Journal of Universal Computer Science, vol. 27, no. 7 (2021), 714-733
submitted: 15/5/2020, accepted: 26/5/2021, appeared: 28/7/2021 CC BY-ND 4.0

Towards a Semantic Graph-based Recommender System. A Case Study of Cultural Heritage

Sara Qassimi
(L2IS Laboratory, Cadi Ayyad University, Marrakesh, Morocco
https://orcid.org/0000-0002-9441-986X, sara.qassimi@uca.ac.ma)

El Hassan Abdelwahed
(LISI Laboratory, Cadi Ayyad University, Marrakesh, Morocco
https://orcid.org/0000-0002-2786-6707, abdelwahed@uca.ac.ma)

Abstract: Research on digital cultural heritage has raised the importance of providing visitors with relevant assistance before and during their visits. With the advent of the social web, the cultural heritage area is affected by the problem of information overload. Indeed, a large number of available resources have emerged coming from the social information systems (SocIS). Therefore, visitors are swamped with enormous choices in their visited cities. SocIS platforms use the features of collaborative tagging, named folksonomy, to commonly contribute to the management of the shared resources. However, collaborative tagging uses uncontrolled vocabulary which semantically weakens the description of resources, consequently decreases their classification, clustering, thereby their recommendation. Therefore, the shared resources have to be pertinently described to ameliorate their recommendations. In this paper, we aim to enhance the cultural heritage visits by suggesting semantically related places that are most likely to interest a visitor. Our proposed approach represents a semantic graph-based recommender system of cultural heritage places through two steps; (1) constructing an emergent semantic description that semantically augments the place and (2) effectively modeling the emerging graphs representing the semantic relatedness of similar cultural heritage places and their related tags. The experimental evaluation shows relevant results attesting the efficiency of the proposed approach.

Keywords: Emergent semantic, Semantic Graphs, Recommender System, Augmentation of Cultural Heritage Places
Categories: H.3.1, H.3.2, H.3.3, H.3.7, H.5.1
DOI: 10.3897/jucs.70330

1 Introduction

Digital cultural heritage is a mature field, in which novel information and communication technologies (ICT) are used in the service of preserving cultural heritage and supporting its discovery by the public [López-Nores et al. 2019]. Research on visitor support systems for museums and cultural sites has been particularly important in recent years. Indeed, the museum is not only a place for the conservation of cultural sites and masterpieces of works showcased in an exhibit but also an institution for the education and entertainment of visitors. The cultural heritage area is particularly affected by the problem of information overload. For example, the Tripadvisor website lists more than 800 points of cultural visits for the city of Marrakesh, Morocco. Consequently, visitors are swamped with the enormous choices in their visited cities. Exploiting a long list of options is confusing and ambiguous for visitors. Therefore, they either spend time sifting through the options that best match their interests or to choose randomly. Besides, visitors do not necessarily know what they should visit or what they could appreciate. They are usually limited to visit the most popular places as in most guided tours. As a result, they may miss cultural heritage sites or places that might have interested them. One of the main research areas related to the problematic of information overload is the field of information retrieval. The general principle is to develop methods and algorithms to respond to the user's requests. However, it is not always easy to know how to express the request especially within the large number of available resources coming from the social information systems (SocIS).

The SocIS are information systems based on social technologies and open collaboration [Tilly et al. 2017]. They have enabled social interaction giving rise to massive shared resources, any identifiable things (e.g. images, articles, videos, blogs, etc.). The SocIS platforms use the features of folksonomy to commonly contribute to the management of the shared resources. The collaborative tagging, named folksonomy, enables a collaborative classification of the shared resources. Folksonomy, unlike a taxonomy, includes terms without a hierarchy [Qassimi et al. 2016]. For example, Instagram provides collaborative tagging, sharing and categorizing photographs and short videos. Different aspects of folksonomy have been explored in information retrieval [Tilly et al. 2019], social network analysis [Feicheng and Yating 2014], decision support systems [Hafidi et al. 2021, Qassimi et al. 2018], community of practice [Hafidi et al. 2019], recommendation systems [Bocanegra et al. 2017] and others. The collaborative tagging has led to an increasing amount of users providing information not only about the shared resources but also about their interest. The resulted growing and rich corpus of social knowledge can be exploited by recommendation technologies [Godoy et al. 2016].

Regardless of its popularity and widespread success, folksonomy is a collection of tags coming from an unsupervised and uncontrolled vocabulary that contains polysemous and synonymous words [Qassimi et al. 2017]. The unsupervised nature of folksonomy's tags semantically weakens the description of resources, consequently hindering their indexing, decreases their classification and clustering. Indeed, a poor classification and clustering of resources decrease their retrieval and suggestion in recommender system (RS). Therefore, the shared resources have to be pertinently described to ameliorate their recommendations.

In this regards, this article represents a semantic graph-based recommender system by (1) semantically augmenting the cultural heritage place and (2) effectively modeling the emerging graphs representing the semantic relatedness of similar cultural heritage places and their related tags to perform graph-based recommendations. The recommender system is based on the emerging graphs representing the semantic relatedness of similar cultural heritage places and their related tags. The emergent graphs represent two related graphs; the graph of cultural heritage places described by the graph of emergent tags. The semantic graphs are a combination of the two graphs that form a multi-relational or multi-layer graph, where nodes represent entities and directed edges represent relationships. The resulting multi-relational graph is referred to as a Knowledge Graph (KG). In this paper, we define the KG as a semantic graph.

This study is a part of a wider range of research project aiming to embrace digital cultural heritage in order to attract the attention of a wider audience. Many research projects have been created to assist the visitors before and during their visit. Therefore, the objective is to promote and enhance the cultural heritage of a touristic city by suggesting its historical places that suit the visitors' interests.

This paper is an extended version of our conference paper that was presented in MEDI 2019 [Qassimi and Abdelwahed 2019 (b)]. It improves the modeling of the semantic enrichment of the graph-based recommender system. The aim is to augment the cultural heritage places to semantically enhance the recommendations. Besides, the paper explores the use of graph databases to deal with the increasing complexity of real-world data, thus gathering powerful data insights enhance recommender systems.

The rest of the paper is organized as follows: Section 2 provides the literature review divided into a background related to recommender systems followed by a discussion of limitations and challenges. Section 3 presents the semantic enhancement of the model of recommender system. Section 4 describes the proposed approach of the semantic graph-based recommender system to suggest the augmented cultural heritage places. The experimental results and evaluation are described in Section 5. Finally, the conclusion and future directions are delineated in Section 6.

2 Literature review

2.1 Background

The Recommender System (RS) is a very active research field both in industry and academia. Information retrieval (IR), machine learning (ML) and human-computer interaction (HCI) and many other fields and disciplines have contributed to recommender systems research. Their common starting concern is to personalize the recommendation to the user's profile and current situation which implies modeling the user's interests. A recommender system is a subclass of information filtering system, deals with this challenge by providing relevant information. Recommender systems assist users to cope with information overload [Konstan and Riedl 2012] and have become one of the most powerful and popular tools in electronic commerce [Ricci et al. 2010]. They are primarily used and implemented by big online applications. Some real-world examples include friends' suggestions on social applications such as Facebook, Twitter, and LinkedIn, suggestions of profiles on Instagram, suggestions for products on Amazon, video recommendations on YouTube, news recommendations on Google News, and so on. These successful implementations of recommendation engines have shown their potential in other areas. The recommender system is able to provide recommendations to users. The RS helps users with four key features; Decide: predicting a rating for a user concerning an item; Compare: rank a list of items in a personalized way for a user; Explore: give items similar to a given target item; Discover: provide a user with unknown items that will be appreciated.

