

REVIEW ARTICLE

Open Access



# Recommender systems to support learners' Agency in a Learning Context: a systematic review

Michelle Deschênes 

Correspondence: [michelle.deschenes.1@ulaval.ca](mailto:michelle.deschenes.1@ulaval.ca)  
Université Laval, 2325 rue de  
l'Université, Québec, QC G1V 0A6,  
Canada

## Abstract

Recommender systems for technology-enhanced learning are examined in relation to learners' agency, that is, their ability to define and pursue learning goals. These systems make it easier for learners to access resources, including peers with whom to learn and experts from whom to learn. In this systematic review of the literature, we apply an Evidence for Policy and Practice Information (EPPI) approach to examine the context in which recommenders are used, the manners in which they are evaluated and the results of those evaluations. We use three databases (two in education and one in applied computer science) and retained articles published therein between 2008 and 2018. Fifty-six articles meeting the requirements for inclusion are analyzed to identify their approach (content-based, collaborative filtering, hybrid, other) and the experiment settings (accuracy, user satisfaction or learning performance), as well as to examine the results and the manner in which they were presented. The results of the majority of the experiments were positive. Finally, given the results introduced in this systematic review, we identify future research questions.

**Keywords:** Recommender systems, Collaborative filtering, Content-based filtering, Systematic review

## Introduction

Recommender systems are “tools and techniques that suggest items that are most likely of interest to a particular user” (Ricci, Rokach, & Shapira, 2015, p. 1). They constitute a powerful method to help users to filter for the products that are most likely to be chosen, out of a very large number of products. They use algorithms that account for, among other elements, the user's browsing patterns, searches, purchases and preferences (Konstan & Riedl, 2012). Research into recommender systems is evolving rapidly, and such systems are being leveraged in more and more specific domains (Ekstrand, Riedl, & Konstan, 2011), particularly in the domain of technology-enhanced learning (TEL) given the digitalisation of learning and the growth of educational data (Drachsler, Verbert, Santos, & Manouselis, 2015).



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Literature reviews published in recent years about recommender systems in education have considered specific approaches and methods. For example, Tarus, Niu, and Mustafa (2018) looked at ontology-based recommenders. Although they identified that ontology-based recommendation combined with other recommendation techniques is widely used to recommend learning resources, they did not closely examine the techniques that could be used with this recommendation. Other reviews considered the application domain. Verbert et al. (2012) presented a context framework for TEL, with different contextual variables: computing, location, time, physical conditions, activity, resource, user, and social relations. They surveyed 22 recommender systems, not papers in which these systems were the subject of experiments, and they did not explain how they chose the systems. Rahayu et al. (2017) conducted a systematic review of recommender systems in a study on the pedagogical use of ePortfolios. Furthermore, others have looked at issues related to recommender systems: Camacho and Alves-Souza (2018) compiled papers using social network data to mitigate the cold-start problem, which is caused by not having enough items or learners to initiate a recommender system (Tang & McCalla, 2009). Erdt, Fernandez, and Rensing (2015) conducted a quantitative survey in which they discussed methods for evaluating recommender systems (type, subject and effects). Although the results of the aforementioned studies are enlightening, it should be noted that there is a lack of information regarding search strategies (e.g., descriptors used for selecting articles) and criteria for inclusion and exclusion. Finally, the chapter by Drachsler et al. (2015) in *The Recommender Systems Handbook*, “Panorama of Recommender Systems to Support Learning,” mainly provided an overview of recommender systems from 2010 to 2014. When they considered more recent publications based on empirical data, they did not mention the search strategy.

Recommender systems are guides that can help teachers find solutions to their documented needs in a context where teachers are responsible for their own professional development and showing agency, a concept we define below (Deschênes & Laferrière, 2019). Teachers have reported that they wish to receive recommendations for resources or people based on their preferences, goals and search topics or recommendations for resources other teachers found interesting and useful. We therefore wanted to examine whether recommender systems that recommend learning resources or people (peers or experts) are a promising solution.

We focused on articles discussing recommender systems that support the recommendation of resources in a learning context. To this end, we identified and analyzed 878 articles and retained 56. This paper is structured as follows: we first describe the recommender systems for TEL, particularly the different techniques and ways to evaluate the systems. Then, we present the methodology we used to perform the systematic review. Finally, we present our results and we discuss the main findings, identify their limitations and provide recommendations for future research.

### **Recommender systems for technology-enhanced learning**

The role of recommender systems is twofold (Ekstrand et al., 2011). Its first task is to predict value: how much will the user value a given resource? The answer is usually presented on the same scale as the one used by the evaluation system (for example, a number of stars). The second task is to recommend resources: what are the N resources most likely to be of high value to a given user? The answer is presented as a

top-N recommendation, or simply sorted by expected value, such as when performing a search in a search engine, with the most relevant results at the top.

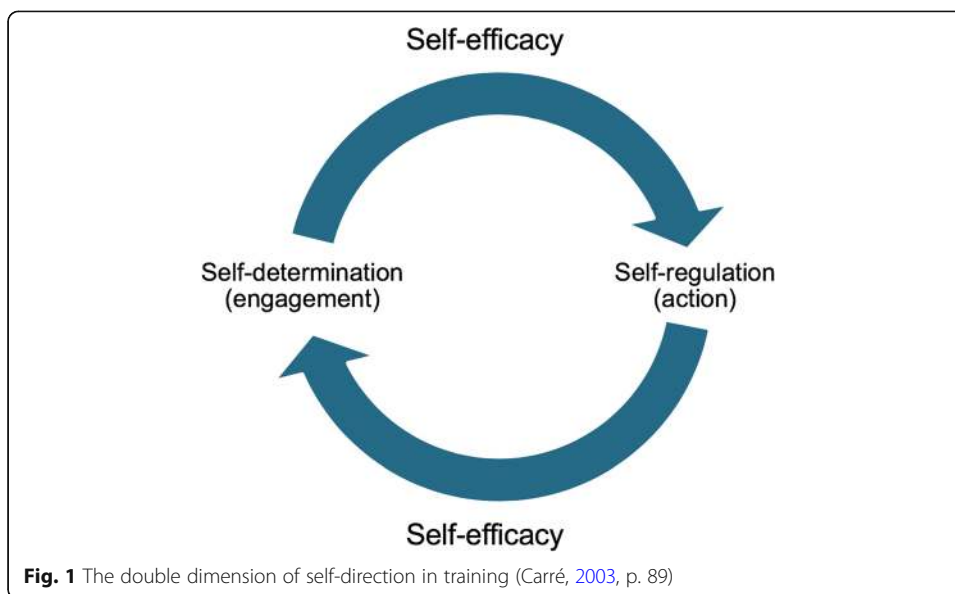
Preferences may be expressed either explicitly or implicitly. Implicit preferences are gathered through the user's actions. These actions are very telling about the user's preferences, even if they are unaware of it: clicking on a link (advertisement, search result or cross-reference), purchasing a product, following someone on a social network, spending time watching a video or listening to a song, etc. For explicit preferences, the system asks the user to rate an item. This can be done using a variety of paradigms, such as a scale of 0–5 stars (with or without half points), positive or negative votes, or only upvotes. These data are harder to collect, since they require user action and thus more effort compared to implicit data.

The rise of the use of recommender systems in various contexts (such as Netflix and Amazon) has also been seen in education, as mentioned in the chapters “Recommender Systems in Technology-Enhanced Learning” and “Panorama of Recommender Systems to Support Learning” in the 2011 and 2015 editions of the *Recommender Systems Handbook*. In the context of TEL, recommender systems can assist in carrying out a learning activity, viewing content, taking a course, joining a community, contacting a user, etc. (Santos & Boticario, 2015). However, the recommendation of resources in a learning context is different from the recommendation of products in a commercial setting (Winoto, Tang, & McCalla, 2012). On this topic, Drachler, Hummel, and Koper (2008) refer to Vygotsky (1978) zone of proximal development to support the theory that a recommender system should suggest resources that are slightly higher than the learner's current level.

### Learners' agency

Agency is defined as “[a] learner's ability to define and pursue learning goals” (Brennan, 2012, p. 24). Agency is manifested through self-directed learning behaviour, that is, “[a] process in which individuals take the initiative, with or without the help of others in diagnosing their learning needs, formulating learning goals, identifying human and material resources for learning, choosing and implementing appropriate learning strategies, and evaluating learning outcomes” (Knowles, 1975, p. 18). Agency exists at the intersection of self-determination (an autonomous, authentic free will to learn) and self-regulation (the exercise of agentic, self-controlled learning activity), a relationship that Jézégou (2013, p. 183) describes as interdependent. Carré (2003, p. 56) represented the articulation of these three concepts (Fig. 1) as follows:

Identifying human and material resources for learning presents a significant challenge: many resources exist, but not all are being used, or are even known. Self-regulated learners, that is, those who take control of their learning, must manage their resources, environment and context, as well as their tools (Butler, 2005; Mandeville, 2001). To assist them in recognizing and managing resources to further the goals they have chosen, we must provide the necessary resources, using the benefits identified by the community. Recommender systems are therefore an interesting avenue to consider in a context where we wish to support the user in their learning process while simultaneously acting within their zone of proximal development.



There is, however, some level of uncertainty regarding the processes involved in supporting a learner's agency. The structure (the rules, roles and resources, both explicit and assumed) required to support agency is a central question (Brennan, 2012). Accordingly, the tension between agency and the provided external structure must be a concern for designers of networked learning environments. Rather than set agency and structure in opposition to each other, Brennan maintains that they are mutually reinforcing concepts, and she proposes using structure to shape agency, converging at the concept of zone of proximal development, while also tying in the concept of "scripting," that is, structuring elements that support agency.

Brennan recommends five strategies that designers of learning environments can adopt to support agency. Among those, "support access to resources" implies making resources available at the right moment, in the right format and at a level fitting the expectations of the learner, whether or not those resources are centralized. Drachsler et al. (2008) argue that, in a context of lifelong learning, personal recommender systems in learning networks are necessary to guide learners in choosing suitable learning activities to follow.

