This is the final peer-reviewed accepted manuscript of:

A. Seravalli *et al*., "Towards Smart Cities for Tourism: the POLIS-EYE Project," *2022 IEEE International Smart Cities Conference (ISC2)*, Pafos, Cyprus, 2022, pp. 1-7.

The final published version is available online at: https://dx.doi.org/10.1109/ISC255366.2022.9922095

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/)

When citing, please refer to the published version.

Towards Smart Cities for Tourism: the POLIS-EYE Project

Alessandro Seravalli, Mariaelena Busani, Simone Venturi*
*Geosmartlab, Via Emilia 69, Imola, 40026, BO Italy
a.seravalli@sis-ter.it, m.busani@sis-ter.it, s.venturi@sis-ter.it
Arianna Brutti, Carlo Petrovich, Angelo Frascella, Fabrizio Paolucci†
†ENEA, Via M.M. Sole 4, Bologna, 40129, Italy
arianna.brutti@enea.it, carlo.petrovich@enea.it, angelo.frascella@enea.it, fabrizio.paolucci@enea.it
Marco Di Felice, Michele Lombardi‡
‡DISI, University of Bologna, Viale del Risorgimento 2, Bologna, 40136, Italy
marco.difelice7@unibo.it, michele.lombardi2@unibo.it
Elena Bellodi§, Riccardo Zese¶, Francesco Bertasi§
§Dipartimento di Ingegneria, Università di Ferrara, Italy
¶Dipartimento di Scienze chimiche, farmaceutiche ed agrarie, Università di Ferrara, Italy
elena.bellodi@unife.it, riccardo.zese@unife.it,francesco.bertasi@unife.it
Elia Balugani, Alket Cecaj, Rita Gamberini, Marco Mamei, Marco Picone‖
‖Dipartimento di Scienze e Metodi dell'Ingegneria
Artificial Intelligence Research and Innovation Center
Università di Modena e Reggio Emilia, Italy
{elia.balugani, alket.cecaj, rita.gamberini, marco.mamei, marco.picone}@unimore.it

***Abstract*—Novel and widespread ICT and Internet of Things (IoT) technology can provide fine-grained real-time information to the tourist sector, both to support the demand side (tourists) and the supply side (managers and organizers). We present the POLIS-EYE project that aims to build decision-support systems helping tourist-managers to organize and optimize policies and resources. In particular, we focus on a service to monitor and forecast people presence in tourist areas by combining heterogeneous datasets with a special focus on data collected from the mobile phone network.**

***Index Terms*—tourist presence and forecasting, mobile phone data analysis, data visualisation**

I. Introduction

The tourism sector is rapidly changing. Multiple global trends are conditioning the way people conceive and experience tourist activities. Digitalisation processes, the global COVID-19 pandemic and the climate crisis are probably the most impactful trends that will shape the tourism sector in future years.

The digitalisation process is deeply changing tourism experiences, also due to a general switch in peoples' expectations towards the way in which services are supplied to the public. The digital fruition of services is now a precondition for a positive user experience, and this is true for the tourism sector too. Its competitiveness is influenced by digital technologies both on the relation with the tourist and the services management.

The global COVID-19 pandemic has stopped most of the tourism activities as we used to know them, but it has also pushed people to discover and find different opportunities, as well as it has accelerated the digital conversion process. The behavioural change of the tourists forces the understanding of new tourists' needs, the space-time dynamics of the travels and the relation with points of interests (POIs), as well as the interaction between involved stakeholders.

At the same time, climate change forces the attention of the public administrations, as well as the operating private actors, on the necessity for an optimised use of resources. This leads to the need of getting to know and of gaining information on current resources requirements in relation to tourism experiences.

The described picture motivates the necessity for stakeholders involved in the tourism planning and operation phases to evolve the approach of services provisioning. The required transition of business and service models has to be supported by reliable information and smart tools.

The digital transition offers here an opportunity in this sense. As the number of data sources describing people and therefore tourist activities and mobility is constantly increasing, new possibilities arise to support data-driven decisions and planning. Data of different nature, produced for different aims, represent the real opportunity to tackle the described situation.

However, even if data availability is growing, the current fragmentation of data formats, spatial and temporal granularity and coverage makes it very difficult to integrate them to produce the relevant information operators are eager for.

Accordingly, one the main required next steps is toward standardisation and integration of the different data features.

In addition, another fundamental leap forward will come from the match between novel data sources and artificial intelligence and machine learning techniques. This will provide high-level, semantically rich information to understand the demand behaviour and needs, and offer personalized solutions that can create a fine-grained match between tourist demands and tourist destination offerings.

