# Enabling a data-informed public sector:


## *From hype to action using the **Big Data Test Infrastructure (BDTI)***

Maria Claudia BODINO- Team Leader Digital, Data and Innovation Policy
Aurora GUALTIERI – Customer Relationship Manager
**EC-BDTI-PILOTS@ec.europa.eu**

European Commission

**Business Owner:**
**DG CNECT**
Directorate-General for Communications Networks, Content and Technology

**Service Provider:**
**DG DIGIT**
Directorate-General for Digital Services

# Public Sector Information and the role of Data analytics

Data is everywhere and growing at an unprecedent pace.
Big Data: 3V - Volume, Variety, Velocity


Data is a key ingredient for services, products, and effective policy making.

There is an ambition to create a single European market for data and make more data available through powerful and trustworthy infrastructures and technologies, in line with EU values and regulations, to support citizens, public sector and companies.

Learn more about the Digital Europe Programme: https://digital-strategy.ec.europa.eu/en/activities/digital-programme

# Policy timeline


## The European Data Strategy
**February 2020**

Data governance to **facilitate data sharing and reuse across sectors** and MS.

## The Open Data Directive
**July 2020**

**Re-use** of **public sector information.** **high-value datasets** *free* via *API.* Geospatial ,Earth observation and environment, Meteorological, Statistics ,Companies and company ownership ,Mobility
>> Cross border dimension.

## Data Governance Act
**November 2020**

Foster the **availability** and **data-sharing** **data intermediaries** **common European data spaces** in strategic domains.
(Essential to train AI models).

## The AI Package
**April 2021**

It provides AI **developers** and **deployers** with clear requirements and obligations regarding specific uses of Ai. The new regulation will apply two years after its entry into force, with some exceptions for specific provisions.

## The Data Act
**February 2022**

**Who** can create value from data
Under **which** conditions
Rules for a fairer access to and reuse data from different sectors
**Agriculture, Smart cities, Health.**

## The Interoperable Act
**November 2022**

It complements the EU data and digital policy landscape on data availability and data exchange, **from a public sector angle.**
It implements **interoperability by design** and fosters the sharing and reuse of interoperable solutions.

## AI Data Act
**December 2023**

It is the **first-ever comprehensive legal framework on AI**. The aim is to **foster trustworthy** AI in Europe by ensuring that AI systems **respect** fundamental rights, safety, and ethical principles and by addressing risks of very powerful and impactful AI models.

Policy vision

https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en

# The Data journey is like cooking a dish

Business challenge/question/problem statement

Ingredients + tools

1. **Find** the data you need → *gather the right ingredients, good quality*
2. **Get, clean** and **prepare** your data → *slice and dice*
3. **Analyse** your data → *mix ingredients together and try different combinations*
4. **Present** the results and **create** knowledge → *serve and consume*

Data→ Information → Knowledge





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


BTDI: Not **only** for big data, but for all **public sector information**

Six months free* of charge service for the EU public administrations

Ready-to-use data analytics stack and support

Cloud platform based on open-source tools

To help the public sector to derive insights from its data and accelerate transition towards data-driven decision making

* The cost of the pilot project must fit within the funding boundaries of the BDTI pilot budget

# Big Data Test Infrastructure Objectives

Objectives

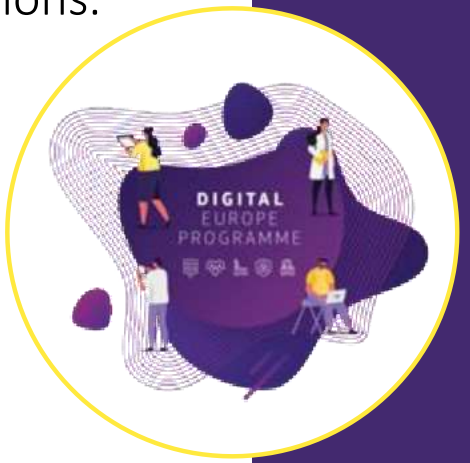
- Increase the easy accessibility, interoperability, quality and usability of public sector information in compliance with the requirement of the Open Data Directive
- Boost the re-use and combination of open public data across the EU for the development of information products and services, including AI applications
- High-value Datasets – Open Data Directive
- Testing Business-to-Government (B2G) data sharing collaborations for the public good
- Data Space Support Centre: explore and experiment with your data*
- BDTI provides a safe testing environment to run big data experiments for data space customers


* https://joinup.ec.europa.eu/collection/semic-support-centre/data-spaces

# About The Big Data Test Infrastructure (BDTI)

The BDTI is funded by the Digital Europe Programme (DEP)
focused on bringing digital technology to businesses, citizens and public administrations.

