



**DATA SHARING AND  
DATA INTEGRATION  
PPP NETWORKING SESSION**

EBDVF2017, November 22<sup>nd</sup> 2017

Tomás Pariente (ATOS) on behalf of the QROWD Consortium

This project has received funding from the European  
Union's Horizon 2020 research and innovation programme



# Motivation



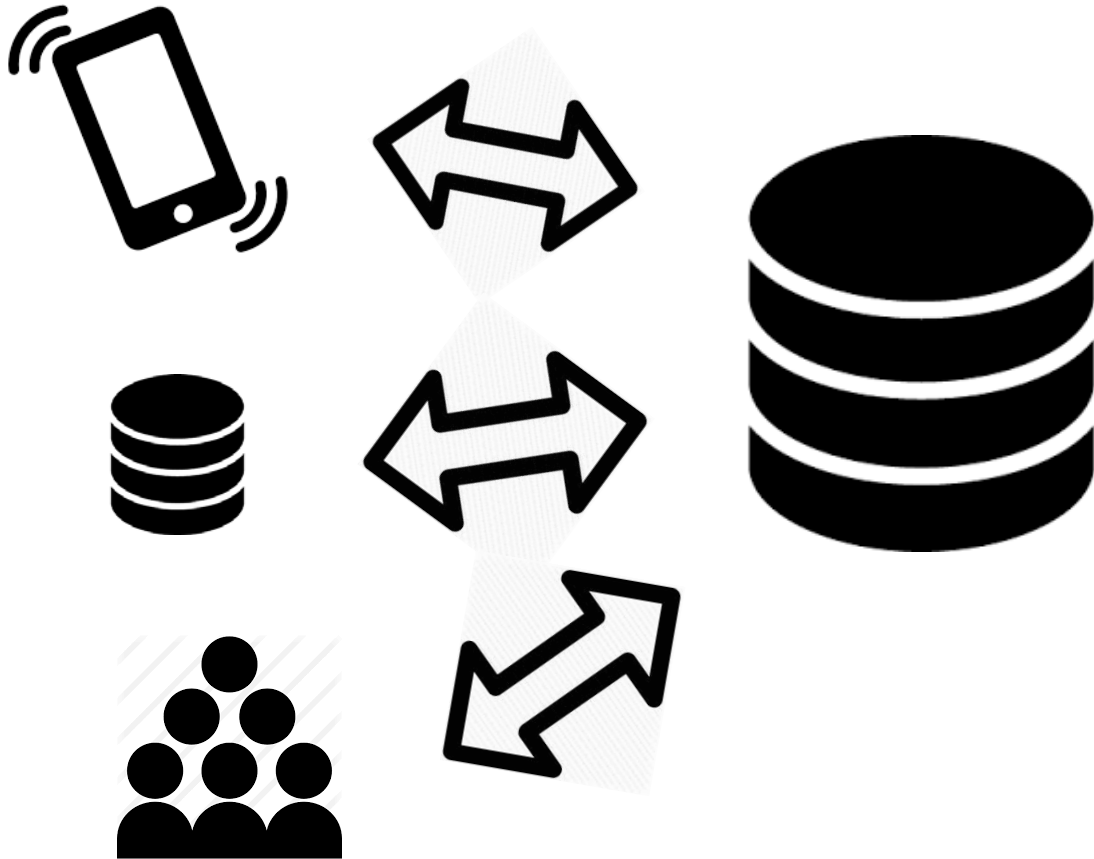
- The world in 2050
  - 9.8 billion people, around 70% living in cities
  - No combustion engine cars– sharing, on-demand, autonomous -
  - No driver license anymore, old jobs related to cars gone
  - Complete change of the city services and life-style
- But we are QROWD, and our horizon is 2020
  - Understand the way people move in cities
  - Helping to take decisions to improve mobility
  - Adding citizens and crowd workers to the loop of city services

# QROWD



Because data integration is  
humanly possible

# For data collection and integration



We deliver tools to:

- Collect data from engaged citizens (i-Log app)
- Integrate data into a data lake from disparate sources (Pentaho Kettle, Karma integration, URI-fication, RDF-ization)
- Data collection from crowd-workers

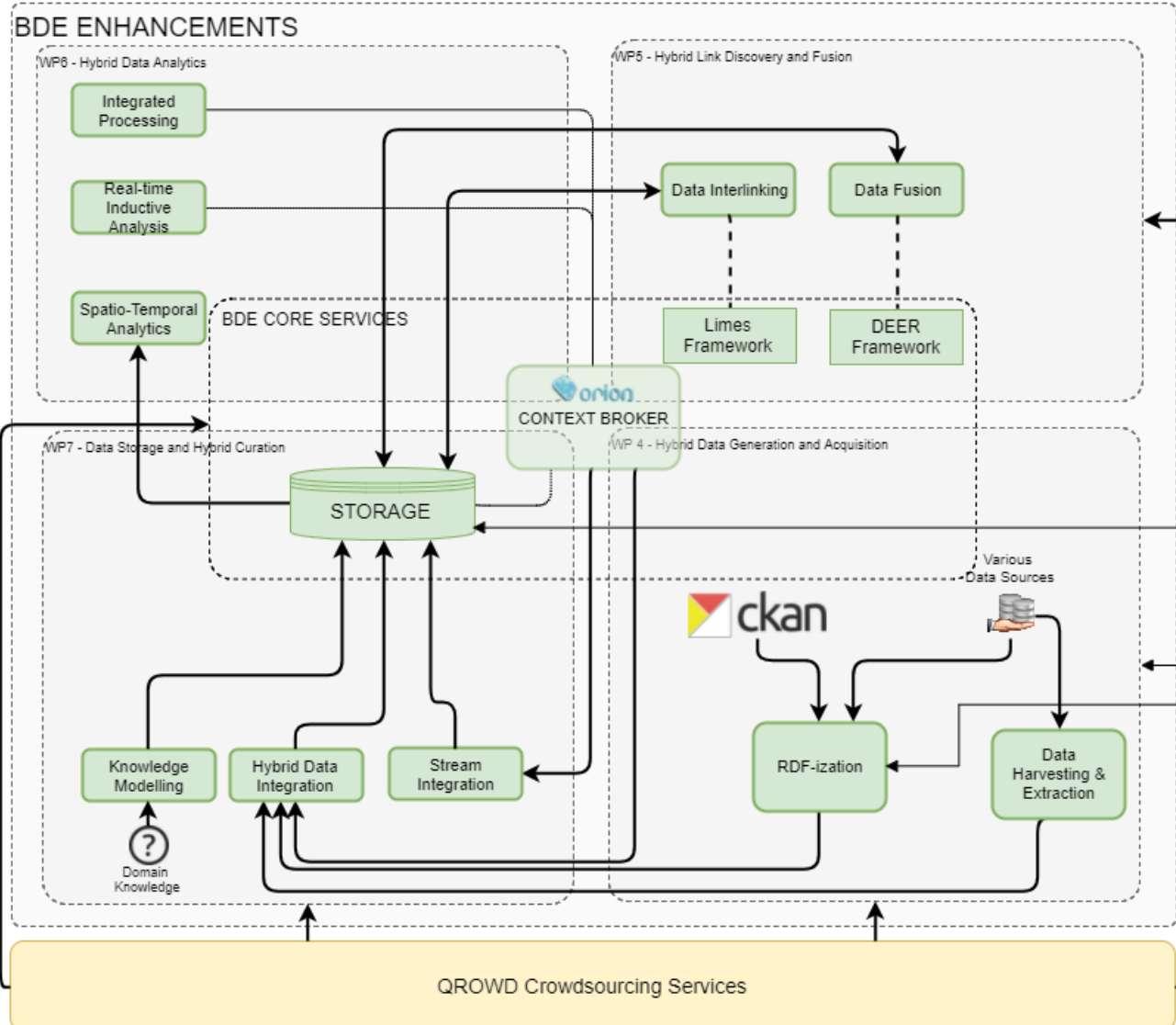
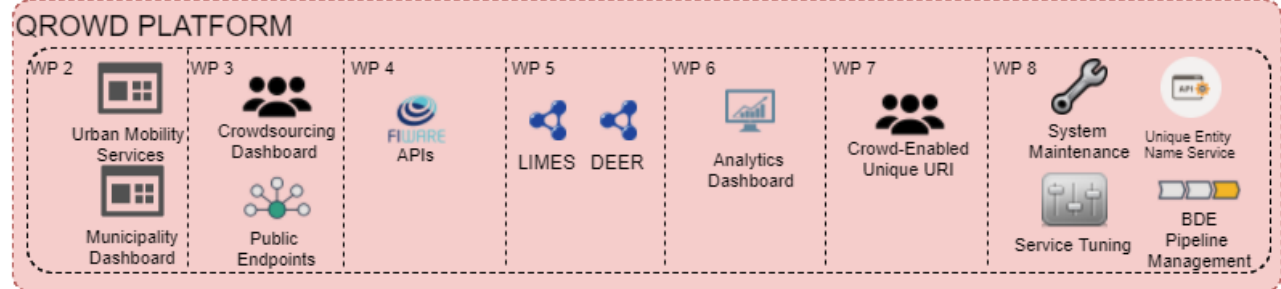
# Main data sources

