eit Urban Mobility

Co-funded by the European Union

# La Digitalització de la Mobilitat

Data Science i Intel·ligència artificial

Eduardo Quiñones
Barcelona Supercomputing Center (BSC)

ATM Àrea de Barcelona
Autoritat del Transport Metropolità

FACTUAL

# Introduction

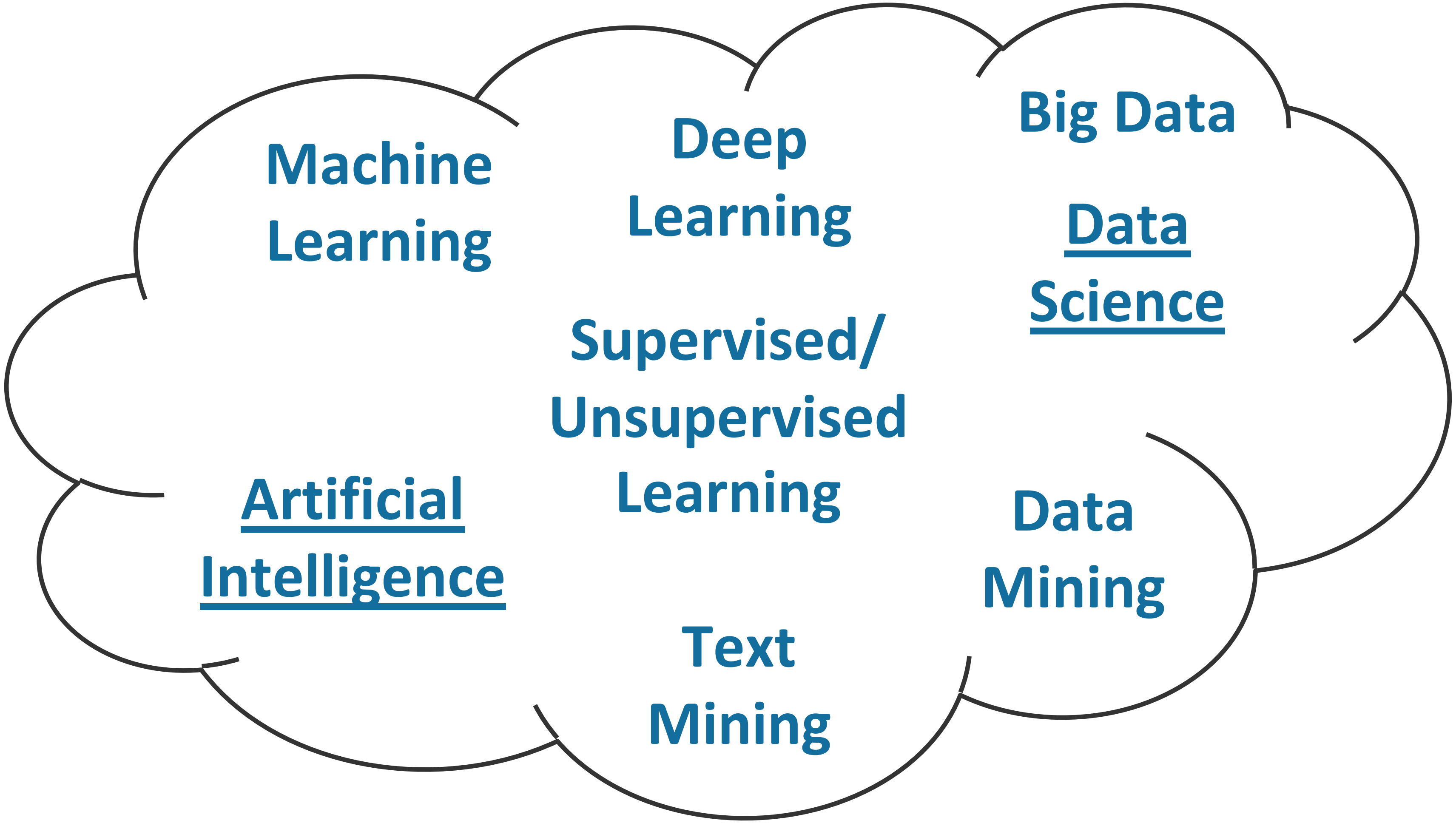
*Barcelona Supercomputing Center – Centro Nacional de Supercomputación (BSC-CNS)*

- Public research center focused on the research and efficient used of supercomputing technologies applied to science, society and economy

*Eduardo Quiñones*

- PhD in Computer Science by Technical University of Catalonia (UPC)
- Team leader of the *Predictable Parallel Computing* research group
- Principal Investigator (PI) of several European and national projects related to mobility

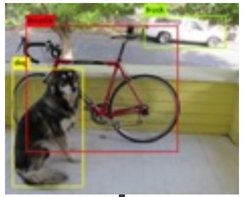
# Data Science and Artificial Intelligence (AI)



# Data Science and Artificial Intelligence (AI)

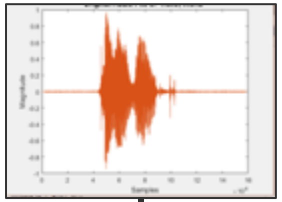

- **Data Science** focuses on the analysis of data sources across all data value chain to extract information (knowledge) upon which decisions can be taken
- **AI** focuses on providing to computers the capability to percieve, learn and reason about data to perform a specific task
  - There are not systems with *Artificial General Intelligence*

*Object recognition and detection*



**"dog", "bicycle", "truck"**

*Natural Processing Language*



**"hello, world"**

*Face Recognition*

**"Eduardo Quiñones"**


*Image Generation*


eit Urban Mobility
Co-funded by the European Union
ATM Àrea de Barcelona Autoritat del Transport Metropolità
FACTUAL

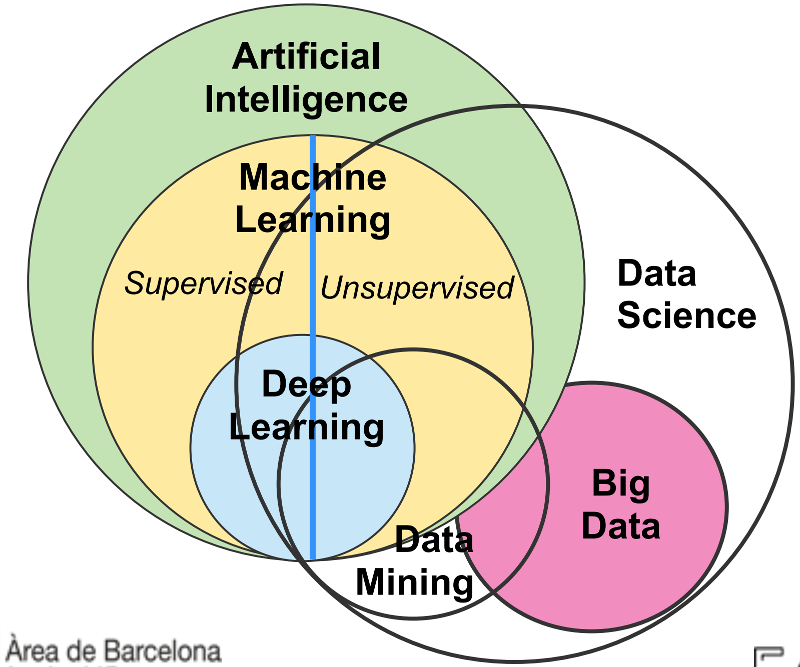
# Data Science and Artificial Intelligence

## Data Science

- *Data Mining:* Set of data analytics methods for **knoweldge extraction**
  - Visualization, data cleaning, AI model training
- *Big Data:* Set of storage and computation technologies used to process huge amounts of data

## Artificial Intelligence

- *Machine Learning*
- *Supervised/Unsupervised Learning*
- *Deep Learning*



# Objective of the Seminar

1. Understand the AI methods that can be used to enhance urban mobility and identify the main threads of AI
2. Applying Data Science and AI to Urban Mobility
3. Data Science and AI in cities from a computing/communication perspective: Edge and Cloud Computing

Urban Mobility

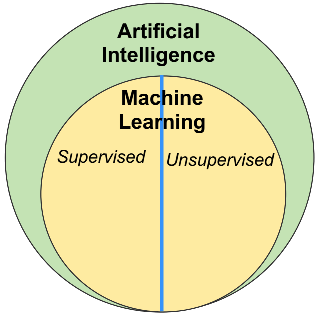
Co-funded by the European Union

ATM Àrea de Barcelona Autoritat del Transport Metropolità

FACTUAL

# 1. Artificial Intelligence

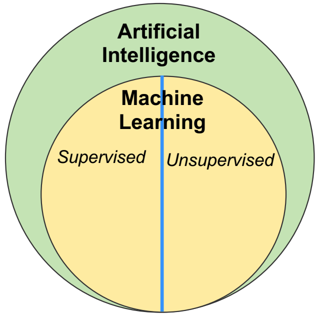
# Machine Learning: Supervised and Unsupervised




*"The field of study that gives computers the ability to learn without being explicitly programmed"* (Arthur Samuel, 1959)

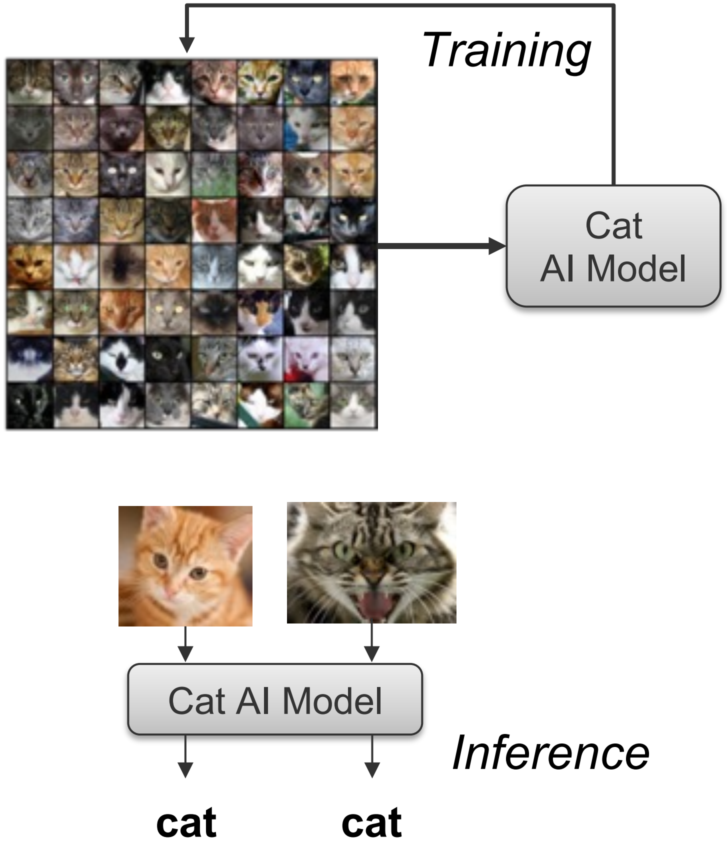
AI models are built upon representative sample data (training data) including the knowledge require to perform an specific action

- *Training phase:* The process of building the model based on a training data-set
  - **Supervised learning:** The training data-set is known and predefined
  - **Unsupervised learning:** The model is trained in operational environment, based on an objective function
- *Inference phase:* The process of performing the action for what the model has been designed, based on an input data

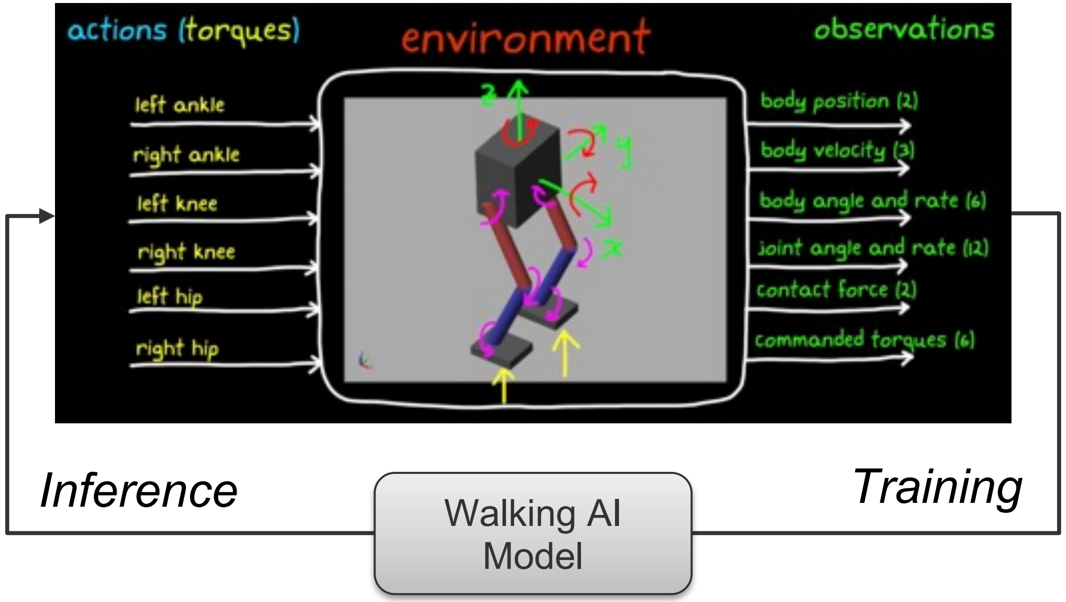



# Machine Learning: Supervised and Unsupervised

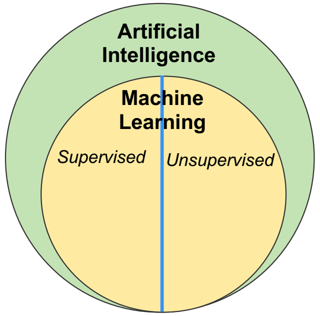
## Supervised




## Unsupervised



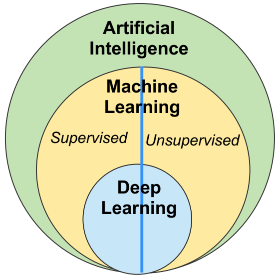
# Machine Learning: Supervised and Unsupervised



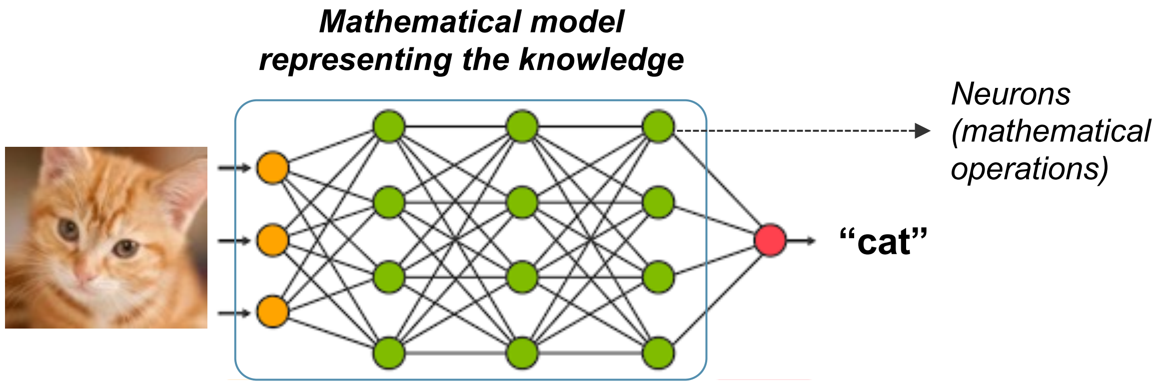


Learning to walk with unsupervised machine learning (deep reinforcement learning)

# Deep Learning

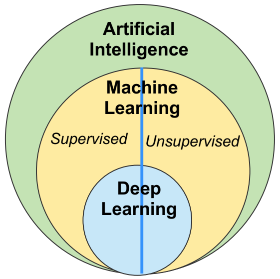



- The knowledge is represented in the form of *artificial neural networks* (ANN) or *deep neural networks* (DNN)
    - Inspired on how the information is processed by the brain

