Expert Systems with Applications 39 (2012) 10059–10072

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A literature review and classification of recommender systems research

Deuk Hee Park, Hyea Kyeong Kim, Il Young Choi, Jae Kyeong Kim *

Department of Management, School of Management, KyungHee University, 1 Hoeki-Dong, Dongdaemoon-Gu, Seoul 130-701, Republic of Korea

ARTICLE INFO

Keywords:
Recommender systems
Literature review
Data mining technique
Classification

ABSTRACT

Recommender systems have become an important research field since the emergence of the first paper on collaborative filtering in the mid-1990s. Although academic research on recommender systems has increased significantly over the past 10 years, there are deficiencies in the comprehensive literature review and classification of that research. For that reason, we reviewed 210 articles on recommender systems from 46 journals published between 2001 and 2010, and then classified those by the year of publication, the journals in which they appeared, their application fields, and their data mining techniques. The 210 articles are categorized into eight application fields (books, documents, images, movie, music, shopping, TV programs, and others) and eight data mining techniques (association rule, clustering, decision tree, k-nearest neighbor, link analysis, neural network, regression, and other heuristic methods). Our research provides information about trends in recommender systems research by examining the publication years of the articles, and provides practitioners and researchers with insight and future direction on recommender systems. We hope that this paper helps anyone who is interested in recommender systems research with insight for future research direction.

1. Introduction

Recommender systems have become an important research area since the emergence of the first research paper on collaborative filtering in the mid-1990s (Resnick, Iakovou, Sushak, Bergstrom, & Riedl, 1994; Shardanand & Maes, 1995). In general, recommender systems directly help users to find content, products, or services (such as books, digital products, movies, music, TV programs, and web sites) by aggregating and analyzing suggestions from other users, which mean reviews from various authorities, and users (Frias-Martinez, Chen, & Liu, 2009; Frias-Martinez, Magoulas, Chen, & Macredie, 2006; Kim, Ji, Ha, & Jo, 2010). These systems use analytic technology to compute the probability that a user will purchase one of the products at each place, so that users will receive recommendations for the right products to purchase.

Recommender systems are generally classified into collaborative filtering (CF) and content-based filtering (CB). In general, CF uses an information filtering technique based on the user's previous evaluation of items or history of previous purchases. However, this technique has been known to reveal two major issues: sparsity problem and the scalability problem (Claypool et al., 1999; Sarwar, Karypis, Konstan, & Riedl, 2000a, 2000b). In contrast, CB analyzes a set of documents rated by an individual user and uses the contents of the documents, as well as the provided ratings, to infer a user profile that can be used to recommend additional items of interest (Basu, Hirsh, & Cohen, 1998). However, the syntactic nature of CB, which detects similarities between items that share the same attribute or characteristic, causes overspecialized recommendations that only include items very similar to those of which the user is already aware (Lopez-Nores, Garca-Duque, Frenandez-Vilas, & Bermejo-Munoz, 2008).

Over the last decade, most of researchers have studied new approaches of recommender systems in order to solve these problems of CF and CB, and to implement them into real world situations. Specifically, applying data mining techniques to recommender systems has been effective in providing personalized information to the user by analyzing his or her preferences.

However, more research is needed to be applicable in real world situations because research fields on recommender systems are still broader and less mature than in other research areas. Therefore, the existing articles on recommender systems must be reviewed with an eye toward the next generation of recommender systems, which will improve recommendation methods to offer more useful and appropriate information to users.

In this research, we reviewed and classified articles on recommender systems that were published in academic journals between 2001 and 2010, in order to gain insight on recommender systems. This research is organized as follows:

(1) The research methodology used in this study is reported.
(2) Criteria for classification of research papers on recommender systems are presented.

* Corresponding author. Tel.: +82 2 961 0508.
E-mail address: jaek@khu.ac.kr (J.K. Kim).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2012.02.038

(3) Research papers on recommender systems are analyzed and the results of their classifications are presented.
(4) Conclusions are presented, and the limitations and implications of this study are discussed.

We hope that this research will accentuate the importance of recommender systems and provide researchers and practitioners with insight on recommender systems research.

2. Research methodology

The purpose of this study is to understand the trend of recommender systems research by examining the published articles, and to afford practitioners and academics with insight and future direction on recommender systems.

Hence, we will verify the distribution of research papers on recommender systems by their year of publication, and classify the research papers by the data mining techniques used for recommendation and by the application fields used. However, considering the nature of research on recommender systems, it would be difficult to confine each paper to a specific discipline. Additional proof of this difficulty can be seen from the fact that research papers on recommender systems are scattered across diverse journals such as marketing, information technology, information science, computer science, and management. As a result, it is necessary to compile the increasing number of research papers on recommender systems systematically. The following electronic journal databases were searched to provide a comprehensive bibliography of research papers on recommender systems:

- ABI/INFORM Database;
- ACM Portal;
- EBSCO Academic Search Premier;
- EBSCO Business Source Premier;
- IEEE/IEE Library;
- Science Direct.

The search process of research papers on recommender systems was performed on the top 125 MIS journals. The search was performed based on five descriptors: "Recommender system", "Recommendation system", "Personalization system", "Collaborative filtering", and "Contents filtering". Two authors reviewed the full text of each research paper, and papers that were not truly related to recommender systems were deleted if the two authors agreed to do so. If the authors' opinions were different, another author reviewed the paper and decided whether to delete it or not. The following research papers, set forth in the description below, were excluded because they were unfit for our research:

- Conference papers, master's and doctoral dissertations, textbooks, unpublished working papers, non-English papers, and news articles were eliminated, Unlike these publications, papers published by academic journals are thought to be reliable and worthy of comment, because they are published after peer review.
- Because research on recommender systems is relatively current, we have only searched research articles published between 2001 and the end of 2010. This 10-year period is considered to be representative of recommender systems research.
- Only research papers that described how recommender systems can be applied were chosen.

We selected 210 research papers on recommender systems from 46 journals. Each research papers was prudently reviewed and classified into one of the eight categories in the application fields and data mining techniques. Although the investigation was not exhaustive, it provides as a comprehensive basis for understanding recommender system research.

3. Classification method

Our classification framework consists of recommendation fields and data mining techniques. In this research, we classify the research papers that were reviewed into eight categories of application fields and eight categories of data mining techniques. The overall graphical classification framework for recommender systems research papers is presented in Fig. 1.

3.1. Classification framework for application fields

Many recommender systems have been used to provide users with information to help them decide which products to purchase (Schafer, Joseph, & Riedl, 2001). However, it is not easy to find papers that classify research papers systematically, even though recommender systems have been applied to diverse business areas. Accordingly, it is meaningful to investigate application fields. Our research adopts the basic classification scheme of Schafer et al., 2001, who have classified recommendation applications by real world, such as books, movies, music, shopping and others. We classify research papers by application fields such as books, documents, images, movies, music, shopping, TV programs and others. Through in-depth reviews of research papers, classifying shopping fields involves online, offline, and mobile shopping product, classifying document fields involves papers, blogs and web pages. Also, other fields involve a minority of recommendation fields such as hotel, travel, and food.

3.2. Classification framework for data mining techniques

In general, data mining techniques are defined as extracting or mining knowledge from data. These techniques are used for the exploration and analysis of large quantities of data in order to discover meaningful patterns and rules (Berry & Linoff, 2004). They can be used to lead decision making and to predict the effect of decisions. Significantly, many researchers have used data mining techniques to improve the performance of recommender systems. Consequently, it is meaningful to classify the research papers according to data mining techniques. We widely classified data mining techniques into the following eight categories: association rule, clustering, decision tree, k-nearest neighbor, link analysis, neural network, regression, and other heuristic methods.

(1) *Association rule:* Association rule mining refers to the discovery of all association rules that are above user-specified minimum support and minimum confidence levels. Given a set of transactions in which each transaction contains a set of items, an association rule applies the form $X \Rightarrow Y$, where X and Y are two sets of items (Cho, Kim, & Kim, 2002).
(2) *Clustering:* The clustering method identifies a finite set of categories or clusters to describe data. Among the clustering methods, the most popular are *K*-means and self-organizing map (SOM). *K*-means takes the input parameter, *K*, and partitions a set of n objects into *K* clusters (Berry & Linoff, 2004). SOM is a method for an unsupervised learning, based on an artificial neurons clustering technique (Lihua, Lu, Jing, & Zongyong, 2005).
(3) *Decision tree:* Most popular classification methods are decision tree induction. Decision tree induction techniques build decision trees to label or categorize cases into a set of known

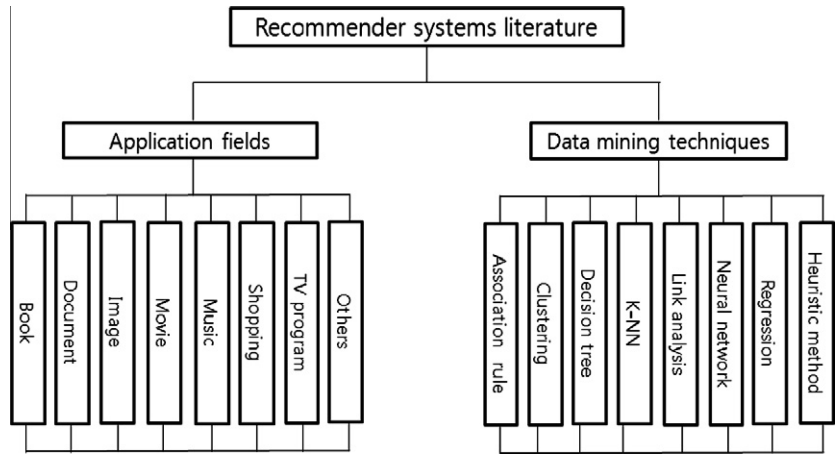

Fig. 1. Classification framework.

classes. The top node in a tree is called as a root node. A decision tree is a tree in which each internal (non-leaf) node represents a test on an attribute, each branch represents an outcome of the test, and each terminal (leaf) node represents a class prediction (Kim, Cho, Kim, Kim, & Suh, 2002).

(4) *k-Nearest* neighbor: The k-NN (k-nearest neighbor) model, a typical traditional CF-based recommender system, makes recommendations according to the following three phases. (1) Recommender systems construct a user profile using the user's preference ratings, which are obtained either directly from explicit ratings of items or indirectly from purchase or usage information. (2) Recommender systems apply statistical or machine learning techniques to discover k users, known as neighbors or recommenders, who in the past have shown similar behaviors. A neighborhood is formed based on the degree of similarity between a mark user and other users. (3) Once a neighborhood is formed for a target user, recommender systems make a top-n item set that the target user is most likely to purchase by analyzing the items in which neighbors have exhibited interest (Kim, Kim, & Ryu, 2009).

(5) *Neural network:* A neural network is a parallel distributed information processing system that is able to learn and self-organize. This system consists of a large number of uncomplicated processing entities which are interconnected to form a network that conducts complex computational tasks (Ibnkahla, 2000). A neural network builds a class of very pliable model that can be used for a diversity of different applications, such as prediction, non-linear regression, or classification (Anders & Korn, 1999).

(6) *Link analysis:* Link analysis discovers relations between domains in large databases. One type of link analysis, social network analysis is a sociological approach for analyzing patterns relationships and interactions between social actors in order to find a fundamental social structure. Also, link analysis has presented great potential in improving the accuracy of web searches. Link analysis consists of PageRank and HITS algorithms. Most link analysis algorithms handle a web page as a single node in the web graph (Cai, He, Wen, & Ma, 2004).

(7) *Regression:* Regression analysis is a powerful process for analyzing associative relationships between dependent variables and one or more independent variables. It has been used for curve fitting, prediction, and testing systematic hypotheses about relationships between variables (Malhotra, 2007).

(8) *Other heuristic methods:* Heuristic methods have been developed by adding new method to existing methods. Heuristic methods include mixture models and the, ontology method.

3.3. Classification process

Each of the selected research papers was reviewed and classified according to the proposed classification framework by two of the four authors of this paper (first team). The other two authors (second team) made a final verification of the classification results. The classification process is composed of the following four steps:

(1) Electronic database search.
(2) Initial classification by one of the two researchers in the first team.
(3) Independent verification of classification results by the other researcher in the first team.
(4) Final verification of classification results discussed by the second team.

The selected criteria and evaluation framework is represented in Fig. 2. The research papers were analyzed by year of publication, by journals in which the research papers were published, and by application fields and data mining techniques.

4. Classification of research papers

We selected a total of 210 research papers from 46 journals and classified them according to the classification framework. The results of our analysis will supply guidelines for future research on recommender systems. The details are described below.

4.1. Distribution by year of publication

The distribution of research papers by year of publication between 2001 and 2010 is shown in Fig. 3. It is apparent that publications related to recommender systems steadily increased between 2000 and 2004, and rapidly increased between 2007 and 2010. The decrease of research papers between 2005 and 2006 is thought to be because recommender systems research apparently extended a new application field between 2005 and 2006. Whereas a majority of recommender systems research between 2005 and 2006 were limited to movie and shopping fields,

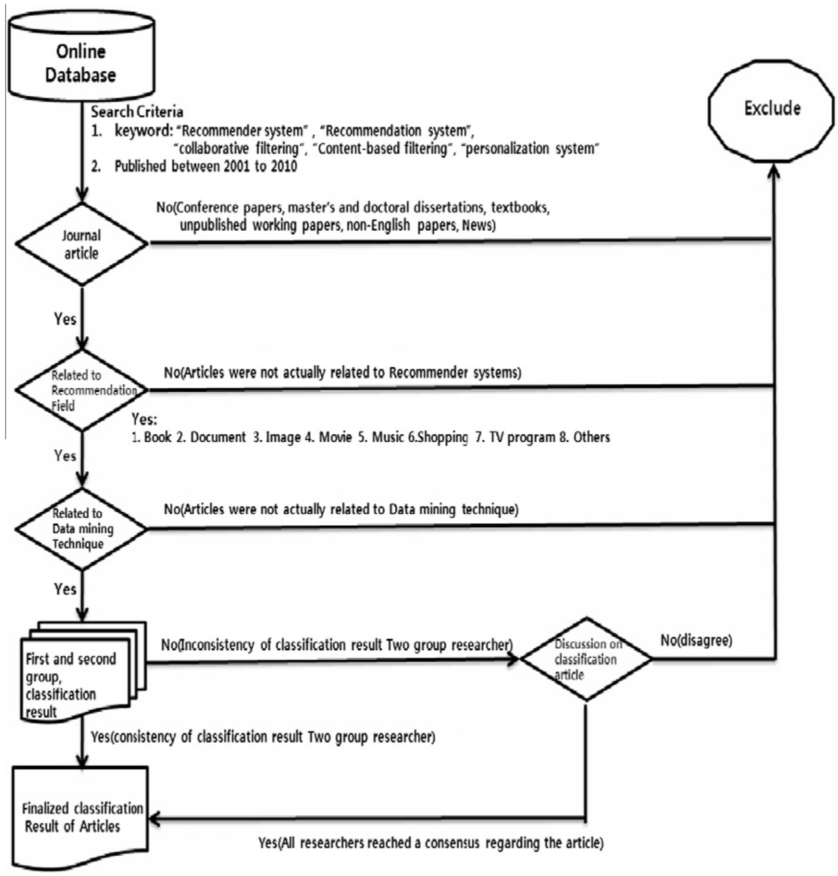

Fig. 2. Selection criteria and evaluation framework.

this research began to extend to other fields such as books, documents, music and other fields in 2007.

4.2. Distribution by journal

Research papers are selected from a total of 46 different journals. Distribution of research papers by journal is presented in Table 1. Expert Systems with Applications published more than 33% (70 out of 210 research papers, or 33.33%) of the total number of research papers. IEEE Intelligent System (21 out of 210 research papers, or 10.00%), along with, Decision Support Systems and ACM Transactions on Information Systems (12 out of 210 research papers, or 5.71%), published the second and third largest percentage of recommender systems-related research papers among the journals. The most research papers were published in Expert Systems with Applications, because this journal focuses on knowledge of the application of expert and intelligent system by industry, governments and universities worldwide (Ngai, Xiu, & Chau, 2009).

4.3. Distribution by application fields and data mining techniques

Distribution of research papers by application fields is represented in Fig. 4. The majority of the research papers were related to movie (53 out of 210 research papers, or 25.2%) and shopping (42 out of 210 research papers, or 20.0%). Because recommender systems in movie and shopping fields have a larger number of practical applications than other fields, it is inferred that although many research papers were published, few of them were related to image fields (7 out of 210 research papers, or 3.3%), and music, and TV program fields (9 out of 210 research papers, or 4.2% respectively). In particular, because the data of MovieLens (www.movielens.org/) are freely accessed, many recommendation methodologies have been proposed and evaluated with MovieLens data, which explains why there is more the recommender systems researches in movie fields than in other fields.

Distribution of research papers by application fields and journal is represented in Table 2. Among the application fields and journals, Expert Systems with Applications included most of the application fields. However, research papers about recommending music and TV programs were usually published in more specific journals. Because music and TV program related papers are usually published at the specific journals.

Distribution of research papers by data mining techniques is shown in Fig. 5, and distribution of the 210 research papers classified by the suggested classification framework is shown in Table 3. Among data mining techniques, the heuristic and k-NN (k-nearest

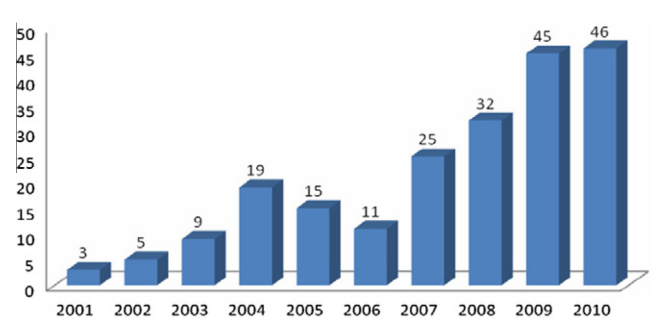

Fig. 3. Distribution of research papers by year of publication.

Table 1
Distribution of research papers by journal in which the research papers were published.

