

Podpora za podatkovno vodeni javni sektor

*Od ideje do izvedbe z uporabo
Big Data Test Infrastructure (BDTI)*

EC-BDTI-PILOTS@ec.europa.eu



GD CNECT

Generalni direktorat za komunikacijska omrežja, vsebine in tehnologijo

DG DIGIT

Generalni direktorat za informatiko



Informacije javnega sektorja in vloga podatkovne analitike (data analytics)

Podatki so **povsod**, njihova količina pa narašča kot še nikoli.

– **Velepodatki (big data)**: 3V – Velikost, Variantnost, Velika hitrost



Podatki so ključni sestavni del **storitev, proizvodov in učinkovitega oblikovanja politik.**

Cilj je vzpostaviti **enotni evropski trg za podatke** in dati na voljo več podatkov prek močnih in zaupanja vrednih infrastruktur in tehnologij, **v skladu z vrednotami in predpisi EU**, da bi podprli državljane, javni sektor in podjetja.



Kaj je Big Data Test Infrastructure (BDTI)?



BDTI: ne le za velepodatke (big data), temveč za vse **informacije javnega sektorja**



Šest mesecev brezplačne* storitve
za javne uprave EU



Nabor storitev za analizo
podatkov, pripravljenih za
uporabo, in podpora



Platforma v oblaku (cloud),
ki temelji na
odprtokodnih (open-source) orodjih



Za pomoč javnemu sektorju pri črpanju spoznanj iz njegovih podatkov in
pospešitev prehoda na **odločanje, ki temelji na podatkih**

* Stroški pilotnega projekta morajo biti združljivi z omejitvami financiranja pilotnega proračuna BDTI.

Cilji Big Data Test Infrastructure

Cilji

- povečati dostopnost, interoperabilnost, kakovost in uporabnost informacij javnega sektorja v skladu z zahtevami **direktive o odprtih podatkih (open data)**
- spodbujati **ponovno uporabo in združevanje odprtih javnih podatkov** po vsej EU za razvoj informacijskih proizvodov in storitev, vključno z aplikacijami umetne inteligence
- nabori podatkov velike vrednosti – direktiva o odprtih podatkih (open data)
- preizkusiti souporabo podatkov **med podjetji in državo** (Business-to-Government, B2G) za **javno dobro**
- podporni center za podatkovni prostor (data space): raziščite in preskusite svoje podatke*
- BDTI zagotavlja varno **preskusno okolje za izvedbo eksperimentov z velepodatki (big data)** za stranke podatkovnih prostorov (data space)

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



Več o Big Data Test Infrastructure (BDTI)

BDTI se financira iz **programa Digitalna Evropa**, ki je usmerjen v uvajanje digitalne tehnologije za podjetja, državljane in javne uprave.

Program Digitalna Evropa zagotavlja strateško financiranje na petih ključnih področjih:

visokozmogljivo računalništvo

umetna inteligenca

(oblak, podatki in umetna inteligenca)

