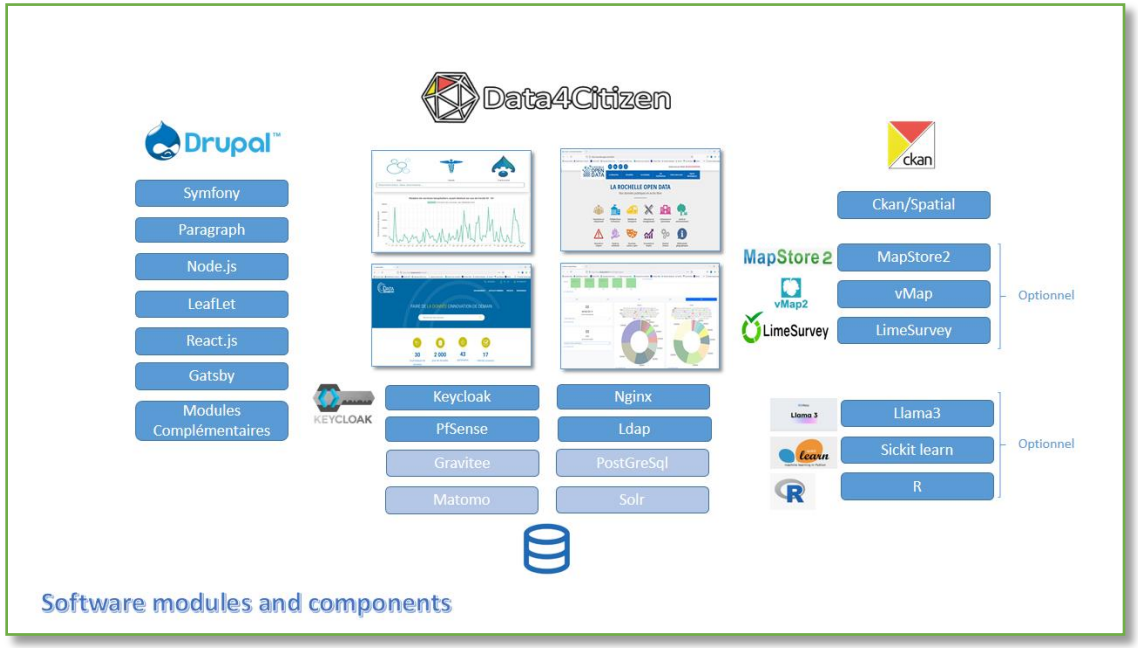
Part 3 : Running AI Platform Powered by LLM

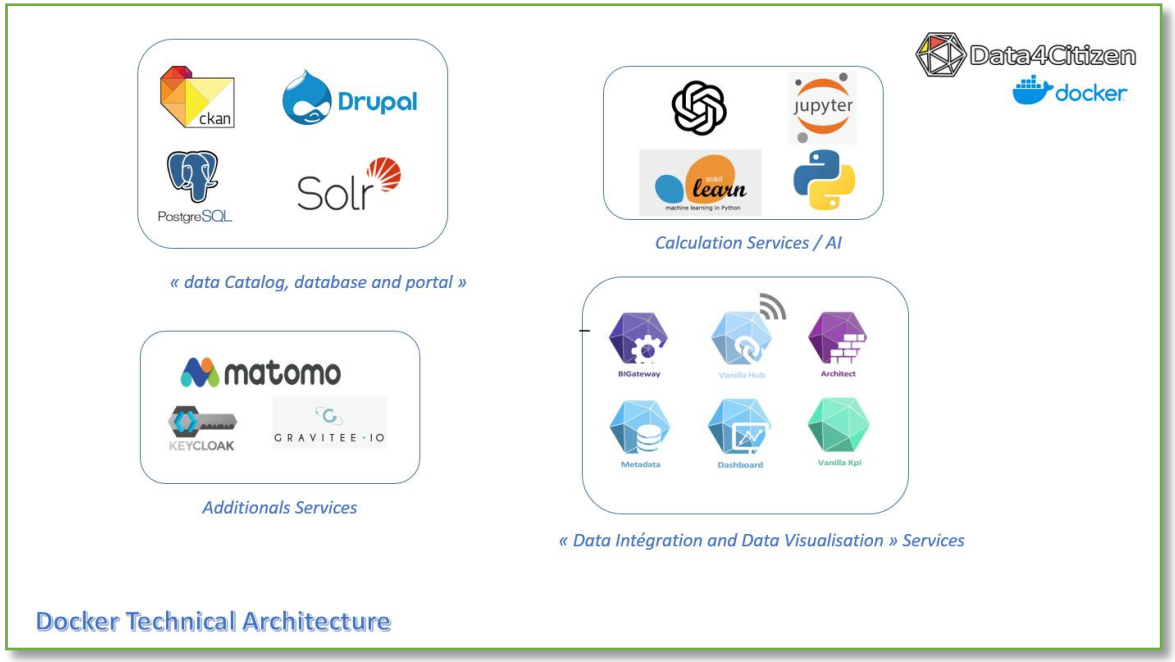
Data4Citizen Platform
Data4Citizen Architecture

