Smart Spaces Based Construction and Personalization of Recommendation Services for Historical e-Tourism

Oksana B. Petrina

Faculty of Mathematics and Information Technology
Petrozavodsk State University
Petrozavodsk, Russia
Email: petrina@cs.karelia.ru

Aleksey G. Varfolomeyev

Faculty of Mathematics and Information Technology
Petrozavodsk State University
Petrozavodsk, Russia
Email: avarf@petrsu.ru

Dmitry G. Korzun

Department of Computer Science
Petrozavodsk State University
Petrozavodsk, Russia
Email: dkorzun@cs.karelia.ru

Aleksandrs Ivanovs

Department of History
Daugavpils University
Daugavpils, Latvia
Research Institute for Regional Studies (REGI)
Rezekne University
Rezekne, Latvia
Email: aleksandrs.ivanovs@du.lv

***Abstract*—The smart spaces approach enables development of advanced digital services referred now as "smart services". In this paper, we discuss the development problem of smart services for the domain of historical e-Tourism. We show the applicability of recommender systems for constructing and personalizing such services within a smart space. The latter is created to accompany a tourist or a group of them. A corpus of historical data distributed among multiple sources in the Internet is collected the tourist's smart space, forming a kind of custom and dynamic historical knowledge. In particular, context-aware semantic relations between historical objects are established and manipulated. This semantics-rich information is an input for recommendation-making. Personalized recommendations with quantitative and qualitative estimates can be constructed. The result then is visualized to assist a historian tourist in historical analysis of points of interests. The contribution consists of the concept definition and system design for creating a smart space with assistance services of personalized recommendations.**

Keywords–Historical e-Tourism; Cultural Heritage; Point of Interest; Recommender Systems; Semantic Network; Smart Spaces; Personalized Services.

I. Introduction

This work extends the results presented in our UBICOMM 2015 paper [1] on the role and smart spaces based design of e-Tourism recommender systems. A lot of research has been already done, especially in the direction of making the services of e-Tourism systems mobile and intelligent [2], [3]. A topical subdomain is historical tourism [4], which we distinguish from more general cultural heritage tourism. In particular, the historical tourism focuses on visiting historical Points of Interest (POI) and on studying their history-aware relation with other historical objects (POIs, events, persons, etc.).

For a historian tourist, a POI is recommended not only if it is nearby and within user's interests. Such a tourist would like to see a spectrum of historically related POIs; some are closely located and some can be faraway. Clustering important POIs for recommendation needs semantic analysis of their relation with historical objects and can be performed by means of ontologies [5]. The content for reasoning about, however, must be extracted from some historical databases or digital archives. Moreover, some historical relations are subjective, e.g., depend on context or personal vision of historical facts.

The study presented in this paper is motivated by the lack of "smart" assistants for historian tourists, although there is a lot of them for mobile e-Tourism in general [2]. Based on our previous work [1], [6]–[9], we expect that development of recommendation services with built-in semantic analysis of historical data can be implemented using ontology-based technologies of the Semantic Web. Furthermore, traditional web-based architectures and mobile standalone applications seem insufficient for this development. We focus on the emerging approach of smart spaces [10], [11]. A ubiquitous computing environment is created where mobile users, multisource data, and various services constructed over these data are intelligently connected based on ontology-driven information sharing and self-generation. Services can be personalized by means of augmentation of personal data to the shared content and customization of required reasoning about the content.

In this paper, we analyze the historical POIs recommendation problem. Our research scope is limited to such important phases of service development as concept definition and system design. To define the concept we provide a reference scenario of recommendation services for historical e-Tourism. Additionally, we explain the concept by a detailed example. The proposed system design consists of a smart space based architecture and operation description for participating software agents. The proposed design supports multiple data sources and adopts various ranking methods. This contribution enables creation of a personalized smart space with services assisting its historian tourists by means of recommendations. We also perform experimental performance evaluation for iterative acquisition of historical knowledge from available data sources in the Internet.

The rest of the paper is organized as follows. Section II discusses the related work motivating our study of POI recommendation services in historical tourism. Section III introduces our reference service scenario for historical e-Tourism. Section IV illustrates by a detailed example the semantic network construction and its use for POI ranking. Section V describes the smart space based system design. Section VI provides details of the software agents that cooperatively construct the recommendation service in the smart space. Section VII analyses the proposed concept definition and system design using experiments and contrasting with the existing solutions. Section VIII concludes the paper.

II. Motivation and Related Work

Historical tourism has distinctive features [4] compared with the general application domain of cultural heritage tourism. The latter embraces both historical and present-day cultural phenomena. According to Nora [12], historical tourism addresses the so-called "sites of memory". They present any material traces of historical events, which sometimes coincide with cultural heritage artifacts. For instance, an architectural monument is directly "involved" in historical developments related to its construction. Another example is any place or a spot associated with a historical event. Traces of historical facts are presented in the multitude of historical sources, including open sources in the Internet.

In general, a point of interest (or attraction) is an actual spot with precise localization on the geographical map (e.g., geo-position coordinates or postal address). Nowadays, POI recommender systems form an important services class in e-Tourism [2]. In addition to POIs, historical tourism takes into consideration a lot of other historical-valued objects such as persons, events, and data sources (written records and narratives, artifacts, alternative information sources, data and knowledge bases available on the Web, etc.). Relations between historical objects define important semantics [13]. Moreover, any historical event might be conditionally defined as a semantic relation between several historical objects [14]. Ontologies become of high application interest for knowledge representation and reasoning in historical research [5] and e-Tourism [3].

In historical tourism, we expect that semantic relations can be effectively represented and manipulated using the technologies of the Semantic Web. To the best of our knowledge, no specialized knowledge base that comprise semantically enriched information about historical objects has been created yet, e.g., see [15]. To a certain extent, a corpus of historical information is represented in the ontological form in such knowledge bases for cultural heritage as DBpedia, Freebase, or YAGO. Additional information can be extracted from web publications of historical sources [6], [16]. In these settings, the methods of web-based systems, mobile programming, and multi-agent systems provide effective means for implementation of data search, access, and reasoning [2], [3].

Semantic Web methods and technologies help to solve the problems of creation, design, enrichment, editing, retrieval, analysis and presentation of historical information [15]. There are mobile services for cultural heritage e-Tourism developed using semantic technologies, e.g., see the review in [17]. For instance, an intelligent tourist guide [18] utilizes cultural heritage information. Nevertheless, the present-day application developments do not take into account the principal peculiarity of historical tourism—semantic relations among historical data. The problem of the semantic-aware retrieval information from multiple data sources can be settled by using semantic technologies, e.g., by parsing the query and ranking the relevance of content [19]. However, effective POI ranking requires the combination of different algorithms. One of the ranking parameters can be the recalls of the users with similar interests that have been posted on a Smart Tourism Website [20]. In addition to users recalls, other parameters should be taken into account, e.g., such context attributes of visited places as time and weather.

Methods of ubiquitous computing and, in particular, the recent progress in communication technologies of the Internet of Things make possible creating environments where diverse devices and computer systems cooperatively construct services surrounding the user [10], [11], [21]. New programming paradigms emerge, such as smart spaces. A smart space supports cooperation by establishing a shared view of resources in the environment. The shared view is ontology-based, applying the technologies of the Semantic Web. For e-Tourism services, a smart space is mobile and personalized, i.e., created around a traveling tourist, attracting appropriate web services and other data sources from the Internet [6], [22], [23]. In particular, smart service attributes and their smart space based implementation were proposed in [9]. As a result, these attributes introduce a new level of adaptation and personalization for e-Tourism services.

