

Datapohjaista päätöksentekoa julkisella sektorilla:

Hypestä käytäntöön:

Big Data Test Infrastructure (BDTI)

EC-BDTI-PILOTS@ec.europa.eu



DG CNECT

Viestinnän pääosasto
Viestintäverkot, sisällöt ja teknologia

DG DIGIT

Tietotekniikan
pääosasto



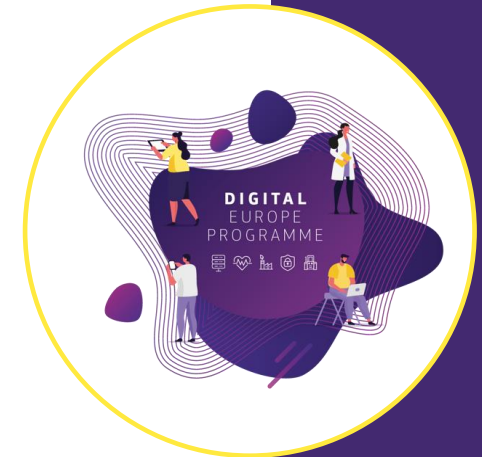
Julkisen sektorin tieto ja data-analytiikka

Dataa on **kaikkialla** ja sen määrä kasvaa ennennäkemättömällä vauhdilla.
- **Massadata**: 3V - Volume, Variety, Velocity



Data on keskeinen tekijä **palveluissa, tuotteissa ja tuloksellisessa päätöksenteossa.**

EU:n tavoitteena on luoda **eurooppalaiset datan sisämarkkinat** ja lisätä datan saatavuutta tuottamalla suorituskykyisiä ja luotettavia infrastruktuureja ja teknologioita. Niillä tuetaan **kansalaisia, julkista sektoria ja yrityksiä EU:n arvojen ja sääntöjen mukaisesti.**



Mikä on Big Data Test Infrastructure (BDTI)?



BDTI: Ei vain massadataa vaan kaikkea **julkisen sektorin tietoa** varten



Kuusi kuukautta ilmaiseksi*
EU:n julkishallinnoille



Käyttövalmis data-analytiikkapino
ja tuki



Avoimeen lähdekoodiin perustuva
pilvialusta



Tavoitteena auttaa julkista sektoria **muuntamaan dataansa informaatioksi**
ja nopeuttaa muutosta kohti **datavetoista päätöksentekoa**

* Pilottihankkeen kustannusten on pysyttävä BDTI-pilottibudjetin rahoituksen puitteissa

Big Data Test Infrastructure: Tavoitteet

Tavoitteet

- Parantaa julkisen sektorin tietojen saatavuutta, yhteentoimivuutta, laatua ja hyödynnettävyyttä **avointa dataa koskevan direktiivin mukaisesti**
- Lisätä **avoimen julkisen datan uudelleenkäyttöä ja yhdistämistä** EU:ssa tietotuotteiden ja -palvelujen kehittämistä varten, mukaan lukien tekoälysovellukset
- Arvokkaat tietoaaineistot – Avointa dataa koskeva direktiivi
- Testissä **yleishyödyllinen Business-to-Government (B2G) -tiedonjakoyhteistyö**
- Data Space -tukikeskus: opi tuntemaan datasi paremmin*
- BDTI tarjoaa data-avaruuksien käyttäjille turvallisen **testausympäristön massadatan käytölle**



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

Faktat: Big Data Test Infrastructure (BDTI)

BDTI:tä rahoitetaan **Digitaalinen Eurooppa -ohjelmasta (DEP)**, jonka ytimessä on digitaalitekniologioiden tuominen yritysten, kansalaisten ja julkishallintojen käyttöön.

