# The role of tags for recommendation: a survey

Antonina Dattolo, Felice Ferrara, Carlo Tasso

Artificial Intelligence Lab,
Department of Mathematics and Computer Science
University of Udine, I-33100 Udine, Italy
{*antonina.dattolo@uniud.it, felice.ferrara, carlo.tasso*}*@.uniud.it*

***Abstract* — Social tagging is an innovative and powerful mechanism introduced with Web 2.0: it shifts the task of classifying resources from a reduced set of knowledge engineers to the wide set of Web users.**
**Users of social tagging systems define personal classifications which can be used by other peers for browsing available resources. However, due to the absence of rules for managing the tagging process, and to the lack of predefined schemas or structures for inserting metadata and relationships among tags, current user generated classifications dop not produce sound taxonomies. This is a strong limitation which prevents an effective and informed resource sharing. For this reason researchers are modeling innovative recommender systems capable to better support tagging, browsing, and searching for new resources.**

**This paper is a survey which discusses the role of tags in recommender systems: starting from social tagging systems, we analyze various techniques for suggesting content and we introduce the approaches exploited for proposing tags for classifying resources, considering both personalized and not-personalized recommendation.**

***Keywords* — Tag, social tagging, recommender system, personalization**

## I. Introduction

Web 2.0 applications are based on a set of technologies which provide users with tools for creating, sharing and promoting new content: users can easily leave the role of passive consumers of resources and become active producers of information. This trend increases both the knowledge on the Web and the number of available resources. Actually, the growing number of resources prevents an effective access to knowledge: a user needs to read the content of resources for evaluating if it is what she really wants to know.

An effective classification of the resources can improve the access to knowledge. Although the manual process usually reaches high quality levels of classification for traditional document collections, it does not scale to the humongous size of the Web, both in terms of cost, time, and expertise of the human personnel required [1].

On the other hand, also automatic classification tools, based on ontologies, show strong limitations which prevent general categorization of resources. In practice, since it is impossible, to build and maintain universal ontologies covering all possible information needs, their use is possible only on well defined domains [2].

Social tagging applications [3] allow users to freely associate labels (known as tags) to resources, distributing the task of classifying document over the set of Web 2.0 users. Social tagging system do not require significant efforts: human classifiers do not have to follow specific rules and may classify only the set of resources they consider interesting.

Social tagging applications have both private and public aspects [4]:

- Users apply tags for personal aims: typically they associate labels to resources in order to find them again.
- Each user can enjoy the classification applied by other users. A user can browse available documents following classifications provided by other users.

However, due to the freedom of social tagging systems the classification process is not rigorous. This means that the classification applied by a user may be not useful to other users. In order to improve the access to resource classified in social tagging systems we needs of technologies able to overcome this limitation adapting the available classification to the specific user.

Recommender systems have been proposed for supporting users in social tagging system: the main aim of these systems is to filter resources which appear relevant to target users according to their interests and their tagging habits [5].

Currently, recommending tags or content is an open challenge: the aim of this paper is to analyze the state of the art related to the role of tags in social tagging systems, specifying how tags are actually used by users, how recommender systems use tags for suggesting resources and how the process of tagging is supported by tag recommendations.

The rest of this paper is organized as follows: Section II introduces social tagging and recommender systems; Section III deepens the discussion on the use of tags for recommending *content*, while Section IV illustrates the use of tags for recommending *tags*. Final considerations and a look to the future close the paper.

## II. Background

In this section we present an overview of social tagging and recommender systems, describing in particular how tags are applied by users, what are limitations connected to the tagging process, and how the recommender systems are classified.

### A. Social tagging systems

By using social tagging systems users can share resources within a community, upload them, and mainly introduce personal classifications, applying on them, specific tags.

A *tag* is a term freely chosen by a user as significant for a resource; it represents a metadata describing the item, so it can be useful as a keyword to identify or to find again later a document. Tags are also the main mechanism used to browse and search new resources in social tagging systems.

The collection of all the tag assignments performed by a user constitutes her *personomy*, while the collection of all personomies constitutes a *folksonomy*.