The discussion above motivates our research focus on POI recommendation services for historical tourism. First, semantic relations between historical objects cannot be bypassed in recommendation making. Second, there is no single source of needed information. The latter is distributed within multiple sources, each represents the information either in ontological form or requires an extraction procedure. Third, a historian tourist needs personalized services, i.e., source information and the result are subject to her/his preferences and context. Last but not least, a recommendation service is "ubiquitous", i.e., the service accompanies a mobile tourist in the anywhere and anytime style. In this paper, we propose the concept model and system design for the smart space based implementation of recommendation services for historical tourism. The key advantage of the proposed solutions is their support for various smart services attributes.

III. Recommendation Scenario

Consider a historical POI recommendation service that provides personal assistance for a tourist during her/his journey. This way, we introduce a reference scenario that defines our concept model of a historical e-Tourism service. Table I summarizes the formal symbol notation.

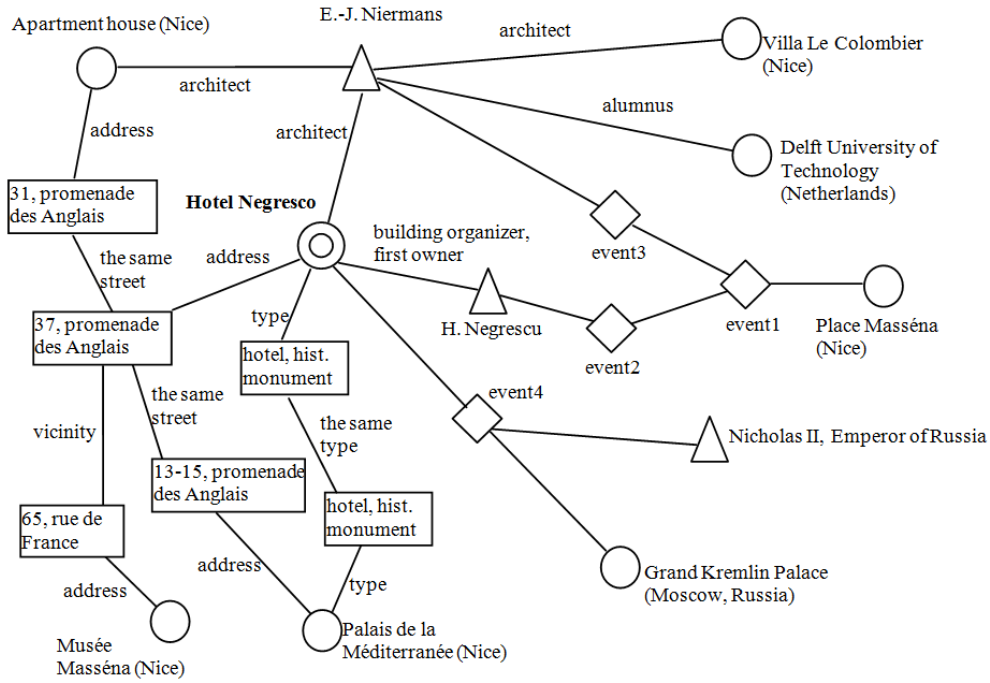

Figure 1. Sample semantic network: historical relations built around Hotel Negresco.

Let P consist of POIs that are accessed from multiple sources. Typical examples of historical POIs are buildings, monuments, fountains, bridges, squares, etc. Spacious objects, such as streets or rivers, are not POIs. As a rule, a historical POI has a particular name. There can be several names associated with a given $p \in P$, e.g., due to historical developments or due to the use of different languages. Each POI has distinctive properties: coordinates and/or address, date, architect, etc. In the service, POIs and other historical objects form the set H.

A historical POI recommendation is essentially based on relations between the elements of H. First, "direct" links exist between historical objects. For instance, a person x is the architect of a building y. Second, links can appear between objects due to similarity. For instance, two buildings have been constructed by the same architect or they are located on the same street. Third, links are a result of involving diverse objects and POIs in a common historical event. Therefore, a semantic network G with nodes H can be constructed.

TABLE I. SYMBOL NOTATION.

Symbol	*Description*
H	The set of all historical objects $H = H_1 \cup \ldots \cup H_n$ acquired from n data sources. Overlapping $H_i \cap H_j \neq \varnothing$ is possible.
$P \subset H$	The set of all POIs $P = P_1 \cup \ldots \cup P_n$, where $P_i \subset H_i$ for any data source i.
G, G_p	Semantic network G where nodes are from H and links are historical relations. In G_p, an initial POI $p \in P$ is fixed.
O	Ontology O describes the historical domain: possible classes and properties of historical objects as well as relations and restrictions for them.
R_p	Star graph R_p is a POI recommendation, where the internal node is the initial POI $p \in P$ and leaves are recommended POIs $q \in P$.
$r_q > 0$	Real-valued rank r_q shows the recommendation degree of $q \in P$ in respect to the initial POI p.
t_q	Annotation t_q summarizes (in a human-readable form) the reason of recommending q if the initial POI is p.

The historical relation semantics have a personalized character: some relations are treated differently by different historians or dependently on a context. For instance, there can be several visions of the role of a person for a certain POI. An important context corresponds to an initial POI; a tourist selects $p \in P$ and considers a semantic graph G_p. The all three types of links mentioned above can be a subject to context consideration and personalization.

Figure 1 shows a sample semantic network that is built manually around Hotel Negresco, one of the most famous buildings of Nice. Small circles are POIs, triangles denote historical persons, text rectangles describe POI properties, and rhombuses represent historical events. The initial POI ($p =$ "Hotel Negresco") is linked with seven other POIs (five of them are located in Nice). The links are based on different properties: one and the same architect, close location, involvement in common historical events, etc. Hotel information is extracted mainly from Wikipedia (see the next section for detailed description).

Based on G_p, a tourist would like to understand which POIs are interesting for her/his personal consideration from historical perspective. An important context for this understanding is that she/he starts from p (e.g., being actually or virtually in this POI). The recommendation result can be represented as a star graph R_p. Its internal node is p and the leaves represent all recommended POIs $q \in P$. Ranks $r_q > 0$ can be associated with the POIs to describe the recommendation degree of q (the higher rank the more recommendable). In a visual representation of G_p the length of edge (p, q) is proportional to the rank. Additional annotations t_q, which describe the reason of recommendation (in an aggregative form), can be also associated. Visual layout of R_p can also take into account the geographical position of q in respect to p (e.g., when q is on the North-East of p).

An example of recommendation is shown in Figure 2 in the form of a star graph derived from the semantic network (it was shown in Figure 1). The example star graph has been built by hand. With this example, we considered the possibility of processing unstructured data sources [8]. If text information is automatically processed, then the star graph will be identical. Geographical positions of the POIs are not reflected.

Now we can formulate our reference scenario of recommendation services for historical e-Tourism as consisting of the following steps.

Step 1: Initial POI selection. It can be made either manually (e.g., pointing out coordinates, a spot on the map, or POI's name) or automatically (e.g., within a definite area pointing out the nearest POI). The area can be either set by the tourist (for instance, on a map), or determined automatically taking into consideration the location of the tourist.