Journal title	Amount	Percentage (%)
Expert Systems with Applications	70	33.33
IEEE Intelligent Systems	21	10.00
ACM Transactions on Information Systems	12	5.71
Decision Support Systems	12	5.71
Knowledge-Based Systems	11	5.24
IEEE Internet Computing	9	4.29
IEEE Transactions on Consumer Electronics	9	4.29
International Journal of Electronic Commerce	7	3.33
Electronic Commerce Research & Applications	6	2.86
IEEE Transactions on Knowledge and Data Engineering	6	2.86
IEEE Transactions on Audio, Speech, and Language Processing	3	1.43
International Journal of Human Computer Studies	3	1.43
Journal of Systems & Software	3	1.43
Behavior & Information Technology	2	0.95
Computers in Human Behavior	2	0.95
Information Processing & Management	2	0.95
IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans	2	0.95
Management Science	2	0.95
ACM Transactions on Computer–Human Interaction	1	0.48
ACM Transactions on Knowledge Discovery from Data	1	0.48
AI Magazine	1	0.48
Communications of the ACM	1	0.48
Computer	1	0.48
Computer Supported Cooperative Work	1	0.48
Computers & Operations Research	1	0.48
Electron Markets	1	0.48
IEEE Circuits and Systems for Video Technology	1	0.48
IEEE Pervasive Computing	1	0.48
IEEE Security & Privacy	1	0.48
IEEE Software	1	0.48
IEEE Spectrum	1	0.48
IEEE Transactions on Fuzzy Systems	1	0.48
IEEE Transactions on Information Forensics and Security	1	0.48
IEEE Transactions on Multimedia	1	0.48
IEEE Transactions on Pattern Analysis and Machine Intelligence	1	0.48
IEEE Transactions on Services Computing	1	0.48
IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews	1	0.48
Information & Management	1	0.48
Information Systems	1	0.48
International Journal of Information Management	1	0.48
International Journal of Technology Management	1	0.48
IT Professional	1	0.48
Journal of Computer Information Systems	1	0.48
Journal of Software Maintenance	1	0.48
Journal of Management Information Systems	1	0.48
Journal of Information Science	1	0.48
Total	210	100.00

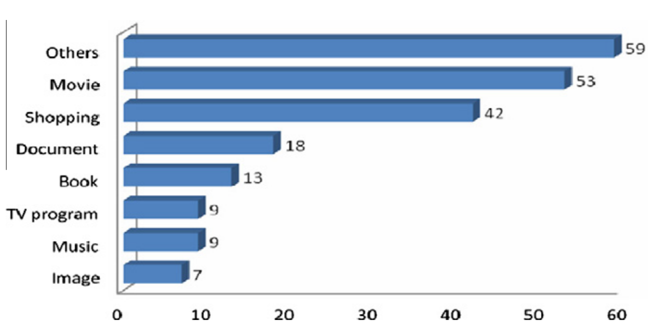

Fig. 4. Distribution of research papers by application fields.

neighbor) models have been used the most often in application fields. Because, the heuristic model is not one method but instead involves adding on new methods to existing diverse methods, it is used to expand advanced research. Also, the CF system is one of the most successful methodologies in recommender systems, and k-NN is a popular type of CF, so k-NN has been applied in most of the application fields.

4.4. Distribution of research papers by publication years and application fields

Distribution of research papers by publication years and application fields is shown in Fig. 6, which shows decreases in most of the application fields during 2006. Until 2006, most recommender systems research was focused on movies and shopping fields. However, the focus of recommender systems research has extended not only to movie and shopping fields, but also to books, documents, music, and other fields beginning in 2007.

4.5. Distribution of research papers by publication years and data mining techniques

Distribution of research papers by publication years and data mining techniques is shown in Fig. 7. Among the data mining techniques, most of the techniques are decreased in 2006, except that the heuristic method increased steadily and reached a peak in 2010. Because the heuristic method is not only one method, but rather involves diverse methods that are not included in other server data mining techniques, its usage has increased annually.

Table 2
Distribution of research papers by recommendation field and journals.

Field	Journal	Amount	
Book	ACM Transactions on Information Systems	2	
	Decision Support Systems	2	
	Electronic Commerce Research & Applications	2	
	IEEE Internet Computing	2	
	Computers in Human Behavior	1	
	Expert Systems with Applications	1	
	International Journal of Information Management	1	
	Knowledge-Based Systems	1	
	Management Science	1	
			13
Document	Expert Systems with Applications	5	
	IEEE Intelligent Systems	3	
	ACM Transactions on Information Systems	2	
	Decision Support Systems	2	
	IEEE Internet Computing	1	
	IEEE Transactions on Information Forensics and Security	1	
	Journal of Computer Information Systems	1	
	Journal of Systems & Software	1	
	Knowledge-Based Systems	1	
	International Journal of Human Computer Studies	1	
			18
Image	Expert Systems with Applications	4	
	Journal of Information Science	1	
	IEEE Intelligent Systems	1	
	IEEE Transactions on Multimedia,	1	
			7
Movie	Expert Systems with Applications	21	
	ACM Transactions on Information Systems	6	
	Knowledge-Based Systems	5	
	International Journal of Electronic Commerce	4	
	IEEE Intelligent Systems	3	
	Electronic Commerce Research & Applications	2	
	IEEE Internet Computing	2	
	IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans	2	
	ACM Transactions on Knowledge Discovery from Data	1	
	Behavior & Information Technology	1	
	Communications of the ACM	1	
	Computer	1	
	Decision Support Systems	1	
	IEEE Circuits and Systems for Video Technology	1	
	IEEE Transactions on Knowledge and Data Engineering	1	
	Information Processing & Management	1	
			53
Music	IEEE Transactions on Audio, Speech, and Language Processing	3	
	Expert Systems with Applications	2	
	ACM Transactions on Information Systems	1	
	IEEE Intelligent Systems	1	
	IEEE Transactions on Consumer Electronics	1	
	Information Processing & Management	1	
			9
Others	Expert Systems with Applications	22	
	IEEE Intelligent Systems	8	
	IEEE Transactions on Knowledge and Data Engineering	5	
	Decision Support Systems	4	
	IEEE Internet Computing	3	
	IEEE Transactions on Consumer Electronics	3	
	International Journal of Electronic Commerce	2	
	Computer Supported Cooperative Work	1	
	Electron Markets	1	
	IEEE Pervasive Computing	1	
	IEEE Security & Privacy	1	
	IEEE Software	1	
	IEEE Spectrum	1	
	IEEE Transactions on Fuzzy Systems	1	
	IEEE Transactions on Pattern Analysis and Machine Intelligence	1	
	IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews	1	
	IT Professional	1	
	Knowledge-Based Systems	1	
	Management Science	1	
			59
Shopping	Expert Systems with Applications	14	
	IEEE Intelligent Systems	5	

Table 2 (*continued*)

Field	Journal	Amount	
	Decision Support Systems	3	
	Electronic Commerce Research & Applications	2	
	International Journal of Human Computer Studies	2	
	Knowledge-Based Systems	2	
	ACM Transaction on Computer–Human Interaction	1	
	ACM Transactions on Information Systems	1	
	AI Magazine	1	
	Behavior & Information Technology	1	
	Computers & Operations Research	1	
	IEEE Transactions on Consumer Electronics	1	
	IEEE Transactions on Services Computing	1	
	Information & Management	1	
	Information Systems	1	
	International Journal of Electronic Commerce	1	
	International Journal of Technology Management	1	
	Journal of Software Maintenance	1	
	Journal of Systems & Software	1	
	Journal of Management Information Systems	1	
			42
TV program	IEEE Transactions on Consumer Electronics	4	
	Computers in Human Behavior	1	
	Expert Systems with Applications	1	
	IEEE Internet Computing	1	
	Journal of Systems & Software	1	
	Knowledge-Based Systems	1	
			9
Total			210

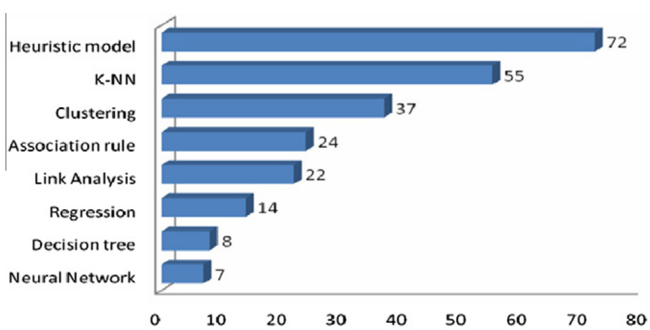

Fig. 5. Distribution of research papers by data mining techniques.

Based on their previous rates of change, more heuristic methods are expected to be used significantly in the future.

5. Conclusion, research implication and future work

Recommender systems have attracted the attention of academics and practitioners. In this research, we have identified 210 research papers on recommender systems, which were published between 2001 and 2010, to understand the trend of recommender systems-related research and to provide practitioners and researchers with insight and future direction on recommender systems. The results represented in this paper have several significant implications:

- Based on previous publication rates, interest in recommender systems related research will grow significantly in the future.
- Fifty-three research papers were related to movie recommendations, whereas image recommendations were identified in only seven research papers. Image field, and Music, and TV program recommendations were identified in nine research papers respectively. Therefore, more research is required to for image, music and TV program recommendations. This result was due to the easy use of the MovieLens data set. Therefore, it looks to be necessary to prepare data sets in other fields.
- Among the 210 research papers, 55 research papers used k-NN and 72 research papers have used heuristic models in the recommender system domain. k-NN creates applied user profile using the user's preference ratings obtained either directly from the user's explicit ratings of items or indirectly from the user's purchase or usage information. Therefore, it is not surprising that the k-NN method has been used in an extensive range of recommender systems domains. Also, because the heuristic model is not a single method, but one that consist of existing diverse methods, its use will be increased.
- Research papers using clustering and association rule techniques rank behind k-NN. From this, we know that both clustering and association rule techniques have been widely used in real business application than other techniques.
- Recently, social network analysis has been used in various applications. However studies on recommender systems using social network analysis are still deficient. Henceforth, we expect that new recommendation approaches using social network analysis will be developed. Therefore, developing the recommendation system research using social network analysis will be an interesting area further research.
- The number of heuristic methods is increasing every year. This result has been caused by the many researchers developing new methodologies and mixed technique model.
- Our research is significant because the majority of recommender systems research has been published in 125 MIS journals, such as ACM, IEEE publications. However, recommender systems research has shifted from the MIS field to various business fields, so we expect to see more recommender systems research published in management and business journals.

Our classification model will provide the practitioner and academic with guideline for future research on recommender systems. However our research has the following limitations: First,

Table 3
Distribution of research papers by application fields and journals.

Recommendation field	Data mining techniques	Reference
Book	Heuristic model	Riedl (2001)
	Clustering	Linden, Smith, and York (2003)
	k-NN	McSherry (2004)
	Link analysis	Huang, Chen, and Zeng (2004)
	Link analysis	Huang, Zeng, and Chen (2007a, 2007b)
	Link analysis	Ziegler and Golbec (2007)
	Regression	Hernández del Olmo and Gaudioso (2008)
	Clustering	Rosaci, Sarné, and Garruzzo (2009)
	k-NN, heuristic model	Kim, Kim, Oh, and Ryu (2010)
	Association rule, k-NN	Kim et al. (2010)
	Heuristic model, link analysis	Hwang, Wei, and Liao (2010)
	Heuristic model	Crespo et al. (2010)
Document	k-NN, neural network, regression	Lee, Hui, and Fong (2002)
	Association rule, clustering	Wang and Shao (2004)
	Heuristic model	Middleton, Shadbolt, and De Roure (2004)
	Clustering, neural network	Lihua et al. (2005)
	Heuristic model	Melamed, Shapira, and Elovici (2007)
	Link analysis	Liang, Yang, Chen, and Ku (2008)
	Heuristic model	Weng and Chang (2008)
	Clustering	Wei, Yang, and Hsiao (2008)
	k-NN, regression	Tang and McCalla (2009)
	Clustering	Lai and Liu (2009)
	Association rule, clustering, Link analysis	Göksedef and Gündüz-Öğüdücü (2010)
	Heuristic model	Champin, Briggs, Coyle and Smyth (2010)
	Heuristic model	Moens, De Beer, Boiy, and Gomez (2010)
	Clustering, heuristic model	Jalali, Mustapha, Sulaiman, and Mamat (2010)
	Link analysis	Dell'Amico and Capra (2010)
Image	Heuristic model	Kwon (2003)
	Heuristic model	Kim, Lee, Cho, and Kim (2004)
	Heuristic model	Boutemedjet and Ziou (2008)
	k-NN	Lee, Park, and Park (2008)
	k-NN, link analysis	Kim, Kim, and Cho (2008)
	k-NN	Lee, Park, and Park (2009)
	Heuristic model, k-NN	Nan Zheng, Li, Liao, and Zhang (2010)
Movie	k-NN	Naren, Benjamin, Batul, Ananth, and George (2001)
	Association rule	Herlocker and Konstan (2001)
	Association rule, decision tree, k-NN	Cheung, Kwok, Law, and Tsui (2003)
	Clustering, k-NN	Roh, Oh, and Han (2003)
	Clustering	Cheung, Tsui, and Liu (2004)
	k-NN	Han, Xie, Yang, and Shen (2004)
	Clustering, k-NN	Weng and Liu (2004)
	k-NN	Zeng, Xing, Zhou, and Zheng (2004)
	k-NN	Herlocker, Konstan, Terveen, and Riedl (2004)
	Link analysis	Miller, Konstan, and Riedl (2004)
	Clustering, k-NN	Min and Han (2005)
	k-NN	Li, Lu, and Xuefeng (2005)
	Clustering	Kim and Yum (2005)
	Regression	Lee, Jun, Lee, and Kim (2005)
	Heuristic model	Adomavicius, Sankaranarayanan, Sen, and Tuzhilin (2005)
	Heuristic model	Salter and Antonopoulos (2006)
	Association rule, k-NN	Du Boucher-Ryan and Bridge (2006)
	Heuristic model	Prangl, Szkaliczki, and Hellwagner (2007)
	k-NN	Hurley, O'Mahony and Silvestre (2007)
	Heuristic model	Im and Hars (2007)
	Clustering, k-NN	Symeonidis, Nanopoulos, and Manolopoulos (2008)
	k-NN	Symeonidis, Nanopoulos, Papadopoulos, and Manolopoulos (2008)
	k-NN	Chen, Cheng, and Chuang (2008)
	Association rule	Leung, Chan, and Chung (2008)
	Heuristic model	Russell and Yoon (2008)
	k-NN	Lee and Olafsson (2009)
	k-NN	Jeong, Lee, and Cho (2009a)
	k-NN	Jeong, Lee, and Cho (2009b)
	Clustering, k-NN	Merve and Arslan (2009)
	k-NN	Koren, Bell, and Volinsky (2009)
	k-NN	Chen, Wang, and Zhang (2009)
	Clustering	Kwon, Cho, and Park (2009)
	Heuristic model	Cho, Kwon, and Park (2009)
	Heuristic model	Yang and Li (2009)
	k-NN	Bobadilla, Serradilla, and Hernando (2009)
	Heuristic model	Julià, Sappa, Lumbreras, Serrat, and López (2009)
	Heuristic model	Koren (2010a)
	Heuristic model	Winoto and Tang (2010)
	Heuristic model	Ahn, Kang, and Lee (2010)

Table 3 (*continued*)

Recommendation field	Data mining techniques	Reference
	Heuristic model, link analysis, regression	Hwang (2010)
	k-NN	Bobadilla, Serradilla, and Bernal (2010)
	Regression	Ozok, Fan, and Norcio (2010)
	Heuristic model, k-NN	Koren (2010b)
Music	k-NN	Ganesan, Garcia-Molina, and Widom (2003)
	Clustering, regression	Zhu, Shi, Kim, and Eom (2006)
	Clustering	Li, Myaeng, and Kim (2007)
	Association rule, k-NN	Yoshii, Goto, Komatani, Ogata, and Okuno (2008)
	Link analysis	Shao, Ogihara, Wang, and Li (2009)
	Clustering, heuristic model	Su, Yeh, Yu, and Tseng (2010)
	Heuristic model	Nanopoulos, Rafailidis, Symeonidis, and Manolopoulos (2010)
	Clustering, neural network	Liu, Hsieh, and Tsai (2010)
Others	Heuristic model	Taab, Werther, Ricci, Zipf, and Gretzel (2002)
	Neural network	Yuan and Tsao (2003)
	Clustering	Chau, Zeng, Chen, Huang, and Hendriawan (2003)
	Heuristic model	Yang, Knoblock, and Wu (2004)
	Heuristic model	Adomavicius and Tuzhilin (2005)
	Heuristic model	Wei, Moreau, and Jennings (2005a)
	Clustering	Ha (2006)
	Heuristic model	McGinty and Smyth (2006)
	Heuristic model	Park, Kang, and Kim (2006)
	Regression	Gretzel and Fesenmaier (2006)
	Heuristic model	Alexander, Gerhard, and Lars (2007)
	Link analysis	Reichling, Veith, and Wulf (2007)
	Association rule	Adda, Valtchev, Missaoui, and Djeraba (2007)
	Clustering, neural network	Martín-Guerrero, Lisboa, Soria-Olivas, Palomares, and Balaguer (2007)
	k-NN, regression	Lee, Ahn, and Han (2007)
	Clustering	Lee and Park (2007)
	Heuristic model	Adomavicius and Kwon (2007)
	Heuristic model	Ricci and Nguyen (2007)
	Link analysis	Zeng, Wang, Zheng, Yuan, and Chen (2008)
	Heuristic model	Lin (2008)
	Heuristic model	Liang (2008)
	Heuristic model	Hernández del Olmo and Gaudioso (2008)
	Link analysis	Malinowski, Weitzel, and Keim (2008)
	Clustering	Linden (2008)
	Regression	Moon and Russell (2008)
	Association rule, k-NN	Hsu (2008)
	Link analysis	Wang and Chiu (2008)
	Decision tree , k-NN	Hernández del Olmo, Gaudioso, and Martin (2009)
	Heuristic model	Hsu (2009)
	Heuristic model	Schiaffino and Amandi (2009)
	Heuristic model	Porcel, López-Herrera, and Herrera-Viedma (2009a)
	Heuristic model	Zhen, Huang, and Jiang (2009a)
	Decision tree	Wang, Chiang, Hsu, Lin, and Lin (2009)
	Association rule	Yang and Wang (2009)
	Heuristic model	Porcel, Moreno, and Herrera-Viedma (2009b)
	Link analysis	Arazy, Kumar, and Shapira (2009)
	Heuristic model	Zhen, Huang, and Jiang (2009b)
	Heuristic model	Kim, Jeong, and Baik (2009)
	Heuristic model, neural network	Han and Chen (2009)
	Heuristic model	Lesk (2009)
	Association rule, Clustering, regression	Kwon and Kim (2009)
	Association rule, k-NN	Schiaffino and Amandi (2009)
	Link analysis	Li and Kao (2009)
	Link analysis	Kuo, Chen, and Liang (2009)
	Heuristic model	Symeonidis, Nanopoulos, and Manolopoulos (2010)
	Heuristic model	Pillonetto, Dinuzzo, and De Nicolao (2010)
	Heuristic model	Zhen, Huang, and Jiang (2010)
	Heuristic model	Jalali et al. (2010)
	Heuristic model	Porcel and Herrera-Viedma (2010)
	Heuristic model	Zhan et al. (2010)
	Heuristic model, k-NN	Munoz-Organero, Ramíez-González, Muñoz-Merino, and Kloos (2010)
	Heuristic model, k-NN	Blanco-Fernandez, Lopez-Nores, Pazos-Arias, Gil-Solla, and Ramos-Cabrer (2010)
	Heuristic model	Yager, Reformat, and Gumrah (2010)
	Heuristic model	Bergamaschi, Guerra, and Leiba (2010)
	Heuristic model	Backhaus et al. (2010)
	Link analysis, regression	Kato, Kashima, Sugiyama, and Asai (2010)
Shopping	Association rule, decision tree	Kim et al. (2002)
	Association rule, decision tree	Cho, Kim & Kim (2002)
	Association rule, clustering	Ha (2002)
	k-NN	Vezina and Militaru (2004)
	Regression	Ant Ozok, Quyin, and Norcio (2004)
	Association rule, k-NN	Wang, Chuang, Hsu, and Keh (2004)

(*continued on next page*)

Table 3 (*continued*)

