TAQE: a Data Modeling Framework for Traffic and Air Quality Applications in Smart Cities

David Martínez[1[0000−0001−6912−9433]], Laura Po[2[0000−0002−3345−176X]], Raquel Trillo-Lado[3[0000−0001−6008−1138]], and José R.R. Viqueira[1[0000−0002−1539−3746]]

[1] CiTIUS, University of Santiago de Compostela, Spain
{david.martinez.casas,jrr.viqueira}@usc.es
[2] University of Modena and Reggio Emilia, Italy
laura.po@unimore.it
[3] I3A, University of Zaragoza, Spain
raqueltl@unizar.es

Abstract. Air quality and traffic monitoring and prediction are critical problems in urban areas. Therefore, in the context of smart cities, many relevant conceptual models and ontologies have already been proposed. However, the lack of standardized solutions boost development costs and hinder data integration between different cities and with other application domains. This paper proposes a classification of existing models and ontologies related to Earth observation and modeling and smart cities in four levels of abstraction, which range from completely general-purpose frameworks to application-specific solutions. Based on such classification and requirements extracted from a comprehensive set of state-of-the-art applications, TAQE, a new data modeling framework for air quality and traffic data, is defined. The effectiveness of TAQE is evaluated both by comparing its expressiveness with the state-of-the-art of the same application domain and by its application in the "TRAFAIR – Understanding traffic flows to improve air quality" EU project.

Keywords: Conceptual Modeling · Smart City · Environmental Data · Air Quality Data · Traffic Data.

1 Introduction

In urban areas, traffic-related pollution is a crucial problem with high environmental impact since more than 40% of emissions of nitrogen oxides come from traffic. As air pollution entails damages on ecosystems, reducing pollutant emissions to the atmosphere is imperative. Environmental decision making in smart cities has to be based on sophisticated traffic and air quality monitoring and modeling infrastructures, which generate large amounts of data, which are complex both in structure and also in semantics [26].

Many specific data modeling solutions and ontologies are being used in different applications to represent such data in smart cities. The lack of a common standardized data modeling framework adds a critical barrier to the data integration [30,31] required for cross-city decision-making. Besides, and more importantly, it does not foster the reuse of conceptual modeling structures. Knowledge and data model reuse reduce development time and resources required for projects. They also facilitate communication among the people involved in the information systems development process and in their maintenance. Furthermore, it has been shown that although a higher level of abstraction increases reusability, specific models are more usable [18].

Based on the above, this paper defines TAQE, a conceptual modeling framework for the representation of air quality and traffic data generated by related smart city monitoring and modeling infrastructures. TAQE is embedded in an information architecture with four levels of abstraction, which is based on international standards of the Open Geospatial Consortium (OGC), the International Organization for Standardization (ISO) and the World Wide Web Consortium (W3C). The design of TAQE is based on a collection of requirements extracted from a comprehensive set of applications. The effectiveness of its data representation capabilities has been tested by two different means. First, TAQE and all the already existing models of a similar level of abstraction found in the literature were analyzed to check the fulfillment of the identified requirements. Second, TAQE was specialized to design the data model needed by the "TRAFAIR – Understanding traffic flows to improve air quality" EU project, to support the representation of its air quality, traffic observation and modeling data.

The paper is organized as follows. Section 2 analyzes well-known data models and ontologies useful in this context, defines the four-level information architecture considered in this proposal, and analyzes different use cases. Section 3 describes the set of requirements that air quality models and traffic models should meet. Section 4 and Section 5 are devoted to the definition of the proposed data modeling framework. The results of the evaluation of TAQE are shown in Section 6. Finally, some conclusions are presented in Section 7.

2 Data Models for Environmental Data and Smart Cities

Data models and semantics are a key aspect of the representation and enhancement of environmental and smart city data. This section classifies available data models into four levels of abstraction, starting from generic frameworks and data models and reaching specific applications.

2.1 Level 1: Data Modeling

A great amount of data infrastructures on the Web are based on Linked Open Data (LOD) best practices [9] and Resource Description Framework (RDF)[4] to facilitate information data integration and interoperability. The vocabularies used in RDF statements to identify objects and properties may also be defined

[4] W3C RDF Prime: `http://www.w3.org/TR/2004/REC-rdf-primer-20040210/`.

in RDF with RDF Schema[5] primitives. The expressiveness of RDF Schema to define vocabularies is extended by the Web Ontology Language (OWL)[6].

So, at this level, several well-known general ontologies, such as the W3C Provenance Ontology (PROV-O), are considered relevant for traffic and air quality applications in smart cities. PROV-O allows for representing provenance information (i.e., information about entities or organizations, activities and people involved in producing data), which can be used to form assessments about the quality, reliability or trustworthiness of the data produced in a variety of application domains by using the PROV Data Model[7], which has a modular design and three main entities: *Entity*, *Activity* and *Agent*. PROV Data Model considers entities and activities, and the time at which they were created, used, or ended; derivations of entities from other entities; and agents bearing responsibility for entities that were generated and activities that happened. Agents can form logical structures for their members.

