

# Ermöglichung eines datengestützten öffentlichen Sektors

*Vom Hype zum Handeln mithilfe der Big Data Test Infrastructure (BDTI)*

[EC-BDTI-PILOTS@ec.europa.eu](mailto:EC-BDTI-PILOTS@ec.europa.eu)



**GD CNECT**  
Generaldirektion Kommunikationsnetze,  
Inhalte und Technologien

**GD DIGIT**  
Generaldirektion  
Informatik



# Informationen des öffentlichen Sektors und die Rolle der Datenanalyse

Daten sind **überall** und die Menge an Daten wächst so schnell wie nie zuvor.

- **Big Data**: 3V - Volume, Variety, Velocity (Volumen, Vielfalt, Schnelligkeit)



Daten sind ein wesentlicher Bestandteil von **Dienstleistungen, Produkten und einer wirksamen Politikgestaltung**.

Das Ziel ist es, einen **europäischen Binnenmarkt für Daten** zu schaffen und mehr Daten durch leistungsstarke und vertrauenswürdige Infrastrukturen und Technologien **im Einklang mit den Werten und Vorschriften der EU** zur Verfügung zu stellen, **um die Bürger, den öffentlichen Sektor und Unternehmen zu unterstützen**.



# Was ist die Big Data Test Infrastructure (BDTI)?



BDTI: Nicht nur für Big Data, sondern für alle **Informationen des öffentlichen Sektors**



Sechs Monate kostenfreier Dienst\*  
für die öffentlichen Verwaltungen  
der EU



Gebrauchsfertiger Datenanalyse-  
Stack und -Support



Cloud-Plattform auf der Grundlage  
von **Open-Source-Tools**



Unterstützung des öffentlichen Sektors bei der **Gewinnung von Erkenntnissen aus seinen Daten** und Beschleunigung des Übergangs zu einer **datengesteuerten Entscheidungsfindung**

\* Die Kosten des Pilotprojekts müssen innerhalb der Fördergrenzen des BDTI-Pilotbudgets liegen.

# Ziele der Big Data Test Infrastructure

## Ziele

- Verbesserung der Zugänglichkeit, Interoperabilität, Qualität und Nutzbarkeit von Informationen des öffentlichen Sektors im Einklang mit den Anforderungen der **Open-Data-Richtlinie**
- Förderung der **Weiterverwendung und der Kombination offener und öffentlicher Daten** in der gesamten EU, um die Entwicklung von Informationsprodukten und -diensten, einschließlich KI-Anwendungen, zu erleichtern
- **Hochwertige Datensätze** – Open-Data-Richtlinie
- Erprobung der Zusammenarbeit im Bereich **Business-to-Government (B2G)** bei der gemeinsamen Nutzung von Daten zum **Wohl der Allgemeinheit**
- Unterstützungszentrum für Datenräume: erkunden Sie Ihre Daten und experimentieren Sie mit ihnen\*
- Die BDTI bietet eine sichere **Testumgebung**, in der Kunden des Data Space Big-Data-Experimente durchführen können.



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

# Über die Big Data Test Infrastructure (BDTI)

Die BDTI wird aus dem **Programm „Digitales Europa“ (DEP)** finanziert, dessen Schwerpunkt auf der **Bereitstellung digitaler Technologien** für Unternehmen, Bürger und **öffentliche Verwaltungen** liegt.

Im Rahmen des DEP werden strategische Mittel in **fünf wichtigen Bereichen** bereitgestellt:

**High performance computing**

**Cybersicherheit**

**Künstliche Intelligenz**

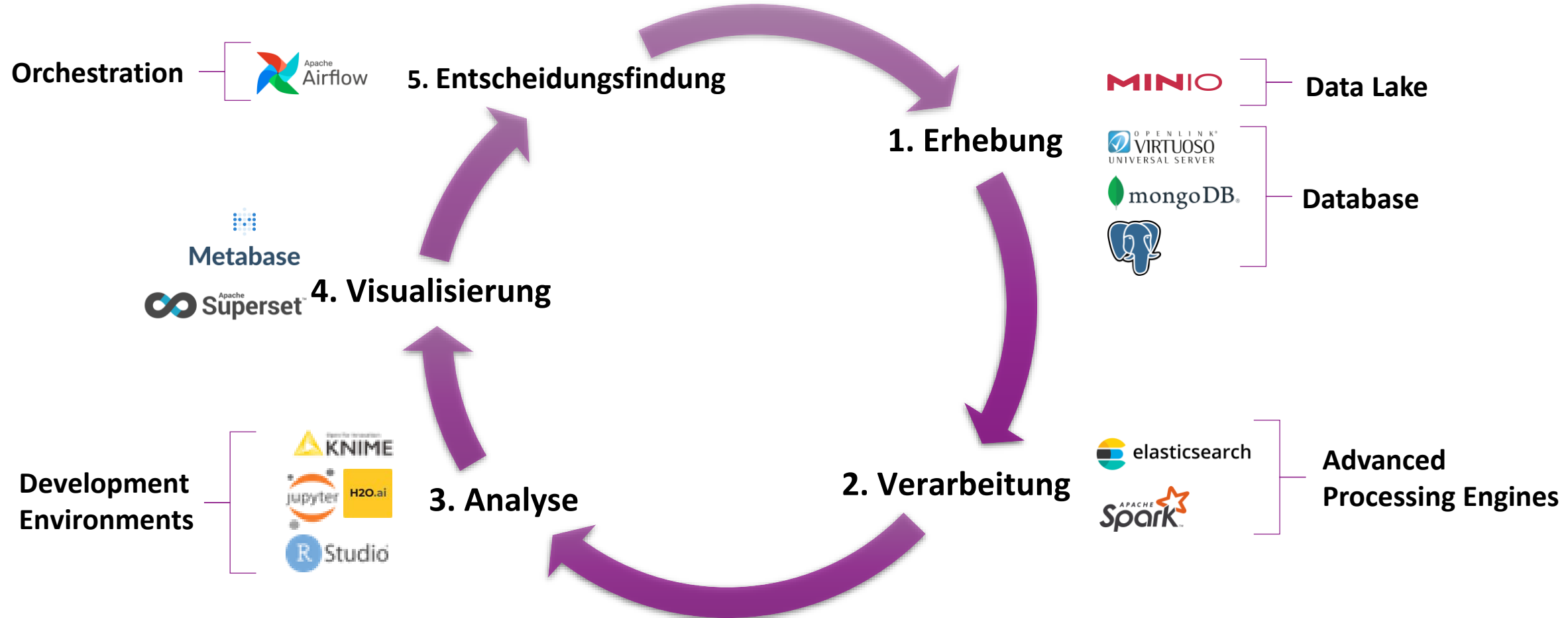
*(Cloud, Daten und KI)*

**Fortgeschrittene digitale  
Kompetenzen**

**Einführung und breite Nutzung digitaler Technologien**



# Open-Source-Tools zur Unterstützung Ihrer Datenreise



# Das BDTI-Portal



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

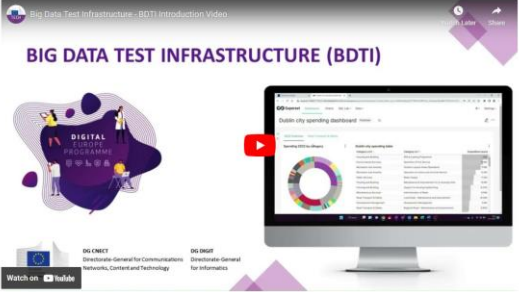
- Home
- My Account
- Service Catalog
- My Services
- My Data

## 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](#)



<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	19.7k	19.7%
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 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	5,894k
2 Low traffic	466.0	3,784k
1 Very low traffic	221.5	1,794k
0 None	0	0

Logout

v0.9.4

# Das BDTI-Portal

100 %   
Open-Source-  
Komponenten



## 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](#)



# Für wen ist die BDTI gedacht?



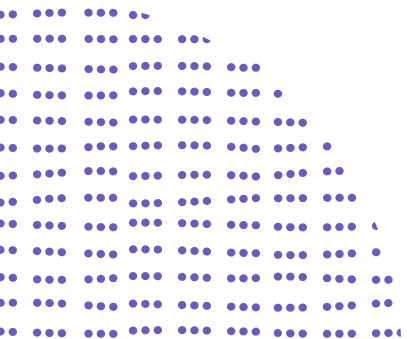
## Europäische öffentliche Verwaltungen

Alle europäischen öffentlichen Verwaltungen auf **lokaler, regionaler und nationaler Ebene** können sich selbstständig für ein BDTI-Pilotprojekt bewerben.



## Partnerschaften mit Hochschulen und der Privatwirtschaft

Hochschulen (Studierende in Master- und PhD-Studiengängen) sowie Startups und Unternehmen (GovTech-Sektor) können sich bewerben, wenn sie mit einer öffentlichen Verwaltung **zusammenarbeiten**.