Recommendation field	Data mining techniques	Reference
	k-NN	Cho and Kim (2004)
	Association rule, k-NN	Liu and Shih (2005a)
	Association rule, k-NN	Liu and Shih (2005b)
	Association rule, clustering	Cho, Cho & Kim (2005)
	k-NN, regression	Kim, Yum, Song, and Kim (2005)
	Decision tree	Yu, Ou, Zhang, and Zhang (2005)
	Heuristic model	Wei, Moreau, and Jennings (2005b)
	Clustering	Choi, Kang, and Jeon (2006)
	Heuristic model	Garfinkel, Gopal, Tripathi, and Yin (2006)
	k-NN	Zanker, Jannach, Gordea, and Jessenitschnig (2007)
	Association rule	Zhang and Jiao (2007)
	Association rule	Pu and Chen (2007)
	Clustering, link analysis	Wang, Dai, and Yuan (2008b)
	Clustering	Kim and Ahn (2008)
	Association rule, k-NN	Wang and Wu (2009)
	k-NN	Albadvi and Shahbazi (2009)
	Heuristic model	Pu and Chen (2009)
	k-NN	Kim et al. (2009)
	Association rule, k-NN	Robillard and Dagenais (2009)
	Heuristic model	Moosavi, Nematbakhsh, and Farsani (2009)
	Heuristic model	Martin-Vicente, Gil-Solla, Ramos-Cabrer, Blanco-Fernandez, and Lopez-Nores (2010)
	Heuristic model	Ochi, Rao, Takayama and Nass (2010)
	Heuristic model	Funk, Rozinat, Karapanos, Alves de Medeiros, and Koca (2010)
	Link analysis	Yuan, Guan, Lee, Lee, and Hur (2010)
	Heuristic model	Taha and Elmasri (2010)
	Heuristic model, k-NN	Wang and Wu (2010)
	Heuristic model	Pathak, Garfinkel, Gopal, Venkatesan, and Yin (2010)
	Association rule, heuristic model	Chen and Pu (2010)
TV program	Decision tree	Lee and Yang (2003)
	Heuristic model, link analysis	Blanco-Fernandez, Pazos-arias, Gil-Solla, Ramos-Cabrer, and Lopez-Nores (2008)
	Heuristic model, k-NN	Martinez et al. (2010)
	Heuristic model, k-NN	Martin-Vicente et al. (2010)
	Clustering, heuristic model	Cantador and Castells (2010)

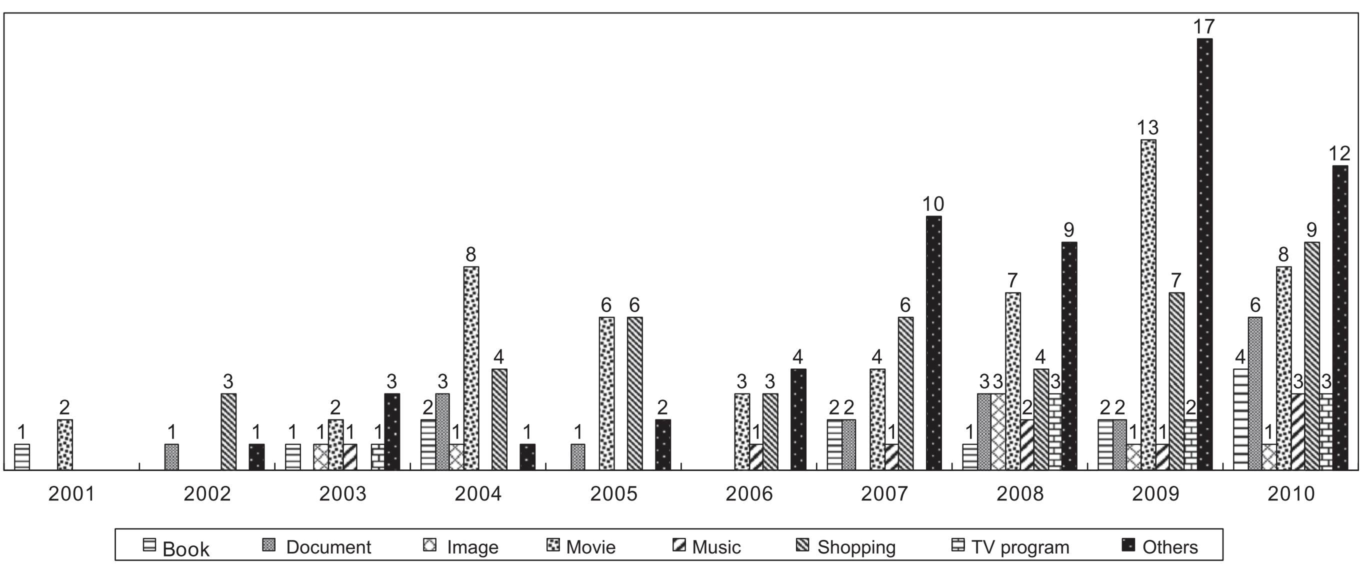

Fig. 6. Distribution of research papers by publication year and application fields.

due to the limitations of time and manpower, we only surveyed research papers published between 2001 and 2010, and our searches were based on the top 125 MIS. Therefore, if the research had been extended to cover other journals such as those focused on computer science and, marketing, the results might have been different. Second, our findings are based on articles that were selected solely from academic journals. If articles from conferences had been included, the results would have been more diverse. Third, our study was conducted based on a search of the following keywords: "Recommender system", "Recommendation system", "Personalization system", "Collaborative filtering", and "Contents filtering". Besides these five keywords, we did not search additional keywords, such as "Hybrid Filtering". Research papers that referred to recommender systems, but did not include any of the five key-words, could not be extracted. We think that recommender systems research also has been published in other lan-

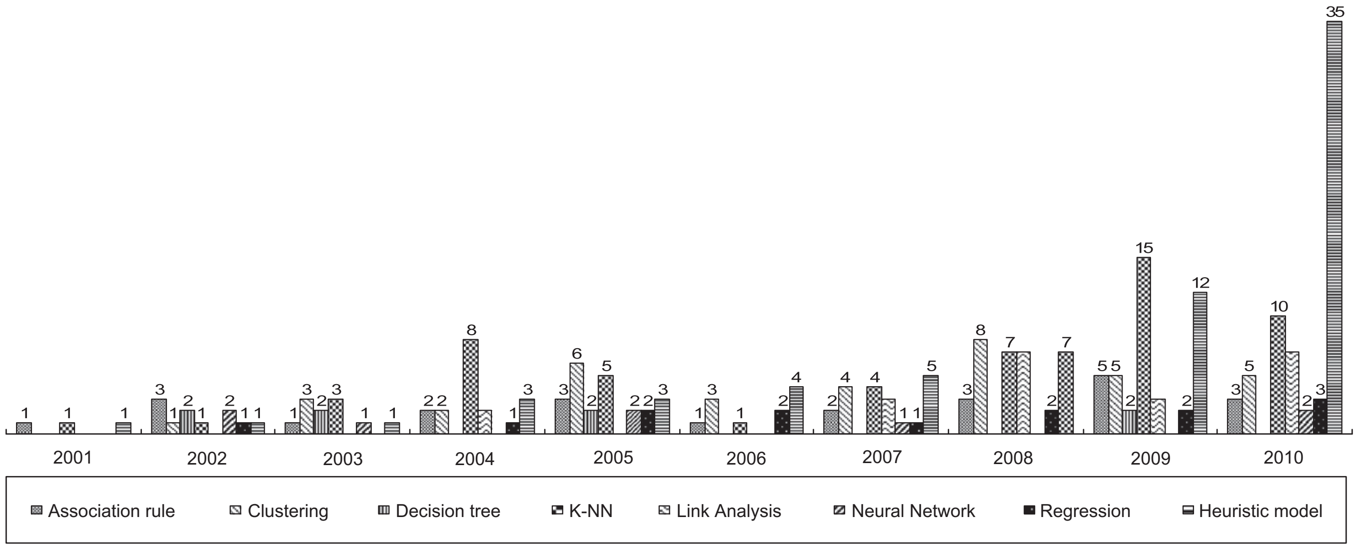

Fig. 7. Distribution of research papers by publication year and data mining technique.

guages. Finally, we classified data mining techniques, but not data mining model.

Accordingly, we will continue to classify articles on an ongoing basis. Moreover, it is also necessary to include conference papers and non-English papers in order to extend our classification model.

Acknowledgement

This research is supported by Ministry of Culture, Sports and Tourism(MCST) and Korea Creative Content Agency(KOCCA) in the Culture Technology(CT) Research & Developement Program 2011.

Reference

Adda, M., Valtchev, P., Missaoui, R., & Djeraba, C. (2007). Toward recommendation based on ontology-powered web-usage mining. *IEEE Internet Computing, 11*, 45–52.

Adomavicius, G., & Kwon, Y. O. (2007). New recommendation techniques for multicriteria rating systems. *IEEE Intelligent Systems, 22*, 48–55.

Adomavicius, G., Sankaranarayanan, R., Sen, S., & Tuzhilin, A. (2005). Incorporating contextual information in recommender systems using a multidimensional approach. *ACM Transactions on Information Systems, 23*, 103–145.

Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering, 6*, 734–749.

Ahn, H. J., Kang, H. J., & Lee, J. P. (2010). Selecting a small number of products for effective user profiling in collaborative filtering. *Expert Systems with Applications, 37*, 3055–3062.

Albadvi, A., & Shahbazi, M. (2009). A hybrid recommendation technique based on product category attributes. *Expert Systems with Applications, 36*, 11480–11488.

Alexander, F., Gerhard, F., & Lars, S. T. (2007). Guest editors' introduction: Recommender systems. *IEEE Intelligent Systems, 22*, 18–21.

Anders, U., & Korn, O. (1999). Model Selection in Neural Networks. *Neural Networks, 12*, 309–323.

Ant Ozok, A., Quyin Fan & Norcio, Anthony F. (2004). Design guidelines for effective recommender system interfaces based on a usability criteria conceptual model: results from a college student population. *Behaviour & Information Technology*.

Arazy, O., Kumar, N., & Shapira, B. (2009). Improving social recommender Systems. *IT Professional, 11*, 38–44.

Backhaus, K., Frohs, M., Weddeling, M., Steiner, M., Becker, J., & Beverungen, D. (2010). Enabling individualized recommendations and dynamic pricing of value-added services through willingness-to-pay data. *Electron Markets, 20*, 131–146.

Basu, C., Hirsh, H., & Cohen, W. (1998). Recommendation as classification : Using social and content-based information in recommendation, *In Proceedings of the 15th National Conference on, Artificial Intelligence*, 714–720.

Bergamaschi, S., Guerra, F., & Leiba, B. (2010). Guest editors' introduction: Information overload. *IEEE Internet Computing, 14*, 10–13.

Berry, M. J. A., & Linoff, J. S. (2004). *Data Mining Techniques for Marketing. Sales and Customer Relationship Management* (2nd ed.,). Wiley.

Blanco-Fernandez, Y., Pazos-arias, J. J., Gil-Solla, A., Ramos-Cabrer, M., & Lopez-Nores, M. (2008). Providing entertainment by content-based filtering and semantic reasoning in intelligent recommender systems. *IEEE Transactions on Consumer Electronics, 54*, 727–735.

Blanco-Fernandez, Y., Lopez-Nores, M., Pazos-Arias, J. J., Gil-Solla, A., & Ramos-Cabrer, M. (2010). Exploiting digital TV users' preferences in a tourism recommender system based on semantic reasoning. *IEEE Transactions on Consumer Electronics, 56*, 904–912.

Bobadilla, J., Serradilla, F., & Hernando, A. (2009). Collaborative filtering adapted to recommender systems of e-learning. *Knowledge-Based Systems, 22*, 261–265.

Bobadilla, J., Serradilla, F., & Bernal, J. (2010). A new collaborative filtering metric that improves the behavior of recommender systems. *Knowledge-Based Systems, 23*, 520–528.

Boutemedjet, S., & Ziou, D. (2008). A Graphical model for context-aware visual content recommendation. *IEEE Transactions on Multimedia, 10*, 52–62.

Cai, D., He, X., Wen, J.R., & Ma, W.Y. (2004). Block-level link analysis, *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in, information retrieval*, 440–447.

Champin, P. A., Briggs, P., Coyle, M., & Smyth, B. (2010). Coping with noisy search experiences. *Knowledge-Based Systems, 23*, 287–294.

Cantador, I., & Castells, P. (2010). Extracting multilayered communities of interest from semantic user profiles: Application to group modeling and hybrid recommendations. *Computers in Human Behavior, 27*, 1321–1336.

Chau, M., Zeng, D., Chen, H., Huang, M., & Hendriawan, D. (2003). Design and evaluation of a multi-agent collaborative Web mining system. *Decision Support Systems, 25*, 167–183.

Chen, G., Wang, F., & Zhang, C. (2009). Collaborative filtering using orthogonal nonnegative matrix tri-factorization. *Information Processing & Management, 45*, 368–379.

Chen, L., & Pu, P. (2010). Experiments on the preference-based organization interface in recommender systems. *ACM Transaction on Computer-Human Interaction, 17*, 1–33.

Chen, Y. L., Cheng, L. C., & Chuang, C. N. (2008). A group recommendation system with consideration of interactions among group members. *Expert Systems with Applications, 34*, 2082–2090.

Cheung, K. W., Kwok, J. T., Law, M. H., & Tsui, K. C. (2003). Mining customer product ratings for personalized marketing. *Decision Support Systems, 35*, 231–243.

Cheung, K. W., Tsui, K. C., & Liu, J. (2004). Extended latent class models for collaborative recommendation. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, 34*, 143–148.

Cho, J. H., Kwon, K. S., & Park, Y. T. (2009). Q-rater: A collaborative reputation system based on source credibility theory. *Expert Systems with Applications, 36*, 3751–3760.

Cho, Y. B., Cho, Y. H., & Kim, S. H. (2005). Mining changes in customer buying behavior for collaborative recommendations. *Expert Systems with Applications, 28*, 359–369.

Cho, Y. H., Kim, J. K., & Kim, S. H. (2002). A personalized recommender system based on web usage mining and decision tree induction. *Expert Systems with Applications, 23*, 329–342.

Cho, Y. H., & Kim, J. K. (2004). Application of web usage mining and product taxonomy to collaborative recommendations in e-commerce. *Expert Systems with Applications, 26*, 233–246.

Choi, S. H., Kang, S. M., & Jeon, Y. J. (2006). Personalized recommendation system based on product specification values. *Expert Systems with Applications, 31*, 607–616.

Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D., & Sartin, M. (1999). Combining content-based and collaborative filters in an online newspaper, *Proceedings of the ACM SIGIR'99 Workshop on Recommender Systems*.

Crespo, R. G., Martínez, O. S., Lovelle, J. M. C., García-Bustelo, B. C. P., Gayo, J. E. L., & de Pablos, P. O. (2010). Recommendation system based on user interaction data applied to intelligent electronic books. *Computers in Human Behavior, 27*, 1445–1449.

Dell'Amico, M., & Capra, L. (2010). Dependable filtering: Philosophy and realizations. *ACM Transactions on Information Systems, 29*, 1–37.

Du Boucher-Ryan, P., & Bridge, D. (2006). Collaborative recommending using formal concept analysis. *Knowledge-Based Systems, 19*, 309–315.

Frias-Martinez, E., Chen, S. Y., & Liu, X. (2009). Evaluation of a personalized digital library based on cognitive styles: Adaptivity vs. adaptability. *International Journal of Information Management, 29*, 48–56.

Frias-Martinez, E., Magoulas, G., Chen, S. Y., & Macredie, R. (2006). Automated user modeling for personalized digital libraries. *International Journal of Information Management, 26*, 234–248.

Funk, M., Rozinat, A., Karapanos, E., Alves de Medeiros, A. K., & Koca, A. (2010). In situ evaluation of recommender systems: Framework and instrumentation. *International Journal of Human – Computer Studies, 68*, 525–547.

Ganesan, P., Garcia-Molina, H., & Widom, J. (2003). Exploiting hierarchical domain structure to compute similarity. *ACM Transactions on Information Systems, 21*, 64–93.

Garfinkel, R., Gopal, R., Tripathi, A., & Yin, F. (2006). Design of a shopbot and recommender system for bundle purchases. *Decision Support Systems, 42*, 1974–1986.

Göksedef, M., & Gündüz-Öğüdücü, S. (2010). Combination of web page recommender systems. *Expert Systems with Applications, 37*, 2911–2922.

Gretzel, U., & Fesenmaier, D. R. (2006). Persuasion in recommender systems. *International Journal of Electronic Commerce, 11*, 81–100.

Ha, S. H. (2002). Helping online customers decide through web personalization. *IEEE Intelligent Systems, 17*, 34–43.

Ha, S. H. (2006). Digital content recommender on the Internet. *IEEE Intelligent Systems, 2*, 70–77.

Han, L., & Chen, G. (2009). HQE: A hybrid method for query expansion. *Expert Systems with Applications, 36*, 7985–7991.

Han, P., Xie, B., Yang, F., & Shen, R. (2004). A scalable P2P recommender system based on distributed collaborative filtering. *Expert Systems with Applications, 27*, 203–210.

Herlocker, J. L., & Konstan, J. A. (2001). Content-independent task-focused recommendation. *IEEE Internet Computing, 5*, 40–47.

Herlocker, J. L., Konstan, J. A., Terveen, L. G., & Riedl, J. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems, 22*, 5–53.

Hernández del Olmo, F., & Gaudioso, E. (2008). Evaluation of recommender systems: A new approach. *Expert Systems with Applications, 35*, 790–804.

Hernández del Olmo, F., Gaudioso, E., & Martin, E. H. (2009). The task of guiding in adaptive recommender systems. *Expert Systems with Applications, 36*, 1972–1977.

Hsu, I. C. (2009). SXRS: An XLink-based recommender system using semantic web technologies. *Expert Systems with Applications, 36*, 3795–3804.

Hsu, M. H. (2008). A personalized English learning recommender system for ESL students. *Expert Systems with Applications, 34*, 683–688.

Huang, Z., Chen, H., & Zeng, D. (2004). Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. *ACM Transactions on Information Systems, 22*, 116–142.

Huang, Z., Zeng, D., & Chen, H. (2007a). A comparison of collaborative-filtering recommendation algorithms for e-commerce. *IEEE Intelligent Systems, 22*, 68–78.

Huang, Z., Zeng, D. D., & Chen, H. (2007b). Analyzing consumer-product graphs: Empirical findings and applications in recommender systems. *Management Science, 53*, 1146–1164.

Hurley, N. J., O'Mahony, M. P., & Silvestre, G. C. M. (2007). Attacking recommender systems: a cost-benefit analysis. *IEEE Intelligent Systems, 22*, 64–68.

Hwang, S. H., Wei, C. P., & Liao, Y. F. (2010). Coauthorship networks and academic literature recommendation. *Electronic Commerce Research and Applications, 9*, 323–334.

Hwang, S. L. (2010). Designing utility-based recommender systems for e-commerce: Evaluation of preference-elicitation methods. *Electronic Commerce Research and Applications*.

Ibnkahla, M. (2000). Applications of neural networks to digital communications-a survey. *Expert Systems with Applications, 80*, 1185–1215.

Im, I., & Hars, A. (2007). Does a one-size recommendation system fit all? the effectiveness of collaborative filtering based recommendation systems across different domains and search modes. *ACM Transactions on Information Systems, 26*, 1–30.

Jalali, M., Mustapha, N., Sulaiman, M., & Mamat, A. (2010). Corrigendum to "WebPUM: A web-based recommendation system to predict user future movements. *Expert Systems with Applications, 37*, 6201–6212.

Jeong, B., Lee, J. W., & Cho, H. B. (2009a). User credit-based collaborative filtering. *Expert Systems with Applications, 36*, 7309–7312.

Jeong, B., Lee, J. W., & Cho, H. B. (2009b). An iterative semi-explicit rating method for building collaborative recommender systems. *Expert Systems with Applications, 36*, 6181–6186.

