Recommender Systems for Science: A Basic Taxonomy

Ali Ghannadrad, Morteza Arezoumandan, Leonardo Candela and Donatella Castelli

Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo" - Consiglio Nazionale delle Ricerche, Via G. Moruzzi, Pisa, 56121, Italy

Abstract

The ever-growing availability of research artefacts of potential interest for users calls for helpers to assist their discovery. Artefacts of interest vary for the typology, e.g. papers, datasets, software. User interests are multifaceted and evolving. This paper analyses and classifies studies on recommender systems exploited to suggest research artefacts to researchers regarding the type of algorithm, users and their representations, item typologies and their representation, and evaluation methods used to assess the effectiveness of the recommendations. This study found that most of the current scientific artefacts recommender system focused only on recommending paper to individual researchers, just a few papers focused on dataset recommendation and software recommender system is unprecedented.

Keywords

Recommender systems, Survey and overview, Systematic literature review, Science artefact

1. Introduction

Open Science is promising to widely extend the already rich set of research artefacts representing human knowledge and research findings. Literature, for decades the primary means to convey research activity and discoveries, is nowadays accompanied and complemented by datasets, software, workflows, and other research artefacts representing items of interest for research activities. The volume of scientific artefacts has increased significantly in recent years, and this trend is destined to grow in volume and velocity. Hence, it becomes difficult for researchers to discover relevant papers, datasets, software, etc. In such a situation, recommender systems come to assist the researchers. Recommender systems are software systems devised to recommend items to users based on their observed interest [1].

Several surveys on recommender systems have been published. Su and Khoshgoftaar [2] proposed a comprehensive review on Collaborative filtering methods, one of the most successful approaches to building recommender systems. Collaborative filtering uses the interests of a group of users to recommend to others with unknown preferences. They present collaborative filtering tasks and their main challenges, like data sparsity, scalability, synonymy, grey sheep, shilling attacks, privacy protection, etc. and the ability to overcome challenges. Furthermore, they introduce three classes of collaborative filtering methods: memory-based, model-based,

and hybrid method. Khan et al. [3] discussed the landscape characterising Cross-Domain Recommender Systems (CDRS), i.e. systems exploiting knowledge built in a domain to recommend items in another domain. The goal of this research is manifold: (*a*) to recognise the most widely used CDRS building-block definitions, (*b*) to identify common features between them, (*c*) to categorise current research in the frame of identified definitions, (*d*) to group together research concerning algorithm types, and (*e*) to present existing problems and suggest future directions for CDRS research. Zhang et al. [4] described the efforts on deep learning-based recommender systems by proposing a taxonomy of deep learning-based recommendation methods. They presented a comprehensive summary of the state of the art and discussed the current trends by providing new viewpoints on this new field. Burke [5] surveyed hybrid recommender system techniques and presented a novel hybrid system that combines knowledge-based recommendation and collaborative filtering. They also show that semantic ratings coming from the knowledge-based part of the system increase the effectiveness of collaborative filtering. These selected surveys tend to focus on "generalist" recommender systems. Beel et al. [6] discussed more than 200 articles about research-paper recommender systems, highlighting advantages and disadvantages and proposing an overview of the most common recommendation concepts and approaches. Bai et al. [7] discussed algorithms, evaluation methods, and current issues in paper recommender systems. The issues were the cold start, sparsity, scalability, privacy, serendipity, and unified scholarly data standards. Färber and Jatowt [8] introduced the research in automatic citation recommendations by discussing the approaches, explaining the evaluation methods, and highlighting the challenges such as self-citations and the possible solutions.

As far as we know, no systematic literature survey has been performed to document the state of the art of recommender systems in science settings. In particular, no systematic study exists in contexts where the items to be recommended are all the scientific artefacts of potential interests in performing research activities. This paper provides a taxonomy regarding the scientific artefacts recommender systems stemming from a systematic mapping study of the current literature. At the first stage, the study identifies all of the literature concerning the recommender system for science with the Systematic Mapping Study approach. Then the studies regarding the intentions of this paper are scrutinised. Finally, it proposed a taxonomy of recommender system algorithms, evaluation methods, users, and items widely used in recommender systems for scientific artefacts.

The paper is organised as follows. Section 2 describes the method used to carry out the study. Section 3 is then dedicated to the experimental analysis and presenting the taxonomy. Section 4 critically discusses the study and the findings. Finally, Section 5 summarises the conclusions achieved from the study and proposes a set of future work.

2. Methodology

A systematic literature review is one of the effective methods of comprehending the state of the art for a given research domain. The monitoring of a systematic and guided approach ensures the reliability of results and eases the process of collecting information. One of the most widely recognised methods for doing systematic literature review (SLR) is the one proposed by Kitchenham et al. [9]. In comparison to the standard literature review, SLR has several

advantages, including a well-defined methodology, a reduction of biases, a more comprehensive range of situations and contexts, etc.; on the other hand, it needs significant effort. Although systematic reviews in software engineering have concentrated on quantitative and empirical investigations, many techniques for synthesising qualitative research results exist [10]. Among these, "A systematic mapping study provides a structure of the type of research reports and results that have been published by categorising them and often gives a visual summary, the map, of its results" [10]. This research was carried out as a Systematic Mapping Study (SMS) offered by Petersen et al. [10]. An SMS consists of five main steps: (*i*) *Defining the research questions*: Research questions (RQs) direct the study; (*ii*) *Conducting search*: The search is typically conducted in different scientific repositories based on keywords extracted from the RQs; (*iii*) *Screening of papers*: Choosing the most relevant papers to the research subject; (*iv*) *Classification Scheme*: Classifying the primary papers; (*v*) *Data Extraction and Mapping Process*: Building the taxonomy.

2.1. Research questions

The definition of the research questions is critical for the SMS because these questions characterise the study's goal. As a result, it is necessary to construct a set of research questions to comprehend the existing literature. These questions guide all the stages of the research. For our study, the following research questions have been defined:

- **RQ1**: How are users (and their interests) represented?
- **RQ2**: What are the items of interest, and how are these items characterised?
- **RQ3**: Which recommender algorithms have been used?
- **RQ4**: Which evaluation methods have been used?

