

# Tourpedia App: A Web Application for Tourists and Accommodation Owners

Angelica Lo Duca<sup>id<sup>a</sup></sup> and Andrea Marchetti<sup>id<sup>b</sup></sup>

*Institute of Informatics and Telematics, National Research Council, via G. Moruzzi, 2, Pisa, Italy*

**Keywords:** Web Applications, Tourism, Open Data.

**Abstract:** We set out a strategy to add missing details to Tourpedia, a knowledge base containing accommodation information completely built on open data. The strategy is based on developing a Web application (called Tourpedia App), which incentivizes accommodation owners to correct and add information about their activity to Tourpedia. Valuable statistics about the accommodation context are returned to every accommodation owner, compiling missing details. Tourists can also use the Tourpedia App to search for accommodation and tourist attractions. The paper describes the strategy implemented to incentivize accommodation owners to release information about their activity. In addition, it describes how Tourpedia App is implemented and how tourists and accommodation owners can use it. The main finding of this study is the implementation of the Tourpedia App, a prototype that demonstrates that it is possible to build real applications based on open data.

## 1 INTRODUCTION

Tourism is one of the most important aspects of the economy of a country such as Italy. To confirm this, a study by The World Travel & Tourism Council (WTTC) states that in 2017 the total contribution of the tourism industry to the Italian economy was 223.2 billion euros, equal to 13% of the national GDP, generating 14.7% of total employment in the country. Furthermore, in Italy, the contribution of the tourism sector to the GDP was higher than both the European (10.3% of GDP) and the global average (10.4% of GDP). This means that for Italy, it is very important to invest resources in this sector.

This paper extends Tourpedia (Lo Duca and Marchetti, 2019), a knowledge base regarding tourism accommodation facilities (hotels, bed & breakfast, ...), built on open data extracted from regional government agencies of three countries: Italy, France, and Spain. Open data are datasets released through a public license, which permits their reuse and redistribution. So the main contribution of Tourpedia has been to build a new larger dataset, originated by integrating trustworthy resources, which can be redistributed.

The specific contribution of this paper involves the following aspects: 1) enriching Tourpedia with tourist attractions; 2) evaluating the quality of data contained in Tourpedia in terms of missing information and errors; and 3) defining a strategy to improve the data quality. This strategy is based on creating a Web application for accommodation owners, termed Tourpedia App, which allows them to check the information associated with their hotels and add any missing details; 4) multilingualism management.

Often, open data released by official government agencies are not exploited by anyone, thus, they remain on the owner's Website. Tourpedia could become a place where official government agencies disseminate their open data regarding tourism. Thus data becomes more accessible, visible, and exploitable.

Hoteliers and general users can use the Tourpedia App. Hoteliers can register their hotels in the Tourpedia App to access some statistics. These statistics include the presence of tourists in the hotels in the same municipality/region, the type of hotels selected by tourists, the number of tourists per travel category, and the number of Internet reservations.

<sup>id<sup>a</sup></sup> <https://orcid.org/0000-0002-5252-6966>

<sup>id<sup>b</sup></sup> <https://orcid.org/0000-0003-4512-1642>

Transport means used to travel and other related information. All these statistics define the accommodation context. As a side effect, the inclusion by the hoteliers of information relating to their activity allows Tourpedia to enrich the dataset relating to accommodation facilities with the missing information.

General users can also use the Tourpedia App to search for hotels close to a particular tourist attraction. Tourpedia App constitutes an alternative system to other commonly used search engines such as Google and Booking because it is based entirely on open data. Open data may contain accommodations and attractions not present on online travel agencies. Thus, users using the Tourpedia App could have additional information. At the moment, the Tourpedia App contains Italian, French, and Spanish accommodations and Italian tourist attractions. Statistics are available only for Italian hoteliers. However, they can be easily extended to other countries, provided that such countries have released open data statistics.

The remainder of the paper is organized as follows: Section 2 describes related work, and Section 3 illustrates the Tourpedia Knowledge Base. Section 4 illustrates the Tourpedia App and its implementation. Finally, Section 6 gives conclusions.