Step 2: Semantic network around the initial POI. The sets H and $P \subset H$ as well as semantic relations among them are searched for and retrieved from available knowledge bases and other data sources. Since the network G_p is potentially infinite, the process is limited. For instance, if the construction reaches another POI q from p, then the search for additional historical objects interconnected with q is terminated. Note that the path from p to q is subject to analysis in order to derive the reason of recommending q (construction of an annotation t_q). This example of limiting the construction process straightforwardly leads to a star graph R_p.

Step 3: POI ranking. Differentiation of recommended POIs can be based on ranks r_q. They are computed based on tourist's preferences. For instance, she/he needs to find a building constructed by the same architect, an edifice that built in the same architectural style, or another historical building located on the same street. Such preferences can be defined in the user's profile. They can be manually defined for the initial POI (before the implementation of the second scenario step). A significant component of the user's profile is the history of the choices of previous initial POIs (e.g., history of visits). For instance, the previously chosen POIs acquire lower ranks, since these POIs should not be repeatedly recommended.

Step 4: Visualization. The recommendation results achieved on Steps 2 and 3 are visually presented in a user friendly way, i.e., by means of a star graph possibly augmented with a map and textual/visual descriptions. For instance, annotations t_q show the reasons of the provided POI recommendations.

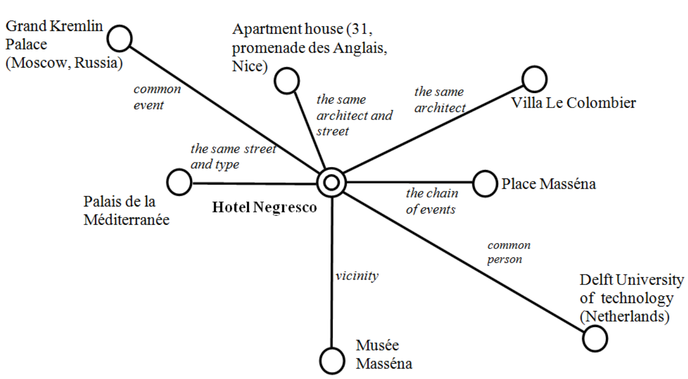

Figure 2. Star graph example: Hotel Negresco and recommendations on its historical surrounding.

Step 5: Feedback. The recommendation process is iterative. Based on the presented results (the star graph with ranks and annotations), the tourist supplements this information by expanding the semantic network G (additional data retrieved from historical sources). The process—supplementing and expanding network—is represented in G: new historical events appear in which both the user and the objects are involved. The user becomes a historical person—a network node in G.

In summary, this reference scenario defines our concept model of a recommendation service for the considered class of recommender systems in historical e-Tourism.

IV. Example of Semantic Network Construction and POI Ranking

The previous section showed the construction of a semantic network around POIs. Let us now illustrate some important details of the construction. We continue the handmade example on Hotel Negresco from Section III. The example reflects the operations to be implemented in the proposed smart space based system design.

The initial POI of the semantic network in Figure 1 is Hotel Negresco. It attracted our attention by its position in the center of Nice and its role of an architectural symbol, the most popular siteseeng in that city. Hotel information is widely represented on several pages in Wikipedia. Not surprisingly, the French version is the most complete one, but it is quite interesting that the same information from different pages is slightly different: for example, there are few differences in the histories how the unique chandelier under the hotel canopy is linked to the Russian Emperor Nicholas II. The Russian version simply states that two chandeliers were made. The first one is for Nicholas II, who placed it in the Grand Kremlin Palace. The second one is for the hotel. The French version says that the chandelier in the hotel was also originally intended to the Russian Tsar. The English version sais that Nicholas II was not able to take the chandelier because of the October Revolution.

By this example, we can see the deficiencies of Wikipedia as a source of historical information. There are no precise references to the sources. The connection between Hotel Negresco, Nicholas II, and the Grand Kremlin Palace is expressed using "event4" event, as it is shown in Figure 1.

From the text of Wikipedia pages, it should be noted that the hotel is connected with two bright personalities: 1) the author of the idea, the customer of the construction, the first owner of the hotel, the Romanian adventurer and businessman Henri Negresco, and 2) Dutch architect Edouard-Jean Niermans. There are other persons as well, for example, Jeanne Augier, the owner of the hotel in the last decades, who has breathed new life into the hotel, aligning it with the art museum. There are also many links away with directors and actors through the movies that were filmed there, or with great people who lived in the hotel. That is, the semantic network around the hotel could be much more extended. We decided to stay at Negresco and Niermans because additional links were discovered between them, and quite an interesting network with no other persons is constructed. In practice, limiting the expansion of the semantic network can be implemented by input of specific parameters, e.g., the network depth or the number of POIs.

Shortly before the start of construction of the hotel Henry Negresco had become the director of the restaurant in Municipal Casino in Nice. In this period, Niermans performed work on the reconstruction of the premises in that building. Apparently, they met there; as a result Niermans was embodied in the idea to build the hotel. Unfortunately, the Municipal Casino of Nice has not survived to the present time; it was demolished in 1979. It was located on Place Massena, now this place is just a square. The existing connection between the Place Massena, Negresco and Niermans is expressed in Figure 1 by three events: “event2” and “event3” link Niermans and Negresco with the Municipal Casino (the historical object that do not exist now), and “event1” connects the Municipal Casino with Massena Square (the event is that there was casino that had been constructed and then disassembled). The last case can also be described as two events.

In addition to Hotel Negresco, Niermans was the architect of many other buildings, but there are few of them in Nice. For example, the apartment house, which is not far from the hotel (on the same street Promenade des Anglais), and Villa Le Colombier, which is quite far from the center (it was built by the architect for his daughter). These objects do not have their own pages on Wikipedia, but the Niermans’ page has references to them. In addition, Niermans’ page describes that Niermans was graduated from Delft University of Technology.

In addition to these connections of Hotel Negresco, the semantic network also embraces other historical monuments, which are located in the nearest neighbourhood like Villa (museum) Massena, or located on the same street as the historic hotel-casino Mediterranean Palace.

Thus, the initial POI is associated with seven related POIs, five of which are in Nice, the sixth one is the Grand Kremlin Palace in Moscow, and the seventh one is Delft University of Technology. Of course, the location (proximity to the initial POI) should influence the rank of an object: the closer the POI, the higher the rank. In addition, it is necessary to take into account the convenience of the route - it is the best when objects are located on the same street. By proximity and convenience the POIs could be ranked as follows:

1) Musée Masséna;
2) Apartment house (31, promenade des Anglais, Nice);
3) Palais de la Méditerranée;
4) Place Masséna;
5) Villa Le Colombier;
6) Delft University of Technology;
7) Grand Kremlin Palace.

It is also necessary to take into account the degree of interest for a historian tourist. This degree can be estimated by the number of connections and the level of facts saturation. For example, the story of Nicholas II and the chandelier is much fuller of facts relevant to the initial POI, than the mention of the fact that the architect Niermans studied in Delft. The same can be said about the Place Massena in comparison with any other POI. If Casino Municipal was preserved, it certainly would have to take the top spot in the ranking. However, as a look at the square, where once there was this casino, still is not as interesting as a building, in the final ranking, it would be put on the 3rd place:

1) Apartment house (31, promenade des Anglais, Nice) – the same street and the same architect;
2) Palais de la Méditerranée – the same street and the same type of building (historical hotel-casino);
3) Place Masséna – rich history closely related with the initial POI, but now it is not so interesting;
4) Musée Masséna – very close to the initial POI, but the relation is rather weak (it is also a historical monument);
5) Villa Le Colombier – it is interesting to see it because it was built by the same architect, but it is very far from the initial POI.