2.2. Conducting search

Searching for papers in scientific repositories has three main steps. The first step is finding the scientific repositories where the searches are conducted. The second step is selecting the keywords to build the search strings. The third step is constructing the search strings and gathering the results. We conducted our search in ACM, IEEEXplore, ScienceDirect, Springer, and Scopus databases. The mentioned data sources cover most conference proceedings and journals for recommender systems topics. The keywords to construct the search strings were driven by the research questions. After identifying the main keywords, we considered the synonym and related concepts of the keywords to have more robust search strings. The keywords and their synonyms are shown in Table 1.

Table 1
Keywords

Keyword	Synonym and related concepts
recommender	recommendation
Scientific products and science	Scientific - Researchers - scientists - articles - papers - datasets

In the third phase, we constructed the search strings with identified keywords to query scientific repositories. In this phase, the papers' titles were searched using the combination of the identified keywords and the boolean operators. After performing our search strategy, we identified 3787 primary papers. Table 2 shows the search queries for each scientific repository and the corresponding results.

Table 2
Search queries and primary results

Repository	Search Query	Primary result
ACM	[Title: recommend*] AND [[Title: papers] OR [Title: articles] OR [Title: datasets] OR [Title: science] OR [Title: researchers] OR [Title: scientific] OR [Title: scientists]]	235
IEEEXplore	("Document Title":recommend*) AND ("Document Title":researchers OR "Document Title":science OR "Document Title":scientific OR "Document Title":scientists OR "Document Title":datasets OR "Document Title":papers OR "Document Title":articles)	279
ScienceDirect	ttl("recommend* science") OR ttl("recommend* scientific") OR ttl("recommend* papers") OR ttl("recommend* articles") OR ttl("recommend* papers") OR ttl("recommend* researchers") OR ttl("recommend* scientists")	177
Springer	Recommend* AND (papers OR articles OR researchers OR datasets OR scientific OR science OR scientists)	332
Scopus	TITLE(Recommend*) AND TITLE(Datasets OR Researchers OR scientific OR science OR Articles OR Papers OR scientists)	2764
Total		3787

2.3. Screening of papers

After achieving the primary results, we had to massage the resulting papers corpus 3787 potential entries to distil into a suitable one. Duplications of the same study stemming from diverse databases were removed. The remaining entries were explored concerning publication type and publication date. As Figure 1a demonstrates, most of the publications are conference proceedings and journal articles. As Figure 1b shows, publications have increased in recent years.

Thus, in our study, we considered articles and conference proceedings as inclusion criteria and publications between 2015 and 2022 in terms of the year of publication. Table 3 shows the inclusion and exclusion criteria.

In the last step, we reviewed the titles and abstracts to remove the papers that were not coherent with our study goals. After reviewing the papers, we reached the final dataset containing 209 papers. The dataset is publicly available [11]. After removing duplicates, applying criteria and reviewing, the resulting corpus is structured as illustrated in Table 4.
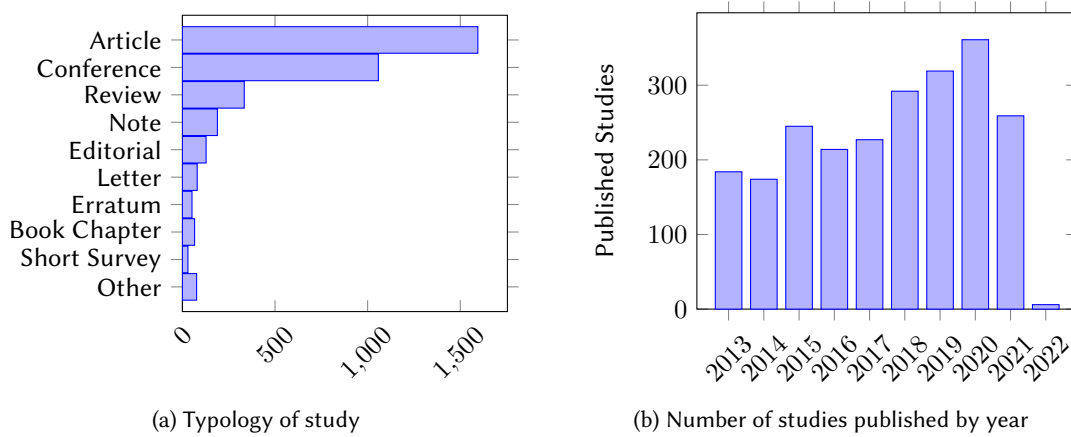
(a) Typology of study

(b) Number of studies published by year

Figure 1: Primary studies per typology and publication year

Table 3

Inclusion and Exclusion Criteria

Criterion	Inclusion Criteria	Exclusion Criteria
Publication Type	Article, proceedings	Thesis, books chapters
Period	Between 2015 and 2022	Before 2015
Language	English	Other languages

Table 4

Inclusion and Exclusion Criteria

	ACM	IEEEXplore	ScienceDirect	Springer	Scopus	Studies
After removing duplicates	114	64	152	40	2205	2575
After applying criteria	64	6	53	11	853	987
After reviewing	8	3	0	6	192	209

3. Analysis

To construct the classification scheme for categorising the studies, we tried to answer the research questions above (c.f. Sec. 2.1). We classified all the studies after studying their content.

The resulting taxonomy is depicted in Figure 2 and discussed in the remainder of the section.

