



# Enhancing Itinerary Recommendation with Linked Open Data

Alessandro Fogli, Alessandro Micarelli, and Giuseppe Sansonetti<sup>(✉)</sup>

Department of Engineering, Roma Tre University,  
Via della Vasca Navale 79, 00146 Rome, Italy  
ale.fogli@stud.uniroma3.it, {micarell,gsansone}@dia.uniroma3.it

**Abstract.** This paper proposes a recommender system that exploits linked open data (LOD) to perform a social context-aware cross-domain recommendation of personalized itineraries integrated with multimedia and textual content. To this aim, the recommendation engine considers the user profile, the context of use, and the features of the points of interest (POIs) extracted from LOD sources. We describe how to extract data and process it to perform hybrid filtering. All recommendations are based on the user's features extracted implicitly and explicitly. These features are used to apply content-based filtering and collaborative filtering, weighing results based on similar users' experience. The preliminary results of an experimental evaluation on a sample of 20 real users show the effectiveness of the proposed system not only in terms of perceived accuracy, but also in terms of novelty, non-obviousness, and serendipity.

**Keywords:** Recommender systems · Points of interest  
Linked open data

## 1 Introduction

Recommender systems (RSs) can play a fundamental role in improving the target user's experience [1,2]. The best-known companies oriented to the sale of items (i.e., products and/or services) have long realized that the personalization of the offer is a fundamental factor to remain competitive on the market. Consequently, great interest is given to the analysis and modeling of users' individual tastes, with the aim of suggesting new products of their possible interest. Nowadays, there are many types of recommender systems that differ based on the domain in which they are deployed or the strategies applied for their design. This article proposes to realize open recommender systems able to take advantage of linked open data to perform a social context-aware cross-domain recommendation of itineraries enriched with multimedia and textual content.

This paper is organized as follows. Section 2 presents some state-of-the-art systems sharing some aspects with our recommender. The architecture underlying the itinerary recommender is detailed in Sect. 3. Section 4 describes the performed experimental tests and the obtained findings. Finally, Sect. 5 reports our conclusions and plans for future work.

## 2 Related Work

In this section, some RSs somehow related to the proposed one are described. Among the itinerary RSs, D'Agostino *et al.* [3] propose a personalized system able to suggest the target user itineraries satisfying not only her preferences and needs, but also her physical and social context. The recommendation process takes into account several aspects: in addition to the popularity of points of interest (POIs) (deducted considering, for example, the number of check-ins on social networking services such as Foursquare<sup>1</sup>), it includes the user profile, the current context of use, and the user's network of social links. The basic idea behind the work of Yoon *et al.* [4] is that planning trips to unknown regions is a difficult task for novice travelers. This burden can be alleviated if the residents of the area offer assistance to them [5]. This system carries out a recommendation of social itineraries realized by learning multiple digital paths generated by users, such as the GPS trajectories of residents and travel experts.

Regarding LOD-based RSs, the system proposed in [6] is a social recommender designed to analyze how data extracted from a user's activities on social networks can be enriched with the semantic knowledge provided by LOD [7]. In particular, such a RS is applied to the artistic and cultural heritage. The work by Heitmann and Hayes [8] demonstrates that LOD can be used to lower the barriers to access the information necessary for a RS. By reducing the data acquisition problem, LOD can indeed be used to fill the gaps of a collaborative filtering algorithm, in particular to mitigate the cold-start problem that occurs when it is needed to provide relevant recommendations for new users and new items [9].

## 3 Itinerary Recommender

In order to plan a personalized itinerary, the system exploits the interaction with the active user and provides her with targeted recommendations that increase her satisfaction with the results returned by the recommender. User data comes from explicit and implicit feedbacks. The first category includes responses collected through questionnaires submitted to the user. The second category includes implicitly collected data during the user's interaction with the application, for example, when choosing a specific route rather than another, or when making a new check-in. This data extends and refines the knowledge base of the user's preferences. Each user is characterized by a vector of weights whose values, between 0 and 1, linearly correspond to the user's interest in a certain category of venues.