Julià, C., Sappa, A. D., Lumbreras, F., Serrat, J., & López, A. (2009). Predicting missing ratings in recommender systems: Adapted factorization approach. *International Journal of Electronic Commerce, 14*, 89–108.

Kato, T., Kashima, H., Sugiyama, M., & Asai, K. (2010). Conic programming for multitask learning. *IEEE Transactions on Knowledge and Data Engineering*.

Kim, C. Y., Lee, J. K., Cho, Y. H., & Kim, D. H. (2004). VISCORS: a visual-content recommender for the mobile web. *IEEE Intelligent Systems, 19*, 32–39.

Kim, D. H., & Yum, B. J. (2005). Collaborative filtering based on iterative principal component analysis. *Expert Systems with Applications, 28*, 823–830.

Kim, H. K., Kim, J. K., & Ryu, Y. U. (2009). Personalized recommendation over a customer network for ubiquitous shopping. *IEEE Transactions on Services Computing, 2*, 140–151.

Kim, H. N., Ji, A. T., Ha, I., & Jo, J. S. (2010). Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation. *Electronic Commerce Research and Applications, 9*, 73–83.

Kim, J. H., Jeong, D. W., & Baik, D. K. (2009). Ontology-based semantic recommendation system in home network environment. *IEEE Transactions on Consumer Electronics, 55*, 1178–1184.

Kim, J. K., Cho, Y. H., Kim, W. J., Kim, J. R., & Suh, J. H. (2002). A personalized recommendation procedure for internet shopping support. *Electronic Commerce Research and Applications, 1*, 301–313.

Kim, J. K., Kim, H. K., & Cho, Y. H. (2008). A user-oriented contents recommendation system in peer-to-peer architecture. *Expert Systems with Applications, 34*, 300–312.

Kim, J. K., Kim, H. K., Oh, H. Y., & Ryu, Y. U. (2010). A group recommendation system for online communities. *International Journal of Information Management, 30*, 212–219.

Kim, K. J., & Ahn, H. C. (2008). A recommender system using GA K-means clustering in an online shopping market. *Expert Systems with Applications, 34*, 1200–1209.

Kim, Y. S., Yum, B. J., Song, J. H., & Kim, S. M. (2005). Development of a recommender system based on navigational and behavioral patterns of customers in e-commerce sites. *Expert Systems with Applications, 28*, 381–393.

Koren, Y. (2010a). Collaborative filtering with temporal dynamics. *Communications of the ACM, 53*.

Koren, Y. (2010b). Factor in the neighbors: Scalable and accurate collaborative filtering. *ACM Transactions on Knowledge Discovery from Data, 4*, 1–24.

Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer, 42*, 30–37.

Kuo, M. H., Chen, L. C., & Liang, C. W. (2009). Building and evaluating a location-based service recommendation system with a preference adjustment mechanism. *Expert Systems with Applications, 36*, 3543–3554.

Kwon, K. S., Cho, J. H., & Park, Y. T. (2009). Multidimensional credibility model for neighbor selection in collaborative recommendation. *Expert Systems with Applications, 36*, 7114–7122.

Kwon, O. B. (2003). "I know what you need to buy": context-aware multimedia-based recommendation system. *Expert Systems with Applications, 25*, 387–400.

Kwon, O. B., & Kim, J. H. (2009). Concept lattices for visualizing and generating user profiles for context-aware service recommendations. *Expert Systems with Applications, 36*, 1893–1902.

Lai, C. H., & Liu, D. R. (2009). Integrating knowledge flow mining and collaborative filtering to support document recommendation. *Journal of Systems and Software, 82*, 2023–2037.

Lee, H. J., & Park, S. J. (2007). MONERS: A news recommender for the mobile web. *Expert Systems with Applications, 32*, 143–150.

Lee, H. Y., Ahn, H. C., & Han, I. G. (2007). VCR: Virtual community recommender using the technology acceptance model and the user's needs type. *Expert Systems with Applications, 33*, 984–995.

Lee, J. S., Jun, C. H., Lee, J. W., & Kim, S. Y. (2005). Classification-based collaborative filtering using market basket data. *Expert Systems with Applications, 29*, 700–704.

Lee, J. S., & Olafsson, S. (2009). Two-way cooperative prediction for collaborative filtering recommendations. *Expert Systems with Applications, 36*, 5353–5361.

Lee, P. Y., Hui, S. C., & Fong, A. C. M. (2002). Neural networks for web content filtering. *IEEE Intelligent Systems, 17*, 48–57.

Lee, T. Q., Park, Y., & Park, Y. T. (2008). A time-based approach to effective recommender systems using implicit feedback. *Expert Systems with Applications, 34*, 3055–3062.

Lee, T. Q., Park, Y., & Park, Y. T. (2009). An empirical study on effectiveness of temporal information as implicit ratings. *Expert Systems with Applications, 36*, 1315–1321.

Lee, W. P., & Yang, T. H. (2003). Personalizing information appliances: a multi-agent framework for TV program recommendations. *Expert Systems with Applications, 25*, 331–341.

Lesk, M. (2009). Reading over your shoulder. *IEEE Security & Privacy, 7*, 78–81.

Leung, C. W., Chan, S. C., & Chung, F. (2008). An empirical study of a cross-level association rule mining approach to cold-start recommendations. *Knowledge-Based Systems, 21*, 515–529.

Li, Q., Myaeng, S. H., & Kim, B. M. (2007). A probabilistic music recommender considering user opinions and audio features. *Information Processing & Management, 43*, 473–487.

Li, Y., Lu, L., & Xuefeng, L. (2005). A hybrid collaborative filtering method for multiple-interests and multiple-content recommendation in E-Commerce. *Expert Systems with Applications, 28*, 67–77.

Li, Y. M., & Kao, C. P. (2009). TREPPS: A Trust-based recommender system for peer production services. *Expert Systems with Applications, 36*, 3263–3277.

Liang, T. P. (2008). Recommendation systems for decision support: An editorial introduction. *Decision Support Systems, 45*, 385–386.

Liang, T. P., Yang, Y. F., Chen, D. N., & Ku, Y. C. (2008). A semantic-expansion approach to personalized knowledge recommendation. *Decision Support Systems, 45*, 401–412.

Lihua, W., Lu, L., Jing, L., & Zongyong, L. (2005). Modeling user multiple interests by an improved GCS approach. *Expert Systems with Applications, 29*, 757–767.

Lin, K. J. (2008). E-commerce technology: back to a prominent future. *IEEE Internet Computing, 12*, 60–65.

Linden, G. (2008). People who read this article also read.... *IEEE Spectrum, 5*.

Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing, 7*, 76–80.

Liu, N. H., Hsieh, S. J., & Tsai, C. F. (2010). An intelligent music playlist generator based on the time parameter with artificial neural networks. *Expert Systems with Applications, 37*, 2815–2825.

Liu, D. R., & Shih, Y. Y. (2005a). Integrating AHP and data mining for product recommendation based on customer lifetime value. *Information & Management, 42*, 387–400.

Liu, D. R., & Shih, Y. Y. (2005b). Hybrid approaches to product recommendation based on customer lifetime value and purchase preferences. *Journal of Systems & Software, 77*, 181–191.

Lopez-Nores, M., Garca-Duque, J., Frenandez-Vilas, R. P., & Bermejo-Munoz, J. (2008). A flexible semantic inference methodology to reason about user preference in knowledge-based recommender systems. *Knowledge-Based systems, 21*, 305–320.

Malhotra, N. K. (2007). *Marketing research: An applied orientation* (5th ed.). Pearson Education Inc.

Malinowski, J., Weitzel, T., & Keim, T. (2008). Decision support for team staffing: An automated relational recommendation approach. *Decision Support Systems, 45*, 429–447.

Martinez, A. B. B., Lopez, M. R., Montenegro, E. C., Fonte, F. A. M., Burguillo, J. C., & Peleterio, A. (2010). Exploiting social tagging in a Web 2.0 recommender system. *IEEE Internet Computing, 14*, 23–30.

Martín-Guerrero, J. D., Lisboa, P. J. G., Soria-Olivas, E., Palomares, A., & Balaguer, E. (2007). An approach based on the Adaptive Resonance Theory for analysing the viability of recommender systems in a citizen Web portal. *Expert Systems with Applications, 33*, 743–753.

Martin-Vicente, M. I., Gil-Solla, A., Ramos-Cabrer, M., Blanco-Fernandez, Y., & Lopez-Nores, M. (2010). A semantic approach to avoiding fake neighborhoods in collaborative recommendation of coupons through digital TV. *IEEE Transactions on Consumer Electronics, 56*, 54–62.

McSherry, D. (2004). Balancing user satisfaction and cognitive load in coverage-optimised retrieval. *Knowledge-Based Systems, 17*, 113–119.

McGinty, L., & Smyth, B. (2006). Adaptive selection: An analysis of critiquing and preference-based feedback in conversational recommender systems. *International Journal of Electronic Commerce, 11*, 35–57.

Melamed, D., Shapira, B., & Elovici, Y. (2007). MarCol: A market-based recommender system. *IEEE Intelligent Systems, 22*, 74–78.

Merve, Acilar. A., & Arslan, A. (2009). A collaborative filtering method based on artificial immune network. *Expert Systems with Applications, 36*, 8324–8332.

Middleton, S. E., Shadbolt, N. R., & De Roure, D. C. (2004). Ontological user profiling in recommender systems. *ACM Transactions on Information Systems, 22*, 54–88.

Miller, B. N., Konstan, J. A., & Riedl, J. (2004). PocketLens: Toward a personal recommender system. *ACM Transactions on Information Systems, 22*, 437–476.

Min, S. H., & Han, I. G. (2005). Detection of the customer time-variant pattern for improving recommender systems. *Expert Systems with Applications, 28*, 189–199.

Moens, M. F., De Beer, J., Boiy, E., & Gomez, J. C. (2010). Identifying and resolving hidden Text salting. *IEEE Transactions on Information Forensics and Security, 5*, 837–847.

Moon, S. K., & Russell, G. J. (2008). Predicting product purchase from inferred customer similarity: An autologistic model approach. *Management Science, 54*, 71–82.

Moosavi, S., Nematbakhsh, M., & Farsani, H. K. (2009). A semantic complement to enhance electronic market. *Expert Systems with Applications, 36*, 5768–5774.

Munoz-Organero, M., Ramíez-González, G. A., Muñoz-Merino, P. J., & Kloos, C. D. (2010). A collaborative recommender system based on space-time similarities. *IEEE Pervasive Computing, 9*, 81–87.

Nanopoulos, A., Rafailidis, D., Symeonidis, P., & Manolopoulos, Y. (2010). MusicBox: personalized music recommendation based on cubic Analysis of social tags. *IEEE Transactions on Audio, Speech, and Language Processing, 18*, 407–412.

Naren, R., Benjamin, J. K., Batul, J. M., Ananth, Y. G., & George, K. (2001). Privacy risks in recommender systems. *IEEE Internet Computing, 5*, 54–62.

Ngai, E. W. T., Xiu, L., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications, 36*, 2592–2602.

Ochi, P., Rao, S., Takayama, L., & Nass, C. (2010). Predictors of user perceptions of web recommender systems: How the basis for generating experience and search product recommendations affect user response. *International Journal of Human – Computer Studies, 68*, 472–482.

Ozok, A. A., Fan, Q., & Norcio, A. F. (2010). Design guidelines for effective recommender system interfaces based on a usability criteria conceptual model: results from a college. *Behaviour & Information Technology, 29*, 57–83.

Park, S. J., Kang, S. G., & Kim, Y. K. (2006). A channel recommendation system in mobile environment. *IEEE Transactions on Consumer Electronics, 52*, 33–39.

Pathak, B., Garfinkel, R., Gopal, R. D., Venkatesan, R., & Yin, F. (2010). Empirical analysis of the impact of recommender systems on sales. *Journal of Management Information Systems, 27*, 159–188.

Pillonetto, G., Dinuzzo, F., & De Nicolao, G. (2010). Bayesian online multitask learning of gaussian processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 32*, 193–205.

Porcel, C., & Herrera-Viedma, E. (2010). Dealing with incomplete information in a fuzzy linguistic recommender system to disseminate information in university digital libraries. *Knowledge-Based Systems, 23*, 32–39.

Porcel, C., López-Herrera, A. G., & Herrera-Viedma, E. (2009a). A recommender system for research resources based on fuzzy linguistic modeling. *Expert Systems with Applications, 36*, 5173–5183.

Porcel, C., Moreno, J. M., & Herrera-Viedma, E. (2009b). A multi-disciplinar recommender system to advice research resources in university digital libraries. *Expert Systems with Applications, 36*, 12520–12528.

Prangl, M., Szkaliczki, T., & Hellwagner, H. (2007). A framework for utility-based multimedia adaptation. *IEEE Circuits and Systems for Video Technology, 17*, 719–728.

Pu, P., & Chen, L. (2007). Trust-inspiring explanation interfaces for recommender systems. *Knowledge-Based Systems, 20*, 542–556.

Pu, P., & Chen, L. (2009). User-involved preference elicitation for product search and recommender systems. *AI Magazine, 29*, 93–104.

Reichling, T., Veith, M., & Wulf, V. (2007). Expert recommender: Designing for a network organization. *Computer Supported Cooperative Work, 16*, 431–465.

Resnick, P., Iakovou, N., Sushak, M., Bergstrom, P., & Riedl, J. (1994). GroupLens: An open architecture for collaborative filtering of netnews. *Computer Supported Cooperative Work Conf.*

Ricci, F., & Nguyen, Q. N. (2007). Acquiring and revising preferences in a critique-based mobile recommender system. *IEEE Intelligent Systems, 22*, 22–29.

Riedl, J. (2001). Personalization and privacy. *IEEE Internet Computing, 5*, 29–31.

Robillard, M. P., & Dagenais, B. (2009). Recommending change clusters to support software investigation: an empirical study. *Journal of software maintenance, 22*, 143–164.

Roh, T. H., Oh, K. J., & Han, I. G. (2003). The collaborative filtering recommendation based on SOM cluster-indexing CBR. *Expert Systems with Applications, 25*, 413–423.

Rosaci, D., Sarné, G. M. L., & Garruzzo, S. (2009). MUADDIB: A distributed recommender system supporting device adaptivity. *ACM Transactions on Information Systems, 27*, 1–41.

Russell, S., & Yoon, V. (2008). Applications of wavelet data reduction in a recommender system. *Expert Systems with Applications, 34*, 2316–2325.

Schafer, J. B., Joseph, A., & Riedl, J. (2001). E-commerce recommendation applications. *Data Mining and Knowledge Discovery, 5*, 115–153.

Schiaffino, S., & Amandi, A. (2009). Building an expert travel agent as a software agent. *Expert Systems with Applications, 36*, 1291–1299.

Salter, J., & Antonopoulos, N. (2006). Cinema screen recommender agent: combining collaborative and content-based filtering. *IEEE Intelligent Systems, 21*, 35–41.

Sarwar, B., Karypis, G., Konstan, J. A., & Riedl, J. (2000a). Application of dimensionality reduction in recommender system- a case study. *Proceedings of the ACM WebKDD-2000 Workshop*.

Sarwar, B., Karypis, G., Konstan, J. A., & Riedl, J. (2000b). Analysis of recommendation algorithms for e-commerce. *Proceedings of the ACM E-Commerce*, 158–167.

Shao, B., Ogihara, M., Wang, D., & Li, T. (2009). Music recommendation based on acoustic features and user access patterns. *IEEE Transactions on Audio, Speech, and Language Processing, 17*, 1602–1611.

Shardanand, U., & Maes, P. (1995). Social information filtering: Algorithms for automating 'Word of Mouth'. *Human Factors in Computing Systems Conf.*

Su, J. H., Yeh, H. H., Yu, P. S., & Tseng, V. S. (2010). Music recommendation using content and context information mining. *IEEE Intelligent Systems, 25*, 16–26.

Symeonidis, P., Nanopoulos, A., & Manolopoulos, Y. (2008). Providing justifications in recommender systems. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, 38*, 1262–1272.

Symeonidis, P., Nanopoulos, A., & Manolopoulos, Y. (2010). A unified framework for providing recommendations in social tagging systems based on ternary semantic analysis. *IEEE Transactions on Knowledge and Data Engineering, 22*, 179–192.

Symeonidis, P., Nanopoulos, A., Papadopoulos, A. N., & Manolopoulos, Y. (2008). Collaborative recommender systems: Combining effectiveness and efficiency. *Expert Systems with Applications, 34*, 2995–3013.

Taab, S., Werther, H., Ricci, F., Zipf, A., & Gretzel, U. (2002). Intelligent systems for tourism. *IEEE Intelligent Systems, 6*, 53–66.

Taha, K., & Elmasri, R. (2010). SPGProfile: Speak group profile. *Information Systems, 35*, 774–790.

Tang, T. Y., & McCalla, G. (2009). A multidimensional paper recommender: Experiments and evaluations. *IEEE Internet Computing, 13*, 34–41.

Wang, H. F., & Wu, C. T. (2009). A mathematical model for product selection strategies in a recommender system. *Expert Systems with Applications, 36*, 7299–7308.

Wang, H. F., & Wu, C. T. (2010). A strategy-oriented operation module for recommender systems in E-commerce. *Computers & Operations Research, 39*, 1837–1849.

Wang, F. H., & Shao, H. M. (2004). Effective personalized recommendation based on time-framed navigation clustering and association mining. *Expert Systems with Applications, 27*, 365–377.

Wang, J. C., & Chiu, C. C. (2008). Recommending trusted online auction sellers using social network analysis. *Expert Systems with Applications, 34*, 1666–1679.

Wang, Y., Dai, W., & Yuan, Y. (2008). Website browsing aid: A navigation graph-based recommendation system. *Decision Support Systems, 45*, 387–400.

Wang, Y. F., Chuang, Y. L., Hsu, M. H., & Keh, H. C. (2004). A personalized recommender system for the cosmetic business. *Expert Systems with Applications, 26*, 427–434.

Wang, Y. F., Chiang, D. A., Hsu, M. H., Lin, C. J., & Lin, I. L. (2009). A recommender system to avoid customer churn: A case study. *Expert Systems with Applications, 36*, 8071–8075.

Wei, C. P., Yang, C. S., & Hsiao, H. W. (2008). A collaborative filtering-based approach to personalized document clustering. *Decision Support Systems, 45*, 413–428.

Wei, Y. Z., Moreau, L., & Jennings, N. R. (2005a). A market-based approach to recommender systems. *ACM Transactions on Information Systems, 23*, 227–266.

Wei, Y. Z., Moreau, L., & Jennings, N. R. (2005b). Learning users' interests by quality classification in market-based recommender systems. *IEEE Transactions on Knowledge and Data Engineering, 17*, 1678–1688.

Weng, S. S., & Chang, H. L. (2008). Using ontology network analysis for research document recommendation. *Expert Systems with Applications, 34*, 1857–1869.

Weng, S. S., & Liu, M. J. (2004). Feature-based recommendations for one-to-one marketing. *Expert Systems with Applications, 26*, 493–508.

Winoto, P., & Tang, T. Y. (2010). The role of user mood in movie recommendations. *Expert Systems with Applications, 37*, 6086–6092.

Vezina, R., & Militaru, D. (2004). Collaborative filtering: theoretical positions and a research agenda in marketing. *International Journal of Technology Management, 28*, 31–45.

Yager, R. R., Reformat, M. Z., & Gumrah, G. (2010). WebPET: An online tool for lexicographic decision making. *IEEE intelligent systems, 25*, 76–83.