After constructing the classification scheme, the relevant papers are categorised into the identified classes. This classification allows determining which areas have been targeted in previous studies and consequently identify gaps and opportunities for future research [10]. In the following sections, we explain each class and highlight the primary studies.
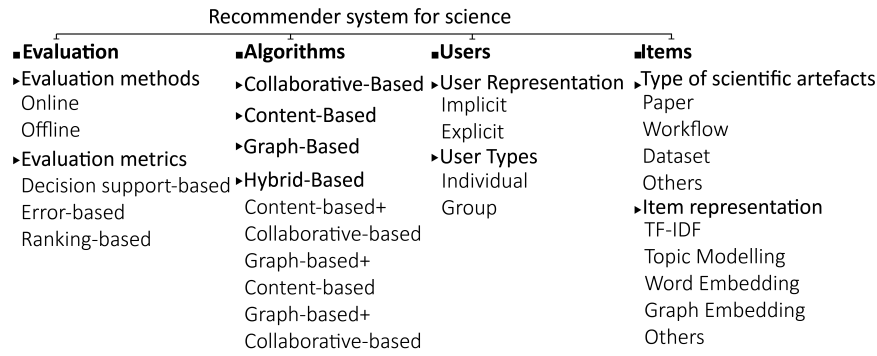
Figure 2: Taxonomy of Recommender system for science

3.1. Users types and Representations

Since group works are essential in scientific projects, we classified users receiving the recommended items as individual users or groups of users.

We identified only four papers having the group of researchers as the target for recommendations, or the interest of the group of users are considered [12, 13, 14, 15]. Asabere et al. [12] proposed a technique to recommend papers published by active participants in a conference to other group profile participants at the same conference based on the similarity of their research interests. They captured users' interests explicitly by asking them to express their preferences and implicitly by looking into the contact duration and contact frequency between active participants and other participants. Zhao et al. [13] used the user's research interest network, an undirected connected graph. The nodes of the graph represent the users, and the edge represents the similarity of interests between two users. A threshold-based method is applied to obtain groups (of users) with similar research interests based on this network. Wang et al. [14] use a hybrid approach in the individual prediction phase and an aggregation technique based on the ER rule in the group aggregation phase. Wang et al. [15] propose a group-oriented paper recommendation method based on probabilistic matrix factorisation and evidential reasoning. The proposed method has three main steps: individual paper recommendation, individual prediction aggregation, and group recommendation. In the individual paper recommendation step, the researcher similarity and paper similarity indicators are calculated. These indicators are integrated into the probabilistic matrix factorisation model to improve performance and produce a more accurate prediction for each group member. In the step of aggregating the individual predictions, the predicted rating of each group member is merged into a group rating by the rule ER. In the final step, group recommendation, the top-k papers with the highest predicted ratings are recommended to the target group.

The majority of studies (205) focus on recommending to the individual researcher.

There are different ways to represent a user (actually user's interests) for recommender systems. We categorised them into two main classes: *explicit representation* and *implicit representation*. In explicit representation, the system relies on the user's input: a query, paper, dataset, etc. To recommend a similar item, the user explicitly enters their preferences to receive the recommended items. For instance, the system proposed in [16] recommends articles based

on given input researchers' queries.

In implicit representation, the system captures users' interests implicitly. Most of the papers we identified relied on implicit representation. Although implicit representation requires more effort, the result is more reliable, and the user has a better experience. We classified implicit representation approaches into three classes: *user profile*, *graph representation*, and *hybrid*. In user profile based approaches, the system constructs a representation of the user preferences (the profile). For instance, the method proposed in [17] relies on relevant and non-relevant documents previously voted by the user. In graph-based approaches, the system relies on "connections" / "links" between nodes representing users. For instance, Huynh et al. [18] proposed an Academic Social Network model for modelling explicit and implicit relations in the academic field. In hybrid approaches, the system merges different representation methods. Hao et al. [19] proposed a method based on the author's periodic interest and academic network structure.

3.2. Items types and Representations

This section discusses the different typologies of items recommended to the researchers in the analysed studies. We identified 16 heterogeneous typologies of artefacts.

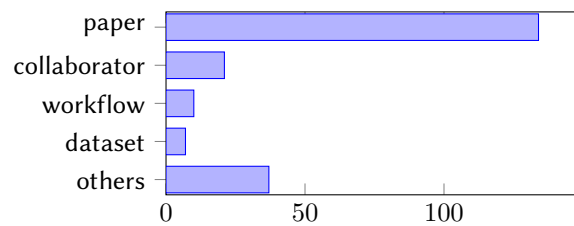As Figure 3 illustrates, out of 209 reviewed papers, 134 of them proposed a paper recommender system.

Figure 3: Typologies of recommended items in selected studies

Only seven papers focused on a method to recommend datasets to the researcher. For example, Altaf et al. [20] proposed a query-based dataset recommendation system that accepts users' interests and recommends relevant datasets. Another example is given in [21], where a three-layered network made of authors, papers, and datasets was used to estimate the proximity between authors and datasets.

We identified ten papers that proposed an algorithm to recommend workflows, i.e. representations of experiments' steps and the data flow across each of these steps [22]. Wen et al. [23] proposed a scientific workflow recommender system by using heterogeneous information networks and tags of scientific workflows. Several studies [24, 25, 26, 27] represent the workflow as layer hierarchy (i.e. specifies hierarchical relations between workflows, its sub-workflows and activities), and the semantic similarity is measured between layers of workflow. A graph-skeleton-based clustering technique is applied to cluster the layer hierarchies, then barycenters in each cluster are identified and used for workflow ranking and recommendation. Another type of workflow recommender system method is presented by [28]. Hou and Wen calculated

the similarity by using the tags besides the workflow descriptions, structures, and hierarchies; then, the workflows are clustered and recommended based on their similarity.

As it seems, the software recommender system for researchers is almost lacking. Only authors in [29] recommend GitHub repositories to the articles. Shao et al. developed a cross-platform recommender system, paper2repo, that can automatically recommend GitHub repositories that match a given paper. They proposed a joint model that incorporates a text encoding technique into a constrained graph convolutional networks formulation to generate joint embeddings of articles and repositories from the two different platforms. In particular, the text encoding technique was used to learn sequence information from paper abstracts and descriptions/tags of repositories.

The literature corpus also contains recommenders not focusing on research artefacts. For instance, we found 21 papers proposing a method to recommend a collaborator or reviewer to the researcher. We also find papers recommending items like "keyword", "tag", research area, paper submission, etc. We classified them as "Others".

