

Podpora verejného sektora založeného na dátach

*Od slov k činom pomocou
Big Data Test Infrastructure (BDTI)*

EC-BDTI-PILOTS@ec.europa.eu



GR CNECT

Generálne riaditeľstvo pre
komunikačné siete, obsah
a technológie

GR DIGIT

Generálne
riaditeľstvo pre
informatiku



Informácie verejného sektora a úloha analýzy dát (Data Analytics)

Dáta sú **všade** a pribúdajú ohromným tempom.

– **Veľké dáta (Big Data)**: 3V – Volume (objem), Variety (rôznorodosť), Velocity (rýchlosť)



Dáta sú kľúčovou zložkou pre **služby, produkty a efektívnu tvorbu politík**.

Snahou je vytvoriť **jednotný európsky trh dát** a sprístupniť viac dát prostredníctvom silných a dôveryhodných infraštruktúr a technológií **v súlade** s hodnotami a regulačnými ustanoveniami EÚ pre potreby občanov, verejného sektora a podnikov.

Čo je to Big Data Test Infrastructure (BDTI)?



BDTI: netýka sa len veľkých dát (Big Data), ale všetkých informácií verejného sektora



Šesť mesiacov bezplatnej* služby
pre orgány verejnej správy EÚ



Nástroje dátovej analýzy a
podpora **prípravené na okamžité**
využitie



Cloudová (Cloud) platforma
založená na nástrojoch s **otvoreným**
zdrojovým kódom (Open-source)



Pomôcť verejnému sektoru **získať poznatky** zo svojich dát a urýchliť
prechod na **rozhodovanie založené na dátach**

* Náklady na pilotný projekt musia byť v súlade s finančným limitom pilotného rozpočtu BDTI.

Ciele Big Data Test Infrastructure

Ciele

- Zvýšiť ľahkú dostupnosť, interoperabilitu, kvalitu a použiteľnosť informácií verejného sektora v súlade s požiadavkou **smernice o otvorených dátach (Open Data)**
- Podporiť **opakované použitie a kombinovanie otvoreného prístupu k dátam verejnej správy** v celej EÚ na vývoj informačných produktov a služieb vrátane aplikácií umelej inteligencie
- Súbory dát s vysokou hodnotou – **smernica o otvorených dátach (Open Data)**
- Testovanie spolupráce v oblasti zdieľania dát medzi **podnikmi a verejným sektorom (Business-to-Government) (B2G)** v záujme **verejného blaha**
- Centrum podpory dátových priestorov (Data Space): **skúmajte svoje dáta a experimentujte s nimi***
- BDTI poskytuje bezpečné **testovacie prostredie na vykonávanie experimentov s veľkými dátami (Big Data)** pre zákazníkov dátového priestoru (Data Space)



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

O Big Data Test Infrastructure (BDTI)

BDTI sa financuje z programu Digitálna Európa (DEP), ktorý je zameraný na poskytovanie digitálnych technológií podnikom, občanom a orgánom verejnej správy.

Program Digitálna Európa poskytuje strategické financovanie v piatich kľúčových oblastiach:

**Vysokovýkonná výpočtová
technika**

Umelá inteligencia
(cloud, dáta a AI)