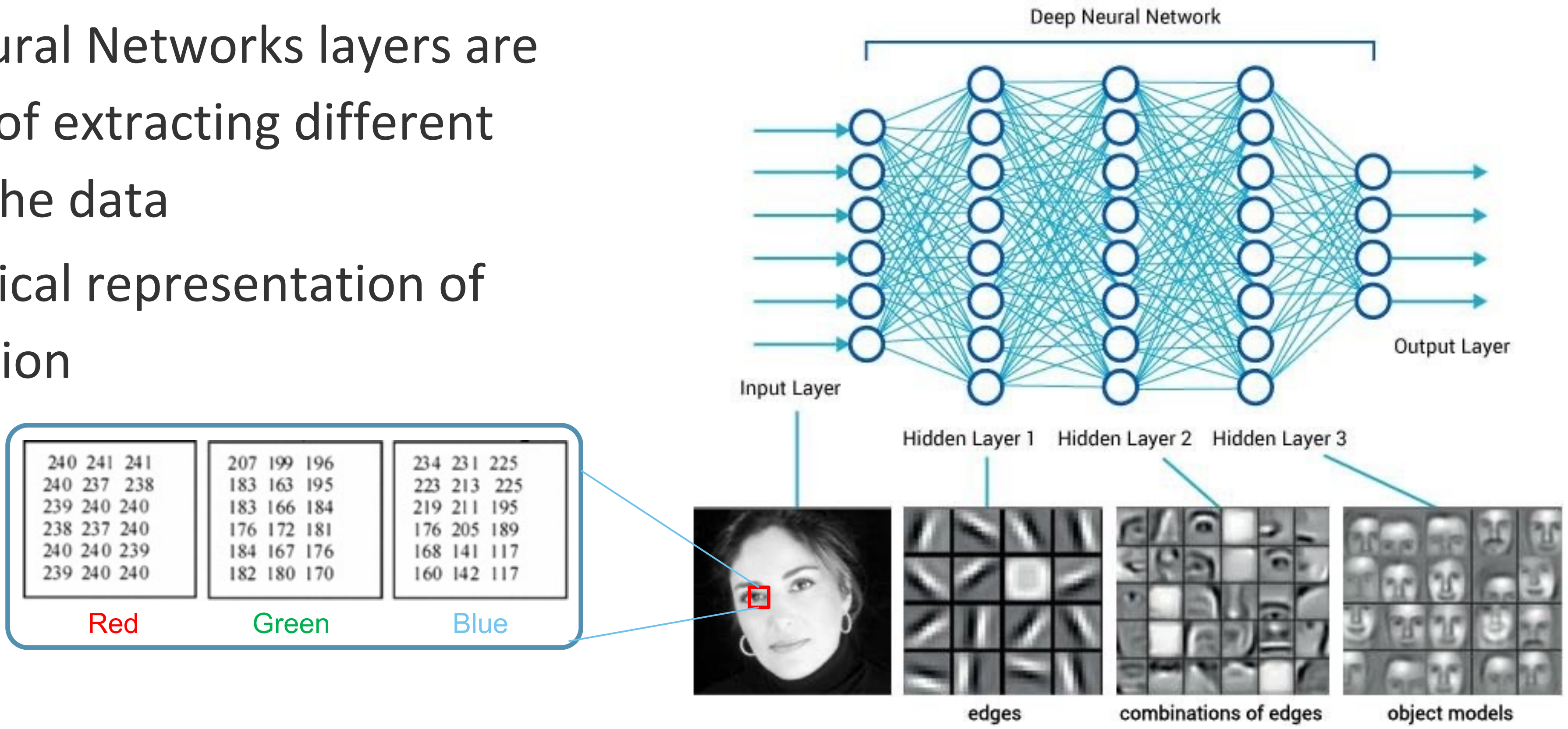



- Firstly proposed in 1967 by Ivakhnenko and Lapa
- In 2010, the use of Graphical Processing Units (GPUs) speed the training processes x100

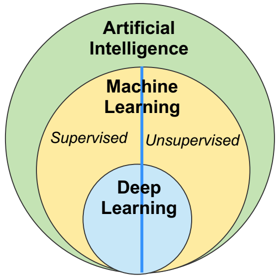
# Deep Learning



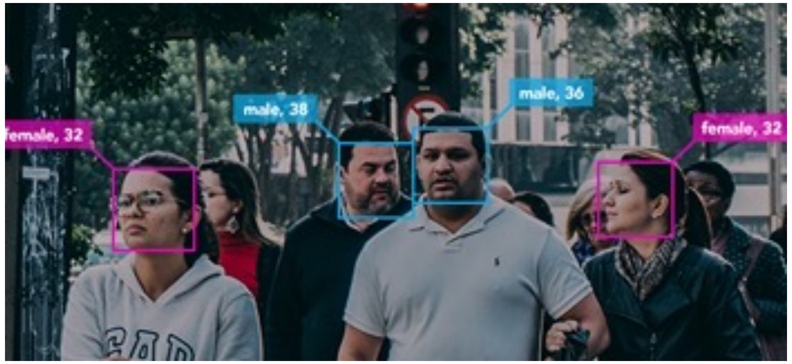

- Artificial Neural Networks layers are responsible of extracting different features of the data
  - Hierarchical representation of information



# Object Detection with Deep Learning



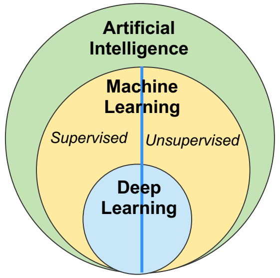

- Key functionality to perceive urban scenarios upon which decisions can be taken
  - Different models are used for different purposes

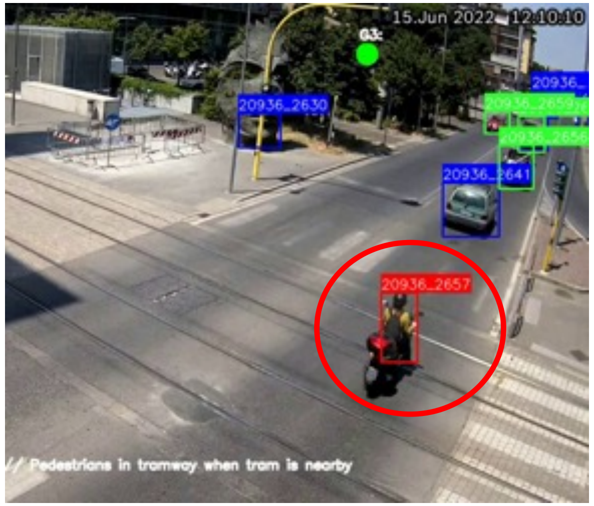


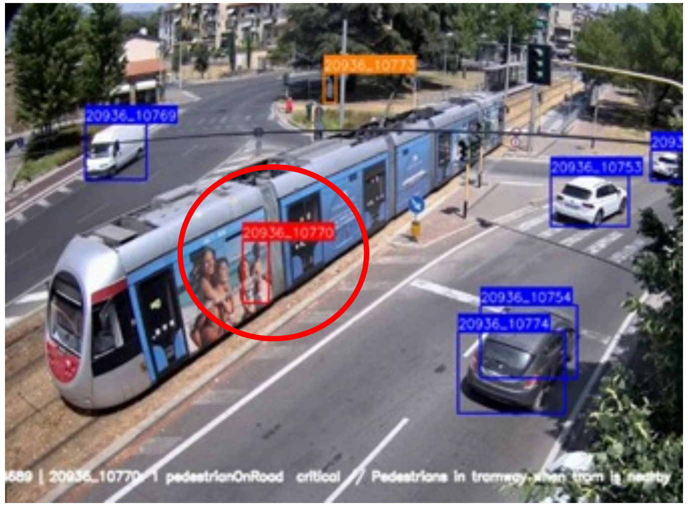
EIT Urban Mobility · Co-funded by the European Union · ATM Àrea de Barcelona Autoritat del Transport Metropolità · FACTUAL

# Object Detection with Deep Learning: Threads and Limitations

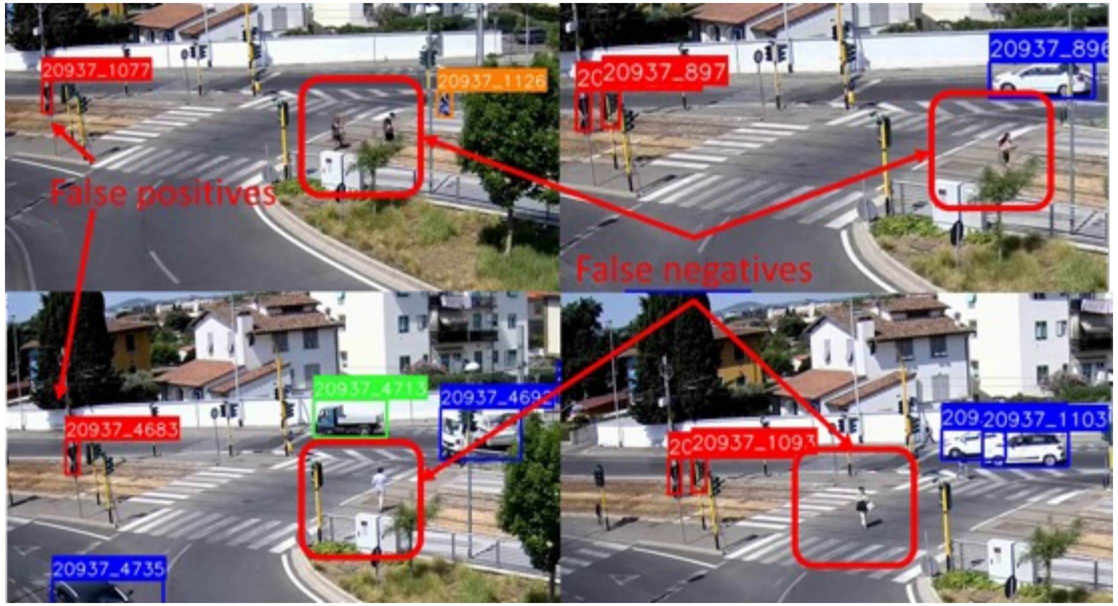



- False positives
  - Motorcycles and bicycles detected as pedestrians
  - Tram commercials detected as pedestrians
  - Traffic signs detected as pedestrians

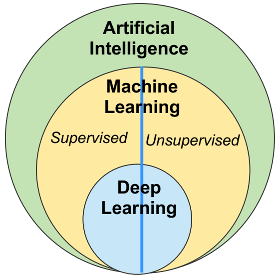


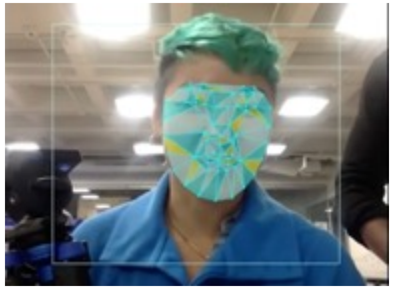

- False negatives
  - Detection failing due to light, obstructions, perspective, distance

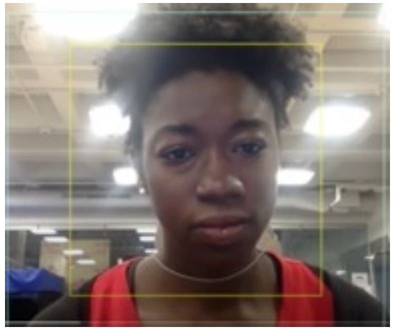


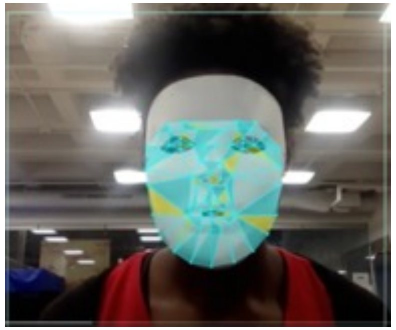
# Object Detection with Deep Learning: Bias/Discrimation




Popular face detection systems <u>are not trained with representative data </u>to detect dark-skin people

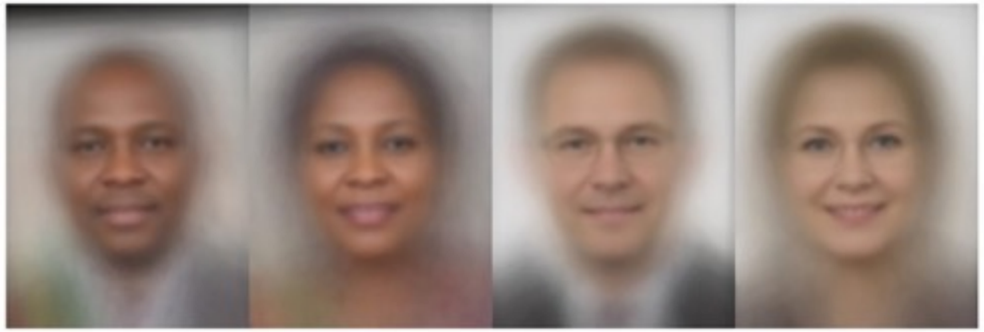




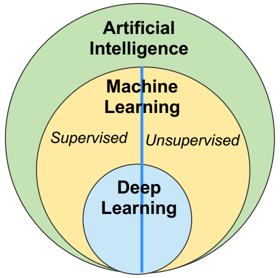


*https://youtu.be/TWWsW1w-BVo*

| Gender Classifier | Darker Male | Darker Female | Lighter Male | Lighter Female | Largest Gap |
|---|---|---|---|---|---|
| Microsoft | 94.0% | 79.2% | 100% | 98.3% | 20.8% |
| FACE++ | 99.3% | 65.5% | 99.2% | 94.0% | 33.8% |
| IBM | 88.0% | 65.3% | 99.7% | 92.9% | 34.4% |



**Gender Shades project** (http://gendershades.org) evaluates the accuracy of AI powered gender classification products




# Object Detection with Deep Learning: Adversarial Attacks

- Impercepcial changes in the input data (not visible by humans)


“Hare”


“Desk”

“Stop sign”

“Max Speed 100”