Demo 1 : Data4Citizen & AI-LLM - Dashboard
Demo 2 : Data4Citizen & DataMining
Demo 3 : Data4Citizen & DM-LLM
Demo 4 : Data4Citizen & Jupyter

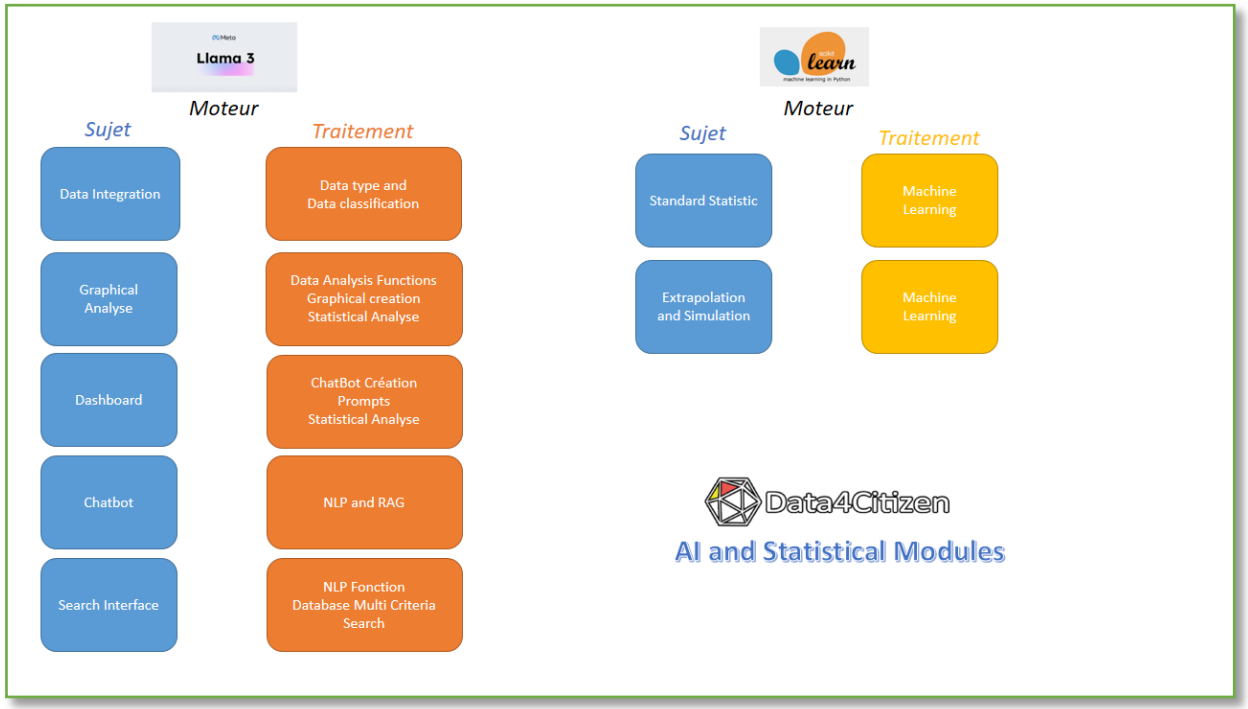
Data4Citizen Platform

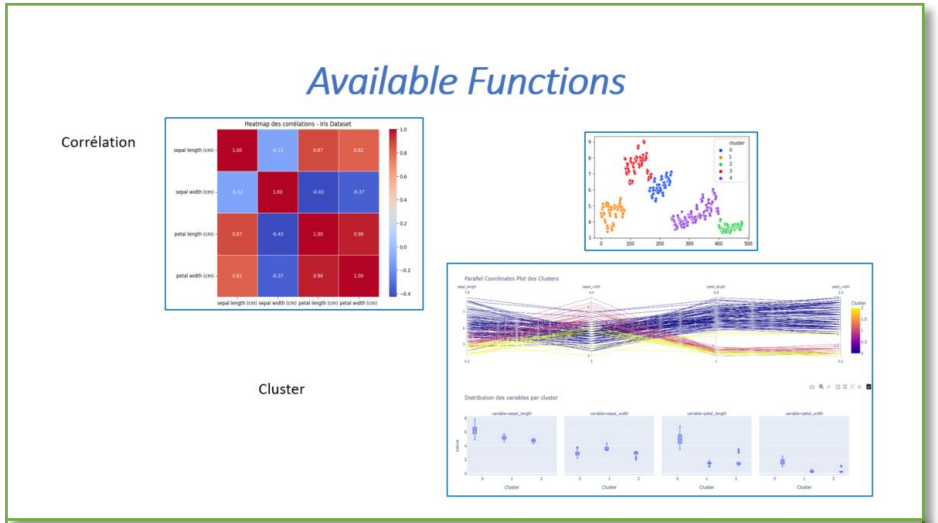
Majority of projects are governmental projects : high recommendation to be GPDR and NIS2 compliant