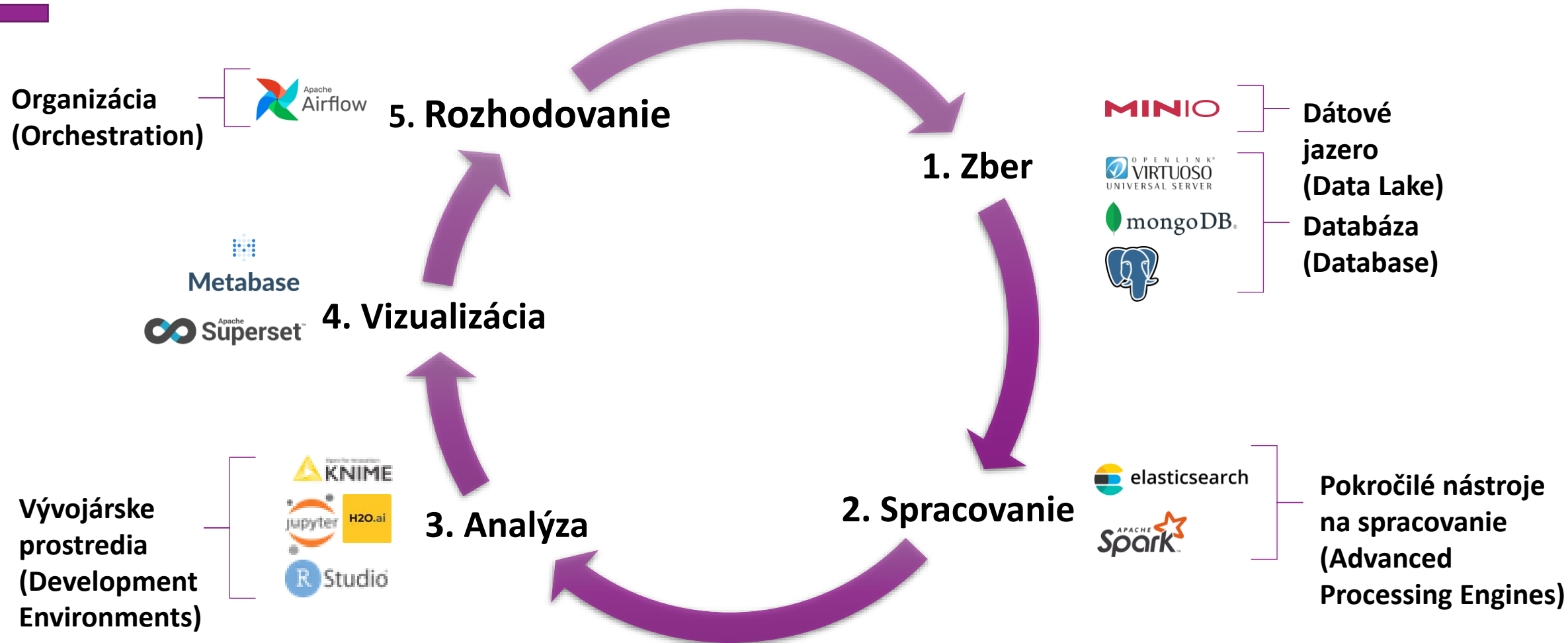
Kybernetická bezpečnosť

Pokročilé digitálne zručnosti

Zavádzanie a rozsiahle využívanie digitálnych technológií



Nástroje s otvoreným zdrojovým kódom (open-source) ako pomoc pri práci s dátami



Portál 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

Main page | Additional information

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	94.4k	2,054k
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

Portál BDTI

100 % 
komponenty s
otvoreným zdrojovým
kódem (Open-source)



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

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](#)

Komu má BDTI slúžiť?



Európskym orgánom verejnej správy

Všetky európske orgány verejnej správy na **miestnej, regionálnej a celoštátnej úrovni** môžu samostatne požiadať o pilotný projekt BDTI



Partnerstvám s akademickou obcou a súkromným sektorom

Akademická obec (študenti magisterského a doktorandského štúdia) a startupy a firmy (sektor GovTech) sa môžu prihlásiť, ak **spolupracujú** s orgánmi verejnej správy

Prečo používať BDTI?



Využite výhodu **šiestich mesiacov zdarma** vrátane **poradenskej a technickej** podpory



Experimentujte s analýzou dát pomocou **vysokovýkonnej infraštruktúry**, ktorá využíva silu pružného cloudu (Elastic Cloud)



Nechajte si poradiť pri prechode od pilotnej fázy k procesu **pripravenému na produkciu**



Otestujte svoj nápad → Získajte hodnotu → Tvorte poznatky

Kto už BDTI využil?



Samospráva autonómneho spoločenstva Valencia potrebovala riešenie na prehľadávanie rozsiahlych vedeckých klinických článkov z rôznych zdrojov, ako sú PubMed.gov a publikácie súvisiace s ochorením COVID-19. Snažila sa získať cenné poznatky a vedomosti z tohto obrovského objemu údajov.