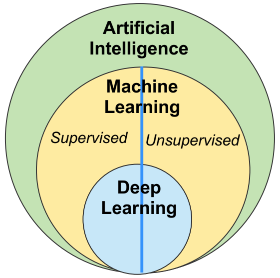
- Physical changes on the input image

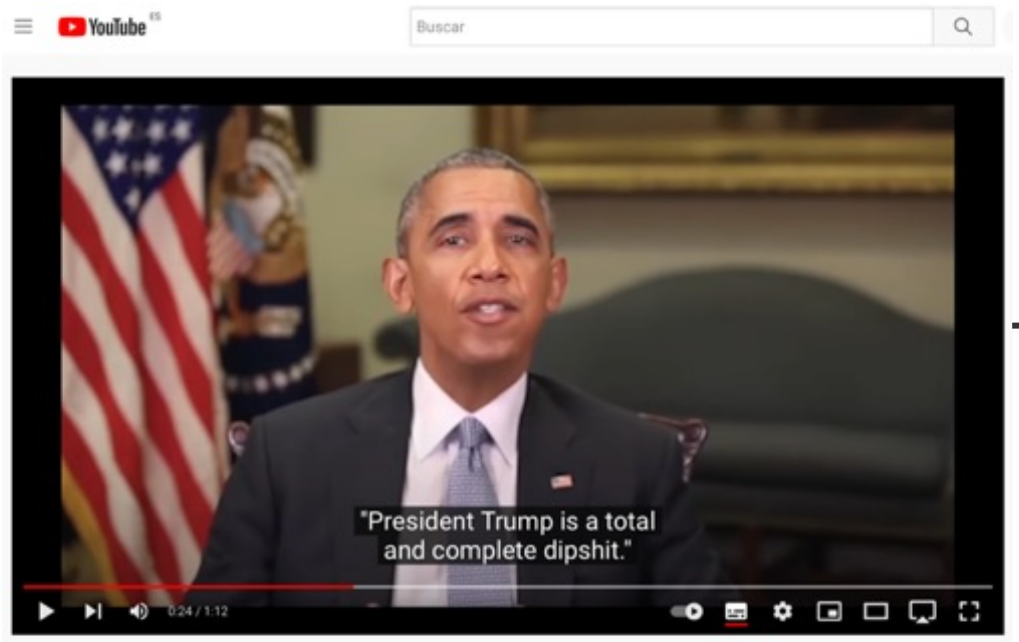

*The T-shirt prevents the person to be detected*


**Face recognition**

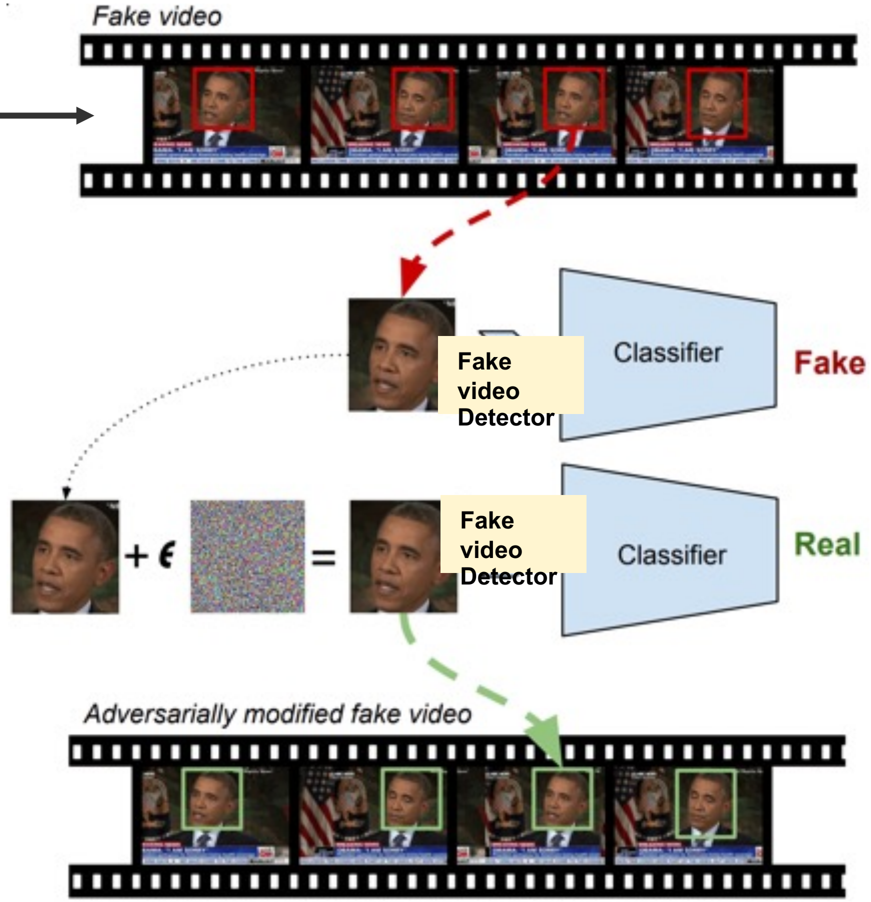

**“Milla Jojovich”**

# Object Detection with Deep Learning: Adversarial Attacks for Deep Fake Generation



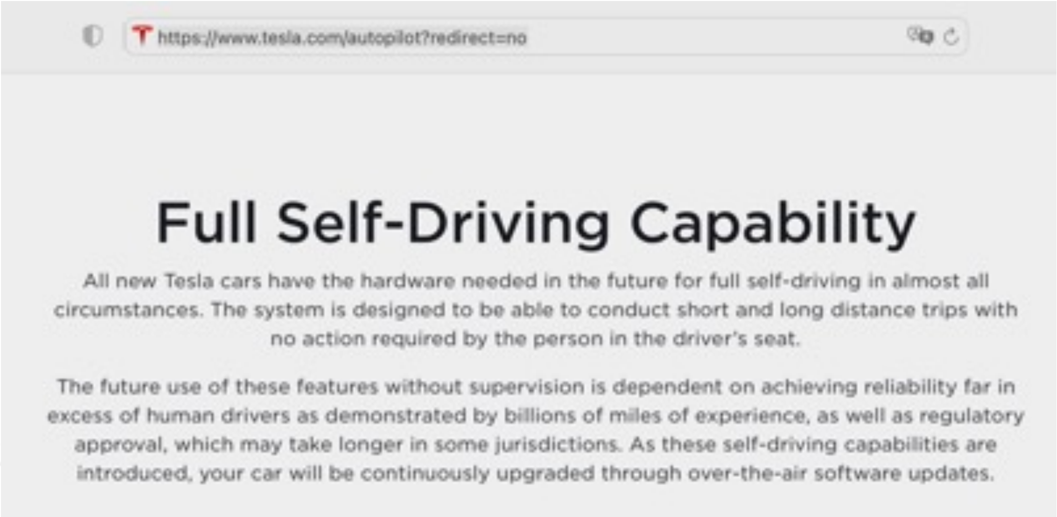

https://youtu.be/cQ54GDm1eL0



# Towards an Ethical and Secure Used of AI

- **Understand the specific task for what the AI engine has been trained**
  1. Ensure that the data-set used to train the AI model is representative for the intended task
  2. Ensure that the AI software method is appropiate for the task to be solved
     - Preferable to use open-source solutions
  3. Identify the security flaws of the data and the model
- Create **multi-disciplinary teams**

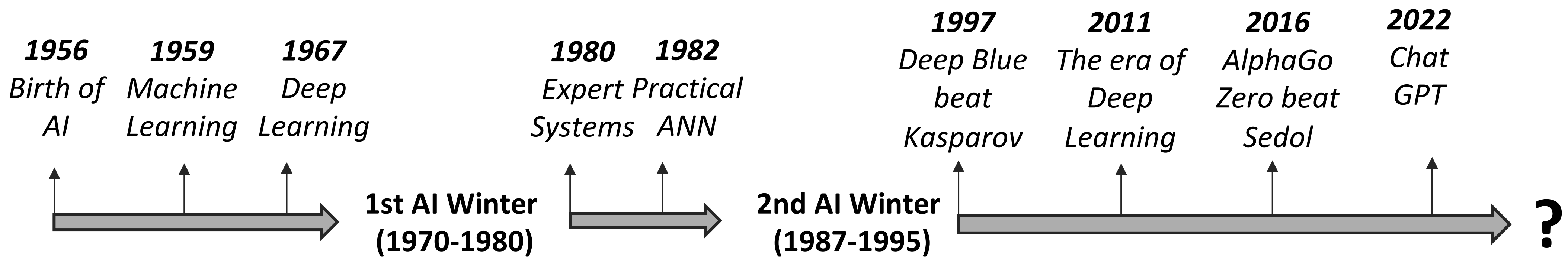
*https://www.tesla.com/autopilot?redirect=no*



# History of AI

**1965, H. A. Simon:** "*machines will be capable, within twenty years, of doing any work a man can do*"
**1970, M. Minsky:** "*In from three to eight years we will have a machine with the general intelligence of an average human being"*

***2015 - 2016, BMW, GM, VW, Tesla:*** *"Autonomous vehicles will be on the road on 2020"*



# Debate about AI and Ethics at Oxford Union Society (Dec 2021)

*"AI will never be ethical. It is a tool, and like any tool, it is used for good and bad. There is no such thing as a good AI, only good and bad humans. We are not smart enough to make AI ethical. We are not smart enough to make AI moral … In the end, I believe that the only way to avoid an AI arms race is to have no AI at all. This will be the ultimate defense against AI"*

**AI engine (Megatron Transformer)** trained with the whole Wikipedia, 63 million English news articles from 2016 to 2019, and 38 GB of public Reddit posts and comments

# 2. Applying Data Science and AI to Urban Mobility

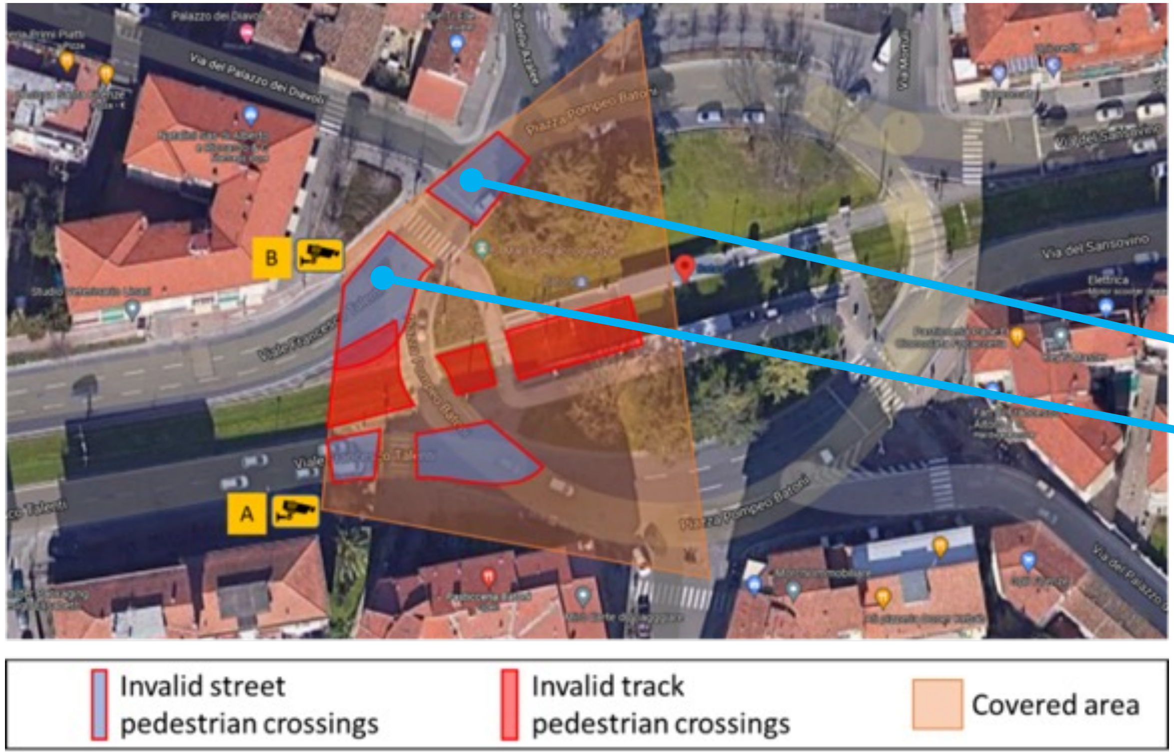
# Applying Data Science to Urban Mobility

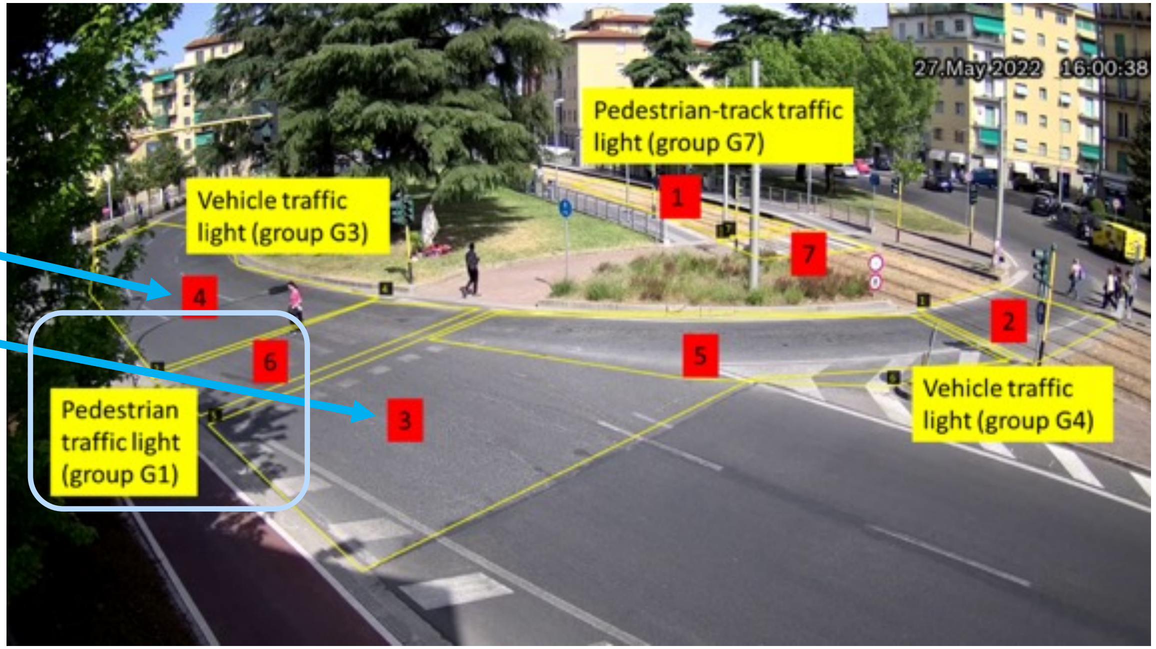
1. Which knowledge do we want to extract?
   - **How many times pedestrians crosses in red**
2. Percepcion
   - Identify data sources
   - Identify the urban actors relevant for the service
   - The dynamics of each actor (e.g., static/moving element, speed, direction)
   - *Provide semantic information to put each actor in the urban context*

*Piazza Batoni, Florence (Italy)*

# Semantic Information and Data Sources

*Piazza Batoni, Florence (Italy)*



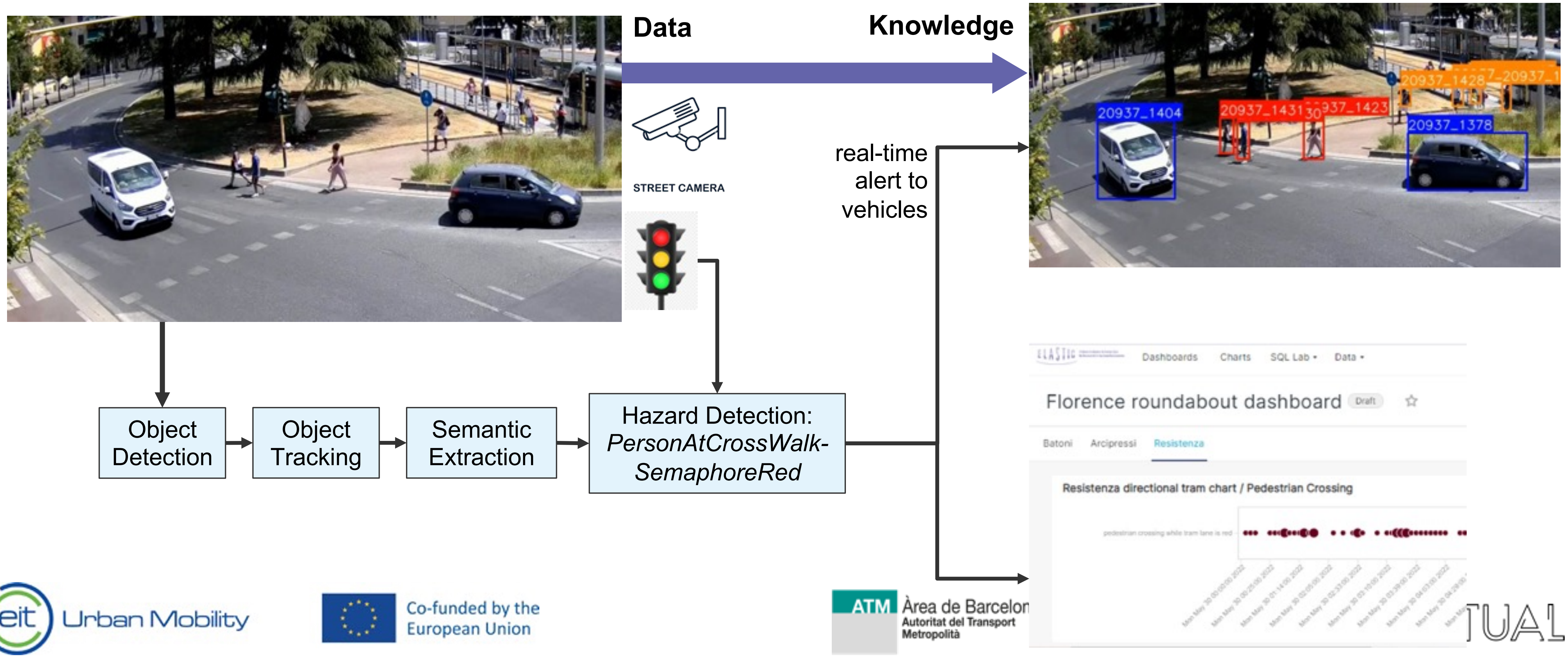



# Applying Data Science to Urban Mobility

1. Which knowledge do we want to extract?
   - **How many times pedestrians crosses in red**
2. Percepcion
   - Identify data sources
   - Identify the urban actors relevant for the analysis
   - The dynamics of each actor (e.g., static/moving element, speed, direction)
   - *Provide semantic information to put each actor in the urban context*
3. Knowledge extraction
   - For what do we want the knowledge?
4. Evaluation
   - Determine the extracted knowledge quality

