



# Big Data Test Infrastructure (BDTI)

*From Hype to Action*

*May 10th, starts at 3:00 PM CET*

## **DG CNECT**

Directorate-General for Communications  
Networks, Content and Technology

## **DG DIGIT**

Directorate-General  
for Informatics



**#DigitalEUProgramme**

# DIGITAL EUROPE PROGRAMME

## Instructions for the live webinar



This is an interactive session. There will be time for questions and answer throughout the presentation. We hope you will share your views.



Click on **\*Connect audio\*** to hear the presenters but please note that your microphones will be muted during the entire session.



Please note that this webinar will be recorded.

An illustration in the top left corner shows four stylized figures in a purple and white color scheme. One figure stands holding a tablet, another sits at a laptop, and two others are positioned around various digital icons including a server rack, a heart rate monitor, a bar chart, a shield, and a building. The background features abstract purple and white lines.

# DIGITAL EUROPE PROGRAMME

## Welcome to the live Webinar

### Agenda for today

- 01** Introduction to the Digital Europe Programme *(5 min)*
- 02** BDTI in a nutshell *(10 min)*
- 03** BDTI Service Architecture *(20 min)*
- 04** BDTI Hands-on use case: **Supporting the Green Deal data space: Creating insights from EU electric transport data** *(30 min)*
- 05** Support & how to apply *(10 min)*
- 06** Q&A session *(15 min)*

# 01

## Introduction to the Digital Europe Programme



**MARIA CLAUDIA BODINO**

BDTI Project Officer  
DG DIGIT – EUROPEAN COMMISSION

#DigitalEUProgramme

# DIGITAL EUROPE PROGRAMME

## Intro to the Digital Europe Programme (DEP)

The **Digital Europe Programme** is a new EU funding programme (7.5 billion) focused on **bringing digital technology** to businesses, citizens and public administrations.

DEP will provide strategic funding in **five crucial areas**:

High performance computing

Cybersecurity

**Artificial intelligence**

*(Cloud, data and AI)*

Advanced digital skills

Deployment and wide use of digital technologies

With a planned overall budget of €7.5 billion, it aims to **accelerate the economic recovery and shape the digital transformation of Europe's society and economy**, bringing benefits to everyone, but in particular to small and medium-sized enterprises.

## DATA for AI - Public Sector Open Data for AI and Open Data Platform

### OBJECTIVES

- Increase the easy availability, quality and usability of public sector information in compliance with the requirement of the **Open Data Directive**
- Boost the **re-use and combination of open public data** across the EU for the development of information products and services, including AI applications.
- contribute to **Green deal applications and other common data spaces** where relevant

### WORK STRAND

- Maintenance and further expansion of the **Big Data Test Infrastructure (BDTI)** Building Block, including the possibility for the public sector to use it for testing **Business-to-Government (B2G)** data sharing collaborations for the **public good**

# DIGITAL EUROPE PROGRAMME

## Context and background

### OPEN DATA DIRECTIVE

The **Open Data Directive** provides common rules for a European market for government-held data and identifies a list of **High Value Dataset** to be published for **free** and available via **API**.

July 2019

### EUROPEAN DATA STRATEGY

The **European strategy for data** aims at creating a **single market for data** that will ensure **Europe's global competitiveness** and **data sovereignty**.

February 2020

### DATA GOVERNANCE ACT

The **EU Data Governance Act** is intended to **foster the availability of data** by increasing trust in data intermediaries and strengthening data sharing across the EU and between sectors.

November 2020

### DATA ACT

The **Data Act** clarifies **who can create value from data** and **under which conditions**.