# Warum die BDTI nutzen?



**Kostenfreie Inanspruchnahme für sechs Monate**, einschließlich **Beratung und technischer Unterstützung**



**Experimentieren mit Datenanalysen** unter Verwendung einer **Hochleistungsinfrastruktur**, die die Leistungsfähigkeit der **Elastic Cloud** nutzt



**Orientierungshilfe für den Übergang** vom Pilotprojekt zu einem **produktionsreifen** Prozess



**Ideen erproben → Wert schöpfen → Wissen schaffen**

# Wer hat die BDTI bereits genutzt?



Die Generalitat Valenciana benötigte eine Lösung, um ein großes Volumen wissenschaftlicher Artikel zu klinischen Studien aus verschiedenen Quellen wie PubMed.gov und Veröffentlichungen im Zusammenhang mit COVID-19 zu sichten. Ihre Herausforderung bestand darin, aus diesem enormen Datenvolumen wertvolle Einblicke und Erkenntnisse zu gewinnen.



Der Europäische Verband der Blutbanken und die GD SANTE haben gemeinsame Anstrengungen unternommen, um eine umfassende frei zugängliche Plattform in Europa zu schaffen. Mit dieser Initiative sollten wichtige Daten für eine gezielte Studie zur COVID-19-Rekonvaleszenzplasma-Therapie gesammelt werden. Die Herausforderung bestand darin, die wirksamsten Bedingungen für diese Behandlung zu ermitteln, indem Datenanalysen genutzt werden, um fundierte Entscheidungen zu treffen und die Forschungsbemühungen zu straffen.



Die Gemeinde stand vor der komplexen Aufgabe, unterschiedliche Datensätze innerhalb der Stadt miteinander in Verbindung zu bringen. Ihr Ziel war es, die Bewegungsmuster der Menschen in verschiedenen Bezirken im Laufe der Zeit zu verstehen. Eine Lösung für diese Problematik zu finden, war von entscheidender Bedeutung, um Einblicke in Mobilitätstrends zu gewinnen und Dienstleistungen neu zu gestalten, mit dem Ziel, kulturelle Aktivitäten und Veranstaltungen wirksam zu fördern.



Fortschrittliche **Datenvisualisierungs- und Textmining-Tools**, durch welche das in den Dokumenten enthaltene **Wissen extrahiert wird**, um so Klinikärzte und Führungskräfte in ihrer klinischen Praxis und bei ihrer täglichen Arbeit zu unterstützen.



Eine gebrauchsfertige, virtuelle Umgebung, in der die **über eine benutzerdefinierte Webseite gesammelten Daten** aufgenommen und anonymisiert werden, um dann mit fortschrittlichen Datenvisualisierungs- und Analysetools analysiert zu werden. Zunächst wurden nur Spenderdaten verarbeitet, dann wurde der Umfang erweitert, um den **gesamten Weg des Blutplasmas von der Spende bis zum Patienten/zur klinischen Studie** zu erfassen.



Prädiktive, deskriptive und Zeitreihenanalysen zu verschiedenen Datensätzen, die **vor, während und nach der Covid-19-Pandemie** erhoben wurden, wie z. B: Daten von Sensoren des öffentlichen lokalen Funknetzes, zu Parkplätzen und georeferenzierte Daten zu Personenbewegungen (d. h. Touristen).



# Bewerbung: ein schnelles und einfaches Verfahren



Machen Sie sich auf unserer [Website](#) mit dem BDTI-Dienst vertraut.



» Definieren Sie den Anwendungsfall für die Datenanalyse mithilfe unseres [BDTI Canva](#) und füllen Sie dann das Antragsformular aus (siehe [Website](#)).



» Senden Sie den Antrag für Ihr Pilotprojekt (Vorlage) per E-Mail

[EC-BDTI-PILOTS@ec.europa.eu](mailto:EC-BDTI-PILOTS@ec.europa.eu)



» Erläutern Sie uns Ihren Anwendungsfall im Rahmen eines persönlichen Treffens.



» Das Pilotprojekt wird genehmigt, wenn es einen Mehrwert schafft, es innerhalb von 6 Monaten durchgeführt werden kann, und ausreichende Ressourcen verfügbar sind (Kompetenzen, Team, Daten).



» Ihre Testumgebung ist eingerichtet.



» Sie können mit dem Pilotprojekt beginnen und Mehrwert schaffen!

# Vielen Dank für Ihre Aufmerksamkeit!



Sind Sie für eine öffentliche Verwaltung tätig, die eine Infrastruktur für die Datenanalyse benötigt? Nehmen Sie Kontakt zu uns auf.



[EC-BDTI-PILOTS@ec.europa.eu](mailto:EC-BDTI-PILOTS@ec.europa.eu)

[BDTI-Website](#)



[Joinup-Seite der BDTI](#)



[BDTI-Newsletter](#)



# Links und Ressourcen

- <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/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://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://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>