



bdva.eu  
bigdatastack.eu  
lbidaas.eu  
trackandknowproject.eu



@BDVA\_PPP  
@BigDataStackEU  
@lbidaas  
@Track&Know



# Big Data Pilot Demo Days

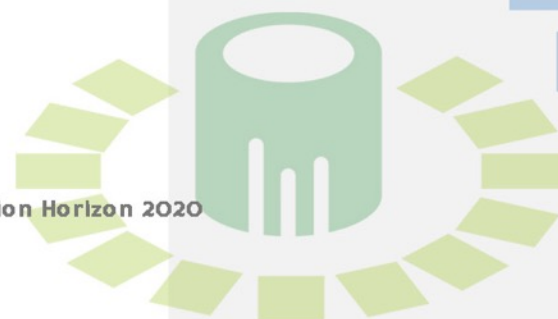
Despina Kopanaki (FORTH) [dkopanaki@ics.forth.gr](mailto:dkopanaki@ics.forth.gr)

Marieke Willems (Trust-IT) [m.willems@trust-itservices.com](mailto:m.willems@trust-itservices.com)

Andrea Schillaci (Trust-IT) [a.schillaci@trust-itservices.com](mailto:a.schillaci@trust-itservices.com)



The below-mentioned project have been funded by the European Commission Horizon 2020  
BigDataStack: grant agreement No 770747  
I-BiDaaS: grant agreement No 780787  
Track and Know: grant agreement No 780754



## BDV PPP Summit 2020 went virtual

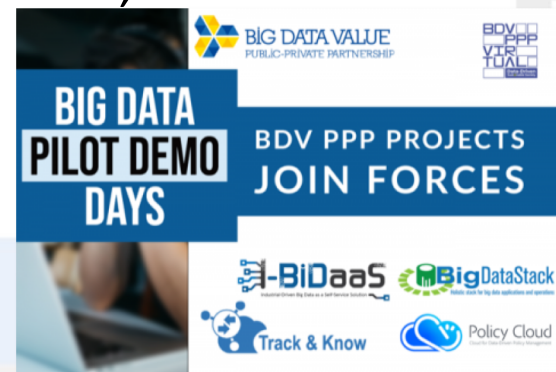


Due to the COVID-19 outbreak, the event is cancelled and some of the activities are going virtual.

Follow us on twitter @BDVA\_PPP to know the latest news and discover our activities.

## Why Big Data Pilot Demo Days?

- The new data-driven industrial revolution highlights the need for **big data technologies to unlock the potential in various application domains.**
- BDV PPP projects I-BiDaaS, BigDataStack and Track & Know and deliver innovative technologies to **address the emerging needs of data operations and applications.**
- To fully exploit the sustainability of the developed technologies, the projects onboarded **pilots that exhibit their applicability in a wide variety of sectors.**
- In their third and final year, the projects are ready **to demonstrate the developed and implemented technologies** to interested end-users from industry as well as technology providers, for further adoption.



## BDV PPP Projects Join Forces



**Holistic stack for big data applications and operations**



**Industrial-Driven Big Data as a Self-Service Solution**

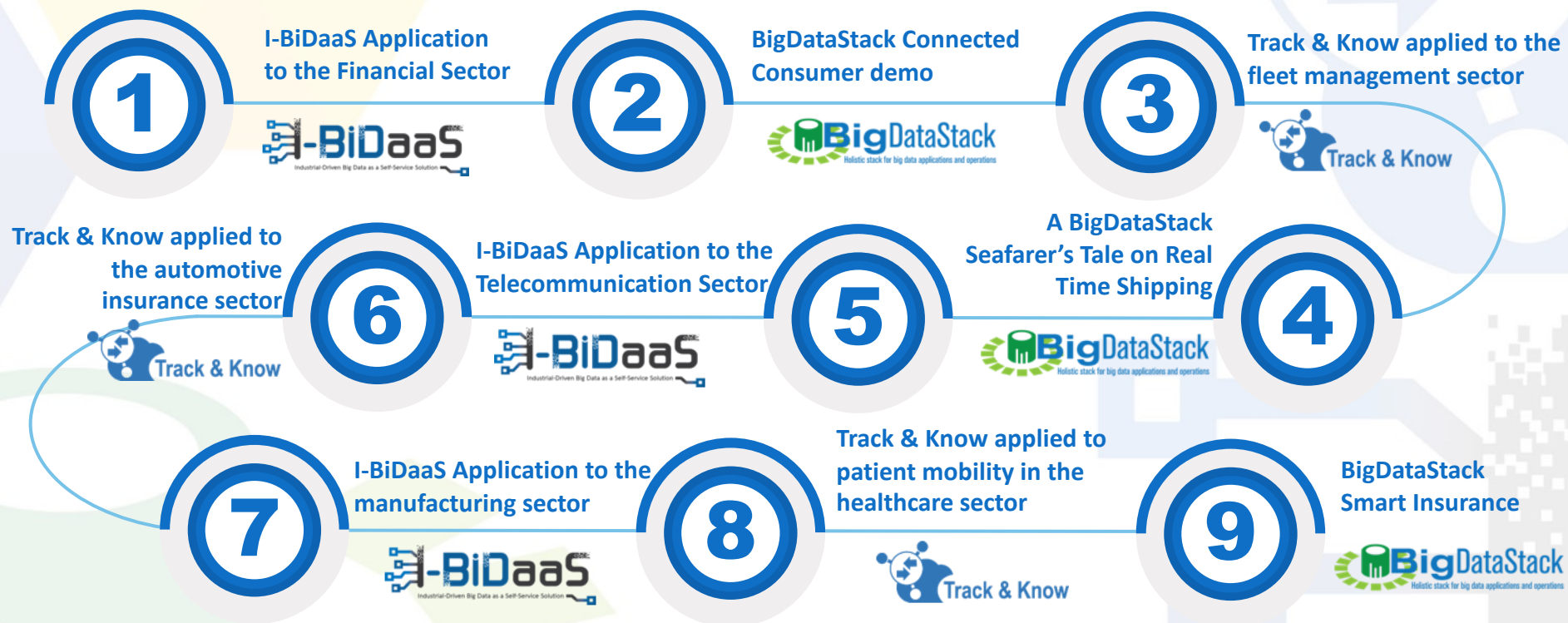


**Big Data for Mobility Tracking Knowledge Extraction in Urban Areas**

BDV PPP projects joining forces to showcase application of innovative technologies in a variety of domains, fostering further adoption, contributing to Europe's digital future.



# Big Data Pilot Demo Days - A Series of Webinars



# Today's Main Topic – Big Data as a Self-Service Solution

- I-BiDaaS overview
- CaixaBank's Pitches: Setting the requirements
- I-BiDaaS architecture: Scientific & Technical view; how it addresses the requirements set by CaixaBank.
- Step by Step demonstration of I-BiDaaS solution and its application to the banking sector.
- Questions & Answers

## Webinar Speakers



**Dr. Dušan Jakovetić**  
*University of Novi Sad, Serbia*

Assistant Professor at the Department of  
Mathematics and Informatics, Faculty of Sciences,  
University of Novi Sad, Serbia

I-BiDaaS Scientific & Technical Manager.



**Dr. Ramon Martin  
de Pozuelo**  
*CaixaBank*

Project Manager at  
Security Innovation & Transformation,  
CaixaBank, Barcelona



[bdva.eu](http://bdva.eu)  
[bigdatastack.eu](http://bigdatastack.eu)  
[lbidaas.eu](http://lbidaas.eu)  
[trackandknowproject.eu](http://trackandknowproject.eu)

 [@BDVA\\_PPP](https://twitter.com/BDVA_PPP)  
[@BigDataStackEU](https://twitter.com/BigDataStackEU)  
[@lbidaas](https://twitter.com/lbidaas)  
[@Track&Know](https://twitter.com/Track&Know)

# Big Data Pilot Demo Days

## I-BiDaaS Application to the Financial Sector

Thursday, May 21, 2020 - 14:00-15:00 CEST



The below-mentioned project have been funded by the European Commission Horizon 2020  
BigDataStack: grant agreement No 770747  
I-BiDaaS: grant agreement No 780787  
Track and Know: grant agreement No 780754





 [bdva.eu](http://bdva.eu)  
[bigdatastack.eu](http://bigdatastack.eu)  
[lbidaas.eu](http://lbidaas.eu)  
[trackandknowproject.eu](http://trackandknowproject.eu)

 [@BDVA\\_PPP](https://twitter.com/BDDVA_PPP)  
[@BigDataStackEU](https://twitter.com/BigDataStackEU)  
[@lbidaas](https://twitter.com/lbidaas)  
[@Track&Know](https://twitter.com/Track&Know)

## *I-BiDaaS Overview*

**Dusan Jakovetic**

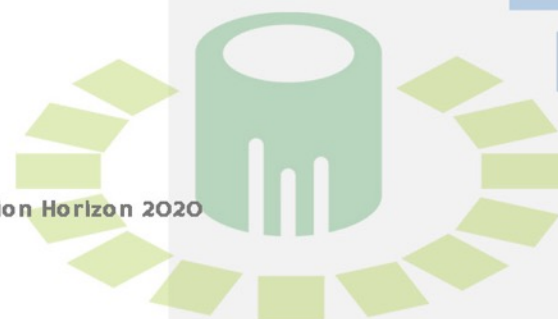
Ass. Professor, University of Novi Sad, Faculty of Sciences, Serbia;  
I-BiDaaS Scientific & Technical Manager

## I-BiDaaS Application to the Financial Sector

Thursday, May 21, 2020 - 14:00-15:00 CEST



The below-mentioned project have been funded by the European Commission Horizon 2020  
BigDataStack: grant agreement No 770747  
I-BiDaaS: grant agreement No 780787  
Track and Know: grant agreement No 780754



## Identity card

**Project Consortium**  
13 partners

**TOTAL BUDGET / TOTAL EC FUNDING**  
€ 4 997 035

**START DATE**  
1 January 2018

**PROJECT NAME**  
Industrial-Driven Big Data as a Self-Service Solution

**PROJECT TYPE**  
RIA

**DURATION**  
36 months



<http://www.ibidaas.eu/>



[@ibidaas](https://twitter.com/ibidaas)



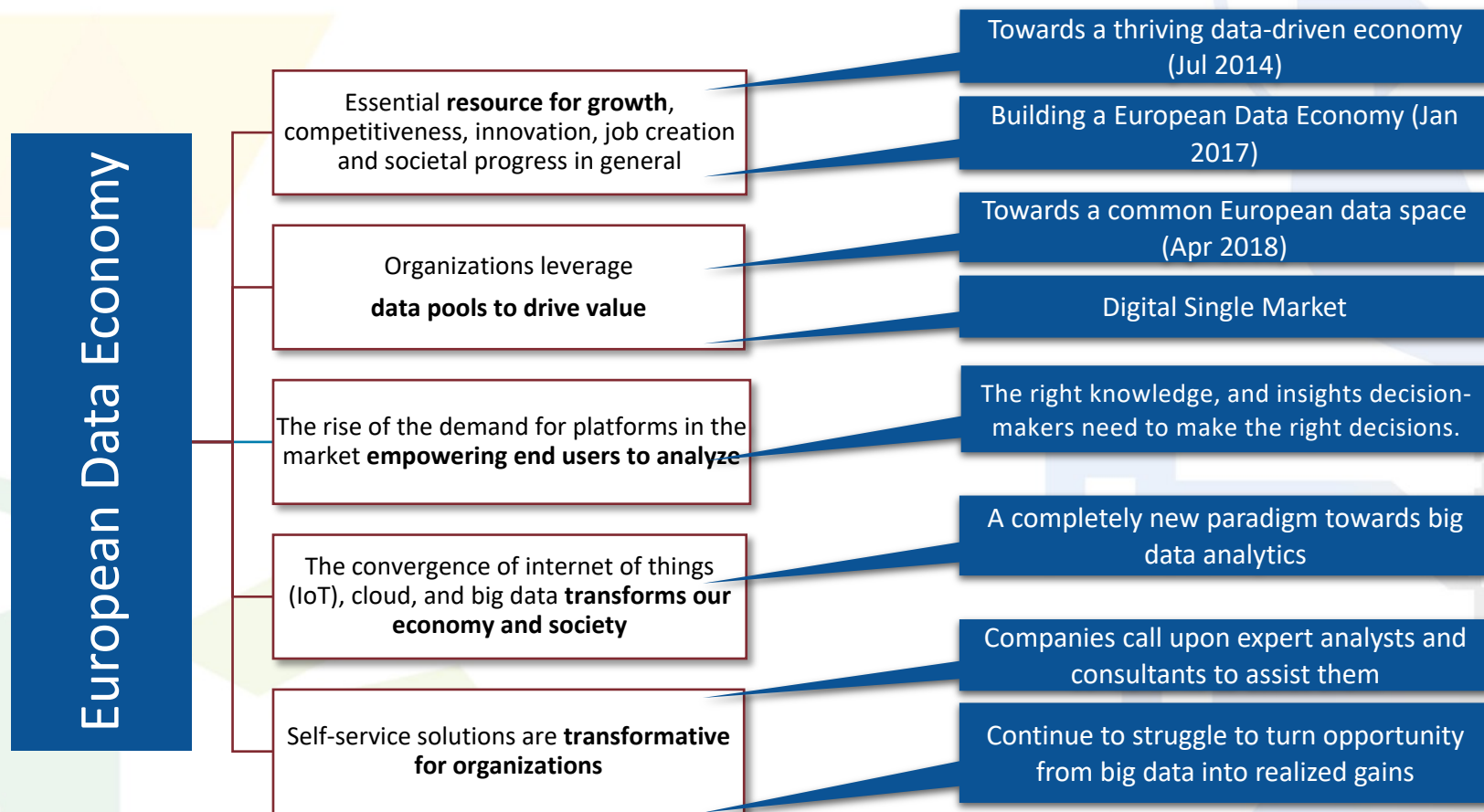
<https://www.linkedin.com/in/i-bidaas/>

## I-BiDaaS Consortium

1. FOUNDATION FOR RESEARCH AND TECHNOLOGY HELLAS (**FORTH**)
2. BARCELONA SUPERCOMPUTING CENTER - CENTRO NACIONAL DE SUPERCOMPUTACION (**BSC**)
3. IBM ISRAEL - SCIENCE AND TECHNOLOGY LTD (**IBM**)
4. CENTRO RICERCHE FIAT SCPA (**CRF**)
5. SOFTWARE AG (**SAG**)
6. CAIXABANK, S.A (**CAIXA**)
7. THE UNIVERSITY OF MANCHESTER (**UNIMAN**)
8. ECOLE NATIONALE DES PONTS ET CHAUSSEES (**ENPC**)
9. ATOS SPAIN SA (**ATOS**)
10. AEGIS IT RESEARCH LTD (**AEGIS**)
11. INFORMATION TECHNOLOGY FOR MARKET LEADERSHIP (**ITML**)
12. University of Novi Sad Faculty of Sciences Serbia (**UNSPMF**)
13. TELEFONICA INVESTIGACION Y DESARROLLO SA (**TID**)