February 2022

# 02

## BDTI in a nutshell



**MARIA CLAUDIA BODINO**

BDTI Project Officer  
DG DIGIT – EUROPEAN COMMISSION

#DigitalEUProgramme



# DIGITAL EUROPE PROGRAMME

## What is the Big Data Test Infrastructure (BDTI)?



**Six months free of charge** service provided by the EC as part of the Digital Europe Programme to EU agencies and EU local, regional and national public administrations



Data analytics stack and support, based on **open-source tools and cloud infrastructure** to boost the re-use and combination of public sector information



Stakeholder onboarding services allowing EU to **test Data analytics and Big Data technologies** and move towards **data-driven decision making for the public good.**

## Is BDTI for me?

If you are a public administration and want to

- opt for **data-driven decision making**
- move from **hype to action**
- test your ideas for **the public good**
- experiment with **data analytics open-source tools**
- in a **safe** and **ready to use cloud environment**.

### Public administrations

- Local
- Regional
- National
- Agencies

### Academia

- Spin-off
- PhD students
- Master students

### Private sector

- Start ups
- GovTech

# DIGITAL EUROPE PROGRAMME

## BDTI pilots

### CONSELLERIA DE SANITAT (CS) - Text Mining

Conselleria de Sanitat, the Health Public Administration of the Comunidad Valenciana Regional Government, needed a tool capable of analysing and extract knowledge from the huge quantity of scientific clinical articles coming from different sources (i.e. PubMed.gov, Covid-19 related clinical articles).



### EU CONVALESCENT PLASMA DATABASE – Data sharing

The European Blood Alliance is working together with the European Commission (DG SANTE) to create and manage an **EU-wide open-access platform** that collects data to support a study on **Covid-19 convalescent plasma therapy**. The aim of the study is to assess in which conditions the convalescent plasma treatment is most effective, in order to take data driven decisions on the therapy and focus the efforts of the research in the most promising directions.



### CITY OF FLORENCE – Mobility data

The main goal of the Municipality is to perform a **cross correlation between the multiple datasets** available within the city to understand how people were and are moving between the different districts, to then derive precious insights about mobility the most and about **how services can be redesigned to foster cultural activities and events**.



## BDTI CONTRIBUTIONS

Advanced **data visualization** and **text mining** tools to help **extracting knowledge contained in the documents**, supporting clinicians and managers in their clinical practices and day-to-day work in fighting the virus.

A ready-to-use, virtual environment in which **data collected through a custom-built website** are ingested and anonymized, to be then analysed with advanced data visualization and analytical tools. Initially, only donation data were processed, then the scope was increased to capture the **end-to-end of blood plasma, from donation to patient/clinical trial**.

Predictive, descriptive and time-series analysis on multiple datasets collected **before, during and after the Covid-19 pandemic** such as: public Wi-Fi sensors, parking and geo-referenced data of people movements (i.e. tourists).

# DIGITAL EUROPE PROGRAMME

## Main benefits



Implement a **free of charge** pilot project for **6 months**



Access to **advisory** and **technical support** during the entire duration of the pilot



Access to a platform with **ready-to-use open-source tools for data analytics**, tailored to your use case



Use high **performance infrastructure** that leverages the power of the elastic **cloud**



**Share and re-use cross border data** across policy domains and organisations, access to **insights on best practices** with big data projects and other pilots



Receive guidance to move from a pilot to a **production-ready** process – **EXIT package ready to use**