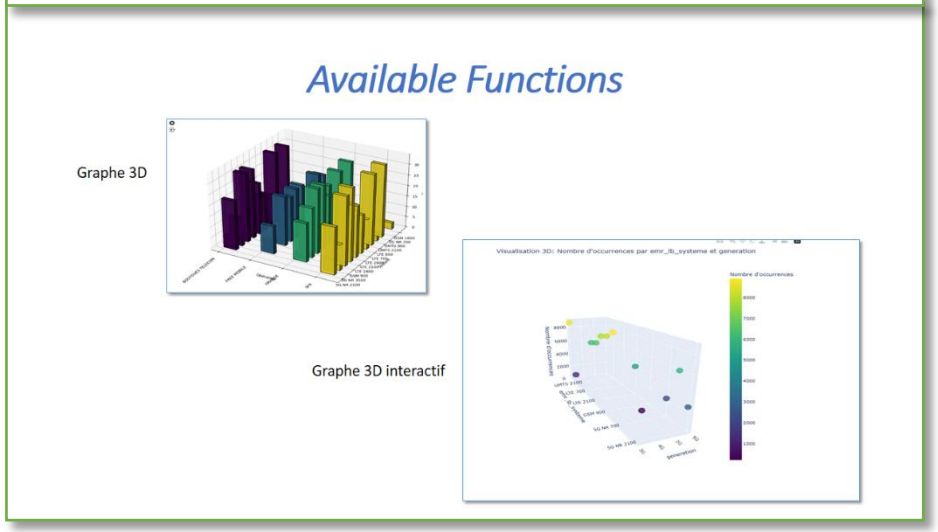




Data4Citizen IA-LLM Architecture

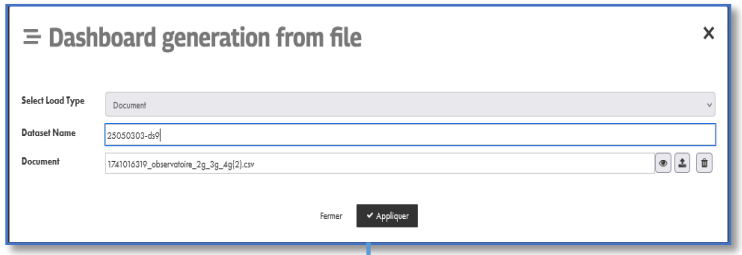


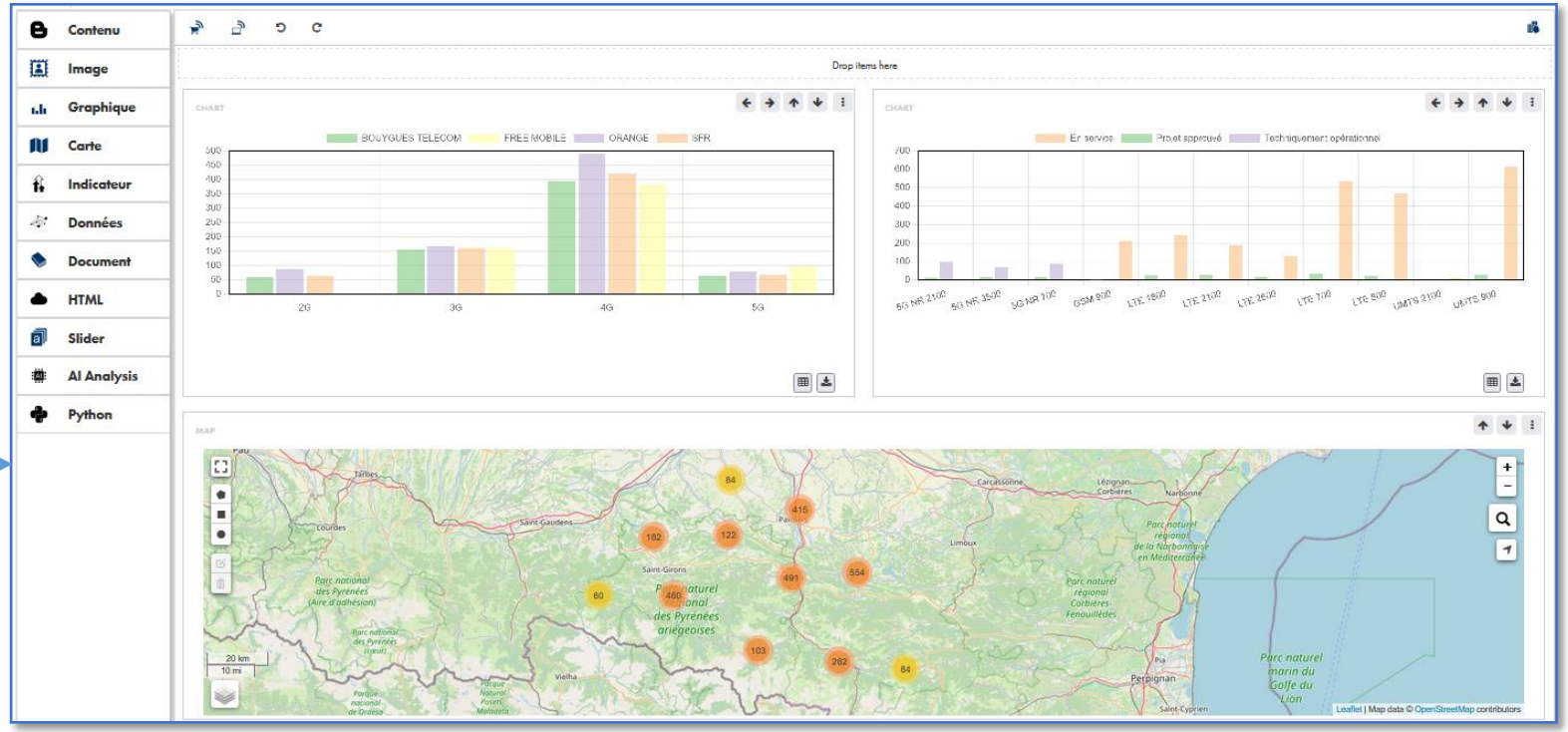




Data4Citizen IA-LLM in Action – Dashboard - 1

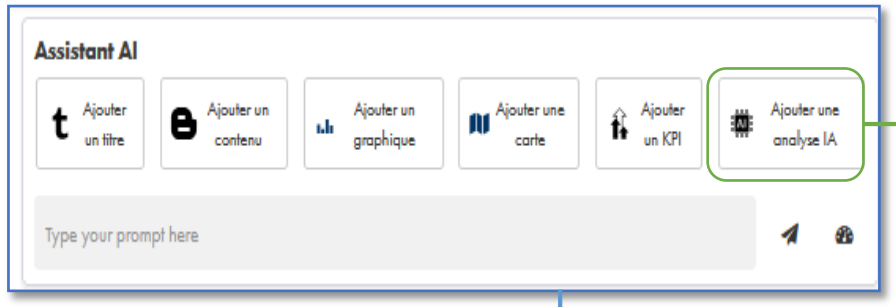
Data Integration – Data Analysis and Dashboard Creation