# Motivation





## Our Vision



A **complete** and **safe environment** for methodological **big data experimentation**



A Big Data as a **Self-Service solution** that helps breaking industrial data silos and boosts EU's data-driven economy



**Increases impact** in research community and contributes to industrial innovation capacity



Tool and services to **increase the quality** of data analytics



Tools and services for **fast ingestion and consolidation** of both realistic and fabricated data



Tools and services for the management of **heterogeneous infrastructures** including elasticity

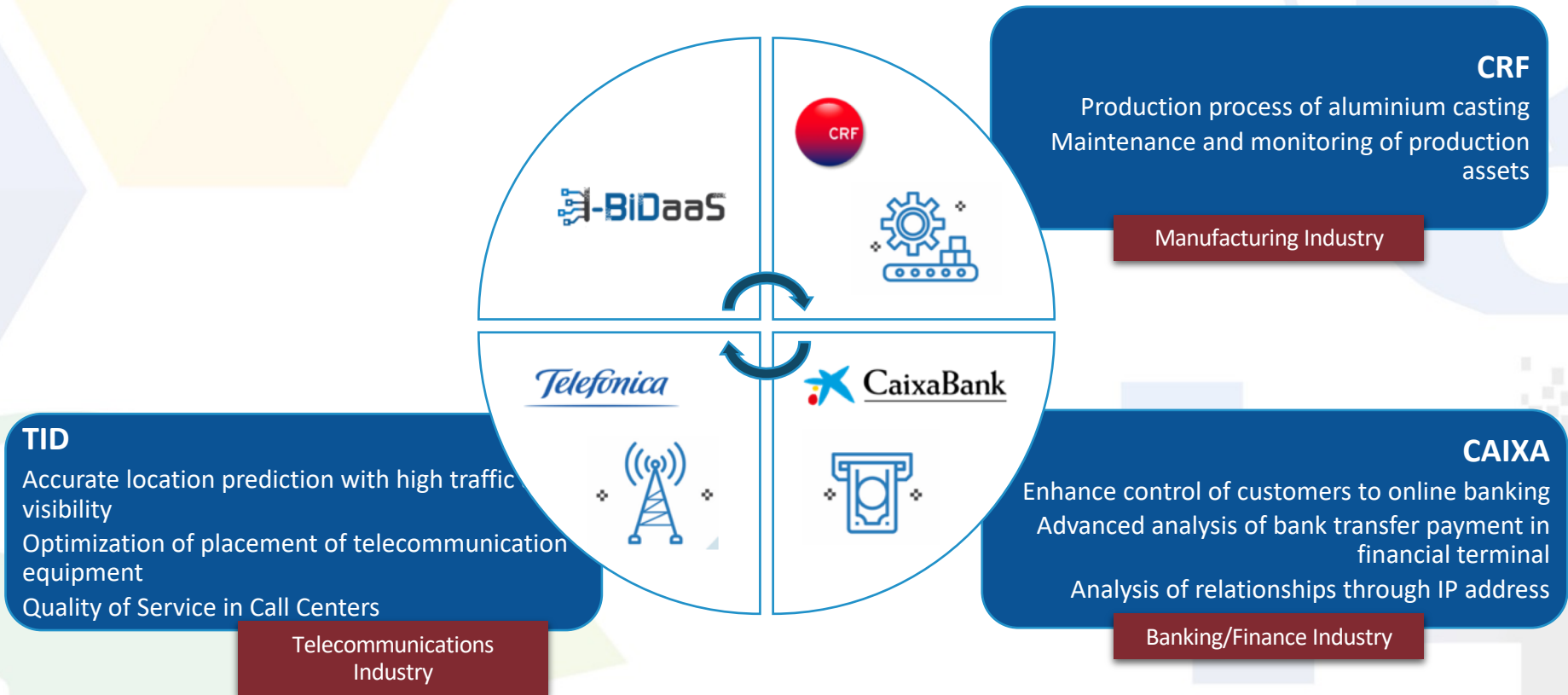
## Project Statement

***I-BiDaaS aims to empower users to easily utilize and interact with big data technologies, by designing, building, and demonstrating, a unified framework that:***

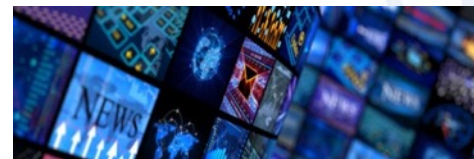
***significantly increases the speed of data analysis while coping with the rate of data asset growth, and facilitates cross-domain data-flow towards a thriving data-driven EU economy.***

***I-BiDaaS will be tangibly validated by three real-world, industry-lead experiments.***

## Application / Experimentation



## I-BiDaaS application domains





# *Setting the Pilot Requirements*

**Ramon Martín de Pozuelo**

Project Manager at Security Innovation & Transformation at  
CaixaBank

## I-BiDaaS Application to the Financial Sector

Thursday, May 21, 2020 - 14:00-15:00 CEST



The below-mentioned project have been funded by the European Commission Horizon 2020  
BigDataStack: grant agreement No 770747  
I-BiDaaS: grant agreement No 780787  
Track and Know: grant agreement No 780754

## CaixaBank and the Use of Data



CaixaBank is the **leading financial group in Spain**, both in banking and insurance and it is developing a strategy of diversification with stakes in international banks and also within leading service companies.



**13.8M** Customers



**5.8M** On-line banking



**4.8M** Mobile Banking



**32K** Employees



**9.1K** ATMs



**4.2K** Branches

In January 2014 CaixaBank created **the Big Data Department**

- We manage 1.247 TB of information only in our Big data
- Department of more than 100 internal people providing Data analytics services to all the Organization.
- Due to Regulation constraints all the infrastructure and the analysis is done internally

# Why do we need I-BiDaaS?

## Current data sharing situation in CaixaBank and I-BiDaaS approach



We have **tons of data** but **confidential**.

We are the data managers but the **real owners of the data are our customers**.



**Lack of agility in our Datapool** for extracting data externally.

**Security procedures and constraints** are necessary but hinder and slow down data sharing processes.

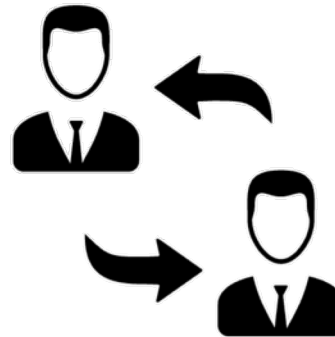
### Requirements

Requirement	Control
<b>Data privacy</b>	Data of customers (e.g., social graph) and external suppliers (e.g., SIEM) can not be accessed or shared with other partners
<b>Regulation Compliance</b>	All activities related to CaixaBank within the project must comply with the relevant regulations (e.g., ISO 27001, GDPR)
<b>Fraud and Security Analytics</b>	Use cases presented will be related to the improvement of security and the prevention of fraud.

**Objective:** Exploit the I-BiDaaS platform to gain agility, efficiency and flexibility in our analytics for security.

## Why do we need I-BiDaaS?

- **Breaking external and cross-sectorial data silos while complying Regulation**
  - *Sharing Data models with other FI*
  - *Sharing Data models with other sectors*
  - *Following ECB & Banco de España constraints, we already have proved with I-BiDaaS that this is possible*



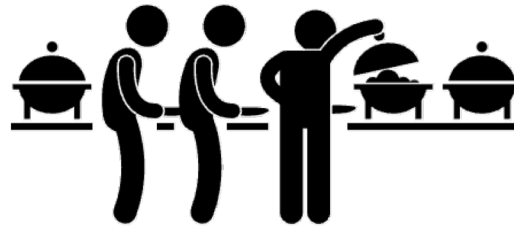
20



## Why do we need I-BiDaaS?

### ■ Secure Self-Service Infrastructure

- *Being able to outsource Big Data Infrastructure preserving privacy & security*
- *To grow in a dedicated and specialized infrastructure*
- *To count on dedicated and specialized specialists*



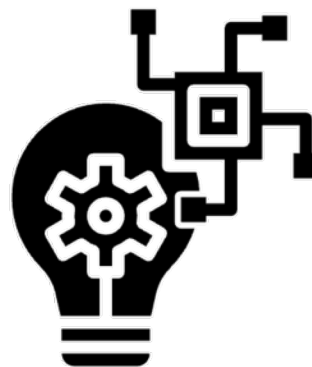
## Why do we need I-BiDaaS?

- **Competitiveness & Innovation**

- *Fast, agile and specialized adaptation to new technology.*
- *Vs Current proprietary infrastructure*

- **Efficiency**

- *Reducing the costs and time of analyzing large datasets*





 [bdva.eu](http://bdva.eu)  
[bigdatastack.eu](http://bigdatastack.eu)  
[lbidaas.eu](http://lbidaas.eu)  
[trackandknowproject.eu](http://trackandknowproject.eu)

 [@BDVA\\_PPP](https://twitter.com/BDVA_PPP)  
[@BigDataStackEU](https://twitter.com/BigDataStackEU)  
[@lbidaas](https://twitter.com/lbidaas)  
[@Track&Know](https://twitter.com/Track&Know)

## Big Data Architecture

**Dusan Jakovetic**

Assistant Professor, University of Novi Sad, Serbia;  
I-BiDaaS Scientific & Technical Manager

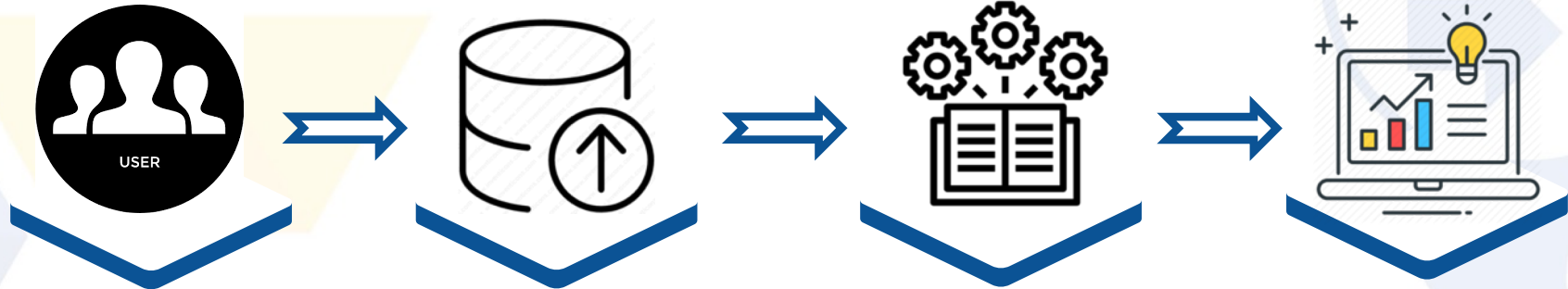
## I-BiDaaS Application to the Financial Sector

Thursday, May 21, 2020 - 14:00-15:00 CEST



The below-mentioned project have been funded by the European Commission Horizon 2020  
BigDataStack: grant agreement No 770747  
I-BiDaaS: grant agreement No 780787  
Track and Know: grant agreement No 780754

# The I-BiDaaS Solution: Front-end



## Users

- Expert mode
- Self-service mode
- Co-develop mode

## Data

- Import your data
- Fabricate Data
- Tokenize data

## Analyze your Data

- Stream & Batch Analytics
- Expert: Upload your code
- Self-service: Select an algorithm from the pool
- Co-develop: custom end-to-end application