## 2 RELATED WORK

### 2.1 Data Quality Improvement

Many strategies exist in the literature to improve the data quality of knowledge bases. According to Heiko Paulheim, there are two main objectives of knowledge graph refinement: a) completion, which involves adding missing information, and error detection, which identifies wrong information (Paulheim 2017). The present paper tries to fulfill the completion objective by applying an alternative technique. Thus in the remainder of the section, only this aspect is investigated.

Completion can be achieved in two ways: a) internal methods, which exploit only the information contained in the knowledge base through Machine Learning algorithms, and b) external methods, which use information extracted from other sources. For a complete survey about complete, please refer to the previously cited paper (Paulheim, 2017). Other work determines completion based on query terms to the knowledge base (Jiang et al. 2019).

This paper describes an external method for knowledge base completion, based on the support of

accommodation owners', who will add missing information directly to the knowledge base.

### 2.2 Hotel Statistics

The problem of providing hoteliers with useful information to improve their business is well-known in the literature (e.g., Torres et al. 2014, Chen et al. 2019, Xia et al., 2019). One of the most important topics in this field regards hotel rating, which concerns the global hotel evaluation, according to some metrics. The literature on this aspect is vast: some works concentrate specifically on hotel rating (e.g., Casalo et al., 2105, Mariani et al., 2018, Narangajavana, 2008), and others analyze the impact of hotel reviews on customer satisfaction (e.g., Vermeulen 2009, Berezina, 2016) and hotel performance (Phillips et al. 2017). There are also many websites providing hotels with statistics about their reputation/rating. Some examples include ReviewPro, TrustYou and Olery. All these vendors elaborate great amounts of reviews from different sources to define their metrics, which define the hotel rank. The main issue with all these platforms regards the secrecy of how the metrics are calculated. Thus, the same hotel may have a different rank on each platform (Mellinas 2019). The Tourpedia App does not provide any ratings or rank about accommodation because it analyses the hotel business from another point of view. In detail, it provides every hotel with its accommodation context, i.e., information about the hotel surroundings (e.g., what tourists desire in the local area where the accommodation is located). The accommodation context is extracted from open data, so there are no secrets in its definition. This accommodation context is not shown to tourists but only to accommodation owners.

Other works focus on the quality of services the hotels provide (Su et al., 2007), such as room cleanliness and politeness of the staff (Rhee et al., 2015). This paper analyses the problem of providing hoteliers with useful information from the point of view of the accommodation context (i.e., statistics about the accommodation surroundings). This aspect could help hoteliers to understand which kind of tourists choose the area where the accommodation is. Thus hoteliers could improve or even add new services to their hotels.

### 2.3 Applications for Tourists

Many applications have been implemented to help tourists choose the right hotel for their trip. Apart from famous Online Travel Agencies, such as

Booking.com and TripAdvisor, which have a very complex business model, there are some less-known academic initiatives. One example was proposed by Al-Ghossein et al. (Al-Ghossein, 2018), who defined a framework that combines hotel reservations with events extracted from open data. The interesting aspect of this framework regards the possibility to select the point of view from which the search is done: hotel-centric, which permits the travelers to search for a hotel and then all the events in the surroundings are shown; event-centric, which permits the travelers to search for an event and then show all the hotels in the surroundings. The Tourpedia App is similar to Al-Ghossein's application because it provides two kinds of searches: hotel-centric and attraction-centric. Another important initiative is DataTourism, which implements a dashboard for tourists to discover information about a hotel (Soualah-Alila et al., 2016). Currently, in the Tourpedia App, the dashboard is associated only with accommodation owners. However, it could be interesting to extend it to tourists. Voyageur is another interesting search engine, giving tourists the possibility to carry out experiential searches, such as "quiet hotel and friendly staff" (Evensen et al. 2019).

