INFINITECH - Pilots Overview
BDVA Event - 7th May

Jose Gato Luis
Associated Head of AI & Robotics
ATOS Research and Innovation

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Categories

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Pilots by categories

- Smart, Reliable and Accurate Risk and Scoring Assessment Pilots
  - 2 pilots
- Personalized Retail and Investment Banking Services
  - 5 pilots

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Pilots by categories

- Financial Crime and Fraud Detection
  - 4 pilots
-

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Pilots by categories

- Personalized Usage Based Insurance Products
  - 2 pilots
- Configurable and Personalized Insurance Products
  - 2 pilots

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Personalized Usage Based Insurance Products

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Categ. Personalized Usage-Based Insurance Pilots

Pilto #11 Personalized insurance products based on IoT connected vehicles

Description	Partners

Improve the risk insurance profiles using the information collected by connected vehicles and applying IoT, HPC, Cloud Computing and Artificial Intelligence technologies	• ATOS
	  ○ Connected Car Platform
	  ○ IA Platform
	• CTAG
	  ○ On board units
	  ○ Real-Time/Historical data
	  ○ 80 vehicles during 4 by day
	• Gradient
	  ○ Anonymization Service
	• Dynamis
	  ○ Requirements
	  ○ insurance company's data

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Infinitech

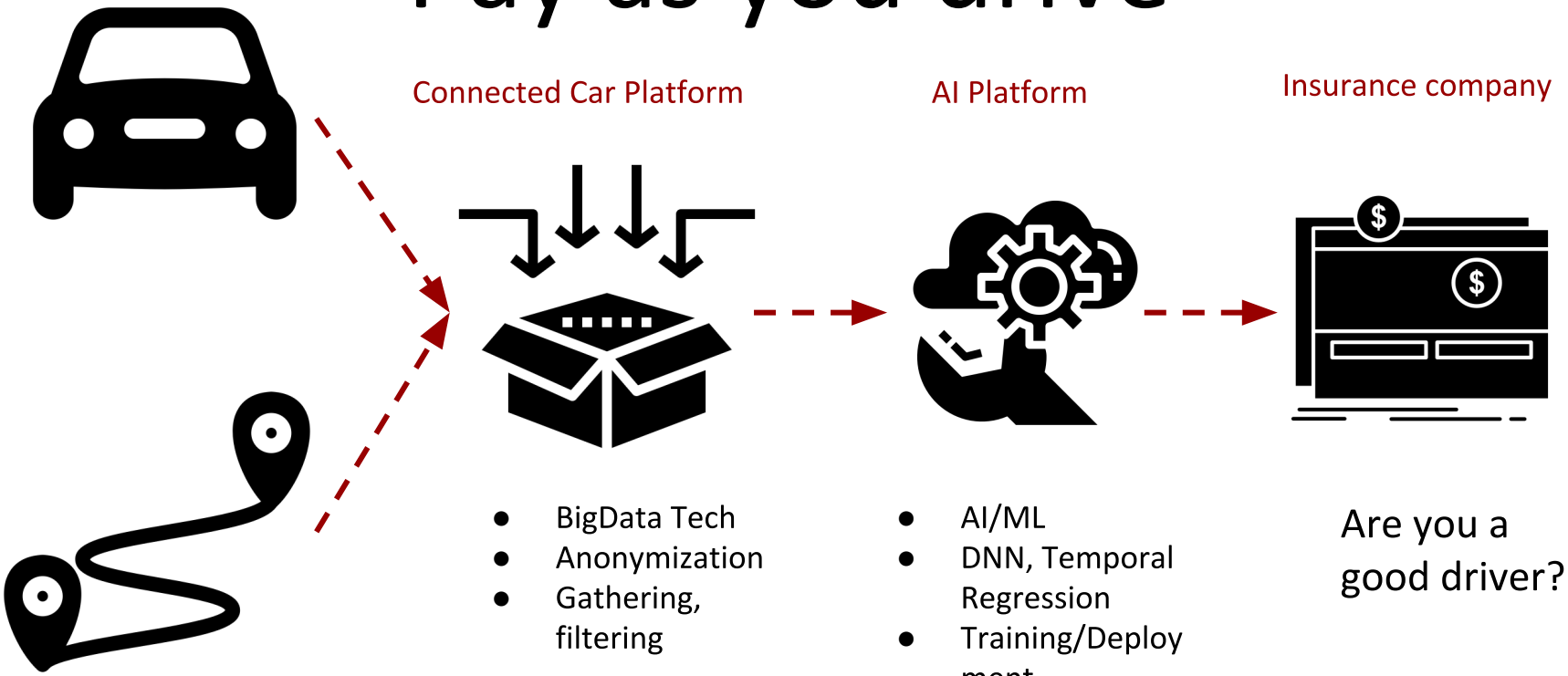
Business Services

Fraud Detection Service	
Description	An End-User service that combines and exploits data from the MLaaS framework, Insurance DB and diverse context information to detect possible frauds on reported events and/or during the daily use of an insurance vehicle.
Owner	ATOS/DYN
BDVA Layer	User Interaction/Visualization
Input	Data Analytics & Data Processing architectures: EASIER.AI Framework, ATOS Connected Vehicles Framework and Data Replicas with context information and insurance datasets
Output	End-User dashboards

