

Dane w sektorze publicznym

*Od hajpu do konkretnych działań przy
użyciu **Big Data Test Infrastructure (BDTI)***

EC-BDTI-PILOTS@ec.europa.eu



DG CNECT

Dyrekcja Generalna ds. Sieci
Komunikacyjnych, Treści i Technologii

DG DIGIT

Dyrekcja Generalna
ds. Informatyki



Informacje sektora publicznego i rola analizy danych

Dane są **wszędzie**, a ich liczba rośnie w niespotykanym tempie.

- **Big Data**: objętość, różnorodność, prędkość



Dane są istotnym składnikiem **usług, produktów i skutecznego kształtowania polityki**.

Pojawiła się ambicja, aby stworzyć **jednolity europejski rynek danych** i udostępnić większą ilość danych za pośrednictwem potężnych i zaufanych infrastruktur i technologii, **zgodnie z wartościami i przepisami UE, dla dobra obywateli, sektora publicznego i przedsiębiorstw**.



Czym jest Big Data Test Infrastructure (BDTI)?



BDTI: stosuje się **nie tylko** do big data, ale do **wszystkich informacji sektora publicznego**



Usługa nieodpłatna przez 6
miesiące*
dla organów administracji publicznej
w UE



Gotowy do użycia stos analizy
danych plus wsparcie



Platforma chmurowa oparta na
narzędziach **open source**



Pomoc dla sektora publicznego w czerpaniu wiedzy z dostępnych mu
danych i przyspieszenie przejścia na oparty na danych proces decyzyjny

* Koszt projektu pilotażowego musi się mieścić w granicach finansowania budżetu pilotażowego BDTI

Cele Big Data Test Infrastructure

Cele

- Zapewnienie łatwego dostępu do informacji sektora publicznego oraz zwiększenie ich interoperacyjności, jakości i użyteczności zgodnie z wymogami **dyrektywy w sprawie otwartych danych**
- Wspieranie **ponownego wykorzystywania i łączenia otwartych danych publicznych** w całej UE na potrzeby rozwoju produktów i usług informacyjnych, w tym zastosowań sztucznej inteligencji
- Zbiory danych o wysokiej wartości – dyrektywa w sprawie otwartych danych
- Testowanie możliwości współpracy przy wymianie danych **biznes-administracja rządowa (B2G)** dla **dobra ogółu**
- Centrum wsparcia przestrzeni danych (data space): wykorzystaj swoje dane – badaj, sprawdzaj eksperymentuj*
- BDTI zapewnia bezpieczne **środowisko testowe do prowadzenia eksperymentów na big data** dla klientów korzystających z przestrzeni danych

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



O Big Data Test Infrastructure (BDTI)

BDTI jest finansowana **ze środków programu „Cyfrowa Europa”**, którego celem jest udostępnianie technologii cyfrowych przedsiębiorstwom, obywatelom i organom administracji publicznej.

„Cyfrowa Europa” - strategiczne finansowanie w 5 kluczowych obszarach:

Obliczenia wielkiej skali

Cyberbezpieczeństwo

Sztuczna inteligencja

(chmura, dane i sztuczna inteligencja)

Zaawansowane umiejętności cyfrowe

Wdrażanie i szerokie wykorzystanie technologii cyfrowych



Narzędzia **open source** paliwem w Twojej podróży z danymi



Portal 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	16.9k	16.9%
Friday, April 21	22.4k	16.3k	16.3%
Saturday, April 22	22.4k	19.9k	19.9%
Sunday, April 23	22.4k	15.7k	15.7%
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 *

Scenario comparison of projected savings and emission reduction after 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	3,789k
1 Very low traffic	221.5	1,795k
0 None	0	0

Portal BDTI

100% 
komponenty open
source



Service Catalog



Home

My Account

Service Catalog

My Services

My Data

Airflow - v2.3.0

Description

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

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

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

ElasticSearch - v7.17.3

Description

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Komu służy BDTI?



Europejskie organy administracji publicznej

Wszystkie europejskie organy administracji publicznej na szczeblu **lokalnym, regionalnym i krajowym** mogą niezależnie ubiegać się o udział w projekcie pilotażowym BDTI



Partnerstwo ze środowiskiem akademickim i sektorem prywatnym

Wnioskodawcami mogą być przedstawiciele środowiska akademickiego (magistrowie, doktoranci) oraz start-upy i przedsiębiorstwa (sektor GovTech) **współpracujące** z administracją publiczną.