2.1.1 Recommender System Approaches

The main commonly distinguished filtering algorithms of recommender systems are collaborative (CF), content-based (CB) and hybrid filtering recommendations approaches [Abbas et al. 2015]. CF and CB are the most popular approaches because of their wide acceptance and extensive usage in RSs field over others. CF relies on identifying preference patterns within a user community. In contrast, CB filtering approach only considers the past preferences of an individual user and builds a preference model using a feature-based representation of the content of recommendable items.

Content-based filtering (CB) approach recommends similar items to the ones that the target user has preferred in the past [Yera and Martínez 2017]. The CB- RS takes into account the data provided by the user both directly and indirectly. It extracts information from the items' features, rather than using users' interactions and past behaviors. The user profile is constructed based on his previously preferred items by filtering metadata describing those items such as the keywords, category, and other descriptive features. For example, if the user has a history of visiting art museums in Paris , then the CB-RS will recommend artistic and cultural museums once the visitor is in Marrakesh. However, the CB-RS does not take into account behavior/data about other users. Hence, if a particular art museum fetches a very low rating or tagging by other users, it will still be recommended to the user. The CB-RS does not take into consideration the user's neighborhood preferences. Hence, it only considers the user's past preferences and the properties/features of the items.

Collaborative filtering (CF) approach recommends the items to the target user on the basis of his past preferences, and filtering out other users with similar behaviors [Villegas et al. 2018, Sinha and Dhanalakshmi 2019]. For the CF recommendation method, the user profile is built by filtering information from the user's behaviors like ratings, assigned tags and comments or implicitly rating by liking the items. User preference is expressed in two categories. Explicit rating, the most direct feedback from users, is a rate given by a user to an item on a sliding scale or assigning tags. Implicit rating is a rate giving by users expressing their preference indirectly, such as page views, clicks, purchase records, likes, and so on. There are two categories of CF: User-based CF that measures the similarity between users' profiles such as the nearest-neighbor method; And Item-based CF that uses the target users' ratings to find similarity between items.

Despite the success of the two filtering techniques, several problems were identified such as the limited content analysis for content-based and the cold start problem in collaborative filtering [Gupta and Goel 2018]. In order to mitigate these problems, many hybrid filtering techniques were proposed by combining content-based, collaborative filtering and other filtering techniques [Isinkaye et al. 2015]. The hybrid-based recommender systems combine two or more filtering techniques. The classical 2-dimensional approaches (users x items) have been extended with the current researches in the recommender system. Hybrid RSs blend basic approaches to achieve some synergy among them [Abbas et al. 2015]; recommender systems harness context-awareness with the personalization to offer the most accurate recommendations. For example, CF and CB approaches can be used together to avoid the new-item problems of CF techniques. Knowledge recommender system [Geng et al. 2018] has emerged with a large amount of generated knowledge. It deals with knowledge overload by filtering the most relevant ones that match the user's preferences. The knowledge recommendation approach is applied to overcome the cold start problem and help users in decision making. Several other approaches were also found in the literature to improve RSs performance that include data mining techniques by extracting knowledge from data; context-aware techniques providing contextual information to enhance the pro-activity of recommendations without getting any explicit request from users; semantic-based techniques using semantic similarity to determine the semantic relations among users and between items presented in domain ontology [Kumar and Sharma 2016].

2.2 Discussion: Limitations and Challenges

The RS approaches come across some challenges that affect the precision of recommendations [Khalid et al. 2014], such as sparsity, cold start, and scalability issues. The sparsity issue occurs due to the insufficient required data to extract descriptive metadata, rating, and contextual information about the items. The CF-based recommender system come across this issue because of its filtering technique depends on the ratings of similar users. For example, the MovieLens data is represented by the user-item matrix that increases its dimension with the users' ratings. This matrix suffers from data sparseness when the majority of users do not rate a large number of items. The cold start problem occurs when a user or an item is new to the system which has insufficient ratings or records at the start. Most CF-based recommender systems lack in offering accurate recommendations because of the challenging cold start issue [Abbas et al. 2015]. It has been addressed by the CB filtering. A scalable system is capable of handling efficiently and effectively a huge volume of data. Indeed, the current recommender systems deal with the scalability issues that increase the time processing and reduce the accuracy of the recommendations [Khalid et al. 2014].

The recommender systems aim to predict unknown ratings of items for users. The method of matrix factorization has solved the task to complete the user-item matrix. The million-dollar Netflix challenge led to a matrix completion that formulates the recommendation concept [Jannach et al. 2016]. However, Netflix prise has overshadowed other challenging and alternative ways of building RS by covering the quality factors of RS. Even Netflix states 'there are much better ways to help people find videos to watch than focusing only on those with a high predicted star rating' [Gomez-Uribe and Hunt 2016]. In many domains, it is insufficient to predict how much a user will like an item. The recommender system's performances need to have a good overview of the interaction of users [Neidhardt et al. 2015]. The challenge is to help the user explore the items and improve interpreting their preferences by interpreting their conversational queries. The key insight is the use of attributed tags on items that will enhance the conversational systems to elicit user preferences. Therefore, tagging can lead users to express well-formed preferences. Then, offering recommendations with a visualization of the user's social graph.

Recommender systems are milestone services for several companies. It is easier and spontaneous to choose the suggested item. Most importantly, the choice is based far more often on requested items than on overall popular items. With the advent of SocIS, the use of social information can prominently enhance the descriptive of items and users' profiles. Besides, the act of tagging not only describes resources, items to be recommended but also informs about the preferences and interests of the creator of tags. Folksonomy provides a complex interaction between users, resources, and their descriptive tags. A set of users annotate a set of resources with a set of tags that might be previously entered in the system by themselves and by other users. It is an emergent collaborative categorization of resources in terms of tags shared by a community. This complex network will not be able to be modeled using the classical recommender system approaches in the literature that rely on techniques of content-based, collaborative filtering and hybrid filtering recommendations using the similarity matrix. For instance, using the similarity matrix (i.e., user-item matrix) for recommending movies is unable to model and merge the explicit taste of a user about its favorite genres with its ratings and tagging for specific movies. It has called the attention of using a flexible recommendation approach that would ease the exploitation of the multiple kinds of possible interactions between different entities (e.g., users, items, tags, genres and types of movies, ratings and so on). Graphs are mathematical structures enabling to encode these interactions. Recent work has shown the efficiency of using graph modeling to improve the effectiveness of recommender systems. Musto et al. [Musto et al. 2018] use Linked Open Data as an external knowledge graph implemented within a hybrid graph data model. A recommend system is designed to suggest jobs by leveraging a directed graph of multi-edge linked jobs, thus it has overcome the great challenges of sparsity and scalability [Shalaby et al. 2017]. Wang et al. [Wang et al. 2019] introduced a knowledge graph for recommendation, so called Knowledgeaware Path Recurrent Network (KPRN) that exploits reasoning on paths to infer the user-item interaction. Chaudhari et al. [Chaudhari et al. 2016] propose a privacy aware recommender system that uses the knowledge graph of NELL, which encodes entities exploits relations present between entities (content from user's history and candidate content) and their relations. Although these approaches use graphs-based recommendations to solve specific issues, our work differs significantly as we aim to enhance the recommendation's performance and diversity by capturing the semantic and social interaction through modelling folksonomy relationship into a semantic graph.

3 Semantically enhancing the model of Recommender System

The research study aims to semantically enhance the model of the recommender system (see Figure 1). The augmentation of the cultural heritage places is achieved by emerging semantic from shared resources that describe them. The emergent semantic of resources is conducted by exploring folksonomy aligned to ontology to extract semantic describing the resources. Besides, we aim to model the linking among these resources by using multi-layer graphs that represent the Folksonomy tripartite relationship of users annotating resources by using tags. The concept of a semantic graph-based recommender system is analogical to the knowledge graphs used to enhance the search engines used by giants companies like Amazon and Facebook that constructed their knowledge graphs to incorporate their large amounts of data [Qassimi et al. 2020].