Pay As You Drive Service	
Description	An End-User service that combines and exploits data from the MLaaS framework, Insurance DB and diverse context information to analyse and compare different Drivers Profiles and Classifications to elaborate the best, adapted and personalised car insurance offerings and bonus.
Owner	ATOS/DYN
BDVA Layer	User Interaction/Visualization
Input	Data Analytics & Data Processing architectures: EASIER.AI Framework, ATOS Connected Vehicles Framework and Data Replicas with context information and insurance datasets
Output	End-User dashboards

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Pay as you drive



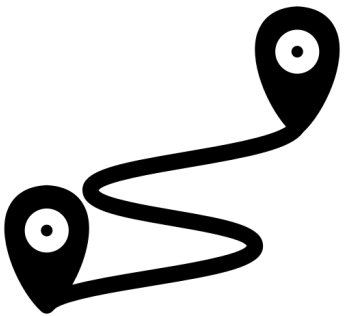
- BigData Tech
- Anonymization
- Gathering, filtering

- AI/ML
- DNN, Temporal Regression
- Training/Deploy ment

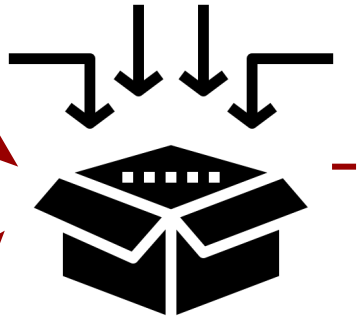
Are you a good driver??

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Fraud Detection

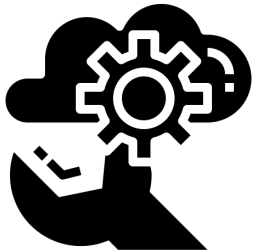


Connected Car Platform



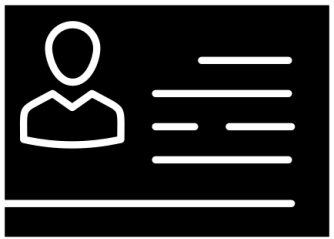
- BigData Tech
- Anonymization
- Gathering, filtering

AI Platform



- AI/ML
- Clustering Model
- Training/Deployment

Insurance company



Policy Agreements

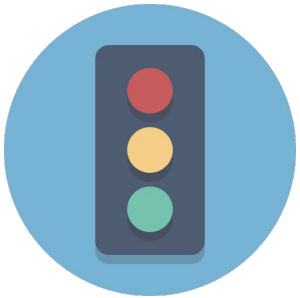
Different drivers

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Main Data Sources

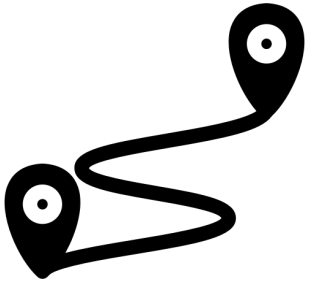
CAN/BUS
80 Vehicles 4h/day
Pre-Historical data

~600 GB



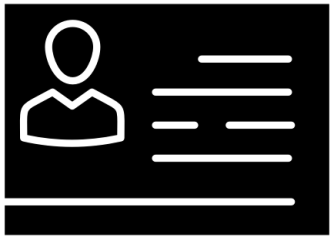
City of Vigo
Traffic Events

~1 TB



Different cities
simulated data

on demand

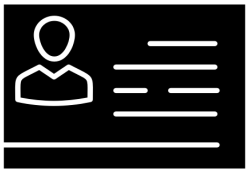


Insurance company data

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Real Life Challenges

Insurance company

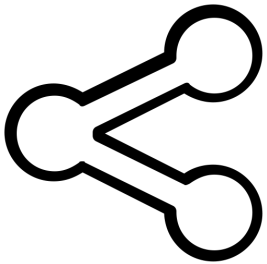


New adapted services

Bad drivers and fraud costs

No data from IoT/Connected

Data providers

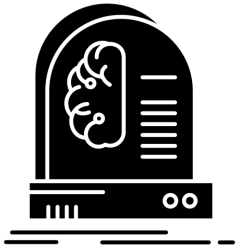


Quality/available data

Combine data sources

Standards protocols

Data intelligence



How to extract intelligence?

What does “good driver” means?