V. Smart Space based Design

The smart spaces paradigm considers computing networked environments equipped with a variety of devices and with access to the Internet [10], [11], [24]. Software agents act as knowledge processors (KPs) running on the devices and interacting via information sharing. A semantic information broker (SIB) is a mediator for information collection and exchange. Each KP produces its share of information and makes it available to others via the SIB. Similarly via the SIB, a KP consumes information of its own interest. The information storage employs RDF (Resource Description Framework) [25]. Agents can apply such advanced Semantic Web technologies as SPARQL Protocol and RDF Query Language or Web Ontology Language (OWL) for shared information maintenance, search, and reasoning [26].

This programming paradigm suits well for the development of e-Tourism services, as recent work [6], [22], [23] indicated. Figure 3 shows the multi-agent system architecture that we adopted from [6] for the case of historical recommendation services. Data source KPs for historical sources provide the smart space with extracted historical data. Client KPs allow the user to participate in the smart space receiving and consuming its services. Semantic network combiner constructs personalized

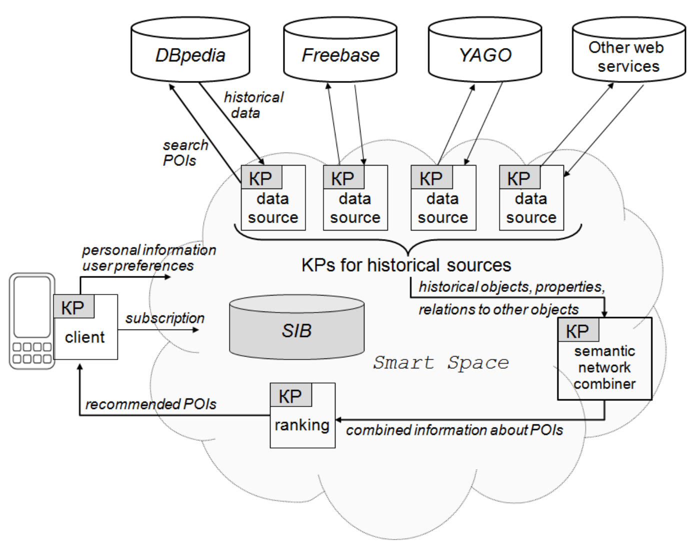

Figure 3. Multi-agent system architecture: historical information from multiple sources and other relevant information are semantically related and collectively processed in the smart space.

structures of POIs and builds the semantic network. Ranking KP makes ranking POIs using different ranking algorithms.

The recommendation service is constructed by cooperative work of multiple KPs on historical data and other information. Consider the properties of the proposed architecture to analyze the advantages that the smart spaces approach provides to the development of recommendation services for historical tourism. Some advantages are valid even for the more general case of e-Tourism.

Popular architectural styles for e-Tourism recommender systems are web-based, agent-based, and mobile [2], [3]. Smart spaces support them to be applied in a composition. SIB is deployed on a host machine in the Internet, similarly as it happens with web services now. Each user (tourist) is mobile, acts using her/his client KP (e.g., on smartphone or tablet), and consumes the service anywhere and anytime. Other KPs produce the information collecting it in the smart space for the use by the service. They can be hosted on the same machine with SIB (web-like solution) or on other computers (agent-based solution). The latter property leads to higher flexibility for system deployment. For instance, some KPs are provided by a travel agency and some KPs are attached from the user side in order to augment the system for personalized operation.

The recommendation service becomes not attached to a fixed source of historical information. A wide pool of available sources is used, where a data source KP is assigned per source (DBpedia, Freebase, YAGO, etc.). Configuration of the pool is flexible and a subject to dynamic inclusion/exclusion. Some data source KPs are set up by system administrators. Some KPs can be attached by the users if the appropriate rights are delegated. Each data source KP has to implement its source-specific interface to access and search for information. Note that a client KP can also provide historical information to the smart space, in addition to her/his preferences, context, and control. The information is further used for personalization.

The function of the data source KPs is to extract historical information from two key types of data sources. The first type is tourism-oriented or universal knowledge bases (e.g., DBpedia). They store many POIs and related information. The POI search is primarily based on coordinates, similar to popular location-based systems. The other type is historical publications (as a rule, in HTML—HyperText Markup Language or in PDF—Portable Document Format) or archival databases records (e.g., in XML—eXtensible Markup Language). XML-files can be mapped into OWL [6]. HTML sources can be processed by means of NLP (Natural Language Processing) tools. In treating data sources of the second type, the main difficulty is that historical objects are usually identified by their names only.

As a result, the smart space contains a representation of the semantic network G, integrating the information extracted from multiple data sources. Their parallel activity is coordinated by combining KPs. The common ontology O is used to represent G in the smart space. The ontology provides a system of classes, relations, and restrictions that collected historical information must confirm. As a result, they constitute a historical semantic network for the reference recommendation scenario. A combining KP reasons over the extracted historical information and establishes semantic relations between historical objects. There can be several combining KPs, and the consistency of G is ensured by O. A combining KP can represent interests of a given tourist, act on behalf of a group of tourists, or perform generic context-aware construction.

Based on the whole semantic network G in the smart space, a ranking KP constructs recommendations. Each recommendation is represented as star graph R_p for a given tourist, initial POI p, and context. Visualization on the client KP can utilize additional information such as ranks r_q and annotations t_q for all recommended POIs from R_p. Importantly that there can be many ranking KPs, each employs own computational method of POI selection for the recommendation. For instance, in POI selection method [6], values of r_q reflect the closeness of q's categories to the categories the initial POI p has. Then an annotation t_q can describe the common categories of p and q.

The proposed architecture reduces the service construction to the interactions of KPs. It follows the principle that a smart space service is knowledge reasoning over the shared content and delivering the result to the users [27]. In our case of historical recommendation services for e-Tourism, the proposed model of KPs interaction is presented in Figure 4.

Smart space content is shared, forming a subject to self-generation. That is, the steps in Figure 4 are performed simultaneously, with event-driven synchronization. An important event to activate data extraction is specifying the initial POI of a tourist (step 1). Content self-generation also supports the service personalization. New historical objects can be found and new semantic relations can be associated with this given POI (the iteration in steps 2a and 2b). Request for additional information about historical objects (step 2b) occur until semantic network around a tourist is being formed. The POI recommendation and ranking are further personalized (the iteration in steps 4a and 4b) when personal information is directly integrated into the rank computation, e.g., the POIs that a tourist has already visited. Updating the user's personal information or preferences (step 4a) requires the rank conversion. This conversion can occur until the user likes the recommended POIs (step 4b).

Table II illustrates this content self-generation model by showing the construction of the semantic network from Figure 1 and, consequently, the star graph from Figure 2. Note

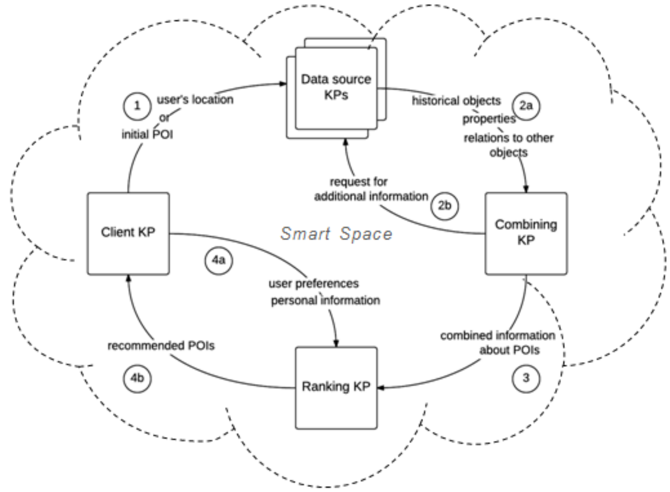

Figure 4. Many KPs interact in a smart space to construct a service.

