

Развитие на основан на данни публичен сектор

От думи към дела —

Big Data Test Infrastructure (BDTI)

(Тестова инфраструктура за големи информационни масиви)

EC-BDTI-PILOTS@ec.europa.eu



DG CNECT

ГД „Съобщителни мрежи,
съдържание и технологии“

DG DIGIT

Генерална
дирекция
„Информатика“



Информация от общественения сектор и роля на анализа на данни (data analytics)

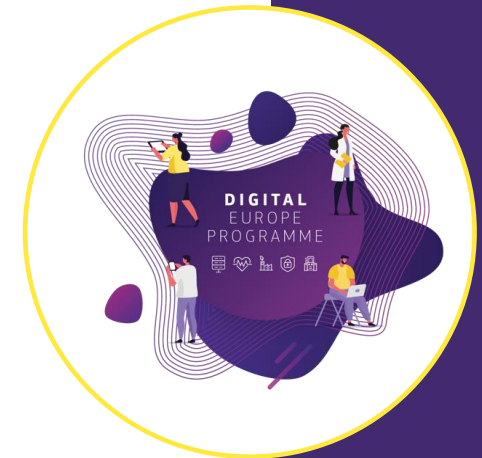
Данните са навсякъде и техният обем нараства както никога.

Big Data (големи информационни масиви) означава обем, разнообразие и скорост!



Данните са ключов елемент за **услугите, продуктите и ефективното изготвяне на политики.**

Защо да нямаме единен европейски пазар на данни и достъп до повече данни чрез мощни и надеждни инфраструктури и технологии, **в съответствие с ценностите и разпоредбите на ЕС, в услуга на гражданите, публичния сектор и дружествата?**



Какво представлява Big Data Test Infrastructure (BDTI)?



BDTI се отнася **не само** за големите информационни масиви, **но** и за всички **данни в общественния сектор**



Шест месеца **безплатно***
обслужване
за публичните администрации в ЕС



Готов за използване пакет
(stack) за анализ на данни и
поддръжка



Облачна платформа (Cloud
platform), базирана на
инструменти с **отворен код**
(open-source)



Публичният сектор ще може да **извлича информация** от данните, с които **разполага**, и да ускори прехода към **основано на данни вземане на решения**.

* Разходите за пилотния проект трябва да са в рамките на възможностите за финансиране на бюджета за пилотни проекти за BDTI

Какви са целите на BDTI?

- По-достъпна, по-съвместима, по-качествена и по-използваема информация от обществения сектор в съответствие с изискването на **Директивата за отворените данни (Open Data Directive)**;
- Насърчаване на **повторното използване и комбинирането на свободно достъпни данни от обществения сектор** в целия ЕС за разработване на информационни продукти и услуги, включително приложения с ИИ;
- **Набори от данни с висока стойност — Директива за отворените данни;**
- Изпитване на сътрудничествата при споделянето на данни от **предприятията към правителствата (B2G)**
- Център за поддръжка на пространствата на данни (Data space): проучвайте и експериментирайте с вашите данни*
- BDTI осигурява безопасна **изпитателна среда за провеждане на изпитвания с големи информационни масиви** за потребители на пространства на данни

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



Относно Big Data Test Infrastructure (BDTI)

BDTI се финансира по Програмата „Цифрова Европа“, чиято цел е да **приблужи цифровите технологии** до предприятията, гражданите и публичните администрации.

Програмата „Цифрова Европа“ предоставя стратегическо финансиране в **пет ключови области**:

**Високопроизводителни
изчислителни технологии**

Киберсигурност

Изкуствен интелект

(Изчислителен облак, данни и изкуствен интелект)

**Усъвършенствани цифрови
умения**

Внедряване и широко използване на цифрови технологии



Инструменти с отворен код за цикъла на вашите данни



Оркестрация
(Orchestration)



5. Вземане
на решения

1. Събиране

MINIO

Езеро с
данни
(Data Lake)



mongoDB

База данни
(Database)



Metabase

4. Визуализация

2. Обработка

elasticsearch

Усъвършенствани
компютри за
обработка
(Advanced Processing
Engines)



Apache
Superset

Среда на
разработване
(Development
Environments)

3. Анализ



Портал на BDTI



portal.p1.bdti.dataplatform.tech.ec.europa.eu

European Commission

- Home
- My Account
- Service Catalog
- My Services
- My Data

Logout

v0.9.4

Welcome

BDTI is a Platform-as-a-Service (PaaS), hosted in the cloud, that offers the necessary managed infrastructure and software frameworks for statistical analysis to data engineers, data scientist, and data analysts for a variety of use cases. The platform enables users to select from different components a deployment suited as a solution for their use case. Standard deployments are readily available, but BDTI allows combining components for a custom solution.

Documentation

[Learn more](#)

BIG DATA TEST INFRASTRUCTURE (BDTI)

Watch on YouTube

<https://youtu.be/fdzNCB1CVUM>

Optimisation of Public Lighting - Dún Laoghaire County

The goal of this dashboard is to support you to achieve savings and CO2 emission reduction by turning lighting off when and where the least necessary. By default, lighting is always on between sunset and sunrise.