Folksonomies [1] substitute traditional hierarchical taxonomies: while taxonomies are defined by a selected set of experts which categorizes resources following a strict hierarchical predefined schema, folksonomies are flat spaces of keywords applied by communities of users. Thanks to the systematic work of experts, taxonomies are more rigorous than folksonomies because the classification is based on a well defined vocabulary. On the other hand, users contributing to a folksonomy are free to add tags without using terms from a specific predefined vocabulary: this allows users to possibly use more than just one term for associating a concept to a resource, providing in such a way a potentially very rich content to folksonomies.

Taxonomies are expensive because they require a systematic work by experts, which have to follow a well defined set of rules. On the other hand, folksonomies are cheaper because the work is distributed among Web 2.0 users.

However, the freedom associated to folksonomies comes with some limitation, which may hinder an effective classification of resources.

- Due to the absence of guidelines, constraints, and control, users can exploit the same tag in different ways: for example, acronyms are a potential cause of ambiguity, or the same tag may be written using different lexical forms (for example: photo, photos, etc.).
- It is frequent to find synonymy, that is different words which describe, more or less, the same concept, or polysemy, that is words which associate multiple meanings to a single word.
- Users classify documents using different levels of expertise and specificity. Since relations among tags are not defined, it is difficult to understand that distinct tags are referring the same concept.

Nevertheless, tags contain very useful, social/semantic information, and their nature can be better understood by analyzing the various motivations/goals that lead a user to perform tagging [1, 6, 4]. Common purposes are:

- **Describe the content**. Tags may be used for summarizing the content of the resource.
- **Describe the type of the document**. Some users use tags for identifying the kind of document. A document may be classified according to its MIME type (for example pdf, doc) or taking into account the publication form (article, blog, book).
- **Describe features and qualities**. Adjectives (such as 'interesting', 'good', and so on) may be used for expressing opinions, emotions or qualitative judges.
- **Associate people to documents**. Tags can report the authors of a document or people involved in a particular task or event. Moreover, tags such as 'my', 'mycomments', 'mystuff', and so on are used to define a relationship between the resources and the tagger.
- **Associate events to documents**. Locations, dates, conferences acronyms are widely used for associating an event to a document.
- **Associate tasks to documents**. Some tags, such as 'mypaper', 'to read', 'jobsearch' reveal personal matters or engagements.

These possible motivations should be considered together with the following two further factors:

1. **Heterogeneity of users**. Taggers have different levels of expertise and goals. This has several consequences: classifications exploited by some users may be not understandable (or acceptable) for other users; different users may describe the content of a resource using distinct vocabularies; different users may have different opinions about a topic; users may not have knowledge about people, events or tasks associated to a resource by other users.

2. **Temporal changes**. Users' knowledge, motivations, and opinions may change over the time. A tag used today for describing an item can be useless in the future: emotions and opinions of people may change; reputation of people evolves; a topic may be not any more interesting to the user.

Currently, tags are used in social communities, social bookmarking applications, and Web 2.0 file sharing systems.

Social communities (both general purpose ones, like Facebook[1], or domain-specific ones, such as aNobii[2]), allow to use tags for expressing opinions and defining relationships among resources and people.

Social bookmarking applications, such as delicious[3], extend traditional bookmarking tools allowing users to upload, label, and access bookmarks from each computer connected to Internet, simplifying the process of content sharing among peers.

Web 2.0 file sharing systems allow users to upload and share with other peers documents. Remarkable examples of these systems are Flickr[4] (photo sharing), Youtube[5] (video sharing), Last.fm[6] (music sharing), BibSonomy[7] (publications sharing).

In [7], authors proposed a taxonomy of these applications which can define tagging rights (who is allowed to tag), can provide different level of tagging support (users can or not see tags applied by others, they could be or not supported by an automatic system) and can support or not social interaction among users.

## *B.* Recommender systems

The increasing volume of information on the Web is the main motivation for recommender systems: they support users while they interact with large information spaces, directing them toward the information they need [8]; these systems model user interests, goals, knowledge and tastes, monitoring and modeling some implicit and/or explicit feedback. User feedback can be acquired by means of ratings for quantifying a relation between the user and an item: the ratings may be explicit, when they require the user evaluation, or implicit when they are automatically generated by the system in terms of measures, such as, for example, the time spent by a user on a Web page. By taking in consideration the ratings provided by a user, a recommender system defines a personalized order of importance for the set of available resources.