DEP tarjoaa strategista rahoitusta viidellä keskeisellä alalla:

Suurteholaskenta

Kyberturvallisuus

Tekoäly

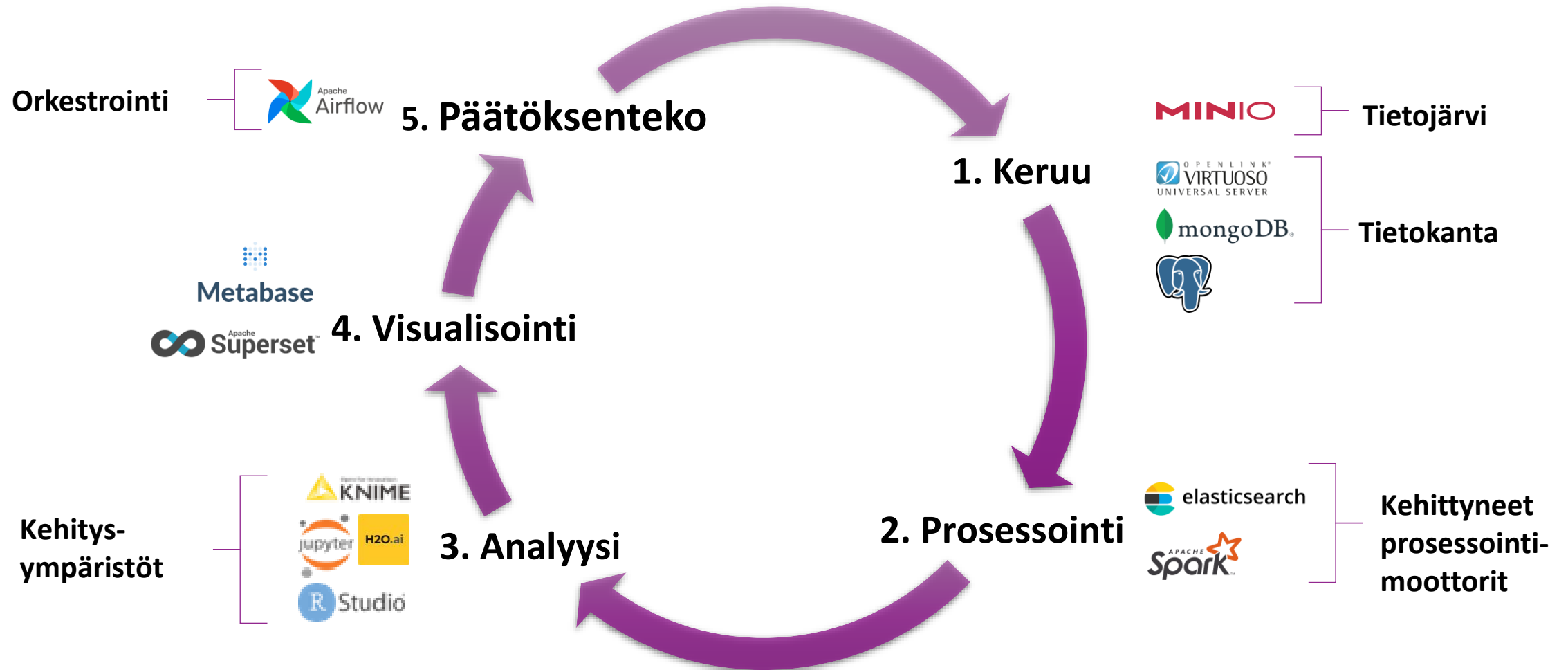
(pilvipalvelut, data ja tekoäly)

Kehittyneet digitaidot

Digitaalisten teknologioiden käyttöönotto ja laaja käyttö



Avoimen lähdekoodin työkalut datamatkaeväiksi



BDTI-portaali



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

BDTI-portaali

100% 
open-source
components



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](#)

Kenelle BDTI on tarkoitettu?



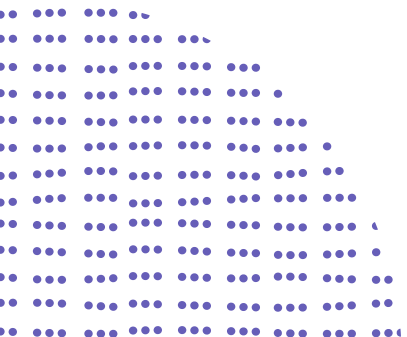
Euroopan julkishallinnot

Kaikki Euroopan julkishallinnot **paikallisella, alueellisella ja valtakunnallisella tasolla** voivat hakea itsenäisesti BDTI-pilottihankkeeseen



Korkeakoulujen ja yksityisen sektorin kumppanuudet

Korkeakoulut (maisteriopiskelijat ja tohtorikoulutettavat) sekä startupit ja yritykset (GovTech-sektori) voivat hakea, jos ne tekevät **yhteistyötä** julkishallinnon kanssa



Miksi käyttää BDTI:tä?