As far as the representation of items is concerned, i.e. approaches aimed at capturing the characteristic features of objects, different methods have been used due to the diversity of scientific artefacts. Since it is possible to have a *text-based characterisation* (be it the artefact itself as for papers, or some metadata or texts accompanying the artefact for other typologies like datasets or workflows), we further analysed and classified text-based representations methods only. Any recommender system algorithm or machine learning algorithm requires transforming the text into numeric or vector representation. This vector representation has to illustrate the characteristics of the text. There are many techniques to represent a text for recommender system algorithms; in this project, we categorised them into five categories: *TF-IDF*, *Topic modelling*, *Word embedding*, *Graph embedding*, and *Mixed*. TF-IDF, Topic modelling, and Word embedding are only applied in Content-based filtering to represent the item's content. Graph embedding can be used in Graph-based algorithms like citation networks. Content-based and Graph-based algorithms are described in Sec. 3.3, respectively. TF-IDF is one of the most used techniques. However, it doesn't perform well in capturing semantic similarity of the documents; thus, Word embedding models can be used to gain more information about the paper [30].

TF-IDF   TF-IDF (term frequency-inverse document frequency) is a statistical measure that considers how relevant a word is to a document in a collection of documents [31].

This measure is computed by multiplying two metrics: (*i*) how many times a word appears in a document and (*ii*) how frequent is a word across a collection of documents. For instance, authors in [32] propose a method based on TF-IDF to represent users' past articles and then recommend articles based on cosine similarity. Patra et al. [33] presented a dataset recommender system; they transformed each dataset into a vector using TF-IDF. The title and summary were preprocessed and normalised for each dataset and converted into a single vector.

Topic modelling   The topic modelling class of approaches includes a class of statistical models to extract topics from a collection of text-based items. A topic is defined as a distribution over a fixed vocabulary [34]. Topic modelling can characterise a set of documents by clustering word groups and similar expressions. Our study found that Latent Dirichlet Allocation (LDA) is one

of the most used Topic modelling algorithms. A few studies apply Latent Semantic Analysis (LSA) to represent the items.

LDA is a probabilistic topic modelling that treats documents as bag-of-words. Training an LDA model requires starting with a set of documents that are represented as a fixed-length vector. The goal of the LDA model is to find the topic and document vectors. Zhao et al. [35] employ LDA to extract implicit topics in a domain and build the concept map using these topics.

Word Embedding   A word embedding is a learned representation for text where words with the same meaning have a similar representation.

The goal of word embeddings is to capture semantic and syntactic regularities, which could be hard to capture with a bag-of-words model like TF-IDF [30]. Words that occur in the same context are represented by vectors close to each other. There are several word embedding models like Word2Vec, Doc2Vec, Glove, etc. Two famous approaches of Word2Vec are CBOW and Skip-gram; the rationale behind both methods is that the neighbour words are semantically related [30]. Zhao et al. [21] propose a model to recommend reviewers to researchers; in this study, the submission or the reviewer is represented by a word embedding method, and the minimum distances between submissions and reviewers are measured. Zhao et al. [36] used Word embedding of the research papers to achieve the sequence of word vectors.

Graph Embedding   In some cases of Graph-based methods, each node represents an item. For example, in some paper recommender systems that use a citation network, nodes represent the papers, and edges represent papers cited by the paper. We called this type of item representation "Graph Embedding" in our classification. Tanner et al. [37] focused on a graph-based approach for recommending academic papers using citation networks incorporating citation relations.

Mixed   Although the item representation falls into one of the categories discussed above in most proposed systems, there are also a few other types of representation. Because they are heterogeneous, we decided to put them all in a category named "mixed". Examples of this category are preference matrices [38] and fuzzy methods [39]. In our systematic mapping study, we faced some papers that combined different item representations to address the shortcomings of each of them. Xia et al. [40] proposed a hybrid method for papers recommendation that combines Collaborative-filtering and Content-based algorithm approaches. The Content-based algorithm approach uses the social tag information to construct profiles for articles and researchers, and the Collaborative-filtering approach exploits the social friend information to help conduct unified probabilistic matrix factorisation. Khadka et al. [41] simultaneously use the citation graph and citation context.

3.3. Algorithms

Recommender system algorithms stemming from the reviewed papers can be classified into four classes: *Collaborative Filtering*, *Content-based*, *Graph-based*, and *Hybrid*. We found that out of 134 paper recommender systems, 56 applied hybrid methods, only 15 papers used Collaborative filtering algorithms, 37 papers proposed Content-based filtering, and 26 studies presented Graph-based algorithms.

Content-based Filtering (CBF)   Content-based filtering utilises item characteristics to recommend similar items to the user's profile that represent the user preferences [42]. In the recommender systems for researchers domain, items are article papers, datasets, software, etc. For using the CBF method, we need to build the profile of the items and researchers' profile which contains researchers' items like the papers written by researchers or the papers that researchers cited. As discussed in Sec. 3.2, recommender systems can use several ways to extract the insights and represent the items, like Word embedding and Topic modelling. After building the user's profile and the item's profile, CBF computes the similarity between the user profile and the item profile. After ranking the items regarding the similarity, the higher ranking items are recommended to researchers. Since the CBF is based on extracting the user profile features and then calculating the similarity between items profile and user profile, the recommended items are recognised better. On the other hand, similarity calculation requires considerable resources in the case of multiple users and items [6].

Collaborative Filtering approaches (CF)   The logic behind Collaborative Filtering (CF) is that researchers are interested in the items that similar researchers like [43]. In CF, two users are considered similar when they rated the same items. The user-item matrix is usually used in CF to represent the users' ratings. CF calculates the similarity between users based on the user-item matrix, finds similar users and eventually recommend items [7]. In comparison to the CBF, CF is independent of the content of the items [6]. In other words, there is no need to extract the features of the recommended item and build the user profile. Moreover, CF can provide serendipitous recommendations because the similarity is calculated between the users' relationships, not based on items and user profiles [6]. Serendipity in a recommendation system is the experience of receiving an unpredictable and surprising recommendation [44]. On the other hand, CF has some disadvantages like sparsity matrix and cold start problem [6].