Yang, H. L., & Wang, C. S. (2009). Recommender system for software project planning one application of revised CBR algorithm. *Expert Systems with Applications, 36*, 8938–8945.

Yang, J. M., & Li, K. F. (2009). Recommendation based on rational inferences in collaborative filtering. *Knowledge-Based Systems, 22*, 105–114.

Yang, Q., Knoblock, C. A., & Wu, X. (2004). Guest editors' introduction: mining actionable knowledge on the web. *IEEE Intelligent Systems, 19*, 30–31.

Yoshii, K., Goto, M., Komatani, K., Ogata, T., & Okuno, H. G. (2008). An efficient hybrid music recommender system using an incrementally trainable probabilistic generative model. *IEEE Transactions on Audio, Speech, and Language Processing, 16*, 435–447.

Yu, J. X., Ou, Y., Zhang, C., & Zhang, S. (2005). Identifying interesting visitors through Web log classification. *IEEE Intelligent Systems, 20*, 55–59.

Yuan, S. T., & Tsao, Y. W. (2003). A recommendation mechanism for contextualized mobile advertising. *Expert Systems with Applications, 29*, 399–414.

Yuan, W., Guan, D., Lee, Y. K., Lee, S. Y., & Hur, S. J. (2010). Improved trust-aware recommender system using small-worldness of trust networks. *Knowledge-Based Systems, 23*, 232–238.

Zanker, M., Jannach, D., Gordea, S., & Jessenitschnig, M. (2007). Comparing recommendation strategies in a commercial context. *IEEE Intelligent Systems, 22*, 69–73.

Zeng, C., Xing, C. X., Zhou, L. Z., & Zheng, X. H. (2004). Similarity measure and instance selection for collaborative filtering. *International Journal of Electronic Commerce, 8*, 115–129.

Zeng, D., Wang, F. U., Zheng, X., Yuan, Y., & Chen, J. (2008). Intelligent-commerce research in china. *IEEE Intelligent Systems, 23*, 14–18.

Zhan, J., Hsieh, C. L., Wang, I. C., Hsu, T. S., Liau, C. J., & Wang, D. W. (2010). Privacy-preserving collaborative recommender systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 40*, 472–476.

Zhang, Y., & Jiao, J. (2007). An associative classification-based recommendation system for personalization in B2C e-commerce applications. *Expert Systems with Applications, 33*, 357–367.

Zhen, L., Huang, G. Q., & Jiang, Z. (2009a). Collaborative filtering based on workflow space. *Expert Systems with Applications, 36*, 7873–7881.

Zhen, L., Huang, G. Q., & Jiang, Z. (2009b). Recommender system based on workflow. *Decision Support Systems, 48*, 237–245.

Zhen, L., Huang, G. Q., & Jiang, Z. (2010). An inner-enterprise knowledge recommender system. *Expert Systems with Applications, 37*, 1703–1712.

Zheng, N., Li, Q., Liao, S., & Zhang, L. (2010). Which photo groups should I choose? A comparative study of recommendation algorithms in Flickr. *Journal of Information Science, 36*, 733–750.

Zhu, X., Shi, Y. Y., Kim, H. G., & Eom, K. W. (2006). An integrated music recommendation system. *IEEE Transactions on Consumer Electronics, 52*, 917–925.

Ziegler, C. N., & Golbec, J. (2007). Investigating interactions of trust and interest similarity. *Decision Support Systems, 43*, 460–475.

CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE
Concurrency Computat.: Pract. Exper. (2015)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/cpe.3449

RESEARCH ARTICLE

A systematic literature review of Linked Data-based recommender systems

Cristhian Figueroa[1,2,*,†], Iacopo Vagliano[1],
Oscar Rodríguez Rocha[1] and Maurizio Morisio[1]

[1]*Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Turin, Italy*
[2]*Universidad del Cauca, Calle 5 No. 4–70, Popayán, Colombia*

SUMMARY

Recommender systems (RS) are software tools that use analytic technologies to suggest different items of interest to an end user. Linked Data is a set of best practices for publishing and connecting structured data on the Web. This paper presents a systematic literature review to summarize the state of the art in RS that use structured data published as Linked Data for providing recommendations of items from diverse domains. It considers the most relevant research problems addressed and classifies RS according to how Linked Data have been used to provide recommendations. Furthermore, it analyzes contributions, limitations, application domains, evaluation techniques, and directions proposed for future research. We found that there are still many open challenges with regard to RS based on Linked Data in order to be efficient for real applications. The main ones are personalization of recommendations, use of more datasets considering the heterogeneity introduced, creation of new hybrid RS for adding information, definition of more advanced similarity measures that take into account the large amount of data in Linked Data datasets, and implementation of testbeds to study evaluation techniques and to assess the accuracy scalability and computational complexity of RS.

Received 4 August 2014; Revised 27 October 2014; Accepted 15 November 2014

KEY WORDS: Linked Data; recommender systems; systematic review; web of data

1. INTRODUCTION

The increasing amount of heterogeneous information available on the Web has led to the difficulty in recommending relevant items that meet the requirements of end users. It has attracted the attention of researchers and has become an interesting research area from the development of the first *recommender systems* (RS) in the mid-1990s [1–3]. In fact, the interest in this area remains high because of the abundance of practical applications that help users to deal with different kinds of information [4].

Nowadays, RS are increasingly common in many application domains, as they use analytic technologies to suggest different items or topics that can be interesting to an end user. However, one of the biggest challenges in these systems is to generate recommendations from the large amount of heterogeneous data that can be extracted from the items. Accordingly, some RS have evolved to exploit the knowledge associated to the relationships between data of items and data obtained from different existing sources [5]. This evolution has been possible, thanks to the rise of the Web

*Correspondence to: Cristhian Figueroa, Politecnico di Torino, Corso Duca degli Abruzzi, 24 - 10129, Turin, Italy.
†E-mail: cristhian.figueroa@polito.it

supported by a set of best practices for publishing and connecting structured data on the Web known as *Linked Data* [6].

Linked Data principles have lead to semantically interlink and connect different resources at data level regardless of the structure, authoring, location, and so on. Data published on the Web using Linked Data have resulted in a global data space called the Web of Data. Moreover, thanks to the efforts of the scientific community and the W3C Linked Open Data (LOD) project‡, more and more data have been published on the Web of Data, helping its growth and evolution.

This work summarizes the state of the art of RS that make use of the structured data published as Linked Data on the Web. We undertook a systematic literature review, which is a form of secondary study that uses a well-defined methodology to identify, analyze, and interpret all available evidences related to specific research questions in a way that is unbiased and (to a degree) repeatable [7, 8]. We considered the most relevant problems that RS intended to solve, the way in which studies addressed these problems using Linked Data, their contributions, application domains, and evaluation techniques that they applied to assess their recommendations. Analyzing these aspects, we deduced current limitations and possible directions of future research. Unlike other works reporting the state of the art in RS [4, 9–11], our systematic literature review is the first to study RS that obtain information from Linked Data in order to generate recommendations.

The remainder of this paper is structured as follows. Section 2 provides a background information about Linked Data and RS. Section 3 summarizes the methodology and defines objectives and research questions. Section 4 outlines the results of the review organized according to each research question defined in Section 3. Section 5 discusses the results as well as the limitations of our systematic literature review. Section 6 contains the conclusions and future work. Finally, we list the selected papers in Appendix A.

2. BACKGROUND

2.1. *Linked Data*

In 1994, Tim Berners-Lee§ uncovered the need of introducing semantics into the Web to extend its capabilities and to publish structured data on it, which became known as *Semantic Web*. The set of good practices or principles for publishing and linking structured data on the Web is known as Linked Data. While the Semantic Web is the goal, Linked Data provides the means to make it a reality [6]. The set of Linked Data principles are as follows:

- Use URI (uniform resource identifiers) as names for things.
- Use HTTP (Hypertext Transfer Protocol) URIs, so that people can look up those names.
- Use of standard mechanisms to provide useful information when someone looks up a URI, for example, RDF (Resource Description Framework) to represent data as graphs and SPARQL (SPARQL Protocol and RDF Query Language) to query Linked Data.
- Include links to other URIs, so that they can discover more things.

The main benefit of using Linked Data as a source for generating recommendations is the large amount of available concepts and the relationships between them that can be used to infer relations more effectively in comparison to derive the same kind of relationships from text [12]. As Linked Data information is machine-readable, it is possible to query datasets on a fine-grained level in order to collect information without having to take manual actions; therefore, information is explicitly represented, which allows for applying reasoning techniques when querying datasets and making implicit knowledge explicit.

2.2. *Recommender systems*

RS are software tools and techniques that provide suggestions of items to a user. These items can belong to different categories or types, for example, songs, places, news, books, films, and events.

‡http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData
§http://www.w3.org/Talks/WWW94Tim

DOI: 10.1002/cpe

According to Adomavicius and Tuzhilin [4], the roots of RS can be traced back to the works in cognitive science, approximation theory, information retrieval, forecasting theories, management science, and consumer choice modeling in marketing.

Nowadays, RS are focused on the recommendation problem of guiding users in a personalized way to interesting items in a large space of possible options [10]. Typically, RS are classified as content based, collaborative filtering, knowledge based, and hybrid [5].

Content-based RS make suggestions that take into account the ratings that users give to items according to their preferences and the content of the items (e.g., extracted keywords, title, pixels, and disk space) [10]. Collaborative-filtering RS generate recommendations of items to a user taking into account ratings that users with similar preferences have given to these items [13]. Knowledge-based RS infer and analyze similarities between user requirements and features of items described in a knowledge base that models users and items according to a specific application domain [14]. Hybrid RS combine one or more of the aforementioned techniques in order to improve recommendations.

With the evolution of the Web toward a global space of connected and structured data, a new kind of knowledge-based RS has emerged known as Linked Data-based RS. This kind of RS suggests items taking into account the knowledge of datasets published under the Linked Data principles. The systematic literature review presented in this paper is focused on this kind of RS.

3. RESEARCH METHODOLOGY

This work studies the state of the art in Linked Data-based RS. It follows the guidelines set out by Kitchenham and Charters [8] for systematic literature reviews in software engineering. These guidelines provide a verifiable method of summarizing existing approaches as well as identifying challenges and future directions in the current research. Figure 1 presents the protocol for our systematic literature review.

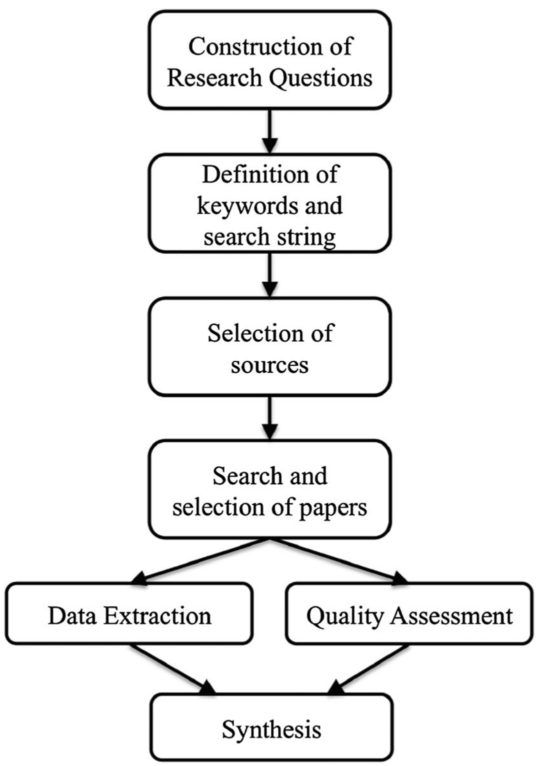

Figure 1. Systematic literature review at a glance.

DOI: 10.1002/cpe

The protocol is defined in order to setup the steps to conduct the systematic literature review. In our work, it was developed by the first and second authors, while the third and fourth authors validated it.

3.1. Construction of research questions, definition of keywords and search string, and selection of sources

The goal of our systematic literature review is to understand how the implicit knowledge, stored in Linked Data datasets and represented as concepts and relations between them, can be exploited to make recommendations. Accordingly, we have defined the following research questions:

RQ1 What studies present RS based on Linked Data?
RQ2 What challenges and problems have been faced by researchers in this area?
RQ3 What contributions have already been proposed (e.g., algorithms, frameworks, and engines)?
RQ4 How is Linked Data used to provide recommendations?
RQ5 What application domains have been considered?
RQ6 What criteria and techniques are used for evaluation?
RQ7 Which directions are the most promising for future research?

Afterwards, a preliminary set of keywords was defined: {*Linked Data, Recommender system*}. This set was then extended by searching for synonyms in order to obtain the final set of keywords used to define a search string. The search string is the query to look for papers in a set of online digital libraries. In this work, the search string that we defined is as follows:

```
("semantic web" OR "linked data" OR "web of data" OR "linked open data") AND (recommendation OR "recommender system" OR "recommendation system" OR "semantic recommendation" OR "semantic recommender").
```

Furthermore, we selected seven scientific digital libraries that represent primary sources for computer science research publications as can be seen in Table I. Other sources like DBLP, CiteSeer, and Google Scholar were not considered as they mainly index data from the primary sources.

3.2. Search and selection

The studies selected in this systematic literature review were identified from the selected sources during March 2014. In Table II, a set of inclusion/exclusion criteria were defined in order to determine whether or not a study should be included.

3.3. Quality assessment, data extraction, and synthesis

We have defined a set of quality criteria that are listed in the checklist provided in Table III. Quality for each question is typically scored with values 1, 0.5, and 0, in order to represent the answers *'yes'*, *'partly'*, and *'no'*.

First and second authors evaluated the selected studies using this checklist. To do this, the total set of selected papers was split into two disjoint subsets, and each author selected only one of these

Table I. Sources selected for the search process.

Source	URL
IEEExplore	http://ieeexplore.ieee.org
SpringerLink	http://link.springer.com
Scopus	http://www.scopus.com
ACM Digital Library	http://dl.acm.org
Science Direct	http://www.sciencedirect.com
ISI Web of Knowledge	http://apps.webofknowledge.com
Wiley Online Library	http://onlinelibrary.wiley.com

DOI: 10.1002/cpe

Table II. Inclusion and exclusion criteria.

Inclusion criteria
Papers presenting recommender systems (RS) using Linked Data to provide recommendations.
Papers addressing exploratory search systems using Linked Data. Exploratory search refers to cognitive consuming search such as learning or topic investigation. Exploratory search systems also recommend relevant topics or concepts, although the key difference with respect to RS is that they still require an input query (commonly a set of keywords).
Papers from conferences and journals.
Papers published from 2004 to 2014. Linked Data is a relative new technology; therefore, RS approaches exploiting it are also recent.
Only papers written in English language.
Short and workshop papers that fulfill the above criteria: we had no reason to believe that they would fail to provide sufficient levels of detail about their studies.
Exclusion criteria
Papers not addressing RS neither exploratory search systems.
Papers addressing RS or exploratory search systems that do not exploit Linked Data to produce recommendations.
Papers addressing similarity measures but not RS. Similarity is a broader topic than RS.
Papers that use Semantic Web techniques (e.g., rule-based or ontology-based reasoning) but not Linked Data.
Papers that report only abstracts or slides of presentations because of the lack of information.
Grey literature. We do not think that technical reports, unpublished studies, and PhD thesis would add much more information with respect to journal and conference papers.

subsets to evaluate the papers. After this evaluation, cross-checking of the assessment was done on arbitrary studies (about 30 % of selected papers) by the third author. Finally, an agreement on differences was reached by discussion.

Data extraction was done in parallel with the quality assessment. We split the set of included studies into two disjoint subsets. First and second authors performed the task on a subset, then the third author cross-checked a random sample of 30% of studies. The data extracted are presented in Table IV.

The synthesis step is based on the methodology for thematic synthesis described by Cruzes and Dybå [15]. This methodology defines codes as descriptive labels applied to segments of text from each study. We defined an initial set of codes based on research questions, and subsequently, we performed a second coding with more precise codes, which were closer to the content of selected papers. The coding was performed by first and second authors: each of them addressed a subset of the papers as for data extraction and quality assessment, because it was done in parallel with them. Then, the third author performed again the coding on a random sample of 30% of papers for cross-checking; afterwards, disagreements were solved by discussion.

4. RESULTS

This section summarizes the relevant information found in the selected studies in order to answer the proposed research questions. A further discussion and analysis of these results are addressed in Section 5.

4.1. Included studies

RQ1 regards the studies that present RS based on Linked Data. We retrieved 69 papers to include in the systematic literature review, corresponding to 52 unique primary studies (a study is a unique research work that can include one or more papers). These studies were published in conferences, workshops, and journals between 2004 and 2014. The criteria for deciding the most significant

DOI: 10.1002/cpe

Table III. Quality assessment checklist.

Question	Score
Q1. Did the study clearly describe the challenges and problems that is addressing?	yes / partly / no (1 / 0.5 / 0)
Q2. Did the study review the related work for the problem?	yes / partly / no (1 / 0.5 / 0)
Q3. Did the study discuss related issues and compare with the alternatives?	yes / partly / no (1 / 0.5 / 0)
Q4. Did the study recommend the further continuous research?	yes / partly / no (1 / 0.5 / 0)
Did the study describe the components or architecture of the proposed recommender system?	yes / partly / no (1 / 0.5 / 0)
Q5. Did the study describe the components or architecture of the proposed recommender system?	yes / partly / no (1 / 0.5 / 0)
Q6. Did the study provide empirical results?	– The study provided an implementation of its work with an empirical evaluation and it was used in real applications, e.g., by other services (1) – The study provided an implementation of its work and an empirical evaluation but was not referred or used in other studies/applications (0.75) – The study provided an implementation only (0.5) – The study did not provide any implementation but it was referred by other works as a base on which start (0.25) – The study did not provide any implementation and was not referred by other works (0)
Q7. Did the study provides a clear description of the context in which the research was carried out?	yes / partly / no (1 / 0.5 / 0)
Q8. Did the study presents a clear statement of findings?	yes / partly / no (1 / 0.5 / 0)

paper for each study were completeness and publication year. The final set of selected papers and corresponding studies can be found in Appendix A.

With regard to the quality assessment, *journals* and *conference* studies have better quality than *workshop* studies as shown in Figure 2. Conference studies have the biggest spread, while journal studies, the lowest. In any case, the quality score is higher than 0.5 for all paper types, that is, rather good according to the quality criteria defined in Section 3.3.

4.2. Research problems

In order to address RQ2, we summarize the main problems involved in the studies considered and regarding the production of accurate recommendations. Table V lists these problems according to the number of studies in which they occurred. The number of studies represents the occurrence of each problem in the selected studies, which may be addressed in more than one study. The same applies for the rest of the results reported in this section.

In the following, we describe each item of Table V:

Lack of semantic information It was the most frequent problem in the selected studies, and it concerns the need for exploiting the rich semantics of information about items. Possible causes of this problem are as follows:

- Data about items are unstructured.
- A categorization of the items is needed.

DOI: 10.1002/cpe

Table IV. Data extraction form.