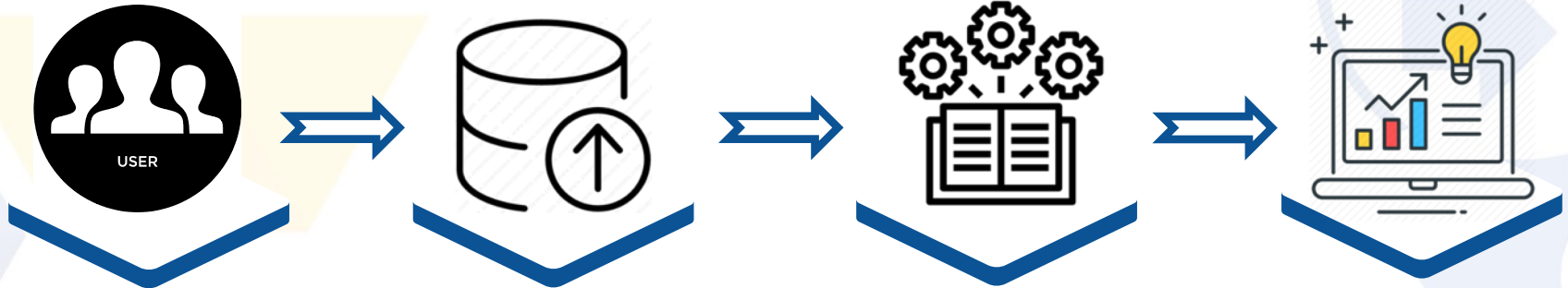
## Results

- Visualize the results
- Share models

## Benefits of using I-BiDaaS



# The I-BiDaaS Solution: Front-end



## Users

- Expert mode
- Self-service mode
- Co-develop mode

## Data

- Import your data
- Fabricate Data
- Tokenize data

## Analyze your Data

- Stream & Batch Analytics
- Expert: Upload your code

## Results

- Visualize the results
- Share models

## Benefits of using I-BiDaaS



Do it  
yourself  
in a flexible  
manner



Break data silos



Safe environment



Interact with Big Data  
technologies



Improve your  
data analysis

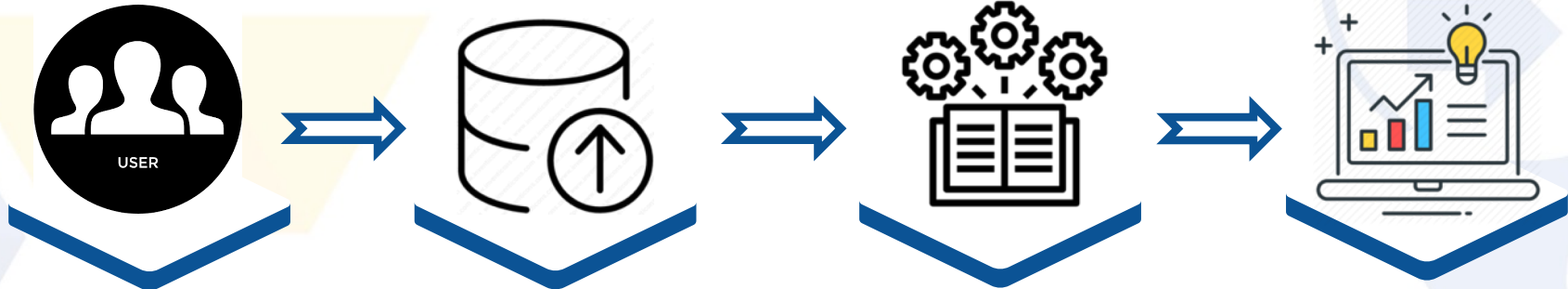
Improve your  
domain data-  
flow

Cope with the rate of  
data asset growth





## The I-BiDaaS Solution: Front-end



### Users

- Expert mode
- Self-service mode
- Co-develop mode

### Data

- Import your data
- Fabricate your data

### Analyze your Data

- Stream & Batch Analytics
- Expert: Upload your code

### Results

- Visualize the results
- Share models

### Benefits of using I-BiDaaS



Do it yourself  
in a flexible manner



Break data silos



Safe environment

Self-service & count on specialists



Interact with big data technologies

Increase speed of data analysis

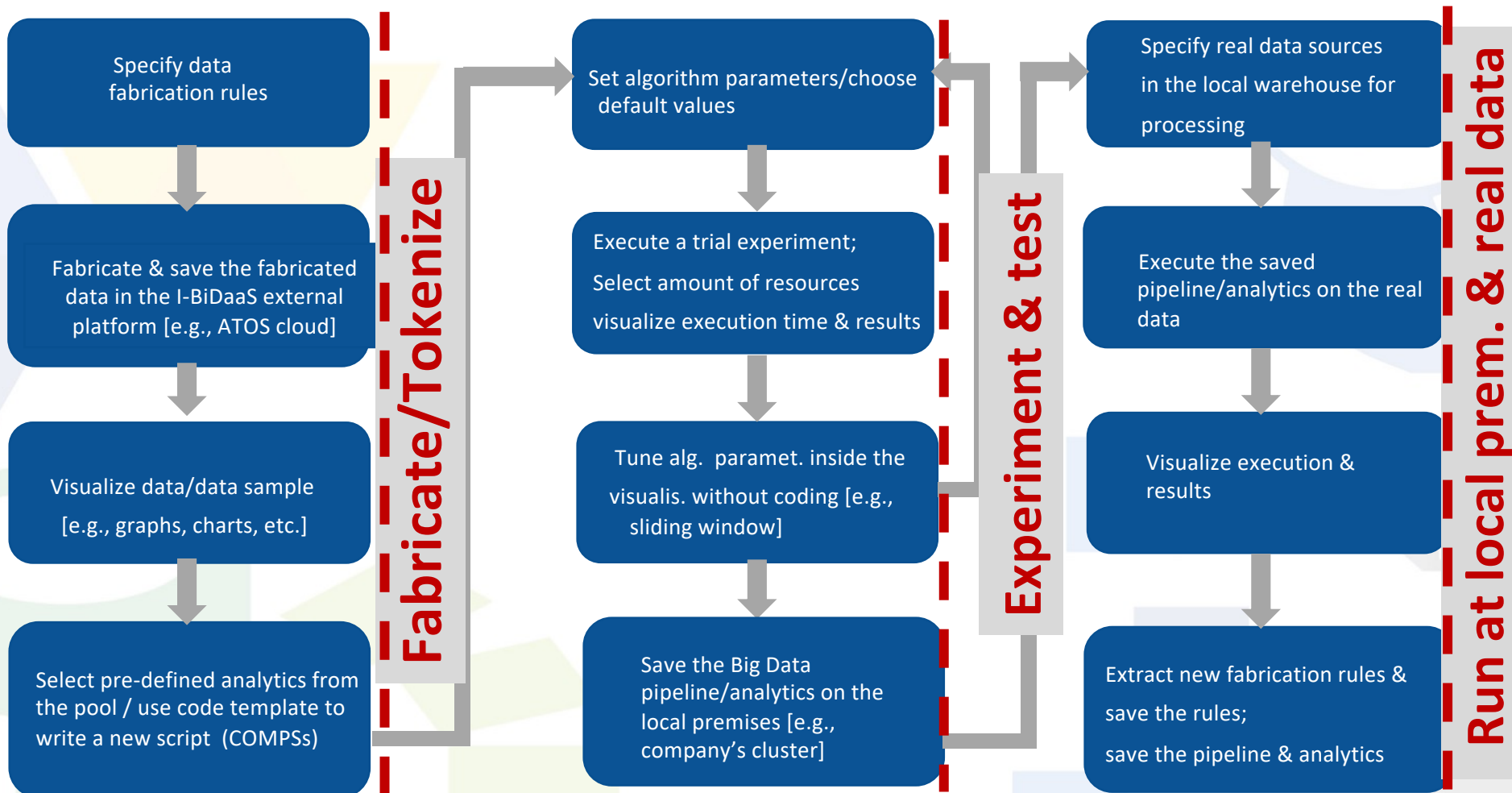


Intra- and inter-domain data-flow

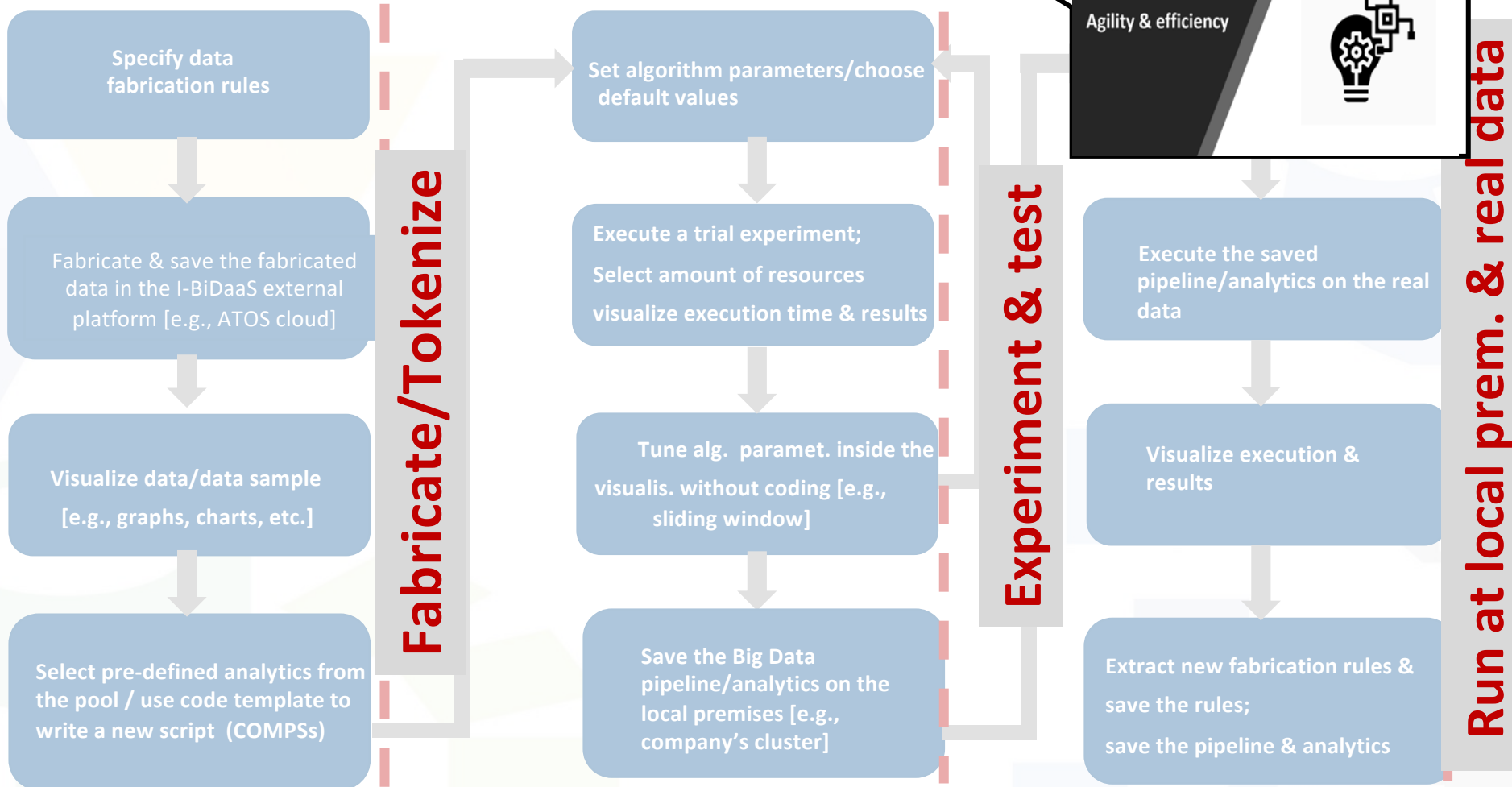


Cope with the rate of data asset growth

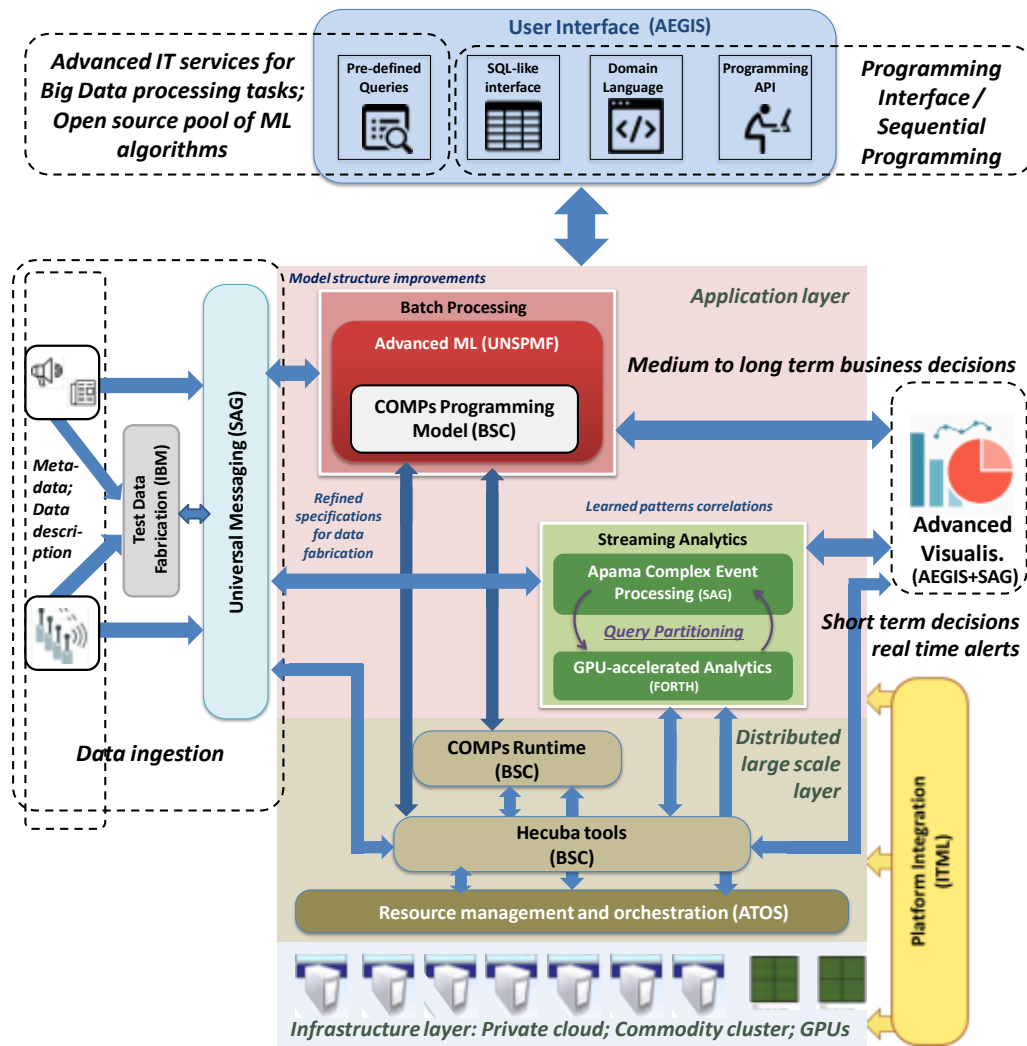
# I-BiDaaS – Prototypical Experimental Workflow