Therefore, of the tasks that recommender systems can support (Drachsler et al., 2015), here we will examine the following: finding good items (content), finding peers, and suggesting learning activities. Tasks that do not fit well in the context of supporting agency, on the other hand, will not be covered in this review. One is the recommendation of learning paths: as agency "accounts for the individual's personal control and responsibility over his or her learning" (Carré, Jézégou, Kaplan, Cyrot, & Denoyel, 2011, p. 14), we wondered whether learning path recommendations would constrain rather than foster agency. Assuming that learners show agency when they determine, influence and personalize their learning paths, which are behaviors that manifest agency (Blaschke, 2018; Klemenčič, 2017), does a system restrain or expand agency when recommending learning paths? One may say, it depends whether the selection and sequencing of resources stimulate or not the learner's agency. That said, we recognize that the sequencing of items is an important part of regulation (Straka, 1999) and that it may have value at least for beginners. In the

learning sciences, this has been debated at length by those favourable to scripting, for example, online collaborative learning (Fischer, Kollar, Stegmann, & Wecker, 2013). This is why, to support the agency of learners, it seems more appropriate to present the range of resources that could allow them to achieve their goal, then let them negotiate and create a meaningful learning path for themselves, rather than to provide a predefined learning path. By doing so, we emphasize the principle of epistemic agency (Scardamalia, 2000; Scardamalia & Bereiter, 2006), which refers to the control people have over the resources they use to achieve their goals. The other task that will not be covered here is predicting learning performance, as the focus of our work does not necessarily take place in a formal context.

### **Recommendation techniques**

The principal recommendation techniques are content-based, collaborative filtering, and hybrid. Content-based recommender systems recommend items that are similar to the ones that the user has liked in the past (Ricci et al., 2015). These systems may use a case-based approach or an attribute-based approach (Drachsler et al., 2008). The first one assumes that if a user likes a certain item, this user will probably also like similar items, and the second one recommends items based on the matching of their attributes to the user profile.

Recommender systems based on collaborative filtering leverage the preferences of other users to provide a recommendation to a particular user. Such systems are called collaborative because they consider two items (book, movie, etc.) to be related based on the fact that many other users prefer these items, rather than by analyzing all of the attributes of the items (Konstan & Riedl, 2012). Many different methods may be used to analyze the items; two of them, user-based and item-based, are described below.

In a user-based approach, users who rated the same item similarly may have the same tastes (Drachsler et al., 2008). One can then recommend to a user items that are well rated, either by similar users or by users they trust. The user-based collaborative filtering concept requires calculating the distance between pairs of users, based on their level of agreement about items they have both rated (Konstan & Riedl, 2012). Systems based on this approach predict a user's appreciation for items by linking this user's preferences with those of a community of users who share the same preferences (Herlocker, Konstan, & Riedl, 2000).

The item-based collaborative filtering approach, on the other hand, recommends items that are preferred by similar users, based on user data instead of ratings (Drachsler et al., 2008). The system calculates the distance between each pair of items based on how closely users who have rated the items agree. This distance between pairs of items tends to be relatively stable over time, such that the distances can be pre-calculated, meaning recommendations can be generated faster (Konstan & Riedl, 2012).

Finally, techniques can be combined in various ways to create a hybrid recommender system. Adomavicius and Tuzhilin (2005, p. 740) illustrate how content-based and collaborative methods can be used to:

- Implement collaborative and content-based methods separately and combine their predictions

- Incorporate some content-based characteristics into a collaborative approach
- Incorporate some collaborative characteristics into a content-based approach
- Constructing a general unifying model that incorporates both content-based and collaborative characteristics

Obviously, other techniques can be combined using the same processes.

### **Evaluation of recommender systems**

Many strategies may be used to evaluate recommender systems. The choice of a strategy should take into account the tasks that the system supports as well as the nature of the data sets (Wan & Okamoto, 2011; Whittaker, Terveen, & Nardi, 2000). Beyond technical considerations, we must account for the needs and characteristics of learners (Manouselis, Drachsler, Vuorikari, Hummel, & Koper, 2011). There are three types of experiments (Gunawardana & Shani, 2015):

- Offline experiments use a protocol and existing data to evaluate the system's performance. Generally, offline experiments use data sets that mimic user behaviour to tweak the parameters of an algorithm, or to compare different approaches.
- User studies require enlisting users, asking them to perform multiple tasks requiring interaction with the system, and collecting data about these interactions. Users are asked a few questions after they complete certain requested tasks.
- Online experiments evaluate the system's performance under real-life conditions, with the users being oblivious to the experiment being conducted. They measure the change in user behaviour as the result of interacting with different recommender systems.

### **Research methodology**

A systematic review is defined as “a review of existing research using explicit, accountable rigorous research methods” (Gough, Oliver, & Thomas, 2017, p. 2). It is a type of review that searches for, appraises and synthesizes the results of research (Grant & Booth, 2009). The Evidence for Policy and Practice Information (EPPI) approach that we have followed is characterized by the use of explicit and transparent methods with the following steps: define the question, search the literature, extract relevant data, analyze data, and interpret and situate findings. Explicit methods improve the internal validity of the process and offer a critical assessment of scientific knowledge that can be leveraged for decision-making (Bertrand, L'Espérance, & Flores-Aranda, 2014). To define the criteria for inclusion and exclusion, as well as the search strategy and evaluation and analysis criteria, we relied on the characteristics of recommender systems, particularly recommender systems for TEL.

### **Research questions**

In this literature review discussing recommender systems that aim to suggest resources in a learning context that supports the learner's agency, we are focusing on the following specific questions:

- RQ1: What are the resources recommended, and what techniques were used to recommend them?
- RQ2: How were experiments performed to evaluate these recommender systems?
- RQ3: What were the results of these experiments and how were they presented?

#### **Criteria for inclusion and exclusion**

The articles we wanted to analyze had to meet a number of criteria. The peer-reviewed articles, written in English between 2008 and 2018, had to deal with a recommender system in a learning context. The systems had to recommend resources and describe the way in which the recommendations were made (algorithms, approaches). As mentioned earlier, we excluded articles that recommended learning paths, since learning paths restrict agency more than they support it. Finally, the articles retained had to describe how the prototype or systems suggested were evaluated, according to either the quality of the resources recommended or the impact of the recommendations. We thus excluded articles that only presented an appraisal of algorithmic performance (speed of execution of algorithms, for example). Finally, the data used had to be real and not simulated. Articles that used generated data sets, or data sets extracted from other contexts (for example, MovieLens) for their evaluations, were therefore excluded.

#### **Search strategy**

To find articles that cover recommendations systems in a learning context, we selected three databases. The two databases selected in education were chosen for their broad scope: Education Source (over 1000 journals) and ERIC (over 250 journals). As the subject of this review is related to technology and informatics, a database in computer science was added: Computers & Applied Sciences Complete (over 480 journals). These databases use different thesauruses, so the search criteria were different depending on the database. Searches in the education databases looked only at recommender systems, since the context is implicitly education. Searches in the applied computer science database looked at the intersection of a search on recommender systems and a search on education.

Searches in the three databases leveraged the union of the sets of results from searches by controlled vocabulary (using descriptors [DE] in Education Source and ERIC and subject items [ZU] in Computers & Applied Sciences Complete) and by free vocabulary (title [TI] and abstract [AB]). The terms used in the controlled vocabulary searches were chosen after an iterative process in which we looked at subjects used in publications related to recommender systems. The searches were performed in January 2019 (Table 1).

#### **Analytical process**

The references were imported into Rayyan<sup>1</sup> in order to identify duplicates and analyze the articles by title and abstract. The references to articles that remained after this step were imported into a reference management application (Mendeley<sup>2</sup>) to allow for annotations when reading the full articles. Though the author screened and evaluated the

---

<sup>1</sup><https://rayyan.qcri.org/>

<sup>2</sup><https://www.mendeley.com/>



**Table 1** Requests by database

Database	Request	Number of results		
		Results	Since 2008	Peer reviewed
Education Source	DE "Recommender systems" OR TI "Recommend* system*" OR TI "Recommender*" OR AB "Recommend* system*" OR AB "Recommender*"	435	376	343
ERIC	DE "Data Analysis" AND DE "Individualized Instruction" OR TI "Recommend* system*" OR TI "Recommender*" OR AB "Recommend* system*" OR AB "Recommender*"	325	221	144
Computers & Applied Sciences Complete	(ZU "recommender systems" OR ZU "recommender systems (information filtering)" OR TI "Recommend* system*" OR TI "Recommender*" OR AB "Recommend* system*" OR AB "Recommender*") AND (ZU "learning" OR ZU "education" OR TI learn* OR TI education* OR AB learn* OR AB education*)	593	550	527

articles, the identification process was performed with the help of a documentation specialist, and the thesis director supervised the conduct of the review. Regarding the interrater reliability of the process, 10% of the articles (88 randomly selected articles) were submitted to a knowledgeable computer scientist who applied the same selection criteria. According to Cohen's kappa, which is a coefficient of interrater agreement for nominal scales (Cohen, 1960), the coding consistency for inclusion or exclusion of articles between the two raters was  $\kappa = .82$ . To ensure a rigorous process, we also examined the intrarater reliability, which is a measure of self-consistency that aims to investigate the reproducibility of the measurements (Gwet, 2014). One year after the first screening, we reconducted the screening for another 10% of the articles. All inclusion/exclusion decisions remained the same, except for one paper that was included in the first screening and then excluded the second time. After verification, it was excluded at the next step (full article reading), so the decision was still the same.

Evaluating the quality of the full articles left us with articles that fit all the parameters stipulated in our research questions and met the inclusion criteria listed previously. To do this, we used a table in which we inventoried the main required elements: the recommendation context, the recommendation approach used, the experimental method and the results of the experiment. If an article did not provide an answer to all these elements, it was excluded.