Another crucial factor in the recommendation process is the current physical context [10, 11]. Almost all this information can be determined without the user's involvement. In fact, the position is detected by the GPS sensor of the mobile

---

<sup>1</sup> <https://foursquare.com/>.

device, just as the means of transport is detected by its accelerometer. Furthermore, the weather conditions are obtained by submitting queries to different meteorological services, based on the user’s current location.

Each single point of interest that composes the final itinerary is created by extracting the data available in the LinkedGeoData project dataset<sup>2</sup>. To retrieve this data, it is necessary to construct a query in the SPARQL language. Query construction takes place dynamically based on the characteristics of the target user and the information extracted from the context. Subsequently, the system searches for the POIs with the greatest number of links inherent to the user’s peculiarities in order to extract only the most similar venues and refine the recommendation. The extracted POIs are, then, filtered by means of a hybrid filtering process that takes into account both the collaborative and the content-based aspects. Representing users through a vector of numerical values makes it possible to cluster and classify them according to their preferences. This allows the system to evaluate the POIs that constitute the itineraries also based on the choices made by other users similar or connected to the target user. The results are, furthermore, filtered according to the context information by removing, for example, outdoor places in case of rain or those closed at that time.

The purpose of this work is to recommend itineraries, namely, paths that maximize the active user’s satisfaction. For this reason, we modeled the whole problem as a directed graph  $G = (V, E)$ , where  $V$  is a set of vertices (or nodes) and  $E$  is a set of edges. Each node represents a venue, that is, one of the points of interest present in the LinkedGeoData triples. Each edge represents a direct connection between two nodes. Information about an edge includes the shortest path to move from one node to another and the travel time, taking into account the user’s means of transport. This information is obtained through the Google Maps API<sup>3</sup>: for each pair of nodes  $(e_i, e_j)$  the system queries APIs for the travel time from  $e_i$  to  $e_j$  and the travel time from  $e_j$  to  $e_i$ , thus creating the edge. Once all the edges have been inferred, a complete graph is obtained from the initial node to the final node. Then, a routing algorithm is executed on it. More specifically, starting from the itinerary consisting of the only points of departure and arrival, further POIs are added gradually until all the time available has been spent. The insertion process is not random, but occurs during the sorting of POIs based on various factors, such as popularity and distance.

The routing algorithm returns many itineraries from the initial node to the final node. To obtain the first  $k$  that maximize the active user’s satisfaction, a scoring function is used. Such a function takes into account: (i) the number of venues that constitute the path; (ii) the path popularity; (iii) the distance from the starting point to the destination point; (iv) the popularity of the itinerary between the user’s friends; (v) the diversity of the venues in the itinerary according to their categories of membership. Once the score has been obtained, the routes are sorted according to it.

---

<sup>2</sup> <http://linkedgeo.org>.

<sup>3</sup> <https://developers.google.com/maps/>.

Once the itinerary preferred by the user has been implicitly obtained, it can be exploited to realize cross-domain recommendations through the use of other LOD sources. Starting from such a route, SPARQL queries are submitted to infer the features related to the POIs that belong to it. Then, the system exploits the extracted features to search for items sharing the same features, through the use of different endpoints. Figure 1 shows an example of screenshot returned by the system, where the active user receives personalized recommendations concerning the itinerary to follow and the related contents (i.e., a book, a movie, and a music album) to enjoy for enhancing her experience.

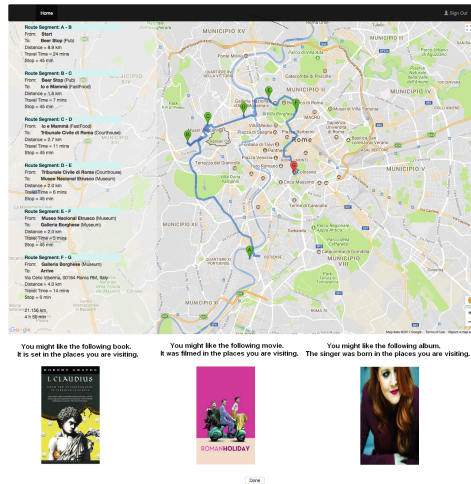


Fig. 1. Example of a system screenshot.

