



Connecting Europe Facility

# Experiment with Big Data Test Infrastructure

4<sup>th</sup> June 2019

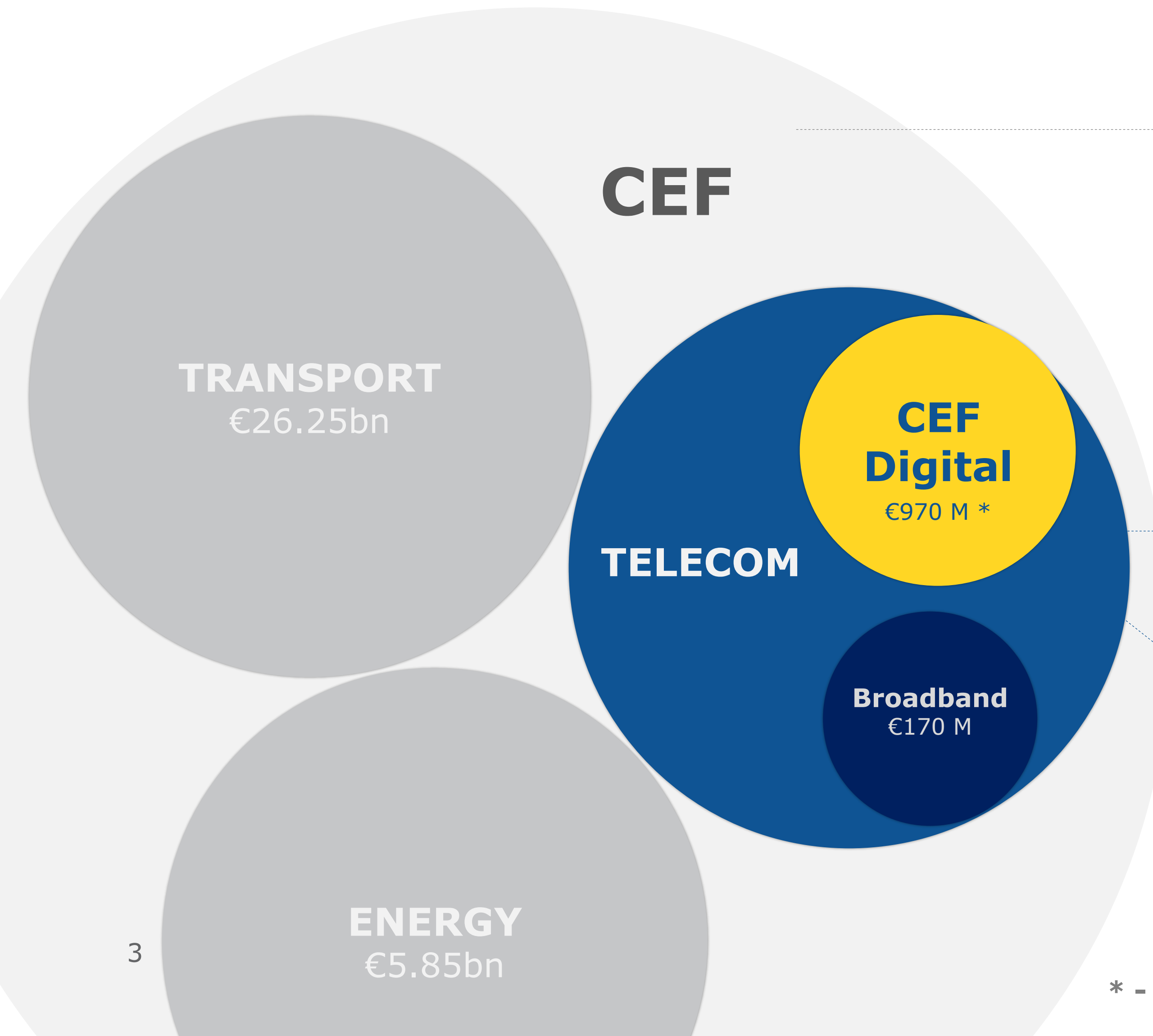


# Agenda

---

1. Presentation of the CEF BDTI building block
2. BDTI Service Offering
3. Application for a BDTI pilot
4. Experiences on BDTI

# The CEF building blocks are funded by the Connecting Europe Facility



## CEF Regulation

Defines how the Commission can finance support for the establishment of trans-European networks to reinforce an interconnected Europe.

## CEF Telecom Guidelines

The CEF Telecom guidelines cover the specific objectives and priorities as well as eligibility criteria for funding of broadband networks and Digital Service Infrastructures (DSIs).

## CEF Work Programmes

Translates the CEF Telecom Guidelines in general objectives and actions planned on a yearly basis.