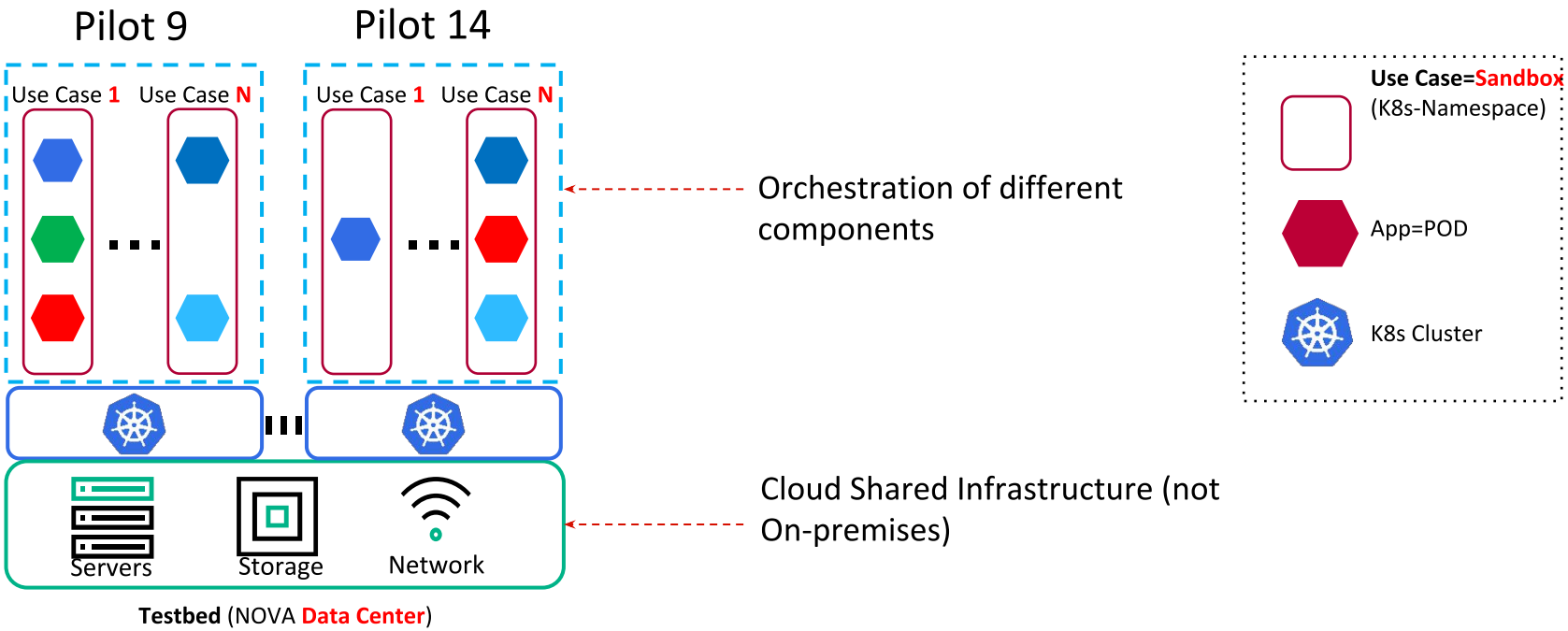
Complexity combination of technologies

Privacy

This project has received funding from the European Union’s horizon 2020 research and innovation programme under grant agreement no 856632

Infinitech

Sandbox and test bed



This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Testbeds	Datasets (external cloud support)		
Connected Vehicle Traffic monitoring framework	Simulated Urban Mobility Dataset	Estimated 368 GB	Data collected from vehicle's CAN Bus (80 vehicles driving 4 h/day 1 year). Historical data comming from exisiting deployments
	Traffic events + Location	Estimated 900 GB	Traffic events published by the city of Vigo and DGT (Historical data related to captured CAN Data)
	NMEA Data for vehicles (Historical)	Estimated 120 GB	Complementary location (GPS, Timestamp, speed, heading...) for Vehicles' CAN Data (Historical data related to captured CAN Data)
	CAN Signals (Live)	Estimated 150 GB	CAN data + Driving dtyle info (revolutions, gear, hard breaking...)+ Parking (close dorrs, windows...) + Maintenance
	Traffic Events (Live)	Estimated 250 GB	Traffic events published by the city of Vigo and DGT
	NMEA Data for vehicles (Livel)	Estimated 50 GB	Complementary location (GPS, Timestamp, speed, heading...) for Vehicles' CAN Signal
SUMO Traffic simulations	Simulated Urban Mobility Dataset	On demand	Simulated Urban mobility data (mainly vehicles CAN Signals)

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Anonymizer	
Description	The anonymization tool modifies data in order to preserve privacy. It is especially indicated in those cases where a dataset contains personal data and it has to be outsourced or shared with a third party. The tool includes different anonymization algorithms that aim at avoiding the appearances of data combinations that could lead to a possible re-identification of the data subjects. It also includes a set of privacy and utility metrics that allow to measure the risk that remains after anonymizing the dataset, and the impact of the anonymization process on the quality of the data.
Owner	GRAD
BDVA Layer	Data Protection

ATOS Connected Car Framework	
Description	A FIWARE-Based framework designed to capture, process and distribute real time traffic and vehicle's information. It will implement Pub/Sub mechanisms and Geolocation and Time series tools.
Owner	ATOS
BDVA Layer	Data Processing Architecture

EASIER.AI Framework	
Description	EASIER-AI is a Hybrid (Cloud/Edge) platform that facilitates to develop, measure, monitor, deploy and update customised AI models based on Machine Learning and Deep Learning techniques. This framework makes easier the deployment and execution of data science tasks, focused on working on Hybrid Infrastructure and exploiting data generated by IoT frameworks
Owner	ATOS
BDVA Layer	Data Analytics

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Personalized Retail and Investment Banking Services

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Categ. Personalized Retail and Investment Banking Services

Pilot #5a Smart and Personalized Pocket Assistant for Personal Financial Management

Description

Smart Services for bank customers

- Smart alerts: prevent possible overdrafts
- Smart automations: identify recurrent payments
- Smart expense advisor: categories compared with other “similar” customers
- Smart recommendations of bank’s products
- Smart sentinel: protection based on alerting on potential anomalies