TABLE II. EXAMPLE OF CONTENT GENERATION.

Step	Interaction	Generated content
1	Initial POI	Hotel Negresco
2	Data source KPs retrieve facts about the initial POI from different data sources.	Hotel Negresco is located at Promenade des Anglais, 37. The architect of Hotel Negresco is E.-J. Niermans. The first owner of Hotel Negresco is H. Negresco.
3	Combining KP collects an information about other POIs which relate with initial POI to create a semantic network.	POIs on the same street: Apartment house, Palais de la Méditerranée. POIs by the same architect: Villa le Colombier, Apartment house. POIs related with H. Negresco: Place Masséna.
4	Ranking KPs differentiate the POIs in the semantic network, selecting the most attracting for the user.	Rank-sorted list: Apartment house, Place Masséna, Palais de la Méditerranée, Villa le Colombier.

that in this paper we do not focus on a particular ranking criterion. Intuitively, the closer the POI and the richer and more appropriate the established relations then the rank becomes higher. For instance, Apartment house receives the highest rank since the POI is a) located on the same street and b) designed by the same architect.

VI. KNOWLEDGE PROCESSORS

The proposed system architecture defines the service-oriented application as a system of KPs interacting in the smart space. We classify these KPs as follows: A) Historical data source KP, B) Semantic network combiner, and C) POI Ranking KP. This section provides our design details of such KPs with an overview of possible algorithms that such KPs can implement.

A. Historical data source processors

Each data source KP provides the smart space with cultural heritage data. As shown in the previous section, different sources — universal knowledge bases (e.g., DBpedia), HTML-pages, databases records etc. can contain historical information about POIs. Consider design of a data source KP for the case of DBpedia knowledge base.

Dbpedia is created by means of automatic transmission of structured information from the Wikipedia pages to a semantic network [28]. Structured information in DBpedia consists of categories related with a Wikipedia page, references to other pages as well as to external resources, and facts presented in the so-called “infoboxes”. They are the fragments of the pages, which provide data structured according to a certain pattern. DBpedia is a source that is conceptualized with a fixed ontology. DBpedia uses RDF for representation of the information, which is automatically extracted from Wikipedia. The scheme of data source KP operation with the DBpedia knowledge base is shown in Figure 5.

TABLE III. DBPEDIA INFORMATION ABOUT HOTEL NEGRESCO.

Property	Value
geo:lat	43.694443
geo:long	7.257500
foaf:homepage	http://www.hotel-negresco-nice.com/
is dbo:knownFor of	dbr:EdouardNiermans(architect)
rdfs:comment	The Hotel Negresco is a famous hotel and site of the equally famous restaurant Le Chantecler, located on the Promenade des Anglais on the Baie des Anges in Nice, France. It was named after Henri Negresco

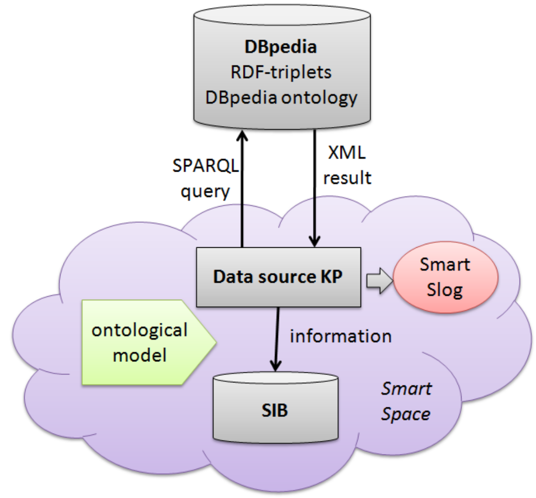

Figure 5. Data source KP operation with DBpedia.

Information representation in the format of RDF triples supports effective searching through the knowledge base using SPARQL queries. The data source KP is oriented to the Dbpedia ontology. The KP implements SPARQL queries to DBpedia, which also provides a certain set of query templates. Application development tool SmartSlog [26] is used to convert the SPARQL query XML result into a local data structure. The KP publishes the information in the smart space using the application ontological model [29]. This ontological model encompasses several important aspects of the tourism problem domain, including tourist places (cities, towns), tourist attractions, tourist events (concerts, shows, etc.), and transport issues. As a result, the KP provides POIs and other objects to construct the semantic network in the form of an RDF graph.

An example of information acquisition about Hotel Negresco from DBpedia is shown in Table III. Experimental performance evaluation of the data source KP operation is presented further in Section VII.

B. Semantic network combiner

Each combining KP is responsible for processing the information extracted from different sources. It aggregates the information received from the data source KPs. A combining KP constructs a personalized structure of POIs according to a given ontological model. The fragment of the ontological model is shown in Figure 6.

The ontology describes the three main entities: POI, Person, Event, and relations between them. Class Location represents POI location in the form of geographical coordinates. Geographical coordinates correspond WC3 Geo with data properties *lat* and *long* to designate geographical latitude and longitude, respectively. Class User is linked to class POI with two object properties *hasVisitPoint* and *hasFavoritePoint*. Thus, the user can mark a POI that she/he has visited and that she/he wants to visit.

If the received information is insufficient for semantic network extension, then the combining KP can send specifying queries to fill information structure of a POI. It is provided with a certain set of query templates that can be put to information sources. Thus, the combining KP establishes semantic relations for the semantic network between POIs and other historical objects. The constructed semantic network is ensured by ontological model; it also takes into account users' preferences and generic context.

In summary, each combining KP publishes information about POIs and the relations between POIs and other historical objects in the smart space. Then the published information is used to determine rank of each POI.

C. POI ranking methods

Ranking KP can use various algorithms to calculate the rank of a POI. In this section some methods are listed that can be used for future research. The proposed system design potentially adopts various ranking methods. Let us consider three classes of ranking algorithms. They show the spectrum of possible solutions, while the implementation and its analysis is the subject of our further research.

The simplest case is when basic information about POI is used, e.g., categories (a particular case of keywords or tags). Computational method in [6] solves the problem of ranking the available POIs by the level of proximity to the tourist's context based on the information about the set of the categories relating to a POI. Different knowledge bases (e.g., DBpedia, Freebase, and YAGO) are found on rich systems of categories. In different knowledge bases, there will be systematic differences in the values of distances between the sets of categories, thus the method used probabilistic approach determining distances between the subsets of a finite set.

Besides the possibility to compare the distances between the sets of categories, the probabilistic distance provides one more advantage: it provides an opportunity to evaluate the value of distance in the terms "long"/"short" before the whole set of the distances has been obtained. In the mobile environments, this extra opportunity is very useful, since it reduces the amount of computation by means of ranking POIs in accordance with the preferences and interests of a user.