Graph-based approaches (GB)   The graph-based method mainly focuses on building the graph representing connections among the concepts of interest. Citation networks (graphs connecting papers by using the references) and social networks (graphs connecting stakeholders) are usually exploited to construct the graph [7]. One of the most used types of graph-based is citation network which can be bibliographic coupling or co-citation. In bibliographic coupling, if two articles have common references are considered relevant. In co-citation, two papers are relevant if cited together in another paper [45]. The recommendation process of the graph-based recommender systems has two phases: the construction of the graph and the generation of recommendations [7]. The graph-based recommender system represents the users and items with a heterogeneous graph [7]. The recommendation progress will be performed as a graph search task in the graph-based model. In the case of the recommender system for scientific domains, one of the main types of Graph-based recommender systems commonly used is the citation network. The citation graph includes papers and the citation relationship between the papers. Papers are represented as graph nodes, and the edges describe the citation relationships between papers. The citation graph illustrates that if two papers have common references or are cited by the same paper, they are supposed to be alike [7].

Hybrid approaches Hybrid approaches propose to combine several approaches to recommend the scientific items to the researchers to enhance the overall accuracy [46]. In addition, hybrid recommender systems can address the shortcomings of individual algorithms. Different recommender system methods can be combined in diverse ways, like CF+CBF, CF+GB, etc. Of the 56 identified hybrid algorithms, the 34 papers propose a combination of CB+GB, 15 papers propose CB+CF, and two papers only apply CF+GB. For example, Berkani et al. [47] proposed a hybrid method that combines CF+CBF to construct a paper recommender system. Firstly, they constructed the profile of researchers and papers using CBF; after that, the social friend information was integrated into the CF method. CBF+GB is another type of combination widely used in recommender systems for papers is CBF+GB. Alshareef et al. [48] combined a paper's content with its bibliometrics to evaluate the citation impact of papers in their surrounding citation network. In addition, we found five studies that merged other types of algorithms. Du et al. [49] used the framework of one-shot learning to overcome the sparse user feedback and then propose an attention-based convolutional neural network model for text similarity to present a paper recommender system.

3.4. Evaluation

In this section, the evaluation methods and metrics which are primarily used in scientific artefacts recommendation systems are discussed.

By reviewing evaluation methods and metrics in papers that proposed paper recommender systems, we have the following result: for methods, 97 papers used offline methods, 27 papers used online methods, 3 papers used both and 7 papers used no evaluation method. Authors and studies prefer offline methods due to the complexity of online evaluation and the scarcity of publicly available datasets. For metrics, decision support metrics have been used in 65 papers, ranking-based evaluation methods have been used in 55 papers, and only four papers exploited error-based methods. We can conclude that decision support metrics and ranking-based evaluations are the most popular evaluation metrics among the scientists conducting research on paper recommender systems.

Evaluation Methods We classified the evaluation methods as the *Online* method and *Offline* method. Online evaluation approaches observe the user interactions regarding the given recommendations [50]. Offline evaluation approaches test the effectiveness of recommender system algorithms on a specific dataset.

Although the online evaluations can estimate the effectiveness of the recommender system in real-world use cases very well, they are costly and need extremely more time than offline evaluations [51]. Our results indicate that most of the scientific artefacts recommendation systems relied on offline evaluation, and just a few of them used both offline and online methods. For instance, Blank et al. [52] performed an offline evaluation and an online survey. The online survey was conducted using Google Doc to obtain user feedback. On the other hand, they tested all their proposed methods on their dataset as an offline evaluation method.

Evaluation Metrics Different evaluation metrics have been used to measure recommender systems' accuracy. In our reviewing process, we identified that decision support metrics

including Precision, Recall, and F1 are the frequent methods to measure the accuracy of the scientific artefacts recommender system. Precision is the fraction of relevant items among the recommended, Recall is the fraction of relevant items that are recommended, and the F1 score is the harmonic mean of the precision and recall [53].

Another type of evaluation metric is error-based such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). MAE is the average of the differences between the true value indicated by the user and the value predicted by the recommender system. MSE and RMSE are MAE variations summing the error after square it [53].

Ranking-based evaluation methods can measure the quality of the item's ranking. Ranking-based evaluation methods aim to evaluate the order of recommended items in terms of their relevance for the users. The most important Ranking based methods are Normalized Discounted Cumulative Gain (nDCG), Mean Reciprocal Rank (MRR), and Average Precision (AP). MRR focuses on where the first relevant item in the recommended list is. MRR for a list with the first relevant item at its third position will be greater than MRR for a list with the first relevant item at 4th position. Precision helps to understand the model's overall performance but doesn't measure the quality of the ranking. Average Precision measures the quality of the selected item's ranking of the recommender model. We found that MRR and nDCG are the most common metrics in ranking-based methods. We identified studies which used to evaluate characteristics such as serendipity of recommendations [54].

4. Discussion

This project proposed a taxonomy of the scientific artefacts recommender system. All in all, our review of the literature indicates several typologies of the studies. As we mentioned before, out of 209 reviewed papers, 134 papers presented a recommender system for paper, ten papers presented an algorithm to recommend workflows to the users, and seven of the identified papers proposed a method for recommending researchers a dataset. Since the greater part of the recommended scientific artefacts is text-based, we provided a classification for their representation. Although TF-IDF is one of the most used techniques, it doesn't perform well in capturing semantic similarity of the documents; thus, Word embedding models can be used to gain more information about the paper. In addition, the size of the TF-IDF vector changes by the number of words in the dictionary, thus calculation of the similarity needs a considerable amount of time given a long representation of the TF-IDF vector. Due to the growth of artificial neural networks, compact word representations using vector embedding have received much interest [30]. In these reviewed papers, only four papers presented a method to recommend scientific artefacts to a group of researchers, and most of them considered individual researchers.