Daily full-lighting expenses versus projected expenses

Date	Full-lighting (Euro)	Projected (Euro)	Savings (%)
Tuesday, April 18	22.4k	17.7k	17.7%
Wednesday, April 19	22.4k	17.5k	17.5%
Thursday, April 20	22.4k	16.5k	16.5%
Friday, April 21	22.4k	16.3k	16.3%
Saturday, April 22	22.4k	19.9k	19.9%
Sunday, April 23	22.4k	16.7k	16.7%
Monday, April 24	22.4k	17.3k	17.3%

Projected savings next week

27.3k Euro


Projected emission reduction next week

3.36 tCO2

Scenario comparison of projected cumulative savings over 1 year *

Traffic level name	Cumulative emission savings (tCO2)	Cumulative savings (Euro)
5 Very high traffic	179k	3,342k
4 High traffic	84.4k	2,024k
3 Medium traffic	726.91	5,809k
2 Low traffic	466.0	3,784k
1 Very low traffic	221.5	1,754k
0 None	0	0

Портал на BDTI

100 % 
КОМПОНЕНТИ С
ОТВОРЕН КОД



Service Catalog

Airflow - v2.3.0

Description

Airflow is a platform created by the community to programmatically author, schedule and monitor workflows.

[Launch](#)

Jupyterlab - lab-3.2.8 - datascience-notebook

Description

The Jupyter Notebook is a web application for creating and sharing documents that contain code, visualizations, and text. It can be used for data science, statistical modeling, machine learning, and much more.

[Launch](#)

Knime - v4.5.3

Description

KNIME Analytics Platform is the open source software for creating data science. Intuitive, open, and continuously integrating new developments, KNIME makes understanding data and designing data science workflows and reusable components accessible to everyone.

[Launch](#)

PgAdmin4 - v6.8

Description

PgAdmin is the most popular and feature rich Open Source administration and development platform for PostgreSQL, the most advanced Open Source database in the world.

[Launch](#)

Apache Superset - v1.0

Description

Apache Superset is a modern data exploration and visualization platform. It is fast, lightweight, intuitive, and loaded with options that make it easy for users of all skill sets to explore and visualize their data, from simple line charts to highly detailed geospatial charts.

[Launch](#)

Jupyterlab - lab-3.4.2 - all-spark-notebook

Description

The Jupyter Notebook is a web application for creating and sharing documents that contain code, visualizations, and text. It can be used for data science, statistical modeling, machine learning, and much more. Used for spark.

[Launch](#)

Knime - v5.1.0

Description

KNIME Analytics Platform is the open source software for creating data science. Intuitive, open, and continuously integrating new developments, KNIME makes understanding data and designing data science workflows and reusable components accessible to everyone.

[Launch](#)

Postgresql - v14.2.0

Description

PostgreSQL is a powerful, open source object-relational database system with over 30 years of active development that has earned it a strong reputation for reliability, feature robustness, and performance.

[Launch](#)

Apache Superset v2.1

Description

Apache Superset is a modern data exploration and visualization platform. It is fast, lightweight, intuitive, and loaded with options that make it easy for users of all skill sets to explore and visualize their data, from simple line charts to highly detailed geospatial charts.

[Launch](#)

Jupyterlab - lab-4.0.4 - all-spark-notebook

Description

The Jupyter Notebook is a web application for creating and sharing documents that contain code, visualizations, and text. It can be used for data science, statistical modeling, machine learning, and much more. Used for spark.

[Launch](#)

Metabase - v0.43.3

Description

Metabase sets up in five minutes, connecting to your database, and bringing its data to life in beautiful visualizations. An intuitive interface makes data exploration feel like second nature—opening data up for everyone, not just analysts and developers.

[Launch](#)

RStudio - v4.1.2

Description

An integrated development environment for R and Python, with a console, syntax-highlighting editor that supports direct code execution, and tools for plotting, history, debugging and workspace management.

[Launch](#)

ElasticSearch - v7.17.3

Description

Elasticsearch is the distributed, RESTful search and analytics engine at the heart of the Elastic Stack.

[Launch](#)

Jupyterlab - lab-4.0.4 - datascience-notebook

Description

The Jupyter Notebook is a web application for creating and sharing documents that contain code, visualizations, and text. It can be used for data science, statistical modeling, machine learning, and much more.

[Launch](#)

MinIO - RELEASE.2022-06-20T23-13-45Z

Description

MinIO offers high-performance, S3 compatible object storage. Native to Kubernetes, MinIO is the only object storage suite available on every public cloud, every Kubernetes distribution, the private cloud and the edge. MinIO is software-defined and is 100% open source under GNU AGPL v3.

[Launch](#)

Spark - v3.2.1

Description

Apache Spark is an open-source unified analytics engine for large-scale data processing. Spark provides an interface for programming clusters with implicit data parallelism and fault tolerance.

[Launch](#)

H2O-3 - v36.1.1

Description

H2O is an in-memory platform for distributed, scalable machine learning. H2O uses familiar interfaces like R, Python, Scala, Java, JSON and the Flow notebook/web interface, and works seamlessly with big data technologies like Hadoop and Spark.