Finally, the last step in the analysis was performed in Excel by the author. Each article retained was submitted to a content analysis based on thematic units. The coding system developed included information about the article, the recommender system (items, supported tasks, technique, presentation of the algorithm), the experiment settings (data, comparison, participants, methods) and the results of the experiment (measurements, results, visualization, comments). The majority of the information was written by the author of the article; therefore, we sometimes had to infer data. For example, we had to deduce the technique used by some recommender systems from their algorithm or illustration. The results are shown in the next section.

A total of 1014 references from three databases were considered; 136 duplicates were removed, and 878 were analyzed by title and abstract. We excluded articles that were not based in a learning context (538 articles), did not discuss a recommender system (112 articles), were not in English (5 articles) or were retracted by their author (1 article). The remaining 222 articles were then read in full and were evaluated and



analyzed. We excluded articles that did not discuss resource recommendation (70 articles). We also excluded articles where the evaluation was incomplete or missing (42 articles), the systems were not recommender systems (39 articles), the system approach or algorithm was not covered (11 articles) or the application was in a context other than learning (13 articles). Figure 2 summarizes the process we followed, Fig. 3 shows the number of included papers by database, and Fig. 4 shows the number of included papers with variation by year: A file with all included references is available at the [Additional file](#).

### Results

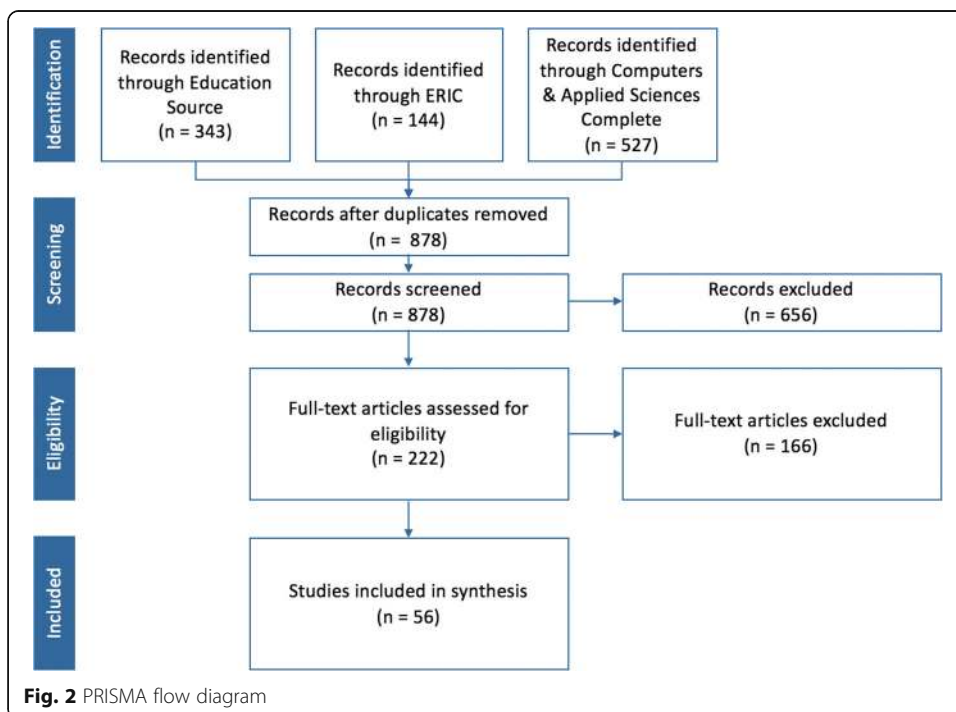
Results are shown according to the way they answer each of the three research questions. The first section shows the supported tasks and the techniques used, the second section shows the experiments documented in the articles, and the last section shows the results of the experiments.

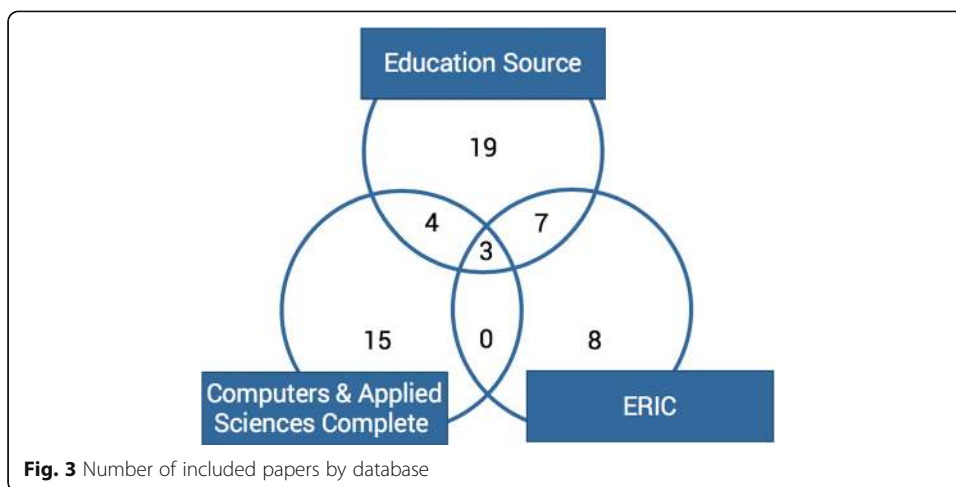
#### Supported tasks and recommendation techniques in use (RQ1)

We first focused on the tasks that the inventoried recommender systems support, and the ways in which these systems express ratings. We also looked at the techniques used and how they were reported.

#### Supported tasks

For this systematic review, the following supported tasks were considered, according to their capacity to support learner agency: finding good items, suggesting learning activities and finding peers. Their distribution and variation over the years are shown in Figs. 5 and 6.





We identified multiple types of “good items” recommended by the systems analyzed: books, learning content, publications (forum or blog posts, articles, etc.), learning material, learning objects, papers and videos. In the large majority of cases, the systems were meant for students.

**Recommendation techniques in use**

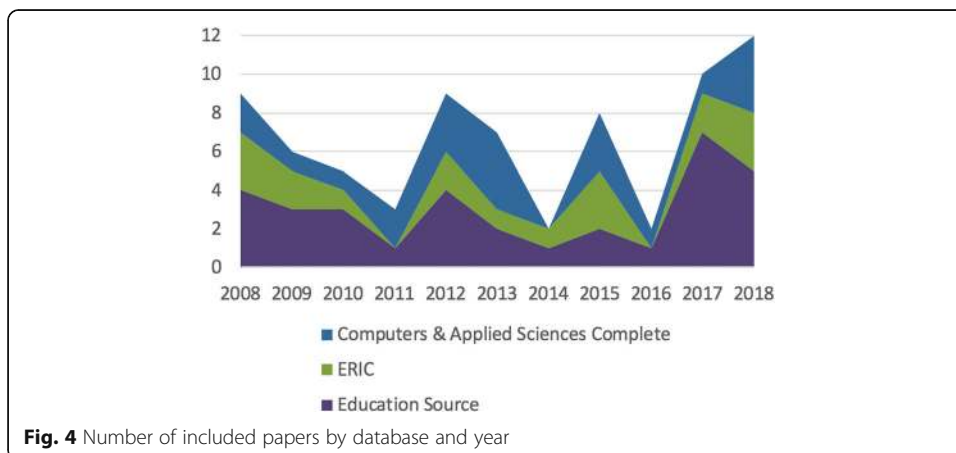
In order to complete tasks, such as finding good items (content), finding peers and suggesting learning activities, the recommender systems documented in the retained articles use different techniques, as shown in Figs. 7 and 8.

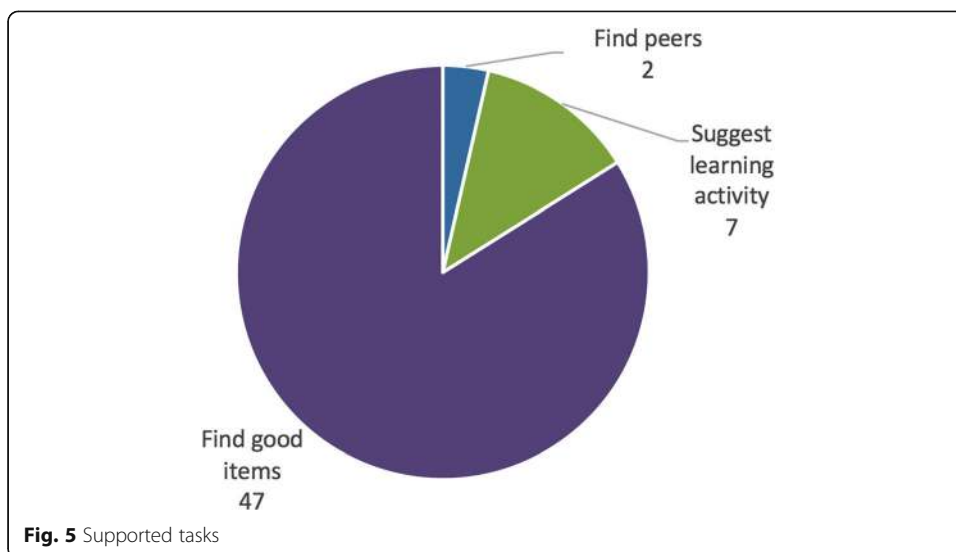
**The conduct of experiments (RQ2)**

To answer the question about how the experiments were conducted, we looked at the experiment goals and settings and at the types of experiments. For all three types of experiments, we looked at the metrics used as well as the participants involved.