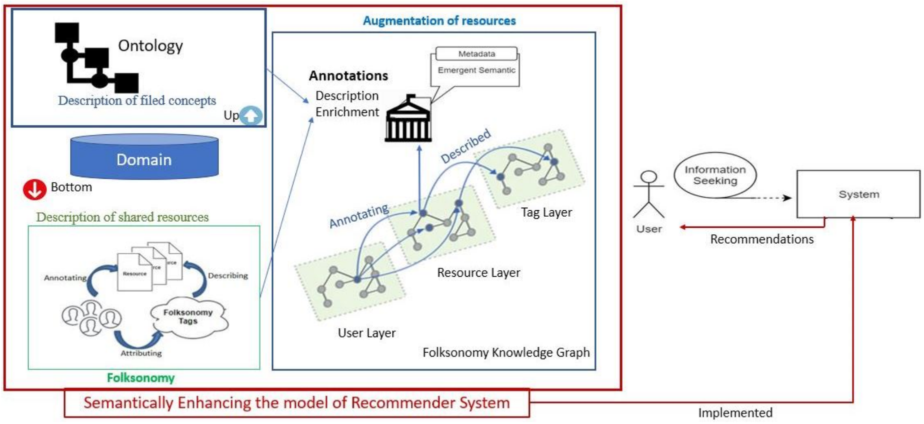

Figure 1: Emergent Semantic of resources towards a Semantic graph-based recommender system

The emergent semantic of resources is modeled using a conceptual model (see Figure 2). A provider (who can be an end-user a professional or a company) post or publish content/resource identified by a URL or URI (the content/resource can be digital, physical or abstract). In the case study, The resources describe cultural heritage places. The resource is described with descriptive metadata provided either by the approach of content-based (extracting main-keyword), controlled vocabulary based (extracting ontology's terms) or folksonomy-based (extracting relevant tags). This descriptive Metadada (main-keyword, ontology's terms and relevant tags) will create an emergent semantic that

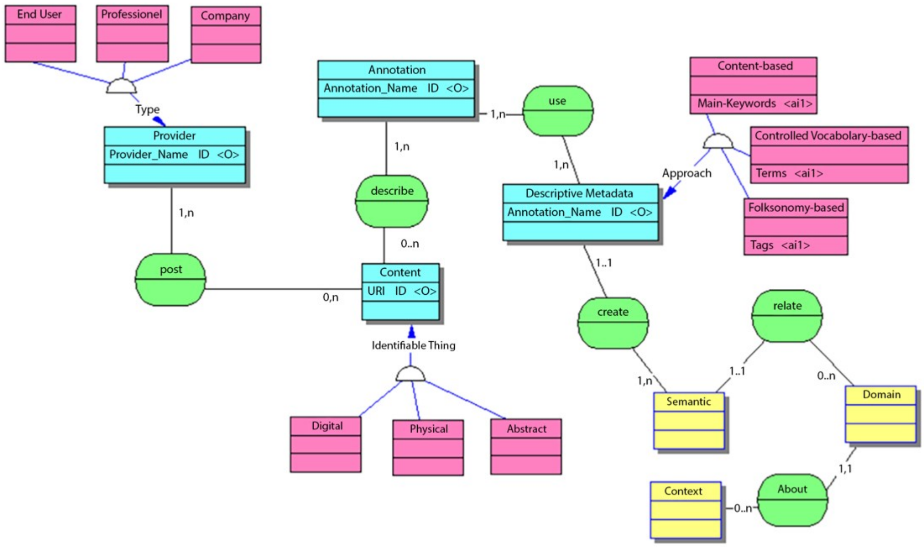

Figure 2: Emergent Semantic of Resources- Conceptual Data Model

relates to a specific domain of a certain context, such as cultural heritage in our case study.

4 Semantic graph-based recommender system of cultural heritage places

Our proposal aims to enhance the cultural heritage visits, i.e., to suggest semantically augmented and related museums of the world that are most likely to interest a visitor [Qassimi and Abdelwahed 2019 (b)]. The proposed approach (see Figure 3) presents the recommendation of cultural places based on graphs by harnessing the emergent semantic tags T={t_i} to augment the places P={p_i} to be recommended.

Step 1: Semantically augmenting the cultural heritage place

The first step mainly focuses on augmenting the cultural heritage places by exploring collaborative tagging to pertinently extract descriptive metadata, a.k.a. Emergent Semantic Tags.

Our previous works [Qassimi and Abdelwahed 2019 (a), Qassimi et al. 2017, Qassimi et al. 2021] present a deep explanation of the concept of the emergent semantic of resources through a combined semantic enrichment approach to augment the resource. The construction of the emergent semantic of resources is realized by extracting a different type of descriptive metadata, namely the relevant tags from the folksonomy, the matching terms from a domain ontology and the extracted content-based main keywords. the emergent semantic of resources augment the cultural heritage place by the pertinently extracted descriptive metadata using collaborative tagging and ontology. Folksonomy is known as collaborative tagging, social indexing or social classification. It refers to a set of keywords ‘tags‘ created by users to describe resources. Each user freely tags the resources using its own keywords, which aids him in categorizing information for

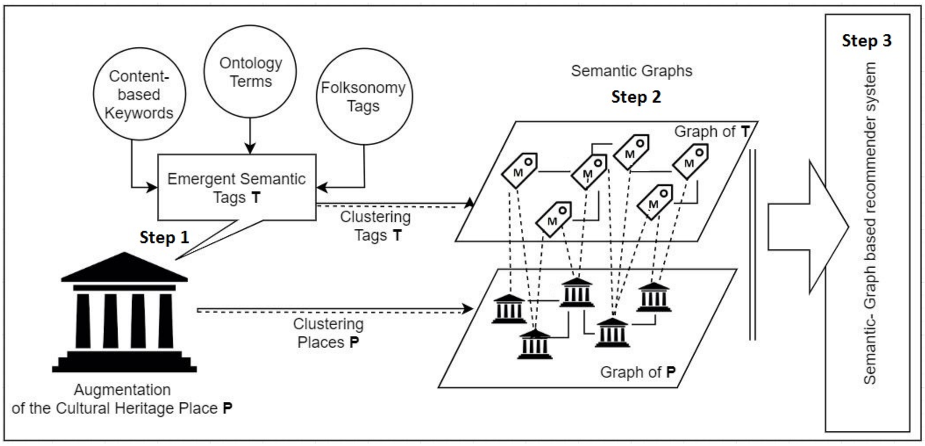

Figure 3: Semantic graph-based recommendation of cultural heritage places

his own personal management and also for sharing with others. For instance, the social annotation services Instagram is based on the exchange of tagged photos and videos. While ontology, the backbone of the semantic web, is a traditional taxonomy built with a controlled vocabulary and maintained by a limited number of experts. It is a representative knowledge structure of a field's descriptive that represents the relationship between mental objects 'concepts'. The emergent semantic Tags T is a combined semantic enrichment of resources performed by extracting different type of descriptive metadata: Relevant tags extracted from the folksonomy; Extracted content-based main keywords; Matching terms from a domain ontology (see Figure 4).