The DEP provides 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**

* https://joinup.ec.europa.eu/collection/semic-support-centre/data-spaces

# Open-source tools to support your data journey


**Orchestration**

Apache Airflow

MAGE

**Visualisation**

Metabase

Apache Superset

**Development Environments**

KNIME

jupyter

H2O.ai

RStudio

QGIS

**5.** Decision-making

**1.** Collection

**2.** Processing

**3.** Analysis

**4.** Visualisation

MINIO — **Data Lake**

OPENLINK VIRTUOSO UNIVERSAL SERVER

mongoDB.

*+ PostGIS extension*

**Database**

doccano — **Data Annotation**

elasticsearch

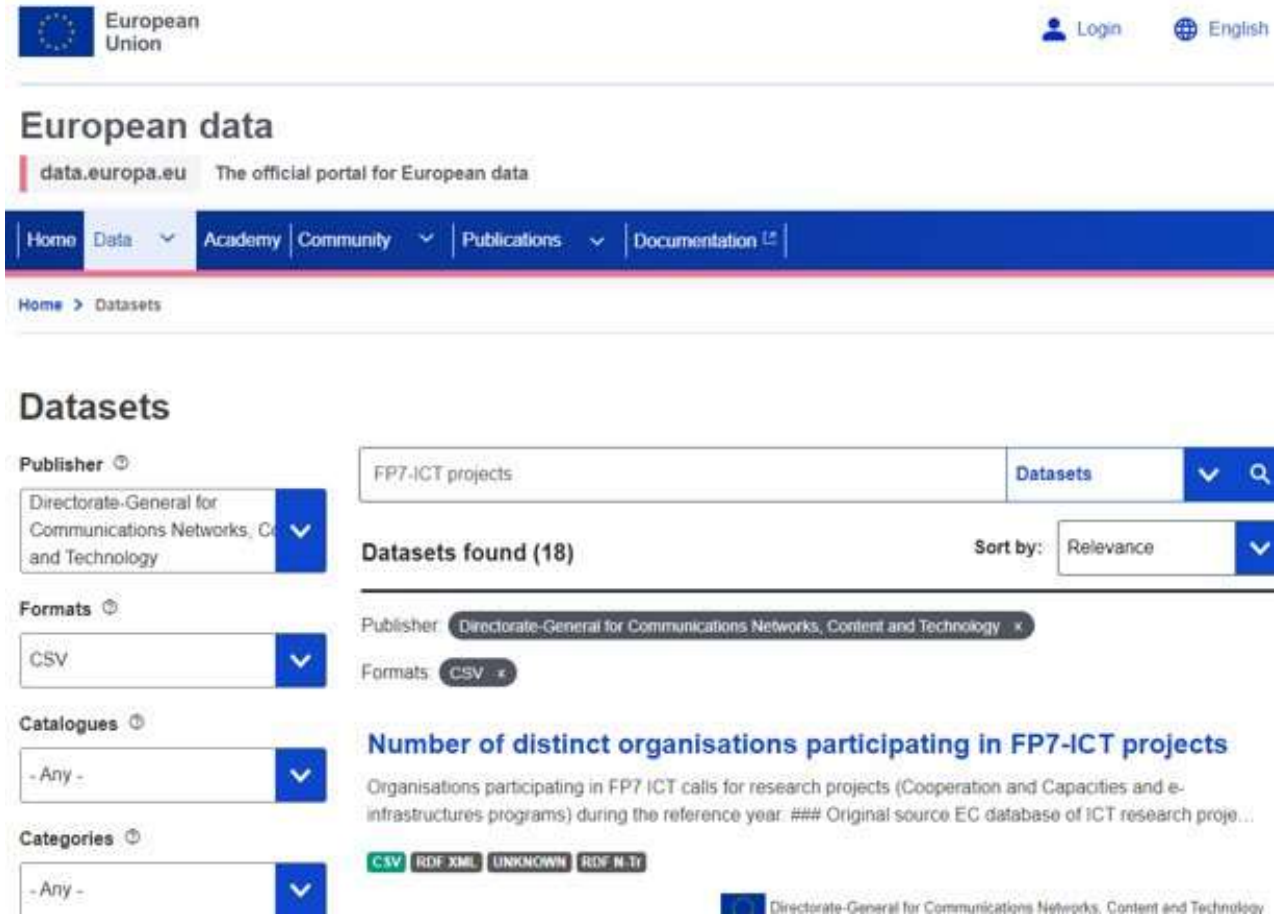