Two main data providers: Municipality of Trento and TomTom

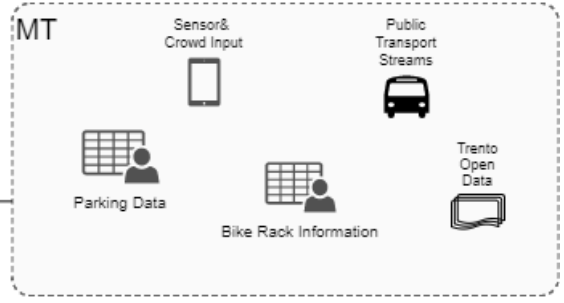
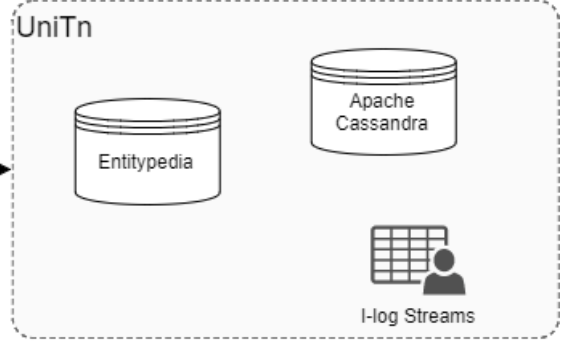
- Open data from Trento Municipality and Region (i.e. city–bus stops, bike racks...-, )
- Stream data (often managed by 3<sup>rd</sup> parties) from some urban traffic sensors and data streams in Trento (i.e. parking occupation, bus positions...).
- Historical data about mobility in the area from Trento and TomTom (i.e. several years of TomTom devices in the area, weather conditions,
- Crowdsourced data

More info in

D4.1 Data Catalog and D1.x and D2.x series



# Architecture view



Using standards and existing tools as much as possible.

- FIWARE and OASC-compliant
- Big Data Europe
- CKAN, RDF

# Human factor on urban mobility Data Value Chain

## Data Acquisition

- **Getting data from citizens**, i.e. by completing data infrastructure (locating bike racks) and measuring occupancy of parking groups that are not

## Data Analysis

- By helping in the **training phase**: i.e. citizens with an extra incentive can provide data with more frequent and precise labels, that can be used as training sets for machine algorithms
- **Confirming the predictions** of transportation mode made by the machine, opening the door to improving the quality of the machine prediction on the go

## Data Curation

- **Curate data on mobility infrastructure**. i.e. in Trento, data about disabled parking spots is incomplete, needs to be verified and curated