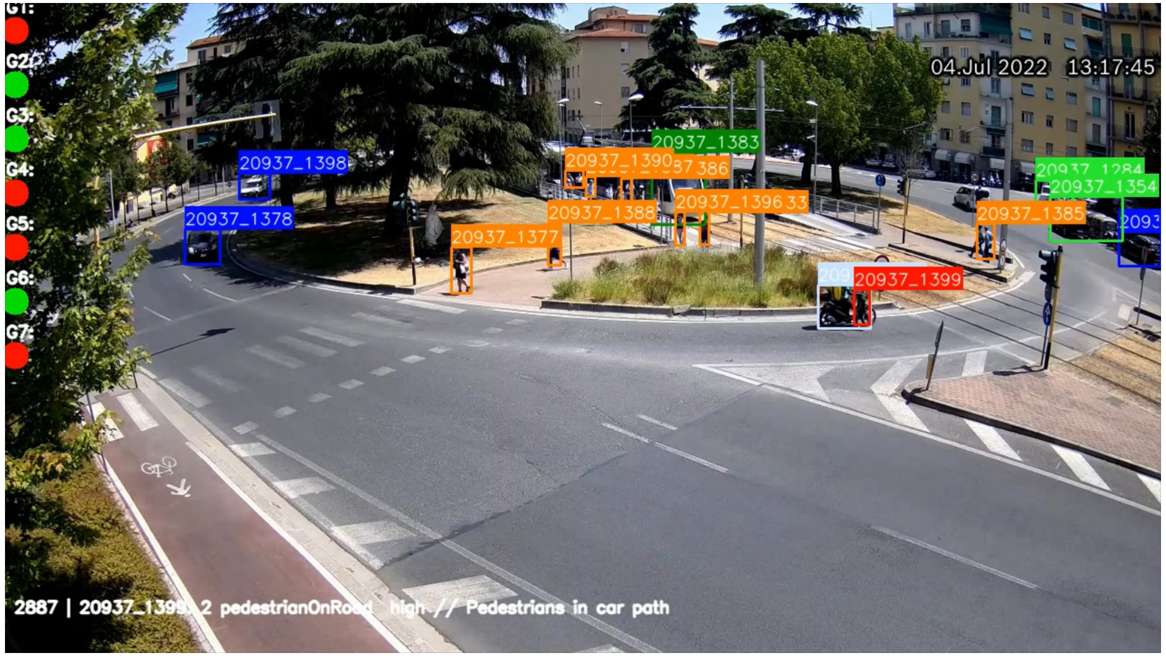
*Piazza Batoni, Florence (Italy)*

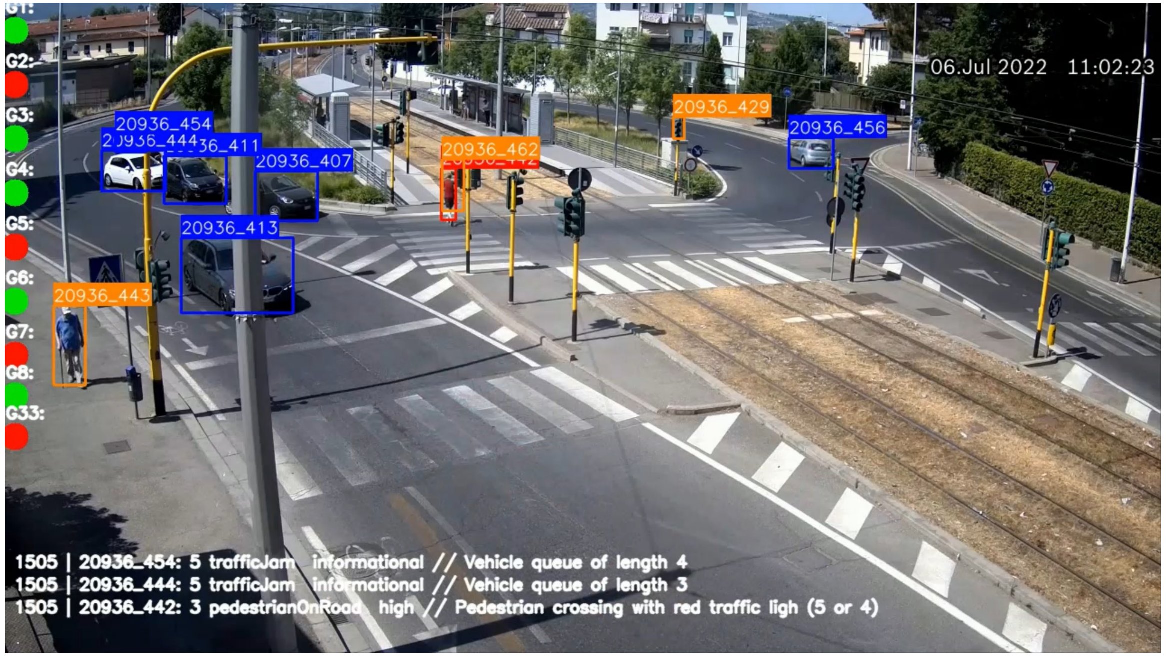
# Pedestrians on the crosswalk when the traffic light is red



# Pedestrians on the crosswalk when the traffic light is red

STREET CAMERA




eit Urban Mobility

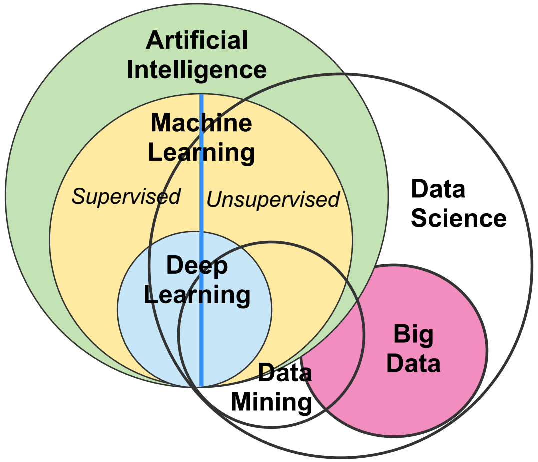
Co-funded by the European Union

ATM Àrea de Barcelona Autoritat del Transport Metropolità

FACTUAL

# How many times pedestrians crosses in red?

**Data Science**



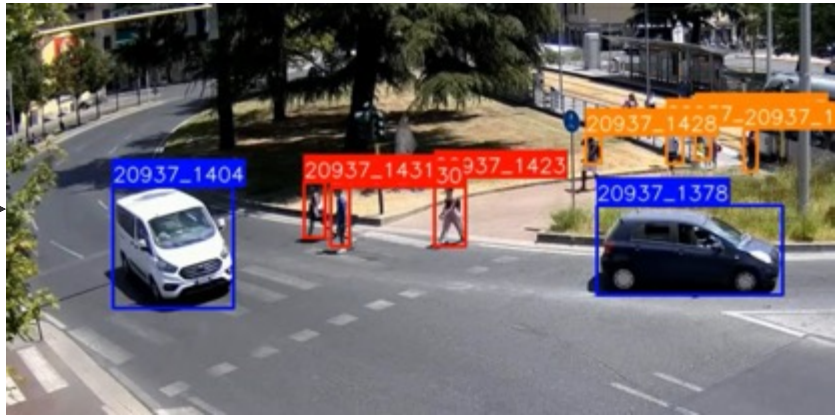

STREET CAMERA

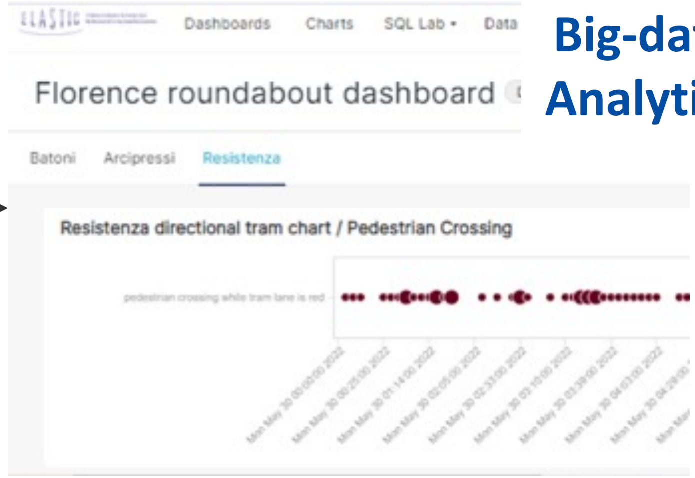
**Data Mining Workflow**

Object Detection → Object Tracking → Semantic Extraction → Hazard Detection: *PersonAtCrossWalk-SemaphoreRed*

***Supervised Deep Learning*** · ***Computer Vision*** · *Programming Logic*

real-time alert to vehicles



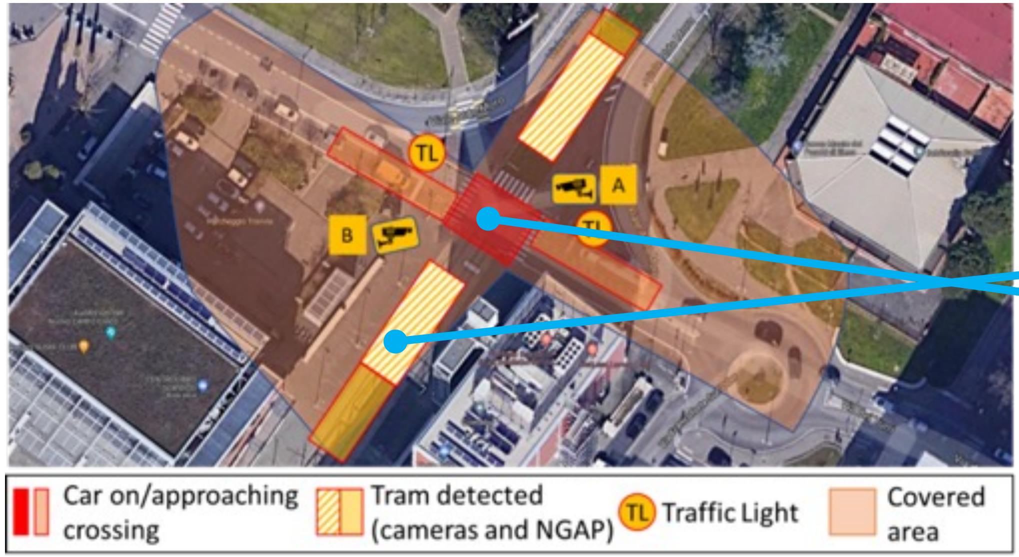



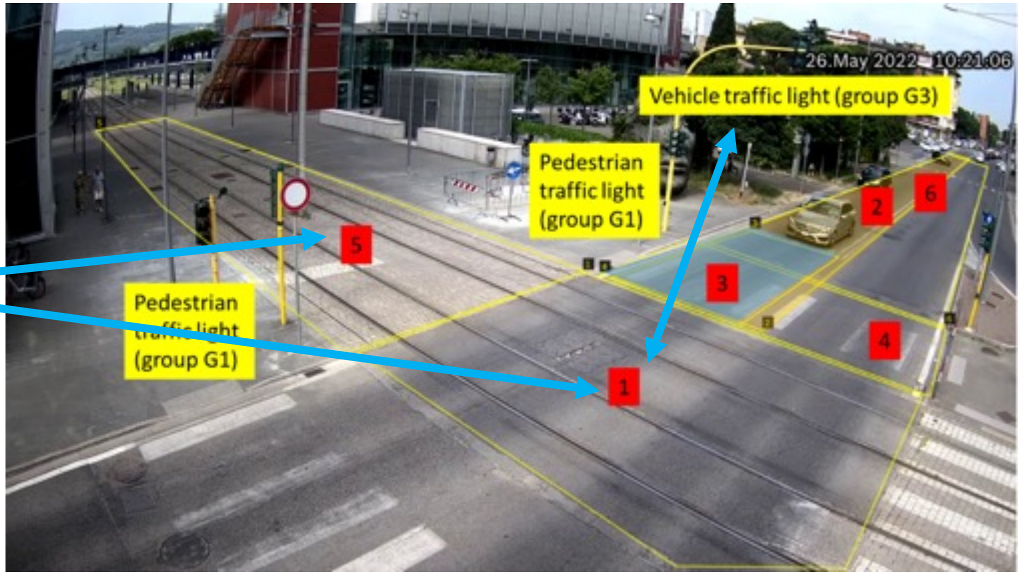

**Big-data Analytics**


eit Urban Mobility
Co-funded by the European Union
ATM Àrea de Barcelona Autoritat del Transport Metropolità
FACTUAL

# Pedestrians/cars crossing the rail tracks when tram approaching

1. Identify data sources
2. Provide semantic information
3. Identify the urban actors
4. Determine dynamics of actors

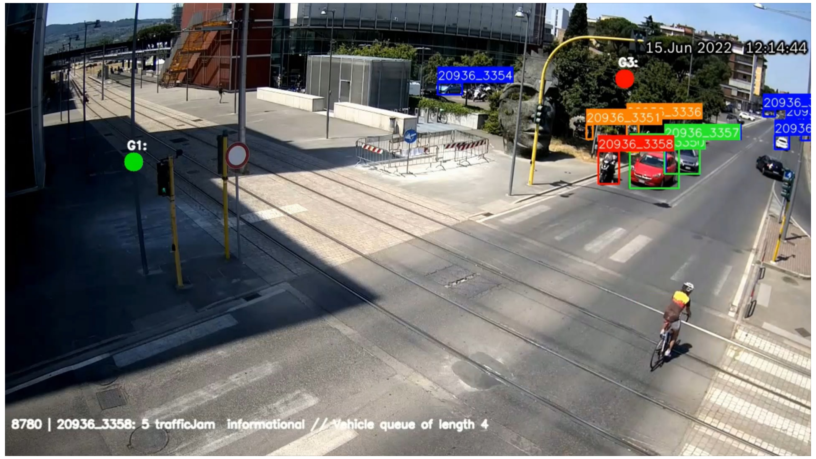



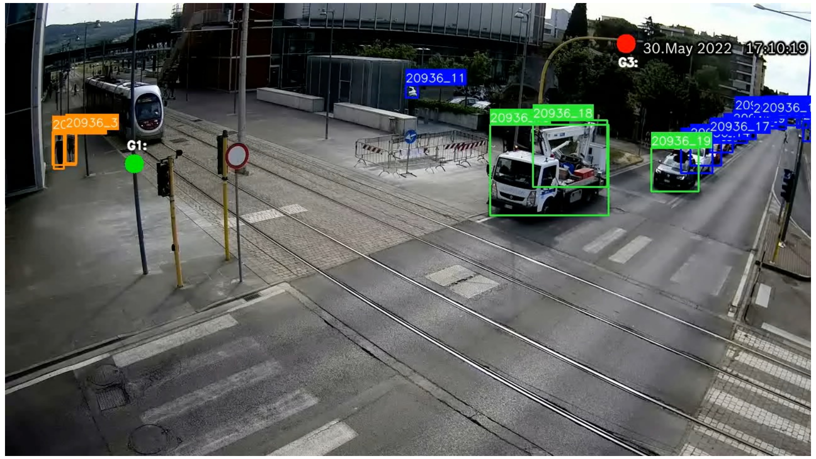




eit Urban Mobility
Co-funded by the European Union
ATM Àrea de Barcelona Autoritat del Transport Metropolità
FACTUAL

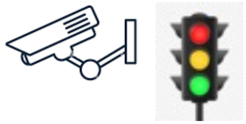
# Pedestrians/cars crossing the rail tracks when tram approaching