According to the strategy used for collecting and evaluating ratings, recommender system can be classified into three main classes:

1. **Collaborative filtering recommender systems**. Collaborative filtering recommender systems [9] implement the idea that a user, that shares interests, knowledge and goals with a community, with good probability may be interested in the documents which appear relevant to that community. For this reason, a collaborative filtering recommender system builds and maintain a profile for each user and computes similarities among users: similar users, known as *neighbors*, are the bridge which allows the user to receive suggestions for new potentially interesting content.

   In turn, collaborative filtering approaches can be classified into two classes [10]:

   - *Model-based* approaches, which build a probabilistic model for predicting, on the basis of user history, her future rating assignments.
   - *Memory-based* approaches, which use statistical techniques for identifying users with common behaviour. When the neighborhood has been defined, neighbors'feedbacks are combined for generating a list of recommendations.

2. **Content-based recommender systems**. Content-based recommender systems [11] analyze the past user activities looking for resources she liked; they model resources extracting some features from documents. The user profile is then defined describing what features are interesting for the user. The relevance of a new resource for a user is calculated matching the representation of the resource to the user profile.

3. **Hybrid recommender systems**. The results returned by collaborative and content-based recommender systems can be combined in a hybrid recommender system [12], applying several different strategies:

   - A *weighted* hybrid recommender system merges results of different techniques. The score of an item is defined as a weighted sum of scores calculated by different recommender systems.
   - A *cascade* hybrid recommender system filters a starting set of resources using a first recommender system. The ranking of these resources is then refined by a second recommender system.
   - A *switching* hybrid recommender system uses one of the available recommender system according to some criterion.

1 http://www.facebook.com/
2 http://www.anobii.com/
3 http://delicious.com/
4 http://www.flickr.com/
5 http://www.youtube.com/
6 http://www.lastfm.it/
7 http://www.bibsonomy.org/

- A *feature combination* hybrid recommender system considers collaborative data as the features used in content-based approaches.

In this work, we distinguish the resources, suggested by recommender systems, in *content* (for example, documents, references, or URL's), and *tag*, and we discuss these two different typologies of recommendations in two separated sections (recommending content in Section III and tags in Section IV).

## III. Recommending content

In this section, we discuss of recommendations related to content which can satisfy the user information needs.

Typically, a recommendation is provided to a user in two different situations:

Tags provided by users implicitly, by queries (sets of keywords), or explicitly, analyzing the interaction of the user with the system, can be used to improve the user modeling process [13], to personalize the behavior of the application [14] and to calculate recommendations. Basic recommendation approaches give higher relevance to a resource if it shares with the query a high number of tags; or if it has been frequently associated to that tags.

These approaches find popular resources; although popularity is a good feature to give confidence in results, other parameters should be considered, such as the personal needs or habits of a user, in order to do not return to all users the same results.

In [5, 15], the authors suggest that tags are a useful information to understand the relationship between a user and one or more items. Following this idea, recently, several researchers proposed some attempts for providing personalized recommendations.

The following two subsections III.A and III.B describe collaborative and content-aware strategies for recommending content using tags.

### *A.* Tag-aware Collaborative Recommender systems

Tag-aware collaborative recommender systems apply collaborative filtering techniques using tags to model user interests. These systems are based on the idea that users with similar interests have a similar tagging history. In particular, tag-aware memory-based approaches have been described in [16], [17] and [18]: in [16], the authors propose a mechanism, called *social ranking*, which analyzes the tagging similarity among users in order to produce a personalized ranking for the available resources. This approach is based on the idea that looking to the users' tagging activity it is possible to quantify the similarity between two users. For this reason, each user is modeled by a vector which counts the number of times that the user applied each tag. Similarly, each tag is described by a vector which counts the number of times the tag have been applied to each resource. The cosine similarity is used for evaluating both the similarity between two users and two tags. The similarity between two tags takes into account the heterogeneity of users which may apply different tags to describe the same concept.