2.2 Level 2: Earth Observation and Modeling and Smart Cities

Regarding sensor data widely used in Earth observation, the most popular ontology is the W3C Semantic Sensor Network Ontology (SSN) [10], which defines a vocabulary to describe sensors and their observations, including both observed values and required metadata of features of interest, observed properties, etc. SSN may be used to annotate sensors and observations, in a way aligned completely to the OGC and ISO Observations and Measurements (O&M) [11] standard, which provides a conceptual schema to represent the results of observation processes and the metadata of these processes, of the entities that are sampled by them (Features of Interest), and of the observation results obtained. Moreover, SSN is also aligned to the Extensible Observation Ontology (OBOE), and PROV Data Model considered in level 1 [21]. The lightweight core of SSN is the SOSA (Sensor Observation, Sample and Actuator) ontology. The core concepts or entities of SOSA are *Procedure*, *Sensor*, *ObservableProperty*, *Observation*, *FeatureOfInterest* and *Result*.

2.3 Level 3: Application Domains

At this level, we focus our attention on two related domains: Air Quality and Traffic Data.

Air Quality Urban air pollution information is usually processed by specialists to monitor, predict, and study air pollution sources within the urban area. Several platforms display statistics or semi-real time air quality measures, like World's Air Pollution: Real-time Air Quality Index[8] and Air quality statistics by

[5] W3C RDF Schema: `http://www.w3.org/TR/rdf-schema/`.
[6] W3C OWL 2: `https://www.w3.org/TR/owl2-overview/`.
[7] PROV Data Model: `https://www.w3.org/TR/prov-dm/`.
[8] Real-time Air Quality Index: `https://waqi.info/`

European Environmental Agency(EEA)[9]. Air quality (AQ) monitoring is usually done using professional AQ monitoring stations; Environmental Agencies are in charge of analyzing AQ conditions and reporting the violations of the concentrations limits. In recent years, it has been explored the potential of low-cost environmental sensors for urban air pollution monitoring. Low-cost sensors can be placed in fixed positions or on vehicles or drones. Thus, they can provide gases concentration in a location, on a path or in a 3D environment. Different studies have taken advantage of a variety of satellites to estimate different air emissions. In addition, air quality models are studied for the representation of observation data in the scope of air quality prediction applications. Air pollution modeling is a complex subject and is linked to 3D city models, meteorological elements, and air tainting information. 3D air pollution models and the associated simulation systems are those that aim at "reconstructing" the environment, its properties, and governing physical laws.

Several ontologies have been defined in a specific context to enrich the air quality measurements and simulations semantically. AIR_POLLUTION_Onto ontology [27] has been conceived for air pollution analysis and control in two case studies. Airbase is the European air quality dataset maintained by the EEA [17]. QBOAirbase, a provenance-augmented version of the Airbase dataset, is multidimensional dataset linked to the Semantic Web. hackAIR ontology has been created within the "hackAIR - Collective awareness platform for outdoor air pollution" EU project [32]. It can store observations from sensors, monitoring stations, AQ related values from fused or forecasted data, or from sky-depicted images, etc.

Models that are used for the study of air quality at an urban scale make use of a grid-based spatial resolution. An ontology of air quality models is defined in [25]. It is linked to 3D city models and other models related to sustainable development ad urban planning. This ontology is devoted to representing pollutant dispersion in urban street canyons, a particular phenomenon that happens in some street bordered by buildings in a specific configuration.

Traffic Data Traffic data plays a key role in a smart city, since enabling efficient transportation and sustainable mobility are important goals that allow enhancing the quality of life of citizens. Therefore, modeling traffic and sharing traffic data is very relevant for many cities [36,13]. Thanks to the collection and modeling of traffic data, public administrations can make informed decisions regarding mobility policies. Besides, offering traffic information to citizens can also help to raise awareness about the importance of choosing suitable mobility options to increase their well-being and reduce pollution. A wide range of different types of sensors can be used to measure traffic [20], such as inductive loops, microwave radars, and video image detection. Moreover, traffic models are studied in order to simulate and predict traffic flows.

[9] Air quality statistics by EEA: `https://www.eea.europa.eu/data-and-maps/dashboards/air-quality-statistics`

The Vocabulary to Represent Data About Traffic[10] has been proposed for the representation of the situation of traffic in a city. This vocabulary extends SSN [10] to represent the intensity of traffic in the different road segments of a city. In particular, it is used to represent road segments, traffic observations, sensor or sensing systems used to obtain a specific measurement, the results of observations (which have values and are produced by a specific sensor or sensing system), and finally instances that represent the type of properties being measured. Authors of this work recommend using this vocabulary in conjunction with the vocabulary available on `http://vocab.linkeddata.es/datosabiertos/def/urbanismo-infraestructuras/callejero` in order to represent city road maps.