STREET CAMERA

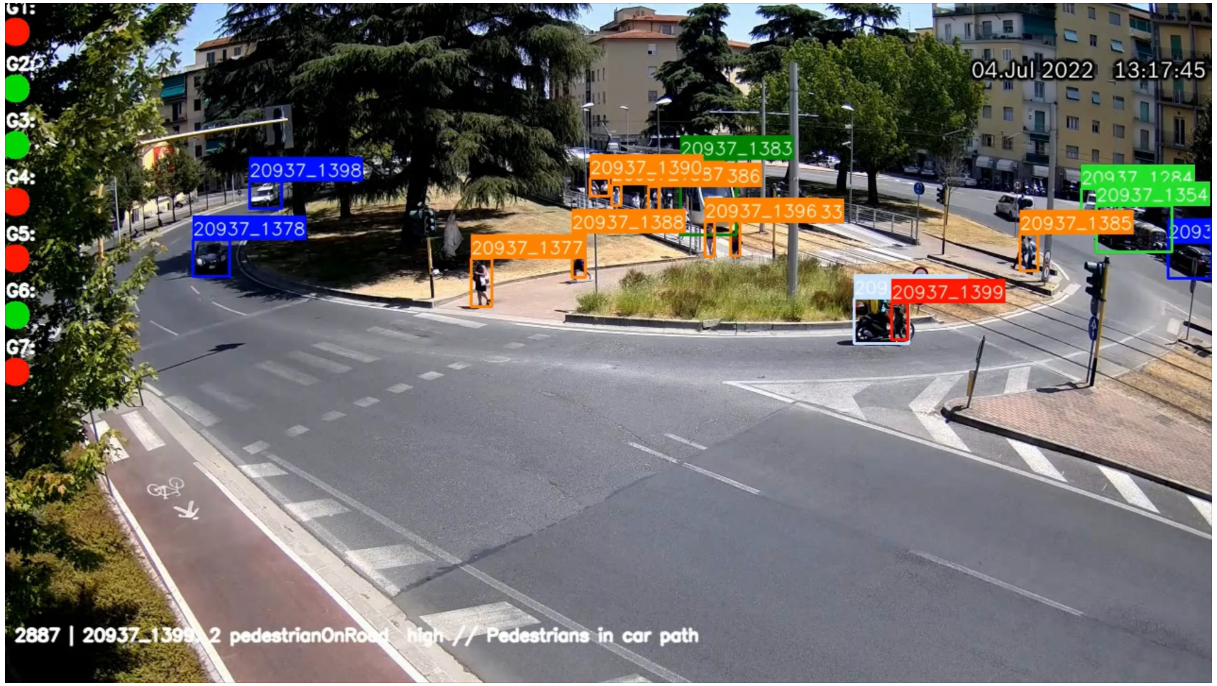



# Pedestrians crossing the street outside designated areas








eit Urban Mobility

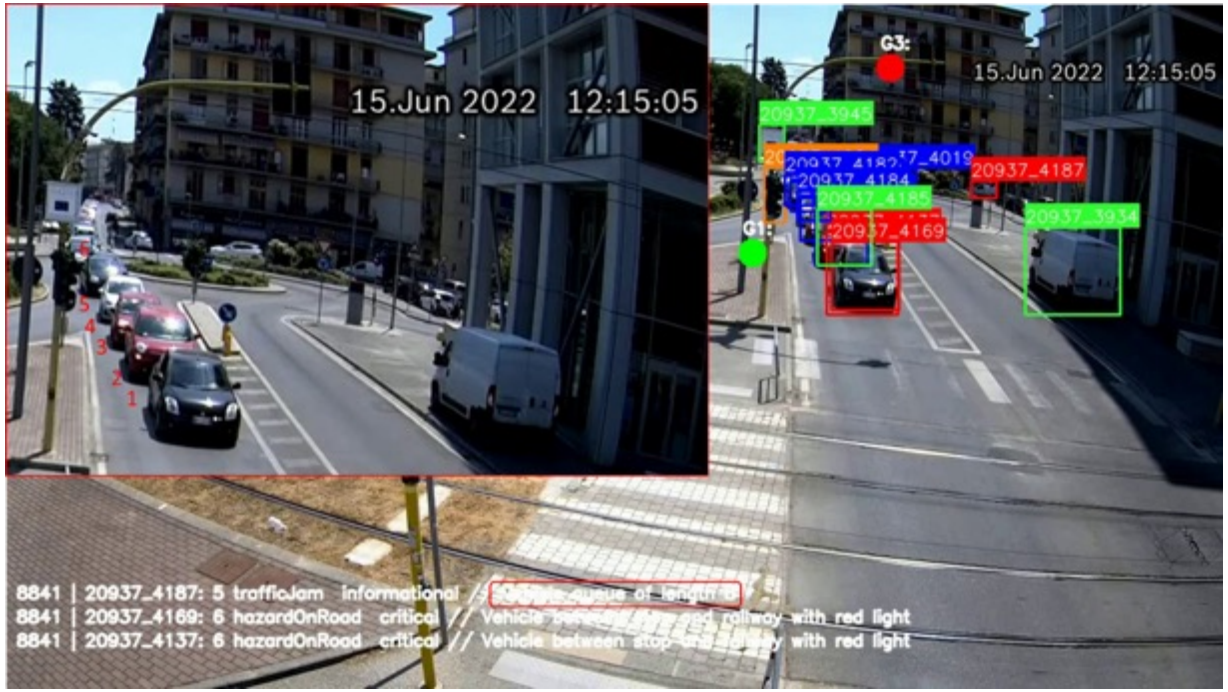
Co-funded by the European Union

ATM Àrea de Barcelona
Autoritat del Transport Metropolità

FACTUAL

# Traffic Queues


**Vehicle traffic queue of length 11**




**Vehicle traffic queue of length 6**

# City-Car Interaction

Co-funded by the European Union

Àrea de Barcelona Autoritat del Transport Metropolità

# Evaluation Methodology

Object Detection → Object Tracking → Semantic Extraction → Hazard Detection: *PersonAtCrossWalk-SemaphoreRed*

## **Confusion matrix** for each <u>event of interest</u>

- *True Positives:* true event detected
- *False Negatives*: true event not detected
- *False Positives*: no event but detected

## Accuracy evaluation

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

Refers to the portion of relevant events correctly detected among all **detected** events

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

Refers to the portion of relevant events correctly detected among all **relevant** events

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

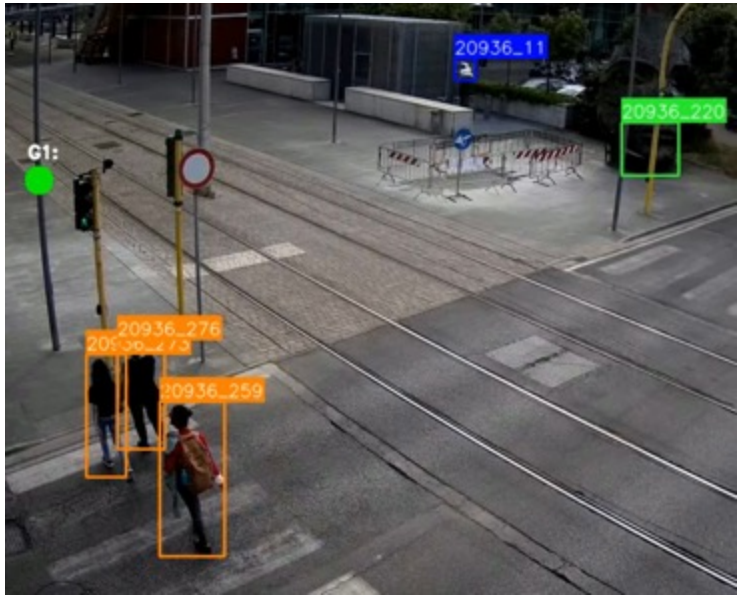
A measure of event detection accuracy:
- F1 = 1 perfect detection
- F1 = 0 no detection

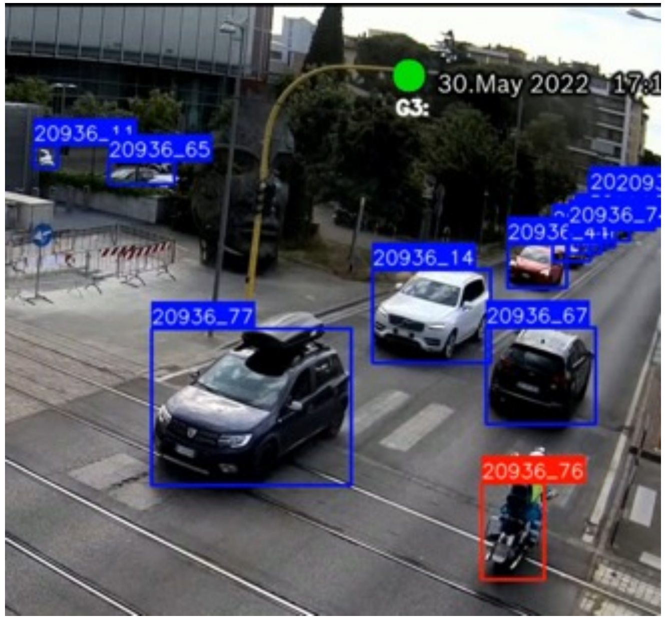
# Evaluation Methodology

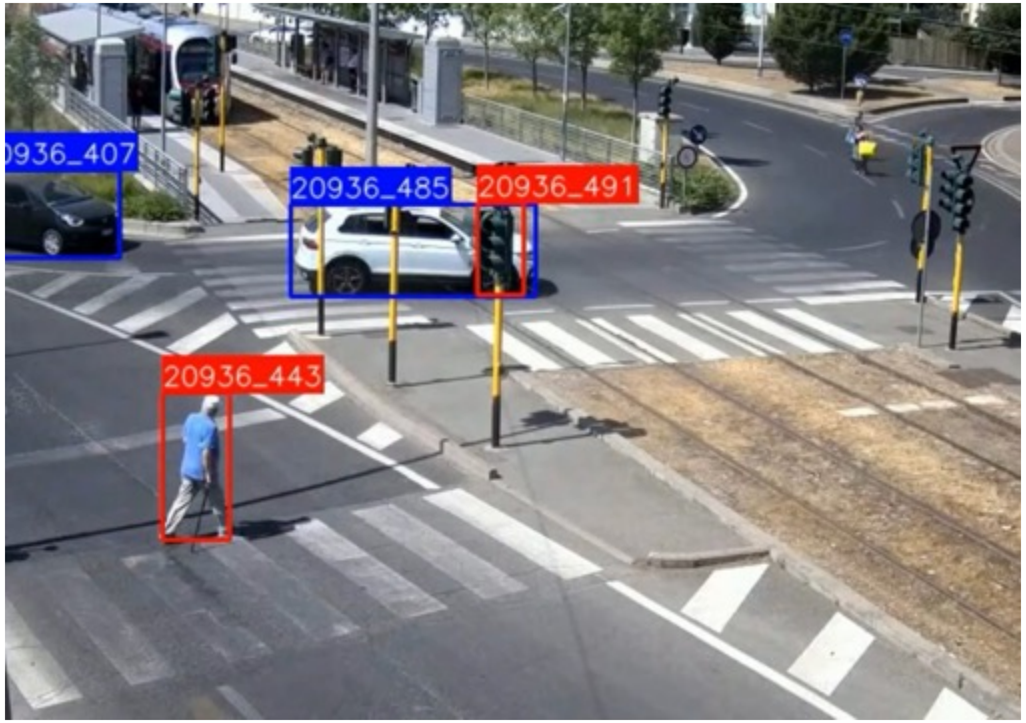
*True Positives:* true event detected

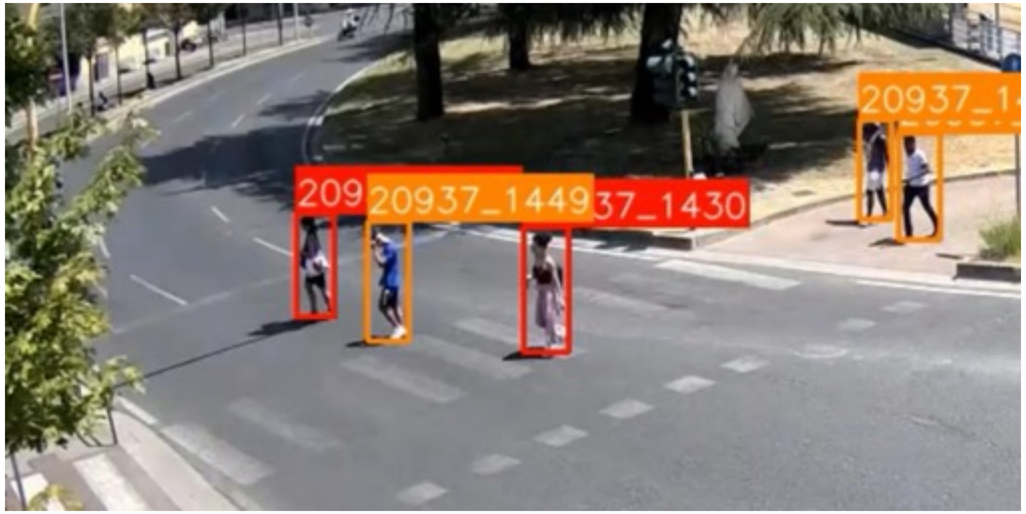
*False Negatives*: true event not detected

*False Positives*: no event but detected




Urban Mobility

Co-funded by the European Union

ATM Àrea de Barcelona
Autoritat del Transport Metropolità

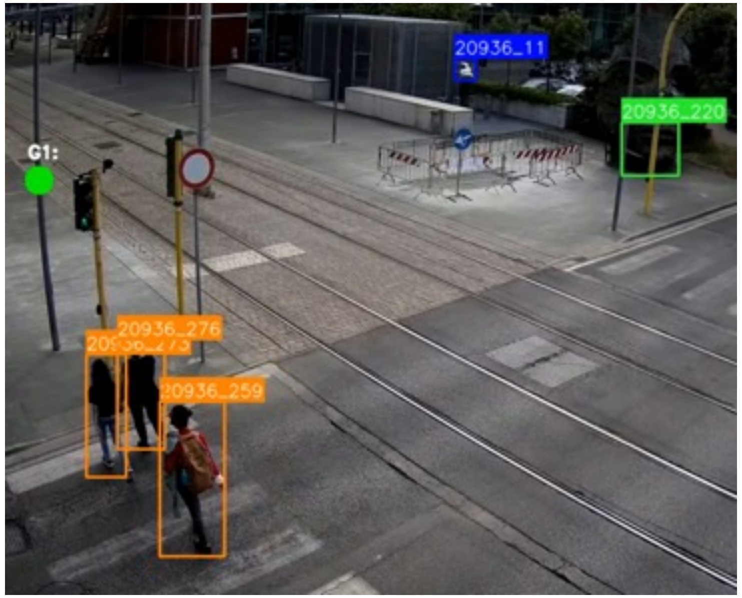
FACTUAL

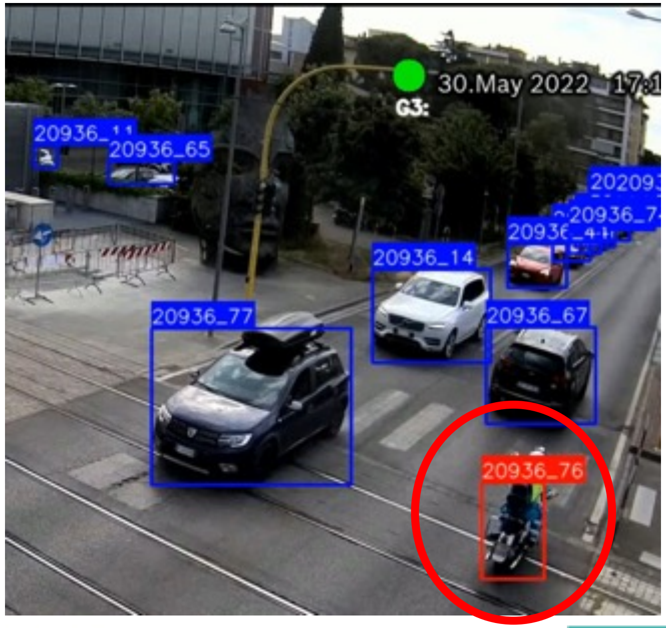
# Evaluation Methodology

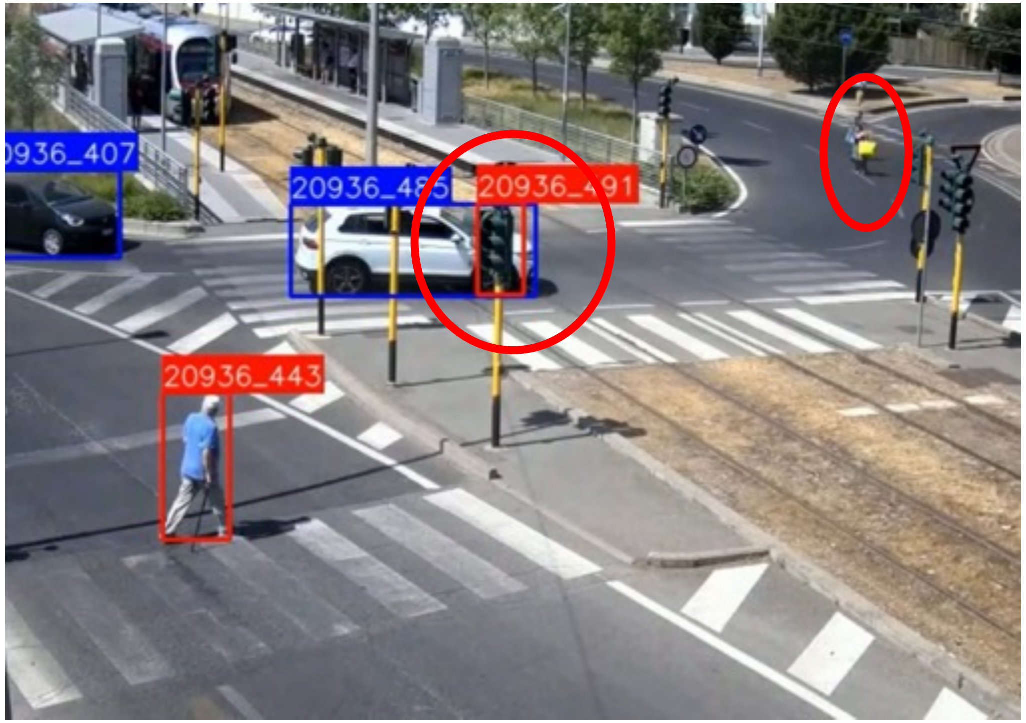
**(3)** *True Positives:* true event detected