[Launch](#)

Kibana - v7.17.3

Description

Kibana is your window into the Elastic Stack. Specifically, it is a browser-based analytics and search dashboard for Elasticsearch.

[Launch](#)

MongoDB - v4.4.13

Description

MongoDB® is a relational open source NoSQL database. Easy to use, it stores data in JSON-like documents. Automated scalability and high-performance. Ideal for developing cloud native applications.

[Launch](#)

Virtuoso - v7.2.7

Description

OpenLink Virtuoso is a next-generation Universal Server that facilitates the development and deployment of a new generation of Enterprise-wide, Internet, Intranet, and Extranet-based solutions, transcending prevalent enterprise challenge areas such as Disparate Databases and Data Sources, Web Service Composition, and Business Process Management.

[Launch](#)



Home

My Account

Service Catalog

My Services

My Data

За кого е предназначена BDTI?



Европейската публична администрация

Всички европейски публични администрации на **местно, регионално и национално равнище** могат да кандидатстват самостоятелно за пилотен проект, свързан с BDTI



Партньорства с академичните среди и частния сектор

Академичните среди (студенти в магистърска и докторска степен) и предприятията, включително стартиращите предприятия (сектор GovTech), могат да кандидатстват, ако си сътрудничат с публична администрация

Защо да използвате BDTI?



Услугата е **безплатна** за шест месеца, включително **консултантската и техническата помощ**



Експериментирайте с анализа на данни посредством **високоэффективна инфраструктура**, която използва мощта на **еластично облачно хранилище за данни (elastic cloud)**



Получавайте насоки как да преминете от пилотния проект към процеса на **подготовка за експлоатация**



Пробвайте идеите си → Извлекете ценното от тях
→ Генерирайте знание

Някой има ли вече опит с BDTI?



Generalitat Valenciana се нуждаеше от решение за пресяване на обемни научни клинични статии от различни източници като PubMed.gov и свързани с COVID-19 публикации. Предизвикателството се състоеше в това да се извлекат ценна информация и знания от този огромен обем данни.



Европейският кръвен алианс и ГД „Здравеопазване и безопасност на храните“ положиха съвместни усилия за създаване на всеобхватна платформа с отворен достъп в Европа. Целта на тази инициатива беше да се съберат ключови данни за целенасочено проучване относно терапията с плазма от лица, възстановили се от COVID-19. Трудността се състоеше в това да се определят най-ефективните условия за това лечение, като се използва анализ на данните, за да се вземат информирани решения и да се рационализират научноизследователските усилия.



Общината бе изправена пред сложната задача да направи връзка между различните набори от данни, с които разполага градът. Целта беше да се разберат моделите на движение на хората в различни райони и в различни периоди от време. Това предизвикателство, свързано с анализа на данни, беше от съществено значение, за разбиране на тенденциите в мобилността и за прекрояване на услугите с цел ефективно насърчаване на културните дейности и прояви.



Усъвършенствани **инструменти за визуализиране на данни и извличане на текст** в помощ на **извличането на знания**, съдържащи се в документите.

Подпомагане на клиничните специалисти и ръководителите в техните клинични практики и ежедневна работа.



Готова за използване виртуална среда, в която **данните, събрани чрез нарочно създаден уебсайт**, се приемат и анонимизират, след което се анализират с помощта на усъвършенствани инструменти за визуализиране и анализ на данните. Първоначално бяха обработени само данни за даряването, след което обхватът беше увеличен, за да се обхване **цялостния път на кръвната плазма — от даряването до пациента/клиничното изпитване**.



Прогнозен и описателен анализ и анализ на динамични редове на множество набори от данни, **събрани преди, по време на и след пандемията от COVID-19**, като например: датчици в публичното пространство, свързани чрез безжична мрежа, данни от паркинги и географски данни за движението на хора (т.е. на туристите).



Как се кандидатства: бърз и опростен процес



Запознавайте се с услугата BDTI на нашия [website](#)



Определяте как ще използвате анализа на данни съгласно с вашия случай на употреба посредством инструмента [BDTI Canva](#) и попълвате формуляра за кандидатстване (вж. [уебсайта](#))



Изпращате вашето искане за пилотен проект (образец) по електронна поща EC-BDTI-PILOTS@ec.europa.eu



Свързвате се с нас, за да опишете своя случай на употреба



Пилотният проект се одобрява, ако:
Създава стойност, може да завърши за 6 месеца, има достатъчно ресурси (умения, екип, данни)



Настройване на Вашата изпитателна среда



Можете да започнете пилотен проект и да генерирате стойност!

Благодарим за вниманието!



Работите за публична администрация, която се нуждае от инфраструктура за анализ на данни? Свържете се с нас!



EC-BDTI-PILOTS@ec.europa.eu

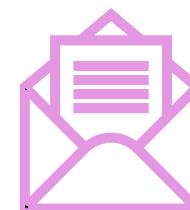
[Уебсайт на BDTI](#)



[Страница на BDTI в Joinup](#)



[Бюлетин на BDTI](#)



Връзки и ресурси

- <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>