In terms of algorithms, most paper recommender systems applied hybrid algorithms to overcome the disadvantages of using a single algorithm like cold start problems.

Concerning evaluation methods, offline methods were used primarily due to the many public datasets.

Our study has some limitations that should be mentioned. First, in identifying the papers, we only searched on the title of the papers and did not consider keywords or the content of the papers. Second, we did not perform snowballing to identify new papers possibly. Third, we

could not compare the papers' proposed methods due to the lack of details and the heterogeneity of terminology and context. Finally, since most of the studied works were focused on suggesting a method to recommend papers to scholars, we classified the algorithms and evaluation methods only in this case.

5. Conclusion and Prospects

This study had a Systematic Mapping Approach on the Recommender system for science. In particular, the study aims at responding to four questions on recommender systems in science cases: users and their interests representation, item typologies and their representation, recommendation algorithms and evaluation. A total of 209 papers of interest have been published between 2015 and 2022.

We found that the paper recommender system is the predominant recommendation class, and there is a considerable gap in recommending other scientific artefacts like datasets and software. Most of the current recommender system recommends items to the individual researcher, thus recommending to the group of researchers can be considered future work. As we identified, the software recommender system for researchers is unprecedented; therefore, one promising direction is to consider the software recommender system for future work. As we identified, most of the authors consider only the accuracy of the recommender systems. Since the interests of the researchers change frequently, other aspects like the diversity of the recommended items and serendipity are critical; recommending serendipitous scientific artefacts can open new doors to the scientists.

The development and diffusion of several research infrastructures represent a valuable context where recommender systems for science will play a role. Such systems could be developed by exploiting known approaches and latest developments borrowed from other application domains, as well as novel and peculiar ones could be devised to overcome known problems like cold-start, sparsity, scalability, and long-tail.

Acknowledgments

References

[1] P. Resnick, H. R. Varian, Recommender systems, Commun. ACM 40 (1997) 56–58. URL: https://doi.org/10.1145/245108.245121. doi:10.1145/245108.245121.

[2] X. Su, T. M. Khoshgoftaar, A survey of collaborative filtering techniques, Adv. in Artif. Intell. 2009 (2009). URL: https://doi.org/10.1155/2009/421425. doi:10.1155/2009/421425.

[3] M. M. Khan, R. Ibrahim, I. Ghani, Cross domain recommender systems: A systematic literature review, ACM Comput. Surv. 50 (2017). URL: https://doi.org/10.1145/3073565. doi:10.1145/3073565.

[4] S. Zhang, L. Yao, A. Sun, Y. Tay, Deep learning based recommender system: A survey and new perspectives, ACM Comput. Surv. 52 (2019). URL: https://doi.org/10.1145/3285029. doi:10.1145/3285029.

[5] R. Burke, Hybrid recommender systems: Survey and experiments, User Modeling and User-Adapted Interaction 12 (2002) 331–370. URL: https://doi.org/10.1023/A:1021240730564. doi:10.1023/A:1021240730564.

[6] J. Beel, B. Gipp, S. Langer, C. Breitinger, Research-paper recommender systems: a literature survey, International Journal on Digital Libraries 17 (2016) 305–338. URL: https://doi.org/10.1007/s00799-015-0156-0. doi:10.1007/s00799-015-0156-0.

[7] X. Bai, M. Wang, I. Lee, Z. Yang, X. Kong, F. Xia, Scientific paper recommendation: A survey, IEEE Access 7 (2019) 9324–9339. doi:10.1109/ACCESS.2018.2890388.

[8] M. Färber, A. Jatowt, Citation recommendation: approaches and datasets, Int J Digit Libr 21 (2020) 375–405. URL: https://doi.org/10.1007/s00799-020-00288-2. doi:10.1007/s00799-020-00288-2.

[9] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, S. Linkman, Systematic literature reviews in software engineering – a systematic literature review, Information and Software Technology 51 (2009) 7–15. URL: https://www.sciencedirect.com/science/article/pii/S0950584908001390. doi:10.1016/j.infsof.2008.09.009, special Section - Most Cited Articles in 2002 and Regular Research Papers.

[10] K. Petersen, R. Feldt, S. Mujtaba, M. Mattsson, Systematic mapping studies in software engineering, in: Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering, EASE'08, BCS Learning & Development Ltd., Swindon, GBR, 2008, pp. 68–77.

[11] M. Arezoumandan, A. Ghannadrad, L. Candela, D. Castelli, Recommender systems for science: A basic taxonomy, 2022. doi:10.5281/zenodo.6006905.

[12] N. Y. Asabere, F. Xia, Q. Meng, F. Li, H. Liu, Scholarly paper recommendation based on social awareness and folksonomy, International Journal of Parallel, Emergent and Distributed Systems 30 (2015) 211–232. URL: https://doi.org/10.1080/17445760.2014.904859. doi:10.1080/17445760.2014.904859. arXiv:https://doi.org/10.1080/17445760.2014.904859.

[13] H. Zhao, R. Zou, H. Duan, Q. Zeng, C. Li, X. Diao, W. Ni, N. Xie, An online paper recommendation system driven by user's interest model and user group, in: Proceedings of the 4th International Conference on Communication and Information Processing, ICCIP '18, Association for Computing Machinery, New York, NY, USA, 2018, pp. 141–144. URL: https://doi.org/10.1145/3290420.3290472. doi:10.1145/3290420.3290472.

[14] G. Wang, H.-R. Wang, Y. Yang, D.-L. Xu, J.-B. Yang, F. Yue, Group article recommendation based on er rule in scientific social networks, Applied Soft Computing 110 (2021) 107631. URL: https://www.sciencedirect.com/science/article/pii/S1568494621005524. doi:https://doi.org/10.1016/j.asoc.2021.107631.

[15] G. Wang, X. Zhang, H. Wang, Y. Chu, Z. Shao, Group-oriented paper recommendation with probabilistic matrix factorization and evidential reasoning in scientific social network, IEEE Transactions on Systems, Man, and Cybernetics: Systems (2021) 1–15. doi:10.1109/TSMC.2021.3072426.

