TEADAL

Pilot 5 - Regional planning for environmental sustainability

TUSCANY REGION, BOX2M

2nd GA

DENIA, MARCH 2023

WWW.TEADAL.EU

Involved partners and datasources

The pilot involves Tuscany Region (RT, public administration) and BOX2M (private company). Also open data provided by other Tuscan organizations (SIR, ARPAT) are used.

SIERT
Sistema Informativo efficienza Energetica Regione Toscana

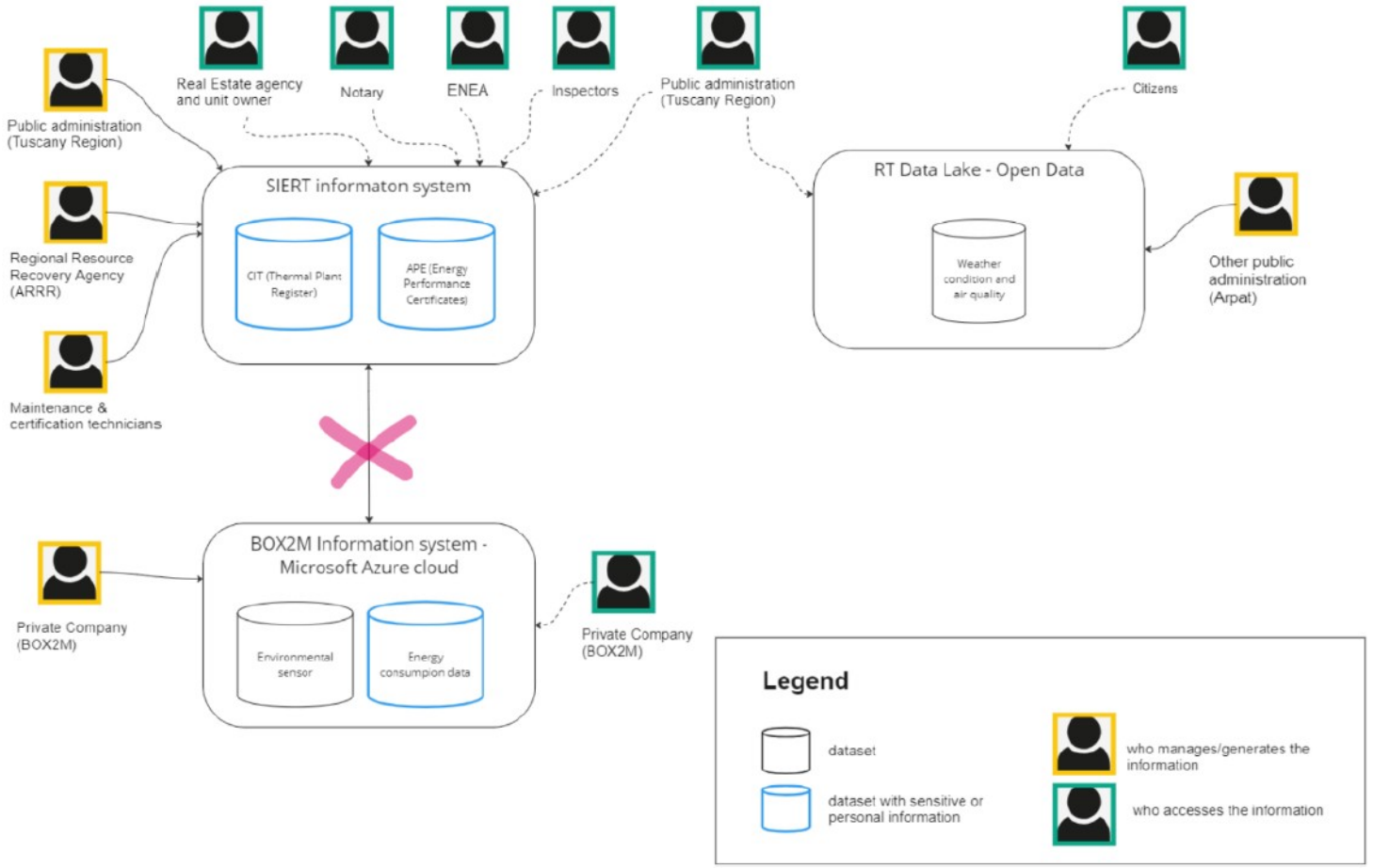
Building's energy performance certificates (APE) and building's thermal and air conditioning systems registry (CIT) managed by Regional Resource Recovery Agency (ARRR) S.p.A.