### 3 THE TOURPEDIA KNOWLEDGE BASE

Tourpedia App exploits data contained in Tourpedia, a Web platform which exploits official open data about tourism provided by government agencies, such as Regions (e.g. Official Open Data released by the Tuscany Region<sup>1</sup>) or cities (e.g. Official Open Data released by the city of Matera<sup>2</sup>). The methodology of this paper involves four aspects: 1) enriching the Tourpedia knowledge base with tourist attractions, 2) evaluating data quality already contained in Tourpedia, 3) defining a strategy to improve data quality through the concept of accommodation context, and 4) multilingualism management. At this stage of implementation, we will not focus on scalability challenges, and we reserve to future work this kind of analysis.

#### 3.1 Current Version of Tourpedia

The main objective of Tourpedia consists of defining, designing, and implementing new technologies, which give a common structure to open data released by government agencies. Open data about tourism are

often distributed through different websites and in different formats or data structures. Tourpedia aims to unify all these open data to provide a single website to access open data regarding tourism. The system architecture behind Tourpedia comprises different modules, which enable a developer to add a new data source easily and without compromising the already imported sources. We described the general architecture of Tourpedia in our previous paper (Lo Duca et al. 2019). Tourpedia also provides a mechanism to map the original schema defined by a data source to the Tourpedia Data Model (TDM), through a simple mapping language called Tourpedia Mapping Language (TML). The TDM is very generic and can be extended easily to add new features. Currently, Tourpedia contains more than 70,000 accommodation facilities, collected by 21 of the official open data websites provided by Italian, French, and Spanish Regions (12 sources are from Italy, 6 from France, and 3 from Spain). All the available data is aggregated, updated continuously, stored in a local database, released under a public license, and can be accessed through a Web API.

#### 3.2 Tourist Attractions

Tourist attractions, as defined by Harris and Howard, are physical or cultural places that meet travelers' specific needs and influence their choice of destination (Harris 1996). These can be natural (climate, culture, landscapes) or site-specific (theater shows, museums). Medlik distinguishes between site attractions (place itself is the draw) and event attractions (festivals, performances), as well as natural (volcano, forest) and artificial attractions (museums) (Medlik 2003). While there is no standardized classification, any element in a destination can be an attraction if it drives visitors to meet personal needs, even unconventional ones like shopping centers. If it is almost impossible to achieve a standard classification of tourist attractions, it is instead possible to use criteria to establish the quality of a tourist attraction. Varra (Varra 2012, p. 21) speaks of three fundamental criteria, namely, the *replicability of the resource*, which is the possibility that it cannot be replicated elsewhere. The *uniqueness of the resource*, that is, the ability to qualify the territory differentiating it from others, and the *importance recognized* by stakeholders in creating value, that is, the contribution associated with it in the global production of value. Considering the criteria defined by Varra, Tourpedia was enriched with data

<sup>1</sup> <http://dati.toscana.it/>

<sup>2</sup> <http://dati.comune.matera.it/>

relating to tourist attractions extracted from open data owned by Italian regions and provinces. In particular, the specific contribution of this paper to *Tourpedia* consists of the following elements: a) addition to Tourpedia of tourist attractions related to the Italian regions and provinces; b) adaptation of the TDM to tourist attractions; c) extension of TML to support the following formats XSL, XML and KML. Since each dataset was provided in a different format, a formal mapping was defined from the original dataset to the Tourpedia schema. In general, considering all the datasets, 19 different types of information were identified to be extracted and inserted as fields in Tourpedia. The entire mapping is specified through TML. The original version of TML has been extended to allow the extraction of complex information, i.e., merging and splitting of two fields and simple correction of error misspelling. Regarding merging, it has been added the possibility to signal that a certain type of information is found in two or more fields of the dataset. Regarding splitting, TML has been extended to allow a field to be split into two or more fields during the extraction process. Regarding error misspellings, such as unformed URLs, TML has been extended to allow their correct insertion in the database.