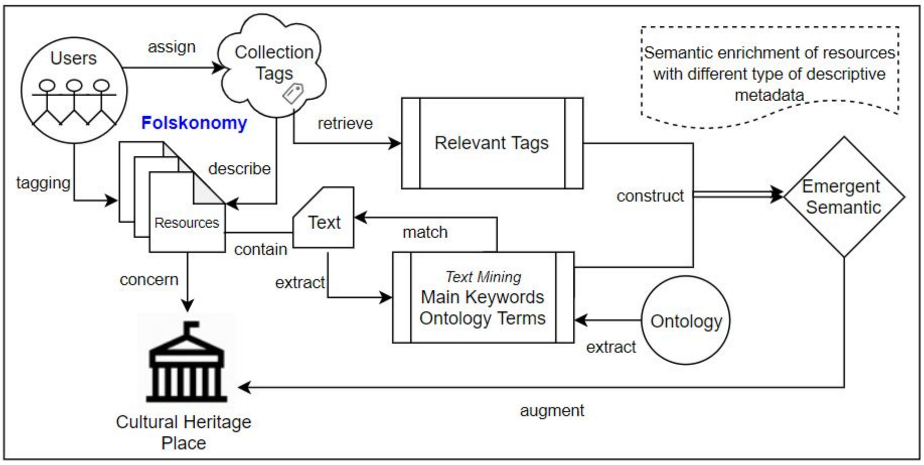

Figure 4: Semantic Augmentation of cultural heritage places

The main steps of the emergent semantic are as follows: Extracting content-based main keywords by generating and filtering candidates keywords. Extracting a set of ontology's terms matching the textual content of a resource. Considering folksonomy tags which are frequently used and understood by many users of a group. Exploring the relevant extracted tags from the folksonomy to enhance the model of the keyword extraction Maui [Maui] that learns the extraction strategy from the manually annotated corpus; i.e., the manually annotated corpus contains also relevant folksonomy's tags to additionally aliment the training data. In a folksonomy, the extraction of relevant folksonomy's tag describing a resource r_k requires to have a high degree of frequency. We compute for each folksonomy's tag t_i a degree of frequency, denoted by $DF(r_k,t_i)$.

$$DF(r_k, t_i) = \sqrt{FR(r_k, t_i)^2 + FU(r_k, t_i)^2} \tag{1}$$

Where: FR (r_k,t_i) is the Frequency of the folksonomy's tag t_i describing the Resource r_k ; FU (r_k,t_i) is the Frequency of Users who use the folksonomy's tag t_i to describe the resource r_k .

$$FR(r_k, t_i) = \frac{Number\ of\ times\ the\ tag\ t_i\ is\ used\ to\ describe\ the\ resource\ r_k}{Number\ of\ tags\ used\ to\ describe\ the\ resource\ r_k} \tag{2}$$

$$FU(r_k, t_i) = \frac{Number\ of\ users\ who\ use\ the\ tag\ t_i\ to\ describe\ the\ resource\ r_k}{Number\ of\ users\ who\ annotate\ the\ resource\ r_k} \tag{3}$$

The emergent semantic is a collection of different type of metadata, so called 'Emergent Semantic Tags T'. It is a shortcut description, summarizations, or abstractive summary of resources describing a particular place. Thus, it represents a semantic augmentation of the cultural heritage place.

$$Emergent\ Semantic\ Tags\ T = \begin{cases} Folksonomy\ Tags \\ Main\ Keywords \\ Ontology\ Terms \end{cases}$$

Step 2: Semantic-graph based Recommender System
Graph of cultural heritage places:
Let G_P=(V_P, E_P), with V_P vertices and E_P edges. G_P is a graph of cultural heritage places P, representing the vertices V_P or nodes as places connected with weighted edges $W(p_i,p_j)$ relating two places p_i and p_j. The two places are semantically similar when their weighted edge $W(p_i,p_j)$ is high.

$$\begin{aligned} W(p_i, p_j) &= \sqrt{W_{semantic}(p_i, p_j)^2 + W_{context}(p_i, p_j)^2} \\ W_{semantic}(p_i, p_j) &= \sqrt{\begin{array}{l} W_{folksonomy-tag}(p_i, p_j)^2 \\ +W_{main-keyword}(p_i, p_j)^2 \\ +W_{ontology-term}(p_i, p_j)^2 \end{array}} \end{aligned} \tag{4}$$

⚠ If the cultural heritage place is not described with textual descriptive, then:

$W_{main-keyword}(p_i, p_j) = 0$ and $W_{ontology-term}(p_i, p_j) = 0$

$$W_{folksonomy-tag}(p_i, p_j) = \frac{Number\ same\ folksonomy\ tags\ describing\ p_i\ \&\ p_j}{Total\ Number\ folksonomy\ tags}$$
$$W_{main-keyword}(p_i, p_j) = \frac{Number\ same\ main\ keywords\ describing\ p_i\ \&\ p_j}{Total\ Number\ main\ keywords}$$
$$W_{ontology-term}(p_i, p_j) = \frac{Number\ same\ ontology\ terms\ describing\ p_i\ \&\ p_j}{Total\ Number\ ontology\ terms}$$

$$W_{context}(p_i, p_j) = \frac{\sum w_c(p_i, p_j)}{Total\ Number\ contextual\ features}$$
$$w_c(r_i, r_j) = \begin{cases} 1 \text{ if } p_i\ and\ p_j\ are\ contextually\ related \\ 0 \text{ else.} \end{cases} \tag{5}$$

The places are clustered based on their same descriptive tags constructing an emergent graph of cultural heritage places. We use the folksonomy relevant tags to semantically relate places, in case, the textual content is absent to perform the extraction of main keywords and matching ontology's term. Besides, the weighted edges $W_{context}(p_i, p_j)$ consider the contextual features (contextual information: spatial, temporal and static) that conjointly describe the two places to enhance their relatedness. The static context characterizes maintained attributes describing the places that are unchanged over time, like the category of a cultural place. The graph of places assembles semantically related places to formalize their similarity. Therefore, the recommender system will explore the place-place relatedness among the graphs of cultural heritage places.

Graph of emergent semantic tags:
Let G_T=(V_T, E_T) where vertices V_T are tags and E_T represents the edges relating the tags. G_T is a semantic graph of tags useful to further find semantic relationships among places annotated with connected tags. Tow tags t_i and t_i are related with weighted edge W(t_i, t_j).

$$W_{place}(t_i, t_j) = \frac{Number\ of\ places\ described\ by\ both\ tags\ t_i\ and\ t_j}{Total\ Number\ of\ tagged\ places} \tag{6}$$

It enables graph-based reasoning about the relationships between tags attributed to describe different places.

Step 3: Recommendation Algorithm

The recommender system will explore the knowledge graphs (KGs) to extract and recommend semantically related places.

Algorithm 1 Graph-based recommendations

$t \in$ T; $t_{p_i} \in T_{p_i}$; p_j,$p_i \in$ P; $Recom_t$,$Recom_p \subset$ P
$Recom_p$: a set of recommended places based on the graph of places
$Recom_t$: a set of recommended places based on the graph of tags
k: a finite number of recommendations.
T_{p_i}: a finite number of tags describing the place p_i.
procedure RecommendationGraphPlaces(p_i)
  for $p_j \in P$ **do**
    if W(p_i, p_j) > 0 **then**
      $Recom_p$ $\leftarrow$list of k-ranked p_j
    end if
  end for
  return $Recom_p$
end procedure
procedure RecommendationGraphTags(p_i,t_{p_i})
  for $t \in T$ **do**
    if W(t_{p_i}, t) > 0 **then**
      $Recom_t$ $\leftarrow$list of k-ranked places annotated with t
    end if
  end for
  return $Recom_t$
end procedure

The graph-based recommendations algorithm suggests cultural heritage places based on the emergent KGs. The graph of cultural heritage places G_p illustrates the intensity of similarity among the related places. Once the recommender system selects the place previously visited, tagged or rated by a user, similar places will be recommended. The similarity of cultural heritage places is computed with the weight W(p_i,p_j). For a place p, the k-highest weights of places related to p are suggested. The graph of tags G_T allows to easily retrieve knowledge about relationships between tags describing different places. Hence, the graph of tags is useful to describe the relationship between places described with related tags. Consequently, the recommender system will explore this tags-knowledge graph to further extract and recommend related places having similar descriptive tags.