03

# BDTI Service Architecture



**WOUTER MAROY**

Technical Lead  
BDTI team

## BDTI improves decision-making

Connected devices

Business reports

Open data

Create insights by analyzing

Web content

Social media

Business processes

in any format

Unstructured

Structured

Semi-structured

or volume

1.000 docs

1.000.000 docs

10 docs

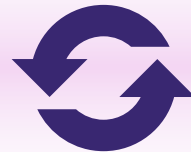
or domain



# DIGITAL EUROPE PROGRAMME

## BDTI helps you doing data analytics

COLLECT  
AND EXPLORE

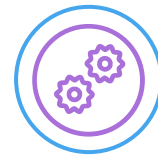


ANALYZE

LEVERAGE AI



VISUALIZE

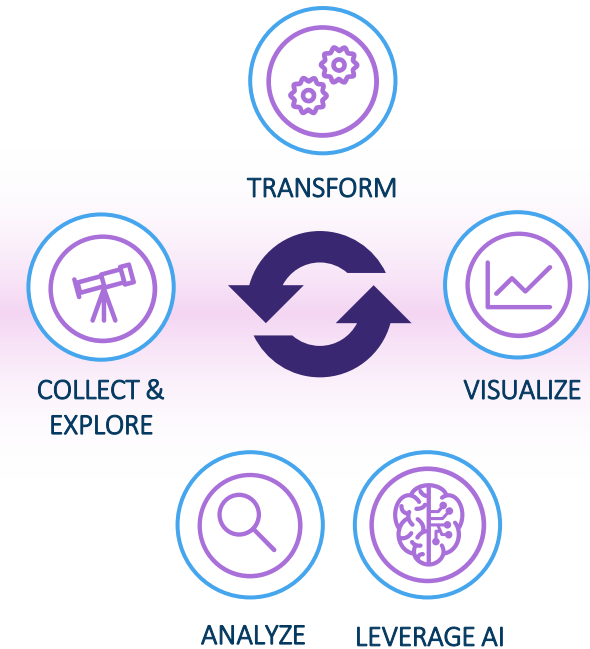


TRANSFORM

# DIGITAL EUROPE PROGRAMME

## BDTI helps you doing data analytics

- Self-service platform with preconfigured tooling
- Computational resources & storage
- Managed service
- Exit package for industrialization

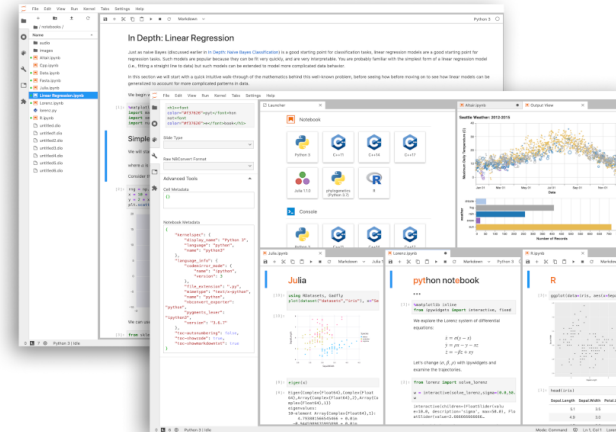
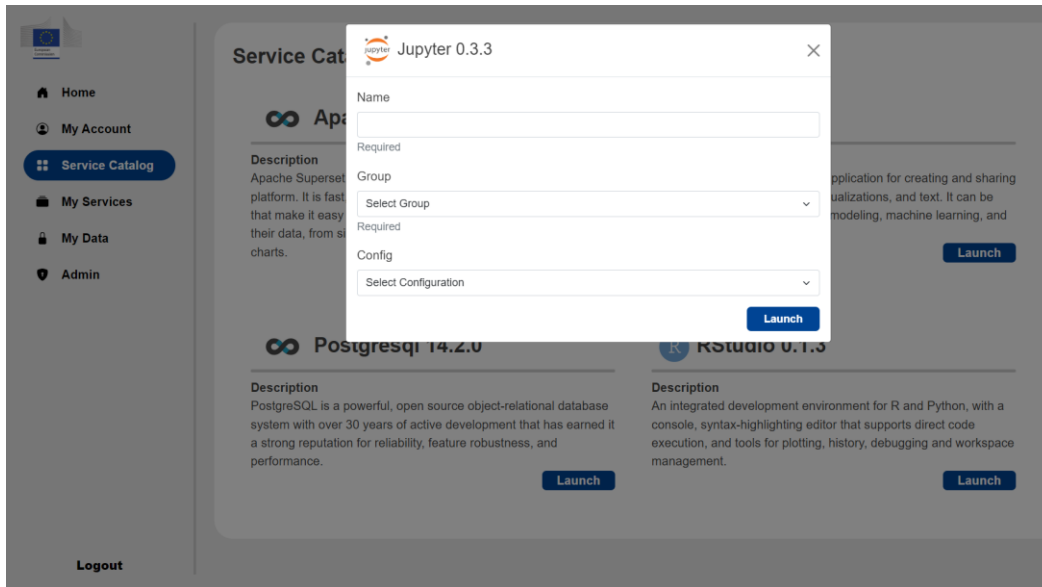