### 3.3 Data Quality Evaluation

One of the most important data-related issues concerns their quality and availability (Pipino et al. 2002). Although a great amount of open data has been collected, in many cases, they are not complete, correct, or even missing. In addition, the distribution of specific information is inhomogeneous: while some accommodations/attractions present a sufficient/abundant quantity of data, in other cases, there are only some essential fields, such as the name, the region, and a single contact. In some cases, even this basic information is not available. In French and Spanish accommodation, geographical coordinates (latitude and longitude) are always available, while in Italy, they are available only in 65.35% of accommodation. In addition, there is a different sensibility among countries regarding the information provided, for example about rooms (completely missing in Spain), number of beds (available only in 49% of Italian accommodation), elevation and chain (available only for about 41% of French accommodation). The described analysis demonstrates how data extracted from Open Data are not complete. This means that a strategy to improve the quality of these data should be defined. A first solution could be based on data enrichment from

other sources, such as Online Travel Agencies (OTA). However, often, data regarding accommodation available on the OTAs are proprietary and cannot be reused for distribution. Another solution, which is exploited in this paper, could be to ask accommodation owners to provide missing information about their activities. In this case, accommodation owners should be incentivized to release their information, for example, by obtaining a reward. This paper defines a simple, rewarding strategy: providing accommodation owners with some statistics about their accommodation context. A comparison among the number of accommodations available on Tourpedia with those available on Booking.com and Tripadvisor, divided by region shows that Booking.com outperforms all the other Websites with 201,174 accommodations, followed by Tourpedia with 82,645 accommodations and then by Tripadvisor with 54,485 accommodations. Booking.com contains a sensible greater number of accommodations than Tourpedia and Tripadvisor for all the regions, but Marche and Comunidad Foral de Navarra, where Tourpedia outperforms the other websites.

### 3.4 Accommodation Context

Accommodation context concerns an analysis of the area where every accommodation is located, with all the tourism information, not directly depending on accommodation. Accommodation context could include tourist statistics, such as their interests in the area or which category of accommodation tourists select. Obviously, all accommodations located in the same area share the same accommodation context, while accommodations located in different areas have different accommodation contexts. In this paper, accommodation context is defined by tourism statistics extracted from open data released by the Italian Istituto Nazionale di Statistica (Istat). The use of open data is justified by the original idea of Tourpedia, which is only open-data-based. Extracted statistics range from a national scope to specifically covering the accommodation. At the moment, statistics are updated to 2016. To update them periodically, the architecture already defined in Tourpedia could be exploited. This aspect and integration with other open data sources will be the object of future studies. Statistics are divided into two categories: statistics on national tourism and specific data on every accommodation. Statistics on national tourism show different data at the national level: a) number of travelers by type of holiday and age group; b) type of accommodation chosen by travelers by type

of trip and by age; c) trend of internet bookings made from 2014 to 2016; d) means of transport most used by travelers. Specific data on every accommodation contain the following statistics: a) the trend of tourists' presence in accommodation at the national, regional, and provincial level; b) the number of tourists divided according to the type of trip; c) air traffic as the number of departing and arriving flights of the main airports in the region where the accommodation is located if they exist.

### 3.5 Multilingualism

Tourpedia extracted data from open government agencies of three different Countries: Italy, France, and Spain, which provide information in their native language (Italian, French, and Spanish). Tourpedia App proposes a mechanism to deal with multilingualism, exploiting the service provided by Geonames, a geographical database that can be downloaded and consulted offline. The proposed mechanism is based on a Resolver, which translates search queries made by users into Geonames IDs. In practice, accommodations contained in Tourpedia have been enriched with the Geonames ID of their associated city. The Geonames ID permits a place independent of the language related to the value stored in the Tourpedia knowledge base. A specific module, named Resolver, has been implemented. For each accommodation, the Resolver takes the city and country fields and searches them in the Geonames database. Then, the Resolver takes the first response of the Geonames database as a result and stores it in Tourpedia as an additional field of accommodation. If a query string does not give any result, the Geonames field is set to zero. As a result, 76.8% of records contained in Tourpedia were mapped to Geonames records. Tourpedia App exploits the Geonames ID to deal with search queries made in different languages (Fig. 1) Currently, search queries can be done only by place (i.e., the accommodation city). In the Tourpedia App, users write the place and the country names (parameters) in whatever language they want (e.g., Parigi, the Italian name of Paris, and France). Tourpedia App asks the Resolver to convert the parameters to their associated Geonames ID through the Geonames database. The obtained Geonames ID is used to access the Tourpedia knowledge base. As a result, Tourpedia returns to Tourpedia the list of records having the specified Geonames ID. Eventually, Tourpedia App gives the list of records to the user.