\* - 100 M Juncker Package

# What is the Big Data Test Infrastructure?

The **Big Data Test Infrastructure** will provide a set of **data and analytics services**, from infrastructure, tools and stakeholder onboarding services, allowing European public organisations to **experiment with Big Data technologies** and move towards **data-driven decision making**





# BDTI – initiative drivers



## **Problem** **Solution**

### **Lack of Big Data technologies**

*Facilitate the prototyping and launching of pilot*

### **Lack of Big Data skills**

*Facilitate Big Data knowledge in public sector*

### **Data sharing among public organisations is not yet a common practice**

*Provide built-in connectors/APIs and foster the sharing of data sources to better support policy-making*

### **Risk of replicating the efforts by implementing similar projects**

*Support public organisations through the creation of a Big Data community for the sharing of good practices, pilot outcomes, etc.*



# BDTI history

---

**2019**

***THE PLATFORM IS UP AND RUNNING!***  
First pilot projects have already started.

**2018**

***BDTI architecture and service design***  
BDTI is now financed by the CEF programme.

**2017**

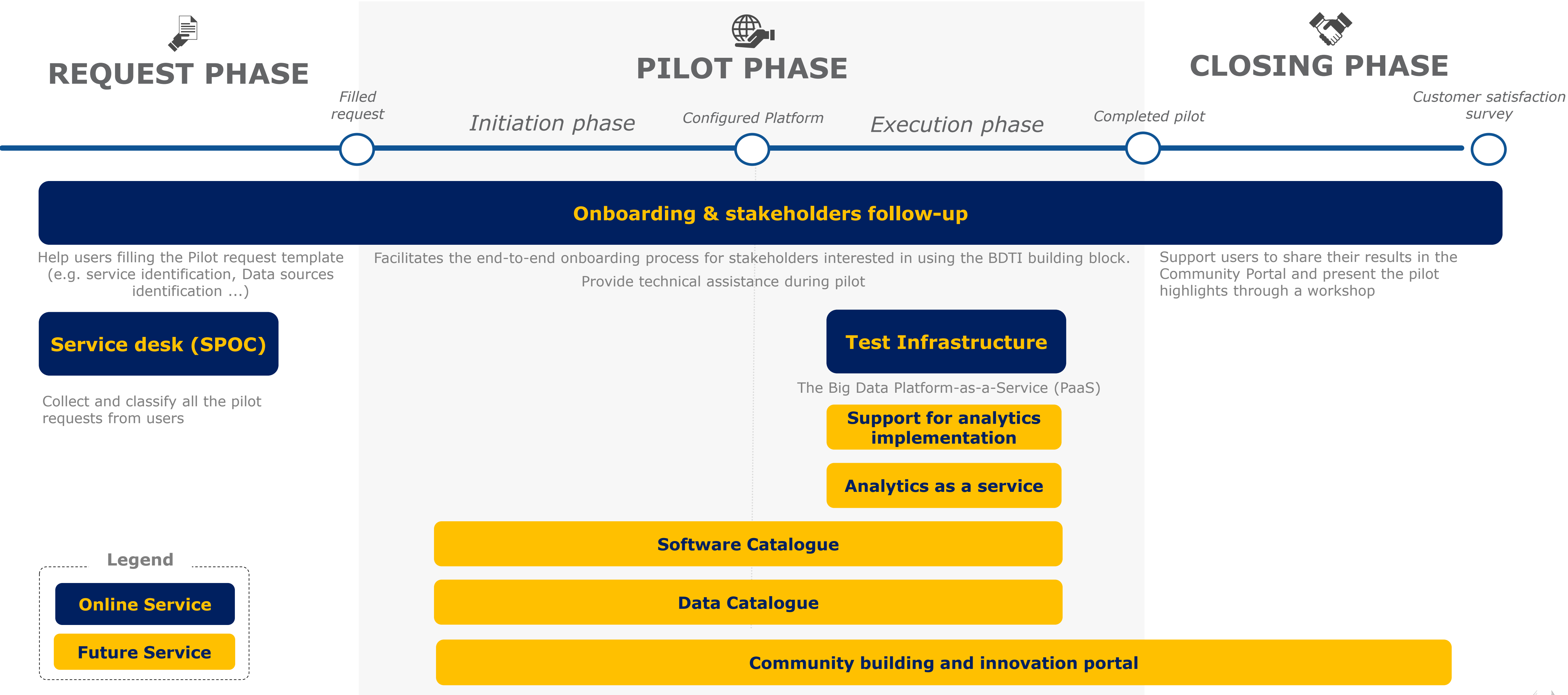
***BDTI requirements definition***  
Collection of business needs from Member States.

**2016**

***Report "Big data analytics for policy making"***  
Big data state of art in European public administrations.



# BDTI service overview



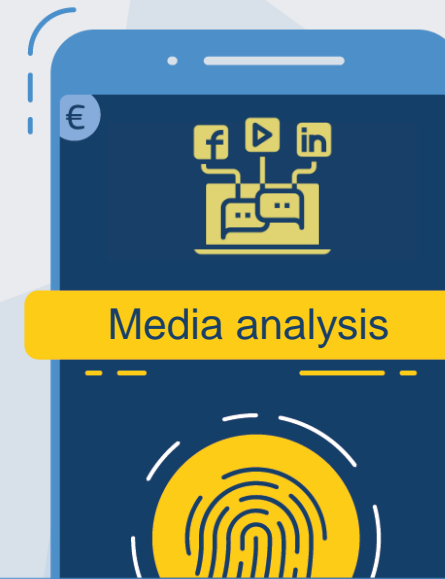
# BDTI applicability



Descriptive analysis

## Descriptive analysis

Use of statistics to quantitatively describe features of a collection of information



Media analysis

## Social Media Analysis

Gather and analyse data from social media to improve business decisions



Time-series analysis

## Time-series Analysis

Analyse time series data in order to extract meaningful statistics and other data characteristics