Data field	Description	Research question
ID	—	—
Title	—	—
Authors	—	—
Year of publication	—	—
Year of conference	—	—
Volume	—	—
Issue	—	—
Location	—	—
Proceeding title	—	—
ISBN	—	—
Publisher	—	—
Examiner	Name of person who performed data extraction	—
Publication source	—	—
Context	Environment in which study was conducted: industry, academic, government	—
Population	Study participants: students, academics, practitioners, etc.	—
Aims	Goals of the study (in our opinion when not clearly reported by authors)	—
Research problem	—	RQ2
Application domain	—	RQ5
Contributions	—	RQ3
Criteria and techniques for evaluation	—	RQ6
Findings	—	—
Limitations	—	RQ7
Future work	—	RQ7
Notes	—	—
Other information	—	—

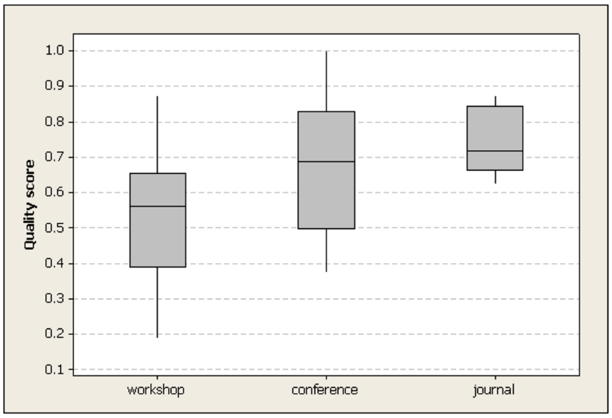

Figure 2. Quality score for different types of study.

- It is necessary to find relationships to link items.
- Social information is lacking.
- It is necessary to acquire content-descriptive metadata.
- Similarity measures that take into account semantic information are needed.

Complexity of information about items It is related to the complexity of information because of noisy metadata about features of items. Other causes for this problem are semantic heterogeneity and distribution of resources. The latter can impact on maintenance of the knowledge bases and can also decrease the accuracy of recommendations.

DOI: 10.1002/cpe

Table V. Distribution of studies according to the problems they addressed.

Problems	Number of studies
Lack of semantic information	13
Complexity of information about items	12
User dependency	8
Cold-start	6
Data quality	6
Computational complexity	5
Data sparsity	5
Domain dependency or specific and limited domain	4
Other problems	2

Table VI. Distribution of studies according to the contributions provided.

Contribution	Number of study
Algorithms	27
Similarity measures	12
Ontologies	8
Information aggregation or enrichment	8
Others	16

User dependency In a number of cases, RS require users to perform manual operations to acquire information about their profiles and interests. Such operations can be user feedback, ratings, filtering, attaching content-descriptive metadata, and semantic annotation of items.

Cold-start It is a well-known problem found mainly on RS based on collaborative-filtering approaches. Cold-start is a situation in which there are not enough ratings for items in order to generate recommendations.

Data quality This problem occurs when the knowledge base used to acquire information for providing recommendations is not reliable. Problems affecting data quality can range from poor reliability (e.g., wrong links between concepts or incorrect representations) to poor quality of recommended items.

Computational complexity It is related to the high computational demand that RS require to produce recommendations because of the large amount of data about items.

Data sparsity This is related to the lack of information about users or items and generates low density of significant data or connections.

Domain dependency It occurs when recommendations are only useful for items in a specific and limited domain without taking into account data that can be obtained from other related domains.

Other problems They include the need for recommending relevant and yet unknown items and the overspecialization of RS.

4.3. Contributions

In order to address RQ3, we classified the contributions provided by each study. Table VI shows the different kind of contributions and the number of studies in which they occurred (each study possibly reports more than one contribution).

The two main contributions are the definition or extension of a similarity measure and the definition or extension of an ontology, accounting for 12 and eight studies respectively. Algorithms are also addressed by 27 studies in total. Finally, information aggregation or enrichment and various other contributions account for eight and 16 studies, respectively. In the following, we describe each item of Table VI:

Algorithms Most of the selected studies proposed new algorithms or extensions of algorithms existing in the literature. In particular, four categories emerged: defining of a new algorithm,

DOI: 10.1002/cpe

adapting an algorithm to Linked Data, combining of algorithms to obtain a new hybrid algorithm, and extending of an existing algorithm. The definition of a new algorithm was the most frequent in 15 studies, while the adaptation of an algorithm to Linked Data, the combination of algorithms to obtain a new hybrid algorithm, and the extension to an algorithm each account for 4 studies. Furthermore, we can group algorithms into two classes:

- Graph-based algorithms, which compute relevance scores for items represented as nodes in a graph. A number of algorithms in this category are (*i*) the weight spreading activation algorithm, which propagates the initial score of a source node through its weighted edges; (*ii*) algorithms that update the scores of its linked nodes; (*iii*) algorithms that explore concepts and relations defined in an RDF graph; (*iv*) topic-based algorithms, which find similar items belonging to the same categories of an initial concept; and (*v*) path-based algorithms to find semantic paths between documents in the RDF graph.
- Algorithms to produce recommendations based on statistical information techniques applied to Linked Data such as support vector machine (SVM), latent Dirichlet allocation (LDA), random indexing (RI), and scaling methods. SVM analyzes and recognizes patterns in RDF triples; LDA is based on the co-occurrence of terms; RI uses distributional statistics to generate high-dimensional vector spaces; and scaling methods take into account the probability that an item could be selected based on its popularity (the number of entities is directly connected with the node). In addition, some algorithms define item-user matrices to compute semantic similarity based on path-lengths.

Similarity measures The selected studies applied a variety of similarity measures. These include pairwise cosine function for vector similarity computation between items, feature-based similarity to evaluate semantic distance on different datasets, rating-based similarity to compute the popularity of items among users, semantic relatedness defined by vocabulary meta-descriptions, content similarity that exploits lexical features, expressivity closeness based on the language constructs adopted, distributional relatedness derived from vocabulary usage, and topic-based similarity that captures the relatedness between items based on the categories they belong to.

Ontologies A number of studies proposed ontologies to assist or improve the recommendation process. New ontologies were proposed to facilitate the process of integration of datasets from a number of domains in order to make RS more flexible to changes, while a combination of existing ontologies described different types of entities such as users and items. Furthermore, it was found that reusing existing ontologies or vocabularies enable interoperability. Ontologies are also used to represent semantic distances, their explanations, user preferences, and item contents. A number of ontologies that are used in selected studies for these purposes are FOAF (Friend Of A Friend), SIOC (Semantically-Interlinked Online Communities), Resource List Ontology and Bibliographic ontology.

Information aggregation or enrichment This refers to the contributions about the aggregation of data to item collections and enrichment of existing ontologies or vocabularies. This is useful, for example, to obtain descriptive information about items and find entities in datasets in order to infer links between them. One contribution of this type is the aggregation of information from a specific domain when items have to be enriched with knowledge contained only on specialized datasets, another is the enrichment databases of RS with shared vocabularies.

Others Other contributions include the integration of other techniques such as opinion aggregators, exploitation of trust in web-based social networks to create predictive RS, and the use of social-based algorithms to improve the performance of the RS.

4.4. Use of Linked Data

Another interesting aspect that we studied was the use of Linked Data in RS, as underlined by RQ4. We classified the selected studies according to the way they used Linked Data to produce recommendations and grouped them into the following:

Linked Data driven RS that rely on the knowledge of the Linked Data to provide recommendations. For example, RS that calculate a semantic similarity based on diverse relationships that can

DOI: 10.1002/cpe

Table VII. Distribution of studies according to the use of Linked Data.

Category	Number of studies
Linked Data driven	37
Hybrid	29
Hybrid and Linked Data driven	21
Linked Data driven only	13
Representation only	10
Hybrid only	6
Exploratory search	4
Exploratory search and Linked Data driven	4
Exploratory search only	0

be found between concepts of Linked Data datasets and are related to features or descriptions of items. Such relationships can be paths, links, or shared topics among a set of items. This category can also include RS that use other techniques applied on data obtained from Linked Data datasets, for example, weight spreading activation, vector space model (VSM), SVM, LDA, and random indexing.

Hybrid RS that exploit Linked Data to perform some operations that can be used or not used to provide recommendations. This means that hybrid RS include Linked Data driven RS, which use recommendation techniques that rely on Linked Data, and RS that use Linked Data in other operations (not necessarily for recommending) that can be preliminary to the recommendation process (e.g., to aggregate more information from other datasets, to describe user profiles, or to annotate raw data in order to extract information to be integrated and used for recommending).

Representation only RS in this category exploit the RDF format to represent data and use at least one vocabulary or ontology to express the underlying semantics. However, no information is extracted from other dataset, and Linked Data are not used to provide recommendations. An example is an RS that represents the information about the users according to FOAF vocabulary but does not exploit Linked Data for other operations.

Exploratory search These systems are not RS, but their main duty is to assist users to explore knowledge and to suggest relevant to a topic or concept. Exploratory search systems and RS use Linked Data in a very similar way, although the key difference is that exploratory search systems still require an explicit input query (commonly a set of keywords). Additionally, users in these systems are not only interested in finding items but also in learning, discovering, and understanding novel knowledge on complex or unknown topics [16].

Each study may be assigned to more than one category; that is, it can be both Linked Data driven and hybrid, or both exploratory search and Linked Data driven. The only exception is for the representation-only category, in which studies cannot belong to other categories.

Table VII shows that most of the studies considered are Linked Data driven, and roughly 60% of them are also hybrid. Only 20% of hybrid studies were hybrid only, while the rest are also Linked Data driven. Moreover, 10 studies are representation only and just four exploratory search systems were included in the systematic literature review. All of the exploratory search studies are also Linked Data driven. This finding is consistent with the focus of the systematic literature review, which is on RS using Linked Data. It is worth noting that exploratory search is a broader topic; in this paper, we only consider the exploratory systems that recommend concepts to users.

The two most interesting categories are Linked Data driven and hybrid. Figure 3 shows the different techniques used by the studies in the first category to provide recommendations. The majority of them rely on datasets or on a similarity measure (about 43% and 35%, respectively), while the remaining 22% adapt natural language processing or content-based techniques or exploit reasoning.

Instead, Figure 4 illustrates the techniques that hybrid studies use together with Linked Data to provide recommendations. Most of them are natural language processing or collaborative-filtering methods (accounting for slightly less than 40% and about 35%, respectively) , and also reasoning or social networks are exploited in some cases.

DOI: 10.1002/cpe

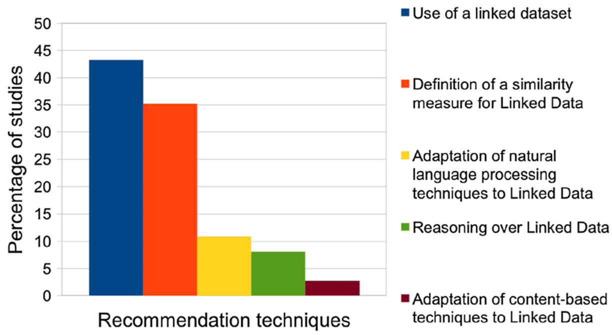

Figure 3. Distribution of Linked Data driven studies according to the recommendation techniques that they exploit (percentages refer to the total number of Linked Data-driven studies).

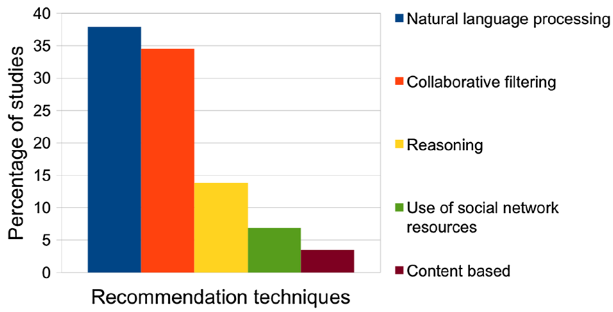

Figure 4. Distribution of hybrid studies according to the recommendation techniques that they exploit (percentages refer to the total number of hybrid studies).

Table VIII. Distribution of studies according to the Linked Data (LD) datasets on which they rely.

	Number of studies				
Dataset	General	LD driven	Hybrid	Hybrid and LD driven	LD driven only
DBpedia	31	28	20	16	12
Freebase	6	6	5	5	1
YAGO	4	3	3	2	1
Wordnet	4	2	3	2	0
DBLP	3	3	3	3	0
Dataset independent	3	3	3	3	0
LinkedMDB	3	3	3	3	0
Geonames	2	1	2	1	0
MusicBrainz	2	1	2	1	0
mySpace	2	2	2	2	0
ACM	1	1	1	1	0
IEEE	1	1	1	1	0
Eventseer2RDF	1	1	1	1	0
LinkedUp	1	1	0	0	1
mEducator	1	1	0	0	1
LinkedGeoData	1	0	1	0	0
LODE	1	1	1	1	0

In addition, we studied which datasets are used and the outcome is presented in Table VIII. It shows how many studies use a dataset overall and also considers the study category. It is possible to notice that DBpedia is used much more than the others. In fact, it is the biggest dataset, and it is the most curated.

DOI: 10.1002/cpe

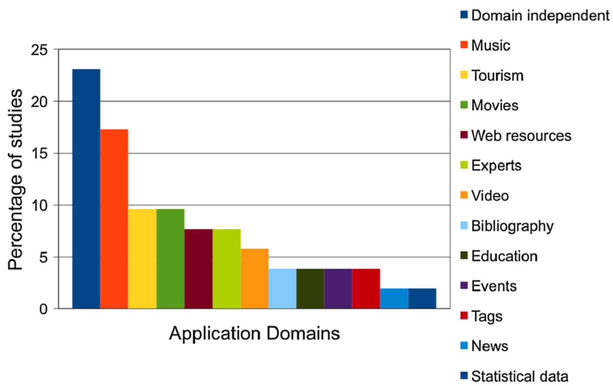

Figure 5. Distribution of studies according to the application domain.

Furthermore, it contains information about many different domains. Other commonly used datasets are Freebase, YAGO, and Wordnet, but the latter is used in just half of the cases by Linked Data-driven studies. In fact, it is also used with natural language processing techniques. On the contrary, the other datasets are used in most cases by Linked Data driven studies and often by studies which are both Linked Data driven and hybrid.

4.5. Application domains

Figure 5 illustrates the application domains considered by the studies selected for the systematic literature review. Most of the studies (about 23%) are not limited to any particular domain and can be used to recommend different kinds of items. Instead, an often occurring domain is music, which represents 17% and is followed by tourism and movies, accounting for roughly 10% each. Then there are web resources, expert recommendations, and video, with between 5% and 7% each, and a number of other domains are considered by the remaining 10% of the studies.

4.6. Evaluation techniques

RQ6 concerns RS evaluation, so we also dealt with this aspect. It is important to note that we focus on RS evaluation; thus, GUI evaluation is not considered, although some of the studies addressed it. RS are commonly evaluated according to their computational complexity and accuracy [17]. The former measures the execution time required to produce recommendations, which depends on the complexity of the algorithms used as well as the runtime of third-party systems needed to produce recommendations. The latter is the capacity of the RS to satisfy the individual user's need for information, and it can be evaluated by means of two techniques: user studies and comparison with similar methods. In this subsection, we detail both of them.

User studies involve users in order to compare recommendations generated by RS with the users' judgements or ratings. In these techniques, the most frequent measures are the following:

- Precision and recall, which evaluate the relevance of an RS taking into account the number of retrieved items, the number of items that evaluators considered as relevant, and the total number of available items.
- User ratings, which are techniques in which a list with results from different RS are presented to users who rate the lists according to their personal criteria [17].
- Ranking quality, which takes into account the retrieval correctness. The latter assigns an output ranking, a performance score based upon the available reference relevance judgments [18]. Common metrics to measure the ranking quality are the normalized discounted cumulated gain, average position, and presence.
- Unexpectedness of a concept suggestion, which is the degree of novelty of a recommendation for the evaluator.

DOI: 10.1002/cpe

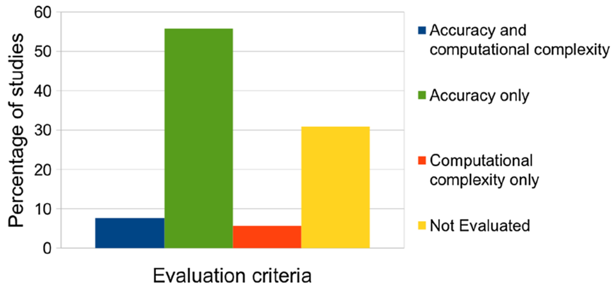

Figure 6. Distribution of studies according to the evaluation criteria (percentages refer to total number of studies).

Table IX. Distribution of studies according to the evaluation techniques.

Type	Technique	Number of papers
User studies	Precision and recall	18
	User ratings	9
	Ranking quality	3
	Unexpectedness	3
Comparison with similar methods	Precision and recall	5
	MAE and RMSE	3
Computational complexity	Execution time	7

MAE, mean absolute error; RMSE, root mean squared error.

In the case of comparisons with similar methods, recommendations generated by a specific RS are compared with well-known similar approaches. In the following, we mention the two main types:

- Precision and recall are measured, but in this case, items recommended by a well-known approach are considered as relevant.
- Mean absolute error (MAE) and root mean squared error (RMSE) are metrics to measure the predictive accuracy of an RS. MAE calculates the average absolute deviation between predicted similarities and similarity values in the real data set, while RMSE pays more attention to large errors [19].

Figure 6 shows the main evaluation techniques found in the selected studies, as well as their classification and their occurrence in these studies. Studies that provided an evaluation accounted for about 70% of the studies included in the systematic literature review. Among these, roughly 55% only used an accuracy technique, while roughly 2% only evaluated the computational complexity, and slightly less than 8% considered both accuracy and computational complexity.

Table IX details the techniques used in the studies included by considering the two types of accuracy evaluation and also computational complexity. The most frequent technique used to evaluate RS is the relevance measured with recall and precision metrics (used by 18 works in user studies and by about five in comparison with similar methods). We expected this result because these metrics are the ones most commonly deployed in information retrieval approaches. Other widely used techniques are user ratings, accounting for nine studies, and execution time, which is exploited by seven studies.

4.7. Future work

RQ7 is related to directions for future research. To address this, we summarized the future work that the selected studies proposed in order to extend or improve their approaches. Specifically, about 67% of studies included in the systematic literature review present diverse proposals for future work. Table X lists the most important, indicating for each one the number of studies in which it was mentioned. A deeper analysis of these results and a discussion of possible directions is presented in Section 5.

In the following, we provide a brief description of each item reported in Table X:

DOI: 10.1002/cpe

Table X. Distribution of studies according to the future work they propose.

Future work	Number of studies
Personalization of recommendations	8
Use more datasets	8
Create hybrid recommender systems	7
Similarity measures	4
Find more semantic relationships (item-user and item-item)	3
Other proposal for future work	3
Consider other domains	2

Personalization of recommendations The idea is to know to what extent personalization can improve recommendations without requiring user profile information or user intervention for manual operations (feedback, filtering, annotation, etc.).

Use more datasets It means to increase the range of data to annotate or match items to be recommended. It can also be useful to explore new domains because of the use of other datasets which can be from diverse domains.

Create hybrid RS This refer to exploring new ways to combine diverse recommendation techniques for creating hybrid approaches and improving the relevancy and quality of recommendations.

Similarity measures It is the creation of new similarity measures or the improvement of existing ones.

Find more semantic relationships It is the possibility of finding more semantic relationships between items and between users and items. It is considered by three studies.

Consider other domains Although domain dependency is one of the problems found in various studies, only two studies took into account exploring new application domains for providing recommendations.

Other proposal for future work This group includes applications in real life contexts, algorithms for categorization of recommendations, improving performance of algorithms, and the study of disambiguation techniques.

4.8. Limitations

The limitations reported in the selected studies are also related to RQ7 as these can help us to uncover the open issues in RS based on Linked Data and their relationships with proposals of future work. They are grouped into four main types: datasets, manual operations, personalization, and computational complexity. We detail each of them in the following:

Datasets This type describes limitations of RS due to the datasets used.