# DIGITAL EUROPE PROGRAMME

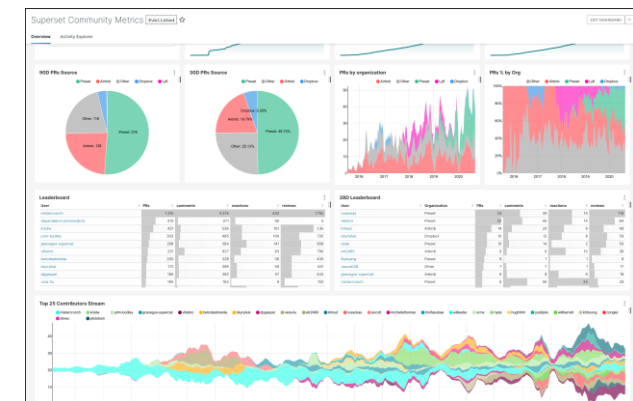
## Supported by tooling and technologies

Easily deploy, manage and use data science tools with a self-service portal.



<https://jupyter.org>

Leverage a variety of useful and popular tools and technologies.



<https://superset.apache.org>

# DIGITAL EUROPE PROGRAMME

## Available tooling

### Data Lake

Object storage

 **MINIO**

### Development Environments

Notebook

 jupyter

Statistical workbench

 R Studio

Analytic workflow engine

 KNIME

### Visualization

BI platform

 Superset

### Databases

Relational database

 PostgreSQL

Document database

 mongoDB

Linked data store

 OPENLINK VIRTUOSO

### Advanced Processing Engines

Distributed processing engine

 APACHE Spark

Search engine

 elasticsearch

### Orchestration

Workflow manager

 Apache Airflow

# 04

## BDTI Hands-on use case

Supporting the Green Deal data space,  
*Creating insights from EU electric transport data*



**KAMIL RADOMSKI**

Data Expert  
BDTI team



# DIGITAL EUROPE PROGRAMME

## Our approach



### Collect & explore



### Transform & analyze



### Visualize



### BDTI benefits



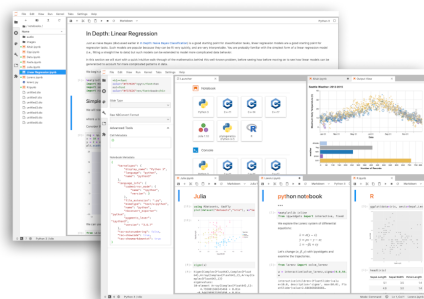
- Intuitive self-service platform that facilitates each step
- On-demand compute and storage resources
- Open-source tooling that is reproduceable in your own environment

# DIGITAL EUROPE PROGRAMME

## Our approach: Collect & explore



Collect & explore



<https://jupyter.org>

### Identified datasets:

- Electric **infrastructure** in Europe:
  - Number of **charging points** (.csv)  
Alternative fuels observatory - <https://alternative-fuels-observatory.ec.europa.eu/>
- Number of **vehicles** per country (.csv)  
Eurostat - [https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=road\\_eqs\\_carhab&lang=en](https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=road_eqs_carhab&lang=en)
- Number of **electric vehicles** per country (.csv)  
Eurostat - [https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=road\\_eqs\\_carpda&lang=en](https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=road_eqs_carpda&lang=en)
- **Land Area** of each European country (.csv)  
Eurostat - [https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=reg\\_area3&lang=en](https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=reg_area3&lang=en)
- Location of **charging points** in **Romania** (.json)  
OpenChargeMap - <https://github.com/openchargemap>