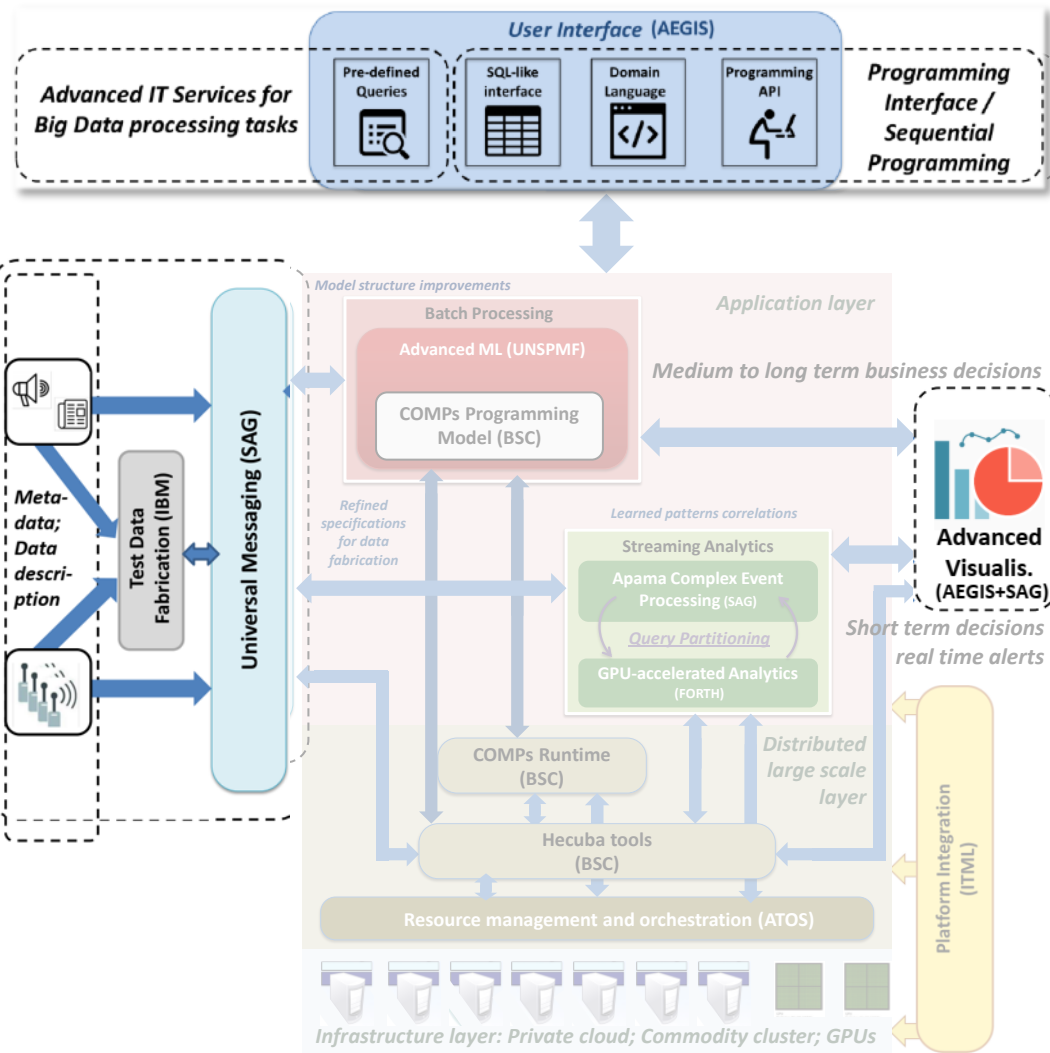


# I-BiDaaS – Experimental Workflow



## The I-BiDaaS Solution: Architecture/back-end





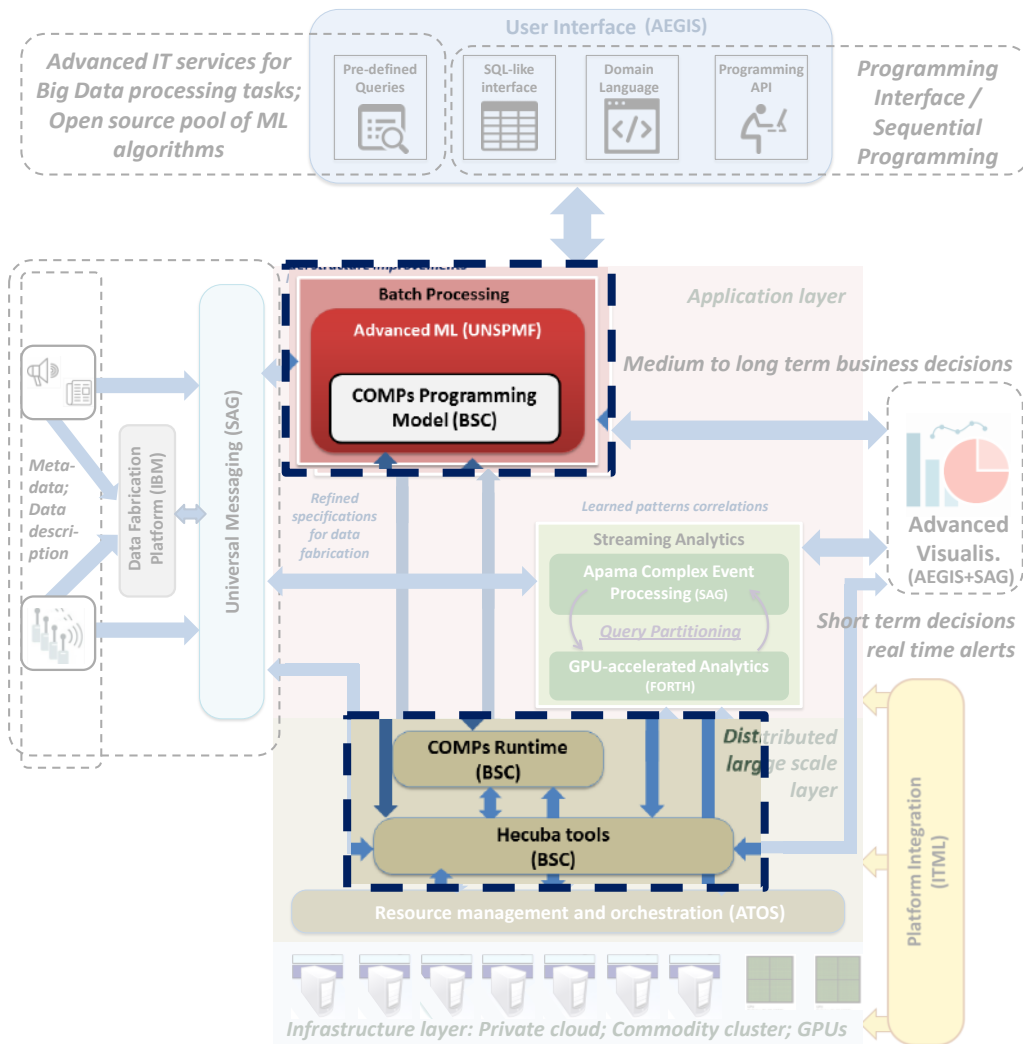
**WP2:**  
*Data, user interface, visualization*

**Technologies:**

- IBM TDF
- SAG UM
- AEGIS AVT

<http://ibidaas.eu/tools>



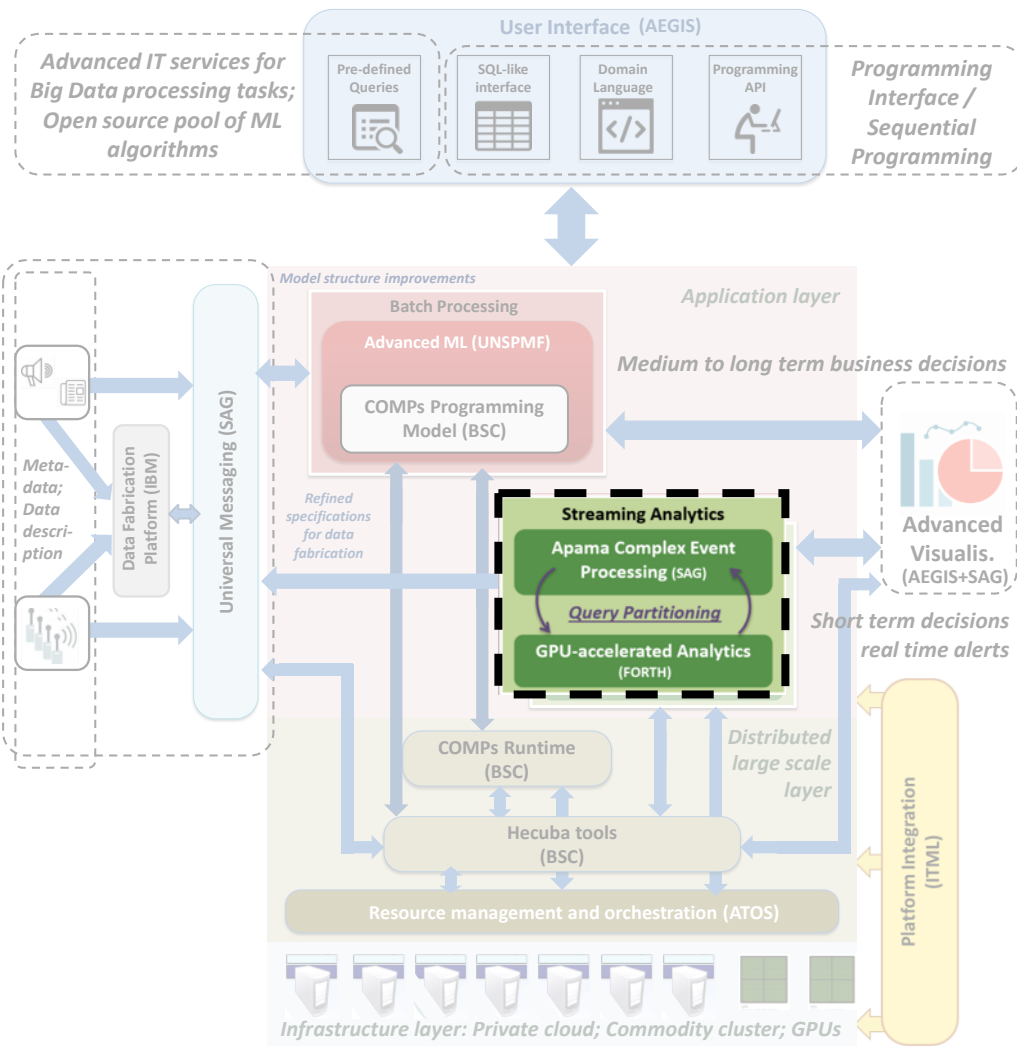


## WP3: Batch analytics

## Technologies:

- BSC COMPSs
- BSC Hecuba
- BSC Qbeast
- Advanced ML (UNSPMF)

<http://ibidaas.eu/tools>

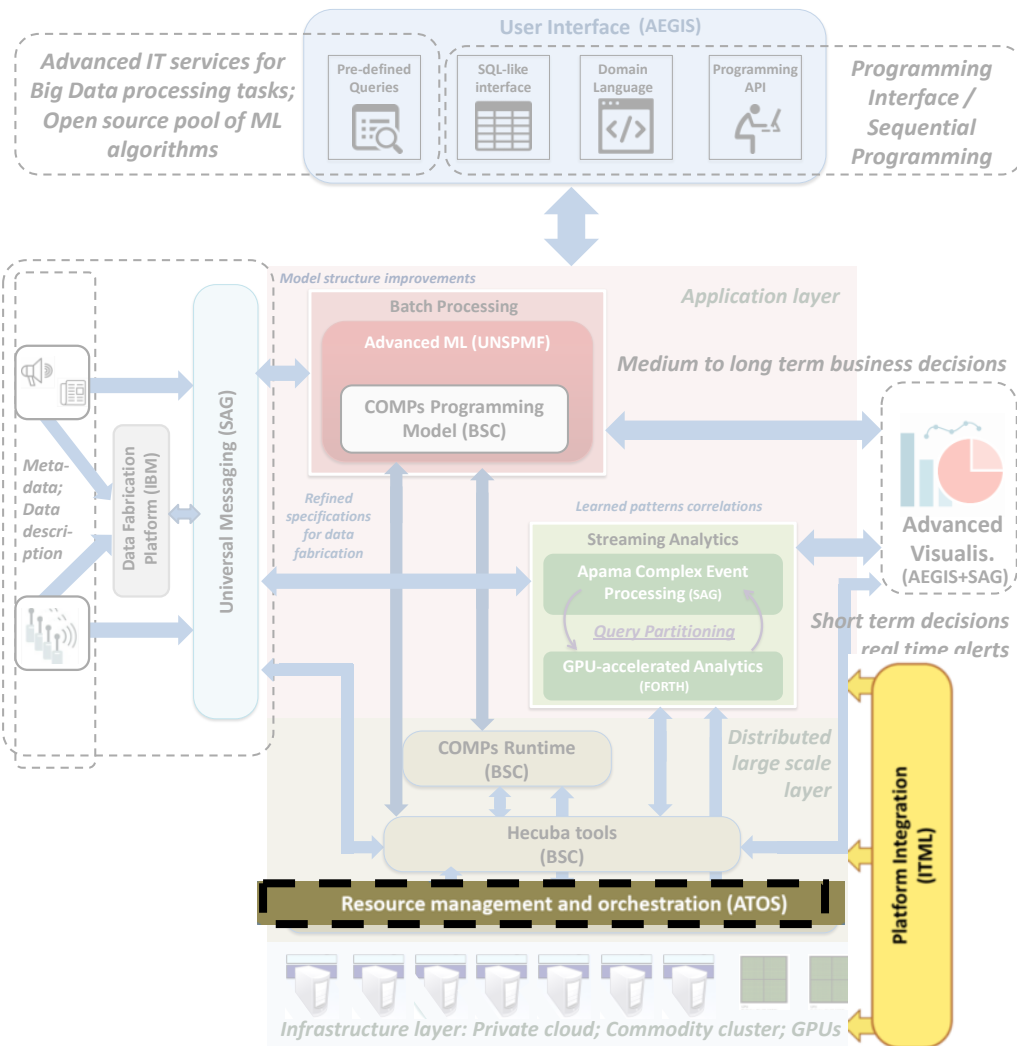


## WP4: Streaming analytics

### Technologies:

- SAG Apama CEP
- FORTH GPU-accel. analytics

<http://ibidaas.eu/tools>



## WP5: Resource mgmt & integration

### Technologies:

- ATOS Resource mgmt
- ITML integration services

<http://ibidaas.eu/tools>

## Key Features & Innovations



Data fabrication capabilities

Solution flexibility

Easy to code programming paradigm

High code reusability

<https://www.ibidaas.eu/deliverables>

## Key Features & Innovations

Data fabrication capabilities

***CAIXA: From 3 months to 1-2 weeks for a new proof-of-concept Big Data technology***

Solution flexibility

Easy to code programming paradigm

High code reusability

<https://www.ibidaas.eu/deliverables>



## Key Features & Innovations

Data fabrication capabilities

Solution flexibility


***CAIXA: Bank transfers use case:  
~3-4 less time to analysis***

Easy to code programming paradigm

High code reusability

<https://www.ibidaas.eu/deliverables>

## Key Features & Innovations (Cont'd)



GPU-accelerated analytics; Synergy of CEP and GPU-accelerated analytics for streaming data