Finally, the obtained results must be conveyed to stakeholders in an effective, flexible and understandable way. Systems for the visualisation of data and AI results represent the practical tool to assist operators with insights and decision support.

The POLIS-EYE (POLIcy Support systEm for smart citY data governancE) project aims to tackle the pictured scenario, developing a solution based on data processing and analysis, that makes tourism related patterns interpretable and actionable. POLIS-EYE has been created with the aim of developing a digital platform able to provide useful information for the regional tourism sector and to become part of the tourism-sector management system in Emilia-Romagna (a major tourist region in Italy). The tool to support public decision makers and key active actors aims to understand the past, current and future dynamics of flows and densities (of people and vehicles), useful for the planning of events and ancillary services (such as public transport).

The objective and contribution of this paper is to present the approach and technologies adopted in the POLIS-EYE project, in particular, we describe:

1) the challenges associated to all the aspects involved in data acquisition and enhancement processes (open and non-open), such as: data management, heterogeneity of data (format, spatial and temporal granularity, etc.), GDPR compliance
2) the integration of artificial intelligence and machine learning models developed with multiple techniques
3) visualisation of information and results of AI models, for an interpretation accessible to multiple stakeholders to support decision making.

II. Related Works

A number of recent research works, applies data mining and artificial intelligence techniques to large-scale tourist data [1] and smart cities [2]. The main reason for this novel/renewed interest is due to the huge amount of data that is available and can be exploited in the tourist domain. Moreover, it is expected in the next future that the amount of data will dramatically grow thanks to Internet of Things systems even more capable of monitoring the world in real time. This context provides the ideal stage for the Business Analytics techniques applied to tourist domain, that can be classified in three categories:

Descriptive Analytics: statistical and AI methods that are used to search and summarize historical data in order to identify some patterns or meaning. For example, in [3] descriptive analytics is used to study tourist arrivals to Catalonia for the out-of-sample period (January 2009 to July 2012).

Predictive Analytics: AI methods (Machine Learning) to build models that are able to classify information, make previsions and identify data structures. For example, in the tourist domain, the system [4] applies predictive analytics to study and predict the Taiwan's inbound tourism arrivals.

Prescriptive Analytics: AI methods (Machine Learning and optimization) to provide recommendations, by automatically solving decision models. In this context, the project [5] proposes, referring to the tourist domain, a context-aware based restaurant recommender system. Similarly the LUME-PLANNER project [6] integrates multiple AI-methods to recommend destination and routes to tourists in a personalized way. More in general, in terms of data collection in a Smart City context, abstract high-level frameworks were presented:

SCIAM (Smart City Infrastructure Architecture Model) [7]: it allows the identification of data exchange interfaces, standard classification and different architectures mapping on the same reference model

SMArc (Smart, Semantic Middleware Architecture Focused on Smart City Energy Management) [8] is a middleware proposal for Smart Grids. Also in this case solutions to integrate multiple data streams (related to the energy sector) are proposed.

U-City (Ubiquitous Eco-City Planning, in Korea) [9]: a project aiming to create a ubiquitous city model. The goal of this project is to define a general infrastructure supporting multiple services in the smart city.

Finally, ITU-T [10] and IES-City [11] are exemplary initiatives to create a platform for IoT and smart sustainable cities stakeholders to exchange knowledge and identify policy and standard needs, there included those services related to the tourist sector.

While all the above proposals tackle challenges similar to the ones of the POLIS-EYE project (e.g., data integration and AI-services applied to the smart city / smart tourist scenario) most of the them do not offer usable actually deployed implementations. POLIS-EYE project combines a proposed data collection/exchange platform with an open interoperability oriented implementation, AI-based Business Analytics services and a user advanced interface, regards to regional tourism.