**Settings**

Experiments were conducted with students or learners (38 articles), users (9 articles), staff members (3 articles), participants (1 article) or readers (1 article), and 4 articles





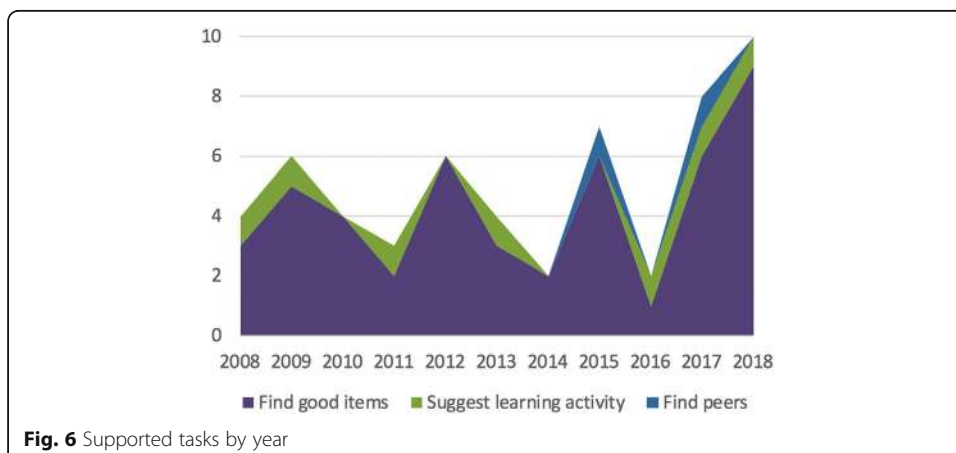
mentioned experiments with multiple groups (students and teachers; students and staff members; students and community members; mentees and experts).

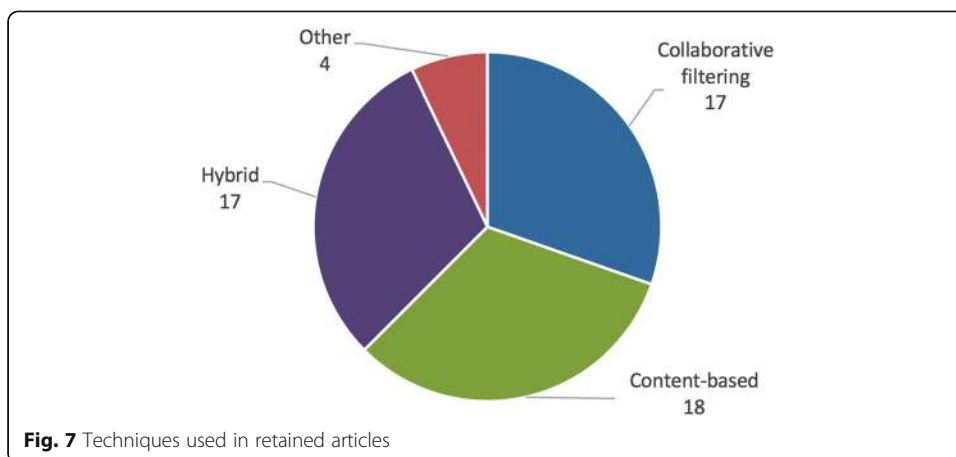
Out of 39 articles (69.6%) showing comparisons, 8 had a comparison group consisting of participants who did not receive recommendations. In 22 cases, the comparison was against other recommendation techniques. In 11 cases, the parameters of the algorithm were compared. In one case, the comparison was between the levels of the participants (the effects on a group of novices compared to the effects on a group of more advanced participants).

Finally, 12 articles described experiments that used a “training set” and “test set” strategy. The training/test ratio used was 80/20 (5 articles), 70/30 (3 articles), 65/35 (2 articles) and 28/1 (1 article). Only one article used a training set smaller than the test set (9/12).

**Types of experiments**

Since making recommendations in a learning context differs from making recommendations in a marketing context (Winoto et al., 2012), many authors, including Fazeli et



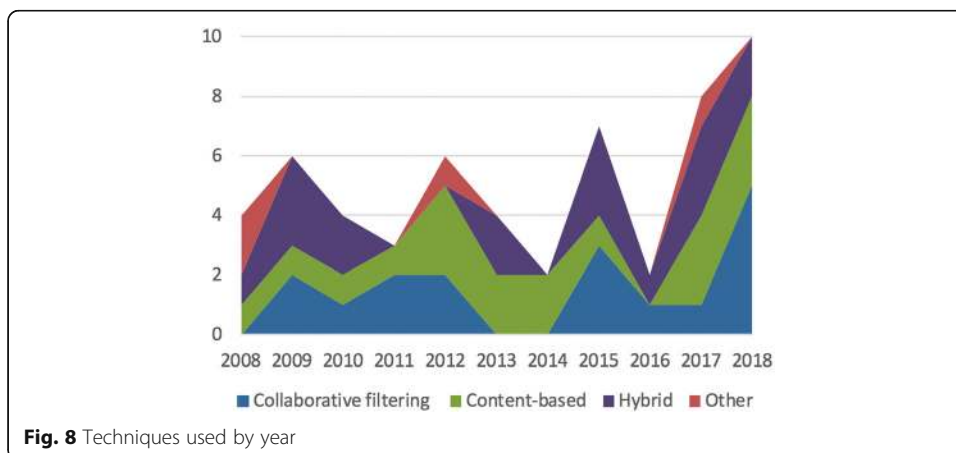


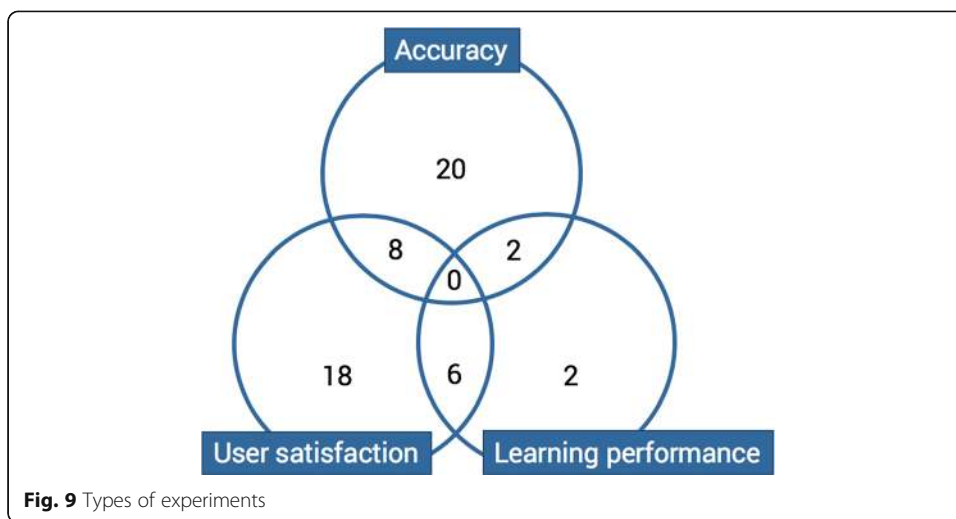
al. (2018), argue that accuracy is not the only metric to consider. Figure 9 shows the types of experiments used in the articles analyzed. Some articles show results that are applicable to more than one type of experiment.

Unsurprisingly, in the 30 articles concerning experiments testing accuracy, the evaluation metrics most commonly used are precision, recall and the F1-measure. In 13 cases, these metrics were used together. Only 2 articles recount experiments using precision without the other two metrics; 2 more used recall alone, and only 1 presents solely F1 results. Table 2 shows the number of articles using each metric where  $p$  is prediction,  $r$  is rating,  $N$  is the number of items,  $Nr$  is the number of relevant items,  $Ns$  is the number of selected items and  $Nrs$  is the number of relevant and selected items:

Rank accuracy metrics can be used, and this was the case for three articles. These articles applied normalized discounted cumulative gain (NDCG), a measure used when graded relevance values are available, and average reciprocal hit rate (ARHR, also known as mean reciprocal rank [MRR]), which is used for evaluating the ranking of the top- $n$  recommender systems.

For user studies (user satisfaction), the data gathered are sometimes quantitative: for example, how many times a recommendation was clicked, or whether a subject looked at a recommendation (determined by tracking eye movement). The data are sometimes qualitative (for example, the subject’s state of mind, or whether the subject enjoyed the





system), often obtained through questionnaires filled out before, during or after a recommendation system is used, which can be advantageous compared to offline and online experiments.

User studies were the type of evaluation most often used in the articles analyzed, as they were used in 32 articles (57.1%). Of those, half indicated that they evaluated user satisfaction. Others mentioned evaluating perceived quality (7), usefulness (3), helpfulness (2), acceptance (2) and motivation (1). In all cases, user feedback was gathered using a survey (often included directly in the application). Among those, one article specified using a feedback form, another specified using two surveys, and two articles mentioned using interviews to gather data on user appreciation.

Finally, one way to carry out online experiments is to redirect a portion of users to different recommender systems (or techniques, or parameters) and record interactions between users and the different systems. In such cases, we consider the effect of the

**Table 2** Evaluation metrics

Measure	Definition	Equation	N	%
Precision	The ratio of relevant items selected to the number of items selected.	$P = \frac{N_s}{N_s}$	20	66.67
Recall	The ratio of relevant items selected to the total number of relevant items available.	$R = \frac{N_s}{N_r}$	20	66.67
F1	The combination of precision and recall measures, because it is possible to increase one at the expense of the other.	$F_1 = \frac{2PR}{P+R}$	14	46.67
Accuracy rate	Ratio of good recommendations to all recommendations.	$AcR = \frac{N_g}{N}$	7	23.33
Mean absolute error	The divergence between prediction and actual opinion.	$MAE = \frac{\sum_{j=1}^n  p_j - r_j }{N}$	5	16.67
Coverage	The ratio of items for which a recommender system can provide recommendations.	$Coverage = \frac{N_c}{N}$	4	13.33
Root mean squared error	Mean Squared Error (MSE) is a metric that penalizes more major errors than minor errors, but does not offer an intuitive scale. RMSE repositions MSE results along a more intuitive scale.	$RMSE = \sqrt{\frac{\sum_{j=1}^n (p_j - r_j)^2}{N}}$	1	3.33
Average rating	The average rating from all the users.	$AvR = \frac{\sum_{j=1}^n r_j}{\#ratings}$	1	3.33

recommender systems on the learner's performance to be behaviour modifications. Only 10 of the 56 articles (17.9%) analyzed included an evaluation of the repercussions of the recommender system on learning. In all cases, the effects of the recommender systems on learners' performance were measured based on tests performed on the students. Nine articles out of ten used two data points (pretest and post-test), while only one used three data points (pretest, midterm exam and final exam). In all 10 articles, authors used statistical tests to compare the results obtained: 5 used t-test, 2 used ANOVA, 1 used ANCOVA, and 2 did not specify the method they used.