The second class of algorithms comes from the recommender systems. Such systems are typically divided into the following three types [22]:

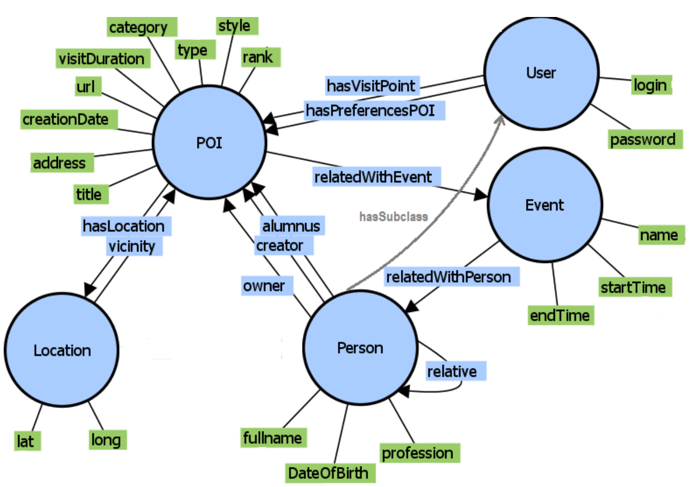

Figure 6. Example fragment of the ontological model.

1) content-based, which are based on similarity of POI's characteristics;
2) collaborative filtering, which are based on similarity of user preferences;
3) hybrid, which are combining 1 and 2.

In particular, a context-aware recommender system from [22] ranks a POI in terms depending on similarity to user preferences. It considers the following context attributes: time, company in which a tourist has visited an attraction, and weather.

The tourist can evaluate the attractiveness of POIs from one to five. Attraction rating estimation for a given tourist is performed in two steps: a group of tourists with ratings similar to the given tourist is determined; rating of attraction is estimated based on ratings of this attraction assigned by the tourists of the group. A recommendation is based on assessment of the similarity between the two tourists depending on the tourist evaluations. Then it calculates the group which the tourist can be classified to. A tourist group is determined by k-Nearest Neighbors method. It composes the resulting list of POIs.

Estimation of rank r_{uq} for POI $q \in P$ and tourist u is based on ratings (scores) of that POI. They are assigned by other tourists of the group with respect to their similarity to the tourist u. The resultant list of attractions L presented to the tourist u is sorted in descending order of:

$$s_q = kr^*_{uq} + (1-k)(1 - \frac{d^w_q}{\max_{i \in L} d^w_q})$$

where $k \in [0, 1]$ is a model parameter reflecting the importance of the POI rating estimation in favor of its reachability; d^w_q is the estimation of time for u to reach q.

Since the semantic network provides rich structural information, we can use the methods of structural ranking. For instance, a variant of the well-known PageRank algorithm can be applied. The basic idea behind PageRank for any directed graph is that a link from a node to another states an endorsement of the latter node, indicating the quality. PageRank takes advantage of the global link structure to order nodes according to their perceived quality. Various algorithms for computing PageRank in general graphs are presented in [30]. Consider the basic idea.

Given a semantic network G_p for $p \in P$. Ranks r_q for all $q \in P$ can be computed iteratively starting from some initial values $r^{(0)}_q$:

$$r^{(i+1)}_q = \alpha \sum_{\forall s: q \to s} p_{sq} r^{(i)}_s + (1-\alpha)\pi_q,$$

where $q \to s$ is a link in G_p, p_{qs} is the weight of the link, α is the damping factor denoting the probability of following the link structure, and π is a personalization vector of damping factors for all POIs.

VII. Comparison and Applicability Analysis

One of the key performance bottlenecks of the proposed smart space based architecture is the need to search and retrieve information from multiple data sources, which are available in the Internet. We experimentally analyze the performance of information acquisition from DBpedia. We assume

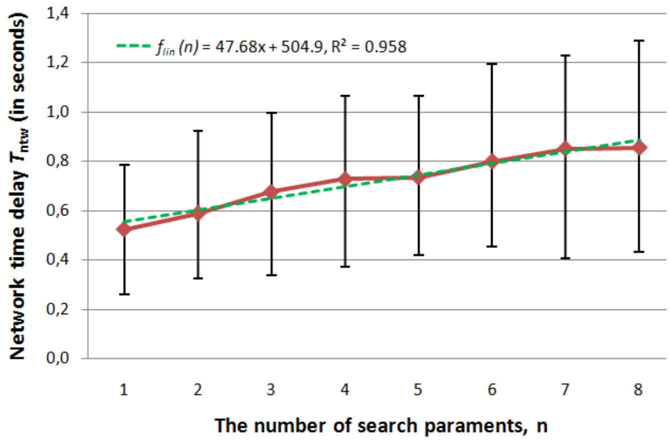

Figure 7. Estimated average time $T_{ntw}(n)$ of sending a query in DBpedia and getting a result.

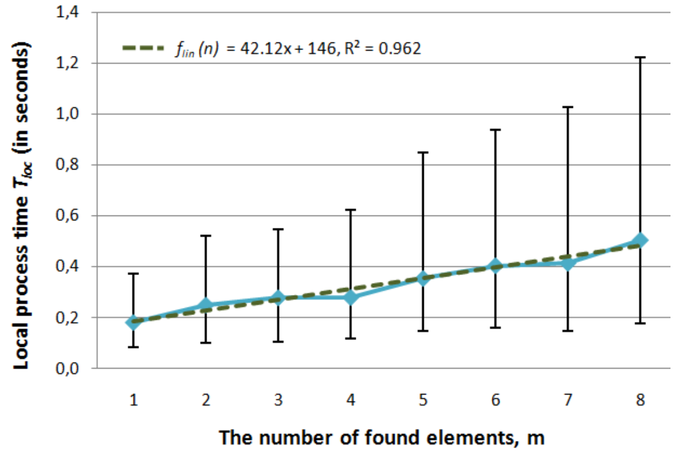

Figure 8. Estimated average time $T_{loc}(m)$ of processing the XML-result and publishing the information in smart space.

that the source has enough information on the studied history-aware problem. The information completeness in open data sources is a subject of our further research.

For this experiment we cannot use Hotel Negresco and its surroundings since it has a modest representation in DBpedia. Let us consider the case of Eiffel Tower to evaluate the performance of the data source KP. DBpedia information about Eiffel Tower is shown in Table IV. The information is much richer than about the Hotel Negresco (see Table III above). There are at least 8 parameters to search with the known coordinates of Eiffel Tower. As a result of the search, relations with 3 persons can be found (Stephen Sauvestre, Emile Nouguier, and Maurice Koechlin) and one historical event (Exposition Universelle in 1989). They define the first relations in the semantic network. Consider data source processing performance under specified conditions.

To retrieve information from DBpedia the data source KP sends a SPARQL query, processes the XML-result, and publishes the found information in the smart space in accordance with the application ontological model. In the experiments, we varied from one to eight query parameters ($n = 8$). The resource-expensive operations for our service are sending a query in DBpedia and receiving its result by network communication (time delay $T_{\rm ntw}$). Then the XML result is processed and the information is published in the smart space (time $T_{\rm loc}$).

The experimental behavior of $T_{\rm ntw}(n)$ and $T_{\rm loc}(m)$ is

TABLE IV. DBPEDIA INFORMATION ABOUT EIFFEL TOWER.