## 4 Experimental Evaluation

In this section, the results of preliminary tests are shown and discussed. The main objective of the performed tests was to evaluate the benefits of LOD in the implementation of recommender systems. In order to compare the advantages of LOD, it was necessary to use other data sources. Therefore, we decided to use the Foursquare API<sup>4</sup> to search for the venues inside the research area. Traditional evaluation approaches for recommender systems are based on offline testing. However, evaluating the perceived quality of an itinerary is an extremely subjective action. For this reason, we decided to use a sample of human testers in order to make the evaluations as impartial as possible. The direct involvement of users is aimed at establishing a qualitative estimation of the system based on user's perceptions. Testers were asked to try two different versions of the system, one enhanced with LOD and one with the Foursquare venues. Users were

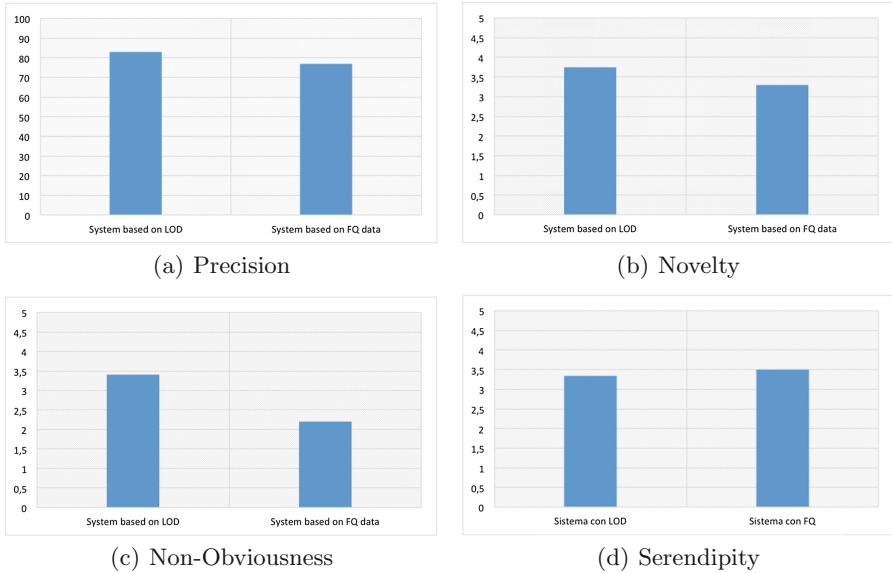
<sup>4</sup> <https://developer.foursquare.com/>.

also asked to create a route within a city that they knew at least in part, so as to be able to assess the proposed final itineraries with full knowledge of the facts. For this reason the system has been deployed in different cities such as Rome, Amsterdam, and San Francisco. The different experimental steps were as follows. Through the registration form the user selects the categories she would like to visit in a hypothetical itinerary. Once the user profile has been created, the system prompts the insertion of some explicit data. Then, the user chooses among the returned itineraries the one that maximizes her level of satisfaction. Once the itinerary has been decided, the user fills in an evaluation form in order to evaluate the quality of the route returned by expressing her agreement or disagreement with some questions. The (dis)agreement scale is represented as a Likert five-point scale.