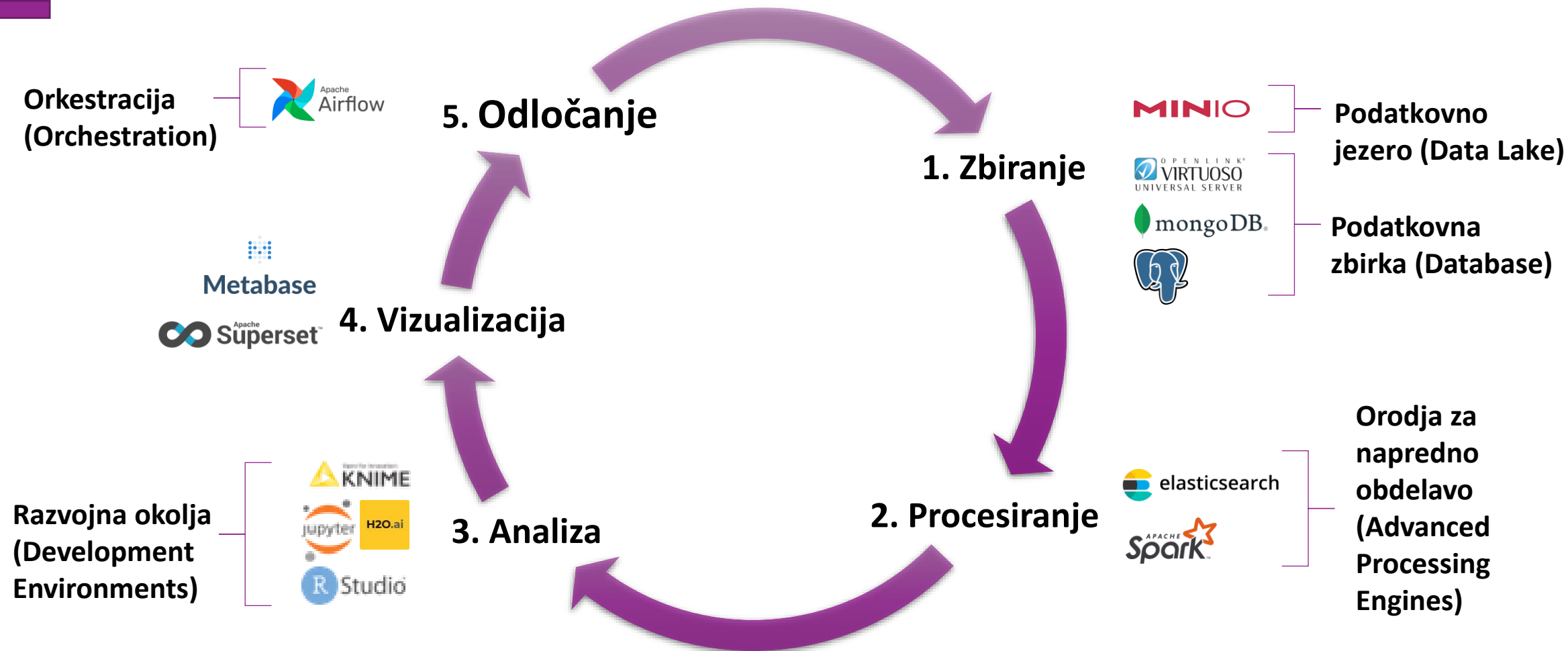
kibernetska varnost

napredne digitalne spretnosti

uvajanje in široka uporaba digitalnih tehnologij



Odprihodna (open-source) orodja za podporo vaši podatkovni poti



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% 
open-source
komponente



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

Komu je namenjena BDTI?



Evropske javne uprave

Vse evropske javne uprave na **lokalni, regionalni in nacionalni ravni** se lahko samostojno prijavijo za pilotni projekt BDTI.



Partnerstva z akademskimi krogi in zasebnim sektorjem

Akademski krogi (magistrski, doktorski študenti), zagonska podjetja (startups) in podjetja (sektor GovTech) se lahko prijavijo, če **sodelujejo** z javno upravo.

Zakaj uporabiti BDTI?



Šestmesečna brezplačna uporaba, vključno s **svetovalno in tehnično** podporo



Eksperimentirajte s podatkovno analitiko z uporabo **visokozmogljive infrastrukture**, ki izkorišča moč elastičnega oblaka (elastic cloud)



Prejmite smernice za prehod s pilotnega postopka na postopek, **pripravljen za uporabo**



Preskusite svojo idejo → Izluščite vrednost → Ustvarite znanje

Kdo je že uporabil BDTI?



Generalitat Valenciana je potrebovala rešitev za pregled znanstvenih kliničnih člankov iz različnih virov, kot so PubMed.gov in publikacije, povezane s COVID-19. Njihov cilj je bil izluščiti dragocena spoznanja in znanja iz te ogromne količine podatkov.



Napredna **vizualizacija podatkov in orodja za besedilno rudarjenje** za pomoč pri **luščenju znanja** iz dokumentov, s čimer podpira klinične zdravnike in menedžerje pri njihovih kliničnih praksah in vsakodnevnem delu.



Evropska zveza za kri in GD za zdravje in varnost hrane sta si skupaj prizadevala za vzpostavitev celovite platforme z odprtim dostopom v Evropi. Namen te pobude je bil zbrati ključne podatke za usmerjeno študijo o zdravljenju s konvalescentno plazmo COVID-19. Izziv je bil opredeliti najučinkovitejše pogoje za to zdravljenje z uporabo analize podatkov za sprejemanje informiranih odločitev in racionalizacijo raziskovalnih prizadevanj.



Virtualno okolje, pripravljeno za uporabo, v katerega se **podatki, zbrani prek prilagojene spletne strani**, vnesejo in se v njem anonimizirajo, nato pa analizirajo z naprednimi orodji za vizualizacijo podatkov in analitičnimi orodji. Sprva so bili obdelani samo podatki o darovanju, nato pa se je področje uporabe povečalo na **celotno pot krvne plazme od darovanja do bolnikov/kliničnega preskušanja**.



Občina se je soočala z zapleteno nalogo vzporejanja različnih naborov podatkov v mestu. Njen cilj je bil razumeti vzorce gibanja ljudi v različnih predelih mesta skozi čas. Ta izziv analize podatkov je bil ključen za pridobitev vpogleda v trende mobilnosti in preoblikovanje storitev za učinkovito spodbujanje kulturnih dejavnosti in dogodkov.



Napovedna in opisna analiza ter analiza časovnih vrst za več naborov podatkov, zbranih **pred in med pandemijo COVID-19 ter po njej**, kot so: senzorji gibanja na osnovi javnih brezžičnih omrežij, podatki o parkiranju in georeferencirani podatki o gibanju ljudi (tj. turistov).

Kako se prijavite: hiter in enostaven postopek



Seznajte se s storitvijo BDTI na našem [spletnem mestu](#)



» Opredelite svoj primer uporabe podatkovne analitike z uporabo [BDTI Canva](#) in nato izpolnite predlogo obrazca zahtevka (glejte [spletno mesto](#))



» Pošljite zahtevek (predlogo) za pilotni projekt po elektronski pošti EC-BDTI-PILOTS@ec.europa.eu



» Srečajte se z nami in podrobneje pojasnite svoj primer uporabe



» Pilotni projekt bo odobren, če:
prinaša vrednost, je izvedljiv v šestih mesecih in so na voljo zadostna sredstva (spretnosti, ekipa, podatki)

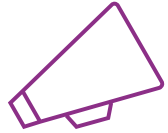


» Vaše testno okolje je pripravljeno



» Začnete s pilotnim projektom in ustvarite vrednost!

Zahvaljujemo se vam za pozornost!



Ali delate za javno upravo, ki potrebuje infrastrukturo za podatkovno analitiko? Stopite v stik z nami



EC-BDTI-PILOTS@ec.europa.eu

[Spletno mesto
BDTI](#)



[Spletna stran za
pristop k BDTI](#)



[Bilten BDTI](#)



Povezave in viri

- <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>