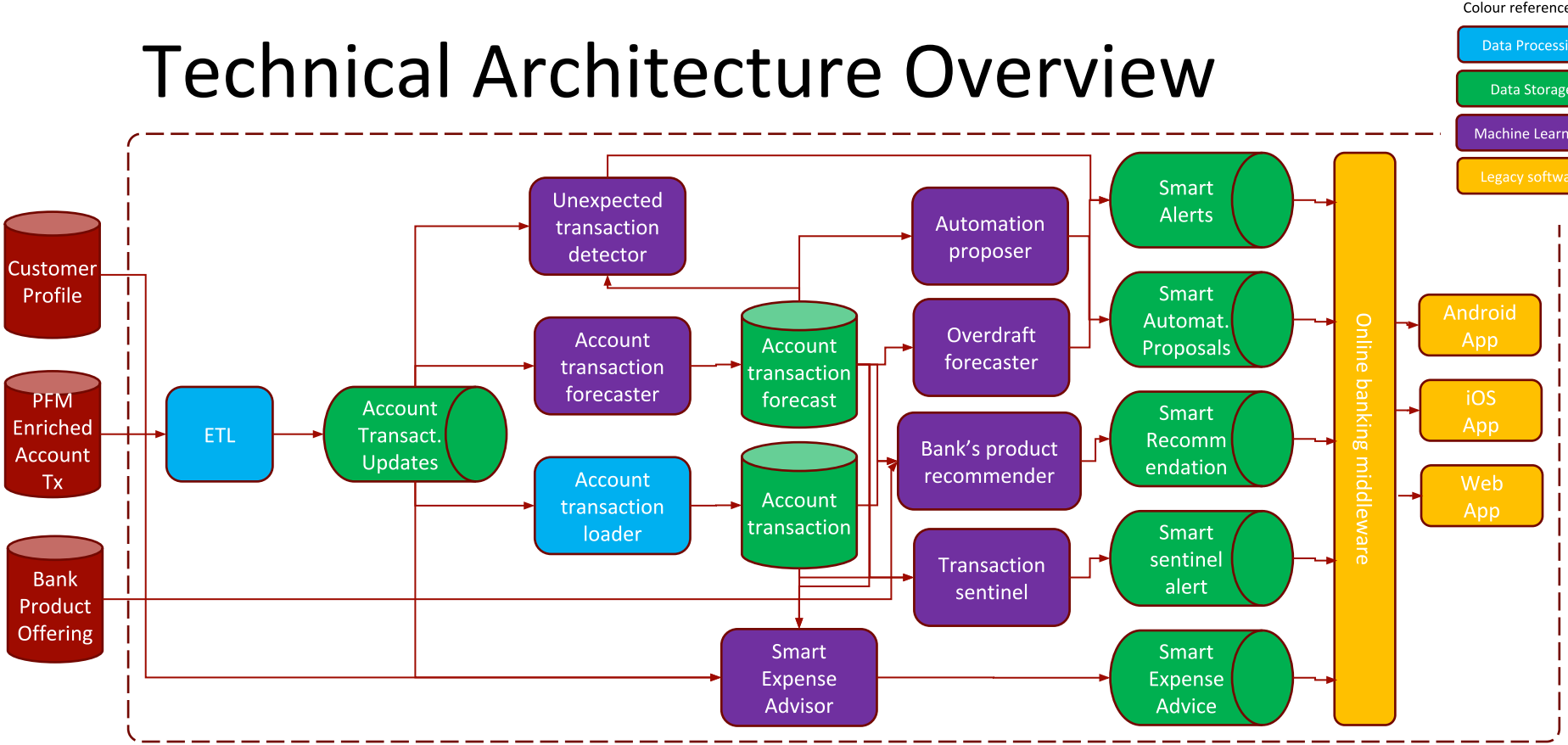
Partners

- Liberbank
  - Final User
  - Data Provider
- GFT Spain
  - Integrator of LIBERBANK
- UNIVERSITY OF PIRAEUS
  - Machine Learning developer
- CrowdPolicy
  - Machine Learning developer

This project has received funding from the European Union’s horizon 2020 research and innovation programme under grant agreement no 856632