- A number of studies required a local copy of the entire dataset in a local server in order to reduce the runtime to produce recommendations. This had to be done as sometimes public datasets offer limited results, restricted access, and high timeout.
- Sometimes data had to be manually curated because of the poor reliability of public datasets.
- A number of RS are limited to the use of only one dataset. This can restrict the knowledge to which the RS can have access, avoiding data from diverse sources and domains being obtained.

Manual Operations It means that RS needed the user to perform manual operations in order to produce recommendations. Among these operations, we found:

- RS requiring manual selection of relevant concepts according to a specific application domain or interests. This is a difficult and tedious task considering the large amount of data that a typical Linked Data dataset can contain.
- RS that did not rank their results, so final users are faced with no priority in the recommendation.

DOI: 10.1002/cpe

Personalization It is about producing recommendations according to the user profile or some personal features.

Computational complexity RS still need to improve the performance because of high computational demand to analyze large amounts of items and information stored into datasets. Another problem is the poor performance of public endpoints to access them.

5. DISCUSSION

In the first part of this section, we present a discussion of the results considering each research question, while in the second part, we mention the limitations of our systematic literature review.

5.1. Specific research questions

This subsection discusses the research questions addressed in this systematic literature review according to the results reported in Section 4.

RQ1 is a general question regarding the studies that describe RS based on Linked Data. To provide an answer, we have followed the steps described in the protocol presented in Section 3 in order to search and select studies in this area. Firstly, we retrieved a total number of 7873 papers (including those duplicated) from scientific digital libraries. After each author filtered papers by title and abstract, we discussed disagreements, and we reach consensus on a final set of 69 papers to include in our study, which correspond to 52 unique studies.

RQ2 deals with research problems in the RS domain that researchers intended to solve by proposing approaches based on Linked Data. We found that the lack of semantic information and its complexity were the most notorious problems in RS.

Lack of semantics regards the need for rich semantic information about items. This is the main reason to devise novel strategies to represent items and user profiles using diverse semantic techniques exploiting several knowledge sources from the Linked Data cloud.

The complexity and heterogeneity of information and the subsequent cost of maintenance of knowledge bases make Linked Data a suitable solution that uses publicly available knowledge bases that are continuously growing and maintained by third parties. However, this poses new challenges, for example, the need for mechanisms to assure the reliability of these knowledge bases that are used to describe user profiles and items and to generate recommendations.

Domain dependency is another problem that has been also addressed by using Linked Data because it allows the possibility to exploit information from different datasets that can be domain-independent or belong to diverse domains. In fact this is one reason why the most used dataset is DBpedia as it is the most generic dataset that can be used for cross-domain RS. Nonetheless, some studies still report this problem as future work.

Computational complexity is a question that has not been widely addressed in the studies considered in this systematic literature review and remains as an open issue because most of the studies have concentrated only on semantic enrichment of items and inclusion of datasets in Linked Data cloud. Computational complexity needs to be addressed more because in RS not only accuracy is important but also scalability and responsiveness. For example, this problem can be critical in RS for mobile scenarios where users demand fast response times.

Other problems such as usability, cold-start, data quality, and data sparsity have been addressed by combining with Linked Data various techniques based on natural language processing, reasoning or social network resources, and creating hybrid RS that exploit both collaborative filtering and content-based approaches.

RQ3 inquires about the contributions proposed in RS based on Linked Data. The analysis showed that the majority of studies are focused not only on providing new algorithms but also on defining or extending a similarity measure of an ontology. Furthermore, adaptation, combination, or extension to algorithms is quite often addressed together with information aggregation

DOI: 10.1002/cpe

or enrichment. Accordingly, we found that Linked Data can be used in RS for several purposes such as the following:

- Defining different similarity functions between items or users by exploiting the large data available in the Linked Data cloud and the vast relationships already established such as properties or context-based categories. In this way, it is possible to extract semantic information from textual descriptions or other textual properties about the items in order to find semantic similarities based on the information stored in interlinked vocabularies of Linked Data. This can be useful in RS based on collaborative filtering to improve the neighborhood formation in user-to-user or item-to-item.
- Generating serendipitous recommendations, for example, to recommend items that are not part of the users' personal data cloud, that is, suggest new, possibly unknown items, to the user; or to guide users in the process of the exploration of the search space giving the possibility for serendipitous discovery of unknown information (for exploratory search systems).
- Offering the explanation of the recommendations given to the users by following the linked-data paths among the recommended items. In this way, users can understand the relationship between the recommended items and why these items were recommended.
- Domain-independency when creating RS as it is possible to access data from Linked Data datasets from different domains.
- Enrichment of information sources such as databases, repositories, and registries with information obtained from dataset in Linked Data cloud which manage huge amounts of data. It offers the possibility to enrich graphs representing users and/or items with new properties in order to improve graph-based recommendation algorithms. Additionally, it helps to mitigate the new-user, new-item, and sparsity problems.
- Annotating items and users with information from multiple sources facilitate RS to suggest items from different sources without changing their inner recommendation algorithms. Using such a semantic-based knowledge representation, recommendation algorithms can be designed independently from the domain of discourse.
- Obtaining hierarchical representation of items because the topic distribution that some datasets in Linked Data cloud offer. In this way, RS can base their recommendation on the exploration of items belonging to similar categories.

RQ4 regards the diverse ways in which Linked Data is used to provide recommendations. First of all, we classified the studies according to the way they exploited Linked Data. As reported in Section 4, four categories were identified: *Linked Data driven RS* rely mainly on Linked Data to perform their tasks, *hybrid RS* use Linked Data and also other techniques, *representation-only RS* do not provide Linked Data-based recommendations but use Linked Data for representing data based on RDF, and finally *exploratory search systems* that are not RS but may help users to find concepts or topics and have some similar features to RS especially in the use of Linked Data.

Table XI describes each category including the most important studies that adopted these strategies, as well as their advantages and disadvantages. The numbers of the studies corresponds to the identifiers in Appendix A.

Most of the studies belong to the first category, and many belong to both the first and the second category. These two categories are also the most interesting as they include RS to better exploit the advantages provided by Linked Data in order to reach best results. We also studied techniques to provide recommendations relying on Linked Data and slightly less than half of Linked Data driven RS used a dataset, almost one third define a similarity measure for Linked Data, while others adapt natural language processing or content-based methods or use reasoning.

With reference to the techniques used together with Linked Data, we found that natural language processing and collaborative filtering are the most used (both account for about one

DOI: 10.1002/cpe

Table XI. Classification of Linked Data-based RS approaches.

Approach	Techniques	Advantages	Disadvantages
Linked Data driven	– *Graph based*: weight spreading activation (S17), semantic exploration in an RDF graph (S29, S10, S3, S9, S19), and projections (S23) – *Reasoning*: (S1, S51) – *Statistical*: Matrix item-user (S29, S35, S31, S13, S37, S10), Scaling methods (S29) and topic discovery (S2)	– Generating serendipitous recommendations – Offering explanations of the recommendations following the linked-data paths – Creating domain-independent RS – Exploiting hierarchical information about items to categorize recommendations	– High cost of exploiting semantic features due to inconsistency of LD datasets – No personalization – No contextual information – High computational complexity – Need for manual operation – Need for dataset customization to address the computational complexity
Hybrid	– *Collaborative Filtering and Linked Data*: (S2, S4, S12, S25, S27, S3, S28, S26, S30, S35) – *Information aggregation and Linked Data*: opinions (S16), ratings (S19), and social tags (S32) – *Statistical methods and Linked Data*: Random Indexing (S10), VSM (S47, S31, S35), LDA (S35), Implicit feedback (S25), SVM (S13), Structure-based statistical semantics (S37)	– Overcoming the data sparsity problem – Allowing collaborative filtering RS to address the cold start problem	– High computational complexity
Representation only	– Item/user information representation using RDF-based ontologies (S36, S38, S20, S40, S14, S15, S42, S46)	– Improving scalability and reusability of ontologies – Easing data integration – Enabling complex queries	– Difficult to reuse the already available knowledge in the Linked Data Cloud
Explorative search	– Set nodes and associated lists (S49, S39, S34) – Spreading activation to typed graphs and graph sampling technique (S11)	– Enabling self-explanation of the recommendations	– No automation of the recommendation because explorative search approaches require frequent interaction with the user

DOI: 10.1002/cpe

third of hybrid RS) as they intended to provide personalized suggestions of items tailored to the preferences of individual users.

Other techniques are less common (less than 15%), and they are reasoning, use of social network resources, and content-based methods. Reasoning has not been widely used as its quality is still insufficient, and its coverage is not broad enough at the level of system components and knowledge elements [20]. Therefore, one solution is to develop RS based on reasoning-oriented natural language processing enriched with multilingual sources and able to support knowledge sources generated largely by people as Linked Data datasets.

As for the datasets used in the selected studies, we found that DBpedia is the most used Linked Data dataset. This is because DBpedia is a generic dataset and most of the studies are domain-independent that need to be evaluated in diverse scenarios. DBpedia is one of the biggest datasets that is frequently updated as it obtains data from Wikipedia that continuously grows into one of the central knowledge sources [21]. It makes Dbpedia multimodal and suitable for RS that need to be domain-independent and for knowledge-based RS where complexity and cost of maintenance of the knowledge base is high. However for RS of a single domain, it is better to use specific datasets but always implementing a linking interface with generic datasets in order to resolve ambiguities or to exploit unknown semantic relationships.

RQ5 concerns the application domains considered by RS based on Linked Data so far. We identified 12 domains, but we found that most of the RS are domain-independent (slightly more than one fifth of the studies). This is because most of the proposed recommendation algorithms can be applied in diverse domains by only changing the dataset or taking only a portion of it in order to obtain the data to generate the recommendations.

However, we also note that items of music, tourism, and movies are the most recommended as these belong to common domains in which there is a large amount of data and state-of-the-art datasets available, which allow the researchers to compare their results with several works developed in the community.

Accordingly, in a number of cases, the domain impacts also on datasets because they require a reduction of information; that is, only a subset of concepts is considered, which requires offline processing and more effort to maintain the dataset even if it improves the performance. For example, Passant developed RS named *dbrec* [22], which required to manually extract a subset of the data of DBpedia related with bands and musical artists.

RQ6 regards the evaluation techniques used to study RS based on Linked Data. We classified them into two types: accuracy and computational complexity. Accuracy evaluates recommendations according to their relevance for final users, while computational complexity measures the execution time required to produce them.

With regard to accuracy, our results demonstrate that researchers are more interested in evaluations made by final users than in comparisons with similar methods. This result was expected because usefulness of recommendations depends more on final user preferences than on comparing with similar approaches where evaluation may be biased as researchers must trust the results obtained. Therefore, future methodologies of evaluation should be user-centered in order to assure the quality of the results of RS.

Additionally as expected, most of the selected studies were more likely to evaluate their recommendations applying traditional methods of information retrieval such as precision and recall that are focused on percentages of true positives, false negatives, and false positives.

Interestingly, we found that few works evaluated the computational complexity of RS, which is a critical factor specially for applications that need responses with short timeouts. Therefore, it is still an open issue considering that accessing Linked Data datasets in most cases is time consuming and requires that researchers download dumps of the datasets to access them in local repositories.

RQ7 aimed to uncover the most promising directions for future research on RS based on Linked Data. To address this issue, we have reported not only future works but also limitations of the selected studies.

DOI: 10.1002/cpe

Section 4.7 summarized the future work reported in the selected studies. We found that the most frequently future works were the personalization of recommendations, the use of more datasets, and the creation of hybrid RS.

The lack of personalization of recommendations is still a common drawback in Linked Data-based RS. It concerns the fact that different users obtain the same set of results with the same input parameters. To solve this drawback, some RS need explicit feed back from users in order to differentiate the results based on information about the user's profile (e.g., browsing history and favorite music genre).

However, these approaches force the user to perform extra work like rating items or building an exhaustive user profiles. Consequently, there is a need of non-invasive personalization approaches supported by Linked Data in order to obtain implicit information from the neighborhood relationships user-to-user, item-to-item, and user-to-item. These relationships can be inferred from the links between concepts of datasets in Linked Data cloud related with properties of items and users.

Using more datasets is needed in order to increase the base of knowledge to produce recommendations. As presented in Section 4.8, there are some limitations of the current Linked Data-based RS with regard to the use of Linked Data datasets such as restricted access, poor reliability, computational complexity, low coverage of languages, domain dependency, and the need for installing a local copy of the dataset. For this reason, it is important to investigate new ways to integrate different datasets in order to (*i*) extend the knowledge base allowing the RS to access to other datasets in case that the main dataset fails or the data are not reliable; (*ii*) create scalable RS because they can be adapted to other domains by only accessing to the appropriate dataset and (*iii*) improve the performance by selecting datasets with better response time.

The creation of hybrid RS is not a new proposal, as could be seen in Section 4.4, combining diverse techniques of recommendation with Linked Data-based approaches is a frequent practice in the selected studies. However, we also found that it is still an open issue because it is necessary to investigate which combinations of techniques are more suitable for RS applied in diverse contexts. For example, combining Linked Data-based RS with social-based RS can be a good choice for applications that require information about the users and their interrelationships. In this way, RS can access information that sometimes is not available in Linked Data datasets such as items rating information, user profiles, and other social information.

The inclusion of user profile information (user profiling) is another aspect that is not widely considered in Linked Data recommender systems. The idea behind the user profiling is to obtain a meaningful concept-driven representation of user preferences in order to enable more precise specifications of user's preferences with less ambiguity. Therefore, this can be also useful to contribute to the personalization of Linked Data-based RS.

The automatic selection of the appropriate dataset according to the type of items or the application domain is another challenge that intend to improve the quality of recommendations. This dynamic process of selection can help the algorithms to choose the best strategy to find candidate items to be a recommender based on the implicit knowledge contained in Linked Data and the relationships with properties of items and users.

As a consequence, it is also important to study new similarity measures and techniques able to automatically combine information from different datasets and to deal with the diversity of data in these datasets. Furthermore, it can be possible to create a statistical models of user interests to overcome the topical diversity of rated items.

Finally, we found that there is still a need for building testbeds in order to allow for rigorous, transparent, and replicable testing and for studying new techniques (or adaptation of those existing) for evaluating the accuracy and computational complexity of RS based on Linked Data. This must also consider that Linked Data-based RS may have access to large amounts of information and that links among items can be unknown to the users. Additionally, large-scale RS should be also evaluated in terms of the ability to scale and provide recommendations with data coming from millions of users and/or items

DOI: 10.1002/cpe

5.2. *Limitations of our systematic literature review*

This section describes the main limitations we faced during our systematic literature review. Firstly, although some of selected papers were initially included because of their title or abstract, in the end they were excluded because we could not access them from our University.

Secondly, we only considered the most relevant paper for each study in order to calculate the frequency of problems, future work, contributions, and evaluation techniques. As a consequence, we could be biased, as some papers belonging to the same study may present a problem or contribution not reported in the most relevant paper.

Finally, we did not perform deep validation. Because of time issues, the majority of studies were read by one researcher, and cross-checking was performed only on about one third of the studies. Nonetheless, for some papers for which assessment was difficult, there was a discussion between the first three authors.

6. CONCLUSIONS

This systematic review has discussed 69 papers reporting 52 primary studies addressing RS that make use of the structured data published as Linked Data. We focused on identifying the most relevant problems that these studies aimed to solve and how they used Linked Data to provide recommendations. Although some of our results are already known, we defined a protocol to support our assumptions. Furthermore, we analyzed contributions, limitations, application domains, evaluation techniques they applied to assess their results, and the proposed directions for future research.

With regard to the research problems, we found that the most relevant ones were the lack of semantic information and the complexity of information about items. In order to overcome the lack of semantics, RS are enriched with diverse Linked Data datasets that are useful to describe users and items while reducing the ambiguity and exploiting the vast amount of links between related concepts stored in these datasets.

The majority of the selected studies have addressed these problems using Linked Data for several purposes, such as (*i*) finding new relationships or similarities based on links, paths, graphs, and created on the basis of Linked Data; (*ii*) generating serendipitous recommendations, that is, recommending items that are not expected by the users because of the links uncovered once the items are enriched with Linked Data; and (*iii*) explaining the recommendations, that is, allowing users to understand the reason of a recommendation by following the paths among items in the Linked Data cloud.

We also provided a classification of the selected studies according to the way they use Linked Data to provide recommendations. In particular, we identified four classes: Linked Data driven RS, which rely on techniques applied on datasets in Linked Data cloud such as categories, paths, number of input, and output links; hybrid RS that combine traditional techniques of recommendation (e.g., collaborative filtering and content based) with Linked Data; representation-only RS that uses Linked Data only to represent items or users but not for recommendations; and exploratory search systems that are not RS but help users to discover content through a guided search and are specially useful for users interested in learning or investigating a topic.

Additionally, we studied the most common datasets that RS use in order to obtain information, and we found that more than a half of these studies rely on DBpedia. This is because DBpedia is considered a central hub for the Linked Data cloud; it is linked to various datasets that gives the possibility to access diverse data from different application domains. Additionally, it makes DBpedia suitable for testing purposes in generic RS.

Concerning the evaluation techniques, the majority of the selected studies are focused on accuracy and rely more often on *user studies* than *comparison with other methods*. Computational complexity is also assessed in few cases; however, we think that it is an important factor to be evaluated especially for applications needing short responses such as RS in mobile environments. Additionally, we found that there is still a need for building testbeds to allow for testing and studying the results of RS based on Linked Data.

DOI: 10.1002/cpe

According to our findings, we identified that two recurrent issues in the selected studies are the high computational demand and the domain dependency. Therefore, we believe that further research is still needed to offer non-invasive personalization, exploit more datasets, and improve performance. Additionally, future work should focus on providing evaluation of RS considering the accuracy and computational complexity. With regard to application domains, music, movies, and tourism items are the most used in RS, and this may be due to the fact that in these domains, there are more datasets that help scientists to assess the results of their RS in comparison with similar approaches.

Finally, it is worth to mention that currently, we are working in the area of RS; in particular, we are developing RS that uses Linked Data as a source of information to recommend items for multiple application domains. The currently obtained results have been presented in [23], in which we describe how the RS based on Linked Data can be applied in the eTourism domain.

APPENDIX A. SELECTED PAPERS

Rows in italics identify papers (P) belonging to a study (S) already reported by other paper (e.g., papers 10, 19, and 54 belong to the same study S10).

Table A.1. Selected papers (P) and corresponding studies (S).