[16] T. Chakraborty, N. Modani, R. Narayanam, S. Nagar, DiSCern: A diversified citation recommendation system for scientific queries, in: 2015 IEEE 31st International Conference on Data Engineering, 2015, pp. 555–566. doi:10.1109/ICDE.2015.7113314.

[17] T. Achakulvisut, D. E. Acuna, T. Ruangrong, K. Kording, Science concierge: A fast content-based recommendation system for scientific publications, PLOS ONE 11 (2016) 1–11. URL: https://doi.org/10.1371/journal.pone.0158423. doi:10.1371/journal.pone.0158423.

[18] T. Huynh, T.-T. Nguyen, H.-N. Tran, Exploiting social relations to recommend scientific publications, in: H. T. Nguyen, V. Snasel (Eds.), Computational Social Networks, Springer International Publishing, Cham, 2016, pp. 182–192.

[19] L. Hao, L. Shijun, L. Pan, Paper recommendation based on author-paper interest and graph structure, in: 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 2021, pp. 256–261. doi:10.1109/CSCWD49262.2021.9437743.

[20] B. Altaf, U. Akujuobi, L. Yu, X. Zhang, Dataset recommendation via variational graph autoencoder, in: 2019 IEEE International Conference on Data Mining (ICDM), 2019, pp. 11–20. doi:10.1109/ICDM.2019.00011.

[21] S. Zhao, D. Zhang, Z. Duan, J. Chen, Y.-p. Zhang, J. Tang, A novel classification method for paper-reviewer recommendation, Scientometrics 115 (2018) 1293–1313. URL: https://doi.org/10.1007/s11192-018-2726-6. doi:10.1007/s11192-018-2726-6.

[22] D. Silva Junior, E. Pacitti, A. Paes, D. de Oliveira, Provenance-and machine learning-based recommendation of parameter values in scientific workflows, PeerJ Computer Science 7 (2021) e606. URL: https://doi.org/10.7717/peerj-cs.606. doi:10.7717/peerj-cs.606.

[23] Y. Wen, J. Hou, Z. Yuan, D. Zhou, Heterogeneous information network-based scientific workflow recommendation for complex applications, Complexity 2020 (2020) 4129063. URL: https://doi.org/10.1155/2020/4129063. doi:10.1155/2020/4129063.

[24] Z. Zhou, Z. Cheng, L.-J. Zhang, W. Gaaloul, K. Ning, Scientific workflow clustering and recommendation leveraging layer hierarchical analysis, IEEE Transactions on Services Computing 11 (2018) 169–183. doi:10.1109/TSC.2016.2542805.

[25] Z. Cheng, Z. Zhou, P. C. Hung, K. Ning, L.-J. Zhang, Layer-hierarchical scientific workflow recommendation, in: 2016 IEEE International Conference on Web Services (ICWS), 2016, pp. 694–699. doi:10.1109/ICWS.2016.97.

[26] Z. Zhou, Z. Cheng, Y. Zhu, Similarity assessment for scientific workflow clustering and recommendation, Science China Information Sciences 59 (2016) 113101. URL: https://doi.org/10.1007/s11432-015-0934-9. doi:10.1007/s11432-015-0934-9.

[27] Z. Cheng, Z. Zhou, X. Wang, Scientific workflow clustering and recommendation, in: 2015 11th International Conference on Semantics, Knowledge and Grids (SKG), 2015, pp.

272–274. doi:10.1109/SKG.2015.52.

[28] J. Hou, Y. Wen, Utilizing tags for scientific workflow recommendation, in: J. H. Abawajy, K.-K. R. Choo, R. Islam, Z. Xu, M. Atiquzzaman (Eds.), International Conference on Applications and Techniques in Cyber Intelligence ATCI 2019, Springer International Publishing, Cham, 2020, pp. 951–958.

[29] H. Shao, D. Sun, J. Wu, Z. Zhang, A. Zhang, S. Yao, S. Liu, T. Wang, C. Zhang, T. Abdelzaher, Paper2repo: Github repository recommendation for academic papers, in: Proceedings of The Web Conference 2020, Association for Computing Machinery, New York, NY, USA, 2020, pp. 629–639. URL: https://doi.org/10.1145/3366423.3380145.

[30] B. Kazemi, A. Abhari, A comparative study on content-based paper-to-paper recommendation approaches in scientific literature, in: Proceedings of the 20th Communications & Networking Symposium, CNS '17, Society for Computer Simulation International, San Diego, CA, USA, 2017.

[31] S.-W. Kim, J.-M. Gil, Research paper classification systems based on tf-idf and lda schemes, Human-centric Computing and Information Sciences 9 (2019) 30. URL: https://doi.org/10.1186/s13673-019-0192-7. doi:10.1186/s13673-019-0192-7.

[32] B. Bulut, B. Kaya, R. Alhajj, M. Kaya, A paper recommendation system based on user's research interests, in: 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2018, pp. 911–915. doi:10.1109/ASONAM.2018.8508313.

[33] B. G. Patra, K. Roberts, H. Wu, A content-based dataset recommendation system for researchers–a case study on Gene Expression Omnibus (GEO) repository, Database 2020 (2020). URL: https://doi.org/10.1093/database/baaa064. doi:10.1093/database/baaa064, baaa064.

[34] C. B. Asmussen, C. Møller, Smart literature review: a practical topic modelling approach to exploratory literature review, Journal of Big Data 6 (2019) 93. URL: https://doi.org/10.1186/s40537-019-0255-7. doi:10.1186/s40537-019-0255-7.

[35] W. Zhao, R. Wu, W. Dai, Y. Dai, Research paper recommendation based on the knowledge gap, in: 2015 IEEE International Conference on Data Mining Workshop (ICDMW), 2015, pp. 373–380. doi:10.1109/ICDMW.2015.40.