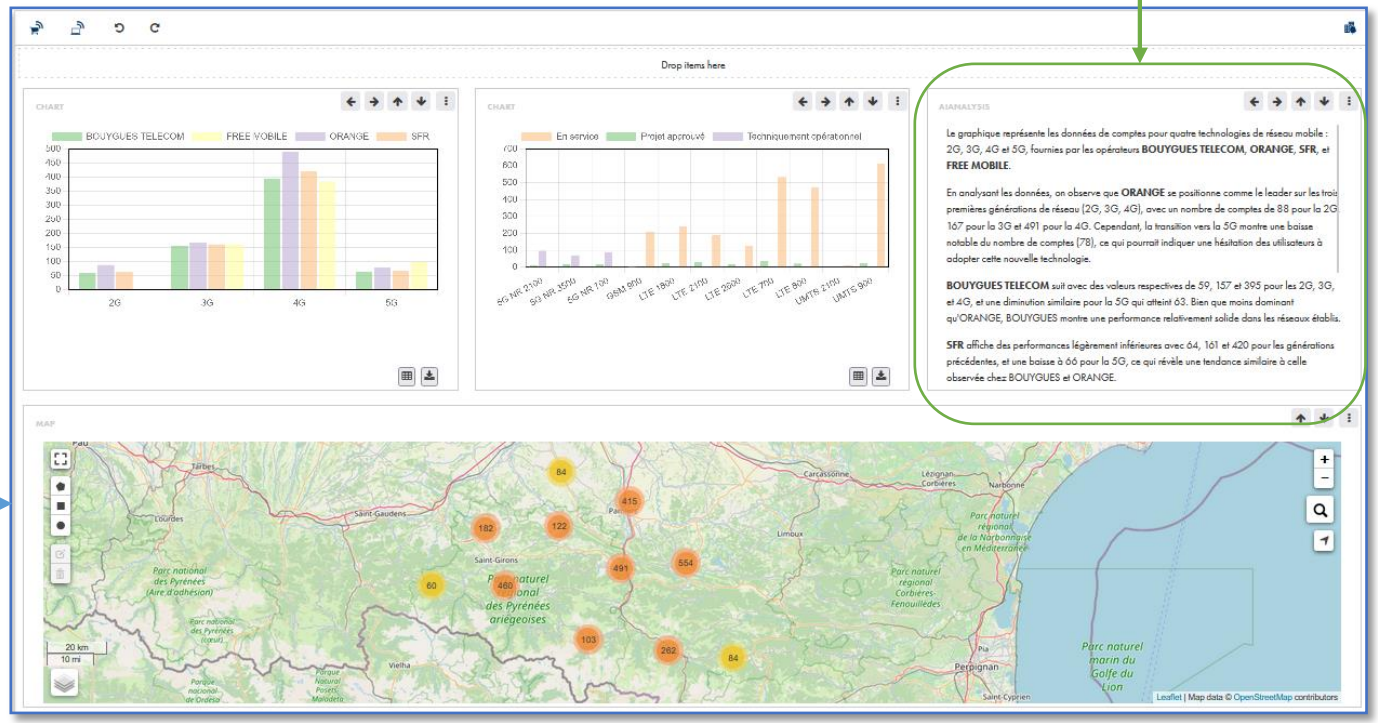




Data4Citizen IA-LLM in Action – Dashboard - 2

Data Analysis inside Dashboard using LLM

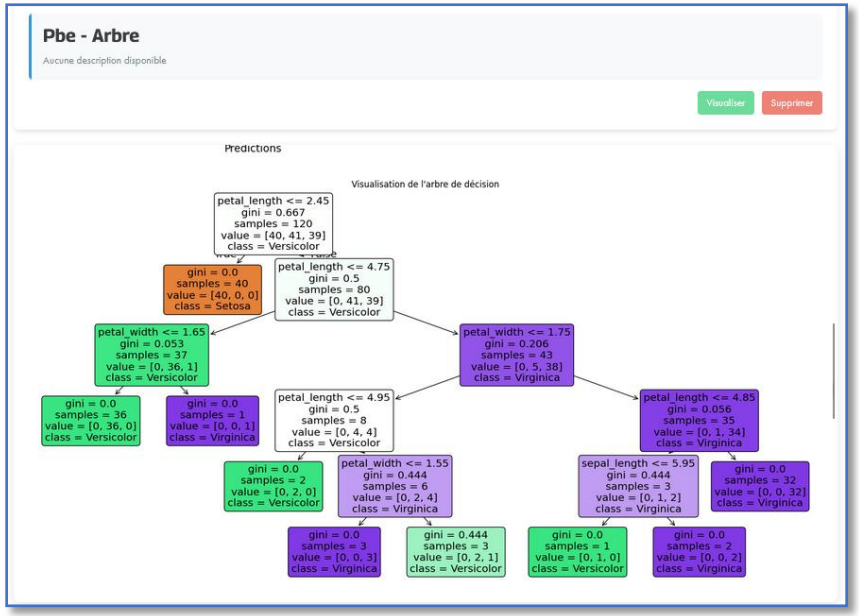


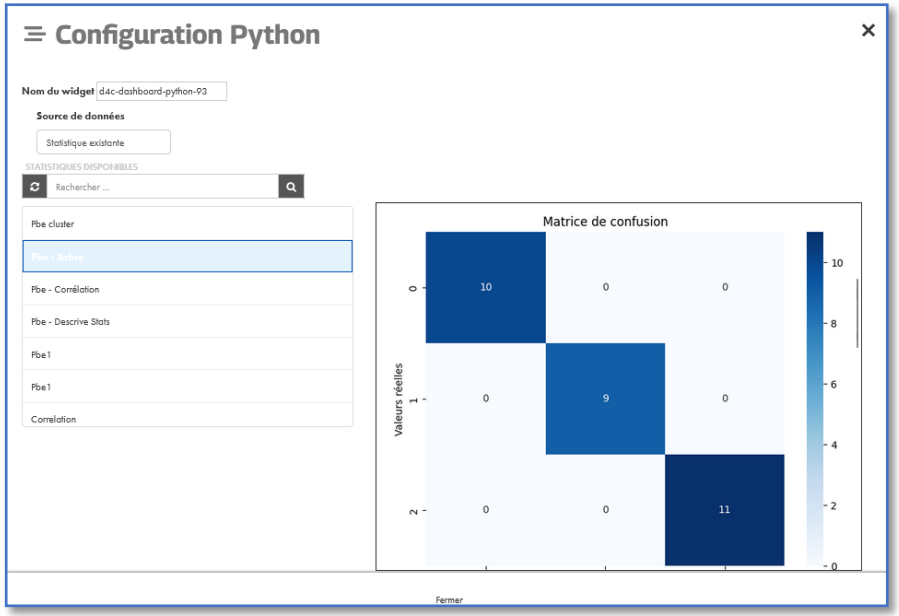


Live Demo

Data4Citizen Just DM in Action – Iris Dataset