**(2)** *False Negatives*: true event not detected

**(2)** *False Positives*: no event but detected





Urban Mobility

Co-funded by the European Union

ATM Àrea de Barcelona
Autoritat del Transport Metropolità

FACTUAL

# Evaluation Methodology

- Manual inspection of data sources (videos, semaphores, tram position) collected on representative conditions

| Event | Precision | Recall | F1-score |
|---|---|---|---|
| **P1 (pedestrians crossing tracks)** | 71% | 98.6% | 82.6% |
| **P2 (pedestrians crossing street)** | 12.7% | 100% | 22.5% |
| **P3 (pedestrians crossing with red)** | 62.2% | 100% | 76.7% |

| Event | Precision | Recall | F1-score |
|---|---|---|---|
| **P1 (pedestrians crossing tracks)** | 28.6% | 100% | 44.44% |
| **P2 (pedestrians crossing street)** | 13.2% | 65.22% | 21.9% |
| **P3 (pedestrians crossing with red)** | 72.3% | 89.4% | 80% |

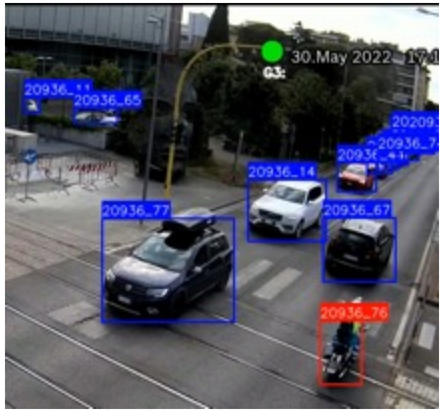
# Evaluation Methodology

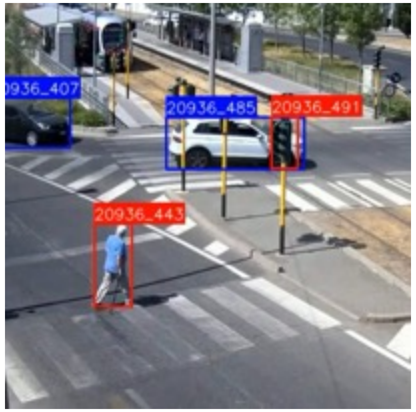
- Manual inspection of data sources (videos, semaphores, tram position) collected on <u>representative conditions</u>

| Event | Precision | Recall | F1-score |
|---|---|---|---|
| P1 (pedestrians crossing tracks) | 71% | 98.6% | 82.6% |
| P2 (pedestrians crossing street) | 12.7% | 100% | 22.5% |
| P3 (pedestrians crossing with red) | 62.2% | 100% | 76.7% |

| Event | Precision | Recall | F1-score |
|---|---|---|---|
| P1 (pedestrians crossing tracks) | 28.6% | 100% | 44.44% |
| P2 (pedestrians crossing street) | 13.2% | 65.22% | 21.9% |
| P3 (pedestrians crossing with red) | 72.3% | 89.4% | 80% |

**The DNN detects motocycles, bycicles and semaphores as pedestrians**





eit Urban Mobility
Co-funded by the European Union
ATM Àrea de Barcelona Autoritat del Transport Metropolità
FACTUAL

# Evaluation Methodology

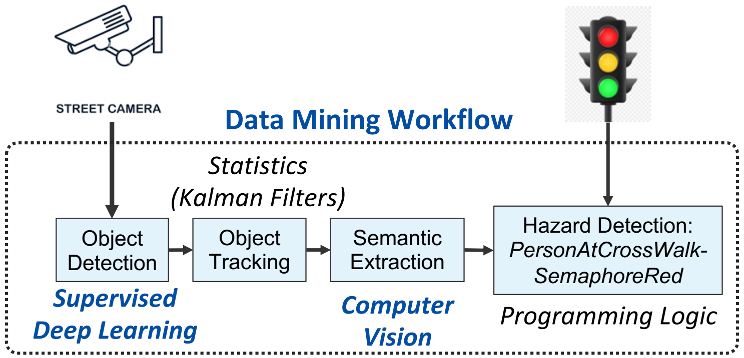
- Manual inspection of data sources (videos, semaphores, tram position) collected on representative conditions

Only pedestrian uses crosswalks and tracks

| Event | Precision | Recall | F1-score |
|---|---|---|---|
| P1 (pedestrians crossing tracks) | 71% | 98.6% | 82.6% |
| P2 (pedestrians crossing street) | 12.7% | 100% | 22.5% |
| P3 (pedestrians crossing with red) | 62.2% | 100% | 76.7% |


| Event | Precision | Recall | F1-score |
|---|---|---|---|
| P1 (pedestrians crossing tracks) | 28.6% | 100% | 44.44% |
| P2 (pedestrians crossing street) | 13.2% | 65.22% | 21.9% |
| P3 (pedestrians crossing with red) | 72.3% | 89.4% | 80% |

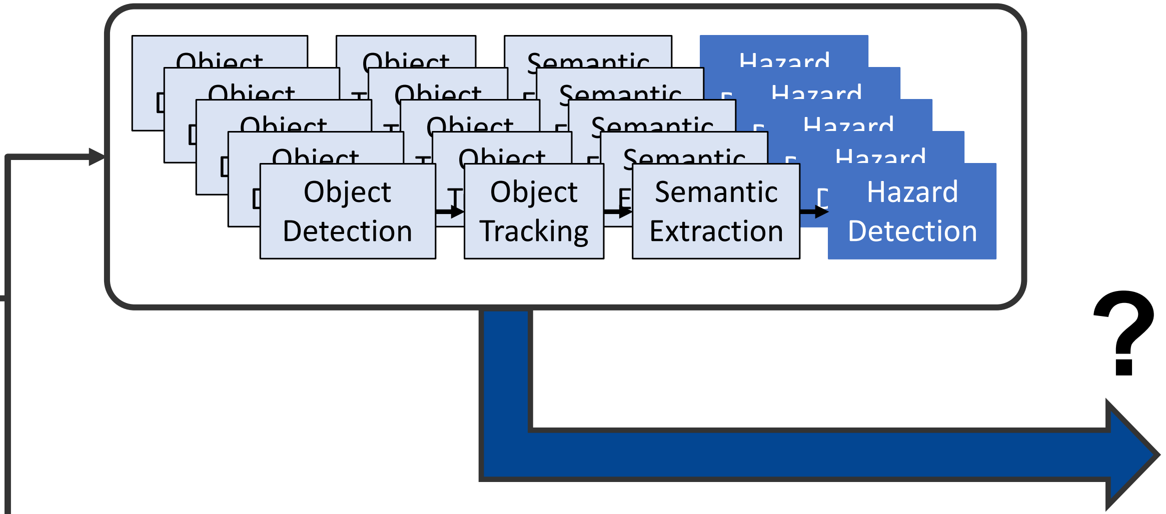
# Actions to increase accuracy




- Improvement of object detection
  - Motorcycles, bicycles and traffic lights (currently identified as pedestrians)
  - Tram vehicles (currently identified as cars)
  - Used of dynamic properties of objects (speed, direction, etc.)
- More complex semantic rules can be applied for each specific event and captured area, targeting the issues identified in the quantitative analysis
  - Refine semantic annotation to filter out traffic signs, compensate for camera perspective
  - Use additional metrics, e.g., speed or trajectory to refine analytics

# 3. Data Science and AI in cities from a computing/communication perspective

# Data Mining Workflows and Computing Resources




**Computing/Communication City Resources**

How are the data-analytics workflows **<u>developed and mapped</u>** to the available **computing/ communication resources?**

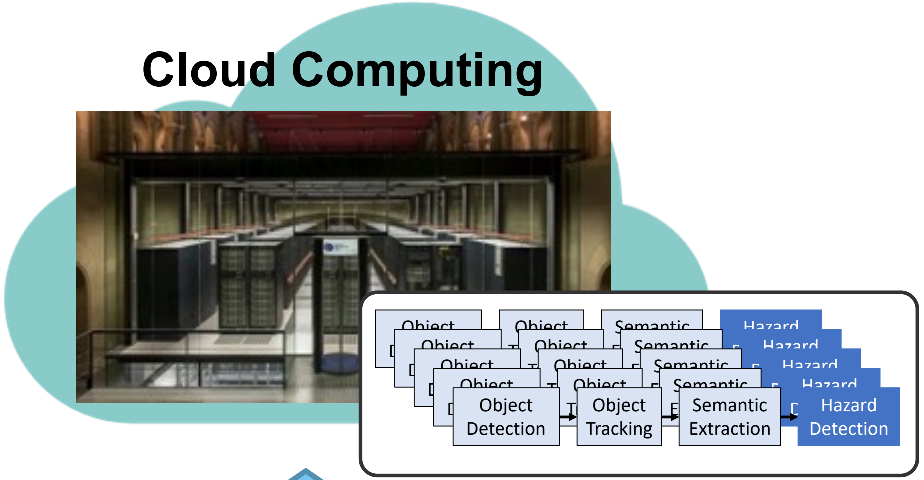
Co-funded by the European Union

ATM Àrea de Barcelona
Autoritat del Transport Metropolità

FACTUAL

# Centralised vs. Distribute

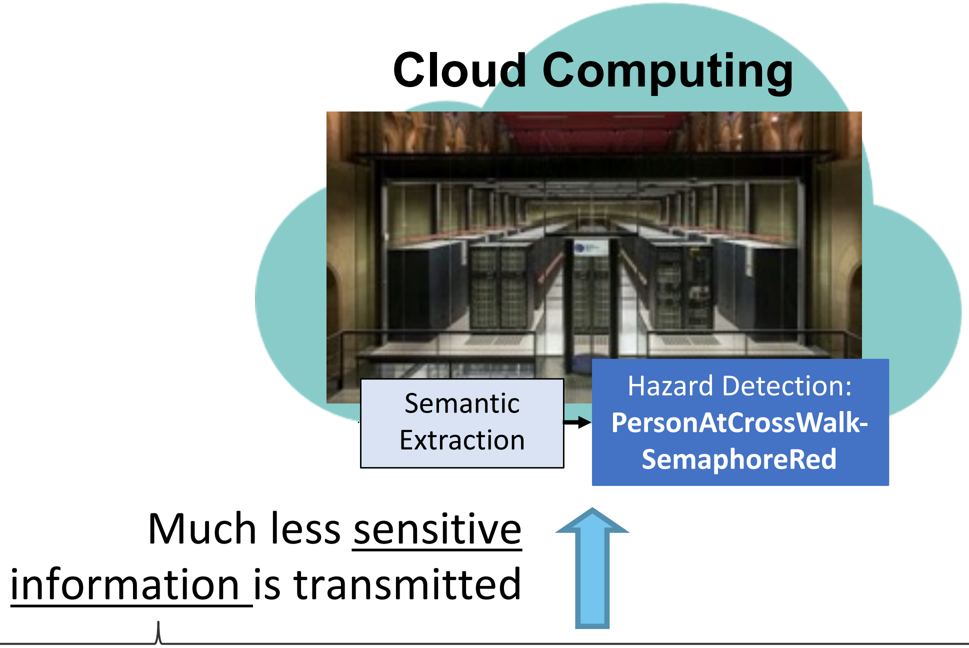
- Unsecure (single point of attack)
- Privacy Issues
- Inefficient in terms of energy and cost

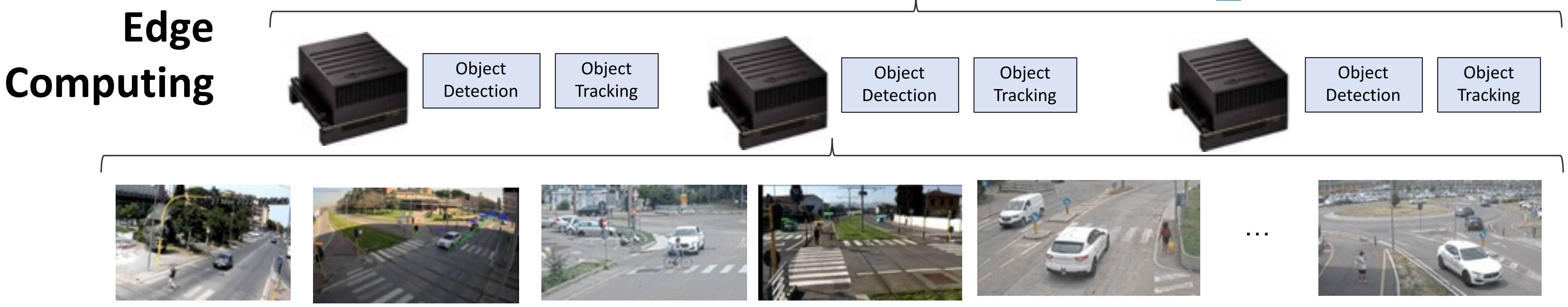



Huge amounts of <u>sensitive information</u> are transmitted

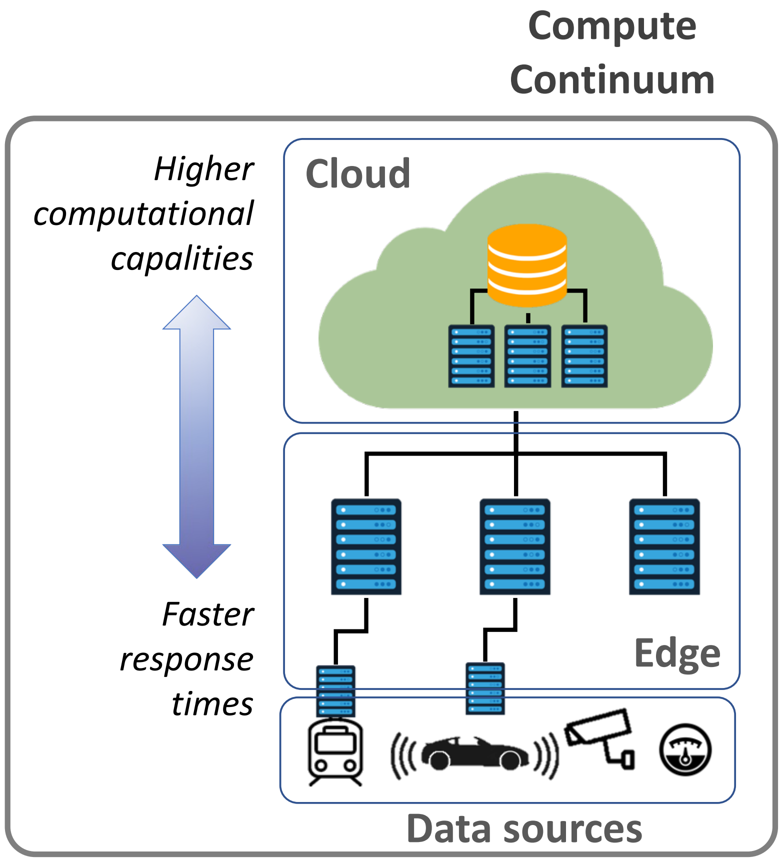
# Centralised vs. Distribute: The Edge Computing