III. The approach of the POLIS-EYE project

Data play a crucial role in the development of platforms and applications for Smart Cities: they are the fuel to feed the analysis and the services, allowing to create new information with added value. Data can as a matter of fact be utilized for new goals which are different from those originally conceived for their collection (e.g. traffic monitoring can be reused for tourism prediction). Collections of data have grown enormously with the acceleration of the digitalization of almost all the services and are nowadays present everywhere: traffic sensors, public transport accesses, position of personal smartphones, diffusion of information (e.g. public events) by websites, weather indicators, purchases, etc. Therefore enormous amounts of data already exist, in digital format, covering almost every sector of our lives. Three critical issues are however present:

1) The *availability* of the data, which are often owned by private or semi-private companies and available only at high costs. Even public companies do not often make available the data. A legislative effort has begun with the Open data approach, also at EU level [12], but still needs to be fully implemented;
2) The *interoperability* of the data: even when the data are available, they often cannot be used in an integrated and harmonized way, having particular and different formats. Source data are created for specific applications and often cannot be reutilized for different applications. Shared and common transmission protocol, formats and standards are therefore necessary to avoid isolated silos.
3) The *privacy compliance* of the data. As the most relevant data to support tourist management deals with information about where and when people go and how they spend their time, strong privacy concerns apply to this kind of data. Therefore, it is important to device GDPR-compliant procedures and algorithms to extract value form the data without undermining the privacy of the individuals.

The objective of the POLIS-EYE project was to avoid the creation of further silos for the Smart City and then to design the development of a platform according to an approach based on interoperability.

A. Interoperability of the data and standardization

Interoperability can be defined as the “capability of two or more networks, systems, devices, applications, or components to exchange and readily use information, securely, effectively, and with little or no inconvenience to the user” [13]; this definition highlights how complex interoperability is and how big the impact is on many aspects (data format, semantic of information, communication protocols, trustworthiness, timing, etc.).

In this context, the POLIS-EYE project adopts the approach proposed by the Smart City Platform project (*https://smartcityplatform.enea.it/#/en/index.html*) [14]. This approach is based on the UrbanDataset interoperability data-exchange format, based on widely used standards (e.g. XML and JSON), that defined common semantic, syntax and structure of the information [15]. It is part of a wider set of national interoperability specification, called Smart City Platform Specification (SCPS), public, freely accessible online and applicable to any Smart City context. In the context of the POLIS-EYE project the UrbanDataset interoperability data-exchange format is based on a set of abstract data models constituting a reference representation of the typologies of data fluxes (e.g. public events, transits and presences of people or transport means), syntax-independent and defining a shared semantic.

In particular, the analysis of the several sources and data types flows considered in POLIS-EYE has highlighted that even data belonging to the same kind of flow can be very heterogeneous. Therefore, in order to provide an unambiguous reference for the algorithms and the forecasting and decision-making services carried out by the project, but also for any future services and new data sources, a set of *Abstract Data Models* has been defined.

The Abstract Data Models provide an univocal, not ambiguous and harmonized representation of the different types of the considered data flows and share the same lists of coded values; they are independent from the format, syntax and implementation structure adopted to real exchange data between systems. They can be used as a basis for defining interoperability specifications for data exchange between data sources/software systems and Smart City Platforms/Applications.

In this perspective, they have been designed to be generic and customizable for different application contexts and sources; for this reason they identify the maximum set of information that may be in a type of data flow, setting only a minimal subset as mandatory (i.e. information that must always be provided).

POLIS-EYE has defined the Abstract Data Models related to four types of data flows:

(i) events: data that identify and describe an event of recreational or cultural interest (e.g. place and day/time, or period, theme, category, etc.); *(ii)* attendance: data on the presence of people or means of transport, in a certain spatial and temporal location; *(iii)* transits: data related to the transit of people or means of transport, detected along a route over a certain period of time; *(iv)* weather conditions: data on detected weather conditions (historical data or current conditions).

The complete specifications of these models are reported in [16], [17]. Thanks to these data models different kind of data can be effectively merged and integrated together.

On the one hand, the use of these common models allows to simplify the design of the POLIS-EYE platform in that all the modules and algorithms can be developed on the basis of this common representation. On the other hand, the fact that the adopted standardization is part of wider national initiatives will push forward the adoption of our choice.

B. The POLIS-EYE Architecture

POLIS-EYE is based on a Smart City Platform able to send and receive data in the UrbanDataset format in a scalable way. The POLIS-EYE architecture is shown in Figure 1 and it is created along the following lines:

- the infrastructure collects the data sent from the different sources
- data are either already expressed in the UrbanDataset format or converted accordingly in compliance with the abstract models. Then they are stored in appropriate data bases

- The platform then integrates AI-based analytics components to extract information from data
- These analyses are collected and displayed in a dashboard, showing the map of the territory and all the results, which can be queried and explored by the administrators and users.