Feedback from analytics to data fabrication

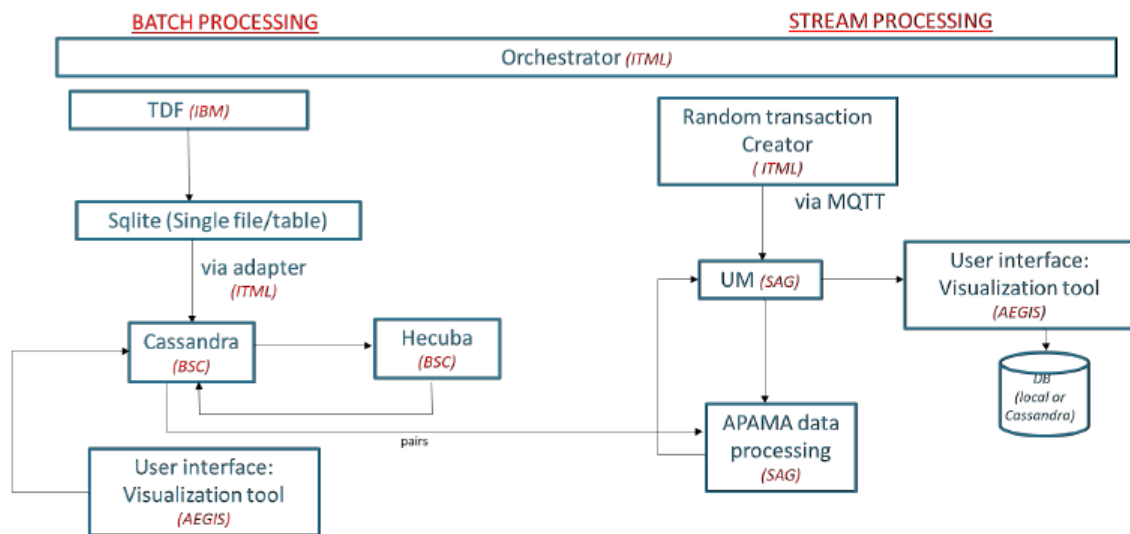
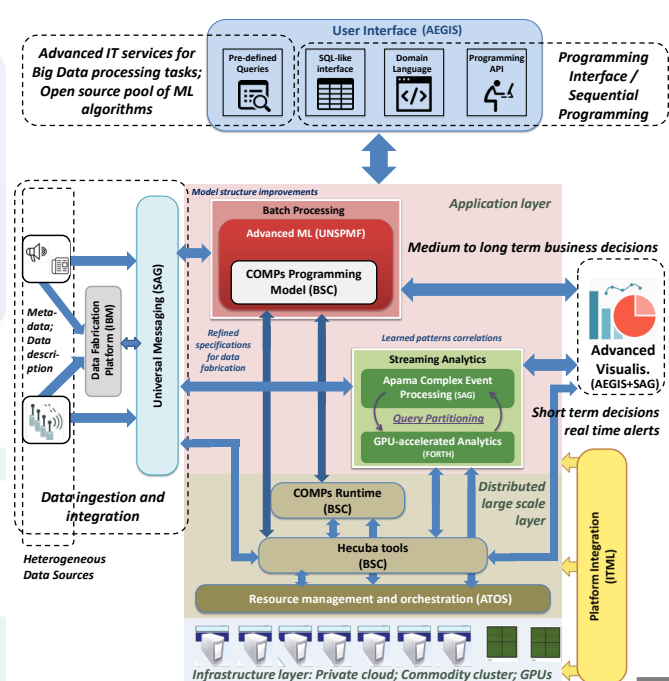
Feedback from analytics to problem modelling

Demonstrated on use cases across 3 different data providers and 3 different industries

<https://www.ibidaas.eu/deliverables>



# I-BiDaaS Solution: A CaixaBank Use Case Example



<https://www.ibidaas.eu/deliverables>

# I-BiDaaS Solution: A CaixaBank Use Case Example

## Data fabrication:

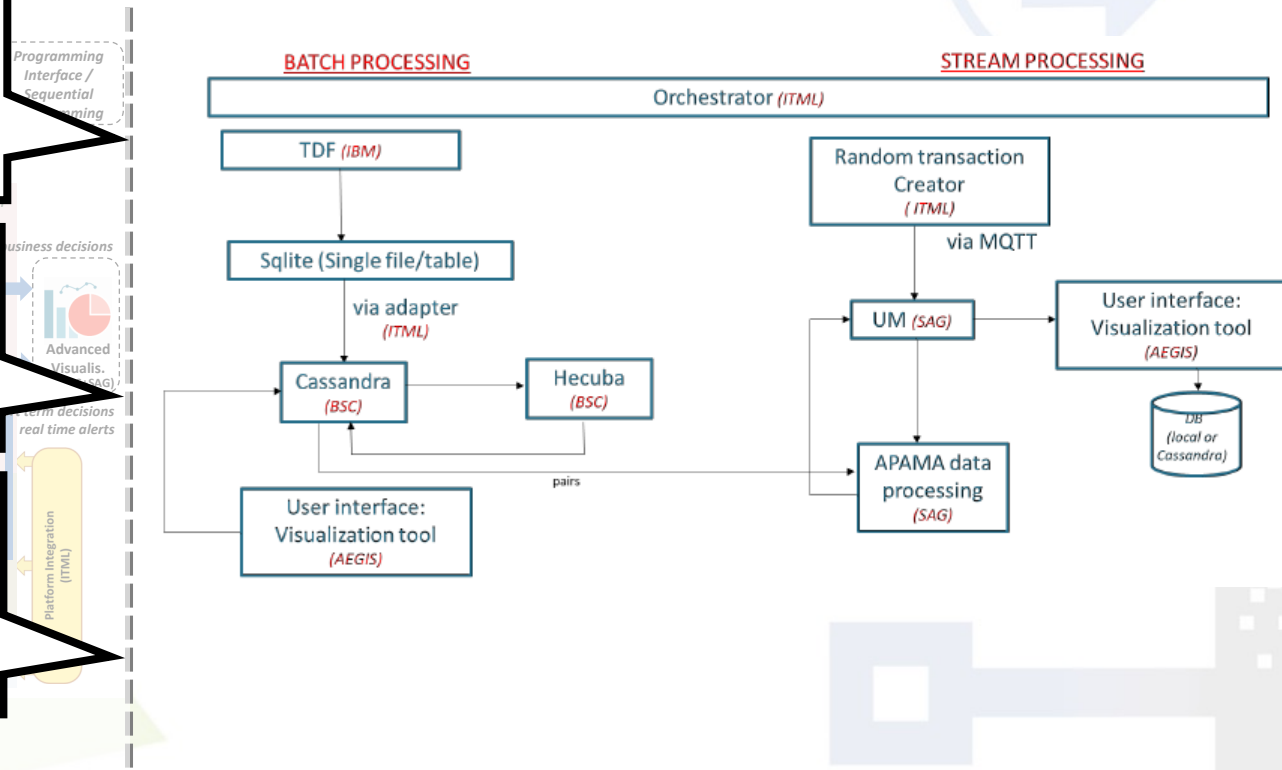
75% time reduction for data access by external stakeholders

## Real time analytics:

Detecting relations of users in real time

## Solution flexibility:

Found more useful results than before the project



<https://www.ibidaas.eu/deliverables>



bdva.eu  
bigdatastack.eu  
lbidaas.eu  
trackandknowproject.eu



@BDVA\_PPP  
@BigDataStackEU  
@lbidaas  
@Track&Know

## *Financial Pilot: step by step*

**Ramon Martín de Pozuelo**

Project Manager at Security Innovation & Transformation at  
CaixaBank

## I-BiDaaS Application to the Financial Sector

Thursday, May 21, 2020 - 14:00-15:00 CEST

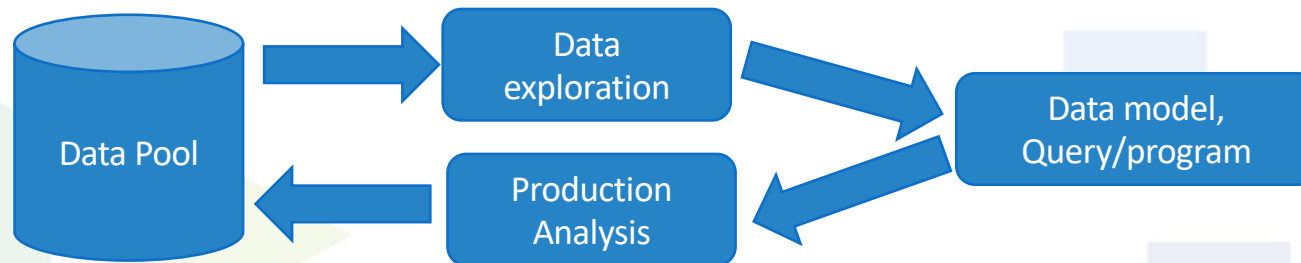


The below-mentioned project have been funded by the European Commission Horizon 2020  
BigDataStack: grant agreement No 770747  
I-BiDaaS: grant agreement No 780787  
Track and Know: grant agreement No 780754

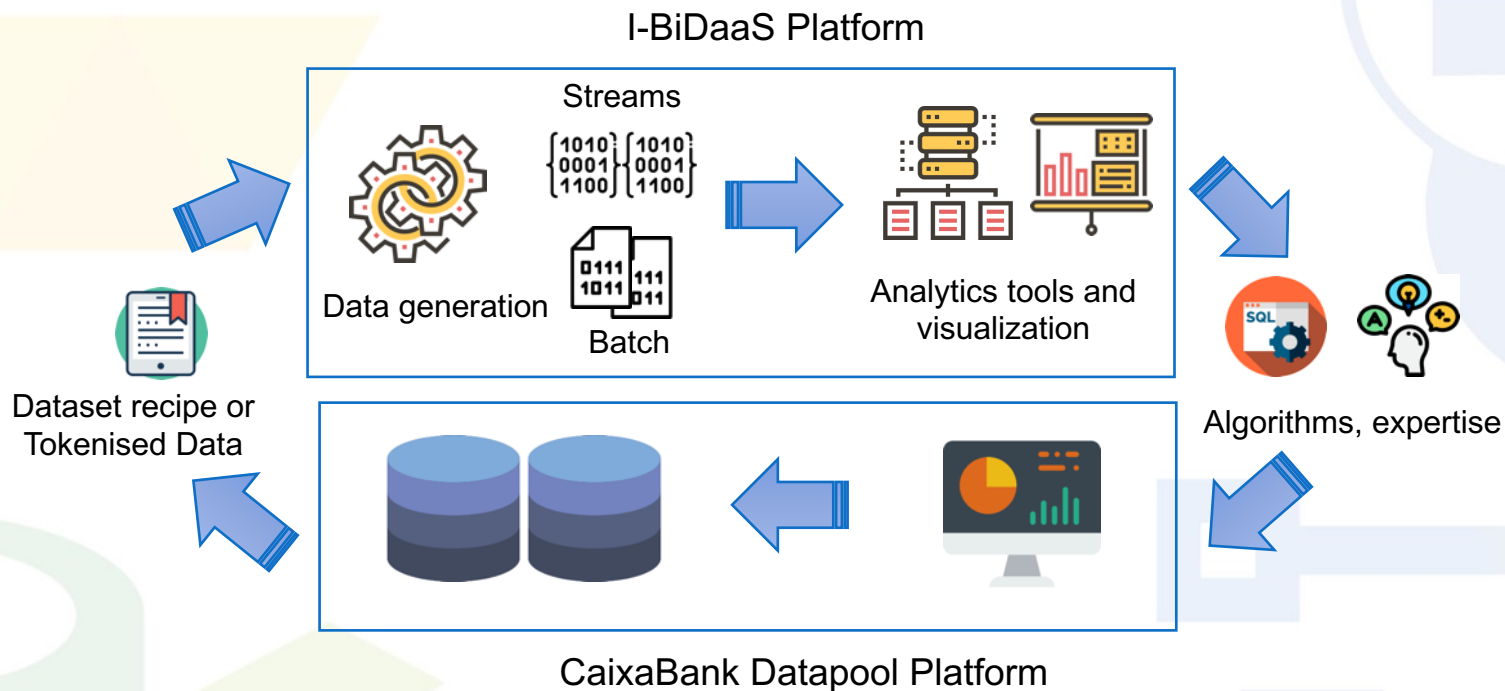


## Current situation

- **Fraud Detection Analytics in CaixaBank**
  - **Currently**, CaixaBank analytics are **executed in-house**.
  - **Analytics lifecycle:**
    - *Security analysis data storage: DataPool (Oracle), Qradar (IBM)*
    - *Data exploration phase to build a model/query (expensive to run)*
    - *Execution of model/query on data in production mode*
  - This process is executed **periodically**.