### **Results of experiments (RQ3)**

In this section, we will analyze the results of the documented experiments for accuracy, user satisfaction and learning performance. We will start out by examining how the authors reported their results.

#### ***Presentation of results***

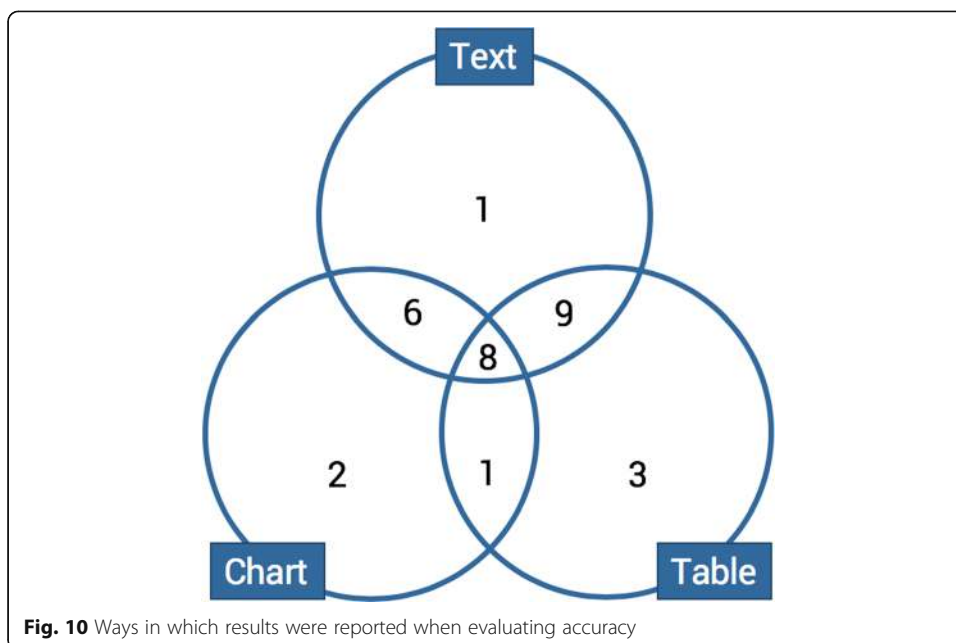
For each result, we noted the way in which it was conveyed, be it through text, charts or tables. The following Venn diagrams show, for every experiment type, the various combinations of methods used to report the results. We can see that for every experiment type, most articles use a combination of text and tables (Figs. 10, 11 and 12).

#### ***Results of the experiments***

Here we will synthesize the results described in the articles studied, first presenting the articles that reported experiments testing accuracy, then those reporting user experiments, and last, those focused on learning performance. Since this is a systematic review and not a meta-analysis, we took a step back and analyzed the conclusions drawn by the authors. As the experiment settings vary and differ, it would be perilous to compare their results.

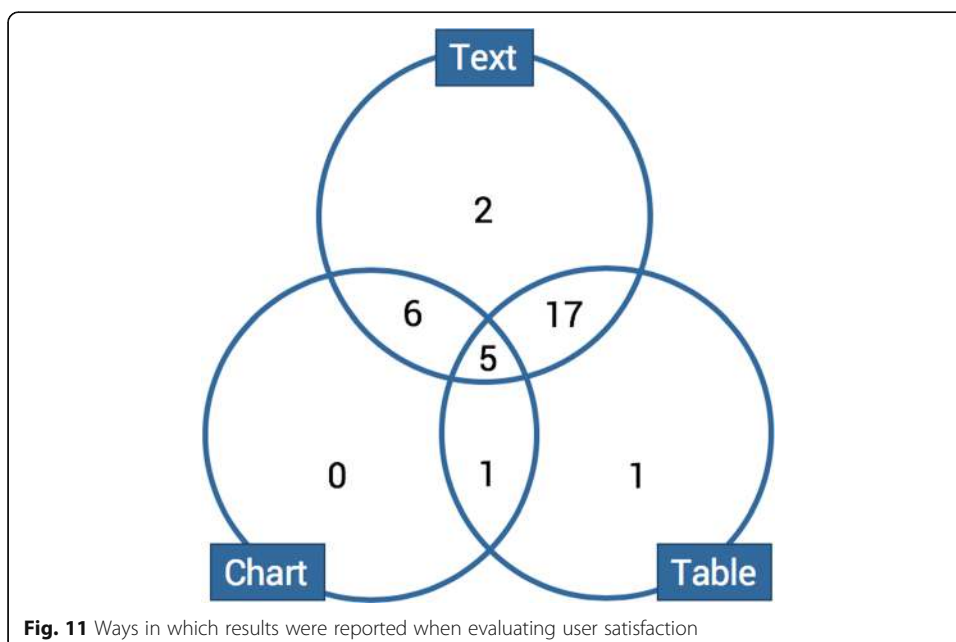
The 30 experiments evaluating accuracy introduce results for each evaluation metric used. Studies comparing different algorithm parameters declared positive results (Crespo & Antunes, 2015; Khribi, Jemni, & Nasraoui, 2009). Two articles were more nuanced: one because the accuracy of all algorithms was under 84% (Booker, 2009), and the other because the proposed algorithm was better with small item lists (Benhamdi, Babouri, & Chiky, 2017).

Studies comparing different recommendation techniques most often concluded by focusing on the proposed approach that obtained better results (Albatayneh, Ghauth, & Chua, 2018; Niemann & Wolpers, 2015). The studies comparing hybridization methods shone a light on the best hybridization techniques (Rodríguez et al., 2017; Zheng et al., 2015). Finally, some articles concluded by emphasizing the performance of the proposed approach, and solutions to problems encountered, like the problem of sparsity (Tadlaoui, Sehaba, George, Chikh, & Bouamrane, 2018), which is caused by a lack of sufficient information to identify similar users (Dascalu et al., 2015). Lastly, some approaches were proposed and evaluated for accuracy but were not compared. Those conclusions were relatively positive in terms of precision and recall (Ferreira-Satler, Romero, Menendez-Dominguez, Zapata, & Prieto, 2012).



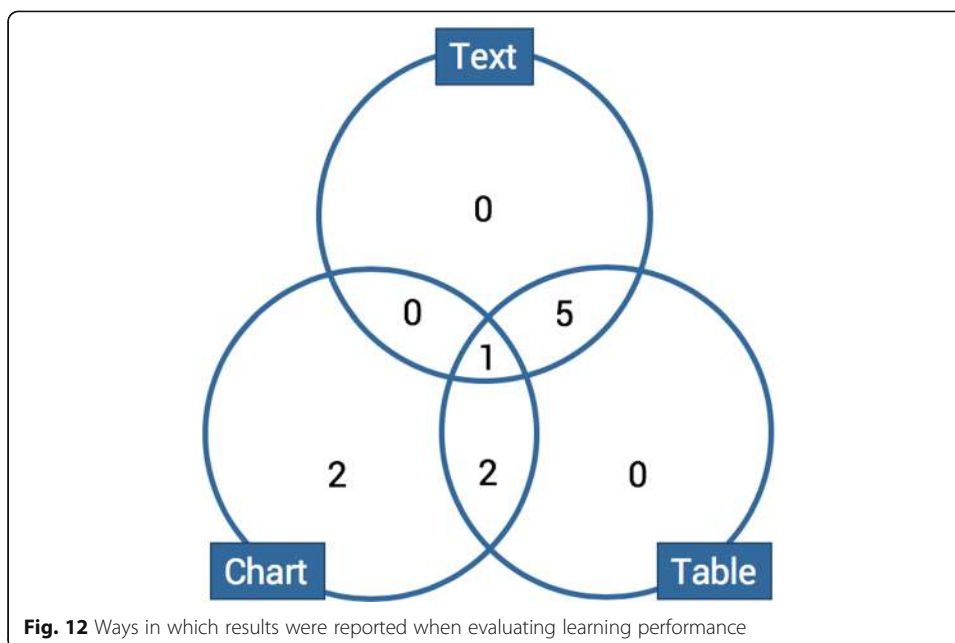
**Fig. 10** Ways in which results were reported when evaluating accuracy

For user studies, the articles that compared groups that did and did not receive personalized recommendations came to positive conclusions. For example, Hsieh, Wang, Su, and Lee (2012) presented results in which most of the experimental group learners said that the recommender system reduced the amount of effort required to search for articles they liked. Others were less positive, like Cabada, Estrada, Hernández, Bustillos, and Reyes-García (2018) and Wang and Yang (2012). Studies comparing algorithm parameters had positive conclusions (Zapata, Menéndez, Prieto, & Romero, 2013). For comparisons between techniques, the results highlighted the techniques that lead to better responses according to the questionnaires, such as in Han, Jo, Ji, and Lim (2016), who concluded that “[t]he proposed CF recommendation, which considers the



**Fig. 11** Ways in which results were reported when evaluating user satisfaction





correlations between learning skills, was observed to be more useful, accurate, and satisfactory” (p. 2282).

The results were also positive for articles that did not describe a comparison (Dascalu et al., 2015; Guangjie, Junmin, Meng, Yumin, & Chen, 2018). Some of them were also prospective, like Drachsler et al. (2010), who presented the participant’s ideas for future developments regarding, among other things, privacy and the possibility of rating the recommendations they received.