Predictive analysis

## Predictive analysis

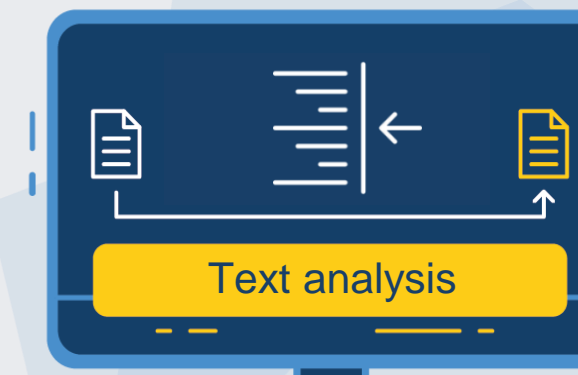
Use statistical techniques that analyse current and historical facts to make predictions about future or unknown events



Network analysis

## Network Analysis

Investigate any structures through the use of network and graph theories



Text analysis

## Text Analysis

Use natural language processing to analyse unstructured text data, to derive pattern and trends



# BDTI future use cases

## WEB ANALYSIS (SCRAPING / MONITORING)

Gather information from websites, involving data scraping (using bot or web-crawler) and data parsing to extract unorganised web data, as well as data from API's, into manageable format.

## IoT & SMART CITY

Gather relevant information on the usage of several interconnected devices (Internet of Things environment) in a Smart City context.

## IoT Security

Safeguard connected devices and networks in the Internet of Things, since security often has not been considered in IoT products design.

## IMAGE PROCESSING

Computational operations using any form of signal processing for which the input is an image, a series of images, or frames of a video; output of image processing may be either an image or a set of characteristics / parameters related to the image

## ROUTE-TRACEABILITY/ FLOW MONITORING

Everything that concerns with tracking and detection of objects through the use of sensors (e.g. GPS, mobile phone signals, road cameras) or any other types of data usable for this purpose.

## APPLYING BIOINFORMATICS TO GENETIC DATA

The use of computational biology, in terms of macromolecules applying "informatics" techniques to understand/organise the information associated to analyse genetic data.

## POPULATION / CUSTOMER SEGMENTATION

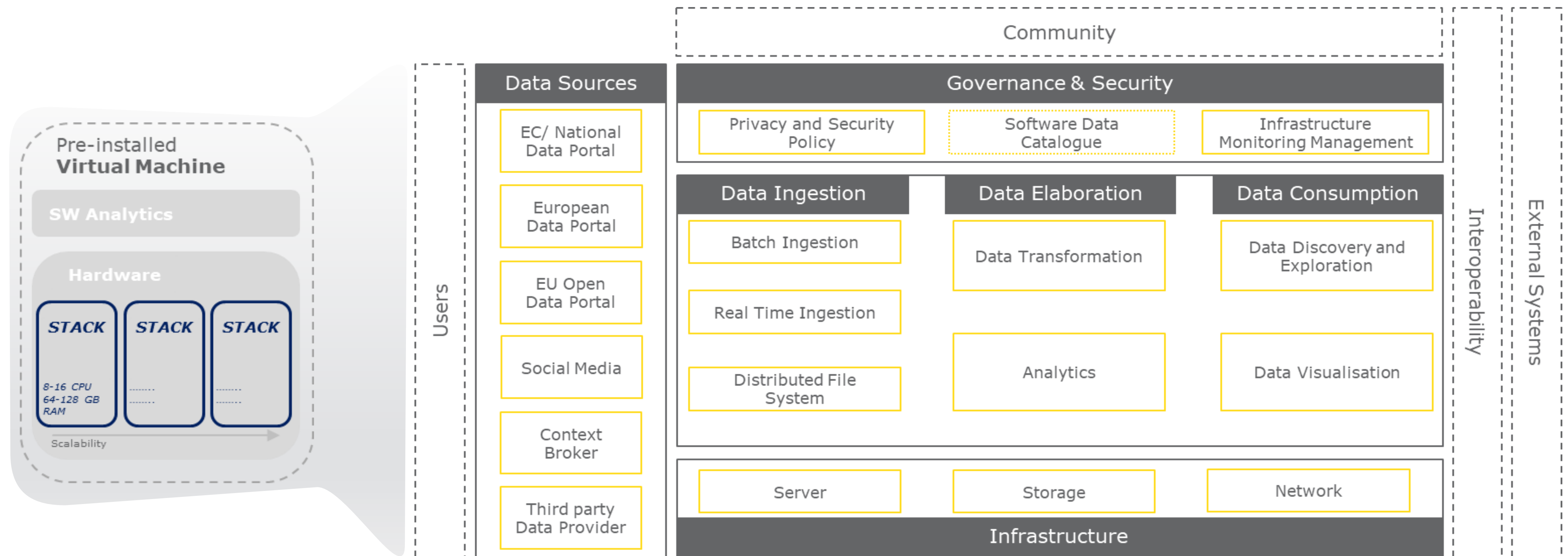
Divide a broad population into sub-groups of consumers based on some types of shared characteristics such as common needs, interests, similar lifestyles or even similar demographic profiles.



# How BDTI works (1/3)

## Technical architecture

The BDTI architecture includes three parts: **the software stack** (i.e., data analytic tools grouped in Data Ingestion, Data Elaboration, Data Consumption and Governance & Security), **the infrastructure** (used through a set of different templates, depending on the pilot) and the different **data sources** to be used by users, currently under analysis.