When a user submits a query, the resources are ranked following two step: query expansion and ranking. The *query expansion* considers a larger set of query tags including tags which appear similar to the starting tags: the $K$ nearest neighbor strategy is used to include top $K$ similar tags for each input tag. The second step defines the final *ranking*: the relevance of a resource depends on the relevance of tags associated to it and on the similarity of users that bookmarked it.

In [17], the authors extract from tags a context for recommendations. In particular, this model is based on the observation that a social bookmarking application allows users to add documents which belong to different domain of interests, and that this is a strong limitation for collaborative recommender systems [9]. So, they propose to use tags for preventing this drawback. Starting from a user's bookmark, the recommendation process takes into account only users that tagged that specific resource. Tags applied on the resource are used for filtering the set of documents and the users to be considered for generating recommendations. Then, the similarity among filtered users is evaluated counting the number of shared bookmarks while the relevance of a resource is calculated taking into account the similarity of users that bookmarked it.

However, this work does not consider limitations connected to the freedom of tagging activity. In [18], the authors address this issue by clustering tags with similar meanings. This approach models each tag by a vector which counts the number of times that the tag has been applied on each resource. The distance between two tags is calculated using the cosine similarity. Using this information the user's tags are grouped in clusters. Tag clusters are then used to split resources tagged by the user into different subcollections. A cluster of tags and related resources defines a topic of interest for the user. Each cluster is extended considering also other similar tags. The approach defines a distinct neighborhood for each topic of interest: given a topic of interest, the neighbor selection is exploited considering users and resources associated to tags which are in the extended cluster of tags. In particular, top K neighbors are chosen selecting users which share the highest number of resources with the target user. Finally, the relevance of a resource for a topic of interest is calculated taking into account the relevance of neighbors that bookmarked it.

A model-based recommender system have been proposed in [19] where the authors describe *TagiCoFi* (Tag informed Collaborative Filtering), a framework that includes the votes and the tags provided by users, and builds on these a mathematical model able to predict user's ratings.

### *B.* Tag aware Content-Based Recommender systems

Recently, two content-based recommender frameworks [20, 21] integrated tags for modeling user interests and providing personalized ranking of unseen items.

In [20], the authors describe a recommender system which uses tags'clustering for fighting redundancy of folksonomies; both users and documents are represented by means vectorial descriptions. In particular, each user is modeled by a vector over the set of tags, counting how many times a user applied each tag. Similarly, each resource is described by a vector over the set of tags: the $tf * idf$ measure (term frequency $*$ inverse document frequency) [22] is used to infer the relevance of each tag for each resource. This approach uses tag clustering to group tags with similar meanings. In order to create tag clusters, the similarity between tags is calculated: each tag is described as a vector which counts how many times a tag has been applied on each resource; the cosine similarity over two vectors defines the similarity between two tags.
A cluster of tags is an intermediary between users and resources; in fact, looking at the user profile it is possible to understand what tag cluster is relevant for the user. On the other side, the vectorial description of resources is used to detect resources relevant for a specific cluster. The recommendation algorithm uses three input, a tag, a user profile and tag clusters, and produces an ordered set of items. The user profile allows to give a score to clusters which contain input tag. In order to assign a relevance value to a resource, the relevance of a cluster is multiplied by the relevance of the resource for the cluster.

In [21], the authors present a different approach where both the textual description of items and tags are used to build the user profile. This approach uses the synsets of Wordnet[8], structures defined as the sets of words with a similar meaning for defining a semantic indexing of documents. A disambiguation strategy associates a synset to each word in the document looking to words that precede and follow it. Similarly, tags are also disambiguated using the textual content of the resource. In this way, a document is defined as a bag-of-sysnsets in opposition to the classical bag-of-words model [22]. Using this descriptive model, a Bayesian classifier considers resources bookmarked by the user to learns about the synsets which are relevant to the user. Matching the sysnset representation of documents with the synsets in the user profile, the recommender system calculates a relevance value for each resource.

8 http://wordnet.princeton.edu/

## IV. Recommending tags

Recommender systems can suggest to the user not only content, by also tags for labeling for a resource: this typology of recommendation aims to improve the usage of social tagging applications [23]:

- Tag suggestions can *increase* the probability of having *tagged resources*. Users just select one or more suggested tags instead of summarizing a resource in order to find representative tags.
- Tag suggestions can promote a *common vocabulary* among users. Proposing a well defined set of tags it is possible to reduce problems connected to the absence of both guidelines and supervised methodologies to the tagging process.