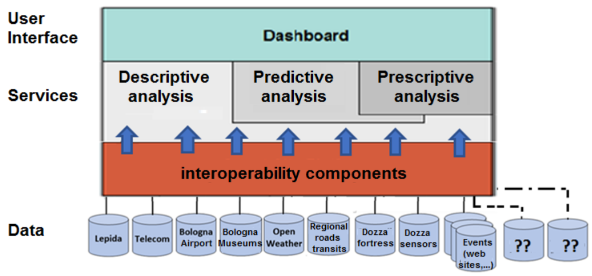

Fig. 1. The POLIS-EYE architecture. Multiple data sources exposes data in the UrbanDataset format. Resulting data feed the analytics services that then populate the dashboard supporting the decision making process.

Accordingly, POLIS-EYE can be easily replicated and applied to other scenarios, both in terms of geographical and application scopes, and relatively to the use of sources or implementation of services other than those currently in use. This approach not only brings benefits to the platform, but also represents an opportunity for those who want to use it. For example, a new entity wishing to use the services of the POLIS-EYE platform on their data, as we will see shortly, could connect to the system simply by providing their data according to the format and interfaces provided, without requiring any intervention either in the own system, nor at the POLIS-EYE platform level.

IV. Main Achievements

A. Data

The POLIS-EYE platform currently integrates with the UrbanDataset format the following data sources:

1) Telecom operators routinely collect data about the number of cell phones connected to a given base transceiver station - BTS (and more specifically to the sectors of a BTS) [18]. This is done transparently from the user perspective (users are not required to install apps or anything) and is compliant with the GDPR rules as the resulting aggregate data are not personal data. We got access to data covering Emilia-Romagna region in Italy in August and September 2019 and 2020. Data have been aggregated spatially at the level of census areas (ACEs) and temporally at 15-minutes intervals (i.e. for each ACE in the region we have the time series of the number of cell phones from the operator present in that ACE at hourly resolution). The number of cell phones can be converted to actual people count taking into account Telecom operators' market share and people demographics - age and gender of people connected to the network, nationality (Italians/foreigners), type of contract (business/regular), type of client (commuter/resident).
2) Data about WiFi access point connections to a major WiFi network spanning the whole region with more than 7700 access points. Data have been aggregated spatially at the level of single access points and temporally at 15-minutes intervals. Extra anonymization measures have been taken into account to make the extracted data GDPR complaint.
3) Data from the Bologna Museums Institution relative to the months from June 2018 to February 2020; for each month, for each visitors' nationality in that month, the total number of admissions to 10 museums is available;
4) Data from the Marconi Bologna Airport, relative to the months from January 2018 to February 2020, reporting the total of arrivals/departures from/to a specific country in a given month.

In addition, data were integrated with those collected from open sources, such as:

1) Daily weather data from the "ilmeteo" website (https://www.ilmeteo.it): minimum, maximum and mean temperature, wind speed, humidity, pressure, rainfall, precipitation type;
2) Data relative to events of different kind obtained from municipal or regional tourist websites, described by the corresponding ACE, size (estimated attendance), category (sport, etc.), topic (football, etc.), duration (in days), and day(s) of the week in which the event took place. A preliminary software, able to automatically query more than 30 selected websites, has been developed for this purpose.

Access to this kind of data required complex negotiations with the data providers, especially with regard to anonymization and GDPR rules. However, the use of the UrbanDataset format and the abstract data models described in the previous section streamlined the data acquisition process making it replicable and expandable to other sources.

These data *per-se* can be very effective to monitor and now-cast tourist presence in the area. In the next section we describe an approach to extend on this and actually forecast people presence so as to support application in need for such information in advance (e.g., resource planning).

B. Descriptive analysis results

For descriptive analytics services, POLIS-EYE applies standard machine learning algorithms (both supervised and unsupervised learning methods) to the collected data, plus an advanced data mining algorithm developed as part of the project for a specific task (trajectory reconstruction).

Regions characteristics and similarities. Clustering was applied for learning descriptive models in terms of clusters of ACE having similar density of foreigners, using data from the mobile phone network in a few months of 2019 combined with daily weather data. The same experiments were also replicated with the same kind of data collected in the same