An Ontology Layer for Intelligent Transportation Systems in order to increase the traffic safety and improve the comfort of drivers is proposed in [16]. The ontology layer is composed of three groups of interrelated concepts: concepts related to vehicles, concepts related to roads, and concepts related to sensors. The concepts related to vehicles describe a taxonomy of vehicles of different types and also allow representing information about their routes and locations. The concepts related to the infrastructure include a taxonomy of different types of roads as well as the representation of other parts of the infrastructure, such as the road segments, traffic lights and traffic signs, lanes, road markings (e.g., painted arrows), tunnels, parking slots, roundabouts, bridges, gas stations, and toll stations). Finally, the concepts related to sensors are based on the use of the SSN ontology. The previous work focuses on long-life elements of the roads, such as traffic signs or road segments, while in the open511 specification[11] and the Road Accident Ontology[12], special situations in roads, such as accidents, special events (e.g., a celebration of a sport event) or particular weather conditions and road conditions (e.g., snow, ice, or fire on the road), are considered.

2.4 Level 4: Use-case specific applications

Several projects have exploited drones for AQ monitoring. Drones can quickly cover vast industrial or rural areas obtaining a complete and detailed pollution map of the target region [33]. From gases concentrations measured by drones real-time AQI maps in both 2D and 3D areas can be produced that describe the AQ conditions of an urban environment.

Sensors can also be placed on vehicles such as taxis or public transport vehicles. Sensors are anchored on top of vehicles, to create a mobile sensor network to increase the number of urban sites monitored. Graphs and heat maps to show an overview of the gasses levels on the taxis/bus routes can be created [22].

Satellite remote sensing of air quality has evolved drastically over the last decade. The satellite retrieved trace gases are useful in analyzing and forecasting events that affect air quality, and they can be used to the inference of surface air quality. Aerosol optical depth (AOD) derived from satellite remote sensing

[10] http://vocab.linkeddata.es/datosabiertos/def/transporte/trafico
[11] open511 specification: `http://www.open511.org/`
[12] Road Accident Ontology: `https://www.w3.org/2012/06/rao.html`

is widely used to estimate surface PM 2.5 concentrations. The satellite provides the total concentration of gases on the column between the surface and the top of the troposphere [23]. To provide an idea of their Spatio-temporal coverage, the ESA Sentinel-5P satellite has an orbital cycle of 16 days. Therefore, it can produce one value every 16 days for each location, around two values per month. Spatial resolution is quite low, above 5km, to be used at an urban scale.

The traditional methods used by public administrations for traffic monitoring are fixed measurement devices (such as inductive loop detectors, radars, video cameras, etc.) that collect data like vehicle presence, vehicle speed, vehicle length and class, lane occupancy. However, these devices can only collect data on the specific section of the road where they are installed. The last few years have seen a dramatic increase in the presence of mobile or aerial devices. These devices have the capability to detect detailed and accurate data over space and time and to cover a dynamic area. Smartphones are deemed valid for traffic sensing purposes since as long as there is a sufficient penetration rate, they will provide accurate measurements of the traffic flow (the number of users with traffic sensors should be at least 2–3% over the total cars which entered the target road). Smartphones can provide location, altitude, and speed [24]. If the reported data from several users in each road segment at a specific time interval are combined, a reasonable estimation of the traffic conditions can be obtained. Drones are used to monitor real-time traffic and also to detect speeding violations or congestion events [15]. They can identify the number of speeding violations, the average duration of the detected speeding violations, the number of congestion events: congestion events and the average period of the detected congestion events.Traffic models provide a representation of the road network in terms of the capacity it gives and the volume of traffic using it.

Traffic models are used to estimate the real traffic conditions in a city and to forecast the impact of policies that modify the viability or also forecast traffic predictions. A traffic model considers data coming from sensors and the traffic demand that can be represented through an Origin-Destination matrix for a different period (e.g., morning and evening peak hours). The output of a traffic model depends on the kind of model [14,19]. For macroscopic models, we have the density (or concentration), the flow (number of vehicles in an interval of time), and the speed. These parameters are expressed using average values.

3 Requirements for a Level 3 Air Quality and Traffic Data Modeling Framework

From the analysis of several specific applications, reported in section 2.4, the following requirements for modeling air quality and traffic emerged.

A generic air quality data model has to provide support for:

A1 in-situ fixed devices (e.g. air quality station). An in-situ sensor collects data at a distance comparable or smaller than any linear dimension of the sensor.

A2 in-situ removable devices (e.g. low-cost air quality sensor). These devices can be moved in different locations, but they can measure at this point once they are in a static position.
A3 remote sensing infrastructures (e.g. those on-board of satellites). Remote sensing is the process of detecting and monitoring characteristics of an area by measuring them at a distance (typically from satellite or aircraft).
A4 ground mobile devices (e.g., AQ devices on buses, always at the same elevation). These devices provide in-situ data along a route.
A5 airborne mobile devices. They provide in-situ data through a 3D trajectory.
A6 in-situ sensors installed at various heights (e.g., application of sensors installed externally on a building on different floors).
A7 static models. Static models do not vary over time; they may be viewed as a "snapshot" of an ecosystem at a particular moment. An example of static air quality models are the interpolated real-time air quality maps that are created by interpolating sensor measurements.
A8 dynamic models. Dynamic models provide means of simulating the time-dependent behavior of systems. Atmospheric dispersion models are dynamic models that use mathematical algorithms to simulate how pollutants disperse in the atmosphere and, in some cases, how they react.