Property	*Value*
geo:lat	48.858223
geo:long	2.294500
dbo:architect	dbr:StephenSauvestre
foaf:homepage	http://www.tour-eiffel.fr
dbo:openingDate	1889-03-31
dbo:thumbnail	wikipedia-en:Special:FilePath/TourEiffelWikimediaCommons.jpg
dbp:height	300
dbp:structuralEngineer	dbr:EmileNouguier dbr:MauriceKoechlin
rdfs:comment	The Eiffel Tower is an iron lattice tower located on the Champ de Mars in Paris, France. It was named after the engineer Alexandre Gustave Eiffel, whose company designed and built the tower.
is dbp:building of	dbr:ExpositionUniverselle (1889)

shown in Figure 7 and Figure 8, respectively. The number of parameters is varied for $1 \leq n \leq 8$. The number of found elements is varied for $1 \leq m \leq 8$. Although the higher number of published RDF triples can be, we consider only ones which are necessary for the semantic network construction. For each n and m we made 50 runs, and the plot shows the average with the standard deviation bars. The presented plots also include the linear regression to show the trend.

We can see that DBpedia lets to get some important information about POI, including, in particular, the coordinates and some relations with other historical objects. Nevertheless, this information is not sufficient to build a rich semantic network around the POI. It is necessary to develop other Data source KPs, which would be able to analyze the information from different types of sources (e.g., HTML pages).

The case of HTML-page analysis is considered in [8]. It describes the extraction of historical facts as RDF-triples from the regional history database "Latgales Dati" (http://latgalesdati.du.lv) by means of PHP Simple HTML DOM Parser. An example of this activity is shown in Table V.

Now let us consider distinctive properties of our system design based on results published in previous works.

The proposed service scenario differs from other proposals. The most close to our work is [18]. The authors explore the use of location aware mobile devices for searching for and browsing a large number of general and cultural heritage information repositories. The application—Mobile Cultural Heritage Guide— searches for POIs in the current tourist's physical location and constructs a "mental map" of nearby POIs within a circular shape. Next semantic crawling is applied

TABLE V. HTML TO RDF TRANSFORMATION.

HTML	*RDF triples in Smart-M3 format*
<div class='page-header'> <h1>Edids Udems</h1> </div>	123, fullName, "Edids Udems"
<div class='person'>	
<div class="birth">Dzimis 1877. gada 15. maijā </div>	123, Birthday, "1877. gada 15 maijā" (15 May 1877)
<div class="info"> Uzņēmējs (tirgotājs). ...</div>	123, profession, "Uzņēmējs (tirgotājs)" (Businessman (trader))

to resemble the process of a human using the Web to find other information relevant to these POIs. Finally, augmented reality is used in combination with facet selection to present this POI-related information to an active tourist on her/his mobile device. Similarly to our scenario, this application aims at dynamic provision of semantically-enriched information in favor of a classical travel guide. In contrast, our scenario introduces both nearby and faraway POIs, which are semantically related within a variety of historical objects, including common historical persons and events. Our scenario supports automation of semantic crawling with POI ranking; the ranks are then used for visual representation of POI recommendations to the user.

Paper [31] considers a prototype application with POI ranking. It supports content-based recommendations for generating personalized routes along cultural heritage assets in outdoor environments (e.g., city tours). The case of indoor environments, such as museums, is studied in recent work [32]. The mobile application helps the visitors to access information concerning exhibits that are of primary interest to them during pre-visit planning, to provide the visitors with relevant information during the visit, and to follow up with post visit memories and reflections. In contrast, our scenario resembles the process of a historian studying historical facts.

Based on recent studies of the Smart-M3 performance, we expect the proposed system design is applicable in real-life setting and has advantages over the other approaches to recommender system development. The applicability of the smart space based architecture, similar to the one shown in Figure 3, is discussed in [10]. A realistic case is a smart space [22] where the number of large data sources, such as web-based databases and repositories, is of the 10-order magnitude and the number of mobile users is of the 10^3-order magnitude.

In contrast to mobile standalone applications, the workload is delegated from a personal mobile device to smart space infrastructure deployed on powerful hosts. Experiments in [33] confirm that this design solution additionally improves reliability and fault tolerance, essentially in wireless network settings. In our scenario, the delegated workload includes the construction of semantic networks for many users and POI ranking over these networks. A personal mobile device visualizes aggregated fragments (e.g., a star graph) of the whole semantic network enriched with derived ranks.

In comparison with web-based recommender system, the proposed smart space based system design provides flexibility in selection of 1) data sources, 2) semantic network construction, 3) POI ranking, and 4) personalization. Although the cost is performance, Semantic Web technologies are now capable to create and maintain relatively large RDF triple stores [25], where the number of RDF triples is of the 10^5-order magnitude and more. In particular, Smart-M3 SIB employs the Redland library for RDF triple store and SPARQL support [11], [24].

VIII. Conclusion

This paper addressed recommendation services development for historical e-Tourism. We studied the problem of historical POI recommendation, the necessity of using semantic relations between historical objects, and the personalized (subjective, contextual) aspect of services. We proposed the smart space based system design for implementing such a recommender system. The proposal provides a concept definition and design solutions for creating a smart space to accompany a historian tourist. Multiple external sources of historical data can be attached to the smart space and used for provision of information relevant to the situation and user's interests. The information is integrated using Semantic Web technologies and analyzed to produce personalized recommendations. The result is visually presented with quantitative (POI ranks) and qualitative (reason annotations) estimates.

Our study makes a step towards concept development for historical e-Tourism. Feasibility study of the proposed concept model, including ontology engineering for integrated representation of historical objects, analysis of POI ranking methods over a semantic network of historical objects, and comprehensive experimental evaluation, is a subject to our further research.

Acknowledgment

This work is financially supported by the Ministry of Education and Science of the Russian Federation within project # 2.2336.2014/K from the project part of state research assignment and project # 14.574.21.0060 (RFMEFI57414X0060) of Federal Target Program "Research and development on priority directions of scientific-technological complex of Russia for 2014–2020". The reported study was partially funded by RFBR according to the research project # 14-07-00252.

References

[1] A. G. Varfolomeyev, A. Ivanovs, D. G. Korzun, and O. B. Petrina, "Smart spaces approach to development of recommendation services for historical e-tourism," in Proc. 9th Int'l Conf. on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM). IARIA XPS Press, Jul. 2015, pp. 56–61.

[2] D. Gavalas, C. Konstantopoulos, K. Mastakas, and G. Pantziou, "Mobile recommender systems in tourism," J. Netw. Comput. Appl., vol. 39, Mar. 2014, pp. 319–333.

[3] J. Borrás, A. Moreno, and A. Valls, "Intelligent tourism recommender systems: A survey," Expert Syst. Appl., vol. 41, no. 16, Nov. 2014, pp. 7370–7389.

[4] V. L. Smith, Hosts and Guests: The Anthropology of Tourism. University of Pennsylvania Press, 1989.

[5] N. Ide and D. Woolner, "Historical ontologies," in Words and Intelligence II: Essays in Honor of Yorick Wilks, ser. Text, Speech and Language Technology, K. Ahmad, C. Brewster, and M. Stevenson, Eds. Springer, 2007, vol. 36, pp. 137–152.

[6] A. Varfolomeyev, D. Korzun, A. Ivanovs, and O. Petrina, "Smart personal assistant for historical tourism," in Recent Advances in Environmental Sciences and Financial Development. Proc. 2nd Int'l Conf. on Environment, Energy, Ecosystems and Development (EEEAD 2014), C. Arapatsakos, M. Razeghi, and V. Gekas, Eds., Nov. 2014, pp. 9–15.

[7] A. Varfolomeyev, D. Korzun, A. Ivanovs, H. Soms, and A. Pashkov, "Smart space services as a teaching aid in history and cultural heritage studies," in ICERI2015 Proceedings. 8th International Conference of Education, Research and Innovation. IATED, Nov. 2015, pp. 1373–1380.