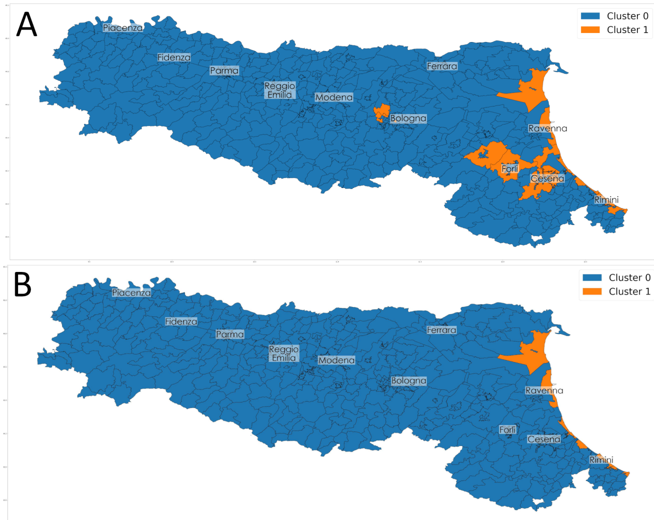

Fig. 2. Clustering of ACEs relative to August-September 2019 (A) and 2020 (B) w.r.t. the distribution of foreigners in the Emilia-Romagna region.

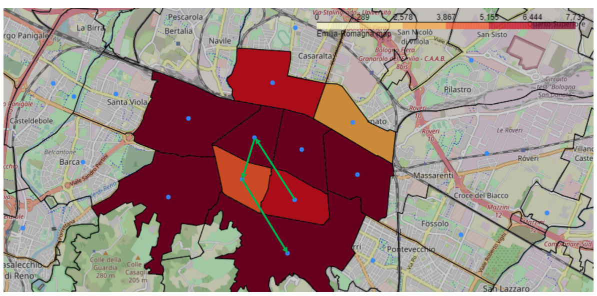

Fig. 3. An example of an aggregated trajectory to reconstruct tourist flows in the city on the basis of presence data only.

months of 2020 with the aim of comparing tourism before and during the COVID-19 pandemic. Before learning the model, data were pre-processed. Firstly, the granularity of the data was changed from a 15-minute sampling to a 1-day sampling. Secondly, weather data were added to each ACE to get a matrix where each row represents an ACE and columns correspond to 33 attributes for each of the 61 days. K-Means was applied by keeping all the parameters except `n_clusters` to their default values. In order to identify the optimal number of clusters, various values were tested in the range [2-15] by calculating the corresponding silhouette coefficients. Besides K-Means, Affinity propagation and Mean-shift clustering algorithms were also tested but with reached lower silhouette coefficients.

Results with 2019 data are shown in Figure 2A: cluster 1 (orange) corresponds to a high density of foreigners at the Marconi Bologna Airport, along the coast and in the province of Forlì-Cesena; cluster 0 (blue) corresponds to a lower density of foreigners in the inland. Results with 2020 data are shown in Figure 2B: with respect to 2019, the internal ACEs disappear. This is probably due to the absence of public events and the reduction in the number of flights. The results of this analysis can be used for a better promotion of tourism, e.g., targeting foreigners or Italians depending on the clusters.

Mobility patterns and trajectories. A custom algorithm was designed in the course of the project to extract aggregated trajectory information from dynamic count data (e.g. the mobile phone network data). The approach operates by 1) identifying candidate paths, as sequence of traverse locations in time; and 2) selecting which paths to use and how many individuals to route along each path so as to approximate the known counts. Since the output consists of aggregated trajectory, it can be considered GDPR compliant.

Formally, given a set of locations V and a number of time steps n, the method approximately solves:

$$\arg\min\{\|c(x) - \hat{c}\|_1 + R(x) \mid x \in \mathbb{R}^{|\Pi|}, x \geq 0\} \quad (1)$$

where $\hat{c}$ is the know count information (for each location and time), x are the amounts of individuals routed along each path, leading to reconstructed counts $c(x)$ for each location and node. $R(x)$ is a regularization terms that takes into account rational behavior (e.g. a preference for shorter paths) and discourages trivial solutions. Π is the set of all possible paths, i.e. the powerset 2^{V^n} of all sequences V^n. Enumeration of Π, which would be prohibitively expensive, is avoided by relying on Column Generation [19]. An example of an extracted (aggregated) trajectory is shown in Figure 3.

C. Predictive analysis results

Tourist patterns of visit. A first goal was being able to predict the amount of people of a certain nationality who might visit the museums in Bologna knowing the total number of arrivals at the airport. Each row of the database used in the experiments is described by month, flight country of origin, total number of passengers from a specific country (Italy included) arrived in Bologna in that month, and total number of visitors to each museum in Bologna in that month from that country. The total number of visitors was then reduced to 5 classes (0-50, 51-100, 101-250, 251-500, over 500 people) in order to get a labelled dataset for multi-class classification. Tree-based classifiers were applied: Random Forest (RF), Decision Trees (DT) and Extra Trees (ET). Data were divided into 75% for training set and 25% for test set; the distribution of the examples among the classes in the training set was: 205 examples for the [0-50] range, 86 for [51-100], 133 for [101-250], 63 for [251-500], 58 for $>$ 500. The overall accuracy obtained over the test set was 0.703 for RF and ET, and 0.67 for DT, while other machine learning performance metrics are reported in Table I for each class (only RF and ET are reported being the best performing models).