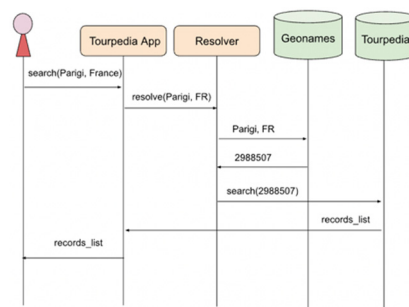


Figure 1: The sequence diagram for managing multilingualism in the Tourpedia App.

## 4 TOURPEDIA APP

Tourpedia App aims to fill in missing information regarding accommodation contained in Tourpedia. This objective is achieved through the implementation of a dashboard for accommodation owners, which gives them the possibility to verify the correctness of information regarding their accommodation and add missing ones.

Accommodation' owner who wants to update data contained in Tourpedia must sign a Tourpedia Data Release Agreement (TDRA), which specifies the terms of data publication. TDRA specifies the following aspects: a) inserted data corresponding to the reality, b) information added to Tourpedia by the accommodation' owner is released as open data through a Creative Commons CC 1.0 license, c) data can be withdrawn at any time. By signing the DRA, the accommodation owner accepts to add their data to the Tourpedia knowledge base and publish them as open data. Standard users can also use the Tourpedia App as a simple consultation tool to search for accommodation and attractions. To incentivize hoteliers to add their information to Tourpedia, the Tourpedia App provides them with the following benefits: a) it is completely free, and b) it contains information regarding the accommodation context.

### 4.1 Users

Tourpedia App is envisaged for two types of users: tourists and accommodation owners. Potentially, tourists are interested in all data contained in Tourpedia. Accommodation owners, instead, are interested only in data related to their activity. Tourists are the main beneficiaries of the Tourpedia App because they exploit data contained in the Web application to organize their trips, select the best accommodation for their holidays, and so on. In

detail, every tourist can search for accommodations or tourist attractions located in a place and select their preferred one based on its geographic position (Fig. 2) As a result, the Tourpedia App returns the list of all accommodations and attractions matching the search, and for each of them, it provides as much information as possible. Once the tourist chooses the best accommodation/attraction, they can contact it through the given Website/social network or contact it directly by telephone/email. A tourist could also be interested in making comparisons, for example, on the accommodation closest to places of interest or attractions, or, vice versa, on the most interesting attractions to visit near the structure where they intend to stay. Currently, Tourpedia App does not provide any mechanism to reserve a room in a hotel because of the lack of financial resources to implement this aspect.

A service provider is a juridical entity that provides any kind of service to tourists. Different categories of service providers can be imagined, such as accommodation owners, museums, archives, etc. Concerning tourists, service providers have different needs related to the management and verification of their data. Fig. 3 describes the flow diagram for service providers. A service provider is mainly interested in checking that the information about his/her service is correct, clear, and comprehensive to attract tourists. In addition, they should be interested in comparing his/her service with those of other competing structures. Tourpedia App provides him/her with an external dashboard capable of allowing him/her to add, modify, and improve data. Regarding accommodation, the dashboard provides additional information related to the accommodation context.

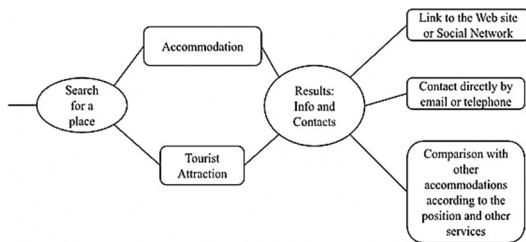


Figure 2: Flow diagram of Tourpedia App for tourists.