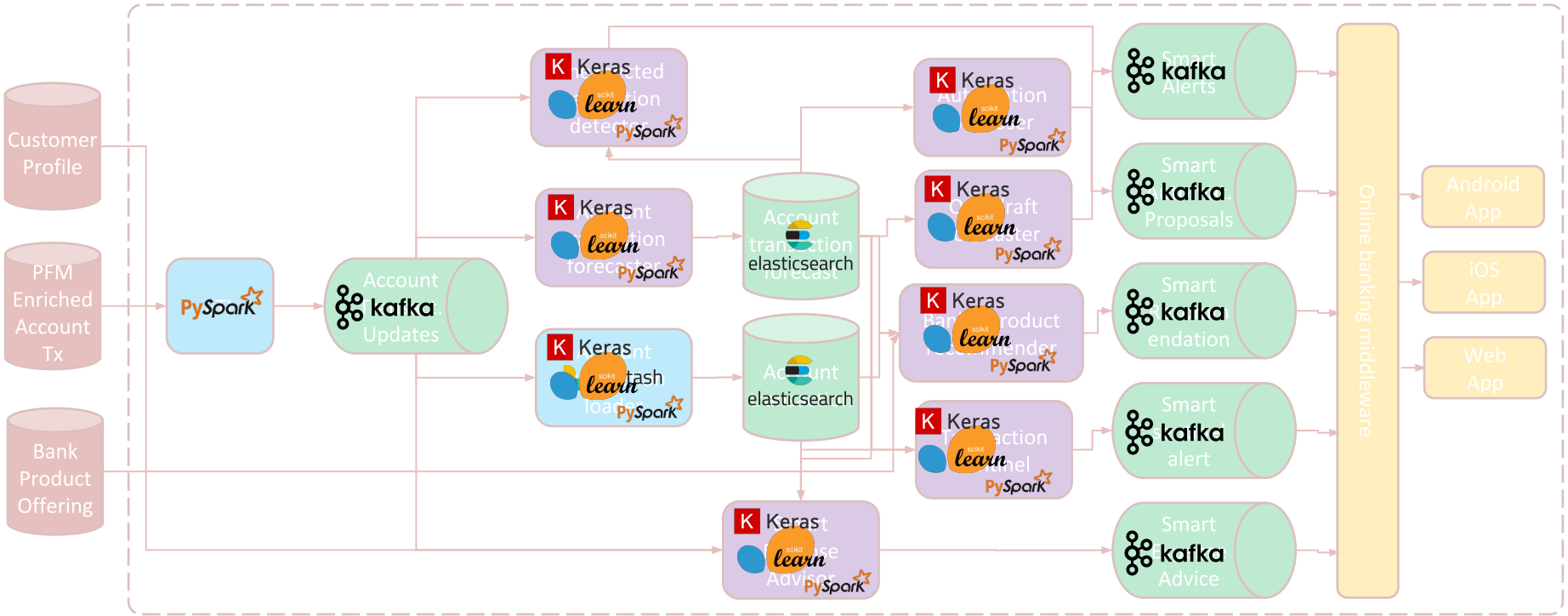
Infinitech

Technical Architecture Overview



This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Software Infrastructure Overview



This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Technologies

Datasources	
Description	• **Customer Profile :** socioeconomic data from each customer (sex, age, postal code of residence, income range, etc) • **PFM Enriched Account Transactions :** account transactions enriched by the PFM (categorisation, geoposition, etc) • **Bank Product Offering :** products and their product categories available to be offered to the customers

ETL	
Description	• ETL : online data extraction, transformation and load for PFM Enriched Account Transactions. This should be performed using the bank's change data capture infrastructure.

Queue Storage	
Description	**Message Broker:** • Account Transaction Updates : a queue topic to store processor pending Account Transaction Updates • Smart Alerts : a queue topic to store processor pending Smart Alerts • Smart Automat. Proposals : a queue topic to store processor pending Smart Automat. Proposals • Smart Recommendation : a queue topic to store processor pending Smart Recommendation • Smart sentinel alert : a queue topic to store processor pending Smart sentinel alert • Smart Expense Advice : a queue topic to store processor pending Smart Expense Advice

Data processors	
Description	• **Unexpected transaction detector :** Machine Learning data processor to identify Unexpected transaction based on transaction forecast and other rules • **Account transaction forecaster :** Machine Learning data processor to forecast future account transactions based on historic data • **Automation proposer :** Machine Learning data processor to identify recurrent operations based on historic data • **Overdraft forecaster :** Machine Learning data processor to identify possible overdraft based on transaction forecast and other rules • **Bank's product recommender :** Machine Learning data processor to recommend bank's products based on historic data and transaction forecast • **Transaction sentinel :** Machine Learning data processor to identify possible frauds

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Infinitech

Technologies

Query Model Storage	
Description	• **Account transaction forecast :** a query optimized database for serving Account transaction forecasts • **Account transactions :** a query optimized database for serving Account actual transactions

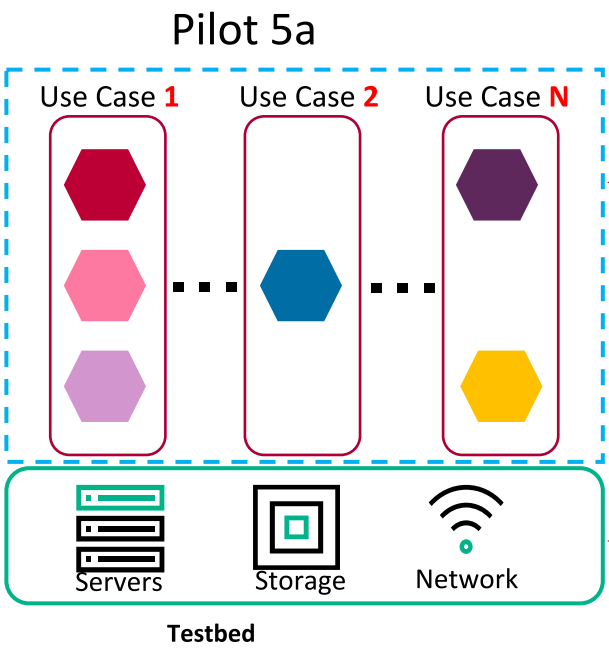
Online banking middleware	
Description	• **Online banking middleware :** responsible for engaging the insights with the customers according the designed user interaction

Apps	
Description	• **Apps :** responsible for the customer interaction with the insights