# How BDTI works (2/3)

Data sources – You can bring your own data or use the data provided by BDTI



## Data Sources

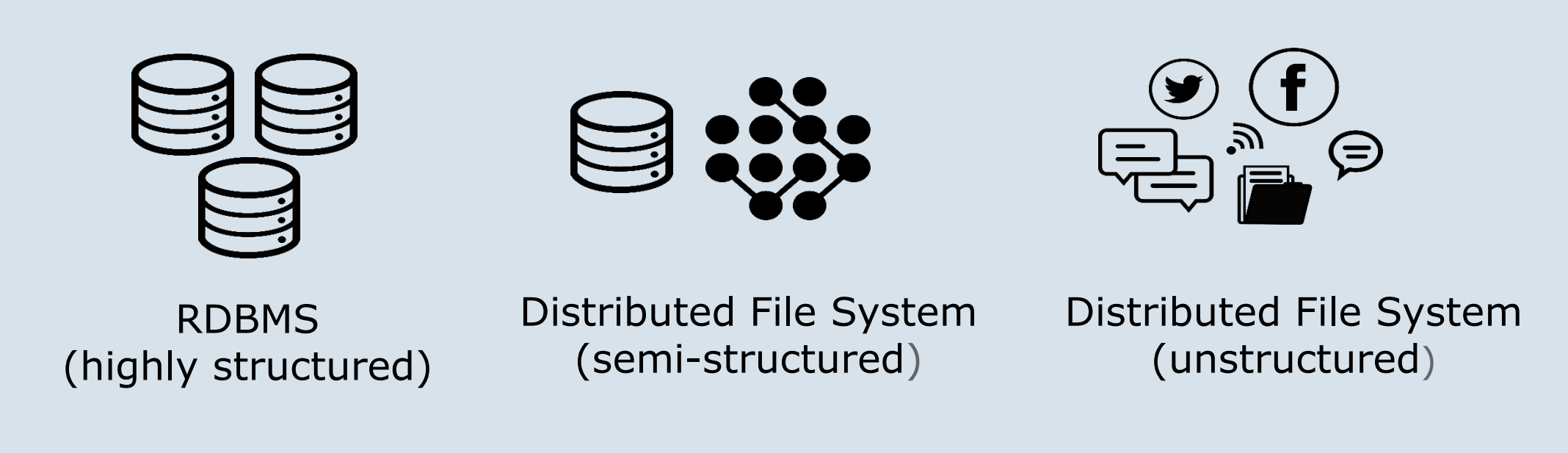
### BDTI (Internal) Data Sources

- EC / National Data Portal
- Social Media
- EU Open Data Portal
- European Data Portal
- Context Broker
- Third party Data Provider

### User (External) Data Sources

Bring your own data  
AND/ OR  
Use open data

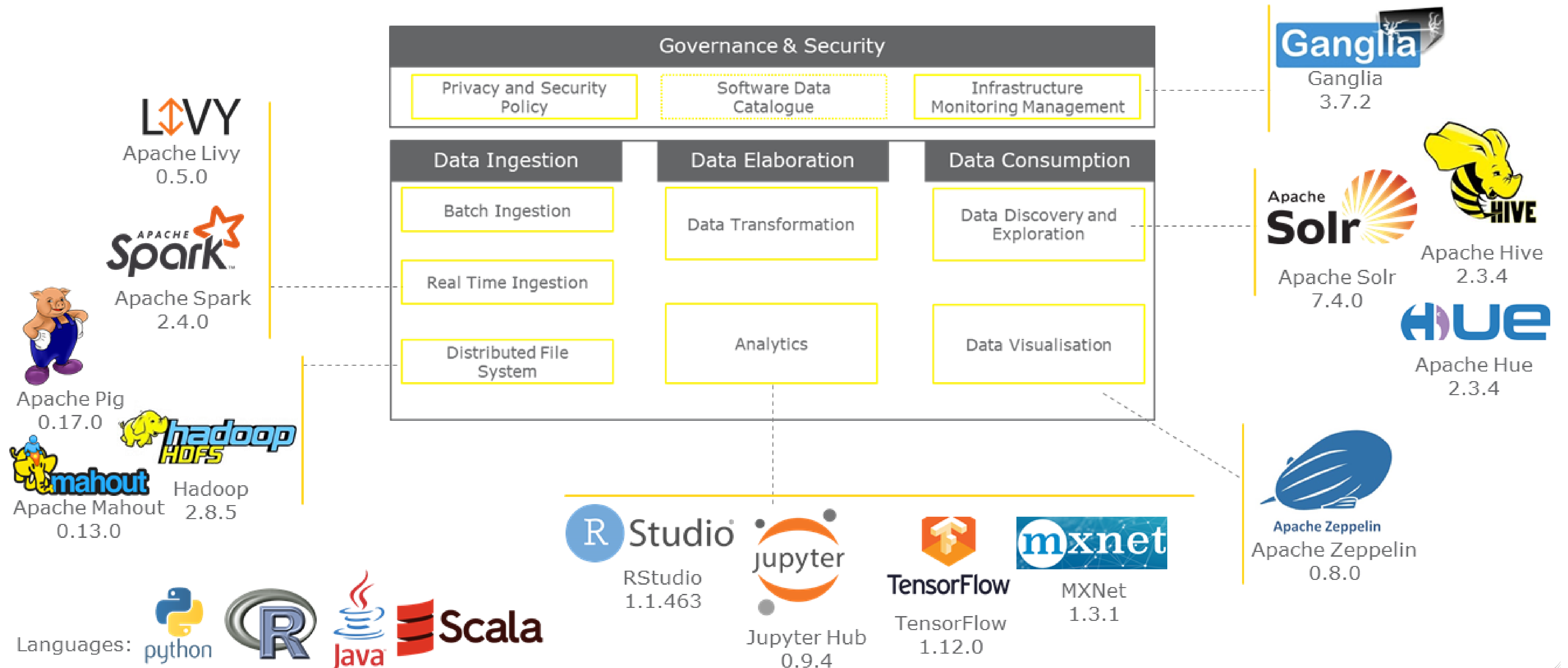
Deploy your data into BDTI Storage  
and/ or  
Access via provided APIs