P	S	Authors	Year	Title	Publication details
1	S1	Fernández-Tobías, I., Cantador, I., Kaminskas, M., Ricci, F.	2011	A generic semantic-based framework for cross-domain recommendation	2nd International Workshop on Information Heterogeneity and Fusion in Recommender Systems - HetRec '11, pp 25 - 32
2	S2	Kabutoya, Y., Sumi, R., Iwata, T., Uchiyama, T., Uchiyama, T.	2012	A Topic Model for Recommending Movies via Linked Open Data	International Conferences on Web Intelligence and Intelligent Agent Technology, pp 625–630
3	S3	Dell'Aglio, D., Celino, I., Cerizza, D.	2010	Anatomy of a Semantic Web-enabled Knowledge-based Recommender System	4th international workshop Semantic Matchmaking and Resource Retrieval in the Semantic Web, at the 9th International Semantic Web Conference, pp 115–130
4	S4	Mannens, E., Coppens, S., Wica, I., Dacquin, H., Van De Walle, R.	2013	Automatic News Recommendations via aggregated Profiling	Journal Multimedia Tools and Applications, 63 (2), pp 407–425
5	S5	Dzikowski, J., Kaczmarek, M.	2012	Challenges in Using Linked Data within a Social Web Recommendation Application to Semantically Annotate and Discover Venues	International Cross Domain Conference and Workshop, pp 360–374
6	S6	Wardhana, A.T.A.; Nugroho, H.T.	2013	Combining FOAF and Music Ontology for Music Concerts Recommendation on Facebook Application	Conference on New Media Studies, pp 1–5
7	S7	Passant, A., Raimond, Y.	2008	Combining Social Music and Semantic Web for music-related recommender systems	First Workshop on Social Data on the Web, pp 19–30
8	S8	Lindley, A., Graf, R.	2011	Computing Recommendations for Long Term Data Accessibility basing on Open Knowledge and Linked Data	5th ACM Conference on Recommender Systems, pp 51–58

DOI: 10.1002/cpe

Table A.1. *Continued.*

P	S	Authors	Year	Title	Publication details
9	S9	Passant, Alexandre	2010	dbrec–Music Recommendations Using DBpedia	The Semantic Web–ISWC 2010, pp 209–224
10	S10	Stankovic, M., Breitfuss, W., Laublet, P.	2011	Discovering Relevant Topics Using DBPedia: Providing Non-obvious Recommendations	2011 International Conferences on Web Intelligence and Intelligent Agent Technology, 1, pp 219–222
11	S11	Marie, N., Gandon, F., Ribière, M., Rodio, F.	2013	Discovery Hub : on-the-fly linked data exploratory	9th International Conference on Semantic Systems, pp 17–24 search
12	S12	Peska, L., Vojtas, P.	2013	Enhancing Recommender System with Linked Open Data	10th International Conference on Flexible Query Answering Systems, pp 483–494
13	S13	Di Noia, T., Mirizzi, R., Ostuni, V. C., Romito, D.	2012	Exploiting the web of data in model-based recommender systems	6th ACM conference on Recommender systems
14	S14	Golbeck, J.	2006	Filmtrust: movie recommendations from semantic web-based social networks	3rd IEEE Consumer Communications and Networking Conference, pp 1314–1315
15	S15	Celma, Ò., Serra, X.	2008	FOAFing the music: Bridging the semantic gap in music recommendation	Web Semantics: Science, Services and Agents on the World Wide Web, 6 (4), 250–256
16	S16	Varga, B., Groza, A.	2011	Integrating DBpedia and SentiWordNet for a tourism recommender system	7th International Conference on Intelligent Computer Communication and Processing, pp 133–136
17	S17	Kaminskas, M., Fernández-Tobías, I., Ricci, F., Cantador, I.	2012	Knowledge-based music retrieval for places of interest	Proceedings of the second international ACM workshop on Music information retrieval with user-centered and multimodal strategies–MIRUM ’12, pp 19–24
18	S18	Dietze, S.	2012	Linked Data as facilitator & practice for TEL recommender systems in research	2nd Workshop on Recommender Systems for Technology Enhanced Learning, pp 7–10
19	*S10*	*Damljanovic, D., Stankovic, M., Laublet, P.*	*2012*	*Linked Data-Based Concept Recommendation : Comparison of Different Methods*	*9th Extended Semantic Web Conference, pp 24–38*
20	S19	Kitaya, K., Huang, H. H., Kawagoe, K.	2012	Music curator recommendations using linked data	Second International Conference on the Innovative Computing Technology, pp 337–339
21	S20	Jung, K., Hwang, M., Kong, H., Kim, P.	2005	RDF Triple Processing Methodology for the Recommendation System Using Personal Information	International Conference on Next Generation Web Services Practices, pp 241–246
22	S21	Calì, A., Capuzzi, S., Dimartino, M. M., Frosini, R.	2013	Recommendation of Text Tags in Social Applications Using Linked Data	ICWE 2013 Workshops
23	*S21*	*Calì, A., Capuzzi, S., Dimartino, M. M., Frosini, R.*	*2013*	*Recommendation of Text Tags Using Linked Data*	*3rd International Workshop on Semantic Search Over the Web, pp 1–3*
24	S22	Meymandpour, R., Davis, J. G.	2012	Recommendations using linked data	5th Ph.D. workshop on Information and knowledge–PIKM ’12, pp 75–82
25	S23	Harispe, S., Ranwez, S., Janaqi, S., Montmain, J.	2013	Semantic Measures Based on RDF Projections: Application to Content-Based Recommendation Systems	On the Move to Meaningful Internet Systems: OTM 2013 Conferences SE–44, pp 606–615

Table A.1. *Continued.*

P	S	Authors	Year	Title	Publication details
26	S24	Hopfgartner, F., Jose, J. M.	2010	Semantic user profiling techniques for personalised multimedia recommendation	Multimedia Systems, 16 (4-5) pp 255–274
27	*S5*	*Łazaruk, S., Dzikowski, J., Kaczmarek, M., Abramowicz, W.*	*2012*	*Semantic Web Recommendation Application*	Federated Conference on Computer Science and Information Systems (FedCSIS), pp 1055–1062
28	S25	Ostuni, V. C., Di Noia, T., Di Sciascio, E., Mirizzi, R.	2013	Top-N recommendations from implicit feedback leveraging linked open data	Proceedings of the 7th ACM conference on Recommender systems, pp 85–92
29	S26	Ahn, J., Amatriain, X.	2010	Towards Fully Distributed and Privacy-Preserving Recommendations via Expert Collaborative Filtering and RESTful Linked Data	International Conference on Web Intelligence and Intelligent Agent Technology, pp 66–73
30	S27	Heitmann, B., Hayes, C.	2010	Using Linked Data to Build Open, Collaborative Recommender Systems	AAAI Spring Symposium: Linked Data Meets Artificial Intelligence, pp 76–81
31	S28	Zarrinkalam, F., Kahani, M.	2012	A multi-criteria hybrid citation recommendation system based on linked data	2nd International eConference on Computer and Knowledge Engineering (ICCKE), 2012, pp 283–288
32	S29	Lommatzsch, A., Kille, B., Kim, J. W., Albayrak, S.	2013	An Adaptive Hybrid Movie Recommender based on Semantic Data	10th Conference on Open Research Areas in Information Retrieval, pp 217–218
33	S30	Torres, D., Skaf-Molli, H., Molli, P.; Díaz, A.	2013	BlueFinder: Recommending Wikipedia Links Using DBpedia Properties	5th Annual ACM Web Science Conference, pp 413–422
34	S31	Ostuni, V. C., Di Noia, T., Mirizzi, R., Romito, D., Di Sciascio, E.	2012	Cinemappy : a Context-aware Mobile App for Movie Recommendations boosted by DBpedia	International Workshop on Semantic Technologies meet Recommender Systems & Big Data SeRSy 2012, pp 37–48
35	S33	Zhang, Y., Wu, H., Sorathia, V., Prasanna, V. K.	2008	Event recommendation in social networks with linked data enablement	15th International Conference on Enterprise Information Systems, pp 371–379
36	S34	Mirizzi, R., Di Noia, T.	2010	From exploratory search to web search and back	3rd workshop on Ph.D. students in information and knowledge management– PIKM '10, pp 39–46
37	S35	Khrouf, H., Troncy, R.	2013	Hybrid event recommendation using linked data and user diversity	Proceedings of the 7th ACM conference on Recommender systems, pp 185–192
38	S36	Bahls, D., Scherp, G., Tochtermann, K., Hasselbring, W.	2012	Towards a Recommender System for Statistical Research Data	2nd International Workshop on Semantic Digital Archives
39	S37	Cheng, Gong; Gong, Saisai; Qu, Yuzhong	2011	An Empirical Study of Vocabulary Relatedness and Its Application to Recommender Systems	10th International Conference on The Semantic Web – Volume Part I, pp 98–113
40	S38	Wang, Y., Stash, N., Aroyo, L., Gorgels, P., Rutledge, L., Schreiber, G.	2008	Recommendations based on semantically enriched museum collections	Web Semantics: Science, Services and Agents on the World Wide Web, 6 (4), 283–290
41	*S11*	*Marie, N., Gandon, F., Legrand, D., Ribière, M.*	*2013*	*Discovery Hub: a discovery engine on the top of DBpedia*	*3rd International* Conference on Web Intelligence, Mining and Semantics
42	S31	Di Noia, T., Mirizzi, R., Ostuni, V. C., Romito, D., Zanker, M.	2012	Linked open data to support content–based recommender systems	8th International Conference on Semantic Systems

Table A.1. *Continued.*

P	S	Authors	Year	Title	Publication details
43	*S31*	*Ostuni, Vito Claudio; Gentile, Giosia; Noia, Tommaso Di; Mirizzi, Roberto; Romito, Davide; Sciascio, Eugenio Di*	*2013*	*Mobile Movie Recommendations with Linked Data*	*International Cross–Domain Conference, pp 400–415*
44	*S31*	*Mirizzi, R., Di Noia, T., Ragone, A., Ostuni, V. C., Di Sciascio, E.*	*2012*	*Movie recommendation with DBpedia*	*3rd Italian Information Retrieval Workshop, pp 101–112*
45	S39	Waitelonis, J., Sack, H.	2011	Towards exploratory video search using linked data	Multimedia Tools and Applications, 59 (2), pp 645–672
46	S40	Li, S., Zhang, Y., Sun, H.	2010	Mashup FOAF for Video Recommendation LightWeight Prototype	7th Web Information Systems and Applications Conference, pp 190–193
47	S41	Hu, Y., Wang, Z., Wu, W., Guo, J., Zhang, M.	2010	Recommendation for Movies and Stars Using YAGO and IMDB	12th International Asia–Pacific Web Conference, pp 123–129
48	S42	Ruotsalo, T., Haav, K., Stoyanov, A., Roche, S., Fani, E., Deliai, R., Mäkelä, E., Kauppinen, T., Hyvönen, E.	2013	SMARTMUSEUM: A mobile recommender system for the Web of Data	Web Semantics: Science, Services and Agents on the World Wide Web, 20, pp 50–67
49	S43	Stankovic, M., Jovanovic, J., Laublet, P.	2011	Linked Data Metrics for Flexible Expert Search on the Open Web	8th Extended Semantic Web Conference, pp 108–123
50	S44	Ozdikis, O., Orhan, F., Danismaz, F.	2011	Ontology-based recommendation for points of interest retrieved from multiple data sources	International Workshop on Semantic Web Information Management, pp 1–6
51	S45	Debattista, J., Scerri, S., Rivera, I., Handschuh, S.	2012	Ontology-based rules for recommender systems	International Workshop on Semantic Technologies meet Recommender Systems & Big Data, pp 49–60
52	S46	Codina, V.; Ceccaroni, L.	2010	Taking Advantage of Semantics in Recommendation Systems	2010 Conference on Artificial Intelligence Research and Development, pp 163–172
53	*S9*	*Passant, A., Decker, S.*	*2010*	*Hey! Ho! LetŠs Go! Explanatory Music Recommendations with dbrec*	*7th Extended Semantic Web Conference, pp 411–415*
54	*S10*	*Stankovic, M.,* Breitfuss, W., Laublet, P.	*2011*	*Linked-data based* suggestion of relevant topics	*7th International* Conference on Semantic Systems, pp 49–55
55	*S9*	*Passant, A.*	*2010*	*Measuring semantic distance on linking data and using it for resources recommendations*	*AAAI Spring Symposium: Linked Data Meets Artificial Intelligence, pp 93–98*
56	*S14*	*Golbeck, J.*	*2006*	*Generating Predictive Movie Recommendations from Trust in Social Network*	*4th International Conference, iTrust 2006, pp 93–104*
57	*S39*	*Sack, H.*	*2009*	*Augmenting Video Search with Linked Open Data*	*International Conference on Semantic Systems, pp 550–558*
58	S47	Baumann, S., Schirru, R., Streit, B.	2011	Towards a Storytelling Approach for Novel Artist Recommendations	8th International Workshop, AMR 2010, Linz, Austria, August 17–18, 2010, Revised Selected Papers, pp 1–15
59	S48	Corallo, A., Lorenzo, G., Solazzo, G.	2006	A Semantic Recommender Engine Enabling an eTourism Scenario	10th International Conference, pp 1092–1101

DOI: 10.1002/cpe

Table A.1. *Continued.*

P	S	Authors	Year	Title	Publication details
60	S49	Nuzzolese, A. G., Presutti, V., Gangemi, A., Musetti, A., Ciancarini, P.	2013	Aemoo: Exploring Knowledge on the Web	Proceedings of the 5th Annual ACM Web Science Conference, pp 272–275
61	*S49*	*Musetti, A., Nuzzolese, A., Draicchio, F., Presutti, V., Blomqvist, E., Gangemi, A., Ciancarini, P.*	*2012*	*Aemoo: Exploratory Search based on Knowledge Patterns over the Semantic Web*	*Semantic Web Challenge*
62	*S47*	*Baumann, S., Schirru, R.*	*2012*	*Using Linked Open Data for Novel Artist Recommendations*	*13th Internal Society for Music Information Retrieval Conference*
63	S50	Cantador, I., Castells, P.	2006	Multilayered Semantic Social Network Modeling by Ontology-Based User Profiles Clustering: Application to Collaborative Filtering	Proceedings of 15th International Conference, pp 334–349
64	*S34*	*Mirizzi, R., Ragone, A., Di Noia, T., Di Sciascio, E.*	*2010*	*Ranking the Linked Data: The Case of DBpedia*	*10th International Conference, pp 337–354*
65	S51	Heitmann, B., Hayes, C.	2010	Enabling Case-Based Reasoning on the Web of Data	The WebCBR Workshop on Reasoning from Experiences on the Web at International Conference on Case-Based Reasoning
66	S52	Alvaro, G., Ruiz, C., Córdoba, C., Carbone, F., Castagnone, M., Gómez-Pérez, J. M., Contreras, J.,	2011	miKrow : Semantic Intra-enterprise Micro-Knowledge Management System	8th Extended Semantic Web Conference, pp 154–168
67	*S50*	*Cantador, I., Castells, P., Bellogín, A.*	*2011*	*An Enhanced Semantic Layer for Hybrid Recommender Systems: Application to News Recommendation*	*Int. J. Semant. Web Inf. Syst., 7 (1), pp 44–78*
68	S32	Cantador, I., Konstas, I., Jose, J. M.	2011	Categorising social tags to improve folksonomy-based recommendations	Web Semantics: Science, Services and Agents on the World Wide Web, 9 (1), pp 1–15
69	*S29*	*Lommatzsch, A., Kille, B., Albayrak, S.*	*2013*	*A Framework for Learning and Analyzing Hybrid Recommenders based on Heterogeneous Semantic Data Categories and Subject Descriptors*	*10th Conference on Open Research Areas in Information Retrieval, pp 137–140*

REFERENCES

1. Hill W, Stead L, Rosenstein M, Furnas G. Recommending and evaluating choices in a virtual community of use. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’95: ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 1995; 194–201.
2. Resnick P, Iacovou N, Suchak M, Bergstrom P, Riedl J. GroupLens: an open architecture for collaborative filtering of netnews. *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work—CSCW ’94*, 1994; 175–186.
3. Shardanand U, Maes P. Social information filtering: algorithms for automating “word of mouth”. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’95, ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 1995; 210–217.
4. Adomavicius G, Tuzhilin A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on* 2005; **17**(6):734–749.

DOI: 10.1002/cpe

5. Ricci F, Rokach L, Shapira B. Introduction to recommender systems handbook. In *Recommender Systems Handbook*. Springer, 2011.
6. Bizer C, Heath T, Berners-Lee T. Linked data—the story so far. *International Journal on Semantic Web and Information Systems* 2009; **5**(3):1–22.
7. Kitchenham B. Procedures for performing systematic reviews, Keele University: Eversleigh, Australia, 2004.
8. Kitchenham B, Charters S. Guidelines for performing systematic literature reviews in software engineering, University of Durham: Durham, UK, 2007.
9. Candillier L, Jack K, Fessant F, Meyer F. State-of-the-art recommender systems. *Collaborative and Social Information Retrieval and Access-Techniques for Improved User Modeling* 2009:1–22.
10. Lops P, De Gemmis M, Semeraro G. Content-based recommender systems: state of the art and trends. In *Recommender Systems Handbook*. Springer, 2011.
11. Bobadilla J, Ortega F, Hernando A, Gutiérrez A. Recommender systems survey. *Knowledge-Based Systems* 2013; **46**(0):109 –132.
12. Damljanovic D, Stankovic M, Laublet P. Linked data-based concept recommendation: comparison of different methods in open innovation scenario. In *The Semantic Web: Research and Applications*, Simperl E, Cimiano P, Polleres A, Corcho O, Presutti V (eds)., Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2012.
13. Felfernig A, Jeran M, Ninaus G, Reinfrank F, Reiterer S. Toward the next generation of recommender systems: applications and research challenges. In *Multimedia Services in Intelligent Environments*. Springer, 2013.
14. Dell'Aglio D, Celino I, Cerizza D. Anatomy of a semantic web-enabled knowledge-based recommender system. *CEUR Workshop Proceedings*, Vol. 667, 2010; 115–130.
15. Cruzes DS, Dybå T. Recommended steps for thematic synthesis in software engineering. *Proceedings of the 2011 International Symposium on Empirical Software Engineering and Measurement*, ESEM '11, IEEE Computer Society, Washington, DC, USA, 2011; 275–284.
16. Mirizzi R, Di Noia T. From exploratory search to web search and back. *Proceedings of the 3rd Workshop on PhD Students in Information and Knowledge Management*, PIKM '10, ACM, New York, NY, USA, 2010; 39–46.
17. Beel J, Langer S, Genzmehr M, Gipp B, Breitinger C, Nürnberger A. Research paper recommender system evaluation: a quantitative literature survey. *Proceedings of the International Workshop on Reproducibility and Replication in Recommender Systems Evaluation*: ACM, 2013; 15–22.
18. Küster U, König-Ries B. Measures for benchmarking semantic web service matchmaking correctness, The Semantic Web: Research and Applications, 2010; 45–59.
19. Baumann S, Schirru R, Streit B. Towards a storytelling approach for novel artist recommendations. In *Adaptive Multimedia Retrieval. Context, Exploration, and Fusion*, Detyniecki M, Knees P, Nrnberger A, Schedl M, Stober S (eds)., Lecture Notes in Computer Science. Springer Berlin Heidelberg, January 2011.
20. McShane M, Nirenburg S, Beale S. NLP with reasoning and for reasoning. In *Ontology and the Lexicon: A Natural Language Processing Perspective*, Huang C, Calzolari N, Gangemi A, Lenci A, Oltramari A, Prevot L (eds). Cambridge University Press: Cambridge, 2010.
21. Urbani J, Maassen J, Drost N, Seinstra F, Bal H. Scalable RDF data compression with MapReduce. *Concurrency and Computation: Practice and Experience* 2013; **25**(1):24–39.
22. Passant A. Dbrec—music recommendations using DBpedia. In *The Semantic Web ISWC 2010*, Patel-Schneider PF, Pan Y, Hitzler P, Mika P, Zhang L, Pan JZ, Horrocks I, Glimm B (eds)., Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2010.
23. Rodríguez Rocha O, Figueroa C, Vagliano I, Moltchanov B. Linked data-driven smart spaces. In *Internet of Things, Smart Spaces, and Next Generation Networks and Systems*, vol. 8638, Balandin S, Andreev S, Koucheryavy Y (eds)., Lecture Notes in Computer Science. Springer International Publishing, 2014 (English).

DOI: 10.1002/cpe