Apache Spark

kibana

**Advanced Processing Engines**

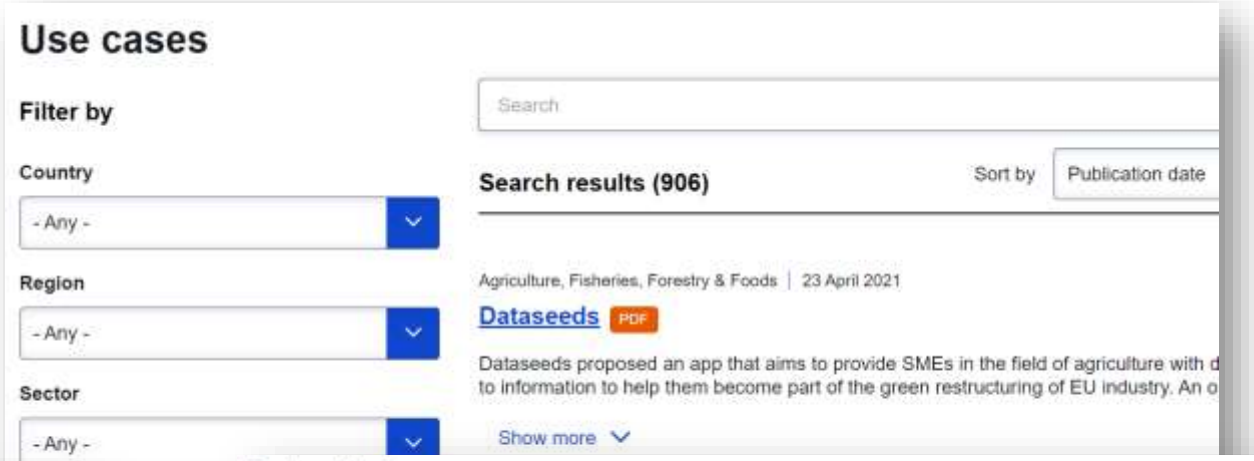
# Data.europa.eu as data hub

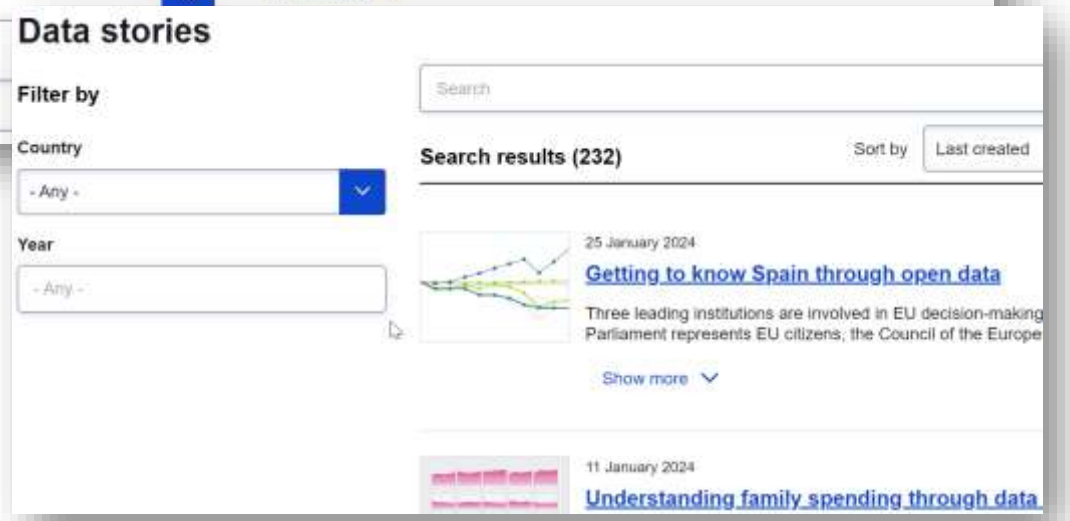

**More than 1.6 million datasets**, grouped in 183 data catalogues
**Navigate or search to get to the data** or catalogue you are looking for
Benefit from **many filters**
**Metadata translations** in all EU languages, machine translations for other text
**Download and transform CSV files** automatically in many different formats
Get **quick visualisations** for geo datasets
Get **feedback for the metadata quality** on how to improve it