# DIGITAL EUROPE PROGRAMME

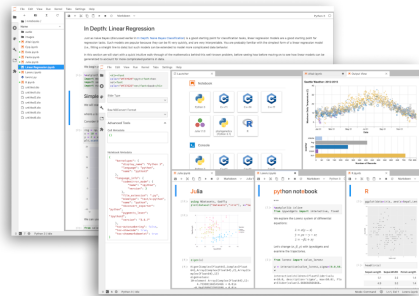
## Our approach: Transform & analyze



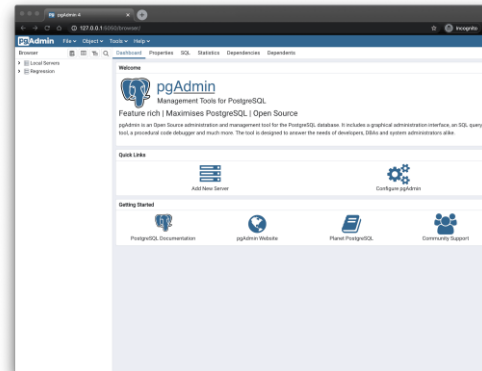
### Transform & analyze

#### Data preparation & manipulations

- **Cleaning & preparing** the data
- **Join** datasets together
- **Load in a database** for serving the data



<https://jupyter.org>



<https://www.pgadmin.org>

# DIGITAL EUROPE PROGRAMME

## Our approach: Visualize



Visualize



<https://superset.apache.org>

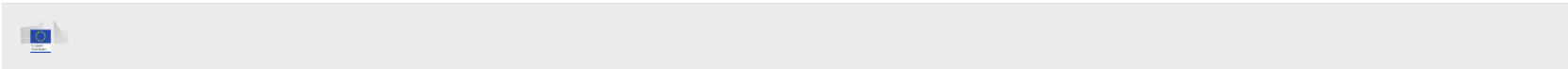
### Data Visualization

- Deliver **insights** through
  - Charts,
  - Maps,
  - Tables
- Build a **story** and make **data-driven decisions**



# DIGITAL EUROPE PROGRAMME

Let's get our hands dirty

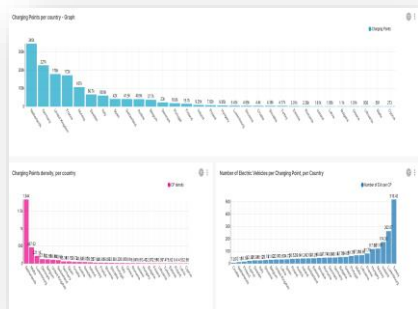


Log In

# DIGITAL EUROPE PROGRAMME

## What did we learn?

### BDTI helps you doing data analytics



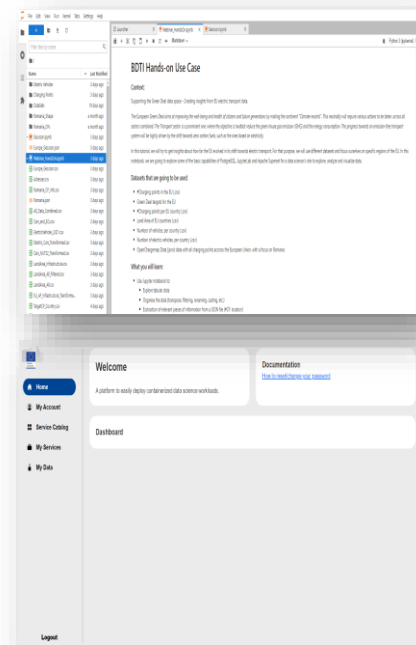
**Delivering insights through visualizations**



**Exploring & preparing various datasets**



**Leveraging BDTI technologies**



# 05

## Support & how to apply



**FRANCESCA VELLA**

Pilot Request Manager  
BDTI team

# DIGITAL EUROPE PROGRAMME

## How BDTI will support your journey

### DISCOVER

Find out the BDTI services and how they leverage insights from data to support the policy-making process and the onboarding of stakeholders.

### START

Support the set up of the pilot and enable users to benefit from advisory, tutorials and knowledge base.

### FINISH

Support in implementing in-house BDTI test infrastructure at the end of the pilot projects. **Exit package for industrialization.**



### EVALUATE

Help to identify the specific business needs and translate these into requirements (techniques, software, infrastructure requirements, reporting tools) and choose the best BDTI offering.