# I-BiDaaS solution



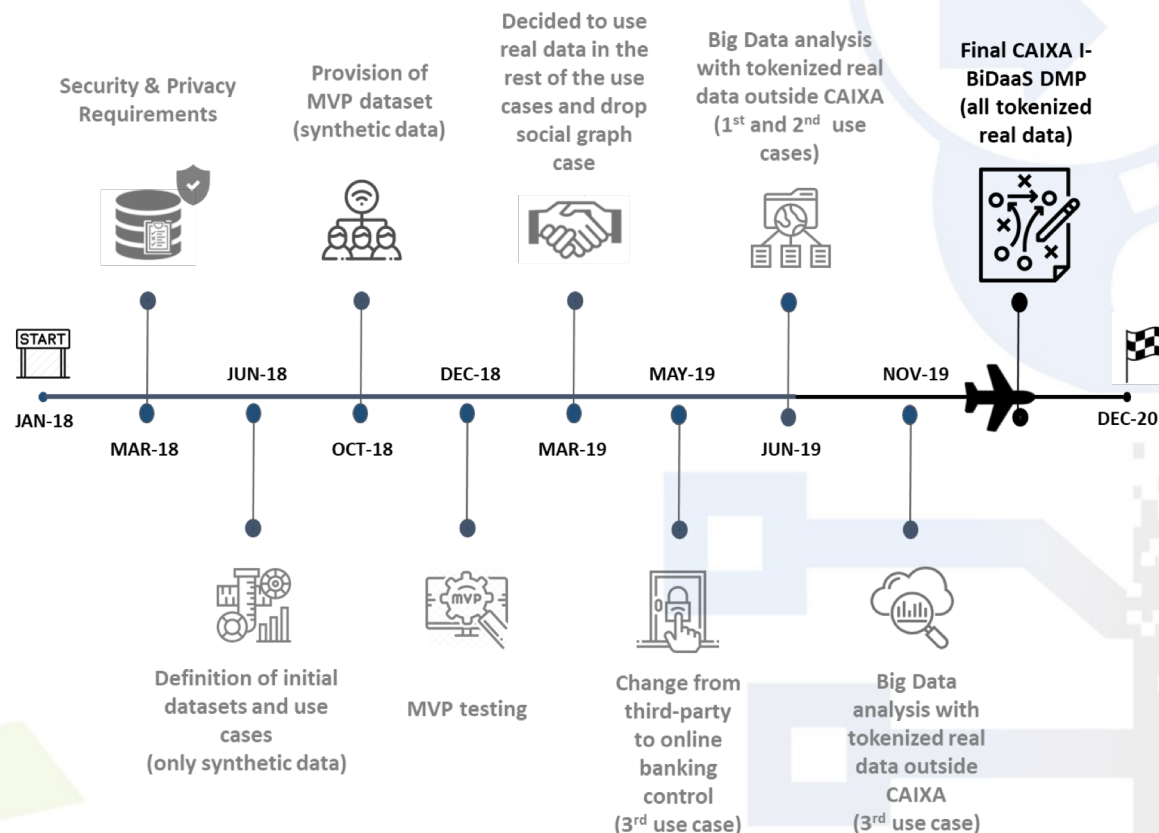
# CaixaBank Data Roadmap

## ■ Synthetic data usage:

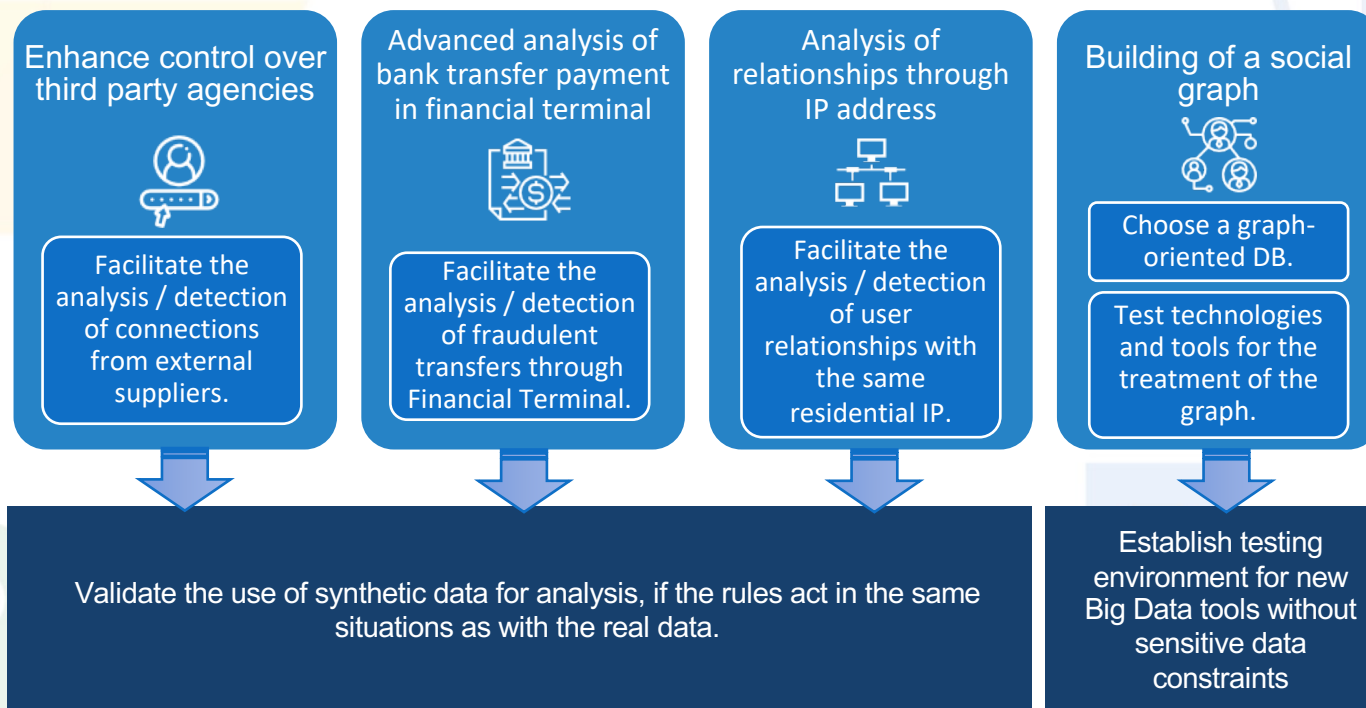
- *We explored the synthetic data solution in the MVP.*

## ■ New opportunities:

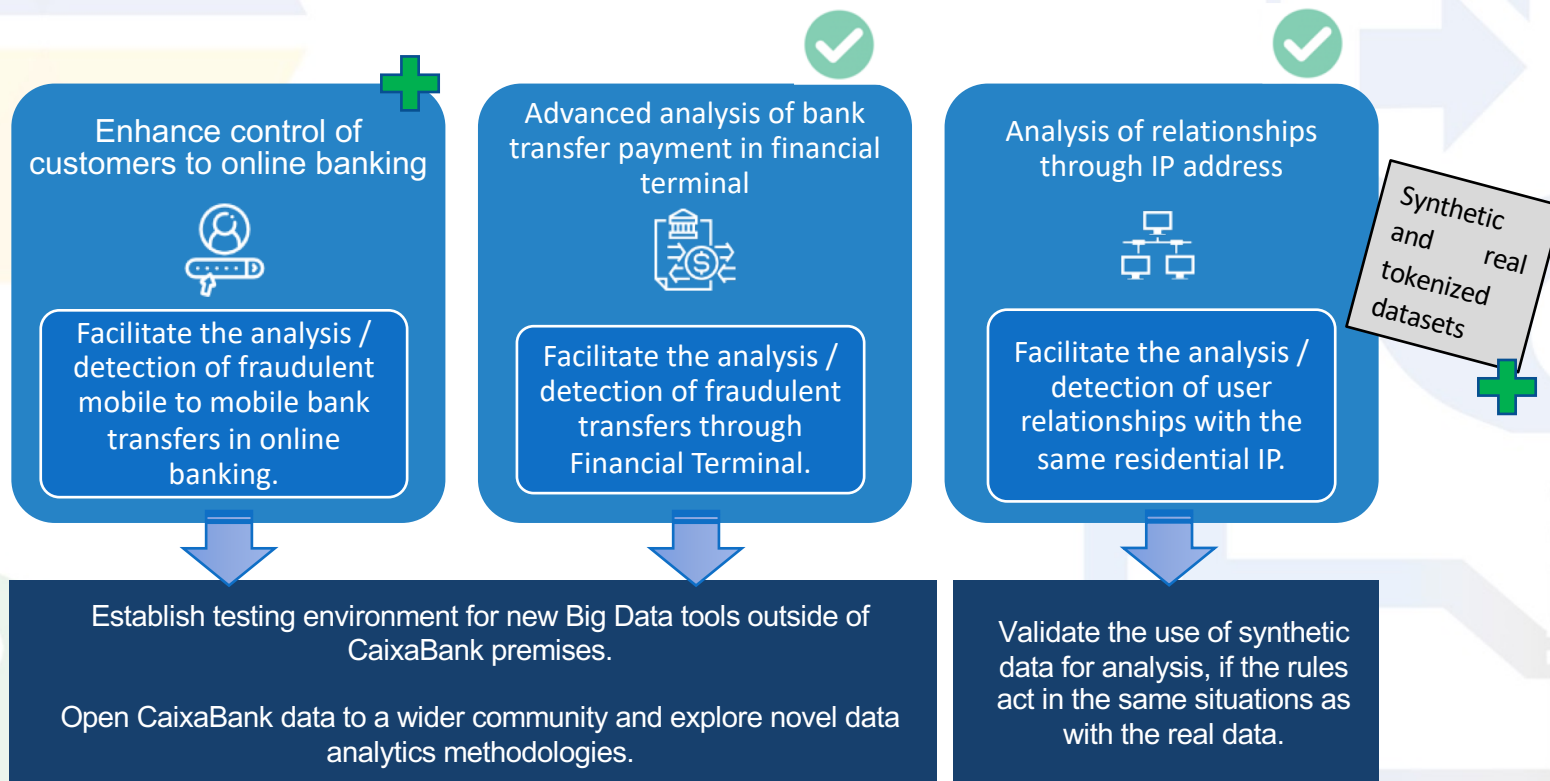
- *The possibility to work with real data outside CaixaBank in a secure way.*
- *We moved from totally synthetic approach to tokenized real datasets.*
- *We include a comparison between synthetic and real data to know better the differences*



## Initial Data & use cases



## Final Data & use cases





CAIXA 1st case

MVP use case



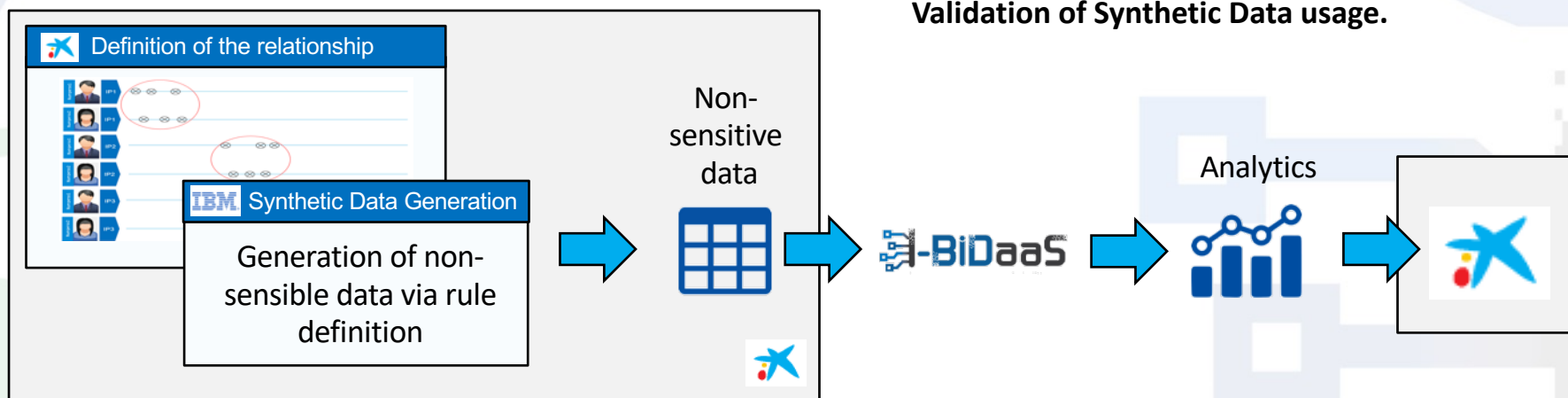
Use Case	I-BiDaaS dataset	Data
Enhance control of customers to online banking	Online banking connections	Real tokenized
Advanced analysis of bank transfer payment in financial terminal	Bank transfers	Real tokenized
Analysis of relationships through IP address	IP address	Synthetic / real tokenized

■ **Business Goal:**

The transaction between two people related by an IP is not Fraud.

■ **Use Case Goal:**

Validation of Synthetic Data usage.



CAIXA 1st case

MVP use case



## EXPERT MODE



Experiment with your own code

Upload your own code (based on pre-defined code templates) and make full use of I-BiDaaS data processing pipeline

## SELF-SERVICE MODE



Experiment with predefined algorithms

Select an algorithm from a pool of available algorithmic implementations and construct a Big Data processing pipeline

Which algorithm should you select?

Please answer the following questions to get a recommendation on which algorithm to use based on your data

Start

## CO-DEVELOP MODE



Customised industrial use cases

Have a look at tailor-made end to end implementations of the I-BiDaaS pipeline for specific industrial use cases in the fields of Banking, Manufacturing and Telecommunications.

CAIXA 1st case

MVP use case



Home

Expert Mode

Self-Service Mode

Co-Develop Mode

Administrator Admin



## I-BiDaaS Use Cases

### IP Address Relations - Batch processing



The CAIXA IP Addresses Use Case is an application of the I-BiDaaS ideas and tools extracting relations between IP addresses. In this use case batch processing of big data emerges as an innovative tool that can be instrumental in fraud detection.

Status: open Upd: 2018-12-12 10:33:38

### IP Address Relations - Stream processing



The CAIXA IP Addresses Use Case is an application of the I-BiDaaS ideas and tools extracting relations between IP addresses. In this use case batch processing of big data emerges as an innovative tool that can be instrumental in fraud detection.