Given the information about the total arrivals by plane from a specific country in a month, the classifier predicts the class of people (e.g. [51-100]) who will visit the museums in Bologna.

Impact of events on tourists. A second goal was building a predictive model that, given a future event for which the number of visitors is unknown, determines the expected percentage increase of people among five classes of increase:

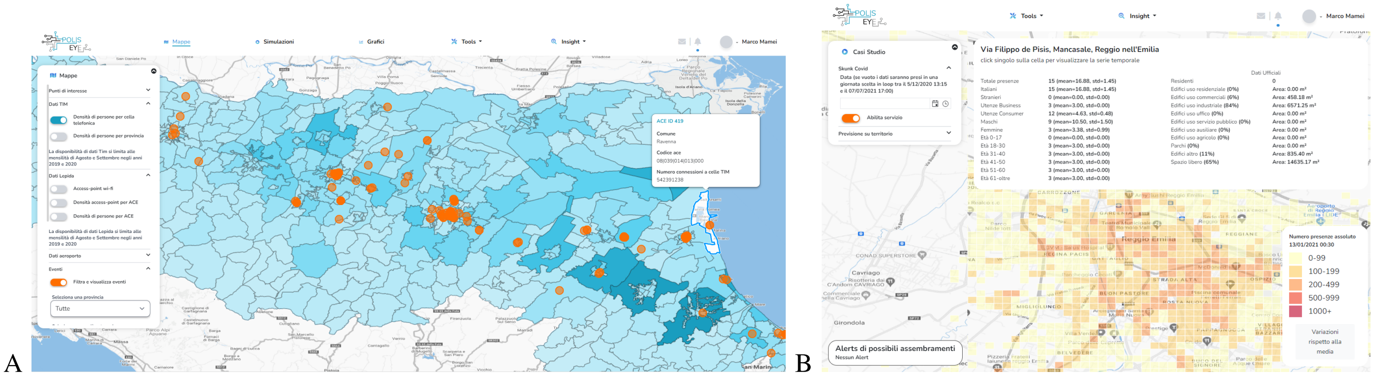

Fig. 4. The POLIS-EYE Dashboard. (A) Data showing people density as measures by cellular network at the ACE level overlaid with events and gatherings in the region. (B) Fine-grained people density in a city as measured by cellular network and WiFi data.

TABLE I
RESULTS OF THE RF AND ET CLASSIFIERS ON THE TEST SET IN TERMS OF PRECISION (P), RECALL (R), F1-SCORE (F1), ACCURACY (ACC) FOR THE PREDICTION OF VISITORS TO THE BOLOGNA MUSEUMS.

		Number of visitors					Acc.
		0-50	51-100	101-250	251-500	>501	
RF	P	0.78	0.69	0.65	0.47	0.69	**0.703**
	R	0.88	0.84	0.52	0.28	0.58	
	F1	0.83	0.76	0.58	0.35	0.63	
ET	P	0.79	0.63	0.63	0.46	0.74	**0.703**
	R	0.94	0.70	0.57	0.21	0.74	
	F1	0.86	0.67	0.60	0.29	0.74	

TABLE II
RESULTS OF THE RF AND ET CLASSIFIERS ON THE TEST SET FOR PREDICTING THE CLASS OF INCREASE OF VISITORS TO A GIVEN EVENT, IN TERMS OF PRECISION (P), RECALL (R), F1-SCORE (F1).