## Data Linking and Integration

- Detection of **inconsistencies**: i.e. when trying to merge several data sources
- **Entity resolution, disambiguation, missing data**: i.e. a mobility point appears at one position in the council's data, in a slightly different in Open Street Map, not at all on TomTom's map

# Example of crowdsourcing for Data generation, acquisition, and sharing

- Streetview experiment



Do the images contains a bike rack?  Yes  
 No

Which type of bike rack?  Type A  
 Type B  
 Type C

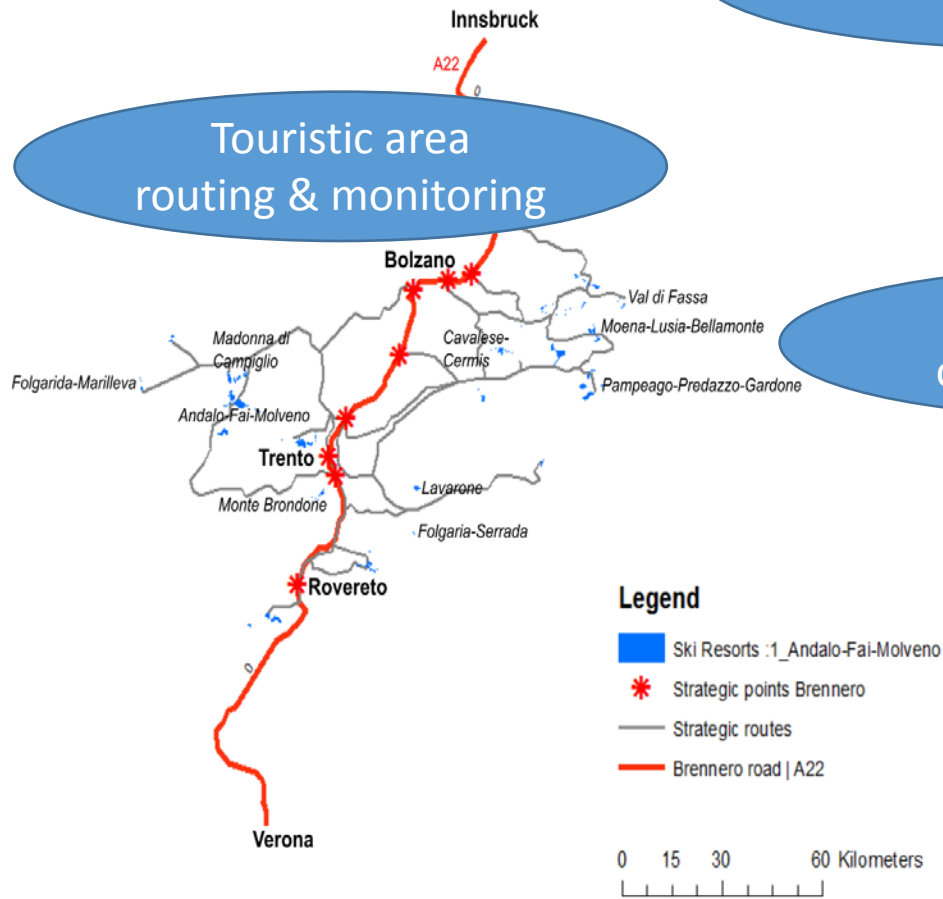
How many bikes can be hosted?  1 - 5  
 6 - 10  
 11 - 20  
 21 +