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

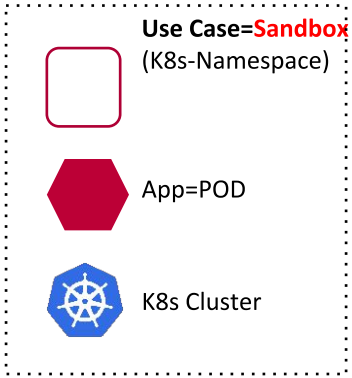
Infinitech

Sandbox and test bed



Orchestration of different components

Cloud or OnPremises (tbd). Maybe mix solution



This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Financial Crime and Fraud Detection

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Categ. Financial Crime and Fraud Detection

Pilot #9 Analyzing Blockchain Transaction Graphs for Fraudulent Activities

Description	Partners
Fraud Detection Blockchain crypto currencies and tokenized assets that are obtained fraudulently. Transactions and tokens: ● Ethereum, Bitcoin (public not regulated) ● Also regulated chains like GUSD A final transaction ends up into a bank product. Holding stable coins that originated from fraudulent. Construction of the massive blockchain transaction graph	• Aktifbank (AKTIF) ○ Responsible for user interfaces and regulations and banking services. • Bogazici Univ. (BOUN) ○ Responsible for HPC software development for big blockchain data and parallel graph analysis.

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Transaction Graph Sizes

- Transaction graph sizes are big and growing.
- Currently transactions-per-second is low on public blockchains: Bitcoin (7 tps) and Ethereum (15 tps). Ethereum performance is expected increase in future releases.
- Hyperledger reported to achieve 3500 transactions-per-second in cloud environment:

  https://www.ibm.com/blogs/research/2018/02/architecture-hyperledger-fabric/
- A parallel / distributed graph system is needed whose performance can scale by simply increasing processing nodes on an HPC cluster.

As of May 2020

Bitcoin transaction count:

527 Million

Source:
https://www.blockchain.com/charts/n-transactions-total

Ethereum transaction count:

700 Million

Source:
https://etherscan.io/chart/tx

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

HPC Requirements

- HPC Cluster with 16-32 nodes with a total of around 1TB memory is expected to handle the *current* transaction sizes.
- As the graph size increase, these requirements will increase and cluster node count and memory size can be scaled.
- HPC cluster supporting MPI (message passing interface) is needed.
- External Metis or Scotch software can be used to partition graphs in order to minimize communication volume between processors.

http://glaros.dtc.umn.edu/gkhome/metis/metis/overview

https://www.labri.fr/perso/pelegrin/scotch/

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Business Services

Public Blockchain Crypto Currency and Token Transaction Dataset and Analysis Service	
Description	Trace customer address and related linked transactions to see if they originate from blacklisted addresses published by authorities.
Owner	BOUN
BDVA Layer	Data Processing Architectures/Data Analytics
Input	Blockchain raw data
Output	Crypto currency and token transaction dataset and analysis results

Blacklisted Blockchain Address Management and Visualization of Graph Traces	
Description	Maintain blacklisted blockchain addresses lists and interact with graph analysis system and provide visualization.
Owner	AKTIF
BDVA Layer	Data Visualization and User Interaction
Input	Results from Transaction Graph Analysis Component
Output	Reports and visualization

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Technologies

Blockchain Transaction Dataset Preparation Component	
Description	Extracts Bitcoin, Ethereum and major ERC20 token transactions (such as Gemini USD (GUSD), Tether USD (USDT), Tether Gold (XAUT), Statis Euro (EURS) and Turkish BiLira (TRYB)) from blockchain.
Owner	BOUN
BDVA Layer	Data Management

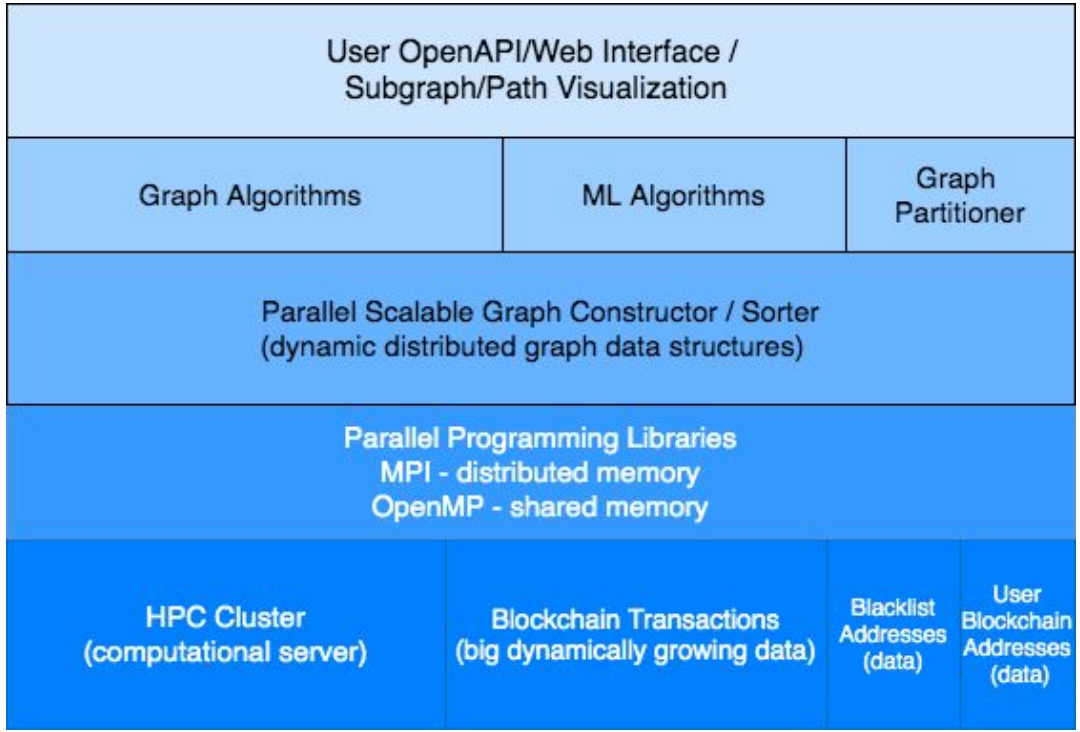
Scalable Transaction Graph Analysis Component	
Description	Constructs distributed/partitioned transaction graph in parallel using MPI. It will utilize graph and machine learning algorithms to analyse fraudulent transactions.
Owner	BOUN
BDVA Layer	Data Processing Architectures/Data Analytics