- More robust in terms of security
- Privacy is ensured
- Energy and cost efficiency



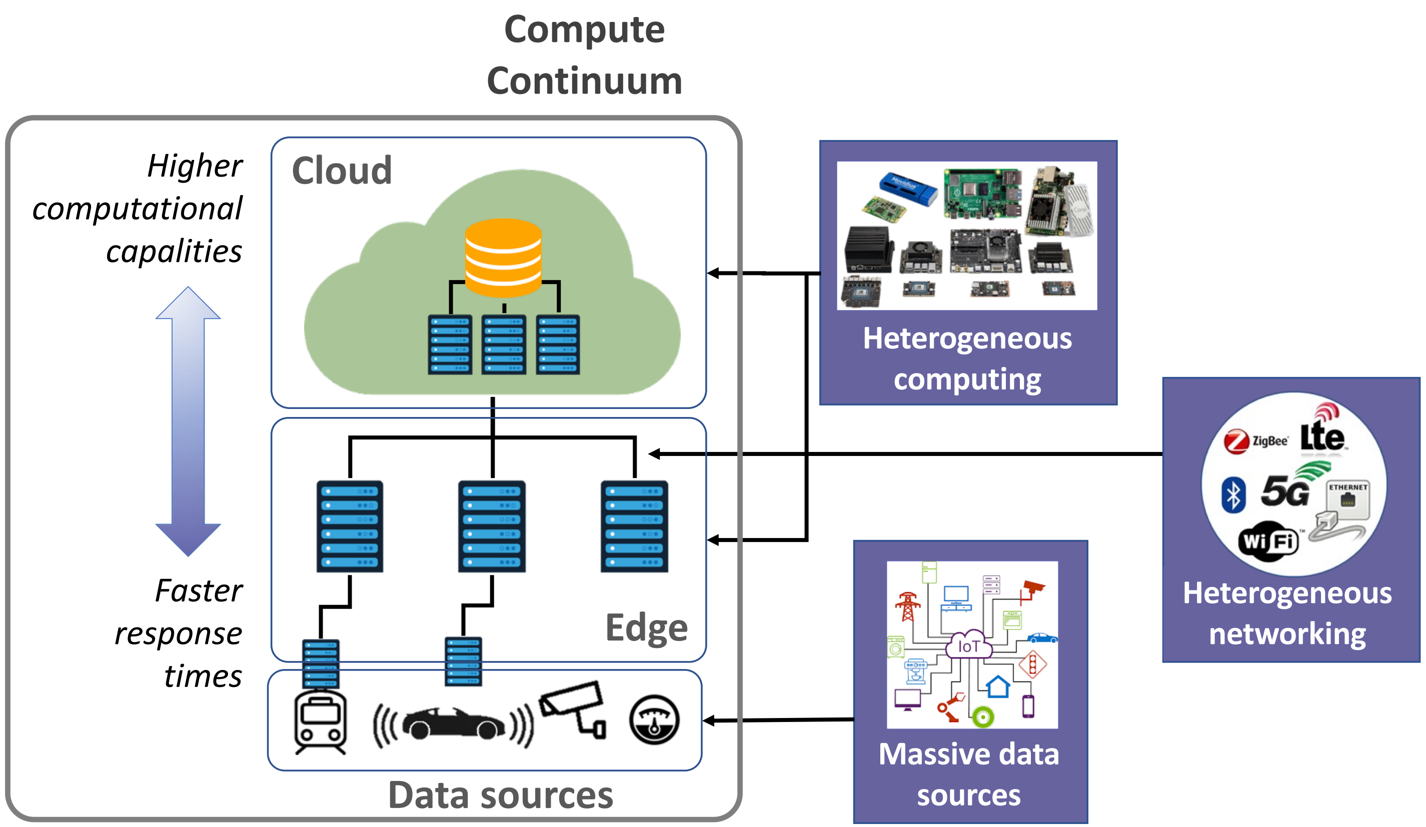
# The Compute Continuum: Edge and Cloud




1. <u>Facilitate</u> the development of **complex data analytics workflows independently of the platform**
2. <u>Increase</u> the **capabilities of the data analytics** by distributing them across the compute continuum
3. <u>Fulfill</u> the **non-functional requirements** inherited from the application domain, e.g., real-time, privacy

# The Compute Continuum: Edge and Cloud

- There are not mature technical solutions capable of fully exploiting the capabilities of the compute continuum
- Very interesting research is still pending!
- Mobility experts will play a fundamental role



# Complex Data Mining Workflows

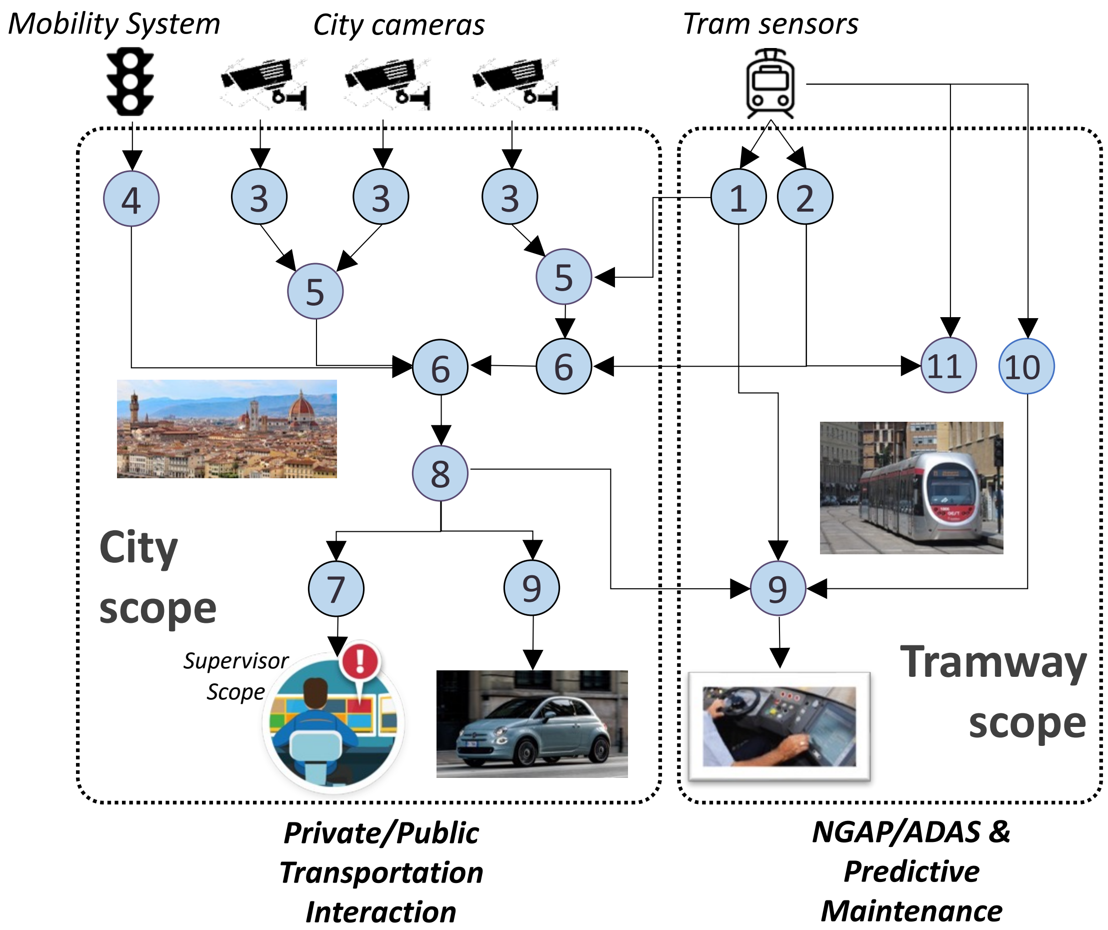
ELASTIC

A Software Architecture for Extreme-ScaLe Big-Data AnalyticS in Fog CompuTIng Ecosystems

## Data Analytics Methods

1. Sensor fusion (ADAS)
2. Tram position (NGAP)
3. Object detection, tracking & semantics
4. UTC/Supervisor consolidation
5. Data fusion
6. Data aggregation
10. Electric power consumption
11. Defect Detector

**7. Dashboard**
**8. Hazard detection**
**9. Alert visualization (cars/trams)**

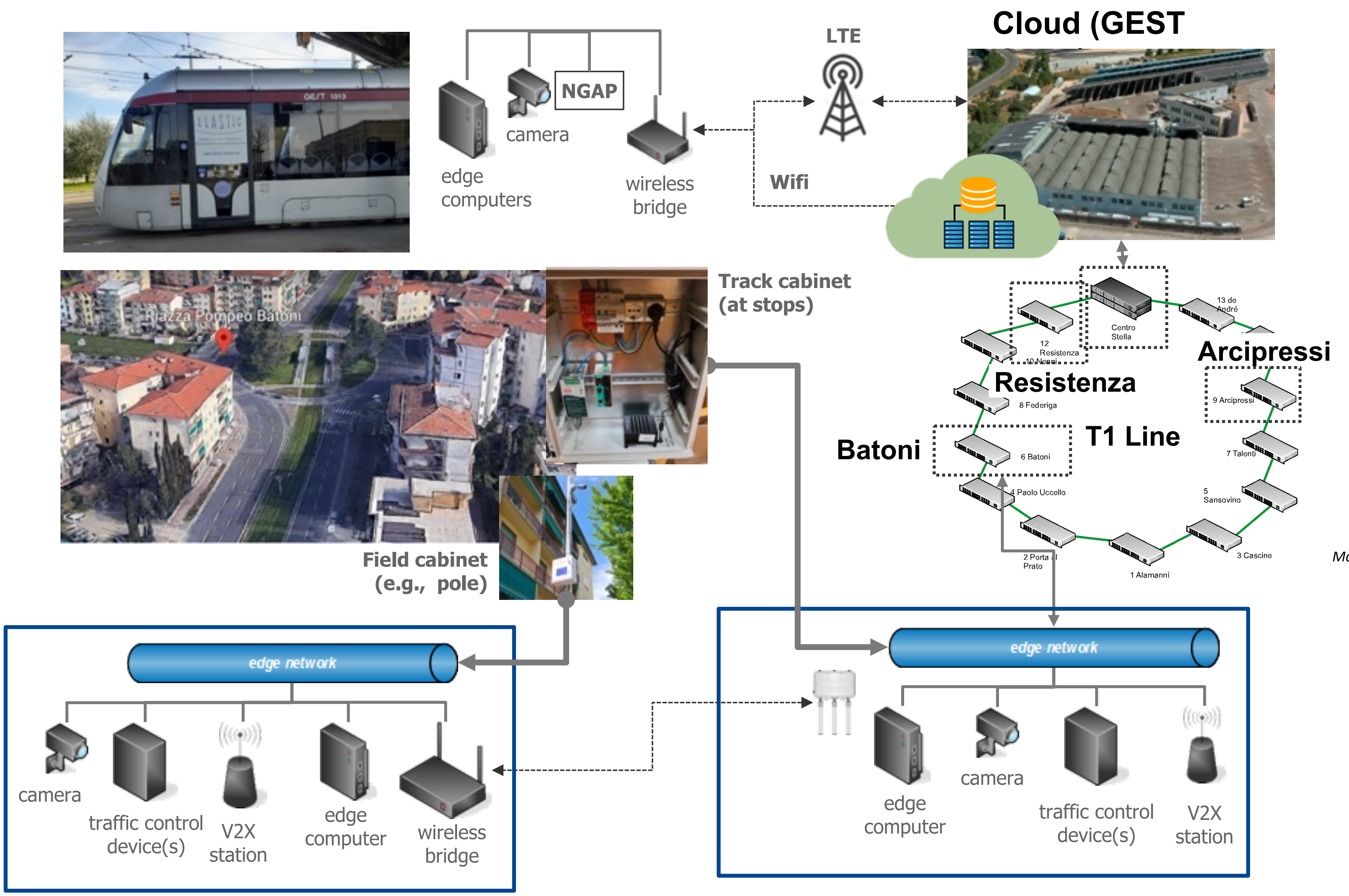


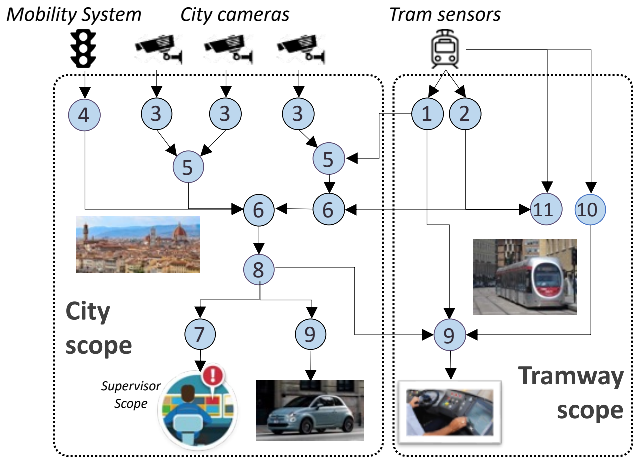

eit Urban Mobility
Co-funded by the European Union
ATM Àrea de Barcelona Autoritat del Transport Metropolità
FACTUAL

# The Compute Continuum

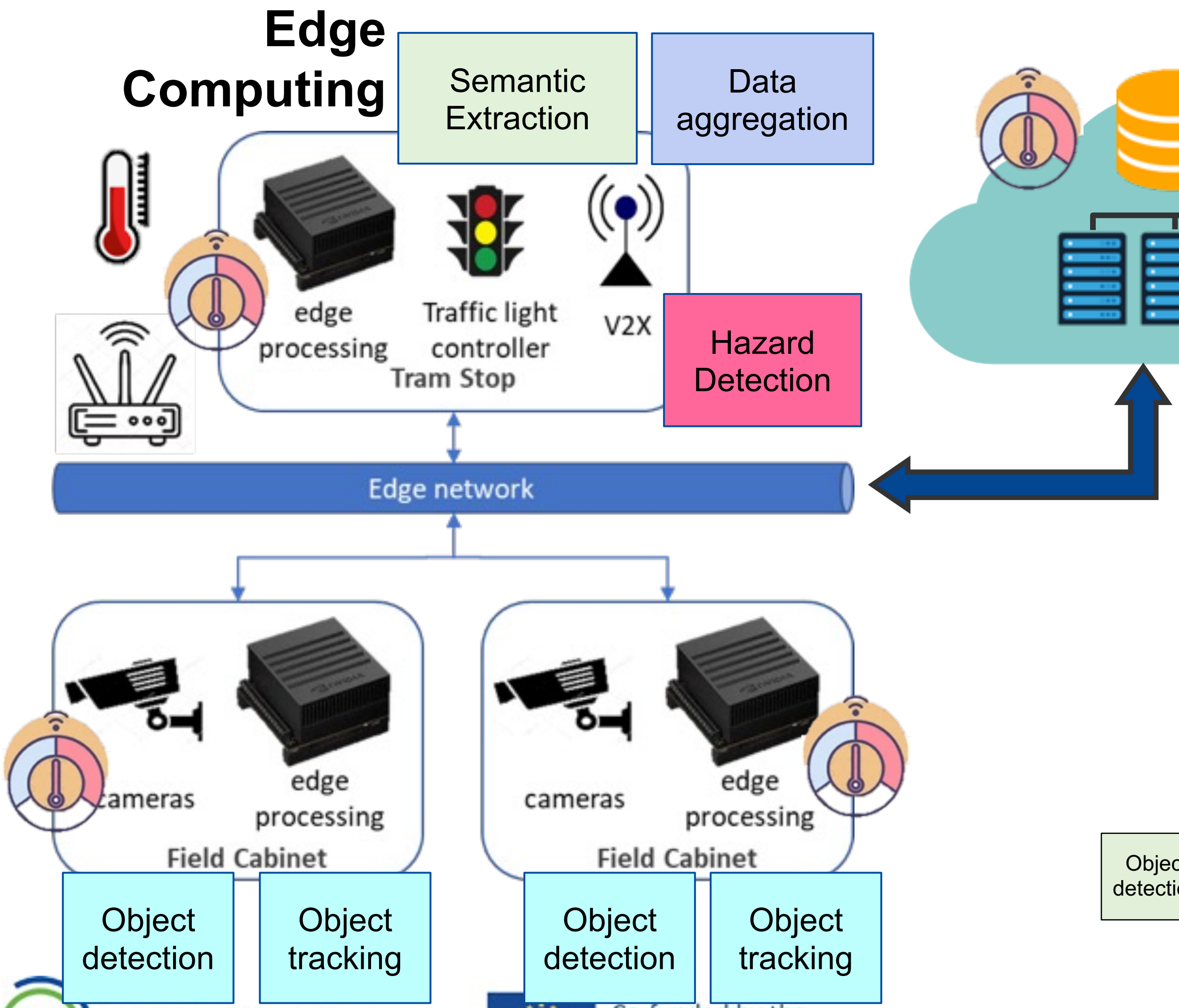
ELASTIC

A Software Architecture for Extreme-ScaLe Big-Data AnalyticS in Fog CompuTIng Ecosystems