5 Evaluation and Results

Recommendation of cultural heritage places

We aim to enhance the cultural heritage visits of a touristic city, like Marrakesh City of Morocco. The semantic graph-based RS suggests cultural heritage places that are most likely to interest visitors of Marrakesh. We collected Wikipedia textual description of each of the list of World Heritage Sites in the Arab States [UNESCO World Heritage List] for training and 21 most cultural places of Marrakesh city for testing. We perform the augmentation of each place by constructing its Emergent Semantic Tags T using the three types of medatada (a.k.a annotations) considering: Relevant tags of the folksonomy TagsFinder [TagsFinder] which are related hashtags used by Instagram and Twitters users; Ontology terms extracted from UNESCO Thesaurus [UNESCO SKOS]; Main keywords extracted from the Wikipedia descriptive content. We surveyed a keyword extraction tool using the Multi-purpose automatic topic indexing Maui [Maui]. Two ensemble machine learning classifiers were used to train Maui: Maui based on the bagging decision trees classifier (Maui Bag); And Maui based on the boosting classifier called AdaBoostM1 using classification trees as single classifiers (Maui Boost). The standard information retrieval measures (Precision P, Recall R, F-Measure F) are used to compare their performance. P is the percentage of correct annotations among those extracted; R is the percentage of correctly extracted annotations among all correct ones; F is the combination of both P and R (see Table 1). The highest measures' values for precision, recall, and F-measure are achieved using Maui Bag. The accuracy of extracting main keywords and ontology terms is improved by training Maui Bag model that learns the annotation extraction strategy based on bagging decision trees classifier.

Extraction of Annotations	**Maui Bag**			**Maui Boost**		
Performance Metrics	**P**	**R**	**F**	**P**	**R**	**F**
Main Keywords	45.14	28.46	34.91	35.78	23.13	28.09
Ontology terms	29.39	28.9	29.14	23.68	22.06	22.84

Table 1: Comparing performances of Main keyword and ontology terms extracting tools

The semantic graph-based recommendations recommend k-top places with similar tags and categories. We compared the performance of our proposed recommender system with the content-based recommender system (CB-RS) [CB-RS source] that computes similarity between all cultural places. The CB-RS algorithm generates a TF-IDF matrix, then computes the cosine similarity between any pair of cultural places based on their descriptions using SciKit Learn's linear_kernel (python library). For example, we compare the recommendations of places related to the ElBadi Palace suggested by our proposed approach and the classic CB-RS (see Figure 5). The CB recommendations of places provides irrelevant recommendations compared to our proposed RS. For example, CB-RS suggests Menara and Agdal Gardens which are not in the same category as ElBadi palace. Besides, it does not recommend three cultural places (Almoradive Koubba, Seven Saints Tombs and Perfume Museum) which have tags and categories in common with ElBadi Palace.

The evaluation of the cultural places recommendation is performed using the automatic or offline evaluation that considers an item as relevant if it has common categories with the target item (the ground truth or true positive Tp, where the number of categories in common >1 and the cosine similarity score >0.015)). Each of the recommending algorithm suggests eleven places, the results are evaluated using the following accuracy metrics (7): Precision P, recall R and F1-measure F, that are calculated from the number of cultural places that are either relevant or not and either recommended or not. Four possible outcomes are shown in the confusion matrix (see Table 2).

$$P = \frac{Tp}{Tp + Fp} \;; R = \frac{Tp}{Tp + Fn} \;; F = \frac{2 \times P \times R}{P + R} \quad (7)$$

The recommendations concerns five cultural heritage places, namely, ElBadi Palace, Bahia Palace, Almoravid Koubba, Marrakech Museum, and Medersa Ben Youssed (see table 3). For each place, the recommendation algorithms of both the proposed RS 'semantic

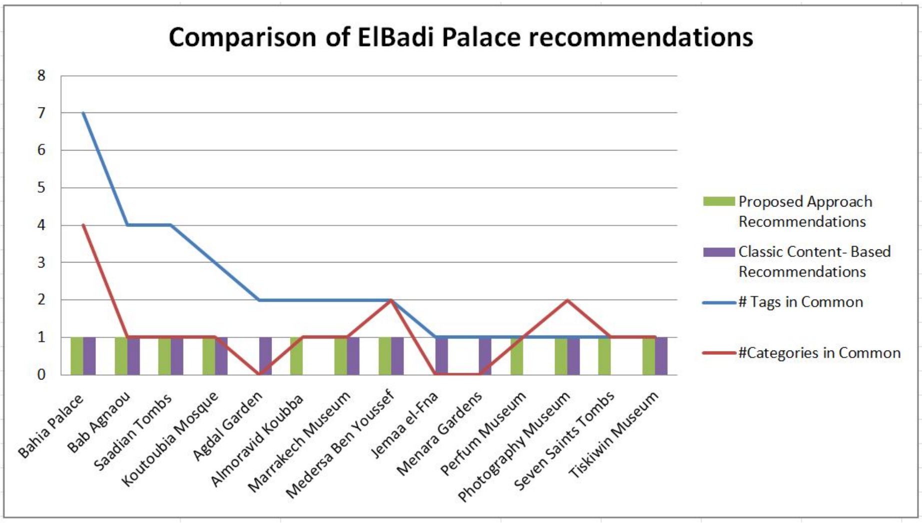

Figure 5: Comparison of ElBadi Palace recommendations

	Relevant	**Irrelevant**
Recommended	True-positive (Tp)	False-positive (Fp)
Not Recommended	False-negative (Fn)	True-negative (Tn)

Table 2: Confusion matrix

graph-based RS‘ and the CB-RS were performed to suggest eleven places. The evaluation of the recommendations compare the proposed approach of semantic graph-based RS with the CB-RS by using the results of F1-measure F which combines precision P and recall R. Thus, our proposed algorithm achieves highest results compared to CB-RS (see Figure 6).

Recommendation of TripAdvisor museums based on Graph Database

The high-performance computing power of graph databases can help reduce the size of large data applications. The growing interest in the use of graph databases has come with the increasing complexity of real-world data and the growing need for graph querying [Arsintescu et al. 2019]. In addition to extracting high-dimensional characteristics, graphs also make it possible to perform a variety of high dimensional analysis (e.g. pattern mining, clustering, etc.) that allows to gather powerful data insights enhancing recommender systems. Recommendations can use graphs to create meaningful clusters of items reducing the dimensionality of the recommendation problem. In the following, we present how graph-oriented database technology leverages the building of flexible recommendation engines. A graph database is characterized by its distinct data model compared to traditional relational databases [Angles and Gutierrez 2008]. In contrast to relational databases that compute data relationships through expensive, high-cost and complex join queries, where join-intensive query performance deteriorates as the size of dataset increases. While a graph database performance tends to remain relatively constant. The graph data model enables to store, process and query connections between data efficiently.