The requirements for traffic data modelling aim at supporting:

T1 in-situ traffic observation. Similarly to **A1** and **A2**, these are traffic sensors located in a specific position.
T2 remote traffic observation at specific locations. Similarly to **A3**, these sensors may observe traffic at various locations at each time instant.
T3 static traffic models. Similarly to **A7**, a static model gives a snapshot of reality.
T4 dynamic traffic models. Similarly to **A8**, dynamic traffic models estimate the evolution concerning the time of traffic variables (flow intensity, occupancy, etc.) during a period.

4 Air quality Model

The air quality data representation capabilities of the TAQE model are described below. The model specializes in the level 2 OGC O&M data model (grey color classes in subsequent figures) with feature (entity) and process types. The data types used to represent the geospatial characteristics of the involved entities are based on those proposed by OGC standards, including feature geometric data types[13] and temporal, spatial and spatio-temporal coverages[14]. The part of the model that represents the observed entities (features of interest in O&M notation) is depicted in Figure 1.

[13] OGC Simple Feature Access: `https://www.opengeospatial.org/standards/sfa`
[14] OGC Coverage Implementation Schema: `http://docs.opengeospatial.org/is/09-146r6/09-146r6.html`

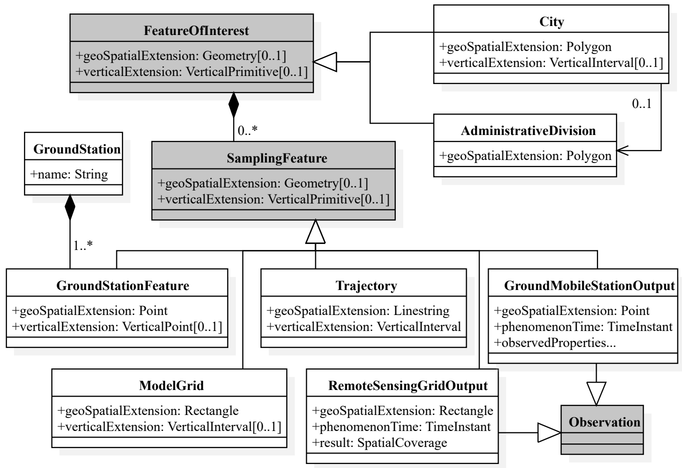

Fig. 1. Air quality feature of interest types in TAQE (Grey background used to show types of the Level 2 OGC O&M data model).

Two subtypes of O&M *FeatureOfInterest* are considered, namely, *City* and *AdministrativeDivision*, to include respectively urban scale applications (smart city scope) and also applications at regional, national and international scale. Estimated properties are often generated at specific samples of the feature of interest (*SamplingFeature*). Various types of sampling features of interest in air quality observation and modeling are shown in Figure 1, including locations of static and mobile ground stations, trajectories of flying platforms like drones and raster grids used by models and remote sensing platforms. Some sampling features are predefined, whereas others are generated by the observation or modeling processes; at the same time, they generate the observed property estimations. This is the case of ground mobile station features and remote sensing grids, which record both sampling features and property estimations, and therefore they inherit from both O&M *SamplingFeature* and O&M *Observation*.The remainder classes of the air quality model, i.e., those related to the representation of observation and modeling processes (O&M Process) and relevant outputs (O&M observation) are shown in Figure 2. Data generation processes are classified according to whether they observe or model properties. Based on the characteristics of their outputs, air quality models are further subdivided into static (*AQStaticModel*) and dynamic (*AQDynamicModel*). The former generates spatial coverages that estimate the observed properties at a specific time element (often real-time). In contrast, the latter provides spatio-temporal coverages that determine the evolution of those properties over a period of time. Air quality observation processes are also subdivided into in-situ (they observe in the surroundings of the process location) and remote sensing (they remotely observe all the places of an output spatial coverage at each time instant). In-situ processes are further classified into mobile ground stations (air quality stations installed in