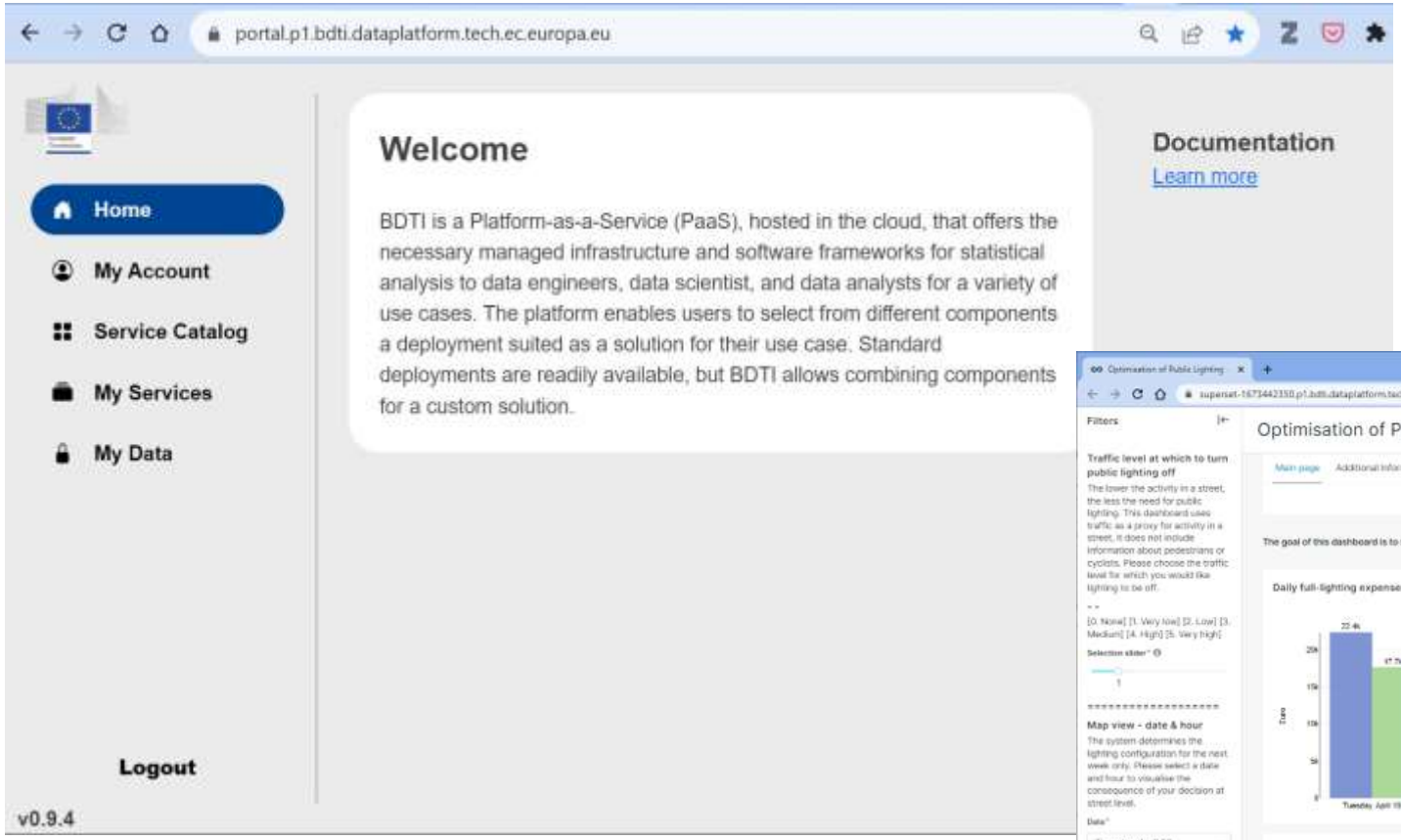
# Data.europa.eu: get inspired

Consult our list of [use cases](#) (more than 900 examples)
Follow re-users through the **use case observatory**
Read our [data stories](#)