DataSet – Sickit-learn Analysis

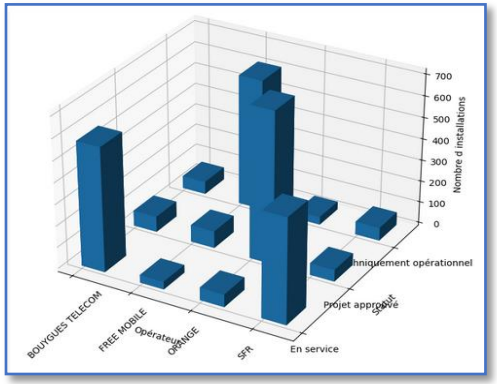


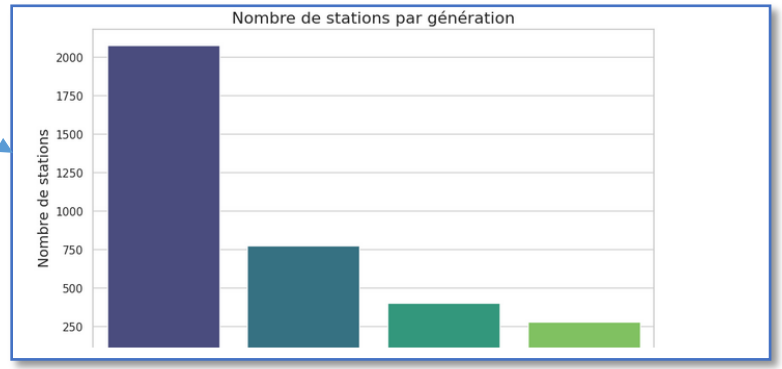


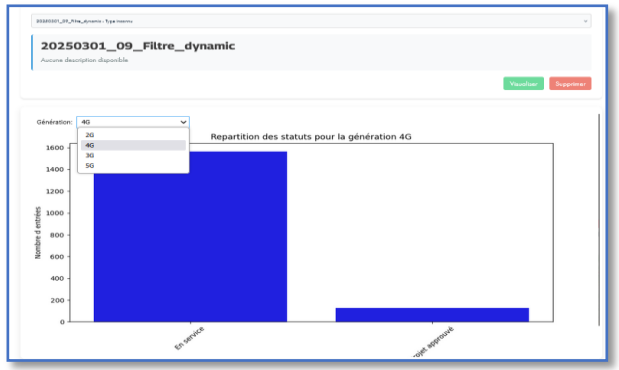
Live Demo

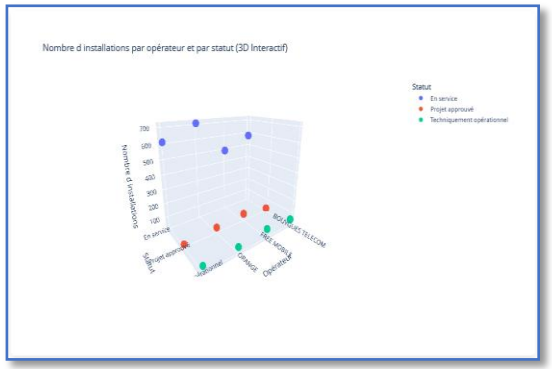
Data4Citizen DM-LLM in Action – Chat AI - 1