Open data on weather conditions provided by Regional Hydrological Service (SIR).

Open data on air quality provided by Regional Agency for Environmental Protection of Tuscany (ARPAT).

Electrical energy related datasets gathered from consumers entities (public buildings, charging stations, lightning infrastructures) together with datasets gathered from micro-production entities (solar roof, building cogeneration power plants)

AS-IS Scenario



Currently data are not shared and not enhanced from public interest perspective.

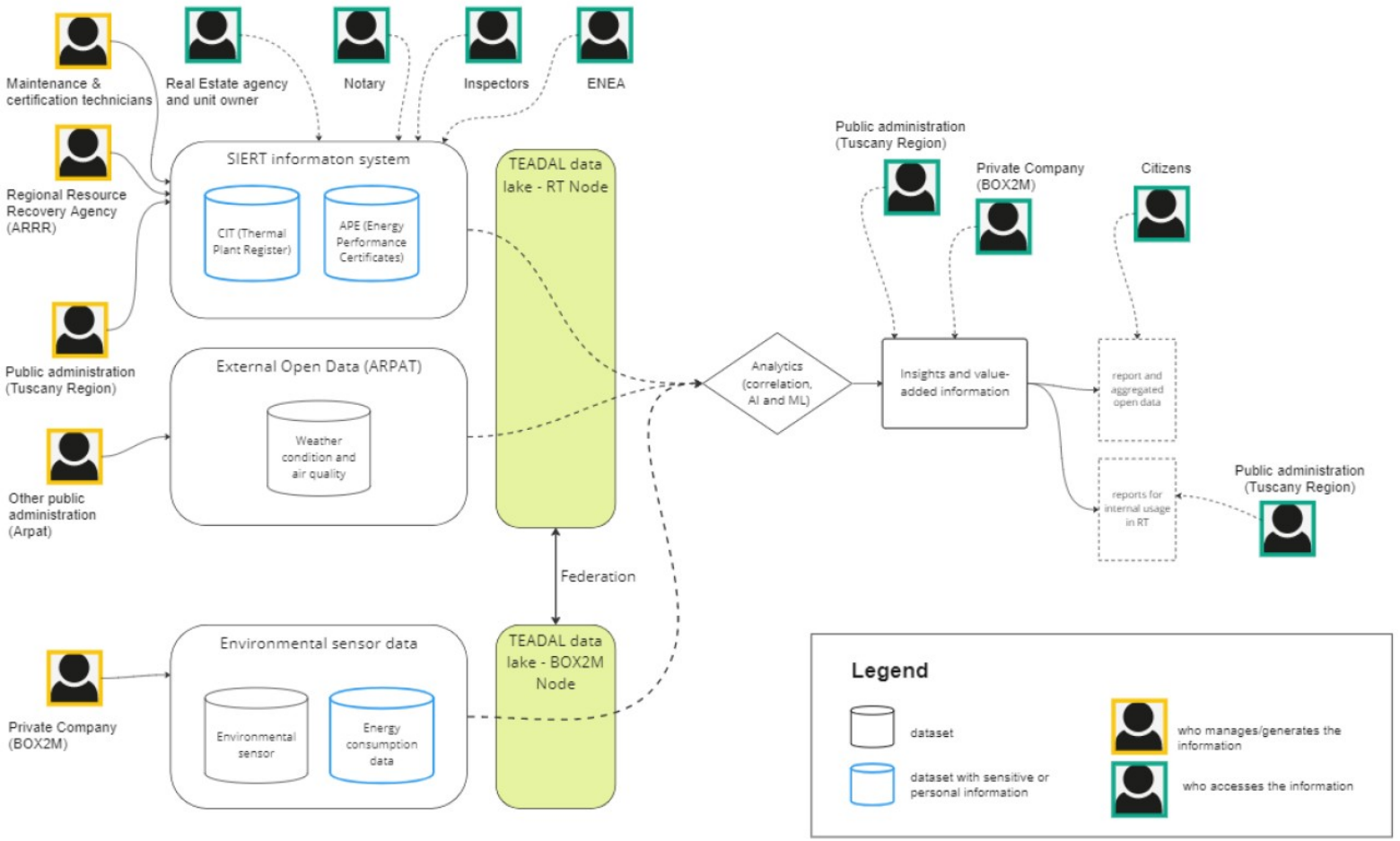
RT cannot share building energy profiles on its own with a private company (BOX2M in this case) for legal reasons.