Status: open Upd: 2018-12-12 12:37:52

### Centro Ricerche FIAT Aluminium Casting (Training)



The use case is concerned with the production process of aluminium casting. The goal is to use Big Data for improving the quality of the production process and operational efficiency, in particular, the quality issues on the automotive component

Status: open Upd: 2019-02-26 09:23:03

### Telefonica CC Sentiment - Stream Processing



The use case is concerned with the sentiment of callers to Telefonica's call centres in Madrid, Barcellona, and Seville.

Status: open Upd: 2019-02-26 09:23:03

Which algorithm should you select?

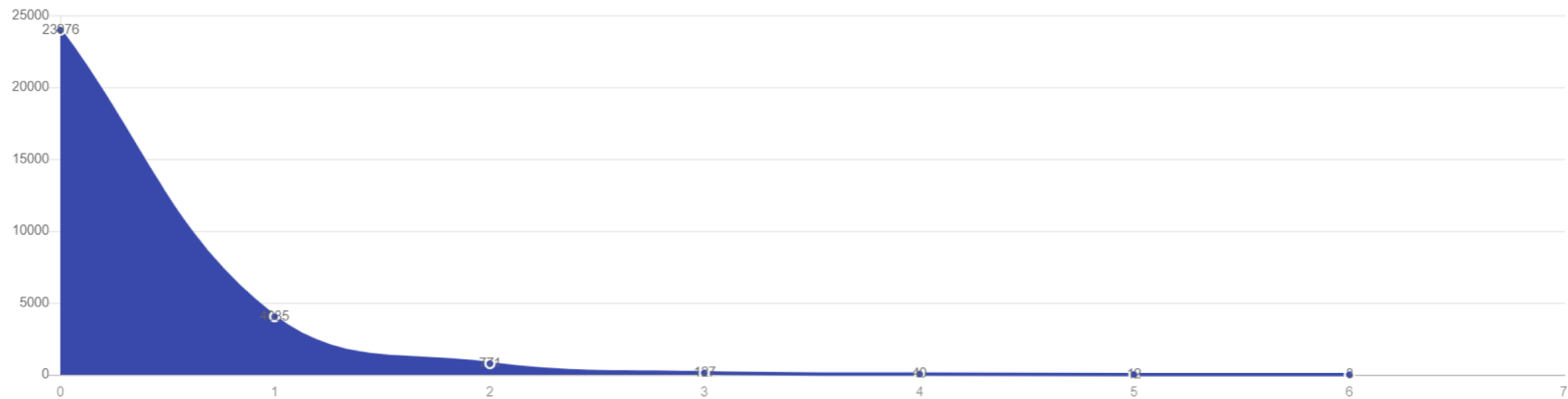
CAIXA 1st case

MVP use case



People Per Relationships

Today



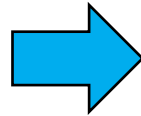
Description:  
Download URL:

BACK

DOWNLOAD FILE

CAIXA 1st case

MVP use case



	A	B	C	D
1	Client_ID_1	Client_ID_2		
2	19440	8810		
3	19440	81427		
4	99560	4423		
5	86814	77216		
6	4897	69743		
7	93297	56347		
8	4659	69543		
9	50790	5677		
10	50790	20044		
11	50790	20548		
12	50790	28930		
13	50790	34807		
14	73358	29298		
15	65647	36884		
16	47305	29446		
17	90833	15211		
18	35403	74935		
19	76443	97593		
20	42385	55955		
21	94502	44981		
22	8959	49501		
23	26040	8404		
24	44281	48842		
25	93004	1841		
26	82004	62775		



CAIXA 1st case

MVP use case



Home

Expert Mode

Self-Service Mode

Co-Develop Mode

Administrator Admin ▾



Run Experiment

### IP Address Relations - Stream processing

Date	↑	Description	Type	UserId 1	UserId 2
15/05/2020 04:27:44		Found related user pair.	Info	12348	23452349
15/05/2020 04:27:48		Found related user pair.	Info	34523454	21523454
15/05/2020 04:27:49		Found related user pair.	Info	34523456	21523456
15/05/2020 04:27:51		Found related user pair.	Info	23452349	12348
15/05/2020 04:28:27		Found related user pair.	Info	3564356	123123

Items per page: 5 ▾

6 - 10 of 10



CAIXA 1st case

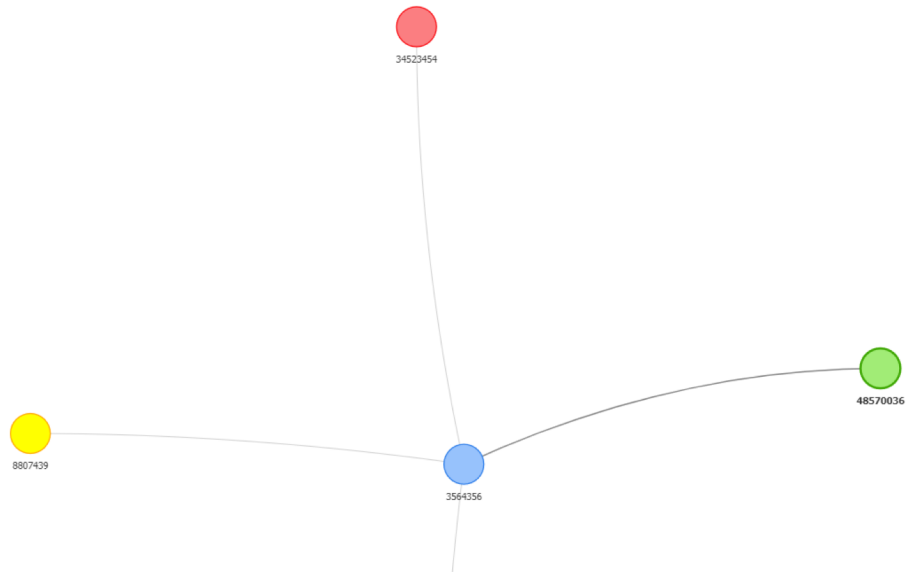
MVP use case



### Network Graph

Select Depth Level

Apply Depth Level

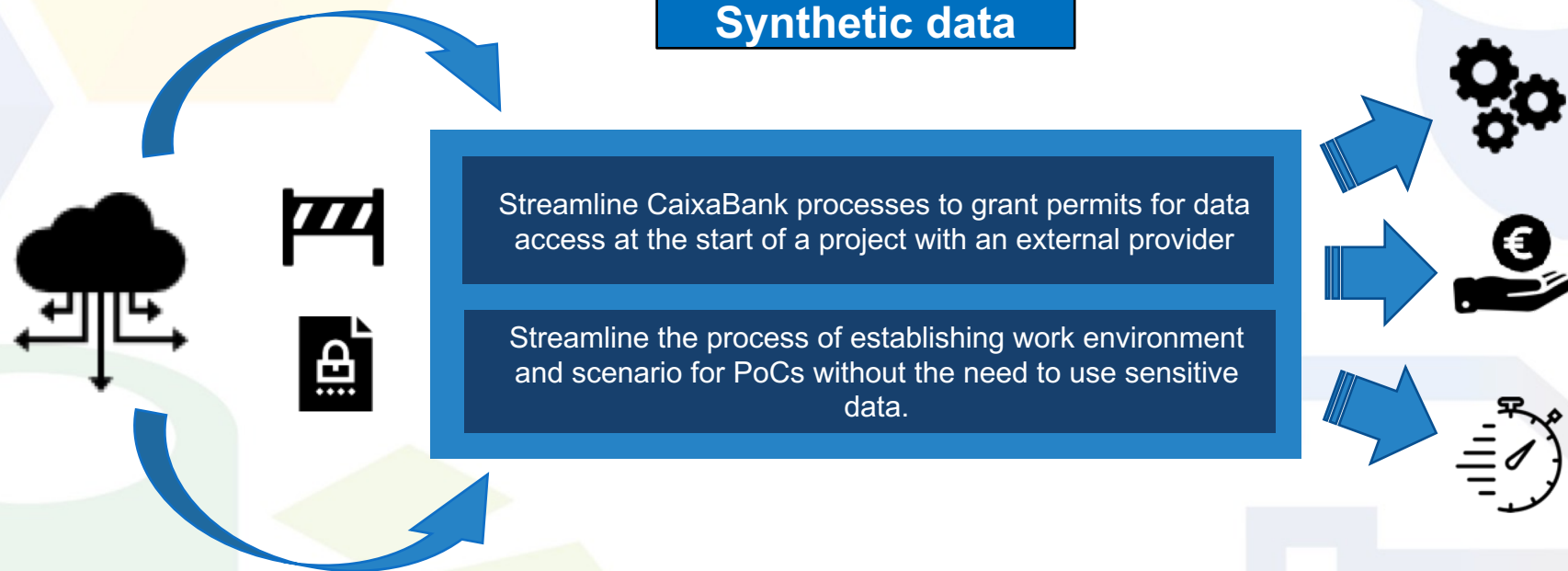


## How I-BiDaaS is helping us?

### Synthetic data

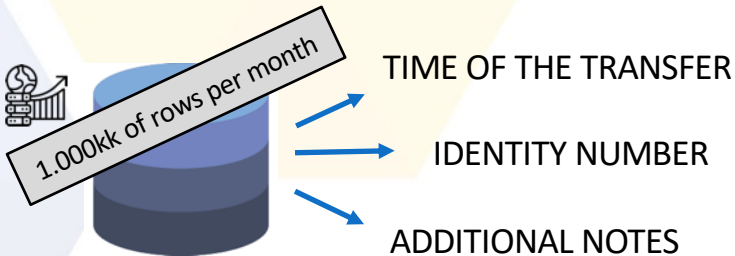
Streamline CaixaBank processes to grant permits for data access at the start of a project with an external provider

Streamline the process of establishing work environment and scenario for PoCs without the need to use sensitive data.



**CAIXA 2nd case**

Analyse bank transfers executed from employees financial terminal in the name of a customer.  
Potential fraudulent transfer or bad practices (e.g. check that the client was present in the time of the movement.)



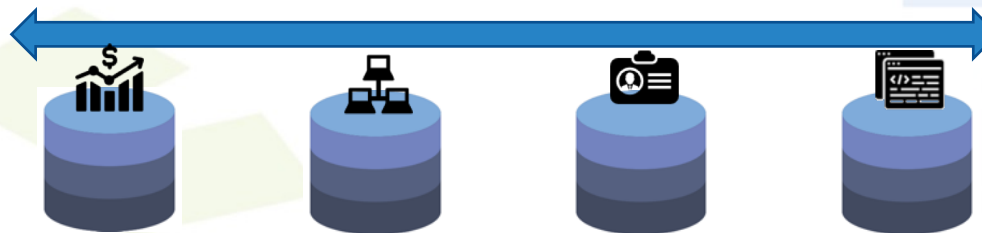
**Ensure the security of our data: Decide what are we going to share clear, tokenized or encrypted.**

The challenge relies on finding the limit of **what and how real data can be shared** to comply to regulation and not lose additional and valuable information for analytics.

**Big Data Analytics objective:** Discover fraudulent scenarios from our data by analysing the presence or not of the client.

**Use case high-level objective: Breaking internal and external silos**

Financial Operation Data, Security Management (SIEM), etc.



CAIXA 2nd case

Analyse bank transfers executed from employees financial terminal in the name of a customer.  
Potential fraudulent transfer or bad practices (e.g. check that the client was present in the time of the movement.)

CONTEXT

Point of view:




- Final User



- Programmer



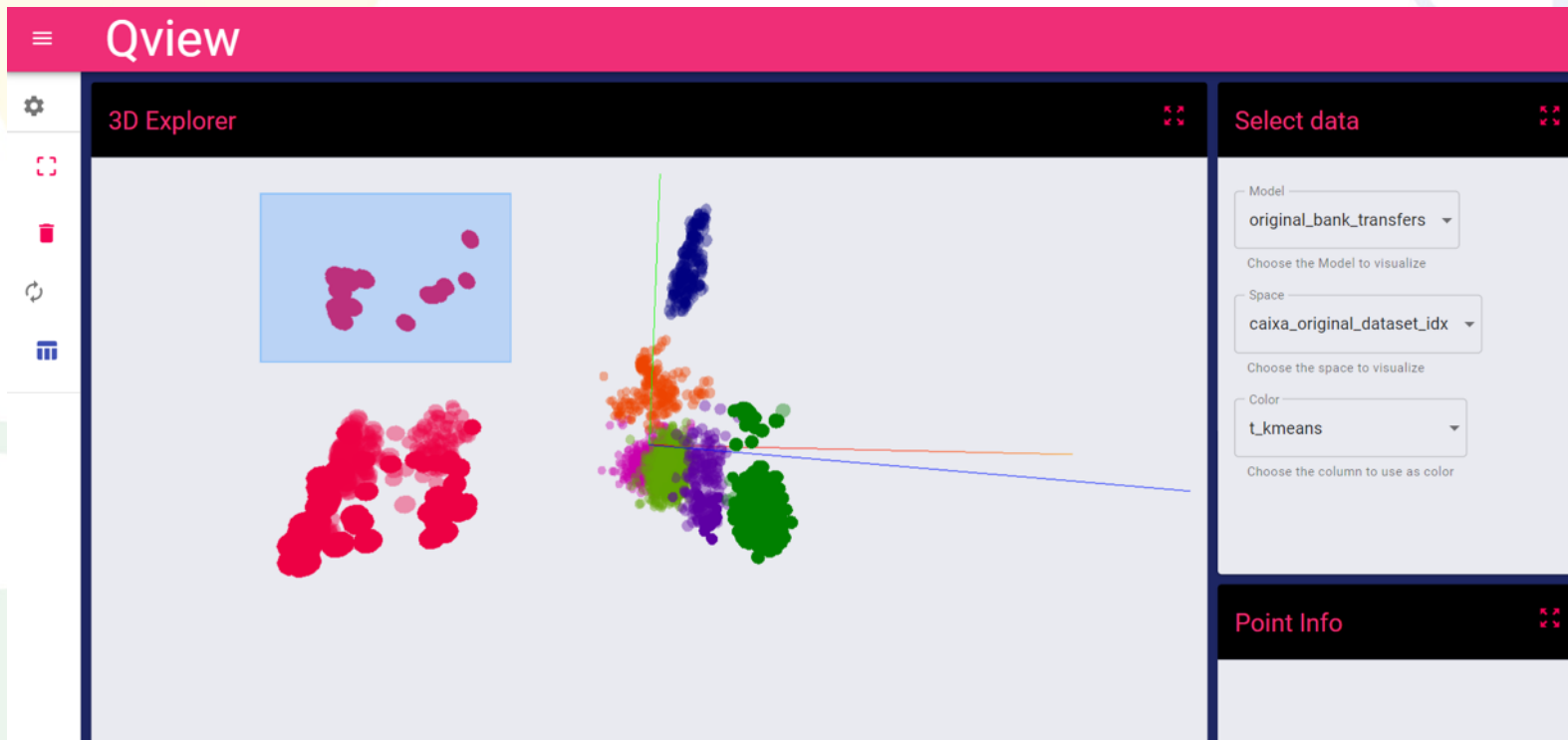
OBJECTIVES

- Identify and glue events, followed by enriching the transfer payment dataset 
- Encrypt the data without lossing value 
- Advanced analytics 



CAIXA 2nd case

Analyse bank transfers executed from employees financial terminal in the name of a customer.  
Potential fraudulent transfer or bad practices (e.g. check that the client was present in the time of the movement.)