User Interface for Blockchain Transaction Reports and Visualization Component	
Description	Will provide user interaction with the Scalable Transaction Graph Analysis component within the bank and collect/manage user as well as annotated blacklisted blockchain addresses . It will utilize OpenAPIs (REST APIs) to submit queries and and provide visualization based on received results using vis.js graph drawing package
Owner	AKTIF
BDVA Layer	Data Visualization and User Interaction

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Infinitech

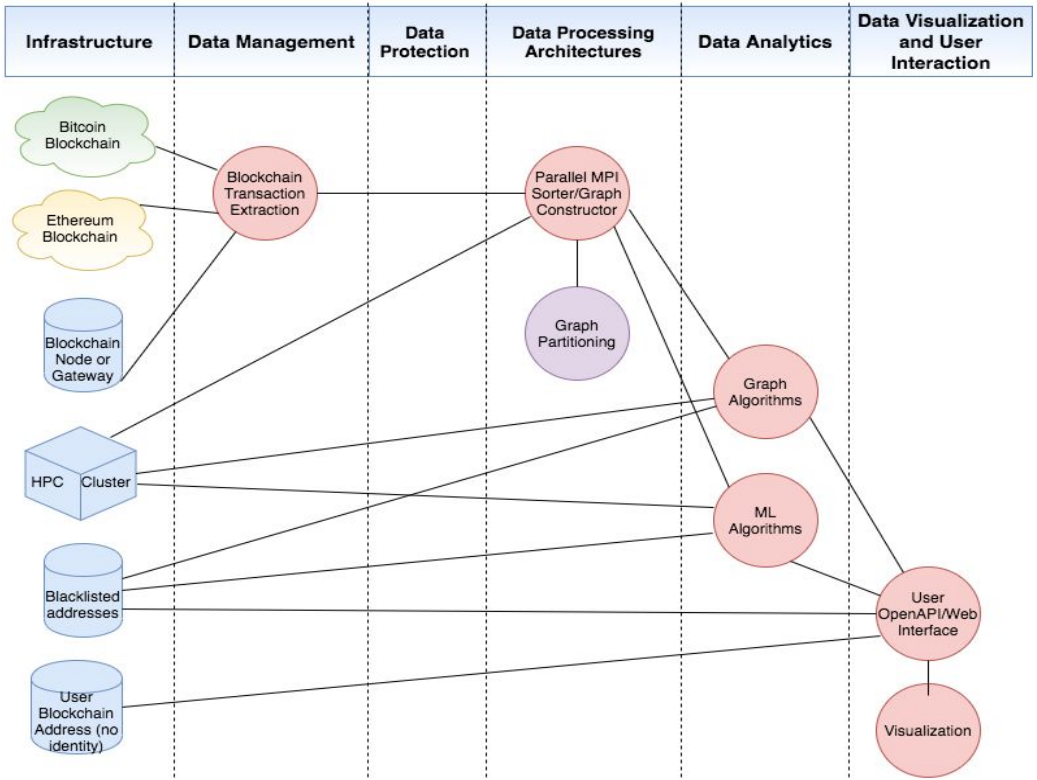
Architecture



This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Infinitech