		% Increase of visitors				
		+[0-10]	+(10-25]	+(25-50]	+(50-75]	+>75
RF	P	1	1.00	0.69	1.00	1.00
	R	1	0.75	1.00	0.25	0.50
	F1	1	0.86	0.82	0.40	0.50
ET	P	1	1.00	0.64	0.00	0.50
	R	1	0.75	1.00	0.00	0.50
	F1	1	0.86	0.78	0.00	0.50

+[0-10]%, +(10-25]%, +(25-50]%, +(50-75]%, +>75%. Data about different events in the region (sport competitions, exhibits, fairs, etc.) were combined with the mobile phone network and weather data. In particular, for each event the 'average peak' and the 'average peak per day' were calculated from the mobile phone data. 'Average peak' is the average, over the 61 available days, of the maximum number of people present each day in the ACE where the event takes place, while the 'average peak per day' is the average, over the 61 available days, of the maximum number of people in the ACE of the event for that specific day of the week. Data were then combined into a 77×11 matrix. Each row corresponds to an event, while columns correspond to average peak, average peak per day, category, topic, duration, day of the week, size, minimum, maximum and mean temperature, and precipitation type. 44 Events registered between 10,000 and 100,000 participants and 33 events between 1,000 and 10,000 participants. The attribute to be predicted is the percentage increase of visitors calculated as $100 \cdot \frac{max\ peak}{average\ peak\ per\ day}$, where $max\ peak$ is the maximum number of people connected to the mobile phone network inside the ACE where the event takes place. The percentage increase was split into the 5 classes above to get a labelled dataset for multi-class classification purposes. Again, RF, DT and ET were tested with the best hyper-parameters tuned by means of a grid search. Overall accuracy over the test set was 0.75 for RF, 0.65 for DT, and 0.7 for ET, while other metrics are reported in Table II, demonstrating that RF has the best performance.

In case of events lasting more than 1 day, every day of the event is treated as an independent event but is discriminated by the attribute 'duration' which is greater than 1. In this context, DT, ET, and RF were also tested as regressors, but were discarded as we did not consider useful returning to the user a single real value.

D. POLIS-EYE Platform

Based on the outlined architecture, the POLIS-EYE dashboard is the final component of the system, useful for the visualisation and interaction processes of good territorial data and of the results of the algorithms. The dashboard is developed through a standard LAPP (Linux + Apache + PostgreSQL + PHP) stack-based architecture using the Laravel framework. The whole stack is wrapped into a Docker container. The solution relies on the use of two databases: the first one is the PostgreSQL shared with the whole system and the second is local, in order to manage the authentication process and authorization functions for accessing resources. This allows the users to access and use the services developed through a web interface (https://poliseye.sis-ter.it).

The dashboard has a responsive and user oriented design (see Fig. 3 for an example). In fact, all the system features have been organised into different sections, in order to maximise the quality of the user experience and to facilitate the interpretation of information. The developed solution works as a decision support system for involved stakeholders, and it offers data visualisation on the territory of interest through graphs and cartographic representations (e.g. heat maps). The elaborated views are categorised on the basis of the type of available data (e.g. wi-fi access point, flights) and on the necessary process for extracting them. Indeed, the dashboard offers two distinct sections to separate the simpler views, in which the raw data are aggregated for individual criteria (e.g. temporally, geographically for ACEs), from the more elaborated views that exploit the services analysis described in the previous paragraphs.

The interaction between the services takes place through REST calls, and the results are stored in the local database, in order to allow the user to retrieve them when needed. This feature has an additional dedicated section and is essential for a decision-making user. In fact, it allows the comparison and interpolation of different processing outputs, triggering critical analysis of information and insight extrapolation.

The platform is interoperable, easily extendable and ready to integrate other services and platforms related to the analysis of the territory.

The platform is currently used in three main case studies representative of the tourism application context: *(i)* Analysis of attendance of Bologna museums by citizens and tourists, *(ii)* analysis of influx of tourists to FICO Eataly World Bologna, *(iii)* supports the analysis of excursion tourism in the village of Dozza.

In these case studies, the sources and data types flows useful for reconstructing and forecasting flows and presences of people in a certain location, in correlation with meteorological events or of recreational or cultural interest, have been identified and analyzed.

The tool provides support to the daily activity of the employees of public bodies and companies in the sector, in the understanding of the dynamics of people and vehicles density and flow and therefore, in the development of conscious strategies and in the planning and operation of related services.

V. Conclusion

The POLIS-EYE project is an innovative and expandable platform to combine multiple data sources (e.g., data from mobile telephone operators, ticketing data, weather information, events, road traffic, access public WiFi) to analyse tourist fluxes and activities. The platform has been employed in a number of case studies and aims to support decision makers and tourist planners to analyse multiple scenarios and plan resources. In our future work, we plan to integrate even more data sources and focus on optimization and what-if analysis to further support advanced decision activities. The availability of the data, with common adopted standards, remains a crucial issue for the success of such projects.