CAIXA 2nd case

Analyse bank transfers executed from employees financial terminal in the name of a customer.  
Potential fraudulent transfer or bad practices (e.g. check that the client was present in the time of the movement.)

**Qview**

**Data Table**

Filter Rows

apellido1_emplea	apellido2_emplea	cod_bloque_o_ed	cod_categoria_ce	cod_festividad	cod_sexo	cod_tipo_persona	cod_tipo_via	contrato_des
[17, 62, 87, 5]	[3, 14, 32, 34, 87]		[None]	[S 'D' 'S' 'D' 'E' 'S]				4718685891
[4, 28, 52, 60]	[0, 41, 47, 53, 58]		[None]	[S 'D' 'S' 'D' 'S' 'D]				[-1]
[1, 19, 23, 40]	[4, 7, 9, 11, 13, 11]		[None]	[S 'D' 'S' 'D' 'S' 'D]				[-1]
[1, 5, 9, 13, 2]	[4, 21, 43, 53, 61]		[None]	[S 'D' 'S' 'D' 'E' 'S]				[-1]
[58, 60, 63, 8]	[5, 17, 20, 62, 82]		[None]	[S 'D' 'S' 'D' 'S' 'D]				[-1]
[58, 60, 63, 8]	[5, 17, 20, 62, 82]		[None]	[S 'D' 'S' 'D' 'S' 'D]				2705684373
[2, 3, 22, 36]	[5, 45, 60, 63, 87]		[None]	[S 'D' 'S' 'D' 'E' 'S]				[-1]
[2, 3, 22, 36]	[5, 45, 60, 63, 87]		[None]	[S 'D' 'S' 'D' 'S' 'D]				9930688468
[58, 60, 63, 8]	[5, 17, 20, 62, 82]		[None]	[S 'D' 'S' 'D' 'S' 'D]				[-1]
[2, 3, 22, 36]	[5, 45, 60, 63, 87]		[None]	[S 'D' 'S' 'D' 'E' 'S]				[-1]
[1, 5, 9, 13, 2]	[4, 21, 43, 53, 61]		[None]	[S 'D' 'S' 'D' 'E' 'S]				[-1]
[58, 60, 63, 8]	[5, 17, 20, 62, 82]		[None]	[S 'D' 'S' 'D' 'S' 'D]				0104688528
[4, 28, 52, 60]	[0, 41, 47, 53, 58]		[None]	[S 'D' 'S' 'D' 'S' 'D]				[-1]

CAIXA 2nd case

Analyse bank transfers executed from employees financial terminal in the name of a customer.  
Potential fraudulent transfer or bad practices (e.g. check that the client was present in the time of the movement.)

Comparing solutions and processes to analyse data outside CaixaBank premises

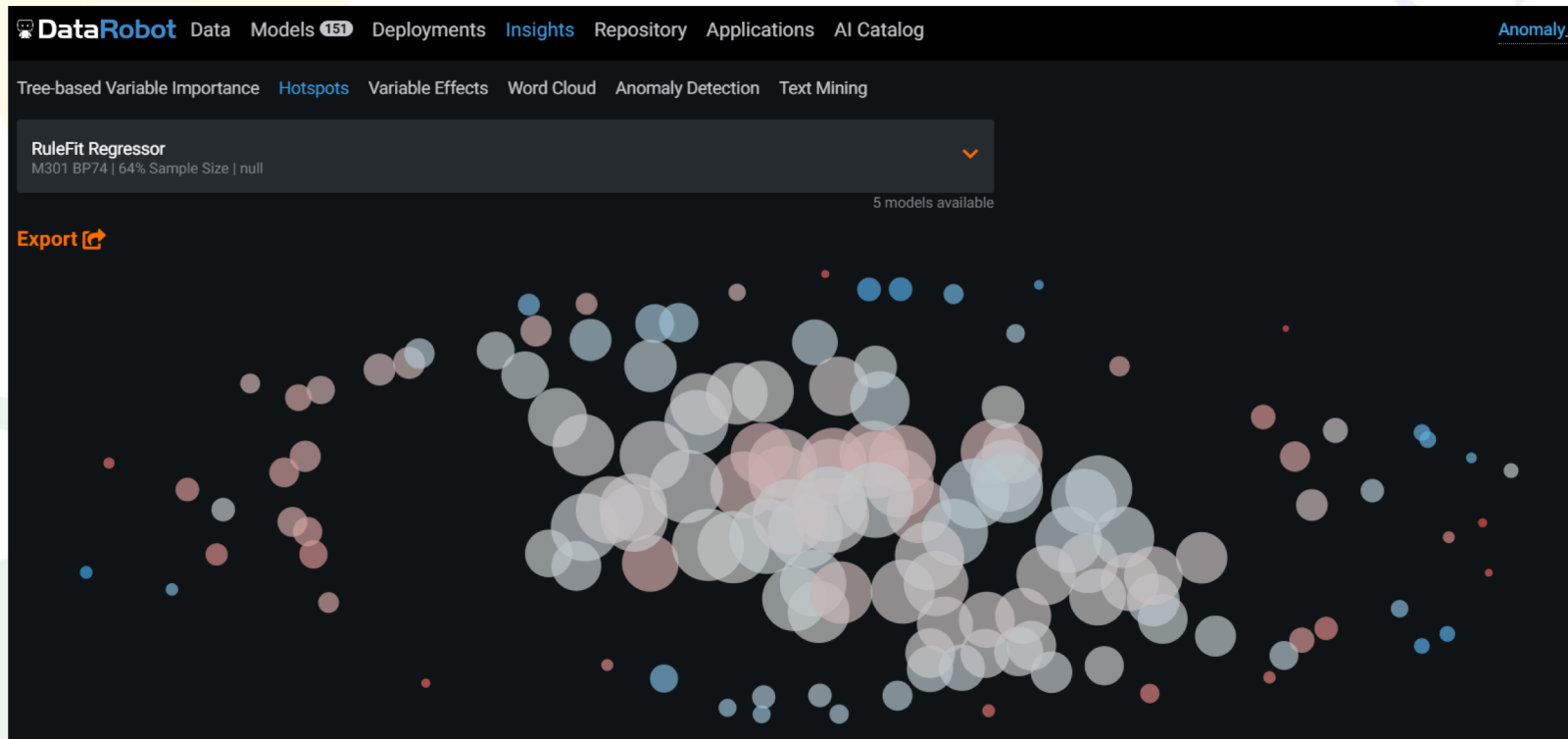


Data Analytics  
comercial products



CAIXA 2nd case

Analyse bank transfers executed from employees financial terminal in the name of a customer.  
Potential fraudulent transfer or bad practices (e.g. check that the client was present in the time of the movement.)



CAIXA 2nd case

Analyse bank transfers executed from employees financial terminal in the name of a customer.  
Potential fraudulent transfer or bad practices (e.g. check that the client was present in the time of the movement.)

**DataRobot** Data Models **151** Deployments Insights Repository Applications AI Catalog Anomaly

Tree-based Variable Importance Hotspots Variable Effects Word Cloud **Anomaly Detection** Text Mining

Isolation Forest Anomaly Detection  
M43 BP97 | 16% Sample Size | null

2 models available

**Columns (43)** [Export](#)

anomalyScore	ID	FK_CENTRO_AP	FK_CONTRATO_PPAL_OPE	FK_EMPLEADO	FK_EMPLEADO_AUT	FK_IMPORTE_PRINCIPAL	FK_NUMPERSO_PRINCIPAL	IDE_NIVEL_PAD
1	1789	5790	['270568800517']	K4970981	K1413164	3500000	8664854	3
0.9896	676	5790	['331368286280']	K5472414	K2924060	9844365.93	8756826	3
0.9787	1362	5790	['270568738363']	K4690877	K4690877	507670.12	696397	1



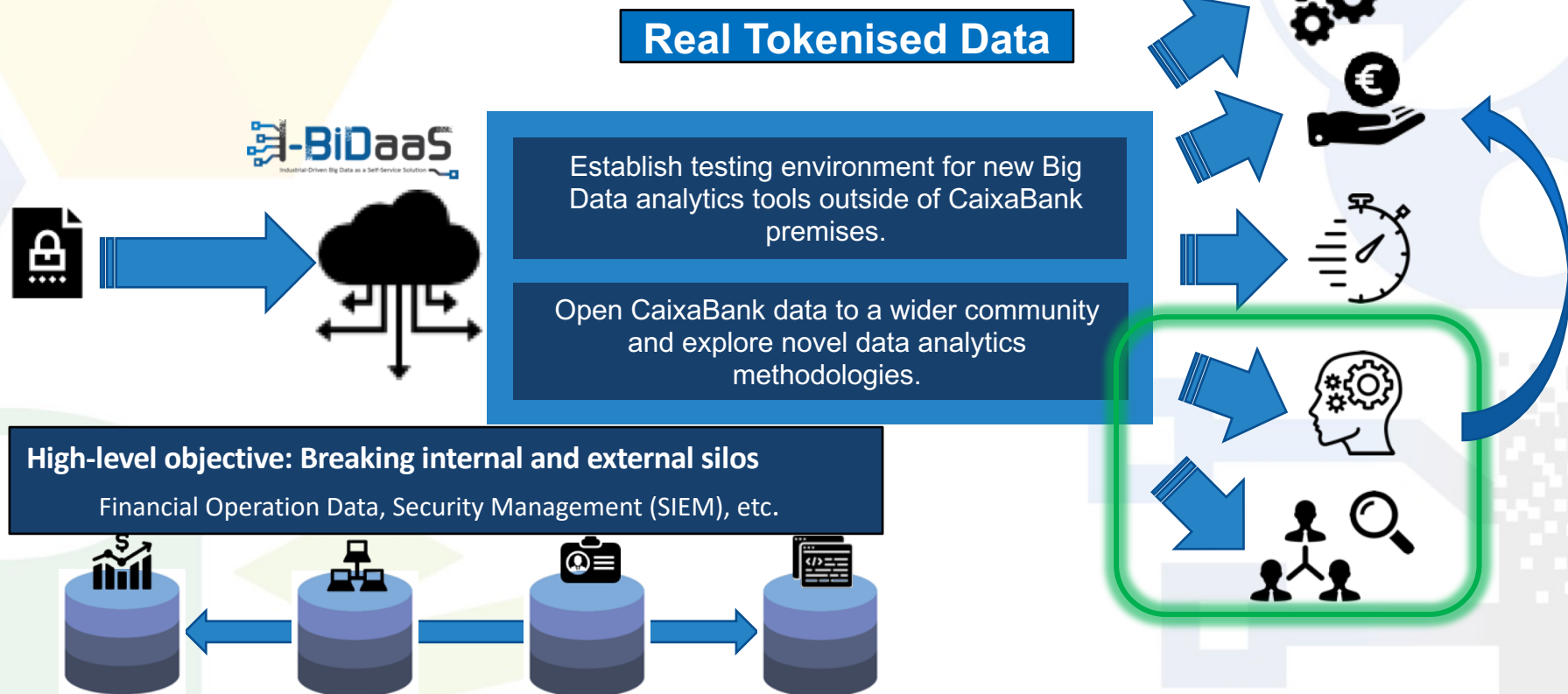
**CAIXA 2nd case**

Analyse bank transfers executed from employees financial terminal in the name of a customer.  
Potential fraudulent transfer or bad practices (e.g. check that the client was present in the time of the movement.)

## DataRobot comparison results

- **Custom solutions:**
  - I-BiDaaS provide more flexibility in the definition of your own code, scoring metrics, etc.
- **Unsupervised learning:**
  - DataRobot has very limited number of unsupervised learning models. I-BiDaaS can provide much more detailed results on unsupervised learning use cases.

## How I-BiDaaS is helping us?



## CaixaBank benefits from I-BiDaaS

Benefits	KPIs
To increase the efficiency and competitiveness in the management of its vast and complex amounts of data.	75% time reduction data access from external stakeholders using synthetic data (From 6 to 1.5 days).
To break data silos not only internally, but also fostering and triggering internal procedures to open data to external stakeholders.	Real data accessed by at least 6 different external entities skipping long-time data access procedures.
To evaluate Big Data analytics tools with real-life use cases of CaixaBank in a much more agile way.	I-BiDaaS overall solution and tools experimentation with 3 different industrial use cases with real data.



bdva.eu  
bigdatastack.eu  
lbidaas.eu  
trackandknowproject.eu



@BDVA\_PPP  
@BigDataStackEU  
@lbidaas  
@Track&Know

# Questions?

## Thank you!

Your feedback is valuable for us!

<https://bit.ly/2Zl19lJ>



The below-mentioned project have been funded by the European Commission Horizon 2020  
BigDataStack: grant agreement No 770747  
I-BiDaaS: grant agreement No 780787  
Track and Know: grant agreement No 780754