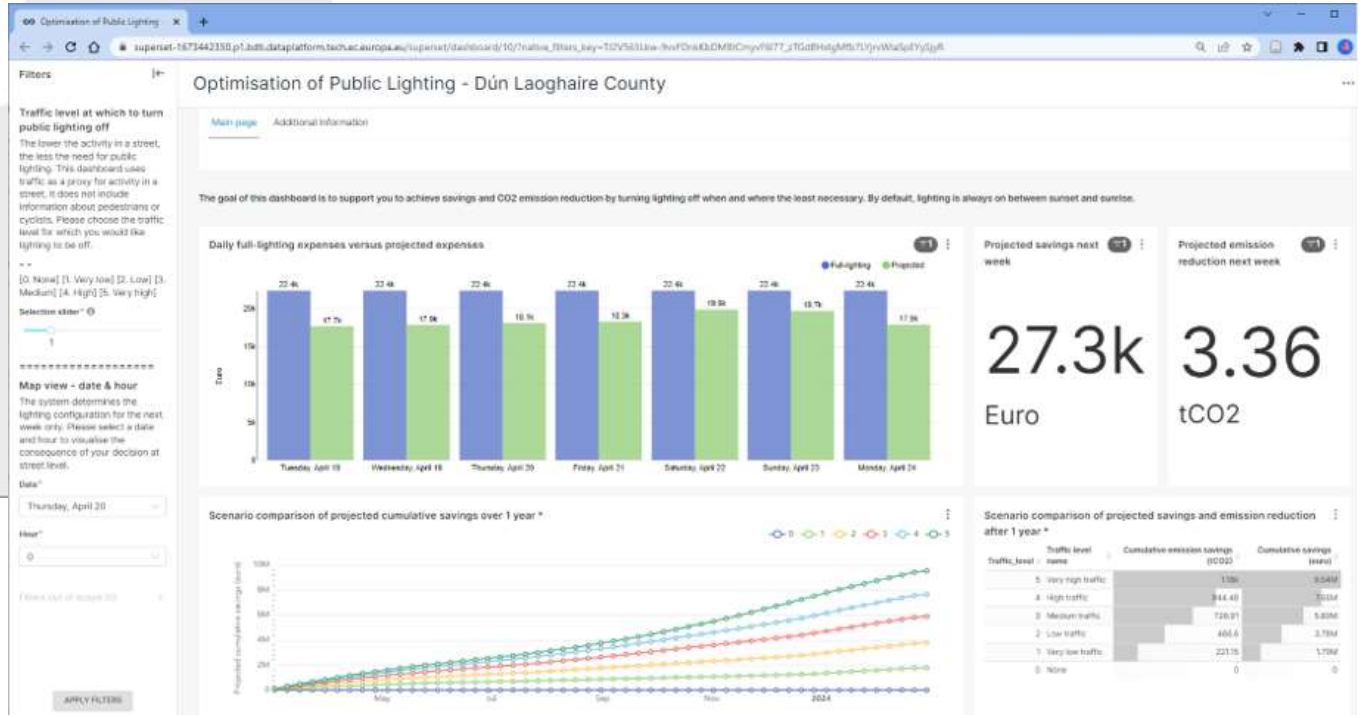



# The BDTI Portal

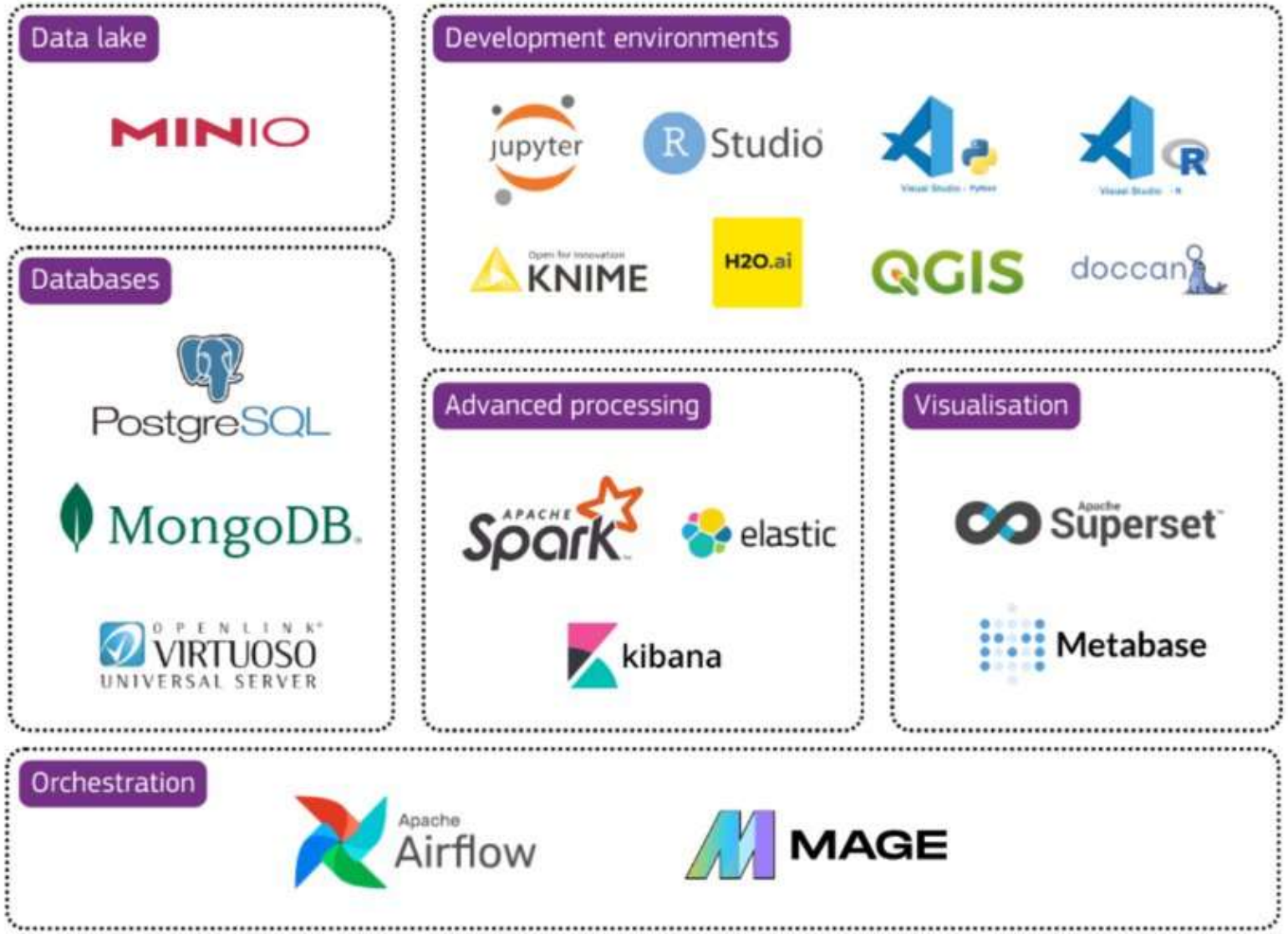


https://youtu.be/fdzNCB1CVUM



# The BDTI portal

**100% ❤ open-source components**

# Who is BDTI for ?


## European Public Administrations

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

## Partnerships with academia and the private sector

Academia (master, PhD students) and startups and companies (GovTech sector)
can apply if they are collaborating with a public administration

# Why use BDTI?


**Benefit of six months free of charge,** including advisory and technical support

**Experiment with data analytics** using high performance infrastructure that leverages the power of the **elastic cloud**

**Receive guidance** to move from a pilot to a production-ready process

**Test your idea → Extract value → Create knowledge**

# Who used it already?


## 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 extracting knowledge from the huge quantity of scientific clinical articles coming from different sources (i.e. PubMed.gov, Covid-19 related clinical articles).

GENERALITAT VALENCIANA
Conselleria de Sanitat Universal i Salut Pública

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

## 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.

eba EUROPEAN BLOOD ALLIANCE

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

## 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 and about **how services can be redesigned to foster cultural activities and events.**

Città di FIRENZE

**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).


## Italy, Portugal and Norway - E-Procurement data

This e-procurement pilot involved Italian, Portuguese, and Norwegian authorities and centred around providing a scalable virtual environment and analytics routines to work on procurement data and support the creation of the procurement data space. The goal was to develop a common framework using open-source software and infrastructure for monitoring EU public procurement through analytical services and tools based on a common data model.