# How BDTI works (3/3)

**Solution Architecture – Platform as a Service, the BDTI offers SW to access and analyse data**



# BDTI Service Offering



# Business Services

## Overview map

First release (Q1 2019)

Second release (end of 2019)

 **Services already implemented**



**Test Infrastructure**



**Onboarding & stakeholders follow-up**



**Service Desk**



**Services to be implemented**



**Community building and Innovation Portal**



**Data Catalogue**



**Big Data and Analytics software catalogue**



**Analytics as a Service**



**Support for Analytics Implementation**



# Test Infrastructure

The **Test Infrastructure** provides the big data platform and all the data analytics tools supplied by the European Commission through a Platform as a Service.

Through this service, **public administrations** can implement their **own pilots project** in the big data field of expertise or **experiment with big data technologies**.

 **Services already implemented**



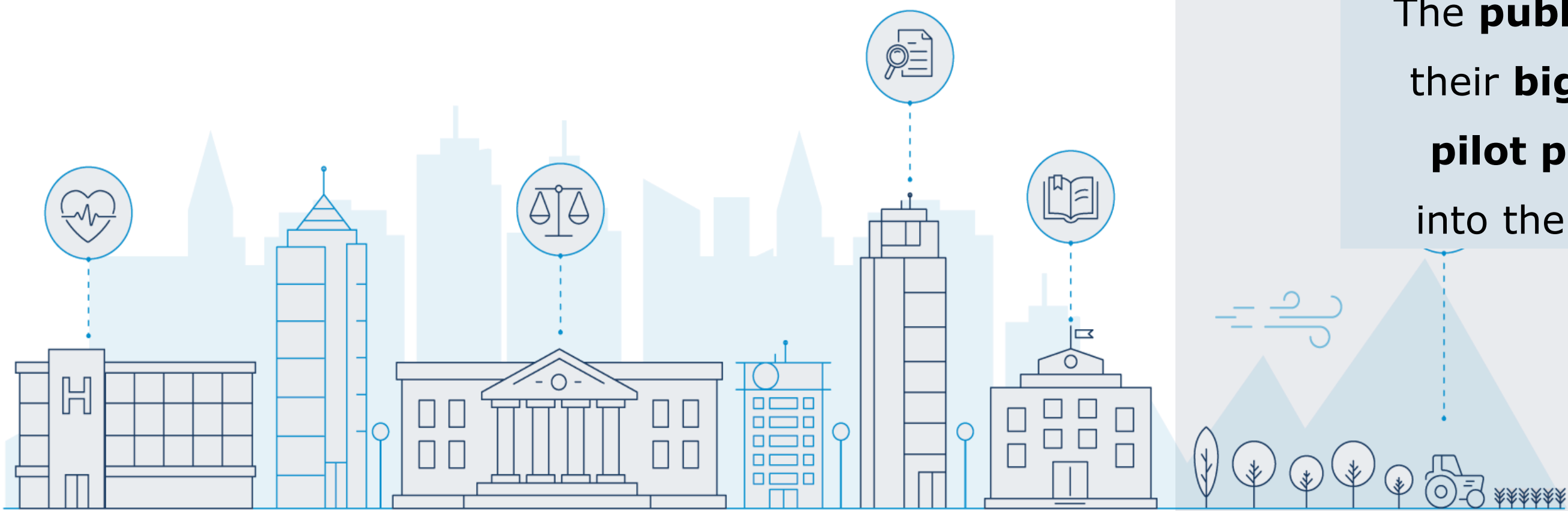
A **public administration** needs an analytical "sandbox" environment to experiment with big data tools and test specific big data use cases.



Test Infrastructure provides a **ready-to-use environment**, respecting privacy policies and using open source tools.



The **public administration** can test their **big data use case** through a **pilot project** before deploying it into their production environment.



# Onboarding & stakeholders follow-up

✓ Services already implemented

## Onboarding & stakeholders follow-up

facilitate the **onboarding process** for stakeholders interested in using the CEF BDTI building block.

Public administrations receive support in the definition of their pilot scope, identification of data sources, or analytical and technical assistance.