Cultural Heritage Places	**Proposed RS**			**CB-RS**		
Performance Metrics	**P**	**R**	**F**	**P**	**R**	**F**
ElBadi Palace	0.27	1	0.42	0.25	1	0.4
Bahia Palace	0.66	0.66	0.66	0.63	0.63	0.63
Almoravid Koubba	0.72	0.88	0.8	0.6	0.75	0.66
Marrakech Museum	0.63	1	0.77	0.63	1	0.77
Medersa Ben Youssef	0.9	0.9	0.9	0.81	0.81	0.81

Table 3: Evaluation of cultural heritage places recommendations

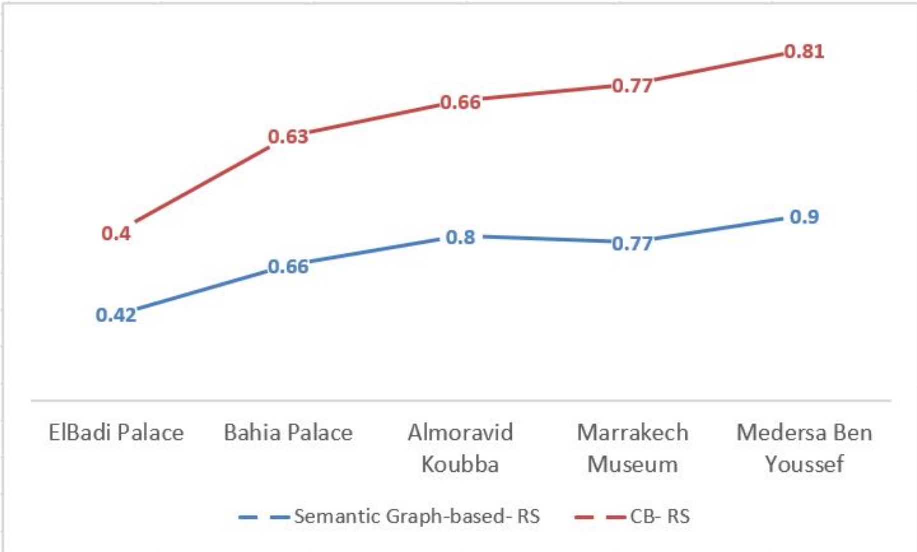

Figure 6: F1-Measure Comparison of our proposed approach with the CB-RS

To conduct cluster analysis and perform the museums recommendation, we used the collection and processed data of 1600 museums scraped from TripAdvisor [Museum TripAdvisor]. To evaluate the graph-based recommendations of TripAdvisior museums, we perform queries with the widely used open-source graph database Neo4j which is characterized as an 'embedded, disk-based, and fully transactional graph database engine' [Park et al. 2014]. The Graph Model in the Neo4j database has the following components: Nodes (equivalent to vertices in graph theory) are the main data elements that are interconnected through relationships. A node can have one or more labels (that describe its role) and properties (i.e. attributes). Labels are used to group nodes, and each node can be assigned by multiple labels. Labels are indexed to speed up finding nodes in a graph. Properties are attributes of both nodes and relationships. Neo4j allows for storing data as key-value pairs, which means properties can have any value (string, number, or boolean). For example, the Cypher query enables to recommend to the target user, having phone number '+1 412-622-3131', museums similar to the previously rated museum. The recommended museums were never rated by the target user (see Figure 7).

Recommendation of 10 other museums semantically similar to the museum 'Carnegie

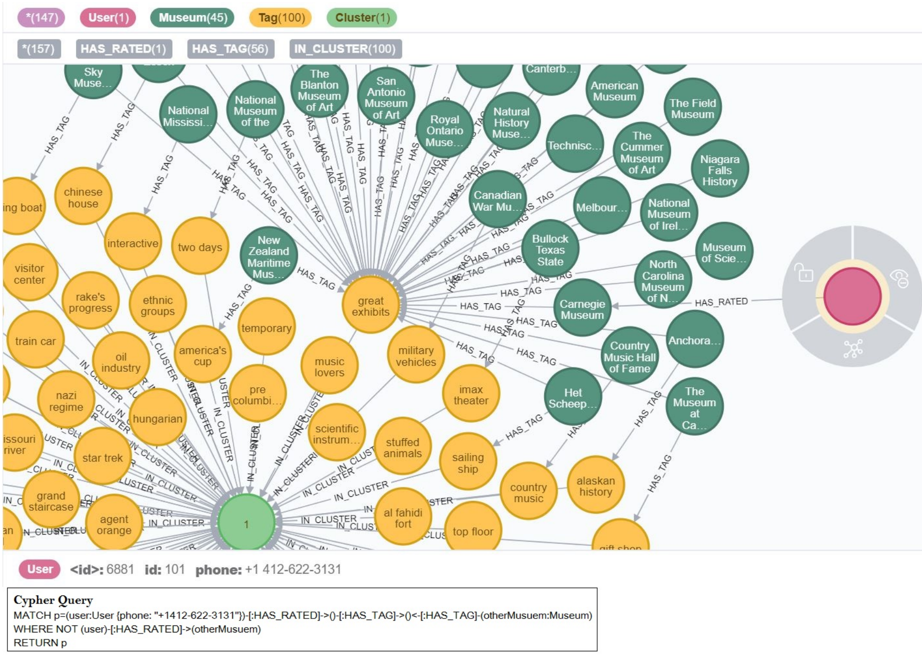

Figure 7: Recommending never rated museums for the target user

Museum of Natural History‘ based on the emergent graph of museums (see Table 4).

The offline evaluation considers the previously rated museums (rating >3) as the ground truth, true positive Tp, interpreted relevant to the user. For each user, we select its previously rated museum to generate recommendations based on graph based recommender system. The experimental evaluation (See Figure 8 and Figure 9) presents the results of the precision and recall metrics evaluating the recommendations performed to randomly selected 10 active users who has previously rated museums. The accuracy metrics evaluate whether the proposed recommender system based on graph database can properly predict the relevant museums that were previously well-rated by the users. The high precision and recall results of the ‘RS based on Graph Database‘ indicate that almost all the recommendations are indeed relevant to the users (Figure 8). We compared the RS based on graph database with the hybrid-based recommender system ‘Hybrid-RS‘ that uses both filtering approaches CB and CF. The algorithm of hybrid-RS recommends museums with similar content to those rated by the 10 active users. Its recommendation process is based also on the similarity of museums with users profiles using collaborative filtering (CF). The model of the Hybrid-RS is trained using the Singular Value Decomposition (SVD) model to extract features and correlation from the user-item matrix. It evaluates its model based on the root mean square error (RMSE) and Mean absolute error (MAE), resulting mean RMSE= 0.2746 and Mean MAE= 0.1702 (using cross validation (n-folds=2); Fold 1: RMSE= 0.2924 , MAE= 0.1700 and Fold 2: RMSE= 0.2569, MAE= 0.1704). The precision-recall curve shows that the precision and the recall of the RS based on graph database are higher than the hybrid based RS (Figure 9).

Recommended Museums	# tags in Common	Tags
Natural History Museum of Utah	8	[great exhibits, gift shop, beautiful building, on display, dinosaur bones, great for kids, all ages, hands on activities]
Natural History Museum of Los Angeles	7	[beautiful building, t rex, on display, mineral collection, great for kids, all ages, gems and minerals]
Smithsonian National Museum of Natural History	7	[great exhibits, t rex, on display, great for kids, all ages, mummies, gems and minerals]
Anchorage Museum at Rasmuson Center	6	[great exhibits, gift shop, on display, great for kids, all ages, hands on activities]
National Mississippi River Museum & Aquarium	6	[great exhibits, gift shop, on display, great for kids, all ages, hands on activities]
High Desert Museum	5	[great exhibits, gift shop, on display, great for kids, all ages]
Illinois State Museum	5	[great exhibits, gift shop, on display, spend an afternoon, all ages]
International Spy Museum	5	[gift shop, on display, great for kids, all ages, hands on activities]
Museum of Geology	5	[gift shop, on display, mineral collection,dinosaur bones, gems and minerals]
Museum of Natural Sciences	5	[t rex, on display, great for kids, all ages, hands on activities]

Table 4: Museum Recommendations

6 Conclusion and perspectives

The need for embracing digital culture has increased in the cultural mediation to attract the attention of a wider audience. However, visitors are swamped with enormous choices in their visited cities. With the advent of social information systems (SocIS), visitors are confronted with several challenges due to the information overload. Choosing the options that best match their interests is time spending and usually let the visitor choose randomly. Even though, SocIS platforms use the features of collaborative tagging, named folksonomy, to commonly organize the shared resources. Folksonomy lacks semantic