**Predictive time-series analysis** and **data transformation** tools help Italian, Portuguese and Norway authorities develop an EU procurement common framework based on a common data model. It would enable public administration, businesses, and citizens in Europe to benefit from **interoperability and cross-border public procurement services.**

## GRNET and University of Macedonia - Linked data

GRNET, together with the University of Macedonia, aimed to transform data from the MITOS API (Greek National Registry of Administrative Public Services, which provides structured descriptions of over 3,000 public services) into Linked Data aligned with EU standards like CPSV-AP and CCCEV. This pilot, called MitosLOD, is periodically collecting data via the MITOS API, transforming them into Linked Data using python and AirFlow and storing them in a Virtuoso RDF store. The aim of the project is also to provide a SPARQL endpoint for data querying and retrieval, with plans to explore data visualisation to assist citizens with public service information.

grnet

UNIVERSITY OF MACEDONIA

**Advanced data visualisation and transformation tools** are being used to convert the data gathered from the MITOS API into Linked Data, aligned with EU standard models such as CPSV-AP and CCCEV. **This transformation improves the accessibility and integration of public service data, enabling better service delivery through linked open data and advanced query mechanisms.**

## CITY OF BOCHUM – Urban data

The City of Bochum currently uses five tree sensors to monitor data such as resistivity, temperature, and humidity, along with other weather-related factors. The goal is to first implement a real-time data visualisation system. Additionally, they aim to develop a machine-learning model that combines this sensor data with information from soil moisture sensors to predict the health of trees in Bochum.