The system evaluation involved a sample of 20 people with different characteristics (see Table 1). To evaluate the systems, the recommendations obtained using LOD were compared with the ones made using the Foursquare APIs. The first evaluation made on the results extracted from the forms filled in by participants concerned the *precision* value of the system. Precision measures the ability of the system to return a route containing highly relevant POIs. The next evaluation carried out concerned the level of *novelty* and *non-obviousness* of the recommended itineraries. This assessment arises from the fact that an itinerary, in order to maximize the active user’s satisfaction, has to not only include highly relevant POIs, but also achieve a significative level of novelty and a low level of obviousness in the routes returned. For assessing novelty and obviousness, we exploited data given by the user when filling in the evaluation form. The last test carried out concerned the *serendipity* metric, which measures how successful and surprising the recommendations are. The obtained experimental results are shown in Fig. 2. It can be noted that the results achieved by the system were encouraging for both versions tested. The graph in Fig. 2(a) shows that the system exploiting LOD has a slightly higher average precision

**Table 1.** Characteristics of testers.

	Item	Frequency	Percentage
Gender	Female	9	45%
	Male	11	55%
Age	18–30	13	65%
	31–50	4	20%
	51–70	3	15%
Profession	Student	10	50%
	Teacher	4	20%
	Employee	3	15%
	Freelancer	1	5%
	Unemployed	2	10%



**Fig. 2.** Comparison analysis between the system based on LOD and the system based on Foursquare (FQ) data.

value than the system exploiting Foursquare (FQ) data. As for novelty and non-obviousness (see Fig. 2(b) and (c)), the LOD-based system still achieved better results than its variant. As for serendipity, the results reported in Fig. 2(d) show that the LOD-based system attained encouraging levels of serendipity, albeit slightly lower than the FQ-based system.

## 5 Conclusions

In this paper, we have described a personalized recommender system of itineraries that exploits linked open data (LOD) and takes into account the physical and social context of the target user. The preliminary experimental results show that LOD is indeed useful for realizing such systems.

Although such findings are encouraging, the possible future developments of this research work are various and can concern each of the modules of the proposed recommender. Moreover, it is certainly necessary to perform further experimental tests with more users and scenarios. Finally, further research activities are necessary for integrating the user profiles with additional information about their personality [12,13], as well as the temporal dynamics [14,15] and the actual nature [16–18] of their interests. Browsing activities on the Web are also rich of relevant information that can be considered in the user modeling process [19,20].