To reach this scope several approaches have been proposed. These methodologies can be classified by means of three classes [24]:

- *Content-based tag recommender systems*. These systems look to the content of resources for suggesting relevant terms to the user.
- *Collaborative tag recommender systems*. Collaborative approaches analyze metadata associated by users to resources inferring the relevance of tags for the specific resource.
- *Graph based tag recommender systems*. These systems build a graph representation of a folksonomy which can be explored for finding tags to be recommended.

Different tag recommender systems use different criteria for evaluating the relevance of tags to be suggested. According to the relevance criteria tag recommender systems can be divided into two classes:

1. Not personalized tag recommender systems. These systems select for each document the best tags to describe the specific resource, ignoring user's tagging habits. Different users will receive the same suggestions for the same resource.
2. Personalized tag recommender systems. These systems suggest the set of the most relevant tags for a resource according to the specific user and her personal way to classify resources.

The following two subsections IV.A and IV.B describe these two classes, classifying them into content-based, collaborative and graph-based approaches.

## *A.* Not Personalized Tag recommendations

Not personalized content-based tag recommender systems are based on techniques well known in information retrieval, natural language processing and machine learning for classifying documents. These approaches split the content of a textual resource into separated phrases (called $n$-grams) and give a score to each of them. Example of this approach are described in [25, 26, 27, 28].

In [25], the authors assign a weight to each term in the text using the $tf * idf$ metric and then recommend top three weighted terms.

KEA [26] is a framework based on a Bayesian classifier which extracts keyphrases (relevant phrases) from a document. It uses a supervised approach, that needs of a training phase in which KEA takes into account two features: the $tf * idf$ measure and the first occurrence of each term.

A Part-Of-Speech (POS) tagger have been used in [27] for filtering the set of extracted keyphrases: authors identified 56 potential pos-patterns in order to reduce the set of candidate keyphrases.

All these methods suggest only terms which appear in the document. For overcoming this limitation in [28] the use of ontologies has been proposed. In this approach a set of terms is extracted from the document and is used for browsing the ontology in order to find other more abstract and conceptual terms. However, the performance of this approach depends on the quality of the available ontology.

On the other hand, collaborative not personalized tag recommender systems infer the relevance of tags observing the behavior of the whole community. Following this idea the simplest approach can suggest to users the most popular tags for a resource. However, due to sparsity of social tagging systems there are resources tagged by only few people and for this reason more sophisticated methods have been proposed [29, 30].

One of the most popular collaborative tag recommender system is *AutoTag* [29], which suggests tags for weblog posts. This framework recommends tags executing three steps: it selects resources similar to the starting document, extracting the tags associated to these resources; associates a weight to each tag according to its frequency; and suggests top ranked tags.

*TagAssist* [30] outperforms *AutoTag* thanks to a preprocessing phase where the Porter's stemmer [31] is used to compress the set of tags.

These collaborative tag recommender systems do not take into account the reputation of users. Graph-based tag recommender systems can take into account also this information representing a folksonomy by a graph where users, tags and resources are nodes connected by edges which describe the tagging activities. An example of this approach is *FolkRank* [32]. *FolkRank* is an algorithm, inspired to the popular *PageRank* [33] algorithm, where the relevance of a resource depend on the relevance of users and tags associated to it. In a similar way, the relevance of tags and resources can be stated simply looking at the relationships among nodes of the graph.

## *B.* Personalized tag recommendations

Personalized collaborative approaches estimate the relevance on a tag considering the user tagging preferences. Personalized collaborative strategies [34, 35]are based on the hypothesis that users with similar tagging habits have a good probability to use similar tags in the future.

In [34], the authors adapt the classical $K$-nearest neighbor algorithm to the task of generating a list of recommended tags. While traditional $K$-nearest neighbor algorithm takes in input a user and produces a set of recommended resources, [34] proposes to use both the user and the resource to calculate a neighborhood of similar users. In particular, the set of neighbors is defined considering users which tagged the specific resource. This approach associates a score to each neighbor: users which bookmarked the same resources and used the same tags obtain a greater score. This score is used for assigning a weight to each tag applied to the resource. Tags assigned by similar users will be more relevant than others.