Európska aliancia pre krv a GR SANTE začali spolupracovať na vytvorení komplexnej platformy s otvoreným prístupom v Európe. Cieľom tejto iniciatívy bolo zhromaždiť kľúčové údaje pre cielenú štúdiu o terapeutickom využití plazmy osôb, ktoré sa zotavili po ochorení COVID-19. Úloha spočívala v zistení najúčinnějších podmienok pre túto liečbu s využitím analýzy dát na prijímanie informovaných rozhodnutí a zefektívnenie výskumného úsilia.



Samospráva čelila zložitej úlohe prepojiť rôzne súbory dát v rámci mesta. Jej cieľom bolo pochopiť vzorce pohybu ľudí v rôznych mestských častiach v priebehu času. Táto výzva v oblasti analýzy dát bola nevyhnutná na získanie poznatkov o vývoji mobility a na prepracovanie služieb na účinnú propagáciu kultúrnych aktivít a podujatí.



Pokročilé nástroje na **vizualizáciu dát a hĺbkovú analýzu textu** s cieľom pomôcť **získať poznatky** obsiahnuté v dokumentoch, podpora klinických lekárov a manažérov v ich klinickej praxi a každodennej práci.



Virtuálne prostredie pripravené na okamžité využitie, v ktorom sa **dáta zozbierané prostredníctvom webovej stránky vytvorenej na mieru** spracovávajú a anonymizujú, aby sa potom mohli analyzovať pomocou pokročilej vizualizácie dát a analytických nástrojov. Pôvodne sa spracúvali len dáta o darcovstve, potom sa rozsah rozšíril s cieľom zachytiť **celú dráhu krvnej plazmy, a to od darovania až po pacienta/klinické skúšanie**.



Analýza, ktorá je prediktívna, opisná a týka sa časových radov viacerých dátových súborov zhromaždených **pred pandémiou COVID-19, počas nej a po nej**, ako sú: verejné Wi-Fi snímače, parkovacie a georeferenčné dáta o pohybe osôb (t. j. turisti).



Ako sa prihlásiť: rýchly a jednoduchý proces



Oboznámte sa so službou BDTI na našej [webovej stránke](#)



» Definujte svoj prípad použitia analýzy dát pomocou nášho [BDTI Canva](#) a potom vyplňte vzorový formulár žiadosti (pozri [webovú stránku](#))



» Predložte svoju pilotnú žiadosť (vzor) e-mailom EC-BDTI-PILOTS@ec.europa.eu



» Spojte sa s nami a opíšte svoj prípad použitia



» Pilotný projekt je schválený, ak: prináša hodnotu, dá sa vyhotoviť do šiestich mesiacov, sú k dispozícii dostatočné zdroje (zručnosti, tím, dáta)

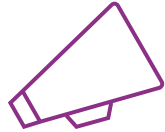


» Vaše testovacie prostredie je nastavené



» Môžete začať pilotný projekt a vytvárať hodnotu!

Ďakujeme za pozornosť!



Pracujete pre orgán verejnej správy, ktorý potrebuje infraštruktúru na analýzu dát? Obráťte sa na nás



EC-BDTI-PILOTS@ec.europa.eu

[Webová stránka BDTI](#)



[Stránka Joinup BDTI](#)



[Newsletter BDTI](#)



Odkazy a zdroje

- https://big-data-test-infrastructure.ec.europa.eu/index_en?etrans=sk
- <https://code.europa.eu/bdti/bdti-demonstrator>
- https://commission.europa.eu/publications/interoperable-europe-act-proposal_sk
- https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_sk
- <https://digital-strategy.ec.europa.eu/sk/policies/legislation-open-data>
- <https://digital-strategy.ec.europa.eu/sk/policies/data-governance-act>
- <https://digital-strategy.ec.europa.eu/sk/policies/european-approach-artificial-intelligence>
- <https://digital-strategy.ec.europa.eu/sk/activities/digital-programme>
- <https://dssc.eu/wp-content/uploads/2023/03/DSSC-Data-Spaces-Glossary-v1.0.pdf>
- <https://digital-strategy.ec.europa.eu/sk/library/staff-working-document-data-spaces>
- https://ec.europa.eu/commission/presscorner/detail/sk/ip_22_1113
- <https://eur-lex.europa.eu/legal-content/SK/TXT/?uri=celex%3A32019L1024>
- <https://joinup.ec.europa.eu/collection/egovernment/solution/big-data-test-infrastructure-bdti>