[36] X. Zhao, H. Kang, T. Feng, C. Meng, Z. Nie, A hybrid model based on LFM and BiGRU toward research paper recommendation, IEEE Access 8 (2020) 188628–188640. doi:10.1109/ACCESS.2020.3031281.

[37] W. Tanner, E. Akbas, M. Hasan, Paper recommendation based on citation relation, in: 2019 IEEE International Conference on Big Data (Big Data), 2019, pp. 3053–3059. doi:10.1109/BigData47090.2019.9006200.

[38] T. T. Chen, M. Lee, Research paper recommender systems on big scholarly data, in: K. Yoshida, M. Lee (Eds.), Knowledge Management and Acquisition for Intelligent Systems, Springer International Publishing, Cham, 2018, pp. 251–260.

[39] S. M. Khatami, M. Maadi, R. Ramezani, A clustering expert system using particle swarm optimization and k-means++ for journal recommendation to publish the papers, Indonesian Journal of Electrical Engineering and Computer Science 12 (2018) 814–823. doi:10.11591/ijeecs.v12.i2.pp814-823.

[40] F. Xia, H. Liu, I. Lee, L. Cao, Scientific article recommendation: Exploiting common author

relations and historical preferences, IEEE Transactions on Big Data 2 (2016) 101–112. doi:`10.1109/TBDATA.2016.2555318`.

[41] A. Khadka, I. Cantador, M. Fernandez, Capturing and exploiting citation knowledge for recommending recently published papers, in: 2020 IEEE 29th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), 2020, pp. 239–244. doi:`10.1109/WETICE49692.2020.00054`.

[42] M. Amami, G. Pasi, F. Stella, R. Faiz, An LDA-based approach to scientific paper recommendation, in: E. Métais, F. Meziane, M. Saraee, V. Sugumaran, S. Vadera (Eds.), Natural Language Processing and Information Systems, Springer International Publishing, Cham, 2016, pp. 200–210. doi:`10.1007/978-3-319-41754-7_17`.

[43] X. Q. Zhang, X. X. Liu, J. Guo, B. Y. Liu, D. G. Gan, A matrix factorization based recommendation algorithm for science and technology resource exploitation, in: 2020 IEEE/ACS 17th International Conference on Computer Systems and Applications (AICCSA), 2020, pp. 1–6. doi:`10.1109/AICCSA50499.2020.9316467`.

[44] D. Kotkov, S. Wang, J. Veijalainen, A survey of serendipity in recommender systems, Knowledge-Based Systems 111 (2016) 180–192. URL: https://www.sciencedirect.com/science/article/pii/S0950705116302763. doi:`10.1016/j.knosys.2016.08.014`.

[45] A. Shahid, M. T. Afzal, A. Alharbi, H. Aljuaid, S. Al-Otaibi, In-text citation's frequencies-based recommendations of relevant research papers, PeerJ Computer Science (2021). doi:`10.7717/peerj-cs.524`.

[46] A. Tsolakidis, E. Triperina, C. Sgouropoulou, N. Christidis, Research publication recommendation system based on a hybrid approach, in: Proceedings of the 20th Pan-Hellenic Conference on Informatics, PCI '16, Association for Computing Machinery, New York, NY, USA, 2016. URL: https://doi.org/10.1145/3003733.3003805. doi:`10.1145/3003733.3003805`.

[47] L. Berkani, R. Hanifi, H. Dahmani, Hybrid recommendation of articles in scientific social networks using optimization and multiview clustering, in: M. Hamlich, L. Bellatreche, A. Mondal, C. Ordonez (Eds.), Smart Applications and Data Analysis, Springer International Publishing, Cham, 2020, pp. 117–132.

[48] A. M. Alshareef, M. F. Alhamid, A. El Saddik, Toward citation recommender systems considering the article impact in the extended nearby citation network, Peer-to-Peer Networking and Applications 12 (2019) 1336–1345. URL: https://doi.org/10.1007/s12083-018-0687-4. doi:`10.1007/s12083-018-0687-4`.

[49] Z. Du, J. Tang, Y. Ding, POLAR: Attention-based CNN for one-shot personalized article recommendation, in: M. Berlingerio, F. Bonchi, T. Gärtner, N. Hurley, G. Ifrim (Eds.), Machine Learning and Knowledge Discovery in Databases, Springer International Publishing, Cham, 2019, pp. 675–690.

[50] S. Vrijenhoek, M. Kaya, N. Metoui, J. Möller, D. Odijk, N. Helberger, Recommenders with a mission: Assessing diversity in news recommendations, in: Proceedings of the 2021 Conference on Human Information Interaction and Retrieval, Association for Computing Machinery, New York, NY, USA, 2021, pp. 173–183. URL: https://doi.org/10.1145/3406522.3446019. doi:`10.1145/3406522.3446019`.

[51] T. Silveira, M. Zhang, X. Lin, Y. Liu, S. Ma, How good your recommender system is? a survey on evaluations in recommendation, International Journal of Machine Learning and Cybernetics 10 (2019) 813–831. URL: https://doi.org/10.1007/s13042-017-0762-9. doi:`10.`

1007/s13042-017-0762-9.

[52] I. Blank, L. Rokach, G. Shani, Leveraging metadata to recommend keywords for academic papers, Journal of the Association for Information Science and Technology 67 (2016) 3073–3091. URL: https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/asi.23571. doi:10.1002/asi.23571.

[53] J. L. Herlocker, J. A. Konstan, L. G. Terveen, J. T. Riedl, Evaluating collaborative filtering recommender systems, ACM Trans. Inf. Syst. 22 (2004) 5–53. URL: https://doi.org/10.1145/963770.963772. doi:10.1145/963770.963772.

[54] C. Nishioka, J. Hauke, A. Scherp, Research paper recommender system with serendipity using tweets vs. diversification, in: A. Jatowt, A. Maeda, S. Y. Syn (Eds.), Digital Libraries at the Crossroads of Digital Information for the Future, Springer International Publishing, Cham, 2019, pp. 63–70. doi:10.1007/978-3-030-34058-2_7.