The main difficulty for personalized collaborative tag recommender systems is to model the ternary relationship which connect users, resources and tags (users-resources-tags). This relationship may be managed splitting it into two dimensional relationships (users-resources, resources-tags, tags-users); or, such as in [35], it is modeled using a 3-order tensor, a geometric entity which expresses a relationship among vectors. Latent semantic analysis is performed on tensors to capture the latent association among users, resources and tags. This approach builds a set of quadruplets $(u, r, t, likeliness)$ where each quadruplet describes the $likeness$ that the user $u$ will tag the resource $r$ with the tag $t$.

Personalized content based strategies analyze the relationship between the content of a resource and the tags applied by a user in order to predict tags for new resources. Examples of this approach are provided in [36, 24]: the system proposed in [36] recommends tags, using a Bayesian classifier for each tag employed by the user. Each classifier is trained using the textual content of documents tagged by a specific tag. In this way the text of a new document can be used for evaluating if a tag can be suggested for other resources.

STaR (Social Tag Recommender System) [24] is based on an approach similar to *AutoTag*. The main difference is that STaR provides personalized tag suggestions. This framework collects two sets of documents similar to the starting resource: a set containing resources tagged by the target user; a set containing documents tagged by other users. Tags applied by the target user are weighted according to the similarity of tagged resource to the starting one. In a similar way a weight is assigned to tags applied by other users. At the end, the two set of tags are merged and the final score of tags is calculated as a linear combination of scores associated to tags.

## V. Final considerations and future work

In this paper we described some important limitations of current social tagging systems, discussing the different methodologies applied by recommender systems for improving the access and the classification of resources. Summarizing, collaborative approaches model the user's tagging behavior and look for users with similar tagging habits: these systems are based on the idea that users with similar interests search and bookmark similar resources. On the other hand, content-based approaches use tags to build a user profile and then suggest resources which appear relevant for the specific user profile. Both collaborative and content approaches have limitations mainly due to ambiguity of tags.

Some recommender systems use clustering for fighting redundancy of folksonomies [18, 20], but also these approaches have some open issues: during the time a user may apply the same tag for expressing different concepts; a tag may be in several clusters.

Other works try to tackle the absence of semantic relationships among tags associating a context to them [17, 21]. However, it is hard to extract a context from some generic tags, such as, for example, *to read*, *my paper*, *job*.

A chance for reducing ambiguity is to support the manual tagging process using tag recommender systems, that suggest representative tags for a resource improving the user satisfaction: users do not need to summarize the resource but they can select among the suggested tags.

Not personalized strategies ignore the heterogeneity of users suggesting the set of tags which appear more representative for a resource according to a global measure. Personalized approaches, on the other hand, tailor the selection of recommended tags considering the user's past tagging activities. However, in literature there is not evidence that personalized tag recommendation approaches outperform not personalized strategies, or vice versa, that not personalized techniques improve the user satisfaction. The evaluation of all these recommender systems is still an open challenge: results of different systems have been evaluated using different dataset and following different evaluation strategies.

Other interesting researches on tags are actually ongoing.

An interesting use of tags have been proposed in [37], where authors introduce the concept of *tagsplanations* which are explanations based on community tags. Explaining the motivations at the basis of a recommendation improves the user satisfaction. *Tagsplanations* take into account two components: tag relevance and tag preference. Tag relevance refers to the representativeness of a tag for describing a resource, while tag preference defines the user's sentiment toward a tag.

Another line of research is concerned with extracting basic semantic relations from folksonomies or adding more ontology-like features to social tagging [28]. For example, *Folk2onto* [38] maps social tags (taken from del.icio.us) to ontological categories (using a Dublin Core-based ontology) in order to classify and give a proper structure to the tagged resources. However, the tasks of associating semantic to tags, extracting semantic relation among tags, is currently far from a real solution.