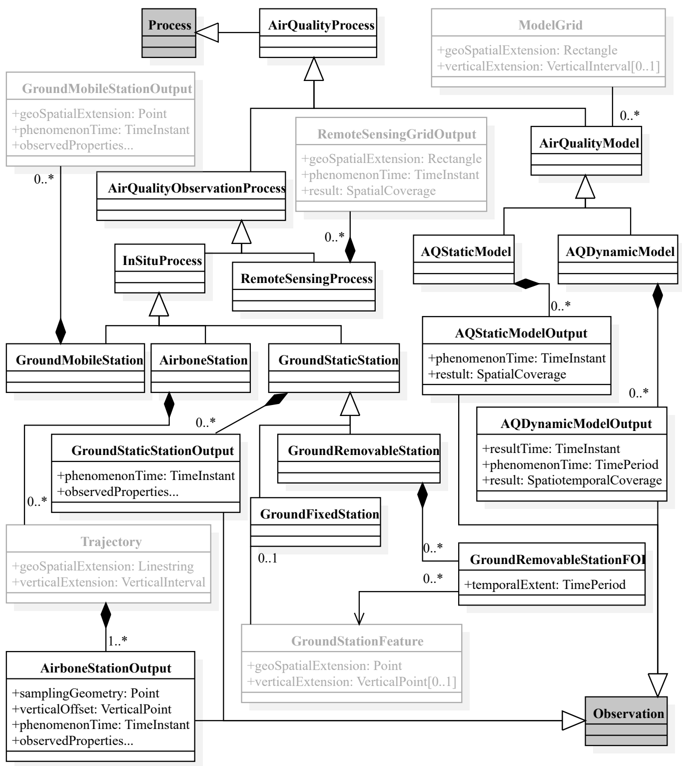

Fig. 2. Air quality process types in TAQE (Grey background used to show types of the Level 2 OGC O&M data model. TAQE feature of interest types, already depicted in Fig. 1, are shown here in light-grey).

land vehicles), airborne stations (mounted on board of flying platforms such as drones) and ground static platforms, which are either air quality stations with a fixed location or air quality stations that may be installed at different locations during their lifetime.

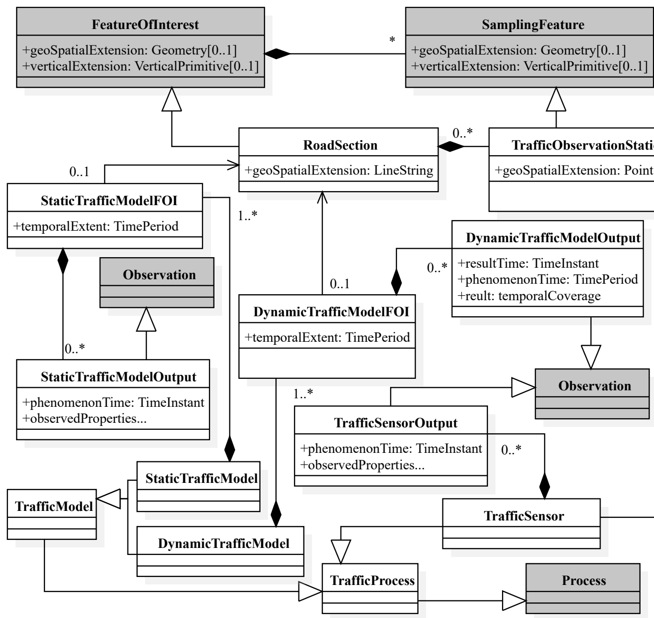

Fig. 3. Traffic data structures in TAQE (Grey background used to show types of the Level 2 OGC O&M data model).

5 Traffic Model

The data structures that enable TAQE to represent the results of traffic observation and modeling processes are graphically depicted in Figure 3. Traffic data properties such as traffic speed, traffic flow intensity, and traffic occupancy are associated with sections of the road network in TAQE. Traffic models are used to provide estimations (real-time estimations, future predictions, etc.) of the properties for a whole road section, whereas observation processes (*TrafficSensor*) are used to perform their measures at specific sampling locations (*TrafficObservationStation*). As in the case of air quality, traffic models may be either static or dynamic (depending on whether they generate estimations for a specific temporal element or they generate an evolution concerning the time of the relevant properties).

6 Models Evaluation

The TAQE data modeling framework has been evaluated, both by testing its use in a specific application and by comparing its capabilities with already existing solutions of a similar level of abstraction.

6.1 Use Case

The “TRAFAIR - Understanding traffic flows to improve air quality” EU project (2017-EU-IA-0167)[29] , funded by the Innovation and Networks Executive Agency under its Connecting Europe Facility program , aims at building an urban scale air quality observation and modeling infrastructure at six cities of Italy and Spain. The main components of the infrastructure are: i) an air quality observation infrastructure based on low-cost sensors, ii) an air quality forecast service based on a lagrangian pollutant dispersion model, iii) a public open data infrastructure based on both OGC and W3C standards, connected to the European Data Portal through the involved public administrations and iv) specific web and mobile applications for public administrations and citizens.

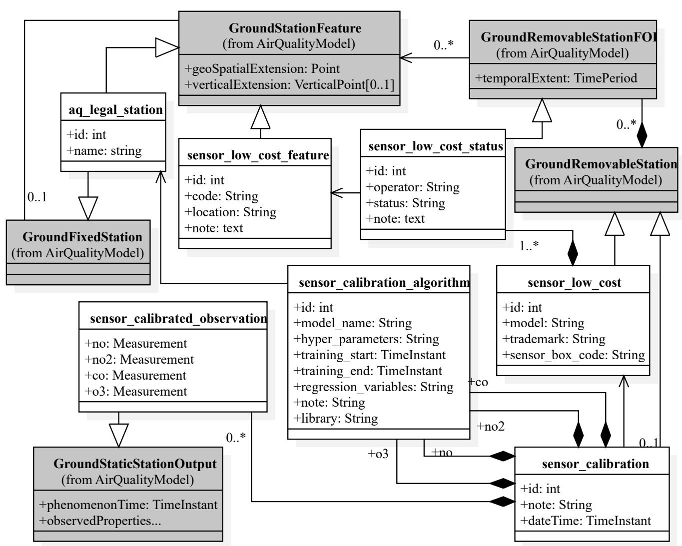

Fig. 4. Illustration of the specialization of TAQE in the TRAFAIR project (TAQE types shown in grey background).