What	Who	Why	How
Information finding	Crowd-workers	Economic	Collection



# Two integrated business cases

## Municipality of Trento and TomTom



Touristic area  
routing & monitoring

Historical analysis

Tom Tom City Trento

On-street and  
off-street parking

Completing city  
infrastructure data

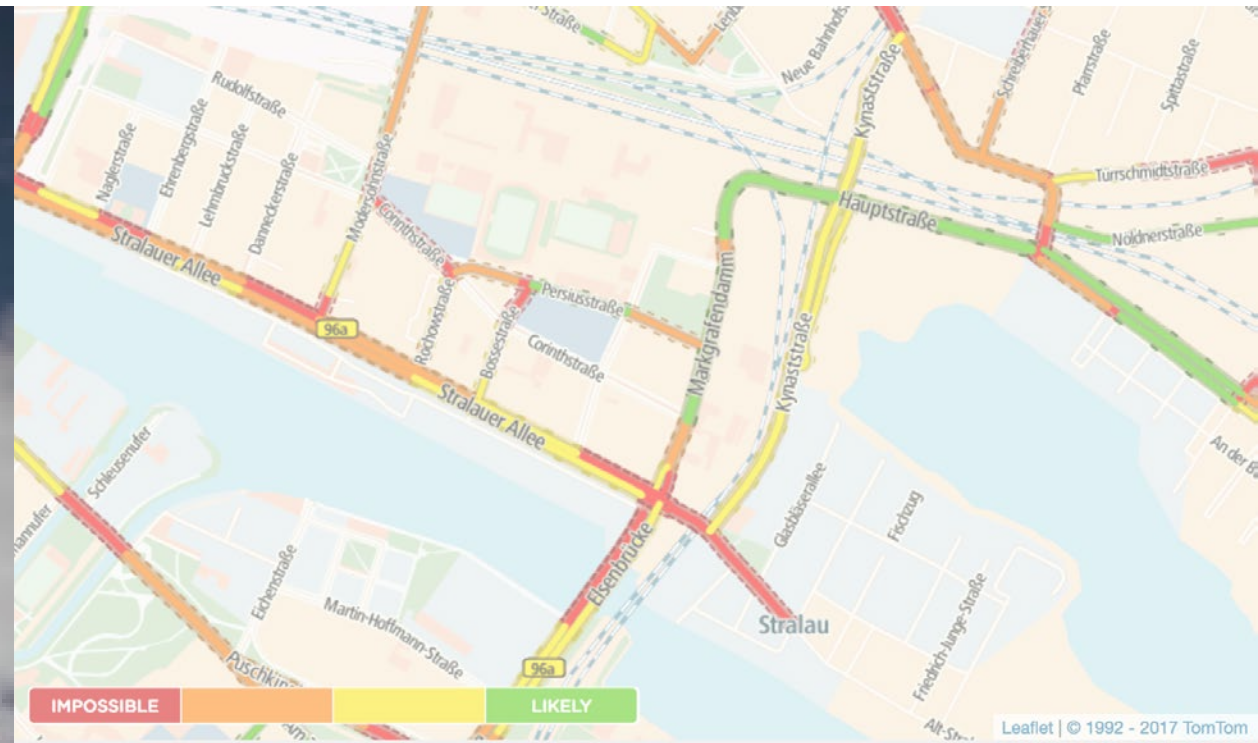
Modal Split



Parking  
probabilities

# Parking probabilities use case

- I'd like to go to the town hall Monday 10am, can I find parking nearby?
  - Car, scooter, or bike
- What is the probability of finding a spot at a given time for a given target location?



# Parking probabilities use case data

## On street paid parking

- Position of parking spaces
- Parking meters stream data
- TomTom historical data

## On street free parking

- TomTom historical data
- Crowd-sense occupancy
- Re-use surveillance cameras

## Bicycle parking

- Bike racks positions
- Bike share API

+ Crowdsourced data

# Data sharing win-win

**Opening models for collaboration between local authorities (i.e. Trento Municipality) and private mobility operators (i.e. TomTom)**

- Council gets extra services: parking probability, visualization, historical data, routing
- Private mobility data operator gets: cleaner sources, assessment of service, exploration new market
- Commercial partners: Integration and consultancy services

# Wrap-up

- Integrating citizens and humans in the loop of urban mobility services
- Create shared data space between public and private actors
- Ongoing work
  - Quantifying the value for each actor to define exploitation model (probably use-case based)
  - Best practices and guidelines for crowdsourced data integration

# Thanks!

Tomás Pariente

tomas.pariantelobo@atos.net

M+34 630966590

@tpariente