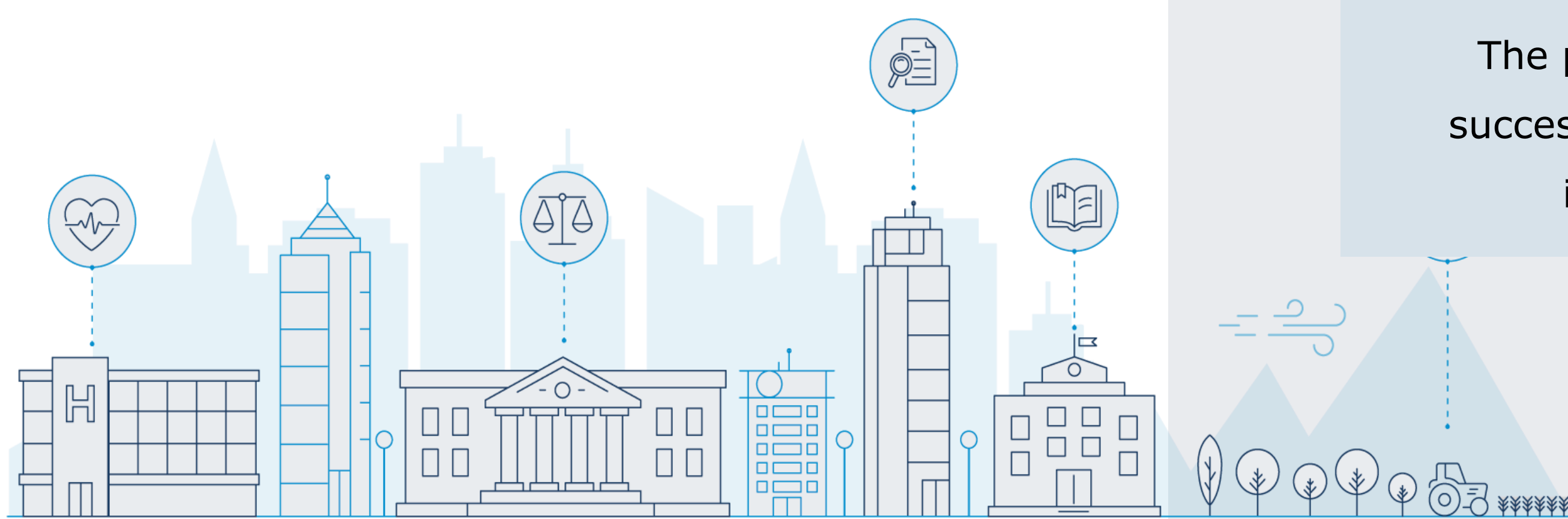
A **public administration** decides to experiment with the test infrastructure and needs guidance during the duration of the pilot.



The onboarding service assists the public administrations throughout the end-to-end pilot development.



The **public administration** successfully completes the test infrastructure pilot.



# Service desk

The service desk acts as a **Single Point of Contact**: collect and classify the tickets and solve them interacting with the users.

During the pilot execution, users can contact the service desk for any kind of technical issue.

 **Services already implemented**



A **public administration** needs **support** using the test infrastructure. Through **CEF digital**, it can create and send a ticket to the Service Desk.



The **Service Desk** takes care of the tickets (e.g., configuration problems, crashes or failures that affects BDTI software) within **8 hours**.



The **public administration** receives updates regarding its issue and BDTI technical team closes the ticket.