To achieve the TRAFAIR objectives, both air quality and traffic have to be observed and modeled, and the relevant data has to be stored and published. Traffic sensors provide traffic flow intensity, average traffic speed, and traffic occupancy at specific locations of the road networks[3,4,12]. Erroneous measurements are detected and filtered out from the big data streams generated with appropriate techniques in near real-time [6,7]. Then, the observed traffic flow intensity is used as input by a static traffic flow model to estimate, in near real-time, the traffic flow intensity at each section of a subset of the city road network[2,5,28].

Regarding air quality, a network of removable low-cost sensors is used to perform measures at specific locations in each city[1,34,35]. Due to the characteristics of these low-cost devices, the quality of their measures disables their direct application. Intelligent regression models (sensor calibrations) are generated for each sensor, using the available ISO certified air quality stations as the ground truth. Those calibrations are next used to transform the voltages generated by the low-cost sensors to specific gas concentrations, enabling the use of these low-cost sensors at points where ISO stations are not located. In addition to the air quality observation, the traffic intensity at each road section, the meteorological conditions, and the city geometry (buildings) are used as input by a dynamic air quality model to generate 48-hour air quality forecasts every day[8]. The use of TAQE to model all the above air quality and traffic data has been successfully tested. A short example of this testing is illustrated in Figure 4, which shows how TAQE classes are specialized to model the air quality data generated by sensor calibrations. In particular, it is shown how TAQE *GroundStationFeature* is used to model both ISO certified air quality station locations (*aq_legal_station*) and removable low cost sensor locations (*sensor_low_cost_feature*). TAQE *GroundRemovableStation* is specialized to model both air quality sensors (*sensor_low_cost*) and sensor calibration models (*sensor_calibration*). Similarly, TAQE *GroundFixedStation* is specialized to model the set of devices installed in ISO certified air quality stations (*aq_legal_station*). Evolution concerning the time of the removable sensor locations and their status is modeled with a specialization of TAQE *GroundRemovableStationFOI*. Finally, the observations generated by sensor calibration models are represented by *sensor_calibrated_observation*, which inherits from TAQE *GroundStaticStationOutput*. All details of mappings between all the other TRAFAIR data structures and their TAQE abstract concepts are not given here due to space limitations.

6.2 Qualitative evaluation

TAQE and previously proposed models for air quality and traffic representation defined in section 2.3 have been evaluated w.r.t. the list of requirements defined in Section 3. Table 1 provide an overview of each model in level 3 related to air quality wrt requirements. While in Table 2 traffic models have been compared.

AIR_POLLUTION_Onto supports in-situ and remote observations, but not models. QBOAirbase extends this view by also adding provenance; this might allow users to store static model outputs. Ontology of air quality models is devoted to representing only air quality models, while the hackAIR ontology integrates sensor measurements and air quality, models.

The model considered in the Vocabulary to Represent Data about Traffic supports in-situ and remote observations, but not models. Nevertheless, a snapshot of traffic in a city can be obtained if it is used in conjunction with vocabularies to

Model	A1	A2	A3	A4	A5	A6	A7	A8
AIR_POLLUTION_Onto [27]	x	x	x					
QBOAirbase [17]	x	x	x				x	
Ontology of air quality models [25]							x	x
hackAIR ontology [32]	x	x	x			x	x	
TAQE model	x	x	x	x	x	x	x	x

Table 1. Evaluation of the AQ models with respect to the requirements

Model	T1	T2	T3	T4
Vocabulary to Represent Data About Traffic	x	x	x	
Ontology Layer for Intelligent Transportation [16]	x		x	
open511 specification				
Road Accident Ontology				
TAQE model	p		x	x

Table 2. Evaluation of the traffic models with respect to the requirements (p = partially, x= completely).

represent city road maps. In [16], sensors can be located into vehicles or as part of an infrastructure element of the roads, but remote sensing is not considered. Besides, a road agent can provide information on the road in only three particular cases: short-term, long-term, or anticipatory; but does not support dynamic traffic models. open511 and Road Accident Ontology provide information about road events (traffic or road accidents in Road Accident Ontology) that can be considered observable properties or features of interest of road sections but do not support either traffic models or data models for sensing.

TAQE fulfills all the requirements. Regarding traffic data, it supports in-situ data only partially since only fixed stations are supported. Remote sensing devices and other observation mechanisms that support the observation of various locations at the same time are not supported yet.