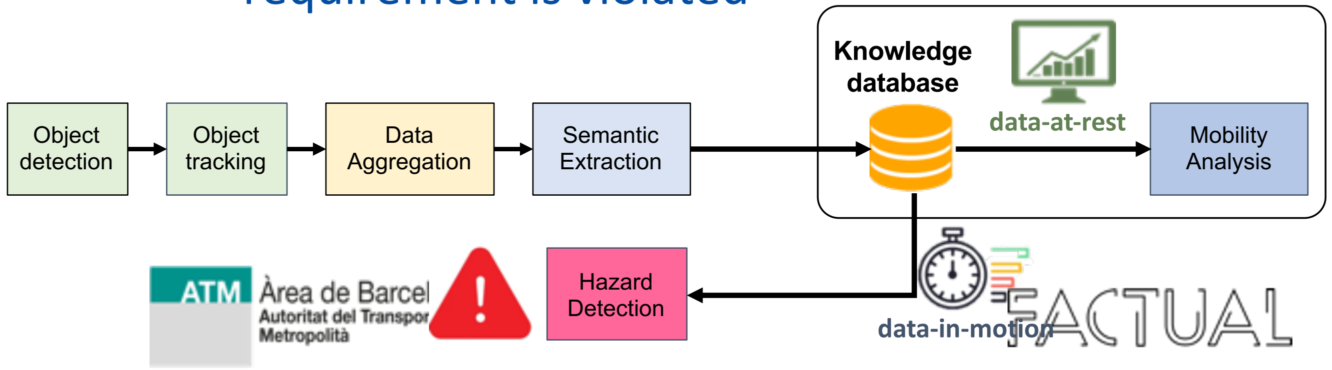



# Towards a unified and secured compute continuum





1. Automated **deployment** and **scheduling** of data-analytics
2. Constant **monitoring** of execution
3. **Re-scheduling** on-the-fly if a requirement is violated




Urban Mobility

Co-funded by the European Union

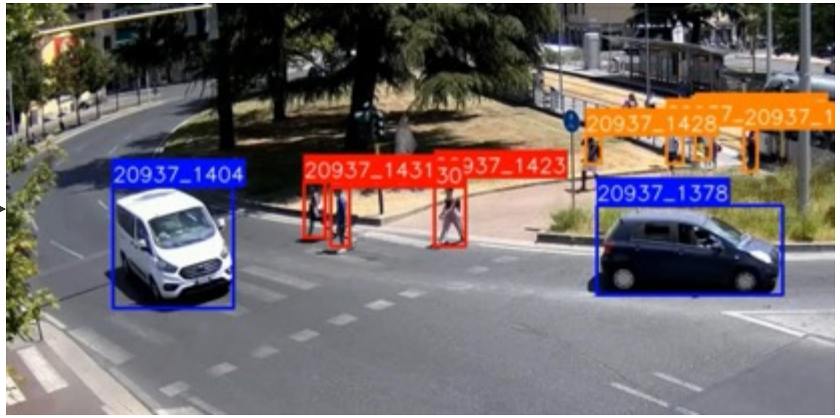
# Towards a unified and secured compute continuum

**Data Science**

**How many times pedestrians crosses in red?**


STREET CAMERA

real-time alert to vehicles




**Data Mining Workflow**

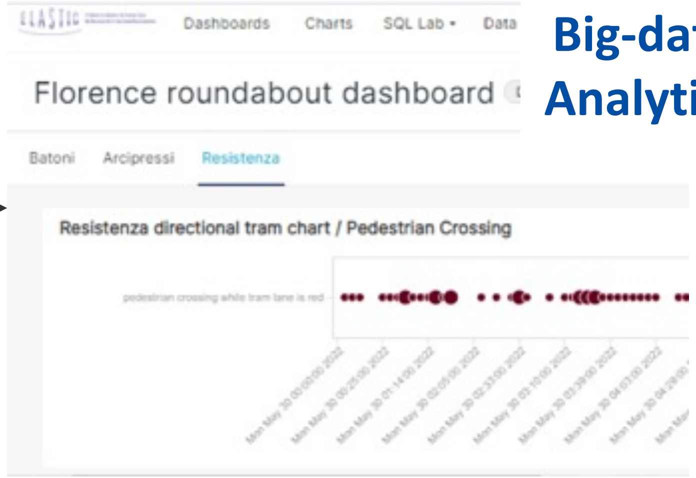
*Statistics (Kalman Filters)*

Object Detection → Object Tracking → Semantic Extraction → Hazard Detection: *PersonAtCrossWalk-SemaphoreRed*

***Supervised Deep Learning***

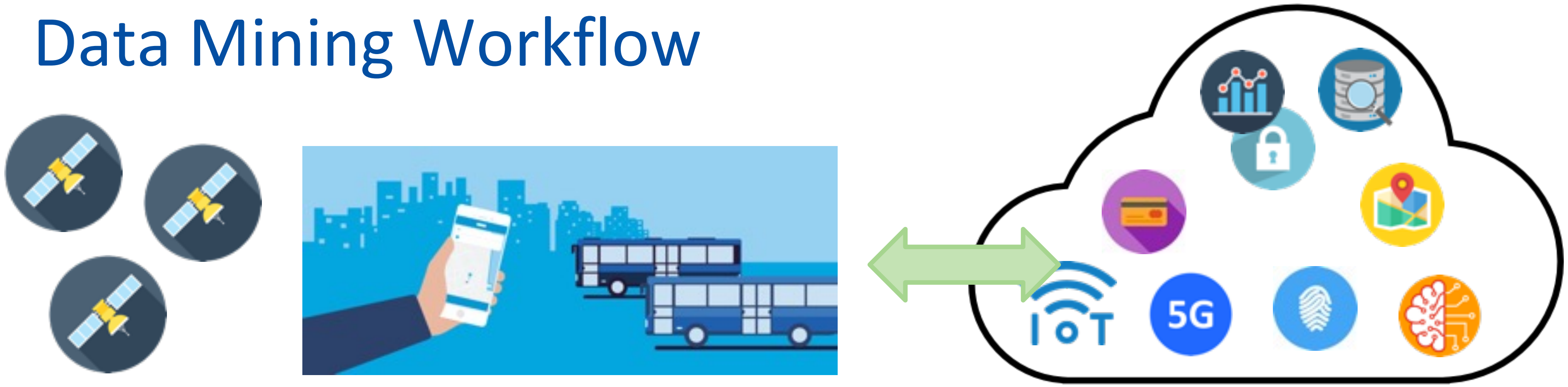
***Computer Vision***

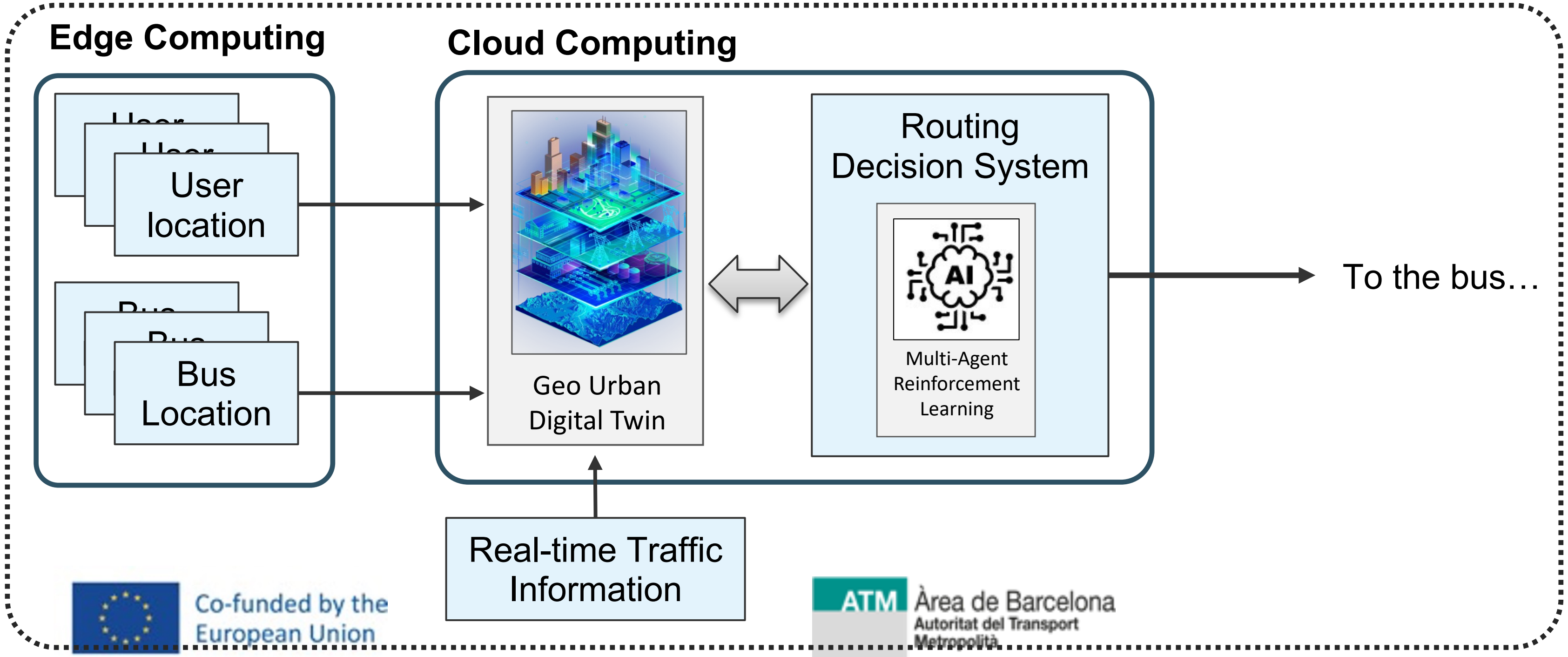
*Programming Logic*




**Big-data Analytics**

# On-demand bus service:
# Data Mining Workflow



# European Research Level

ELASTIC

A Software Architecture for Extreme-ScaLe
Big-Data AnalyticS in Fog CompuTIng Ecosystems

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

sixsq

CITTÀ METROPOLITANA DI FIRENZE

ikerlan GEST isep ICE Information Catalyst THALES

www.elastic-project.eu

CLASS

COORDINATING EDGE AND CLOUD
FOR BIG DATA ANALYTICS


Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Comune di Modena

IBM

MASERATI

UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

Atos

www.class-project.eu

eit Urban Mobility

Co-funded by the European Union

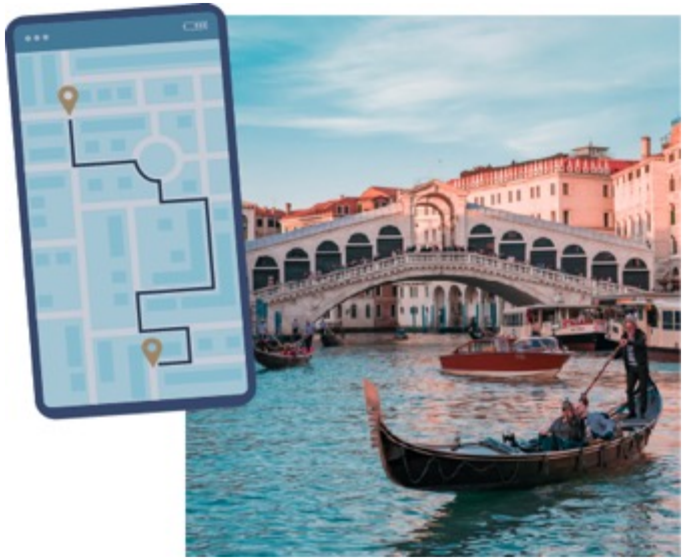
ATM Àrea de Barcelona
Autoritat del Transport Metropolità

FACTUAL

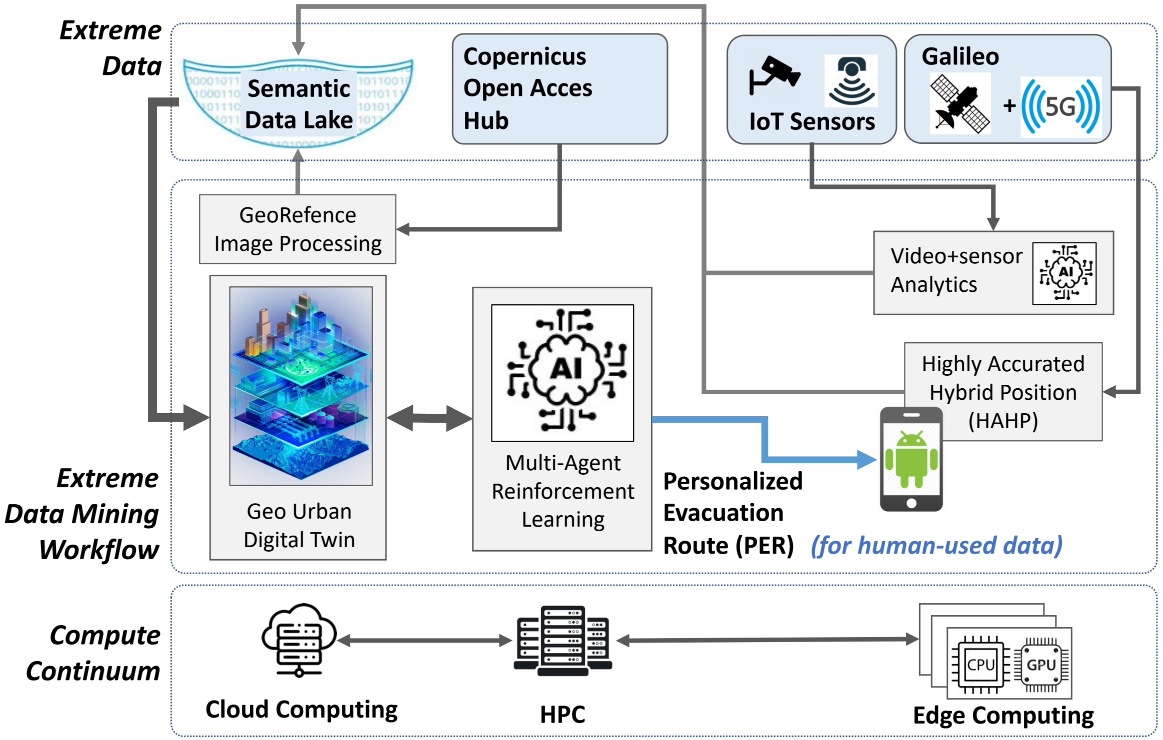
# European Research Level

A distributed data-mining software platform for extreme data across the compute continuum

https://extract-project.eu



Personalized Evacuation Routing System to guide citizens in the **city of Venice** through a safe route in real time



# Next Research Projects

## PROXIMITY (National Project)(2022 – 2025)

- Create an integrated 5G and edge ecosystem to facilitate real-time connectivity between end users, sensors and compute continuum to increase the performance of mobility services
- Real-time safety application for road users, focused around tramway intersections in Barcelona
- *Involved partners: BSC, Barcelona City Council and Tramway operator*

## AIRURBAN (National Project) (2022 – 2025)

- Improve accuracy and time-granularity of vehicle-related emission and air quality models, incorporating real-time traffic information
- Pilot infrastructure in Barcelona to monitor traffic around existing air quality monitoring stations, to develop and validate microscopic emission models and enhanced air quality prediction
- *Involved partners: BSC, Barcelona City Council and Generalitat*


eit Urban Mobility

Co-funded by the European Union

Co-funded by the European Union

# La Digitalització de la Mobilitat

Data Science i Intel·ligència artificial

Eduardo Quiñones
Barcelona Supercomputing Center (BSC)