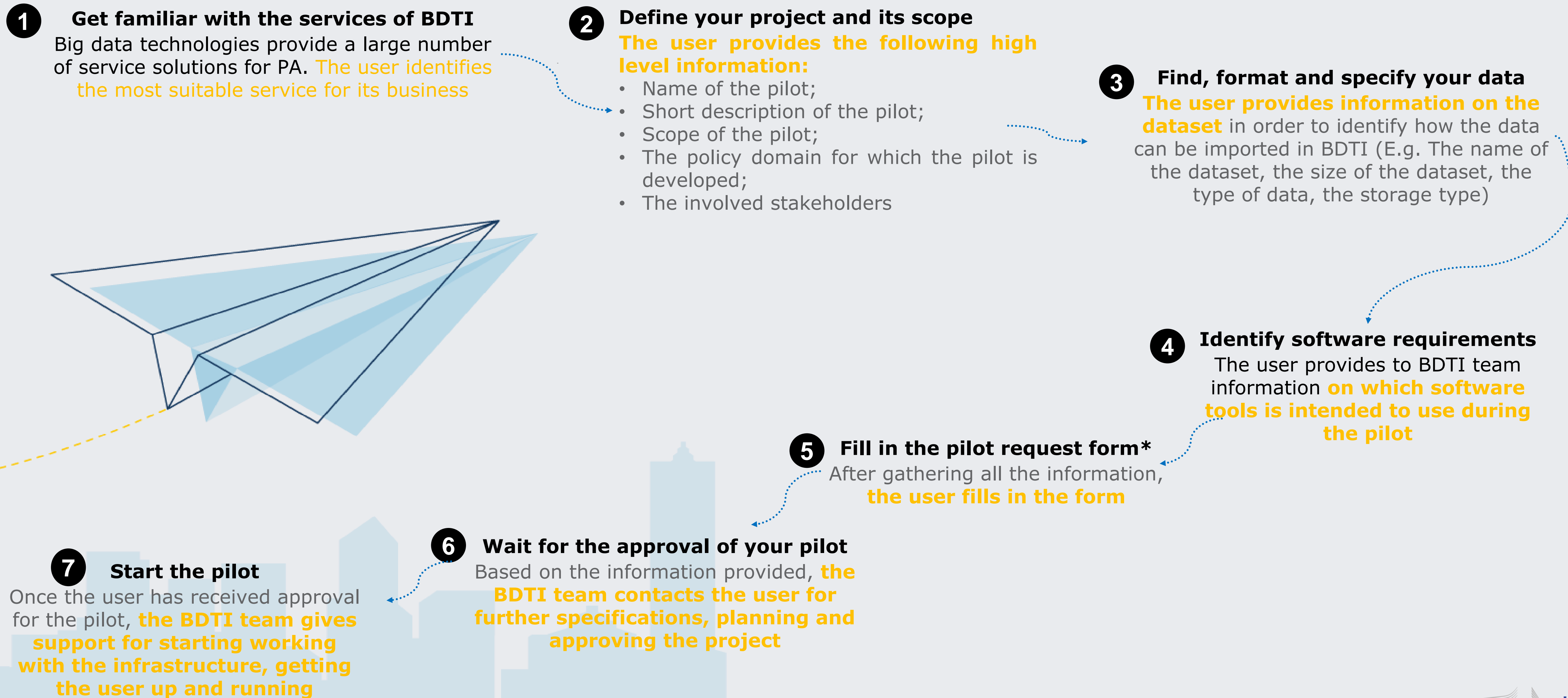




# Application for a BDTI pilot



# Ready to get started?



# Use cases acceptance criteria

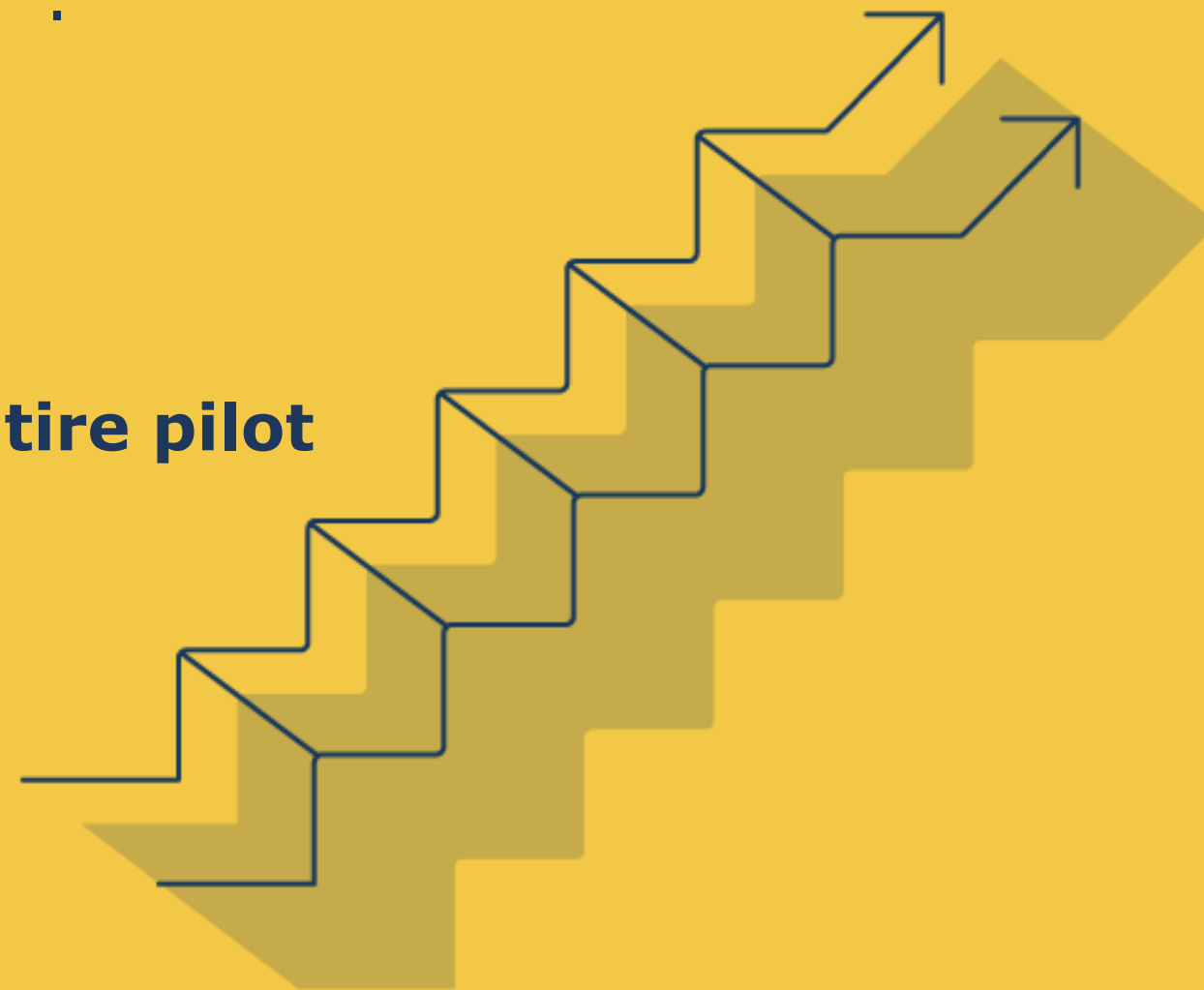
## Business criteria



**Potential users:** Member State or public administration at national level

**Clear value added:** Business and technical

**Clear contact point for the entire pilot**



## Functional criteria



**Pilot duration:** 6 months

**Pilot use cases:** (only use case in scope\*)

**Pilot BDTI geographical distribution/ resource allocation**

**Resource usage limit:** based on CEF budget

**Skills/Maturity level:** adequate skilled resources and/or level of maturity on the big data subject

\*Predictive analysis, Route-traceability / flow monitoring, Web analysis (scraping / monitoring), Text analysis, Descriptive analysis, Time-series analysis, Social media analysis, Network analysis, Population / customer segmentation