### References


[1] Dattolo, A., Tomasi, F., Vitali, F. In: Towards disambiguating social tagging systems. Volume 1, Chapter 20. IGI-Global (2010) 349–369

[2] Shirky, C.: Ontology is overrated: Categories, links, and tags. `http://www.shirky.com/writings/ontology_overrated.html` (2005)

[3] Mathes, A.: Folksonomies cooperative classification and communication through shared metadata. computer mediated communication - lis590cmc. `http://www.adammathes.com/academic/computer-mediatedcommunication/folksonomies.html` (2004)

[4] Golder, S., Huberman, A.: The structure of collaborative tagging systems. Journal of Information Science **32**(2) (2006) 198–208

[5] Sen, S., Vig, J., Riedl, J.: Tagommenders: connecting users to items through tags. In: Proceedings of the 18th international conference on World wide web, New York, NY, USA, ACM (2009) 671–680

[6] Xu, Z., Fu, Y., Mao, J., Su, D.: Towards the semantic web: Collaborative tag suggestions. In: Proceedings of the Collaborative Web Tagging Workshop at the World Wide Web Conference. (2006)

[7] Marlow, C., Naaman, M., Boyd, D., Davis, M.: Ht06, tagging paper, taxonomy, flickr,academic article, to read. In: Proceedings of the Seventeenth Conference on Hypertext and Hypermedia, New York, NY, USA, ACM (2006) 31–40

[8] Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transaction on Knowledge and Data Engeneering **17**(6) (2005) 734–749

[9] Schafer, J., Frankowski, D., Herlocker, J., Sen, S.: Collaborative filtering recommender systems. In Brusilovsky, P., Kobsa, A., Nejdl, W., eds.: The Adaptive Web. LNCS (4321). Springer-Verlag, Berlin, Germany (May 2007) 291–324

[10] Breese, J., Heckerman, D., Kadie, C.: Empirical analysis of predictive algorithms for collaborative filtering. In: Proceeding of the Fourteenth Conference on Uncertainty in Artificial Intelligence (UAI), Morgan Kaufmann (1998) 43–52

[11] Pazzani, M.J., Billsus, D.: Content-based recommendation systems. In Brusilovsky, P., Kobsa, A., Nejdl, W., eds.: The Adaptive Web. LNCS (4321). Springer-Verlag, Berlin, Germany (May 2007) 325–341

[12] Burke, R.: Hybrid web recommender systems. In Brusilovsky, P., Kobsa, A., Nejdl, W., eds.: The Adaptive Web. LNCS (4321). Springer-Verlag, Berlin, Germany (May 2007) 377–408

[13] Carmagnola, F., Cena, F., Console, L., Cortassa, O., Gena, C., Goy, A., Torre, I., Toso, A., Vernero, F.: Tag-based user modeling for social multi-device adaptive guides. User Modeling and User-Adapted Interaction **18**(3) (2008) 497–538

[14] Zhao, S., Du, N., Nauerz, A., Zhang, X., Yuan, Q., Fu, R.: Improved recommendation based on collaborative tagging behaviors. In: IUI '08: Proceedings of the 13th international conference on Intelligent user interfaces, New York, NY, USA, ACM (2008) 413–416

[15] Van Setten, M., Brussee, R., Van Vliet, H., Gazendam, L., Van Houten, Y., Veenstra, M.: On the importance of who tagged what. In: Proceedings of the Workshop on the Social Navigation and Community based Adaptation Technologies in conjunction with Adaptive Hypermedia and Adaptive Web-Based Systems 2006. (2006) 552–561

[16] Zanardi, V., Capra, L.: Social ranking: Finding relevant content in web 2.0. In: Proceeding of the 2nd ACM International Conference on Recommender Systems, Lausanne, Switzerland (October 2008) 51–58

[17] Nakamoto, R., Nakajima, S., Miyazaki, J., Uemura, S.: Tag-based contextual collaborative filtering. Journal of Intelligent Information Systems **34**(2) (2007)

[18] Dattolo, A., Ferrara, F., Tasso, C.: Neighbor selection and recommendations in social bookmarking tools. In: Proceeding of the Ninth International Conference on Intelligent Systems Design and Applications, Pisa, Italy (30 November - 2 December 2009)