BOX2M isn't able to physically share data gathered by sensors due to their huge volumes and velocity

Data sharing requires formal agreement.

Sharing data on private buildings sensors should be authorised by the building owner.

TO-BE Scenario



The TO-BE solution consists in creating a **data lake federation** where data feeds a virtual data mart supporting the analysis.

The architecture should include two data lakes, one for RT and one for BOX2M.

The goal is to correlate building energy profiles managed by RT with sensors data deployed by BOX2M and further data freely available.

Data sharing and use have to be forbidden or limited in case of personal, sensitive or reserved commercial data.

Goal of the use case

Discover correlation between building energy profiles and dynamic data coming from sensors measuring environmental parameters and energy consumption.

Make territorial analysis on buildings energy behaviour related both to building energy profiles and context data about weather conditions and air-quality.

Understand territorial patterns and support decisions on environment policies.

The business purpose is to use static and dynamic information of each building or area to reduce energy consumption, negotiate prices in bulk and decrease pollution.

Also impacts on data management energy efficiency has to be evaluated, compared with transferring and duplicating the same data storage between different organisations.

Data models and synthetic data generation

Open data

Weather conditions data of concrete interest are historical records of maximum and minimum temperatures across the multiple subregions and municipalities of the Tuscany region from 2009 to 2023, in a structured JSON format.

Open daily air quality values in some Tuscany specific subregions are provided. The downloaded measurements include PM10 and PM2.5 daily concentration values from 2018 to 2023, in CSV format.

There is no need for data synthesis, as the information is easily crawled and downloaded

Close data

Thermal and energy performance certificates databases include information about the combustion and energy efficiency, or rating, of registered buildings, plants and thermal units. This data is not publicly accessible, but presents a defined structure and comprehensible schemas. All this data shall be synthesised.

Environmental and energy consumption datasets are related to addresses, timestamps and the sensor-measured values. Data will be provided by experimental API in a structured JSON format.

Focus on CIT dataset

used to segment visualization with other datalake (e.g. Florence datalake can visualize data limited to its territory)

Field Name	TIPO (format)	Permited values – coding	Description	NOTES
cadastre_code	character varying(10)	-	codice regionale a 10 cifre, già presente sul cit, indentifica in modo univoco l'impianto	PK
thermal_unit	character varying(4)	-	format a 4 cifre (ESEMPIO GT01). INSIEME AL CODICE CATASTO IDENTIFICA in modo univoco IL GENERATORE. il GRUPPO TERMICO è Già assegnato dal CIT	PK
plant_address	character varying(400)	via, numero, palazzo, scala, INTERNO	include via, numero, palazzo, scala, INTERNO dove è ubicato l'impianto, separati dalla virgola e riportati solo se presenti (esempio: VIA ROMA, 5)	GDPR
municipality	character varying(6)	-	codice istat a 6 cifre del comune in cui è ubicato l'impianto	OK, SEGM
installation_date	date	-	FORMAT AAAA-MM-GG	YYYY-MM
check_date	date	-	FORMAT AAAA-MM-GG, non saranno accettati rcee passati 60 giorni dalla loro redazione	YYYY-MM
remarks	character varying(400)	-	sezione osservazioni del rcee	NO
recommendations	character varying(400)	-	sezione RACCOMANDAZIONI del rcee	NO
prescriptions	character varying(400)	-	sezione PRESCRIZIONI del rcee	NO
arrival_time	character varying(5)	-	FORMAT HH:mm (esempio: 09:15)	NO
departure_time	character varying(5)	-	FORMAT HH:mm (esempio: 09:15)	NO
manufacturer	character varying(500)	-	indicare il fabbricante del generatore	COMM
template	character varying(500)	-	INDICARE IL modello DEL GENERATORE	COMM
freshman	character varying(200)	-	INDICARE la matricola del generatore	COMM
plant_manager	character varying(200)	Cognome Nome	Cognome e Nome del responsabile impianto SEPARATI DA UNO SPAZIO (ESEMPIO: ROSSI MARIO)	GDPR
fiscal_Code	character varying(16)	-	codice fiscale del responsabile impianto	GDPR
business_name	character varying(200)	-	EVENTUALE RAGIONE SOCIALE QUANDO L'impianto e' di responsabilità di una persona giuridica	GDPR
VAT_number	character varying(11)	-	EVENTUALE partita iva QUANDO L'IMPIANTO E' DI RESPONSABILITÀ DI UNA PERSONA GIURIDICA	GDPR
total_power	double precision	-	potenza nominale del generatore. separatore decimale punto	OK
declaration_conformity	integer	SI=3 4=NO	applicare la codifica, assegnando il valore 3 quando la dichiarazione di conformità è uguale a si ed il valore 4 quando è no	OK
booklet_use_maintenance	integer	SI=3 4=NO	APPLICARE LA CODIFICA IN ANALOGIA A QUANTO VISTO PER LA DICHIARAZIONE DI CONFORMITà	OK
booklet_air conditioning	integer	SI=3 4=NO	APPLICARE LA CODIFICA IN ANALOGIA A QUANTO VISTO PER LA DICHIARAZIONE DI CONFORMITÀ	OK
booklet_air conditioning_completed_correctly	integer	SI=3 4=NO	APPLICARE LA CODIFICA IN ANALOGIA A QUANTO VISTO PER LA DICHIARAZIONE DI CONFORMITÀ	OK
water_hardness	INTEGER	0 - 99 °fr	sono ammessi solo VALORI INTERI DA 0 A 99 °fr	OK
heating_treatment	integer	20 = Condizionamento chimico	APPLICARE LA CODIFICA, riportata a fianco, al campo indicante la tipologia di trattamento acqua sul circuito di riscaldamento	OK
acs_treatment	INTEGER	20 = Condizionamento chimico	APPLICARE LA CODIFICA, RIPORTATA A FIANCO, AL CAMPO INDICANTE LA TIPOLOGIA DI TRATTAMENTO ACQUA SUL CIRTCUITO PER la produzione di acqua calda sanitaria	OK