# Experiences on BDTI

---





# European Big Data Hackathon



**How can innovative data reduce response burden + extend the time use survey data?**



**BDTI provided big data cloud environments for data analysis and on-site support**

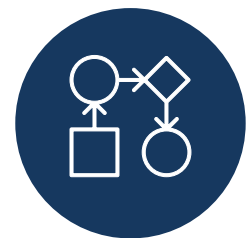


**Results were presented during the 2019 NTTS Conference**





# Big data cloud environment



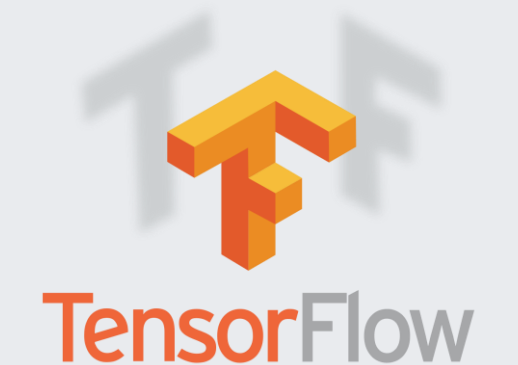
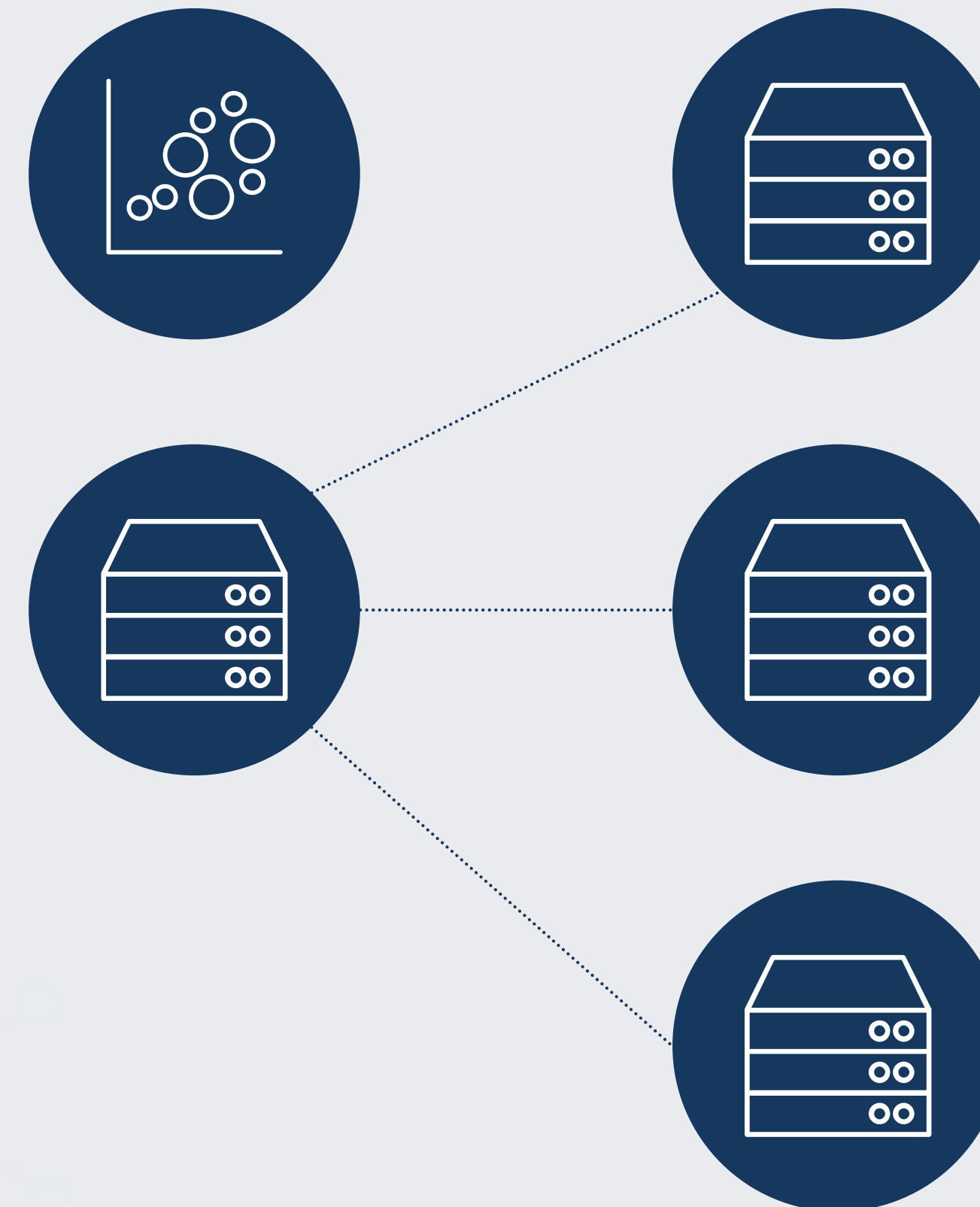
**Big data compute cluster**  
5 nodes: 1 master, 3 cores, 1 visualisation



**Total available memory:** 768 GiB  
**Total available cores:** 192 vCPUs



**State of the art big data software packages**  
Completely extendable, modifiable



# Find our more!

How can the CEF building blocks help you achieve your objectives?



## Visit our website

Learn more on how to get started with the building blocks and get access to specific content such as our Success Stories, tech articles, sample software / specifications, etc.

<http://ec.europa.eu/cefdigital>

## Contact us

Do you want to use the building blocks for your project? Do you want to tell us more about your project?

Contact us: [DIGIT-BDTI-CEF-SUPPORT@ec.europa.eu](mailto:DIGIT-BDTI-CEF-SUPPORT@ec.europa.eu)

[cef-building-block@ec.europa.eu](mailto:cef-building-block@ec.europa.eu)