### USE

Support the implementation of the pilot and give technical assistance during the pilot.

### SHARE

Support users in sharing their results and in presenting the pilot highlights through a workshop (Feedback collection).

# DIGITAL EUROPE PROGRAMME

## Who can apply for BDTI



### EUROPEAN PUBLIC ADMINISTRATIONS

All European Public Administrations at local, regional and national level can independently apply for a BDTI pilot project



### EUROPEAN AGENCIES

All European Agencies can independently apply for a BDTI pilot project



### ECOSYSTEM WITH ACADEMIA AND PRIVATE SECTOR

Academia, spin-off, startups can apply for develop pilot projects as long as there is a **clear collaboration** with a Public Administration/European Agency that will represent the main point of contact for the project

# DIGITAL EUROPE PROGRAMME

## How to get started with BDTI



### GET FAMILIAR WITH BDTI SERVICES

BDTI offers a safe environment to experiment with public sector information and open-source tools before deploying them in your own production environment



### DEFINE YOUR DATA ANALYTICS USE CASE

Build your BDTI use case. The support team is available to guide the users through the process of defining their organisation's use case, as well as to clarify any preliminary questions



### SUBMIT YOUR "BDTI PILOT REQUEST"

Request to use BDTI by submitting an online form. The user provides information on the online form in order to clarify the pilot objective (general summary, short description and any supporting evidence)



### ELABORATE YOUR BUSINESS AND TECHNICAL NEED

The functional team will work with you on the elaboration of your business case to ensure that it fits within the requirements, while our technical team will design your BDTI test environment which is tailored for your specific technical needs



### PILOT PROJECT APPROVAL

The European Commission will give a final validation to run the pilot project.



### TEST ENVIRONMENT SET UP

The BDTI technical team will set-up your tailored cloud-based data analytics environment so that you can start with your big data experiments.



### START THE PILOT AND GAIN INSIGHT

The users can experiment with data on the test infrastructure provided by BDTI

An illustration in the top left corner shows four stylized figures: a woman in a yellow top holding a tablet, a man in a white lab coat, a man in a white shirt holding a smartphone, and a man in a yellow shirt sitting at a laptop. They are surrounded by various digital icons including a server rack, a heart with a pulse line, a bar chart, a shield, and a factory. The background consists of purple and white wavy lines and dots.

# DIGITAL EUROPE PROGRAMME

## How to request a BDTI pilot

**DROP US A LINE TO REQUEST A BDTI PILOT**

BDTI Functional Mailbox:

**[EC-BDTI-PILOTS@ec.europa.eu](mailto:EC-BDTI-PILOTS@ec.europa.eu)**

# DIGITAL EUROPE PROGRAMME

## Use cases acceptance criteria

### BUSINESS CRITERIA

#### Potential users

EU public administration at national, regional or local level and EU agencies, in potential collaboration with:

- Academia, spin-off
- Startups and companies

#### Clear value added

Business and technical

#### Clear contact point

for the entire pilot



### FUNCTIONAL CRITERIA

#### Pilot duration

6 months

#### Pilot use cases

Only use case in scope

#### Resource usage limit

Based on EC budget

#### Skills/Maturity level

Adequate skilled resources in data literacy and public sector information





**Q&A time**

**#DigitalEUProgramme**

# Thank you for your attention!

## DROP US A LINE

BDTI Functional Mailbox: [EC-BDTI-PILOTS@ec.europa.eu](mailto:EC-BDTI-PILOTS@ec.europa.eu)

**Learn more about the Big Data Test Infrastructure**  
<https://digital-strategy.ec.europa.eu/en/policies/bdti>

**Learn more about the BDTI success stories**  
<https://digital-strategy.ec.europa.eu/en/related-content?topic=190&type=185>

**Big Data Test Infrastructure FAQ**  
<https://digital-strategy.ec.europa.eu/en/faqs/big-data-test-infrastructure>

**Recording of the last BDTI webinar on March 2nd**  
[https://www.youtube.com/watch?v=Vyig44\\_fv88](https://www.youtube.com/watch?v=Vyig44_fv88)