DataSet – Chat AI using Python, Sickit-learn and LLM



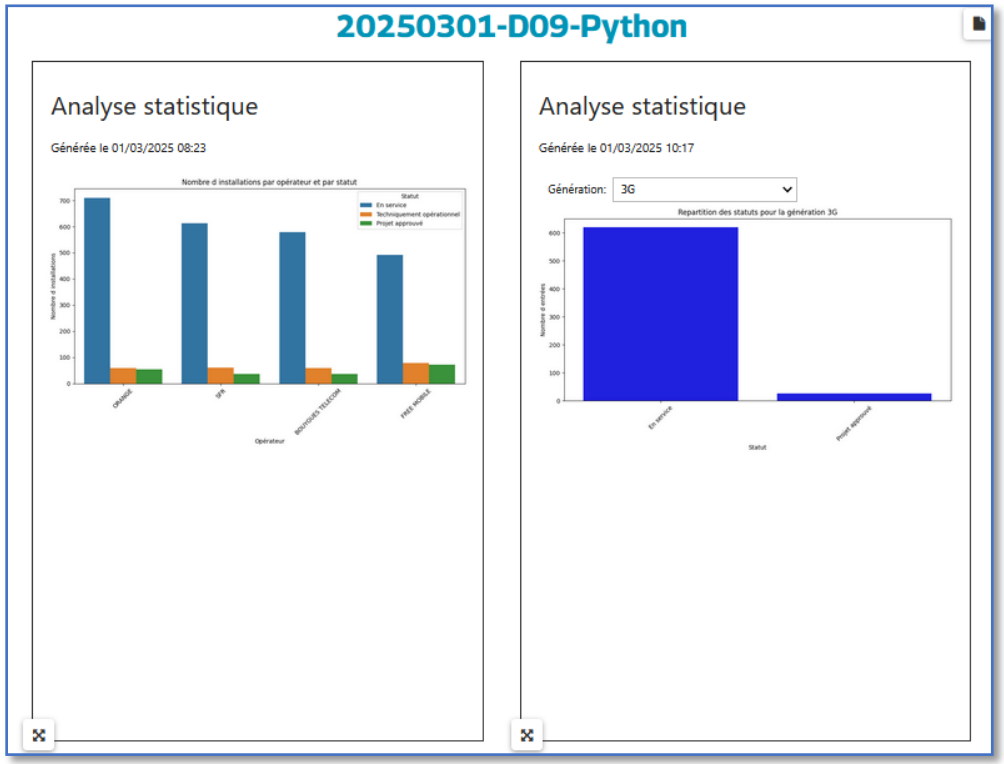


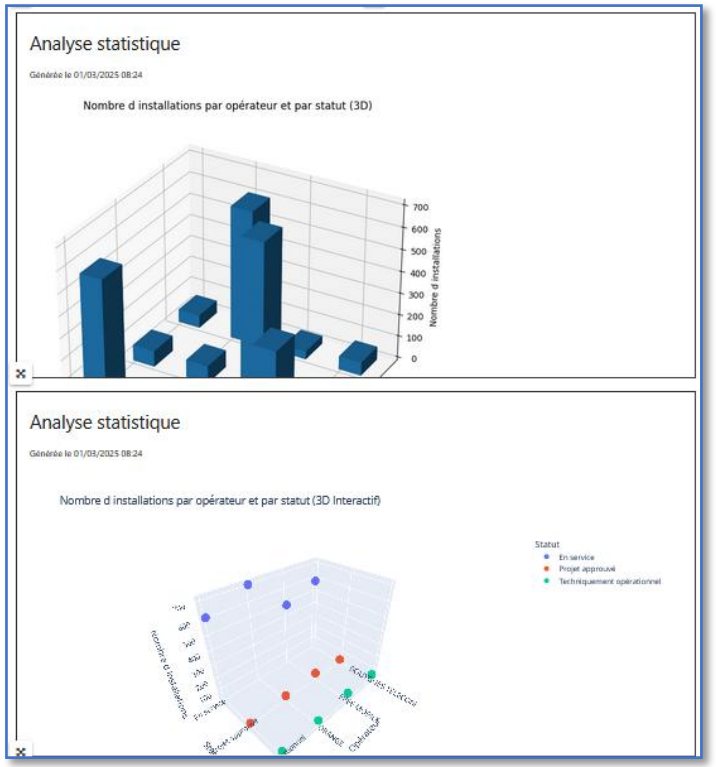




Data4Citizen DM-LLM in Action – Chat AI - 1

When AI-Analysis turns back inside a Dashboard

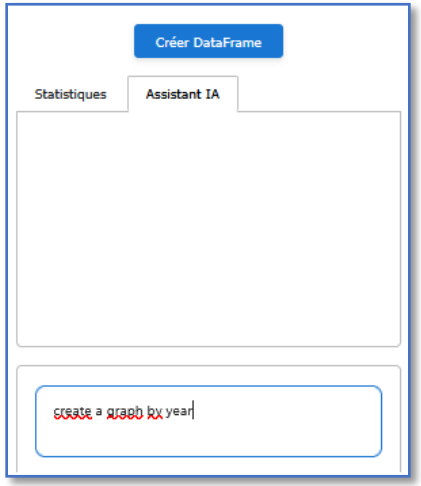