Kuusi kuukautta ilmaiseksi, mukaan lukien **neuvonta ja tekninen tuki**



Data-analytiikan testaaminen suurtehoisessa infrastruktuurissa, joka hyödyntää **elastisen pilven** tehoa



Opastus siirtymisessä pilottivaiheesta **tuotantovalmiiseen** prosessiin



Testaa ideasi → Uuta arvoa → Luo tietämystä

Esimerkkejä käytöstä



Generalitat Valencianan piti käydä läpi mittava määrä tieteellisiä artikkeleita eri lähteistä, kuten PubMed.gov ja covid-19-aiheiset julkaisut. Haasteena oli louhia valtavasta datamäärästä merkityksellistä tietämystä.



Kehittyneet **datan visualisointi- ja tekstinlouhintavälineet** avuksi asiakirjojen sisältämän **tietämyksen uuttamiseen**, tavoitteena tukea lääkäreiden ja vastaanottojen esihenkilöiden jokapäiväistä työtä



European Blood Alliance ja DG SANTE perustivat yhdessä laajan avoimen alustan Eurooppaan. Projektissa haluttiin kerätä kriittistä tietoa covid-19-toipilasplasmaterapiaa koskevaa selvitystä varten. Tehtävänä oli määrittää hoitomuodon suotuisimmat edellytykset data-analyysillä, joka mahdollistaa tietoon perustuvat päätökset ja sujuvoittaa tutkimustyötä.



Käyttövalmis virtuaaliympäristö, jossa **räätälöidyn verkkosivuston kautta kerättyä dataa** otetaan talteen ja anonymisoidaan analysoitavaksi kehittyneillä datan visualisointi- ja analysointivälineillä. Ensin kerättiin vain verenluovutusdataa, mutta kohdealaa laajennettiin **veriplasmaan, verenluovutuksesta klinisiin kokeisiin asti**.



Firenzen kaupungin haasteellisena tehtävänä oli korreloida kaupungin erilaisia tietoaineistoja. Tavoitteena oli ymmärtää ihmisten liikkumistapoja eri alueilla ajan mittaan. Tämä data-analyysi oli olennainen edellytys liikkumistrendien selvittämiseksi ja palvelujen parantamiselle kulttuuritoimintojen ja -tapahtumien edistämiseksi.



Ennustava, kuvaileva ja aikasarjapohjainen analyysi tietoaineistoista, jotka kerättiin **ennen covid-19-pandemiaa, sen aikana ja jälkeen**, kuten: yleiset wifi-anturit, pysäköinti- ja georeferoitu data ihmisten (lähinnä turistien) liikkeistä.



Nopea ja helppo hakuprosessi



Tutustu BDTI-palveluun [verkkosivullamme](#)



» Määrittele data-analytiikan käyttötapaaksesi [BDTI Canva](#)-välineellämme ja täytä hakulomake (ks. [verkkosivu](#))



» Lähetä pilottipyynnösi (lomake) sähköpostitse EC-BDTI-PILOTS@ec.europa.eu



» Esittele meille käyttötapaustasi



» Pilottihanke hyväksytään, jos se:

tuottaa arvoa; voidaan tehdä 6 kuukaudessa; omaa riittävät resurssit (osaaminen, tiimi, data)

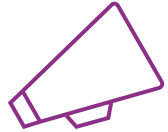


» Testiympäristösi on valmis



» Voit aloittaa pilotin ja luoda arvoa!

Kiitos mielenkiinnostasi!



Työskenteletkö julkishallinnossa, joka tarvitsee infrastruktuurin data-analytiikkaa varten? Ota yhteyttä!



EC-BDTI-PILOTS@ec.europa.eu

[BDTI:n verkkosivusto](#)



[BDTI:n Joinup-sivu](#)



[BDTI:n uutiskirje](#)



Linkkejä

- <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>