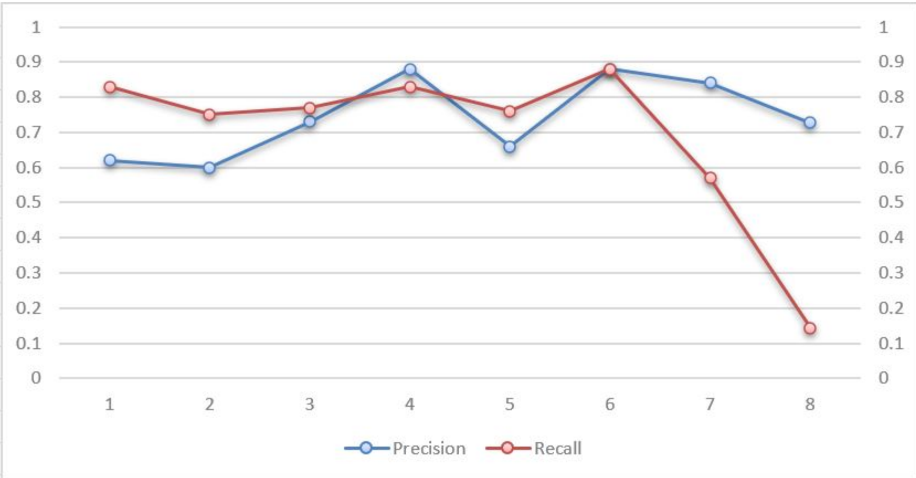

Figure 8: Evaluation of the recommendation of TripAdvisor Museums

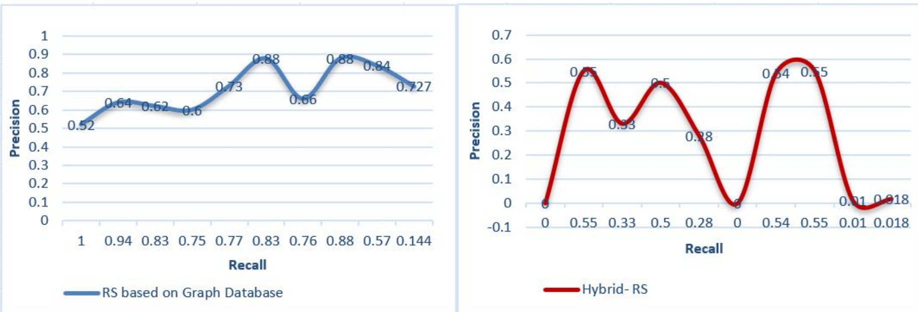

Figure 9: Precision-Recall Curve: Comparison of the RS based on graph database with the Hybrid-RS

which reduces its effectiveness of organizing resources. It decreases their findability and discoverability, thereby their recommendation. In this paper, our research study aims to enhance the visitor support system by recommending cultural sites that are most likely to interest visitors. Our proposed approach represents a semantic graph-based recommender system of cultural heritage places by (1) constructing an emergent semantic description that semantically augments the place and (2) effectively modeling the emerging graphs representing the semantic relatedness of similar cultural heritage places and their related tags. The experimental evaluation shows relevant results attesting the efficiency of our proposal applied to recommend cultural heritage places of Marrakesh city. Future perspectives will focus on creating a real-world application enabling the recommendation of personalized content to visitors of cultural heritage sites in a context-aware manner. The real-world mobile application will integrate the augmented reality (AR) to directly deliver recommended information (i.e., cultural heritage sites, videos describing a place to visit, its most popular tags, and so on). It will include a semantic graph-based context-aware recommender system (CARS) that rises in value the cultural heritage of the touristic city by suggesting historical places that suit the visitor's interests. The CARS will filter similar users by analyzing their similar profiles (e.g., ethnicity, age, country of provenance, previous tags describing the visited historical places, etc.) to suggest fitting historical places with their descriptions. In future works, we aim to investigate the potential of using the Linked Open Data LOD to increase semantics relatedness of resources describing places in the emergent semantic graphs to obtain a specific knowledge graph representative of the cultural heritage domain. Future directions will also investigate the network analysis, Linked Open Data LOD, considering scalability and performance of graph data processing algorithms, supporting complex analytics operating in real-time processing. Other perspectives aim to perform online evaluations as a primary mean to assess the accuracy and reliability of recommendations.

References

[López-Nores et al. 2019] López-Nores, M., Kuflik, T., Wallace, M. et al.: "Preface to the UMUAI special issue on personalized delivery of cultural heritage content: perspectives on 7 years of progress in the field"; User Model User-Adap Inter 29, 1–7 (2019).10.1007/s11257-019-09230-x

[Tilly et al. 2017] Tilly, R., Posegga, O., Fischbach, K., Schoder, D.: "Towards a Conceptualization of Data and Information Quality in Social Information Systems"; Business and Information Systems Engineering, 59, 3-21.(2017)

[Tilly et al. 2019] Kumar, K. P., Srivastava, A. and Geethakumari, G.: "A psychometric analysis of information propagation in online social networks using latent trait theory"; Computing 98, 583-607 (2016).10.1007/s00607-015-0472-7

[Feicheng and Yating 2014] Feicheng, Ma. and Yating, Li.: "Utilising social network analysis to study the characteristics and functions of the co-occurrence network of online tags"; Online Information Review 38: 232-247 (2014).10.1108/OIR-11-2012-0124

[Qassimi et al. 2018] Qassimi, S., Abdelwahed, E.H., Hafidi, M., Lamrani, R.: "A Graph-Based Model for Tag Recommendations in Clinical Decision Support System"; In: Abdelwahed E., Bellatreche L., Golfarelli M., Méry D., Ordonez C. (eds) Model and Data Engineering. MEDI 2018. Lecture Notes in Computer Science, vol 11163. Springer, Cham. (2018)

[Qassimi et al. 2019] Qassimi, S., Abdelwahed, E.H., Hafidi, M., Lamrani R.: "The Role of Recommender System of Tags in Clinical Decision Support"; In: Ezziyyani M. (eds) Advanced Intelligent Systems for Sustainable Development (AI2SD'2018). AI2SD 2018. Advances in Intelligent Systems and Computing, vol 914. Springer, Cham.(2019)

[Hafidi et al. 2021] Hafidi, M., Abdelwahed, E.H., Qassimi, S.: "Graph□based tag recommendations using clusters of patients in clinical decision support system"; Concurrency Computat Pract Exper. 33:e5624.(2021)

[Hafidi et al. 2019] Hafidi, M., Abdelwahed, E.H., Qassimi, S., Lamrani R.: "Clustering and Social Recommendation Applied in Health Community of Practice"; In: Ezziyyani M. (eds) Advanced Intelligent Systems for Sustainable Development (AI2SD'2018). AI2SD 2018. Advances in Intelligent Systems and Computing, vol 914. Springer, Cham. (2019)

[Bocanegra et al. 2017] Bocanegra, C.L., Ramos, J.L., Rizo, C., Civit, A., Fernández-Luque, L.: "HealthRecSys: A semantic content-based recommender system to complement health videos"; BMC Medical Informatics and Decision Making, 17.(2017).

[Godoy et al. 2016] Godoy, D. and Corbellini, A.: "Folksonomy-Based Recommender Systems: A State-of-the-Art Review"; Int. J. Intell. Syst. 31 314-346 (2016).10.1002/int.21753

[Qassimi et al. 2017] Qassimi, S., Abdelwahed, E.H., Hafidi, M. , Lamrani, R.: "Towards an emergent semantic of web resources using collaborative tagging"; In: Ouhammou Y, Ivanovic M, Abelló A, Bellatreche L (eds), Model and data engineering. MEDI 2017. Lecture notes in computer science, vol 10563. Springer, Cham (2017).