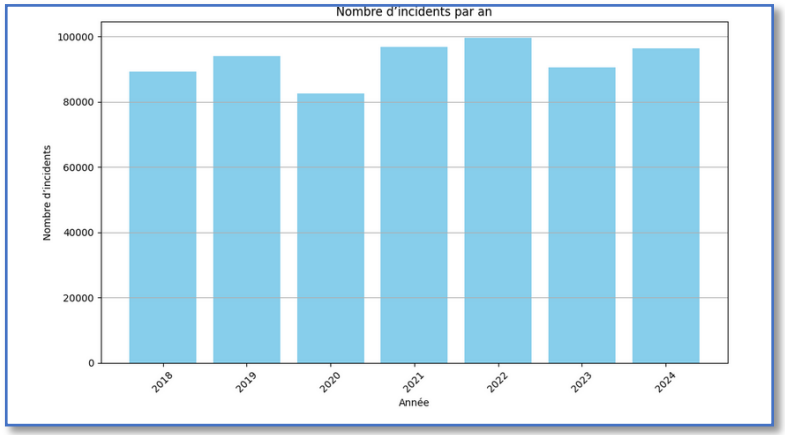


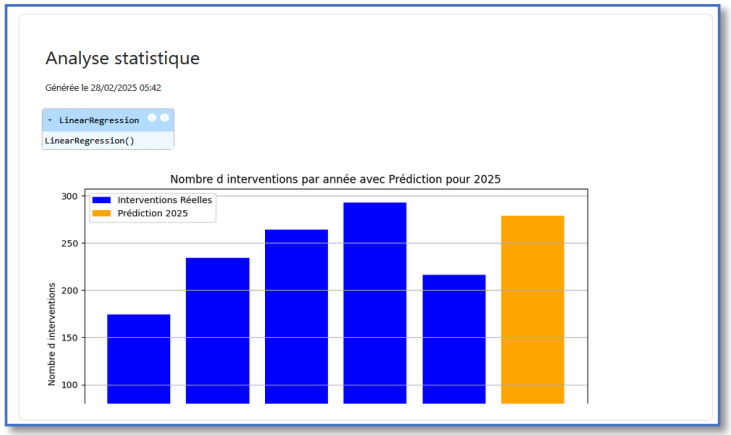
Live Demo

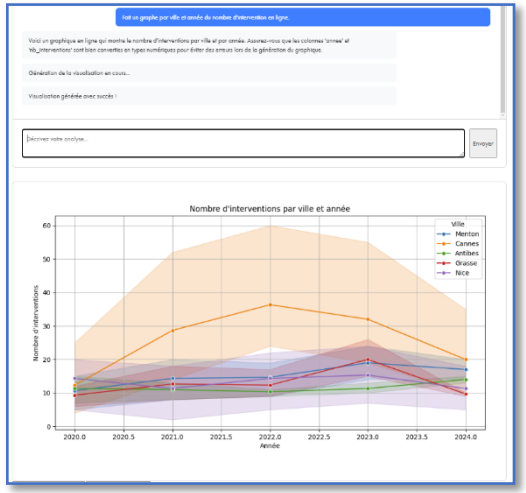
Data4Citizen DM-LLM in Action – Jupyter

Jupyter – Chat AI using Python, Sickit-learn and LLM