BOCHUM

**Predictive, time-series and data visualisation analysis on multiple datasets collected from five Tree sensors**. The team wants firstly to have a system that visualises this data in real time and additionally to create a Machine Learning model that considers the data and, by combining them with additional soil moisture sensors, **to calculate and predict the tree health.**


## CITY OF TURKU – Mobility data

The Municipality of Turku, in collaboration with the University of Turku, is working on a pilot to analyse traffic flows and improve public transport efficiency. They are combining various mobility data sources and geodata with BDTI's tools. The long-term objective includes determining suitable locations for dedicated bus lanes and assessing how these changes would impact traffic flow and bus connection speed. The project also aims to address public transport capacity and Park & Ride hotspot locations.

TURKU
ÅBO

**Predictive, descriptive and time-series analysis** on advanced data transformation and visualisation, including datasets from public Wi-Fi sensors, parking systems, and geo-referenced movement patterns such as tourist activity, leading **to enhanced mobility solutions in the future.**

## CITY OF NAPLES – Mobility data

The Municipality of Naples is using advanced analytics on public space and mobility data to support urban planning. The pilot seeks to redesign public spaces and improve citizen participation in mobility strategies, with a focus on climate resilience, such as relief hubs for extreme heat events. By integrating urban morphology, mobility opinions, and green capital data, it aims to enhance planning decisions. Open data sources, such as OpenStreetMap and Urban Heat Island data, are used to address gaps in unpublished and outdated data.

COMUNE DI NAPOLI

**Predictive, descriptive and time-series analysis** on multiple datasets collected related to public spaces and mobility in order to streamline citizens' participation and to build a transportation dataset **tailored for Public Administration** in compliance **with the open data directive.**

From Hype to Action


# The BDTI Canva

by the BTDI Team

The BDTI Canva aims to help you build a strong data use case through a series of questions.

For more information, visit the BDTI website

Contact us by emai: EC-BDTI-PILOTS@ec.europa.eu

**Context:**
Who are you? Who are your stakeholders?

**Objective(s):**
What is the problem you are trying to address?
What is your timeframe?

**Data's added value:**
Which information helps you address the problem? From which sector and or domain?

**Data's availability:**
Does the data you need exist?If it doesn't exist, can you collect it? From whom can you get the data you need? Can you reuse the data? What license applies to the data you'd like to use? How is the quality of the data you'd like to use? Are the different datasets interoperable? Do you know how to connect the dots?

**Data's risk(s):**
What could go wrong when using data to address this objective? Are there legal and ethical considerations to make? Are you dealing with personal data?

**Data's processing:**
What do you need to gather, process and analyze the data (i.e., tools, software, computing power, ...)? Do you already have them? If you do not, where can you get them (e.g., applying to the BDTI)?

**Data skills:**
What data literacy and skills do you need (i.e., data engineering, data analysis, data science, data visualization)? Do you already have these available within your team/organization?

**Your solution**
Combine what you've learned from the elements above into a statement describing your solution

# How to apply: a fast and simple process


Get familiar with the BDTI service on our website

Define your data analytics use case using our BDTI Canva and then fill in the template request form (see website)

Submit your pilot request (template) by email

Meet with us to elaborate on your use case

Pilot Project is approved if:

- Brings value
- It can be done in 6 months
- Sufficient resources available (skills, team)

Your test environment is set up

You can start piloting and create value!

# BDTI Translated in the 24 EU Official Languages

**Home -> Resources -> Promotional Material**



Translations are available here : https://big-data-test-infrastructure.ec.europa.eu/resources_en

# The BDTI Kitchen: Baking data Newsletter


## The BDTI Kitchen: Baking Data

European Commission


### In this issue



## Data literacy corner

### Guide to navigating datasets

The data.europa.eu search guide provides a comprehensive look at navigating European datasets, enabling public sector workers to find relevant data faster and more effectively.

more

## BDTI latest news

### Arezzo pilot launches

The municipality of Arezzo, Italy, has launched a pilot under the BDTI initiative to enhance urban planning through time-series analysis of public space and accident data. This pilot project is part of BDTI's commitment to supporting European public administrations in experimenting with open-source data analytics tools to foster safer and more sustainable cities.

more

## Upcoming events

### Harnessing climate data for tourism [live workshop]

This workshop will focus on climate data analysis techniques to help regions understand and predict tourism trends amid changing environmental conditions. Learn to make accurate predictions and informed decisions by uncovering patterns and trends in historical data in this classification and predictive analytics workshop. Secure your spot now!