BDVA Reference Model

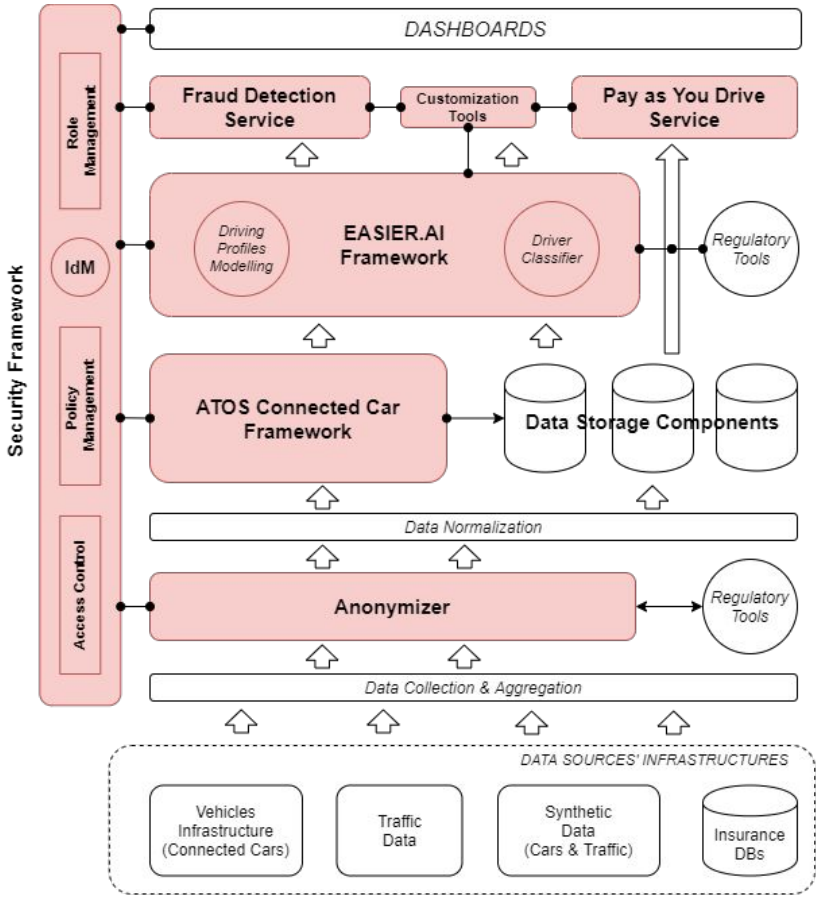


BDVA RA

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

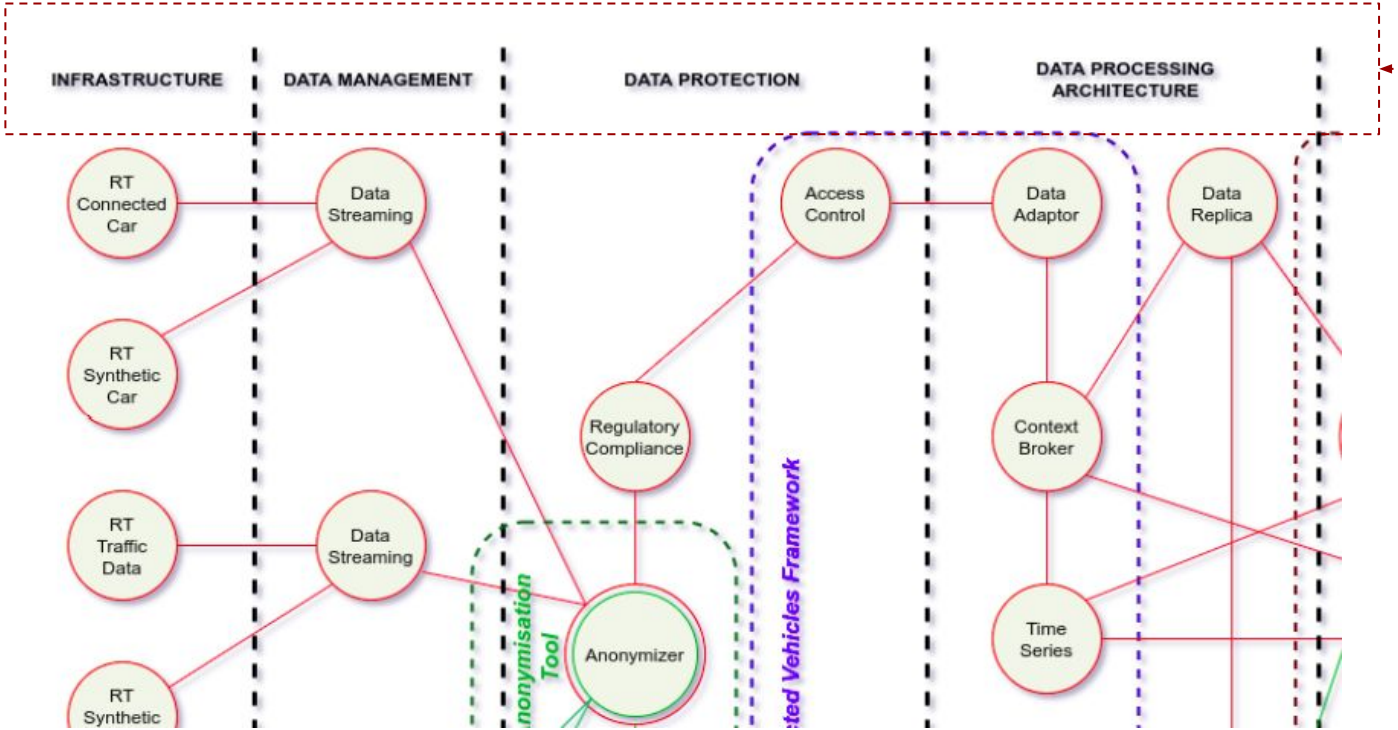
Architecture and technologies for each pilot

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632



This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Infinitech



BDVA RA

- From Generic Architecture to INFINITECH RA
- Microservices components available in the Sandbox

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Jose Gato Luis
Associated Head of AI & Robotics
ATOS Research and Innovation

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632