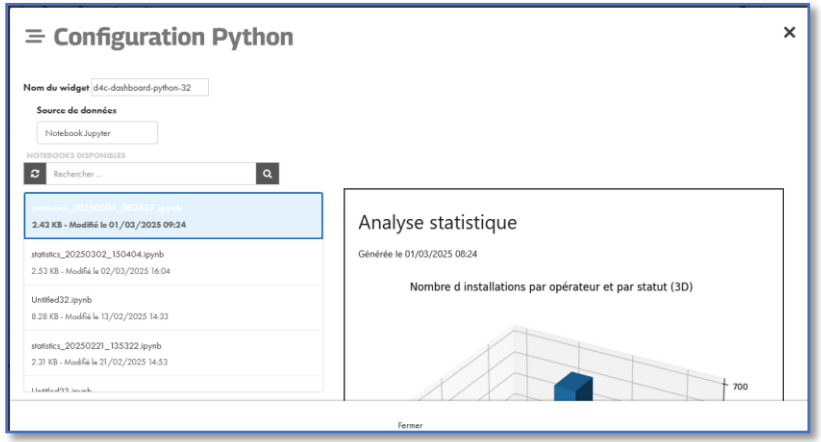


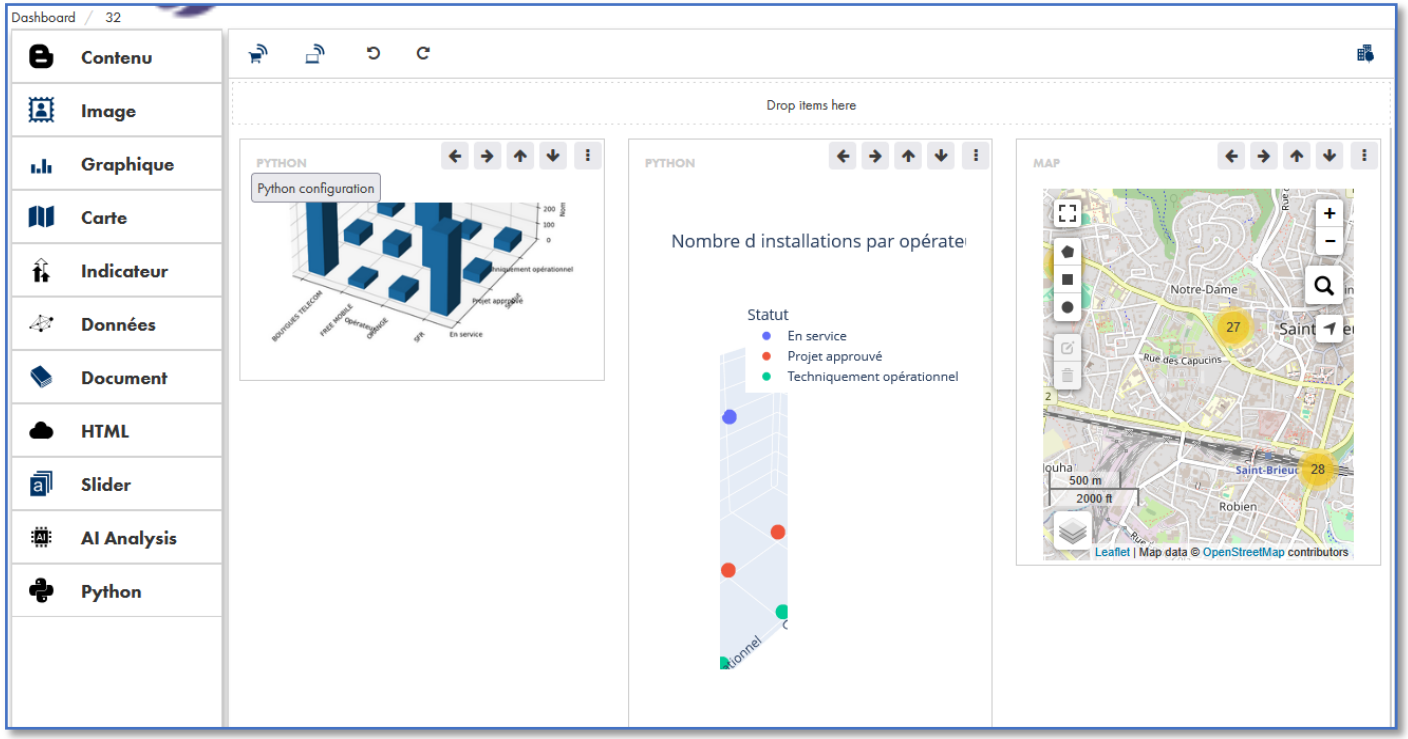




Data4Citizen DM-LLM in Action – Jupyter

When Notebook turns back inside a Dashboard





Live Demo

BPM-Conseil

kata containers

Data4Citizen

BPM-Conseil

kata
containers

Data4Citizen