## References

1. Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (eds.): *Recommender Systems Handbook*. Springer, Boston (2011). <https://doi.org/10.1007/978-0-387-85820-3>
2. Biancalana, C., Gasparetti, F., Micarelli, A., Sansonetti, G.: An approach to social recommendation for context-aware mobile services. *ACM Trans. Intell. Syst. Technol.* **4**(1), 10:1–10:31 (2013)
3. D’Agostino, D., Gasparetti, F., Micarelli, A., Sansonetti, G.: A social context-aware recommender of itineraries between relevant points of interest. In: Stephanidis, C. (ed.) *HCI 2016, Part II. CCIS*, vol. 618, pp. 354–359. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-40542-1\\_58](https://doi.org/10.1007/978-3-319-40542-1_58)
4. Yoon, H., Zheng, Y., Xie, X., Woo, W.: Social itinerary recommendation from user-generated digital trails. *Pers. Ubiquit. Comput.* **16**(5), 469–484 (2012)
5. D’Aniello, G., Gaeta, A., Gaeta, M., Loia, V., Reformat, M.Z.: Collective awareness in smart city with fuzzy cognitive maps and fuzzy sets. In: 2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1554–1561, July 2016
6. De Angelis, A., Gasparetti, F., Micarelli, A., Sansonetti, G.: A social cultural recommender based on linked open data. In: *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization. UMAP 20117*. ACM, New York, pp. 329–332 (2017)
7. Biancalana, C., Gasparetti, F., Micarelli, A., Sansonetti, G.: Social semantic query expansion. *ACM Trans. Intell. Syst. Technol.* **4**(4), 60:1–60:43 (2013)
8. Heitmann, B., Hayes, C.: Using linked data to build open, collaborative recommender systems. In: *Linked Data Meets Artificial Intelligence*, Papers from the 2010 AAAI Spring Symposium, Technical report SS-10-07, Stanford, California, USA, 22–24 March 2010, Palo Alto, California, USA, pp. 76–81. AAAI Press (2010)
9. D’Aniello, G., Gaeta, M., Loia, V., Orciuoli, F., Tomasiello, S.: A dialogue-based approach enhanced with situation awareness and reinforcement learning for ubiquitous access to linked data. In: *2014 International Conference on Intelligent Networking and Collaborative Systems*, pp. 249–256, September 2014
10. Biancalana, C., Flamini, A., Gasparetti, F., Micarelli, A., Millevolte, S., Sansonetti, G.: Enhancing traditional local search recommendations with context-awareness. In: Konstan, J.A., Conejo, R., Marzo, J.L., Oliver, N. (eds.) *UMAP 2011*. LNCS, vol. 6787, pp. 335–340. Springer, Heidelberg (2011). [https://doi.org/10.1007/978-3-642-22362-4\\_29](https://doi.org/10.1007/978-3-642-22362-4_29)
11. Biancalana, C., Gasparetti, F., Micarelli, A., Miola, A., Sansonetti, G.: Context-aware movie recommendation based on signal processing and machine learning. In: *Proceedings of the 2nd Challenge on Context-Aware Movie Recommendation. CAMRa 2011*, pp. 5–10. ACM, New York (2011)
12. Bologna, C., De Rosa, A.C., De Vivo, A., Gaeta, M., Sansonetti, G., Viserta, V.: Personality-based recommendation in e-commerce. In: *CEUR Workshop Proceedings*, vol. 997. CEUR-WS.org, Aachen (2013)
13. Onori, M., Micarelli, A., Sansonetti, G.: A comparative analysis of personality-based music recommender systems. In: *CEUR Workshop Proceedings*, vol. 1680, pp. 55–59. CEUR-WS.org, Aachen (2016)
14. Arru, G., Gurini, D.F., Gasparetti, F., Micarelli, A., Sansonetti, G.: Signal-based user recommendation on Twitter. In: *Proceedings of the 22nd International Conference on World Wide Web, WWW 2013 Companion*, pp. 941–944. ACM, New York (2013)

15. Caldarelli, S., Gurini, D.F., Micarelli, A., Sansonetti, G.: A signal-based approach to news recommendation. In: CEUR Workshop Proceedings, vol. 1618. CEUR-WS.org, Aachen (2016)
16. Gurini, D.F., Gasparetti, F., Micarelli, A., Sansonetti, G.: A sentiment-based approach to twitter user recommendation. In: CEUR Workshop Proceedings, vol. 1066. CEUR-WS.org, Aachen (2013)
17. Gurini, D.F., Gasparetti, F., Micarelli, A., Sansonetti, G.: iSCUR: interest and sentiment-based community detection for user recommendation on Twitter. In: Dimitrova, V., Kuflik, T., Chin, D., Ricci, F., Dolog, P., Houben, G.-J. (eds.) UMAP 2014. LNCS, vol. 8538, pp. 314–319. Springer, Cham (2014). [https://doi.org/10.1007/978-3-319-08786-3\\_27](https://doi.org/10.1007/978-3-319-08786-3_27)
18. Gurini, D.F., Gasparetti, F., Micarelli, A., Sansonetti, G.: Temporal people-to-people recommendation on social networks with sentiment-based matrix factorization. *Future Gener. Comput. Syst.* **78**, 430–439 (2018)
19. Gasparetti, F., Micarelli, A., Sansonetti, G.: Exploiting web browsing activities for user needs identification. In: 2014 International Conference on Computational Science and Computational Intelligence, vol. 2, pp. 86–89. IEEE Computer Society, Los Alamitos, March 2014
20. Bonifacio, A., Biancalana, C., Gasparetti, F., Micarelli, A., Sansonetti, G.: Implicit evaluation of user’s expertise in scientific domains. In: Stephanidis, C. (ed.) HCI 2017. CCIS, vol. 713, pp. 420–427. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-58750-9\\_58](https://doi.org/10.1007/978-3-319-58750-9_58)