7 Conclusions

This paper presented a data modeling framework, called TAQE, for the representation of observation and modeled data in the air quality and traffic application domains. The model is integrated into level 3 of the information architecture of four levels of abstraction, which is based on the extensive use of reputed international standards. Both TAQE and all the other models found in the literature for air quality and traffic data were evaluated concerning a collection of requirements extracted from a comprehensive set of applications. Besides, the application of TAQE to a real use case in the scope of the TRAFAIR EU project was also undertaken, showing the utility of the model and its potential to both reduce development costs and to ease semantic data integration between different areas in the same application domain, but also between different application domains, enabling this way the implementation of tools of a more general purpose. Future work is related to the extension of the traffic model to support removable in-situ devices and remote sensing.

Acknowledgement

This research was supported by the TRAFAIR project (2017-EU-IA-0167), co-financed by the Connecting Europe Facility of the European Union, and by the NEAT-Ambience project (Next-gEnerATion dAta Management to foster suitable Behaviors and the resilience of cItizens against modErN ChallEnges PID2020-113037RB-I00 / AEI / 10.13039/501100011033). We thank reviewers who provided insight to improve the final version of this paper.

References

1. C. Bachechi, F. Desimoni, L. Po, and D. M. Casas. Visual analytics for spatio-temporal air quality data. In E. B. et. al., editor, *24th International Conference on Information Visualisation, IV 2020, Melbourne, Australia, September 7-11, 2020*, pages 460–466. IEEE, 2020.
2. C. Bachechi and L. Po. Implementing an urban dynamic traffic model. In P. M. Barnaghi, G. Gottlob, Y. Manolopoulos, T. Tzouramanis, and A. Vakali, editors, *2019 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2019, Thessaloniki, Greece, October 14-17, 2019*, pages 312–316. ACM, 2019.
3. C. Bachechi and L. Po. Traffic analysis in a smart city. In B. P., G. G., K. D., M. Y., P. R., T. T., and V. A., editors, *Proceedings - 2019 IEEE/WIC/ACM International Conference on Web Intelligence Workshops, WI 2019 Companion*, page 275 – 282. Association for Computing Machinery, Inc, 2019.
4. C. Bachechi, L. Po, and F. Rollo. Big data analytics and visualization in traffic monitoring. *Big Data Research*, 27, 2022.
5. C. Bachechi, F. Rollo, F. Desimoni, and L. Po. Using real sensors data to calibrate a traffic model for the city of modena. *Advances in Intelligent Systems and Computing*, 1131 AISC:468 – 473, 2020.
6. C. Bachechi, F. Rollo, and L. Po. Real-time data cleaning in traffic sensor networks. In *17th IEEE/ACS International Conference on Computer Systems and Applications, AICCSA 2020, Antalya, Turkey, November 2-5, 2020*, pages 1–8. IEEE, 2020.
7. C. Bachechi, F. Rollo, and L. Po. Detection and classification of sensor anomalies for simulating urban traffic scenarios. *Cluster Computing*, 2021.
8. A. Bigi, G. Veratti, S. Fabbi, L. Po, and G. Ghermandi. Forecast of the impact by local emissions at an urban micro scale by the combination of lagrangian modelling and low cost sensing technology: The trafair project. In *19th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Harmo 2019*, 2019.
9. C. Bizer, T. Heath, and T. Berners-Lee. Linked data - the story so far. *Int. J. Semantic Web Inf. Syst.*, 5(3):1–22, 2009.