Dlaczego korzystać z BDTI?



Skorzystaj z możliwości nieodpłatnego użytkowania przez 6 miesięcy, w tym doradztwa i wsparcia technicznego



Eksperymentuj z analizą danych przy użyciu infrastruktury wysokiej wydajności, która pozwala wykorzystać moc chmury elastycznej (elastic cloud)



Dowiedz się, jak przejść z procesu pilotażowego do procesu produkcyjnego



Przetestuj swój pomysł → Zyskaj wartość → Twórz wiedzę

Kto już korzystał z BDTI?



Hiszpańska organizacja samorządowa Generalitat Valenciana potrzebowała rozwiązania, które umożliwiłoby filtrowanie informacji z obszernych artykułów klinicznych pochodzących z różnych źródeł, takich jak PubMed.gov i publikacje dotyczące COVID-19. Chodziło o ekstrakcję wartościowych informacji i wiedzy z olbrzymiego wolumenu danych.



Europejski Sojusz na rzecz Krwi i DG SANTE podjęły wspólne wysiłki na rzecz utworzenia w Europie kompleksowej i otwartej platformy. Celem było zebranie kluczowych danych na potrzeby ukierunkowanych badań nad terapią osoczem ozdrowieńców COVID-19. Zadanie polegało na określeniu najskuteczniejszych warunków dla tej metody leczenia, wykorzystaniu analizy danych w procesie decyzyjnym i usprawnieniu wysiłków badawczych.



Gmina stanęła przed złożonym zadaniem korelowania zróżnicowanych zbiorów danych w obrębie miasta. Chodziło o zrozumienie schematów przemieszczania się mieszkańców w czasie z jednej dzielnicy do drugiej. Miało to zasadnicze znaczenie dla zrozumienia tendencji w dziedzinie mobilności i odpowiedniego przeprojektowania usług, by móc skuteczniej promować działania i wydarzenia kulturalne.



Zaawansowane narzędzia wizualizacji danych i eksploracji tekstu pomagające w **wyodrębnianiu wiedzy** zawartej w dokumentach wspieranie klinicystów i menedżerów w praktyce klinicznej i codziennej pracy.



Gotowe do użycia środowisko wirtualne, w którym **dane gromadzone za pośrednictwem specjalnie utworzonej strony internetowej**, są importowane i anonimizowane, a następnie analizowane za pomocą zaawansowanych narzędzi wizualizacji i analizy danych. Początkowo przetwarzano jedynie dane dawcy, następnie rozszerzono ich zakres o **całą ścieżkę osocza od dawcy do pacjenta/próby klinicznej**.



Analiza predykcyjna, deskryptywna i analiza szeregów czasowych wykonana na wielu zbiorach danych zgromadzonych **przed pandemią COVID-19, w jej trakcie i po jej zakończeniu**, takich jak: publiczne czujniki Wi-Fi, dane parkowania i dane georeferencyjne dotyczące przemieszczania się osób (tj. turystów).



Jak stosować: szybki i prosty proces



Poznaj usługę BDTI na naszej [stronie internetowej](#)



» Określ swój przypadek użycia analizy danych dzięki narzędziu [BDTI Canva](#), wypełnij formularz wniosku (zob. [strona internetowa](#))



» Wyślij wniosek pilotażowy (wzór) pocztą elektroniczną

EC-BDTI-PILOTS@ec.europa.eu



» Spotkaj się z nami, żeby omówić Twój przypadek użycia



» Projekt pilotażowy zatwierdza się, jeśli: przynosi wartość, jest wykonalny w 6 miesięcy, są wystarczające zasoby (umiejętności, zespół, dane)

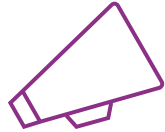


» Twoje środowisko testowe jest skonfigurowane



» Możesz rozpocząć testy pilotażowe i tworzyć wartość!

Dziękujemy za uwagę!



Pracujesz dla organu administracji publicznej, który potrzebuje infrastruktury do analizy danych? Skontaktuj się z nami



EC-BDTI-PILOTS@ec.europa.eu

[Strona internetowa BDTI](#)



[Strona Joinup BDTI](#)



[Biuletyn BDTI](#)



Linki i zasoby

- <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>