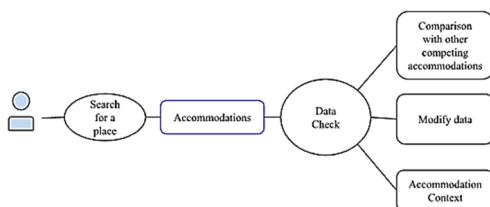


Figure 3: Flow diagram of Tourpedia App for service providers.

## 4.2 Navigation

A user can search for a location by filling in the text input. As output, the Tourpedia App shows on a geographic map all the accommodations and attractions located near that location. Accommodations and attractions are represented through markers (beds and amphitheatres, respectively). The nearest markers are grouped into clusters to make viewing and browsing easier and faster. On the left part of the screen, there is a list of accommodations and attractions. The two lists are maintained separately and can be displayed only one at a time. However, it is possible to switch from one to the other with a click at any time, determining, among other things, a change of color theme in the entire window. The lists are organized into numbered pages, each containing a maximum of six previews. These are boxes within which only the essential information for each accommodation or attraction is provided: name, address, category (only for attractions), and an image, replaced by a Google Street View snapshot in case it is not in the database. Such an image is calculated using the latitude and longitude of the accommodation/attraction.

If the user clicks on one of the elements of the lists (or marker), all the information about that accommodation/attraction is shown. Information is organized in two sections: the first one describes the name, the associated image, and the category (only for attractions); the second one instead, contains details (address, telephone number, fax, e-mail, website). Only for accommodation, there is a further link, which permits a user to claim to be the accommodation owner. This link connects the Web application to the Dashboard, described in the next section. Accommodation also has a third section, where additional information is listed and divided into categories. After a brief description of the accommodation, it is possible to find data relating to the services offered within it, such as the presence of Wi-Fi or private parking, the number of rooms, suites, or beds available, and languages spoken by the staff. The section also has a slideshow showing the main attractions in its vicinity, navigable by two arrows or by scrolling horizontally.

## 4.3 Dashboard

To mitigate the problem of missing information contained in Tourpedia, the Tourpedia App implements a dashboard for service providers to allow them to contribute to the improvement and enrichment of their data. Currently, the dashboard is envisaged only for accommodation owners.

However, the service could be easily enlarged to other service providers. Since the goal of the Tourpedia App is to have data as correct and complete as possible, only service providers are allowed to modify/add information. This means that, at the moment, simple users cannot suggest changes or additions to the data precisely because the Tourpedia App currently has no resources to manage this aspect.

Since each accommodation owner should be allowed to change only and exclusively the data relating to his/her accommodation, they are provided with specific credentials (username and password), which are randomly generated and sent by email through a procedure described later in the paper. The dashboard can be accessed directly from the details associated with the accommodation on the left part of the map.

Once logged, the dashboard displays a page where the user can enter or modify the data relating to their accommodation. Such a page is divided into two sections. In the first, the blue one, there are the fields related to the basic data, i.e., name, category, address, stars, telephone, email, country, region, province, code postal, city, latitude, and longitude. The second section is longer than the previous one and contains all the additional information that an accommodation owner may decide to insert, i.e., website, fax, possible locations or parts of the city, social contacts, opening period, number of beds, rooms and any suites.

The second part consists of four panels: accommodation facilities, languages spoken by the staff, a photo of the accommodation, and a description box. Finally, the accommodation owner must sign a TDRA. The accommodation owner can withdraw information related to their activity simply by updating their profile with empty information. Once the user submits the form, some statistics about his/her accommodation context are shown, as described in Section 4.1. This is done to encourage accommodation owners to add information about their accommodation in Tourpedia. To advertise the dashboard, an automatic mail-sending system has been set up, accompanied by a random account generation algorithm. The system takes incoming emails from the accommodation structures contained in the database and automatically builds an email for each of them, containing a unique username and password to access the dashboard. At the moment, emails have not been sent yet. Thus we defer to future work on analyzing this aspect, although the system is ready.