Finally, for each of the 10 articles on experiments that tested learning performance, the studies demonstrated that post-test results were statistically better for groups exposed to recommender systems (Ghauth & Abdullah, 2010). Of those studies, six made sure that the groups were of similar levels at the outset, pointing out that pre-test results did not show significant difference (Wang, 2008). Two more articles mentioned that the groups were of similar levels without indicating whether statistical tests were performed on pre-test results, and two more did not mention pre-test scores. Finally, one article mentioned that “the experimental group outperformed the control group in terms of overall quality of summary writing in the final exam. However, such a difference was not revealed between the two groups with respect to their performance in listening comprehension, reading comprehension, grammar, and vocabulary tests” (Wang & Yang, 2012, p. 633).

## Discussion

Our systematic review of recommender systems in education, particularly in a context where resource recommendation supports agency, leads us to establish the relevance of accounting for publication in both education databases and applied computer science databases.

While studying the context in which recommender systems are used in education, we established that, in most cases, the systems aimed to find good items and suggest learning activities, while only two articles introduced systems for finding peers. These ratios

are similar to those identified by Drachsler et al. (2015), who identified 61 articles about systems that aimed to find good items, 9 that aimed to find peers, and 4 that aimed to recommend learning activities. We noticed that the main recommendation techniques (content-based, collaborative filtering, hybrid) that were used followed the same ratios. We have not identified any clear trend regarding either the tasks accomplished by the recommendation systems or the recommendations techniques that have been used over the years.

Content-based recommender systems generate recommendations based on the products' features and their ratings from the user (Zapata et al., 2013). Their preferences are usually represented with an attribute or keyword preference vector. Each attribute has a dimension, and each item has a position in that space, as defined by the vector. Each user has a preference profile, also defined by a vector. Vectors may be represented in a number of ways – for example, 0 or 1 representing whether or not the attribute is present (true or false), or the number of occurrences to include an intensity value.

In the articles, different calculations were used to establish the similarity between two vectors (the user's preferences and the features of the item). One of the measures used is cosine similarity, which corresponds to the angle between these vectors (Oduwobi & Ojokoh, 2015). Another measure used in the articles is Euclidean distance (Bauman & Tuzhilin, 2018).

The vast majority of the articles describing collaborative filtering systems employ a user-based approach (15 articles). Lau, Lee, and Singh (2015) state that “to date, as far as we have surveyed, item-based CF [collaborative filtering] is yet to be employed prevalently in the e-learning domain in any significant way” (p. 85). The similarity metrics used vary, but the Pearson correlation is the most common. Other similarity metrics are Jaccard similarity,  $k$ -means algorithms and the Markov chain model.

Of the four articles that did not use content-based or collaborative filtering techniques, three used the association rule and one presented a knowledge-based system. The association rule is frequently used in sales, to suggest items complementary to those already in the shopping cart. It is the probability that a user will choose two products divided by the multiplication of the probability that the user will choose product X by the probability that the user will choose product Y (Wang, 2008). This approach can be used in a learning context to suggest complementary resources when the user identifies a resource that may help them reach their goal. Knowledge-based recommender systems suggest items based on deductions about the needs and preferences of users (Zapata et al., 2013): “[k]nowledge-based approaches use knowledge about how a particular item meets a particular user need, and can therefore reason about the relationship between a need and a possible recommendation” (Rodríguez et al., 2017, p. 20).

Finally, systems labelled “hybrid” use a combination of two or more of the techniques discussed above (Zapata et al., 2013). They use advantages from one technique to offset the disadvantages of another (Morales-del-Castillo, Peis, Moreno, & Herrera-Viedma, 2009). For example, content-based techniques can be combined with collaborative filtering techniques to handle the cold-start problem (Benhamdi et al., 2017).

The limitations of each approach were documented in previous research, and many of the authors of the selected articles also mentioned them, particularly in reference to the cold-start problem (Tang & McCalla, 2009) and the sparsity problem (Dascalu et al., 2015).

By looking at experiment settings, we observed that most articles focus only on one type of experiment (accuracy, user satisfaction, learning performance). This can be explained by the fact that different types of experiments lend themselves to being performed at different stages of the development process, from prototyping to large-scale implementation.

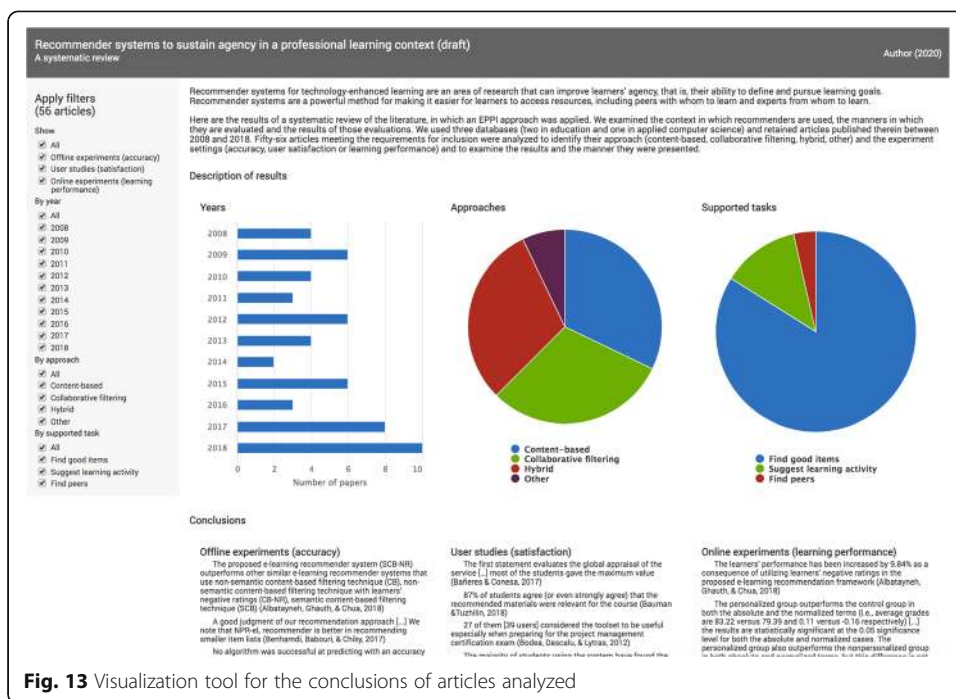
User studies were the type of evaluation most often used in the articles analyzed. However, user studies are very expensive to conduct in terms of both cost and time, and it can be difficult to recruit a sufficient number of participants. The biases inherent in this type of experiment are the same as those noted in experiments in other contexts (representativeness of samples, desirability, etc.). It is from this perspective that Fazeli et al. (2018) proposed a study in which they “traded loss of experimental control (which would have been obtained by working with fake users and fake problems) for increased ecological validity (which is obtained by working with real users, real problems, and real resources)” (p. 304). They used five user-centred metrics, explicitly establishing a parallel with the learning domain by using Vygotsky’s zone of proximal development: accuracy, novelty, diversity, serendipity and usefulness.

The results of the evaluation of the system documented are overwhelmingly positive. The results are also sometimes prospective, looking to improve the system or continue conducting experiments. Given that we were looking for a way to make better resources available for learners to pursue a goal, the results confirm that recommender systems can be a powerful method to support their agency. The positive results in terms of accuracy, learning performance and user satisfaction solidify our hypothesis that recommender systems can suggest good items, learning activities, peers and experts based on learners’ preferences, goals and search topics.

The results are presented as text and tables, as is conventional for scientific articles. Some of the results would benefit from using alternative data visualization techniques; in online publications, they could even be represented dynamically. For example, in order to let the reader view the results of each study, we developed a visualization tool that allows users to filter conclusions according to various parameters: the experiment type, the year the article was published, the supported tasks and the approach used (content-based, collaborative filtering, hybrid, other). The tool (see screen capture in Fig. 13) is available at <http://mdeschenes.com/recsys/>.

Though we refrained from comparing the results of the different experiments, it is observed that all the reported results were positive according to the different measures applied. It seems that researchers did not publish “negative results.” Could it be that some researchers wait to find an algorithm whose experimentation is acceptable before publishing, a tendency known as publication bias? This bias has already been documented in systematic reviews on recommender systems (Gasparic & Janes, 2016).

Some initiatives seek to reduce this bias, such as the LAK Failathon, the goal of which is to offer an explicit and structured space for researchers and practitioners to share their failures and to help them learn from each other’s mistakes (Clow, Ferguson, Macfadyen, Prinsloo, & Slade, 2016). However, there is still a long way to go before we can compare different algorithms, in different contexts, with different users.



**Limitations**

Even though systematic reviews attempt to bring together all the knowledge in a given topic, they are not free from limitations (Grant & Booth, 2009). In the identification phase, we may have missed interesting articles inventoried in databases other than those used. To mitigate this risk, we used three databases, even though using two databases is considered acceptable. The search terms as well as the various thesauruses used may be considered a limitation; however, we made sure to attempt multiple iterations in order to refine the search.

In a similar vein, while it is preferable not to limit the literature to a certain period, we choose an 11-year period, which we justify by the fact that the topic is of a technological nature, and technology is a rapidly evolving field. There are undoubtedly interesting articles in other languages that were not retained; we only considered articles in English.

In the screening, eligibility and inclusion phases, the main limitation is that the decisions were made by the author alone. To reduce bias, we asked a computer scientist to screen a number of articles to verify the interrater reliability and conducted an intrarater reliability analysis 1 year after the first screening. As well, some article authors did not explicitly name the approach used (content-based, collaborative filtering, hybrid, other). In those cases, we had to deduce the method from the explanation they provided.