[Konstan and Riedl 2012] Konstan, J.A., Riedl, J.: "Recommender systems: from algorithms to user experience"; User Model User-Adapt Interact ;22:101–23 (2012).

[Isinkaye et al. 2015] Isinkaye, F.O., Folajimi, Y.O., Ojokoh, B.a.: "Recommendation systems: principles, methods and evaluation"; Egypt Inform J (2015).

[Yera and Martínez 2017] Yera, R., Martínez, L.: "Fuzzy Tools in Recommender Systems: A Survey"; International Journal of Computational Intelligence Systems 10 (1), pp. 776-803 (2017).

[Villegas et al. 2018] Villegas, NM., Sánchez, C., Díaz-Cely, J., Tamura, G.: "Characterizing context-aware recommender systems: a systematic literature review"; Knowl Based Syst 140:173–200 (2018).

[Angles and Gutierrez 2008] Angles , R. and Gutierrez, C.: "Survey of graph database models"; ACM Comput. Surv. 40, 1, (2008).10.1145/1322432.1322433

[Park et al. 2014] Park, Y., Shankar, M. , Park, B. and Ghosh, J.: "Graph databases for large-scale healthcare systems: A framework for efficient data management and data services"; 2014 IEEE 30th International Conference on Data Engineering Workshops, Chicago, IL, pp. 12-19 (2014).10.1109/ICDEW.2014.6818295

[Qassimi and Abdelwahed 2019 (a)] Qassimi, S., Abdelwahed, E.H.: "The role of collaborative tagging and ontologies in emerging semantic of web resources";Computing (2019).

[Qassimi et al. 2016] Qassimi,S., Abdelwahed, E.H., Hafidi, M., Lamrani,R.: "Enrichment of ontology by exploiting collaborative tagging systems: A contextual semantic approach"; Third International Conference on Systems of Collaboration (SysCo), IEEE Conference Publications, pp: 1 - 6 (2016).

[Maui] Multi-purpose automatic topic indexing, http://www.medelyan.com/software, Last accessed 04 June 2021

[UNESCO World Heritage List] World Heritage Centre - World Heritage List, https://whc.unesco.org/pg.cfm?cid=31&l=en&&&mode=table&order=regi, Last accessed 04 June 2021

[Museum TripAdvisor] Museum Reviews Collected from TripAdvisor, https://www.kaggle.com/annecool37/museum-data, Last accessed 04 June 2021

[TagsFinder] Folksonomy TagsFinder, https://www.tagsfinder.com/, Last accessed 04 June 2021

[UNESCO SKOS] UNESCO Thesaurus SKOS, https://skos.um.es/unescothes/ , Last accessed 04 June 2021

[CB-RS source] simple content-based recommendation engine using python, https://www.kaggle.com/cclark/simple-content-based-recommendation-engine, Last accessed 04 June 2021

[Qassimi et al. 2020] Qassimi, S., Abdelwahed, E.H. and Hafidi, M.: "Folksonomy Graphs Based Context-Aware Recommender System Using Spectral Clustering"; International Journal of Machine Learning and Computing vol. 10, no. 1, pp. 63-68, (2020).

[Qassimi and Abdelwahed 2019 (b)] Qassimi, S., Abdelwahed, E.H.: "Semantic Graph-Based Recommender System. Application in Cultural Heritage"; In: Attiogbé C., Ferrarotti F., Maabout S. (eds) New Trends in Model and Data Engineering. MEDI 2019. Communications in Computer and Information Science, vol 1085. Springer, Cham,(2019).

[Ricci et al. 2010] Ricci, F., Rokach, L. and Shapira, B.: "Recommender Systems Handbook"; (2010) 10.1007/978-0-387-85820-3_1.

[Abbas et al. 2015] Abbas, A., Zhang, L., Khan, S.U.: "A survey on context-aware recommender systems based on computational intelligence techniques"; Computing 97, 667–690 (2015). 10.1007/s00607-015-0448-7

[Sinha and Dhanalakshmi 2019] Sinha, B.B., Dhanalakshmi, R.: "Evolution of recommender system over the time"; Soft Comput 23, 12169–12188 (2019). 10.1007/s00500-019-04143-8

[Geng et al. 2018] Geng, J., Zhang, S., Hui, W.: "Knowledge Recommendation Method for Concept Development of Manufacturing Technology Using Morphological Similarity"; International Journal of Software Engineering and Knowledge Engineering 28 (2018): 1121-1150.

[Gupta and Goel 2018] Gupta S., Goel S.: "Handling User Cold Start Problem in Recommender Systems Using Fuzzy Clustering"; In: Mishra D., Nayak M., Joshi A. (eds) Information and Communication Technology for Sustainable Development. Lecture Notes in Networks and Systems, vol 10. Springer, Singapore.(2018)

[Kumar and Sharma 2016] Kumar, B. and Sharma, N.: "Approaches, Issues and Challenges in Recommender Systems: A Systematic Review"; Indian Journal of Science and Technology. 9. (2016). 10.17485/ijst/2015/v8i1/94892.

[Khalid et al. 2014] Khalid, O., Khan, M., Khan S. and Zomaya, A.: "OmniSuggest: A Ubiquitous Cloud-Based Context-Aware Recommendation System for Mobile Social Networks"; IEEE Transactions on Services Computing, vol. 7, no. 03, pp. 401-414, (2014).

[Jannach et al. 2016] Jannach, D., Resnick, P., Tuzhilin, A., and Zanker, M.: "Recommender systems — beyond matrix completion";Commun. ACM 59, 11 (October 2016), 94–102. 10.1145/2891406

[Gomez-Uribe and Hunt 2016] Gomez-Uribe, C., Hunt, N.: "The Netflix Recommender System: Algorithms, Business Value, and Innovation"; ACM Trans. Manage. Inf. Syst. 6, 4, Article 13 (2016), 19 pages. 10.1145/2843948

[Neidhardt et al. 2015] Neidhardt, J., Seyfang, L., Schuster, R., Werthner, H.: " A picture-based approach to recommender systems. Information Technology and Tourism"; 15, 49-69. (2015).

[Arsintescu et al. 2019] Arsintescu, B., Deo, S., Harris, W.: "PathQuery Pregel: high-performance graph query with bulk synchronous processing"; Pattern Analysis and Applications, 23, 1493-1504. (2019).

[Musto et al. 2018] Musto, C., Lops, P., Degemmis, M., Semeraro, G.: "Semantics-aware Recommender Systems Exploiting Linked Open Data and Graph-based Features"; Companion Proceedings of the The Web Conference 2018. (2018).

[Shalaby et al. 2017] Shalaby, W., AlAila, B., Korayem, M., Pournajaf, L., AlJadda, K., Quinn, S.P., Zadrozny, W.: "Help me find a job: A graph-based approach for job recommendation at scale"; 2017 IEEE International Conference on Big Data (Big Data) (2017): 1544-1553.

[Wang et al. 2019] Wang, X., Wang, D., Xu, C., He, X., Cao, Y., Chua, T. S.: "Explainable reasoning over knowledge graphs for recommendation"; In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 33, pp. 5329-5336) (2019).

[Chaudhari et al. 2016] Chaudhari, S., Azaria, A., Mitchell, T.M.: "An Entity Graph Based Recommender System"; AI Commun. 30: 141-149.(2016)

[Qassimi et al. 2021] Qassimi, S., Abdelwahed, E., Hafidi, M., Qazdar, A.: "Towards a folksonomy graph-based context-aware recommender system of annotated books"; Journal of Big Data, 8, 1-17.(2021).