#### 4.4 Evaluation

Many criteria exist to evaluate a Web site about tourism (Law, 2010). They can be summarized in the

following six criteria: authority, coverage, currency, objectivity, and accuracy (Dalhousie University).

Authority specifies that the person responsible for a site (e.g., person, institution, organization) has the qualifications and knowledge to do so. Since the Tourpedia App exploits data deriving from other sources, its authority regards data source authority. Data are extracted from open data released by official governments and institutions, the Tourpedia App can be considered an authoritative source.

Purpose defines the objective of the Web site. In the case of the Tourpedia App, the purpose is to help tourists to search for accommodation in a given place. The Google web interface helps tourists to understand how the application works. Although the purpose is very clear, the Tourpedia App has many limitations because it is only a demonstrative prototype, demonstrating that open data can be used to build applications about tourism. This means that the purpose of the Tourpedia App is not entirely achieved.

Coverage is difficult to measure because it defines how a Web site covers a topic. From the point of view of the Tourpedia App, coverage regards two aspects: data coverage and application coverage. Data coverage regards the number of accommodations available in the Tourpedia App. This aspect depends on data available in Tourpedia, limited to three Countries: Italy, France, and Spain. Thus data coverage is very limited. Application coverage can be defined as the number of features the application provides, such as the types of searches. In the Tourpedia App, only searches about a place can be done. Thus application coverage is limited.

Objectivity should define whether or not there are some biases or errors. In the case of the Tourpedia App, objectivity is directly connected to data objectivity, and this aspect does not depend directly on the application but on the data providers. Through the dashboard, the Tourpedia App tries to mitigate this problem. Accuracy defines how the Web site is accurate, i.e., information is precise. Similarly to objectivity, accuracy also depends on the accuracy defined by data. Since data is derived from different sources, there is a difference in data accuracy. Thus there is some data very accurate and others with many errors. Also, in this case, the Tourpedia App tries to mitigate this problem through the dashboard.

## 5 CONCLUSIONS

This paper aimed at fulfilling four main objectives. Firstly, it described the concept of tourist attraction and how Tourpedia was enriched through tourist

attractions extracted from open data provided by official Italian municipalities. Secondly, it evaluated the quality of data contained in Tourpedia. The analysis demonstrated that extracted open data are never complete. Thirdly, to add missing information to Tourpedia, a strategy to boost accommodation owners' release of such information about their activities was defined. This strategy was based on implementing the Tourpedia App, a Web application that provides accommodation owners with a dashboard to consult the accommodation context associated with their activity. The dashboard encourages accommodation owners to update their profile on the application. Finally, a strategy to manage multilingualism was proposed based on exploiting an external database (i.e., Geonames).

Tourpedia could pave the way towards an alternative way of thinking, not based on the proprietary market but on the sharing of common knowledge. This could create a more shared system, which is not only in the hands of a few people. We know that the Tourpedia App is only a prototype, which cannot compete with other systems exploiting proprietary data. However, the implementation of this application is to demonstrate that it is possible to build Web applications based entirely on open data.