As for the quality of the 56 studies used, we kept all those that answered our three research questions even though some offered few details. Moreover, one might argue against five of the studies retained since they did not include enough participants. They document user-centric evaluations that had fewer than 20 participants, the number of participants considered to be the minimum for user-centric evaluation of recommender systems (Knijnenburg, Willemsen, & Kobsa, 2011).

## Conclusion

In this literature review, we looked at articles about recommender systems that recommend resources in a learning context. It is not surprising, then, that most of the experiments were conducted on students. However, from an agency point of view, it would be beneficial to conduct research with learners in a less formal setting. We argue that these results could also be applicable to both lifelong learning and professional development, including professional development for teachers.

We suggest that this systematic review has shown that there is a need to develop peer recommender systems for TEL. Those peers could be other learners *with* whom we can learn, or experts *from* whom we can learn. It could be applied to learners in a formal context, and, again, it could be applied to lifelong learning.

As other authors have done before us (including Drachsler et al., 2015; Fazeli et al., 2018), we maintain that evaluation prototypes should not be limited to accuracy. They could also consider user studies and online experiments.

When it comes to the need to use online evaluations to investigate the effects of recommender systems on learning, we suggest not limiting the measurements of effects to the learners' grades. As for educational success, which includes but is not limited to academic achievement, we suggest conducting research that considers aspects other than grades, such as learners' engagement in the learning process and achievement of the goals they have set.

With this in mind, we also suggest not waiting until the end of the development process to conduct user studies and online experiments. We suggest borrowing the principle of iteration from design-based research: "During formative evaluation, iterative cycles of development, implementation, and study allow the designer to gather information about how an intervention is or is not succeeding in ways that might lead to better design" (Design-Based Research Collective, 2003, p. 7). Thus, ways of documenting variation in user satisfaction throughout the iterations should be planned, while accounting for the context in which recommender systems are used.

We also suggest borrowing from co-design and participatory design the idea of involving the users in the design process. This means no longer designing *for* users (or on behalf of users), but *with* users (Spinuzzi, 2005). Our aim is therefore to fulfill the needs expressed by the target audience and to find solutions to problems, not always to generalize.

## Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s41239-020-00219-w>.

Additional file 1

## Acknowledgements

The author thanks Dr. Thérèse Laferrière, director of her Ph.D. project. She also thanks Dr. Séverine Parent, Catherine Lamy as well as her colleagues from the Faculty of Education at Laval University for their advice about systematic literature reviews. Thanks to Nancy Deschênes, who provided interrater agreement data and translated this article.

## Authors' contributions

The author read and approved the final manuscript.

## Funding

This article was written as part of a doctoral project and will be included in the author's thesis. Funding for the research project was provided by the Fonds de recherche du Québec – Société et Culture (FRQSC).

### Availability of data and materials

The data that support the findings of this study comprise academic papers that are available from the respective publishers. We have included full references to all of the papers at the [Additional file](#).

### Competing interests

The author declared that she has no competing interest.

Received: 5 February 2020 Accepted: 25 May 2020

Published online: 21 October 2020

### References

#### References marked with an asterisk are the 56 articles retained for this systematic review

- 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*, 17(6), 734–749 <https://doi.org/10.1109/tkde.2005.99>.
- Albatayneh, N. A., Ghauth, K. I., & Chua, F.-F. (2018). Utilizing learners' negative ratings in semantic content-based recommender system for e-learning forum. *Educational Technology & Society*, 21(1), 112–125 [https://doi.org/10.1007/978-3-319-07692-8\\_35](https://doi.org/10.1007/978-3-319-07692-8_35).
- Bauman, K., & Tuzhilin, A. (2018). Recommending remedial learning materials to students by filling their knowledge gaps. *MIS Quarterly*, 42(1), 313–3A7 <https://doi.org/10.25300/misq/2018/13770>.
- Benhamdi, S., Babouri, A., & Chiky, R. (2017). Personalized recommender system for e-learning environment. *Education and Information Technologies*, 22(4), 1455–1477 <https://doi.org/10.1007/s10639-016-9504-y>.
- Bertrand, K., L'Espérance, N., & Flores-Aranda, J. (2014). La méthode de la revue systématique: illustration provenant du domaine de la toxicomanie et des troubles mentaux concomitants chez les jeunes. Méthodes qualitatives, quantitatives et mixtes dans la recherche en sciences humaines, sociales et de la santé, (pp. 145–163).
- Blaschke, L. M. (2018). Self-determined learning (heutagogy) and digital media creating integrated educational environments for developing lifelong learning skills. In *The digital turn in higher education*, (pp. 129–140). Wiesbaden: Springer VS.
- Booker, Q. E. (2009). Automating “word of mouth” to recommend classes to students: An application of social information filtering algorithms. *Journal of College Teaching & Learning*, 6(3), 39–44 <https://doi.org/10.19030/tlc.v6i3.1162>.
- Brennan, K. (2012). *Best of both worlds: Issues of structure and agency in computational creation, in and out of school* (Ph.D. Thesis). Cambridge: Massachusetts Institute of Technology.
- Butler, D. L. (2005). L'autorégulation de l'apprentissage et la collaboration dans le développement professionnel des enseignants. *Revue des Sciences de l'Éducation*, 31(1), 55–78 <https://doi.org/10.7202/012358ar>.
- Cabada, R. Z., Estrada, M. L. B., Hernández, F. G., Bustillos, R. O., & Reyes-García, C. A. (2018). An affective and web 3.0-based learning environment for a programming language. *Telematics and Informatics*, 35(3), 611–628 <https://doi.org/10.1016/j.tele.2017.03.005>.
- Camacho, L. A. G., & Alves-Souza, S. N. (2018). Social network data to alleviate cold-start in recommender system: A systematic review. *Information Processing & Management*, 54(4), 529–544 <https://doi.org/10.1016/j.ipm.2018.03.004>.
- Carré, P. (2003). La double dimension de l'apprentissage autodirigé contribution à une théorie du sujet social apprenant. *La Revue Canadienne pour l'étude de l'Éducation des Adultes*, 17, 66–91.
- Carré, P., Jézégou, A., Kaplan, J., Cyrot, P., & Denoyel, N. (2011). “L'autoformation”. The state of research on self-directed learning in France. *International Journal of Self-Directed Learning*, 8(1), 7–17.
- Clow, D., Ferguson, R., Macfadyen, L., Prinsloo, P., & Slade, S. (2016). LAK failathon. In Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, (pp. 509–511).
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46 <https://doi.org/10.1177/001316446002000104>.
- Crespo, P. T., & Antunes, C. (2015). Predicting teamwork results from social network analysis. *Expert Systems*, 32(2), 312–325 <https://doi.org/10.1111/exsy.12038>.
- Dascalu, M.-I., Bodea, C.-N., Moldoveanu, A., Mohora, A., Lytras, M., & de Pablos, P. O. (2015). A recommender agent based on learning styles for better virtual collaborative learning experiences. *Computers in Human Behavior*, 45, 243–253 <https://doi.org/10.1016/j.chb.2014.12.027>.
- Deschênes, M., & Laferrière, T. (2019). Le codesign d'une plateforme numérique fondé sur des principes au service de l'agentivité des enseignantes et des enseignants en contexte de développement professionnel. *Canadian Journal of Learning and Technology*, 45(1), 1–20 <https://doi.org/10.21432/cjlt27798>.
- Design-Based Research Collective (2003). Design-based research: an emerging paradigm for educational inquiry. *Educational Researcher*, 32(1), 5–8 <https://doi.org/10.3102/0013189x032001005>.
- Drachsler, H., Hummel, H., & Koper, R. (2008). Personal recommender systems for learners in lifelong learning: requirements, techniques and model. *International Journal of Learning Technology*, 3(4), 404–423 <https://doi.org/10.1504/ijlt.2008.019376>.
- Drachsler, H., Pecceu, D., Arts, T., Hutten, E., Rutledge, L., van Rosmalen, P., & Koper, R. (2010). ReMashed – an usability study of a recommender system for mash-ups for learning. *International Journal of Emerging Technologies in Learning*, 51, 7–11 <https://doi.org/10.3991/ijet.v5i1.1191>.
- Drachsler, H., Verbert, K., Santos, O. C., & Manouselis, N. (2015). Panorama of recommender systems to support learning. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), *Recommender systems handbook*, (pp. 421–451). Boston: Springer [https://doi.org/10.1007/978-1-4899-7637-6\\_12](https://doi.org/10.1007/978-1-4899-7637-6_12).
- Ekstrand, M. D., Riedl, J. T., & Konstan, J. A. (2011). Collaborative filtering recommender systems. *Foundations and Trends in Human-Computer Interaction*, 4(2), 81–173 <https://doi.org/10.1561/1100000009>.
- Erdt, M., Fernandez, A., & Rensing, C. (2015). Evaluating recommender systems for technology enhanced learning: a quantitative survey. *IEEE Transactions on Learning Technologies*, 8(4), 326–344 <https://doi.org/10.1109/tlt.2015.2438867>.