Acknowledgement

Work supported by: the POR-FESR 2014-2020 Project: POLIcy Support systEm for smart citY data governancE - POLIS-EYE. PG/2018/631990. CUP E21F18000200007;

References

[1] A. Kontogianni and E. Alepis, “Smart tourism: State of the art and literature review for the last six years,” *Array*, vol. 6, p. 100020, 2020.

[2] A. Achilleos, C. Markides, A. Konstantinidisa, I. Giorkas, G. M. Kapitsaki, C. Mettouris, and G. A. Papadopoulos, “Adopting an open smart city platform: a survey,” in *IEEE International Smart Cities Conference (ISC2)*, Casablanca Morocco, 2019.

[3] O. Claveria, E. Monte, and S. Torra, “A new forecasting approach for the hospitality industry,” *International Journal of Contemporary Hospitality Management*, vol. 27, no. 7, pp. 1520–1538, Oct. 2015. [Online]. Available: https://doi.org/10.1108/ijchm-06-2014-0286

[4] C. W. S. Chen, M. C. Cheng, and S. Sriboonchitta, “Predictive analytics of taiwan inbound tourism from ASEAN 5,” *International Journal of Tourism Sciences*, vol. 18, no. 2, pp. 124–138, Apr. 2018. [Online]. Available: https://doi.org/10.1080/15980634.2018.1471876

[5] K. Achmad, L. Nugroho, A. Djunaedi, and W. Widyawan, “Context-aware based restaurant recommender system: A prescriptive analytics,” *Journal of Engineering Science and Technology*, vol. 14, pp. 2847–2864, 10 2019.

[6] “Lume planner,” http://www.lumeplanner.it.

[7] M. Gottschalk, M. Uslar, and C. Delfs, *The Use Case and Smart Grid Architecture Model Approach*. Springer International Publishing, 2017. [Online]. Available: https://doi.org/10.1007/978-3-319-49229-2

[8] J. Rodríguez-Molina, J.-F. Martínez, P. Castillejo, and R. de Diego, “SMArc: A proposal for a smart, semantic middleware architecture focused on smart city energy management,” *International Journal of Distributed Sensor Networks*, vol. 9, no. 12, p. 560418, Dec. 2013. [Online]. Available: https://doi.org/10.1155/2013/560418

[9] M. Jang and S.-T. Suh, “U-city: New trends of urban planning in korea based on pervasive and ubiquitous geotechnology and geoinformation,” in *Computational Science and Its Applications – ICCSA 2010*. Springer Berlin Heidelberg, 2010, pp. 262–270.

[10] “Itu-t, smart sustainable cities,” https://www.itu.int/en/ITU-T/ssc/Pages/default.aspx.

[11] “IES-City Framework,” https://pages.nist.gov/smartcitiesarchitecture/.

[12] Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1561563110433&uri=CELEX:32019L1024

[13] G. Locke and P. Gallagher, *NIST Special Publication 1108 - NIST Framework and Roadmap for Smart Grid Interoperability Standards*, release 1.0 ed. NIST, 2010.

[14] A. Brutti, P. De Sabbata, A. Frascella, N. Gessa, R. Ianniello, C. Novelli, S. Pizzuti, and G. Ponti, *Smart City Platform Specification: A Modular Approach to Achieve Interoperability in Smart Cities*. In: Cicirelli, F., Guerrieri, A., Mastroianni, C., Spezzano, G., Vinci, A. (eds) The Internet of Things for Smart Urban Ecosystems. Internet of Things. Springer, Cham., 2019. [Online]. Available: https://doi.org/10.1007/978-3-319-96550-5_2

[15] A. Brutti, A. Frascella, N. Gessa, P. De Sabbata, and C. Novelli, “Interoperability in the Smart City: A semantic approach for merging flexibility with strictness,” in *Workshop SSC 2018: Fourth IEEE International Workshop on Sensors and Smart Cities, Special session in Proceedings 2018 IEEE International Conference on Smart Computing*, 2018.

[16] POLIS-EYE - Abstract models for people flows. [Online]. Available: http://www.cross-tec.enea.it/tecnopolo/imple/pgcl.asp?p=750&lingua=en&attiva_debug=0

[17] SCPS UrbanDataset Web Library . [Online]. Available: http://smartcityplatform.bologna.enea.it/UDWebLibrary /en/urbandataset

[18] M. Rahnema, “Overview of the gsm system and protocol architecture,” *IEEE Communications*, vol. 31, no. 4, pp. 92–100, 1993.

[19] G. Desaulniers, J. Desrosiers, and M. M. Solomon, *Column generation*. Springer Science & Business Media, 2006, vol. 5.