[8] A. Varfolomeyev, D. Korzun, A. Ivanovs, H. Soms, and O. Petrina, "Smart space based recommendation service for historical tourism," in ICTE in Regional Development 2015 Valmiera, Latvia, ser. Procedia Computer Science, E. Ginters and M. Schumann, Eds. Elsevier, 2015, vol. 77, pp. 85–91.

[9] K. Kulakov, O. Petrina, D. Korzun, and A. Varfolomeev, "Towards an understanding of smart service: The case study for cultural heritage e-tourism," in Proc. 18th Conf. Open Innovations Framework Program FRUCT. ITMO Univeristy, Apr. 2016, pp. 145–152.

[10] D. Korzun, S. Balandin, and A. Gurtov, "Deployment of Smart Spaces in Internet of Things: Overview of the design challenges," in Proc. 13th Int'l Conf. Next Generation Wired/Wireless Networking and 6th Conf. on Internet of Things and Smart Spaces (NEW2AN/ruSMART 2013), LNCS 8121, S. Balandin, S. Andreev, and Y. Koucheryavy, Eds. Springer, Aug. 2013, pp. 48–59.

[11] J. Kiljander, A. D'Elia, F. Morandi, P. Hyttinen, J. Takalo-Mattila, A. Ylisaukko-oja, J.-P. Soininen, and T. S. Cinotti, "Semantic interoperability architecture for pervasive computing and Internet of Things," IEEE Access, vol. 2, Aug. 2014, pp. 856–874.

[12] P. Nora, "Between memory and history: Les lieux de mémoire," Representations, no. 26, 1989, pp. 7–24, Special Issue: Memory and Counter-Memory.

[13] M. Kalus, "Semantic networks and historical knowledge management: Introducing new methods of computer-based research," The Journal of the Association for History and Computing, vol. 10, Dec. 2007.

[14] A. Ivanovs and A. Varfolomeyev, "Computer technologies in local history studies: Towards a new model of region research," Acta Humanitarica Universitatis Saulensis, vol. 19, 2014, pp. 97–107.

[15] A. Meroño-Peñuela, A. Ashkpour, M. van Erp, K. Mandemakers, L. Breure, A. Scharnhorst, S. Schlobach, and F. van Harmelen, "Semantic technologies for historical research: A survey," Semantic Web journal, vol. 6, no. 6, 2015, pp. 539–564.

[16] E. Ahonen and E. Hyvonen, "Publishing historical texts on the semantic web - a case study," in Proc. 2009 IEEE Int'l Conf. on Semantic Computing (ICSC '09). IEEE Computer Society, 2009, pp. 167–173.

[17] L. Ardissono, T. Kuflik, and D. Petrelli, "Personalization in cultural heritage: The road travelled and the one ahead," User Modeling and User-Adapted Interaction, vol. 22, no. 1-2, Apr. 2012, pp. 73–99.

[18] C. van Aart, B. Wielinga, and W. R. van Hage, "Mobile cultural heritage guide: Location-aware semantic search," in Proc. 17th Int'l Conf. on Knowledge Engineering and Management by the Masses (EKAW'10), LNCS 6317, P. Cimiano and H. S. Pinto, Eds. Berlin, Heidelberg: Springer, 2010, pp. 257–271.

[19] F. Zhao, Z. Sun, and H. Jin, "Topic-centric and semantic-aware retrieval system for internet of things," Information Fusion, vol. 23, 2015, pp. 33–42.

[20] A. Becheru, C. Badica, and M. Antonie, "Towards social data analytics for smart tourism: A network science perspective," in Linguistic Linked Open Data. Springer International Publishing, 2016, pp. 35–48.

[21] J. Augusto, V. Callaghan, D. Cook, A. Kameas, and I. Satoh, "Intelligent environments: a manifesto," Human-centric Computing and Information Sciences, vol. 3, no. 1, 2013.

[22] A. Smirnov, A. Kashevnik, A. Ponomarev, N. Teslya, M. Shchekotov, and S. Balandin, "Smart space-based tourist recommendation system," in Proc. 14th Int'l Conf. Next Generation Wired/Wireless Networking and 7th Conf. on Internet of Things and Smart Spaces (NEW2AN/ruSMART 2014), LNCS 8638, S. Balandin, S. Andreev, and Y. Koucheryavy, Eds. Springer, Aug. 2014, pp. 40–51.

[23] K. Kulakov and A. Shabaev, "An approach to creation of smart space-based trip planning service," in Proc. 16th Conf. of Open Innovations Association FRUCT. ITMO Univeristy, Oct. 2014, pp. 38–44.

[24] I. Galov, A. Lomov, and D. Korzun, "Design of semantic information broker for localized computing environments in the Internet of Things," in Proc. 17th Conf. of Open Innovations Association FRUCT. ITMO Univeristy, IEEE, Apr. 2015, pp. 36–43.

[25] C. Gutierrez, C. A. Hurtado, A. O. Mendelzon, and J. Pérez, "Foundations of semantic web databases," J. Comput. Syst. Sci., vol. 77, no. 3, May 2011, pp. 520–541.

[26] D. G. Korzun, A. A. Lomov, P. I. Vanag, J. Honkola, and S. I. Balandin, "Multilingual ontology library generator for Smart-M3 information sharing platform," International Journal on Advances in Intelligent Systems, vol. 4, no. 3&4, 2011, pp. 68–81.

[27] D. Korzun and S. Balandin, "A peer-to-peer model for virtualization and knowledge sharing in smart spaces," in Proc. 8th Int'l Conf. on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM 2014). IARIA XPS Press, Aug. 2014, pp. 87–92.

[28] C. Bizer, J. Lehmann, G. Kobilarov, S. Auer, C. Becker, R. Cyganiak, and S. Hellmann, "DBpedia – a crystallization point for the Web of Data," Web Semantics: Science, Services and Agents on the World Wide Web, vol. 7, 2009, pp. 154–165.

[29] K. Kulakov and O. Petrina, "Ontological model for storage historical and trip planning information in smart space," in Proc. 17th Conf. Open Innovations Framework Program FRUCT. ITMO Univeristy, Apr. 2015, pp. 96–103.

[30] A. D. Sarma, A. R. Molla, G. Pandurangan, and E. Upfal, Fast Distributed PageRank Computation, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2013.

[31] N. Stash, L. Veldpaus, P. D. Bra, and A. P. Rodeirs, "Creating personalized city tours using the CHIP prototype," in Late-Breaking Results, Project Papers and Workshop Proceedings of the 21st Conf. on User Modeling, Adaptation, and Personalization (UMAP 2013): Proc. Workshop on Personal Access to Cultural Heritage (PATCH 2013), ser. CEUR Workshop Proceedings, S. Berkovsky, E. Herder, P. Lops, and O. C. Santos, Eds., vol. 997, Jun. 2013, pp. 68–79.

[32] T. Kuflik, A. Wecker, J. Lanir, and O. Stock, "An integrative framework for extending the boundaries of the museum visit experience: linking the pre, during and post visit phases," Information Technology & Tourism, vol. 15, no. 1, 2015, pp. 17–47.

[33] I. Galov and D. Korzun, "Fault tolerance support of Smart-M3 application on the software infrastructure level," in Proc. 16th Conf. of Open Innovations Association FRUCT. ITMO Univeristy, Oct. 2014, pp. 16–23.