10. M. Compton, P. Barnaghi, L. Bermudez, R. García Castro, O. Corcho, S. Cox, J. Graybeal, M. Hauswirth, C. Henson, A. Herzog, V. Huang, K. Janowicz, D. Kelsey, D. Phuoc, L. Lefort, M. Leggieri, H. Neuhaus, A. Nikolov, K. Page, and K. Taylor. The ssn ontology of the w3c semantic sensor network incubator group. *Web Semantics: Science, Services and Agents on the World Wide Web*, 17, 12 2012.
11. S. Cox. *ISO 19156:2011 - Geographic information – Observations and measurements.* International Organization for Standardization, 01 2011.
12. F. Desimoni, S. Ilarri, L. Po, F. Rollo, and R. Trillo-Lado. Semantic traffic sensor data: The trafair experience. *Applied Sciences (Switzerland)*, 10(17), 2020.
13. S. Djahel, R. Doolan, G.-M. Muntean, and J. Murphy. A communications-oriented perspective on Traffic Management Systems for smart cities: Challenges & innovative approaches. *IEEE Communications Surveys & Tutorials*, 17(1):125–151, 2015.
14. A. Dombalyan, V. Kocherga, E. Semchugova, and N. Negrov. Traffic forecasting model for a road section. *Transportation Research Procedia*, 20:159 – 165, 2017. 12th International Conference on Organization and Traffic Safety Management in large cities, SPbOTSIC-2016, 28-30 September 2016, St. Petersburg, Russia.
15. M. Elloumi, R. Dhaou, B. Escrig, H. Idoudi, and L. A. Saidane. Monitoring road traffic with a uav-based system. In *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6, 2018.
16. S. Fernandez, R. Hadfi, T. Ito, I. Marsa-Maestre, and J. Velasco. Ontology-based architecture for intelligent transportation systems using a traffic sensor network. *Sensors*, 16(8):1287, August 2016.
17. L. Galárraga, K. A. M. Mathiassen, and K. Hose. Qboairbase: The european air quality database as an rdf cube. In *International Semantic Web Conference*, 2017.
18. A. Gomez-Perez, M. Fernández-López, and O. Corcho. *Ontological Engineering: With Examples from the Areas of Knowledge Management, E-Commerce and the Semantic Web.* Springer, 01 2004.
19. Y. Hou, P. Edara, and C. Sun. Traffic flow forecasting for urban work zones. *IEEE Transactions on Intelligent Transportation Systems*, 16(4):1761–1770, 2015.
20. S. Ilarri, O. Wolfson, and T. Delot. Collaborative sensing for urban transportation. *IEEE Data Engineering Bulletin*, 37(4):3–14, December 2014. Special Issue on Urban Informatics.
21. K. Janowicz, A. Haller, S. J. Cox, D. L. Phuoc], and M. Lefrançois. Sosa: A lightweight ontology for sensors, observations, samples, and actuators. *Journal of Web Semantics*, 56:1 – 10, 2019.
22. S. Kaivonen and E. C.-H. Ngai. Real-time air pollution monitoring with sensors on city bus. *Digital Communications and Networks*, 6(1):23 – 30, 2020.
23. R. V. Martin. Satellite remote sensing of surface air quality. *Atmospheric Environment*, 42(34):7823 – 7843, 2008.
24. J. Martín, E. J. Khatib, P. Lázaro, and R. Barco. Traffic monitoring via mobile device location. *Sensors*, 19(20):4505, Oct 2019.
25. C. Métral, G. Falquet, and K. D. Karatzas. Ontologies for the integration of air quality models and 3d city models. *ArXiv*, abs/1201.6511, 2012.
26. P. Nesi, L. Po, J. R. R. Viqueira, and R. Trillo-Lado. An integrated smart city platform. In J. Szymański and Y. Velegrakis, editors, *Semantic Keyword-Based Search on Structured Data Sources*, pages 171–176, Cham, 2018. Springer International Publishing.

27. M. M. Oprea. Air_pollution_onto: an ontology for air pollution analysis and control. In Iliadis, Maglogiann, Tsoumakasis, Vlahavas, and Bramer, editors, *Artificial Intelligence Applications and Innovations III*, pages 135–143, Boston, MA, 2009. Springer US.
28. L. Po, F. Rollo, C. Bachechi, and A. Corni. From sensors data to urban traffic flow analysis. In *2019 IEEE International Smart Cities Conference, ISC2 2019, Casablanca, Morocco, October 14-17, 2019*, pages 478–485. IEEE, 2019.
29. L. Po, F. Rollo, J. R. R. Viqueira, R. T. Lado, A. Bigi, J. C. López, M. Paolucci, and P. Nesi. TRAFAIR: understanding traffic flow to improve air quality. In *2019 IEEE International Smart Cities Conference, ISC2*, pages 36–43, 2019.
30. M. A. Regueiro, J. R. Viqueira, C. Stasch, and J. A. Taboada. Semantic mediation of observation datasets through sensor observation services. *Future Generation Computer Systems*, 67:47–56, 2017.
31. M. A. Regueiro, J. R. Viqueira, J. A. Taboada, and J. M. Cotos. Virtual integration of sensor observation data. *Computers & Geosciences*, 81:12–19, 2015.
32. M. Riga, A. Moumtzidou, S. Vrochidis, I. Kompatsiaris, and P. Syropoulou. D4.2: Semantic integration and reasoning of environmental data. Technical report, CERTH, 2017.
33. G. Rohi, O. Ejofodomi, and G. Ofualagba. Autonomous monitoring, analysis, and countering of air pollution using environmental drones. *Heliyon*, 2020.
34. F. Rollo and L. Po. Senseboard: Sensor monitoring for air quality experts. In C. Costa and E. Pitoura, editors, *Proceedings of the Workshops of the EDBT/ICDT 2021 Joint Conference, Nicosia, Cyprus, March 23, 2021*, volume 2841 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2021.
35. F. Rollo, B. Sudharsan, L. Po, and J. G. Breslin. Air quality sensor network data acquisition, cleaning, visualization, and analytics: A real-world iot use case. In A. Doryab, Q. Lv, and M. Beigl, editors, *UbiComp/ISWC '21: 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and 2021 ACM International Symposium on Wearable Computers, Virtual Event, September 21-25, 2021*, pages 67–68. ACM, 2021.
36. A. Sharif, J. Li, M. Khalil, R. Kumar, M. I. Sharif, and A. Sharif. Internet of Things — smart traffic management system for smart cities using Big Data analytics. In *14th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP 2017)*. IEEE, December 2017.