[19] Zhen, Y., Li, W., Yeung, D.: Tagicofi: tag informed collaborative filtering. In: Proceedings of the third ACM conference on Recommender systems, New York, NY, USA, ACM (2009) 69–76

[20] Shepitsen, A., Gemmell, J., Mobasher, B., Burke, R.: Personalized recommendation in collaborative tagging systems using hierarchical clustering. In: Proc. of the 2nd International Conference on Recommender Systems, Lausanne, Switzerland (October 2008)

[21] De Gemmis, M., Lops, P., Semeraro, G., Basile, P.: Integrating tags in a semantic content-based recommender. In: Proceedings of the ACM conference on Recommender systems, New York, NY, USA, ACM (2008) 163–170

[22] Manning, C., Raghavan, P., Schutze, H.: Introduction to Information Retrieval. Cambridge University Press (2008)

[23] Jäschke, R., Marinho, L., Hotho, A., Schmidt-Thieme, L., Stumme, G.: Tag recommendations in social bookmarking systems. AI Communications **21**(4) (2008) 231–247

[24] Musto, C., Narducci, F., De Gemmis, M., Lops, P., Semeraro, G.: Star : a social tag recommender system. In: Proceedings of the ECML/PKDD 2009 Discovery Challenge Workshop. (2009)

[25] Brooks, H., Montanez, N.: Improved annotation of the blogosphere via autotagging and hierarchical clustering. In: Proceedings of the 15th International conference on World Wide Web, New York, NY, USA, ACM (2006) 625–632

[26] Witten, I., Paynter, G., Frank, E., Gutwin, C., Nevill-Manning, C.: Kea: Practical automatic keyphrase extraction. In: Proceedings of the 4th ACM International Conference on Digital Libraries, ACM Press (1999) 254–255

[27] Hulth, A.: Improved automatic keyword extraction given more linguistic knowledge. In: Empirical methods in natural language processing, Morristown, NJ, USA, Association for Computational Linguistics (2003) 216–223

[28] Baruzzo, A., Dattolo, A., Pudota, N., Tasso, C.: Recommending new tags using domain-ontologies. In: Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology. (2009) 409–412

[29] Mishne, G.: Autotag: a collaborative approach to automated tag assignment for weblog posts. In: Proceedings of 15th International Conference on World Wide Web, ACM (2006) 953–954

[30] Sood, S., Owsley, S., Hammond, K., L., B.: Tagassist: Automatic tag suggestion for blog posts. In: Proceedings of the International Conference on Weblogs and Social Media. (2007)

[31] Porter, M.: An algorithm for suffix stripping. Program **14**(3) (1980) 130–137

[32] Jäschke, R., , Hotho, A., Schmidt-Thieme, L., Stumme, G.: Folkrank: A ranking algorithm for folksonomies. In: Proceedings of FGIR. (2006)

[33] Page, L., Brin, S., Motwani, R., Winograd, T.: The pagerank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab (1999)

[34] Gemmell, J., Schimoler, T., Ramezani, M., Mobasher, B.: Adapting k-nearest neighbor for tag recommendation in folksonomies. In: Proceedings of the 7th Workshop on Intelligent Techniques for Web Personalization and Recommender Systems in in conjunction with the 21st International Joint Conference on Artificial Intelligence. (July 2009)

[35] Symeonidis, P., Nanopoulos, A., Manolopoulos, Y.: Tag recommendations based on tensor dimensionality reduction. In: Proceedings of the 2008 ACM conference on Recommender systems, New York, NY, USA, ACM (2008) 43–50

[36] Basile, P., Gendarmi, D., Lanubile, F., Semeraro, G.: Recommending smart tags in a social bookmarking system. In: Bridging the Gep between Semantic Web and Web 2.0. (2007) 22–29

[37] Vig, J., Sen, S., Riedl, J.: Tagsplanations: explaining recommendations using tags. In: IUI '09: Proceedings of the 13th International Conference on Intelligent User Interfaces, New York, NY, USA, ACM (2009) 47–56

[38] Sotomayor, B.: folk2onto: Mapping social tags into ontological categories. `www.deli.deusto.es/Resources/Documents/folk2onto.pdf` (2006)