## REFERENCES

- Al-Ghossein, M., Abdessalem, T., & Barré, A. (2018). Open data in the hotel industry: leveraging forthcoming events for hotel recommendation. *Information Technology & Tourism*, 20(1-4), 191-216.
- Berezina, K., Bilgihan, A., Cobanoglu, C., & Okumus, F. (2016). Understanding satisfied and dissatisfied hotel customers: text mining of online hotel reviews. *Journal of Hospitality Marketing & Management*, 25(1), 1-24.
- Casalo, L. V., Flavian, C., Guinaliu, M., & Ekinci, Y. (2015). Do online hotel rating schemes influence booking behaviors?. *International Journal of Hospitality Management*, 49, 28-36.
- Chen, M.-C., Hsiao, Y.-H., Chang, K.-C. and Lin, M.-K. (2019), "Applying big data analytics to support Kansei engineering for hotel service development", *Data Technologies and Applications*, Vol. 53 No. 1, pp. 33-57. Dalhousie University. Online Document. 6 criteria for Web sites. Available online: [https://cdn.dal.ca/content/dam/dalhousie/pdf/library/CoreSkills/6\\_Criteria\\_for\\_Web\\_sites.pdf](https://cdn.dal.ca/content/dam/dalhousie/pdf/library/CoreSkills/6_Criteria_for_Web_sites.pdf) (Last access 2020/01/31).
- Evensen, S., Feng, A., Halevy, A., Li, J., Li, V., Li, Y., ... & Pavlovic, E. (2019, May). Voyageur: An experiential travel search engine. In *The World Wide Web Conference* (pp. 3511-5).
- Harris, R. and Howard, J. (1996). *Dictionary of Travel, Tourism and Hospitality Terms*. Elsternwick, Hospitality Press Pty Ltd.
- Jiang, S., Liang, J., Xiao, Y., Tang, H., Huang, H., & Tan, J. (2019). Towards the Completion of a Domain-Specific Knowledge Base with Emerging Query Terms. In *2019 IEEE 35th International Conference on Data Engineering (ICDE)* (pp. 1430-1441). IEEE.
- Law, R., Qi, S., & Buhalis, D. (2010). Progress in tourism management: A review of website evaluation in tourism research. *Tourism management*, 31(3), 297-313.
- Lo Duca, A. and Marchetti, A. (2019), "Open data for tourism: the case of Tourpedia", *Journal of Hospitality and Tourism Technology*, Vol. 10 No. 3, pp. 351-368.
- Mariani, M. M., & Borghi, M. (2018). Effects of the Booking.com rating system: Bringing hotel class into the picture. *Tourism Management*, 66, 47-52.
- Medlik, S. 2003. *Dictionary of Travel, Tourism and Hospitality*. Burlington, Elsevier Science.
- Narangajavana, Y., & Hu, B. (2008). The relationship between the hotel rating system, service quality improvement, and hotel performance changes: A canonical analysis of hotels in Thailand. *Journal of quality assurance in hospitality & tourism*, 9(1), 34-56.
- Juan Pedro Mellinas, Sofia Reino, (2019) "Average scores integration in official star rating scheme", *Journal of Hospitality and Tourism Technology*, <https://doi.org/10.1108/JHTT-07-2017-0050>.
- Paulheim, H. (2017). Knowledge graph refinement: A survey of approaches and evaluation methods. *Semantic web*, 8(3), 489-508.
- Pipino, L. L., Lee, Y. W., & Wang, R. Y. (2002). Data quality assessment. *Communications of the ACM*, 45(4), 211-218.
- Phillips, P., Barnes, S., Zigan, K., & Schegg, R. (2017). Understanding the impact of online reviews on hotel performance: an empirical analysis. *Journal of Travel Research*, 56(2), 235-249.
- Rhee, H. T. and Yang, S. B. (2015). "How does hotel attribute importance vary among different travelers? An exploratory case study based on a conjoint analysis". *Electronic markets*, 25(3), 211-226.
- Soualah-Alila, F., Coustaty, M., Rempulski, N., & Doucet, A. (2016). DataTourism: designing an architecture to process tourism data. In *Information and communication technologies in tourism 2016* (pp. 751-763). Springer, Cham.
- Su, C. S., & Sun, L. H. (2007). Taiwan's hotel rating system: A service quality perspective. *Cornell Hotel and Restaurant Administration Quarterly*, 48(4), 392-401.
- Torres, E.N, Adler, H., Behnke, C. (2014) Stars, diamonds, and other shiny things: The use of expert and consumer feedback in the hotel industry, *Journal of Hospitality and Tourism Management*, Volume 21, 2014, Pages 34-43, ISSN 1447-6770
- Varra, Lucia (a cura di). 2012. *Dal dato diffuso alla conoscenza condivisa*. Firenze, University Press.
- Vermeulen, I. E., & Seegers, D. (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. *Tourism management*, 30(1), 123-127.
- Xia, H., Vu, H.Q., Lan, Q., Law, R. & Li, G. (2019) Identifying hotel competitiveness based on hotel feature ratings, *Journal of Hospitality Marketing & Management*, 28:1, 81-100.