… 67 fields

primary key

filed access must be limited to a year-month format

not useful

protected due to commercial restrictions

GDPR protected

useful for analysis

Computations and insights to be performed

- Define and calculate common building and geographical area identifiers among each static, dynamic and open dataset
- Join CIT, APE, sensors data and open data by the building identifiers and/or by geographical area identifiers that has been previously defined
- Return a dashboard with dynamic choropleth and time series charts based on main available metrics
- Calculate a statistical model in which, for example, both dynamic energy metrics and air-quality metrics depends on buildings energy profiles and weather conditions

[RT_CIT].[plant_address] (dimension)
[RT_CIT].[municipality] (dimension)
[RT_CIT].[combustion_efficiency] (measure)

[RT_APE].[address] (dimension)
[RT_APE].[municipality] (dimension)
[RT_APE].[energy_rating] (measure)

[SIR].[municipality] (dimension)
[SIR].[date] (dimension)
[SIR].[value] (measure)

[ARPAT].[municipality] (dimension)
[ARPAT].[observation_date] (dimension)
[ARPAT].[PM10] (measure)
[ARPAT].[PM2dot5] (measure)

[BOX2M].[locationName] (dimension)
[BOX2M].[municipality] (dimension)
[BOX2M].[timestampLocal] (dimension)
[BOX2M].[value] (measure)

Main goals and KPIs

P6_G1: evaluation of the impact of the solution in terms of management energy efficiency

P6_G2: reduction of the time needed to integrate new data sources and perform analysis

P6_G3: reduction of the processing costs for data storage and for performing analysis

P6_G4: evaluation of solution in terms of security

KPI code	Category	Related Goals	Description
P6-KPI1	Time to market	P6_G2	Time (in days) to perform analysis (retrieve necessary datasets and integrate them) in absence and in presence of Teadal.
P6-KPI2	Data storage	P6_G3, P6_G1	Data storage space (e.g. in Gb) necessary to replicate data useful to perform analysis in presence and in absence of Teadal.
P6-KPI3	Data storage	P6_G3, P6_G1	Number of records to be stored in absence and in presence of Teadal.
P6-KPI4	Data storage	P6_G3	The costs in euro to be sustained in order to maintain data storage in absence and in presence of Teadal.
P6-KPI5	Processing cost	P6_G1	Number of procedures to be built and maintained to ingest and process all necessary data in absence and in presence of Teadal.
P6-KPI6	Security	P6_G4	According to DPIA strategy, changes in risk gravity/impact and risk probability in absence and in presence of Teadal.

TEADAL

THANKS

TEADAL.EU

@TEADAL_eu

@TEADAL

TEADAL project is funded by the EU's Horizon Europe programme under Grant Agreement number 101070186