## Spotlight on public sector initiatives

### Greece's MitosLOD pilot up for Best Cases Award

The BDTI-powered pilot MitosLOD, a project leveraging linked open data for public sector interoperability in Greece, has been nominated for a best practices award. This pilot exemplifies how public sector data can be transformed into actionable insights for improved governance.

more

# BDTI Pilot Stories


**Telling** your data story

**Inspiring** others to seek data-driven insights

## The MitosLOD Pilot Story: Transforming Greek Public Services with Linked Open Data

GRNET and UoM aim to transform MITOS, which provides structured descriptions of over 3,000 public services, into Linked Open Data.

View draft | **View published** | Edit draft | History | Translate | Unpublish


Data-driven decision-making is transforming the public sector. Armed with this knowledge, Greece's National Infrastructures for Research and Technology (GRNET S.A.) and the University of Macedonia (UoM) have embarked on a ground-breaking pilot project with the Big Data Test Infrastructure (BDTI) as the backbone. Their mission is to leverage the BDTI to convert the Greek National Registry of Administrative Public Services (MITOS) into Linked Open Data, thereby enhancing transparency, efficiency, and accessibility of public services. This article delves into the journey of the MitosLOD project, highlighting its innovative approach and the significant role of the BDTI environment.

## Norwegian Digitalisation Agency uses Commission Big Data tool to optimise public procurement

Digdir, the Norwegian Digitalisation Agency, used the Big Data Test Infrastructure (BDTI) to help optimise public procurement in Norway.

Digdir has been helping Norwegian public administrations in their digital transformation for six years. They have already implemented a PEPPOL network, based on the **eDelivery** and **eInvoicing** building blocks, to digitalise invoicing in public procurement. This network allows public and private organisations to seamlessly exchange and automatically process digital invoices.

Digdir was looking to help improve digital public procurement, or eProcurement, even further by gathering and analysing big datasets on transactions in this area.

© metamorworks - iStock Getty Images Plus

## Experimenting with BDTI led to new dimensions in labour market intelligence

Establishing objective information about the labour market is important for policy and decision makers to ensure Europe's competitiveness and that the skills of the workforce meet demand.

With a significant share of job offers advertised online, the data contained within them hold valuable pieces of information about the current job market. Tapping into this data with big data technologies can help to understand, for example, trends in existing and emerging skills required and the number of job vacancies. Establishing objective information about the labour market is important for policy and decision makers to ensure Europe's competitiveness and that the skills of the workforce meet demand. When working on new data-driven solutions, the European Commission's big data testing tool helps to experiment with real data before starting system development.

© Smederevac - iStock Getty Images Plus

# BDTI Skills Studio

## Past Events


Did you miss these workshops?
Watch the recording here:
https://code.europa.eu/bdti/bdti-skills-studio

## What's Next?


Register for upcoming events

# Thank you for your attention!


Are you working for a public administration in need of infrastructure for data analytics? Get in touch

EC-BDTI-PILOTS@ec.europa.eu

***BDTI website***

***BDTI's Joinup page***

***BDTI's* newsletter**

# Links and resource


- https://big-data-test-infrastructure.ec.europa.eu/
- https://code.europa.eu/bdti/bdti-demonstrator
- https://commission.europa.eu/publications/interoperable-europe-act-proposal_en
- https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en

  https://digital-strategy.ec.europa.eu/en/policies/legislation-open-data
- https://digital-strategy.ec.europa.eu/en/policies/data-governance-act
- https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence
- https://digital-strategy.ec.europa.eu/en/activities/digital-programme
- https://dssc.eu/wp-content/uploads/2023/03/DSSC-Data-Spaces-Glossary-v1.0.pdf
- https://digital-strategy.ec.europa.eu/en/library/staff-working-document-data-spaces
- https://ec.europa.eu/commission/presscorner/detail/en/ip_22_1113
- https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32019L1024
- https://joinup.ec.europa.eu/collection/egovernment/solution/big-data-test-infrastructure-bdti