- Fazeli, S., Drachslar, H., Bitter-Rijpkema, M., Brouns, F., van der Vegt, W., & Sloep, P. B. (2018). User-centric evaluation of recommender systems in social learning platforms: accuracy is just the tip of the iceberg. *IEEE Transactions on Learning Technologies*, 11(3), 294–306 <https://doi.org/10.1109/tlt.2017.2732349>.
- Ferreira-Satler, M., Romero, F., Menendez-Dominguez, V., Zapata, A., & Prieto, M. (2012). Fuzzy ontologies-based user profiles applied to enhance e-learning activities. *Soft Computing – A Fusion of Foundations, Methodologies & Applications*, 16(7), 1129–1141 <https://doi.org/10.1007/s00500-011-0788-y>.
- Fischer, F., Kollar, K., Stegmann, K., & Wecker, C. (2013). Toward a script theory of guidance in computer-supported collaborative learning. *Educational Psychologist*, 48(1), 56–66.
- Gasparic, M., & Janes, A. (2016). What recommendation systems for software engineering recommend: a systematic literature review. *Journal of Systems and Software*, 113, 101–113.
- Ghauth, K. I., & Abdullah, N. A. (2010). Measuring learner's performance in e-learning recommender systems. *Australasian Journal of Educational Technology*, 26(6), 764–774 <https://doi.org/10.14742/ajet.1041>.
- Gough, D., Oliver, S., & Thomas, J. (2017). Introducing systematic reviews. In D. Gough, S. Oliver, & J. Thomas (Eds.), *An introduction to systematic reviews*, (2nd ed., pp. 1–18). London: Sage.
- Grant, M. J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, 26(2), 91–108 <https://doi.org/10.1111/j.1471-1842.2009.00848.x>.
- Guangjie, L., Junmin, L., Meng, S., Yumin, L., & Chen, W. (2018). Topic-aware staff learning material generation in complaint management systems. *International Journal of Innovation & Learning*, 24(1), 93–103 <https://doi.org/10.1504/ijil.2018.10009636>.
- Gunawardana, A., & Shani, G. (2015). Evaluating recommender systems. In *Recommender systems handbook*, (pp. 265–308). Boston: Springer [https://doi.org/10.1007/978-1-4899-7637-6\\_8](https://doi.org/10.1007/978-1-4899-7637-6_8).
- Gwet, K. L. (2014). *Handbook of inter-rater reliability: the definitive guide to measuring the extent of agreement among raters*. Gaithersburg: Advanced Analytics, LLC.
- Han, J., Jo, J., Ji, H., & Lim, H. (2016). A collaborative recommender system for learning courses considering the relevance of a learner's learning skills. *Cluster Computing*, 19(4), 2273–2284 <https://doi.org/10.1007/s10586-016-0670-x>.
- Herlocker, J. L., Konstan, J. A., & Riedl, J. (2000). Explaining collaborative filtering recommendations. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, (pp. 241–250). ACM <https://doi.org/10.1145/358916.358995>.
- Hsieh, T.-C., Wang, T.-I., Su, C.-Y., & Lee, M.-C. (2012). A fuzzy logic-based personalized learning system for supporting adaptive english learning. *Journal of Educational Technology & Society*, 15(1), 273–288.
- Jézégou, A. (2013). The influence of the openness of an E-learning situation on adult students' self-regulation. *The International Review of Research in Open and Distance Learning*, 14(3), 182–201.
- Khribi, M. K., Jemni, M., & Nasraoui, O. (2009). Automatic recommendations for E-learning personalization based on web usage mining techniques and information retrieval. *Part of a Special Issue: New Directions in Advanced Learning Technologies*, 12(4), 30–42 <https://doi.org/10.1109/icalt.2008.198>.
- Klemenčič, M. (2017). From student engagement to student agency: conceptual considerations of European policies on student-centered learning in higher education. *Higher Education Policy*, 30(1), 69–85.
- Knijnenburg, B. P., Willemsen, M. C., & Kobsa, A. (2011). A pragmatic procedure to support the user-centric evaluation of recommender systems. In *Proceedings of the fifth ACM conference on Recommender systems*, (pp. 321–324) <https://doi.org/10.1145/2043932.2043993>.
- Knowles, M. S. (1975). *Self-directed learning: a guide for learners and teachers*. New York: Association Press.
- Konstan, J. A., & Riedl, J. (2012). Deconstructing recommender systems. *IEEE Spectrum*, 10, 1–7.
- Lau, S. B.-Y., Lee, C.-S., & Singh, Y. P. (2015). A folksonomy-based lightweight resource annotation metadata schema for personalized hypermedia learning resource delivery. *Interactive Learning Environments*, 23(1), 79–105 <https://doi.org/10.1080/10494820.2012.745429>.
- Mandeville, L. (2001). Apprendre par l'expérience : un modèle de formation continue. In D. Raymond (Ed.), *Nouveaux espaces de développement professionnel et organisationnel*, (pp. 151–164). Sherbrooke: Éditions du CRP.
- Manouselis, N., Drachslar, H., Vuorikari, R., Hummel, H., & Koper, R. (2011). Recommender systems in technology enhanced learning. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), *Recommender systems handbook*, (pp. 387–415). Boston: Springer [https://doi.org/10.1007/978-0-387-85820-3\\_12](https://doi.org/10.1007/978-0-387-85820-3_12).
- Morales-del-Castillo, J. M., Peis, E., Moreno, J. M., & Herrera-Viedma, E. (2009). D-fussion: a semantic selective dissemination of information service for the research community in digital libraries. *Information Research: An International Electronic Journal*, 14(2).
- Niemann, K., & Wolpers, M. (2015). Creating usage context-based object similarities to boost recommender systems in technology enhanced learning. *IEEE Transactions on Learning Technologies*, 8(3), 274–285 <https://doi.org/10.1109/tlt.2014.2379261>.
- Oduwobi, O., & Ojokoh, B. A. (2015). Providing personalized services to users in a recommender system. *International Journal of Web-Based Learning and Teaching Technologies*, 10(2), 26–48 <https://doi.org/10.4018/ijwltt.2015040103>.
- Rahayu, P., Senseuse, D. I., Purwandari, B., Budi, I., Khalid, F., & Zulkarnaim, N. (2017). A systematic review of recommender system for e-portfolio domain. In *Proceedings of the 5th International Conference on Information and Education Technology*, (pp. 21–26) <https://doi.org/10.1145/3029387.3029420>.
- Ricci, F., Rokach, L., & Shapira, B. (2015). Recommender systems: introduction and challenges. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), *Recommender systems handbook*, (pp. 1–34). Boston: Springer [https://doi.org/10.1007/978-1-4899-7637-6\\_1](https://doi.org/10.1007/978-1-4899-7637-6_1).
- Rodríguez, P., Heras, S., Palanca, J., Poveda, J. M., Duque, N., & Julián, V. (2017). An educational recommender system based on argumentation theory. *AI Communications*, 30(1), 19–36 <https://doi.org/10.3233/aic-170724>.
- Santos, O. C., & Boticario, J. G. (2015). User-centred design and educational data mining support during the recommendations elicitation process in social online learning environments. *Expert Systems*, 32(2), 293–311 <https://doi.org/10.1111/exsy.12041>.
- Scardamalia, M. (2000). Can schools enter a knowledge society? In M. Selinger, & J. Wynn (Eds.), *Educational technology and the impact on teaching and learning*, (pp. 6–10). Abingdon: Research Machines.



- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge handbook of the learning sciences*, (pp. 97–118). New York: Cambridge University Press.
- Spinuzzi, C. (2005). The methodology of participatory design. *Technical Communication*, 52(2), 163–174 [https://doi.org/10.1207/s15427625tcq0604\\_4](https://doi.org/10.1207/s15427625tcq0604_4).
- Straka, G. A. (1999). Perceived work conditions and self-directed learning in the process of work. *International Journal of Training and Development*, 3(4), 240–249.
- Tadlaoui, M., Sehaba, K., George, S., Chikh, A., & Bouamrane, K. (2018). Social recommender approach for technology-enhanced learning. *International Journal of Learning Technology*, 13(1), 61–89 <https://doi.org/10.1504/ijlt.2018.091631>.
- Tang, T. Y., & McCalla, G. (2009). A multidimensional paper recommender. *IEEE Internet Computing*, 13(4), 34–41 <https://doi.org/10.1109/mic.2009.73>.
- Tarus, J. K., Niu, Z., & Mustafa, G. (2018). Knowledge-based recommendation: a review of ontology-based recommender systems for e-learning. *Artificial Intelligence Review*, 50(1), 21–48 <https://doi.org/10.1007/s10462-017-9539-5>.
- Verbert, K., Manouselis, N., Ochoa, X., Wolpers, M., Drachler, H., Bosnic, I., & Duval, E. (2012). Context-aware recommender systems for learning: a survey and future challenges. *IEEE Transactions on Learning Technologies*, 5(4), 318–335.
- Vygotsky, L. S. (1978). *Mind in society: the development of higher psychological processes*. Cambridge: Harvard University Press.
- Wan, X., & Okamoto, T. (2011). Utilizing learning process to improve recommender system for group learning support. *Neural Computing & Applications*, 20(5), 611–621 <https://doi.org/10.1007/s00521-009-0283-x>.
- Wang, F.-H. (2008). Content recommendation based on education-contextualized browsing events for web-based personalized learning. *Educational Technology & Society*, 11(4), 94–112.
- Wang, P.-Y., & Yang, H.-C. (2012). Using collaborative filtering to support college students' use of online forum for English learning. *Computers & Education*, 59(2), 628–637 <https://doi.org/10.1016/j.compedu.2012.02.007>.
- Whittaker, S., Terveen, L., & Nardi, B. A. (2000). Let's stop pushing the envelope and start addressing it: a reference task agenda for HCI. *Human Computer Interaction*, 15(2–3), 75–106 [https://doi.org/10.1207/s15327051hci1523\\_2](https://doi.org/10.1207/s15327051hci1523_2).
- Winoto, P., Tang, T. Y., & McCalla, G. (2012). Contexts in a paper recommendation system with collaborative filtering. *International Review of Research in Open and Distance Learning*, 13(5), 56–75 <https://doi.org/10.19173/irrodl.v13i5.1243>.
- Zapata, A., Menéndez, V. H., Prieto, M. E., & Romero, C. (2013). A framework for recommendation in learning object repositories: an example of application in civil engineering. *Advances in Engineering Software*, 56, 1–14 <https://doi.org/10.1016/j.advengsoft.2012.10.005>.
- Zheng, X.-L., Chen, C.-C., Hung, J.-L., He, W., Hong, F.-X., & Lin, Z. (2015). A hybrid trust-based recommender system for online communities of practice. *IEEE Transactions on Learning Technologies*, 8(4), 345–356 <https://doi.org/10.1109/tlt.2015.2419262>.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---